

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

(19) World Intellectual Property

Organization

International Bureau

(43) International Publication Date

27 August 2020 (27.08.2020)



(10) International Publication Number

WO 2020/168407 A1

(51) International Patent Classification:

C12N 15/52 (2006.01) C12P 7/42 (2006.01)

C12N 9/02 (2006.01)

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).

(21) International Application Number:

PCT/BR2020/050052

(22) International Filing Date:

20 February 2020 (20.02.2020)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/808,247 20 February 2019 (20.02.2019) US

(71) Applicant: **BRASKEM S.A.** [BR/BR]; Rua Eteno, 1.561, Polo Petroquímico, 42810-000 Camaçari/BA (BR).

(72) Inventors: **KOCH, Daniel Johannes**; c/o Braskem S.A., Rua Eteno, 1.561, Polo Petroquímico, 42810-000 Camaçari/BA (BR). **PEDERSON PARIZZI, Lucas**; c/o Braskem S.A., Rua Eteno, 1.561, Polo Petroquímico, 42810-000 Camaçari/BA (BR). **GALZERANI, Felipe**; c/o Braskem S.A., Rua Eteno, 1.561, Polo Petroquímico, 42810-000 Camaçari/BA (BR).

(74) Agent: **DANNEMANN, SIEMSEN, BIGLER & IPANEMA MOREIRA**; Rua Marquês de Olinda, 70 - Botafogo, 22251-040 Rio de Janeiro - RJ (BR).

(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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

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

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))
- in black and white; the international application as filed contained color or greyscale and is available for download from PATENTSCOPE

(54) Title: MICROORGANISMS AND METHODS FOR THE PRODUCTION OF OXYGENATED COMPOUNDS FROM HEXOSES

(57) Abstract: The present application relates to recombinant microorganisms useful in the biosynthesis of monoethylene glycol (MEG), or optionally MEG and one or more co-product, from one or more hexose feedstock. The present application also relates to recombinant microorganisms useful in the biosynthesis of glycolic acid (GA), or optionally GA and one or more co-product, from one or more hexose feedstock. The present application relates to recombinant microorganisms useful in the biosynthesis of xylitol, or optionally xylitol and one or more co-product, from one or more hexose feedstock. Also provided are methods of producing MEG (or GA or xylitol), or optionally MEG (or GA or xylitol) and one or more co-product, from one or more hexose feedstock using the recombinant microorganisms, as well as compositions comprising the recombinant microorganisms and/or the products MEG (or GA or xylitol), or optionally MEG (or GA or xylitol) and one or more co-product.



WO 2020/168407 A1

MICROORGANISMS AND METHODS FOR THE PRODUCTION OF OXYGENATED COMPOUNDS FROM HEXOSES

5 CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to U.S. Provisional Application No. 62/808,247 filed February 20, 2019, entitled "MICROORGANISMS AND METHODS FOR THE PRODUCTION OF OXYGENATED COMPOUNDS FROM HEXOSES", the disclosures of which are incorporated by reference herein.

10 TECHNICAL FIELD

[0002] This application relates to recombinant microorganisms useful in the biosynthesis of monoethylene glycol or monoethylene glycol and one or more co-product from one or more hexose feedstock. This application additionally relates to recombinant microorganisms useful in the biosynthesis of glycolic acid or glycolic acid and one or more co-product from one or more hexose feedstock. The application further relates to methods of producing monoethylene glycol or monoethylene glycol and one or more co-product from one or more hexose feedstock using the recombinant microorganisms, as well as methods of producing glycolic acid or glycolic acid and one or more co-product from one or more hexose feedstock using the recombinant microorganisms. The application further relates to compositions comprising one or more of these compounds and/or the recombinant microorganisms.

STATEMENT REGARDING SEQUENCE LISTING

[0003] The Sequence Listing associated with this application is provided in text format in lieu of a paper copy, and is hereby incorporated by reference into the specification. The name of the text file containing the Sequence Listing is BRSK-004_02WO_ST25.txt. The text file is about 616 KB, was created on February 18, 2020, and is being submitted electronically via EFS-Web.

BACKGROUND

[0004] A large number of chemical compounds are currently derived from petrochemicals. Compounds such as monoethylene glycol (MEG), glycolic acid, acetone, isopropanol (IPA), propene, serine, glycine, monoethanolamine, and ethylenediamine are valuable as raw material in the production of products like polyethylene terephthalate (PET) resins (from MEG), plastic polypropylene (from propene), polyglycolic acid and other biocompatible copolymers (from glycolic acid) and polyurethane fibers (from ethylenediamine). Alkenes (such as ethylene, propylene, different butenes, and pentenes,

for example) are used in the plastics industry, fuels, and in other areas of the chemical industry. For example, isobutene is a small, highly reactive molecule that is used extensively as a platform chemical to manufacture a wide variety of products including fuel additives, rubber and rubber additives, and specialty chemicals.

5 **[0005]** However, the compounds are currently produced from precursors that originate from fossil fuels, which contribute to climate change. To develop more environmentally friendly processes for the production of MEG, researchers have engineered microorganisms with biosynthetic pathways to produce MEG. However, these pathways are challenging to implement, with loss of product yield, redox balance and excess biomass formation being
10 some major obstacles to overcome.

[0006] Thus there exists a need for improved biosynthesis pathways for the production of MEG and other chemical compounds useful in industrial and pharmaceutical applications.

SUMMARY OF THE DISCLOSURE

15 **[0007]** The present application relates to recombinant microorganisms having one or more biosynthesis pathways for the production of monoethylene glycol (MEG) or glycolic acid (GA), or optionally, MEG (or GA) and one or more co-product from one or more hexose feedstock.

20 **[0008]** The recombinant microorganisms and methods of the present disclosure combine the advantages of glucose based, fermentative MEG production and xylose based, fermentative MEG production. In some embodiments, the recombinant microorganisms and methods of the present disclosure combine the advantages of xylose degradation biochemistry for high yielding MEG (or GA), or optionally, MEG (or GA) and one or more co-product, formation with the advantages of readily available pure hexose sugar feedstocks.

25 **[0009]** In some embodiments, the recombinant microorganisms and methods of the present disclosure solves the problem of xylose feedstock availability. In some embodiments, the recombinant microorganisms and methods of the present disclosure solves the problem of non-affordable xylose feedstock price. In some embodiments, the recombinant microorganisms and methods of the present disclosure solves the problem of
30 xylose feedstock impurities. In some embodiments, the recombinant microorganisms and methods of the present disclosure solves the problem of inefficient xylose uptake by a microorganism. In some embodiments, the recombinant microorganisms and methods of the present disclosure solves the problem of glucose induced inhibition of xylose utilization. In

some embodiments, the recombinant microorganisms and methods of the present disclosure solves the problem of a shortage of ATP in MEG (or GA) production pathways. In some embodiments, the recombinant microorganisms and methods of the present disclosure solves the problem of excess NADH in MEG (or GA) production pathways. In some
5 embodiments, the recombinant microorganisms and methods of the present disclosure solves the problem of low overall product yield potential.

[0010] In some embodiments, the recombinant microorganisms and methods of the present disclosure provide a lossless conversion of one or more hexose feedstock to one or more pentose-5-phosphate intermediate. In some embodiments, the one or more pentose-5-phosphate intermediate is used for the production of MEG (or GA), or optionally, MEG (or
10 GA) and one or more co-product, by one or more xylose based fermentation methods. In some embodiments, glucose flux is funneled into the pentose phosphate pathway instead of the glycolysis pathway.

[0011] In one aspect, the present disclosure provides a recombinant microorganism comprising one or more biochemical pathway that produces monoethylene glycol (MEG) (or glycolic acid) from one or more hexose feedstock via one or more pentose-5-phosphate intermediate. In one embodiment, one or more co-product is co-produced with MEG (or glycolic acid). In another embodiment, the one or more pentose-5-phosphate intermediate is one or more of D-xylulose-5-phosphate, D-ribulose-5-phosphate or D-ribose-5-phosphate.
15

[0012] Therefore, in one embodiment, the application relates to a recombinant microorganism comprising one or more biochemical pathway comprising the expression of at least one enzyme having an activity that converts one or more hexose feedstock in a lossless conversion to one or more pentose-5-phosphate intermediate.
20

[0013] In some embodiments, the recombinant microorganism comprises the expression of at least one enzyme having transketolase activity. In some embodiments, the enzyme having transketolase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to tktA from *E. coli*. In other embodiments, the enzyme having transketolase activity is tktA from *E. coli*. In some embodiments, the enzyme having transketolase activity is encoded by an amino acid sequence having at least 70% sequence identity at least 80% sequence identity, or at least 90% sequence identity to tktB from *E. coli*. In other embodiments, the enzyme having transketolase activity is tktB from *E. coli*. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having transketolase activity comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 148 and 150. In
25 some embodiments, the one or more nucleic acid molecules encoding the enzyme having
30
35

transketolase activity is tktA, or homolog thereof. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having transketolase activity is tktB, or homolog thereof. In a further embodiment, the one or more nucleic acid molecule encoding the enzyme having transketolase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 147 and 149.

[0014] In some embodiments, the recombinant microorganism comprises the expression of at least one enzyme having transaldolase activity. In some embodiments, the enzyme having transaldolase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to talA or talB from *E. coli*. In some embodiments, the enzyme having transaldolase activity is talA from *E. coli*. In other embodiments, the enzyme having transaldolase activity is talB from *E. coli*. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having transaldolase activity comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 152 and 154. In a further embodiment, the one or more nucleic acid molecule encoding the enzyme having transaldolase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 151 and 153.

[0015] In some embodiments, the recombinant microorganism comprises the expression of at least one enzyme having ribulose-5-phosphate 3-epimerase activity. In some embodiments, the enzyme having ribulose-5-phosphate 3-epimerase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to rpe from *E. coli*. In other embodiments, the enzyme having ribulose-5-phosphate 3-epimerase activity is rpe from *E. coli*. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having ribulose-5-phosphate 3-epimerase activity comprise an amino acid sequence set forth in SEQ ID NO: 158. In a further embodiment, the one or more nucleic acid molecule encoding the enzyme having ribulose-5-phosphate 3-epimerase activity is encoded by a nucleic acid sequence set forth in SEQ ID NO: 157.

[0016] In some embodiments, the recombinant microorganism comprises the expression of at least one enzyme having ribose-5-phosphate isomerase activity. In some embodiments, the enzyme having ribose-5-phosphate isomerase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to rpiA from *E. coli*. In other embodiments, the enzyme having ribose-5-phosphate isomerase activity is rpiA from *E. coli*. In other embodiments, the enzyme having ribose-5-phosphate isomerase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to rpiB from *E. coli*. In other embodiments, the enzyme having ribose-

5-phosphate isomerase activity is rpiB from *E. coli*. In another embodiment, the one or more nucleic acid molecules encoding enzyme having ribose-5-phosphate isomerase activity comprise an amino acid sequence set forth in SEQ ID NO: 156. In a further embodiment, the one or more nucleic acid molecule encoding the enzyme having ribose-5-phosphate isomerase activity is encoded by a nucleic acid sequence set forth in SEQ ID NO: 155.

[0017] In some embodiments, the recombinant microorganism comprises the expression of at least one enzyme having a transketolase activity, a transaldolase activity, a ribulose-5-phosphate 3-epimerase activity, and a ribose-5-phosphate isomerase activity. In other embodiments, the recombinant microorganism further comprises a deleted or diminished activity in one or more endogenous enzymes selected from glyceraldehyde 3-phosphate dehydrogenase, phosphoglycerate kinase and phosphoglycerate mutase. In some embodiments, the endogenous glyceraldehyde 3-phosphate dehydrogenase enzyme is gapA, the phosphoglycerate kinase is pgk and the phosphoglycerate mutase is gpmA and/or gpmM.

[0018] In some embodiments, the recombinant microorganism comprises the expression of at least one enzyme having fructose-6-phosphate phosphoketolase activity. In some embodiments, an enzyme having fructose-6-phosphate phosphoketolase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having fructose-6-phosphate phosphoketolase activity selected from the group consisting of *Bifidobacterium dentium* BDP_1006, *Bifidobacterium lactis* xfp, *Lactobacillus paraplantarum* xpkA and *Bifidobacterium breve* xfp. In a preferred embodiment, an enzyme having fructose-6-phosphate phosphoketolase activity is selected from the group consisting of *Bifidobacterium dentium* BDP_1006, *Bifidobacterium lactis* xfp, *Lactobacillus paraplantarum* xpkA and *Bifidobacterium breve* xfp. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having fructose-6-phosphate phosphoketolase activity comprise an amino acid sequence selected from the group consisting of SEQ ID NOs: 212, 214, 216 and 218. In a further embodiment, the one or more nucleic acid molecule encoding the enzyme having fructose-6-phosphate phosphoketolase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 211, 213, 215 and 217.

[0019] In some embodiments, the recombinant microorganism comprises the expression of at least one enzyme having phosphate acetyltransferase activity. In some embodiments, an enzyme having phosphate acetyltransferase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having phosphate acetyltransferase activity selected from *E. coli* pta and *Clostridium acetobutylicum* pta. In a preferred embodiment, an enzyme

having phosphate acetyltransferase activity is selected from *E. coli* pta and *Clostridium acetobutylicum* pta. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having phosphate acetyltransferase activity comprise an amino acid sequence selected from SEQ ID NOs: 220 and 222. In a further embodiment, the one or more nucleic acid molecule encoding the enzyme having phosphate acetyltransferase activity is encoded by a nucleic acid sequence selected from SEQ ID NOs: 219 and 221.

[0020] In some embodiments, the recombinant microorganism comprises the expression of at least one enzyme having a fructose-6-phosphate phosphoketolase activity, a phosphate acetyltransferase activity, a transketolase activity, a transaldolase activity, a ribulose-5-phosphate 3-epimerase activity, and a ribose-5-phosphate isomerase activity. In other embodiments, the recombinant microorganism further comprises a deleted or diminished activity in an endogenous 6-phosphofructokinase enzyme. In some embodiments, the endogenous 6-phosphofructokinase enzyme is pfkA and/or pfkB.

[0021] In another embodiment, the one or more pentose-5-phosphate intermediate produced in the lossless conversion of one or more hexose feedstock can be connected with any one of the known C2 MEG or glycolic acid production pathways by conversion of the one or more pentose-5-phosphate intermediate to one or more pentose and/or pentose-1-phosphate intermediate. In some embodiments, the recombinant microorganism comprises the expression of at least one enzyme having a pentose phosphatase activity, an arabinol phosphate dehydrogenase activity, and/or a phosphopentomutase activity. In some embodiments, the phosphopentomutase is (Pgm3). In some embodiments, the phosphopentosemutase is Pgm3 from *Saccharomyces cerevisiae*. In some embodiments, the phosphopentosemutase is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity or at least 90% sequence identity to the Pgm3 from *Saccharomyces cerevisiae*.

[0022] In some embodiments, the recombinant microorganism comprises the expression of at least one enzyme having pentose phosphatase activity. In other embodiments, the at least one enzyme having pentose phosphatase activity is selected from one or more of an enzyme having D-pentose-5-phosphatase activity, an enzyme having D-xylulose-5-phosphatase activity, an enzyme having D-ribose-5-phosphatase activity, and an enzyme having D-ribulose-5-phosphatase activity. In some embodiments, the pentose phosphatase is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to a D-pentose-5-phosphatase selected from the group consisting of *E. coli* phoA, *E. coli* yfbT and *E. coli* yidA. In some embodiments, the enzyme having D-xylulose-5-phosphatase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at

least 90% sequence identity to *Bacillus subtilis* araL. In some embodiments, the enzyme having D-ribose-5-phosphatase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having D-ribose-5-phosphatase activity selected from the group consisting of *Arabidopsis thaliana* SGPP, *Pseudomonas fluorescens* PFLU_2693, and *E. coli* ybiV. In some embodiments, the enzyme having D-ribulose-5-phosphatase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to *Plasmodium falciparum* PF10_0325. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having D-pentose-5-phosphatase activity is selected from the group consisting of SEQ ID NOs: 159, 161, 163, 165, 167, 169, 171 and 173. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having D-pentose-5-phosphatase activity comprise an amino acid sequence selected from the group consisting of SEQ ID NOs: 160, 162, 164, 166, 168, 170, 172 and 174.

[0023] In some embodiments, the recombinant microorganism comprises the expression of at least one enzyme having arabitol phosphate dehydrogenase activity. In some embodiments, the enzyme having arabitol phosphate dehydrogenase activity is selected from one or more of an enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity, an enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity, an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity and an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity.

[0024] In some embodiments, an enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to APDH from *Enterococcus avium*. In a preferred embodiment, an enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity is APDH from *Enterococcus avium*. In some embodiments, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity is set forth in SEQ ID NO: 175. In another embodiment, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity comprise an amino acid sequence set forth in SEQ ID NO: 176.

[0025] In some embodiments, an enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to APDH from *Enterococcus avium*. In a preferred embodiment, an enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity is APDH from *Enterococcus avium*. In some

embodiments, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity is set forth in SEQ ID NO: 175. In another embodiment, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity comprise an amino acid sequence set forth in SEQ ID NO: 176.

[0026] In some embodiments, an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity selected from the group consisting of *Candida albicans* ARD1, *Candida tropicalis* ARD1, *Scheffersomyces stipitis* ARDH, *Bacillus subtilis* egsA (araM), *Aeropyrum pernix* egsA, *E. coli* gpsA and *Saccharomyces cerevisiae* GPD1. In a preferred embodiment, an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity is selected from the group consisting of *Candida albicans* ARD1, *Candida tropicalis* ARD1, *Scheffersomyces stipitis* ARDH, *Bacillus subtilis* egsA (araM), *Aeropyrum pernix* egsA, *E. coli* gpsA and *Saccharomyces cerevisiae* GPD1. In some embodiments, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity is selected from the group consisting of SEQ ID NOs: 177, 179, 181, 189, 191, 193 and 195. In another embodiment, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity comprise an amino acid sequence selected from the group consisting of SEQ ID NOs: 178, 180, 182, 190, 192, 194 and 196.

[0027] In some embodiments, an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity selected from the group consisting of *Pseudomonas fluorescens* mtID, *Klebsiella pneumoniae* dalD, *Ralstonia solanacearum* dalD, *Bacillus subtilis* egsA (araM), *Aeropyrum pernix* egsA, *E. coli* gpsA and *Saccharomyces cerevisiae* GPD1. In a preferred embodiment, an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity is selected from the group consisting of *Pseudomonas fluorescens* mtID, *Klebsiella pneumoniae* dalD, *Ralstonia solanacearum* dalD, *Bacillus subtilis* egsA (araM), *Aeropyrum pernix* egsA, *E. coli* gpsA and *Saccharomyces cerevisiae* GPD1. In some embodiments, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity is selected from the group consisting of SEQ ID NOs: 183, 185, 187, 189, 191, 193 and 195. In another embodiment, the one or more nucleic acid molecules encoding an enzyme having D-arabitol

5-phosphate 4-dehydrogenase activity comprise an amino acid sequence selected from the group consisting of SEQ ID NOs: 184, 186, 188, 190, 192, 194 and 196.

[0028] In some embodiments, the recombinant microorganism comprises the expression of at least one enzyme having phosphopentomutase activity. In some embodiments, an enzyme having phosphopentomutase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having phosphopentomutase activity selected from the group consisting of *E. coli* deoB, *E. coli* pgm, *Bacillus subtilis* pgcA, *Lactococcus lactis* pgmB, *E. coli* ycjU, *Pseudomonas aeruginosa* algC, and *E. coli* cpsG. In a preferred embodiment, an enzyme having phosphopentomutase activity is selected from the group consisting of *E. coli* deoB, *E. coli* pgm, *Bacillus subtilis* pgcA, *Lactococcus lactis* pgmB, *E. coli* ycjU, *Pseudomonas aeruginosa* algC, and *E. coli* cpsG. In some embodiments, the one or more nucleic acid molecules encoding an enzyme having phosphopentomutase activity is selected from the group consisting of SEQ ID NOs: 197, 199, 201, 203, 205, 207 and 209. In another embodiment, the one or more nucleic acid molecules encoding an enzyme having phosphopentomutase activity comprise an amino acid sequence selected from the group consisting of SEQ ID NOs: 198, 200, 202, 204, 206, 208 and 210.

[0029] In some embodiments of any of the recombinant microorganisms described above, the recombinant microorganism further comprises a diminished activity in one or more endogenous enzymes selected from glucose 6-phosphate-1-dehydrogenase, 6-phosphogluconolactonase, and 6-phosphogluconate dehydrogenase. In further embodiments, the glucose 6-phosphate-1-dehydrogenase is zwf, the 6-phosphogluconolactonase is pgl, and the 6-phosphogluconate dehydrogenase is gnd.

[0030] In some embodiments, the application provides for a recombinant microorganism that co-produces MEG (or glycolic acid) and one or more co-product selected from acetone, isopropanol, propene, isobutene and one or more serine pathway compounds. In some preferred embodiments, the one or more serine pathway compounds is selected from serine, glycine, monoethanolamine (MEA) and ethylenediamine (EDA).

[0031] In some embodiments, the one or more hexose feedstock is selected from glucose or oligomers of glucose thereof. In other embodiments, the oligomers of glucose are selected from fructose, sucrose, starch, cellobiose, maltose, lactose and cellulose.

[0032] In some embodiments, the expression of an enzyme having transketolase activity or an enzyme having fructose-6-phosphate phosphoketolase activity in the recombinant microorganism enables a lossless conversion of one or more hexose feedstock to one or more pentose-5-phosphate intermediate.

[0033] In some embodiments, the recombinant microorganism produces MEG or glycolic acid (GA) through the conversion of glycolaldehyde in a C2 pathway and through the conversion of dihydroxyacetone phosphate (DHAP) or pyruvate in a C3 pathway. In other embodiments, the glycolaldehyde is oxidized to glycolic acid by a glycolaldehyde dehydrogenase.

[0034] In some embodiments, the at least one enzyme for the production of MEG or GA through the conversion of DHAP or pyruvate in a C3 pathway are selected from at least one enzyme having an activity selected from a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine transaminase activity, a hydroxypyruvate decarboxylase activity, a 3-phosphohydroxypyruvate reductase activity, a glycolaldehyde reductase activity, a glycolaldehyde dehydrogenase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a serine decarboxylase activity, an ethanolamine aminotransferase or ethanolamine oxidoreductase (deaminating) activity, a glycerate decarboxylase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, and a glycerate 2-kinase activity.

[0035] In some embodiments, the recombinant microorganism produces MEG or glycolic acid (GA) through the conversion of glycolaldehyde in a C2 pathway and produces one or more co-product through the conversion of dihydroxyacetone phosphate (DHAP) or pyruvate in a C3 pathway. In other embodiments, the one or more co-product is selected from acetone, isopropanol, propene, isobutene and one or more serine pathway compounds. In some preferred embodiments, the one or more serine pathway compounds is selected from serine, glycine, monoethanolamine (MEA) and ethylenediamine (EDA).

[0036] In some embodiments, the at least one enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from at least one enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, and an acetoacetate decarboxylase activity, and the one or more co-product comprises acetone.

[0037] In some embodiments, the at least one enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from at least one enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-

CoA hydrolase activity, an acetoacetate decarboxylase activity, and a secondary alcohol dehydrogenase activity, and the one or more co-product comprises isopropanol.

[0038] In some embodiments, the at least one enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from at least one enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, a secondary alcohol dehydrogenase activity, and a dehydratase activity, and the one or more co-product comprises propene.

[0039] In some embodiments, the at least one enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from at least one enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, a 3-hydroxyisovalerate (3HIV) synthase activity, a hydroxymethylglutaryl-CoA synthase activity, a methylglutaconyl-CoA hydratase activity, a methylcrotonyl-CoA carboxylase activity, a methylcrotonyl-CoA hydratase activity, a 3-hydroxyisovaleryl-CoA thioesterase activity, a 3HIV kinase activity, a 3HIV-3-phosphate decarboxylase activity, and a 3HIV decarboxylase activity, and the one or more co-product comprises isobutene.

[0040] In some embodiments, the at least one enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from at least one enzyme having an activity selected from a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, and a glycerate 2-kinase activity, and wherein the one or more co-product comprises L-serine.

[0041] In some embodiments, the at least one enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from at least one enzyme having an activity selected from a serine hydroxymethyltransferase activity, a transferase activity, a formaldehyde dehydrogenase activity, a formate dehydrogenase activity, an activity associated with glycine cleavage system, a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a

serine transaminase activity, a hydroxypyruvate decarboxylase activity, a serine oxidoreductase (deaminating) activity, a serine decarboxylase activity, an ethanolamine aminotransferase or ethanolamine oxidoreductase (deaminating) activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, a glycerate 2-kinase activity, a glycolaldehyde dehydrogenase activity, a glycolate dehydrogenase activity, an alanine-glyoxylate aminotransferase activity, an alanine transaminase activity, an NAD(P)H dependent glutamate dehydrogenase activity, and wherein the one or more co-product comprises glycine. In another embodiment, the activity associated with glycine cleavage system comprise an enzyme or protein selected from a glycine decarboxylase (P protein), an aminomethyltransferase (T protein), a dihydrolipoamide dehydrogenase (L protein), and an H protein.

[0042] In some embodiments, the at least one enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from at least one enzyme having an activity selected from a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a transaminase activity, a hydroxypyruvate decarboxylase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a serine decarboxylase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, a glycerate 2-kinase activity, an acetaldehyde dehydrogenase activity, and an ethanolamine ammonia lyase activity, and the one or more co-product comprises monoethanolamine (MEA).

[0043] In some embodiments, the at least one enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from at least one enzyme having an activity selected from a serine dehydrogenase activity, a 2-aminomalonate semialdehyde decarboxylase activity, an aminoacetaldehyde transaminase activity, a 2-aminomalonate semialdehyde transaminase activity, a 2,3-diaminopropanoate decarboxylase activity, a serine decarboxylase activity, an ethanolamine dehydrogenase activity, a serine hydroxymethyltransferase activity, an aldehyde oxidase activity, an N-acetyl transferase or O-acetyl transferase activity, an N-acetylserine dehydrogenase activity, a transaminase activity, a deacetylase activity, a serine aminase activity, and a 2,3-diaminopropanoate ammonia lyase activity, and the one or more co-product comprises ethylenediamine (EDA).

[0044] In some embodiments, the at least one enzyme for the production of MEG or GA from glycolaldehyde in a C2 pathway are selected from at least one enzyme having an

activity selected from a D-tagatose 3-epimerase activity, a D-ribulokinase activity, a D-ribulose-1-phosphate aldolase activity, a D-xylulose 1-kinase activity, a D-xylulose-1-phosphate aldolase activity, a xylose reductase or aldose reductase activity, a xylitol dehydrogenase activity, a xylose isomerase activity, a xylose dehydrogenase activity, a
 5 xylonolactonase activity, a xylonate dehydrogenase activity, a 2-keto-3-deoxy-D-pentionate aldolase activity, a glycolaldehyde reductase activity and a glycolaldehyde dehydrogenase activity. In some embodiments, the enzyme having D-xylulose-1-phosphate aldolase activity is aldolB. In some embodiments, the enzyme having D-ribulose-1-phosphate aldolase activity is fucA.

10 **[0045]** In some embodiments of any of the recombinant microorganisms described above, the recombinant microorganism further comprises one or more modifications to diminish or delete activity in a glycolaldehyde dehydrogenase, a lactate dehydrogenase, a xylose isomerase, a xylulose kinase or combination thereof.

15 **[0046]** In one embodiment, at least a portion of the excess NADH produced in the C3 pathway is used as a source of reducing equivalents in the C2 pathway. In another embodiment, at least a portion of the excess NADH produced in the C3 pathway is used to produce ATP.

[0047] In one embodiment, excess biomass formation is minimized and production of MEG (or glycolic acid) or MEG (or glycolic acid) and one or more co-products is maximized.

20 **[0048]** In another aspect, the application provides for a method of producing MEG or glycolic acid (GA) using a recombinant microorganism of any of the above embodiments, wherein the method comprises cultivating the recombinant microorganism in a culture medium containing one or more hexose feedstock providing a carbon source until the MEG or GA is produced. In some embodiments, one or more co-product is co-produced with MEG
 25 or GA. In further embodiments, the one or more co-product is selected from acetone, isopropanol, propene, isobutene and one or more serine pathway compound. In yet further embodiments, the one or more serine pathway compound is selected from serine, glycine, monoethanolamine (MEA) and ethylenediamine (EDA).

[0049] In yet another aspect, a method of producing a recombinant microorganism that
 30 produces or accumulates MEG or glycolic acid (GA) from one or more exogenous hexose feedstock via one or more pentose-5-phosphate intermediate, comprising: introducing into or expressing in the recombinant microorganism one or more enzyme for the conversion of the one or more hexose feedstock to the one or more pentose-5-phosphate intermediate; introducing into or expressing in the recombinant microorganism one or more enzyme for the
 35 conversion of the one or more pentose-5-phosphate intermediate to one or more pentose or

pentose-1-phosphate intermediate; introducing into or expressing in the recombinant microorganism one or more C2 pathway comprising one or more enzyme for the production of MEG or GA from glycolaldehyde; and introducing into or expressing in the recombinant microorganism one or more C3 pathway comprising one or more enzyme for the production of MEG or GA derived from DHAP or pyruvate; and culturing the recombinant microorganism in a culture medium containing the one or more hexose feedstock to produce or accumulate MEG or GA, wherein glycolaldehyde and DHAP (or pyruvate) are intermediates produced in the C2 pathway, and wherein MEG or GA are produced in both the C2 and C3 pathways.

[0050] In some embodiments, the application provides for a method of producing a recombinant microorganism that produces or accumulates MEG or glycolic acid (GA) and one or more co-product from one or more exogenous hexose feedstock via one or more pentose-5-phosphate intermediate, comprising: introducing into or expressing in the recombinant microorganism one or more enzyme for the conversion of the one or more hexose feedstock to the one or more pentose-5-phosphate intermediate; introducing into or expressing in the recombinant microorganism one or more enzyme for the conversion of the one or more pentose-5-phosphate intermediate to one or more pentose or pentose-1-phosphate intermediate; introducing into or expressing in the recombinant microorganism one or more C2 pathway comprising one or more enzyme for the production of MEG or GA from glycolaldehyde; and introducing into or expressing in the recombinant microorganism one or more C3 pathway comprising one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate; and culturing the recombinant microorganism in a culture medium containing the one or more hexose feedstock to produce or accumulate MEG or GA and one or more co-product, wherein glycolaldehyde and DHAP (or pyruvate) are intermediates produced in the C2 pathway, and wherein MEG or GA is produced in the one or more C2 pathway and the one or more co-product is produced in the one or more C3 pathway.

BRIEF DESCRIPTION OF DRAWINGS

[0051] Illustrative embodiments of the disclosure are illustrated in the drawings, in which:

[0052] FIG. 1 illustrates lossless transformation of glucose to a pentose-phosphate. The symbol ▼ means enzymes to be potentially down regulated or inactivated/abolished, ie. respective gene potentially attenuated or deleted

[0053] FIG. 2 illustrates MEG and possible co-production pathways via D-xylulose-1-phosphate.

[0054] FIG. 3 illustrates MEG and possible co-production pathways via D-xylonate.

[0055] FIG. 4 illustrates MEG and possible co-production pathways via D-ribulose-1-phosphate.

5 **[0056]** FIG. 5 illustrates a scheme of options for connecting pentose phosphate pathway and MEG production pathways via pentose phosphatases (PP).

[0057] FIG. 6 illustrates a scheme for producing pentose-1-phosphates and derivatives from the pentose phosphate pathway via arabitol phosphate dehydrogenases (APD).

[0058] FIG. 7 illustrates the isomerization of pentose phosphates mediated by arabitol phosphate dehydrogenases (APD).

10 **[0059]** FIG. 8 illustrates a scheme for producing pentose-1-phosphates and derivatives from the pentose phosphate pathway via phosphopentomutases (PPM).

[0060] FIG. 9 illustrates lossless transformation of glucose to pentose phosphates and an acetyl-CoA. The symbol $\blacktriangledown$ means enzymes to be potentially down regulated or inactivated/abolished, ie. respective gene potentially attenuated or deleted.

15 **[0061]** FIG. 10 illustrates an overview of the MEG and Ser, Gly, MEA, EDA co-production pathways.

[0062] FIG. 11 illustrates published EDA production pathways. From WO 2014/049382. Reaction F: direct L-serine amination via L-serine aminase. Reaction G: direct pyruvate amination via 2,3-diaminopropionate ammonia lyase.

20 **[0063]** FIG. 12 illustrates the biosynthesis pathway of xylitol from glucose with the use of phosphatase. The symbol $\blacktriangledown$ means enzymes to be potentially down regulated or inactivated/abolished, ie. respective gene potentially attenuated or deleted.

[0064] FIG. 13 illustrates the biosynthesis pathway of xylitol from glucose with the use of arabitol-P dehydrogenase (ADP). The symbol $\blacktriangledown$ means enzymes to be potentially down regulated or inactivated/abolished, ie. respective gene potentially attenuated or deleted.

25 **[0065]** FIG. 14 illustrates the biosynthesis pathway of xylitol from glucose with the use of phosphopentomutase (PPM). The symbol $\blacktriangledown$ means enzymes to be potentially down regulated or inactivated/abolished, ie. respective gene potentially attenuated or deleted.

[0066] FIG. 15 is a scheme for an *in vitro* phosphoglucomutase assay on its natural substrate, glucose-1P.

[0067] FIG. 16 is a scheme for *in vitro* conversion of pentose to glycolic acid precursors through key pentose-5P intermediates acted on by phosphoglucomutase enzymes.

5 [0068] FIG. 17 is a scheme for *in vivo* screening of key enzyme candidates for facilitating conversion of pentose sugars to glycolic acid.

[0069] FIG. 18 is a scheme for *in vitro* conversion of pentose to glycolic acid precursors through key pentose-1P intermediates acted on by aldolase enzymes.

10 DETAILED DESCRIPTION

Definitions

[0070] The following definitions and abbreviations are to be used for the interpretation of the disclosure.

15 [0071] As used herein and in the appended claims, the singular forms "a," "an," and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "an enzyme" includes a plurality of such enzymes and reference to "the microorganism" includes reference to one or more microorganisms, and so forth.

20 [0072] As used herein, the terms "comprises," "comprising," "includes," "including," "has," "having," "contains," "containing," or any other variation thereof, are intended to cover a non-exclusive inclusion. A composition, mixture, process, method, article, or apparatus that comprises a list of elements is not necessarily limited to only those elements but may include other elements not expressly listed or inherent to such composition, mixture, process, method, article, or apparatus. Further, unless expressly stated to the contrary, "or" refers to an inclusive "or" and not to an exclusive "or."

25 [0073] The terms "about" and "around," as used herein to modify a numerical value, indicate a close range surrounding that explicit value. If "X" were the value, "about X" or "around X" would indicate a value from 0.9X to 1.1X, or, in some embodiments, a value from 0.95X to 1.05X. Any reference to "about X" or "around X" specifically indicates at least the values X, 0.95X, 0.96X, 0.97X, 0.98X, 0.99X, 1.01X, 1.02X, 1.03X, 1.04X,
30 and 1.05X. Thus, "about X" and "around X" are intended to teach and provide written description support for a claim limitation of, e.g., "0.98X."

[0074] As used herein, the terms "microbial," "microbial organism," and "microorganism" include any organism that exists as a microscopic cell that is included within the domains of archaea, bacteria or eukarya, the latter including yeast and filamentous fungi, protozoa, algae, or higher Protista. Therefore, the term is intended to encompass prokaryotic or eukaryotic cells or organisms having a microscopic size and includes bacteria, archaea, and eubacteria of all species as well as eukaryotic microorganisms such as yeast and fungi. Also included are cell cultures of any species that can be cultured for the production of a chemical.

[0075] As described herein, in some embodiments, the recombinant microorganisms are prokaryotic microorganism. In some embodiments, the prokaryotic microorganisms are bacteria. "Bacteria", or "eubacteria", refers to a domain of prokaryotic organisms. Bacteria include at least eleven distinct groups as follows: (1) Gram-positive (gram+) bacteria, of which there are two major subdivisions: (1) high G+C group (*Actinomyces*, *Mycobacteria*, *Micrococcus*, others) (2) low G+C group (*Bacillus*, *Clostridia*, *Lactobacillus*, *Staphylococci*, *Streptococci*, *Mycoplasmas*); (2) Proteobacteria, e.g., Purple photosynthetic +non-photosynthetic Gram-negative bacteria (includes most "common" Gram-negative bacteria); (3) Cyanobacteria, e.g., oxygenic phototrophs; (4) Spirochetes and related species; (5) Planctomyces; (6) Bacteroides, Flavobacteria; (7) Chlamydia; (8) Green sulfur bacteria; (9) Green non-sulfur bacteria (also anaerobic phototrophs); (10) Radioresistant micrococci and relatives; (11) Thermotoga and Thermosipho thermophiles.

[0076] "Gram-negative bacteria" include cocci, nonenteric rods, and enteric rods. The genera of Gram-negative bacteria include, for example, *Neisseria*, *Spirillum*, *Pasteurella*, *Brucella*, *Yersinia*, *Francisella*, *Haemophilus*, *Bordetella*, *Escherichia*, *Salmonella*, *Shigella*, *Klebsiella*, *Proteus*, *Vibrio*, *Pseudomonas*, *Bacteroides*, *Acetobacter*, *Aerobacter*, *Agrobacterium*, *Azotobacter*, *Spirilla*, *Serratia*, *Vibrio*, *Rhizobium*, *Chlamydia*, *Rickettsia*, *Treponema*, and *Fusobacterium*.

[0077] "Gram positive bacteria" include cocci, nonsporulating rods, and sporulating rods. The genera of gram positive bacteria include, for example, *Actinomyces*, *Bacillus*, *Clostridium*, *Corynebacterium*, *Erysipelothrix*, *Lactobacillus*, *Listeria*, *Mycobacterium*, *Myxococcus*, *Nocardia*, *Staphylococcus*, *Streptococcus*, and *Streptomyces*.

[0078] The term "recombinant microorganism" and "recombinant host cell" are used interchangeably herein and refer to microorganisms that have been genetically modified to express or to overexpress endogenous enzymes, to express heterologous enzymes, such as those included in a vector, in an integration construct, or which have an alteration in expression of an endogenous gene. By "alteration" it is meant that the expression of the

gene, or level of a RNA molecule or equivalent RNA molecules encoding one or more polypeptides or polypeptide subunits, or activity of one or more polypeptides or polypeptide subunits is up regulated or down regulated, such that expression, level, or activity is greater than or less than that observed in the absence of the alteration. For example, the term "alter" can mean "inhibit," but the use of the word "alter" is not limited to this definition. It is understood that the terms "recombinant microorganism" and "recombinant host cell" refer not only to the particular recombinant microorganism but to the progeny or potential progeny of such a microorganism. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

[0079] The term "expression" with respect to a gene sequence refers to transcription of the gene and, as appropriate, translation of the resulting mRNA transcript to a protein. Thus, as will be clear from the context, expression of a protein results from transcription and translation of the open reading frame sequence. The level of expression of a desired product in a host cell may be determined on the basis of either the amount of corresponding mRNA that is present in the cell, or the amount of the desired product encoded by the selected sequence. For example, mRNA transcribed from a selected sequence can be quantitated by qRT-PCR or by Northern hybridization (see Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press (1989)). Protein encoded by a selected sequence can be quantitated by various methods, e.g., by ELISA, by assaying for the biological activity of the protein, or by employing assays that are independent of such activity, such as western blotting or radioimmunoassay, using antibodies that recognize and bind the protein. See Sambrook *et al.*, 1989, *supra*.

[0080] The term "polynucleotide" is used herein interchangeably with the term "nucleic acid" and refers to an organic polymer composed of two or more monomers including nucleotides, nucleosides or analogs thereof, including but not limited to single stranded or double stranded, sense or antisense deoxyribonucleic acid (DNA) of any length and, where appropriate, single stranded or double stranded, sense or antisense ribonucleic acid (RNA) of any length, including siRNA. The term "nucleotide" refers to any of several compounds that consist of a ribose or deoxyribose sugar joined to a purine or a pyrimidine base and to a phosphate group, and that are the basic structural units of nucleic acids. The term "nucleoside" refers to a compound (as guanosine or adenosine) that consists of a purine or pyrimidine base combined with deoxyribose or ribose and is found especially in nucleic acids. The term "nucleotide analog" or "nucleoside analog" refers, respectively, to a nucleotide or nucleoside in which one or more individual atoms have been replaced with a

different atom or with a different functional group. Accordingly, the term polynucleotide includes nucleic acids of any length, DNA, RNA, analogs and fragments thereof. A polynucleotide of three or more nucleotides is also called nucleotidic oligomer or oligonucleotide.

5 **[0081]** It is understood that the polynucleotides described herein include "genes" and that the nucleic acid molecules described herein include "vectors" or "plasmids." Accordingly, the term "gene", also called a "structural gene" refers to a polynucleotide that codes for a particular sequence of amino acids, which comprise all or part of one or more proteins or enzymes, and may include regulatory (non-transcribed) DNA sequences, such as
10 promoter sequences, which determine for example the conditions under which the gene is expressed. The transcribed region of the gene may include untranslated regions, including introns, 5'-untranslated region (UTR), and 3'-UTR, as well as the coding sequence.

[0082] The term "enzyme" as used herein refers to any substance that catalyzes or promotes one or more chemical or biochemical reactions, which usually includes enzymes
15 totally or partially composed of a polypeptide or polypeptides, but can include enzymes composed of a different molecule including polynucleotides.

[0083] As used herein, the term "non-naturally occurring," when used in reference to a microorganism organism or enzyme activity of the disclosure, is intended to mean that the microorganism organism or enzyme has at least one genetic alteration not normally
20 found in a naturally occurring strain of the referenced species, including wild-type strains of the referenced species. Genetic alterations include, for example, modifications introducing expressible nucleic acids encoding metabolic polypeptides, other nucleic acid additions, nucleic acid deletions and/or other functional disruption of the microorganism's genetic material. Such modifications include, for example, coding regions and functional fragments
25 thereof, for heterologous, homologous, or both heterologous and homologous polypeptides for the referenced species. Additional modifications include, for example, non-coding regulatory regions in which the modifications alter expression of a gene or operon. Exemplary non-naturally occurring microorganism or enzyme activity includes the hydroxylation activity described above.

30 **[0084]** The term "exogenous" as used herein with reference to various molecules, *e.g.*, polynucleotides, polypeptides, enzymes, etc., refers to molecules that are not normally or naturally found in and/or produced by a given yeast, bacterium, organism, microorganism, or cell in nature.

[0085] On the other hand, the term "endogenous" or "native" as used herein with
35 reference to various molecules, *e.g.*, polynucleotides, polypeptides, enzymes, etc., refers to

molecules that are normally or naturally found in and/or produced by a given yeast, bacterium, organism, microorganism, or cell in nature.

[0086] The term “heterologous” as used herein in the context of a modified host cell refers to various molecules, e.g., polynucleotides, polypeptides, enzymes, etc., wherein at least one of the following is true: (a) the molecule(s) is/are foreign (“exogenous”) to (i.e., not naturally found in) the host cell; (b) the molecule(s) is/are naturally found in (e.g., is “endogenous to”) a given host microorganism or host cell but is either produced in an unnatural location or in an unnatural amount in the cell; and/or (c) the molecule(s) differ(s) in nucleotide or amino acid sequence from the endogenous nucleotide or amino acid sequence(s) such that the molecule differing in nucleotide or amino acid sequence from the endogenous nucleotide or amino acid as found endogenously is produced in an unnatural (e.g., greater than naturally found) amount in the cell.

[0087] The term “homolog,” as used herein with respect to an original enzyme or gene of a first family or species, refers to distinct enzymes or genes of a second family or species which are determined by functional, structural, or genomic analyses to be an enzyme or gene of the second family or species which corresponds to the original enzyme or gene of the first family or species. Homologs most often have functional, structural, or genomic similarities. Techniques are known by which homologs of an enzyme or gene can readily be cloned using genetic probes and PCR. Identity of cloned sequences as homologs can be confirmed using functional assays and/or by genomic mapping of the genes.

[0088] A protein has “homology” or is “homologous” to a second protein if the amino acid sequence encoded by a gene has a similar amino acid sequence to that of the second gene. Alternatively, a protein has homology to a second protein if the two proteins have “similar” amino acid sequences. Thus, the term “homologous proteins” is intended to mean that the two proteins have similar amino acid sequences. In certain instances, the homology between two proteins is indicative of its shared ancestry, related by evolution. The terms “homologous sequences” or “homologs” are thought, believed, or known to be functionally related. A functional relationship may be indicated in any one of a number of ways, including, but not limited to: (a) degree of sequence identity and/or (b) the same or similar biological function. Preferably, both (a) and (b) are indicated. The degree of sequence identity may vary, but in one embodiment, is at least 50% (when using standard sequence alignment programs known in the art), at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least 98.5%, or at least about 99%, or at least 99.5%, or at least

99.8%, or at least 99.9%. Homology can be determined using software programs readily available in the art, such as those discussed in *Current Protocols in Molecular Biology* (F.M. Ausubel *et al.*, eds., 1987) Supplement 30, section 7.718, Table 7.71. Some alignment programs are MacVector (Oxford Molecular Ltd, Oxford, U.K.) and ALIGN Plus (Scientific and Educational Software, Pennsylvania). Other non-limiting alignment programs include Sequencher (Gene Codes, Ann Arbor, Michigan), AlignX, and Vector NTI (Invitrogen, Carlsbad, CA). A similar biological function may include, but is not limited to: catalyzing the same or similar enzymatic reaction; having the same or similar selectivity for a substrate or co-factor; having the same or similar stability; having the same or similar tolerance to various fermentation conditions (temperature, pH, etc.); and/or having the same or similar tolerance to various metabolic substrates, products, by-products, intermediates, etc. The degree of similarity in biological function may vary, but in one embodiment, is at least 1%, at least 2%, at least 3%, at least 4%, at least 5%, at least 6%, at least 7%, at least 8%, at least 9%, at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least 98.5%, or at least about 99%, or at least 99.5%, or at least 99.8%, or at least 99.9%, according to one or more assays known to one skilled in the art to determine a given biological function.

[0089] The term "variant" refers to any polypeptide or enzyme described herein. A variant also encompasses one or more components of a multimer, multimers comprising an individual component, multimers comprising multiples of an individual component (e.g., multimers of a reference molecule), a chemical breakdown product, and a biological breakdown product. In particular, non-limiting embodiments, an enzyme may be a "variant" relative to a reference enzyme by virtue of alteration(s) in any part of the polypeptide sequence encoding the reference enzyme. A variant of a reference enzyme can have enzyme activity of at least 10%, at least 30%, at least 50%, at least 80%, at least 90%, at least 100%, at least 105%, at least 110%, at least 120%, at least 130% or more in a standard assay used to measure enzyme activity of a preparation of the reference enzyme. In some embodiments, a variant may also refer to polypeptides having at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the full-length, or unprocessed enzymes of the present disclosure. In some embodiments, a variant may also refer to polypeptides having at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%,

at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the mature, or processed enzymes of the present disclosure.

[0090] The term "signal sequence" as used herein refers to an amino acid sequence that targets peptides and polypeptides to cellular locations or to the extracellular environment.

5 Signal sequences are typically at the N-terminal portion of a polypeptide and are typically removed enzymatically. Polypeptides that have their signal sequences are referred to as being full-length and/or unprocessed. Polypeptides that have had their signal sequences removed are referred to as being mature and/or processed.

[0091] The term "yield potential" as used herein refers to a yield of a product from a biosynthetic pathway. In one embodiment, the yield potential may be expressed as a percent
10 by weight of end product per weight of starting compound.

[0092] The term "thermodynamic maximum yield" as used herein refers to the maximum yield of a product obtained from fermentation of a given feedstock, such as glucose, based on the energetic value of the product compared to the feedstock. In a normal fermentation,
15 without use of additional energy sources such as light, hydrogen gas or methane or electricity, for instance, the product cannot contain more energy than the feedstock. The thermodynamic maximum yield signifies a product yield at which all energy and mass from the feedstock is converted to the product. This yield can be calculated and is independent of a specific pathway. If a specific pathway towards a product has a lower yield than the
20 thermodynamic maximum yield, then it loses mass and can most likely be improved upon or substituted with a more efficient pathway towards the product.

[0093] The term "redox balanced" refers to a set of reactions, which taken together produce as much redox cofactors as they consume. Designing metabolic pathways and engineering an organism such that the redox cofactors are balanced or close to being
25 balanced usually results in a more efficient, higher yield production of the desired compounds. Redox reactions always occur together as two half-reactions happening simultaneously, one being an oxidation reaction and the other a reduction reaction. In redox processes, the reductant transfers electrons to the oxidant. Thus, in the reaction, the reductant or reducing agent loses electrons and is oxidized, and the oxidant or oxidizing
30 agent gains electrons and is reduced. In one embodiment, the redox reactions take place in a biological system. Biological energy is frequently stored and released by means of redox reactions. Photosynthesis involves the reduction of carbon dioxide into sugars and the oxidation of water into molecular oxygen. The reverse reaction, respiration, oxidizes sugars to produce carbon dioxide and water. As intermediate steps, the reduced carbon compounds

are used to reduce nicotinamide adenine dinucleotide (NAD⁺), which then contributes to the creation of a proton gradient, which drives the synthesis of adenosine triphosphate (ATP) and is maintained by the reduction of oxygen. The term redox state is often used to describe the balance of GSH/GSSG, NAD⁺/NADH and NADP⁺/NADPH in a biological system such as a cell or organ. The redox state is reflected in the balance of several sets of metabolites (e.g., lactate and pyruvate, beta-hydroxybutyrate, and acetoacetate), whose interconversion is dependent on these ratios. An abnormal redox state can develop in a variety of deleterious situations, such as hypoxia, shock, and sepsis.

[0094] The terms “C2 pathway”, “C2 branch pathway”, “C2 biochemical pathway” or “C2 stream” as used herein refers to a biochemical pathway wherein MEG can be produced via glycolaldehyde.

[0095] The terms “C3 pathway”, “C3 branch pathway”, “C3 biochemical pathway” or “C3 stream” as used herein refers to a biochemical pathway wherein MEG and/or one or more co-product such as acetone, isopropanol, propene, isobutene and/or serine pathway compounds can be produced via pyruvate, acetyl-CoA or dihydroxyacetonephosphate (DHAP).

Introduction

[0096] This disclosure combines the advantages of xylose degradation biochemistry for high yielding MEG (or glycolic acid) or MEG (or glycolic acid) and co-product formation with the advantages of readily available pure hexose sugar feedstocks. It does so by providing a lossless conversion of the hexose glucose to the intermediate D-xylulose-5-phosphate, and further to the pentoses D-xylulose or D-ribulose, to be used as intermediates for the production of MEG (or glycolic acid) or MEG (or glycolic acid) and a co-product by any one of the previously described D-xylose based methods. Apart from glucose, other hexoses such as fructose or hexose oligosaccharides such as starch or sucrose or cellobiose can be used. In some embodiments, hexoses may be selected from D-allose, D-altrose, D-glucose, D-mannose, D-gulose, D-idose, D-galactose, D-talose, D-tagatose, D-sorbose, D-fructose, D-psicose, and other hexoses known in the art. In some embodiments, pentoses may be selected from D-xylose, D-ribose, D-arabinose, D-lyxose, D-xylulose, D-ribulose, and other pentoses known in the art. In some embodiments, the hexoses and pentoses may be selected from the levorotary or dextrorotary enantiomer of any of the hexoses and pentoses disclosed herein.

[0097] Compared to other glucose based MEG or glycolic acid production methods, the present methods solve the following problems: ATP shortage, if co-production is utilized; large NADH excess; low overall product yield potential.

[0098] Compared to other xylose based MEG or glycolic acid production methods, the present methods solve the following challenges and problems: a process depending on xylose (availability/ market limitations, price, purity); glucose induced inhibition of xylose utilization.

5 **[0099]** Fermentative MEG production is described in WO2010/076324 (or US2011/0294178; Metabolic Explorer), which is herein incorporated in its entirety. This application suggested diol production via 2-ketoacid decarboxylation and reduction, including a serine biosynthesis based pathway to the intermediate hydroxypyruvate and further to ethyleneglycol. However, the disclosed pathway has a reduced total yield potential of 0.69
10 **g_MEG/ g_glucose, while the thermodynamic maximum yield for a glucose → MEG** conversion is 0.82 g/g. This pathway is also not redox balanced and has a high excess of 2 mol NADH per mol of consumed glucose, all of which needs to be re-oxidized for the cell to be viable. In an aerobic fermentation, this NADH can be used to generate ATP, which
15 **however would be in high excess (2 NADH → 6 ATP), leading to excess biomass formation** during the production phase and therefore reduced product formation and yield.

[00100] Thus, the fermentative MEG production pathway disclosed in WO2010/076324 has an ATP shortage (-1ATP per MEG), excess NADH (+1 NADH per MEG), low yield potential ($y_{max} = 0.69 \text{ g\_MEG/ g\_glucose}$) and is a challenging pathway that has not been demonstrated at high efficiency/productivity.

20 **[00101]** The disclosure of WO2011/130378A1 (or US2011/0312049; Genomatica) proposes an approach similar to WO2010/076324 to produce MEG from glucose via hydroxypyruvate, but also mentions pathway variations with alternative, but related key intermediates glycerate or ethanolamine.

[00102] The disclosure of WO2011/130378A1 has the same drawbacks as
25 WO2010/076324, except for ATP shortage. ATP can be +0 or +1 per MEG, depending on utilized enzymes.

[00103] In the case of glycolic acid, the described pathways from glucose also go through 3-phosphoglycerate and serine pathway reactions, or via the glyoxylate shunt. In both cases, one CO₂ is lost per glycolic acid, leading to a maximum yield (0.84 g/g) much lower than the
30 thermodynamic maximum yield potential (1.7 g/g).

[00104] A demonstrated fermentative production of MEG from xylose (WO2013/126721), via ribulose-1-phosphate, has a high yield potential (0.82 g_MEG/g_xylose) which equals the thermodynamic maximum yield. It produces MEG via two different pathways which are active in parallel, a 2-carbon stream (via glycolaldehyde) and a 3-carbon stream (via

dihydroxyacetonephosphate). The C2 stream is easy to implement, but the C3 stream is difficult to implement at high efficiency via metabolic engineering. The C3 stream utilizes the pathways presented in WO2010/076324 or WO2011/130378.

[00105] Assuming a typically ATP driven xylose import, the overall process is at least ATP neutral. Thus, some xylose and therefore yield will be lost in order to obtain some surplus ATP required for cell growth and maintenance.

[00106] However, the uptake of xylose is not as efficient and fast as that of glucose, the preferred carbon source of most microorganisms. Also, presence of glucose in the media usually inhibits utilization of other sugars such as xylose. For a more efficient process, the organism's regulation leading to this preferential consumption needs to be disrupted and the strain adapted towards xylose preference or sugar co-consumption.

[00107] The key challenge, however, is obtaining xylose as affordable and clean feedstock. Xylose as pure chemical is expensive and not available in bulk quantities. Xylose in hemicellulose hydrolysates is available in larger quantities and at potentially lower cost than glucose, but is accompanied by many impurities and substances that inhibit fermentations.

[00108] Therefore, the fermentative production of MEG (or glycolic acid) from xylose (WO2013/126721) poses a challenge with respect to using xylose as feedstock (availability, price, purity, inhibition of xylose utilization by glucose) and to using a C3 pathway, which has not been demonstrated at high efficiency/productivity. Moreover, there is an ATP shortage, +0ATP (or -1ATP if not using glycerate kinase), which is not sufficient for cell maintenance.

[00109] A further demonstrated fermentative production of MEG from xylose (Alkim *et al.*, Microb Cell Fact (2015) 14:127), via xylulose-1-phosphate, is very similar to the route described by WO2013/126721. It has the same high yield potential (0.82 g/g), difficult to implement C3 stream for MEG production via DHAP, ATP shortage and feedstock challenges.

[00110] Another demonstrated fermentative production of MEG from xylose (WO2013/119020), via xylonate, shares similarities to the route described by WO2013/126721. It produces glycolaldehyde and pyruvate as key intermediates, allowing MEG production from glycolaldehyde with a yield potential of 0.41 g/g. This represents a high relative yield since it is achieved with only half the flux. However, no pathway to convert the remaining pyruvate to MEG is presented in WO2013/119020 or elsewhere. Currently, no realistic and efficient pathway is known to convert pyruvate to MEG. While pyruvate as a co-product itself would enable a redox neutral overall process (+0 NADH), it is not an

economically interesting product and the process would lack 1 ATP (probably ~2 ATP more for pyruvate export). Thus, ideally a pyruvate derived, economically interesting co-product at high yield is required that delivers surplus ATP. Therefore, the fermentative production of MEG from xylose (WO2013/119020), via xylonate, poses a challenge with respect to using xylose as feedstock (availability, price, purity, inhibition of xylose utilization by glucose), low absolute yield of MEG, ATP shortage (depending on co-product, it could be -1 to -3 ATP with pyruvate), and to requiring a pyruvate derived co-product with high yield potential and surplus ATP.

[00111] U.S. Application No. 62/305,814, U.S. Application No. 62/430,742 and U.S. Application No. 62/406,684, each of which is herein incorporated in its entirety, describe high yielding, easy to implement pathways for the co-production of MEG and compounds such as acetone, isopropanol (IPA), propene or isobutene and offer solutions to most of the challenges described for the above mentioned MEG production methods. They furthermore offer solutions for challenges encountered for previously described IPA or isobutene productions from glucose.

[00112] All currently known MEG (or glycolic acid) production methods using glucose as feedstock have low yield potential. This is an intrinsic drawback of the biochemistry of how glucose is degraded to MEG, with one decarboxylation occurring per produced MEG (or glycolic acid) molecule for all the proposed and known pathways. However, one decarboxylation per MEG is too much to achieve redox-neutral and therefore optimal yield.

[00113] All the presented MEG (or glycolic acid) production methods using xylose as feedstock have high or high relative yield potential, due to utilization of specific, beneficial xylose degradation biochemistry, leading to or close to an ideal 0.5 decarboxylation reactions per MEG molecule. However, all these methods share the challenges of using xylose as a feedstock, such as its market limitations and technical challenges from feedstock impurity.

[00114] The present disclosure combines the advantages of xylose degradation biochemistry for high yielding MEG (or glycolic acid), or optionally, MEG (or glycolic acid) and one or more co-product, formation with the advantages of readily available pure hexose sugar feedstocks. It does so by providing a lossless conversion of the hexose glucose to the intermediate D-xylulose-5-phosphate, and further to the pentoses D-xylulose or D-ribulose, to be used as intermediates for the production of MEG (or glycolic acid), or optionally, MEG (or glycolic acid) and one or more co-product, by any one of the previously described D-xylose based methods. Glycolic acid (GA) can be produced instead of MEG by oxidation of

glycolaldehyde. Apart from glucose, other hexoses such as fructose or hexose oligosaccharides such as starch or sucrose can be used.

[00115] Compared to other glucose based MEG (or glycolic acid) production methods, the recombinant microorganisms and methods of the present disclosure solves: the problem of ATP shortage, if co-production is utilized; the problem of large NADH excess; the problem of low overall product yield potential.

[00116] Compared to other xylose based MEG (or glycolic acid) production methods, the recombinant microorganisms and methods of the present disclosure solves the challenges of a process depending on xylose (availability/ market limitations, price, purity) and the problem of glucose induced inhibition of xylose utilization.

[00117] The present disclosure relates to the production of MEG, or optionally, MEG and one or more co-product, from a hexose, preferably in *E. coli*. Alternatively, glycolic acid (GA) can be produced instead of MEG by oxidation of glycolaldehyde.

[00118] If glucose is utilized by glycolysis, the standard degradation pathway of *E. coli* and most other organisms, it will be degraded via 3-phosphoglycerate, the common key intermediate for all glucose to MEG methods described so far. However, this 3-carbon compound is degraded to one 2-carbon compound (MEG), losing one CO₂ per MEG, which is true for all described pathway variations, where NADH is in excess and means significant loss of yield potential (only 0.69 g_MEG per gram of sugar, vs 0.82 g_MEG thermodynamic maximum yield potential).

[00119] *Utilization of a non-oxidative entry into pentose phosphate pathway*

[00120] In the present disclosure, glucose flux is funneled into the pentose phosphate pathway instead of the glycolysis pathway (**FIG. 1**). A transketolase, such as encoded by *tktA* or *tktB* from *E. coli*, is used as a non-oxidative entry into the pentose phosphate pathway to transform the glycolysis intermediates fructose-6-phosphate and glyceraldehyde-3-phosphate into D-xylulose-5-phosphate and D-erythrose-4-phosphate (**FIG. 1**). This produces the key intermediate D-xylulose-5-phosphate.

[00121] Alternatively, a fructose 6-phosphate phosphoketolase (Fpk) and a phosphate acetyltransferase (PTA) can be used as entry into the pentose phosphate pathway, making one erythrose-4-phosphate and one acetyl-CoA from fructose-6-phosphate (**FIG. 9**).

[00122] *Utilization of pentose phosphate pathway for full conversion into pentose intermediates*

[00123] For a lossless conversion of all hexose carbons into a pentose intermediate, D-erythrose-4-phosphate needs to be further processed. A transaldolase, such as encoded by talA or talB from *E. coli*, catalyzes the conversion of D-erythrose-4-phosphate and D-fructose-6-phosphate to generate D-seduheptulose-7-phosphate and D-glyceraldehyde-3-phosphate. These intermediates are further processed by a transketolase, such as encoded by tktA or tktB from *E. coli*, to generate D-ribose-5-phosphate and D-xylulose-5-phosphate. D-ribose-5-phosphate can be readily converted into D-ribulose-5-phosphate and further to D-xylulose-5-phosphate via ribose-5-phosphate isomerase (such as encoded by rpiA or rpiB in *E. coli*) and ribulose-5-phosphate 3-epimerase (such as encoded by rpe in *E. coli*), respectively. Thus, all glucose or fructose can be converted completely into D-xylulose-5-phosphate. The overall stoichiometry is:

[00124] 2.5 glucose + 2.5 ATP + 0.5 phosphate → 2 D-xylulose-5-phosphate + D-ribose-5-phosphate

[00125] The bidirectional isomerase and epimerase reactions can of course also transform the two produced D-xylulose-5-phosphates and one D-ribose-5-phosphate molecule into three D-ribulose-5-phosphates. The overall net transformation depends lastly on which intermediate, D-xylulose-5-phosphate or D-ribulose-5-phosphate, is actually consumed by the following pathway.

[00126] In case of the alternative entry into the pentose phosphate pathway via Fpk (FIG. 9), the stoichiometry is:

[00127] 2 glucose + 2 ATP + CoA → 2 D-xylulose-5-phosphate + 1 acetyl-CoA

[00128] *Optimization of flux towards non-oxidative entry into pentose phosphate pathway*

[00129] Deactivation of oxidative branch of pentose phosphate pathway

[00130] The common pathway in *E. coli*, the oxidative entry into the pentose phosphate pathway via 6-phospho D-glucono-1,5-lactone and oxidative decarboxylation to D-ribulose-5-phosphate, should not be utilized, since it transforms the hexose glucose into a pentose with loss of one carbon (FIG. 1). It is advantageous to inhibit at least one or more enzymes catalyzing one or more appropriate reactions in the oxidative branch of the pentose phosphate pathway, namely glucose 6-phosphate-1-dehydrogenase, 6-phosphogluconolactonase, and 6-phosphogluconate dehydrogenase, by deleting one or more of the genes responsible. For example, in *E. coli*, the genes are zwf (glucose 6-phosphate-1-dehydrogenase), pgl (6-phosphogluconolactonase), and gnd (6-phosphogluconate dehydrogenase).

[00131] Downregulation of glycolysis downstream reactions

[00132] The upper part of glycolysis is needed to transform 2.5 glucose or fructose into the key intermediates 2x fructose-6-phosphate and 1x glyceraldehyde-3-phosphate. To reduce or eliminate further flux through the lower part of glycolysis, i.e. the oxidative phosphorylation of glyceraldehyde-3-phosphate to 1,3-bisphospho D-glycerate and its subsequent conversion to 3-phospho-D-glycerate and 2-phospho-D-glycerate, activity of glyceraldehyde 3-phosphate dehydrogenase, phosphoglycerate kinase and phosphoglycerate mutase, by gapA, pgk and gpmA/gpmM, respectively, in *E. coli* can be diminished.

[00133] However, if the alternative entry into the pentose phosphate pathway via fructose 6-phosphate phosphoketolase (Fpk) is utilized, then no glyceraldehyde 3-phosphate is needed and the appropriate 6-phosphofructokinase activity can be diminished or deleted (pfkA and/or pfkB genes in *E. coli*).

[00134] *Connecting pentose-phosphate intermediates with MEG production routes*

[00135] To connect the pentose phosphate pathway intermediate D-xylulose-5-phosphate with any one of the known MEG or glycolic acid production pathways, the D-xylulose-5-phosphate intermediate needs to be dephosphorylated by a pentose 5-phosphatase to generate D-xylulose (**FIG. 2** and **FIG. 5**). Similarly, the D-ribulose-5-phosphate intermediate may be connected to any one of the known MEG or glycolic acid production pathways through dephosphorylation by a pentose 5-phosphatase to generate D-ribulose (**FIG. 4** and **FIG. 5**).

[00136] In the case of a xylonate based pathway to MEG, the xylose isomerase function, such as XylA from *E. coli*, is further required to transform the generated D-xylulose into D-xylose (**FIG. 3**).

[00137] *Alternatives and variations in connecting pentose-phosphate pathway with MEG production routes*

[00138] In the case of a D-ribulose-1-phosphate based pathway, the D-xylulose-5-phosphate intermediate can be connected to the pentose phosphate pathway as described via D-xylulose formation, following epimerization and phosphorylation (dte and fucK mediated) (**FIG. 4**). Alternatively, D-ribulose-5-phosphate can be used as an entry point and degraded to D-ribulose by utilizing a D-ribulose 5-phosphatase. This way, the necessary isomerization reaction is performed at the level of RPE rather than DTE (**FIG. 4**).

[00139] In some embodiments, pentose-5-phosphates can be transformed into pentose-1-phosphates using two arabitol dehydrogenase reactions in a row (**FIG. 6**). Here, the 2-keto position of D-xylulose-5-phosphate or D-ribulose-5-phosphate is reduced, then the 4-hydroxy

position is oxidized to a keto group, turning the pentose-5-phosphates into pentose-1-phosphates (FIG. 7).

[00140] In further embodiments, pentose-5-phosphates can be transformed directly into pentose-1-phosphates through the action of a pentose phosphomutase (also known as phosphopentomutase or PPM). Similar to a phosphoglucomutase or phosphomannomutase, it transfers the phosphate residue from the last to the 1-position of the phospho sugar.

[00141] *Hexose utilization*

[00142] The entry molecules of the pathway in the present disclosure are fructose-6-phosphate and glyceraldehyde-3-phosphate, both of which are obtained from normal glycolytic degradation of glucose or fructose in most organisms. If an organism has the ability to consume starch or sucrose or cellulose, for instance, via expression of a sucrose invertase, it will still generate glucose (and fructose), enabling it to utilize the methods of the present disclosure in the same way and to the same extent, yielding the same benefits.

[00143] Monoethylene glycol (MEG)

[00143] Monoethylene glycol (MEG) is an important raw material for industrial applications. A primary use of MEG is in the manufacture of polyethylene terephthalate (PET) resins, films and fibers. In addition, MEG is important in the production of antifreezes, coolants, aircraft anti-icer and deicers and solvents. MEG is also known as ethane-1,2-diol.

[00144] Ethylene glycol is also used as a medium for convective heat transfer in, for example, automobiles and liquid cooled computers.

[00145] Because of its high boiling point and affinity for water, ethylene glycol is a useful desiccant. Ethylene glycol is widely used to inhibit the formation of natural gas clathrates (hydrates) in long multiphase pipelines that convey natural gas from remote gas fields to a gas processing facility. Ethylene glycol can be recovered from the natural gas and reused as an inhibitor after purification treatment that removes water and inorganic salts.

[00146] Minor uses of ethylene glycol include in the manufacture of capacitors, as a chemical intermediate in the manufacture of 1,4-dioxane, and as an additive to prevent corrosion in liquid cooling systems for personal computers. Ethylene glycol is also used in the manufacture of some vaccines; as a minor ingredient in shoe polish, inks and dyes; as a rot and fungal treatment for wood; and as a preservative for biological specimens.

[00147] Glycolic acid

[00147] Glycolic acid is used in the textile industry as a dyeing and tanning agent, in food processing as a flavoring agent and as a preservative, and in the pharmaceutical industry as a skin care agent. It is also used in adhesives and plastics. Glycolic acid is often included into emulsion polymers, solvents and additives for ink and paint in order to improve flow properties and impart gloss. It is used in surface treatment products that increase the coefficient of friction on tile flooring.

[00148] Due to its excellent capability to penetrate skin, glycolic acid finds applications in skin care products to improve the skin's appearance and texture. It can be used as a chemical peel performed by a dermatologist in concentrations of 20 to 70% or at-home kits in lower concentrations between 10 and 20%. In addition to concentration, pH also plays a large part in determining the potency of glycolic acid in solution.

[00149] Glycolic acid can be synthesized in various ways. The predominant approach uses a catalyzed reaction of formaldehyde with synthesis gas (carbonylation of formaldehyde), for its low cost. It is also prepared by the reaction of chloroacetic acid with sodium hydroxide followed by re-acidification. Other methods, not noticeably in use, include hydrogenation of oxalic acid, and hydrolysis of the cyanohydrin derived from formaldehyde. Some of today's glycolic acids are formic acid-free. Glycolic acid can be isolated from natural sources, such as sugarcane, sugar beets, pineapple, cantaloupe and unripe grapes.

[00150] Glycolic acid is a useful intermediate for organic synthesis, in a range of reactions including: oxidation-reduction, esterification and long chain polymerization. It is used as a monomer in the preparation of polyglycolic acid and other biocompatible copolymers (e.g. PLGA). Commercially, important derivatives include the methyl (CAS# 96-35-5) and ethyl (CAS# 623-50-7) esters which are readily distillable. The butyl ester is a component of some varnishes, being desirable because it is nonvolatile and has good dissolving properties.

Acetone

[00151] Acetone (also known as propanone) is an organic compound with the formula $(CH_3)_2CO$. It is a colorless, volatile, flammable liquid, and is the simplest ketone.

[00152] Acetone is miscible with water and serves as an important solvent, typically for cleaning purposes in the laboratory. Over 6.7 million tonnes are produced worldwide, mainly for use as a solvent and production of methyl methacrylate and bisphenol A. It is a common building block in organic chemistry. Familiar household uses of acetone are as the active ingredient in nail polish remover and as paint thinner.

Isopropanol

[00153] Isopropyl alcohol (IUPAC name 2-propanol), also called isopropanol, is a compound with the chemical formula C_3H_8O or C_3H_7OH or $CH_3CHOHCH_3$. It is a colorless, flammable chemical compound with a strong odor. It is the simplest example of a secondary alcohol, where the alcohol carbon atom is attached to two other carbon atoms sometimes shown as $(CH_3)_2CHOH$. It is a structural isomer of propanol. It has a wide variety of industrial and household uses.

[00154] Propene, also known as propylene or methyl ethylene, is an unsaturated organic compound having the chemical formula C_3H_6 . It has one double bond, and is the second simplest member of the alkene class of hydrocarbons.

[00155] Propene is produced from fossil fuels—petroleum, natural gas, and, to a much lesser extent, coal. Propene is a byproduct of oil refining and natural gas processing.

Isobutene

[00156] Isobutene (also known as isobutylene or 2-methylpropene) is a hydrocarbon of industrial significance. It is a four-carbon branched alkene (olefin), one of the four isomers of butylene (butene). At standard temperature and pressure it is a colorless flammable gas.

[00157] Isobutene is used as an intermediate in the production of a variety of products. It is reacted with methanol and ethanol in the manufacture of the gasoline oxygenates methyl tert-butyl ether (MTBE) and ethyl tert-butyl ether (ETBE), respectively. Alkylation with butane produces isooctane, another fuel additive. Isobutene is also used in the production of methacrolein. Polymerization of isobutene produces butyl rubber (polyisobutene). Antioxidants such as butylated hydroxytoluene (BHT) and butylated hydroxyanisole (BHA) are produced by Friedel-Crafts alkylation of phenols using isobutene.

[00158] Polymer and chemical grade isobutene is typically obtained by dehydrating tertiary butyl alcohol or catalytic dehydrogenation of isobutane. Gasoline oxygenates MTBE and ETBE are generally produced by reacting methanol or ethanol with isobutene contained in butene streams from olefin steam crackers or refineries. Isobutene is not isolated before the reaction as separating the ethers from the remaining butenes is simpler.

Serine pathway compounds

[00159] Compounds that may be co-produced with MEG (or glycolic acid) include serine pathway compounds, for example, serine, glycine, monoethanolamine (MEA) and ethylenediamine (EDA).

[00160] Serine is a non-essential amino acid that can be synthesized in the human body. Being highly water soluble, serine finds application as moisturizer in lotions of pharma and cosmetic industry. Further, there is a huge market for serine in the chemical industry because it can be converted into other chemicals such as plastics, detergents, dietary supplements and a variety of other products. In fact, serine has been mentioned as one of the 30 most promising biological substances to replace chemicals from the oil industry.

[00161] The α -decarboxylation of serine yields ethanolamine, an industrial product used as an intermediate in the herbicide, textile, metal, detergent, plastics, and personal care products industries with a production volume running into several hundreds of kilotonnes per annum (Scott, E. *et al.* (2007) Biomass in the manufacture of industrial products—the use of proteins and amino acids. *Appl Microbiol Biotechnol.* 75(4): 751–762).

[00162] Glycine, the simplest amino acid, is valuable for pharmaceutical and industrial applications. It is included as an additive in pet food and animal feed. For humans, glycine is sold as a sweetener/taste enhancer. Certain food supplements and protein drinks contain glycine. Certain drug formulations include glycine to improve gastric absorption of the drug. Glycine serves as a buffering agent in antacids, analgesics, antiperspirants, cosmetics, and toiletries. Many miscellaneous products use glycine or its derivatives, such as the production of rubber sponge products, fertilizers and metal complexants. Glycine is also valuable as an intermediate in the synthesis of a variety of chemical products. It is used in the manufacture of the herbicide glyphosate. Glycine can be converted to oxalic acid, which is used as a bleaching agent in the textile and pulp industries and wastewater treatment. Glycine is also extensively used in laboratory research, for example, in gel electrophoresis.

[00163] Ethylenediamine (EDA) (1,2-diaminoethane, $C_2H_4(NH_2)_2$) is used in large quantities for production of many industrial chemicals. It forms derivatives with carboxylic acids (including fatty acids), nitriles, alcohols (at elevated temperatures), alkylating agents, carbon disulfide, and aldehydes and ketones. Because of its bifunctional nature, having two amines, it readily forms heterocycles such as imidazolidines. A most prominent derivative of ethylenediamine is the chelating agent EDTA, which is derived from ethylenediamine via a Strecker synthesis involving cyanide and formaldehyde. Hydroxyethylethylenediamine is another commercially significant chelating agent. Numerous bio-active compounds and drugs contain the N-CH₂-CH₂-N linkage, including some antihistamines. Salts of ethylenebisdithiocarbamate are commercially significant fungicides under the brand names Maneb, Mancozeb, Zineb, and Metiram. Some imidazoline-containing fungicides are derived from ethylenediamine. Ethylenediamine is an ingredient in the common bronchodilator drug aminophylline, where it serves to solubilize the active ingredient theophylline.

Ethylenediamine has also been used in dermatologic preparations. When used as a pharmaceutical excipient, after oral administration its bioavailability is about 0.34, due to a substantial first-pass effect. Less than 20% is eliminated by urinal excretion. Ethylenediamine, because it contains two amine groups, is a widely used precursor to various polymers. Condensates derived from formaldehyde are plasticizers. It is widely used in the production of polyurethane fibers. The PAMAM class of dendrimers are derived from ethylenediamine. The bleaching activator tetraacetythylenediamine is generated from ethylenediamine. The derivative N,N-ethylenebis(stearamide) (EBS) is a commercially significant mold-release agent and a surfactant in gasoline and motor oil.

[00164] Ethylenediamine is also used as: a solvent to solubilize proteins such as albumins and casein; certain electroplating baths; corrosion inhibitor in paints and coolants; chemicals for color photography developing, binders, adhesives, fabric softeners, curing agents for epoxys, and dyes. Ethylenediamine dihydroiodide (EDDI) is added to animal feeds as a source of iodide.

Xylitol

[00165] Xylitol is a chemical compound, a sugar alcohol, of considerable value as a sweetener. It is detected by the human palate as approximately as sweet as sucrose, and it is non-toxic and non-cariogenic.

[00166] One method of producing xylitol utilizes xylan, a hemicellulose, which is extracted from hardwoods and corncobs. Xylans can be hydrolyzed into xylose, which is then catalytically hydrogenated into xylitol. Catalytic routes to xylitol are cost- and energy-intensive due to the use of extensive separation and purification steps involved in the process and have low overall yield. Another method of producing xylitol includes utilizing fermentative and biocatalytic processes in bacteria, fungi, and/or yeast cells.

Enzymes

[00167] Exemplary enzymes that may be used in the MEG (or glycolic acid), or optionally, MEG (or glycolic acid) and one or more co-product, biosynthesis pathways disclosed herein are listed in **Table 1**.

[00168] Table 1

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
Isomerases that may be used in D-xylonate pathways of the present disclosure									
D-xylose + NAD(P)H => Xylitol + NAD(P)+	1.1.1.307	xylose reductase	xyl1	<i>Scheffersomyces stipitis</i>	D-xylose reductase	GeneID: 4839234	82, 83	P31867	84
D-xylose + NAD(P)H => Xylitol + NAD(P)+	1.1.1.307	xylose reductase	GRE3	<i>Saccharomyces cerevisiae</i>	aldose reductase	GeneID:856504	85, 86	P38715	87
Xylitol + NAD+ => D-xylulose + NADH	1.1.1.9	xylitol dehydrogenase	xyl2	<i>Scheffersomyces stipitis</i>	D-xylulose reductase	GeneID: 4852013	88, 89	P22144	90
Xylitol + NAD+ => D-xylulose + NADH	1.1.1.9	xylitol dehydrogenase	xdh1	<i>Trichoderma reesei</i>	Xylitol dehydrogenase	ENA Nr.: AF428150.1	91	Q876R2	92
D-xylopyranose => D-xylulose	5.3.1.5	xylose isomerase	xylA	<i>Pyromyces</i> sp.	xylose isomerase	ENA Nr.: CAB76571.1	93, 94	Q9P8C9	95
D-xylopyranose => D-xylulose	5.3.1.5	xylose isomerase	xylA	<i>Escherichia coli</i>	xylose isomerase	GeneID: 948141	143	P00944	144
Glycolaldehyde reductases that may be used in all MEG pathways									
glycolaldehyde + NAD(P)H => monoethylene glycol + NAD(P)+	1.1.1.-	glycolaldehyde reductase	gldA	<i>Escherichia coli</i>	glycerol dehydrogenase	GeneID:12933659	12	P0A9S5	13

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
glycolaldehyde + NAD(P)H $\rightleftharpoons$ monoethylene glycol + NAD(P) ⁺	1.1.1. -	glycolaldehyde reductase	GRE2	Saccharomyces cerevisiae	methylglyoxal reductase	GeneID:854014	14	Q12068	15
glycolaldehyde + NAD(P)H $\rightleftharpoons$ monoethylene glycol + NAD(P) ⁺	1.1.1. -	glycolaldehyde reductase	GRE3	Saccharomyces cerevisiae	aldose reductase	GeneID:856504	16	P38715	17
glycolaldehyde + NAD(P)H $\rightleftharpoons$ monoethylene glycol + NAD(P) ⁺	1.1.1. -	glycolaldehyde reductase	yqhD*	Escherichia coli	Alcohol dehydrogenase	GeneID:947493	18, 19	Modified version of Q46856; G149E	20
glycolaldehyde + NAD(P)H $\rightleftharpoons$ monoethylene glycol + NAD(P) ⁺	1.1.1. -	glycolaldehyde reductase	yqhD	Escherichia coli	Alcohol dehydrogenase	GeneID:947493	21, 22	Q46856	23
glycolaldehyde + NAD(P)H $\rightleftharpoons$ monoethylene glycol + NAD(P) ⁺	1.1.1. -	glycolaldehyde reductase	ydjg	Escherichia coli	methylglyoxal reductase	GeneID:12930149	24	P77256	25

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
glycolaldehyde + NAD(P)H $\rightleftharpoons$ monoethylene glycol + NAD(P) ⁺	1.1.1. -	glycolaldehyde reductase	fucO	Escherichia coli	lactaldehyde reductase	GeneID:947273	26, 27	P0A9S1	28
glycolaldehyde + NAD(P)H $\rightleftharpoons$ monoethylene glycol + NAD(P) ⁺	1.1.1. -	glycolaldehyde reductase	yafB (dkgB)	Escherichia coli	methylglyoxal reductase [multifunctional]	545778205	29	P30863	30
glycolaldehyde + NAD(P)H $\rightleftharpoons$ monoethylene glycol + NAD(P) ⁺	1.1.1. -	glycolaldehyde reductase	yqhE (dkgA)	Escherichia coli	2,5-diketo-D-gluconic acid reductase A	GeneID: 947495	31	Q46857	32
Enzymes that may be used in D-ribulose-1-phosphate pathway to MEG									
D-xylulose $\rightleftharpoons$ D-ribulose	5.1.3. -	D-ribulose-3-epimerase	DTE	Pseudomonas cichorii	D-tagatose 3-epimerase	ENA Nr.: BAA24429.1	1, 2	O50580	3
D-xylulose $\rightleftharpoons$ D-ribulose	5.1.3. -	D-ribulose-3-epimerase	C1KKR1	Rhodobacter sphaeroides	D-tagatose 3-epimerase	ENA Nr.: FJ851309.1	4	C1KKR1	5

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
D-ribulose +ATP <=> D-ribulose-1-phosphate +ADP	2.7.1. -	D-ribulose-1-kinase	fucK	Escherichia coli	L-fuculokinase	GeneID: 946022	6, 7, 257	P11553	8
D-ribulose-1-phosphate <=> glyceraldehyde +dihydroxyacetonephosphate	4.1.2. -	D-ribulose-1-phosphate aldolase	fucA	Escherichia coli	L-fuculose phosphate aldolase	GeneID: 947282	9, 10	P0AB87	11
Enzymes that may be used in D-xylulose-1-phosphate pathway to MEG									
D-xylulose +ATP <=> D-xylulose-1-phosphate +ADP	2.7.1. -	D-xylulose 1-kinase	khk-C (cDNA)	Homo sapiens	ketoheokinase C	GenBank: CR456801.1	53, 54, 256	P50053	55
D-xylulose-1-phosphate <=> glyceraldehyde +dihydroxyacetonephosphate	4.1.2. -	D-xylulose-1-phosphate aldolase	aldoB (cDNA)	Homo sapiens	Fructose-bisphosphate aldolase B	CCDS6756.1	56, 57	P05062	58
Enzymes that may be used in D-xylonate pathway to MEG									

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
D-xylose +NAD+ $\rightleftharpoons$ D-xylonolactone +NADH, or D-xylose +NAD+ $\rightleftharpoons$ D-xylonate +NADH	1.1.1.175	xylose dehydrogenase	xylB	Caulobacter crescentus	D-xylose 1-dehydrogenase	GeneID: 7329904	59, 60	B8H1Z0	61
D-xylose +NADP+ $\rightleftharpoons$ D-xylonolactone +NADPH, or D-xylose +NADP+ $\rightleftharpoons$ D-xylonate +NADPH	1.1.1.179	xylose dehydrogenase	xdh1, HVO_B0028	Haloferax volcanii	D-xylose 1-dehydrogenase	GeneID: 8919161	62	D4GP29	63
D-xylose +NADP+ $\rightleftharpoons$ D-xylonolactone +NADPH, or D-xylose +NADP+ $\rightleftharpoons$ D-xylonate +NADPH	1.1.1.179	xylose dehydrogenase	xyd1	Trichoderma reesei	D-xylose 1-dehydrogenase	ENA Nr.: EF136590.1	64	A0A024S MV2	65
D-xylonolactone +H ₂ O $\rightleftharpoons$ D-xylonate	3.1.1.68	xylonolactonase	xylC	Caulobacter crescentus	Xylonolactonase	GeneID: 7329903	66	A0A0H3 C6P8	67
D-xylonate $\rightleftharpoons$ 2-keto-3-deoxy-xylonate +H ₂ O	4.2.1.82	xylonate dehydratase	xylD	Caulobacter crescentus	xylonate dehydratase	GeneID: 7329902	68	A0A0H3 C6H6	69

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
D-xylonate <=> 2-keto-3-deoxy-xylonate +H ₂ O	4.2.1.82	xylonate dehydratase	yjhG	Escherichia coli	xylonate dehydratase	GeneID: 946829	70, 71	P39358	72
D-xylonate <=> 2-keto-3-deoxy-xylonate +H ₂ O	4.2.1.82	xylonate dehydratase	yagF	Escherichia coli	xylonate dehydratase	GeneID: 944928	73, 74	P77596	75
2-keto-3-deoxy-xylonate <=> glycolaldehyde + pyruvate	4.1.2.-	2-keto-3-deoxy-D-pentonate aldolase	yjhH	Escherichia coli	Uncharacterized lyase	GeneID: 948825	76, 77	P39359	78
2-keto-3-deoxy-xylonate <=> glycolaldehyde + pyruvate	4.1.2.-	2-keto-3-deoxy-D-pentonate aldolase	yagE	Escherichia coli	Probable 2-keto-3-deoxy-galactonate aldolase	GeneID: 944925	79, 80	P75682	81
Enzymes that may be used in 2-propanol (IPA) pathway via acetone or in acetone pathway to isobutene									
2 acetyl-CoA -> acetoacetyl-CoA + CoA	2.3.1.9	acetyl coenzyme A acetyltransferase	thIA	Clostridium acetobutylicum	acetyl coenzyme A acetyltransferase	3309200	33, 34	P45359	35

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
2 acetyl-CoA -> acetoacetyl-CoA + CoA	2.3.1.9	acetyl coenzyme A acetyltransferase	atoB	Escherichia coli	acetyl coenzyme A acetyltransferase	GeneID: 946727	36	P76461	37
2 acetyl-CoA -> acetoacetyl-CoA + CoA	2.3.1.9	acetyl coenzyme A acetyltransferase	ERG10	Saccharomyces cerevisiae	acetyl coenzyme A acetyltransferase	856079	38	P41338	39
acetoacetyl-CoA + acetate -> acetoacetate + acetyl-CoA	2.8.3.8	Acetyl-CoA:acetoacetate-CoA transferase subunit	atoA	Escherichia coli	Acetyl-CoA:acetoacetate-CoA transferase subunit	48994873	41, 42	P76459	43
acetoacetyl-CoA + acetate -> acetoacetate + acetyl-CoA	2.8.3.8	Acetyl-CoA:acetoacetate-CoA transferase subunit	atoD	Escherichia coli	Acetyl-CoA:acetoacetate-CoA transferase subunit	48994873	44, 45	P76458	46
acetoacetate -> acetone + CO ₂	4.1.1.4	acetoacetate decarboxylase	adc	Clostridium acetobutylicum	acetoacetate decarboxylase	6466901	47, 48	P23670	49

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
acetoacetate -> acetone + CO ₂	4.1.1.4	acetoacetate decarboxylase	adc	Clostridium beijerinckii	acetoacetate decarboxylase	149901357	50, 51	A6M020	52
acetone + acetyl-CoA + H ₂ O <-> 3-hydroxy-isovalerate	2.3.3.-	3-hydroxy-isovalerate synthase	Hmgcs1	Mus musculus	hydroxymethylglutaryl-CoA synthase	CCDS56901.1; GeneID:208715	104	Q3UWQ9	105
acetone + acetyl-CoA + H ₂ O <-> 3-hydroxy-isovalerate	2.3.3.-	3-hydroxy-isovalerate synthase	ERG13	Saccharomyces cerevisiae	hydroxymethylglutaryl-CoA synthase	GeneID: 854913	106	P54839	107
acetone + acetyl-CoA + H ₂ O <-> 3-hydroxy-isovalerate	2.3.3.-	3-hydroxy-isovalerate synthase	PksG	Lactobacillus crispatus ST1	hydroxymethylglutaryl-CoA synthase/ polyketide intermediate transferase	GeneID: 9107446	108	AEL95_01455	109
acetone + acetyl-CoA + H ₂ O <-> 3-hydroxy-isovalerate	2.3.3.-	3-hydroxy-isovalerate synthase	Pnap_0477	Polaromonas naphthalenivorans	hydroxymethylglutaryl-CoA lyase	ABM35799.1	110	A1VJH1	111

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
3-hydroxy-isovalerate + ATP <-> ADP + H(+) + 3-phosphonoxysovalerate	2.7.1. -	hydroxyisovalerate kinase	TA1305	Thermoplasma acidophilum	mevalonate-diphosphate decarboxylase/ mevalonate-monophosphate decarboxylase	GeneID: 1456782	112	Q9HIN1	113
3-hydroxy-isovalerate + ATP <-> ADP + H(+) + 3-phosphonoxysovalerate	2.7.1. -	hydroxyisovalerate kinase	TA1305* (L200E)	Thermoplasma acidophilum	mevalonate-diphosphate decarboxylase/ mevalonate-monophosphate decarboxylase	GeneID: 1456782	114	Modified version of Q9HIN1; L200E	115
3-hydroxy-isovalerate + ATP <-> ADP + H(+) + 3-phosphonoxysovalerate	2.7.1. -	hydroxyisovalerate kinase	PTO1356	Picrophilus torridus	mevalonate-diphosphate decarboxylase	GeneID: 2845209	116	Q6KZB1	117
3-phosphonoxysovalerate -> CO(2) + isobutene	4.1.1. -	3-phosphonoxysovalerate decarboxylase	smi_1746	Streptococcus mitis	mevalonate-diphosphate decarboxylase	Genbank: CBJ22986.1	118	D3HAT7	119

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
3-phosphonoxisovalerate → CO ₂ + isobutene	4.1.1. -	3-phosphonoxisovalerate decarboxylase	mvaD	Streptococcus gordonii	mevalonate-diphosphate decarboxylase	GeneID: 25051665	120	A8AUU9	121
3-hydroxy-isovalerate → CO ₂ + isobutene	4.1.1. -	hydroxyisovalerate decarboxylase	TA1305	Thermoplasma acidophilum	mevalonate-diphosphate decarboxylase	GeneID: 1456782	112	Q9HIN1	113
3-hydroxy-isovalerate → CO ₂ + isobutene	4.1.1. -	hydroxyisovalerate decarboxylase	PTO1356	Picrophilus torridus	mevalonate-diphosphate decarboxylase	GeneID: 2845209	116	Q6KZB1	117
3-hydroxy-isovalerate → CO ₂ + isobutene	4.1.1. -	hydroxyisovalerate decarboxylase	mvaD	Streptococcus gordonii	mevalonate-diphosphate decarboxylase	GeneID: 25051665	120	A8AUU9	121
Hydrolases that may be used in improved acetone pathway to isobutene									
Acetoacetyl-CoA + H ₂ O ⇌ CoA + acetoacetate	3.1.2. 11	acetate:acetoacetyl-CoA hydrolase	ctfA	Clostridium acetobutylicum	butyrate-acetoacetate CoA-transferase, complex A	NCBI- GeneID: 1116168	96	P33752	97

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
Acetoacetyl-CoA + H ₂ O ⇌ CoA + acetoacetate	3.1.2.11	acetate:acetoacetyl-CoA hydrolase	ctfB	Clostridium acetobutylicum	butyrate-acetoacetate CoA-transferase, subunit B	NCBI- GeneID: 1116169	98	P23673	99
Acetoacetyl-CoA + H ₂ O ⇌ CoA + acetoacetate	3.1.2.11	acetate:acetoacetyl-CoA hydrolase	atoA	Escherichia coli (strain K12)	Acetyl-CoA:acetoacetate-CoA transferase subunit	GeneID: 946719	100	P76459	101
Acetoacetyl-CoA + H ₂ O ⇌ CoA + acetoacetate	3.1.2.11	acetate:acetoacetyl-CoA hydrolase	atoD	Escherichia coli (strain K12)	Acetyl-CoA:acetoacetate-CoA transferase subunit	GeneID: 947525	102	P76458	103
Enzymes that may be used in HMG-CoA pathway to isobutene									
acetyl-CoA + H ₂ O + acetoacetyl-CoA ⇌ (S)-3-hydroxy-3-methylglutaryl-CoA + CoA	2.3.3.10	HMG-CoA synthase	hmgS	Saccharomyces cerevisiae	HMG-CoA synthase	GeneID: 854913	122	P54839	123

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
(S)-3-hydroxy-3-methylglutaryl-CoA $\rightleftharpoons$ trans-3-methylglutaconyl-CoA + H ₂ O	4.2.1.18	methylglutaconyl-CoA hydratase	liuC	Pseudomonas putida	methylglutaconyl-CoA hydratase	GeneID: 1041856	124	Q88FM3	125
ADP + phosphate + 3-methylglutaconyl-CoA $\rightleftharpoons$ ATP + 3-methylcrotonoyl-CoA + HCO ₃ ⁻	6.4.1.4	methylcrotonyl-CoA carboxylase	liuB	Pseudomonas aeruginosa	methylcrotonyl-CoA carboxylase subunit beta	GeneID: 878244	126	Q9I297	127
ADP + phosphate + 3-methylglutaconyl-CoA $\rightleftharpoons$ ATP + 3-methylcrotonoyl-CoA + HCO ₃ ⁻	6.4.1.4	methylcrotonyl-CoA carboxylase	liuD	Pseudomonas aeruginosa	methylcrotonyl-CoA carboxylase subunit alpha	GeneID: 879012	128	Q9I299	129
trans-2(or 3)-enoyl-CoA + H ₂ O $\rightleftharpoons$ (3S)-3-hydroxyacyl-CoA	4.2.1.17	methylcrotonyl-CoA hydratase	fadA	E. coli	fatty acid oxidation complex, 3-ketoacyl-CoA thiolase	GeneID: 948324	130	P21151	131

