



(51) International Patent Classification:

C12N 9/90 (2006.01) C12N 1/19 (2006.01)
C12N 15/52 (2006.01) C12P 17/18 (2006.01)
C12N 15/63 (2006.01)

(21) International Application Number:

PCT/US2018/045222

(22) International Filing Date:

03 August 2018 (03.08.2018)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/541,038 03 August 2017 (03.08.2017) US

(71) Applicant: ANTHEIA, INC. [US/US]; 1505 O'Brien Drive, Suite B1, Menlo Park, California 94025 (US).

(72) Inventors: SMOLKE, Christina D.; 1505 O'Brien Drive, Suite B1, Menlo Park, California 94025 (US). WELLS, Derek H.; 1505 O'Brien Drive, Suite B1, Menlo Park, California 94025 (US).

(74) Agent: SMITH, Kristin D.; Wilson Sonsini Goodrich & Rosati, 650 Page Mill Road, Palo Alto, California 94304-1050 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA,

SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

Published:

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

(54) Title: ENGINEERED BENZYLISOQUINOLINE ALKALOID EPIMERASES AND METHODS OF PRODUCING BENZYLISOQUINOLINE ALKALOIDS

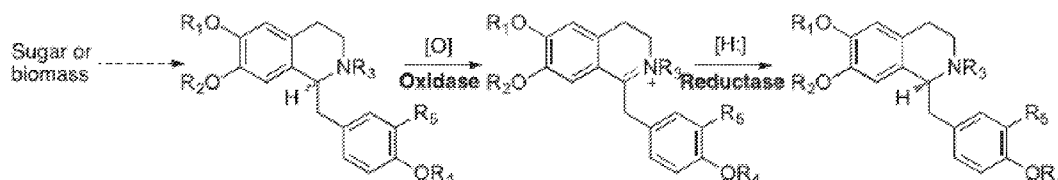


FIG. 3

(57) Abstract: The present disclosure provides systems and methods for increasing production of an alkaloid product through the epimerization of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid via an engineered epimerase in an engineered host cell. A (*S*)-1-benzylisoquinoline alkaloid is contacted with said engineered epimerase. Contacting said (*S*)-1-benzylisoquinoline alkaloid with said engineered epimerase converts said (*S*)-1-benzylisoquinoline alkaloid to said (*R*)-1-benzylisoquinoline alkaloid.



ENGINEERED BENZYLISOQUINOLINE ALKALOID EPIMERASES AND METHODS OF PRODUCING BENZYLISOQUINOLINE ALKALOIDS

CROSS-REFERENCE

[0001] This application claims the benefit of U.S. Provisional Application No. 62/541,038, filed August 3, 2017, which application is incorporated herein by reference.

SUMMARY OF THE INVENTION

[0002] The present disclosure provides methods for the production of diverse benzyloquinoline alkaloids (BIAs) in engineered host cells comprising engineered epimerases. Additionally, the present disclosure provides methods for the production of an engineered epimerase. Engineered epimerases may comprise engineered fused epimerases and engineered split epimerases. Engineered fused epimerases may have one or more modifications to an oxidase portion of the engineered fused epimerase. These one or more modifications may increase activity of the epimerase. In some embodiments, an engineered fused enzyme may have an oxidase enzyme portion and a reductase enzyme portion that originated from different parent enzymes, respectively.

[0003] Engineered epimerases may also comprise engineered split epimerases. In some embodiments, a split epimerase may comprise an epimerase system having two separate components. In some embodiments, a split epimerase may comprise a first protein comprising an oxidase component and a second, separate protein comprising a reductase component. In some cases, the oxidase component and the reductase component of a split epimerase are encoded by separate mRNAs. In some cases, the oxidase component and the reductase component of a split epimerase are encoded by a same mRNA and are translated from separate translation initiation sites. In some cases, the oxidase component of a split epimerase is encoded by a first mRNA, and the reductase component of the same split epimerase is encoded by a second, separate mRNA. In some cases, the oxidase component of a split epimerase comprises a first molecule and the reductase component of the same split epimerase comprises a second molecule. In some cases, a split epimerase comprises a first protein encoded by a first transcription unit in a first genomic location and a second protein encoded by a second transcription unit in a second genomic location.

[0004] In some embodiments, an engineered split epimerase may comprise two separate enzymes encoding an oxidase and a reductase, in engineered host cells. In some embodiments, an engineered split epimerase may further comprise an engineered oxidase portion having one or more modifications to increase activity of the epimerase. In some embodiments, an engineered split epimerase may have an oxidase enzyme and reductase enzyme split from a parent fused enzyme. In some embodiments, an engineered split epimerase may have an oxidase enzyme split from a first parent fused enzyme and a reductase enzyme split from a second parent fused enzyme that is different than the first parent fused enzyme. In some embodiments, an engineered split epimerase may have an oxidase enzyme that

originates from a portion of a first parent enzyme and a reductase enzyme that originates from a portion of a second parent enzyme.

[0005] In some embodiments, the disclosure provides methods for increasing production of an alkaloid product through the epimerization of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid via an engineered split epimerase in an engineered host cell in comparison to the epimerization of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid via a fused epimerase in similar conditions. In some methods, the methods comprise contacting said (*S*)-1-benzylisoquinoline alkaloid with said engineered split epimerase in said engineered host cell, thereby increasing production of an alkaloid product, wherein contacting said (*S*)-1-benzylisoquinoline alkaloid with said engineered split epimerase converts said (*S*)-1-benzylisoquinoline alkaloid to said (*R*)-1-benzylisoquinoline alkaloid within said engineered host cell.

[0006] In some embodiments, the disclosure provides methods for producing an alkaloid product through the epimerization of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid via an engineered split epimerase in an engineered host cell. In some methods, the methods comprise contacting said (*S*)-1-benzylisoquinoline alkaloid with said engineered split epimerase in said engineered host cell, thereby increasing production of an alkaloid product, wherein contacting said (*S*)-1-benzylisoquinoline alkaloid with said engineered split epimerase converts said (*S*)-1-benzylisoquinoline alkaloid to said (*R*)-1-benzylisoquinoline alkaloid within said engineered host cell, wherein said engineered split epimerase comprises an oxidase component and a reductase component, wherein said oxidase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 17, and wherein said reductase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 18.

[0007] In some embodiments, the disclosure provides engineered host cells which comprise an engineered split epimerase, wherein said engineered split epimerase increases conversion of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid relative to conversion of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid via a fused epimerase under similar conditions.

[0008] In some embodiments, the disclosure provides methods for increasing production of diverse alkaloid products through the epimerization of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid via engineered epimerases in an engineered host cell. In further embodiments, the present disclosure provides methods for increasing production of diverse alkaloid products through the epimerization of (*S*)-reticuline to (*R*)-reticuline via an engineered epimerase comprising two separate enzymes encoding an oxidase and a reductase compared to the production of diverse alkaloid products through the epimerization of (*S*)-reticuline to (*R*)-reticuline via a wild-type epimerase.

[0009] An additional aspect of the invention provides a method of increasing the efficiency of epimerizing a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid relative to a wild-

type epimerase by utilizing one or more engineered epimerases. The method comprises contacting the (*S*)-1-benzylisoquinoline alkaloid with one or more engineered epimerases. In some embodiments, an engineered split epimerase comprises two separate enzymes encoding an oxidase and a reductase. Contacting the (*S*)-1-benzylisoquinoline alkaloid with the oxidase enzyme of the engineered split epimerase and, separately, the reductase enzyme of the engineered split epimerase, converts more of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid relative to a wild-type epimerase.

[0010] While engineered split epimerases may be composed of a separate oxidase enzyme and reductase enzyme that originate from a parent or wild-type epimerase, engineered epimerases may also comprise a separate oxidase enzyme and reductase enzyme that originate from separate parent or wild-type epimerases. Examples of parent epimerases having an oxidase and reductase component comprise amino acid sequences selected from the group consisting of: SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, and 15, as listed in **Table 1**.

[0011] A further aspect of the invention provides a process for converting a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid comprising contacting the (*S*)-1-benzylisoquinoline alkaloid with an engineered oxidase, comprising at least one modification, and a reductase in an amount sufficient to convert the (*S*)-1-benzylisoquinoline alkaloid to (*R*)-1-benzylisoquinoline alkaloid. The engineered oxidase may have one or more modifications that improve conversion of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid. In some embodiments, the engineered oxidase may be within an engineered split epimerase, where the engineered oxidase enzyme and the reductase enzyme are separate. In some embodiments, the engineered oxidase may be within an engineered fused epimerase, where the engineered oxidase enzyme and the reductase enzyme are fused.

[0012] In another aspect of the invention, a method of epimerizing a stereocenter of a 1-benzylisoquinoline alkaloid is provided. The method comprises contacting the 1-benzylisoquinoline alkaloid with at least one engineered epimerase. Contacting the 1-benzylisoquinoline alkaloid with an engineered split epimerase inverts the stereochemistry of the stereocenter of the 1-benzylisoquinoline alkaloid to the opposite stereochemistry of the stereocenter of the 1-benzylisoquinoline alkaloid more efficiently than using a wild-type epimerase.

[0013] In some examples, an engineered non-plant cell comprises a plurality of coding sequences each encoding an enzyme that is selected from the group of enzymes listed in **Table 3**. In some examples, the heterologous coding sequences may be operably connected. Heterologous coding sequences that are operably connected may be within the same pathway of producing a particular benzylisoquinoline alkaloid product via an engineered epimerase activity, comprising engineered oxidase and reductase enzymes.

[0014] In an additional aspect of the invention, an engineered non-plant cell that produces an increased amount of a bisbenzylisoquinoline alkaloid product using an engineered split epimerase in contrast to the use of a wild-type epimerase is provided. In a further aspect of the invention, an engineered non-plant

cell that produces an increased amount of bisbenzylisoquinoline product using an engineered fused epimerase in contrast to the use of a wild-type epimerase is provided. The bisbenzylisoquinoline alkaloid is produced using a coupling enzyme that is present within the engineered non-plant cell. Additionally, the engineered non-plant cell comprises at least two heterologous coding sequences encoding an engineered split epimerase, having a separate oxidase enzyme and a separate reductase enzyme, used in the production of at least one benzylisoquinoline alkaloid monomer within the engineered non-plant cell. Further, the at least one coupling enzyme dimerizes two benzylisoquinoline alkaloid monomers within the engineered non-plant cell, thereby forming the bisbenzylisoquinoline alkaloid.

[0015] In an additional aspect of the invention, an engineered non-plant cell that produces an increased amount of a pro-morphinan or morphinan alkaloid product using an engineered split epimerase in contrast to the use of a wild-type epimerase is provided. In a further aspect of the invention, an engineered non-plant cell that produces an increased amount of a pro-morphinan or morphinan alkaloid product using an engineered fused epimerase in contrast to the use of a wild-type epimerase is provided. The engineered non-plant cell comprises at least two heterologous coding sequences encoding an engineered split epimerase, having a separate oxidase enzyme and a separate reductase enzyme, used in the production of at least one morphinan alkaloid within the engineered non-plant cell.

[0016] In an additional aspect of the invention, an engineered non-plant cell that produces an increased amount of a nal-opioid or nor-opioid product using an engineered split epimerase in contrast to the use of a wild-type epimerase is provided. In a further aspect of the invention, an engineered non-plant cell that produces an increased amount of a nal-opioid or nor-opioid product using an engineered fused epimerase in contrast to the use of a wild-type epimerase is provided. The engineered non-plant cell comprises at least two heterologous coding sequences encoding an engineered split epimerase, having a separate oxidase enzyme and a separate reductase enzyme, used in the production of at least one nal-opioid or nor-opioid within the engineered non-plant cell.

[0017] In some embodiments, the disclosure provides methods for increasing production of diverse alkaloid products through the conversion of a promorphinan alkaloid to a morphinan alkaloid via thebaine synthases in an engineered host cell. In further embodiments, the present disclosure provides methods for increasing production of diverse alkaloid products through the conversion of salutaridinol-7-*O*-acetate to thebaine via a thebaine synthase. Examples of parent thebaine synthases comprise amino acid sequences selected from the group consisting of: SEQ ID NOs: 30, 31, 32, 33, and 34, as listed in **Table 2**.

[0018] In some embodiments, the disclosure provides methods for increasing production of diverse alkaloid products through the conversion of a promorphinan alkaloid to a morphinan alkaloid via engineered thebaine synthases in an engineered host cell. In further embodiments, the present disclosure provides methods for increasing production of diverse alkaloid products through the conversion of salutaridinol-7-*O*-acetate to thebaine via an engineered thebaine synthase.

[0019] In some embodiments, the engineered thebaine synthase is a fusion enzyme. In further embodiments, the thebaine synthase is fused to an acetyl transferase enzyme. In further embodiments, the thebaine synthase is encoded within an acetyl transferase enzyme. In other embodiments, the thebaine synthase is fused to a reductase enzyme.

INCORPORATION BY REFERENCE

[0020] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

[0021] The novel features of the invention are set forth with particularity in the appended claims. A better understanding of the features and advantages of the invention will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the invention are utilized, and the accompanying drawings of which:

[0022] **FIG. 1** illustrates examples of synthesis, recycling, and salvage pathways of tetrahydrobiopterin, in accordance with embodiments of the invention.

[0023] **FIG. 2** illustrates a biosynthetic scheme for conversion of glucose to 4-HPA, dopamine, and 3,4-DHPA, in accordance with embodiments of the invention.

[0024] **FIG. 3** illustrates a schematic example of (*R*)-1-benzylisoquinoline alkaloid formation, in accordance with embodiments of the invention.

[0025] **FIG. 4** illustrates an amino acid sequence of a parent DRS-DRR fusion enzyme, in accordance with embodiments of the invention.

[0026] **FIG. 5** illustrates amino acid sequences of a DRS enzyme and a DRR enzyme, respectively, that are derived from a parent fusion enzyme illustrated in **FIG. 4**, in accordance with embodiments of the invention.

[0027] **FIG. 6** illustrates an enzyme having opioid 3-*O*-demethylase activity, in accordance with embodiments of the invention.

[0028] **FIG. 7** illustrates an enzyme having opioid *N*-demethylase activity, in accordance with embodiments of the invention.

[0029] **FIG. 8** illustrates an enzyme having *N*-methyltransferase activity, in accordance with embodiments of the invention.

[0030] **FIG. 9** illustrates the functional expression of BM3 variants, in accordance with embodiments of the invention.

[0031] **FIG. 10** illustrates a biosynthesis scheme for conversion of L-tyrosine to a nor-opioid or nal-opioid in a microbial cell, in accordance with embodiments of the invention.

[0032] **FIG. 11** illustrates plasmid/YAC vectors comprising codon-optimized amino acid sequence of DRS-DRR and DRS and DRR enzymes for enzyme expression and engineering, in accordance with embodiments of the invention.

[0033] **FIG. 12** illustrates a biosynthetic scheme for conversion of L-tyrosine to reticuline via norcoclaurine, in accordance with embodiments of the invention.

[0034] **FIG. 13** illustrates a biosynthetic scheme for conversion of L-tyrosine to reticuline via norlaudanosoline, in accordance with embodiments of the invention.

[0035] **FIG. 14** illustrates a biosynthetic scheme for conversion of L-tyrosine to morphinan alkaloids, in accordance with embodiments of the invention.

[0036] **FIG. 15** illustrates a biosynthetic scheme for production of semi-synthetic opioids, in accordance with embodiments of the invention.

[0037] **FIG. 16** illustrates a biosynthetic scheme for production of opioids, in accordance with embodiments of the invention.

[0038] **FIG. 17** illustrates (A) a biosynthetic scheme for conversion of L-tyrosine to bisBIAs and (B) yeast strains engineered to biosynthesize bisBIAs, in accordance with embodiments of the invention.

[0039] **FIG. 18** illustrates expressing separated DRS and DRR produced greater amounts of salutaridine in comparison to the fused parent, in accordance with embodiments of the invention.

[0040] **FIG. 19** illustrates the concentration of nearly all pathway intermediates was increased in the strain expressing separated DRS and DRR enzymes relative to the strain expressing the fused parent, in accordance with embodiments of the invention.

[0041] **FIG. 20** illustrates chiral separation of reticuline, in accordance with embodiments of the invention.

[0042] **FIG. 21** illustrates when DRS expression is high (TDH3 promoter), and DRR expression is low (either CYC1 promoter), a maximum amount of salutaridine around 20 μ M was produced, in accordance with embodiments of the invention.

[0043] **FIGs. 22 and 23** each illustrate a homology model of DRS based on 4NKX, in accordance with embodiments of the invention.

[0044] **FIG. 24** illustrates an alignment between PbDRS-DRR, PrDRS, and PrDRR, in accordance with embodiments of the invention.

[0045] FIG. 25 illustrates expressing separated variants of DRS and DRR, wherein the DRS variant is engineered with enhancing mutations produced greater amounts of thebaine in comparison to the parent DRS variant, in accordance with embodiments of the invention.

[0046] FIG. 26 illustrates yeast platform strains for the production of reticuline from L-tyrosine, in accordance with embodiments of the invention.

[0047] FIG. 27 illustrates yeast strains for the production of thebaine and hydrocodone from L-tyrosine, in accordance with embodiments of the invention.

[0048] FIG. 28 illustrates the general ring closure reaction converting a tetracyclic scaffold to a pentacyclic scaffold, in accordance with embodiments of the invention.

[0049] FIG. 29 illustrates the production of the morphinan alkaloid thebaine from sugar and L-tyrosine from an engineered yeast strain, in accordance with embodiments of the invention.

[0050] FIG. 30 illustrates the production of promorphinan alkaloids and a morphinan alkaloid thebaine from sugar and L-tyrosine from an engineered yeast strain, in accordance with embodiments of the invention.

DETAILED DESCRIPTION OF THE INVENTION

[0051] The disclosure provides methods for increased production of diverse alkaloid products, in comparison to the use of a fused DRS-DRR epimerase, through the independent expression of dehydroreticuline synthase (DRS) and dehydroreticuline reductase (DRR) enzymes and the use of DRS and DRR enzymes in epimerization of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid. The present disclosure also provides methods for the production of diverse benzylisoquinoline alkaloids (BIAs) in engineered host cells using engineered epimerases. Additionally, the present disclosure provides methods for the production of a set of engineered epimerases in engineered host cells. In some cases, the DRS and DRR enzymes may be derived from a parent epimerase. In some cases, the DRS and DRR enzymes may each be derived from a separate parent epimerase. In some cases, the DRS and DRR enzymes may be derived from one or more parent epimerases. As discussed herein, the independent expression of DRS and DRR, when compared to the expression of fused DRS-DRR, substantially improves the conversion of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid, including (*S*)-reticuline to (*R*)-reticuline. In particular cases, the disclosure provides methods for producing bisbenzylisoquinoline, promorphinan, morphinan, nal-opioid, and nor-opioid alkaloid products through the increased conversion of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid in an engineered host cell. In further particular cases, the disclosure provides methods for producing promorphinan, morphinan, nal-opioid, and nor-opioid alkaloid products through the increased conversion of (*S*)-reticuline to (*R*)-reticuline in an engineered host cell. In further particular cases, the present disclosure provides methods for producing diverse alkaloid products through the increased conversion of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid.

[0052] The conversion of (*S*)-reticuline to (*R*)-reticuline is an important step in the production of morphinan alkaloids in opium poppies. In recent years, several different academic groups independently identified fused DRS-DRR variants that carry out this step (Winzer et al. 2015, Farrow et al. 2015, Galanie et al. 2015). A notable feature of this enzyme is that in the morphinan alkaloid-producing plants, it is found as a translational fusion between two enzymes, a cytochrome P450 and an aldo-keto reductase. In closely related plants that do not produce these compounds, these enzymes are not fused and are thus expressed separately. It is thought that the fusion of the two enzymes allows for improved substrate channeling, since the product of the first enzyme, which converts (*S*)-reticuline to dehydroreticuline, is the direct substrate of the second enzyme, which converts dehydroreticuline to (*R*)-reticuline (Winzer et al. 2015).

[0053] This application describes protein engineering improvements to DRS-DRR variants that resulted in substantial increases in activity in comparison to a parental molecule. Experimental results show that the separation and independent expression of two enzymes, a cytochrome P450 called dehydroreticuline synthase (DRS), and an aldo-keto reductase called dehydroreticuline reductase (DRR), derived from an evolutionarily fused parent enzyme, substantially improved conversion of (*S*)- to (*R*)-reticuline in a heterologous yeast host cell. Improvements to DRS via mutagenesis and screening also improved activity of the enzyme.

[0054] The examples discussed herein show that the separation and expression of independent DRS and DRR enzymes, derived from one or more parental epimerases, in a strain harboring a benzyloquinoline alkaloid (BIA) pathway, resulted in a substantial increase in salutaridine production, when salutaridine synthase (SalSyn) activity was not limiting relative to the amount of (*R*)-reticuline produced. The strain expressing DRS and DRR separately further showed a quantitative increase in all BIA pathway intermediates up to salutaridine. Upon chiral separation of reticuline from yeast, cells that expressed DRS and DRR separately produced predominantly (*R*)-reticuline, in comparison to a parental DRS-DRR fusion enzyme, which produced equivalent amounts of (*S*)- and (*R*)-reticuline. To investigate which of the independent DRS and DRR activities was limiting, a series of vectors with varying promoter strengths was constructed. Differential expression of DRS and DRR showed that intracellular DRS activity was limiting relative to DRR. To better understand the function of DRS in yeast, and to assign function to beneficial residues identified by protein engineering improvements, a structural homology model was built based on the existing structure of human steroidogenic cytochrome P450 CYP17A1. To improve intracellular DRS activity, the ORF of DRS was randomly mutagenized and nearly 5,000 independent mutant variants were screened for salutaridine production. From these variants, a combinatorial variant containing 6 different mutations was identified and selected for introduction and expression in BIA production strains.

[0055] In particular cases, the disclosure provides methods for producing promorphinan, morphinan, nal-opioid and nor-opioid alkaloid products through the increased conversion of a promorphinan alkaloid to a morphinan alkaloid in an engineered host cell. In further particular cases, the disclosure provides

methods for producing morphinan, nal-opioid and nor-opioid alkaloid products through the increased conversion of a promorphinan alkaloid to a morphinan alkaloid in an engineered host cell. In further particular cases, the present disclosure provides methods for producing diverse alkaloid products through the increased conversion of a promorphinan alkaloid to a morphinan alkaloid.

Benzylisoquinoline Alkaloids (BIAs) of Interest

[0056] Host cells which produce BIAs of interest are provided. In some examples, engineered strains of host cells such as the engineered strains of the invention provide a platform for producing benzylisoquinoline alkaloids of interest and modifications thereof across several structural classes including, but not limited to, precursor BIAs, benzylisoquinolines, promorphinans, morphinans, bisbenzylisoquinolines, nal-opioids, nor-opioids, and others. Each of these classes is meant to include biosynthetic precursors, intermediates, and metabolites thereof, of any convenient member of an engineered host cell biosynthetic pathway that may lead to a member of the class. Non-limiting examples of compounds are given below for each of these structural classes. In some cases, the structure of a given example may or may not be characterized itself as a benzylisoquinoline alkaloid. The present chemical entities are meant to include all possible isomers, including single enantiomers, racemic mixtures, optically pure forms, mixtures of diastereomers, and intermediate mixtures.

[0057] Benzylisoquinoline alkaloid precursors may include, but are not limited to, norcoclaurine (NC) and norlaudanoline (NL), as well as NC and NL precursors, such as tyrosine, tyramine, 4-hydroxyphenylacetaldehyde (4-HPA), 4-hydroxyphenylpyruvic acid (4-HPPA), L-3,4-dihydroxyphenylalanine (L-DOPA), 3,4-dihydroxyphenylacetaldehyde (3,4-DHPA), and dopamine. In some embodiments, the one or more BIA precursors are 3,4-dihydroxyphenylacetaldehyde (3,4-DHPA) and dopamine. In certain instances, the one or more BIA precursors are 4-hydroxyphenylacetaldehyde (4-HPA) and dopamine. In particular, NL and NC may be synthesized, respectively, from precursor molecules via a Pictet-Spengler condensation reaction, where the reaction may occur spontaneously or may be catalyzed by any convenient enzymes.

[0058] Benzylisoquinolines may include, but are not limited to, norcoclaurine, norlaudanoline, coclaurine, 3'-hydroxycoclaurine, 4'-*O*-methylnorlaudanoline, 4'-*O*-methyl-laudanosoline, *N*-methylnorcoclaurine, laudanoline, *N*-methylcoclaurine, 3'-hydroxy-*N*-methylcoclaurine, reticuline, norreticuline, papaverine, laudanine, laudanoline, tetrahydropapaverine, 1,2-dihydropapaverine, and orientaline.

[0059] Promorphinans may include, but are not limited to, salutaridine, salutaridinol, and salutaridinol-7-*O*-acetate.

[0060] Morphinans may include, but are not limited to, thebaine, codeinone, codeine, morphine, morphinone, oripavine, neopinone, neopine, neomorphine, hydrocodone, dihydrocodeine, 14-hydroxycodeinone, oxycodone, 14-hydroxycodone, morphinone, hydromorphone, dihydromorphone,

dihydroetorphine, ethylmorphine, etorphine, metopon, buprenorphine, pholcodine, heterocodeine, and oxymorphone.

[0061] Bisbenzylisoquinolines may include, but are not limited to, berbaminine, guattegaumerine, dauricine, and liensinine.

[0062] Nal-opioids may include, but are not limited to, naltrexone, naloxone, nalmeferine, nalorphine, nalorphine, nalodeine, naldemedine, naloxegol, 6 β -naltrexol, naltrindole, methylnaltrexone, methylsamidorphan, alvimopan, axelopran, bevenpran, dinicotinate, levallorphan, samidorphan, buprenorphine, dezocine, eptazocine, butorphanol, levorphanol, nalbuphine, pentazocine, phenazocine, norbinaltorphimine, and diprenorphine.

[0063] Nor-opioids may include, but are not limited to, norcodeine, noroxycodone, northebaine, norhydrocodone, nordihydro-codeine, nor-14-hydroxy-codeine, norcodeinone, nor-14-hydroxy-codeinone, normorphine, noroxymorphone, nororipavine, norhydro-morphine, nordihydro-morphine, nor-14-hydroxy-morphine, normorphinone, and nor-14-hydroxy-morphinone.

[0064] In certain embodiments, the engineered strains of the invention may provide a platform for producing compounds related to tetrahydrobiopterin synthesis including, but not limited to, dihydroneopterin triphosphate, 6-pyruvoyl tetrahydropterin, 5,6,7,8-tetrahydrobiopterin, 7,8-dihydrobiopterin, tetrahydrobiopterin 4a-carbinolamine, quinonoid dihydrobiopterin, and biopterin.

Host Cells

[0065] Any convenient cells may be utilized in the subject host cells and methods. In some cases, the host cells are non-plant cells. In some instances, the host cells may be characterized as microbial cells. In certain cases, the host cells are insect cells, mammalian cells, bacterial cells, fungal cells, or yeast cells. Any convenient type of host cell may be utilized in producing the subject BIA-producing cells, see, e.g., US2008/0176754, US2014/0273109, PCT/US2014/063738, PCT/US2016/030808, PCT/US2015/060891, PCT/US2016/031506, and PCT/US2017/057237, the disclosures of which are incorporated by reference in their entirety. Host cells of interest include, but are not limited to, bacterial cells, such as *Bacillus subtilis*, *Escherichia coli*, *Streptomyces*, *Anabaena*, *Arthrobacter*, *Acetobacter*, *Acetobacterium*, *Bacillus*, *Bifidobacterium*, *Brachybacterium*, *Brevibacterium*, *Carnobacterium*, *Clostridium*, *Corynebacterium*, *Enterobacter*, *Escherichia*, *Gluconacetobacter*, *Gluconobacter*, *Hafnia*, *Halomonas*, *Klebsiella*, *Kocuria*, *Lactobacillus*, *Leuconostoc*, *Macrococcus*, *Methylobacter*, *Methylocella*, *Methylococcus*, *Microbacterium*, *Micrococcus*, *Microcystis*, *Moorella*, *Oenococcus*, *Pediococcus*, *Prochlorococcus*, *Propionibacterium*, *Proteus*, *Pseudoalteromonas*, *Pseudomonas*, *Psychrobacter*, *Rhodobacter*, *Rhodococcus*, *Rhodopseudomonas*, *Serratia*, *Staphylococcus*, *Streptococcus*, *Streptomyces*, *Synechococcus*, *Synechocystis*, *Tetragenococcus*, *Weissella*, *Zymomonas*, and *Salmonella typhimurium* cells, insect cells such as *Drosophila melanogaster* S2 and *Spodoptera frugiperda* Sf9 cells, and yeast cells such as *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and *Pichia pastoris* cells. In some examples, the host cells are yeast cells or *E. coli* cells. In some cases, the host cell is a yeast cell. In some

instances, the host cell is from a strain of yeast engineered to produce a BIA of interest, such as a (*R*)-1-benzylisoquinoline alkaloid. In some instances, the host cell is from a strain of yeast engineered to produce enzymes of interest. In some instances, the host cell is from a strain of yeast engineered to produce an engineered epimerase. In some embodiments, an engineered epimerase may be an engineered split epimerase. In some embodiments, an engineered epimerase may be an engineered fused epimerase. In some embodiments, epimerase activity may be encoded by separate oxidase and reductase enzymes. Additionally, in some embodiments an engineered epimerase may be able to more efficiently convert a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid relative to a parent epimerase. In some embodiments, a parent epimerase may be a wild-type epimerase. In some embodiments, a parent epimerase may be substantially similar to a wild-type epimerase. In some cases, a parent epimerase that is substantially similar to a wild-type epimerase may have an amino acid sequence that is at least 75% or more, 80% or more, 81% or more, 82% or more, 83% or more, 84% or more, 85% or more, 86% or more, 87% or more, 88% or more, 89% or more, 90% or more, 91% or more, 92% or more, 93% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more, or 99% or more similar to an amino acid sequence of a wild-type epimerase. In some embodiments, an engineered epimerase may be separated into smaller enzymes that exhibit oxidase and reductase activities that more efficiently convert a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid relative to its corresponding parent epimerase.

[0066] In some instances, the host cell is from a strain of yeast engineered to produce a thebaine synthase. The thebaine synthase may be able to more efficiently convert a salutaridinol-7-*O*-acetate to a thebaine relative to a spontaneous reaction. In some instances, the host cell is from a strain of yeast engineered to produce an engineered thebaine synthase. In some embodiments, an engineered thebaine synthase may be an engineered fusion enzyme. Additionally, the engineered thebaine synthase may be able to more efficiently convert a salutaridinol-7-*O*-acetate to a thebaine relative to a parent thebaine synthase. In some embodiments, the parent thebaine synthase may be a wild-type thebaine synthase. In some embodiments, a parent thebaine synthase may be substantially similar to a wild-type thebaine synthase. In some cases, a parent thebaine synthase that is substantially similar to a wild-type thebaine synthase may have an amino acid sequence that is at least 75% or more, 80% or more, 81% or more, 82% or more, 83% or more, 84% or more, 85% or more, 86% or more, 87% or more, 88% or more, 89% or more, 90% or more, 91% or more, 92% or more, 93% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more, or 99% or more similar to an amino acid sequence of a wild-type thebaine synthase. The engineered thebaine synthase may be engineered as a fusion enzyme to another enzyme to more efficiently convert a salutaridinol-7-*O*-acetate to a thebaine relative to the parent thebaine synthase.

[0067] Any of the host cells described in US2008/0176754, US2014/0273109, PCTUS2014/063738, PCT/US2016/030808, PCT/US2015/060891, PCT/US2016/031506, PCT/US2017/057237, and US Provisional Application No. 62/627,264 by Smolke et al. may be adapted for use in the subject cells and methods. In certain embodiments, the yeast cells may be of the species *Saccharomyces cerevisiae* (*S*).

cerevisiae). In certain embodiments, the yeast cells may be of the species *Schizosaccharomyces pombe*. In certain embodiments, the yeast cells may be of the species *Pichia pastoris*. Yeast is of interest as a host cell because cytochrome P450 proteins are able to fold properly into the endoplasmic reticulum membrane so that their activity is maintained. In examples, cytochrome P450 proteins are involved in some biosynthetic pathways of interest. In additional examples, cytochrome P450 proteins are involved in the production of BIAs of interest. In further examples, cytochrome P450 proteins are involved in the production of an enzyme of interest.

[0068] Yeast strains of interest that find use in the invention include, but are not limited to, CEN.PK (Genotype: *MATa/α ura3-52/ura3-52 trp1-289/trp1-289 leu2-3_112/leu2-3_112 his3 Δ1/his3 Δ1 MAL2-8C/MAL2-8C SUC2/SUC2*), S288C, W303, D273-10B, X2180, A364A, Σ1278B, AB972, SK1, and FL100. In certain cases, the yeast strain is any of S288C (*MATα; SUC2 mal mel gal2 CUP1 flo1 flo8-1 hap1*), BY4741 (*MATα; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0*), BY4742 (*MATα; his3Δ1; leu2Δ0; lys2Δ0; ura3Δ0*), BY4743 (*MATa/MATα; his3Δ1/his3Δ1; leu2Δ0/leu2Δ0; met15Δ0/MET15; LYS2/lys2Δ0; ura3Δ0/ura3Δ0*), and WAT11 or *W(R)*, derivatives of the W303-B strain (*MATa; ade2-1; his3-11, -15; leu2-3, -112; ura3-1; canR; cyr+*) which express the *Arabidopsis thaliana* NADPH-P450 reductase ATR1 and the yeast NADPH-P450 reductase CPR1, respectively. In another embodiment, the yeast cell is W303alpha (*MATα; his3-11,15 trp1-1 leu2-3 ura3-1 ade2-1*). The identity and genotype of additional yeast strains of interest may be found at EUROSCARF (web.uni-frankfurt.de/fb15/mikro/euroscarf/col_index.html).

[0069] In some instances, the host cell is a fungal cell. In certain embodiments, the fungal cells may be of the *Aspergillus* species and strains include *Aspergillus niger* (ATCC 1015, ATCC 9029, CBS 513.88), *Aspergillus oryzae* (ATCC 56747, RIB40), *Aspergillus terreus* (NIH 2624, ATCC 20542) and *Aspergillus nidulans* (FGSC A4).

[0070] In certain embodiments, heterologous coding sequences may be codon optimized for expression in *Aspergillus* sp. and expressed from an appropriate promoter. In certain embodiments, the promoter may be selected from phosphoglycerate kinase promoter (PGK), MbfA promoter, cytochrome c oxidase subunit promoter (CoxA), SrpB promoter, TvdA promoter, malate dehydrogenase promoter (MdhA), beta-mannosidase promoter (ManB). In certain embodiments, a terminator may be selected from glucoamylase terminator (GlaA) or TrpC terminator. In certain embodiments, the expression cassette consisting of a promoter, heterologous coding sequence, and terminator may be expressed from a plasmid or integrated into the genome of the host. In certain embodiments, selection of cells maintaining the plasmid or integration cassette may be performed with antibiotic selection such as hygromycin or nitrogen source utilization, such as using acetamide as a sole nitrogen source. In certain embodiments, DNA constructs may be introduced into the host cells using established transformation methods such as protoplast transformation, lithium acetate, or electroporation. In certain embodiments, cells may be cultured in liquid ME or solid MEA (3 % malt extract, 0.5 % peptone, and ±1.5 % agar) or in Vogel's minimal medium with or without selection.

[0071] In some instances, the host cell is a bacterial cell. The bacterial cell may be selected from any bacterial genus. Examples of genres from which the bacterial cell may come include *Anabaena*, *Arthrobacter*, *Acetobacter*, *Acetobacterium*, *Bacillus*, *Bifidobacterium*, *Brachybacterium*, *Brevibacterium*, *Carnobacterium*, *Clostridium*, *Corynebacterium*, *Enterobacter*, *Escherichia*, *Gluconacetobacter*, *Gluconobacter*, *Hafnia*, *Halomonas*, *Klebsiella*, *Kocuria*, *Lactobacillus*, *Leuconostoc*, *Macrococcus*, *Methylobacter*, *Methylocella*, *Methylococcus*, *Microbacterium*, *Micrococcus*, *Microcystis*, *Moorella*, *Oenococcus*, *Pediococcus*, *Prochlorococcus*, *Propionibacterium*, *Proteus*, *Pseudoalteromonas*, *Pseudomonas*, *Psychrobacter*, *Rhodobacter*, *Rhodococcus*, *Rhodopseudomonas*, *Serratia*, *Staphylococcus*, *Streptococcus*, *Streptomyces*, *Synechococcus*, *Synechocystis*, *Tetragenococcus*, *Weissella*, and *Zymomonas*. Examples of bacterial species which may be used with the methods of this disclosure include *Arthrobacter nicotianae*, *Acetobacter aceti*, *Arthrobacter arilaitensis*, *Bacillus cereus*, *Bacillus coagulans*, *Bacillus licheniformis*, *Bacillus pumilus*, *Bacillus sphaericus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, *Bifidobacterium adolescentis*, *Brachybacterium tyrofermentans*, *Brevibacterium linens*, *Carnobacterium divergens*, *Corynebacterium flavescens*, *Enterococcus faecium*, *Gluconacetobacter europaeus*, *Gluconacetobacter johannae*, *Gluconobacter oxydans*, *Hafnia alvei*, *Halomonas elongata*, *Kocuria rhizophila*, *Lactobacillus acidifarinae*, *Lactobacillus jensenii*, *Lactococcus lactis*, *Lactobacillus yamanashiensis*, *Leuconostoc citreum*, *Macrococcus caseolyticus*, *Microbacterium foliorum*, *Micrococcus lylae*, *Oenococcus oeni*, *Pediococcus acidilactici*, *Propionibacterium acidipropionici*, *Proteus vulgaris*, *Pseudomonas fluorescens*, *Psychrobacter celer*, *Staphylococcus condimenti*, *Streptococcus thermophilus*, *Streptomyces griseus*, *Tetragenococcus halophilus*, *Weissella cibaria*, *Weissella koreensis*, *Zymomonas mobilis*, *Corynebacterium glutamicum*, *Bifidobacterium bifidum/breve/longum*, *Streptomyces lividans*, *Streptomyces coelicolor*, *Lactobacillus plantarum*, *Lactobacillus sakei*, *Lactobacillus casei*, *Pseudoalteromonas citrea*, *Pseudomonas putida*, *Clostridium ljungdahlii/acetikum/acetobutylicum/beijerinckii/butyricum*, and *Moorella themocellum/thermoacetica*.

[0072] In certain embodiments, the bacterial cells may be of a strain of *Escherichia coli*. In certain embodiments, the strain of *E. coli* may be selected from BL21, DH5 α , XL1-Blue, HB101, BL21, and K12. In certain embodiments, heterologous coding sequences may be codon optimized for expression in *E. coli* and expressed from an appropriate promoter. In certain embodiments, the promoter may be selected from T7 promoter, tac promoter, trc promoter, tetracycline-inducible promoter (tet), lac operon promoter (lac), lacO1 promoter. In certain embodiments, the expression cassette consisting of a promoter, heterologous coding sequence, and terminator may be expressed from a plasmid or integrated into the genome. In certain embodiments, the plasmid is selected from pUC19 or pBAD. In certain embodiments, selection of cells maintaining the plasmid or integration cassette may be performed with antibiotic selection such as kanamycin, chloramphenicol, streptomycin, spectinomycin, gentamycin, erythromycin or ampicillin. In certain embodiments, DNA constructs may be introduced into the host cells using established transformation methods such as conjugation, heat shock chemical transformation, or

electroporation. In certain embodiments, cells may be cultured in liquid Luria–Bertani (LB) media at about 37°C with or without antibiotics.

[0073] In certain embodiments, the bacterial cells may be a strain of *Bacillus subtilis*. In certain embodiments, the strain of *B. subtilis* may be selected from 1779, GP25, RO-NN-1, 168, BSo5, BEST195, 1A382, and 62178. In certain embodiments, heterologous coding sequences may be codon optimized for expression in *Bacillus* sp. and expressed from an appropriate promoter. In certain embodiments, the promoter may be selected from *grac* promoter, *p43* promoter, or *trnQ* promoter. In certain embodiments, the expression cassette consisting of the promoter, heterologous coding sequence, and terminator may be expressed from a plasmid or integrated into the genome. In certain embodiments, the plasmid is selected from pHP13 pE194, pC194, pHT01, or pHT43. In certain embodiments, integrating vectors such as pDG364 or pDG1730 may be used to integrate the expression cassette into the genome. In certain embodiments, selection of cells maintaining the plasmid or integration cassette may be performed with antibiotic selection such as erythromycin, kanamycin, tetracycline, and spectinomycin. In certain embodiments, DNA constructs may be introduced into the host cells using established transformation methods such as natural competence, heat shock, or chemical transformation. In certain embodiments, cells may be cultured in liquid Luria–Bertani (LB) media at 37°C or M9 medium plus glucose and tryptophan.

Genetic Modifications to Host Cells

[0074] The host cells may be engineered to include one or more modifications (such as two or more, three or more, four or more, five or more, or even more modifications) that provide for the production of BIAs of interest. Additionally or alternatively, the host cells may be engineered to include one or more modifications (such as two or more, three or more, four or more, five or more, or even more modifications) that provide for the production of enzymes of interest. In some cases, a modification is a genetic modification, such as a mutation, addition, or deletion of a gene or fragment thereof, or transcription regulation of a gene or fragment thereof. As used herein, the term "mutation" refers to a deletion, insertion, or substitution of an amino acid(s) residue or nucleotide(s) residue relative to a reference sequence or motif. The mutation may be incorporated as a directed mutation to the native gene at the original locus. In some cases, the mutation may be incorporated as an additional copy of the gene introduced as a genetic integration at a separate locus, or as an additional copy on an episomal vector such as a 2 μ or centromeric plasmid. In certain instances, the substrate inhibited copy of the enzyme is under the native cell transcriptional regulation. In some instances, the substrate inhibited copy of the enzyme is introduced with engineered constitutive or dynamic regulation of protein expression by placing it under the control of a synthetic promoter. In some examples, the object of one or more modifications may be a native gene. In some examples, the object of one or more modifications may be a non-native gene. In some examples, a non-native gene may be inserted into a host cell. In further examples, a non-native gene may be altered by one or more modifications prior to being inserted into a host cell.

[0075] An engineered host cell may overproduce one or more BIAs of interest. By overproduce is meant that the cell has an improved or increased production of a BIA molecule of interest relative to a control cell (e.g., an unmodified cell). By improved or increased production is meant both the production of some amount of the BIA of interest where the control has no BIA of interest production, as well as an increase of about 10% or more, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, such as 2-fold or more, such as 5-fold or more, including 10-fold or more in situations where the control has some BIA of interest production.

[0076] An engineered host cell may overproduce one or more (*S*)-1-benzylisoquinoline alkaloids. In some cases, the engineered host cell may produce some amount of the (*S*)-1-benzylisoquinoline alkaloid of interest where the control has no (*S*)-1-benzylisoquinoline alkaloid production, as well as an increase of about 10% or more, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, such as 2-fold or more, such as 5-fold or more, including 10-fold or more in situations where the control has some (*S*)-1-benzylisoquinoline alkaloid of interest production.

[0077] An engineered host cell may further overproduce one or more (*R*)-1-benzylisoquinoline alkaloids. In some cases, the engineered host cell may produce some amount of the (*R*)-1-benzylisoquinoline alkaloid of interest where the control has no (*R*)-1-benzylisoquinoline alkaloid production, as well as an increase of about 10% or more, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, such as 2-fold or more, such as 5-fold or more, including 10-fold or more in situations where the control has some (*R*)-1-benzylisoquinoline alkaloid of interest production. An engineered host cell may further overproduce one or more of 1-benzylisoquinoline alkaloids.

[0078] An engineered host cell may further overproduce one or more morphinan alkaloids. In some cases, the engineered host cell may produce some amount of the morphinan alkaloid of interest where the control has no morphinan alkaloid production, as well as an increase of about 10% or more, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, such as 2-fold or more, such as 5-fold or more, including 10-fold or more in situations where the control has some morphinan alkaloid of interest production. In some cases, the morphinan alkaloid is formed from a 1-benzylisoquinoline alkaloid product, or derivative thereof, of an epimerization reaction catalyzed by an engineered epimerase within an engineered host cell. The engineered epimerase may comprise two separate enzymes that work to produce an epimerase reaction. An engineered host cell may further overproduce one or more of promorphinan, nor-opioid, or nal-opioid alkaloids.

[0079] In some cases, the engineered host cell having an engineered split epimerase is capable of producing an increased amount of (*R*)-reticuline relative to a host cell having an engineered fused

epimerase. In some cases, the engineered host cell having modifications to an oxidase portion of an engineered epimerase is capable of producing an increased amount of (*R*)-reticuline relative to a control host cell that lacks the one or more modifications to the oxidase portion of the engineered epimerase (e.g., as described herein). In certain instances, the increased amount of (*R*)-reticuline is about 10% or more relative to the control host cell, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, about 2-fold or more, about 5-fold or more, or even about 10-fold or more relative to the control host cell. In some cases, (*R*)-reticuline is the product of an epimerization reaction catalyzed by at least one engineered epimerase within an engineered host cell. In these cases, (*S*)-reticuline may be the substrate of the epimerization reaction.

[0080] In some cases, the engineered host cell is capable of producing an increased amount of thebaine relative to a control host cell that lacks the one or more modifications (e.g., as described herein). In some cases, the engineered host cell having a thebaine synthase is capable of producing an increased amount of thebaine relative to a host cell that lacks a thebaine synthase. In some cases, the engineered host cell having an engineered thebaine synthase is capable of producing an increased amount of thebaine relative to a host cell having a parent thebaine synthase (e.g., as described herein). In certain instances, the increased amount of thebaine is about 10% or more relative to the control host cell, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, about 2-fold or more, about 5-fold or more, or even about 10-fold or more relative to the control host cell. In some cases, thebaine is the product of a thebaine synthase reaction within an engineered host cell. In some cases, thebaine is the product of a thebaine synthase reaction catalyzed by at least one engineered thebaine synthase within an engineered host cell. In these cases, salutaridinol-7-*O*-acetate may be the substrate of the thebaine synthase reaction.

[0081] Additionally, an engineered host cell may overproduce one or more enzymes of interest. By overproduce is meant that the cell has an improved or increased production of an enzyme of interest relative to a control cell (e.g., an unmodified cell). By improved or increased production is meant both the production of some amount of the enzyme of interest where the control has no production, as well as an increase of about 10% or more, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, such as 2-fold or more, such as 5-fold or more, including 10-fold or more in situations where the control has some enzyme of interest production.

[0082] An engineered host cell may overproduce one or more engineered DRS-DRR enzymes. In some cases, the engineered host cell may produce some amount of the engineered DRS-DRR epimerase where the control has no DRS-DRR enzyme production, or where the control has a same level of production of wild-type epimerases in comparison to the engineered host cell, as well as an increase of about 10% or more, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, such as 2-fold or more, such as 5-fold or more,

including 10-fold or more in situations where the control has some DRS-DRR enzyme production. In some cases, an engineered DRS-DRR epimerase may be an engineered split epimerase. In some cases, an engineered DRS-DRR epimerase may be an engineered fused epimerase.

[0083] An engineered host cell may overproduce one or more thebaine synthase enzymes. In some cases, the engineered host cell may produce some amount of the thebaine synthase enzyme where the control has no thebaine synthase enzyme production, as well as an increase of about 10% or more, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, such as 2-fold or more, such as 5-fold or more, including 10-fold or more in situations where the control has some thebaine synthase enzyme production.

[0084] An engineered host cell may overproduce one or more engineered thebaine synthase enzymes. In some cases, the engineered host cell may produce some amount of the engineered thebaine synthase where the control has no thebaine synthase enzyme production, or where the control has a same level of production of wild-type thebaine synthase in comparison to the engineered host cell, as well as an increase of about 10% or more, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, such as 2-fold or more, such as 5-fold or more, including 10-fold or more in situations where the control has some thebaine synthase enzyme production. In some cases, an engineered thebaine synthase may be an engineered fusion enzyme.

[0085] An engineered host cell may further overproduce one or more enzymes that are derived from the thebaine synthase enzyme. In some cases, the engineered host cell may produce some amount of the enzymes that are derived from the thebaine synthase enzyme, where the control has no production of enzymes that are derived from the thebaine synthase enzyme, as well as an increase of about 10% or more, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, such as 2-fold or more, such as 5-fold or more, including 10-fold or more in situations where the control has some production of enzymes that are derived from the thebaine synthase enzyme.

[0086] Additionally, an engineered host cell may overproduce one or more bisbenzylisoquinoline alkaloids (bisBIAs). In particular, an engineered host cell is capable of producing an increased amount of bisbenzylisoquinoline alkaloids (bisBIAs) relative to a control host cell that lacks the one or more modifications (e.g., as described herein), including modifications related to harboring an engineered epimerase. In certain instances, the increased amount of bisBIAs is about 10% or more relative to the control host cell, such as about 20% or more, about 30% or more, about 40% or more, about 50% or more, about 60% or more, about 80% or more, about 100% or more, about 2-fold or more, about 5-fold or more, or even about 10-fold or more relative to the control host cell. In some cases, the bisBIA is formed from at least one BIA monomer that is the product, or derivative thereof, of an epimerization reaction catalyzed by an engineered epimerase within an engineered host cell. The engineered epimerase may

comprise two separate enzymes that work to produce an epimerase reaction. An engineered host cell may further overproduce one or more of cepharanthine, fangchinoline, liensinine, neferine, tubocurarine, dauricine, tetrandrine, curine, berbaminine, guattegaumerine, 2'-norberbaminine, and berbamine.

[0087] In some cases, the one or more (such as two or more, three or more, or four or more) modifications may be selected from: a localization mutation; a cytochrome P450 reductase interaction mutation; an accessibility mutation; an activity enhancing mutation; an engineered fused epimerase modification; an engineered fused thebaine synthase modification; and an engineered split epimerase modification. A cell that includes one or more modifications may be referred to as an engineered cell.

