



- (51) International Patent Classification:
C12N 9/88 (2006.01)
- (21) International Application Number:
PCT/US2013/039151
- (22) International Filing Date:
1 May 2013 (01.05.2013)
- (25) Filing Language:
English
- (26) Publication Language:
English
- (30) Priority Data:
61/641,823 2 May 2012 (02.05.2012) US
13/801,971 13 March 2013 (13.03.2013) US
- (71) Applicants: **DANISCO US INC.** [US/US]; 925 Page Mill Road, Palo Alto, CA 94304 (US). **THE GOODYEAR TIRE & RUBBER COMPANY** [US/US]; 1144 East Market Street, Akron, OH 44316-0001 (US).
- (72) Inventors: **BECK, Zachary, Q.**; 925 Page Mill Road, Palo Alto, CA 94304-1013 (US). **ESTELL, David, A.**; 925 Page Mill Road, Palo Alto, CA 94304-1013 (US). **MILLER, Jeffrey, V.**; 925 Page Mill Road, Palo Alto, CA 94304-1013 (US). **NGAI, James**; 925 Page Mill Road, Palo Alto, CA 94304-1013 (US). **RIFE, Christopher, L.**; 925 Page Mill Road, Palo Alto, CA 94304-1013 (US). **WELLS, Derek, H.**; 925 Page Mill Road, Palo Alto, CA 94304-1013 (US).

(74) Agents: **SHIEH-NEWTON, Terri** et al.; Morrison & Forster LLP, 755 Page Mill Road, Palo Alto, CA 94304-1018 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

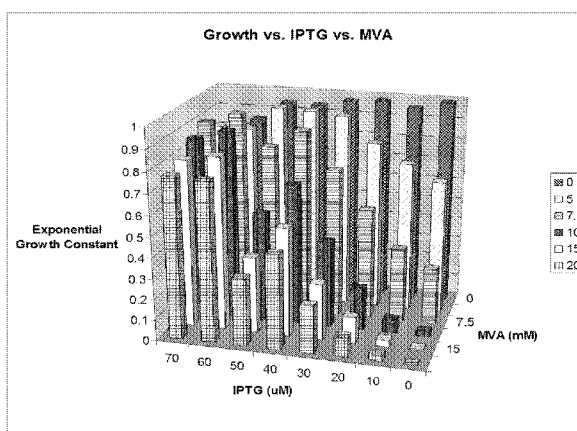
Published:

— without international search report and to be republished upon receipt of that report (Rule 48.2(g))

[Continued on next page]

(54) Title: IDENTIFICATION OF ISOPRENE SYNTHASE VARIANTS WITH IMPROVED PROPERTIES FOR THE PRODUCTION OF ISOPRENE

Figure 3



(57) Abstract: The invention provides for compositions and methods for producing isoprene using isoprene synthase variants with improved properties.

WO 2013/166211 A2

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

IDENTIFICATION OF ISOPRENE SYNTHASE VARIANTS WITH IMPROVED PROPERTIES FOR THE PRODUCTION OF ISOPRENE

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to U.S. Provisional Patent Application No. 61/641,823, filed May 2, 2012, and U.S. Utility Patent Application No. 13/801,971, filed March 13, 2013, the disclosures of which are incorporated by reference herein in their entirety.

INCORPORATION BY REFERENCE

[0002] The content of the following submission on ASCII text file (ST.25 text format) is incorporated herein by reference in its entirety: a computer readable form (CRF) of the Sequence Listing (file name is "643842004540_Sequence_Listing.txt"; date recorded: April 24, 2013; and the size of the ASCII text file in bytes is 77,654 bytes).

FIELD OF THE INVENTION

[0003] The present invention provides methods and compositions comprising isoprene synthase variants. In particular, the present invention provides isoprene synthase variants for increased isoprene production in host cells.

BACKGROUND OF THE INVENTION

[0004] Isoprene (2-methyl-1,3-butadiene) is a volatile hydrocarbon that is insoluble in water and soluble in alcohol. Commercially viable quantities of isoprene can be obtained by direct isolation from petroleum C5 cracking fractions or by dehydration of C5 isoalkanes or isoalkenes (Weissermel and Arpe, Industrial Organic Chemistry, 4th ed., Wiley-VCH, pp. 117-122, 2003). The C5 skeleton can also be synthesized from smaller subunits. It would be desirable, however, to have a commercially viable method of producing isoprene that was independent of nonrenewable resources.

[0005] Biosynthetic production of isoprene occurs by two distinct metabolic pathways (Julsing et al., Appl Microbiol Biotechnol, 75:1377-1384, 2007). In eukaryotes and archae, isoprene is formed via the mevalonate (MVA) pathway, while some eubacteria and higher plants produce isoprene via the methylerythritol phosphate (MEP) pathway. Isoprene emissions from

plants are light and temperature-dependent with increases linked to leaf development. An isoprene-producing enzyme, isoprene synthase, has been identified in Aspen trees (Silver and Fall, *Plant Physiol*, 97:1588-1591, (1991); and Silver and Fall, *J Biol Chem*, 270:13010-13016, (1995)) and is believed to be responsible for the *in vivo* production of isoprene from whole leaves. Bacterial production of isoprene has also been described (Kuzma et al., *Curr Microbiol*, 30:97-103, (1995); and Wilkins, *Chemosphere*, 32:1427-1434, (1996)), and varies in amount with the phase of bacterial growth and the nutrient content of the culture medium (U.S. Patent No. 5,849,970 to Fall et al.; and Wagner et al., *J Bacteriol*, 181:4700-4703, (1999)). The levels of isoprene obtainable through bacterial systems of the prior art, however, are insufficient for commercial uses. Thus what the art needs is an efficient, large scale isoprene biological production process to produce isoprene. The invention described herein addresses these problems and provides additional benefits as well.

BRIEF SUMMARY OF THE INVENTION

[0006] The invention provides, *inter alia*, compositions and methods for the increased production of isoprene by using polypeptides having isoprene synthase activity, wherein the polypeptides have one or more amino acid substitution(s) at one or more residues, such that the polypeptides have improved kinetic properties for the production of isoprene.

[0007] Accordingly in one aspect, the invention provides for isolated polypeptides having isoprene synthase activity, wherein the polypeptides comprise one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X323F, X118E, X36W, X22K, X228Y, X448L, X488F, X467H, X443S, X331P, X453I, X71K, X71L, X448I, X71M, X392Y, X448V, X282H, X383Y, X323Y, X511Y, X448E, X376M, X488L, X120E, X461A, X414I, X282W, X071K, X493E, X392S, X448Q, X282Y, X537N, X447Y, X240C, X443Q, X538R, and X510C, and wherein the polypeptides have improved K_{cat} compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the amino acid substitution is selected from the group consisting of: D323F, A118E, K36W, S22K, M228Y, A448L, E488F, E467H, A443S, C331P, A453I, R71K, R71L, A448I, R71M, W392Y, A448V, S282H, T383Y, D323Y, H511Y, A448E, L376M, E488L, S120E, R461A, K414I, S282W, R071K, S493E, W392S, A448Q, S282Y, E537N, I447Y, T240C, A443Q, P538R, and S510C.

[0008] In another aspect, the invention provides for isolated polypeptides having isoprene synthase activity, wherein the polypeptides comprise one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X389D, X383H, X36E, X436Y, X447V, X392F, X161R, X99D, X99E, X161C, X414S, X537T, X393V, X443G, X510V, X36N, X392A, X161A, X254R, X472C, X161Q, X36Y, X537C, X41Y, X161M, X43L, X120A, X087M, X36S, X36H, X348Y, X467W, X288T, X447T, X025N, and wherein the polypeptides have improved Kcat compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the amino acid substitution is selected from the group consisting of: G389D, T383H, K36E, L436Y, I447V, W392F, K161R, G99D, G99E, K161C, K414S, E537T, K393V, A443G, S510V, K36N, W392A, K161A, H254R, E472C, K161Q, K36Y, E537C, E41Y, K161M, R43L, S120A, G087M, K36S, K36H, K348Y, E467W, S288T, I447T, and D025N.

[0009] In another aspect, the invention provides for isolated polypeptides having isoprene synthase activity, wherein the polypeptides comprise one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X268I, X488M, X480I, X240V, X460A, X071L, X36Q, X118P, X36T, X282I, X409I, X071I, X409T, X436F, X348F, X118Q, X342Y, X74Q, X41P, X437Y, X392T, X89D, X41M, X510E, X36P, X393I, X71I, X381M, X374Y, X58Y, X526Q, X543F, X240M, X111S, X463T, X120M, and X135G, and wherein the polypeptides have improved Kcat compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the amino acid substitution is selected from the group consisting of: V268I, E488M, E480I, T240V, M460A, R071L, K36Q, A118P, K36T, S282I, V409I, R071I, V409T, L436F, K348F, A118Q, I342Y, S74Q, E41P, C437Y, W392T, F89D, E41M, S510E, K36P, K393I, R71I, T381M, K374Y, E58Y, L526Q, E543F, T240M, G111S, K463T, S120M, and E135G.

[0010] In another aspect, the invention provides for isolated polypeptides having isoprene synthase activity, wherein the polypeptides comprise one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X22K, X348F, X392V, X392F, X488C, X22R, X71V, X443R, X234R, X453V, X437Y, X392C, X463F, X538K, X393L, X254C, X436Y, X21R, X437L, X444D, X374Y, X363L, X447V, X444E, X71I, X504F, X488W, X71H, X36P, X381I, X436F, X460A, X443G, X288A,

X392T, X537I, X374Y, X242G, X437M, X436I, X376I, X288Y, and X392M, and wherein the polypeptides have improved K_M compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the amino acid substitution is selected from the group consisting of: S22K, K348F, W392V, W392F, E488C, S22R, R71V, A443R, Q234R, A453V, C437Y, W392C, K463F, P538K, K393L, H254C, L436Y, S21R, C437L, S444D, K374Y, A363L, I447V, S444E, R71I, I504F, E488W, R71H, K36P, T381I, L436F, M460A, A443G, S288A, W392T, E537I, K374Y, R242G, C437M, L436I, L376I, S288Y, and W392M.

[0011] In another aspect, the invention provides for isolated polypeptides having isoprene synthase activity, wherein the polypeptides comprise one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X342Y, X414W, X461A, X443S, X392A, X463T, X488T, X526Q, X502F, X502M, X488M, X36Y, X414R, X071I, X36W, X502L, X481Y, X472R, X36H, X36T, X415Y, X415H, X58Y, X381M, X481V, X89E, X480I, X36Q, X89D, X254R, X161N, X071L, X424P, X415V, X22K, X58L, X392S, X268I, X392I, X526E, X537V, and X36D and wherein the polypeptides have improved K_M compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the amino acid substitution is selected from the group consisting of: I342Y, K414W, R461A, A443S, W392A, K463T, E488T, L526Q, T502F, T502M, E488M, K36Y, K414R, R071I, K36W, T502L, T481Y, E472R, K36H, K36T, E415Y, E415H, E58Y, T381M, T481V, F89E, E480I, K36Q, F89D, H254R, K161N, R071L, H424P, E415V, S22K, E58L, W392S, V268I, W392I, L526E, E537V, and K36D.

[0012] In another aspect, the invention provides for isolated polypeptides having isoprene synthase activity, wherein the polypeptides comprise one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X393I, X41M, X025N, X36N, X118Q, X389E, X282W, X348Y, X383L, X510V, X323Y, X323F, X383H, X41P, X288T, X36S, X453I, X447T, X472I, X376M, X161E, X488L, X381L, X071K, X409T, X392Y, X135G, X542L, X228Y, X118P, X543F, and X36E, and wherein the polypeptides have improved K_M compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the amino acid substitution is selected from the group consisting of: K393I, E41M, D025N, K36N, A118Q, G389E, S282W, K348Y, T383L, S510V, D323Y, D323F, T383H, E41P, S288T, K36S, A453I, I447T, E472I,

L376M, K161E, E488L, T381L, R071K, V409T, W392Y, E135G, F542L, M228Y, A118P, E543F, and K36E.

[0013] In another aspect, the invention provides for isolated polypeptides having isoprene synthase activity, wherein the polypeptides comprise one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X472R, X389E, X242G, X376I, X161N, X288C, X240V, X481Y, X463F, and X393L, and wherein the polypeptides have improved K_{IDMAPP} compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the amino acid substitution is selected from the group consisting of: E472R, G389E, R242G, L376I, K161N, S288C, T240V, T481Y, K463F, and K393L.

[0014] In another aspect, the invention provides for isolated polypeptides having isoprene synthase activity, wherein the polypeptides comprise one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X120Q, X58L, X240M, and X453V, and wherein the polypeptides have improved K_{IDMAPP} compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the amino acid substitution is selected from the group consisting of: S120Q, E58L, T240M, and A453V.

[0015] In another aspect, the invention provides for isolated polypeptides having isoprene synthase activity, wherein the polypeptides comprise one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X543F, X74Q, X414R, X415H, and X415V, and wherein the polypeptides have improved K_{IDMAPP} compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the amino acid substitution is selected from the group consisting of: E543F, S74Q, K414R, E415H, and E415V.

[0016] In another aspect, the invention provides for isolated polypeptides having isoprene synthase activity, wherein the polypeptides comprise one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X510E, X161M, X120Q, X120E, X161R, X120A, X493E, X448E, X240V, X240M, and X537T, and wherein the polypeptides have improved specific productivity compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some

embodiments, the amino acid substitution is selected from the group consisting of: S510E, K161M, S120Q, S120E, K161R, S120A, S493E, A448E, T240V, T240M, and E537T.

[0017] In another aspect, the invention provides for isolated polypeptides having isoprene synthase activity, wherein the polypeptides comprise one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X161N, X447Y, X502M, X537N, X348Y, X543F, X331P, X537C, X120M, X288C, and X161Q, and wherein the polypeptides have improved specific productivity compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the amino acid substitution is selected from the group consisting of: K161N, I447Y, T502M, E537N, K348Y, E543F, C331P, E537C, S120M, S288C, and K161Q.

[0018] In another aspect, the invention provides for isolated polypeptides having isoprene synthase activity, wherein the polypeptide comprises one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X161A, X481V, X087M, X467W, X448V, X502L, and X376M, and wherein the polypeptide has improved specific productivity compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the amino acid substitution is selected from the group consisting of: K161A, T481V, G087M, E467W, A448V, T502L, and L376M.

[0019] In any of the aspects and/or embodiments herein, recombinant host cells comprise the polypeptide of any of the above. In some embodiments, the host cell is selected from the group consisting of a bacterial, algal, fungal, yeast, cyanobacterial, or Clostridial cell. In some embodiments, the host cell is a bacterial cell. In some embodiments, the bacterial cell is a gram-positive bacterial cell or gram-negative bacterial cell. In some embodiments, the bacterial cell is selected from the group consisting of *E. coli*, *L. acidophilus*, *P. citrea*, *B. subtilis*, *B. licheniformis*, *B. lentus*, *B. brevis*, *B. stearothermophilus*, *B. alkalophilus*, *B. amyloliquefaciens*, *B. clausii*, *B. halodurans*, *B. megaterium*, *B. coagulans*, *B. circulans*, *B. lautus*, *B. thuringiensis*, *S. albus*, *S. lividans*, *S. coelicolor*, *S. griseus*, *Pseudomonas sp.*, *P. alcaligenes*, *Clostridium sp.*, *Corynebacterium sp.*, and *C. glutamicum* cells. In some embodiments, the host cell is an algal cell. In some embodiments, the algal cell is selected from the group consisting of green algae, red algae, glaucophytes, chlorarachniophytes, euglenids, chromista, or dinoflagellates. In some embodiments, the host cell is a fungal cell. In some embodiments, the fungal cell is a

filamentous fungi. In some embodiments, the host cell is a yeast cell. In some embodiments, the yeast cell is selected from the group consisting of *Saccharomyces sp.*, *Schizosaccharomyces sp.*, *Pichia sp.*, or *Candida sp.* In some embodiments, the yeast cell is a *Saccharomyces cerevisiae* cell.

[0020] In any of the aspects and/or embodiments herein, methods for producing isoprene comprise (a) culturing the host cells as described herein, including those described above, under conditions suitable for the production of isoprene and (b) producing isoprene. In some embodiments, the method further comprises recovering the isoprene.

BRIEF DESCRIPTION OF THE DRAWINGS

[0021] Figure 1 shows the relationship between growth and DMAPP concentration in assay strain. DW425 was grown in the presence of various concentrations of mevalonate (0, 10, 20, 30, 40, 50 mM mevalonate in exp1, and 0, 2.5, 5, 10, 20 mM mevalonate in exp2) and IPTG (0 μ M IPTG and 50 μ M IPTG). Cells were harvested and collected for metabolite analysis upon completion of the growth experiment.

[0022] Figure 2 shows DW425 growth in the presence of varying concentrations of IPTG (0, 10, 20, 30, 40, 50, 60, and 70 μ M) and mevalonate (0, 5, 7.5, 10, 15, 20 mM).

[0023] Figure 3 shows growth as a function of [IPTG] and [mevalonate].

[0024] Figure 4 shows a map of pCL201.

[0025] Figure 5 shows a view of Wild Type IspS showing the location of Phe495 with respect to Gly491 and Leu494, all in stick representation.

[0026] Figure 6 shows a view of Wild Type IspS showing the location of Gln509, in stick representation, with respect to the active site. The magnesium ions and (+)-bornyl diphosphate in the active site are modeled based on a structural alignment with PDB 1N24.

[0027] Figure 7 shows a view of Wild Type IspS showing residue Phe287 in stick representation. (+)-bornyl diphosphate (BP) is placed based on a structural alignment with PDB 1N24.

[0028] Figure 8 shows a view of Wild Type IspS showing residue Gly397 in stick representation.

[0029] Figure 9 shows an active site view of Wild Type IspS showing residue Asn438 in stick representation. (+)-bornyl diphosphate (BP) and Mg²⁺ (spheres) are placed based on a structural alignment with PDB 1N24.

[0030] Figure 10 shows a view of Wild Type IspS showing residue Glu451 in stick representation.

[0031] Figure 11 shows an active site view of Wild Type IspS showing residue Tyr514 in stick representation. (+)-bornyl diphosphate (BP) and Mg²⁺ (spheres) are placed based on a structural alignment with PDB 1N24.

[0032] Figure 12 shows the MVA and DXP metabolic pathways for isoprene (based on F. Bouvier *et al.*, Progress in Lipid Res. 44: 357-429, 2005). The following description includes alternative names for each polypeptide in the pathways and a reference that discloses an assay for measuring the activity of the indicated polypeptide (each of these references are each hereby incorporated by reference in their entireties, particularly with respect to assays for polypeptide activity for polypeptides in the MVA and DXP pathways). **Mevalonate Pathway: AACT**; Acetyl-CoA acetyltransferase, MvaE, EC 2.3.1.9. Assay: J. Bacteriol., 184: 2116–2122, 2002; **HMGS**; Hydroxymethylglutaryl-CoA synthase, MvaS, EC 2.3.3.10. Assay: J. Bacteriol., 184: 4065–4070, 2002; **HMGR**; 3-Hydroxy-3-methylglutaryl-CoA reductase, MvaE, EC 1.1.1.34. Assay: J. Bacteriol., 184: 2116–2122, 2002; **MVK**; Mevalonate kinase, ERG12, EC 2.7.1.36. Assay: Curr Genet 19:9-14, 1991. **PMK**; Phosphomevalonate kinase, ERG8, EC 2.7.4.2, Assay: Mol Cell Biol., 11:620–631, 1991; **DPMDC**; Diphosphomevalonate decarboxylase, MVD1, EC 4.1.1.33. Assay: Biochemistry, 33:13355-13362, 1994; **IDI**; Isopentenyl-diphosphate delta-isomerase, IDI1, EC 5.3.3.2 or isopentenyl pyrophosphate. Assay: J. Biol. Chem. 264:19169-19175, 1989. **DXP Pathway: DXS**; 1-Deoxyxylulose-5-phosphate synthase, dxs, EC 2.2.1.7. Assay: PNAS, 94:12857-62, 1997; **DXR**; 1-Deoxy-D-xylulose 5-phosphate reductoisomerase, dxr, EC 2.2.1.7. Assay: Eur. J. Biochem. 269:4446–4457, 2002; **MCT**; 4-Diphosphocytidyl-2C-methyl-D-erythritol synthase, IspD, EC 2.7.7.60. Assay: PNAS, 97: 6451–6456, 2000; **CMK**; 4-Diphosphocytidyl-2-C-methyl-D-erythritol kinase, IspE, EC 2.7.1.148. Assay: PNAS, 97:1062-1067, 2000; **MCS**; 2C-Methyl-D-erythritol 2,4-cyclodiphosphate synthase, IspF, EC 4.6.1.12.

Assay: PNAS, 96:11758-11763, 1999; **HDS**; 1-Hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase, ispG, EC 1.17.4.3. Assay: J. Org. Chem., 70:9168 -9174, 2005; **HDR**; 1-Hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate reductase, IspH, EC 1.17.1.2. Assay: JACS, 126:12847-12855, 2004.

[0033] Figure 13 shows a monomer view of wild type IspS showing the location of sites where substitutions are not tolerated.

[0034] Figure 14 shows the location of residues in the active site of IspS that do not tolerate substitution. Mg²⁺ (spheres) are placed based on a structural alignment with PDB 1N24.

[0035] Figure 15 shows the location of buried sites in IspS that do not tolerate substitution. Mg²⁺ (spheres) are placed based on a structural alignment with PDB 1N24.

[0036] Figure 16 shows residues located at the dimer interface of IspS that do not tolerate substitution. Chain A is light gray and chain B is dark gray.

[0037] Figure 17 shows the location of N-terminally located or interacting sites of IspS that do not tolerate substitution. Mg²⁺ (spheres) and residues 1-50 are modeled based on a structural alignment with PDB 1N24.

[0038] Figure 18 shows the location of N-terminally located or interacting sites of IspS that do not tolerate substitution. Mg²⁺ (spheres) are placed based on a structural alignment with PDB 1N24.

[0039] Figure 19 shows the proposed substrate capture loop positions of IspS that do not tolerate substitution. Mg²⁺ (spheres) are placed based on a structural alignment with PDB 1N24.

[0040] Figure 20 shows the monomer view of wild type IspS showing the location of sites where variants demonstrate improved specific activity.

[0041] Figure 21 shows the location of buried sites that are in or interact with the N-terminus of IspS, where variants demonstrated increased specific activity.

[0042] Figure 22 shows the location of buried sites that are in or interact with the C-terminus of IspS, where variants demonstrated increased specific activity. Mg^{2+} (spheres) are placed based on a structural alignment with PDB 1N24

[0043] Figure 23 shows the dimer interface of IspS, with chain A in light gray and chain B in dark gray. Variants at position 247 demonstrated improved specific activity.

[0044] Figure 24 shows the location of N-terminal sites of IspS where variants demonstrated improved specific activity.

[0045] Figure 25 shows the surface loop positions of IspS where variants demonstrated improved specific activity. Mg^{2+} (spheres) are placed based on a structural alignment with PDB 1N24.

[0046] Figure 26 shows the surface positions of IspS where variants demonstrated improved specific activity. Mg^{2+} (spheres) are placed based on a structural alignment with PDB 1N24.

[0047] Figure 27 shows the proposed substrate capture loop positions of IspS where variants demonstrated improved specific activity. Mg^{2+} (spheres) are placed based on a structural alignment with PDB 1N24.

[0048] Figure 28 shows positions of IspS demonstrating improved activity. Mg^{2+} (spheres) are placed based on a structural alignment with PDB 1N24.

[0049] Figure 29 shows positions of IspS demonstrating improved activity. Mg^{2+} (spheres) are placed based on a structural alignment with PDB 1N24.

[0050] Figure 30 shows monomer view of IspS showing N-terminal and Surface Loop positions demonstrating improved growth. Mg^{2+} (spheres) are placed based on a structural alignment with PDB 1N24.

[0051] Figure 31 shows monomer view of IspS showing N-terminal Helix positions demonstrating improved growth.

[0052] Figure 32 shows monomer view of wild type IspS showing the location of sites from Table 15 (A). Mg^{2+} (spheres) are placed based on a structural alignment with PDB 1N24. B) Close-up view of sites in A.

[0053] Figure 33 shows monomer view of IspS with location of sites where variants demonstrate improved growth colored dark gray. Mg^{2+} (spheres) are placed based on a structural alignment with PDB 1N24.

[0054] Figure 34 shows view of IspS with location of sites where variants demonstrate improved growth shown as sticks. A and B are separated by 180°.

[0055] Figure 35 shows an alignment of N-terminally truncated *P. alba* IspS molecules (SEQ ID NOS:31, 22, 27, and 29 respectively).

[0056] Figure 36 shows a plasmid map of pDW207 harboring the *P. alba* IspS MAR variant.

[0057] Figure 37 shows a plasmid map of pDW208 harboring the *P. alba* IspS MRR variant.

[0058] Figure 38A-B show the alignment of N-terminally cleaved *P. alba* IspS sequences (SEQ ID NOS:31, 32, and 22 respectively).

[0059] Figure 39A-C: Figure 39A shows the correlation of K_{IDMAPP} vs. specific isoprene productivity for isoprene synthase variants. Figure 39B shows the correlation of K_M vs. specific isoprene productivity for isoprene synthase variants. Figure 39C shows the correlation of k_{cat} vs. specific isoprene productivity for isoprene synthase variants. All values are normalized to parent molecule performance.

DETAILED DESCRIPTION

[0060] The present invention provides for, *inter alia*, isolated polypeptides having isoprene synthase activity wherein the variant has one or more improved properties, such as 1) k_{cat} (2) K_M (3) K_i and (4) specific productivity. The invention also provides for methods and compositions comprising at least one isoprene synthase variant. The variant can include one or more amino acid residue substitution(s) from a parent isoprene synthase polypeptide, wherein the parent isoprene synthase may be a wild type or non-wild type sequence. The invention provides amino acid residue substitutions at particular positions within the polypeptide, wherein

the substitution may result in at least one improved property as compared to its parent sequence or a reference sequence. In particular, the present invention provides isoprene synthase variants, host cells and systems for increased isoprene production. Biosynthetically produced isoprene of the present invention finds use in the manufacture of rubber, polymers, and elastomers.

I. General Techniques

[0061] The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology (including recombinant techniques), microbiology, cell biology, biochemistry, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature, “*Molecular Cloning: A Laboratory Manual*”, second edition (Sambrook et al., 1989); “*Oligonucleotide Synthesis*” (M. J. Gait, ed., 1984); “*Animal Cell Culture*” (R. I. Freshney, ed., 1987); “*Methods in Enzymology*” (Academic Press, Inc.); “*Current Protocols in Molecular Biology*” (F. M. Ausubel et al., eds., 1987, and periodic updates); “*PCR: The Polymerase Chain Reaction*”, (Mullis et al., eds., 1994). Singleton et al., “*Dictionary of Microbiology and Molecular Biology*” 2nd ed., J. Wiley & Sons (New York, N.Y. 1994), Baltz et al., “*Manual of Industrial Microbiology and Biotechnology*” 3rd ed., (Washington, DC : ASM Press, 2010), and March, *Advanced Organic Chemistry Reactions, Mechanisms and Structure* 4th ed., John Wiley & Sons (New York, N.Y. 1992) provide one skilled in the art with a general guide to many of the terms used in the present application.

II. Definitions

[0062] Unless defined otherwise herein, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention pertains. Although any methods and materials similar or equivalent to those described herein find use in the practice of the present invention, the preferred methods and materials are described herein. Furthermore, the headings provided herein are not limitations of the various aspects or embodiments of the invention, which can be had by reference to the specification as a whole. Accordingly, the terms defined immediately below are more fully described by reference to the specification as a whole. Nonetheless, in order to facilitate understanding of the invention, a number of terms are defined below.

[0063] “X” refers to any amino acid residue. However, when in the context of an amino acid substitution (e.g. “X003C”), it is to be understood that “X” refers to an amino acid residue other than the amino acid residue resulting from the substitution (e.g., X is an amino acid residue other than C).

[0064] As used herein, the term “polypeptides” includes polypeptides, proteins, peptides, fragments of polypeptides, and fusion polypeptides.

[0065] As used herein, an “isolated polypeptide” is not part of a library of polypeptides, such as a library of 2, 5, 10, 20, 50 or more different polypeptides and is separated from at least one component with which it occurs in nature. An isolated polypeptide can be obtained, for example, by expression of a recombinant nucleic acid encoding the polypeptide.

[0066] By “heterologous polypeptide” is meant a polypeptide encoded by a nucleic acid sequence derived from a different organism, species, or strain than the host cell. In some aspects, a heterologous polypeptide is not identical to a wild-type polypeptide that is found in the same host cell in nature.

[0067] As used herein, a “nucleic acid” refers to two or more deoxyribonucleotides and/or ribonucleotides covalently joined together in either single or double-stranded form. It is to be understood that mutations, including single nucleotide mutations, can occur within a nucleic acid as defined herein.

[0068] By “recombinant nucleic acid” is meant a nucleic acid of interest that is free of one or more nucleic acids (e.g., genes) which, in the genome occurring in nature of the organism from which the nucleic acid of interest is derived, flank the nucleic acid of interest. The term therefore includes, for example, a recombinant DNA which is incorporated into a vector, into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate molecule (e.g., a cDNA, a genomic DNA fragment, or a cDNA fragment produced by PCR or restriction endonuclease digestion) independent of other sequences. A recombinant nucleic acid may be obtained using molecular biology techniques that are known in the art, or part or all of a recombinant nucleic acid may be chemically synthesized.

[0069] By “heterologous nucleic acid” is meant a nucleic acid sequence derived from a different organism, species or strain than the host cell. In some aspects, the heterologous nucleic acid is not identical to a wild-type nucleic acid that is found in the same host cell in nature. In some aspects, a heterologous nucleic acid is not identical to a wild-type nucleic acid that is found in the same host cell in nature.

[0070] An “endogenous nucleic acid” is a nucleic acid whose nucleic acid sequence is naturally found in the host cell. In some embodiments, an endogenous nucleic acid is identical to a wild-type nucleic acid that is found in the host cell in nature. In some embodiments, one or more copies of endogenous nucleic acids are introduced into a host cell.

[0071] A nucleic acid or protein of the invention may be in isolated or purified form. As used herein, “isolated,” with respect to nucleic acid or protein, means separated from other components, such as, but not limited to a cell or cell culture. It is preferably in a homogeneous state although it can be in either a dry or aqueous solution. Purity and homogeneity are typically determined using analytical chemistry techniques, such as polyacrylamide gel electrophoresis or high performance liquid chromatography. A protein or nucleic acid that is the predominant species present in a preparation is substantially purified. The term “purified” denotes that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. Particularly, “purified” means that when isolated, the isolate contains at least 90%, at least 95%, at least 98%, or more preferably at least 99% of nucleic acid or protein by weight of the isolate.

[0072] Purified polypeptides may be obtained by a number of methods including, for example, laboratory synthesis, chromatography, preparative electrophoresis, gel electrophoresis, centrifugation, precipitation, affinity purification, etc. (see, generally, R Scopes, Protein Purification, Springer-Verlag, N.Y. (1982), Deutscher, Methods in Enzymology Vol. 182: Guide to Protein Purification, Academic Press, Inc. N.Y. (1990)).

[0073] “Polypeptides” includes polypeptides, proteins, peptides, fragments of polypeptides, and fusion polypeptides. It is also understood that a polypeptide may be coded for by more than one nucleotide sequence due to the degeneracy of the genetic code.

[0074] A “heterologous polypeptide” is a polypeptide encoded by a heterologous nucleic acid. In some embodiments, the sequence is not identical to that of another polypeptide encoded

by a nucleic acid naturally found in the same host cell. Examples of heterologous proteins include enzymes such as isoprene synthases. In some embodiments, the genes encoding the proteins are naturally occurring genes, while in other embodiments mutated and/or synthetic genes are used.

[0075] An “endogenous polypeptide” is a polypeptide whose amino acid sequence is naturally found in the host cell. In some embodiments, an endogenous polypeptide is identical to a wild-type polypeptide that is found in the host cell in nature.

[0076] As used herein, an “expression control sequence” means a nucleic acid sequence that directs transcription of a nucleic acid of interest. An expression control sequence can be a promoter, such as a constitutive or an inducible promoter, or an enhancer. An expression control sequence can be “native” or heterologous. A native expression control sequence is derived from the same organism, species, or strain as the gene being expressed. A heterologous expression control sequence is derived from a different organism, species, or strain as the gene being expressed. An “inducible promoter” is a promoter that is active under environmental or developmental regulation.

[0077] By “operably linked” is meant a functional linkage between a nucleic acid expression control sequence (such as a promoter) and a second nucleic acid sequence, wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence.

[0078] The term “isoprene” refers to 2-methyl-1,3-butadiene (CAS# 78-79-5). It can be the direct and final volatile C5 hydrocarbon product from the elimination of pyrophosphate from DMAPP. It may not involve the linking or polymerization of IPP molecules to DMAPP molecules. The term “isoprene” is not generally intended to be limited to its method of production unless indicated otherwise herein.

[0079] As used herein, the terms “isoprene synthase,” “isoprene synthase variant”, and “IspS,” refer to enzymes that catalyze the elimination of pyrophosphate from dimethylallyl diphosphate (DMAPP) to form isoprene. An “isoprene synthase” may be a wild type sequence or an isoprene synthase variant.

[0080] An “isoprene synthase variant” indicates a non-wild type polypeptide having isoprene synthase activity. One skilled in the art can measure isoprene synthase activity using known methods. See, for example, by GC-MS (see, e.g., WO 2009/132220, Example 3) or Silver et al., *J. Biol. Chem.* 270:13010-13016, (1995). Variants may have substitutions, additions, deletions, and/or truncations from a wild type isoprene synthase sequence. Variants may have substitutions, additions, deletions, and/or truncations from a non-wild type isoprene synthase sequence. The variants described herein contain at least one amino acid residue substitution from a parent isoprene synthase polypeptide. In some embodiments, the parent isoprene synthase polypeptide is a wild type sequence. In some embodiments, the parent isoprene synthase polypeptide is a non-wild type sequence. In various embodiments, the variant will have at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 100%, at least about 110%, at least about 120%, at least about 130%, at least about 140%, at least about 150%, at least about 160%, at least about 170%, at least about 180%, at least about 190%, at least about 200% of the activity of a wild type isoprene synthase. In various embodiments, the variant will have at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 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%, at least about 99% sequence identity to a wild type isoprene synthase. In various embodiments, the number of differing amino acid residues between the variant and the wild type may be one or more, e.g. 1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50, or more amino acid residues. Wild type isoprene synthases can include any isoprene synthases from plants, for example, kudzu isoprene synthases, poplar isoprene synthases, English oak isoprene synthases, and willow isoprene synthases.

[0081] As used herein, the term “naturally-occurring” refers to anything (e.g., proteins, amino acids, or nucleic acid sequences) that is found in nature (e.g., has not been manipulated by means of recombinant or chemical methods). As used herein, the term “non-naturally occurring” refers to anything that is not found in nature (e.g., recombinantly produced or chemically synthesized proteins, amino acids, or nucleic acid sequences produced in the laboratory).

[0082] As used herein, an amino acid residue of an amino acid sequence of interest that “corresponds to” or is “corresponding to” or in “correspondence with” an amino acid residue of a reference amino acid sequence indicates that the amino acid residue of the sequence of interest is at a location homologous or equivalent to an enumerated residue in the reference amino acid sequence. One skilled in the art can determine whether a particular amino acid residue position in a polypeptide corresponds to that of a homologous reference sequence. For example, the sequence of an isoprene synthase polypeptide may be aligned with that of a reference sequence (e.g. *P. alba* MEA isoprene synthase, SEQ ID NO:1) using known techniques (e.g., basic local alignment search tool (BLAST), ClustalW2, Structure based sequences alignment program (STRAP), or the like). In addition, crystal structure coordinates of a reference sequence may be used as an aid in determining a homologous polypeptide residue’s three dimensional structure (see, for example, WO 2010/124146 or US 2011/0076743). In another aspect, equivalent residues may be identified by determining homology at the level of tertiary structure. Using such methods, the amino acid residues of an isoprene synthase polypeptide or isoprene synthase variant may be numbered according to the corresponding amino acid residue position numbering of the reference sequence. For example, the amino acid sequence of SEQ ID NO:1 may be used for determining amino acid residue position numbering of each amino acid residue of an isoprene synthase variant of interest.

[0083] The term “identical” in the context of two nucleic acids or polypeptide sequences refers to the residues in the two sequences that are the same when aligned for maximum correspondence, as measured using one of the following sequence comparison or analysis algorithms.

[0084] As used herein, “homology” refers to sequence similarity or identity, with identity being preferred. Homology may be determined using standard techniques known in the art (see, e.g., Smith and Waterman, *Adv. Appl. Math.* 2:482 (1981); Needleman and Wunsch, *J. Mol. Biol.* 48:443 (1970); Pearson and Lipman, *Proc. Natl. Acad. Sci. USA* 85:2444 (1988); software programs such as GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package (Genetics Computer Group, Madison, WI); and Devereux et al., *Nucl. Acid Res.* 12:387-395 (1984)). One example of a useful algorithm is PILEUP. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pair-wise alignments. It can also plot a tree showing the clustering relationships used to create the alignment. PILEUP

uses a simplification of the progressive alignment method of Feng and Doolittle (see Feng and Doolittle, *J. Mol. Evol.* 35:351-360 (1987)). The method is similar to that described by Higgins and Sharp (see Higgins and Sharp, *CABIOS* 5:151-153 (1989)). Useful PILEUP parameters including a default gap weight of 3.00, a default gap length weight of 0.10, and weighted end gaps. Another example of a useful algorithm is the BLAST algorithm, described by Altschul et al., (see Altschul et al., *J. Mol. Biol.* 215:403-410 (1990); and Karlin et al. *Proc. Natl. Acad. Sci. USA* 90:5873-5787 (1993)). A particularly useful BLAST program is the WU-BLAST-2 program (see Altschul et al., *Meth. Enzymol.* 266:460-480 (1996)). WU-BLAST-2 uses several search parameters, most of which are set to the default values. The adjustable parameters are set with the following values: overlap span =1, overlap fraction = 0.125, word threshold (T) = 11. The HSP S and HSP S2 parameters are dynamic values and are established by the program itself depending upon the composition of the particular sequence and composition of the particular database against which the sequence of interest is being searched. However, the values may be adjusted to increase sensitivity.

[0085] The percent sequence identity between a reference sequence and a test sequence of interest may be readily determined by one skilled in the art. The percent identity shared by polynucleotide or polypeptide sequences is determined by direct comparison of the sequence information between the molecules by aligning the sequences and determining the identity by methods known in the art. An example of an algorithm that is suitable for determining sequence similarity is the BLAST algorithm, (see Altschul, et al., *J. Mol. Biol.*, 215:403-410 (1990)). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence that either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. These initial neighborhood word hits act as starting points to find longer HSPs containing them. The word hits are expanded in both directions along each of the two sequences being compared for as far as the cumulative alignment score can be increased. Extension of the word hits is stopped when: the cumulative alignment score falls off by the quantity X from a maximum achieved value; the cumulative score goes to zero or below; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLAST program uses as defaults a wordlength (W) of 11, the BLOSUM62 scoring matrix (see Henikoff and Henikoff, *Proc. Natl. Acad. Sci. USA*

89:10915 (1992)) alignments (B) of 50, expectation (E) of 10, M⁵, N⁻⁴, and a comparison of both strands.

[0086] The BLAST algorithm then performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin and Altschul, *supra*). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a isoprene synthase nucleic acid of this invention if the smallest sum probability in a comparison of the test nucleic acid to a isoprene synthase nucleic acid is less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001. Where the test nucleic acid encodes an isoprene synthase polypeptide, it is considered similar to a specified isoprene synthase nucleic acid if the comparison results in a smallest sum probability of less than about 0.5, and more preferably less than about 0.2.

[0087] Percent “identical” or “identity” in the context of two or more nucleic acid or polypeptide sequences refers to two or more sequences that are the same or have a specified percentage of nucleic acid residues or amino acid residues, respectively, that are the same, when compared and aligned for maximum similarity, as determined using a sequence comparison algorithm or by visual inspection. “Percent sequence identity” or “% identity” or “% sequence identity” or “% amino acid sequence identity” of a subject amino acid sequence to a reference amino acid sequence means that the subject amino acid sequence is identical (i.e., on an amino acid-by-amino acid basis) by a specified percentage to the reference amino acid sequence over a comparison length when the sequences are optimally aligned. Thus, 80% amino acid sequence identity or 80% identity with respect to two amino acid sequences means that 80% of the amino acid residues in two optimally aligned amino acid sequences are identical.

[0088] “Percent sequence identity” or “% identity” or “% sequence identity” of a subject nucleic acid sequence to a reference nucleic acid sequence means that the subject nucleic acid sequence is identical (i.e., on a nucleotide-by-nucleotide basis for a polynucleotide sequence) by a specified percentage to the reference sequence over a comparison length when the sequences are optimally aligned. Thus, 80% nucleotide sequence identity or 80% identity with respect to two nucleic acid sequences means that 80% of the nucleotide residues in two optimally aligned nucleic acid sequences are identical.

[0089] The “percent sequence identity” or “% sequence identity” or “% identity” of a subject sequence to a reference sequence can be calculated by optimally aligning the two sequences and comparing the two optimally aligned sequences over the comparison length. The number of positions in the optimal alignment at which identical residues occur in both sequences is determined, thereby providing the number of matched positions, and the number of matched positions is then divided by the total number of positions of the comparison length (which, unless otherwise specified, is the length of the reference sequence). The resulting number is multiplied by 100 to yield the percent sequence identity of the subject sequence to the reference sequence.

[0090] “Optimal alignment” or “optimally aligned” refers to the alignment of two (or more) sequences giving the highest percent identity score. For example, optimal alignment of two polypeptide sequences can be achieved by manually aligning the sequences such that the maximum number of identical amino acid residues in each sequence are aligned together or by using software programs or procedures described herein or known in the art. Optimal alignment of two nucleic acid sequences can be achieved by manually aligning the sequences such that the maximum number of identical nucleotide residues in each sequence are aligned together or by using software programs or procedures described herein or known in the art.

[0091] Two sequences (e.g., polypeptide sequences) may be deemed “optimally aligned” when they are aligned using defined parameters, such as a defined amino acid substitution matrix, gap existence penalty (also termed gap open penalty), and gap extension penalty, so as to achieve the highest similarity score possible for that pair of sequences. The BLOSUM62 scoring matrix (see Henikoff and Henikoff, *supra*) is often used as a default scoring substitution matrix in polypeptide sequence alignment algorithms (e.g., BLASTP). The gap existence penalty is imposed for the introduction of a single amino acid gap in one of the aligned sequences, and the gap extension penalty is imposed for each residue position in the gap. Exemplary alignment parameters employed are: BLOSUM62 scoring matrix, gap existence penalty=11, and gap extension penalty=1. The alignment score is defined by the amino acid positions of each sequence at which the alignment begins and ends (e.g., the alignment window), and optionally by the insertion of a gap or multiple gaps into one or both sequences, so as to achieve the highest possible similarity score.

[0092] Optimal alignment between two or more sequences can be determined manually by visual inspection or by using a computer, such as, but not limited to e.g., the BLASTP program for amino acid sequences and the BLASTN program for nucleic acid sequences (see, e.g., Altschul et al., *Nucleic Acids Res.* 25(17):3389-3402 (1997); see also the National Center for Biotechnology Information (NCBI) website) or CLUSTALW program.

[0093] A polypeptide of interest may be said to be “substantially identical” to a reference polypeptide if the polypeptide of interest comprises an amino acid sequence having at least about 60%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 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%, at least about 99%, or at least about 99.5% sequence identity to the amino acid sequence of the reference polypeptide. The percent identity between two such polypeptides can be determined manually by inspection of the two optimally aligned polypeptide sequences or by using software programs or algorithms (e.g., BLAST, ALIGN, CLUSTAL) using standard parameters. One indication that two polypeptides are substantially identical is that the first polypeptide is immunologically cross-reactive with the second polypeptide. Typically, polypeptides that differ by conservative amino acid substitutions are immunologically cross-reactive. Thus, a polypeptide is substantially identical to a second polypeptide, e.g., where the two peptides differ only by a conservative amino acid substitution or one or more conservative amino acid substitutions.

[0094] A nucleic acid of interest may be said to be “substantially identical” to a reference nucleic acid if the nucleic acid of interest comprises a nucleotide sequence having at least about 60%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 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%, at least about 99%, or at least about 99.5% sequence identity to the nucleotide sequence of the reference nucleic acid. The percent identity between two such nucleic acids can be determined manually by inspection of the two optimally aligned nucleic acid sequences or by using software programs or algorithms (e.g., BLAST, ALIGN, CLUSTAL) using standard parameters. One indication that two nucleic acid sequences are substantially identical is that the two nucleic acid molecules hybridize to each other under stringent conditions (e.g., within a range of medium to high stringency).

[0095] As used herein, the term “mass yield” refers to the mass of the product produced by the recombinant (*e.g.*, bacterial) cells divided by the mass of the glucose consumed by the recombinant cells multiplied by 100, or expressed as a percentage.

[0096] By “specific productivity,” it is meant the mass of the product produced by the recombinant (*e.g.*, bacterial) cell divided by the product of the time for production, the cell density, and the volume of the culture.

[0097] By “titer,” it is meant the mass of the product produced by the recombinant (*e.g.*, bacterial) cells divided by the volume of the culture.

[0098] As used herein, the term “cell productivity index (CPI)” refers to the mass of the product produced by the recombinant (*e.g.*, bacterial) cells divided by the mass of the recombinant cells produced in the culture.

[0100] Unless defined otherwise herein, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention pertains.

[0101] As used herein, the singular terms “a,” “an,” and “the” include the plural reference unless the context clearly indicates otherwise.

[0102] Unless otherwise indicated, nucleic acids are written left to right in 5' to 3' orientation; amino acid sequences are written left to right in amino to carboxy orientation, respectively.

[0103] Reference to “about” a value or parameter herein also includes (and describes) embodiments that are directed to that value or parameter per se.

[0104] It is understood that all aspects and embodiments of the invention described herein include “comprising,” “consisting,” and “consisting essentially of” aspects and embodiments. It is to be understood that methods or compositions “consisting essentially of” the recited elements include only the specified steps or materials and those that do not materially affect the basic and novel characteristics of those methods and compositions.

[0105] It is to be understood that this invention is not limited to the particular methodology, protocols, and reagents described, as these may vary, depending upon the context they are used by those of skill in the art.

[0106] It is intended that every maximum numerical limitation given throughout this specification includes every lower numerical limitation, as if such lower numerical limitations were expressly written herein. Every minimum numerical limitation given throughout this specification will include every higher numerical limitation, as if such higher numerical limitations were expressly written herein. Every numerical range given throughout this specification will include every narrower numerical range that falls within such broader numerical range, as if such narrower numerical ranges were all expressly written herein.

Isoprene Synthase Variants with Improved Kinetic Properties

[0107] Isoprene (2-methyl-1,3-butadiene) is the monomer of natural rubber and also the critical starting material for a variety of synthetic polymers, most notably synthetic rubbers. Isoprene monomer is employed in the manufacture of polyisoprene and various copolymers (with isobutylene, butadiene, styrene, or other monomers). Building a host cell strain (prokaryotic or eukaryotic) capable of producing commercially viable levels of isoprene requires optimization of the biological pathway, for example isoprene synthase and/or the DXP pathway and/or the MVA pathway. A key enzyme in the pathway is isoprene synthase (IspS), which converts the precursor DMAPP to isoprene. Isoprene synthases (IspS) that have been identified include those from plants such as poplar, English oak and kudzu vine. Some of the plant IspS enzymes identified have been partially characterized in part by expression in *E. coli* and some of the kinetic parameters of these enzymes have been determined *in vitro* with purified protein. However, the kinetic parameters (K_m , rate, *etc.*) of the native IspS enzymes are insufficient for commercial production of isoprene in a biological host. Thus, one problem to be solved is the provision of isoprene synthase variants (e.g. with substitutions at specific residues) which have improved properties (e.g., kinetic parameters) such that a greater amount of isoprene can be biologically produced.

[0108] As described in more detail herein, the inventors have solved this problem by constructing polypeptides having isoprene synthase activity that are engineered with one or more substitutions (e.g., amino acid substitutions) such that it has one or more improved properties,

such as 1) kcat (2) KM (3) Ki and (4) specific productivity. Host cells can be used to express such polypeptides for production of isoprene at a level that is commercially relevant.

[0109] As indicated above, optimizing kinetic properties of interest include, but are not limited to: 1) kcat (2) KM (3) Ki and (4) specific productivity. Without wishing to be bound by theory, in one embodiment, isoprene synthase variants that display increased kcat values (e.g., greater than 1 as compared to a reference isoprene synthase, such as *P. alba* MEA isoprene synthase, SEQ ID NO:1) can catalyze the conversion of DMAPP to isoprene more efficiently. In another embodiment, isoprene synthase variants that display decreased KM values (e.g., lower than 1 as compared to a reference isoprene synthase, such as *P. alba* MEA isoprene synthase, SEQ ID NO:1) can maintain decreased concentrations of DMAPP. In another embodiment, isoprene synthase variants that display increased K_{iDMAPP} values (e.g., greater than 1 as compared to a reference isoprene synthase, such as *P. alba* MEA isoprene synthase, SEQ ID NO:1) can catalyze the conversion of DMAPP to isoprene more efficiently. In another embodiment, isoprene synthase variants have increased specific productivity (e.g., greater than 1 as compared to a reference isoprene synthase, such as *P. alba* isoprene synthase). The sequence of *P. alba* MEA isoprene synthase is as follows:

MEARRSANYEPNSWDYDYLLSSDTDESIEVYKDKAKKLEAEVRREINNEKAEFLTLEL
 IDNVQRLGLGYRFESDIRGALDRFVSSGGFDAVTKTSLHGTALSFRLLRQHGFVVSQEAF
 SGFKDQNGNFLENLKEDIKAILSLYEASFLALEGENILDEAKVFAISHLKELSEEKIGKEL
 AEQVNHAELEPLHRRTQRLEAVWSIEAYRKKEDANQVLELAILDYNMIQSVYQRDLR
 ETSRWRRVGLATKLHFARDRLIESFYWAVGVAFEPQYSDCRNSVAKMFSFVTIIDDY
 DVYGTLDLELFTDAVERWDVNAINDLDPYMKLCFLALYNTINEIAYDNLKDKGENILP
 YLTKAWADLCNAFLQEAKWLYNKSTPTFDDYFGNAWKSSSGPLQLVFAYFAVVQNIK
 KEEIENLQKYHDTISRPSHIFRLCNDLASASAEIARGETANSVSCYMRKGISEELATESV
 MNLIDETWKKMNKEKLGGS�FAKPFVETAINLARQSHCTYHNGDAHTSPDELTRKRVL
 SVITEPILPFER (SEQ ID NO:1).

[0110] These properties and parameters can be assessed by the conversion of DMAPP to isoprene *in vitro* with purified or partially purified isoprene synthase or *in vivo* in the context of a host organism such as *E. coli* expressing the DXP pathway, the MVA pathway, or both. It is contemplated that enzymes having various degrees of stability, solubility, activity, and/or

expression level in one or more of test conditions will find use in the present invention for the production of isoprene in a diversity of hosts.

[0111] The invention features compositions and methods for the production of increased amounts of isoprene. In particular, these compositions and methods may increase the rate of isoprene production and the total amount of isoprene that is produced. The biosynthetic processes for isoprene production described herein are a desirable alternative to using natural rubber. As discussed further below, the amount of isoprene produced by cells can be greatly increased by introducing a heterologous nucleic acid encoding an isoprene synthase (IspS) variant into the cells.

[0112] As exemplified in the Examples section, one of skill in the art can determine *in vitro* kinetic constants (k_{cat} , K_M and K_{iDMAPP}) for the ability of the variant isoprene synthases to catalyze the conversion of DMAPP to isoprene by fitting data to the following rate equation:

$$\frac{\text{rate}}{[\text{isoprene synthase}]} = \frac{k_{cat} \cdot [\text{DMAPP}]}{K_M + [\text{DMAPP}] \left(1 + \frac{[\text{DMAPP}]}{K_{iDMAPP}}\right)}$$

[0113] In one aspect, polypeptides having isoprene synthase activity or isoprene synthase variants with k_{cat} values greater than about 1 with respect to a reference sequence (such as SEQ ID NO:1) are selected. In other aspects, isoprene synthase variants with k_{cat} values greater than about 1.05, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3.0, 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 4.0, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 5.0, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9 or 6.0 are selected. In other aspects, isoprene synthase variants with k_{cat} values greater than about 6.5, 7.0, 7.5, 8.0, 8.5, 9.0, 9.5, or 10 are selected. In some embodiments, the variants have one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1: X323F, X118E, X36W, X22K, X228Y, X448L, X488F, X467H, X443S, X331P, X453I, X71K, X71L, X448I, X71M, X392Y, X448V, X282H, X383Y, X323Y, X511Y, X448E, X376M, X488L, X120E, X461A, X414I, X282W, X071K, X493E, X392S, X448Q, X282Y, X537N, X447Y, X240C, X443Q, X538R, and X510C. In some embodiments, the polypeptide has improved K_{cat} compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the variants have one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1: D323F, A118E, K36W, S22K, M228Y, A448L, E488F, E467H, A443S, C331P,

A453I, R71K, R71L, A448I, R71M, W392Y, A448V, S282H, T383Y, D323Y, H511Y, A448E, L376M, E488L, S120E, R461A, K414I, S282W, R071K, S493E, W392S, A448Q, S282Y, E537N, I447Y, T240C, A443Q, P538R, and S510C.

[0114] In some embodiments, the variants have one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X389D, X383H, X36E, X436Y, X447V, X392F, X161R, X99D, X99E, X161C, X414S, X537T, X393V, X443G, X510V, X36N, X392A, X161A, X254R, X472C, X161Q, X36Y, X537C, X41Y, X161M, X43L, X120A, X087M, X36S, X36H, X348Y, X467W, X288T, X447T, X025N. In some embodiments, the polypeptide has improved K_{cat} compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the variants have one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1: G389D, T383H, K36E, L436Y, I447V, W392F, K161R, G99D, G99E, K161C, K414S, E537T, K393V, A443G, S510V, K36N, W392A, K161A, H254R, E472C, K161Q, K36Y, E537C, E41Y, K161M, R43L, S120A, G087M, K36S, K36H, K348Y, E467W, S288T, I447T, and D025N.

[0115] In some embodiments, the variants have one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X268I, X488M, X480I, X240V, X460A, X071L, X36Q, X118P, X36T, X282I, X409I, X071I, X409T, X436F, X348F, X118Q, X342Y, X74Q, X41P, X437Y, X392T, X89D, X41M, X510E, X36P, X393I, X71I, X381M, X374Y, X58Y, X526Q, X543F, X240M, X111S, X463T, X120M, and X135G. In some embodiments, the polypeptide has improved K_{cat} compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the variants have one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1: V268I, E488M, E480I, T240V, M460A, R071L, K36Q, A118P, K36T, S282I, V409I, R071I, V409T, L436F, K348F, A118Q, I342Y, S74Q, E41P, C437Y, W392T, F89D, E41M, S510E, K36P, K393I, R71I, T381M, K374Y, E58Y, L526Q, E543F, T240M, G111S, K463T, S120M, and E135G.

[0116] In another aspect, polypeptides having isoprene synthase activity or isoprene synthase variants with K_M values less than about 1 with respect to a reference sequence (such as SEQ ID NO:1) are selected. In other aspects, isoprene synthase variants with K_M values less than about 0.95, 0.9, 0.8., 0.7, 0.5, 0.4, 0.3, 0.2 or 0.1 are selected. In some embodiments, the

variants have one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1: X22K, X348F, X392V, X392F, X488C, X22R, X71V, X443R, X234R, X453V, X437Y, X392C, X463F, X538K, X393L, X254C, X436Y, X21R, X437L, X444D, X374Y, X363L, X447V, X444E, X71I, X504F, X488W, X71H, X36P, X381I, X436F, X460A, X443G, X288A, X392T, X537I, X374Y, X242G, X437M, X436I, X376I, X288Y, and X392M. In some embodiments, the polypeptide has improved K_M compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the amino acid substitution is selected from the group consisting of: S22K, K348F, W392V, W392F, E488C, S22R, R71V, A443R, Q234R, A453V, C437Y, W392C, K463F, P538K, K393L, H254C, L436Y, S21R, C437L, S444D, K374Y, A363L, I447V, S444E, R71I, I504F, E488W, R71H, K36P, T381I, L436F, M460A, A443G, S288A, W392T, E537I, K374Y, R242G, C437M, L436I, L376I, S288Y, and W392M.

[0117] In some embodiments, the variants have one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1: X342Y, X414W, X461A, X443S, X392A, X463T, X488T, X526Q, X502F, X502M, X488M, X36Y, X414R, X071I, X36W, X502L, X481Y, X472R, X36H, X36T, X415Y, X415H, X58Y, X381M, X481V, X89E, X480I, X36Q, X89D, X254R, X161N, X071L, X424P, X415V, X22K, X58L, X392S, X268I, X392I, X526E, X537V, and X36D. In some embodiments, the polypeptide has improved K_M compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the amino acid substitution is selected from the group consisting of: I342Y, K414W, R461A, A443S, W392A, K463T, E488T, L526Q, T502F, T502M, E488M, K36Y, K414R, R071I, K36W, T502L, T481Y, E472R, K36H, K36T, E415Y, E415H, E58Y, T381M, T481V, F89E, E480I, K36Q, F89D, H254R, K161N, R071L, H424P, E415V, S22K, E58L, W392S, V268I, W392I, L526E, E537V, and K36D.

[0118] In some embodiments, the variants have one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X393I, X41M, X025N, X36N, X118Q, X389E, X282W, X348Y, X383L, X510V, X323Y, X323F, X383H, X41P, X288T, X36S, X453I, X447T, X472I, X376M, X161E, X488L, X381L, X071K, X409T, X392Y, X135G, X542L, X228Y, X118P, X543F, and X36E. In some embodiments, the polypeptide has improved K_M compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the amino acid

substitution is selected from the group consisting of: K393I, E41M, D025N, K36N, A118Q, G389E, S282W, K348Y, T383L, S510V, D323Y, D323F, T383H, E41P, S288T, K36S, A453I, I447T, E472I, L376M, K161E, E488L, T381L, R071K, V409T, W392Y, E135G, F542L, M228Y, A118P, E543F, and K36E.

[0119] In one aspect, polypeptides having isoprene synthase activity or isoprene synthase variants with K_{IDMAPP} values greater than about 1 with respect to a reference sequence (such as SEQ ID NO:1) are selected. In other aspects, isoprene synthase variants with K_{IDMAPP} values greater than about 1.05, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3.0, 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 4.0, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 5.0, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9 or 6.0 are selected. In other aspects, isoprene synthase variants with K_{IDMAPP} values greater than about 6.5, 7.0, 7.5, 8.0, 8.5, 9.0, 9.5, or 10 are selected. In some embodiments, the variants have one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1: X472R, X389E, X242G, X376I, X161N, X288C, X240V, X481Y, X463F, and X393L. In some embodiments, the variants have improved K_{IDMAPP} compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the amino acid substitution is selected from the group consisting of: E472R, G389E, R242G, L376I, K161N, S288C, T240V, T481Y, K463F, and K393L.

[0120] In another aspect of the invention, the polypeptides have isoprene synthase activity and have one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X120Q, X58L, X240M, and X453V. In some embodiments, the polypeptide has improved K_{IDMAPP} compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the amino acid substitution is selected from the group consisting of: S120Q, E58L, T240M, and A453V.

[0121] In another aspect of the invention, the polypeptides have isoprene synthase activity and have one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X543F, X74Q, X414R, X415H, and X415V. In some embodiments, the polypeptide has improved K_{IDMAPP} compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the

amino acid substitution is selected from the group consisting of: E543F, S74Q, K414R, E415H, and E415V.

[0122] In another aspect, polypeptides having isoprene synthase activity or isoprene synthase variants with specific productivity greater than about 1 with respect to a reference sequence (such as SEQ ID NO:1) are selected. In other aspects, isoprene synthase variants with specific productivity greater than about 1.05, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3.0, 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 4.0, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 5.0, 5.1, 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, 5.8, 5.9 or 6.0 are selected. In other aspects, isoprene synthase variants with specific productivity greater than about 6.5, 7.0, 7.5, 8.0, 8.5, 9.0, 9.5, or 10 are selected. In some embodiments, the variants have one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1: X510E, X161M, X120Q, X120E, X161R, X120A, X493E, X448E, X240V, X240M, and X537T. In some embodiments, the polypeptides have improved specific productivity compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the amino acid substitution is selected from the group consisting of: S510E, K161M, S120Q, S120E, K161R, S120A, S493E, A448E, T240V, T240M, and E537T.

[0123] In other aspects, polypeptides have isoprene synthase activity and have one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X161N, X447Y, X502M, X537N, X348Y, X543F, X331P, X537C, X120M, X288C, and X161Q. In some embodiments, the polypeptide has improved specific productivity compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the amino acid substitution is selected from the group consisting of: K161N, I447Y, T502M, E537N, K348Y, E543F, C331P, E537C, S120M, S288C, and K161Q.

[0124] In other aspects, polypeptides have isoprene synthase activity and have one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X161A, X481V, X087M, X467W, X448V, X502L, and X376M. In some embodiments, the polypeptide has improved specific productivity compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s). In some embodiments, the amino acid substitution is selected from the group consisting of: K161A, T481V, G087M, E467W, A448V, T502L, and L376M.

[0125] In some aspects of the invention, polypeptides having isoprene synthase activity or the isoprene synthase variants have increased k_{cat} or decreased K_M or increased K_{iDMAPP} or increased specific activity. In another aspect of the invention, the variant has any two of these four properties. In another aspect of the invention, the variant has any three of these four properties. In another aspect of the invention, the variant has all four of these properties.

[0126] The invention also contemplates methods for screening for isoprene synthase variants, comprising: (a) contacting a host cell with a medium comprising about 10 μM to about 70 μM IPTG, and about 5 mM to about 20 mM mevalonic acid (MVA), wherein the host cell comprises a nucleic acid encoding an isoprene synthase variant in operable combination with a promoter; and (b) measuring the growth rate of the host cell. The variant growth rate may be compared to that of a reference isoprene synthase (e.g. a parent isoprene synthase, a wild-type isoprene synthase, or MEA *P. alba* isoprene synthase). The methods may be used to screen for variants having a particular property of interest, for example, one or more of the properties described herein. In some embodiments, an increased growth rate indicates an isoprene synthase variant with an increased ability to convert DMAPP to isoprene within the host cell synthase. Growth rates may be analyzed, for example, according to methods known in the art, or as exemplified in the Examples below. In some embodiments, the method further comprises determining a growth index for the variant. In some embodiments, the method further comprises determining a performance index for the variant. Growth rate of the cells in exponential phase and/or final density of the cells may be taken into consideration as factors when selecting variants. As exemplified below, for the variants shown in the examples, the growth rate of the cells in exponential phase was a consideration. In addition, growth rate and final density was also taken into consideration when selecting for variants described herein.

[0127] In some embodiments, the IPTG is present in the medium at a concentration from about 10 μM to about 60 μM . In some embodiments, the IPTG is present in the medium at a concentration from about 20 μM to about 60 μM . In some embodiments, the IPTG is present in the medium at a concentration from about 40 μM to about 60 μM . In some embodiments, the IPTG is present in the medium at a concentration of about 50 μM . In some embodiments, the MVA is present in the medium at a concentration of about 5 mM to about 20 mM. In some embodiments, the MVA is present in the medium at a concentration of about 7 mM to about 15 mM. In some embodiments, the MVA is present in the medium at a concentration of about 8

mM to about 12 mM. In some embodiments, the MVA is present in the medium at a concentration of about 10 mM. In some embodiments, the host cell is MD09-170.

Isoprene Synthase Parent Sequences

[0128] Isoprene synthase variants may be generated from a parent isoprene synthase, wherein the parent isoprene synthase may be an isoprene synthase as described herein, including wild type and non-wild type isoprene synthases. Exemplary parent isoprene synthase nucleic acids include nucleic acids that encode a polypeptide, fragment of a polypeptide, peptide, or fusion polypeptide that has at least one activity of an isoprene synthase polypeptide. Exemplary parent isoprene synthase polypeptides and nucleic acids include naturally-occurring polypeptides and nucleic acids from any of the source organisms described herein as well as variant polypeptides and nucleic acids derived from any of the source organisms described herein.

[0129] In some embodiments, the parent isoprene synthase is from the family Fabaceae, the family Salicaceae, or the family Fagaceae. In some embodiments, the parent isoprene synthase polypeptide or nucleic acid is a naturally-occurring polypeptide or nucleic acid from *Pueraria montana* (kudzu) (Sharkey et al., Plant Physiology 137: 700-712, (2005)), poplar (such as *Populus alba x tremula* CAC35696, Miller et al., Planta 213: 483-487, (2001)) or *Populus alba*, aspen (such as *Populus tremuloides*) Silver et al., JBC 270(22): 13010-1316, (1995)), or English Oak (*Quercus robur*) (Zimmer et al., WO 98/02550). Suitable parent isoprene synthases include, but are not limited to, those identified by GenBank Accession Nos. AY341431, AY316691, AB198180, AJ294819.1, EU693027.1, EF638224.1, AM410988.1, EF147555.1, AY279379, AJ457070, and AY182241. Additional parent sequences are described in PCT/US2009/041581 and PCT/US2010/032134.

[0130] In various embodiments, the parent isoprene synthase has at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 92%, at least about 94%, at least about 96%, at least about 98%, at least about 99% sequence identity with *P. alba* MEA (SEQ ID NO:1). In other embodiments, the parent isoprene synthase has at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 92%, at least about 94%, at least about 96%, at least about 98%, at least about 99% sequence identity with full-length *P. alba* or complete *P. alba* (see, e.g., Figure 38A-38B).

[0131] Several methods are known in the art that are suitable for generating variants of the enzymes of the present invention, including but not limited to site-saturation mutagenesis, scanning mutagenesis, insertional mutagenesis, random mutagenesis, site-directed mutagenesis, and directed-evolution, as well as various other recombinatorial approaches.

[0132] Standard methods can be used to determine whether a polypeptide has isoprene synthase polypeptide activity by measuring the ability of the polypeptide to convert DMAPP into isoprene *in vitro*, in a cell extract, or *in vivo*. Isoprene synthase polypeptide activity in the cell extract can be measured, for example, as described in Silver et al., J. Biol. Chem. 270:13010-13016, (1995). In one exemplary assay, DMAPP (Sigma) can be evaporated to dryness under a stream of nitrogen and rehydrated to a concentration of 100 mM in 100 mM potassium phosphate buffer pH 8.2 and stored at -20 °C. To perform the assay, a solution of 5 µL of 1M MgCl₂, 1 mM (250 µg/ml) DMAPP, 65 µL of Plant Extract Buffer (PEB) (50 mM Tris-HCl, pH 8.0, 20 mM MgCl₂, 5% glycerol, and 2 mM DTT) can be added to 25 µL of cell extract in a 20 ml Headspace vial with a metal screw cap and teflon coated silicon septum (Agilent Technologies) and cultured at 37°C for 15 minutes with shaking. The reaction can be quenched by adding 200 µL of 250 mM EDTA and quantified by GC/MS.

[0133] In some aspects, the isoprene synthase polypeptide is a plant isoprene synthase polypeptide or a variant thereof. In some aspects, the isoprene synthase polypeptide is an isoprene synthase from *Pueraria* or a variant thereof. In some aspects, the isoprene synthase polypeptide is an isoprene synthase from *Populus* or a variant thereof. In some aspects, the isoprene synthase polypeptide is a poplar isoprene synthase polypeptide or a variant thereof. In some aspects, the isoprene synthase polypeptide is a kudzu isoprene synthase polypeptide or a variant thereof. In some aspects, the isoprene synthase polypeptide is a polypeptide from *Pueraria* or *Populus* or a hybrid, *Populus alba* x *Populus tremula*, or a variant thereof. In some aspects, the isoprene synthase polypeptide is a willow isoprene synthase polypeptide or a variant thereof. In some aspects, the isoprene synthase polypeptide is a eucalyptus isoprene synthase polypeptide or a variant thereof. In some aspects, the isoprene synthase polypeptide is from *Robinia*, *Salix*, or *Melaleuca* or variants thereof.

[0134] In some aspects, the isoprene synthase polypeptide or nucleic acid is from the family Fabaceae, such as the Faboideae subfamily. In some aspects, the isoprene synthase polypeptide

or nucleic acid is a polypeptide or nucleic acid from *Pueraria montana* (kudzu) (Sharkey et al., Plant Physiology 137: 700-712, (2005)), *Pueraria lobata*, poplar (such as *Populus alba*, *Populus nigra*, *Populus trichocarpa*, or *Populus alba x tremula* (CAC35696) (Miller et al., Planta 213:483-487, (2001)), aspen (such as *Populus tremuloides*) (Silver et al., JBC 270(22):13010-1316, (1995)), English Oak (*Quercus robur*) (Zimmer et al., WO 98/02550), or a variant thereof. In some aspects, the isoprene synthase polypeptide is an isoprene synthase from *Pueraria montana*, *Pueraria lobata*, *Populus tremuloides*, *Populus alba*, *Populus nigra*, or *Populus trichocarpa* or a variant thereof. In some aspects, the isoprene synthase polypeptide is an isoprene synthase from *Populus alba* or a variant thereof. In some aspects, the isoprene synthase is *Populus balsamifera* (Genbank JN173037), *Populus deltoides* (Genbank JN173039), *Populus fremontii* (Genbank JN173040), *Populus granididenta* (Genbank JN173038), *Salix* (Genbank JN173043), *Robinia pseudoacacia* (Genbank JN173041), *Wisteria* (Genbank JN173042), *Eucalyptus globulus* (Genbank AB266390) or *Melaleuca alterniflora* (Genbank AY279379) or variant thereof. In some aspects, the nucleic acid encoding the isoprene synthase (e.g., isoprene synthase from *Populus alba* or a variant thereof) is codon optimized.

[0135] In some aspects, the isoprene synthase nucleic acid or polypeptide is a naturally-occurring polypeptide or nucleic acid (e.g., naturally-occurring polypeptide or nucleic acid from *Populus*). In some aspects, the isoprene synthase nucleic acid or polypeptide is not a wild-type or naturally-occurring polypeptide or nucleic acid. In some aspects, the isoprene synthase nucleic acid or polypeptide is a variant of a wild-type or naturally-occurring polypeptide or nucleic acid (e.g., a variant of a wild-type or naturally-occurring polypeptide or nucleic acid from *Populus*).

[0136] In some aspects, the isoprene synthase polypeptide is a variant. In some aspects, the isoprene synthase polypeptide is a variant of a wild-type or naturally occurring isoprene synthase. In some aspects, the variant has improved activity such as improved catalytic activity compared to the wild-type or naturally occurring isoprene synthase. The increase in activity (e.g., catalytic activity) can be at least about any of 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 95%. In some aspects, the increase in activity such as catalytic activity is at least about any of 1 fold, 2 folds, 5 folds, 10 folds, 20 folds, 30 folds, 40 folds, 50 folds, 75 folds, or 100 folds. In some aspects, the increase in activity such as catalytic activity is about 10% to about 100 folds (e.g., about 20% to about 100 folds, about 50% to about 50 folds, about 1 fold to

about 25 folds, about 2 folds to about 20 folds, or about 5 folds to about 20 folds). In some aspects, the variant has improved solubility compared to the wild-type or naturally occurring isoprene synthase. The increase in solubility can be at least about any of 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 95%. The increase in solubility can be at least about any of 1 fold, 2 folds, 5 folds, 10 folds, 20 folds, 30 folds, 40 folds, 50 folds, 75 folds, or 100 folds. In some aspects, the increase in solubility is about 10% to about 100 folds (e.g., about 20% to about 100 folds, about 50% to about 50 folds, about 1 fold to about 25 folds, about 2 folds to about 20 folds, or about 5 folds to about 20 folds). In some aspects, the isoprene synthase polypeptide is a variant of naturally occurring isoprene synthase and has improved stability (such as thermostability) compared to the naturally occurring isoprene synthase.

