



(86) Date de dépôt PCT/PCT Filing Date: 2011/09/28
(87) Date publication PCT/PCT Publication Date: 2012/04/19
(85) Entrée phase nationale/National Entry: 2013/03/26
(86) N° demande PCT/PCT Application No.: US 2011/053787
(87) N° publication PCT/PCT Publication No.: 2012/050931
(30) Priorité/Priority: 2010/09/28 (US61/387,435)

(51) Cl.Int./Int.Cl. *C12N 9/00* (2006.01),
C12N 15/52 (2006.01), *C12N 15/63* (2006.01),
C12P 5/02 (2006.01)
(71) Demandeur/Applicant:
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA,
US
(72) Inventeurs/Inventors:
FORTMAN, JEFFREY L., US;
KATZ, LEONARD, US;
STEEN, ERIC J., US;
KEASLING, JAY D., US
(74) Agent: FETHERSTONHAUGH & CO.

(54) Titre : PRODUCTION D'ALPHA-OLEFINES EN UTILISANT DES POLYCETIDES SYNTHASES
(54) Title: PRODUCING ALPHA-OLEFINS USING POLYKETIDE SYNTHASES

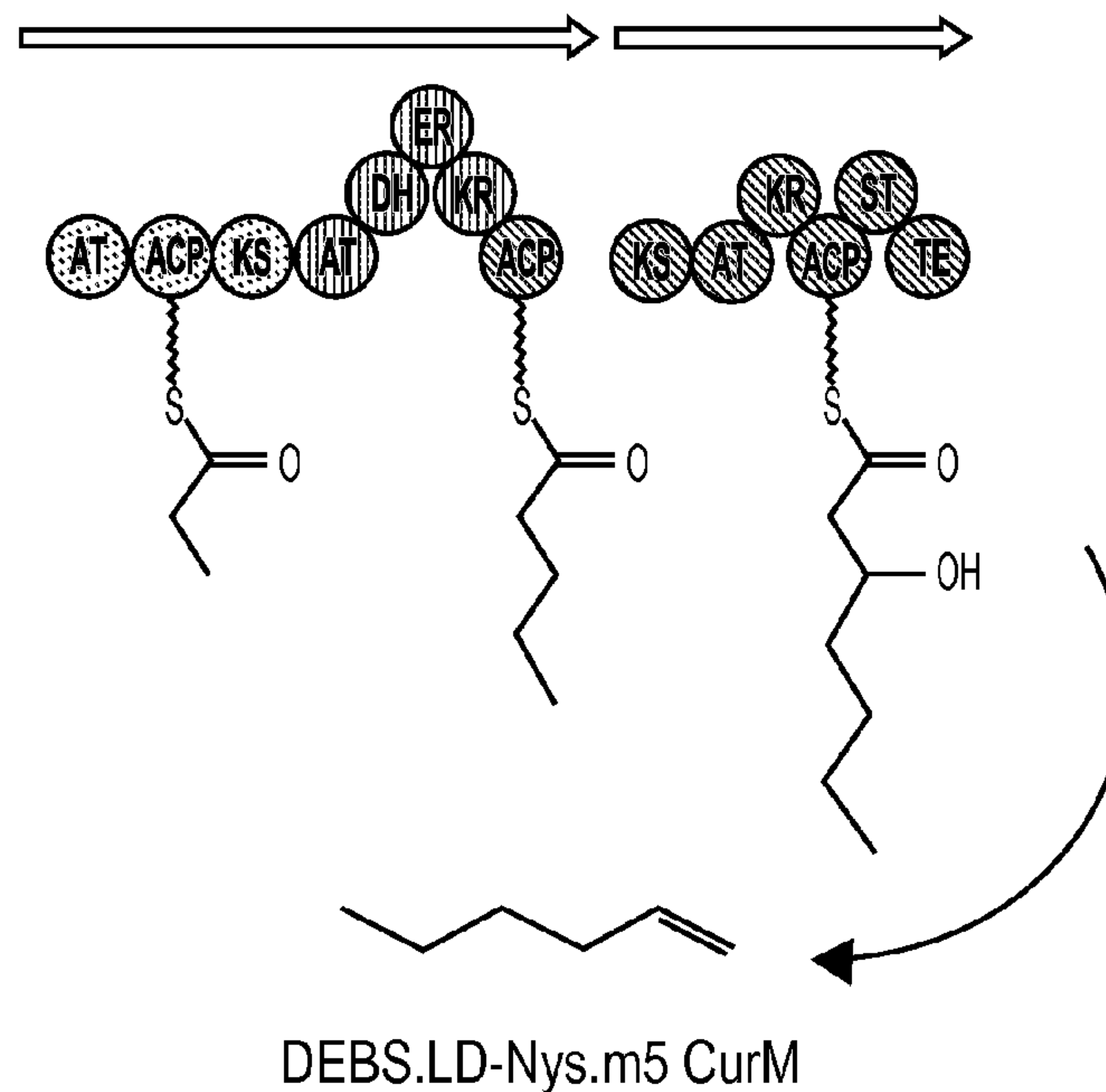


FIG. 1

(57) Abrégé/Abstract:

The present invention provides for a polyketide synthase (PKS) capable of synthesizing an α -olefin, such as 1-hexene or butadiene. The present invention also provides for a host cell comprising the PKS and when cultured produces the α -olefin.



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

(19) World Intellectual Property
Organization
International Bureau



(10) International Publication Number

WO 2012/050931 A3

(43) International Publication Date
19 April 2012 (19.04.2012)

(51) International Patent Classification:
C12P 5/02 (2006.01)

(21) International Application Number:
PCT/US2011/053787

(22) International Filing Date:
28 September 2011 (28.09.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/387,435 28 September 2010 (28.09.2010) US

(71) Applicant (for all designated States except US): **THE REGENTS OF THE UNIVERSITY OF CALIFORNIA** [US/US]; 1111 Franklin Street, 12th Floor, Oakland, CA 94607-5200 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **FORTMAN, Jeffrey, L.** [US/US]; 1534 Innes Avenue, San Francisco, CA 94124 (US). **KATZ, Leonard** [US/US]; 6389 Longcroft Drive, Oakland, CA 94611 (US). **STEEN, Eric, J.** [US/US]; 952-1/2 Haight Street, San Francisco, CA 94117 (US). **KEASLING, Jay, D.** [US/US]; 1160 Sterling Avenue, Berkeley, CA 94708 (US).

(74) Agent: **CHIANG, Robin, C.**; Lawrence Berkeley National Laboratory, Technology Transfer and Intellectual Property Mgmt., One Cyclotron Road, MS 56A-0120, Berkeley, CA 94720-8127 (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, 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, 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, MD, 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:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))

(88) Date of publication of the international search report:
21 June 2012

(54) Title: PRODUCING ALPHA-OLEFINS USING POLYKETIDE SYNTHASES

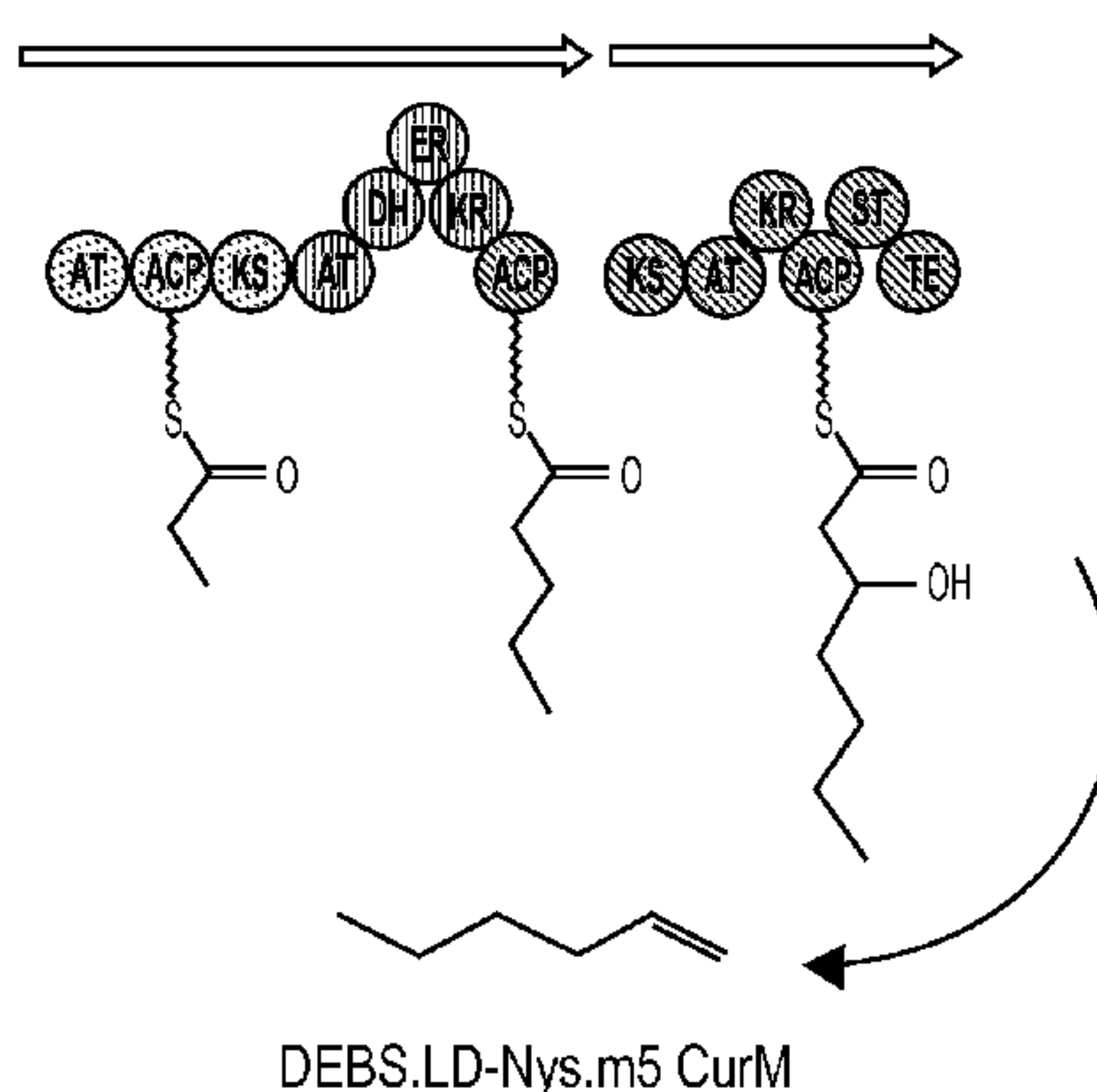


FIG. 1

(57) Abstract: The present invention provides for a polyketide synthase (PKS) capable of synthesizing an α -olefin, such as 1-hexene or butadiene. The present invention also provides for a host cell comprising the PKS and when cultured produces the α -olefin.

Producing alpha-olefins Using Polyketide Synthases

Inventors: Jeffrey L. Fortman, Leonard Katz, Eric J. Steen, Jay D. Keasling

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to U.S. Provisional Patent Application Ser. No. 61/387,435, filed September 28, 2010, which is hereby incorporated by reference in its entirety.

STATEMENT OF GOVERNMENTAL SUPPORT

[0002] This invention was made with government support under Contract No. DE-AC02-05CH11231 awarded by the U.S. Department of Energy and Award No. 0540879 awarded by the National Science Foundation. The government has certain rights in the invention

FIELD OF THE INVENTION

[0003] This invention relates generally to α -olefin production using polyketide synthases and so relates to the fields of chemistry, microbiology, and molecular biology.

BACKGROUND OF THE INVENTION

[0004] Type I polyketide synthases (PKSs) are programmable, multifunctional enzymes capable of possessing all of the catalytic capacity of fatty-acid synthases (FASs). However, unlike the FAS enzyme, which iteratively extends and fully reduces the β -carbonyl generated with each extension of the hydrocarbon backbone, PKS systems utilize discrete sets of enzymatic domains for each extension and reduction of the nascent chain. These sets, commonly referred to as modules, can incorporate a variety of extenders units resulting in different side chains. They also can encode between zero and three of the reducing domains associated with FASs, respectively leading to a ketone, hydroxy, double bond, or fully saturated carbon at the beta position of the growing polyketide chain (Hopwood and Sherman. 1990. Annual Review of Genetics 24:37-66).

[0005] Due to their modularity, PKS systems have been extensively explored for production of "unnatural" natural products (Weissman and Leadlay. 2005. Nature Reviews Microbiology 3:925-936). Hundreds of these molecules have been produced, ranging from basic lactones to modified versions of drugs and drug-like compounds.

SUMMARY OF THE INVENTION

[0006] The present invention provides polyketide synthases (PKSs) capable of synthesizing α -olefins, recombinant expression vectors for producing them, recombinant host cells that express them and produce the desired alpha olefin, methods for making alpha olefins, and alpha olefins produced by the methods. The PKSs of the invention are not naturally occurring and so are referred to as “recombinant” PKS enzymes. In some embodiments of the invention, the α -olefin is not a compound synthesized by a naturally occurring PKS. In some embodiments of the invention, the PKS is a hybrid PKS comprising modules and/or portions thereof, from two, three, four or more naturally occurring PKSs. A hybrid PKS can contain naturally occurring modules from two or more naturally occurring PKSs and/or it can contain one or more modules composed of portions, including intact domains, of two or more modules from the same naturally occurring PKS or from two or more naturally occurring PKS, or both. In some embodiments of the invention, a recombinant nucleic acid comprising a CurM module or portion thereof, which may be either naturally occurring or recombinant, is employed.

[0007] The present invention provides recombinant nucleic acids that encode PKSs of the invention. The recombinant nucleic acids include nucleic acids that include a portion or all of a PKS of the invention, nucleic acids that further include regulatory sequences, such as promoter and translation initiation and termination sequences, and can further include sequences that facilitate stable maintenance in a host cell, i.e., sequences that provide the function of an origin of replication or facilitate integration into host cell chromosomal or other DNA by homologous recombination. In some embodiments, the recombinant nucleic acid is stably integrated into a chromosome of a host cell. In some embodiments, the recombinant nucleic acid is a plasmid. Thus, the present invention also provides vectors, including expression vectors, comprising a recombinant nucleic acid of the present invention. The present invention also provides host cells comprising any of the recombinant nucleic acid and/or PKS of the present invention. In some embodiments, the host cell, when cultured under suitable conditions, is capable of producing the α -olefin. These host cells include, for example and without limitation, prokaryotes such as *E. coli* species, *Bacillus* species, *Streptomyces* species, Myxobacterial species, as well as eukaryotes including but not limited to yeast and fungal strains.

[0008] Thus, the present invention provides a wide variety of host cell comprising one or

more of the recombinant nucleic acids and/or PKSs of the present invention. In some embodiments, the host cell, when cultured, is capable of producing an α -olefin that it otherwise does not produce, or produces at a lower level, in the absence of a nucleic acid of the invention.

[0009] The present invention provides methods for producing α -olefins, said methods generally comprising: providing a host cell of the present invention, and culturing said host cell in a suitable culture medium under suitable conditions such that the α -olefin is produced.

[0010] The present invention also provides compositions comprising an α -olefin from a host cell in which the α -olefin was produced, and in some embodiments may include trace residues and/or other components of the host cell. Such trace residues and/or other components may include, for example, cellular material produced by the lysis of the host cell. The present invention also provides methods of purifying α -olefins and methods for converting them to other useful products.

BRIEF DESCRIPTION OF THE DRAWINGS

[0011] The foregoing aspects and embodiments of the invention as well as others will be readily appreciated by the skilled artisan from the following detailed description of illustrative embodiments when read in conjunction with the accompanying drawings.

[0012] Figure 1 shows an illustrative example of the modular organization of a biosynthetic pathway suitable for synthesizing 1-hexene in accordance with the invention. In this illustration, the proposed modules are sourced from the loading module of DEBS1 from the erythromycin PKS, module 5 from the nystatin PKS NysC, and CurM the terminal module of the curacin PKS. In another embodiment of the invention, the nystatin PKS module 5 is replaced with portions of modules 9 and 10 from the indanomycin PKS; this alternative embodiment has actually been used to produce 1-hexene.

[0013] Figure 2 shows types of modules employed and corresponding precursors utilized for incorporation into polyketide chains. The loading module is designated S. While any suitable loading domain can be used (such as those loading acetate and benzoic acid), only two examples are illustrated in this figure. The remaining compounds represent the structures incorporated into the growing polyketide chain employing extender modules A - P. The

dashed line indicates the C-C bond formed through Claisen condensation; atoms to the right of the bond and the C atom at the left of the dashed line represent the structures determined by the module employed. The R group represents the existing acyl chain prior to incorporation determined by the module.

[0014] Figure 3 shows: (A) a PKS system that can be used to produce 1-hexene in accordance with the invention, (B) how additional modules can be added to yield longer, even-chain α -olefins, and (C) how changing the loading module to incorporate acetate (from malonyl-CoA) will allow access to the saturated, linear, odd-chain α -olefins in accordance with the methods of the invention.

[0015] Figure 4 shows an embodiment of the invention that illustrates utilization of the avermectin PKS loading module. The side chains illustrated are merely examples and do not constitute the entire pool of side chains that can be incorporated using the avermectin loading module (or similar loading modules) in accordance with the methods and teaching of the invention.

[0016] Figure 5 shows, in part (A), an example of an illustrative pathway to 3-methylenepent-4-enoic acid, an example of the carboxylated butadiene derivatives accessible using PKSs in accordance with the methods of the invention and how the distance between the diene and carboxylate moieties can be increased via the use of additional PKS modules. Figure 5, part (B), shows the proposed mechanism of the exomethylene biosynthesis from the jamaicamide pathway (see Edwards et al. 2004. Chem Biol. 11(6):817-33; incorporated herein by reference).

[0017] Figure 6 shows a PKS for producing butadiene in accordance with the methods of the invention. While this invention is not to be limited in any manner by any proposed mechanism of action recited or shown herein, this figure, for simplicity, illustrates loss of the hydroxyl group as a water molecule, the enzymatic mechanism utilizes sulfate as a leaving group.

[0018] Figure 7 shows a PKS for producing butadiene in accordance with the methods of the invention. Parts (A) and (B) show the loading of the acrylyl-CoA using the DEBS propionyl-CoA specific loading domain modified to accept acrylyl-CoA. Part (C) shows the thiotransfer of the acrylate moiety to KS domain. Part (D) shows the binding of the malonyl-CoA and transfer to ACP domain. Part (E) shows KS catalyzing the condensation of the

moiety with release of CO₂. Part (F) shows KR catalyzing the reduction of the β -carbonyl group. Part (G) shows the final step and the release of the butadiene, CO₂, and water (as in Figure 6, the loss of the hydroxyl group is illustrated with a water molecule, but the enzymatic mechanism utilizes sulfate as a leaving group).

[0019] Figure 8 shows an enzymatic pathway accessible by the methods and materials of the invention to produce acrylyl-CoA comprising exogenously supplying propionate, and expressing PrpE and acyl-CoA dehydrogenase activities. A host cell comprising this system would be provided with propionate, which could be exogenously fed to, if not produced endogenously by, the host cell selected for production.

[0020] Figure 9 shows an enzymatic pathway accessible by the methods and materials of the invention to produce acrylyl-CoA comprising exogenously supplying propionate and glucose. A host cell comprising this system would be provided with propionate, either through exogenous feeding or the introduction of propionate biosynthesis pathway, as above, and a suitable organic molecule that the host cell can directly or indirectly convert into a pyruvate. For example, if the host cell is *E. coli*, the suitable organic molecule can be glucose. This pathway utilizes the central metabolic intermediate pyruvate to produce lactate via a lactate dehydrogenase. Lactate is then converted to lactoyl-CoA by a lactate CoA transferase, utilizing propionyl-CoA as a cofactor and releasing propionate. Lactoyl-CoA is then dehydrated using a lactoyl-CoA dehydratase to yield acrylyl-CoA. One embodiment of this invention includes the lactate dehydrogenase, LdhA, from *E. coli*, the lactate CoA transferase, Pct, from *Clostridium proponicum*, and the lactoyl-CoA dehydratase enzymes, EI and EII, from *C. proponicum*. The introduction of this pathway into *E. coli* or yeast for diene (such as butadiene) production represents a novel application of these enzymes. An embodiment of this invention is use of this pathway for PKS-based acrylate production.

[0021] Figure 10 shows an enzymatic pathway accessible by the methods and materials of the invention to produce acrylyl-CoA starting from the common metabolic precursor malonyl-CoA. This pathway generates malonyl-CoA using an acetyl-CoA carboxylase, acetyl-CoA and CO₂. Malonyl-CoA is then reduced by a malonyl-CoA reductase releasing malonyl semialdehyde. Malonyl semialdehyde is converted to 3-hydroxypropionate using a substrate specific oxidoreductase. A 3-hydroxypropionate CoA ligase catalyzes the formation of 3-hydroxypropionyl-CoA. This intermediate is then dehydrated to acrylyl-CoA by the reverse reaction of 3-hydroxypropionyl-CoA hydratase. In one embodiment of the invention,

these enzymes are the acetyl-CoA carboxylase complex (AccA/AccD) from *E. coli*, the malonyl-CoA reductase (The introduction of this pathway into *E. coli* or yeast for diene (e.g. butadiene) production represents a novel application of these enzymes and is a unique embodiment of this invention. An embodiment of this invention is use of this pathway for PKS-based acrylate production.

[0022] Figure 11 shows an enzymatic pathway accessible by the methods and materials of the invention to produce isoprene via the mevalonate pathway.

[0023] Figure 12 shows an example of an illustrative pathway accessible by a PKS provided by the invention for producing isoprene. (A) shows the loading of the acrylyl-CoA using the DEBS propionyl-CoA specific loading domain modified to accept acrylyl-CoA, and extension with malonyl-CoA to form the beta-keto ACP bound intermediate. Parts (B)-(D) show the HMG-CoA-like mechanism involved in the replacement of the β -carbonyl group with a methyl group using PKS enzymes from the PKSX (Bacillaene) cluster from *Bacillus subtilis* (Butcher, et al. 2007. Proc Natl Acad Sci U S A. 104(5):1506-9; incorporated herein by reference). This invention is not to be limited by any proposed mechanism shown herein. In this embodiment, the penultimate product is released as the free acid and subsequently decarboxylated to isoprene in accordance with the methods of the invention by either a decarboxylase, or extracellular chemical catalysis/pyrolysis.

[0024] Figure 13 shows a PKS provided by the invention for producing (*E*)-penta-1,3-diene. This figure illustrates loss of the hydroxyl group as a water molecule, but the enzymatic mechanism utilizes sulfate as a leaving group.

[0025] Figure 14 shows precursor supply pathways in *E. coli* for producing acrylyl-CoA, as described in previous figures, and [2*S*]-methylmalonyl-CoA. Each enzymes depicted can be expressed in a host cell wherein each enzyme can be independently either endogenous or native to the host cell, or introduced into recombinant

[0026] Figure 15 shows methods and materials provided by the invention for maximizing precursor supply pathways in *E. coli*. The means to maximizing acrylyl-CoA can comprise one or more of “knocking out” (eliminating or reducing the expression of) PrpC activity, knocking out YgfH activity, exogenously feeding propionate (or producing propionate endogenously), overexpressing PrpE activity to increase cytosolic pools of propionyl-CoA. From this intermediate, the introduction of the propionyl-CoA carboxylase complex

(AccA/PccB) will yield methylmalonyl-CoA (Pfeifer, et al. Science. 2001 Mar 2;291(5509):1790-2; incorporated herein by reference). This pool of propionyl-CoA can also be utilized in the pathways described in Figures 8 and 9.

[0027] Figure 16 shows an illustrative PKS provided by the invention to produce 3-hydroxy-1-octene. The PKS comprises the following elements: (i) Load module and KS1 from PikA1 (pikromycin), followed by (ii) Module 1 and KS2: AT - ACP segment from Module 5 and KS6 domain from the Nystatin PKS, (iii) Module 2: the hydroxymalonate-specific AT and contiguous ACP domains from ZmaA (zwittermicin PKS) from *Bacillus cereus*, DH, ER and KR domains from nanchangmycin PKS Module 2, and (iv) Module 3: AT-TE segment of the CurM module (curacin PKS). For the production of the precursor hydroxymalonyl-ACP, enzymes ZmaD, ZmaG, and ZmaE are also produced by or provided to the host strain. This figure illustrates loss of the hydroxyl group as a water molecule, however, it should be noted that the enzymatic mechanism utilizes sulfate as a leaving group.

[0028] Figure 17 shows an illustrative PKS provided by the invention to produce 1-decene. The PKS comprises the following elements: (i) Load module and KS1 from PikA1, followed by (ii) Module 1 and KS2: AT - ACP segment from Module 5 and KS6 domain from the nystatin PKS, (iii) Module 2 and KS3: AT - ACP segment from Module 15 and KS16 domain from the nystatin PKS, (iv) Module 3 and KS4: AT - ACP segment from Module 3 and K4 domain from the oligomycin PKS, and (v) Module 4: AT - TE segment from CurM. This figure illustrates loss of the hydroxyl group as a water molecule, however, it should be noted that the enzymatic mechanism utilizes sulfate as a leaving group.

[0029] Figure 18 shows an illustrative PKS provided by the invention to produce 1-octene. The PKS comprises the following elements: (i) Loading Module and KS1 from PikA1, followed by (ii) Module 1 and KS2: AT - ACP segment from Module 5 and KS6 domain from the nystatin PKS Module 2, (iii) and KS3: AT - ACP segment from Module 15 and KS16 domain from the nystatin PKS, and then (iv) Module 3: AT - ST segment from the CurM module. This figure illustrates loss of the hydroxyl group as a water molecule, however, it should be noted that the enzymatic mechanism utilizes sulfate as a leaving group.

DETAILED DESCRIPTION

[0030] This invention is not limited to particular embodiments described, as such may, of course, vary. The terminology used herein is for the purpose of describing particular

embodiments only, and is not intended to be limiting, because the scope of the present invention will be limited only by the appended claims.

[0031] Where a range of values is provided, each intervening value, to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between the upper and lower limits of that range is also specifically disclosed. Each smaller range between any stated value or intervening value in a stated range and any other stated or intervening value in that stated range is encompassed within the invention.

[0032] Unless defined otherwise, all technical and scientific terms used herein have the meaning commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in practicing the present invention, suitable methods and materials are now described. All publications cited are incorporated herein by reference to disclose and describe the methods and/or materials and/or results therein.

[0033] As used herein and in the appended claims, the singular forms "a", "and", and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "an α -olefin" includes a plurality of such α -olefins, and so forth.

[0034] The term "even-chain α -olefin" refers to an α -olefin with a carbon backbone, which, disregarding any functional groups or substituents, has an even number of carbon atoms.

[0035] The term "odd-chain α -olefin" refers to an α -olefin with a carbon backbone, which, disregarding any functional groups or substituents, has an odd number of carbon atoms.

[0036] The term "functional variant" describes an enzyme that has a polypeptide sequence that is at least 70%, 75%, 80%, 85%, 90%, 95% or 99% identical to an enzyme described herein. A "functional variant" enzyme may retain amino acids residues recognized as conserved for the enzyme in nature, and/or may have non-conserved amino acid residues. Amino acids can be, relative to the native enzyme, substituted (different), inserted, or deleted, but the variant has generally similar enzymatic activity as compared to an enzyme described herein. A "functional variant" enzyme may be found in nature or be an engineered mutant (recombinant) thereof.

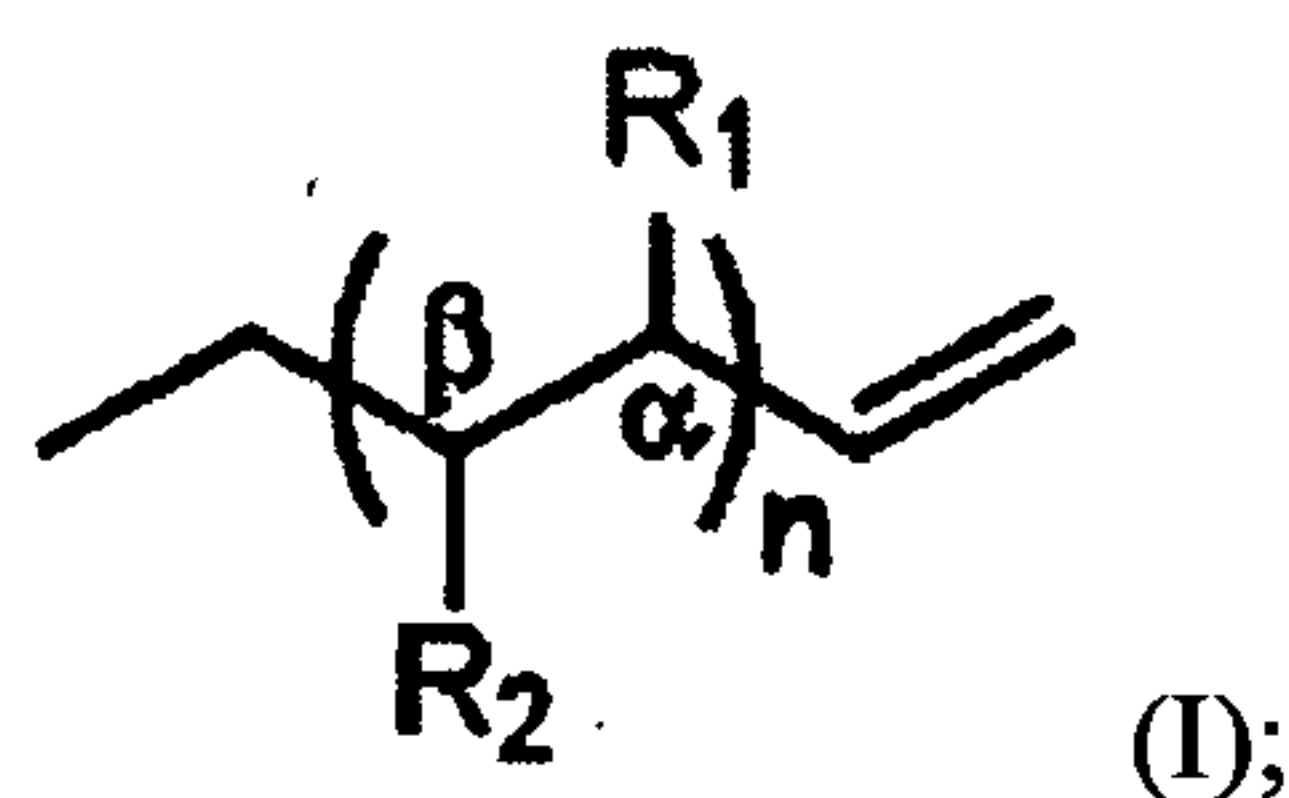
[0037] The objects, advantages, and features of the invention will become more apparent to

those persons skilled in the art upon reading the details of the invention as more fully described below.

Polyketide Synthases (PKS)

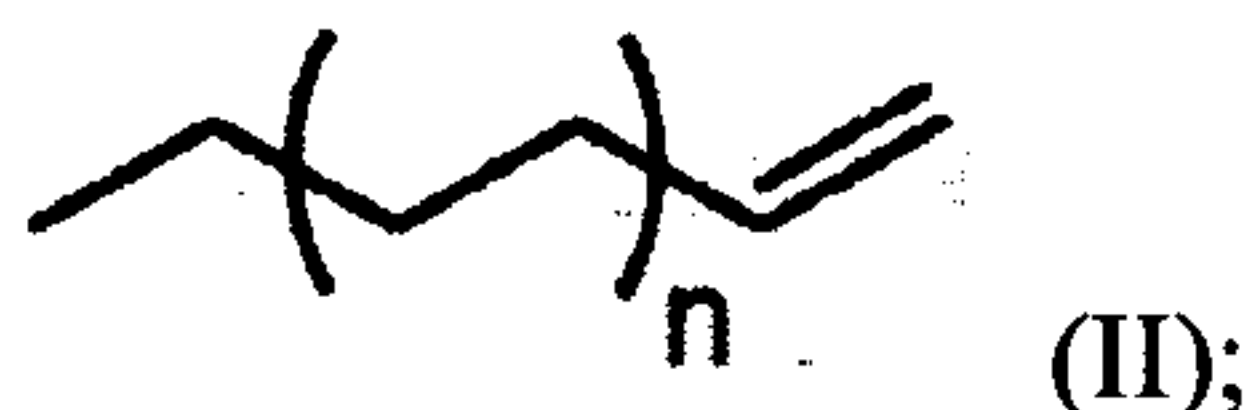
[0038] The present invention provides recombinant polyketide synthase (PKS) enzymes capable of synthesizing an α -olefin. The PKS enzymes of the invention are not naturally occurring PKS. In some embodiments of the invention, the α -olefin is not a compound synthesized by a naturally occurring PKS. In some embodiments of the invention, the PKS is a hybrid PKS comprising modules, domains, and/or portions thereof, or functional variants thereof, from two or more PKSs. Such α -olefins include the diketides and triketides, and polyketides of more than three ketide units, such as 4, 5, or 6 or more ketide units. The α -olefin can further include one or more functional groups in addition to the double bond that characterizes them. Such functional groups include, but are not limited to, ethyl, methyl and hydroxy side chains, internal olefins, and ketones.

[0039] In some embodiments of the invention, the α -olefin is an even-chain α -olefin having the following chemical structure:



wherein each R_1 is independently -H or -CH₃, each R_2 is independently -H or -OH, n is an integer, and $\alpha\beta$ is a single or double bond, with the proviso that when an $\alpha\beta$ is a double bond then the corresponding R_2 is H. In some embodiments of the invention, n is an integer from 1 to 10. n indicates the number of two-carbon-chain subunits in the carbon backbone of the α -olefin. The R_1 , R_2 , and $\alpha\beta$ within each two-carbon-subunit of a multiple subunit α -olefin is independent of the R_1 , R_2 , and $\alpha\beta$ of any other two-carbon-subunit in the molecule. In some embodiments, however, one or more, up to all, subunits have identical R_1 , R_2 , and $\alpha\beta$.

[0040] In some embodiments of the invention, the α -olefin has the following chemical structure:



wherein n is an integer from 0 to 10.

[0041] In one embodiment, the invention provides methods, host cells, and nucleic acids for making the C3-alpha olefins propylene (propene) and polymers and products derived therefrom, including but not limited to: polypropylene, acrylonitrile, propylene oxide, alcohols, cumene, acrylic acid, injection molded plastics, electronics, electrical appliances, housewares, bottle caps, toys, luggage, films, fibers, carpets, clothing, ropes, pipes, conduit, wire, cable, elastomeric polymers, acrylic fibers, nitrile rubber, acrylonitrile-butadiene-styrene (ABS) resins, styrene-acrylonitrile (SAN) resins, acrylamide, adiponitrile, polyether polyols, polyurethanes, flexible foams, rigid foams, insulation, propylene glycol, polyester resins, antifreeze, de-icing fluids, propylene glycol ethers, paints, coatings, inks, resins, cleaners, isopropanol, cosmetics, pharmaceuticals, food, ink, adhesives, 2-ethylhexanol, phthalate plasticizers, phenol, acetone, polycarbonate, phenolic resins, epoxy resins, methyl methacrylate (MMA), and acrylic esters.

[0042] In one embodiment, the invention provides methods, host cells, and nucleic acids for making the C4-alpha olefin butene and polymers and products derived therefrom, including but not limited to: polybutylene, copolymers with ethylene and/or propene, hot-melt adhesives, synthetic rubber, diesel fuel, and jet fuel.

[0043] In one embodiment, the invention provides methods, host cells, and nucleic acids for making the C4 diolefin butadiene and polymers and products derived therefrom, including but not limited to: styrene butadiene rubber (SBR), polybutadiene rubber, acrylonitrile butadiene styrene (ABS), styrene butadiene (SB) copolymer latex, nitrile rubber, adiponitrile, chloroprene, butanediol, tetrahydrofuran, tires, adhesives, coatings, high impact polystyrene, thermoplastic resins, engineering nylons (from C12 lactam), paper coating, gaskets and seals, hoses, gloves, nylon fibers, polymers, wet suits, electrical insulation, polybutylene terephthalate, spandex, and binders. Butadiene has the following chemical structure:



[0044] In one embodiment, the invention provides methods, host cells, and nucleic acids for

making the C5 α olefin: 1-pentene and polymers and products derived therefrom, including but not limited to: gasoline, polymers, adhesives, sealants, diesel fuel, and jet fuel.