Substrate Inhibition Alleviating Mutations

[0088] In some instances, the engineered host cells are cells that include one or more substrate inhibition alleviating mutations (such as two or more, three or more, four or more, five or more, or even more) in one or more biosynthetic enzyme genes of the cell. In some examples, the one or more biosynthetic enzyme genes are native to the cell (e.g., is present in an unmodified cell). In some examples, the one or more biosynthetic enzyme genes are non-native to the cell. As used herein, the term “substrate inhibition alleviating mutation” refers to a mutation that alleviates a substrate inhibition control mechanism of the cell.

[0089] A mutation that alleviates substrate inhibition reduces the inhibition of a regulated enzyme in the cell of interest relative to a control cell and provides for an increased level of the regulated compound or a downstream biosynthetic product thereof. In some cases, by alleviating inhibition of the regulated enzyme is meant that the IC_{50} of inhibition is increased by 2-fold or more, such as by 3-fold or more, 5-fold or more, 10-fold or more, 30-fold or more, 100-fold or more, 300-fold or more, 1000-fold or more, or even more. By increased level is meant a level that is 110% or more of that of the regulated compound in a control cell or a downstream product thereof, such as 120% or more, 130% or more, 140% or more, 150% or more, 160% or more, 170% or more, 180% or more, 190% or more, or 200% or more, such as at least 3-fold or more, at least 5-fold or more, at least 10-fold or more or even more of the regulated compound in the engineered host cell or a downstream product thereof.

[0090] A variety of substrate inhibition control mechanisms and biosynthetic enzymes in the engineered host cell that are directed to regulation of levels of BIAs of interest, or precursors thereof, may be targeted for substrate inhibition alleviation. The engineered host cell may include one or more substrate inhibition alleviating mutations in one or more biosynthetic enzyme genes. The one or more mutations may be located in any convenient biosynthetic enzyme genes where the biosynthetic enzyme is subject to regulatory control. In some embodiments, the one or more biosynthetic enzyme genes encode one or more tyrosine hydroxylase enzymes. In certain instances, the one or more substrate inhibition alleviating mutations are present in a biosynthetic enzyme gene that is TyrH. In some embodiments, the engineered host cell may include one or more substrate inhibition alleviating mutations in one or more biosynthetic enzyme genes such as one of those genes described in **Table 3**.

[0091] In certain embodiments, the one or more substrate inhibition alleviating mutations are present in the TyrH gene. The TyrH gene encodes tyrosine hydroxylase, which is an enzyme that converts tyrosine to L-DOPA. However, TyrH is inhibited by its substrate, tyrosine. Mammalian tyrosine hydroxylase activity, such as that seen in humans or rats, can be improved through mutations to the TyrH gene that relieve substrate inhibition. In particular, substrate inhibition from tyrosine can be relieved by a point mutation W166Y in the TyrH gene. The point mutation W166Y in the TyrH gene may also improve the binding of the cosubstrate of tyrosine hydroxylase, BH₄, to catalyze the reaction of tyrosine to L-DOPA. The mutants of TyrH, when expressed in yeast strains to produce BIAs from sugar (such as those described in United States Provisional Patent Application Serial No. 61/899,496) can significantly improve the production of BIAs.

[0092] Any convenient numbers and types of mutations may be utilized to alleviate a substrate inhibition control mechanism. In certain embodiments, the engineered host cells of the present invention may include 1 or more, 2 or more, 3 or more, 4 or more, 5 or more, 6 or more, 7 or more, 8 or more, 9 or more, 10 or more, 11 or more, 12 or more, 13 or more, 14 or more, or even 15 or more substrate inhibition alleviating mutations, such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15 substrate inhibition alleviating mutations in one or more biosynthetic enzyme genes within the engineered host cell.

Cofactor Recovery Promoting Mechanisms

[0093] In some instances, the engineered host cells are cells that include one or more cofactor recovery promoting mechanisms (such as two or more, three or more, four or more, five or more, or even more) in one or more biosynthetic enzyme genes of the cell. In some examples, the one or more biosynthetic enzyme genes are native to the cell (e.g., is present in an unmodified cell). In some examples, the one or more biosynthetic enzyme genes are non-native to the cell. As used herein, the term “cofactor recovery promoting mechanism” refers to a mechanism that promotes a cofactor recovery control mechanism of the cell.

[0094] A variety of cofactor recovery control mechanisms and biosynthetic enzymes in the engineered host cell that are directed to regulation of levels of BIAs of interest, or precursors thereof, may be targeted for cofactor recovery promotion. The engineered host cell may include one or more cofactor recovery promoting mechanism in one or more biosynthetic enzyme genes. In examples, the engineered host cell may include a heterologous coding sequence that encodes dihydrofolate reductase (DHFR). When DHFR is expressed, it may convert 7,8-dihydrobiopterin (BH₂) to the tetrahydrobiopterin (BH₄), thereby recovering BH₄ as a TyrH cosubstrate. In some examples, the engineered host cell may include one or more cofactor recovery promoting mechanisms in one or more biosynthetic enzyme genes such as one of those genes described in **Table 2**.

[0095] Any convenient numbers and types of mechanisms may be utilized to promote a cofactor recovery control mechanism. In certain embodiments, the engineered host cells of the present invention may include 1 or more, 2 or more, 3 or more, 4 or more, 5 or more, 6 or more, 7 or more, 8 or more, 9 or

more, 10 or more, 11 or more, 12 or more, 13 or more, 14 or more, or even 15 or more cofactor recovery promoting mechanisms such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15 cofactor recovery promoting mechanisms in one or more biosynthetic enzyme genes within the engineered host cell.

Product Inhibition Alleviating Mutations

[0096] In some instances, the engineered host cells are cells that include one or more product inhibition alleviating mutations (such as two or more, three or more, four or more, five or more, or even more) in one or more biosynthetic enzyme genes of the cell. In some examples, the one or more biosynthetic enzyme genes are native to the cell (e.g., is present in an unmodified cell). In some examples, the one or more biosynthetic enzyme genes are non-native to the cell. As used herein, the term “product inhibition alleviating mutation” refers to a mutation that alleviates a short term and/or long term product inhibition control mechanism of an engineered host cell. Short term product inhibition is a control mechanism of the cell in which there is competitive binding at a cosubstrate binding site. Long term product inhibition is a control mechanism of the cell in which there is irreversible binding of a compound away from a desired pathway.

[0097] A mutation that alleviates product inhibition reduces the inhibition of a regulated enzyme in the cell of interest relative to a control cell and provides for an increased level of the regulated compound or a downstream biosynthetic product thereof. In some cases, by alleviating inhibition of the regulated enzyme is meant that the IC_{50} of inhibition is increased by 2-fold or more, such as by 3-fold or more, 5-fold or more, 10-fold or more, 30-fold or more, 100-fold or more, 300-fold or more, 1000-fold or more, or even more. By increased level is meant a level that is 110% or more of that of the regulated compound in a control cell or a downstream product thereof, such as 120% or more, 130% or more, 140% or more, 150% or more, 160% or more, 170% or more, 180% or more, 190% or more, or 200% or more, such as at least 3-fold or more, at least 5-fold or more, at least 10-fold or more or even more of the regulated compound in the engineered host cell or a downstream product thereof.

[0098] A variety of product inhibition control mechanisms and biosynthetic enzymes in the engineered host cell that are directed to regulation of levels of BIAs of interest may be targeted for product inhibition alleviation. The engineered host cell may include one or more product inhibition alleviating mutations in one or more biosynthetic enzyme genes. The mutation may be located in any convenient biosynthetic enzyme genes where the biosynthetic enzyme is subject to regulatory control. In some embodiments, the one or more biosynthetic enzyme genes encode one or more tyrosine hydroxylase enzymes. In certain instances, the one or more product inhibition alleviating mutations are present in a biosynthetic enzyme gene that is TyrH. In some embodiments, the engineered host cell includes one or more product inhibition alleviating mutations in one or more biosynthetic enzyme genes such as one of those genes described in **Table 3**.

[0099] In certain embodiments, the one or more product inhibition alleviating mutations are present in the TyrH gene. The TyrH gene encodes tyrosine hydroxylase, which is an enzyme that converts tyrosine

to L-DOPA. TyrH requires tetrahydrobiopterin (BH₄) as a cosubstrate to catalyze the hydroxylation reaction. Some microbial strains, such as *Saccharomyces cerevisiae*, do not naturally produce BH₄, but can be engineered to produce this substrate through a four-enzyme synthesis and recycling pathway, as illustrated in **FIG. 1**. **FIG. 1** illustrates examples of synthesis, recycling, and salvage pathways of tetrahydrobiopterin, in accordance with embodiments of the invention. **FIG. 1** provides the use of the enzymes PTPS, pyruvoyl tetrahydropterin synthase; SepR, sepiapterin reductase; PCD, pterin 4a-carbinolamine dehydratase; QDHPR, dihydropteridine reductase; and DHFR, dihydrofolate reductase. Of the enzymes that are illustrated in **FIG. 1**, yeast synthesizes an endogenous GTP cyclohydrolase I. GTP and dihydroneopterin triphosphate are naturally synthesized in yeast. Additionally, other metabolites in **FIG. 1** are not naturally produced in yeast.

[00100] TyrH is inhibited by its product L-DOPA, as well as other catecholamines, particularly dopamine. Mammalian tyrosine hydroxylase activity, such as from humans or rats, can be improved through mutations that relieve product inhibition. For example, short term product inhibition, such as competitive binding at the cosubstrate binding site, can be relieved by a point mutation W166Y on the TyrH gene. In particular, the point mutation W166Y on the TyrH gene may improve binding of the cosubstrate. Additionally, short term product inhibition to relieve competitive binding at the cosubstrate binding site may be improved by a point mutation S40D on the TyrH gene. Short term product inhibition may also be improved by the joint mutations of R37E, R38E on the TyrH gene. In particular, R37E, R38E mutations may together specifically improve tyrosine hydroxylase activity in the presence of dopamine.

[00101] Additionally, long term product inhibition may be relieved by point mutations on the TyrH gene. Long term product inhibition relief may include the irreversible binding of catecholamine to iron in the active site such that there is less catecholamine present to act as a product inhibitor of tyrosine hydroxylase activity. Long term product inhibition can be relieved by the mutations E332D and Y371F, respectively, in the TyrH gene.

[00102] Combinations of the mutations can be made (such as two or three or more mutations at once) to relieve multiple types of substrate and product inhibition to further improve the activity of TyrH. The mutants of TyrH, when expressed in yeast strains to produce BIAs from sugar (such as those described in United States Provisional Patent Application Serial No. 61/899,496) can significantly improve the production of BIAs.

[00103] Any convenient numbers and types of mutations may be utilized to alleviate a product inhibition control mechanism. In certain embodiments, the engineered host cells of the present invention may include 1 or more, 2 or more, 3 or more, 4 or more, 5 or more, 6 or more, 7 or more, 8 or more, 9 or more, 10 or more, 11 or more, 12 or more, 13 or more, 14 or more, or even 15 or more product inhibition alleviating mutations, such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15 product inhibition alleviating mutations in one or more biosynthetic enzyme genes within the engineered host cell.

Feedback Inhibition Alleviating Mutations

[00104] In some instances, the engineered host cells are cells that include one or more feedback inhibition alleviating mutations (such as two or more, three or more, four or more, five or more, or even more) in one or more biosynthetic enzyme genes of the cell. In some cases, the one or more biosynthetic enzyme genes are native to the cell (e.g., is present in an unmodified cell). Additionally or alternatively, in some examples the one or more biosynthetic enzyme genes are non-native to the cell. As used herein, the term “feedback inhibition alleviating mutation” refers to a mutation that alleviates a feedback inhibition control mechanism of an engineered host cell. Feedback inhibition is a control mechanism of the cell in which an enzyme in the synthetic pathway of a regulated compound is inhibited when that compound has accumulated to a certain level, thereby balancing the amount of the compound in the cell. A mutation that alleviates feedback inhibition reduces the inhibition of a regulated enzyme in the engineered host cell relative to a control cell. In this way, engineered host cell provides for an increased level of the regulated compound or a downstream biosynthetic product thereof. In some cases, by alleviating inhibition of the regulated enzyme is meant that the IC_{50} of inhibition is increased by 2-fold or more, such as by 3-fold or more, 5-fold or more, 10-fold or more, 30-fold or more, 100-fold or more, 300-fold or more, 1000-fold or more, or even more. By increased level is meant a level that is 110% or more of that of the regulated compound in a control cell or a downstream product thereof, such as 120% or more, 130% or more, 140% or more, 150% or more, 160% or more, 170% or more, 180% or more, 190% or more, or 200% or more, such as at least 3-fold or more, at least 5-fold or more, at least 10-fold or more or even more of the regulated compound in the host cell or a downstream product thereof.

[00105] A variety of feedback inhibition control mechanisms and biosynthetic enzymes that are directed to regulation of levels of BIAs of interest may be targeted for alleviation in the host cell. The host cell may include one or more feedback inhibition alleviating mutations in one or more biosynthetic enzyme genes native to the cell. The one or more mutations may be located in any convenient biosynthetic enzyme genes where the biosynthetic enzyme is subject to regulatory control. In some embodiments, the one or more biosynthetic enzyme genes may encode one or more enzymes selected from a 3-deoxy-d-arabinose-heptulosonate-7-phosphate (DAHP) synthase and a chorismate mutase. In some embodiments, the one or more biosynthetic enzyme genes encode a 3-deoxy-d-arabinose-heptulosonate-7-phosphate (DAHP) synthase. In some instances, the one or more biosynthetic enzyme genes may encode a chorismate mutase. In certain instances, the one or more feedback inhibition alleviating mutations may be present in a biosynthetic enzyme gene selected from ARO4 and ARO7. In certain instances, the one or more feedback inhibition alleviating mutations may be present in a biosynthetic enzyme gene that is ARO4. In certain instances, the one or more feedback inhibition alleviating mutations are present in a biosynthetic enzyme gene that is ARO7. In some embodiments, the engineered host cell may include one or more feedback inhibition alleviating mutations in one or more biosynthetic enzyme genes such as one of those genes described in **Table 3**.

[00106] Any convenient numbers and types of mutations may be utilized to alleviate a feedback inhibition control mechanism. As used herein, the term "mutation" refers to a deletion, insertion, or substitution of an amino acid(s) residue or nucleotide(s) residue relative to a reference sequence or motif. The mutation may be incorporated as a directed mutation to the native gene at the original locus. In some cases, the mutation may be incorporated as an additional copy of the gene introduced as a genetic integration at a separate locus, or as an additional copy on an episomal vector such as a 2 μ or centromeric plasmid. In certain instances, the feedback inhibited copy of the enzyme is under the native cell transcriptional regulation. In some instances, the feedback inhibited copy of the enzyme is introduced with engineered constitutive or dynamic regulation of protein expression by placing it under the control of a synthetic promoter.

[00107] In certain embodiments, the one or more feedback inhibition alleviating mutations may be present in the ARO4 gene. ARO4 mutations of interest may include, but are not limited to, substitution of the lysine residue at position 229 with a leucine, a substitution of the glutamine residue at position 166 with a lysine residue, or a mutation as described by Hartmann M, et al. ((2003) Proc Natl Acad Sci U S A 100(3):862-867) or Fukuda et al. ((1992) J Ferment Bioeng 74(2):117-119). In some instances, mutations for conferring feedback inhibition may be selected from a mutagenized library of enzyme mutants. Examples of such selections may include rescue of growth of o-fluoro-D,L-phenylalanine or growth of aro3 mutant yeast strains in media with excess tyrosine as described by Fukuda et al. ((1990) Breeding of Brewing Yeast Producing a Large Amount of Beta-Phenylethyl Alcohol and Beta-Phenylethyl Acetate. Agr Biol Chem Tokyo 54(1):269-271).

[00108] In certain embodiments, the engineered host cells of the present invention may include 1 or more, 2 or more, 3 or more, 4 or more, 5 or more, 6 or more, 7 or more, 8 or more, 9 or more, 10 or more, 11 or more, 12 or more, 13 or more, 14 or more, or even 15 or more feedback inhibition alleviating mutations, such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 or 15 feedback inhibition alleviating mutations in one or more biosynthetic enzyme genes within the engineered host cell.

Transcriptional Modulation Modifications

[00109] The host cells may include one or more transcriptional modulation modifications (such as two or more, three or more, four or more, five or more, or even more modifications) of one or more biosynthetic enzyme genes of the cell. In some examples, the one or more biosynthetic enzyme genes are native to the cell. In some examples, the one or more biosynthetic enzyme genes are non-native to the cell. Any convenient biosynthetic enzyme genes of the cell may be targeted for transcription modulation. By transcription modulation is meant that the expression of a gene of interest in a modified cell is modulated, e.g., increased or decreased, enhanced or repressed, relative to a control cell (e.g., an unmodified cell). In some cases, transcriptional modulation of the gene of interest includes increasing or enhancing expression. By increasing or enhancing expression is meant that the expression level of the gene of interest is increased by 2-fold or more, such as by 5-fold or more and sometimes by 25-, 50-, or 100-fold

or more and in certain embodiments 300-fold or more or higher, as compared to a control, i.e., expression in the same cell not modified (e.g., by using any convenient gene expression assay). Alternatively, in cases where expression of the gene of interest in a cell is so low that it is undetectable, the expression level of the gene of interest is considered to be increased if expression is increased to a level that is easily detectable. In certain instances, transcriptional modulation of the gene of interest includes decreasing or repressing expression. By decreasing or repressing expression is meant that the expression level of the gene of interest is decreased by 2-fold or more, such as by 5-fold or more and sometimes by 25-, 50-, or 100-fold or more and in certain embodiments 300-fold or more or higher, as compared to a control. In some cases, expression is decreased to a level that is undetectable. Modifications of host cell processes of interest that may be adapted for use in the subject host cells are described in U.S. Publication No. 20140273109 (14/211,611) by Smolke et al., the disclosure of which is herein incorporated by reference in its entirety.

[00110] Any convenient biosynthetic enzyme genes may be transcriptionally modulated, and include but are not limited to, those biosynthetic enzymes described in **FIG. 2**. In particular, **FIG. 2** illustrates a biosynthetic scheme for conversion of glucose to 4-HPA, dopamine, and 3,4-DHPA, in accordance with embodiments of the invention. Examples of enzymes described in **FIG. 2** include ARO3, ARO4, ARO1, ARO7, TYR1, TYR, TyrH, DODC, MAO, ARO10, ARO9, ARO8, and TKL. In some instances, the one or more biosynthetic enzyme genes may be selected from ARO10, ARO9, ARO8, and TKL. In some cases, the one or more biosynthetic enzyme genes may be ARO10. In certain instances, the one or more biosynthetic enzyme genes may be ARO9. In some embodiments, the one or more biosynthetic enzyme genes may be TKL. In some embodiments, the host cell includes one or more transcriptional modulation modifications to one or more genes such as one of those genes described in **Table 3**.

[00111] In some embodiments, the transcriptional modulation modification may include a substitution of a strong promoter for a native promoter of the one or more biosynthetic enzyme genes or the expression of an additional copy(ies) of the gene or genes under the control of a strong promoter. The promoters driving expression of the genes of interest may be constitutive promoters or inducible promoters, provided that the promoters may be active in the host cells. The genes of interest may be expressed from their native promoters. Additionally or alternatively, the genes of interest may be expressed from non-native promoters. Although not a requirement, such promoters may be medium to high strength in the host in which they are used. Promoters may be regulated or constitutive. In some embodiments, promoters that are not glucose repressed, or repressed only mildly by the presence of glucose in the culture medium, may be used. There are numerous suitable promoters, examples of which include promoters of glycolytic genes such as the promoter of the *B. subtilis* *tsr* gene (encoding fructose biphosphate aldolase) or GAPDH promoter from yeast *S. cerevisiae* (coding for glyceraldehyde-phosphate dehydrogenase) (Bitter G. A., *Meth. Enzymol.* 152:673 684 (1987)). Other strong promoters of interest include, but are not limited to, the ADHI promoter of baker's yeast (Ruohonen L., et al, *J. Biotechnol.* 39:193 203 (1995)), the phosphate-starvation induced promoters such as the PHO5 promoter

of yeast (Hinnen, A., et al, in *Yeast Genetic Engineering*, Barr, P. J., et al. eds, Butterworths (1989), the alkaline phosphatase promoter from *B. licheniformis* (Lee, J. W. K., et al., *J. Gen. Microbiol.* 137:1127-1133 (1991)), GPD1, and TEF1. Yeast promoters of interest include, but are not limited to, inducible promoters such as Gal1-10, Gal1, GalL, GalS, repressible promoter Met25, tetO, and constitutive promoters such as glyceraldehyde 3-phosphate dehydrogenase promoter (GPD), alcohol dehydrogenase promoter (ADH), translation-elongation factor-1-alpha promoter (TEF), cytochrome c-oxidase promoter (CYC1), MRP7 promoter, etc. In some instances, the strong promoter is GPD1. In certain instances, the strong promoter is TEF1. Autonomously replicating yeast expression vectors containing promoters inducible by hormones such as glucocorticoids, steroids, and thyroid hormones are also known and include, but are not limited to, the glucocorticoid responsive element (GRE) and thyroid hormone responsive element (TRE), see e.g., those promoters described in U.S. Pat. No. 7,045,290. Vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. Additionally any promoter/enhancer combination (as per the Eukaryotic Promoter Data Base EPDB) could also be used to drive expression of genes of interest. It is understood that any convenient promoters specific to the host cell may be selected, e.g., *E. coli*. In some cases, promoter selection may be used to optimize transcription, and hence, enzyme levels to maximize production while minimizing energy resources.

Inactivating Mutations

[00112] The engineered host cells may include one or more inactivating mutations to an enzyme of the cell (such as two or more, three or more, four or more, five or more, or even more). The inclusion of one or more inactivating mutations may modify the flux of a synthetic pathway of an engineered host cell to increase the levels of a BIA of interest or a desirable enzyme or precursor leading to the same. In some examples, the one or more inactivating mutations are to an enzyme native to the cell. Additionally or alternatively, the one or more inactivating mutations are to an enzyme non-native to the cell. As used herein, by “inactivating mutation” is meant one or more mutations to a gene or regulatory DNA sequence of the cell, where the mutation(s) inactivates a biological activity of the protein expressed by that gene of interest. In some cases, the gene is native to the cell. In some instances, the gene encodes an enzyme that is inactivated and is part of or connected to the synthetic pathway of a BIA of interest produced by the host cell. In some instances, an inactivating mutation is located in a regulatory DNA sequence that controls a gene of interest. In certain cases, the inactivating mutation is to a promoter of a gene. Any convenient mutations (e.g., as described herein) may be utilized to inactivate a gene or regulatory DNA sequence of interest. By “inactivated” or “inactivates” is meant that a biological activity of the protein expressed by the mutated gene is reduced by 10% or more, such as by 20% or more, 30% or more, 40% or more, 50% or more, 60% or more, 70% or more, 80% or more, 90% or more, 95% or more, 97% or more, or 99% or more, relative to a control protein expressed by a non-mutated control gene. In some cases, the protein is an enzyme and the inactivating mutation reduces the activity of the enzyme.

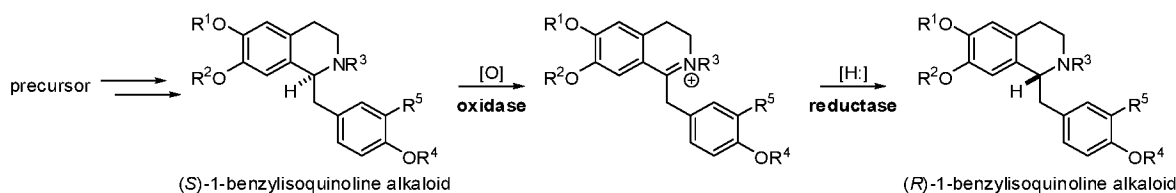
[00113] In some examples, the engineered host cell includes an inactivating mutation in an enzyme native to the cell. Any convenient enzymes may be targeted for inactivation. Enzymes of interest may include, but are not limited to those enzymes, described in **Table 3** whose action in the synthetic pathway of the engineered host cell tends to reduce the levels of a BIA of interest. In some cases, the enzyme has glucose-6-phosphate dehydrogenase activity. In certain embodiments, the enzyme that includes an inactivating mutation is ZWF1. In some cases, the enzyme has alcohol dehydrogenase activity. In some embodiments, the enzyme that includes an inactivating mutation is selected from ADH2, ADH3, ADH4, ADH5, ADH6, ADH7, and SFA1. In certain embodiments, the enzyme that includes an inactivating mutation(s) is ADH2. In certain embodiments, the enzyme that includes an inactivating mutation(s) is ADH3. In certain embodiments, the enzyme that includes an inactivating mutation(s) is ADH4. In certain embodiments, the enzyme that includes an inactivating mutation(s) is ADH5. In certain embodiments, the enzyme that includes an inactivating mutation(s) is ADH6. In certain embodiments, the enzyme that includes an inactivating mutation(s) is ADH7. In some cases, the enzyme has aldehyde oxidoreductase activity. In certain embodiments, the enzyme that includes an inactivating mutation is selected from ALD2, ALD3, ALD4, ALD5, and ALD6. In certain embodiments, the enzyme that includes an inactivating mutation(s) is ALD2. In certain embodiments, the enzyme that includes an inactivating mutation(s) is ALD3. In certain embodiments, the enzyme that includes an inactivating mutation(s) is ALD4. In certain embodiments, the enzyme that includes an inactivating mutation(s) is ALD5. In certain embodiments, the enzyme that includes an inactivating mutation(s) is ALD6. In some cases, the enzyme has aryl-alcohol dehydrogenase activity. In some embodiments, the enzyme that includes an inactivating mutation is selected from AAD4, AAD6, AAD10, AAD14, AAD15, AAD16. In certain embodiments, the enzyme that includes an inactivating mutation(s) is AAD4. In certain embodiments, the enzyme that includes an inactivating mutation(s) is AAD6. In certain embodiments, the enzyme that includes an inactivating mutation(s) is AAD10. In certain embodiments, the enzyme that includes an inactivating mutation(s) is AAD14. In certain embodiments, the enzyme that includes an inactivating mutation(s) is AAD15. In certain embodiments, the enzyme that includes an inactivating mutation(s) is AAD16. In some embodiments, the host cell includes one or more inactivating mutations to one or more genes described in **Table 3**.

Epimerization Modifications

[00114] Some methods, processes, and systems provided herein describe the conversion of (*S*)-1-benzylisoquinoline alkaloids to (*R*)-1-benzylisoquinoline alkaloids. Some of these methods, processes, and systems may comprise an engineered host cell. In some examples, the conversion of (*S*)-1-benzylisoquinoline alkaloids to (*R*)-1-benzylisoquinoline alkaloids is a key step in the conversion of a substrate to a diverse range of alkaloids. In some examples, the conversion of (*S*)-1-benzylisoquinoline alkaloids to (*R*)-1-benzylisoquinoline alkaloids comprises an epimerization reaction via an engineered epimerase. In some cases, epimerization of a substrate alkaloid may be performed by oxidizing an (*S*)-

substrate to the corresponding Schiff base or imine intermediate, then stereospecifically reducing this intermediate to an (*R*)-product as provided in **FIG. 3** and as represented generally in **Scheme 1**. As provided in **Scheme 1**, R₁, R₂, R₃, and R₄ may be H or CH₃. R₅ may be H, OH, or OCH₃.

Scheme 1



[00115] In some examples, the conversion of the (*S*)-substrate to the (*R*)-product may involve at least one oxidation reaction and at least one reduction reaction. In some cases, an oxidation reaction is optionally followed by a reduction reaction. In some cases, at least one of the oxidation and reduction reactions is carried out in the presence of an enzyme. In some cases, at least one of the oxidation and reduction reactions is catalyzed by an engineered epimerase. In some cases, the oxidation and reduction reactions are both carried out in the presence of an engineered fused epimerase. In some cases, the oxidation and reduction reactions are both carried out in the presence of an engineered split epimerase having a separately expressed oxidase component and reductase component, respectively. In some cases, an engineered epimerase is useful to catalyze the oxidation and reduction reactions. The oxidation and reduction reactions may be catalyzed by the same engineered epimerase.

[00116] In some methods, processes and systems described herein, an oxidation reaction may be performed in the presence of an enzyme that is part of an engineered epimerase. In some examples, the engineered epimerase may have an oxidase component. In some cases, the oxidase component may be a component of an engineered fused epimerase. In some case, the oxidase component may be independently expressed as part of an engineered split epimerase. The oxidase may use a (*S*)-1-benzylisoquinoline as a substrate. The oxidase may convert the (*S*)-substrate to a corresponding imine or Schiff base derivative. The oxidase may be referred to as 1,2-dehydroreticuline synthase (DRS). Non-limiting examples of enzymes suitable for oxidation of (*S*)-1-benzylisoquinoline alkaloids in this disclosure include a cytochrome P450 oxidase, a 2-oxoglutarate-dependent oxidase, and a flavoprotein oxidase. For example, (*S*)-tetrahydroprotoberberine oxidase (STOX, E.C 1.3.3.8) may oxidize (*S*)-norreticuline and other (*S*)-1-benzylisoquinoline alkaloids to 1,2-dehydronorreticuline and other corresponding 1,2-dehydro products. In some examples, a protein that comprises an oxidase domain of any one of the preceding examples may perform the oxidation. In some examples, the oxidase may catalyze the oxidation reaction within a host cell, such as an engineered host cell, as described herein. In some cases, the oxidase may have one or more activity-increasing components as discussed herein and as described in **Examples 6 and 8**.

[00117] In some examples, a reduction reaction may follow the oxidation reaction. The reduction reaction may be performed by an enzyme that is part of an engineered epimerase. In some examples, the reductase

may use an imine or Schiff base derived from a 1-benzylisoquinoline as a substrate. The reductase may convert the imine or Schiff base derivative to a (*R*)-1-benzylisoquinoline. The reductase may be referred to as 1,2-dehydroreticuline reductase (DRR). Non-limiting examples of enzymes suitable for reduction of an imine or Schiff base derived from an (*S*)-1-benzylisoquinoline alkaloid include an aldo-keto reductase (e.g., a codeinone reductase-like enzyme (EC 1.1.1.247)) and a short chain dehydrogenase (e.g., a salutaridine reductase-like enzyme (EC 1.1.1.248)). In some examples, a protein that comprises a reductase domain of any one of the preceding examples may perform the reduction. In a further embodiment, the reduction is stereospecific. In some examples, the reductase may catalyze the reduction reaction within a host cell, such as an engineered host cell, as described herein.

[00118] An example of an enzyme that can perform an epimerization reaction that converts (*S*)-1-benzylisoquinoline alkaloids to (*R*)-1-benzylisoquinoline alkaloids includes an epimerase having an oxidase domain and a reductase domain. In particular, the epimerase may have a cytochrome P450 oxidase 82Y2-like domain. Additionally, the epimerase may have a codeinone reductase-like domain. An epimerase having a cytochrome P450 oxidase 82Y2-like domain and also having a codeinone reductase-like domain may be referred to as a DRS-DRR enzyme. In particular, a DRS-DRR enzyme may be a fusion enzyme that is a fusion epimerase. Further, when a DRS-DRR enzyme is modified by at least one activity-increasing modification, the fusion enzyme may be an engineered fusion epimerase. Examples of activity-increasing modifications are discussed in **Examples 6 and 8** below.

[00119] An example of an amino acid sequence of a DRS-DRR enzyme that may be used to perform the conversion of (*S*)-1-benzylisoquinoline alkaloids to (*R*)-1-benzylisoquinoline alkaloids is provided in **FIG. 4**. In particular, **FIG. 4** illustrates an amino acid sequence of a DRS-DRR enzyme that has been codon-optimized, in accordance with embodiments of the invention. Further, **FIG. 5** illustrates a split of an oxidase portion and reductase portion, each of the DRS-DRR enzyme of **FIG. 4**. Additional amino acid sequences of a DRS-DRR enzyme are set forth in **Table 1**. An amino acid sequence for an epimerase that is utilized in converting an (*S*)-1-benzylisoquinoline alkaloid to an (*R*)-1-benzylisoquinoline alkaloid may be 75% or more identical to a given amino acid sequence as listed in **Table 1**. For example, an amino acid sequence for such an epimerase may comprise an amino acid sequence that is at least 75% or more, 80% or more, 81% or more, 82% or more, 83% or more, 84% or more, 85% or more, 86% or more, 87% or more, 88% or more, 89% or more, 90% or more, 91% or more, 92% or more, 93% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more, or 99% or more identical to an amino acid sequence as provided herein. Additionally, in certain embodiments, an "identical" amino acid sequence contains at least 80%-99% identity at the amino acid level to the specific amino acid sequence. In some cases an "identical" amino acid sequence contains at least about 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94% and more in certain cases, at least 95%, 96%, 97%, 98% and 99% identity, at the amino acid level. In some cases, the amino acid sequence may be identical but the DNA sequence is altered such as to optimize codon usage for the host organism, for example.

[00120] Amino acid residues of homologous epimerases may be referenced according to the numbering scheme of SEQ ID NO. 16, and this numbering system is used throughout the disclosure to refer to specific amino acid residues of epimerases which are homologous to SEQ ID NO. 16. Epimerases homologous to SEQ ID NO. 16 may have at least about 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% sequence identity to SEQ ID NO. 16. In some cases, an amino acid referred to as position 50 in a homologous epimerase may not be the 50th amino acid in the homologous epimerase, but would be the amino acid which corresponds to the amino acid at position 50 in SEQ ID NO. 16 in a protein alignment of the homologous epimerase with SEQ ID NO. 16. In some cases, homologous enzymes may be aligned with SEQ ID NO. 16 either according to primary sequence, secondary structure, or tertiary structure.

[00121] An engineered host cell may be provided that produces an engineered epimerase that converts (*S*)-1-benzylisoquinoline alkaloid to (*R*)-1-benzylisoquinoline alkaloid, wherein the epimerase comprises an amino acid sequence selected from the group consisting of: SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, and 18, and having one or more activity-enhancing modifications as described in **Table 13**. The epimerase that is produced within the engineered host cell may be recovered and purified so as to form a biocatalyst. In some cases, the epimerase may be split into one or more enzymes. Additionally, one or more enzymes that are produced by splitting the epimerase may be recovered from the engineered host cell. These one or more enzymes that result from splitting the epimerase may also be used to catalyze the conversion of (*S*)-1-benzylisoquinoline alkaloids to (*R*)-1-benzylisoquinoline alkaloids. Additionally, the use of an engineered split epimerase may be used to increase the production of benzylisoquinoline alkaloid products within a cell when compared to the production of benzylisoquinoline alkaloid products within a cell utilizing a fused epimerase.

[00122] In additional cases, the one or more enzymes that are recovered from the engineered host cell that produces the epimerase may be used in a process for converting a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid. The process may include contacting the (*S*)-1-benzylisoquinoline alkaloid with an epimerase in an amount sufficient to convert said (*S*)-1-benzylisoquinoline alkaloid to (*R*)-1-benzylisoquinoline alkaloid. In examples, the (*S*)-1-benzylisoquinoline alkaloid may be contacted with a sufficient amount of the one or more enzymes such that at least 5% of said (*S*)-1-benzylisoquinoline alkaloid is converted to (*R*)-1-benzylisoquinoline alkaloid. In further examples, the (*S*)-1-benzylisoquinoline alkaloid may be contacted with a sufficient amount of the one or more enzymes such that at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 82%, at least 84%, at least 86%, at least 88%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, at least 99.7%, or 100% of said (*S*)-1-benzylisoquinoline alkaloid is converted to (*R*)-1-benzylisoquinoline alkaloid.

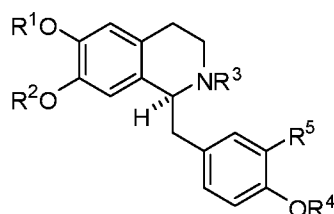
[00123] The one or more enzymes that may be used to convert a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid may contact the (*S*)-1-benzylisoquinoline alkaloid in vitro.

Additionally, or alternatively, the one or more enzymes that may be used to convert a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid may contact the (*S*)-1-benzylisoquinoline alkaloid in vivo. Additionally, the one or more enzymes that may be used to convert a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid may be provided to a cell having the (*S*)-1-benzylisoquinoline alkaloid within, or may be produced within an engineered host cell.

[00124] In some examples, the methods provide for engineered host cells that produce an alkaloid product, wherein the epimerization of a (*S*)-substrate to a (*R*)-product may comprise a key step in the production of an alkaloid product. In some examples, the alkaloid produced is a (*R*)-1-benzylisoquinoline alkaloid. In still other embodiments, the alkaloid produced is derived from a (*R*)-1-benzylisoquinoline alkaloid, including, for example, 4-ring promorphinan and 5-ring morphinan alkaloids. In another embodiment, a (*S*)-1-benzylisoquinoline alkaloid is an intermediate toward the product of the engineered host cell. In still other embodiments, the alkaloid product is selected from the group consisting of 1-benzylisoquinoline, morphinan, promorphinan, nor-opioid, nal-opioid, or bisbenzylisoquinoline alkaloids.

[00125] In some examples, the (*S*)-substrate is a (*S*)-1-benzylisoquinoline alkaloid selected from the group consisting of (*S*)-norreticuline, (*S*)-reticuline, (*S*)-tetrahydropapaverine, (*S*)-norcoclaurine, (*S*)-coclaurine, (*S*)-*N*-methylococlaurine, (*S*)-3'-hydroxy-*N*-methylococlaurine, (*S*)-norisoorientaline, (*S*)-orientaline, (*S*)-isoorientaline, (*S*)-norprotosinomenine, (*S*)-protosinomenine, (*S*)-norlaudanoline, (*S*)-laudanoline, (*S*)-4'-*O*-methyllaudanoline, (*S*)-6-*O*-methyllaudanoline, (*S*)-4'-*O*-methyllaudanoline.

[00126] In some examples, the (*S*)-substrate is a compound of Formula I:



Formula I,

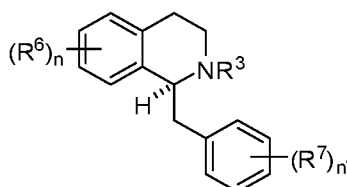
or a salt thereof, wherein:

R^1 , R^2 , R^3 , and R^4 are independently selected from hydrogen and methyl; and

R^5 is selected from hydrogen, hydroxy, and methoxy.

[00127] In some other examples, at least one of R^1 , R^2 , R^3 , R^4 , and R^5 is hydrogen.

[00128] In still other examples, the (*S*)-substrate is a compound of Formula II:



Formula II,

or a salt thereof, wherein:

R³ is selected from hydrogen and C₁-C₄ alkyl;

R⁶ and R⁷ are independently selected at each occurrence from hydroxy, fluoro, chloro, bromo, carboxaldehyde, C₁-C₄ acyl, C₁-C₄ alkyl, and C₁-C₄ alkoxy;

n is 0, 1, 2, 3, or 4; and

n' is 0, 1, 2, 3, 4 or 5.

[00129] When a bond is drawn across a ring, it means substitution may occur at a non-specific ring atom or position. For example, in Formula II shown above, the hydrogen of any –CH– in the 6-membered ring may be replaced with R⁷ to form –CR⁷–.

[00130] In some examples, R⁶ and R⁷ are independently methyl or methoxy. In some other examples, n and n' are independently 1 or 2. In still other embodiments, R³ is hydrogen or methyl.

[00131] In some examples, the methods provide for engineered host cells that produce alkaloid products from (*S*)-reticuline. The epimerization of (*S*)-reticuline to (*R*)-reticuline may comprise a key step in the production of diverse alkaloid products from a precursor. In some examples, the precursor is L-tyrosine or a sugar (e.g., glucose). The diverse alkaloid products can include, without limitation, 1-benzylisoquinoline, morphinan, promorphinan, nor-opioid, or nal-opioid alkaloids.

[00132] Any suitable carbon source may be used as a precursor toward an epimerized 1-benzylisoquinoline alkaloid. Suitable precursors can include, without limitation, monosaccharides (e.g., glucose, fructose, galactose, xylose), oligosaccharides (e.g., lactose, sucrose, raffinose), polysaccharides (e.g., starch, cellulose), or a combination thereof. In some examples, unpurified mixtures from renewable feedstocks can be used (e.g., cornsteep liquor, sugar beet molasses, barley malt, biomass hydrolysate). In still other embodiments, the carbon precursor can be a one-carbon compound (e.g., methanol, carbon dioxide) or a two-carbon compound (e.g., ethanol). In yet other embodiments, other carbon-containing compounds can be utilized, for example, methylamine, glucosamine, and amino acids (e.g., L-tyrosine). In some examples, a 1-benzylisoquinoline alkaloid may be added directly to an engineered host cell of the invention, including, for example, norlaudanosoline, laudanosoline, norreticuline, and reticuline. In still further embodiments, a 1-benzylisoquinoline alkaloid may be added to the engineered host cell as a single enantiomer (e.g., a (*S*)-1-benzylisoquinoline alkaloid), or a mixture of enantiomers, including, for example, a racemic mixture.

[00133] In some examples, the methods provide for the epimerization of a stereocenter of a 1-benzylisoquinoline alkaloid, or a derivative thereof, using an engineered epimerase. In a further embodiment, the method comprises contacting the 1-benzylisoquinoline alkaloid with an engineered epimerase. The engineered epimerase may invert the stereochemistry of a stereocenter of a 1-benzylisoquinoline alkaloid, or derivative thereof, to the opposite stereochemistry. In some examples, the engineered epimerase converts a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid. In some examples of this conversion of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline

alkaloid utilizing the engineered epimerase, the (*S*)-1-benzylisoquinoline alkaloid is selected from the group consisting of (*S*)-norreticuline, (*S*)-reticuline, (*S*)-tetrahydropapaverine, (*S*)-norcoclaurine, (*S*)-coclaurine, (*S*)-*N*-methylococlaurine, (*S*)-3'-hydroxy-*N*-methylococlaurine, (*S*)-norisoorientaline, (*S*)-orientaline, (*S*)-isoorientaline, (*S*)-norprotosinomenine, (*S*)-protosinomenine, (*S*)-norlaudanosoline, (*S*)-laudanosoline, (*S*)-4'-*O*-methyllaudanosoline, (*S*)-6-*O*-methylnorlaudanosoline, and (*S*)-4'-*O*-methylnorlaudanosoline.

[00134] In still other embodiments, the 1-benzylisoquinoline alkaloid that is epimerized using an engineered epimerase may comprise two or more stereocenters, wherein only one of the two or more stereocenters is inverted to produce a diastereomer of the substrate (e.g., (*S*, *R*)-1-benzylisoquinoline alkaloid converted to (*R*, *R*)-1-benzylisoquinoline alkaloid). In examples where only one stereocenter of a 1-benzylisoquinoline alkaloid is inverted when contacted with the at least one enzyme, the product is referred to as an epimer of the 1-benzylisoquinoline alkaloid.

[00135] In some examples, the 1-benzylisoquinoline alkaloid is presented to the enzyme as a single stereoisomer. In some other examples, the 1-benzylisoquinoline alkaloid is presented to the enzyme as a mixture of stereoisomers. In still further embodiments, the mixture of stereoisomers may be a racemic mixture. In some other examples, the mixture of stereoisomers may be enriched in one stereoisomer as compared to another stereoisomer.

[00136] In some examples, a 1-benzylisoquinoline alkaloid, or a derivative thereof, is recovered. In some examples, the 1-benzylisoquinoline alkaloid is recovered from a cell culture. In still further embodiments, the recovered 1-benzylisoquinoline alkaloid is enantiomerically enriched in one stereoisomer as compared to the original mixture of 1-benzylisoquinoline alkaloids presented to the enzyme. In still further embodiments, the recovered 1-benzylisoquinoline alkaloid has an enantiomeric excess of at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 82%, at least 84%, at least 86%, at least 88%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, at least 99.7%, or 100%.

[00137] In some examples, a promorphinan, or a derivative thereof, is recovered. In some examples, the promorphinan is recovered from a cell culture. In still further embodiments, the recovered promorphinan has an enantiomeric excess of at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 82%, at least 84%, at least 86%, at least 88%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, at least 99.7%, or 100%.

[00138] In some examples, a morphinan, or a derivative thereof, is recovered. In some examples, the morphinan is recovered from a cell culture. In still further embodiments, the recovered morphinan has an enantiomeric excess of at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 82%, at least 84%, at least 86%, at least 88%, at least 90%, at least 91%, at least 92%,

at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, at least 99.7%, or 100%.

[00139] In some examples, a bisbenzylisoquinoline, or a derivative thereof, is recovered. In some examples, the bisbenzylisoquinoline is recovered from a cell culture. In still further embodiments, the recovered bisbenzylisoquinoline is enantiomerically enriched in one stereoisomer as compared to the original mixture of bisbenzylisoquinoline presented to the enzyme. In still further embodiments, the recovered bisbenzylisoquinoline has an enantiomeric excess of at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 82%, at least 84%, at least 86%, at least 88%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, at least 99.7%, or 100%.

[00140] In some examples, a nal-opioid, or a derivative thereof, is recovered. In some examples, the nal-opioid is recovered from a cell culture. In still further embodiments, the recovered nal-opioid has an enantiomeric excess of at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 82%, at least 84%, at least 86%, at least 88%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, at least 99.7%, or 100%.

[00141] In some examples, a nor-opioid, or a derivative thereof, is recovered. In some examples, the nor-opioid is recovered from a cell culture. In still further embodiments, the recovered nor-opioid has an enantiomeric excess of at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 82%, at least 84%, at least 86%, at least 88%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, at least 99.7%, or 100%.

[00142] "Isomers" are different compounds that have the same molecular formula. "Stereoisomers" are isomers that differ only in the way the atoms are arranged in space. "Enantiomers" are a pair of stereoisomers that are non-superimposable mirror images of each other. A 1:1 mixture of a pair of enantiomers is a "racemic" mixture. "Diastereoisomers" or "diastereomers" are stereoisomers that have at least two asymmetric atoms but are not mirror images of each other. The term "epimer" as used herein refers to a compound having the identical chemical formula but a different optical configuration at a particular position. For example, the (*R,S*) and (*S,S*) stereoisomers of a compound are epimers of one another. In some examples, a 1-benzylisoquinoline alkaloid is converted to its epimer (e.g., epi-1-benzylisoquinoline alkaloid). The absolute stereochemistry is specified according to the Cahn-Ingold-Prelog R-S system. When a compound is a pure enantiomer, the stereochemistry at each chiral carbon can be specified by either R or S. Resolved compounds whose absolute configuration is unknown can be designated (+) or (-) depending on the direction (dextro- or levorotatory) in which they rotate plane polarized light at the wavelength of the sodium D line. Certain compounds described herein contain one

or more asymmetric centers and can thus give rise to enantiomers, diastereomers, and other stereoisomeric forms that can be defined, in terms of absolute stereochemistry, as (*R*)- or (*S*)-.

Table 1. Example amino acid sequences of DRS-DRR enzymes, split DRS and DRR enzymes, and other nucleotide sequences.