[0137] In some aspects, the variant comprises a mutation in the wild-type or naturally occurring isoprene synthase. In some aspects, the variant has at least one amino acid substitution, at least one amino acid insertion, and/or at least one amino acid deletion. In some aspects, the variant has at least one amino acid substitution. In some aspects, the number of differing amino acid residues between the variant and wild-type or naturally occurring isoprene synthase can be one or more, e.g. 1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50, or more amino acid residues. Naturally occurring isoprene synthases can include any isoprene synthases from plants, for example, kudzu isoprene synthases, poplar isoprene synthases, English oak isoprene synthases, and willow isoprene synthases. In some aspects, the variant is a variant of isoprene synthase from *Populus alba*. In some aspects, the variant of isoprene synthase from *Populus alba* has at least one amino acid substitution, at least one amino acid insertion, and/or at least one amino acid deletion. In some aspects, the variant is a truncated *Populus alba* isoprene synthase. In some aspects, the nucleic acid encoding variant (e.g., variant of isoprene synthase from *Populus alba*) is codon optimized (for example, codon optimized based on host cells where the heterologous isoprene synthase is expressed).

[0138] The isoprene synthase polypeptide provided herein can be any of the isoprene synthases or isoprene synthase variants described in WO 2009/132220, WO 2010/124146, and U.S. 2010/0086978, the contents of which are expressly incorporated herein by reference in their entirety with respect to the isoprene synthases and isoprene synthase variants.

[0139] Any one of the promoters described herein (e.g., promoters described herein and identified in the Examples of the present disclosure including inducible promoters and

constitutive promoters) can be used to drive expression of any of the isoprene synthases described herein.

[0140] Suitable isoprene synthases include, but are not limited to, those identified by Genbank Accession Nos. AY341431, AY316691, AY279379, AJ457070, and AY182241. Types of isoprene synthases which can be used in any one of the compositions or methods including methods of making microorganisms encoding isoprene synthase described herein are also described in International Patent Application Publication Nos. WO 2009/076676, WO 2010/003007, WO 2009/132220, WO 2010/031062, WO 2010/031068, WO 2010/031076, WO 2010/013077, WO 2010/031079, WO 2010/148150, WO 2010/124146, WO 2010/078457, and WO 2010/148256.

Nucleic Acids Encoding Isoprene Synthase Variants

[0141] The invention also provides for nucleic acids encoding any of the isoprene synthase variants described herein. In various embodiments, the nucleic acid is a recombinant nucleic acid. For instance, in some embodiments, an isoprene synthase variant nucleic acid is operably linked to another nucleic acid encoding all or a portion of another polypeptide such that the recombinant nucleic acid encodes a fusion polypeptide that includes an isoprene synthase variant and all or part of another polypeptide (*e.g.*, a peptide that facilitates purification or detection of the fusion polypeptide, such as a His-tag). In some embodiments, part or all of a recombinant nucleic acid is chemically synthesized. In some aspects, the nucleic acid is a heterologous nucleic acid. By “heterologous nucleic acid” is meant a nucleic acid whose nucleic acid sequence is not identical to that of another nucleic acid naturally found in the same host cell.

[0142] In some embodiments, the nucleic acid includes at least or about 50, 100, 150, 200, 300, 400, 500, 600, 700, 800, or more contiguous nucleotides from a naturally-occurring isoprene synthase nucleic acid. In some aspects, the nucleic acid has one or more mutations compared to the sequence of a wild-type (*i.e.*, a sequence occurring in nature) isoprene synthase nucleic acid. In some embodiments, the nucleic acid has one or more mutations (*e.g.*, a silent mutation) that increase the transcription or translation of isoprene synthase nucleic acid. In some embodiments, the nucleic acid is a degenerate variant of any nucleic acid encoding an isoprene synthase polypeptide.

[0143] An isoprene synthase nucleic acid can be incorporated into a vector, such as an expression vector, using standard techniques known to one of skill in the art. Methods used to ligate the DNA construct comprising a nucleic acid of interest such as isoprene synthase, a promoter, a terminator, and other sequences and to insert them into a suitable vector are well known in the art. Additionally, vectors can be constructed using known recombination techniques (*e.g.*, Invitrogen Life Technologies, Gateway Technology).

[0144] In some embodiments, it may be desirable to over-express isoprene synthase nucleic acids at levels far higher than currently found in naturally-occurring cells. This result may be accomplished by the selective cloning of the nucleic acids encoding those polypeptides into multicopy plasmids or placing those nucleic acids under a strong inducible or constitutive promoter. Methods for over-expressing desired polypeptides are common and well known in the art of molecular biology and examples may be found in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor, (2001).

[0145] In other embodiments, the nucleic acids encoding for polypeptides having isoprene synthase activity is integrated into the chromosome of the host cell using molecular biology techniques readily available to one of skill in the art.

[0146] As is discussed in greater detail below, host cells engineered to contain isoprene synthase variants with improved kinetic properties can also be used in conjunction with one or more MVA pathway polypeptides, one or more DXP pathway polypeptides, IDI, and other components to maximize the isoprene production. Exemplary schematic of the MVA pathway and DXP pathway is shown in Figure 12.

MVA pathway nucleic acids and polypeptides

[0147] The complete MVA pathway can be subdivided into two groups: an upper and lower pathway. In the upper portion of the MVA pathway, acetyl Co-A produced during cellular metabolism is converted to mevalonate via the actions of polypeptides having either: (a) (i) thiolase activity or (ii) acetoacetyl-CoA synthase activity, (b) HMG-CoA reductase, and (c) HMG-CoA synthase enzymatic activity. First, acetyl Co-A is converted to acetoacetyl CoA via the action of a thiolase or an acetoacetyl-CoA synthase (which utilizes acetyl-CoA and malonyl-CoA). Next, acetoacetyl-CoA is converted to 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) by

the enzymatic action of HMG-CoA synthase. This Co-A derivative is reduced to mevalonate by HMG-CoA reductase. In the lower MVA pathway, mevalonate is then converted into mevalonate-5-phosphate via the action of mevalonate kinase which is subsequently transformed into 5-diphosphomevalonate by the enzymatic activity of phosphomevalonate kinase. Finally, IPP is formed from 5-diphosphomevalonate by the activity of the enzyme mevalonate-5-pyrophosphate decarboxylase. The mevalonate-dependent biosynthetic pathway is particularly important for the production of the isoprenoid precursor molecules dimethylallyl diphosphate (DMAPP) and isopentenyl pyrophosphate (IPP).

[0148] Exemplary MVA pathway polypeptides include, but are not limited to: 3-hydroxy-3-methylglutaryl-CoA synthase (HMG-CoA synthase) polypeptides (*e.g.*, an enzyme encoded by *mvaS*), 3-hydroxy-3-methylglutaryl-CoA reductase (HMG-CoA reductase) polypeptides (*e.g.*, enzyme encoded by *mvaR* or enzyme encoded by *mvaE* that has been modified to be thiolase-deficient but still retains its reductase activity), mevalonate kinase (MVK) polypeptides, phosphomevalonate kinase (PMK) polypeptides, diphosphomevalonate decarboxylase (MVD) polypeptides, phosphomevalonate decarboxylase (PMDC) polypeptides, isopentenyl phosphate kinase (IPK) polypeptides, IPP isomerase polypeptides, IDI polypeptides, and polypeptides (*e.g.*, fusion polypeptides) having an activity of two or more MVA pathway polypeptides. In particular, MVA pathway polypeptides include polypeptides, fragments of polypeptides, peptides, and fusions polypeptides that have at least one activity of an MVA pathway polypeptide. Exemplary MVA pathway nucleic acids include nucleic acids that encode a polypeptide, fragment of a polypeptide, peptide, or fusion polypeptide that has at least one activity of an MVA pathway polypeptide. Exemplary MVA pathway polypeptides and nucleic acids include naturally-occurring polypeptides and nucleic acids from any of the source organisms described herein. In addition, variants of MVA pathway polypeptide that confer the result of better isoprene production can also be used as well.

[0149] Non-limiting examples of MVA pathway polypeptides which can be used are described in International Patent Application Publication No. WO 2009/076676; WO 2010/003007 and WO 2010/148150.

Nucleic acids encoding polypeptides of the upper MVA pathway

[0150] The upper portion of the MVA pathway uses acetyl Co-A produced during cellular metabolism as the initial substrate for conversion to mevalonate via the actions of polypeptides having either: (a) (i) thiolase activity or (ii) acetoacetyl-CoA synthase activity, (b) HMG-CoA reductase, and (c) HMG-CoA synthase enzymatic activity. First, acetyl Co-A is converted to acetoacetyl CoA via the action of a thiolase or an acetoacetyl-CoA synthase (which utilizes acetyl-CoA and malonyl-CoA). Next, acetoacetyl-CoA is converted to 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) by the enzymatic action of HMG-CoA synthase. This Co-A derivative is reduced to mevalonate by HMG-CoA reductase, which is a rate-limiting step of the mevalonate pathway of isoprene production.

[0151] Non-limiting examples of upper MVA pathway polypeptides include: acetyl-CoA acetyltransferase (AA-CoA thiolase) polypeptides, acetoacetyl-CoA synthase polypeptides, 3-hydroxy-3-methylglutaryl-CoA synthase (HMG-CoA synthase) polypeptides, 3-hydroxy-3-methylglutaryl-CoA reductase (HMG-CoA reductase) polypeptides. Upper MVA pathway polypeptides can include polypeptides, fragments of polypeptides, peptides, and fusions polypeptides that have at least one activity of an upper MVA pathway polypeptide. Exemplary upper MVA pathway nucleic acids include nucleic acids that encode a polypeptide, fragment of a polypeptide, peptide, or fusion polypeptide that has at least one activity of an upper MVA pathway polypeptide. Exemplary MVA pathway polypeptides and nucleic acids include naturally-occurring polypeptides and nucleic acids from any of the source organisms described herein. Thus, it is contemplated herein that any gene encoding an upper MVA pathway polypeptide can be used in the present invention.

[0152] In certain embodiments, various options of *mvaE* and *mvaS* genes from *L. grayi*, *E. faecium*, *E. gallinarum*, *E. casseliflavus* and/or *E. faecalis* alone or in combination with one or more other *mvaE* and *mvaS* genes encoding proteins from the upper MVA pathway are contemplated within the scope of the invention. In other embodiments, an acetoacetyl-CoA synthase gene is contemplated within the scope of the present invention in combination with one or more other genes encoding: (i) 3-hydroxy-3-methylglutaryl-CoA synthase (HMG-CoA synthase) polypeptides and 3-hydroxy-3-methylglutaryl-CoA reductase (HMG-CoA reductase) polypeptides. Thus, in certain aspects, any of the combinations of genes contemplated herein can be expressed in recombinant cells in any of the ways described herein.

[0153] Additional non-limiting examples of upper MVA pathway polypeptides which can be used herein are described in International Patent Application Publication Nos. WO 2009/076676; WO 2010/003007 and WO 2010/148150.

Acetoacetyl-CoA synthase nucleic acids and polypeptides

[0154] The acetoacetyl-CoA synthase gene (aka *nphT7*) is a gene encoding an enzyme having the activity of synthesizing acetoacetyl-CoA from malonyl-CoA and acetyl-CoA and having minimal activity (*e.g.*, no activity) of synthesizing acetoacetyl-CoA from two acetyl-CoA molecules. *See, e.g.*, Okamura et al., PNAS Vol 107, No. 25, pp. 11265-11270 (2010), the contents of which are expressly incorporated herein for teaching about *nphT7*. An acetoacetyl-CoA synthase gene from an actinomycete of the genus *Streptomyces* CL190 strain was described in JP Patent Publication (Kokai) No. 2008-61506 A and US 2010/0285549. Acetoacetyl-CoA synthase can also be referred to as acetyl CoA:malonyl CoA acyltransferase. A representative acetoacetyl-CoA synthase (or acetyl CoA:malonyl CoA acyltransferase) that can be used is Genbank AB540131.1.

[0155] In one embodiment, acetoacetyl-CoA synthase of the present invention synthesizes acetoacetyl-CoA from malonyl-CoA and acetyl-CoA via an irreversible reaction. The use of acetoacetyl-CoA synthase to generate acetyl-CoA provides an additional advantage in that this reaction is irreversible while acetoacetyl-CoA thiolase enzyme's action of synthesizing acetoacetyl-CoA from two acetyl-CoA molecules is reversible. Consequently, the use of acetoacetyl-CoA synthase to synthesize acetoacetyl-CoA from malonyl-CoA and acetyl-CoA can result in significant improvement in productivity for isoprene compared with using thiolase to generate the end same product.

[0156] Furthermore, the use of acetoacetyl-CoA synthase to produce isoprene provides another advantage in that acetoacetyl-CoA synthase can convert malonyl CoA to acetyl CoA via decarboxylation of the malonyl CoA. Thus, stores of starting substrate are not limited by the starting amounts of acetyl CoA. The synthesis of acetoacetyl-CoA by acetoacetyl-CoA synthase can still occur when the starting substrate is only malonyl-CoA. In one embodiment, the pool of starting malonyl-CoA is increased by using host strains that have more malonyl-CoA. Such increased pools can be naturally occurring or be engineered by molecular manipulation. *See, for*

example Fowler, et al., Applied and Environmental Microbiology, Vol. 75, No. 18, pp. 5831-5839 (2009).

[0157] In any of the aspects or embodiments described herein, an enzyme that has the ability to synthesize acetoacetyl-CoA from malonyl-CoA and acetyl-CoA can be used. Non-limiting examples of such an enzyme are described herein. In certain embodiments described herein, an acetoacetyl-CoA synthase gene derived from an actinomycete of the genus *Streptomyces* having the activity of synthesizing acetoacetyl-CoA from malonyl-CoA and acetyl-CoA can be used.

[0158] An example of such an acetoacetyl-CoA synthase gene is the gene encoding a protein having the amino acid sequence In another aspect, the acetoacetyl-CoA synthase gene encodes a protein having the amino acid sequence of:

MTDVRFRIIGTGAYVPERIVSNDEVGAPAGVDDDWITRKTGIRQRRWAADDQATSDLA
TAAGRAALKAAGITPEQLTVIAVATSTPDRPQPPTAAYVQHHLGATGTAAFDVNAVCS
GTVFALSSVAGTLVYRGGYALVIGADLYSRILNPADRKTVVLFGDGAGAMVLGPTSTG
TGPIVRRVALHTFGGLTDLIRVPAGGSRQPLDTDGLDAGLQYFAMDGREVRRFVTEHLP
QLIKGFLHEAGVDAADISHFVPHQANGVMLDEVFGELHLPRATMHRTVETYGNTGAAS
IPITMDAAVRAGSFRPGE LVLLAGFGGGMAASFALIEW (SEQ ID NO:2). Such a protein
having the amino acid sequence of SEQ ID NO: 2 corresponds to an acetoacetyl-CoA synthase
having activity of synthesizing acetoacetyl-CoA from malonyl-CoA and acetyl-CoA and having
no activity of synthesizing acetoacetyl-CoA from two acetyl-CoA molecules.

[0159] In one embodiment, the gene encoding a protein having the amino acid sequence of SEQ ID NO: 2 can be obtained by a nucleic acid amplification method (e.g., PCR) with the use of genomic DNA obtained from an actinomycete of the *Streptomyces* sp. CL190 strain as a template and a pair of primers that can be designed with reference to JP Patent Publication (Kokai) No. 2008-61506 A.

[0160] As described herein, an acetoacetyl-CoA synthase gene for use in the present invention is not limited to a gene encoding a protein having the amino acid sequence of SEQ ID NO: 2 from an actinomycete of the *Streptomyces* sp. CL190 strain. Any gene encoding a protein having the ability to synthesize acetoacetyl-CoA from malonyl-CoA and acetyl-CoA and which does not synthesize acetoacetyl-CoA from two acetyl-CoA molecules can be used in the presently described methods. In certain embodiments, the acetoacetyl-CoA synthase gene can

be a gene encoding a protein having an amino acid sequence with high similarity or substantially identical to the amino acid sequence of SEQ ID NO: 2 and having the function of synthesizing acetoacetyl-CoA from malonyl-CoA and acetyl-CoA. The expression “highly similar” or “substantially identical” refers to, for example, at least about 80% identity, at least about 85%, at least about 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%, and at least about 99% identity. As used above, the identity value corresponds to the percentage of identity between amino acid residues in a different amino acid sequence and the amino acid sequence of SEQ ID NO: 2, which is calculated by performing alignment of the amino acid sequence of SEQ ID NO: 2 and the different amino acid sequence with the use of a program for searching for a sequence similarity.

[0161] In other embodiments, the acetoacetyl-CoA synthase gene may be a gene encoding a protein having an amino acid sequence derived from the amino acid sequence of SEQ ID NO:2 by substitution, deletion, addition, or insertion of 1 or more amino acid(s) and having the function of synthesizing acetoacetyl-CoA from malonyl-CoA and acetyl-CoA. Herein, the expression “more amino acids” refers to, for example, 2 to 30 amino acids, preferably 2 to 20 amino acids, more preferably 2 to 10 amino acids, and most preferably 2 to 5 amino acids.

[0162] In still other embodiments, the acetoacetyl-CoA synthase gene may consist of a polynucleotide capable of hybridizing to a portion or the entirety of a polynucleotide having a nucleotide sequence complementary to the nucleotide sequence encoding the amino acid sequence of SEQ ID NO:2 under stringent conditions and capable of encoding a protein having the function of synthesizing acetoacetyl-CoA from malonyl-CoA and acetyl-CoA. Herein, hybridization under stringent conditions corresponds to maintenance of binding under conditions of washing at 60 °C 2x SSC. Hybridization can be carried out by conventionally known methods such as the method described in Sambrook, J. et al. *Molecular Cloning, A Laboratory Manual*, 3rd Ed., Cold Spring Harbor Laboratory (2001).

[0163] As described herein, a gene encoding an acetoacetyl-CoA synthase having an amino acid sequence that differs from the amino acid sequence of SEQ ID NO:2 can be isolated from potentially any organism, for example, an actinomycete that is not obtained from the *Streptomyces* sp. CL190 strain. In addition, acetoacetyl-CoA synthase genes for use herein can be obtained by modifying a polynucleotide encoding the amino acid sequence of SEQ ID NO: 2

by a method known in the art. Mutagenesis of a nucleotide sequence can be carried out by a known method such as the Kunkel method or the gapped duplex method or by a method similar to either thereof. For instance, mutagenesis may be carried out with the use of a mutagenesis kit (*e.g.*, product names; Mutant-K and Mutant-G (TAKARA Bio)) for site-specific mutagenesis, product name; an LA PCR in vitro Mutagenesis series kit (TAKARA Bio), and the like.

[0164] The activity of an acetoacetyl-CoA synthase having an amino acid sequence that differs from the amino acid sequence of SEQ ID NO:2 can be evaluated as described below. Specifically, a gene encoding a protein or polypeptide to be evaluated is first introduced into a host cell such that the gene can be expressed therein, followed by purification of the protein by a technique such as chromatography. Malonyl-CoA and acetyl-CoA are added as substrates to a buffer containing the obtained protein to be evaluated, followed by, for example, incubation at a desired temperature (*e.g.*, 10°C to 60°C). After the completion of reaction, the amount of substrate lost and/or the amount of product (acetoacetyl-CoA) produced are determined. Thus, it is possible to evaluate whether or not the protein being tested has the function of synthesizing acetoacetyl-CoA from malonyl-CoA and acetyl-CoA and to evaluate the degree of synthesis. In such case, it is possible to examine whether or not the protein has the activity of synthesizing acetoacetyl-CoA from two acetyl-CoA molecules by adding acetyl-CoA alone as a substrate to a buffer containing the obtained protein to be evaluated and determining the amount of substrate lost and/or the amount of product produced in a similar manner.

Genes encoding *mvaE* and *mvaS* polypeptides

[0165] In certain embodiments, various options of *mvaE* and *mvaS* genes from *L. grayi*, *E. faecium*, *E. gallinarum*, *E. casseliflavus* and/or *E. faecalis* alone or in combination with one or more other *mvaE* and *mvaS* genes encoding proteins from the upper MVA pathway are contemplated within the scope of the invention. In some embodiments, the *mvaE* gene in *L. grayi*, *E. faecium*, *E. gallinarum*, *E. casseliflavus*, and *E. faecalis* encodes a polypeptide that possesses both thiolase and HMG-CoA reductase activities (Hedl, et al., J. Bacteriol. 184(8):2116–2122 (April 2002)). In fact, the *mvaE* gene product represented the first bifunctional enzyme of IPP biosynthesis found in eubacteria and the first example of HMG-CoA reductase fused to another protein in nature. The *mvaS* gene, on the other hand, can encode a polypeptide having an HMG-CoA synthase activity.

[0166] Accordingly, recombinant cells (*e.g.*, *E. coli*) can be engineered to express one or more *mvaE* and *mvaS* genes from *L. grayi*, *E. faecium*, *E. gallinarum*, *E. casseliflavus* and/or *E. faecalis* to produce isoprene. The one or more *mvaE* and *mvaS* genes can be expressed on a multicopy plasmid. The plasmid can be a high copy plasmid, a low copy plasmid, or a medium copy plasmid. Alternatively, the one or more *mvaE* and *mvaS* genes can be integrated into the host cell's chromosome. For both heterologous expression of the one or more *mvaE* and *mvaS* genes on a plasmid or as an integrated part of the host cell's chromosome, expression of the genes can be driven by either an inducible promoter or a constitutively expressing promoter. The promoter can be a strong driver of expression, it can be a weak driver of expression, or it can be a medium driver of expression of the one or more *mvaE* and *mvaS* genes.

Exemplary mvaE polypeptides and nucleic acids

[0167] The *mvaE* gene encodes a polypeptide that possesses both thiolase and HMG-CoA reductase activities. The thiolase activity of the polypeptide encoded by the *mvaE* gene converts acetyl Co-A to acetoacetyl CoA whereas the HMG-CoA reductase enzymatic activity of the polypeptide converts 3-hydroxy-3-methylglutaryl-CoA to mevalonate. Exemplary *mvaE* polypeptides and nucleic acids include naturally-occurring polypeptides and nucleic acids from any of the source organisms described herein as well as mutant polypeptides and nucleic acids derived from any of the source organisms described herein that have at least one activity of a *mvaE* polypeptide.

[0168] Mutant *mvaE* polypeptides include those in which one or more amino acid residues have undergone an amino acid substitution while retaining *mvaE* polypeptide activity (*i.e.*, the ability to convert acetyl Co-A to acetoacetyl CoA as well as the ability to convert 3-hydroxy-3-methylglutaryl-CoA to mevalonate). The amino acid substitutions can be conservative or non-conservative and such substituted amino acid residues can or cannot be one encoded by the genetic code. The standard twenty amino acid "alphabet" has been divided into chemical families based on similarity of their side chains. Those families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). A

“conservative amino acid substitution” is one in which the amino acid residue is replaced with an amino acid residue having a chemically similar side chain (*i.e.*, replacing an amino acid having a basic side chain with another amino acid having a basic side chain). A “non-conservative amino acid substitution” is one in which the amino acid residue is replaced with an amino acid residue having a chemically different side chain (*i.e.*, replacing an amino acid having a basic side chain with another amino acid having an aromatic side chain).

[0169] Amino acid substitutions in the *mvaE* polypeptide can be introduced to improve the functionality of the molecule. For example, amino acid substitutions that increase the binding affinity of the *mvaE* polypeptide for its substrate, or that improve its ability to convert acetyl CoA to acetoacetyl CoA and/or the ability to convert 3-hydroxy-3-methylglutaryl-CoA to mevalonate can be introduced into the *mvaE* polypeptide. In some aspects, the mutant *mvaE* polypeptides contain one or more conservative amino acid substitutions.

[0170] In one aspect, *mvaE* proteins that are not degraded or less prone to degradation can be used for the production of isoprene. Examples of gene products of *mvaEs* that are not degraded or less prone to degradation which can be used include, but are not limited to, those from the organisms *E. faecium*, *E. gallinarum*, *E. casseliflavus*, *E. faecalis*, and *L. grayi*. One of skill in the art can express *mvaE* protein in *E. coli* BL21 (DE3) and look for absence of fragments by any standard molecular biology techniques. For example, absence of fragments can be identified on Safestain stained SDS-PAGE gels following His-tag mediated purification or when expressed in isoprene-producing *E. coli* BL21 using the methods of detection described herein.

[0171] Standard methods, such as those described in Hedl et al. *J. Bacteriol.* 184(8): 2116–2122 (April 2002)) can be used to determine whether a polypeptide has *mvaE* activity, by measuring acetoacetyl-CoA thiolase as well as HMG-CoA reductase activity. In an exemplary assay, acetoacetyl-CoA thiolase activity is measured by spectrophotometer to monitor the change in absorbance at 302 nm that accompanies the formation or thiolysis of acetoacetyl-CoA. Standard assay conditions for each reaction to determine synthesis of acetoacetyl-CoA, are 1 mM acetyl-CoA, 10 mM MgCl₂, 50 mM Tris, pH 10.5 and the reaction is initiated by addition of enzyme. Assays can employ a final volume of 200 μl. For the assay, 1 enzyme unit (eu) represents the synthesis or thiolysis in 1 min of 1 μmol of acetoacetyl-CoA. In another exemplary assay, of HMG-CoA reductase activity can be monitored by spectrophotometer by

the appearance or disappearance of NADP(H) at 340 nm. Standard assay conditions for each reaction measured to show reductive decylation of HMG-CoA to mevalonate are 0.4 mM NADPH, 1.0 mM (*R,S*)-HMG-CoA, 100 mM KCl, and 100 mM K₂PO₄, pH 6.5. Assays employ a final volume of 200 µl. Reactions are initiated by adding the enzyme. For the assay, 1 eu represents the turnover, in 1 min, of 1 µmol of NADP(H). This corresponds to the turnover of 0.5 µmol of HMG-CoA or mevalonate.

[0172] Exemplary *mvaE* nucleic acids include nucleic acids that encode a polypeptide, fragment of a polypeptide, peptide, or fusion polypeptide that has at least one activity of a *mvaE* polypeptide. Exemplary *mvaE* polypeptides and nucleic acids include naturally-occurring polypeptides and nucleic acids from any of the source organisms described herein as well as mutant polypeptides and nucleic acids derived from any of the source organisms described herein. Exemplary *mvaE* nucleic acids include, for example, *mvaE* nucleic acids isolated from *Listeria grayi*_DSM 20601, *Enterococcus faecium*, *Enterococcus gallinarum* EG2, *Enterococcus faecalis*, and/or *Enterococcus casseliflavus*. The *mvaE* nucleic acid encoded by the *Listeria grayi*_DSM 20601 *mvaE* gene can have at least about 99%, 98%, 97%, 96%, 95%, 95%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, or 85% sequence identity to SEQ ID NO:7. The *mvaE* nucleic acid encoded by the *Enterococcus faecium* *mvaE* gene can have at least about 99%, 98%, 97%, 96%, 95%, 95%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, or 85% sequence identity to SEQ ID NO:8. The *mvaE* nucleic acid encoded by the *Enterococcus gallinarum* EG2 *mvaE* gene can have at least about 99%, 98%, 97%, 96%, 95%, 95%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, or 85% sequence identity to SEQ ID NO:9. The *mvaE* nucleic acid encoded by the *Enterococcus casseliflavus* *mvaE* gene can have at least about 99%, 98%, 97%, 96%, 95%, 95%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, or 85% sequence identity to SEQ ID NO:10. The *mvaE* nucleic acid encoded by the *Enterococcus faecalis* *mvaE* gene can have at least about 99%, 98%, 97%, 96%, 95%, 95%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, or 85% sequence identity to the *mvaE* gene previously disclosed in *E. coli* to produce mevalonate (see US 2005/0287655 A1; Tabata, K. and Hashimoto, S.-I. Biotechnology Letters 26:1487–1491, (2004)). The *mvaE* nucleic acid can be expressed in a recombinant cell on a multicopy plasmid. The plasmid can be a high copy plasmid, a low copy plasmid, or a medium copy plasmid. Alternatively, the *mvaE* nucleic acid can be integrated into the host cell's chromosome. For both heterologous expression of an *mvaE* nucleic acid on a plasmid or as an integrated part of the host cell's chromosome, expression of the nucleic acid can be driven by either an inducible

promoter or a constitutively expressing promoter. The promoter can be a strong driver of expression, it can be a weak driver of expression, or it can be a medium driver of expression of the *mvaE* nucleic acid.

Sequence of Listeria grayi DSM 20601 mvaE

atggttaaagacattgtaataattgatgccctccgtactcccacgtaagtagccggtcagctctcaaagatgacggcggtggaattggga
accgcagttacaaggctctgttcgagaagaacgaccaggtcaaagaccatgtagaacaagtcatttttggcaacgttttacaggcagggga
acggccagaatcccggcgtcagatcgccttaattctggcctgtccgcagagataccggcttcgactattaaccaggtgtgtggtctggc
ctgaaagcaataagcatggcgcgccaacagatcctactcggagaagcggaaagtaatgtagcaggaggtatcgaatccatgacgaatgc
gcccaggtattacatattataataaagaagaagacacccctctcaaagcctgttctacgatgaccttcgatggtctgaccgacgcgtttagcgg
aaagattatgggttaacagccgaaaatggtccgaacagtacggcgatcacgtgagggccaggacgcctttgcgtatggatcgcagatg
aaagcagcaaaaggccaagaacagggcattttcgcagctgaaatactgcctcttgaataggggacgaagttattactcaggacgagggg
gttcgtcaagagaccacctcgaaaaattaagtctgcttcggaccattttaaagaagatggtactgttacagcgggcaacgcctcaacgatc
aatgatggcgcctcagccgtgatcattgcatcaaaggagtttctgagacaaaccagattcctaccttgcgatcgtacatgatattacagag
ataggcattgatccatcaataatgggcattgtcccgtgagtgatcaataaactgatcgtacgtaaccaaaattagcatggaagaaatcga
ctctttgaaattaatgaggcatttgcagcatcctcgggtgtagttcaaaaagagttaagcattcccgatgaaaagatcaatattggcggttccg
gtattgcactaggccatcctcttggcggccacaggagcgcgcattgtaaccacctagcgcaccagttgaaacgtacacacggacgctatgg
tattgcctccctgtgattggcgggtggccttggcctagcaatattaatagaagtgcctcaggaaagatcagccggttaaaaaatttatcaattgg
cccgtgaggaccgtctggctagacttcaggagcaagccgtgatcagcccagctacaaaacatgtactggcagaaatgacacttctgaag
atattgccgacaatctgatcgaaaaatcaaatatctgaaatggaaatccctcttgggtgtggtttgaaatctgagggtcaatgataagattatacc
atcccactagcaactgaggaaccgagtgtaatcgtgcctgtaataatggtgcaaaaatggcaaacacctgggcgggtttcagtcagaatt
aaaagatgggttctgcgtggcgaattgtacttatgaaactcaaaagaaccgcaactatcgagcatacgatecggcagagaaagcggca
attttctgtccgcagcgcagtcacatccatcattgtgaaacgaggtgggggtctaaaagagatagtagtgctacgttcgatgatgatccg
acgttctgtctattgatctgatagttgatactaaagacgcaatggcgctaacatcattaacaccattctcaggggtgtagccggctttctgag
ggaaatccttaccgaagaaattctgttctctatcttattacgcaaccgaatcaattgtgaccgccagctgtcgcataccttacgaagcact
gagtaaaaaagggtgatggtaaacgaatcgtgaaaaagtggtgctgctgcatctaaattgcccagttagatccttatcgagctgcaaccaca
acaaaggattatgaatggtattgaggccgtcgttttggcctcaggaaatgacacacgggcggcgcggcagccgcacatgcgtatgcttc
acgcgatcagcactatcggggcttaagccagtggtgcagaggtgcagaaggcgcgttacacggggagatcagctaccacttgcactcggca
gcgttggcgggtgcaattgaggtcttgcctaaagcgaaggcggcattcgaatcatggggatcacagaggcgaaggagctggcagaagtc
acagctgcggtagggctggcgcaaacctggcggcgttaagagcgtttagtgaaggaatacagcaaggtcacatgtcgtcctcaggct
cgctctcttcattatcggtaggtgctacaggaaggaagttgaaatcctggccgaaaattacagggtctcgtatgaatcaggcgaacgc
tcagaccatactgcagagatcagatcgcaaaaagttgaattgtga (SEQ ID NO:7)

Sequence of *Enterococcus faecium mvaE*

atgacatgaacgttggaatcgataaaatgcattcttggccaccttactttgtggacatgactgatctggcagtagcacgggatgfcgatcc
caataagtttctgattggtattggccaggaccagatggcagttaatccgaaaacgcaggatattgtgacattgccacaaatgctgccaaaa
cactactgtcagctgaggaccttgataaaattgatatggcctatagtcggcaccgagagtggatcgaatccaaagcgagtgccgtagtgc
ttcacagggtgctcggtatccagaagttgctcgctccttgaaatcaaagaagcctgttatgggggtaccgcggctttacagttcgtgtaaac
cacattaggaatcatcctgaatcaaaggttctgtagttgcatcagatatcgcgaaatacggcctggcttctggagggtgaaccaacgcaaggt
gcaggcgtgtggctatgctcgtcactgaccctaagatcattgcttcaacgacgatagcctcgcgcttacacaagatatctatgacttct
ggcgaccagttggacatgactatcctatgctcgacgggcctcttagtacagagacctacatccagtcatttcagaccgatggcaggaatac
acaaaacggtcgcagcatgactggcagactttgctgcccttagctttcatatcccgtataactaaaatgggcaaaaaggcgtgcttgcaatc
cttgaaggcgaatcagaggaggctcagaaccgtatactagcaaaaatgaaaagagtatagcctactccagaaaaggcggtaacctgtata
ccggtagcctgtatctaggacttattcacttctggaaaatgcagaagacctaaagctggtgatttaataggcctcttttcttacggtccgggtg
ctgttgcggagttttctcaggaaggctggtgaggactatcaggaacagctactaaacaaaacatgccgaacagctggcccatagaaag
caactgacaatcgaggagtacgaaacgatgttctccgatcgttggacgtggacaaagacgccgaatacgaagacacattagcttatagca
ttctgctagtcgaaacaccgtacgtgagtacaggagtga (SEQ ID NO:8)

Sequence of *Enterococcus gallinarum EG2 mvaE*

atgaaagaagtggttatgattgatcgcgctcgcacaccattgggaaatacagaggtagtcttagtccttttacagcgggtggagctggggac
actggtcacgaaagggtgctggataaaacaaagcttaagaaagacaagatagaccaagtgatattcggcaatgtgcttcaggcaggaaa
cggacaaaacgttgaagacaaatagccctgaacagtggcttaccagttgacgtgccggcgatgactattaacgaagtttgggggtccgga
atgaaagcgggtatttagcccgccagttatacagttaggggagcagagtgtgctcattgcaggggttacggagtcaatgcaaacgac
cctgctgaaaccttaccagtcagagaccaacgaatacggagagccgatcatcaatggttaatgacgggctgacggatgcttttccaat
gctcacatgggtcttactgccgaaaagggtggcgacccagtttctcagtgctgcgcgaggaacaagaccgggtacgcattgtccagccaattga
aagcagcgcacgcggttgaagccggggtgttctcagaagagattattccgggtaagattagcgcagaggatgcttgagtgaagacgagg
cagtaagaggcaacagcactttgaaaaactgggcaccttgcggacggtgttttctgaagaggcaggttaccgctggcaatgcttacc
gctgaatgacggcgctagtgtcgtgattcttgcataaaaagaatacgcggaaaacaataatctgccttacctggcgacgataaaaggaggtg
cggaaagttggtatcgtcctctatcatgggtattgccccataaaaggccattcaaaagttaacagatcggtcgggatgaacctgtccacga
ttgatctgttgaaattaatgaagcattcgcggcatctagcattgttctcaagagctgcaattggacgaagaaaaagtgaatatctatggc
ggggcgatagctttagccatccaatcggcgcaagcggagcccggatactgacaaccttagcatacggcctctcgtgagcaaaaagcg
ttatggtattgctcattatgtatcggcgggtgttctgtgctgtgtaagaagtaaatggagcagaccacaagacgttcagaag
aaaaagtttaccagcttaccctccgagcggagatcgcagcttatcagaagaacgttctgactcaagaaacggcacttatttccagga
gcagacggttccgaagaactgtccgatcagatgattgagaatcaggtctccgaagtggaaattccaatgggaattgcacaaaatttccagat
taatggcaagaaaaatggattcctatggcgactgaagaaccttcagtaatagcggcagcatcgaacggcggcaaatctcggggaacatt

tgcgcggaaacgcctcagcggcttatgcgcggcagattgtcctgtctggcaaatcagaatatcaagccgtgataaatgccgtgaatcatc
gcaaagaagaactgattctttgcgcaaacgagtcgtaaccgagattgttaaacgcgggggagggtgttcaggatattctacgcgggagtta
tgggttctttcacgcgtattatcaatcgactttctggtggacgtcaaggacgcaatgggggcaaacatgatcaactctattctcgaaagcgtt
gcaaataaactgcgtgaatggtcccgaagaggaaatactgttctcatcctgtcaaacctcgtacggagtcctggcatctgcatgttgc
gagattcctttgaaagacttggcgtgtaacaaagaaattggtgaacagatgccaaagaaattcaacaggcaggggaatatgctaagcttga
cccttaccgcgcggcaaccataacaaggggattatgaacggatcgaagccgtcgttggcgaacgggaaacgacacacgggctgttcc
cgcttctattcacgcatacgcgccgtaatggcttgtaccaagggttaacggattggcagatcaaggcgataaactggttgtaaat AAC
agtcccactggctgtggcgactgtcggcgcgtcgaacatattacaaaagccaaagctccctcggcatgctggatattgattccgcaa
aagaactggcccaagtgatcggcggtaggttagcacagaatctggcggcgttacgtgattagtagcagaaggcattcagaaaggac
acatgggcttcaagcacgttctttagcgatttcgataggtgccatcggtaggagatagagcaagtcgcgaaaaaactgcgtgaagctga
aaaaatgaatcagcaaacggcaatacagattttgaaaaaattcgcgagaaatga (SEQ ID NO:9)

Sequence of *Enterococcus casseliflavus mvaE*

atgaaatcggattgaccgtctgtccttctcatcccgaattgtatttgacatgactgagctggcagaatcacgcggggatgatccagcta
aatatcatattggaatcgacaagatcagatggcagtgaaatcgcgcaaacgaggacatcataaactgggtgcaaacgctgcgagtaaga
tcgtgacagagaaagaccgcgagttgattgatatggaatcgttggcacggaatcaggaattgaccactcaaagcaagcggcgtgattatt
caccatctccttaaaatcagtcgttcgccgttttcgaggtaaaagaagcttgcctatggcggaaactgctgcctgacatggcgaaggag
tatgtcaaaaatcatccggagcgttaaggtcttggtaattgcgtcagacatcgcgcgttatggttggccagcggaggagaagtactcaagg
cgtggggggccgtagccatgatgattacacaaaacccccgatttttcgattgaagacgatagtgttttctcacagaggatatctatgattct
ggcggcctgattactccgagttccctgtagtggacgggcccttcaactcaacgtatatagagagtttcagaaagtttgaaccggcaca
aggaattgtccggaagagggtggaagattatcaagctattgctttcacataccctatacgaagatgggtaagaaagcgtccagagtgtt
tagacaaaccgatgaagataaccaggagcgttaatggctagatagaggagtctattcgtatagccggagaattgtaacctgtacaca
ggcagctgtaccttggcttacaagcttgttggaaaactctaaaagttacaaccgggagatcggatcggcctctttcctatggcagtggtg
cgggtgcccaggttctttaccgggtatttagaagaaaattaccaagagtacctgttcgctcaaaagccatcaagaaatgctggatagccggactc
ggattacggtcgatgaatacagaccatctttcagagactctgccagaacatggtgaatgccgaatatacagcgcacgtccccttttcta
taaccaagattgagaacgacattcgttattataaaatctga (SEQ ID NO:10)

Exemplary *mvaS* polypeptides and nucleic acids

[0173] The *mvaS* gene encodes a polypeptide that possesses HMG-CoA synthase activity. This polypeptide can convert acetoacetyl CoA to 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA). Exemplary *mvaS* polypeptides and nucleic acids include naturally-occurring polypeptides and nucleic acids from any of the source organisms described herein as well as

mutant polypeptides and nucleic acids derived from any of the source organisms described herein that have at least one activity of a *mvaS* polypeptide.

[0174] Mutant *mvaS* polypeptides include those in which one or more amino acid residues have undergone an amino acid substitution while retaining *mvaS* polypeptide activity (*i.e.*, the ability to convert acetoacetyl CoA to 3-hydroxy-3-methylglutaryl-CoA). Amino acid substitutions in the *mvaS* polypeptide can be introduced to improve the functionality of the molecule. For example, amino acid substitutions that increase the binding affinity of the *mvaS* polypeptide for its substrate, or that improve its ability to convert acetoacetyl CoA to 3-hydroxy-3-methylglutaryl-CoA can be introduced into the *mvaS* polypeptide. In some aspects, the mutant *mvaS* polypeptides contain one or more conservative amino acid substitutions.

[0175] Standard methods, such as those described in Quant et al. (Biochem J., 262:159-164 (1989)), can be used to determine whether a polypeptide has *mvaS* activity, by measuring HMG-CoA synthase activity. In an exemplary assay, HMG-CoA synthase activity can be assayed by spectrophotometrically measuring the disappearance of the enol form of acetoacetyl-CoA by monitoring the change of absorbance at 303 nm. A standard 1 ml assay system containing 50 mM-Tris/HCl, pH 8.0, 10 mM-MgCl₂ and 0.2 mM-dithiothreitol at 30 °C; 5 mM-acetyl phosphate, 10, M-acetoacetyl- CoA and 5 μl samples of extracts can be added, followed by simultaneous addition of acetyl-CoA (100 μM) and 10 units of PTA. HMG-CoA synthase activity is then measured as the difference in the rate before and after acetyl-CoA addition. The absorption coefficient of acetoacetyl-CoA under the conditions used (pH 8.0, 10 mM-MgCl₂), is $12.2 \times 10^3 \text{ M}^{-1} \text{ cm}^{-1}$. By definition, 1 unit of enzyme activity causes 1 μmol of acetoacetyl-CoA to be transformed per minute.

[0176] Alternatively, production of mevalonate in recombinant cells can be measured by, without limitation, gas chromatography (see U.S. Patent Application Publication No.: US 2005/0287655 A1) or HPLC (See U.S. Patent Application Publication No.: 2011/0159557 A1). As an exemplary assay, cultures can be inoculated in shake tubes containing LB broth supplemented with one or more antibiotics and incubated for 14h at 34°C at 250 rpm. Next, cultures can be diluted into well plates containing TM3 media supplemented with 1% Glucose, 0.1% yeast extract, and 200 μM IPTG to final OD of 0.2. The plate are then sealed with a Breath Easier membrane (Diversified Biotech) and incubated at 34°C in a shaker/incubator at 600 rpm for 24 hours. 1 mL of each culture is then centrifuged at 3,000 x g for 5 min.

Supernatant is then added to 20% sulfuric acid and incubated on ice for 5 min. The mixture is then centrifuged for 5 min at 3000 x g and the supernatant was collected for HPLC analysis. The concentration of mevalonate in samples is determined by comparison to a standard curve of mevalonate (Sigma). The glucose concentration can additionally be measured by performing a glucose oxidase assay according to any method known in the art. Using HPLC, levels of mevalonate can be quantified by comparing the refractive index response of each sample versus a calibration curve generated by running various mevalonate containing solutions of known concentration.

[0177] Exemplary *mvaS* nucleic acids include nucleic acids that encode a polypeptide, fragment of a polypeptide, peptide, or fusion polypeptide that has at least one activity of a *mvaS* polypeptide. Exemplary *mvaS* polypeptides and nucleic acids include naturally-occurring polypeptides and nucleic acids from any of the source organisms described herein as well as mutant polypeptides and nucleic acids derived from any of the source organisms described herein. Exemplary *mvaS* nucleic acids include, for example, *mvaS* nucleic acids isolated from *Listeria grayi*_DSM 20601, *Enterococcus faecium*, *Enterococcus gallinarum* EG2, *Enterococcus faecalis*, and/or *Enterococcus casseliflavus*. The *mvaS* nucleic acid encoded by the *Listeria grayi*_DSM 20601 *mvaS* gene can have at least about 99%, 98%, 97%, 96%, 95%, 95%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, or 85% sequence identity to SEQ ID NO:11. The *mvaS* nucleic acid encoded by the *Enterococcus faecium* *mvaS* gene can have at least about 99%, 98%, 97%, 96%, 95%, 95%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, or 85% sequence identity to SEQ ID NO:12. The *mvaS* nucleic acid encoded by the *Enterococcus gallinarum* EG2 *mvaS* gene can have at least about 99%, 98%, 97%, 96%, 95%, 95%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, or 85% sequence identity to SEQ ID NO:13. The *mvaS* nucleic acid encoded by the *Enterococcus casseliflavus* *mvaS* gene can have at least about 99%, 98%, 97%, 96%, 95%, 95%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, or 85% sequence identity to SEQ ID NO:14. The *mvaS* nucleic acid encoded by the *Enterococcus faecalis* *mvaS* gene can have at least about 99%, 98%, 97%, 96%, 95%, 95%, 93%, 92%, 91%, 90%, 89%, 88%, 87%, 86%, or 85% sequence identity to the *mvaE* gene previously disclosed in *E. coli* to produce mevalonate (see US 2005/0287655 A1; Tabata, K. and Hashimoto, S.-I. Biotechnology Letters 26: 1487–1491, (2004)).

Sequence of *Listeria grayi* DSM 20601 mvaS

atggaagaagtggtaattatagatgcacgtcggactccgattggtaaatatcacgggtcgttgaagaagtttcagcgggtggcgtggggac
ggccgtggctaaagacatgttcgaacgcaaccagaaaatcaagaggagatcgcgcaggcgcataattggaatgtcttcagggcaggaaa
tggccagaaccccgcgcggcaagttgctcttcaatcagggtgtccgttgacattcccgttctacaattaacgaggtttgtgggtctggttg
aaagctatcttgatgggcatggacaatccaactcggcaaacgcaagtagtgcaggcggcggcattgaatcaatgacaatgcgcca
agcctgtcccactataacaaggcggaggatagatgtgccagtgtcgcagcatgacactggatggtctgacagacgcattttctagtaa
acctatgggattaacagcggaaaacgtcgcacagcgtacggatctcccgtgaggcgcagatcaattcgcataatctcagatgaaa
gcagcaaaagcgcaggcagaaaacaatcgttaagaaattgtccactggcgggtgaaactaaaaccatcacagctgacgaagggat
cagatcccaacaacgatggagaaactggcaagtctcaaacctgttttaaaaccgatggcactgtaaccgcagggaatgctagcaccatt
aatgacggggccgcccttgtgctgctgctgcaaaaactactgcgaaactaatgacataccgtaccttgcgacaatcaagaattgttga
gttggaatcgcagcggagattatgggcatctccgataaaagcgatacaaacattgttacaatacaaaaagttagcctcgaagatattgga
gttttgaataaatgaagccttggcgaagtagcatagtggtgaaatctgagttgggattagatccggcctaaagttaaccgttatgggggtg
gtatatccttagtcatgcaattggggcaaccggcgtcgcctggccacttactggtgatcaaatgcaggagatacaagcacgttatgta
ttgcgagcctgtgctgtggtggacttgactggcaatgcttttagaacgtccaactattgagaaggctaaaccgacagacaaaaagtct
atgaattgacaccagctgaacggtgcaagagctggaaaatacaagaaaatcagttctgaaactaaacagcagttatctcagatgatgcttg
ccgaggacactgcaaaccatttgatagaaaatcaaatatcagagattgaaactcccaatgggctcgggatgaacctgaaggttgatgggaa
agcctatgtgtgccaatggcgcagggagagccgtccgtcgcggccatgtctaatgtgccaataatggccggcgaattcacactcag
tcgaaagaacggctgctcagaggtcagattgtttcagcgcgaagaatccgaatgaaatcgaacagagaatagctgagaaccaagctttga
tttcgaacgtgccgaacagtcctatccttccattgtgaaaagagagggtctccgccgattgcacttgcattttctgccgattctcag
caggagtctcggaccagtccacattttatcagtgaccctttttagatgtgaaagacgcgatgggggcaaatatcataatgcaacttgc
agggcgtcgcagccctgttcgcgaatggttccccaatgaggaaattcttttctattctcgaacttggctacggagagcttagtcacggct
gtttgtgaagtccatttagtgcacttagcaagagaggtggtgcaacggtggcccagaaaattgtgcaggcgtcgtcttcgaaaagacag
accataaccgcgcagtgaccacaacaagggtatgaacggtgtagaggctgttatgcttggcacaggcaacgacacgcgcgcagctct
cagccgcttgcagatgacgcagcgcaccggtagctatcagggtctgactaactggacgattgagtcggatgcctggttaggcgaga
taactgcccgtggccatcgtcagttggagcgcgtaccaaaagtgtgcccaggctcaagcggcactggagattagtgatgttactct
tctcaagagcttgacgcttagcggcgtcagtaggttagtacaataatctcgcggccctgcgcgactggttccgaaggtatacaaaaagg
gcacatgtccatgcaagcccgtctctcgaatcgcggctggtgctgaaaagccgagatcgagcaggtcggcaaaaagttgcggcaga
accgccaatgaatcagcagcaggcgtccgttttcttgcgagatccgcgaacaatga (SEQ ID NO:11)

Sequence of *Enterococcus faecium* mvaS

atgaacgtcggcattgacaaaattaatttttccgtccaccgtattatctggatattggtcgcactggcccacgcacgcgaagtggaccgaac
aaattacaattggaattggacaggatcagatggctgtgagcaaaaagacgcacgatatgtaacattcgggctagtccgcgaagggaaa
ttttagaacctgaggacttgaagctatagacatggttatagttgtaccgaatcgggcattgacgagagcaaaatccgcggtcgttttac

atcgtttgttgggcgtacaaccttcgctcgcagtttgaattaaagaagcctgttacggggcaaccgcaggcattcagttgccaaactca
 tatacaagcgaaccggagagcaaggtcctgtaattgcaagcgaatagctcggatggtcttcggcaggtggagagcccacacaagg
 cgcaggggcagttgctatgcttctcacggcaaatcccagaatcctgacctcgaaacgacaatctgatgttaacgcaggatatttatgacttc
 tggagaccacttggtcacgcttacctatggtagatggccaccttccaatcaagtctatattgacagtttaagaaggtctggcaagcacattg
 cgaacgcaatcaagcttataatccgactatgccgcgattagtttcatattccgtatacaaaaatgggtaagaaagccctgctcgtgttttgc
 agatgaagtggaaactgaacaggaacgcggtatggcacggtatgaagagtctatcgtatattcacccggatcggcaactgtatacgggat
 cattgtacctggggctgataccttattgaaaacagttctcacctgtcggcgggcgaccggataggattgttagttatgggagtgccgctgt
 cagcgaattttctccggctgttagtggcaggctatgaaaatcaattgaacaaagaggcgcataccagctcctggatcagcgtcagaagc
 tttccatcgaagagtatgagggcattttacagattccttagaaattgatcaggatgcagcgttctcggatgacctgcatattccatccgcgag
 ataaaaaacacgattcggactataaggagagctga (SEQ ID NO:12)

Sequence of *Enterococcus gallinarum* EG2 mvaS

atggaagaaagtgtcatcattgacgcactgcgtactccaataggaaagtaccacggctcgtgaaagattacacagctgttgaactggggac
 agtagcagcaaaggcgttctggcacgaaatcagcaagcaaaagaacacatagcgaagttattattggcaacgtcctgcaagccggaa
 gtgggcagaatccagccgacaagtcagttacagtcaggattgtctctgatatccccgctagcacgatcaatgaaagtgtggctcgggt
 atgaaagcattctgatgggtatggagcaaatcagctgaacaaagcctctgtggtcttaacaggcgaattgaaagcatgaccaacgcgc
 cgctgttagttattacaacaaggctgaggatcaatattcggcgggtagcacaatgatgcacgatggtctaacagatgctttcagttccaa
 accaatgggcttaaccgcagagaccgtcgtgagagatatggaattacgcgtaaggaaacaagatgaaattgcttatcacttcaaatgaagg
 cggccaaagcccaggcggcgaaaaagtgtgatcaggaaattgtaccctgacggaaaaatccggaacggttctccaggacgaaggcatc
 agagccgcgacaacagtcgagaagctagctgagcttaaacgggtgtcaaaaaagacggaacagttacagcgggtaacgcctctacgat
 aatgatggcgtgctatggtattaatagcatcaaatcttattgcaagaacaccagattccttatctggccgttataaaggagatcgttag
 gtgggtttgccccgaaataatgggtatttccccattaaggtatagacaccctgctgaaaaatcaagcactgacctagaggatatagga
 atattgagattaatgaagccttctgctgcgagttcgttggtagaacgcgagttggcctggacccccaaaaagttaatcgtatggcgggt
 ggtatatcactcggccacgcaattggggcgacgggagctcgcattgcgacgaccgttcttatcagctgaaagataccaggagcgtac
 ggtatagcttcttatgcgttgggtgggggtcttgattggcgtatgcttctggaaaaccatcggccactgcctcacaactaattttgatagg
 aatctgctccgaaaaactgagaagaagaagttttatgcgctagctcctaacgaacgcttagcgtttttggaagcccaaggcgtattaccg
 ctgctgaaaccctggtctccaggagatgacctaaacaagagacagccaatcacttaatcgaaaaccaatcagcgaagttgaaattcctt
 taggcgtggcctgaacttacaggtgaatgggaaagcgtataatgttctctggccacggaggaaccgtccgttatcgtcgcgatgtcgaat
 ggcgcaaaaatggctgtcctattacaacaacaagtcaggagaggtgttacggggtcagattgttctcatggacgtacaggaccagaa
 gcaatattagcgaagttgaatccgagcaagctaccatttctcggtggcaaatgaaacataccctctatcgtgaaaagaggaggaggtct
 gcgtagagtcattggcaggaatttcagtcggccgaaagtgacttagccacggcgtatgtatcaattgacctgatgtagatgtaaggatgc
 aatgggtgctaataatcatcaatagatcctagaaggtgttgcggaattgttagaaaatggttccagaagaagaatctgttctcaattctctc
 caatctcgcgacagaaagtctgtaacggcgacgtgctcagttccgtttgataaattgtccaaaactgggaatgctcgacaagtagctgta

aaatagtcacgcggcgactttgctaagatagatccatacagagctgccacacacaataaaggattatgaatggcgttgaagcgtaatct
tagccaccggaatgacacccgtgcggtgctggctcatgccacggttacgcggcacgcaatgggcgaatgaagggcttacctcttga
cgattatcgaagatcgctgatagctctatcacattacctttggctattgcgacagtgggggtgccacaaaaatcttgcacaaagcacag
gccgccctggcgtaactggcgttgagacggcgtcggaaactggccagcctggcggcgagtgtgggattagtcaaaattggccgcttac
gagcactagtgagcgaggcattcagcaagggcacatgagtatgcaagctagatccctggccattagcgtaggtgcgaaaggtactgaa
atagagcaactagctgcgaagctgagggcagcgcgcaaatgaatcaggagcaggctcgtaaattctgaccgaaataagaaattaa
(SEQ ID NO:13)

Sequence of *Enterococcus casseliflavus mvaS*

atgaacgttgggaattgataaaatcaatTTTTTcgttccgcctatttcattgatatggtgatctcgctcatgcaagagaagttgacccaacaag
ttactataggaataggccaagatcagatggcagtaacaagaacgcagatatacgtacgttcgcgatgcacgccggaaggatattc
tgactaaggaagatttacaggccatagatatgtaaatagtggggactgagctggggtcgcagagcaaggcaagtgtctgtattgcat
cggcttttaggtattcagcctttgcgcgctccttgaattaaggaggcatgctatggggccactgccggccttcagttgcaaaagctcatgt
gcaggctaatacccagagcaaggtcctggtgtagcttccgatatagcacgctacggactggcatccggaggagaaccgactcaaggtgt
aggtgctgtggcaatgttgatttccgctgatccagctatctgcagttagaaaatgataatctcatgttgaccaagatatatacatttttggcg
cccggctgggcatcaatacctatggtagacggccatctgtctaatacgcgtctatagacagctttaaacaagtctggcaagcacattgcga
gaaaaaccaacggactgctaaagattatgctgcattgtcgttccatattccgtacacgaaaatgggtaagaaagctctgttagcggTTTTTgcg
gaggaagatgagacagaacaaaagcggtaaatggcagttatgaagaatcaattgtatacagctcgtcggactggaaatctgtatactggctc
actctatctgggcctgatttcttactggagaatagtagcagtttacaggcgaacgatcgcataaggtctgttttagctatggttcaggggccgtt
cggaaTTTTTcagtgccctcttgggtaccgggttacgagaacaattagcgaagctgccatcaagctctcttgacgaccggcaaaaactg
actatcgagagtagcaagccatgtttaatgaaaccattgatattgatcaggaccagtcatttgaggatgacttactgtactccatcagagaga
tcaaaaacactattcgtactataacgaggagaatgaataa (SEQ ID NO:14)

[0178] The *mvaS* nucleic acid can be expressed in a recombinant cell on a multicopy plasmid. The plasmid can be a high copy plasmid, a low copy plasmid, or a medium copy plasmid. Alternatively, the *mvaS* nucleic acid can be integrated into the host cell's chromosome. For both heterologous expression of an *mvaS* nucleic acid on a plasmid or as an integrated part of the host cell's chromosome, expression of the nucleic acid can be driven by either an inducible promoter or a constitutively expressing promoter. The promoter can be a strong driver of expression, it can be a weak driver of expression, or it can be a medium driver of expression of the *mvaS* nucleic acid.

[0179] Compositions of recombinant cells as described herein are contemplated within the scope of the invention as well. It is understood that recombinant cells also encompass progeny cells as well.

Nucleic acids encoding polypeptides of the lower MVA pathway

[0180] In some aspects of the invention, the cells described in any of the compositions or methods described herein further comprise one or more nucleic acids encoding a lower mevalonate (MVA) pathway polypeptide(s). In some aspects, the lower MVA pathway polypeptide is an endogenous polypeptide. In some aspects, the endogenous nucleic acid encoding a lower MVA pathway polypeptide is operably linked to a constitutive promoter. In some aspects, the endogenous nucleic acid encoding a lower MVA pathway polypeptide is operably linked to an inducible promoter. In some aspects, the endogenous nucleic acid encoding a lower MVA pathway polypeptide is operably linked to a strong promoter. In a particular aspect, the cells are engineered to over-express the endogenous lower MVA pathway polypeptide relative to wild-type cells. In some aspects, the endogenous nucleic acid encoding a lower MVA pathway polypeptide is operably linked to a weak promoter.

[0181] The lower mevalonate biosynthetic pathway comprises mevalonate kinase (MVK), phosphomevalonate kinase (PMK), and diphosphomevalonate decarboxylase (MVD). In some aspects, the lower MVA pathway can further comprise isopentenyl diphosphate isomerase (IDI). Cells provided herein can comprise at least one nucleic acid encoding isoprene synthase, one or more upper MVA pathway polypeptides, and/or one or more lower MVA pathway polypeptides. Polypeptides of the lower MVA pathway can be any enzyme (a) that phosphorylates mevalonate to mevalonate 5-phosphate; (b) that converts mevalonate 5-phosphate to mevalonate 5-pyrophosphate; and (c) that converts mevalonate 5-pyrophosphate to isopentenyl pyrophosphate. More particularly, the enzyme that phosphorylates mevalonate to mevalonate 5-phosphate can be from the group consisting of *M. mazei* mevalonate kinase, *Lactobacillus* mevalonate kinase polypeptide, *Lactobacillus sakei* mevalonate kinase polypeptide, yeast mevalonate kinase polypeptide, *Saccharomyces cerevisiae* mevalonate kinase polypeptide, *Streptococcus* mevalonate kinase polypeptide, *Streptococcus pneumoniae* mevalonate kinase polypeptide, *Streptomyces* mevalonate kinase polypeptide, and *Streptomyces CL190* mevalonate kinase polypeptide. In another aspect, the enzyme that phosphorylates mevalonate to mevalonate 5-phosphate is *M. mazei* mevalonate kinase.

[0182] In some aspects, the lower MVA pathway polypeptide is a heterologous polypeptide. In some aspects, the cells comprise more than one copy of a heterologous nucleic acid encoding a lower MVA pathway polypeptide. In some aspects, the heterologous nucleic acid encoding a lower MVA pathway polypeptide is operably linked to a constitutive promoter. In some aspects, the heterologous nucleic acid encoding a lower MVA pathway polypeptide is operably linked to an inducible promoter. In some aspects, the heterologous nucleic acid encoding a lower MVA pathway polypeptide is operably linked to a strong promoter. In some aspects, the heterologous nucleic acid encoding a lower MVA pathway polypeptide is operably linked to a weak promoter. In some aspects, the heterologous lower MVA pathway polypeptide is a polypeptide from *Saccharomyces cerevisiae*, *Enterococcus faecalis*, or *Methanosarcina mazei*.

[0183] The nucleic acids encoding a lower MVA pathway polypeptide(s) can be integrated into a genome of the cells or can be stably expressed in the cells. The nucleic acids encoding a lower MVA pathway polypeptide(s) can additionally be on a vector.

[0184] Exemplary lower MVA pathway polypeptides are also provided below: (i) mevalonate kinase (MVK); (ii) phosphomevalonate kinase (PMK); (iii) diphosphomevalonate decarboxylase (MVD); and (iv) isopentenyl diphosphate isomerase (IDI). In particular, the lower MVK polypeptide can be from the genus *Methanosarcina* and, more specifically, the lower MVK polypeptide can be from *Methanosarcina mazei*. Additional examples of lower MVA pathway polypeptides can be found in U.S. Patent Application Publication 2010/0086978 the contents of which are expressly incorporated herein by reference in their entirety with respect to lower MVK pathway polypeptides and lower MVK pathway polypeptide variants.

[0185] Lower MVA pathway polypeptides include polypeptides, fragments of polypeptides, peptides, and fusions polypeptides that have at least one activity of a lower MVA pathway polypeptide. Exemplary lower MVA pathway nucleic acids include nucleic acids that encode a polypeptide, fragment of a polypeptide, peptide, or fusion polypeptide that has at least one activity of a lower MVA pathway polypeptide. Exemplary lower MVA pathway polypeptides and nucleic acids include naturally-occurring polypeptides and nucleic acids from any of the source organisms described herein. In addition, variants of lower MVA pathway polypeptides that confer the result of better isoprene production can also be used as well.

[0186] In some aspects, the lower MVA pathway polypeptide is a polypeptide from *Saccharomyces cerevisiae*, *Enterococcus faecalis*, or *Methanosarcina mazei*. In some aspects, the MVK polypeptide is selected from the group consisting of *Lactobacillus* mevalonate kinase polypeptide, *Lactobacillus sakei* mevalonate kinase polypeptide, yeast mevalonate kinase polypeptide, *Saccharomyces cerevisiae* mevalonate kinase polypeptide, *Streptococcus* mevalonate kinase polypeptide, *Streptococcus pneumoniae* mevalonate kinase polypeptide, *Streptomyces* mevalonate kinase polypeptide, *Streptomyces* CL190 mevalonate kinase polypeptide, *Methanosarcina mazei* mevalonate kinase polypeptide, and *M. burtonii* mevalonate kinase polypeptide. Any one of the promoters described herein (*e.g.*, promoters described herein and identified in the Examples of the present disclosure including inducible promoters and constitutive promoters) can be used to drive expression of any of the MVA polypeptides described herein.

[0187] Any one of the cells described herein can comprise IDI nucleic acid(s) (*e.g.*, endogenous or heterologous nucleic acid(s) encoding IDI). Isopentenyl diphosphate isomerase polypeptides (isopentenyl-diphosphate delta-isomerase or IDI) catalyzes the interconversion of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) (*e.g.*, converting IPP into DMAPP and/or converting DMAPP into IPP). Exemplary IDI polypeptides include polypeptides, fragments of polypeptides, peptides, and fusions polypeptides that have at least one activity of an IDI polypeptide. Standard methods (such as those described herein) can be used to determine whether a polypeptide has IDI polypeptide activity by measuring the ability of the polypeptide to interconvert IPP and DMAPP *in vitro*, in a cell extract, or *in vivo*. Exemplary IDI nucleic acids include nucleic acids that encode a polypeptide, fragment of a polypeptide, peptide, or fusion polypeptide that has at least one activity of an IDI polypeptide. Exemplary IDI polypeptides and nucleic acids include naturally-occurring polypeptides and nucleic acids from any of the source organisms described herein as well as mutant polypeptides and nucleic acids derived from any of the source organisms described herein.

DXP pathway nucleic acids and polypeptides

[0188] In some aspects of the invention, the recombinant cells described in any of the compositions or methods described herein (including host cells that have been modified as described herein) further comprise one or more heterologous nucleic acids encoding a DXS polypeptide and/or other DXP pathway polypeptides. In some aspects, the cells further comprise

a chromosomal copy of an endogenous nucleic acid encoding a DXS polypeptide and/or other DXP pathway polypeptides. In some aspects, the *E. coli* cells further comprise one or more nucleic acids encoding an IDI polypeptide and a DXS polypeptide and/or other DXP pathway polypeptides. In some aspects, one nucleic acid encodes the isoprene synthase polypeptide, IDI polypeptide, and DXS polypeptide and/or other DXP pathway polypeptides. In some aspects, one plasmid encodes the isoprene synthase polypeptide, IDI polypeptide, and DXS polypeptide and/or other DXP pathway polypeptides. In some aspects, multiple plasmids encode the isoprene synthase polypeptide, IDI polypeptide, and DXS polypeptide or other DXP pathway polypeptides.

[0189] Exemplary DXS polypeptides include polypeptides, fragments of polypeptides, peptides, and fusions polypeptides that have at least one activity of a DXS polypeptide. Standard methods (such as those described herein) can be used to determine whether a polypeptide has DXS polypeptide activity by measuring the ability of the polypeptide to convert pyruvate and D-glyceraldehyde-3-phosphate into 1-deoxy-D-xylulose-5-phosphate *in vitro*, in a cell extract, or *in vivo*. Exemplary DXS polypeptides and nucleic acids and methods of measuring DXS activity are described in more detail in International Publication No. WO 2009/076676, U.S. Patent Application No. 12/335,071 (US Publ. No. 2009/0203102), WO 2010/003007, US Publ. No. 2010/0048964, WO 2009/132220, and US Publ. No. 2010/0003716.

[0190] Exemplary DXP pathway polypeptides include, but are not limited to any of the following polypeptides: DXS polypeptides, DXR polypeptides, MCT polypeptides, CMK polypeptides, MCS polypeptides, HDS polypeptides, HDR polypeptides, and polypeptides (*e.g.*, fusion polypeptides) having an activity of one, two, or more of the DXP pathway polypeptides. In particular, DXP pathway polypeptides include polypeptides, fragments of polypeptides, peptides, and fusions polypeptides that have at least one activity of a DXP pathway polypeptide. Exemplary DXP pathway nucleic acids include nucleic acids that encode a polypeptide, fragment of a polypeptide, peptide, or fusion polypeptide that has at least one activity of a DXP pathway polypeptide. Exemplary DXP pathway polypeptides and nucleic acids include naturally-occurring polypeptides and nucleic acids from any of the source organisms described herein as well as mutant polypeptides and nucleic acids derived from any of the source organisms described herein. Exemplary DXP pathway polypeptides and nucleic acids and

methods of measuring DXP pathway polypeptide activity are described in more detail in International Publication No.: WO 2010/148150.

[0191] Exemplary DXS polypeptides include polypeptides, fragments of polypeptides, peptides, and fusions polypeptides that have at least one activity of a DXS polypeptide. Standard methods (such as those described herein) can be used to determine whether a polypeptide has DXS polypeptide activity by measuring the ability of the polypeptide to convert pyruvate and D-glyceraldehyde-3-phosphate into 1-deoxy-D-xylulose-5-phosphate *in vitro*, in a cell extract, or *in vivo*. Exemplary DXS polypeptides and nucleic acids and methods of measuring DXS activity are described in more detail in International Publication No. WO 2009/076676, U.S. Patent Application No. 12/335,071 (US Publ. No. 2009/0203102), WO 2010/003007, US Publ. No. 2010/0048964, WO 2009/132220, and US Publ. No. 2010/0003716.

[0192] In particular, DXS polypeptides convert pyruvate and D-glyceraldehyde 3-phosphate into 1-deoxy-D-xylulose 5-phosphate (DXP). Standard methods can be used to determine whether a polypeptide has DXS polypeptide activity by measuring the ability of the polypeptide to convert pyruvate and D-glyceraldehyde 3-phosphate *in vitro*, in a cell extract, or *in vivo*.

[0193] DXR polypeptides convert 1-deoxy-D-xylulose 5-phosphate (DXP) into 2-C-methyl-D-erythritol 4-phosphate (MEP). Standard methods can be used to determine whether a polypeptide has DXR polypeptides activity by measuring the ability of the polypeptide to convert DXP *in vitro*, in a cell extract, or *in vivo*.

[0194] MCT polypeptides convert 2-C-methyl-D-erythritol 4-phosphate (MEP) into 4-(cytidine 5'-diphospho)-2-methyl-D-erythritol (CDP-ME). Standard methods can be used to determine whether a polypeptide has MCT polypeptides activity by measuring the ability of the polypeptide to convert MEP *in vitro*, in a cell extract, or *in vivo*.

[0195] CMK polypeptides convert 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol (CDP-ME) into 2-phospho-4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol (CDP-MEP). Standard methods can be used to determine whether a polypeptide has CMK polypeptides activity by measuring the ability of the polypeptide to convert CDP-ME *in vitro*, in a cell extract, or *in vivo*.

[0196] MCS polypeptides convert 2-phospho-4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol (CDP-MEP) into 2-C-methyl-D-erythritol 2, 4-cyclodiphosphate (ME-CPP or

cMEPP). Standard methods can be used to determine whether a polypeptide has MCS polypeptides activity by measuring the ability of the polypeptide to convert CDP-MEP *in vitro*, in a cell extract, or *in vivo*.

[0197] HDS polypeptides convert 2-C-methyl-D-erythritol 2, 4-cyclodiphosphate into (E)-4-hydroxy-3-methylbut-2-en-1-yl diphosphate (HMBPP or HDMAPP). Standard methods can be used to determine whether a polypeptide has HDS polypeptides activity by measuring the ability of the polypeptide to convert ME-CPP *in vitro*, in a cell extract, or *in vivo*.

[0198] HDR polypeptides convert (E)-4-hydroxy-3-methylbut-2-en-1-yl diphosphate into isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). In one embodiment, the ispH gene can be used to encode for HDR polypeptides. IspH is also known as 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate reductase, 4Fe-4S protein, ECK0030, JW0027, lytB, yaaE, and b0029. Standard methods can be used to determine whether a polypeptide has HDR polypeptides activity by measuring the ability of the polypeptide to convert HMBPP *in vitro*, in a cell extract, or *in vivo*.

Source organisms for isoprene synthase, MVA pathway, DXP pathway and IDI polypeptides

[0199] Isoprene synthase, IDI, DXP pathway, and/or MVA pathway nucleic acids (and their encoded polypeptides) can be obtained from any organism that naturally contains isoprene synthase, IDI, DXP pathway, and/or MVA pathway nucleic acids. Isoprene is formed naturally by a variety of organisms, such as bacteria, yeast, plants, and animals. Some organisms contain the MVA pathway for producing isoprene. Isoprene synthase nucleic acids can be obtained, *e.g.*, from any organism that contains an isoprene synthase. MVA pathway nucleic acids can be obtained, *e.g.*, from any organism that contains the MVA pathway. IDI and DXP pathway nucleic acids can be obtained, *e.g.*, from any organism that contains the IDI and DXP pathway.

[0200] The nucleic acid sequence of the isoprene synthase, DXP pathway, IDI, and/or MVA pathway nucleic acids can be isolated from a bacterium, fungus, plant, algae, or cyanobacterium. Exemplary source organisms include, for example, yeasts, such as species of *Saccharomyces* (*e.g.*, *S. cerevisiae*), bacteria, such as species of *Escherichia* (*e.g.*, *E. coli*), or species of *Methanosarcina* (*e.g.*, *Methanosarcina mazei*), *M. burtonii*, plants, such as kudzu or poplar (*e.g.*,

Populus alba or *Populus alba x tremula* CAC35696) or aspen (e.g., *Populus tremuloides*).