[0045] In one embodiment, the invention provides methods, host cells, and nucleic acids for making the C6 α -olefin (see Figure 1, example): 1-hexene and polymers and products derived therefrom, including but not limited to comonomer, polyethylene, polymer, high density polyethylene (HDPE), linear low density polyethene (LLDPE), 1-heptanal, heptanoic acid, resin, film, plastic pipe, containers, diesel fuel, and jet fuel. 1-hexene has the following chemical structure:



and an illustration of a 1-hexene producing PKS is provided in

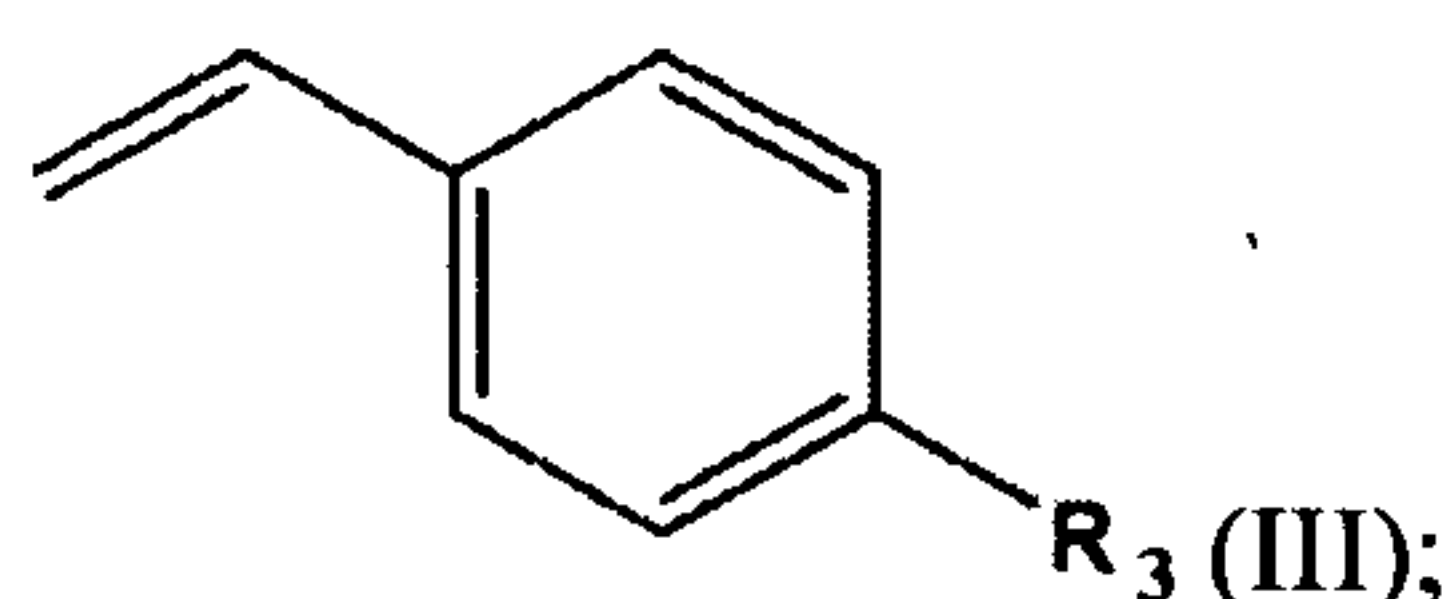
Figure 1.

[0046] In one embodiment, the invention provides methods, host cells, and nucleic acids for making the C10 α -olefin: 1-decene and polymers and products derived therefrom, including but not limited to: detergent formulations, linear alkyl benzene (LAB), linear alkyl benzene sulfonate (LABS), polyalphaolefin synthetic lubricant basestocks (PAO), heatshrink materials, electrical insulation sleeves, rash guards in clothing, polyolefin elastomers (POE), flexible foams, footwear, seat cushions, armrests, pillows, radar coolants, strings, polyol esters, detergent alcohols, plasticizer alcohols, specialty chemicals, epoxides, derivatives thereof, comonomer, intermediate in production of epoxides, amines, oxo alcohols, synthetic lubricants, synthetic fatty acids, alkylated aromatics, emulsifiers, performance waxes, cosmetic formulations, viscosity controller, solvent, decene butene copolymer, binder, film forming, decene / PVP copolymer, food additives, glazing agent, anti-foaming agent, anti-dusting agent, white mineral oil substitute, polishing agent, well fluids, alpha olefin oligomers, and the like. 1-decene has the following chemical structure:



[0047] In one embodiment, the invention provides methods, host cells, and nucleic acids for making the C8 aromatic α -olefin: styrene and polymers and products derived therefrom,

including but not limited to: homopolymers, copolymers, polystyrene, expandable polystyrene (EPS), acrylonitrile-butadiene-styrene (ABS), resins, styrene-acrylonitrile (SAN), acrylonitrile-styrene-acrylate (ASA), styrene butadiene, styrene butadiene rubber, copolymer with maleic anhydride, terephthalate, unsaturated polyester resins, containers, closures, lids and vending cups, construction; electrical and electronic parts; domestic appliances and housings; household goods and home furnishings; and toys, sporting goods and recreational articles, packaging, thermoplastics, cutlery, CDs, insulating materials, polymer bonded explosives, consumer products, renewable plastics, renewable products, hardhats, tires, etc. In some embodiments of the invention, the aromatic α -olefin has the following chemical structure:



wherein R_3 is $-H$, $-OH$, $-NH_3$, or $-NO_2$.

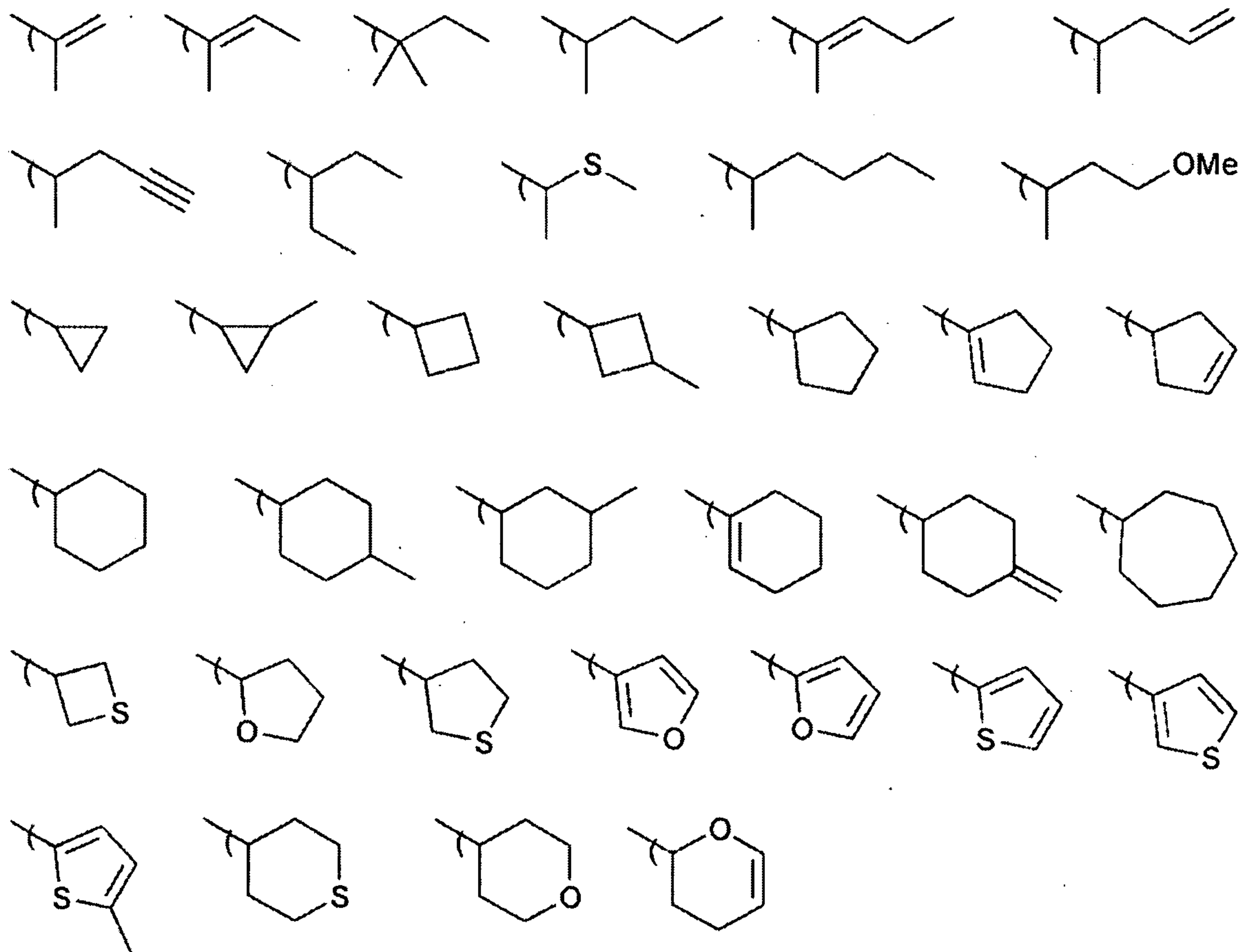
[0048] Alpha olefins are commonly used in the cosmetics and skin care industry, and the present invention therefore provides useful starting materials for making cosmetics and skin care products. For example, alpha olefin sulfonate, sulfate free personal cleaners, soap, copolymer maleic acid, and the like are all used in these industries and provided by the invention. Alpha olefins provided by the invention can also be used in the flavor and fragrance industry. For example, 3-hydroxy-1-octene and 3-oxo-1-octene can be made using the methods and materials of the invention and are used in applications where a mushroom flavor/fragrance is desired.

[0049] The present invention can also be used to generate intermediates useful in the synthesis of pharmaceuticals. These olefins can be coupled via olefin metathesis to one another or other olefin intermediates obtained via traditional chemical syntheses to yield bioactive molecules useful as drugs.

[0050] In some embodiments, the α -olefin produced in accordance with the invention is (*E*)-deca-1,5-diene, which has the following chemical structure:

[0051] In some embodiments, the α -olefin produced in accordance with the invention has the

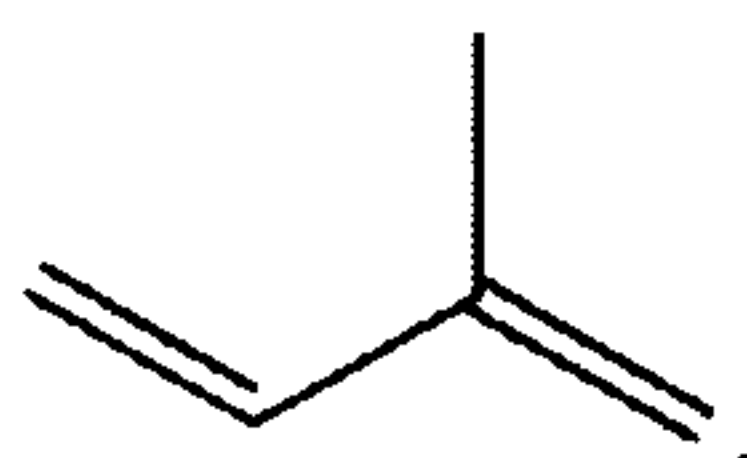
following chemical structure:  (IV); wherein R is one of the following structures:



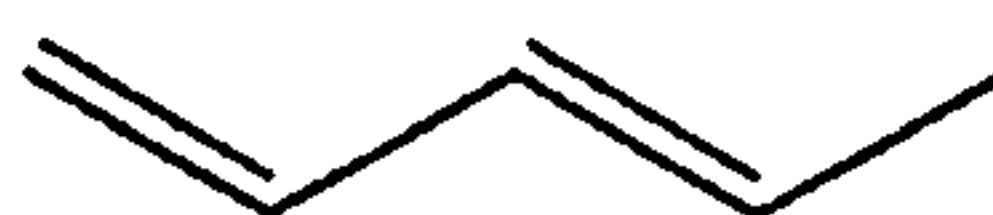
[0052] In some embodiments, the α -olefin produced in accordance with the invention is a polyolefin having chemical structure (I) and comprising at least two, three, four, five, or more C-C double bonds. Such α -olefins include, but are not limited to, diolefins, such as diolefins with two C-C double bonds on the carbon backbone. Such diolefins include, but are not limited to, butadiene, isoprene, and penta-1,3-diene. Butadiene has the chemical structure shown in [0043], above.

[0053] In some embodiments, the α -olefin produced in accordance with the invention is

isoprene, which has the following chemical structure:



[0054] In some embodiments, the α -olefin produced in accordance with the invention is penta-1,3-diene, which has the following chemical structure:



[0055] Complex polyketides comprise a large class of natural products that are synthesized in bacteria (mainly members of the actinomycete family; e.g. *Streptomyces*), fungi and plants. Polyketides form the macrolactone component of a large number of clinically important drugs, such as antibiotics (e.g. erythromycin, tylosin), antifungal agents (e.g. nystatin), anticancer agents (e.g. epothilone), immunosuppressives (e.g. rapamycin), etc. Though these compounds do not resemble each other either in their structure or their mode of action, they share a common basis for their biosynthesis, which is carried out by a group of enzymes designated polyketide synthases.

[0056] Polyketide synthases (PKS) employ short chain fatty acyl CoAs in Claisen condensation reactions to produce polyketides. Unlike fatty acid synthases that utilize acetyl CoA as the starter and malonyl CoA as the extender units, and use a single module iteratively to produce the nascent acyl chains, PKSs are composed of discrete modules, each catalyzing the chain growth of a single step. Modules can differ from each other in composition, so that, overall, a number of different starters (e.g. acetyl CoA, propionyl CoA) and extenders, some of which contain stereospecific methyl (or ethyl) side chains can be incorporated into a polyketide. In addition, PKS modules do not always reduce the 3-carbonyl formed from condensation but may leave it either unreduced (ketone), partially reduced (hydroxyl, 2,3-ene), or fully reduced (3-methylene). Many PKSs employ malonyl CoA or [S]-2-methylmalonyl CoA as the starter for polyketide synthesis. In such cases, the terminal carboxyl group is usually removed by a decarboxylase domain present at the N-terminus of the loading domain of the PKS. Thus, the structure (and chirality) of the α -carbon and β -carbonyl is determined by the module of the PKS employed in the synthesis of the growing chain at each particular step. Because of the correspondence between the modules used in the

synthesis and the structure of the polyketide produced, it is possible to program PKS synthesis to produce a compound of desired structure by selection and genetic manipulation of polyketide synthases.

[0057] Figure 2 shows the various modules and the precursor utilized by each module for incorporation into the corresponding nascent acyl (polyketide) chain to give rise to a range of compounds of interest. Table 1, below, provides illustrative PKS sources for each module in Figure 2. Each PKS source (amino acid sequence and corresponding coding sequence) is well-known to one skilled in the art and readily available. In addition, for each module in Table 1, there are other modules from other PKS (or from recombinant DNA technology) that can be used. In addition, other structures can be incorporated in the ketide or polyketide that are not shown in Table 1 and Figure 2. For example, useful loading modules include the benzoate loading module of soraphen PKS, the isobutyrate loading module of the lipomycin PKS and bafilomycin PKS, and the acrylate loading module from the difacidin pathway. The acrylate loading module from the difacidin PKS loads and dehydrates a hydroxypropionate molecule by the use of enzymes difA-E to yield a PKS with an aryl-ACP (Chen, 2006, J. Bact. 188:4024-4036; incorporated herein by reference).

[0058] The present invention also contemplates the use of functional variants of PKS modules, domains, and portions thereof. In one important embodiment, the invention provides a variety of recombinant modules that carry out the same enzymatic reactions conducted by the CurM module.

[0059] Table 1. PKS sources of the various modules.

Module	PKS Source
S1	Spiramycin PKS Loading Domain (with and without inactivation or deletion of the KS ^Q domain)
S2	Pikromycin PKS Loading Domain (with and without inactivation or deletion of the KS ^Q domain)
S3	Spiramycin PKS Loading Domain
S4	Erythromycin PKS Loading Domain
A	Rifamycin PKS Module 2
B	Oligomycin PKS Module 1

C	Spiramycin PKS Module 1
D	Pikromycin PKS Module 2
E	Oligomycin PKS Module 3
F	Erythromycin PKS Module 3
G	Oligomycin PKS Module 5
H	Primaricin PKS Module 7
I	Tylosin PKS Module 1
J	Erythromycin PKS Module 1
K	Avermectin PKS Module 7
L	Rapamycin PKS Module 1
M	Erythromycin PKS Module 4
N	Pederin Module 2
O	Ascomycin Module 4
P	FK506 Module 4
	Curacin A Chain Termination
Q	Module (CurM)

[0060] All extender modules carry the β -acyl ACP synthase (commonly called the ketosynthase or KS) domain, which conducts the decarboxylative condensation step between the extender and the growing polyketide chain, and the acyl carrier protein (ACP) domain that carries the growing acyl chain and presents it to any cognate reductive domains for reduction of the β -carbonyl. Modules can differ from each other in composition so that a number of different starter and extender units, some of which contain stereospecific side chains (e.g. methyl, ethyl, propylene) can be incorporated. The acyltransferase (AT) domain of each module determines the extender unit (e.g. malonyl CoA, methylmalonyl CoA, and the like) incorporated. In addition, PKS modules do not always reduce the β -carbonyl formed from condensation but may leave it either unreduced (ketone), partially reduced (hydroxyl, 2,3-ene) or fully reduced (3-methylene), as shown in Figure 2. The ketoreductase (KR) domain reduces the ketone to the OH function (stereospecifically); the dehydratase (DH) domain removes water from the α and β carbons leaving an α,β *trans*-double bond; the enoylreductase (ER) domain reduces the double bond to a β -methylene center; the reductive state of the β -carbonyl, therefore, is determined by the presence of functional reductive domains in the corresponding module. Less commonly, modules may contain an additional C-methylation domain (yielding an additional α -methyl side chain, as in epothilone).

[0061] The Curacin A Chain Termination Module is annotated as CurM. CurM catalyzes an extension of the nascent polyketide molecule with acetate (from malonyl-CoA). The resulting beta carbonyl is reduced to a hydroxyl group by a KR domain. The resulting beta hydroxyl

group is then sulfonated by the ST domain (from the common metabolic precursor 3'-phosphoadenosine-5'-phosphosulfate). The TE domain releases the 3-sulfo polyketide which then undergoes loss of sulfate and a decarboxylation to form a terminal olefin moiety. The chain termination module of the PKS of the present invention can comprise the ST and TE domains of the CurM Chain Termination Module and variants thereof with similar activity. Additional PKS modules carrying the combination of a sulfotransferase (pfam00685)/thioesterase have been identified in nature and can be used in additional embodiments of the invention. One such olefination module (Ols) has been characterized from *Synechococcus* sp. strain PCC 7002 (Mendez-Perez et al. 2011. Appl. Env. Microbiol. 77:4264-4267 2011). Others include, but are not limited to, PKS enzymes from *Cyanothece* sp. PCC 7424, *Cyanothece* sp. PCC 7822, *Prochloron didemni* P1-Palau, *Pseudomonas entomophila* L48, and *Haliangium ochraceum* DSM 14365. The present invention also provides consensus sequences that differ from these naturally occurring sequences but encode similar enzymatic activities.

[0062] The makeup of the PKS, therefore, determines the choice of starter and extender acyl units incorporated, the extent of reduction at each condensation step, and the total number of units added to the chain. The wide diversity of structures of polyketides seen in nature is thus attributable to the diversity in PKS enzymes.

[0063] A partial list of PKS amino acid and corresponding nucleic acid coding sequences that can be used in the PKSs of the present invention includes, for illustration and not limitation, Ambruticin (U.S. Pat. No. 7,332,576); Avermectin (U.S. Pat. No. 5,252,474; MacNeil et al., 1993, Industrial Microorganisms: Basic and Applied Molecular Genetics, Baltz, Hegeman, & Skatrud, eds. (ASM), pp. 245-256; MacNeil et al., 1992, Gene 115: 119-25); Candicidin (FRO008) (Hu et al., 1994, Mol. Microbiol. 14: 163-72); Curacin A (Chang et al., 2004, *J. Nat. Prod.*, 67 (8), pp 1356-1367; Gu et al., 2009, *J. Am. Chem. Soc.*, 131 (44), pp 16033-16035); Epothilone (U.S. Pat. No. 6,303,342); Erythromycin (WO 93/13663; U.S. Pat. No. 5,824,513; Donadio et al., 1991, Science 252:675-79; Cortes et al., 1990, Nature 348:176-8); FK506 (Motamedi et al., 1998, Eur. J. Biochem. 256:528-34; Motamedi et al., 1997, Eur. J. Biochem. 244:74-80); FK520 or ascomycin (U.S. Pat. No. 6,503,737; see also Nielsen et al., 1991, Biochem. 30:5789-96); Jerangolid (U.S. Pat. No. 7,285,405); Leptomycin (U.S. Pat. No. 7,288,396); Lovastatin (U.S. Pat. No. 5,744,350); Nemadectin (MacNeil et al., 1993, supra); Niddamycin (Kakavas et al., 1997, J. Bacteriol. 179:7515-22); Oleandomycin (Swan

et al., 1994, Mol. Gen. Genet. 242:358-62; U.S. Pat. No. 6,388,099; Olano et al., 1998, Mol. Gen. Genet. 259:299-308); Pederin (PCT publication no. WO 2003/044186); Pikromycin (Xue et al., 2000, Gene 245:203-211); Pimaricin (PCT publication no. WO 2000/077222); Platenolide (EP Pat. App. 791,656); Rapamycin (Schwecke et al., 1995, Proc. Natl. Acad. Sci. USA 92:7839-43); Aparicio et al., 1996, Gene 169:9-16); Rifamycin (August et al., 1998, Chemistry & Biology, 5: 69-79); Soraphen (U.S. Pat. No. 5,716,849; Schupp et al., 1995, J. Bacteriology 177: 3673-79); Spiramycin (U.S. Pat. No. 5,098,837); and Tylosin (EP 0 791,655; Kuhstoss et al., 1996, Gene 183:231-36; U.S. Pat. No. 5,876,991); each of the foregoing references is incorporated herein by reference.. Additional suitable PKS coding are readily available to one skilled in the art (e.g., by cloning and sequencing of DNA from polyketide producing organisms or by reference to GenBank).

[0064] Of the more than one hundred PKSs studies and reported on in the scientific literature, the correspondence between the modules used in the biosynthesis of, and the structure of, the polyketide produced is understood both at the level of the protein sequence of the PKS and the DNA sequence of the corresponding genes. The organization of modules and correspondence with polyketide structure can be identified by amino acid and/or nucleic acid sequence determination. One can thus clone (or synthesize) DNA sequences corresponding to desired modules and transfer them as fully functioning units to heterologous hosts, including otherwise non-polyketide producing hosts such as *E. coli* (Pfeifer, et al., *Science* 291, 1790 (2001); incorporated herein by reference), and polyketide-producing hosts, such as *Streptomyces* (Kao et al., *Science* 265, 509 (1994); incorporated herein by reference).

[0065] Additional genes employed in polyketide biosynthesis have also been identified. Genes that determine phosphopantetheine:protein transferase (PPTase) that transfer the 4-phosphopantetheine co-factor of the ACP domains, commonly present in polyketide producing hosts, have been cloned in *E. coli* and other hosts (Weissman et al., *ChemBiochem* 5, 116 (2004); incorporated herein by reference). While it is possible to re-program polyketide biosynthesis to produce a compound of desired structure by either genetic manipulation of a single PKS or by construction of a hybrid PKS composed of modules from two or more sources (see Weissman et al., *supra*), the present invention provides the first means for making an alpha-olefin by a recombinant PKS.

[0066] Recombinant methods for manipulating modular PKS genes to make the PKSs of the present invention are described in U.S. Pat. Nos. 5,672,491; 5,843,718; 5,830,750; 5,712,146;

and 6,303,342; and in PCT publication nos. WO 98/49315 and WO 97/02358; each of which is incorporated herein by reference. A number of genetic engineering strategies have been used with various PKSs to demonstrate that the structures of polyketides can be manipulated to produce novel polyketides (see the patent publications referenced supra and Hutchinson, 1998, Curr. Opin. Microbiol. 1:319-329, and Baltz, 1998, Trends Microbiol. 6:76-83; incorporated herein by reference). In some embodiments, the components of the hybrid PKS are arranged onto polypeptides having interpolypeptide linkers that direct the assembly of the polypeptides into the functional PKS protein, such that it is not required that the PKS have the same arrangement of modules in the polypeptides as observed in natural PKSs. Suitable interpolypeptide linkers to join polypeptides and intrapolypeptide linkers to join modules within a polypeptide are described in PCT publication No. WO 00/47724, incorporated herein by reference.

[0067] The vast number of polyketide pathways that have been elucidated to date and the present invention in combination provide a variety of different options to produce α -olefins in accordance with the invention. While the products can be vastly different in size and functionality, all employ similar methods for preparing the PKS and corresponding coding sequence and for producing the desired α -olefin. The interfaces between non-cognate enzyme partners can be optimized on a case-by-case basis. ACP-linker-KS and ACP-linker-TE regions from the proteins of interest will be aligned to examine the least disruptive fusion point for the hybrid synthase. Genetic constructions will employ sequence and ligation independent cloning (SLIC), or other sequence independent cloning techniques, so as to eliminate the incorporation of genetic "scarring".

[0068] In some embodiments, the PKS that produces the α -olefin of interest comprises the sulfotransferase (ST)-thioesterase (TE) domains from *Lyngbya majuscula* CurM or similar domains from another naturally occurring PKS or one of the recombinant domains provided by the invention. The α -olefins capable of being produced by the invention include, but are not limited to, the diketides propylene, 1-butene, and styrene and the triketides 1-hexene and 1-pentene. In one aspect of the invention, the host cell is fed or exogenously provided or endogenously produces acrylate and so produces diolefins such as 1,5-hexadiene and butadiene. In another aspect, feeding or exogenously providing or endogenous production of benzoic acid to the host cell comprising a PKS of the invention enables the production of styrene derivatives.

[0069] In some embodiments, host cells that are capable of producing diolefins are also capable of producing acrylate or acrylyl-CoA/ACP, thus eliminating the need for exogenous acrylate. By coupling one of many PKS thioesterase domains to the module loading acrylate from these precursor pathways, the PKS system is capable of producing acrylic acid. Acrylic acid can also be obtained from acrylyl-CoA or acrylyl-ACP by use of a non-PKS hydrolase in accordance with the invention. In some embodiments of the invention, host cells that are capable of producing diolefins are also capable of producing benzoate.

[0070] *L. majuscula* CurM ST-TE domains comprise the following amino acid sequence:

```
F ILSSPRSGST LLRVMLAGHS SLFSPPELHL LPFNTMKERQ EQLNLSYLGE
GLQKTFMEVK NLDATASQAL IKDLESQNLS IQQVYGMLQE NIAPRLLVDK
SPTYAMEPTI LERGEALFAN SKYIYLVRHP YSVIESFVRM RMQKLVGLGE
ENPYRVAEQV WAKSNQNILN FLSQLEPERQ HQIRYEDLVK KPQQVLSQLC
DFLNVFFEPE LLQPYQGDRM TGGVHQKSLS ISDPNFLKHN TIDESLADKW
KTIQLPYPLK SETQRIASQL SYELPNLVTT PTNQQPQVST TPSTEQPIME
EKFLEFGGNQ ICLCSWGSPE HPVVLCIHGI LEQGLAWQEV ALPLAAQGYR
VVAPDLFGHG RSSHLEMVTS YSSLTFLAQI DRVIQELPDQ PLLLVGHSMG
AMLATAIASV RPKKIKELIL VELPLPAEES KKEHAVNQLT TCLDYLSTP
QHPIFPDVAT AASRLRQAIP SLSEEF SYIL AQRITQPNQG GVRWSWDAIL
RTRSILGLNN LPGGRSQYLE MLKSIQVPTT LUYGDSSKLN RPEDLQQQKM
TMTQAKRVFL SGGHNLHIDA AAALASLILT S (SEQ ID NO:10)
```

[0071] *L. majuscula* CurM ST domain comprises the following amino acid sequence:

```
F ILSSPRSGST LLRVMLAGHS SLFSPPELHL LPFNTMKERQ EQLNLSYLGE
GLQKTFMEVK NLDATASQAL IKDLESQNLS IQQVYGMLQE NIAPRLLVDK
SPTYAMEPTI LERGEALFAN SKYIYLVRHP YSVIESFVRM RMQKLVGLGE
ENPYRVAEQV WAKSNQNILN FLSQLEPERQ HQIRYEDLVK KPQQVLSQLC
DFLNVFFEPE LLQPYQGDRM TGGVHQKSLS ISDPNFLKHN TIDESLADKW
KTIQLPYPLK (SEQ ID NO:11)
```

[0072] *L. majuscula* CurM TE domain comprise the following amino acid sequence:

```
EKFLEFGGNQ ICLCSWGSPE HPVVLCIHGI LEQGLAWQEV ALPLAAQGYR
VVAPDLFGHG RSSHLEMVTS YSSLTFLAQI DRVIQELPDQ PLLLVGHSMG
AMLATAIASV RPKKIKELIL VELPLPAEES KKEHAVNQLT TCLDYLSTP
QHPIFPDVAT AASRLRQAIP SLSEEF SYIL AQRITQPNQG GVRWSWDAIL
RTRSILGLNN LPGGRSQYLE MLKSIQVPTT LUYGDSSKLN RPEDLQQQKM
TMTQAKRVFL SGGHNLHIDA AAALASLILT S (SEQ ID NO:12)
```

[0073] In some embodiments, the PKS of the present invention comprises a naturally occurring sulfotransferase – thioesterase (ST-TE) domains, or ST or TE domain, functionally similar, but not identical, to *L. majuscula* CurM. In some embodiments, the PKS of the present invention comprises the amino acid sequences of the ST and/or TE of any of the

proteins/peptides described in Tables 2-4, or functionally variants thereof. One skilled in the art can identify such *L. majuscula* CurM-like ST and/or TE domains using available bioinformatics programs. For example, the *L. majuscula* CurM ST-TE can be split in two separately functional portions by relying on its crystal structure and annotation of catalytic boundaries with programs like protein BLAST, and the sequences can be homology-modeled to get a better grasp of the boundary of catalytic domains, using *L. majuscula* CurM ST-TE as an anchoring template. Together, such methods can be employed to make solid predictions about catalytic activity and responsible amino acid regions within a larger protein.

[0074] In some embodiments, ST and/or TE domains, or functionally variants thereof, comprise one or more of the following amino acid residues (using *L. majuscula* CurM as a reference sequence): R205, H266, S100, E124, N211, and N267. In some embodiments, ST and/or TE domains, or functionally variants thereof, comprise the following amino acid residues (using *L. majuscula* CurM as a reference sequence): R205 and H266, and optionally one or more of S100, E124, N211, and N267. In some embodiments, the PKS comprises a ST domain and a TE domains that are derived or obtained from two different organisms or sources.

[0075] Table 2. List of proteins/peptides comprising CurM-like ST-TE domains.

Ref. Protein/peptide [organism or source]	No. of amino acid residues	Accession No.
1. polyketide synthase module [Lyngbya majuscula 3L]	2211 aa protein	ZP_08432359.1 GI:332712433
2. CurM [Lyngbya majuscula]	2147 aa protein	AAT70108.1 GI:50082961
3. beta-ketoacyl synthase [Cyanospora sp. PCC 7424]	2762 aa protein	YP_002377174.1 GI:218438845
4. beta-ketoacyl synthase [Cyanospora sp. PCC 7822]	2775 aa protein	YP_003887107.1 GI:307151723
5. polyketide synthase [Prochloron didemni Pl-Palau]	2999 aa protein	AEH57210.1 GI:335387269
6. Chain A, Thioesterase Domain From Curacin Biosynthetic Pathway	286 aa protein	3QIT_A GI:325534050
7. polyketide synthase module [Lyngbya majuscula 3L]	2277 aa protein	ZP_08425908.1 GI:332705832
8. polyketide synthase [Synechococcus sp. PCC 7002]	2720 aa protein	YP_001734428.1 GI:170077790
9. polyketide synthase [Pseudomonas entomophila L48]	1217 aa protein	YP_610919.1 GI:104784421
10. KR domain-containing protein [Haliangium ochraceum DSM 14365]	3045 aa protein	YP_003265308.1 GI:262194099
11. OciA [Planktothrix agardhii NIES-205]	2858 aa protein	ABW84363.1 GI:158954787

12. OciA [Planktothrix agardhii NIVA-CYA 116]	3477 aa protein	ABI26077.1 GI:112824006
13. CurM [Burkholderia pseudomallei 668]	358 aa protein	YP_001062692.1 GI:126444569
14. amino acid adenylation domain-containing protein [Cyanothecce sp. PCC 8802]	1470 aa protein	YP_003137597.1 GI:257059709
15. amino acid adenylation domain-containing protein [Cyanothecce sp. PCC 8801]	1470 aa protein	YP_002372038.1 GI:218246667
16. polyketide synthase [Ostreococcus lucimarinus CCE9901]	18193 aa protein	XP_001416378.1 GI:145343541

[0076] Table 3. List of proteins/peptides comprising a CurM-like ST domain.