Sequence	Description	SEQ. ID NO.
MELQYISYFQPTSSVVALLLALVSILSSVVVLRKTFLNNYSSSPASSTKTA VLSHQRQQSCALPISGLLHIFMNKNGLIHVTLGNMADKYGPIFSFPTGSH RTL VVSSWEMVKECFTGNNDTAFSNRPIPLAFKTIFYACGGIDSYGLSSV PYGKYWRELKRVCHNLLSNQQLLKFRHLIISQVDTSFNKLYELCKNSE DNHGNYTTTTTTAAGMVRIDDWLAELSFNVIGRIVCGFQSGPKTGAPSR VEQFKEAINEASYFMSTSPVSDNVPMLGWIDQLTGLTRNMKHCGKKLD LVVESIINDHRQKRRFSRTKGGDEKDDEQDDFIDICLSIMEQPQLPGNNN PSQIPIKSIVLDMIGGGTDTTKLTTIWTLSLLLNNPHVLDKAKQEVD AHF RTKRRSTNDAAA VVDFDDIRNLVYIQAIKESMRLYPASPVERLSGE DCVVGGFHVPA GTRLWANVWKMQRDPK VWDDPLVFRPDRFLSDEQK MVDVRGQNYELLFPGAGRRVCPGVSFSLDLMQLVLTRLILEFEMKSPSG KVDMTATPGLMSYKVIPLDILLTHRRKPCVQSAASERDMESSGVPVITL GSGKVMPVLGMGTFEKVVGKGERERLAILKAIEVGYRYFDTAAAYETE EVLGEAIAEALQLGLVKSRDELFISSMLWCTDAHADRVLALQNSLRNL KLEYVDLYMLPFPASLKPGKITMDIPEEDICRMDYRSVWAAMEECQNL GFTKSIGVSNFCKKLQELMATANIPPAVNQVEMSPA FQKKLREYCNA NNILVSAISVLGSNGTPWGSNAVLGSEVLKKIAMAKGKSVAQVSMRWV YEQGASLVVKS FSEERLRENLNIFDWELTKEDHEKIGEIPQCRILSAYFL VSPNGPFKSQEELWDDEA*	<i>P. somniferum</i> plant source; full-length amino acid sequence >RQNK- 2062398 (also <u>FPYZ-</u> <u>2037562</u> , <u>BMRX-</u> <u>2007040</u> , and <u>MLPX-</u> <u>2016197</u>)	SEQ. ID NO. 1
MELQYISYFQPTSSVVALLLALVSILSSVVVLRKTFLNNYSSSPASSTKTA VLSHQRQQSCALPISGLLHIFMNKNGLIHVTLGNMADKYGPIFSFPTGSH RTL VVSSWEMVKECFTGNNDTAFSNRPIPLAFKTIFYACGGIDSYGLSSV PYGKYWRELKRVCHNLLSNQQLLKFRHLIISQVDTSFNKLYELCKNSE DNHGNYTTXLLL PQLAWRQPWKLYYXTTTTAAGMVRIDDWLAELSFN VIGRIVCGFQSGPKTGAPSRVEQFKEAINEASYFMSTSPVSDNVPMLGWI DQLTGLTRNMKHCGKKLDLVVESIINDHRQKRRFSRTKGGDEKDDEQD DFIDICLSIMEQPQLPGNNNPSQIPIKSIVLDMIGGGTDTTKLTTIWTLSLL LNNPHVLDKAKQEVD AHFRTKRRSTNDAAA VVDFDDIRNLVYIQAI	<i>P. somniferum</i> plant source; full-length amino acid sequence >KKCW- 2026866	SEQ. ID NO. 2

KESMRLYPASPVVERLSGEDCVVGGFHVPAAGTRLWANVWKMQRDPK VWDDPLVFRPDRFLSDEQKMVDVRGQNYELLPGAGRRVCPGVSFSLD LMQLVLTRLILEFEMKSPSGKVDMTATPGLMSYKVIPLDILLTHRRIKPC VQSAASERDMESSGVPVITLGSQKVMVPLGMGTFEKVGKGSERERLAIL KAIEVGYRYFDTAAAYETEEVLGEAIAEALQLGLVKSRDELFISSMLWC TDAHADRVLALQNSLRNLKLEYVDLYMLPFPASLKP KITMDIPEEDI CRMDYRSVWAAMEECQNLGFTKSIGVSNFSCKKLQELMATANIPPAVN QVEMSPAQQKKLREYCNANNILVSAISVLGSNGTPWGSNAVLGSEVL KKIAMAKGKSAQVSMRWVYEQASLVVKSFSERLRENLNIFDWELT KEDHEKIGEIPQCRILSAYFLVSPNGPFKSQEELWDDEA*	(also <u>FPYZ-</u> <u>2037562</u>, <u>MLPX-</u> <u>2016197</u>)	
MELQYISYFQPTSSVALLALVSILSSVVLRKTFLNNYSSSPASSTKTA VLSHQRRQSCALPISGLLHIFMNKNGLIHVTLGNMADKYGPIFSFPTGSH RTLVSWSWEMVKECFTGNNDTAFSNRPIPLAKTIFYACGGIDSYGLSSV PYGKYWRELKRVCHNLLSNQQLKFRHLIISQVDTSFNKLYELCKNSE DNHGNYYYYTTTTAAGMVRIDDWLAELSFNVIGRIVCGFQSGPKTGAPSR VEQFKEAINEASYFMSTSPVSDNVPMLGWIDQLTGLTRNMKHCCKLD LVVESIINDHRQKRRFSRTKGGDEKDDEQDDFIDICLSIMEQPQLPGNNN PSQIPIKSIVLDMIGGGTDTTKLTTIWTLSSLLNNPHVLDKAKQEVDHAF RTKRRSTNDAAA VVDFDDIRNLVYIQAIKESMRLYPASPVVERLSGE DCVVGGFHVPAAGTRLWANVWKMQRDPK VWDDPLVFRPDRFLSDEQK MVDVRGQNYELLPGAGRRVCPGVSFSLDLMQLVLTRLILEFEMKSPSG KVDMTATPGLMSYKVIPLDILLTHRRIKPCVQSAASERDMESSGVPVITL GSGKVMVPLGMGTFEKVGKGSERERLAILKAIEVGYRYFDTAAAYETE EVLGEAIAEALQLGLVKSRDELFISSMLWCTDAHADRVLALQNSLRNL KLEYVDLYMLPFPASLKP KITMDIPEEDICRMDYRXVSKPWLH*	<i>P.</i> <i>somniferum</i> plant source; partial- length amino acid sequence >SUF- 2025636	SEQ. ID NO. 3
MRWHRXIDSYGLSSVPYGYWRELKRVCHNLLSNQQLKFRHLIISQ VDTSFNKLYELCKNSEDNQGNYP TTTTAAGMVRIDDWLAELSFNVIGRI VCGFQSGPKTGAPSRVEQFKEAINEASYFMSTSPVSDNVPMLGWIDQLT GLTRNMKHCCKLDLVVESIINDHRQKRRFSRTKGGDEKDDEQDDFIDI CLSIMEQPQLPGNNNPSQIPIKSIVLDMIGGGTDTTKLTTIWTLSSLLNNP HVLDKAKQEVDHAFRTKRRSTNDAAA VVDFDDIRNLVYIQAIKESM RLYPASPVVERLSGEDCVVGGFHVPAAGTRLWANVWKMQRDPK VWDD PLVFRPDRFLSDEQKMVDVRGQNYELLPGAGRRVCPGVSFSLDLMQL VLTRLILEFEMKSPSGKVDMTATPGLMSYKVIPLDILLTHRRIKPCVQSA ASERDMESSGVPVITLGSQKVMVPLGMGTFEKVGKGSERERLAILKAIE VGYRYFDTAAAYETEEVLGEAIAEALQLGLVKSRDELFISSMLWCTDA	<i>P.</i> <i>somniferum</i> plant source; partial- length amino acid sequence >MIKW- 2013651	SEQ. ID NO. 4

HADRVLLALQNSLRNLKLEYVDLYMLPFPASLKPGKITMDIPEEDICRM DYRSVWAAMEECQNLGFTKSIGVSNFSCKKLQELMATANIPPAVNQVE MSPAFQQKKLREYCNANNILVSAISVLGSNGTPWGSNAVLGSEVLKKIA MAKGKSAQVSMRWVYEQASLVVKSFSSEERLRENLNIFDWELTKED HEKIGEIPQCRILSAYFLVSPNGPFKSQEELWDDEA*		
MELQYISYFQPTSSVVALLLALVSILSSVVVLRLKTLNNYSSSPASSTKTA VLSHQRQQSCALPISGLLHIFMNKNGLIHVTLGNMADKYGPIFSFPTGSH RTL VVSSWEMVKECFTGNNDTAFSNRPIPLAFKTIFYACGGIDSYGLSSV PYGKYWRELKRVCHNLLSNQQLKFRHLIISQVDTSFNKLYELCKNSE DNQGNYYYYTTTAAGMVRIDDWLAELSFNVIGRIVCGFQSGPKTGAPSRV EQFKEAINEASYFMSTSPVSDNVPMLGWIDQLTGLTRNMKHCGKKLDL VVESIINDHRQKRRFSRTKGGDEKDDEQDDFIDICLSIMEQPQLPGNNNP SQIPIKSIVLDMIGGGTDTTKLT TIWTL SLLLNNPHVLDKAKQEVD AHR TKRRSTNDAAA AVDFDDIRNLVYIQAIKESMRLYPASPVERLSGED CVVGGFHVPAGTRLWANVWKMQRDPKVWDDPLVFRPDRFLSDEQKM VDVRGQNYELL PFGAGRRVCPGVSFSLDLMQLVLTRLILEFEMKSPSGK VDMTATPGLMSYKVIPLDILLTHRRIKPCVQSAASERDMESSGVPVITLG SGKVMPVLGMGTFEKVGKGSERERLAILKAIEVGYRYFDTAAYETEE VLGEAIAEALQLGLVKSRLDELFISSMLWCTDAHADRVLLALQNSLRNL KLEYVDLYMLPFPASLKPGKITMDIPEEDICRM DYRSVWAAMEECQNL GFTKSIGVSNFSCKKLQELMATANIPPAVNQVEMSPAFQQKKLREYCN ANNILVSAISVLGSNGTPWGSNAVLGSEVLKKIAMAKGKSAQVSMRWV YEQASLVVKSFSSEERLRENLNIFDWELTKEDHEKIGEIPQCRILSAYFL VSPNGPFKSQEELWDDEA*	<i>P. setigerum</i> plant source; full-length amino acid sequence >EPRK- 2027940 (also <u>FPYZ-</u> <u>2037562</u> , <u>STDO-</u> <u>2019715</u> , <u>FNXH-</u> <u>2029312</u> , <u>MLPX-</u> <u>2016196</u> , <u>MLPX-</u> <u>2016197</u>)	SEQ. ID NO. 5
MELQYISYFQPTSSVVALLLALVSILSSVVVLRLKTLNNYSSSPASSTKTA VLSHQRQQSCALPISGLLHIFMNKNGLIHVTLGNMADKYGPIFSFPTGSH RTL VVSSWEMVKECFTGNNDTAFSNRPIPLAFKTIFYACGGIDSYGLSSV PYGKYWRELKRVCHNLLSNQQLKFRHLIISQVDTSFNKLYELCKNSE DNQGNYYYYTTTAAGMVRIDDWLAELSFNVIGRIVCGFQSGPKTGAPSRV EQFKEAINEASYFMSTSPVSDNVPMLGWIDQLTGLTRNMKHCGKKLDL VVESIINDHRQKRRFSRTKGGDEKDDEQDDFIDICLSIMEQPQLPGNNNP SQIPIKSIVLDMIGGGTDTTKLT TIWTL SLLLNNPHVLDKAKQEVD AHR TKRRSTNDAAA AVDFDDIRNLVYIQALYPASPVERLSGEDCVVGGF HVPAGTRLWANVWKMQRDPKVWDDPLVFRPDRFLSDEQKMVDVRGQ	<i>P. setigerum</i> plant source; partial- length amino acid sequence >QCOU- 2000833	SEQ. ID NO. 6

NYELLPGAGRRVCPGVSFSLDLMQLVLTRLILEFEMKSPSGKVDMTAT PGLMSYKVIPLDILLTHRIKPCVQSAASERDMESSGVPVITLGSGKVM VLGMGTFEKVGKGSEERLAILKAIEVGYRYFDATAAYETEEVLGEAIA EALQLGLVKSDELFISSMLWCTDAHADRVLALQNSLRNLKLEYVDL YMLPFPASLKPGKITMDIPEEDICRMDYRSVWAAMEE		
MELQYFSYFQPTSSVVALLLALVSILFSVVVLRKTFSNNYSSPASSTETA VLCHQRQQSCALPISGLLHVFMNKNGLIHVTLGNMADKYGPIFSFPTGS HRTLVSWSWEMVKECFTGNNDTAFSNRPIPLAFQTIFYACGGIDSYGLSS VPYGYWRELKVCVHNLLSNQQLLKFRHLIISQVDTSFNKLYELCKNS EDNQGMVRMDDWLAQLSFNVIGRIVCGFQSDPKTGAPSRVEQFKEVIN EASYFMSTSPVSDNVPMLGWIDQLTGLTRNMKHCCKLDLVVESIIKD HRQKRRFSRTKGGDEKDDEQDDFIDICLSIMEQPQLPGNNSPPQIPIKSIV LDMIGGGTDTTKLTTIWTL SLLLNNPHVLDKAKQEVD AFRKKRRSTD DAAA AVDFDDIRNLVYIQAIIKESMRLYPASPVERLSGEDCVVGGFH VPAGTRLWANVWKMQRDPKVDWDDPLVFRPERFLSDEQKMVDVRGQN YELLPGAGRRICPGVSFSLDLMQLVLTRLILEFEMKSPSGKVDMTATPG LMSYKVVPLDILLTHRIKSCVQLASSERDMESSGVPVITLSSGKVMPL GMGTFEKVGKGSEERLAILKAIEVGYRYFDATAAYETEEVLGEAIAEA LQLGLIESRDELFISSMLWCTDAHDPDRVLALQNSLRNLKLEYLDLYML PFPASLKPGKITMDIPEEDICRMDYRSVWSAMEECQNLGFTKSIGVSNFS SKKLQELMATANIPPAVNQVEMSPAQQKLRKYCNANNILVSAVSILG SNGTPWGSNAVLGSEVLKQIAMAKGKSAQVSMRWVYEQGASLVVK SFSEERLRENLNIFDWELTKEDNEKIGEIPQCRILTAYFLVSPNGPFSQE ELWDDKA*	<i>P.</i> <i>bracteatum</i> plant source; full-length amino acid sequence >SSDU- 2015634 (also <u>SSDU-</u> <u>2015636</u>, <u>ZSNV-</u> <u>2027701</u>, <u>RRID-</u> <u>2004435</u>)	SEQ. ID NO. 7
MELQYFSYFQPTSSVVALLLALVSILFSVVVLRKTFSNNYSSPASSTETA VLCHQRQQSCALPISGLLHVFMNKNGLIHVTLGNMADKYGPIFSFPTGS HRTLVSWSWEMVKECFTGNNDTAFSNRPIPLAFQTIFYACGGIDSYGLSS VPYGYWRELKVCVHNLLSNQQLLKFRHLIISQVDTSFNKLYELCKNS EDNQGMVRMDDWLAQLSFNVIGRIVCGFQSDPKTGAPSRVEQFKEVIN EASYFMSTSPVSDNVPMLGWIDQLTGLTRNMKHCCKLDLVVESIIKD HRQKRRFSRTKGGDEKDDEQDDFIDICLSIMEQPQLPGNNSPPQIPIKSIV LDMIGGGTDTTKLTTIWTL SLLLNNPHVLDKAKQEVD AFRKKRRSTD DAAA AVDFDDIRNLVYIQAIIKESMRLYPASPVERLSGEDCVVGGFH VPAGTRLWANVWKMQRDPKVDWDDPLVFRPERFLSDEQKMVDVRGQN YELLPGAGRRICPGVSFSLDLMQLVLTRLILEFEMKSPSGKVDMTATPG LMSYKVVPLDILLTHRIKSCVQLASSERDMESSGVPVITLSSGKVMPL	<i>P.</i> <i>bracteatum</i> plant source; full-length amino acid sequence >TMWO- 2027322 (also <u>RRID-</u> <u>2004435</u>)	SEQ. ID NO. 8

GMGTFEKVGKGSERERLAILKAIEVGYRYFDTAAAYETEEVLGEAIAEA LQLGLIESRDELFISSMLWCTDAHPDRVLLALQNSLRNLKLEYLDLYML PFPASLKPGKITMDIPEEDICRMDYRSVWSAMEECQNLGFTKSIGVSNFS CKKLQELMATANIPPAVNQVEMSPAQQKKLREYCNANNILVSAVSILG SNGTPWGSNAVLGSEVLKQIAMAKGKSAQVSMRWVVEQGASLVVK SFSEERLRENLNIFDWELTKEDNEKIGEIPQCRILTAYFLVSPNGPFKSQE ELWDDKA*		
SSPASSTETA VLCHQRQQSCALPISGLLHIFMNKNGLIHVTLGNMADKY GPIFSFPTGSHRILVVSSWEMVKECFTGNNDTAFSNRPIPLAFKTIFYACR GIDSYGLSSVPYGYWRELKVCVHNLLSNQQLLKFRHLIISQVDTSFN KLYELCKNSEDNQGMVRMDDWLAQLSFSVIGRIVCGFQSDPKTGAPSR VEQFKEAINEASYFMSTSPVSDNVPMLGWIDQLTGLTRNMTHCGKKLD LVVESIINDHRQKRRFSRTKGGDEKDDEQDDFIDICLSIMEQPQLPGNNN PPKIPIKSIVLDMIGAGTDTTKLTIIWTL SLLLNNPNVLAKAKQEVD AHFE TKKRSTNEASVVVDFDDIGNLVYIQAIIKESMRLYPVSPVVERLSSDCV VGGFHVPA GTRLWANVWKMQRDPK VWD DPLVFRPERFLSDEQKMVD VRGQNYELL PFGAGRRICPGVSFSLDLMQLVLTRLILEFEMKSPSGKVD MTATPGLMSYKVVPLDILLTHRRIKSCVQLASSERDMESSGVPVITLRS KVMPVLGMGTFEKAGKGSERERLAILKAIEVGYRYFDTAAAYETEEVL GEAIAEALQLGLIKSRDELFISSMLWCTDAHPDRVLLALQNSLRNLKLE YVDLYMLPFPASLKPGKITMDIPEEDICPMDYRSVWSAMEECQNLGLTK SIGVSNFSCKKLEELMATANIPPAVNQVEMSPAQQKKLREYCNANNIL VSAVSILGSNGTPWGSNAVLGSEVLKQIAMAKGKSAQVSMRWVVEQ GASLVVKSFSSEERLRENLNIFDWQLTKEDNEKIGEIPQCRILSAYFLVSPK GPFKSQEELWDDKA*	<i>P.</i> <i>bracteatum</i> plant source; partial- length amino acid sequence >pbr.PBRS T1PF_8940 5	SEQ. ID NO. 9
SSPASSTETA VLCHQRQQSCALPISGLLHIFMNKNGLIHVTLGNMADKY GPIFSFPTGSHRILVVSSWEMVKECFTGNNDTFFSNRPIPLAFKIIFYAGG VDSYGLALVPYGYWRELKICVHNLLSNQQLLKFRHLIISQVDTSFNK LYELCKNSEDNQGMVRMDDWLAQLSFSVIGRIVCGFQSDPKTGAPSRV EQFKEAINEASYFMSTSPVSDNVPMLGWIDQLTGLTRNMTHCGKKLDL VVESIINDHRQKRRFSRTKGGDEKDDEQDDFIDICLSIMEQPQLPGNNNP PKIPIKSIVLDMIGGGTDTTKLTTIWTL SLLLNNPHVLDKAKQEVD AHFL TKKRSTNDAAVVDFDDIRNLVYIQAIIKESMRLYPASPVERLSGEDCV VGGFHVPA GTRLWVNVWKMQRDPNVWADPMVFRPERFLSHGQKKM VDVRGKNYELL PFGAGRRICPGISFSLDLMQLVLTRLILEFEMKSPSGKV DMTATPGLMSYKVVPLDILLTHRRIKSCVQLASSERDMESSGVPVITLRS	<i>P.</i> <i>bracteatum</i> plant source; partial- length amino acid sequence >pbr.PBRS T1PF_4328	SEQ. ID NO. 10

GKVMPLVGMGTFEKAGKGSERERLAILKAIEVGYRYFDTAAAYETEEV LGEAIAEALQLGLIKSRDELFISSMLWCTDAHPDRVLLALQNSLRNLKLE YVDLYMLPFPASLKPGKITMDIPEEDICPMDYRSVWSAMEECQNLGLTK SIGVSNFSCCKLEELMATANIPPAVNQVEMSPAQQKKLREYCNANNIL VSAVSILGSNGTPWGSNAVLGSEVLKKIAMAKGKSVAQVSMRWVYEQ GASLVVKSFSSEERLRENLNIFDWQLTKEDNEKIGEIPQCRILSAYFLVSPK GPFKSQEELWDDKA*		
SSPASSTETAVLCHQRQQSCALPISGLLHIFMNKNGLIHVTLGNMADKY GPIFSFPTGSHRILVVSSWEMVKECFTGNNDTFFSNRPIPLAFKIIFYAGG VDSYGLALVPYGKYWRELKICVHNLLSNQQLNFRHLIISQVDTSFNK LYDLSNKKKNTTTDSGTVRMDDWLAQLSFNVIGRIVCGFQTHETATS SVERFTEAIDEASRFMSIATVSDTFPWLGWIDQLTGLTRKMKHYGKKLD LVVESIIEDHRQNRISGKQGDDFIDICLSIMEQPQIIPGNNDPPRQPIKS IVLDMIGGGTDTTKLTTTWTLSTLLNPHVLEKAREEVDAHFGTKRRPT NDDAVMVEFDDIRNLVYIQAIHESMRLYPASPVVERLSGEDCVVGGFH VPAGTRLWVNVWKMQRDPNVWADPMVFRPERFLSDEQKMVDVRGQ NYELLFPAGARRICPGVSFSLDLMQLVLTRLILEFEMKSPSGKVDMTATP GLMSYKVVPLDILLTHRIKSCVQLASSERDMESSGVPVITLRSKGKMP VLGMGTFEKAGKGSERERLAILKAIEVGYRYFDTAAAYETEEVLGEAIA EALQLGLIKSRDELFISSMLWCTDAHPDRVLLALQNSLRNLKLEYVDLY MLPFPASLKPGKITMDIPEEDICPMDYRSVWSAMEECQNLGLTKSIGVS NFSCCKLEELMATANIPPAVNQVEMSPAQQKKLREYCNANNILVSAV SILGSNGTPWGSNAVLGSEVLKKIAMAKGKSVAQVSMRWVYEQGASL VVKSFSSEERLRENLNIFDWQLTKEDNEKIGEIPQCRILSAYFLVSPKGPFK SQEELWDDKA*	<i>P.</i> <i>bracteatum</i> plant source; partial- length amino acid sequence >pbr.PBRS T1PF_1218 0	SEQ. ID NO. 11
VALRKKILKNYYSSSSSTATAVSHQWPKASRALPLIDLLHVFFNKTDLM HVTLGNMADKFGPIFSFPTGSHRTLWVSSWEKAKECFTGNNDIVFSGRP LPLAFKLIFYAGGIDSYGISQVPYGKKWRELNRNICVHNILSNQQLLKFRH LMISQVDNSFNKLYEVCNSNKDEGDSATSTTAAGIVRMDDWLGLAFD VIARIVCGFQSQTETSTSSMERFTEAMDEASRFMSVTAVSDTVPWLGW IDQLTGLKRNMKHCGKKLNLVKSIEDHRQKRRLSSTKKGDENIDED EQDDFIDICLSIMEQPQLPGNNNPPKIPIKSIVLDMIGGGTDTTKLTTIWT SLLLNNPHVLDKAKQEVDHFLTKRRSTNDAAVDFDDIRNLVYIQAIH KESMRLYPASPVVERLSGEDCVVGGFHVPAGTRLWVNVWKMQRDPN VWADPMVFRPERFLSDEQKMVDVRGQNYELLFPAGARRICPGVSFSLD LMQLVLTRLILEFEMKSPSGKVDMTATPGLMSYKVVPLDILLTHRIKS	<i>P.</i> <i>bracteatum</i> plant source; partial- length amino acid sequence >pbr.PBRS T1PF_4329	SEQ. ID NO. 12

CVQLASSERDMESSGVPVITLRSGKVMPLVGMGTFEKAGKGSEERERLAI LKAIEVGYRYFDTAAAYETEEVLGEAIAEALQLGLIKSRDELFISSMLWC TDAHPDRVLLALQNSLRNLKLEYVDLYMLPFPASLKPGRITMDIPEEDIC PMDYRSVWSAMEECQNLGLTKSIGVSNFSCKKLEELMATANIPPAVNQ VEMSPAQQKKLREYCNANNILVSAVSILGSNGTPWGSNAVLGSEVLK KIAMAKGKSVAQVSMRWVVEQASLVVKSFSSEERLRENLNIFDWQLT KEDNEKIGEIPQCRILSAYFLVSPKGPFSQEELWDDKA*		
MELQYFSYFQPTSSVVALLLALVSILFSVVVLKTFSSNNYSSPASSTETA VLCHQRQQSCALPISGLLHVFMNKNGLIHVTLGNMADKYGPISFPTGS HRTLVSWSWEMVKECFTGNNDTAFSNRPIPLAFQTIFYACGGIDSYGLSS VPYGYWRELKVCVHNLLSNQQLLKFRHLIISQVDTSFNKLYELCKNS EDNQGMVRMDDWLAQLSFNVIGRIVCGFQSDPKTGAPSRVEQFKEVIN EASYFMSTSPVSDNVPMLGWIDQLTGLTRNMKHCCKLDLVVESIHKD HRQKRRFSRTKGGDEKDDEQDDFIDICLSIMEQPQLPGNNSPPQIPIKSIV LDMIGGGTDTTKLTTIWTLSLLLNNPHVLDKAKQEVDHFRKKRRSTD DAAA AVDFDDIRNLVYIQAIIKESMRLYPASPVERLSGEDCVVGGFH VPAGTRLWANVWKMQRDPKVVDDPLVFRPERFLSDEQKMVDVRGQN YELLPGAGRRICPGVSFSLDMLQLVLRLLILEFEMKSPSGKVDMTATPG LMSYKVVPLDILLTHRIKSCVQLASSERDMESSGVPVITLSSGKVMPLV GMGTFEKVGKSEERERLAILKAIEVGYRYFDTAAAYETEEVLGEAIAEA LQLGLIESRDELFISSMLWCTDAHPDRVLLALQNSLRNLKLEYLDLYML PFPASLKPGRITMDIPEEDICRMDYRSVWSAMEECQNLGFTKSIGVSNFS SKKLQELMATANIPPAVNQVEMSPAQQKKLREYCNANNILVSAVSILG SNGTPWGSNAVLGSEVLKQIAMAKGKSVAQVSMRWVXKFSAYAIVWS LFFGHRICITLYSFLIRNVAYICITY*	<i>P.</i> <i>bracteatum</i> plant source; partial- length amino acid sequence >SSDU- 2015635	SEQ. ID NO. 13
MELQYFSYFQPTSSVVALLLALVSILFSVVVLKTFSSNNYSSPASSTETA VLCHQRQQSCALPISGLLHVFMNKNGLIHVTLGNMADKYGPISFPTGS HRTLVSWSWEMVKECFTGNNDTAFSNRPIPLAFQTIFYACGGIDSYGLSS VPYGYWRELKVCVHNLLSNQQLLKFRHLIISQVDTSFNKLYELCKNS EDNQGMVRMDDWLAQLSFNVIGRIVCGFQSDPKTGAPSRVEQFKEVIN EASYFMSTSPVSDNVPMLGWIDQLTGLTRNMKHCCKLDLVVESIHKD HRQKRRFSRTKGGDEKDDEQDDFIDICLSIMEQPQLPGNNSPPQIPIKSIV LDMIGGGTDTTKLTTIWTLSLLLNNPHVLDKAKQEVDHFRKKRRSTD DAAA AVDFDDIRNLVYIQAIIKESMRLYPASPVERLSGEDCVVGGFH VPAGTRLWANVWKMQRDPKVVDDPLVFRPERFLSDEQKMVDVRGQN YELLPGAGRRICPGVSFSLDMLQLVLRLLILEFEMKSPSGKVDMTATPG	<i>P.</i> <i>bracteatum</i> plant source; partial- length amino acid sequence >SSDU-	SEQ. ID NO. 14

LMSYKVVPLDILLTHRIKSCVQLASSERDMESSGVPVITLSSGKVMPLV GMGTFEKVGKGSERERLAILKAIEVGYRYFDTAAAYETEEVLGEAIAEA LQLGLIESRDELFISSMLWCTDAHPDRVLLALQNSLRQVFLMQIRLIYIC TYQQVHLNIYFQINEFVLCDMYRNKLEY	2015637	
LNNYSSSPASSTKTAVLSHQRRQSCALPISGLLHIFMNKNGLIHVTLGNM ADKYGPISFPTGSHRTLVSSEWEMVKECFTGNNDTAFSNRPIPLAFKTI FYACGGIDSYGLSSVPYGYWRELKVCVHNLLSNQQLLKFRHLIISQV DTSFNKLYELCKNSDNQGNYPPTTTTAAGMVRIDDWLAELSFNVIGRIV CGFQSGPKTGAPSRVEQFKEAINEASYFMSTSPVSDNVPMLGWIDQLTG LTRNMKHCCKLDLVVESIINDHRQKRRFSRTKGGDEKDDEQDDFIDIC LSIMEQPQLPGNNPSQIPIKSIVLDMIGGGTDTTKLTTIWTLSLLLNNPH VLDKAKQEVDAAHFRTKRRSTNDAAA VVDFDDIRNLVYIQAIKESMR LYPASPVVERLSGEDCVVGGFHVPAAGTRLWANVWKMQRDPKVVDDP LVFRPDRFLSDEQKMVDVRGQNYELLPGAGRRVC PGVSFSLDLMQLV LTRLILEFEMKSPSGKVDMTATPG LMSYKVIPLDILLTHRIKPCVQSAA SERDMESSGVPVITLSSGKVMPLVGMGTFEKVGKGSERERLAFLKAIEV GYRYFDTAAAYETEEVLGEAIAEALQLGLIKSRDELFISSMLWPCDAHPD LVVPALQNSLRNLKLEYVDLYMLPFPASLKPGKITMDIPEEDICRMDYR SVWAAMEECQNLGFTKSGVSNFSCCKLQELMATANIPPAVNQVEMSP AFQQKKLREYCNANNILVSAISVLGSGNGTPWGSNAVLGSEVLKKIAMA KGKSAQVSMRWVYEQGASLVVKSSEERLRENLNIFDWELTKEDHEK IGEIPQCRILSAYFLVSPNGPFKSQEELWDDEA*	<i>C. majus</i> plant source; partial- length amino acid sequence >chm.CMA ST2PF_149 84	SEQ. ID NO. 15
MELQYFSYFQPTSSVVALLLALVSILFSVVVLRKTFSNNYSSPASSTETA VLCHQRRQSCALPISGLLHVF MNKNGLIHVTLGNMADKYGPISFPTGS HRTLVSSEWEMVKECFTGNNDTAFSNRPIPLAFQTIFYACGGIDSYGLSS VPYGYWRELKVCVHNLLSNQQLLKFRHLIISQVDTSFNKLYELCKNS EDNQGMVRMDDWLAQLSFNVIGRIVCGFQSDPKTGAPSRVEQFKEVIN EASYFMSTSPVSDNVPMLGWIDQLTGLTRNMKHCCKLDLVVESIHKD HRQKRRFSRTKGGDEKDDEQDDFIDICLSIMEQPQLPGNNSPPQIPIKSIV LDMIGGGTDTTKLTTIWTLSLLLNNPHVLDKAKQEVDAAHFRTKRRSTD DAAA VVDFDDIRNLVYIQAIKESMRLYPASPVVERLSGEDCVVGGFH VPAGTRLWANVWKMQRDPKVVDDPLVFRPERFLSDEQKMVDVRGQN YELLPGAGRRICPGVSFSLDLMQLV LTRLILEFEMKSPSGKVDMTATPG LMSYKVVPLDILLTHRIKSCVQLASSERDMESSGVPVITLSSGKVMPLV GMGTFEKVGKGSERERLAILKAIEVGYRYFDTAAAYETEEVLGEAIAEA LQLGLIESRDELFISSMLWCTDAHPDRVLLALQNSLRNLKLEYLDLYML	<u>P.</u> <u>bracteatum</u> <u>DRS-DRR</u>	SEQ. ID NO. 16

PFPA SLKPGKITMDIPEEDICRMDYRSVWSAMEECQNLGFTK SIGVSNFS CKKLQELMATANIPPAVNQVEMSPA FQQKKLREYCNANNILVSAVSILG SNGTPWGSNAVLGSEVLKQIAMAKGK SVAQVSMRWVYE QGASLVVK SFSEERLREN LNIFDWELTKEDNEKIGEIPQCRILTAYFLVSPNGPFSQE ELWDDKA*		
MELQYFSYFQPTSSVVALLLALVSILFSVVVL RKTFSNNYSSPASSTETA VLCHQRQQSCALPISGLLHVFMNKNGLIHVTLGNMADKYGP IFSFPTGS HRTLVS SSWEMVKECFTGNNDTAFSNRPIPLAFQTIFYACGGIDSYGLSS VPY GKYWREL RKVCVHNLLSNQQLLKFRHLIISQVDT SFNKLYELCKNS EDNQGMVRMDDWLAQLSFNVIGRIVCGFQSDPKTGAPSRVEQFKEVIN EASYFMSTSPVSDNPMLGWIDQLTGLTRNMKHC GK KLDLVVESIIKD HRQKRRFSRTKGGDEKDDEQDDFIDICLSIMEQPQLPGNNSPPQIPIKSIV LDMIGGGTDTTKLTTIWTLSLLLNNPHVLDKAKQEVDAHFRKKRRSTD DAAA AVDFDDIRNLVYIQAIKESMRLYPASP VVERLSGEDCVVGGFH VPAGTRLWANVWKMQRDPKVWDDPLVFRPERFLSDEQKMVDVRGQN YELL PFGAGRRICPGVSFSLDLMQLVLTRLILEFEMKSPSGKVDMTATPG LMSYKV VPLDILLTHRRIKSCVQLASSERD	<u>P.</u> <u>bracteatum</u> <u>DRS</u>	SEQ. ID NO. 17
MESSGVPVITLSSGKVMPVLGMGT FEKVGKG SERERLAILKAIEVGYRY FDTAAYETEEVLGEAIAEALQLGLIESRDELFISSMLWCTDAHPDRVLL ALQNSLRNLKLEYLDLYMLPFPASL KPGKITMDIPEEDICRMDYRSVWS AMEECQNLGFTK SIGVSNFSCKKLQELMATANIPPAVNQVEMSPA FQQ KKLREYCNANNILVSAVSILGSNGTPWGSNAVLGSEVLKQIAMAKGKS VAQVSMRWVYE QGASLVVKS FSEERLREN LNIFDWELTKEDNEKIGEIP QCRILTAYFLVSPNGPFSQEELWDDKA*	<u>P.</u> <u>bracteatum</u> <u>DRR</u>	SEQ. ID NO. 18
TTCAGTTCGAGTTTATCATTATCAATACTGCCATTTC AAAGAATACGT AAATAATTAATAGTAGTGATTTTCCTAACTTTATT TAGTCAAAAAATT AGCCTTTTAATTCTGCTGTAACCCGTACATGCCCAAAATAGGGGGCG GGTTACACAGAATATATAACATCGTAGGTGTCTGGGTGAACAGTTTA TTCCTGGCATCCACTAAATATAATGGAGCCCGCTTTTAAAGCTGGCA TCCAGAAAAAAAAGAATCCCAGCACCAAAATATTGTTTTCTTCACC AACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAGAACA GGGGCACA AACAGGCAAAAAACGGGCACAACCTCAATGGAGTGATG CAACCTGCCTGGAGTAAATGATGACACAAGGCAATTGACCCACGCA TGTATCTATCTCATTTTCTTACACCTTCTATTACCTTCTGCTCTCTCTG ATTTGGAAAAAGCTGAAAAAAAAGGTTGAAACCAGTTCCTGAAAT TATCCCTACTTGACTAATAAGTATATAAAGACGGTAGGTATTGAT	<u>TDH3</u> <u>Promoter</u>	SEQ. ID NO. 19

TGTAATTCTGTAAATCTATTTCTTAAACTTCTTAAATTCTACTTTTATA GTTAGTCTTTTTTTTAGTTTTTAAAACACCAAGAACTTAGTTTCGAATA AACACACATAAAACAAACAAA		
GAGCGTTGGTTGGTGGATCAAGCCCACGCGTAGGCAATCCTCGAGC AGATCCGCCAGGCGTGTATATATAGCGTGGATGGCCAGGCAACTTTA GTGCTGACACATACAGGCATATATATATGTGTGCGACGACACATGAT CATATGGCATGCATGTGCTCTGTATGTATATAAACTCTTGTTTTCTT CTTTTCTCTAAATATTCTTTCCTTATACATTAGGACCTTTGCAGCATA AATTACTATACTTCTATAGACACACAAACACAAATACACACACTAAA TTAATA	<u>CYC1</u> <u>Promoter</u>	SEQ. ID NO. 20
CATAGCTTCAAATGTTTCTACTCCTTTTTTACTCTTCCAGATTTTCTC GGACTCCGCGCATCGCCGTACCACTTCAAAAACACCCAAGCACAGCAT ACTAAATTTCCCCTCTTTCTTCTCTAGGGTGTCGTTAATTACCCGTA CTAAAGGTTTGGAAGAAAAAGAGACCGCCTCGTTTCTTTTTCTT CGTCGAAAAAGGCAATAAAAAATTTTATCACGTTTCTTTTTCTTGAA AATTTTTTTTTTTGATTTTTTTTCTTTTCGATGACCTCCCATTGATATT TAAGTTAATAAACGGTCTTCAATTTCTCAAGTTTCAGTTTCATTTTTTC TTGTTCTATTACAACCTTTTTTTACTTCTTGCTCATTAGAAAGAAAGCA TAGCAATCTAATCTAAGTTTAAATTACAAA	<u>TEF1</u> <u>Promoter</u>	SEQ. ID NO. 21
ACAGGCCCTTTTCCTTTGTGCGATATCATGTAATTAGTTATGTCACGC TTACATTCACGCCCTCCTCCCACATCCGCTCTAACCGAAAAGGAAGG AGTTAGACAACCTGAAGTCTAGGTCCCTATTTATTTTTTTAATAGTT ATGTTAGTATTAAGAACGTTATTTATATTTCAAATTTTTCTTTTTTTC TGTACAAACGCGTGTACGCATGTAACATTATACTGAAAACCTTGCTT GAGAAGGTTTTGGGACGCTCGAAGGCTTAAATTTG	<u>CYC1</u> <u>Terminator</u>	SEQ. ID NO. 22
GCGAATTTCTTATGATTTATGATTTTTATTATTAAATAAGTTATAAAA AAAATAAGTGTATACAAATTTTAAAGTGAAGTCTTAGGTTTTAAACG AAAATTCTTATTCTTGAGTAACTCTTTCCTGTAGGTCAGGTTGCTTTC TCAGGTA	<u>ADH1</u> <u>Terminator</u>	SEQ. ID NO. 23
CCTCGCCGCAGTTAATTAAAGTCAGTGAGCGAGGAAGCGCGTAACT ATAACGGTCCTAAGGTAGCGAATCCTGATGCGGTATTTTCTCCTTAC GCATCTGTGCGGTATTTACACCCGCATAGATCGGCAAGTGCACAAAC AATACTTAAATAAATACTACTCAGTAATAACCTATTTCTTAGCATTTT TGACGAAATTTGCTATTTTGTAGAGTCTTTTACACCATTTGTCTCCA	<u>pDW10</u>	SEQ. ID NO. 24

CACCTCCGCTTACATCAACACCAATAACGCCATTTAATCTAAGCGCA TCACCAACATTTTCTGGCGTCAGTCCACCAGCTAACATAAAATGTAA GCTTTCGGGGCTCTCTTGCCTTCCAACCCAGTCAGAAATCGAGTTCC AATCCAAAAGTTCACCTGTCCCACCTGCTTCTGAATCAAACAAGGGA ATAAACGAATGAGGTTTCTGTGAAGCTGCACTGAGTAGTATGTTGCA GTCTTTTGGAAATACGAGTCTTTTAATAACTGGCAAACCGAGGAACT CTTGGTATTCTTGCCACGACTCATCTCCATGCAGTTGGACGATATCA ATGCCGTAATCATTGACCAGAGCCAAAACATCCTCCTTAAGTTGATT ACGAAACACGCCAACCAAGTATTTTCGGAGTGCCTGAACTATTTTTAT ATGCTTTTACAAGACTTGAAATTTTCCTTGCAATAACCGGGTCAATT GTTCTCTTTCTATTGGGCACACATATAATAACCCAGCAAGTCAGCATC GGAATCTAGAGCACATTCTGCGGCCTCTGTGCTCTGCAAGCCGCAAA CTTTCACCAATGGACCAGAACTACCTGTGAAATTAATAACAGACATA CTCCAAGCTGCCTTTGTGTGCTTAATCACGTATACTCACGTGCTCAAT AGTCACCAATGCCCTCCCTCTTGGCCCTCTCCTTTTCTTTTTTCGACC GAATTAATTCTTAATCGGCAAAAAAAGAAAAGCTCCGGATCAAGAT TGTACGTAAGGTGACAAGCTATTTTTCAATAAAGAATATCTTCCACT ACTGCCATCTGGCGTCATAACTGCAAAGTACACATATATTACGATGC TGTCTATTAAATGCTTCCTATATTATATATATAGTAATGTCGTGATC TATGGTGCCTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGC CAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTG TCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGA GCTGCATGTGTCAGAGGTTTTACCGTCATCACCGAAACGCGCGAGA CGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGAT AATAATGGTTTCTTAGACGGATCGCTTGCCTGTAACCTACACGCGCC TCGTATCTTTTAATGATGGAATAATTTGGGAATTTACTCTGTGTTTAT TTATTTTTATGTTTTGTATTTGGATTTTAGAAAGTAAATAAAGAAGGT AGAAGAGTTACGGAATGAAGAAAAAAAATAAACAAAGGTTTAAA AAATTTCAACAAAAGCGTACTTTACATATATATTTATTAGACAAGA AAAGCAGATTAAATAGATATACATTCGATTAACGATAAGTAAAATG TAAAATCACAGGATTTTCGTGTGTGGTCTTCTACACAGACAAGGTGA ACAATTCGGCATTAAATACCTGAGAGCAGGAAGAGCAAGATAAAAG GTAGTATTTGTTGGCGATCCCCCTAGAGTCTTTTACATCTTCGAAAAA CAAAAACCTATTTTTTCTTTAATTTCTTTTTTTACTTTCTATTTTTAATTT ATATATTTATATTAATAAATTTAAATTATAATTATTTTTATAGCACGT GATGAAAAGGACCCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAA CCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCA		
--	--	--

TGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAA GAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGC GGCATTTTGCCTTCCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGT AAAAGATGCTGAAGATCAGTTGGGACGCGTAGTCTAGACCAGCCAG GACAGAAATGCCTCGACTTCGCTGCTACCCAAGGTTGCCGGGTGACG CACACCGTGGAAACGGATGAAGGCACGAACCCAGTGGACATAAGCC TGTCGGTTCGTAAGCTGTAATGCAAGTAGCGTATGCGCTCACGCAA CTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGTG GCGGTTTTTCATGGCTTGTTATGACTGTTTTTTTGGGGTACAGTCTATG CCTCGGGCATCCAAGCAGCAAGCGCGTTACGCCGTGGGTTCGATGTTT GATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCCCT AAAACAAAGTTAAACATTATGAGGGAAGCGGTGATCGCCGAAGTAT CGACTCAACTATCAGAGGTAGTTGGCGCCATCGAGCGCCATCTCGAA CCGACGTTGCTGGCCGTACATTTGTACGGCTCCGCAGTGGATGGCGG CCTGAAGCCACACAGTGATATTGATTTGCTGGTTACGGTGACCGTAA GGCTTGATGAAACAACGCGGCGAGCTTTGATCAACGACCTTTTGAA ACTTCGGCTTCCCCTGGAGAGAGCGAGATTCTCCGCGCTGTAGAAGT CACCATTGTTGTGCACGACGACATCATTCCGTGGCGTTATCCAGCTA AGCGCGAACTGCAATTTGGAGAATGGCAGCGCAATGACATTCTTGC AGGTATCTTCGAGCCAGCCACGATCGACATTGATCTGGCTATCTTGC TGACAAAAGCAAGAGAACATAGCGTTGCCTTGGTAGGTCCAGCGGC GGAGGAACTCTTTGATCCGGTTCCTGAACAGGATCTATTTGAGGCGC TAAATGAAACCTTAACGCTATGGAACTCGCCGCCCCGACTGGGCTGGC GATGAGCGAAATGTAGTGCTTACGTTGTCCCGCATTGTTGACAGCGC AGTAACCGGCAAAATCGCGCCGAAGGATGTCGCTGCCGGCTGGGCA ATGGAGCGCCTGCCGGCCCAGTATCAGCCCCTCATACTTGAAGCTAG ACAGGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTCGCGCGCA GATCAGTTGGAAGAATTTGTCCACTACGTGAAAGGCGAGATCACCA AGGTAGTCGGCAAATAACCCTCGAGCATTCAAGGCGCCTTGATTATT TGACGTGGTTTTGATGGCCTCCACGCACGTTGTGATATGTAGATGATT CAGTTCGAGTTTATCATTATCAATACTGCCATTTCAAAGAATACGTA AATAATTAATAGTAGTGATTTTCCTAACTTTATTTAGTCAAAAAATTA GCCTTTTAATTCTGCTGTAACCCGTACATGCCCAAATAGGGGGCGG GTTACACAGAATATATAACATCGTAGGTGTCTGGGTGAACAGTTTAT TCCTGGCATCCACTAAATATAATGGAGCCCGCTTTTTAAGCTGGCAT CCAGAAAAAAAAGAATCCCAGCACCAAAATATTGTTTTCTTCACCA ACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAGAACAG		
---	--	--

GGGCACAAACAGGCAAAAAACGGGCACAACCTCAATGGAGTGATGC AACCTGCCTGGAGTAAATGATGACACAAGGCAATTGACCCACGCAT GTATCTATCTCATTTTCTTACACCTTCTATTACCTTCTGCTCTCTCTGA TTTGGA AAAAGCTGAAAAAAAAGGTTGAAACCAGTTCCTGAAATT ATTCCCCTACTTGACTAATAAGTATATAAAGACGGTAGGTATTGATT GTAATTCTGTAAATCTATTTCTTAAACTTCTTAAATTCTACTTTTATA GTTAGTCTTTTTTTTAGTTTTAAAACACCAAGAACTTAGTTTCGAATA AACACACATAAAACAAACAAAATGGAACCTCAGTACTTCTCCTATTTT CAACCCACTTCATCTGTCGTAGCCCTACTACTAGCACTAGTGAGTAT TTTTATTTAGCGTAGTTGTTTTGAGGAAGACTTTCAGTAACAATTACTC CAGCCCCGCGTCAAGTACGGAAACCGCTGTGCTGTGTCATCAGAGGC AACAGAGTTGCGCCCTACCTATCAGCGGCCTTCTTCACGTGTTCATG AATAAGAACGGCCTGATTCATGTCACCTTGGGAAATATGGCTGACAA ATATGGCCCTATCTTCAGTTTTCCGACAGGCAGCCACCGTACTTTAGT AGTCAGTTCCTGGGAAATGGTGAAAGAGTGTTTCACCGGTAATAACG ACACGGCATTCTCCAACAGACCAATCCCTTTGGCTTTTCAAACCATA TTCTACGCCTGTGGCGGCATTGATTCTTACGGTTTAAGTAGTGTCCTG TATGGTAAATACTGGAGGGAGTTGAGAAAGGTGTGTGTTCAACACCT GCTGAGTAATCAGCAATTGCTGAAGTTCAGACATCTTATAATCTCCC AAGTGGATACGTCTTTTAAACAAGTTGTATGAGCTGTGTAAGAACTCT GAAGATAATCAAGGTATGGTAAGGATGGATGATTGGCTAGCTCAAC TTTCCTTTAACGTCATCGGTAGGATCGTTTGCGGATTCCAGTCTGACC CAAAGACGGGTGCACCTTCAAGGGTAGAACAGTTTAAGGAAGTCAT AAATGAGGCGTCATATTTTATGTCAACAAGTCCAGTCTCCGATAACG TACCAATGTTGGGATGGATCGACCAATTGACCGGTCTGACGAGGAA CATGAAGCATTGTGGGAAGAAGCTTGACTTAGTAGTGGAGTCAATTA TCAAGGACCATAGGCAAAAGAGACGTTTTTCACGTACAAAAGGTGG CGATGAGAAGGATGACGAACAGGACGACTTTATTGATATTTGCTTGA GCATCATGGAGCAGCCACAGTTGCCCGGGAACAATTCTCCCCCTCAA ATTCCGATCAAATCTATCGTGCTAGACATGATTGGGGGTGGTACCGA CACTACGAAACTTACAACCATATGGACCCTATCACTTTTGTTGAACA ATCCTCACGTGTTAGATAAAGCTAAACAAGAGGTGACGCTCACTTT CGTAAAAAGAGAAGATCAACAGATGACGCAGCAGCGGCAGTCGTTG ATTTTGACGACATAAGAAATTTAGTATACATCCAAGCCATCATTA GAAAGTATGAGGCTTTATCCAGCCAGCCCGGTGGTTGAGCGTCTTTC CGGCGAGGATTGCGTTGTTGGAGGTTTTACGTGCCTGCTGGTACGA GACTATGGGCTAACGTTTGAAGATGCAAAGAGATCCCAAAGTTTG		
--	--	--

GGACGATCCTCTAGTATTCAGACCTGAAAGGTTTTTGAGCGACGAGC AAAAGATGGTAGACGTTTCGTGGCCAAAACCTATGAACTTCTGCCATTC GGCGCAGGAAGAAGAATCTGTCCAGGCGTTTCCTTTAGTCTTGACCT TATGCAACTTGTCTAACCAGGTTAATCCTAGAGTTGAAATGAAGT CCCCCTCCGGCAAGGTAGATATGACCGCAACTCCAGGACTAATGTCT TACAAGGTGGTTCCATTGGACATATTGCTGACTCACCGTCGTATCAA GTCATGCGTTCAATTGGCGTCTTCTGAACGTGATATGGAAAGTTCTG GGGTGCCTGTGATCACATTGTCCTCAGGTAAAGTAATGCCCCTACTG GGCATGGGAACCTTCGAAAAGGTGGGTAAAGGGTCTGAACGTGAGC GTTTAGCCATTCTTAAAGCGATCGAAGTTGGTTACCGTTACTTTGATA CCGCAGCGGCATATGAAACGGAAGAAGTTCTAGGGGAAGCCATTGC TGAAGCTTTACAATTGGGTCTGATAGAGAGCCGTGACGAGCTGTTCA TCAGCTCAATGCTTTGGTGCACCGACGCACATCCAGACCGTGTGCTA CTTGCTCTGCAAAACAGTCTGAGAAATCTAAAACCTGAATATCTAGA CCTATATATGTTGCCGTTTCCTGCCAGCCTTAAGCCGGGCAAAATTA CGATGGATATTCTGAGGAGGATATTTGCCGTATGGATTATCGTTCA GTCTGGAGCGCCATGGAAGAGTGTCAAAACCTTAGGATTTACTAAAA GTATTGGTGTAAAGCAACTTTTCTTGCAAGAAATTACAAGAATTAATG GCCACTGCAAATATCCCGCCCGCGGTAAATCAAGTAGAGATGTCACC AGCTTTCCAACAGAAAAAACTGAGGGAATATTGTAACGCAAACAAC ATATTGGTATCCGCAGTAAGCATTCTGGGATCAAACGGGACGCCCTG GGGTAGTAATGCTGTTCTTGGAAGCGAAGTTTTGAAACAGATCGCGA TGCGGAAAGGCAAAAGCGTTGCGCAAGTCAGTATGAGGTGGGTCTA TGAGCAGGGCGCGTCTTTAGTAGTCAAGAGTTTCTCTGAAGAACGTT TAAGAGAAAACCTGAATATTTTTGACTGGGAGCTTACGAAAGAAGA CAATGAGAAGATAGGCGAAATCCCGCAATGTAGAATCCTTACTGCG TACTTCCTTGTCTCCCCGAACGGCCCGTTTAAATCTCAGGAAGAGCT TTGGGATGACAAGGCAtaaACAGGCCCTTTTCCTTTGTCGATATCATG TAATTAGTTATGTCACGCTTACATTCACGCCCTCCTCCACATCCGCT CTAACCGAAAAGGAAGGAGTTAGACAACCTGAAGTCTAGGTCCCTA TTTATTTTTTTAATAGTTATGTTAGTATTAAGAACGTTATTTATATT CAAATTTTTCTTTTTTTCTGTACAAACGCGTGTACGCATGTAACATT ATACTGAAAACCTTGCTTGAGAAGGTTTTGGGACGCTCGAAGGCTTT AATTTGTAATCATTATCACTTTACGGGTCCTTTCCGGTGATCCGACAG GTTACGGGGCGGCGACCTCGCGGGTTTTCGCTATTTATGAAAATTTT CCGGTTTAAGGCGTTTCCGTTCTTCTTCGTCATAACTTAATGTTTTTA TTTAAATACCTCGCGAGTGGCAACACTGAAAATACCCATGGAGCG		
--	--	--