Exemplary sources for isoprene synthases, IDI, and/or MVA pathway polypeptides which can be used are also described in International Patent Application Publication Nos. WO 2009/076676, WO 2010/003007, WO 2009/132220, WO 2010/031062, WO 2010/031068, WO 2010/031076, WO 2010/013077, WO 2010/031079, WO 2010/148150, WO 2010/078457, and WO 2010/148256.

[0201] In some aspects, the source organism is a yeast, such as *Saccharomyces sp.*, *Schizosaccharomyces sp.*, *Pichia sp.*, or *Candida sp.*

[0202] In some aspects, the source organism is a bacterium, such as strains of *Bacillus* such as *B. licheniformis* or *B. subtilis*, strains of *Pantoea* such as *P. citrea*, strains of *Pseudomonas* such as *P. alcaligenes*, strains of *Streptomyces* such as *S. lividans* or *S. rubiginosus*, strains of *Escherichia* such as *E. coli*, strains of *Enterobacter*, strains of *Streptococcus*, or strains of *Archaea* such as *Methanosarcina mazei*.

[0203] As used herein, “the genus *Bacillus*” includes all species within the genus “*Bacillus*,” as known to those of skill in the art, including but not limited to *B. subtilis*, *B. licheniformis*, *B. lentus*, *B. brevis*, *B. stearothermophilus*, *B. alkalophilus*, *B. amyloliquefaciens*, *B. clausii*, *B. halodurans*, *B. megaterium*, *B. coagulans*, *B. circulans*, *B. lautus*, and *B. thuringiensis*. It is recognized that the genus *Bacillus* continues to undergo taxonomical reorganization. Thus, it is intended that the genus include species that have been reclassified, including but not limited to such organisms as *B. stearothermophilus*, which is now named “*Geobacillus stearothermophilus*.” The production of resistant endospores in the presence of oxygen is considered the defining feature of the genus *Bacillus*, although this characteristic also applies to the recently named *Alicyclobacillus*, *Amphibacillus*, *Aneurinibacillus*, *Anoxybacillus*, *Brevibacillus*, *Filobacillus*, *Gracilibacillus*, *Halobacillus*, *Paenibacillus*, *Salibacillus*, *Thermobacillus*, *Ureibacillus*, and *Virgibacillus*.

[0204] In some aspects, the source organism is a gram-positive bacterium. Non-limiting examples include strains of *Streptomyces* (e.g., *S. lividans*, *S. coelicolor*, or *S. griseus*) and *Bacillus*. In some aspects, the source organism is a gram-negative bacterium, such as *E. coli* or *Pseudomonas sp.* In some aspects, the source organism is *L. acidophilus*.

[0205] In some aspects, the source organism is a plant, such as a plant from the family Fabaceae, such as the Faboideae subfamily. In some aspects, the source organism is kudzu, poplar (such as *Populus alba x tremula* CAC35696), aspen (such as *Populus tremuloides*), or *Quercus robur*.

[0206] In some aspects, the source organism is an algae, such as a green algae, red algae, glaucophytes, chlorarachniophytes, euglenids, chromista, or dinoflagellates.

[0207] In some aspects, the source organism is a cyanobacteria, such as cyanobacteria classified into any of the following groups based on morphology: *Chroococcales*, *Pleurocapsales*, *Oscillatoriales*, *Nostocales*, or *Stigonematales*.

Phosphoketolase nucleic acids and polypeptides

[0208] Phosphoketolase enzymes catalyze the conversion of xylulose 5-phosphate to glyceraldehyde 3-phosphate and acetyl phosphate and/or the conversion of fructose 6-phosphate to erythrose 4-phosphate and acetyl phosphate. In certain embodiments, the phosphoketolase enzyme is capable of catalyzing the conversion of xylulose 5-phosphate to glyceraldehyde 3-phosphate and acetyl phosphate. In other embodiments, the phosphoketolase enzyme is capable of catalyzing the conversion of fructose 6-phosphate to erythrose 4-phosphate and acetyl phosphate. In further embodiments, the phosphoketolase polypeptide catalyzes the conversion of sedoheptulose-7-phosphate to a product (e.g., ribose-5-phosphate) and acetyl phosphate. Thus, without being bound by theory, the expression of phosphoketolase as set forth herein can result in an increase in the amount of acetyl phosphate produced from a carbohydrate or carbon source. This acetyl phosphate can be converted into acetyl-CoA which can then be utilized by the enzymatic activities of the MVA pathway to produce isoprene. Thus the amount of these compounds produced from a carbohydrate substrate may be increased. Alternatively, production of Acetyl-P and AcCoA can be increased without the increase being reflected in higher intracellular concentration. In certain embodiments, intracellular acetyl-P or acetyl-CoA concentrations will remain unchanged or even decrease, even though the phosphoketolase reaction is taking place.

[0209] In some aspects of the invention, the recombinant cells described in any of the compositions or methods described herein further comprise one or more nucleic acids encoding

a phosphoketolase polypeptide or a polypeptide having phosphoketolase activity. In some aspects, the phosphoketolase polypeptide is an endogenous polypeptide. In some aspects, the endogenous nucleic acid encoding a phosphoketolase polypeptide is operably linked to a constitutive promoter. In some aspects, the endogenous nucleic acid encoding a phosphoketolase polypeptide is operably linked to an inducible promoter. In some aspects, the endogenous nucleic acid encoding a phosphoketolase polypeptide is operably linked to a strong promoter. In some aspects, more than one endogenous nucleic acid encoding a phosphoketolase polypeptide is used (e.g, 2, 3, 4, or more copies of an endogenous nucleic acid encoding a phosphoketolase polypeptide). In a particular aspect, the cells are engineered to overexpress the endogenous phosphoketolase polypeptide relative to wild-type cells. In some aspects, the endogenous nucleic acid encoding a phosphoketolase polypeptide is operably linked to a weak promoter.

[0210] Exemplary phosphoketolase nucleic acids include nucleic acids that encode a polypeptide, fragment of a polypeptide, peptide, or fusion polypeptide that has at least one activity of a phosphoketolase polypeptide. Exemplary phosphoketolase polypeptides and nucleic acids include naturally-occurring polypeptides and nucleic acids from any of the source organisms described herein as well as mutant polypeptides and nucleic acids derived from any of the source organisms described herein. In some aspects, a nucleic acid encoding a phosphoketolase is from *Clostridium acetobutylicum*, *Lactobacillus reuteri*, *Lactobacillus plantarum*, *Lactobacillus paraplantarum*, *Bifidobacterium longum*, *Bifidobacterium animalis*, *Bifidobacterium breve*, *Enterococcus gallinarum*, *Gardnerella vaginalis*, *Ferrimonas balearica*, *Mucilaginibacter paludis*, *Nostoc punctiforme*, *Nostoc punctiforme PCC 73102*, *Pantoea*, *Pedobactor saltans*, *Rahnella aquatilis*, *Rhodopseudomonas palustris*, *Streptomyces griseus*, *Streptomyces avermitilis*, *Nocardiosis dassonvillei*, and/or *Thermobifida fusca*. Additional examples of phosphoketolase enzymes which can be used herein are described in U.S. 7,785,858 and International Patent Application Publication No. WO 2011/159853 which are incorporated by reference herein.

[0211] Standard methods can be used to determine whether a polypeptide has phosphoketolase peptide activity by measuring the ability of the peptide to convert D-fructose 6-phosphate or D-xylulose 5-phosphate into acetyl-P. Acetyl-P can then be converted into ferryl acetyl hydroxamate, which can be detected spectrophotometrically (Meile et al., J. Bact.

183:2929-2936, (2001)). Any polypeptide identified as having phosphoketolase peptide activity as described herein is suitable for use in the present invention.

[0212] In any of the embodiments herein, the recombinant cells can be further engineered to increase the activity of one or more of the following genes selected from the group consisting of ribose-5-phosphate isomerase (*rpiA* and/or *rpiB*), D-ribulose-5-phosphate 3-epimerase (*rpe*), transketolase (*tktA* and/or *tktB*), transaldolase B (*tal B*), phosphate acetyltransferase (*pta* and/or *eutD*). In another embodiment, the recombinant cells can be further engineered to decrease the activity of one or more genes of the following genes including glucose-6-phosphate dehydrogenase (*zwf*), 6-phosphofructokinase-1 (*pfkA* and/or *pfkB*), fructose bisphosphate aldolase (*fba*, *fbaA*, *fbaB*, and/or *fbaC*), glyceraldehyde-3-phosphate dehydrogenase (*gapA* and/or *gapB*), acetate kinase (*ackA*), citrate synthase (*gltA*), EI (*ptsI*), EIICB^{Glc} (*ptsG*), EIIA^{Glc} (*crr*), and/or HPr (*ptsH*).

Recombinant Cells Capable of Producing Isoprene

[0213] Isoprene (2-methyl-1,3-butadiene) is an important organic compound used in a wide array of applications. For instance, isoprene is employed as an intermediate or a starting material in the synthesis of numerous chemical compositions and polymers, including in the production of synthetic rubber. Isoprene is also an important biological material that is synthesized naturally by many plants and animals.

[0214] As described herein, the present invention provides recombinant cells capable of producing of isoprene, wherein the cells comprise one or more nucleic acids encoding one or more polypeptides of the MVA pathway, and a heterologous nucleic acid encoding an isoprene synthase polypeptide, wherein culturing the cells in a suitable media provides for the production of isoprene. In a further embodiment, the recombinant cells further comprise one or more nucleic acids encoding an isopentenyl diphosphate isomerase (IDI) polypeptide. In certain embodiments, the present invention provides recombinant cells capable of isoprene production, wherein the cells comprise one or more nucleic acids encoding one or more polypeptides of the MVA pathway, and a heterologous nucleic acid encoding an isoprene synthase polypeptide, wherein the cells produce increased amounts of isoprene compared to isoprene-producing cells that do not comprise a nucleic acid encoding a polypeptide having phosphomevalonate decarboxylase activity and/or a nucleic acid encoding a polypeptide having isopentenyl kinase

activity. In a further embodiment, the recombinant cells further comprise one or more nucleic acids encoding an isopentenyl diphosphate isomerase (IDI) polypeptide.

[0215] Production of isoprene can also be made by using any of the recombinant host cells described herein further comprising one or more of the enzymatic pathways manipulations wherein enzyme activity is modulated to increase carbon flow towards mevalonate production. The recombinant cells described herein that have various enzymatic pathways manipulated for increased carbon flow to mevalonate production can be used to produce isoprene. In one embodiment, the recombinant cells further comprise a nucleic acid encoding a phosphoketolase. In another embodiment, the recombinant cells can be further engineered to increase the activity of one or more of the following genes selected from the group consisting of ribose-5-phosphate isomerase (*rpiA* and/or *rpiB*), D-ribulose-5-phosphate 3-epimerase (*rpe*), transketolase (*tktA* and/or *tktB*), transaldolase B (*tal B*), phosphate acetyltransferase (*pta* and/or *eutD*). In another embodiment, these recombinant cells can be further engineered to decrease the activity of one or more genes of the following genes including glucose-6-phosphate dehydrogenase (*zwf*), 6-phosphofruktokinase-1 (*pfkA* and/or *pfkB*), fructose biphosphate aldolase (*fba*, *fbaA*, *fbaB*, and/or *fbaC*), glyceraldehyde-3-phosphate dehydrogenase (*gapA* and/or *gapB*), acetate kinase (*ackA*), citrate synthase (*glcA*), EI (*ptsI*), EIICB^{Glc} (*ptsG*), EIIA^{Glc} (*crr*), and/or HPr (*ptsH*).

Isoprene synthase nucleic acids and polypeptides

[0216] In some aspects of the invention, the cells described in any of the compositions or methods described herein (including host cells that have been modified as described herein) further comprise one or more nucleic acids encoding an isoprene synthase polypeptide or a polypeptide having isoprene synthase activity. In some aspects, the isoprene synthase polypeptide is an endogenous polypeptide. In some aspects, the endogenous nucleic acid encoding an isoprene synthase polypeptide is operably linked to a constitutive promoter. In some aspects, the endogenous nucleic acid encoding an isoprene synthase polypeptide is operably linked to an inducible promoter. In some aspects, the endogenous nucleic acid encoding an isoprene synthase polypeptide is operably linked to a strong promoter. In a particular aspect, the cells are engineered to overexpress the endogenous isoprene synthase pathway polypeptide relative to wild-type cells. In some aspects, the endogenous nucleic acid encoding an isoprene synthase polypeptide is operably linked to a weak promoter.

[0217] In some aspects, the isoprene synthase polypeptide is a heterologous polypeptide. In some aspects, the cells comprise more than one copy of a heterologous nucleic acid encoding an isoprene synthase polypeptide. In some aspects, the heterologous nucleic acid encoding an isoprene synthase polypeptide is operably linked to a constitutive promoter. In some aspects, the heterologous nucleic acid encoding an isoprene synthase polypeptide is operably linked to an inducible promoter. In some aspects, the heterologous nucleic acid encoding an isoprene synthase polypeptide is operably linked to a strong promoter. In some aspects, the heterologous nucleic acid encoding an isoprene synthase polypeptide is operably linked to a weak promoter. In some aspects, the isoprene synthase polypeptide is a polypeptide or variant thereof from *Pueraria* or *Populus* or a hybrid such as *Populus alba* x *Populus tremula*. In some aspects, the isoprene synthase polypeptide is a polypeptide or variant thereof from *Pueraria montana* or *Pueraria lobata*, *Populus tremuloides*, *Populus alba*, *Populus nigra*, and *Populus trichocarpa*. In some aspects, the isoprene synthase polypeptide is from *Eucalyptus*.

[0218] The nucleic acids encoding an isoprene synthase polypeptide(s) can be integrated into a genome of the host cells or can be stably expressed in the cells. The nucleic acids encoding an isoprene synthase polypeptide(s) can additionally be on a vector.

[0219] Exemplary isoprene synthase nucleic acids include nucleic acids that encode a polypeptide, fragment of a polypeptide, peptide, or fusion polypeptide that has at least one activity of an isoprene synthase polypeptide. Isoprene synthase polypeptides convert dimethylallyl diphosphate (DMAPP) into isoprene. Exemplary isoprene synthase polypeptides include polypeptides, fragments of polypeptides, peptides, and fusions polypeptides that have at least one activity of an isoprene synthase polypeptide. Exemplary isoprene synthase polypeptides and nucleic acids include naturally-occurring polypeptides and nucleic acids from any of the source organisms described herein. In addition, variants of isoprene synthase can possess improved activity such as improved enzymatic activity. In some aspects, an isoprene synthase variant has other improved properties, such as improved stability (*e.g.*, thermo-stability), and/or improved solubility.

Pathways involving the Entner-Doudoroff pathway

[0220] The Entner-Doudoroff (ED) pathway is an alternative to the Emden-Meyerhoff-Parnass (EMP –glycolysis) pathway. Some organisms, like *E. coli*, harbor both the ED and EMP

pathways, while others have only one or the other. *Bacillus subtilis* has only the EMP pathway, while *Zymomonas mobilis* has only the ED pathway (Peekhaus and Conway. J. Bact. 180:3495-3502 (1998); Stulke and Hillen. Annu. Rev. Microbiol. 54, 849–880 (2000); Dawes et al. Biochem. J. 98:795-803 (1966)). Fructose biphosphate aldolase (fba, fbaA, fbaB, and/or fbaC) interacts with the Entner-Doudoroff pathway and reversibly catalyzes the conversion of fructose 1,6-bisphosphate into dihydroxyacetone phosphate (DHAP) and glyceraldehyde 3-phosphate (GAP) (Baldwin S. A., et. al., Biochem J. 169(3):633-41 (1978)).

[0221] Phosphogluconate dehydratase (edd) removes one molecule of H₂O from 6-phospho-D-gluconate to form 2-dehydro-3-deoxy-D-gluconate 6-phosphate, while 2-keto-3-deoxygluconate 6-phosphate aldolase (eda) catalyzes an aldol cleavage (Egan et al. J. Bact. 174:4638-4646 (1992)). The two genes are in an operon.

[0222] Metabolites that can be directed into the phosphoketolase pathway can also be diverted into the ED pathway. To avoid metabolite loss to the ED-pathway, phosphogluconate dehydratase gene (*e.g.*, the endogenous phosphogluconate dehydratase gene) and/or an 2-keto-3-deoxygluconate 6-phosphate aldolase gene (*e.g.*, the endogenous 2-keto-3-deoxygluconate 6-phosphate aldolase gene) activity is attenuated. One way of achieving attenuation is by deleting phosphogluconate dehydratase (edd) and/or 2-keto-3-deoxygluconate 6-phosphate aldolase (eda). This can be accomplished by replacing one or both genes with a chloramphenicol or kanamycin cassette followed by looping out of the cassette. Without these enzymatic activities, more carbon can flux through the phosphoketolase enzyme, thus increasing the yield of isoprene.

[0223] The activity of phosphogluconate dehydratase (edd) and/or 2-keto-3-deoxygluconate 6-phosphate aldolase (eda) can also be decreased by other molecular manipulations of the enzymes. The decrease of enzyme activity can be any amount of reduction of specific activity or total activity as compared to when no manipulation has been effectuated. In some instances, the decrease of enzyme activity is decreased by at least about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%.

[0224] In some cases, attenuating the activity of the endogenous phosphogluconate dehydratase gene and/or the endogenous 2-keto-3-deoxygluconate 6-phosphate aldolase gene results in more carbon flux into the mevalonate dependent biosynthetic pathway in comparison

to cells that do not have attenuated endogenous phosphogluconate dehydratase gene and/or endogenous acetate kinase 2-keto-3-deoxygluconate 6-phosphate aldolase gene expression.

Pathways involving the oxidative branch of the pentose phosphate pathway

[0225] *E. coli* uses the pentose phosphate pathway to break down hexoses and pentoses and to provide cells with intermediates for various anabolic pathways. It is also a major producer of NADPH. The pentose phosphate pathway is composed from an oxidative branch (with enzymes like glucose 6-phosphate 1-dehydrogenase (*zwf*), 6-phosphogluconolactonase (*pgl*) or 6-phosphogluconate dehydrogenase (*gnd*)) and a non-oxidative branch (with enzymes such as transketolase (*tktA*), transaldolase (*talA* or *talB*), ribulose-5-phosphate-epimerase and (or) ribose-5-phosphate epimerase) (Sprenger. Arch. Microbiol.164:324-330 (1995)).

[0226] In order to direct carbon towards the phosphoketolase enzyme, the non-oxidative branch of the pentose phosphate pathway (transketolase, transaldolase, ribulose-5-phosphate-epimerase and (or) ribose-5-phosphate epimerase) expression can be modulated (*e.g.*, increase enzyme activity) to allow more carbon to flux towards fructose 6-phosphate and xylulose 5-phosphate, thereby increasing the eventual production of isoprene. Increase of transketolase, transaldolase, ribulose-5-phosphate-epimerase and (or) ribose-5-phosphate epimerase activity can be any amount of increase of specific activity or total activity as compared to when no manipulation has been effectuated. In some instances, the enzyme activity is increased by at least about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%. In some aspects, the activity of transketolase, transaldolase, ribulose-5-phosphate-epimerase and (or) ribose-5-phosphate epimerase is modulated by increasing the activity of an endogenous transketolase, transaldolase, ribulose-5-phosphate-epimerase and (or) ribose-5-phosphate epimerase. This can be accomplished by replacing the endogenous transketolase, transaldolase, ribulose-5-phosphate-epimerase and (or) ribose-5-phosphate epimerase gene promoter with a synthetic constitutively high expressing promoter. The genes encoding transketolase, transaldolase, ribulose-5-phosphate-epimerase and (or) ribose-5-phosphate epimerase can also be cloned on a plasmid behind an appropriate promoter. The increase of the activity of transketolase, transaldolase, ribulose-5-phosphate-epimerase and (or) ribose-5-phosphate epimerase can result in more carbon flux into the mevalonate dependent biosynthetic

pathway in comparison to cells that do not have increased expression of transketolase, transaldolase, ribulose-5-phosphate-epimerase and (or) ribose-5-phosphate epimerase.

Pathways involving phosphofructokinase

[0227] Phosphofructokinase is a crucial enzyme of glycolysis which catalyzes the phosphorylation of fructose 6-phosphate. *E. coli* has two isozymes encoded by *pfkA* and *pfkB*. Most of the phosphofructokinase activity in the cell is due to *pfkA* (Kotlarz et al. *Biochim. Biophys. Acta* 381:257-268 (1975)).

[0228] In order to direct carbon towards the phosphoketolase enzyme, phosphofructokinase expression can be modulated (*e.g.*, decrease enzyme activity) to allow more carbon to flux towards fructose 6-phosphate and xylulose 5-phosphate, thereby increasing the eventual production of isoprene. Decrease of phosphofructokinase activity can be any amount of reduction of specific activity or total activity as compared to when no manipulation has been effectuated. In some instances, the decrease of enzyme activity is decreased by at least about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%. In some aspects, the activity of phosphofructokinase is modulated by decreasing the activity of an endogenous phosphofructokinase. This can be accomplished by replacing the endogenous phosphofructokinase gene promoter with a synthetic constitutively low expressing promoter. The gene encoding phosphofructokinase can also be deleted. The decrease of the activity of phosphofructokinase can result in more carbon flux into the mevalonate dependent biosynthetic pathway in comparison to cells that do not have decreased expression of phosphofructokinase.

Vectors

[0229] Suitable vectors can be used for any of the compositions and methods described herein. For example, suitable vectors can be used to optimize the expression of one or more copies of a gene encoding an isoprene synthase, an acetoacetyl co-A synthase, an MVA pathway enzyme, a DXP pathway enzyme, a phosphoketolase, and/or a polyprenyl pyrophosphate synthase, in a cell. In some aspects, the vector contains a selective marker. Examples of selectable markers include, but are not limited to, antibiotic resistance nucleic acids (*e.g.*, kanamycin, ampicillin, carbenicillin, gentamicin, hygromycin, phleomycin, bleomycin,

neomycin, or chloramphenicol) and/or nucleic acids that confer a metabolic advantage, such as a nutritional advantage on the host cell. In some aspects, one or more copies of an isoprene synthase, an acetoacetyl co-A synthase, an MVA pathway enzyme, a DXP pathway enzyme, a phosphoketolase, and/or a polyprenyl pyrophosphate synthases nucleic acid(s) integrate into the genome of host cells without a selective marker.

[0230] Any one of the vectors characterized herein or used in the Examples of the present disclosure can be used in the present invention.

Exemplary host cells

[0231] One of skill in the art will recognize that expression vectors are designed to contain certain components which optimize gene expression for certain host strains. Such optimization components include, but are not limited to origin of replication, promoters, and enhancers. The vectors and components referenced herein are described for exemplary purposes and are not meant to narrow the scope of the invention.

[0232] Any microorganism or progeny thereof that can be used to heterologously express genes can be used to express one or more copies of a nucleic acid encoding an isoprene synthase, acetoacetyl co-A synthase, MVA pathway enzyme, DXP pathway enzyme, phosphoketolase, and/or a polyprenyl pyrophosphate synthase polypeptides in a cell. Exemplary host cells include, for example, yeasts, such as species of *Saccharomyces* (e.g., *S. cerevisiae*), bacteria, such as species of *Escherichia* (e.g., *E. coli*), archaea, such as species of *Methanosarcina* (e.g., *Methanosarcina mazei*), plants, such as kudzu or poplar (e.g., *Populus alba* or *Populus alba x tremula* CAC35696) or aspen (e.g., *Populus tremuloides*).

[0233] Bacteria cells, including gram positive or gram negative bacteria can be used to express any of the nucleic acids or polypeptides described above. In some embodiments, the host cell is a gram-positive bacterium. Non-limiting examples include strains of *Streptomyces* (e.g., *S. lividans*, *S. coelicolor*, *S. rubiginosus*, or *S. griseus*), *Streptococcus*, *Bacillus* (e.g., *B. licheniformis* or *B. subtilis*), *Listeria* (e.g., *L. monocytogenes*), *Corynebacteria*, or *Lactobacillus* (e.g., *L. spp*). In some embodiments, the source organism is a gram-negative bacterium. Non-limiting examples include strains of *Escherichia* (e.g., *E. coli*), *Pseudomonas* (e.g., *P. alcaligenes*), *Pantoea* (e.g., *P. citrea*), *Enterobacter*, or *Helicobacter* (*H. pylori*). In particular,

one or more copies of a nucleic acid encoding an isoprene synthase, acetoacetyl co-A synthase, MVA pathway enzyme, DXP pathway enzyme, phosphoketolase, and/or a polyprenyl pyrophosphate synthase polypeptides can be expressed in any one of *P. citrea*, *B. subtilis*, *B. licheniformis*, *B. lentus*, *B. brevis*, *B. stearothermophilus*, *B. alkalophilus*, *B. amyloliquefaciens*, *B. clausii*, *B. halodurans*, *B. megaterium*, *B. coagulans*, *B. circulans*, *B. lautus*, *B. thuringiensis*, *S. albus*, *S. lividans*, *S. coelicolor*, *S. griseus*, *Pseudomonas sp.*, and *P. alcaligenes* cells.

[0234] There are numerous types of anaerobic cells that can be used as host cells in the compositions and methods of the present invention. In one aspect of the invention, the cells described in any of the compositions or methods described herein are obligate anaerobic cells and progeny thereof. Obligate anaerobes typically do not grow well, if at all, in conditions where oxygen is present. It is to be understood that a small amount of oxygen may be present, that is, there is some tolerance level that obligate anaerobes have for a low level of oxygen. In one aspect, obligate anaerobes engineered to produce isoprene can serve as host cells for any of the methods and/or compositions described herein and are grown under substantially oxygen-free conditions, wherein the amount of oxygen present is not harmful to the growth, maintenance, and/or fermentation of the anaerobes.

[0235] In another aspect of the invention, the host cells described and/or used in any of the compositions or methods described herein are facultative anaerobic cells and progeny thereof. Facultative anaerobes can generate cellular ATP by aerobic respiration (*e.g.*, utilization of the TCA cycle) if oxygen is present. However, facultative anaerobes can also grow in the absence of oxygen. This is in contrast to obligate anaerobes which die or grow poorly in the presence of greater amounts of oxygen. In one aspect, therefore, facultative anaerobes can serve as host cells for any of the compositions and/or methods provided herein and can be engineered to produce isoprene. Facultative anaerobic host cells can be grown under substantially oxygen-free conditions, wherein the amount of oxygen present is not harmful to the growth, maintenance, and/or fermentation of the anaerobes, or can be alternatively grown in the presence of greater amounts of oxygen.

[0236] The host cell can additionally be a filamentous fungal cell and progeny thereof. (*See, e.g.*, Berka & Barnett, *Biotechnology Advances*, 7(2):127-154 (1989)). In some aspects, the filamentous fungal cell can be any of *Trichoderma longibrachiatum*, *T. viride*, *T. koningii*, *T. harzianum*, *Penicillium sp.*, *Humicola insolens*, *H. lanuginosa*, *H. grisea*, *Chrysosporium sp.*, *C.*

lucknowense, *Gliocladium sp.*, *Aspergillus sp.*, such as *A. oryzae*, *A. niger*, *A. sojae*, *A. japonicus*, *A. nidulans*, or *A. awamori*, *Fusarium sp.*, such as *F. roseum*, *F. gramineum*, *F. cerealis*, *F. oxysporum*, or *F. venenatum*, *Neurospora sp.*, such as *N. crassa*, *Hypocrea sp.*, *Mucor sp.*, such as *M. miehei*, *Rhizopus sp.* or *Emericella sp.* In some aspects, the fungus is *A. nidulans*, *A. awamori*, *A. oryzae*, *A. aculeatus*, *A. niger*, *A. japonicus*, *T. reesei*, *T. viride*, *F. oxysporum*, or *F. solani*. In certain embodiments, plasmids or plasmid components for use herein include those described in U.S. Patent Pub. No. US 2011/0045563.

[0237] The host cell can also be a yeast, such as *Saccharomyces sp.*, *Schizosaccharomyces sp.*, *Pichia sp.*, or *Candida sp.* In some aspects, the *Saccharomyces sp.* is *Saccharomyces cerevisiae* (See, e.g., Romanos et al., *Yeast*, 8(6):423-488 (1992)). In certain embodiments, plasmids or plasmid components for use herein include those described in U.S. pat. No. 7,659,097 and U.S. Patent Pub. No. US 2011/0045563.

[0238] The host cell can also be a species of plant, such as a plant from the family Fabaceae, such as the Faboideae subfamily. In some aspects, the host cell is kudzu, poplar (such as *Populus alba x tremula* CAC35696), aspen (such as *Populus tremuloides*), or *Quercus robur*.

[0239] The host cell can additionally be a species of algae, such as a green algae, red algae, glaucophytes, chlorarachniophytes, euglenids, chromista, or dinoflagellates. (See, e.g., Saunders & Warmbrodt, "Gene Expression in Algae and Fungi, Including Yeast," (1993), National Agricultural Library, Beltsville, MD). In certain embodiments, plasmids or plasmid components for use herein include those described in U.S. Patent Pub. No. US 2011/0045563. In some aspects, the host cell is a cyanobacterium, such as cyanobacterium classified into any of the following groups based on morphology: *Chlorococcales*, *Pleurocapsales*, *Oscillatoriales*, *Nostocales*, or *Stigonematales* (See, e.g., Lindberg et al., *Metab. Eng.*, 12(1):70-79 (2010)). In certain embodiments, plasmids or plasmid components for use herein include those described in U.S. Patent Pub. No.: US 2010/0297749; US 2009/0282545 and PCT Pat. Appl. No. WO 2011/034863.

[0240] *E. coli* host cells can be used to express one or more isoprene synthase (which would also include variants), acetoacetyl co-A synthase, MVA pathway enzyme, DXP pathway enzyme, phosphoketolase, and/or a polyprenyl pyrophosphate synthase polypeptides in the compositions and methods described herein. In one aspect, the host cell is a recombinant cell of

an *Escherichia coli* (*E. coli*) strain, or progeny thereof, capable of producing isoprene that expresses one or more nucleic acids encoding isoprene synthase variant, acetoacetyl co-A synthase, MVA pathway enzyme, DXP pathway enzyme, phosphoketolase, and/or IDI polypeptides. The *E. coli* host cells can produce isoprene, in amounts, peak titers, and cell productivities greater than that of the same cells lacking one or more heterologously expressed nucleic acids encoding isoprene synthase, acetoacetyl co-A synthase, MVA pathway enzyme, DXP pathway enzyme, phosphoketolase, and/or IDI polypeptides. In addition, the one or more heterologously expressed nucleic acids encoding isoprene synthase, acetoacetyl co-A synthase, MVA pathway enzyme, DXP pathway enzyme, phosphoketolase, and/or IDI polypeptides in *E. coli* can be chromosomal copies (e.g., integrated into the *E. coli* chromosome). In other aspects, the *E. coli* cells are in culture.

[0241] In other aspects, the host cell can be a species of yeast other than *S. cerevisiae* such as, but not limited to, a *Pichia spp.*, a *Candida spp.*, a *Hansenula spp.*, a *Kluyveromyces spp.*, a *Kluyveromyces spp.*, or a *Schizosaccharomyces spp.* In still other aspects, the host cell can be a species of bacterium including, but not limited to, an *Arthrobacter spp.*, a *Zymomonas spp.*, a *Brevibacterium spp.*, a *Clostridium spp.*, an *Aerococcus spp.*, a *Bacillus spp.*, an *Actinobacillus spp.* (such as, but not limited to, *A. succinogens*), a *Carbobacterium spp.*, a *Corynebacterium spp.*, an *Enterococcus spp.*, an *Erysipelothrix spp.*, a *Gemella spp.*, a *Geobacillus spp.*, a *Globicatella spp.*, a *Lactobacillus spp.* (such as, but not limited to, *L. lactis* and *L. rhammosus*), a *Lactococcus spp.*, a *Leuconostoc spp.*, a *Pediococcus spp.*, a *Streptococcus spp.*, a *Tetragenococcus spp.*, an *Actinobacillus spp.*, or a *Vagococcus spp.* In other aspects, the fermenting organism can be a fungus such as, but not limited to, a *Rhizopus spp.*

[0242] In other aspects, the host cell can be a lactic acid bacteria, such as those of the genera *Aerococcus*, *Bacillus*, *Carbobacterium*, *Enterococcus*, *Erysipelothrix*, *Gemella*, *Globicatella*, *Lactobacillus*, *Lactococcus*, *Leuconostoc*, *Pediococcus*, *Streptococcus*, *Tetragenococcus* and *Vagococcus*. For example, other bacteria of the genus *Lactobacillus* which may be substituted include, but are not limited to, *L. heiveticus*, *L. delbrueckii*, *L. casei*, *L. acidophilus*, *L. amylovorus*, *L. leichmanii* or *L. bulgaricus*. *L. amylovorus*, and *L. pentosus*.

Additional Host cell Mutations

[0243] The invention also contemplates additional host cell mutations that increase carbon flux through the MVA pathway. By increasing the carbon flow, more isoprene can be produced. The recombinant cells comprising acetoacetyl-CoA synthase as described herein can also be engineered for increased carbon flux towards mevalonate production wherein the activity of one or more enzymes from the group consisting of: (a) citrate synthase, (b) phosphotransacetylase; (c) acetate kinase; (d) lactate dehydrogenase; (e) NADP-dependent malic enzyme; (f) pyruvate dehydrogenase; (g) 6-phosphogluconolactonase; (h) phosphoenolpyruvate carboxylase; (i) the inhibitor of RssB activity during magnesium starvation protein; (j) the *acrA* component of the multidrug efflux pump *acrAB-TolC*; and (k) the fumarate and nitrate reduction sRNA (FNR) is modulated.

Citrate synthase pathway

[0244] Citrate synthase catalyzes the condensation of oxaloacetate and acetyl-CoA to form citrate, a metabolite of the Tricarboxylic acid (TCA) cycle (Ner, S. et al. *Biochemistry* 22: 5243-5249 (1983); Bhayana, V. and Duckworth, H., *Biochemistry* 23: 2900-2905 (1984)). In *E. coli*, this enzyme, encoded by *gltA*, behaves like a trimer of dimeric subunits. The hexameric form allows the enzyme to be allosterically regulated by NADH. This enzyme has been widely studied (Wiegand, G., and Remington, S. *Annual Rev. Biophysics Biophys. Chem.* 15: 97-117 (1986); Duckworth et al. *Biochem Soc Symp.* 54:83-92 (1987); Stockell, D. et al. *J. Biol. Chem.* 278: 35435-35443 (2003); Maurus, R. et al. *Biochemistry.* 42:5555-5565 (2003)). To avoid allosteric inhibition by NADH, replacement by or supplementation with the *Bacillus subtilis* NADH-insensitive citrate synthase has been considered (Underwood et al. *Appl. Environ. Microbiol.* 68:1071-1081 (2002); Sanchez et al. *Met. Eng.* 7:229-239 (2005)).

[0245] The reaction catalyzed by citrate synthase is directly competing with the thiolase catalyzing the first step of the mevalonate pathway, as they both have acetyl-CoA as a substrate (Hedl et al. *J. Bact.* 184:2116-2122 (2002)). Therefore, one of skill in the art can modulate citrate synthase expression (*e.g.*, decrease enzyme activity) to allow more carbon to flux into the mevalonate pathway, thereby increasing the eventual production of mevalonate and isoprene. Decrease of citrate synthase activity can be any amount of reduction of specific activity or total activity as compared to when no manipulation has been effectuated. In some instances, the

decrease of enzyme activity is decreased by at least about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% or 100%. In some aspects, the activity of citrate synthase is modulated by decreasing the activity of an endogenous citrate synthase gene. This can be accomplished by chromosomal replacement of an endogenous citrate synthase gene with a transgene encoding an NADH-insensitive citrate synthase or by using a transgene encoding an NADH-insensitive citrate synthase that is derived from *Bacillus subtilis*. The activity of citrate synthase can also be modulated (*e.g.*, decreased) by replacing the endogenous citrate synthase gene promoter with a synthetic constitutively low expressing promoter. The gene encoding citrate synthase can also be deleted. The decrease of the activity of citrate synthase can result in more carbon flux into the mevalonate dependent biosynthetic pathway in comparison to microorganisms that do not have decreased expression of citrate synthase.

Pathways involving phosphotransacetylase and/or acetate kinase

[0246] Phosphotransacetylase (encoded in *E. coli* by (i) (*pta*) (Shimizu et al. Biochim. Biophys. Acta 191: 550-558 (1969)) or (ii) (*eutD*) (Bologna et al. J. Microbiology 48:629-636 (2010)) catalyzes the reversible conversion between acetyl-CoA and acetylphosphate (acetyl-P), while acetate kinase (encoded in *E. coli* by *ackA*) (Kakuda, H. et al. J. Biochem. 11:916-922 (1994)) uses acetyl-P to form acetate. These genes can be transcribed as an operon in *E. coli*. Together, they catalyze the dissimilation of acetate, with the release of ATP. Thus, one of skill in the art can increase the amount of available acetyl Co-A by attenuating the activity of phosphotransacetylase gene (*e.g.*, the endogenous phosphotransacetylase gene) and/or an acetate kinase gene (*e.g.*, the endogenous acetate kinase gene). In certain embodiments, enhancement is achieved by placing an upregulated promoter upstream of the gene in the chromosome, or to place a copy of the gene behind an adequate promoter on a plasmid. One way of achieving attenuation is by deleting phosphotransacetylase (*pta*) and/or acetate kinase (*ackA*). This can be accomplished by replacing one or both genes with a chloramphenicol cassette followed by looping out of the cassette. In some aspects, the activity of acetate kinase is modulated by decreasing the activity of an endogenous acetate kinase. This can be accomplished by replacing the endogenous acetate kinase gene promoter with a synthetic constitutively low expressing promoter. In certain embodiments, the attenuation of the acetate kinase gene disrupts the expression of the phosphotransacetylase (*pta*) gene. Acetate is produced by *E. coli* for a variety

of reasons (Wolfe, A. *Microb. Mol. Biol. Rev.* 69:12-50 (2005)). Without being bound by theory, since ackA-pta use acetyl-CoA, deleting those genes might allow carbon not to be diverted into acetate and to increase the yield of mevalonate and/or isoprene.

[0247] In some aspects, the recombinant microorganism produces decreased amounts of acetate in comparison to microorganisms that do not have attenuated endogenous phosphotransacetylase gene and/or endogenous acetate kinase gene expression. Decrease in the amount of acetate produced can be measured by routine assays known to one of skill in the art. The amount of acetate reduction is at least about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% as compared when no molecular manipulations are done.

[0248] The activity of phosphotransacetylase (pta and/or eutD) can be increased by other molecular manipulations of the enzymes. The increase of enzyme activity can be and increase in any amount of specific activity or total activity as compared to when no manipulation has been effectuated. In some instances, the increase of enzyme activity is increased by at least about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%. In one embodiment the activity of pta is increased by altering the promoter and/or rbs on the chromosome, or by expressing it from a plasmid. In any aspects of the invention, provided herein are recombinant cells comprising one or more heterologously expressed nucleic acids encoding monophosphate decarboxylase and/or isopentenyl kinase polypeptides as disclosed herein and further engineered to increase the activity of phosphotransacetylase (pta and/or eutD). Activity modulation (e.g., increased) of phosphotransacetylase isozymes is also contemplated herein. In any aspects of the invention, provided herein are recombinant cells comprising one or more expressed nucleic acids encoding monophosphate decarboxylase and/or isopentenyl kinase polypeptides as disclosed herein and further engineered to increase the activity of a phosphotransacetylase (pta and/or eutD) isozyme.

[0249] The activity of phosphotransacetylase (pta) and/or acetate kinase (ackA) can also be decreased by other molecular manipulation of the enzymes. The decrease of enzyme activity can be any amount of reduction of specific activity or total activity as compared to when no manipulation has been effectuated. In some instances, the decrease of enzyme activity is decreased by at least about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%,

30%, 35%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%. In any aspects of the invention, provided herein are recombinant cells comprising one or more heterologously expressed nucleic acids encoding phosphoketolase polypeptides as disclosed herein and further engineered to decrease the activity of acetate kinase (*ackA*). Activity modulation (e.g., decreased) of acetate kinase isozymes is also contemplated herein. In any aspects of the invention, provided herein are recombinant cells comprising one or more expressed nucleic acids encoding monophosphate decarboxylase and/or isopentenyl kinase polypeptides as disclosed herein and further engineered to decrease the activity of a acetate kinase isozyme.

[0250] In some cases, attenuating the activity of the endogenous phosphotransacetylase gene and/or the endogenous acetate kinase gene results in more carbon flux into the mevalonate dependent biosynthetic pathway in comparison to microorganisms that do not have attenuated endogenous phosphotransacetylase gene and/or endogenous acetate kinase gene expression.

Pathways involving glyceraldehyde 3-phosphate

[0251] Glyceraldehyde 3-phosphate dehydrogenase (*gapA* and/or *gapB*) is a crucial enzyme of glycolysis catalyzes the conversion of glyceraldehyde 3-phosphate into 1,3-biphospho-D-glycerate (Branlant G. and Branlant C. Eur. J. Biochem. 150:61-66 (1985)).

[0252] In certain aspects, recombinant cells comprising one or more expressed nucleic acids encoding monophosphate decarboxylase and/or isopentenyl kinase polypeptides as disclosed herein further comprise one more nucleic acids encoding a phosphoketolase polypeptide. In order to direct carbon towards the phosphoketolase enzyme, glyceraldehyde 3-phosphate dehydrogenase expression can be modulated (e.g., decrease enzyme activity) to allow more carbon to flux towards fructose 6-phosphate and xylulose 5-phosphate, thereby increasing the eventual production of isoprene. Decrease of glyceraldehyde 3-phosphate dehydrogenase activity can be any amount of reduction of specific activity or total activity as compared to when no manipulation has been effectuated. In some instances, the decrease of enzyme activity is decreased by at least about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%. Or 100%. In some aspects, the activity of glyceraldehyde 3-phosphate dehydrogenase is modulated by decreasing the activity of an endogenous glyceraldehyde 3-

phosphate dehydrogenase. This can be accomplished by replacing the endogenous glyceraldehyde 3-phosphate dehydrogenase gene promoter with a synthetic constitutively low expressing promoter. The gene encoding glyceraldehyde 3-phosphate dehydrogenase can also be deleted. The gene encoding glyceraldehyde 3-phosphate dehydrogenase can also be replaced by a *Bacillus* enzyme catalyzing the same reaction but producing NADPH rather than NADH. The decrease of the activity of glyceraldehyde 3-phosphate dehydrogenase can result in more carbon flux into the mevalonate-dependent biosynthetic pathway in comparison to cells that do not have decreased expression of glyceraldehyde 3-phosphate dehydrogenase. In any aspects of the invention, provided herein are recombinant cells comprising one or more expressed nucleic acids encoding monophosphate decarboxylase and/or isopentenyl kinase polypeptides as disclosed herein and further engineered to decrease the activity of glyceraldehyde 3-phosphate dehydrogenase (*gapA* and/or *gapB*). Activity modulation (e.g., decreased) of glyceraldehyde 3-phosphate dehydrogenase isozymes is also contemplated herein. In any aspects of the invention, provided herein are recombinant cells comprising one or more heterologously expressed nucleic acids encoding monophosphate decarboxylase and/or isopentenyl kinase polypeptides as disclosed herein and further engineered to decrease the activity of a glyceraldehyde 3-phosphate dehydrogenase (*gapA* and/or *gapB*) isozyme.

Pathways involving lactate dehydrogenase

[0253] In *E. coli*, D-Lactate is produced from pyruvate through the enzyme lactate dehydrogenase (LdhA) (Bunch, P. et al. Microbiol. 143:187-195 (1997)). Production of lactate is accompanied with oxidation of NADH, hence lactate is produced when oxygen is limited and cannot accommodate all the reducing equivalents. Thus, production of lactate could be a source for carbon consumption. As such, to improve carbon flow through to mevalonate production (and isoprene production, if desired), one of skill in the art can modulate the activity of lactate dehydrogenase, such as by decreasing the activity of the enzyme.

[0254] Accordingly, in one aspect, the activity of lactate dehydrogenase can be modulated by attenuating the activity of an endogenous lactate dehydrogenase gene. Such attenuation can be achieved by deletion of the endogenous lactate dehydrogenase gene. Other ways of attenuating the activity of lactate dehydrogenase gene known to one of skill in the art may also be used. By manipulating the pathway that involves lactate dehydrogenase, the recombinant microorganism produces decreased amounts of lactate in comparison to microorganisms that do

not have attenuated endogenous lactate dehydrogenase gene expression. Decrease in the amount of lactate produced can be measured by routine assays known to one of skill in the art. The amount of lactate reduction is at least about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% as compared when no molecular manipulations are done.

[0255] The activity of lactate dehydrogenase can also be decreased by other molecular manipulations of the enzyme. The decrease of enzyme activity can be any amount of reduction of specific activity or total activity as compared to when no manipulation has been effectuated. In some instances, the decrease of enzyme activity is decreased by at least about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%.

[0256] Accordingly, in some cases, attenuation of the activity of the endogenous lactate dehydrogenase gene results in more carbon flux into the mevalonate dependent biosynthetic pathway in comparison to microorganisms that do not have attenuated endogenous lactate dehydrogenase gene expression.

Pathways involving malic enzyme

[0257] Malic enzyme (in *E. coli* *sfcA* and *maeB*) is an anaplerotic enzyme that catalyzes the conversion of malate into pyruvate (using NAD⁺ or NADP⁺) by the equation below:



[0258] Thus, the two substrates of this enzyme are (S)-malate and NAD(P)⁺, whereas its 3 products are pyruvate, CO₂, and NADPH.

[0259] Expression of the NADP-dependent malic enzyme (*maeB*) (Iwikura, M. et al. J. Biochem. 85: 1355-1365 (1979)) can help increase mevalonate and/or isoprene yield by 1) bringing carbon from the TCA cycle back to pyruvate, direct precursor of acetyl-CoA, itself direct precursor of the mevalonate pathway and 2) producing extra NADPH which could be used in the HMG-CoA reductase reaction (Oh, M. K. et al. J. Biol. Chem. 277: 13175-13183 (2002); Bologna, F. et al. J. Bact. 189:5937-5946 (2007)).

[0260] As such, more starting substrate (pyruvate or acetyl-CoA) for the downstream production of mevalonate and/or isoprene can be achieved by modulating, such as increasing, the activity and/or expression of malic enzyme. The NADP-dependent malic enzyme gene can be an endogenous gene. One non-limiting way to accomplish this is by replacing the endogenous NADP-dependent malic enzyme gene promoter with a synthetic constitutively expressing promoter. Another non-limiting way to increase enzyme activity is by using one or more heterologous nucleic acids encoding an NADP-dependent malic enzyme polypeptide. One of skill in the art can monitor the expression of maeB RNA during fermentation or culturing using readily available molecular biology techniques.

[0261] Accordingly, in some embodiments, the recombinant microorganism produces increased amounts of pyruvate in comparison to microorganisms that do not have increased expression of an NADP-dependent malic enzyme gene. In some aspects, increasing the activity of an NADP-dependent malic enzyme gene results in more carbon flux into the mevalonate dependent biosynthetic pathway in comparison to microorganisms that do not have increased NADP-dependent malic enzyme gene expression.

[0262] Increase in the amount of pyruvate produced can be measured by routine assays known to one of skill in the art. The amount of pyruvate increase can be at least about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% as compared when no molecular manipulations are done.

[0263] The activity of malic enzyme can also be increased by other molecular manipulations of the enzyme. The increase of enzyme activity can be any amount of increase of specific activity or total activity as compared to when no manipulation has been effectuated. In some instances, the increase of enzyme activity is at least about 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%.

Pathways involving pyruvate dehydrogenase complex

[0264] The pyruvate dehydrogenase complex, which catalyzes the decarboxylation of pyruvate into acetyl-CoA, is composed of the proteins encoded by the genes aceE, aceF and

lpdA. Transcription of those genes is regulated by several regulators. Thus, one of skill in the art can increase acetyl-CoA by modulating the activity of the pyruvate dehydrogenase complex. Modulation can be to increase the activity and/or expression (*e.g.*, constant expression) of the pyruvate dehydrogenase complex. This can be accomplished by different ways, for example, by placing a strong constitutive promoter, like PL.6
(aattcatataaaaaacatacagataaacatctgcggtgataaattatctctggcggtgttgacataaataccactggcggtgatactgagcac
atcagcaggacgcaactgaccacatgaaggtg (SEQ ID NO:3) lambda promoter, GenBank NC_001416),
in front of the operon or using one or more synthetic constitutively expressing promoters.

[0265] Accordingly, in one aspect, the activity of pyruvate dehydrogenase is modulated by increasing the activity of one or more genes of the pyruvate dehydrogenase complex consisting of (a) pyruvate dehydrogenase (E1), (b) dihydrolipoyl transacetylase, and (c) dihydrolipoyl dehydrogenase. It is understood that any one, two or three of these genes can be manipulated for increasing activity of pyruvate dehydrogenase. In another aspect, the activity of the pyruvate dehydrogenase complex can be modulated by attenuating the activity of an endogenous pyruvate dehydrogenase complex repressor gene, further detailed below. The activity of an endogenous pyruvate dehydrogenase complex repressor can be attenuated by deletion of the endogenous pyruvate dehydrogenase complex repressor gene.

[0266] In some cases, one or more genes of the pyruvate dehydrogenase complex are endogenous genes. Another way to increase the activity of the pyruvate dehydrogenase complex is by introducing into the microorganism one or more heterologous nucleic acids encoding one or more polypeptides from the group consisting of (a) pyruvate dehydrogenase (E1), (b) dihydrolipoyl transacetylase, and (c) dihydrolipoyl dehydrogenase.

[0267] By using any of these methods, the recombinant microorganism can produce increased amounts of acetyl Co-A in comparison to microorganisms wherein the activity of pyruvate dehydrogenase is not modulated. Modulating the activity of pyruvate dehydrogenase can result in more carbon flux into the mevalonate dependent biosynthetic pathway in comparison to microorganisms that do not have modulated pyruvate dehydrogenase expression.

Combinations of mutations

[0268] It is understood that for any of the enzymes and/or enzyme pathways described herein, molecular manipulations that modulate any combination (two, three, four, five or six) of the enzymes and/or enzyme pathways described herein is expressly contemplated. For ease of the recitation of the combinations, citrate synthase (*gltA*) is designated as A, phosphotransacetylase (*ptaB*) is designated as B, acetate kinase (*ackA*) is designated as C, lactate dehydrogenase (*ldhA*) is designated as D, malic enzyme (*sfcA* or *maeB*) is designated as E, pyruvate decarboxylase (*aceE*, *aceF*, and/or *lpdA*) is designated as F, 6-phosphogluconolactonase (*ybhE*) is designated as G, and phosphoenolpyruvate carboxylase (*ppl*) is designated as H. As discussed above, *aceE*, *aceF*, and/or *lpdA* enzymes of the pyruvate decarboxylase complex can be used singly, or two of three enzymes, or three of three enzymes for increasing pyruvate decarboxylase activity.

[0269] Accordingly, for combinations of any two of the enzymes A-H, non-limiting combinations that can be used are: AB, AC, AD, AE, AF, AG, AH, BC, BD, BE, BF, BG, BH, CD, CE, CF, CG, CH, DE, DF, DG, DH, EF, EG, EH, and GH. For combinations of any three of the enzymes A-H, non-limiting combinations that can be used are: ABC, ABD, ABE, ABF, ABG, ABH, BCD, BCE, BCF, BCG, BCH, CDE, CDF, CDG, CDH, DEF, DEH, ACD, ACE, ACF, ACG, ACH, ADE, ADF, ADG, ADH, AEF, AEG, AEH, BDE, BDF, BDG, BDH, BEF, BEG, BEH, CEF, CEG, CEH, CFG, CFH, and CGH. For combinations of any four of the enzymes A-H, non-limiting combinations that can be used are: ABCD, ABCE, ABCF, ABCG, ABCH, ABDE, ABDF, ABDG, ABDH, ABEF, ABEG, ABEH, BCDE, BCDF, BCDG, BCDH, CDEF, CDEG, CDEH, ACDE, ACDF, ACDG, ACDH, ACEF, ACEG, ACEH, BCEF, BDEF, BGEF, BHEF, ADEF. For combinations of any five of the enzymes A-H, non-limiting combinations that can be used are: ABCDE, ABCDF, ABCDG, ABCDH, ABDEF, ABDEG, ABDEH, BCDEF, BCDEG, BCDEH, ACDEF, ACDEG, ACEDH, ABCEF, ABCEG, and ABCEH. For combinations of any six of the enzymes A-H, non-limiting combinations that can be used are: ABCDEF, ABCDEG, ABCDEH, BCDEFG, BCDEFH, and CDEFGH. For combinations of any seven of the enzymes A-H, non-limiting combinations that can be used are: ABCDEFG, ABCDEFH, BCDEFGH. In another aspect, all eight enzyme combinations are used ABCDEFGH.

[0270] Accordingly, the recombinant microorganism as described herein can achieve increased mevalonate production that is increased compared to microorganisms that are not grown under conditions of tri-carboxylic acid (TCA) cycle activity, wherein metabolic carbon flux in the recombinant microorganism is directed towards mevalonate production by modulating the activity of one or more enzymes from the group consisting of (a) citrate synthase, (b) phosphotransacetylase; (c) acetate kinase; (d) lactate dehydrogenase; (e) NADP-dependent malic enzyme; (f) pyruvate dehydrogenase; (g) 6-phosphogluconolactonase; and (h) phosphoenolpyruvate carboxylase.

Other regulators and factors for increased isoprene production

[0271] Other molecular manipulations can be used to increase the flow of carbon towards isoprene production. One method is to reduce, decrease or eliminate the effects of negative regulators for pathways that feed into the mevalonate pathway. For example, in some cases, the genes *aceEF-lpdA* are in an operon, with a fourth gene upstream *pdhR*. *pdhR* is a negative regulator of the transcription of its operon. In the absence of pyruvate, it binds its target promoter and represses transcription. It also regulates *ndh* and *cyoABCD* in the same way (Ogasawara, H. et al. *J. Bact.* 189:5534-5541 (2007)). In one aspect, deletion of *pdhR* regulator can improve the supply of pyruvate, and hence the production mevalonate and/or isoprene.

[0272] In other embodiments, any of the resultant strains described above can be further engineered to modulate the activity of the Entner-Doudoroff pathway. The gene coding for phosphogluconate dehydratase or aldolase can be attenuated or deleted. In other embodiments, any of the resultant strains described above may also be engineered to decrease or remove the activity of acetate kinase or citrate synthase. In other embodiments, any of the strains the resultant strain may also be engineered to decrease or remove the activity of phosphofructokinase. In other embodiments, any of the resultant strains described above may also be engineered to modulate the activity of glyceraldehyde-3-phosphate dehydrogenase. The activity of glyceraldehyde-3-phosphate dehydrogenase can be modulated by decreasing its activity. In other embodiments, the enzymes from the non-oxidative branch of the pentose phosphate pathway, such as transketolase, transaldolase, ribulose-5-phosphate-epimerase and (or) ribose-5-phosphate epimerase can be overexpressed.

[0273] In other aspects, the host cells can be further engineered to increase intracellular acetyl-phosphate concentrations by introducing heterologous nucleic acids encoding sedoheptulose-1,7-bisphosphatase/fructose-1,6-bisphosphate aldolase and sedoheptulose-1,7-bisphosphatase/fructose-1,6-bisphosphate phosphatase. In certain embodiments, the host cells having these molecular manipulations can be combined with attenuated or deleted transaldolase (talB) and phosphofructokinase (pfkA and/or pfkB) genes, thereby allowing faster conversion of erythrose 4-phosphate, dihydroxyacetone phosphate, and glyceraldehyde 3-phosphate into sedoheptulose 7-phosphate and fructose 1-phosphate.

[0274] In other aspects, the introduction of 6-phosphogluconolactonase (PGL) into microorganisms (such as various *E. coli* strains) which lack PGL can be used to improve production of mevalonate and/or isoprene. PGL may be introduced using chromosomal integration or extra-chromosomal vehicles, such as plasmids. In other aspects, PGL may be deleted from the genome of microorganisms (such as various *E. coli* strains) which express an endogenous PGL to improve production of mevalonate and/or isoprene. In some aspects, deletion of PGL results in any of about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100%, inclusive, including any values in between these percentages, higher percent yield of isoprene in comparison to microorganisms that express PGL. In other aspects, deletion of PGL results in any of about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100%, inclusive, including any values in between these percentages, higher instantaneous percent yield of isoprene in comparison to microorganisms that express PGL. In other aspects, deletion of PGL results in any of about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100%, inclusive, including any values in between these percentages, higher cell productivity index for isoprene in comparison to microorganisms that express PGL. In other aspects, deletion of PGL results in any of about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100%, inclusive, including any values in between these percentages, higher volumetric productivity of isoprene in comparison to microorganisms that express PGL. In other aspects, deletion of PGL results in any of about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 100%, inclusive, including any values in between these percentages, higher peak specific productivity of isoprene in comparison to microorganisms that express PGL. In some aspects the deletion of PGL results in peak specific productivity being maintained for a longer period of time in comparison to microorganisms that express PGL.

Exemplary Transformation Methods

[0275] Isoprene synthase, DXS, IDI, and/or MVA pathway nucleic acids or vectors containing them can be inserted into a host cell (*e.g.*, a plant cell, a fungal cell, a yeast cell, or a bacterial cell described herein) using standard techniques for expression of the encoded isoprene synthase, DXS, IDI, and/or MVA pathway polypeptide. Introduction of a DNA construct or vector into a host cell can be performed using techniques such as transformation, electroporation, nuclear microinjection, transduction, transfection (*e.g.*, lipofection mediated or DEAE-Dextrin mediated transfection or transfection using a recombinant phage virus), incubation with calcium phosphate DNA precipitate, high velocity bombardment with DNA-coated microprojectiles, and protoplast fusion. General transformation techniques are known in the art (see, *e.g.*, Current Protocols in Molecular Biology F. M. Ausubel et al. (eds) Chapter 9, (1987); Sambrook et al., Molecular Cloning: A Laboratory Manual, 3rd ed., Cold Spring Harbor, (2001); and Campbell et al., Curr Genet, 16:53-56, (1989), which are each hereby incorporated by reference in their entirety, particularly with respect to transformation methods). The introduced nucleic acids may be integrated into chromosomal DNA or maintained as extrachromosomal replicating sequences. Transformants can be selected by any method known in the art. Suitable methods for selecting transformants are described in International Publication No. WO 2009/076676, U.S. Patent Publ. No. 2009/0203102, WO 2010/003007, US Publ. No. 2010/0048964, WO 2009/132220, and US Publ. No. 2010/0003716.

[0276] Other exemplary transformation methods that can be used are described in US Pub. 2009/0203102, WO 2009/076676, WO 2010/003007, WO 2009/132220, WO 2010/031062, WO 2010/031068, WO 2010/031076, WO 2010/031077, and WO 2010/031079.

Exemplary Cell Culture Media

[0277] As used herein, the terms “minimal medium” or “minimal media” refer to growth media containing the minimum nutrients possible for cell growth, generally, but not always, without the presence of one or more amino acids (*e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more amino acids). Minimal medium typically contains: (1) a carbon source for host cell (*e.g.*, bacterial cell) growth; (2) various salts, which can vary among host cell species and growing conditions; and (3) water. The carbon source can vary significantly, from simple sugars like glucose to more complex hydrolysates of other biomass, such as yeast extract, as discussed in more detail

below. The salts generally provide essential elements such as magnesium, nitrogen, phosphorus, and sulfur to allow the cells to synthesize proteins and nucleic acids. Minimal medium can also be supplemented with selective agents, such as antibiotics, to select for the maintenance of certain plasmids and the like. For example, if a microorganism is resistant to a certain antibiotic, such as ampicillin or tetracycline, then that antibiotic can be added to the medium in order to prevent cells lacking the resistance from growing. Medium can be supplemented with other compounds as necessary to select for desired physiological or biochemical characteristics, such as particular amino acids and the like.

[0278] Any minimal medium formulation can be used to cultivate the host cells. Exemplary minimal medium formulations include, for example, M9 minimal medium and TM3 minimal medium. Each liter of M9 minimal medium contains (1) 200 ml sterile M9 salts (64 g $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$, 15 g KH_2PO_4 , 2.5 g NaCl , and 5.0 g NH_4Cl per liter); (2) 2 ml of 1 M MgSO_4 (sterile); (3) 20 ml of 20% (w/v) glucose (or other carbon source); and (4) 100 μl of 1 M CaCl_2 (sterile). Each liter of TM3 minimal medium contains (1) 13.6 g K_2HPO_4 ; (2) 13.6 g KH_2PO_4 ; (3) 2 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$; (4) 2 g Citric Acid Monohydrate; (5) 0.3 g Ferric Ammonium Citrate; (6) 3.2 g $(\text{NH}_4)_2\text{SO}_4$; (7) 0.2 g yeast extract; and (8) 1 ml of 1000X Trace Elements solution; pH is adjusted to ~ 6.8 and the solution is filter sterilized. Each liter of 1000X Trace Elements contains: (1) 40 g Citric Acid Monohydrate; (2) 30 g $\text{MnSO}_4 \cdot \text{H}_2\text{O}$; (3) 10 g NaCl ; (4) 1 g $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$; (4) 1 g $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$; (5) 1 g $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$; (6) 100 mg $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$; (7) 100 mg H_3BO_3 ; and (8) 100 mg $\text{NaMoO}_4 \cdot 2\text{H}_2\text{O}$; pH is adjusted to ~ 3.0 .

[0279] An additional exemplary minimal media includes (1) potassium phosphate K_2HPO_4 , (2) Magnesium Sulfate $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, (3) citric acid monohydrate $\text{C}_6\text{H}_8\text{O}_7 \cdot \text{H}_2\text{O}$, (4) ferric ammonium citrate $\text{NH}_4\text{FeC}_6\text{H}_5\text{O}_7$, (5) yeast extract (from biospringer), (6) 1000X Modified Trace Metal Solution, (7) sulfuric acid 50% w/v, (8) foamblast 882 (Emerald Performance Materials), and (9) Macro Salts Solution 3.36ml. All of the components are added together and dissolved in deionized H_2O and then heat sterilized. Following cooling to room temperature, the pH is adjusted to 7.0 with ammonium hydroxide (28%) and q.s. to volume. Vitamin Solution and spectinomycin are added after sterilization and pH adjustment.

[0280] Any carbon source can be used to cultivate the host cells. The term "carbon source" refers to one or more carbon-containing compounds capable of being metabolized by a host cell

or organism. For example, the cell medium used to cultivate the host cells may include any carbon source suitable for maintaining the viability or growing the host cells.

[0281] In some embodiments, the carbon source is a carbohydrate (such as monosaccharide, disaccharide, oligosaccharide, or polysaccharides), invert sugar (e.g., enzymatically treated sucrose syrup), glycerol, glycerine (e.g., a glycerine byproduct of a biodiesel or soap-making process), dihydroxyacetone, one-carbon source, fatty acid (e.g., a saturated fatty acid, unsaturated fatty acid, or polyunsaturated fatty acid), lipid, phospholipid, glycerolipid, monoglyceride, diglyceride, triglyceride, polypeptide (e.g., a microbial or plant protein or peptide), renewable carbon source (e.g., a biomass carbon source such as a hydrolyzed biomass carbon source; beet sugar or cane sugar molasses), yeast extract, component from a yeast extract, polymer, acid, alcohol, aldehyde, ketone, amino acid, succinate, lactate, acetate, ethanol, or any combination of two or more of the foregoing. In some embodiments, the carbon source is a product of photosynthesis, including, but not limited to, glucose.

[0282] Exemplary monosaccharides include glucose and fructose; exemplary oligosaccharides include lactose and sucrose, and exemplary polysaccharides include starch and cellulose. Exemplary carbohydrates include C6 sugars (e.g., fructose, mannose, galactose, or glucose) and C5 sugars (e.g., xylose or arabinose). In some embodiments, the cell medium includes a carbohydrate as well as a carbon source other than a carbohydrate (e.g., glycerol, glycerine, dihydroxyacetone, one-carbon source, fatty acid, lipid, phospholipid, glycerolipid, monoglyceride, diglyceride, triglyceride, renewable carbon source, or a component from a yeast extract). In some embodiments, the cell medium includes a carbohydrate as well as a polypeptide (e.g., a microbial or plant protein or peptide). In some embodiments, the microbial polypeptide is a polypeptide from yeast or bacteria. In some embodiments, the plant polypeptide is a polypeptide from soy, corn, canola, jatropha, palm, peanut, sunflower, coconut, mustard, rapeseed, cottonseed, palm kernel, olive, safflower, sesame, or linseed.

[0283] In some embodiments, the concentration of the carbohydrate is at least or about 5 grams per liter of broth (g/L, wherein the volume of broth includes both the volume of the cell medium and the volume of the cells), such as at least or about 10, 15, 20, 30, 40, 50, 60, 80, 100, 150, 200, 300, 400, or more g/L. In some embodiments, the concentration of the carbohydrate is between about 50 and about 400 g/L, such as between about 100 and about 360 g/L, between about 120 and about 360 g/L, or between about 200 and about 300 g/L. In some embodiments,

this concentration of carbohydrate includes the total amount of carbohydrate that is added before and/or during the culturing of the host cells.

[0284] Exemplary lipids are any substance containing one or more fatty acids that are C4 and above fatty acids that are saturated, unsaturated, or branched.

[0285] Exemplary fatty acids include compounds of the formula R-COOH, where "R" is a hydrocarbon. Exemplary unsaturated fatty acids include compounds where "R" includes at least one carbon-carbon double bond. Exemplary unsaturated fatty acids include, but are not limited to, oleic acid, vaccenic acid, linoleic acid, palmitelaidic acid, and arachidonic acid. Exemplary polyunsaturated fatty acids include compounds where "R" includes a plurality of carbon-carbon double bonds. Exemplary saturated fatty acids include compounds where "R" is a saturated aliphatic group. In some embodiments, the carbon source includes one or more C12-C22 fatty acids, such as a C12 saturated fatty acid, a C14 saturated fatty acid, a C16 saturated fatty acid, a C18 saturated fatty acid, a C20 saturated fatty acid, or a C22 saturated fatty acid. In an exemplary embodiment, the fatty acid is palmitic acid. In some embodiments, the carbon source is a salt of a fatty acid (*e.g.*, an unsaturated fatty acid), a derivative of a fatty acid (*e.g.*, an unsaturated fatty acid), or a salt of a derivative of fatty acid (*e.g.*, an unsaturated fatty acid). Suitable salts include, but are not limited to, lithium salts, potassium salts, sodium salts, and the like. Di- and triglycerols are fatty acid esters of glycerol.

[0286] In some embodiments, the concentration of the lipid, fatty acid, monoglyceride, diglyceride, or triglyceride is at least or about 1 gram per liter of broth (g/L, wherein the volume of broth includes both the volume of the cell medium and the volume of the cells), such as at least or about 5, 10, 15, 20, 30, 40, 50, 60, 80, 100, 150, 200, 300, 400, or more g/L. In some embodiments, the concentration of the lipid, fatty acid, monoglyceride, diglyceride, or triglyceride is between about 10 and about 400 g/L, such as between about 25 and about 300 g/L, between about 60 and about 180 g/L, or between about 75 and about 150 g/L. In some embodiments, the concentration includes the total amount of the lipid, fatty acid, monoglyceride, diglyceride, or triglyceride that is added before and/or during the culturing of the host cells. In some embodiments, the carbon source includes both (i) a lipid, fatty acid, monoglyceride, diglyceride, or triglyceride and (ii) a carbohydrate, such as glucose. In some embodiments, the ratio of the lipid, fatty acid, monoglyceride, diglyceride, or triglyceride to the carbohydrate is about 1:1 on a carbon basis (*i.e.*, one carbon in the lipid, fatty acid, monoglyceride, diglyceride,

or triglyceride per carbohydrate carbon). In particular embodiments, the amount of the lipid, fatty acid, monoglyceride, diglyceride, or triglyceride is between about 60 and 180 g/L, and the amount of the carbohydrate is between about 120 and 360 g/L.

[0287] Exemplary microbial polypeptide carbon sources include one or more polypeptides from yeast or bacteria. Exemplary plant polypeptide carbon sources include one or more polypeptides from soy, corn, canola, jatropha, palm, peanut, sunflower, coconut, mustard, rapeseed, cottonseed, palm kernel, olive, safflower, sesame, or linseed.

[0288] Exemplary renewable carbon sources include cheese whey permeate, cornsteep liquor, sugar beet molasses, barley malt, and components from any of the foregoing. Exemplary renewable carbon sources also include glucose, hexose, pentose and xylose present in biomass, such as corn, switchgrass, sugar cane, cell waste of fermentation processes, and protein by-product from the milling of soy, corn, or wheat. In some embodiments, the biomass carbon source is a lignocellulosic, hemicellulosic, or cellulosic material such as, but are not limited to, a grass, wheat, wheat straw, bagasse, sugar cane bagasse, soft wood pulp, corn, corn cob or husk, corn kernel, fiber from corn kernels, corn stover, switch grass, rice hull product, or a by-product from wet or dry milling of grains (*e.g.*, corn, sorghum, rye, triticale, barley, wheat, and/or distillers grains). Exemplary cellulosic materials include wood, paper and pulp waste, herbaceous plants, and fruit pulp. In some embodiments, the carbon source includes any plant part, such as stems, grains, roots, or tubers. In some embodiments, all or part of any of the following plants are used as a carbon source: corn, wheat, rye, sorghum, triticale, rice, millet, barley, cassava, legumes, such as beans and peas, potatoes, sweet potatoes, bananas, sugarcane, and/or tapioca. In some embodiments, the carbon source is a biomass hydrolysate, such as a biomass hydrolysate that includes both xylose and glucose or that includes both sucrose and glucose.

[0289] In some embodiments, the renewable carbon source (such as biomass) is pretreated before it is added to the cell culture medium. In some embodiments, the pretreatment includes enzymatic pretreatment, chemical pretreatment, or a combination of both enzymatic and chemical pretreatment (*see*, for example, Farzaneh et al., *Bioresource Technology* 96 (18): 2014-2018, (2005); U.S. Patent No. 6,176,176; U.S. Patent No. 6,106,888). In some embodiments, the renewable carbon source is partially or completely hydrolyzed before it is added to the cell culture medium.

[0290] In some embodiments, the renewable carbon source (such as corn stover) undergoes ammonia fiber expansion (AFEX) pretreatment before it is added to the cell culture medium (*see, for example, Farzaneh et al., Bioresource Technology 96 (18): 2014-2018, (2005)*). During AFEX pretreatment, a renewable carbon source is treated with liquid anhydrous ammonia at moderate temperatures (such as about 60 to about 100 °C) and high pressure (such as about 250 to about 300 psi) for about 5 minutes. Then, the pressure is rapidly released. In this process, the combined chemical and physical effects of lignin solubilization, hemicellulose hydrolysis, cellulose decrystallization, and increased surface area enables near complete enzymatic conversion of cellulose and hemicellulose to fermentable sugars. AFEX pretreatment has the advantage that nearly all of the ammonia can be recovered and reused, while the remaining serves as nitrogen source for microbes in downstream processes. Also, a wash stream is not required for AFEX pretreatment. Thus, dry matter recovery following the AFEX treatment is essentially 100%. AFEX is basically a dry-to-dry process. The treated renewable carbon source is stable for long periods and can be fed at very high solid loadings in enzymatic hydrolysis or fermentation processes. Cellulose and hemicellulose are well preserved in the AFEX process, with little or no degradation. There is no need for neutralization prior to the enzymatic hydrolysis of a renewable carbon source that has undergone AFEX pretreatment. Enzymatic hydrolysis of AFEX-treated carbon sources produces clean sugar streams for subsequent fermentation use.

[0291] In some embodiments, the concentration of the carbon source (*e.g., a renewable carbon source*) is equivalent to at least or about 0.1, 0.5, 1, 1.5, 2, 3, 4, 5, 10, 15, 20, 30, 40, or 50% glucose (w/v). The equivalent amount of glucose can be determined by using standard HPLC methods with glucose as a reference to measure the amount of glucose generated from the carbon source. In some embodiments, the concentration of the carbon source (*e.g., a renewable carbon source*) is equivalent to between about 0.1 and about 20% glucose, such as between about 0.1 and about 10% glucose, between about 0.5 and about 10% glucose, between about 1 and about 10% glucose, between about 1 and about 5% glucose, or between about 1 and about 2% glucose.