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
trans-2(or 3)-enoyl-CoA + H ₂ O $\rightleftharpoons$ (3S)-3-hydroxyacyl-CoA	4.2.1.17	methylcrotonyl-CoA hydratase	fadB	E. coli	fatty acid oxidation complex, enoyl-CoA hydratase	GeneID: 948336	132	P21177	133
3-hydroxy-isovaleryl-CoA + H ₂ O $\rightleftharpoons$ 3-hydroxy-isovalerate + CoA	3.1.2.-	3-hydroxy-isovaleryl-CoA thioesterase	tesB	E. coli	acyl-CoA thioesterase	GeneID: 945074	134	P0AGG2	135
Enzymes that may be used in 2-propanol (IPA) pathway via acetone									
acetone + NAD(P)H -> 2-propanol + NAD(P) ⁺	1.1.1.2	secondary alcohol dehydrogenase	adh	Clostridium beijerinckii	secondary alcohol dehydrogenase	60592972	136, 137	P25984	138
acetone + NAD(P)H -> 2-propanol + NAD(P) ⁺	1.1.1.2	secondary alcohol dehydrogenase	adh	Clostridium carboxidivorans	alcohol dehydrogenase	308066805	139	C6PZV5	140
NADH + NADP ⁺ $\rightleftharpoons$ NAD ⁺ + NADPH	1.6.1.1	Soluble pyridine nucleotide transhydrogenase	udhA	Escherichia coli	Soluble pyridine nucleotide transhydrogenase	GeneID: 948461	141	P27306	142
Enzymes that may be used in hexose to pentose-5-phosphate pathways									

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
ATP + D-xylulose = ADP + D-xylulose 5-phosphate	2.7.1.17	Xylulose 5-kinase	xylB	E. coli	Xylulose kinase	GeneID: 948133	145	P09099	146
D-erythrose 4-phosphate + D-xylulose 5-phosphate ↔ β-D-fructofuranose 6-phosphate + D-glyceraldehyde 3-phosphate	2.2.1.1	Transketolase	tktA	E. coli	Transketolase	GeneID: 947420	147	P27302	148
D-sedoheptulose 7-phosphate + D-glyceraldehyde 3-phosphate ↔ D-ribose 5-phosphate + D-xylulose 5-phosphate	2.2.1.1	Transketolase	tktA	E. coli	Transketolase	GeneID: 947420	147	P27302	148
D-erythrose 4-phosphate + D-xylulose 5-phosphate ↔ β-D-fructofuranose 6-phosphate + D-glyceraldehyde 3-phosphate	2.2.1.1	Transketolase	tktB	E. coli	Transketolase	GeneID: 945865	149	P33570	150

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
D-sedoheptulose 7-phosphate + D-glyceraldehyde 3-phosphate ↔ D-ribose 5-phosphate + D-xylulose 5-phosphate	2.2.1.1	Transketolase	tktB	E. coli	Transketolase	GeneID: 945865	149	P33570	150
D-sedoheptulose 7-phosphate + D-glyceraldehyde 3-phosphate ↔ β-D-fructofuranose 6-phosphate + D-erythrose 4-phosphate	2.2.1.2	Transaldolase	talA	E. coli	Transaldolase	GeneID: 947006	151	P0A867	152
D-sedoheptulose 7-phosphate + D-glyceraldehyde 3-phosphate ↔ β-D-fructofuranose 6-phosphate + D-erythrose 4-phosphate	2.2.1.2	Transaldolase	talB	E. coli	Transaldolase	GeneID: 944748	153	P0A870	154

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
D-ribose 5-phosphate ↔ D-ribose 5-phosphate	5.3.1.6	Ribose-5-phosphate isomerase	rpiA	E. coli	Ribose-5-phosphate isomerase	GeneID: 947407	155	P0A7Z0	156
D-ribose 5-phosphate ↔ D-ribose 5-phosphate	5.3.1.6	Ribose-5-phosphate isomerase	rpiB	E. coli	Ribose-5-phosphate isomerase	GeneID: 948602	254	P37351	253
D-ribulose 5-phosphate ↔ D-xylulose 5-phosphate	5.1.3.1	Ribulose-5-phosphate 3-epimerase	rpe	E. coli	Ribulose-5-phosphate 3-epimerase	GeneID: 947896	157	P0AG07	158
D-pentose 5-phosphate + H ₂ O → D-pentose + phosphate	3.1.3.1	D-pentose-5-phosphatase (PP)	phoA	E. coli	alkaline phosphatase	GeneID: 945041	159	P00634	160
D-ribose 5-phosphate + H ₂ O → D-ribose + phosphate	3.1.3.23	D-ribose-5-phosphatase (PP)	SGPP	Arabidopsis thaliana	Haloacid dehalogenase-like hydrolase	GeneID: 818456	161	Q9ZVJ5	162
D-xylulose 5-phosphate + H ₂ O → D-xylulose + phosphate	3.1.3.23	D-xylulose-5-phosphatase (PP)	araL	Bacillus subtilis	sugar phosphatase	GeneID: 937431	163	P94526	164

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
D-ribose 5-phosphate + H ₂ O -> D-ribose + phosphate	3.1.3.23	D-ribose-5-phosphatase (PP)	PFLU_2693	Pseudomonas fluorescens	haloacid dehalogenase-like enzyme	GeneID: 7821858	165	C3K9U8	166
D-pentose 5-phosphate + H ₂ O -> D-pentose + phosphate	3.1.3.23	D-pentose-5-phosphatase (PP)	yfbT	E. coli	sugar phosphatase	GeneID: 946777	167	P77625	168
D-ribose 5-phosphate + H ₂ O -> D-ribose + phosphate	3.1.3.23	D-ribose-5-phosphatase (PP)	ybiV	E. coli	sugar phosphatase	GeneID: 945432	169	P75792	170
D-pentose 5-phosphate + H ₂ O -> D-pentose + phosphate	3.1.3.23	D-pentose-5-phosphatase (PP)	yidA	E. coli	sugar phosphatase	GeneID: 948204	171	P0A8Y5	172
D-ribulose 5-phosphate + H ₂ O -> D-ribulose + phosphate	3.1.3.23	D-ribulose-5-phosphatase (PP)	PF10_0325	Plasmodium falciparum	Haloacid dehalogenase-like hydrolase	GeneID: 810482	173	Q8IJ74	174

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
D-arabitol 1-phosphate + NAD ⁺ ↔ D-xylulose 5-phosphate + NADH + H ⁺	1.1.1.301	D-arabitol 1-phosphate 4-dehydrogenase (APD)	APDH	Enterococcus avium	D-arabitol-phosphate dehydrogenase	GenBank: AY078980.1	175	Q8KQL2	176
D-arabitol 5-phosphate + NAD ⁺ ↔ D-ribulose 5-phosphate + NADH + H ⁺	1.1.1.-	D-arabitol 5-phosphate 2-dehydrogenase (APD)	APDH	Enterococcus avium	D-arabitol-phosphate dehydrogenase	GenBank: AY078980.1	175	Q8KQL2	176
D-arabitol 1-phosphate + NAD ⁺ ↔ D-ribulose 1-phosphate + NADH + H ⁺	1.1.1.-	D-arabitol 1-phosphate 2-dehydrogenase (APD)	ARD1	Candida albicans (strain WO-1)	D-arabitol 2-dehydrogenase	EMBL Translation : AAC37430.1	177	P43066	178
D-arabitol 1-phosphate + NAD ⁺ ↔ D-ribulose 1-phosphate + NADH + H ⁺	1.1.1.-	D-arabitol 1-phosphate 2-dehydrogenase (APD)	ARD1	Candida tropicalis	D-arabitol 2-dehydrogenase	EMBL Translation : AAA66355.1	179	P50166	180

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
D-arabitol 1-phosphate + NAD ⁺ ↔ D-ribulose 1-phosphate + NADH + H ⁺	1.1.1. -	D-arabitol 1-phosphate 2-dehydrogenase (APD)	ARDH	Scheffersomyces stipitis (strain ATCC 58785)	D-arabitol 2-dehydrogenase	EMBL Translation : CAA86939.1	181	P50167	182
D-arabitol 5-phosphate + NAD ⁺ ↔ D-xylulose 1-phosphate + NADH + H ⁺	1.1.1. -	D-arabitol 5-phosphate 4-dehydrogenase (APD)	mtID	Pseudomonas fluorescens	D-arabitol 4-dehydrogenase	GenBank: AF007800.1	183	GenBank : AAC04472.1	184
D-arabitol 5-phosphate + NAD ⁺ ↔ D-xylulose 1-phosphate + NADH + H ⁺	1.1.1. -	D-arabitol 5-phosphate 4-dehydrogenase (APD)	dalID	Klebsiella pneumoniae	D-arabitol 4-dehydrogenase	NCBI: NZ_FLCX01000003.1	185	O52720	186
D-arabitol 5-phosphate + NAD ⁺ ↔ D-xylulose 1-phosphate + NADH + H ⁺	1.1.1. -	D-arabitol 5-phosphate 4-dehydrogenase (APD)	dalID	Ralstonia solanacearum (strain GMI1000)	D-arabitol 4-dehydrogenase	GeneID: 1220971	187	P58708	188

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
D-arabitol 5-phosphate + NAD ⁺ ↔ D-xylulose 1-phosphate + NADH + H ⁺	1.1.1. -	D-arabitol 5-phosphate 4-dehydrogenase (APD)	egsA (araM)	Bacillus subtilis	Glycerol-1-phosphate dehydrogenase	GeneID: 938011	189	P94527	190
D-arabitol 1-phosphate + NAD ⁺ ↔ D-ribulose 1-phosphate + NADH + H ⁺	1.1.1. -	D-arabitol 1-phosphate 2-dehydrogenase (APD)	egsA (araM)	Bacillus subtilis	Glycerol-1-phosphate dehydrogenase	GeneID: 938011	189	P94527	190
D-arabitol 5-phosphate + NAD ⁺ ↔ D-xylulose 1-phosphate + NADH + H ⁺	1.1.1. -	D-arabitol 5-phosphate 4-dehydrogenase (APD)	egsA	Aeropyrum pernix (strain ATCC 700893)	Glycerol-1-phosphate dehydrogenase	EMBL: BAA79484.3	191	Q9YER2	192
D-arabitol 1-phosphate + NAD ⁺ ↔ D-ribulose 1-phosphate + NADH + H ⁺	1.1.1. -	D-arabitol 1-phosphate 2-dehydrogenase (APD)	egsA	Aeropyrum pernix (strain ATCC 700893)	Glycerol-1-phosphate dehydrogenase	EMBL: BAA79484.3	191	Q9YER2	192

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
D-arabitol 5-phosphate + NAD ⁺ ↔ D-xylulose 1-phosphate + NADH + H ⁺	1.1.1. -	D-arabitol 5-phosphate 4-dehydrogenase (APD)	gpsA	E. coli	glycerol-3-phosphate dehydrogenase	GeneID: 948125	193	P0A6S7	194
D-arabitol 1-phosphate + NAD ⁺ ↔ D-ribulose 1-phosphate + NADH + H ⁺	1.1.1. -	D-arabitol 1-phosphate 2-dehydrogenase (APD)	gpsA	E. coli	glycerol-3-phosphate dehydrogenase	GeneID: 948125	193	P0A6S7	194
D-arabitol 5-phosphate + NAD ⁺ ↔ D-xylulose 1-phosphate + NADH + H ⁺	1.1.1. -	D-arabitol 5-phosphate 4-dehydrogenase (APD)	GPD1	Saccharomyces cerevisiae	glycerol-3-phosphate dehydrogenase	GeneID: 851539	195	Q00055	196
D-arabitol 1-phosphate + NAD ⁺ ↔ D-ribulose 1-phosphate + NADH + H ⁺	1.1.1. -	D-arabitol 1-phosphate 2-dehydrogenase (APD)	GPD1	Saccharomyces cerevisiae	glycerol-3-phosphate dehydrogenase	GeneID: 851539	195	Q00055	196
D-xylulose 5-phosphate ↔ D-xylulose 1-phosphate	5.4.2. -	phosphopentomutase (PPM)	deoB	E. coli	phosphopentomutase	GeneID: 948910	197	P0A6K6	198

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
D-ribulose 5-phosphate ↔ D-ribulose 1-phosphate	5.4.2. -	phosphopentomutase (PPM)	deoB	E. coli	phosphopentomutase	GeneID: 948910	197	P0A6K6	198
D-xylulose 5-phosphate ↔ D-xylulose 1-phosphate	5.4.2. -	phosphopentomutase (PPM)	pgm	E. coli	alpha-phosphoglucosomutase	GeneID: 945271	199	P36938	200
D-ribulose 5-phosphate ↔ D-ribulose 1-phosphate	5.4.2. -	phosphopentomutase (PPM)	pgm	E. coli	alpha-phosphoglucosomutase	GeneID: 945271	199	P36938	200
D-xylulose 5-phosphate ↔ D-xylulose 1-phosphate	5.4.2. -	phosphopentomutase (PPM)	pgcA	Bacillus subtilis	alpha-phosphoglucosomutase	GeneID: 936247	201	P18159	202
D-ribulose 5-phosphate ↔ D-ribulose 1-phosphate	5.4.2. -	phosphopentomutase (PPM)	pgcA	Bacillus subtilis	alpha-phosphoglucosomutase	GeneID: 936247	201	P18159	202
D-xylulose 5-phosphate ↔ D-xylulose 1-phosphate	5.4.2. -	phosphopentomutase (PPM)	pgmB	Lactococcus lactis	beta-phosphoglucosomutase	GeneID: 1114041	203	P71447	204

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
D-ribulose 5-phosphate ↔ D-ribulose 1-phosphate	5.4.2. -	phosphopentomutase (PPM)	pgmB	Lactococcus lactis	beta-phosphoglucumutase	GeneID: 1114041	203	P71447	204
D-xylulose 5-phosphate ↔ D-xylulose 1-phosphate	5.4.2. -	phosphopentomutase (PPM)	ycjU	E. coli	beta-phosphoglucumutase	GeneID: 945891	205	P77366	206
D-ribulose 5-phosphate ↔ D-ribulose 1-phosphate	5.4.2. -	phosphopentomutase (PPM)	ycjU	E. coli	beta-phosphoglucumutase	GeneID: 945891	205	P77366	206
D-ribulose 5-phosphate ↔ D-ribulose 1-phosphate	5.4.2. 7	Phosphopentomutase (PPM)	PGM3 (PRM15)	Saccharomyces cerevisiae	phosphoribomutase	GenBank: NM_001182785.1 Gene ID: 855321	255	Q03262	258
D-xylulose 5-phosphate ↔ D-xylulose 1-phosphate	5.4.2. -	phosphopentomutase (PPM)	algC	Pseudomonas aeruginosa	phosphomannose mutase	GenBank: M60873.1	207	P26276	208
D-ribulose 5-phosphate ↔ D-ribulose 1-phosphate	5.4.2. -	phosphopentomutase (PPM)	algC	Pseudomonas aeruginosa	phosphomannose mutase	GenBank: M60873.1	207	P26276	208

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
D-xylulose 5-phosphate ↔ D-xylulose 1-phosphate	5.4.2. -	phosphopentomutase (PPM)	cpsG	E. coli	phosphomannose mutase	GeneID: 946574	209	P24175	210
D-ribulose 5-phosphate ↔ D-ribulose 1-phosphate	5.4.2. -	phosphopentomutase (PPM)	cpsG	E. coli	phosphomannose mutase	GeneID: 946574	209	P24175	210
D-fructose 6-phosphate + acetyl phosphate ↔ D-erythrose 4-phosphate + H ₂ O	4.1.2. 22	fructose 6-phosphate phosphoketolase (Fpk)	BDP_1006	Bifidobacterium dentium	fructose 6-phosphate phosphoketolase	EMBL: ADB09649.1	211	D2QA13	212
D-fructose 6-phosphate + acetyl phosphate ↔ D-erythrose 4-phosphate + H ₂ O	4.1.2. 22	fructose 6-phosphate phosphoketolase (Fpk)	xfp	Bifidobacterium lactis	Xylulose-5-phosphate/fructose-6-phosphate phosphoketolase	GeneID: 29696432	213	Q9AEM9	214

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
D-fructose 6-phosphate + phosphate ↔ acetyl phosphate + D-erythrose 4-phosphate + H ₂ O	4.1.2.22	fructose 6-phosphate phosphoketolase (Fpk)	xpkA	Lactobacillus paraplantarum	phosphoketolase	EMBL: AAQ64626.2	215	Q6UPD8	216
D-fructose 6-phosphate + phosphate ↔ acetyl phosphate + D-erythrose 4-phosphate + H ₂ O	4.1.2.22	fructose 6-phosphate phosphoketolase (Fpk)	xfp	Bifidobacterium breve	phosphoketolase	EMBL: ADF97524.1	217	D6PAH1	218
Acetyl-CoA + phosphate ↔ CoA + acetyl phosphate.	2.3.1.8	Phosphate acetyltransferase	pta	E. coli	Phosphate acetyltransferase	GeneID: 946778	219	P0A9M8	220
Acetyl-CoA + phosphate ↔ CoA + acetyl phosphate.	2.3.1.8	Phosphate acetyltransferase	pta	Clostridium acetobutylicum	Phosphate acetyltransferase	GeneID: 1117925	221	P71103	222
hydroxypyruvate + H ⁺ ↔ CO ₂ + glyceraldehyde	4.1.1.-	Hydroxypyruvate decarboxylase	kivd	Lactococcus lactis	α-ketoisovalerate decarboxylase	GenBank: AJ746364.1	223	Q684J7	224

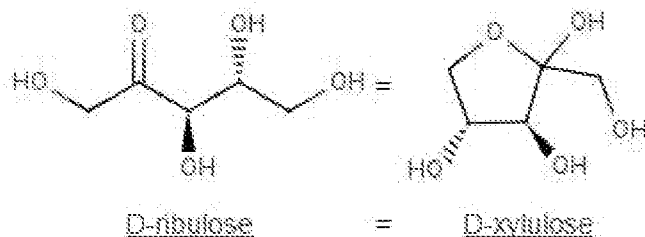
Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
hydroxypyruvate + H ⁺ ↔ CO ₂ + glycolaldehyde	4.1.1. -	Hydroxypyruvate decarboxylase	sucA	E. coli	2-oxoglutarate decarboxylase	GeneID: 945303	225	P0AFG3	226
3-phospho-D-glycerate + NAD ⁺ ↔ 3-phospho-hydroxypyruvate + NADH + H ⁺	1.1.1.95	D-3-phosphoglycerate dehydrogenase	serA	E. coli	D-3-phosphoglycerate dehydrogenase	GeneID: 945258	227	P0A9T0	228
3-phospho-L-serine + 2-oxoglutarate ↔ L-glutamate + 3-phospho-hydroxypyruvate	2.6.1.52	Phosphoserine aminotransferase	serC	E. coli	Phosphoserine aminotransferase	GeneID: 945527	229	P23721	230
3-phospho-hydroxypyruvate + H ₂ O → hydroxypyruvate + phosphate	-	3-phosphohydroxypyruvate phosphatase	yeaB (nudL)	E. coli	putative CoA pyrophosphohydrolase	GeneID: 946330	231	P43337	232
3-phospho-L-serine + H ₂ O → L-serine + phosphate	3.1.3.3	Phosphoserine phosphatase	serB	E. coli	Phosphoserine phosphatase	GeneID: 948913	233	P0AGB0	234

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/ annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
L-serine + H ⁺ → ethanolamine + CO ₂	4.1.1.65	serine decarboxylase	AtSDC (AT1G43710)	Arabidopsis thaliana	serine decarboxylase	GeneID: 840958	235	Q9MA74	236
ethanolamine + oxygen + H ₂ O → ammonium + hydrogen peroxide + glycolaldehyde	1.4.3.8	Ethanolamine oxidase	tynA	E. coli	amine oxidase	GeneID: 945939	237	P46883	238
ethanolamine + 2-oxoglutarate → glycolaldehyde + L-glutamate	2.6.1.-	Ethanolamine aminotransferase	alaA	E. coli	glutamate-pyruvate aminotransferase	GeneID: 946772	239	P0A959	240
D-glycerate + NAD(P) ⁺ ↔ hydroxypyruvate + NAD(P)H + H ⁺	1.1.1.-	Hydroxypyruvate reductase	ghrB	E. coli	glyoxylate reductase	GeneID: 948074	241	P37666	242
pyruvate + L-serine ↔ L-alanine + hydroxypyruvate	2.6.1.51	serine-pyruvate aminotransferase	AGXT1	Homo sapiens	serine-pyruvate aminotransferase	Gene ID: 189, CCDS2543.1	243	P21549	244
3-phospho-D-glycerate + H ₂ O → D-glycerate + phosphate	3.1.3.38	3-phosphoglycerate phosphatase	phoA	E. coli	phosphoglycerate phosphatase	GeneID: 945041	245	P00634	246

Described Reaction	EC no.	Required enzyme activity	Gene candidate	Source Organism	Natural/annotated function	Gene Identifier (nt)	SEQ ID NO (nt)	Uniprot ID	SEQ ID NO (AA)
2-phospho-D-glycerate + H ₂ O → D-glycerate + phosphate	3.1.3.20	2-phosphoglycerate phosphatase	phoA	E. coli	phosphoglycerate phosphatase	GeneID: 945041	245	P00634	246
D-glycerate + ATP ↔ 3-phospho-D-glycerate + ADP + H ⁺	2.7.1.31	Glycerate 3-kinase	GLYK	Arabidopsis thaliana	Glycerate 3-kinase	GeneID: 844378, mRNA NM_179581.2	247	Q94414	248
D-glycerate + ATP ↔ 2-phospho-D-glycerate + ADP + H ⁺	2.7.1.165	Glycerate 2-kinase	glxK	E. coli	Glycerate 2-kinase	GeneID: 945129	249	P77364	250
D-glycerate + ATP ↔ 2-phospho-D-glycerate + ADP + H ⁺	2.7.1.165	Glycerate 2-kinase	garK	E. coli	Glycerate 2-kinase	GeneID: 947632	251	P23524	252

D-tagatose 3-epimerase (EC 5.1.3.31)

[00169] The present disclosure describes enzymes that can catalyze the epimerization of various ketoses at the C-3 position, interconverting D-fructose and D-psicose, D-tagatose and D-sorbose, D-ribulose and D-xylulose, and L-ribulose and L-xylulose. The specificity depends on the species. The enzymes from *Pseudomonas cichorii* and *Rhodobacter sphaeroides* require Mn^{2+} . In one embodiment, the enzyme is D-tagatose 3-epimerase (dte). In another embodiment, the D-tagatose 3-epimerase catalyzes the conversion of D-xylulose to D-ribulose.

**[00170]**

[00171] In one embodiment, the D-tagatose 3-epimerase is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Pseudomonas sp.*, *Mesorhizobium sp.* and *Rhodobacter sp.* In some embodiments, the D-tagatose 3-epimerase is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Pseudomonas cichorii*, *Pseudomonas sp.* ST-24, *Mesorhizobium loti* and *Rhodobacter sphaeroides*. In some embodiments, the one or more nucleic acid molecules is dte and/or FJ851309.1, or homolog thereof. In a further embodiment, the D-tagatose 3-epimerase comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 3 and 5. In yet a further embodiment, the D-tagatose 3-epimerase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1, 2 and 4.

[00172] D-tagatose 3-epimerase may also be known as L-ribulose 3-epimerase or ketose 3-epimerase.

D-ribulokinase (EC 2.7.1.16)

[00173] The present disclosure describes enzymes that can catalyze the following reactions:

[00174] L-fucose + ATP → L-fucose 1-phosphate + ADP + H⁺

[00175] D-ribulose + ATP → D-ribulose 1-phosphate + ADP + H⁺

[00176] D-ribulokinase may also be known as L-fuculokinase, fuculokinase, ATP: L-fuculose 1-phosphotransferase or L-fuculose kinase.

[00177] Thus, in some embodiments, the disclosure provides for an enzyme that plays roles in the fucose degradation pathway, the super pathway of fucose and rhamnose degradation and/or the D-arabinose degradation I pathway.

[00178] In some embodiments, the enzyme can function as both an L-fuculokinase and a D-ribulokinase, the second enzyme of the L-fucose and D-arabinose degradation pathways, respectively.

[00179] In particular embodiments, the enzyme converts D-ribulose to D-ribulose-1-phosphate. In one embodiment, the D-ribulokinase is encoded by one or more nucleic acid molecules obtained from *E. coli*. In some embodiments, the one or more nucleic acid molecules is fucK, or homolog thereof. In a further embodiment, the D-ribulokinase comprises an amino acid sequence set forth in SEQ ID NO: 8. In yet a further embodiment, the D-ribulokinase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 6 and 7.

D-ribulose-1-phosphate aldolase (EC 4.1.2.17)

[00180] The present disclosure describes enzymes that can catalyze the following reversible reactions:

[00181] L-fuculose 1-phosphate ⇌ (S)-lactaldehyde + dihydroxy acetone phosphate (DHAP)

[00182] D-ribulose 1-phosphate ⇌ glycolaldehyde + dihydroxy acetone phosphate (DHAP)

[00183] D-ribulose-1-phosphate aldolase may also be known as L-fuculose-phosphate aldolase, L-fuculose 1-phosphate aldolase or L-fuculose-1-phosphate (S)-lactaldehyde-lyase.

[00184] Thus, in some embodiments, the disclosure provides for an enzyme that plays roles in the fucose degradation pathway, the super pathway of fucose and rhamnose degradation and/or the D-arabinose degradation I pathway. In one embodiment, the enzyme may use Zn²⁺ as a cofactor. In another embodiment, an inhibitor of this enzyme may be phosphoglycolohydroxamate.

[00185] In some embodiments, the enzyme can function as both an L-fuculose-phosphate aldolase and a D-ribulose-phosphate aldolase, the third enzyme of the L-fucose and D-arabinose degradation pathways, respectively.

[00186] The substrate specificity of the enzyme has been tested with a partially purified preparation from an *E. coli* strain.

[00187] Crystal structures of the enzyme and a number of point mutants have been solved. The combination of structural data and enzymatic activity of mutants allowed modelling and refinement of the catalytic mechanism of the enzyme. The enantiomeric selectivity of the enzyme has been studied.

[00188] In particular embodiments, the enzyme converts D-ribulose-1-phosphate to glycolaldehyde and DHAP. In one embodiment, the D-ribulose-1-phosphate aldolase is encoded by one or more nucleic acid molecules obtained from *E. coli*. In some embodiments, the one or more nucleic acid molecules is fucA, or homolog thereof. In a further embodiment, the D-ribulose-1-phosphate aldolase comprises an amino acid sequence set forth in SEQ ID NO: 11. In yet a further embodiment, the D-ribulose-1-phosphate aldolase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 9 and 10.

Glycolaldehyde reductase (EC 1.1.1.77)

[00189] The present disclosure describes enzymes that can catalyze the following reversible reactions:

[00190] ethylene glycol + NAD⁺ $\rightleftharpoons$ glycolaldehyde + NADH + H⁺

[00191] (S)-propane-1,2-diol + NAD⁺ $\rightleftharpoons$ (S)-lactaldehyde + NADH + H⁺

[00192] Glycolaldehyde reductase may also be known as lactaldehyde reductase, propanediol oxidoreductase, (R) [or(S)]-propane-1,2-diol:NAD⁺ oxidoreductase or L-1,2-propanediol oxidoreductase.

[00193] Thus, in some embodiments, the disclosure provides for an enzyme that plays roles in the ethylene glycol degradation pathway, the super pathway of glycol metabolism and degradation, the anaerobic L-lactaldehyde degradation pathway and/or the super pathway of fucose and rhamnose degradation. In one embodiment, the enzyme may use Fe²⁺ as a cofactor.

[00194] L-1,2-propanediol oxidoreductase is an iron-dependent group III dehydrogenase. It anaerobically reduces L-lactaldehyde, a product of both the L-fucose and L-rhamnose catabolic pathways, to L-1,2-propanediol, which is then excreted from the cell.

[00195] Crystal structures of the enzyme have been solved, showing a domain-swapped dimer in which the metal, cofactor and substrate binding sites could be located. An aspartate and three conserved histidine residues are required for Fe²⁺ binding and enzymatic activity.

[00196] *In vitro*, the enzyme can be reactivated by high concentrations of NAD⁺ and efficiently inactivated by a mixture of Fe³⁺ and ascorbate or Fe²⁺ and H₂O₂. Metal-catalyzed oxidation of the conserved His277 residue is proposed to be the cause of the inactivation.

[00197] Expression of FucO enables engineered one-turn reversal of the β -oxidation cycle. FucO activity contributes to the conversion of isobutyraldehyde to isobutanol in an engineered strain.

[00198] In particular embodiments, the enzyme converts glycolaldehyde to MEG. In some embodiments, the glycolaldehyde reductase is from *Escherichia coli*. In some embodiments, the glycolaldehyde reductase is encoded by the fucO gene.

[00199] In one embodiment, the glycolaldehyde reductase is encoded by one or more nucleic acid molecules obtained from a microorganism selected from *E. coli* and *S. cerevisiae*. In another embodiment, the one or more nucleic acid molecules is selected from gldA, GRE2, GRE3, yqhD, ydjG, fucO, yafB (dkgB), and/or yqhE (dkgA), or homolog thereof. In another embodiment, the one or more nucleic acid molecules is yqhD. In some embodiments, the yqhD comprises a G149E mutation. In a further embodiment, the glycolaldehyde reductase comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 13, 15, 17, 20, 23, 25, 28, 30 and 32. In yet a further embodiment, the glycolaldehyde reductase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 12, 14, 16, 18, 19, 21, 22, 24, 26, 27, 29 and 31.

Aldehyde reductases

[00200] A number of aldehyde reductases may be used to convert glycolaldehyde to MEG.

[00201] An NADPH-dependent aldehyde reductase (YqhD) can catalyze the following reactions:

[00202] acetol + NADP⁺ $\rightleftharpoons$ methylglyoxal + NADPH + H⁺ (reversible, EC 1.1.1.-)

[00203] an alcohol + NADP+ $\rightleftharpoons$ an aldehyde + NADPH + H+ (reversibility unspecified, EC 1.1.1.2)

[00204] an aldehyde + NADP+ + H₂O $\rightarrow$ a carboxylate + NADPH + 2 H+ (EC 1.2.1.4)

[00205] 1,3-propanediol + NADP+ $\rightleftharpoons$ 3-hydroxypropionaldehyde + NADPH + H+ (reversibility unspecified, EC 1.1.1.-)

[00206] D-3,4-dihydroxybutanal + NADPH $\rightleftharpoons$ 1,3,4-butanetriol + NADP+ (reversibility unspecified)

[00207] YqhD is an NADPH-dependent aldehyde reductase that may be involved in glyoxal detoxification and/or be part of a glutathione-independent response to lipid peroxidation.

[00208] It has been reported that various alcohols, aldehydes, amino acids, sugars and α -hydroxy acids have been tested as substrates for YqhD. The purified protein only shows NADP-dependent alcohol dehydrogenase activity, with a preference for alcohols longer than C(3), but with K_m values in the millimolar range, suggesting that they are not the physiological substrates. In contrast, YqhD does exhibit short-chain aldehyde reductase activity with substrates such as propanaldehyde, acetaldehyde, and butanaldehyde, as well as acrolein and malondialdehyde. In a metabolically engineered strain, phenylacetaldehyde and 4-hydroxyphenylacetaldehyde are reduced to 2-phenylethanol and 2-(4-hydroxyphenyl)ethanol by the endogenous aldehyde reductases YqhD, YjgB, and YahK.

[00209] Overexpression of YqhD increases 1,3-propanediol oxidoreductase activity of the cell. *E. coli* has been engineered to express YqhD for the industrial production of 1,3-propanediol. YqhD activity contributes to the production of isobutanol, 1,2-propanediol, 1,2,4-butanetriol and acetol as well. Mutation of yqhD enables production of butanol by an engineered one-turn reversal of the β -oxidation cycle.

[00210] YqhD has furfural reductase activity, which appears to cause growth inhibition due to depletion of NADPH in metabolically engineered strains that produce alcohol from lignocellulosic biomass.

[00211] The crystal structure of YqhD has been solved at 2 Å resolution. YqhD is an asymmetric dimer of dimers, and the active site contains a Zn²⁺ ion. The NADPH cofactor is modified by hydroxyl groups at positions 5 and 6 in the nicotinamide ring.

[00212] Overexpression of yqhD leads to increased resistance to reactive oxygen-generating compounds such as hydrogen peroxide, paraquat, chromate and potassium tellurite. A yqhD deletion mutant shows increased sensitivity to these compounds and to

glyoxal, and contains increased levels of reactive aldehydes that are generated during lipid peroxidation. Conversely, yqhD deletion leads to increased furfural tolerance.

[00213] In particular embodiments, an NADPH-dependent aldehyde reductase converts glycolaldehyde to MEG. In some embodiments, the NADPH-dependent aldehyde reductase is from *Escherichia coli*. In some embodiments, the NADPH-dependent aldehyde reductase is encoded by the yqhD gene.

[00214] A multi-functional methylglyoxal reductase (DkgA) can catalyze the following reactions:

[00215] acetol + NADP⁺ $\rightleftharpoons$ methylglyoxal + NADPH + H⁺ (the reaction is physiologically favored in the opposite direction, EC 1.1.1.-)

[00216] isobutanol + NADP⁺ $\rightleftharpoons$ isobutanal + NADPH + H⁺ (reversibility unspecified, EC 1.1.1.-)

[00217] ethyl-(2R)-methyl-(3S)-hydroxybutanoate + NADP⁺ $\rightleftharpoons$ ethyl-2-methylacetoacetate + NADPH + H⁺ (reversibility unspecified, EC 1.1.1.-)

[00218] 2-keto-L-gulonate + NADP⁺ $\leftarrow$ 2,5-didehydro-D-gluconate + NADPH + H⁺ (the reaction is favored in the opposite direction, EC 1.1.1.346)

[00219] DkgA (YqhE) belongs to the aldo-keto reductase (AKR) family and has been shown to have methylglyoxal reductase and beta-keto ester reductase activity.

[00220] dkgA is reported to encode a 2,5-diketo-D-gluconate reductase (25DKGR) A, one of two 25DKG reductases in *E. coli*. The enzyme uses NADPH as the preferred electron donor and is thought to be involved in ketogluconate metabolism. The specific activity of the enzyme towards 2,5-diketo-D-gluconate is reported to be almost 1000-fold lower than its activity towards methylglyoxal.

[00221] Due to its low K_m for NADPH, reduction of furans by DkgA may deplete NADPH pools and thereby limit cellular biosynthesis. A broad survey of aldehyde reductases showed that DkgA was one of several endogenous aldehyde reductases that contribute to the degradation of desired aldehyde end products of metabolic engineering.

[00222] A crystal structure of DkgA has been solved at 2.16 Å resolution.

[00223] In particular embodiments, a multi-functional methylglyoxal reductase converts glycolaldehyde to MEG. In some embodiments, the multi-functional methylglyoxal reductase is from *Escherichia coli*. In some embodiments, the multi-functional methylglyoxal reductase is encoded by the dkgA gene.

[00224] A multi-functional methylglyoxal reductase (DkgB) can catalyze the following reactions:

[00225] acetol + NADP+ $\rightleftharpoons$ methylglyoxal + NADPH + H+ (the reaction is physiologically favored in the opposite direction, EC 1.1.1.-)

5 **[00226]** 4-nitrobenzyl alcohol + NADP+ $\rightleftharpoons$ 4-nitrobenzaldehyde + NADPH + H+ (reversibility unspecified, EC 1.1.1.91)

[00227] 2-keto-L-gulonate + NADP+ $\leftarrow$ 2,5-didehydro-D-gluconate + NADPH + H+ (the reaction is favored in the opposite direction, EC 1.1.1.346)

[00228] DkgB (YafB) is a member of the aldo-keto reductase (AKR) subfamily 3F. DkgB
10 was shown to have 2,5-diketo-D-gluconate reductase, methylglyoxal reductase and 4-nitrobenzaldehyde reductase activities.

[00229] dkgB is reported to encode 2,5-diketo-D-gluconate reductase (25DKGR) B, one
of two 25DKG reductases in *E. coli*. The enzyme uses NADPH as the preferred electron
donor and is thought to be involved in ketogluconate metabolism. However, the specific
15 activity of the enzyme towards 2,5-diketo-D-gluconate is reported to be almost 1000-fold
lower than its activity towards methylglyoxal.

[00230] In particular embodiments, a multi-functional methylglyoxal reductase converts
glycolaldehyde to MEG. In some embodiments, the multi-functional methylglyoxal reductase
is from *Escherichia coli*. In some embodiments, the multi-functional methylglyoxal reductase
20 is encoded by the dkgB gene.

[00231] A methylglyoxal reductase (YeaE) can catalyze the following reaction:

[00232] acetol + NADP+ $\rightleftharpoons$ methylglyoxal + NADPH + H+ (the reaction is physiologically
favored in the opposite direction, EC 1.1.1.-)

[00233] YeaE has been shown to have methylglyoxal reductase activity.

25 **[00234]** The subunit structure of YeaE has not been determined, but its amino acid
sequence similarity to the aldo-keto reductases DkgA (YqhE) and DkgB (YafB) suggests that
it may be monomeric.

[00235] In particular embodiments, a methylglyoxal reductase converts glycolaldehyde to
MEG. In some embodiments, the methylglyoxal reductase is from *Escherichia coli*. In some
30 embodiments, the methylglyoxal reductase is encoded by the yeaE gene.

[00236] A L-glyceraldehyde 3-phosphate reductase (yghZ) can catalyze the following
reactions:

[00237] L-glyceraldehyde 3-phosphate + NADPH + H⁺ → sn-glycerol 3-phosphate + NADP⁺ (EC 1.1.1.-)

[00238] acetol + NADP⁺ ⇌ methylglyoxal + NADPH + H⁺ (the reaction is physiologically favored in the opposite direction, EC 1.1.1.-)

5 **[00239]** YghZ is an L-glyceraldehyde 3-phosphate (L-GAP) reductase. The enzyme is also able to detoxify methylglyoxal at a low rate. YghZ defines the AKR14 (aldo-keto reductase 14) protein family.

[00240] L-GAP is not a natural metabolite and is toxic to *E. coli*. L-GAP is a substrate of both the glycerol-3-phosphate and hexose phosphate transport systems of *E. coli* K-12. It has
10 been postulated that the physiological role of YghZ is the detoxification of L-GAP, which may be formed by non-enzymatic racemization of GAP or by an unknown cellular process.

[00241] The crystal structure of the *E. coli* enzyme has been determined and is suggested to be a tetramer. However, others have found that the protein forms an octamer based on gel filtration and electron microscopy studies.

15 **[00242]** In particular embodiments, a L-glyceraldehyde 3-phosphate reductase converts glycolaldehyde to MEG. In some embodiments, the L-glyceraldehyde 3-phosphate reductase is from *Escherichia coli*. In some embodiments, the L-glyceraldehyde 3-phosphate reductase is encoded by the yghZ gene.

[00243] An L-1,2-propanediol dehydrogenase/glycerol dehydrogenase (GldA) can
20 catalyze the following reactions:

[00244] (S)-propane-1,2-diol + NAD⁺ ⇌ acetol + NADH + H⁺ (reversible reaction)

[00245] aminoacetone + NADH + H⁺ → (R)-1-aminopropan-2-ol + NAD⁺ (EC 1.1.1.75)

[00246] glycerol + NAD⁺ ⇌ dihydroxyacetone + NADH + H⁺ (reversible reaction, EC 1.1.1.6)

25 **[00247]** The physiological function of the GldA enzyme has long been unclear. The enzyme was independently isolated as a glycerol dehydrogenase and a D-1-amino-2-propanol:NAD⁺ oxidoreductase. At that time, D-1-amino-2-propanol was thought to be an intermediate for the biosynthesis of vitamin B12, and although *E. coli* is unable to synthesize vitamin B12 *de novo*, enzymes catalyzing the synthesis of this compound were sought. It
30 was later found that GldA was responsible for both activities.

[00248] The primary *in vivo* role of GldA was recently proposed to be the removal of dihydroxyacetone by converting it to glycerol. However, a dual role in the fermentation of

glycerol has also recently been established. Glycerol dissimilation in *E. coli* can be accomplished by two different pathways. The glycerol and glycerophosphodiester degradation pathway requires the presence of a terminal electron acceptor and utilizes an ATP-dependent kinase of the Glp system, which phosphorylates glycerol to glycerol-3-phosphate. However, upon inactivation of the kinase and selection for growth on glycerol, it was found that an NAD⁺-linked dehydrogenase, GldA, was able to support glycerol fermentation. Recently, it was shown that GldA was involved in glycerol fermentation both as a glycerol dehydrogenase, producing dihydroxyacetone, and as a 1,2-propanediol dehydrogenase, regenerating NAD⁺ by producing 1,2-propanediol from acetol.

[00249] The enzyme is found in two catalytically active forms, a large form of eight subunits and a small form of two subunits. The large form appears to be the major species.

[00250] In particular embodiments, an L-1,2-propanediol dehydrogenase/glycerol dehydrogenase converts glycolaldehyde to MEG. In some embodiments, the L-1,2-propanediol dehydrogenase/glycerol dehydrogenase is from *Escherichia coli*. In some embodiments, the L-1,2-propanediol dehydrogenase/glycerol dehydrogenase is encoded by the *gldA* gene.

[00251] An NADPH-dependent methylglyoxal reductase (GRE2) from *Saccharomyces cerevisiae* can catalyze the following reactions:

[00252] (S)-lactaldehyde + NADP⁺ $\rightleftharpoons$ methylglyoxal + NADPH

[00253] 3-methylbutanol + NAD(P)⁺ $\rightleftharpoons$ 3-methylbutanal + NAD(P)H

[00254] Gre2 is a versatile enzyme that catalyzes the stereoselective reduction of a broad range of substrates including aliphatic and aromatic ketones, diketones, as well as aldehydes, using NADPH as the cofactor.

[00255] The crystal structures of Gre2 from *S. cerevisiae* in an apo-form at 2.00Å and NADPH-complexed form at 2.40Å resolution have been solved. Gre2 forms a homodimer, each subunit of which contains an N-terminal Rossmann-fold domain and a variable C-terminal domain, which participates in substrate recognition. The induced fit upon binding to the cofactor NADPH makes the two domains shift toward each other, producing an interdomain cleft that better fits the substrate. Computational simulation combined with site-directed mutagenesis and enzymatic activity analysis enabled characterization of a potential substrate-binding pocket that determines the stringent substrate stereoselectivity for catalysis.

[00256] Gre2 catalyzes the irreversible reduction of the cytotoxic compound methylglyoxal (MG) to (S)-lactaldehyde as an alternative to detoxification of MG by glyoxalase I GLO1. MG

is synthesized via a bypath of glycolysis from dihydroxyacetone phosphate and is believed to play a role in cell cycle regulation and stress adaptation. GRE2 also catalyzes the reduction of isovaleraldehyde to isoamylalcohol. The enzyme serves to suppress isoamylalcohol-induced filamentation by modulating the levels of isovaleraldehyde, the signal to which cells respond by filamentation. GRE2 is also involved in ergosterol metabolism.

[00257] In particular embodiments, an NADPH-dependent methylglyoxal reductase converts glycolaldehyde to MEG. In some embodiments, the NADPH-dependent methylglyoxal reductase is from *S. cerevisiae*. In some embodiments, the NADPH-dependent methylglyoxal reductase is encoded by the GRE2 gene.

Thiolase/Acetyl coenzyme A acetyltransferase (EC 2.3.1.9)

[00258] The present disclosure describes enzymes that can catalyze the following reaction:

[00259] $2 \text{ acetyl-CoA} \rightleftharpoons \text{acetoacetyl-CoA} + \text{coenzyme A}$ (reversible reaction)

[00260] Thiolase/Acetyl coenzyme A acetyltransferase may also be known as acetyl-CoA-C-acetyltransferase, acetoacetyl-CoA thiolase, acetyl-CoA:acetyl-CoA C-acetyltransferase or thiolase II.

[00261] Thus, in some embodiments, the disclosure provides for an enzyme that plays a role in acetoacetate degradation (to acetyl CoA). In one embodiment, an inhibitor of this enzyme may be acetoacetyl-CoA.

[00262] In particular embodiments, the enzyme converts acetyl-CoA to acetoacetyl-CoA. In one embodiment, the thiolase or acetyl coenzyme A acetyltransferase is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Clostridium sp.*, *Bacillus sp.*, *E. coli*, *Saccharomyces sp.* and *Marinobacter sp.* In some embodiments, the thiolase or acetyl coenzyme A acetyltransferase is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Clostridium acetobutylicum*, *Clostridium thermosaccharolyticum*, *Bacillus cereus*, *E. coli*, *Saccharomyces cerevisiae* and *Marinobacter hydrocarbonoclasticus*. In some embodiments, the one or more nucleic acid molecules is thIA, atoB and/or ERG10, or homolog thereof. In a further embodiment, the thiolase or acetyl coenzyme A acetyltransferase comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 35, 37 and 40. In yet a further embodiment, the thiolase or acetyl coenzyme A acetyltransferase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 33, 34, 36, 38 and 39.

Acetyl-CoA:acetoacetate-CoA transferase (EC 2.8.3.-)

[00263] The present disclosure describes enzymes that can catalyze the following reaction:

5 **[00264]** acetoacetate + acetyl-CoA $\rightleftharpoons$ acetoacetyl-CoA + acetate (reversible reaction, EC 2.8.3.-)

[00265] Acetyl-CoA:acetoacetate-CoA transferase may also be known as acetate:acetoacetyl-CoA transferase or acetoacetyl-CoA transferase.

10 **[00266]** Thus, in some embodiments, the disclosure provides for an enzyme that plays a role in acetoacetate degradation (to acetyl CoA). In one embodiment, inhibitors of this enzyme may include acetyl-CoA and coenzyme A.

[00267] The growth of *E. coli* on short-chain fatty acids (C3-C6) requires the activation of the acids to their respective thioesters. This activation is catalyzed by acetoacetyl-CoA transferase. The reaction takes place in two half-reactions which involves a covalent
15 enzyme-CoA. The enzyme undergoes two detectable conformational changes during the reaction. It is thought likely that the reaction proceeds by a ping-pong mechanism. The enzyme can utilize a variety of short-chain acyl-CoA and carboxylic acid substrates but exhibits maximal activity with normal and 3-keto substrates.

[00268] In particular embodiments, the enzyme converts acetoacetyl-CoA to
20 acetoacetate. In some embodiments, the acetyl-CoA:acetoacetate-CoA transferase is from *Clostridium spp.* In some embodiments, the acetyl-CoA:acetoacetate-CoA transferase is from *Clostridium acetobutylicum*. In some embodiments, the acetyl-CoA:acetoacetate-CoA transferase is from *Escherichia coli*. In some embodiments, the acetyl-CoA:acetoacetate-CoA transferase is encoded by the *atoA* and *atoD* genes. In another embodiment, the
25 subunit composition of acetoacetyl-CoA transferase is [(AtoA)₂][(AtoD)₂], with (AtoA)₂ being the β complex and (AtoD)₂ being the α complex. In one embodiment, the acetyl-CoA:acetoacetate-CoA transferase is a fused acetyl-CoA:acetoacetate-CoA transferase: α subunit/ β subunit. In another embodiment, the acetyl-CoA:acetoacetate-CoA transferase is encoded by the *ydiF* gene.

30

Acetate:Acetoacetyl-CoA hydrolase (EC 3.1.2.11)

[00269] The present disclosure describes enzymes that can catalyze the following reaction:

[00270] acetoacetyl-CoA + H₂O $\rightleftharpoons$ CoA + acetoacetate

[00271] Acetoacetyl-CoA hydrolase may also be known as acetoacetyl coenzyme A hydrolase, acetoacetyl CoA deacylase or acetoacetyl coenzyme A deacylase.

[00272] This enzyme belongs to the family of hydrolases, specifically those acting on thioester bonds.

[00273] In particular embodiments, the enzyme converts acetoacetyl-CoA to acetoacetate. In some embodiments, the acetate:acetoacetyl-CoA hydrolase is from *Clostridium spp.* In some embodiments, the acetate:acetoacetyl-CoA hydrolase is from *Clostridium acetobutylicum*. In another embodiment, the Acetoacetyl-CoA hydrolase is encoded by the ctfA (subunit A) and/or ctfB (subunit B) genes.

[00274] In a further embodiment, the acetyl-CoA:acetoacetate-CoA transferase or acetate:acetoacetyl-CoA hydrolase comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 43, 46, 97, 99, 101 and 103. In yet a further embodiment, the acetyl-CoA:acetoacetate-CoA transferase or acetate:acetoacetyl-CoA hydrolase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 41, 42, 44, 45, 96, 98, 100 and 102.

Acetoacetate decarboxylase (EC 4.1.1.4)

[00275] The present disclosure describes enzymes that can catalyze the following reaction:

[00276] acetoacetate + H⁺ → acetone + CO₂

[00277] Acetoacetate decarboxylase may also be known as ADC, AADC or acetoacetate carboxy-lyase.

[00278] Thus, in some embodiments, the disclosure provides for an enzyme that plays roles in isopropanol biosynthesis, pyruvate fermentation to acetone, the super pathway of *Clostridium acetobutylicum* acidogenic and solventogenic fermentation and/or the super pathway of *Clostridium acetobutylicum* solventogenic fermentation.

[00279] Acetoacetate decarboxylase (ADC) plays a key role in solvent production in *Clostridium acetobutylicum*. During the acidogenic phase of growth, acids accumulate causing a metabolic shift to solvent production. In this phase acids are re-assimilated and metabolized to produce acetone, butanol and ethanol.

[00280] Preliminary purification and crystallization of the enzyme has revealed that a lysine residue is implicated in the active site. The enzyme is a large complex composed of 12 copies of a single type of subunit.

[00281] The enzyme of *Clostridium acetobutylicum* ATCC 824 has been purified and the adc gene encoding it cloned. The enzyme has also been purified from the related strain *Clostridium acetobutylicum* DSM 792 and the gene cloned and sequenced. The decarboxylation reaction proceeds by the formation of a Schiff base intermediate.

[00282] ADC is a key enzyme in acid uptake, effectively pulling the CoA-transferase reaction in the direction of acetoacetate formation.

[00283] In particular embodiments, the enzyme converts acetoacetate to acetone. In one embodiment, the acetoacetate decarboxylase is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Clostridium sp.*, *Bacillus sp.*, *Chromobacterium sp.* and *Pseudomonas sp.* In another embodiment, the acetoacetate decarboxylase is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Clostridium acetobutylicum*, *Clostridium beijerinckii*, *Clostridium cellulolyticum*, *Bacillus polymyxa*, *Chromobacterium violaceum* and *Pseudomonas putida*. In some embodiments, the one or more nucleic acid molecules encoding the acetoacetate decarboxylase is adc, or homolog thereof. In a further embodiment, the acetoacetate decarboxylase comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 49 and 52. In yet another embodiment, the acetoacetate decarboxylase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 47, 48, 50 and 51.

Alcohol Dehydrogenase (EC 1.1.1.-)

[00284] The present disclosure describes enzymes that can catalyze the reversible oxidation of primary or secondary alcohols to aldehydes or ketones, respectively. In one embodiment, the enzyme is a secondary alcohol dehydrogenase (S-ADH) and catalyzes the reduction of ketones such as acetone into secondary alcohols such as 2-propanol (isopropanol).

[00285] In some embodiments the S-ADH is from *Burkholderia sp.* In some embodiments, the S-ADH is from *Burkholderia sp.* AIU 652. In some embodiments, the S-ADH is from *Alcaligenes sp.* In some embodiments, the S-ADH is from *Alcaligenes eutrophus*. In some embodiments, the S-ADH is from *Clostridium sp.* In some embodiments, the S-ADH is from *Clostridium ragsdalei*. In some embodiments, the S-ADH is from *Clostridium beijerinckii*. In

some embodiments, the S-ADH is from *Thermoanaerobacter sp.* In some embodiments, the S-ADH is from *Thermoanaerobacter Brockii*. In some embodiments, the S-ADH is from *Thermoanaerobacter ethanolicus* (*Clostridium thermohydrosulfuricum*). In some embodiments, the S-ADH is encoded by the adhB gene. In some embodiments, the S-ADH is from the trypanosomatid *Phytomonas sp.* In some embodiments, the S-ADH is from *Rhodococcus sp.* In some embodiments, the S-ADH is from *Rhodococcus ruber*. In some embodiments, the S-ADH is from *Methanobacterium palustre*. In some embodiments, the S-ADH is from methanogenic archaea *Methanogenium liminatans*. In some embodiments, the S-ADH is from the parasitic protist *Entamoeba histolytica* (EhAdh1). In some embodiments, the S-ADH is from parasitic protozoan *Tritrichomonas foetus*. In some embodiments, the S-ADH is from human parasite *Trichomonas vaginalis*.

[00286] In some embodiments, the S-ADH is predicted from homology and can be from *Thermoanaerobacter mathranii*, *Micrococcus luteus*, *Nocardiopsis alba*, *Mycobacterium hassiacum*, *Helicobacter suis*, *Candida albicans*, *Candida parapsilosis*, *Candida orthopsilosis*, *Candida metapsilosis*, *Grosmannia clavigera* and *Scheffersomyces stipitis*.

[00287] In some embodiments, the alcohol dehydrogenase has at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity with an alcohol dehydrogenase from *Clostridium sp.* In other embodiments, the alcohol dehydrogenase is an alcohol dehydrogenase selected from *Clostridium beijerinckii* adh and *Clostridium carboxidivorans* adh. In a further embodiment, the alcohol dehydrogenase comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 138 and 140. In yet another embodiment, the alcohol dehydrogenase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 136, 137, and 139.

Dehydratase (EC 4.2.1.-)

[00288] The present disclosure describes enzymes that can catalyze the following reactions:

[00289] isopropanol $\rightleftharpoons$ propene + H₂O

D-xylulose 1-kinase (EC 2.7.1.-)

[00290] The present disclosure describes enzymes that can catalyze the conversion of D-xylulose to D-xylulose-1-phosphate. In some embodiments, the conversion can be catalyzed by a human ketohexokinase C (khk-C), also known as fructokinase.

[00291] Ketohexokinase, or fructokinase, phosphorylates fructose to fructose-1-phosphate. The enzyme is involved in fructose metabolism, which is part of carbohydrate metabolism. It is found in the liver, intestine and kidney cortex.

[00292] In human liver, purified fructokinase, when coupled with aldolase, has been discovered to contribute to an alternative mechanism to produce oxalate from xylitol. In coupled sequence, fructokinase and aldolase produce glycolaldehyde, a precursor to oxalate, from D-xylulose via D-xylulose 1-phosphate.

[00293] In particular embodiments, the enzyme converts D-xylulose to D-xylulose-1-phosphate. In one embodiment, the D-xylulose 1-kinase is encoded by one or more nucleic acid molecules obtained from *Homo sapiens*. In some embodiments, the one or more nucleic acid molecules encoding the D-xylulose 1-kinase is ketohexokinase C (khk-C), or homolog thereof. In another embodiment, the one or more nucleic acid molecules encoding the D-xylulose 1-kinase comprises an amino acid sequence set forth in SEQ ID NOs: 55 or 256. In a further embodiment, the one or more nucleic acid molecules encoding the D-xylulose 1-kinase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 53 and 54.

D-xylulose-1-phosphate aldolase (EC 4.1.2.-)

[00294] The present disclosure describes enzymes that can catalyze the conversion of D-xylulose-1-phosphate to glycolaldehyde and DHAP. In some embodiments, the conversion can be catalyzed by a human aldolase B, which is also known as fructose-bisphosphate aldolase B or liver-type aldolase.

[00295] Aldolase B is one of three isoenzymes (A, B, and C) of the class I fructose 1,6-bisphosphate aldolase enzyme (EC 4.1.2.13), and plays a key role in both glycolysis and gluconeogenesis. The generic fructose 1,6-bisphosphate aldolase enzyme catalyzes the reversible cleavage of fructose 1,6-bisphosphate (FBP) into glyceraldehyde 3-phosphate and dihydroxyacetone phosphate (DHAP) as well as the reversible cleavage of fructose 1-phosphate (F1P) into glyceraldehyde and dihydroxyacetone phosphate. In mammals, aldolase B is preferentially expressed in the liver, while aldolase A is expressed in muscle and erythrocytes and aldolase C is expressed in the brain. Slight differences in isozyme structure result in different activities for the two substrate molecules: FBP and fructose 1-phosphate. Aldolase B exhibits no preference and thus catalyzes both reactions, while aldolases A and C prefer FBP.

[00296] Aldolase B is a homotetrameric enzyme, composed of four subunits. Each subunit has a molecular weight of 36 kDa and contains an eight-stranded α/β barrel, which encloses lysine 229 (the Schiff-base forming amino acid that is key for catalysis).

[00297] In particular embodiments, the enzyme converts D-xylulose-1-phosphate to glycolaldehyde and DHAP. In one embodiment, the D-xylulose-1-phosphate aldolase is encoded by one or more nucleic acid molecules obtained from *Homo sapiens*. In another embodiment, the one or more nucleic acid molecules encoding the D-xylulose-1-phosphate aldolase is aldolase B (aldoB), or homolog thereof. In some embodiments, the one or more nucleic acid molecules encoding the D-xylulose-1-phosphate aldolase comprises an amino acid sequence set forth in SEQ ID NO: 58. In some embodiments, the one or more nucleic acid molecules encoding the D-xylulose-1-phosphate aldolase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 56 and 57.

D-xylose isomerase (EC 5.3.1.5)

[00298] The present disclosure describes enzymes that can catalyze the following reversible reaction:

[00299] D-xylopyranose $\rightleftharpoons$ D-xylulose

[00300] D-xylose isomerase may also be known as xylose isomerase or D-xylose ketol-isomerase.

[00301] Thus, in some embodiments, the disclosure provides for an enzyme that plays a role in xylose degradation.

[00302] Xylose isomerase catalyzes the first reaction in the catabolism of D-xylose.

[00303] Two conserved histidine residues, H101 and H271, were shown to be essential for catalytic activity. The fluorescence of two conserved tryptophan residues, W49 and W188, is quenched during binding of xylose, and W49 was shown to be essential for catalytic activity. The presence of Mg^{2+} , Mn^{2+} or Co^{2+} protects the enzyme from thermal denaturation.

[00304] The subunit composition has not been established experimentally.

[00305] In particular embodiments, the enzyme converts D-xylose to D-xylulose. In one embodiment, the recombinant microorganism further comprises an endogenous or exogenous xylose isomerase that catalyzes the conversion of D-xylose to D-xylulose. In one embodiment, the xylose isomerase is exogenous. In another embodiment, the xylose isomerase is encoded by one or more nucleic acid molecules obtained from *Pyromyces sp*

or *E. coli*. In another embodiment, the one or more nucleic acid molecules encoding the xylose isomerase is xylA, or homolog thereof. In yet another embodiment, the one or more nucleic acid molecules encoding the xylose isomerase comprises an amino acid sequence selected from SEQ ID NOs: 95 and 144. In a further embodiment, the one or more nucleic acid molecules encoding the xylose isomerase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 93, 94 and 143.

[00306] In some embodiments, a recombinant microorganism producing MEG or GA, or optionally, MEG or GA and one or more co-product, comprises a deletion, insertion, or loss of function mutation in a gene encoding a D-xylose isomerase to prevent conversion of D-xylose to D-xylulose and instead shunt the reaction toward the conversion of D-xylose to D-xylonate.

D-xylulose 5-kinase/xylulokinase

[00307] The present disclosure describes enzymes that can catalyze the following reactions:

[00308] $\text{D-xylulose} + \text{ATP} \rightarrow \text{D-xylulose 5-phosphate} + \text{ADP} + \text{H}^+ \text{ (EC 2.7.1.17)}$

[00309] $\text{ATP} + \text{1-deoxy-D-xylulose} \rightarrow \text{1-deoxy-D-xylulose 5-phosphate} + \text{ADP} + \text{H}^+ \text{ (EC 2.7.1.-)}$

[00310] D-xylulose 5-kinase may also be known as xylulose kinase or xylulokinase.

[00311] Xylulokinase catalyzes the phosphorylation of D-xylulose, the second step in the xylose degradation pathway, producing D-xylulose-5-phosphate, an intermediate of the pentose phosphate pathway.

[00312] In the absence of substrate, xylulokinase has weak ATPase activity. Xylulokinase can also catalyze the phosphorylation of 1-deoxy-D-xylulose. This would allow a potential salvage pathway for generating 1-deoxy-D-xylulose 5-phosphate for use in the biosynthesis of terpenoids, thiamine and pyridoxal. The rate of phosphorylation of 1-deoxy-D-xylulose is 32-fold lower than the rate of phosphorylation of D-xylulose.

[00313] The kinetic mechanism of the bacterial enzyme has been studied, suggesting a predominantly ordered reaction mechanism. The enzyme undergoes significant conformational changes upon binding of the substrate and of ATP. Two conserved aspartate residues, D6 and D233, were found to be essential for catalytic activity, and a catalytic mechanism has been proposed.

[00314] Crystal structures of bacterial xylulokinase in the apo form and bound to D-xylulose have been determined at 2.7 and 2.1 Å resolution, respectively.

[00315] In particular embodiments, the enzyme converts D-xylulose to D-xylulose-5-phosphate. In some embodiments, the D-xylulose 5-kinase is from *Escherichia coli*. In some
5 embodiments, the D-xylulose 5-kinase is encoded by the xylB gene. In some embodiments, the D-xylulose 5-kinase is from *Saccharomyces cerevisiae*. In some embodiments the D-xylulose 5-kinase is encoded by the XKS1 gene. In some embodiments, the D-xylulose 5-kinase is from *Pichia stipitis*. In some embodiments the D-xylulose 5-kinase is encoded by the XYL3 gene.

[00316] In some embodiments, the D-xylulose 5-kinase is encoded by an amino acid
10 sequence having at least 70% sequence identity to xylB from *E. coli*. In a further embodiment, the D-xylulose 5-kinase is encoded by an amino acid sequence having at least 80% sequence identity to xylB from *E. coli*. In yet a further embodiment, the D-xylulose 5-kinase is encoded by an amino acid sequence having at least 90% sequence identity to xylB
15 from *E. coli*. In other embodiments, the D-xylulose 5-kinase is xylB from *E. coli*.

[00317] In one embodiment, the D-xylulose 5-kinase is encoded by one or more nucleic acid molecules obtained from *E. coli*. In some embodiments, the one or more nucleic acid molecules encoding the D-xylulose 5-kinase is xylB, or homolog thereof. In another
20 embodiment, the one or more nucleic acid molecules encoding the D-xylulose 5-kinase comprises an amino acid sequence set forth in SEQ ID NO: 146. In a further embodiment, the one or more nucleic acid molecules encoding the D-xylulose 5-kinase is encoded by a nucleic acid sequence set forth in SEQ ID NO: 145.

Xylose dehydrogenase (EC 1.1.1.175 or EC 1.1.1.179)

[00318] The present disclosure describes enzymes that can catalyze the following
25 reactions:

[00319] aldehydo-D-xylose + NAD⁺ + H₂O → D-xylonate + NADH + 2 H⁺

[00320] α-D-xylopyranose + NAD⁺ ⇌ D-xylonolactone + NADH + H⁺ (reversibility unspecified, EC 1.1.1.175)

[00321] Xylose dehydrogenase may also be known as D-xylose dehydrogenase, D-xylose
30 1-dehydrogenase, (NAD⁺)-linked D-xylose dehydrogenase, NAD⁺-D-xylose dehydrogenase, D-xylose:NAD⁺ 1-oxidoreductase

[00322] D-xylose dehydrogenase catalyzes the NAD⁺-dependent oxidation of D-xylose to D-xylonolactone. This is the first reaction in the oxidative, non-phosphorylative pathway for the degradation of D-xylose in *Caulobacter crescentus*. This pathway is similar to the pathway for L-arabinose degradation in *Azospirillum brasilense*. The amino acid sequence of the *C. crescentus* enzyme is unrelated to that of xylose dehydrogenase from the archaeon *Haloarcula marismortui*, or the L-arabinose 1-dehydrogenase of *Azospirillum brasilense*.

[00323] D-xylose is the preferred substrate for recombinant D-xylose dehydrogenase from *Caulobacter crescentus*. The enzyme can use L-arabinose, but it is a poorer substrate. The *K_m* for L-arabinose is 166 mM. Other substrates such as D-arabinose, L-xylose, D-ribose, D-galactose, D-glucose and D-glucose-6-phosphate showed little or no activity in the assay, as measured by NADH production. *C. crescentus* D-xylose dehydrogenase can convert D-xylose to D-xylonate directly.

[00324] Partially purified, native D-xylose dehydrogenase from *C. crescentus* had a *K_m* of 70 μ M for D-xylose. This value was lower than the *K_m* of 760 μ M for the recombinant, His-tagged enzyme.

[00325] In some embodiments, the D-xylose dehydrogenase is from the halophilic archaeon *Haloferax volcanii*. The *Haloferax volcanii* D-xylose dehydrogenase catalyzes the first reaction in the oxidative xylose degradation pathway of the halophilic archaeon *Haloferax volcanii*. The *H. volcanii* D-xylose dehydrogenase shows 59% amino acid sequence identity to a functionally characterized xylose dehydrogenase from *Haloarcula marismortui* and 56% identity to an ortholog in *Halorubrum lacusprofundi*, but is only 11% identical to the bacterial NAD⁺-dependent xylose dehydrogenase from *Caulobacter crescentus* CB15.

[00326] In particular embodiments, the enzyme converts D-xylose to D-xylonolactone. In one embodiment, the xylose dehydrogenase is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Caulobacter sp.*, *Haloarcula sp.*, *Haloferax sp.*, *Halorubrum sp.* and *Trichoderma sp.* In another embodiment, the xylose dehydrogenase is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Caulobacter crescentus*, *Haloarcula marismortui*, *Haloferax volcanii*, *Halorubrum lacusprofundi* and *Trichoderma reesei*. In some embodiments, the one or more nucleic acid molecules encoding the xylose dehydrogenase is selected from xylB, xdh1 (HVO_B0028) and/or xyd1, or homolog thereof. In a further embodiment, the one or more nucleic acid molecules encoding the xylose dehydrogenase comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 61, 63 and 65. In yet another embodiment, the one or

more nucleic acid molecules encoding the xylose dehydrogenase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 59, 60, 62 and 64.

Xylonolactonase (3.1.1.68)

5 **[00327]** The present disclosure describes enzymes that can catalyze the following reaction:

[00328] D-xylono-1,4-lactone + H₂O $\rightleftharpoons$ D-xylonate

[00329] This enzyme belongs to the family of hydrolases, specifically those acting on carboxylic ester bonds. This enzyme participates in pentose and glucuronate
10 interconversions.

[00330] Xylonolactonase may also be known as D-xylonolactonase, xylono-1,4-lactonase, xylono-gamma-lactonase or D-xylono-1,4-lactone lactonohydrolase.

[00331] In particular embodiments, the enzyme converts D-xylonolactone to D-xylonate. In one embodiment, the xylonolactonase is encoded by one or more nucleic acid molecules
15 obtained from a microorganism selected from *Caulobacter sp.* and *Haloferax sp.* In another embodiment, the xylonolactonase is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Caulobacter crescentus*, *Haloferax volcanii* and *Haloferax gibbonsii*. In some embodiments, the one or more nucleic acid molecules encoding the xylonolactonase is xylC, or homolog thereof. In a
20 further embodiment, the one or more nucleic acid molecules encoding the xylonolactonase comprises an amino acid sequence set forth in SEQ ID NO: 67. In yet another embodiment, the one or more nucleic acid molecules encoding the xylonolactonase is encoded by a nucleic acid sequence set forth in SEQ ID NO: 66.

25 **Xylonate dehydratase (EC 4.2.1.82)**

[00332] The present disclosure describes enzymes that can catalyze the following reaction:

[00333] D-xylonate $\rightleftharpoons$ 2-keto-3-deoxy-D-xylonate + H₂O

[00334] This enzyme belongs to the family of lyases, specifically the hydro-lyases, which
30 cleave carbon-oxygen bonds. This enzyme participates in pentose and glucuronate interconversions.

[00335] Xylonate dehydratase may also be known as D-xylonate hydro-lyase, D-xylonate aldolate dehydratase or D-xylonate dehydratase.

[00336] In particular embodiments, the enzyme converts D-xylonate to 2-keto-3-deoxy-D-xylonate. In one embodiment, the xylonate dehydratase is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Caulobacter sp.*, *Sulfolobus sp.* and *E. coli*. In another embodiment, the xylonate dehydratase is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Caulobacter crescentus*, *Sulfolobus solfataricus* and *E. coli*. In some embodiments, the one or more nucleic acid molecules encoding the xylonate dehydratase is selected from xylD, yjhG and/or yagF, or homolog thereof. In a further embodiment, the one or more nucleic acid molecules encoding the xylonate dehydratase comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 69, 72 and 75. In yet another embodiment, the one or more nucleic acid molecules encoding the xylonate dehydratase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 68, 70, 71, 73 and 74.

2-keto-3-deoxy-D-pentionate aldolase (4.1.2.28)

[00337] The present disclosure describes enzymes that can catalyze the following reaction:

[00338] 2-dehydro-3-deoxy-D-pentionate $\rightleftharpoons$ glycolaldehyde + pyruvate (reversibility unspecified)

[00339] This enzyme belongs to the family of lyases, specifically the aldehyde-lyases, which cleave carbon-carbon bonds. This enzyme participates in pentose and glucuronate interconversions.

[00340] 2-keto-3-deoxy-D-pentionate aldolase may also be known as 2-dehydro-3-deoxy-D-pentionate glycolaldehyde-lyase (pyruvate-forming), 2-dehydro-3-deoxy-D-pentionate aldolase, 3-deoxy-D-pentulosonic acid aldolase, and 2-dehydro-3-deoxy-D-pentionate glycolaldehyde-lyase.

[00341] YjhH appears to be a 2-dehydro-3-deoxy-D-pentionate aldolase. Genetic evidence suggests that YagE may also function as a 2-dehydro-3-deoxy-D-pentionate aldolase. yagE is part of the prophage CP4-6.

[00342] A *yjhH yagE* double mutant cannot use D-xylonate as the sole source of carbon, and crude cell extracts do not contain 2-dehydro-3-deoxy-D-pentionate aldolase activity. Both phenotypes are complemented by providing *yjhH* on a plasmid.

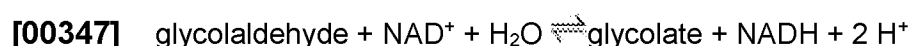
[00343] ArcA appears to activate *yjhH* gene expression under anaerobiosis. Two putative ArcA binding sites were identified 211 and 597 bp upstream of this gene, but no promoter upstream of it has been identified.

[00344] The crystal structure of YagE suggests that the protein is a homotetramer. Co-crystal structures of YagE in the presence of pyruvate and 2-keto-3-deoxygalactonate have been solved.

[00345] In particular embodiments, the enzyme converts 2-keto-3-deoxy-xylonate to glycolaldehyde and pyruvate. In one embodiment, the 2-keto-3-deoxy-D-pentionate aldolase is encoded by one or more nucleic acid molecules obtained from a microorganism selected from *Pseudomonas sp.* and *E. coli*. In another embodiment, the 2-keto-3-deoxy-D-pentionate aldolase is encoded by one or more nucleic acid molecules obtained from *E. coli*. In some embodiments, the one or more nucleic acid molecules encoding the 2-keto-3-deoxy-D-pentionate aldolase is selected from *yjhH* and/or *yagE*, or homolog thereof. In a further embodiment, the one or more nucleic acid molecules encoding the 2-keto-3-deoxy-D-pentionate aldolase comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 78 and 81. In yet another embodiment, the one or more nucleic acid molecules encoding the 2-keto-3-deoxy-D-pentionate aldolase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 76, 77, 79 and 80.

Glycolaldehyde dehydrogenase (1.2.1.21)

[00346] The present disclosure describes enzymes that can catalyze the following reaction:



[00348] This enzyme belongs to the family of oxidoreductases, specifically those acting on the aldehyde or oxo group of donor with NAD⁺ or NADP⁺ as acceptor. This enzyme participates in glyoxylate and dicarboxylate metabolism.

[00349] Glycolaldehyde dehydrogenase may also be known as glycolaldehyde:NAD⁺ oxidoreductase or glycol aldehyde dehydrogenase.

[00350] In *E. coli* aldehyde dehydrogenase A (AldA) is an enzyme of relatively broad substrate specificity for small α -hydroxyaldehyde substrates. It is thus utilized in several metabolic pathways.

[00351] L-fucose and L-rhamnose are metabolized through parallel pathways which converge after their corresponding aldolase reactions yielding the same products: dihydroxy-acetone phosphate and L-lactaldehyde. Aerobically, aldehyde dehydrogenase A oxidizes L-lactaldehyde to L-lactate.

[00352] In parallel pathways utilizing the same enzymes, D-arabinose and L-xylose can be metabolized to dihydroxy-acetone phosphate and glycolaldehyde, which is oxidized to glycolate by aldehyde dehydrogenase A.

[00353] Crystal structures of the enzyme alone and in ternary and binary complexes have been solved.

[00354] Aldehyde dehydrogenase A is only present under aerobic conditions and is most highly induced by the presence of fucose, rhamnose or glutamate. The enzyme is inhibited by NADH, which may act as a switch to shift from oxidation of lactaldehyde to its reduction by propanediol oxidoreductase. AldA is upregulated during short-term adaptation to glucose limitation.

[00355] Based on sequence similarity, AldA was predicted to be a succinate-semialdehyde dehydrogenase.

[00356] Regulation of aldA expression has been investigated. The gene is regulated by catabolite repression, repression under anaerobic conditions via ArcA, and induction by the carbon source.

[00357] In particular embodiments, the enzyme converts glycolaldehyde to glycolate. In some embodiments, the glycolaldehyde dehydrogenase is from *Escherichia coli*. In some embodiments, the glycolaldehyde dehydrogenase is encoded by the aldA gene.

[00358] In some embodiments, a recombinant microorganism producing MEG, or optionally, MEG and one or more co-product, comprises a deletion, insertion, or loss of function mutation in a gene encoding a glycolaldehyde dehydrogenase to prevent the production of glycolic acid from glycolaldehyde and instead shunt the reaction toward conversion of glycolaldehyde to MEG. In some embodiments, the deletion, disruption, mutation, and/or reduction in the activity of a glycolaldehyde dehydrogenase that catalyzes the conversion of glycolaldehyde to glycolic acid is partial, wherein an amount of glycolic acid is still produced.

[00359] In other embodiments, a recombinant microorganism producing glycolic acid comprises or expresses at least one nucleic acid molecule encoding a glycolaldehyde dehydrogenase.

5 **Lactate dehydrogenase (1.1.1.28)**

[00360] The present disclosure describes enzymes that can catalyze the following reaction:

[00361] (R)-lactate + NAD⁺ ← pyruvate + NADH + H⁺

10 **[00362]** Lactate dehydrogenase (LDH) is an enzyme found in nearly all living cells such as in animals, plants and prokaryotes. LDH catalyzes the conversion of lactate to pyruvic acid and back, as it converts NADH to NAD⁺ and back. A dehydrogenase is an enzyme that transfers a hydride from one molecule to another.

15 **[00363]** LDH exist in four distinct enzyme classes. The most common one is NAD(P)-dependent L-lactate dehydrogenase. Other LDHs act on D-lactate and/or are dependent on cytochrome c: D-lactate dehydrogenase (cytochrome) and L-lactate dehydrogenase (cytochrome).

[00364] LDH has been of medical significance because it is found extensively in body tissues, such as blood cells and heart muscle. Because it is released during tissue damage, it is a marker of common injuries and disease such as heart failure.

20 **[00365]** Lactate dehydrogenase may also be known as lactic acid dehydrogenase, (R)-lactate:NAD⁺ oxidoreductase or D-lactate dehydrogenase—fermentative.

[00366] In *E. coli*, lactate dehydrogenase (LdhA) is a soluble NAD-linked lactate dehydrogenase (LDH) that is specific for the production of D-lactate. LdhA is a homotetramer and shows positive homotropic cooperativity under higher pH conditions.

25 **[00367]** *E. coli* contains two other lactate dehydrogenases: D-lactate dehydrogenase and L-lactate dehydrogenase. Both are membrane-associated flavoproteins required for aerobic growth on lactate.

30 **[00368]** LdhA is present under aerobic conditions but is induced when *E. coli* is grown on a variety of sugars under anaerobic conditions at acidic pH. Unlike most of the genes involved in anaerobic respiration, LdhA is not activated by Fnr; rather the ArcAB system and several genes involved in the control of carbohydrate metabolism (csrAB and mlc) appear to regulate expression. The expression of LdhA is negatively affected by the transcriptional regulator ArcA. LdhA belongs to the σ^{32} regulon.

[00369] The *ldhA* gene is a frequent target for mutations in metabolic engineering, most often to eliminate production of undesirable fermentation side products, but also to specifically produce D-lactate.

[00370] In particular embodiments, the enzyme converts pyruvate to lactate. In some embodiments, the lactate dehydrogenase is from *Escherichia coli*. In some embodiments, the lactate dehydrogenase is encoded by the *ldhA* gene.

[00371] In some embodiments, a recombinant microorganism producing MEG or GA, or optionally, MEG or GA and one or more co-product, comprises a deletion, insertion, or loss of function mutation in a gene encoding a lactate dehydrogenase to prevent the production of lactate from pyruvate and instead shunt the reaction toward production of one or more co-products.

Xylose reductase or aldose reductase (EC 1.1.1.21)

[00372] The present disclosure describes enzymes that can catalyze the following reactions:

[00373] α -D-xylose + NADPH + H⁺ $\rightleftharpoons$ xylitol + NADP

[00374] an alditol + NAD(P)⁺ $\rightleftharpoons$ NAD(P)H + aldose

[00375] Aldose reductase may also be known as alditol:NAD(P)⁺ 1-oxidoreductase, polyol dehydrogenase or aldehyde reductase.

[00376] Aldose reductase is a cytosolic oxidoreductase that catalyzes the reduction of a variety of aldehydes and carbonyls, including monosaccharides.

[00377] Aldose reductase may be considered a prototypical enzyme of the aldo-keto reductase enzyme superfamily. The enzyme comprises 315 amino acid residues and folds into a β/α -barrel structural motif composed of eight parallel β strands. Adjacent strands are connected by eight peripheral α -helical segments running anti-parallel to the β sheet. The catalytic active site is situated in the barrel core. The NADPH cofactor is situated at the top of the β/α barrel, with the nicotinamide ring projecting down in the center of the barrel and pyrophosphate straddling the barrel lip.

[00378] The reaction mechanism of aldose reductase in the direction of aldehyde reduction follows a sequential ordered path where NADPH binds, followed by the substrate. Binding of NADPH induces a conformational change (Enzyme•NADPH $\rightarrow$ Enzyme*•NADPH) that involves hinge-like movement of a surface loop (residues 213-217) so as to cover a portion of the NADPH in a manner similar to that of a safety belt. The

alcohol product is formed via a transfer of the pro-R hydride of NADPH to the face of the substrate's carbonyl carbon. Following release of the alcohol product, another conformational change occurs ($E^* \cdot NAD(P)^+ \rightarrow E \cdot NAD(P)^+$) in order to release $NADP^+$. Kinetic studies have shown that reorientation of this loop to permit release of $NADP^+$ appears to represent the rate-limiting step in the direction of aldehyde reduction. As the rate of coenzyme release limits the catalytic rate, it can be seen that perturbation of interactions that stabilize coenzyme binding can have dramatic effects on the maximum velocity (V_{max}).

[00379] D-xylose-fermenting *Pichia stipitis* and *Candida shehatae* were shown to produce one single aldose reductase (ALR) that is active both with NADPH and NADH. Other yeasts such as *Pachysolen tannophilus* and *C. tropicalis* synthesize multiple forms of ALR with different coenzyme specificities. The significant dual coenzyme specificity distinguishes the *P. stipitis* and the *C. shehatae* enzymes from most other ALRs so far isolated from mammalian or microbial sources. The yeast *Candida tenuis* CBS 4435 produces comparable NADH- and NADPH-linked aldehyde-reducing activities during growth on D-xylose.

[00380] In particular embodiments, the enzyme converts D-xylose to xylitol. In some embodiments, the xylose reductase or aldose reductase is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Hypocrea sp.*, *Scheffersomyces sp.*, *Saccharomyces sp.*, *Pachysolen sp.*, *Pichia sp.*, *Candida sp.*, *Aspergillus sp.*, *Neurospora sp.*, and *Cryptococcus sp.* In some embodiments, the xylose reductase or aldose reductase is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Hypocrea jecorina*, *Scheffersomyces stipitis*, *Saccharomyces cerevisiae*, *Pachysolen tannophilus*, *Pichia stipitis*, *Pichia quercuum*, *Candida shehatae*, *Candida tenuis*, *Candida tropicalis*, *Aspergillus niger*, *Neurospora crassa* and *Cryptococcus lactativorus*. In another embodiment, the one or more nucleic acid molecules encoding the xylose reductase or aldose reductase is xyl1 and/or GRE3 or homolog thereof. In some embodiments, the one or more nucleic acid molecules encoding the xylose reductase or aldose reductase comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 84 and 87. In some embodiments, the one or more nucleic acid molecules encoding the xylose reductase or aldose reductase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 82, 83, 85 and 86.

Xylitol dehydrogenase (1.1.1.9)

[00381] The present disclosure describes enzymes that can catalyze the following reaction:

[00382] xylitol + NAD⁺ $\rightleftharpoons$ D-xylulose + NADH + H⁺

[00383] Xylitol dehydrogenase may also be known as D-xylulose reductase, NAD⁺-dependent xylitol dehydrogenase, erythritol dehydrogenase, 2,3-cis-polyol(DPN) dehydrogenase (C3-5), pentitol-DPN dehydrogenase, xylitol-2-dehydrogenase or xylitol:NAD⁺ 2-oxidoreductase (D-xylulose-forming).

[00384] Xylitol dehydrogenase (XDH) is one of several enzymes responsible for assimilating xylose into eukaryotic metabolism and is useful for fermentation of xylose contained in agricultural byproducts to produce ethanol. For efficient xylose utilization at high flux rates, cosubstrates should be recycled between the NAD⁺-specific XDH and the NADPH-preferring xylose reductase, another enzyme in the pathway.

[00385] In particular embodiments, the enzyme converts xylitol to D-xylulose. In one embodiment of any aspect disclosed above, the xylitol dehydrogenase is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Scheffersomyces* sp., *Trichoderma* sp., *Pichia* sp., *Saccharomyces* sp., *Gluconobacter* sp., *Galactocandida* sp., *Neurospora* sp., and *Serratia* sp. In another embodiment, the xylitol dehydrogenase is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Scheffersomyces stipitis*, *Trichoderma reesei*, *Pichia stipitis*, *Saccharomyces cerevisiae*, *Gluconobacter oxydans*, *Galactocandida mastotermitis*, *Neurospora crassa* and *Serratia marcescens*. In another embodiment, the one or more nucleic acid molecules encoding the xylitol dehydrogenase is xyl2 and/or xdh1, or homolog thereof. In some embodiments, the one or more nucleic acid molecules encoding the xylitol dehydrogenase comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 90 and 92. In some embodiments, the one or more nucleic acid molecule encoding the xylitol dehydrogenase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 88, 89 and 91. In another embodiment, the enzyme converts D-xylulose to xylitol. In another embodiment, the enzyme has high activity for the conversion of D-xylulose to xylitol and preferably low activity or no activity for the reverse reaction, i.e. for the conversion of xylitol to D-xylulose. In another embodiment, this is achieved through enzyme engineering.

Soluble pyridine nucleotide transhydrogenase (EC 1.6.1.1.)

[00386] The present disclosure describes enzymes that can catalyze the following reaction:

[00387] NADH + NADP⁺ $\rightleftharpoons$ NAD⁺ + NADPH

[00388] Soluble pyridine nucleotide transhydrogenase may also be known as NAD(P)+ transhydrogenase (B-specific), STH, pyridine nucleotide transhydrogenase, or transhydrogenase.

[00389] *E. coli* contains both a soluble and a membrane-bound pyridine nucleotide transhydrogenase. The soluble pyridine nucleotide transhydrogenase is the *sthA* or *udhA* gene product; its primary physiological role appears to be the reoxidation of NADPH. The membrane-bound proton-translocating transhydrogenase is the *pntAB* gene product; PntAB is a major source of NADPH.

[00390] UdhA contains noncovalently bound FAD and is present in a form consisting of seven or eight monomers.

[00391] Moderate overexpression of UdhA (SthA) allows an increased maximal growth rate of a phosphoglucose isomerase mutant, and a *pgi sthA* double mutant is not viable. These phenotypes may be due to the ability of UdhA to restore the cellular redox balance under conditions of excess NADPH formation. Mutations in *sthA* appear during adaptation of a *pgi* mutant strain to growth on glucose minimal medium.

[00392] Transcription of *sthA* is downregulated by growth on glycerol.

[00393] In some embodiments, expression of a transhydrogenase can increase activity of a NADPH-dependent alcohol dehydrogenase, leading to improved acetone to 2-propanol conversion. In one embodiment, the soluble pyridine nucleotide transhydrogenase is encoded by one or more nucleic acid molecules obtained from *E. coli*. In another embodiment, the one or more nucleic acid molecules encoding the soluble pyridine nucleotide transhydrogenase is *udhA*, or homolog thereof. In some embodiments, the one or more nucleic acid molecules encoding the soluble pyridine nucleotide transhydrogenase comprises an amino acid sequence set forth in SEQ ID NO: 142. In some embodiments, the one or more nucleic acid molecules encoding the soluble pyridine nucleotide transhydrogenase is encoded by a nucleic acid sequence set forth in SEQ ID NO: 141.

Hydroxymethylglutaryl-CoA synthase (EC 2.3.3.-)

[00394] The present disclosure describes enzymes that can catalyze the following reaction:

[00395] acetoacetyl-CoA + acetyl-CoA + H₂O ↔ (S)-3-hydroxy-3-methylglutaryl-CoA + coenzyme A + H⁺

[00396] Hydroxymethylglutaryl-CoA synthase may also be known as (S)-3-hydroxy-3-methylglutaryl-CoA acetoacetyl-CoA-lyase (CoA-acetylating), 3-hydroxy-3-methylglutaryl CoA synthetase, 3-hydroxy-3-methylglutaryl coenzyme A synthase, 3-hydroxy-3-methylglutaryl coenzyme A synthetase, 3-hydroxy-3-methylglutaryl-CoA synthase, 3-hydroxy-3-methylglutaryl-coenzyme A synthase, β -hydroxy- β -methylglutaryl-CoA synthase, HMG-CoA synthase, acetoacetyl coenzyme A transacetase, hydroxymethylglutaryl coenzyme A synthase, and hydroxymethylglutaryl coenzyme A-condensing enzyme.

[00397] Hydroxymethylglutaryl-CoA synthase catalyzes the condensation of acetyl-CoA with acetoacetyl-CoA to form (S)-3-hydroxy-3-methylglutaryl-CoA, an early stage in the synthesis of (R)-mevalonate, a precursor of cholesterol.

[00398] The enzyme catalyzes a complex reaction that can be divided into four steps. The first step involves the formation of an enzyme acetyl-CoA binary complex, followed by the transfer of the acetyl group from the CoA thioester to a cysteine residue on the enzyme, forming a thioester acyl-enzyme intermediate. In the next step the now reduced CoA dissociates, and the second substrate, acetoacetyl-CoA, binds the enzyme. The third step involves the formation of a carbanion by removal of a proton from the methyl of the acetylcysteine. The activated acetylcysteine then undergoes a Claisen-like condensation with the γ -carbon of the acetoacetyl-CoA ligand, which forms the HMG-CoA while retaining the thioester bond to the enzyme. The last step comprises the hydrolysis of this bond, resulting in free HMG-CoA.

[00399] The HMGCS1 gene from *Homo sapiens* has been cloned and sequenced (Russ AP et al. (1992) Amplification and direct sequencing of a cDNA encoding human cytosolic 3-hydroxy-3-methylglutaryl-coenzyme A synthase. *Biochim Biophys Acta* 1132(3): 329-31). The gene was expressed in *Escherichia coli*, and the recombinant protein was purified and characterized (Rokosz LL et al. (1994) Human cytoplasmic 3-hydroxy-3-methylglutaryl coenzyme A synthase: expression, purification, and characterization of recombinant wild-type and Cys129 mutant enzymes. *Arch Biochem Biophys* 312(1): 1-13). The enzyme is a homodimer of 120 kDa. Catalysis proceeds by formation of a covalent acetyl-enzyme intermediate. Kinetic data suggest that the two substrates (acetyl-CoA and acetoacetyl-CoA) compete for binding to the same site.

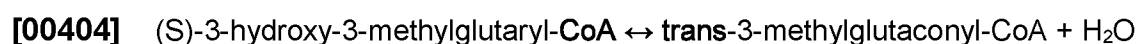
[00400] In one embodiment, the hydroxymethylglutaryl-CoA synthase can have a 3-hydroxyisovalerate (3HIV) synthase activity and can catalyze the following reaction:

[00401] acetone + acetyl-CoA + H₂O $\leftrightarrow$ 3-hydroxyisovalerate

[00402] In one embodiment, the 3HIV synthase is encoded by one or more nucleic acid molecules obtained from a microorganism selected from *Mus sp.*, *Saccharomyces sp.*, *Lactobacillus sp.* and *Polaromonas sp.* In another embodiment, the 3HIV synthase is encoded by one or more nucleic acid molecules obtained from a microorganism selected from *Mus musculus*, *Saccharomyces cerevisiae*, *Lactobacillus crispatus* and *Polaromonas naphthalenivorans*. In some embodiments, the one or more nucleic acid molecules encoding the 3HIV synthase is selected from Hmgcs1, ERG13, PksG and/or Pnap_0477, or homolog thereof. In a further embodiment, the one or more nucleic acid molecules encoding the 3HIV synthase comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 105, 107, 109 and 111. In yet another embodiment, the one or more nucleic acid molecules encoding the 3HIV synthase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 104, 106, 108 and 110. In some embodiments, the one or more nucleic acid molecules encoding the hydroxymethylglutaryl-CoA synthase is hmgS, or homolog thereof. In a further embodiment, the one or more nucleic acid molecules encoding the hydroxymethylglutaryl-CoA synthase comprises an amino acid sequence set forth in SEQ ID NO: 123. In yet another embodiment, the one or more nucleic acid molecules encoding the hydroxymethylglutaryl-CoA synthase is encoded by a nucleic acid sequence set forth in SEQ ID NO: 122.

Methylglutaconyl-CoA hydratase (EC 4.2.1.18)

[00403] The present disclosure describes enzymes that can catalyze the following reaction:



[00405] This enzyme catalyzes the *syn*-hydration of 3-methylglutaconyl-CoA to (S)-3-hydroxy-3-methylglutaryl-CoA in the leucine degradation pathway. The bacterial enzyme has been characterized in *Pseudomonas putida*. It differs from the mammalian enzyme in having only one glutamyl residue in its active site rather than two, resulting in a different reaction mechanism. These enzymes are members of the crotonase superfamily (Wong BJ and Gerlt JA (2004) Evolution of function in the crotonase superfamily: (3S)-methylglutaconyl-CoA hydratase from *Pseudomonas putida*. Biochemistry 43(16): 4646-4654) and reviewed in (Hamed RB et al. (2008) Mechanisms and structures of crotonase superfamily enzymes--how nature controls enolate and oxyanion reactivity. Cell Mol Life Sci 65(16): 2507-2527).

[00406] Recombinant enzyme was expressed in *Escherichia coli*, purified and characterized. The apparent molecular mass of the 10-His-tagged polypeptide was

determined to be 32.251 kDa by ESI-MS. The 10-His-tag was subsequently removed before characterization of the enzyme (Wong and Gerlt 2004).

[00407] In one embodiment, the methylglutaconyl-CoA hydratase is encoded by one or more nucleic acid molecules obtained from *Pseudomonas sp.* In another embodiment, the methylglutaconyl-CoA hydratase is encoded by one or more nucleic acid molecules obtained from *Pseudomonas putida*. In some embodiments, the one or more nucleic acid molecules encoding the methylglutaconyl-CoA hydratase is liuC, or homolog thereof. In a further embodiment, the one or more nucleic acid molecules encoding the methylglutaconyl-CoA hydratase comprises an amino acid sequence set forth in SEQ ID NO: 125. In yet another embodiment, the one or more nucleic acid molecules encoding the methylglutaconyl-CoA hydratase is encoded by a nucleic acid sequence set forth in SEQ ID NO: 124.

Methylcrotonyl-CoA carboxylase (EC 6.4.1.4)

[00408] The present disclosure describes enzymes that can catalyze the following reaction:

[00409] $\text{ADP} + \text{phosphate} + \text{H}^+ + 3\text{-methylglutaconyl-CoA} \leftrightarrow \text{ATP} + 3\text{-methylcrotonoyl-CoA} + \text{HCO}_3^-$

[00410] The enzyme activity is associated with the 3-methylcrotonyl-CoA carboxylase complex. This enzyme is a biotin-containing, biotin-dependent carboxylase involved in the L-leucine (and isovalerate) degradation pathway of *Pseudomonas aeruginosa* PAO1. This pathway is also the last phase of the acyclic terpene utilization pathway (citronellol degradation and cis-geranyl-CoA degradation pathways). The enzyme is not expressed in citronellol or citronellate grown cells, but is expressed in isovalerate grown cells. Genes liuB and liuD encode the two subunits of 3-methylcrotonyl-CoA carboxylase. The subunits are encoded in the liuRABCDE gene cluster of this organism (Hoschle B et al. (2005) Methylcrotonyl-CoA and geranyl-CoA carboxylases are involved in leucine/isovalerate utilization (Liu) and acyclic terpene utilization (Atu), and are encoded by liuB/liuD and atuC/atuF, in *Pseudomonas aeruginosa*. Microbiology 151(Pt 11): 3649-3656; Forster-Fromme K and Jendrossek D (2010). Catabolism of citronellol and related acyclic terpenoids in pseudomonads. Appl Microbiol Biotechnol 87(3): 859-869).