Ref. Protein/peptide [organism or source]	No. of amino acid residues	Accession No.
1. polyketide synthase module [Lyngbya majuscula 3L]	2211 aa protein	ZP_08432359.1 GI:332712433
2. CurM [Lyngbya majuscula]	2147 aa protein	AAT70108.1 GI:50082961
3. beta-ketoacyl synthase [Cyanothecce sp. PCC 7424]	2762 aa protein	YP_002377174.1 GI:218438845
4. beta-ketoacyl synthase [Cyanothecce sp. PCC 7822]	2775 aa protein	YP_003887107.1 GI:307151723
5. polyketide synthase [Prochloron didemni P1-Palau]	2999 aa protein	AEH57210.1 GI:335387269
6. polyketide synthase module [Lyngbya majuscula 3L]	2277 aa protein	ZP_08425908.1 GI:332705832
7. polyketide synthase [Synechococcus sp. PCC 7002]	2720 aa protein	YP_001734428.1 GI:170077790
8. polyketide synthase [Pseudomonas entomophila L48]	1217 aa protein	YP_610919.1 GI:104784421
9. OciA [Planktothrix agardhii NIES-205]	2858 aa protein	ABW84363.1 GI:158954787
10. OciA [Planktothrix agardhii NIVA-CYA 116]	3477 aa protein	ABI26077.1 GI:112824006
11. CurM [Burkholderia pseudomallei 668]	358 aa protein	YP_001062692.1 GI:126444569
12. KR domain-containing protein [Haliangium ochraceum DSM 14365]	3045 aa protein	YP_003265308.1 GI:262194099
16. COG3321: Polyketide synthase modules and related proteins (ISS) [Ostreococcus tauri]	11541 aa protein	XP_003074830.1 GI:308800098
17. polyketide synthase [Ostreococcus lucimarinus CCE9901]	18193 aa protein	XP_001416378.1 GI:145343541
18. modular polyketide synthase type I [Micromonas sp. RCC299]	14149 aa protein	XP_002507643.1 GI:255071123
19. hypothetical protein RBXJA2T_11932 [Rubrivivax benzoatilyticus JA2]	301 aa protein	ZP_08402700.1 GI:332526592
20. hypothetical protein Dshi_1965 [Dinoroseobacter shibae DFL 12]	310 aa protein	YP_001533306.1 GI:159044512
21. hypothetical protein glr1901		

[Gloeobacter violaceus PCC 7421]	301 aa protein	NP_924847.1 GI:37521470
22. hypothetical protein Sros_9233 [Streptosporangium roseum DSM 43021]	290 aa protein	YP_003344594.1 GI:271970398
23. hypothetical protein SAV_2309 [Streptomyces avermitilis MA-4680]	299 aa protein	NP_823485.1 GI:29828851
24. conserved hypothetical protein [Streptomyces viridochromogenes DSM 40736]	289 aa protein	ZP_07307763.1 GI:302555421
25. sulfotransferase [Frankia sp. Eu11c]	332 aa protein	YP_004017900.1 GI:312197839
26. hypothetical protein Nit79A3_2110 [Nitrosomonas sp. Is79A3]	304 aa protein	YP_004695298.1 GI:339483512
27. sulfotransferase [Thermobispora bispora DSM 43833]	346 aa protein	YP_003651260.1 GI:296268628
28. sulfotransferase [Frankia sp. Eu11c]	264 aa protein	YP_004017815.1 GI:312197754
29. hypothetical protein gll1899 [Gloeobacter violaceus PCC 7421]	320 aa protein	NP_924845.1 GI:37521468
30. SecC motif-containing protein [Shewanella loihica PV-4]	359 aa protein	YP_001093305.1 GI:127512108
31. predicted protein [Micromonas pusilla CCMP1545]	507 aa protein	XP_003055946.1 GI:303273170
32. Putative protein-tyrosine sulfotransferase [Plesiocystis pacifica SIR-1]	305 aa protein	ZP_01905212.1 GI:149916710
33. hypothetical protein PB2503_07444 [Parvularcula bermudensis HTCC2503]	310 aa protein	YP_003854688.1 GI:304321045
34. hypothetical protein PPE_01162 [Paenibacillus polymyxa E681]	422 aa protein	YP_003869548.1 GI:308067943
35. putative sulfotransferase [Streptomyces griseus subsp. griseus NBRC 13350]	339 aa protein	YP_001822417.1 GI:182434698
36. hypothetical protein LYNGBM3L_54590 [Lyngbya majuscula 3L]	318 aa protein	ZP_08430625.1 GI:332710682
37. SecC motif-containing protein [Shewanella sediminis HAW-EB3]	336 aa protein	YP_001473003.1 GI:157374403
38. sulfotransferase domain protein [Plesiocystis pacifica SIR-1]	329 aa protein	ZP_01905835.1 GI:149917336
39. sulfotransferase [Streptomyces cf. griseus XylebKG-1]	339 aa protein	ZP_08234477.1 GI:326775212
40. hypothetical protein Sros_1208 [Streptosporangium roseum DSM 43021]	347 aa protein	YP_003336949.1 GI:271962753
41. sulfotransferase [Trichodesmium erythraeum IMS101]	430 aa protein	YP_722743.1 GI:113476682
42. hypothetical protein Sros_1207 [Streptosporangium roseum DSM 43021]	336 aa protein	YP_003336948.1 GI:271962752
43. sulfotransferase [Flexistipes sinusarabici DSM 4947]	322 aa protein	YP_004602630.1 GI:336322663
44. Protein-tyrosine sulfotransferase [Harpegnathos saltator]	381 aa protein	EFN77815.1 GI:307196126
45. sulfotransferase domain protein [Rhodobacterales bacterium HTCC2083]	346 aa protein	ZP_05076043.1 GI:254462627
46. PREDICTED: similar to Transport and Golgi organization [Tribolium castaneum]	382 aa protein	13 CG32632-PB XP_968004.1 GI:91090216

47. sulfotransferase [Psychromonas ingrahamii 37]	335 aa protein	YP_943667.1 GI:119945987
48. hypothetical protein sll5046 [Synechocystis sp. PCC 6803]	316 aa protein	NP_942202.1 GI:38505581
49. sulfotransferase domain-containing protein [Roseobacter denitrificans OCh 114]	340 aa protein	YP_680960.1 GI:110677953
50. putative sulfotransferase [Burkholderia multivorans CGD2M]	271 aa protein	ZP_03574132.1 GI:221201092
51. sulfotransferase [Kangiella koreensis DSM 16069]	317 aa protein	YP_003146110.1 GI:256822147
52. Protein-tyrosine sulfotransferase [Acromyrmex echinator]	381 aa protein	EGI65733.1 GI:332025570
53. sulfotransferase [Flexistipes sinusarabici DSM 4947]	325 aa protein	YP_004602624.1 GI:336322657
54. nodulation protein noeE [Magnetospirillum magneticum AMB-1]	433 aa protein	YP_420424.1 GI:83310160
55. glycosyl transferase family 2 [Paenibacillus polymyxa SC2]	421 aa protein	YP_003945463.1 GI:310640705
56. protein-tyrosine sulfotransferase 2 [Paenibacillus polymyxa M1]	423 aa protein	CCC84070.1 GI:343095861
57. sulfotransferase [Pedobacter sp. BAL39]	346 aa protein	ZP_01884054.1 GI:149277914
58. sulfotransferase:SEC-C motif protein [Shewanella sp. HN-41]	333 aa protein	ZP_08568256.1 GI:336313314
59. sulfotransferase domain protein [Lyngbya majuscula 3L]	344 aa protein	ZP_08430778.1 GI:332710841
60. sulfotransferase domain-containing protein [Rhodopirellula baltica WH47]	325 aa protein	EGF25969.1 GI:327539348
61. hypothetical protein MettrDRAFT_3778 [Methylosinus trichosporium OB3b]	396 aa protein	ZP_06890062.1 GI:296448163
62. PREDICTED: MGC82552 protein-like [Saccoglossus kowalevskii]	436 aa protein	XP_002733820.1 GI:291227699
63. Sulfotransferase domain superfamily [Microcoleus chthonoplastes PCC 7420]	318 aa protein	ZP_05029988.1 GI:254416234
64. tyrosylprotein sulfotransferase-2 [Ictalurus punctatus]	356 aa protein	NP_001187093.1 GI:318064902
65. family 2 glycosyl transferase [Nitrosococcus halophilus Nc4]	1043 aa protein	YP_003528062.1 GI:292492623
66. PREDICTED: protein-tyrosine sulfotransferase [Apis mellifera]	380 aa protein	XP_624657.2 GI:328780257
67. tyrosine sulfotransferase [Culex quinquefasciatus]	395 aa protein	XP_001864662.1 GI:170057846
68. hypothetical protein L8106_07576 [Lyngbya sp. PCC 8106]	281 aa protein	ZP_01622104.1 GI:119489297
69. hypothetical protein NIDE3002 [Candidatus Nitrospira defluvii]	375 aa protein	YP_003798623.1 GI:302038301
70. SecC motif-containing protein [Shewanella frigidimarina NCIMB 400]	342 aa protein	YP_749583.1 GI:114562070
71. sulfotransferase [Acaryochloris marina MBIC11017]	294 aa protein	YP_001519982.1 GI:158338805
72. protein-tyrosine sulfotransferase 2 [Danio rerio]	356 aa protein	NP_956713.1 GI:41056257

73. PREDICTED: protein-tyrosine sulfotransferase-like [Bombus terrestris]	380 aa protein	XP_003394254.1 GI:340711379
74. GG17794 [Drosophila erecta]	504 aa protein	XP_001978257.1 GI:194895457
75. hypothetical protein Swit_1252 [Sphingomonas wittichii RW1]	308 aa protein	YP_001261755.1 GI:148554173
76. putative enzyme [Lyngbya sp. PCC 8106]	305 aa protein	ZP_01618957.1 GI:119484340
77. sulfotransferase [Trichodesmium erythraeum IMS101]	344 aa protein	YP_722339.1 GI:113476278
78. transport and golgi organization 13, isoform C [Drosophila melanogaster]	346 aa protein	NP_001096973.1 GI:161077803
79. hypothetical protein BRAFLDRAFT_89531 [Branchiostoma floridae]	454 aa protein	XP_002590725.1 GI:260791416
80. hypothetical protein L8106_19296 [Lyngbya sp. PCC 8106]	318 aa protein	ZP_01620966.1 GI:119487094
81. TyrosylProtein SulfoTransferase family member (tpst-1) [Caenorhabditis elegans]	380 aa protein	NP_499646.3 GI:71992370
82. methionine biosynthesis protein MetW, putative [Nitrosococcus oceanus ATCC 27777]	1039 aa protein	ZP_05048086.1 GI:254434578
83. tyrosylprotein sulfotransferase 2 [Xenopus laevis]	375 aa protein	NP_001088427.1 GI:148235112
84. glycosyl transferase family protein [Nitrosococcus oceanus ATCC 27777]	1037 aa protein	YP_343263.1 GI:77164738
85. GM17618 [Drosophila sechellia]	478 aa protein	XP_002042697.1 GI:195352394
86. GE17090 [Drosophila yakuba]	508 aa protein	XP_002100486.1 GI:195478327
87. GI14854 [Drosophila mojavensis]	459 aa protein	XP_002010160.1 GI:195131443
88. GD17135 [Drosophila simulans]	501 aa protein	XP_002106871.1 GI:195566606
89. GF22576 [Drosophila ananassae]	498 aa protein	XP_001965589.1 GI:194766965
90. transport and golgi organization 13, isoform B [Drosophila melanogaster]	499 aa protein	NP_727717.1 GI:24641809
91. GL20242 [Drosophila persimilis]	515 aa protein	XP_002023360.1 GI:195165053
92. PREDICTED: protein-tyrosine sulfotransferase-like [Acyrtosiphon pisum]	392 aa protein	XP_001942867.2 GI:328706076
93. PREDICTED: similar to MGC82552 protein [Gallus gallus]	379 aa protein	XP_415794.2 GI:118100226
94. protein-tyrosine sulfotransferase A [Loa loa]	384 aa protein	XP_003139556.1 GI:312073519
95. GA26942 [Drosophila pseudoobscura pseudoobscura]	521 aa protein	XP_001354726.2 GI:198468492
96. GK16105 [Drosophila willistoni]	466 aa protein	XP_002067611.1 GI:195439384
97. PREDICTED: protein-tyrosine sulfotransferase-like [Nasonia vitripennis]	396 aa protein	XP_001606792.1 GI:156543274
98. PREDICTED: protein-tyrosine sulfotransferase 1-like		

[Monodelphis domestica]	450 aa protein	XP_001362570.2 GI:334324786
99. predicted protein [Nematostella vectensis]	272 aa protein	XP_001630972.1 GI:156378079
100. AGAP000900-PA [Anopheles gambiae str. PEST]	392 aa protein	EAA12079.6 GI:333469474

[0077] Table 4. List of proteins/peptides comprising a CurM-like TE domain.

Ref. Protein/peptide [organism or source]	No. of amino acid residues	Accession No.
37. hypothetical protein lpl0509 [Legionella pneumophila str. Lens]	282 aa protein	YP_125875.1 GI:54293460
38. putative lipase LipA [Legionella pneumophila 130b]	282 aa protein	AAM73852.1 GI:21666982
39. Alpha/beta hydrolase [gamma proteobacterium IMCC2047]	294 aa protein	ZP_08648517.1 GI:339055924
40. hypothetical protein lpp0533 [Legionella pneumophila str. Paris]	282 aa protein	YP_122871.1 GI:54296502
41. Putative hydrolase or acyltransferase of alpha/beta superfamily [Rheinheimera sp. A13L]	279 aa protein	ZP_08572335.1 GI:336317483
42. lipase A [Legionella pneumophila subsp. pneumophila str. Philadelphia 1]	283 aa protein	YP_094512.1 GI:52840713
43. alpha/beta fold family hydrolase [Shewanella oneidensis MR-1]	288 aa protein	NP_718168.1 GI:24374125
44. alpha/beta hydrolase [Azotobacter vinelandii DJ]	300 aa protein	YP_002798221.1 GI:226943148
45. lipase [Pseudomonas syringae pv. mori str. 301020]	284 aa protein	EGH20834.1 GI:330888173
46. alpha/beta hydrolase fold protein [Dechlorosoma suillum PS]	288 aa protein	EGW60665.1 GI:345129761
47. hydrolase, alpha/beta fold family [Moritella sp. PE36]	291 aa protein	ZP_01900040.1 GI:149911422
48. lipase [Pseudomonas syringae pv. lachrymans str. M301315]	284 aa protein	EGH85562.1 GI:330987459
49. lipase [Pseudomonas syringae pv. tabaci ATCC 11528]	284 aa protein	EGH90697.1 GI:331010641
50. Alpha/beta hydrolase fold [Pseudomonas syringae pv. syringae B728a]	284 aa protein	YP_235108.1 GI:66045267
51. serine hydrolase-like 2 [Xenopus laevis]	304 aa protein	NP_001079604.1 GI:147899135
52. Alpha/beta hydrolase fold protein [Pseudomonas syringae pv. japonica str. M301072PT]	284 aa protein	EGH30625.1 GI:330899206
53. alpha/beta hydrolase fold protein [Shewanella halifaxensis HAW-EB4]	291 aa protein	YP_001674055.1 GI:167623761
54. lipase [Pseudomonas syringae pv. aesculi str. 0893_23]	284 aa protein	EGH02171.1 GI:330867462

55. Alpha/beta hydrolase fold protein [Pseudomonas syringae pv. morsprunorum str. M302280PT] 284 aa protein	EGH09818.1 GI:330875669
56. putative hydrolase [marine gamma proteobacterium HTCC2143] 308 aa protein	ZP_01616002.1 GI:119475649
57. alpha/beta fold family hydrolase [Pseudomonas syringae pv. actinidiae str. M302091] 284 aa protein	EGH65633.1 GI:330965373
58. putative alpha/beta hydrolase [Sphingobium japonicum UT26S] 290 aa protein	YP_003545632.1 GI:294012172
59. Alpha/beta hydrolase fold protein [Pseudomonas syringae pv. aptata str. DSM 50252] 284 aa protein	EGH78853.1 GI:330980750
60. Alpha/beta hydrolase fold protein [Pseudomonas syringae pv. pisi str. 1704B] 284 aa protein	EGH43451.1 GI:330940344
61. Alpha/beta hydrolase fold protein [Pseudomonas syringae pv. syringae FF5] 284 aa protein	ZP_06500867.1 GI:289679977
62. hydrolase [gamma proteobacterium NOR51-B] 299 aa protein	ZP_04957287.1 GI:254282319
63. predicted Hydrolase or acyltransferase (alpha/beta hydrolase superfamily) protein [Marinobacter algicola DG893] 305 aa protein	ZP_01893017.1 GI:149375245
64. lipase A [Legionella longbeachae D-4968] 281 aa protein	ZP_06187778.1 GI:270159122
65. alpha/beta hydrolase fold protein [Pseudomonas fluorescens WH6] 284 aa protein	ZP_07774389.1 GI:312959874
66. alpha/beta hydrolase fold protein [Alicyclophilus denitrificans K601] 300 aa protein	YP_004387755.1 GI:330824452
67. hydrolase [Sorangium cellulosum 'So ce 56'] 322 aa protein	YP_001615653.1 GI:162453286
68. alpha/beta fold family hydrolase [Shewanella violacea DSS12] 296 aa protein	YP_003557221.1 GI:294141243
69. lipase [Pseudomonas syringae pv. phaseolicola 1448A] 284 aa protein	YP_274221.1 GI:71736540
70. putative esterase [uncultured microorganism] 293 aa protein	BAI49930.1 GI:269913831
71. alpha/beta fold family hydrolase [Cellvibrio japonicus Ueda107] 298 aa protein	YP_001982425.1 GI:192359104
72. alpha/beta hydrolase fold protein [Pseudomonas fulva 12-X] 287 aa protein	YP_004474637.1 GI:333900764
73. alpha/beta hydrolase [gamma proteobacterium HdN1] 283 aa protein	YP_003810829.1 GI:304311231
74. putative hydrolase [Nocardia farcinica IFM 10152] 249 aa protein	YP_117116.1 GI:54022874
75. alpha/beta hydrolase fold protein [Haliangium ochraceum DSM 14365] 289 aa protein	YP_003269090.1 GI:262197881
76. Alpha/beta hydrolase fold protein [Pseudomonas syringae pv. syringae 642] 284 aa protein	ZP_07264074.1 GI:302187401
77. Alpha/beta hydrolase fold protein [Limnobacter sp. MED105] 310 aa protein	ZP_01916760.1 GI:149928530
78. putative lipase LipA [Legionella drancourtii LLAP12] 280 aa protein	ZP_05108808.1 GI:254495899
79. hydrolase, alpha/beta fold family [Pseudomonas syringae pv. tomato T1] 284 aa protein	ZP_03396309.1 GI:213968164

80. alpha/beta hydrolase fold protein [Rhodopseudomonas palustris CGA009]	340 aa protein	NP_946347.1 GI:39934071
81. lipase, putative [Pseudomonas syringae pv. glycinea str. B076]	284 aa protein	EFW80985.1 GI:320324913
82. alpha/beta fold family hydrolase [Pseudomonas syringae pv. tomato str. DC3000]	284 aa protein	NP_792039.1 GI:28869420
83. hydrolase or acytransferase [Azoarcus sp. BH72]	300 aa protein	YP_933620.1 GI:119898407
84. alpha/beta superfamily hydrolase [Shewanella piezotolerans WP3]	293 aa protein	YP_002311621.1 GI:212635096
85. putative hydrolase [Aromatoleum aromaticum EbN1]	292 aa protein	YP_158988.1 GI:56477399
86. putative hydrolase [Pseudomonas fluorescens SBW25]	284 aa protein	YP_002871482.1 GI:229589363
87. alpha/beta hydrolase fold protein [Rhodopseudomonas palustris TIE-1]	289 aa protein	YP_001990203.1 GI:192289598
88. hydrolase [Pseudomonas aeruginosa PA01]	286 aa protein	NP_250313.1 GI:15596819
89. putative hydrolase [Pseudomonas aeruginosa 39016]	286 aa protein	ZP_07792860.1 GI:313106637
90. hydrolase, alpha/beta fold family protein [marine gamma proteobacterium HTCC2207]	286 aa protein	ZP_01223987.1 GI:90416054
91. PA1622 [synthetic construct]	287 aa protein	AAT50924.1 GI:49083048
92. lipase [Pseudomonas syringae pv. maculicola str. ES4326]	284 aa protein	EGH59298.1 GI:330959038
93. hydrolase [Pseudomonas stutzeri DSM 4166]	282 aa protein	AEA83820.1 GI:327480510
94. alpha/beta hydrolase [Chromohalobacter salexigens DSM 3043]	289 aa protein	YP_574516.1 GI:92114588
95. alpha/beta hydrolase fold [Marinobacter aquaeolei VT8]	306 aa protein	YP_958843.1 GI:120554492
96. hydrolase [Pseudomonas stutzeri A1501]	285 aa protein	YP_001172415.1 GI:146282262
97. Hydrolase or acetyltransferase [Leptospira biflexa serovar Patoc strain 'Patoc 1 (Ames)']	287 aa protein	YP_001964755.1 GI:189912866
98. hypothetical hydrolase/acyltransferase [Photobacterium profundum 3TCK]	299 aa protein	ZP_01219631.1 GI:90411621
99. putative hydrolase [Pseudomonas aeruginosa PA7]	286 aa protein	YP_001349005.1 GI:152984242
100. hydrolase [Oxalobacter formigenes HOxBLS]	289 aa protein	ZP_04576152.1 GI:237745672

[0078] In some embodiments of the invention, a precursor molecule, such as propionate or acrylate, is provided to the PKS to produce a polyketide of interest. The precursor molecule can be fed or exogenously provided to or endogenously produced by the host cell comprising the PKS, or the host cell can produce the enzymes capable of biosynthesizing the precursor

molecule from a simpler molecule that can be fed or exogenously provided to the host cell or the host cell naturally endogenously produces. For example, *Streptomyces* species produces propionyl-CoA as part of its innate metabolism, thus eliminating the need for exogenous propionate provision.

[0079] In some embodiments of the invention, the PKS capable of producing an α -olefin of interest comprises CurM, the terminal PKS from the curacin biosynthesis pathway (Chang, 2004) or a similar module. CurM is a monomodular PKS protein containing an unusual sulfotransferase domain. This domain sulfonates the beta hydroxyl group of the penultimate product and the combination of the ST-TE domains catalyze a decarboxylation and functional dehydration (with sulfate as the leaving group) to yield the terminal olefin. Figure 1 illustrates how domains from CurM can be coupled to other PKS enzymes to produce an α -olefin, such as 1-hexene, in accordance with the methods and materials of the invention. In the example shown in Figure 1, first PKS ORF encodes a loading module specific for propionate (via the CoA) and an extension module that incorporates acetate (via malonyl-CoA) and fully reduces the β -carbonyl. In this example shown in Figure 1, the loading domain is from the erythromycin PKS (Donadio et al. 1991. Science 675-679; incorporated herein by reference) and module 5 is from the nystatin PKS (Brautaset et al. 2000. Chemistry & Biology 7:395-403; incorporated herein by reference), but there are other modules that can be used to provide the same product. The second and third proteins that constitute the multi-subunit PKS in this example come from the curacin PKS and corresponding gene cluster (see Chang et al. 2004. Journal of Natural Products 67:1356-1367; sequence updated in Gu et al. J Am Chem Soc. 2009 Nov 11;131(44):16033-5; both of which are incorporated herein by reference). Using this PKS, the first two modules can be replaced with any of several well characterized modules to yield several dozen different α -olefins. Increasing the number of upstream modules to three or more increases the number of different products into the hundreds and higher.

[0080] To ensure appropriate interactions between the two PKS proteins in this and related examples, one can use the acyl-carrier protein (ACP) and C-terminus from CurM's native enzyme partner, CurL. In general, native C- and N- terminal docking partners can be used in the combinatorial PKS enzymes of the invention. Other cognate domains from different PKS enzymes can also be used.

[0081] Figure 3A shows an exemplary PKS for producing a triketide α -olefin. Figures 3B

and 3C show exemplary extensions of this model, demonstrating how additional modules can be employed to yield longer, fully saturated, linear α -olefins.

[0082] Incorporation of the avermectin loading domain into a PKS of the invention provides access to a number of other α -olefins. Some examples of this aspect of the invention to make both known and novel α -olefins are shown in Figure 4.

[0083] In some embodiments, the PKS of the invention produces a butadiene with a pendant acid moiety, such that the butadiene is suitable for subsequent crosslinking. Figure 5 shows such a PKS that comprises a set of enzymes comprising an HMG-like system found in several PKS enzymes and corresponding gene clusters. This system converts the β -carbonyl to a number of different chemical moieties, most pertinently an exomethylene. Briefly, one of the previously described systems for incorporating an acrylate starter (DEBS (Donadio et al. 1991. *supra*) or difficidin (Chen et al. 2006. *J Bacteriol.* 188(11):4024-36; incorporated herein by reference) loading module) can be fused to an HMG-like module, such as JamE from the jamaicamide cluster (Edwards et al. 2004. *Chem Biol.* 11(6):817-33; incorporated herein by reference), and a TE domain at the C-terminus. Such a bimodular PKS enzyme can be co-expressed with the genes encoding accessory proteins required for the incorporation of the desired chemistry. In this example these enzymes are JamH, JamG and JamI (Edwards et al., 2004, *supra*).

[0084] In some embodiments of the invention, the PKS comprises a CurM chain termination module of *Lyngbya majuscula* CurM or functionally equivalent module. In some embodiments of the invention, the PKS comprises the ST and TE domains of the curM chain termination module and sequences derived from a different CurM module or another PKS entirely. In some embodiments of the invention, the PKS comprises the KR, ACP, ST and TE domains of the CurM chain termination module and sequences derived from a different CurM module or another PKS entirely. In some embodiments of the invention, the PKS comprises the AT, KR, ACP, ST and TE domains of the CurM chain termination module and sequences derived from a different CurM module or another PKS entirely.

[0085] In some embodiments, the PKS of the invention comprises an acrylate loading module, such as the acrylate loading module from the difficidin PKS (Chen et al., 2006, *supra*), which incorporates the acrylyl moiety from a hydroxypropionate precursor involving the enzymes difA-E.

[0086] The present invention also provides a PKS comprising an acrylate loading module coupled to a thioesterase domain, wherein the PKS is capable of producing acrylate (acrylic acid). The erythromycin PKS, for example and without limitation, includes suitable such modules and domains.

[0087] The following depict the amino acid sequences of SEQ ID NO:1-9

[0088] *L. majuscula* CurM (GenBank: ACV42478.1) has the following amino acid sequence (SEQ ID NO:1):

```

1 msnvskttqg dvssqevlqv lqemrsrlea vnkaktepia ivgmacrfpg gandpstywr
61 llhdgidait pvpphrwdvn ahyepnpeip gkaytkqggf ieqvdqfdpl ffgispreai
121 sldpqyrlll evtwealena gqtwtnlkns ktsvfmgvst ddyaslsnpi linnrslgvg
181 rishllglqg sniqldtacs sslvaihlac qslrsgesnl alvggvnlil spistigrct
241 mkalspdgrc ktfdaaangy ggaegcgvvv lkrlsdaitd gdlisalirg sainhdgpss
301 gltvpngmaq kqviqqalsn arlephqvsy Leahgtgtal gdpieieala aiygknrpvd
361 qplyvgsvkt nighleaaag vsalikvvla lqhqeipphl hlkqpnpyvd wdklpikipt
421 slmpwnceak priagissfg isgtnahlll eevpelikgq kakgksendl erplhiltls
481 tktekaleel vsryqnhwet ypelaidsvc ytantgraqf nhrlaviasg seeltqklrq
541 htageevvgv fsgkvpsngs eskvaflftg qgsqylmng r glyetqptfr qaldtcdhil
601 rpyldnplle ilypqdaqks ndspldqtgy tqpalfsiey allklweswg ikpnavmghs
661 vgeyvaatva gvfsledglk liaargrlmq glpaggemvs vmaseskvle tlkamsledk
721 vaiaaingpe sivilsgeaea iramathles vgiktkqlqv shafhsplme pmlaeafeava
781 nqityhqpri piisnvtgtk adksiataqy wvnhrqpvr faqgmatlhq qgyetfleig
841 akpillgmkg qclspdvgvw lpslrhgvde wqqilsslqg lyvqgakvdw sgfdrdysre
901 kvvlptypfq rerywvetsi nqqqvcsge pnlqgtpegt sttivkllsq gntkelaekv
961 ektsdlppeq klldpdllas lsqqhqgela rlttkkwfyk vqwisqaikp qrnksnnqvc
1021 hwliltdskg lgkslathlq qlgnecsvvy qadnyqnyep giyhinpshp qefeqvyqti
1081 fengklplqk vihlwsltda seqdlttetl eqaqlwgcgs tlhlqltlvk npnstppklw
1141 mitrgtqpvl sptekltvat splwglgrti asehpqlwgg lvdldpqqse devevllqqi
1201 idsqkedhla vrnrkiyvar llkhipqesq plslrsdaty litgglgalg lktaawmaek
1261 garnlvlisr rqpseqaqqt iqsleelgtq kvvlsadisv esdvanileq iqtslppllg
1321 vihaagvldd gllqqtnwer ftkvmapkvn gtwnlhkltq hlsldffvcf ssmssllgsp
1381 gqgnyaaana fmdavvhyrr emglpglsin wggwsegma trlasqhqr mqttagisli
1441 peqqiqvlee lvrtqstaqv gvlpvdwsvl akqfssanps sllelllqqe tssektderi
1501 lekqlaapit erqdilknyi qlvvaktlgi npskistddn fvelgmddslm gmevvnklsg
1561 dldfiiypre fyerptidsl tqylsaelse dnlatqpspt sleifatks psgnsarpas
1621 vssrlpgiif ilssprsgst llrvmlaghs slfsspelhl lpfntmkerq eqnlslsylge
1681 glqktfmevk nldatasqal ikdlesqnls iqqvygmlqe niaprllvdk sptyamepti
1741 lergealfan skyiylvrhp ysviesfvr rmqklvglge enpyrvaeqv waksqnlnl
1801 flsqleperq hqiryedlvk kpqqvlsqcl dflnvpfepe llqpyqgdr m tggvqhksls
1861 isdpnflkhn tidesladkw ktiqlpyplk setqriasql syelpnlvtt ptnqqpqvst
1921 tpsteqpime ekflefggnq iclcswgsp hpvvlcihgi legglawqev alplaaqgyr
1981 vvapdlfghg rsshlemvts yssltflaqi drviqelpdq plllvghsmg amlataiasv
2041 rpkkikelil velplpaees kkesavnqlt tcldylsstp qhpifpdvat aasrlrqaip
2101 slseefsyil aqritqpnnq gvrwsdaii rtrsilgl n lpgrsqyle mlksiqvptt
2161 lvygdsskln rpedlqqqkm tmtqakrvfl sgghnlhida aaalasli t s

```

[0089] HexORF1 has the following amino acid sequence (SEQ ID NO:2):

```

MADLSKLSDSRTAQPGRIVRPWPLSGCNESALRARARQLRAHLDRFPDAG
VEGVGAALAHDEQADAGPHRAVVVASSTSELLDGLAAVADGRPHASVVRG
VARPSAPVVVFVFPQGGAQWAGMAGELLGESRVFAAAMDACARAFEVTDW
TLAQVLDSPEQSRREVVQPALFAVQTSALALWRSFGVTPDAVVGHSIGE

```

LAAAHVCGAAGAADAARAAALWSREMIPLVGNGDMAAVALSADEIEPRIA
RWDDDVVLAVNGPRSVLLTGSPEPVARRVQELSAEGVRAQVINVSMAAH
SAQVDDIAEGMRSALAWFAPGGSEVPFYASLTGGAVDTRELVADYWRRSF
RLPVRFDEAIRSALEVGPFTFVEASPHPVLAALQQTLD AEGSSAAVVPT
LQRGQGGMRFLAAQAFTGGVAVDWTAAYDDVGAEPGSLPEFAPAE
DEPAESGVDWNAPPHVLRERLLAVVNGETAALAGREADAEATFRELGLDS
VLAAQLRAKVSAAIGREVNIALLYDHPTPRALAEALAGTEVAQRETRAR
TNEAAPGEPVAVVAMACRLPGGVSTPEEFWELLSEGRDAVAGLPTDRGWD
LDSLFHPDPTRSGTAHQGGGFLTEATAFDPAFFGMSPREALAVDPQQRL
MLELSWEVLERAGIPPTSLQASPTGVFVGLIPQEYGPRLAEGGEGVEGYL
MTGTTTSVASGRIAYTLGLEGPAISVDTACSSSLVAVHLACQSLRRGESS
LAMAGGVTVMPTPGMLVDFSRMNSLAPDGRCKAFSAGANGFGMAEGAGML
LLERLSDARRNGHPVLAVLRGTAVNSDGASNGLSAPNGRAQVRVIQQALA
ESGLGPADIDAVEAHGTGTRLGDPLEARALFEAYGRDREQPLHLGSVKSN
LGHTQAAAGVAGVIKMLAMRAGTLPRTLHASERSKEIDWSSGAISLLDE
PEPWPAGARPRRAGVSSFGVSGTNAHVIVEEAPESSADAVAESGVRVPVP
VVPWVVSARSAEGLAAQAERLARFVGERSDQDPVDIGFSLVRSRSLLEHR
AVVLGKGRDDLAVAGLASLASDGSATGVVSGVARGRARVAFGFSGQGAQRV
GMGAELASVYPVFAEALAEVTGALGLDPEVFGDVDRLGRTEVTQAALFAF
EVAVVRLLESFGVRPDVLIGHSIGEIAAAYVAGVFSLGDAALVGARGRL
MQALPAGGVMVAVQAGEAEVVAALEGFADRVSLAAVNGPSSVVVSGEAEA
VEQVVARLGKVKSKRLRVSHAFHSPLMEMPLADFRQVAEQITYNEPQLPV
VSNVSGRLAEPGELTTPDYWVRHVREAVRFGDGVRLAADGVGVLEVGP
DSVLTALARES LDGEDGLRAVPLLKDRPEPETLLTGVAQAFTHG VQVDW
PALLPGGRRVELPTYAFQRRRYWLEDADPTGGDPAALGLTAADHPLLGA
VPLAEDQGIVITSRLSLRTHPWLADHEIGGTVLLPGAGLVEIALRAGDEV
GCGRVEELTLEIPLVVPQEGGVTVQIRVGAPDESGWRPMTVHSRTDPEEE
WTRHVSGVLSPDVPTERYDLGAWPPAGATPVELDGFYEAYARLGAYGPS
FQGLRAAWRRGDEVFAEVSLPVEEQETAGRFTLHPALLDAALQSAGAGAF
FDSGGSMRLPFAWSGVSVFAAGASTVRVRLSPAGPDAVTVALADPTGAPV
ALVERLLIPEMSPEQLERVRGEEKEAPYVLDWVPVEVPADDLVRPERWTL
LGGADAGVGLDVAGAFASLEPSDGAPEFVVLPCVPPTSPTRAADVRQSTL
QALTVLQNWVTDERHADSRLVLVTRRAVGVAHDDVPDLTHAALWGLVRS
AQTENPGRFLLVDLDEGAELAEVLPGALGSGESQVAVRAGRVLAAARLARS
GSGGAELVPPAGAPWRLDTTSPGTLENLALVPSAEPLGLDVRVSVRAA
GLNFRDVLIALGMPGDARMGGEGAGVVTDVGSEVTTLAPGDRVMGMLSS
AFGPTAVSDHRALVRVPDDWSFEQAASVPTVFATAYYGLVDLAELRAGQS
VLVHAAAGGVGMAAVQLARHLGAEVFGTASTGKWDSL RAGGLDAEHIASS
RTVEFEETFLAATAGRGVDVLD SLAGEFVDASLRLPRGGRFVEMGKAD
IRDAERVAADHPGV TYRSFDLLEAGLDRFQEILTEVVRLFERGVLRHLPV
TAWDVRRAAEAFRFVSQARHVGKNVLMVRVWDRDGTVLITGGTGALGAL
VARHLVAEHGMRNVLLAGRRGVDAPGARELLAELETAGAQVSVVACDVAD
RDAVAELIAKVPVEHPLTAVVHTAGVVADATLTALDAERVDTVLRAKVDA
VLHLHEATRGLDLAGFVLFSSASGIFGSPGQGN YAAANSFIDAFAHHRA
QGLPALSLAWGLWARTSGMAGQLGHDDVARISRTGLAPITDDQGMALLDA
ALGAGRPLLVPVRLDRAALRSQATAGTLPPILRGLVRATVRRRAASTAAQ
GPSLAERLAGLPVTEHERIVVELVRADLAAVLGHSSSAGIDPGRAFQDMG
IDSLTAVELRNRLNGATGLRLAASLVFDYPTPNALATHILDELALDTAGA
GAAGEPDGPAPAPADEARFRRVINSIPLDRIRRAGLLDALLGLAGTSADT
AASDDFDQEEDGPAIASMDVDDLVRIALGESD TTADITEGTD RS*

[0090] HexORF1' has the following amino acid sequence (SEQ ID NO:3):

MADLSKLSDSRTAQPGRIVRPWPLSGCNESALRARARQLRAHLDRFPDAG
VEGVGAALAHDEQADAGPHRAVVVASSTSELLDGLAAVADGRPHASVVRG
VARPSAPVVFVFPQGGAQWAGMAGELLGESRVFAAAMDACARAFEPVTDW
TLAQVLDSPEQSRREVVQPALFAVQTSALALWRSFGVTPDAVVGHSIGE
LAAAHVCGAAGAADAARAAALWSREMIPLVGNGDMAAVALSADEIEPRIA
RWDDDDVVLAVNGPRSVLLTGSPEPVARRVQELSAEGVRAQVINVSMAAH
SAQVDDIAEGMRSALAWFAPGGSEVPFYASLTGGAVDTRELVADYWRRSF
RLPVRFDEAIRSALEVPGTFVEASPHVLAALQQTLD AEGSSAAVVPT
LQRGQGMRRFLLAAAQAFTGGVAVDWTAAAYDDVGAEPGSLPEFAPAE
DEPAESGVDWNAPPHVLRERLLAVVNGETAALAGREADAEATFRELGLDS
VLAAQLRAKVSAAIGREVNIALLYDHPTPRALAEALAGTEVAQRETRAR
TNEAAPGEPVAVVAMACRLPGGVSTPEEFWELLSEGRDAVAGLPTDRGWD
LDSL FHPDPTRSGTAHQRG GGF L TEATAFDPAFFGMSPREALAVDPQQRL
MLELSWEVLERAGIPPTSLQASPTGVFVGLIPQEYGPRLAEGGEGVEGYL
MTGTTTSVASGRIAYTLGLEGPAISVDTACSSSLVAVHLACQSLRRGESS
LAMAGGVTVMPTPGMLVDFSRMNSLAPDGRCKAFSAGANGFGMAEGAGML
LLERLS DARRNGHPVLAVLRGTAVNSDGASNGLSAPNGRAQVRVIQQALA
ESGLGPADIDAVEAHGTGTRLGDPIEARALFEAYGRDREQPLHLGSKSN
LGHTQAAAGVAGVIKMLAMRAGTLPRTLHASERSKEIDWSSGAISLLDE
PEPWPAGARPRRAGVSSFGVSGTNAHVIVEEAPESSADAVAESGVRVPVP
VVPWVVSARSAEGLAAQAERLARFVGERSDQDPVDIGFSLVRSRSLLEHR
AVVLGKGRDDLAVAGLASLASDGSATGVVSGVARGRARVAFGFSGQGAQRV
GMGAELASVYPVFAEALAEVTGALGLDPEVFGDVDRLGRTEVTQAALFAF
EVAVVRLLESFGVRPDVLIGHTSIGEIAAAYVAGVFSLGDAALVGARGRL
MQALPAGGVMVAVQAGEAEVVAALEGFADRVSLAAVNGPSSVVVSGEAEA
VEQVVARLGKVKSKRLRVSHAFHSPLMEPMLADFRQVAEQITYNEPQLPV
VSNVSGRLAEPGELTTPDYWVRHVREAVRFGDGVRLAADGVGV LVEVGP
DSVLTALARES LDGEDGLRAVPLLRKDRPEPETLLTGVAQAFTHGVQVDW
PALLPGGRRVELPTYAFQRRRYWLEDADPTGGDPAALGLTAADHPLLGA
VPLAEDQGIVITSRLSLRTHPWLADHEIGGTVLLPGAGLVEIALRAGDEV
GCGRVEELTLEIPLVVPQEGGVTVQIRVGAPDESGWRPMTVHSRTDPEEE
WTRHVSGVLSPDVPTERYDLGAWPPAGATPVELDGFYEAYARLGAYGPS
FQGLRAAWRRGDEVFAEVSLPVEEQETAGRFTLHPALLDAALQSAGAGAF
FDSGGSMRLPFAWSGVSVFAAGASTVRVRLSPAGPDAVTVALADPTGAPV
ALVERLLIPEMSPEQLERVRGEEKEAPYVLDWVPVEVPADDLVRPERWTL
LGGADAGVGLDVAGAFASLEPSDGAPEFVVLPCVPPTSPTRAADV RQSTL
QALTVLQNWVTDERHADSRLVLVTRRAVGVG AHDDVPDLTHAALWGLVRS
AQTENPGRFLLVDLDEGAELAEVLPGALGSGESQVAVRAGRVL AARLARS
GSGGAELVPPAGAPWRLDTTSPGTLENLALVPSAEEPLGPLDVRVSVRAA
GLNFRDVLIALGMYPGDARMGGEGAGVVTDVGSEVTTLAPGDRVMGMLSS
AFGPTAVSDHRALVRVPDDWSFEQAASVPTVFATAYYGLVDLAE L RAGQS
VLVHAAAGGVGMAAVQLARHLGAEVFGTASTGKWDSL RAGGLDAEHIASS
RTVEFEETFLAATAGRGVDVLD SLAGEFVDASLRLLPRGGRFVEMGKAD
IRDAERVAADHPGV TYRSFDLLEAGLDRFQEILTEVVRLFERGVLRHLPV
TAWDVRRAAEAFRFVSQARHVGKNVLMPRVWDRDGTVLITGGTGALGAL
VARHLVAEHGMRNVLLAGRRGVDAPGARELLAELETAGAQVSVVACDVAD
RDAVAELIAKVPVEHPLTAVVHTAGVVADATLTALDAERVDTVLRKVDA
VLHLHEATRGLDLAGFVLFSSASGIFGSPGQGN YAAANSFIDAF AHHRRRA