GCGTAACCGTTCGCACAGgatctaggtgaagatccttttgataatctcatgacccaaatcccttaacg tgagtttctgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttctgcgctaa tctgctgcttcaaacaacccaccgctaccagcgggtggttgttggcgatcaagagctaccaactcttttc cgaaggtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagccgtagttaggccaccacttc aagaactctgtagcaccgcctacatacctcgtctgtctaactcgttaccagtggtcgtgccagtggcgataagtc gtgtcttaccgggttgactcaagacgatgttaccggataaggcgcagcggcggcgtgaacggggggttcgt gcacacagcccagcttgagcgaacgacctacaccgaactgagatactacagcgtgagctatgagaaagcgc cacgctcccgaaggagaaagcggacaggtatccggaagcggcagggcgggaacaggagagcgcacga gggagctccaggggaaacgcctggtatctttatagtcctgtcgggttcgccacctctgacttgagcgtcgatttt gtgatctcgtcagggggcggagcctatggaaaaacgccagcaacgcggcagtggaacTGCATTAT GAATTAGTTACGCTAGGGATAACAGGGTAATATAGAACCCGAACGA CCGAGCGCAGCGGCGGCCGCGCTGATACCGCCGC		
CCTCGCCGCAGTTAATTAAAGTCAGTGAGCGAGGAAGCGCGTA ATAACGGTCCTAAGGTAGCGAATCCTGATGCGGTATTTTCTCCTTAC GCATCTGTGCGGTATTTACACCCGCATAGATCGGCAAGTGCACAAAC AATACTTAAATAAATACTACTCAGTAATAACCTATTTCTTAGCATTTT TGACGAAATTTGCTATTTTGTAGAGTCTTTTACACCATTTGTCTCCA CACCTCCGCTTACATCAACACCAATAACGCCATTTAATCTAAGCGCA TCACCAACATTTTCTGGCGTCAGTCCACCAGCTAACATAAAATGTAA GCTTTCGGGGCTCTCTTGCCTTCCAACCCAGTCAGAAATCGAGTTCC AATCCAAAAGTTCACCTGTCCCACCTGCTTCTGAATCAAACAAGGGA ATAAACGAATGAGGTTTCTGTGAAGCTGCACTGAGTAGTATGTTGCA GTCTTTTGGAAATACGAGTCTTTTAATAACTGGCAAACCGAGGAACT CTTGGTATTCTTGCCACGACTCATCTCCATGCAGTTGGACGATATCA ATGCCGTAATCATTGACCAGAGCCAAAACATCCTCCTTAAGTTGATT ACGAAACACGCCAACCAAGTATTTTCGGAGTGCCTGAACTATTTTTAT ATGCTTTTACAAGACTTGAAATTTTCCTTGCAATAACCGGGTCAATT GTTCTCTTTCTATTGGGCACACATATAATACCCAGCAAGTCAGCATC GGAATCTAGAGCACATTCTGCGGCCTCTGTGCTCTGCAAGCCGCAAA CTTTCACCAATGGACCAGAATACTACCTGTGAAATTAATAACAGACATA CTCCAAGCTGCCTTTGTGTGCTTAATCACGTATACTCACGTGCTCAAT AGTCACCAATGCCCTCCCTCTTGGCCCTCTCCTTTTCTTTTTTCGACC GAATTAATTCTTAATCGGCAAAAAAGAAAAGCTCCGGATCAAGAT TGTACGTAAGGTGACAAGCTATTTTCAATAAAGAATATCTTCCACT ACTGCCATCTGGCGTCATAACTGCAAAGTACACATATATTACGATGC TGTCTATTAAATGCTTCCTATATTATATATATAGTAATGTCGTGATC TATGGTGCACCTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGC	pDW18	SEQ. ID NO. 25

CAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTG TCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGA GCTGCATGTGTCAGAGGTTTTACCGTCATCACCGAAACGCGCGAGA CGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTAAATGTCATGAT AATAATGGTTTTCTTAGACGGATCGCTTGCCTGTAACCTACACGCGCC TCGTATCTTTTAATGATGGAATAATTTGGGAATTTACTCTGTGTTTAT TTATTTTTATGTTTTGTATTTGGATTTTAGAAAAGTAAATAAAGAAGGT AGAAGAGTTACGGAATGAAGAAAAAAAATAAACAAAGGTTTAAA AAATTTCAACAAAAAGCGTACTTTACATATATATTTATTAGACAAGA AAAGCAGATTAAATAGATATACATTCGATTAACGATAAGTAAAATG TAAAATCACAGGATTTTCGTGTGTGGTCTTCTACACAGACAAGGTGA ACAATTCGGCATTAAATACCTGAGAGCAGGAAGAGCAAGATAAAAG GTAGTATTTGTTGGCGATCCCCCTAGAGTCTTTTACATCTTCGGAAAA CAAAAACCTATTTTTTCTTTAATTTCTTTTTTTACTTTCTATTTTAAATTT ATATATTTATATTAATAAAATTTAAATTATAATTATTTTTATAGCACGT GATGAAAAGGACCCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAA CCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCA TGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAA GAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGC GGCATTTTGCCTTCCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGT AAAAGATGCTGAAGATCAGTTGGGACGCGTAGTCTAGACCAGCCAG GACAGAAATGCCTCGACTTCGCTGCTACCCAAGGTTGCCGGGTGACG CACACCGTGGAAACGGATGAAGGCACGAACCCAGTGGACATAAGCC TGTTCCGTTCTGTAAGCTGTAATGCAAGTAGCGTATGCGCTCACGCAA CTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGTG GCGGTTTTTCATGGCTTGTTATGACTGTTTTTTTGGGGTACAGTCTATG CCTCGGGCATCCAAGCAGCAAGCGCGTTACGCCGTGGGTTCGATGTTT GATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCCCT AAAACAAAGTTAAACATTATGAGGGAAGCGGTGATCGCCGAAGTAT CGACTCAACTATCAGAGGTAGTTGGCGCCATCGAGCGCCATCTCGAA CCGACGTTGCTGGCCGTACATTTGTACGGCTCCGCAGTGGATGGCGG CCTGAAGCCACACAGTGATATTGATTTGCTGGTTACGGTGACCGTAA GGCTTGATGAAACAACGCGGCGAGCTTTGATCAACGACCTTTTGGAA ACTTCGGCTTCCCCTGGAGAGAGCGAGATTCTCCGCGCTGTAGAAGT CACCATTGTTGTGCACGACGACATCATTCCGTGGCGTTATCCAGCTA AGCGCGAACTGCAATTTGGAGAATGGCAGCGCAATGACATTCTTGC AGGTATCTTCGAGCCAGCCACGATCGACATTGATCTGGCTATCTTGC		
---	--	--

TGACAAAAGCAAGAGAACATAGCGTTGCCTTGGTAGGTCCAGCGGC GGAGGAACTCTTTGATCCGGTTCCTGAACAGGATCTATTTGAGGCGC TAAATGAAACCTTAACGCTATGGAACCTCGCCGCCCCGACTGGGCTGGC GATGAGCGAAATGTAGTGCTTACGTTGTCCCGCATTTGGTACAGCGC AGTAACCGGCAAAATCGCGCCGAAGGATGTCGCTGCCGGCTGGGCA ATGGAGCGCCTGCCGGCCCAGTATCAGCCCGTCATACTTGAAGCTAG ACAGGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTCGCGCGCA GATCAGTTGGAAGAATTTGTCCACTACGTGAAAGGCGAGATCACCA AGGTAGTCGGCAAATAACCCTCGAGCATTCAAGGCGCCTTGATTATT TGACGTGGTTTGATGGCCTCCACGCACGTTGTGATATGTAGATGAGA GCGTTGGTTGGTGGATCAAGCCCACGCGTAGGCAATCCTCGAGCAG ATCCGCCAGGCGTGTATATATAGCGTGGATGGCCAGGCAACTTTAGT GCTGACACATACAGGCATATATATATGTGTGCGACGACACATGATCA TATGGCATGCATGTGCTCTGTATGTATATAAACTCTTGTTCCTTCT TTTCTCTAAATATTCTTTCCTTATACATTAGGACCTTTGCAGCATAAA TTACTATACTTCTATAGACACACAAACACAAATACACACACTAAATT AATAATGGAACCTCAGTACTTCTCCTATTTTCAACCCACTTCATCTGT CGTAGCCCTACTACTAGCACTAGTGAGTATTTTATTTAGCGTAGTTGT TTTGAGGAAGACTTTCAGTAACAATTACTCCAGCCCCGCGTCAAGTA CGGAAACCGCTGTGCTGTGTCATCAGAGGCAACAGAGTTGCGCCCTA CCTATCAGCGGCCTTCTTCACGTGTTTATGAATAAGAACGGCCTGAT TCATGTCACCTTGGGAAATATGGCTGACAAATATGGCCCTATCTTCA GTTTTCCGACAGGCAGCCACCGTACTTTAGTAGTCAGTTCCTGGGAA ATGGTGAAAGAGTGTTTTACCGGTAATAACGACACGGCATTCTCCAA CAGACCAATCCCTTTGGCTTTTCAAACCATATTCTACGCCTGTGGCG GCATTGATTCTTACGGTTTAAGTAGTGTCCTGATGGTAAATACTGG AGGGAGTTGAGAAAGGTGTGTGTTTACAACCTGCTGAGTAATCAGC AATTGCTGAAGTTCAGACATCTTATAATCTCCCAAGTGGATACGTCT TTTAACAAGTTGTATGAGCTGTGTAAGAAGTCTGAAGATAATCAAGG TATGGTAAGGATGGATGATTGGCTAGCTCAACTTTCCTTTAACGTCA TCGGTAGGATCGTTTGCGGATTCCAGTCTGACCCAAAGACGGGTGCA CCTTCAAGGGTAGAACAGTTTAAAGGAAGTCATAAATGAGGCGTCAT ATTTTATGTCAACAAGTCCAGTCTCCGATAACGTACCAATGTTGGGA TGGATCGACCAATTGACCGGTCTGACGAGGAACATGAAGCATTGTG GGAAGAAGCTTGACTTAGTAGTGAGTCAATTATCAAGGACCATAG GCAAAAGAGACGTTTTTACGTACAAAAGGTGGCGATGAGAAGGAT GACGAACAGGACGACTTTATTGATATTTGCTTGAGCATCATGGAGCA		
---	--	--

GCCACAGTTGCCCGGGAACAATTCTCCCCCTCAAATTCGGATCAAAT CTATCGTGCTAGACATGATTGGGGGTGGTACCGACACTACGAAACTT ACAACCATATGGACCCTATCACTTTTGTGTAACAATCCTCACGTGTT AGATAAAGCTAAACAAGAGGTCGACGCTCACTTTCGTAAAAAGAGA AGATCAACAGATGACGCAGCAGCGGCAGTCGTTGATTTTGACGACA TAAGAAATTTAGTATACATCCAAGCCATCATTAAAGAAAGTATGAG GCTTTATCCAGCCAGCCCGGTGGTTGAGCGTCTTCCGGCGAGGATT GCGTTGTTGGAGGTTTTACGTCGCTGGTACGAGACTATGGGCT AACGTTTGAAGATGCAAAGAGATCCCAAAGTTTGGGACGATCCTCT AGTATTCAGACCTGAAAGGTTTTTGAGCGACGAGCAAAAGATGGTA GACGTTTCGTGGCCAAACTATGAACTTCTGCCATTCGGCGCAGGAAG AGAATCTGTCCAGGCGTTTCCTTTAGTCTTGACCTTATGCAACTGT CCTAACCAGGTTAATCCTAGAGTTCGAAATGAAGTCCCCGTCCGGCA AGGTAGATATGACCGCAACTCCAGGACTAATGTCTTACAAGGTGGTT CCATTGGACATATTGCTGACTACCGTCGTATCAAGTCATGCGTTCA ATTGGCGTCTTCTGAACGTGATATGGAAAGTTCTGGGGTGCCTGTGA TCACATTGTCCTCAGGTAAAGTAATGCCCGTACTGGGCATGGGAACC TTCGAAAAGGTGGGTAAGGGGTCTGAACGTGAGCGTTTAGCCATTCT TAAAGCGATCGAAGTTGGTTACCGTTACTTTGATACCGCAGCGGCAT ATGAAACGGAAGAAGTTCTAGGGGAAGCCATTGCTGAAGCTTTACA ATTGGGTCTGATAGAGAGCCGTGACGAGCTGTTTCATCAGCTCAATGC TTTGGTGCACCGACGCACATCCAGACCGTGTGCTACTTGCTCTGCAA AACAGTCTGAGAAATCTAAACTTGAATATCTAGACCTATATATGTT GCCGTTTCTGCCAGCCTTAAGCCGGGCAAAATTACGATGGATATTC CTGAGGAGGATATTTGCCGTATGGATTATCGTTCAGTCTGGAGCGCC ATGGAAGAGTGTCAAACTTAGGATTTACTAAAAGTATTGGTGTAAG CACTTTTCTTGCAAGAAATTACAAGAATTAATGGCCACTGCAAATA TCCCGCCCGCGGTAAATCAAGTAGAGATGTCACCAGCTTTCCAACAG AAAAAACTGAGGGAATATTGTAACGCAAACAACATATTGGTATCCG CAGTAAGCATTCTGGGATCAAACGGGACGCCCTGGGGTAGTAATGC TGTTCTTGGAAGCGAAGTTTTGAAACAGATCGCGATGGCGAAAGGC AAAAGCGTTGCGCAAGTCAGTATGAGGTGGGTCTATGAGCAGGGCG CGTCTTTAGTAGTCAAGAGTTTCTCTGAAGAACGTTTAAGAGAAAAC CTGAATATTTTTGACTGGGAGCTTACGAAAGAAGACAATGAGAAGA TAGGCGAAATCCCGCAATGTAGAATCCTTACTGCGTACTTCCTTGTC TCCCCGAACGGCCCGTTTAAATCTCAGGAAGAGCTTTGGGATGACAA GGCA _{taa} ACAGGCCCTTTTCTTTGTGCGATATCATGTAATTAGTTATG		
--	--	--

TCACGCTTACATTACGCCCCTCCTCCACATCCGCTCTAACCGAAAA GGAAGGAGTTAGACAACCTGAAGTCTAGGTCCCTATTTATTTTTTTT AATAGTTATGTTAGTATTAAGAACGTTATTTATATTTCAAATTTTTCT TTTTTTTCTGTACAAACGCGTGTACGCATGTAACATTATACTGAAAA CCTTGCTTGAGAAGGTTTTGGGACGCTCGAAGGCTTTAATTTGTAAT CATTATCACTTTACGGGTCTTTCCGGTGATCCGACAGGTTACGGGG CGGCGACCTCGCGGGTTTTCGCTATTTATGAAAATTTTCCGGTTTAAG GCGTTTCCGTTCTTCTTCGTCATAACTTAATGTTTTTATTTAAAATAC CTCGCGAGTGGCAACACTGAAAATACCCATGGAGCGGCGTAACCGT CGCACAGgatctagtgaaagatccttttgataatctcatgacaaaatcccttaacgtgagtttcgttccactg agcgtcagaccccgtagaaaagatcaaaggatcttcttgatcctttttctgcgcgtaatctgctgcttgcaca aaaaaccaccgctaccagcgggtgttgttggcgatcaagactacaaactctttccgaagtaactggctt cagcagagcgcagatacacaatactgtccttctagtgtagccgtagtaggcccaccacttcaagaactctgtagca ccgctacatacctcgtctgctaactctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttg gactcaagacgatatgtaccggataaggcgagcggctgggctgaacggggggtcgtgcacacagcccagct tggagcgaacgacctacccgaactgagatacctacagcgtgagctatgagaaagcggcacgcttcccgaagg gagaaaggcggacaggtatccggtaagcggcagggtcggacaggagagcgcacgaggagcttccaggg ggaacgcctggtatctttatagctctgctgggttcgccacctctgacttgagcgtcgattttgtgatgctcgtcag ggggcgaggcctatgaaaaacgccagcaacgcggcagtggaacTGCATTATGAATTAGT TACGCTAGGGATAACAGGGTAATATAGAACCCGAACGACCGAGCGC AGCGGCGGCCGCGCTGATACCGCCGC		
CCTCGCCGCAGTTAATTAAAGTCAGTGAGCGAGGAAGCGCGTAACT ATAACGGTCCTAAGGTAGCGAATCCTGATGCGGTATTTTCTCCTTAC GCATCTGTGCGGTATTTACACCCGCATAGATCGGCAAGTGCACAAAC AATACTTAAATAAATACTACTCAGTAATAACCTATTTCTTAGCATTTT TGACGAAATTTGCTATTTTGTAGAGTCTTTTACACCATTTGTCTCCA CACCTCCGCTTACATCAACACCAATAACGCCATTTAATCTAAGCGCA TCACCAACATTTTCTGGCGTCAGTCCACCAGCTAACATAAAATGTAA GCTTTCGGGGCTCTCTTGCCTTCCAACCCAGTCAGAAATCGAGTTCC AATCCAAAAGTTCACCTGTCCCACCTGCTTCTGAATCAAACAAGGGA ATAAACGAATGAGGTTTCTGTGAAGCTGCACTGAGTAGTATGTTGCA GTCTTTTGGAAATACGAGTCTTTTAATAACTGGCAAACCGAGGAACT CTTGGTATTCTTGCCACGACTCATCTCCATGCAGTTGGACGATATCA ATGCCGTAATCATTGACCAGAGCCAAAACATCCTCCTTAAGTTGATT ACGAAACACGCCAACCAAGTATTTTCGGAGTGCCTGAACTATTTTTAT ATGCTTTTACAAGACTTGAAATTTTCCTTGCAATAACCGGGTCAATT GTTCTCTTTCTATTGGGCACACATATAATAACCCAGCAAGTCAGCATC	pDW21	SEQ. ID NO. 26

GGAATCTAGAGCACATTCTGCGGCCTCTGTGCTCTGCAAGCCGCAAA CTTTCACCAATGGACCAGAACTACCTGTGAAATTAATAACAGACATA CTCCAAGCTGCCTTTGTGTGCTTAATCACGTATACTCACGTGCTCAAT AGTCACCAATGCCCTCCCTCTTGGCCCTCTCCTTTTCTTTTTTCGACC GAATTAATTCTTAATCGGCAAAAAAAGAAAAGCTCCGGATCAAGAT TGTACGTAAGGTGACAAGCTATTTTTCAATAAAGAATATCTTCCACT ACTGCCATCTGGCGTCATAACTGCAAAGTACACATATATTACGATGC TGTTCTATTAAATGCTTCCTATATTATATATATAGTAATGTCGTGATC TATGGTGCACCTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGC CAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTG TCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGA GCTGCATGTGTCAGAGGTTTTACCGTCATCACCGAAACGCGCGAGA CGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGAT AATAATGGTTTTCTTAGACGGATCGCTTGCCTGTAACCTACACGCGCC TCGTATCTTTTAATGATGGAATAATTTGGGAATTTACTCTGTGTTTAT TTATTTTTATGTTTTGTATTTGGATTTTAGAAAAGTAAATAAAGAAGGT AGAAGAGTTACGGAATGAAGAAAAAAAAAATAAACAAAGGTTTAAA AAATTTCAACAAAAGCGTACTTTACATATATATTTATTAGACAAGA AAAGCAGATTAAATAGATATACATTCGATTAAACGATAAGTAAAATG TAAAATCACAGGATTTTCGTGTGTGGTCTTCTACACAGACAAGGTGA ACAATTCGGCATTAAATACCTGAGAGCAGGAAGAGCAAGATAAAAAG GTAGTATTTGTTGGCGATCCCCCTAGAGTCTTTTACATCTTCGGAAAA CAAAAACCTATTTTTTCTTTAATTTCTTTTTTTACTTTCTATTTTTAATTT ATATATTTATATTAATAAATTTAAATTATAATTATTTTTATAGCACGT GATGAAAAGGACCCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAA CCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCA TGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAA GAGTATGAGTATTCAACATTTCCGTGTGCGCCCTATTCCCTTTTTTGC GGCATTTTGCCTTCCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGT AAAAGATGCTGAAGATCAGTTGGGACGCGTAGTCTAGACCAGCCAG GACAGAAATGCCCTCGACTTCGCTGCTACCCAAGGTTGCCGGGTGACG CACACCGTGGAAACGGATGAAGGCACGAACCCAGTGGACATAAGCC TGTTTCGGTTCGTAAGCTGTAATGCAAGTAGCGTATGCGCTCACGCAA CTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGTG GCGGTTTTTCATGGCTTGTTATGACTGTTTTTTTGGGGTACAGTCTATG CCTCGGGCATCCAAGCAGCAAGCGCGTTACGCCGTGGGTTCGATGTTT GATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCCCT		
---	--	--

AAAACAAAGTTAAACATTATGAGGGAAGCGGTGATCGCCGAAGTAT CGACTCAACTATCAGAGGTAGTTGGCGCCATCGAGCGCCATCTCGAA CCGACGTTGCTGGCCGTACATTTGTACGGCTCCGCAGTGGATGGCGG CCTGAAGCCACACAGTGATATTGATTTGCTGGTTACGGTGACCGTAA GGCTTGATGAAACAACGCGGCGAGCTTTGATCAACGACCTTTTGAA ACTTCGGCTTCCCCTGGAGAGAGCGAGATTCTCCGCGCTGTAGAAGT CACCATTGTTGTGCACGACGACATCATTCCGTGGCGTTATCCAGCTA AGCGCGAACTGCAATTTGGAGAATGGCAGCGCAATGACATTCTTGC AGGTATCTTCGAGCCAGCCACGATCGACATTGATCTGGCTATCTTGC TGACAAAAGCAAGAGAACATAGCGTTGCCTTGGTAGGTCCAGCGGC GGAGGAACTCTTTGATCCGGTTCCTGAACAGGATCTATTTGAGGCGC TAAATGAAACCTTAACGCTATGGAACTCGCCGCCCGACTGGGCTGGC GATGAGCGAAATGTAGTGCTTACGTTGTCCCGCATTTGGTACAGCGC AGTAACCGGCAAAATCGCGCCGAAGGATGTCGCTGCCGGCTGGGCA ATGGAGCGCCTGCCGCCCCAGTATCAGCCCGTCATACTTGAAGCTAG ACAGGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTCGCGCGCA GATCAGTTGGAAGAATTTGTCCACTACGTGAAAGGCGAGATCACCA AGGTAGTCGGCAAATAACCCCTCGAGCATTCAAGGCGCCTTGATTATT TGACGTGGTTTGATGGCCTCCACGCACGTTGTGATATGTAGATGAGA GCGTTGGTTGGTGGATCAAGCCCACGCGTAGGCAATCCTCGAGCAG ATCCGCCAGGCGTGTATATATAGCGTGGATGGCCAGGCAACTTTAGT GCTGACACATACAGGCATATATATATGTGTGCGACGACACATGATCA TATGGCATGCATGTGCTCTGTATGTATATAAACTCTTGTTTTCTTCT TTTCTCTAAATATTCTTTCCCTTATACATTAGGACCTTTGCAGCATAAA TTACTATACTTCTATAGACACACAAACACAAATACACACACTAAATT AATAATGGAACCTCAGTACTTCTCCTATTTTCAACCCACTTCATCTGT CGTAGCCCTACTACTAGCACTAGTGAGTATTTTATTTAGCGTAGTTGT TTTGAGGAAGACTTTCAGTAACAATTACTCCAGCCCCGCGTCAAGTA CGGAAACCGCTGTGCTGTGTCATCAGAGGCAACAGAGTTGCGCCCTA CCTATCAGCGGCCTTCTTACGTGTTTCATGAATAAGAACGGCCTGAT TCATGTACCTTGGGAAATATGGCTGACAAATATGGCCCTATCTTCA GTTTTCCGACAGGCAGCCACCGTACTTTAGTAGTCAGTTCCTGGGAA ATGGTGAAAGAGTGTTTACCGGTAATAACGACACGGCATTCTCCAA CAGACCAATCCCTTTGGCTTTTCAAACCATATTCTACGCCTGTGGCG GCATTGATTCTTACGGTTTAAAGTAGTGTCGCGTATGGTAAATACTGG AGGGAGTTGAGAAAGGTGTGTGTTTCAACCTGCTGAGTAATCAGC AATTGCTGAAGTTCAGACATCTTATAATCTCCCAAGTGGATACGTCT		
---	--	--

TTTAACAAGTTGTATGAGCTGTGTAAGAAGCTCTGAAGATAATCAAGG TATGGTAAGGATGGATGATTGGCTAGCTCAACTTTTCCTTTAACGTCA TCGGTAGGATCGTTTGCGGATTCCAGTCTGACCCAAAGACGGGTGCA CCTTCAAGGGTAGAACAGTTTAAGGAAGTCATAAATGAGGCGTCAT ATTTTATGTCAACAAGTCCAGTCTCCGATAACGTACCAATGTTGGGA TGGATCGACCAATTGACCGGTCTGACGAGGAACATGAAGCATTGTG GGAAGAAGCTTGACTTAGTAGTGGAGTCAATTATCAAGGACCATAG GCAAAAGAGACGTTTTTTCACGTACAAAAGGTGGCGATGAGAAGGAT GACGAACAGGACGACTTTATTGATATTTGCTTGAGCATCATGGAGCA GCCACAGTTGCCCGGGAACAATTCTCCCCCTCAAATTCGGATCAAAT CTATCGTGCTAGACATGATTGGGGGTGGTACCGACACTACGAAACTT ACAACCATATGGACCCTATCACTTTTGTGTAACAATCCTCACGTGTT AGATAAAGCTAAACAAGAGGTCGACGCTCACTTTCGTAAAAAGAGA AGATCAACAGATGACGCAGCAGCGGCAGTCGTTGATTTTGACGACA TAAGAAATTTAGTATACATCCAAGCCATCATTAAAGAAAGTATGAG GCTTTATCCAGCCAGCCCGGTGGTTGAGCGTCTTTCCGGCGAGGATT GCGTTGTTGGAGGTTTTTCACGTGCCTGCTGGTACGAGACTATGGGCT AACGTTTGAAGATGCAAAGAGATCCCAAAGTTTGGGACGATCCTCT AGTATTCAGACCTGAAAGGTTTTTGTAGCGACGAGCAAAAGATGGTA GACGTTTCGTGGCCAAAACCTATGAACTTCTGCCATTTCGGCGCAGGAAG AGAATCTGTCCAGGCGTTTCCTTTAGTCTTGACCTTATGCAACTTGT CCTAACCAGGTTAATCCTAGAGTTCGAAATGAAGTCCCGTCCGGCA AGGTAGATATGACCGCAACTCCAGGACTAATGTCCTTACAAGGTGGTT CCATTGGACATATTGCTGACTACCGTCGTATCAAGTCATGCGTTCA ATTGGCGTCTTCTGAACGTGAT ^{taa} GCGAATTTCTTATGATTTATGATT TTTATTATTAAATAAGTTATAAAAAAATAAGTGTATACAAATTTTA AAGTGACTCTTAGGTTTTAAAACGAAAATTCCTTATTCTTGAGTAACT CTTTCCTGTAGGTCAGGTTGCTTTCCTCAGGTACATAGCTTCAAAATGT TTCTACTCCTTTTTTACTCTTCCAGATTTTCTCGGACTCCGCGCATCGC CGTACCACCTTCAAAACACCCAAGCACAGCATACTAAATTTCCCCTCT TTCTTCCTCTAGGGTGTGTTAATTACCCGTAATAAGGTTTGGA AGAAAAAAGAGACCGCCTCGTTTCTTTTCTTCGTCGAAAAAGGCAA TAAAAATTTTTATCACGTTTCTTTTCTTGAAAATTTTTTTTTTTGATT TTTTTCTTTTCGATGACCTCCCATTTGATATTTAAGTTAATAAACGGT CTTCAATTTCTCAAGTTTCAGTTTCATTTTTCTTGTTCTATTACA TTTACTTCTTGCTCATTAGAAAGAAAGCATAGCAATCTAATCTAA GTTTAAATTACAAAATGGAAAGTTCTGGGGTGCCTGTGATCACATTG		
--	--	--

TCCTCAGGTAAAGTAATGCCCCGTAAGGGCATGGGAACCTTCGAAAA GGTGGGTAAGGGGTCTGAACGTGAGCGTTTAGCCATTCTTAAAGCGA TCGAAGTTGGTTACCGTTACTTTGATACCGCAGCGGCATATGAAACG GAAGAAGTTCTAGGGGAAGCCATTGCTGAAGCTTTACAATTGGGTCT GATAGAGAGCCGTGACGAGCTGTTTCATCAGCTCAATGCTTTGGTGCA CCGACGCACATCCAGACCGTGTGCTACTTGCTCTGCAAAACAGTCTG AGAAATCTAAACTTGAATATCTAGACCTATATATGTTGCCGTTTCC TGCCAGCCTTAAGCCGGGCAAAATTACGATGGATATTCTTGAGGAG GATATTTGCCGTATGGATTATCGTTCAGTCTGGAGCGCCATGGAAGA GTGTCAAAACTTAGGATTTACTAAAAGTATTGGTGTAAGCAACTTTT CTTGCAAGAAATTACAAGAATTAATGGCCACTGCAAATATCCCGCCC GCGGTAAATCAAGTAGAGATGTCACCAGCTTTCCAACAGAAAAAAC TGAGGGAATATTGTAACGCAACAACATATTGGTATCCGCAGTAAG CATTCTGGGATCAAACGGGACGCCCTGGGGTAGTAATGCTGTTCTTG GAAGCGAAGTTTTGAAACAGATCGCGATGGCGAAAGGCCAAAAGCGT TGCGCAAGTCAGTATGAGGTGGGTCTATGAGCAGGGCGCGTCTTTAG TAGTCAAGAGTTTCTCTGAAGAACGTTTAAGAGAAAACCTGAATATT TTTGACTGGGAGCTTACGAAAGAAGACAATGAGAAGATAGGCGAAA TCCCGCAATGTAGAATCCTTACTGCGTACTTCCTTGTCTCCCCGAACG GCCCCGTTTAAATCTCAGGAAGAGCTTTGGGATGACAAGGCAtaaACAG GCCCCTTTTCTTTGTGATATCATGTAATTAGTTATGTCACGCTTAC ATTCACGCCCTCCTCCCACATCCGCTCTAACCGAAAAGGAAGGAGTT AGACAACCTGAAGTCTAGGTCCCTATTTATTTTTTTTAATAGTTATGT TAGTATTAAGAACGTTATTTATATTTCAAATTTTTCTTTTTTTCTGTA CAAACGCGTGTACGCATGTAACTTATACTGAAAACCTTGCTTGAGA AGGTTTTGGGACGCTCGAAGGCTTTAATTTGTAATCATTATCACTTTA CGGGTCCTTTCCGGTGATCCGACAGGTTACGGGGCGGCGACCTCGCG GGTTTTCGCTATTTATGAAAATTTCCGGTTTAAGGCGTTTCCGTTCT TCTTCGTCATAACTTAATGTTTTTATTTAAATACCTCGCGAGTGCCA ACACTGAAAATACCCATGGAGCGGCGTAACCGTCGCACAGgatctaggtg aagatccttttgataatctcatgacaaaaatccctaacgtgagtttctgtccactgagcgtcagaccccgtagaaa agatcaaaggatcttcttgagatcctttttctgcgcgtaactctgctgcttgcaacaaaaaaccaccgctaccagc ggtggtttgttgcggatcaagagctaccaactcttttccgaaggtactggttcagcagagcgagataccaa atactgtccttctagtgtagccgtagttaggccaccactcaagaactctgtagcaccgcctacacctcgtctgct aatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttgactcaagacgatagttaccgg ataaggcgagcggtcgggctgaacggggggtcgtgcacacagcccagcttgagcgaacgacacacccg aactgagatacctacagcgtgagctatgagaaagccacgcttccgaaggagaaaggcgacaggtatcc		
--	--	--

ggtaagcggcagggctcgaacaggagagcgcacagggagctccaggggaaacgcctggatctttatagt cctgtcgggtttcggcaccctctgacttgagcgtcgatgtttgtgatgctcgtcagggggcgagcctatggaaaaa cgccagcaacgcggcagtggaacgTGCATTATGAATTAGTTACGCTAGGGATAAC AGGGTAATATAGAACCCGAACGACCGAGCGCAGCGGCGGCCGCGCT GATACCGCCGC		
CCTCGCCGCAGTTAATTAAAGTCAGTGAGCGAGGAAGCGCGTAACT ATAACGGTCCTAAGGTAGCGAATCCTGATGCGGTATTTTCTCCTTAC GCATCTGTGCGGTATTTACACCCGCATAGATCGGCAAGTGCACAAAC AATACTTAAATAAATACTACTCAGTAATAACCTATTTCTTAGCATTTT TGACGAAATTTGCTATTTTGTAGAGTCTTTTACACCATTTGTCTCCA CACCTCCGCTTACATCAACACCAATAACGCCATTTAATCTAAGCGCA TCACCAACATTTTCTGGCGTCAGTCCACCAGCTAACATAAAATGTAA GCTTTCGGGGCTCTCTTGCCTTCCAACCCAGTCAGAAATCGAGTTCC AATCCAAAAGTTCACCTGTCCACCTGCTTCTGAATCAAACAAGGGA ATAAACGAATGAGGTTTCTGTGAAGCTGCACTGAGTAGTATGTTGCA GTCTTTTGGAAATACGAGTCTTTTAATAACTGGCAAACCGAGGAACT CTTGGTATTCTTGCCACGACTCATCTCCATGCAGTTGGACGATATCA ATGCCGTAATCATTGACCAGAGCCAAAACATCCTCCTTAAGTTGATT ACGAAACACGCCAACCAAGTATTTTCGGAGTGCCTGAACTATTTTTAT ATGCTTTTACAAGACTTGAAATTTTCCTTGCAATAACCGGGTCAATT GTTCTCTTTCTATTGGGCACACATATAATAACCCAGCAAGTCAGCATC GGAATCTAGAGCACATTCTGCGGCCTCTGTGCTCTGCAAGCCGCAAA CTTTCACCAATGGACCAGAACTACCTGTGAAATTAATAACAGACATA CTCCAAGCTGCCTTTGTGTGCTTAATCACGTATACTCACGTGCTCAAT AGTCACCAATGCCCTCCCTCTTGGCCCTCTCCTTTTCTTTTTTCGACC GAATTAATTCTTAATCGGCAAAAAAAGAAAAGCTCCGGATCAAGAT TGTACGTAAGGTGACAAGCTATTTTCAATAAAGAATATCTTCCACT ACTGCCATCTGGCGTCATAACTGCAAAGTACACATATATTACGATGC TGTCTATTAAATGCTTCCTATATTATATATATAGTAATGTCGTGATC TATGGTGCCTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGC CAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTG TCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGA GCTGCATGTGTCAGAGGTTTTACCGTCATCACCGAAACGCGCGAGA CGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGAT AATAATGGTTTCTTAGACGGATCGCTTGCCTGTAACCTACACGCGCC TCGTATCTTTTAATGATGGAATAATTTGGGAATTTACTCTGTGTTTAT TTATTTTTATGTTTTGTATTTGGATTTTAGAAAGTAAATAAAGAAGGT	pJL29	SEQ. ID NO. 27

AGAAGAGTTACGGAATGAAGAAAAAAAAAATAAACAAAGGTTTAAA AAATTTCAACAAAAGCGTACTTTACATATATATTTATTAGACAAGA AAAGCAGATTAAATAGATATACATTCGATTAACGATAAGTAAAATG TAAAATCACAGGATTTTCGTGTGTGGTCTTCTACACAGACAAGGTGA AACAAATTCGGCATTAAATACCTGAGAGCAGGAAGAGCAAGATAAAAG GTAGTATTTGTTGGCGATCCCCCTAGAGTCTTTTACATCTTCGGAAAA CAAAACTATTTTTTCTTTAATTTCTTTTTTTACTTTCTATTTTTAATTT ATATATTTATATTAATAAATTAAATTATAATTATTTTTATAGCACGT GATGAAAAGGACCCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAA CCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCA TGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAA GAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGC GGCATTTTGCCTTCCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGT AAAAGATGCTGAAGATCAGTTGGGACGCGTAGTCTAGACCAGCCAG GACAGAAATGCCTCGACTTCGCTGCTACCCAAGGTTGCCGGGTGACG CACACCGTGGAAACGGATGAAGGCACGAACCCAGTGGACATAAGCC TGTTTCGGTTCGTAAGCTGTAATGCAAGTAGCGTATGCGCTCACGCAA CTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGTG GCGGTTTTTCATGGCTTGTTATGACTGTTTTTTTGGGGTACAGTCTATG CCTCGGGCATCCAAGCAGCAAGCGCGTTACGCCGTGGGTTCGATGTTT GATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCCCT AAAACAAAGTTAAACATTATGAGGGAAGCGGTGATCGCCGAAGTAT CGACTCAACTATCAGAGGTAGTTGGCGCCATCGAGCGCCATCTCGAA CCGACGTTGCTGGCCGTACATTTGTACGGCTCCGCAGTGGATGGCGG CCTGAAGCCACACAGTGATATTGATTTGCTGGTTACGGTGACCGTAA GGCTTGATGAAACAACGCGGCGAGCTTTGATCAACGACCTTTTGAA ACTTCGGCTTCCCCTGGAGAGAGCGAGATTCTCCGCGCTGTAGAAGT CACCATTGTTGTGCACGACGACATCATTCCGTGGCGTTATCCAGCTA AGCGCGAACTGCAATTTGGAGAATGGCAGCGCAATGACATTCTTGC AGGTATCTTCGAGCCAGCCACGATCGACATTGATCTGGCTATCTTGC TGACAAAAGCAAGAGAACATAGCGTTGCCTTGGTAGGTCCAGCGGC GGAGGAACTCTTTGATCCGGTTCCTGAACAGGATCTATTTGAGGCGC TAAATGAAACCTTAACGCTATGGAACTCGCCGCCGACTGGGCTGGC GATGAGCGAAATGTAGTGCTTACGTTGTCCCGCATTTGGTACAGCGC AGTAACCGGCAAAATCGCGCCGAAGGATGTCGCTGCCGGCTGGGCA ATGGAGCGCCTGCCGGCCCAGTATCAGCCCGTCATACTTGAAGCTAG ACAGGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTCGCGCGCA		
--	--	--

GATCAGTTGGAAGAATTTGTCCACTACGTGAAAGGCGAGATCACCA AGGTAGTCGGCAAATAACCCTCGAGCATTCAAGGCGCCTTGATTATT TGACGTGGTTTGATGGCCTCCACGCACGTTGTGATATGTAGATGATT CAGTTCGAGTTTATCATTATCAATACTGCCATTTCAAAGAATACGTA AATAATTAATAGTAGTGATTTTCCTAACTTTATTTAGTCAAAAAATTA GCCTTTTAATTCTGCTGTAACCCGTACATGCCCAAATAGGGGGCGG GTTACACAGAATATATAACATCGTAGGTGTCTGGGTGAACAGTTTAT TCCTGGCATCCACTAAATATAATGGAGCCCGCTTTTTAAGCTGGCAT CCAGAAAAAAAAGAATCCCAGCACCAAAATATTGTTTTCTTCACCA ACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAGAACAG GGGCACAAACAGGCAAAAAACGGGCACAACCTCAATGGAGTGATGC AACCTGCCTGGAGTAAATGATGACACAAGGCAATTGACCCACGCAT GTATCTATCTCATTTTCTTACACCTTCTATTACCTTCTGCTCTCTCTGA TTTGGA AAAAAGCTGAAAAAAAAGGTTGAAACCAGTTCCTGAAATT ATTCCCCTACTTGACTAATAAGTATATAAAGACGGTAGGTATTGATT GTAATTCTGTAAATCTATTTCTTAAACTTCTTAAATTCTACTTTTATA GTTAGTCTTTTTTTTAGTTTTAAAACACCAAGAACTTAGTTTCGAATA AACACACATAAAACAAACAAAATGGAACCTCAGTACTTCTCCTATTTT CAACCCACTTCATCTGTCGTAGCCCTACTACTAGCACTAGTGAGTAT TTTATTTAGCGTAGTTGTTTTGAGGAAGACTTTCAGTAACAATTACTC CAGCCCCGCGTCAAGTACGGAAACCGCTGTGCTGTGTCATCAGAGGC AACAGAGTTGCGCCCTACCTATCAGCGGCCTTCTTCACGTGTTTCATG AATAAGAACGGCCTGATTCATGTCACCTTGGGAAATATGGCTGACAA ATATGGCCCTATCTTCAGTTTTCCGACAGGCAGCCACCGTACTTTAGT AGTCAGTTCCTGGGAAATGGTGAAAGAGTGTTTCACCGGTAAATAACG ACACGGCATTCTCCAACAGACCAATCCCTTTGGCTTTTCAAACCATA TTCTACGCCTGTGGCGGCATTGATTCTTACGGTTTAAGTAGTGTCGCG TATGGTAAATACTGGAGGGAGTTGAGAAAGGTGTGTGTTCACAACTT GCTGAGTAATCAGCAATTGCTGAAGTTCAGACATCTTATAATCTCCC AAGTGGATACGTCTTTTAAACAAGTTGTATGAGCTGTGTAAGAACTCT GAAGATAATCAAGGTATGGTAAGGATGGATGATTGGCTAGCTCAAC TTTCCTTTAACGTCATCGGTAGGATCGTTTGCGGATTCCAGTCTGACC CAAAGACGGGTGCACCTTCAAGGGTAGAACAGTTTAAAGGAAGTCAT AAATGAGGCGTCATATTTTATGTCAACAAGTCCAGTCTCCGATAACG TACCAATGTTGGGATGGATCGACCAATTGACCGGTCTGACGAGGAA CATGAAGCATTGTGGGAAGAAGCTTGACTTAGTAGTGGAGTCAATTA TCAAGGACCATAGGCAAAAGAGACGTTTTTTCACGTACAAAAGGTGG		
---	--	--

CGATGAGAAGGATGACGAACAGGACGACTTTATTGATATTTGCTTGA GCATCATGGAGCAGCCACAGTTGCCCGGGAACAATTCTCCCCCTCAA ATTCCGATCAAATCTATCGTGCTAGACATGATTGGGGGTGGTACCGA CACTACGAAACTTACAACCATATGGACCCTATCACTTTTGTTGAACA ATCCTCACGTGTTAGATAAAGCTAAACAAGAGGTCGACGCTCACTTT CGTAAAAAGAGAAGATCAACAGATGACGCAGCAGCGGCAGTCGTTG ATTTTGACGACATAAGAAATTTAGTATACATCCAAGCCATCATTA GAAAGTATGAGGCTTTATCCAGCCAGCCCGGTGGTTGAGCGTCTTTC CGGCGAGGATTGCGTTGTTGGAGGTTTTACGTGCCTGCTGGTACGA GACTATGGGCTAACGTTTGAAGATGCAAAGAGATCCCAAAGTTTG GGACGATCCTCTAGTATTCAGACCTGAAAGGTTTTTGAGCGACGAGC AAAAGATGGTAGACGTTCTGGCCAAAACCTATGAACTTCTGCCATTC GGCGCAGGAAGAAGAATCTGTCCAGGCGTTTCCTTTAGTCTTGACCT TATGCAACTTGTCTAACCAGGTTAATCCTAGAGTTCGAAATGAAGT CCCCGTCCGGCAAGGTAGATATGACCGCAACTCCAGGACTAATGTCT TACAAGGTGGTTCCATTGGACATATTGCTGACTCACCGTCGTATCAA GTCATGCGTTCAATTGGCGTCTTCTGAACGTGAT^{taa}GCGAATTTCTTA TGATTTATGATTTTTATTATTAATAAGTTATAAAAAAATAAGTGT ATACAAATTTTAAAGTGACTCTTAGGTTTTAAAACGAAAATTCCTAT TCTTGAGTAACTCTTTCCTGTAGGTCAGGTTGCTTCTCAGGTACATA GCTTCAAAATGTTTCTACTCCTTTTTTACTCTTCCAGATTTTCTCGGAC TCCGCGCATCGCCGTACCACTTCAAAACACCCAAGCACAGCATACTA AATTTCCCTCTTTCTTCTCTAGGGTGTCTTAATTACCCGTAATAA AGGTTTGGAAAAGAAAAAGAGACCGCCTCGTTTCTTTTCTTCGTC GAAAAAGGCAATAAAAATTTTTATCACGTTTCTTTTCTTGAAAATTT TTTTTTTTGATTTTTTTCTCTTTCGATGACCTCCCATTGATATTTAAGT TAATAAACGGTCTTCAATTTCTCAAGTTTCAGTTTCATTTTTCTTGTTT TATTACAACTTTTTTTACTTCTTGCTCATTAGAAAGAAAGCATAGCAA TCTAATCTAAGTTTTAATTACAAAATGGAAAGTTCTGGGGTGCCTGT GATCACATTGTCCTCAGGTAAAGTAATGCCCCTACTGGGCATGGGAA CCTTCGAAAAGGTGGGTAAGGGGTCTGAACGTGAGCGTTTAGCCATT CTTAAAGCGATCGAAGTTGGTTACCGTTACTTTGATACCGCAGCGGC ATATGAAACGGAAGAAGTTCTAGGGGAAGCCATTGCTGAAGCTTTA CAATTGGGTCTGATAGAGAGCCGTGACGAGCTGTTTCATCAGCTCAAT GCTTTGGTGCACCGACGCACATCCAGACCGTGTGCTACTTGCTCTGC AAAACAGTCTGAGAAATCTAAAACCTGAATATCTAGACCTATATATG TTGCCGTTTCTGCCAGCCTTAAGCCGGGCAAAATTACGATGGATAT		
--	--	--

TCCTGAGGAGGATATTTGCCGTATGGATTATCGTTCAGTCTGGAGCG CCATGGAAGAGTGTCAAAACTTAGGATTTACTAAAAGTATTGGTGTA AGCAACTTTTCTTGCAAGAAATTACAAGAATTAATGGCCACTGCAAA TATCCCGCCCGCGGTAAATCAAGTAGAGATGTCACCAGCTTTCCAAC AGAAAAAACTGAGGGAATATTGTAACGCAAACAACATATTGGTATC CGCAGTAAGCATTCTGGGATCAAACGGGACGCCCTGGGGTAGTAAT GCTGTTCTTGGAAGCGAAGTTTTGAAACAGATCGCGATGGCGAAAG GCAAAAGCGTTGCGCAAGTCAGTATGAGGTGGGTCTATGAGCAGGG CGCGTCTTTAGTAGTCAAGAGTTTCTCTGAAGAACGTTTAAGAGAAA ACCTGAATATTTTTGACTGGGAGCTTACGAAAGAAGACAATGAGAA GATAGGCGAAATCCCGCAATGTAGAATCCTTACTGCGTACTTCCTTG TCTCCCGAACGGCCCGTTTAAATCTCAGGAAGAGCTTTGGGATGAC AAGGCAtaaACAGGCCCTTTTCCTTTGTCGATATCATGTAATTAGTTA TGTCACGCTTACATTCACGCCCTCCTCCACATCCGCTCTAACCGAA AAGGAAGGAGTTAGACAACCTGAAGTCTAGGTCCCTATTTATTTTTT TTAATAGTTATGTTAGTATTAAGAACGTTATTTATATTTCAAATTTTT CTTTTTTTTCTGTACAAACGCGTGTACGCATGTAAACATTATACTGAAA ACCTTGCTTGAGAAGGTTTTGGGACGCTCGAAGGCTTTAATTTGTAA TCATTATCACTTTACGGGTCTTTCCGGTGATCCGACAGGTTACGGG GCGGCGACCTCGCGGGTTTTTCGCTATTTATGAAAATTTTCCGGTTTAA GGCGTTTCCGTTCTTCTTCGTCATAACTTAATGTTTTTATTTAAAATA CCTCGCGAGTGGCAACACTGAAAATACCCATGGAGCGGCGTAACCG TCGCACAGgatctaggtgaagatccttttgataatctcatgacaaaatccctaacgtgagtttctgtccac tgagcgtcagaccccgtagaaaagatcaaaggatctcttgagatcctttttctgcgcgtaatctgctgcttgc caaaaaaaccaccgctaccagcgggtggttgggtccggatcaagagctaccaactcttttccgaaggtactggc ttcagcagagcgcagataccaaatactgtccttctagttagccgtagttaggccaccactcaagaactctgtagc accgcctacatactcgtctgctaactcgttaccagtggctgctgccagtggcgataagtcgtgtctaccgggtt ggactcaagacgatagttaccggataagggcagcgggtcgggctgaacggggggttcgtgcacacagcccag cttggagcgaacgacctacaccgaactgagatactacagcgtgagctatgagaaagcgcacgcttcccgaag ggagaaaggcggacaggtatccggtaagcggcagggcgggaacaggagagcgcacgaggagcttccagg gggaaacgcctggtatcttatagtcctgtcgggttcgccacctctgacttgagcgtcgattttgtgatgctcgtca ggggggcggagccctatggaaaaacgccagcaacgcggcagtggaacgTGCATTATGAATTAGT TACGCTAGGGATAACAGGGTAATATAGAACCCGAACGACCGAGCGC AGCGGCGGCCGCGCTGATACCGCCGC		
CCTCGCCGCAGTTAATTAAAGTCAGTGAGCGAGGAAGCGCGTAACT ATAACGGTCCTAAGGTAGCGAATCCTGATGCGGTATTTTCTCCTTAC GCATCTGTGCGGTATTTACACCCGCATAGATCGGCAAGTGCACAAAC	pJL32	SEQ. ID NO. 28