[0292] In some embodiments, the carbon source includes yeast extract or one or more components of yeast extract. In some aspects, the concentration of yeast extract is 0.1% (w/v), 0.09% (w/v), 0.08% (w/v), 0.07% (w/v), 0.06% (w/v), 0.05% (w/v), 0.04% (w/v), 0.03% (w/v),

0.02% (w/v), or 0.01% (w/v) yeast extract. In some embodiments, the concentration of yeast extract is at least 1 gram of yeast extract per liter of broth (g/L, wherein the volume of broth includes both the volume of the cell medium and the volume of the cells), such as at least or about 5, 10, 15, 20, 30, 40, 50, 60, 80, 100, 150, 200, 300, or more g/L. In some embodiments, the concentration of yeast extract is between about 1 and about 300 g/L, such as between about 1 and about 200 g/L, between about 5 and about 200 g/L, between about 5 and about 100 g/L, or between about 5 and about 60 g/L. In some embodiments, the concentration includes the total amount of yeast extract that is added before and/or during the culturing of the host cells. In some embodiments, the carbon source includes both yeast extract (or one or more components thereof) and another carbon source, such as glucose. In some embodiments, the ratio of yeast extract to the other carbon source is about 1:5, about 1:10, or about 1:20 (w/w).

[0293] Additionally the carbon source may also be one-carbon substrates such as carbon dioxide, or methanol. Glycerol production from single carbon sources (*e.g.*, methanol, formaldehyde, or formate) has been reported in methylotrophic yeasts (Yamada et al., *Agric. Biol. Chem.*, 53(2):541-543, (1989)) and in bacteria (Hunter et. al., *Biochemistry*, 24:4148-4155, (1985)). These organisms can assimilate single carbon compounds, ranging in oxidation state from methane to formate, and produce glycerol. The pathway of carbon assimilation can be through ribulose monophosphate, through serine, or through xylulose-momophosphate (Gottschalk, *Bacterial Metabolism*, Second Edition, Springer-Verlag: New York, (1986), which is hereby incorporated by reference in its entirety, particularly with respect to carbon sources). The ribulose monophosphate pathway involves the condensation of formate with ribulose-5-phosphate to form a six carbon sugar that becomes fructose and eventually the three carbon product glyceraldehyde-3-phosphate. Likewise, the serine pathway assimilates the one-carbon compound into the glycolytic pathway via methylenetetrahydrofolate.

[0294] In addition to one and two carbon substrates, methylotrophic organisms are also known to utilize a number of other carbon containing compounds such as methylamine, glucosamine and a variety of amino acids for metabolic activity. For example, methylotrophic yeast is known to utilize the carbon from methylamine to form trehalose or glycerol (Bellion et al., *Microb. Growth Cl Compd., Int. Symp.*, 7th ed., 415-432., Murrell et al. (eds), Intercept, Andover, UK, (1993)). Similarly, various species of *Candida* metabolize alanine or oleic acid (Sulter et al., *Arch. Microbiol.* 153(5), 485-9, (1990)).

[0295] In some embodiments, cells are cultured in a standard medium containing physiological salts and nutrients (*see, e.g.*, Pourquie, J. et al., *Biochemistry and Genetics of Cellulose Degradation*, Aubert et al. (eds), Academic Press, pp. 71-86, (1988); and Ilmen et al., *Appl. Environ. Microbiol.* 63:1298-1306, (1997)). Exemplary growth media are common commercially prepared media such as Luria Bertani (LB) broth, Sabouraud Dextrose (SD) broth, or Yeast medium (YM) broth. One skilled in the art of microbiology or fermentation science would know other defined or synthetic growth media that may also be used, and the appropriate medium for growth of particular host cells.

[0296] In some aspects, the cells described herein are capable of using syngas as a source of energy and/or carbon. In some embodiments, the syngas includes at least carbon monoxide and hydrogen. In some embodiments, the syngas further additionally includes one or more of carbon dioxide, water, or nitrogen. In some embodiments, the molar ratio of hydrogen to carbon monoxide in the syngas is 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 3.0, 4.0, 5.0, or 10.0. In some embodiments, the syngas comprises 10, 20, 30, 40, 50, 60, 70, 80, or 90% by volume carbon monoxide. In some embodiments, the syngas comprises 10, 20, 30, 40, 50, 60, 70, 80, or 90% by volume hydrogen. In some embodiments, the syngas comprises 10, 20, 30, 40, 50, 60, 70, 80, or 90% by volume carbon dioxide. In some embodiments, the syngas comprises 10, 20, 30, 40, 50, 60, 70, 80, or 90% by volume water. In some embodiments, the syngas comprises 10, 20, 30, 40, 50, 60, 70, 80, or 90% by volume nitrogen.

[0297] Synthesis gas may be derived from natural or synthetic sources. The source from which the syngas is derived is referred to as a “feedstock.” In some embodiments, the syngas is derived from biomass (e.g., wood, switch grass, agriculture waste, municipal waste) or carbohydrates (e.g., sugars). In other embodiments, the syngas is derived from coal, petroleum, kerogen, tar sands, oil shale, or natural gas. In other embodiments, the syngas is derived from rubber, such as from rubber tires.

[0298] Syngas can be derived from a feedstock by a variety of processes, including methane reforming, coal liquefaction, co-firing, fermentative reactions, enzymatic reactions, and biomass gasification. Biomass gasification is accomplished by subjecting biomass to partial oxidation in a reactor at temperatures above about 700 °C in the presence of less than a stoichiometric amount of oxygen. The oxygen is introduced into the bioreactor in the form of air, pure oxygen,

or steam. Gasification can occur in three main steps: 1) initial heating to dry out any moisture embedded in the biomass; 2) pyrolysis, in which the biomass is heated to 300-500 °C in the absence of oxidizing agents to yield gas, tars, oils and solid char residue; and 3) gasification of solid char, tars and gas to yield the primary components of syngas. Co-firing is accomplished by gasification of a coal/biomass mixture. The composition of the syngas, such as the identity and molar ratios of the components of the syngas, can vary depending on the feedstock from which it is derived and the method by which the feedstock is converted to syngas.

[0299] Synthesis gas can contain impurities, the nature and amount of which vary according to both the feedstock and the process used in production. Fermentations may be tolerant to some impurities, but there remains the need to remove from the syngas materials such as tars and particulates that might foul the fermentor and associated equipment. It is also advisable to remove compounds that might contaminate the isoprene product such as volatile organic compounds, acid gases, methane, benzene, toluene, ethylbenzene, xylenes, H₂S, COS, CS₂, HCl, O₃, organosulfur compounds, ammonia, nitrogen oxides, nitrogen-containing organic compounds, and heavy metal vapors. Removal of impurities from syngas can be achieved by one of several means, including gas scrubbing, treatment with solid-phase adsorbents, and purification using gas-permeable membranes.

[0300] In addition to an appropriate carbon source, the cell medium desirably contains suitable minerals, salts, cofactors, buffers, and other components known to those skilled in the art suitable for the growth of the cultures or the enhancement of isoprene production (*see*, for example, WO 2004/033646 and references cited therein and WO 96/35796 and references cited therein). In some embodiments where an isoprene synthase, DXS, IDI, and/or MVA pathway nucleic acid is under the control of an inducible promoter, the inducing agent (*e.g.*, a sugar, metal salt or antimicrobial), is desirably added to the medium at a concentration effective to induce expression of an isoprene synthase, DXS, IDI, and/or MVA pathway polypeptide. In some embodiments, cell medium has an antibiotic (such as kanamycin) that corresponds to the antibiotic resistance nucleic acid (such as a kanamycin resistance nucleic acid) on a vector that has one or more DXS, IDI, or MVA pathway nucleic acids.

[0301] Other exemplary cell culture media that can be used are described in US Pub. 2009/0203102, WO 2009/076676, WO 2010/003007, WO 2009/132220, WO 2010/031062, WO 2010/031068, WO 2010/031076, WO 2010/031077, and WO 2010/031079.

Methods for the Production of Isoprene

[0302] Provided herein are methods of producing isoprene by culturing any of the recombinant cells described herein under conditions suitable for production of isoprene. In one aspect, isoprene can be produced by culturing recombinant cells expressing one or more nucleic acids encoding: (a) polypeptide having isoprene synthase activity and improved kinetic properties, wherein polypeptide is encoded by a heterologous nucleic acid; and (b) one or more mevalonate (MVA) pathway polypeptides in culture media. In one aspect, one or more heterologous nucleic acids encoding a thiolase, a HMG-CoA reductase, a lower MVA pathway polypeptide, and an isoprene synthase polypeptide can be used. In another aspect, isoprene can be produced by culturing recombinant cells comprising one or more heterologous nucleic acids encoding a thiolase, a HMG-CoA reductase and HMG-CoA synthase, a lower MVA pathway polypeptide, and an isoprene synthase polypeptide. In yet another aspect, one or more heterologous nucleic acids encoding one or more upper MVA pathway polypeptides, one or more lower MVA pathway polypeptides, and/or one or more DXP pathway polypeptides can be used. In some aspects, the recombinant cells described herein exhibit any of about 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 95%, or 100%, inclusive, including any value in between these percentages, increased isoprene production in comparison to cells which do not comprise one or more heterologous nucleic acids encoding an isoprene synthase polypeptide and one or more MVA pathway polypeptides. The isoprene can be produced from any of the cells described herein and according to any of the methods described herein. Any of the cells can be used for the purpose of producing isoprene using the fermentable sugar feedstocks produced by any of the methods described herein as a carbon source.

[0303] The cells can further express one or more nucleic acid molecules encoding the lower MVA pathway polypeptide(s) described above (*e.g.*, MVK, PMK, MVD, and/or IDI), any of the upper MVA pathways polypeptide(s) described above (*e.g.*, a thiolase, an acetoacetyl-CoA synthase, an HMG-CoA reductase, and/or an HMG-CoA synthase) and/or any of the isoprene synthase polypeptide(s) described above (*e.g.* *P. alba* isoprene synthase). In some aspects, the recombinant (*e.g.*, bacterial) cells can be any of the cells described herein. Any of the isoprene synthases or variants thereof described herein, any of the bacterial strains described herein, any of the promoters described herein, and/or any of the vectors described herein can also be used to

produce isoprene. In some aspects, the method of producing isoprene further comprises a step of recovering the isoprene.

[0304] In some aspects, the amount of isoprene produced is measured at a productivity time point. In some aspects, the productivity for the cells is about any of the amounts of isoprene disclosed herein. In some aspects, the cumulative, total amount of isoprene produced is measured. In some aspects, the cumulative total productivity for the cells is about any of the amounts of isoprene disclosed herein.

[0305] In some aspects, any of the cells described herein (for examples the cells in culture) produce isoprene at greater than about any of or about any of 1, 10, 25, 50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1,000, 1,250, 1,500, 1,750, 2,000, 2,500, 3,000, 4,000, 5,000, or more nmole of isoprene/gram of cells for the wet weight of the cells/hour (nmole/g_{wcm}/hr). In some aspects, the amount of isoprene is between about 2 to about 5,000 nmole/g_{wcm}/hr, such as between about 2 to about 100 nmole/g_{wcm}/hr, about 100 to about 500 nmole/g_{wcm}/hr, about 150 to about 500 nmole/g_{wcm}/hr, about 500 to about 1,000 nmole/g_{wcm}/hr, about 1,000 to about 2,000 nmole/g_{wcm}/hr, or about 2,000 to about 5,000 nmole/g_{wcm}/hr. In some aspects, the amount of isoprene is between about 20 to about 5,000 nmole/g_{wcm}/hr, about 100 to about 5,000 nmole/g_{wcm}/hr, about 200 to about 2,000 nmole/g_{wcm}/hr, about 200 to about 1,000 nmole/g_{wcm}/hr, about 300 to about 1,000 nmole/g_{wcm}/hr, or about 400 to about 1,000 nmole/g_{wcm}/hr.

[0306] In some aspects, the cells in culture produce isoprene at greater than or about 1, 10, 25, 50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1,000, 1,250, 1,500, 1,750, 2,000, 2,500, 3,000, 4,000, 5,000, 10,000, 100,000, or more ng of isoprene/gram of cells for the wet weight of the cells/hr (ng/g_{wcm}/h). In some aspects, the amount of isoprene is between about 2 to about 5,000 ng/g_{wcm}/h, such as between about 2 to about 100 ng/g_{wcm}/h, about 100 to about 500 ng/g_{wcm}/h, about 500 to about 1,000 ng/g_{wcm}/h, about 1,000 to about 2,000 ng/g_{wcm}/h, or about 2,000 to about 5,000 ng/g_{wcm}/h. In some aspects, the amount of isoprene is between about 20 to about 5,000 ng/g_{wcm}/h, about 100 to about 5,000 ng/g_{wcm}/h, about 200 to about 2,000 ng/g_{wcm}/h, about 200 to about 1,000 ng/g_{wcm}/h, about 300 to about 1,000 ng/g_{wcm}/h, or about 400 to about 1,000 ng/g_{wcm}/h.

[0307] In some aspects, the cells in culture produce a cumulative titer (total amount) of isoprene at greater than about any of or about any of 1, 10, 25, 50, 100, 150, 200, 250, 300, 400,

500, 600, 700, 800, 900, 1,000, 1,250, 1,500, 1,750, 2,000, 2,500, 3,000, 4,000, 5,000, 10,000, 50,000, 100,000, or more mg of isoprene/L of broth (mg/L_{broth}, wherein the volume of broth includes the volume of the cells and the cell medium). In some aspects, the amount of isoprene is between about 2 to about 5,000 mg/L_{broth}, such as between about 2 to about 100 mg/L_{broth}, about 100 to about 500 mg/L_{broth}, about 500 to about 1,000 mg/L_{broth}, about 1,000 to about 2,000 mg/L_{broth}, or about 2,000 to about 5,000 mg/L_{broth}. In some aspects, the amount of isoprene is between about 20 to about 5,000 mg/L_{broth}, about 100 to about 5,000 mg/L_{broth}, about 200 to about 2,000 mg/L_{broth}, about 200 to about 1,000 mg/L_{broth}, about 300 to about 1,000 mg/L_{broth}, or about 400 to about 1,000 mg/L_{broth}.

[0308] In some aspects, the isoprene produced by the cells in culture (such as any of the recombinant cells described herein) comprises at least about 1, 2, 5, 10, 15, 20, or 25% by volume of the fermentation offgas. In some aspects, the isoprene comprises between about 1 to about 25% by volume of the offgas, such as between about 5 to about 15 %, about 15 to about 25%, about 10 to about 20%, or about 1 to about 10 %.

[0309] In some aspects, any of the methods described herein further include a step of recovering isoprene produced by any of the recombinant cells disclosed herein. In some aspects, the isoprene is recovered by absorption stripping (*See, e.g.*, U.S. Patent Application Publication No. 2011/0178261 A1). Other purification methods which can be used are described in more detail in U.S. Patent Application Publication No. US2010/0196977 A1.

Exemplary Production of Isoprene

[0310] In some embodiments, the cells are cultured in a culture medium under conditions permitting the production of isoprene by the cells. In some embodiments, the cells in culture produce isoprene at greater than or about 1, 10, 25, 50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1,000, 1,250, 1,500, 1,750, 2,000, 2,500, 3,000, 4,000, 5,000, or more nmole of isoprene/gram of cells for the wet weight of the cells/hour (nmole/gwcm/hr). In some embodiments, the amount of isoprene is between about 2 to about 5,000 nmole/gwcm/hr, such as between about 2 to about 100 nmole/gwcm/hr, about 100 to about 500 nmole/gwcm/hr, about 150 to about 500 nmole/gwcm/hr, about 500 to about 1,000 nmole/gwcm/hr, about 1,000 to about 2,000 nmole/gwcm/hr, or about 2,000 to about 5,000 nmole/gwcm/hr. The amount of isoprene in units of nmole/gwcm/hr can be measured as disclosed in U.S. Patent No. 5,849,970.

For example, two mL of headspace (*e.g.*, headspace from a culture such as 2 mL of culture cultured in sealed vials at 32°C with shaking at 200 rpm for approximately 3 hours) are analyzed for isoprene using a standard gas chromatography system, such as a system operated isothermally (85°C) with an n-octane/porasil C column (Alltech Associates, Inc., Deerfield, IL) and coupled to a RGD2 mercuric oxide reduction gas detector (Trace Analytical, Menlo Park, CA) (*see, e.g.*, Greenberg et al, *Atmos. Environ.* 27A: 2689-2692, (1993); Silver et al., *Plant Physiol.* 97:1588-1591, (1991)). The gas chromatography area units are converted to nmol isoprene via a standard isoprene concentration calibration curve. In some embodiments, the value for the grams of cells for the wet weight of the cells is calculated by obtaining the A_{600} value for a sample of the cell culture, and then converting the A_{600} value to grams of cells based on a calibration curve of wet weights for cell cultures with a known A_{600} value. In some embodiments, the grams of the cells is estimated by assuming that one liter of broth (including cell medium and cells) with an A_{600} value of 1 has a wet cell weight of 1 gram. The value is also divided by the number of hours the culture has been incubating for, such as three hours.

[0311] In some embodiments, the cells in culture produce isoprene at greater than or about 1, 10, 25, 50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1,000, 1,250, 1,500, 1,750, 2,000, 2,500, 3,000, 4,000, 5,000, 10,000, 100,000, or more ng of isoprene/gram of cells for the wet weight of the cells/hr (ng/gwcm/h). In some embodiments, the amount of isoprene is between about 2 to about 5,000 ng/gwcm/h, such as between about 2 to about 100 ng/gwcm/h, about 100 to about 500 ng/gwcm/h, about 500 to about 1,000 ng/gwcm/h, about 1,000 to about 2,000 ng/gwcm/h, or about 2,000 to about 5,000 ng/gwcm/h. The amount of isoprene in ng/gwcm/h can be calculated by multiplying the value for isoprene production in the units of nmole/gwcm/hr discussed above by 68.1 (as described in Equation 5 below).

[0312] In some embodiments, the cells in culture produce a cumulative titer (total amount) of isoprene at greater than or about 1, 10, 25, 50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1,000, 1,250, 1,500, 1,750, 2,000, 2,500, 3,000, 4,000, 5,000, 10,000, 50,000, 100,000, or more mg of isoprene/L of broth (mg/L broth, wherein the volume of broth includes the volume of the cells and the cell medium). In some embodiments, the amount of isoprene is between about 2 to about 5,000 mg/L broth, such as between about 2 to about 100 mg/L broth, about 100 to about 500 mg/L broth, about 500 to about 1,000 mg/L broth, about 1,000 to about 2,000 mg/L broth, or about 2,000 to about 5,000 mg/L broth. The specific productivity of

isoprene in mg of isoprene/L of headspace from shake flask or similar cultures can be measured by taking a 1 ml sample from the cell culture at an OD_{600} value of approximately 1.0, putting it in a 20 mL vial, incubating for 30 minutes, and then measuring the amount of isoprene in the headspace. If the OD_{600} value is not 1.0, then the measurement can be normalized to an OD_{600} value of 1.0 by dividing by the OD_{600} value. The value of mg isoprene/L headspace can be converted to mg/Lbroth/hr/ OD_{600} of culture broth by multiplying by a factor of 38. The value in units of mg/Lbroth/hr/ OD_{600} can be multiplied by the number of hours and the OD_{600} value to obtain the cumulative titer in units of mg of isoprene/L of broth.

[0313] The instantaneous isoprene production rate in mg/L broth/hr in a fermentor can be measured by taking a sample of the fermentor off-gas, analyzing it for the amount of isoprene (in units such as mg of isoprene per L of gas), and multiplying this value by the rate at which off-gas is passed through each liter of broth (*e.g.*, at 1 vvm (volume of air/volume of broth/minute) this is 60 Lgas per hour). Thus, an off-gas level of 1 mg/Lgas corresponds to an instantaneous production rate of 60 mg/Lbroth/hr at air flow of 1 vvm. If desired, the value in the units mg/Lbroth/hr can be divided by the OD_{600} value to obtain the specific rate in units of mg/Lbroth/hr/ OD . The average value of mg isoprene/Lgas can be converted to the total product productivity (grams of isoprene per liter of fermentation broth, mg/Lbroth) by multiplying this average off-gas isoprene concentration by the total amount of off-gas sparged per liter of fermentation broth during the fermentation. Thus, an average off-gas isoprene concentration of 0.5 mg/Lbroth/hr over 10 hours at 1 vvm corresponds to a total product concentration of 300 mg isoprene/Lbroth.

[0314] In some embodiments, the cells in culture convert greater than or about 0.0015, 0.002, 0.005, 0.01, 0.02, 0.05, 0.1, 0.12, 0.14, 0.16, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.2, 1.4, or 1.6% of the carbon in the cell culture medium into isoprene. In some embodiments, the percent conversion of carbon into isoprene is between about 0.002 to about 1.6%, such as about 0.002 to about 0.005%, about 0.005 to about 0.01%, about 0.01 to about 0.05%, about 0.05 to about 0.15%, 0.15 to about 0.2%, about 0.2 to about 0.3%, about 0.3 to about 0.5%, about 0.5 to about 0.8%, about 0.8 to about 1.0%, or about 1.0 to about 1.6%. The percent conversion of carbon into isoprene (also referred to as “% carbon yield”) can be measured by dividing the moles carbon in the isoprene produced by the moles carbon in the carbon source (such as the

moles of carbon in batched and fed glucose and yeast extract). This number is multiplied by 100% to give a percentage value (as indicated in Equation 1).

Equation 1

% Carbon Yield = (moles carbon in isoprene produced)/(moles carbon in carbon source) * 100

[0315] For this calculation, yeast extract can be assumed to contain 50% w/w carbon.

Equation 2

% Carbon Yield = (39.1 g isoprene * 1/68.1 mol/g * 5 C/mol)/[(181221 g glucose * 1/180 mol/g * 6 C/mol) + (17780 g yeast extract * 0.5 * 1/12 mol/g)] * 100 = 0.042%

[0316] One skilled in the art can readily convert the rates of isoprene production or amount of isoprene produced into any other units. Exemplary equations are listed below for interconverting between units.

Units for Rate of Isoprene production (total and specific)

Equation 3

1 g isoprene/L_{broth}/hr = 14.7 mmol isoprene/L_{broth}/hr (total volumetric rate)

Equation 4

1 nmol isoprene /g_{wcm}/hr = 1 nmol isoprene /L_{broth}/hr/OD₆₀₀ (This conversion assumes that one liter of broth with an OD₆₀₀ value of 1 has a wet cell weight of 1 gram.)

Equation 5

1 nmol isoprene/g_{wcm}/hr = 68.1 ng isoprene/g_{wcm}/hr (given the molecular weight of isoprene)

Equation 6

1 nmol isoprene/L_{gas} O₂/hr = 90 nmol isoprene/L_{broth}/hr (at an O₂ flow rate of 90 L/hr per L of culture broth)

Equation 7

$1 \mu\text{g isoprene}/L_{\text{gas}} \text{ isoprene in off-gas} = 60 \mu\text{g isoprene}/L_{\text{broth}}/\text{hr}$ at a flow rate of 60
 L_{gas} per L_{broth} (1 vvm)

Units for Titer (total and specific)

Equation 8

$1 \text{ nmol isoprene}/\text{mg cell protein} = 150 \text{ nmol isoprene}/L_{\text{broth}}/\text{OD}_{600}$ (This conversion assumes that one liter of broth with an OD_{600} value of 1 has a total cell protein of approximately 150 mg) (specific productivity)

Equation 9

$1 \text{ g isoprene}/L_{\text{broth}} = 14.7 \text{ mmol isoprene}/L_{\text{broth}}$ (total titer)

[0317] If desired, Equation 10 can be used to convert any of the units that include the wet weight of the cells into the corresponding units that include the dry weight of the cells.

Equation 10

Dry weight of cells = (wet weight of cells)/3.3

[0318] In some embodiments encompassed by the invention, a cell comprising a heterologous nucleic acid encoding an isoprene synthase variant polypeptide produces an amount of isoprene that is at least or about 2-fold, 3-fold, 5-fold, 10-fold, 25-fold, 50-fold, 100-fold, 150-fold, 200-fold, 400-fold, or greater than the amount of isoprene produced from a corresponding cell grown under essentially the same conditions without the heterologous nucleic acid encoding the isoprene synthase variant polypeptide.

[0319] In some embodiments encompassed by the invention, a cell comprising a heterologous nucleic acid encoding an isoprene synthase variant polypeptide and one or more heterologous nucleic acids encoding a DXS, IDI, and/or MVA pathway polypeptide produces an amount of isoprene that is at least or about 2-fold, 3-fold, 5-fold, 10-fold, 25-fold, 50-fold, 100-fold, 150-fold, 200-fold, 400-fold, or greater than the amount of isoprene produced from a

corresponding cell grown under essentially the same conditions without the heterologous nucleic acids.

Exemplary Isoprene Purification Methods

[0320] In some embodiments, any of the methods described herein further include recovering the isoprene. For example, the isoprene produced using the compositions and methods of the invention can be recovered using standard techniques, such as gas stripping, membrane enhanced separation, fractionation, adsorption/desorption, pervaporation, thermal or vacuum desorption of isoprene from a solid phase, or extraction of isoprene immobilized or absorbed to a solid phase with a solvent (*see*, for example, U.S. Patent Nos. 4,703,007 and 4,570,029, which are each hereby incorporated by reference in their entireties, particularly with respect to isoprene recovery and purification methods). In one aspect, the isoprene is recovered by absorption stripping (*see, e.g.*, US Pub. No. 2011/0178261). In particular aspects, extractive distillation with an alcohol (such as ethanol, methanol, propanol, or a combination thereof) is used to recover the isoprene. In some embodiments, the recovery of isoprene involves the isolation of isoprene in a liquid form (such as a neat solution of isoprene or a solution of isoprene in a solvent). Gas stripping involves the removal of isoprene vapor from the fermentation off-gas stream in a continuous manner. Such removal can be achieved in several different ways including, but not limited to, adsorption to a solid phase, partition into a liquid phase, or direct condensation. In some embodiments, membrane enrichment of a dilute isoprene vapor stream above the dew point of the vapor resulting in the condensation of liquid isoprene. In some embodiments, the recovery is performed as described in U.S. 2011/0178261, which is incorporated by reference, in particular for the teaching for purification and/or recovery of isoprene.

[0321] The recovery of isoprene may involve one step or multiple steps. In some embodiments, the removal of isoprene vapor from the fermentation off-gas and the conversion of isoprene to a liquid phase are performed simultaneously. For example, isoprene can be directly condensed from the off-gas stream to form a liquid. In some embodiments, the removal of isoprene vapor from the fermentation off-gas and the conversion of isoprene to a liquid phase are performed sequentially. For example, isoprene may be adsorbed to a solid phase and then extracted from the solid phase with a solvent. In one aspect, the isoprene is recovered by using absorption stripping as described in U.S. Appl. No. 12/969,440 (US Publ. No. 2011/0178261).

[0322] In some embodiments, any of the methods described herein further include purifying the isoprene. For example, the isoprene produced using the compositions and methods of the invention can be purified using standard techniques. Purification refers to a process through which isoprene is separated from one or more components that are present when the isoprene is produced. In some embodiments, the isoprene is obtained as a substantially pure liquid. Examples of purification methods include (i) distillation from a solution in a liquid extractant and (ii) chromatography. As used herein, “purified isoprene” means isoprene that has been separated from one or more components that are present when the isoprene is produced. *See, e.g.* U.S. Patent Application Publication No. 2009/0203102, PCT Publication No. WO 2009/076676 and U.S. Patent Application Serial No. 12/496,573. In some embodiments, the isoprene is at least about 20%, by weight, free from other components that are present when the isoprene is produced. In various embodiments, the isoprene is at least or about 25%, 30%, 40%, 50%, 60%, 70%, 75%, 80%, 90%, 95%, or 99%, by weight, pure. Purity can be assayed by any appropriate method, *e.g.*, by column chromatography, HPLC analysis, or GC-MS analysis. Suitable purification methods are described in more detail in U.S. Patent Application Publication 2010/0196977 A1.

[0323] Throughout this specification, various patents, patent applications and other types of publications (*e.g.*, journal articles, nucleotide and protein sequence database accession numbers) are referenced. The disclosure of all patents, patent applications, sequences and publications cited herein are hereby incorporated by reference in their entirety for all purposes.

[0324] The invention can be further understood by reference to the following examples, which are provided by way of illustration and are not meant to be limiting.

EXAMPLES

[0325] Various modifications and variations of the described composition, method and/or system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific embodiments or aspects, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. It is also to be understood that one, some, or all of the properties of the various embodiments described herein may be combined to form

other embodiments of the present invention. These and other aspects of the invention will become apparent to one of skill in the art.

[0326] In the experimental disclosure which follows, the following abbreviations apply: °C (degrees Centigrade); rpm (revolutions per minute); H₂O (water); diH₂O (deionized water); aa and AA (amino acid); bp (base pair); kb (kilobase pair); kD (kilodaltons); gm (grams); µg and ug (micrograms); mg (milligrams); ng (nanograms); µl and ul (microliters); ml (milliliters); mm (millimeters); qs (quantity sufficient); nm (nanometers); µm and um (micrometer); M (molar); mM (millimolar); µM and uM (micromolar); pM (picomolar); U (units); MW (molecular weight); sec (seconds); min (minute/minutes); hr (hour/hours); OD₆₀₀ (optical density at 600 nm); BSA (bovine serum albumin); DMAPP (dimethylallyl diphosphate); DTT (dithiothreitol); EtOH (ethanol); IPTG (isopropyl-beta-D-thiogalactopyranoside); isoprene (2-methyl-1,3-butadiene); IspS (isoprene synthase); PAGE (polyacrylamide gel electrophoresis); PBS (phosphate buffered saline [150 mM NaCl, 10 mM sodium phosphate buffer, pH 7.2]); and SDS (sodium dodecyl sulfate).

[0327] The following abbreviations apply to companies whose products or services may have been referred to in the experimental examples: Agilent (Agilent Technologies, Santa Clara, CA); Becton Coulter (Becton Coulter, Inc., Fullerton, CA); Bio-Rad (Bio-Rad Laboratories, Hercules, CA); Cayman Chemical (Cayman Chemical Co., Ann Arbor, MI); CTC Analytics (CTC Analytics A.G., Zwingen, Switzerland); EMS (Electron Microscopy Supply, Hatfield, PA); Epicentre (Epicentre Biotechnologies, Madison, WI); Integrated DNA Technologies (Integrated DNA Technologies, Coralville, IA); Invitrogen (Invitrogen Corp., Carlsbad, CA); Molecular Dynamics (Molecular Dynamics, Sunnyvale, CA); Novagen (Novagen, Inc., Madison, WI); Perkin Elmer (Perkin Elmer, Waltham, MA); Roche (Roche Applied Science, Indianapolis, IN); Sigma (Sigma-Aldrich, St. Louis, MO); Stratagene (Stratagene Cloning Systems, La Jolla, CA); Qiagen (Qiagen, Inc., Valencia, CA); Takara (Takara Bio USA, Madison, WI); Thomson Instrument (Thomson Instrument Co., Oceanside, CA); V&P Scientific (V&P Scientific, Inc., San Diego, CA); and Zinsser (Zinsser North America, Northridge, CA).

Example 1: Isoprene Synthase Growth Screen: Validation, Optimization and Limitations

[0328] This Example describes the development of an *in vivo* screen to select improved variants of isoprene synthase. The inventors have found that the *in vivo* screen can be used to

select cells that contain less isoprene synthase activity than a control (in our case, an isoprene synthase MEA-Poplar Alba). In addition, the *in vivo* screen can be used to select cells that contain more isoprene synthase activity than a control (in our case, an isoprene synthase MEA-Poplar Alba).

Methods

[0329] Strains: The screening strain contained a constitutively expressed lower pathway and variants of isoprene synthase expressed from a pET plasmid. The screening strains were DW425-positive control.

[0330] Assay conditions: Strains were grown overnight in LB medium containing 50 μ M kanamycin at 34°C. The overnight cultures were diluted to approximately 0.2 OD₆₀₀ in TM3 media containing 1% glucose, 0.1% yeast extract, 8 mM MgSO₄ and one of the following concentrations of IPTG: 0, 10, 20, 30, 40, 50, 60, or 70 μ M. Cells were grown for approximately 2 hours post-induction and transferred to a 96-well clear bottom microtiter plate containing various concentrations of mevalonate (0, 5, 7.5, 10, 15, 20 mM) and the same media used in the day culture to a final OD₆₀₀ of 0.2-0.3. The plates were monitored in kinetic mode on a Spectramax UV-Vis spectrophotometer. The experiment was monitored at 34 C for 3 hours with shaking for 1 minute prior to each measurement (taken every 5 minutes).

[0331] Data analysis: All data were transferred to Excel. The absorbance measurements were converted to their natural log. A line was then fit to the series using the function “LINEST” to yield the exponential growth constant (growth rate).

[0332] Metabolite analysis was performed by using the following protocol for methanol/water extraction of metabolites at small-scale (MVA, DXP):

1. Samples from small-scale experiment was quenched; commonly 1 ml sample was spun down, the supernatant was discarded, 100 μ l pure methanol was loaded onto the pellet, and the samples was stored at -80C until there was time for metabolite extraction and analysis.
2. Samples were taken from storage in -80C; pellets were resuspended (recommended to break pellet with glass capillary tubes).
3. The sample was spun down in a refrigerated microcentrifuge at 14000g (rfc) for 4 min.
4. The supernatant was placed into clean 1.5 mL Eppendorf tubes.

5. The pellet was resuspended in 100 μ l 6:1 MeOH/5 mM NH₄OAc pH 8.0. Centrifuge at 14000 g (rfc) for 4 min. The samples may be extracted in 6:1 MeOH/5 mM NH₄OAc pH 7.0 if the metabolites of interest are not stable at pH 8.0 (for example, DXP metabolites, or CoA-containing metabolites).
6. The supernatant was combined with the supernatant from step 4.
7. Steps 5-6 were repeated, extracting with 100 μ l 1:1 MeOH/5 mM NH₄OAc pH 8.0 (or pH 7.0, see above). Sample pellets can be discarded after taking the supernatant. 1.5 ml Eppendorf tubes containing accumulated supernatant fractions were closed and extract was mixed by vortexing.
8. In order to remove suspended debris, 1.5 ml Eppendorf tubes were centrifuged at 14000 g (rfc) for 4 min.
9. ~200 μ l extract were placed into LC/MS vials containing conical inserts. The remaining extracts were stored at -20 °C.

[0333] Without being bound by theory, it is recommended to use repeat pipettors for dispensing 2% formic acid (for fast pipetting and consistent volumes). Repeat pipettors significantly improve time efficiency over standard pipettors, and because they are technically positive displacement pipettes, they are quite precise (and accurate, assuming good calibration and proper maintenance). Further recommendations include, but are not limited to: keep Eppendorf tubes on ice (at 0 °C) whenever possible, the microcentrifuge should be set at -9 °C; allow ~20 min. for the centrifuge to cool, for resuspending pellets, the use of the glass capillary tubes is recommended. Mechanical breaking of the cell pellet is usually very fast with just a little physical assistance. It is not recommended to vortex the resuspended pellets, as the cell mass ends up on the sides of the tubes very easily, potentially causing significant experimental error due to the low volumes of the samples.

[0334] Without being bound by theory, the following recommendations are given for conducting LC/MS analysis:

1. The LC/MS vials should be kept on the tray at 4 °C during the analysis. The column should be at room temperature. The tray/column temperatures will be set automatically after starting the sequence in Xcalibur, but it is better to set the tray temperature in advance.
2. Use standards prepared as in the attached spreadsheet for calibration. Record standard preparation date as labeled on each tube.

3. LC/MS method for isoprenoids and MVA pathway metabolites (currently on new TSQ Quantum Access) – Method file: IPS_BioBasic100_090316 (or similar, see latest date extension); HPLC column: Macherey-Nagel Nucleodex beta-OH EC 2 mm x 100 mm (particle size 5 μm , pore size 100 \AA), C/N 720351.20; Guard column: 721460.40 (2 mm guard column not currently available). LC/MS method for DXP pathway metabolites: Method development C18-ion pair\Metabolites_C18_TBAip_11, with tributylammonium acetate as an ion-pair reagent; HPLC column: C18 Phenomenex Synergi 4 μ Hydro-RP 80A 150x2.0 mm, C/N 00F-4375-B0; Guard column: Security Guard Cartridges AQ C18 4 x 2.0mm, C/N AJ0-7510. For LC/MS method for detection of CoA-containing metabolites, see “Protocol for acidic extraction of metabolites at small-scale (CoAs, etc.)”
4. After analysis, samples should be stored at -20 °C; standards should be stored at -80 °C.
5. Metabolite quantitation can be determined using LCQuan software package. After back-calculation of all dilutions (including initial methanol quench), concentrations should be normalized to OD and converted to intracellular concentrations, utilizing the assumption that the intracellular volume of 1L of fermentation broth at 200 OD is ~50 mL.

Results

[0335] A system has previously been developed to select DMAPP utilizing enzymes from pools of plasmids that express unknown proteins (Appl Environ Microbiol. 2007 October; 73(19): 6277–6283). The inventors have refined and optimized the screening protocol to enable selection cells that contain isoprene synthase activity. The screen is based on experimental results concluding that the concentration of DMAPP in *E. coli* correlates with cell growth rate (Figure 1). Therefore, without being bound by theory, the growth rate of these cells can be thought of as a biosensor for intracellular DMAPP concentrations. Without being bound by theory, the underlying rationale of this screen was that the concentration of DMAPP in a cell can be decreased by increasing the enzyme activity of DMAPP consuming enzymes (isoprene synthase) and, thus, would result in increased growth rate.

[0336] To test this hypothesis, DW425 cells were grown in media containing a matrix of IPTG and mevalonate concentrations (Figure 2 and Figure 3). Cells grown in the presence of 60 μM IPTG, and greater, resulted in impaired growth without addition of mevalonate (compared to non-induced cells). At IPTG concentrations of 0-50 μM IPTG, growth was unimpaired, compared to non-induced cells. Cell growth was inhibited with all concentrations of

mevalonate in uninduced cells. Increasing the concentration of IPTG resulted in increased growth rate for any given mevalonate concentration screened (Figure 2 and Figure 3). Earlier studies were performed to determine that the concentration of enzyme expressed in the cells correlates with the concentration of IPTG present. Therefore, increased isoprene synthase expression/activity in these strains results in improved growth.

Example 2: Analysis of *P. alba* Isoprene Synthase SELs by DMAPP Toxicity Relief

[0337] There is a strong correlation between increased intracellular DMAPP levels and growth inhibition of *E. coli*, which can be alleviated by the expression of *P. alba* isoprene synthase (IspS). Without being bound by theory, increased levels of IspS activity should therefore allow for better growth due to more rapid conversion of DMAPP to isoprene. By monitoring the growth rates of *E. coli* expressing variants of IspS under these conditions, the inventors can identify mutant IspS enzymes that display increased ability to convert DMAPP to isoprene within the cell.

Methods:

1) Plasmid and Strain Construction:

[0338] SEL Plasmid Backbone – The plasmid backbone used to generate SELs was constructed by QuikChange (Stratagene) PCR on the template pDu39 (see Table 1 for primer sequences). The PCR product was treated with 1 μ l DpnI (Roche) for 3 hours, and then 1 μ l of the entire reaction was transformed into chemically competent *E. coli* Top10 cells (Invitrogen) according to the manufacturer's recommended protocol. Cells were recovered and plated on LB medium containing 50 μ g/ml kanamycin. The next day, positive colonies were chosen for growth, plasmid purification (Qiagen) and sequencing (Quintara Biosciences). Plasmids which harbored the correct base changes were selected for sequencing of the entire open reading frame to confirm the integrity of the coding sequence. One of these plasmids, pCL201 (see Figure 4), was selected as the backbone for construction of SELs (by Verdezyne and DNA2.0).

Table 1. QuikChange and Sequencing Primers

MEA Hairpin Disrupt (pET) F	ggagatatacatatggaagcacgtcgcctctgcgaactacgaacctaa	(SEQ ID NO:4)
MEA Hairpin Disrupt	ttaggttcgtagttcgcagagcgcgacgtgcttccatgtatatctcc	(SEQ ID NO:5)

cttctaataacctggaatgctgtttccggggatcgcagtggtgagtaacctgcatcatcaggagtaggataaaatgcttgatggtcggaa
 gaggcataaattccgtcagccagtttagtctgaccatctcatctgtaacatcattggcaacgctacctttgccatgtttcagaacaactctggc
 gcatcgggcttcccatacaategatagattgtcgacctgattcccacattatcgcgagcccatttataccatataaatcagcatccatgtt
 ggaatttaatecggcctagagcaagacgtttccgtgaatatggctcataacacccctgtattactgtttatgtaagcagacagttttattgtt
 catgacaaaaatcccttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttctgcg
 cgtaatctgctgcttgaacaaaaaaaccaccgctaccagcgggtgtttgttggccggatcaagagctaccaactcttttccgaaggtaact
 ggcttcagcagagcgcagataccaataactgtccttctagttagccgtagtttaggccaccactcaagaactctgtagcaccgcctacatac
 ctgctctgctaactctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttgactcaagacgatagttaccggataag
 gcgcagcggctgggctgaacggggggtcgtgcacacagcccagcttgagcgaacgacctacaccgaactgagatacctacagcgtg
 agctatgagaaagcggcacgcttcccgaaggagaaagggcggacaggtatccggtaagcggcagggctcggaaacaggagagcgcacg
 agggagcttccaggggaaacgcctggtatctttatagctctgctgggttccacactctgacttgagcgtcgattttgtgatgctcgtcag
 gggggcggagcctatgaaaaacgccagcaacgcggccttttacggttctggccttttctgctcactgttcttctcgtcgtta
 tccctgattctgtgataaccgtattaccgcctttgagtgagctgataccgctcggcgcagccgaacgaccgagcgcagcagtcagtga
 gcgaggaagcggaaagagcgcctgatgcggtattttctccttacgcatctgtgcggatttcacaccgcataatggtgcactctcagtacaat
 ctgctctgatccgcatagttaagccaagtataactccgctatcgtactgactgggtcatggctgcgccccgacaccgccaacaccg
 ctgacgcgcctgacgggcttctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggtttca
 ccgcatcaccgaaacgcgcgagggcagctgcggtaagctcatcagcgtggtcgtgaagcgattcacagatgtctgcctgttcatccgcgt
 ccagctcgttgagtttccagaagcgttaatgtctgcttctgataaagcgggccatgtaaggcgggttttctgtttggtcactgatgcct
 ccgtgtaagggggatttctgttcatgggggtaatgataccgatgaaacgagagaggatgctcacgatacgggttactgatgatgaacatgcc
 cgttactggaacggtgtgagggtaaacaactggcggatggatgcggcgggaccagagaaaaatcactcagggtcaatgccagcgttc
 gftaatacagatgtaggtgtccacagggtagccagcagcatcctgcgatgcagatccggaacataatggtgcagggcgtgacttccgcg
 tttccagactttacgaaacacggaaaccgaagaccattcatgttgtgctcaggtcgcagacgtttgcagcagcagtcgcttcacgttcgctc
 gcgtatcgggtgattcattctgctaaccagtaaggcaaccccgccagcctagccgggtcctcaacgacaggagcacgatcatgcgaccccg
 tggggccgcatgccggcgataatggcctgcttctgccgaaacgtttggtgcccgggaccagtgacgaaggcttgagcggaggcgtgca
 agattccgaataaccgcaagcgcagggccgatcatcgtcgcgctccagcgaagcggctcctcggcgaatgacctagagcgtgcccgg
 cacctgtcctacgagttgcatgataaagaagacagtcataagtgcggcgacgatagtcatgccccgcgccaccggaaggagctgactg
 ggtgaaaggctcgaaggcctcggctcagatcccgggtgcctaatgagtgagctaaacttacattaattcggttgcgctcactgcccgtttcc
 agtcgggaaacctgtcgtccagctgcattaatgaatcggccaacgcgcggggagagggcgtttgcgtattgggcgccagggtgtgttttc
 tttaccagtgagacgggcaacagctgattgcccttaccgctggccctgagagagttgcagcaagcgggtccacgctggtttgcccag
 caggcgaaaatcctgtttgatggtggttaacggcgggatataacatgagctgtcttcgggtatcgtcgtatcccactaccgatatccgacc
 aacgcgcagcccggactcggtaatggcgcgcattgcgcccagcgcctatctgatcgttggaaccagcatcagtgggaaacgatgcct
 cattcagcatttgcaggtttgtgaaaaccggacatggcactccagtcgccttcccgttccgctatcggctgaatttgattgcgagtgagatat

ttatgccagccagccagacgcagacgccgagacagaacttaatgggcccgctaacagcgcgatttgctggtgaccaatgcgaccag
atgctccacgccagtcgctaccgtctcatgggagaaaataactgttgatgggtgctggtcagagacatcaagaaataacgccggaa
cattagtgcaggcagctccacagcaatggcactcctggtcatccagcggatagttaatgatcagcccactgacgcgttgccgagaaatt
gtgcaccgccgctttacaggttcgacgccgcttcgttaccatcgcaccaccacgctggcaccagttgatcggcgcgagattaatcg
ccgcgacaatttgcgacggcgcgtgagggccagactggaggtggcaacgccaatcagcaacgactgtttgcccgccagttgttgcc
acgcggttgggaatgtaattcagctccgccatcggcgttccacttttccgcgtttcgcagaaacgtggctggcctggtcaccacgcgg
gaaacggtctgataagagacaccggcactctgcgacatcgtataacgttactggtttcacattcaccacctgaattgactcttccgggc
gctatcatgccataccgcgaaaggtttgcgccattcgtggtgctccgggatctcgcgctccttatgcgactcctgcattaggaaagcag
cccagtagtaggttagggcgttagcaccgccgccaaggaatggtgatgcaaggagatggcgcccaacagtccccggccacg
gggcctgccaccatacccacgccgaaacaagcgtcatgagcccgaagtggcgaagccgatcttccccatcgggatgctggcgatata
ggcgccagcaaccgcacctgtggcgccggtgatccggccacgatgcgtccggcgtagaggatcagatctcgcgaaattaat
acgactcactataggggaattgtgagcggataacaattcccctctagaataattttgttaactttaagaaggagatatacatatggaagcac
gtcgtctcgaactacgaacctaacagctgggactatgattacctgctctccgacacggacgagtcctcgaagtatacaaaagacaa
agcgaanaagctggaagccgaagttcgtcgcgagattaataacgaaaaagcagaatttctgacctgctggaactgattgacaacgtccag
cgcctgggcctgggttaccgttcgagctgatatccgtggtgcgctggatcgttcgttctccggcggttcgatgcggttaaccaagactt
ccctgcacggtacggcactgtcttccgtctgctgcgtcaaacgggtttgaggtttctcaggaagcgttcagcgggttcaaaagacaaaacg
gcaacttctggagaacctgaaggaagatatcaaaactatcctgacgtgacgagccagcttccgtgctctggaagggcgaacatcct
ggacgagggcgaaggtttcgaatctctcatctgaaagaactgctgaaagaaagatcgtaaaagctggcagaacaggtgaacctgc
actggaactgccactgcacgccgtactcagcgtctggaagcagatggtctatcagggcctaccgtaaaaaggaggacgcgaatcaggtt
ctgctggagctggcaattctggattacaacatgatccagctctgataaccagcgtgatctcgtgaaacgtcccgttgggtgctggtgggt
ctggcgacaaaactgcactttgctcgtgaccgcctgattgagagcttctactgggccgtgggtgtagcattcgaaccgcaactcctgactg
ccgtaactccgtcgcaaaaatgtttcttctgtaaccattatcgcgatatctacgatgtatacggcacctggacgaactggagctgttactg
atgcagttgagcgttgggacgtaaacccatcaacgacctgccggattacatgaaactgtgctttctgctctgtataacactatfaacgaaat
cgcctacgacaacctgaaagataaaggtgagaacatcctgccgtatctgaccaaagcctgggctgacctgtgcaacgcttctcgaagaa
gccaagtggctgtacaacaaatctactccgacctttgacgactacttcggcaacgcattggaatcctcttctggcccgtgcaactggtgttc
gcttactcgtcgtcgtcagaacattaaaaaggagagatcgaaacctgcaaaaatccatgacaccatctctcgtccttccatattctcc
gtctgtgcaatgacctggctagcgcgtctcggaaattgcgctggtgaaaccgcaaatagcgtttctgttacatgcgactaaaggtatct
ccgaagaactggctaccgaaagcgtgatgaatctgatcgtgaaacctggaaaaagatgaacaaggaaaaactgggtgtagcctgttcg
cgaaaccgttcgtggaaccgcgatcaacctggcacgtcaatctcactgcacttatcataacggcgacgcgcatacctctccgatgagct
gacctgcaaacgcgttctgctgtaatcactgaaccgattctgccgttgaacgctaaagatccgaattcagctccgtcgacaagcttgcg
gccgactcgcgaccaccaccaccactgagatccggctgctaacaagcccgaaggaagctgagttggctgctgccaccgctg

agcaataactagcataaccacctggggcctctaaacgggtcttgaggggttttttctgaaaggaggaactatatccggat (SEQ ID NO:21)

Amino Acid Sequence of *P. alba* IspS:

MEARRSANYEPNSWDYDYLLSSDTDESIEVYKDKAKKLEAEVRREINNEKAEFLTLEL
IDNVQRLGLGYRFESDIRGALDRFVSSGGFDAVTKTSLHGTALSFRLLRQHGFVVSQEAF
SGFKDQNGNFLENLKEDIKAILSLEYEASFLALEGENILDEAKVFAISHLKELSEEKIGKEL
AEQVNHAELEPLHRRTQRLEAVWSIEAYRKKEDANQVLELAILDYNMIQSVYQRDLR
ETSRWWRRVGLATKLHFARDRLIESFYWAVGVAFEPQYSDCRNSVAKMFSFVTIIDDIY
DVYGTLDLELELFTDAVERWDVNAINDLDPYMKLCFLALYNTINEIAYDNLKDKGENILP
YLTAWADLCNAFLQEAKWLYNKSTPTFDDYFGNAWKSSSGPLQLVFAYFAVVQNIK
KEEIEENLQKYHDTISRPSHIFRLCNDLASASAEIARGETANSVSCYMRTKGISEELATESV
MNLIDETWKKMNEKLGGSLEAKPFVETAINLARQSHCTYHNGDAHTSPDELTRKRVL
SVITEPILPFER (SEQ ID NO:22)

[0339] SEL Expression Host – A P1 lysate of MCM521 (described herein) was made and transduced into BL21(DE3) according to standard molecular biology techniques (Miller, A Short Course in Bacterial Genetics). Transductants were selected on LB medium plates containing 20 µg/ml of kanamycin. Positive colonies were further verified by PCR to confirm the presence of PL.2-mKKDyI in the BL21 DE3 strain. 1 µl of pCP20 plasmid was then transformed into this strain and positive colonies were selected for on LB containing 50 µg/ml carbenicillin and incubated overnight at 30°C. Positive transformants were streaked on an LB plate and incubated at 37°C to induce loss of the pCP20 plasmid. To confirm the loss of the neomycin (kanamycin) resistance marker, colonies that grew at 37°C were patched onto LB medium containing either 20 µg/ml of kanamycin, 50 µg/ml of carbenicillin, or no antibiotics. The strains with integrated PL.2 mKKDyI without the kanamycin resistance marker that have lost pCP20 should be sensitive to kanamycin and carbenicillin. Four colonies sensitive to kanamycin and carbenicillin were checked by PCR for the presence of mKKDyI in BL21(DE3) with the parental BL21(DE3) strain as a control. The resulting strain, MD09-170, was used for expression of IspS variants in the growth assay on SELs described below. Control strains for the growth assay harbored either the empty pET24a+ vector or pCL201 as negative and positive controls for IspS expression, respectively (see Table 2).

Table 2. Strains

Strain	Plasmid	Description
MD09-170		BL21 (DE3) PL.2-mKKDyI
DW424	pET24a+	BL21 (DE3) PL.2-mKKDyI + vector (negative control)
DW425	pCL201	BL21 (DE3) PL.2-mKKDyI + <i>P. alba</i> IspS (wild type control)

2) Construction of SELs:

[0340] 25 Site Evaluation Libraries (SELs) of IspS were analyzed previously for specific activity. Table 3 lists the residues included in these libraries. For the growth assay, plasmids harboring variant IspS enzymes in these libraries were purified and transformed into the expression host MD09-170: Original libraries were replicated directly from glycerol stocks and grown at 30°C overnight in LB containing 50 µg/ml kanamycin in 96 deep-well plates (VWR). Cells from the overnight growth were harvested by centrifugation (Eppendorf 5804 R), and supernatants were discarded. Plasmid purification was performed on cell pellets using a Hamilton Microlab STAR robot using the Nucleospin Multi-96 Plus Plasmid purification kit (Macherey Nagel) according to the manufacturers' recommended protocols. 3 µl of the resulting plasmid DNA for each variant was transformed into chemically competent MD09-170 cells in a flat-bottom 96-well polystyrene plate (Falcon) using an Eppendorf Thermomixer R set to 42°C. Cells were recovered for 2 hours in LB medium, and then diluted and incubated overnight in LB medium containing 50 µg/ml kanamycin. Glycerol stocks of plates containing MD09-170 cells with variants from all 25 original libraries were made and stored at -80°C prior to analysis by growth assay.

[0341] A second set of 80 SELs was ordered and manufactured by DNA2.0. These libraries were transformed directly into the screening host MD09-170. Table 4 lists all 80 residues that were chosen for this set. Sites were picked primarily based upon their location in the recently solved crystal structure of *P. alba* IspS. Strains DW425 and DW424 (see Table 2) were seeded back into 96-well plates for the wild type and negative controls, respectively.

Table 3. Sites selected for 25 SELs

Description	Sites
Surface hydrophobic residues	I28, V30, L130, G153, V299, L303, L469, L494
Hinge region	R198, I229, L260
Negatively charged area	D311, D323

Flexible loops	A443, A453, N454, H515, A519, E525
Active site	F388, N438, E451
Miscellaneous	D345, R528, T536

Table 4. Sites selected for 80 SELs

Position	Residue	Selection Criteria
3	A	surface hydrophobic
7	A	symmetry contact
9	Y	symmetry contact
12	N	conservation
13	S	conservation
16	Y	N-terminal loop
18	Y	N-terminal loop
20	L	conservation
23	D	conservation
25	D	surface hydrophilic
26	E	symmetry contact
27	S	surface hydrophilic
33	D	symmetry contact
36	K	symmetry contact
44	R	symmetry contact
50	K	surface hydrophilic
53	F	conservation
59	L	surface hydrophobic
69	G	conservation
74	S	surface hydrophilic
78	G	conservation
81	D	surface hydrophilic
87	G	surface hydrophobic
99	G	conservation
116	Q	conservation
117	E	symmetry contact
120	S	surface hydrophilic
121	G	surface loop
125	Q	surface hydrophilic
127	G	conservation
139	A	conservation
165	I	surface hydrophobic
173	E	surface hydrophilic
174	E	symmetry contact
177	G	conservation
179	E	surface hydrophilic
194	R	conservation
197	Q	conservation
202	V	conservation
216	Q	conservation

240	T	conservation
246	R	symmetry contact
251	T	surface hydrophilic
254	H	conservation
287	F	active site
290	V	active site
308	L	surface hydrophobic
376	L	flexible loops
377	Y	symmetry contact
379	K	conservation
389	G	conservation
397	G	active site
400	Q	hydrophobic pocket
403	F	active site
421	Q	conservation
426	T	conservation
430	P	misc
434	F	active site
445	A	surface hydrophobic
448	A	conservation
457	S	flexible loops
462	T	conservation
476	N	surface hydrophilic
487	K	surface loop
488	E	surface loop
489	K	surface loop
490	L	surface loop
491	G	surface loop
492	G	surface loop
493	S	surface loop
495	F	surface loop
496	A	surface loop
497	K	surface loop
498	P	conservation
509	Q	conservation
514	Y	active site
521	T	conservation
539	I	C-terminal
540	L	surface hydrophobic
544	R	conservation

3) Growth Assay for Increased IspS Activity

[0342] For the growth assay, glycerol stocks of SELs were inoculated into 200 μ l LB medium containing 50 μ g/ml kanamycin in flat bottom microtiter plates (Cellstar) and grown

overnight at 30°C using the System Duetz (EnzyScreen BV). For pre-induction, 7 µl of the overnight culture from each well was inoculated into 100 µl of TM3 medium containing 50 µM IPTG and 50 µg/ml kanamycin, and plates were grown for 2 hours at 30°C. Pre-induced cultures were then diluted 1:10 into TM3 medium containing 11 mM mevalonic acid, 50 µM IPTG and 50 µg/ml kanamycin in glass-bottom 96 square-well microtiter plates (Matrical). Cultures were grown at 34°C and shaken at 225 rpm for approximately 10 hours in a Growth Profiler 1152 (EnzyScreen). Growth curves were generated for each IspS variant according to the manufacturer's recommended protocol. Negative controls were strains harboring the empty pET24a+ vector (DW424), and positive controls were strains expressing wild type *P. alba* IspS (DW425) grown either with or without MVA.

[0343] For data analysis, the relative growth rate of each variant over a given period of time was measured against wild type controls. Specifically, a line was fit to the series using the "LINEST" function in Microsoft Excel to yield the exponential growth constant (growth rate). These values were then divided by the average of 4 (in most cases) growth constants from the positive controls to yield a "Growth Index" number for each variant. The Growth Index values for variants in all 105 SEL libraries are listed in Table 5. In some cases, a particular variant was either absent from the glycerol stock, did not grow in the overnight LB culture, or was not transferred to the final plate for growth assay. Values for these particular wells are listed as ND (Not Determined). In the instances where a particular variant was not generated in initial mutagenesis of *P. alba* IspS, the wild type residue was substituted.

Table 5. Growth Index Ranking for Plate 001.

Variant	GI
V30L	1.61
V30K	1.52
I28T	1.46
V30Y	1.45
V30W	1.39
I28S	1.38
G153W	1.34
I28R	1.31
L130Y	1.29
V30S	1.26
V30V	1.25
V30F	1.22
L130L	1.20
L130K	1.17
I28I	1.15
L130G	1.15
L130I	1.14
V30R	1.13
G153Y	1.12
L130L	1.12
V30I	1.11
I28Y	1.10
L130V	1.09
L130L	1.09
V30P	1.08
V30E	1.07
G153G	1.07
L130R	1.05
V30T	1.05
L130L	1.04
V30Q	1.04
I28P	1.03
G153S	1.02
G153R	1.01
G153C	1.00
L130D	0.97
G153G	0.97
L130E	0.96
G153A	0.96
V30V	0.95

V30M	0.94
I28I	0.92
G153G	0.92
L130L	0.91
G153T	0.90
V30G	0.89
L130W	0.88
L130Q	0.88
G153Q	0.86
G153G	0.86
G153M	0.83
G153N	0.83
G153D	0.81
L130P	0.76
V30N	0.60
I28I	0.57
I28I	0.55
I28L	0.52
I28W	0.42
I28D	0.38
G153K	0.32
G153L	0.32
G153P	0.32
V30C	0.31
G153V	0.31
I28E	0.29
L130A	0.28
I28G	0.28
I28F	0.27
V30A	0.26
I28A	ND
L130M	0.22
I28M	0.15
L130C	0.15
I28N	0.14
I28C	0.10
I28V	ND
V30D	ND
L130S	ND
G153H	ND

Table 5. Growth Index Ranking for Plate 002.

Variant	GI		
R198R	1.11	L260G	0.34
R198R	1.08	V299T	0.29
I229L	1.00	L260C	0.25
R198R	0.98	V299A	0.25
I229C	0.94	R198G	0.24
L260L	0.89	R198N	0.23
L260L	0.88	L260A	0.22
R198R	0.87	V299M	0.22
I229I	0.86	L260S	ND
R198K	0.84	V299S	0.15
V299V	0.82	L260D	0.15
I229T	0.80	V299N	0.14
I229V	0.78	L260E	0.11
I229M	0.78	I229Y	0.11
L260M	0.77	V299R	0.10
L260L	0.75	R198L	0.10
L260W	0.74	I229K	0.09
R198C	0.74	R198F	0.08
V299L	0.72	I229W	0.08
V299V	0.69	R198H	0.07
R198M	0.68	V299E	0.05
L260Y	0.66	V299K	0.03
R198V	0.64	R198Y	0.03
R198A	0.62	V299W	0.02
I229A	0.62	R198D	0.02
V299V	0.62	I229D	0.02
L260Q	0.61	I229R	0.01
V299V	0.60	V299Y	0.01
I229H	0.57	V299G	0.01
L260I	0.54	I229E	0.01
R198I	0.52	L260R	0.01
I229N	0.49	R198P	ND
I229F	0.45	L260K	0.00
I229S	0.45	L260T	0.00
L260H	0.45	V299I	ND
R198T	0.42	I229P	-0.02
L260P	0.40	V299H	-0.03
L260V	0.38	V299P	-0.04
I229G	0.36	V299D	-0.05
I229Q	0.35		
R198S	0.35		

Table 5. Growth Index Ranking for Plate 003.

Variant	GI
L303L	1.10
L303L	1.10
D311D	1.08
L303L	1.06
L303L	0.99
D311W	0.99
D311L	0.98
D311K	0.98
L303W	0.96
D311F	0.93
D311S	0.92
D311G	0.90
D323K	0.89
L303T	0.88
D311R	0.87
L303R	0.85
D311T	0.84
D311V	0.82
L303L	0.81
L303V	0.81
L303I	0.80
D311D	0.80
D311A	0.80
L303S	0.79
D311I	0.78
L303M	0.77
D323L	0.77
D323R	0.75
D311D	0.75
L303E	0.74
D323Y	0.74
L303C	0.74
D323S	0.74
D345L	0.73
D311Q	0.73
L303G	0.72
D323G	0.71
D311D	0.71
D345G	0.70
D323T	0.69

D311E	0.69
D323M	0.69
D323N	0.68
D345W	0.68
D311D	0.67
L303Q	0.67
D345D	0.66
L303A	0.66
D323E	0.66
L303P	0.66
D323A	0.66
D345M	0.66
D345A	0.65
D323D	0.64
D323V	0.63
D323C	0.62
L303H	0.62
L303D	0.62
D323F	0.61
D345S	0.61
D323I	0.61
D345V	0.59
D323Q	0.58
D345T	0.58
D323W	0.57
D345Y	0.57
D345R	0.56
D345N	0.55
D345K	0.53
D345Q	0.53
D345I	0.53
D345D	0.53
D345E	0.52
D345C	0.51
D345D	0.41
D345P	0.37
D311M	ND
D323P	ND
D323H	ND
D311P	-0.02

Table 5. Growth Index Ranking for Plate 004.

Variant	GI
E451E	1.48
A443A	1.36
A443A	1.34
A443A	1.29
F388F	1.28
N438N	1.27
A443A	1.20
F388V	1.20
A443A	1.17
A443A	1.15
A443A	1.13
E451E	1.13
A443A	1.12
F388Q	0.96
F388T	0.89
F388F	0.89
A443S	0.86
A443H	0.85
A443L	0.84
A443N	0.82
F388R	0.74
F388S	0.68
A443Q	0.66
A443I	0.63
A443G	0.50
A443R	0.44
F388D	0.42
A443V	0.40
A443F	0.28
A443P	0.06
E451Q	0.06
A443T	0.03
F388H	ND
F388L	0.00
F388F	-0.01
F388F	-0.01
E451W	-0.01
E451K	-0.01
F388A	-0.02
N438A	-0.02

E451G	-0.02
E451H	-0.02
F388P	-0.02
E451D	-0.02
E451R	-0.03
F388Y	ND
F388K	-0.03
E451Y	-0.03
E451F	-0.03
F388C	-0.03
E451I	-0.03
N438F	-0.04
E451V	-0.04
N438H	-0.04
E451A	-0.04
E451C	-0.04
E451T	-0.04
N438T	-0.04
N438D	-0.04
N438E	-0.04
E451L	-0.05
N438I	-0.05
N438W	ND
E451P	-0.05
N438V	-0.05
F388F	-0.05
N438M	-0.05
N438L	-0.05
N438C	-0.05
N438K	-0.06
E451S	-0.06
F388F	-0.06
E451N	-0.06
N438R	-0.06
N438Y	-0.06
N438Q	-0.06
N438P	-0.06
N438G	-0.07
F388G	-0.08
N438S	-0.09

Table 5. Growth Index Ranking for Plate 005.

Variant	GI
N454N	1.10
L494L	1.06
L469W	1.06
L494K	1.03
L469L	1.00
L469R	0.98
L469L	0.98
L494L	0.98
L494W	0.97
L494Y	0.95
L469V	0.95
L494G	0.94
L469F	0.93
L494I	0.93
L469I	0.92
A453T	0.91
A453A	0.90
L494V	0.90
L494R	0.89
A453W	0.88
L469N	0.88
A453C	0.88
L469T	0.88
L469S	0.87
L469H	0.87
N454N	0.86
L494C	0.86
L494D	0.86
L494E	0.86
L469C	0.85
A453S	0.85
A453A	0.85
N454S	0.84
A453A	0.84
L469P	0.83
L494S	0.83
N454G	0.83
A453N	0.83
L494H	0.83
L469A	0.83

L469L	0.82
L494N	0.81
L469Q	0.79
L494L	0.78
A453E	0.77
L469L	0.77
N454H	0.76
A453L	0.76
A453H	0.75
L469G	0.75
L494Q	0.74
N454A	0.73
L494P	0.73
N454T	0.73
L469L	0.73
N454E	0.72
A453I	0.70
L494A	0.69
N454Y	0.67
N454W	0.66
A453V	0.64
N454V	0.63
A453F	0.62
N454Q	0.62
N454L	0.61
A453R	0.59
N454D	0.56
A453D	0.54
N454C	0.53
N454F	0.49
A453K	0.45
L494T	0.42
N454I	0.31
A453G	0.26
N454M	0.22
L469Y	0.19
A453P	0.02
N454P	0.00
A453Y	ND
N454R	-0.03

Table 5. Growth Index Ranking for Plate 006.

Variant	GI		
A519W	1.27	E525N	0.73
E525K	1.13	E525Q	0.72
H515Y	1.07	R528M	0.70
R528K	1.05	H515N	0.70
E525E	1.01	E525A	0.69
A519K	0.99	R528F	0.69
H515K	0.98	E525E	0.67
H515W	0.97	R528H	0.67
H515R	0.96	A519S	0.66
E525W	0.94	A519L	0.66
R528R	0.92	H515E	0.66
R528R	0.92	E525S	0.64
A519A	0.91	H515S	0.63
E525L	0.89	A519T	0.61
A519R	0.89	R528Y	0.61
A519A	0.87	R528T	0.60
R528V	0.85	R528L	0.59
A519A	0.85	H515A	0.58
E525V	0.85	A519C	0.50
A519G	0.85	R528A	0.49
H515G	0.85	A519V	0.45
A519A	0.85	A519Q	0.44
R528R	0.85	R528S	0.43
E525R	0.83	A519E	0.39
E525E	0.83	E525P	0.37
E525D	0.83	R528N	0.35
E525T	0.82	R528E	0.34
E525C	0.81	A519D	0.33
E525G	0.80	R528G	0.33
E525M	0.79	R528C	0.31
H515F	0.78	R528W	0.30
H515M	0.78	H515L	0.15
H515V	0.78	R528D	0.08
E525F	0.77	A519P	0.00
A519H	0.76	H515H	-0.01
A519F	0.76	H515H	ND
H515Q	0.73	H515P	-0.02
A519Y	0.73	R528P	-0.02
H515T	0.73	H515H	ND
E525H	0.73	H515H	ND

Table 5. Growth Index Ranking for Plate 007.

Variant	GI
T536L	0.62
T536T	0.59
T536Y	0.58
T536I	0.49
T536H	0.46
T536F	0.45
T536G	0.42
T536T	0.40
T536V	0.38
T536S	0.36
T536K	0.31
T536T	0.30
T536M	0.29
T536N	0.29
T536C	0.27
T536A	0.26
T536D	0.18
T536R	0.15
T536E	0.14
T536P	0.01

Table 5. Growth Index Ranking for Plate 008.

Variant	GI		
G78L	1.33	S74Q	0.77
G78Y	1.31	G78P	0.76
G78W	1.27	S74P	0.76
G78I	1.26	G78E	0.74
S74K	1.24	S74N	0.72
K36L	1.23	G78D	0.72
S74W	1.22	K36N	0.72
S74I	1.20	S74E	0.72
K36K	1.20	S74D	0.70
K36I	1.19	S74A	0.70
G78V	1.19	G78M	0.70
K36V	1.19	A7A	0.69
G78H	1.18	K36C	0.69
G78T	1.18	G78N	0.68
S74V	1.17	G78Q	0.68
K36Y	1.17	G78A	0.66
S74L	1.14	S74M	0.65
S74T	1.14	K36Q	0.65
S74Y	1.13	K36P	0.61
G78K	1.13	K36D	0.58
K36R	1.11	K36A	0.56
S74H	1.10	K36E	0.54
K36T	1.07	A7W	0.50
G78F	1.01	A7V	0.39
K36W	1.01	A7C	0.34
S74G	0.98	A7H	0.31
K36S	0.98	A7P	0.31
G78R	0.97	A7Y	0.30
S74R	0.97	A7S	0.30
K36H	0.97	A7T	0.29
K36F	0.95	A7I	0.29
S74S	0.94	A7F	0.20
G78G	0.91	A7Q	0.19
G78S	0.91	A7E	0.17
S74F	0.86	A7N	0.16
G78C	0.84	A7L	0.09
K36G	0.83	A7D	0.07
S74C	0.78	A7R	0.07
K36M	0.77	A7M	0.06
A7G	0.77	A7K	-0.02

Table 5. Growth Index Ranking for Plate 009.

Variant	GI		
R44H	1.30	R44E	0.82
R44T	1.27	Q216C	0.81
R44F	1.21	R44Q	0.80
G121I	1.17	R44W	0.79
R44V	1.17	R44P	0.78
G121L	1.15	Q216G	0.78
G121H	1.14	R44R	0.78
R44Y	1.13	E488A	0.77
R44I	1.10	E488L	0.76
G121F	1.09	G121D	0.75
G121T	1.07	Q216E	0.75
G121Y	1.07	G121S	0.74
R44K	1.04	Q216S	0.74
G121V	1.03	G121N	0.74
G121K	1.02	G121Q	0.72
G121W	1.00	E488V	0.72
R44C	1.00	Q216M	0.72
R44A	1.00	Q216N	0.71
Q216H	0.99	G121E	0.71
R44M	0.99	E488W	0.70
R44L	0.99	Q216D	0.70
R44D	0.96	Q216Q	0.68
R44N	0.96	E488F	0.67
R44S	0.95	E488C	0.66
Q216I	0.94	E488E	0.66
G121A	0.94	E488G	0.66
Q216T	0.92	E488D	0.65
G121G	0.91	E488T	0.65
Q216V	0.90	E488H	0.64
G121M	0.90	E488Q	0.61
Q216K	0.89	Q216P	0.60
G121C	0.88	Q216R	0.60
E488I	0.88	E488R	0.54
Q216F	0.87	R44G	0.53
Q216L	0.87	E488M	0.52
G121R	0.87	E488Y	0.44
Q216W	0.86	E488S	0.43
Q216Y	0.84	E488K	0.32
Q216A	0.83	E488N	0.25
G121P	0.83	E488P	0.17

Table 5. Growth Index Ranking for Plate 010.

Variant	GI		
E179L	1.46	T251G	0.92
E179I	1.40	G177S	0.92
T251T	1.36	T251Q	0.91
E179K	1.35	G177R	0.90
E179H	1.33	R246S	0.90
E179W	1.26	G177C	0.88
G177V	1.26	R246N	0.87
G177T	1.25	R246Q	0.85
E179T	1.23	E179D	0.84
G177L	1.22	E179P	0.82
E179V	1.22	G177F	0.81
G177I	1.18	R246E	0.80
R246R	1.17	E179N	0.79
E179Y	1.14	T251E	0.79
G177H	1.14	R246D	0.79
G177K	1.13	G177D	0.79
R246K	1.13	T251D	0.78
E179F	1.11	T251C	0.77
G177P	1.11	T251M	0.76
T251N	1.10	R246A	0.75
T251Y	1.09	E179Q	0.74
E179S	1.07	G177E	0.74
E179G	1.05	T251A	0.74
T251H	1.04	T251W	0.73
R246T	1.04	T251V	0.71
G177Y	1.03	T251F	0.71
R246H	1.03	T251L	0.69
G177A	1.03	R246M	0.65
T251K	1.01	R246Y	0.61
G177M	1.01	R246C	0.61
E179M	1.01	R246L	0.54
G177W	1.00	R246V	0.53
T251S	1.00	R246W	0.53
E179A	0.96	T251I	0.52
T251R	0.94	G177Q	0.44
E179R	0.94	R246I	0.40
E179C	0.94	T251P	0.37
R246G	0.94	R246F	0.30
G177G	0.93	R246P	0.01
G177N	0.93	E179E	0.00

Table 5. Growth Index Ranking for Plate 011.