[00411] The enzyme was purified from cell extracts by avidin-affinity chromatography and the SDS-gel-isolated subunits were subjected to trypsin fingerprint analysis and ESI-MS which allowed identification of their corresponding genes (Hoschle et al. 2005).

[00412] The 3-methylcrotonyl-CoA carboxylase of *Pseudomonas citronellolis* was characterized in earlier work (Hector ML and Fall RR (1976) Multiple acyl-coenzyme A carboxylases in *Pseudomonas citronellolis*. Biochemistry 15(16): 3465-3472; Fall RR and Hector ML (1977) Acyl-coenzyme A carboxylases. Homologous 3-methylcrotonyl-CoA and geranyl-CoA carboxylases from *Pseudomonas citronellolis*. Biochemistry 16(18): 4000-4005; Fall RR (1981) 3-Methylcrotonyl-CoA and geranyl-CoA carboxylases from *Pseudomonas citronellolis*. Methods Enzymol 71 Pt C: 791-799).

[00413] In one embodiment, the methylcrotonyl-CoA carboxylase is encoded by one or more nucleic acid molecules obtained from *Pseudomonas sp.* In another embodiment, the methylcrotonyl-CoA carboxylase is encoded by one or more nucleic acid molecules obtained from *Pseudomonas aeruginosa*. In some embodiments, the one or more nucleic acid molecules encoding the methylcrotonyl-CoA carboxylase is selected from liuB and/or liuD, or homologs thereof. In a further embodiment, the one or more nucleic acid molecules encoding the methylcrotonyl-CoA carboxylase comprises an amino acid sequence selected from SEQ ID NOs: 127 and 129. In yet another embodiment, the one or more nucleic acid molecules encoding the methylcrotonyl-CoA carboxylase is encoded by a nucleic acid sequence selected from SEQ ID NOs: 126 and 128.

Methylcrotonyl-CoA hydratase (EC 4.2.1.17)

[00414] The present disclosure describes enzymes that can catalyze the following reaction:

[00415] $\text{trans-2(or 3)-enoyl-CoA} + \text{H}_2\text{O} \leftrightarrow \text{(3S)-3-hydroxyacyl-CoA}$

[00416] An exemplary enzyme is a 3-ketoacyl-CoA thiolase. It is involved in the degradation of fatty acids via the β -oxidation cycle. It has broad chain-length specificity for substrates although it exhibits its highest activity with medium-chain substrates. It is part of a multienzyme complex and is coded for by the fadA gene (Yang SY et al (1990) Nucleotide sequence of the fadA gene. Primary structure of 3-ketoacyl-coenzyme A thiolase from *Escherichia coli* and the structural organization of the fadAB operon. J Biol Chem 265(18): 10424-10429;).

[00417] 3-ketoacyl-CoA thiolase may also be known as acetyl-CoA C-acyltransferase, β -ketothiolase, acetyl-CoA acyltransferase and acyl-CoA:acetyl-CoA C-acyltransferase.

[00418] Another exemplary enzyme is an enoyl-CoA hydratase. The alpha subunit has four enzymatic activities associated with it. It is part of a multienzyme complex. Two of the activities, enoyl-CoA hydratase (EC 4.2.1.17) and 3-OH acyl-CoA epimerase (EC 5.1.2.3)

are carried out by the same N terminal active site (Yang SY and Elzinga M (1993) Association of both enoyl coenzyme A hydratase and 3-hydroxyacyl coenzyme A epimerase with an active site in the amino-terminal domain of the multifunctional fatty acid oxidation protein from *Escherichia coli*. J Biol Chem 268(9): 6588-6592).

5 **[00419]** In one embodiment, the methylcrotonyl-CoA hydratase is a 3-ketoacyl-CoA thiolase. In another embodiment, the methylcrotonyl-CoA hydratase is encoded by one or more nucleic acid molecules obtained from *Escherichia coli*. In some embodiments, the one or more nucleic acid molecules encoding the methylcrotonyl-CoA hydratase is fadA, or homolog thereof. In a further embodiment, the one or more nucleic acid molecules encoding
10 the methylcrotonyl-CoA hydratase comprises an amino acid sequence set forth in SEQ ID NO: 131. In yet another embodiment, the one or more nucleic acid molecules encoding the methylcrotonyl-CoA hydratase is encoded by a nucleic acid sequence set forth in SEQ ID NO: 130.

[00420] In one embodiment, the methylcrotonyl-CoA hydratase is an enoyl-CoA
15 hydratase. In another embodiment, the methylcrotonyl-CoA hydratase is encoded by one or more nucleic acid molecules obtained from *Escherichia coli*. In some embodiments, the one or more nucleic acid molecules encoding the methylcrotonyl-CoA hydratase is fadB, or homolog thereof. In a further embodiment, the one or more nucleic acid molecules encoding the methylcrotonyl-CoA hydratase comprises an amino acid sequence set forth in SEQ ID
20 NO: 133. In yet another embodiment, the one or more nucleic acid molecules encoding the methylcrotonyl-CoA hydratase is encoded by a nucleic acid sequence set forth in SEQ ID NO: 132.

3-hydroxy-isovaleryl-CoA thioesterase (EC 3.1.2.-)

25 **[00421]** The present disclosure describes enzymes that can catalyze the following reactions:

[00422] 3-hydroxyisovaleryl-CoA + H₂O ↔ 3-hydroxyisovalerate + CoA

[00423] an acyl-CoA + H₂O → a carboxylate + coenzyme A + H⁺

[00424] An exemplary acyl-CoA thioesterase is TesB. Thioesterase II (TesB) is one of a
30 number of thioesterases present in *E. coli*. The enzyme has relatively broad substrate specificity, cleaving medium- and long-chain acyl-CoA substrates; the best tested substrate was 3,5-tetradecadienoyl-CoA (Nie L et al. (2008) A novel paradigm of fatty acid beta-oxidation exemplified by the thioesterase-dependent partial degradation of conjugated linoleic acid that fully supports growth of *Escherichia coli*. Biochemistry 47(36): 9618-9626).

Thioesterase II is one of the thioesterases supporting growth on oleate or conjugated linoleic acid as the sole source of carbon (Nie et al. 2008).

[00425] A crystal structure of the enzyme has been solved at 1.9 Å resolution. The D204 residue was predicted to be in the active site; its importance was confirmed by kinetic analysis of mutants (Li J et al. (2000) Crystal structure of the *Escherichia coli* thioesterase II, a homolog of the human Nef binding enzyme. Nat Struct Biol 7(7): 555-559).

[00426] Strains either lacking or overproducing tesB have no obvious defect (Narasimhan ML et al. (1986) Genetic and biochemical characterization of an *Escherichia coli* K-12 mutant deficient in acyl-coenzyme A thioesterase II. J Bacteriol 165(3): 911-917; Naggert J et al. (1991) Cloning, sequencing, and characterization of *Escherichia coli* thioesterase II. J Biol Chem 266(17): 11044-11050). Overproduction of TesB relieves inhibition of fatty acid synthesis by long-chain acyl-ACP molecules that accumulate upon glycerol starvation (Jiang P and Cronan JE (1994) Inhibition of fatty acid synthesis in *Escherichia coli* in the absence of phospholipid synthesis and release of inhibition by thioesterase action. J Bacteriol 176(10): 2814-2821).

[00427] In one embodiment, the 3-hydroxy-isovaleryl-CoA thioesterase is encoded by one or more nucleic acid molecules obtained from *E. coli*. In some embodiments, the one or more nucleic acid molecules encoding the 3-hydroxy-isovaleryl-CoA thioesterase is tesB, or homolog thereof. In a further embodiment, the one or more nucleic acid molecules encoding the 3-hydroxy-isovaleryl-CoA thioesterase comprises an amino acid sequence set forth in SEQ ID NO: 135. In yet another embodiment, the one or more nucleic acid molecules encoding the 3-hydroxy-isovaleryl-CoA thioesterase is encoded by a nucleic acid sequence set forth in SEQ ID NO: 134.

Mevalonate-3-kinase (EC 2.7.1.-)

[00428] The present disclosure describes enzymes that can catalyze the following reaction:

[00429] 3-hydroxyisovalerate + ATP $\rightleftharpoons$ ADP + H(+) + 3-phosphonoxisovalerate

[00430] (R)-mevalonate + ATP $\rightarrow$ (R)-mevalonate 3-phosphate + ADP + H+

[00431] Mevalonate-3-kinase may also be known as (R)-MVA 3-phosphotransferase or 3-hydroxyisovalerate (3HIV) kinase.

[00432] The subunit structure of this enzyme from *Thermoplasma acidophilum* has not been reported.

[00433] The mevalonate-3-kinase from the thermophilic archaeon *Thermoplasma acidophilum* is thought to participate in a variant of the mevalonate pathway found in archaea Azami Y et al. (2014) (R)-Mevalonate 3-Phosphate Is an Intermediate of the Mevalonate Pathway in *Thermoplasma acidophilum*. J Biol Chem 289(23): 15957-15967;
 5 Vinokur JM et al. (2014) Evidence of a Novel Mevalonate Pathway in Archaea. Biochemistry 53(25): 4161–4168).

[00434] Recombinant His-tagged enzyme was expressed in *Escherichia coli*, purified and characterized. Despite its homology with diphosphomevalonate decarboxylase, it showed no decarboxylase activity (Azami et al. 2014; Vinokur et al. 2014). The enzyme showed weak
 10 phosphomevalonate kinase activity, producing small amounts of (R)-mevalonate diphosphate (Azami et al. 2014). It had no mevalonate-5-kinase activity (Vinokur et al. 2014).

[00435] In one embodiment, the 3HIV kinase is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Thermoplasma sp.* and *Picrophilus sp.* In another embodiment, the 3HIV kinase is encoded
 15 by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Thermoplasma acidophilum* and *Picrophilus torridus*. In some embodiments, the one or more nucleic acid molecules encoding the 3HIV kinase is TA1305 and/or PTO1356, or homolog thereof. In some embodiments, the TA1305 comprises a L200E mutation. In a further embodiment, the one or more nucleic acid molecules encoding
 20 the 3HIV -kinase comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 113, 115 and 117. In yet another embodiment, the one or more nucleic acid molecules encoding the 3HIV kinase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 112, 114 and 116.

Mevalonate diphosphate decarboxylase (EC 4.1.1.-)

[00436] The present disclosure describes enzymes that can catalyze the following reactions:

[00437] (R)-mevalonate diphosphate + ATP → isopentenyl diphosphate + CO₂ + ADP + phosphate

[00438] 3-phosphonoxyisovalerate → CO₂ + isobutene

[00439] 3-hydroxyisovalerate → CO₂ + isobutene

[00440] Mevalonate diphosphate decarboxylase may also be known as pyrophosphomevalonate decarboxylase, mevalonate-5-pyrophosphate decarboxylase,

pyrophosphomevalonic acid decarboxylase, 5-pyrophosphomevalonate decarboxylase, mevalonate 5-diphosphate decarboxylase, and ATP:(R)-5-diphosphomevalonate carboxy-lyase (dehydrating), 3-phosphonoxyisovalerate decarboxylase, 3-hydroxyisovalerate-3-phosphate decarboxylase, 3HIV-3-phosphate decarboxylase, 3-hydroxyisovalerate decarboxylase and 3HIV decarboxylase.

[00441] This enzyme converts mevalonate 5-diphosphate (MVAPP) to isopentenyl diphosphate (IPP) through ATP dependent decarboxylation. The two substrates of this enzyme are ATP and mevalonate 5-diphosphate, whereas its four products are ADP, phosphate, isopentenyl diphosphate, and CO₂.

[00442] Mevalonate diphosphate decarboxylase catalyzes the final step in the mevalonate pathway. The mevalonate pathway is responsible for the biosynthesis of isoprenoids from acetate. This pathway plays a key role in multiple cellular processes by synthesizing sterol isoprenoids, such as cholesterol, and non-sterol isoprenoids, such as dolichol, heme A, tRNA isopentenyltransferase, and ubiquinone. This enzyme belongs to the family of lyases, specifically the carboxy-lyases, which cleave carbon-carbon bonds.

[00443] Mevalonate diphosphate decarboxylase recognizes and binds two substrates: ATP and mevalonate 5-diphosphate. After binding, the enzyme performs three types of reactions that can be separated into two main stages. First, phosphorylation occurs. This creates a reactive intermediate, which in the second stage undergoes concerted dephosphorylation and decarboxylation.

[00444] In one embodiment, the enzyme that catalyzes the reaction 3-phosphonoxyisovalerate → CO₂ + isobutene is a 3HIV-3-phosphate decarboxylase. In another embodiment, the 3HIV-3-phosphate decarboxylase is encoded by one or more nucleic acid molecules obtained from *Streptococcus* sp. In some embodiments, the microorganism is selected from *Streptococcus mitis* and/or *Streptococcus gordonii*. In some embodiments, the one or more nucleic acid molecules encoding the 3HIV-3-phosphate decarboxylase comprises an amino acid sequence selected from SEQ ID NOs: 119 and 121. In further embodiments, the one or more nucleic acid molecule encoding the 3HIV-3-phosphate decarboxylase is encoded by a nucleic acid sequence selected from SEQ ID NOs: 118 and 120.

[00445] In one embodiment, the enzyme that catalyzes the reaction 3-hydroxyisovalerate → CO₂ + isobutene is a 3HIV decarboxylase. In another embodiment, the 3HIV decarboxylase is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Streptococcus* sp., *Thermoplasma* sp. and *Picrophilus* sp. In another embodiment, the 3HIV decarboxylase is encoded by one or

more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Streptococcus gordonii*, *Thermoplasma acidophilum* and *Picrophilus torridus*. In some embodiments, the one or more nucleic acid molecules encoding the 3HIV decarboxylase comprises mvaD, TA1305 and/or PTO1356, or homolog thereof. In a further embodiment, the one or more nucleic acid molecules encoding the 3HIV decarboxylase comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 113, 117 and 121. In yet another embodiment, the one or more nucleic acid molecules encoding the 3HIV decarboxylase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 112, 116 and 120.

Transketolase (EC 2.2.1.1)

[00446] The present disclosure describes enzymes that can catalyze the following reactions:

[00447] D-erythrose 4-phosphate + D-xylulose 5-phosphate $\leftrightarrow$ β -D-fructofuranose 6-phosphate + D-glyceraldehyde 3-phosphate

[00448] D-sedoheptulose 7-phosphate + D-glyceraldehyde 3-phosphate $\leftrightarrow$ D-ribose 5-phosphate + D-xylulose 5-phosphate

[00449] Transketolase may also be known as glycolaldehydetransferase.

[00450] Transketolase catalyzes the reversible transfer of a ketol group between several donor and acceptor substrates. This key enzyme is a reversible link between glycolysis and the pentose phosphate pathway. The enzyme is involved in the catabolism of pentose sugars, the formation of D-ribose 5-phosphate, and the provision of D-erythrose 4-phosphate, a precursor of aromatic amino acids and PLP. *E. coli* contains two transketolase isozymes, TktA and TktB. TktA is responsible for the major transketolase activity.

[00451] In addition to its function in central carbon metabolism, transketolase appears to also have an unexpected role in chromosome structure; a tktA mutant affects chromosome topology.

[00452] Crystal structures of TktA in complex with donor and acceptor substrates have been solved, elucidating the reaction mechanism and mode of action of transketolase. A computational model of transketolase activity using a quantum mechanical/molecular mechanical method has been proposed, defining a new route for thiamine diphosphate activation. Transketolase I (TktA) is homodimeric. The urea denaturation pathways of wild type and active site mutants of TktA have been investigated, and the effects of temperature

and pH on the structure, stability, aggregation and activity of transketolase have been determined. The acceptor specificity of TktA has been investigated.

[00453] TktA abundance is affected by the SOS inducer and mutagen 7-methoxy-2-nitronaphtho[2,1-b]furan (R7000). tktA is negatively regulated during entry into stationary phase. The effect by RpoS is likely indirect and might be mediated by an intermediate regulator that itself is directly regulated by RpoS.

[00454] The subunit structure of transketolase II (TktB) has not been explicitly determined. Overproduction of TktB suppresses the tktA mutant phenotype. Expression of tktB is increased in a tyrR mutant in the presence of phenylalanine. tktB expression is increased in stationary phase and positively regulated by RpoS and ppGpp. Levels of TktB protein increase during osmotic stress under aerobic, but not anaerobic growth conditions. TktB appears to be associated with the degradosome and may connect carbon metabolism to replication.

[00455] Expression of tktA and tktB is complementary, resulting in approximately constant levels of transketolase expression throughout growth.

[00456] In some embodiments, the transketolase is encoded by an amino acid sequence having at least 70% sequence identity to tktA from *E. coli*. In a further embodiment, the transketolase is encoded by an amino acid sequence having at least 80% sequence identity to tktA from *E. coli*. In yet a further embodiment, the transketolase is encoded by an amino acid sequence having at least 90% sequence identity to tktA from *E. coli*. In other embodiments, the transketolase is tktA from *E. coli*. In some embodiments, the transketolase is encoded by an amino acid sequence having at least 70% sequence identity to tktB from *E. coli*. In a further embodiment, the transketolase is encoded by an amino acid sequence having at least 80% sequence identity to tktB from *E. coli*. In yet a further embodiment, the transketolase is encoded by an amino acid sequence having at least 90% sequence identity to tktB from *E. coli*. In other embodiments, the transketolase is tktB from *E. coli*.

[00457] In some embodiments, the one or more nucleic acid molecules encoding the transketolase is tktA, or homolog thereof. In some embodiments, the one or more nucleic acid molecules encoding the transketolase is tktB, or homolog thereof. In another embodiment, the one or more nucleic acid molecules encoding the transketolase comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 148 and 150. In a further embodiment, the one or more nucleic acid molecule encoding the transketolase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 147 and 149.

Transaldolase (EC 2.2.1.2)

[00458] The present disclosure describes enzymes that can catalyze the following reaction:

5 **[00459]** D-sedoheptulose 7-phosphate + D-glyceraldehyde 3-phosphate $\leftrightarrow$ β -D-fructofuranose 6-phosphate + D-erythrose 4-phosphate

[00460] Transaldolase may also be known as dihydroxyacetone transferase; dihydroxyacetone synthase; formaldehyde transketolase.

10 **[00461]** Transaldolase B is an enzyme of the non-oxidative branch of the pentose phosphate pathway. Along with transketolase, transaldolase creates a reversible link between the pentose phosphate pathway and glycolysis. It catalyzes the interconversion of glyceraldehyde-3-phosphate and sedoheptulose-7-phosphate to fructose-6-phosphate and erythrose-4-phosphate. The reversibility of this reaction and carbon flux through the pentose phosphate pathway has been addressed both experimentally and theoretically.

15 **[00462]** There are two closely related transaldolases in *E. coli*, encoded by talA and talB. Only transaldolase B has been biochemically characterized. TalB is a dimer in solution and in the crystal structure. Mutation of the R300 residue leads to the formation of catalytically active monomers. Catalytically important active site residues have been identified by site-directed mutagenesis.

20 **[00463]** Crystal structures of transaldolase B have been determined, confirming the presence of a Schiff-base intermediate at the active site and leading to a proposed reaction mechanism.

[00464] A talB null mutant has no growth defect on minimal media with glucose as the carbon source.

25 **[00465]** In some embodiments, the transaldolase is encoded by an amino acid sequence having at least 70% sequence identity to talA or talB from *E. coli*. In a further embodiment, the transaldolase is encoded by an amino acid sequence having at least 80% sequence identity to talA or talB from *E. coli*. In yet a further embodiment, the transaldolase is encoded by an amino acid sequence having at least 90% sequence identity to talA or talB from *E. coli*.
30 **[00466]** In other embodiments, the transaldolase is talA from *E. coli*. In yet further embodiments, the transaldolase is talB from *E. coli*.

[00466] In some embodiments, the one or more nucleic acid molecules encoding the transaldolase is talA, or homolog thereof. In some embodiments, the one or more nucleic

acid molecules encoding the transaldolase is talB, or homolog thereof. In another embodiment, the one or more nucleic acid molecules encoding the transaldolase comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 152 and 154. In a further embodiment, the one or more nucleic acid molecule encoding the transaldolase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 151 and 153.

Ribose-5-phosphate isomerase (EC 5.3.1.6)

[00467] The present disclosure describes enzymes that can catalyze the following reaction:

[00468] D-ribose 5-phosphate $\leftrightarrow$ D-ribulose 5-phosphate

[00469] Ribose-5-phosphate isomerase may also be known as phosphopentosisomerase; phosphoriboisomerase; ribose phosphate isomerase; 5-phosphoribose isomerase; D-ribose 5-phosphate isomerase; D-ribose-5-phosphate ketol-isomerase.

[00470] There are two physically and genetically distinct ribose-5-phosphate isomerases present in *E. coli*. The constitutive ribose-5-phosphate isomerase A (rpiA) normally accounts for more than 99% of the ribose-5-phosphate isomerase activity in the cell and functions in the pentose phosphate pathway (non-oxidative branch). The inducible ribose-5-phosphate isomerase B (rpiB) can substitute for rpiA's function if its expression is induced. There is no sequence similarity between the two enzymes.

[00471] Crystal structures of rpiA have been solved and active site residues and an acid-base catalytic mechanism were predicted. An rpiA mutant requires ribose for growth.

[00472] In some embodiments, the ribose-5-phosphate isomerase is encoded by an amino acid sequence having at least 70% sequence identity to rpiA or rpiB from *E. coli*. In a further embodiment, the ribose-5-phosphate isomerase is encoded by an amino acid sequence having at least 80% sequence identity to rpiA or rpiB from *E. coli*. In yet a further embodiment, the ribose-5-phosphate isomerase is encoded by an amino acid sequence having at least 90% sequence identity to rpiA or rpiB from *E. coli*. In other embodiments, the ribose-5-phosphate isomerase is rpiA or rpiB from *E. coli*.

[00473] In some embodiments, the one or more nucleic acid molecules encoding the ribose-5-phosphate isomerase is rpiA, or homolog thereof. In another embodiment, the one or more nucleic acid molecules encoding the ribose-5-phosphate isomerase comprise an amino acid sequence set forth in SEQ ID NO: 156. In a further embodiment, the one or more

nucleic acid molecules encoding the ribose-5-phosphate isomerase is encoded by a nucleic acid sequence set forth in SEQ ID NO: 155. In some embodiments, the one or more nucleic acid molecules encoding the ribose-5-phosphate isomerase is *rpiB*, or homolog thereof. In another embodiment, the one or more nucleic acid molecules encode the ribose-5-phosphate comprising an amino acid sequence set forth in SEQ ID NO: 253. In a further embodiment, the one or more nucleic acid molecules encoding the ribose-5-phosphate isomerase is encoded by a nucleic acid sequence set forth in SEQ ID NO: 254.

Ribulose-5-phosphate 3-epimerase (EC 5.1.3.1)

[00474] The present disclosure describes enzymes that can catalyze the following reaction:

[00475] D-ribulose 5-phosphate ↔ D-xylulose 5-phosphate

[00476] Ribulose-5-phosphate 3-epimerase may also be known as ribulose-phosphate 3-epimerase; phosphoribulose epimerase; erythrose-4-phosphate isomerase; phosphoketopentose 3-epimerase; xylulose phosphate 3-epimerase; phosphoketopentose epimerase; D-ribulose phosphate-3-epimerase; D-ribulose 5-phosphate epimerase; D-ribulose-5-P 3-epimerase; D-xylulose-5-phosphate 3-epimerase; pentose-5-phosphate 3-epimerase.

[00477] Ribulose-5-phosphate 3-epimerase (Rpe) is an enzyme of the non-oxidative branch of the pentose phosphate pathway.

[00478] Rpe requires ferrous iron for activity and is vulnerable to damage by H₂O₂ due to Fenton chemistry. Mn²⁺, Co²⁺ and Zn²⁺ can substitute for Fe²⁺ to varying degrees, and Rpe containing these alternative cations is not vulnerable to H₂O₂. Induction of the manganese transporter can protect Rpe from H₂O₂ damage.

[00479] In some embodiments, the ribulose-5-phosphate 3-epimerase is encoded by an amino acid sequence having at least 70% sequence identity to *rpe* from *E. coli*. In a further embodiment, the ribulose-5-phosphate 3-epimerase is encoded by an amino acid sequence having at least 80% sequence identity to *rpe* from *E. coli*. In yet a further embodiment, the ribulose-5-phosphate 3-epimerase is encoded by an amino acid sequence having at least 90% sequence identity to *rpe* from *E. coli*. In other embodiments, the ribulose-5-phosphate 3-epimerase is *rpe* from *E. coli*.

[00480] In some embodiments, the one or more nucleic acid molecules encoding the ribulose-5-phosphate 3-epimerase is *rpe*, or homolog thereof. In another embodiment, the

one or more nucleic acid molecules encoding the ribulose-5-phosphate 3-epimerase comprise an amino acid sequence set forth in SEQ ID NO: 158. In a further embodiment, the one or more nucleic acid molecule encoding the ribulose-5-phosphate 3-epimerase is encoded by a nucleic acid sequence set forth in SEQ ID NO: 157.

5

Fructose 6-phosphate phosphoketolase (Fpk, EC 4.1.2.22)

[00481] The present disclosure describes enzymes that can catalyze the following reaction:

[00482] D-fructose 6-phosphate + phosphate $\leftrightarrow$ acetyl phosphate + D-erythrose 4-phosphate + H₂O

10

[00483] The phosphoketolase reaction by which β -D-fructofuranose 6-phosphate is converted to D-erythrose 4-phosphate and acetyl phosphate is one of the key reactions in the Bifidobacterium shunt. There is evidence for the existence of two distinct F6P-phosphoketolase enzymes in bifidobacteria. One is specific solely for F6P while the other is able to utilize both F6P and D-xylulose 5-phosphate (EC: 4.1.2.9), a reaction that appears later in the Bifidobacterium shunt. The enzyme encoded by the xfp gene, originally discovered in *Bifidobacterium animalis* lactis, is the dual-specificity enzyme. A phosphoketolase has also been purified from *Leuconostoc nesenteroides* (LEUM_1961).

15

[00484] In some embodiments, an enzyme having fructose-6-phosphate phosphoketolase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having fructose-6-phosphate phosphoketolase activity selected from the group consisting of *Bifidobacterium dentium* BDP_1006, *Bifidobacterium lactis* xfp, *Lactobacillus paraplantarum* xpkA and *Bifidobacterium breve* xfp. In a preferred embodiment, an enzyme having fructose-6-phosphate phosphoketolase activity is selected from the group consisting of *Bifidobacterium dentium* BDP_1006, *Bifidobacterium lactis* xfp, *Lactobacillus paraplantarum* xpkA and *Bifidobacterium breve* xfp. In another embodiment, the one or more nucleic acid molecules encoding the fructose-6-phosphate phosphoketolase comprise an amino acid sequence selected from the group consisting of SEQ ID NOs: 212, 214, 216 and 218. In a further embodiment, the one or more nucleic acid molecule encoding the fructose-6-phosphate phosphoketolase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 211, 213, 215 and 217.

20

25

30

Phosphate acetyltransferase (EC 2.3.1.8)

[00485] The present disclosure describes enzymes that can catalyze the following reaction:

[00486] $\text{Acetyl-CoA} + \text{phosphate} \leftrightarrow \text{CoA} + \text{acetyl phosphate}$

[00487] Phosphate acetyltransferase (Pta) catalyzes the reversible conversion between acetyl-CoA and acetylphosphate, a step in the metabolism of acetate. Both pyruvate and phosphoenolpyruvate activate the enzyme in the direction of acetylphosphate synthesis and inhibit the enzyme in the direction of acetyl-CoA synthesis. The acetate formation from acetyl-CoA pathway has been the target of metabolic engineering to reduce the flux to acetate and increase the production of commercially desired end products. It has also been studied using systems biology approaches such as metabolic modeling and flux balance analysis.

[00488] Pta is composed of three domains; only the C-terminal domain is required for phosphate acetyltransferase activity. The N-terminal domain is involved in stabilization of the native quaternary structure and metabolic regulation.

[00489] Pta may be able to utilize both acetyl-CoA and propionyl-CoA. An ack pta double mutant has reduced levels of propionate from L-threonine, suggesting that the enzyme is part of the anaerobic pathway metabolizing L-threonine to propionate. A pta mutant does not grow on acetate as the sole source of carbon. Both pta and pta ackA mutants are impaired in their ability to survive glucose starvation. The growth defect of a pta mutant appears to be due to perturbation of acetyl-CoA flux. pta mutants produce large amounts of lactate when grown on glucose as the carbon source under microaerophilic conditions. The effect of a pta mutation on metabolism, enzyme activity and gene expression has been thoroughly studied recently. pta and recBC mutants are synthetically growth inhibited.

[00490] Levels of Pta are decreased by growth on acetate and under low pH conditions. pta belongs to the CreBC regulon. FNR has a slightly positive effect on pta expression. The growth-rate dependent expression pattern of pta-ackA was measured.

[00491] The pta gene that encodes the enzyme has been cloned from *Clostridium acetobutylicum*, sequenced and expressed in *Escherichia coli*. The gene is adjacent to the ackA gene, which encodes the enzyme that catalyzes the second step - acetate kinase. Enzyme activity assays performed on cell extracts from *Escherichia coli* and *Clostridium acetobutylicum* harboring the subclone showed elevated activity. The enzymes shows a decrease in specific activity when the organism reaches the solvent formation stage.

[00492] Enzymes having phosphate acetyltransferase activity or phosphate acetyltransferase genes have also been identified or measured from *E. coli* (eutD, pta),

Roseovarius nubinhibens ISM, *Clostridium kluyveri*, *Chlamydomonas reinhardtii* (PAT2), *Dasytricha ruminantium*, *Pelobacter acetylenicus*, *Gottschalkia acidurici*, *Lactobacillus sanfranciscensis*, *Paracoccus denitrificans* NKNIS, *Eubacterium oxidoreducens* G41, *Mycoplasma pneumoniae* M129, *Thermotoga maritima*, *Moorella thermoacetica*,
 5 *Methanosarcina thermophila*, *Clostridium propionicum* and *Fusobacterium nucleatum*.

[00493] In some embodiments, an enzyme having phosphate acetyltransferase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having phosphate acetyltransferase activity selected from *E. coli* pta and *Clostridium acetobutylicum* pta. In a
 10 preferred embodiment, an enzyme having phosphate acetyltransferase activity is selected from *E. coli* pta and *Clostridium acetobutylicum* pta. In another embodiment, the one or more nucleic acid molecules encoding the phosphate acetyltransferase comprise an amino acid sequence selected from SEQ ID NOs: 220 and 222. In a further embodiment, the one or
 15 more nucleic acid molecule encoding the phosphate acetyltransferase is encoded by a nucleic acid sequence selected from SEQ ID NOs: 219 and 221.

Pentose phosphatase (EC 3.1.3.23 and EC 3.1.3.1)

[00494] The present disclosure describes enzymes that can catalyze the following reactions:

20 **[00495]** D-pentose 5-phosphate + H₂O → D-pentose + phosphate

[00496] D-ribose 5-phosphate + H₂O → D-ribose + phosphate

[00497] D-xylulose 5-phosphate + H₂O → D-xylulose + phosphate

[00498] D-ribulose 5-phosphate + H₂O → D-ribulose + phosphate

[00499] a sugar phosphate + H₂O → a sugar + phosphate

25 **[00500]** *E. coli* YbiV is a sugar phosphatase belonging to the family of type II haloacid dehalogenase (HAD)-like hydrolases. It shows a low level of discrimination between its preferred substrates. In addition, YbiV appears to have a low level of phosphotransferase activity using monophosphates as the phosphate donor. The phosphatase activity of YbiV was also discovered in a high-throughput screen of purified proteins. Crystal structures of
 30 YbiV have been solved, and a catalytic mechanism was suggested. YbiV may exist as a homodimer in solution.

[00501] *E. coli* YidA is a promiscuous sugar phosphatase belonging to the superfamily of haloacid dehalogenase (HAD)-like hydrolases. Its preferred substrate is erythrose-4-

phosphate. YidA selectively hydrolyzes α -D-glucose-1-phosphate and has no activity with the β form. The reaction proceeds via the canonical phosphomonoester hydrolase mechanism, which involves breakage of the P-O bond, not the C1-O bond. The phosphatase activity of YidA was first discovered in a high-throughput screen of purified proteins.

5 Mutagenesis of the predicted catalytic Asp residue in YidA results in loss of phosphatase activity. YidA does not catalyze phosphoryl transfer to a sugar acceptor.

[00502] *E. coli* alkaline phosphatase (phoA) is a periplasmic, homodimeric enzyme that catalyses the hydrolysis and transphosphorylation of a wide variety of phosphate monoesters. The reaction proceeds through a phosphoseryl intermediate with the
10 subsequent release of inorganic phosphate and alcohol. The transphosphorylation reaction results in the transfer of a phosphoryl group to the alcohol of acceptors such as Tris or ethanolamine. Alkaline phosphatase is a metalloenzyme, binding two zinc atoms and one magnesium ion per monomer. Alkaline phosphatase occurs in three major forms designated isozymes 1, 2 and 3 whose relative proportions are dependent on the growth conditions. The
15 isozymes are differentiated by the presence or absence of an NH₂-terminal arginine residue: present in both subunits of isozyme 1, absent in both subunits of isozyme 3 and heterogenous in isozyme 2. Removal of the N-terminal arginine is catalysed by the membrane-associated, proteolytic enzyme Iap. The precursor polypeptide is secreted across the inner membrane to the periplasmic space concomitant with removal of the
20 signal sequence. Folding of PhoA *in vivo* is catalysed by the periplasmic protein, DsbA and is thought to occur as the polypeptide elongates from the ribosome. phoA is part of the phosphate regulon; its expression is positively regulated by the PhoB transcriptional regulator.

[00503] In some embodiments, the pentose phosphatase is an alkaline phosphatase. In
25 some preferred embodiments, the alkaline phosphatase may be evolved to preferably act on targeted pentose 5-phosphate. In some embodiments, the pentose phosphatase is a sugar phosphatase. In other embodiments, the sugar phosphatase may be evolved to preferably act on targeted pentose 5-phosphate. In some embodiments, the pentose phosphatase is a haloacid dehalogenase-like hydrolase. In another embodiment, the
30 haloacid dehalogenase-like hydrolase may be evolved to preferably act on targeted pentose 5-phosphate.

[00504] In some embodiments, the pentose phosphatase is selected from one or more of a D-pentose-5-phosphatase, a D-xylulose-5-phosphatase, a D-ribose-5-phosphatase, and a D-ribulose-5-phosphatase. In some embodiments, the pentose phosphatase is encoded by
35 an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to a D-pentose-5-phosphatase selected from the

group consisting of *E. coli* phoA, *E. coli* yfbT and *E. coli* yidA. In some embodiments, the D-xylulose-5-phosphatase is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to *Bacillus subtilis* araL. In some embodiments, the D-ribose-5-phosphatase is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to a D-ribose-5-phosphatase selected from the group consisting of *Arabidopsis thaliana* SGPP, *Pseudomonas fluorescens* PFLU_2693, and *E. coli* ybiV. In some embodiments, the D-ribulose-5-phosphatase is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to *Plasmodium falciparum* PF10_0325.

[00505] In some embodiments, the one or more nucleic acid molecules encoding the D-pentose-5-phosphatase is selected from the group consisting of SEQ ID NOs: 159, 161, 163, 165, 167, 169, 171 and 173. In another embodiment, the one or more nucleic acid molecules encoding the D-pentose-5-phosphatase comprise an amino acid sequence selected from the group consisting of SEQ ID NOs: 160, 162, 164, 166, 168, 170, 172 and 174.

Arabitol phosphate dehydrogenase (EC 1.1.1.-)

[00506] The present disclosure describes enzymes that can catalyze the following reactions:

[00507] D-arabitol 1-phosphate + NAD⁺ ↔ D-xylulose 5-phosphate + NADH + H⁺

[00508] D-arabitol 5-phosphate + NAD⁺ ↔ D-ribulose 5-phosphate + NADH + H⁺

[00509] D-arabitol 1-phosphate + NAD⁺ ↔ D-ribulose 1-phosphate + NADH + H⁺

[00510] D-arabitol 5-phosphate + NAD⁺ ↔ D-xylulose 1-phosphate + NADH + H⁺

[00511] D-arabitol-phosphate dehydrogenase APDH from *Enterococcus avium* has been purified to homogeneity. The protein, which forms a homotetramer, catalyzes the dehydrogenation of both D-arabitol 1-phosphate and D-arabitol 5-phosphate, producing D-xylulose 5-phosphate and D-ribulose 5-phosphate, respectively. Maximal velocity with D-arabitol 1-phosphate was 10-fold than with D-arabitol 5-phosphate. The purified protein was partially sequenced, and the APDH gene encoding it was cloned. The enzyme requires Mn²⁺ and can not utilize Zn²⁺ for activity. Both NAD(+) and NADP(+) were accepted as cofactors, but reaction rates with NAD⁺/NADH were about 14 times higher than with NADP⁺/NADPH. The enzyme catalyzes a reversible reaction, but the rate of the reduction reaction is much higher than the oxidative reaction. Kinetic data suggests that the enzyme

forms a ternary complex with its substrate and NADH. Both biochemical evidence and protein sequence homology comparisons indicate that similar enzymes are widespread among the Gram-positive bacteria and participate in arabitol catabolism.

[00512] In some embodiments, the pentol dehydrogenase activity may be evolved to apply to pentol phosphate.

[00513] In some embodiments, the arabitol phosphate dehydrogenase is selected from one or more of an enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity, an enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity, an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity and an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity.

[00514] In some embodiments, an enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to APDH from *Enterococcus avium*. In a preferred embodiment, an enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity is APDH from *Enterococcus avium*. In some embodiments, the one or more nucleic acid molecules encoding a D-arabitol 1-phosphate 4-dehydrogenase is set forth in SEQ ID NO: 175. In another embodiment, the one or more nucleic acid molecules encoding a D-arabitol 1-phosphate 4-dehydrogenase comprise an amino acid sequence set forth in SEQ ID NO: 176.

[00515] In some embodiments, an enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to APDH from *Enterococcus avium*. In a preferred embodiment, an enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity is APDH from *Enterococcus avium*. In some embodiments, the one or more nucleic acid molecules encoding a D-arabitol 5-phosphate 2-dehydrogenase is set forth in SEQ ID NO: 175. In another embodiment, the one or more nucleic acid molecules encoding a D-arabitol 5-phosphate 2-dehydrogenase comprise an amino acid sequence set forth in SEQ ID NO: 176.

[00516] In some embodiments, an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity selected from the group consisting of *Candida albicans* ARD1, *Candida tropicalis* ARD1, *Scheffersomyces stipitis* ARDH, *Bacillus subtilis* egsA (araM), *Aeropyrum pernix* egsA, *E. coli* gpsA and *Saccharomyces cerevisiae* GPD1. In a preferred embodiment, an enzyme having D-arabitol

1-phosphate 2-dehydrogenase activity is selected from the group consisting of *Candida albicans* ARD1, *Candida tropicalis* ARD1, *Scheffersomyces stipitis* ARDH, *Bacillus subtilis* egsA (araM), *Aeropyrum pernix* egsA, *E. coli* gpsA and *Saccharomyces cerevisiae* GPD1. In some embodiments, the one or more nucleic acid molecules encoding a D-arabitol 1-phosphate 2-dehydrogenase is selected from the group consisting of SEQ ID NOs: 177, 179, 181, 189, 191, 193 and 195. In another embodiment, the one or more nucleic acid molecules encoding a D-arabitol 1-phosphate 2-dehydrogenase comprise an amino acid sequence selected from the group consisting of SEQ ID NOs: 178, 180, 182, 190, 192, 194 and 196.

[00517] In some embodiments, an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity selected from the group consisting of *Pseudomonas fluorescens* mtlD, *Klebsiella pneumoniae* dalD, *Ralstonia solanacearum* dalD, *Bacillus subtilis* egsA (araM), *Aeropyrum pernix* egsA, *E. coli* gpsA and *Saccharomyces cerevisiae* GPD1. In a preferred embodiment, an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity is selected from the group consisting of *Pseudomonas fluorescens* mtlD, *Klebsiella pneumoniae* dalD, *Ralstonia solanacearum* dalD, *Bacillus subtilis* egsA (araM), *Aeropyrum pernix* egsA, *E. coli* gpsA and *Saccharomyces cerevisiae* GPD1. In some embodiments, the one or more nucleic acid molecules encoding a D-arabitol 5-phosphate 4-dehydrogenase is selected from the group consisting of SEQ ID NOs: 183, 185, 187, 189, 191, 193 and 195. In another embodiment, the one or more nucleic acid molecules encoding a D-arabitol 5-phosphate 4-dehydrogenase comprise an amino acid sequence selected from the group consisting of SEQ ID NOs: 184, 186, 188, 190, 192, 194 and 196.

Phosphopentomutase (EC 5.4.2.-)

[00518] The present disclosure describes enzymes that can catalyze the following reactions:

[00519] D-xylulose 5-phosphate ↔ D-xylulose 1-phosphate

[00520] D-ribulose 5-phosphate ↔ D-ribulose 1-phosphate

[00521] α-D-ribose-1-phosphate ↔ D-ribose 5-phosphate

[00522] 2-deoxy-α-D-ribose 1-phosphate ↔ 2-deoxy-D-ribose 5-phosphate

[00523] *E. coli* phosphopentomutase deoB is a catabolic enzyme which catalyzes the transfer of a phosphate group between the C1 and the C5 carbon atoms of ribose and deoxyribose, respectively.

[00524] A mutation in deoB suppresses the high thymine requirement for growth of thy mutants and improves the survival of thyA mutants in stationary phase. Transposon insertion mutations in deoB suppress the growth defect of a tktA tktB mutant. Deletion of deoB increases glycerol consumption as well as hydrogen and ethanol production compared to wild type, and increases lycopene production in an engineered strain. The deo operon has a complex pattern of regulation. Expression of deoB is downregulated by nitrogen starvation. The *E. coli* phosphopentomutase appears to be biochemically and structurally distinct from mammalian phosphopentomutase, making it a potential target for antibiotic development.

[00525] In some embodiments, the phosphopentomutase (PPM), also named here as phosphosugarmutase (PSM) also has phosphoglucomutase activity, and is *Saccharomyces cerevisiae* phosphoribomutase PRM15 (also known as PGM3) (SEQ ID NO. 255 or SEQ ID NO: 258). In some embodiments, the phosphopentosemutase is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity or at least 90% sequence identity to the Pgm3 from *Saccharomyces cerevisiae*.

[00526] In some embodiments, the phosphopentomutase is *Homo sapiens* phosphoglucomutase-2 (PGM2).

[00527] In some embodiments, a phosphopentomutase activity may be evolved to apply to D-xylulose 5-phosphate. In other embodiments, a phosphopentomutase activity may be evolved to apply to D-ribulose 5-phosphate. In some embodiments, an alpha-phosphoglucomutase activity may be evolved to apply to D-ribulose 5-phosphate or D-xylulose 5-phosphate. This enzyme from the class EC 5.4.2.2 is reported to require α -glucose 1,6-bisphosphate as co-factor. In further embodiments, a beta-phosphoglucomutase activity may be evolved to apply to D-ribulose 5-phosphate or D-xylulose 5-phosphate. Enzymes from the class EC 5.4.2.6 are reported to be able to autophosphorylate themselves, thus not requiring external glucose 1,6-bisphosphate as co-factor. In yet further embodiments, a phosphomannomutase activity may be evolved to apply to D-ribulose 5-phosphate or D-xylulose 5-phosphate. This enzyme from the class EC 5.4.2.8 is reported to require α -glucose 1,6-bisphosphate or α -D-Mannose 1,6-bisphosphate as co-factor.

[00528] In some embodiments, an enzyme having phosphopentomutase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having phosphopentomutase activity selected from the group consisting of *E. coli* deoB, *E. coli* pgm,

Bacillus subtilis pgcA, *Lactococcus lactis* pgmB, *E. coli* ycjU, *Pseudomonas aeruginosa* algC, and *E. coli* cpsG. In a preferred embodiment, an enzyme having phosphopentomutase activity is selected from the group consisting of *E. coli* deoB, *E. coli* pgm, *Bacillus subtilis* pgcA, *Lactococcus lactis* pgmB, *E. coli* ycjU, *Pseudomonas aeruginosa* algC, *S. cerevisiae* PGM3, and *E. coli* cpsG. In some embodiments, the one or more nucleic acid molecules encoding a phosphopentomutase is selected from the group consisting of SEQ ID NOs: 197, 199, 201, 203, 205, 207 and 209. In another embodiment, the one or more nucleic acid molecules encoding a phosphopentomutase comprise an amino acid sequence selected from the group consisting of SEQ ID NOs: 198, 200, 202, 204, 206, 208 and 210.

Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)

[00529] The present disclosure describes enzymes that can catalyze the following reaction:

[00530] D-glucose 6-phosphate + NADP⁺ ↔ 6-phospho-D-glucono-1,5-lactone + NADPH

[00531] Glucose-6-phosphate 1-dehydrogenase may also be known as glucose-6-phosphate dehydrogenase (NADP⁺); NADP-glucose-6-phosphate dehydrogenase; Zwischenferment; D-glucose 6-phosphate dehydrogenase; glucose 6-phosphate dehydrogenase (NADP); NADP-dependent glucose 6-phosphate dehydrogenase; 6-phosphoglucose dehydrogenase; Entner-Doudoroff enzyme; G6PDH; GPD; glucose-6-phosphate dehydrogenase.

[00532] Glucose-6-phosphate dehydrogenase (G6PDH) is the first enzyme of the pentose phosphate pathway and provides a large fraction of the NADPH needed for anabolism.

[00533] The *E. coli* G6PDH shows a strong preference for NADP⁺ over NAD⁺. The structural basis for this preference was studied using molecular simulations, kinetic characterization of site-directed mutants and phylogenetic analyses.

[00534] Metabolic flux through the pathways of central carbon metabolism was measured using GC-MS and ¹³C labeling and 2D NMR spectroscopy. Regulation of these pathways under different growth conditions was measured at the level of enzyme expression and activity.

[00535] Substitution of an NADH-producing glucose-6-phosphate dehydrogenase for the native NADPH-producing enzyme reduced the growth rate of otherwise wild-type cells, while increasing the growth rate of a Δpgi mutant. This suggests that whether production of NADH

by G6PDH is beneficial or detrimental *in vivo* depends on the operation of the upper Embden-Meyerhof pathway.

[00536] In addition to its role in central carbon metabolism, G6PDH was found to be the source of a linear peptide with the amino acid sequence Asn-Asn-Trp-Asn-Asn (NNWNN) that acts as an "extracellular death factor" (EDF) for MazEF-mediated cell death. The peptide acts by increasing the endoribonuclease activity of the toxins MazF and ChpBK. EDF production under stress conditions is due to cleavage of the *zwf* mRNA at specific ACA sites by MazF, generating a leaderless truncated mRNA. The location of the EDF-coding region with respect to the MazF cleavage sites is important, and the trans-translation system is required.

[00537] *zwf* is one of the most consistently flux-coupled genes, which are genes whose expression transition patterns upon perturbations are correlated with their corresponding flux values. Expression of *zwf* is growth rate-regulated at the transcriptional level. G6PDH activity is greater in rapidly growing cells, and is greater under nitrogen-limited compared to carbon-limited growth conditions. *zwf* is a part of the SoxRS regulon which responds to superoxide stress. Additional regulators have been shown to activate transcription of *zwf*. Exposure to tellurite activates transcription of *zwf* and thereby increases the synthesis of NADPH.

[00538] A *zwf* null mutation does not affect the growth rate significantly. However, central carbon metabolism and metabolic flux is changed. A *pgi zwf* double mutant does not grow on glucose as the sole source of carbon. In the presence of glucose, it accumulates high levels of glucose-6-phosphate, which inhibits fructose-1,6-bisphosphatase I activity. Deletion of *zwf* reduces the organic solvent tolerance of *E. coli* JM109.

[00539] In some embodiments, a recombinant microorganism producing MEG (or glycolic acid), or optionally, MEG (or glycolic acid) and one or more co-product, comprises a deletion, insertion, or loss of function mutation in a gene encoding a glucose-6-phosphate dehydrogenase to prevent the flux of glucose-6-phosphate through the oxidative branch of the pentose phosphate pathway and instead shunt glucose-6-phosphate through the non-oxidative branch of the pentose phosphate pathway to produce one or more pentose-5-phosphate intermediate.

6-phosphogluconolactonase (EC 3.1.1.31)

[00540] The present disclosure describes enzymes that can catalyze the following reaction:

[00541] 6-phospho D-glucono-1,5-lactone + H₂O → D-gluconate 6-phosphate + H⁺

[00542] 6-phosphogluconolactonase may also be known as phosphogluconolactonase; 6-PGL.

[00543] 6-phosphogluconolactonase is an enzyme of the oxidative pentose phosphate pathway.

5 [00544] A *pgl* mutant strain grows only slightly slower than wild type on glucose as the sole source of carbon. Growth on glucose may be due to non-enzymatic hydrolysis of 6-phospho D-glucono-1,5-lactone or a bypass pathway that involves dephosphorylation and export of gluconolactone, hydrolysis to gluconate, followed by gluconate re-import and phosphorylation. When grown on maltose medium, strains lacking Pgl activity turn blue after
10 iodine treatment. The phenotype of a *pgl* deletion strain can be complemented by expression of the *pgl* gene from *Pseudomonas putida*, although there is no detectable similarity between the two genes.

[00545] A strategy for metabolic engineering of *E. coli* for the production of riboflavin included overexpression of *pgl*, leading to an increase in riboflavin titer.

15 [00546] *pgl* is part of a genomic region that is deleted in the *E. coli* B strain BL21, but is present in the K-12 strain MG1655.

[00547] In some embodiments, a recombinant microorganism producing MEG (or glycolic acid), or optionally, MEG (or glycolic acid) and one or more co-product, comprises a deletion, insertion, or loss of function mutation in a gene encoding a 6-
20 phosphogluconolactonase to prevent the flux of glucose-6-phosphate through the oxidative branch of the pentose phosphate pathway and instead shunt glucose-6-phosphate through the non-oxidative branch of the pentose phosphate pathway to produce one or more pentose-5-phosphate intermediate.

25 **6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)**

[00548] The present disclosure describes enzymes that can catalyze the following reaction:

[00549] D-gluconate 6-phosphate + NADP⁺ → D-ribulose 5-phosphate + CO₂ + NADPH

[00550] 6-phosphogluconate dehydrogenase may also be known as phosphogluconate
30 dehydrogenase (NADP⁺-dependent, decarboxylating); phosphogluconic acid dehydrogenase; 6-phosphogluconic dehydrogenase; 6-phosphogluconic carboxylase; 6-phospho-D-gluconate dehydrogenase; glyceraldehyde 3-phosphate dehydrogenase.

[00551] 6-phosphogluconate dehydrogenase is an enzyme of the oxidative branch of the pentose phosphate pathway.

[00552] Three crystal structures of the enzyme in complex with substrate and cosubstrate compounds have been solved. Binding of NADP⁺ may induce a conformational change in the enzyme. A catalytic mechanism has been proposed.

[00553] *gnd* is a highly polymorphic gene within *E. coli* populations, likely due to interstrain transfer and recombination. This may be a result of its proximity to the *rfb* region, which determines O antigen structure.

[00554] Expression of 6-phosphogluconate dehydrogenase is growth rate-regulated. Most of the growth rate-dependent increase in Gnd levels is due to increased transcription, leading to higher mRNA levels. Posttranscriptional regulation involves a secondary structure element between codons 67 and 78 of the *gnd* mRNA. This region may function by sequestration of the translation initiation region into an mRNA secondary structure, thus reducing the efficiency of translation initiation. However, the effector of this regulatory mechanism has apparently not yet been identified. *truA* (*hisT*) mutants reduce the growth rate-dependent increase of Gnd expression by post-transcriptional regulation. Growth under acidic conditions upregulates expression of *gnd*. *gnd* is one of the most consistently flux-coupled genes (FCGs), which are genes whose expression transition patterns upon perturbations are correlated with their corresponding flux values.

[00555] Certain growth conditions selected for a deletion mutation in the promoter region that results in increased transcription of *gnd* and increased enzyme activity. An *edd gnd* double mutant is unable to grow on gluconate. A null mutation in *gnd* does not significantly alter the growth rate. However, cellular metabolism and metabolic flux is changed; succinate production is increased during growth on glucose or glycerol. A *gnd* deletion mutant shows enhanced ethanol and H₂ production compared to wild type during anaerobic growth on glycerol, whereas in a different, heavily engineered strain, overexpression of *gnd* increases ethanol and H₂ production.

[00556] In some embodiments, a recombinant microorganism producing MEG (or glycolic acid), or optionally, MEG (or glycolic acid) and one or more co-product, comprises a deletion, insertion, or loss of function mutation in a gene encoding a 6-phosphogluconate dehydrogenase to prevent the flux of glucose-6-phosphate through the oxidative branch of the pentose phosphate pathway and instead shunt glucose-6-phosphate through the non-oxidative branch of the pentose phosphate pathway to produce one or more pentose-5-phosphate intermediate.

Glyceraldehyde 3-phosphate dehydrogenase, phosphorylating (EC 1.2.1.12)

[00557] The present disclosure describes enzymes that can catalyze the following reaction:

5 **[00558]** D-glyceraldehyde 3-phosphate + NAD⁺ + phosphate ↔ 1,3-bisphospho-D-glycerate + NADH + H⁺

[00559] Glyceraldehyde 3-phosphate dehydrogenase may also be known as glyceraldehyde-3-phosphate dehydrogenase (phosphorylating); triosephosphate dehydrogenase; dehydrogenase, glyceraldehyde phosphate; phosphoglyceraldehyde dehydrogenase; 3-phosphoglyceraldehyde dehydrogenase; NAD⁺-dependent glyceraldehyde phosphate dehydrogenase; glyceraldehyde phosphate dehydrogenase (NAD⁺); glyceraldehyde-3-phosphate dehydrogenase (NAD⁺); NADH-glyceraldehyde phosphate dehydrogenase; glyceraldehyde-3-P-dehydrogenase.

15 **[00560]** Glyceraldehyde 3-phosphate dehydrogenase A catalyzes the reversible oxidative phosphorylation of D-glyceraldehyde-3-phosphate to 1,3-bisphospho-D-glycerate in the presence of NAD⁺ and phosphate during glycolysis and gluconeogenesis in *E. coli*. The enzyme is also found in many other organisms and its properties have been extensively studied.

20 **[00561]** *E. coli* is unusual in having two glyceraldehyde-3-phosphate dehydrogenase (GAPDH) activities encoded by gapA and epd (gapB). However, the gapA encoded enzyme has a highly efficient phosphorylating glyceraldehyde-3-phosphate dehydrogenase activity and a low phosphorylating erythrose-4-phosphate dehydrogenase activity, whereas the epd encoded enzyme has an efficient non-phosphorylating erythrose-4-phosphate dehydrogenase activity and a very low phosphorylating glyceraldehyde-3-phosphate dehydrogenase activity.

[00562] The GapA protein has a sequence that is more similar to eukaryotic sequences than to the thermophilic bacterial enzymes, and to prokaryotic enzymes in general. The gapA product is required for glycolysis, while the epd product is not. Both enzymes may be involved in production of pyridoxal 5'-phosphate (PLP).

30 **[00563]** Early studies of gapA mutants from *E. coli* K-10 implicated its role in glycolysis and demonstrated some of its catalytic properties. A gapA mutant exhibits a growth defect and also exhibits increased aggregation and lysis phenotypes that are rescued by high-salt media.

[00564] Regulation of gapA gene expression has been studied. The regulation of the fkpA, gapA, and hslT genes is affected by evolution under conditions of chronic heat stress.

[00565] The *E.coli* sequence contains several amino acids that are conserved in all GAPDHs and are postulated to be involved in NAD⁺ binding, or the catalytic mechanism.

5 **[00566]** The crystal structure of the wild-type enzyme in the presence of NAD⁺ has been determined at 1.80 Å resolution and was similar to those of other GAPDHs. The crystal structure of a N313T mutant was also determined at 2.17 Å resolution. Several other *E. coli* GAPDH crystal structures have been reported with and without bound NAD⁺, and in the hemiacetal intermediate state.

10 **[00567]** Molecular factors responsible for the NAD⁺ cofactor stereospecificity have been studied using site-directed mutagenesis. The enzyme is a B-specific dehydrogenase that catalyzes transfer of the pro-S hydrogen and binds NAD(H) in the syn nicotinamide orientation. Refolding of denatured *E. coli* GAPDH in the presence of chaperone protein Tig; trigger factor has been studied.

15 **[00568]** ADP-ribosylated GAPDH is a secreted virulence factor in some fungi and Gram-positive pathogens, as well as in pathogenic strains of *E. coli*. Non-pathogenic *E. coli* do not secrete GAPDH. Evidence suggests that *E. coli* GAPDH is also involved in DNA repair.

[00569] A series of vectors inducibly expressing paired-terminus antisense RNAs was constructed to silence central carbon metabolism in host *E. coli* K-12 MG1655. A vector that
20 silenced gapA at 93% efficacy caused severe growth inhibition. Regulating the expression of an engineered *E. coli* gapA through changes in temperature has been demonstrated to control glycolysis.

[00570] In some embodiments, a recombinant microorganism producing MEG (or glycolic acid), or optionally, MEG (or glycolic acid) and one or more co-product, comprises a
25 deletion, insertion, or loss of function mutation in a gene encoding a glyceraldehyde 3-phosphate dehydrogenase to prevent the conversion of glyceraldehyde 3-phosphate to 1,3-bisphospho-D-glycerate and instead allow D-glyceraldehyde 3-phosphate to be converted to D-xylulose-5-phosphate (with a concurrent conversion of fructose-6-phosphate to D-erythrose-4-phosphate) by a transketolase, and thus produce a pentose-5-phosphate
30 intermediate needed for the production of MEG (or glycolic acid), or optionally, MEG (or glycolic acid) and one or more co-product and provide more D-erythrose-4-phosphate for the non-oxidative branch of the pentose phosphate pathway to further produce one or more pentose-5-phosphate intermediate.

6-phosphofructokinase (EC 2.7.1.11)

[00571] Phosphofructokinase (Pfk) catalyzes the phosphorylation of fructose-6-phosphate on the C1 carbon during glycolysis. *E. coli* contains two Pfk isozymes, Pfk-1 (pfkA) and Pfk-2 (pfkB), which do not share sequence similarity. More than 90% of the phosphofructokinase activity present in wild type *E. coli* can be attributed to Pfk-1.

[00572] PfkA catalyzes the phosphorylation of fructose-6-phosphate and is a key enzyme regulating the glycolysis pathway. The enzyme cannot catalyze the reverse reaction *in vivo*. The enzyme shows cooperative kinetics with the substrate fructose-6-phosphate, but not with the other substrate, ATP. Recently, it was shown that PfkA also catalyzes phosphorylation of sedoheptulose-7-phosphate as part of the sedoheptulose bisphosphate bypass. Crystal structures of PfkA have been solved with and without activators and inhibitors. Based on sequence similarity, PfkA was predicted to be an NAD⁺ kinase.

[00573] PfkB is a member of the ribokinase family of sugar kinases. PfkB, unlike PfkA, does not show cooperative interaction with fructose-6-phosphate, inhibition by PEP or activation by ADP. MgATP²⁻ is the true substrate of the enzyme. PfkB can also use tagatose-6-phosphate as a substrate. This reaction is part of the galactitol catabolism pathway. A crystal structure of PfkB in the tetrameric form inhibited by MgATP has been solved at 1.98 Å resolution. Comparison of this structure with a crystal structure of PfkB in complex with fructose-6-phosphate suggests negative interplay between fructose-6-phosphate binding and MgATP binding.

[00574] In some embodiments, a recombinant microorganism producing MEG (or glycolic acid), or optionally, MEG (or glycolic acid) and one or more co-product, comprises a deletion, insertion, or loss of function mutation in a gene encoding a 6-phosphofructokinase to prevent the conversion of fructose-6-phosphate to 1,6-bisphosphate and instead allow fructose-6-phosphate to be converted to erythrose-4-phosphate and acetyl-phosphate by a fructose-6-phosphate phosphoketolase, and provide more erythrose-4-phosphate for the non-oxidative branch of the pentose phosphate pathway to further produce one or more pentose-5-phosphate intermediate needed for the production of MEG (or GA), or optionally, MEG (or GA) and one or more co-product. In some embodiments, the 6-phosphofructokinase is pfkA and/or pfkB.

Hydroxypyruvate decarboxylase, 2-oxoglutarate decarboxylase, 2-keto acid decarboxylase (EC 4.1.1.-)

[00575] The present disclosure describes enzymes that can catalyze the following reactions:

[00576] hydroxypyruvate + H⁺ ↔ CO₂ + glycolaldehyde

[00577] 2-oxoglutarate + coenzyme A + NAD⁺ → succinyl-CoA + CO₂ + NADH

5 **[00578]** 4-methyl-2-oxopentanoate + H⁺ → 3-methylbutanal + CO₂

[00579] 3-methyl-2-oxobutanoate + H⁺ → isobutanal + CO₂

[00580] Hydroxypyruvate decarboxylase may also be known as hydroxypyruvate carboxy-lyase.

10 **[00581]** 2-oxoglutarate decarboxylase may also be known as oxoglutarate decarboxylase; alpha-ketoglutarate decarboxylase; alpha-ketoglutaric decarboxylase; pre-2-oxoglutarate decarboxylase; 2-oxoglutarate carboxy-lyase.

[00582] *E. coli* SucA is responsible for the 2-oxoglutarate decarboxylase activity of the 2-oxoglutarate dehydrogenase multienzyme complex (OGDHC) that catalyzes the conversion of 2-oxoglutarate (2-ketoglutarate) to succinyl-CoA and CO₂, with the production of NADH.

15 **[00583]** The OGDHC is a member of the 2-oxo acid dehydrogenase family. Members of this family contain multiple copies of three enzymatic components: 2-oxoglutarate decarboxylase (E1), lipoamide acyltransferase (E2) and lipoamide dehydrogenase (E3). In most Gram-positive bacteria and in mitochondria the E1 component is a heterodimer composed of two subunits, while in most (but not all) Gram-negative bacteria it is made up of
20 a single type of subunit. In both cases multiple copies of the E1 component along with multiple copies of the E3 component are assembled around an E2 core of 24 subunits with octahedral symmetry, or 60 subunits with icosahedral symmetry (depending on which complex and species). In *E. coli* the E3 component is shared with the pyruvate dehydrogenase and glycine cleavage multi-enzyme complexes. E1 and E2 differ slightly for
25 the 2-oxoglutarate and pyruvate dehydrogenase complexes, and are designated (o) and (p) to distinguish them.

[00584] The *E. coli* OGDHC contains 12 units of the E1(o) component 2-oxoglutarate decarboxylase, thiamine-requiring encoded by *sucA*, 24 units of the E2(o) component dihydrolipoyltranssuccinylase encoded by *sucB*, and 2 units of the E3 component lipoamide
30 dehydrogenase encoded by *lpd*. The 24 E2(o) units form the octahedral core of the complex. They contain lipoyllysine and binding sites for dimers of the E1(o) and E3 subunits. Electron cryotomography showed that they are flexibly tethered to the E2 core.

[00585] During the OGDHC reaction cycle, 2-oxoglutarate is bound and decarboxylated by SucA, a thiamin-diphosphate cofactor containing enzyme. The crystal structure of a truncated, apo form of SucA lacking the N-terminal 77 residues has been determined at 2.6 Å resolution. The structure of the holo form with thiamin diphosphate and Mg^{2+} was determined at 3.5 Å resolution. The truncated form retained decarboxylase activity but did not assemble with E2(o) into an OGDH complex. Data also suggested the presence of an AMP binding site. An oxygen-dependent thiamin free radical was demonstrated in the OGDHC, which was generated by a side reaction with O_2 .

[00586] Studies of engineered SucA prepared by saturation mutagenesis of His260 and His298 suggested that His260 is required for substrate recognition, but His298 could be replaced by hydrophobic residues of similar size. Data also suggested that E2(o) has a role in specificity.

[00587] The *sucA* gene was cloned and sequenced in earlier work and regulation of *sucABCD* was studied. The *sucAB* and *sucCD* genes were shown to be mutually essential, with either pair sufficient to produce succinyl-CoA, but simultaneous deletion of *sucAB* and *sucCD* was not viable.

[00588] α -ketoisovalerate decarboxylase catalyzes the decarboxylation of 3-methyl-2-oxobutanoate to isobutanal. The enzyme is highly specific for 3-methyl-2-oxobutanoate, but also shows activity with other branched-chain 2-keto acids (4-methyl-2-oxopentanoate, 22.7% relative activity; (S)-3-methyl-2-oxopentanoate, 16.7%, 2-oxo-3-phenylpropanoate, 7.1% and 4-(methylthio)-2-oxobutanoate, 5.8%.

[00589] The enzyme is a homo-tetramer, encoded by the *kivd* gene, which has been sequenced and cloned. The deduced protein sequence shares 98.6% identity (over its first 438 amino acids) with an *L. lactis* strain IL1403 protein, encoded by the *ipd* gene (this gene is interrupted at position L439 by the insertion of an IS983 element). The *kivd* gene does not have any homology with any gene(s) in the sequenced genomes of *L. lactis* strains MG1363 and SK11.

[00590] A study of the Kivd activity testing 156 lactic acid bacteria strains (*Lactococcus*, *Lactobacillus*, *Leuconostoc*) indicated that only *L. lactis* strains possess the activity, and even within lactococcal strains, only 7 out of 45 strains had the activity.

[00591] A homologous protein has been described from *L. lactis* strain B1157 as branched-chain α -keto acid decarboxylase. That protein shows 89.8% identity with Kivd and also has a preference for 2-keto-isovalerate.

[00592] In some embodiments, an enzyme having 2-keto acid decarboxylase activity, an enzyme having hydroxypyruvate decarboxylase activity or an enzyme having 2-oxoglutarate decarboxylase activity converts hydroxypyruvate to glycolaldehyde. In some embodiments, the enzyme that converts hydroxypyruvate to glycolaldehyde is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to Kivd or SucA. In some embodiments, the enzyme having 2-keto acid decarboxylase activity is Kivd. In some embodiments, the enzyme having 2-oxoglutarate decarboxylase activity is SucA.

[00593] In some embodiments, the one or more nucleic acid molecules encoding the enzyme having 2-oxoglutarate decarboxylase activity is sucA, or homolog thereof. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having 2-keto acid decarboxylase activity is Kivd, or homolog thereof. In another embodiment, the one or more nucleic acid molecules encoding an enzyme having 2-keto acid decarboxylase activity, an enzyme having hydroxypyruvate decarboxylase activity or an enzyme having 2-oxoglutarate decarboxylase activity comprise an amino acid sequence selected from SEQ ID NOs: 224 and 226. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having 2-keto acid decarboxylase activity, an enzyme having hydroxypyruvate decarboxylase activity or an enzyme having 2-oxoglutarate decarboxylase activity is encoded by a nucleic acid sequence selected from SEQ ID NOs: 223 and 225.

2-oxoglutarate reductase, 3-phospho-hydroxypyruvate reductase, 3-phosphoglycerate dehydrogenase (EC 1.1.1.-)

[00594] The present disclosure describes enzymes that can catalyze the following reactions:

[00595] (S)-2-hydroxyglutarate + NAD⁺ ↔ 2-oxoglutarate + NADH + H⁺

[00596] 3-phospho-D-glycerate + NAD⁺ ↔ 3-phospho-hydroxypyruvate + NADH + H⁺

[00597] (R)-2-hydroxyglutarate + NAD⁺ ↔ 2-oxoglutarate + NADH + H⁺

[00598] 3-phosphoglycerate dehydrogenase may also be known as phosphoglycerate dehydrogenase; PHGDH (gene name); D-3-phosphoglycerate:NAD⁺ oxidoreductase; alpha-phosphoglycerate dehydrogenase; 3-phosphoglyceric acid dehydrogenase; D-3-phosphoglycerate dehydrogenase; glycerate 3-phosphate dehydrogenase; glycerate-1,3-phosphate dehydrogenase; phosphoglycerate oxidoreductase; phosphoglyceric acid

dehydrogenase; SerA; 3-phosphoglycerate:NAD⁺ 2-oxidoreductase; SerA 3PG dehydrogenase; 3PHP reductase.

[00599] 3-phosphoglycerate dehydrogenase catalyzes the first committed step in the biosynthesis of L-serine. The enzyme is regulated by allosteric end-product inhibition that shows cooperativity. Inhibition by serine acts primarily through reduction of catalytic velocity and has only a small effect on the K_m s of the substrates; SerA is thus classified as a type V allosteric enzyme.

[00600] The basis for allosteric and cooperative inhibition by serine has been studied extensively. Occupation of two of the four serine binding sites in the homotetramer results in 85% inhibition of activity. Further binding of serine shows negative cooperativity. Phosphate is able to reduce the site-to-site cooperative effects on serine binding; the effect was mainly due to the presence of intrinsically bound NADH. A Trp139Gly mutation results in a homodimeric enzyme that has lost cooperativity in serine binding and allosteric inhibition. Site-directed mutagenesis of residues within the effector binding site, the regulatory interface between subunits, and a flexible hinge region support a model where movement of adjacent domains is involved in inhibition of the enzymatic activity. Transient kinetic analysis showed that the cooperativity of inhibition of catalytic activity results from a conformational change due to serine binding. An enzyme missing the regulatory domain is no longer inhibited by serine, but other kinetic parameters remain the same. Hybrid tetramers provided further insight into the mechanism of allosteric inhibition.

[00601] Site-directed mutagenesis has allowed the identification of residues within the active site that contribute to substrate binding and catalysis. Mutations in the hinge region between the substrate and nucleotide binding domains affect the k_{cat} of the enzyme; certain mutations uncouple serine binding and catalytic inhibition.

[00602] Extensive site-directed mutagenesis and structural studies have contributed to a detailed view of the interactions between allosteric regulation, cooperativity and catalytic activity. Further insight into the catalytic pathway was provided by stopped-flow kinetic analysis, indicating that the rate-limiting step in both catalytic directions is a conformational change of the enzyme. Serine binding appears to lead to the formation of a dead-end quaternary complex between the enzyme, coenzyme, substrate, and effector that eliminates the conformational change subsequent to substrate binding.

[00603] The enzyme has been shown to also have an α -ketoglutarate reductase activity, producing 2-hydroxyglutarate. While the metabolic role of this reaction is not yet known, it is thought that it may play a role in regulating serine biosynthesis and in recycling NADH back to NAD⁺, especially during anaerobiosis.

[00604] Crystal structures of the wild type enzyme and various mutants have been solved. The structure showed that each subunit of the homotetramer consists of three distinct domains, a nucleotide binding domain, a substrate binding domain, and a regulatory/serine binding domain.

5 **[00605]** serA is essential for growth on glycerol minimal medium; the growth defect can be rescued by addition of serine.

[00606] In some embodiments, the enzyme having 3-phosphoglycerate dehydrogenase activity can be an enzyme having 3-phospho-hydroxypyruvate reductase activity or an enzyme having 2-oxoglutarate reductase activity. In some embodiments, an enzyme having
10 3-phosphoglycerate dehydrogenase activity, an enzyme having 3-phospho-hydroxypyruvate reductase activity, or an enzyme having 2-oxoglutarate reductase activity catalyzes the conversion of glycerate 3-phosphate to 3-phosphohydroxypyruvate. In some embodiments, an enzyme having 3-phosphoglycerate dehydrogenase activity, an enzyme having 3-phospho-hydroxypyruvate reductase activity, or an enzyme having 2-oxoglutarate reductase
15 activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to serA. In some embodiments, the enzyme having 3-phosphoglycerate dehydrogenase activity, the enzyme having 3-phospho-hydroxypyruvate reductase activity, or the enzyme having 2-oxoglutarate reductase activity is serA.

20 **[00607]** In some embodiments, the one or more nucleic acid molecules encoding the enzyme having 3-phosphoglycerate dehydrogenase activity is serA, or homolog thereof. In another embodiment, the one or more nucleic acid molecules encoding an enzyme having 3-phosphoglycerate dehydrogenase activity, an enzyme having 3-phospho-hydroxypyruvate reductase activity, or an enzyme having 2-oxoglutarate reductase activity comprises an
25 amino acid sequence set forth in SEQ ID NO: 228. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having 3-phosphoglycerate dehydrogenase activity, an enzyme having 3-phospho-hydroxypyruvate reductase activity, or an enzyme having 2-oxoglutarate reductase activity is encoded by a nucleic acid sequence set forth in SEQ ID NO: 227.

30

Phosphoserine aminotransferase, serine aminotransferase, L-serine transaminase (EC 2.6.1.52)

[00608] The present disclosure describes enzymes that can catalyze the following reactions:

[00609] 3-phospho-L-serine + 2-oxoglutarate $\leftrightarrow$ L-glutamate + 3-phospho-hydroxypyruvate

[00610] (3R)-3-hydroxy-2-oxo-4 phosphonooxybutanoate + L-glutamate $\leftrightarrow$ 4-phospho-hydroxy-L-threonine + 2-oxoglutarate

5 **[00611]** 2-oxoglutarate + N-succinyl-L,L-2,6-diaminopimelate $\leftrightarrow$ L-glutamate + N-succinyl-2-amino-6-ketopimelate

[00612] Phosphoserine aminotransferase may also be known as phosphoserine transaminase; PSAT; 3-phosphoserine aminotransferase; hydroxypyruvic phosphate-glutamic transaminase; L-phosphoserine aminotransferase; phosphohydroxypyruvate
10 transaminase; phosphohydroxypyruvic-glutamic transaminase; 3-O-phospho-L-serine:2-oxoglutarate aminotransferase; SerC; PdxC; 3PHP transaminase

[00613] The serC-encoded enzyme, phosphoserine/phosphohydroxythreonine aminotransferase, functions in the biosynthesis of both serine and pyridoxine, by using different substrates. Pyridoxal 5'-phosphate is a cofactor for both enzyme activities,
15 suggesting that it can act in an autocatalytic fashion, stimulating its own biosynthesis.

[00614] The redundancy and promiscuity among aminotransferase enzymes has been investigated. No activity could be observed with non-phosphorylated substrates; however, 3-hydroxypyruvate was able to be used as the substrate for an assay of SerC enzymatic activity. In addition, genetic experiments showed that SerC is a minor alanine transaminase.

20 **[00615]** The normal activities of two enzymes, ArgD and SerC, are sufficient for succinyl diaminopimelate (SDAP) and lysine biosynthesis; a third enzyme, AstC, is sufficient for SDAP biosynthesis, but alone can not fulfill the cell's requirement for lysine. Additional enzymes, including GabT and PuvE, may be able to contribute to SDAP biosynthesis. Expression of argD, astC, serC, aspC, gabT, hisC, ilvE, patA, puvE, or tyrB from a plasmid
25 **allows growth of the triple Δ argD serC astC mutant on minimal medium.**

[00616] Crystal structures of the enzyme in the unligated form and in complex with the substrate analog α -methyl-L-glutamate have been solved, and a molecular reaction mechanism was proposed.

[00617] serC is essential for growth on glycerol minimal medium; the growth defect can
30 be rescued by addition of serine and pyridoxol/pyridoxine.

[00618] In some embodiments, the enzyme having phosphoserine aminotransferase activity can be an enzyme having L-serine transaminase activity or an enzyme having serine aminotransferase activity. In some embodiments, an enzyme having phosphoserine

aminotransferase activity, an enzyme having L-serine transaminase activity or an enzyme having serine aminotransferase activity catalyzes the conversion of L-serine to hydroxypyruvate. In some embodiments, the enzyme having phosphoserine aminotransferase activity, the enzyme having L-serine transaminase activity or the enzyme having serine aminotransferase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to serC. In some embodiments, the enzyme having phosphoserine aminotransferase activity, the enzyme having L-serine transaminase activity or the enzyme having serine aminotransferase activity is serC.

[00619] In some embodiments, the one or more nucleic acid molecules encoding the enzyme having phosphoserine aminotransferase activity is serC, or homolog thereof. In another embodiment, the one or more nucleic acid molecules encoding an enzyme having phosphoserine aminotransferase activity, an enzyme having L-serine transaminase activity, or an enzyme having serine aminotransferase activity comprise an amino acid sequence set forth in SEQ ID NO: 230. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having phosphoserine aminotransferase activity, an enzyme having L-serine transaminase activity or an enzyme having serine aminotransferase activity is encoded by a nucleic acid sequence set forth in SEQ ID NO: 229.

3-phospho-hydroxypyruvate phosphatase

[00620] The present disclosure describes enzymes that can catalyze the following reaction:

[00621] 3-phospho-hydroxypyruvate + H₂O → hydroxypyruvate + phosphate

[00622] YeaB (NudL) belongs to the Nudix family of hydrolases and was predicted to have CoA pyrophosphohydrolase activity.

[00623] *yeaB (nudL)* was isolated as a multicopy suppressor of the repression of *flhDC* transcription in a *pgsA* mutant. The suppression may be due to the reduction of σ^S expression in cells that overexpress *nudL*.

[00624] *yeaB (nudL)* was also isolated as a multicopy suppressor of the PLP auxotrophy of a *pdxB* deletion strain. NudL was found to be part of a serendipitous metabolic pathway that produces an intermediate of the pyridoxal 5'-phosphate biosynthesis I pathway, 4-phospho-hydroxy-L-threonine, that lies downstream of PdxB. The pathway diverts 3-phosphohydroxypyruvate from serine biosynthesis. With a K_{cat} of 5.7×10^{-5} , NudL is an

inefficient catalyst of the conversion of 3-phosphohydroxypyruvate to hydroxypyruvate, but its activity appears to be sufficient for production of PLP.

[00625] In some embodiments, the enzyme having 3-phospho-hydroxypyruvate phosphatase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to yeaB. In some embodiments, the enzyme having 3-phospho-hydroxypyruvate phosphatase activity is yeaB.

[00626] In some embodiments, the one or more nucleic acid molecules encoding the enzyme having 3-phospho-hydroxypyruvate phosphatase activity is yeaB, or homolog thereof. In another embodiment, the one or more nucleic acid molecules encoding an enzyme having 3-phospho-hydroxypyruvate phosphatase activity comprise an amino acid sequence set forth in SEQ ID NO: 232. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having 3-phospho-hydroxypyruvate phosphatase activity is encoded by a nucleic acid sequence set forth in SEQ ID NO: 231.

Phosphoserine phosphatase (EC 3.1.3.3)

[00627] The present disclosure describes enzymes that can catalyze the following reaction:

[00628] 3-phospho-L-serine + H₂O → L-serine + phosphate

[00629] Phosphoserine phosphatase catalyzes the last step in serine biosynthesis. The enzyme belongs to the superfamily of haloacid dehalogenase (HAD)-like hydrolases. Enzymatic studies were originally performed using partially purified enzyme from *E. coli* strain W; assays of the purified enzyme were performed as part of an investigation of the HAD superfamily of enzymes.

[00630] serB is essential for growth on glycerol minimal medium; the growth defect can be rescued by addition of serine. Gph, HisB and YtjC were identified as multicopy **suppressors of the conditional ΔserB phenotype. Directed evolution experiments identified mutations that increased fitness and enzymatic activity of these suppressors.**

[00631] In some embodiments, the enzyme having phosphoserine phosphatase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to serB. In some embodiments, the enzyme having phosphoserine phosphatase activity is serB.

[00632] In some embodiments, the one or more nucleic acid molecules encoding the enzyme having phosphoserine phosphatase activity is serB, or homolog thereof. In another embodiment, the one or more nucleic acid molecules encoding an enzyme having phosphoserine phosphatase activity comprise an amino acid sequence set forth in SEQ ID NO: 234. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having phosphoserine phosphatase activity is encoded by a nucleic acid sequence set forth in SEQ ID NO: 233.

Serine-pyruvate aminotransferase (EC 2.6.1.51)

[00633] The present disclosure describes enzymes that can catalyze the following reactions:

[00634] pyruvate + L-serine $\leftrightarrow$ L-alanine + hydroxypyruvate

[00635] glyoxylate + L-alanine $\leftrightarrow$ glycine + pyruvate

[00636] Serine-pyruvate aminotransferase may also be known as alanine-glyoxylate aminotransferase.

[00637] The peroxisomal serine--pyruvate aminotransferase (AGXT1) and mitochondrially localized alanine-glyoxylate aminotransferase 2 (AGXT2) both catalyze the conversion of glyoxylate to glycine using alanine as the amino donor. Unlike AGXT2, AGXT1 cannot utilize asymmetric dimethylarginine (ADMA) as an amino donor.

[00638] Peroxisomal serine--pyruvate aminotransferase is a pyridoxal phosphate dependent liver specific enzyme composed of a homodimer. Its location in the peroxisome is crucial for proper enzyme activity. A peroxisomal targeting sequence (PTS1) at the C-terminus is required for translocation into peroxisomes.

[00639] Dysfunction or mistargeting of serine--pyruvate aminotransferase leading to absence in hepatic peroxisomes, causes glyoxylate to escape into the cytosol where it is further metabolized to oxalate and glycolate. Oxalate cannot be further metabolized in humans and leads to the formation of insoluble calcium oxalate in the kidney and urinary tract. Mutations in the AGXT1 gene leads to improper peroxisomal targeting and causes the autosomal recessive metabolic disorder, primary hyperoxaluria type 1, which results in irreversible kidney damage. One-third of primary hyperoxaluria type 1 patients have a unique protein sorting defect in which the hepatic peroxisomal enzyme is mistargeted to the mitochondria.

[00640] In some embodiments, the enzyme having serine-pyruvate aminotransferase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *Homo sapiens* AGXT1. In some embodiments, the enzyme having serine-pyruvate aminotransferase activity is *Homo sapiens* AGXT1.

[00641] In some embodiments, the one or more nucleic acid molecules encoding the enzyme having serine-pyruvate aminotransferase activity is AGXT1, or homolog thereof. In some embodiments, the one or more nucleic acid molecules encoding an enzyme having serine-pyruvate aminotransferase activity comprise an amino acid sequence set forth in SEQ ID NO: 244. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having serine-pyruvate aminotransferase activity is encoded by a nucleic acid sequence set forth in SEQ ID NO: 243.

Serine decarboxylase (EC 4.1.1.65)

[00642] The present disclosure describes enzymes that can catalyze the following reaction:

[00643] a 3-O-*sn*-phosphatidyl-L-serine + H⁺ → an L-1-phosphatidylethanolamine + CO₂

[00644] Serine decarboxylase may also be known as phosphatidylserine decarboxylase; PS decarboxylase; phosphatidyl-L-serine carboxy-lyase.

[00645] Phosphatidylserine decarboxylase is one of a small class of enzymes that use a covalently bound pyruvoyl prosthetic group. The pyruvoyl group is thought to act analogously to pyridoxal phosphate cofactor by forming a Schiff base with the amino group of the substrate and then serving as an electron sink to facilitate the decarboxylation.

[00646] Four of these enzymes, histidine decarboxylase (E.C. 4.1.1.22), phosphatidylserine decarboxylase, aspartate 1-decarboxylase, and S-adenosylmethionine decarboxylase are decarboxylases forming important biological amines. All of these enzymes are known to have the pyruvoyl prosthetic group attached via an amide linkage to the amino terminus of the α subunit. Two other enzymes in this group are D-proline reductase and glycine reductase (E.C. 1.2.1.4.2).

[00647] Pyruvoyl-containing enzymes are expressed as a zymogen which is processed post-translationally by a self-maturation cleavage called serinolysis. In this process the pyruvoyl group is formed from a serine residue, splitting the precursor protein into two parts which become the α and β subunits. In some cases additional subunits may be involved.

[00648] This enzyme differs from other pyruvoyl-dependent decarboxylases composed of nonidentical subunits in that the pyruvate prosthetic group is associated with the smaller subunit. The enzyme is a multimer of unknown number of the heterodimer.

[00649] In some embodiments, an enzyme having serine decarboxylase activity catalyzes the conversion of L-serine to ethanolamine.

[00650] In some embodiments, the enzyme having serine decarboxylase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *Arabidopsis thaliana* SDC. In some embodiments, the enzyme having serine decarboxylase activity is *Arabidopsis thaliana* SDC.

[00651] In some embodiments, the one or more nucleic acid molecules encoding the enzyme having serine decarboxylase activity is SDC, or homolog thereof. In some embodiments, the one or more nucleic acid molecules encoding an enzyme having serine decarboxylase activity comprise an amino acid sequence set forth in SEQ ID NO: 236. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having serine decarboxylase activity is encoded by a nucleic acid sequence set forth in SEQ ID NO: 235.

Ethanolamine oxidoreductase (deaminating) (EC 1.4.3.8), Ethanolamine aminotransferase (EC 2.6.1.-)

[00652] The present disclosure describes enzymes that can catalyze the following reactions:

[00653] ethanolamine + oxygen + H₂O → ammonium + hydrogen peroxide + glycolaldehyde

[00654] ethanolamine + 2-oxoglutarate → glycolaldehyde + L-glutamate

[00655] Ethanolamine oxidoreductase (deaminating) may also be known as ethanolamine oxidase. This enzyme belongs to the family of oxidoreductases, specifically those acting on the CH-NH₂ group of donors with oxygen as acceptor.

[00656] In some embodiments, an ethanolamine oxidase or an ethanolamine aminotransferase catalyzes the conversion of ethanolamine to glycolaldehyde.

[00657] In some embodiments, the enzyme having ethanolamine oxidase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least

80% sequence identity, or having at least 90% sequence identity to *E. coli* tynA. In some embodiments, the enzyme having ethanolamine oxidase activity is *E. coli* tynA.

[00658] In some embodiments, the one or more nucleic acid molecules encoding the enzyme having ethanolamine oxidase activity is tynA, or homolog thereof. In some
5 embodiments, the one or more nucleic acid molecules encoding an enzyme having ethanolamine oxidase activity comprise an amino acid sequence set forth in SEQ ID NO: 238. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having ethanolamine oxidase activity is encoded by a nucleic acid sequence set forth in SEQ ID NO: 237.

[00659] In some embodiments, the enzyme having ethanolamine aminotransferase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *E. coli* alaA. In some embodiments, the enzyme having ethanolamine aminotransferase activity is *E. coli* alaA.

[00660] In some embodiments, the one or more nucleic acid molecules encoding the enzyme having ethanolamine aminotransferase activity is alaA, or homolog thereof. In some
15 embodiments, the one or more nucleic acid molecules encoding an enzyme having ethanolamine aminotransferase activity comprise an amino acid sequence set forth in SEQ ID NO: 240. In a further embodiment, the one or more nucleic acid molecule encoding an
20 enzyme having ethanolamine aminotransferase activity is encoded by a nucleic acid sequence set forth in SEQ ID NOs: 239.

Hydroxypyruvate reductase (EC 1.1.1.-)

[00661] The present disclosure describes enzymes that can catalyze the following
25 reactions:

[00662] D-glycerate + NAD(P)⁺ ↔ hydroxypyruvate + NAD(P)H + H⁺

[00663] Hydroxypyruvate reductase may also be known as beta-hydroxypyruvate reductase; NADH:hydroxypyruvate reductase; D-glycerate dehydrogenase.

[00664] Hydroxypyruvate reductase is an enzyme found in higher plants, algae,
30 mammalian tissues and bacteria. In most cases it has been postulated to convert hydroxypyruvate to glycerate. However, most enzymes also carry the reduction of glyoxylate to glycolate.

[00665] In the serine cycle methylotrophs, hydroxypyruvate reductase plays a key role in the assimilation of carbon. It catalyzes the conversion of hydroxypyruvate to glycerate, a key step of the serine cycle, but it also plays an important role in the metabolism of C2 compounds, by interconverting glyoxylate and glycolate.

5 **[00666]** In some embodiments, hydroxypyruvate reductase catalyzes the conversion of glycerate to hydroxypyruvate.

[00667] In some embodiments, the enzyme having hydroxypyruvate reductase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *E. coli* ghrB. In some
10 embodiments, the enzyme having hydroxypyruvate reductase activity is *E. coli* ghrB.

[00668] In some embodiments, the one or more nucleic acid molecules encoding the hydroxypyruvate reductase is ghrB, or homolog thereof. In some embodiments, the one or more nucleic acid molecules encoding an enzyme having hydroxypyruvate reductase activity comprise an amino acid sequence set forth in SEQ ID NO: 242. In a further embodiment, the
15 one or more nucleic acid molecule encoding an enzyme having hydroxypyruvate reductase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NO: 241.

Glycerate decarboxylase

20 **[00669]** The present disclosure describes enzymes that can catalyze the following reaction:

[00670] D-glycerate + H⁺ → ethylene glycol + CO₂

[00671] In some embodiments, an enzyme having glycerate decarboxylase activity catalyzes the conversion of glycerate to ethylene glycol.

25

3-Phosphoglycerate phosphatase (EC 3.1.3.38) or 2-phosphoglycerate phosphatase (EC 3.1.3.20)

[00672] The present disclosure describes enzymes that can catalyze the following reaction:

30 **[00673]** 3-phospho-D-glycerate + H₂O → D-glycerate + phosphate

[00674] 2-phospho-D-glycerate + H₂O → D-glycerate + phosphate

[00675] 3-phosphoglycerate phosphatase may also be known as D-3-phosphoglycerate phosphatase; 3-PGA phosphatase. 2-phosphoglycerate phosphatase may also be known as D-2-phosphoglycerate phosphatase; 2-PGA phosphatase. These enzymes belong to the family of hydrolases, specifically those acting on phosphoric monoester bonds.

5 **[00676]** In some embodiments, the enzyme having 3-phosphoglycerate phosphatase activity or the enzyme having 2-phosphoglycerate phosphatase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *E. coli* phoA. In some embodiments, the enzyme having 3-phosphoglycerate phosphatase activity or enzyme having 2-
10 phosphoglycerate phosphatase activity is *E. coli* phoA.

[00677] In some embodiments, the one or more nucleic acid molecules encoding the enzyme having 3-phosphoglycerate phosphatase activity or enzyme having 2-phosphoglycerate phosphatase activity is phoA, or homolog thereof. In some embodiments, the one or more nucleic acid molecules encoding an enzyme having 3-phosphoglycerate
15 phosphatase activity or enzyme having 2-phosphoglycerate phosphatase activity comprise an amino acid sequence set forth in SEQ ID NO: 246. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having 3-phosphoglycerate phosphatase activity or enzyme having 2-phosphoglycerate phosphatase activity is encoded by a nucleic acid sequence set forth in SEQ ID NO: 245.

20

Glycerate kinase (EC 2.7.1.31)

[00678] The present disclosure describes enzymes that can catalyze the following reaction:

[00679] D-glycerate + ATP $\leftrightarrow$ 3-phospho-D-glycerate + ADP + H⁺

25 **[00680]** D-glycerate + ATP $\leftrightarrow$ 2-phospho-D-glycerate + ADP + H⁺

[00681] Glycerate kinase may also be known as glycerate 3-kinase; glycerate kinase (phosphorylating) (ambiguous); D-glycerate 3-kinase; D-glycerate kinase (ambiguous); glycerate-kinase (ambiguous); GK (ambiguous); D-glyceric acid kinase (ambiguous); ATP:(R)-glycerate 3-phosphotransferase.

30 **[00682]** This enzyme belongs to the family of transferases, specifically those transferring phosphorus-containing groups (phosphotransferases) with an alcohol group as acceptor. This enzyme participates in 3 metabolic pathways: serine/glycine/threonine metabolism, glycerolipid metabolism, and glyoxylate-dicarboxylate metabolism.

[00683] In some embodiments, an enzyme having glycerate kinase activity catalyzes the conversion of 3-phosphoglycerate to glycerate. In other embodiments, an enzyme having glycerate kinase activity catalyzes the conversion of 2-phosphoglycerate to glycerate.

[00684] In some embodiments, the enzyme having glycerate kinase activity is a glycerate 3-kinase. In some embodiments, the enzyme having glycerate 3-kinase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *Arabidopsis thaliana* GLYK. In some embodiments, the glycerate 3-kinase is *Arabidopsis thaliana* GLYK.

[00685] In some embodiments, the one or more nucleic acid molecules encoding the enzyme having glycerate 3-kinase activity is GLYK, or homolog thereof. In some embodiments, the one or more nucleic acid molecules encoding an enzyme having glycerate 3-kinase activity comprise an amino acid sequence set forth in SEQ ID NO: 248. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having glycerate 3-kinase activity is encoded by a nucleic acid sequence set forth in SEQ ID NO: 247.

[00686] In some embodiments, the enzyme having glycerate kinase activity is an enzyme having glycerate 2-kinase activity. In some embodiments, the enzyme having glycerate 2-kinase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *E. coli* glxK. In some embodiments, the enzyme having glycerate 2-kinase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *E. coli* garK. In other embodiments, the enzyme having glycerate 2-kinase activity is *E. coli* glxK. In some embodiments, the glycerate 2-kinase is *E. coli* garK.

[00687] In some embodiments, the one or more nucleic acid molecules encoding the enzyme having glycerate 2-kinase activity is glxK, or homolog thereof. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having glycerate 2-kinase activity is garK, or homolog thereof. In some embodiments, the one or more nucleic acid molecules encoding an enzyme having glycerate 2-kinase activity comprise an amino acid sequence selected from SEQ ID NOs: 250 and 252. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having glycerate 2-kinase activity is encoded by a nucleic acid sequence selected from SEQ ID NOs: 249 and 251.

Transferase that transfers one-carbon group (EC 2.1.2.-)

[00688] The present disclosure describes enzymes that can catalyze the following reaction:

[00689] $M\text{-THF} + H_2O \leftrightarrow THF + \text{formaldehyde}$

[00690] Transferases such as the hydroxymethyl-, formyl- and related transferases may be used. Examples of hydroxymethyl-, formyl- and related transferases include glycine hydroxymethyltransferase, phosphoribosylglycinamide formyltransferase, phosphoribosylaminoimidazolecarboxamide formyltransferase, glycine formimidoyltransferase, glutamate formiminotransferase, D-alanine 2-hydroxymethyltransferase, deoxycytidylate 5-hydroxymethyltransferase, methionyl-tRNA formyltransferase, aminomethyltransferase, 3-methyl-2-oxobutanoate hydroxymethyltransferase and UDP-4-amino-4-deoxy-L-arabinose formyltransferase.