QGLPALSLAWGLWARTSGMAGQLGHDDVARISRTGLAPITDDQGMALLDA
 ALGAGRPLLVPVRLDRAALRSQATAGTLPPILRGLVRATVRRRAASTAAQ
 GPSLAERLAGLPVTEHERIVVELVRADLA AVLGHASAERV PADQAF AELG
 VDSLTAVELRNRLNGATGLRLAASLVFDYPTPNALATHILDELALDTAGA
 GAAGEPDGPAPAPADEARFRRVINSIPLDRIRRAGLLDALLGLAGTSADT
 AASDDFDQEEDGPAIASMDVDDLVRIALGESD TTADITEG TDRS*

[0091] HexORF2 has the following amino acid sequence (SEQ ID NO:4):

MSSASSEKIVEALRASLTENERLRRLNQELAAAAHEPVAIVSMACRFPGGVESPE
 DFWDLISEGRDAVSGLPDNRGWDLDALYDPDPEAQGKTYVREGAFLYDAAEFDAE
 LFGISPREALAMDPQQRLLMETSWEVLERAGIRPDSLRGKPVGVFTGGITS DYVT
 RHYASGTAPQLPSGVESHFMTGSAGSVFSGRIAYTYGFEGPAVTVDTACSSSLVA
 LHMAAQSLRQGECSLAFAGGVAVLPNPGTFVGF SRQRALSPDGRCKAFSADADGT
 GWGEGAGLV LLEKLS DARRNGHPVLAILRGS AVNQDGASNGLTAPNGPSQQRVIR
 AALANARLSPDDVDVVEAHGTGTPLGDPIEAQALQATYGRSRSR AERPLWLGSVKS
 NVAHAQAAAGVASVIKVMALRHRLLPKTLHADERSPHIDWHSGAVELLTEAREW
 SRTEGRARRAGVSSFGISGTNAHVIIIEAPELIKQKAKGKSENDERPLHILTL
 STKTEKALEELVSR YQNHWE TYPELAISDVCYTANTGRAQFNHRLAVIASGSEEL
 TQKLRQHTAGEEVGVFSGKVPNSGSESKVAFLFTGQGSQYLNMG RQLYETQPTF
 RQALDTCDHILRPYLDNPLLEILYPQDAQKSNDSP LDQTGYTQPALFSIEYALLK
 LWESWGIKPNVVMGHSVGEYVAATVAGVFSLEDGLKLIAARGRLMQGLPAGGEMV
 SVMASESKVLETLKAMSLEDKVAIAAINGPESIVISGEAEAIRAMATHLESVGIK
 TKQLQVSHAFHSP LMEPMLAEFEAVANQITYHQPRIPIISNVTG TKADKSIATAQ
 YWVNHRQPVRF AQGMATLHQQGYETFLEIGAKPILLGMGKQCLSPDVGVWLPSL
 RHGVDEWQQILSSLGQLYVQGA KVDWSGFDRDYSREKVVLPTYPFQRERYWVETS
 INQQQVVCSGEPNLQGTPEGTSTTIVKLLSQGNTKELAEKVEKTS DLPPEQLKLL
 PDLLASLSQQHQQELARLT TTKWFYKVQWISQAIKPQRNKSNNQVCHWLILTDSK
 GLGKSLATHLQQLGNECSVVYQADNYQNYEPGIYHINPSHPQEFEQVYQTI FENG
 KLPLQKVIHLWSLDTASEQDLTTETLEQAQLWGCGSTLHLLQTLVKNPNSTPPKL
 WMITRGTQPVLSPT EKLT VATSPLWGLGRTIASEHPQLWGGLVDLDPQGS EDEVE
 VLLQQIIDSQKEDHLAVRN RKIYVARLLKHIPQESQPLSLRSDATY LITGGLGAL
 GLKTAAWMAEKGARNLV LISRRQPSEQAQQT IQSLEELGTQVKVLSADISVESDV
 ANILEQIQTS LPPLLGVIHAAGVLD DGLLQQTNWERFTKVMAPKVNGTWN LHKL T
 QHLSLDFFVCFSSMSSLLGSPGQGN YAAANAFMDAVVHYRREMGLPGLSINWGGW
 SEGGMATRLASQHQNRMQTAGISLISPEQGIQVLEELVRTQSTAQVGVL PVDWSV
 LAKQFSSANPSSLLLELLQQETSSEKTDERILEKLQAAPITERQDILKNYIQLVV
 AKTLGINPSKISTDDNFVELGMDSLMGMEVVNKLSGDLDFI IYPREFYERPTIDS
 LTQYLSAELSEDNLATQPSPTSLEIFATKSSPSGNSARPASVSSRLPGIIFILSS
 PRSGSTLLRVMLAGHSSLFSPPELHLLPFNTMKERQEQLNLSYLGEGLQKTFMEV
 KNLDATASQALIKDLESQNLSIQQVYGMLQENIAPRLLVDKSPTYAMEPTILERG
 EALFANSKYIYLVRHPYSVIESFVRMRMQKLVGLGEENPYRVAEQVWAKSNQNIL
 NFLSQLEPERQH QIRYEDLVKKPQQVLSQLCDFLNVPFEPELLQPYQGDRMTGGV
 HQKSLSISDPNFLKHNTIDESLADKWKT IQLPYPLKSETQRIASQLSYELPNLVT
 TPTNQQPQVSTTPSTEQPI ME EK FLEFGGNQICLSWGSPEHPVVLCIHGILEQG
 LAWQEV ALPLAAQGYRVVAPDLFGHGRSSHLEMTSYSSLTFLAQIDRVIQELPD
 QPLLLVGHS MGAMLATAIASVRPKKIKELILVELPLPAEESKKESAVNQ LTTCLD
 YLSSTPQHPIFPDVATAASRLRQAIPSLSEEF SYILAQRITQPNQGGVRWSWDAI
 IRTSRILGLNNLPGGRSQYLEMLKSIQVPTTLVYGDSSKLNRPEDLQQQKMTMTQ
 AKRVFLSGGHN LHIDAAAALASLILTS*

[0092] The amino acid sequence of a PKS capable of producing 1-butene has the following amino acid sequence (SEQ ID NO:5):

MADLSKLSDSRTAQPGRIVRPWPLSGCNESALRARARQLRAHLDRFPDAGVEGVG
AALAHDEQADAGPHRAVVVASSTSELLDGLAAVADGRPHASVVRGVARPSAPVVF
VFPGQGAQWAGMAGELLGESRVFAAAMDACARAFEPVTDWTLAQVLDSPEQSRRV
EVVQPALFAVQTSALALWRSFGVTPDAVVGHSIGELAAAHVCGAAGAADAARAAA
LWSREMIPLVGNGDMAAVALSADEIEPRIARWDDDVVLAVNGPRSVLLTGSPEP
VARRVQELSAEGVRAQVINVSMAAHSQAQVDDIAEGMRSALAWFAPGGSEVPFYAS
LTGGAVDTRELVDYWRRSFRLPVRFDEAIRSALEVGPGTFVEASPHPVLAAALQ
QTLDAEGSSAAVVPTLQRGQGGMRFLAAAQAFTGGVAVDWTAAYDDVGAEPGS
LPEFAPAEEDDEPAESGVDWNAPPHVLRERLLAVVNGETAALAGREADAEATFRE
LGLDSVLAAQLRAKVSAAIGREVNIALLYDHPTPRALAEALSSGTEVAQRETRAR
TNEAAPGEPIAVVAMACRLPGGVSTPEEFWELLSEGRDAVAGLPTDRGWDLDSLF
HPDPTRSGTAHQRGGGFLTEATAFDPAFFGMSPREALAVDPQQRLMLELSWEVLE
RAGIPPTSLQASPTGVFVGLIPQEYGPRLAEGGEGVEGYLMTGTTTTSVASGRIAY
TLGLEGPAISVDTACSSSLVAVHLACQSLRRGESSLAMAGGVTVMPTPGMLVDFS
RMNSLAPDGRCKAFSAGANGFGMAEGAGMLLLERLSDARRNGHPVLAVLRGTAVN
SDGASNGLSAPNGRAQVRVVIQQALESGLGPADIDAVEAHGTGTRLGDPPEARAL
FEAYGRDREQPLHLGSKSNLGHQTAAAGVAGVIKMLAMRAGTLPRTLHASERS
KEIDWSSGAISLLDEPEPWPAGARPRRAGVSSFGISGTNAHAIIEEAPELIKGQK
AKGKSENDLERPLHILTLSTKTEKALEELVSRVQNHWEYPELAISDVCYTANTG
RAQFNHRLAVIASGSEELTQKLRQHTAGEEVGVFSGKVPNSGSESKVAFLFTGQ
GSQYLNMGRLQYETQPTFRQALDTCDHILRPYLDNPLLEILYPQDAQKSNDSPLD
QTGYTQPALFSIEYALLKLWESWGIKPNVVMGHSVGEYVAATVAGVFSLEDGLKL
IAARGRLMQGLPAGGEMVSVMASESKEVLETLKAMSLEDKVAIAAINGPESIVISG
EAEAIRAMATHLESVGIKTKQLQVSHAFHSPLMEPMLAEFEAVANQITYHQPRIP
IISNVTGKADKSIATAQYVWNHVRQPVRFQAQGMATLHQQGYETFLEIGAKPILL
GMGKQCLSPDVGVWLPSLRHGVDEWQQIILSSLGQLYVQGAQVDWSGFDRDYSREK
VVLPTYPFQRRERYWVETSINQQQVVCSEPNLQGTPEGTSTTIVKLLSQGNTKEL
AEKVEKTSDDLPEQLKLLPDLLASLSQQHQQELARLTTKKWFYKVQWISQAIPQ
RNKSNNQVCHWLILTDKGLGKSLATHLQQLGNECSVVYQADNYQNYEPGIYHIN
PSHPQEFQVYQTI FENGKLPLOKVIHLWSLDTASEQDLTTETLEQAQLWGCGST
LHLLQTLVKNPNSTPPKLWMITRGTQPVLSPTTEKLTVATSPLWGLGRTIASEHPQ
LWGGLVDLDPQGSSEDEVEVLLQQIIDSQKEDHLAVRNRIYVARLLKHIPQESQP
LSLRSDATYELITGGLGALGLKTAAWMAEKGARNLVLSRRQPSEQAQQTIQSLEE
LGTQVKVLSADISVESDVANILEQIQTSPLPPLGVIHAAGVLDDGLLQQTNWERF
TKVMAKPVNGTWNLHKLTQHLSLDFVCFSSMSSLLGSPGQGNVYAAANAFMDAVV
HYRREMGLPGLSINWGGWSEGGMATRLASQHQNRMQTAGISLISPEQGIQVLEEL
VRTQSTAQVGVLVDWVSLAKQFSSANPSSLLLELLQQETSSEKTDERILEKLQA
APITERQDILKNYIQLVVAKTLGINPSKIISTDDNFVELGMDSLMGMEVVNKLSGD
LDFIIPREFYERPTIDSLTQYLSAELSEDNLATQPSPTSLEIFATKSSPSGNSA
RPASVSSRLPGIIFILSSPRSGSTLLRVMLAGHSSLFSPPELHLLPFNTMKERQE
QLNLSYLGEGLQKTFMEVKNLDATASQALIKDLESQNLISIQQVYGMLQENIAPRL
LVDKSPTYAMEPTILERGEALFANSKYIYLVRHPYSVIESFVRMRMQKLVGLGEE
NPYRVAEQVWAKSNQNILNFLSQLEPERQHQIRYEDLVKKPQQVLSQLCDFLNVP
FEPELLQPYQGDRMTGGVHQKSLISDPNFLKHNTIDESLADKWKTIQLPYPLKS
ETQRIASQLSYELPNLVTTPTNQPPQVSTTPSTEQPIMEEEKFLEFGGNQICLCSW
GSPEHPVVLCHIGILEQGLAWQEVALPLAAQGYRVVAPDLFGHGRSSHLEMVTSY

SSLTFLAQIDRVIQELPDQPLLLVGHSMGAMLATAIASVRPKKIKELILVELPLP
 AEESKKESAVNQLTTCLDYLSTPQHPIFPDVATAASRLRQAIPSLSEEFYSYILA
 QRITQPNQGGVRWSWDAIIRTRSILGLNNLPGGRSQYLEMLKSIQVPTTLVYGDS
 SKLNRPEDLQQQKMTMTQAKRVFLSGGHNHIDAAAALASLILTS*

[0093] The amino acid sequence of a PKS capable of producing propene has the following amino acid sequence (SEQ ID NO:6):

MAGHGDATAQKAQDAEKSEGDSDAIAVIGMSCRFPGAPGTAEFWQLLSSGADAVVTAADGRR
 RGTIDAPADFDAFFGMSPREAAATDPQQRLVLELGWEALEDAGIVPESLRGEAASVFGAM
 NDDYATLLHRAGAPTDYATATGLQHSMIANRLSYFLGLRGPSLVVDTGQSSSLVAVALAVES
 LRGGTSGIALAGGVNLVLAEEGSAAMERVGALSPDGRCHTFDARANGYVRGEGGAIVVLKPL
 ADALADGDRVYCVVRGVATGNDGGGPGLTVPDRAGQEAVLRAACDQAGVRPADVRFVELHGT
 GTPAGDPVEAEALGAVYGTGRPANELLVGSVKTNIHLEGAAGIAGFVKAALCLHERALPA
 SLNFETPNPAIPLERLRLKVQTAHAALQPGTGGGPLLAVSAFGMGGTNCHVVLEETPGGRQ
 PAETGQADACLFSSASPMLLLSARSEQALRAQAARLREHLEDGADPLDIAYSLATTRTRFEH
 RAAVPCGDPDRLSSALAALAAGQTPRGVRIGSTDADGRLALLFTGQGAQHPPGMGQELYTTDP
 HFAAALDEVCEELQRCGTQNLREVMFTPDQPDLLDRTEYTQPALFALQTALYRTLRTARGETQA
 HLVLGHVSGEITAAGIAGVLDLPDAARLITARAHVMGQLPHGGAMLSVQAAEHDLDQLAHTH
 GVEIAAVNGPTHCVLSGPRTALEETAQHLREQNVRHTWLKVSHAFHSALMDPMLGAFRDTLN
 TLNYQPPTIPLISNLTGQIADPNHLCTPDYWDHARHTVRFADAVQTAHHQGTTTYLEIGPH
 PTLTLLHHTLDNPTTIPTLHRERPEPETLTQAIAAVGVRTDGDIDWAVLCGASRPRRVELPT
 YAFQRRTHWAPGLTPNHAPADRPAAEPQRAMAVGPVSREALVRLVGETTASVLGLDGPDEVA
 LDRPFTSQGLDSMTAVELAGLLGTAAGVALDPTLVYELPTPRAVADHLAKTLLGESAADADQ
 EVNGRTGEAEAKAGDPPIAVIGIGCRFPGGVATPDDLWELVASGTDASTFPTDRGWDLDGLY
 DPDPSTPGKSYVRHGGFLHDAQAQFDAEFFGISPREATAMDPQQRLLLETSWEALERAGVVPE
 SLRGGRTGVFVGTTAPEYGPRLHEGTDGYEGFLLTGTTASVASGRIAYALGTRGPALTVDTA
 CSSLVALHLAVQSLRRGECDLALAGGTTVMMSGPMFVEFSRQRGLAPDGRCKAFSADADGT
 AWAEGVGMLLVERLSDAERLGHRLAVVRGTAVNQDGASNGLTAPSGPAQQQVIRDALS DAG
 LSADDIDAVEAHGTGTALGDPIEAGALLATYGHKPKRQTPVWLGLSKSNIGHTQAAAGIAGII
 KMOVQALRHDTLPRTLHADHPSSKVDWDAGPLQLLTDARWPADPDRPRRAGISAFGVSGTNA
 HVVLEEPPELIKQKAKGKSENDERPLHILTLSTKTEKALEELVSRYQNHWETYPELAISD
 VCYTANTGRAQFNHRLAVIASGSEELTQKLRQHTAGEEVGVFSGKVPNSGSESKVAFLFTG
 QGSQYLMGRQLYETQPTFRQALDTCDHILRPYLDNPLLEILYPQDAQKSNDSPLDQTYTQ
 PALFSIEYALLKLWESWGIKPNVVMGHSVGEYVAATVAGVFSLEDGLKLI AARGRLMQGLPA
 GGEMVSVMASESKVLETLKAMSLEDKVAIAAINGPESIVISGEAEAIRAMATHLESVGIKTK
 QLQVSHAFHSPLMEPMLAEFEAVANQITYHQPRIPISNVTGTKADKSIATAQYWNHVRQP
 VRFAQGMATLHQQGYETFLEIGAKPILLGMGKQCLSPDVGVWLPSLRHGVDEWQQILSSLGQ
 LYVQGAQVDWSGFDRDYSREKVVLPITYPFQRRERYWVETSINQQQVVCSEPNLQGTPEGTST
 TIVKLLSQGNTKELAEKVEKTS DLPPEQLKLLPDLLASLSQQHQQELARLTTKKWFYKVQWI
 SQAIKPQRNKSNNQVCHWLILTD SKGLGKSLATHLQQLGNECSVVYQADNYQNYEPGIYHIN
 PSHPQEFQVYQTI FENGKLPLQKVIHLWSLDTASEQDLTTETLEQAQLWGCGSTLHLLQTL
 VKNPNSTPPKLWMITRGTQPVLSPTTEKLTVATSPLWGLGRTIASEHPQLWGGLVDLDPQGSE
 DEVEVLLQQIIDSQKEDHLAVRNRKIYVARLLKHIPQESQPLSLRSDATYELITGGLGALGLK
 TAAWMAEKGARNLVLSRRQPSEQAQQTIQSLEELGTQVKVLSADISVESDVANILEQIQTS
 LPPLLGVIIHAAGVLDDGLLQQTNWERFTKVMAPKVNGTWNHKLTOHLSLDDFFVCFSSMSSL
 LGSPGQGNIAAANAFMDAVVHYRREMGLPGLSINWGGWSEGGMATRLASQHQNRMQTAGISL
 ISPEQGIQVLEELVRTQSTAQVGVLPVDWSVLAKQFSSANPSSLLLELLQQETSSEKTDERI
 LEKLQAAPITERQDILKNYIQLVVAKTLGINPSKISTDDNFVELGMDSLMGMEVVNKLSGDL
 DFIIYPREFYERPTIDSLTQYLSAELSEDNLATQPSPTSLEIFATKSSPSGNSARPASVSSR

LPGIIFILSSPRSGSTLLRVMLAGHSSLFSPPELHLLPFNTMKERQEQLNLSYLGEGLQKTF
 MEVKNLDATASQALIKDLESQNLSIQQVYGMLQENIAPRLLVDKSPTYAMEPTILERGEALF
 ANSKYIYLVRHPYSVIESFVRMRMQKLVGLGEENPYRVAEQVWAKSNQNILNFLSQLEPERQ
 HQIRYEDLVKKPQQVLSQLCDFLNVPFEPPELLQPYQGDRMTGGVHQKSLSISDPNFKHNTI
 DESLADKWKTIQLPYPLKSETQRIASQLSYELPNLVTTPTNQQPQVSTTPSTEQPIIMEEKFL
 EFGGNQICLCSWGSPEHPVVLCHGILEQQGLAWQEVALPLAAQGYRVVAPDLFGHGRSSHLE
 MVTSYSSLTFLAQIDRVIQELPDQPLLLVGHSMGAMLATAIASVRPKKIKELILVELPLPAE
 ESKKESAVNQLTTCLDYLSSTPQHPIFPDVATAASRLRQAIPSLSEEF SYILAQRITQPNQG
 GVRWSWDAIIRTRSILGLNNLPGGRSQYLEMLKSIQVPTTLVYGDSSKLNRPEDLQQQKMTM
 TQAKRVFLSGGHNHIDAAAALASLILTS

[0094] The amino acid sequence of a PKS capable of producing styrene has the following
 amino acid sequence (SEQ ID NO:7):

MTKEYTRPQSAPLTEGDLTLIVAHLAERLRMDARFIDVHEPF SRHGLDSRGAVDLVVDLRT
 ALGRPLSPVVWQHPTPDALARHLAGGADAREGQARADSAYERPGAPNEPIAIVGMACRFPG
 APDVDSYWRLLSGGVDAVTEVPAGRWDMDAFYDRDPRSLGDVSTLRGGFIDDVDRFDAMFFG
 ISPREAVSMDPQQRLMLELAWEALEDAGIVAERLKESTGVFFGCIWDDYVTLIHQGRGAI
 AQHTVTGNHRSIIANRVSYTLDLRGPSMTVDSACSSALVTIHMACESLRSGESTLALAGGVN
 LNIAPESTIGVHKFGGLSPDGRCFTFDARANGYVRGEGGGVVVLKRLSSAIADGDPIICVIR
 GSAVNNDGASNGLTGPNPLAQEAVLRTAYERAGVNPADVQYVELHGTGTQLGDPVEASALGA
 VLGKRRPAERPLLVGSAKTNVGHLEGAAGIVGLLKAALCLKHKQLAPNLNFETPNPHIPFAE
 LNLKVQOGALGPWPDMDRPLVCGVSSFGLGGTNAHVVLSEWASLEAELHPLAAESPEALREEV
 QRRLLTMTSLVGRAPLSFLCGRSAAQRSAKEHRLAVTARSFEELKQROLLGFLEHEKHVSVA
 GRVDLGAAPKVVFVFAQGGAQWFGMGRALLQREPVRTTIEQCSSFQQNLGWSLLDELMTD
 RESSRLDEIDVSLPAIIISIEIALAAQWRAWGVEPAFVVGHSTGEIAAAHVAGVLSIEDAMRT
 ICAYGRIIRKLRGKGGMGLVALSWEDAGKELTGYEGRLFRAIEHSADSTVLAGEPDALDALL
 QALERKNVFCRRVAMDVAPHCPQVDCLRDELFDALREVRPNKAQIPIVSEVTGTALDGERFD
 ASHWVRNFGDPALFSTAIDHLLQEGFDIFLELTPHPLALPAIESNLRRSGRRGVVLP SLRN
 EDERGVMLDTLGVLYVRGAPVRWDNVYPAAFESMPLPSTAGGGKPLPPMPLLI SARTDAALA
 AQAARLRAHLDSHLDLELVDVAYS LAATRTHFERRAVVVARDRAGILDGLDALAHGGS AALL
 GRSAAHGKLAILFTGQGSQRPTMGRALYDAFPVFRGALDAAAHLDRDLDRPLRDVLFAPDG
 SEQAARLDQTAFTQPALFALEVALFELLQSFGLKPALLLGH SIGELVAAHVAGVLSLQDACT
 LVAARAKLMQALPQGGAMVTLQASEQEARDLLQAAEGRVSLAAVNGHLSTVVAGDEDAVLKI
 ARQVEALGRKATRLRVSHAFHSPHMDGMLDDFRRVAQGLTFHPARIP IISNVTGARATDQEL
 ASPETWVRHVRDTRVFLDGVRTLHAEGARAFLELGPHPVLSALA QDALGHDEGPSPCAFLPT
 LRKGRDDAEAF TAALGALHAAGLTPDWN AFFAPFAPCKVPLPTYTFQRRERFWLDASTAHAAS
 ATPAAALEGRFWQAVESGDI DTLSSSELHVDGDEQRAALALVLPTLSSFRH KRQE QSTVD AWR
 YRVTWKPLTTAATPADLAGTWLLVVP SALGDDALLATL TEALTRRGARVLALRVSDI HIGRS
 ALVEHLREALAETAPLRGVLSLLALDEHRLADRSALPAGLALS LALVQGLDDLAIEAPLWLF
 TRGAVSIGHSDPITHPTQAMIWGLGRVVGLEH PERWGG LVDVSAGVDES AVGRLLPALAQRH
 DEDQLALRPAGLYARRIVRAPLGDAPPAREFRPRGTILITGGTGALGAHVARW LARQGA EHL
 ILISRRGAEAPGASELHAELNALGVRTTLAACDVADRSALQALLDSIPSDCPLTAVFHTAGA
 RDDGLIGDMTPERIERVLAPKLD SALHLHELTKNSALDAFVLYASLSGVLGNPGQANYAAAN
 AFLDALAEHRRSLGLTATSVAWGGWGGGGMATERVAAQLQQRGLLQMAPSLALAA LAQALQQ
 DETTITVADIDWSRFAPAFS VARQRPLLRDLPEAQRALQASEGASSEHGPATGLLDELRSRS
 ESEQDLLLATLVRGETATVLGHAEASHVDPDKGFMDLGLDSLMTVELRRRLQKATGVKLPPT
 LAFDHPSPHRVAFFLRDSLSEDNLATQPSPTSLEIFATKSSPSGNSARPASVSSRLPGIIFI
 LSSPRSGSTLLRVMLAGHSSLFSPPELHLLPFNTMKERQEQLNLSYLGEGLQKTFMEVKNLD
 ATASQALIKDLESQNLSIQQVYGMLQENIAPRLLVDKSPTYAMEPTILERGEALFANSKYIY

LVRHPYSVIESFVRMRMQKLVGLGEENPYRVAEQVWAKSNQNI LNFLSQLEPERQHQIRYED
LVKKPQQVLSQLCDFLNVPFEPELLQPYQGDRMTGGVHQKSLSISDPNFLKHNTIDESLADK
WKTIQLPYPLKSETQRIASQLSYELPNLVTTPTNQPPQVSTTPSTEQPIMEEFLEFGGNQI
CLCSWGSPEHPVVLCHGILEQGLAWQEVALPLAAQGYRVVAPDLFGHGRSSHLEMVTSYSS
LTFLAQIDRVIQELPDQPLLLVGHSMGAMLATAIASVRPKKIKELILVELPLPAEESKKESA
VNQLTTCLDYLSSTPQHPIFPDVATAASRLRQAIPSLSEEF SYILAQRITQPNQGGVRWSWD
AIIRTRSILGLNNLPGGRSQYLEMLKSIQVPTTLVYGDSSKLNRPEDLQQQKMTMTQAKRVF
LSGGHNLHIDAAAALASLIITS

[0095] The amino acid sequence of a PKS (pentene ORF1) capable of producing pentene has the following amino acid sequence (SEQ ID NO:8):

MRAPYGNRQVNRRFLREFRAKRPHCVSPLHFLAEFSESRTTGSAGVTAPI DRPGVSMAPKS
GAQRSSDIAVVGMSCLPGAPGIDFWHLLTTGGS AIERRADGTWRGSLDGAADFDAFFDM
TPRQAAAADPQQRLMLELGWTALENAGIVPGSLAGTDTGVFVGIAADDYAALLHRSATPVSG
HTATGLSRGMAANRLSYLLGLRGPSLAVDSAQSSSLVAVHLACESLRRGESDLAIVGGVSLI
LAEDSTAGMELMGALSPDGRCHTFDARANGYVRGEGGACVVLKPLERALADGDRVHCVVRGS
AVNNDGGGSTLTTPHREAQA AVLRAAYERAGVGPDQVSYVELHGTGTPVGDPVEAAALGAVL
GTAHGRNAPLSVGSVKTNVGHLEAAAGLVGFVKAALCVREGVPPSLNHATPNPAI PMDRLN
LRVPTRLEPWPHPDDRATGRLRLAGVSSFGMGGTNAHVVEEAPLPEAGEPVGAGVPLAVVP
VVVSGRSAGAVAE LASRLNESVRSDRLVDVGLSSVVSRSVFEHRSVVLGDSAE LSAGLDAL
AADGVSPVLVSGVASVGGGRSVFVFPAGGVKWAGMALGLWAESAVFAESMARCEAA FAGLVE
WRLADV LGDGAALEREDVVQPASFVMSLAALWRS LGVVPDAVVGHSQGEIAAAVVAGGLS
LEDGARVVVLRARVAEEVL SGGGIASVRLSRAEVEERLAGGGGGLSVAVVNAPSSTVVAGEL
GDLDRFVAACEAEGVRARRLEFGYASHSRFVEPVRRERLLEGLADV RPVRGRIPFYSTVEAAE
FDTAGLDAEYWFGNLRRPVR FQETVERLLADGFRVFVECGAHPVLTGAVQETAETAGREICS
VGSLRRDEGGLRRFLTSAAEAFVQGEVSWPVLF DGTGARTVDLPTYPFQRRHHWAPDGSAS
AAPTRDIRPDETA AVPADTMDLAGQLRADVASLP TTEQIARLLDQVRDGVATVLGLDARDEV
RAEATFKELGVESLTGVELKNHLRARTGLHVPTSLIYDCPTPLAAAHYLRDELLGRPAEQAV
VPAGIPVDEPIAIVGMGCRLPGGVSSPEGLWDLVASGVDAVSPFPTDRGWDVGGLFDPEPGV
PGRSYVREGGFLHEAGEFDAGFFGISPREALAMPQQRL LLET SWEALERAGIDPHTLRGSR
TGVYAGVMAQEYGPRLHEGADGYEGYLLTGSSSSSVASGRISYVLGLEGP AVTVDTACSSSLV
ALHLAVRALRSGECDLALAGGATVMAEPGMFVEFSRQRGLSAHGRCKAYS SDSADGTGWAEGA
GVLLVERLSDAVRHGRRLAVVRGSAVNQDGASNGLTAPNGRSQSRLIRQALADARLG VADV
DVVEGHGTGTRLGDPIEAQALLATY GQRDAGRPLRLGSLKSNVGHTQAAAGVAGVIKMVMAM
RHGVLPKTLHVDEPTAEVDWSAGAVSLLREQEAWPRGERVRRAGVSSFGVSGTNAHVILEQP
PGVPSQSAGPGSGSVVDVPVVPWMVSGKTPEALSAQATALMTYLDERP DVSSLDVGYSALT
RSALDERAVVLGSDRETLLCGVKALSAGHEASGLVTG SVGAGGRIGFVFSGQGGQWLGMGRG
LYRAFPVFAAAFDEACAELDAHLGQEIGVREVVS GSDAQLLDRTLWAQSGLFALQVGLLKL
DSWGV RPSVVLGH SVGELAAAFAGVVSLSGAARLVAGRARLMQALPSGGGMLAVPAGEELL
WSLLADQGDRVGIAAVNAAGSVVLSGDRDVLDDLAGRLDGQGIRSRWL RVSHAFHSYRMDPM
LAEFAELARTVDYRRCEVP IIVSTLTGDLDDAGRMSGPDYWVRQVREPVRFADGVQALVEHDV
ATVVELGPDGALSALI QECVAASDHAGRLSAVPAMRRNQDEAQKVM TALAHVHVRRGGAVDWR
SFFAGTGAKQIELPTYAFQRQRYWLVP SDSGDVTGAGLAGAEHPLLGA VVPVAGGDEVLLTG
RISVRTHPWLAEHRVLGEVIVAGTALLEIALHAGERLGCERVEELTLEAPLVLP ERGAIQVQ
LRVGAPENSGRRPMALYSRPEGAAEHDWTRHATGRLAPGRGEAAGDLADWPAPGALPVDLDE
FYRDLAELGLEYGPIFQGLKAAWRQGDEVYAE AALPGTEDSGFGVHPALLDAALHATAVRDM
DDARLPFQWEGVSLHAKAAPALRVRVVPAGDDAKSLLVCDGTGRPVISVDRLV LRSAAARRT
GARRQAHQARLYRLSWPTVQLPTS AQPPSCVLLGTSEVSADIQVYPDLRSLTAALDAGAEPP
GVVIAPTPPGGGRTADVRETT RHALDLVQGWLS DQRLNESRLLLVTQGAVAVEPGE PVTDLA

QAALWGLLRSTQTEHPDRFVLVDVPEPAQLLPALPGVLACGEPQLALRRGGAHAPRLAGLGS
 DDVLPVPDGTGWRLEATRPGSLDGLALVDEPTATAPLGDEVRIAMRAAGVNFRDALIALGM
 YPGVASLGSEGAGVVETGPGVTGLAPGDRVMGMI PKAFGLAVADHRMVTRIPAGWSFARA
 ASVPIVFLTAYYALVDLAGLRPGESLLVHSAAGGVGMAAIQLARHLGAEVYATASEDKWQAV
 ELSREHLASSRTCDFEQQFLGATGGRGVDDVNLNLSLAGEFADASLRMLPRGGRFLELGKTDVR
 DPVEVADAHPGVSYQAFDTVEAGPQRIGEMLHELVELFEGRVLEPLPVTAWDVRQAPEALRH
 LSQARHVGKLVLTMPVWDAAGTVLVTGGTGALGAEVARHLVIERGVRNLVLVSRRGPAASG
 AAELVAQLTAYGAEVSLQACDVADRETAKVLASIPDEHPLTAVVHAAGVLDDGVSESLTVE
 RLDQVLRPKVDGARNLLELIDPDVALVLFSSVSGVLGSGGQGNVYAAANSFLDALAQQRQSRG
 LPTRSLAWGPWAEHGMASLTREAEQDRLARSGLLPISTEEGLSQFDAACGGAHTVVAPVRFS
 RLSDGNAIKFSVLQGLVGPHRVNKAATADDAESLRKRLGRLPDAEQHRILLDLVRMHVA AVL
 GFAGSQEITADGTFKVLGFDSTLVVELNRNIRINGATGLRLPATLVFNYPPTDALAAHLVTALS
 ADRLAGTFEELDRWAANLPTLARDEATRAQITTRLQAILQSLADVSGGTGGGSVPDRLRSAT
 DDELFQLLDNDLELP

[0096] The amino acid sequence of a PKS (pentene ORF2) capable of producing pentene has the following amino acid sequence (SEQ ID NO:9):

MSNEEKLREYLRRALVDLHQARERLHEAESGEREPIAIVAMGCRYPGGVQDPEGLWKLVASG
 GDAIGEFPADRGWHLDELYDPDPDQPGTCYTRHGGFLHDAGEFDAGFFDISPREALAMPQQ
 RLLLEISWETVESAGMDPRSLRGSRTGVFAGLMYEGYDTGAHRAGEGVEGYLGTGNAGSVAS
 GRVAYAFGFEGPAVTVDTACSSSLVALHLACQSLRQGECDLALAGGVTVMTSTPERFVEFSRQ
 RGLAPDGRCKSFAAADGTGWGEGAGLVLLERLSDARRNGHRVLAVVRGSAVNQDGASNGLT
 APNGLAQERVIQQVLTSAAGLSASDVDAVEAHGTGTRLGDPIEAQALIAAYGQDRDRDRPLWL
 GSVKSNIGHTQAAAGVAGVIKVMAMRHGELPRTLHVDEPNSHVDWSAGAVRLLTENIRWPG
 TGTRRAGVSSFGVSGTNAHVIVGDYAQQKSPLAPPATQDRPWHLTLTSAKNAQALNALQKSY
 GDYLAQHPSVDPRDLCLSANTGRSPLKERRFFVFKQVADLQQTNLNQDFLAQPRLSPPAKIAF
 LFTGQGSQYYGMGQQLYQTSFVFRQVLDECDRLWQTSPEAPALTDLLYGNHNPDLVHETVY
 TQPLLFAVEYAIQWLWSGVTPDFCMGHSVGEYVAACLAGVFSLADGMKLITARGKLMHAL
 PSNGSMAAVFADKTVIKPYLSEHLTVGAENGSHLVLSGKTPCLEASIHKLQSQGIKTKPLKV
 SHAFHSPLMAPMLAEFREIAEQITFHPPRIPLISNVTGGQIEAEIAQADYWVKHVSQPVKFV
 QSIQTLAQAGVNVYLEIGVKPVLLSMGRHCLAEQEAVWLPSLRPHSEPWPEILTSLGKLYEQ
 GLNIDWQTV EAGDRRRKLILPTYPFQRQRYWFNQGSWQTVETESVNP GPDDLNDWLYQVAWT
 PLDTLPPEPSAKLWLILGDRHDHQPIEAQFKNAQRVYLGQSNHFPTNAPWEVSADALDNL
 FTHVGSQNLGILYLCPPGEDPEDLDEIQKQTSGFALQLIQTLTYQQKIAPCWVFWTHQSQRV
 LETDAVTGFAQGGWLGLAQAI ALEHP ELWGGI IDVDDSLPNFAQICQQRQVQQLAVRHQKLY
 GAQLKKQPSLPQKNLQIQPQQT YLVTGGLGAIGRKIAQWLAAAGAEKVILVSRRAPAADQQT
 LPTNAV VYPCDLADAAQVAKLFQTYPHIKGIFHAAGTLADGLLQQQTWQKFQTVAAAKMKGT
 WHLHRHSQKLDLDFVLFSSVAGVLGSPGQGNVYAAANRGMAAIAQYRQAQGLPALAIHWGPW
 AEGGMANSLSNQNLAWLPPPQGLTILEKVLGAQGE MGVPKPDWQNLAKQFPEFAKTHYFAAV
 IPSAEAVPPTASIFDKLINLEASQRADYLLDYLRSSVAQILKLEIEQIQSHDSLLDLGMDSL
 MIMEAIASLKQDLQLMLYPREIYERPRLDVLTAYLAAEFTKAHDSEAATAAAAI PSQSLSVK
 TKKQWQKPDHKNPNPIAFILSSPRSGSTLLRVMLAGHPGLYSPPELHLLPFETMGDRHQELG
 LSHLGEGLQRALMDLENLTPEASQAKVNQWVKANTPIADIYAYLQRQAEQRLLIDKSPSYGS
 DRHILDHSEILFDQAKYIHLVRHPYAVIESFTRLRMDKLLGAEQQNPNYALAESIWRTSNRNI
 LDLGRTVGADRYLQVIYEDLVRDPRKVLTNICDFLGVD FDEALLNPYSGDRLTDGLHQQSMG
 VGDPNFLQHKTIDPALADKWSITLPAALQLDTIQLAETFAYDLPQEPQLTPQTQSLPSMVE
 RFVTVRGLETCLCEWGRHQPLVLLLHGILEQGASWQLIAPQLAAQGYWVVPDLRGHGKSA
 HAQSYSMLDFLADVDALAKQLGDRPFTLVGHSMGSIIGAMYAGIRQTQVEKLILVETIVPND
 IDDAETGNHLTTHLDYLAAPPQHPIFPSLEVAARRLRQATPQLPKDLSAFLTQRSTKSVEKG

VQWRWDAFLRTRAGIEFNGISRRRYLALLKDIQAPITLIYGDQSEFNRPADLQAIQAALPQA
QRLTVAGGHNLFENPQAIAQIVYQQLQTPVPKTQ

Nucleic Acids encoding the PKS

[0097] The present invention provides recombinant nucleic acids that encode the PKSs of the invention. The recombinant nucleic acids include double-stranded and single-stranded DNAs and RNA derived therefrom. The recombinant nucleic acids of the invention include those that encode an open reading frame (ORF) of a PKS of the present invention. The recombinant nucleic acids of the invention also include, in a variety of embodiments, promoter sequences for transcribing the ORF in a suitable host cell. The recombinant nucleic acids of the invention include, in some embodiments, sequences sufficient for having the recombinant nucleic acid stably replicate in a host cell, such as sequences that provide a replicon capable of stable maintenance in a host cell or sequences that direct homologous recombination of the nucleic acid into a chromosome of the host cell. In some embodiments, the nucleic acid is a plasmid, including but not limited to plasmids containing an origin of replication. The present invention also provides vectors, such as expression vectors, comprising another recombinant nucleic acid of the present invention. The present invention provides host cell comprising any of the recombinant nucleic acids and/or capable of expressing a PKS of the present invention. In some embodiments, the host cell, when cultured under suitable conditions, is capable of producing an α -olefin of the invention.