AACTTAAATAAATACTACTCAGTAATAACCTATTTCTTAGCATTTT TGACGAAATTTGCTATTTTGTAGAGTCTTTTACACCATTTGTCTCCA CACCTCCGCTTACATCAACACCAATAACGCCATTTAATCTAAGCGCA TCACCAACATTTTCTGGCGTCAGTCCACCAGCTAACATAAAATGTAA GCTTTCGGGGCTCTCTTGCCTTCCAACCCAGTCAGAAATCGAGTTCC AATCCAAAAGTTCACCTGTCCACCTGCTTCTGAATCAAACAAGGGA ATAAACGAATGAGGTTTCTGTGAAGCTGCACTGAGTAGTATGTTGCA GTCTTTTGGAAATACGAGTCTTTTAATAACTGGCAAACCGAGGAACT CTTGGTATTCTTGCCACGACTCATCTCCATGCAGTTGGACGATATCA ATGCCGTAATCATTGACCAGAGCCAAAACATCCTCCTTAAGTTGATT ACGAAACACGCCAACCAAGTATTTTCGGAGTGCCTGAACTATTTTTAT ATGCTTTTACAAGACTTGAAATTTTCCTTGCAATAACCGGGTCAATT GTTCTCTTTCTATTGGGCACACATATAATAACCCAGCAAGTCAGCATC GGAATCTAGAGCACATTCTGCGGCCTCTGTGCTCTGCAAGCCGCAAA CTTTCACCAATGGACCAGAACTACCTGTGAAATTAATAACAGACATA CTCCAAGCTGCCTTTGTGTGCTTAATCACGTATACTCACGTGCTCAAT AGTCACCAATGCCCTCCCTCTTGGCCCTCTCCTTTTCTTTTTTCGACC GAATTAATTCTTAATCGGCAAAAAAAGAAAAGCTCCGGATCAAGAT TGTACGTAAGGTGACAAGCTATTTTCAATAAAGAATATCTTCCACT ACTGCCATCTGGCGTCATAACTGCAAAGTACACATATATTACGATGC TGTCTATTAAATGCTTCCTATATTATATATATAGTAATGTCGTGATC TATGGTGCACCTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGC CAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTG TCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGA GCTGCATGTGTCAGAGGTTTTACCGTCATCACCGAAACGCGCGAGA CGAAAGGGCCTCGTGATACGCCATTTTTTATAGGTAAATGTCATGAT AATAATGGTTTCTTAGACGGATCGCTTGCCTGTAACCTACACGCGCC TCGTATCTTTTAATGATGGAATAATTTGGGAATTTACTCTGTGTTTAT TTATTTTTATGTTTTGTATTTGGATTTTAGAAAGTAAATAAAGAAGGT AGAAGAGTTACGGAATGAAGAAAAAAAATAAACAAAGGTTTAAA AAATTTCAACAAAAGCGTACTTTACATATATATTTATTAGACAAGA AAAGCAGATTAAATAGATATACATTCGATTAAACGATAAGTAAAATG TAAAATCACAGGATTTTCGTGTGTGGTCTTCTACACAGACAAGGTGA ACAATTCGGCATTAAATACCTGAGAGCAGGAAGAGCAAGATAAAAG GTAGTATTTGTTGGCGATCCCCCTAGAGTCTTTTACATCTTCGAAAAA CAAAAACCTATTTTTTCTTTAATTTCTTTTTTTACTTTCTATTTTTAATTT ATATATTTATATTAAAAAATTTAAATTATAATTATTTTTATAGCACGT		
--	--	--

GATGAAAAGGACCCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAA CCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCA TGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAA GAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGC GGCATTTCCTTCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGT AAAAGATGCTGAAGATCAGTTGGGACGCGTAGTCTAGACCAGCCAG GACAGAAATGCCTCGACTTCGCTGCTACCCAAGGTTGCCGGGTGACG CACACCGTGGAAACGGATGAAGGCACGAACCCAGTGGACATAAGCC TGTTCCGTTTCGTAAGCTGTAATGCAAGTAGCGTATGCGCTCACGCAA CTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGTG GCGGTTTTTCATGGCTTGTTATGACTGTTTTTTTGGGGTACAGTCTATG CCTCGGGCATCCAAGCAGCAAGCGCGTTACGCCGTGGGTTCGATGTTT GATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCCCT AAAACAAAGTTAAACATTATGAGGGAAGCGGTGATCGCCGAAGTAT CGACTCAACTATCAGAGGTAGTTGGCGCCATCGAGCGCCATCTCGAA CCGACGTTGCTGGCCGTACATTTGTACGGCTCCGCAGTGGATGGCGG CCTGAAGCCACACAGTGATATTGATTTGCTGGTTACGGTGACCGTAA GGCTTGATGAAACAACGCGGCGAGCTTTGATCAACGACCTTTTGAA ACTTCGGCTTCCCCTGGAGAGAGCGAGATTCTCCGCGCTGTAGAAGT CACCATTGTTGTGCACGACGACATCATTCCGTGGCGTTATCCAGCTA AGCGCGAACTGCAATTTGGAGAATGGCAGCGCAATGACATTCTTGC AGGTATCTTCGAGCCAGCCACGATCGACATTGATCTGGCTATCTTGC TGACAAAAGCAAGAGAACATAGCGTTGCCTTGGTAGGTCCAGCGGC GGAGGAACTCTTTGATCCGGTTCCTGAACAGGATCTATTTGAGGCGC TAAATGAAACCTTAACGCTATGGAACTCGCCGCCCGACTGGGCTGGC GATGAGCGAAATGTAGTGCTTACGTTGTCCCGCATTGTTGACAGCGC AGTAACCGGCAAAATCGCGCCGAAGGATGTCGCTGCCGGCTGGGCA ATGGAGCGCCTGCCGGCCAGTATCAGCCCGTCATACTTGAAGCTAG ACAGGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTCGCGCGCA GATCAGTTGGAAGAATTTGTCCACTACGTGAAAGGCGAGATCACCA AGGTAGTCGGCAAATAACCCTCGAGCATTCAAGGCGCCTTGATTATT TGACGTGGTTTGATGGCCTCCACGCACGTTGTGATATGTAGATGAGA GCGTTGGTTGGTGGATCAAGCCCACGCGTAGGCAATCCTCGAGCAG ATCCGCCAGGCGTGTATATATAGCGTGGATGGCCAGGCAACTTTAGT GCTGACACATACAGGCATATATATATGTGTGCGACGACACATGATCA TATGGCATGCATGTGCTCTGTATGTATATAAACTCTTGTTTTCTTCT TTTCTCTAAATATTCTTTCTTATACATTAGGACCTTTGCAGCATAAA		
--	--	--

TTACTATACTTCTATAGACACACAAACACAAATACACACACTAAATT AATAATGGAACCTTCAGTACTTCTCCTATTTTCAACCCACTTCATCTGT CGTAGCCCTACTACTAGCACTAGTGAGTATTTTATTTAGCGTAGTTGT TTTGAGGAAGACTTTTCAGTAACAATTACTCCAGCCCCGCGTCAAGTA CGGAAACCGCTGTGCTGTGTCATCAGAGGCAACAGAGTTGCGCCCTA CCTATCAGCGGCCTTCTTCACGTGTTTCATGAATAAGAACGGCCTGAT TCATGTCACCTTGGGAAATATGGCTGACAAATATGGCCCTATCTTCA GTTTTCCGACAGGCAGCCACCGTACTTTAGTAGTCAGTTCCTGGGAA ATGGTGAAAGAGTGTTTTACCGGTAATAACGACACGGCATTCTCCAA CAGACCAATCCCTTTGGCTTTTCAAACCATATTCTACGCCTGTGGCG GCATTGATTCTTACGGTTTAAGTAGTGTCCTGATGGTAAATACTGG AGGGAGTTGAGAAAGGTGTGTGTTACAAACCTGCTGAGTAATCAGC AATTGCTGAAGTTCAGACATCTTATAATCTCCCAAGTGGATACGTCT TTTAACAAGTTGTATGAGCTGTGTAAGAACTCTGAAGATAATCAAGG TATGGTAAGGATGGATGATTGGCTAGCTCAACTTTCCTTTAACGTCA TCGGTAGGATCGTTTGCGGATTCCAGTCTGACCCAAAGACGGGTGCA CCTTCAAGGGTAGAACAGTTTAAGGAAGTCATAAATGAGGCGTCAT ATTTTATGTCAACAAGTCCAGTCTCCGATAACGTACCAATGTTGGGA TGGATCGACCAATTGACCGGTCTGACGAGGAACATGAAGCATTGTG GGAAGAAGCTTGACTTAGTAGTGAGTCAATTATCAAGGACCATAG GCAAAAGAGACGTTTTTTCACGTACAAAAGGTGGCGATGAGAAGGAT GACGAACAGGACGACTTTATTGATATTTGCTTGAGCATCATGGAGCA GCCACAGTTGCCCCGGGAACAATTCTCCCCCTCAAATTCCGATCAAAT CTATCGTGCTAGACATGATTGGGGGTGGTACCGACACTACGAAACTT ACAACCATATGGACCCTATCACTTTTGTGTAACAATCCTCACGTGTT AGATAAAGCTAAACAAGAGGTGCGACGCTCACTTTCGTAAAAAGAGA AGATCAACAGATGACGCAGCAGCGGCAGTCGTTGATTTTGACGACA TAAGAAATTTAGTATACATCCAAGCCATCATTAAGAAAGTATGAG GCTTTATCCAGCCAGCCCGGTGGTTGAGCGTCTTCCGGCGAGGATT GCGTTGTTGGAGGTTTTACGTGCCTGCTGGTACGAGACTATGGGCT AACGTTTGGAAAGATGCAAAGAGATCCCAAAGTTTGGGACGATCCTCT AGTATTCAGACCTGAAAGGTTTTTGAGCGACGAGCAAAAGATGGTA GACGTTTCGTGGCCAAAACCTATGAACTTCTGCCATTCGGCGCAGGAAG AAGAATCTGTCCAGGCGTTTCCTTTAGTCTTGACCTTATGCAACTTGT CCTAACCAGGTTAATCCTAGAGTTCGAAATGAAGTCCCCGTCCGGCA AGGTAGATATGACCGCAACTCCAGGACTAATGTCTTACAAGGTGGTT CCATTGGACATATTGCTGACTCACCGTCGTATCAAGTCATGCGTTCA		
---	--	--

ATTGGCGTCTTCTGAACGTGATtaaGCGAATTTCTTATGATTTATGATT TTTATTATTAAATAAGTTATAAAAAAATAAGTGTATACAAATTTTA AAGTGACTCTTAGGTTTTAAAACGAAAATTCTTATTCTTGAGTAACT CTTTCCTGTAGGTCAGGTTGCTTCTCAGGTATTCAGTTCGAGTTTAT CATTATCAATACTGCCATTTCAAAGAATACGTAAATAATTAATAGTA GTGATTTTCTAACTTTATTTAGTCAAAAAATTAGCCTTTTAATTCTG CTGTAACCCGTACATGCCCAAATAGGGGGCGGGTTACACAGAATA TATAACATCGTAGGTGTCTGGGTGAACAGTTTATTCCTGGCATCCAC TAAATATAATGGAGCCCGCTTTTTAAGCTGGCATCCAGAAAAAAA AGAATCCCAGCACCAAAATATTGTTTTCTTCACCAACCATCAGTTCA TAGGTCCATTCTCTTAGCGCAACTACAGAGAACAGGGGCACAAACA GGCAAAAAACGGGCACAACCTCAATGGAGTGATGCAACCTGCCTGG AGTAAATGATGACACAAGGCAATTGACCCACGCATGTATCTATCTCA TTTTCTTACACCTTCTATTACCTTCTGCTCTCTCTGATTTGGAAAAAG CTGAAAAAAAAGGTTGAAACCAGTTCCTGAAATTATTCCCCTACTT GACTAATAAGTATATAAAGACGGTAGGTATTGATTGTAATTCTGTAA ATCTATTTCTTAACTTCTTAAATTCTACTTTTATAGTTAGTCTTTTTT TTAGTTTTAAAACACCAAGAACTTAGTTTTGAATAAACACACATAAA CAAACAAAATGGAAAGTTCTGGGGTGCCTGTGATCACATTGTCTCTCA GGTAAAGTAATGCCCCGTAAGTGGGCATGGGAACCTTCGAAAAGGTGG GTAAGGGGTCTGAACGTGAGCGTTTAGCCATTCTTAAAGCGATCGAA GTTGGTTACCGTTACTTTGATACCGCAGCGGCATATGAAACGGAAGA AGTTCTAGGGGAAGCCATTGCTGAAGCTTTACAATTGGGTCTGATAG AGAGCCGTGACGAGCTGTTTCATCAGCTCAATGCTTTGGTGCACCGAC GCACATCCAGACCGTGTGCTACTTGCTCTGCAAAACAGTCTGAGAAA TCTAAAACTTGAATATCTAGACCTATATATGTTGCCGTTTCCTGCCAG CCTTAAGCCGGGCAAAATTACGATGGATATTCCTGAGGAGGATATTT GCCGTATGGATTATCGTTTCAGTCTGGAGCGCCATGGAAGAGTGTCAA AACTTAGGATTTACTAAAAGTATTGGTGTAAGCAACTTTTCTTGCAA GAAATTACAAGAATTAATGGCCACTGCAAATATCCCGCCCGCGGTA AATCAAGTAGAGATGTCACCAGCTTTCCAACAGAAAAAACTGAGGG AATATTGTAACGCAAACAACATATTGGTATCCGCAGTAAGCATTCTG GGATCAAACGGGACGCCCTGGGGTAGTAATGCTGTTCTTGGAAGCG AAGTTTTGAAACAGATCGCGATGGCGAAAGGCAAAAGCGTTGCGCA AGTCAGTATGAGGTGGGTCTATGAGCAGGGCGCGTCTTTAGTAGTCA AGAGTTTCTCTGAAGAACGTTTAAGAGAAAACCTGAATATTTTTGAC TGGGAGCTTACGAAAGAAGACAATGAGAAGATAGGCGAAATCCCGC		
---	--	--

AATGTAGAATCCTTACTGCGTACTTCCTTGTCTCCCCGAACGGCCCCGT TTAAATCTCAGGAAGAGCTTTGGGATGACAAGGCAtaaACAGGCCCCCT TTTCCTTTGTGATATCATGTAATTAGTTATGTCACGCTTACATTAC GCCCTCCTCCACATCCGCTCTAACCGAAAAGGAAGGAGTTAGACA ACCTGAAGTCTAGGTCCCTATTTATTTTTTTTAATAGTTATGTTAGTA TTAAGAACGTTATTTATATTTCAAATTTTCTTTTTTTTTCTGTACAAAC GCGTGTACGCATGTAACATTATACTGAAAACCTTGCTTGAGAAGGTT TTGGGACGCTCGAAGGCTTTAATTTGTAATCATTATCACTTTACGGGT CCTTTCCGGTGATCCGACAGGTTACGGGGCGGCGACCTCGCGGGTTT TCGCTATTTATGAAAATTTTCCGGTTTAAGGCGTTTCCGTTCTTCTTC GTCATAACTTAATGTTTTTATTTAAAATACCTCGCGAGTGGCAACAC TGAAAATACCCATGGAGCGGCGTAACCGTCGCACAGgatctaggtgaagatcc ttttgataatctcatgacaaaaatccctaacgtgagtttctggtccactgagcgtcagaccccgtagaaaagatcaa aggatcttctgagatcctttttctgcgcgtaatctgctgcttcaaacaaaaaaaccaccgctaccagcgggtggtt tgttgccggatcaagagctaccaactcttttccgaaggtaactggcttcagcagagcgcagataccaaatactgt ccttctagtgtagccgtagtttagccaccactcaagaactctgtagcaccgcctacatacctcgtctgctaactct gttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttgactcaagacgatagttaccggataag gcgcagcggctgggctgaacggggggtctgtcacacagcccagcttggagcgaacgacctacaccgaactg agatacctacagcgtgagctatgagaaagcggcagcgttcccgaaggagaaaggcggacaggtatccggtaa gcggcagggctcggaacaggagagcgcacgagggagcttcagggggaacgcctggtatctttatagtcctgt cgggttccgccacctctgacttgagcgtcgattttgtgatgctcgtcagggggcgaggcctatgaaaaacgcc agcaacgcggcagtggaacgTGCATTATGAATTAGTTACGCTAGGGATAACAG GGTAATATAGAACCCGAACGACCGAGCGCAGCGGCGGCCGCGCTGA TACCGCCGC		
CCTCGCCGCAGTTAATTAAAGTCAGTGAGCGAGGAAGCGCGTAACT ATAACGGTCCTAAGGTAGCGAATCCTGATGCGGTATTTTCTCCTTAC GCATCTGTGCGGTATTTACACCCGCATAGATCGGCAAGTGCACAAAC AATACTTAAATAAATACTACTCAGTAATAACCTATTTCTTAGCATTTT TGACGAAATTTGCTATTTTGTAGAGTCTTTTACACCATTGTCTCCA CACCTCCGCTTACATCAACACCAATAACGCCATTTAATCTAAGCGCA TCACCAACATTTTCTGGCGTCAGTCCACCAGCTAACATAAAATGTAA GCTTTCGGGGCTCTCTTGCTTCCAACCCAGTCAGAAATCGAGTTCC AATCCAAAAGTTCACCTGTCCCACCTGCTTCTGAATCAAACAAGGGA ATAAACGAATGAGGTTTCTGTGAAGCTGCACTGAGTAGTATGTTGCA GTCTTTTGGAAATACGAGTCTTTTAATAACTGGCAAACCGAGGAACT CTTGGTATTCTTGCCACGACTCATCTCCATGCAGTTGGACGATATCA ATGCCGTAATCATTGACCAGAGCCAAAACATCCTCCTTAAGTTGATT	pJL35	SEQ. ID NO. 29

ACGAAACACGCCAACCAAGTATTTTCGGAGTGCCTGAACTATTTTTAT ATGCTTTTACAAGACTTGAAATTTTCCTTGCAATAACCGGGTCAATT GTTCTCTTTCTATTGGGCACACATATAATACCCAGCAAGTCAGCATC GGAATCTAGAGCACATTCTGCGGCCTCTGTGCTCTGCAAGCCGCAAA CTTTCACCAATGGACCAGAACTACCTGTGAAATTAATAACAGACATA CTCCAAGCTGCCTTTGTGTGCTTAATCACGTATACTCACGTGCTCAAT AGTCACCAATGCCCTCCCTCTTGGCCCTCTCCTTTTCTTTTTTCGACC GAATTAATTCTTAATCGGCAAAAAAAGAAAAGCTCCGGATCAAGAT TGTACGTAAGGTGACAAGCTATTTTTCAATAAAGAATATCTTCCACT ACTGCCATCTGGCGTCATAACTGCAAAGTACACATATATTACGATGC TGTCTATTAAATGCTTCCTATATTATATATATAGTAATGTCGTGATC TATGGTGCACCTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGC CAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTG TCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGA GCTGCATGTGTCAGAGGTTTTACCGTCATCACCGAAACGCGCGAGA CGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTAAATGTCATGAT AATAATGGTTTCTTAGACGGATCGCTTGCCTGTAACCTACACGCGCC TCGTATCTTTTAATGATGGAATAATTTGGGAATTTACTCTGTGTTTAT TTATTTTTATGTTTTGTATTTGGATTTTAGAAAGTAAATAAAGAAGGT AGAAGAGTTACGGAATGAAGAAAAAAAATAAACAAAGGTTTAAA AAATTTCAACAAAAAGCGTACTTTACATATATATTTATTAGACAAGA AAAGCAGATTAAATAGATATACATTCGATTAACGATAAGTAAATG TAAATCACAGGATTTTCGTGTGTGGTCTTCTACACAGACAAGGTGA ACAATTCGGCATTAAATACCTGAGAGCAGGAAGAGCAAGATAAAAG GTAGTATTTGTTGGCGATCCCCCTAGAGTCTTTTACATCTTCGGAAAA CAAAAACATTTTTTTCTTTAATTTCTTTTTTTACTTTCTATTTTTAATTT ATATATTTATATTAATAAAATTTAAATTATAATTATTTTTATAGCACGT GATGAAAAGGACCCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAA CCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCA TGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAA GAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGC GGCATTTTGCCTTCCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGT AAAAGATGCTGAAGATCAGTTGGGACGCGTAGTCTAGACCAGCCAG GACAGAAATGCCTCGACTTCGCTGCTACCCAAGGTTGCCGGGTGACG CACACCGTGGAACGGATGAAGGCACGAACCCAGTGACATAAGCC TGTTTCGGTTCGTAAGCTGTAATGCAAGTAGCGTATGCGCTCACGCAA CTGGTCCAGAACCTTGACCGAACGCAGCGGTGGTAACGGCGCAGTG		
---	--	--

GCGGTTTTTCATGGCTTGTTATGACTGTTTTTTTTGGGGTACAGTCTATG CCTCGGGCATCCAAGCAGCAAGCGCGTTACGCCGTGGGTTCGATGTTT GATGTTATGGAGCAGCAACGATGTTACGCAGCAGGGCAGTCGCCCT AAAACAAAGTTAAACATTATGAGGGAAGCGGTGATCGCCGAAGTAT CGACTCAACTATCAGAGGTAGTTGGCGCCATCGAGCGCCATCTCGAA CCGACGTTGCTGGCCGTACATTTGTACGGCTCCGCAGTGGATGGCGG CCTGAAGCCACACAGTGATATTGATTTGCTGGTTACGGTGACCGTAA GGCTTGATGAAACAACGCGGCGAGCTTTGATCAACGACCTTTTGGA ACTTCGGCTTCCCCTGGAGAGAGCGAGATTCTCCGCGCTGTAGAAGT CACCATTGTTGTGCACGACGACATCATTCCGTGGCGTTATCCAGCTA AGCGCGAACTGCAATTTGGAGAATGGCAGCGCAATGACATTCTTGC AGGTATCTTCGAGCCAGCCACGATCGACATTGATCTGGCTATCTTGC TGACAAAAGCAAGAGAACATAGCGTTGCCTTGGTAGGTCCAGCGGC GGAGGAACTCTTTGATCCGGTTCCTGAACAGGATCTATTTGAGGCGC TAAATGAAACCTTAACGCTATGGAATCGCCGCCCCGACTGGGCTGGC GATGAGCGAAATGTAGTGCTTACGTTGTCCCGCATTGTTGACAGCGC AGTAACCGGCAAAATCGCGCCGAAGGATGTCGCTGCCGGCTGGGCA ATGGAGCGCCTGCCGGCCCAGTATCAGCCCGTCATACTTGAAGCTAG ACAGGCTTATCTTGGACAAGAAGAAGATCGCTTGGCCTCGCGCGCA GATCAGTTGGAAGAATTTGTCCACTACGTGAAAGGCGAGATCACCA AGGTAGTCGGCAAATAACCCTCGAGCATTCAAGGCGCCTTGATTATT TGACGTGGTTTGATGGCCTCCACGCACGTTGTGATATGTAGATGATT CAGTTCGAGTTTATCATTATCAATACTGCCATTTCAAAGAATACGTA AATAATTAATAGTAGTGATTTTCCTAACTTTATTTAGTCAAAAAATTA GCCTTTTAATTCTGCTGTAACCCGTACATGCCCAAAATAGGGGGCGG GTTACACAGAATATATAACATCGTAGGTGTCTGGGTGAACAGTTTAT TCCTGGCATCCACTAAATATAATGGAGCCCGCTTTTTAAGCTGGCAT CCAGAAAAAAAAAGAATCCCAGCACCAAAATATTGTTTTCTTCACCA ACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAGAACAG GGGCACAACAGGCAAAAAACGGGCACAACCTCAATGGAGTGATGC AACCTGCCTGGAGTAAATGATGACACAAGGCAATTGACCCACGCAT GTATCTATCTCATTTTCTTACACCTTCTATTACCTTCTGCTCTCTCTGA TTTGGAAAAAGCTGAAAAAAAAGGTTGAAACCAGTTCCTGAAATT ATTCCCCTACTTGACTAATAAGTATATAAAGACGGTAGGTATTGATT GTAATTCTGTAAATCTATTTCTTAACTTCTTAAATTCTACTTTTATA GTTAGTCTTTTTTTTAGTTTTTAAACACCAAGAACTTAGTTTTGAATA AACACACATAAAACAAACAAAATGGAACCTCAGTACTTCTCCTATTTT		
--	--	--

CAACCCACTTCATCTGTCGTAGCCCTACTACTAGCACTAGTGAGTAT TTTATTTAGCGTAGTTGTTTTGAGGAAGACTTTCAGTAACAATTACTC CAGCCCCGCGTCAAGTACGGAAACCGCTGTGCTGTGTCATCAGAGGC AACAGAGTTGCGCCCTACCTATCAGCGGCCTTCTTCACGTGTTTCATG AATAAGAACGGCCTGATTCATGTCACCTTGGGAAATATGGCTGACAA ATATGGCCCTATCTTCAGTTTTCCGACAGGCAGCCACCGTACTTTAGT AGTCAGTTCCTGGGAAATGGTGAAAGAGTGTTTCACCGGTAATAACG ACACGGCATTCTCCAACAGACCAATCCCTTTGGCTTTTCAAACCATA TTCTACGCCTGTGGCGGCATTGATTCTTACGGTTTAAGTAGTGTCCCG TATGGTAAATACTGGAGGGAGTTGAGAAAGGTGTGTGTTCAACACCT GCTGAGTAATCAGCAATTGCTGAAGTTCAGACATCTTATAATCTCCC AAGTGGATACGTCTTTTAAACAAGTTGTATGAGCTGTGTAAGAACTCT GAAGATAATCAAGGTATGGTAAGGATGGATGATTGGCTAGCTCAAC TTTCCTTTAACGTCATCGGTAGGATCGTTTGC GGATTCCAGTCTGACC CAAAGACGGGTGCACCTTCAAGGGTAGAACAGTTTAAAGGAAGTCAT AAATGAGGCGTCATATTTTATGTCAACAAGTCCAGTCTCCGATAACG TACCAATGTTGGGATGGATCGACCAATTGACCGGTCTGACGAGGAA CATGAAGCATTGTGGGAAGAAGCTTGACTTAGTAGTGGAGTCAATTA TCAAGGACCATAGGCCAAAAGAGACGTTTTTCACGTACAAAAGGTGG CGATGAGAAGGATGACGAACAGGACGACTTTATTGATATTTGCTTGA GCATCATGGAGCAGCCACAGTTGCCCGGGAACAATTCTCCCCCTCAA ATTCCGATCAAATCTATCGTGCTAGACATGATTGGGGGTGGTACCGA CACTACGAAACTTACAACCATATGGACCCTATCACTTTTGTTGAACA ATCCTCACGTGTTAGATAAAGCTAAACAAGAGGTGACGCTCACTTT CGTAAAAAGAGAAGATCAACAGATGACGCAGCAGCGGCAGTCGTTG ATTTTGACGACATAAGAAATTTAGTATACATCCAAGCCATCATTA GAAAGTATGAGGCTTTATCCAGCCAGCCCGGTGGTTGAGCGTCTTTC CGGCGAGGATTGCGTTGTTGGAGGTTTTACGTGCCTGCTGGTACGA GACTATGGGCTAACGTTTGGAAGATGCAAAGAGATCCCAAAGTTTG GGACGATCCTCTAGTATTCAGACCTGAAAGGTTTTTGAGCGACGAGC AAAAGATGGTAGACGTTTCGTGGCCAAAACCTATGAACTTCTGCCATTC GGCGCAGGAAGAAGAATCTGTCCAGGCGTTTCCTTTAGTCTTGACCT TATGCAACTTGTCCTAACCAGGTAAATCCTAGAGTTCGAAATGAAGT CCCCGTCCGGCAAGGTAGATATGACCGCAACTCCAGGACTAATGTCT TACAAGGTGGTTCCATTGGACATATTGCTGACTACCGTCGTATCAA GTCATGCGTTCAATTGGCGTCTTCTGAACGTGAT ^{taa} GCGAATTTCTTA TGATTTATGATTTTTATTATTAAATAAGTTATAAAAAAATAAGTGT		
---	--	--

ATACAAATTTTAAAGTGACTCTTAGGTTTTAAAACGAAAATTCCTTAT TCTTGAGTAACTCTTTCCTGTAGGTCAGGTTGCTTCTCAGGTAGAGC GTTGGTTGGTGGATCAAGCCACGCGTAGGCAATCCTCGAGCAGATC CGCCAGGCGTGTATATATAGCGTGGATGGCCAGGCAACTTTAGTGCT GACACATACAGGCATATATATATGTGTGCGACAACACATGATCATAT GGCATGCATGTGCTCTGTATGTATATAAACTCTTGTTTTCTTCTTTT CTCTAAATATTCTTTCCTTATACATTAGGACCTTTGCAGCATAAATTA CTATACTTCTATAGACACACAAACACAAATACACACACTAAATTAAT AATGGAAAGTTCTGGGGTGCCTGTGATCACATTGTCCTCAGGTAAAG TAATGCCCGTACTGGGCATGGGAACCTTCGAAAAGGTGGGTAAAGG GTCTGAACGTGAGCGTTTAGCCATTCTTAAAGCGATCGAAGTTGGTT ACCGTTACTTTGATACCGCAGCGGCATATGAAACGGAAGAAGTTCTA GGGGAAGCCATTGCTGAAGCTTTACAATTGGGTCTGATAGAGAGCC GTGACGAGCTGTTTCATCAGCTCAATGCTTTGGTGCACCGACGCACAT CCAGACCGTGTGCTACTTGCTCTGCAAAACAGTCTGAGAAATCTAAA ACTTGAATATCTAGACCTATATATGTTGCCGTTTCCTGCCAGCCTTAA GCCGGGCAAATTACGATGGATATTCCTGAGGAGGATATTTGCCGTA TGGATTATCGTTCAGTCTGGAGCGCCATGGAAGAGTGTCAAACTTA GGATTTACTAAAAGTATTGGTGTAAGCAACTTTTCTTGCAAGAAATT ACAAGAATTAATGGCCACTGCAAATATCCCGCCCGCGGTAAATCAA GTAGAGATGTCACCAGCTTTC AACAGAAAAAACTGAGGGAATATT GTAACGCAAACAACATATTGGTATCCGCGAGTAAGCATTCTGGGATCA AACGGGACGCCCTGGGGTAGTAATGCTGTTCTTGGAAGCGAAGTTTT GAAACAGATCGCGATGGCGAAAGGCAAAGCGTTGCGCAAGTCAGT ATGAGGTGGGTCTATGAGCAGGGCGCGTCTTTAGTAGTCAAGAGTTT CTCTGAAGAACGTTTAAAGAGAAAACCTGAATATTTTGACTGGGAGC TTACGAAAGAAGACAATGAGAAGATAGGCGAAATCCCGCAATGTAG AATCCTTACTGCGTACTTCCTTGTCTCCCGAACGGCCCGTTTAAATC TCAGGAAGAGCTTTGGGATGACAAGGCA^{taa}ACAGGCCCTTTTCCTT TGTCGATATCATGTAATTAGTTATGTCACGCTTACATTCACGCCCTCC TCCCACATCCGCTCTAACC GAAAAGGAAGGAGTTAGACAACCTGAA GTCTAGGTCCCTATTTATTTTTTTTAAATAGTTATGTTAGTATTAAGAA CGTTATTTATATTTCAAATTTTTCTTTTTTTCTGTACAAACGCGTGTA CGCATGTAACATTATACTGAAAACCTTGCTTGAGAAGGTTTTGGGAC GCTCGAAGGCTTTAATTTGTAATCATTATCACTTTACGGGTCCTTTCC GGTGATCCGACAGGTTACGGGGCGGCGACCTCGCGGGTTTTCGCTAT TTATGAAAATTTCCGGTTTAAAGGCGTTTCCGTTCTTCTTCGTCATAA		
--	--	--

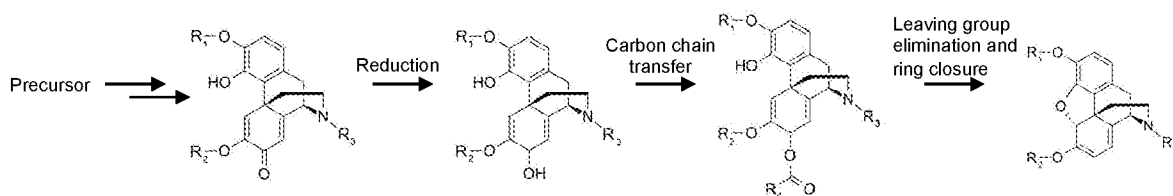
CTTAATGTTTTTATTTAAATACCTCGCGAGTGGCAACACTGAAAAT ACCCATGGAGCGGCGTAACCGTCGCACAGGatctaggtgaagatccttttgataatctc atgacaaaaatccctaacgtgagtttctgtccactgagcgtcagaccccgtagaaaagatcaaaggatcttctga gatcctttttctgcgcgtaatctgctgcttgcaacaaaaaaccaccgctaccagcgggtgttggttgccggtac aagagctaccaactcttttccgaagtaactggcttcagcagagcgcagataccaaatactgtccttctagtgtagc cgtagttagggccaccacttcaagaactctgtagcaccgcctacatacctcgtctgctaactcgttaccagtggctg ctgccagtggcgataagtcgtgtcttaccgggtggactcaagacgatagttaccggataaggcgcagcggtcgg gctgaacggggggtcgtgcacacagcccagcttgagcgaacgacctacaccgaactgagatacctacagcg tgagctatgagaaagcgccacgcttcccgaaggagaaaggcggacaggtatccggttaagcggcagggtcgg aacaggagagcgcacgagggagcttcaggggaaacgcctggtatctttagtctgtcgggttcgccacct ctgacttgagcgtcgattttgtgatgctcgtcagggggcggagcctatgaaaaacgccagcaacgcggcagt ggaacgTGCATTATGAATTAGTTACGCTAGGGATAACAGGGTAATATAG AACCCGAACGACCGAGCGCAGCGGCGGCCGCGCTGATACCGCCGC		
---	--	--

Morphinan Alkaloid Generating Modifications

[00143] Some methods, processes, and systems provided herein describe the conversion of promorphinan alkaloids to morphinan alkaloids. Some of the methods, processes, and systems describe the conversion of a tetracyclic scaffold to a pentacyclic scaffold (**FIG. 28**). Some of the methods, processes, and systems may comprise an engineered host cell. In some examples, the production of pentacyclic thebaine, or a morphinan alkaloid, from a tetracyclic precursor, or a promorphinan alkaloid is described. In some examples, the conversion of promorphinan alkaloids to thebaine are key steps in the conversion of a substrate to a diverse range of benzylisoquinoline alkaloids.

[00144] In some examples, the tetracyclic precursor may be salutaridine, salutaridinol, or salutaridinol-7-*O*-acetate. The tetracyclic precursor may be converted to pentacyclic thebaine by closure of an oxide bridge between C-4 and C-5. In some examples, the tetracyclic precursor salutaridine may be prepared for ring closure by stepwise hydroxylation and *O*-acetylation at C-7. Ring closure may be activated by elimination of an acetate leaving group. In some examples, the allylic elimination and oxide ring closure that generates thebaine occurs spontaneously. In other examples, the ring closure reaction that generates pentacyclic thebaine is promoted by factors such as pH or solvent. In other examples, the thebaine-generating ring closure reaction is promoted by contact with a protein or enzyme. These conversion steps are provided in **FIG. 14** and represented generally in **Scheme 2**. R₁, R₂, and R₃ may be H or CH₃. R₄ may be CH₃, CH₃CH₂, CH₃CH₂CH₂, or other appropriate alkyl group. In some cases, R₁, R₂, R₃, and R₄ may be CH₃ as provided in **FIG. 14**.

Scheme 2



[00145] In some examples, the first enzyme that prepares the tetracyclic precursor is salutaridine reductase (SalR). In some cases, SalR hydroxylates the substrate salutaridine at the C-7 position (see Formula III). The product of this reaction may be one or more salutaridinol epimers. In some examples, the product is (7S)-salutaridinol. In some examples, the salutaridine reductase may catalyze the reduction reaction within a host cell, such as an engineered host, as described herein.

[00146] In some examples, the second enzyme that prepares the tetracyclic precursor is salutaridinol 7-*O*-acetyltransferase (SalAT). In some cases, SalAT transfers the acetyl from acetyl-CoA to the 7-OH of salutaridinol (see Formula IV). In other cases, SalAT may utilize a novel cofactor such as *n*-propionyl-CoA and transfer the propionyl to the 7-OH of salutaridinol. In some examples, the product of SalAT is (7S)-salutaridinol-7-*O*-acetate. In some examples, the salutaridinol 7-*O*-acetyltransferase may catalyze the acetyl transfer reaction within a host cell, such as an engineered host, as described herein.

[00147] In some examples, the tetracyclic precursor of thebaine is (7S)-salutaridinol-7-*O*-acetate. In some examples (7S)-salutaridinol-7-*O*-acetate is unstable and spontaneously eliminates the acetate at C-7 and closes the oxide bridge between C-4 and C-5 to form thebaine (see Formula V). In some examples, the rate of elimination of the acetate leaving group is promoted by pH. In some examples, the allylic elimination and oxide bridge closure is catalyzed by an enzyme with thebaine synthase activity, or a thebaine synthase. In some examples, this enzyme is a Bet v 1-fold protein. In some examples, this enzyme is an engineered thebaine synthase, an engineered SalAT, a dirigent (DIR) protein, or a chalcone isomerase (CHI). In some examples, the enzyme encoding thebaine synthase activity may catalyze the ring closure reaction within a host cell, such as an engineered host, as described herein.

[00148] In some examples, the salutaridine reductase enzyme may be SalR or a SalR-like enzyme from plants in the *Ranunculales* order that biosynthesize thebaine, for example *Papaver somniferum*. In other examples, the enzyme with salutaridine reductase activity may be from mammals or any other vertebrate or invertebrate that biosynthesizes endogenous morphine.

[00149] In some examples, the salutaridinol 7-*O*-acetyltransferase enzyme may be SalAT or a SalAT-like enzyme from plants in the *Ranunculales* order that biosynthesize thebaine, for example *P. somniferum*. In other examples, the enzyme with salutaridinol 7-*O*-acetyltransferase activity may be from mammals or any other vertebrate or invertebrate that biosynthesizes endogenous morphine.

[00150] In some examples, the thebaine synthase enzyme may be a Bet v 1 fold protein from plants in the *Ranunculales* order that biosynthesize thebaine, for example *P. somniferum*. In some examples, the Bet v 1 protein includes the following domains in order from the N-terminus to C-terminus: a β -strand, one or two α -helices, six β -strands, and one or two α -helices. The protein is organized such that it has a Bet v 1 fold and an active site that accepts large, bulky, hydrophobic molecules, such as morphinan alkaloids. This protein may be any plant Bet v 1 protein, pathogenesis-related 10 protein (PR-10), a major latex protein (MLP), fruit or pollen allergen, plant hormone binding protein (e.g., binding to cytokinin or brassinosteroids), plant polyketide cyclase-like protein, or norcoclaurine synthase (NCS)-related protein that has a Bet v 1 fold. Other non-plant examples of the Bet v 1 fold protein are polyketide cyclases, activator of Hsp90 ATPase homolog 1 (AHA1) proteins, SMU440-like proteins (e.g., from *Streptococcus mutans*), PA1206-related proteins (e.g., from *Pseudomonas aeruginosa*), CalC calicheamicin resistance protein (e.g., from *Micromonospora echinospora*), and the CoxG protein from carbon monoxide metabolizing *Oligotropha carboxidovorans*. Further examples from Bet v 1-related families include START lipid transfer proteins, phosphatidylinositol transfer proteins, and ring hydroxylases.

[00151] In some examples, the thebaine synthase enzyme may be a dirigent protein from plants in the *Ranunculales* order that biosynthesize thebaine, for example *P. somniferum*. In other examples, the enzyme may be any dirigent protein from plants.

[00152] In some examples, the thebaine synthase enzyme may be a chalcone isomerase protein from plants in the *Ranunculales* order that biosynthesize thebaine, for example *P. somniferum*. In other examples, the enzyme may be any chalcone isomerase protein from plants.

[00153] In some examples, the thebaine synthase enzyme may be a SalAT-like enzyme from plants in the *Ranunculales* order that biosynthesize thebaine, for example *P. somniferum*. In other examples, the enzyme may be any SalAT-like protein from plants.

[00154] In some examples, the enzyme with thebaine synthase activity may be from mammals or any other vertebrate or invertebrate that biosynthesizes endogenous morphine.

[00155] In some examples, any combination of the above enzymes together with additional accessory proteins may function to convert any tetracyclic precursor into thebaine. In some examples, these enzymes catalyze the reactions within a host cell, such as an engineered host, as described herein.

[00156] Examples of amino acid sequences for thebaine synthase activity are set forth in **Table 2**. An amino acid sequence for a thebaine synthase that is utilized in a tetracyclic precursor to thebaine may be 75% or more identical to a given amino acid sequence as listed in **Table 2**. For example, an amino acid sequence for such a thebaine synthase may comprise an amino acid sequence that is at least 75% or more, 80% or more, 81% or more, 82% or more, 83% or more, 84% or more, 85% or more, 86% or more, 87% or more, 88% or more, 89% or more, 90% or more, 91% or more, 92% or more, 93% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more, or 99% or more identical to an amino acid

sequence as provided herein. Additionally, in certain embodiments, an "identical" amino acid sequence contains at least 80%-99% identity at the amino acid level to the specific amino acid sequence. In some cases an "identical" amino acid sequence contains at least about 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94% and more in certain cases, at least 95%, 96%, 97%, 98% and 99% identity, at the amino acid level. In some cases, the amino acid sequence may be identical but the DNA sequence is altered such as to optimize codon usage for the host organism, for example.

[00157] An engineered host cell may be provided that produces a salutaridine reductase, salutaridinol 7-*O*-acetyltransferase, and thebaine synthase that converts a tetracyclic precursor into thebaine, wherein the thebaine synthase comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 30, 31, 32, 33, 34, 35, 36, and 37. In some cases, the thebaine synthase may form a fusion protein with other enzymes. The enzymes that are produced within the engineered host cell may be recovered and purified so as to form a biocatalyst. These one or more enzymes may also be used to catalyze the conversion of a tetracyclic promorphinan precursor to thebaine.

[00158] In other examples, the thebaine synthase comprises an amino acid sequence selected from the group consisting of SEQ ID NOs: 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, and 61.

[00159] In additional cases, the one or more enzymes that are recovered from the engineered host cell may be used in a process for converting a tetracyclic promorphinan precursor to a thebaine. The process may include contacting the tetracyclic promorphinan precursor with the recovered enzymes in an amount sufficient to convert said tetracyclic promorphinan precursor to thebaine. In examples, the tetracyclic promorphinan precursor may be contacted with a sufficient amount of the one or more enzymes such that at least 5% of said tetracyclic promorphinan precursor is converted to thebaine. In further examples, the tetracyclic promorphinan precursor may be contacted with a sufficient amount of the one or more enzymes such that at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 80%, at least 82%, at least 84%, at least 86%, at least 88%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, at least 99.7%, or 100% of said tetracyclic promorphinan precursor is converted to thebaine.

[00160] In some examples, process conditions are implemented to support the formation of thebaine in engineered host cells. In some cases, engineered host cells are grown at pH 3.3, and once high cell density is reached the pH is adjusted to pH 8.0 to support continued production of thebaine at higher pH. In some cases, the engineered host cells produce additional enzymes to convert sugar and other simple precursors, such as tyrosine, to thebaine. In some cases, the SalAT enzyme has been engineered to exhibit higher activity at pH 8.0 and is expressed from a late stage promoter.

[00161] In some examples, one or more of the enzymes converting a tetracyclic promorphinan precursor to a thebaine are localized to cellular compartments. In some examples, SalR, SalAT, and Bet v 1 may be

modified such that they encode targeting sequences that localize them to the endoplasmic reticulum membrane of the engineered host cell. In particular, in certain instances, the host cell may be engineered to increase production of salutaridinol or thebaine or products for which thebaine is a precursor from reticuline or its precursors by localizing Bet v 1 and/or SalR and/or SalAT to organelles in the yeast cell. Bet v 1 and/or SalR and/or SalAT may be localized to the yeast endoplasmic reticulum in order to decrease the spatial distance between Bet v 1 and/or SalR and/or SalAT and CYP2D2 or CYP2D6 or SalSyn or an engineered cytochrome P450 enzyme that catalyzes the conversion of reticuline to salutaridine. By increased production is meant both the production of some amount of the compound of interest where the control has no production of the compound of interest, as well as an increase of 10% or more, such as 50% or more, including 2-fold or more, e.g., 5-fold or more, such as 10-fold or more in situations where the control has some production of the compound of interest.

[00162] In other examples, SalAT and Bet v 1 may be co-localized in to a single protein fusion. In some examples, the fusion is created between SalAT and Bet v 1 by one of several methods, including, direct fusion, co-localization to a yeast organelle, or by enzyme co-localization tools such as leucine zippers, protein scaffolds that utilize adaptor domains, or RNA scaffolds that utilize aptamers. Co-localizing the thebaine synthesis enzyme may facilitate substrate channeling between the active sites of the enzymes and limit the diffusion of unstable intermediates such as salutaridinol-7-*O*-acetate.

[00163] In some examples, an engineered salutaridinol 7-*O*-acetyltransferase (SalAT) enzyme is used in converting a tetracyclic promorphinan precursor to a thebaine. In some examples, a SalAT enzyme is engineered to combine two functions: (1) the transfer of an acyl group from acetyl-CoA to the 7-OH of salutaridinol, and (2) the subsequent elimination of the acetyl group and closure of an oxide bridge between carbons C4 and C5 to form thebaine.

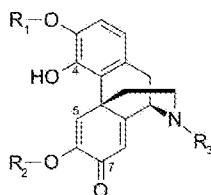
[00164] In some examples, an enzyme with salutaridinol 7-*O*-acetyltransferase activity is fused to a peptide with a Bet v 1 fold. In some examples, salutaridinol 7-*O*-acetyltransferase enzyme and the Bet v 1 fold protein may be fused in any order from N-terminus to C-terminus, C-terminus to N-terminus, N-terminus to N-terminus, or C-terminus to C-terminus. In some examples, the two protein sequences may be fused directly or fused through a peptide linker region.

[00165] In some examples, an enzyme with salutaridinol 7-*O*-acetyltransferase activity is fused to a peptide with a Bet v 1 fold by circular permutation. In some cases, the N- and C-termini of SalAT are fused and the Bet v 1 sequence is then inserted randomly within this sequence. In some cases, the resulting fusion protein library is screened for thebaine production. In other cases, a circular permutation SalAT library is first screened for activity in the absence of Bet v 1. In other cases, the N- and C-termini of SalAT are fused and the enzyme is digested and blunt end cloned. In other cases, this library of circularly permuted SalAT is screened for salutaridinol 7-*O*-acetyltransferase activity. In other cases, active variants from the circularly permuted SalAT library are then used to design protein fusions with a peptide with a Bet v 1 fold.

[00166] The one or more enzymes that may be used to convert a tetracyclic promorphinan precursor to a thebaine may contact the tetracyclic promorphinan precursor *in vitro*. Additionally, or alternatively, the one or more enzymes that may be used to convert a tetracyclic promorphinan precursor to thebaine may contact the tetracyclic promorphinan precursor *in vivo*. Additionally, the one or more enzymes that may be used to convert a tetracyclic promorphinan precursor to thebaine may be provided to a cell having the tetracyclic promorphinan precursor within, or may be produced within an engineered host cell.

[00167] In some examples, the methods provide for engineered host cells that produce an alkaloid product, wherein the conversion of a tetracyclic promorphinan precursor to a thebaine may comprise a key step in the production of an alkaloid product. In some examples, the alkaloid product is a thebaine. In still other embodiments, the alkaloid product is derived from a thebaine, including for example, downstream morphinan alkaloids. In another embodiment, a tetracyclic promorphinan precursor is an intermediate toward the product in of the engineered host cell. In still other embodiments, the alkaloid product is selected from the group consisting of morphinan, nor-opioid, or nal-opioid alkaloids.

[00168] In some examples, the substrate of the reduction reaction is a compound of Formula III:



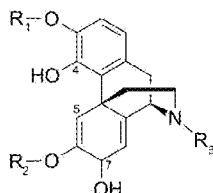
Formula III,

or a salt thereof, wherein:

R₁, R₂, and R₃ are independently selected from hydrogen and methyl.

[00169] In some other examples, R₁, R₂, and R₃ are methyl, and the reduction reaction is catalyzed by a salutaridine reductase.

[00170] In some examples, the substrate of the carbon chain transfer reaction is a compound of Formula IV:



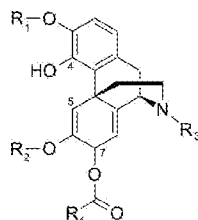
Formula IV,

or a salt thereof, wherein:

R₁, R₂, and R₃ are independently selected from hydrogen and methyl.

[00171] In some other examples, R₁, R₂, and R₃ are methyl, and the carbon chain transfer reaction is catalyzed by a salutaridinol 7-*O*-acetyltransferase.

[00172] In some examples, the substrate of thebaine synthase is a compound of Formula V:



Formula V,

or a salt thereof, wherein:

R₁, R₂, and R₃ are independently selected from hydrogen and methyl; and

R₄ is selected from methyl, ethyl, propyl, and other appropriate alkyl group.

[00173] In some other examples, R₁, R₂, R₃, and R₄ are methyl, and the ring closure reaction is catalyzed by a thebaine synthase. In some examples, the thebaine synthase is a Bet v 1 protein.

[00174] In some examples, the methods provide for engineered host cells that produce alkaloid products from salutaridine. The conversion of salutaridine to thebaine may comprise a key step in the production of diverse alkaloid products from a precursor. In some examples, the precursor is L-tyrosine or a sugar (e.g., glucose). The diverse alkaloid products can include, without limitation, morphinan, nor-opioid, or nal-opioid alkaloids.