Serine hydroxymethyltransferase (EC 2.1.2.1)

[00691] The present disclosure describes enzymes that can catalyze the following reaction:

[00692] $L\text{-serine} + \text{tetrahydrofolate (THF)} \leftrightarrow \text{Glycine} + 5,10\text{-methylenetetrahydrofolate (M-THF)}$

[00693] Serine hydroxymethyltransferase (GlyA) converts serine to glycine, transferring a methyl group to tetrahydrofolate, thus forming 5,10-methylene-tetrahydrofolate (M-THF). M-THF is the major source of C1 units in the cell, making GlyA a key enzyme in the biosynthesis of purines, thymidine, methionine, choline and lipids.

[00694] The enzyme also catalyzes several side reactions including hydrolysis of 5,10-methenylTHF to 5-formylTHF and the reversible cleavage of 3-hydroxy amino acids (L-threonine, allothreonine, 3-phenylserine) to glycine and an aldehyde. D-alanine inactivates the enzyme by reacting with the pyridoxal phosphate prosthetic group to form pyridoxamine phosphate.

[00695] The Thr226 residue within a conserved region of the enzyme appears to be involved in substrate discrimination. The His228 residue plays a role in determining reaction specificity. Lys229 does not appear to play a catalytic role. Arg363 appears to be the binding site for the carboxyl group of the amino acid substrate. The hydroxyl group of Tyr65 may be involved in the conversion of the active site from a closed to an open conformation. Both Tyr55 and Arg235 are required for the transaldimination reaction.

[00696] Studies on refolding of the enzyme indicate that pyridoxal 5'-phosphate (PLP) only binds to the dimeric apoenzyme at the end of the folding pathway. The mechanism of PLP addition has been investigated further. At high concentrations of PLP, a second molecule of PLP can bind at Lys346. A conserved hydrophobic contact area is involved in stability of the PLP binding site. Tyr55 is required for correct positioning of the PLP cofactor.

[00697] Crystal structures of wild type and mutant serine hydroxymethyltransferase have been solved.

[00698] glyA mutants cannot use glycine as the sole source of nitrogen. A glyA mutant is auxotrophic for glycine; glyA was later shown to be essential for growth on glycerol minimal medium.

[00699] Sequences 3' to the structural gene within the glyA mRNA are required for mRNA stability.

[00700] In some embodiments, the enzyme having serine hydroxymethyltransferase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *E. coli* glyA. In some embodiments, the one or more nucleic acid molecule encoding an enzyme having serine hydroxymethyltransferase comprises an amino acid sequence set forth in UniProt ID P0A825. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having serine hydroxymethyltransferase is encoded by a nucleic acid sequence set forth in Gene ID 947022.

Formaldehyde dehydrogenase (EC 1.2.1.46 and EC 1.2.1.-)

[00701] The present disclosure describes enzymes that can catalyze the following reaction:

[00702] Formaldehyde + NAD⁺ + H₂O → formate + NADH + 2 H⁺

[00703] Formaldehyde dehydrogenase may also be known as NAD-linked formaldehyde dehydrogenase, NAD-dependent formaldehyde dehydrogenase or formaldehyde:NAD⁺ oxidoreductase.

[00704] Most of the formaldehyde dehydrogenases found in animals, plants and bacteria belong to a group called class III alcohol dehydrogenase group and require the addition of glutathione for activity. As a matter of fact, the true substrate for these enzymes was shown to be not formaldehyde, but S-hydroxymethylglutathione, which is formed nonenzymatically from formaldehyde and glutathione.

[00705] Unlike those enzymes, the enzyme isolated from *Pseudomonas putida* catalyzes the irreversible oxidation of formaldehyde to formate without the addition of glutathione. Since its substrate is formaldehyde, in an essence this is the "genuine" formaldehyde dehydrogenase. Like other formaldehyde dehydrogenases, the *P. putida* FDH is a zinc-containing metalloenzyme. It also requires NAD⁺ as the electron acceptor. However, unlike the enzymes that belong to the class III alcohol dehydrogenase group, it is sensitive to 4-methylpyrazole. In one embodiment, the formaldehyde dehydrogenase is from *Pseudomonas putida*. In some embodiments, the formaldehyde dehydrogenase is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *P. putida* fdhA. In some embodiments, the one or more nucleic acid molecule encoding a formaldehyde dehydrogenase comprises an amino acid sequence set forth in GenBank Accession BAA04743.1. In a further embodiment, the one or more nucleic acid molecule encoding a formaldehyde dehydrogenase is encoded by a nucleic acid sequence set forth in GenBank Accession D21201.1.

[00706] In the industrially important actinomycete *Corynebacterium glutamicum* ATCC 13032, evidence suggests that two enzymes contribute to the degradation of toxic formaldehyde, mycothiol-dependent formaldehyde dehydrogenase encoded by gene fadH, and to a lesser extent acetaldehyde dehydrogenase encoded by gene ald (acetaldehyde dehydrogenase). A mutant lacking both of these enzymes was unable to grow in formaldehyde-containing medium. It also did not grow in vanillate-containing medium because the oxidation of vanillate produces intracellular formaldehyde. Detoxification of formaldehyde is necessary when this soil bacterium encounters formaldehyde in its habitat, or when formaldehyde is generated during metabolism of environmental compounds such as vanillate. The formate produced by FadH can be further oxidized to CO₂ by the formate dehydrogenase encoded by gene fdhF.

[00707] An ald mutant showed a reduction in formaldehyde degradation of about 30% as compared with wild-type. Inactivation of the chromosomal ald gene resulted in loss of acetaldehyde dehydrogenase activity and loss of the ability of this organism to grow on or utilize ethanol, suggesting a two step oxidation of ethanol to acetate. Expression of gene ald is dependent on the transcriptional regulator RamA, whereas RamB has a slightly negative effect on expression. In one embodiment, the formaldehyde dehydrogenase is from *Corynebacterium glutamicum* ATCC 13032. In some embodiments, the formaldehyde dehydrogenase is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *Corynebacterium glutamicum* ATCC 13032 ald. In some embodiments, the one or more

nucleic acid molecule encoding a formaldehyde dehydrogenase comprises an amino acid sequence set forth in UniProt ID Q8NLZ0. In a further embodiment, the one or more nucleic acid molecule encoding a formaldehyde dehydrogenase is encoded by a nucleic acid sequence set forth in Gene ID 1020739.

5 **[00708]** In some embodiments, the enzyme having formaldehyde dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *Pseudomonas oleovorans* alkH.

[00709] In *Saccharomyces cerevisiae*, two tandem-repeated genes ALD2 and ALD3
10 encode two cytoplasmic stress-inducible isoforms of aldehyde dehydrogenase. The expression of these isoforms is dependent on the general-stress transcription factors Msn2 and Msn4 but is independent of the HOG MAP kinase pathway. Both forms can use the cofactor NAD⁺ much more efficiently than NADP⁺, and are not activated by any cations. While ALD3 is induced by a variety of stresses, including osmotic shock, heat shock,
15 glucose exhaustion, oxidative stress and drugs, ALD2 is only induced by osmotic stress and glucose exhaustion.

[00710] In some embodiments, the enzyme having formaldehyde dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *Saccharomyces*
20 *cerevisiae* ALD2. In other embodiments, the enzyme having formaldehyde dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *Saccharomyces cerevisiae* ALD3. In some embodiments, the one or more nucleic acid molecule encoding an enzyme having formaldehyde dehydrogenase activity comprises an
25 amino acid sequence selected from UniProt ID P47771 and UniProt ID P54114. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having formaldehyde dehydrogenase activity is encoded by a nucleic acid sequence selected from Gene ID 855206 and Gene ID 855205.

[00711] In some embodiments, the enzyme having formaldehyde dehydrogenase activity
30 is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *Homo sapiens* ALDH3A2. In other embodiments, the enzyme having formaldehyde dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *Homo sapiens*
35 ALDH9A1. In some embodiments, the one or more nucleic acid molecule encoding an

enzyme having formaldehyde dehydrogenase activity comprises an amino acid sequence selected from UniProt ID P51648 and UniProt ID P49189. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having formaldehyde dehydrogenase activity is encoded by a nucleic acid sequence selected from Gene ID 224 and Gene ID 223.

5

Formate dehydrogenase (EC 1.2.1.-)

[00712] The present disclosure describes enzymes that can catalyze the following reactions:

[00713] formate + an oxidized electron acceptor + H^+ $\rightarrow$ CO_2 + a reduced electron acceptor

10

[00714] formate + H^+ $\rightarrow$ CO_2 + H_2 (catalyzed by complex)

[00715] formate + an oxidized hydrogenase 3 $\rightarrow$ CO_2 + a reduced hydrogenase 3

15

[00716] Formate dehydrogenase-H is one of three membrane-associated formate dehydrogenase isoenzymes in *E. coli*. All are functional in the anaerobic metabolism of the organism.

20

[00717] Formate dehydrogenase-H (FDH-H) is located in the cytoplasm and together with hydrogenase-3, FDH-H forms the formate-hydrogen lyase complex. The enzyme is oxygen sensitive and contains selenium as selenocysteine incorporated cotranslationally at the position of an in-frame UGA stop codon in the FdhF open reading frame. A crystal structure of FDH-H has been solved at 2.3 Å resolution, confirming the presence of a [4Fe-4S] cluster, coordination of the Mo cofactor by selenocysteine, and the position of the binding site for the inhibitor nitrate. Expression of fdhF is induced by formate and the absence of external electron acceptors, and is repressed by nitrate, nitrite, trimethylamine N-oxide, and oxygen. Formate can overcome repression by nitrate but not by oxygen. Inhibition of DNA gyrase enhances expression of fdhF.

25

[00718] fdnGHI encodes membrane bound formate dehydrogenase N (FDH-N) - a respiratory enzyme that catalyses the oxidation of formate to carbon dioxide, donating electrons to the quinone pool for the reduction of anaerobic respiratory substrates such as nitrate and trimethylamine N-oxide. FDH-N is a member of the complex iron sulfur molybdoenzyme (CISM) family. The oxidation of formate by FDH-N is electrogenic ($H^+/e^- = 1$); oxidation of formate in the periplasm is accompanied by menaquinone reduction at the cytoplasmic face of the inner membrane. Expression of formate dehydrogenase-N is induced by nitrate and anaerobiosis, mediated by NarL and Fnr, respectively. Purified FDH-N

30

contains three subunits, designated α (FdnG), β (FdnH) and γ (FdnI). A crystal structure, resolved at 1.6 Å, indicates that this subcomplex is further organised into physiologically relevant trimers with the α and β subunits located towards the periplasmic face of the inner membrane and the γ subunits located towards the cytoplasm. Electrons are transferred from the site of formate oxidation in the α subunit across the membrane to the site of menaquinone reduction in the γ subunit. Protons are taken up from the cytoplasm at the menaquinone reduction site.

[00719] *fdoGHI* encodes formate dehydrogenase O (FDH-O) - a respiratory molybdoenzyme that catalyses the oxidation of formate to carbon dioxide, donating electrons to the membrane soluble quinone pool for the reduction of nitrate. FDH-O and nitrate reductase Z participate in a formate to nitrate electron transport pathway that is active when cells are shifted from aerobic to anaerobic conditions. The pathway operates with either menaquinone or ubiquinone. FDH-O appears to be expressed constitutively; unlike formate dehydrogenase N (FDH-N), it is not regulated by Fnr or NarL. Expression of FDH-O is increased under aerobic conditions; under anaerobic conditions, nitrate stimulates expression slightly; the global regulators H-NS and CRP may play a role in regulation of FDH-O expression. FDH-O may contribute to the cells ability to rapidly adapt to anaerobiosis while levels of FDH-N are still insufficient. FDH-O is a heterotrimeric complex consisting of an α (FdoG), a β (FdoH) and a γ (FdoI) subunit - it shares extensive sequence similarity and immunological properties with the anaerobically expressed FDH-N.

[00720] *Candida boidinii* formate dehydrogenase FDH1 is an NAD-dependent enzyme that mediates the detoxification of formate and is strongly inhibited by Cu^{2+} , Hg, p-chloromercuribenzoate, cyanide, azide, thiocyanate and cyanate. The inhibition of cyanide is reversible and competes with formate. Protein expression is induced by methanol and repressed by glucose. Since the enzymatic reaction catalyzed by this formate dehydrogenase can regenerate NADH, it has been cloned into *E.coli* to optimize NADH requiring engineered pathways.

[00721] In the industrially important actinomycete *Corynebacterium glutamicum* ATCC 13032 evidence suggests that a formate dehydrogenase catalyzes the oxidation of formate to CO_2 . Both formate and toxic formaldehyde are present in the environment and can be dissimilated by this soil bacterium via the oxidation of formaldehyde to formate. This can be accomplished via FadH and Ald. Formate is then converted to CO_2 by FdhF. FdhF is a molybdenum cofactor-dependent formate dehydrogenase that is active under oxic conditions, and was speculated to be involved in the stress response. The exact electron acceptor used by FdhF has not been defined. Gene *fdhF* is part of a gene cluster containing related genes *fdhD* and *cg0617* that were shown by mutant analysis to be required for

formate dehydrogenase activity. The growth of *Corynebacterium glutamicum* ATCC 13032 is inhibited to some extent in the presence of formate and strains lacking formate dehydrogenase activity show increased inhibition. Radiotracer experiments showed that when *Corynebacterium glutamicum* ATCC 13032 was grown with glucose and ^{13}C -formate, it was metabolized to ^{13}C -carbon dioxide. An fdhF deletion mutant could not metabolize formate. Growth studies also demonstrated the requirement for Mo^{2+} . Protein sequence analysis suggested that FdhF is not an integral membrane protein, but is likely either cytosolic or membrane-associated. Putative orthologs have been identified in a variety of other soil bacteria.

[00722] When *Cupriavidus oxalaticus* is grown on formate as the main carbon and energy source, NAD $^{+}$ -dependent formate dehydrogenase is the key enzyme that generates NADH and CO_2 . The latter enters the ribulose diphosphate carboxylase reaction. The enzyme has been purified to homogeneity from cells grown with formate. The enzyme is a complex flavoprotein containing 2 FMN (flavin mononucleotide), 18-25 non-heme iron atoms and 15-20 acid-labile sulfides. The specific activity was 42 units/mg. The enzyme is specific toward its natural substrate formate, but can accept multiple nonphysiological electron acceptors, including methylviologen, phenazine methosulfate, methylene blue, nitro blue tetrazolium salt, FMN, FAD, riboflavin, and oxygen. It has been shown that the enzyme can also catalyze the reaction in the opposite direction. However, under the conditions employed the enzyme catalyzed the oxidation of formate about 30 times faster than CO_2 reduction.

[00723] NAD $^{+}$ -dependent formate dehydrogenase from *Gottschalkia acidurici* catalyzes a reversible reaction. While it catalyzes the oxidation of formate to CO_2 , it also catalyzes the reduction of the latter to formate, which is then converted to acetate. The enzyme has been partially purified and found to be a large enzyme complex (molecular weight of at least 200 kDa) that is very sensitive to oxygen and light. The enzyme contains a L-selenocysteine. Crude preparations of the enzyme could be coupled to NAD reduction during formate oxidation through ferredoxin. When the artificial electron acceptor methyl viologen was used instead of NAD, ferredoxin was not required. Cyanide inhibited the enzyme 90%. Basal formate oxidation activity in cell extracts was 0.85 $\mu\text{mol}/\text{min}/\text{mg}$ protein, but increased 12-fold upon the addition of tungstate and selenite. Interestingly, the enzyme from the related organism *Clostridium cylindrosporium*, while having a similar requirement for selenite, requires molybdate rather than tungstate, which has an antagonistic effect on it.

[00724] The tungsten-containing NAD $^{+}$ -dependent formate dehydrogenase from *Methylobacterium extorquens* is a heterodimer containing iron-sulfur clusters, FMN and tungsten. It is somewhat unusual to find a tungsten-containing enzyme in aerobic bacteria, although several other examples have been found. The smaller β subunit appears to be a

fusion protein, with its N-terminal domain related to NueE-like subunits, and its C-terminal domain related to NuoF-like subunits of known NADH-ubiquinone oxidoreductases.

[00725] Two different forms of formate dehydrogenase FDH have been purified from *Methylosinus trichosporium* OB3b independently by two groups, and the two proteins were found to have different properties. This protein is composed of two types of subunits in an apparent $\alpha_2\beta_2$ arrangement, with a total size of 315 kDa. It contains nonheme iron and sulfide, and no other metals, and appeared to require FMN.

[00726] The *Moraxella* sp. NAD⁺-dependent formate dehydrogenase fdh is a relatively simple dimeric protein, with no prosthetic groups.

[00727] In some embodiments, the enzyme having formate dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to a formate dehydrogenase selected from the group consisting of *E.coli* fdhF (chlF, FDH-H), *E. coli* FDH-N, *E. coli* FDH-O, *Candida boidinii* FDH1, *Corynebacterium glutamicum* fdhF, *Cupriavidus oxalaticus* NAD⁺-dependent formate dehydrogenase, *Gottschalkia acidurici* NAD⁺-dependent formate dehydrogenase, *Methylobacterium extorquens* Fdh1, *Methylosinus trichosporium* formate dehydrogenase, and *Moraxella* sp. NAD⁺-dependent formate dehydrogenase fdh. In some embodiments, the one or more nucleic acid molecule encoding a formate dehydrogenase or formate dehydrogenase subunit comprises an amino acid sequence selected from UniProt ID P07658, UniProt ID P0AEK7, UniProt ID P0AAJ3, UniProt ID P24183, UniProt ID P32176, UniProt ID P0AAJ5, UniProt ID P0AEL0, UniProt ID O13437, UniProt ID Q8NSY6, UniProt ID Q8KTI7, UniProt ID Q8KTI8, and UniProt ID O08375. In a further embodiment, the one or more nucleic acid molecule encoding a formate dehydrogenase or formate dehydrogenase subunit is encoded by a nucleic acid sequence selected from Gene ID 948584, Gene ID 946038, Gene ID 948794, Gene ID 946035, Gene ID 948394, Gene ID 948395, Gene ID 948383, GenBank accession AJ011046.2, Gene ID 1021531, GenBank accession AF489516, and GenBank accession Y13245.1.

Formate hydrogenlyase complex

[00728] The present disclosure describes enzymes that can catalyze the following reaction:

[00729] formate + H⁺ → CO₂ + H₂ (catalyzed by complex)

[00730] The component enzymes of formate hydrogenlyase complex are: 1) formate dehydrogenase H (also known as ChlF, FdhF, FDH-H), 2) hydrogenase 3, which has multiple subunits (hycBCDEFG gene: hycB, hycC, hycD, hycE, hycF, hycG).

[00731] Formate dehydrogenase-H (FDH-H) is described above.

5 **[00732]** Microbial hydrogenases catalyse the reversible reduction of protons to molecular hydrogen. *E. coli* hydrogenase 3, encoded by the *hyc* genes (*hycD*, *hycC*, *hycF*, *hycG*, *hycB* and *hycE*), is a multisubunit enzyme that forms part of the formate hydrogenlyase (FHL) complex responsible for the fermentative or anaerobic oxidation of formic acid to carbon dioxide and molecular hydrogen.

10 **[00733]** Hydrogenase 3 functions primarily in the production of H₂ and is important for H₂ production at acidic pH. Hydrogen uptake in a strain lacking hydrogenase 1 and hydrogenase 2 is further reduced by the incorporation of a *hycE* mutation, suggesting that hydrogenase 3 can also function in hydrogen uptake. Hydrogenase 3 shows a high tolerance to product (H₂) inhibition.

15 **[00734]** Hydrogenase 3 is a membrane associated H₂ evolving respiratory [NiFe] hydrogenase. It contains the large (HycE) and small (HycG) subunits that are characteristic of 'standard' NiFe hydrogenases plus two additional hydrophilic subunits (HycB and HycF) and two inner membrane subunits (HycC and HycD). Fe-S prosthetic groups located in the hydrophilic part of the complex may form the electron transport pathway. Isolation of FHL
20 using affinity chromatography indicates the presence of a core complex containing HycE, HycB HycF HycG and FdhF which has formate hydrogenlyase activity *in vitro*; a larger complex containing the membrane associated subunits HycC and HycD is isolated in the presence of detergent.

25 **[00735]** Formate oxidation in an anaerobically grown fermenting *E. coli* strain lacking hydrogenase 1 and hydrogenase 2 enzymes generates membrane potential.

[00736] Sequence similarity between the genes encoding hydrogenase 3 and those encoding subunits that form the core of energy conserving NADH:quinone oxidoreductase (complex I) has been reported and an evolutionary relationship between the two has been proposed.

30 **[00737]** Strains with insertion mutations of genes within the *hyc* operon are defective for hydrogenase activity.

[00738] Hydrogenase 3 is a nickel containing Fe-S protein.

[00739] The *hyc* operon is regulated coordinately with the structural gene for formate dehydrogenase H. Expression is repressed by oxygen and by nitrate and induced by formate under fermentative growth conditions. Formate is an obligate inducer of the formate hydrogenlyase complex genes.

5 **[00740]** *E. coli* K-12 contains three other hydrogenases: hydrogenase 1 and hydrogenase 2 (respiratory enzymes that function in H₂ uptake) and hydrogenase 4 (less well characterised; possibly silent).

[00741] In some embodiments, the hydrogenase 3 is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or
10 having at least 90% sequence identity to a hydrogenase 3 selected from the group consisting of *E. coli* hycB, *E. coli* hycC, *E. coli* hycD, *E. coli* hycE, *E. coli* hycF, and *E. coli* hycG. In some embodiments, the one or more nucleic acid molecule encoding a hydrogenase 3 or formate hydrogenlyase complex comprises an amino acid sequence selected from UniProt ID P0AAK1, UniProt ID P16429, UniProt ID P16430, UniProt ID
15 P16431, UniProt ID P16432, and UniProt ID 16433. In a further embodiment, the one or more nucleic acid molecule encoding a formate hydrogenlyase complex or formate hydrogenlyase subunit is encoded by a nucleic acid sequence selected from Gene ID 948002, Gene ID 945327, Gene ID 948994, Gene ID 947396, Gene ID 947048, and Gene ID 947191.

20

Glycine Cleavage System

[00742] The glycine cleavage system is composed of four proteins: three enzymes and a carrier protein. In animals, the system is loosely bound to the mitochondrial inner membrane. The enzymes are i) P-protein (a pyridoxal phosphate-containing protein) or glycine
25 dehydrogenase (decarboxylating) (EC1.4.4.2), ii) T-protein or aminomethyl-transferase (EC2.1.2.10), and iii) L-protein or dihydrolipoamide dehydrogenase (EC1.8.1.4). The carrier protein is called H-protein (a lipoic acid-containing protein).

[00743] The glycine cleavage reaction catalyzes the following reversible reaction:

[00744] Glycine + THF + NAD⁺ ↔ M-THF + CO₂ + NH₃ + NADH + H⁺

30 **[00745]** The system is partitioned into three partial reactions. The reaction is completely reversible, and in both glycine cleavage and glycine synthesis an aminomethyl moiety bound to the lipoic acid of H-protein represents an intermediate that is subsequently degraded to, or can be formed from, methylene-tetrahydrofolate (M-THF) and ammonia by the action of T-

protein. Possibly the reaction may involve a ternary complex of P-protein, aminomethyl moiety of glycine and H-protein, as a crucial intermediary state.

[00746] Reaction catalyzed by P-protein

[00747] The first partial reaction of the glycine degradation is the decarboxylation catalyzed by P-protein (a glycine decarboxylase). H-protein serves as a co-substrate. One of the most characteristic properties of the glycine cleavage reaction is that, although P-protein should belong to a class of pyridoxal phosphate-dependent amino acid decarboxylases, P-protein requires H-protein to catalyze the decarboxylation of glycine significantly. The reaction proceeds via a sequential random mechanism where the carboxyl carbon of glycine is converted to carbon dioxide. The remnant of the glycine molecule is transferred to one of the sulfhydryl groups formed by the reductive cleavage of disulfide in lipoate attached to H-protein.

[00748] P-protein, a pyridoxal phosphate-containing protein of about 200 kDa, is either a homodimer or a dimer of heterodimers. The former has one molecule of pyridoxal phosphate per subunit, and the latter has one molecule of the cofactor per dimer on the β subunit. The pyridoxal cofactor is attached to a specific lysine residue. The pyridoxal cofactor interacts with the active-site pocket noncovalently. The active site of *T. thermophilus* P-protein is connected to the molecular surface by a channel with a broad entrance facing the solvent. The molecular surface around the channel is composed of several positively-charged amino acid residues, which are possibly involved in the complex formation with H-protein.

[00749] In some embodiments, the enzyme having glycine decarboxylase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *E. coli* gcvP. In some embodiments, the one or more nucleic acid molecule encoding an enzyme having glycine decarboxylase activity comprises an amino acid sequence set forth in UniProt ID P33195. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having glycine decarboxylase activity is encoded by a nucleic acid sequence set forth in Gene ID 947394.

[00750] Reaction catalyzed by T-protein

[00751] The decarboxylated moiety of glycine attached to H-protein is subjected to further degradation catalyzed by T-protein (an aminomethyltransferase). The reaction requires THF and yields ammonia, M-THF, and H-protein with reduced lipoate. In the absence of THF, formaldehyde is produced instead of M-THF, but the reaction rate is less than 0.05% of that measured in the presence of THF. In the reverse reaction, T-protein catalyzes the formation

of the H-protein-bound aminomethyl lipoate intermediate from M-THF, ammonia, and H-protein with reduced lipoate via an ordered Ter Bi mechanism, in which H-protein is the first substrate to bind followed by M-THF and ammonia. The order of the product release is THF and the methylamine-loaded H-protein.

5 **[00752]** T-protein is a monomer of about 40 kDa and forms a 1:1 complex with H-protein. A cross-linking study employing *E. coli* proteins revealed that the interaction of H-protein with T-protein causes a conformational change of T-protein. Intermolecular contact between Lys-288 of T-protein and Asp-43 of H-protein was found. The N-terminal region of T-protein is essential for the interaction with H-protein and for holding T-protein in a compact form. The
10 crystal structure of human T-protein has been analyzed in a free form and that bound to N5-methyl-tetrahydrofolate, an analogue of M-THF. The overall structure consists of three cloverleaf-like structure with the central cavity where the THF cofactor is bound with the pteridin ring deeply buried into the hydrophobic pocket and the glutamyl group pointed to the C-terminal side surface. The structure resembles those of bacterial T-protein from
15 *Thermotoga naritima*, *E. coli*, and *Pyrococcus horikoshii* OT3. Structural and mutational analyses of human T-protein indicated that the invariant Asp-101 might play a key role in the initiation of the catalysis by increasing the nucleophilic character of the N10 atom of the folate substrate.

[00753] Residues involved in binding of folate have been identified by crosslinking and
20 site-directed mutagenesis. The N-terminal region of GvcT is important for the proper conformation of GvcT, allowing interaction with the H-protein.

[00754] In some embodiments, the enzyme having aminomethyltransferase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *E. coli* gcvT. In some
25 embodiments, the one or more nucleic acid molecule encoding an enzyme having aminomethyltransferase activity comprises an amino acid sequence set forth in UniProt ID P27248. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having aminomethyltransferase activity is encoded by a nucleic acid sequence set forth in Gene ID 947390.

30 **[00755]** Reaction catalyzed by L-protein

[00756] The last step of the glycine cleavage reaction is the reoxidation of the reduced lipoate attached to H-protein catalyzed by L-protein. L-protein is well known as dihydrolipoamide dehydrogenase, lipoamide dehydrogenase, dihydrolipoyl dehydrogenase, or E3 protein component of 2-oxoacid (pyruvate, 2-oxoglutarate, and branched-chain 2-

oxoacid) dehydrogenase multienzyme complexes. It catalyzes the transfer of electrons to the ultimate acceptor, NAD.

[00757] Experiments employing pea L-protein and H-protein showed that the oxidation of dihydrolipoyl H-protein was not affected by the presence of structurally related analogues such as apoH-protein or octanoylated H-protein. The structural interaction between L-protein and H-protein may not be essential for the oxidation reaction.

[00758] Kinetics of the reaction have been studied and suggest a modified ping-pong mechanism. Site-directed mutagenesis was used to identify and characterize the redox-active disulfide and a charged residue influencing the redox potential of the FAD cofactor.

The insertion of the FAD cofactor is essential for dimerization and full activity.

[00759] An *lpd* null mutant produces more pyruvate and L-glutamate under aerobic conditions. Metabolic flux analysis shows that the Entner-Doudoroff pathway I and the glyoxylate shunt are activated. Another dihydrolipoate dehydrogenase activity has been detected in *E. coli* *lpd* mutants; thus, an isozyme may exist.

[00760] A mutation in the *lpd* gene in *E. coli* causes the pyruvate dehydrogenase complex to be less sensitive to NADH inhibition and active during anaerobic growth. Amino acid substitutions at Glu354 that lowered the sensitivity of the enzyme to NADH inhibition were proposed to act by restricting the movement of NADH.

[00761] Suppressor mutations in *lpd* have been shown to restore growth to a redox-defective mutant that lacks both the thioredoxin and glutathione/glutaredoxin reduction pathways. The suppressor mutations reduced *Lpd* activity resulting in dihydrolipoamide accumulation, which could then serve as an electron donor via reduction of glutaredoxins. The reoxidation of *Lpd* restored TCA cycle function.

[00762] *lpd* shows differential codon adaptation, resulting in differential translation efficiency signatures, in aerotolerant compared to obligate anaerobic microbes. It was therefore predicted to play a role in the oxidative stress response. An *lpd* deletion mutant was shown to be more sensitive than wild-type specifically to hydrogen peroxide exposure, but not other stresses.

[00763] In some embodiments, the enzyme having dihydrolipoamide dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *E. coli* *lpd* (*lpdA*, E3 subunit). In some embodiments, the one or more nucleic acid molecule encoding an enzyme having dihydrolipoamide dehydrogenase activity comprises an amino acid sequence set forth in UniProt ID P0A9P0. In a further embodiment, the one or more nucleic

acid molecule encoding an enzyme having dihydrolipoamide dehydrogenase activity is encoded by a nucleic acid sequence set forth in Gene ID 944854.

[00764] H-protein

[00765] H-protein is a monomeric and heat-stable protein of about 14 kDa. Vertebrate H-protein is composed of 125 amino acid residues, and lipoic acid is covalently linked to Lys-59. The X-ray crystal structure of the lipoylated pea leaf H-protein (131 residues) revealed that the lipoyl-lysine was localized on the surface of the protein. As mentioned above, the lipoyllysine arm on H-protein shuttles the reaction intermediate and reducing equivalents between the active sites of the components of the glycine cleavage system. The mechanism is analogous to that found in 2-oxoacid dehydrogenase complexes.

[00766] Lipoylation of H-protein as well as acyltransferase (E2) components of 2-oxoacid dehydrogenase complexes is catalyzed by lipoate-protein ligase A (LplA) in *E. coli*. The enzyme catalyzes both the formation of lipoyl-AMP from lipoate and ATP and the transfer of the lipoyl-moiety of lipoyl-AMP to H-protein and E2 components. The X-ray crystallographic study showed that LplA consists of a large N-terminal domain and a small C-terminal domain with a substrate-binding pocket at the interface between the two domains.

[00767] In mammals, lipoylation is an intramitochondrial event. Lipoic acid is first activated to lipoyl-GMP by lipoate activating enzyme, employing GTP as a high-energy compound. Lipoate activating enzyme is the same protein already known as xenobiotic-metabolizing medium-chain fatty acid: CoA ligase-III. Lipoate is then transferred from lipoyl-GMP to apoproteins by the action of lipoyltransferase.

[00768] The H-protein, coded for by the *gcvH* gene in *E. coli*, is a lipoylprotein that is reduced as it shuttles the methylamine group of glycine from the P-protein to the T-protein and is reoxidized by the dihydrolipoamide dehydrogenase. GcvH functions as a substrate for the three enzymes of the *gcv* complex.

[00769] Residues 61-65 are predicted to contain the lipoyl modification (on lysine), based on conservation of these residues and their correspondence to the lipoate attachment site of the *Pisum sativum* protein.

[00770] The interaction between GcvH and GcvT has been examined. Interaction between the two proteins may be necessary to form the folate binding site, in which the folate polyglutamyl region binds, exposing the pteridine ring. The GcvT N terminus is important for interaction with GcvH, probably by mediating a conformational change, and residue D43 of GcvH is proximal to GcvT in the GcvH-GcvT complex.

[00771] In some embodiments, the H-protein is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *E. coli* gcvH. In some embodiments, the one or more nucleic acid molecule encoding an H-protein comprises an amino acid sequence set forth in UniProt ID P0A6T9. In a further embodiment, the one or more nucleic acid molecule encoding an H-protein is encoded by a nucleic acid sequence set forth in Gene ID 947393.

Glycolate dehydrogenase or glycolate oxidase (EC 1.1.99.14)

[00772] The present disclosure describes enzymes that can catalyze the following reaction:

[00773] glycolate + an oxidized electron acceptor → glyoxylate + a reduced electron acceptor

[00774] Glycolate dehydrogenase may also be known as glycolate oxidase, glycolate oxidoreductase and glycolate:(acceptor) 2-oxidoreductase.

[00775] Glycolate oxidase catalyzes the first step in the utilization of glycolate as the sole source of carbon. The enzyme may be membrane-associated. A cytoplasmic membrane-associated glycolate oxidoreductase activity from *E. coli* ATCC11775 (serovar O1:K1:H7) has been isolated, and the GlcF subunit itself could only be detected in the membrane fraction. The physiological electron acceptor is unknown. Crude extracts from an *E. coli* strain expressing glcDEF from a multicopy plasmid contain glycolate oxidase activity. Insertion mutants in either glcD, glcE, or glcF abolish this activity, suggesting that all three gene products are subunits of a glycolate oxidase complex. Expression of the glcDEFGB operon is induced by growth on glycolate.

[00776] A putative glycolate oxidase in *Arabidopsis thaliana* is a mitochondrial homodimeric protein that binds one FAD per monomer and is expressed in leaves, stems, flowers and roots. Enzyme activity is inhibited by cyanide ions. It catalyzes the oxidation of D-lactate to pyruvate stereospecifically, mediating the detoxification of methylglyoxal and D-lactate.

[00777] In some embodiments, the enzyme having glycolate dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to an enzyme having glycolate dehydrogenase activity selected from *E. coli* glycolate dehydrogenase GLC and *Arabidopsis thaliana* glycolate dehydrogenase. In some embodiments, the one or more nucleic acid molecule encoding a glycolate dehydrogenase or glycolate dehydrogenase

subunit comprises an amino acid sequence selected from UniProt ID P0AEP9, UniProt ID P52073, UniProt ID P52074, and UniProt ID Q94AX4. In a further embodiment, the one or more nucleic acid molecule encoding a glycolate dehydrogenase or glycolate dehydrogenase subunit is encoded by a nucleic acid sequence selected from Gene ID 947353, Gene ID 2847718, Gene ID 2847717, and GenBank accession Y13245.1.

Alanine-glyoxylate aminotransferase (EC 2.6.1.44)

[00778] The present disclosure describes enzymes that can catalyze the following reaction:

[00779] glyoxylate + L-alanine $\leftrightarrow$ glycine + pyruvate

[00780] In *Saccharomyces cerevisiae* alanine--glyoxylate aminotransferase subunit is one of three different enzymes used for glycine synthesis. The AGX1 gene, encoding this enzyme, was identified by comparing enzyme specific activities in knockout strains. When placed in a background deficient for the other enzymes responsible for glycine synthesis, the mutation in AGX1 produced complete glycine auxotrophy. The enzymes was subsequently purified and characterized. The enzyme, which contains pyridoxal 5'-phosphate as cofactor, is a homodimer of about 80 kDa, and is highly specific for L-alanine and glyoxylate.

[00781] The mitochondrially localized *Homo sapiens* alanine--glyoxylate aminotransferase 2 (AGXT2) and peroxisomal serine--pyruvate aminotransferase (AGXT1, see above) both catalyze the conversion of glyoxylate to glycine using alanine as the amino donor. However, AGXT2, but not AGXT1, can also utilize asymmetric dimethylarginine (ADMA) as an amino donor, leading to the formation of α -keto- δ -(NN-dimethylguanidino) valeric acid (DMGV). ADMA is a potent endogenous inhibitor of nitric-oxide (NO) synthase. ADMA levels are also controlled by cytosolic dimethylarginine dimethylaminohydrolases (DDAHs) that hydrolyze ADMA to citrulline and dimethylamine. Elevated levels of ADMA are associated with diabetes, hypertension, congestive heart failure, and atherosclerosis. AGXT2 is a pyridoxal phosphate dependent enzyme that is expressed primarily in the kidney, and its activity is one mechanism by which the kidney regulates blood pressure.

[00782] In *Arabidopsis thaliana*, alanine transaminases with four aminotransferase activities have been identified. AOAT1 (GGAT1) is peroxisomal located. Knock-out plants have reduced activity of AOAT, GPAT (glutamate:pyruvate aminotransferase), AGAT (alanine:glyoxylate aminotransferase) and GGAT (glutamate:glyoxylate aminotransferase). The GGAT and AGAT activities were reduced most drastically. These indicate AOAT1 is primarily involved in photorespiration. Similarly, AOAT2 (GGAT2), which is predicted to be

located in the peroxisome, is likely involved in photorespiration. *In vitro* assay of the recombinant proteins indicated that GGAT1 and GGAT2 have four aminotransferase activities, namely GGAT, AGAT, GPAT and AOAT. The two recombinant proteins exhibited very similar *K_m* values towards amino acid substrates glutamate and alanine, as well as the oxoacid substrates glyoxylate, pyruvate and 2-oxoglutarate.

[00783] In some embodiments, the enzyme having alanine-glyoxylate aminotransferase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to an enzyme having alanine-glyoxylate aminotransferase activity selected from *Saccharomyces cerevisiae* AGX1, *Homo sapiens* AGXT2, *Arabidopsis thaliana* AOAT1 and *Arabidopsis thaliana* AOAT2. In some embodiments, the one or more nucleic acid molecule encoding an enzyme having alanine-glyoxylate aminotransferase activity comprises an amino acid sequence selected from UniProt ID P43567, UniProt ID Q9BYV1, UniProt ID Q9LR30 and UniProt ID Q9S7E9. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having alanine-glyoxylate aminotransferase activity is encoded by a nucleic acid sequence selected from Gene ID 850514, Gene ID 64902, TAIR accession AT1G23310 and TAIR accession AT1G70580.

Alanine transaminase (EC 2.6.1.2)

[00784] The present disclosure describes enzymes that can catalyze the following reaction:

[00785] 2-oxoglutarate + L-alanine ↔ L-glutamate + pyruvate

[00786] In some embodiments, the alanine transaminase is a glutamate-pyruvate aminotransferase. AlaA is one of three major alanine-synthesizing transaminases in *E. coli*. AlaA and AlaC together account for 90% of glutamic-pyruvic transaminase (GPT) activity in the cell. A crystal structure of AlaA has been solved at 2.11 Å resolution. The structure shows a symmetric α₂ homodimer typical of fold type I aminotransferases. An *alaA* deletion strain has no growth defect, but an *alaA avtA* double mutant forms small colonies on agar plates. An *alaA avtA alaC* triple mutant has a slow growth phenotype in liquid medium. The defects of the double and triple mutants can be rescued by addition of alanine. Fitness and competitive growth experiments were performed under different growth conditions. Particularly under oxygen-limiting conditions, the doubling time of the Δ*alaA* strain in minimal media is increased compared to growth in rich media. Under competitive growth conditions, the Δ*alaA* mutation confers a disadvantage compared to wild type even in rich media. *alaA* was identified in a screen for genes that reduce the lethal effects of stress. An *alaA* insertion

mutant is more sensitive than wild type to mitomycin C and other stresses and less sensitive to 10% SDS. The *alaA* gene was first identified as a mutant with a leaky requirement for alanine or valine.

[00787] AlaB is one of three major alanine-synthesizing transaminases in *E. coli*. AlaB catalyzes a glutamate-pyruvate aminotransferase reaction, generating alanine from pyruvate with glutamate as the amino donor. This activity has been assayed in crude cell extracts, and the gene encoding AlaB had been isolated on a plasmid. Expression of *alaB* from a multicopy plasmid partially suppresses the growth defect of an *ilvE alaA* mutant strain.

[00788] AlaC is one of three major alanine-synthesizing transaminases in *E. coli*. A homology model of the enzyme based on the crystal structure of AlaA has been generated. An *alaC* deletion strain has no growth defect, but an *alaA avtA alaC* triple mutant has a slow growth phenotype in liquid medium. The defect can be rescued by addition of alanine. Fitness and competitive growth experiments were performed under different growth conditions. Particularly under oxygen-limiting conditions, the doubling time of the Δ *alaC* strain in minimal media is increased compared to growth in rich media; unlike for the *alaA* and *avtA* mutants, addition of L-alanine returns the doubling time to that observed in DMEM medium. Under competitive growth conditions, the Δ *alaC* mutation confers a disadvantage compared to wild type even in rich media. Expression of *alaC* is activated by the transcriptional regulator SgrR. AlaC may thus play a role in glucose-phosphate stress. However, an *alaC* deletion mutant does not show altered sensitivity to α -methylglucoside, which induces sugar-phosphate stress.

[00789] In *Homo sapiens*, alanine aminotransferase is a cytoplasmic enzyme that catalyzes the reversible transamination between L-alanine and 2-oxoglutarate to form pyruvate and L-glutamate. The interconversion of these four major metabolic intermediates gives this enzyme both degradative and biosynthetic roles. It participates in the alanine-glucose cycle of skeletal muscle and liver, gluconeogenesis, and glutamate generation in the brain. Alanine aminotransferase is expressed in kidney, skeletal muscle, adipose tissue and heart. There are two isoforms of the enzyme: alanine aminotransferase 1 (GPT) and alanine aminotransferase 2 (GPT2). Human alanine aminotransferase 1 (GPT) was purified from liver. Recombinant human alanine aminotransferase 2 (GPT2) from adipose tissue was expressed in *Escherichia coli* and the molecular mass was determined by SDS-PAGE. The enzyme is expressed at high levels in adipose tissue, muscle, brain and kidney and at lower levels in breast and liver.

[00790] The subunit structure of an alanine transaminase from the marine polychaete annelid *Arenicola marina* (lugworm) has not been reported. It has a native apparent

molecular mass of 91 kDa as determined by gel filtration chromatography. The gene encoding it in this organism has not been identified. In marine annelids and mollusks this reaction participates in an anaerobic energy generation pathway that operates during periods of hypoxia or anoxia. Alanine transaminase (glutamate pyruvate transaminase) from this organism has been partially purified from body wall muscle and characterized. High alanine transaminase activity was found in this tissue. Specific, reversible, L-glutamate - dependent L-alanine transaminase and L-aspartate transaminase activities have also been demonstrated in tissues of the mussel *Mytilus edulis*.

[00791] The *Arabidopsis thaliana* tryptophan aminotransferase TAA1 protein is involved in the formation of indole-3-pyruvate, a precursor to indole-3-acetate (IAA), a biologically important auxin that acts as a phytohormone in many plant species. *In vitro* assays reveal that this pyridoxal 5'-phosphate (PLP)-dependent aminotransferase can act on a number of different L-amino acids, including L-phenylalanine, L-tyrosine, L-leucine, L-alanine, L-methionine, and L-glutamine using either pyruvate or 2-oxoglutarate as a cosubstrate. However, enzymatic assays, *in silico* docking experiments, and mutant phenotypic analyses all suggest that L-tryptophan is the *in vivo* substrate for this enzyme. TAA1 has a K_m of 0.29 mM and a V_{max} of 12.9 $\mu\text{M}/\text{min}$ when tested with L-tryptophan and pyruvate. It is unclear whether pyruvate or 2-oxoglutarate is the more biologically relevant cosubstrate for this enzyme.

[00792] The gene encoding this protein was identified in screens for shade avoidance mutants and mutants with a weak ethylene insensitivity suggesting that the auxin synthesized through a TAA1-mediated pathway is particularly important for the responses to specific environmental and hormonal stimuli. In addition, normal developmental processes, such as embryogenesis, that require proper auxin levels, are disrupted when TAA1 and one or more of its closely related family members (i.e. TAR1 and TAR2) are knocked-out in *Arabidopsis* plants. This enzyme activity appears to be widely distributed in the plant kingdom, based on the ability of enzymatic extracts from 30 different species distributed among 16 families, to catalyze the formation of IPA in a transaminase reaction. Three species of algae also have this activity.

[00793] In *Arabidopsis thaliana*, alanine transaminases with four aminotransferase activities have been identified. AOAT1 (GGAT1) is peroxisomal located. Knock-out plants have reduced activity of AOAT, GPAT (glutamate:pyruvate aminotransferase), AGAT (alanine:glyoxylate aminotransferase) and GGAT (glutamate:glyoxylate aminotransferase). The GGAT and AGAT activities were reduced most drastically. These indicate AOAT1 is primarily involved in photorespiration. Similarly, AOAT2 (GGAT2), which is predicted to be located in the peroxisome, is likely involved in photorespiration. *In vitro* assay of the

recombinant proteins indicated that GGAT1 and GGAT2 have four aminotransferase activities, namely GGAT, AGAT, GPAT and AOAT. The two recombinant proteins exhibited very similar Km values towards amino acid substrates glutamate and alanine, as well as the oxoacid substrates glyoxylate, pyruvate and 2-oxoglutarate.

[00794] Alanine aminotransferase activities have also been described from *Candida maltosa*, *Clostridium propionicum*, *Pyrococcus furiosus*, *Megathyrus maximus* and *Panicum miliaceum*. The *Panicum miliaceum* putative alanine aminotransferase was cloned from this NAD-ME type C4 plant and found to express in both mesophyll and bundle sheath cells, and the gene expression was light-inducible. The mRNA accumulation increased dramatically during greening in both cell types which is in agreement with its predicted role in C4 photosynthesis.

[00795] In some embodiments, the enzyme having alanine transaminase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to an enzyme having alanine transaminase activity selected from *E. coli* glutamate-pyruvate aminotransferase alaA, *E. coli* glutamate-pyruvate aminotransferase alaB, *E. coli* glutamate-pyruvate aminotransferase alaC, *Homo sapiens* alanine aminotransferase 1 (GPT), *Homo sapiens* alanine aminotransferase 2 (GPT2), *Arenicola marina* alanine transaminase, *Arabidopsis thaliana* tryptophan aminotransferase TAA1, *Arabidopsis thaliana* AOAT1, *Arabidopsis thaliana* AOAT2, *Candida maltosa* alanine aminotransferase, *Clostridium propionicum* alanine aminotransferase, *Pyrococcus furiosus* alanine aminotransferase aat, *Megathyrus maximus* alanine transaminase, and *Panicum miliaceum* alanine transaminase AlaAT-2. In some embodiments, the one or more nucleic acid molecule encoding enzyme having alanine transaminase activity comprises an amino acid sequence selected from UniProt ID P0A959, UniProt ID P77434, UniProt ID P24298, UniProt ID Q8TD30, UniProt ID Q9S7N2, UniProt ID Q9LR30, UniProt ID Q9S7E9, UniProt ID Q9P9M8, and UniProt ID P34106. In a further embodiment, the one or more nucleic acid molecule encoding enzyme having alanine transaminase activity is encoded by a nucleic acid sequence selected from Gene ID 946772, Gene ID 946850, Gene ID 2875, Gene ID 84706, Gene ID 843393, TAIR accession AT1G23310, TAIR accession AT1G70580, GenBank accession AF163769.1 and GenBank accession X69421.1.

Glutamate dehydrogenase (EC 1.4.1.2 and EC 1.4.1.3)

[00796] The present disclosure describes enzymes that can catalyze the following reactions:

[00797] L-glutamate + NAD⁺ + H₂O ↔ 2-oxoglutarate + ammonium + NADH + H⁺

[00798] L-glutamate + NAD(P)⁺ + H₂O ↔ 2-oxoglutarate + ammonium + NAD(P)H + H⁺

[00799] In some embodiments, the glutamate dehydrogenase is an NAD-dependent glutamate dehydrogenase from *Saccharomyces cerevisiae* (GDH2) that degrades glutamate to ammonia and alpha-ketoglutarate. Expression of GDH2 is sensitive to nitrogen catabolite repression and intracellular ammonia levels.

[00800] There are two NAD-dependent glutamate dehydrogenase (GDH) genes in *Arabidopsis*, GDH1 and GDH2, encoding the alpha- and beta-subunits, respectively. Seven hexameric isoforms of GDH have been detected which are composed of different ratios of the alpha and beta subunits. Different isoforms are distributed in different tissues under different environmental and physiological conditions. The enzyme activity of GDH is controlled in part at the transcriptional level.

[00801] Glutamate dehydrogenase from *Peptoniphilus asaccharolyticus* catalyzes the NAD-dependent, oxidative deamination of L-glutamate to 2-oxoglutarate. The reaction is highly substrate specific. No activity was observed in the presence of D-glutamate, D- or L-aspartate, glutamine or when NADP replaced NAD. Using sucrose gradient density centrifugation, researchers estimated that the molecular weight of the native enzyme was between 300-340 kDa, suggesting that the enzyme may be a homohexamer. Other researchers, on the other hand, estimated that the molecular weight of the native enzyme was approximately 226 kDa, using gel filtration. The K_M values for L-glutamate and NAD⁺ were 1.3 mM and 0.25 mM, respectively.

[00802] *Halobacterium salinarum* is one of the organisms reported to have more than one form of GDH, with different forms utilizing different cofactors (NAD and NADP). It was eventually shown that the organism has four genes encoding four different glutamate dehydrogenase enzymes. Two of these gene products have been purified and characterized biochemically. One of the genes, *gdhA1*, which was originally predicted to encode an NADP-specific form, was found to encode an NAD-specific enzyme.

[00803] *Homo sapiens* glutamate dehydrogenases (GDHs) are homohexameric mitochondrial matrix enzymes that catalyze the reversible oxidative deamination of L-glutamate to 2-oxoglutarate and ammonia. Mammalian GDHs are unusual enzymes, in that they are able to use either NAD or NADP as a co-factor. Humans express two GDH isoenzymes. Glutamate dehydrogenase 1 (GLUD1) is expressed at high levels in liver, brain, pancreas and kidney. Glutamate dehydrogenase 2 (GLUD2) is encoded by an X chromosome-linked intronless gene and is expressed in retina, testis, and brain. Mutations in

GLUD1 that lead to enzyme overactivity, result in hyperinsulinemia. Allosteric control of mammalian GDH activity by positive effectors like ADP and L-leucine and negative effectors like GTP have been extensively studied.

[00804] *Bacillus subtilis* PCI 219 has a single glutamate dehydrogenase (GDH) with dual coenzyme specificity for NAD(H) and NADP(H). Its molecular weight was estimated to be 250,000+/-20,000 by gel filtration, and 270,000+/-30,000 by zone centrifugation in a sucrose density gradient. The subunit size was about 57,000, suggesting that it is a homotetramer.

[00805] A cDNA clone was isolated from *Solanum lycopersicum* tissues. The cDNA encoded a protein that shares identity with glutamate dehydrogenase (GDH) of plants. Expression analysis of this protein showed that it is expressed in stems, roots and leaves but is absent in fruit tissues. The study also showed that the two subunits of tomato GDH1 were encoded by a single gene. This enzyme converts 2-oxoglutarate to L-glutamate.

[00806] Glutamate dehydrogenase activities have also been described from *Clostridium propionicum* and *Thermotoga maritima*.

[00807] In some embodiments, the enzyme having glutamate dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to a glutamate dehydrogenase selected from *Saccharomyces cerevisiae* NAD-dependent glutamate dehydrogenase GDH2, *Arabidopsis thaliana* NAD-dependent glutamate dehydrogenase GDH2, *Arabidopsis thaliana* NAD-dependent glutamate dehydrogenase GDH1, *Peptoniphilus asaccharolyticus* NAD-dependent glutamate dehydrogenase gdhA, *Halobacterium salinarum* NAD-dependent glutamate dehydrogenase gdhA, *Thermotoga maritima* glutamate dehydrogenase, *Homo sapiens* glutamate dehydrogenase 1 (GLUD1), *Homo sapiens* glutamate dehydrogenase 2 (GLUD2), *Bacillus subtilis* glutamate dehydrogenase and *Solanum lycopersicum* glutamate dehydrogenase GDH1. In some embodiments, the one or more nucleic acid molecule encoding an enzyme having glutamate dehydrogenase activity comprises an amino acid sequence selected from UniProt ID P33327, UniProt ID Q38946, UniProt ID Q38946, UniProt ID P28997, UniProt ID P29051, UniProt ID P00367, UniProt ID P49448 and UniProt ID P93541. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having glutamate dehydrogenase activity is encoded by a nucleic acid sequence selected from Gene ID 461927, TAIR accession AT5G07440, TAIR accession AT5G18170, GenBank accession M76403.1, GenBank accession X63837.1, Gene ID 2746, Gene ID 2747 and GenBank accession U48695.1.

Acetaldehyde dehydrogenase (EC 1.2.1.10)

[00808] The present disclosure describes enzymes that can catalyze the following reaction:

[00809] acetaldehyde + coenzyme A + NAD⁺ ↔ **acetyl**-CoA + NADH + H⁺

- 5 **[00810]** *E. coli* mhpF encodes an acylating acetaldehyde dehydrogenase. MhpF is active as a monomer; the rate-limiting step of the reaction appears to be transthioesterification. MhpF is involved in synthesis of n-butanol in an engineered reversal of the β-oxidation pathway. The expression of MhpE is translationally coupled to MhpF, and interaction between the two proteins appears to be required for solubility of MhpE.
- 10 **[00811]** *E. coli* AdhE is a homopolymeric protein with three Fe²⁺-dependent catalytic functions: alcohol dehydrogenase, coenzyme A-dependent acetaldehyde dehydrogenase, and pyruvate formate-lyase deactivase. However, the existence of the pyruvate formate-lyase deactivase activity of AdhE has been debated. The homopolymeric structure of AdhE is unusual in that 20-60 subunits are helically arranged to form rod-like ultrastructures. Under
- 15 fermentative conditions AdhE catalyzes the reduction of acetyl-CoA to acetaldehyde and the latter compound to ethanol. Aerobically, in the reverse direction, AdhE can catalyze the oxidation of acetaldehyde to acetyl-CoA. Expression of adhE appears to be regulated at the transcriptional and translational levels, and possibly at the posttranslational level. Expression of adhE is approximately 10-fold higher during anaerobic growth than during aerobic growth.
- 20 The AdhE from *E. coli* B was partially purified and characterized in early work. It was later purified to homogeneity from *E. coli* B and its coenzyme A-linked aldehyde dehydrogenase activity was subjected to detailed kinetic analysis. A bi-uni-uni-uni ping-pong mechanism was proposed. AdhE from *Salmonella enterica* subsp. *enterica* serovar Typhimurium showed 70% amino acid sequence identity to that of *E. coli*. It had a lower K_m for alcohol substrates
- 25 and some differences in substrate specificity as compared to the *E. coli* enzyme. In the metabolic engineering field, deletion of adhE is a determinant in the production of compounds such as succinate, D-lactate, and polyhydroxyalkanoates.

[00812] *Chlamydomonas reinhardtii* ADH1 encodes a dual function alcohol dehydrogenase / acetaldehyde dehydrogenase. It appears to be active under anoxic

30 conditions and participates in two different anaerobic ethanol production pathways in *Chlamydomonas*.

[00813] DmpF is an acylating acetaldehyde dehydrogenase from *Pseudomonas* sp. The final two steps of the meta-cleavage pathway in *Pseudomonas* sp. CF600 involve the conversion of (S)-4-hydroxy-2-oxopentanoate to pyruvate and acetyl-CoA by the enzymes 4-

hydroxy-2-oxovalerate aldolase and acetaldehyde dehydrogenase (acylating). Biochemical studies demonstrated that these two enzymes comprise a bifunctional aldolase-dehydrogenase heterodimer, and suggest that the product of the aldolase reaction, acetaldehyde, is transferred to the dehydrogenase active site via a channeling mechanism.

5 This minimizes the risk to the cells posed by the toxic acetaldehyde.

[00814] The presence of the *todI* gene product in *Pseudomonas putida* F1 was suggested by the protein expression pattern of plasmid constructs. The predicted amino acid sequence of the *todI* gene product showed very high identity with other bacterial acylating aldehyde dehydrogenase gene products.

10 **[00815]** The *adhE* gene of *Clostridium acetobutylicum* ATCC 824 encodes a multifunctional enzyme that has both alcohol dehydrogenase and acetaldehyde dehydrogenase activities. Both activities are necessary for the formation of butan-1-ol and ethanol during solventogenesis. The *adhE* gene is part of the *sol* operon, which is located on the pSOL1 megaplasmid. Expression of the gene from a plasmid in *Clostridium*
15 *acetobutylicum* ATCC 824 resulted in elevated activities of NADH-dependent butanol dehydrogenase, NAD-dependent acetaldehyde dehydrogenase and butyraldehyde dehydrogenase, and a small increase in NADH-dependent ethanol dehydrogenase. Complementation of a mutant deficient in butyraldehyde dehydrogenase, acetoacetate decarboxylase, and acetoacetyl-coenzyme A:acetate/butyrate:coenzyme A-transferase
20 activities, which produces neither butanol nor acetone, by the *adhE* gene resulted in restored butanol formation without any acetone formation or any significant increase in ethanol production, suggesting that the primary role of the enzyme is in butanol formation, providing both a butanal dehydrogenase activity (converting butanoyl-CoA to butan-1-al) and butanol dehydrogenase activity. In addition, inactivation of the gene drastically reduced butanol
25 production (by 85%), supporting this role. Another gene from the pSOL1 plasmid, *adhE2*, encodes a second multifunctional aldehyde/alcohol dehydrogenase involved in butanol production. However, that enzyme is produced only under alcohologenic conditions, and is not expressed under solventogenic conditions. The gene from strain ATCC 824 was originally called *aad*.

30 **[00816]** Acetaldehyde dehydrogenase activities have also been described from *Leuconostoc mesenteroides*, *Pelobacter acetylenicus* and *Pseudomonas putida*. The partially purified *Leuconostoc mesenteroides* CoA-dependent aldehyde dehydrogenase could not be separated from an NAD-linked alcohol dehydrogenase that co-purified with it. The enzyme was specific to NAD and could not use NADP. While acetaldehyde and 1-
35 propanal were the best substrates, the enzyme could also use butan-1-al (31% of activity with acetaldehyde) and isobutanol (14%).

[00817] In some embodiments, the enzyme having acetaldehyde dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to an enzyme having acetaldehyde dehydrogenase activity selected from *E. coli* mhpF, *E. coli* AdhE, *Chlamydomonas reinhardtii* ADH1, *Leuconostoc mesenteroides* CoA-dependent acetaldehyde dehydrogenase, *Pelobacter acetylenicus* acetaldehyde dehydrogenase, *Pseudomonas sp.* dmpF, *Pseudomonas putida* acylating aldehyde dehydrogenase todI, *Pseudomonas putida* acetaldehyde dehydrogenase cmhH and *Clostridium acetobutylicum* alcohol/aldehyde dehydrogenase AdhE. In some embodiments, the one or more nucleic acid molecule encoding an enzyme having acetaldehyde dehydrogenase activity comprises an amino acid sequence selected from UniProt ID P77580, UniProt P0A9Q7, UniProt ID A8JI07, UniProt ID Q52060, UniProt ID Q51949 and UniProt ID P33744. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having acetaldehyde dehydrogenase activity is encoded by a nucleic acid sequence selected from Gene ID 945008, Gene ID 945837, Gene ID 5729132, GenBank accession X60835.1, GenBank accession U09250.1 and Gene ID 1116167.

Ethanolamine ammonia lyase (EC 4.3.1.7)

[00818] The present disclosure describes enzymes that can catalyze the following reaction:

[00819] ammonium + acetaldehyde ↔ ethanolamine

[00820] Ethanolamine ammonia-lyase (EAL) allows *E. coli* to utilize ethanolamine as the sole source of nitrogen and carbon in the presence of external vitamin B12. EAL is an adenosylcobalamin-dependent enzyme that is spontaneously inactivated by its substrate and can be reactivated by EutA. The enzyme was first studied in the non-K-12 strain NCIB 8114. Crystal structures of an N-terminally truncated, but active form of the enzyme both in binary and ternary complexes with the cofactor and substrate have been solved. The enzyme is composed of a hexamer of ($\alpha\beta$)₂ dimers, with the α subunit holding the active site and the cobalamin cofactor bound at the interface between the α and β subunits. The authors propose a reaction mechanism that is consistent with a previously described mechanism for adenosylcobalamin-dependent rearrangements. The stereochemical course of the reaction has been modeled on the basis of crystal structures, accounting for the apparent lack of stereospecificity of the enzyme. Production of EAL is catabolite repressed and is induced by the simultaneous presence of ethanolamine and the adenosylcobalamin cofactor. Ethanolamine ammonia-lyase comprises two subunits, α (EutB) and β (EutC).

[00821] In some embodiments, the enzyme having ethanolamine ammonia lyase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to an *E. coli* enzyme having ethanolamine ammonia lyase activity. In some embodiments, the one or more nucleic acid molecule encoding an ethanolamine ammonia lyase subunit comprises an amino acid sequence selected from UniProt ID P0AEJ6 and UniProt ID P19636. In a further embodiment, the one or more nucleic acid molecule encoding an ethanolamine ammonia lyase subunit is encoded by a nucleic acid sequence selected from Gene ID 946924 and Gene ID 946925.

Serine aminase (EC 2.6.1.-)

[00822] The present disclosure describes enzymes that can catalyze the following reaction:

[00823] L-Serine + NH₄⁺ → (S)-2,3-diaminopropanoate

[00824] In some embodiments, an enzyme having serine aminase activity can be a serine-glyoxylate transaminase. In other embodiments, an enzyme having serine aminase activity can be a serine-pyruvate transaminase (see above).

[00825] Serine-glyoxylate aminotransferase catalyzes the transfer of the α-amino group of L-serine to glyoxylate, forming glycine and hydroxypyruvate. In the serine-cycle methylotrophs this enzyme plays two important roles: the formation of an acceptor (glycine) for a one-carbon unit, and the conversion of L-serine to hydroxypyruvate in the assimilatory pathway. This is the first microbial serine-glyoxylate aminotransferase to be purified, and a few years later the gene encoding it was identified as well.

[00826] Serine: glyoxylate aminotransferase encoded by *Arabidopsis thaliana* AGT1 is a homodimer. The purified recombinant protein has the highest activity with the serine: glyoxylate transamination. It also catalyzed alanine: glyoxylate transamination and serine: pyruvate transamination with much lower specific activity.

[00827] *sgaA* is presumed to be the gene encoding serine-glyoxylate aminotransferase on the *Methylobacterium extorquens* chromosome. While the product of the cloned gene has not been expressed, mutation complementation was used to investigate its role. Mutations in this gene have abolished the ability to grow on C1 compounds, and complementation of the mutants by a cloned intact version of the gene fragment has restored activity. In addition, *sgaA* sequence is similar to a number of aminotransferases.

[00828] In further embodiments, an enzyme having serine aminase activity can be an enzyme having phosphoserine aminotransferase activity. In some embodiments, the one or more nucleic acid molecules encoding the phosphoserine aminotransferase is serC, or homolog thereof. In another embodiment, the one or more nucleic acid molecules encoding an enzyme having phosphoserine aminotransferase activity, an enzyme having L-serine transaminase activity, or an enzyme having serine aminotransferase activity comprise an amino acid sequence set forth in SEQ ID NO: 230. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having phosphoserine aminotransferase activity, an enzyme having L-serine transaminase activity, or an enzyme having serine aminotransferase activity is encoded by a nucleic acid sequence set forth in SEQ ID NO: 229. In some embodiments, the enzyme having phosphoserine aminotransferase activity is *Homo sapiens* PSAT1, or homolog thereof. In one embodiment, the one or more nucleic acid molecule encoding an enzyme having phosphoserine aminotransferase activity comprises an amino acid sequence set forth in UniProt ID Q9Y617. In another embodiment, the one or more nucleic acid molecule encoding an enzyme having phosphoserine aminotransferase activity is encoded by a nucleic acid sequence set forth in Gene ID 29968.

[00829] In some embodiments, the enzyme having serine aminase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to an *Arabidopsis thaliana* serine-glyoxylate aminotransferase AGT1, *Hyphomicrobium methylovorum* GM2 serine-glyoxylate aminotransferase sgaA and *Methylobacterium extorquens* sgaA. In some embodiments, the one or more nucleic acid molecule encoding an enzyme having serine-glyoxylate aminotransferase activity comprises an amino acid sequence selected from UniProt ID Q56YA5 and UniProt ID O08374. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having serine-glyoxylate aminotransferase activity is encoded by a nucleic acid sequence selected from TAIR accession AT2G13360, GenBank accession D86125.1 and GenBank accession L27235.1.

[00830] In some embodiments, an enzyme having serine aminase activity is an enzyme having serine-pyruvate aminotransferase activity. In some embodiments, the enzyme having serine-pyruvate aminotransferase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *Homo sapiens* AGXT1. In some embodiments, the enzyme having serine-pyruvate aminotransferase activity is *Homo sapiens* AGXT1. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having serine-pyruvate aminotransferase activity is AGXT1, or homolog thereof. In some embodiments, the one or more nucleic acid molecules encoding an enzyme having serine-pyruvate aminotransferase

activity comprise an amino acid sequence set forth in SEQ ID NO: 244. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having serine-pyruvate aminotransferase activity is encoded by a nucleic acid sequence set forth in SEQ ID NO: 243.

5 **[00831]** In some embodiments, an enzyme having serine aminase activity is an enzyme having alanine-glyoxylate aminotransferase activity. In some embodiments, the enzyme having alanine-glyoxylate aminotransferase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to an enzyme having alanine-glyoxylate aminotransferase
 10 activity selected from *Saccharomyces cerevisiae* AGX1, *Homo sapiens* AGXT2, *Arabidopsis thaliana* AOAT1 and *Arabidopsis thaliana* AOAT2. In some embodiments, the one or more nucleic acid molecule encoding an enzyme having alanine-glyoxylate aminotransferase activity comprises an amino acid sequence selected from UniProt ID P43567, UniProt ID Q9BYV1, UniProt ID Q9LR30 and UniProt ID Q9S7E9. In a further embodiment, the one or
 15 more nucleic acid molecule encoding an enzyme having alanine-glyoxylate aminotransferase activity is encoded by a nucleic acid sequence selected from Gene ID 850514, Gene ID 64902, TAIR accession AT1G23310 and TAIR accession AT1G70580.

[00832] In some embodiments, an enzyme having serine aminase activity is an enzyme having alanine transaminase activity. In some embodiments, the enzyme having alanine
 20 transaminase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to an enzyme having alanine transaminase activity selected from *E. coli* glutamate-pyruvate aminotransferase alaA, *E. coli* glutamate-pyruvate aminotransferase alaB, *E. coli* glutamate-pyruvate aminotransferase alaC, *Homo sapiens* alanine aminotransferase 1 (GPT), *Homo*
 25 *sapiens* alanine aminotransferase 2 (GPT2), *Arenicola marina* alanine transaminase, *Arabidopsis thaliana* tryptophan aminotransferase TAA1, *Arabidopsis thaliana* AOAT1, *Arabidopsis thaliana* AOAT2, *Candida maltosa* alanine aminotransferase, *Clostridium propionicum* alanine aminotransferase, *Pyrococcus furiosus* alanine aminotransferase aat, *Megathyrus maximus* alanine transaminase, and *Panicum miliaceum* alanine transaminase
 30 AlaAT-2. In some embodiments, the one or more nucleic acid molecule encoding an enzyme having alanine transaminase activity comprises an amino acid sequence selected from UniProt ID P0A959, UniProt ID P77434, UniProt ID P24298, UniProt ID Q8TD30, UniProt ID Q9S7N2, UniProt ID Q9LR30, UniProt ID Q9S7E9, UniProt ID Q9P9M8, and UniProt ID P34106. In a further embodiment, the one or more nucleic acid molecule encoding an
 35 enzyme having alanine transaminase activity is encoded by a nucleic acid sequence selected from Gene ID 946772, Gene ID 946850, Gene ID 2875, Gene ID 84706, Gene ID

843393, TAIR accession AT1G23310, TAIR accession AT1G70580, GenBank accession AF163769.1 and GenBank accession X69421.1.

5 **Diaminobutyrate decarboxylase (EC 4.1.1.86)**

[00833] The present disclosure describes enzymes that can catalyze the following reaction:

[00834] L-2,4-diaminobutanoate $\leftrightarrow$ propane-1,3-diamine + CO₂

10 **[00835]** The systematic name of this enzyme class is L-2,4-diaminobutanoate carboxy-lyase (propane-1,3-diamine-forming). Diaminobutyrate decarboxylase may also be known as DABA DC, L-2,4-diaminobutyrate decarboxylase, L-2,4-diaminobutanoate carboxy-lyase.

15 **[00836]** This enzyme belongs to the family of transferases, specifically those transferring phosphorus-containing groups (phosphotransferases) with an alcohol group as acceptor. This enzyme participates in 3 metabolic pathways: serine/glycine/threonine metabolism, glycerolipid metabolism, and glyoxylate-dicarboxylate metabolism.

20 **[00837]** In some embodiments, a gene encoding a L-2,4-diaminobutyrate decarboxylase is the rhbB gene of *Sinorhizobium meliloti* Rm2011. The rhbB gene is homologous to the ddc gene of *Acinetobacter baumannii*, which are involved in the production of propane-1,3-diamine by that organism. Thus, it is proposed that these genes encode diaminobutyrate--2-oxoglutarate transaminase and L-2,4-diaminobutyrate decarboxylase, respectively.

25 **[00838]** In some embodiments, the diaminobutyrate decarboxylase is the L-2,4-diaminobutyrate decarboxylase. In some embodiments, the L-2,4-diaminobutyrate decarboxylase is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *Sinorhizobium meliloti* Rm2011 rhbB. In some embodiments, the L-2,4-diaminobutyrate decarboxylase is *Sinorhizobium meliloti* Rm2011 rhbB.

30 **[00839]** In some embodiments, the one or more nucleic acid molecules encoding the L-2,4-diaminobutyrate decarboxylase is rhbB, or homolog thereof. In some embodiments, the one or more nucleic acid molecule encoding L-2,4-diaminobutyrate decarboxylase comprises an amino acid sequence set forth in UniProt ID Q9Z3R2. In a further embodiment, the one or more nucleic acid molecule encoding L-2,4-diaminobutyrate decarboxylase is encoded by a nucleic acid sequence set forth in Gene ID 1236295.

Ornithine decarboxylase (EC 4.1.1.17)

[00840] The present disclosure describes enzymes that can catalyze the following reaction:

[00841] $\text{L-ornithine} + \text{H}^+ \leftrightarrow \text{CO}_2 + \text{putrescine}$

[00842] The systematic name of this enzyme class is L-ornithine carboxy-lyase (putrescine-forming). Ornithine decarboxylase may also be known as SpeC and L-ornithine carboxy-lyase. *E. coli* encodes two ornithine decarboxylase enzymes, the biosynthetic (constitutive) SpeC and the degradative (inducible) SpeF. SpeF is activated by the guanosine nucleotides GTP, GDP, pppGpp and ppGpp. SpeF is overproduced from a *speF*-containing plasmid when cells are grown at pH 5.2, but not at pH 7. Overexpression of RNase III from a plasmid increases expression from the *speF* promoter, perhaps by processing of the 5' UTR.

[00843] When ornithine is available, *E. coli* is able to synthesize putrescine via two pathways. One pathway yields putrescine directly from ornithine by a reaction catalyzed by ornithine decarboxylase. The other pathway involves two reactions using arginine. In addition, *E. coli* possesses two forms of ornithine decarboxylase, a biosynthetic (or constitutive) form apparently present in all strains, and a biodegradative (or inducible) form which is present in some strains of *E. coli*. These two types of ornithine decarboxylases have been characterized in *E. coli* K-12. The biodegradative ornithine decarboxylase is induced by low pH and by the presence of ornithine in the growth medium. The activity of the biosynthetic ornithine decarboxylases is modulated by a number of positive and negative effectors. The positive effectors include nucleotides, GTP being more effective in activating ornithine decarboxylase, while ppGpp reacts as a negative effector of ornithine decarboxylase. The accumulation of ppGpp leads to the cessation of stable RNA synthesis and appears to be related to the fidelity of protein synthesis.

[00844] Ornithine decarboxylase regulation was reported to occur by a protein inhibitor named antizyme that binds to ornithine decarboxylase and non-competitively inhibits its activity. The antizyme is induced by polyamines. The gene encoding *E. coli* antizyme was cloned and identified as a bifunctional antizyme/transcriptional regulator. However another publication concluded that the *E. coli* antizyme is not a direct counterpart of mammalian antizyme genes. The product of *E. coli* gene *atoC* (see AtoC transcriptional activator; Az protein inhibitor of ODC) was later identified as the bifunctional protein that has both antizyme activity (posttranslational inhibition of polyamine biosynthetic enzymes) and activity as a transcriptional regulator of genes involved in short-chain fatty acid catabolism. Polyamines activated transcription of *atoC*. AtoC along with sensory histidine kinase AtoS is part of a two-component signal transduction system. Although AtoS-catalyzed

phosphorylation of AtoC is essential for transcriptional activation of the *atoDAEB* operon for catabolism of short chain fatty acids, it is not yet known whether phosphorylation of antizyme plays a role in regulation of ornithine decarboxylase.

[00845] In some embodiments, genes encoding an ornithine decarboxylase are the *speC* and *speF* genes of *Escherichia coli*. In some embodiments, the ornithine decarboxylase is the L-ornithine decarboxylase. In some embodiments, the ornithine decarboxylase is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *Escherichia coli speC* or *speF*. In some embodiments, the ornithine decarboxylase is *Escherichia coli speC* or *speF*.

[00846] In some embodiments, the one or more nucleic acid molecules encoding the ornithine decarboxylase is *speC*, or homolog thereof. In some embodiments, the one or more nucleic acid molecules encoding the ornithine decarboxylase is *speF*, or homolog thereof. In some embodiments, the one or more nucleic acid molecules encoding ornithine decarboxylase comprise an amino acid sequence selected from UniProt ID P21169 and UniProt ID P24169. In a further embodiment, the one or more nucleic acid molecule encoding ornithine decarboxylase is encoded by a nucleic acid sequence selected from Gene ID 947457 and Gene ID 945297.

[00847] 2,3-diaminopropionate ammonia-lyase (EC 4.3.1.15)

[00848] The present disclosure describes enzymes that can catalyze the following reaction:

[00849] 2 ammonium + pyruvate $\leftrightarrow$ 2,3-diaminopropanoate + H₂O

[00850] 2,3-Diaminopropionate ammonia-lyase is not stereospecific and catalyzes the α,β -elimination of both the D and L stereoisomer of 2,3-diaminopropionate. The enzyme also exhibits weak activity toward D-serine, and does not exhibit activity toward L-serine, D- β -Cl-alanine, or L- β -Cl-alanine.

[00851] The enzyme is homodimeric and contains a pyridoxal 5'-phosphate prosthetic group, belonging to the fold-type II family of PLP-containing enzymes. Crystal structures of the apo- and holoenzyme and the enzyme in complex with a reaction intermediate and substrate have been solved. Kinetic properties of mutants in active site residues were analyzed, and a reaction mechanism was proposed.

[00852] In some embodiments, an enzyme having 2,3-diaminopropionate ammonia-lyase activity is encoded by an amino acid sequence having at least 70% sequence identity,

having at least 80% sequence identity, or having at least 90% sequence identity to *E. coli* 2,3-diaminopropionate ammonia-lyase ygeX. In other embodiments, the enzyme having 2,3-diaminopropionate ammonia-lyase activity is *E. coli* ygeX. In some embodiments, the one or more nucleic acid molecule encoding an enzyme having 2,3-diaminopropionate ammonia-lyase activity comprises an amino acid sequence set forth in UniProt ID P66899. In further
5 embodiments, the one or more nucleic acid molecule encoding an enzyme having 2,3-diaminopropionate ammonia-lyase activity is encoded by a nucleic acid sequence set forth in Gene ID 947012.

10 Glyoxylate shunt

[00853] The glyoxylate cycle, a variation of the tricarboxylic acid cycle, is an anabolic pathway occurring in plants, bacteria, protists, and fungi. The glyoxylate cycle centers on the conversion of acetyl-CoA to succinate for the synthesis of carbohydrates. In microorganisms, the glyoxylate cycle allows cells to utilize simple carbon compounds as a carbon source
15 when complex sources such as glucose are not available. The cycle is generally assumed to be absent in animals, with the exception of nematodes at the early stages of embryogenesis. In recent years, however, the detection of malate synthase and isocitrate lyase, key enzymes involved in the glyoxylate cycle, in some animal tissue has raised questions regarding the evolutionary relationship of enzymes in bacteria and animals and suggests
20 that animals encode alternative enzymes of the cycle that differ in function from known malate synthase and isocitrate lyase in non-metazoan species.

[00854] The glyoxylate cycle utilizes five of the eight enzymes associated with the tricarboxylic acid cycle: citrate synthase, aconitase, succinate dehydrogenase, fumarase, and malate dehydrogenase. The two cycles differ in that in the glyoxylate cycle, isocitrate is
25 **converted into glyoxylate and succinate by isocitrate lyase instead of into α -ketoglutarate.** This bypasses the decarboxylation steps that take place in the TCA cycle, allowing simple carbon compounds to be used in the later synthesis of macromolecules, including glucose. Glyoxylate is subsequently combined with acetyl-CoA to produce malate, catalyzed by malate synthase. Malate is also formed in parallel from succinate by the action of succinate
30 dehydrogenase and fumarase.

[00855] Fatty acids from lipids are commonly used as an energy source by vertebrates as fatty acids are degraded through beta oxidation into acetate molecules. This acetate, bound to the active thiol group of coenzyme A, enters the citric acid cycle (TCA cycle) where it is fully oxidized to carbon dioxide. This pathway thus allows cells to obtain energy from fat. To

utilize acetate from fat for biosynthesis of carbohydrates, the glyoxylate cycle, whose initial reactions are identical to the TCA cycle, is used.

[00856] Cell-wall containing organisms, such as plants, fungi, and bacteria, require very large amounts of carbohydrates during growth for the biosynthesis of complex structural polysaccharides, such as cellulose, glucans, and chitin. In these organisms, in the absence of available carbohydrates (for example, in certain microbial environments or during seed germination in plants), the glyoxylate cycle permits the synthesis of glucose from lipids via **acetate generated in fatty acid β -oxidation**.

[00857] The glyoxylate cycle bypasses the steps in the citric acid cycle where carbon is lost in the form of CO₂. The two initial steps of the glyoxylate cycle are identical to those in **the citric acid cycle: acetate → citrate → isocitrate. In the next step, catalyzed by the first** glyoxylate cycle enzyme, isocitrate lyase, isocitrate undergoes cleavage into succinate and glyoxylate (the latter gives the cycle its name). Glyoxylate condenses with acetyl-CoA (a step catalyzed by malate synthase), yielding malate. Both malate and oxaloacetate can be converted into phosphoenolpyruvate, which is the product of phosphoenolpyruvate carboxykinase, the first enzyme in gluconeogenesis. The net result of the glyoxylate cycle is therefore the production of glucose from fatty acids. Succinate generated in the first step can enter into the citric acid cycle to eventually form oxaloacetate.

Biosynthesis of MEG (or glycolic acid), or optionally, MEG (or glycolic acid) and one or more co-product using a recombinant microorganism

[00858] As discussed above, in one aspect, the present disclosure provides a recombinant microorganism comprising one or more biochemical pathway that produces monoethylene glycol (MEG) (or glycolic acid) from one or more hexose feedstock via one or more pentose-5-phosphate intermediate. In one embodiment, one or more co-product is co-produced with MEG (or glycolic acid). In another embodiment, the one or more pentose-5-phosphate intermediate is one or more of D-xylulose-5-phosphate, D-ribulose-5-phosphate or D-ribose-5-phosphate.

[00859] Therefore, in one embodiment, the application relates to a recombinant microorganism comprising one or more biochemical pathway comprising at least one enzyme having an activity that converts one or more hexose feedstock in a lossless conversion to one or more pentose-5-phosphate intermediate.

[00860] In some embodiments, the recombinant microorganism comprises expression of at least one enzyme having transketolase activity. In some embodiments, the enzyme

having transketolase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to tktA from *E. coli*. In other embodiments, the enzyme having transketolase activity is tktA from *E. coli*. In some embodiments, the enzyme having transketolase activity is encoded by an amino acid sequence having at least 70% sequence identity at least 80% sequence identity, or at least 90% sequence identity to tktB from *E. coli*. In other embodiments, the enzyme having transketolase activity is tktB from *E. coli*. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having transketolase activity comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 148 and 150. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having transketolase activity is tktA, or homolog thereof. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having transketolase activity is tktB, or homolog thereof. In a further embodiment, the one or more nucleic acid molecule encoding the enzyme having transketolase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 147 and 149.

[00861] In some embodiments, the recombinant microorganism comprises expression of at least one enzyme having transaldolase activity. In some embodiments, the enzyme having transaldolase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to talA or talB from *E. coli*. In some embodiments, the enzyme having transaldolase activity is talA from *E. coli*. In other embodiments, the enzyme having transaldolase activity is talB from *E. coli*. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having transaldolase activity comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 152 and 154. In a further embodiment, the one or more nucleic acid molecule encoding the enzyme having transaldolase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 151 and 153.

[00862] In some embodiments, the recombinant microorganism comprises expression of at least one enzyme having ribulose-5-phosphate 3-epimerase activity. In some embodiments, the enzyme having ribulose-5-phosphate 3-epimerase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to rpe from *E. coli*. In other embodiments, the enzyme having ribulose-5-phosphate 3-epimerase activity is rpe from *E. coli*. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having ribulose-5-phosphate 3-epimerase activity comprise an amino acid sequence set forth in SEQ ID NO: 158. In a further embodiment, the one or more nucleic acid molecule encoding the enzyme

having ribulose-5-phosphate 3-epimerase activity is encoded by a nucleic acid sequence set forth in SEQ ID NO: 157.

[00863] In some embodiments, the recombinant microorganism comprises expression of at least one enzyme having ribose-5-phosphate isomerase activity. In some embodiments, the enzyme having ribose-5-phosphate isomerase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to rpiA or rpiB from *E. coli*. In other embodiments, the enzyme having ribose-5-phosphate isomerase activity is rpiA or rpiB from *E. coli*. In another embodiment, the one or more nucleic acid molecules encoding the rpiA enzyme having ribose-5-phosphate isomerase activity comprise an amino acid sequence set forth in SEQ ID NO: 156. In a further embodiment, the one or more nucleic acid molecule encoding the rpiA enzyme having ribose-5-phosphate isomerase activity is encoded by a nucleic acid sequence set forth in SEQ ID NO: 155. In another embodiment, the one or more nucleic acid molecules encoding the rpiB enzyme having ribose-5-phosphate isomerase activity comprise an amino acid sequence set forth in SEQ ID NO: 253. In a further embodiment, the one or more nucleic acid molecule encoding the rpiB enzyme having ribose-5-phosphate isomerase activity is encoded by a nucleic acid sequence set forth in SEQ ID NO: 254.

[00864] In some embodiments, the recombinant microorganism comprises expression of at least one enzyme having a transketolase activity, a transaldolase activity, a ribulose-5-phosphate 3-epimerase activity, and a ribose-5-phosphate isomerase activity. In other embodiments, the recombinant microorganism further comprises a deleted or diminished activity in one or more endogenous enzymes selected from glyceraldehyde 3-phosphate dehydrogenase, phosphoglycerate kinase and phosphoglycerate mutase. In some embodiments, the endogenous glyceraldehyde 3-phosphate dehydrogenase enzyme is gapA, the phosphoglycerate kinase is pgk and the phosphoglycerate mutase is gpmA and/or gpmM.

[00865] In some embodiments, the recombinant microorganism comprises expression of at least one enzyme having fructose-6-phosphate phosphoketolase activity. In some embodiments, an enzyme having fructose-6-phosphate phosphoketolase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having fructose-6-phosphate phosphoketolase activity selected from the group consisting of *Bifidobacterium dentium* BDP_1006, *Bifidobacterium lactis* xfp, *Lactobacillus paraplantarum* xpkA and *Bifidobacterium breve* xfp. In a preferred embodiment, an enzyme having fructose-6-phosphate phosphoketolase activity is selected from the group consisting of *Bifidobacterium dentium* BDP_1006, *Bifidobacterium lactis* xfp, *Lactobacillus paraplantarum* xpkA and

Bifidobacterium breve xfp. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having fructose-6-phosphate phosphoketolase activity comprise an amino acid sequence selected from the group consisting of SEQ ID NOs: 212, 214, 216 and 218. In a further embodiment, the one or more nucleic acid molecule encoding the enzyme having fructose-6-phosphate phosphoketolase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 211, 213, 215 and 217.

[00866] In some embodiments, the recombinant microorganism comprises expression of at least one enzyme having phosphate acetyltransferase activity. In some embodiments, an enzyme having phosphate acetyltransferase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having phosphate acetyltransferase activity selected from *E. coli* pta and *Clostridium acetobutylicum* pta. In a preferred embodiment, an enzyme having phosphate acetyltransferase activity is selected from *E. coli* pta and *Clostridium acetobutylicum* pta. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having phosphate acetyltransferase activity comprise an amino acid sequence selected from SEQ ID NOs: 220 and 222. In a further embodiment, the one or more nucleic acid molecule encoding the enzyme having phosphate acetyltransferase activity is encoded by a nucleic acid sequence selected from SEQ ID NOs: 219 and 221.

[00867] In some embodiments, the recombinant microorganism comprises expression of at least one enzyme having a fructose-6-phosphate phosphoketolase activity, a phosphate acetyltransferase activity, a transketolase activity, a transaldolase activity, a ribulose-5-phosphate 3-epimerase activity, and a ribose-5-phosphate isomerase activity. In other embodiments, the recombinant microorganism further comprises a deleted or diminished activity in an endogenous 6-phosphofructokinase enzyme. In some embodiments, the endogenous 6-phosphofructokinase enzyme is pfkA and/or pfkB.

[00868] In another embodiment, the one or more pentose-5-phosphate intermediate produced in the lossless conversion of one or more hexose feedstock can be connected with any one of the known C2 MEG or glycolic acid production pathways by conversion of the one or more pentose-5-phosphate intermediate to one or more pentose and/or pentose-1-phosphate intermediate. In some embodiments, the recombinant microorganism comprises expression of at least one enzyme having a pentose phosphatase activity, an arabinol phosphate dehydrogenase activity, and/or a phosphopentomutase activity.

[00869] In some embodiments, the recombinant microorganism comprises expression of at least one enzyme having pentose phosphatase activity. In other embodiments, the at least one enzyme having pentose phosphatase activity is selected from one or more of an

enzyme having D-pentose-5-phosphatase activity, an enzyme having D-xylulose-5-phosphatase activity, an enzyme having D-ribose-5-phosphatase activity, and an enzyme having D-ribulose-5-phosphatase activity. In some embodiments, the pentose phosphatase is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to a D-pentose-5-phosphatase selected from the group consisting of *E. coli* phoA, *E. coli* yfbT and *E. coli* yidA. In some embodiments, the enzyme having D-xylulose-5-phosphatase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to *Bacillus subtilis* araL. In some embodiments, the enzyme having D-ribose-5-phosphatase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having D-ribose-5-phosphatase activity selected from the group consisting of *Arabidopsis thaliana* SGPP, *Pseudomonas fluorescens* PFLU_2693, and *E. coli* ybiV. In some embodiments, the enzyme having D-ribulose-5-phosphatase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to *Plasmodium falciparum* PF10_0325. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having D-pentose-5-phosphatase activity is selected from the group consisting of SEQ ID NOs: 159, 161, 163, 165, 167, 169, 171 and 173. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having D-pentose-5-phosphatase activity comprise an amino acid sequence selected from the group consisting of SEQ ID NOs: 160, 162, 164, 166, 168, 170, 172 and 174.

[00870] In some embodiments, the recombinant microorganism comprises expression of at least one enzyme having arabitol phosphate dehydrogenase activity. In some embodiments, the enzyme having arabitol phosphate dehydrogenase activity is selected from one or more of an enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity, an enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity, an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity and an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity.

[00871] In some embodiments, an enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to APDH from *Enterococcus avium*. In a preferred embodiment, an enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity is APDH from *Enterococcus avium*. In some embodiments, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity is set forth in SEQ ID NO: 175. In another

embodiment, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity comprise an amino acid sequence set forth in SEQ ID NO: 176.

[00872] In some embodiments, an enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to APDH from *Enterococcus avium*. In a preferred embodiment, an enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity is APDH from *Enterococcus avium*. In some embodiments, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity is set forth in SEQ ID NO: 175. In another embodiment, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity comprise an amino acid sequence set forth in SEQ ID NO: 176.

[00873] In some embodiments, an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity selected from the group consisting of *Candida albicans* ARD1, *Candida tropicalis* ARD1, *Scheffersomyces stipitis* ARDH, *Bacillus subtilis* egsA (araM), *Aeropyrum pernix* egsA, *E. coli* gpsA and *Saccharomyces cerevisiae* GPD1. In a preferred embodiment, an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity is selected from the group consisting of *Candida albicans* ARD1, *Candida tropicalis* ARD1, *Scheffersomyces stipitis* ARDH, *Bacillus subtilis* egsA (araM), *Aeropyrum pernix* egsA, *E. coli* gpsA and *Saccharomyces cerevisiae* GPD1. In some embodiments, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity is selected from the group consisting of SEQ ID NOs: 177, 179, 181, 189, 191, 193 and 195. In another embodiment, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity comprise an amino acid sequence selected from the group consisting of SEQ ID NOs: 178, 180, 182, 190, 192, 194 and 196.

[00874] In some embodiments, an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity selected from the group consisting of *Pseudomonas fluorescens* mtID, *Klebsiella pneumoniae* dalD, *Ralstonia solanacearum* dalD, *Bacillus subtilis* egsA (araM), *Aeropyrum pernix* egsA, *E. coli* gpsA and *Saccharomyces cerevisiae* GPD1. In a preferred embodiment, an enzyme having D-arabitol

5-phosphate 4-dehydrogenase activity is selected from the group consisting of *Pseudomonas fluorescens* mtlD, *Klebsiella pneumoniae* dalD, *Ralstonia solanacearum* dalD, *Bacillus subtilis* egsA (araM), *Aeropyrum pernix* egsA, *E. coli* gpsA and *Saccharomyces cerevisiae* GPD1. In some embodiments, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity is selected from the group consisting of SEQ ID NOs: 183, 185, 187, 189, 191, 193 and 195. In another embodiment, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity comprise an amino acid sequence selected from the group consisting of SEQ ID NOs: 184, 186, 188, 190, 192, 194 and 196.

[00875] In some embodiments, the recombinant microorganism comprises expression of at least one enzyme having phosphopentomutase activity. In some embodiments, an enzyme having phosphopentomutase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having phosphopentomutase activity selected from the group consisting of *E. coli* deoB, *E. coli* pgm, *Bacillus subtilis* pgcA, *Lactococcus lactis* pgmB, *E. coli* ycjU, *Pseudomonas aeruginosa* algC, and *E. coli* cpsG. In a preferred embodiment, an enzyme having phosphopentomutase activity is selected from the group consisting of *E. coli* deoB, *E. coli* pgm, *Bacillus subtilis* pgcA, *Lactococcus lactis* pgmB, *E. coli* ycjU, *Pseudomonas aeruginosa* algC, and *E. coli* cpsG. In some embodiments, the one or more nucleic acid molecules encoding an enzyme having phosphopentomutase activity is selected from the group consisting of SEQ ID NOs: 197, 199, 201, 203, 205, 207 and 209. In another embodiment, the one or more nucleic acid molecules encoding an enzyme having phosphopentomutase activity comprise an amino acid sequence selected from the group consisting of SEQ ID NOs: 198, 200, 202, 204, 206, 208 and 210.

[00876] In some embodiments of any of the recombinant microorganisms described above, the recombinant microorganism further comprises a diminished activity in one or more endogenous enzymes selected from glucose 6-phosphate-1-dehydrogenase, 6-phosphogluconolactonase, and 6-phosphogluconate dehydrogenase. In further embodiments, the glucose 6-phosphate-1-dehydrogenase is zwf, the 6-phosphogluconolactonase is pgl, and the 6-phosphogluconate dehydrogenase is gnd.

[00877] In some embodiments, the application provides for a recombinant microorganism that co-produces MEG (or glycolic acid) and one or more co-product selected from acetone, isopropanol, propene, isobutene and one or more serine pathway compounds. In some preferred embodiments, the one or more serine pathway compounds is selected from serine, glycine, monoethanolamine (MEA) and ethylenediamine (EDA).

[00878] In some embodiments, the one or more hexose feedstock is selected from glucose or oligomers of glucose thereof. In other embodiments, the oligomers of glucose are selected from fructose, sucrose, starch, cellobiose, maltose, lactose and cellulose.

[00879] In some embodiments, the expression of an enzyme having transketolase activity or an enzyme having fructose-6-phosphate phosphoketolase activity in the recombinant microorganism enables a lossless conversion of one or more hexose feedstock to one or more pentose-5-phosphate intermediate.

[00880] In some embodiments, the recombinant microorganism produces MEG or glycolic acid (GA) through the conversion of glycolaldehyde in a C2 pathway and through the conversion of dihydroxyacetone phosphate (DHAP) or pyruvate in a C3 pathway. In other embodiments, the glycolaldehyde is oxidized to glycolic acid by a glycolaldehyde dehydrogenase.

[00881] In some embodiments, the at least one enzyme for the production of MEG or GA through the conversion of DHAP or pyruvate in a C3 pathway are selected from at least one enzyme having an activity selected from a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine transaminase activity, a hydroxypyruvate decarboxylase activity, a 3-phosphohydroxypyruvate reductase activity, a glycolaldehyde reductase activity, a glycolaldehyde dehydrogenase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a serine decarboxylase activity, an ethanolamine aminotransferase or ethanolamine oxidoreductase (deaminating) activity, a glycerate decarboxylase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, and a glycerate 2-kinase activity.