[0098] It will be apparent to one of skill in the art that a variety of recombinant vectors can be utilized in the practice of the invention. As used herein, "vector" refers to polynucleotide elements that are used to introduce recombinant nucleic acid into cells for either expression or replication (or both). Selection and use of such vectors generally is routine in the art. An "expression vector" is a recombinant nucleic acid capable of expressing (producing proteins encoded by) DNA coding sequences (and corresponding mRNA) that are operatively linked with regulatory sequences, such as promoters. Thus, an expression vector refers to a recombinant DNA or RNA construct, such as a plasmid, a phage, a recombinant virus, or other vector that, upon introduction into an appropriate host cell that when cultured under appropriate conditions, results in expression of the DNA coding sequence. Appropriate expression vector elements suitable for use in accordance with the present invention are well known to those of skill in the art and include those that are replicable in eukaryotic cells and/or prokaryotic cells and those that remain episomal as well as those that integrate into the

host cell genome.

[0099] The vectors of the invention include those chosen to contain control sequences operably linked to the resulting coding sequences in a manner that expression of the coding sequences may be effected in an appropriate host. Suitable control sequences include those that function in eukaryotic and prokaryotic host cells. If the cloning vectors employed to obtain PKS genes lack control sequences for expression operably linked to the PKS-encoding nucleotide sequences, the nucleotide sequences are inserted into appropriate expression vectors. This can be done individually, or using a pool of isolated encoding nucleotide sequences, which can be inserted into "host" vectors, the resulting vectors transformed or transfected into host cells, and the resulting cells plated out into individual colonies. Suitable control sequences for single cell cultures of various types of organisms are well known in the art. Control systems for expression in yeast are widely available and are routinely used. Control elements include promoters, optionally containing operator sequences, and other elements depending on the nature of the host, such as ribosome binding sites. Particularly useful promoters for prokaryotic hosts include those from PKS gene clusters that result in the production of polyketides as secondary metabolites, including those from Type I or aromatic (Type II) PKS gene clusters. Examples are act promoters, tcm promoters, spiramycin promoters, and the like. However, other bacterial promoters, such as those derived from the genes encoding sugar metabolizing enzymes, such as those that metabolize galactose, lactose (lac) and maltose, are also useful. Additional examples include promoters derived from the genes encoding biosynthetic enzymes such as those that encode the enzymes for tryptophan (trp) biosynthesis, the β -lactamase (bla) gene promoter, bacteriophage lambda PL promoter, and the T5 promoter. In addition, synthetic promoters, such as the tac promoter (U.S. Pat. No. 4,551,433; incorporated herein by reference), can be used to construct an expression vector of the invention.

[00100] As noted, particularly useful control sequences are those which themselves, or with suitable regulatory systems, activate expression during transition from growth to stationary phase in the vegetative mycelium. Illustrative control sequences, vectors, and host cells of these types include the modified *Streptomyces coelicolor* CH999 and vectors described in PCT publication No. WO 96/40968 and similar strains of *Streptomyces lividans*. See U.S. Pat. Nos. 5,672,491; 5,830,750; 5,843,718; and 6,177,262, each of which is hereby incorporated by reference. Other regulatory sequences may also be desirable; these include

those that allow for regulation of expression of the PKS sequences relative to the growth of the host cell. Regulatory sequences are known to those of skill in the art, and examples include those which cause the expression of a gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Other types of regulatory elements may also be present in the vector, for example, enhancer sequences.

[00101] Selectable markers can also be included in the recombinant expression vectors of the invention. A variety of markers are known that are useful in selecting for transformed cell lines; these generally are any gene whose expression confers a selectable phenotype on transformed cells when the cells are grown in an appropriate selective medium. Such markers include, for example, genes that confer antibiotic resistance or sensitivity to a host cell.

[00102] The various PKS nucleotide sequences, or a mixture of such sequences, can be cloned into one or more recombinant vectors as individual cassettes, with separate control elements or under the control of a single promoter. The PKS encoding subunits or components can include flanking restriction sites to allow for the easy deletion and insertion of other PKS encoding subunits. The design of such restriction sites is known to those of skill in the art and can be accomplished using the techniques described in the scientific literature, such as site-directed mutagenesis and PCR. Methods for introducing the recombinant vectors of the present invention into suitable hosts are known to those of skill in the art and include the use of CaCl_2 or other agents, such as other divalent cations, lipofection, DMSO, protoplast transformation, conjugation, and electroporation.

Host Cells comprising the PKS

[00103] The present invention provides host cells comprising the recombinant nucleic acid and/or PKS of the present invention. In many embodiments, the host cell, when cultured, is capable of producing an α -olefin. The host cell can be a eukaryotic or a prokaryotic cell. Suitable eukaryotic cells include yeast cells, such as from the genus *Saccharomyces*, *Candida*, or *Schizosaccharomyces*. A suitable species from the genus *Saccharomyces* is *Saccharomyces cerevisiae*. A suitable species from the genus *Schizosaccharomyces* is *Schizosaccharomyces pombe*. Suitable prokaryotic cells include, but are not limited to, the gram negative *Escherichia coli* and the gram positive *Streptomyces* species, such as *S. coelicolor* and *S. lividans*.

[00104] The PKSs of the invention can be in a host cell, and can isolated and purified.

The PKS can synthesize the α -olefin *in vivo* (in a host cell) or *in vitro* (in a cell extract or where all necessary chemical components or starting materials are provided). The present invention provides methods of producing the α -olefin using any of these *in vivo* or *in vitro* means.

[00105] In some embodiments of the invention, the host cell comprises a PKS which produces butadiene comprising a loading module comprising an acrylyl-ACP starter, such as a DEBS propionyl-CoA specific loading domain which is modified to accept acrylyl-CoA, and one or more nucleic acids encoding and capable of expressing biosynthetic enzymes for synthesizing acrylyl-CoA from propionate (see Figure 8). An example of a set of such enzymes is propionyl-CoA ligase (synthetase) (EC 6.2.1.17) and acyl-CoA dehydrogenase (mammalian) (EC 1.3.99.3), and functional variants thereof. In some embodiments, the host cell comprises a nucleic acid encoding and capable of expressing an enzyme, or functional variant thereof, capable of converting propionate into propionyl-CoA, and an enzyme, or functional variant thereof, capable of converting propionyl-CoA into acrylyl-CoA. An enzyme capable of converting propionate into propionyl-CoA is the propionyl-CoA ligase encoded by the *prpE* gene of *Salmonella typhimurium*. An enzyme capable of converting propionyl-CoA into acrylyl-CoA is the mammalian acyl-CoA dehydrogenase. A host cell comprising this system is provided with propionate, such as by exogenously feeding propionate to the host cell, or produces propionate endogenously.

[00106] The amino acid sequence of the propionyl-CoA ligase encoded by the *prpE* gene in *Salmonella typhimurium* (GenBank accession no. NP_459366) comprises:

```

1 msfsefyqrs inepeafwae qarridwrqp ftqtlldhsrp pfarwfcggt tnlchnavdr
61 wrdkqpeala liavssetde ertftfsqlh devnivaaml lslgvqrgdr vlvympmie
121 aqitllacar igaihsvvfg gfashsvaar iddarpaliv sadagarggk ilpykklldd
181 aiaqaqhqp hvllvdrcla kmawvdgrdl dfatlrqghl gasvpvawle snetscilyt
241 sgttgkpkgv qrdvgyava latsmdtifg gkaggvffca sdigwvghs yivyapllag
301 mativyeqlp typdcgvwwk ivekyqvnrm fsaptairvl kkfptaqrn hdlsslealy
361 lagepldept aswvtetlgv pvidnywqte sgwpimalar alddrpsrlg spgvpmgyn
421 vqllnevtge pcginekgml viegplppgc iqtigddar fvktywslfn rvyatfdwg
481 irdaegyyfi lgrtddvini aghrlgtrei eesissypnv aevavvgikd alkgqvavaf
541 vipkqsdsla dreaardeen aimalvdnqi ghfgrpahvw fvsqpktrs gkmlrrtiqa
601 egrdpdgl ttiddpaslq qirqaiee (SEQ ID NO:13)

```

[00107] The amino acid sequence of the acyl-CoA dehydrogenase of *Mus musculus* (GenBank accession no. Q07417) comprises:

```

1 maaallalar gplrralgvr dwrrlhtvyq svelpethqm lrqtcrdfae kelvpiaaql
61 drehlftaq vkkmgelgl amdvpelsg aglgylaysi aleeisraca stgvimsvnn
121 slylgpilkf gsaqqkqwi tpftngdkig cfalsepgng sdagaastta reegdswvln

```

```

181 gtkawitnsw easatvvfas tdrsrqnkgi saflvpmpptp gltlgkkekdk lgirasstan
241 lifedcripk enllgepgmg fkiamqtdm grigiasqal giaqasldca vkyaenrnaf
301 gapltrlqni qfkladmala lesarlltwr aamlkdnkpk ftkesamrkl aaseaatais
361 aiqilgsm gyvtempaer yyrdaritei yegtseiqlr viaghllrsy rs
      (SEQ ID NO:14)

```

[00108] In some embodiments, the host cell of the invention comprises a PKS which produces butadiene and comprises a loading module comprising an acrylyl-ACP starter, such as a DEBS propionyl-CoA specific loading domain which is modified to accept acrylyl-CoA, and one or more nucleic acids encoding and capable of expressing biosynthetic enzymes for synthesizing acrylyl-CoA from pyruvate (see Figure 9). An example of a set of such enzymes is lactate dehydrogenase (EC 6.2.1.17), lactate CoA transferase (EC 2.8.3.1), propionyl-CoA ligase (synthetase) (EC 6.2.1.17), and lactoyl-CoA dehydratase (EC 4.2.164), or functional variants thereof. In some embodiments, the host cell comprises one or more nucleic acids encoding and capable of expressing these four enzymes. A host cell comprising this system is provided with propionate and a suitable organic molecule that the host cell can directly or indirectly convert into a pyruvate, and these compounds are either exogenously fed to or produced by the host cell. For example, if the host cell is *E. coli*, a suitable organic molecule is glucose.

[00109] The amino acid sequence of the lactate dehydrogenase encoded by the *ldhA* gene of *E. coli* (GenBank accession no. CAQ31881) comprises:

```

1 taktangcea vcifvnddgs rpvleelkkh gvkyialrca gfnvdlldaa kelglkvrv
61 paydpeavae haigmmmtln rrihrayqrt rdanfslegl tgftmygkta gvigtgkigv
121 amlrilkgfg mrlldfpyp saaalelgve yvdlptlfse sdvislhcp1 tpenyhllne
181 aafdqmknvg mivntsergal idsqaaieal knqkigslgm dvyenerdlf fedksndviq
241 ddvfrrlsac hnvlfthgha fltaealtsi sqttlqnlsn lekgetcpne lv
      (SEQ ID NO:15)

```

[00110] The amino acid sequence of the lactate CoA transferase encoded by the *pct* gene of *Clostridium proponicum* (GenBank accession no. CAB77207) comprises:

```

1 mrkvpiitad eaaklikdgd tvttsgfvgn aipealdrav ekrfletgep knityvycgs
61 qgnrdgrgae hfaheglkr yiaghwatvp algkmamenk meaynvsqga lchlfrdias
121 hkpgvftkvg igtfdprng ggkvnditke divelveikg qeylfypafp ihvalirgty
181 adesgnitfe kevaplegts vcqavknsgg ivvvqvervv kagtdprhv kvpgiyvdyv
241 vvadpedhqq sldceydpal sgehrpevv geplplsakk vigrrgaiel ekdvavnlgv
301 gapeyvasva deegivdfmt ltaesgaigg vpaggvrfga synadalidq gyqfdyydgg
361 gldlcylgla ecdekgninv srfgpriagc ggfinitqnt pkvffcgftt agglkvkied
421 gkviivqegk qkkflkaveq itfngdvala nkqqvtyite rcvfllkedg lhlseiapgi
481 qtqildvm dfapiidrda ngqiklmdaa lfaeglmglk emks
      (SEQ ID NO:16)

```

[00111] The amino acid sequence of the lactoyl-CoA dehydratase encoded by the *pct* gene of *Clostridium proponicum* (GenBank accession no. CAB77206) comprises:

```

1 efkiaivddd laqesrqirv dvldgeggpl yrmakawqqm ygcslatdtk kgrgrmlink
61 tiqtgadaiv vammkfc dpe ewdypvmyre feekgvkslm ievdqvsssf eqiktrlqsf
121 veml
      (SEQ ID NO:17)

```

[00112] In some embodiments, the host cell of the invention comprises a PKS which produces butadiene and comprises a loading module comprising an acrylyl-ACP starter, such as a DEBS proprionyl-CoA specific loading domain which is modified to accept acrylyl-CoA, and one or more nucleic acids encoding and capable of expressing biosynthetic enzymes for synthesizing acrylyl-CoA from acetyl CoA (see Figure 10). An example of a set of such enzymes is acetyl-CoA carboxylase (EC 6.4.1.2), malonyl-CoA reductase (EC 1.2.1.75), 3-hydroxypropionate:CoA ligase, and 3-hydroxypropionyl-CoA hydratase (EC 4.2.1.17), or functional variants thereof. In some embodiments, the host cell comprises one or more nucleic acids encoding and capable of expressing the four enzymes described. A host cell comprising this system can be engineered to produce increased titers of acetyl-CoA.

[00113] In some embodiments, the host cell of the invention comprises a PKS which produces butadiene and comprises a loading module comprising an acrylyl-ACP starter, such as a DEBS proprionyl-CoA specific loading domain which is modified to accept acrylyl-CoA, and one or more nucleic acids encoding and capable of expressing biosynthetic enzymes for synthesizing acrylyl-CoA from propionate (see Figure 15). An example of a set of such enzymes is propionyl-CoA ligase (synthetase) (EC 6.2.1.17), lactoyl-CoA dehydratase (EC 4.2.1.64), lactate dehydrogenase (EC 6.2.1.17), lactate CoA transferase (EC 2.8.3.1), or functional variants thereof. A host cell comprising this system is provided with propionate, either by exogenous feeding or by endogenous production.

[00114] In some embodiments, the host cell of the invention comprises a PKS which produces 3-methyl-2,4-pentadienoic acid and comprises the modules shown in Figure 12 including a loading module comprising an acrylyl-ACP starter, such as a DEBS proprionyl-CoA specific loading domain which is modified to accept acrylyl-CoA, and one or more nucleic acids encoding and capable of expressing biosynthetic enzymes for synthesizing acrylyl-CoA (see Figure 12). There are several methods for enabling a host cell to synthesize acrylyl-CoA described above. 3-methyl-2,4-pentadienoic acid can be enzymatically, catalytically, or pyrolytically converted to isoprene.

Methods of Producing α -olefins Using the PKS

[00115] The present invention provides a method of producing an α -olefin comprising: providing a host cell of the present invention, and culturing said host cell in a suitable culture medium such that the α -olefin is produced. The method can further comprise isolating said α -olefin from the host cell and/or the culture medium. The method can further comprise polymerizing the α -olefin to itself and/or any other suitable organic molecule(s), including but not limited to other compounds comprising a C-C double bond. A variety of methods for heterologous expression of PKS genes and host cells suitable for expression of these genes and production of polyketides are described, for example, in U.S. Pat. Nos. 5,843,718; 5,830,750 and 6,262,340; WO 01/31035, WO 01/27306, and WO 02/068613; and U.S. Patent Application Pub. Nos. 20020192767 and 20020045220; each of which is incorporated herein by reference.

[00116] The present invention provides for a composition comprising an α -olefin isolated from a host cell from which the α -olefin is produced, and trace residues and/or contaminants of the host cell. Such trace residues and/or contaminants include cellular material produced by the lysis of the host cell. The present invention also provides α -olefins in substantially pure form.

[00117] Certain α -olefins produced by the PKSs of the present invention can be used as fuels. In some embodiments of the invention, an α -olefin produced in accordance with the invention can be used as a “green” jet fuel. The α -olefin can be catalytically oligomerized, including but not limited to dimerized, and optionally purified. The resulting products can be dimerized again to yield a mixture of branched molecules that are then catalytically hydrogenated. For example, 1-butene produced by a PKS of the present invention can be catalytically dimerized and purified. The resulting octene products can be dimerized again to yield a mixture of branched C16 molecules that are then catalytically hydrogenated. Oligomers of butene have been validated by the US Navy as both jet and diesel fuel replacements (Harvey, 2011. Journal of Chemical Technology and Biotechnology 86(1): 2–9.). Additional benefit may come from making branched, or aromatic, α -olefins using the avermectin (or other) loading modules, as described herein.

[00118] Thus, among others, the present invention has one or more of the following advantages: (1) it reduces the dependence on oil for producing certain chemicals, and (2) it

serves as a means of capture and sequestration of carbon from the atmosphere.

[00119] The invention having been described, the following examples are offered to illustrate and not limit the subject invention.

Examples

[00120] , Constructs can be conveniently designed at the amino acid level and then, back translation and DNA synthesis, such as that offered commercially by service providers such as DNA 2.0, can be conducted to yield the desired nucleic acid, which may be optimized for expression in a particular host cell type. Subsequent plasmid assembly can be conducted using standard molecular biology techniques.

EXAMPLE 1

Production of 1-hexene using a PKS-based enzyme system

[00121] In one embodiment of the invention, PKS modules from three different organisms were used to construct a tri-ketide pathway designed for the production of 1-hexene. In this embodiment, the 1-hexene synthase consists of two ORFs. HexORF1 combines EryA1 loading module + KS1 and AT-ACP from IdmO, HexORF2 utilizes the KS domain from IdmP and AT-TE domains from CurM. In another embodiment of this invention the loop I region of the indanomycin sourced ACP in HexORF1, SSSAGIDPGRAFQDMGI, is swapped with ASAERV PADQAFAELGV, the segment of ACP directly following EryA1. Both of these designs were back translated using software designed to optimize expression in *E. coli*. The genes are synthesized and ligated into two pairs of compatible, *E. coli* expression vectors that are subsequently transformed into *E. coli* BAP-1. The amino acid sequences for HexORF1, HexORF1', and HexORF2 are provided as SEQ ID NOs: 2-4, respectively.

[00122] Experiments have been performed demonstrating *E. coli* BAP-1 utilizing exogenously added propionate. In both examples of 1-hexene production, overnight cultures of a pBbA7C-HexORF1' (or pBbA7C-HexORF1) + pBbS7k-HexORF2 cotransformed strain were grown from a single colony and used to inoculate (1% v/v) three 50-mL cultures of LB medium supplemented with 0.5% glucose and 10% glycerol in 250mL screw cap (unsealed) flask. Cultures were grown to an OD600 of 1.0 to 1.2, induced with 50uM IPTG and grown at (30°C) for an additional 3 hours. Then 100 mM propionate was supplemented to the

culture and a Teflon septum was used to seal the cap. The cultures were then grown at 20°C for 24 hours after which 1-butene was detectable in the headspace of the culture using solid phase micro extraction followed by GC-MS.

EXAMPLE 2

Production of butadiene using a PKS-based enzyme system

[00123] An example of a PKS system for producing butadiene is shown in Figures 6 and 7. The system built to produce 1-butene, described below, was fed acrylate to produce butadiene. Otherwise, all experimental details are as described as that for 1-butene production, above. While the limits of detection were inadequate to determine productivity in this instance, the invention provides several routes to increasing butadiene productivity, including: adding an acrylate specific CoA ligase to the host strain, adding an acrylate importer to the host strain, and/or utilizing a host strain less sensitive to acrylate toxicity.

EXAMPLE 3

Production of 1-butene using a PKS-based enzyme system

[00124] An illustrative PKS system for producing 1-butene was constructed using the AT-TE PKS domains from CurM and the loading module for propionyl-CoA + KS1 from EryA1. For in vivo 1-butene production, overnight cultures of *E. coli* BAP1 carrying pBbS7k-Butene (PKS protein sequence provided as SEQ ID NO:5) were grown from a single colony and used to inoculate (1% v/v) three 50-mL cultures of LB medium supplemented with 0.5% glucose and 10% glycerol in 250mL screw cap (unsealed) flask. Cultures were grown to an OD600 of 1.0 to 1.2, induced with 50uM IPTG and grown at (30°C) for an additional 3 hours. Then 100 mM propionate was supplemented to the culture and a Teflon septum was used to seal the cap. The cultures were then grown at 20°C for 24 hours after which 1-butene was detectable in the headspace of the culture using solid phase micro extraction followed by GC-MS.

EXAMPLE 4

Production of isoprene using a PKS-based enzyme system

[00125] An example of a PKS system for producing isoprene is shown in Figure 12.

EXAMPLE 5

Production of (*E*)-penta-1,3-diene using a PKS-based enzyme system

[00126] An example of a PKS system for producing (*E*)-penta-1,3-diene is shown in Figure 13.

EXAMPLE 6

Production of propene (propylene) using a PKS-based enzyme system

An illustrative propene synthase of the invention is a single enzyme consisting of the loading module + KS1 from the niddamycin PKS (Kakavas, 1997) fused to AT-TE domains from CurM (Chang, 2004; Gu, 2009) (the amino acid sequence for this construct is provided as SEQ ID NO:6). For in vivo propene production, overnight cultures of *E. coli* BAP1 carrying pBbS7k-propene were grown from a single colony and used to inoculate (1% v/v) three 50-mL cultures of LB medium supplemented with 0.5% glucose and 10% glycerol in 250mL screw cap (unsealed) flask. Cultures were grown to an OD600 of 1.0 to 1.2, induced with 50uM IPTG and grown at (30°C) for an additional 3 hours. Then a Teflon septum was used to seal the cap. The cultures were then grown at 20°C for 24 hours after which propene was detectable in the headspace of the culture using solid phase micro extraction followed by GC-MS.

EXAMPLE 7

Production of Styrene using a PKS-based enzyme system

An illustrative styrene synthase of the invention was constructed by fusing the ST and TE domains from CurM onto the loading and first extension modules from the soraphen PKS (Schupp, 1995; Wilkinson, 2001). The amino acid sequence for this construct is provided as SEQ ID NO:7. For styrene biosynthesis in the system illustrated, a pool of benzoyl-CoA is provided. To facilitate production of this essential precursor, the styrene synthase construct was coexpressed with an *E. coli* codon optimized gene encoding benzoate-CoA ligase, *badA*, from *Rhodopseudomonas palustris* (Egland, et al., J Bacteriol. 1995 Nov;177(22):6545-51.) and fed exogenous benzoate. For in vivo styrene production, overnight cultures of *E. coli* BAP1 carrying pBbS7k-SS1(SS1 encodes SEQ ID NO:7) were grown from a single colony and used to inoculate (1% v/v) three 50-mL cultures of LB medium supplemented with 0.5%

glucose and 10% glycerol in 250mL screw cap (unsealed) flask. Cultures were grown to an OD600 of 1.0 to 1.2, induced with 50uM IPTG and grown at (30°C) for an additional 3 hours. Then 100 mM benzoic acid was supplemented to the culture and a Teflon septum was used to seal the cap. The cultures were then grown at 20°C for 24 hours after which styrene was detectable in the headspace of the culture using solid phase micro extraction followed by GC-MS.

EXAMPLE 8

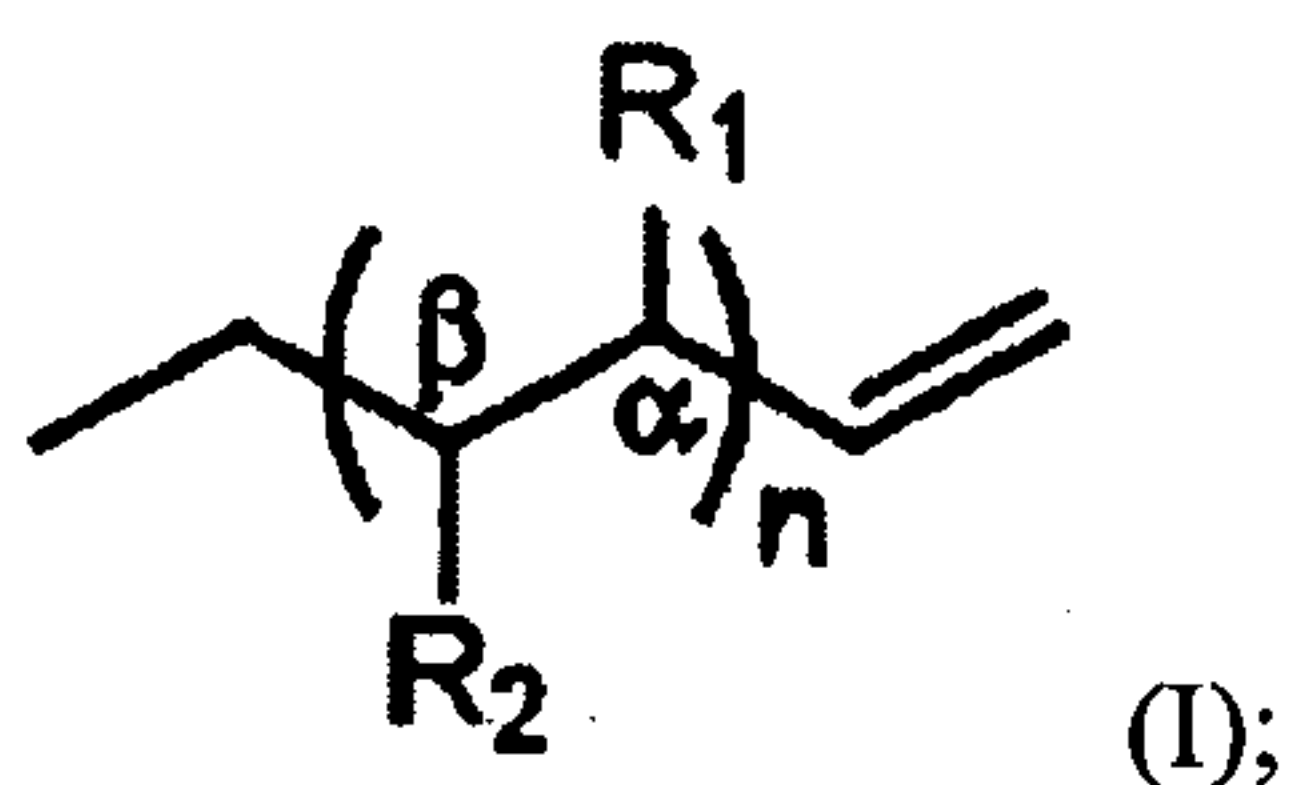
Production of pentene using a PKS-based enzyme system

[00127] An illustrative pentene synthase of the invention was designed as two ORFs. The first ORF is built using the chalcomycin PKS loading module +KS1 fused to spinosad AT-ACP PKS module two. The second ORF is the KS from spinosad M3 fused to the olefination module (Ols) from *Synechococcus* sp. PCC7002 (Mendez-Perez, et al., Appl Environ Microbiol. 2011 Jun;77(12):4264-7). The amino acid sequences for these chimeric proteins are provided as SEQ ID NO:8 and 9.

[00128] While the present invention has been described with reference to the specific embodiments thereof, it should be understood by those skilled in the art that various changes may be made and equivalents may be substituted without departing from the true spirit and scope of the invention. In addition, many modifications may be made to adapt a particular situation, material, composition of matter, process, process step or steps, to the objective, spirit and scope of the present invention. All such modifications are intended to be within the scope of the claims appended hereto.

What we claim is:

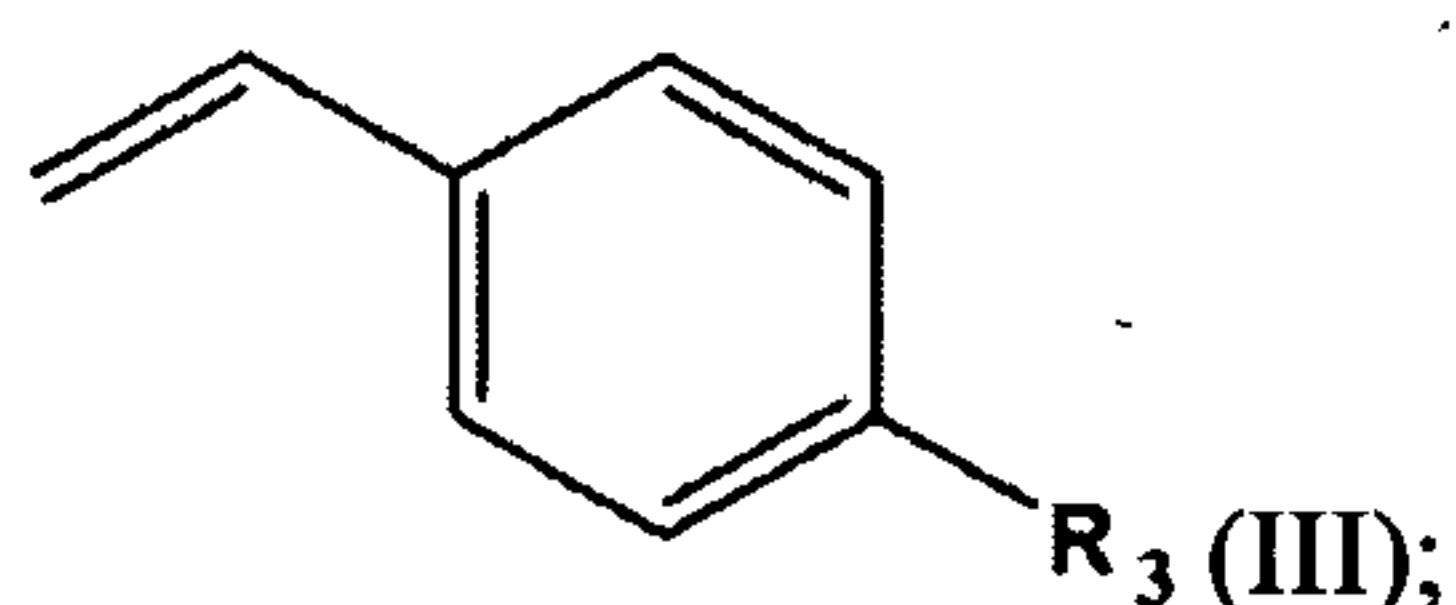
1. A non-naturally occurring polyketide synthase (PKS), or functional variant thereof, capable of synthesizing an α -olefin.
2. The PKS of claim 1, wherein the α -olefin is not a compound synthesized by a naturally occurring PKS.
3. The PKS of claim 1, wherein the PKS is a hybrid PKS comprising modules, domains, and/or portions thereof, or functional variant thereof, from two or more PKSs.
4. The PKS of claim 3, wherein the PKS comprises a terminal module comprising ST, and TE, or functional variant thereof, of *Lyngbya majuscula* CurM.
5. The PKS of claim 4, wherein the PKS comprises a terminal module comprising KR, ACP, ST, and TE, or functional variant thereof, of *Lyngbya majuscula* CurM.
6. The PKS of claim 5, wherein the PKS comprises a terminal module comprising AT, KR, ACP, ST, and TE, or functional variant thereof, of *Lyngbya majuscula* CurM.
7. The PKS of claim 3, wherein the PKS comprises at least one terminal module comprising a ST and a TE, or functional variant thereof, described in Tables 2-4.
8. The PKS of claim 3, wherein the PKS comprises a loading module which incorporates acrylyl-CoA.
9. The PKS of claim 3, wherein the PKS comprises the loading module is a DEBS propionyl-CoA specific loading domain modified to accept acrylyl-CoA.
10. The PKS of claim 9, wherein the PKS comprises a module in which a terminal carbon is incorporated as a methyl group which is later oxidized.
11. The PKS of claim 1, wherein the α -olefin has the following chemical structure:



wherein each R_1 is independently -H or -CH₃, each R_2 is independently -H or -OH, n is an integer from 1 to 10, $\alpha\beta$ is (1) a single bond or (2) a double bond, with the proviso that when an $\alpha\beta$ is a double bond then the corresponding R_2 is H.

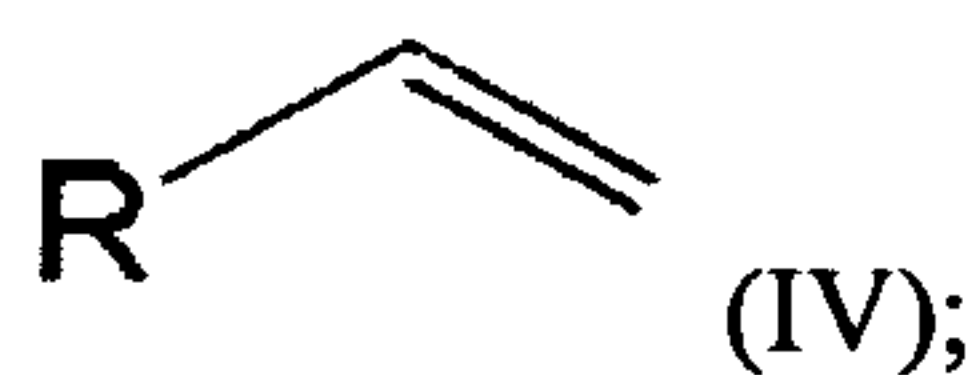
12. The PKS of claim 11, wherein the α -olefin is 1-hexene, 1-decene, or (*E*)-deca-1,5-diene.