[00175] Any suitable carbon source may be used as a precursor toward a pentacyclic morphinan alkaloid. Suitable precursors can include, without limitation, monosaccharides (e.g., glucose, fructose, galactose, xylose), oligosaccharides (e.g., lactose, sucrose, raffinose), polysaccharides (e.g., starch, cellulose), or a combination thereof. In some examples, unpurified mixtures from renewable feedstocks can be used (e.g., cornsteep liquor, sugar beet molasses, barley malt, biomass hydrolysate). In still other embodiments, the carbon precursor can be a one-carbon compound (e.g., methanol, carbon dioxide) or a two-carbon compound (e.g., ethanol). In yet other embodiments, other carbon-containing compounds can be utilized, for example, methylamine, glucosamine, and amino acids (e.g., L-tyrosine). In some examples, a 1-benzylisoquinoline alkaloid may be added directly to an engineered host cell of the invention, including, for example, norlaudanoline, laudanoline, norreticuline, and reticuline.

[00176] In some examples, the benzylisoquinoline alkaloid product, or a derivative thereof, is recovered. In some examples, the benzylisoquinoline alkaloid product is recovered from a cell culture. In some examples, the benzylisoquinoline alkaloid product is a morphinan, nor-opioid, or nal-opioid alkaloid.

Table 2. Example amino acid sequences of morphinan alkaloid generating enzymes.

Sequence Name	Description	Sequence	SEQ. ID. NO.
Bet v1			
	<i>P. bracteatum</i>	MAPRGVSGLVGKLSTELDVNCDAEKYYNMYKNGED VQKAVPHLCMDVKVISGDATRSGCIKEWNVNIDGKTI RSVEETTHNDETKTLRHRVFEGDMMKDYKKFDTIME VNPDPGNGC VVTRSIEYEKVNENSPTPFDY LQFGHQ AMEDMNKY	SEQ. ID. NO. 30
	<i>P. setigerum</i>	MLVGKLSTELEVDCDAEKYYNMYKHGEDVKKALCV DVKVISGDPTRSGCIKEWNVNIDGKTIRSVEETTHNDE TKTLRHRVFEGDMMKDFKKFDTIMVVNPKPDGNGC VVTRSIEYEKTNENSPTPFDY LQFGHQAIEDMNKYL	SEQ. ID. NO. 31
	<i>P. setigerum</i>	MLVGKLSTELEVDCDAEKYYNMYKHGEDKRQCVDV KVISGDPTRSGCIKEWNVNIDGKTIRSVEETTHNDETK TLRHRVFEGDMMKDFKKFDTIMVVNPKPDGNGC VV TRSIEYEKTNENSPTPFDY LQFGHQAIEDMNKY	SEQ. ID. NO. 32
	<i>P. setigerum</i>	MLVGKLSTELEVDCDAEKYYNMYKHGEDVKKAVPH LCVDVKIISGDP TSSGCIKEWNVNIDGKTIRSVEETTH DDETKTLRHRVFEGDVMKDFKKFDTIMVVNPKPDGN GCVVTRSIEYEKTNENSPTPFDY LQFGHQAIEDMNKY L	SEQ. ID. NO. 33
	<i>P. setigerum</i>	MVKIISGDP TSSGCIKEWNVNIDGKTIRSVEETTHDDE TKTLRHRVFEGDVMKDFKKFDTIMVVNPKPDGNGCV VTRSIEYEKTNENSPTPFDY LQFGHQAIEDMNKYL	SEQ. ID. NO. 34
	<i>P. somniferum</i>	MDSINSSIYFCAYFRELIHKLLMAPPGVSGLVGKLSTEL EVNCDAEKYYNMYKHGEDVQKAVPHLCVDVKVISG DPTRSGCIKEWNVNIDGKTIRSVEETTHNDETKTLRHR VFEGDVMKDFKKFDTIMVVNPKPDGNGC VVTRSIEY EKTNDNSPTPFDY LQFGHQAIEDMNKYLRDSE	SEQ. ID. NO. 35
	<i>P. somniferum</i>	MNFFIKDHL YICLVGKLSTELEVDCDAEKYYNMYKH GEDVKKAVPHLCVDVKIISGDP TSSGCIKEWNVNIDG KTIRSVEETTHDDETKTLRHRVFEGDVMKDFKKFDTI MVVNPKPDGNGC VVTRSIEYEKTNENSPTPFDY LQFG HQAIEDMNKYLRDSESN	SEQ. ID. NO. 36
	<i>P. somniferum</i>	MAPLGVSGLVGKLSTELEVDCDAEKYYNMYKHGED VKKAVPHLCVDVKIISGDP TSSGCIKEWNVNIDGKTIR	SEQ. ID. NO. 37

		SVEETTHDDETKTLRHRVFEGDVMKDFKKFDTIMVV NPKPDGNGC VVTRSIEYEKTNENSPTPFDYLQFGHQAI EDMNKYLRDSESN	
SalAT			
	<i>P. somniferum</i>	MMKVCVSSREKIKPSRPTPGHLKTHKLSFLDQVAARI YVPLLLYYAGNKENVDTDRCNIIKKSLAETLTKFYIL AGKIVNDEIERFVNCNDDGVDFCVTKVSNCQLFQVIK RPDIFDQVTLFLPFDPCDNEITASGDILLSVQVNVFED CRGMVIGLCINHKVADASSITTFVNYWATIARGLVLN VDDRQIQDPCFQVQSIFPQKEKGIGFKISSSSIDGTLVT KKFGFEASKLAELKERCKFAGATEDIRGGYKPNRVEA LSTFLWKCfididQAKTKAAAPARVYLASNANIRSR VPQLPTSSFGNMVAITDAIFTVNSNENNGINDPYYPKL VQKFRDAVKRVDGEYIEALQSTDLLLNNVTCLFKHIL NGQTLISFTSWCRFPFYDTDLLD	SEQ. ID. NO. 38
	<i>P. somniferum</i>	MKVQVISKELIKPSTPTPPRLRNFKLSLLDQLLPPFYVP IIIFYPANDDHESNNNDQCIKANILKKSLSETLTRFYPIA GRIRDKILVECNDEGVHYIEAKVNAVMSDFMSLDVIH QLHPSYITLDDLAEEAQLAVQVTMFDCGGIALSICSSH KIIDGCTSTTFLNSWAATARAPSNPEIVYPTDAAAIFF AQPSGVQVSTLESDDRLQGENVVTKRFLFSASKITAL RARIAESRSSNILSKYPSRSEAVSALVWKSFMETSRVK VTREHTFSAEASTKPIVRSIANFVVNLRLNPLPNVS FGNIIMDATAESLIIDNGENTLGFVETLDGLISQLRLGV TKMDDEYVRKLREDDVEFLKSLDEASHPSNGEGDGN GERV	SEQ. ID. NO. 39
	<i>P. setigerum</i>	MNDTMKIEVVSKEKPSYPTPNLKIHNLSNLDQLIP AFYMDHILYYP SLDSNDSSLGDDEEDKKMIFSASSRH RCDVVKKSLAETLTRYYP LAGRIKDEKSVECNDEGV DYIEARVVGITVSQVIQLASSDIEVMEPFLPYEPYGGT GSAFRRAGIHSNSKPLLKIQVNVFDCGGMVICLSGSH KVIDATSILNFVNDWAATARGGFDTHDDELKVA VVD KPCYIFSSMFPPTSFGNQEEKDTADQAQLVPDRIEIVT KRFVFKDSSIAKLKKKCIHVNTNNGSDHQVDKQEHN MQQMPSRIEALTSLIWMCFMDVDRRFRVKQIDDAVS PVNTVNEVSLPKQVQYVAGFAINLRTRTIQPLPTNSFG NMTDTAIAEVTNLNLTGSDHFNNEKGIRDQSQNYPELV	SEQ. ID. NO. 40

		SKIKDSIKLVDNKHIEAMKRNLAISCNNIKMHQMMKE STFDQNTRELLMFSSWCRFPIYEADFGWGKPSWASIT KLLYKNCVMFLDTSSGDGIEAWVSLKEEDMVEFERH EELVALAS	
	<i>P. somniferum</i>	MKVQVISKEIHKPSSPTPPHLRNFKLSLLDQILPPFYVPI VMFYFAGDDYVTNNNIHDQSSKSEFLKKSLSLSETLTRF YPIAGRIKDNILDCNNEGVDYIEAKVNGIMSDFMSVD VVHQLHPSHIMLDDVAKEAQLAVQVNLFDCCGIAISI SMCHKIVDACTAITFINGWAATARAAPKQEIVCPTFDS AAIFPALPPGVQVSSLESDDSVQGVNVVTKMFAFTAP KIASLRARIAELRSSDGLSKYPTRTEALSALVWKSFIR TSRVKAARKYSLSPASTKPKVSVANYAVNLRTRLNP PLPQVSFGNILMDATAESTTTIDDDDSHEFADTLAGLI GQLRLGVSRINGDYIRKLQEGDLAFLKSLDEASHDSN GEKVQICWISSLCRFPFYEADFGWGKPSWVALNTNAE YKNSLFLMDTKCGTGIEAWVSLEEDDMAIFEEDQDLL QCVKSIN	SEQ. ID. NO. 41
	<i>P. setigerum</i>	MENMKVEVVVKQTIKSTQTPLHSKTFNLSFLDQHLG PPIYIPFTLYYESGDVNNKNNHCDGYKNNLEEACEHR VSVIKQSLSETLARYYPLAGRMKEDNLAVECNDEGV EYFETRVS DVRLSQVIKRSFNHNSVLRKFLPPCISSCD NSMSIPFDYGFKSKTLLAIQVNIFECGGIVIGMCMAHR LADASTMFTFITDWAATARGAIEDIKGPSFDFS YTLFP QKDVINNFKPFDPMLTREEDLVTKYFVFPASKIVELKR RNVNNIVCQDTSQQNTSPCTRVEAVTSFMWKRYMDS VRAKNQTQATSVEKYGALYTVNLRSRITPPLPANSFG NIYFTTIALSTPSDENDIDDGLRKDVSSPNDLNLVGKV RDAIKKIDDKYTRKLQSSSEDEL VNDVKPLTSGEAIFLG FSSWCRFPIYEADFGWGKPTWVSIGTMALRNTVFLM DTKSGDGIEAFVNMAKEDMDNFEVKLLADQ	SEQ. ID. NO. 42
	<i>P. setigerum</i>	MENMKVEVVLEQTIKSTQTPLHSKTFNLSFLDQHLG PPIYIPFTLYYESGDVNNKNNHCDGYKNNLEEVC EHR VSVIKQSLSETLARYYPLAGRMKEDNLAVECNDEGV EYFETRVS DVRLSQVIKRSFNHNSVLRKFLPPCISSCD NSMSIPFDYGFKSKTLLAIQVNIFECGGIVIGMCMAHR LADASTMFTFITDWAATARGAIEDIKGPSFDFS YTLFP QKDVINNFKPFDPMLTREEDLVTKYFVFPASKIVELKR	SEQ. ID. NO. 43

		RNVNNIVCQDTSQQNTSPCTRVEAVTSFMWKRYMDS VRAKNQTQATSVEKYGALYTVNLRSRITPPLPANSFG NIYFTIALSTPSDENDIDDGLRKDVSSPNDLNLVGKV RDAIKKIDDKYTRKLQSSSEDELVNDVKPLTSGEAIFLG FSSWCRFPIYEADFGWGKPTWVSIGTMALRNTVFLM DTKSGDGIEAFVNMAKEDMDNFEVKLLADQLLHVHP TV	
	<i>P. setigerum</i>	MSSTVEVISKQTIKPSTPTPIQRKNHSLSLIDQHFAPIYI PIVLFPAAAVNDTGNVQHGDNTCVLKRSLSETLVHF YPLAGRMKDNIVVDCNDQGVEFTEVKVSGTMCDFL MKPDEQLSGLLPSEAVCMNFVREAQVMIQVNTFDCG SKAISLCVSHKIADASTITTFSRCWAETTIAVSKSTSAV TPIVSSKFHPTFDAASLFPPIKQLISPSGVTPALPELIPSE ESKFGKIISKRFIFSATTINSVREKLSALMADKLKYRR LTRVEVVSALIWNSFDKLATTGSVAVMVKHAVNLRK RIDPPLPDVSFGNILEFTKAVVGEAAANTTTQGTVGSS SKLLEELSEFAGQLREPVSKMNKGDHDFDMENTDYE ERDLWMSSWCNYGLYDIDFGCGKPVWVTTVATMYP YSDGFFMNDTRCGQGIEVWGNLVEEDMANFQLNLSE LLDRI	SEQ. ID. NO. 44
	<i>P. somniferum</i>	MMKVCVSSREKIKPSRPTPGHLKTHKLSFLDQVAARI YVPLLLYYAGNKENVDTDRCNIIKKSLETTLTKFYIL AGKIVNDEIERFVNCNDDGVDFCVTKVSNCQLFQVIK RPDIFDQVTLFLPFDPCDNEITASGDILLSVQVNVFED CRGMVIGLCINHKVADASSITTFVNYWATIARGLVLN VDDRQIQDPCFQVQSIFPQKEKGIGFKISSSSIDGTLVT KKFGFEASKLAELKERCKFTTEPEDGYKPTRVEALSA FLWKCFIDIDQAKLKGVARTKVYLATNAVNMRSMV PQLPTSSFGNIISITDAVFSINDDSTGINDPYYPKLVRK FRDAIKKIDRDYIEALRSTDLLLNNMMKLIEHVLSGHT LSIYFSSWCRFPLYETDFGWGKPIWVSTCTIPQKNVIV LMDSNSSADGIEAYVTLAKEDMGELEHHEELLALIS	SEQ. ID. NO. 45
Dirigent proteins			
	<i>P. somniferum</i>	MGAMKFFSFLAVAMVLSLAHIQAQQGNWGDETVPY TMGPEKITKLRFYFHDIIVTGNNPTAVQIAQATGTNSSS TLFGALFMIDDPLTEGPDPSRLVGRAQGFYGSAGQN	SEQ. ID. NO. 46

		EAALILGMSLVFTGNEKFNGSTISVLSRNPVTHTEREF AIVGGTGYFQFARGFISAKTYSLVGPNAVVEYNCTIV HPSSVSESGKSNSSPGKSDSNSGSQISLGSNLVFSVIA YVTIILSL	
	<i>P. setigerum</i>	MVLSMSHSQAQEGNWGDESVPYTMGPEKMTKLRFY FHDITGNSPTAVQIAQATGTNTSATMFGALMMIDDPL TEGPDNSRLVGRAQGFYGSAGQNELALILGMSLVFT GNEKFNGSTISVLSRNPVMHTEREF AIVGGTGYFQFA RGFISAKTYSLVGPNAVVEYNCTIVHPSSVSESGKSDS SSGKSDSSSGSQISLGTNLVFLSVIAFVTIIVSPQHFSW	SEQ. ID. NO. 47
Chalcone isomerase			
	<i>P. somniferum</i>	MTKTVLVDDIPFPQNITTVTTEKQLPLLGGGITDMEIH FLQIKFTAIGTAIGVYLEPEIASHLQQWKGKTGAELSQ DDEFFAAVVSASVEKYVRVVVIKEIKGSQYMLQLES WVRDELAADKYEDEEEESLDKVIEFFQSKYLKQLSF IPSHFSATTPAVAEIGLEIEGQKDLKIKVENGNVIEMIQ KWYLGGRGVSPSTTQSLATSL	SEQ. ID. NO. 48
	<i>P. somniferum</i>	MPFLKAIEIEGCKFRPFVTPPGSTQILFLAGSGVKEEFG DSKSMKYSSCAIYLQPTCILYLAKAWAQKSVDITQS LNFFMDIATGPFKEYCRITMLETAKGEDYAAMITKNC EEMLTNSKRYSETAKAALTKFSEAFNGRTLASGSSIH VTVSTNSVTLAFTEDGSTPKQGDVTLDCKEVGEAFL MSTISLHTTIRESMGSRISGLYK	SEQ. ID. NO. 49
	<i>P. setigerum</i>	MAPMAQLSEIQVEQFVFPPTMTPPSSTESLFLGGAGVR GLQIQDRFIKFTAIGVYLAEEAIPSLSPKWKSPEELT DDVEFFMDIVTGPFEKFKITMILPLTGDQYAEKV TEN CIQYLKSKDMYTDAAEAKAVERFIEIFKNEMFPASSIL FTISPAGSLTVGF*	SEQ. ID. NO. 50
	<i>P. rhoeas</i>	MVYLEPEIATHLKQWKGKTGAELSQDDDDFFSAVVSA PVEKYVRVVVIKEIKGSQYMLQLESWVRDELAADK YEDEEEESLDKVIEFFQSKYLKQHSVIITFHFSATTPAV AEIGLEIEGQKDLKIKVENGNVEMIQKWYLGGRGV SPSTTQSLATSL	SEQ. ID. NO. 51
	<i>P. bracteatum</i>	MTKMVLVDDIPFPQNITTATTAKQLPLLGGGITDMEIH FLQIKFTAIGVYLEPEIASHLKQWKGKTGAELSQDDEF FSAIVSAPVEKYVRVVVIKEIKGSQYMLQLESWVRDE	SEQ. ID. NO. 52

		LAAADKYEDEEEESLEKVIEFFQSKYLKQHSVIPFHFS ATTPAVAEIGLEIEGHKDLKMKVENGNVEMIQKWY LAGTRGVSPSTTQSLATSL	
	<i>P. bracteatum</i>	MAPMAQLSEIQVEQFVFPPTMTPPSSTESLFLGGAGVR GLQIQDRFIKFTAIGVYLAEEAIPSLSPKWWSKTPEELT NDVEFFMDIVTGPFEKFKITMILPLTGDQYAEKV TEN CVEYLKSKDLYTDAEAKAVERFIEIFKNEMFPPASSIL FTISPTGSLTVGFSKDTSIPEARNAVIENKALSESILESII GKNGVSPAACKQSLAERISELLK	SEQ. ID. NO. 53
Other			
	<i>P. ginseng</i>	MGLTGKCLICQTGIKSDGDVFHELFGTRPHHVPNITPAN IQGCDLHEGEFGKVGSVVIWNYSIDGNAMIAKEEIVAI DEEDKSVTFKVVEGHLFEEFKSIVFSVHVDTKGEDNL VTWSIDYEKLNESVKDPTSYLDLFLSVTRDIEAHHLPK	SEQ. ID. NO. 54
	<i>A. hypogaea</i>	MGVFTFEDEITSTVPPAKLYNAMKDADSITPKIIDVVK SVEIVEGNNGPGTIKKLTIVEDGETKFILHKVESIDEAN YAYNYSVVGVALPPTAEKITFETKLVEGPNGGSIGK LTLKYHTKGDAKPDEEELKKGKAKGEGLFRAIEGYV LANPTQY	SEQ. ID. NO. 55
	<i>H. perforatum</i>	MGIDPFTMAAYTIVKEEESPIAPHRLFKALVLERHQVL VKAQPHVFKSGEIIEGDGGVGT VTKITFVDGHPLTYM LHKFDEIDAANFYCKYTLFEGDVLRDNIEKVVEVVKL EAVGGGSKGKITVTYHPKPGCTVNEEEVKIGEEKKAYE FYKQVEEYLAANPEVFA	SEQ. ID. NO. 56
	<i>L. luteus</i>	MGVFTFQDEYTSTIAPAKLYKALVTDADIIPKAVETI QSVEIVEGNNGPGTIKKLTFIEGGESKYVLHKIEAIDEA NLGYNYSIVGGVGLPDTIEKISFETKLVEGANGGSIGK VTIKIETKGDAQPNEEEGKAAKARGDAFFKAIESYLSA HPDYN	SEQ. ID. NO. 57
	<i>Strawberry (Fragaria x ananassa)</i>	MAGVFTYETEFTSVIPPPRLFKAFILDADNLIPKIAPQA VKCAEIIEGDGGVGTIKKITFGEGSQFGSVTHKIDGIDK ENFVYSYSLIEGDALSDKIEKISYETKL VSSSDGGSIIKS TSNYHTKGDVEIKEEHVKAGKEKFSHLFKLVEGYLLA NPNEYC	SEQ. ID. NO. 58
	<i>A. deliciosa</i>	MDLSGKMVKQVEILSDGIVFYEIFRYRLYLISEMSPVN IQGVLDLLEGNWGT VGSVIFFKYTIDGKEKTAKDIVEAI	SEQ. ID. NO. 59

		DEETKSVTFKIVEGDLMEYKTFIIIVQVDTKGEHNSV TWTFFHYEKLKEDVEEPNTLMNFCIEITKDIETYHLK	
	<i>T. flavum</i>	MGIINQVSTVTKVIHHELEVAASADDIWTVYSWPGLA KHLPDLLPGAFAFEKLEIIGDGGVGTILDMTFVPGEFPHE YKEKFILVDNEHRLKKVQMIEGGYLDLGVTYYMDTI HVVPTGKDSCVIKSSTEYHVKPEFVKIVEPLITTGPLA AMADAISKLVLEHKS	SEQ. ID. NO. 60
	<i>V. radiata</i>	MVKEFNTQTELSVRLEALWAVLSKDFITVVPKVLPHI VKDVQLIEGDGGVGTILIFNFLPEVSPSYQREEITEFDE SSHEIGLQVIEGGYLSQGLSYYKTTFKLSEIEEDKTLV NVKISYDHDSIEEKVTPKTSQSTLMYLRRLERYLSN GSA	SEQ. ID. NO. 61

BisBIA Generating Modifications

[00177] Some methods, processes, and systems provided herein describe the increased production of bisbenzylisoquinoline alkaloids (bisBIAs) by utilizing two separate epimerase enzymes derived from a parent epimerase enzyme when compared to production of the bisBIAs by utilizing a corresponding fused enzyme. In examples, a corresponding fused enzyme comprises a fused epimerase having corresponding oxidase and reductase regions to the two separate epimerase enzymes. In examples, the two separate epimerase enzymes may comprise an oxidase and a reductase. BisBIAs are dimeric molecules that may be formed by coupling reactions between two BIA monomers. In examples, bisBIAs may be formed by carbon-oxygen coupling reactions. In other examples, bisBIAs may be formed by carbon-carbon coupling reactions. In some examples, the bisBIA dimeric molecule is a homodimer, comprising two identical BIA monomers. In examples, an engineered host cell may produce one BIA monomer. In these examples, the BIA monomers may form homodimers when contacted with one or more coupling enzymes. In other examples, the bisBIA dimeric molecule is a heterodimer, comprising two different BIA monomers. For example, a bisBIA may be a heterodimer that comprises BIA monomers that are enantiomers of each other. In some examples, an engineered host cell may produce two or more BIA monomers. In these examples, the BIA monomers may form homodimers and heterodimers when contacted with one or more coupling enzymes.

[00178] Some of these methods, processes, and systems that describe the production of bisBIAs may comprise an engineered host cell. In some examples, the engineered host cell may be engineered to produce BIA monomers which, in turn, may be used as building block molecules for forming bisBIAs. Examples of BIA monomers that may be used to form bisBIAs include coclaurine, *N*-methylocclaurine, laudanine, norcoclaurine, norlaudanoline, 6-*O*-methyl-norlaudanoline, 3'-hydroxy-*N*-methylocclaurine, 3'-hydroxycoclaurine, reticuline, norreticuline, norlaudanine, laudanoline, and

papaverine. In particular, engineered host cells may synthesize BIA monomers from norcoclaurine or norlaudanosoline by expression of heterologous enzymes including *O*-methyltransferases, *N*-methyltransferases, and 3'-hydroxylases. Examples of *O*-methyltransferases may include norcoclaurine 6-*O*-methyltransferase (6OMT). Further examples of *O*-methyltransferases may include catechol *O*-methyltransferase (COMT). Further examples of *N*-methyltransferases may include coclaurine *N*-methyltransferase (CNMT). Examples of 3'-hydroxylases may include *N*-methylcoclaurine 3'-hydroxylase (CYP80B1).

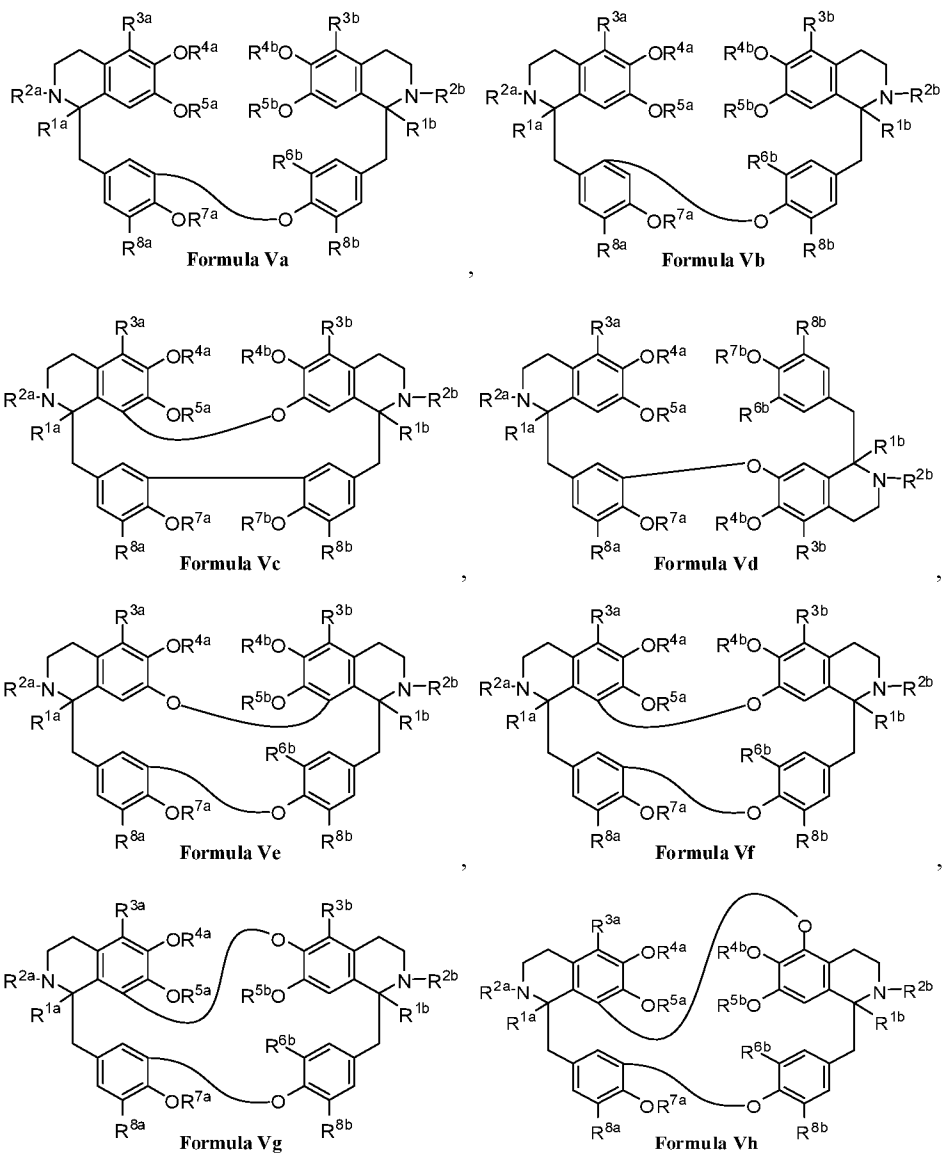
[00179] The engineered host cells may produce either (*S*) or (*R*) enantiomers of any given BIA monomer. Additionally or alternatively, the engineered host cells may produce a mixture of both enantiomers. The ratio of (*S*) and (*R*) enantiomers may be determined by the substrate and product specificities of the one or more enzymes that synthesize the BIA monomers. Alternatively, the amount of each enantiomer present may be modified by the expression and engagement of the two separate oxidase and reductase enzymes of the engineered epimerase that performs the epimerization of one stereoisomer into another. In some cases, the amount of each enantiomer present may be modified by the expression and engagement of the engineered fused epimerase that performs the epimerization of one stereoisomer into another.

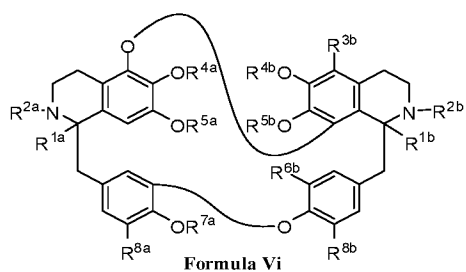
[00180] These BIA monomers may be fused into a dimeric bisBIA scaffold. In particular, the BIA monomers may be fused into a dimeric bisBIA scaffold utilizing one or more enzymes that are produced by the engineered host cell. Additionally or alternatively, the BIA monomers may be fused into a dimeric bisBIA scaffold utilizing one or more enzymes that are provided to the BIA monomers from a source that is external to the engineered host cell. The one or more enzymes may be used to form carbon-oxygen and/or carbon-carbon coupling reactions to fuse two BIA monomers at one, two, or three positions. In some examples, two BIA monomers may be linked by an ether bridge. In some examples, a direct carbon-carbon bond may be used to connect the two BIA monomers. In some examples, a bisBIA that is formed by fusing two BIA monomers may comprise one diphenyl ether linkage. In some examples, two BIA monomers may be fused to form a bisBIA that comprises two diphenyl ether linkages. In some examples, a bisBIA that is formed from two BIA monomers may comprise three diphenyl ether linkages. In some examples, the bisBIA may comprise one diphenyl ether linkage and one benzyl phenyl ether linkage. In some cases, the bisBIA may comprise one benzyl phenyl ether linkage and two diphenyl ether linkages.

[00181] In examples, the BIA monomers may be contacted with a sufficient amount of the one or more enzymes that may be used to form coupling reactions to fuse two BIA monomers such that at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 82%, at least 84%, at least 86%, at least 88%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, at least 99.7%, or 100% of said BIA monomers are converted to bisBIAs. The one or more enzymes that may be used to dimerize the BIA monomers into bisBIAs may contact the BIA monomers *in vitro*. Additionally, or

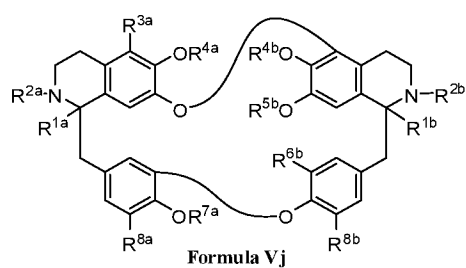
alternatively, the one or more enzymes that may be used to dimerize the BIA monomers into bisBIAs may contact the BIA monomers *in vivo*. Additionally, the one or more bisBIA dimerizing enzyme may be expressed in a host cell that produces BIA monomers. Alternatively, the BIA monomers may be provided to the engineered host cell that expresses the bisBIA dimerizing enzyme. Alternatively, the one or more bisBIA dimerizing enzymes may be provided to a cell having BIA monomers within.

[00182] In some examples, the bisbenzylisoquinoline alkaloid is a compound of any one of Formulas Va-Vu:

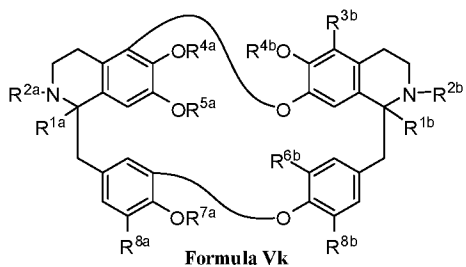




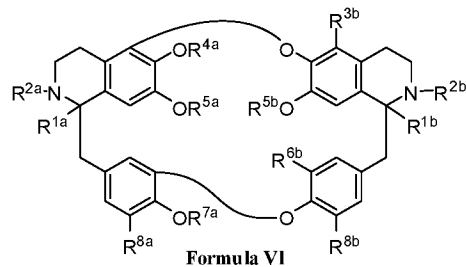
,



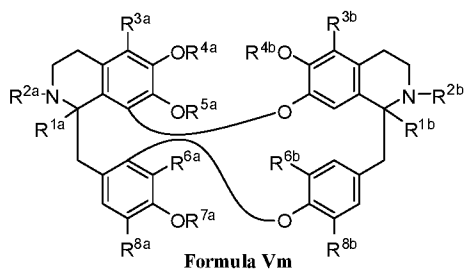
,



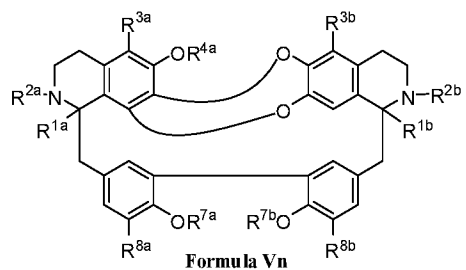
,



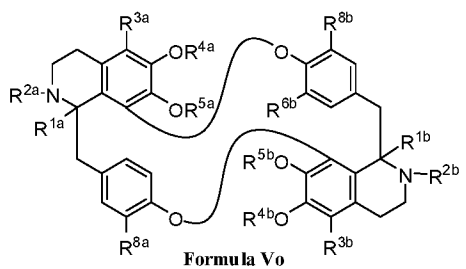
,



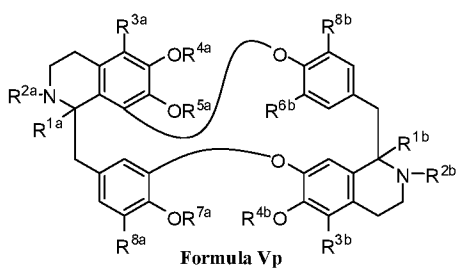
,



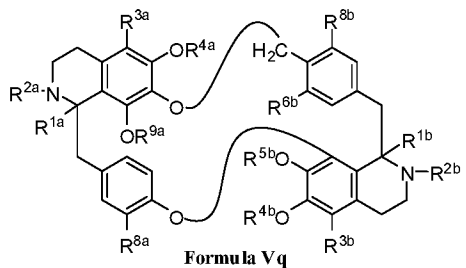
,



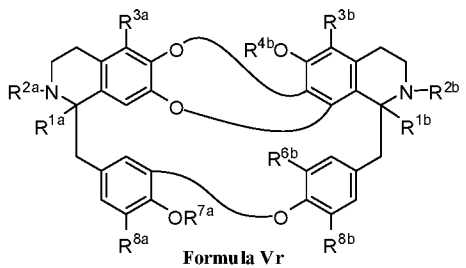
,



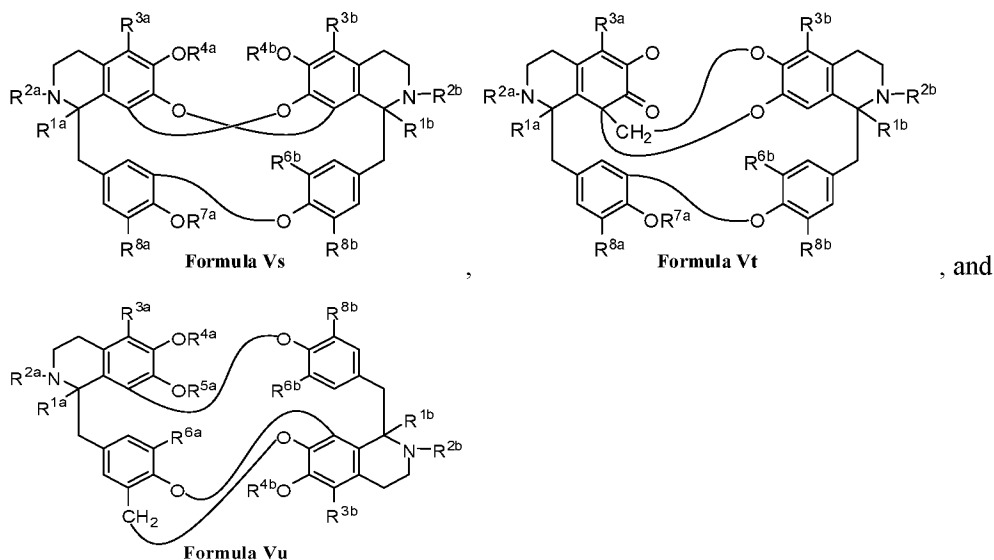
,



,



,



or a salt thereof, wherein:

R^{1a} , R^{1b} , R^{2a} , and R^{2b} are independently selected from hydrogen and C_1 - C_4 alkyl;

R^{3a} , R^{3b} , R^{6a} , R^{6b} , R^{8a} , and R^{8b} are independently selected from hydrogen, hydroxy, fluoro, chloro, bromo, carboxaldehyde, C_1 - C_4 acyl, C_1 - C_4 alkyl, and C_1 - C_4 alkoxy;

R^{4a} and R^{5a} are independently selected from hydrogen and C_1 - C_4 alkyl, or R^{4a} and R^{5a} together form a methylene bridge;

R^{4b} and R^{5b} are independently selected from hydrogen and C_1 - C_4 alkyl, or R^{4b} and R^{5b} together form a methylene bridge; and

R^{7a} , R^{7b} , and R^{9a} are independently selected from hydrogen and C_1 - C_4 alkyl.

[00183] In some examples, R^{1a} and R^{1b} are each hydrogen; R^{2a} and R^{2b} are each methyl; R^{3a} and R^{3b} are each hydrogen; R^{4a} and R^{5a} are independently hydrogen or methyl; R^{4b} and R^{5b} are independently hydrogen or methyl, or R^{4b} and R^{5b} together form a methylene bridge; R^{6a} , R^{6b} , R^{8a} , and R^{8b} are each hydrogen; and R^{7a} , R^{7b} , and R^{9a} are independently hydrogen or methyl.

[00184] As illustrated above, the bisBIA compounds of Formulas Va, Vb, and Vd are formed by fusing two BIA monomers using a carbon-oxygen coupling reaction. Additionally, the bisBIA compounds of Formulas Vc, Vf, and Vh are formed by fusing two BIA monomers using both a carbon-oxygen coupling reaction and a carbon-carbon coupling reaction. Further, the bisBIA compounds of Formulas Ve, Vg, Vi, Vj, Vk, Vl, Vm, Vo, Vp, and Vq are formed by fusing two BIA monomers using two carbon-oxygen coupling reactions. The bisBIA compound of Formula Vn is formed by fusing two BIA monomers using two carbon-oxygen coupling reactions and a carbon-carbon coupling reaction. Additionally, the bisBIA compound of Formula Vr is formed by fusing two BIA monomers using three carbon-oxygen coupling reactions.

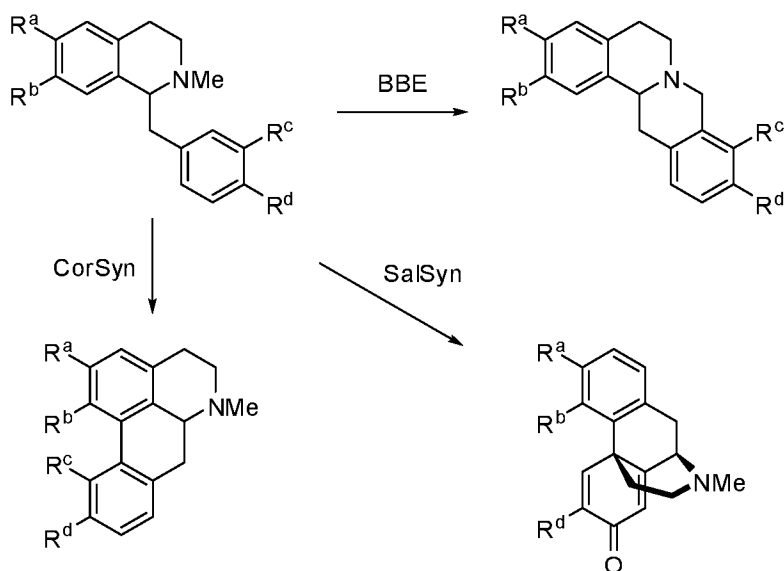
[00185] The one or more enzymes that may be used to form the coupling reactions may include known cytochrome P450s such as *Berberis stolonifera* CYP80A1 or similar cytochrome P450 enzymes from

other plants that naturally synthesize these compounds. Alternatively, the coupling reaction may be performed by an enzyme that is not a cytochrome P450. The one or more enzymes that may be used to form the coupling reactions may be engineered to accept non-native substrates. Accordingly, the one or more enzymes that may be used to form the coupling reactions may be used to generate non-natural bisBIA molecules. In examples, the one or more enzymes may fuse a natural BIA monomer with a non-natural BIA monomer to produce a non-natural bisBIA molecule. In other examples, the one or more enzymes may fuse two non-natural BIA monomers to produce a non-natural bisBIA molecule. Enzyme engineered strategies may be used to identify one or more enzymes that may be used to form the coupling reactions that fuse BIA monomers to produce bisBIAs. In examples, enzyme engineering strategies may include site directed mutagenesis, random mutagenesis and screening, DNA shuffling, and screening.

[00186] Once bisBIAs are formed, the bisBIAs may be further derivatized or modified. The bisBIAs may be derivatized or modified utilizing one or more enzymes that are produced by the engineered host cell. In particular, the bisBIAs may be derivatized or modified by contacting the bisBIAs with one or more enzymes that are produced by the engineered host cell. Additionally or alternatively, the bisBIAs may be derivatized or modified by contacting the bisBIAs with one or more enzymes that are provided to the bisBIAs from a source that is external to the engineered host cell. The one or more enzymes that may be used to derivatize or modify the bisBIAs may be used to perform tailoring reactions. Examples of tailoring reactions include oxidation, reduction, *O*-methylation, *N*-methylation, *O*-demethylation, acetylation, methylenedioxybridge formation, and *O,O*-demethylenation. A bisBIA may be derivatized or modified using one or more tailoring reactions.

[00187] Examples of tailoring reactions are provided in **Table 9**. In some examples, tailoring enzymes may be used to catalyze carbon-carbon coupling reactions performed on a bisBIA, or a derivative thereof. Examples of tailoring enzymes that may be used to catalyze carbon-carbon coupling reactions include a Berberine bridge enzyme (BBE) from *Papaver somniferum*, *Eschscholzia californica*, *Coptis japonica*, *Berberis stolonifer*, *Thalictrum flavum*, or another species; Salutaridine synthase (SalSyn) from *Papaver somniferum* or another species; and Corytuberine synthase (CorSyn) from *Coptis japonica* or another species. Non-limiting examples of reactions that can be catalyzed by tailoring enzymes are shown in **Scheme 3**, wherein R^a , R^b , R^c , and R^d are independently selected from hydrogen, hydroxy, fluoro, chloro, bromo, carboxaldehyde, C_1 - C_4 acyl, C_1 - C_4 alkyl, and C_1 - C_4 alkoxy. In some examples, R^a , R^b , and the carbon atoms to which they are attached optionally form a carbocycle or heterocycle. In some examples, R^c , R^d , and the carbon atoms to which they are attached optionally form a carbocycle or heterocycle.

Scheme 3



[00188] In some examples, tailoring enzymes may be used to catalyze oxidation reactions performed on a bisBIA, or a derivative thereof. Examples of tailoring enzymes that may be used to catalyze oxidation reactions include a Tetrahydropprotoberberine oxidase (STOX) from *Coptis japonica*, *Argemone mexicana*, *Berberis wilsonae*, or another species; Dihydrobenzophenanthridine oxidase (DBOX) from *Papaver somniferum* or another species; Methylstylophine hydroxylase (MSH) from *Papaver somniferum* or another species; and Protopine 6-hydroxylase (P6H) from *Papaver somniferum*, *Eschscholzia californica*, or another species.

[00189] Tailoring enzymes may also be used to catalyze methylenedioxy bridge formation reactions performed on a bisBIA, or a derivative thereof. Examples of tailoring enzymes that may be used to catalyze methylenedioxy bridge formation reactions include a Stylophine synthase (StySyn) from *Papaver somniferum*, *Eschscholzia californica*, *Argemone mexicana*, or another species; Cheilanthifoline synthase (CheSyn) from *Papaver somniferum*, *Eschscholzia californica*, *Argemone mexicana*, or another species; and Canadine synthase (CAS) from *Thalictrum flavum*, *Coptis chinensis*, or another species.

[00190] In other examples, tailoring enzymes may be used to catalyze *O*-methylation reactions performed on a bisBIA, or a derivative thereof. Examples of tailoring enzymes that may be used to catalyze *O*-methylation reactions include a Norcoclaurine 6-*O*-methyltransferase (6OMT) from *Papaver somniferum*, *Thalictrum flavum*, *Coptis japonica*, *Papaver bracteatum*, or another species; 3'-hydroxy-*N*-methylcoclaurine 4'-*O*-methyltransferase (4'OMT) from *Papaver somniferum*, *Thalictrum flavum*, *Coptis japonica*, *Coptis chinensis*, or another species; Reticuline 7-*O*-methyltransferase (7OMT) from *Papaver somniferum*, *Eschscholzia californica*, or another species; and Scoulerine 9-*O*-methyltransferase (9OMT) from *Papaver somniferum*, *Thalictrum flavum*, *Coptis japonica*, *Coptis chinensis*, or another species.

[00191] Additionally, tailoring enzymes may be used to catalyze *N*-methylation reactions performed on a bisBIA, or a derivative thereof. Examples of tailoring enzymes that may be used to catalyze *N*-methylation reactions include Coclaurine *N*-methyltransferase (CNMT) from *Papaver somniferum*,

Thalictrum flavum, *Coptis japonica*, or another species; Tetrahydroprotoberberine *N*-methyltransferase (TNMT) from *Papaver somniferum*, *Eschscholzia californica*, *Papaver bracteatum*, or another species.

[00192] Further, tailoring enzymes may be used to catalyze *O*-demethylation reactions performed on a bisBIA, or a derivative thereof. Examples of tailoring enzymes that may be used to catalyze *O*-demethylation reactions include Thebaine demethylase (T6ODM) from *Papaver somniferum* or another species; and Codeine demethylase (CODM) from *Papaver somniferum*, or another species.

[00193] Tailoring enzymes may also be used to catalyze reduction reactions performed on a bisBIA, or a derivative thereof. Examples of tailoring enzymes that may be used to catalyze reduction reactions include Salutaridine reductase (SalR) from *Papaver somniferum*, *Papaver bracteatum*, or another species; Codeinone reductase (COR) from *Papaver somniferum* or another species; and Sanguinarine reductase (SanR) from *Eschscholzia californica* or another species. In other examples, tailoring enzymes may be used to catalyze acetylation reactions performed on a bisBIA, or a derivative thereof. An example of a tailoring enzyme that may be used to catalyze acetylation reactions includes Salutaridine acetyltransferase (SalAT) from *Papaver somniferum* or another species.

O-Demethylation Modifications

[00194] Some methods, processes, and systems provided herein describe the conversion of a first benzyloquinoline alkaloid to a second benzyloquinoline alkaloid by the removal of an *O*-linked methyl group. Some of these methods, processes, and systems may comprise an engineered host cell. In some examples, the conversion of a first benzyloquinoline alkaloid to a second benzyloquinoline alkaloid is a key step in the conversion of a substrate to a nor-opioids or nal-opioids. In some examples, the conversion of a first alkaloid to a second alkaloid comprises a demethylase reaction.

[00195] FIG. 6 illustrates an enzyme having opioid 3-*O*-demethylase (ODM) activity, in accordance with embodiments of the invention. Specifically, the enzyme may act on any morphinan alkaloid structure to remove the methyl group from the oxygen bound to carbon 3.

[00196] Examples of amino acid sequences of ODM enzymes are set forth in Table 4. An amino acid sequence for an ODM that is utilized in converting a first alkaloid to a second alkaloid may be 75% or more identical to a given amino acid sequence as listed in Table 4. For example, an amino acid sequence for such an epimerase may comprise an amino acid sequence that is at least 75% or more, 80% or more, 81% or more, 82% or more, 83% or more, 84% or more, 85% or more, 86% or more, 87% or more, 88% or more, 89% or more, 90% or more, 91% or more, 92% or more, 93% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more, or 99% or more identical to an amino acid sequence as provided herein. Additionally, in certain embodiments, an "identical" amino acid sequence contains at least 80%-99% identity at the amino acid level to the specific amino acid sequence. In some cases an "identical" amino acid sequence contains at least about 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94% and more in certain cases, at least 95%, 96%, 97%, 98% and 99% identity, at the amino acid

level. In some cases, the amino acid sequence may be identical but the DNA sequence is altered such as to optimize codon usage for the host organism, for example.

[00197] An engineered host cell may be provided that produces an ODM that converts a first alkaloid to a second alkaloid, wherein the ODM comprises a given amino acid sequence as listed in **Table 4**. An engineered host cell may be provided that produces one or more ODM enzymes. The ODM that is produced within the engineered host cell may be recovered and purified so as to form a biocatalyst. The process may include contacting the first alkaloid with an ODM in an amount sufficient to convert said first alkaloid to a second alkaloid. In examples, the first alkaloid may be contacted with a sufficient amount of the one or more enzymes such that at least 5% of said first alkaloid is converted to a second alkaloid. In further examples, the first alkaloid may be contacted with a sufficient amount of the one or more enzymes such that at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 82%, at least 84%, at least 86%, at least 88%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, at least 99.7%, or 100% of said first alkaloid is converted to a second alkaloid.

[00198] The one or more enzymes that may be used to convert a first alkaloid to a second alkaloid may contact the first alkaloid in vitro. Additionally, or alternatively, the one or more enzymes that may be used to convert a first alkaloid to a second alkaloid may contact the first alkaloid in vivo. In some examples, the one or more enzymes that may be used to convert a first alkaloid to a second alkaloid may be provided to a cell having the first alkaloid within. In some examples, the one or more enzymes that may be used to convert a first alkaloid to a second alkaloid may be produced within an engineered host cell.

[00199] In some examples, the methods provide for engineered host cells that produce an alkaloid product, wherein the *O*-demethylation of a substrate to a product may comprise a key step in the production of an alkaloid product. In some examples, the alkaloid produced is a nor-opioid or a nal-opioid. In still other embodiments, the alkaloid produced is derived from a nor-opioid or a nal-opioid. In another embodiment, a first alkaloid is an intermediate toward the product of the engineered host cell. In still other embodiments, the alkaloid product is selected from the group consisting of morphine, oxymorphone, oripavine, hydromorphone, dihydromorphone, 14-hydroxymorphine, morphinone, and 14-hydroxymorphinone.