[00882] In some embodiments, the recombinant microorganism produces MEG or glycolic acid (GA) through the conversion of glycolaldehyde in a C2 pathway and produces one or more co-product through the conversion of dihydroxyacetone phosphate (DHAP) or pyruvate in a C3 pathway. In other embodiments, the one or more co-product is selected from acetone, isopropanol, propene, isobutene and one or more serine pathway compounds. In some preferred embodiments, the one or more serine pathway compounds is selected from serine, glycine, monoethanolamine (MEA) and ethylenediamine (EDA).

[00883] In some embodiments, the at least one enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from at least one enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-

CoA hydrolase activity, and an acetoacetate decarboxylase activity, and the one or more co-product comprises acetone.

[00884] In some embodiments, the at least one enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from at least one enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, and a secondary alcohol dehydrogenase activity, and the one or more co-product comprises isopropanol.

[00885] In some embodiments, the at least one enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from at least one enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, a secondary alcohol dehydrogenase activity, and a dehydratase activity, and the one or more co-product comprises propene.

[00886] In some embodiments, the at least one enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from at least one enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, a 3-hydroxyisovalerate (3HIV) synthase activity, a hydroxymethylglutaryl-CoA synthase activity, a methylglutaconyl-CoA hydratase activity, a methylcrotonyl-CoA carboxylase activity, a methylcrotonyl-CoA hydratase activity, a 3-hydroxyisovaleryl-CoA thioesterase activity, a 3HIV kinase activity, a 3HIV-3-phosphate decarboxylase activity, and a 3HIV decarboxylase activity, and the one or more co-product comprises isobutene.

[00887] In some embodiments, the at least one enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from at least one enzyme having an activity selected from a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, and a glycerate 2-kinase activity, and wherein the one or more co-product comprises L-serine.

[00888] In some embodiments, the at least one enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from at least one enzyme having an activity selected from a serine hydroxymethyltransferase activity, a transferase activity, a formaldehyde dehydrogenase activity, a formate dehydrogenase activity, an activity associated with glycine cleavage system, a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine transaminase activity, a hydroxypyruvate decarboxylase activity, a serine oxidoreductase (deaminating) activity, a serine decarboxylase activity, an ethanolamine aminotransferase or ethanolamine oxidoreductase (deaminating) activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, a glycerate 2-kinase activity, a glycolaldehyde dehydrogenase activity, a glycolate dehydrogenase activity, an alanine-glyoxylate aminotransferase activity, an alanine transaminase activity, an NAD(P)H dependent glutamate dehydrogenase activity, and wherein the one or more co-product comprises glycine. In another embodiment, the activity associated with glycine cleavage system comprise an enzyme or protein selected from a glycine decarboxylase (P protein), an aminomethyltransferase (T protein), a dihydrolipoamide dehydrogenase (L protein), and an H protein.

[00889] In some embodiments, the at least one enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from at least one enzyme having an activity selected from a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a transaminase activity, a hydroxypyruvate decarboxylase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a serine decarboxylase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, a glycerate 2-kinase activity, an acetaldehyde dehydrogenase activity, and an ethanolamine ammonia lyase activity, and the one or more co-product comprises monoethanolamine (MEA).

[00890] In some embodiments, the at least one enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from at least one enzyme having an activity selected from a serine dehydrogenase activity, a 2-aminomalonate semialdehyde decarboxylase activity, an aminoacetaldehyde transaminase activity, a 2-aminomalonate semialdehyde transaminase activity, a 2,3-diaminopropanoate decarboxylase activity, a serine decarboxylase activity, an ethanolamine dehydrogenase

activity, a serine hydroxymethyltransferase activity, an aldehyde oxidase activity, an N-acetyl transferase or O-acetyl transferase activity, an N-acetylserine dehydrogenase activity, a transaminase activity, a deacetylase activity, a serine aminase activity, and a 2,3-diaminopropanoate ammonia lyase activity, and the one or more co-product comprises ethylenediamine (EDA).

[00891] In some embodiments, the at least one enzyme for the production of MEG or GA from glycolaldehyde in a C2 pathway are selected from at least one enzyme having an activity selected from a D-tagatose 3-epimerase activity, a D-ribulokinase activity, a D-ribulose-1-phosphate aldolase activity, a D-xylulose 1-kinase activity, a D-xylulose-1-phosphate aldolase activity, a xylose reductase or aldose reductase activity, a xylitol dehydrogenase activity, a xylose isomerase activity, a xylose dehydrogenase activity, a xylonolactonase activity, a xylonate dehydrogenase activity, a 2-keto-3-deoxy-D-pentonate aldolase activity, a glycolaldehyde reductase activity and a glycolaldehyde dehydrogenase activity.

[00892] In some embodiments of any of the recombinant microorganisms described above, the recombinant microorganism further comprises one or more modifications to diminish or delete activity in a glycolaldehyde dehydrogenase, a lactate dehydrogenase, a xylose isomerase, a xylulose kinase or combination thereof.

[00893] In one embodiment, at least a portion of the excess NADH produced in the C3 pathway is used as a source of reducing equivalents in the C2 pathway. In another embodiment, at least a portion of the excess NADH produced in the C3 pathway is used to produce ATP.

[00894] In one embodiment, excess biomass formation is minimized and production of MEG (or glycolic acid) or MEG (or glycolic acid) and one or more co-products is maximized.

[00895] *Hexose to pentose-5-phosphate intermediate*

[00896] In the present disclosure, glucose flux is funneled into the pentose phosphate pathway instead of the glycolysis pathway by using a non-oxidative entry into the pentose phosphate pathway.

[00897] **[A]** Therefore, in one embodiment, the application relates to a recombinant microorganism capable of producing one or more pentose-5-phosphate intermediate from one or more hexose feedstock, wherein the recombinant microorganism expresses one or more of the following from (a) to (d):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having transketolase activity that catalyzes a reversible conversion of fructose-6-

phosphate and glyceraldehyde-3-phosphate to erythrose-4-phosphate and D-xylulose-5-phosphate, respectively, and/or that catalyzes a reversible conversion of glyceraldehyde-3-phosphate from (b) and seduheptulose-7-phosphate from (b) to D-ribose-5-phosphate and D-xylulose-5-phosphate, respectively;

5 (b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having transaldolase activity that catalyzes a reversible conversion of fructose-6-phosphate and erythrose-4-phosphate from (a) to glyceraldehyde-3-phosphate and seduheptulose-7-phosphate, respectively;

10 (c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having ribose-5-phosphate isomerase activity that catalyzes an interconversion of D-ribose-5-phosphate from (a) and D-ribulose-5-phosphate;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having ribulose-5-phosphate 3-epimerase activity that catalyzes an interconversion of D-xylulose-5-phosphate from (a) and D-ribulose-5-phosphate from (c);

15 wherein the recombinant microorganism optionally further comprises a deletion, insertion, or loss of function mutation in a gene encoding a glyceraldehyde 3-phosphate dehydrogenase;

wherein the one or more hexose feedstock is converted to fructose-6-phosphate and glyceraldehyde-3-phosphate through an endogenous glycolysis pathway in the recombinant microorganism,

20 and wherein the one or more pentose-5-phosphate intermediate produced is one or more of D-ribose-5-phosphate, D-xylulose-5-phosphate or D-ribulose-5-phosphate.

[00898] [B] In another embodiment, the application relates to a recombinant microorganism capable of producing one or more pentose-5-phosphate intermediate from one or more hexose feedstock, wherein the recombinant microorganism expresses one or
25 more of the following from (a) to (f):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having fructose-6-phosphate phosphoketolase activity that catalyzes a reversible conversion of fructose-6-phosphate to erythrose-4-phosphate and acetyl-phosphate;

30 (b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having phosphate acetyltransferase activity that catalyzes a reversible conversion of acetyl-phosphate from (a) to acetyl-CoA;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having transaldolase activity that catalyzes a reversible conversion of fructose-6-

phosphate and erythrose-4-phosphate from (a) to glyceraldehyde-3-phosphate and seduheptulose-7-phosphate, respectively;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having transketolase activity that catalyzes a reversible conversion of glyceraldehyde-3-phosphate from (c) and seduheptulose-7-phosphate from (c) to D-ribose-5-phosphate and D-xylulose-5-phosphate, respectively;

(e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having ribose-5-phosphate isomerase activity that catalyzes an interconversion of D-ribose-5-phosphate from (d) and D-ribulose-5-phosphate;

(f) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having ribulose-5-phosphate 3-epimerase activity that catalyzes an interconversion of D-xylulose-5-phosphate from (d) and D-ribulose-5-phosphate from (e);

wherein the recombinant microorganism optionally further comprises a deletion, insertion, or loss of function mutation in a gene encoding a 6-phosphofructokinase;

wherein the one or more hexose feedstock is converted to fructose-6-phosphate through an endogenous glycolysis pathway in the recombinant microorganism,

wherein the acetyl-CoA produced in step (b) can be used to produce one or more co-products selected from acetone, isopropanol, propene, isobutene, and serine pathway compounds;

and wherein the one or more pentose-5-phosphate intermediate produced is one or more of D-ribose-5-phosphate, D-xylulose-5-phosphate or D-ribulose-5-phosphate.

[00899] In some embodiments, the oxidative branch of the pentose phosphate pathway is deleted or inactivated to optimize flux towards the non-oxidative entry into the pentose phosphate pathway.

[00900] [C] Therefore, in one embodiment, the recombinant microorganism of embodiment **[A]** or embodiment **[B]** optionally further comprises one or more modifications selected from the group consisting of:

(i) a deletion, insertion, or loss of function mutation in a gene encoding a glucose 6-phosphate-1-dehydrogenase that catalyzes the conversion of glucose-6-phosphate to 6-phospho-D-glucono-1,5-lactone;

(ii) a deletion, insertion, or loss of function mutation in a gene encoding a 6-phosphogluconolactonase that catalyzes the conversion of 6-phospho-D-glucono-1,5-lactone to gluconate-6-phosphate; and

(iii) a deletion, insertion, or loss of function mutation in a gene encoding a 6-phosphogluconate dehydrogenase that catalyzes the conversion of gluconate-6-phosphate to D-ribulose-5-phosphate.

[00901] In some embodiments, the enzyme having transketolase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *E. coli* tktA. In other embodiments, the transketolase is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *E. coli* tktB. In some embodiments, the transketolase is *E. coli* tktA. In other
10 embodiments, the transketolase is *E. coli* tktB.

[00902] In some embodiments, the one or more nucleic acid molecules encoding the enzyme having transketolase activity is tktA, or homolog thereof. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having transketolase activity is tktB, or homolog thereof. In another embodiment, the one or more nucleic acid molecules
15 encoding the enzyme having transketolase activity comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 148 and 150. In a further embodiment, the one or more nucleic acid molecule encoding the enzyme having transketolase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 147 and 149.

[00903] In some embodiments, the enzyme having transaldolase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to talA or talB from *E. coli*. In some embodiments, the enzyme having transaldolase activity is talA from *E. coli*. In other embodiments, the enzyme having transaldolase activity is talB from *E. coli*. In another embodiment, the one or
20 more nucleic acid molecules encoding the enzyme having transaldolase activity comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 152 and 154. In a further embodiment, the one or more nucleic acid molecule encoding the enzyme having transaldolase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 151 and 153.

[00904] In one embodiment, the enzyme having ribose-5-phosphate isomerase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *E. coli* rpiA or rpiB. In some embodiments, the enzyme having ribose-5-phosphate isomerase activity is *E. coli* rpiA or rpiB.
30

[00905] In some embodiments, the one or more nucleic acid molecules encoding the enzyme having ribose-5-phosphate isomerase activity is *rpiA*, or homolog thereof. In another embodiment, the one or more nucleic acid molecules encoding an enzyme having ribose-5-phosphate isomerase activity comprise an amino acid sequence set forth in SEQ ID NO: 156. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having ribose-5-phosphate isomerase activity is encoded by a nucleic acid sequence set forth in SEQ ID NO: 155.

[00906] In one embodiment, the enzyme having ribulose-5-phosphate 3-epimerase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *E. coli* *rpe*. In some embodiments, the enzyme having ribulose-5-phosphate 3-epimerase activity is *E. coli* *rpe*.

[00907] In some embodiments, the one or more nucleic acid molecules encoding the enzyme having ribulose-5-phosphate 3-epimerase activity is *rpe*, or homolog thereof. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having ribulose-5-phosphate 3-epimerase activity comprise an amino acid sequence set forth in SEQ ID NO: 158. In a further embodiment, the one or more nucleic acid molecule encoding the enzyme having ribulose-5-phosphate 3-epimerase activity is encoded by a nucleic acid sequence set forth in SEQ ID NO: 157.

[00908] In one embodiment, the enzyme having fructose-6-phosphate phosphoketolase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having fructose-6-phosphate phosphoketolase activity selected from the group consisting of *Bifidobacterium dentium* BDP_1006, *Bifidobacterium lactis* xfp, *Lactobacillus paraplantarum* xpkA and *Bifidobacterium breve* xfp. In other embodiments, the enzyme having fructose-6-phosphate phosphoketolase activity is selected from the group consisting of *Bifidobacterium dentium* BDP_1006, *Bifidobacterium lactis* xfp, *Lactobacillus paraplantarum* xpkA and *Bifidobacterium breve* xfp.

[00909] In some embodiments, the one or more nucleic acid molecules encoding the enzyme having fructose-6-phosphate phosphoketolase activity is selected from BDP_1006, xfp, xpkA, or homolog thereof. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having fructose-6-phosphate phosphoketolase activity comprise an amino acid sequence selected from the group consisting of SEQ ID NOs: 212, 214, 216 and 218. In a further embodiment, the one or more nucleic acid molecule encoding

the enzyme having fructose-6-phosphate phosphoketolase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 211, 213, 215 and 217.

[00910] In one embodiment, the enzyme having phosphate acetyltransferase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having phosphate acetyltransferase activity selected from *E. coli* pta and *Clostridium acetobutylicum* pta. In other embodiments, the enzyme having phosphate acetyltransferase activity is selected from *E. coli* pta and *Clostridium acetobutylicum* pta.

[00911] In some embodiments, the one or more nucleic acid molecules encoding the enzyme having phosphate acetyltransferase activity is pta, or homolog thereof. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having phosphate acetyltransferase activity comprise an amino acid sequence selected from SEQ ID NOs: 220 and 222. In a further embodiment, the one or more nucleic acid molecule encoding the enzyme having phosphate acetyltransferase activity is encoded by a nucleic acid sequence selected from SEQ ID NOs: 219 and 221.

[00912] *Connecting pentose phosphate pathway and MEG (or glycolic acid) production pathways*

[00913] In another aspect, a pentose-5-phosphate intermediate produced in embodiment [A] or embodiment [B] (and optionally comprising embodiment [C]) can be connected with any one of the known MEG or glycolic acid production pathways by pentose phosphatases.

[00914] [D] Therefore, in one embodiment, the application relates to a recombinant microorganism capable of producing D-xylulose from D-xylulose-5-phosphate, wherein the recombinant microorganism expresses one or more enzyme having pentose-5-phosphatase activity, wherein the D-xylulose-5-phosphate is produced by non-oxidative branch of pentose phosphate pathway, and wherein the D-xylulose can be used to produce MEG (or glycolic acid) and optionally, one or more co-product.

[00915] In one embodiment, the enzyme having pentose-5-phosphatase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having pentose-5-phosphatase activity selected from the group consisting of *E. coli* phoA, *E. coli* yfbT and *E. coli* yidA. In further embodiments, the enzyme having pentose-5-phosphatase activity is selected from *E. coli* phoA, *E. coli* yfbT and *E. coli* yidA.

[00916] In some embodiments, the one or more nucleic acid molecules encoding the enzyme having pentose-5-phosphatase activity is phoA, yfbT, yidA or homolog thereof. In

another embodiment, the one or more nucleic acid molecules encoding the enzyme having pentose-5-phosphatase activity comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 160, 168 and 172. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having pentose-5-phosphatase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 159, 167 and 171.

[00917] In another embodiment, the pentose-5-phosphatase is a D-xylulose-5-phosphatase. In another embodiment, the D-xylulose-5-phosphatase is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to *Bacillus subtilis* araL. In further embodiments, the D-xylulose-5-phosphatase is *Bacillus subtilis* araL.

[00918] In some embodiments, the one or more nucleic acid molecules encoding the enzyme having D-xylulose-5-phosphatase activity is araL, or homolog thereof. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having D-xylulose-5-phosphatase activity comprises an amino acid sequence set forth in SEQ ID NO: 164. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having D-xylulose-5-phosphatase activity is encoded by a nucleic acid sequence set forth in SEQ ID NO: 163.

[00919] [E] In one embodiment, the application relates to a recombinant microorganism capable of producing D-ribulose from D-ribulose-5-phosphate or from D-ribose-5-phosphate, wherein the recombinant microorganism expresses one or more of the following from (a) to (c):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having D-ribulose-5-phosphatase activity and/or an enzyme having pentose-5-phosphatase activity that catalyzes a reversible conversion of D-ribulose-5-phosphate to D-ribulose;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having D-ribose-5-phosphatase activity and/or an enzyme having pentose-5-phosphatase activity that catalyzes a reversible conversion of D-ribose-5-phosphate to D-ribose;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having D-ribose isomerase activity that catalyzes a reversible conversion of D-ribose from (b) to D-ribulose;

wherein the D-ribulose-5-phosphate and/or D-ribose-5-phosphate is produced by non-oxidative branch of pentose phosphate pathway, and wherein the D-ribulose can be used to produce MEG (or glycolic acid) and optionally, one or more co-product.

[00920] In one embodiment, the enzyme having pentose-5-phosphatase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having pentose-5-phosphatase activity selected from the group consisting of *E. coli* phoA, *E. coli* yfbT and *E. coli* yidA. In further embodiments, the enzyme having pentose-5-phosphatase activity is selected from *E. coli* phoA, *E. coli* yfbT and *E. coli* yidA.

[00921] In some embodiments, the one or more nucleic acid molecules encoding the enzyme having pentose-5-phosphatase activity is phoA, yfbT, yidA or homolog thereof. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having pentose-5-phosphatase activity comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 160, 168 and 172. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having pentose-5-phosphatase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 159, 167 and 171.

[00922] In another embodiment, the enzyme having D-ribulose-5-phosphatase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to *Plasmodium falciparum* PF10_0325. In further embodiments, the enzyme having D-ribulose-5-phosphatase activity is *Plasmodium falciparum* PF10_0325.

[00923] In some embodiments, the one or more nucleic acid molecules encoding the enzyme having D-ribulose-5-phosphatase activity is PF10_0325, or homolog thereof. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having D-ribulose-5-phosphatase activity comprises an amino acid sequence set forth in SEQ ID NO: 174. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having D-ribulose-5-phosphatase activity comprises a nucleic acid sequence set forth in SEQ ID NO: 173.

[00924] In one embodiment, the enzyme having D-ribose-5-phosphatase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having D-ribose-5-phosphatase activity selected from the group consisting of *Arabidopsis thaliana* SGPP, *Pseudomonas fluorescens* PFLU_2693, and *E. coli* ybiV. In some embodiments, the enzyme

having D-ribose-5-phosphatase activity is selected from the group consisting of *Arabidopsis thaliana* SGPP, *Pseudomonas fluorescens* PFLU_2693, and *E. coli* ybiV.

[00925] In some embodiments, the one or more nucleic acid molecules encoding the enzyme having D-ribose-5-phosphatase activity is selected from SGPP, PFLU_2693, ybiV, or homolog thereof. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having D-ribose-5-phosphatase activity comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 162, 166 and 170. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having D-ribose-5-phosphatase activity comprises a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 161, 165 and 169.

[00926] In another aspect, a pentose-5-phosphate intermediate produced in embodiment **[A]** or embodiment **[B]** (and optionally comprising embodiment **[C]**) can be connected with any one of the known MEG or glycolic acid production pathways by arabitol phosphate dehydrogenases.

[00927] **[F]** In one embodiment, the application relates to a recombinant microorganism capable of producing D-ribulose-1-phosphate from D-xylulose-5-phosphate or producing D-xylulose-1-phosphate from D-ribulose-5-phosphate, wherein the recombinant microorganism expresses one or more arabitol phosphate dehydrogenase, wherein the one or more arabitol phosphate dehydrogenase catalyzes one or more of the following from (a) to (d):

- (a) a reversible conversion of D-xylulose-5-phosphate to D-arabitol-1-phosphate;
- (b) a reversible conversion of D-arabitol-1-phosphate from (a) to D-ribulose-1-phosphate;
- (c) a reversible conversion of D-ribulose-5-phosphate to D-arabitol-5-phosphate;
- (d) a reversible conversion of D-arabitol-5-phosphate from (c) to D-xylulose-1-phosphate,

wherein the D-xylulose-5-phosphate and/or D-ribulose-5-phosphate is produced by non-oxidative branch of pentose phosphate pathway, and wherein the D-ribulose-1-phosphate and/or D-xylulose-1-phosphate can be used to produce MEG (or glycolic acid) and optionally, one or more co-product.

[00928] In some embodiments, the enzyme having arabitol phosphate dehydrogenase activity is selected from one or more of an enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity, an enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity, an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity and an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity.

[00929] In some embodiments, an enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to APDH from *Enterococcus avium*. In a preferred embodiment, an enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity is APDH from *Enterococcus avium*. In some embodiments, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity is set forth in SEQ ID NO: 175. In another embodiment, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity comprise an amino acid sequence set forth in SEQ ID NO: 176.

[00930] In some embodiments, an enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to APDH from *Enterococcus avium*. In a preferred embodiment, an enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity is APDH from *Enterococcus avium*. In some embodiments, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity is set forth in SEQ ID NO: 175. In another embodiment, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity comprise an amino acid sequence set forth in SEQ ID NO: 176.

[00931] In some embodiments, an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity selected from the group consisting of *Candida albicans* ARD1, *Candida tropicalis* ARD1, *Scheffersomyces stipitis* ARDH, *Bacillus subtilis* egsA (araM), *Aeropyrum pernix* egsA, *E. coli* gpsA and *Saccharomyces cerevisiae* GPD1. In a preferred embodiment, an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity is selected from the group consisting of *Candida albicans* ARD1, *Candida tropicalis* ARD1, *Scheffersomyces stipitis* ARDH, *Bacillus subtilis* egsA (araM), *Aeropyrum pernix* egsA, *E. coli* gpsA and *Saccharomyces cerevisiae* GPD1. In some embodiments, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity is selected from the group consisting of SEQ ID NOs: 177, 179, 181, 189, 191, 193 and 195. In another embodiment, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity comprise an amino acid sequence selected from the group consisting of SEQ ID NOs: 178, 180, 182, 190, 192, 194 and 196.

[00932] In some embodiments, an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity selected from the group consisting of *Pseudomonas fluorescens* mtlD, *Klebsiella pneumoniae* dalD, *Ralstonia solanacearum* dalD, *Bacillus subtilis* egsA (araM), *Aeropyrum pernix* egsA, *E. coli* gpsA and *Saccharomyces cerevisiae* GPD1. In a preferred embodiment, an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity is selected from the group consisting of *Pseudomonas fluorescens* mtlD, *Klebsiella pneumoniae* dalD, *Ralstonia solanacearum* dalD, *Bacillus subtilis* egsA (araM), *Aeropyrum pernix* egsA, *E. coli* gpsA and *Saccharomyces cerevisiae* GPD1. In some embodiments, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity is selected from the group consisting of SEQ ID NOs: 183, 185, 187, 189, 191, 193 and 195. In another embodiment, the one or more nucleic acid molecules encoding an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity comprise an amino acid sequence selected from the group consisting of SEQ ID NOs: 184, 186, 188, 190, 192, 194 and 196.

[00933] In another aspect, a pentose-5-phosphate intermediate produced in embodiment [A] or embodiment [B] (and optionally comprising [C]) can be connected with any one of the known MEG or glycolic acid production pathways by phosphopentomutases.

[00934] [G] In one embodiment, the application relates to a recombinant microorganism capable of producing D-xylulose-1-phosphate from D-xylulose-5-phosphate or producing D-ribulose-1-phosphate from D-ribulose-5-phosphate, wherein the recombinant microorganism expresses one or more phosphopentomutases, wherein the D-xylulose-5-phosphate and/or D-ribulose-5-phosphate is produced by non-oxidative branch of pentose phosphate pathway, and wherein the D-ribulose-1-phosphate and/or D-xylulose-1-phosphate can be used to produce MEG (or glycolic acid) and optionally, one or more co-product.

[00935] In some embodiments, an enzyme having phosphopentomutase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having phosphopentomutase activity selected from the group consisting of *E. coli* deoB, *E. coli* pgm, *Bacillus subtilis* pgcA, *Lactococcus lactis* pgmB, *E. coli* ycjU, *Pseudomonas aeruginosa* algC, and *E. coli* cpsG. In a preferred embodiment, an enzyme having phosphopentomutase activity is selected from the group consisting of *E. coli* deoB, *E. coli* pgm, *Bacillus subtilis* pgcA, *Lactococcus lactis* pgmB, *E. coli* ycjU, *Pseudomonas aeruginosa* algC, and *E. coli* cpsG. In some embodiments, the one or more nucleic acid molecules encoding an enzyme

having phosphopentomutase activity is selected from the group consisting of SEQ ID NOs: 197, 199, 201, 203, 205, 207 and 209. In another embodiment, the one or more nucleic acid molecules encoding an enzyme having phosphopentomutase activity comprise an amino acid sequence selected from the group consisting of SEQ ID NOs: 198, 200, 202, 204, 206, 208 and 210.

[00936] *MEG (or glycolic acid), or optionally MEG (or glycolic acid) and co-product production pathways*

[00937] In some embodiments, the pentose or pentose-1-phosphate intermediates produced in embodiments **[D]**, **[E]**, **[F]**, and **[G]** are used in known MEG (or glycolic acid) C2 production pathways, which are coupled to C3 pathways, as described below, to co-produce additional MEG (or glycolic acid) and/or one or more co-products.

[00938] In some embodiments, MEG (or glycolic acid) is produced via a C2 pathway that uses D-xylulose-1-phosphate.

[00939] **[H]** In one embodiment, the application relates to a recombinant microorganism capable of producing monoethylene glycol (MEG) (or glycolic acid) from one or more hexose feedstock, wherein the recombinant microorganism from embodiment **[A]** or from embodiment **[B]** (and optionally comprising embodiment **[C]**), and comprising additionally one or more of embodiments **[D]**, **[F]**, and **[G]**, further expresses one or more of the following from (a) to (d):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having D-xylulose 1-kinase activity that catalyzes the conversion of D-xylulose from embodiment **[D]** to D-xylulose-1-phosphate;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having D-xylulose-1-phosphate aldolase activity that catalyzes the conversion of D-xylulose-1-phosphate from (a), from embodiment **[F]** and/or from embodiment **[G]**, to glycolaldehyde and dihydroxyacetonephosphate (DHAP);

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde reductase activity or an enzyme having aldehyde reductase activity that catalyzes the conversion of glycolaldehyde from (b) to MEG;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde dehydrogenase activity that catalyzes the conversion of glycolaldehyde from (b) to glycolic acid;

wherein MEG (or glycolic acid) and DHAP are produced.

[00940] In one embodiment, the enzyme having D-xylulose 1-kinase activity is encoded by one or more nucleic acid molecules obtained from *Homo sapiens*. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having D-xylulose 1-kinase activity is ketohexokinase C (khk-C), or homolog thereof. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having D-xylulose 1-kinase activity comprises an amino acid sequence set forth in SEQ ID NO: 55. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having D-xylulose 1-kinase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 53 and 54.

[00941] In one embodiment, the enzyme having D-xylulose-1-phosphate aldolase activity is encoded by one or more nucleic acid molecules obtained from *Homo sapiens*. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having D-xylulose-1-phosphate aldolase activity is aldolase B (aldoB), or homolog thereof. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having D-xylulose-1-phosphate aldolase activity comprises an amino acid sequence set forth in SEQ ID NO: 58. In some embodiments, the one or more nucleic acid molecule encoding the enzyme having D-xylulose-1-phosphate aldolase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 56 and 57.

[00942] In some embodiments, MEG (or glycolic acid) is produced via a C2 pathway that uses D-ribulose-1-phosphate.

[00943] [I] In one embodiment, the application relates to a recombinant microorganism capable of producing monoethylene glycol (MEG) (or glycolic acid) from one or more hexose feedstock, wherein the recombinant microorganism from embodiment **[A]** or from embodiment **[B]** (and optionally comprising embodiment **[C]**), and comprising additionally one or more of embodiments **[D]**, **[E]**, **[F]**, and **[G]**, further expresses one or more of the following from (a) to (e):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having D-tagatose 3-epimerase activity that catalyzes the conversion of D-xylulose from embodiment **[D]** to D-ribulose;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having D-ribulokinase activity that catalyzes the conversion of D-ribulose from (a) and/or from embodiment **[E]** to D-ribulose-1-phosphate;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having D-ribulose-1-phosphate aldolase activity that catalyzes the conversion of D-

ribulose-1-phosphate from (b), from embodiment [F] and/or from embodiment [G] to glycolaldehyde and dihydroxyacetonephosphate (DHAP);

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde reductase activity that catalyzes the conversion of glycolaldehyde from (c) to MEG;

(e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde dehydrogenase activity that catalyzes the conversion of glycolaldehyde from (c) to glycolic acid;

wherein MEG (or glycolic acid) and DHAP are produced.

[00944] In one embodiment, the enzyme having D-tagatose 3-epimerase activity is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Pseudomonas sp.*, *Mesorhizobium sp.* and *Rhodobacter sp.* In some embodiments, the enzyme having D-tagatose 3-epimerase activity is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Pseudomonas cichorii*, *Pseudomonas sp.* ST-24, *Mesorhizobium loti* and *Rhodobacter sphaeroides*. In some embodiments, the one or more nucleic acid molecules is dte and/or FJ851309.1, or homolog thereof. In a further embodiment, the enzyme having D-tagatose 3-epimerase activity comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 3 and 5. In yet a further embodiment, the enzyme having D-tagatose 3-epimerase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 1, 2 and 4.

[00945] In one embodiment, the enzyme having D-ribulokinase activity is encoded by one or more nucleic acid molecules obtained from *E. coli*. In some embodiments, the one or more nucleic acid molecules is fucK, or homolog thereof. In a further embodiment, the enzyme having D-ribulokinase activity comprises an amino acid sequence set forth in SEQ ID NO: 8. In yet a further embodiment, the enzyme having D-ribulokinase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 6, 7, and 257.

[00946] In one embodiment, the enzyme having D-ribulose-1-phosphate aldolase activity is encoded by one or more nucleic acid molecules obtained from *E. coli*. In some embodiments, the one or more nucleic acid molecules is fucA, or homolog thereof. In a further embodiment, the enzyme having D-ribulose-1-phosphate aldolase activity comprises an amino acid sequence set forth in SEQ ID NO: 11. In yet a further embodiment, the

enzyme having D-ribulose-1-phosphate aldolase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 9 and 10.

[00947] In some embodiments, MEG (or glycolic acid) is produced via a C2 pathway that uses D-xylonate.

5 **[00948]** **[J]** In one embodiment, the application relates to a recombinant microorganism capable of producing monoethylene glycol (MEG) (or glycolic acid) from one or more hexose feedstock, wherein the recombinant microorganism from embodiment **[A]** or from embodiment **[B]** (and optionally comprising embodiment **[C]**), and comprising additionally embodiment **[D]**, further expresses one or more of the following from (a) to (d):

10 (a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having xylose isomerase activity that catalyzes the conversion of D-xylulose from embodiment **[D]** to D-xylose;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having xylose dehydrogenase activity that catalyzes the conversion of D-xylose
15 from (a) to D-xylonolactone;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having xylonolactonase activity that catalyzes the conversion of D-xylonolactone from (b) to D-xylonate;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an
20 enzyme having xylose dehydrogenase activity that catalyzes the conversion of D-xylose from (a) to D-xylonate;

wherein the recombinant microorganism further expresses one or more of the following from (e) to (h):

(e) at least one endogenous or exogenous nucleic acid molecule encoding an
25 enzyme having xylonate dehydratase activity that catalyzes the conversion of D-xylonate from (c) or (d) to 2-keto-3-deoxy-xylonate;

(f) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 2-keto-3-deoxy-D-pentonate aldolase activity that catalyzes the conversion of 2-keto-3-deoxy-xylonate from (e) to glycolaldehyde and pyruvate;

30 (g) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde reductase activity that catalyzes the conversion of glycolaldehyde from (f) to MEG;

(h) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde dehydrogenase activity that catalyzes the conversion of glycolaldehyde from (f) to glycolic acid;

wherein MEG (or glycolic acid) and pyruvate are produced.

5 **[00949]** In some embodiments, the recombinant microorganism comprises an endogenous or exogenous enzyme having xylose isomerase activity that catalyzes the conversion of D-xylulose to D-xylose. In one embodiment, the enzyme having xylose isomerase activity is exogenous. In another embodiment, the enzyme having xylose isomerase activity is encoded by one or more nucleic acid molecules obtained from
 10 *Pyromyces sp.* In a further embodiment, the enzyme having xylose isomerase activity is encoded by one or more nucleic acid molecules obtained from *E. coli*. In another embodiment, the one or more nucleic acid molecules encoding the enzyme having xylose isomerase activity is xylA, or homolog thereof. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having xylose isomerase activity comprises an
 15 amino acid sequence selected from SEQ ID NOs: 95 and 144. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having xylose isomerase activity comprises a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 93, 94 and 143.

[00950] In one embodiment, the enzyme having xylose dehydrogenase activity is
 20 encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Caulobacter sp.*, *Haloarcula sp.*, *Haloferax sp.*, *Halorubrum sp.* and *Trichoderma sp.* In another embodiment, the enzyme having xylose dehydrogenase activity is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Caulobacter crescentus*, *Haloarcula marismortui*,
 25 *Haloferax volcanii*, *Halorubrum lacusprofundi* and *Trichoderma reesei*. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having xylose dehydrogenase activity is selected from xylB, xdh1 (HVO_B0028) and/or xyd1, or homolog thereof. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having xylose dehydrogenase activity comprises an amino acid sequence selected
 30 from the group consisting of SEQ ID NOs: 61, 63 and 65. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having xylose dehydrogenase activity comprises a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 59, 60, 62 and 64.

[00951] In one embodiment, the enzyme having xylonolactonase activity is encoded by
 35 one or more nucleic acid molecules obtained from a microorganism selected from

Caulobacter sp. and *Haloferax sp.* In another embodiment, the enzyme having xylonolactonase activity is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Caulobacter crescentus*, *Haloferax volcanii* and *Haloferax gibbonsii*. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having xylonolactonase activity is xylC, or homolog thereof. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having xylonolactonase activity comprises an amino acid sequence set forth in SEQ ID NO: 67. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having xylonolactonase activity comprises a nucleic acid sequence set forth in SEQ ID NO: 66.

[00952] In one embodiment, the enzyme having xylonate dehydratase activity is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Caulobacter sp.*, *Sulfolobus sp.* and *E. coli*. In another embodiment, the enzyme having xylonate dehydratase activity is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Caulobacter crescentus*, *Sulfolobus solfataricus* and *E. coli*. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having xylonate dehydratase activity is selected from xylD, yjhG and/or yagF, or homolog thereof. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having xylonate dehydratase activity comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 69, 72 and 75. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having xylonate dehydratase activity comprises a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 68, 70, 71, 73 and 74.

[00953] In one embodiment, the enzyme having 2-keto-3-deoxy-D-pentionate aldolase activity is encoded by one or more nucleic acid molecules obtained from a microorganism selected from *Pseudomonas sp.* and *E. coli*. In another embodiment, the enzyme having 2-keto-3-deoxy-D-pentionate aldolase activity is encoded by one or more nucleic acid molecules obtained from *E. coli*. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having 2-keto-3-deoxy-D-pentionate aldolase activity is selected from yjhH and/or yagE, or homolog thereof. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having 2-keto-3-deoxy-D-pentionate aldolase activity comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 78 and 81. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having 2-keto-3-deoxy-D-pentionate aldolase activity comprises a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 76, 77, 79 and 80.

[00954] Co-production of MEG (or glycolic acid) via a C2 pathway and MEG (or glycolic acid) via a C3 pathway

[00955] In one aspect, MEG (or glycolic acid) is produced from one or more hexose feedstock by the lossless transformation of the one or more hexose feedstock to one or more pentose-5-phosphate intermediate, followed by a conversion of the one or more pentose-5-phosphate intermediate to one or more pentose and/or pentose-1-phosphate intermediate, followed by a conversion of the one or more pentose and/or pentose-1-phosphate intermediate to MEG (or glycolic acid) and DHAP or pyruvate via a C2 pathway, and a conversion of DHAP or pyruvate to MEG via a C3 pathway.

[00956] In some embodiments, the application relates to a recombinant microorganism capable of producing MEG (or glycolic acid) from one or more hexose feedstock, wherein the recombinant microorganism from embodiment **[A]** or from embodiment **[B]** (and optionally comprising embodiment **[C]**), and having additionally one or more of embodiments **[D]**, **[E]**, **[F]**, and **[G]**, further comprises one or more C2 biosynthesis pathway selected from embodiments **[H]**, **[I]** and **[J]** for production of MEG (or glycolic acid) and one or more C3 biosynthesis pathway for the production of MEG (or glycolic acid). The C3 biosynthesis pathways for the production of MEG are, for example, as described in WO 2010/076324 (Metabolic Explorer), herein incorporated by reference in its entirety.

[00957] In some embodiments, the C3 biosynthesis pathway for the production of MEG comprises three enzymatic reactions starting with transformation of the 3-phosphohydroxypyruvate precursor (precursor for serine). First, a phosphatase activity allows conversion of phosphohydroxypyruvate into hydroxypyruvate. Hydroxypyruvate is then transformed into glycolaldehyde with a 2-keto acid decarboxylase activity. Finally, a hydroxy aldehyde reductase activity allows the conversion of glycolaldehyde into ethylene glycol. Another pathway for the production of ethylene glycol starts from L-serine as precursor. First a transaminase or an amino acid oxidase activity allows conversion of serine into hydroxypyruvate. The next two steps to convert hydroxypyruvate into glycolaldehyde and then to MEG are similar to the first pathway described above.

[00958] In a preferred embodiment, the disclosure provides for a recombinant microorganism comprising one or more C3 biosynthesis pathways for production of MEG. In some embodiments, the recombinant microorganism, particularly a bacterium, contains at least one gene encoding a polypeptide with 2-keto acid decarboxylase activity and one gene encoding a polypeptide with hydroxy aldehyde reductase activity. These genes can be exogenous or endogenous, and can be expressed chromosomally or extrachromosomally.

[00959] In a further embodiment of the disclosure, the recombinant microorganism, particularly a bacterium, comprises modifications in which the availability of the intermediate 3-phosphoglycerate is increased. Preferably, the increase is achieved by attenuating the level of expression of genes encoding phosphoglycerate mutases, in particular one or both genes *gpmA* and *pgmM*. This can be done by replacing the wild-type promoter of these genes by a weaker promoter, or by the use of an element destabilizing the corresponding messenger RNA or the protein. If needed, complete attenuation of the genes can also be achieved by the deletion of the corresponding DNA sequences.

[00960] In another embodiment, the recombinant microorganism, particularly a bacterium, comprises modifications in which flux into the serine biosynthesis pathway is stimulated. This can be achieved by increasing the level of expression of 3-phosphoglycerate dehydrogenase and/or phosphoserine aminotransferase, encoded by the *serA* and *serC* genes, respectively. Increasing the level of expression of the 3-phosphoglycerate dehydrogenase and/or phosphoserine aminotransferase can be accomplished by introducing artificial promoters that drive the expression of the *serA* and/or *serC* genes, by increasing the number of copies in the cell or by introducing mutations into the *serA* and/or *serC* genes that increase the activity of the corresponding proteins. The expression of the *serA* gene can also be increased by replacing the wild type *lrp* gene (encoding the leucine- responsive regulatory protein) by an *lrp* mutated allele (such as the *lrp*-1 allele corresponding to a GLU 114 ASP substitution in the *lrp* protein) leading to the constitutive activation of the transcription of the *serA* gene.

[00961] In a particular embodiment of the disclosure mutations can be introduced into the *serA* gene that reduce the sensitivity of the SerA protein to the feed-back inhibitor serine (feed-back desensitized alleles) and thus permit an increased activity in the presence of serine. Examples of desensitized alleles, i.e. feed-back insensitive alleles, have been described in EP 0 931 833 (Ajinomoto) or EP 0 620 853 (Wacker).

[00962] In another embodiment, the recombinant microorganism, particularly a bacterium, comprises modifications in which flux into the hydroxypyruvate biosynthesis pathway is stimulated. This result can be achieved by increasing the level of expression of serine transaminase or serine oxidase (for the pathway starting from serine as precursor), or by increasing the expression of 3-phosphohydroxypyruvate phosphatase. Increasing the level of expression of serine oxidase can be accomplished by introducing and overexpressing the gene coding for L-amino acid oxidase from *R. opacus*, or by introducing mutations into the gene that increase the activity of the corresponding protein. An increase in the expression of serine transaminase can be accomplished by introducing artificial promoters that drive the

expression of the serC gene of *E. coli*, by increasing the number of copies in the cell or by introducing mutations into the serC gene that increase the activity of the corresponding protein. An increase of the expression of 3-phosphohydroxypyruvate phosphatase can be accomplished by introducing artificial promoters that drive the expression of the yeaB gene or serB gene of *E. coli*, by increasing the number of copies in the cell or by introducing mutations into the yeaB gene or the serB gene that increase the activity of the corresponding proteins. An increase of the expression of 3-phosphohydroxypyruvate phosphatase can also be accomplished by introducing and overexpressing the gene GPP2 from *S. cerevisiae*, or by introducing mutations into the GPP2 gene that increase the activity of the corresponding protein.

[00963] In a further embodiment of the disclosure, the recombinant microorganism, particularly a bacterium, comprises modifications to present an attenuated level of glycolaldehyde conversion to other compounds than ethylene glycol. This may be achieved by attenuating the level of glycolaldehyde consuming enzymes like hydroxythreonine aldolase (encoded by UaE) or glycolaldehyde dehydrogenase (encoded by aldA, aldB). Attenuation of these genes can be done by replacing the natural promoter by a lower strength promoter or by elements destabilizing the corresponding messenger RNA or the protein. If needed, complete attenuation of the gene can also be achieved by a deletion of the corresponding DNA sequence.

[00964] In a further embodiment of the disclosure, the efficiency of sugar import is increased, either by using a sugar import system not relying on phosphoenolpyruvate (PEP) as phosphodonator like galP that is known to transport glucose, or by providing more phosphoenolpyruvate (PEP) to the sugar-phosphotransferase system. Various means exist that may be used to increase the availability of PEP in a microorganism. In particular, this can be accomplished by attenuating the reaction $\text{PEP} \rightarrow \text{pyruvate}$. Preferentially, at least one gene selected among pykA and pykF, encoding pyruvate kinase, is attenuated in said strain in order to obtain this result. Another way to increase the availability of PEP is to favour the reaction $\text{pyruvate} \rightarrow \text{PEP}$. This can be accomplished by increasing the activity of phosphoenolpyruvate synthase which catalyzes the above reaction. This enzyme is encoded by the ppsA gene. Therefore, in the microorganism, the expression of the ppsA gene is preferentially increased. Both modifications can be present in the microorganism simultaneously.

[00965] In a further embodiment of the disclosure, the recombinant microorganism, particularly a bacterium, comprises modifications to present an attenuated level of serine conversion to other compounds than ethylene glycol. This result may be achieved by attenuating the level of serine consuming enzymes like serine deaminases (encoded by

sdaA, sdaB and/or tdcG), serine transacetylase (encoded by cysE), tryptophan synthase (encoded by trpAB) or serine hydroxymethyltransferase (encoded by glyA). These genes can be attenuated by replacing the natural promoter by a lower strength promoter or by elements destabilizing the corresponding messenger RNA or the protein. If needed, complete attenuation of the gene can also be achieved by a deletion of the corresponding DNA sequence.

[00966] In a further embodiment of the disclosure, the recombinant microorganism, particularly a bacterium, comprises modifications to present an attenuated level of hydroxypyruvate conversion to other compounds than glycolaldehyde. This result may be achieved by attenuating the level of hydroxypyruvate consuming enzymes like hydroxypyruvate reductase (encoded by ghrA) or hydroxypyruvate isomerase (encoded by hyi). These genes can be attenuated by replacing the natural promoter by a lower strength promoter or by elements destabilizing the corresponding messenger RNA or the protein. If needed, complete attenuation of the gene can also be achieved by a deletion of the corresponding DNA sequence.

[00967] In some embodiments, the application relates to a recombinant microorganism capable of producing MEG (or glycolic acid) from one or more hexose feedstock, wherein the recombinant microorganism from embodiment **[A]** or from embodiment **[B]** (and optionally comprising embodiment **[C]**), and having additionally one or more of embodiments **[D]**, **[E]**, **[F]**, and **[G]**, further comprises one or more C3 biosynthesis pathway for the production of MEG (or glycolic acid). The C3 biosynthesis pathways for the production of MEG are, for example, as described in as described in WO 2011/130378 (Genomatica), herein incorporated by reference in its entirety.

[00968] In some embodiments, the disclosure provides a recombinant microorganism comprising an ethylene glycol pathway having at least one exogenous nucleic acid encoding an ethylene glycol pathway enzyme expressed in a sufficient amount to produce ethylene glycol, the ethylene glycol pathway including a serine aminotransferase, a serine oxidoreductase (deaminating), a hydroxypyruvate decarboxylase, a glycolaldehyde reductase, a serine decarboxylase, an ethanolamine aminotransferase, an ethanolamine oxidoreductase (deaminating), a hydroxypyruvate reductase or a glycerate decarboxylase.

[00969] In some embodiments, the recombinant microorganism comprises an ethylene glycol pathway having at least one exogenous nucleic acid encoding ethylene glycol pathway enzymes expressed in a sufficient amount to produce ethylene glycol, the ethylene glycol pathway including a serine aminotransferase or a serine oxidoreductase (deaminating); a hydroxypyruvate decarboxylase, and a glycolaldehyde reductase.

[00970] In some embodiments, the recombinant microorganism comprises an ethylene glycol pathway having at least one exogenous nucleic acid encoding ethylene glycol pathway enzymes expressed in a sufficient amount to produce ethylene glycol, the ethylene glycol pathway including a serine aminotransferase or a serine oxidoreductase (deaminating); a hydroxypyruvate reductase, and a glycerate decarboxylase.

[00971] In some embodiments, the recombinant microorganism comprises an ethylene glycol pathway having at least one exogenous nucleic acid encoding ethylene glycol pathway enzymes expressed in a sufficient amount to produce ethylene glycol, the ethylene glycol pathway including a serine decarboxylase; an ethanolamine aminotransferase or an ethanolamine oxidoreductase (deaminating), and a glycolaldehyde reductase.

[00972] In some embodiments, the disclosure provides a recombinant microorganism comprising an ethylene glycol pathway having at least one exogenous nucleic acid encoding an ethylene glycol pathway enzyme expressed in a sufficient amount to produce ethylene glycol, the ethylene glycol pathway including a hydroxypyruvate decarboxylase, glycolaldehyde reductase, a hydroxypyruvate reductase, a glycerate decarboxylase, a 3-phosphoglycerate phosphatase, and a glycerate kinase.

[00973] In some embodiments, the recombinant microorganism comprises an ethylene glycol pathway having at least one exogenous nucleic acid encoding ethylene glycol pathway enzymes expressed in a sufficient amount to produce ethylene glycol, the ethylene glycol pathway including a hydroxypyruvate reductase; a hydroxypyruvate decarboxylase, and a glycolaldehyde reductase.

[00974] In some embodiments, the recombinant microorganism comprises an ethylene glycol pathway having at least one exogenous nucleic acid encoding ethylene glycol pathway enzymes expressed in a sufficient amount to produce ethylene glycol, the ethylene glycol pathway including a 3-phosphoglycerate phosphatase or a glycerate kinase; a hydroxypyruvate reductase; a hydroxypyruvate decarboxylase, and a glycolaldehyde reductase.

[00975] In some embodiments, the recombinant microorganism comprises an ethylene glycol pathway having at least one exogenous nucleic acid encoding an ethylene glycol pathway enzyme expressed in a sufficient amount to produce ethylene glycol, the ethylene glycol pathway including a glycerate decarboxylase.

[00976] In some embodiments, the recombinant microorganism comprises an ethylene glycol pathway having at least one exogenous nucleic acid encoding ethylene glycol

pathway enzymes expressed in a sufficient amount to produce ethylene glycol, the ethylene glycol pathway including a 3-phosphoglycerate phosphatase or a glycerate kinase and a glycerate decarboxylase.

[00977] In some embodiments, the disclosure provides a recombinant microorganism comprising an ethylene glycol pathway, wherein the recombinant microorganism comprises at least one exogenous nucleic acid encoding an enzyme or protein that converts a substrate to a product selected from the group consisting of serine to hydroxypyruvate, hydroxypyruvate to glycolaldehyde, glycolaldehyde to ethylene glycol, serine to ethanolamine, ethanolamine to glycolaldehyde, 3-phosphoglycerate to glycerate, glycerate to hydroxypyruvate, hydroxypyruvate to glycerate, and glycerate to ethylene glycol. One skilled in the art will understand that these are merely exemplary and that any of the substrate-product pairs disclosed herein suitable to produce a desired product and for which an appropriate activity is available for the conversion of the substrate to the product can be readily determined by one skilled in the art based on the teachings herein. Thus, the disclosure provides a recombinant microorganism containing at least one exogenous nucleic acid encoding an enzyme or protein, where the enzyme or protein converts the substrates and products of an ethylene glycol pathway.

[00978] While generally described herein as a recombinant microorganism that contains an ethylene glycol pathway, it is understood that the disclosure additionally provides a recombinant microorganism comprising at least one exogenous nucleic acid encoding an ethylene glycol pathway enzyme expressed in a sufficient amount to produce an intermediate of an ethylene glycol pathway. Therefore, in addition to a recombinant microorganism containing an ethylene glycol pathway that produces ethylene glycol, the disclosure additionally provides a recombinant microorganism comprising at least one exogenous nucleic acid encoding an ethylene glycol pathway enzyme, where the microbial organism produces an ethylene glycol pathway intermediate, for example, hydroxypyruvate, ethanolamine, glycolaldehyde, or glycerate.

[00979] In some embodiments, a serine aminotransferase or serine oxidoreductase (deaminating) catalyzes the conversion of serine to hydroxypyruvate. In some embodiments, a hydroxypyruvate decarboxylase catalyzes the conversion of hydroxypyruvate to glycolaldehyde. In some embodiments, a glycolaldehyde reductase catalyzes the conversion of glycolaldehyde to ethylene glycol. In some embodiments, a serine decarboxylase catalyzes the conversion of serine to ethanolamine. In some embodiments, an ethanolamine aminotransferase or ethanolamine oxidoreductase (deaminating) catalyzes the conversion of ethanolamine to glycolaldehyde. In some embodiments, a hydroxypyruvate reductase

catalyzes the conversion of glycerate to hydroxypyruvate. In some embodiments, a glycerate decarboxylase catalyzes the conversion of glycerate to ethylene glycol. In some embodiments, a 3-phosphoglycerate phosphatase or glycerate kinase catalyzes the conversion of 3-phosphoglycerate to glycerate.

5 **[00980]** In some embodiments, MEG (or glycolic acid) is produced from the lossless transformation of one or more hexose feedstock to one or more pentose-5-phosphate intermediate, followed by a conversion of the one or more pentose-5-phosphate intermediate to one or more pentose and/or pentose-1-phosphate intermediate, followed by a conversion of the one or more pentose and/or pentose-1-phosphate intermediate to MEG (or glycolic acid) and DHAP or pyruvate via a C2 pathway, and a conversion of DHAP or pyruvate to MEG (or glycolic acid) via a C3 pathway.

[00981] In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[H]** and the C3 pathway for production of MEG (or glycolic acid) comprises embodiment **[K]**. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[H]** and the C3 pathway for production of MEG (or glycolic acid) comprises embodiment **[L]**. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[I]** and the C3 pathway for production of MEG (or glycolic acid) comprises embodiment **[K]**. In further embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[I]** and the C3 pathway for production of MEG (or glycolic acid) comprises embodiment **[L]**. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[J]** and the C3 pathway for production of MEG (or glycolic acid) comprises embodiment **[K]**. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[J]** and the C3 pathway for production of MEG (or glycolic acid) comprises embodiment **[L]**.

[00982] **[K]** In one embodiment, the application relates to a recombinant microorganism capable of producing monoethylene glycol (MEG) (or glycolic acid) from one or more hexose feedstock, wherein the recombinant microorganism from embodiment **[A]** or from embodiment **[B]** (and optionally comprising embodiment **[C]**), and comprising additionally one or more of embodiments **[D]**, **[E]**, **[F]**, and **[G]**, further comprising one or more of embodiments **[H]**, **[I]**, and **[J]**, further expresses one or more of the following from (a) to (h):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphoglycerate dehydrogenase activity that catalyzes the conversion of 3-phosphoglycerate to 3-phosphohydroxypyruvate;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having phosphoserine aminotransferase activity that catalyzes the conversion of 3-phosphohydroxypyruvate from (a) to phospho-L-serine;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphohydroxypyruvate phosphatase activity that catalyzes the conversion of 3-phosphohydroxypyruvate from (a) to hydroxypyruvate;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having phosphoserine phosphatase activity that catalyzes the conversion of phospho-L-serine from (b) to L-serine;

(e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having L-serine transaminase or serine oxidase activity that catalyzes the conversion of L-serine from (d) to hydroxypyruvate;

(f) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxypyruvate decarboxylase activity that catalyzes the conversion of hydroxypyruvate from (c) or (e) to glycolaldehyde;

(g) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde reductase activity that catalyzes the conversion of glycolaldehyde from (f) to MEG;

(h) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde dehydrogenase activity that catalyzes the conversion of glycolaldehyde from (f) to glycolic acid;

wherein the produced intermediate DHAP from embodiments [H] and/or [I] and/or pyruvate from embodiment [J] is converted to 3-phosphoglycerate through endogenous glycolysis or gluconeogenesis, respectively, in the microorganism, and wherein MEG (or glycolic acid) is produced.

[00983] [L] In one embodiment, the application relates to a recombinant microorganism capable of producing monoethylene glycol (MEG) (or glycolic acid) from one or more hexose feedstock, wherein the recombinant microorganism from embodiment [A] or from embodiment [B] (and optionally comprising embodiment [C]), and comprising additionally one or more of embodiments [D], [E], [F], and [G], further comprising one or more of embodiments [H], [I], and [J], further expresses one or more of the following from (a) to (j):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 2-phosphoglycerate phosphatase activity and/or an enzyme having glycerate-2-kinase activity that catalyzes the conversion of 2-phosphoglycerate to glycerate;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphoglycerate phosphatase activity or an enzyme having glycerate-3-kinase activity that catalyzes the conversion of 3-phosphoglycerate to glycerate;

5 (c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxypyruvate reductase activity that catalyzes the conversion of glycerate from (a) and/or (b) to hydroxypyruvate;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine aminotransferase activity or an enzyme having serine oxidoreductase (deaminating) activity that catalyzes the conversion of L-serine to hydroxypyruvate;

10 (e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having L-serine decarboxylase activity that catalyzes the conversion of L-serine to ethanolamine;

(f) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxypyruvate decarboxylase activity that catalyzes the conversion of
15 hydroxypyruvate from (c) and/or (d) to glycolaldehyde;

(g) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having ethanolamine aminotransferase or ethanolamine oxidoreductase (deaminating) activity that catalyzes the conversion of ethanolamine from (e) to glycolaldehyde;

20 (h) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycerate decarboxylase activity that catalyzes the conversion of glycerate from (a) and/or (b) to MEG;

(i) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde reductase activity that catalyzes the conversion of
25 glycolaldehyde from (f) and/or (g) to MEG;

(j) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde dehydrogenase activity that catalyzes the conversion of glycolaldehyde from (f) and/or (g) to glycolic acid;

wherein the produced intermediate DHAP from embodiments [H] and/or [I] and/or pyruvate
30 from embodiment [J] is converted to 3-phosphoglycerate and/or 2-phosphoglycerate through endogenous glycolysis or gluconeogenesis in the microorganism, and wherein MEG (or glycolic acid) is produced.

[00984] In some embodiments, a 2-keto acid decarboxylase, a hydroxypyruvate decarboxylase or a 2-oxoglutarate decarboxylase converts hydroxypyruvate to glycolaldehyde. In some embodiments, the enzyme that converts hydroxypyruvate to glycolaldehyde is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to Kivd, dxs, or SucA. In some embodiments, the 2-keto acid decarboxylase is Kivd from *Lactococcus lactis*. In yet another embodiment, the one or more nucleic acid molecules encoding the 2-keto acid decarboxylase comprises an amino acid sequence set forth in SEQ ID NO: 224. In a further embodiment, the one or more nucleic acid molecules encoding the 2-keto acid decarboxylase comprises by a nucleic acid sequence set forth in SEQ ID NO: 223. In some embodiments, the 2-oxoglutarate decarboxylase is SucA from *E. coli*. In yet another embodiment, the one or more nucleic acid molecules encoding the 2-oxoglutarate decarboxylase comprises an amino acid sequence set forth in SEQ ID NO: 226. In a further embodiment, the one or more nucleic acid molecules encoding the 2-oxoglutarate decarboxylase comprises a nucleic acid sequence set forth in SEQ ID NO: 225.

[00985] In some embodiments, the hydroxy aldehyde reductase can be a glycolaldehyde reductase. In some embodiments, the enzyme having glycolaldehyde reductase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to yqhD or FucO. In some embodiments, the enzyme having glycolaldehyde reductase activity is encoded by one or more nucleic acid molecules obtained from a microorganism selected from *E. coli* and *S. cerevisiae*. In another embodiment, the one or more nucleic acid molecules is selected from gldA, GRE2, GRE3, yqhD, ydjG, fucO, yafB (dkgB), and/or yqhE (dkgA), or homolog thereof. In another embodiment, the one or more nucleic acid molecules is yqhD. In some embodiments, the yqhD comprises a G149E mutation. In a further embodiment, the enzyme having glycolaldehyde reductase activity comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 13, 15, 17, 20, 23, 25, 28, 30 and 32. In yet a further embodiment, the enzyme having glycolaldehyde reductase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 12, 14, 16, 18, 19, 21, 22, 24, 26, 27, 29 and 31.

[00986] In some embodiments, the 3-phosphoglycerate dehydrogenase can be a 3-phospho-hydroxypyruvate reductase or a 2-oxoglutarate reductase. In some embodiments, the enzyme having 3-phospho-hydroxypyruvate reductase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to serA. In some embodiments, the enzyme having 3-phospho-hydroxypyruvate reductase activity is serA from *E. coli*. In yet

another embodiment, the one or more nucleic acid molecules encoding the enzyme having 3-phospho-hydroxypyruvate reductase activity comprises an amino acid sequence set forth in SEQ ID NO: 228. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having 3-phospho-hydroxypyruvate reductase activity comprises a nucleic acid sequence set forth in SEQ ID NO: 227.

[00987] In some embodiments, the phosphoserine aminotransferase can be an L-serine transaminase, a serine aminotransferase or a serine-pyruvate aminotransferase. In some embodiments, the enzyme having phosphoserine aminotransferase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to serC. In some embodiments, the enzyme having phosphoserine aminotransferase activity is serC from *E. coli*. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having phosphoserine aminotransferase activity comprises an amino acid sequence set forth in SEQ ID NO: 230. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having phosphoserine aminotransferase activity comprises a nucleic acid sequence set forth in SEQ ID NO: 229.

[00988] In some embodiments, the enzyme having serine-pyruvate aminotransferase activity is AGXT1 from *Homo sapiens*. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having serine-pyruvate aminotransferase activity comprises an amino acid sequence set forth in SEQ ID NO: 244. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having serine-pyruvate aminotransferase activity comprises a nucleic acid sequence set forth in SEQ ID NO: 243.

[00989] In some embodiments, the enzyme having 3-phospho-hydroxypyruvate phosphatase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to yeaB (nudL). In some embodiments, the enzyme having 3-phospho-hydroxypyruvate phosphatase activity is yeaB (nudL) from *E. coli*. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having 3-phospho-hydroxypyruvate phosphatase activity comprises an amino acid sequence set forth in SEQ ID NO: 232. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having 3-phospho-hydroxypyruvate phosphatase activity comprises a nucleic acid sequence set forth in SEQ ID NO: 231.

[00990] In some embodiments, the enzyme having phosphoserine phosphatase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to serB. In some

embodiments, the enzyme having phosphoserine phosphatase activity is serB from *E. coli*. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having phosphoserine phosphatase activity comprises an amino acid sequence set forth in SEQ ID NO: 234. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having phosphoserine phosphatase activity comprises a nucleic acid sequence set forth in SEQ ID NO: 233.

[00991] In some embodiments, an enzyme having 2-phosphoglycerate phosphatase activity or an enzyme having glycerate-2-kinase activity converts 2-phosphoglycerate to glycerate. In some embodiments, the enzyme that converts 2-phosphoglycerate to glycerate is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to phoA, glxK or garK. In some embodiments, the enzyme having 2-phosphoglycerate phosphatase activity is phoA from *E. coli*. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having 2-phosphoglycerate phosphatase activity comprises an amino acid sequence set forth in SEQ ID NO: 246. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having 2-phosphoglycerate phosphatase activity comprises a nucleic acid sequence set forth in SEQ ID NO: 245. In some embodiments, the enzyme having glycerate-2-kinase activity is glxK from *E. coli*. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having glycerate-2-kinase activity comprises an amino acid sequence set forth in SEQ ID NO: 250. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having glycerate-2-kinase activity comprises a nucleic acid sequence set forth in SEQ ID NO: 249. In some embodiments, the enzyme having glycerate-2-kinase activity is garK from *E. coli*. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having glycerate-2-kinase activity comprises an amino acid sequence set forth in SEQ ID NO: 252. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having glycerate-2-kinase activity comprises a nucleic acid sequence set forth in SEQ ID NO: 251.

[00992] In some embodiments, an enzyme having 3-phosphoglycerate phosphatase activity or an enzyme having glycerate-3-kinase activity converts 3-phosphoglycerate to glycerate. In some embodiments, the enzyme that converts 3-phosphoglycerate to glycerate is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to phoA or GLYK. In some embodiments, the enzyme having 3-phosphoglycerate phosphatase activity is phoA from *E. coli*. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having 3-phosphoglycerate phosphatase activity comprises an amino acid

sequence set forth in SEQ ID NO: 246. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having 3-phosphoglycerate phosphatase activity comprises a nucleic acid sequence set forth in SEQ ID NO: 245. In some embodiments, the enzyme having glycerate-3-kinase activity is GLYK from *Arabidopsis thaliana*. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having glycerate-3-kinase activity comprises an amino acid sequence set forth in SEQ ID NO: 248. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having glycerate-3-kinase activity comprises a nucleic acid sequence set forth in SEQ ID NO: 247.

[00993] In some embodiments, an enzyme having hydroxypyruvate reductase activity converts glycerate to hydroxypyruvate. In some embodiments, the enzyme that converts glycerate to hydroxypyruvate is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to ghrB. In some embodiments, the enzyme having hydroxypyruvate reductase activity is ghrB from *E. coli*. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme hydroxypyruvate reductase activity comprises an amino acid sequence set forth in SEQ ID NO: 242. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme hydroxypyruvate reductase activity comprises a nucleic acid sequence set forth in SEQ ID NO: 241.

[00994] In some embodiments, an enzyme having serine decarboxylase activity converts L-serine to ethanolamine. In some embodiments, the enzyme that converts L-serine to ethanolamine is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to SDC. In some embodiments, the enzyme having serine decarboxylase activity is SDC from *Arabidopsis thaliana*. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having serine decarboxylase activity comprises an amino acid sequence set forth in SEQ ID NO: 236. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having serine decarboxylase activity comprises a nucleic acid sequence set forth in SEQ ID NO: 235.

[00995] In some embodiments, an enzyme having ethanolamine aminotransferase activity or an enzyme having ethanolamine oxidoreductase (deaminating) activity converts ethanolamine to glycolaldehyde. In some embodiments, the enzyme that converts ethanolamine to glycolaldehyde is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to alaA or tyrA. In some embodiments, the enzyme having ethanolamine

aminotransferase activity is *alaA* from *E. coli*. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having ethanolamine aminotransferase activity comprises an amino acid sequence set forth in SEQ ID NO: 240. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having ethanolamine aminotransferase activity comprises a nucleic acid sequence set forth in SEQ ID NO: 239. In some embodiments, the enzyme having ethanolamine oxidoreductase (deaminating) activity is *tynA* from *E. coli*. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having ethanolamine oxidoreductase (deaminating) activity comprises an amino acid sequence set forth in SEQ ID NO: 238. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having ethanolamine oxidoreductase (deaminating) activity comprises a nucleic acid sequence set forth in SEQ ID NO: 237.

[00996] In another aspect, MEG (or glycolic acid) is produced from the lossless transformation of one or more hexose feedstock to one or more pentose-5-phosphate intermediate, followed by a conversion of the one or more pentose-5-phosphate intermediate to one or more pentose and/or pentose-1-phosphate intermediate, followed by a conversion of the one or more pentose and/or pentose-1-phosphate intermediate to MEG (or glycolic acid) and DHAP or pyruvate via a C2 pathway, and a conversion of DHAP or pyruvate to one or more co-product via a C3 pathway. In some embodiments, the C2 pathway comprises embodiment **[H]**. In further embodiments, the C2 pathway comprises embodiment **[I]**. In yet further embodiments, the C2 pathway comprises embodiment **[J]**.

[00997] Co-production of MEG (or glycolic acid) via a C2 pathway and acetone, isopropanol, propene and/or isobutene via a C3 pathway

[00998] In some embodiments, MEG (or glycolic acid) is produced from the lossless transformation of one or more hexose feedstock to one or more pentose-5-phosphate intermediate, followed by a conversion of the one or more pentose-5-phosphate intermediate to one or more pentose and/or pentose-1-phosphate intermediate, followed by a conversion of the one or more pentose and/or pentose-1-phosphate intermediate to MEG (or glycolic acid) and DHAP or pyruvate via a C2 pathway, and a conversion of DHAP or pyruvate to acetone via a C3 pathway. In some embodiments, the C2 pathway comprises embodiment **[H]**. In further embodiments, the C2 pathway comprises embodiment **[I]**. In yet further embodiments, the C2 pathway comprises embodiment **[J]**.

[00999] In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[H]** and the C3 pathway for production of acetone comprises embodiment **[M]**. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[I]** and the C3 pathway for production of acetone comprises

embodiment **[M]**. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[J]** and the C3 pathway for production of acetone comprises embodiment **[M]**.

[001000] **[M]** In one embodiment, the application relates to a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and acetone, from one or more hexose feedstock, wherein the recombinant microorganism from embodiment **[A]** or from embodiment **[B]** (and optionally comprising embodiment **[C]**), and comprising additionally one or more of embodiments **[D]**, **[E]**, **[F]** and **[G]**, further comprising one or more of embodiments **[H]**, **[I]**, and **[J]**, further expresses one or more of the following from (a) to (c):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having thiolase or acetyl coenzyme A acetyltransferase activity that catalyzes the conversion of acetyl-CoA to acetoacetyl-CoA;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having acetyl-CoA:acetoacetate-CoA transferase activity or an enzyme having acetate:acetoacetyl-CoA hydrolase activity that catalyzes the conversion of acetoacetyl-CoA from (a) to acetoacetate; and/or

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having acetoacetate decarboxylase activity that catalyzes the conversion of acetoacetate from (b) to acetone;

wherein the produced intermediate DHAP from embodiments **[H]** and/or **[I]** and/or pyruvate from embodiment **[J]** is converted to acetyl-CoA through endogenous glycolysis in the microorganism, and wherein MEG (or glycolic acid) and acetone are co-produced.

[001001] In some embodiments, MEG (or glycolic acid) is produced from the lossless transformation of one or more hexose feedstock to one or more pentose-5-phosphate intermediate, followed by a conversion of the one or more pentose-5-phosphate intermediate to one or more pentose and/or pentose-1-phosphate intermediate, followed by a conversion of the one or more pentose and/or pentose-1-phosphate intermediate to MEG (or glycolic acid) and DHAP or pyruvate via a C2 pathway, and a conversion of DHAP or pyruvate to isobutene via a C3 pathway. In some embodiments, the C2 pathway comprises embodiment **[H]**. In further embodiments, the C2 pathway comprises embodiment **[I]**. In yet further embodiments, the C2 pathway comprises embodiment **[J]**.

[001002] In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[H]** and the C3 pathway for production of isobutene comprises

embodiment [N]. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [I] and the C3 pathway for production of isobutene comprises embodiment [N]. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [J] and the C3 pathway for production of isobutene comprises embodiment [N].

[001003] [N] In some embodiments, the application relates to a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and isobutene via acetone or HMG-CoA, from one or more hexose feedstock, wherein the recombinant microorganism from embodiment [A] or from embodiment [B] (and optionally comprising embodiment [C]), and comprising additionally one or more of embodiments [D], [E], [F] and [G], further comprising one or more of embodiments [H], [I], and [J], further expresses one or more of the following from (a) to (d):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having thiolase or acetyl coenzyme A acetyltransferase activity that catalyzes the conversion of acetyl-CoA to acetoacetyl-CoA;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having acetyl-CoA:acetoacetate-CoA transferase activity or an enzyme having acetate:acetoacetyl-CoA hydrolase activity that catalyzes the conversion of acetoacetyl-CoA from (a) to acetoacetate;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having acetoacetate decarboxylase activity that catalyzes the conversion of acetoacetate from (b) to acetone;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-hydroxyisovalerate synthase activity that catalyzes the conversion of acetone from (c) and acetyl-CoA to 3-hydroxyisovalerate (3HIV);

or

wherein the recombinant microorganism expresses one or more of the following from (e) to (j):

(e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having thiolase or acetyl coenzyme A acetyltransferase activity that catalyzes the conversion of acetyl-CoA to acetoacetyl-CoA;

(f) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxymethylglutaryl-CoA synthase activity that catalyzes the conversion of acetoacetyl-CoA from (e) and acetyl-CoA to 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA);

(g) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having methylglutaconyl-CoA hydratase activity that catalyzes the conversion of HMG-CoA from (f) to 3-methylglutaconyl-CoA;

(h) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having methylcrotonyl-CoA carboxylase activity that catalyzes the conversion of 3-methylglutaconyl-CoA from (g) to 3-methylcrotonyl-CoA;

(i) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having methylcrotonyl-CoA hydratase activity that catalyzes the conversion of 3-methylcrotonyl-CoA from (h) to 3-hydroxyisovaleryl-CoA;

(j) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-hydroxyisovaleryl-CoA thioesterase activity that catalyzes the conversion of 3-hydroxyisovaleryl-CoA from (i) to 3HIV;

wherein the recombinant microorganism further expresses (a1) and (a2), and/or (b1) selected from:

(a1) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3HIV kinase activity that catalyzes the conversion of 3HIV from (d) or (j) to 3HIV-3-phosphate;

(a2) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3HIV-3-phosphate decarboxylase activity that catalyzes the conversion of 3HIV-3-phosphate from (a1) to isobutene;

(b1) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3HIV decarboxylase activity that catalyzes the conversion of 3HIV from (d) or (j) to isobutene;

wherein the produced intermediate DHAP from embodiments [H] and/or [I] and/or pyruvate from embodiment [J] is converted to acetyl-CoA through endogenous glycolysis in the microorganism, and wherein MEG (or glycolic acid) and isobutene are co-produced.

[001004] In some embodiments, MEG (or glycolic acid) is produced from the lossless transformation of one or more hexose feedstock to one or more pentose-5-phosphate intermediate, followed by a conversion of the one or more pentose-5-phosphate intermediate to one or more pentose and/or pentose-1-phosphate intermediate, followed by a conversion of the one or more pentose and/or pentose-1-phosphate intermediate to MEG (or glycolic acid) and DHAP or pyruvate via a C2 pathway, and a conversion of DHAP or pyruvate to isopropanol via a C3 pathway. In some embodiments, the C2 pathway comprises

embodiment [H]. In further embodiments, the C2 pathway comprises embodiment [I]. In yet further embodiments, the C2 pathway comprises embodiment [J].

[001005] [O] In one embodiment, the recombinant microorganisms from embodiment [M] and/or [N] (optionally comprising embodiment [EE]), optionally further express at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having secondary alcohol dehydrogenase activity that catalyzes the conversion of acetone to isopropanol.

[001006] In some embodiments, the alcohol dehydrogenase has at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity with an alcohol dehydrogenase from *Clostridium* sp. In other embodiments, the alcohol dehydrogenase is an alcohol dehydrogenase selected from *Clostridium beijerinckii* adh and *Clostridium carboxidivorans* adh. In a further embodiment, the alcohol dehydrogenase comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 138 and 140. In yet another embodiment, the alcohol dehydrogenase is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 136, 137, and 139.

[001007] In some embodiments, MEG (or glycolic acid) is produced from the lossless transformation of one or more hexose feedstock to one or more pentose-5-phosphate intermediate, followed by a conversion of the one or more pentose-5-phosphate intermediate to one or more pentose and/or pentose-1-phosphate intermediate, followed by a conversion of the one or more pentose and/or pentose-1-phosphate intermediate to MEG (or glycolic acid) and DHAP or pyruvate via a C2 pathway, and a conversion of DHAP or pyruvate to propene via a C3 pathway. In some embodiments, the C2 pathway comprises embodiment [H]. In further embodiments, the C2 pathway comprises embodiment [I]. In yet further embodiments, the C2 pathway comprises embodiment [J].

[001008] [P] In another embodiment, the recombinant microorganisms from embodiment [O] (optionally comprising embodiment [EE]), optionally further expresses at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having dehydratase activity that catalyzes the conversion of isopropanol to propene.

[001009] In one embodiment of any pathway disclosed above for isobutene co-production, the enzyme having 3-hydroxyisovalerate synthase activity is encoded by one or more nucleic acid molecules obtained from a microorganism selected from *Mus* sp., *Saccharomyces* sp., *Lactobacillus* sp. and *Polaromonas* sp. In another embodiment, the enzyme having 3-hydroxyisovalerate synthase activity is encoded by one or more nucleic acid molecules obtained from a microorganism selected from *Mus musculus*, *Saccharomyces cerevisiae*, *Lactobacillus crispatus* and *Polaromonas naphthalenivorans*. In some embodiments, the one

or more nucleic acid molecules encoding the enzyme having 3-hydroxyisovalerate synthase activity is selected from Hmgcs1, ERG13, PksG and/or Pnap_0477, or homolog thereof. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having 3-hydroxyisovalerate synthase activity comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 105, 107, 109 and 111. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having 3-hydroxyisovalerate synthase activity comprises a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 104, 106, 108 and 110.

[001010] In one embodiment of any pathway disclosed above for isobutene co-production, the enzyme having hydroxymethylglutaryl-CoA synthase activity is encoded by one or more nucleic acid molecules obtained from *Saccharomyces* sp. In another embodiment, the enzyme having hydroxymethylglutaryl-CoA synthase activity is encoded by one or more nucleic acid molecules obtained from *Saccharomyces cerevisiae*. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having hydroxymethylglutaryl-CoA synthase activity is HmgS, or homolog thereof. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having hydroxymethylglutaryl-CoA synthase activity comprises an amino acid sequence set forth in SEQ ID NO: 123. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having hydroxymethylglutaryl-CoA synthase activity comprises a nucleic acid sequence set forth in SEQ ID NO: 122.

[001011] In one embodiment of any pathway disclosed above for isobutene co-production, the enzyme having methylglutaconyl-CoA hydratase activity is encoded by one or more nucleic acid molecules obtained from *Pseudomonas* sp. In another embodiment, the enzyme having methylglutaconyl-CoA hydratase activity is encoded by one or more nucleic acid molecules obtained from *Pseudomonas putida*. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having methylglutaconyl-CoA hydratase activity is liuC, or homolog thereof. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having methylglutaconyl-CoA hydratase activity comprises an amino acid sequence set forth in SEQ ID NO: 125. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having methylglutaconyl-CoA hydratase activity comprises a nucleic acid sequence set forth in SEQ ID NO: 124.

[001012] In one embodiment of any pathway disclosed above for isobutene co-production, the enzyme having methylcrotonyl-CoA carboxylase activity is encoded by one or more nucleic acid molecules obtained from *Pseudomonas* sp. In another embodiment, the enzyme having methylcrotonyl-CoA carboxylase activity is encoded by one or more nucleic acid molecules obtained from *Pseudomonas aeruginosa*. In some embodiments, the one or more

nucleic acid molecules encoding the enzyme having methylcrotonyl-CoA carboxylase activity is liuB, and/or liuD, or homologs thereof. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having methylcrotonyl-CoA carboxylase activity comprises an amino acid sequence selected from SEQ ID NOs: 127 and 129. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having methylcrotonyl-CoA carboxylase activity comprises a nucleic acid sequence selected from SEQ ID NOs: 126 and 128.

[001013] In one embodiment of any pathway disclosed above for isobutene co-production, the enzyme having methylcrotonyl-CoA hydratase activity is a 3-ketoacyl-CoA thiolase. In another embodiment, the enzyme having methylcrotonyl-CoA hydratase activity is an enoyl-CoA hydratase. In another embodiment, the enzyme having methylcrotonyl-CoA hydratase activity is encoded by one or more nucleic acid molecules obtained from *E. coli*. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having methylcrotonyl-CoA hydratase activity is fadA, and/or fadB, or homologs thereof. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having methylcrotonyl-CoA hydratase activity comprises an amino acid sequence selected from SEQ ID NOs: 131 and 133. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having methylcrotonyl-CoA hydratase activity comprises a nucleic acid sequence selected from SEQ ID NOs: 130 and 132.

[001014] In one embodiment of any pathway disclosed above for isobutene co-production, the enzyme having 3-hydroxyisovaleryl-CoA thioesterase activity is encoded by one or more nucleic acid molecules obtained from *E. coli*. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having 3-hydroxyisovaleryl-CoA thioesterase activity is tesB, or homolog thereof. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having 3-hydroxyisovaleryl-CoA thioesterase activity comprises an amino acid sequence set forth in SEQ ID NO: 135. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having 3-hydroxyisovaleryl-CoA thioesterase activity comprises a nucleic acid sequence set forth in SEQ ID NO: 134.

[001015] In one embodiment of any pathway disclosed above for isobutene co-production, the enzyme having 3HIV kinase activity is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Thermoplasma sp.* and *Picrophilus sp.* In another embodiment, the enzyme having 3HIV kinase activity is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Thermoplasma acidophilum* and *Picrophilus torridus*. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having 3HIV kinase activity is TA1305 and/or PTO1356, or homolog thereof. In some embodiments, the

TA1305 comprises a L200E mutation. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having 3HIV kinase activity comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 113, 115 and 117. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having 3HIV kinase activity comprises a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 112, 114 and 116.

[001016] In one embodiment of any pathway disclosed above for isobutene co-production, the enzyme having 3HIV-3-phosphate decarboxylase activity is encoded by one or more nucleic acid molecules obtained from *Streptococcus sp.* In another embodiment, the enzyme having 3HIV-3-phosphate decarboxylase activity is encoded by one or more nucleic acid molecules obtained from a microorganism selected from *Streptococcus mitis* and *Streptococcus gordonii*. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having 3HIV-3-phosphate decarboxylase activity comprises smi_1746 and/or mvaD, or homolog thereof. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having 3HIV-3-phosphate decarboxylase activity comprises an amino acid sequence selected from SEQ ID NOs: 119 and 121. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having 3HIV-3-phosphate decarboxylase activity comprises a nucleic acid sequence selected from SEQ ID NOs: 118 and 120.

[001017] In one embodiment of any pathway disclosed above for isobutene co-production, the enzyme having 3HIV decarboxylase activity is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Streptococcus sp.*, *Thermoplasma sp.* and *Picrophilus sp.* In another embodiment, the enzyme having 3HIV decarboxylase activity is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Streptococcus gordonii*, *Thermoplasma acidophilum* and *Picrophilus torridus*. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having 3HIV decarboxylase activity comprises mvaD, TA1305 and/or PTO1356, or homolog thereof. In a further embodiment, the one or more nucleic acid molecules encoding the enzyme having 3HIV decarboxylase activity comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 113, 117 and 121. In yet another embodiment, the one or more nucleic acid molecules encoding the enzyme having 3HIV decarboxylase activity comprises a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 112, 116 and 120.

[001018] Co-production of MEG (or glycolic acid) via a C2 pathway and one or more serine pathway compound via a C3 pathway

[001019] In some embodiments, the production of MEG via a C2 pathway is coupled to the production of one or more serine pathway compound via a C3 pathway. In one embodiment, one or more serine pathway compound relates to the L-serine biosynthetic pathway. In another embodiment, the one or more serine pathway compound is L-serine (Ser), glycine (Gly), monoethanolamine (MEA), and/or ethylenediamine (EDA) (**FIG. 10**).

[001020] Production of Monoethylene glycol (MEG) or glycolic acid via one or more C2 pathway

[001021] To produce MEG, one or more hexose sugar or hexose feedstock is converted to one or more pentose-5-phosphate intermediates via a non-oxidative entry into the pentose phosphate pathway in a manner that preserves yield potential, as described above. The one or more pentose-5-phosphate intermediates are then converted to one or more pentose-1-phosphate or pentose intermediates by pentose-5-phosphatases, D-arabitol phosphate dehydrogenases and/or phosphopentomutases (as described above). The one or more pentose-1-phosphate or pentose intermediates are then degraded by known C2 pathways (D-xylulose-1-phosphate pathway (Alkim et al., Microb Cell Fact (2015) 14:127), D-ribulose-1-phosphate pathway (WO2013/126721), or D-xylonate pathway (WO2013/119020)) into the C2 carbon compound glycolaldehyde and the C3 carbon compound dihydroxy acetonephosphate (DHAP) or pyruvate by an aldolase. Glycolaldehyde is reduced to MEG, consuming an NADH. Alternatively, glycolaldehyde can be oxidized by a glycolaldehyde dehydrogenase to glycolic acid. The C3 compound DHAP is further oxidized to one or more of the L-serine pathway compounds Ser, Gly, MEA, or EDA, producing NADH.

[001022] Ammonia (NH_3)

[001023] Ammonia is a compound of nitrogen and hydrogen with the formula NH_3 , which serves as a precursor to food and fertilizers as well as a building block for the synthesis of pharmaceutical products and commercial cleaning products. Ammonia is present as an ammonium cation when it is positively charged, whose chemical formula is NH_4^+ .

[001024] The present disclosure teaches that ammonia or a similar compound is utilized as nitrogen source. When ammonia or a similar compound is used as a nitrogen source, it is fixed into organic matter, for example, glutamate. In one embodiment, when 2-oxoglutarate is usually formed into glutamate, the ammonia is fixed and this process consumes one NADH.

[001025] Production of L-serine (Ser)

[001026] Ser is produced by the natural pathway via phosphoserine or a variation thereof. In one embodiment, DHAP produced from the production of MEG (or glycolic acid) via a C2

pathway is converted by endogenous glycolysis in the microorganism to 3-phospho-D-glycerate (3-phosphoglycerate). 3-phosphoglycerate is converted to 3-phosphohydroxypyruvate by a D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95). The 3-phosphohydroxypyruvate is then converted to 3-phosphoserine by a phosphoserine aminotransferase (EC 2.6.1.52). The 3-phosphoserine is then converted to L-serine by a phosphoserine phosphatase (EC 3.1.3.3).

[001027] In some embodiments the reaction from 3-phosphoglycerate to L-serine is the following:

[001028] 3-phospho-D-glycerate + NAD⁺ + L-glutamate + H₂O → L-serine + NADH + 2-oxoglutarate + phosphate

[001029] Considering the production of two NADH from DHAP to 3-phosphohydroxypyruvate and one NADH required for fixation of NH₃, Ser production produces exactly one excess NADH, which is needed for the equimolar co-production of one MEG (or glycolic acid).

[001030] In some embodiments, the production of MEG (or glycolic acid) and L-serine is very close to the thermodynamic maximum yield potential using the lossless conversion of one or more hexose to one or more pentose-5-phosphate intermediates, followed by the conversion of the one or more pentose-5-phosphate intermediates to one or more pentose-1-phosphate and/or pentose intermediates, and then followed by the co-production of MEG (or glycolic acid) via a C2 pathway and L-serine via a C3 pathway from the one or more pentose-1-phosphate and/or pentose intermediates. In some embodiments, the thermodynamic yield potential is 14% better for co-production of MEG (or glycolic acid) and L-serine via the pathways disclosed in the present application compared to production of L-serine made from glucose by the natural, standard C3 pathway.

[001031] Co-production: hexose + NH₃ → MEG (or glycolic acid) + Ser + 0 ATP

[001032] Y(pathway) = (0.371+0.629) g/g = 1.00 g (MEG (or glycolic acid)+Ser)/g(hexose+NH₃), 96% of Y(max)(heat of combustion) = 1.044 g/g

[001033] Standard pathway: glucose + 2 NH₃ → 2 Ser + 2 NADH + 0 ATP

[001034] Y(pathway) = 0.981 g (Ser)/g(glucose+2NH₃), 84% of Y(max)(heat of combustion) = 1.164 g/g

[001035] In some embodiments, MEG (or glycolic acid) and serine are co-produced from the lossless transformation of one or more hexose feedstock to one or more pentose-5-phosphate intermediate, followed by a conversion of the one or more pentose-5-phosphate

intermediate to one or more pentose and/or pentose-1-phosphate intermediate, followed by a conversion of the one or more pentose and/or pentose-1-phosphate intermediate to MEG (or glycolic acid) and DHAP or pyruvate via a C2 pathway, and a conversion of DHAP or pyruvate to serine via a C3 pathway. In some embodiments, the C2 pathway comprises
 5 embodiment [H]. In further embodiments, the C2 pathway comprises embodiment [I]. In yet further embodiments, the C2 pathway comprises embodiment [J].

[001036] In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [H] and the C3 pathway for production of serine comprises embodiment [Q]. In some embodiments, the C2 pathway for production of MEG (or glycolic
 10 acid) comprises embodiment [I] and the C3 pathway for production of serine comprises embodiment [Q]. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [J] and the C3 pathway for production of serine comprises embodiment [Q].

[001037] [Q] In one embodiment, the application relates to a recombinant microorganism
 15 capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and serine, from one or more hexose feedstock and a nitrogen source, wherein the recombinant microorganism from embodiment [A] or from embodiment [B] (and optionally comprising embodiment [C]), and comprising additionally one or more of embodiments [D], [E], [F] and [G], further comprising one or more of embodiments [H], [I], and [J], further expresses one or more of
 20 the following from (a) to (h):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphoglycerate dehydrogenase activity that catalyzes the conversion of 3-phosphoglycerate to 3-phosphohydroxypyruvate;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an
 25 enzyme having 3-phosphoglycerate phosphatase activity and/or an enzyme having glycerate 3-kinase activity that catalyzes the conversion of 3-phosphoglycerate to glycerate;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 2-phosphoglycerate phosphatase activity and/or an enzyme having glycerate 2-kinase activity that catalyzes the conversion of 2-phosphoglycerate to glycerate;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an
 30 enzyme having phosphoserine aminotransferase activity that catalyzes the conversion of 3-phosphohydroxypyruvate from (a) to phospho-L-serine;

(e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphohydroxypyruvate phosphatase activity that catalyzes the conversion of 3-phosphohydroxypyruvate from (a) to hydroxypyruvate;

(f) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having phosphoserine phosphatase activity that catalyzes the conversion of phospho-L-serine from (d) to L-serine;

(g) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxypyruvate reductase activity that catalyzes the conversion of glycerate from (b) and/or (c) to hydroxypyruvate;

(h) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine-pyruvate aminotransferase activity that catalyzes the conversion of hydroxypyruvate from (e) and/or (g) to L-serine;

wherein the produced intermediate DHAP from embodiments [H] and/or [I] and/or pyruvate from embodiment [J] is converted to 3-phosphoglycerate and/or 2-phosphoglycerate through endogenous glycolysis or gluconeogenesis in the microorganism, and wherein MEG (or glycolic acid) and serine are produced.

[001038] Production of glycine (Gly)

[001039] Gly is produced by natural pathways, for instance via Ser, or variations thereof.

[001040] In one embodiment, Gly is produced via a Ser based pathway. In the Ser based pathway, 5,10-methylene tetrahydrofolate (M-THF) is produced from THF and L-serine is converted to glycine by a serine hydroxymethyltransferase. M-THF is utilized in the biosynthesis of various cellular compounds, for instance in methylation reactions.

[001041] In a preferred embodiment, M-THF can also be used to produce two more NADH or one NADH and one H₂:

[001042] $M\text{-THF} + \text{H}_2\text{O} \leftrightarrow \text{THF} + \text{formaldehyde}$ (EC 2.1.2.-; transferases such as hydroxymethyl-, formyl- and related transferases that transfers one-carbon group), subsequent oxidation of formaldehyde (EC 1.2.1.46; formaldehyde dehydrogenase) to formate and NADH, and further oxidation of formate to CO₂ and NADH (formate dehydrogenase, FDH) or CO₂ and H₂ (formate hydrogenlyase complex).

[001043] In some embodiments, this reconstitution of THF, if done partially, can be used to generate just enough NADH to perform another, distinct glycine biosynthesis route. In one embodiment, the route via the glycine cleavage system, which utilizes M-THF, NH₃, CO₂ and

NADH to synthesize glycine directly, can be used together with the serine based glycine production to utilize the generated excess M-THF to generate more glycine.

[001044] The glycine-cleavage system (GCV) is a multienzyme complex that catalyzes the reversible oxidation of glycine, yielding carbon dioxide, ammonia, 5,10-methylenetetrahydrofolate (M-THF) and a reduced pyridine nucleotide:

[001045] Glycine + THF + NAD⁺ ↔ M-THF + CO₂ + NH₃ + NADH + H⁺

[001046] Tetrahydrofolate serves as a recipient for one-carbon units generated during glycine cleavage to form the methylene group. The GCV system consists of four protein components, the P protein, H protein, T protein, and L protein. P protein (EC 1.4.4.2, glycine dehydrogenase (decarboxylating)) catalyzes the pyridoxal phosphate-dependent liberation of CO₂ from glycine, leaving a methylamine moiety. The methylamine moiety is transferred to the lipoic acid group of the H protein, which is bound to the P protein prior to decarboxylation of glycine. The T protein (EC 2.1.2.10, aminomethyltransferase) catalyzes the release of NH₃ from the methylamine group and transfers the remaining C1 unit to THF, forming 5,10-mTHF. The L protein (EC 1.8.1.4, dihydrolipoyl dehydrogenase) then oxidizes the lipoic acid component of the H protein and transfers the electrons to NAD⁺, forming NADH.

[001047] The same set of enzymes is sometimes referred to as glycine synthase when it runs in the reverse direction to form glycine. In the anaerobic bacteria, *Clostridium acidurici*, the glycine cleavage system runs mostly in the direction of glycine synthesis.

[001048] Alternatively, in another embodiment, glycine can also be produced via transamination of glyoxylate by alanine-glyoxylate aminotransferase (EC 2.6.1.44). Though in organisms like *Saccharomyces cerevisiae*, this pathway is only expressed during growth on non-fermentable carbon sources like ethanol or acetate, this pathway can be readily overexpressed in any microorganism. The amino group donor alanine gets reconstituted by transamination with glutamate (alanine transaminase, EC 2.6.1.2), which in turn gets reconstituted by fixation of ammonia by a NADH or NADPH dependent glutamate dehydrogenase (EC 1.4.1.2 or EC 1.4.1.3). In one embodiment, this glycine pathway avoids the intermediate L-serine and does not lead to the production of M-THF, but rather directly produces two reducing equivalents such as NADH.

[001049] In some embodiments, the production of MEG (or glycolic acid) and glycine is closer to the thermodynamic maximum yield potential using the lossless conversion of one or more hexose to one or more pentose-5-phosphate intermediates, followed by the conversion of the one or more pentose-5-phosphate intermediates to one or more pentose-1-phosphate and/or pentose intermediates, and then followed by the co-production of MEG

(or glycolic acid) via a C2 pathway and glycine via a C3 pathway from the one or more pentose-1-phosphate and/or pentose intermediates. In some embodiments, the thermodynamic yield potential is 37% better for co-production of MEG (or glycolic acid) and glycine via the pathways disclosed in the present application compared to production of glycine made from glucose by the natural, standard C3 pathway.

[001050] Co-production, serine pathway: hexose + NH₃ + THF → MEG (or glycolic acid) + Gly + M-THF + 0 ATP

[001051] Co-production, serine pathway, assuming THF reconstitution via formate and using FDH:

[001052] hexose + NH₃ → MEG (or glycolic acid) + Gly + 2 NADH + 0 ATP

[001053] Co-production, glyoxylate pathway: hexose + NH₃ → MEG (or glycolic acid) + Gly + 2 NAD(P)H + 0 ATP

[001054] $Y(\text{pathway}) = (0.371 + 0.449) \text{ g/g} = 0.820 \text{ g(MEG (or glycolic acid) + Gly) / g(hexose + NH}_3\text{)}$, 81% of $Y(\text{max})(\text{heat of combustion}) = 1.013 \text{ g/g}$

[001055] Co-production, serine pathway, assuming partial THF reconstitution via formate and FDH plus glycine synthesis via glycine cleavage system:

[001056] Partial THF reconstitution through NADH production: M-THF → 2/3 M-THF + 1/3 THF + 2/3 NADH

[001057] Utilization of remaining M-THF and generated NADH for glycine cleavage system:

[001058] 2/3 M-THF + 2/3 NADH + 2/3 NH₃ + 2/3 CO₂ → 2/3 Gly

[001059] hexose + 5/3 NH₃ + 2/3 CO₂ → MEG (or glycolic acid) + 5/3 Gly + 0 ATP

[001060] $y(\text{pathway}) = (0.348 + 0.698) \text{ g/g} = 1.046 \text{ g/g}$, 97% of $y(\text{max}) = 1.076 \text{ g/g}$

[001061] Standard pathway: glucose + 2 NH₃ + 2 THF → 2 Gly + 2 M-THF + 2 NADH + 0 ATP

[001062] $Y(\text{pathway}) = 0.701 \text{ g(Gly) / g(glucose + 2NH}_3\text{)}$, 59% of $Y(\text{max})(\text{heat of combustion}) = 1.197 \text{ g/g}$

[001063] In some embodiments, MEG (or glycolic acid) and glycine are co-produced from the lossless transformation of one or more hexose feedstock to one or more pentose-5-phosphate intermediate, followed by a conversion of the one or more pentose-5-phosphate intermediate to one or more pentose and/or pentose-1-phosphate intermediate, followed by a conversion of the one or more pentose and/or pentose-1-phosphate intermediate to MEG

(or glycolic acid) and DHAP or pyruvate via a C2 pathway, and a conversion of DHAP or pyruvate to glycine via a C3 pathway. In some embodiments, the C2 pathway comprises embodiment [H]. In further embodiments, the C2 pathway comprises embodiment [I]. In yet further embodiments, the C2 pathway comprises embodiment [J].