Variant	GI		
H254H	1.32	F287M	0.26
H254K	1.20	F287V	0.21
L308H	1.19	L308P	0.20
L308I	1.18	L308A	0.19
H254F	1.17	F287Y	0.19
H254V	1.16	V290C	0.18
H254W	1.15	L308T	0.17
H254I	1.12	V290A	0.17
H254T	1.11	V290S	0.16
H254R	1.07	V290G	0.15
L308W	1.07	L308K	0.15
L308L	1.05	F287A	0.10
H254Y	0.97	F287H	0.09
L308R	0.94	F287K	0.08
V290I	0.92	L308F	0.06
L308Y	0.92	V290H	0.04
V290V	0.89	F287T	0.04
H254G	0.89	V290K	0.04
H254L	0.88	F287Q	0.03
H254S	0.87	V290W	0.03
L308C	0.85	F287S	0.02
H254D	0.85	V290Y	0.02
L308G	0.84	L308M	0.02
H254E	0.81	F287G	0.02
H254A	0.80	F287I	0.01
H254C	0.78	F287N	0.01
L308D	0.76	F287R	0.01
F287F	0.74	V290R	0.00
L308N	0.71	H254P	ND
L308E	0.71	V290F	0.00
L308S	0.69	F287C	0.00
H254M	0.68	F287P	-0.01
H254Q	0.68	V290M	-0.01
L308Q	0.65	V290Q	-0.02
L308V	0.59	F287D	-0.02
H254N	0.59	V290E	-0.02
V290T	0.59	F287E	-0.02
V290L	0.48	V290N	-0.03
F287W	0.42	V290D	-0.03
F287L	0.39	V290P	-0.03

Table 5. Growth Index Ranking for Plate 012.

Variant	GI		
Q421R	1.39	T426Y	0.48
Q421H	1.26	T426P	0.46
Q421E	1.24	F434V	0.46
T426I	1.19	F434M	0.45
Q421G	1.16	F434T	0.44
Q421V	1.13	T426K	0.40
Q421K	1.12	T426R	0.38
P430S	1.10	P430G	0.30
P430T	1.10	F434F	0.24
Q421I	1.09	F434C	0.23
Q421T	1.08	F434A	0.17
Q421Q	1.08	F434Y	0.14
P430V	1.06	F434S	0.10
T426V	1.05	F434H	0.09
Q421W	1.05	F434K	0.05
Q421L	1.05	F434N	0.05
Q421M	0.92	T426T	0.04
Q421A	0.91	F434Q	0.04
Q421Y	0.91	F434G	0.03
T426S	0.86	P430H	0.02
T426G	0.86	P430I	0.02
Q421D	0.86	P430W	0.02
T426H	0.86	F434P	0.01
Q421S	0.83	P430K	0.01
T426F	0.82	P430Q	0.01
T426L	0.82	Q421F	0.01
P430A	0.78	F434E	0.01
Q421N	0.77	F434D	0.01
T426C	0.76	F434W	0.01
Q421P	0.74	Q421C	0.00
T426W	0.73	P430R	0.00
P430C	0.72	F434R	0.00
T426A	0.68	P430M	0.00
T426M	0.66	P430D	0.00
T426D	0.66	P430E	0.00
T426N	0.60	P430F	0.00
T426Q	0.59	P430P	0.00
T426E	0.59	P430L	0.00
F434I	0.55	P430Y	0.00
F434L	0.48	P430N	-0.01

Table 5. Growth Index Ranking for Plate 013.

Variant	GI		
A445H	1.18	A445I	0.42
F403V	1.06	G397A	0.40
G397G	1.05	F403Q	0.38
F403T	1.03	Q400S	0.27
F403I	0.93	Q400V	0.20
A445S	0.92	F403E	0.18
F403Y	0.91	G397V	0.16
F403L	0.91	G397I	0.09
A445K	0.86	Q400G	0.09
F403S	0.81	G397M	0.08
F403H	0.78	F403W	0.05
F403F	0.76	G397C	0.05
A445Q	0.74	Q400E	0.02
Q400Q	0.74	Q400A	0.01
F403A	0.72	F403D	0.01
F403G	0.71	Q400P	0.01
F403M	0.70	Q400I	0.00
A445T	0.70	G397Q	0.00
A445R	0.68	G397R	0.00
A445A	0.68	G397P	0.00
F403C	0.67	G397S	0.00
A445E	0.67	G397Y	0.00
A445C	0.65	G397N	0.00
A445N	0.64	F403P	-0.01
A445G	0.62	F403K	-0.01
A445D	0.62	G397L	-0.01
A445F	0.61	Q400R	-0.01
Q400L	0.61	Q400F	-0.01
A445M	0.60	Q400K	-0.02
F403N	0.58	G397E	-0.02
A445L	0.53	Q400D	-0.02
A445Y	0.53	G397F	-0.02
A445P	0.52	G397H	-0.02
Q400H	0.52	F403R	-0.02
A445V	0.51	Q400Y	-0.02
Q400T	0.51	G397D	-0.02
Q400C	0.46	Q400W	-0.03
A445W	0.44	G397K	-0.03
Q400N	0.43	G397W	-0.03
Q400M	0.43	G397T	-0.03

Table 5. Growth Index Ranking for Plate 014.

Variant	GI		
D33T	1.49	G99W	1.03
D33V	1.46	G99F	1.02
F53L	1.44	D33G	0.97
G99T	1.43	F53Q	0.96
D33Y	1.39	G99Q	0.94
K50K	1.35	K50S	0.91
F53T	1.34	F53E	0.90
K50I	1.33	F53P	0.87
K50L	1.33	D33E	0.83
G99V	1.32	K50E	0.83
D33W	1.30	D33P	0.82
F53W	1.27	D33D	0.81
G99I	1.27	G99P	0.80
F53V	1.26	G99E	0.78
F53I	1.26	K50G	0.77
D33H	1.26	K50Q	0.74
G99K	1.24	F53D	0.71
F53Y	1.24	K50P	0.70
D33I	1.23	G99D	0.70
G99Y	1.23	K50D	0.68
G99L	1.23	D33N	0.64
D33S	1.22	G99N	0.57
K50W	1.22	D33C	0.50
D33K	1.21	G99C	0.46
D33R	1.18	F53C	0.44
K50H	1.18	F53N	0.43
K50Y	1.17	K50C	0.43
F53R	1.17	G99M	0.43
K50T	1.14	K50N	0.35
D33L	1.13	G99H	ND
F53S	1.12	K50A	0.27
K50V	1.11	G99A	0.22
G99S	1.11	F53A	0.16
G99R	1.09	K50M	0.11
F53K	1.08	D33A	0.04
D33Q	1.07	K50R	ND
D33F	1.07	G99G	ND
F53F	1.06	K50F	ND
F53H	1.05	D33M	-0.03
F53G	1.05	F53M	-0.07

Table 5. Growth Index Ranking for Plate 015.

Variant	GI		
D23H	1.04	E26V	0.65
D23V	0.96	E26K	0.64
D23T	0.94	S27T	0.64
E26G	0.93	D81Q	0.63
D81T	0.93	S27M	0.63
D23F	0.92	D81W	0.62
D23S	0.91	D81R	0.62
D23I	0.89	D81K	0.61
D23K	0.88	S27G	0.59
D23W	0.88	S27C	0.59
D23E	0.86	S27V	0.58
D81V	0.85	E26L	0.58
D81H	0.85	E26T	0.55
D81L	0.84	E26Q	0.55
D81F	0.83	E26E	0.54
D23G	0.82	S27N	0.54
D23L	0.82	S27S	0.54
D23D	0.82	S27R	0.54
D23R	0.81	D23A	0.53
D81A	0.81	E26S	0.52
D23C	0.81	S27E	0.51
D81S	0.80	D81E	0.51
S27H	0.80	S27A	0.49
D81Y	0.80	E26N	0.49
D81G	0.80	E26P	0.49
S27L	0.78	E26D	0.47
D23P	0.78	E26M	0.46
D23M	0.77	S27F	0.46
D23N	0.76	E26R	0.46
D23Y	0.75	E26C	0.45
S27Q	0.73	S27D	0.44
D81C	0.71	D81P	0.42
D81M	0.70	S27P	0.41
D81N	0.69	S27Y	0.38
D81I	0.69	E26W	0.34
E26H	0.68	E26F	0.33
S27K	0.67	S27W	0.30
S27I	0.67	E26Y	0.29
D23Q	0.66	E26A	0.24
E26I	0.65	D81D	0.01

Table 5. Growth Index Ranking for Plate 016.

Variant	GI		
G69S	1.21	Y16H	0.86
G69M	1.16	Y16F	0.85
L20M	1.14	G69D	0.84
G69G	1.13	Y18R	0.83
G69K	1.12	Y18F	0.82
G69T	1.11	Y18T	0.80
G69I	1.10	Y18V	0.78
G69L	1.10	Y16T	0.77
Y16M	1.09	L20H	0.77
Y16I	1.09	Y18M	0.76
L20V	1.08	L20N	0.76
G69V	1.08	Y18L	0.75
L20S	1.08	L20G	0.75
G69H	1.08	L20Q	0.75
G69R	1.06	Y18Q	0.74
G69Q	1.05	L20C	0.74
Y16L	1.05	Y16W	0.73
G69N	1.04	Y18S	0.72
Y18G	1.03	G69P	0.72
L20Y	1.03	Y18D	0.67
G69A	1.03	L20E	0.66
L20L	1.03	Y18K	0.65
Y16Y	1.01	Y18E	0.61
L20T	1.01	Y16E	0.53
Y18C	1.00	L20R	0.53
L20W	1.00	Y16C	0.50
G69E	1.00	Y18P	0.49
G69C	0.99	Y16Q	0.48
Y18H	0.97	Y16P	0.48
L20F	0.97	Y18A	0.45
L20I	0.96	L20D	0.45
G69F	0.96	Y16S	0.44
Y18N	0.96	Y16N	0.41
Y18Y	0.95	Y16D	0.40
L20A	0.95	Y16K	0.39
G69Y	0.94	L20K	0.36
G69W	0.92	L20P	0.36
Y18I	0.91	Y16G	0.36
Y18W	0.91	Y16A	0.34
Y16V	0.91	Y16R	0.21

Table 5. Growth Index Ranking for Plate 017.

Variant	GI		
A3T	1.39	S13A	0.57
A3H	1.21	Y9W	0.56
A3K	1.13	S13C	0.56
A3F	1.10	A3S	0.51
A3Y	1.09	S13E	0.50
A3I	1.05	N12H	0.48
S13L	1.04	N12I	0.44
A3R	1.04	N12P	0.43
A3Q	0.97	N12R	0.41
A3E	0.94	N12Q	0.38
S13T	0.90	S13D	0.35
A3N	0.89	N12G	0.33
S13H	0.87	Y9H	0.31
S13K	0.85	N12K	0.29
Y9F	0.84	N12D	0.27
N12S	0.83	N12L	0.25
A3L	0.83	N12F	0.25
S13I	0.82	Y9I	0.22
S13V	0.81	Y9V	0.20
A3D	0.79	S13P	0.18
N12T	0.79	A3M	0.16
A3G	0.78	Y9T	0.15
S13R	0.76	N12E	0.15
S13G	0.76	A3V	0.14
A3P	0.76	N12W	0.14
N12C	0.73	Y9R	0.11
A3A	0.73	Y9P	0.10
S13Y	0.72	Y9L	0.09
N12A	0.72	Y9A	0.09
A3W	0.70	Y9S	0.08
S13M	0.69	Y9C	0.05
S13S	0.68	Y9N	0.04
A3C	0.68	Y9Y	0.04
N12N	0.68	Y9K	0.03
S13N	0.67	Y9Q	0.03
S13F	0.67	Y9M	0.02
S13W	0.64	Y9G	0.01
S13Q	0.62	Y9D	-0.01
N12M	0.61	Y9E	-0.01
N12V	0.59	N12Y	-0.01

Table 5. Growth Index Ranking for Plate 018.

Variant	GI		
A139T	1.29	Q197C	0.87
Q197T	1.19	G127W	0.86
G127T	1.19	Q197R	0.86
R194H	1.16	G127D	0.85
Q197V	1.15	A139M	0.84
G127H	1.13	R194W	0.84
G127F	1.08	Q197G	0.83
A139P	1.07	Q197A	0.83
A139H	1.07	A139A	0.83
A139S	1.07	A139N	0.83
A139V	1.07	Q197W	0.83
G127V	1.05	R194F	0.82
Q197H	1.04	Q197P	0.82
A139I	1.04	A139F	0.82
Q197I	1.04	R194K	0.81
A139C	1.03	G127M	0.80
G127I	1.03	A139W	0.80
G127L	1.03	Q197Q	0.75
Q197M	1.01	G127N	0.74
G127Y	1.01	R194C	0.72
Q197F	0.99	G127Q	0.71
A139Q	0.98	G127A	0.70
A139G	0.97	R194M	0.68
Q197Y	0.97	R194I	0.65
Q197S	0.97	R194Q	0.61
G127S	0.96	G127R	0.59
Q197E	0.95	G127K	0.58
A139E	0.95	A139Y	0.56
R194R	0.93	R194V	0.43
R194Y	0.92	A139R	0.30
Q197L	0.91	A139K	0.30
Q197N	0.90	R194A	0.27
Q197K	0.90	R194N	0.18
Q197D	0.89	R194T	0.15
A139L	0.88	R194S	0.08
G127E	0.88	R194E	0.03
A139D	0.87	R194D	0.02
G127P	0.87	G127G	0.00
G127C	0.87	R194P	-0.01
R194L	0.87	R194G	-0.03

Table 5. Growth Index Ranking for Plate 019.

Variant	GI		
L59H	1.96	Q125P	0.99
L59K	1.49	E117V	0.99
Q116V	1.47	L59Y	0.97
L59I	1.46	Q125N	0.97
E117I	1.42	Q116D	0.97
L59R	1.41	Q125V	0.96
Q116W	1.36	E117A	0.95
Q116T	1.36	E117Y	0.94
Q116Y	1.32	Q125A	0.93
E117L	1.29	Q125L	0.93
L59G	1.28	Q116E	0.92
E117W	1.27	Q116H	0.92
L59F	1.26	Q125F	0.92
L59A	1.25	Q116P	0.91
Q116I	1.22	Q116R	0.91
L59M	1.22	E117C	0.90
E117F	1.21	Q125C	0.90
L59C	1.21	Q125R	0.89
Q116F	1.20	Q125K	0.89
L59E	1.19	Q116G	0.86
L59T	1.18	E117T	0.86
L59D	1.12	Q125S	0.85
Q116L	1.12	Q125D	0.83
Q116K	1.09	E117N	0.82
L59L	1.09	Q125G	0.82
E117H	1.08	Q116Q	0.81
Q125M	1.06	Q116N	0.80
L59W	1.06	E117R	0.80
Q125I	1.05	E117S	0.80
L59N	1.05	Q116M	0.76
Q125W	1.05	L59S	0.74
L59V	1.05	Q125E	0.73
L59Q	1.04	E117G	0.70
Q125H	1.02	E117D	0.70
Q116S	1.02	Q125Q	0.69
Q125Y	1.01	E117Q	0.66
Q125T	1.01	E117K	0.55
Q116A	1.01	E117P	0.44
E117M	1.00	E117E	0.26
Q116C	0.99	L59P	0.03

Table 5. Growth Index Ranking for Plate 020.

Variant	GI		
Y377W	1.26	G389Q	0.80
G389K	1.24	K379N	0.79
L376L	1.24	K379C	0.78
Y377Y	1.23	K379M	0.77
L376I	1.20	K379E	0.77
G389T	1.18	Y377S	0.77
Y377V	1.17	G389E	0.77
G389L	1.16	G389A	0.75
G389H	1.13	G389D	0.73
G389V	1.12	K379D	0.70
K379R	1.12	Y377G	0.69
Y377H	1.12	K379F	0.69
Y377L	1.10	L376M	0.65
K379V	1.10	Y377P	0.61
L376Y	1.09	Y377K	0.60
K379T	1.09	L376V	0.60
G389R	1.06	Y377M	0.59
K379H	1.04	Y377D	0.58
G389I	1.03	L376W	0.57
Y377I	1.03	Y377N	0.57
G389Y	1.01	Y377E	0.56
G389M	1.01	Y377C	0.55
G389S	1.01	Y377A	0.54
Y377T	0.98	Y377Q	0.49
G389N	0.96	L376H	0.49
Y377F	0.95	L376T	0.48
K379L	0.95	L376Q	0.48
G389G	0.94	L376C	0.43
K379G	0.93	L376A	0.23
L376F	0.92	Y377R	0.23
K379W	0.92	L376K	0.22
G389F	0.88	L376S	0.17
K379S	0.88	L376N	0.16
G389W	0.87	L376G	0.14
K379I	0.86	L376R	0.12
K379A	0.86	L376P	0.11
K379P	0.83	L376D	0.08
G389C	0.83	L376E	0.08
G389P	0.82	K379Y	0.00
K379Q	0.81	K379K	-0.01

Table 5. Growth Index Ranking for Plate 021.

Variant	GI		
I165Y	1.46	I165G	0.99
E173H	1.25	V202W	0.98
V202H	1.24	I165W	0.98
I165H	1.24	I165S	0.95
E173W	1.23	I165N	0.95
E174I	1.23	E174M	0.94
E173T	1.23	E174A	0.94
E173V	1.20	E173A	0.94
E174H	1.20	V202C	0.93
E173L	1.19	V202L	0.93
E173I	1.19	V202M	0.93
V202T	1.18	V202A	0.92
E173Y	1.18	I165M	0.91
V202R	1.16	E173Q	0.90
E174L	1.16	I165Q	0.90
I165K	1.16	V202Q	0.90
E174V	1.15	E174G	0.90
V202I	1.15	E173N	0.88
E174T	1.14	V202N	0.88
I165T	1.14	E174S	0.87
E173F	1.14	E174N	0.86
E174K	1.13	E173P	0.86
E173K	1.12	V202S	0.85
I165F	1.11	V202E	0.85
E174F	1.11	V202D	0.84
V202K	1.09	E173C	0.83
E173R	1.08	E174C	0.83
E173G	1.08	I165E	0.83
E174R	1.07	E173E	0.81
E174W	1.05	V202G	0.81
I165R	1.05	I165C	0.79
I165I	1.04	E174Q	0.72
I165D	1.03	E174D	0.70
I165A	1.02	E174E	0.69
V202F	1.02	E174P	0.68
V202Y	1.02	I165P	0.67
E173M	1.01	E173D	0.65
E173S	1.01	I165V	0.61
I165L	1.00	V202P	0.35
E174Y	1.00	V202V	0.32

Table 5. Growth Index Ranking for Plate 022.

Variant	GI		
K489K	1.65	K489A	0.94
G491I	1.57	L490K	0.94
L490I	1.52	G491E	0.94
L490H	1.46	G491Q	0.93
G491H	1.41	K489F	0.93
L490V	1.41	G491G	0.91
I539V	1.40	L490R	0.90
G491V	1.36	L490A	0.90
G491L	1.36	I539S	0.90
K489W	1.35	I539H	0.89
L490W	1.34	I539C	0.88
G491W	1.33	L490C	0.88
G491T	1.33	K489S	0.88
G491Y	1.31	L490G	0.86
G491K	1.29	K489G	0.84
I539T	1.26	K489Y	0.83
L490L	1.25	L490D	0.78
K489R	1.24	I539A	0.77
L490T	1.21	L490E	0.75
K489I	1.21	K489C	0.68
I539L	1.21	I539Y	0.67
K489T	1.20	K489Q	0.66
I539I	1.17	I539F	0.65
G491R	1.16	I539R	0.64
G491M	1.14	I539W	0.64
K489V	1.14	I539Q	0.64
L490F	1.14	L490Q	0.62
G491F	1.13	K489D	0.62
K489H	1.08	K489N	0.61
G491A	1.08	K489E	0.58
G491C	1.05	I539P	0.53
K489L	1.05	K489P	0.50
L490M	1.04	I539E	0.36
G491S	1.03	I539N	0.32
I539K	1.00	I539G	0.28
G491N	1.00	K489M	0.23
L490S	0.98	G491P	0.21
G491D	0.96	L490P	0.17
I539M	0.96	I539D	-0.01
L490N	0.95	L490Y	-0.01

Table 5. Growth Index Ranking for Plate 023.

Variant	GI		
Q509T	1.49	Q509L	0.16
K487T	1.39	K487P	0.15
Q509V	1.34	T521M	0.13
Q509I	1.30	T521H	0.02
K487H	1.19	Y514H	0.01
K487K	1.18	Y514T	0.01
K487V	1.12	Q509F	0.01
K487L	1.11	Y514V	0.00
Y514Y	1.08	T521T	0.00
K487C	1.06	Q509Y	0.00
K487W	1.05	T521A	-0.01
Q509S	1.03	Y514I	-0.01
K487R	1.03	Y514W	-0.01
K487F	1.03	T521N	-0.01
K487A	0.99	Q509R	-0.01
K487G	0.97	Y514D	-0.02
K487S	0.93	Y514S	-0.02
Q509G	0.91	T521F	-0.02
Q509Q	0.90	T521C	-0.02
K487E	0.90	Q509P	-0.02
Q509C	0.90	T521Y	-0.02
K487I	0.90	Y514P	-0.02
K487M	0.87	T521K	-0.02
Q509A	0.86	Y514R	-0.02
K487Q	0.81	Y514A	-0.02
Q509M	0.80	Y514L	-0.02
T521G	0.77	Y514N	-0.03
T521E	0.77	T521I	-0.03
Q509N	0.74	T521R	-0.03
T521S	0.73	Q509W	-0.03
K487D	0.69	Y514Q	-0.03
K487Y	0.61	T521P	-0.03
T521L	0.60	Y514E	-0.03
Q509H	0.53	Y514K	-0.03
T521Q	0.51	Y514C	-0.03
K487N	0.44	T521W	-0.03
Q509E	0.40	T521D	-0.03
T521V	0.36	Y514F	-0.04
Q509D	0.28	Y514G	-0.04
Q509K	0.26	Y514M	-0.04

Table 5. Growth Index Ranking for Plate 024.

Variant	GI		
L540V	1.73	L540E	1.01
L540T	1.58	L540A	1.00
T462T	1.43	R544M	0.98
A448T	1.35	L540N	0.98
L540H	1.34	L540P	0.96
R544T	1.33	R544F	0.95
T462V	1.31	R544P	0.95
L540I	1.31	T462R	0.94
R544W	1.28	L540D	0.93
L540K	1.28	T462A	0.93
A448H	1.26	R544G	0.91
T462K	1.26	L540C	0.90
T462H	1.26	T462M	0.89
R544V	1.26	A448K	0.89
R544S	1.23	R544N	0.89
A448V	1.23	A448Y	0.88
L540Y	1.20	A448G	0.87
R544K	1.19	R544R	0.86
T462I	1.18	R544Q	0.86
L540W	1.18	A448L	0.85
A448R	1.15	T462C	0.84
T462W	1.14	R544E	0.84
T462Y	1.14	T462D	0.84
L540R	1.14	R544A	0.83
L540S	1.12	R544D	0.83
L540L	1.12	T462Q	0.83
R544H	1.11	A448C	0.83
A448S	1.11	A448Q	0.82
L540Q	1.09	T462N	0.81
L540M	1.09	A448N	0.79
T462F	1.08	A448M	0.78
T462S	1.06	A448F	0.78
L540G	1.05	A448A	0.77
R544I	1.05	A448D	0.74
R544C	1.04	A448W	0.72
T462L	1.04	A448E	0.72
L540F	1.03	A448P	0.70
R544L	1.02	T462P	0.60
T462G	1.02	R544Y	0.50
A448I	1.01	T462E	-0.08

Table 5. Growth Index Ranking for Plate 025.

Variant	GI		
P498H	1.26	P498M	0.86
S457S	1.23	P498D	0.86
A496H	1.20	P498E	0.86
A496T	1.17	K497R	0.85
S457H	1.16	K497A	0.84
P498T	1.16	A496M	0.84
A496I	1.14	S457K	0.84
K497T	1.13	K497W	0.82
K497V	1.12	A496G	0.82
P498R	1.12	A496P	0.82
K497K	1.12	S457N	0.82
P498I	1.11	K497D	0.81
P498K	1.11	P498Q	0.81
P498V	1.10	S457F	0.80
S457R	1.10	P498N	0.80
P498Y	1.10	A496E	0.80
P498L	1.09	A496A	0.76
S457T	1.09	A496Q	0.76
A496L	1.06	S457M	0.76
A496R	1.06	S457D	0.75
S457Q	1.06	K497Q	0.75
A496V	1.04	P498C	0.75
K497H	1.04	A496D	0.74
P498S	1.03	S457W	0.74
A496K	1.03	S457E	0.74
K497L	1.03	K497M	0.74
K497I	1.02	K497N	0.73
P498F	1.02	K497E	0.72
A496Y	1.01	A496C	0.70
P498G	1.01	K497P	0.69
K497Y	0.99	A496N	0.66
S457Y	0.99	S457G	0.65
A496W	0.98	P498A	0.64
A496F	0.97	K497C	0.64
K497G	0.96	S457L	0.62
K497F	0.94	S457V	0.59
A496S	0.93	S457C	0.47
K497S	0.92	S457I	0.43
P498W	0.90	S457A	0.33
P498P	0.88	S457P	0.14

Table 5. Growth Index Ranking for Plate 026.

Variant	GI		
G87K	1.34	S120E	0.88
S120L	1.34	G87A	0.88
G87T	1.24	D25G	0.88
G87L	1.24	S120D	0.88
S120I	1.20	D25F	0.88
S120K	1.18	S120G	0.87
G87V	1.17	D25W	0.86
G87Y	1.17	S120R	0.84
S120T	1.16	S120A	0.82
T240T	1.15	D25E	0.82
S120W	1.14	G87E	0.81
G87H	1.12	T240M	0.81
G87W	1.11	G87N	0.80
D25H	1.11	D25S	0.80
G87I	1.10	S120N	0.80
D25L	1.10	D25R	0.77
D25I	1.09	T240A	0.76
S120H	1.08	G87Q	0.75
S120Y	1.07	S120Q	0.74
D25T	1.06	D25Q	0.73
S120V	1.05	S120S	0.69
G87M	1.04	T240N	0.69
G87R	1.03	T240C	0.67
D25V	1.02	T240Q	0.67
S120F	1.01	D25C	0.66
D25K	1.00	S120P	0.60
D25Y	1.00	T240S	0.56
T240L	0.99	T240G	0.35
G87F	0.99	T240H	0.15
G87P	0.98	T240Y	0.09
D25P	0.97	T240E	0.08
S120M	0.96	T240W	0.08
G87S	0.96	T240F	0.05
T240I	0.96	T240D	0.03
T240V	0.95	T240K	0.02
G87D	0.94	T240P	0.01
D25N	0.93	T240R	0.01
D25D	0.93	G87G	-0.01
D25M	0.90		
D25A	0.89		
G87C	0.89		
S120C	0.88		

Table 5. Growth Index Ranking for Plate 027.

Variant	GI		
F495L	1.84	F495F	1.07
F495T	1.56	G492Y	1.07
S493T	1.53	G492R	1.06
G492A	1.48	S493C	1.04
F495Y	1.45	S493A	1.03
F495K	1.45	F495G	1.02
F495W	1.44	F495A	1.02
S493V	1.40	G492N	0.99
F495S	1.40	F495N	0.98
G492T	1.38	F495D	0.96
F495V	1.38	S493H	0.95
G492H	1.38	F495C	0.95
G492V	1.37	F495P	0.95
S493R	1.35	G492F	0.94
F495H	1.31	F495E	0.94
G492K	1.30	G492Q	0.93
S493W	1.29	S493P	0.93
F495R	1.28	G492M	0.89
S493E	1.27	G492P	0.89
G492I	1.26	G492S	0.86
S493L	1.26	N476Q	0.83
F495M	1.25	S493Q	0.83
G492E	1.25	N476S	0.82
G492D	1.24	S493F	0.81
N476Y	1.24	S493N	0.79
S493K	1.23	N476D	0.73
S493G	1.21	N476M	0.71
S493I	1.20	N476P	0.64
S493S	1.20	N476E	0.62
G492L	1.18	N476A	0.02
G492W	1.18	N476N	0.01
F495Q	1.16	S493D	0.00
F495I	1.15	N476I	-0.01
N476W	1.15	N476H	-0.01
N476V	1.14	N476G	-0.02
N476R	1.13	N476F	-0.04
G492G	1.11	N476C	-0.04
G492C	1.11	N476K	-0.05
N476T	1.09	N476L	-0.08
S493M	1.09		
S493Y	1.08		

Results/Discussion

[0344] Table 6 lists all variants identified that displayed growth indices of 1.2 or higher. Without being bound by theory, mutations at these positions may result in increased intracellular activity of IspS by several different means. Without being bound by theory, increased intracellular activity could be a result of one or a combination of any of the following properties of IspS: increased cellular viability, increased kcat, decreased Km, increased specific activity, increased solubility, decreased insolubility, improved ribosome binding, increased translation initiation rate, increased translation elongation rate, increased transcription initiation rate, increased transcription elongation rate, decreased secondary structure of DNA, decreased secondary structure of RNA, increased secondary structure of DNA, increased secondary structure of RNA, increased folding rates, increased affinity for intracellular chaperones, increased stability, decreased protein turnover, decreased exposure to intracellular protease, decreased affinity for intracellular protease, decreased localization to the periplasm, improved localization to the cytoplasm, decreased inclusion body formation, decreased membrane localization, increased expression due to a more favorable codon, increased DNA stability, increased RNA stability, and decreased RNA degradation.

[0345] Without being bound by theory, any mutation that has a positive effect on the properties of nucleic acid sequences (DNA and RNA) encoding or expressing IspS, or the biochemical properties of the IspS enzyme itself, could allow for greater activity within the cell. All variants with a growth index of 1.2 or higher are subjected to secondary growth assays in a matrix of mevalonic acid and IPTG. These variants are also pooled together and subjected to several rounds of enrichment under IPTG induction and mevalonic acid pathway flux to determine which enzymes allow for the best growth in competition experiments. The most promising variants are examined further for benefits to specific productivity in isoprene producing strains.

[0346] Residues listed in Table 7 are necessary for growth under high DMAPP pressure and thus, can be considered in one embodiment as immutable residues. Substitution of the wild type amino acid with any other residue results in minimal to no growth under the growth rate assay conditions. Growth index values for each

position are shown in Table 7. Phenylalanine 287 (F287) is located in the active site, and defines the bottom of the active site cavity (Figure 7). Based on structural alignments with other terpene synthases, F287 determines the length of the substrate that can be accommodated into the active site, thereby preventing access to the active site by isoprenoids with more than five carbons.

[0347] Glycine 397 (G397) is located on the side of the active site cavity (Figure 8). This residue occurs at a kink in an alpha-helix, suggesting that the conformational flexibility of glycine (and other small amino acids) may be required at this position to allow the helix to bend. The bend in the helix is adjacent to the putative substrate binding position in the active site.

[0348] Asparagine 438 (N438) is positioned at the top of the active site (Figure 9). Structural alignments with other terpene synthases indicate that N428 may be involved directly in coordination of the magnesium ions, as well as having possible interactions with the substrate.

[0349] Glutamate 451 (E451) is on a substrate access loop located above the active site (Figure 10). Based on homology modeling and structure-based alignments with other terpene synthases, these loops may have an open position for substrate capture, and then close over the active site once substrate is bound. Residue E451 is proposed to have a role in coordinating one or more magnesium ions during this process.

[0350] Tyrosine 514 (Y514) is in the active site, below N438 (Figure 11). Y514 may be involved in substrate binding, or it may play a direct role in catalysis.

Table 6. Variants with Growth Index value equal to or greater than 1.2.

Variant	GI
L59H	1.96
F495L	1.84
L540V	1.73
V30L	1.61
L540T	1.58
G491I	1.57
F495T	1.56
S493T	1.53
V30K	1.52
L490I	1.52
D33T	1.49
Q509T	1.49
L59K	1.49
G492A	1.48
Q116V	1.47
L59I	1.46
E179L	1.46
L490H	1.46
I28T	1.46
D33V	1.46
I165Y	1.46
F495Y	1.45
F495K	1.45
V30Y	1.45
F53L	1.44
F495W	1.44
G99T	1.43
T462T	1.43
E117I	1.42
G491H	1.41
L59R	1.41
L490V	1.41
S493V	1.40
E179I	1.40
I539V	1.40
F495S	1.40
D33Y	1.39
Q421R	1.39
V30W	1.39
A3T	1.39
K487T	1.39

G492T	1.38
F495V	1.38
G492H	1.38
I28S	1.38
G492V	1.37
A443A	1.36
Q116W	1.36
G491V	1.36
Q116T	1.36
G491L	1.36
A448T	1.35
S493R	1.35
K489W	1.35
E179K	1.35
G153W	1.34
L490W	1.34
L540H	1.34
G87K	1.34
S120L	1.34
F53T	1.34
Q509V	1.34
G78L	1.33
K50I	1.33
K50L	1.33
R544T	1.33
E179H	1.33
G491W	1.33
G491T	1.33
Q116Y	1.32
G99V	1.32
F495H	1.31
G491Y	1.31
G78Y	1.31
T462V	1.31
L540I	1.31
I28R	1.31
G492K	1.30
Q509I	1.30
D33W	1.30
R44H	1.30
G491K	1.29
A139T	1.29

L130Y	1.29
E117L	1.29
S493W	1.29
F495R	1.28
R544W	1.28
L540K	1.28
L59G	1.28
A519W	1.27
S493E	1.27
F53W	1.27
E117W	1.27
G99I	1.27
G78W	1.27
R44T	1.27
E179W	1.26
A448H	1.26
F53V	1.26
Y377W	1.26
F53I	1.26
G492I	1.26
T462K	1.26
T462H	1.26
S493L	1.26
G177V	1.26
G78I	1.26
P498H	1.26
L59F	1.26
D33H	1.26
V30S	1.26
Q421H	1.26
I539T	1.26
R544V	1.26
F495M	1.25
E173H	1.25
G492E	1.25
L59A	1.25
G177T	1.25
G87T	1.24
G99K	1.24
V202H	1.24
S74K	1.24
G389K	1.24
G87L	1.24
G492D	1.24
Q421E	1.24

N476Y	1.24
I165H	1.24
F53Y	1.24
K489R	1.24
K36L	1.23
D33I	1.23
R544S	1.23
E179T	1.23
A448V	1.23
G99Y	1.23
E173W	1.23
S493K	1.23
G99L	1.23
E174I	1.23
E173T	1.23
D33S	1.22
G177L	1.22
V30F	1.22
K50W	1.22
Q116I	1.22
E179V	1.22
S74W	1.22
L59M	1.22
D33K	1.21
L490T	1.21
A3H	1.21
K489I	1.21
I539L	1.21
R44F	1.21
E117F	1.21
L59C	1.21
S493G	1.21
G69S	1.21
E173V	1.20
L540Y	1.20
S493I	1.20
S120I	1.20
Q116F	1.20
A496H	1.20
H254K	1.20
S74I	1.20
K489T	1.20
L376I	1.20
E174H	1.20

Table 7. Growth index values for the immutable sites.

Variant	GI
F287	
F287A	0.10
F287C	0.00
F287D	-0.02
F287E	-0.02
F287F	0.74
F287G	0.02
F287H	0.09
F287I	0.01
F287K	0.08
F287L	0.39
F287M	0.26
F287N	0.01
F287P	-0.01
F287Q	0.03
F287R	0.01
F287S	0.02
F287T	0.04
F287V	0.21
F287W	0.42
F287Y	0.19

G397M	0.08
G397N	0.00
G397P	0.00
G397Q	0.00
G397R	0.00
G397S	0.00
G397T	-0.03
G397V	0.16
G397W	-0.03
G397Y	0.00

Variant	GI
G397	
G397A	0.40
G397C	0.05
G397D	-0.02
G397E	-0.02
G397F	-0.02
G397G	1.05
G397H	-0.02
G397I	0.09
G397K	-0.03
G397L	-0.01

Variant	GI
N438	
N438A	-0.02
N438C	-0.05
N438D	-0.04
N438E	-0.04
N438F	-0.04
N438G	-0.07
N438H	-0.04
N438I	-0.05
N438K	-0.06
N438L	-0.05
N438M	-0.05
N438N	1.27
N438P	-0.06
N438Q	-0.06
N438R	-0.06
N438S	-0.09
N438T	-0.04
N438V	-0.05
N438W	-0.05
N438Y	-0.06

Variant	GI
E451	
E451A	-0.04
E451C	-0.04
E451D	-0.02
E451E	1.48
E451E	1.13
E451F	-0.03
E451G	-0.02
E451H	-0.02
E451I	-0.03
E451K	-0.01
E451L	-0.05
E451N	-0.06
E451P	-0.05
E451Q	0.06
E451R	-0.03
E451S	-0.06
E451T	-0.04
E451V	-0.04
E451W	-0.01
E451Y	-0.03

Variant	GI
Y514	
Y514A	-0.02
Y514C	-0.03
Y514D	-0.02
Y514E	-0.03
Y514F	-0.04
Y514G	-0.04
Y514H	0.01
Y514I	-0.01
Y514K	-0.03
Y514L	-0.02
Y514M	-0.04
Y514N	-0.03
Y514P	-0.02
Y514Q	-0.03
Y514R	-0.02
Y514S	-0.02
Y514T	0.01
Y514V	0.00
Y514W	-0.01
Y514Y	1.08

Example 3: Primary Specific Activity Assay for *P. alba* IspS Complete Site Evaluation Libraries (SELs)

[0351] Site Evaluation Libraries (SELs) of the entire *P. alba* isoprene synthase (MEA *P. alba*) backbone (544 amino acids) were built in the parent vector pCL201 (Figure 4), and screened for specific activity to identify isoprene synthase (IspS) molecules with improved characteristics. In most cases, the SEL at a given position contained all 20 possible amino acid substitutions including the wild type. Numbering of each library corresponds to the ORF of MEA *P. alba* (SEQ ID NO:1), where the starting methionine is position 1.

Individual strains, built in the MD09-170 background, that contained variants for expression were arrayed into microtiter plates such that each well corresponded to a specific amino acid substitution at the given position in MEA *P. alba*. Microtiter plates contained four SELs, or four positions in MEA *P. alba* with all possible substitutions. Remaining wells were used for control strains. Plates were grown, induced, and lysed, in order to measure the amount of isoprene produced per the specific amount of IspS protein in each sample. Specific activity values were calculated for all variants in the entire set of SELs.

Methods

Cell Growth and Lysis

[0352] Glycerol stocks of MEA *P. alba* IspS libraries were thawed briefly and inoculated into microtiter plates (Cellstar) containing liquid LB medium with kanamycin at a concentration of 20 ug/ml. Cultures were grown overnight at 250 rpm, 30°C to saturation in a shaking incubator using the EnzyScreen clamp system (EnzyScreen). The next day, cultures were removed and inoculated using a Liquidator96 pipettor (Rainin Instruments) at a ratio of 1:10 into TM3-glucose medium containing 50 ug/ml kanamycin and 50 uM IPTG. Wild type controls were grown separately and inoculated into each microtiter plate containing TM3-glucose with a titration of IPTG concentrations, from 30 uM to 65 uM, in separate wells. Plates were returned to the shaking incubator at 250 rpm, 30°C and induced for five hours. Plates were then removed from the incubator and cultures were harvested into polypropylene microtiter plates (Nunc) by centrifugation at 3700 rpm for 20 minutes at 4° C in a tabletop centrifuge. The supernatant was removed and pellets were stored at -80°C prior to lysis, DMAPP assay, and protein determination.

[0353] Prior to cell lysis, plates were removed from the -80°C freezer and thawed on the bench for 10 minutes. Pellets were thoroughly resuspended in 200 ul of lysis buffer (100 mM Tris, 100 mM NaCl pH 7.6 buffer, 1 mg/ml BSA, 50U/ul Epicentre readylyse lysozyme, 0.1 mg/ml DNase, 0.5 mM PMSF/AEBSF, 5 mM MgCl₂) using a Biomek automated workstation (Beckman Coulter), removed, and shaken at room temperature at 450 rpm for 30 minutes. Lysates were then spun at 3200 rpm at 4°C for 10 minutes, and 150 ul of the supernatant was transferred using a Biomek to a new microtiter plate for DMAPP and dot blot assay.

DMAPP Assay

[0354] For the DMAPP assay, 25 ul of lysate was added to 75 ul of DMAPP assay buffer (100mM Tris/100mM NaCl pH 7.6, 1 mg/ml BSA, 50mM MgCl₂, 1mM DMAPP) in a 96-well glass block (Zinser) using a Liquidator96 pipettor (Rainin Instruments). Glass blocks were sealed with aluminum foil seals (Beckman Coulter) and incubated at 450 rpm for one minute at room temperature. Blocks were then incubated at 34°C in a water bath for 30 minutes, and the reaction was stopped by incubation at 70°C for two minutes. Blocks were cooled briefly before loading onto the GC-MS.

[0355] Sealed glass blocks were loaded onto an Agilent 7890a Gas Chromatography (GC) System equipped with a flame ionization detector (FID) and a CTC CombiPAL autosampler. The GC FID method parameters are described below:

Column: ZB-5ms

Dimensions: 15m x 0.25mm x 0.25 μ m

Oven:

Ramp (°C/min)	Temperature (°C)	Hold Time (min)
0	37	28

Total Run Time:	28 minutes
Front Inlet Temperature:	110°C
Split Ratio:	50:1
Flow Rate:	3.4 mL/min
Injection Volume:	100 μ L
Headspace Syringe Volume:	1 mL
Detector Temperature	160°C
Hydrogen Flow:	40 mL/min
Air Flow:	400 mL/min
Makeup Flow:	0.1 mL/min
Makeup Gas Type:	Helium

[0356] Chemstation software (version E.02.00.493) was used to control the GC and Cycle Composer software (version 1.5.2) was used to control the CTC autosampler. The Cycle Composer software was programmed to continuously inject one sample after another in sequence for a total of 48 injections. 0.2% v/v isoprene balanced with nitrogen gas from Air Liquide was used as the standard for determining calibration response factors. Three separate 2 mL vials were filled with the calibration gas and analyzed using the method described above to determine an average response factor. Calculated response factors allowed for the conversion of individual sample peak area counts to isoprene concentrations using Microsoft Excel.

Protein Determination

[0357] Prior to the protein determination assay, several wild type samples from each plate were analyzed by GC-MS for isoprene, and protein concentration was back-calculated from the known specific activity of MEA *P. alba* to determine the average amount of IspS for all samples in the microtiter plate. For the dot blot assay, nitrocellulose membranes (Invitrogen) were soaked in 1X PBS buffer (10mM Sodium Phosphate, 150mM NaCl, PH7.8 +/- 0.2) and equilibrated for at least 5 minutes. Lysates were then diluted in 1X PBS using a Hamilton

MicroLab STAR liquid handling workstation to achieve loading concentration between 0.025-0.5ug of *P. alba* IspS. Purified standards were added at concentrations between 0.025-1 ug. The blotting unit (Minifold-1, Whatman) was assembled according to the manufacturer's recommended protocol. Vacuum was applied briefly to remove excess 1X PBS buffer. Samples (approximately 200 ul of each) were transferred to the Minifold-1, and vacuum was applied at 20 kPa. After samples were filtered completely, wells were washed once with 200 ul of 1X PBS buffer. After the wash buffer passed completely through the membrane, the vacuum was removed, and membranes were removed carefully with forceps, labeled, and dried on clean filter paper.

[0358] Immunodetection of *P. alba* IspS molecules at each position on nitrocellulose membranes was carried out using the WesternBreeze kit from Invitrogen. Primary monoclonal or polyclonal antibody (anti-mouse against purified *P. alba* IspS, Prosci Incorporated) was diluted 1:5000 in blocking solution, and secondary antibody (Alexa Fluor 488 goat anti-mouse IgG (H+L), Invitrogen) was diluted to a concentration of 2 ug/ml in blocking solution. Fluorescent spots were quantified using a Storm 860 Molecular Imager (GMI, Inc.) and ImageQuant software (GE Healthcare), according to the manufacturers' recommended protocols, and specific protein concentrations for each sample were determined by comparison to known standards using Microsoft Excel.

Results

[0359] Specific activity values were calculated for every variant in the entire set of SELs by dividing the molar amount of isoprene produced in a given amount of time by the specific amount of protein in each sample. Performance index (PI) was calculated by dividing the specific activity of any given variant by the average of several WT specific activity measurements from the same microtiter plate. A variant that displayed a PI value of 1.5 for specific activity, for example, was 50% improved over WT. PIs for protein concentration and isoprene produced were also calculated in the same fashion, and these measurements were used for detailed data analysis.

[0360] Table 8 provides precise definitions for locations of the residues listed in Tables 9 and 10. For example, a residue listed as "N-term" in Table 9 or 10, is between residue 1 and 215 of reference sequence MEA *P. alba* IspS (SEQ ID NO:1).

Table 8. Definitions of locations of MEA *P. alba* IspS amino acid positions

Location	Definition
N-term	Residues 1-215
Hinge Region	Residues 216-245
C-term	Residues 246-544
N-term helices	Residues 134-179
Dimer Interface	Residues 240-255 and 316-353
Substrate Capture Loops	Residues 441-454 and 515-527
Active Site	Residues within 8Å of active site based on PDB 3N0G
Buried	Surface Accessibility below 35% and located internally
Surface	Residues located on the surface of the structure
Surface Loop	Residues in loops located on the surface of the structure

[0361] Surface accessibilities and putative functions of amino acids of interest in the structure of MEA *P. alba* are also listed in Tables 9 and 10. Surface accessibility was calculated using the program MOE, which is written and supported by the Chemical Computing Group, Inc. An estimate for the water-accessible surface area of each residue was determined using a probe having a specified radius. The estimate was then compared against a library of peptides and the ratio between them was reported as the percent surface accessibility. Tables 9 and 10 also list the putative functions of each residue. For example, functions include but aren't limited to metal binding (in the active site), substrate capture, altered loop shape, alternative interactions in a pocket, and dimer formation.

[0362] From the primary data, positions in MEA *P. alba* which tolerated no changes from the wild type residue were identified (see Table 9, Figures 13-19). MEA *P. alba* variants with amino acid substitutions other than wild type at these positions displayed specific activity no higher than 30% ($PI \leq 0.3$) of the value of WT, and were therefore functionally inactive. Wild type residues at these positions represent a minimum set that is required for the efficient conversion of DMAPP to isoprene by MEA *P. alba*. Many of these positions map to or near the active site of MEA *P. alba* (see Figure 14), and are putatively involved in but are not limited to metal binding (for substrate orientation), substrate capture, substrate binding, and catalysis. Figure 15, for example, shows a position that has an unknown role in enzyme function. Figure 16 shows positions that may be involved in IspS dimer formation, and Figures 17 and 18 show positions that are in or interact with the N-terminus and may be involved in loop closure or function of the active site. Figure 19 shows positions located in the substrate capture loops that tolerate no substitutions.

[0363] Variants that displayed specific activity higher than WT in the primary in vitro assay were selected for retesting. Variants were tested following the methods described above, except that polyclonal antibody in addition to monoclonal antibody was used for immunodetection, according to standard biochemical practices. Table 10 lists a set of retested variants that showed higher specific activity (a PI > 1.3) than WT. Figure 20 shows all positions in the crystal structure of IspS where variants displayed increased specific activity upon retest. In comparison to the wild type enzyme, these variants confer specific activity benefits to IspS by alteration/enhancement of the putative functions listed in Table 10. Figure 21 shows buried positions in or that interact with the N-terminus of IspS and where variants displayed increased specific activity. Figure 22 shows buried positions that are in or interact with the C-terminus of IspS, and where variants displayed increased specific activity. Figure 23 shows position 247, where improved variants may positively affect dimerization of IspS, and Figure 24 shows additional sites at the N-terminus where variants display clear specific activity benefits. Figure 25 shows the locations of positions on surface loops where variants displayed high specific activity, and Figure 26 shows positions on the surface of the enzyme that are not in loops, and where variants displayed increased specific activity. Figure 27 shows positions in proposed substrate capture loops where variants displayed increased specific activity relative to wild type. Particular positions in this region have variants which displayed increased activity, whereas adjacent positions are immutable (see Table 9 and Figure 19). This indicates that the proposed “substrate capture loops” in IspS are critical in the enzymatic conversion of DMAPP to isoprene, and are highly sensitive to perturbation, which can result in either negative or positive effects on activity. All variants listed in Table 10 and shown in Figures 20 through 27 or any combination thereof represent mutations in IspS that allow the enzyme to more efficiently convert DMAPP to isoprene.

Table 9. Positions in *MEA P. alba* displaying PI specific activity values ≤ 0.3 for all non-WT amino acid substitutions.

Residue	Position	Location	% Surface Accessibility	Function
R	4	N-term	Not calc.	Conserved twin R presumably needed for N-term loop closure
Y	9	N-term	65	Based on model, this points into active site and interacts with D295 (part of DDxxD)
W	243	Dimer Interface	9	Base of dimer interface, in a hydrophobic pocket
D	258	N-term	44	Near active site; possible interaction with closed N-term tail (W244)
R	259	N-term	29	Near active site; possible interaction with closed N-

				term tail
E	262	N-term	24	Near active site; possible interaction with closed N-term tail
W	266	Active Site	3	Bottom of active site
R	280	Buried	1	Interacts with backbone of P274, possible loop stabilization
D	294	Active Site	32	Part of DDxxD motif
D	295	Active Site	12	Part of DDxxD motif
D	298	Active Site	55	Part of DDxxD motif
E	305	N-term	19	Possible interaction with R004 when N-term loop is closed
Y	387	Substrate Capture Loops	7	Substrate capture loop; role unknown
S	396	Active Site	0	points into active site
G	397	Active Site	0	in active site; allows for kink in helix
R	435	Active Site	13	points into active site
D	439	Active Site	5	likely metal binding
E	446	Active Site	17	likely metal binding
R	449	Substrate Capture Loops	64	Substrate capture loop; role unknown
G	450	Substrate Capture Loops	78	Substrate capture loop; role unknown, though likely structural
Y	514	Active Site	28	probably catalytic residue
D	518	Substrate Capture Loops	Not calc.	Substrate capture loop; role unknown

Table 10. Retested variants of MEA *P. alba* displaying PI specific activity values > 1.3.

Position	Mutation	Location	Charge Change	% Surface Accessibility	Function
2	V	N-term	1	Not calc.	alternate surface interactions
22	K	N-term	1	54	alternate surface interactions
22	R	N-term	1	54	alternate surface interactions
36	D	N-term	-2	33	alternate interactions with nearby residues
36	E	N-term	-2	33	alternate interactions with nearby residues
36	H	N-term	-1	33	alternate interactions with nearby residues
36	W	N-term	-1	33	alternate interactions with nearby residues
43	E	N-term	-2	52	alternate interactions with nearby residues
58	F	N-term	1	20	alternate interactions with nearby residues
71	I	Buried	-1	3	alternate interactions in hydrophobic pocket
89	D	Buried	-1	5	alternate interactions in pocket
89	E	Buried	-1	5	alternate interactions in pocket
118	E	Buried	-1	10	alternate interactions in pocket
118	P	Buried	0	10	alternate interactions in pocket
120	M	surface loop	0	52	alternate surface interactions
120	Q	surface loop	0	52	alternate surface interactions
151	F	surface loop	0	14	alternate interactions in pocket
151	Y	surface loop	0	14	alternate interactions in pocket
153	P	surface loop	0	40	affect shape of loop
161	C	Buried	-1	10	alternate interactions in pocket
228	Y	Buried	0	5	alternate interactions in pocket
234	R	Hinge region	1	24	alternate surface interactions
247	I	dimer interface	0	5	alternate interactions in pocket
247	L	dimer interface	0	5	alternate interactions in pocket
247	M	dimer interface	0	5	alternate interactions in pocket
254	C	surface loop	0	89	alternate surface interactions
268	I	Buried	0	2	alternate interactions in pocket
282	H	Buried	0	8	alternate interactions in pocket
282	W	Buried	0	8	alternate interactions in pocket
288	A	Buried	0	0	alternate interactions in pocket
288	T	Buried	0	0	alternate interactions in pocket
288	Y	Buried	0	0	alternate interactions in pocket
331	P	Buried	0	0	alternate interactions in pocket
348	Y	Dimer interface	-1	16	alternate surface interactions
376	L	surface	0	63	alternate surface interactions
380	E	surface loop	-1	33	alternate surface interactions
389	D	surface	-1	18	alternate surface interactions
391	G	Buried	0	0	alternate interactions in pocket
392	C	Buried	0	7	alternate interactions in pocket
392	F	Buried	0	7	alternate interactions in pocket
392	M	Buried	0	7	alternate interactions in pocket
392	S	Buried	0	7	alternate interactions in pocket

392	V	Buried	0	7	alternate interactions in pocket
392	Y	Buried	0	7	alternate interactions in pocket
409	T	surface loop	0	14	alternate interactions in pocket
437	L	Buried	0	7	alternate interactions in pocket
437	M	Buried	0	7	alternate interactions in pocket
443	G	Substrate Capture Loops	0	4	affect shape of loop
443	Q	Substrate Capture Loops	0	4	affect shape of loop
443	R	Substrate Capture Loops	1	4	affect shape of loop
443	S	Substrate Capture Loops	0	4	affect shape of loop
444	D	Substrate Capture Loops	-1	56	affect shape of loop
444	E	Substrate Capture Loops	-1	56	affect shape of loop
447	T	Substrate Capture Loops	0	23	alternate interactions in pocket
447	V	Substrate Capture Loops	0	23	alternate interactions in pocket
448	V	Substrate Capture Loops	0	77	alternate surface interactions
460	A	Buried	0	21	alternate interactions in pocket
461	A	Buried	-1	32	alternate interactions in pocket
481	Y	Buried	0	0	alternate interactions in pocket
488	L	Buried	1	13	alternate interactions in pocket
502	F	Buried	0	1	alternate interactions in pocket
502	M	Buried	0	1	alternate interactions in pocket

Example 4: Growth Assay on 1024 Variants Selected from the Primary Screen

[0364] The primary *in vitro* specific activity screen identified variants of MEA *P. alba* that enhance the ability of the enzyme to catalyze the conversion DMAPP to isoprene. Since IspS must function inside a living cell, it was additionally necessary to measure the capability of the enzyme to convert DMAPP to isoprene *in vivo*. Examples 1 and 2 describe the methodology in determining the *in vivo* effectiveness of IspS. Essentially, by conversion of DMAPP to isoprene, IspS relieves the toxic effect of DMAPP on the growth of *E. coli*. Increased performance in comparison to wild type during the course of a growth curve indicates improved isoprene synthase function within a given strain. IspS variants that display both improved specific activity and the best growth performance are indicative of the enzymes best suited to improved isoprene production during fermentation.

Methods

Growth Assay and Specific Activity Measurements

[0365] 1024 variants from the primary specific activity screen were selected for growth study and to confirm increased specific activity increases relative to the wild type MEA *P. alba* enzyme. Variants at positions that displayed high mutability (tolerance to mutation), and increased performance for both specific activity and expression not significantly less than wild type, were chosen for this study. Individual variants were isolated from their original glycerol stock plates and re-arrayed for the growth assay. Variants were induced at both low and high levels of IPTG, and their growth curves were determined in the presence of mevalonic acid (MVA). In these strains, MVA is taken up and drives flux through the mevalonic acid pathway to DMAPP, which is toxic to cell growth. Expression of functional *P. alba* IspS molecules allows for the conversion of DMAPP to isoprene, and the relief of growth inhibition. In these assays, better performing IspS molecules more effectively convert DMAPP to isoprene and result in improved growth.

[0366] Glycerol stocks of MEA *P. alba* IspS libraries were thawed briefly and inoculated into microtiter plates containing liquid LB with kanamycin at a concentration of 20 µg/ml. Cultures were grown overnight at 250 rpm, 30°C to saturation in a shaking incubator. The next day, cultures were removed and inoculated at a ratio of 1:10 into TM3-glucose medium containing 50 µg/ml kanamycin and 40 or 100 µM IPTG (Sigma). Wild type controls were grown separately and inoculated into each microtiter plate containing TM3-glucose with a titration of IPTG concentrations, from 30 µM to 65 µM (for cultures induced at 40 µM), or 40 to 200 µM (for cultures induced at 100 µM), in separate wells. Plates were returned to the shaking incubator at 250 rpm, 30°C and pre-induced for two hours. Cultures were then diluted into TM3-glucose medium containing 50 µg/ml kanamycin, 40 or 100 µM IPTG, and 20 mM MVA at a ratio of 1:10 in microtiter plates (Matrical). WT controls with or without MVA were included, as well as appropriate controls with titrations of IPTG. Plates were transferred to a Growth Profiler 1152 (EnzyScreen) and growth curves and optical densities (ODs) were determined according to the manufacturer's recommendation over a ten-hour time course. Performance indices (PIs) for growth of each strain were determined by comparison to four replicate WT strains induced at either 40 or 100 µM IPTG. PI values for OD at 300 minutes, Max OD, and area under the curve were calculated. Specific activity for

all variants in this study, induced at the 40 μ M IPTG level, was also determined according to the methods described in the previous example. Samples were isolated from the same pre-induction plate as the samples used in the growth assays.

Results

[0367] Table 8 provides definitions for locations of the residues listed in Tables 11 through 15. Table 11 lists all variants that displayed a PI value for specific activity greater than 1.4. Locations, surface accessibilities, and putative functions are also listed. Table 11 lists several variants that enhance the enzymatic efficiency of IspS, either alone or in combination. Locations of positions where variants displayed improved specific activity are shown in Figures 28 and 29. Variants with improved specific activity may allow for more efficient conversion of DMAPP to isoprene, and allow for improved cellular production of isoprene during fermentation.

[0368] Tables 12 and 13 list variants with improved growth at the 40 μ M and 100 μ M induction level, respectively. While several different growth parameters were measured, all correlated well with each other, so only PI values for maximum OD (OD Max) were examined for variants listed in Tables 12 and 13. Listed variants displayed OD Max values that were 50% better than WT (PI of 1.5 or greater) at the given induction level. Variants that displayed improved growth (a PI value greater than 1.3 for OD Max) in both 40 μ M and 100 μ M IPTG induction conditions are listed in Table 14, and shown in Figures 30 and 31. These variants represent mutations that may allow for the highest overall growth performance and conversion of DMAPP to isoprene in cells expressing IspS. Several of these variants map near or within a particular N-terminal helix region of MEA *P. alba*, spanning residues 134 to 179. Several changes at or near this location (“N-term helices” in Tables 8, 12 through 15) displayed a growth benefit at either or both growth conditions. Not only do multiple variants map to this location in MEA *P. alba*, but the variants that displayed the largest benefit to growth are facing outwards from the helix and are located on the surface of the enzyme (see Figs. 31, 33, and 34).

[0369] Table 15 lists variants that displayed improved performance (PI greater than 1.2) for all three parameters of specific activity, OD Max at 40 μ M, and 100 μ M IPTG. Several of these variants are listed in Table 14, and the majority of them are also located in or near the N-terminal helix described above, with the exceptions of V30K and V84T (see Figure 32).

This indicates that alterations at the helix spanning residues 150 to 172 are critical not only for improved growth of the host cell, but also for improved enzymatic activity. Since there is no obvious catalytic role for the N-terminal helix (see Fig. 33), these variants may influence IspS activity either intra-molecularly through conformational change of the enzyme structure, or inter-molecularly via the above-mentioned interactions with an unidentified enzyme, cellular process or structure. MEA *P. alba* enzymes harboring variants at this particular location, either alone or in combination with other variants conferring beneficial properties such as improved catalytic rates, likely will allow for improved growth rate of host strains, and improved isoprene production during fermentation.

Table 11. Retested variants of MEA *P. alba* that displayed PI values for specific activity > 1.4.

Position	Mutation	Location	Charge Change	% Surface Accessibility	Function
18	E	N-term	-1	64	alternate interactions with nearby residues
18	D	N-term	-1	64	alternate interactions with nearby residues
18	S	N-term	0	64	alternate interactions with nearby residues
36	P	N-term	-1	33	alternate interactions with nearby residues
82	Q	N-term	-1	27	alternate interactions with nearby residues
87	S	surface loop	0	35	alternate surface interactions
87	N	surface loop	0	35	alternate surface interactions
87	R	surface loop	1	35	alternate surface interactions
137	C	N-term helices	0	15	alternate interactions in pocket
143	N	N-term helices	0	1	alternate interactions in hydrophobic pocket
163	I	N-term helices	0	10	alternate interactions in pocket
163	Q	N-term helices	0	10	alternate interactions in pocket
170	G	N-term helices	1	79	alternate surface interactions
242	T	Dimer interface	-1	35	improve dimer interface interactions
251	E	surface	-1	74	alternate interactions with nearby residues
409	S	surface loop	0	14	alternate interactions in pocket
437	M	Buried	0	7	alternate interactions in pocket
437	K	Buried	1	7	alternate interactions in pocket
443	G	Substrate Capture Loops	0	4	affect shape of loop
443	S	Substrate Capture Loops	0	4	affect shape of loop
443	Q	Substrate Capture Loops	0	4	affect shape of loop

444	P	Substrate Capture Loops	0	56	affect shape of loop
447	Q	Substrate Capture Loops	0	23	alternate interactions in pocket
447	T	Substrate Capture Loops	0	23	alternate interactions in pocket
447	M	Substrate Capture Loops	0	23	alternate interactions in pocket
447	E	Substrate Capture Loops	-1	23	alternate interactions in pocket
447	S	Substrate Capture Loops	0	23	alternate interactions in pocket
447	R	Substrate Capture Loops	1	23	alternate interactions in pocket
460	Q	Buried	0	21	alternate interactions in pocket
460	S	Buried	0	21	alternate interactions in pocket
460	G	Buried	0	21	alternate interactions in pocket
460	A	Buried	0	21	alternate interactions in pocket
461	D	Buried	-2	32	alternate interactions in pocket
461	S	Buried	-1	32	alternate interactions in pocket
461	T	Buried	-1	32	alternate interactions in pocket
461	E	Buried	-2	32	alternate interactions in pocket
542	N	surface loop	0	17	alternate interactions with nearby residues

Table 12. Variants of MEA *P. alba* that displayed PI values > 1.5 for OD Max at 40 mM IPTG

Position	Mutation	Location	Charge Change	% Surface Accessibility	Function
134	P	N-term helices	-1	37	loop stabilization
138	C	N-term helices	-1	53	alternate interactions with nearby residues
143	F	N-term helices	0	1	alternate interactions in hydrophobic pocket
143	V	N-term helices	0	1	alternate interactions in hydrophobic pocket
156	G	N-term helices	0	13	allow more flexibility in loop
159	G	N-term helices	1	32	alternate interactions with nearby residues
159	Q	N-term helices	1	32	alternate interactions with nearby residues
163	C	N-term helices	0	10	alternate interactions with nearby residues
163	E	N-term helices	-1	10	alternate interactions with nearby residues
163	Q	N-term helices	0	10	alternate interactions with nearby residues
163	V	N-term helices	0	10	alternate interactions with nearby residues
163	Y	N-term helices	0	10	alternate interactions with nearby residues

166	C	N-term helices	0	46	alternate surface interactions
166	D	N-term helices	-1	46	alternate surface interactions
166	G	N-term helices	0	46	alternate surface interactions
166	P	N-term helices	0	46	alternate surface interactions
166	V	N-term helices	0	46	alternate surface interactions
167	M	N-term helices	0	21	alternate interactions with nearby residues
170	G	N-term helices	1	79	alternate surface interactions
170	H	N-term helices	1	79	alternate surface interactions
170	K	N-term helices	2	79	alternate surface interactions
170	N	N-term helices	1	79	alternate surface interactions
170	R	N-term helices	2	79	alternate surface interactions
170	S	N-term helices	1	79	alternate surface interactions
170	W	N-term helices	1	79	alternate surface interactions
414	F	surface loop	-1	70	alternate surface interactions
414	G	surface loop	-1	70	alternate surface interactions
414	N	surface loop	-1	70	alternate surface interactions
414	P	surface loop	-1	70	alternate surface interactions
421	R	surface loop	1	23	alternate surface interactions
491	Q	surface loop	0	58	alternate surface interactions
491	V	surface loop	0	58	alternate surface interactions
491	Y	surface loop	0	58	alternate surface interactions

Table 13. Variants of MEA *P. alba* that displayed PI values > 1.5 for OD Max at 100 mM IPTG

Position	Mutation	Location	Charge Change	% Surface Accessibility	Function
29	N	N-term	1	34	alternate surface interactions
47	V	surface loop	0	16	affect shape of loop
86	C	surface loop	0	59	alternate interactions
94	A	surface loop	-1	73	affect shape of loop
131	F	N-term helices	1	58	improved interaction with neighboring his
134	E	N-term helices	-2	37	loop stabilization
134	P	N-term helices	-1	37	loop stabilization
156	G	N-term helices	0	13	allow more flexibility in loop
162	P	N-term helices	0	60	loop stabilization
169	C	N-term helices	-1	60	alternate surface interactions
178	E	N-term helices	-2	73	alternate surface interactions
179	T	N-term helices	1	46	alternate surface interactions
231	D	hinge region	-1	33	alternate surface interactions
231	K	hinge region	1	33	alternate surface interactions
231	R	hinge region	1	33	alternate surface interactions
231	T	hinge region	0	33	alternate surface interactions
231	V	hinge region	0	33	alternate surface interactions
242	N	dimer interface	-1	35	improve dimer interface interactions
242	I	dimer interface	-1	35	improve dimer interface interactions
369	C	active site	0	2	alter active site cavity
414	C	surface loop	-1	70	alternate surface interactions

414	F	surface loop	-1	70	alternate surface interactions
414	G	surface loop	-1	70	alternate surface interactions
414	N	surface loop	-1	70	alternate surface interactions
421	D	surface loop	-1	23	alternate surface interactions

Table 14. Variants of MEA *P. alba* that displayed PI values > 1.3 for OD Max at 40 and 100 mM IPTG

Position	Mutation	Location	Charge Change	% Surface Accessibility	Function
50	S	N-term	-1	86	alternate surface interactions
81	F	N-term	1	48	alternate surface interactions
134	E	N-term helices	-2	37	improved interaction with neighboring his
134	P	N-term helices	-1	37	loop stabilization
137	N	N-term helices	0	15	alternate interactions in pocket
143	V	N-term helices	0	1	alternate interactions in hydrophobic pocket
156	G	N-term helices	0	13	allow more flexibility in loop
159	D	N-term helices	0	32	alternate interactions with nearby residues
159	G	N-term helices	1	32	allow more flexibility in loop
159	Q	N-term helices	1	32	alternate interactions with nearby residues
166	C	N-term helices	0	46	alternate surface interactions
166	W	N-term helices	0	46	alternate surface interactions
167	M	N-term helices	0	21	alternate interactions with nearby residues
167	N	N-term helices	0	21	alternate interactions with nearby residues
169	C	N-term helices	-1	60	alternate surface interactions
170	H	N-term helices	1	79	alternate surface interactions
170	K	N-term helices	2	79	alternate surface interactions
170	W	N-term helices	1	79	alternate surface interactions
414	C	surface loop	-1	70	alternate surface interactions
414	F	surface loop	-1	70	alternate surface interactions
414	G	surface loop	-1	70	alternate surface interactions
414	N	surface loop	-1	70	alternate surface interactions
414	P	surface loop	-1	70	alternate surface interactions

Table 15. Variants of MEA *P. alba* that displayed PI values > 1.2 for Specific Activity, OD Max at 40 μ M IPTG, and 100 μ M IPTG

Position	Mutation	Location	Charge Change	% Surface Accessibility	Function
30	K	N-term	1	60	replace surface hydrophobic w/ charged

84	T	N-term	0	30	polar residue for better solvent interactions
134	C	N-term helices	-1	37	size reduction, improved interaction with neighboring His
134	D	N-term helices	-2	37	improved interaction with neighboring his
134	E	N-term helices	-2	37	improved interaction with neighboring his
140	S	N-term helices	0	1	may improve interactions with other polar residues in pocket
140	T	N-term helices	0	1	may improve interactions with other polar residues in pocket
143	F	N-term helices	0	1	alternate interactions in hydrophobic pocket
143	I	N-term helices	0	1	alternate interactions in hydrophobic pocket
143	M	N-term helices	0	1	alternate interactions in hydrophobic pocket
143	V	N-term helices	0	1	alternate interactions in hydrophobic pocket
163	I	N-term helices	0	10	alternate interactions in pocket
163	M	N-term helices	0	10	alternate interactions in pocket
166	P	N-term helices	0	45	stabilize helix
166	V	N-term helices	0	45	alternate surface interactions
169	Q	N-term helices	-1	60	alternate surface interactions
170	H	N-term helices	1	80	alternate surface interactions
170	K	N-term helices	2	80	alternate surface interactions
172	V	N-term helices	0	47	affect shape of loop

Example 5: Specific Activity and Growth Assays on Combinatorial Libraries

[0370] Single variants of MEA *P. alba* IspS improved for specific activity, growth, or both traits were selected for combination into three seven-member libraries.

Methods

[0371] Libraries were constructed in the pCL201 vector and transformed into the MD09-170 screening strain (DNA2.0). 160 individual variants, representing approximately 80 to 90% of the 128 possible combinations in each library, were screened for both specific activity and growth following the methods described in the previous examples. Table 16 lists the variants chosen for combinatorial libraries, their locations in the crystal structure, surface accessibilities, and selection criteria (either specific activity, growth or both).

Table 16. Variants chosen for combinatorial libraries

Residue	Position	Mutation	Library	Location	% Surface Accessibility	Selection Criteria
S	288	C	1/2	C-term	0	Growth (Solubility)
S	22	R	1	N-term	54	Specific Activity
R	71	I	1	N-term	3	Specific Activity
S	444	D	1/2	SubLoop	56	Specific Activity
M	460	A	1	SubLoop	21	Specific Activity
A	443	G	1	SubLoop	4	Specific Activity
T	502	M	1	Buried	1	Specific Activity
V	409	T	2	surface loop	14	Growth/Specific Activity
R	242	N	2	DimInt	35	Growth/Specific Activity
K	414	F	2	surface loop	70	Growth
V	162	P	2	N-term helix	60	Growth
G	87	R	2	surface loop	35	Growth/Specific Activity
S	288	T	3	C-term	0	Growth/Specific Activity
N	47	V	3	surface loop	16	Growth
I	447	T	3	SubLoop	23	Specific Activity
E	170	H	3	N-term helix	79	Growth
S	231	T	3	hinge region	33	Growth
K	414	N	3	surface loop	70	Growth
I	156	G	3	N-term helix	13	Growth

Results

[0372] Combinatorial variants displaying significantly improved specific activity and/or growth performance were identified. Table 17 contains a list of combinatorial variants that displayed performance index (PI) values for specific activity greater than 2.6. The left hand column lists the variant number, and the subsequent columns list the genotype for the 7 different positions in that library. Variants with improved specific activity allow for more efficient enzymatic conversion of DMAPP to isoprene, likely by improved kinetic

parameters. Table 18 contains a list of combinatorial variants that displayed PI values greater than 1.3 for specific activity, OD Max at 40 uM, and OD Max at 100 uM. IspS variants improved for both specific activity and growth parameters also convert DMAPP to isoprene more efficiently than the WT enzyme, and likely are beneficial to growth of the host strain by mediation of a deleterious effect of IspS within the host.

[0373] Since each library member contained any combination of seven possible mutations, effects of variants were observed multiple times in different configurations. This provided a robust internal control to help identify the most effective combinations present in each library. For example, the S444D mutation was present in the highest specific activity combinatorial variants, suggesting that this variant, in combination with other improved variants, is particularly beneficial for in vitro activity. M460A, A443G, and I447T also displayed this type of effect. Similarly, the V162P mutation was present in all combinatorial variants with improved properties for both specific activity and growth, suggesting that V162P works well in combination with other variants and may be an ideal mutation for more efficient conversion of DMAPP to isoprene within the host cell, for the reasons described in the previous paragraph. I156G and E170H also showed this effect. The variants G087R, R242N, and S288T also displayed improved specific activity in combination with other variants, but were not always among the highest performers for the in vitro assay. The combinatorial variants listed in Tables 17 and 18 may represent significantly improved IspS enzymes that allow for optimal conversion of DMAPP into isoprene during fermentation of host cells. The presence or absence of individual mutations in these particular combinatorial variants additionally may signify the best overall mutations that can be combined in future variants of IspS, critical to the optimization of isoprene production by fermentation of microorganisms.