[00200] In some examples, the substrate alkaloid is an opioid selected from the group consisting of codeine, oxycodone, thebaine, hydrocodone, dihydrocodeine, 14-hydroxycodone, codeinone, and 14-hydroxycodeinone.

N-Demethylation Modifications

[00201] Some methods, processes, and systems provided herein describe the conversion of a first alkaloid to a second alkaloid by the removal of an *N*-linked methyl group. Some of these methods, processes, and systems may comprise an engineered host cell. In some examples, the conversion of a first alkaloid to a second alkaloid is a key step in the conversion of a substrate to a nor-opioids or nal-opioids. In some examples, the conversion of a first alkaloid to a second alkaloid comprises a demethylase reaction.

[00202] FIG. 7 illustrates an enzyme having opioid *N*-demethylase activity, in accordance with embodiments of the invention. Specifically, the enzyme may act on any morphinan alkaloid structure to remove the methyl group from the nitrogen.

[00203] Examples of an amino acid sequence of an *N*-demethylase (NDM) enzyme that may be used to perform the conversion a first alkaloid to a second alkaloid are provided in Table 5. An amino acid sequence for an NDM that is utilized in converting a first alkaloid to a second alkaloid may be 75% or more identical to a given amino acid sequence as listed in Table 5. For example, an amino acid sequence for such an epimerase may comprise an amino acid sequence that is at least 75% or more, 80% or more, 81% or more, 82% or more, 83% or more, 84% or more, 85% or more, 86% or more, 87% or more, 88% or more, 89% or more, 90% or more, 91% or more, 92% or more, 93% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more, or 99% or more identical to an amino acid sequence as provided herein. Additionally, in certain embodiments, an "identical" amino acid sequence contains at least 80%-99% identity at the amino acid level to the specific amino acid sequence. In some cases an "identical" amino acid sequence contains at least about 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94% and more in certain cases, at least 95%, 96%, 97%, 98% and 99% identity, at the amino acid level. In some cases, the amino acid sequence may be identical but the DNA sequence is altered such as to optimize codon usage for the host organism, for example.

[00204] An engineered host cell may be provided that produces an NDM that converts a first alkaloid to a second alkaloid, wherein the NDM comprises an amino acid sequence as listed in Table 5. An engineered host cell may be provided that produces one or more NDM enzymes. The NDM that is produced within the engineered host cell may be recovered and purified so as to form a biocatalyst. The process may include contacting the first alkaloid with an NDM in an amount sufficient to convert said first alkaloid to a second alkaloid. In examples, the first alkaloid may be contacted with a sufficient amount of the one or more enzymes such that at least 5% of said first alkaloid is converted to a second alkaloid. In further examples, the first alkaloid may be contacted with a sufficient amount of the one or more enzymes such that at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 82%, at least 84%, at least 86%, at least 88%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, at least 99.7%, or 100% of said first alkaloid is converted to a second alkaloid.

[00205] The one or more enzymes that may be used to convert a first alkaloid to a second alkaloid may contact the first alkaloid in vitro. Additionally, or alternatively, the one or more enzymes that may be used to convert a first alkaloid to a second alkaloid may contact the first alkaloid in vivo. In some examples, the one or more enzymes that may be used to convert a first alkaloid to a second alkaloid may be provided to a cell having the first alkaloid within. In some examples, the one or more enzymes that may be used to convert a first alkaloid to a second alkaloid may be produced within an engineered host cell.

[00206] In some examples, the methods provide for engineered host cells that produce an alkaloid product, wherein the *N*-demethylation of a substrate to a product may comprise a key step in the production of an alkaloid product. In some examples, the alkaloid produced is a nor-opioid or a nal-opioid. In still other embodiments, the alkaloid produced is derived from a nor-opioid or a nal-opioid. In another embodiment, a first alkaloid is an intermediate toward the product of the engineered host cell. In still other embodiments, the alkaloid product is selected from the group consisting of norcodeine, noroxycodone, northebaine, norhydrocodone, nordihydro-codeine, nor-14-hydroxy-codeine, norcodeinone, nor-14-hydroxy-codeinone, normorphine, noroxymorphine, nororipavine, norhydro-morphine, nordihydro-morphine, nor-14-hydroxy-morphine, normorphinone, and nor-14-hydroxy-morphinone.

[00207] In some examples, the substrate alkaloid is an opioid selected from the group consisting of codeine, oxycodone, thebaine, hydrocodone, dihydrocodeine, 14-hydroxycodone, codeinone, 14-hydroxycodeinone, morphine, oxymorphine, oripavine, hydromorphine, dihydromorphine, 14-hydroxy-morphine, morphinone, and 14-hydroxy-morphinone.

N-Methyltransferase Modifications

[00208] Some methods, processes, and systems provided herein describe the conversion of a first alkaloid to a second alkaloid by the addition of an *N*-linked sidechain group. Some methods, processes, and systems provided herein describe the conversion of a first alkaloid to a second alkaloid by the transfer of a sidechain group from a cosubstrate to the first alkaloid. Some of these methods, processes, and systems may comprise an engineered host cell. In some examples, the conversion of a first alkaloid to a second alkaloid is a key step in the conversion of a substrate to a nal-opioid. In some examples, the conversion of a first alkaloid to a second alkaloid comprises a methyltransferase reaction.

[00209] FIG. 8 illustrates an enzyme having *N*-methyltransferase (NMT) activity, in accordance with embodiments of the invention. Specifically, the enzyme may act on any morphinan alkaloid structure to add a methyl group or other carbon moiety to the nitrogen. S-Adenosyl methionine (SAM) may act as the donor of the functional group (methyl, allyl, cyclopropylmethyl, or other).

[00210] Examples of amino acid sequences of NMT enzymes are set forth in Table 6. An amino acid sequence for an NMT that is utilized in converting a first alkaloid to a second alkaloid may be 75% or more identical to a given amino acid sequence as listed in Table 6. For example, an amino acid sequence

for such an epimerase may comprise an amino acid sequence that is at least 75% or more, 80% or more, 81% or more, 82% or more, 83% or more, 84% or more, 85% or more, 86% or more, 87% or more, 88% or more, 89% or more, 90% or more, 91% or more, 92% or more, 93% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more, or 99% or more identical to an amino acid sequence as provided herein. Additionally, in certain embodiments, an "identical" amino acid sequence contains at least 80%-99% identity at the amino acid level to the specific amino acid sequence. In some cases an "identical" amino acid sequence contains at least about 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94% and more in certain cases, at least 95%, 96%, 97%, 98% and 99% identity, at the amino acid level. In some cases, the amino acid sequence may be identical but the DNA sequence is altered such as to optimize codon usage for the host organism, for example.

[00211] An engineered host cell may be provided that produces an NMT that converts a first alkaloid to a second alkaloid, wherein the NMT comprises an amino acid sequence as provided in **Table 6**. An engineered host cell may be provided that produces one or more NMT enzymes. The NMT that is produced within the engineered host cell may be recovered and purified so as to form a biocatalyst. The process may include contacting the first alkaloid with an NMT in an amount sufficient to convert said first alkaloid to a second alkaloid. In examples, the first alkaloid may be contacted with a sufficient amount of the one or more enzymes such that at least 5% of said first alkaloid is converted to a second alkaloid. In further examples, the first alkaloid may be contacted with a sufficient amount of the one or more enzymes such that at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 82%, at least 84%, at least 86%, at least 88%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, at least 99.7%, or 100% of said first alkaloid is converted to a second alkaloid.

[00212] The one or more enzymes that may be used to convert a first alkaloid to a second alkaloid may contact the first alkaloid in vitro. Additionally, or alternatively, the one or more enzymes that may be used to convert a first alkaloid to a second alkaloid may contact the first alkaloid in vivo. In some examples, the one or more enzymes that may be used to convert a first alkaloid to a second alkaloid may be provided to a cell having the first alkaloid within. In some examples, the one or more enzymes that may be used to convert a first alkaloid to a second alkaloid may be produced within an engineered host cell.

[00213] In some examples, the methods provide for engineered host cells that produce an alkaloid product, wherein the *N*-methyltransferase of a substrate to a product may comprise a key step in the production of an alkaloid product. In some examples, the alkaloid produced is a nal-opioid. In still other embodiments, the alkaloid produced is derived from a nor-opioid or a nal-opioid. In another embodiment, a first alkaloid is an intermediate toward the product of the engineered host cell. In still other embodiments, the alkaloid product is selected from the group including naloxone, naltrexone, and nalmefene.

[00214] In some examples, the substrate alkaloid is an opioid selected from the group consisting of norcodeine, noroxycodone, northebaine, norhydrocodone, nordihydro-codeine, nor-14-hydroxy-codeine, norcodeinone, nor-14-hydroxy-codeinone, normorphine, noroxymorphone, nororipavine, norhydro-morphone, nordihydro-morphine, nor-14-hydroxy-morphine, normorphinone, and nor-14-hydroxy-morphinone. In some examples, the cosubstrate is S-adenosylmethionine, allyl-S-adenosylmethionine, or cyclopropylmethyl-S-adenosylmethionine.

Heterologous Coding Sequences

[00215] In some instances, the engineered host cells harbor one or more heterologous coding sequences (such as two or more, three or more, four or more, five or more) which encode activity(ies) that enable the engineered host cells to produce desired enzymes of interest and/or BIAs of interest, e.g., as described herein. As used herein, the term "heterologous coding sequence" is used to indicate any polynucleotide that codes for, or ultimately codes for, a peptide or protein or its equivalent amino acid sequence, e.g., an enzyme, that is not normally present in the host organism and may be expressed in the host cell under proper conditions. As such, "heterologous coding sequences" includes multiple copies of coding sequences that are normally present in the host cell, such that the cell is expressing additional copies of a coding sequence that are not normally present in the cells. The heterologous coding sequences may be RNA or any type thereof, e.g., mRNA, DNA or any type thereof, e.g., cDNA, or a hybrid of RNA/DNA. Coding sequences of interest include, but are not limited to, full-length transcription units that include such features as the coding sequence, introns, promoter regions, 3'-UTRs, and enhancer regions.

[00216] In examples, the engineered host cells may comprise a plurality of heterologous coding sequences each encoding an enzyme, such as an enzyme listed in **Table 3**. In some examples, the plurality of enzymes encoded by the plurality of heterologous coding sequences may be distinct from each other. In some examples, some of the plurality of enzymes encoded by the plurality of heterologous coding sequences may be distinct from each other and some of the plurality of enzymes encoded by the plurality of heterologous coding sequences may be duplicate copies.

[00217] In some examples, the heterologous coding sequences may be operably connected. Heterologous coding sequences that are operably connected may be within the same pathway of producing a particular benzyloquinoline alkaloid product and/or epimerase product. In some examples, the operably connected heterologous coding sequences may be directly sequential along the pathway of producing a particular benzyloquinoline alkaloid product and/or epimerase product. In some examples, the operably connected heterologous coding sequences may have one or more native enzymes between one or more of the enzymes encoded by the plurality of heterologous coding sequences. In some examples, the heterologous coding sequences may have one or more heterologous enzymes between one or more of the enzymes encoded by the plurality of heterologous coding sequences. In some examples, the heterologous coding sequences may have one or more non-native enzymes between one or more of the enzymes encoded by the plurality of heterologous coding sequences.

[00218] The engineered host cells may also be modified to possess one or more genetic alterations to accommodate the heterologous coding sequences. Alterations of the native host genome include, but are not limited to, modifying the genome to reduce or ablate expression of a specific protein that may interfere with the desired pathway. The presence of such native proteins may rapidly convert one of the intermediates or final products of the pathway into a metabolite or other compound that is not usable in the desired pathway. Thus, if the activity of the native enzyme were reduced or altogether absent, the produced intermediates would be more readily available for incorporation into the desired product.

[00219] Heterologous coding sequences include but are not limited to sequences that encode enzymes, either wild-type or equivalent sequences, that are normally responsible for the production of BIAs of interest in plants. In some cases, the enzymes for which the heterologous sequences code may be any of the enzymes in the 1-benzylisoquinoline alkaloid pathway, and may be from any convenient source. The choice and number of enzymes encoded by the heterologous coding sequences for the particular synthetic pathway may be selected based upon the desired product. In certain embodiments, the host cells of the invention may include 1 or more, 2 or more, 3 or more, 4 or more, 5 or more, 6 or more, 7 or more, 8 or more, 9 or more, 10 or more, 11 or more, 12 or more, 13 or more, 14 or more, or even 15 or more heterologous coding sequences, such as 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 heterologous coding sequences.

[00220] As used herein, the term "heterologous coding sequences" also includes the coding portion of the peptide or enzyme, i.e., the cDNA or mRNA sequence, of the peptide or enzyme, as well as the coding portion of the full-length transcriptional unit, i.e., the gene including introns and exons, as well as "codon optimized" sequences, truncated sequences or other forms of altered sequences that code for the enzyme or code for its equivalent amino acid sequence, provided that the equivalent amino acid sequence produces a functional protein. Such equivalent amino acid sequences may have a deletion of one or more amino acids, with the deletion being N-terminal, C-terminal, or internal. Truncated forms are envisioned as long as they have the catalytic capability indicated herein. Fusions of two or more enzymes are also envisioned to facilitate the transfer of metabolites in the pathway, provided that catalytic activities are maintained.

[00221] Operable fragments, mutants, or truncated forms may be identified by modeling and/or screening. In some cases, this is achieved by deletion of, for example, N-terminal, C-terminal, or internal regions of the protein in a step-wise fashion, followed by analysis of the resulting derivative with regard to its activity for the desired reaction compared to the original sequence. If the derivative in question operates in this capacity, it is considered to constitute an equivalent derivative of the enzyme proper.

[00222] In examples, some heterologous proteins may show occurrences where they are incorrectly processed when expressed in a recombinant host. For example, plant proteins such as cytochrome P450 enzymes expressed in microbial production hosts may have occurrences of incorrect processing. In particular, salutaridine synthase may undergo N-linked glycosylation when heterologously expressed in

yeast. This N-linked glycosylation may not be observed in plants, which may be indicative of incorrect N-terminal sorting of the nascent SalSyn transcript so as to reduce the activity of the enzyme in the heterologous microbial host. In such examples, protein engineering directed at correcting N-terminal sorting of the nascent transcript so as to remove the N-linked glycosylation pattern may result in improved activity of the salutaridine synthase enzyme in the recombinant production host.

[00223] Aspects of the invention also relate to heterologous coding sequences that code for amino acid sequences that are equivalent to the native amino acid sequences for the various enzymes. An amino acid sequence that is "equivalent" is defined as an amino acid sequence that is not identical to the specific amino acid sequence, but rather contains at least some amino acid changes (deletions, substitutions, inversions, insertions, etc.) that do not essentially affect the biological activity of the protein as compared to a similar activity of the specific amino acid sequence, when used for a desired purpose. The biological activity refers to, in the example of an epimerase, its catalytic activity. Equivalent sequences are also meant to include those which have been engineered and/or evolved to have properties different from the original amino acid sequence. Mutable properties of interest include catalytic activity, substrate specificity, selectivity, stability, solubility, localization, etc.

[00224] In some instances, the expression of each type of enzyme is increased through additional gene copies (i.e., multiple copies), which increases intermediate accumulation and/or BIA of interest production. Embodiments of the invention include increased BIA of interest production in a host cell through simultaneous expression of multiple species variants of a single or multiple enzymes. In some cases, additional gene copies of a single or multiple enzymes are included in the host cell. Any convenient methods may be utilized including multiple copies of a heterologous coding sequence for an enzyme in the host cell.

[00225] In some examples, the engineered host cell includes multiple copies of a heterologous coding sequence for an enzyme, such as 2 or more, 3 or more, 4 or more, 5 or more, or even 10 or more copies. In certain embodiments, the engineered host cell includes multiple copies of heterologous coding sequences for one or more enzymes, such as multiple copies of two or more, three or more, four or more, etc. In some cases, the multiple copies of the heterologous coding sequence for an enzyme are derived from two or more different source organisms as compared to the host cell. For example, the engineered host cell may include multiple copies of one heterologous coding sequence, where each of the copies is derived from a different source organism. As such, each copy may include some variations in explicit sequences based on inter-species differences of the enzyme of interest that is encoded by the heterologous coding sequence.

[00226] In certain embodiments, the engineered host cell includes multiple copies of heterologous coding sequences for one or more enzymes, such as multiple copies of two or more, three or more, four or more, etc. In some cases, the multiple copies of the heterologous coding sequence for an enzyme are derived from two or more different source organisms as compared to the host cell. For example, the engineered

host cell may include multiple copies of one heterologous coding sequence, where each of the copies is derived from a different source organism. As such, each copy may include some variations in explicit sequences based on inter-species differences of the enzyme of interest that is encoded by the heterologous coding sequence.

[00227] The engineered host cell medium may be sampled and monitored for the production of BIAs of interest. The BIAs of interest may be observed and measured using any convenient methods. Methods of interest include, but are not limited to, LC-MS methods (e.g., as described herein) where a sample of interest is analyzed by comparison with a known amount of a standard compound. Additionally, there are other ways that BIAs of interest may be observed and/or measured. Examples of alternative ways of observing and/or measuring BIAs include GC-MS, UV-vis spectroscopy, NMR, LC-NMR, LC-UV, TLC, capillary electrophoresis, among others. Identity may be confirmed, e.g., by m/z and MS/MS fragmentation patterns, MRM transitions, and quantitation or measurement of the compound may be achieved via LC trace peaks of known retention time and/or EIC MS peak analysis by reference to corresponding LC-MS analysis of a known amount of a standard of the compound. In some cases, identity may be confirmed via multiple reaction monitoring using mass spectrometry.

[00228] Additionally, a culture of the engineered host cell may be sampled and monitored for the production of enzymes of interest, such as a DRS-DRR enzyme. The enzymes of interest may be observed and measured using any convenient methods. Methods of interest include enzyme activity assays, polyacrylamide gel electrophoresis, carbon monoxide spectroscopy, and western blot analysis.

METHODS

Methods for Culturing Host Cells for BIA production

[00229] As summarized above, some aspects of the invention include methods of preparing benzyloquinoline alkaloids (BIAs) of interest. Additionally, some aspects of the invention include methods of preparing enzymes of interest. As such, some aspects of the invention include culturing an engineered host cell under conditions in which the one or more host cell modifications (e.g., as described herein) are functionally expressed such that the cell converts starting compounds of interest into product enzymes and/or BIAs of interest. Also provided are methods that include culturing an engineered host cell under conditions suitable for protein production such that one or more heterologous coding sequences are functionally expressed and convert starting compounds of interest into product enzymes or BIAs of interest. In examples, the method is a method of preparing a benzyloquinoline alkaloid (BIA) that includes culturing an engineered host cell (e.g., as described herein); adding a starting compound to the cell culture; and recovering the BIA from the cell culture. In some examples, the method is a method of preparing an enzyme that includes culturing an engineered host cell (e.g., as described herein); adding a starting compound to the cell culture; and recovering the enzyme from the cell culture.

[00230] Fermentation media may contain suitable carbon substrates. The source of carbon suitable to perform the methods of this disclosure may encompass a wide variety of carbon containing substrates.

Suitable substrates may include, without limitation, monosaccharides (e.g., glucose, fructose, galactose, xylose), oligosaccharides (e.g., lactose, sucrose, raffinose), polysaccharides (e.g., starch, cellulose), or a combination thereof. In some cases, unpurified mixtures from renewable feedstocks may be used (e.g., cornsteep liquor, sugar beet molasses, barley malt). In some cases, the carbon substrate may be a one-carbon substrate (e.g., methanol, carbon dioxide) or a two-carbon substrate (e.g., ethanol). In other cases, other carbon containing compounds may be utilized, for example, methylamine, glucosamine, and amino acids.

[00231] Any convenient methods of culturing engineered host cells may be employed for producing the enzymes and/or BIAs of interest. The particular protocol that is employed may vary, e.g., depending on the engineered host cell, the heterologous coding sequences, the enzymes of interest, the BIAs of interest, etc. The cells may be present in any convenient environment, such as an environment in which the cells are capable of expressing one or more functional heterologous enzymes. In some embodiments, the cells are cultured under conditions that are conducive to enzyme expression and with appropriate substrates available to allow production of enzymes and/or BIAs of interest *in vivo*. In some embodiments, the functional enzymes are extracted from the engineered host for production of enzymes and/or BIAs of interest under *in vitro* conditions. In some instances, the engineered host cells are placed back into a multicellular host organism. The engineered host cells are in any phase of growth, including, but not limited to, stationary phase and log-growth phase, etc. In addition, the cultures themselves may be continuous cultures or they may be batch cultures.

[00232] Cells may be grown in an appropriate fermentation medium at a temperature between 14-40°C. Cells may be grown with shaking at any convenient speed (e.g., 200 rpm). Cells may be grown at a suitable pH. Suitable pH ranges for the fermentation may be between pH 5-9. Fermentations may be performed under aerobic, anaerobic, or microaerobic conditions. Any suitable growth medium may be used. Suitable growth media may include, without limitation, common commercially prepared media such as synthetic defined (SD) minimal media or yeast extract peptone dextrose (YEPD) rich media. Any other rich, defined, or synthetic growth media appropriate to the microorganism may be used.

[00233] Cells may be cultured in a vessel of essentially any size and shape. Examples of vessels suitable to perform the methods of this disclosure may include, without limitation, multi-well shake plates, test tubes, flasks (baffled and non-baffled), and bioreactors. The volume of the culture may range from 10 microliters to greater than 10,000 liters.

[00234] The addition of agents to the growth media that are known to modulate metabolism in a manner desirable for the production of alkaloids may be included. In a non-limiting example, cyclic adenosine 2'3'-monophosphate may be added to the growth media to modulate catabolite repression.

[00235] Any convenient cell culture conditions for a particular cell type may be utilized. In certain embodiments, the host cells that include one or more modifications are cultured under standard or readily optimized conditions, with standard cell culture media and supplements. As one example, standard

growth media when selective pressure for plasmid maintenance is not required may contain 20 g/L yeast extract, 10 g/L peptone, and 20 g/L dextrose (YPD). Host cells containing plasmids are grown in synthetic complete (SC) media containing 1.7 g/L yeast nitrogen base, 5 g/L ammonium sulfate, and 20 g/L dextrose supplemented with the appropriate amino acids required for growth and selection.

Alternative carbon sources which may be useful for inducible enzyme expression include, but are not limited to, sucrose, raffinose, and galactose. Cells are grown at any convenient temperature (e.g., 30°C) with shaking at any convenient rate (e.g., 200 rpm) in a vessel, e.g., in test tubes or flasks in volumes ranging from 1-1000 mL, or larger, in the laboratory.

[00236] Culture volumes may be scaled up for growth in larger fermentation vessels, for example, as part of an industrial process. The industrial fermentation process may be carried out under closed-batch, fed-batch, or continuous chemostat conditions, or any suitable mode of fermentation. In some cases, the cells may be immobilized on a substrate as whole cell catalysts and subjected to fermentation conditions for alkaloid production.

[00237] A batch fermentation is a closed system, in which the composition of the medium is set at the beginning of the fermentation and not altered during the fermentation process. The desired organism(s) are inoculated into the medium at the beginning of the fermentation. In some instances, the batch fermentation is run with alterations made to the system to control factors such as pH and oxygen concentration (but not carbon). In this type of fermentation system, the biomass and metabolite compositions of the system change continuously over the course of the fermentation. Cells typically proceed through a lag phase, then to a log phase (high growth rate), then to a stationary phase (growth rate reduced or halted), and eventually to a death phase (if left untreated). In additional cases, the batch fermentation system may be opened at certain times to add additional substrates for fermenting the desired organism. In particular, in some cases, a fermentation system may include a fed batch reactor.

[00238] A continuous fermentation is an open system, in which a defined fermentation medium is added continuously to the bioreactor and an equal amount of fermentation media is continuously removed from the vessel for processing. Continuous fermentation systems are generally operated to maintain steady state growth conditions, such that cell loss due to medium being removed must be balanced by the growth rate in the fermentation. Continuous fermentations are generally operated at conditions where cells are at a constant high cell density. Continuous fermentations allow for the modulation of one or more factors that affect target product concentration and/or cell growth.

[00239] The liquid medium may include, but is not limited to, a rich or synthetic defined medium having an additive component described above. Media components may be dissolved in water and sterilized by heat, pressure, filtration, radiation, chemicals, or any combination thereof. Several media components may be prepared separately and sterilized, and then combined in the fermentation vessel. The culture medium may be buffered to aid in maintaining a constant pH throughout the fermentation.

[00240] Process parameters including temperature, dissolved oxygen, pH, stirring, aeration rate, and cell density may be monitored or controlled over the course of the fermentation. For example, temperature of a fermentation process may be monitored by a temperature probe immersed in the culture medium. The culture temperature may be controlled at the set point by regulating the jacket temperature. Water may be cooled in an external chiller and then flowed into the bioreactor control tower and circulated to the jacket at the temperature required to maintain the set point temperature in the vessel.

[00241] Additionally, a gas flow parameter may be monitored in a fermentation process. For example, gases may be flowed into the medium through a sparger. Gases suitable for the methods of this disclosure may include compressed air, oxygen, and nitrogen. Gas flow may be at a fixed rate or regulated to maintain a dissolved oxygen set point.

[00242] The pH of a culture medium may also be monitored. In examples, the pH may be monitored by a pH probe that is immersed in the culture medium inside the vessel. If pH control is in effect, the pH may be adjusted by acid and base pumps which add each solution to the medium at the required rate. The acid solutions used to control pH may be sulfuric acid or hydrochloric acid. The base solutions used to control pH may be sodium hydroxide, potassium hydroxide, or ammonium hydroxide.

[00243] Further, dissolved oxygen may be monitored in a culture medium by a dissolved oxygen probe immersed in the culture medium. If dissolved oxygen regulation is in effect, the oxygen level may be adjusted by increasing or decreasing the stirring speed. The dissolved oxygen level may also be adjusted by increasing or decreasing the gas flow rate. The gas may be compressed air, oxygen, or nitrogen.

[00244] Stir speed may also be monitored in a fermentation process. In examples, the stirrer motor may drive an agitator. The stirrer speed may be set at a consistent rpm throughout the fermentation or may be regulated dynamically to maintain a set dissolved oxygen level.

[00245] Additionally, turbidity may be monitored in a fermentation process. In examples, cell density may be measured using a turbidity probe. Alternatively, cell density may be measured by taking samples from the bioreactor and analyzing them in a spectrophotometer. Further, samples may be removed from the bioreactor at time intervals through a sterile sampling apparatus. The samples may be analyzed for alkaloids produced by the host cells. The samples may also be analyzed for other metabolites and sugars, the depletion of culture medium components, or the density of cells.

[00246] In another example, a feed stock parameter may be monitored during a fermentation process. In particular, feed stocks including sugars and other carbon sources, nutrients, and cofactors that may be added into the fermentation using an external pump. Other components may also be added during the fermentation including, without limitation, anti-foam, salts, chelating agents, surfactants, and organic liquids.

[00247] Any convenient codon optimization techniques for optimizing the expression of heterologous polynucleotides in host cells may be adapted for use in the subject host cells and methods, see e.g.,

Gustafsson, C. et al. (2004) *Trends Biotechnol*, 22, 346-353, which is incorporated by reference in its entirety.

[00248] The subject method may also include adding a starting compound to the cell culture. Any convenient methods of addition may be adapted for use in the subject methods. The cell culture may be supplemented with a sufficient amount of the starting materials of interest (e.g., as described herein), e.g., a mM to μ M amount such as between about 1-5 mM of a starting compound. It is understood that the amount of starting material added, the timing and rate of addition, the form of material added, etc., may vary according to a variety of factors. The starting material may be added neat or pre-dissolved in a suitable solvent (e.g., cell culture media, water, or an organic solvent). The starting material may be added in concentrated form (e.g., 10x over desired concentration) to minimize dilution of the cell culture medium upon addition. The starting material may be added in one or more batches, or by continuous addition over an extended period of time (e.g., hours or days).

Methods for Isolating Products from the Fermentation Medium

[00249] The subject methods may also include recovering the enzymes and/or BIAs of interest from the cell culture. Any convenient methods of separation and isolation (e.g., chromatography methods or precipitation methods) may be adapted for use in the subject methods to recover the enzymes and/or BIAs of interest from the cell culture. Filtration methods may be used to separate soluble from insoluble fractions of the cell culture. In some cases, liquid chromatography methods (e.g., reverse phase HPLC, size exclusion, normal phase chromatography) may be used to separate the BIA of interest from other soluble components of the cell culture. In some cases, extraction methods (e.g., liquid extraction, pH based purification, solid phase extraction, affinity chromatography, ion exchange, etc.) may be used to separate the enzymes and/or BIAs of interest from other components of the cell culture.

[00250] The produced alkaloids may be isolated from the fermentation medium using methods known in the art. A number of recovery steps may be performed immediately after (or in some instances, during) the fermentation for initial recovery of the desired product. Through these steps, the alkaloids (e.g., BIAs) may be separated from the cells, cellular debris and waste, and other nutrients, sugars, and organic molecules may remain in the spent culture medium. This process may be used to yield a BIA-enriched product.

[00251] In an example, a product stream having a benzylisoquinoline alkaloid (BIA) product is formed by providing engineered yeast cells and a feedstock including nutrients and water to a batch reactor. In particular, the engineered yeast cells may be subjected to fermentation by incubating the engineered yeast cells for a time period of at least about 5 minutes to produce a solution comprising the BIA product and cellular material. Once the engineered yeast cells have been subjected to fermentation, at least one separation unit may be used to separate the BIA product from the cellular material to provide the product stream comprising the BIA product. In particular, the product stream may include the BIA product as

well as additional components, such as a clarified yeast culture medium. Additionally, a BIA product may comprise one or more BIAs of interest, such as one or more BIA compounds.

[00252] Different methods may be used to remove cells from a bioreactor medium that include an enzyme and/or BIA of interest. In examples, cells may be removed by sedimentation over time. This process of sedimentation may be accelerated by chilling or by the addition of fining agents such as silica. The spent culture medium may then be siphoned from the top of the reactor or the cells may be decanted from the base of the reactor. Alternatively, cells may be removed by filtration through a filter, a membrane, or other porous material. Cells may also be removed by centrifugation, for example, by continuous flow centrifugation or by using a continuous extractor.

[00253] If some valuable enzymes and/or BIAs of interest are present inside the cells, the cells may be permeabilized or lysed and the cell debris may be removed by any of the methods described above. Agents used to permeabilize the cells may include, without limitation, organic solvents (e.g., DMSO) or salts (e.g., lithium acetate). Methods to lyse the cells may include the addition of surfactants such as sodium dodecyl sulfate, or mechanical disruption by bead milling or sonication.

[00254] Enzymes and/or BIAs of interest may be extracted from the clarified spent culture medium through liquid-liquid extraction by the addition of an organic liquid that is immiscible with the aqueous culture medium. In examples, the use of liquid-liquid extraction may be used in addition to other processing steps. Examples of suitable organic liquids include, but are not limited to, isopropyl myristate, ethyl acetate, chloroform, butyl acetate, methylisobutyl ketone, methyl oleate, toluene, oleyl alcohol, ethyl butyrate. The organic liquid may be added to as little as 10% or as much as 100% of the aqueous medium. The organic liquid may be as little as 10%, may be 100%, may be 200%, may be 300%, may be 400%, may be 500%, may be 600%, may be 700%, may be 800%, may be 900%, may be 1000%, may be more than 1000%, or may be a percentage in between those listed herein of the volume of the aqueous liquid.

[00255] In some cases, the organic liquid may be added at the start of the fermentation or at any time during the fermentation. This process of extractive fermentation may increase the yield of enzymes and/or BIAs of interest from the host cells by continuously removing enzymes and/or BIAs to the organic phase.

[00256] Agitation may cause the organic phase to form an emulsion with the aqueous culture medium. Methods to encourage the separation of the two phases into distinct layers may include, without limitation, the addition of a demulsifier or a nucleating agent, or an adjustment of the pH. The emulsion may also be centrifuged to separate the two phases, for example, by continuous conical plate centrifugation.

[00257] Alternatively, the organic phase may be isolated from the aqueous culture medium so that it may be physically removed after extraction. For example, the solvent may be encapsulated in a membrane.

[00258] In examples, enzymes and/or BIAs of interest may be extracted from a fermentation medium using adsorption methods. In examples, BIAs of interest may be extracted from clarified spent culture medium by the addition of a resin such as Amberlite® XAD4 or another agent that removes BIAs by adsorption. The BIAs of interest may then be released from the resin using an organic solvent. Examples of suitable organic solvents include, but are not limited to, methanol, ethanol, ethyl acetate, or acetone.

[00259] BIAs of interest may also be extracted from a fermentation medium using filtration. At high pH, the BIAs of interest may form a crystalline-like precipitate in the bioreactor. This precipitate may be removed directly by filtration through a filter, membrane, or other porous material. The precipitate may also be collected by centrifugation and/or decantation.

[00260] The extraction methods described above may be carried out either *in situ* (in the bioreactor) or *ex situ* (e.g., in an external loop through which media flows out of the bioreactor and contacts the extraction agent, then is recirculated back into the vessel). Alternatively, the extraction methods may be performed after the fermentation is terminated using the clarified medium removed from the bioreactor vessel.

Methods for Purifying Products from Alkaloid-Enriched Solutions

[00261] Subsequent purification steps may involve treating the post-fermentation solution enriched with BIA product(s) of interest using methods known in the art to recover individual product species of interest to high purity.

[00262] In one example, BIAs of interest extracted in an organic phase may be transferred to an aqueous solution. In some cases, the organic solvent may be evaporated by heat and/or vacuum, and the resulting powder may be dissolved in an aqueous solution of suitable pH. In a further example, the BIAs of interest may be extracted from the organic phase by addition of an aqueous solution at a suitable pH that promotes extraction of the BIAs of interest into the aqueous phase. The aqueous phase may then be removed by decantation, centrifugation, or another method.

[00263] The BIA-containing solution may be further treated to remove metals, for example, by treating with a suitable chelating agent. The BIA of interest-containing solution may be further treated to remove other impurities, such as proteins and DNA, by precipitation. In one example, the BIA of interest-containing solution is treated with an appropriate precipitation agent such as ethanol, methanol, acetone, or isopropanol. In an alternative example, DNA and protein may be removed by dialysis or by other methods of size exclusion that separate the smaller alkaloids from contaminating biological macromolecules.

[00264] In further examples, the solution containing BIAs of interest may be extracted to high purity by continuous cross-flow filtration using methods known in the art.

[00265] If the solution contains a mixture of BIAs of interest, it may be subjected to acid-base treatment to yield individual BIA of interest species using methods known in the art. In this process, the pH of the aqueous solution is adjusted to precipitate individual BIAs.

[00266] For high purity, small-scale preparations, the BIAs may be purified in a single step by liquid chromatography.

Liquid Chromatography Mass Spectrometry (LCMS) Method

[00267] The BIA compounds of interest, including 1-benzylisoquinoline alkaloids, bisbenzylisoquinoline alkaloids, promorphinan alkaloids, morphinan alkaloids, nal-opioids, and nor-opioids, may be separated using liquid chromatography, and detected and quantified using mass spectrometry. Compound identity may be confirmed by characteristic elution time, mass-to-charge ratio (m/z) and fragmentation patterns (MS/MS). Quantitation may be performed by comparison of compound peak area to a standard curve of a known reference standard compound. Additionally, BIAs of interest may be detected by alternative methods such as GC-MS, UV-vis spectroscopy, NMR, LC-NMR, LC-UV, TLC, and capillary electrophoresis.

Purpald Assay Method

[00268] For high throughput screening of demethylation reactions a purpald assay may be used. For example, demethylation catalyzed by 2-oxoglutarate dependent dioxygenases produces formaldehyde as a product as shown in the generalized chemical equation: [substrate] + 2-oxoglutarate + O₂ $\xrightarrow{\text{enzyme}}$ [product] + formaldehyde + succinate + CO₂. Purpald reagent in alkaline conditions undergoes a color change in the presence of formaldehyde that can be quantified to concentrations as low as 1 nM with a spectrophotometer at 510 nm.

Yeast-Derived Alkaloid APIs Versus Plant-Derived APIs

[00269] The clarified yeast culture medium (CYCM) may contain a plurality of impurities. The clarified yeast culture medium may be dehydrated by vacuum and/or heat to yield an alkaloid-rich powder. This product is analogous to the concentrate of poppy straw (CPS) or opium, which is exported from poppy-growing countries and purchased by API manufacturers. For the purposes of this invention, CPS is a representative example of any type of purified plant extract from which the desired alkaloids product(s) may ultimately be further purified. **Table 10** and **Table 11** highlight the impurities in these two products that may be specific to either CYCM or CPS or may be present in both. While some BIAs may have a pigment as an impurity, other BIAs may be categorized as pigments themselves. Accordingly, these BIAs may be assessed for impurities based on non-pigment impurities. By analyzing a product of unknown origin for a subset of these impurities, a person of skill in the art could determine whether the product originated from a yeast or plant production host.

[00270] API-grade pharmaceutical ingredients are highly purified molecules. As such, impurities that could indicate the plant- or yeast-origin of an API (such as those listed in **Table 10** and **Table 11**) may not be present at the API stage of the product. Indeed, many of the API products derived from yeast strains of the present invention may be largely indistinguishable from the traditional plant-derived APIs.

In some cases, however, conventional alkaloid compounds may be subjected to chemical modification using chemical synthesis approaches, which may show up as chemical impurities in plant-based products that require such chemical modifications. For example, chemical derivatization may often result in a set of impurities related to the chemical synthesis processes. In certain situations, these modifications may be performed biologically in the yeast production platform, thereby avoiding some of the impurities associated with chemical derivation from being present in the yeast-derived product. In particular, these impurities from the chemical derivation product may be present in an API product that is produced using chemical synthesis processes but may be absent from an API product that is produced using a yeast-derived product. Alternatively, if a yeast-derived product is mixed with a chemically-derived product, the resulting impurities may be present but in a lesser amount than would be expected in an API that only or primarily contains chemically-derived products. In this example, by analyzing the API product for a subset of these impurities, a person of skill in the art could determine whether the product originated from a yeast production host or the traditional chemical derivatization route.

[00271] Non-limiting examples of impurities that may be present in chemically-derivatized morphinan APIs but not in biosynthesized APIs include a codeine-*O*(6)-methyl ether impurity in API codeine; 8,14-dihydroxy-7,8-dihydrocodeinone in API oxycodone; and tetrahydrothebaine in API hydrocodone. The codeine-*O*(6)-methyl ether may be formed by chemical over-methylation of morphine. The 8,14-dihydroxy-7,8-dihydrocodeinone in API oxycodone may be formed by chemical over-oxidation of thebaine. Additionally, the tetrahydrothebaine in API hydrocodone may be formed by chemical over-reduction of thebaine.

[00272] However, in the case where the yeast-derived compound and the plant-derived compound are both subjected to chemical modification through chemical synthesis approaches, the same impurities associated with the chemical synthesis process may be expected in the products. In such a situation, the starting material (e.g., CYCM or CPS) may be analyzed as described above.

Host Cell Derived Nal-opioids vs Chemically Derived Nal-opioids

[00273] Nal-opioids produced by chemical synthesis may contain a plurality of impurities. These impurities may arise from many different causes, for example, unreacted starting materials, incomplete reactions, the formation of byproducts, persistence of intermediates, dimerization, or degradation. An example of an unreacted starting material could be oxymorphone remaining in a preparation of naltrexone. An example of an impurity arising from an incomplete reaction could be 3-*O*-Methylbuprenorphine resulting from the incomplete 3-*O*-demethylation of thebaine. Chemical modification can result in the addition or removal of functional groups at off-target sites. For example, the oxidation of C10 to create 10-hydroxynaltrexone and 10-ketonaltrexone during naltrexone synthesis, or the removal of the 6-*O*-methyl group to give 6-*O*-desmethylbuprenorphine during buprenorphine synthesis. Impurities may arise from the persistence of reaction intermediates, for example the persistence of *N*-oxides like oxymorphone *N*-oxide formed during the *N*-demethylation process. Another

source of impurities is dimerization, the conjugation of two opioid molecules, for example two buprenorphine molecules (2,2'-bisbuprenorphine), two naltrexone molecules (2,2'-bisnaltrexone), or two naloxone molecules (2,2'-bisnaloxone). Impurities may arise from degradation of starting materials, reaction intermediates, or reaction products. The extreme physical conditions used in chemical syntheses may make the presence of degradation more likely. An example of an impurity that may arise from degradation is dehydrobuprenorphine produced by oxidizing conditions during buprenorphine synthesis.

[00274] Nal-opioids produced by enzyme catalysis in a host cell may contain different impurities than nal-opioids produced by chemical synthesis. Nal-opioids produced by enzyme catalysis in a host cell may contain fewer impurities than nal-opioids produced by chemical synthesis. Nal-opioids produced by enzyme catalysis in a host cell may lack certain impurities that are found in nal-opioids produced by chemical synthesis. In examples, key features of enzyme synthesis may include, (1) enzymes target a specific substrate and residue with high fidelity; (2) enzymes perform reactions in the mild physiological conditions within the cell which do not compromise the stability of the molecules; and (3) enzymes are engineered to be efficient catalysts that drive reactions to completion.

[00275] Table 12 highlights some of the impurities that may be specific to chemically produced nal-opioids. Accordingly, nal-opioids may be assessed for impurities to determine the presence or absence of any impurity from Table 12. By analyzing a product of unknown origin for a subset of these impurities, a person of skill in the art could determine whether the product originated from a chemical or enzymatic synthesis.

Methods of Engineering Host Cells

[00276] Also included are methods of engineering host cells for the purpose of producing enzymes and/or BIAs of interest. Inserting DNA into host cells may be achieved using any convenient methods. The methods are used to insert the heterologous coding sequences into the engineered host cells such that the host cells functionally express the enzymes and convert starting compounds of interest into product enzymes and/or BIAs of interest.

[00277] Any convenient promoters may be utilized in the subject engineered host cells and methods. The promoters driving expression of the heterologous coding sequences may be constitutive promoters or inducible promoters, provided that the promoters are active in the engineered host cells. The heterologous coding sequences may be expressed from their native promoters, or non-native promoters may be used. Such promoters may be low to high strength in the host in which they are used. Promoters may be regulated or constitutive. In certain embodiments, promoters that are not glucose repressed, or repressed only mildly by the presence of glucose in the culture medium, are used. Promoters of interest include but are not limited to, promoters of glycolytic genes such as the promoter of the *B. subtilis* *tsr* gene (encoding the promoter region of the fructose biphosphate aldolase gene) or the promoter from yeast *S. cerevisiae* gene coding for glyceraldehyde 3-phosphate dehydrogenase (GPD, GAPDH, or TDH3), the ADH1 promoter of baker's yeast, the phosphate-starvation induced promoters such as the PHO5 promoter of

yeast, the alkaline phosphatase promoter from *B. licheniformis*, yeast inducible promoters such as Gal1-10, Gal1, GalL, GalS, repressible promoter Met25, tetO, and constitutive promoters such as glyceraldehyde 3-phosphate dehydrogenase promoter (GPD), alcohol dehydrogenase promoter (ADH), translation-elongation factor-1- α promoter (TEF), cytochrome c-oxidase promoter (CYC1), MRP7 promoter, etc. Autonomously replicating yeast expression vectors containing promoters inducible by hormones such as glucocorticoids, steroids, and thyroid hormones may also be used and include, but are not limited to, the glucocorticoid responsive element (GRE) and thyroid hormone responsive element (TRE). These and other examples are described in U.S. Pat. No. 7,045,290, which is incorporated by reference, including the references cited therein. Additional vectors containing constitutive or inducible promoters such as α factor, alcohol oxidase, and PGH may be used. Additionally any promoter/enhancer combination (as per the Eukaryotic Promoter Data Base EPDB) could also be used to drive expression of genes. Any convenient appropriate promoters may be selected for the host cell, e.g., *E. coli*. One may also use promoter selection to optimize transcript, and hence, enzyme levels to maximize production while minimizing energy resources.

[00278] Any convenient vectors may be utilized in the subject engineered host cells and methods. Vectors of interest include vectors for use in yeast and other cells. The types of yeast vectors may be broken up into 4 general categories: integrative vectors (YIp), autonomously replicating high copy-number vectors (YE_p or 2 μ plasmids), autonomously replicating low copy-number vectors (YC_p or centromeric plasmids) and vectors for cloning large fragments (YACs). Vector DNA is introduced into prokaryotic or eukaryotic cells via any convenient transformation or transfection techniques. DNA of another source (e.g. PCR-generated double stranded DNA product, or synthesized double stranded or single stranded oligonucleotides) may be used to engineer the yeast by integration into the genome. Any single transformation event may include one or several nucleic acids (vectors, double stranded or single stranded DNA fragments) to genetically modify the host cell. **FIG. 11** illustrates examples of convenient vectors. Further, **Table 8** provides examples of base plasmid sequences from which portions of sequences represented in **FIG. 11** may be derived. In some embodiments, plasmid sequences in **Table 8** may be utilized in integrating aspects of embodiments disclosed herein.

UTILITY

[00279] The engineered host cells and methods of the invention, e.g., as described above, find use in a variety of applications. Applications of interest include, but are not limited to: research applications and therapeutic applications. Methods of the invention find use in a variety of different applications including any convenient application where the production of enzymes and/or BIAs is of interest.

[00280] The subject engineered host cells and methods find use in a variety of therapeutic applications. Therapeutic applications of interest include those applications in which the preparation of pharmaceutical products that include BIAs is of interest. The engineered host cells described herein produce BIAs of interest and enzymes of interest. Reticuline is a major branch point intermediate of interest in the

synthesis of BIAs including engineering efforts to produce end products such as opioid products. The subject host cells may be utilized to produce BIAs of interest from simple and inexpensive starting materials that may find use in the production of BIAs of interest, including reticuline, and BIA end products. As such, the subject host cells find use in the supply of therapeutically active BIAs of interest.

[00281] In some instances, the engineered host cells and methods find use in the production of commercial scale amounts of BIAs thereof where chemical synthesis of these compounds is low yielding and not a viable means for large-scale production. In certain cases, the host cells and methods are utilized in a fermentation facility that would include bioreactors (fermenters) of e.g., 5,000-200,000 liter capacity allowing for rapid production of BIAs of interest thereof for therapeutic products. Such applications may include the industrial-scale production of BIAs of interest from fermentable carbon sources such as cellulose, starch, and free sugars.

[00282] The subject engineered host cells and methods find use in a variety of research applications. The subject host cells and methods may be used to analyze the effects of a variety of enzymes on the biosynthetic pathways of a variety of enzymes and/or BIAs of interest. In addition, the engineered host cells may be engineered to produce enzymes and/or BIAs of interest that find use in testing for bioactivity of interest in as yet unproven therapeutic functions. In some cases, the engineering of host cells to include a variety of heterologous coding sequences that encode for a variety of enzymes elucidates the high yielding biosynthetic pathways towards enzymes and/or BIAs of interest. In certain cases, research applications include the production of enzymes and/or BIAs of interest for therapeutic molecules of interest that may then be further chemically modified or derivatized to desired products or for screening for increased therapeutic activities of interest. In some instances, host cell strains are used to screen for enzyme activities that are of interest in such pathways, which may lead to enzyme discovery via conversion of BIA metabolites produced in these strains.

[00283] The subject engineered host cells and methods may be used as a production platform for plant specialized metabolites. The subject host cells and methods may be used as a platform for drug library development as well as plant enzyme discovery. For example, the subject engineered host cells and methods may find use in the development of natural product based drug libraries by taking yeast strains producing interesting scaffold molecules, such as guattegaumerine, and further functionalizing the compound structure through combinatorial biosynthesis or by chemical means. By producing drug libraries in this way, any potential drug hits are already associated with a production host that is amenable to large-scale culture and production. As another example, these subject engineered host cells and methods may find use in plant enzyme discovery. The subject host cells provide a clean background of defined metabolites to express plant EST libraries to identify new enzyme activities. The subject host cells and methods provide expression methods and culture conditions for the functional expression and increased activity of plant enzymes in yeast.

KITS AND SYSTEMS

[00284] Aspects of the invention further include kits and systems, where the kits and systems may include one or more components employed in methods of the invention, e.g., engineered host cells, starting compounds, heterologous coding sequences, vectors, culture medium, etc., as described herein. In some embodiments, the subject kit includes an engineered host cell (e.g., as described herein), and one or more components selected from the following: starting compounds, a heterologous coding sequence and/or a vector including the same, vectors, growth feedstock, components suitable for use in expression systems (e.g., cells, cloning vectors, multiple cloning sites (MCS), bi-directional promoters, an internal ribosome entry site (IRES), etc.), and a culture medium.

[00285] Any of the components described herein may be provided in the kits, e.g., host cells including one or more modifications, starting compounds, culture medium, etc. A variety of components suitable for use in making and using heterologous coding sequences, cloning vectors and expression systems may find use in the subject kits. Kits may also include tubes, buffers, etc., and instructions for use. The various reagent components of the kits may be present in separate containers, or some or all of them may be pre-combined into a reagent mixture in a single container, as desired.

[00286] Also provided are systems for producing enzymes and/or BIAs of interest, where the systems may include engineered host cells including one or more modifications (e.g., as described herein), starting compounds, culture medium, a fermenter and fermentation equipment, e.g., an apparatus suitable for maintaining growth conditions for the host cells, sampling and monitoring equipment and components, and the like. A variety of components suitable for use in large scale fermentation of yeast cells may find use in the subject systems.