13. The PKS of claim 1, wherein the α -olefin is an aromatic α -olefin having the following chemical structure:

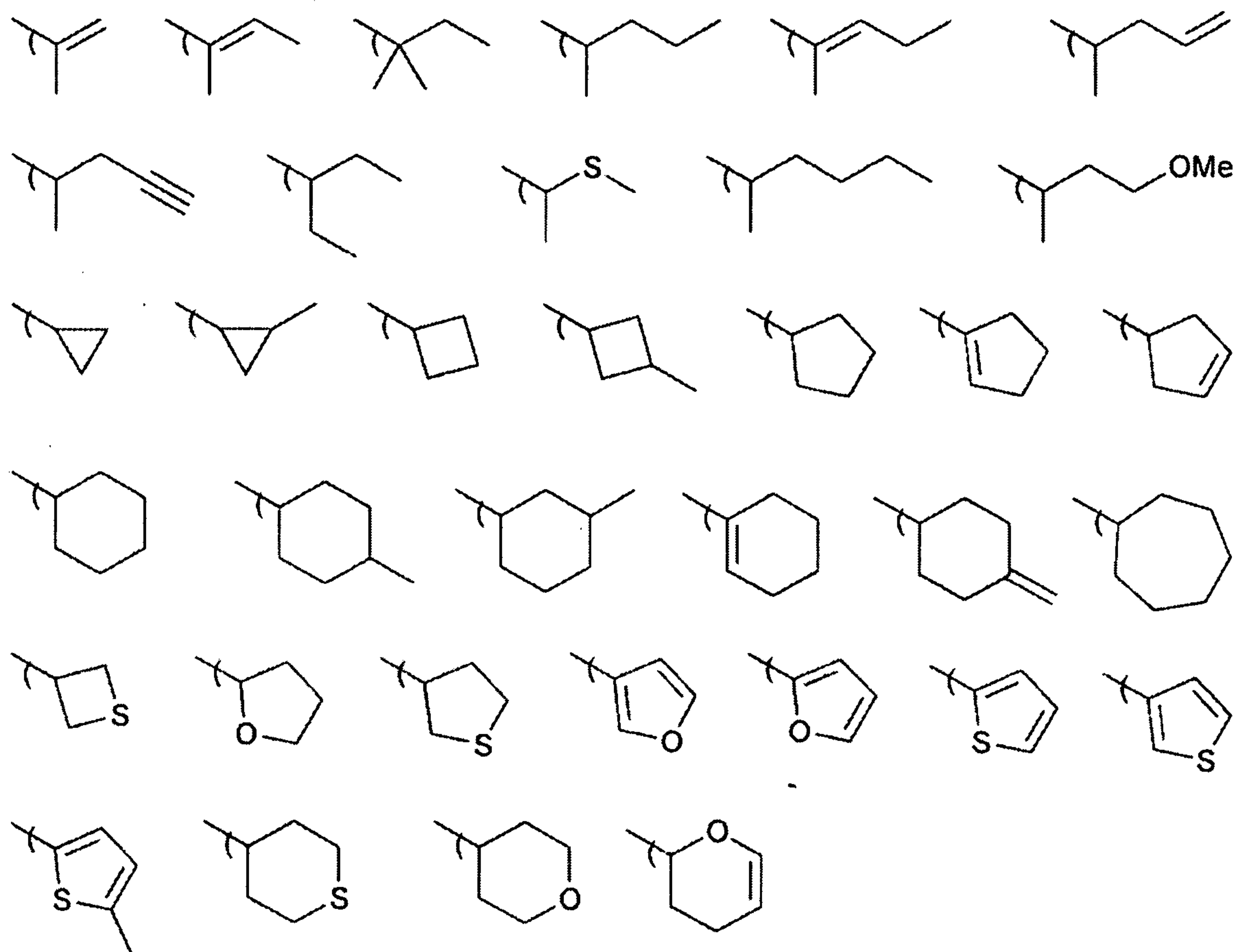


wherein R_3 is -H, -OH, -NH₃, or -NO₂.

14. The PKS of claim 1, wherein the α -olefin has the following chemical structure:



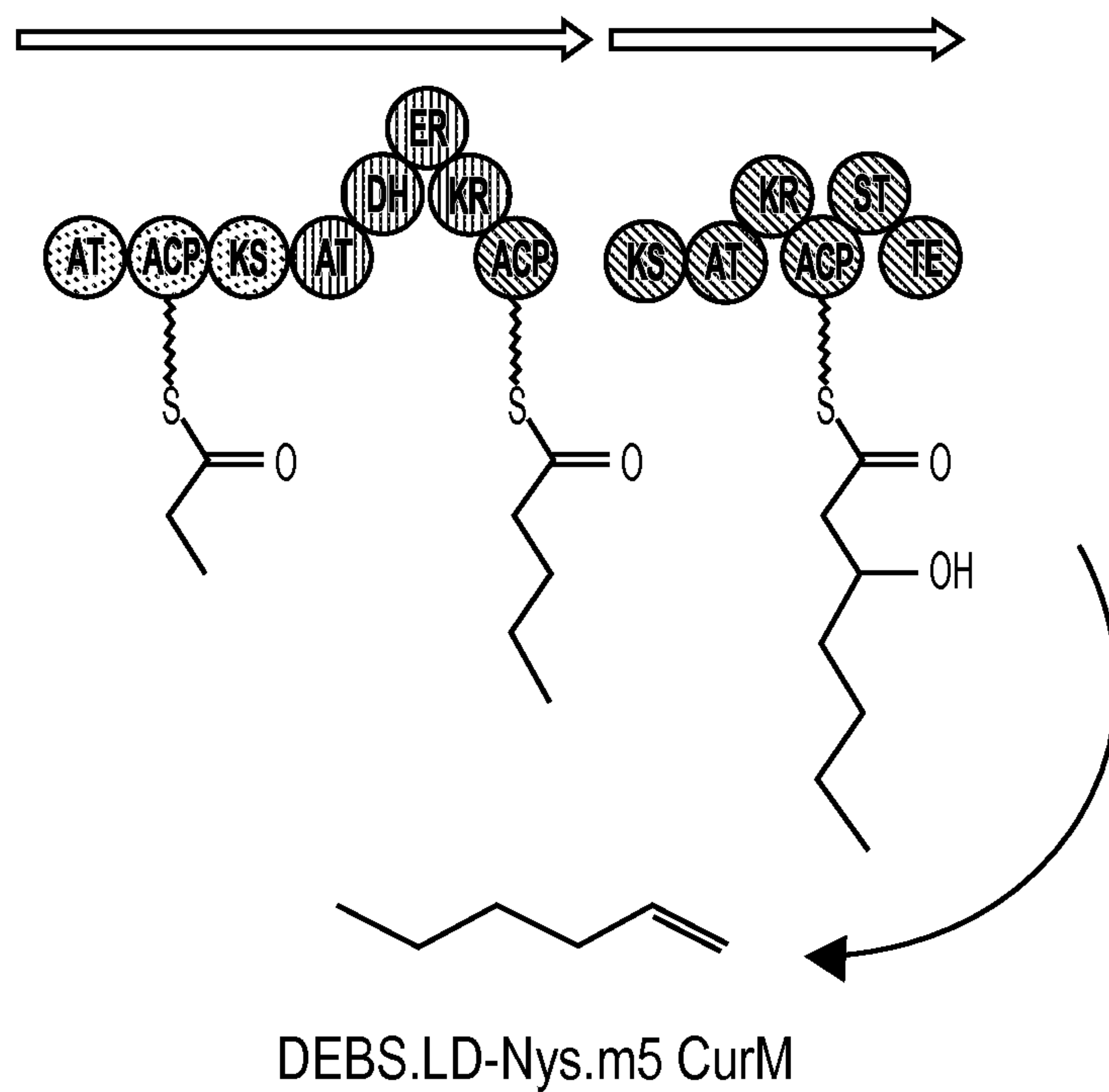
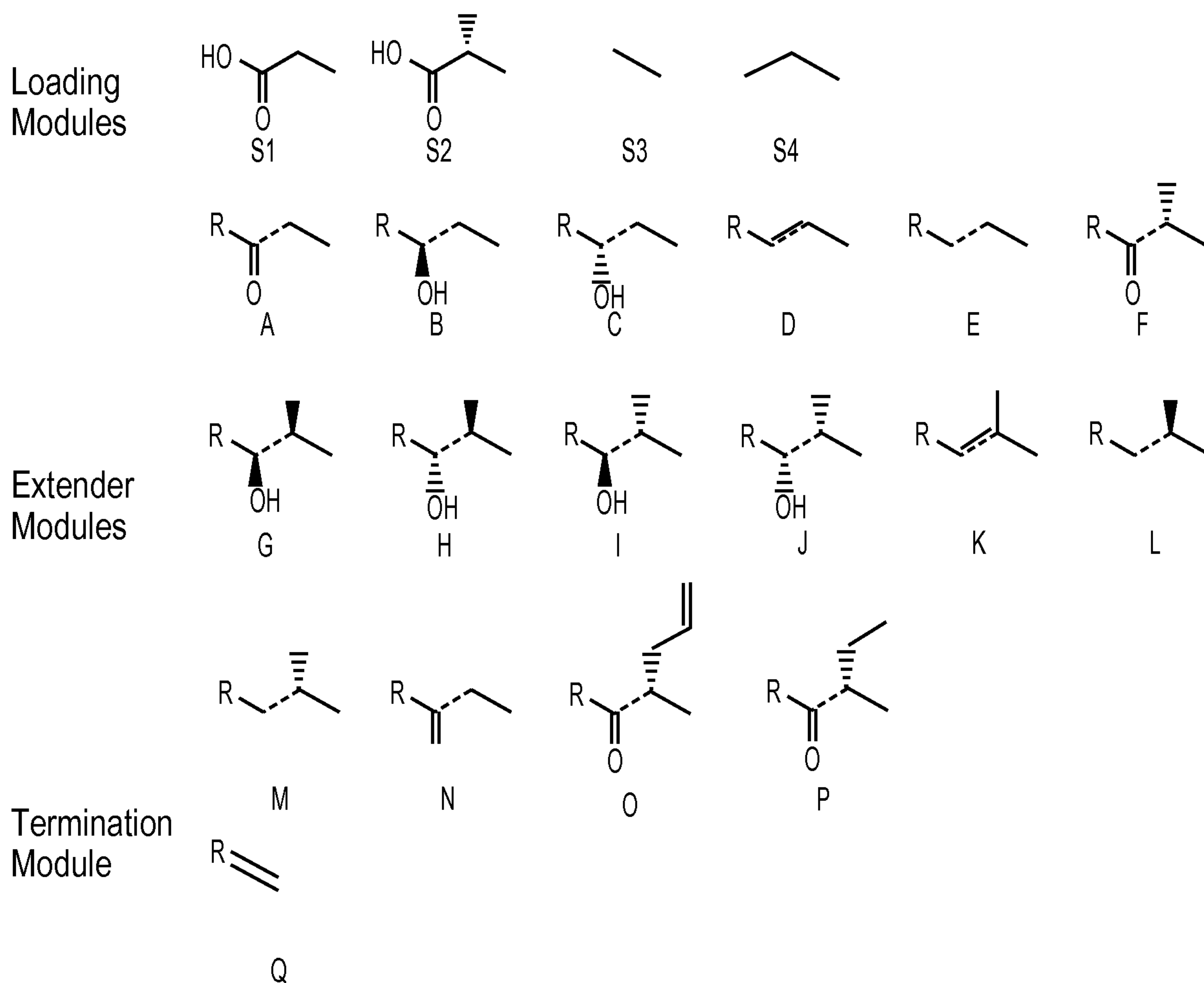
wherein R is one of the following structures



15. A recombinant nucleic acid encoding the polyketide synthase (PKS) of claim 1-13.
16. A replicon comprising the recombinant nucleic acid 14, wherein the replicon is capable of stable maintenance in a host cell.
17. The replicon of claim 16, wherein the replicon is a plasmid or vector.
18. The replicon of claim 17, wherein the vector is an expression vector.
19. A host cell comprising the recombinant nucleic acid of claim 15.
20. A host cell comprising the replicon of claim 16-18.
21. The host cell of claim 19 or 20, wherein the host cell when cultured produces the α -olefin.
22. The host cell of claim 21, wherein the host cell comprises a PKS which produces butadiene comprising a loading module comprising an acrylyl-ACP starter, and

- further comprises one or more nucleic acids encoding propionyl-CoA ligase (synthetase) (EC 6.2.1.17) and acyl-CoA dehydrogenase (mammalian) (EC 1.3.99.3), or functional variants thereof, wherein when cultured and optionally provided with propionate the PKS is capable of synthesizing acrylyl-CoA and producing butadiene.
23. The host cell of claim 21, wherein the host cell comprises a PKS which produces butadiene comprising a loading module comprising an acrylyl-ACP starter, and further comprises one or more nucleic acids encoding lactate dehydrogenase (EC 6.2.1.17), lactate CoA transferase (EC 2.8.3.1), propionyl-CoA ligase (synthetase) (EC 6.2.1.17), and lactoyl-CoA dehydratase (EC 4.2.164), or functional variants thereof, wherein when cultured and optionally provided with propionate and glucose the PKS is capable of synthesizing acrylyl-CoA and producing butadiene.
24. The host cell of claim 21, wherein the host cell comprises a PKS which produces butadiene comprising a loading module comprising an acrylyl-ACP starter, and further comprises one or more nucleic acids encoding acetyl-CoA carboxylase (EC 6.4.1.2), malonyl-CoA reductase (EC 1.2.1.-), 3-hydroxypropionate:CoA ligase, and 3-hydroxypropionyl-CoA hydratase (EC 4.2.1.17), or functional variants thereof, wherein when cultured the PKS is capable of synthesizing acrylyl-CoA and producing butadiene
25. The host cell of claim 21, wherein the host cell comprises a PKS which produces butadiene comprising a loading module comprising an acrylyl-ACP starter, and further comprises one or more nucleic acids encoding propionyl-CoA ligase (synthetase) (EC 6.2.1.17), the AccA/PccB gene products and lactoyl-CoA dehydratase (EC 4.2.164), and lactate dehydrogenase (EC 6.2.1.17), lactate CoA transferase (EC 2.8.3.1), propionyl-CoA ligase (synthetase) (EC 6.2.1.17), and lactoyl-CoA dehydratase (EC 4.2.164), or acyl-CoA dehydrogenase, or functional variants thereof, wherein when cultured and optionally provided with propionate the PKS is capable of synthesizing acrylyl-CoA and producing butadiene.
26. A method of producing an α -olefin, comprising: providing a host cell of any one of claims 21-25, and culturing said host cell in a suitable culture medium such that the α -olefin is produced.
27. The method of claim 26, further comprising isolating the α -olefin.

28. The method of claim 27, further comprising polymerizing the α -olefin to produce a polymer.
29. The method of any one of claims 26-28, wherein the α -olefin is a polyolefin.
30. The method of claim 29, wherein the polyolefin is butadiene, isoprene, or (*E*)-deca-1,5-diene.
31. A composition comprising an α -olefin isolated from the host cell of any one of claims 21-25, and trace residues and/or contaminants of the host cell.
32. The composition of claim 31, wherein the α -olefin is polyolefin.
33. The composition of claim 32, wherein the polyolefin is butadiene, isoprene, or (*E*)-deca-1,5-diene

**FIG. 1****FIG. 2**

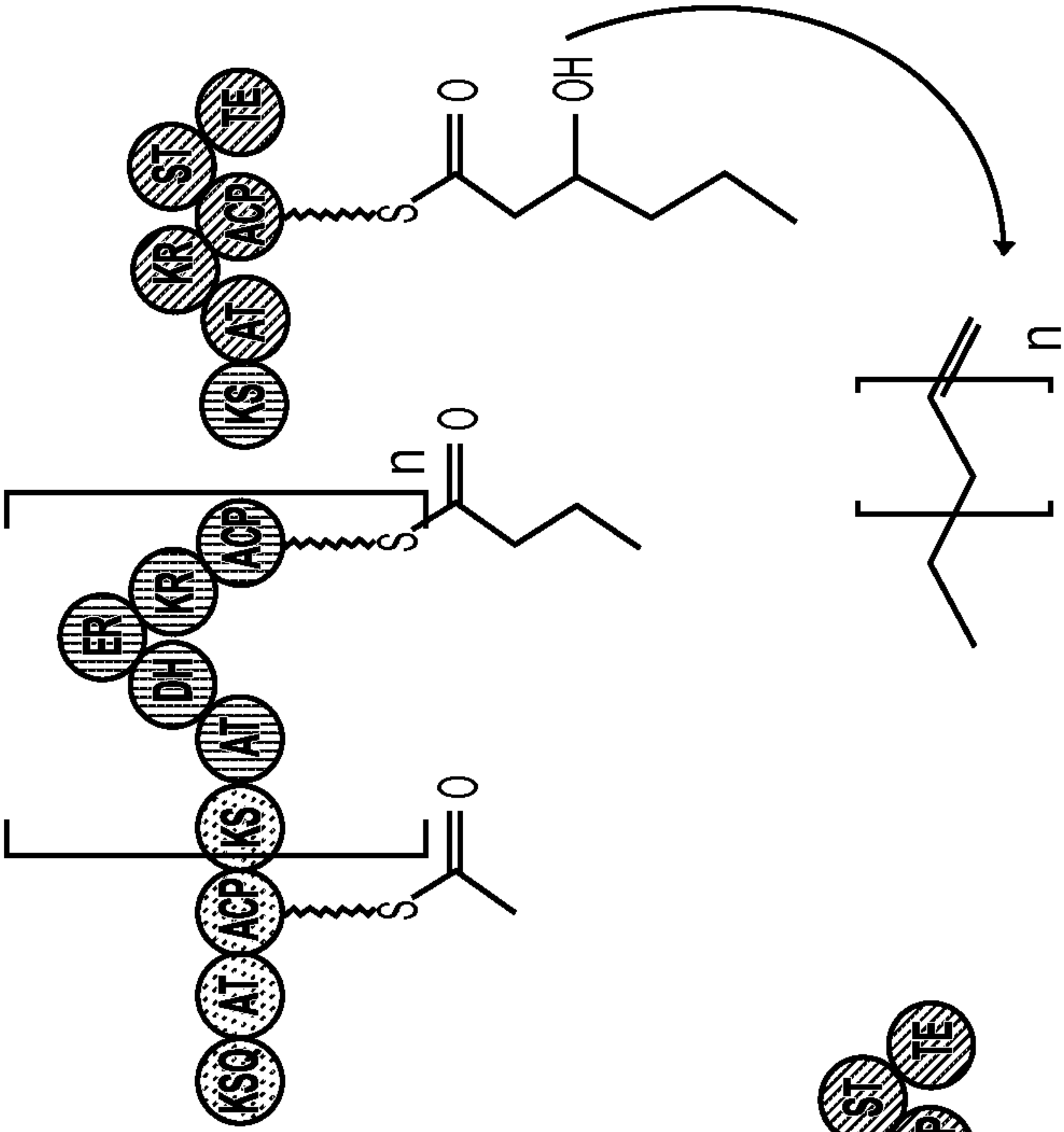


FIG. 3C

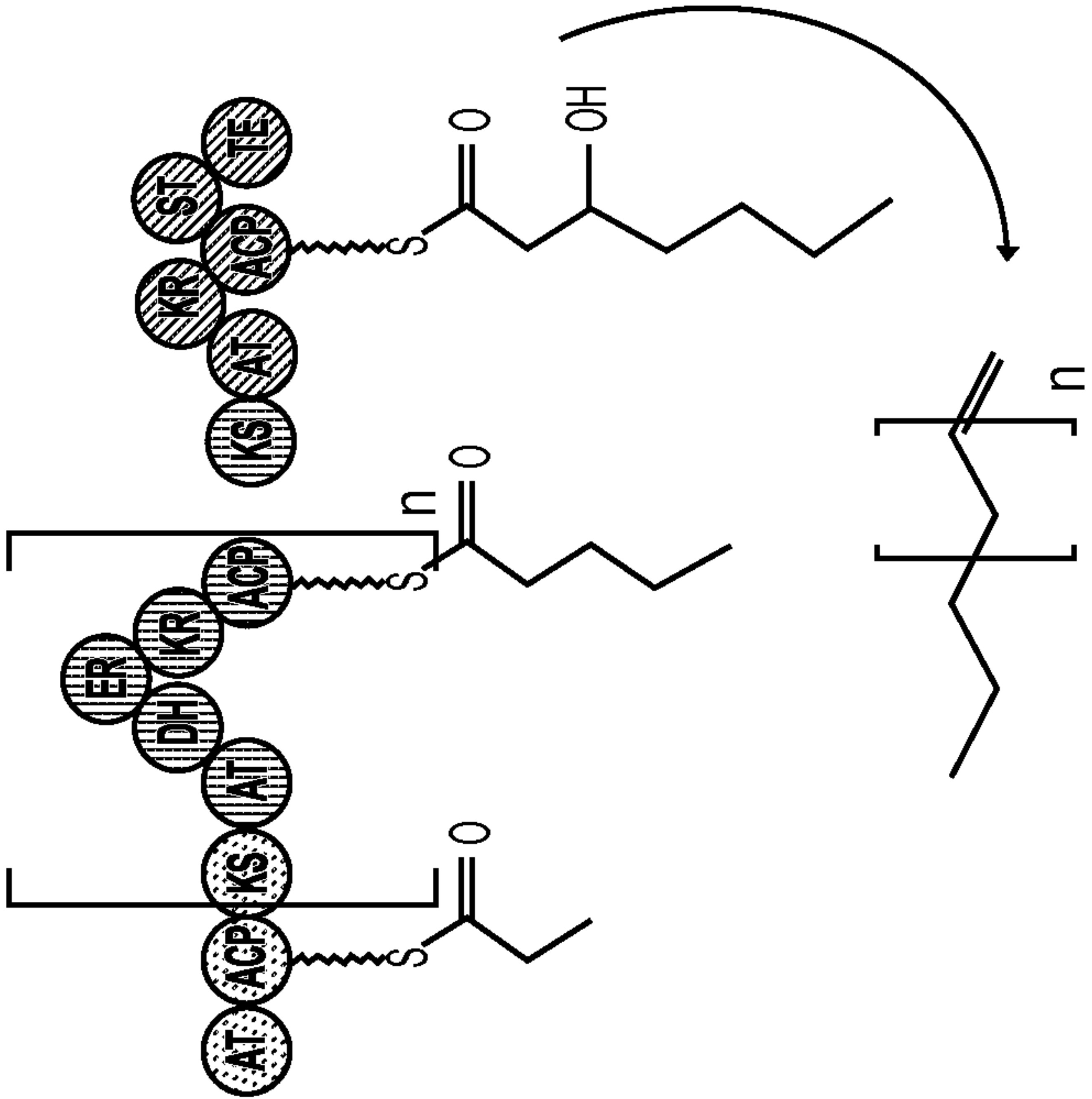


FIG. 3B

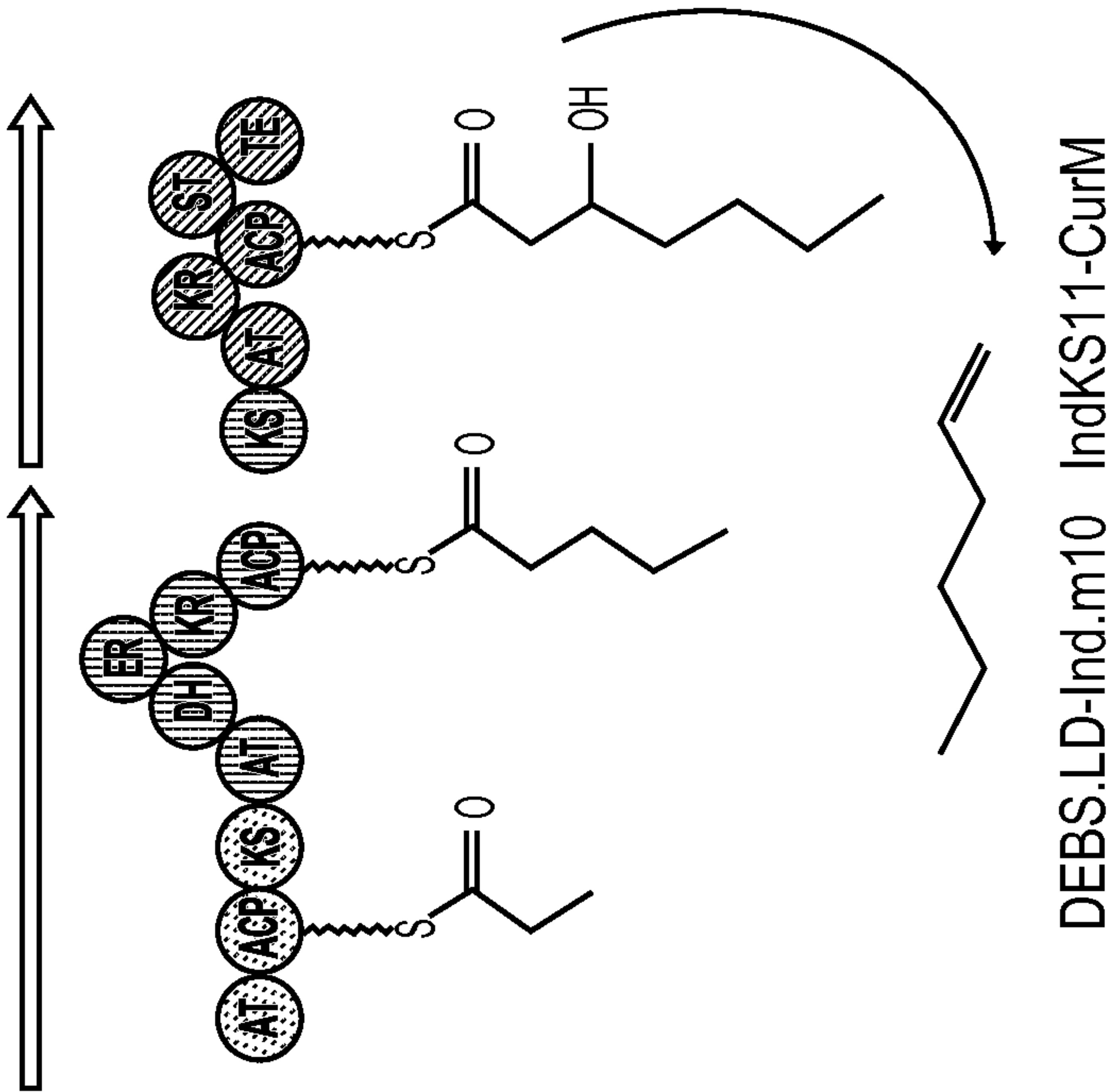


FIG. 3A

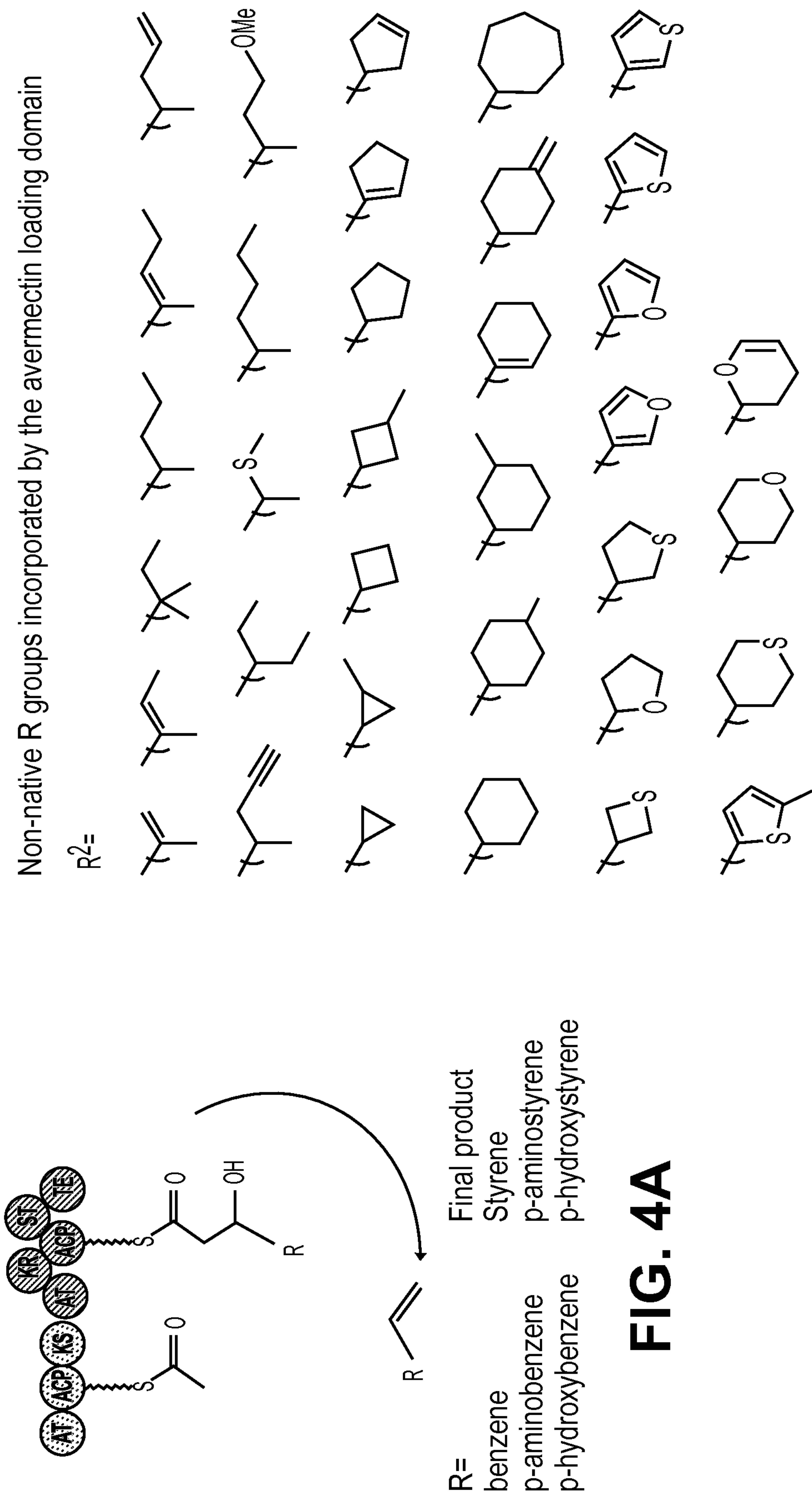
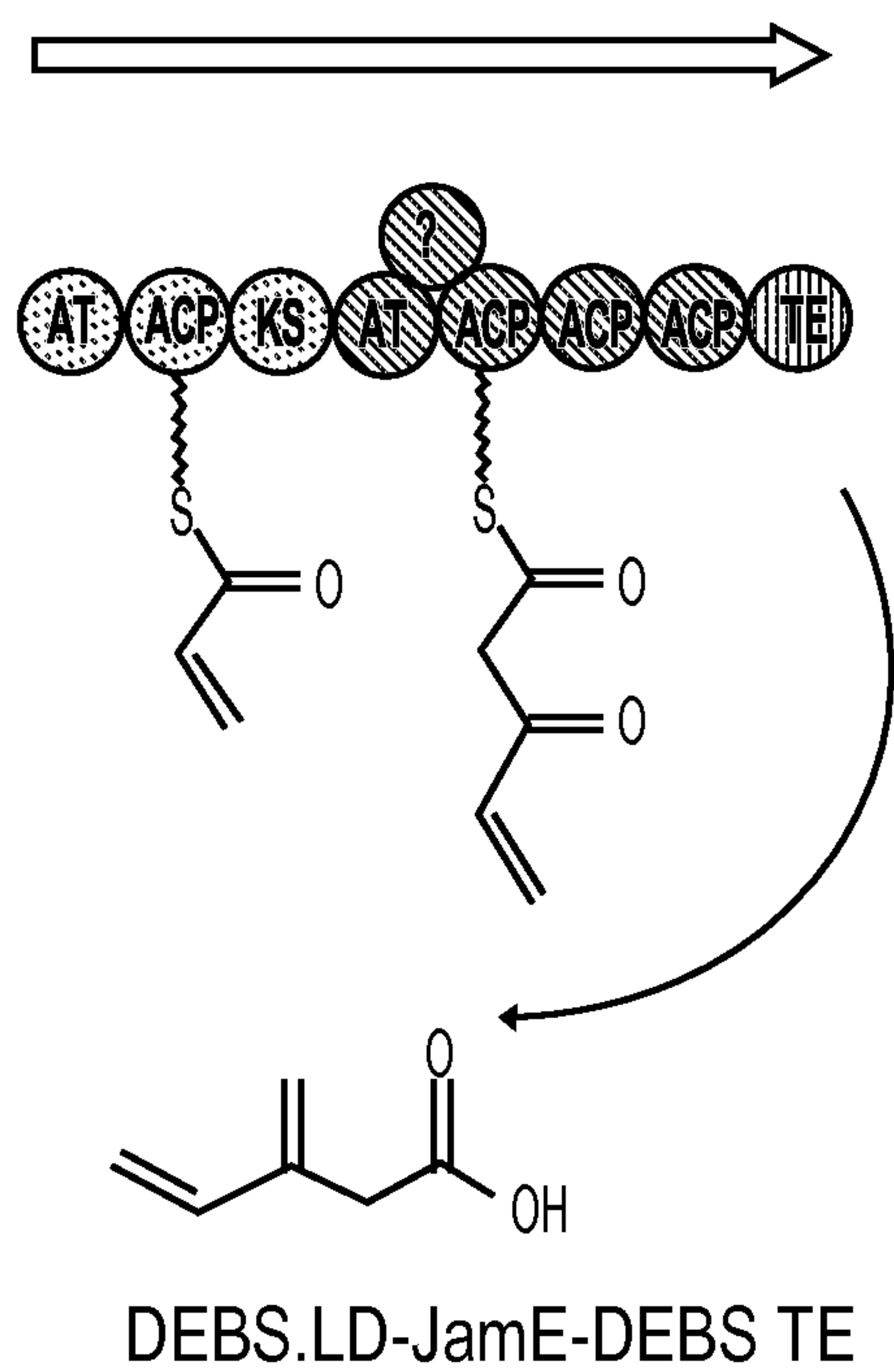
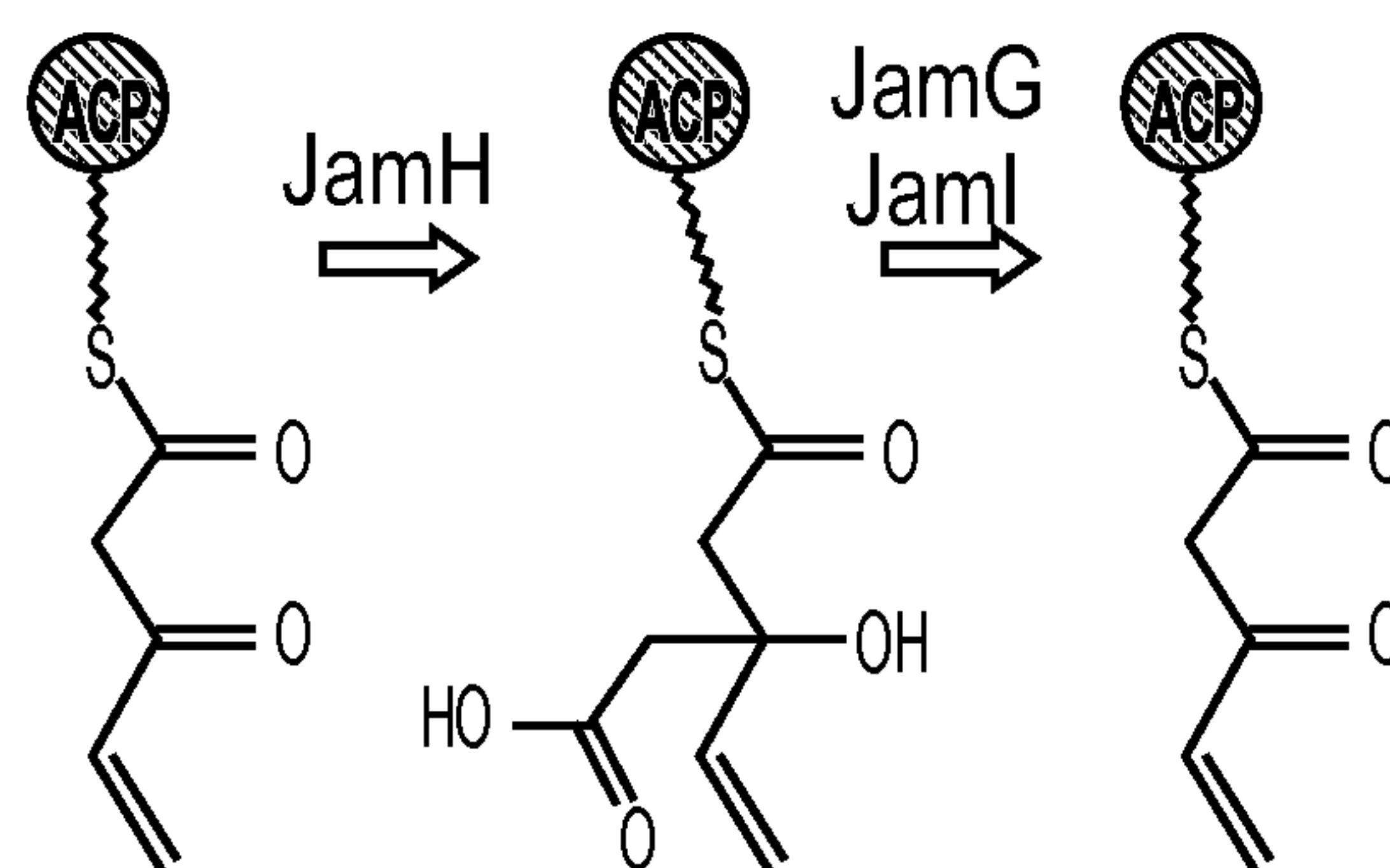
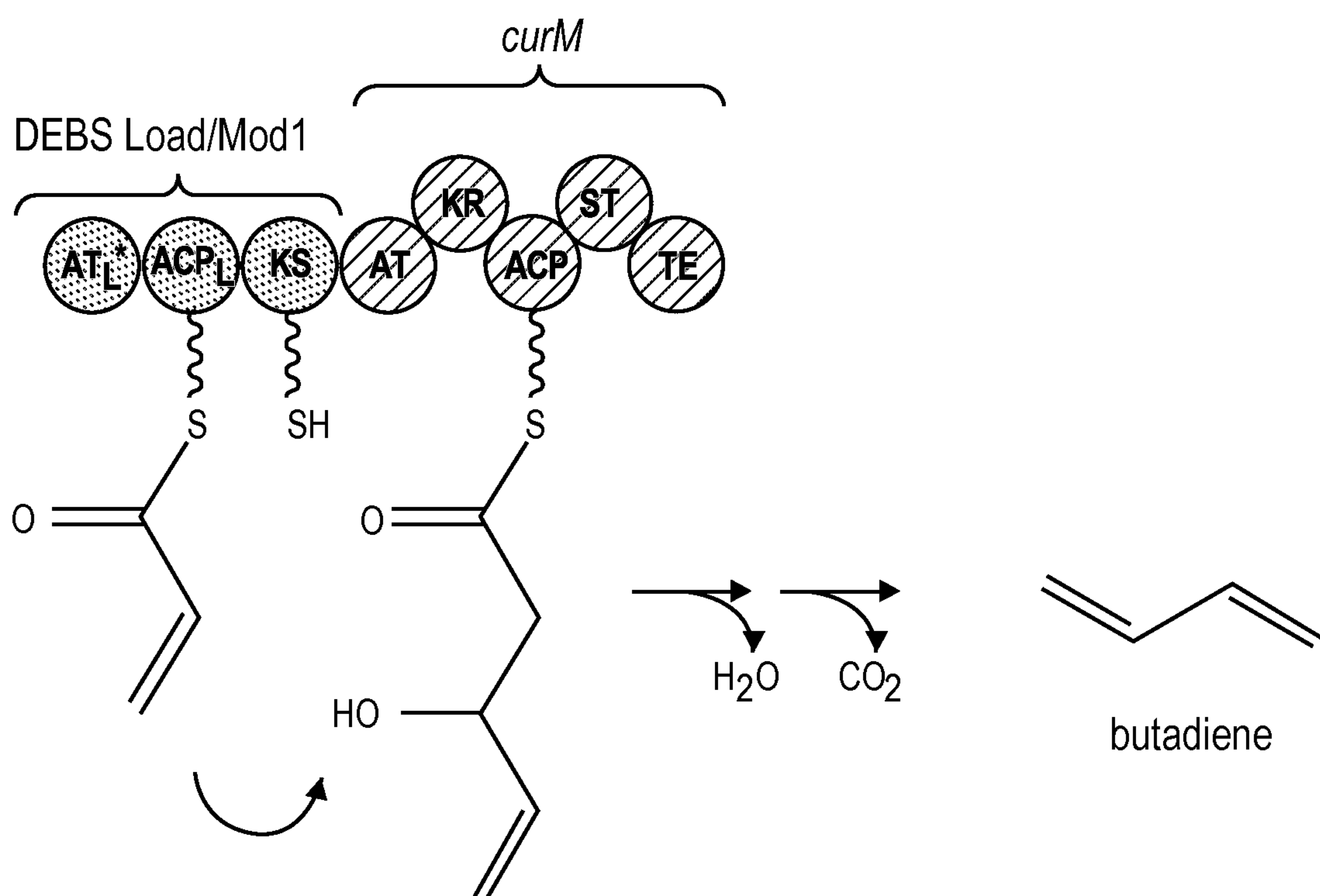