Table 17. Combinatorial variants of MEA *P. alba* displaying PI specific activity values > 2.6.

Variant	Pos 1	Pos 2	Pos 3	Pos 4	Pos 5	Pos 6	Pos 7
1	022R	071I	288C	443A	444D	460A	502M
2	022R	071I	288C	443G	444D	460A	502M
3	022S	071I	288C	443A	444S	460A	502T
4	022S	071R	288C	443A	444D	460A	502T
5	087G	162P	242N	288S	409T	414F	444D
6	087G	162P	242N	288S	409T	414K	444D
7	087G	162P	242R	288C	409T	414F	444D
8	087G	162V	242N	288C	409V	414F	444D
9	087G	162V	242R	288C	409T	414F	444D

10	087G	162V	242R	288C	409V	414K	444D
11	087R	162P	242N	288C	409T	414F	444D
12	087R	162P	242N	288C	409T	414K	444D
13	087R	162P	242N	288S	409T	414K	444D
14	087R	162P	242N	288S	409V	414F	444D
15	087R	162P	242N	288S	409V	414K	444D
16	087R	162P	242R	288C	409T	414F	444D
17	087R	162P	242R	288C	409V	414F	444D
18	087R	162P	242R	288C	409V	414K	444D
19	087R	162P	242R	288S	409T	414F	444D
20	087R	162P	242R	288S	409V	414K	444D
21	087R	162V	242N	288C	409V	414F	444D
22	087R	162V	242N	288S	409T	414F	444D
23	087R	162V	242N	288S	409T	414K	444D
24	087R	162V	242N	288S	409V	414F	444D
25	087R	162V	242R	288C	409T	414K	444D
26	087R	162V	242R	288C	409V	414K	444D
27	087R	162V	242R	288S	409V	414F	444D

Table 18. Combinatorial variants of MEA *P. alba* displaying PI >1.3 for specific activity, OD Max at 40 μ M IPTG, and OD Max at 100 μ M IPTG.

Variant	Pos 1	Pos 2	Pos 3	Pos 4	Pos 5	Pos 6	Pos 7
1	047N	156G	170H	231S	288T	414K	447I
2	047V	156G	170H	231S	288T	414K	447I
3	047V	156I	170H	231S	288T	414F	447I
4	047V	156I	170H	231T	288T	414K	447I
5	087G	162P	242N	288C	409T	414F	444S
6	087G	162P	242N	288C	409V	414K	444S
7	087G	162P	242N	288S	409T	414F	444S
8	087G	162P	242R	288C	409T	414K	444S
9	087G	162P	242R	288S	409T	414K	444S
10	087R	162P	242N	288C	409T	414F	444S
11	087R	162P	242N	288C	409V	414K	444S
12	087R	162P	242R	288C	409V	414K	444S

Example 6: Specific Activity Determination of N-terminal Truncations of MEA *P. alba* IspS

[0374] Isoprene synthase contains tandem arginine residues at the N-terminus that are necessary for proper enzymatic conversion of DMAPP to isoprene. As a truncated variant, MEA *P. alba* demonstrates high specific activity in comparison to enzymes with longer N-terminal regions, up to the naturally occurring chloroplast targeting peptide. The MEA *P. alba* enzyme has only two residues upstream of the tandem arginine residues (see Figure 35), yet the function of these residues with regard to enzyme activity was not reported. N-terminal truncations of the MEA *P. alba* enzyme therefore were generated and assayed to determine if further truncations confer a specific activity benefit to IspS.

Methods

[0375] The two truncations of MEA *P. alba* were constructed by QuikChange (Stratagene) PCR on the template pCL201 (see Table 19 for primer sequences) as previously described following the manufacturer's recommended protocol. The PCR product was treated with 1 μ l DpnI (Roche) for 3 hours, and then 1 μ l of the entire reaction was transformed into chemically competent *E. coli* Top10 cells (Invitrogen) according to the manufacturer's recommended protocol. Cells were recovered and plated on LB medium containing 50 μ g/ml kanamycin. The next day, positive colonies were chosen for growth,

plasmid purification (Qiagen) and sequencing (Quintara Biosciences). Plasmids which harbored the correct truncations were selected for sequencing of the entire open reading frame to confirm the integrity of the coding sequence. These plasmids, pDW207 (see Figure 36) and pDW208 (see Figure 37), were transformed by electroporation into the expression strain MD09-170 for determination of specific activity (see Table 20). Specific activity was determined as previously described. At least 30 replicates of each truncation were analyzed in comparison to MEA *P. alba*.

Table 19. Primers used for QuikChange Mutagenesis

HgS MRR Forward	TATACATATGCGTCGCTCTGCGAACTACGA	(SEQ ID NO:23)
HgS MRR Reverse	CAGAGCGACGCATATGTATATCTCCTTCTT	(SEQ ID NO:24)
HgS MAR Forward	TATACATATGGCACGTCGCTCTGCGAACTA	(SEQ ID NO:25)
HgS MAR Reverse	AGCGACGTGCCATATGTATATCTCCTTCTT	(SEQ ID NO:26)

Table 20. Strains with N-terminal truncations

Strain	Plasmid	Description
DW618	pDW207	BL21 (DE3) PL.2-mKKDyI + <i>P. alba</i> IspS MAR (-1 from MEA <i>P. alba</i>)
DW619	pDW208	BL21 (DE3) PL.2-mKKDyI + <i>P. alba</i> IspS MRR (-2 from MEA <i>P. alba</i>)

Results

[0376] Specific activities of the truncated molecules of *P. alba* IspS expressed in strains DW618 (MAR) or DW619 (MRR) were either not improved or slightly lower, respectively, than the parental MEA *P. alba* enzyme. Table 21 shows performance index values for both the MAR and MRR truncations of *P. alba* IspS. The MAR truncation displayed specific activity that was approximately equivalent to the control MEA *P. alba* molecule, and the MRR truncation displayed specific activity that was approximately 81% of the control.

Although these truncations did not have increased specific activity in comparison to MEA *P. alba*, they retained sufficient activity to be of potential future use in fermentation strains that convert DMAPP to isoprene via an IspS enzyme, where complete removal of the N-terminus up to but not including the tandem arginine residues is required.

Table 21. Performance index values for truncated variants of *P. alba* IspS

Strain	Variant	PI Specific Activity	Standard Deviation
DW618	MAR	0.983189	0.091889
DW619	MRR	0.813857	0.072938

Amino Acid Sequence of *P. alba* IspS MAR

MARRSANYEPNSWDYDYLLSSDTDESIEVYKDKAKKLEAEVRREINNEKAEFLTLL
 LIDNVQRLGLGYRFESDIRGALDRFVSSGGFDAVTKTSLHGTALSFRLLRQHGFVSQ
 EAFSGFKDQNGNFLENLKEDIKAILSLEYEASFLALEGENILDEAKVFAISHLKELSEEKI
 GKELAEQVNHAELEPLHRRRTQRLEAVWSIEAYRKKEDANQVLELAILDYNMIQSV
 YQRDLRETSRWWRRVGLATKLHFARDRLIESFYWAVGVAFEPQYSDCRNSVAKMF
 SFVTIIDDYDVYGTLDLELELFTDAVERWDVNAINDLPDYMKLCFLALYNTINEIAYD
 NLKDKGENILPYLTKAWADLCNAFLQEAKWLYNKSTPTFDDYFGNAWKSSSGPLQL
 VFAYFAVVQNIKKEEIEENLQKYHDTISRPSHIFRLCNDLASASAEIARGETANSVSCY
 MRTKGISEELATESVMNLIDETWKKMNKEKLGGS�FAKPFVETAINLARQSHCTYHN
 GDAHTSPDELTRKRVLSVITEPILPFER (SEQ ID NO:27)

DNA Sequence of plasmid pDW207

tggcgaatgggacgcgccctgtagcggcgcattaagcgcggcgggtgtgtgttacgcgcagcgtgaccgctacactgccagcgc
 ccctagcgcggcctccttcgctttctccctctctctcgcacgttcgccgcttccccgtcaagctctaaatcgggggctcccttag
 ggtccgatttagtgctttacggcacctcgacccccaaaaacttgattaggggtgatggtcacgtagtgggccatcgcctgatagacgg
 ttttcgcccttgacgttgagtcacgttcttaatagtggactctgttccaaactggaacaactcaaccctatctcggctctattctttg
 attataagggtatttgccgatttcggcctattggttaaaaatgagctgatttaacaaaatgacgcgaatttaacaaaatattaacgtt
 acaatttcaggtggcacttttcggggaatgtgcgcggaaccctattgtttattttctaaatacattcaaatatgtatccgctcatgaatta
 attcttagaaaaactcatcgagcatcaaatgaaactgcaatttattcatatcaggattatcaataccatattttgaaaaagccgttctgtaat
 gaaggagaaaaactcaccgagcagttccataggatggcaagatcctggtatcggctcgcattccgactgtccaacatcaatacaac
 ctattaattcccctcgtcaaaaataaggttatcaagtgagaaatcccatgagtgacgactgaatccgggtgagaatggcaaaagtattatg

cattttttccagacttgttcaacagccagccattacgctcgtcatcaaaatcactcgcacccaaccggttattcattcgtgattgcgc
ctgagcgagacgaaatagcgcgatcgtgtttaaaggacaattacaacaggaatcgaatgcaaccggcgcaggaacactgccagc
gcatcaacaatattttcacctgaatcaggatatttcttaataacctggaatgctgtttcccggggatcgcagtggtgagtaacctgcatc
atcaggagtacggataaatgcttgatggtcggagaggcataaattccgctcagccagtttagtctgaccatctcatctgtaacatcattg
gcaacgctacctttgccatgtttcagaacaactctggcgcacgggcttccatacaatcgatagattgctgcacctgattgcccacat
tatcgcgagcccatttataccataataatcagcatccatgttggaatttaatcgcggcctagagcaagacgtttcccgttgaatatggctc
ataacacccttgtattactgtttatgtaagcagacagttttattgttcatgacaaaatccctaacgtgagtttccgttccactgagcgtcag
accccgtagaaaagatcaaggatcttcttgagatcctttttctgcgcgtaactgctgcttgcacacaaaaaacaccgctaccagc
ggtggtttgttccggatcaagagctaccaactctttccgaagtaactggcttcagcagagcgcagatacacaactgtccttcta
gtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgtctgctaactctgttaccagtggctgctgcc
agtggcgataagtcgtgtcttaccgggttgactcaagacgatagttaccggataaggcgcagcggctgggctgaacggggggttcg
tgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgcacgcttcccga
agggagaaaaggcggacaggtatccggttaagcggcaggggtcggaaacaggagagcgcacgagggaactccaagggggaaacgcc
tggtatctttatagctctgctgggttccacactctgacttgagcgtcgtttttgtgatgctcgtcagggggcggagcctatggaaaa
cgccagcaacgcggccttttacggttctggcctttgctggcctttgctcactgttcttctgctgttaccctgattctgttgataac
cgtattaccgctttgagtgagctgataccgctcggcgcagccgaacgaccgagcgcagcagtgagcaggaagcgggaag
agcgcctgatcgggtattttctcttacgcatctgtcgggtatttcaaccgcatatggtgcactctcagtacaactctgctctgatgccgc
atagttaagccagtatacactccgctatcgtactggtactgggtcatggctgcgccccgacaccgccaacaccgctgacgcgcct
gacgggcttgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagagggtttaccgctatca
ccgaaacgcgcgagggcagctcggtaagctcatcagcgtggtcgtgaagcgattcacagatgtctgctgttaccgcgctcagct
cgttgagtttctccagaagcgttaatgtctggcttctgataaagcgggcatgtaaggcgggttttctgtttggtcactgatgcctccg
tgaagggggatttctgttcatgggggtaatgataccgatgaaacgagagaggatgctcacgatacgggttactgatgatgaacatgcc
cgttactggaacggtgtgagggttaacaactggcggatggatgcggcgggaccagagaaaaatcactcagggtcaatgccagcg
cttcgttaatacagatgtaggtgttccacagggtagccagcagcatctgcgatgcagatccggaacataatggtgcagggcgtgac
ttccgctttccagactttacgaaacacggaaaccgaagaccattcatgttgttctcaggtcgcagacgtttgcagcagcagtcgcttc
acgttcgctcgcgtatcgggtattctgctaaccagtaaggcaacccgccagcctagccgggtcctcaacgacaggagcacgat
catgcgcacccgtggggccgcatccggcgataatggcctgcttctcgcgaaacgtttggtggcgggaccagtgacgaaggctt
gagcaggggctgcaagattccgaataaccgaaagcagagccgatcatcgtcgcgctccagcgaagcggctctcggcgaat
gaccagagcgtgccggcactgtcttacgattgcatgataaagaagacagtcataagtgcggcgacgatagtcatgccccgcg
cccaccggaaggagctgactgggtgaaggctctcaaggcctcggctcagatcccgggtcctaatgagtgagctaacttacattaat
tgcgttgcgctcactgcccgtttccagtcgggaaacctgctgtgccagctgcattaatgaatcggccaacgcgcggggagaggcgg
tttgcgtattgggcgccagggtggtttttttaccagtgagacgggcaacagctgattgcccttaccgctggcctgagagagtt
gcagcaagcgggtccacgctggttggcccagcagggcaaaatcctgtttgatggtggttaacggcgggatataacatgagctgtcttcg
gtatcgtctatcccactaccgagataccgaccaacgcgcagcccggactcggtaatggcgcgcattgcgccagcgcctatctga

tcgttggcaaccagcatcgcagtgggaacgatgccctcattcagcatttgcattggttggtaaaccggacatggcactccagtcgcct
tccggtccgctatcggctgaatttgattgcgagtgagatattatgccagccagaccagacgcagacgcgccgagacagaacttaatgg
gcccgctaacagcgcgatttgctggtgacccaatgcgaccagatgctccacgccagtcgcgtaccgtcttcatgggagaaaataata
ctgttgatgggtgctggtcagagacatcaagaaataacgccggaacattagtgcaggcagctccacagcaatggcatcctggtcatc
cagcggatagttaatgatcagcccactgacgcgttgcgcgagaagattgtgcaccgccgtttacaggcttcgacgccgttcgttcta
ccatcgacaccaccacgctggcaccagttgatcggcgcgagatattaatgccgcgacaatttgcgacggcgcgtgcagggccaga
ctggaggtggcaacgccaatcagcaacgactgtttgcccgccagttgtgtgccacgcggttgggaatgtaattcagctccgccatcg
ccgcttccacttttccgcgttttcgagaaacgtggctggcctggttaccacgcgggaaacggctctgataagagacaccggcatac
tctgcgacatcgataacgttactggtttcacattcaccaccctgaattgactctctccgggcgctatcatgccataaccgcgaaaggttt
gcgccattc gatggtgtccgggatctcgacgctctcccttatgcgactcctgcattaggaaagcagcccagtagtaggttgaggccgttg
agcaccgccgccgcaaggaatggtgatgcaaggagatggcgcccaacagtccccggccacggggcctgccaccataccaccg
ccgaaacaagcgtcatgagcccgaagtggcgagcccgatcttccccatcggtgatgtcggcgatataggcgccagcaaccgcacc
tgtggcgccggtgatgccggccacgatgcgtccggcgtagaggatcgagatctc gatccgcgaaattaatac gactactatagg
gaattgtgagcggataacaattcccctctagaaataattttgttaacttaagaaggagatatacatatggcacgtcgtctcgaactac
gaacctaacagctgggactatgattacctgctgtctccgacacggacgagtcaccatcgaagtatacaaaagacaaagcgaaaaagctg
gaagccgaagttcgtcgcgagattaataacgaaaaagcagaatttctgacctgctggaactgattgacaacgtccagcgcctggggcc
tgggtaccgtttcagctctgatatccgtggtgcgctggatcgtctgttctccggcggttcgatgcggttaaccaagacttccctgca
cggtagcggcactgtctttccgtctgctgcgtaaacacggttttgaggtttctcaggaagcgttcagcggcttcaaagacaaaacggca
acttctggagaacctgaaggaagatatcaaaactatcctgagcctgtacgagccagcttctgctctggaagggcgaaaacatcct
ggacgagggcaaggttttcgcaatctctcatctgaaagaactgctgaaagaaagatcgtaaaagagctggcagaacaggtgaacca
tgcactggaactgccactgcatcggctactcagcgtctggaagcagatggtctatcgaggcctaccgtaaaaaggaggacgcgaa
tcaggttctgctggagctggcaattctggattacaacatgatccagtctgtataccagcgtgatctgcgtgaaacgtcccgttgggtggcgt
cgtgtgggtctggcgaccaaaactgcactttgctcgtgaccgcctgattgagagcttctactgggcccgtgggtgtagcattcgaaccgca
atactccgactgccgtaactccgtcgaaaaatgtttcttctgtaaccattatcgacgatattctacgatgtatacggcaccctggacgaac
tggagctgtttactgatgcagttgagcgttgggacgtaaacgcatcaacgacctgccggattacatgaaactgtgctttctggctctgta
taacactattaacgaaatcgcctacgacaacctgaaagataaaggtgagaacatcctgccgtatctgaccaaaagcctgggctgacctgt
gcaacgctttctgcaagaagccaagtggctgtacaacaaatctactccgacctttgacgactacttcggcaacgcatggaatcctctt
ctggcccgtgcaactggtgttcgcttactcgtctcgtgcgagaacattaaaaaggaagagatcgaaaacctgcaaaaataccatgac
accatctctcgtccttccatattctccgtctgtgcaatgacctggctagcgcgtctcgggaaattgcgcgtggtgaaaccgcaaatagc
gtttctgttcatcgcgactaaaggtatctccgaagaactggctaccgaaagcgtgatgaatctgatcgaatcctggaaaaagatg
aacaaggaaaaactgggtgtagcctgttcgcgaaaccgttcgtggaaccgcgatcaacctggcacgtcaatctcactgcacttate
ataacggcgacgcgcatacctctccggatgagctgacctgcaaacgcgttctgtctgtaatcactgaaccgattctgccgttgaacgc
taaggatccgaattcagctccgtcgacaagcttgcggccgactcagcaccaccaccaccactgagatccggctgctaaca

ggtggtttgttccggatcaagagctaccaactcttttccgaaggtaactggcttcagcagagcgcagataccaaactgtccttcta
 gtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgtctgtctaatctgttaccagtggctgtgcc
 agtggcgataagtcgtgtcttaccgggttgactcaagacgatagttaccggataaggcgcagcggcgggctgaacggggggttcg
 tgcacacagcccagcttgagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcggccacgctcccga
 agggagaaaggcggacaggtatccggtaaagcggcagggcggaaacaggagagcgcacgagggaacttccagggggaaacgcc
 tggatctttatagctctgtcgggttcgccacctctgactgagcgtcgatftttgtgatgctcgtcagggggcggagcctatgaaaaa
 cgccagcaacgcggccttttacggctctggccttttctgctcaccatgctcttctcgttaccctgattctgtggataac
 cgtattaccgcctttgagtgagctgataccgctcggcgcagccgaacgaccgagcgcagcagtgagcaggaagcgggaag
 agcgcctgatgcggatftttctccttacgcatctgtcgggtatfttccacccgatatatggtgactctcagtaaatctgctctgatgccg
 atagttaagccagatatacactccgctatcgtactgactgggtcattggctgcgccccgacaccgccaacaccgctgacgcgcct
 gacgggctgtctgctcccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggtttaccgctatca
 ccgaaacgcgcgagggcagctgcggtaaagctcatcagcgtggtcgtgaagcgttaccagatgtctgctgttaccgctccagct
 cgttgagtttccagaagcgttaagtctggttctgataaagcgggcatgtaaggcgggttttctggttgactgatgctccg
 tgaagggggatttctgttcatgggggtaagtataccgatgaaacgagagaggtgctcacgatacgggttactgatgatgaacatgcc
 cggttactggaacgttgtagggtaaacactggcggatggatgcggcgggaccagagaaaaatcactcagggtcaatgccagcg
 ctctgtaatacagatgtaggtgtccacagggtagccagcagcatctcgtatgcagatccggaacataatggtgagggcgtgac
 ttccgcgttccagactttacgaaacacggaaaccgaagaccattcatgttgtgctcaggtcgcagacgttttgcagcagcagtcgctt
 acgttcgctcgcgtatcgggtattcattctgtaaccagtaaggcaaccccgccagcctagccgggtcctcaacgacaggagcagat
 catgcgcaccggtggggcccatccggcgataatggcctgcttctcggcaaacgtttggtggcgggaccagtgacgaaggctt
 gagcagggcgtgcaagattccgaataccgcaagcgacaggccgatcatcgtcgcgctccagcgaagcggctctcggcaaat
 gaccagagcgtgccggcactgtctacgagttgcatgataaagaagacagtcataagtgcggcgacgatagtcatgccccgcg
 cccaccggaaggagctgactgggtgaaggctctcaaggcgtcggcagatccgggtcctaatgagtgagctaacctacattaat
 tgcgttgcgtcactgcccgtttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggccaacgcgcggggagaggcgg
 tttgctattgggcgccagggtggttttttaccagtgagacgggcaacagctgattgcccttaccgcctggcctgagagagtt
 gcagcaagcggccacgctggttggccagcagggcaaacctgtttgatggtggttaacggcgggatataacatgagctgtcttcg
 gtatcgtctatcccactaccgagataccgcaccaacgcgcagcccggactcggtaatggcgcgattgcgccagcggccatctga
 tcgttggcaaccagcatcagtgagggaacgatgccctcattcagcatttgcattggttggaaaccggacatggcactccagtcgct
 tcccgttccgctatcggctgaatttgattgcgagtgagatattatgccagccagccagacgcagacgcgccgagacagaacttaatgg
 gcccgtaacagcgcgatttctggtgaccaatgcgaccagatgctccagcccagtcgctaccgttctatgggagaaaataata
 ctgttgaggggtgtctgtcagagacatcaagaaataacgccggaacattagtcagggcagcttccacagcaatggcatcctggtcact
 cagcggatagftaatgatcagcccactgacgcgttgcgcgagaagattgtgaccgccgctttacaggcttcagccgcttctgctta
 ccatcgacaccaccgctggcaccagttgatcggcgcgagatttaatgccgcgacaatttgcgacggcgcgtgcagggccaga
 ctggaggtggcaacgccaatcagcaacgactgtttgcccgcaattgtgtgccacgcggttgggaatgtaattcagctccgcatcg
 ccgcttccacttttccgcgtttctgcagaaacgtggctggcctggttaccacgcgggaaacggtctgataagagacaccggcatac

tctgcgacatcgtataacgttactggttcacattcaccacctgaattgactctctccggcgctatcatgccataaccgcgaaagggtttt
 gcgccattc gatggtgtccgggatctcgacgctctcccttatgcgactcctgcattagggaagcagcccagtagtaggttgaggccgttg
 agcaccgccgccgaaggaatggtgcatgcaaggagatggcgcccaacagtccccggccacggggcctgccaccatacccacg
 ccgaaacaagcgctcatgagcccgaagtggcgagcccgatctccccatcgggtgatgtcggcgatataggcgccagcaaccgcacc
 tgtggcgccggtgatgccggccacgatgcgtccggcgtagaggatcgagatctcgatcccggaattaatacactactataggg
 gaattgtgagcggataacaattcccctctagaataatttgttaactttaagaaggagatatacatatgcgtcgtctgcgaactacgaa
 cctaacagctgggactatgattacctgctgtcctccgacacggacgagtcctcgaagtatacaaaagacaaaagcgaaaaagctggaa
 gccgaagttcgtcgcgagattaataacgaaaaagcagaatttctgacctgctggaactgattgacaacgtccagcgcctgggcctgg
 gttaccgttctgagtctgataccgtggtgcgctggatcgcttcgtttctccggcggttcgatgcggttaaccaagacttccctgcacgg
 tacggcactgctttccgtctgctgcgtcaacacgggtttgagggttctcaggaagcgttcagcggcttcaaaagacaaaacggcaactt
 cctggagaacctgaaggaaagatatcaaagctatcctgagcctgtacgaggccagcttctggctctggaaggcgaaaacatcctgga
 cgaggcgaagggttttcgcaatctctcatctgaaagaactgtctgaagaaaagatcggtaaagagctggcagaacaggtgaacctgca
 ctggaactgccactgcatcgccgtactcagcgtctggaagcagtatggtctatcgaggcctaccgtaaaaaggaggacgcgaatcag
 gttctgctggagctggcaattctggattacaacatgatccagtctgtataccagcgtgatctgcgtgaaacgtcccgttggtggcgtcgtg
 tgggtctggcgacaaaactgcactttgctgtgaccgcctgattgagagcttctactgggccgtgggtgtagcattcgaaccgcaact
 ccgactgccgtaactccgtcgcaaaaatgtttcttctgtaaccattatcgacgatatctacgatgtatacggcacctggacgaactgga
 gctgtttactgatgcagttgagcgttgggacgtaaacgccatcaacgacctgccggattacatgaaactgtgctttctggctctgtataac
 actattaacgaaatcgctacgacaacctgaaagataaagggtgagaacatcctgccgtatctgacaaaagcctgggctgacctgtgca
 acgctttctgcaagaagccaagtggctgtacaacaaatctactccgacctttgacgactacttcggcaacgeatggaaatcctcttctg
 gccgctgcaactgggttctgcttacttcgctgtcgtgcagaacattaaaagggaagagatcgaaaacctgcaaaaataccatgacacc
 atctctcgtccttccatattctccgtctgtgcaatgacctggctagcgcgtctgcggaattgcgctggtgaaaccgcaaatagcgttt
 ctgtttacatgcgcactaaaggatctccgaagaactggctaccgaaagcgtgatgaatctgatcgaatcggaaaaagatgaa
 aaggaaaaactgggtgtagcctgttcgcgaaaccgttcgtgaaaccgcgatcaacctggcacgtcaatctcactgacttatcataa
 cggcgacgcgcatacctctccgatgagctgacctgcaaacgcgttctgtctgtaatcactgaaccgattctgccgtttgaaacgctaag
 gatccgaattcagctccgtcgacaagcttgcggccgactcagcaccaccaccaccactgagatccggctgctaacaagc
 ccgaaaggaagctgagttggctgctgccaccgctgagcaataactagcataacccttggggcctctaaacgggtcttgagggtttt
 tctgtaaaggaggaaactatatccggat (SEQ ID NO:30)

Example 7: Productive Mutations, Combinable Mutations and Suitability Score

[0377] Productive positions are described as those positions within a molecule that are most useful for making combinatorial variants exhibiting an improved characteristic, where the position itself allows for at least one combinable mutation. Combinable mutations can be described as those substitutions in a molecule that can be used to make selected combinatorial

variants. Combinable mutations do not significantly decrease expression, specific activity or growth, while at the same time improving at least one desired characteristic of the molecule such as growth or specific activity. Positions in IspS containing all combinable mutations were determined using performance index (PI) values resulting from the DMAPP assay for specific activity and protein determination, as described in Example 3. Productive positions are the positions which have shown a certain degree of tolerance for multiple substitutions, while at the same time meeting a set of criteria for combinability as set forth below.

[0378] When evaluating the data set, the most productive positions were determined when the following criteria were applied:

[0379] Positions containing substitutions where the minimum performance indices (PI) relative to wild type IspS for specific activity and expression are greater than or equal to a PI of 0.9 and where at least one PI relative to wild type IspS for specific activity or growth is greater than or equal to a PI of 1.0 (Group A).

[0380] Positions containing substitutions where the minimum performance indices (PI) relative to wild type IspS for specific activity and expression are greater than or equal to a PI of 0.8 and where at least one PI relative to wild type IspS for specific activity or growth is greater than or equal to a PI of 1.2 (Group B).

[0381] Positions containing substitutions where the minimum performance indices (PI) relative to wild type IspS for specific activity and expression are greater than or equal to a PI of 0.5 and where at least one PI relative to wild type IspS for specific activity or growth is greater than or equal to a PI of 1.5 (Group C).

[0382] Groups A, B, and C further contain positions that have differing degrees of tolerance for multiple substitutions. To measure this degree of substitutions tolerated, a Rank was assigned to each position. The Rank was assigned according to the percentage of the substitutions within each position that fall within groups A, B, or C. Combinable positions and substitutions are shown in Table 23.

[0383] The criteria to determine the Rank for productive positions are as follows:

[0384] Positions where less than 15% but greater than 0% of the substitutions at a given position fall within groups A, B, or C are given a Rank of "1".

[0385] Positions where less than 30%, but greater than, or equal to 15% of the substitutions at a given position fall within groups A, B, or C are given a Rank of “2”.

[0386] Positions where less than 50%, but greater than, or equal to 30% of the substitutions at a given position fall within groups A, B, or C are given a Rank of “3”.

[0387] Positions where greater than, or equal to 50% of the substitutions at a given position fall within groups A, B, or C are given a Rank of “4”.

[0388] Substitutions are further assigned a Suitability Score based on the group(s) the substitution is a member of, and where a higher score represents a substitution more suitable for use in making combinatorial variants. Suitability scores are represented and defined in Table 23. Suitability scores and Rank for individual substitutions of IspS that fit the above criteria are represented in Table 23.

Table 22. Suitability Score for the defined groups.

Substitutions Occur in Group(s):	Suitability Score
A, B and C	+++++
A and B	++++
A or (B and C)	+++
B	++
C	+

Table 23. Suitability Score and Rank of positions in IspS with substitutions within the positions that are combinable.

POSITION	RANK	VARIANTS SUITABILITY SCORE (+)	VARIANTS SUITABILITY SCORE (++)	VARIANTS SUITABILITY SCORE (+++) WT AA 1ST	VARIANTS SUITABILITY SCORE (++++)	VARIANTS SUITABILITY SCORE (+++++)
2	4	Q	FGLR	EHIS	CDNTV	AKP
3	2			AEGKNQRT		
6	1			S	NT	
13	2	N	M	SQT		
17	1	T		DE		
18	3	H		YFMN	AQR	DEKS
19	1			LF		Y
20	2			LIV	T	
21	1	M		S		W
22	3	AHKNRTY		S	Q	
23	2	K	EG	DT	N	
24	2	AMS		T	C	LV
25	2		NQ	DAES	T	
26	4	Y	L	EGNQRT	DHKMRSV	C
27	4		RY	SEFKV	ACGHILMPQ	DN
28	3		CW	IEFMP	DN	
29	2		G	EDPRT	Q	N
30	3			VNQ	ADEMRT	
31	2	L	D	YQW	N	
32	2		C	KDGNR	E	
33	2	K	E	DN		
34	2			KDEQS		
36	4	HTWY	GS	KFR	ACDEMNPQ	
37	4		WY	KFI	AEGHMNRT	CDPQS
40	3			ACDEFMNPQV		

41	3	M		ECDFNQSV		
42	3	C	L	VAST	FI	M
43	2		G	RIQ		
44	2			RADKMY	NQ	
45	2			ECMNQ		
46	1			IFV		
47	4	C	F	NEIKRV	AGHMQTW	DS
48	4			NACEFLQRS	HIK	DGT
49	3			EGHIRSW	AC	LV
50	4		IQV	KCGMNPR	ADEFHSY	
51	2			AEGLQT		
52	1		D	E		
53	3		CGMS	FD	EHNPQV	
54	2			LACEHIQ	M	
55	3		PW	TAHNQSY	CDE	
56	4			LHQRS	CN	EFGIKTVY
57	1			LI		
58	3	ADFGHRSY	C	E	N	
59	3		C	LFMSVY	HT	
60	1			ICV		
62	1			NV		
63	1	C		V		
68	1			L	I	
69	1		AQ	G		
71	2	HLV		RI	KM	
73	1			ED		
74	3	Q	AC	SGMP	DENY	
75	1			DE		
77	2			RANTV	L	
78	4			GEIKNPQVW	ADFLM	
79	3	C	LN	AMRY	QT	
81	4		W	DCEHLN	AFGMRSTV	Q
82	4		D	RCFGLW	AEHIKMQS	NTVY
83	2			FGHILV	W	
84	4			VFHLNQRSTWY	A	M
85	2		GM	SCLNR		
86	2		GQ	SCN	ADM	
87	4		HMQSV	GCEFKNLT	DP	
88	2			GCDIVWY		
89	2		HM	FCI		
91	4		F	ACDEGHLRSTVY	KW	
92	3		D	VACEFGILQW		

93	4		M	THIQVW	ADEGLNPY	CFRS
94	4		FQ	KCVY	ADEHILMNRS T	GP
95	4		R	TCHKM	AEPQSVY	DFGINW
96	2	CG	DNQ	S		
97	2			LAMP	F	
98	4		P	HCSTVW	ADFGILMNQ	
99	4		DNS	GACHPQT	EFM	
100	2		C	TAILMV		
101	1			AS		
102	1			LM		
103	2		M	SACGL		
107	2		A	LCF		
109	2			QCNS	E	
110	1		A	H		
111	1			GA		
113	2		MPQ	ECHV		
114	2		AM	VC		
115	2			SDY	A	
116	4			QGHLSTV	ACDEIP	
117	3		QW	EADI	CFLMV	
118	2			AIV	M	
119	1			FLM		
120	4	IL		SADEFKNRWY	HTV	CGMQ
121	2		FM	GDLVW		
123	2			KISWY	LT	V
124	1			DCE		
125	4		CP	QADGHKLNSTVW	EIY	
126	4	FR	GSY	N	ACDMTV	E
127	3	L	EMNS	GDFW	C	
128	3		ELMST	NA	CDPQ	
129	2		W	FLY		
130	3			LACDQVY	E	
131	4		N	EDFGR	ACPQSV	HKLMTWY
132	4		QS	N	CDFHLRWY	IP
133	3			LEGIPQTVY	D	
134	4		CFP	KDGHILNRWY	EMQSTV	A
135	2	G		EHS		
136	1			DN	E	
137	4			IACDGPQSV	EHN	T
138	4	L	FMQT	KCDEPRSV	IN	
139	3	DG		APSTV	N	CQ
140	3	C	A	INQSTV	MW	

143	4		I	LAFGNRW	S	CDEHKMQTVY
146	1			AM		
149	1		W	L		
150	1		GS	A		
151	4			LEGMNQRSTVW	CHI	AF
152	2			EADIMP		
153	1			GD	C	
155	4			NEKM	ITVY	ACGHQRSW
156	4	V	ACGMQS	IEKLRY	DNT	
158	1			DE		
159	2	D	C	E	M	
160	2		GY	AFHS	I	
161	3	M		KLRSY	ACNQ	
162	3		GH	VDFNPT	S	
163	3			FCHIMVWY	EQ	
164	2	C	S	A	T	
165	1		D	I		
166	4		Y	SCEHKPQVW	ADG	N
167	4		Y	HCLP	AEGKMRSTW	FINQV
169	4		P	KEGR	DIMST	ACHNQV
170	4			EGINR	HKMQTV	LSWY
171	4			LCEGIMW	HKRS	ANQTVY
172	3	V	HKT	SGNQR	AC	
173	1		Q	E		
175	4			KAGHNPTV	S	CFIQR
176	2			IACNQV	M	
177	3	S		GDEHNPT	AC	
178	4		C	KDEGILMNPQVY	AFRST	
179	4			EGIPQSTVWY	ACLMN	D
180	4	KMR	E	LFHVW	CQT	AI
181	4	GIK	CTY	AFMNW	HQSV	
182	2			EHN	S	
183	2		M	QAL		
185	1		A	ND		
187	1			ACS		
188	1		M	L		
190	2	C		L	IM	
193	1		Q	HW		
194	2	HWY		RI	L	
196	1			TV		
197	3	AEHIPT		QG	S	CDN
203	1		F	W		

204	2			SAFMWY	C	
208	1		A	Y		
210	1			KM		
211	4	L	C	KDEFGHIMRSTV	ANQ	
212	2			EADMPQT		
215	2		W	NDY	CH	
216	2			QAEN		
217	3	AM	D	VCEKNPQT	I	
218	1	M		LV		
219	2	A		LIMV	C	
220	1			EDN		
221	1			L	M	
222	1			AS		
223	1			IC		
224	2			LACTV		
226	1			YF		
228	2	L		MHR	FY	
229	2			IA	V	C
231	4	M	CI	SDGHRV	KQT	A
232	1			VQ	I	
234	1		R	Q		
235	2		E	RADN	K	
238	1	K		R		
240	2	ILM		TV		C
241	2	L		SC	AMT	
242	4			RKL	ADEHIMNQS T	G
245	4	AEGN		R	IL	CKMQTV
246	2		E	RHQ	DK	N
247	2	I		V	T	LM
249	1	I		L		
251	3			TH	AGKR	DENPQS
253	1			LM		
254	2	R	CET	H	D	
260	1			LM		
268	1			VI		
270	1			VI		
271	2			AS	CV	T
272	2	HS		FQ	DGPW	
275	1			QE		
276	2		M	YFHQ		
278	4		FHIK	DLMRY	AENQSTVW	
279	1			C	A	

282	3	HIW	R	SC	AQ	Y
296	1	M		I	V	
297	1		F	Y		
302	1			T	H	
306	1			L		C
307	1			EQR		
308	1		S	L		
314	1		F	EH		
315	2	M		RGK		
317	2	C		DS	EQ	N
319	1			N	F	M
320	2			ANT	C	
321	1			ILM		
322	1		D	N		
323	1			DIT		
327	2	NQ		Y	M	
328	1			ML		
329	2	A		KGQR		
331	1			CT	P	
332	1			FY		
333	1			LF		
336	1		M	Y		
343	1			AIV		
345	1			DY		
346	1			NA		
348	2	F		K	RY	
350	2			KHWY		
351	2			GEM	DN	
352	2			EFIMV		
356	1			PMS		
357	2		N	YE	M	
358	1	M	A	L		
359	1	A		T		
360	1			KQ		
361	2	M		AQSV	T	
363	1		C	AS		
364	2			DNT	EV	
365	2		IV	L	CM	
366	1			CA		
367	2			NDEM		
368	2	LY		ADQ	N	
369	4		KW	FHQ	MNRTV	CDEGS

370	4			LADEFHNRSTV	GQ	
371	4			QGHINPRTWY	CS	F
373	1		S	A	G	
376	1			L		IM
377	2	FHV		Y	W	
378	1			ND		
379	4	CDNP	AHT	KERS		GQ
380	3		GR	SKN	ACDQTV	E
383	1			TQ	S	
386	2			DKS	EN	
389	4	F	L	GCMPRT	HI	ADEKNQSV
390	1		T	NS		
392	3	C	A	WFM	ISTV	Y
393	3	LM		KHR	Q	CITV
402	2			VFIL		
403	2	ACT		F		
405	1			YF		
407	1			A	G	
408	2		GN	VQS	I	T
409	2		R	VCQS	HI	T
410	4		PY	QEGHIR	CDKLMT	
411	2	CDE		N	G	
413	1			KP		
414	3		A	KCHIQ	EGLNP	
415	2	MNP	S	E		
418	1			EN		
421	2	PW		Q		H
422	4	Y	CESV	KGHQR	ANT	D
423	4	ACDEFHMTV		YG	Q	NS
424	3		CN	HDGIST	EPQV	
425	4	ACEFKLMNQR STV		DP		
426	2	GY		TAMQ		
428	2	AMN	R	SV	EQ	
429	4	M	SV	RACDGHKN	ILTWY	EFQ
431	1			SG		
432	2		QT	HAM	E	
436	2	I		L	MY	
437	3	EFLW	H	CN	KT	M
440	1	M		L	I	
443	2	G		A	R	Q
444	2		K	SNQT	P	DE
445	1	QS		A		

447	3	VY		IKR	AEMQS	T
448	3	L		AHST	EMNPQV	
453	1	V		A		
455	1	G		S		A
457	2		Q	SD	NT	
458	1			C		T
460	2		C	MAEG	QRS	
461	3	F		RN	DEGQST	A
462	1			TS	Q	
463	2	F		KGN	ADE	
464	4	K	T	GADEFHVY	LR	CMNQS
465	2			I	ACGST	
466	3		A	SEGKNT	P	D
467	1			EN		
468	2			EANPQ	D	
469	2	Y		LAN		
470	2			A	M	IL
471	2	I		TN	EHQ	
472	2	R	Q	EAGN	DS	
473	2			S	LV	I
475	1			MI	T	
476	1	E		N		
480	2	DV	MQ	E	N	
481	1	N	V	T		
484	2	CH		KA		
487	1	L		K		
488	1	C	A	E		
489	4	ADEFGMNSTV	Q	KR		
490	4	CRS	NV	LIY	ADEFHM	
491	4	IL	D	GACMNQ	EKSTVY	
492	2		A	GTV	C	
493	3	EM	APQ	SCGKV		
494	2			LGIQV	D	
495	1	Y		F		
496	2	IK	EQS	A	PT	
497	2	L	GNR	KMT		
498	2	A	ES	P		
499	1	M		F		
500	3		HQ	VIY	LM	AC
501	1			EN	D	
502	3	FL	QS	TH	ACRV	M
503	2	F		ALM	I	

504	1			IL		
506	2			LIV		M
509	2	STV		QA		
510	2			ST	CV	
511	2		Y	HIM		
512	1	Q		C		
513	4	AEILMQRY		TS	V	CGKN
515	2	AGV		HQ	N	
516	1		DR	N		
517	1			GP		
519	2	W		AC	ST	
522	1			SAK		
525	3		G	E	ACPQS	FR
528	1			RK		
529	1			KA		
531	4	CD		VGN	AMT	EHKQRS
534	2	GM		VAS		
535	2	AGM		ICST		
536	2	K		TM	AFG	
537	2	IV		EHNQ	KT	
538	1	K		P		
539	1			IV		
540	2		D	LEQRV	AP	
541	1	L		P	M	
542	2	IL	QY	FM	P	
543	1		D	E		
544	2			RGNPQS	C	

Example 8: Less combinable improved variants with enhanced specific activity or growth activity

[0389] Table 24 lists variants that were either in suitability groups B or C, or not listed in Table 23. These “less combinable” variants did not fit the criteria for combinability as described above, yet displayed improved performance for either specific activity or growth upon retest.

Table 24. Positions in MEA *P. alba* with less combinable improved mutations that displayed PI specific activity values > 1.3

Residue	Position	Mutation	Location	Charge Change	% Surface Accessibility	Function
S	22	K	N-term	1	54	alternate surface interactions

S	22	R	N-term	1	54	alternate surface interactions
K	36	H	N-term	-1	33	alternate interactions with nearby residues
K	36	W	N-term	-1	33	alternate interactions with nearby residues
R	43	E	N-term	-2	52	alternate interactions with nearby residues
E	58	F	N-term	1	20	alternate interactions with nearby residues
G	87	S	surface loop	0	35	alternate surface interactions
G	87	R	surface loop	1	35	alternate surface interactions
F	89	D	Buried	-1	5	alternate interactions in pocket
F	89	E	Buried	-1	5	alternate interactions in pocket
A	118	E	Buried	-1	10	alternate interactions in pocket
A	118	P	Buried	0	10	alternate interactions in pocket
L	151	Y	surface loop	0	14	alternate interactions in pocket
G	153	P	surface loop	0	40	affect shape of loop
Q	234	R	Hinge Region	1	24	alternate surface interactions
V	247	I	dimer interface	0	5	alternate interactions in pocket
H	254	C	surface loop	0	89	alternate surface interactions
S	282	H	Buried	0	8	alternate interactions in pocket
S	282	W	Buried	0	8	alternate interactions in pocket
S	288	A	Buried	0	0	alternate interactions in pocket
S	288	T	Buried	0	0	alternate interactions in pocket
S	288	Y	Buried	0	0	alternate interactions in pocket
A	391	G	Buried	0	0	alternate interactions in pocket
W	392	C	Buried	0	7	alternate interactions in pocket
C	437	L	Buried	0	7	alternate interactions in pocket
A	443	G	SubLoop	0	4	affect shape of loop
A	443	S	SubLoop	0	4	affect shape of loop
I	447	V	SubLoop	0	23	alternate interactions in pocket
T	481	Y	Buried	0	0	alternate interactions in pocket
E	488	L	Buried	1	13	alternate interactions in pocket
T	502	F	Buried	0	1	alternate interactions in pocket
F	542	N	surface loop	0	17	alternate interactions with nearby residues

Table 25. Positions in MEA *P. alba* with less combinable improved mutations that displayed PI values > 1.3 for OD Max at 40 μ M and 100 μ M IPTG

Residue	Position	Mutation	Location	Charge Change	% Surface Accessibility	Function
V	30	K	N-term	1	60	replace surface hydrophobic w/ charged
K	134	C	N-term helix	-1	37	size reduction, improved interaction with neighboring His

K	134	P	N-term helix	-1	37	loop stabilization
L	143	I	N-term helix	0	1	alternate interactions in hydrophobic pocket
I	156	G	N-term helix	0	13	allow more flexibility in loop
E	159	D	N-term helix	0	32	alternate interactions with nearby residues
E	159	G	N-term helix	1	32	alternate interactions with nearby residues
E	159	Q	N-term helix	1	32	alternate interactions with nearby residues
S	172	V	N-term helix	0	47	affect shape of loop
K	414	F	surface loop	-1	70	alternate surface interactions
Q	421	R	surface loop	1	23	alternate surface interactions
Q	421	D	surface loop	-1	23	alternate surface interactions

Example 9: Identification of Isoprene Synthase Variants with Improved Kinetic Properties

[0390] Isoprene synthase variants were previously selected based on improved specific activity in defined conditions compared to isoprene synthase variant MEA-*P.alba*. In this study, in vitro kinetic constants (k_{cat} , K_M and K_{iDMAPP}) for the ability of the variant isoprene synthases to catalyze the conversion of DMAPP to isoprene were determined by fitting data to the following rate equation:

$$\frac{rate}{[Isoprene\ synthase]} = \frac{k_{cat} * [DMAPP]}{K_M + [DMAPP] * (1 + \frac{[DMAPP]}{K_{iDMAPP}})}$$

[0391] Without being by theory, isoprene synthase variants that display increased k_{cat} values have the potential ability to catalyze the conversion of DMAPP to isoprene more efficiently (if all other kinetic parameters are held constant). Isoprene synthase variants that display decreased K_M values with respect to the parent isoprene synthase molecule have the potential to maintain decreased concentrations of DMAPP compared to the parent isoprene synthase (if all other kinetic parameters are held constant). Isoprene synthase variants that display increased K_{iDMAPP} values have the potential to catalyze the conversion of DMAPP to isoprene more efficiently (if all other kinetic parameters are held constant). The described improvements in enzymatic properties can potentially benefit the production of isoprene for commercial applications.

Methods: Cell growth

[0392] All chemicals were purchased from Sigma unless otherwise specified. All SEL strains were obtained from the isoprene synthase SEL screen as having desirable properties as described above. 96-well deep well plates (Axygen) containing 0.6 mL LB and 50 µg/mL kanamycin were inoculated with the *E. coli* strains and grown overnight at 34°C at 600 rpm. The next day, 48-well plates (Axygen) containing 2 mL TM3 and 0.1% yeast extract, 1% glucose, 50 µg/mL kanamycin and 8 mM MgSO₄ was inoculated with 40 µL of each overnight culture. Cells were grown at 34°C to OD₆₀₀ of approximately 0.5 and induced with 40 µM IPTG for all strains containing isoprene synthase variant genes and 25-60 µM IPTG (in 5 µM steps) for cells containing the parent isoprene synthase gene. Cells were grown for one additional hour and then 10 mM mevalonate was added to the cultures. The absorbance at 600 nm was measured every 1 h, starting at induction, using a Spectramax M5 UV-Vis spectrophotometer. 4 hours after mevalonate addition, the specific isoprene productivity was measured and then the cells were centrifuged for 20 minutes at 3000 RPM in a Sorvall Legend RT equipped with a 6445 swinging bucket rotor. Supernatant was removed and the pellets were frozen and stored at -80°C until cell lysis was performed.

Specific Productivity Assay

[0393] 100 µL of cells were transferred to a Zinsser 96-well glass block, sealed with aluminum foil lids, and incubated for 30 minutes at 34°C in a Shel Lab shaking incubator at 600 RPM. The glass block was then transferred to an 80°C water bath for 2 minutes followed by analysis of the headspace gas by GC-MS. The specific productivity of each culture was determined by combining the absorbance data for the cultures (immediately prior to the glass block procedure) with the isoprene production rate measured by GC-MS.

GC-MS measurement of isoprene

[0394] The amount of isoprene in units of nmole/g_{wcm}/hr can be measured as disclosed in U.S. Pat. No. 5,849,970, which is hereby incorporated by reference in its entirety, particularly with respect to the measurement of isoprene production. For example, two mL of headspace (e.g., headspace from a culture such as 2 mL of culture cultured in sealed vials at 32°C with shaking at 200 rpm for approximately 3 hours) are analyzed for isoprene using a standard gas chromatography system, such as a system operated isothermally (85°C) with an n-

octane/porasil C column (Alltech Associates, Inc., Deerfield, Ill.) and coupled to a RGD2 mercuric oxide reduction gas detector (Trace Analytical, Menlo Park, Calif.) (see, for example, Greenberg et al, Atmos. Environ. 27A: 2689-2692, 1993; Silver et al., Plant Physiol. 97:1588-1591, 1991, which are each hereby incorporated by reference in their entirety, particularly with respect to the measurement of isoprene production). The gas chromatography area units are converted to nmol isoprene via a standard isoprene concentration calibration curve. In some embodiments, the value for the grams of cells for the wet weight of the cells is calculated by obtaining the A_{600} value for a sample of the cell culture, and then converting the A_{600} value to grams of cells based on a calibration curve of wet weights for cell cultures with a known A_{600} value. In some embodiments, the grams of the cells is estimated by assuming that one liter of broth (including cell medium and cells) with an A_{600} value of 1 has a wet cell weight of 1 gram. The value is also divided by the number of hours the culture has been incubating for, such as three hours.

Cell Lysis

[0395] Materials: Tris/ NaCl pH 7.6, $MgCl_2$, 4-(2-Aminoethyl) benzenesulfonyl fluoride Hydrochloride (AEBSF), DNase I, DMAPP Triammonium salt(Cayman chemicals), Lysozyme (Epicentre), 96-well Zinsser Glass Block, Seal & Sample Aluminum foil lids (Part No:538619)(Beckman coulter), Nunc* MicroWell* 96-Well Plates, Polypropylene, High Volume (Part No:2449946), (Thermo Scientific) (VWR).

[0396] Procedure: 200 μ L lysis buffer (100 mM Tris, 100 mM NaCl pH 7.6, 1 mg/ml BSA, 50U/ μ l Epicentre readylyse lysozyme, 0.1 mg/ml DNAase, 0.5 mM AEBSF, 5 mM $MgCl_2$) was added to each well of the 48-well plate, the pellets were re-suspended, and the plate was sealed with aluminum foil lids. The resuspended cells were re-frozen at -80°C for 10 minutes . The plate was then removed from the -80°C freezer and mixed at 450 rpm for 30 minutes at room temperature using the Thermomixer (Eppendorf). The mixture was then centrifuged at 3200 rpm for 10 minutes at 4°C . The supernatant was collected for use in enzymatic activity assays.

Isoprene synthase enzymatic activity assay

[0397] 25 μ L of *E.coli* lysate, containing isoprene synthase, was incubated with 0.25, 0.5, 1, 3, 5, 7, 10 and 20 mM DMAPP, in 100 μ L reactions containing 50 mM $MgCl_2$ and 100 mM Tris/NaCl in a Zinsser 96-well glass block sealed with aluminum foil lids for 30 minutes

at 34 °C. The glass blocks were then transferred to an 80 °C water bath for 2 minutes. Next, the glass blocks were analyzed by GC-FID (see below) to determine the concentration of isoprene generated in the reactions.

GC-FID Analysis

[0398] *Equipment and Materials:* Gas chromatograph (GC), 7890 (Agilent Technologies), Flame ionization detector (FID) 7890 (Agilent Technologies), HP-5ms column, 5%-phenyl-methylpolysiloxane, 15 m x 0.25 mm x 0.25 μm (Agilent Technologies), CTC autosampler (Leap Technologies), 0.2% v/v isoprene, balance nitrogen (Air Liquide), Chemstation with Enhanced Data Analysis (D.03.00.611)

[0399] **Procedure:** 96-well glass blocks were analyzed using GC-FID with the following parameters:

Oven:

Rate (°C/min)	Temperature (°C)	Time (min)
0	37	28

Run Time: 28 minutes

FRONT INLET	
Front Inlet Temperature	110°C
Flow Rate	3.4 mL/min
Flow Mode	Constant Flow
Split Ratio	50:1
Carrier Gas	Helium
FLAME IONIZATION DETECTOR	
Detector Temperature	160°C
Hydrogen Flow	40 mL/min
Air Flow	400 mL/min
Makeup Flow	0.1 mL/min
Makeup Gas Type	Helium

SYRINGE CYCLE

syringe:1.0ml-HS
 JN(37,100,0,1000,GC Inj2,500,50,50,10,600)
 [MACRO JN]
 Syringe Temperature (°C);50;30;150
 Fill Speed (μl/s);SYR.Fill Speed;SYR.Min Speed;SYR.Max Speed
 Fill Strokes ();SYR.Fill Strokes;0;99
 Pullup Delay (ms);SYR.Pullup Del;0;10000
 Inject to;INJECTOR

```

Injection Speed (µl/s);SYR.Inject Speed;SYR.Min Speed;SYR.Max Speed
Pre Inject Delay (ms);500;0;99000
Post Inject Delay (ms);500;0;99000
Flush Time (s);10;0;600
GC Runtime (s);600;30;86400
SET_TEMP(SYR,Syringe Temperature,0,)
MOVETO_OBJECT(Home,,,)
WAIT_FOR_DS()
WAIT_SYNC_SIG(Start,)
GET_SAMPLE(SL.tray,1,SL.volume,0,,,Fill Speed,Pullup Delay,,Fill Strokes,Off,,,)
INJ_SAMPLE(Inject to,Inject,Injected,,,Pre Inject Delay,Injection Speed,Post Inject
Delay,1,)
MOVETO_OBJECT(Home,,,)
START_FLUSH(Flush Time,)
WAIT(Flush Time,)
STOP_FLUSH()
GET_SAMPLE(SL.tray,2,SL.volume,0,,,Fill Speed,Pullup Delay,,Fill Strokes,Off,,,)
INJ_SAMPLE(Inject to,Inject,Injected,,,Pre Inject Delay,Injection Speed,Post Inject
Delay,1,)
MOVETO_OBJECT(Home,,,)
START_FLUSH(Flush Time,)
WAIT(Flush Time,)
STOP_FLUSH()
GET_SAMPLE(SL.tray,3,SL.volume,0,,,Fill Speed,Pullup Delay,,Fill Strokes,Off,,,)
INJ_SAMPLE(Inject to,Inject,Injected,,,Pre Inject Delay,Injection Speed,Post Inject
Delay,1,)
MOVETO_OBJECT(Home,,,)
START_FLUSH(Flush Time,)
WAIT(Flush Time,)
STOP_FLUSH()

```

Calculations/data analysis

[0400] The peak areas were converted to isoprene concentrations by dividing the peak area by the response factor calculated from 0.2% v/v isoprene in nitrogen calibration standards.

HPLC analysis

[0401] *Equipment and Materials:* 96-well polypropylene plates, 46600-666 (Thermo Scientific), Zone Free Films, plate covers, Z721646 (Sigma Aldrich), High Performance Liquid Chromatography, 1200 (Agilent Technologies), Diode Array Detector, G1315D (Agilent Technologies), ProSwift RP-2H column, 4.6 x 50 mm (Dionex), In-line filter with 0.46 µm stainless steel frits¹, 5067-1562 (Agilent Technologies), Trifluoroacetic Acid, 1 mL

ampules, 91707 (Sigma Aldrich), Acetonitrile, 9017-03 (J.T. Baker), Isopropanol, 9079-05 (J.T. Baker), Software, Chemstation (B.04.01).

[0402] *Procedure:* 10 μ L of 2% TFA in water was added to 200 μ L of each sample and the samples were transferred to a 96-well polypropylene plate (see above). A Zone Free Film plate cover was used on the 96-well plate. Mobile phase A consisted of 0.1% TFA in 55:25:20 water:acetonitrile:isopropanol. Mobile phase B consisted of 0.1% TFA in isopropanol. Samples were analyzed by HPLC-UV with the following parameters:

[0403] Gradient: A: 0.1% TFA in 55:25:20 water:acetonitrile:isopropanol. B: 0.1% TFA in isopropanol

Time (min)	%B	Flow Rate (mL/min)
0	0	1.5
0.2	0	1.5
2.0	30	1.5
2.5	100	1.5
4.5	100	1.5
5.0	0	1.5
6.0	0	1.5

Run Time	6 minutes
Injection Volume	40 μ L
Column Temperature	70°C
UV Detection Wavelength	220 nm

Calculations and data analysis

[0404] Chemstation was used to capture and process the data.

Calculation of isoprene synthase kinetic parameters:

[0405] Data from the isoprene synthase kinetic assays were fit to the following modified version of the Henri-Michaelis-Menten equation that takes into account substrate inhibition using Kaleidagraph 4.0 (Synergy Software) to determine K_M , k_{cat} and k_{IDMAPP} values for each isoprene synthase analyzed:

$$\frac{rate}{[Isoprene\ synthase]} = \frac{k_{cat} * [DMAPP]}{K_M + [DMAPP] * \left(1 + \frac{[DMAPP]}{K_{IDMAPP}}\right)}$$

[0406] All K_M , k_{cat} and k_{iDMAPP} data were normalized to the average value calculated from 43 replicates of the isoprene synthase parent molecule.

[0407] Table 26 shows exemplary isoprene synthase variants that display improved k_{cat} values for the conversion of DMAPP to isoprene. All k_{cat} values are normalized to the k_{cat} of the parent isoprene synthase molecule. Isoprene synthase variants with K_{cat} of 1.240593 to 4.720585 are shown in Table 26 and are designated as ++++. Isoprene synthase variants with K_{cat} of 1.130066 to 1.239277 are shown in Table 26 and are designated as ++. Isoprene synthase variants with K_{cat} of 1.001492 to 1.128037 are shown in Table 26 and are designated as +.

Table 26

Variant	k_{cat}
D323F	+++
A118E	+++
K36W	+++
S22K	+++
M228Y	+++
A448L	+++
E488F	+++
E467H	+++
A443S	+++
C331P	+++
A453I	+++
R71K	+++
R71L	+++
A448I	+++
R71M	+++
W392Y	+++
A448V	+++
S282H	+++
T383Y	+++
D323Y	+++
H511Y	+++
A448E	+++
L376M	+++
E488L	+++
S120E	+++
R461A	+++
K414I	+++
S282W	+++
R071K	+++
S493E	+++
W392S	+++
A448Q	+++
S282Y	+++
E537N	+++
I447Y	+++
T240C	+++
A443Q	+++
P538R	+++

S510C	+++
G389D	++
T383H	++
K36E	++
L436Y	++
I447V	++
W392F	++
K161R	++
G99D	++
G99E	++
K161C	++
K414S	++
E537T	++
K393V	++
A443G	++
S510V	++
K36N	++
W392A	++
K161A	++
H254R	++
E472C	++
K161Q	++
K36Y	++
E537C	++
E41Y	++
K161M	++
R43L	++
S120A	++
G087M	++
K36S	++
K36H	++
K348Y	++
E467W	++
S288T	++
I447T	++
D025N	++
V268I	+

E488M	+
E480I	+
T240V	+
M460A	+
R071L	+
K36Q	+
A118P	+
K36T	+
S282I	+
V409I	+
R071I	+
V409T	+
L436F	+
K348F	+
A118Q	+
I342Y	+
S74Q	+
E41P	+
C437Y	+

W392T	+
F89D	+
E41M	+
S510E	+
K36P	+
K393I	+
R71I	+
T381M	+
K374Y	+
E58Y	+
L526Q	+
E543F	+
T240M	+
G111S	+
K463T	+
S120M	+
E135G	+
WT	1

[0408] Table 27 shows isoprene synthase variants that display K_M values less than 1 compared to the parent isoprene synthase molecule for the conversion of DMAPP to isoprene. All K_M values are normalized to the K_M of the parent isoprene synthase molecule. The variants with K_M values from 0.105 to 0.587 are designated as +++. The variants with K_M values from 0.608 to 0.844 are designated as ++. The variants with K_M values from 0.846 to 0.998 are designated as +.

[0409]

Table 27

Variant	KM
S22K	+++
K348F	+++
W392V	+++
W392F	+++
E488C	+++
S22R	+++
R71V	+++
A443R	+++
Q234R	+++
A453V	+++
C437Y	+++
W392C	+++
K463F	+++
P538K	+++
K393L	+++
H254C	+++
L436Y	+++
S21R	+++
C437L	+++
S444D	+++
K374Y	+++
A363L	+++
I447V	+++
S444E	+++
R71I	+++
I504F	+++
E488W	+++
R71H	+++
K36P	+++
T381I	+++
L436F	+++
M460A	+++
A443G	+++
S288A	+++
W392T	+++
E537I	+++
K374Y	+++

R242G	+++
C437M	+++
L436I	+++
L376I	+++
S288Y	+++
W392M	+++
I342Y	++
K414W	++
R461A	++
A443S	++
W392A	++
K463T	++
E488T	++
L526Q	++
T502F	++
T502M	++
E488M	++
K36Y	++
K414R	++
R071I	++
K36W	++
T502L	++
T481Y	++
E472R	++
K36H	++
K36T	++
E415Y	++
E415H	++
E58Y	++
T381M	++
T481V	++
F89E	++
E480I	++
K36Q	++
F89D	++
H254R	++
K161N	++

R071L	++
H424P	++
E415V	++
S22K	++
E58L	++
W392S	++
V268I	++
W392I	++
L526E	++
E537V	++
K36D	++
K393I	+
E41M	+
D025N	+
K36N	+
A118Q	+
G389E	+
S282W	+
K348Y	+
T383L	+
S510V	+
D323Y	+

D323F	+
T383H	+
E41P	+
S288T	+
K36S	+
A453I	+
I447T	+
E472I	+
L376M	+
K161E	+
E488L	+
T381L	+
R071K	+
V409T	+
W392Y	+
E135G	+
F542L	+
M228Y	+
A118P	+
E543F	+
K36E	+
WT	1.000

[0410] Table 28 shows isoprene synthase variants that display K_{iDMAPP} values greater than 1 compared to the parent isoprene synthase molecule for the conversion of DMAPP to isoprene. All K_{iDMAPP} values are normalized to the K_{iDMAPP} of the parent isoprene synthase molecule. The variants with K_i from 1.506 to 5.745 and even no substrate inhibition are designated as +++. The variants with K_i from 1.078 to 1.422 are designated as ++. The variants with K_i from 1.002 to 1.074 are designated as +.

[0411]

Table 28

Variant	Ki
E472R	+++
G389E	+++
R242G	+++
L376I	+++
K161N	+++
S288C	+++
T240V	+++
T481Y	+++
K463F	+++
K393L	+++

S120Q	++
E58L	++
T240M	++
A453V	++
E543F	+
S74Q	+
K414R	+
E415H	+
E415V	+
WT	1.000

Example 10: Correlation of isoprene synthase kinetic parameters with specific isoprene productivity

[0412] The specific isoprene productivity for *E. coli* BL21 strains that express isoprene synthase variants was plotted against the corresponding kinetic parameters for each isoprene synthase variant to determine if specific properties of isoprene synthase variants could be used to predict increased specific isoprene productivity. The specific isoprene productivity was positively correlated with decreased substrate inhibition (K_{IDMAPP}) (Figure 39A, Table 29). In addition, the specific isoprene productivity was positively correlated with decreased K_M values (Figure 39B, Table 29). There was no correlation between the specific isoprene productivity and k_{cat} values (Figure 39C, Table 29).

Table 29

Variant	Specific Isoprene Productivity
S510E	+++
K161M	+++
S120Q	+++
S120E	+++
K161R	+++
S120A	+++
S493E	+++
A448E	+++
T240V	+++
T240M	+++
E537T	+++
K161N	++
I447Y	++

T502M	++
E537N	++
K348Y	++
E543F	++
C331P	++
E537C	++
S120M	++
S288C	++
K161Q	++
K161A	+
T481V	+
G087M	+
E467W	+
A448V	+
T502L	+
L376M	+

CLAIMS

What is claimed is:

1. An isolated polypeptide having isoprene synthase activity, wherein the polypeptide comprises one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X323F, X118E, X36W, X22K, X228Y, X448L, X488F, X467H, X443S, X331P, X453I, X71K, X71L, X448I, X71M, X392Y, X448V, X282H, X383Y, X323Y, X511Y, X448E, X376M, X488L, X120E, X461A, X414I, X282W, X071K, X493E, X392S, X448Q, X282Y, X537N, X447Y, X240C, X443Q, X538R, and X510C, and wherein the polypeptide has improved Kcat compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s).
2. The polypeptide of claim 1 wherein the amino acid substitution is selected from the group consisting of: D323F, A118E, K36W, S22K, M228Y, A448L, E488F, E467H, A443S, C331P, A453I, R71K, R71L, A448I, R71M, W392Y, A448V, S282H, T383Y, D323Y, H511Y, A448E, L376M, E488L, S120E, R461A, K414I, S282W, R071K, S493E, W392S, A448Q, S282Y, E537N, I447Y, T240C, A443Q, P538R, and S510C.
3. An isolated polypeptide having isoprene synthase activity, wherein the polypeptide comprises one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X389D, X383H, X36E, X436Y, X447V, X392F, X161R, X99D, X99E, X161C, X414S, X537T, X393V, X443G, X510V, X36N, X392A, X161A, X254R, X472C, X161Q, X36Y, X537C, X41Y, X161M, X43L, X120A, X087M, X36S, X36H, X348Y, X467W, X288T, X447T, X025N, and wherein the polypeptide has improved Kcat compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s).
4. The polypeptide of claim 3 wherein the amino acid substitution is selected from the group consisting of: G389D, T383H, K36E, L436Y, I447V, W392F, K161R, G99D, G99E, K161C, K414S, E537T, K393V, A443G, S510V, K36N, W392A, K161A, H254R, E472C, K161Q, K36Y, E537C, E41Y, K161M, R43L, S120A, G087M, K36S, K36H, K348Y, E467W, S288T, I447T, and D025N.

5. An isolated polypeptide having isoprene synthase activity, wherein the polypeptide comprises one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X268I, X488M, X480I, X240V, X460A, X071L, X36Q, X118P, X36T, X282I, X409I, X071I, X409T, X436F, X348F, X118Q, X342Y, X74Q, X41P, X437Y, X392T, X89D, X41M, X510E, X36P, X393I, X71I, X381M, X374Y, X58Y, X526Q, X543F, X240M, X111S, X463T, X120M, and X135G, and wherein the polypeptide has improved K_{cat} compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s).

6. The polypeptide of claim 5 wherein the amino acid substitution is selected from the group consisting of: V268I, E488M, E480I, T240V, M460A, R071L, K36Q, A118P, K36T, S282I, V409I, R071I, V409T, L436F, K348F, A118Q, I342Y, S74Q, E41P, C437Y, W392T, F89D, E41M, S510E, K36P, K393I, R71I, T381M, K374Y, E58Y, L526Q, E543F, T240M, G111S, K463T, S120M, and E135G.

7. An isolated polypeptide having isoprene synthase activity, wherein the polypeptide comprises one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X22K, X348F, X392V, X392F, X488C, X22R, X71V, X443R, X234R, X453V, X437Y, X392C, X463F, X538K, X393L, X254C, X436Y, X21R, X437L, X444D, X374Y, X363L, X447V, X444E, X71I, X504F, X488W, X71H, X36P, X381I, X436F, X460A, X443G, X288A, X392T, X537I, X374Y, X242G, X437M, X436I, X376I, X288Y, and X392M, and wherein the polypeptide has improved K_M compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s).

8. The polypeptide of claim 7 wherein the amino acid substitution is selected from the group consisting of: S22K, K348F, W392V, W392F, E488C, S22R, R71V, A443R, Q234R, A453V, C437Y, W392C, K463F, P538K, K393L, H254C, L436Y, S21R, C437L, S444D, K374Y, A363L, I447V, S444E, R71I, I504F, E488W, R71H, K36P, T381I, L436F, M460A, A443G, S288A, W392T, E537I, K374Y, R242G, C437M, L436I, L376I, S288Y, and W392M.

9. An isolated polypeptide having isoprene synthase activity, wherein the polypeptide comprises one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X342Y, X414W,

X461A, X443S, X392A, X463T, X488T, X526Q, X502F, X502M, X488M, X36Y, X414R, X071I, X36W, X502L, X481Y, X472R, X36H, X36T, X415Y, X415H, X58Y, X381M, X481V, X89E, X480I, X36Q, X89D, X254R, X161N, X071L, X424P, X415V, X22K, X58L, X392S, X268I, X392I, X526E, X537V, and X36D and wherein the polypeptide has improved K_M compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s).

10. The polypeptide of claim 9 wherein the amino acid substitution is selected from the group consisting of: I342Y, K414W, R461A, A443S, W392A, K463T, E488T, L526Q, T502F, T502M, E488M, K36Y, K414R, R071I, K36W, T502L, T481Y, E472R, K36H, K36T, E415Y, E415H, E58Y, T381M, T481V, F89E, E480I, K36Q, F89D, H254R, K161N, R071L, H424P, E415V, S22K, E58L, W392S, V268I, W392I, L526E, E537V, and K36D.

11. An isolated polypeptide having isoprene synthase activity, wherein the polypeptide comprises one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X393I, X41M, X025N, X36N, X118Q, X389E, X282W, X348Y, X383L, X510V, X323Y, X323F, X383H, X41P, X288T, X36S, X453I, X447T, X472I, X376M, X161E, X488L, X381L, X071K, X409T, X392Y, X135G, X542L, X228Y, X118P, X543F, and X36E, and wherein the polypeptide has improved K_M compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s).

12. The polypeptide of claim 11 wherein the amino acid substitution is selected from the group consisting of: K393I, E41M, D025N, K36N, A118Q, G389E, S282W, K348Y, T383L, S510V, D323Y, D323F, T383H, E41P, S288T, K36S, A453I, I447T, E472I, L376M, K161E, E488L, T381L, R071K, V409T, W392Y, E135G, F542L, M228Y, A118P, E543F, and K36E.

13. An isolated polypeptide having isoprene synthase activity, wherein the polypeptide comprises one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X472R, X389E, X242G, X376I, X161N, X288C, X240V, X481Y, X463F, and X393L, and wherein the polypeptide has improved K_{IDMAPP} compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s).

14. The polypeptide of claim 13 wherein the amino acid substitution is selected from the group consisting of: E472R, G389E, R242G, L376I, K161N, S288C, T240V, T481Y, K463F, and K393L.

15. An isolated polypeptide having isoprene synthase activity, wherein the polypeptide comprises one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X120Q, X58L, X240M, and X453V, and wherein the polypeptide has improved K_{iDMAPP} compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s).

16. The polypeptide of claim 15 wherein the amino acid substitution is selected from the group consisting of: S120Q, E58L, T240M, and A453V.

17. An isolated polypeptide having isoprene synthase activity, wherein the polypeptide comprises one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X543F, X74Q, X414R, X415H, and X415V, and wherein the polypeptide has improved K_{iDMAPP} compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s).

18. The polypeptide of claim 17 wherein the amino acid substitution is selected from the group consisting of: E543F, S74Q, K414R, E415H, and E415V.

19. An isolated polypeptide having isoprene synthase activity, wherein the polypeptide comprises one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X510E, X161M, X120Q, X120E, X161R, X120A, X493E, X448E, X240V, X240M, and X537T, and wherein the polypeptide has improved specific productivity compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s).

20. The polypeptide of claim 19 wherein the amino acid substitution is selected from the group consisting of: S510E, K161M, S120Q, S120E, K161R, S120A, S493E, A448E, T240V, T240M, and E537T.

21. An isolated polypeptide having isoprene synthase activity, wherein the polypeptide comprises one or more amino acid substitution(s) at one or more residues

corresponding to SEQ ID NO:1 selected from the group consisting of: X161N, X447Y, X502M, X537N, X348Y, X543F, X331P, X537C, X120M, X288C, and X161Q, and wherein the polypeptide has improved specific productivity compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s).