5 **[001064]** In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [H] and the C3 pathway for production of glycine comprises embodiment [R]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [H] and the C3 pathway for production of glycine comprises embodiment [S]. In further embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [H] and the C3 pathway for production of glycine comprises embodiment [T]. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [I] and the C3 pathway for production of glycine comprises embodiment [R]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [I] and the C3 pathway for production of glycine comprises embodiment [S]. In further embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [I] and the C3 pathway for production of glycine comprises embodiment [T]. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [J] and the C3 pathway for production of glycine comprises embodiment [R]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [J] and the C3 pathway for production of glycine comprises embodiment [S]. In further embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [J] and the C3 pathway for production of glycine comprises embodiment [T].

[001065] [R] In one embodiment, the application relates to a recombinant microorganism capable of producing monoethylene glycol (MEG) (or glycolic acid) and glycine, from one or more hexose feedstock and a nitrogen source, wherein the recombinant microorganism from embodiment [A] or from embodiment [B] (and optionally comprising embodiment [C]), and comprising additionally one or more of embodiments [D], [E], [F] and [G], further comprising one or more of embodiments [H], [I], and [J], further expresses one or more of the following from (a) to (e):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine hydroxymethyltransferase activity that catalyzes the conversion of L-serine and tetrahydrofolate (THF) to glycine and 5,10-methylene tetrahydrofolate (M-THF);

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having transferase activity that catalyzes the conversion of M-THF from (a) to formaldehyde;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having formaldehyde dehydrogenase activity that catalyzes the conversion of formaldehyde from (b) to formate and NADH;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having formate dehydrogenase activity that catalyzes the conversion of formate from (c) to CO₂ and NADH;

(e) at least one endogenous or exogenous nucleic acid molecule encoding a protein of the glycine cleavage system that catalyze the conversion of M-THF from (a), CO₂, NH₃ and NADH from (c) or (d) to glycine and THF;

wherein THF is reconstituted from steps (b) through (e), wherein optionally formate from (c) is further oxidized to CO₂ and H₂ by a formate hydrogenaselyase complex, and wherein MEG (or glycolic acid) and glycine are produced.

[001066] [S] In one embodiment, the application relates to a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and glycine, from one or more hexose feedstock and a nitrogen source, wherein the recombinant microorganism from embodiment **[A]** or from embodiment **[B]** (and optionally comprising embodiment **[C]**), and comprising additionally one or more of embodiments **[D]**, **[E]**, **[F]** and **[G]**, further comprising one or more of embodiments **[H]**, **[I]**, and **[J]**, further expresses one or more of the following from (a) to (k):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphoglycerate dehydrogenase activity that catalyzes the conversion of 3-phosphoglycerate to 3-phosphohydroxypyruvate;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having phosphoserine aminotransferase activity that catalyzes the conversion of 3-phosphohydroxypyruvate from (a) to phospho-L-serine;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphohydroxypyruvate phosphatase activity that catalyzes the conversion of 3-phosphohydroxypyruvate from (a) to hydroxypyruvate;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having phosphoserine phosphatase activity that catalyzes the conversion of phospho-L-serine from (b) to L-serine;

(e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having L-serine transaminase or serine oxidase activity that catalyzes the conversion of L-serine from (d) to hydroxypyruvate;

(f) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxypyruvate decarboxylase activity that catalyzes the conversion of hydroxypyruvate from (c) or (e) to glycolaldehyde;

(g) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde dehydrogenase activity that catalyzes the conversion of glycolaldehyde from (f) to glycolic acid;

(h) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolate dehydrogenase activity that catalyzes the conversion of glycolic acid from (g) to glyoxylate;

(i) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having alanine-glyoxylate aminotransferase activity that catalyzes the conversion of glyoxylate from (h) and alanine to glycine and pyruvate;

(j) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having alanine transaminase activity that catalyzes the conversion of pyruvate from (i) and glutamate to alanine and 2-oxoglutarate;

(k) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having NAD(P)H dependent glutamate dehydrogenase activity that catalyzes the conversion of 2-oxoglutarate from (j) and ammonia to glutamate;

wherein the produced intermediate DHAP from embodiments [H] and/or [I] and/or pyruvate from embodiment [J] is converted to 3-phosphoglycerate and/or 2-phosphoglycerate through endogenous glycolysis or gluconeogenesis in the microorganism, wherein the glyoxylate for step (i) optionally comes from glyoxylate shunt in the microorganism, wherein alanine and glutamate are reconstituted from steps (j) and (k), and wherein MEG (or glycolic acid) and glycine are co-produced.

[001067] [T] In one embodiment, the application relates to a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and glycine, from one or more hexose feedstock and a nitrogen source, wherein the recombinant microorganism from embodiment [A] or from embodiment [B] (and optionally comprising embodiment [C]), and comprising additionally one or more of embodiments [D], [E], [F] and [G], further comprising one or more of embodiments [H], [I], and [J], further expresses one or more of the following from (a) to (l):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 2-phosphoglycerate phosphatase activity and/or an enzyme having glycerate-2-kinase activity that catalyzes the conversion of 2-phosphoglycerate to glycerate;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphoglycerate phosphatase activity and/or an enzyme having glycerate-3-kinase activity that catalyzes the conversion of 3-phosphoglycerate to glycerate;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxypyruvate reductase activity that catalyzes the conversion of glycerate from (a) and/or (b) to hydroxypyruvate;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine aminotransferase or serine oxidoreductase (deaminating) activity that catalyzes the conversion of L-serine to hydroxypyruvate;

(e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having L-serine decarboxylase activity that catalyzes the conversion of L-serine to ethanolamine;

(f) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxypyruvate decarboxylase activity that catalyzes the conversion of hydroxypyruvate from (c) and/or (d) to glycolaldehyde;

(g) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having ethanolamine aminotransferase activity or an enzyme having ethanolamine oxidoreductase (deaminating) activity that catalyzes the conversion of ethanolamine from (e) to glycolaldehyde;

(h) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde dehydrogenase activity that catalyzes the conversion of glycolaldehyde from (f) and/or (g) to glycolic acid;

(i) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolate dehydrogenase activity that catalyzes the conversion of glycolic acid from (g) to glyoxylate;

(j) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having alanine-glyoxylate aminotransferase activity that catalyzes the conversion of glyoxylate from (i) and alanine to glycine and pyruvate;

(k) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having alanine transaminase activity that catalyzes the conversion of pyruvate from (j) and glutamate to alanine and 2-oxoglutarate;

(l) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having NAD(P)H dependent glutamate dehydrogenase activity that catalyzes the conversion of 2-oxoglutarate from (k) and ammonia to glutamate;

wherein the produced intermediate DHAP from embodiments [H] and/or [I] and/or pyruvate from embodiment [J] is converted to 3-phosphoglycerate and/or 2-phosphoglycerate through endogenous glycolysis or gluconeogenesis in the microorganism, wherein the glyoxylate for step (j) optionally comes from glyoxylate shunt in the microorganism, wherein alanine and glutamate are reconstituted from steps (k) and (l), and wherein MEG (or glycolic acid) and glycine are co-produced.

[001068] In some embodiments, an enzyme having serine hydroxymethyltransferase activity converts L-serine to glycine. In some embodiments, the enzyme that converts L-serine to glycine is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *E. coli* glyA. In some embodiments, the one or more nucleic acid molecule encoding an enzyme having serine hydroxymethyltransferase activity comprises an amino acid sequence set forth in UniProt ID P0A825. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having serine hydroxymethyltransferase activity is encoded by a nucleic acid sequence set forth in Gene ID 947022.

[001069] In some embodiments, an enzyme having transferase activity that transfers one-carbon groups is used to convert M-THF to formaldehyde. Transferases such as the hydroxymethyl-, formyl- and related transferases may be used. Examples of hydroxymethyl-, formyl- and related transferases include glycine hydroxymethyltransferase, phosphoribosylglycinamide formyltransferase, phosphoribosylaminoimidazolecarboxamide formyltransferase, glycine formimidoyltransferase, glutamate formiminotransferase, D-alanine 2-hydroxymethyltransferase, deoxycytidylate 5-hydroxymethyltransferase, methionyl-tRNA formyltransferase, aminomethyltransferase, 3-methyl-2-oxobutanoate hydroxymethyltransferase and UDP-4-amino-4-deoxy-L-arabinose formyltransferase.

[001070] In some embodiments, an enzyme having formaldehyde dehydrogenase activity is used to convert formaldehyde to formate and NADH. In some embodiments, the enzyme having formaldehyde dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to a formaldehyde dehydrogenase selected from *Saccharomyces cerevisiae* ALD2, *Saccharomyces cerevisiae* ALD3, *Homo sapiens* ALDH3A2, and *Homo sapiens* ALDH9A1. In some embodiments, the one or more nucleic acid molecule encoding an enzyme having formaldehyde dehydrogenase activity comprises an amino acid sequence

selected from UniProt ID P47771, UniProt ID P54114, UniProt ID P51648 and UniProt ID P49189. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having formaldehyde dehydrogenase activity is encoded by a nucleic acid sequence selected from Gene ID 855206, Gene ID 855205, Gene ID 224 and Gene ID 223.

- 5 **[001071]** In some embodiments, an enzyme having formate dehydrogenase activity is used to convert formate to CO₂ and NADH. In some embodiments, the enzyme having formate dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to an enzyme having formate dehydrogenase activity selected from the group
- 10 consisting of *E.coli* fdhF (chIF, FDH-H), *E. coli* FDH-N, *E. coli* FDH-O, *Candida boidinii* FDH1, *Corynebacterium glutamicum* fdhF, *Cupriavidus oxalaticus* NAD⁺-dependent formate dehydrogenase, *Gottschalkia acidurici* NAD⁺-dependent formate dehydrogenase, *Methylobacterium extorquens* Fdh1, *Methylosinus trichosporium* formate dehydrogenase, and *Moraxella* sp. NAD⁺-dependent formate dehydrogenase fdh. In some embodiments, the
- 15 one or more nucleic acid molecule encoding an enzyme or subunit of an enzyme associated with formate dehydrogenase activity comprises an amino acid sequence selected from UniProt ID P07658, UniProt ID P0AEK7, UniProt ID P0AAJ3, UniProt ID P24183, UniProt ID P32176, UniProt ID P0AAJ5, UniProt ID P0AEL0, UniProt ID O13437, UniProt ID Q8NSY6, UniProt ID Q8KTI7, UniProt ID Q8KTI8, and UniProt ID O08375. In a further embodiment,
- 20 the one or more nucleic acid molecule encoding an enzyme or subunit of an enzyme associated with formate dehydrogenase activity comprises a nucleic acid sequence selected from Gene ID 948584, Gene ID 946038, Gene ID 948794, Gene ID 946035, Gene ID 948394, Gene ID 948395, Gene ID 948383, GenBank accession AJ011046.2, Gene ID 1021531, GenBank accession AF489516, and GenBank accession Y13245.1.
- 25 **[001072]** In some embodiments, one or more proteins of the glycine cleavage system are used to produce glycine from M-THF, CO₂, NH₃ and NADH. In some embodiments, the proteins of the glycine cleavage system comprise: i) P-protein (a pyridoxal phosphate-containing protein) or glycine decarboxylase (EC 1.4.4.2), ii) T-protein or aminomethyl-transferase (EC 2.1.2.10), iii) L-protein or dihydrolipoamide dehydrogenase (EC1.8.1.4), and
- 30 iv) a carrier protein called H-protein (a lipoic acid-containing protein). In some embodiments, the enzyme having glycine decarboxylase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *E. coli* gcvP. In some embodiments, the one or more nucleic acid molecule encoding an enzyme having glycine decarboxylase activity comprises an
- 35 amino acid sequence set forth in UniProt ID P33195. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having glycine decarboxylase activity

comprises a nucleic acid sequence set forth in Gene ID 947394. In some embodiments, the enzyme having aminomethyltransferase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *E. coli* gcvT. In some embodiments, the one or more nucleic acid molecule encoding an enzyme having aminomethyltransferase activity comprises an amino acid sequence set forth in UniProt ID P27248. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having aminomethyltransferase activity comprises a nucleic acid sequence set forth in Gene ID 947390. In some embodiments, the enzyme having dihydrolipoamide dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *E. coli* lpd (lpdA, E3 subunit). In some embodiments, the one or more nucleic acid molecule encoding an enzyme having dihydrolipoamide dehydrogenase activity comprises an amino acid sequence set forth in UniProt ID P0A9P0. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having dihydrolipoamide dehydrogenase activity comprises a nucleic acid sequence set forth in Gene ID 944854. In some embodiments, the H-protein is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *E. coli* gcvH. In some embodiments, the one or more nucleic acid molecule encoding an H-protein comprises an amino acid sequence set forth in UniProt ID P0A6T9. In a further embodiment, the one or more nucleic acid molecule encoding an H-protein is encoded by a nucleic acid sequence set forth in Gene ID 947393.

[001073] In some embodiments, an enzyme having glycolate dehydrogenase activity is used to convert glycolic acid to glyoxylate. In some embodiments, the enzyme having glycolate dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to an enzyme having glycolate dehydrogenase activity selected from *E. coli* glycolate dehydrogenase GLC and *Arabidopsis thaliana* glycolate dehydrogenase. In some embodiments, the one or more nucleic acid molecule encoding an enzyme or enzyme subunit associated with glycolate dehydrogenase activity comprises an amino acid sequence selected from UniProt ID P0AEP9, UniProt ID P52073, UniProt ID P52074, and UniProt ID Q94AX4. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme or enzyme subunit associated with glycolate dehydrogenase activity comprises a nucleic acid sequence selected from Gene ID 947353, Gene ID 2847718, Gene ID 2847717, and GenBank accession Y13245.1.

[001074] In some embodiments, an enzyme having alanine-glyoxylate aminotransferase activity is used to convert glyoxylate to glycine. In some embodiments, the enzyme having alanine-glyoxylate aminotransferase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to an enzyme having alanine-glyoxylate aminotransferase activity selected from *Saccharomyces cerevisiae* AGX1, *Homo sapiens* AGXT2, *Arabidopsis thaliana* AOAT1 and *Arabidopsis thaliana* AOAT2. In some embodiments, the one or more nucleic acid molecule encoding an enzyme having alanine-glyoxylate aminotransferase activity comprises an amino acid sequence selected from UniProt ID P43567, UniProt ID Q9BYV1, UniProt ID Q9LR30 and UniProt ID Q9S7E9. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having alanine-glyoxylate aminotransferase activity comprises a nucleic acid sequence selected from Gene ID 850514, Gene ID 64902, TAIR accession AT1G23310 and TAIR accession AT1G70580.

[001075] In some embodiments, an enzyme having alanine transaminase activity is used to reconstitute alanine from pyruvate and glutamate. In some embodiments, the enzyme having alanine transaminase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to an enzyme having alanine transaminase activity selected from *E. coli* glutamate-pyruvate aminotransferase alaA, *E. coli* glutamate-pyruvate aminotransferase alaB, *E. coli* glutamate-pyruvate aminotransferase alaC, *Homo sapiens* alanine aminotransferase 1 (GPT), *Homo sapiens* alanine aminotransferase 2 (GPT2), *Arenicola marina* alanine transaminase, *Arabidopsis thaliana* tryptophan aminotransferase TAA1, *Arabidopsis thaliana* AOAT1, *Arabidopsis thaliana* AOAT2, *Candida maltosa* alanine aminotransferase, *Clostridium propionicum* alanine aminotransferase, *Pyrococcus furiosus* alanine aminotransferase aat, *Megathyrus maximus* alanine transaminase, and *Panicum miliaceum* alanine transaminase AlaAT-2. In some embodiments, the one or more nucleic acid molecule encoding an enzyme having alanine transaminase activity comprises an amino acid sequence selected from UniProt ID P0A959, UniProt ID P77434, UniProt ID P24298, UniProt ID Q8TD30, UniProt ID Q9S7N2, UniProt ID Q9LR30, UniProt ID Q9S7E9, UniProt ID Q9P9M8, and UniProt ID P34106. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having alanine transaminase activity comprises a nucleic acid sequence selected from Gene ID 946772, Gene ID 946850, Gene ID 2875, Gene ID 84706, Gene ID 843393, TAIR accession AT1G23310, TAIR accession AT1G70580, GenBank accession AF163769.1 and GenBank accession X69421.1.

[001076] In some embodiments, an enzyme having glutamate dehydrogenase activity is used to reconstitute glutamate from ammonia and 2-oxoglutarate. In some embodiments, the

enzyme having glutamate dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to an enzyme having glutamate dehydrogenase activity selected from *Saccharomyces cerevisiae* NAD-dependent glutamate dehydrogenase GDH2, *Arabidopsis thaliana* NAD-dependent glutamate dehydrogenase GDH2, *Arabidopsis thaliana* NAD-dependent glutamate dehydrogenase GDH1, *Peptoniphilus asaccharolyticus* NAD-dependent glutamate dehydrogenase gdhA, *Halobacterium salinarum* NAD-dependent glutamate dehydrogenase gdhA, *Thermotoga maritima* glutamate dehydrogenase, *Homo sapiens* glutamate dehydrogenase 1 (GLUD1), *Homo sapiens* glutamate dehydrogenase 2 (GLUD2), *Bacillus subtilis* glutamate dehydrogenase and *Solanum lycopersicum* glutamate dehydrogenase GDH1. In some embodiments, the one or more nucleic acid molecule encoding an enzyme having glutamate dehydrogenase activity comprises an amino acid sequence selected from UniProt ID P33327, UniProt ID Q38946, UniProt ID Q38946, UniProt ID P28997, UniProt ID P29051, UniProt ID P00367, UniProt ID P49448 and UniProt ID P93541. In a further embodiment, the one or more nucleic acid molecule encoding an enzyme having glutamate dehydrogenase activity is encoded by a nucleic acid sequence selected from Gene ID 461927, TAIR accession AT5G07440, TAIR accession AT5G18170, GenBank accession M76403.1, GenBank accession X63837.1, Gene ID 2746, Gene ID 2747 and GenBank accession U48695.1.

[001077] Production of monoethanolamine (MEA)

[001078] MEA can be produced via decarboxylation of Ser or transamination of glycolaldehyde.

[001079] In some preferred embodiments, serine decarboxylases are utilized, since they are found naturally, i.e. in choline biosynthesis pathways in plants, and the transamination of a glycolaldehyde intermediate would create a cross talk between the MEG and MEA pathway.

[001080] Alternatively, in another embodiment, MEA may be formed by ethanolamine ammonia lyase (EC 4.3.1.7) from acetaldehyde and ammonia:

[001081] Acetaldehyde + NH₃ ↔ ethanolamine

[001082] In this case, MEA is not formed via the Ser biosynthesis pathway, but rather from acetyl-CoA and its reduction to acetaldehyde by acetaldehyde dehydrogenase. While the redox situation does not change, this pathway yields +1 ATP versus the Ser based pathway. It also avoids the toxic intermediate Ser, but has the toxic and volatile intermediate acetaldehyde.

[001083] In some embodiments, the production of MEG (or glycolic acid) and MEA is very close to the thermodynamic maximum yield potential using the lossless conversion of one or more hexose to one or more pentose-5-phosphate intermediates, followed by the conversion of the one or more pentose-5-phosphate intermediates to one or more pentose-1-phosphate and/or pentose intermediates, and then followed by the co-production of MEG (or glycolic acid) via a C2 pathway and MEA via a C3 pathway from the one or more pentose-1-phosphate and/or pentose intermediates. In some embodiments, the thermodynamic yield potential is 15% better for co-production of MEG (or glycolic acid) and MEA via the pathways disclosed in the present application compared to production of MEA made from glucose by natural or published similar pathways.

[001084] Co-production, Ser pathway: hexose +NH₃ -> MEG (or glycolic acid) +MEA +0 ATP

[001085] Co-production, acetaldehyde pathway: hexose +NH₃ -> MEG (or glycolic acid) +MEA +1 ATP

[001086] $Y(\text{pathway}) = (0.371 + 0.365) \text{ g/g} = 0.736 \text{ g(MEG (or glycolic acid) + MEA) / g(hexose + NH}_3\text{)}$, 98% of $Y(\text{max})(\text{heat of combustion}) = 0.749 \text{ g/g}$

[001087] Standard pathway: glucose +2 NH₃ -> 2 MEA +2 NADH +0 ATP

[001088] $Y(\text{pathway}) = 0.570 \text{ g(MEA) / g(glucose + 2NH}_3\text{)}$, 85% of $Y(\text{max})(\text{heat of combustion}) = 0.669 \text{ g/g}$

[001089] In some embodiments, MEG (or glycolic acid) and MEA are co-produced from one or more hexose feedstock by the lossless transformation of the one or more hexose feedstock to one or more pentose-5-phosphate intermediate, followed by a conversion of the one or more pentose-5-phosphate intermediate to one or more pentose and/or pentose-1-phosphate intermediate, followed by a conversion of the one or more pentose and/or pentose-1-phosphate intermediate to MEG (or glycolic acid) and DHAP or pyruvate via a C2 pathway, and a conversion of DHAP or pyruvate to MEA via a C3 pathway.

[001090] In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[H]** and the C3 pathway for production of MEA comprises embodiment **[U]**. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[H]** and the C3 pathway for production of MEA comprises embodiment **[V]**. In further embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[H]** and the C3 pathway for production of MEA comprises embodiment **[W]**. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[I]** and the C3 pathway for production of MEA comprises

embodiment [U]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [I] and the C3 pathway for production of MEA comprises embodiment [V]. In further embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [I] and the C3 pathway for production of MEA comprises embodiment [W]. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [J] and the C3 pathway for production of MEA comprises embodiment [U]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [J] and the C3 pathway for production of MEA comprises embodiment [V]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [J] and the C3 pathway for production of MEA comprises embodiment [W].

[001091] [U] In one embodiment, the application relates to a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and monoethanolamine (MEA), from one or more hexose feedstock and a nitrogen source, wherein the recombinant microorganism from embodiment [A] or from embodiment [B] (and optionally comprising embodiment [C]), and comprising additionally one or more of embodiments [D], [E], [F] and [G], further comprising one or more of embodiments [H], [I], and [J], further expresses one or more of the following from (a) to (i):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphoglycerate dehydrogenase activity that catalyzes the conversion of 3-phosphoglycerate to 3-phosphohydroxypyruvate;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphoglycerate phosphatase activity and/or an enzyme having glycerate 3-kinase activity that catalyzes the conversion of 3-phosphoglycerate to glycerate;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 2-phosphoglycerate phosphatase activity and/or an enzyme having glycerate 2-kinase activity that catalyzes the conversion of 2-phosphoglycerate to glycerate;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having phosphoserine aminotransferase activity that catalyzes the conversion of 3-phosphohydroxypyruvate from (a) to phospho-L-serine;

(e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphohydroxypyruvate phosphatase activity that catalyzes the conversion of 3-phosphohydroxypyruvate from (a) to hydroxypyruvate;

(f) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having phosphoserine phosphatase activity that catalyzes the conversion of phospho-L-serine from (d) to L-serine;

(g) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxypyruvate reductase activity that catalyzes the conversion of glycerate from (b) and/or (c) to hydroxypyruvate;

(h) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine-pyruvate aminotransferase activity that catalyzes the conversion of hydroxypyruvate from (e) and/or (g) to L-serine;

(i) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine decarboxylase activity that catalyzes the conversion of L-serine from (f) and/or (h) to MEA;

wherein the produced intermediate DHAP from embodiments [H] and/or [I] and/or pyruvate from embodiment [J] is converted to 3-phosphoglycerate and/or 2-phosphoglycerate through endogenous glycolysis or gluconeogenesis in the microorganism, and wherein MEG (or glycolic acid) and MEA are co-produced.

[001092] [V] In one embodiment, the application relates to a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and monoethanolamine (MEA) from one or more hexose feedstock and a nitrogen source, wherein the recombinant microorganism from embodiment [A] or from embodiment [B] (and optionally comprising embodiment [C]), and comprising additionally one or more of embodiments [D], [E], [F], and [G], further comprising one or more of embodiments [H], [I], and [J], further expresses one or more of the following from (a) to (f):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 2-phosphoglycerate phosphatase activity and/or an enzyme having glycerate-2-kinase activity that catalyzes the conversion of 2-phosphoglycerate to glycerate;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphoglycerate phosphatase activity and/or an enzyme having glycerate-3-kinase activity that catalyzes the conversion of 3-phosphoglycerate to glycerate;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxypyruvate reductase activity that catalyzes the conversion of glycerate from (a) and/or (b) to hydroxypyruvate;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine-pyruvate aminotransferase or serine oxidoreductase (deaminating) activity that catalyzes the conversion of L-serine to hydroxypyruvate;

(e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxypyruvate decarboxylase activity that catalyzes the conversion of hydroxypyruvate from (c) and/or (d) to glycolaldehyde;

(f) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having transaminase activity that catalyzes the conversion of glycolaldehyde from (e) to MEA;

wherein the produced intermediate DHAP from embodiments [H] and/or [I] and/or pyruvate from embodiment [J] is converted to 3-phosphoglycerate and/or 2-phosphoglycerate through endogenous glycolysis or gluconeogenesis in the microorganism, and wherein MEG (or glycolic acid) and MEA are co-produced.

[001093] [W] In one embodiment, the application relates to a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and monoethanolamine (MEA) from one or more hexose feedstock and a nitrogen source, wherein the recombinant microorganism from embodiment [A] or from embodiment [B] (and optionally comprising embodiment [C]), and comprising additionally one or more of embodiments [D], [E], [F], and [G], further comprising one or more of embodiments [H], [I], and [J], further expresses one or more of the following from (a) to (b):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having acetaldehyde dehydrogenase activity that catalyzes the conversion of acetyl-CoA to acetaldehyde;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having ethanolamine ammonia lyase activity that catalyzes the conversion of acetaldehyde and ammonia to MEA;

wherein the produced intermediate DHAP from embodiments [H] and/or [I] and/or pyruvate from embodiment [J] is converted to acetyl-CoA through endogenous glycolysis in the microorganism, and wherein MEG (or glycolic acid) and MEA are co-produced.

[001094] In some embodiments, an enzyme having acetaldehyde dehydrogenase activity is used to reduce acetyl-CoA to acetaldehyde. In some embodiments, the enzyme having acetaldehyde dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to an enzyme having acetaldehyde dehydrogenase activity selected from

E. coli mhpF, *E. coli* AdhE, *Chlamydomonas reinhardtii* ADH1, *Leuconostoc mesenteroides* CoA-dependent acetaldehyde dehydrogenase, *Pelobacter acetylenicus* acetaldehyde dehydrogenase, *Pseudomonas* sp. dmpF, *Pseudomonas putida* acylating aldehyde dehydrogenase todI, *Pseudomonas putida* acetaldehyde dehydrogenase cmtH and
 5 *Clostridium acetobutylicum* alcohol/aldehyde dehydrogenase AdhE. In some embodiments, the one or more nucleic acid molecule encoding an enzyme having acetaldehyde dehydrogenase activity comprises an amino acid sequence selected from UniProt ID P77580, UniProt P0A9Q7, UniProt ID A8JI07, UniProt ID Q52060, UniProt ID Q51949 and UniProt ID P33744. In a further embodiment, the one or more nucleic acid molecule
 10 encoding an enzyme having acetaldehyde dehydrogenase activity comprises a nucleic acid sequence selected from Gene ID 945008, Gene ID 945837, Gene ID 5729132, GenBank accession X60835.1, GenBank accession U09250.1 and Gene ID 1116167.

[001095] In some embodiments, an enzyme having ethanolamine ammonia lyase activity is used to convert acetaldehyde and ammonia to MEA. In some embodiments, the enzyme
 15 having ethanolamine ammonia lyase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to an *E. coli* ethanolamine ammonia lyase. In some embodiments, the one or more nucleic acid molecule encoding an ethanolamine ammonia lyase subunit comprises an amino acid sequence selected from UniProt ID P0AEJ6 and UniProt ID
 20 P19636. In a further embodiment, the one or more nucleic acid molecule encoding an ethanolamine ammonia lyase subunit is encoded by a nucleic acid sequence selected from Gene ID 946924 and Gene ID 946925.

[001096] Production of ethylenediamine (EDA)

[001097] EDA can be produced by any of the pathways A through E described in WO
 25 2014/049382, which is herein incorporated in its entirety (**FIG. 11**).

[001098] In some embodiments, MEG (or glycolic acid) and EDA are co-produced from one or more hexose feedstock by the lossless transformation of the one or more hexose feedstock to one or more pentose-5-phosphate intermediate, followed by a conversion of the one or more pentose-5-phosphate intermediate to one or more pentose and/or pentose-1-
 30 phosphate intermediate, followed by a conversion of the one or more pentose and/or pentose-1-phosphate intermediate to MEG (or glycolic acid) and DHAP or pyruvate via a C2 pathway, and a conversion of DHAP or pyruvate to EDA via a C3 pathway.

[001099] In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[H]** and the C3 pathway for production of EDA comprises
 35 embodiment **[X]**. In other embodiments, the C2 pathway for production of MEG (or glycolic

acid) comprises embodiment [H] and the C3 pathway for production of EDA comprises embodiment [Y]. In further embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [H] and the C3 pathway for production of EDA comprises embodiment [Z]. In yet further embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [H] and the C3 pathway for production of EDA comprises embodiment [AA]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [H] and the C3 pathway for production of EDA comprises embodiment [BB].

[001100] In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [I] and the C3 pathway for production of EDA comprises embodiment [X]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [I] and the C3 pathway for production of EDA comprises embodiment [Y]. In further embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [I] and the C3 pathway for production of EDA comprises embodiment [Z]. In yet further embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [I] and the C3 pathway for production of EDA comprises embodiment [AA]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [I] and the C3 pathway for production of EDA comprises embodiment [BB].

[001101] In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [J] and the C3 pathway for production of EDA comprises embodiment [X]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [J] and the C3 pathway for production of EDA comprises embodiment [Y]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [J] and the C3 pathway for production of EDA comprises embodiment [Z]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [J] and the C3 pathway for production of EDA comprises embodiment [AA]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [J] and the C3 pathway for production of EDA comprises embodiment [BB].

[001102] [X] In one embodiment, the application relates to a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and ethylenediamine (EDA), from one or more hexose feedstock and a nitrogen source, wherein the recombinant microorganism from embodiment [A] or from embodiment [B] (and optionally comprising embodiment [C]), and comprising additionally one or more of embodiments [D], [E], [F] and

[G], further comprising one or more of embodiments [H], [I], and [J], further expresses one or more of the following from (a) to (c):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine dehydrogenase activity that catalyzes the conversion of L-serine to 2-aminomalonate semialdehyde;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 2-aminomalonate semialdehyde decarboxylase activity that catalyzes the conversion of 2-aminomalonate semialdehyde from (a) to aminoacetaldehyde;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having aminoacetaldehyde transaminase activity that catalyzes the conversion of aminoacetaldehyde from (b) to EDA;

wherein 2-aminomalonate semialdehyde may optionally be converted to aminoacetaldehyde by a spontaneous reaction, and wherein MEG (or glycolic acid) and EDA are co-produced.

[001103] According to this aspect of the disclosure, the recombinant microorganism overexpresses at least one of the genes encoding enzymes exhibiting activity of serine dehydrogenase and aminoacetaldehyde transaminase. These genes may be endogenous genes or exogenous genes.

[001104] The first reaction of the conversion of L-serine into 2-aminomalonate semialdehyde is catalysed by a serine dehydrogenase enzyme. This enzyme belongs to the large enzyme family of alcohol dehydrogenases also called aldehyde reductases. Several enzymes are known to exhibit serine dehydrogenase activity. In one embodiment of the disclosure, these enzymes are encoded by genes chosen among a list of genes well known in the art (Chowdhury et al., 1996, Yao et al., 2010, Tchigvintsev et al., 2012, Fujisawa et al., 2003, Hawes et al., 1996 and Lokanath et al., 2005), including but not limited to the genes listed here: *mmsB* from *Pseudomonas putida*, from *Synechococcus* PCC6301 or from *Bacillus cereus*; *hibdh* from *Pseudomonas putida* E23; PA0743 from *Pseudomonas aeruginosa*; *ydfG* from *Escherichia coli* or from *Bacillus brevis* or from *Bacillus subtilis*; *sdh* from *Agrobacterium tumefaciens*; *hibadh* from *Rattus norvegicus* or from *Thermus thermophilus* HB8; *yiaY* from *Escherichia coli*.

[001105] In a preferred embodiment of the disclosure, the serine dehydrogenase is encoded by *ydfG* from *Escherichia coli* or *mmsB* from *Pseudomonas putida*, or *yiaY* from *Escherichia coli*. Preferably, these enzymes are optimized by mutating the encoding genes in order to improve their catalytic efficiency of L-serine into 2-aminomalonate semialdehyde.

[001106] In another embodiment of the disclosure the serine dehydrogenase enzyme is obtained by evolving enzymes in order to modify their substrate specificity and/or their catalytic efficiency to obtain an enzyme which exhibits specificity for serine and activity of serine dehydrogenase. These enzymes are selected among the group of enzymes having the same type of catalytic activity on substrates chemically similar to L-serine. Preferably these enzymes may be chosen among 3-hydroxyisobutyrate dehydrogenases and serine dehydrogenases. More preferably they are encoded by genes chosen among a list of genes well known in the art, including but not limited to the genes here: *gldA* from *Escherichia coli* or from *Leuconostoc citreum* or from *Symbiobacterium thermophilum*; *yqhE* from *Escherichia coli*; *yafB* from *Escherichia coli*; *alr* from *Leishmania donovani*; *sakR1* from *Synechococcus sp.*; *yhdN* from *Bacillus subtilis*; *ytbE* from *Bacillus subtilis*; *AKR4C9* from *Arabidopsis thaliana*; *fucO* from *Escherichia coli*. Any polypeptide having at least 90% sequence identity to any of the polypeptides encoded by these genes may be used.

[001107] Evolution of these enzymes is carried out by means and methods well known by one skilled in the art in order to obtain enzyme having improved specificity for the substrate L-serine and/or enabling to convert it into 2-aminomalonate semialdehyde with an improved activity. The selection of the evolved enzymes is performed by expressing the evolved enzymes in the microorganism of the disclosure or *in vitro* with L-serine as substrate and by detecting the product 2-aminomalonate semialdehyde.

[001108] The second reaction of the conversion of 2-aminomalonate semialdehyde into aminoacetaldehyde is performed spontaneously in the cell (Fujisawa et al., 2003). In another embodiment of the disclosure, the second reaction of conversion of 2-aminomalonate semialdehyde into aminoacetaldehyde is catalysed by an enzyme having 2-aminomalonate semialdehyde decarboxylase activity. This enzyme is not encountered naturally. Therefore it is obtained by evolution of known enzyme or by screening metagenomic libraries. The 2-aminomalonate semialdehyde decarboxylase activity is performed with an evolved amino acid decarboxylase or an evolved keto-acid decarboxylase which catalyses the decarboxylation of amino acids or keto acids. Preferably an evolved amino acid decarboxylase is chosen. More preferably the evolved amino acid decarboxylase is chosen among histidine decarboxylase, serine decarboxylase, aspartate decarboxylase, diaminobutanoate decarboxylase, ornithine decarboxylase. These enzymes are encoded by genes chosen among a list of genes well known in the art, including but not limited to the genes listed here: *sdc* from *Arabidopsis thaliana*; *panD* from *Aquifex aeolicus* or from *Bacillus subtilis*; *GAD* or *GAD2* or *GAD3* or *GAD4* or *GAD5* from *Arabidopsis thaliana*; *GAD* or *GAD2* or *OAZ1* or *ODC1* from *Bos Taurus*; *gadA* or *gadB* or *panD* or *speC* or *speF* from *Escherichia coli*; *SCC105.13* from *Streptomyces coelicolor*; *gadB* from *Mannheimia*

succiniciproducens; bdb from *Haloferax volcanii*; odc1 from *Lactobacillus* sp.; kivD from *Lactococcus lactis* subsp. *Lactis*; kdcA from *Lactococcus lactis*; OAZ1 or ODC1 from *Bos Taurus*; speC or speF from *Escherichia coli*; SPE1 from *Saccharomyces cerevisiae*. Any polypeptide having at least 90% sequence identity to any of the polypeptides encoded by these genes may be used. Preferably, the sdc gene from *Arabidopsis thaliana* is used for obtaining the 2-aminomalonate semialdehyde decarboxylase activity.

[001109] Evolution of these enzymes is carried out by means and methods well known by one skilled in the art in order to obtain enzymes having specificity for the substrate 2-aminomalonate semialdehyde and enabling to convert it into aminoacetaldehyde. The selection of the evolved enzyme is performed by expressing the evolved enzyme in the microorganism of the invention or *in vitro* with 2-aminomalonate semialdehyde as substrate and by quantifying the product aminoacetaldehyde.

[001110] The last reaction of the conversion of aminoacetaldehyde into ethylenediamine is catalysed by an aminoacetaldehyde transaminase. This enzyme is not encountered naturally. Therefore it is obtained by evolution of a known enzyme or by screening metagenomic libraries. In one embodiment of the invention, the aminoacetaldehyde transaminase activity is performed with an evolved transaminase or aminotransferase which catalyses the exchange of an amino group of one molecule with an oxo group on another molecule. Preferably, the evolved aminotransferase is chosen among phosphoserine aminotransferase or aspartate aminotransferase or glutamate aminotransferase. More preferably, the evolved aminotransferase is chosen among aminotransferases using glutamate as amino group donor. These enzymes are encoded by genes chosen among a list of genes well known in the art, including but not limited to the genes listed here: serC from *Escherichia coli* or from *Bacillus subtilis* or from *Corynebacterium glutamicum*; GOT1 from *Sus scrofa*; patA from *Escherichia coli*; ygjG from *Brucella canis*; rocD from *Rhizobium* NGR 234 or from *Streptomyces avermitilis*; SCO1284 from *Streptomyces coelicolor*; AGT or AGT2 or AGT3 or GGT1 from *Arabidopsis thaliana*; AGXT from *Bos taurus*. Any polypeptide having at least 90% sequence identity to any of the polypeptides encoded by these genes may be used. Preferably, genes serC from *Escherichia coli* or GOT1 from *Sus scrofa* are used for obtaining the aminoacetaldehyde transaminase activity.

[001111] Evolution of these enzymes is carried out by means and methods well known by one skilled in the art in order to obtain enzymes having specificity for the substrate aminoacetaldehyde and enabling to convert it into ethylenediamine. The selection of the evolved enzyme is done by expressing the evolved enzyme in the microorganism of the invention or *in vitro* with aminoacetaldehyde as substrate and by detecting the product ethylenediamine.

[001112] In another embodiment of the disclosure, aminoacetaldehyde transaminase enzymes can be isolated from strains growing on ethylenediamine as sole carbon and nitrogen source. For this purpose enrichment cultures from environmental samples on ethylenediamine are cultivated on minimal medium with ethylenediamine as sole nitrogen and carbon source. Metagenomic libraries are generated from these cultures and screened for the presence of aminoacetaldehyde transaminase enzymes. This approach allows isolating the gene corresponding to the enzymatic activity and is well-known to the expert in the field.

[001113] According to a specific aspect of the disclosure, the microorganism from embodiment **[X]** is engineered to overexpress at least one of the following genes: ydfG gene or mmsB gene or yiaY gene, encoding for the serine dehydrogenase; and/or an evolved serC gene or GOT1 gene, encoding for the aminoacetaldehyde transaminase activity.

[001114] **[Y]** In one embodiment, the application relates to a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and ethylenediamine (EDA), from one or more hexose feedstock and a nitrogen source, wherein the recombinant microorganism from embodiment **[A]** or from embodiment **[B]** (and optionally comprising embodiment **[C]**), and comprising additionally one or more of embodiments **[D]**, **[E]**, **[F]** and **[G]**, further comprising one or more of embodiments **[H]**, **[I]**, and **[J]**, further expresses one or more of the following from (a) to (c):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine dehydrogenase activity that catalyzes the conversion of L-serine to 2-aminomalonate semialdehyde;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 2-aminomalonate semialdehyde transaminase activity that catalyzes the conversion of 2-aminomalonate semialdehyde from (a) to 2,3-diaminopropanoate;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 2,3-diaminopropanoate decarboxylase activity that catalyzes the conversion of 2,3-diaminopropanoate from (b) to EDA;

wherein MEG (or glycolic acid) and EDA are co-produced.

[001115] According to this aspect of the disclosure, the recombinant microorganism overexpresses at least one of the genes encoding enzymes exhibiting activity of serine dehydrogenase, 2-aminomalonate semialdehyde transaminase and 2,3-diaminopropanoate decarboxylase. These genes may be endogenous genes or exogenous genes.

[001116] The first reaction of conversion of L-serine into 2-aminomalonate semialdehyde is catalysed by a serine dehydrogenase enzyme. This enzyme belongs to the large enzyme family of alcohol dehydrogenases also called aldehyde reductases. Several enzymes are known to exhibit serine dehydrogenase activity. In one embodiment of the disclosure, the serine dehydrogenase is chosen among these known enzymes. These enzymes are encoded by genes chosen among a list of genes well known in the art (Chowdhury et al., 1996, Yao et al., 2010, Tchigvintsev et al., 2012, Fujisawa et al., 2003, Hawes et al., 1996 and Lokanath et al., 2005), including but not limited to the genes listed here: *mmsB* from *Pseudomonas putida*, from *Synechococcus* PCC6301 or from *Bacillus cereus*; *hibdh* from *Pseudomonas putida* E23; PA0743 from *Pseudomonas aeruginosa*; *ydfG* from *Escherichia coli* or from *Bacillus brevis* or from *Bacillus subtilis*; *sdh* from *Agrobacterium tumefaciens*; *hibadh* from *Rattus norvegicus* or from *Thermus thermophilus* HB8; *yiaY* from *Escherichia coli*.

[001117] In a preferred embodiment of the disclosure, the serine dehydrogenase is encoded by *ydfG* from *Escherichia coli* or *mmsB* from *Pseudomonas putida*, or *yiaY* from *Escherichia coli*. Preferably, these enzymes are optimized by mutating the encoding genes in order to improve their catalytic efficiency of L-serine into 2-aminomalonate semialdehyde.

[001118] In another embodiment of the disclosure the serine dehydrogenase enzyme is obtained by evolving enzymes in order to modify their substrate specificity and/or their catalytic efficiency to obtain an enzyme which exhibits specificity for serine and activity of serine dehydrogenase. These enzymes are selected among the group of enzymes having the same type of catalytic activity on substrates chemically similar to L-serine. Preferably these enzymes may be chosen among 3-hydroxyisobutyrate dehydrogenases and serine dehydrogenases. More preferably they are encoded by genes chosen among a list of genes well known in the art, including but not limited to the genes here: *gldA* from *Escherichia coli* or from *Leuconostoc citreum* or from *Symbiobacterium thermophilum*; *yqhE* from *Escherichia coli*; *yafB* from *Escherichia coli*; *alr* from *Leishmania donovani*; *sakR1* from *Synechococcus sp.*; *yhdN* from *Bacillus subtilis*; *ytbE* from *Bacillus subtilis*; AKR4C9 from *Arabidopsis thaliana*; *fucO* from *Escherichia coli*. Any polypeptide having at least 90% sequence identity to any of the polypeptides encoded by these genes may be used.

[001119] Evolution of these enzymes is carried out by means and methods well known by one skilled in the art in order to obtain enzyme having improved specificity for the substrate L-serine and/or enabling to convert it into 2-aminomalonate semialdehyde with an improved activity. The selection of the evolved enzymes is performed by expressing the evolved

enzymes in the microorganism of the disclosure or *in vitro* with L-serine as substrate and by detecting the product 2-aminomalonate semialdehyde.

[001120] The second reaction of conversion of 2-aminomalonate semialdehyde into 2,3-diaminopropanoate is catalysed by a 2-aminomalonate semialdehyde transaminase. This enzyme is not encountered naturally. Therefore it is obtained by evolution of a known enzyme or by screening metagenomic libraries. The 2-aminomalonate semialdehyde transaminase activity is performed with an evolved transaminase or aminotransferase which catalyses the exchange of an amino group of one molecule with an oxo group of another molecule. Preferably the evolved aminotransferase is chosen among phosphoserine aminotransferase or aspartate aminotransferase or glutamate aminotransferase. More preferably, the evolved aminotransferase is chosen among aminotransferase using glutamate as amino group donor. These enzymes are encoded by genes chosen among a list of genes well known in the art, including but not limited to the genes listed here: serC from *Escherichia coli* or from *Bacillus subtilis* or from *Corynebacterium glutamicum*; GOT1 from *Sus scrofa*; patA from *Escherichia coli*; ygjG from *Brucella canis*; rocD from *Rhizobium* NGR 234 or from *Streptomyces avermitilis*; SCO1284 from *Streptomyces coelicolor*; AGT or AGT2 or AGT3 or GGT1 from *Arabidopsis thaliana*; AGXT from *Bos taurus*. Any polypeptide having at least 90% sequence identity to any of the polypeptides encoded by these genes may be used. Preferably, genes serC from *Escherichia coli* or GOT1 from *Sus scrofa* are used for obtaining the 2-aminomalonate semialdehyde transaminase activity.

[001121] Evolution of these enzymes is carried out by means and methods well known by the man skilled in the art in order to obtain enzymes having improved specificity for the substrate 2-aminomalonate semialdehyde and/or enabling to convert it into 2,3-diaminopropanoate with an improved activity. The selection of the evolved enzyme is done by expressing the evolved enzyme in the microorganism of the invention or *in vitro* with 2-aminomalonate semialdehyde as substrate and by quantifying the product 2,3-diaminopropanoate.

[001122] The third reaction of conversion of 2,3-diaminopropanoate into ethylenediamine is catalysed by an enzyme having 2,3-diaminopropanoate decarboxylase activity. This enzyme is not encountered naturally. Therefore it is obtained by evolution of known enzyme or by screening metagenomic libraries. The 2,3-diaminopropanoate decarboxylase activity is performed with an evolved amino acid decarboxylase or an evolved keto-acid decarboxylase which catalyses the decarboxylation of amino acids or keto-acids. Preferably an evolved amino acid decarboxylase is chosen. More preferably the evolved amino acid decarboxylase is chosen among histidine decarboxylase, serine decarboxylase, aspartate decarboxylase, diaminobutanoate decarboxylase, ornithine decarboxylase. These enzymes are encoded by

genes chosen among a list of genes well known in the art, including but not limited to the genes listed here: *sdc* from *Arabidopsis thaliana*; *padC* or *yclB* from *Bacillus subtilis*; *ubiD* from *Campylobacter jejuni* or from *Escherichia coli*; *PAD1* or *GAD1* or *SPE1* from *Saccharomyces cerevisiae*; *panD* from *Aquifex aeolicus* or from *Bacillus subtilis*; *GAD* or *GAD2* or *GAD3* or *GAD4* or *GAD5* from *Arabidopsis thaliana*; *GAD* or *GAD2* or *OAZ1* or *ODC1* from *Bos Taurus*; *gadA* or *gadB* or *panD* or *speC* or *speF* from *Escherichia coli*; *SCC105.13* from *Streptomyces coelicolor*; *gadB* from *Mannheimia succiniciproducens*; *bdb* from *Haloferax volcanii*; *odc1* from *Lactobacillus sp.*; *kivD* from *Lactococcus lactis subsp. Lactis*; *kdcA* from *Lactococcus lactis*; *OAZ1* or *ODC1* from *Bos Taurus*; *speC* or *speF* from *Escherichia coli*; *SPE1* from *Saccharomyces cerevisiae*. Any polypeptide having at least 90% sequence identity to any of the polypeptides encoded by these genes may be used. Preferably, the *sdc* gene from *Arabidopsis thaliana* is used for obtaining the 2,3-diaminopropanoate decarboxylase activity.

[001123] Evolution of these enzymes is carried out by means and methods well known by the man skilled in the art in order to obtain enzymes having improved specificity for the substrate 2,3-diaminopropanoate and/or enabling to convert it into ethylenediamine with an improved activity. The selection of the evolved enzyme is performed by expressing the evolved enzyme in the microorganism of the invention or *in vitro* with 2,3-diaminopropanoate as substrate and by quantifying the product ethylenediamine. According to a specific aspect of the disclosure, the microorganism from embodiment **[Y]** is engineered to overexpress: *ydfG* gene, *yiaY* gene or *mmsB* gene, encoding for the serine dehydrogenase; and/or evolved *serC* gene or *GOT1* gene, encoding for 2-aminomalonate semialdehyde transaminase activity; and/or evolved *sdc* gene from *Arabidopsis thaliana*, coding for the 2,3-diaminopropanoate decarboxylase.

[001124] **[Z]** In one embodiment, the application relates to a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and ethylenediamine (EDA), from one or more hexose feedstock and a nitrogen source, wherein the recombinant microorganism from embodiment **[A]** or from embodiment **[B]** (and optionally comprising embodiment **[C]**), and comprising additionally one or more of embodiments **[D]**, **[E]**, **[F]** and **[G]**, further comprising one or more of embodiments **[H]**, **[I]**, and **[J]**, further expresses one or more of the following from (a) to (c):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine decarboxylase activity that catalyzes the conversion of L-serine to ethanolamine;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having ethanolamine dehydrogenase activity that catalyzes the conversion of ethanolamine from (a) to aminoacetaldehyde;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having aminoacetaldehyde transaminase activity that catalyzes the conversion of aminoacetaldehyde from (b) to EDA;

wherein MEG (or glycolic acid) and EDA are co-produced.

[001125] According to this aspect of the disclosure, the recombinant microorganism overexpresses at least one of the genes encoding enzymes exhibiting activity of serine decarboxylase, ethanolamine dehydrogenase and aminoacetaldehyde transaminase. These genes may be endogenous genes or exogenous genes.

[001126] The first reaction of conversion of L-serine into ethanolamine is catalyzed by an enzyme having serine decarboxylase activity. This group of enzymes catalyses the decarboxylation of L-serine into ethanolamine. In a preferred embodiment of the invention, the serine decarboxylase is encoded by *sdC* from *Arabidopsis thaliana* (Rontein et al., 2001, W02007/144346). The conversion of L-serine into ethanolamine by the serine decarboxylase, encoded by *sdC* from *Arabidopsis thaliana*, is disclosed in particular in patent application W02007/144364, which is incorporated by reference herein. Preferably, these enzymes are optimized by mutating the encoding genes in order to improve their conversion efficiency of L-serine into ethanolamine.

[001127] In one embodiment of the disclosure, the serine decarboxylase activity is obtained by evolving enzymes in order to modify their substrate specificity and/or their catalytic efficiency to obtain an enzyme which exhibits specificity for serine and improved activity of serine decarboxylase. These enzymes are selected among the group of enzymes having the same type of catalytic activity on substrates chemically similar to ethanolamine. These enzymes are encoded by genes chosen among a list of genes well known in the art, including but not limited to the genes listed here: GAD1 or SPE1 from *Saccharomyces cerevisiae*; panD from *Aquifex aeolicus* or from *Bacillus subtilis*; GAD or GAD2 or GAD3 or GAD4 or GAD5 from *Arabidopsis thaliana*; GAD or GAD2 or OAZ1 or ODC1 from *Bos Taurus*; gadA or gadB or panD or speC or speF from *Escherichia coli*; SCC105.13 from *Streptomyces coelicolor*; gadB from *Mannheimia succiniciproducens*; bdb from *Haloferax volcanii*; odc1 from *Lactobacillus sp.*; OAZ1 or ODC1 from *Bos Taurus*; speC or speF from *Escherichia coli*; SPE1 from *Saccharomyces cerevisiae*. Any polypeptide having at least 90% sequence identity to any of the polypeptides encoded by those genes may be used.

[001128] Evolution of these enzymes is carried out by means and methods well known by one skilled in the art in order to obtain enzymes having improved specificity for the substrate L-serine and enabling to convert it into ethanolamine with an improved activity. The selection of the evolved enzyme is performed by expressing the evolved enzyme in the microorganism of the invention or *in vitro* with L-serine as substrate and by quantifying the product ethanolamine.

[001129] The second reaction of conversion of ethanolamine into aminoacetaldehyde is catalysed by an ethanolamine dehydrogenase enzyme. Natural enzymes having this activity are not disclosed in prior art; however some enzymes have low catalytic activity. Therefore it is advantageous to evolve these enzymes with low catalytic activity towards evolved enzymes with improved activity. Useful enzymes can also be obtained by screening metagenomic libraries.

[001130] In one embodiment of the disclosure, the ethanolamine dehydrogenase activity is obtained by evolving enzymes in order to modify their substrate specificity and/or their catalytic efficiency to obtain an enzyme which exhibits specificity for ethanolamine and activity of ethanolamine dehydrogenase. These enzymes are selected among the group of enzymes having the same type of catalytic activity on substrates chemically similar to ethanolamine. These enzymes are encoded by genes chosen among a list of genes well known in the art, including but not limited to the genes listed here: *mmsB* from *Pseudomonas putida*, or from *Synechococcus* PCC6301, or from *Bacillus cereus*; *hibdh* from *Pseudomonas putida* E23; PA0743 from *Pseudomonas aeruginosa*; *ydfG* from *Escherichia coli* or from *Bacillus brevis* or from *Bacillus subtilis*; *sdh* from *Agrobacterium tumefaciens*; *hibadh* from *Rattus norvegicus* or from *Thermus thermophilus* HB8; *gldA* from *Escherichia coli* or from *Leuconostoc citreum* or from *Symbiobacterium thermophilum*; *yqhE* from *Escherichia coli*; *yafB* from *Escherichia coli*; *aladh* from *Enterobacter aerogenes*; *alr* from *Leishmania donovani*; *sakR1* from *Synechococcus* sp.; *yhdN* from *Bacillus subtilis*; *ytbE* from *Bacillus subtilis*; *yiaY* from *Escherichia coli*; AKR4C9 from *Arabidopsis thaliana*; *fucO* from *Escherichia coli*. Any polypeptide having at least 90% sequence identity to any of the polypeptides encoded by these genes may be used. Preferably genes *fucO* from *Escherichia coli* or *yiaY* from *Escherichia coli* are used for obtaining the ethanolamine dehydrogenase activity.

[001131] Evolution of these enzymes is carried out by means and methods well known by one skilled in the art in order to obtain enzymes having improved specificity for the substrate ethanolamine and/or enabling to convert it into aminoacetaldehyde with an improved activity. The selection of evolved enzymes is done by expressing the evolved enzymes in the

microorganism of the invention or *in vitro* with ethanolamine as substrate and by quantifying the product aminoacetaldehyde.

[001132] The last reaction of conversion of aminoacetaldehyde into ethylenediamine is catalysed by an aminoacetaldehyde transaminase. This enzyme is not encountered naturally. Therefore it is obtained by evolution of known enzyme or by screening metagenomic libraries. In one embodiment of the disclosure, the aminoacetaldehyde transaminase activity is performed with an evolved transaminase or aminotransferase which catalyses the exchange of an amino group of one molecule with an oxo group on another molecule. Preferably, the evolved aminotransferase is chosen among phosphoserine aminotransferase or aspartate aminotransferase or glutamate aminotransferase. More preferably the evolved aminotransferase is chosen among aminotransferase using glutamate as amino group donor. These enzymes are encoded by genes chosen among a list of genes well known in the art, including but not limited to the genes listed here: serC from *Escherichia coli* or from *Bacillus subtilis* or from *Corynebacterium glutamicum*; GOT1 from *Sus scrofa*; patA from *Escherichia coli*; ygjG from *Brucella canis*; rocD from *Rhizobium* NGR 234 or from *Streptomyces avermitilis*; SCO1284 from *Streptomyces coelicolor*; AGT or AGT2 or AGT3 or GGT1 from *Arabidopsis thaliana*; AGXT from *Bos taurus*. Any polypeptide having at least 90% sequence identity to any of the polypeptides encoded by these genes may be used. Preferably, genes serC from *Escherichia coli* or GOT1 from *Sus scrofa* are used.

[001133] Evolution of these enzymes is carried out by means and methods well known by one skilled in the art in order to obtain enzymes having improved specificity for the substrate aminoacetaldehyde and/or enabling to convert it into ethylenediamine with an improved activity. The selection of the evolved enzyme is done by expressing the evolved enzyme in the microorganism of the invention or *in vitro* with aminoacetaldehyde as substrate and by detecting the product ethylenediamine.

[001134] In another embodiment of the disclosure, aminoacetaldehyde transaminase enzymes can be isolated from strains growing on ethylenediamine as sole carbon and nitrogen source. For this purpose enrichment cultures from environmental samples on ethylenediamine are cultivated on minimal medium with ethylenediamine as sole nitrogen and carbon source. Metagenomic libraries are generated from these cultures and screened for the presence of aminoacetaldehyde transaminase enzymes. This approach allows isolating the gene corresponding to the enzymatic activity and is well-known to the expert in the field.

[001135] According to a specific aspect of the disclosure, the microorganism from embodiment **[Z]** is engineered to overexpress: an sdc gene from *Arabidopsis thaliana*,

encoding a serine decarboxylase; and/or fucO or yiaY genes from *Escherichia coli*, encoding for the ethanolamine dehydrogenase activity; and/or an evolved serC gene or GOT1 gene, encoding for aminoacetaldehyde transaminase activity.

[001136] [AA] In one embodiment, the application relates to a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and ethylenediamine (EDA), from one or more hexose feedstock and a nitrogen source, wherein the recombinant microorganism from embodiment **[A]** or from embodiment **[B]** (and optionally comprising embodiment **[C]**), and comprising additionally one or more of embodiments **[D]**, **[E]**, **[F]** and **[G]**, further comprising one or more of embodiments **[H]**, **[I]**, and **[J]**, further expresses one or more of the following from (a) to (c):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine hydroxymethyltransferase activity that catalyzes the conversion of L-serine to glycine;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having aldehyde oxidase activity that catalyzes the conversion of glycine from (a) to aminoacetaldehyde;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having aminoacetaldehyde transaminase activity that catalyzes the conversion of aminoacetaldehyde from (b) to EDA;

wherein MEG (or glycolic acid) and EDA are co-produced.

[001137] Preferably gene glyA from *Escherichia coli* is used for obtaining the serine hydroxymethyltransferase activity.

[001138] The aldehyde oxidase enzyme is obtained by evolving enzymes in order to modify their substrate specificity and/or their catalytic efficiency to obtain an enzyme which exhibits specificity for glycine and activity of aldehyde oxidase. These enzymes are selected among the group of enzymes having the same type of catalytic activity on substrates chemically similar to glycine. These enzymes are encoded by genes chosen among a list of genes well known in the art, including but not limited to the genes listed here: aldH1 from *Aquifex aeolicus*; dhaS from *Anoxybacillus flavithermus*; Aldh from *Apis mellifera*; aldX, aldY, dhaS, ycbD, yfmT or ywdH from *Bacillus subtilis*; prr from *Escherichia coli*; ALD2, ALD3, ALD4, ALD5, ALD6 from *Saccharomyces cerevisiae*; betB from *Roseobacter denitrificans*; AAur_0650 from *Arthrobacter aurescens*. Any polypeptide having at least 90% sequence identity to any of the polypeptides encoded by these genes may be used.

[001139] Conversion of aminoacetaldehyde into ethylenediamine by an enzyme having aminoacetaldehyde transaminase activity. This enzyme is obtained by evolving enzymes in order to modify their substrate specificity and/or their catalytic efficiency to obtain an enzyme which exhibits specificity for aminoacetaldehyde and activity of transaminase. These enzymes are selected among the group of enzymes having the same type of catalytic activity on substrates chemically similar to aminoacetaldehyde. These enzymes are encoded by genes chosen among a list of genes well known in the art, including but not limited to the genes listed here: serC from *Escherichia coli* or from *Bacillus subtilis* or from *Corynebacterium glutamicum*; GOT1 from *Sus scrofa*; patA from *Escherichia coli*; yggG from *Brucella canis*; rocD from *Rhizobium* NGR 234 or from *Streptomyces avermitilis*; SCO1284 from *Streptomyces coelicolor*; AGT or AGT2 or AGT3 or GGT1 from *Arabidopsis thaliana*; AGXT from *Bos taurus*. Any polypeptide having at least 90% sequence identity to any of the polypeptides encoded by these genes may be used.

[001140] According to a specific aspect of the disclosure, the microorganism from embodiment **[AA]** is engineered to overexpress an evolved serC gene or GOT1 gene, encoding for aminoacetaldehyde transaminase activity.

[001141] [BB] In one embodiment, the application relates to a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and ethylenediamine (EDA), from one or more hexose feedstock and a nitrogen source, wherein the recombinant microorganism from embodiment **[A]** or from embodiment **[B]** (and optionally comprising embodiment **[C]**), and comprising additionally one or more of embodiments **[D]**, **[E]**, **[F]** and **[G]**, further comprising one or more of embodiments **[H]**, **[I]**, and **[J]**, further expresses one or more of the following from (a) to (e):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having an amino acid N-acetyl transferase activity or O-acetyl transferase activity that catalyzes the conversion of L-serine to N-acetylserine;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having N-acetylserine dehydrogenase activity that catalyzes the conversion of N-acetylserine from (a) to N-acetylmalonate semialdehyde;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having transaminase activity that catalyzes the conversion of N-acetylmalonate semialdehyde from (b) to acetylaminopropanoate;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having deacetylase activity that catalyzes the conversion of acetylaminopropanoate from (c) to 2,3-diaminopropanoate;

(e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 2,3-diaminopropanoate decarboxylase activity that catalyzes the conversion of 2,3-diaminopropanoate from (d) to EDA;

wherein MEG (or glycolic acid) and EDA are co-produced.

5 **[001142]** The first conversion step of embodiment **[BB]** may be an amino acid N-acetyl transferase activity or O-acetyl transferase activity, since the transformation of O to N is spontaneous. Preferably gene *argA* from *Escherichia coli* is used for obtaining the amino acid N-acetyl transferase activity.

[001143] An enzyme having N-acetylserine dehydrogenase activity is obtained by evolving
10 enzymes in order to modify their substrate specificity and/or their catalytic efficiency to obtain an enzyme which exhibits specificity for N-acetylserine and activity of N-acetylserine dehydrogenase. These enzymes are selected among the group of enzymes having the same type of catalytic activity on substrates chemically similar to N-acetylserine. These enzymes are encoded by genes chosen among a list of genes well known in the art,
15 including but not limited to the genes listed here: *mmsB* from *Pseudomonas putida*, or from *Synechococcus* PCC6301, or from *Bacillus cereus*; *hibdh* from *Pseudomonas putida* E23; PA0743 from *Pseudomonas aeruginosa*; *ydfG* from *Escherichia coli* or from *Bacillus brevis* or from *Bacillus subtilis*; *sdh* from *Agrobacterium tumefaciens*; *hibadh* from *Rattus norvegicus* or from *Thermus thermophilus* HB8; *gldA* from *Escherichia coli* or from
20 *Leuconostoc citreum* or from *Symbiobacterium thermophilum*; *yqhE* from *Escherichia coli*; *yafB* from *Escherichia coli*; *aladh* from *Enterobacter aerogenes*; *alr* from *Leishmania donovani*; *sakR1* from *Synechococcus* sp.; *yhdN* from *Bacillus subtilis*; *ytbE* from *Bacillus subtilis*; *yiaY* from *Escherichia coli*; AKR4C9 from *Arabidopsis thaliana*; *fucO* from *Escherichia coli*. Any polypeptide having at least 90% sequence identity to any of the
25 polypeptides encoded by these genes may be used.

[001144] An enzyme having a transaminase activity to convert N-acetylmalonate semialdehyde to acetylaminopropanoate may be obtained by evolving enzymes in order to modify their substrate specificity and/or their catalytic efficiency to obtain an enzyme which exhibits specificity for N-acetylmalonate semialdehyde and activity of N-acetylmalonate
30 semialdehyde transaminase. These enzymes are selected among the group of enzymes having the same type of catalytic activity on substrates chemically similar to N-acetylmalonate semialdehyde. These enzymes are encoded by genes chosen among a list of genes well known in the art, including but not limited to the genes listed here: *serC* from *Escherichia coli* or from *Bacillus subtilis* or from *Corynebacterium glutamicum*; GOT1 from
35 *Sus scrofa*; *patA* from *Escherichia coli*; *ygjG* from *Brucella canis*; *rocD* from *Rhizobium* NGR

234 or from *Streptomyces avermitilis*; 2SCG18.31c from *Streptomyces coelicolor*; AGT or AGT2 or AGT3 or GGT1 from *Arabidopsis thaliana*; AGXT from *Bos taurus*. Any polypeptide having at least 90% sequence identity to any of the polypeptides encoded by these genes may be used.

- 5 **[001145]** Preferably gene *argE* from *Escherichia coli* is used for obtaining the deacetylase activity to convert acetylaminopropanoate into 2,3-diaminopropanoate.

[001146] An enzyme having amino-acid decarboxylase activity or keto acid decarboxylase activity to convert 2,3-diaminopropanoate into ethylenediamine may be obtained by evolving enzymes in order to modify their substrate specificity and/or their catalytic efficiency to obtain
 10 an enzyme which exhibits specificity for 2,3-diaminopropanoate and activity of amino acid decarboxylase or keto acid decarboxylase. These enzymes are selected among the group of enzymes having the same type of catalytic activity on substrates chemically similar to 2,3-diaminopropanoate. These enzymes are encoded by gene chosen among a list of genes well known in the art, including but not limited to the genes listed here: *sdc* from *Arabidopsis*
 15 *thaliana*; *padC* or *yclB* from *Bacillus subtilis*; *ubiD* from *Campylobacter jejuni* or from *Escherichia coli*; *PAD1* or *GAD1* or *SPE1* from *Saccharomyces cerevisiae*; *panD* from *Aquifex aeolicus* or from *Bacillus subtilis*; *GAD* or *GAD2* or *GAD3* or *GAD4* or *GAD5* from *Arabidopsis thaliana*; *GAD* or *GAD2* or *OAZ1* or *ODC1* from *Bos Taurus*; *gadA* or *gadB* or *panD* or *speC* or *speF* from *Escherichia coli*; *SCC105.13* from *Streptomyces coelicolor*;
 20 *gadB* from *Mannheimia succiniciproducens*; *bdb* from *Haloferax volcanii*; *odc1* from *Lactobacillus sp.*; *kivD* from *Lactococcus lactis subsp. Lactis*; *kdcA* from *Lactococcus lactis*; *OAZ1* or *ODC1* from *Bos Taurus*; *speC* or *speF* from *Escherichia coli*; *SPE1* from *Saccharomyces cerevisiae*. Any polypeptide having at least 90% sequence identity to any of the polypeptides encoded by these genes may be used.

- 25 **[001147]** In a further embodiment of the disclosure, the method is performed with a microorganism wherein serine biosynthesis is optimized. This optimization is disclosed in particular in patent application WO 2007/144346, which is incorporated by reference herein.

[001148] Alternatively, in another embodiment, EDA can be produced by the following process: Ser can be directly aminated to (S)-2,3-diaminopropanoate by serine aminase (EC
 30 2.6.1.-), then decarboxylated to EDA, for instance by an enzyme from the family of L-2,4-diaminobutyrate or ornithine decarboxylases (**FIG. 11**). However, if such an enzyme with (S)-2,3-diaminopropanoate decarboxylase activity is not specific, it may also act on other amino acids or serine itself.

[001149] In some embodiments, EDA can be produced by the following process: the intermediate (S)-2,3-diaminopropanoate may also be produced by direct amination of pyruvate using (S)-2,3-diaminopropanoate ammonia lyase (EC 4.3.1.15) (**FIG. 11**):

[001150] 2,3-diaminopropanoate + H₂O ↔ 2 NH₃ + pyruvate

5 **[001151]** In some embodiments, the production of MEG (or glycolic acid) and EDA is very close to the thermodynamic maximum yield potential using the lossless conversion of one or more hexose to one or more pentose-5-phosphate intermediates, followed by the conversion of the one or more pentose-5-phosphate intermediates to one or more pentose-1-phosphate and/or pentose intermediates, and then followed by the co-production of MEG (or glycolic acid) via a C2 pathway and EDA via a C3 pathway from the one or more pentose-1-phosphate and/or pentose intermediates. In some embodiments, the thermodynamic yield potential is 14% better for co-production of MEG (or glycolic acid) and EDA via the pathways disclosed in the present application compared to production of EDA made from glucose by natural or published similar pathways.

15 **[001152]** Co-production: hexose + 2 NH₃ → MEG (or glycolic acid) + EDA + 0 ATP

[001153] Y(pathway) = (0.337 + 0.326) g/g = 0.663 g(MEG (or glycolic acid) + EDA)/g(hexose + 2NH₃), 97% of Y(max)(heat of combustion) = 0.687 g/g

[001154] Standard pathway: glucose + 4 NH₃ → 2 EDA + 2 NADH + 0 ATP

20 **[001155]** Y(pathway) = 0.484 g(EDA)/g(glucose + 4NH₃), 85% of Y(max)(heat of combustion) = 0.571 g/g

[001156] In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[H]** and the C3 pathway for production of EDA comprises embodiment **[CC]**. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[H]** and the C3 pathway for production of EDA comprises embodiment **[DD]**. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[I]** and the C3 pathway for production of EDA comprises embodiment **[CC]**. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[I]** and the C3 pathway for production of EDA comprises embodiment **[DD]**. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[J]** and the C3 pathway for production of EDA comprises embodiment **[CC]**. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[J]** and the C3 pathway for production of EDA comprises embodiment **[DD]**.

[001157] [CC] In one embodiment, the application relates to a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and ethylenediamine (EDA), from one or more hexose feedstock and a nitrogen source, wherein the recombinant microorganism from embodiment **[A]** or from embodiment **[B]** (and optionally comprising embodiment **[C]**), and comprising additionally one or more of embodiments **[D]**, **[E]**, **[F]** and **[G]**, further comprising one or more of embodiments **[H]**, **[I]**, and **[J]**, further expresses one or more of the following from (a) to (b):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having a serine aminase activity that catalyzes the conversion of L-serine to (S)-2,3-diaminopropanoate;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having an (S)-2,3-diaminopropanoate decarboxylase activity that catalyzes the conversion of (S)-2,3-diaminopropanoate from (a) to EDA;

wherein MEG (or glycolic acid) and EDA are co-produced.

[001158] [DD] In one embodiment, the application relates to a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and ethylenediamine (EDA), from one or more hexose feedstock and a nitrogen source, wherein the recombinant microorganism from embodiment **[A]** or from embodiment **[B]** (and optionally comprising embodiment **[C]**), and comprising additionally one or more of embodiments **[D]**, **[E]**, **[F]** and **[G]**, further comprising one or more of embodiments **[H]**, **[I]**, and **[J]**, further expresses one or more of the following from (a) to (b):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having an (S)-2,3-diaminopropanoate ammonia lyase activity that catalyzes the conversion of pyruvate and ammonium to (S)-2,3-diaminopropanoate;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having (S)-2,3-diaminopropanoate decarboxylase activity that catalyzes the conversion of (S)-2,3-diaminopropanoate from (a) to EDA;

wherein pyruvate is produced from endogenous glycolysis, and wherein MEG (or glycolic acid) and EDA are co-produced.

[001159] In some embodiments, an enzyme having 2,3-diaminopropionate ammonia-lyase activity is used to convert pyruvate and ammonium to (S)-2,3-diaminopropanoate. In some embodiments, an enzyme having 2,3-diaminopropionate ammonia-lyase activity is encoded by an amino acid sequence having at least 70% sequence identity, having at least 80% sequence identity, or having at least 90% sequence identity to *E. coli* 2,3-diaminopropionate

ammonia-lyase ygeX. In other embodiments, the enzyme having 2,3-diaminopropionate ammonia-lyase activity is *E. coli* ygeX. In some embodiments, the one or more nucleic acid molecule encoding an enzyme having 2,3-diaminopropionate ammonia-lyase activity comprises an amino acid sequence set forth in UniProt ID P66899. In further embodiments, the one or more nucleic acid molecule encoding an enzyme having 2,3-diaminopropionate ammonia-lyase activity is encoded by a nucleic acid sequence set forth in Gene ID 947012.

[001160] In one embodiment of any aspect disclosed above, the enzyme having glycolaldehyde reductase activity is encoded by one or more nucleic acid molecules obtained from a microorganism selected from *E. coli* and *S. cerevisiae*. In another embodiment, the one or more nucleic acid molecules is selected from gldA, GRE2, GRE3, yqhD, ydjG, fucO, yafB (dkgB), and/or yqhE (dkgA), or homolog thereof. In another embodiment, the one or more nucleic acid molecules is yqhD. In some embodiments, the yqhD comprises a G149E mutation. In a further embodiment, the enzyme having glycolaldehyde reductase activity comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 13, 15, 17, 20, 23, 25, 28, 30 and 32. In yet a further embodiment, the enzyme having glycolaldehyde reductase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 12, 14, 16, 18, 19, 21, 22, 24, 26, 27, 29 and 31.

[001161] In one embodiment of any aspect disclosed above, the enzyme having thiolase or acetyl coenzyme A acetyltransferase activity is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Clostridium* sp., *Bacillus* sp., *E. coli*, *Saccharomyces* sp. and *Marinobacter* sp. In some embodiments, the enzyme having thiolase or acetyl coenzyme A acetyltransferase activity is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Clostridium acetobutylicum*, *Clostridium thermosaccharolyticum*, *Bacillus cereus*, *E. coli*, *Saccharomyces cerevisiae* and *Marinobacter hydrocarbonoclasticus*. In some embodiments, the one or more nucleic acid molecules is thlA, atoB and/or ERG10, or homolog thereof. In a further embodiment, the enzyme having thiolase or acetyl coenzyme A acetyltransferase activity comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 35, 37 and 40. In yet a further embodiment, the enzyme having thiolase or acetyl coenzyme A acetyltransferase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 33, 34, 36, 38 and 39.

[001162] In one embodiment of any aspect disclosed above, the enzyme having acetyl-CoA:acetoacetate-CoA transferase or acetate:acetoacetyl-CoA hydrolase activity is encoded by one or more nucleic acid molecules obtained from a microorganism selected from *Clostridium* sp. and *E. coli*. In another embodiment, the enzyme having acetyl-

CoA:acetoacetate-CoA transferase or acetate:acetoacetyl-CoA hydrolase activity is encoded by one or more nucleic acid molecules obtained from *E. coli*. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having acetyl-CoA:acetoacetate-CoA transferase activity is *atoA* and/or *atoD*, or homolog thereof. In another embodiment, the enzyme having acetyl-CoA:acetoacetate-CoA transferase or acetate:acetoacetyl-CoA hydrolase activity is encoded by one or more nucleic acid molecules obtained from *Clostridium acetobutylicum*. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having acetate:acetoacetyl-CoA hydrolase activity is *ctfA* and/or *ctfB*, or homolog thereof. In a further embodiment, the enzyme having acetyl-CoA:acetoacetate-CoA transferase or acetate:acetoacetyl-CoA hydrolase activity comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 43, 46, 97, 99, 101 and 103. In yet a further embodiment, the enzyme having acetyl-CoA:acetoacetate-CoA transferase or acetate:acetoacetyl-CoA hydrolase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 41, 42, 44, 45, 96, 98, 100 and 102.

[001163] In one embodiment of any aspect disclosed above, the enzyme having acetoacetate decarboxylase activity is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Clostridium sp.*, *Bacillus sp.*, *Chromobacterium sp.* and *Pseudomonas sp.* In another embodiment, the enzyme having acetoacetate decarboxylase activity is encoded by one or more nucleic acid molecules obtained from a microorganism selected from the group consisting of *Clostridium acetobutylicum*, *Clostridium beijerinckii*, *Clostridium cellulolyticum*, *Bacillus polymyxa*, *Chromobacterium violaceum* and *Pseudomonas putida*. In some embodiments, the one or more nucleic acid molecules encoding the enzyme having acetoacetate decarboxylase activity is *adc*, or homolog thereof. In a further embodiment, the enzyme having acetoacetate decarboxylase activity comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 49 and 52. In yet another embodiment, the enzyme having acetoacetate decarboxylase activity is encoded by a nucleic acid sequence selected from the group consisting of SEQ ID NOs: 47, 48, 50 and 51.

[001164] [EE] In another embodiment, the recombinant microorganism selected from embodiment **[H]**, embodiment **[I]**, and embodiment **[J]**, optionally further comprises one or more modifications selected from the group consisting of:

(i) a deletion, insertion, or loss of function mutation in a gene encoding a xylose isomerase that catalyzes the conversion of D-xylulose to D-xylose;

(ii) a deletion, insertion, or loss of function mutation in a gene encoding a glycolaldehyde dehydrogenase that catalyzes the conversion of glycolaldehyde to glycolic acid; and

(iii) a deletion, insertion, or loss of function mutation in a gene encoding a lactate dehydrogenase that catalyzes the conversion of pyruvate to lactate.

[001165] In some embodiments, a recombinant microorganism producing MEG (or glycolic acid) or, optionally, MEG (or glycolic acid) and one or more co-products, comprises a deletion, insertion, or loss of function mutation in a gene encoding a xylose isomerase to prevent the conversion of D-xylulose to D-xylose and instead shunt the reaction toward conversion of D-xylulose to D-xylulose-1-phosphate, D-xylulose-5-phosphate, or D-ribulose. In some embodiments, the xylose isomerase is from *Escherichia coli*. In some embodiments, the xylose isomerase is encoded by the xylA gene, or homolog thereof.

[001166] In some embodiments, a recombinant microorganism producing MEG (or glycolic acid) or, optionally, MEG (or glycolic acid) and one or more co-products, comprises a deletion, insertion, or loss of function mutation in a gene encoding a glycolaldehyde dehydrogenase to prevent the production of glycolic acid from glycolaldehyde and instead shunt the reaction toward conversion of glycolaldehyde to MEG. In some embodiments, the glycolaldehyde dehydrogenase is from *Escherichia coli*. In some embodiments, the glycolaldehyde dehydrogenase is encoded by the aldA gene, or homolog thereof. In some embodiments, the deletion, insertion, or loss of function mutation in a gene encoding a glycolaldehyde dehydrogenase is partial, wherein some glycolaldehyde dehydrogenase function is still present and an amount of glycolic acid is still produced.

[001167] In some embodiments, a recombinant microorganism producing MEG (or glycolic acid) or, optionally, MEG (or glycolic acid) and one or more co-products, comprises a deletion, insertion, or loss of function mutation in a gene encoding a lactate dehydrogenase to prevent the production of lactate from pyruvate and instead shunt the reaction toward production of one or more co-products. In some embodiments, the lactate dehydrogenase is from *Escherichia coli*. In some embodiments, the lactate dehydrogenase is encoded by the ldhA gene, or homolog thereof.

[001168] Non-limiting combinations of any of the recombinant microorganism and methods embodiments described herein are included as part of this disclosure.

Recombinant Microorganism

[001169] The disclosure provides microorganisms that can be engineered to express various endogenous or exogenous enzymes.

[001170] In various embodiments described herein, the recombinant microorganism is a eukaryotic microorganism. In some embodiments, the eukaryotic microorganism is a yeast.

5 In exemplary embodiments, the yeast is a member of a genus selected from the group consisting of *Yarrowia*, *Candida*, *Saccharomyces*, *Pichia*, *Hansenula*, *Kluyveromyces*, *Issatchenkia*, *Zygosaccharomyces*, *Debaryomyces*, *Schizosaccharomyces*, *Pachysolen*, *Cryptococcus*, *Trichosporon*, *Rhodotorula*, and *Myxozyma*.

[001171] In some embodiments, the recombinant microorganism is a prokaryotic microorganism. In exemplary embodiments, the prokaryotic microorganism is a member of a genus selected from the group consisting of *Escherichia*, *Clostridium*, *Zymomonas*, *Salmonella*, *Rhodococcus*, *Pseudomonas*, *Bacillus*, *Lactobacillus*, *Enterococcus*, *Alcaligenes*, *Klebsiella*, *Paenibacillus*, *Arthrobacter*, *Corynebacterium*, and *Brevibacterium*.

[001172] In some embodiments, the recombinant microorganism is used to produce
15 monoethylene glycol (MEG) or glycolic acid (GA), or optionally, MEG (or GA) and one or more co-product, disclosed herein.

[001173] Accordingly, in another aspect, the present inventions provide a method of producing MEG or GA, or optionally, MEG (or GA) and one or more co-product, using a recombinant microorganism described herein. In one embodiment, the method comprises
20 cultivating the recombinant microorganism in a culture medium containing a feedstock providing a carbon source until MEG or GA, or optionally, MEG (or GA) and one or more co-product, is produced. In a further embodiment, the MEG (or GA), or optionally, MEG (or GA) and one or more co-product, is recovered. Recovery can be by methods known in the art, such as distillation, membrane-based separation gas stripping, solvent extraction, and
25 expanded bed adsorption.

[001174] In some embodiments, the feedstock comprises a carbon source. In various embodiments described herein, the carbon source may be selected from sugars, glycerol, alcohols, organic acids, alkanes, fatty acids, lignocellulose, proteins, carbon dioxide, and carbon monoxide. In an exemplary embodiment, the carbon source is a sugar. In a further
30 exemplary embodiment, the sugar is glucose or oligomers of glucose thereof. In other embodiments, the oligomers of glucose are selected from fructose, sucrose, starch, cellobiose, maltose, lactose and cellulose.

Biosynthesis of Xylitol, or optionally, Xylitol and one or more co-product using a recombinant microorganism

[001175] In one aspect, the present disclosure provides a recombinant microorganism comprising one or more biochemical pathway that produces xylitol from one or more hexose feedstock. In one aspect, the present disclosure provides a recombinant microorganism comprising one or more biochemical pathway to produce xylitol from hexose, such as glucose, through the nonoxidative branch of the PPP (i.e., with the use of 1) transketolase, 2) transaldolase, 3) ribulose-5P epimerase and 4) ribose-5P isomerase. In another aspect, the present disclosure preferably provides a recombinant microorganism that has reduced or eliminated activity of at least one enzyme of the oxidative branch of the PPP (i.e., deletion or inactivation of the gene “zwf” in *E. coli* that codifies glucose-6-phosphate 1-dehydrogenase; and/or the gene “pgl” in *E. coli* that codifies the 6-phosphogluconolactonase; and/or gene “gnd” in *E. coli* that codifies the 6-phosphogluconate dehydrogenase, decarboxylating) in order to reduce carbon loss, resulting in an increased yield of xylitol. In one embodiment, one or more co-product is co-produced with xylitol.

[001176] Therefore, in one embodiment, the application relates to a recombinant microorganism comprising one or more biochemical pathway comprising at least one enzyme having an activity that converts one or more hexose feedstock in a lossless conversion to xylitol.

[001177] In one aspect, a microorganism of the present disclosure is capable of converting a hexose or an oligomer comprising a hexose into xylitol as a single end product. In one aspect, a microorganism of the present disclosure is capable of converting a hexose or an oligomer comprising a hexose into xylitol as a co-product.

[001178] In one aspect, a microorganism of the present disclosure is capable of (1) converting a hexose, such as glucose, to D-Ribose-5P and/or (2) D-xylulose-5P and/or D-ribulose-5P. In one aspect, a microorganism of the present disclosure is capable of converting D-Ribose-5P to D-Ribulose-5P via expression of *rpiA*. In one aspect, a microorganism of the present disclosure is capable of converting D-Ribulose-5P to D-Xylulose-5P via expression of *rpe*. In one aspect, a microorganism of the present disclosure is capable of converting D-Xylulose-5P to D-Ribulose-5P via expression of *rpe*. In one aspect, a microorganism of the present disclosure is capable of converting D-Xylulose-5P to D-Xylulose via expression of a phosphatase. In one aspect, the microorganism of the present disclosure comprises a phosphatase that has a high specificity and/or high activity on D-ribulose-5P and/or D-xylulose-5P.

[001179] In one aspect, a microorganism of the present disclosure is capable of converting D-Ribulose-5P to D-Ribulose via expression of a phosphatase. In one aspect, a microorganism of the present disclosure is capable of converting D-Ribulose to D-Xylulose. In one aspect, a microorganism of the present disclosure is capable of converting D-Xylulose to Xylitol via expression of, but not limited to, EC 1.1.1.9. In some aspects, the microorganism of the present disclosure is capable of performing any one or more of the above reactions. In some aspects, the microorganism of the present disclosure is capable of performing any one or more of the reactions described in FIG. 12. In some aspects, the microorganism of the present disclosure does not have arabitol and/or xylitol-5P as an intermediate of the metabolic pathway.

[001180] In one aspect, a microorganism of the present disclosure is capable of (1) converting a hexose, such as glucose, to D-Ribose-5P and/or (2) D-xylulose-5P and/or D-ribulose-5P. In one aspect, a microorganism of the present disclosure is capable of converting D-Ribose-5P to D-Ribulose-5P via expression of *rpiA*. In one aspect, a microorganism of the present disclosure is capable of converting D-Ribulose-5P to D-Xylulose-5P via expression of *rpe*. In one aspect, a microorganism of the present disclosure is capable of converting D-Xylulose-5P to D-Ribulose-5P via expression of *rpe*. In one aspect, a microorganism of the present disclosure is capable of converting (1) D-Ribulose-5P to D-Arabitol-5P via expression of ADP, (2) D-Arabitol-5P to D-Xylulose-1P via expression of ADP, (3) Xylulose-5P to D-Arabitol-1P via expression of ADP, and/or (4) D-Arabitol-1P to D-Ribulose-1P via expression of ADP. In one aspect, a microorganism of the present disclosure is capable of converting D-Ribulose-1P into D-Ribulose via expression of, but not limited to, EC 3.1.3.- and/or EC 2.7.1.17. In one aspect, a microorganism of the present disclosure is capable of converting D-Ribulose into D-Xylulose via expression of, but not limited to, EC 5.1.3.30 and/or EC 5.1.3.31. In one aspect, a microorganism of the present disclosure is capable of converting D-Xylulose-1P into D-Xylulose via expression of, but not limited to, EC 3.1.3.- and/or EC 2.7.1.17. In one aspect, a microorganism of the present disclosure is capable of converting D-Xylulose to Xylitol via expression of, but not limited to, EC 1.1.1.9. In some aspects, the microorganism of the present disclosure is capable of performing any one or more of the above reactions. In some aspects, the microorganism of the present disclosure is capable of performing any one or more of the reactions described in FIG. 13.

[001181] In one aspect, a microorganism of the present disclosure is capable of (1) converting a hexose, such as glucose, to D-Ribose-5P and/or (2) D-xylulose-5P and/or D-ribulose-5P. In one aspect, a microorganism of the present disclosure is capable of converting D-Ribose-5P to D-Ribulose-5P via expression of *rpiA*. In one aspect, a

microorganism of the present disclosure is capable of converting D-Ribulose-5P to D-Xylulose-5P via expression of rpe. In one aspect, a microorganism of the present disclosure is capable of converting D-Xylulose-5P to D-Ribulose-5P via expression of rpe. In one aspect, a microorganism of the present disclosure is capable of converting D-Ribulose-5P to D-Ribulose-1P via expression of PPM. In one aspect, a microorganism of the present disclosure is capable of converting D-Xylulose-5P to D-Xylulose-1P via expression of PPM. In one aspect, a microorganism of the present disclosure is capable of converting D-Ribulose-1P to D-Ribulose via expression of, but not limited to, EC 3.1.3.- and/or EC 2.7.1.17. In one aspect, a microorganism of the present disclosure is capable of converting D-Xylulose-1P to D-Xylulose via expression of, but not limited to, EC 3.1.3.- and/or EC 2.7.1.17. In one aspect, a microorganism of the present disclosure is capable of converting D-Ribulose to D-Xylulose via expression of, but not limited to, EC 5.1.3.31 and/or EC 5.1.3.30. In one aspect, a microorganism of the present disclosure is capable of converting D-Xylulose to Xylitol via expression of, but not limited to, EC 1.1.1.9. In some aspects, the microorganism of the present disclosure is capable of performing any one or more of the above reactions. In some aspects, the microorganism of the present disclosure is capable of performing any one or more of the reactions described in FIG. 14.

[001182] In some aspects, the enzymes indicated above include xylitol hydrogenase (EC 1.1.1.9) which is involved in the hydrogenation of D-xylulose to xylitol. In some aspects, the enzymes indicated above include sugar phosphatase (EC 3.1.3.-, or EC 3.1.3.23) or xylulokinase (EC 2.7.1.17) which are involved in the conversion of D-xylulose-1P to D-xylulose and/or of D-ribulose-1P to D-ribulose. In some aspects, the enzymes indicated above include D-psicose 3-epimerase (EC 5.1.3.30) or D-tagatose 3-epimerase (EC 5.1.3.31), which are involved in the conversion of D-ribulose to D-xylulose.

[001183] In some aspects, the one or more enzymes are heterologous to the microorganism. In some aspects, some of the enzymes are heterologous to the microorganism and other enzymes are native to the microorganism.

Methods of producing a recombinant microorganism that produces or accumulates MEG (or glycolic acid), or optionally, MEG (or glycolic acid) and one or more co-product

[001184] As described above, in one aspect, the present application provides methods of producing a recombinant microorganism that produces or accumulates MEG (or glycolic acid) from one or more exogenous hexose feedstock via one or more pentose-5-phosphate intermediate, comprising: introducing into or expressing in the recombinant microorganism one or more enzyme for the conversion of the one or more hexose feedstock to the one or

more pentose-5-phosphate intermediate; introducing into or expressing in the recombinant microorganism one or more enzyme for the conversion of the one or more pentose-5-phosphate intermediate to one or more pentose or pentose-1-phosphate intermediate; introducing into or expressing in the recombinant microorganism a C2 pathway comprising one or more enzyme for the production of MEG or GA from glycolaldehyde; and introducing into or expressing in the recombinant microorganism a C3 pathway comprising one or more enzyme for the production of MEG or GA derived from DHAP or pyruvate; and culturing the recombinant microorganism in a culture medium containing the one or more hexose feedstock to produce or accumulate MEG or GA, wherein glycolaldehyde and DHAP (or pyruvate) are intermediates produced in the C2 pathway, and wherein MEG or GA are produced in both the C2 and C3 pathways. In some embodiments, the glycolaldehyde is oxidized to glycolic acid by a glycolaldehyde dehydrogenase.

[001185] In another aspect, the present application provides methods of producing a recombinant microorganism that produces or accumulates MEG (or glycolic acid) and one or more co-product from one or more exogenous hexose feedstock via one or more pentose-5-phosphate intermediate, comprising: introducing into or expressing in the recombinant microorganism one or more enzyme for the conversion of the one or more hexose feedstock to the one or more pentose-5-phosphate intermediate; introducing into or expressing in the recombinant microorganism one or more enzyme for the conversion of the one or more pentose-5-phosphate intermediate to one or more pentose or pentose-1-phosphate intermediate; introducing into or expressing in the recombinant microorganism a C2 pathway comprising one or more enzyme for the production of MEG or GA from glycolaldehyde; and introducing into or expressing in the recombinant microorganism a C3 pathway comprising one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate; and culturing the recombinant microorganism in a culture medium containing the one or more hexose feedstock to produce or accumulate MEG or GA and one or more co-product, wherein glycolaldehyde and DHAP (or pyruvate) are intermediates produced in the C2 pathway, and wherein MEG or GA is produced in the C2 pathway and the one or more co-product is produced in the C3 pathway. In some embodiments, the glycolaldehyde is oxidized to glycolic acid by a glycolaldehyde dehydrogenase.

[001186] In some embodiments, the enzymes for the conversion of the one or more hexose feedstock to the one or more pentose-5-phosphate intermediate are selected from one or more enzyme having a transketolase activity, a transaldolase activity, a ribulose-5-phosphate 3-epimerase activity, and a ribose-5-phosphate isomerase activity. In further embodiments, the method further comprises introducing into the recombinant microorganism one or more modifications to diminish or delete activity in one or more endogenous enzymes

selected from glyceraldehyde 3-phosphate dehydrogenase (gapA), phosphoglycerate kinase (pgk) and phosphoglycerate mutase (gpmA and/or gpmM).

[001187] In some embodiments, the enzymes for the conversion of the one or more hexose feedstock to the one or more pentose-5-phosphate intermediate are selected from one or more enzyme having a fructose-6-phosphate phosphoketolase activity, a phosphate acetyltransferase activity, a transketolase activity, a transaldolase activity, a ribulose-5-phosphate 3-epimerase activity, and a ribose-5-phosphate isomerase activity. In further embodiments, the method further comprises introducing into the recombinant microorganism one or more modifications to diminish or delete activity in an endogenous 6-phosphofructokinase (pfkA and/or pfkB) enzyme.

[001188] In yet further embodiments, the method further comprises: introducing into the recombinant microorganism one or more modifications to delete or diminish activity in one or more endogenous enzymes selected from glucose 6-phosphate-1-dehydrogenase, 6-phosphogluconolactonase, and 6-phosphogluconate dehydrogenase. In some embodiments, the glucose 6-phosphate-1-dehydrogenase is zwf, the 6-phosphogluconolactonase is pgl, and the 6-phosphogluconate dehydrogenase is gnd.

[001189] In some embodiments, the one or more enzyme for the conversion of the one or more pentose-5-phosphate intermediate to one or more pentose or pentose-1-phosphate intermediate are selected from one or more enzyme having a pentose phosphatase activity, a D-ribulose-5-phosphatase activity, a D-ribose-5-phosphatase activity, a D-ribose isomerase activity, an arabitol phosphate dehydrogenase activity, and a phosphopentomutase activity.