FIG. 4B

**FIG. 5A****FIG. 5B****FIG. 6**

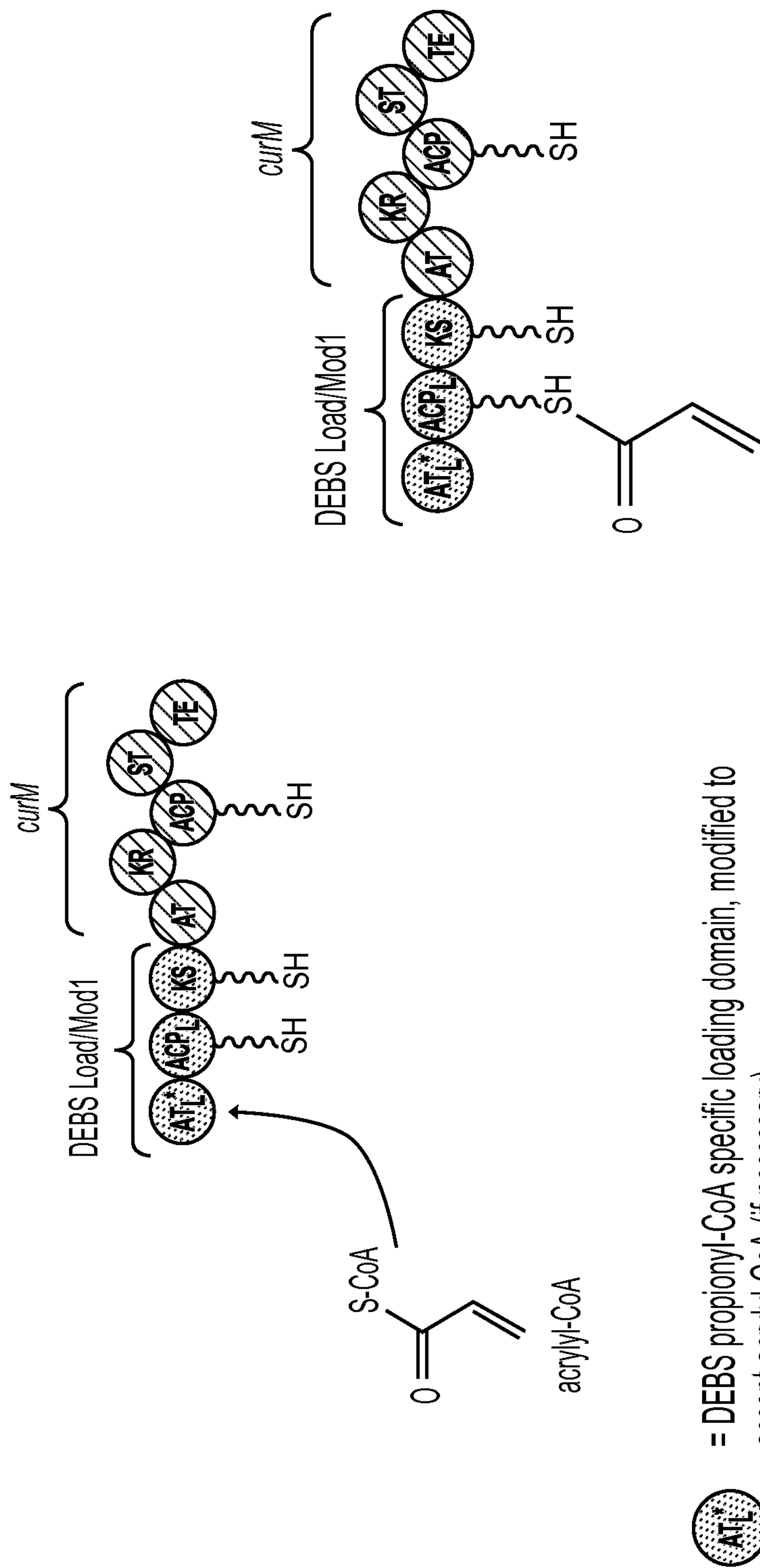
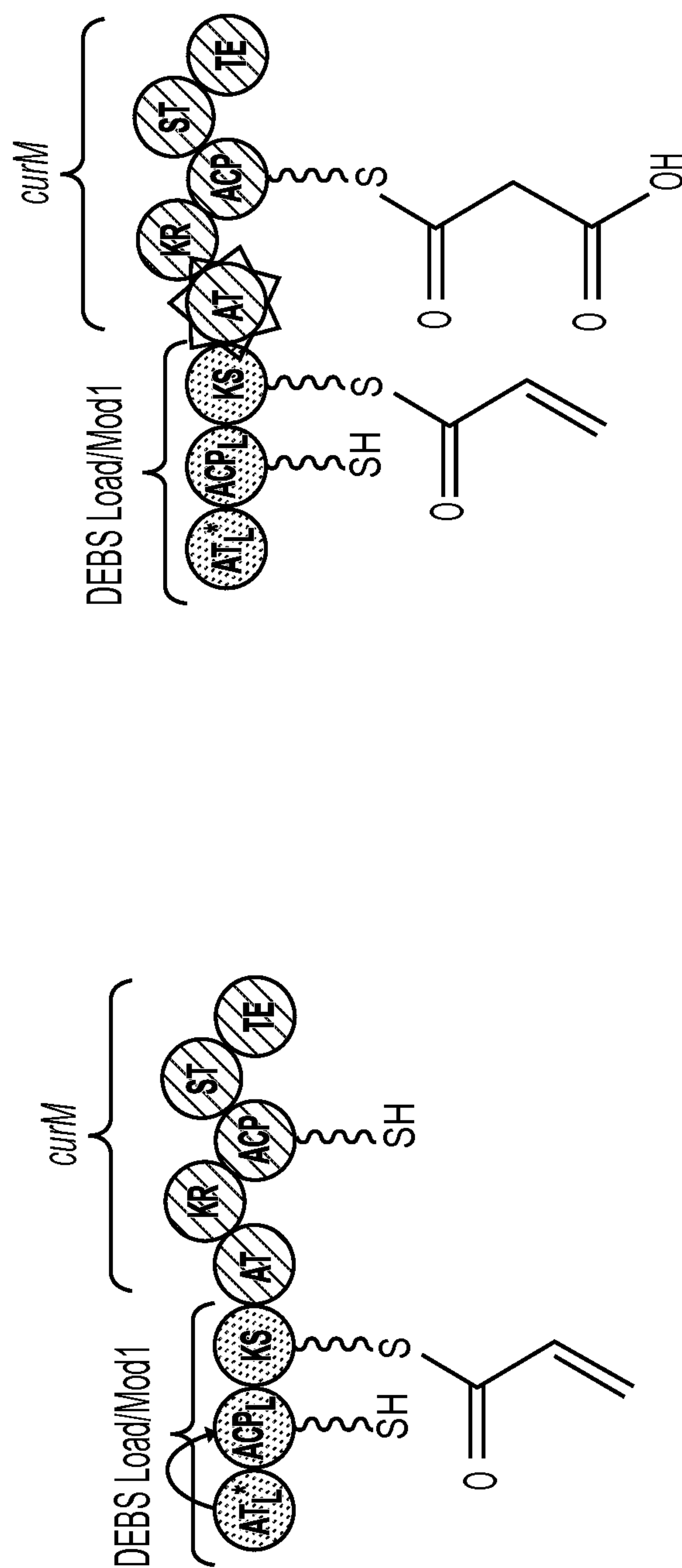


FIG. 7B

FIG. 7A

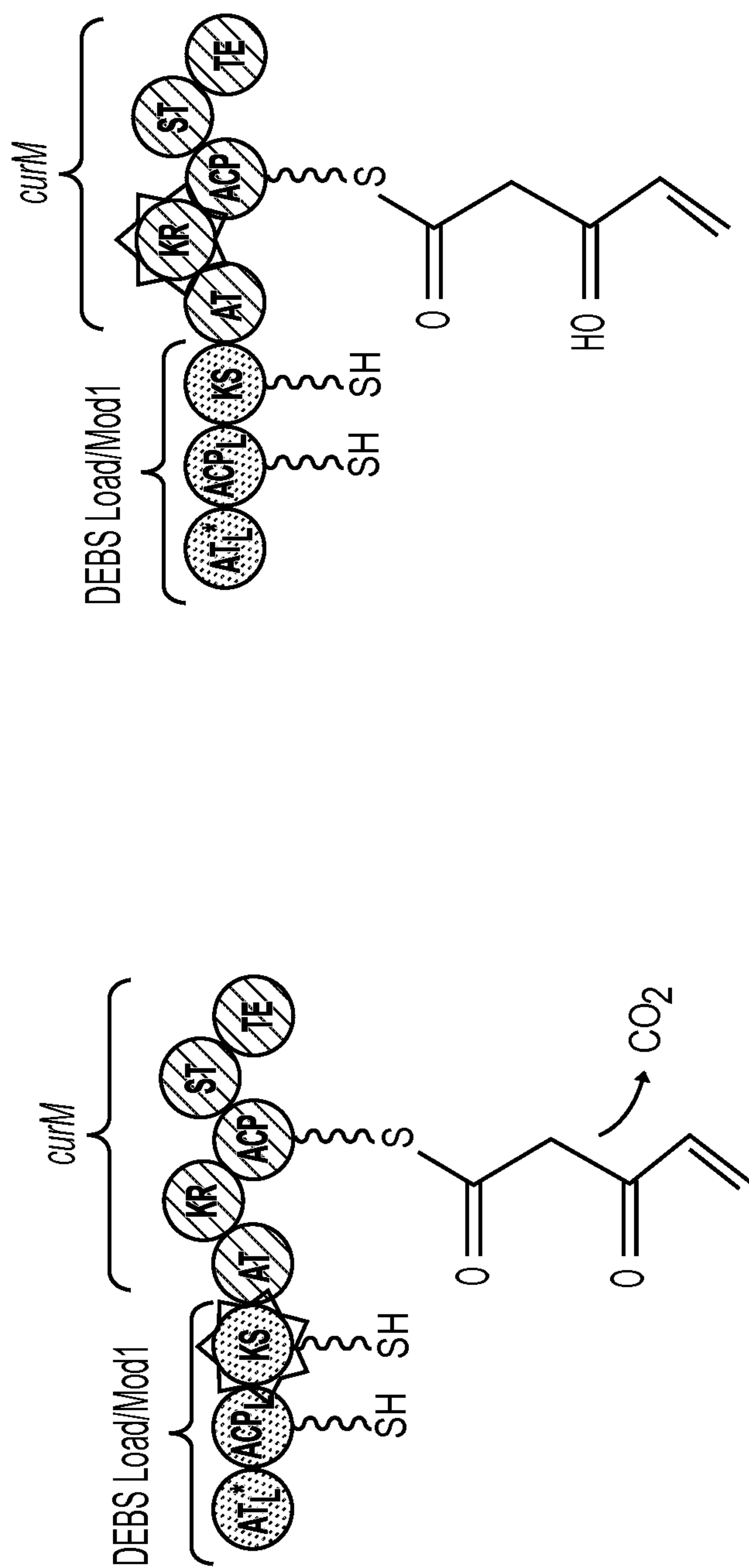


Thiotransfer of acrylate moiety to KS domain

AT Specifies binding of malonyl-CoA and transfer to ACP domain

FIG. 7C

FIG. 7D

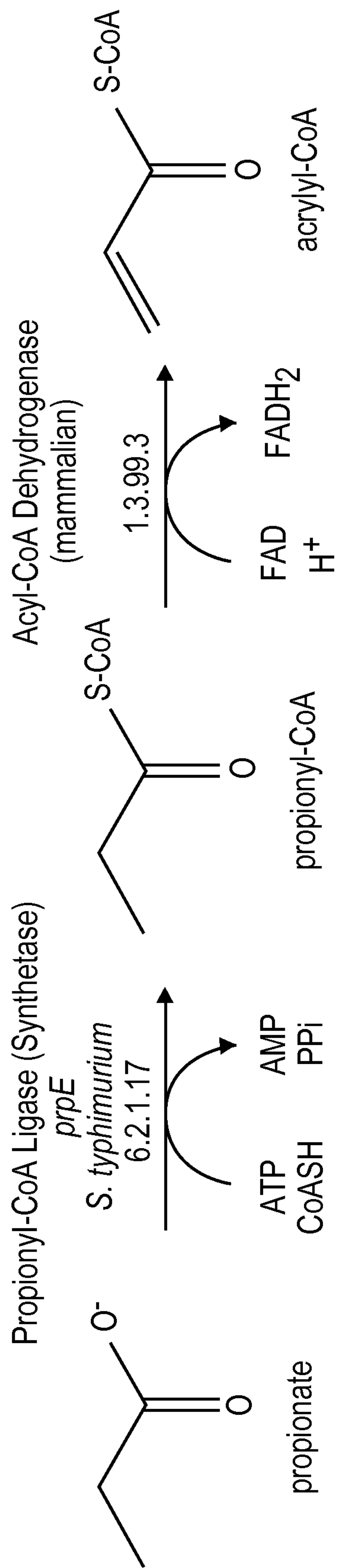
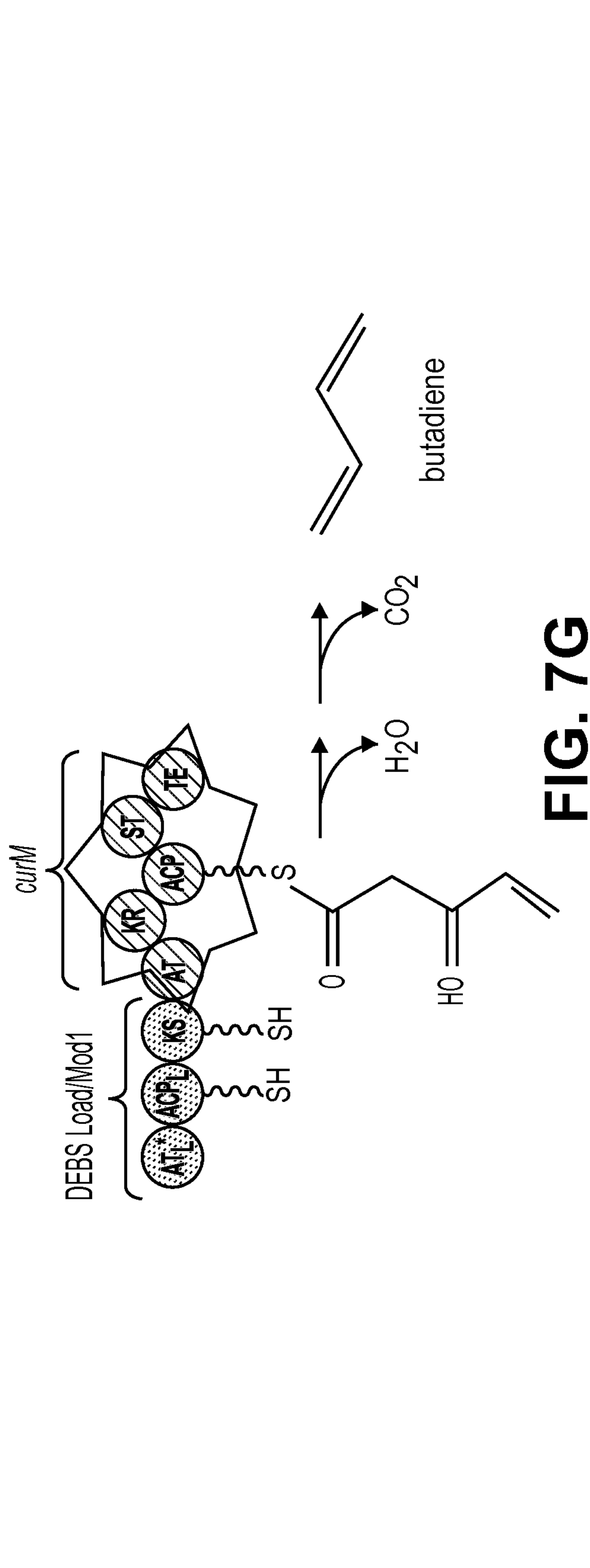