22. The polypeptide of claim 21 wherein the amino acid substitution is selected from the group consisting of: K161N, I447Y, T502M, E537N, K348Y, E543F, C331P, E537C, S120M, S288C, and K161Q.

23. An isolated polypeptide having isoprene synthase activity, wherein the polypeptide comprises one or more amino acid substitution(s) at one or more residues corresponding to SEQ ID NO:1 selected from the group consisting of: X161A, X481V, X087M, X467W, X448V, X502L, and X376M, and wherein the polypeptide has improved specific productivity compared to a parent polypeptide which does not comprise said one or more amino acid substitution(s).

24. The polypeptide of claim 23 wherein the amino acid substitution is selected from the group consisting of: K161A, T481V, G087M, E467W, A448V, T502L, and L376M.

25. A recombinant host cell comprising the polypeptide of any one of claims 1-24.

26. The host cell of claim 25 wherein the host cell is selected from the group consisting of a bacterial, algal, fungal, yeast, cyanobacterial, or Clostridial cell.

27. The host cell of claim 26 wherein the host cell is a bacterial cell.

28. The host cell of claim 27 wherein the bacterial cell is a gram-positive bacterial cell or gram-negative bacterial cell.

29. The host cell of claim 28 wherein the bacterial cell is selected from the group consisting of *E. coli*, *L. acidophilus*, *P. citrea*, *B. subtilis*, *B. licheniformis*, *B. lentus*, *B. brevis*, *B. stearothermophilus*, *B. alkalophilus*, *B. amyloliquefaciens*, *B. clausii*, *B. halodurans*, *B. megaterium*, *B. coagulans*, *B. circulans*, *B. lautus*, *B. thuringiensis*, *S. albus*, *S. lividans*, *S. coelicolor*, *S. griseus*, *Pseudomonas sp.*, *P. alcaligenes*, *Clostridium sp.*, *Corynebacterium sp.*, and *C. glutamicum* cells.

30. The host cell of claim 26 wherein the host cell is an algal cell.

31. The host cell of claim 30 wherein the algal cell is selected from the group consisting of green algae, red algae, glaucophytes, chlorarachniophytes, euglenids, chromista, or dinoflagellates.
32. The host cell of claim 26 wherein the host cell is a fungal cell.
33. The host cell of claim 32 wherein the fungal cell is a filamentous fungi.
34. The host cell of claim 26 wherein the host cell is a yeast cell.
35. The host cell of claim 34 wherein the yeast cell is selected from the group consisting of *Saccharomyces sp.*, *Schizosaccharomyces sp.*, *Pichia sp.*, or *Candida sp.*
36. The host cell of claim 35 wherein the yeast cell is a *Saccharomyces cerevisiae* cell.
37. A method for producing isoprene comprising (a) culturing the host cells of any one of claims 25-36 under conditions suitable for the production of isoprene and (b) producing isoprene.
38. The method of claim 37 further comprising recovering the isoprene.

Figure 1

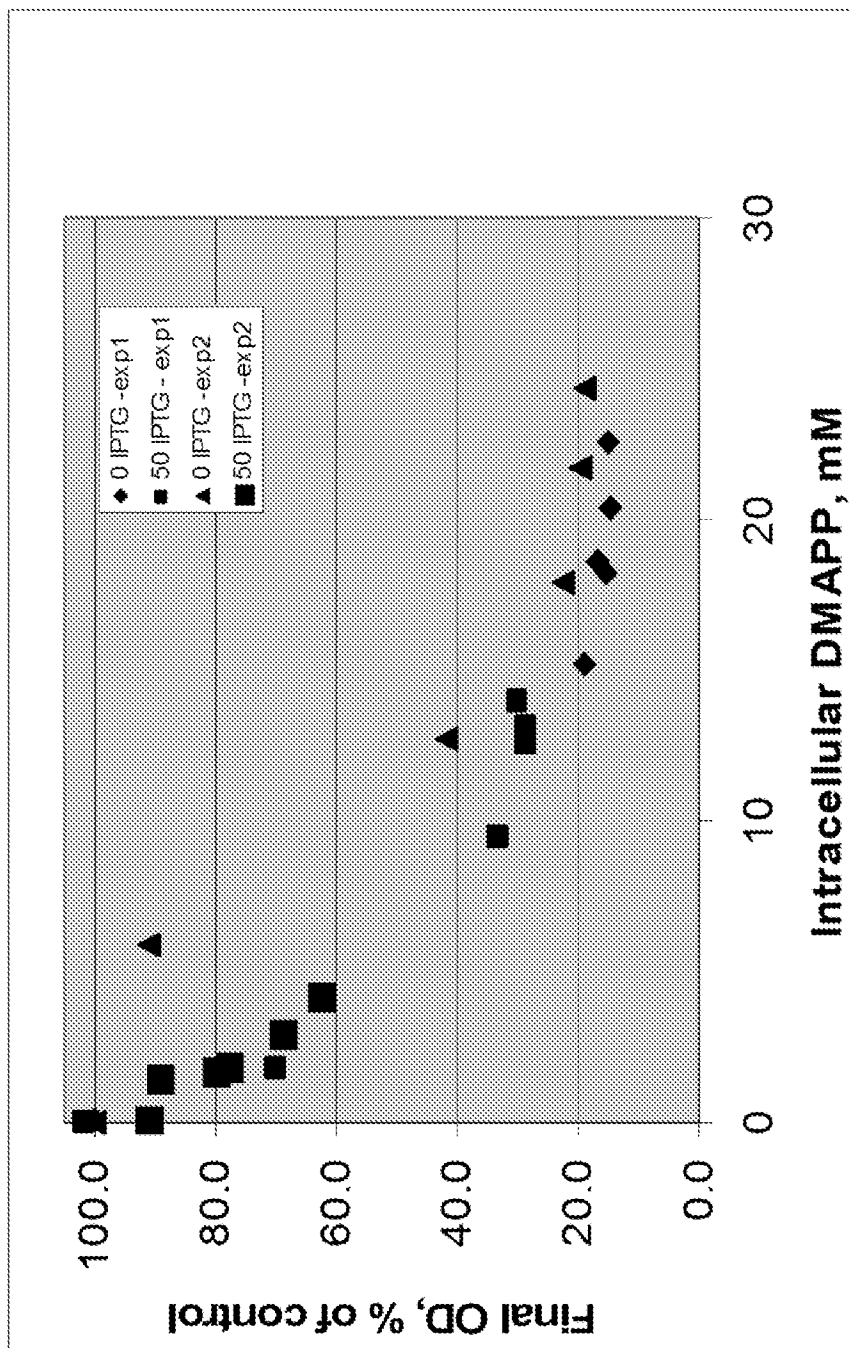


Figure 2

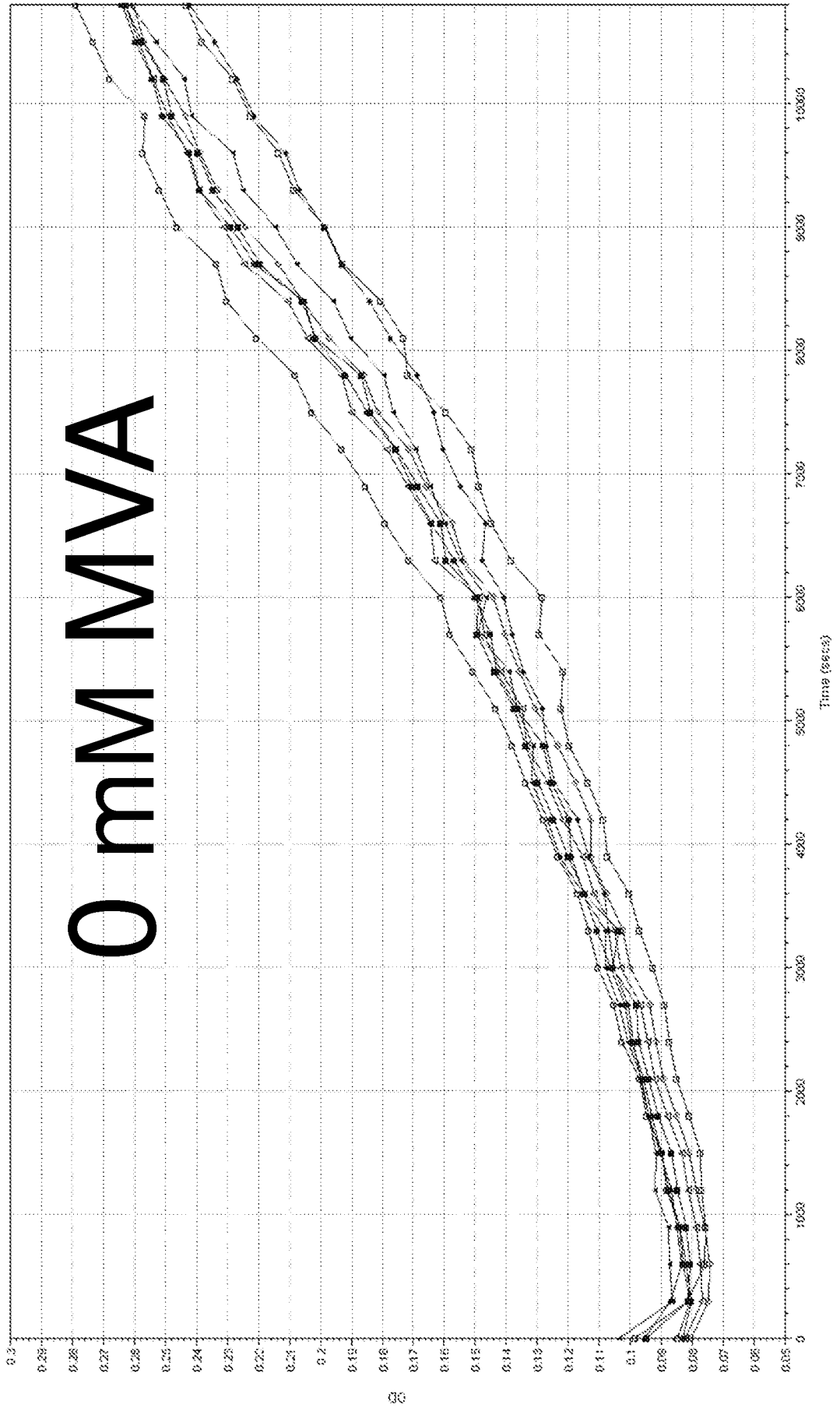


Figure 2 (cont.)

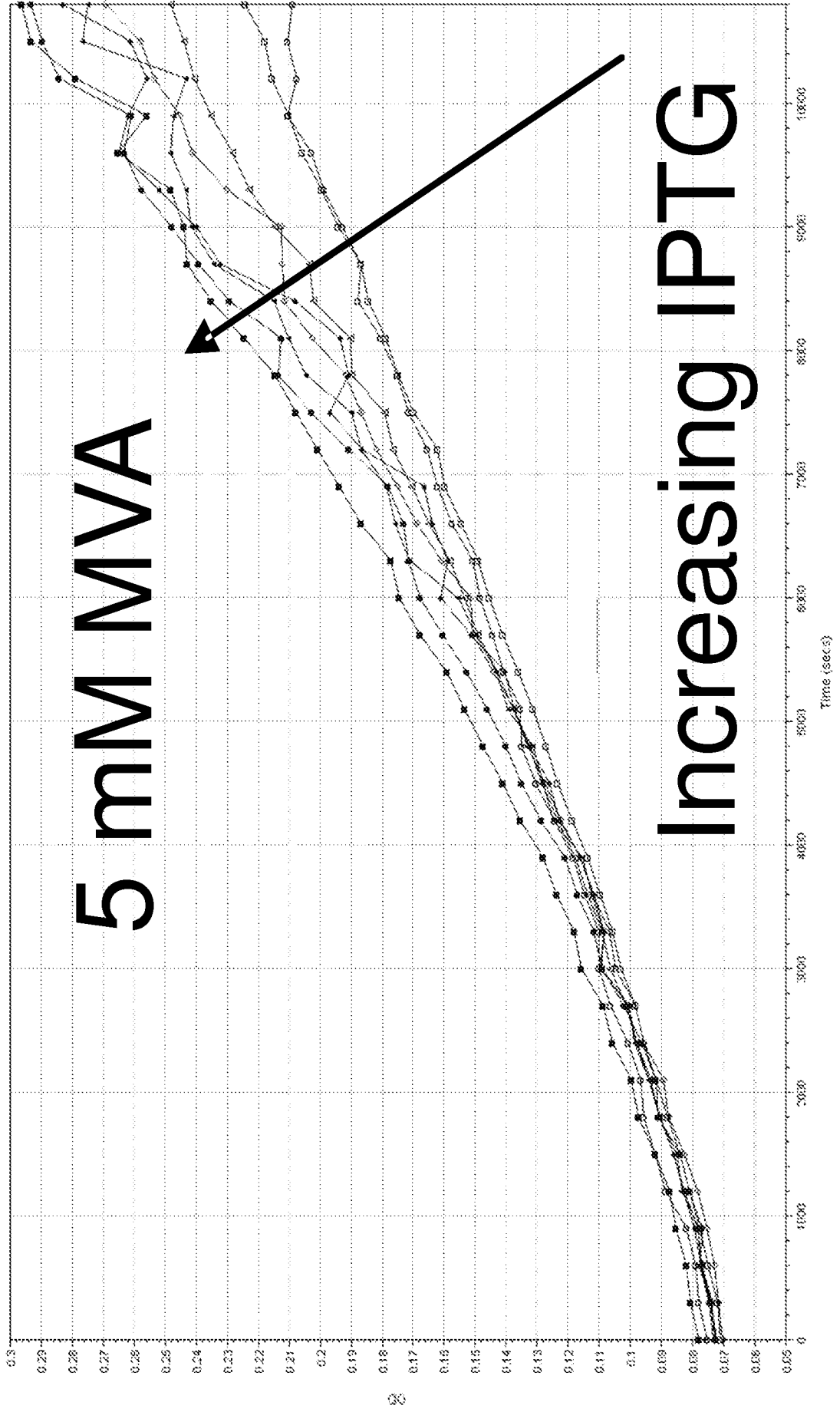


Figure 2 (cont.)

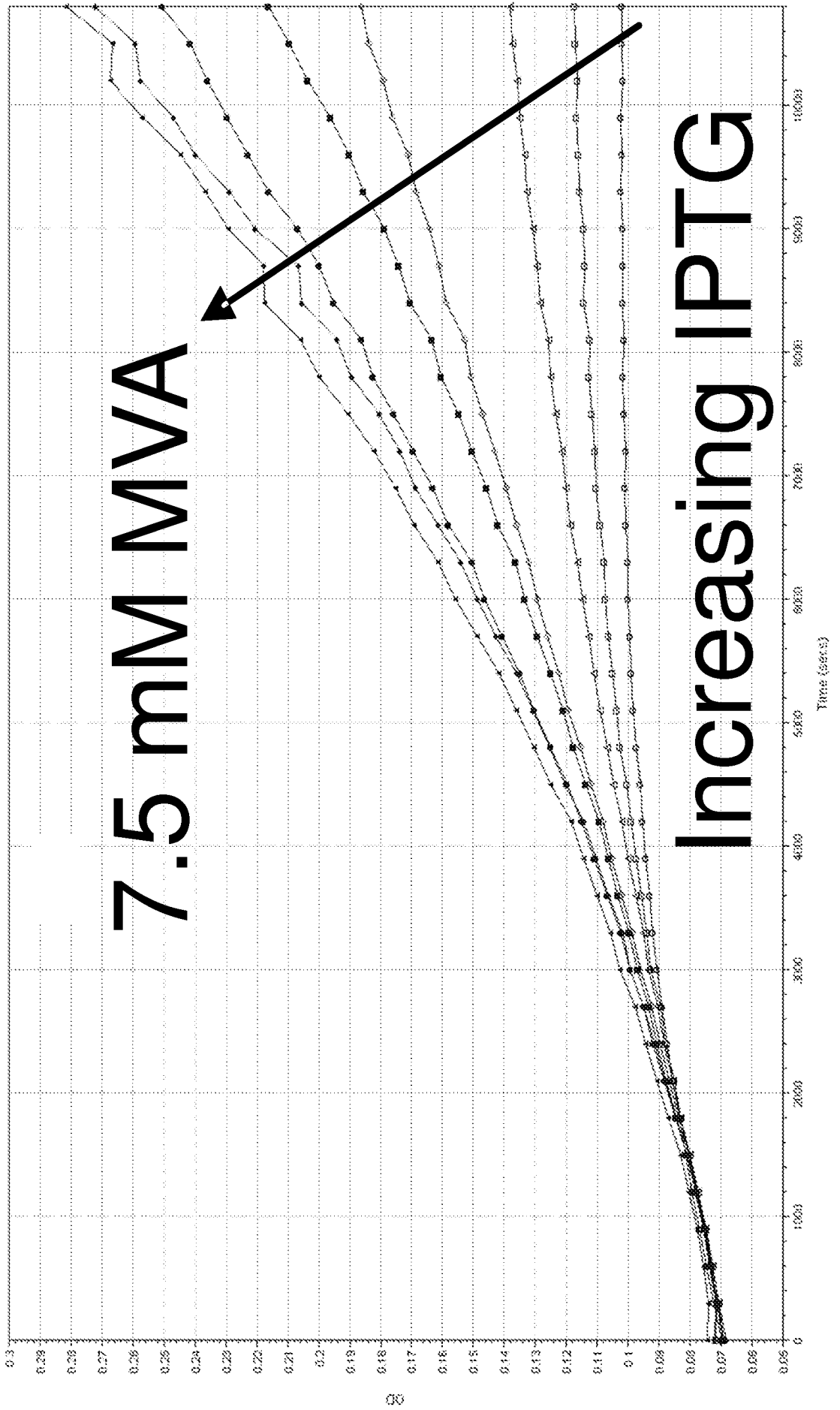


Figure 2 (cont.)

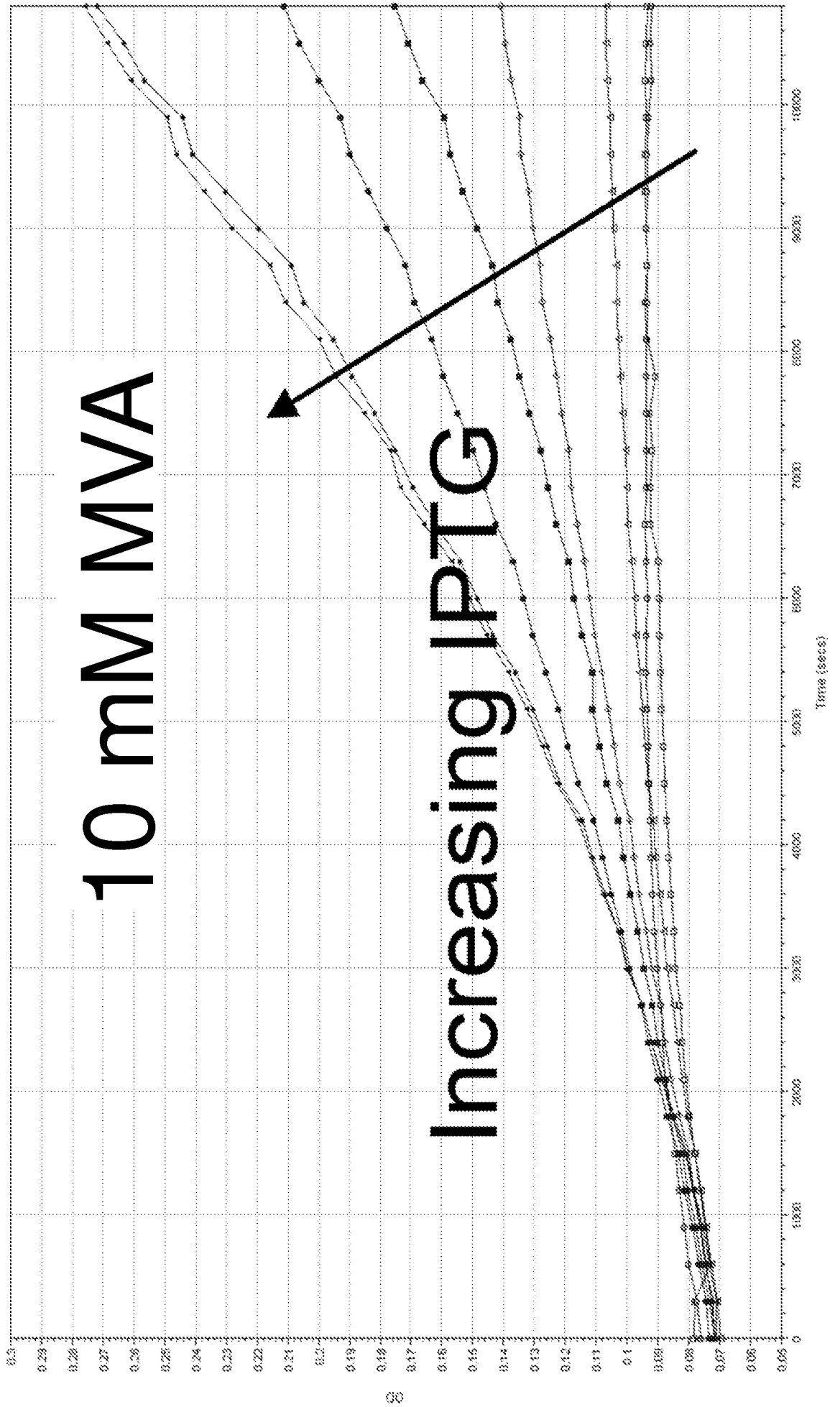


Figure 2 (cont.)

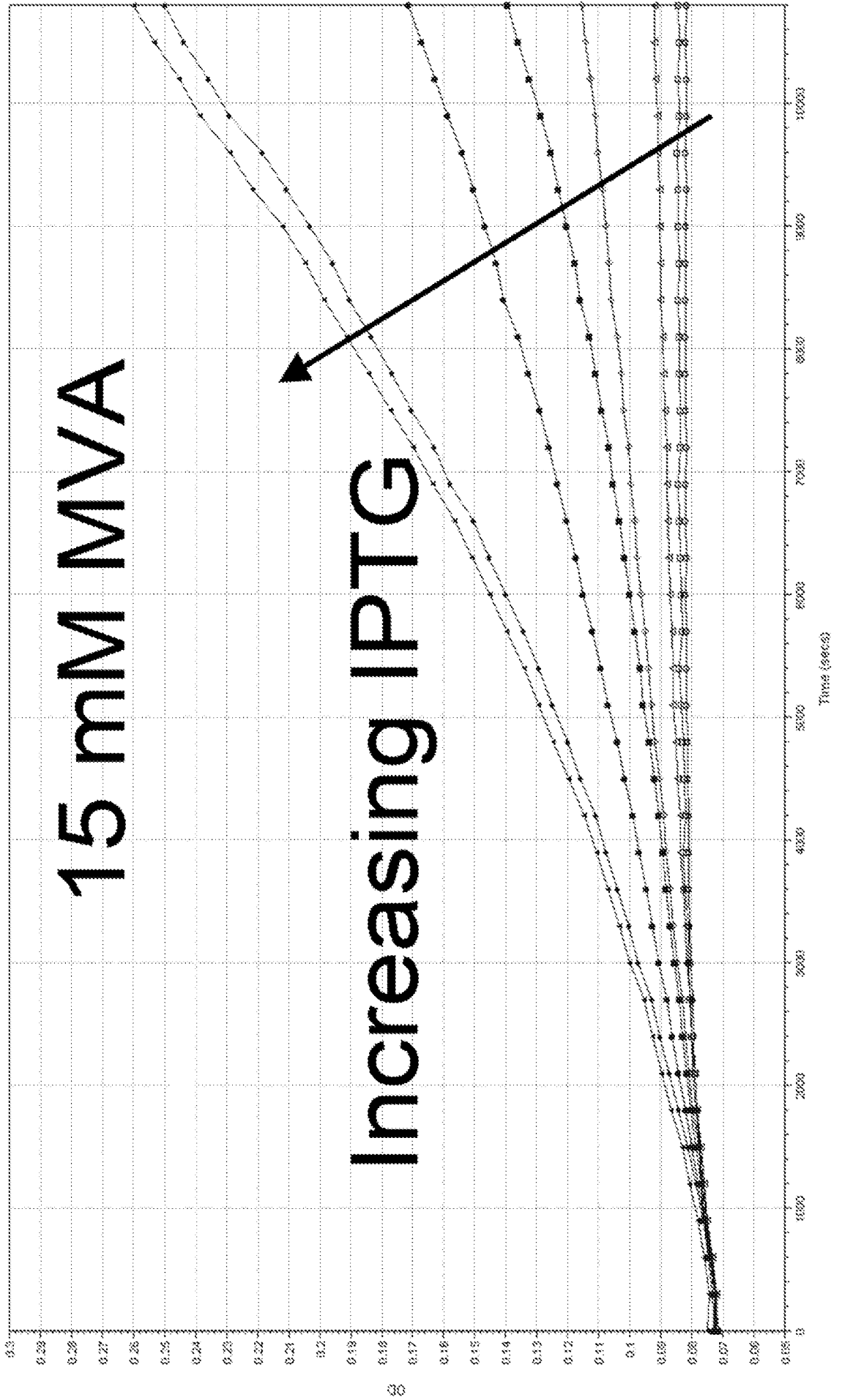


Figure 2 (cont.)

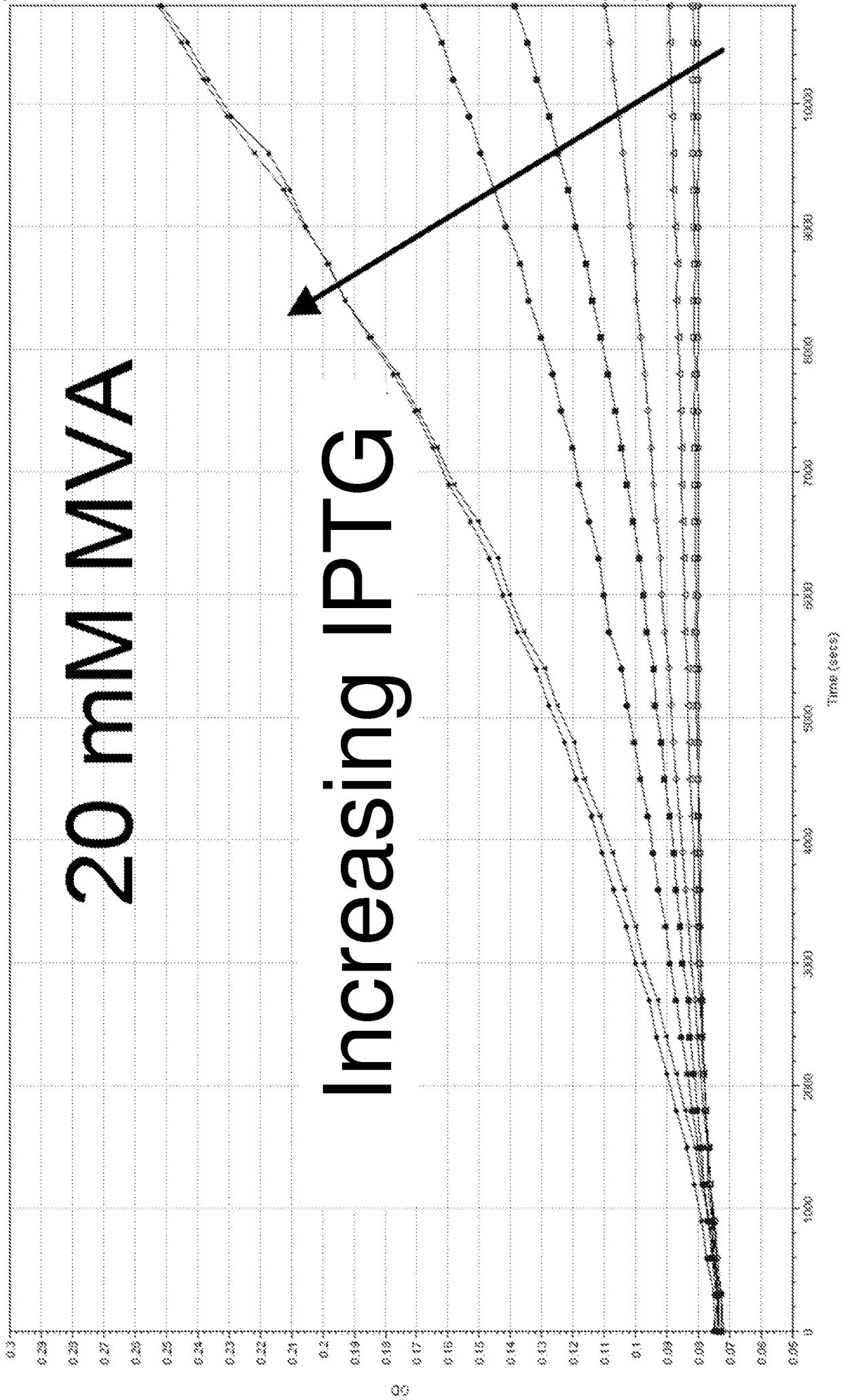


Figure 3

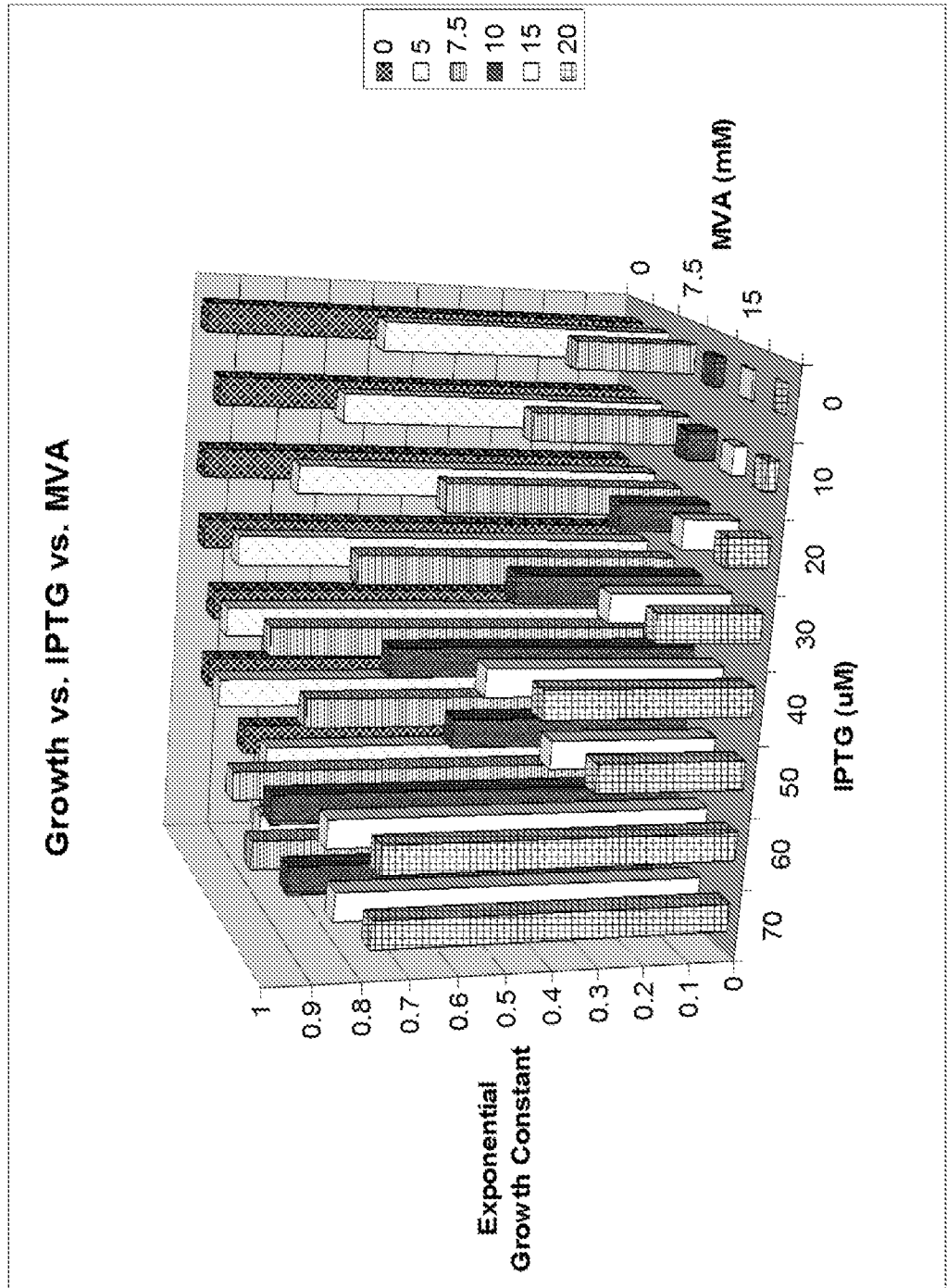
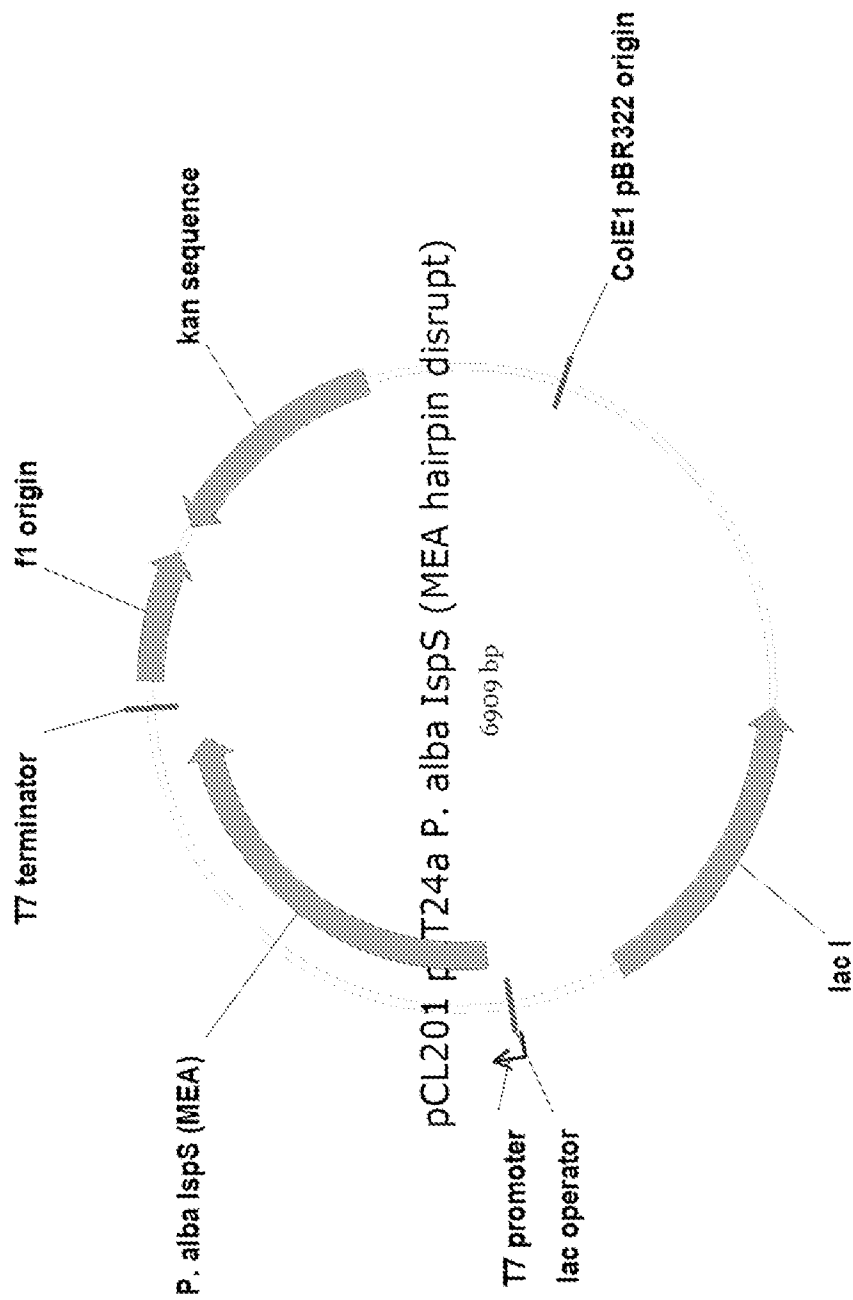