[001190] In some embodiments, the one or more enzyme for the production of MEG or GA derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine transaminase activity, a hydroxypyruvate decarboxylase activity, a 3-phosphohydroxypyruvate reductase activity, a glycolaldehyde reductase activity, a glycolaldehyde dehydrogenase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a serine decarboxylase activity, an ethanolamine aminotransferase or ethanolamine oxidoreductase (deaminating) activity, a glycerate decarboxylase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, and a glycerate 2-kinase activity.

[001191] In some embodiments, the one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, and an acetoacetate decarboxylase activity, and the one or more co-product comprises acetone.

[001192] In some embodiments, the one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, and a secondary alcohol dehydrogenase activity, and the one or more co-product comprises isopropanol.

[001193] In some embodiments, the one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, a secondary alcohol dehydrogenase activity, and a dehydratase activity, and the one or more co-product comprises propene.

[001194] In some embodiments, the one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, a 3-hydroxyisovalerate (3HIV) synthase activity, a hydroxymethylglutaryl-CoA synthase activity, a methylglutaconyl-CoA hydratase activity, a methylcrotonyl-CoA carboxylase activity, a methylcrotonyl-CoA hydratase activity, a 3-hydroxyisovaleryl-CoA thioesterase activity, a 3HIV kinase activity, a 3HIV-3-phosphate decarboxylase activity, and a 3HIV decarboxylase activity, and the one or more co-product comprises isobutene.

[001195] In some embodiments, the one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a

glycerate 3-kinase activity, and a glycerate 2-kinase activity, and the one or more co-product comprises L-serine.

[001196] In some embodiments, the one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a serine hydroxymethyltransferase activity, a transferase activity, a formaldehyde dehydrogenase activity, a formate dehydrogenase activity, an activity associated with glycine cleavage system, a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine transaminase activity, a hydroxypyruvate decarboxylase activity, a serine oxidoreductase (deaminating) activity, a serine decarboxylase activity, an ethanolamine aminotransferase or ethanolamine oxidoreductase (deaminating) activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, a glycerate 2-kinase activity, a glycolaldehyde dehydrogenase activity, a glycolate dehydrogenase activity, an alanine-glyoxylate aminotransferase activity, an alanine transaminase activity, and an NAD(P)H dependent glutamate dehydrogenase activity, and the one or more co-product comprises glycine. In other embodiments, the activity associated with glycine cleavage system comprise an enzyme or protein selected from a glycine decarboxylase (P protein), an aminomethyltransferase (T protein), a dihydrolipoamide dehydrogenase (L protein), and an H protein.

[001197] In some embodiments, the one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a transaminase activity, a hydroxypyruvate decarboxylase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a serine decarboxylase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, a glycerate 2-kinase activity, an acetaldehyde dehydrogenase activity, and an ethanolamine ammonia lyase activity, and the one or more co-product comprises monoethanolamine (MEA).

[001198] In some embodiments, the one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a serine dehydrogenase activity, a 2-aminomalonate semialdehyde decarboxylase activity, an aminoacetaldehyde transaminase

activity, a 2-aminomalonate semialdehyde transaminase activity, a 2,3-diaminopropanoate decarboxylase activity, a serine decarboxylase activity, an ethanolamine dehydrogenase activity, a serine hydroxymethyltransferase activity, an aldehyde oxidase activity, an N-acetyl transferase or O-acetyl transferase activity, an N-acetylserine dehydrogenase activity, a
 5 transaminase activity, a deacetylase activity, a serine aminase activity, and a 2,3-diaminopropanoate ammonia lyase activity, and the one or more co-product comprises ethylenediamine (EDA).

[001199] In some embodiments, the one or more enzyme for the production of MEG or GA in the C2 pathway are selected from one or more enzyme having an activity selected from a
 10 D-tagatose 3-epimerase activity, a D-ribulokinase activity, a D-ribulose-1-phosphate aldolase activity, a D-xylulose 1-kinase activity, a D-xylulose-1-phosphate aldolase activity, a xylose reductase or aldose reductase activity, a xylitol dehydrogenase activity, a xylose isomerase activity, a xylose dehydrogenase activity, a xylonolactonase activity, a xylonate dehydrogenase activity, a 2-keto-3-deoxy-D-pentonate aldolase activity, a glycolaldehyde
 15 reductase activity and a glycolaldehyde dehydrogenase activity.

[001200] In some embodiments, any of the methods of producing a recombinant microorganism described above further comprises: introducing into the recombinant microorganism one or more modifications to diminish or delete activity in a glycolaldehyde dehydrogenase, a lactate dehydrogenase, a xylose isomerase, a xylulose kinase or
 20 combination thereof.

Hexose to pentose-5-phosphate intermediate

[001201] In the present disclosure, glucose flux is funneled into the pentose phosphate pathway instead of the glycolysis pathway by using a non-oxidative entry into the pentose phosphate pathway.

[001202] [mA] In one embodiment, the application relates to a method of producing a recombinant microorganism capable of producing one or more pentose-5-phosphate intermediate from one or more hexose feedstock, comprising introducing into and/or overexpressing in the recombinant microorganism one or more of the following from (a) to
 (d):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having transketolase activity that catalyzes a reversible conversion of fructose-6-phosphate and glyceraldehyde-3-phosphate to erythrose-4-phosphate and D-xylulose-5-phosphate, respectively, and/or that catalyzes a reversible conversion of glyceraldehyde-3-phosphate from (b) and seduheptulose-7-phosphate from (b) to D-ribose-5-phosphate and
 35 D-xylulose-5-phosphate, respectively;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having transaldolase activity that catalyzes a reversible conversion of fructose-6-phosphate and erythrose-4-phosphate from (a) to glyceraldehyde-3-phosphate and seduheptulose-7-phosphate, respectively;

5 (c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having ribose-5-phosphate isomerase activity that catalyzes an interconversion of D-ribose-5-phosphate from (a) and D-ribulose-5-phosphate;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having ribulose-5-phosphate 3-epimerase activity that catalyzes an interconversion
10 of D-xylulose-5-phosphate from (a) and D-ribulose-5-phosphate from (c);

wherein the method optionally further comprises introducing a deletion, insertion, or loss of function mutation in a gene encoding a glyceraldehyde 3-phosphate dehydrogenase;

wherein the one or more hexose feedstock is converted to fructose-6-phosphate and glyceraldehyde-3-phosphate through an endogenous glycolysis pathway in the recombinant
15 microorganism,

and wherein the one or more pentose-5-phosphate intermediate produced is one or more of D-ribose-5-phosphate, D-xylulose-5-phosphate or D-ribulose-5-phosphate.

[001203] [mB] In another embodiment, the application relates to a method of producing a recombinant microorganism capable of producing one or more pentose-5-phosphate
20 intermediate from one or more hexose feedstock, comprising introducing into and/or overexpressing in the recombinant microorganism one or more of the following from (a) to (f):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having fructose-6-phosphate phosphoketolase activity that catalyzes a reversible
25 conversion of fructose-6-phosphate to erythrose-4-phosphate and acetyl-phosphate;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having phosphate acetyltransferase activity that catalyzes a reversible conversion of acetyl-phosphate from (a) to acetyl-CoA;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an
30 enzyme having transaldolase activity that catalyzes a reversible conversion of fructose-6-phosphate and erythrose-4-phosphate from (a) to glyceraldehyde-3-phosphate and seduheptulose-7-phosphate, respectively;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having transketolase activity that catalyzes a reversible conversion of glyceraldehyde-3-phosphate from (c) and seduheptulose-7-phosphate from (c) to D-ribose-5-phosphate and D-xylulose-5-phosphate, respectively;

5 (e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having ribose-5-phosphate isomerase activity that catalyzes an interconversion of D-ribose-5-phosphate from (d) and D-ribulose-5-phosphate;

(f) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having ribulose-5-phosphate 3-epimerase activity that catalyzes an interconversion
10 of D-xylulose-5-phosphate from (d) and D-ribulose-5-phosphate from (e);

wherein the method optionally further comprises introducing a deletion, insertion, or loss of function mutation in a gene encoding a 6-phosphofructokinase;

wherein the one more more hexose feedstock is converted to fructose-6-phosphate through an endogenous glycolysis pathway in the recombinant microorganism,

15 wherein the acetyl-CoA produced in step (b) can be used to produce one or more co-products selected from acetone, isopropanol, propene, isobutene, and serine pathway compounds;

and wherein the one or more pentose-5-phosphate intermediate produced is one or more of D-ribose-5-phosphate, D-xylulose-5-phosphate or D-ribulose-5-phosphate.

20 **[001204]** In some embodiments, the oxidative branch of the pentose phosphate pathway is deleted or inactivated to optimize flux towards the non-oxidative entry into the pentose phosphate pathway.

[001205] [mC] Therefore, in one embodiment, the method of embodiment **[mA]** or embodiment **[mB]** optionally further comprises introducing into the recombinant
25 microorganism one or more modifications selected from the group consisting of:

(i) a deletion, insertion, or loss of function mutation in a gene encoding a glucose 6-phosphate-1-dehydrogenase that catalyzes the conversion of glucose-6-phosphate to 6-phospho-D-glucono-1,5-lactone;

(ii) a deletion, insertion, or loss of function mutation in a gene encoding a 6-
30 phosphogluconolactonase that catalyzes the conversion of 6-phospho-D-glucono-1,5-lactone to gluconate-6-phosphate; and

(iii) a deletion, insertion, or loss of function mutation in a gene encoding a 6-phosphogluconate dehydrogenase that catalyzes the conversion of gluconate-6-phosphate to D-ribulose-5-phosphate.

[001206] *Connecting pentose phosphate pathway and MEG (or glycolic acid) production pathways*

[001207] In another aspect, a pentose-5-phosphate intermediate produced by the method of embodiment **[mA]** or embodiment **[mB]** (and optionally comprising embodiment **[mC]**) can be connected with any one of the known C2 MEG or glycolic acid production pathways by pentose phosphatases.

[001208] **[mD]** Therefore, in one embodiment, the application relates to a method of producing a recombinant microorganism capable of producing D-xylulose from D-xylulose-5-phosphate, wherein the method comprises introducing into or expressing in the recombinant microorganism one or more of a pentose-5-phosphatase, wherein the D-xylulose-5-phosphate is produced by non-oxidative branch of pentose phosphate pathway, and wherein the D-xylulose can be used to produce MEG (or glycolic acid) and optionally, one or more co-product.

[001209] **[mE]** In one embodiment, the application relates to a method of producing a recombinant microorganism capable of producing D-ribulose from D-ribulose-5-phosphate or from D-ribose-5-phosphate, wherein the method further comprises introducing into or expressing in the recombinant microorganism one or more of the following from (a) to (c):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having D-ribulose-5-phosphatase activity and/or an enzyme having pentose-5-phosphatase activity that catalyzes a reversible conversion of D-ribulose-5-phosphate to D-ribulose;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having D-ribose-5-phosphatase activity and/or an enzyme having pentose-5-phosphatase activity that catalyzes a reversible conversion of D-ribose-5-phosphate to D-ribose;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having D-ribose isomerase activity that catalyzes a reversible conversion of D-ribose from (b) to D-ribulose;

wherein the D-ribulose-5-phosphate and/or D-ribose-5-phosphate is produced by non-oxidative branch of pentose phosphate pathway, and wherein the D-ribulose can be used to produce MEG (or glycolic acid) and optionally, one or more co-product.

[001210] In another aspect, a pentose-5-phosphate intermediate produced by the method of embodiment **[mA]** or embodiment **[mB]** (and optionally comprising embodiment **[mC]**) can be connected with any one of the known C2 MEG or glycolic acid production pathways by arabitol phosphate dehydrogenases.

5 **[001211] [mF]** In one embodiment, the application relates to a method of producing a recombinant microorganism capable of producing D-ribulose-1-phosphate from D-xylulose-5-phosphate or producing D-xylulose-1-phosphate from D-ribulose-5-phosphate, wherein the method comprises introducing into or expressing in the recombinant microorganism one or more arabitol phosphate dehydrogenase, wherein the one or more arabitol phosphate
10 dehydrogenase catalyzes one or more of the following from (a) to (d):

(a) a reversible conversion of D-xylulose-5-phosphate to D-arabitol-1-phosphate;

(b) a reversible conversion of D-arabitol-1-phosphate from (a) to D-ribulose-1-phosphate;

(c) a reversible conversion of D-ribulose-5-phosphate to D-arabitol-5-phosphate;

15 (d) a reversible conversion of D-arabitol-5-phosphate from (c) to D-xylulose-1-phosphate,

wherein the D-xylulose-5-phosphate and/or D-ribulose-5-phosphate is produced by non-oxidative branch of pentose phosphate pathway, and wherein the D-ribulose-1-phosphate and/or D-xylulose-1-phosphate can be used to produce MEG (or glycolic acid) and
20 optionally, one or more co-product.

[001212] In another aspect, a pentose-5-phosphate intermediate produced by the method of embodiment **[mA]** or embodiment **[mB]** (and optionally comprising embodiment **[mC]**) can be connected with any one of the known C2 MEG or glycolic acid production pathways by phosphopentomutases.

25 **[001213] [mG]** In one embodiment, the application relates to a method of producing a recombinant microorganism capable of producing D-xylulose-1-phosphate from D-xylulose-5-phosphate or producing D-ribulose-1-phosphate from D-ribulose-5-phosphate, wherein the method comprises introducing into or expressing in the recombinant microorganism one or more phosphopentomutases, wherein the D-xylulose-5-phosphate and/or D-ribulose-5-phosphate is produced by non-oxidative branch of pentose phosphate pathway, and wherein
30 the D-ribulose-1-phosphate and/or D-xylulose-1-phosphate can be used to produce MEG (or glycolic acid) and optionally, one or more co-product.

[001214] *MEG (or glycolic acid), or optionally MEG (or glycolic acid) and one or more co-product production pathways*

[001215] In some embodiments, the pentose or pentose-1-phosphate intermediates produced in embodiments **[mD]**, **[mE]**, **[mF]**, and **[mG]** are used in known MEG (or glycolic acid) C2 production pathways, which are coupled to C3 pathways, as described below, to co-produce additional MEG (or glycolic acid) or one or more co-products.

- 5 **[001216]** In some embodiments, MEG (or glycolic acid) is produced via a C2 pathway that uses D-xylulose-1-phosphate.

[001217] [mH] In one embodiment, the application relates to a method of producing a recombinant microorganism capable of producing monoethylene glycol (MEG) (or glycolic acid) from one or more hexose feedstock, wherein the method of embodiment **[mA]** or
 10 embodiment **[mB]** (and optionally comprising embodiment **[mC]**), and comprising additionally one or more of embodiments **[mD]**, **[mF]**, and **[mG]**, further comprises introducing into or reexpressing in the recombinant microorganism one or more of the following from (a) to (d):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an
 15 enzyme having D-xylulose 1-kinase activity that catalyzes the conversion of D-xylulose from embodiment **[mD]** to D-xylulose-1-phosphate;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having D-xylulose-1-phosphate aldolase activity that catalyzes the conversion of D-xylulose-1-phosphate from (a), from embodiment **[mF]** and/or from embodiment **[mG]**, to
 20 glycolaldehyde and dihydroxyacetonephosphate (DHAP);

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde reductase activity or an an enzyme having aldehyde reductase activity that catalyzes the conversion of glycolaldehyde from (b) to MEG;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an
 25 enzyme having glycolaldehyde dehydrogenase activity that catalyzes the conversion of glycolaldehyde from (b) to glycolic acid;

wherein MEG (or glycolic acid) and DHAP are produced.

[001218] In some embodiments, MEG (or glycolic acid) is produced via a C2 pathway that uses D-ribulose-1-phosphate.

30 **[001219] [mI]** In one embodiment, the application relates to a method of producing a recombinant microorganism capable of producing monoethylene glycol (MEG) (or glycolic acid) from one or more hexose feedstock, wherein the method of embodiment **[mA]** or embodiment **[mB]** (and optionally comprising embodiment **[mC]**), and comprising

additionally one or more of embodiments **[mD]**, **[mE]**, **[mF]**, and **[mG]**, further comprises introducing into or expressing in the recombinant microorganism one or more of the following from (a) to (e):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having D-tagatose 3-epimerase activity that catalyzes the conversion of D-xylulose from embodiment **[mD]** to D-ribulose;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having D-ribulokinase activity that catalyzes the conversion of D-ribulose from (a) and/or from embodiment **[mE]** to D-ribulose-1-phosphate;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having D-ribulose-1-phosphate aldolase activity that catalyzes the conversion of D-ribulose-1-phosphate from (b), from embodiment **[mF]** and/or from embodiment **[mG]** to glycolaldehyde and dihydroxyacetonephosphate (DHAP);

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde reductase activity that catalyzes the conversion of glycolaldehyde from (c) to MEG;

(e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde dehydrogenase activity that catalyzes the conversion of glycolaldehyde from (c) to glycolic acid;

wherein MEG (or glycolic acid) and DHAP are produced.

[001220] In some embodiments, MEG (or glycolic acid) is produced via a C2 pathway that uses D-xylonate.

[001221] [mJ] In one embodiment, the application relates to a method of producing a recombinant microorganism capable of producing monoethylene glycol (MEG) (or glycolic acid) from one or more hexose feedstock, wherein the method of embodiment **[mA]** or embodiment **[mB]** (and optionally comprising embodiment **[mC]**), and comprising additionally embodiment **[mD]**, further comprises introducing into and/or expressing in the recombinant microorganism one or more of the following from (a) to (d):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having xylose isomerase activity that catalyzes the conversion of D-xylulose from embodiment **[mD]** to D-xylose;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having xylose dehydrogenase activity that catalyzes the conversion of D-xylose from (a) to D-xylonolactone;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having xylonolactonase activity that catalyzes the conversion of D-xylonolactone from (b) to D-xylonate;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having xylose dehydrogenase activity that catalyzes the conversion of D-xylose from (a) to D-xylonate;

wherein the method further comprises introducing into or expressing in the recombinant microorganism one or more of the following from (e) to (h):

(e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having xylonate dehydratase activity that catalyzes the conversion of D-xylonate from (c) or (d) to 2-keto-3-deoxy-xylonate;

(f) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 2-keto-3-deoxy-D-pentonate aldolase activity that catalyzes the conversion of 2-keto-3-deoxy-xylonate from (e) to glycolaldehyde and pyruvate;

(g) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde reductase activity that catalyzes the conversion of glycolaldehyde from (f) to MEG;

(h) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde dehydrogenase activity that catalyzes the conversion of glycolaldehyde from (f) to glycolic acid;

wherein MEG (or glycolic acid) and pyruvate are produced.

[001222] In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[mH]** and the C3 pathway for production of MEG (or glycolic acid) comprises embodiment **[mK]**. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[mH]** and the C3 pathway for production of MEG (or glycolic acid) comprises embodiment **[mL]**. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[mI]** and the C3 pathway for production of MEG (or glycolic acid) comprises embodiment **[mK]**. In further embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[mI]** and the C3 pathway for production of MEG (or glycolic acid) comprises embodiment **[mL]**. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises

embodiment **[mJ]** and the C3 pathway for production of MEG (or glycolic acid) comprises embodiment **[mK]**. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[mJ]** and the C3 pathway for production of MEG (or glycolic acid) comprises embodiment **[mL]**.

- 5 **[001223] [mK]** In one embodiment, the application relates to a method of producing a recombinant microorganism capable of producing monoethylene glycol (MEG) (or glycolic acid) from one or more hexose feedstock, wherein the method of embodiment **[mA]** or embodiment **[mB]** (and optionally comprising embodiment **[mC]**), and comprising additionally one or more of embodiments **[mD]**, **[mE]**, **[mF]**, and **[mG]**, further comprising
 10 one or more of embodiments **[mH]**, **[mI]**, and **[mJ]**, further comprises introducing into and/or expressing in the recombinant microorganism one or more of the following from (a) to (h):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphoglycerate dehydrogenase activity that catalyzes the conversion of 3-phosphoglycerate to 3-phosphohydroxypyruvate;

- 15 (b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having phosphoserine aminotransferase activity that catalyzes the conversion of 3-phosphohydroxypyruvate from (a) to phospho-L-serine;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphohydroxypyruvate phosphatase activity that catalyzes the
 20 conversion of 3-phosphohydroxypyruvate from (a) to hydroxypyruvate;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having phosphoserine phosphatase activity that catalyzes the conversion of phospho-L-serine from (b) to L-serine;

- (e) at least one endogenous or exogenous nucleic acid molecule encoding an
 25 enzyme having L-serine transaminase or serine oxidase activity that catalyzes the conversion of L-serine from (d) to hydroxypyruvate;

(f) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxypyruvate decarboxylase activity that catalyzes the conversion of hydroxypyruvate from (c) or (e) to glycolaldehyde;

- 30 (g) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde reductase activity that catalyzes the conversion of glycolaldehyde from (f) to MEG;

(h) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde dehydrogenase activity that catalyzes the conversion of glycolaldehyde from (f) to glycolic acid;

wherein the produced intermediate DHAP from embodiments [mH] and/or [mI] and/or pyruvate from embodiment [mJ] is converted to 3-phosphoglycerate through endogenous glycolysis or gluconeogenesis, respectively, in the microorganism, and wherein MEG (or glycolic acid) is produced.

[001224] [mL] In one embodiment, the application relates to a method of producing a recombinant microorganism capable of producing monoethylene glycol (MEG) (or glycolic acid) from one or more hexose feedstock, wherein the method of embodiment [mA] or embodiment [mB] (and optionally comprising embodiment [mC]), and comprising additionally one or more of embodiments [mD], [mE], [mF], and [mG], further comprising one or more of embodiments [mH], [mI], and [mJ], further comprises introducing into and/or expressing in the recombinant microorganism one or more of the following from (a) to (j):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 2-phosphoglycerate phosphatase activity and/or an enzyme having glycerate-2-kinase activity that catalyzes the conversion of 2-phosphoglycerate to glycerate;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphoglycerate phosphatase activity or an enzyme having glycerate-3-kinase activity that catalyzes the conversion of 3-phosphoglycerate to glycerate;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxypyruvate reductase activity that catalyzes the conversion of glycerate from (a) and/or (b) to hydroxypyruvate;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine aminotransferase activity or an enzyme having serine oxidoreductase (deaminating) activity that catalyzes the conversion of L-serine to hydroxypyruvate;

(e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having L-serine decarboxylase activity that catalyzes the conversion of L-serine to ethanolamine;

(f) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxypyruvate decarboxylase activity that catalyzes the conversion of hydroxypyruvate from (c) and/or (d) to glycolaldehyde;

(g) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having ethanolamine aminotransferase or ethanolamine oxidoreductase

(deaminating) activity that catalyzes the conversion of ethanolamine from (e) to glycolaldehyde;

(h) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycerate decarboxylase activity that catalyzes the conversion of glycerate from (a) and/or (b) to MEG;

(i) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde reductase activity that catalyzes the conversion of glycolaldehyde from (f) and/or (g) to MEG;

(j) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde dehydrogenase activity that catalyzes the conversion of glycolaldehyde from (f) and/or (g) to glycolic acid;

wherein the produced intermediate DHAP from embodiments [mH] and/or [mI] and/or pyruvate from embodiment [mJ] is converted to 3-phosphoglycerate and/or 2-phosphoglycerate through endogenous glycolysis or gluconeogenesis in the microorganism, and wherein MEG (or glycolic acid) is produced.

[001225] In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mH] and the C3 pathway for production of acetone comprises embodiment [mM]. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mI] and the C3 pathway for production of acetone comprises embodiment [mM]. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mJ] and the C3 pathway for production of acetone comprises embodiment [mM].

[001226] [mM] In one embodiment, the application relates to a method of producing a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and acetone, from one or more hexose feedstock, wherein the method of embodiment [mA] or embodiment [mB] (and optionally comprising embodiment [mC]), and comprising additionally one or more of embodiments [mD], [mE], [mF] and [mG], further comprising one or more of embodiments [mH], [mI], and [mJ], further comprises introducing into and/or expressing in the recombinant microorganism one or more of the following from (a) to (c):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having thiolase or acetyl coenzyme A acetyltransferase activity that catalyzes the conversion of acetyl-CoA to acetoacetyl-CoA;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having acetyl-CoA:acetoacetate-CoA transferase activity or an enzyme having

acetate:acetoacetyl-CoA hydrolase activity that catalyzes the conversion of acetoacetyl-CoA from (a) to acetoacetate; and/or

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having acetoacetate decarboxylase activity that catalyzes the conversion of acetoacetate from (b) to acetone;

wherein the produced intermediate DHAP from embodiments [mH] and/or [mI] and/or pyruvate from embodiment [mJ] is converted to acetyl-CoA through endogenous glycolysis in the microorganism, and wherein MEG (or glycolic acid) and acetone are co-produced.

[001227] In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mH] and the C3 pathway for production of isobutene comprises embodiment [mN]. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mI] and the C3 pathway for production of isobutene comprises embodiment [mN]. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mJ] and the C3 pathway for production of isobutene comprises embodiment [mN].

[001228] [mN] In some embodiments, the application relates to a method of producing a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and isobutene via acetone or HMG-CoA, from one or more hexose feedstock, wherein the method of embodiment [mA] or embodiment [mB] (and optionally comprising embodiment [mC]), and comprising additionally one or more of embodiments [mD], [mE], [mF] and [mG], further comprising one or more of embodiments [mH], [mI], and [mJ], further comprises introducing into and/or expressing in the recombinant microorganism one or more of the following from (a) to (d):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having thiolase or acetyl coenzyme A acetyltransferase activity that catalyzes the conversion of acetyl-CoA to acetoacetyl-CoA;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having acetyl-CoA:acetoacetate-CoA transferase activity or an enzyme having acetate:acetoacetyl-CoA hydrolase activity that catalyzes the conversion of acetoacetyl-CoA from (a) to acetoacetate;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having acetoacetate decarboxylase activity that catalyzes the conversion of acetoacetate from (b) to acetone;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-hydroxyisovalerate synthase activity that catalyzes the conversion of acetone from (c) and acetyl-CoA to 3-hydroxyisovalerate (3HIV);

or

5 wherein the the method comprises introducing into or expressing in the recombinant microorganism one or more of the following from (e) to (j):

(e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having thiolase or acetyl coenzyme A acetyltransferase activity that catalyzes the conversion of acetyl-CoA to acetoacetyl-CoA;

10 (f) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxymethylglutaryl-CoA synthase activity that catalyzes the conversion of acetoacetyl-CoA from (e) and acetyl-CoA to 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA);

(g) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having methylglutaconyl-CoA hydratase activity that catalyzes the conversion of
15 HMG-CoA from (f) to 3-methylglutaconyl-CoA;

(h) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having methylcrotonyl-CoA carboxylase activity that catalyzes the conversion of 3-methylglutaconyl-CoA from (g) to 3-methylcrotonyl-CoA;

(i) at least one endogenous or exogenous nucleic acid molecule encoding an
20 enzyme having methylcrotonyl-CoA hydratase activity that catalyzes the conversion of 3-methylcrotonyl-CoA from (h) to 3-hydroxyisovaleryl-CoA;

(j) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-hydroxyisovaleryl-CoA thioesterase activity that catalyzes the conversion of 3-hydroxyisovaleryl-CoA from (i) to 3HIV;

25 wherein the method further introducing into or expressing in the recombinant microorganism (a1) and (a2), and/or (b1) selected from:

(a1) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3HIV kinase activity that catalyzes the conversion of 3HIV from (d) or (j) to 3HIV-3-phosphate;

30 (a2) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3HIV-3-phosphate decarboxylase activity that catalyzes the conversion of 3HIV-3-phosphate from (a1) to isobutene;

(b1) at least one endogenous or exogenous nucleic acid molecule encoding a an enzyme having 3HIV decarboxylase activity that catalyzes the conversion of 3HIV from (d) or (j) to isobutene;

wherein the produced intermediate DHAP from embodiments [mH] and/or [mI] and/or pyruvate from embodiment [mJ] is converted to acetyl-CoA through endogenous glycolysis in the microorganism, and wherein MEG (or glycolic acid) and isobutene are co-produced.

[001229] [mO] In one embodiment, the method of embodiment [mM] and/or [mN] (optionally comprising embodiment [EE]), optionally further comprise introducing into and/or expressing in the recombinant microorganism at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having secondary alcohol dehydrogenase activity that catalyzes the conversion of acetone to isopropanol.

[001230] [mP] In another embodiment, the method of embodiment [mO] (optionally comprising embodiment [EE]), optionally further comprises introducing into and/or expressing in the recombinant microorganism at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having dehydratase activity that catalyzes the conversion of isopropanol to propene.

[001231] In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mH] and the C3 pathway for production of serine comprises embodiment [mQ]. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mI] and the C3 pathway for production of serine comprises embodiment [mQ]. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mJ] and the C3 pathway for production of serine comprises embodiment [mQ].

[001232] [mQ] In one embodiment, the application relates to a method of producing a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and serine, from one or more hexose feedstock and a nitrogen source, wherein the method of embodiment [mA] or embodiment [mB] (and optionally comprising embodiment [mC]), and comprising additionally one or more of embodiments [mD], [mE], [mF] and [mG], further comprising one or more of embodiments [mH], [mI], and [mJ], further comprises introducing into and/or expressing in the recombinant microorganism one or more of the following from (a) to (h):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphoglycerate dehydrogenase activity that catalyzes the conversion of 3-phosphoglycerate to 3-phosphohydroxypyruvate;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphoglycerate phosphatase activity and/or an enzyme having glycerate 3-kinase activity that catalyzes the conversion of 3-phosphoglycerate to glycerate;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 2-phosphoglycerate phosphatase activity and/or an enzyme having glycerate 2-kinase activity that catalyzes the conversion of 2-phosphoglycerate to glycerate;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having phosphoserine aminotransferase activity that catalyzes the conversion of 3-phosphohydroxypyruvate from (a) to phospho-L-serine;

(e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphohydroxypyruvate phosphatase activity that catalyzes the conversion of 3-phosphohydroxypyruvate from (a) to hydroxypyruvate;

(f) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having phosphoserine phosphatase activity that catalyzes the conversion of phospho-L-serine from (d) to L-serine;

(g) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxypyruvate reductase activity that catalyzes the conversion of glycerate from (b) and/or (c) to hydroxypyruvate;

(h) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine-pyruvate aminotransferase activity that catalyzes the conversion of hydroxypyruvate from (e) and/or (g) to L-serine;

wherein the produced intermediate DHAP from embodiments [mH] and/or [mI] and/or pyruvate from embodiment [mJ] is converted to 3-phosphoglycerate and/or 2-phosphoglycerate through endogenous glycolysis or gluconeogenesis in the microorganism, and wherein MEG (or glycolic acid) and serine are produced.

[001233] In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mH] and the C3 pathway for production of glycine comprises embodiment [mR]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mH] and the C3 pathway for production of glycine comprises embodiment [mS]. In further embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mH] and the C3 pathway for production of glycine comprises embodiment [mT]. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mI] and the C3 pathway for production of glycine comprises embodiment [mR]. In other embodiments, the C2 pathway for production of MEG

(or glycolic acid) comprises embodiment [mI] and the C3 pathway for production of glycine comprises embodiment [mS]. In further embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mI] and the C3 pathway for production of glycine comprises embodiment [mT]. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mJ] and the C3 pathway for production of glycine comprises embodiment [mR]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mJ] and the C3 pathway for production of glycine comprises embodiment [mS]. In further embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mJ] and the C3 pathway for production of glycine comprises embodiment [mT].

[001234] [mR] In one embodiment, the application relates to a method of producing a recombinant microorganism capable of producing monoethylene glycol (MEG) (or glycolic acid) and glycine, from one or more hexose feedstock and a nitrogen source, wherein the method of embodiment [mA] or embodiment [mB] (and optionally comprising embodiment [mC]), and comprising additionally one or more of embodiments [mD], [mE], [mF] and [mG], further comprising one or more of embodiments [mH], [mI], and [mJ], further comprises introducing into or expressing in the recombinant microorganism one or more of the following from (a) to (e):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine hydroxymethyltransferase activity that catalyzes the conversion of L-serine and tetrahydrofolate (THF) to glycine and 5,10-methylene tetrahydrofolate (M-THF);

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having transferase activity that catalyzes the conversion of M-THF from (a) to formaldehyde;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having formaldehyde dehydrogenase activity that catalyzes the conversion of formaldehyde from (b) to formate and NADH;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having formate dehydrogenase activity that catalyzes the conversion of formate from (c) to CO₂ and NADH;

(e) at least one endogenous or exogenous nucleic acid molecule encoding a protein of the glycine cleavage system that catalyze the conversion of M-THF from (a), CO₂, NH₃ and NADH from (c) or (d) to glycine and THF;

wherein THF is reconstituted from steps (b) through (e), wherein optionally formate from (c) is further oxidized to CO₂ and H₂ by a formate hydrogenaselyase complex, and wherein MEG (or glycolic acid) and glycine are produced.

[001235] [mS] In one embodiment, the application relates to a method of producing a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and glycine, from one or more hexose feedstock and a nitrogen source, wherein the method of embodiment **[mA]** or embodiment **[mB]** (and optionally comprising embodiment **[mC]**), and comprising additionally one or more of embodiments **[mD]**, **[mE]**, **[mF]** and **[mG]**, further comprising one or more of embodiments **[mH]**, **[mI]**, and **[mJ]**, further comprises introducing into or expressing in the recombinant microorganism one or more of the following from (a) to (k):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphoglycerate dehydrogenase activity that catalyzes the conversion of 3-phosphoglycerate to 3-phosphohydroxypyruvate;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having phosphoserine aminotransferase activity that catalyzes the conversion of 3-phosphohydroxypyruvate from (a) to phospho-L-serine;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphohydroxypyruvate phosphatase activity that catalyzes the conversion of 3-phosphohydroxypyruvate from (a) to hydroxypyruvate;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having phosphoserine phosphatase activity that catalyzes the conversion of phospho-L-serine from (b) to L-serine;

(e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having L-serine transaminase or serine oxidase activity that catalyzes the conversion of L-serine from (d) to hydroxypyruvate;

(f) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxypyruvate decarboxylase activity that catalyzes the conversion of hydroxypyruvate from (c) or (e) to glycolaldehyde;

(g) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde dehydrogenase activity that catalyzes the conversion of glycolaldehyde from (f) to glycolic acid;

(h) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolate dehydrogenase activity that catalyzes the conversion of glycolic acid from (g) to glyoxylate;

(i) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having alanine-glyoxylate aminotransferase activity that catalyzes the conversion of glyoxylate from (h) and alanine to glycine and pyruvate;

(j) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having alanine transaminase activity that catalyzes the conversion of pyruvate from (i) and glutamate to alanine and 2-oxoglutarate;

(k) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having NAD(P)H dependent glutamate dehydrogenase activity that catalyzes the conversion of 2-oxoglutarate from (j) and ammonia to glutamate;

wherein the produced intermediate DHAP from embodiments [mH] and/or [mI] and/or pyruvate from embodiment [mJ] is converted to 3-phosphoglycerate and/or 2-phosphoglycerate through endogenous glycolysis or gluconeogenesis in the microorganism, wherein the glyoxylate for step (i) optionally comes from glyoxylate shunt in the microorganism, wherein alanine and glutamate are reconstituted from steps (j) and (k), and wherein MEG (or glycolic acid) and glycine are co-produced.

[001236] [mT] In one embodiment, the application relates to a method of producing a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and glycine, from one or more hexose feedstock and a nitrogen source, wherein the method of embodiment [A] or embodiment [B] (and optionally comprising embodiment [C]), and comprising additionally one or more of embodiments [D], [E], [F] and [G], further comprising one or more of embodiments [H], [I], and [J], further comprises introducing into or expressing in the recombinant microorganism one or more of the following from (a) to (l):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 2-phosphoglycerate phosphatase activity and/or an enzyme having glycerate-2-kinase activity that catalyzes the conversion of 2-phosphoglycerate to glycerate;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphoglycerate phosphatase activity and/or an enzyme having glycerate-3-kinase activity that catalyzes the conversion of 3-phosphoglycerate to glycerate;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxypyruvate reductase activity that catalyzes the conversion of glycerate from (a) and/or (b) to hydroxypyruvate;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine aminotransferase or serine oxidoreductase (deaminating) activity that catalyzes the conversion of L-serine to hydroxypyruvate;

(e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having L-serine decarboxylase activity that catalyzes the conversion of L-serine to ethanolamine;

(f) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxypyruvate decarboxylase activity that catalyzes the conversion of hydroxypyruvate from (c) and/or (d) to glycolaldehyde;

(g) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having ethanolamine aminotransferase activity or an enzyme having ethanolamine oxidoreductase (deaminating) activity that catalyzes the conversion of ethanolamine from (e) to glycolaldehyde;

(h) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolaldehyde dehydrogenase activity that catalyzes the conversion of glycolaldehyde from (f) and/or (g) to glycolic acid;

(i) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having glycolate dehydrogenase activity that catalyzes the conversion of glycolic acid from (g) to glyoxylate;

(j) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having alanine-glyoxylate aminotransferase activity that catalyzes the conversion of glyoxylate from (i) and alanine to glycine and pyruvate;

(k) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having alanine transaminase activity that catalyzes the conversion of pyruvate from (j) and glutamate to alanine and 2-oxoglutarate;

(l) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having NAD(P)H dependent glutamate dehydrogenase activity that catalyzes the conversion of 2-oxoglutarate from (k) and ammonia to glutamate;

wherein the produced intermediate DHAP from embodiments [mH] and/or [mI] and/or pyruvate from embodiment [mJ] is converted to 3-phosphoglycerate and/or 2-phosphoglycerate through endogenous glycolysis or gluconeogenesis in the microorganism, wherein the glyoxylate for step (j) optionally comes from glyoxylate shunt in the microorganism, wherein alanine and glutamate are reconstituted from steps (k) and (l), and wherein MEG (or glycolic acid) and glycine are co-produced.

[001237] In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[mH]** and the C3 pathway for production of MEA comprises embodiment **[mU]**. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[mH]** and the C3 pathway for production of MEA comprises embodiment **[mV]**. In further embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[mH]** and the C3 pathway for production of MEA comprises embodiment **[mW]**. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[mI]** and the C3 pathway for production of MEA comprises embodiment **[mU]**. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[mI]** and the C3 pathway for production of MEA comprises embodiment **[mV]**. In further embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[mI]** and the C3 pathway for production of MEA comprises embodiment **[mW]**. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[mJ]** and the C3 pathway for production of MEA comprises embodiment **[mU]**. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[mJ]** and the C3 pathway for production of MEA comprises embodiment **[mV]**. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[mJ]** and the C3 pathway for production of MEA comprises embodiment **[mW]**.

[001238] **[mU]** In one embodiment, the application relates to a method of producing a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and monoethanolamine (MEA), from one or more hexose feedstock and a nitrogen source, wherein the method of embodiment **[mA]** or embodiment **[mB]** (and optionally comprising embodiment **[mC]**), and comprising additionally one or more of embodiments **[mD]**, **[mE]**, **[mF]** and **[mG]**, further comprising one or more of embodiments **[mH]**, **[mI]**, and **[mJ]**, further comprises introducing into or expressing in the recombinant microorganism one or more of the following from (a) to (i):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphoglycerate dehydrogenase activity that catalyzes the conversion of 3-phosphoglycerate to 3-phosphohydroxypyruvate;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphoglycerate phosphatase activity and/or an enzyme having glycerate 3-kinase activity that catalyzes the conversion of 3-phosphoglycerate to glycerate;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 2-phosphoglycerate phosphatase activity and/or an enzyme having glycerate 2-kinase activity that catalyzes the conversion of 2-phosphoglycerate to glycerate;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having phosphoserine aminotransferase activity that catalyzes the conversion of 3-phosphohydroxypyruvate from (a) to phospho-L-serine;

(e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphohydroxypyruvate phosphatase activity that catalyzes the conversion of 3-phosphohydroxypyruvate from (a) to hydroxypyruvate;

(f) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having phosphoserine phosphatase activity that catalyzes the conversion of phospho-L-serine from (d) to L-serine;

(g) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxypyruvate reductase activity that catalyzes the conversion of glycerate from (b) and/or (c) to hydroxypyruvate;

(h) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine-pyruvate aminotransferase activity that catalyzes the conversion of hydroxypyruvate from (e) and/or (g) to L-serine;

(i) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine decarboxylase activity that catalyzes the conversion of L-serine from (f) and/or (h) to MEA;

wherein the produced intermediate DHAP from embodiments [mH] and/or [mI] and/or pyruvate from embodiment [mJ] is converted to 3-phosphoglycerate and/or 2-phosphoglycerate through endogenous glycolysis or gluconeogenesis in the microorganism, and wherein MEG (or glycolic acid) and MEA are co-produced.

[001239] [mV] In one embodiment, the application relates to a method of producing a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and monoethanolamine (MEA) from one or more hexose feedstock and a nitrogen source, wherein the method of embodiment [mA] or embodiment [mB] (and optionally comprising embodiment [mC]), and comprising additionally one or more of embodiments [mD], [mE], [mF], and [mG], further comprising one or more of embodiments [mH], [mI], and [mJ], further comprises introducing into or expressing in the recombinant microorganism one or more of the following from (a) to (f):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 2-phosphoglycerate phosphatase activity and/or an enzyme having glycerate-2-kinase activity that catalyzes the conversion of 2-phosphoglycerate to glycerate;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 3-phosphoglycerate phosphatase activity and/or an enzyme having glycerate-3-kinase activity that catalyzes the conversion of 3-phosphoglycerate to glycerate;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxypyruvate reductase activity that catalyzes the conversion of glycerate from (a) and/or (b) to hydroxypyruvate;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine-pyruvate aminotransferase or serine oxidoreductase (deaminating) activity that catalyzes the conversion of L-serine to hydroxypyruvate;

(e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having hydroxypyruvate decarboxylase activity that catalyzes the conversion of hydroxypyruvate from (c) and/or (d) to glycolaldehyde;

(f) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having transaminase activity that catalyzes the conversion of glycolaldehyde from (e) to MEA;

wherein the produced intermediate DHAP from embodiments [mH] and/or [mI] and/or pyruvate from embodiment [mJ] is converted to 3-phosphoglycerate and/or 2-phosphoglycerate through endogenous glycolysis or gluconeogenesis in the microorganism, and wherein MEG (or glycolic acid) and MEA are co-produced.

[001240] [mW] In one embodiment, the application relates to a method of producing a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and monoethanolamine (MEA) from one or more hexose feedstock and a nitrogen source, wherein the method of embodiment [mA] or embodiment [mB] (and optionally comprising embodiment [mC]), and comprising additionally one or more of embodiments [mD], [mE], [mF], and [mG], further comprising one or more of embodiments [mH], [mI], and [mJ], further comprises introducing into or expressing in the recombinant microorganism one or more of the following from (a) to (b):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having acetaldehyde dehydrogenase activity that catalyzes the conversion of acetyl-CoA to acetaldehyde;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having ethanolamine ammonia lyase activity that catalyzes the conversion of acetaldehyde and ammonia to MEA;

wherein the produced intermediate DHAP from embodiments [mH] and/or [mI] and/or pyruvate from embodiment [mJ] is converted to acetyl-CoA through endogenous glycolysis in the microorganism, and wherein MEG (or glycolic acid) and MEA are co-produced.

[001241] In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mH] and the C3 pathway for production of EDA comprises embodiment [mX]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mH] and the C3 pathway for production of EDA comprises embodiment [mY]. In further embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mH] and the C3 pathway for production of EDA comprises embodiment [mZ]. In yet further embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mH] and the C3 pathway for production of EDA comprises embodiment [mAA]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mH] and the C3 pathway for production of EDA comprises embodiment [mBB].

[001242] In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mI] and the C3 pathway for production of EDA comprises embodiment [mX]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mI] and the C3 pathway for production of EDA comprises embodiment [mY]. In further embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mI] and the C3 pathway for production of EDA comprises embodiment [mZ]. In yet further embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mI] and the C3 pathway for production of EDA comprises embodiment [mAA]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mI] and the C3 pathway for production of EDA comprises embodiment [mBB].

[001243] In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mJ] and the C3 pathway for production of EDA comprises embodiment [mX]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mJ] and the C3 pathway for production of EDA comprises embodiment [mY]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mJ] and the C3 pathway for production of EDA comprises embodiment [mZ]. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment [mJ] and the C3 pathway for production of EDA comprises embodiment [mZ].

acid) comprises embodiment **[mJ]** and the C3 pathway for production of EDA comprises embodiment **[mAA]**. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[mJ]** and the C3 pathway for production of EDA comprises embodiment **[mBB]**.

5 **[001244] [mX]** In one embodiment, the application relates to a method of producing a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and ethylenediamine (EDA), from one or more hexose feedstock and a nitrogen source, wherein the method of embodiment **[mA]** or embodiment **[mB]** (and optionally comprising embodiment **[mC]**), and comprising additionally one or more of embodiments
10 **[mD]**, **[mE]**, **[mF]** and **[mG]**, further comprising one or more of embodiments **[mH]**, **[mI]**, and **[mJ]**, further comprises introducing into or expressing in the recombinant microorganism one or more of the following from (a) to (c):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine dehydrogenase activity that catalyzes the conversion of L-serine to 2-aminomalonate semialdehyde;
15

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 2-aminomalonate semialdehyde decarboxylase activity that catalyzes the conversion of 2-aminomalonate semialdehyde from (a) to aminoacetaldehyde;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an
20 enzyme having aminoacetaldehyde transaminase activity that catalyzes the conversion of aminoacetaldehyde from (b) to EDA;

wherein 2-aminomalonate semialdehyde may optionally be converted to aminoacetaldehyde by a spontaneous reaction, and wherein MEG (or glycolic acid) and EDA are co-produced.

[001245] [mY] In one embodiment, the application relates to a method of producing a
25 recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and ethylenediamine (EDA), from one or more hexose feedstock and a nitrogen source, wherein the method of embodiment **[mA]** or embodiment **[mB]** (and optionally comprising embodiment **[mC]**), and comprising additionally one or more of embodiments **[mD]**, **[mE]**, **[mF]** and **[mG]**, further comprising one or more of embodiments **[mH]**, **[mI]**,
30 and **[mJ]**, further comprises introducing into or expressing in the recombinant microorganism one or more of the following from (a) to (c):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine dehydrogenase activity that catalyzes the conversion of L-serine to 2-aminomalonate semialdehyde;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 2-aminomalonate semialdehyde transaminase activity that catalyzes the conversion of 2-aminomalonate semialdehyde from (a) to 2,3-diaminopropanoate;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 2,3-diaminopropanoate decarboxylase activity that catalyzes the conversion of 2,3-diaminopropanoate from (b) to EDA;

wherein MEG (or glycolic acid) and EDA are co-produced.

[001246] [mZ] In one embodiment, the application relates to a method of producing a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and ethylenediamine (EDA), from one or more hexose feedstock and a nitrogen source, wherein the method of embodiment **[mA]** or embodiment **[mB]** (and optionally comprising embodiment **[mC]**), and comprising additionally one or more of embodiments **[mD]**, **[mE]**, **[mF]** and **[mG]**, further comprising one or more of embodiments **[mH]**, **[mI]**, and **[mJ]**, further comprises introducing into or expressing in the recombinant microorganism one or more of the following from (a) to (c):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine decarboxylase activity that catalyzes the conversion of L-serine to ethanolamine;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having ethanolamine dehydrogenase activity that catalyzes the conversion of ethanolamine from (a) to aminoacetaldehyde;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having aminoacetaldehyde transaminase activity that catalyzes the conversion of aminoacetaldehyde from (b) to EDA;

wherein MEG (or glycolic acid) and EDA are co-produced.

[001247] [mAA] In one embodiment, the application relates to a method of producing a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and ethylenediamine (EDA), from one or more hexose feedstock and a nitrogen source, wherein the method of embodiment **[mA]** or embodiment **[mB]** (and optionally comprising embodiment **[mC]**), and comprising additionally one or more of embodiments **[mD]**, **[mE]**, **[mF]** and **[mG]**, further comprising one or more of embodiments **[mH]**, **[mI]**, and **[mJ]**, further comprises introducing into or expressing in the recombinant microorganism one or more of the following from (a) to (c):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having serine hydroxymethyltransferase activity that catalyzes the conversion of L-serine to glycine;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having aldehyde oxidase activity that catalyzes the conversion of glycine from (a) to aminoacetaldehyde;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having aminoacetaldehyde transaminase activity that catalyzes the conversion of aminoacetaldehyde from (b) to EDA;

wherein MEG (or glycolic acid) and EDA are co-produced.

[001248] [mBB] In one embodiment, the application relates to a method of producing a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and ethylenediamine (EDA), from one or more hexose feedstock and a nitrogen source, wherein the method of embodiment **[mA]** or embodiment **[mB]** (and optionally comprising embodiment **[mC]**), and comprising additionally one or more of embodiments **[mD]**, **[mE]**, **[mF]** and **[mG]**, further comprising one or more of embodiments **[mH]**, **[mI]**, and **[mJ]**, further comprises introducing into or expressing in the recombinant microorganism one or more of the following from (a) to (e):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having an amino acid N-acetyl transferase activity or O-acetyl transferase activity that catalyzes the conversion of L-serine to N-acetylserine;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having N-acetylserine dehydrogenase activity that catalyzes the conversion of N-acetylserine from (a) to N-acetylmalonate semialdehyde;

(c) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having transaminase activity that catalyzes the conversion of N-acetylmalonate semialdehyde from (b) to acetylaminopropanoate;

(d) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having deacetylase activity that catalyzes the conversion of acetylaminopropanoate from (c) to 2,3-diaminopropanoate;

(e) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having 2,3-diaminopropanoate decarboxylase activity that catalyzes the conversion of 2,3-diaminopropanoate from (d) to EDA;

wherein MEG (or glycolic acid) and EDA are co-produced.

[001249] In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[mH]** and the C3 pathway for production of EDA comprises embodiment **[mCC]**. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[mH]** and the C3 pathway for production of EDA comprises embodiment **[mDD]**. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[mI]** and the C3 pathway for production of EDA comprises embodiment **[mCC]**. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[mI]** and the C3 pathway for production of EDA comprises embodiment **[mDD]**. In some embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[mJ]** and the C3 pathway for production of EDA comprises embodiment **[mCC]**. In other embodiments, the C2 pathway for production of MEG (or glycolic acid) comprises embodiment **[mJ]** and the C3 pathway for production of EDA comprises embodiment **[mDD]**.

[001250] **[mCC]** In one embodiment, the application relates to a method of producing a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and ethylenediamine (EDA), from one or more hexose feedstock and a nitrogen source, wherein the method of embodiment **[mA]** or embodiment **[mB]** (and optionally comprising embodiment **[mC]**), and comprising additionally one or more of embodiments **[mD]**, **[mE]**, **[mF]** and **[mG]**, further comprising one or more of embodiments **[mH]**, **[mI]**, and **[mJ]**, further comprises introducing into or expressing in the recombinant microorganism one or more of the following from (a) to (b):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having a serine aminase activity that catalyzes the conversion of L-serine to (S)-2,3-diaminopropanoate;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having an (S)-2,3-diaminopropanoate decarboxylase activity that catalyzes the conversion of (S)-2,3-diaminopropanoate from (a) to EDA;

wherein MEG (or glycolic acid) and EDA are co-produced.

[001251] **[mDD]** In one embodiment, the application relates to a method of producing a recombinant microorganism capable of co-producing monoethylene glycol (MEG) (or glycolic acid) and ethylenediamine (EDA), from one or more hexose feedstock and a nitrogen source, wherein the method of embodiment **[mA]** or embodiment **[mB]** (and optionally comprising embodiment **[mC]**), and comprising additionally one or more of embodiments **[mD]**, **[mE]**, **[mF]** and **[mG]**, further comprising one or more of embodiments **[mH]**, **[mI]**,

and **[mJ]**, further comprises introducing into or expressing in the recombinant microorganism one or more of the following from (a) to (b):

(a) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having an (S)-2,3-diaminopropanoate ammonia lyase activity that catalyzes the conversion of pyruvate and ammonium to (S)-2,3-diaminopropanoate;

(b) at least one endogenous or exogenous nucleic acid molecule encoding an enzyme having (S)-2,3-diaminopropanoate decarboxylase activity that catalyzes the conversion of (S)-2,3-diaminopropanoate from (a) to EDA;

[001252] wherein pyruvate is produced from endogenous glycolysis, and wherein MEG (or glycolic acid) and EDA are co-produced.

[001253] [mEE] In another embodiment, the method selected from embodiment **[mH]**, embodiment **[mI]**, and embodiment **[mJ]**, optionally further comprises introducing into the recombinant microorganism one or more modifications selected from the group consisting of:

(i) a deletion, insertion, or loss of function mutation in a gene encoding a xylose isomerase that catalyzes the conversion of D-xylulose to D-xylose;

(ii) a deletion, insertion, or loss of function mutation in a gene encoding a glycolaldehyde dehydrogenase that catalyzes the conversion of glycolaldehyde to glycolic acid; and

(iii) a deletion, insertion, or loss of function mutation in a gene encoding a lactate dehydrogenase that catalyzes the conversion of pyruvate to lactate.

Enzyme Engineering

[001254] The enzymes in the recombinant microorganism can be engineered to improve one or more aspects of the substrate to product conversion. Non-limiting examples of enzymes that can be further engineered for use in methods of the disclosure include a transketolase, a transaldolase, a pentose-5-phosphatase, an arabitol phosphate dehydrogenase, a phosphopentomutase, an aldehyde reductase, an acetoacetyl coenzyme A hydrolase, a xylose isomerase, a 3-phosphoglycerate dehydrogenase, a phosphoserine aminotransferase, a 3-phosphohydroxypyruvate phosphatase, a phosphoserine phosphatase, a serine transaminase, a hydroxypyruvate decarboxylase, a 3-phosphohydroxypyruvate reductase, a glycolaldehyde dehydrogenase, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase, a serine decarboxylase, an ethanolamine aminotransferase or ethanolamine oxidoreductase (deaminating), a

glycerate decarboxylase, a hydroxypyruvate reductase, a 3-phosphoglycerate phosphatase, a 2-phosphoglycerate phosphatase, a glycerate 3-kinase, a glycerate 2-kinase, a mevalonate diphosphate decarboxylase, and combinations thereof. These enzymes can be engineered for improved catalytic activity, improved selectivity, improved stability, improved tolerance to various fermentation conditions (temperature, pH, etc.), or improved tolerance to various metabolic substrates, products, by-products, intermediates, etc. The term “improved catalytic activity” as used herein with respect to a particular enzymatic activity refers to a higher level of enzymatic activity than that measured relative to a comparable non-engineered enzyme.

[001255] Directed evolution is a term used to describe the entire range of molecular biology techniques that allow natural evolutionary processes to be mimicked in the laboratory. For enzymes, this generally involves the random mutagenesis of one or more starting genes, followed by a screening or selection step to isolate or enrich for enzyme variants with improvements in one or more desirable properties. The process can be iterated until the desired level of change is reached, or until no further change is elicited. A wide range of tools and techniques have been developed over more than two decades to shorten the process from the millions of years taken by nature, to just weeks or months in the laboratory. The most common strategies mimic the mechanisms of evolution that occur in nature, such as error-prone PCR (epPCR), which introduces random point mutations in a population of DNA products, and DNA shuffling techniques, which allow random recombination typically between parent genes with >70% homology. Later techniques accessed a wider range of amino acids through saturation or cassette mutagenesis targeted to pre-chosen sites or at randomly distributed sites, and enabled the random recombination of non-homologous genes. Further techniques can create random insertions and deletions of codons, shuffle domains or exons, or loop regions, and produce a library of random truncations.

[001256] For example, engineering methods have been used to alter the stability, substrate specificity and stereospecificity of aldolases to produce excellent enzymes for biocatalytic processes. The thermostability and solvent tolerance of fructose-1,6-bisphosphate aldolase (FBP-aldolase) was increased using family DNA shuffling of the *fda* genes from *Escherichia coli* and *Edwardsiella ictaluri*. A fourth generation variant was identified which displayed an average 280-fold higher half-life at 53°C than either parent. The same variant also displayed enhanced activity in various polar and non-polar organic solvents (Hao and Berry 2004 Protein Eng Des Sel 17:689-697).

[001257] As another example, acetoacetyl coenzyme A hydrolase can convert acetoacetyl-CoA to acetoacetate. However, the hydrolase is unspecific in that it also reacts with the

same magnitude of order with acetyl-CoA, which is the substrate required for acetoacetyl-CoA formation by the enzyme thiolase. Thus, to create more efficient acetoacetyl-CoA hydrolases, these enzymes have been engineered to have at least 10x higher activity for the acetoacetyl-CoA substrate than for acetyl-CoA substrate by replacing several glutamic acid residues in the enzyme beta subunit that is important for catalysis (WO 2015/042588).

[001258] As another example, the *E. coli* YqhD enzyme is a broad substrate aldehyde reductase with NADPH-dependent reductase activity for more than 10 aldehyde substrates and is a useful enzyme to produce biorenewable fuels and chemicals (Jarboe 2010 *Applied Microbiology and Biotechnology* 89:249). Though YqhD enzyme activity is beneficial through its scavenging of toxic aldehydes, the enzyme is also NADPH-dependent and contributes to NADPH depletion and growth inhibition of organisms. Error-prone PCR of YqhD was performed in order to improve 1,3-propanediol production from 3-hydroxypropionaldehyde (3-HPA). This directed engineering yielded two mutants, D99QN147H and Q202A, with decreased K_m and increased k_{cat} for certain aldehydes, particularly 3-HPA (Li et al. 2008 Prog. Nat. Sci. 18 (12):1519-1524). The improved catalytic activity of the D99QN147H mutant is consistent with what is known about the structure of YqhD (Sulzenbacher et al. 2004 J. Mol. Biol. 342 (2):489-502), as residues Asp99 and Asn147 both interact with NADPH. Use of the D99QN147H mutant increased 1,3-propanediol production from 3-HPA 2-fold. Mutant YqhD enzymes with increased catalytic efficiency (increased K_{cat}/K_m) toward NADPH have also been described in WO 2011012697 A2, which is herein incorporated in its entirety.

[001259] As another example, xylose isomerase is a metal-dependent enzyme that catalyzes the interconversion of aldose and ketose sugars, primarily between xylose to xylulose and glucose to fructose. It has lower affinity for lyxose, arabinose and mannose sugars. The hydroxyl groups of sugars may define the substrate preference of sugar isomerases. The aspartate at residue 256 of *Thermus thermophilus* xylose isomerase was replaced with arginine (Patel et al. 2012 Protein Engineering, Design & Selection vol. 25 no. 7 pp. 331–336). This mutant xylose isomerase exhibited an increase in specificity for D-lyxose, L-arabinose and D-mannose. The catalytic efficiency of the D256R xylose isomerase mutant was also higher for these 3 substrates compared to the wild type enzyme. It was hypothesized that the arginine at residue 256 in the mutant enzyme may play a role in the catalytic reaction or influence changes in substrate orientation.

[001260] As another example, the enzyme xylitol dehydrogenase plays a role in the utilization of xylose along with xylose reductase. Xylose reductase (XR) reduces xylose to xylitol and then xylitol dehydrogenase (XDH) reoxidizes xylitol to form xylulose. However, since XR prefers NADPH as cosubstrate, while XDH exclusively uses NAD^+ as cosubstrate,

a cosubstrate recycling problem is encountered. One solution is to engineer XDH such that its cosubstrate specificity is altered from NAD⁺ to NADP⁺ (Ehrensberger et al. 2006 Structure 14: 567-575). A crystal structure of the *Gluconobacter oxydans* holoenzyme revealed that Asp38 is largely responsible for the NAD⁺ specificity of XDH. Asp38 interacts with the hydroxyls of the adenosine ribose, and Met39 stacks under the purine ring and is also located near the 2' hydroxyl. A double mutant (D38S/M39R) XDH was constructed that exclusively used NADP⁺ without loss of enzyme activity.

[001261] As another example, the enzyme mevalonate diphosphate decarboxylase (MVD) is an ATP-dependent enzyme which catalyzes the phosphorylation/decarboxylation of (R)-mevalonate-5-diphosphate to isopentenyl pyrophosphate (IPP) in the mevalonate (MVA) pathway. In the classical MVA pathway, MVD catalyzes the final step, where it produces IPP from (R)-mevalonate-5-diphosphate (MVAPP) in an irreversible reaction dependent upon ATP. MVAPP is phosphorylated first, and consequent decarboxylation occurs with the concomitant release of inorganic phosphate. With the same mechanism, classical MVDs also catalyze the conversion of the nonphosphorylated 3-hydroxyisovalerate (3-HIV) to isobutene. Mevalonate diphosphate (MDP) decarboxylase variants having improved activity in converting 3-phosphonoxisovalerate into isobutene are disclosed in, for example, WO 2012052427 and WO 2015004211, each of which is herein incorporated in its entirety.

Metabolic Engineering – Enzyme Overexpression or enzyme downregulation/deletion for Increased Pathway Flux

[001262] In various embodiments described herein, the exogenous and endogenous enzymes in the recombinant microorganism participating in the biosynthesis pathways described herein may be overexpressed.

[001263] The terms “overexpressed” or “overexpression” refers to an elevated level (e.g., aberrant level) of mRNAs encoding for a protein(s), and/or to elevated levels of protein(s) in cells as compared to similar corresponding unmodified cells expressing basal levels of mRNAs or having basal levels of proteins. In particular embodiments, mRNA(s) or protein(s) may be overexpressed by at least 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 8-fold, 10-fold, 12-fold, 15-fold or more in microorganisms engineered to exhibit increased gene mRNA, protein, and/or activity.

[001264] In some embodiments, a recombinant microorganism of the disclosure is generated from a host that contains the enzymatic capability to synthesize substrates such as D-xylulose, D-ribulose, D-ribulose-1-phosphate, D-xylulose-1-phosphate, D-ribulose-5-phosphate, D-xylulose-5-phosphate, D-xylonolactone, D-xylonate, 2-keto-3-deoxy-xylonate,

glycolaldehyde, DHAP, pyruvate, acetoacetyl-CoA, acetoacetate or 3-hydroxyisovalerate. In some embodiments, it can be useful to increase the synthesis or accumulation of, for example, D-xylulose, D-ribulose, D-ribulose-1-phosphate, D-xylulose-1-phosphate, D-ribulose-5-phosphate, D-xylulose-5-phosphate, D-ribose-5-phosphate, D-xylonolactone, D-xylonate, 2-keto-3-deoxy-xylonate, glycolaldehyde, DHAP, pyruvate, acetoacetyl-CoA, acetoacetate or 3-hydroxyisovalerate, to increase the production of MEG (or GA), or optionally, MEG (or GA) and one or more co-product.

[001265] In some embodiments, it may be useful to increase the expression of endogenous or exogenous enzymes involved in the MEG (or GA), or optionally, MEG (or GA) and one or more co-product, biosynthesis pathways to increase flux from, for example, D-xylulose, D-ribulose, D-ribulose-1-phosphate, D-xylulose-1-phosphate, D-ribulose-5-phosphate, D-xylulose-5-phosphate, D-ribose-5-phosphate, D-xylonolactone, D-xylonate, 2-keto-3-deoxy-xylonate, glycolaldehyde, DHAP, pyruvate, acetoacetyl-CoA, acetoacetate or 3-hydroxyisovalerate, thereby resulting in increased synthesis or accumulation of MEG (or GA), or optionally, MEG (or GA) and one or more co-product.

[001266] Increased synthesis or accumulation can be accomplished by, for example, overexpression of nucleic acids encoding one or more of the above-described MEG (or GA), or optionally, MEG (or GA) and one or more co-product, biosynthesis pathway enzymes. Overexpression of a MEG (or GA), or optionally, MEG (or GA) and one or more co-product biosynthesis pathway enzyme or enzymes can occur, for example, through increased expression of an endogenous gene or genes, or through the expression, or increased expression, of an exogenous gene or genes. Therefore, naturally occurring organisms can be readily modified to generate non-natural, MEG (or GA), or optionally, MEG (or GA) and one or more co-product, producing microorganisms through overexpression of one or more nucleic acid molecules encoding a MEG (or GA), or optionally, MEG (or GA) and one or more co-product, biosynthesis pathway enzyme. In addition, a non-naturally occurring organism can be generated by mutagenesis of an endogenous gene that results in an increase in activity of an enzyme in the MEG (or GA), or optionally, MEG (or GA) and one or more co-product, biosynthesis pathways.

[001267] Equipped with the present disclosure, the skilled artisan will be able to readily construct the recombinant microorganisms described herein, as the recombinant microorganisms of the disclosure can be constructed using methods well known in the art as exemplified above to exogenously express at least one nucleic acid encoding a MEG (or GA), or optionally, MEG (or GA) and one or more co-product, biosynthesis pathway enzyme in sufficient amounts to produce MEG (or GA), or optionally, MEG (or GA) and one or more co-product.

[001268] Methods for constructing and testing the expression levels of a non-naturally occurring MEG (or GA)-producing, or optionally, MEG (or GA) and one or more co-product producing, host can be performed, for example, by recombinant and detection methods well known in the art. Such methods can be found described in, for example, Sambrook et al.,
5 *Molecular Cloning: A Laboratory Manual*, Third Ed., Cold Spring Harbor Laboratory, New York (2001); Ausubo et al., *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, Md. (1999).

[001269] A variety of mechanisms known in the art can be used to express, or overexpress, exogenous or endogenous genes. For example, an expression vector or
10 vectors can be constructed to harbor one or more MEG (or GA), or optionally, MEG (or GA) and one or more co-product, biosynthesis pathway enzymes encoding nucleic acids as exemplified herein operably linked to expression control sequences functional in the host organism. Expression vectors applicable for use in the microbial host organisms of the invention include, for example, plasmids, phage vectors, viral vectors, episomes and artificial
15 chromosomes, including vectors and selection sequences or markers operable for stable integration into a host chromosome. Selectable marker genes also can be included that, for example, provide resistance to antibiotics or toxins, complement auxotrophic deficiencies, or supply critical nutrients not in the culture media. Expression control sequences can include constitutive and inducible promoters, transcription enhancers, transcription terminators, and
20 the like which are well known in the art. When two or more exogenous encoding nucleic acids are to be co-expressed, both nucleic acids can be inserted, for example, into a single expression vector or in separate expression vectors. For single vector expression, the encoding nucleic acids can be operationally linked to one common expression control sequence or linked to different expression control sequences, such as one inducible
25 promoter and one constitutive promoter. The transformation of exogenous nucleic acid sequences involved in a metabolic or synthetic pathway can be confirmed using methods well known in the art.

[001270] As will be understood by those of skill in the art, it can be advantageous to modify a coding sequence to enhance its expression in a particular host. The genetic code is
30 redundant with 64 possible codons, but most organisms typically use a subset of these codons. The codons that are utilized most often in a species are called optimal codons, and those not utilized very often are classified as rare or low-usage codons. Codons can be substituted to reflect the preferred codon usage of the host, a process sometimes called "codon optimization" or "controlling for species codon bias."

[001271] Optimized coding sequences containing codons preferred by a particular prokaryotic or eukaryotic host (see also, Murray et al. (1989) Nucl. Acids Res. 17:477-508) can be prepared, for example, to increase the rate of translation or to produce recombinant RNA transcripts having desirable properties, such as a longer half-life, as compared with transcripts produced from a non-optimized sequence. Translation stop codons can also be modified to reflect host preference. For example, typical stop codons for *S. cerevisiae* and mammals are UAA and UGA, respectively. The typical stop codon for monocotyledonous plants is UGA, whereas insects and *E. coli* commonly use UAA as the stop codon (Dalphin et al. (1996) Nucl. Acids Res. 24: 216-218).

[001272] Those of skill in the art will recognize that, due to the degenerate nature of the genetic code, a variety of nucleic acid sequences can be used to encode a given enzyme of the disclosure. The nucleic acid sequences encoding the biosynthetic enzymes are referenced herein merely to illustrate an embodiment of the disclosure, and the disclosure includes any nucleic acid sequences that encode the amino acid sequences of the polypeptides and proteins of the enzymes of the present disclosure. In similar fashion, a polypeptide can typically tolerate one or more amino acid substitutions, deletions, and insertions in its amino acid sequence without loss or significant loss of a desired activity. The disclosure includes such polypeptides with different amino acid sequences than the specific proteins described herein so long as the modified or variant polypeptides have the enzymatic anabolic or catabolic activity of the reference polypeptide. Furthermore, the amino acid sequences encoded by the nucleic acid sequences shown herein merely illustrate embodiments of the disclosure.

[001273] Expression control sequences are known in the art and include, for example, promoters, enhancers, polyadenylation signals, transcription terminators, internal ribosome entry sites (IRES), and the like, that provide for the expression of the polynucleotide sequence in a host cell. Expression control sequences interact specifically with cellular proteins involved in transcription (Maniatis et al., *Science*, 236: 1237-1245 (1987)). Exemplary expression control sequences are described in, for example, Goeddel, *Gene Expression Technology: Methods in Enzymology*, Vol. 185, Academic Press, San Diego, Calif. (1990).

[001274] In various embodiments, an expression control sequence may be operably linked to a polynucleotide sequence. By “operably linked” is meant that a polynucleotide sequence and an expression control sequence(s) are connected in such a way as to permit gene expression when the appropriate molecules (e.g., transcriptional activator proteins) are bound to the expression control sequence(s). Operably linked promoters are located

upstream of the selected polynucleotide sequence in terms of the direction of transcription and translation. Operably linked enhancers can be located upstream, within, or downstream of the selected polynucleotide.

[001275] In some embodiments, the recombinant microorganism is manipulated to delete, disrupt, mutate, and/or reduce the activity of one or more endogenous enzymes that catalyzes a reaction in a pathway that competes with the biosynthesis pathway for the production of MEG (or GA), or optionally, MEG (or GA) and one or more co-product.

[001276] In some embodiments, the recombinant microorganism is manipulated to delete, disrupt, mutate, and/or reduce the activity of one or more endogenous enzymes in the oxidative branch of the pentose phosphate pathway. In some embodiments, the manipulation prevents the conversion of glucose-6-phosphate through the oxidative branch of the pentose phosphate pathway and instead shunts glucose-6-phosphate through the non-oxidative branch of the pentose phosphate pathway to produce one or more pentose-5-phosphate intermediate needed for the production of MEG (or GA), or optionally, MEG (or GA) and one or more co-product. In some such embodiments, the one or more endogenous enzyme is selected from glucose 6-phosphate-1-dehydrogenase, 6-phosphogluconolactonase, and 6-phosphogluconate dehydrogenase. In further embodiments, the glucose 6-phosphate-1-dehydrogenase is zwf, the 6-phosphogluconolactonase is pgl, and the 6-phosphogluconate dehydrogenase is gnd.

[001277] In some embodiments, the recombinant microorganism is manipulated to delete, disrupt, mutate, and/or reduce the activity of glyceraldehyde 3-phosphate dehydrogenase and/or phosphoglycerate kinase and/or phosphoglycerate mutase. In some embodiments, the manipulation prevents the conversion of glyceraldehyde 3-phosphate to 1,3-bisphospho-D-glycerate and subsequent intermediates and instead allow glyceraldehyde 3-phosphate to be converted to xylulose-5-phosphate (with a concurrent conversion of fructose-6-phosphate to erythrose-4-phosphate) by a transketolase, and thus produce a pentose-5-phosphate intermediate needed for the production of MEG, or optionally, MEG and one or more co-product, and provide more erythrose-4-phosphate for the non-oxidative branch of the pentose phosphate pathway to further produce one or more pentose-5-phosphate intermediate. In some embodiments, the glyceraldehyde 3-phosphate dehydrogenase is gapA, the phosphoglycerate kinase is pgk and the phosphoglycerate mutase is gpmA and/or gpmM.

[001278] In some embodiments, the recombinant microorganism is manipulated to delete, disrupt, mutate, and/or reduce the activity of 6-phosphofructokinase. In some embodiments, the manipulation prevents the conversion of fructose-6-phosphate to 1,6-bisphosphate and

instead allow fructose-6-phosphate to be converted to erythrose-4-phosphate and acetyl-phosphate by a fructose-6-phosphate phosphoketolase, and provide more erythrose-4-phosphate for the non-oxidative branch of the pentose phosphate pathway to further produce one or more pentose-5-phosphate intermediate needed for the production of MEG (or GA), or optionally, MEG (or GA) and one or more co-product. In some embodiments, the 6-phosphofructokinase is pfkA and/or pfkB.

[001279] In some embodiments, the recombinant microorganism is manipulated to delete, disrupt, mutate, and/or reduce the activity of one or more endogenous enzymes that catalyzes the conversion of glycolaldehyde to glycolic acid. In some such embodiments, the enzyme that catalyzes the conversion of glycolaldehyde to glycolic acid is a glycolaldehyde dehydrogenase. In some embodiments, the glycolaldehyde dehydrogenase is from *Escherichia coli*. In some embodiments, the glycolaldehyde dehydrogenase is encoded by the aldA gene or homologs thereof. In some embodiments, the manipulation prevents the production of glycolic acid from glycolaldehyde and instead shunts the reaction toward conversion of glycolaldehyde to MEG. In some embodiments, the deletion, disruption, mutation, and/or reduction in the activity of one or more endogenous enzymes that catalyzes the conversion of glycolaldehyde to glycolic acid is partial, wherein some glycolaldehyde dehydrogenase function is still present and an amount of glycolic acid is still produced.

[001280] In some embodiments, the recombinant microorganism is manipulated to delete, disrupt, mutate, and/or reduce the activity of one or more endogenous enzymes that catalyzes the conversion of pyruvate to lactate. In some such embodiments, the enzyme that catalyzes the conversion of pyruvate to lactate is a lactate dehydrogenase. In some embodiments, the lactate dehydrogenase is from *Escherichia coli*. In some embodiments, the lactate dehydrogenase is encoded by the ldhA gene or homologs thereof. In some embodiments, the manipulation prevents the production of lactate from pyruvate and instead shunts the reaction toward production of isobutene.

[001281] In some embodiments, the recombinant microorganism is manipulated to delete, disrupt, mutate, and/or reduce the activity of one or more endogenous enzymes that catalyzes the conversion of D-xylulose to D-xylose. In some such embodiments, the enzyme that catalyzes the conversion of D-xylulose to D-xylose is a D-xylose isomerase. In some embodiments, the D-xylose isomerase is from *E. coli*. In some embodiments, the D-xylose isomerase is encoded by the xylA gene or homologs thereof. In some embodiments, the manipulation prevents conversion of D-xylulose to D-xylose and instead shunts the reaction toward the conversion of D-xylulose to D-xylulose-1-phosphate, D-xylulose-5-phosphate, or D-ribulose.

[001282] EXAMPLES**Example 1****[001283] In vitro enzymatic assay for PGM3 protein**

[001284] The enzymatic assay of a phosphoglucomutase, which is a phosphopentomutase, encoded by ScPGM3 (SEQ ID NO: 255 or SEQ ID NO: 258)(PGM3 from *Saccharomyces cerevisiae*) was measured on its natural substrate, ribose-1-phosphate and glucose-1-phosphate in the presence of glucose-1,6-biphosphate, by monitoring the reduction of NADP⁺ into NADPH at OD 340nm for 10 minutes (FIG. 15). Both glucose molecules were purchased from Merck. The reaction mixture contained HEPES 60mM, pH 7.5, KCl 60mM, MgCl₂ 3mM, purified PGM3 protein, NADP⁺ 0.5mM, glucose-1,6-bisphosphate 5μM, glucose-1-phosphate 8mM, and commercially available Glucose-6-Phosphate dehydrogenase at 3U/mL (from Merck).

[001285] The activity on pentose-5P substrates (ribulose-5P and xylulose-5) was determined according to the scheme in FIG. 16. The reaction mixture contained HEPES 60mM, pH 7.5, KCl 60mM, MgCl₂ 3mM, NADH 0.25mM, glucose-1,6-biphosphate 5μM, pentose (ribulose or xylulose) 5mM, ATP 4mM, purified PGM3 protein, purified ribokinase or ribulokinase, and aldolB obtained as previously described (Cam Y. *et al.* (2015) Engineering of a Synthetic Metabolic Pathway for the Assimilation of (d)-Xylose into Value-Added Chemicals. ACS Synth Biol), glycerol-3-phosphate dehydrogenase (GPDH) at 2U/mL (Merck). Results showed xylulose-5P PGM3 specific activity was about 0.015 μmol min⁻¹ mg⁻¹, and ribulose-5P PGM3 specificity was about 0.016 ± 0.002 μmol min⁻¹ mg⁻¹.

Example 2**[001286] In vivo enzymatic assay for PGM3 protein**

[001287] To assess the feasibility of using PGM3 for glycolic acid (GA) production *in vivo*, an assay was developed based on the use of a screening strain with the genotype MG1655 **ΔtktA-ΔtktB**. As described in FIG. 17, such a strain cannot grow on xylose by itself, because of the deletion of transketolase genes tktA (GenBank Gene ID: 947420) and tktB (GenBank Gene ID: 945865). However, a strain expressing an active PGM3 protein could use pentoses phosphorylated on carbon 5 to produce pentose phosphorylated on carbon 1, and subsequently convert pentose-1P to glycolaldehyde and DHAP by aldolase activity. Glycolaldehyde and glyceraldehyde-3-phosphate can enter the glyoxylate shunt and glycolysis and support strain growth. Consequently, *in vivo* PGM3 activity in the screening strain is directly correlated with cell growth.

[001288] Screening strain PGM3 (MG1655 Δ tktA- Δ tktB) was transformed by electroporation with plasmid carrying both *S. cerevisiae* PGM3 and proper aldolase (aldoB or fucA) using standard procedure (Woodall C.A. *E. coli* Plasmid Vectors. Methods in Molecular Biology™. 2003. vol 235). Resulting strains were grown in M9 xylose medium (20 g/L xylose) for 50 hours. Kanamycin was added with a final concentration of 100 µg/mL. Growth was monitored by OD600 (FIG. 17).

[001289] Because pentose intermediates phosphorylated on carbon 5, such as ribulose-5P and xylulose-5P, can be naturally obtained from glucose via PPP pathways, active PGM3 from *S. cerevisiae* is also suitable to demonstrate in-vivo production of glycolic acid from glucose, in a similar approach. No *E. coli* MG1655 Δ tktA- Δ tktB cell growth is observed without PGM3.Sc overexpression plasmid. On the other hand, *E. coli* MG1655 Δ tktA- Δ tktB restored its ability to growth on xylose when the active PGM3 from *S. cerevisiae* was able to convert D-xylulose-5P and D-ribulose-5P into D-xylulose-1P and D-ribulose-1P, respectively, and so the pentose-1P intermediates converted into glycoaldehyde and G3P by aldolase activity (aldoB on D-xylulose-1P and fucA on D-ribulose-1P).

Example 3

[001290] Enzymatic assay of D-xylulose-1-phosphate aldolase (aldoB) (Maylay AD, *et al.* (2002) Arch Biochem Biophys, 408:295-304) and D-ribulose-1-phosphate aldolase (FucA) (Elsinghorst EA, *et al.* (1994) J Bacteriol, 176:7223-7232) activities were tested according to the scheme in FIG. 18 on their natural substrates, xylulose-1P and ribulose-1P, respectively. The reaction mixture contained HEPES 60mM, pH 7.5, KCl 60mM, MgCl₂ 3mM, NADH 0.25mM, ATP 4mM, purified kinase, and aldolase, GPDH at 2U/mL. Furthermore, the kinases used to produce pentose-1P were fucK (SEQ ID NO: 257)(LeBlanc DJ, *et al.* (1971) Metabolism of D-arabinose: origin of a D-ribulokinase activity in Escherichia coli. J Bacteriol, 106:82-89) for ribulose-1P, and khk-C (SEQ ID NO: 256) (Asipu A, *et al.* (2003) Properties of normal and mutant recombinant human ketohexokinases and implications for the pathogenesis of essential fructosuria. Diabetes, 52:2426-2432) for xylulose-1P. Both proteins were tested as previously described for kinases producing pentose-5P.

HPLC analysis.

[001291] Pentose, glycolate and glycerol quantifications were determined by high-performance liquid chromatography using Thermo Fisher Scientific system (Courtaboeuf, France) equipped with a RI detector, with a UV detector at 205 nm and a Phenomenex column (Rezex H+) at 50°C using 2.5 mM H₂SO₄ as mobile phase at 0.5 mL/min.

[001292] NUMBERED EMBODIMENTS OF THE DISCLOSURE

1. A recombinant microorganism comprising one or more biochemical pathways that produces monoethylene glycol (MEG) or glycolic acid (GA) from one or more hexose feedstock via one or more pentose-5-phosphate intermediate.

2. The recombinant microorganism of any one of the preceding embodiments, wherein one or more co-product is co-produced with MEG or GA.

3. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more pentose-5-phosphate intermediate is one or more of D-ribose-5-phosphate, D-xylulose-5-phosphate or D-ribulose-5-phosphate.

4. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more biochemical pathway comprises at least one enzyme having transketolase activity.

5. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having transketolase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to tktA or tktB from *E. coli*.

6. The recombinant microorganism of any one of the preceding embodiments, wherein the at least enzyme having transketolase activity comprises tktA or tktB from *E. coli*.

7. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more biochemical pathway comprises at least one enzyme having transaldolase activity.

8. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having transaldolase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to talA or talB from *E. coli*.

9. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having transaldolase activity is selected from *E. coli* talA and *E. coli* talB.

10. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more biochemical pathway comprises at least one enzyme having ribulose-5-phosphate 3-epimerase activity.

11. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having ribulose-5-phosphate 3-epimerase activity is

encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to rpe from *E. coli*.

12. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having ribulose-5-phosphate 3-epimerase activity comprises rpe from *E. coli*.

13. The recombinant microorganism of any one of embodiments 1-12, wherein the one or more biochemical pathway comprises at least one enzyme having ribose-5-phosphate isomerase activity.

14. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having ribose-5-phosphate isomerase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to rpiA or rpiB from *E. coli*.

15. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having ribose-5-phosphate isomerase activity comprises rpiA or rpiB from *E. coli*.

16. The recombinant microorganism of any one of the preceding embodiments, wherein the recombinant microorganism further comprises a deleted or diminished activity in one or more endogenous enzymes selected from glyceraldehyde 3-phosphate dehydrogenase, phosphoglycerate kinase and phosphoglycerate mutase.

17. The recombinant microorganism of any one of the preceding embodiments, wherein the glyceraldehyde 3-phosphate dehydrogenase is gapA, the phosphoglycerate kinase is pgk and the phosphoglycerate mutase is gpmA and/or gpmM.

18. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more biochemical pathway comprises at least one enzyme having fructose-6-phosphate phosphoketolase activity.

19. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having fructose-6-phosphate phosphoketolase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having fructose-6-phosphate phosphoketolase activity selected from the group consisting of *Bifidobacterium dentium* BDP_1006, *Bifidobacterium lactis* xfp, *Lactobacillus paraplantarum* xpkA and *Bifidobacterium breve* xfp.

20. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having fructose-6-phosphate phosphoketolase activity is selected from the group consisting of *Bifidobacterium dentium* BDP_1006, *Bifidobacterium lactis* xfp, *Lactobacillus paraplantarum* xpkA and *Bifidobacterium breve* xfp.

5 21. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more biochemical pathway comprises at least one enzyme having phosphate acetyltransferase activity.

10 22. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having phosphate acetyltransferase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having phosphate acetyltransferase activity selected from *E. coli* pta and *Clostridium acetobutylicum* pta.

15 23. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having phosphate acetyltransferase activity is selected from *E. coli* pta and *Clostridium acetobutylicum* pta.

24. The recombinant microorganism of any one of the preceding embodiments or 18-23, wherein the recombinant microorganism further comprises a deleted or diminished activity in an endogenous 6-phosphofructokinase enzyme.

20 25. The recombinant microorganism of any one of the preceding embodiments, wherein the 6-phosphofructokinase is pfkAB.

26. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more biochemical pathway comprises at least one enzyme having pentose phosphatase activity.

25 27. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having pentose phosphatase activity is selected from one or more of an enzyme having D-pentose-5-phosphatase activity, an enzyme having D-xylulose-5-phosphatase activity, an enzyme having D-ribose-5-phosphatase activity, and an enzyme having D-ribulose-5-phosphatase activity.

30 28. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having D-pentose-5-phosphatase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having D-pentose-5-phosphatase activity selected from the group consisting of *E. coli* phoA, *E. coli* yfbT and *E. coli* yidA.

29. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having D-pentose-5-phosphatase activity is selected from the group consisting of *E. coli* phoA, *E. coli* yfbT and *E. coli* yidA.

30. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having D-xylulose-5-phosphatase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to *Bacillus subtilis* araL.

31. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having D-xylulose-5-phosphatase activity is *Bacillus subtilis* araL.

32. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having D-ribose-5-phosphatase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having D-ribose-5-phosphatase activity selected from the group consisting of *Arabidopsis thaliana* SGPP, *Pseudomonas fluorescens* PFLU_2693, and *E. coli* ybiV.

33. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having D-ribose-5-phosphatase activity is selected from the group consisting of *Arabidopsis thaliana* SGPP, *Pseudomonas fluorescens* PFLU_2693, and *E. coli* ybiV.

34. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having D-ribulose-5-phosphatase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to *Plasmodium falciparum* PF10_0325.

35. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having D-ribulose-5-phosphatase activity is *Plasmodium falciparum* PF10_0325.

36. The recombinant microorganism of any one of any one of the preceding embodiments, wherein the one or more biochemical pathway comprises at least one enzyme having arabitol phosphate dehydrogenase activity.

37. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having arabitol phosphate dehydrogenase activity is selected from one or more of an enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity, an enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity, an enzyme

having D-arabitol 1-phosphate 2-dehydrogenase activity and an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity.

38. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to APDH from *Enterococcus avium*.

39. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity is APDH from *Enterococcus avium*.

40. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to APDH from *Enterococcus avium*.

41. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity is APDH from *Enterococcus avium*.

42. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity selected from the group consisting of *Candida albicans* ARD1, *Candida tropicalis* ARD1, *Scheffersomyces stipitis* ARDH, *Bacillus subtilis* egsA (araM), *Aeropyrum pernix* egsA, *E. coli* gpsA and *Saccharomyces cerevisiae* GPD1.

43. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity is selected from the group consisting of *Candida albicans* ARD1, *Candida tropicalis* ARD1, *Scheffersomyces stipitis* ARDH, *Bacillus subtilis* egsA (araM), *Aeropyrum pernix* egsA, *E. coli* gpsA and *Saccharomyces cerevisiae* GPD1.

44. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity selected from the group consisting of *Pseudomonas fluorescens* mtlD, *Klebsiella pneumoniae* dalD, *Ralstonia solanacearum* dalD, *Bacillus*

subtilis egsA (araM), *Aeropyrum pernix* egsA, *E. coli* gpsA and *Saccharomyces cerevisiae* GPD1.

45. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity is selected from the group consisting of *Pseudomonas fluorescens* mtlD, *Klebsiella pneumoniae* dalD, *Ralstonia solanacearum* dalD, *Bacillus subtilis* egsA (araM), *Aeropyrum pernix* egsA, *E. coli* gpsA and *Saccharomyces cerevisiae* GPD1.

46. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more biochemical pathway comprises at least one enzyme having phosphopentomutase activity.

47. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having phosphopentomutase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having phosphopentomutase activity selected from the group consisting of *E. coli* deoB, *E. coli* pgm, *Bacillus subtilis* pgcA, *Lactococcus lactis* pgmB, *E. coli* ycjU, *Pseudomonas aeruginosa* algC, and *E. coli* cpsG.

48. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having phosphopentomutase activity is selected from the group consisting of *E. coli* deoB, *E. coli* pgm, *Bacillus subtilis* pgcA, *Lactococcus lactis* pgmB, *E. coli* ycjU, *Pseudomonas aeruginosa* algC, and *E. coli* cpsG.

49. The recombinant microorganism of any one of any one of the preceding embodiments, wherein the recombinant microorganism further comprises a deleted or diminished activity in one or more endogenous enzymes selected from glucose 6-phosphate-1-dehydrogenase, 6-phosphogluconolactonase and 6-phosphogluconate dehydrogenase.

50. The recombinant microorganism of any one of the preceding embodiments, wherein the glucose 6-phosphate-1-dehydrogenase is zwf, the 6-phosphogluconolactonase is pgl and the 6-phosphogluconate dehydrogenase is gnd.

51. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more co-product is selected from acetone, isopropanol, propene, isobutene and one or more serine pathway compounds.

52. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more serine pathway compounds is selected from serine, glycine, monoethanolamine (MEA) and ethylenediamine (EDA).

53. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more hexose feedstock is selected from glucose or oligomers of glucose thereof.

54. The recombinant microorganism of any one of the preceding embodiments, wherein the oligomers of glucose are selected from fructose, sucrose, starch, cellobiose, maltose, lactose and cellulose.

55. The recombinant microorganism of any one of the preceding embodiments, wherein the expression of an enzyme having transketolase activity or an enzyme having fructose-6-phosphate phosphoketolase activity enables a lossless conversion of one or more hexose feedstock to the one or more pentose-5-phosphate intermediate.

56. The recombinant microorganism of any one of the preceding embodiments, wherein MEG or GA is produced through the conversion of glycolaldehyde in a C2 pathway and through the conversion of dihydroxyacetone phosphate (DHAP) or pyruvate in a C3 pathway.

57. The recombinant microorganism of any one of the preceding embodiments, wherein GA is produced by the oxidation of glycolaldehyde by a glycolaldehyde dehydrogenase.

58. The recombinant microorganism of any one of the preceding embodiments, wherein one or more enzyme for the production of MEG or GA through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine transaminase activity, a hydroxypyruvate decarboxylase activity, a 3-phosphohydroxypyruvate reductase activity, a glycolaldehyde reductase activity, a glycolaldehyde dehydrogenase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a serine decarboxylase activity, an ethanolamine aminotransferase or ethanolamine oxidoreductase (deaminating) activity, a glycerate decarboxylase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, and a glycerate 2-kinase activity.

59. The recombinant microorganism of any one of the preceding embodiments, wherein MEG or GA is produced through the conversion of glycolaldehyde in a C2 pathway and one or more co-product is produced through the conversion of dihydroxyacetone phosphate (DHAP) or pyruvate in a C3 pathway.

60. The recombinant microorganism of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, and an acetoacetate decarboxylase activity, and wherein the one or more co-product comprises acetone.

61. The recombinant microorganism of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, and a secondary alcohol dehydrogenase activity, and wherein the one or more co-product comprises isopropanol.

62. The recombinant microorganism of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, a secondary alcohol dehydrogenase activity, and a dehydratase activity, and wherein the one or more co-product comprises propene.

63. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, a 3-hydroxyisovalerate (3HIV) synthase activity, a hydroxymethylglutaryl-CoA synthase activity, a methylglutaconyl-CoA hydratase activity, a methylcrotonyl-CoA carboxylase activity, a methylcrotonyl-CoA hydratase activity, a 3-hydroxyisovaleryl-CoA thioesterase activity, a 3HIV kinase activity, a 3HIV-3-phosphate decarboxylase activity, and a 3HIV decarboxylase activity, and wherein the one or more co-product comprises isobutene.

64. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected

from a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, and a glycerate 2-kinase activity, and wherein the one or more co-product comprises L-serine.

65. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a serine hydroxymethyltransferase activity, a transferase activity, a formaldehyde dehydrogenase activity, a formate dehydrogenase activity, an activity associated with glycine cleavage system, a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine transaminase activity, a hydroxypyruvate decarboxylase activity, a serine oxidoreductase (deaminating) activity, a serine decarboxylase activity, an ethanolamine aminotransferase or ethanolamine oxidoreductase (deaminating) activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, a glycerate 2-kinase activity, a glycolaldehyde dehydrogenase activity, a glycolate dehydrogenase activity, an alanine-glyoxylate aminotransferase activity, an alanine transaminase activity, an NAD(P)H dependent glutamate dehydrogenase activity, and wherein the one or more co-product comprises glycine.

66. The method of any one of the preceding embodiments, wherein the activity associated with glycine cleavage system comprise an enzyme or protein selected from a glycine decarboxylase (P protein), an aminomethyltransferase (T protein), a dihydrolipoamide dehydrogenase (L protein), and an H protein.

67. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a transaminase activity, a hydroxypyruvate decarboxylase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a serine decarboxylase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, a glycerate 2-kinase activity, an acetaldehyde dehydrogenase activity, and an

ethanolamine ammonia lyase activity, and wherein the one or more co-product comprises monoethanolamine (MEA).

68. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a serine dehydrogenase activity, a 2-aminomalonate semialdehyde decarboxylase activity, an aminoacetaldehyde transaminase activity, a 2-aminomalonate semialdehyde transaminase activity, a 2,3-diaminopropanoate decarboxylase activity, a serine decarboxylase activity, an ethanolamine dehydrogenase activity, a serine hydroxymethyltransferase activity, an aldehyde oxidase activity, an N-acetyl transferase or O-acetyl transferase activity, an N-acetylserine dehydrogenase activity, a transaminase activity, a deacetylase activity, a serine aminase activity, and a 2,3-diaminopropanoate ammonia lyase activity, and wherein the one or more co-product comprises ethylenediamine (EDA).

69. The recombinant microorganism of any one of the preceding embodiments, wherein one or more enzyme for the production of MEG or GA from glycolaldehyde in a C2 pathway are selected from one or more enzyme having an activity selected from a D-tagatose 3-epimerase activity, a D-ribulokinase activity, a D-ribulose-1-phosphate aldolase activity, a D-xylulose 1-kinase activity, a D-xylulose-1-phosphate aldolase activity, a xylose reductase or aldose reductase activity, a xylitol dehydrogenase activity, a xylose isomerase activity, a xylose dehydrogenase activity, a xylonolactonase activity, a xylonate dehydrogenase activity, a 2-keto-3-deoxy-D-pentionate aldolase activity, a glycolaldehyde reductase activity and a glycolaldehyde dehydrogenase activity.

70. The recombinant microorganism of any one of the preceding embodiments, wherein the recombinant microorganism further comprises one or more modifications to diminish or delete activity in a glycolaldehyde dehydrogenase, a lactate dehydrogenase, a xylose isomerase, a xylulose kinase or combination thereof.

71. The recombinant microorganism of any one of the preceding embodiments, wherein at least a portion of the excess NADH produced in the C3 pathway is used as a source of reducing equivalents in the C2 pathway.

72. The recombinant microorganism of any one of the preceding embodiments, wherein at least a portion of the excess NADH produced in the C3 pathway is used to produce ATP.