 Catalyzes condensation with release of CO₂

 Catalyzes reduction of the β -carbonyl group

FIG. 7E

FIG. 7F



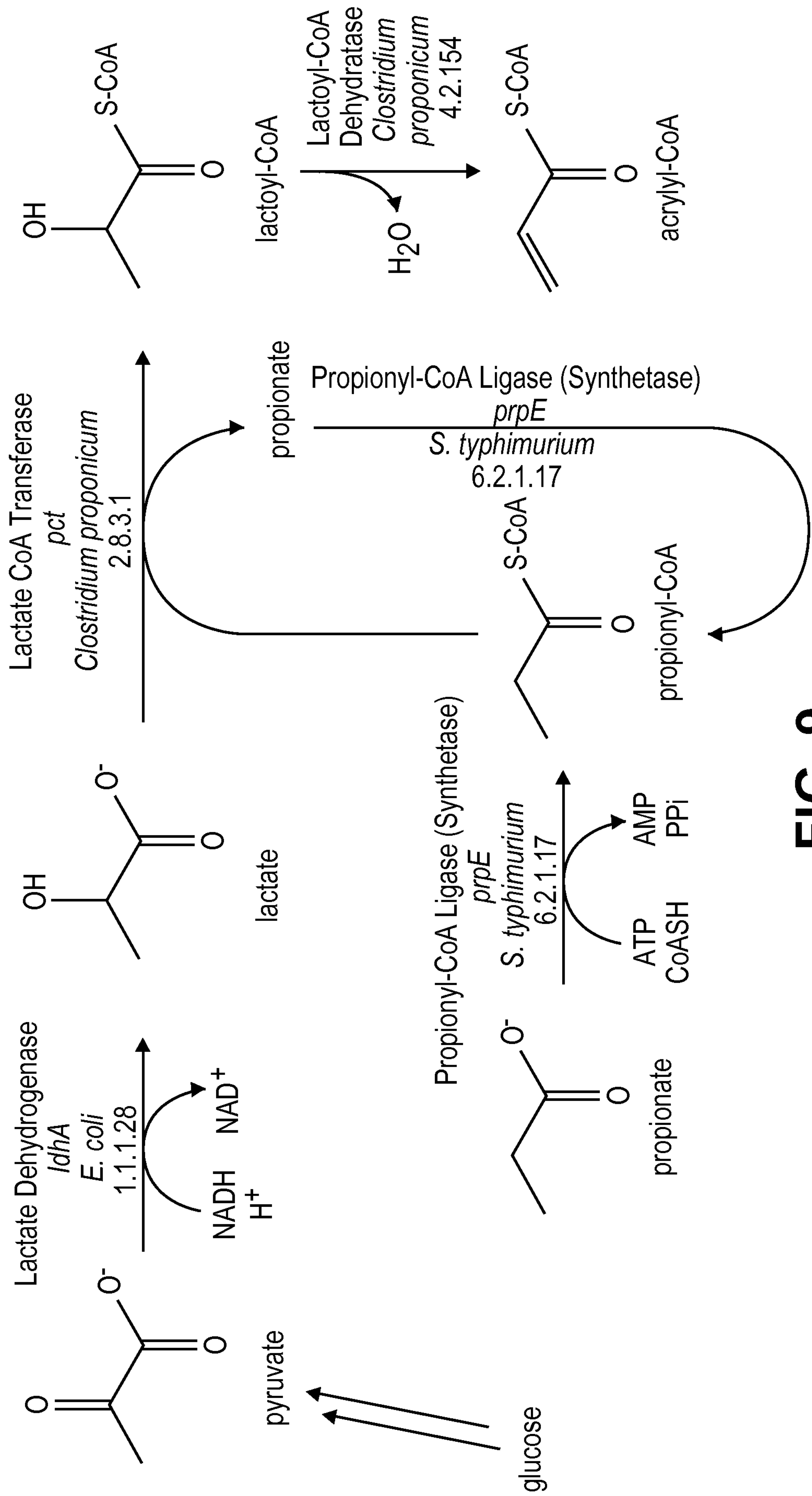


FIG. 9

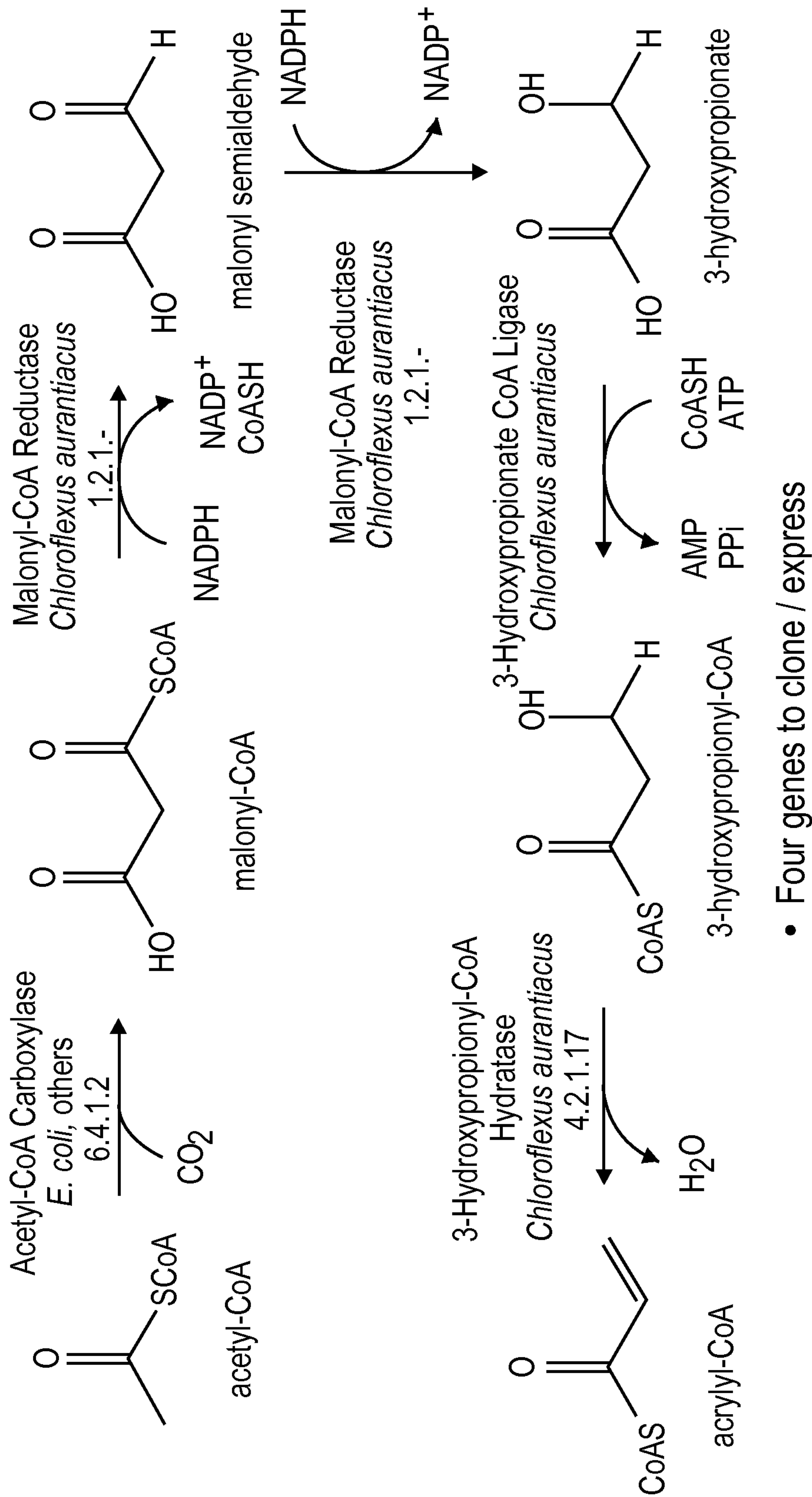


FIG. 10

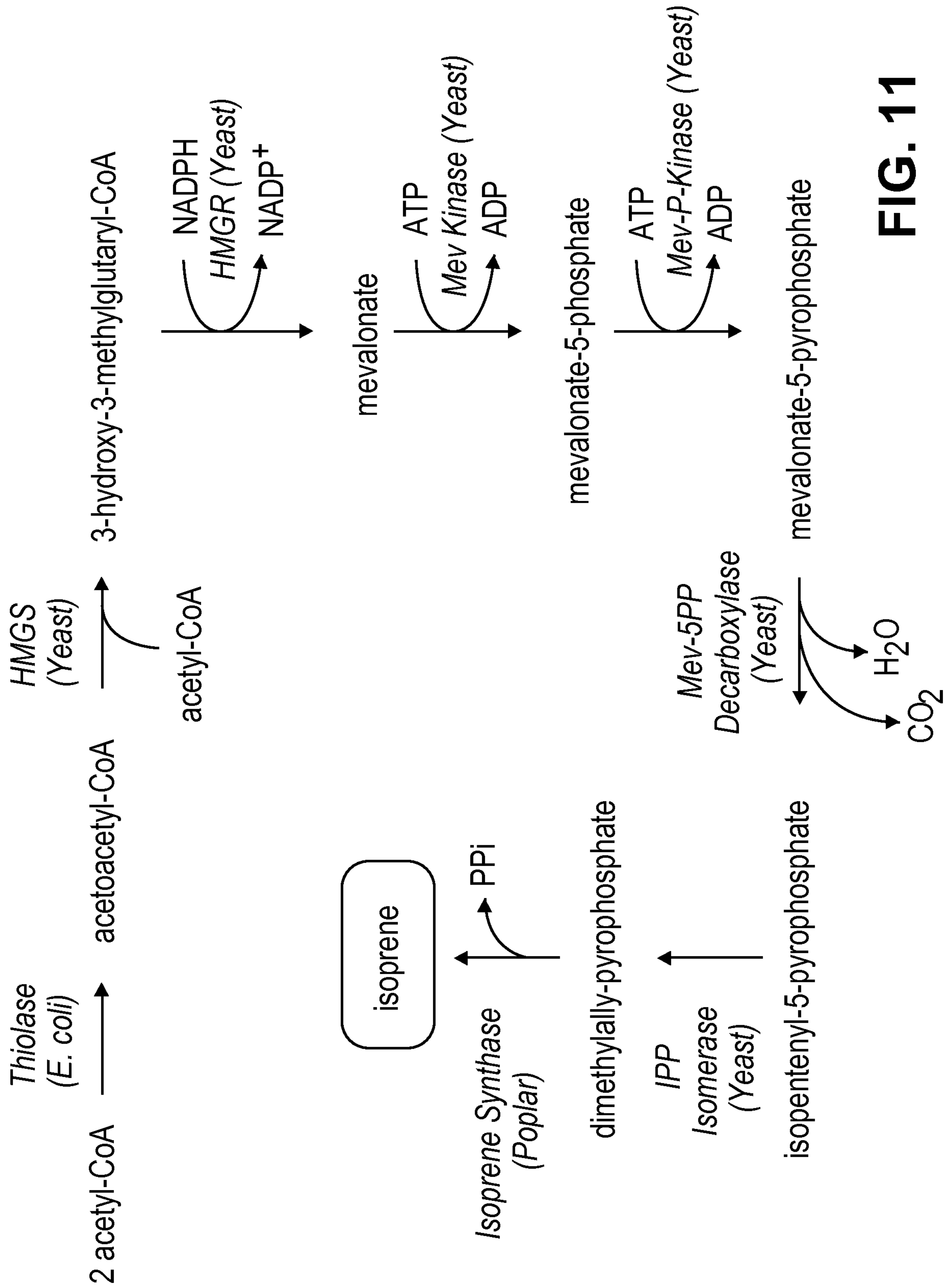


FIG. 11

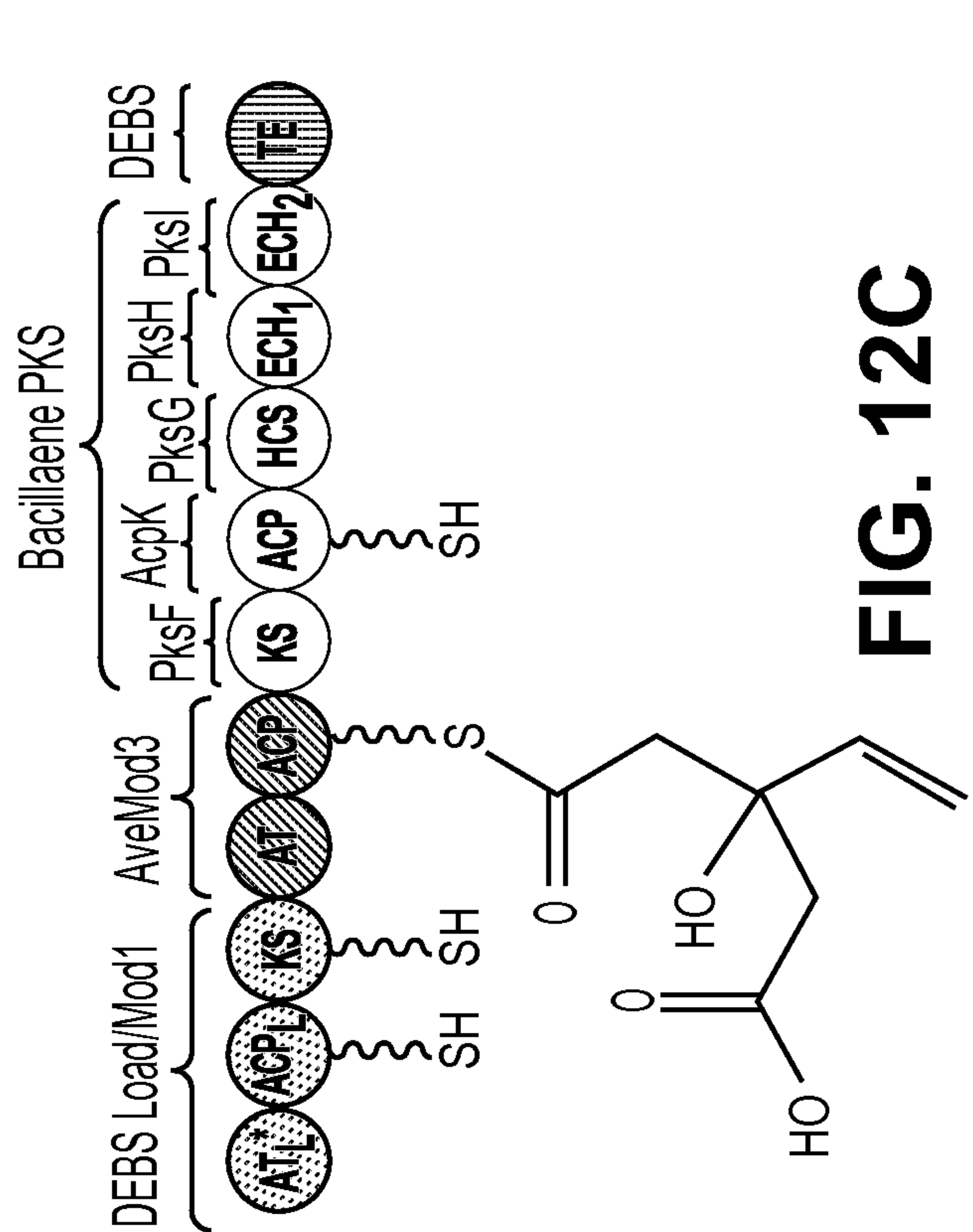
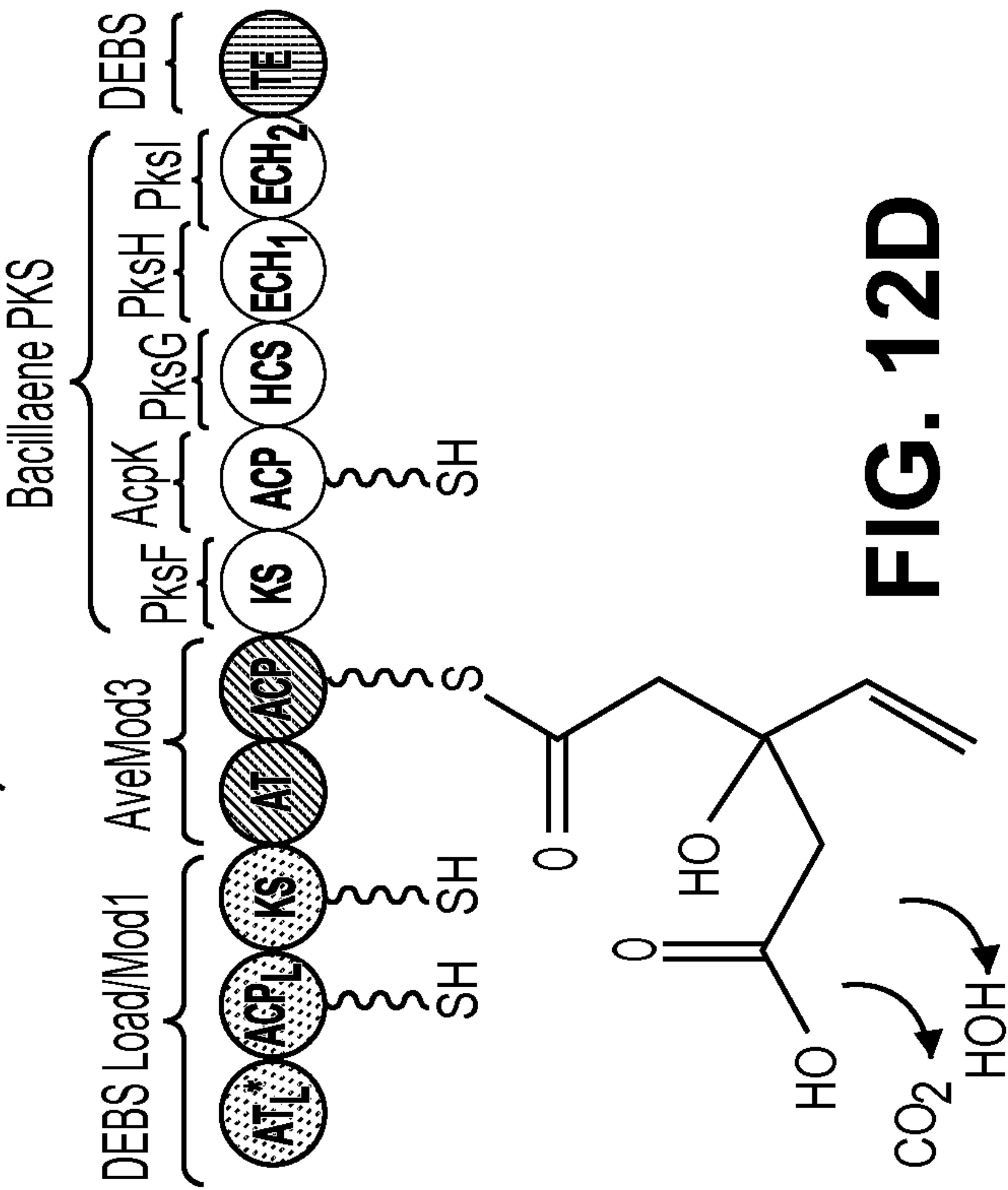
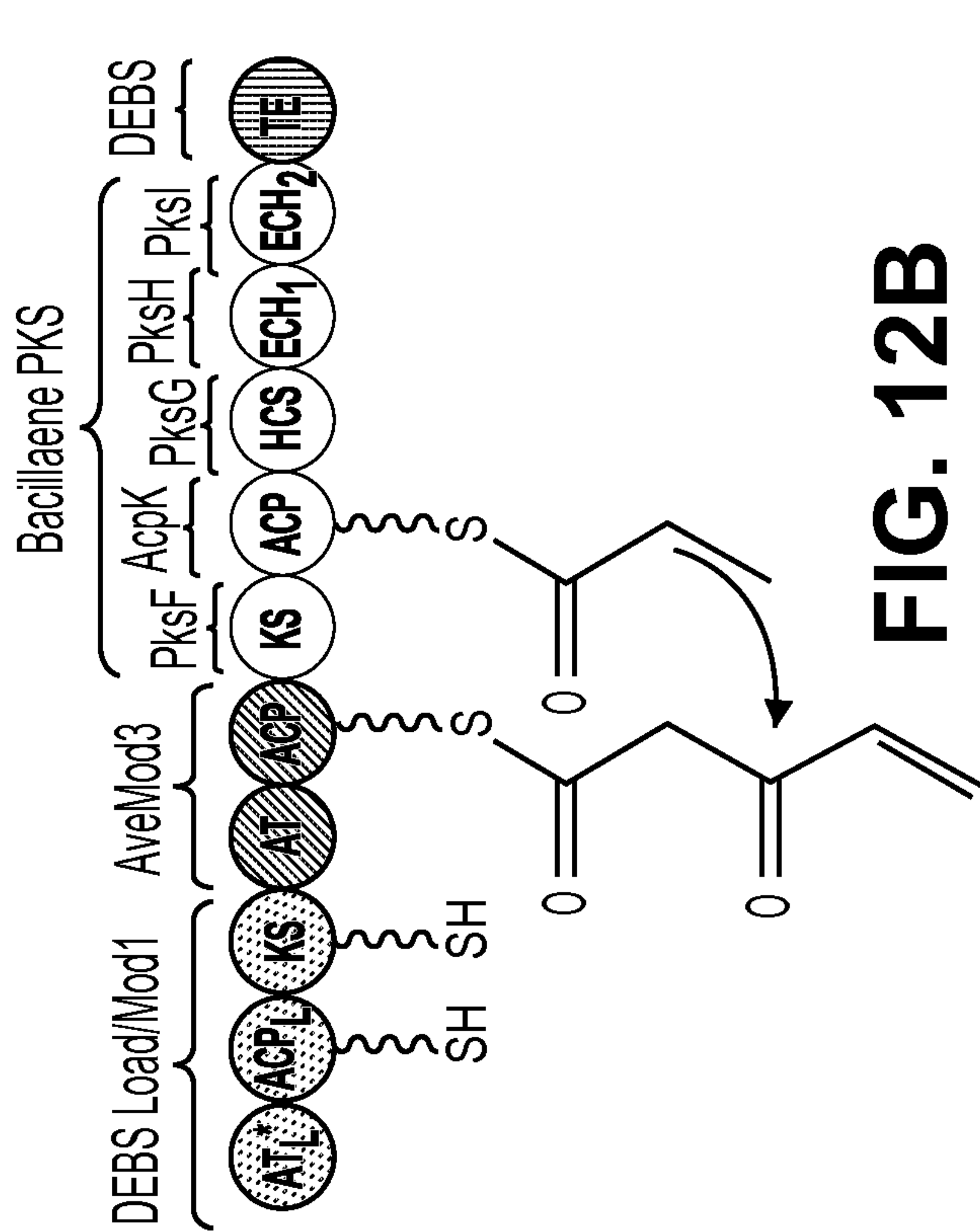


FIG. 12D

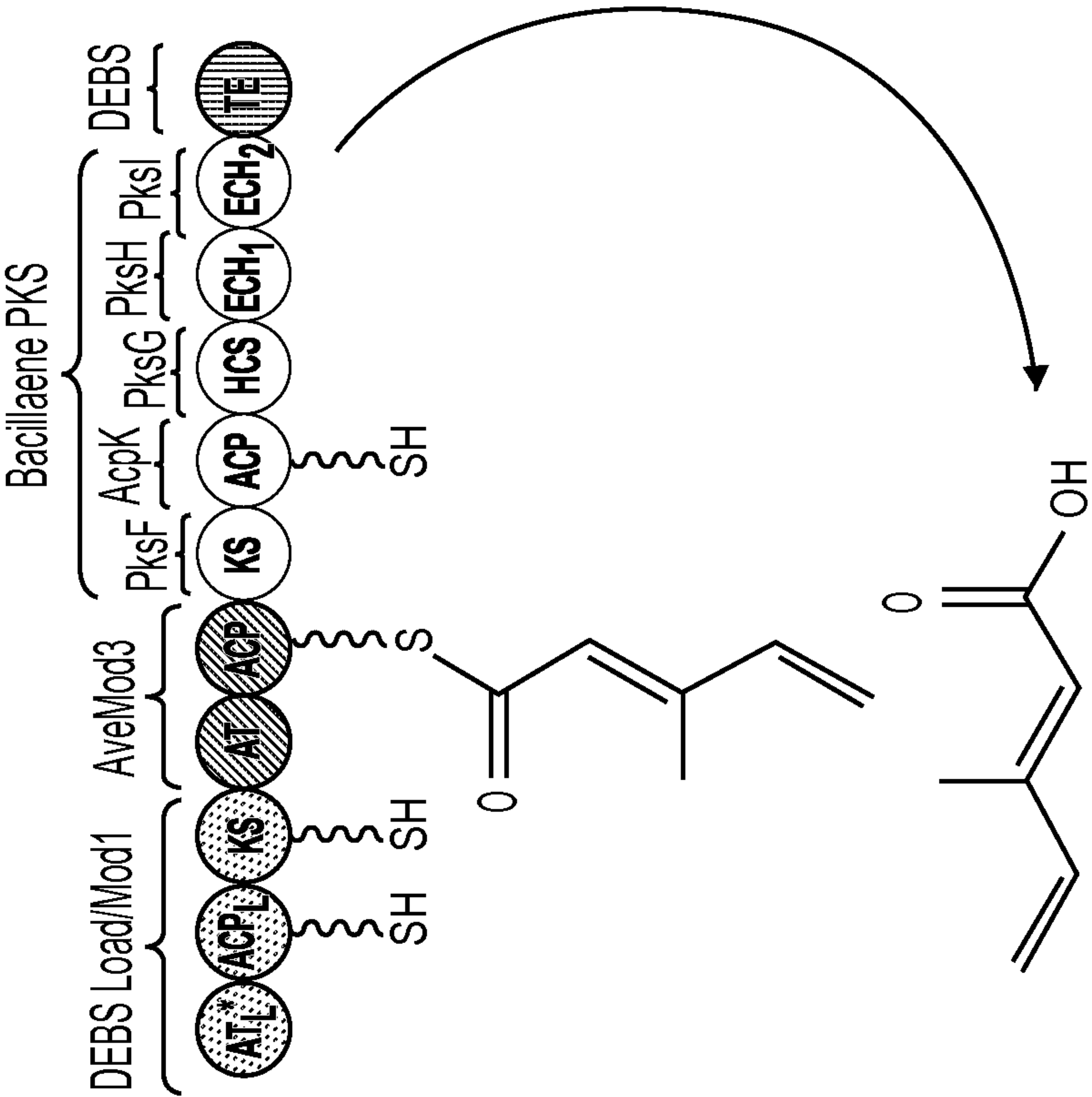


FIG. 12E

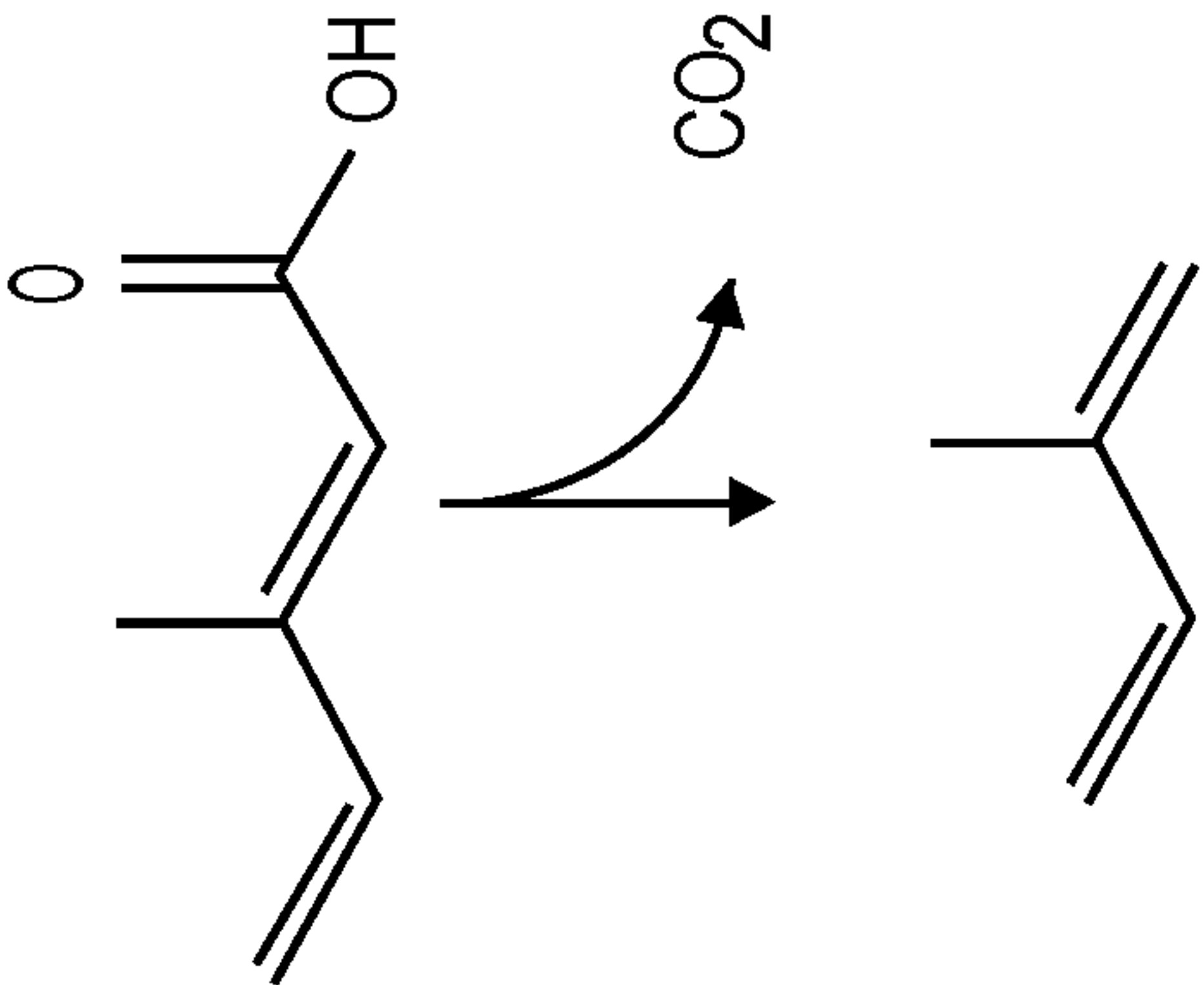


FIG. 12F

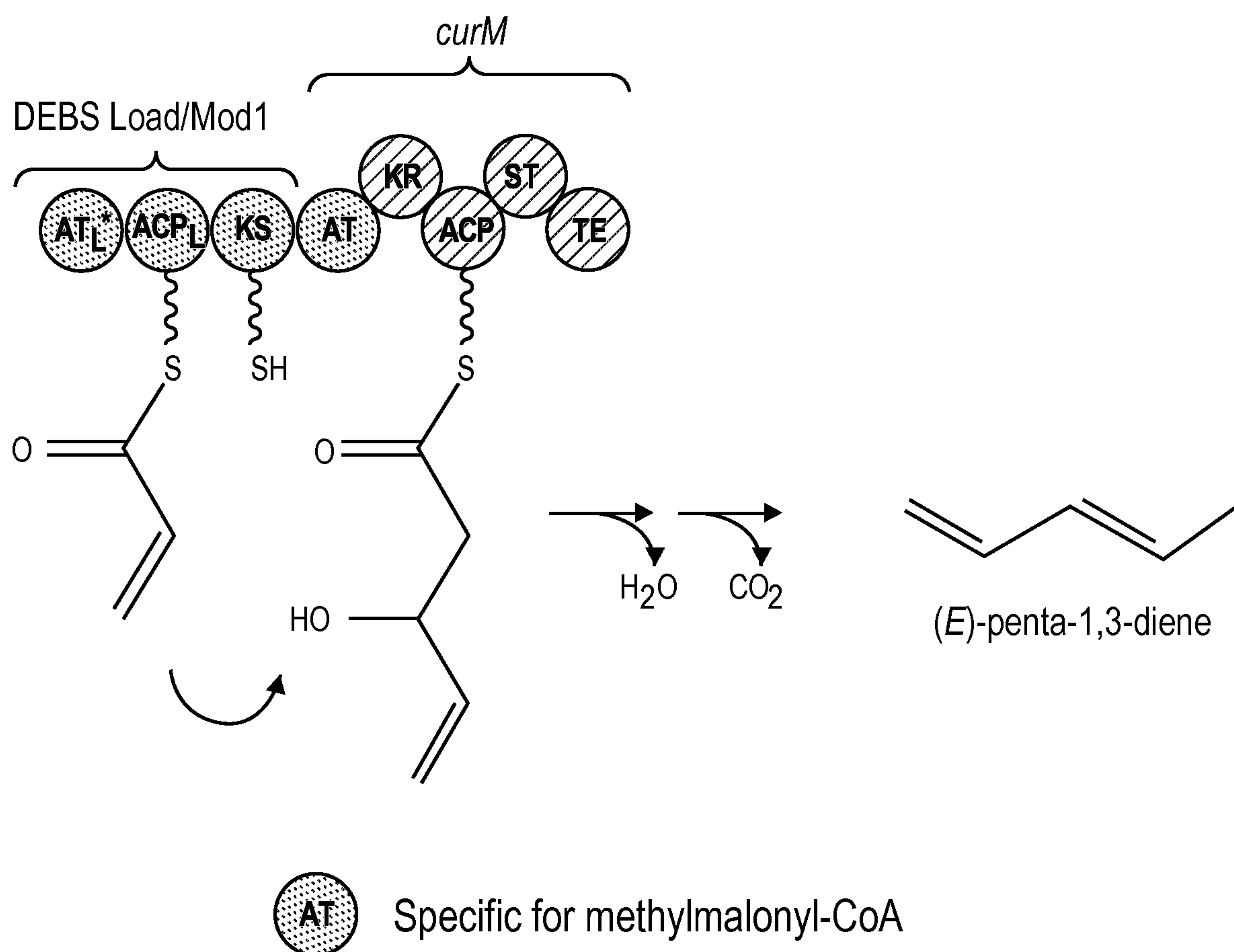
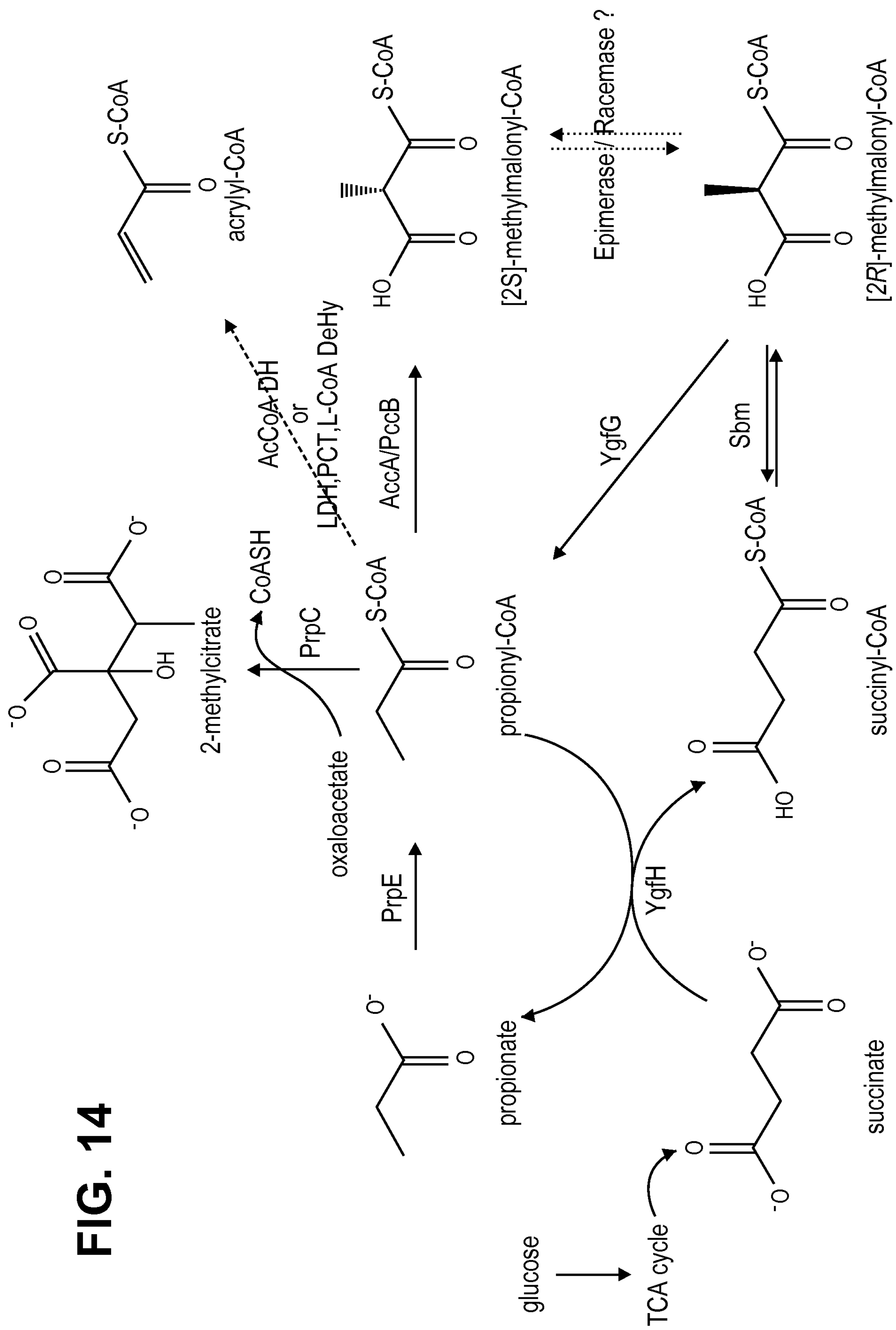
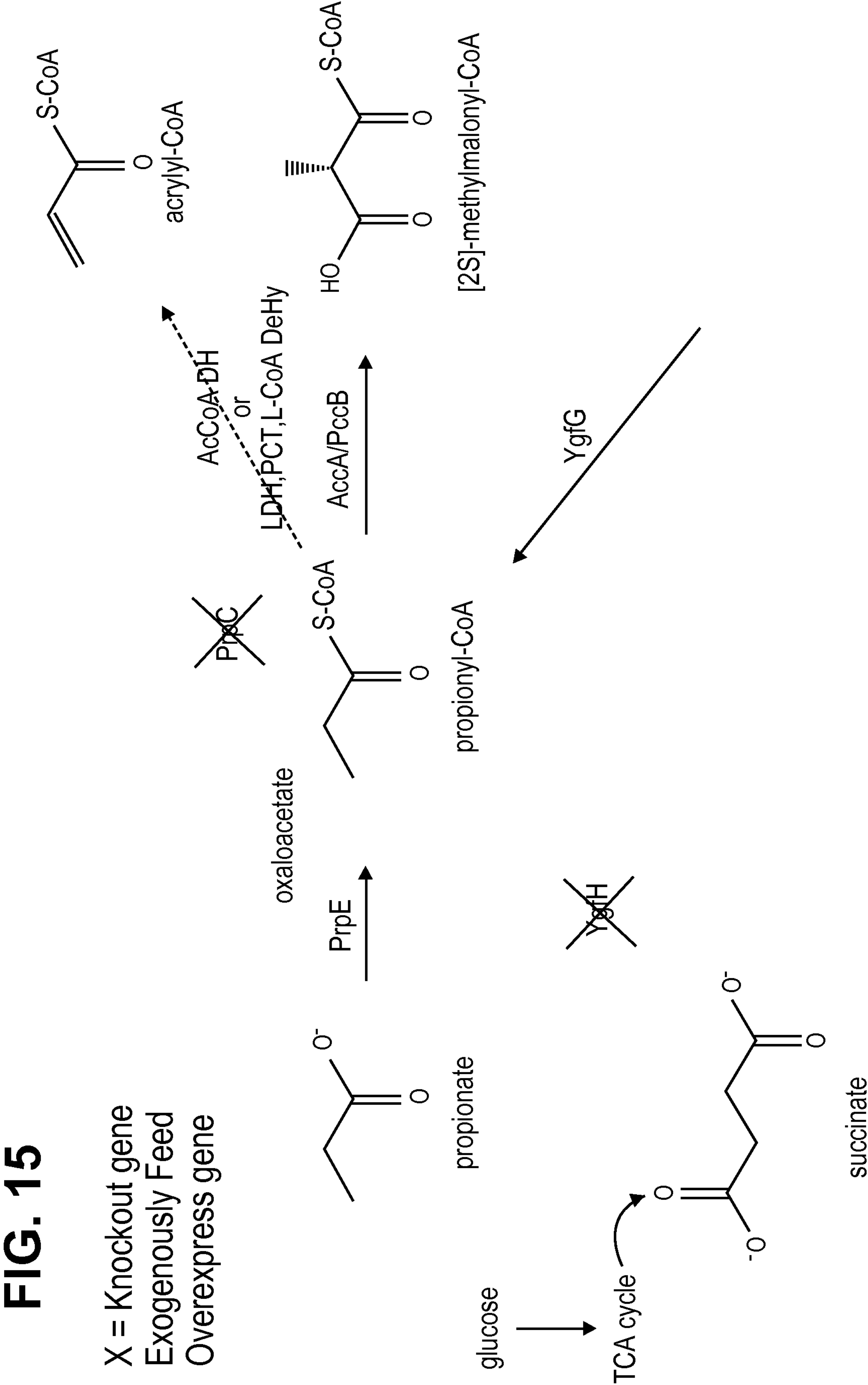
**FIG. 13**

FIG. 14





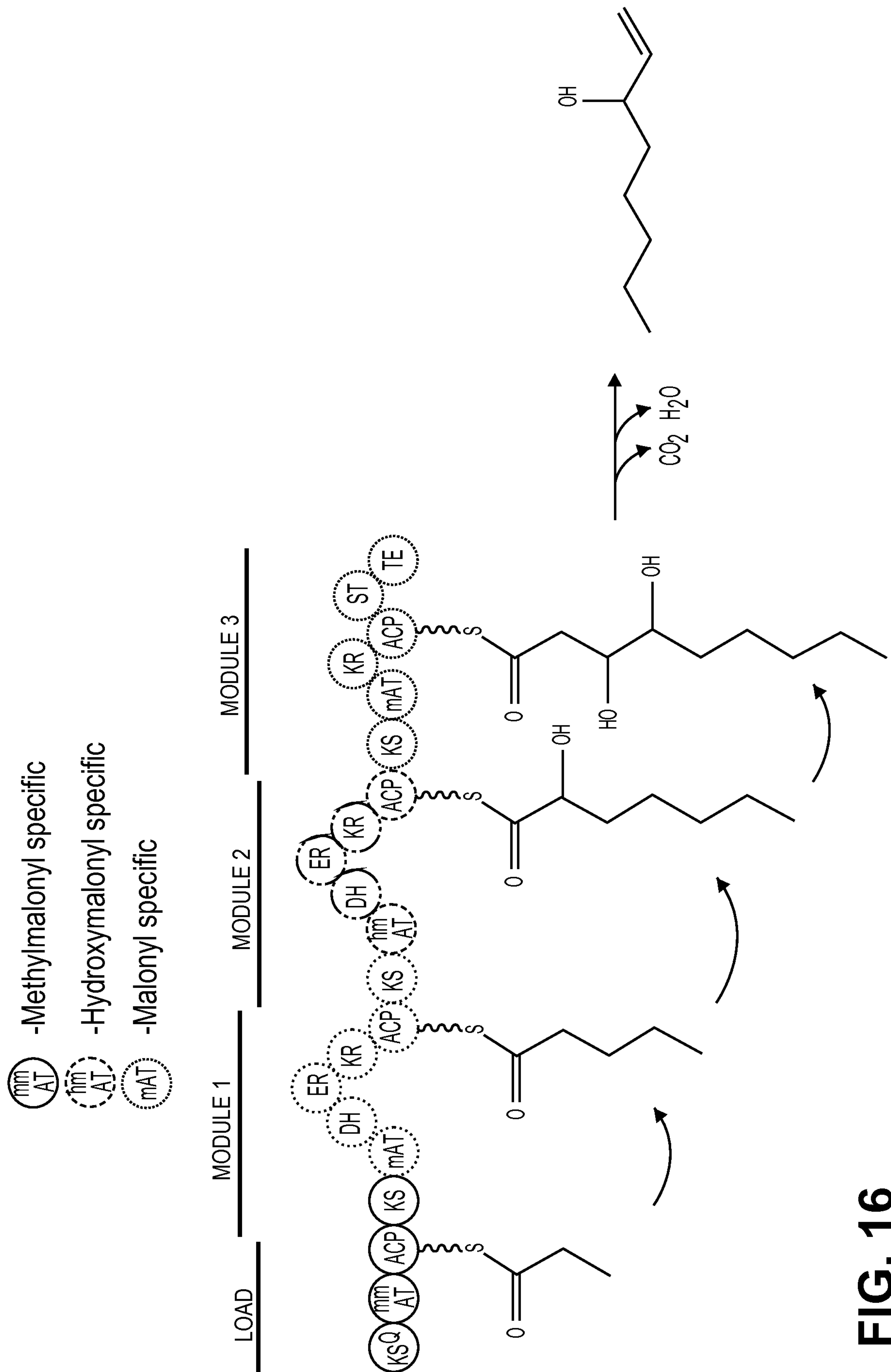


FIG. 16

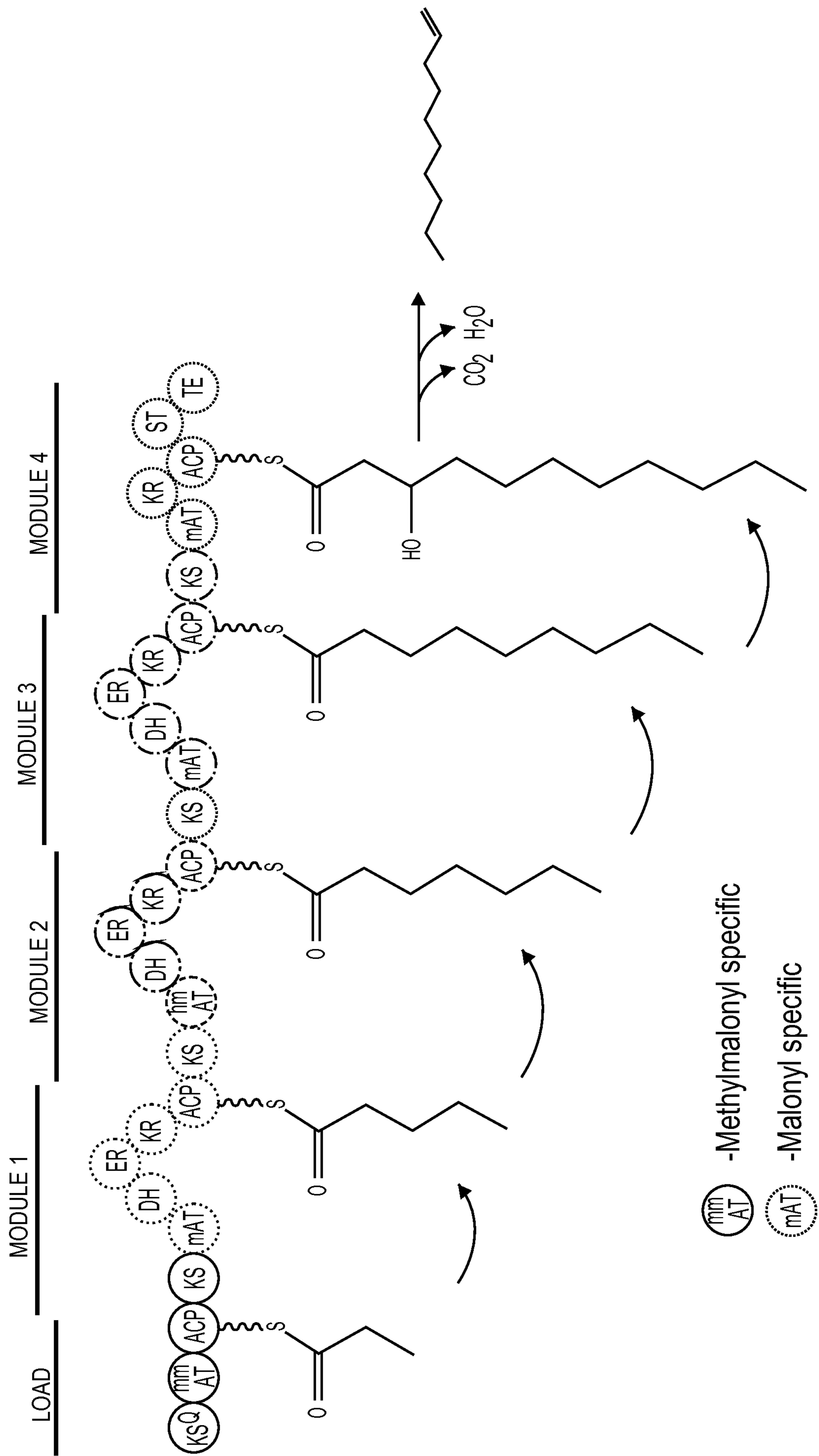


FIG. 17

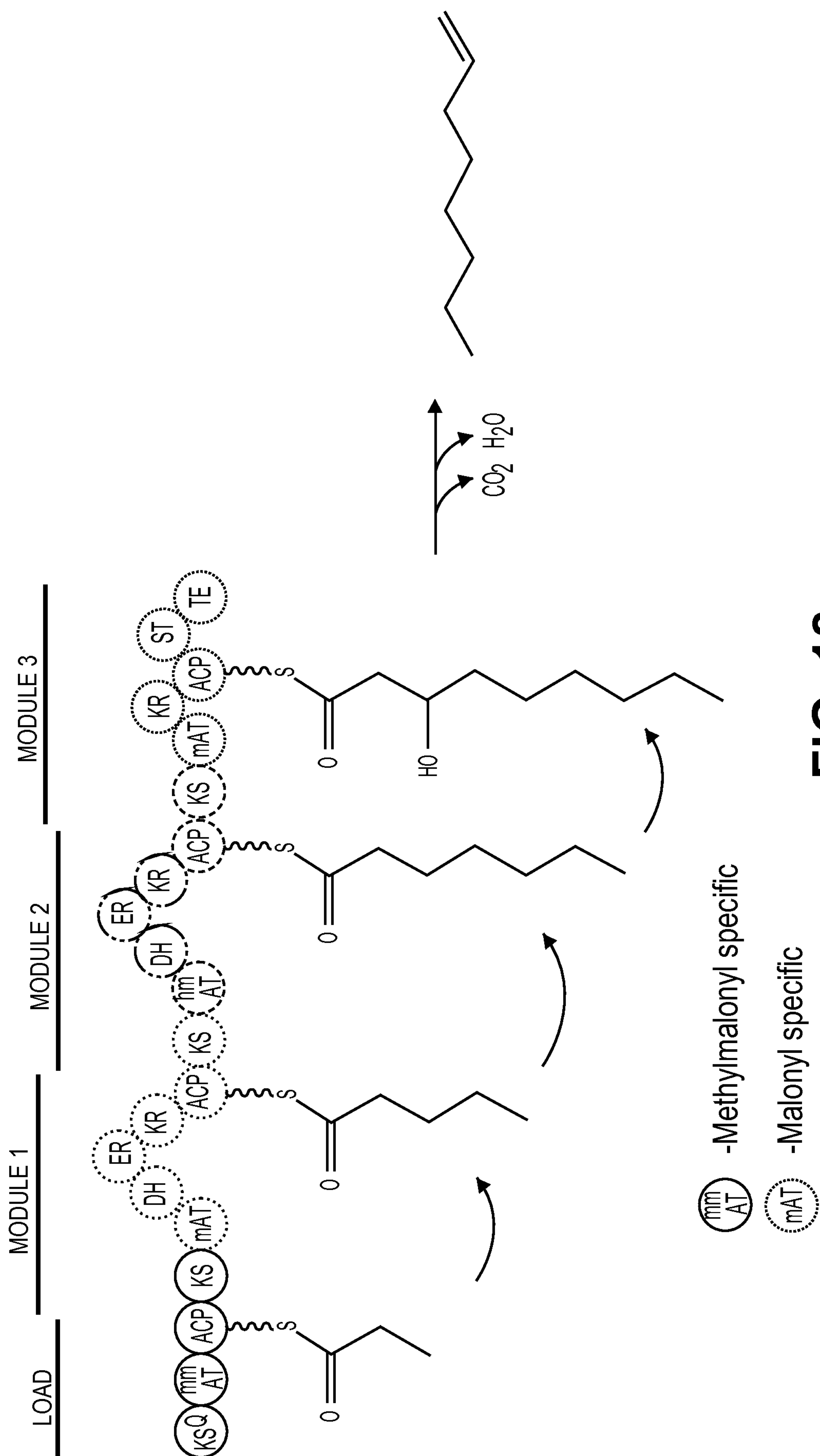


FIG. 18