Figure 4



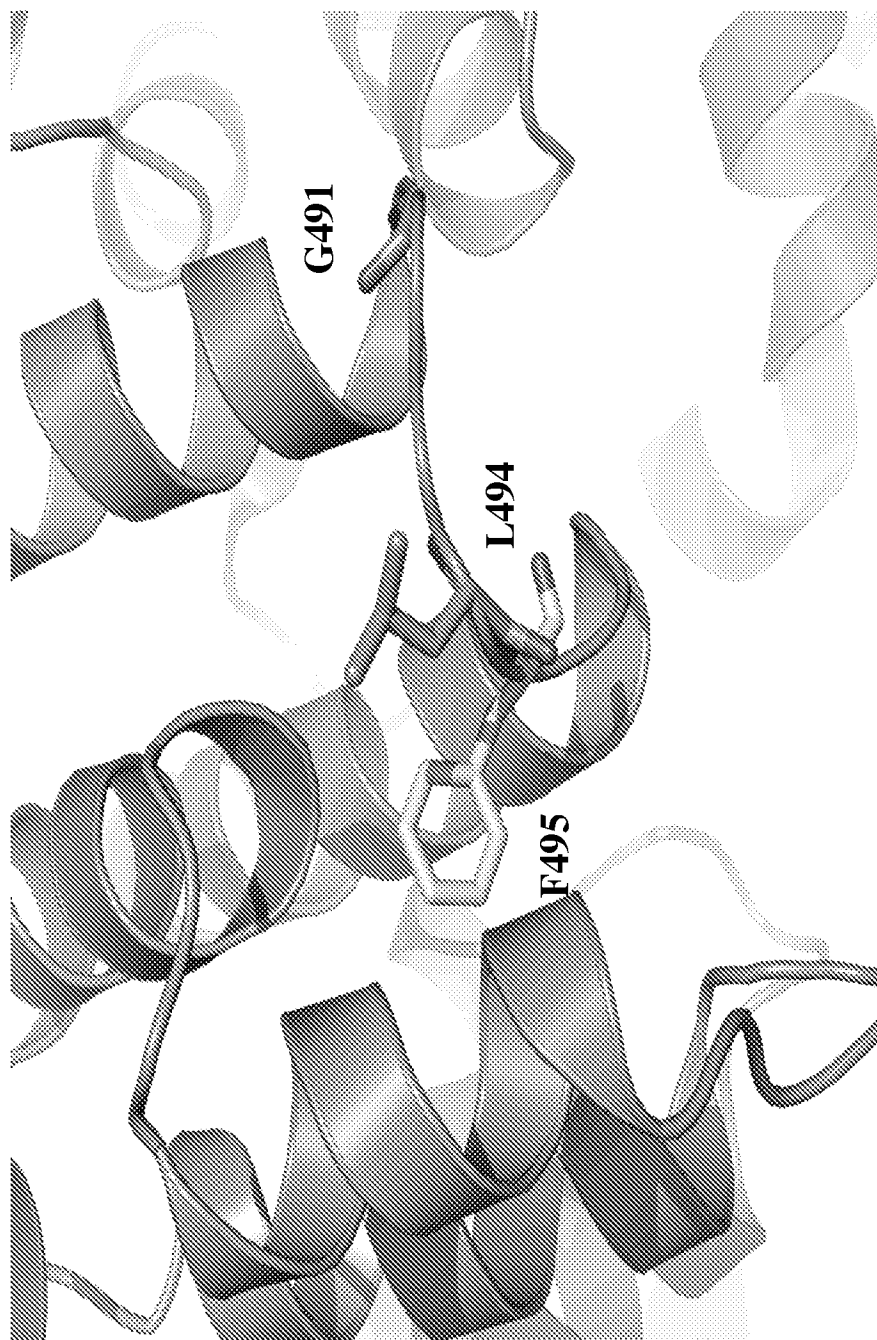


Figure 5

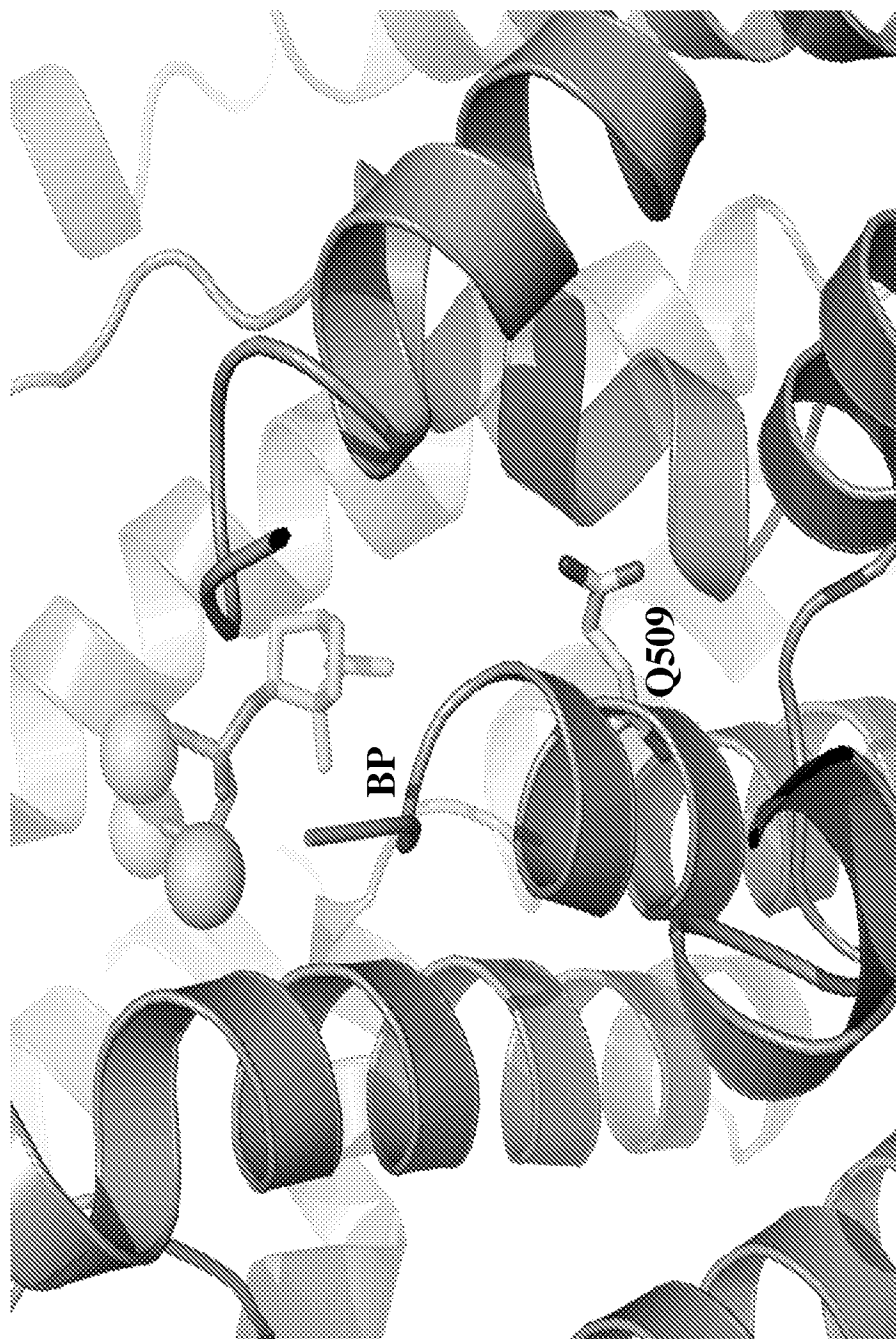


Figure 6

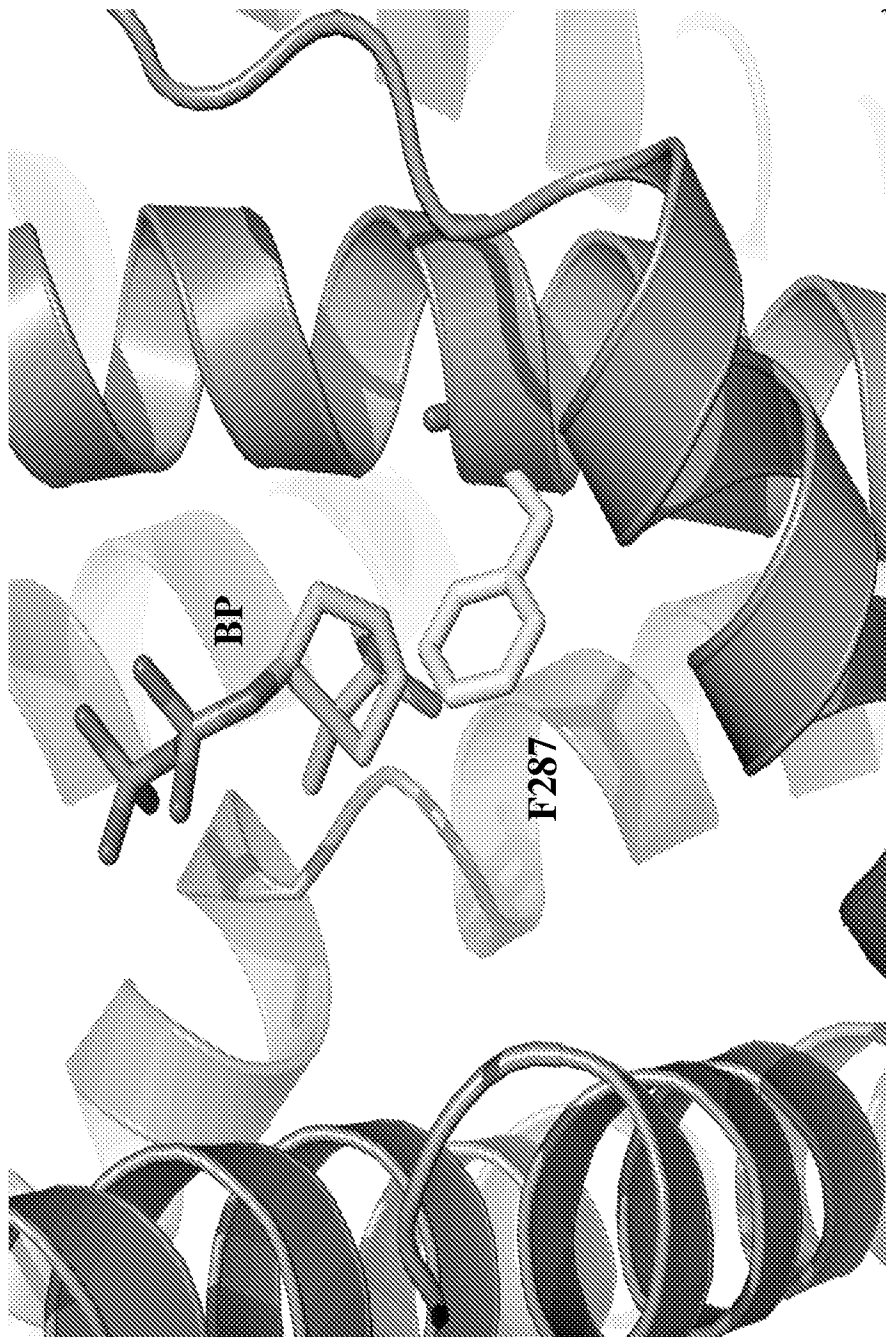


Figure 7

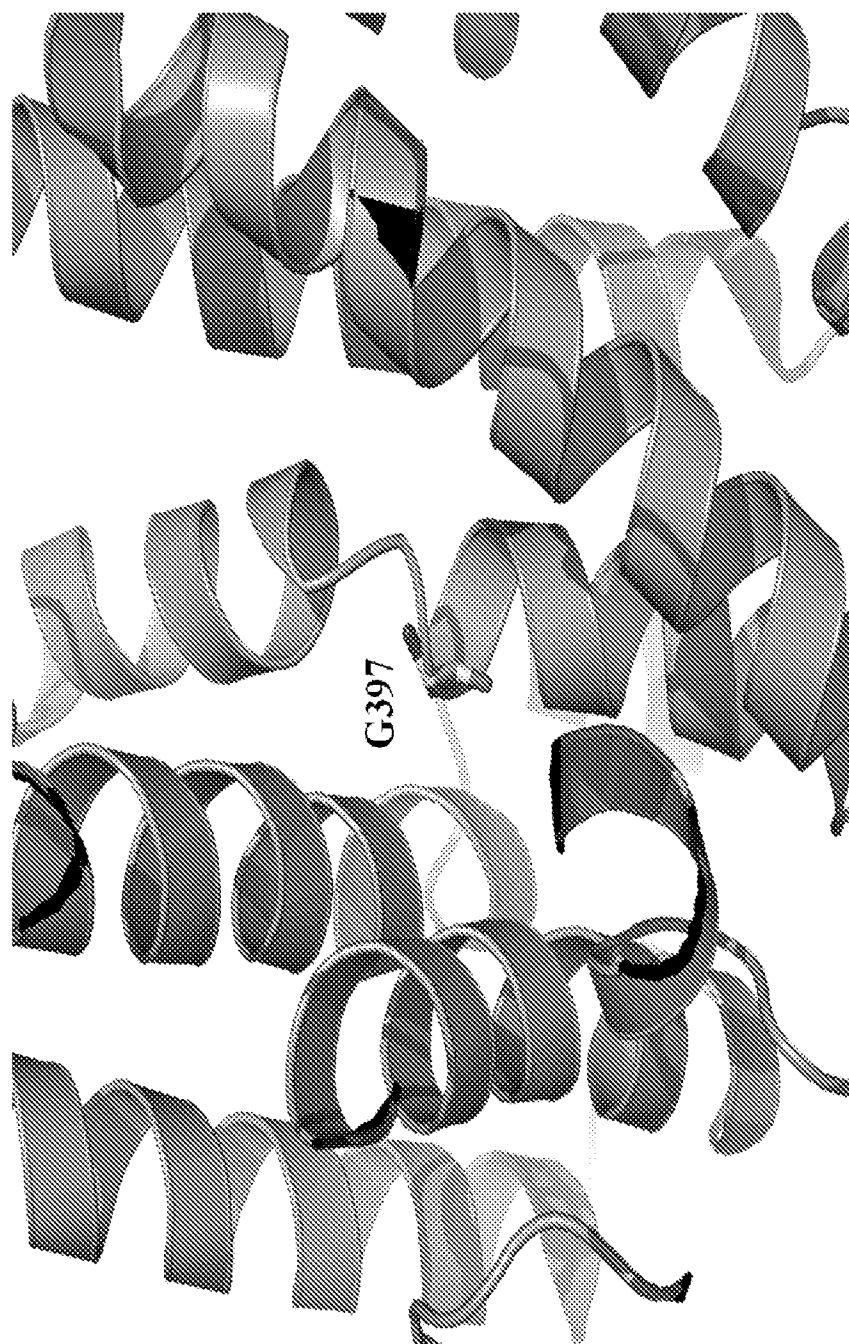


Figure 8

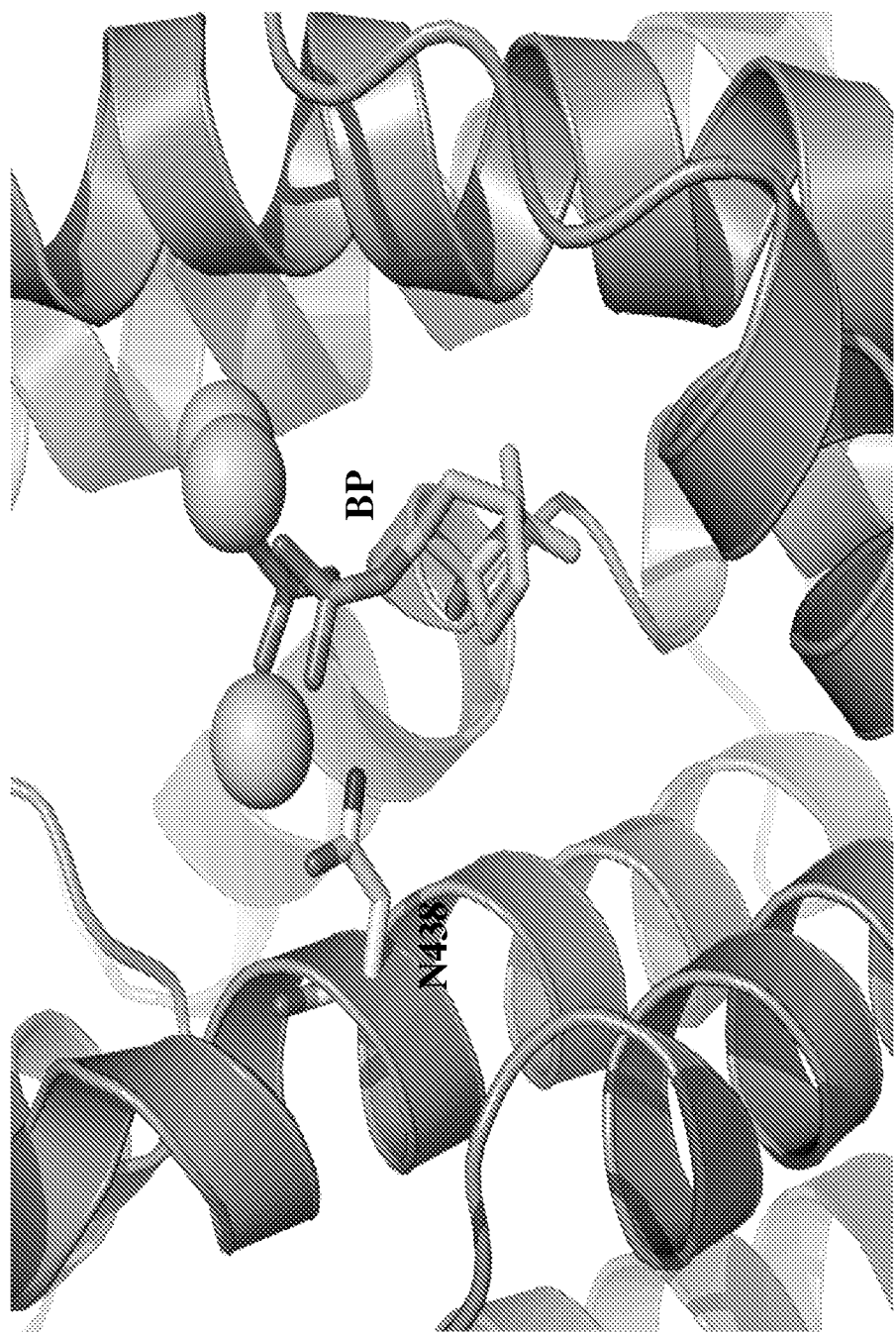


Figure 9

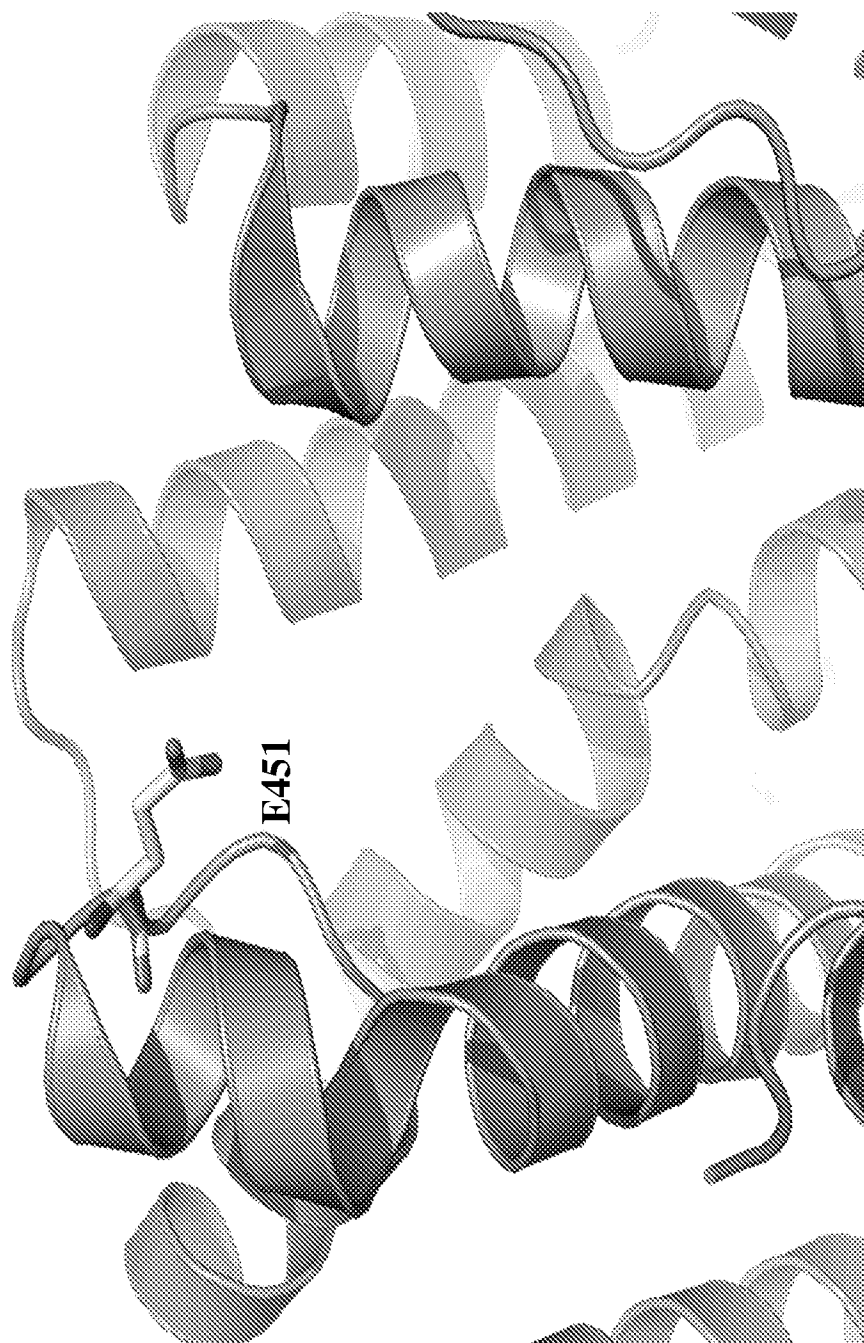


Figure 10

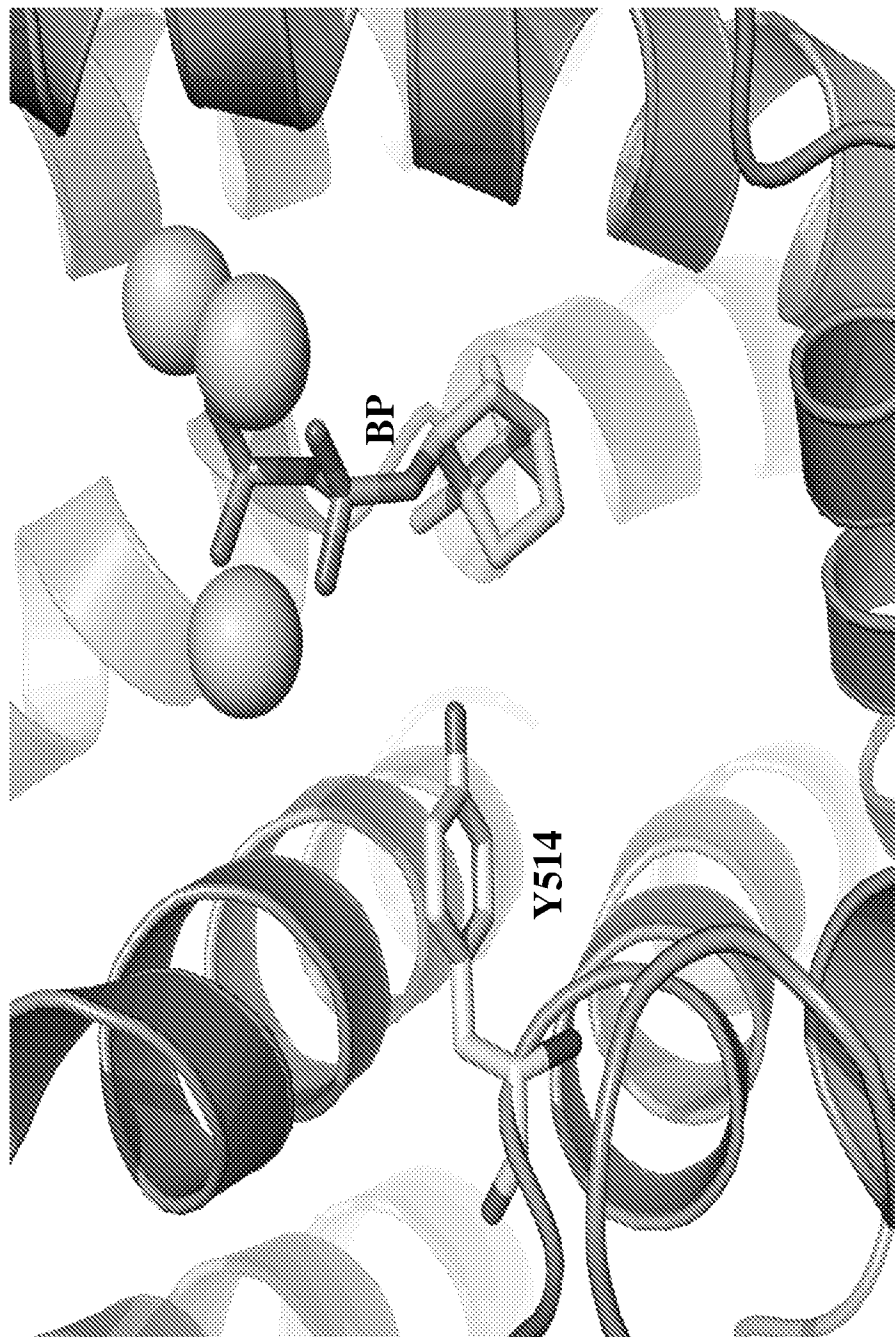


Figure 11

Figure 12

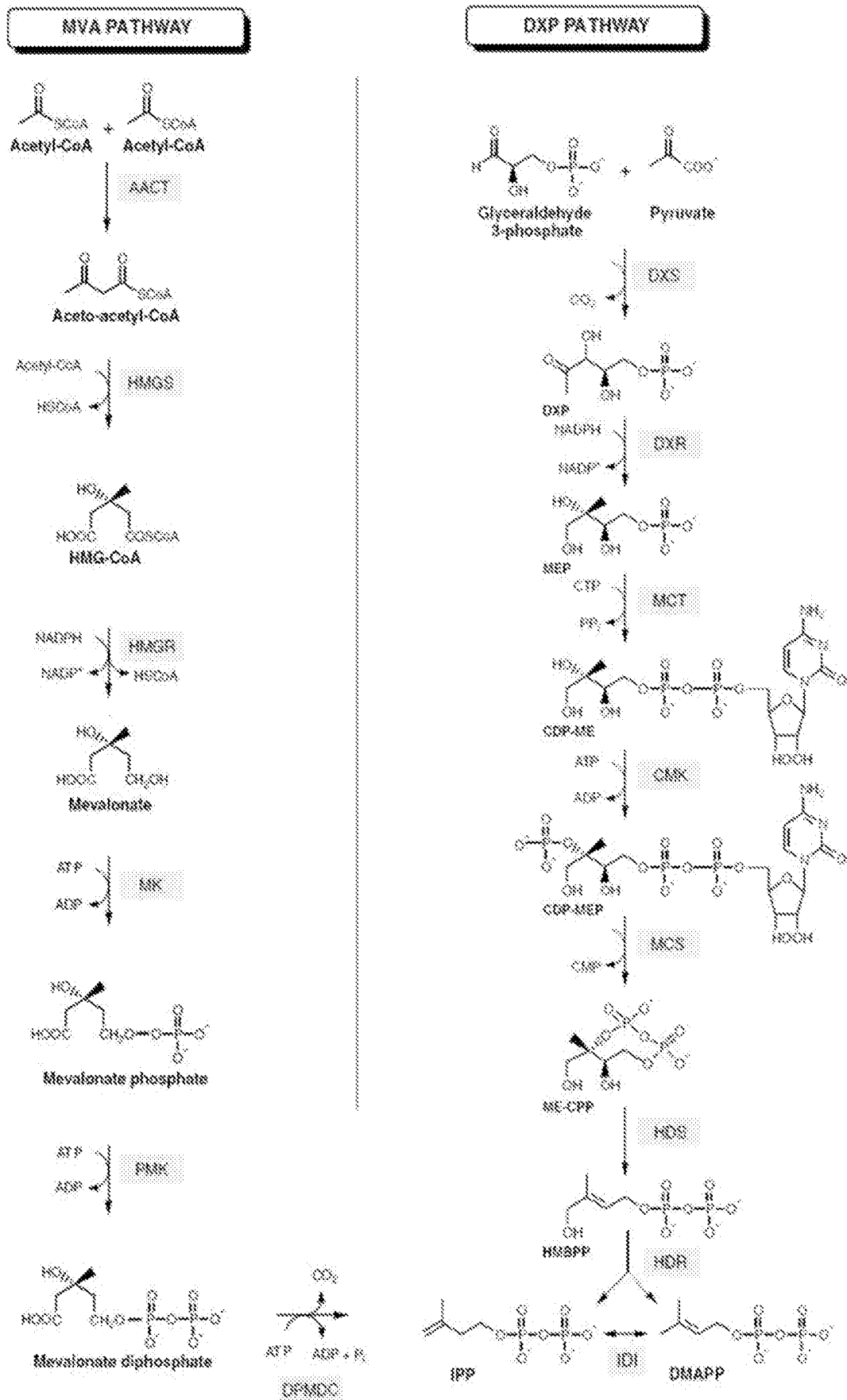




Figure 13

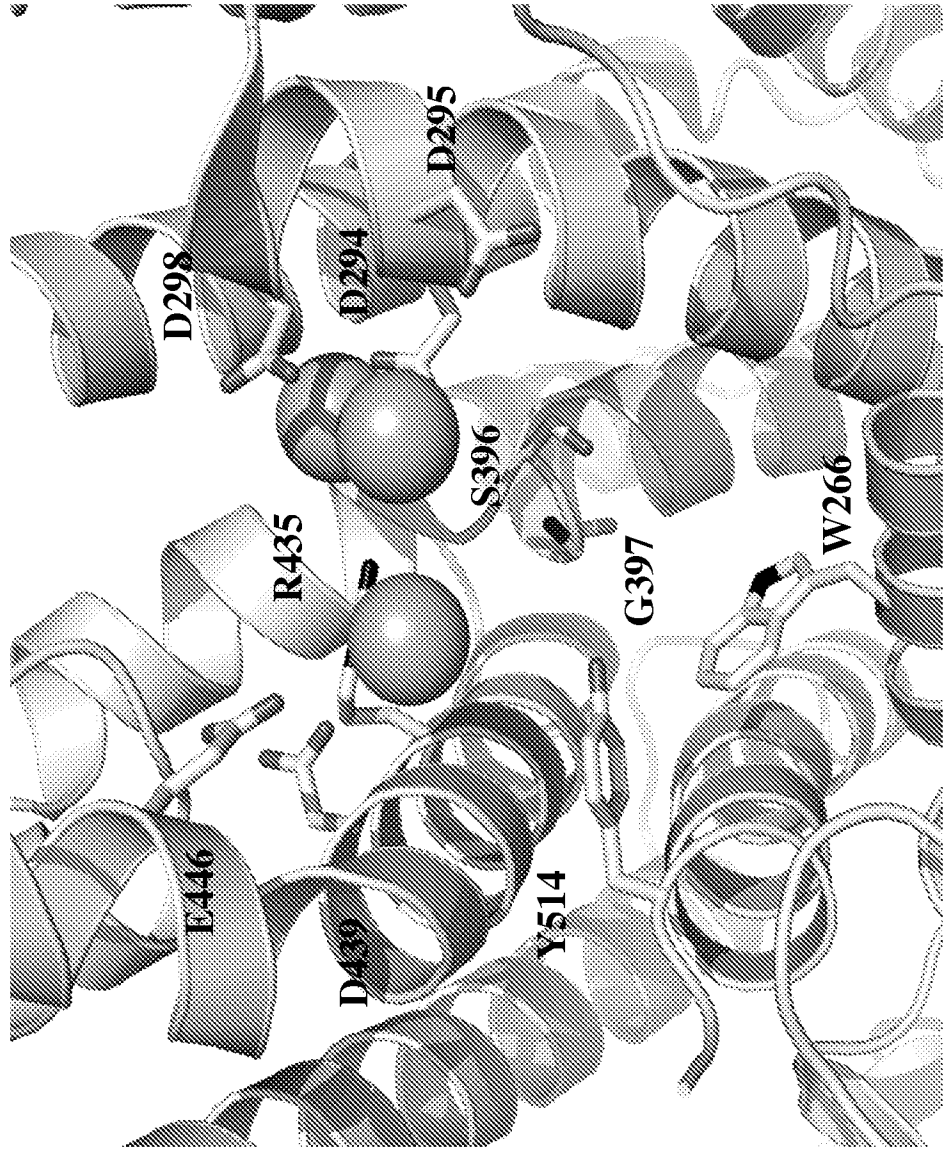


Figure 14

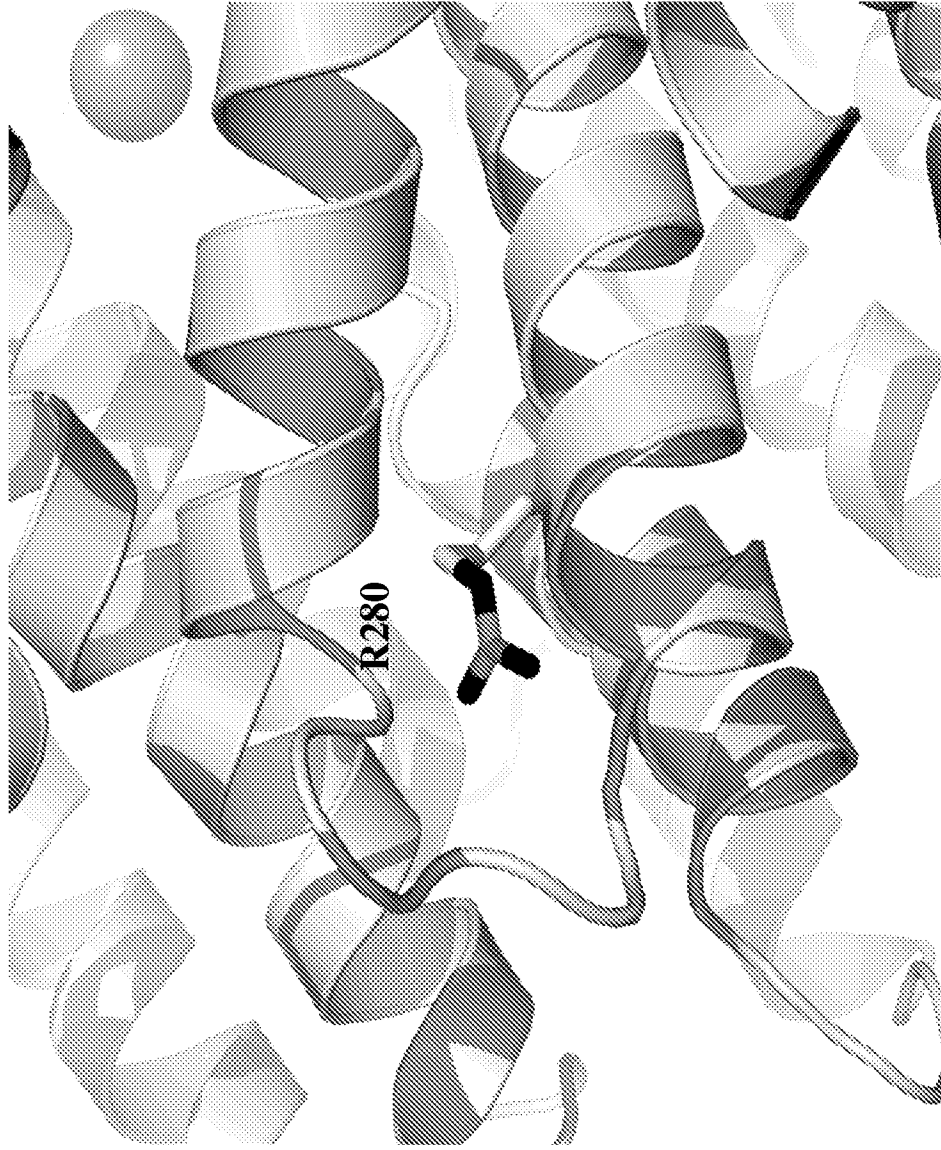


Figure 15

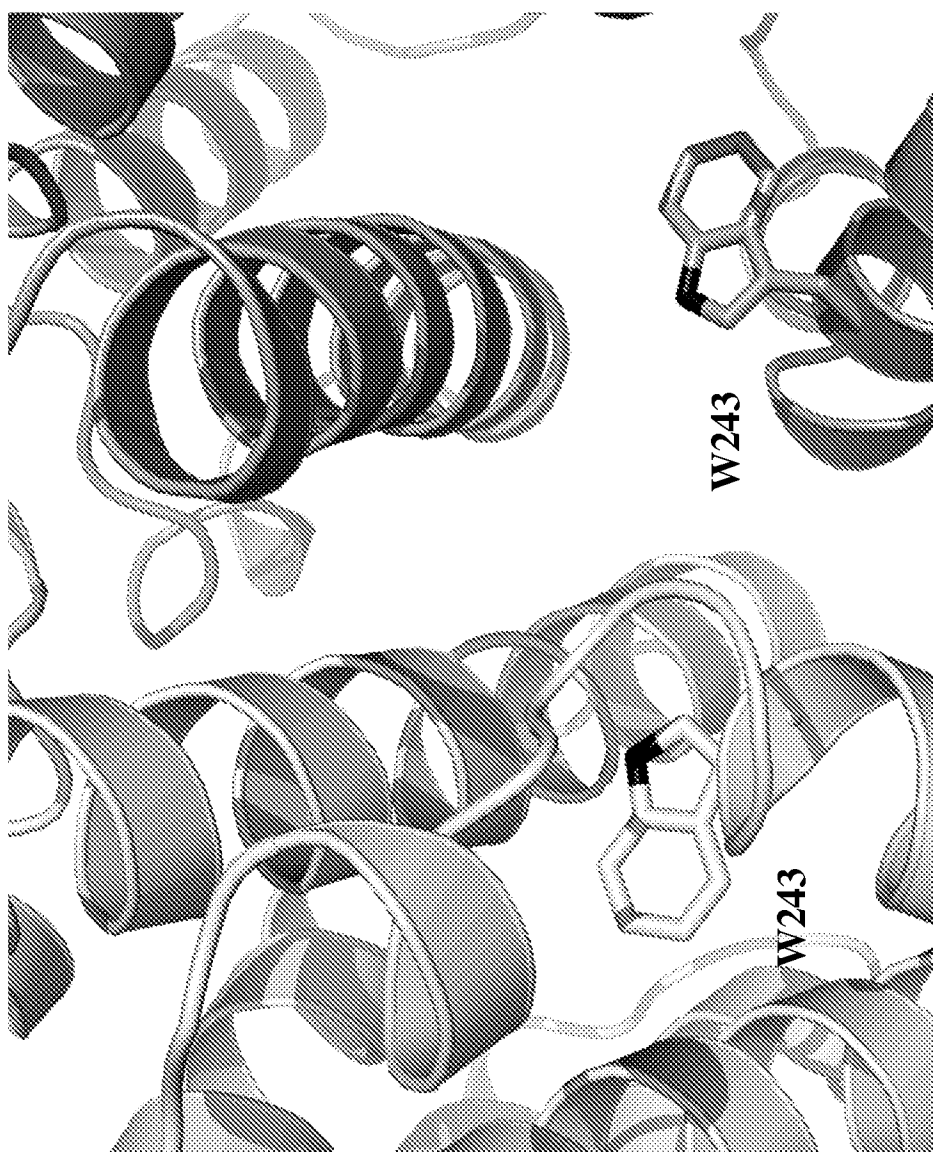


Figure 16

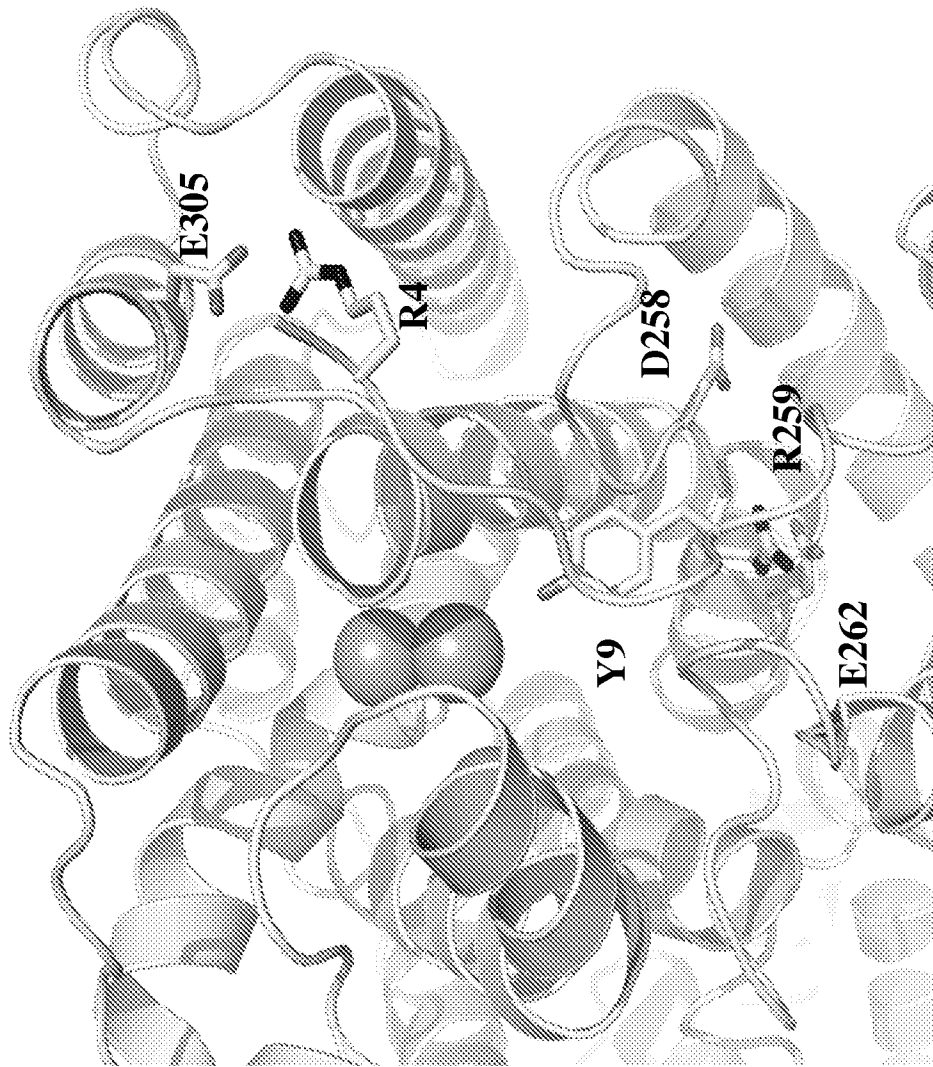


Figure 17

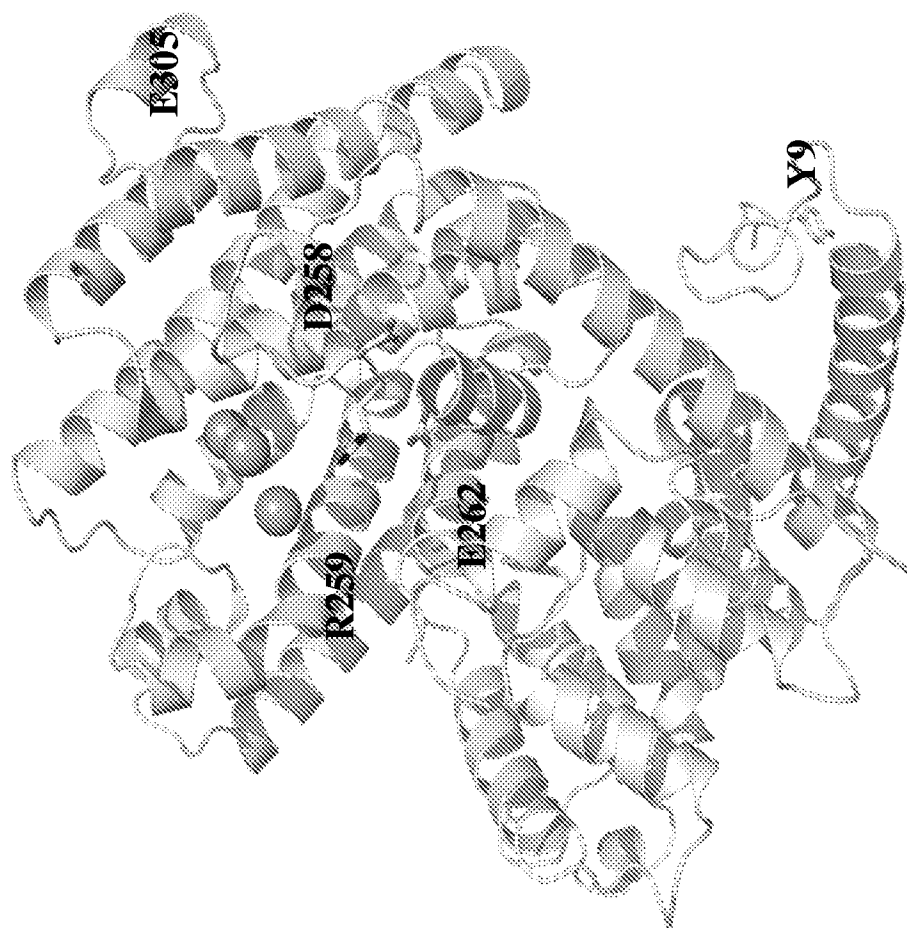


Figure 18

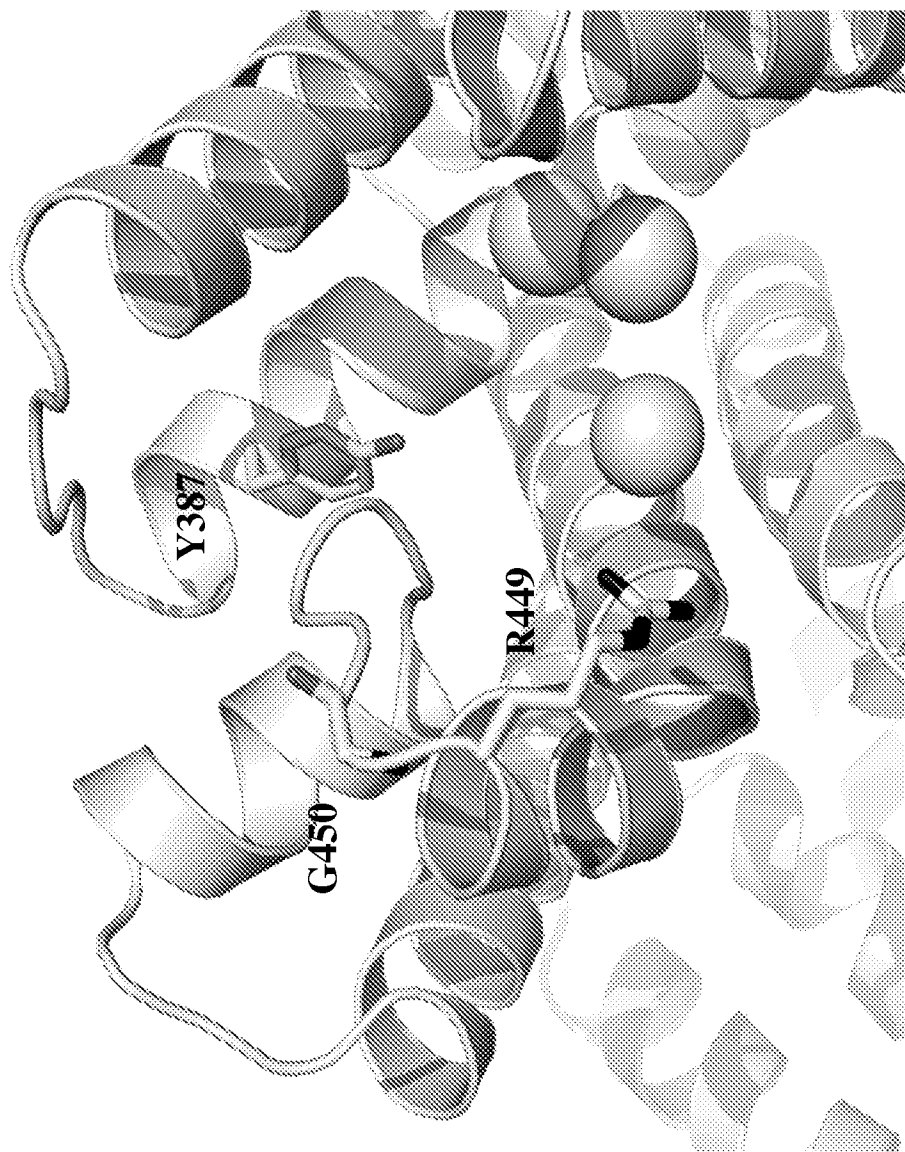


Figure 19



Figure 20

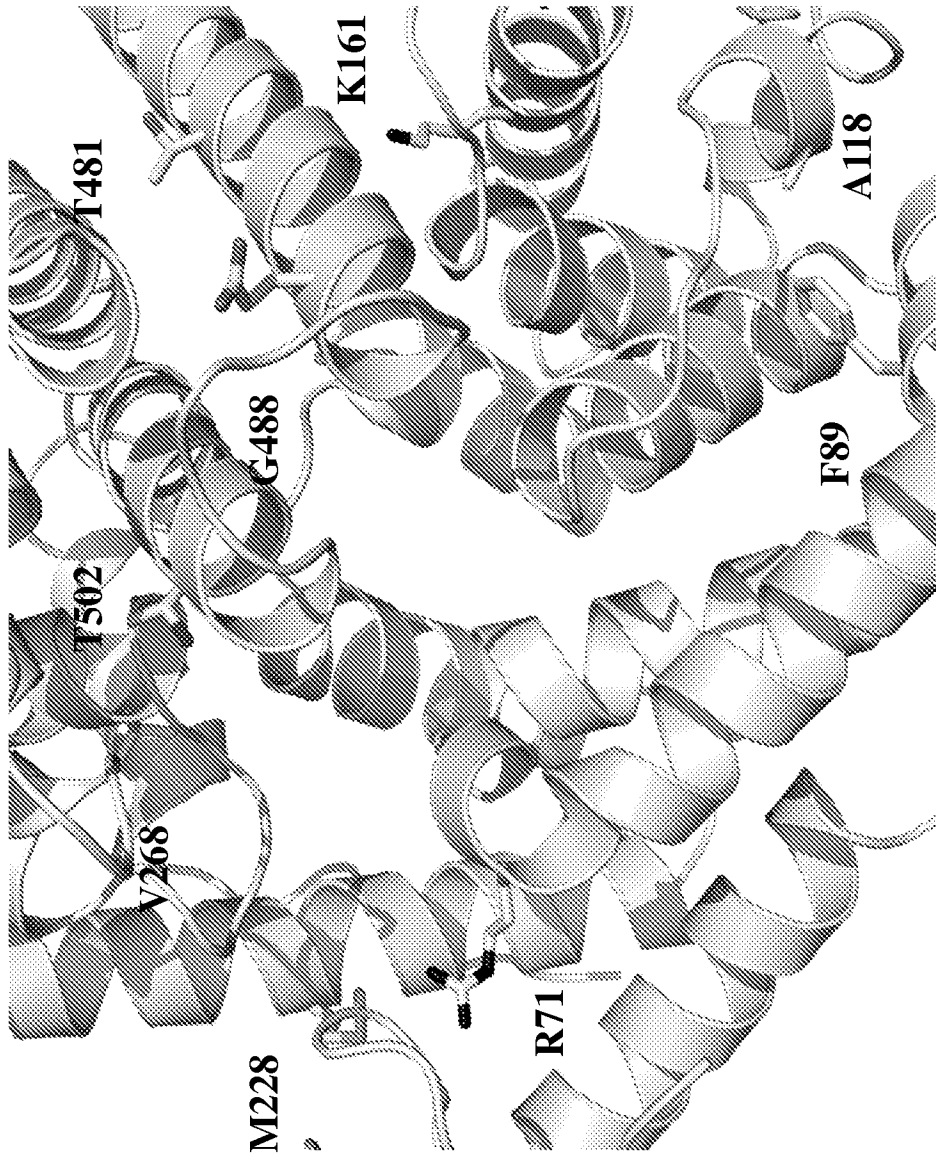


Figure 21

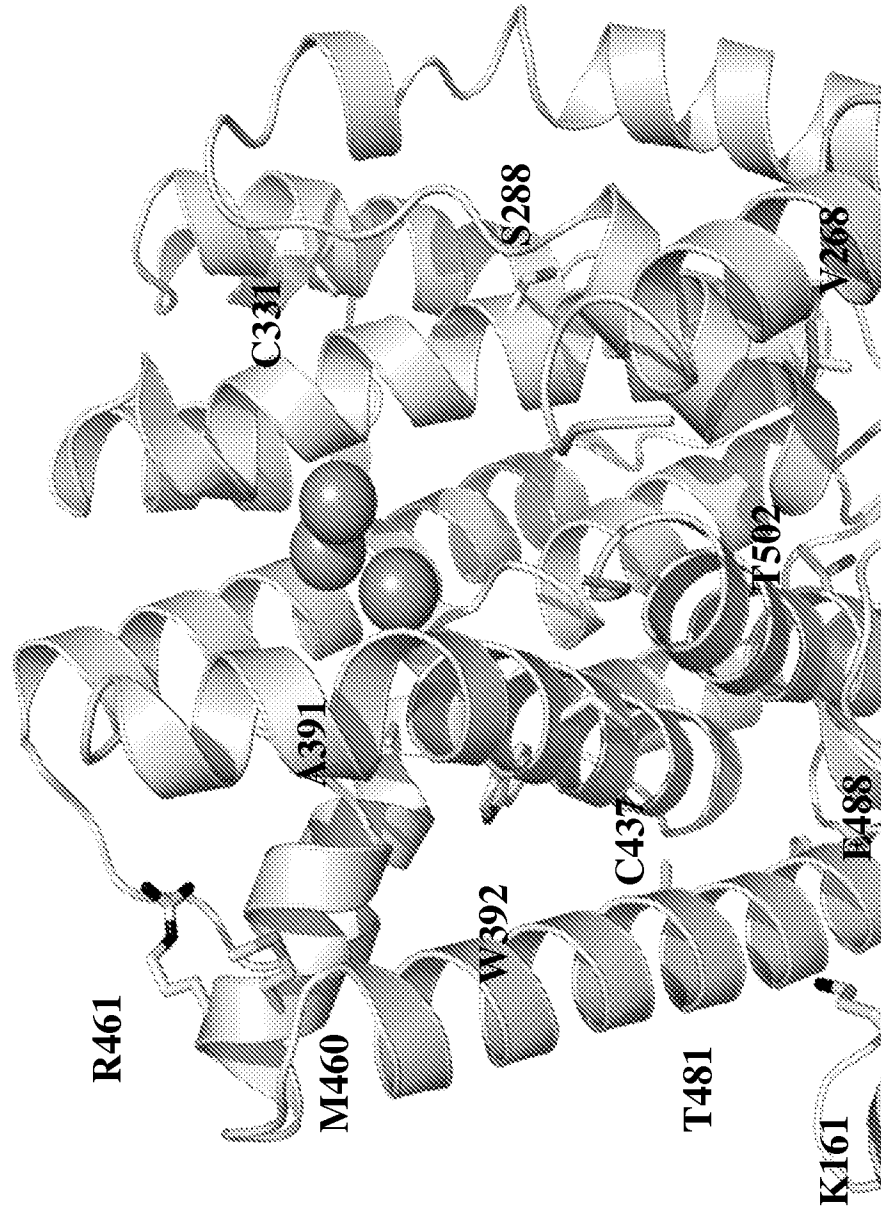


Figure 22

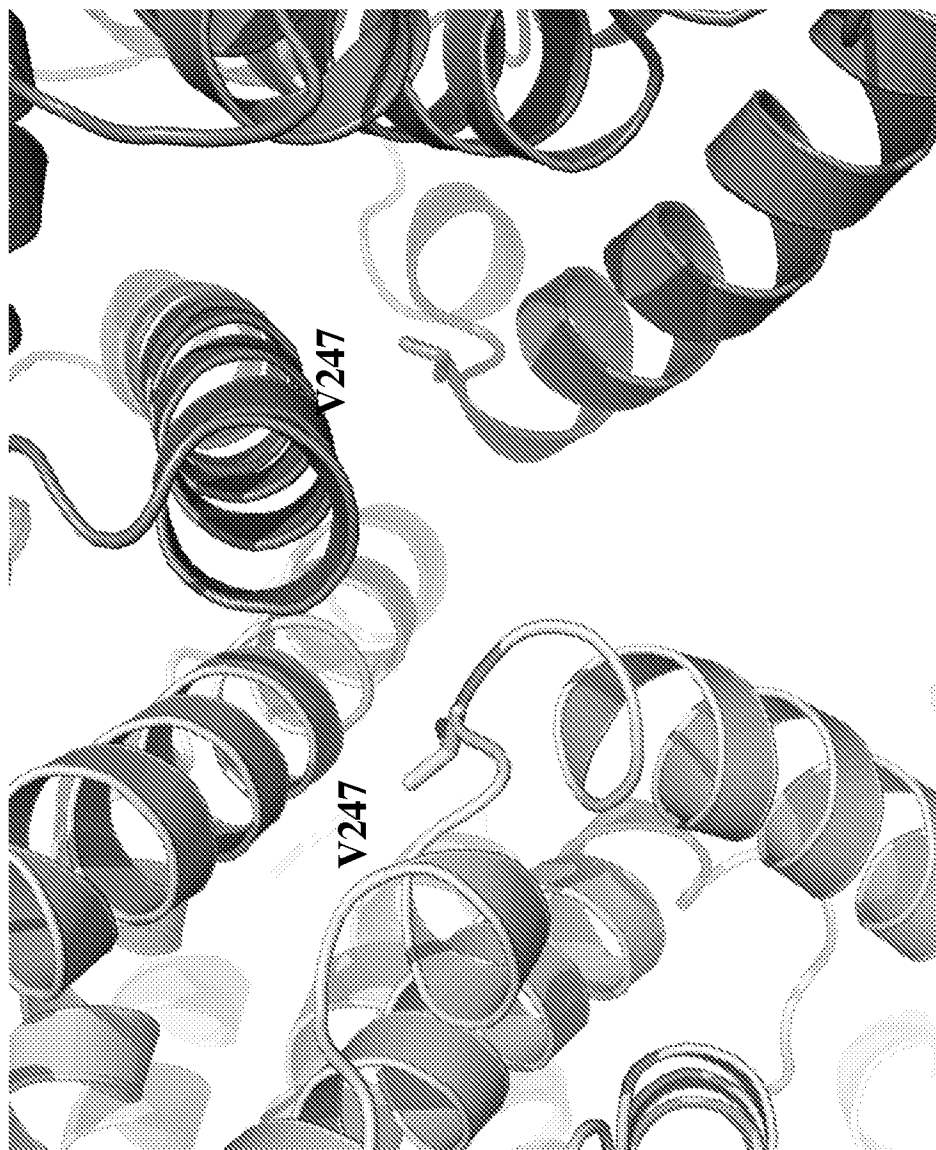


Figure 23

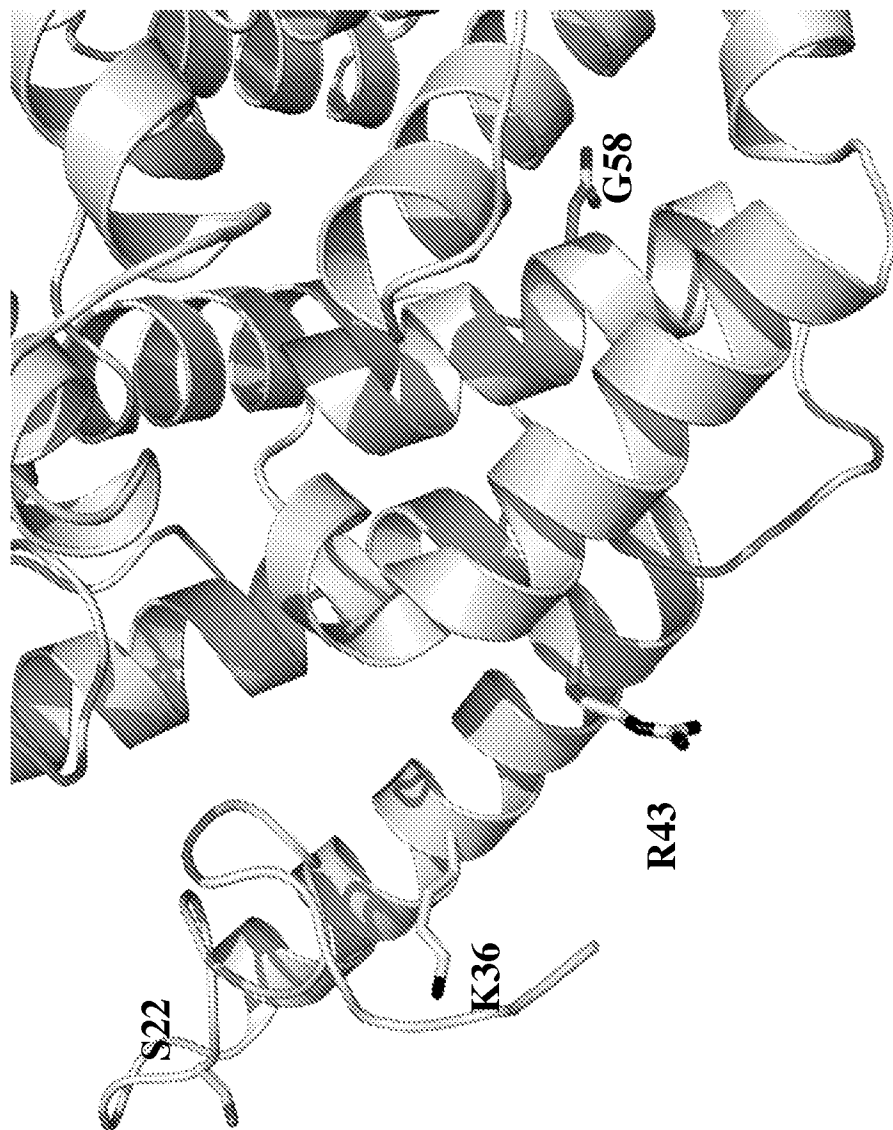


Figure 24

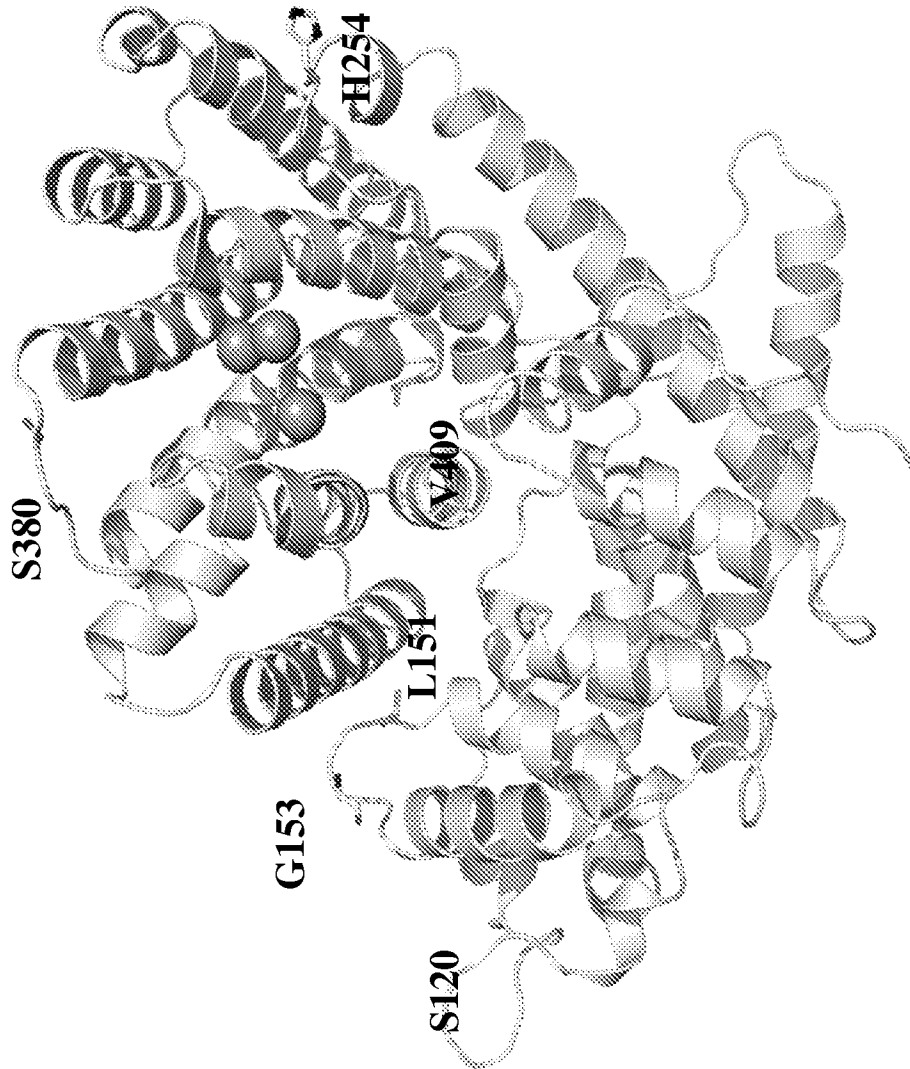


Figure 25

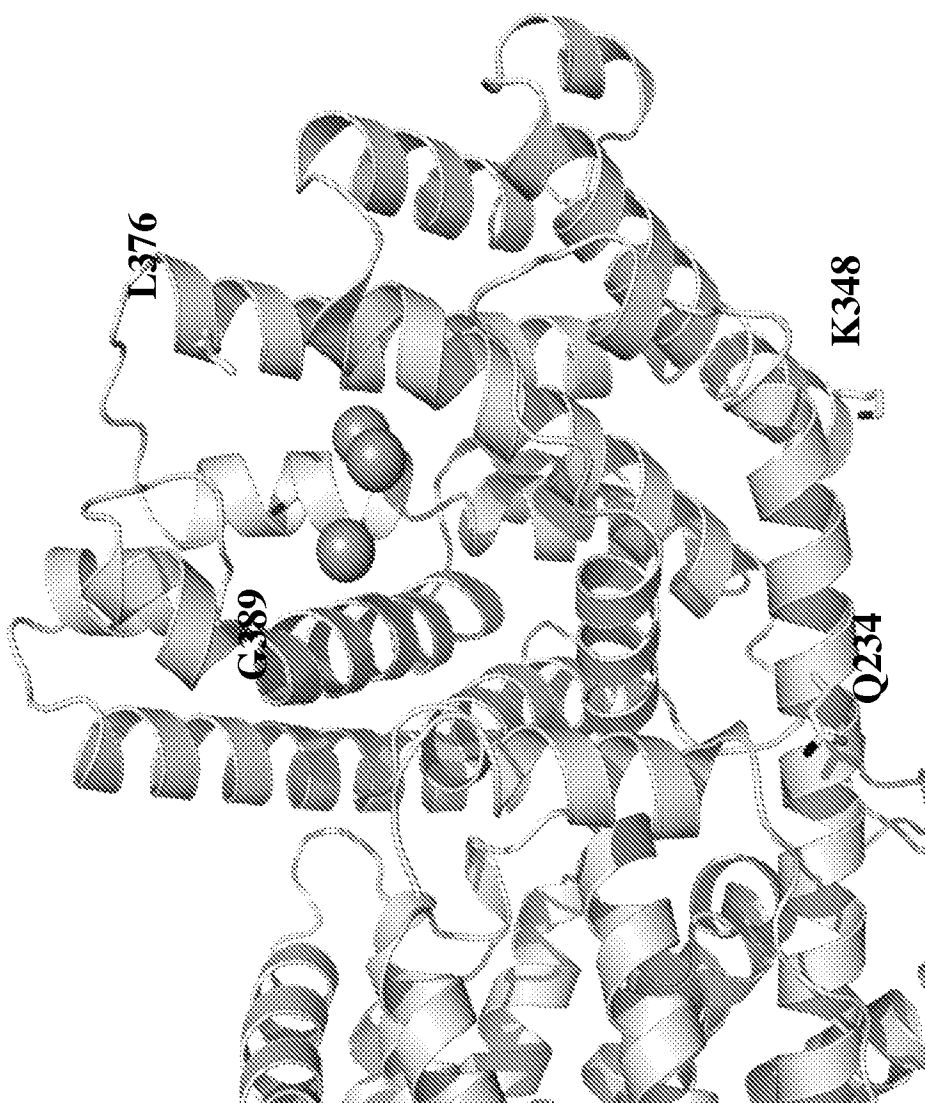


Figure 26

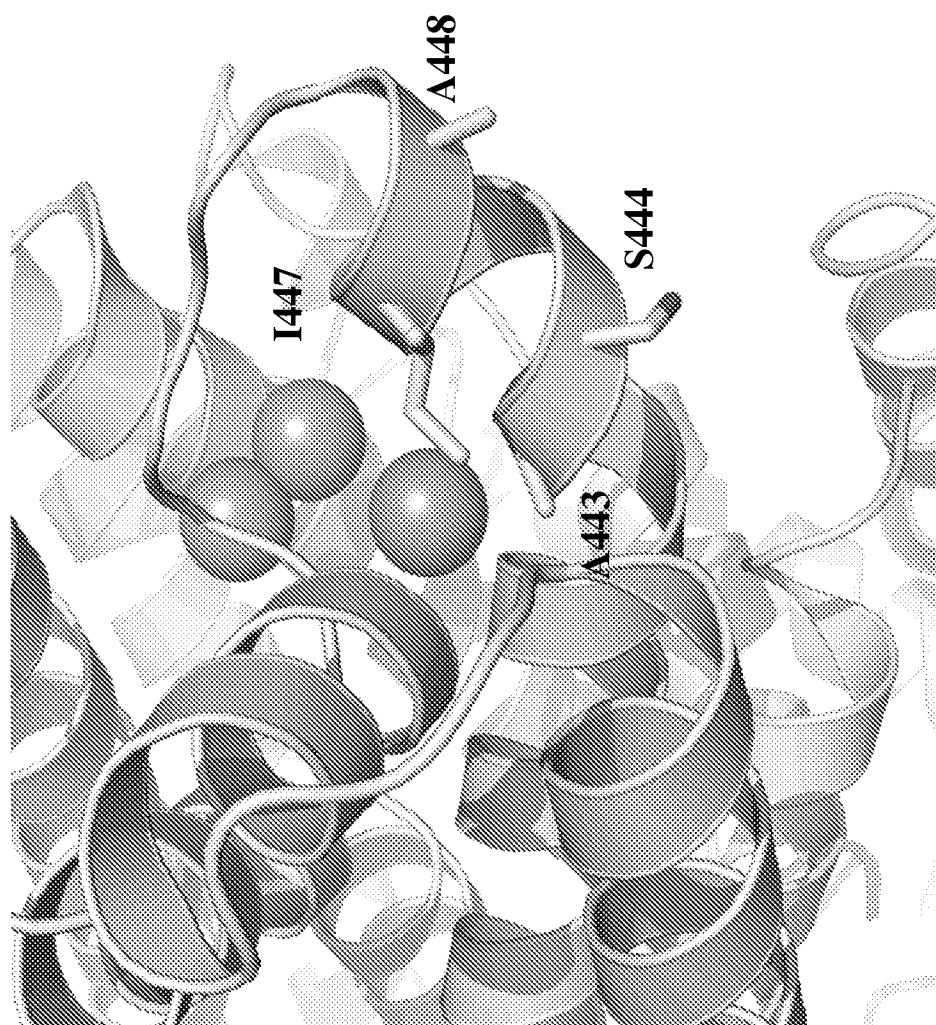


Figure 27

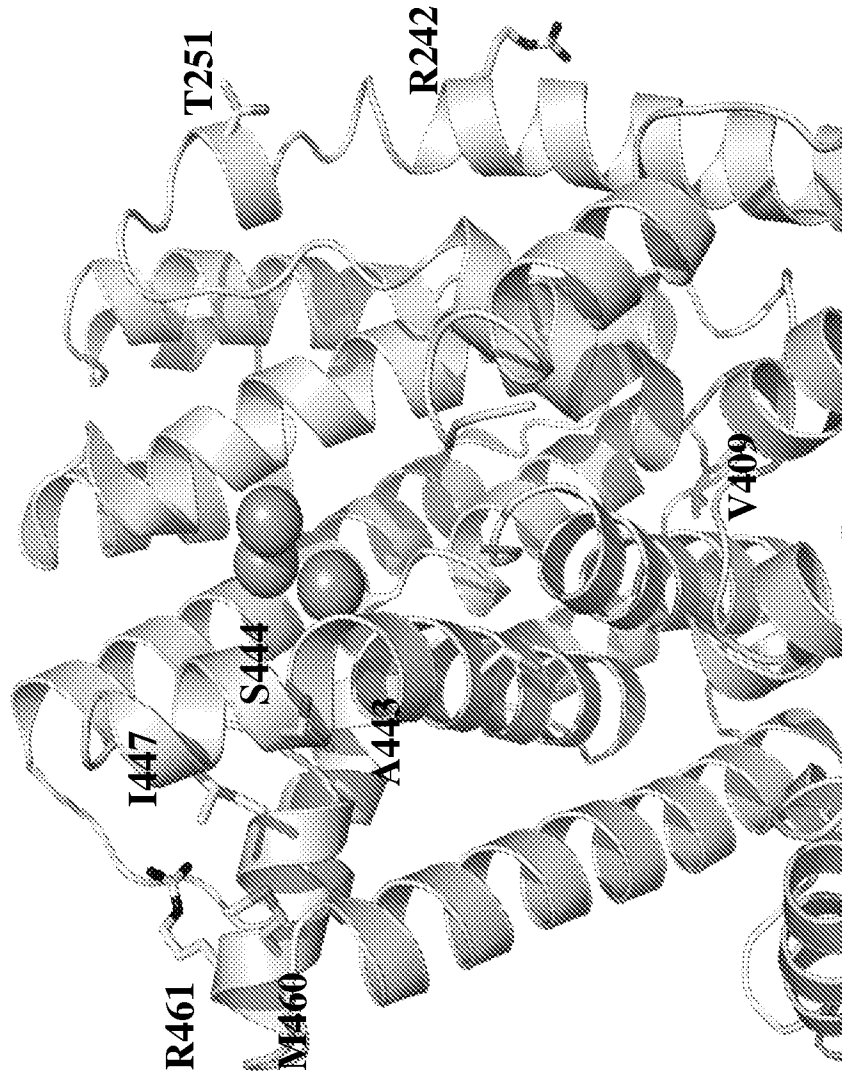


Figure 28

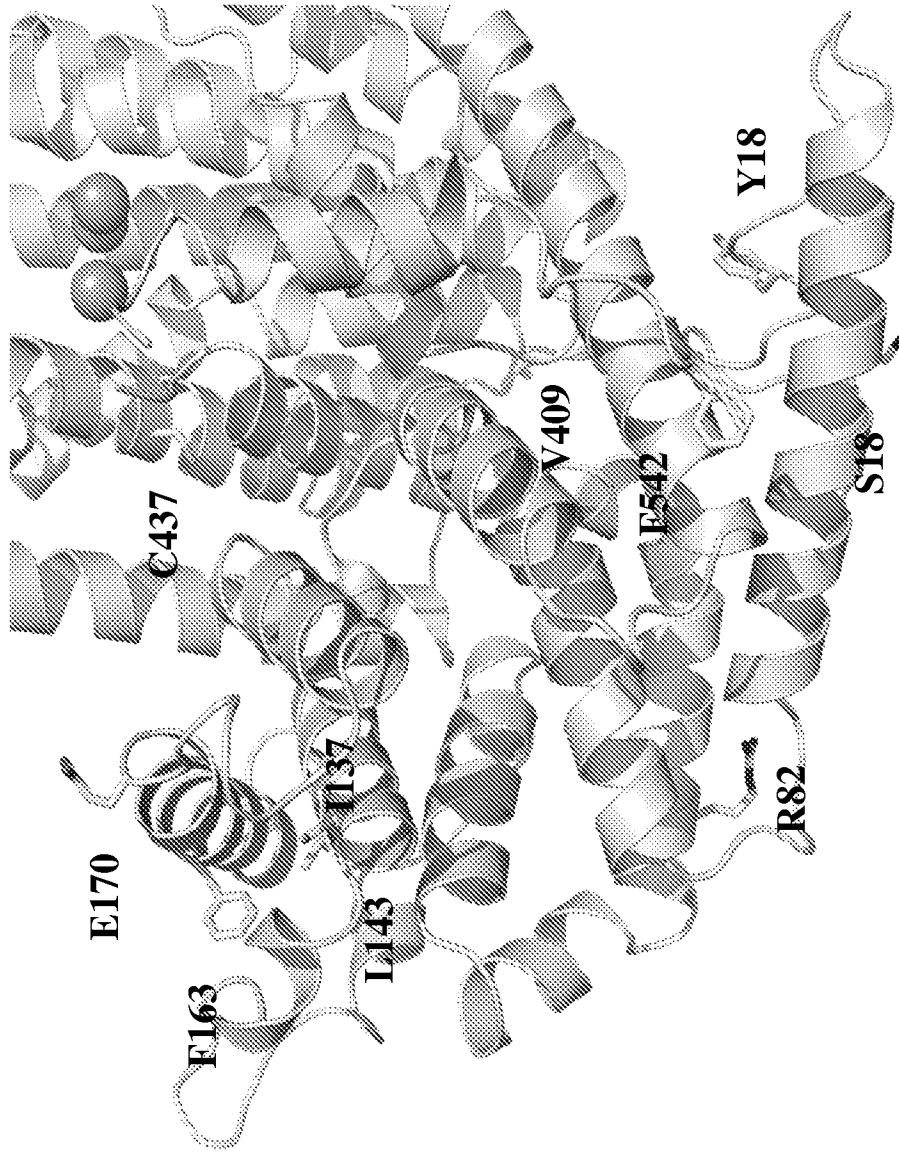


Figure 29

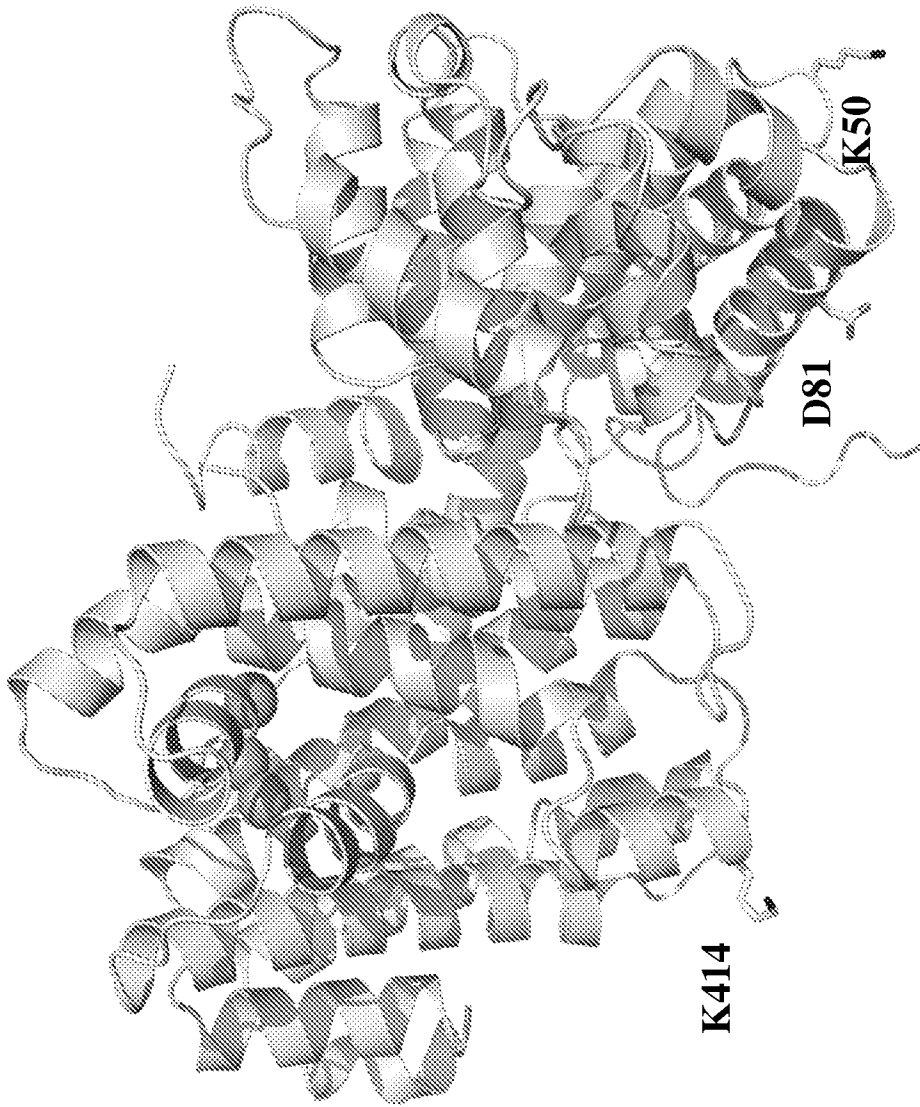


Figure 30

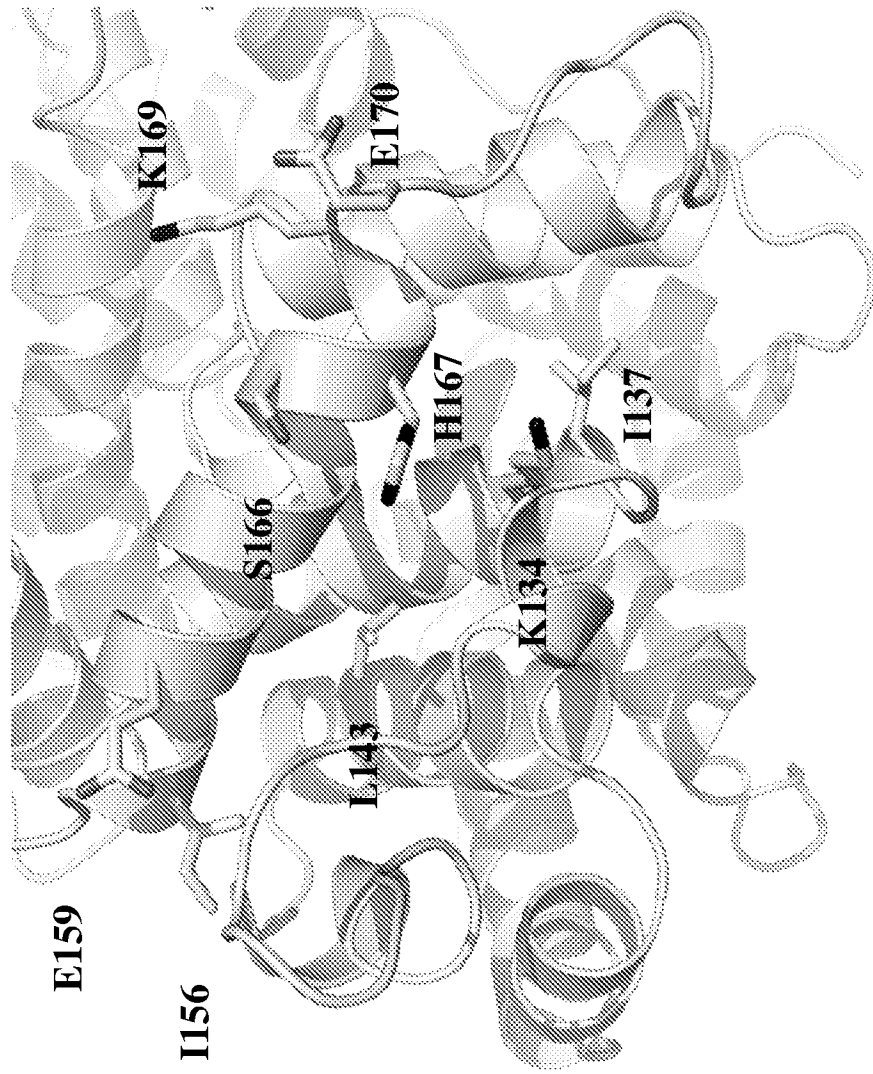


Figure 31

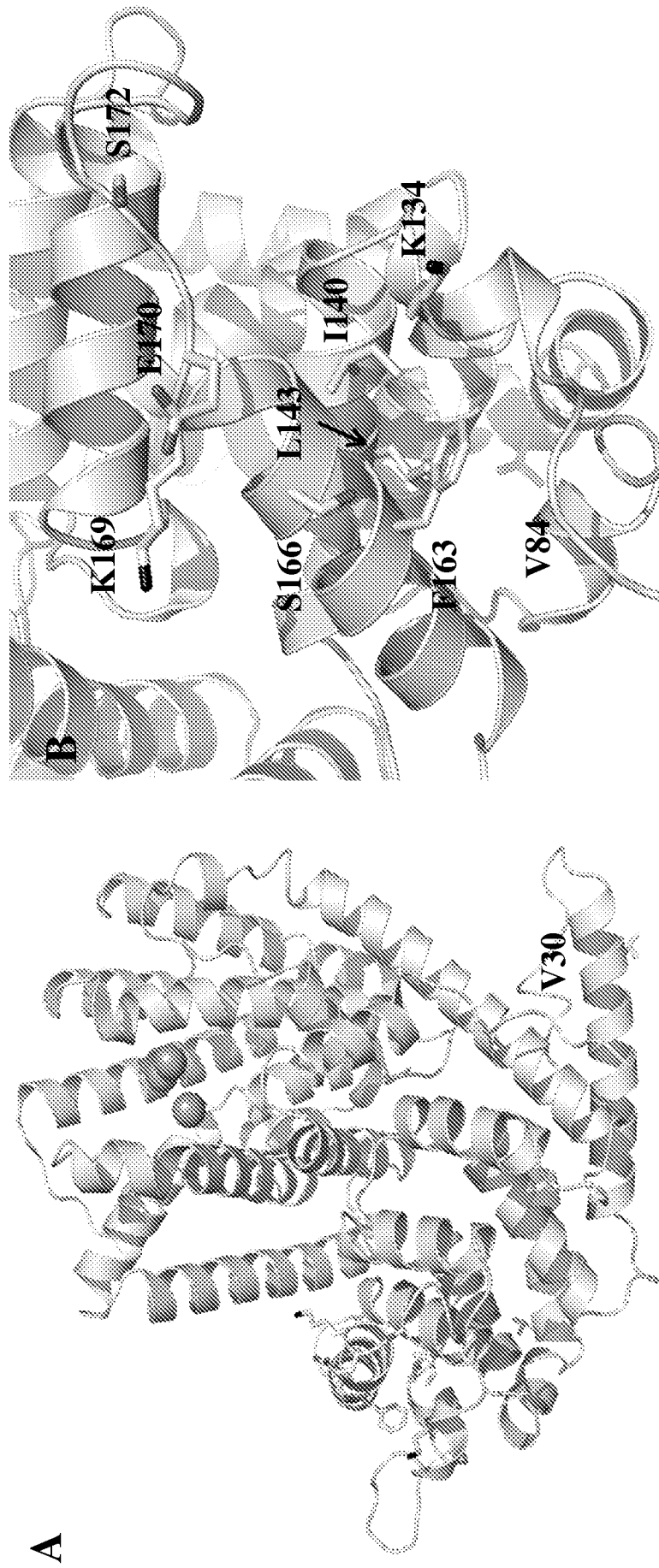


Figure 32

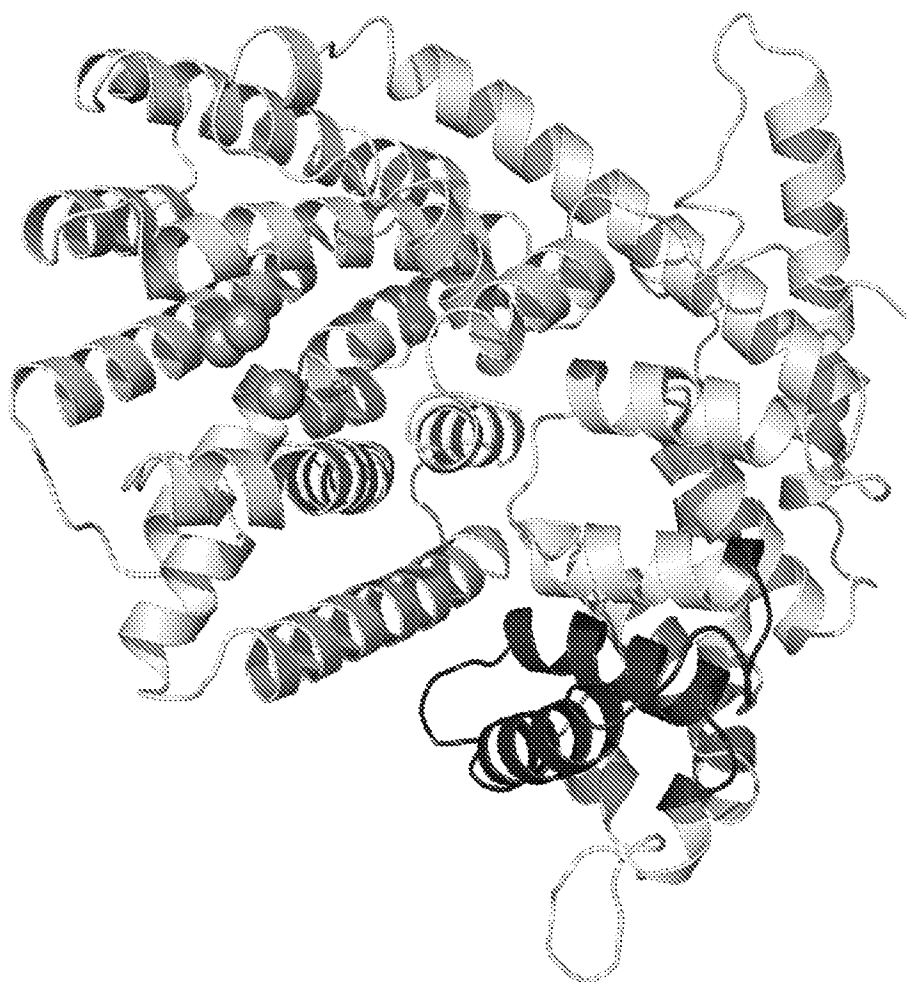


Figure 33

Figure 34

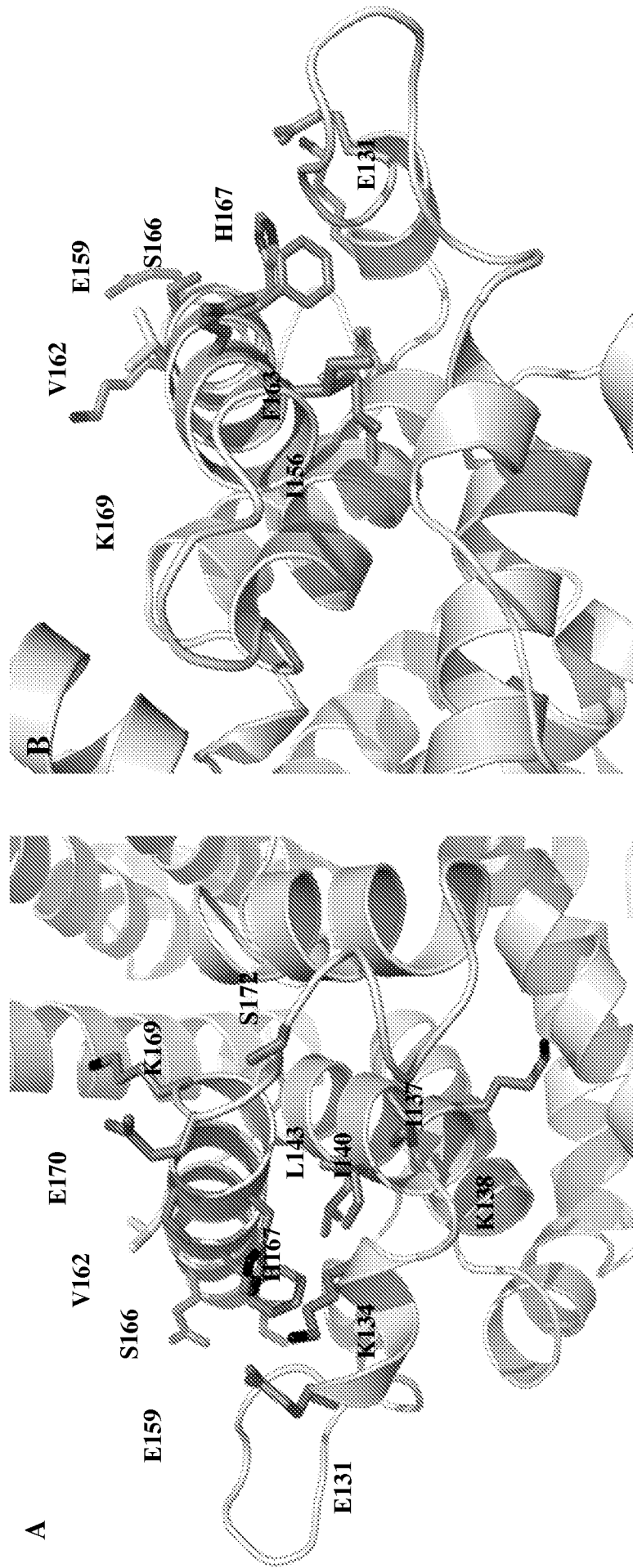


Figure 36

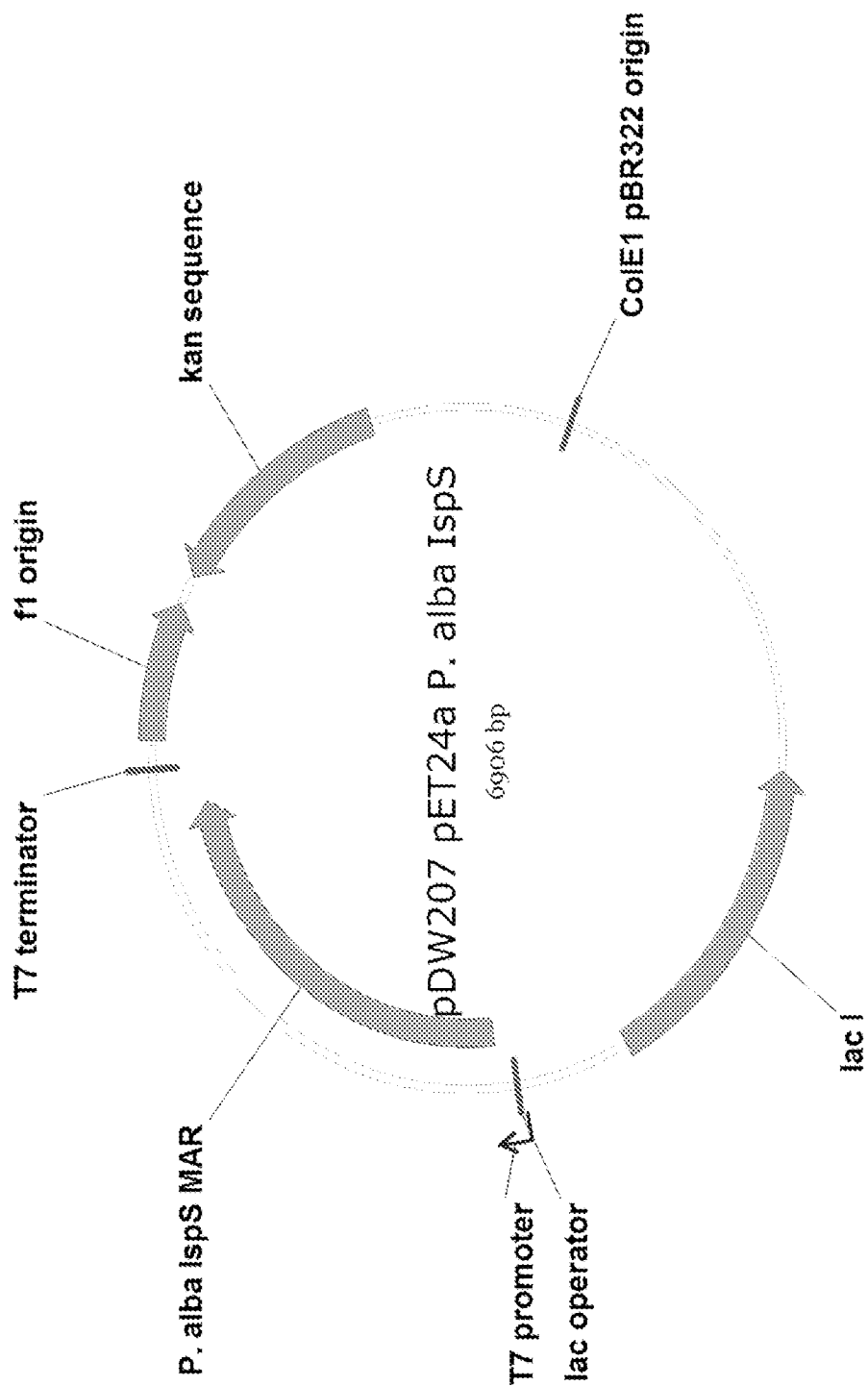


Figure 37

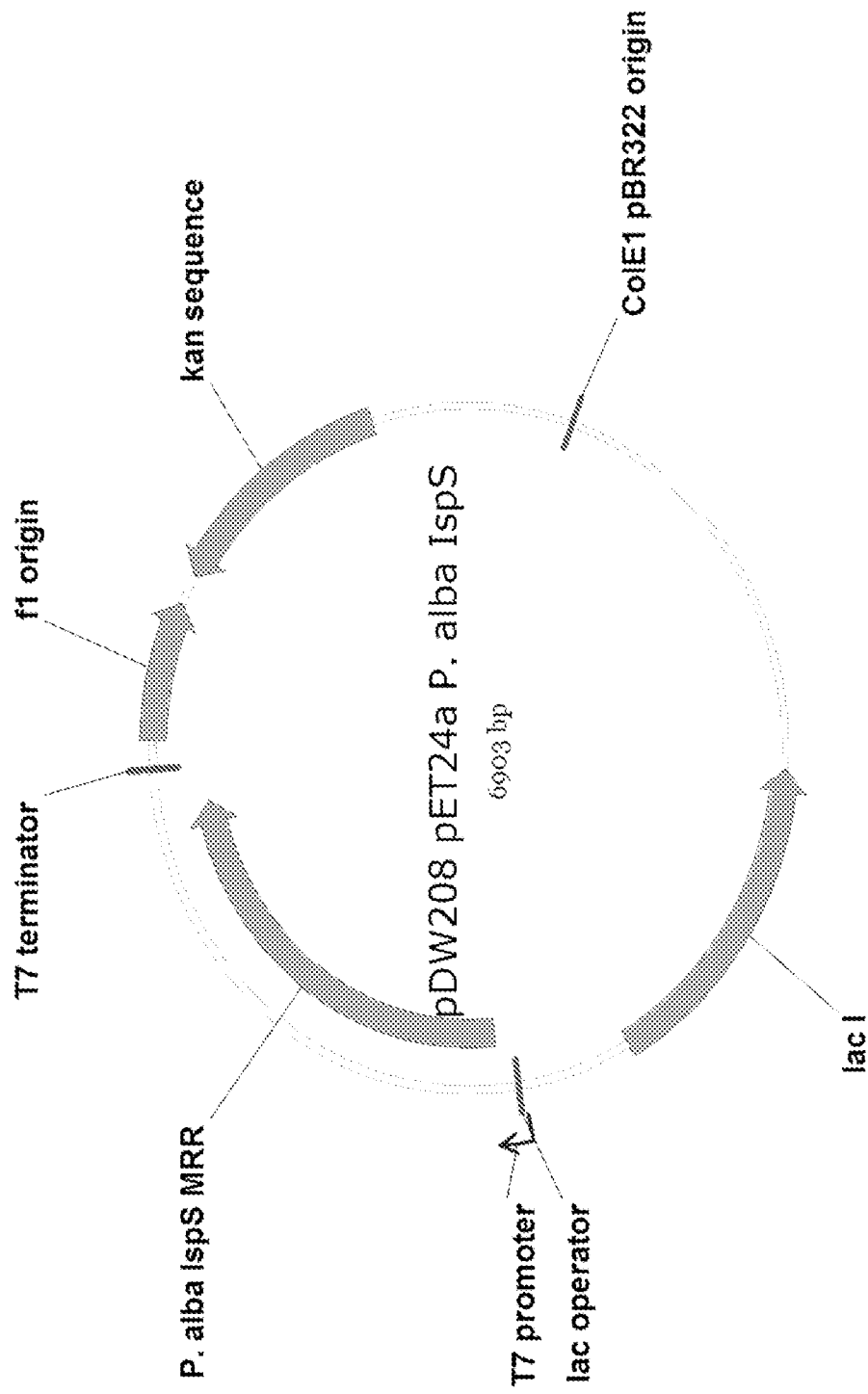


Figure 38A

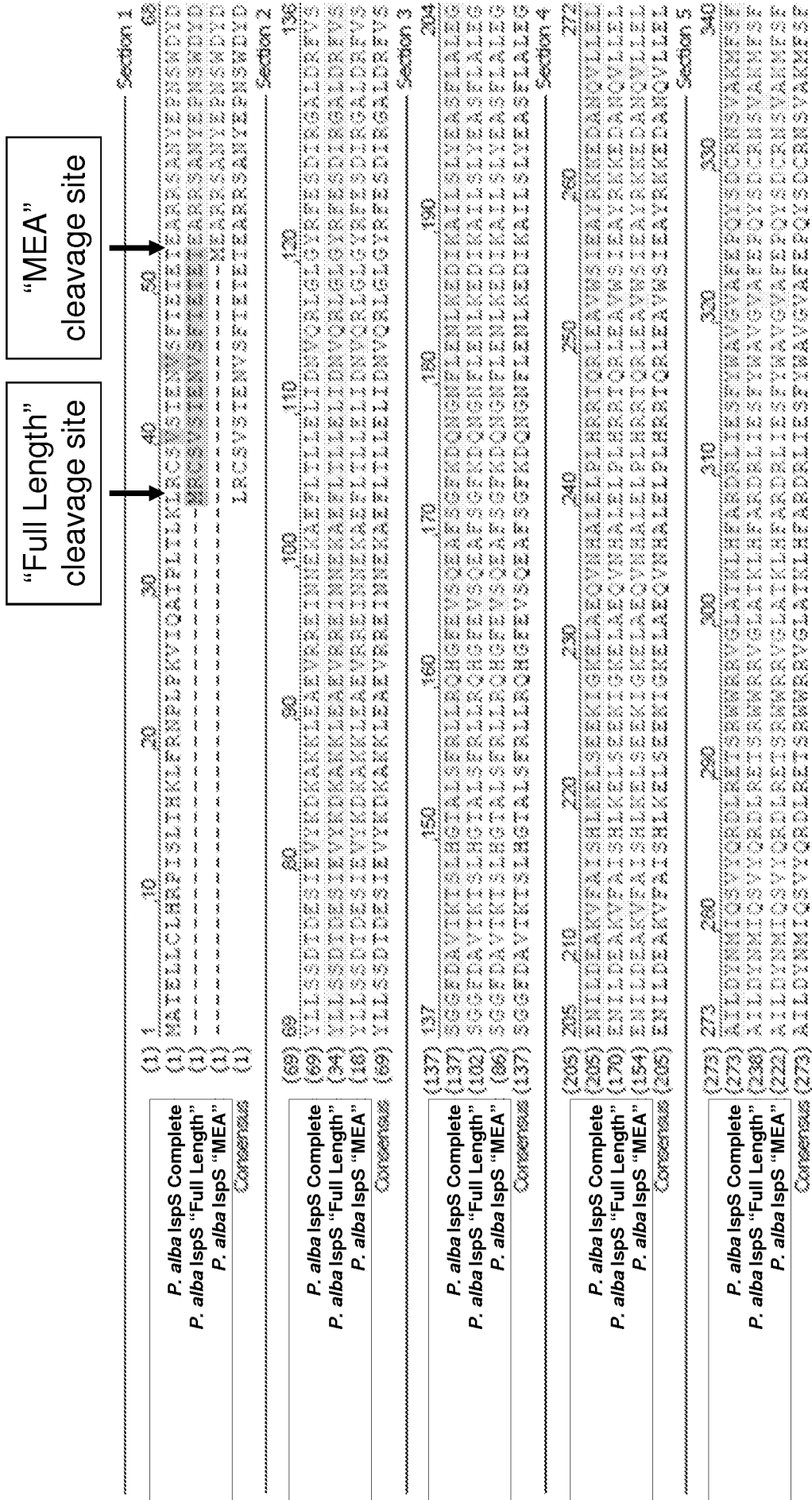


Figure 38B

	350	360	370	380	390	Section 6
<i>P. alba</i> lspS Complete	VIIDDIYDVYCGILDELELEFD	DAVERNDVNAINDLFDYMKL	CFALALYNTINEIAYDNL	KDKGENILFY	400	Section 6
<i>P. alba</i> lspS "Full Length"	VIIDDIYDVYCGILDELELEFD	DAVERNDVNAINDLFDYMKL	CFALALYNTINEIAYDNL	KDKGENILFY	400	
<i>P. alba</i> lspS "MEA"	VIIDDIYDVYCGILDELELEFD	DAVERNDVNAINDLFDYMKL	CFALALYNTINEIAYDNL	KDKGENILFY	400	
Consensus	VIIDDIYDVYCGILDELELEFD	DAVERNDVNAINDLFDYMKL	CFALALYNTINEIAYDNL	KDKGENILFY	400	
	409	420	430	440	450	Section 7
<i>P. alba</i> lspS Complete	LINAWADLCHNAFLQCAKWL	YNNKSTFFDDYFGNAWKS	SSSGPLQLVFA	YFAYVQINIKKEE	ENLQKXHD	Section 7
<i>P. alba</i> lspS "Full Length"	LINAWADLCHNAFLQCAKWL	YNNKSTFFDDYFGNAWKS	SSSGPLQLVFA	YFAYVQINIKKEE	ENLQKXHD	
<i>P. alba</i> lspS "MEA"	LINAWADLCHNAFLQCAKWL	YNNKSTFFDDYFGNAWKS	SSSGPLQLVFA	YFAYVQINIKKEE	ENLQKXHD	
Consensus	LINAWADLCHNAFLQCAKWL	YNNKSTFFDDYFGNAWKS	SSSGPLQLVFA	YFAYVQINIKKEE	ENLQKXHD	
	477	490	500	510	520	Section 8
<i>P. alba</i> lspS Complete	TISRPSSHIFRLCNDLASASA	BIARGETANSVSCYMR	IKGISEELATE	SVNNLIDETWKKNNK	EKLGGS	Section 8
<i>P. alba</i> lspS "Full Length"	TISRPSSHIFRLCNDLASASA	BIARGETANSVSCYMR	IKGISEELATE	SVNNLIDETWKKNNK	EKLGGS	
<i>P. alba</i> lspS "MEA"	TISRPSSHIFRLCNDLASASA	BIARGETANSVSCYMR	IKGISEELATE	SVNNLIDETWKKNNK	EKLGGS	
Consensus	TISRPSSHIFRLCNDLASASA	BIARGETANSVSCYMR	IKGISEELATE	SVNNLIDETWKKNNK	EKLGGS	
	545	550	560	570	580	Section 9
<i>P. alba</i> lspS Complete	LEAKPFVETA	INLARQSHCTYHNGDA	HISPEELIRKRVLSV	ITTEPILPFER	~	Section 9
<i>P. alba</i> lspS "Full Length"	LEAKPFVETA	INLARQSHCTYHNGDA	HISPEELIRKRVLSV	ITTEPILPFER	~	
<i>P. alba</i> lspS "MEA"	LEAKPFVETA	INLARQSHCTYHNGDA	HISPEELIRKRVLSV	ITTEPILPFER	~	
Consensus	LEAKPFVETA	INLARQSHCTYHNGDA	HISPEELIRKRVLSV	ITTEPILPFER	~	

Figure 39