73. The recombinant microorganism of any one of the preceding embodiments, wherein excess biomass formation is minimized and production of MEG (or GA) or MEG (or GA) and one or more co-product is maximized.

5 74. A method of producing MEG or glycolic acid (GA) using a recombinant microorganism of any of the preceding claims, wherein the method comprises cultivating the recombinant microorganism in a culture medium containing one or more hexose feedstock providing a carbon source until the MEG or GA is produced.

75. The method of any one of the preceding embodiments, wherein one or more co-product is co-produced with MEG or GA.

10 76. The method of any one of the preceding embodiments, wherein the one or more co-product is selected from acetone, isopropanol, propene, isobutene and one or more serine pathway compound.

15 77. The method of any one of the preceding embodiments, wherein the one or more serine pathway compound is selected from serine, glycine, monoethanolamine (MEA) and ethylenediamine (EDA).

78. A method of producing a recombinant microorganism that produces or accumulates MEG or glycolic acid (GA) from one or more exogenous hexose feedstock via one or more pentose-5-phosphate intermediate, comprising:

20 introducing into the recombinant microorganism one or more enzyme for the conversion of the one or more hexose feedstock to the one or more pentose-5-phosphate intermediate;

introducing into the recombinant microorganism one or more enzyme for the conversion of the one or more pentose-5-phosphate intermediate to one or more pentose or pentose-1-phosphate intermediate;

25 introducing into the recombinant microorganism a C2 pathway comprising one or more enzyme for the production of MEG or GA from glycolaldehyde; and

introducing into the recombinant microorganism a C3 pathway comprising one or more enzyme for the production of MEG or GA derived from DHAP or pyruvate; and

30 culturing the recombinant microorganism in a culture medium containing the one or more hexose feedstock to produce or accumulate MEG or GA,

wherein glycolaldehyde and DHAP (or pyruvate) are intermediates produced in the C2 pathway, and wherein MEG or GA are produced in both the C2 and C3 pathways.

79. A method of producing a recombinant microorganism that produces or accumulates MEG or glycolic acid (GA) and one or more co-product from one or more exogenous hexose feedstock via one or more pentose-5-phosphate intermediate, comprising:

5 introducing into the recombinant microorganism one or more enzyme for the conversion of the one or more hexose feedstock to the one or more pentose-5-phosphate intermediate;

10 introducing into the recombinant microorganism one or more enzyme for the conversion of the one or more pentose-5-phosphate intermediate to one or more pentose or pentose-1-phosphate intermediate;

introducing into the recombinant microorganism a C2 pathway comprising one or more enzyme for the production of MEG or GA from glycolaldehyde; and

15 introducing into the recombinant microorganism a C3 pathway comprising one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate; and

20 culturing the recombinant microorganism in a culture medium containing the one or more hexose feedstock to produce or accumulate MEG or GA and one or more co-product, wherein glycolaldehyde and DHAP (or pyruvate) are intermediates produced in the C2 pathway, and wherein MEG or GA is produced in the C2 pathway and the one or more co-product is produced in the C3 pathway.

80. The method of any one of the preceding embodiments, wherein the glycolaldehyde is oxidized to glycolic acid by a glycolaldehyde dehydrogenase.

25 81. The method of any one of the preceding embodiments, wherein the one or more enzyme for the conversion of the one or more hexose feedstock to the one or more pentose-5-phosphate intermediate are selected from one or more enzyme having a transketolase activity, a transaldolase activity, a ribulose-5-phosphate 3-epimerase activity, and a ribose-5-phosphate isomerase activity.

30 82. The method of any one of the preceding embodiments, wherein the method further comprises introducing into the recombinant microorganism one or more modifications to diminish or delete activity in one or more endogenous enzymes selected from glyceraldehyde 3-phosphate dehydrogenase (gapA), phosphoglycerate kinase (pgk) and phosphoglycerate mutase (gpmA and/or gpmM).

83. The method of any one of the preceding embodiments, wherein the one or more enzyme for the conversion of the one or more hexose feedstock to the one or more

pentose-5-phosphate intermediate are selected from one or more enzyme having a fructose-6-phosphate phosphoketolase activity, a phosphate acetyltransferase activity, a transketolase activity, a transaldolase activity, a ribulose-5-phosphate 3-epimerase activity, and a ribose-5-phosphate isomerase activity.

5 84. The method of any one of the preceding embodiments, wherein the method further comprises introducing into the recombinant microorganism one or more modifications to diminish or delete activity in an endogenous 6-phosphofructokinase (pfkAB) enzyme.

 85. The method of any one of the preceding embodiments, wherein the method further comprises: introducing into the recombinant microorganism one or more
10 modifications to diminish or delete activity in one or more endogenous enzymes selected from glucose 6-phosphate-1-dehydrogenase, 6-phosphogluconolactonase, and 6-phosphogluconate dehydrogenase.

 86. The method of any one of the preceding embodiments, wherein the glucose 6-phosphate-1-dehydrogenase is zwf, the 6-phosphogluconolactonase is pgl, and the 6-
15 phosphogluconate dehydrogenase is gnd.

 87. The method of any one of the preceding embodiments, wherein the one or more enzyme for the conversion of the one or more pentose-5-phosphate intermediate to one or more pentose or pentose-1-phosphate intermediate are selected from one or more enzyme having a pentose phosphatase activity, a D-ribulose-5-phosphatase activity, a D-
20 ribose-5-phosphatase activity, a D-ribose isomerase activity, an arabitol phosphate dehydrogenase activity, and a phosphopentomutase activity.

 88. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of MEG or GA derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a 3-
25 phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine transaminase activity, a hydroxypyruvate decarboxylase activity, a 3-phosphohydroxypyruvate reductase activity, a glycolaldehyde reductase activity, a glycolaldehyde dehydrogenase activity, a serine oxidoreductase (deaminating) or serine-
30 pyruvate aminotransferase activity, a serine decarboxylase activity, an ethanolamine aminotransferase or ethanolamine oxidoreductase (deaminating) activity, a glycerate decarboxylase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, and a glycerate 2-kinase activity.

89. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, and an acetoacetate decarboxylase activity, and wherein the one or more co-product comprises acetone.

90. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, and a secondary alcohol dehydrogenase activity, and wherein the one or more co-product comprises isopropanol.

91. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, a secondary alcohol dehydrogenase activity, and a dehydratase activity, and wherein the one or more co-product comprises propene.

92. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, a 3-hydroxyisovalerate (3HIV) synthase activity, a hydroxymethylglutaryl-CoA synthase activity, a methylglutaconyl-CoA hydratase activity, a methylcrotonyl-CoA carboxylase activity, a methylcrotonyl-CoA hydratase activity, a 3-hydroxyisovaleryl-CoA thioesterase activity, a 3HIV kinase activity, a 3HIV-3-phosphate decarboxylase activity, and a 3HIV decarboxylase activity, and wherein the one or more co-product comprises isobutene.

93. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-

phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, and a glycerate 2-kinase activity, and wherein the one or more co-product comprises L-serine.

94. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a serine hydroxymethyltransferase activity, a transferase activity, a formaldehyde dehydrogenase activity, a formate dehydrogenase activity, an activity associated with glycine cleavage system, a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine transaminase activity, a hydroxypyruvate decarboxylase activity, a serine oxidoreductase (deaminating) activity, a serine decarboxylase activity, an ethanolamine aminotransferase or ethanolamine oxidoreductase (deaminating) activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, a glycerate 2-kinase activity, a glycolaldehyde dehydrogenase activity, a glycolate dehydrogenase activity, an alanine-glyoxylate aminotransferase activity, an alanine transaminase activity, and an NAD(P)H dependent glutamate dehydrogenase activity, and wherein the one or more co-product comprises glycine.

95. The method of any one of the preceding embodiments, wherein the activity associated with glycine cleavage system comprise an enzyme or protein selected from a glycine decarboxylase (P protein), an aminomethyltransferase (T protein), a dihydrolipoamide dehydrogenase (L protein), and an H protein.

96. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a transaminase activity, a hydroxypyruvate decarboxylase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a serine decarboxylase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, a glycerate 2-kinase activity, an acetaldehyde dehydrogenase activity, and an ethanolamine ammonia lyase activity, and wherein the one or more co-product comprises monoethanolamine (MEA).

97. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a serine dehydrogenase activity, a 2-aminomalonate semialdehyde decarboxylase activity, an aminoacetaldehyde transaminase activity, a 2-aminomalonate semialdehyde transaminase activity, a 2,3-diaminopropanoate decarboxylase activity, a serine decarboxylase activity, an ethanolamine dehydrogenase activity, a serine hydroxymethyltransferase activity, an aldehyde oxidase activity, an N-acetyl transferase or O-acetyl transferase activity, an N-acetylserine dehydrogenase activity, a transaminase activity, a deacetylase activity, a serine aminase activity, and a 2,3-diaminopropanoate ammonia lyase activity, and wherein the one or more co-product comprises ethylenediamine (EDA).

98. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of MEG or GA in the C2 pathway are selected from one or more enzyme having an activity selected from a D-tagatose 3-epimerase activity, a D-ribulokinase activity, a D-ribulose-1-phosphate aldolase activity, a D-xylulose 1-kinase activity, a D-xylulose-1-phosphate aldolase activity, a xylose reductase or aldose reductase activity, a xylitol dehydrogenase activity, a xylose isomerase activity, a xylose dehydrogenase activity, a xylonolactonase activity, a xylonate dehydrogenase activity, a 2-keto-3-deoxy-D-pentionate aldolase activity, a glycolaldehyde reductase activity and a glycolaldehyde dehydrogenase activity.

99. The method of any one of the preceding embodiments, wherein the method further comprises: introducing into the recombinant microorganism one or more modifications to diminish or delete activity in a glycolaldehyde dehydrogenase, a lactate dehydrogenase, a xylose isomerase, xylulose kinase or combination thereof.

100. A recombinant microorganism comprising one or more biochemical pathway that produces monoethylene glycol (MEG) or glycolic acid (GA) from one or more hexose feedstock via one or more pentose-5-phosphate intermediate.

101. The recombinant microorganism of any one of the preceding embodiments, wherein one or more co-product is co-produced with MEG or GA.

102. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more pentose-5-phosphate intermediate is one or more of D-ribose-5-phosphate, D-xylulose-5-phosphate or D-ribulose-5-phosphate.

103. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more biochemical pathway comprises expression of at least one enzyme having transketolase activity.

104. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having transketolase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to tktA tktB from *E. coli*.

105. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more biochemical pathway comprises expression of at least one enzyme having transaldolase activity.

106. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having transaldolase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to talB or talB from *E. coli*.

107. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more biochemical pathway comprises expression of at least one enzyme having ribulose-5-phosphate 3-epimerase activity.

108. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having ribulose-5-phosphate 3-epimerase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to rpe from *E. coli*.

109. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more biochemical pathway comprises expression of at least one enzyme having ribose-5-phosphate isomerase activity.

110. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having ribose-5-phosphate isomerase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to rpiA or rpiB from *E. coli*.

111. The recombinant microorganism of any one of the preceding embodiments, wherein the recombinant microorganism further comprises a deleted or diminished activity in at least one endogenous enzyme selected from glyceraldehyde 3-phosphate dehydrogenase, phosphoglycerate kinase and/or phosphoglycerate mutase.

112. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more biochemical pathway comprises expression of at least one enzyme having fructose-6-phosphate phosphoketolase activity.

113. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having fructose-6-phosphate phosphoketolase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having fructose-6-phosphate phosphoketolase activity selected from the group consisting of *Bifidobacterium dentium* BDP_1006, *Bifidobacterium lactis* xfp, *Lactobacillus paraplantarum* xpkA and *Bifidobacterium breve* xfp.

114. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more biochemical pathway comprises expression of at least one enzyme having phosphate acetyltransferase activity.

115. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having phosphate acetyltransferase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having phosphate acetyltransferase activity selected from *E. coli* pta and *Clostridium acetobutylicum* pta.

116. The recombinant microorganism of any one of the preceding embodiments, wherein the recombinant microorganism further comprises a deleted or diminished activity in an endogenous 6-phosphofructokinase enzyme.

117. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more biochemical pathway comprises at least one enzyme having pentose phosphatase activity.

118. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having pentose phosphatase activity is selected from one or more of an enzyme having D-pentose-5-phosphatase activity, an enzyme having D-xylulose-5-phosphatase activity, an enzyme having D-ribose-5-phosphatase activity, and an enzyme having D-ribulose-5-phosphatase activity.

119. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more biochemical pathway comprises at least one enzyme having arabitol phosphate dehydrogenase activity.

120. The recombinant microorganism of any one of the preceding embodiments, wherein the at least one enzyme having arabitol phosphate dehydrogenase activity is

selected from one or more of an enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity, an enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity, an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity and an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity.

5 121. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more biochemical pathway comprises at least one enzyme having phosphopentomutase activity.

122. The recombinant microorganism of any one of the preceding embodiments, wherein the recombinant microorganism further comprises a deleted or diminished activity in
10 one or more endogenous enzymes selected from glucose 6-phosphate-1-dehydrogenase, 6-phosphogluconolactonase and 6-phosphogluconate dehydrogenase.

123. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more co-product is selected from acetone, isopropanol, propene, isobutene and one or more serine pathway compounds.

15 124. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more serine pathway compounds is selected from serine, glycine, monoethanolamine (MEA) and ethylenediamine (EDA).

125. The recombinant microorganism of any one of the preceding embodiments, wherein the one or more hexose feedstock is selected from glucose or oligomers of glucose
20 thereof.

126. The recombinant microorganism of any one of the preceding embodiments, wherein the oligomers of glucose are selected from fructose, sucrose, starch, cellobiose, maltose, lactose and cellulose.

127. The recombinant microorganism of any one of the preceding embodiments,
25 wherein the expression of an enzyme having transketolase activity or an enzyme having fructose-6-phosphate phosphoketolase activity enables a lossless conversion of one or more hexose feedstock to the one or more pentose-5-phosphate intermediate.

128. The recombinant microorganism of any one of the preceding embodiments, wherein MEG or GA is produced through the conversion of glycolaldehyde in a C2 pathway
30 and through the conversion of dihydroxyacetone phosphate (DHAP) or pyruvate in a C3 pathway.

129. The recombinant microorganism of any one of the preceding embodiments, wherein GA is produced by the oxidation of glycolaldehyde by a glycolaldehyde dehydrogenase.

130. The recombinant microorganism of any one of the preceding embodiments, wherein one or more enzyme for the production of MEG or GA through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine transaminase activity, a hydroxypyruvate decarboxylase activity, a 3-phosphohydroxypyruvate reductase activity, a glycolaldehyde reductase activity, a glycolaldehyde dehydrogenase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a serine decarboxylase activity, an ethanolamine aminotransferase or ethanolamine oxidoreductase (deaminating) activity, a glycerate decarboxylase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, and a glycerate 2-kinase activity.

131. The recombinant microorganism of any one of the preceding embodiments, wherein MEG or GA is produced through the conversion of glycolaldehyde in a C2 pathway and one or more co-product is produced through the conversion of dihydroxyacetone phosphate (DHAP) or pyruvate in a C3 pathway.

132. The recombinant microorganism of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, and an acetoacetate decarboxylase activity, and wherein the one or more co-product comprises acetone.

133. The recombinant microorganism of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, and a secondary alcohol dehydrogenase activity, and wherein the one or more co-product comprises isopropanol.

134. The recombinant microorganism of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, a secondary alcohol dehydrogenase activity, and a dehydratase-isomerase activity, and wherein the one or more co-product comprises propene.

135. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, a 3-hydroxyisovalerate (3HIV) synthase activity, a hydroxymethylglutaryl-CoA synthase activity, a methylglutaconyl-CoA hydratase activity, a methylcrotonyl-CoA carboxylase activity, a methylcrotonyl-CoA hydratase activity, a 3-hydroxyisovaleryl-CoA thioesterase activity, a 3HIV kinase activity, a 3HIV-3-phosphate decarboxylase activity, and a 3HIV decarboxylase activity, and wherein the one or more co-product comprises isobutene.

136. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, and a glycerate 2-kinase activity, and wherein the one or more co-product comprises L-serine.

137. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a serine hydroxymethyltransferase activity, a transferase activity, a formaldehyde dehydrogenase activity, a formate dehydrogenase activity, an activity associated with glycine cleavage system, a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a

phosphoserine phosphatase activity, a serine transaminase activity, a hydroxypyruvate decarboxylase activity, a serine oxidoreductase (deaminating) activity, a serine decarboxylase activity, an ethanolamine aminotransferase or ethanolamine oxidoreductase (deaminating) activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate
 5 phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, a glycerate 2-kinase activity, a glycolaldehyde dehydrogenase activity, a glycolate dehydrogenase activity, an alanine-glyoxylate aminotransferase activity, an alanine transaminase activity, an NAD(P)H dependent glutamate dehydrogenase activity, and wherein the one or more co-product comprises glycine.

10 138. The method of any one of the preceding embodiments, wherein the activity associated with glycine cleavage system comprise an enzyme or protein selected from a glycine decarboxylase (P protein), an aminomethyltransferase (T protein), a dihydrolipoamide dehydrogenase (L protein), and an H protein.

15 139. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a transaminase activity, a hydroxypyruvate decarboxylase activity, a serine
 20 oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a serine decarboxylase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, a glycerate 2-kinase activity, an acetaldehyde dehydrogenase activity, and an ethanolamine ammonia lyase activity, and wherein the one or more co-product comprises
 25 monoethanolamine (MEA).

140. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a serine dehydrogenase activity, a 2-aminomalonate semialdehyde decarboxylase
 30 activity, an aminoacetaldehyde transaminase activity, a 2-aminomalonate semialdehyde transaminase activity, a 2,3-diaminopropanoate decarboxylase activity, a serine decarboxylase activity, an ethanolamine dehydrogenase activity, a serine hydroxymethyltransferase activity, an aldehyde oxidase activity, an N-acetyl transferase or O-acetyl transferase activity, an N-acetylserine dehydrogenase activity, a transaminase
 35 activity, a deacetylase activity, a serine aminase activity, and a 2,3-diaminopropanoate

ammonia lyase activity, and wherein the one or more co-product comprises ethylenediamine (EDA).

141. The recombinant microorganism of any one of the preceding embodiments, wherein one or more enzyme for the production of MEG or GA from glycolaldehyde in a C2 pathway are selected from one or more enzyme having an activity selected from a D-tagatose 3-epimerase activity, a D-ribulokinase activity, a D-ribulose-1-phosphate aldolase activity, a D-xylulose 1-kinase activity, a D-xylulose-1-phosphate aldolase activity, a xylose reductase or aldose reductase activity, a xylitol dehydrogenase activity, a xylose isomerase activity, a xylose dehydrogenase activity, a xylonolactonase activity, a xylonate dehydrogenase activity, a 2-keto-3-deoxy-D-pentonate aldolase activity, a glycolaldehyde reductase activity and a glycolaldehyde dehydrogenase activity.

142. The recombinant microorganism of any one of the preceding embodiments, wherein the recombinant microorganism further comprises one or more modifications to diminish or delete activity in a glycolaldehyde dehydrogenase, a lactate dehydrogenase, a xylose isomerase, a xylulose kinase or combination thereof.

143. The recombinant microorganism of any one of the preceding embodiments, wherein at least a portion of the excess NADH produced in the C3 pathway is used as a source of reducing equivalents in the C2 pathway.

144. The recombinant microorganism of any one of the preceding embodiments, wherein at least a portion of the excess NADH produced in the C3 pathway is used to produce ATP.

145. The recombinant microorganism of any one of the preceding embodiments, wherein excess biomass formation is minimized and production of MEG (or GA) or MEG (or GA) and one or more co-product is maximized.

146. A method of producing MEG or glycolic acid (GA) using a recombinant microorganism of any of the preceding Embodiments, wherein the method comprises cultivating the recombinant microorganism in a culture medium containing one or more hexose feedstock providing a carbon source until the MEG or GA is produced.

147. The method of any one of the preceding embodiments, wherein one or more co-product is co-produced with MEG or GA.

148. The method of any one of the preceding embodiments, wherein the one or more co-product is selected from acetone, isopropanol, propene, isobutene and one or more serine pathway compound.

149. The method of any one of the preceding embodiments, wherein the one or more serine pathway compound is selected from serine, glycine, monoethanolamine (MEA) and ethylenediamine (EDA).

150. A method of producing a recombinant microorganism that produces or accumulates MEG or glycolic acid (GA) from one or more exogenous hexose feedstock via one or more pentose-5-phosphate intermediate, comprising:

introducing into or expressing in the recombinant microorganism one or more enzyme for the conversion of the one or more hexose feedstock to the one or more pentose-5-phosphate intermediate;

introducing into or expressing in the recombinant microorganism one or more enzyme for the conversion of the one or more pentose-5-phosphate intermediate to one or more pentose or pentose-1-phosphate intermediate;

introducing into or expressing in the recombinant microorganism a C2 pathway comprising one or more enzyme for the production of MEG or GA from glycolaldehyde; and

introducing into or expressing in the recombinant microorganism a C3 pathway comprising one or more enzyme for the production of MEG or GA derived from DHAP or pyruvate; and

culturing the recombinant microorganism in a culture medium containing the one or more hexose feedstock to produce or accumulate MEG or GA,

wherein glycolaldehyde and DHAP (or pyruvate) are intermediates produced in the C2 pathway, and wherein MEG or GA are produced in both the C2 and C3 pathways.

151. A method of producing a recombinant microorganism that produces or accumulates MEG or glycolic acid (GA) and one or more co-product from one or more exogenous hexose feedstock via one or more pentose-5-phosphate intermediate, comprising:

introducing into or expressing in the recombinant microorganism one or more enzyme for the conversion of the one or more hexose feedstock to the one or more pentose-5-phosphate intermediate;

introducing into or expressing in the recombinant microorganism one or more enzyme for the conversion of the one or more pentose-5-phosphate intermediate to one or more pentose or pentose-1-phosphate intermediate;

introducing into or expressing in the recombinant microorganism a C2 pathway comprising one or more enzyme for the production of MEG or GA from glycolaldehyde; and

introducing into or expressing in the recombinant microorganism a C3 pathway comprising one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate; and

culturing the recombinant microorganism in a culture medium containing the one or more hexose feedstock to produce or accumulate MEG or GA and one or more co-product, wherein glycolaldehyde and DHAP (or pyruvate) are intermediates produced in the C2 pathway, and wherein MEG or GA is produced in the C2 pathway and the one or more co-product is produced in the C3 pathway.

152. The method of any one of the preceding embodiments, wherein the glycolaldehyde is oxidized to glycolic acid by a glycolaldehyde dehydrogenase.

153. The method of any one of the preceding embodiments, wherein the one or more enzyme for the conversion of the one or more hexose feedstock to the one or more pentose-5-phosphate intermediate are selected from one or more enzyme having a transketolase activity, a transaldolase activity, a ribulose-5-phosphate 3-epimerase activity, and a ribose-5-phosphate isomerase activity.

154. The method of any one of the preceding embodiments, wherein the method further comprises introducing into the recombinant microorganism one or more modifications to diminish or delete activity in one or more endogenous enzymes selected from glyceraldehyde 3-phosphate dehydrogenase (gapA), phosphoglycerate kinase (pgk) and phosphoglycerate mutase (gpmA and/or gpmM).

155. The method of any one of the preceding embodiments, wherein the one or more enzyme for the conversion of the one or more hexose feedstock to the one or more pentose-5-phosphate intermediate are selected from one or more enzyme having a fructose-6-phosphate phosphoketolase activity, a phosphate acetyltransferase activity, a transketolase activity, a transaldolase activity, a ribulose-5-phosphate 3-epimerase activity, and a ribose-5-phosphate isomerase activity.

156. The method of any one of the preceding embodiments, wherein the method further comprises introducing into the recombinant microorganism one or more modifications to diminish or delete activity in an endogenous 6-phosphofructokinase (pfkA and/or pfkB) enzyme.

157. The method of any one of the preceding embodiments, wherein the method further comprises: introducing into the recombinant microorganism one or more modifications to diminish or delete activity in one or more endogenous enzymes selected from glucose 6-phosphate-1-dehydrogenase, 6-phosphogluconolactonase, and 6-phosphogluconate dehydrogenase.

158. The method of any one of the preceding embodiments, wherein the one or more enzyme for the conversion of the one or more pentose-5-phosphate intermediate to one or more pentose or pentose-1-phosphate intermediate are selected from one or more enzyme having a pentose phosphatase activity, a D-ribulose-5-phosphatase activity, a D-ribose-5-phosphatase activity, a D-ribose isomerase activity, an arabitol phosphate dehydrogenase activity, and a phosphopentomutase activity.

159. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of MEG or GA derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine transaminase activity, a hydroxypyruvate decarboxylase activity, a 3-phosphohydroxypyruvate reductase activity, a glycolaldehyde reductase activity, a glycolaldehyde dehydrogenase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a serine decarboxylase activity, an ethanolamine aminotransferase or ethanolamine oxidoreductase (deaminating) activity, a glycerate decarboxylase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, and a glycerate 2-kinase activity.

160. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, and an acetoacetate decarboxylase activity, and wherein the one or more co-product comprises acetone.

161. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase

activity, and a secondary alcohol dehydrogenase activity, and wherein the one or more co-product comprises isopropanol.

162. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, a secondary alcohol dehydrogenase activity, and a dehydratase activity, and wherein the one or more co-product comprises propene.

163. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, a 3-hydroxyisovalerate (3HIV) synthase activity, a hydroxymethylglutaryl-CoA synthase activity, a methylglutaconyl-CoA hydratase activity, a methylcrotonyl-CoA carboxylase activity, a methylcrotonyl-CoA hydratase activity, a 3-hydroxyisovaleryl-CoA thioesterase activity, a 3HIV kinase activity, a 3HIV-3-phosphate decarboxylase activity, and a 3HIV decarboxylase activity, and wherein the one or more co-product comprises isobutene.

164. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, and a glycerate 2-kinase activity, and wherein the one or more co-product comprises L-serine.

165. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a serine hydroxymethyltransferase activity, a transferase activity, a formaldehyde dehydrogenase activity, a formate dehydrogenase activity, an activity associated with glycine cleavage system, a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase

activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine transaminase activity, a hydroxypyruvate decarboxylase activity, a serine oxidoreductase (deaminating) activity, a serine decarboxylase activity, an ethanolamine aminotransferase or ethanolamine oxidoreductase (deaminating) activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, a glycerate 2-kinase activity, a glycolaldehyde dehydrogenase activity, a glycolate dehydrogenase activity, an alanine-glyoxylate aminotransferase activity, an alanine transaminase activity, and an NAD(P)H dependent glutamate dehydrogenase activity, and wherein the one or more co-product comprises glycine.

166. The method of any one of the preceding embodiments, wherein the activity associated with glycine cleavage system comprise an enzyme or protein selected from a glycine decarboxylase (P protein), an aminomethyltransferase (T protein), a dihydrolipoamide dehydrogenase (L protein), and an H protein.

167. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a transaminase activity, a hydroxypyruvate decarboxylase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a serine decarboxylase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, a glycerate 2-kinase activity, an acetaldehyde dehydrogenase activity, and an ethanolamine ammonia lyase activity, and wherein the one or more co-product comprises monoethanolamine (MEA).

168. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a serine dehydrogenase activity, a 2-aminomalonate semialdehyde decarboxylase activity, an aminoacetaldehyde transaminase activity, a 2-aminomalonate semialdehyde transaminase activity, a 2,3-diaminopropanoate decarboxylase activity, a serine decarboxylase activity, an ethanolamine dehydrogenase activity, a serine hydroxymethyltransferase activity, an aldehyde oxidase activity, an N-acetyl transferase or O-acetyl transferase activity, an N-acetylserine dehydrogenase activity, a transaminase activity, a deacetylase activity, a serine aminase activity, and a 2,3-diaminopropanoate ammonia lyase activity, and wherein the one or more co-product comprises ethylenediamine (EDA).

169. The method of any one of the preceding embodiments, wherein one or more enzyme for the production of MEG or GA in the C2 pathway are selected from one or more enzyme having an activity selected from a D-tagatose 3-epimerase activity, a D-ribulokinase activity, a D-ribulose-1-phosphate aldolase activity, a D-xylulose 1-kinase activity, a D-xylulose-1-phosphate aldolase activity, a xylose reductase or aldose reductase activity, a xylitol dehydrogenase activity, a xylose isomerase activity, a xylose dehydrogenase activity, a xylonolactonase activity, a xylonate dehydrogenase activity, a 2-keto-3-deoxy-D-pentionate aldolase activity, a glycolaldehyde reductase activity and a glycolaldehyde dehydrogenase activity.

170. The method of any one of the preceding embodiments, wherein the method further comprises: introducing into the recombinant microorganism one or more modifications to diminish or delete activity in a glycolaldehyde dehydrogenase, a lactate dehydrogenase, a xylose isomerase, a xylulose kinase or combination thereof.

[001293] The foregoing detailed description has been given for clearness of understanding only and no unnecessary limitations should be understood there from as modifications will be obvious to those skilled in the art.

[001294] While the disclosure has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications and this application is intended to cover any variations, uses, or adaptations of the disclosure following, in general, the principles of the disclosure and including such departures from the present disclosure as come within known or customary practice within the art to which the disclosure pertains and as may be applied to the essential features hereinbefore set forth and as follows in the scope of the appended claims.

[001295] The disclosures, including the claims, figures and/or drawings, of each and every patent, patent application, and publication cited herein are hereby incorporated herein by reference in their entireties

INCORPORATION BY REFERENCE

[001296] All references, articles, publications, patents, patent publications, and patent applications cited herein are incorporated by reference in their entireties for all purposes. However, mention of any reference, article, publication, patent, patent publication, and patent application cited herein is not, and should not be taken as, an acknowledgment or any form of suggestion that they constitute valid prior art or form part of the common general knowledge in any country in the world. Further, the following references are hereby incorporated by reference. See PCT publication WO2001053306; U.S. Patent No.

7,226,761B2; and EPO Publication No. EP2957640.

WHAT IS CLAIMED IS:

1. A recombinant microorganism comprising one or more biochemical pathway that produces monoethylene glycol (MEG) or glycolic acid (GA) from one or more hexose feedstock via one or more pentose-5-phosphate intermediates.
- 5 2. The recombinant microorganism of claim 1, wherein one or more co-product is co-produced with MEG or GA.
3. The recombinant microorganism of claim 1 or 2, wherein the one or more pentose-5-phosphate intermediate is one or more of D-ribose-5-phosphate, D-xylulose-5-phosphate or D-ribulose-5-phosphate.
- 10 4. The recombinant microorganism of any one of claims 1-3, wherein the one or more biochemical pathway comprises expression of at least one enzyme having transketolase activity.
5. The recombinant microorganism of claim 4, wherein the at least one enzyme having transketolase activity is encoded by an amino acid sequence having at least 70% sequence
15 identity, at least 80% sequence identity, or at least 90% sequence identity to tktA tktB from *E. coli*.
6. The recombinant microorganism of any one of claims 1-5, wherein the one or more biochemical pathway comprises expression of at least one enzyme having transaldolase activity.
- 20 7. The recombinant microorganism of claim 6, wherein the at least one enzyme having transaldolase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to talB or talB from *E. coli*.
8. The recombinant microorganism of any one of claims 1-7, wherein the one or more
25 biochemical pathway comprises expression of at least one enzyme having ribulose-5-phosphate 3-epimerase activity.
9. The recombinant microorganism of claim 8, wherein the at least one enzyme having ribulose-5-phosphate 3-epimerase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence
30 identity to rpe from *E. coli*.
10. The recombinant microorganism of any one of claims 1-9, wherein the one or more biochemical pathway comprises expression of at least one enzyme having ribose-5-phosphate isomerase activity.

11. The recombinant microorganism of claim 10, wherein the at least one enzyme having ribose-5-phosphate isomerase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to rpiA or rpiB from *E. coli*.

12. The recombinant microorganism of any one of claims 1-11, wherein the recombinant microorganism further comprises a deleted or diminished activity in at least one endogenous enzyme selected from glyceraldehyde 3-phosphate dehydrogenase, phosphoglycerate kinase and/or phosphoglycerate mutase.

13. The recombinant microorganism of any one of claims 1-11, wherein the one or more biochemical pathway comprises expression of at least one enzyme having fructose-6-phosphate phosphoketolase activity.

14. The recombinant microorganism of claim 13, wherein the at least one enzyme having fructose-6-phosphate phosphoketolase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having fructose-6-phosphate phosphoketolase activity selected from the group consisting of *Bifidobacterium dentium* BDP_1006, *Bifidobacterium lactis* xfp, *Lactobacillus paraplantarum* xpkA and *Bifidobacterium breve* xfp.

15. The recombinant microorganism of claim 13, wherein the one or more biochemical pathway comprises expression of at least one enzyme having phosphate acetyltransferase activity.

16. The recombinant microorganism of claim 15, wherein the at least one enzyme having phosphate acetyltransferase activity is encoded by an amino acid sequence having at least 70% sequence identity, at least 80% sequence identity, or at least 90% sequence identity to an enzyme having phosphate acetyltransferase activity selected from *E. coli* pta and *Clostridium acetobutylicum* pta.

17. The recombinant microorganism of any one of claims 1-11 or 13-16, wherein the recombinant microorganism further comprises a deleted or diminished activity in an endogenous 6-phosphofructokinase enzyme.

18. The recombinant microorganism of any one of claims 1-17, wherein the one or more biochemical pathway comprises at least one enzyme having pentose phosphatase activity.

19. The recombinant microorganism of claim 18, wherein the at least one enzyme having pentose phosphatase activity is selected from one or more of an enzyme having D-pentose-5-phosphatase activity, an enzyme having D-xylulose-5-phosphatase activity, an enzyme

having D-ribose-5-phosphatase activity, and an enzyme having D-ribulose-5-phosphatase activity.

20. The recombinant microorganism of any one of claims 1-19, wherein the one or more biochemical pathway comprises at least one enzyme having arabitol phosphate dehydrogenase activity.

21. The recombinant microorganism of claim 20, wherein the at least one enzyme having arabitol phosphate dehydrogenase activity is selected from one or more of an enzyme having D-arabitol 1-phosphate 4-dehydrogenase activity, an enzyme having D-arabitol 5-phosphate 2-dehydrogenase activity, an enzyme having D-arabitol 1-phosphate 2-dehydrogenase activity and an enzyme having D-arabitol 5-phosphate 4-dehydrogenase activity.

22. The recombinant microorganism of any one of claims 1-21, wherein the one or more biochemical pathway comprises at least one enzyme having phosphopentomutase activity.

23. The recombinant microorganism of any one of claims 1-22, wherein the recombinant microorganism further comprises a deleted or diminished activity in one or more endogenous enzymes selected from glucose 6-phosphate-1-dehydrogenase, 6-phosphogluconolactonase and 6-phosphogluconate dehydrogenase.

24. The recombinant microorganism of any one of claims 1-23, wherein the one or more co-product is selected from acetone, isopropanol, propene, isobutene and one or more serine pathway compounds.

25. The recombinant microorganism of claim 24, wherein the one or more serine pathway compounds is selected from serine, glycine, monoethanolamine (MEA) and ethylenediamine (EDA).

26. The recombinant microorganism of any one of claims 1-25, wherein the one or more hexose feedstock is selected from glucose or oligomers of glucose thereof.

27. The recombinant microorganism of claim 26, wherein the oligomers of glucose are selected from fructose, sucrose, starch, cellobiose, maltose, lactose and cellulose.

28. The recombinant microorganism of any one of claims 1-27, wherein the expression of an enzyme having transketolase activity or an enzyme having fructose-6-phosphate phosphoketolase activity enables a lossless conversion of one or more hexose feedstock to the one or more pentose-5-phosphate intermediate.

29. The recombinant microorganism of any one of claims 1-28, wherein MEG or GA is produced through the conversion of glycolaldehyde in a C2 pathway and through the conversion of dihydroxyacetone phosphate (DHAP) or pyruvate in a C3 pathway.

30. The recombinant microorganism of claim 29, wherein GA is produced by the oxidation of glycolaldehyde by a glycolaldehyde dehydrogenase.

31. The recombinant microorganism of claim 29, wherein one or more enzyme for the production of MEG or GA through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine transaminase activity, a hydroxypyruvate decarboxylase activity, a 3-phosphohydroxypyruvate reductase activity, a glycolaldehyde reductase activity, a glycolaldehyde dehydrogenase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a serine decarboxylase activity, an ethanolamine aminotransferase or ethanolamine oxidoreductase (deaminating) activity, a glycerate decarboxylase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, and a glycerate 2-kinase activity.

32. The recombinant microorganism of any one of claims 1-31, wherein MEG or GA is produced through the conversion of glycolaldehyde in a C2 pathway and one or more co-product is produced through the conversion of dihydroxyacetone phosphate (DHAP) or pyruvate in a C3 pathway.

33. The recombinant microorganism of claim 32, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, and an acetoacetate decarboxylase activity, and wherein the one or more co-product comprises acetone.

34. The recombinant microorganism of claim 32, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, and a secondary alcohol dehydrogenase activity, and wherein the one or more co-product comprises isopropanol.

35. The recombinant microorganism of claim 32, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, a secondary alcohol dehydrogenase activity, and a dehydratase-isomerase activity, and wherein the one or more co-product comprises propene.

36. The method of claim 32, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, a 3-hydroxyisovalerate (3HIV) synthase activity, a hydroxymethylglutaryl-CoA synthase activity, a methylglutaconyl-CoA hydratase activity, a methylcrotonyl-CoA carboxylase activity, a methylcrotonyl-CoA hydratase activity, a 3-hydroxyisovaleryl-CoA thioesterase activity, a 3HIV kinase activity, a 3HIV-3-phosphate decarboxylase activity, and a 3HIV decarboxylase activity, and wherein the one or more co-product comprises isobutene.

37. The method of claim 32, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, and a glycerate 2-kinase activity, and wherein the one or more co-product comprises L-serine.

38. The method of claim 32, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a serine hydroxymethyltransferase activity, a transferase activity, a formaldehyde dehydrogenase activity, a formate dehydrogenase activity, an activity associated with glycine cleavage system, a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine transaminase activity, a hydroxypyruvate decarboxylase activity, a serine oxidoreductase (deaminating) activity, a serine decarboxylase activity, an ethanolamine aminotransferase or ethanolamine oxidoreductase (deaminating) activity, a hydroxypyruvate

reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, a glycerate 2-kinase activity, a glycolaldehyde dehydrogenase activity, a glycolate dehydrogenase activity, an alanine-glyoxylate aminotransferase activity, an alanine transaminase activity, an NAD(P)H dependent glutamate dehydrogenase activity, and wherein the one or more co-product comprises glycine.

39. The method of claim 38, wherein the activity associated with glycine cleavage system comprise an enzyme or protein selected from a glycine decarboxylase (P protein), an aminomethyltransferase (T protein), a dihydrolipoamide dehydrogenase (L protein), and an H protein.

40. The method of claim 32, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a transaminase activity, a hydroxypyruvate decarboxylase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a serine decarboxylase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, a glycerate 2-kinase activity, an acetaldehyde dehydrogenase activity, and an ethanolamine ammonia lyase activity, and wherein the one or more co-product comprises monoethanolamine (MEA).

41. The method of claim 32, wherein one or more enzyme for the production of one or more co-product through the conversion of DHAP or pyruvate in a C3 pathway are selected from one or more enzyme having an activity selected from a serine dehydrogenase activity, a 2-aminomalonate semialdehyde decarboxylase activity, an aminoacetaldehyde transaminase activity, a 2-aminomalonate semialdehyde transaminase activity, a 2,3-diaminopropanoate decarboxylase activity, a serine decarboxylase activity, an ethanolamine dehydrogenase activity, a serine hydroxymethyltransferase activity, an aldehyde oxidase activity, an N-acetyl transferase or O-acetyl transferase activity, an N-acetylserine dehydrogenase activity, a transaminase activity, a deacetylase activity, a serine aminase activity, and a 2,3-diaminopropanoate ammonia lyase activity, and wherein the one or more co-product comprises ethylenediamine (EDA).

42. The recombinant microorganism of any one of claims 29-41, wherein one or more enzyme for the production of MEG or GA from glycolaldehyde in a C2 pathway are selected from one or more enzyme having an activity selected from a D-tagatose 3-epimerase

activity, a D-ribulokinase activity, a D-ribulose-1-phosphate aldolase activity, a D-xylulose 1-kinase activity, a D-xylulose-1-phosphate aldolase activity, a xylose reductase or aldose reductase activity, a xylitol dehydrogenase activity, a xylose isomerase activity, a xylose dehydrogenase activity, a xylonolactonase activity, a xylonate dehydrogenase activity, a 2-keto-3-deoxy-D-pentonate aldolase activity, a glycolaldehyde reductase activity and a glycolaldehyde dehydrogenase activity.

43. The recombinant microorganism of any one of claims 1-42, wherein the recombinant microorganism further comprises one or more modifications to diminish or delete activity in a glycolaldehyde dehydrogenase, a lactate dehydrogenase, a xylose isomerase, a xylulose kinase or combination thereof.

44. The recombinant microorganism of any one of claims 1-43, wherein at least a portion of the excess NADH produced in the C3 pathway is used as a source of reducing equivalents in the C2 pathway.

45. The recombinant microorganism of any one of claims 1-43, wherein at least a portion of the excess NADH produced in the C3 pathway is used to produce ATP.

46. The recombinant microorganism of any one of claims 1-45, wherein excess biomass formation is minimized and production of MEG (or GA) or MEG (or GA) and one or more co-product is maximized.

47. A method of producing MEG or glycolic acid (GA) using a recombinant microorganism of any of the preceding claims, wherein the method comprises cultivating the recombinant microorganism in a culture medium containing one or more hexose feedstock providing a carbon source until the MEG or GA is produced.

48. The method of claim 47, wherein one or more co-product is co-produced with MEG or GA.

49. The method of claim 48, wherein the one or more co-product is selected from acetone, isopropanol, propene, isobutene and one or more serine pathway compound.

50. The method of claim 49, wherein the one or more serine pathway compound is selected from serine, glycine, monoethanolamine (MEA) and ethylenediamine (EDA).

51. A method of producing a recombinant microorganism that produces or accumulates MEG or glycolic acid (GA) from one or more exogenous hexose feedstock via one or more pentose-5-phosphate intermediate, comprising:

introducing into or expressing in the recombinant microorganism one or more enzyme for the conversion of the one or more hexose feedstock to the one or more pentose-5-phosphate intermediate;

5 introducing into or expressing in the recombinant microorganism one or more enzyme for the conversion of the one or more pentose-5-phosphate intermediate to one or more pentose or pentose-1-phosphate intermediate;

introducing into or expressing in the recombinant microorganism a C2 pathway comprising one or more enzyme for the production of MEG or GA from glycolaldehyde; and
10 introducing into or expressing in the recombinant microorganism a C3 pathway comprising one or more enzyme for the production of MEG or GA derived from DHAP or pyruvate; and

culturing the recombinant microorganism in a culture medium containing the one or more hexose feedstock to produce or accumulate MEG or GA,

wherein glycolaldehyde and DHAP (or pyruvate) are intermediates produced in the C2 pathway, and wherein MEG or GA are produced in both the C2 and C3 pathways.

15 52. A method of producing a recombinant microorganism that produces or accumulates MEG or glycolic acid (GA) and one or more co-product from one or more exogenous hexose feedstock via one or more pentose-5-phosphate intermediate, comprising:

introducing into or expressing in the recombinant microorganism one or more enzyme for the conversion of the one or more hexose feedstock to the one or more pentose-5-phosphate intermediate;
20

introducing into or expressing in the recombinant microorganism one or more enzyme for the conversion of the one or more pentose-5-phosphate intermediate to one or more pentose or pentose-1-phosphate intermediate;

25 introducing into or expressing in the recombinant microorganism a C2 pathway comprising one or more enzyme for the production of MEG or GA from glycolaldehyde; and

introducing into or expressing in the recombinant microorganism a C3 pathway comprising one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate; and

30 culturing the recombinant microorganism in a culture medium containing the one or more hexose feedstock to produce or accumulate MEG or GA and one or more co-product, wherein glycolaldehyde and DHAP (or pyruvate) are intermediates produced in the C2 pathway, and wherein MEG or GA is produced in the C2 pathway and the one or more co-product is produced in the C3 pathway.

53. The method of claim 51 or 52, wherein the glycolaldehyde is oxidized to glycolic acid by a glycolaldehyde dehydrogenase.

54. The method of any one of claims 51-53, wherein the one or more enzyme for the conversion of the one or more hexose feedstock to the one or more pentose-5-phosphate intermediate are selected from one or more enzyme having a transketolase activity, a
5 transaldolase activity, a ribulose-5-phosphate 3-epimerase activity, and a ribose-5-phosphate isomerase activity.

55. The method of any one of claims 51-54, wherein the method further comprises introducing into the recombinant microorganism one or more modifications to diminish or
10 delete activity in one or more endogenous enzymes selected from glyceraldehyde 3-phosphate dehydrogenase (gapA), phosphoglycerate kinase (pgk) and phosphoglycerate mutase (gpmA and/or gpmM).

56. The method of any one of claims 51-53, wherein the one or more enzyme for the conversion of the one or more hexose feedstock to the one or more pentose-5-phosphate
15 intermediate are selected from one or more enzyme having a fructose-6-phosphate phosphoketolase activity, a phosphate acetyltransferase activity, a transketolase activity, a transaldolase activity, a ribulose-5-phosphate 3-epimerase activity, and a ribose-5-phosphate isomerase activity.

57. The method of claim 56, wherein the method further comprises introducing into the
20 recombinant microorganism one or more modifications to diminish or delete activity in an endogenous 6-phosphofructokinase (pfkA and/or pfkB) enzyme.

58. The method of any one of claims 51-57, wherein the method further comprises: introducing into the recombinant microorganism one or more modifications to diminish or
delete activity in one or more endogenous enzymes selected from glucose 6-phosphate-1-
25 dehydrogenase, 6-phosphogluconolactonase, and 6-phosphogluconate dehydrogenase.

59. The method of any one of claims 51-58, wherein the one or more enzyme for the conversion of the one or more pentose-5-phosphate intermediate to one or more pentose or
pentose-1-phosphate intermediate are selected from one or more enzyme having a pentose
phosphatase activity, a D-ribulose-5-phosphatase activity, a D-ribose-5-phosphatase activity,
30 a D-ribose isomerase activity, an arabitol phosphate dehydrogenase activity, and a phosphopentomutase activity.

60. The method of claim 51, wherein one or more enzyme for the production of MEG or GA derived from DHAP or pyruvate in the C3 pathway are selected from one or more
enzyme having an activity selected from a 3-phosphoglycerate dehydrogenase activity, a

phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine transaminase activity, a hydroxypyruvate decarboxylase activity, a 3-phosphohydroxypyruvate reductase activity, a glycolaldehyde reductase activity, a glycolaldehyde dehydrogenase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a serine decarboxylase activity, an ethanolamine aminotransferase or ethanolamine oxidoreductase (deaminating) activity, a glycerate decarboxylase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, and a glycerate 2-kinase activity.

61. The method of claim 52, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, and an acetoacetate decarboxylase activity, and wherein the one or more co-product comprises acetone.

62. The method of claim 52, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, and a secondary alcohol dehydrogenase activity, and wherein the one or more co-product comprises isopropanol.

63. The method of claim 52, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, a secondary alcohol dehydrogenase activity, and a dehydratase activity, and wherein the one or more co-product comprises propene.

64. The method of claim 52, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a thiolase or acetyl coenzyme A acetyltransferase activity, an acetyl-CoA:acetoacetate transferase or acetate:acetoacetyl-CoA hydrolase activity, an acetoacetate decarboxylase activity, a 3-hydroxyisovalerate (3HIV) synthase activity, a hydroxymethylglutaryl-CoA synthase activity, a methylglutaconyl-CoA hydratase activity, a methylcrotonyl-CoA carboxylase activity, a methylcrotonyl-CoA

hydratase activity, a 3-hydroxyisovaleryl-CoA thioesterase activity, a 3HIV kinase activity, a 3HIV-3-phosphate decarboxylase activity, and a 3HIV decarboxylase activity, and wherein the one or more co-product comprises isobutene.

65. The method of claim 52, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, and a glycerate 2-kinase activity, and wherein the one or more co-product comprises L-serine.

66. The method of claim 52, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a serine hydroxymethyltransferase activity, a transferase activity, a formaldehyde dehydrogenase activity, a formate dehydrogenase activity, an activity associated with glycine cleavage system, a 3-phosphoglycerate dehydrogenase activity, a phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a serine transaminase activity, a hydroxypyruvate decarboxylase activity, a serine oxidoreductase (deaminating) activity, a serine decarboxylase activity, an ethanolamine aminotransferase or ethanolamine oxidoreductase (deaminating) activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, a glycerate 2-kinase activity, a glycolaldehyde dehydrogenase activity, a glycolate dehydrogenase activity, an alanine-glyoxylate aminotransferase activity, an alanine transaminase activity, and an NAD(P)H dependent glutamate dehydrogenase activity, and wherein the one or more co-product comprises glycine.

67. The method of claim 66, wherein the activity associated with glycine cleavage system comprise an enzyme or protein selected from a glycine decarboxylase (P protein), an aminomethyltransferase (T protein), a dihydrolipoamide dehydrogenase (L protein), and an H protein.

68. The method of claim 52, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from 3-phosphoglycerate dehydrogenase activity, a

phosphoserine aminotransferase activity, a 3-phosphohydroxypyruvate phosphatase activity, a phosphoserine phosphatase activity, a transaminase activity, a hydroxypyruvate decarboxylase activity, a serine oxidoreductase (deaminating) or serine-pyruvate aminotransferase activity, a serine decarboxylase activity, a hydroxypyruvate reductase activity, a 3-phosphoglycerate phosphatase activity, a 2-phosphoglycerate phosphatase activity, a glycerate 3-kinase activity, a glycerate 2-kinase activity, an acetaldehyde dehydrogenase activity, and an ethanolamine ammonia lyase activity, and wherein the one or more co-product comprises monoethanolamine (MEA).

69. The method of claim 52, wherein one or more enzyme for the production of one or more co-product derived from DHAP or pyruvate in the C3 pathway are selected from one or more enzyme having an activity selected from a serine dehydrogenase activity, a 2-aminomalonate semialdehyde decarboxylase activity, an aminoacetaldehyde transaminase activity, a 2-aminomalonate semialdehyde transaminase activity, a 2,3-diaminopropanoate decarboxylase activity, a serine decarboxylase activity, an ethanolamine dehydrogenase activity, a serine hydroxymethyltransferase activity, an aldehyde oxidase activity, an N-acetyl transferase or O-acetyl transferase activity, an N-acetylserine dehydrogenase activity, a transaminase activity, a deacetylase activity, a serine aminase activity, and a 2,3-diaminopropanoate ammonia lyase activity, and wherein the one or more co-product comprises ethylenediamine (EDA).

70. The method of any one of claims 51-69, wherein one or more enzyme for the production of MEG or GA in the C2 pathway are selected from one or more enzyme having an activity selected from a D-tagatose 3-epimerase activity, a D-ribulokinase activity, a D-ribulose-1-phosphate aldolase activity, a D-xylulose 1-kinase activity, a D-xylulose-1-phosphate aldolase activity, a xylose reductase or aldose reductase activity, a xylitol dehydrogenase activity, a xylose isomerase activity, a xylose dehydrogenase activity, a xylonolactonase activity, a xylonate dehydrogenase activity, a 2-keto-3-deoxy-D-pentionate aldolase activity, a glycolaldehyde reductase activity and a glycolaldehyde dehydrogenase activity.

71. The method of any one of claims 51-70, wherein the method further comprises: introducing into the recombinant microorganism one or more modifications to diminish or delete activity in a glycolaldehyde dehydrogenase, a lactate dehydrogenase, a xylose isomerase, a xylulose kinase or combination thereof.

FIG. 1

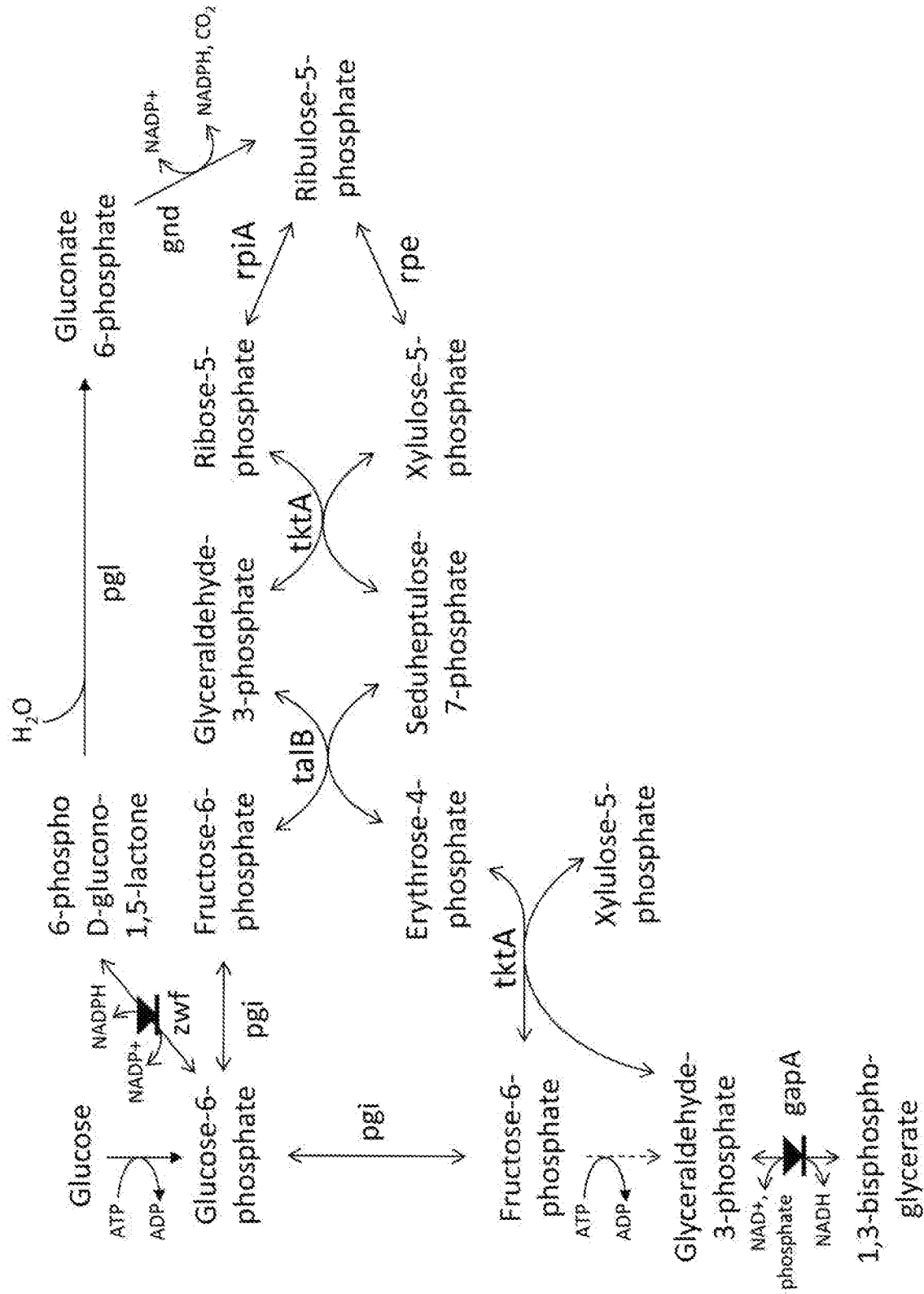


FIG. 3

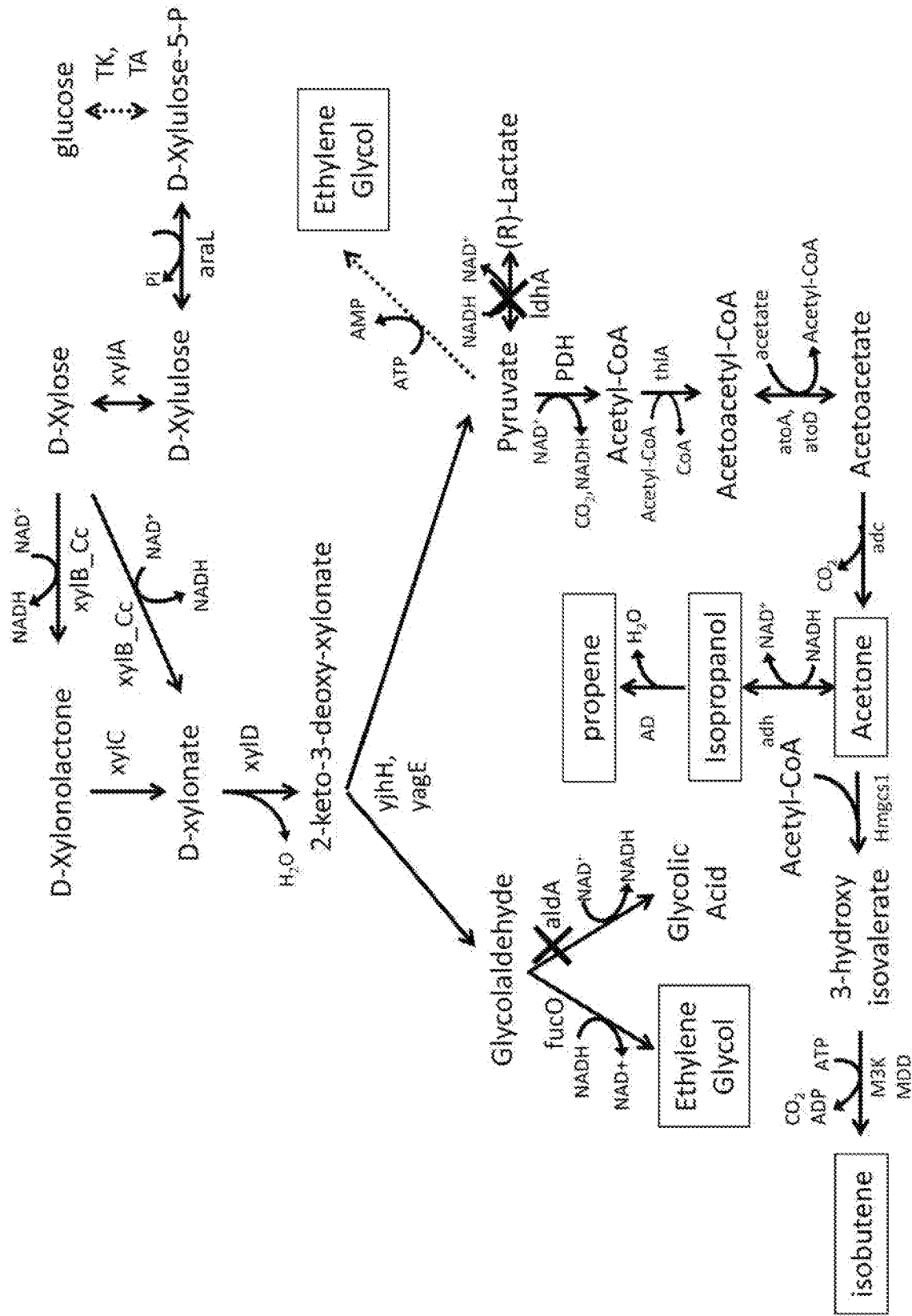
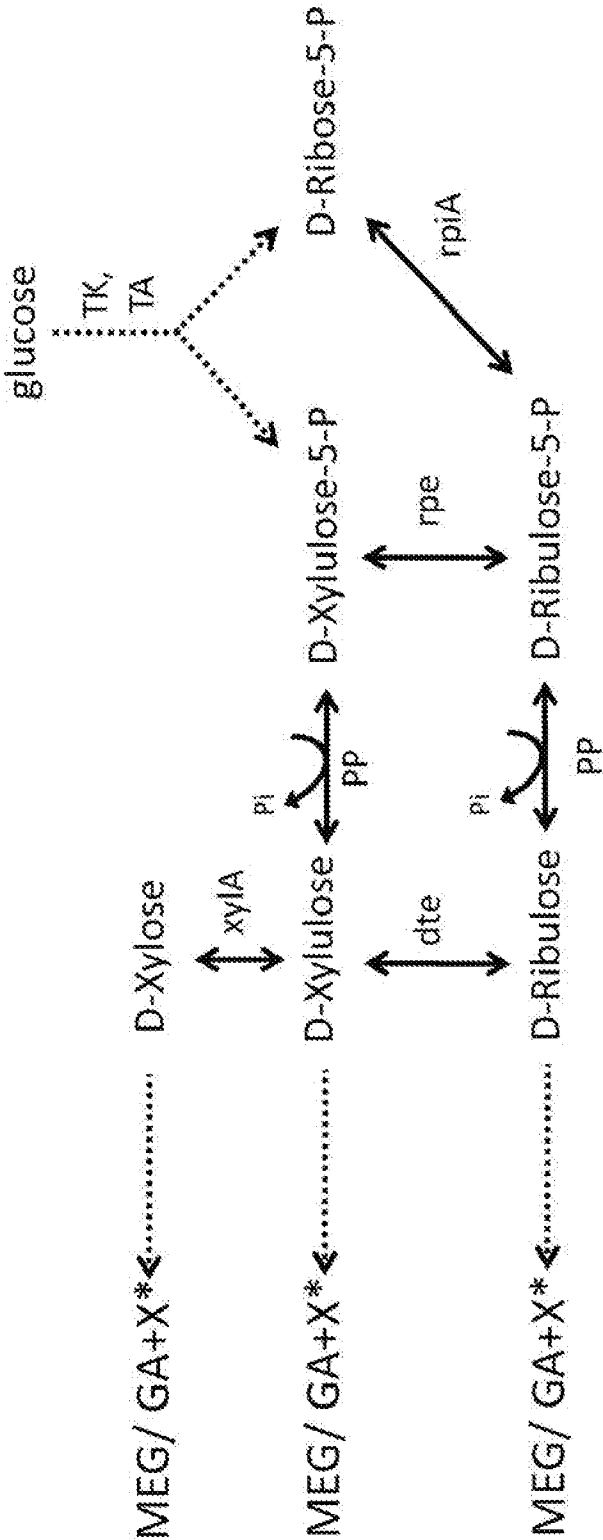
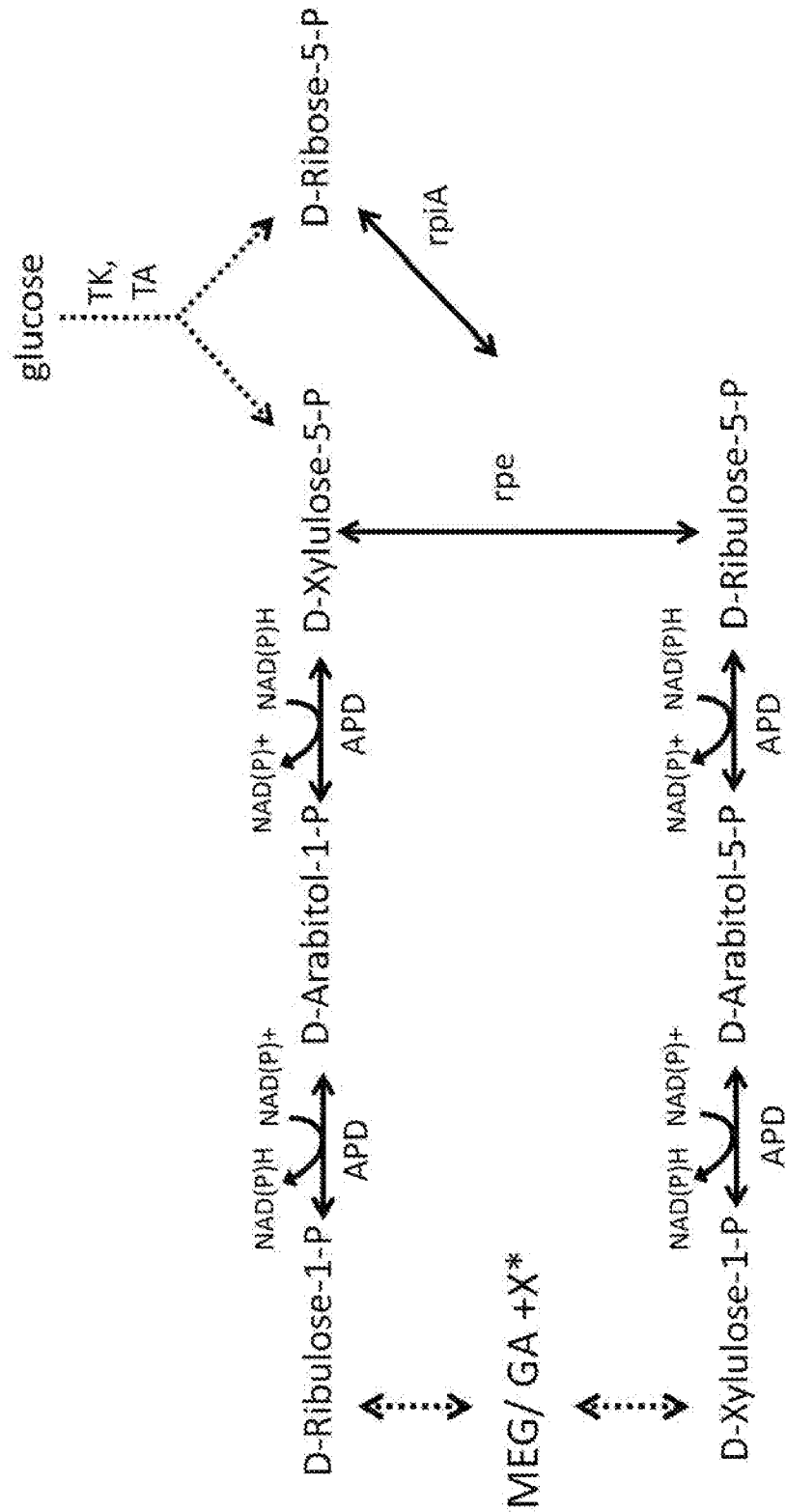


FIG. 5



*X = MEG/ GA/ acetone/ isopropanol/ propene/ isobutene

FIG. 6



*X = MEG/ GA/ acetone/ isopropanol/ propene/ isobutene

FIG. 7

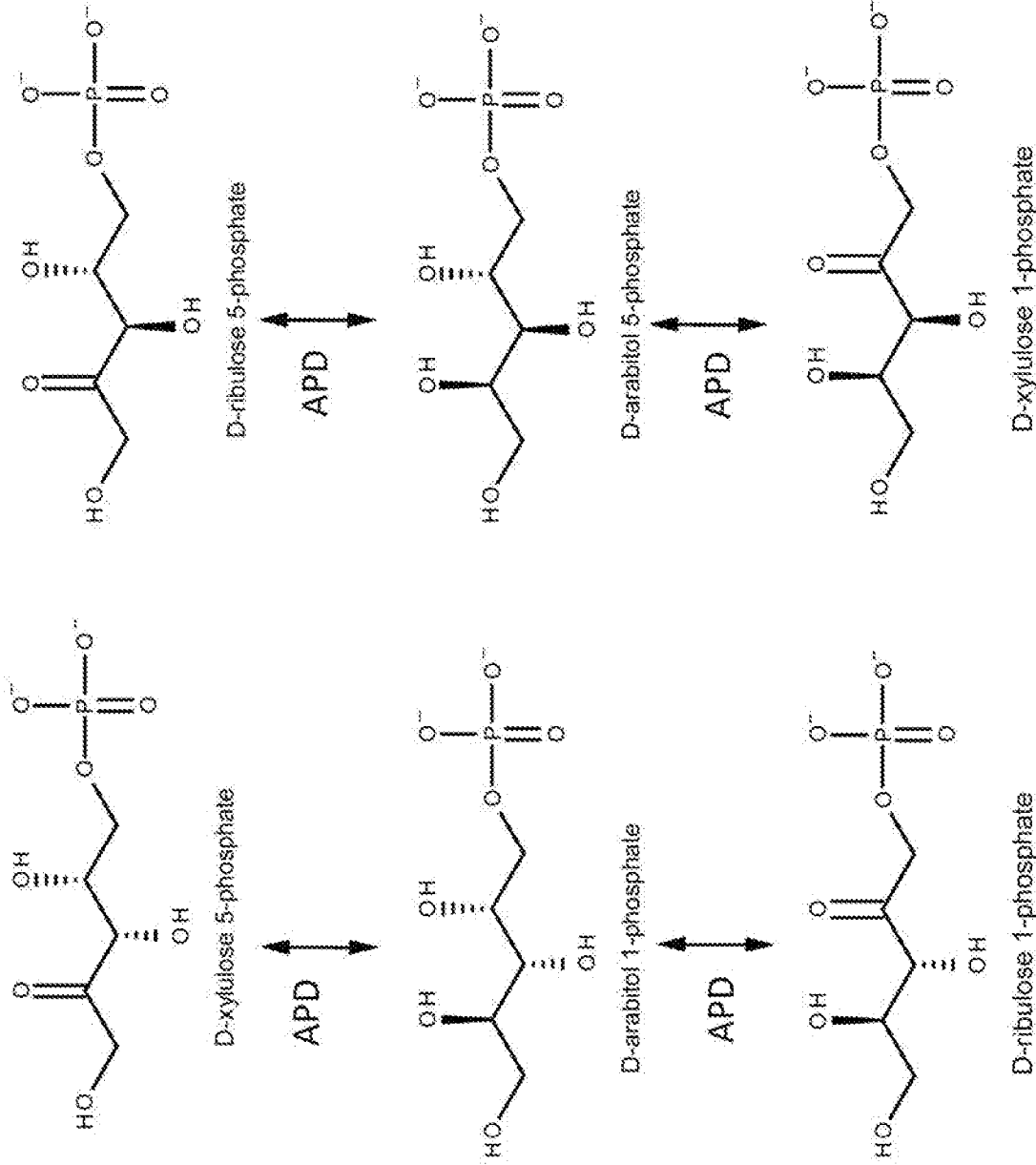
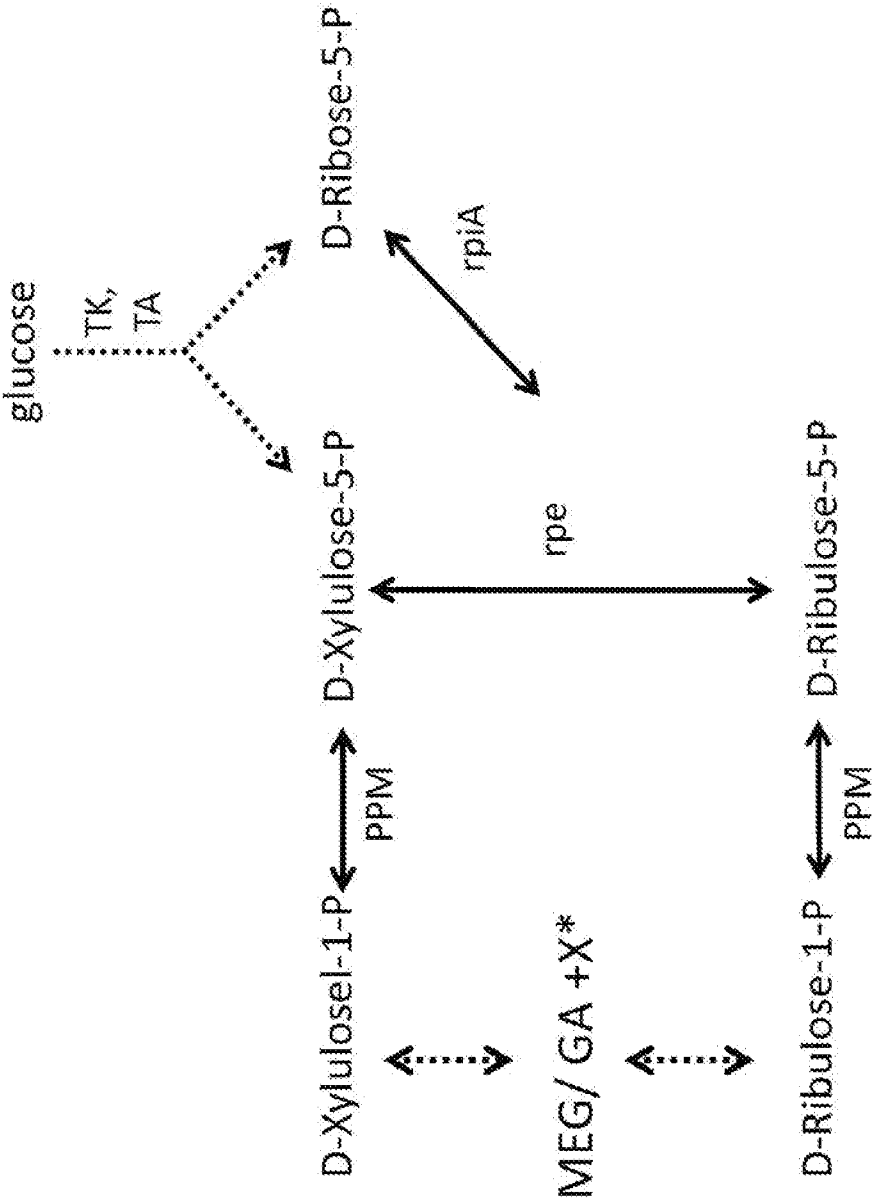
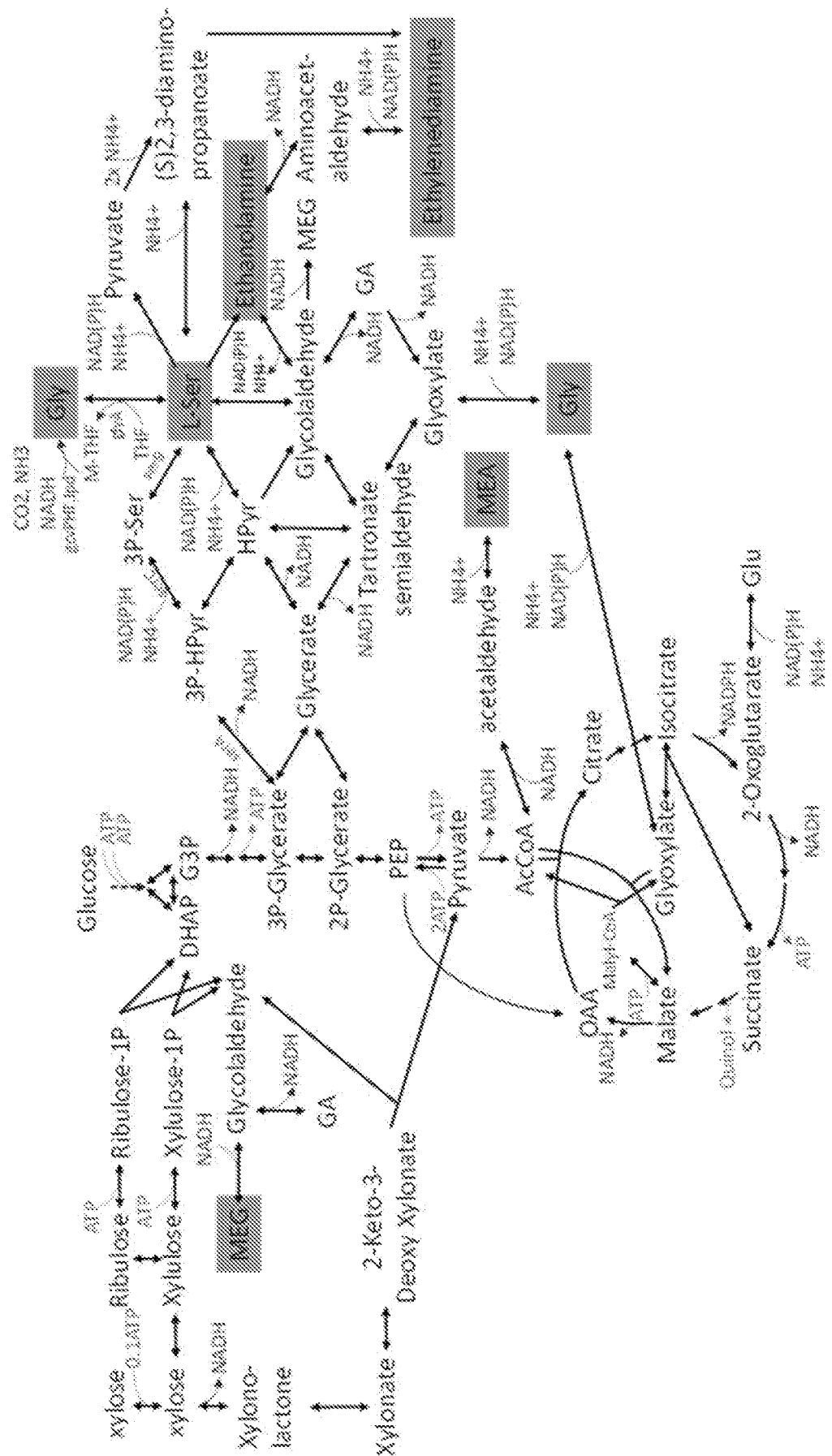


FIG. 8



*X = MEG/ GA/ acetone/ isopropanol/ propene/ isobutene

OFFICE



ॐ

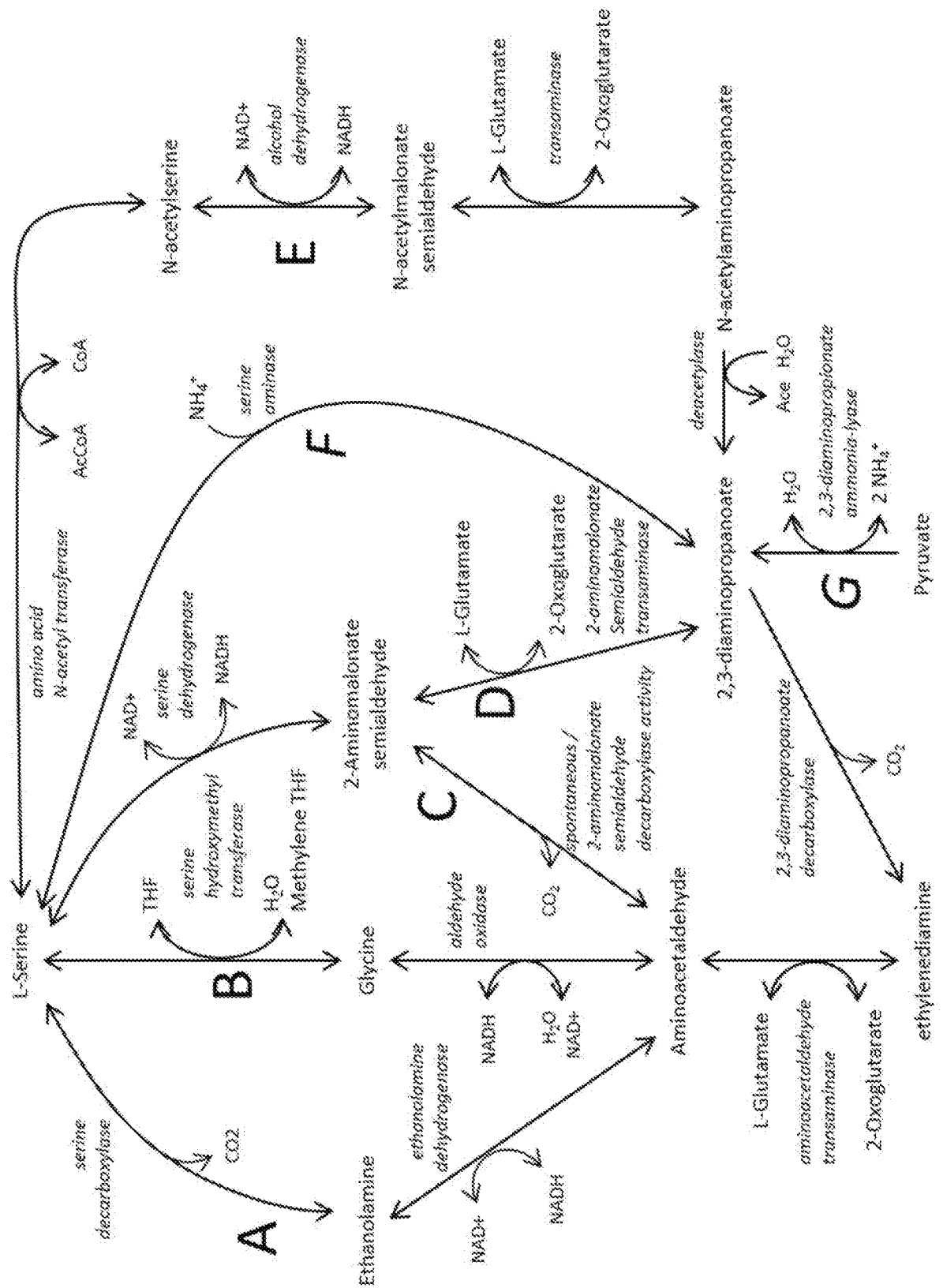


FIG. 12

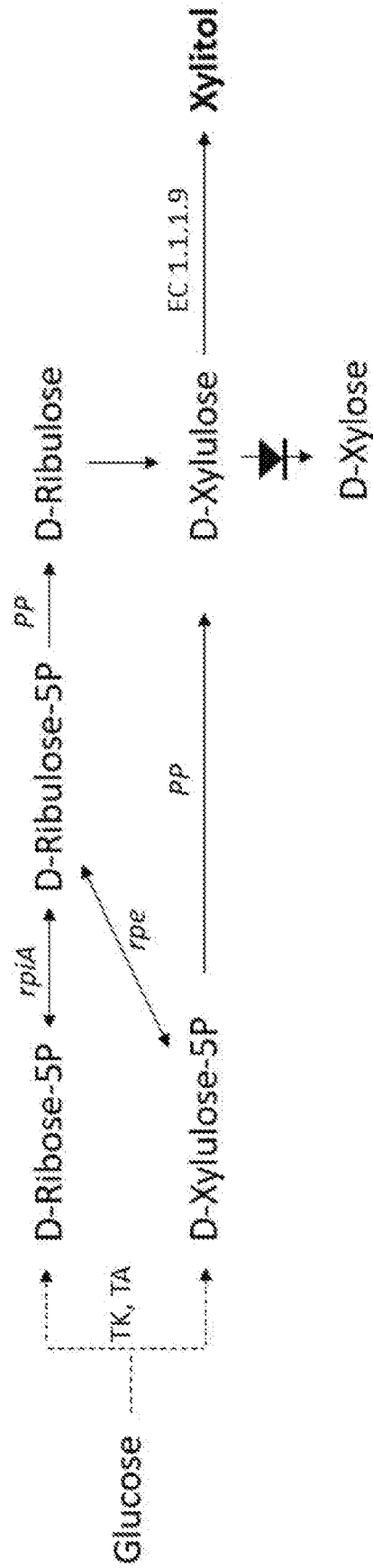


FIG. 13

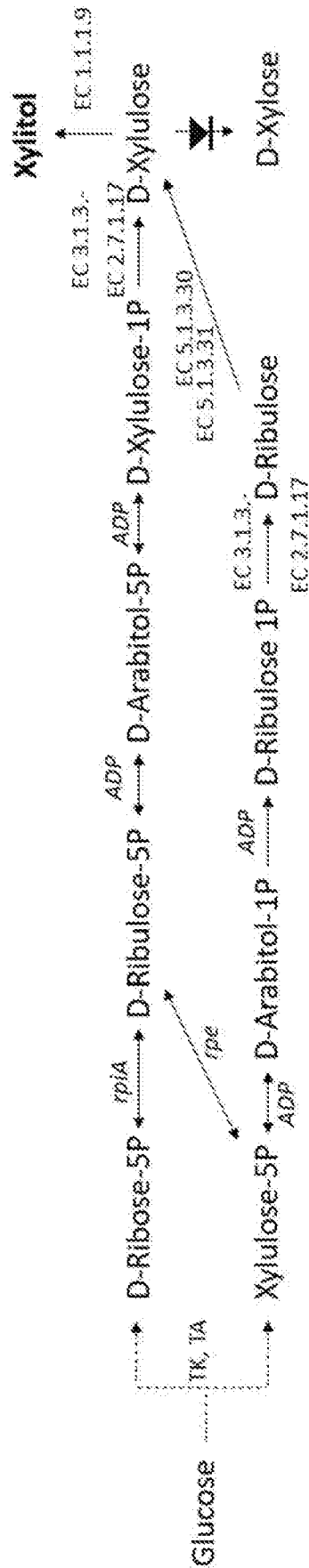


FIG. 14

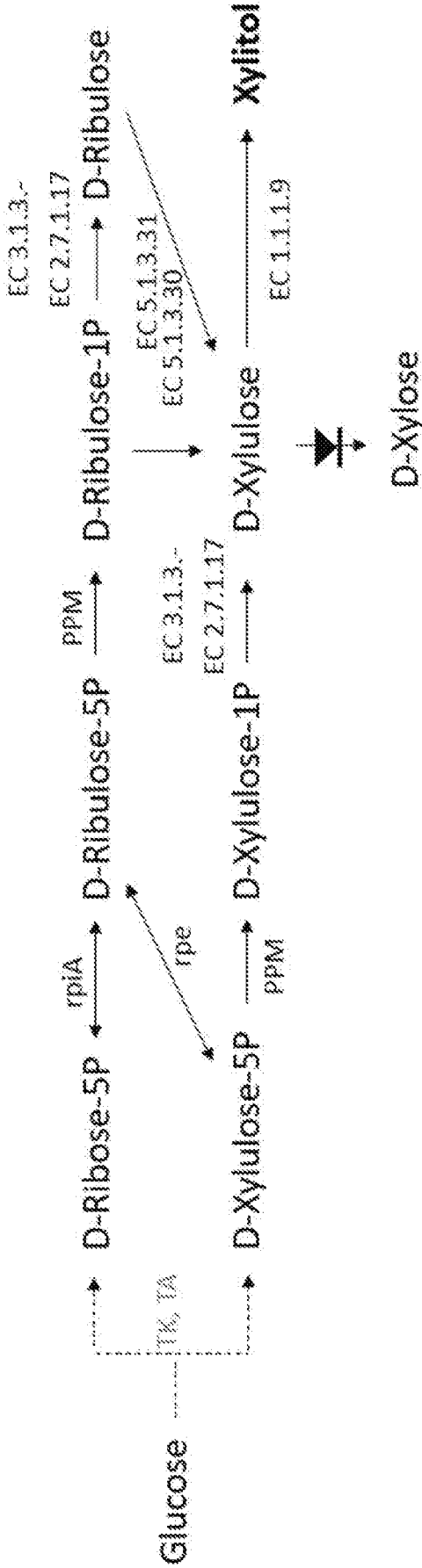


FIG. 15

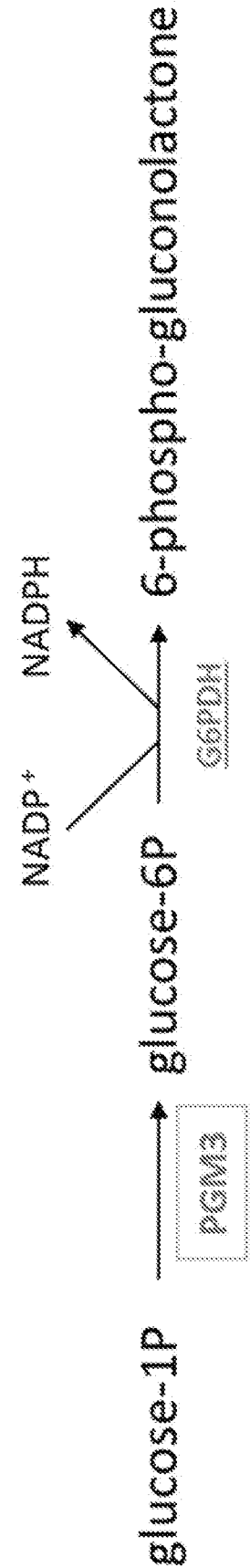


FIG. 16

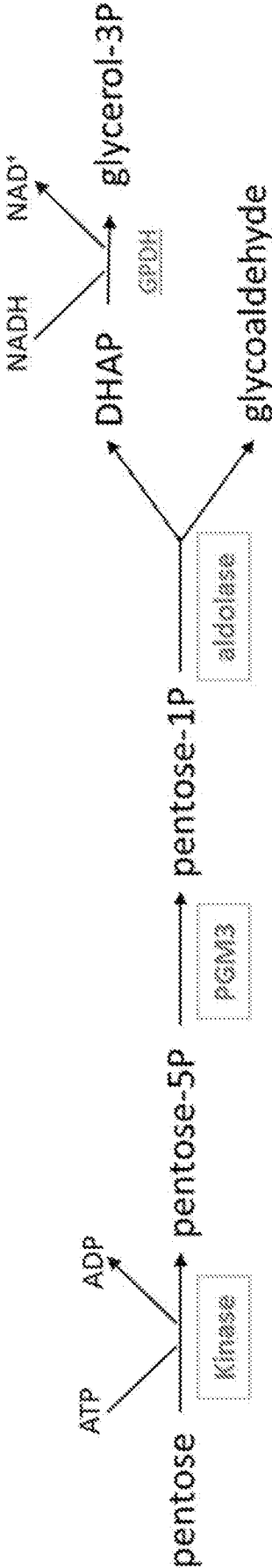
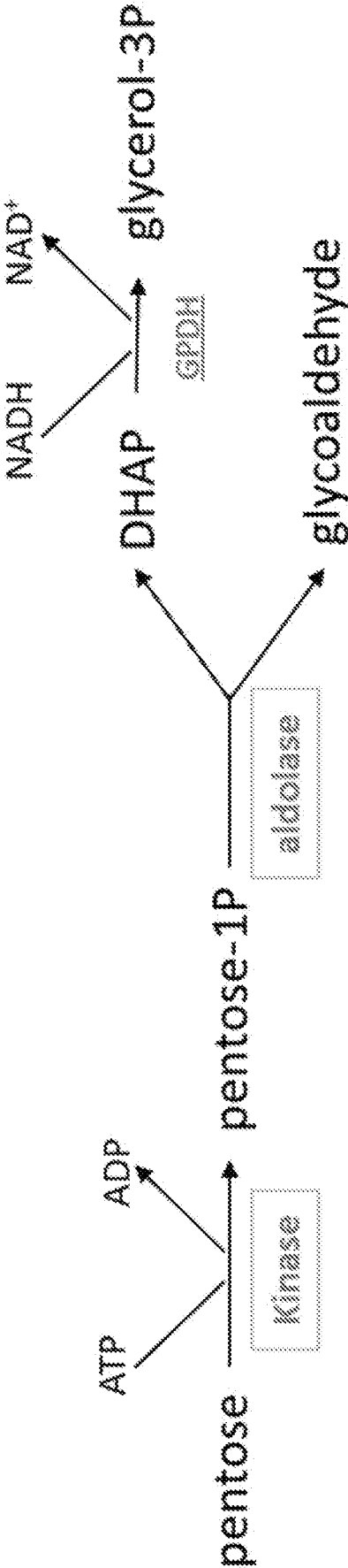


FIG. 18



INTERNATIONAL SEARCH REPORT

International application No
PCT/BR2020/050052

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N15/52 C12N9/02 C12P7/42
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 C12P

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.
X	WO 2014/187357 A1 (TIANJIN INST IND BIOTECHNOLOGY CAS [CN]) 27 November 2014 (2014-11-27)	1-11, 26-41, 51-54, 56,57
A	paragraph [0022] - paragraph [0052]; claims 3, 4 paragraph [0383] - paragraph [0397] paragraph [0038] paragraph [0050] paragraph [0605] ----- -/--	12-25, 42-50, 55,58-71



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

21 April 2020

Date of mailing of the international search report

30/04/2020

Name and mailing address of the ISA/

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

Authorized officer

Stoyanov, Borislav

INTERNATIONAL SEARCH REPORT

International application No
PCT/BR2020/050052

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2013/066568 A1 (DANISCO US INC [US]; GOODYEAR TIRE & RUBBER [US]) 10 May 2013 (2013-05-10)	1-11, 26-41, 51-57
A	paragraph [0111] - paragraph [0112]; figures 72, 73 paragraph [0125] paragraph [0197] - paragraph [0205] paragraph [0224] paragraph [0278] - paragraph [0279] paragraph [0347] - paragraph [0367]; example 24	12-25, 42-50, 58-71
A	----- EP 2 411 502 A1 (METABOLIC EXPLORER SA [FR]) 1 February 2012 (2012-02-01) claims 1-11 -----	1-71

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/BR2020/050052

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2014187357	A1	27-11-2014	CA 2913197 A1 27-11-2014
		CN 104178443 A 03-12-2014	
		EP 3006556 A1 13-04-2016	
		JP 6191061 B2 06-09-2017	
		JP 2016518145 A 23-06-2016	
		US 2016097064 A1 07-04-2016	
		WO 2014187357 A1 27-11-2014	
WO 2013066568	A1	10-05-2013	AU 2012332996 A1 24-04-2014
		BR 112014008278 A2 18-04-2017	
		CA 2851265 A1 10-05-2013	
		CN 104053765 A 17-09-2014	
		DK 2764084 T3 09-04-2018	
		EP 2764084 A1 13-08-2014	
		EP 3339426 A1 27-06-2018	
		HK 1201071 A1 21-08-2015	
		JP 6420147 B2 07-11-2018	
		JP 6639606 B2 05-02-2020	
		JP 2014531904 A 04-12-2014	
		JP 2019030315 A 28-02-2019	
		SG 11201401249P A 29-05-2014	
		US 2013089906 A1 11-04-2013	
		US 2015225744 A1 13-08-2015	
		US 2017198311 A1 13-07-2017	
		WO 2013066568 A1 10-05-2013	
EP 2411502	A1	01-02-2012	AR 075927 A1 04-05-2011
		BR PI1012693 A2 15-09-2015	
		CA 2754949 A1 30-09-2010	
		CN 102365357 A 29-02-2012	
		EP 2233562 A1 29-09-2010	
		EP 2411502 A1 01-02-2012	
		ES 2593084 T3 05-12-2016	
		JP 5800795 B2 28-10-2015	
		JP 2012521206 A 13-09-2012	
		KR 20120002593 A 06-01-2012	
		MY 164623 A 30-01-2018	
		US 2012315682 A1 13-12-2012	
		WO 2010108909 A1 30-09-2010	