[00287] In some cases, the system includes components for the large scale fermentation of engineered host cells, and the monitoring and purification of enzymes and/or BIA compounds produced by the fermented host cells. In certain embodiments, one or more starting compounds (e.g., as described herein) are added to the system, under conditions by which the engineered host cells in the fermenter produce one or more desired BIA products of interest. In some instances, the host cells produce a BIA of interest (e.g., as described herein). In certain cases, the BIA products of interest are opioid products, such as thebaine, codeine, neopine, morphine, neomorphine, hydrocodone, oxycodone, hydromorphone, dihydrocodeine, 14-hydroxycodeine, dihydromorphine, and oxymorphone. In some cases, the BIA products of interest are nal-opioids, such as naltrexone, naloxone, nalmeferine, nalorphine, nalorphine, nalodeine, naldemedine, naloxegol, 6 β -naltrexol, naltrindole, methylnaltrexone, methylsamidorphan, alvimopan, axelopran, bevenpran, dinicotinate, levallorphan, samidorphan, buprenorphine, dezocine, eptazocine, butorphanol, levorphanol, nalbuphine, pentazocine, phenazocine, norbinaltorphimine, and diprenorphine. In some cases, the BIA products of interest are nor-opioids, such as norcodeine, noroxycodone, northebaine,

norhydrocodone, nordihydro-codeine, nor-14-hydroxy-codeine, norcodeinone, nor-14-hydroxy-codeinone, normorphine, noroxymorphine, nororipavine, norhydro-morphine, nordihydro-morphine, nor-14-hydroxy-morphine, normorphinone, and nor-14-hydroxy-morphinone. In some cases, the BIA products are bisbenzylisoquinoline products, such as berbaminine, guattegaumerine, dauricine, and liensinine.

[00288] In some cases, the system includes processes for monitoring and or analyzing one or more enzymes and/or BIAs of interest compounds produced by the subject host cells. For example, a LC-MS analysis system as described herein, a chromatography system, or any convenient system where the sample may be analyzed and compared to a standard, e.g., as described herein. The fermentation medium may be monitored at any convenient times before and during fermentation by sampling and analysis. When the conversion of starting compounds to enzymes and/or BIA products of interest is complete, the fermentation may be halted and purification of the BIA products may be done. As such, in some cases, the subject system includes a purification component suitable for purifying the enzymes and/or BIA products of interest from the host cell medium into which it is produced. The purification component may include any convenient means that may be used to purify the enzymes and/or BIA products of interest produced by fermentation, including but not limited to, silica chromatography, reverse-phase chromatography, ion exchange chromatography, HIC chromatography, size exclusion chromatography, liquid extraction, and pH extraction methods. In some cases, the subject system provides for the production and isolation of enzyme and/or BIA fermentation products of interest following the input of one or more starting compounds to the system.

[00289] The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how to make and use the present invention, and are not intended to limit the scope of what the inventors regard as their invention nor are they intended to represent that the experiments below are all or the only experiments performed. Efforts have been made to ensure accuracy with respect to numbers used (e.g. amounts, temperature, etc.), but some experimental errors and deviations should be accounted for. Unless indicated otherwise, parts are parts by weight, molecular weight is weight average molecular weight, temperature is in degrees Centigrade, and pressure is at or near atmospheric.

Discussion of Enzyme List

[00290] The host cells may be engineered to include one or more modifications (such as two or more, three or more, four or more, five or more, or even more modifications) that provide for the production of BIAs of interest and/or enzymes of interest. **Table 3** provides a list of exemplary genes that may be acted upon by one or more modifications so as to provide for the production of BIAs of interest and/or enzymes of interest in an engineered host cell.

[00291] Modifications of genes as provided in **Table 3** may be used to produce BIAs of interest from engineered host cells that are supplied with a medium containing the minimal nutrients required for

growth. This minimal medium may contain a carbon source, a nitrogen source, amino acids, vitamins, and salts. For example, modifications of genes as provided in **Table 3** may be used to produce BIAs of interest from engineered host cells that are fed sugar. Additionally, modifications of one or more genes as provided in **Table 3** may be used to augment the biosynthetic processes of host cells that may be engineered for drug production.

[00292] Additionally, the use of these modifications to provide for the production of BIAs of interest and/or enzymes of interest in engineered host cells is not readily apparent from the mere identification of enzymes that may be produced by the genes. In particular, synthetic pathways that have been reconstructed in host cells, such as yeast cells, as described herein comprise a variety of enzymes that do not act together in nature within a single organism. Additionally, some of the enzymes discussed herein do not act for BIA biosynthesis in their natural context. Further, some of the enzymes described herein are not evolved to function in particular host cells, such as yeast cells, and are not evolved to function together. In these cases, it would not be obvious that the enzymes would exhibit sufficient activity in the context of the synthetic BIA pathway in a host cell, such as yeast, to have sufficient flux through the pathway to produce downstream BIA end products.

[00293] For example, plant enzymes are often difficult to functionally express in heterologous microbial hosts, such as yeast. In many cases the enzymes may be misfolded, not correctly localized within the host cell, and/or incorrectly processed. The differences in protein translation and processing between yeast and plants can lead to these enzymes exhibiting substantially reduced to no detectable activities in the yeast host. These challenges arise commonly for endomembrane localized enzymes, such as cytochrome P450s, which are strongly represented in the BIA pathways. Even reduced enzyme activities may pose a substantial challenge to engineering yeast to produce complex BIAs, which requires sufficient activity at each step to ensure high-level accumulation of the desired BIA products.

[00294] Additionally, there are endogenous enzymes/pathways in some host cells, such as yeast, that may act on many of the early precursors in the BIA pathway (i.e., intermediates from tyrosine to norcoclaurine), and thus it may not be readily apparent that there would be sufficient flux through the heterologous pathway to achieve substantial BIA production given these competing endogenous pathways. For example, the Erlich pathway (Hazelwood, et al. 2008. *Appl. Environ. Microbiol.* 74: 2259-66; Larroy, et al. 2003. *Chem. Biol. Interact.* 143-144: 229-38; Larroy, et al. 2002. *Eur. J. Biochem.* 269: 5738-45) in yeast is the main endogenous pathway that would act to convert many of the intermediates in the early BIA pathway to undesired products and divert flux from the synthetic pathway.

[00295] Further, many of the enzymes as discussed herein, and as provided in **Table 3**, may function under very specific regulation strategies, including spatial regulation, in the native plant hosts, which may be lost upon transfer to the heterologous yeast host. In addition, plants present very different biochemical environments than yeast cells under which the enzymes are evolved to function, including pH, redox state, and substrate, cosubstrate, coenzyme, and cofactor availabilities. Given the differences in

biochemical environments and regulatory strategies between the native hosts and the heterologous yeast hosts, it is not obvious that the enzymes would exhibit substantial activities when in the context of the yeast environment and further not obvious that they would work together to direct simple precursors such as sugar to complex BIA compounds. Maintaining the activities of the enzymes in the yeast host is particularly important as many of the pathways have many reaction steps (>10), such that if these steps are not efficient then one would not expect accumulation of desired downstream products.

[00296] In addition, in the native plant hosts, the associated metabolites in these pathways may be localized across different cell and tissue types. In several examples, there are cell types that may be specialized for biosynthesis and cell types that may be synthesized for metabolite accumulation. This type of cell specialization may be lost when expressing the pathways within a heterologous yeast host, and may play an important role in controlling the toxicity of these metabolites on the cells. Thus, it is not obvious that yeast could be successfully engineered to biosynthesize and accumulate these metabolites without being harmed by the toxicity of these compounds.

[00297] As one example, in the native plant hosts, the enzyme BBE is reported to have dynamic subcellular localization. In particular, the enzyme BBE initially starts in the ER and then is sorted to the vacuole (Bird and Facchini. 2001. *Planta*. 213: 888-97). It has been suggested that the ER-association of BBE in plants (Alcantara, et al. 2005. *Plant Physiol*. 138: 173-83) provides the optimal basic pH (pH ~8.8) for BBE activity (Ziegler and Facchini. 2008. *Annu. Rev. Plant Biol.* 59: 735-69). As another example, there is evidence that sanguinarine biosynthesis occurs in specialized vesicles within plant cells (Amann, et al. 1986. *Planta*. 167: 310-20), but only some of the intermediates accumulate in the vesicles. This may occur so as to sequester them from other enzyme activities and/or toxic effects.

[00298] As another example, the biosynthetic enzymes in the morphinan pathway branch are all localized to the phloem, which is part of the vascular tissue in plants. In the phloem, the pathway enzymes may be further divided between two cell types: the sieve elements common to all plants, and the laticifer which is a specialized cell type present only in certain plants which make specialized secondary metabolites. The upstream enzymes (i.e., from NCS through to SalAT) are predominantly in the sieve elements, and the downstream enzymes (i.e., T6ODM, COR, CODM) are mostly in the laticifer (Onoyovwe, et al. 2013. *Plant Cell*. 25: 4110-22). Additionally, it was discovered that the final steps in the noscapine biosynthetic pathway take place in the laticifer (Chen and Facchini. 2014. *Plant J.* 77: 173-84). This compartmentalization is thought to be highly important for regulating biosynthesis by isolating or trafficking intermediates, providing optimal pH, enhancing supply of cofactors, although the nature of the poppy laticifer microenvironment is still under investigation (Ziegler and Facchini. 2008. *Annu. Rev. Plant Biol.* 59: 735-69). Further, it is predicted that several of the enzymes may function as multi-enzyme complexes or metabolic channels common to plant secondary metabolism (Kempe, et al. 2009. *Phytochemistry*. 70: 579-89; Allen, et al. 2004. *Nat. Biotechnol.* 22: 1559-66). When biosynthetic enzymes are combined from different hosts and/or expressed recombinantly in a heterologous yeast cell it is not clear that these complexes or channels will form as they would in the native host. In an additional

example, in *Coptis japonica*, berberine is biosynthesized in root tissues and then accumulated within the rhizome via the action of specialized ATP-binding cassette transport proteins (Shitan, et al. 2013. Phytochemistry. 91: 109-16). In opium poppy, morphinan alkaloids are accumulated within the latex (cytoplasm of laticifer cells) (Martin, et al. 1967. Biochemistry. 6: 2355-63).

[00299] Further, even without these considerations, it is also the case that the plant enzymes for several of the steps in the pathways described herein have not yet been characterized. For example, the conversion of tyrosine to the early benzyloquinoline alkaloid scaffold norcoclaurine has not yet been characterized. Thus, for several of the steps in the pathways described herein, alternative biosynthetic scheme were produced by bringing together enzyme activities that do not normally occur together in nature for the biosynthesis of BIAs or identifying new enzyme activities from genome sequence information to use in the reconstructed pathways.

[00300] For example, the two-step conversion of tyrosine to dopamine may be achieved by combining at least 5 mammalian enzymes and 1 bacterial enzyme, which do not naturally occur together and were not evolved to function in the context of this pathway or with plant enzymes. In these instances, it may not be obvious to utilize these enzymes for the biosynthesis of compounds they were not evolved for in nature and that they would function effectively in the context of a heterologous microbial host and this pathway.

[00301] As another example, until recent years the enzyme responsible for the conversion of (*S*)-reticuline to (*R*)-reticuline was unknown. Even when a fused epimerase enzyme was discovered, evolutionary analysis suggested that morphine-producing poppies evolved a fusion enzyme between the oxidase and reductase for an epimerase reaction, which was in contrast to non-morphine producing poppies where the epimerase enzymes were non-fused. Based on this analysis, some scholars believed the fusion of the oxidase and reductase portions was necessary to efficiently catalyze the conversion of (*S*)-reticuline to (*R*)-reticuline. Novel methods of using engineered split epimerases as discussed herein may perform this epimerization reaction in yeast and in the context of the synthetic BIA pathway, and may perform this epimerization with greater efficiency than performing an epimerization with a wild-type epimerase.

[00302] Examples of the genes that are the object of modifications so as to produce BIAs of interest and/or enzymes of interest are discussed below. Additionally, the genes are discussed in the context of a series of Figures that illustrate pathways that are used in generating BIAs of interest and/or enzymes of interest.

[00303] [TKL1] In some examples, the engineered host cell may modify the expression of the enzyme transketolase. Transketolase is encoded by the TKL1 gene. In examples, transketolase catalyzes the reaction of fructose-6-phosphate + glyceraldehyde-3-phosphate $\leftrightarrow$ xylulose-5-phosphate + erythrose-4-phosphate, as referenced in **FIG. 2**. An engineered host cell may be modified to include constitutive overexpression of the TKL1 gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the TKL1 gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy,

copies, or additional copies, of the TKL1 gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the TKL1 gene within the engineered host cell. The TKL1 gene may be derived from *Saccharomyces cerevisiae* or another species. In some examples, the TKL1 gene may be 100% similar to the naturally occurring gene.

[00304] [ZWF1] In some examples, the engineered host cell may modify the expression of the enzyme glucose-6-phosphate dehydrogenase. Glucose-6-phosphate dehydrogenase is encoded by the ZWF1 gene. In examples, glucose-6-phosphate dehydrogenase catalyzes the reaction of glucose-6-phosphate $\rightarrow$ 6-phosphogluconolactone, as referenced in **FIG. 2**. An engineered host cell may be modified to delete the coding region of the ZWF1 gene in the engineered host cell. Alternatively, the engineered host cell may be modified to disable the functionality of the ZWF1 gene, such as by introducing an inactivating mutation.

[00305] [ARO4] In some examples, the engineered host cell may modify the expression of the enzyme 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP) synthase. DAHP synthase is encoded by the ARO4 gene. In examples, DAHP synthase catalyzes the reaction of erythrose-4-phosphate + phosphoenolpyruvic acid $\rightarrow$ DAHP, as referenced in **FIG. 2**. An engineered host cell may modify the ARO4 gene to incorporate one or more feedback inhibition alleviating mutations. In particular, a feedback inhibition alleviating mutation (e.g., ARO4^{FBR}) may be incorporated as a directed mutation to a native ARO4 gene at the original locus; as an additional copy introduced as a genetic integration at a separate locus; or as an additional copy on an episomal vector such as a 2- μ m or centromeric plasmid. The identifier “FBR” in the mutation ARO4^{FBR} refers to feedback resistant mutants and mutations. The feedback inhibited copy of the DAHP synthase enzyme may be under a native yeast transcriptional regulation, such as when the engineered host cell is a yeast cell. Alternatively, the feedback inhibited copy of the DAHP synthase enzyme may be introduced to the engineered host cell with engineered constitutive or dynamic regulation of protein expression by placing it under the control of a synthetic promoter. In some cases, the ARO4 gene may be derived from *Saccharomyces cerevisiae*. In some cases, the ARO4 gene may be 100% similar to the naturally occurring gene. Examples of modifications to the ARO4 gene include a feedback inhibition resistant mutation, K229L, or Q166K.

[00306] [ARO7] In some examples, the engineered host cell may modify the expression of the enzyme chorismate mutase. Chorismate mutase is encoded by the ARO7 gene. In examples, chorismate mutase catalyzes the reaction of chorismate $\rightarrow$ prephenate, as referenced in **FIG. 2**. An engineered host cell may modify the ARO7 gene to incorporate one or more feedback inhibition alleviating mutations. In particular, a feedback inhibition alleviating mutation (e.g., ARO7^{FBR}) may be incorporated as a directed mutation to a native ARO7 gene at the original locus; as an additional copy introduced as a genetic integration at a separate locus; or as an additional copy on an episomal vector such as a 2- μ m or centromeric plasmid. The identifier “FBR” in the mutation ARO7^{FBR} refers to feedback resistant mutants and mutations. The feedback inhibited copy of the chorismate mutase enzyme may be under a native

yeast transcriptional regulation, such as when the engineered host cell is a yeast cell. Alternatively, the feedback inhibited copy of the chorismate mutase enzyme may be introduced to the engineered host cell with engineered constitutive or dynamic regulation of protein expression by placing it under the control of a synthetic promoter. In some cases, the ARO7 gene may be derived from *Saccharomyces cerevisiae*. In some cases, the ARO7 gene may be 100% similar to the naturally occurring gene. Examples of modifications to the ARO7 gene include a feedback inhibition resistant mutation or T226I.

[00307] [ARO10] In some examples, the engineered host cell may modify the expression of the enzyme phenylpyruvate decarboxylase. Phenylpyruvate decarboxylase is encoded by the ARO10 gene. In examples, phenylpyruvate decarboxylase catalyzes the reaction of hydroxyphenylpyruvate $\rightarrow$ 4-hydroxyphenylacetate (4HPA), as referenced in **FIG. 2**. An engineered host cell may be modified to include constitutive overexpression of the ARO10 gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the ARO10 gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the ARO10 gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the ARO10 gene within the engineered host cell. The ARO10 gene may be derived from *Saccharomyces cerevisiae* or another species. In some examples, the ARO10 gene may be 100% similar to the naturally occurring gene.

[00308] [ADH2-7, SFA1] In some examples, the engineered host cell may modify the expression of alcohol dehydrogenase enzymes. Alcohol dehydrogenase enzymes may be encoded by one or more of the ADH2, ADH3, ADH4, ADH5, ADH6, ADH7, and SFA1 genes. In examples, alcohol dehydrogenase catalyzes the reaction of 4HPA $\rightarrow$ tyrosol. An engineered host cell may be modified to delete the coding region of one or more of the ADH2, ADH3, ADH4, ADH5, ADH6, ADH7, and SFA1 genes in the engineered host cell. Alternatively, the engineered host cell may be modified to disable the functionality of one or more of the ADH2, ADH3, ADH4, ADH5, ADH6, ADH7, and SFA1 genes, such as by introducing an inactivating mutation.

[00309] [ALD2-6] In some examples, the engineered host cell may modify the expression of aldehyde oxidase enzymes. Aldehyde oxidase enzymes may be encoded by one or more of the ALD2, ALD3, ALD4, ALD5, and ALD6 genes. In examples, aldehyde oxidase catalyzes the reaction of 4HPA $\rightarrow$ hydroxyphenylacetic acid. An engineered host cell may be modified to delete the coding region of one or more of the ALD2, ALD3, ALD4, ALD5, and ALD6 genes in the engineered host cell. Alternatively, the engineered host cell may be modified to disable the functionality of one or more of the ALD2, ALD3, ALD4, ALD5, and ALD6 genes, such as by introducing an inactivating mutation.

[00310] [AAD4], [AAD6], [AAD10], [AAD14], [AAD15], [AAD16] In some examples, the engineered host cell may modify the expression of aryl-alcohol dehydrogenase enzymes. Aryl-alcohol dehydrogenase enzymes may be encoded by one or more of AAD4, AAD6, AAD10, AAD14, AAD15,

and AAD16 genes. In examples, aryl-alcohol dehydrogenase catalyzes the reaction of aromatic aldehyde + $\text{NAD}^+ \rightarrow$ aromatic alcohol + NADH.

[00311] [ARO9] In some examples, the engineered host cell may modify the expression of the enzyme aromatic aminotransferase. Aromatic aminotransferase is encoded by the ARO9 gene. In examples, aromatic aminotransferase catalyzes the reaction of hydroxyphenylpyruvate + L-alanine $\leftrightarrow$ tyrosine + pyruvate, as referenced in **FIG. 2**. An engineered host cell may be modified to include constitutive overexpression of the ARO9 gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the ARO9 gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the ARO9 gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the ARO9 gene within the engineered host cell. The ARO9 gene may be derived from *Saccharomyces cerevisiae* or another species. In some examples, the ARO9 gene may be 100% similar to the naturally occurring gene.

[00312] [ARO8] In some examples, the engineered host cell may modify the expression of the enzyme aromatic aminotransferase. Aromatic aminotransferase is encoded by the ARO8 gene. In examples, aromatic aminotransferase catalyzes the reaction of hydroxyphenylpyruvate + glutamate $\leftrightarrow$ tyrosine + alpha-ketoglutarate, as referenced in **FIG. 2**. An engineered host cell may be modified to include constitutive overexpression of the ARO8 gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the ARO8 gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the ARO8 gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the ARO8 gene within the engineered host cell. The ARO8 gene may be derived from *Saccharomyces cerevisiae* or another species. In some examples, the ARO8 gene may be 100% similar to the naturally occurring gene.

[00313] [TYR1] In some examples, the engineered host cell may modify the expression of the enzyme prephenate dehydrogenase. Prephenate dehydrogenase is encoded by the TYR1 gene. In examples, prephenate dehydrogenase catalyzes the reaction of prephenate + $\text{NADP}^+ \rightarrow$ 4-hydroxyphenylpyruvate + CO_2 + NADPH, as referenced in **FIG. 2**. An engineered host cell may be modified to include constitutive overexpression of the TYR1 gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the TYR1 gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the TYR1 gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the TYR1 gene within the engineered host cell. The TYR1 gene may be derived from *Saccharomyces*

cerevisiae or another species. In some examples, the TYR1 gene may be 100% similar to the naturally occurring gene.

[00314] [TYR] In some examples, the engineered host cell may modify the expression of the enzyme tyrosinase. Tyrosinase is encoded by the TYR gene. In examples, tyrosinase catalyzes the reaction of tyrosine $\rightarrow$ L-DOPA, as referenced in **FIGs. 2, 12, and 13**. In other examples, tyrosinase catalyzes the reaction of L-DOPA $\rightarrow$ dopaquinone. An engineered host cell may be modified to include constitutive expression of the TYR gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the TYR gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the TYR gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the TYR gene within the engineered host cell. The TYR gene may be derived from *Ralstonia solanacearum*, *Agaricus bisporus*, or another species. In some examples, the TYR gene may be 100% similar to the naturally occurring gene.

[00315] [TyrH] In some examples, the engineered host cell may modify the expression of the enzyme tyrosine hydroxylase. Tyrosine hydroxylase is encoded by the TyrH gene. In examples, tyrosine hydroxylase catalyzes the reaction of tyrosine $\rightarrow$ L-DOPA, as referenced in **FIGs. 2, 12, and 13**. An engineered host cell may be modified to include constitutive expression of the TyrH gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the TyrH gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the TyrH gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the TyrH gene within the engineered host cell. The TyrH gene may be derived from *Homo sapiens*, *Rattus norvegicus*, *Mus musculus*, or another species. In some examples, the TyrH gene may be 100% similar to the naturally occurring gene.

[00316] [DODC] In some examples, the engineered host cell may modify the expression of the enzyme L-DOPA decarboxylase. L-DOPA decarboxylase is encoded by the DODC gene. In examples, L-DOPA decarboxylase catalyzes the reaction of L-DOPA $\rightarrow$ dopamine, as referenced in **FIGs. 2, 12, and 13**. An engineered host cell may be modified to include constitutive expression of the DODC gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the DODC gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the DODC gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the DODC gene within the engineered host cell. The DODC gene may be derived from *Pseudomonas putida*, *Rattus norvegicus*, or another species. In some examples, the DODC gene may be 100% similar to the naturally occurring gene.

[00317] [TYDC] In some examples, the engineered host cell may modify the expression of the enzyme tyrosine/DOPA decarboxylase. Tyrosine/DOPA decarboxylase is encoded by the TYDC gene. In examples, tyrosine/DOPA decarboxylase catalyzes the reaction of L-DOPA $\rightarrow$ dopamine, as referenced in **FIGs. 2, 12, and 13**. An engineered host cell may be modified to include constitutive expression of the TYDC gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the TYDC gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the TYDC gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the TYDC gene within the engineered host cell. The TYDC gene may be derived from *Papaver somniferum* or another species. In some examples, the TYDC gene may be 100% similar to the naturally occurring gene.

[00318] [MAO] In some examples, the engineered host cell may modify the expression of the enzyme monoamine oxidase. Monoamine oxidase is encoded by the MAO gene. In examples, monoamine oxidase catalyzes the reaction of dopamine $\rightarrow$ 3,4-DHPA, as referenced in **FIGs. 2 and 13**. An engineered host cell may be modified to include constitutive expression of the MAO gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the MAO gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the MAO gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the MAO gene within the engineered host cell. In some cases, the MAO gene may be codon optimized for expression in *Saccharomyces cerevisiae*. The MAO gene may be derived from *Escherichia coli*, *Homo sapiens*, *Micrococcus luteus*, or another species. In some examples, the MAO gene may be 77% similar to the naturally occurring gene.

[00319] [NCS] In some examples, the engineered host cell may modify the expression of the enzyme norcoclaurine synthase. Norcoclaurine synthase is encoded by the NCS gene. In examples, norcoclaurine synthase catalyzes the reaction of 4HPA + dopamine $\rightarrow$ (S)-norcoclaurine, as referenced in **FIGs. 12 and 13**. In particular, **FIG. 12** illustrates a biosynthetic scheme for conversion of L-tyrosine to reticuline via norcoclaurine, in accordance with embodiments of the invention. **FIG. 12** provides the use of the enzymes TyrH, tyrosine hydroxylase; DODC, DOPA decarboxylase; NCS, norcoclaurine synthase, as discussed herein; 6OMT, 6-O-methyltransferase; CNMT, coclaurine N-methyltransferase; CYP80B1, cytochrome P450 80B1; CPR, cytochrome P450 NADPH reductase; 4'OMT, 3'hydroxy-N-methylcoclaurine 4'-O-methyltransferase. L-DOPA, L-3,4-dihydroxyphenylalanine; and 4-HPA, 4-hydroxyphenylacetaldehyde. Of the enzymes that are illustrated in **FIG. 12**, 4-HPA and L-tyrosine are naturally synthesized in yeast. All other listed metabolites are not naturally produced in yeast. Additionally, although TyrH may catalyze the conversion of L-tyrosine to L-DOPA, other enzymes may also be used to perform this step as described in the specification. For example, tyrosinases may also be used to perform the conversion of L-tyrosine to L-DOPA. In addition, other enzymes such as cytochrome

P450 oxidases may also be used to perform the conversion of L-tyrosine to L-DOPA. Such enzymes may exhibit oxidase activity on related BIA precursor compounds including L-DOPA and L-tyrosine.

[00320] Additionally, norcoclaurine synthase catalyzes the reaction of 3,4-DHPA + dopamine $\rightarrow$ (S)-norlaudanosoline, as referenced in **FIG. 13**. In particular, **FIG. 13** illustrates a biosynthetic scheme for conversion of L-tyrosine to reticuline via norlaudanosoline, in accordance with embodiments of the invention. **FIG. 13** provides the use of the enzymes TyrH, tyrosine hydroxylase; DODC, DOPA decarboxylase; maoA, monoamine oxidase; NCS, norcoclaurine synthase; 6OMT, 6-O-methyltransferase; CNMT, coclaurine N-methyltransferase; 4'OMT, 3'-hydroxy-N-methylcoclaurine 4'-O-methyltransferase. L-DOPA, L-3,4-dihydroxyphenylalanine; and 3,4-DHPA, 3,4-dihydroxyphenylacetaldehyde. Of the enzymes that are illustrated in **FIG. 13**, L-tyrosine is naturally synthesized in yeast. Other metabolites that are shown in **FIG. 13** are not naturally produced in yeast.

[00321] An engineered host cell may be modified to include constitutive expression of the NCS gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the NCS gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the NCS gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the NCS gene within the engineered host cell. Additionally, the norcoclaurine synthase may have an N-terminal truncation. In some cases, the NCS gene may be codon optimized for expression in *Saccharomyces cerevisiae*. The NCS gene may be derived from *Coptis japonica*, *Papaver somniferum*, *Papver bracteatum*, *Thalictrum flavum*, *Corydalis saxicola*, or another species. In some examples, the NCS gene may be 80% similar to the naturally occurring gene.

[00322] [6OMT] In some examples, the engineered host cell may modify the expression of the enzyme norcoclaurine 6-O-methyltransferase. Norcoclaurine 6-O-methyltransferase is encoded by the 6OMT gene. In some examples, norcoclaurine 6-O-methyltransferase catalyzes the reaction of norcoclaurine $\rightarrow$ coclaurine, as referenced in **FIG. 12**. In other examples, norcoclaurine 6-O-methyltransferase catalyzes the reaction of norlaudanosoline $\rightarrow$ 3'-hydroxycoclaurine, as well as other reactions detailed herein, such as those provided in **FIG. 13**. Additionally, the engineered host cell may be modified to include constitutive expression of the 6OMT gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the 6OMT gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the 6OMT gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the 6OMT gene within the engineered host cell. The 6OMT gene may be derived from *P. somniferum*, *T. flavum*, *Coptis japonica*, or another species. In some examples, the 6OMT gene may be 100% similar to the naturally occurring gene.

[00323] [CNMT] In some examples, the engineered host cell may modify the expression of the enzyme coclaurine-*N*-methyltransferase. Coclaurine-*N*-methyltransferase is encoded by the CNMT gene. In some examples, coclaurine-*N*-methyltransferase catalyzes the reaction of coclaurine $\rightarrow$ *N*-methylcoclaurine, as referenced in **FIG. 12**. In other examples, the coclaurine-*N*-methyltransferase enzyme may catalyze the reaction of 3'-hydroxycoclaurine $\rightarrow$ 3'-hydroxy-*N*-methylcoclaurine. In other examples, coclaurine-*N*-methyltransferase may catalyze other reactions detailed herein, such as those provided in **FIG. 13**. Additionally, the engineered host cell may be modified to include constitutive expression of the CNMT gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the CNMT gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the CNMT gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the CNMT gene within the engineered host cell. The CNMT gene may be derived from *P. somniferum*, *T. flavum*, *Coptis japonica*, or another species. In some examples, the CNMT gene may be 100% similar to the naturally occurring gene.

[00324] [4'OMT] In some examples, the engineered host cell may modify the expression of the enzyme 4'-*O*-methyltransferase. 4'-*O*-methyltransferase is encoded by the 4'OMT gene. In some examples, 4'-*O*-methyltransferase catalyzes the reaction of 3'-hydroxy-*N*-methylcoclaurine $\rightarrow$ reticuline, as referenced in **FIG. 12**. In other examples, 4'-*O*-methyltransferase catalyzes other reactions detailed herein, such as those provided in **FIG. 13**. Additionally, the engineered host cell may be modified to include constitutive expression of the 4'OMT gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the 4'OMT gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the 4'OMT gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the 4'OMT gene within the engineered host cell. The 4'OMT gene may be derived from *P. somniferum*, *T. flavum*, *Coptis japonica*, or another species. In some examples, the 4'OMT gene may be 100% similar to the naturally occurring gene.

[00325] [CYP80B1] In some examples, the engineered host cell may modify the expression of the enzyme cytochrome P450 80B1. Cytochrome P450 80B1 is encoded by the CYP80B1 gene. In examples, cytochrome P450 80B1 catalyzes the reaction of *N*-methylcoclaurine $\rightarrow$ 3'-hydroxy-*N*-methylcoclaurine, as referenced in **FIG. 12**. An engineered host cell may be modified to include constitutive expression of the cytochrome P450 80B1 gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the cytochrome P450 80B1 gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the cytochrome P450 80B1 gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of

a strong promoter element for the overexpression of the cytochrome P450 80B1 gene within the engineered host cell. In some cases, the CYP80B1 gene may be codon optimized for expression in *Saccharomyces cerevisiae*. The cytochrome P450 80B1 gene may be derived from *P. somniferum*, *E. californica*, *T. flavum*, or another species. In some examples, the P450 80B1 gene may be 77% similar to the naturally occurring gene.

[00326] [FOL2] In some examples, the engineered host cell may modify the expression of the enzyme GTP cyclohydrolase. GTP cyclohydrolase is encoded by the FOL2 gene. In some examples, GTP cyclohydrolase catalyzes the reaction of $\text{GTP} \rightarrow \text{dihydroneopterin triphosphate}$, as referenced in **FIG. 1**. The engineered host cell may be modified to include constitutive overexpression of the FOL2 gene in the engineered host cell. The engineered host cell may also be modified to include native regulation. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the FOL2 gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the FOL2 gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the FOL2 gene within the engineered host cell. The FOL2 gene may be derived from *Saccharomyces cerevisiae*, *Homo sapiens*, *Mus musculus*, or another species. In some examples, the FOL2 gene may be 100% similar to the naturally occurring gene.

[00327] [PTPS] In some examples, the engineered host cell may modify the expression of the enzyme 6-pyruvoyl tetrahydrobiopterin (PTP) synthase. Pyruvoyl tetrahydrobiopterin synthase is encoded by the PTPS gene. In some examples, 6-pyruvoyl tetrahydrobiopterin synthase catalyzes the reaction of $\text{dihydroneopterin triphosphate} \rightarrow \text{PTP}$, as referenced in **FIG. 1**. The engineered host cell may be modified to include constitutive expression of the PTPS gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the PTPS gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the PTPS gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the PTPS gene within the engineered host cell. In some cases, the PTPS gene may be codon optimized for expression in *Saccharomyces cerevisiae*. The PTPS gene may be derived from *Rattus norvegicus*, *Homo sapiens*, *Mus musculus*, or another species. In some examples, the PTPS gene may be 80% similar to the naturally occurring gene.

[00328] [SepR] In some examples, the engineered host cell may modify the expression of the enzyme sepiapterin reductase. Sepiapterin reductase is encoded by the SepR gene. In some examples, sepiapterin reductase catalyzes the reaction of $\text{PTP} \rightarrow \text{BH}_4$, as referenced in **FIG. 1**. The engineered host cell may be modified to include constitutive expression of the SepR gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the SepR gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the SepR gene. Additionally or alternatively, the

engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the SepR gene within the engineered host cell. In some cases, the SepR gene may be codon optimized for expression in *Saccharomyces cerevisiae*. The SepR gene may be derived from *Rattus norvegicus*, *Homo sapiens*, *Mus musculus*, or another species. In some examples, the SepR gene may be 72% similar to the naturally occurring gene.

[00329] [PCD] In some examples, the engineered host cell may modify the expression of the enzyme 4a-hydroxytetrahydrobiopterin (pterin-4 α -carbinolamine) dehydratase. 4a-hydroxytetrahydrobiopterin dehydratase is encoded by the PCD gene. In some examples, 4a-hydroxytetrahydrobiopterin dehydratase catalyzes the reaction of 4a-hydroxytetrahydrobiopterin $\rightarrow$ H₂O + quinonoid dihydropteridine, as referenced in **FIG. 1**. The engineered host cell may be modified to include constitutive expression of the PCD gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the PCD gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the PCD gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the PCD gene within the engineered host cell. In some cases, the PCD gene may be codon optimized for expression in *Saccharomyces cerevisiae*. The PCD gene may be derived from *Rattus norvegicus*, *Homo sapiens*, *Mus musculus*, or another species. In some examples, the PCD gene may be 79% similar to the naturally occurring gene.

[00330] [QDHPR] In some examples, the engineered host cell may modify the expression of the enzyme quinonoid dihydropteridine reductase. Quinonoid dihydropteridine reductase is encoded by the QDHPR gene. In some examples, quinonoid dihydropteridine reductase catalyzes the reaction of quinonoid dihydropteridine $\rightarrow$ BH₄, as referenced in **FIG. 1**. The engineered host cell may be modified to include constitutive expression of the QDHPR gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the QDHPR gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the QDHPR gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the QDHPR gene within the engineered host cell. In some cases, the QDHPR gene may be codon optimized for expression in *Saccharomyces cerevisiae*. The QDHPR gene may be derived from *Rattus norvegicus*, *Homo sapiens*, *Mus musculus*, or another species. In some examples, the QDHPR gene may be 75% similar to the naturally occurring gene.

[00331] [DHFR] In some examples, the engineered host cell may modify the expression of the enzyme dihydrofolate reductase. Dihydrofolate reductase is encoded by the DHFR gene. In some examples, dihydrofolate reductase catalyzes the reaction of 7,8-dihydrobiopterin (BH₂) $\rightarrow$ 5,6,7,8-tetrahydrobiopterin (BH₄), as referenced in **FIG. 1**. This reaction may be useful in recovering BH₄ as a co-substrate for the conversion of tyrosine to L-DOPA, as illustrated in **FIG. 12**. The engineered host cell may be modified to include constitutive expression of the DHFR gene in the engineered host cell.

Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the DHFR gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the DHFR gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the DHFR gene within the engineered host cell. In some cases, the DHFR gene may be codon optimized for expression in *Saccharomyces cerevisiae*. The DHFR gene may be derived from *Rattus norvegicus*, *Homo sapiens*, or another species. In some examples, the DHFR gene may be 77% similar to the naturally occurring gene.

[00332] [DRS-DRR] As discussed above with regard to epimerizing 1-BIAs, the engineered host cell may modify the expression of a BIA epimerase. The BIA epimerase is encoded by the DRS-DRR gene. In some examples, DRS-DRR may also be referred to as CYP-COR. In some examples, an engineered split version, or an engineered fused version, of a BIA epimerase catalyzes the conversion of (*S*)-1-BIA → (*R*)-1-BIA, as referenced in **FIG. 14**. In particular, **FIG. 14** illustrates a biosynthetic scheme for conversion of L-tyrosine to morphinan alkaloids, in accordance with embodiments of the invention. **FIG. 14** provides the use of the enzymes CPR, cytochrome P450 reductase; DRS-DRR, dehydreticuline synthase and dehydreticuline reductase; SalSyn, salutaridine synthase; SalR, salutaridine reductase; SalAT, salutaridinol 7-*O*-acetyltransferase; T6ODM, thebaine 6-*O*-demethylase; COR, codeinone reductase; and CODM, codeine-*O*-demethylase.

[00333] The engineered host cell may be modified to include constitutive expression of the engineered DRS-DRR gene in the engineered host cell. In some cases, the engineered DRS-DRR gene may encode an engineered fusion epimerase. In some cases, the engineered DRS-DRR gene may encode an engineered split epimerase. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the DRS-DRR gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the DRS-DRR gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the DRS-DRR gene within the engineered host cell. The DRS-DRR gene may be derived from *Papaver bracteatum*, *Papaver somniferum*, *Papaver setigerum*, *Chelidonium majus*, or another species. In some examples, the DRS-DRR gene may be 77% similar to the naturally occurring gene.

[00334] [CPR] In some examples, the engineered host cell may modify the expression of the enzyme cytochrome P450 reductase. The cytochrome P450 reductase is encoded by the CPR gene. In some examples, the cytochrome P450 reductase catalyzes the reaction of (*R*)-reticuline → salutaridine, as referenced in **FIG. 14**. Additionally, the cytochrome P450 reductase catalyzes other reactions such as those described in FIGs. throughout the application. The engineered host cell may be modified to include constitutive expression of the CPR gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the CPR gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy,

copies, or additional copies, of the CPR gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the CPR gene within the engineered host cell. The CPR gene may be derived from *E. californica*, *P. somniferum*, *H. sapiens*, *S. cerevisiae*, *A. thaliana*, or another species. In some examples, the CPR gene may be 100% similar to the naturally occurring gene.

[00335] [SalSyn] In some examples, the engineered host cell may modify the expression of the enzyme salutaridine synthase. The salutaridine synthase is encoded by the SalSyn gene. In some examples, the salutaridine synthase catalyzes the reaction of (*R*)-reticuline → salutaridine, as referenced in **FIG. 14**. The engineered host cell may be modified to include constitutive expression of the SalSyn gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the SalSyn gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the SalSyn gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the SalSyn gene within the engineered host cell. In some cases, the SalSyn gene may be codon optimized for expression in *Saccharomyces cerevisiae*. In some examples the SalSyn may be modified at the N-terminus. The SalSyn gene may be derived from *Papaver somniferum*, *Papaver spp.*, *Chelidonium majus*, or another species. In some examples, the SalSyn gene may be 78% similar to the naturally occurring gene.

[00336] [SalR] In some examples, the engineered host cell may modify the expression of the enzyme salutaridine reductase. Salutaridine reductase is encoded by the SalR gene. In some examples, salutaridine reductase reversibly catalyzes the reaction of salutaridinol → salutaridine, as referenced in **FIG. 14**. The engineered host cell may be modified to include constitutive expression of the SalR gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the SalR gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the SalR gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the SalR gene within the engineered host cell. In some cases, the SalR gene may be codon optimized for expression in *Saccharomyces cerevisiae*. The SalR gene may be derived from *Papaver somniferum*, *Papaver bracteatum*, *Papaver spp.*, *Chelidonium majus*, or another species. In some examples, the SalR gene may be 80-100% similar to the naturally occurring gene.

[00337] [SalAT] In some examples, the engineered host cell may modify the expression of the enzyme acetyl-CoA:salutaridinol 7-*O*-acetyltransferase. Acetyl-CoA:salutaridinol 7-*O*-acetyltransferase is encoded by the SalAT gene. In some examples, acetyl-CoA:salutaridinol 7-*O*-acetyltransferase catalyzes the reaction of acetyl-CoA + salutaridinol → CoA + 7-*O*-acetylsalutaridinol, as referenced in **FIG. 14**. The engineered host cell may be modified to include constitutive expression of the SalAT gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to

synthetically regulate the expression of the SalAT gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the SalAT gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the SalAT gene within the engineered host cell. In some cases, the SalAT gene may be codon optimized for expression in *Saccharomyces cerevisiae*. The SalAT gene may be derived from *Papaver somniferum*, *Papaver bracteatum*, *Papaver orientale*, *Papaver spp.*, or another species. In some examples, the SalAT gene may be 77-80% similar to the naturally occurring gene.

[00338] [TS] In some examples, the engineered host cell may modify the expression of the enzyme thebaine synthase. Thebaine synthase is encoded by the TS gene. In some examples, a thebaine synthase or an engineered thebaine synthase catalyzes the reaction of 7-*O*-acetyl salutaridinol → thebaine + acetate, as referenced in **FIG. 14**. In some examples, the reaction of 7-*O*-acetyl salutaridinol → thebaine + acetate occurs spontaneously, but thebaine synthase catalyzes some portion of this reaction. In particular, **FIG. 14** illustrates a biosynthetic scheme for conversion of L-tyrosine to morphinan alkaloids, in accordance with embodiments of the invention. **FIG. 14** provides the use of the enzymes CPR, cytochrome P450 reductase; DRS-DRR, dehydroreticuline synthase and dehydroreticuline reductase; SalSyn, salutaridine synthase; SalR, salutaridine reductase; SalAT, salutaridinol 7-*O*-acetyltransferase; TS, thebaine synthase; T6ODM, thebaine 6-*O*-demethylase; COR, codeinone reductase; and CODM, codeine-*O*-demethylase.

[00339] The engineered host cell may be modified to include constitutive expression of the TS gene or the engineering TS gene in the engineered host cell. In some cases, the engineered TS gene may encode an engineered fusion enzyme. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the TS gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the TS gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the TS gene within the engineered host cell. In some cases, the TS gene may be codon optimized for expression in *Saccharomyces cerevisiae*. The TS gene may be derived from *Papaver somniferum*, *Papaver bracteatum*, *Papaver orientale*, *Papaver spp.*, or another species. In some examples, the TS gene may be 75-80% similar to the naturally occurring gene.

[00340] [T6ODM] In some examples, the engineered host cell may modify the expression of the enzyme thebaine 6-*O*-demethylase. Thebaine 6-*O* demethylase is encoded by the T6ODM gene. In some examples, thebaine 6-*O*-demethylase catalyzes the reaction of thebaine → neopinone, as referenced in **FIG. 14**. Once the neopinone has been produced, the neopinone may be converted to codeinone. The conversion of neopinone → codeinone may occur spontaneously. Alternatively, the conversion of neopinone → codeinone may occur as a result of a catalyzed reaction. In other examples, the T6ODM enzyme may catalyze the *O*-demethylation of substrates other than thebaine. For example, T6ODM may *O*-demethylate oripavine to produce morphinone. Alternatively, T6ODM may catalyze the *O*-demethylation of BIAs within the 1-benzylisoquinoline, protoberberine, or protopine classes such as

papaverine, canadine, and allocryptopine, respectively. The engineered host cell may be modified to include constitutive expression of the T6ODM gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the T6ODM gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the T6ODM gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the T6ODM gene within the engineered host cell. In some cases, the T6ODM gene may be codon optimized for expression in *Saccharomyces cerevisiae*. The T6ODM gene may be derived from *Papaver somniferum*, or another species. In some examples, the T6ODM gene may be 76.2% similar to the naturally occurring gene.

[00341] [COR] In some examples, the engineered host cell may modify the expression of the enzyme codeinone reductase. Codeinone reductase is encoded by the COR gene. In some examples, codeinone reductase catalyzes the reaction of codeinone to codeine, as referenced in **FIG. 14**. In some cases, codeinone reductase can catalyze the reaction of neopinone to neopine. In other examples, COR can catalyze the reduction of other morphinans including hydrocodone → dihydrocodeine, 14-hydroxycodeinone → 14-hydroxycodeine, and hydromorphone → dihydromorphone. The engineered host cell may be modified to include constitutive expression of the COR gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the COR gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the COR gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the COR gene within the engineered host cell. In some cases, the COR gene may be codon optimized for expression in *Saccharomyces cerevisiae*. Additionally or alternatively, the COR gene may be modified with the addition of targeting sequences for mitochondria, vacuole, endoplasmic reticulum, or a combination thereof. The COR gene may be derived from *Papaver somniferum*, or another species. In some examples, the COR gene may be 76-78% similar to the naturally occurring gene. In examples, the COR gene may be 76.8%, 77.0%, 77.3%, or 77.7% similar to the naturally occurring gene.

[00342] [CODM] In some examples, the engineered host cell may modify the expression of the enzyme codeine *O*-demethylase. Codeine *O*-demethylase is encoded by the CODM gene. In some examples, codeine *O*-demethylase catalyzes the reaction of codeine to morphine, as referenced in **FIG. 14**. Codeine *O*-demethylase can also catalyze the reaction of neopine to neomorphine. Codeine *O*-demethylase can also catalyze the reaction of thebaine to oripavine. In other examples, CODM may catalyze the *O*-demethylation of BIAs within the 1-benzylisoquinoline, aporphine, and protoberberine classes such as reticuline, isocorydine, and scoulerine, respectively. In other examples, the CODM enzyme may catalyze an *O,O*-demethylation reaction to cleave the methylenedioxy bridge structures in protopines. The engineered host cell may be modified to include constitutive expression of the CODM gene in the

engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the CODM gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the CODM gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the CODM gene within the engineered host cell. In some cases, the CODM gene may be codon optimized for expression in *Saccharomyces cerevisiae*. Additionally or alternatively, the CODM gene may be modified with the addition of targeting sequences for mitochondria. The CODM gene may be derived from *Papaver somniferum*, *Papaver spp.*, or another species. In some examples, the CODM gene may be 75% similar to the naturally occurring gene. In examples, the CODM gene may be 75.2% similar to the naturally occurring gene.

[00343] [BBE] In some examples, the engineered host cell may modify the expression of the enzyme berberine bridge enzyme. The berberine bridge enzyme is encoded by the BBE gene. In some examples, berberine bridge enzyme catalyzes the reaction of (*S*)-reticuline $\rightarrow$ (*S*)-scoulerine. The engineered host cell may be modified to include constitutive expression of the BBE gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the BBE gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the BBE gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the BBE gene within the engineered host cell. The BBE gene may be derived from *Papaver somniferum*, *Argemone mexicana*, *Eschscholzia californica*, *Berberis stolonifera*, *Thalictrum flavum subsp. glaucum*, *Coptis japonica*, *Papaver spp.*, or another species. In some examples, the BBE gene may be 99% similar to the naturally occurring gene.

[00344] [CYP2D6] In some examples, the engineered host cell may modify the expression of cytochrome P450, family 2, subfamily D, polypeptide 6. This particular cytochrome P450 is encoded by the CYP2D6 gene. This particular cytochrome P450 enzyme may be characterized as a promiscuous oxidase. In some examples, this particular cytochrome P450 enzyme may catalyze the reaction of (*R*)-reticuline + NADPH + H⁺ + O₂ $\rightarrow$ salutaridine + NADP⁺ + 2 H₂O, among other reactions.

[00345] [S9OMT] In some examples, the engineered host cell may modify the expression of the enzyme *S*-adenosyl-L-methionine:(*S*)-scoulerine 9-*O*-methyltransferase. *S*-adenosyl-L-methionine:(*S*)-scoulerine 9-*O*-methyltransferase is encoded by the S9OMT gene. In some examples, *S*-adenosyl-L-methionine:(*S*)-scoulerine 9-*O*-methyltransferase catalyzes the reaction of *S*-adenosyl-L-methionine + (*S*)-scoulerine $\rightarrow$ *S*-adenosyl-L-homocysteine + (*S*)-tetrahydrocolumbamine. The engineered host cell may be modified to include constitutive expression of the S9OMT gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the S9OMT gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the S9OMT gene. Additionally or alternatively, the

engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the S9OMT gene within the engineered host cell. In some cases, the S9OMT gene may be codon optimized for expression in *Saccharomyces cerevisiae*. The S9OMT gene may be derived from *Thalictrum flavum subsp. glaucum*, *Coptis japonica*, *Coptis chinensis*, *Papaver somniferum*, *Thalictrum spp.*, *Coptis spp.*, *Papaver spp.*, or another species. In some examples, the S9OMT gene may be 100% similar to the naturally occurring gene. In examples, the S9OMT gene may be 80% similar to the naturally occurring gene.

[00346] [CAS] In some examples, the engineered host cell may modify the expression of the enzyme (*S*)-canadine synthase. (*S*)-canadine synthase is encoded by the CAS gene. In some examples, (*S*)-canadine synthase catalyzes the reaction of (*S*)-tetrahydrocolumbamine $\rightarrow$ (*S*)-canadine. The engineered host cell may be modified to express the CAS gene in the engineered host cell. The engineered host cell may be modified to include constitutive expression of the CAS gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the CAS gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the CAS gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the CAS gene within the engineered host cell. The CAS gene may be derived from *Thalictrum flavum subsp. glaucum*, *Coptis japonica*, *Thalictrum spp.*, *Coptis spp.*, or another species. In some examples, the CAS gene may be 100% similar to the naturally occurring gene.

[00347] [STOX] In some examples, the engineered host cell may modify the expression of the enzyme (*S*)-tetrahydroprotoberberine oxidase. (*S*)-tetrahydroprotoberberine oxidase is encoded by the STOX gene. In some examples, (*S*)-tetrahydroprotoberberine oxidase catalyzes the reaction of (*S*)-tetrahydroberberine + 2 O₂ $\rightarrow$ berberine + 2 H₂O₂. The engineered host cell may be modified to include constitutive expression of the STOX gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the STOX gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the STOX gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the STOX gene within the engineered host cell. In some examples the STOX may be modified at the N-terminus. In some cases, the STOX gene may be codon optimized for expression in *Saccharomyces cerevisiae*. The STOX gene may be derived from *Berberis wilsonae*, *Coptis japonica*, *Berberis spp.*, *Coptis spp.*, or another species. In some examples, the STOX gene may be 78% similar to the naturally occurring gene.

[00348] [TNMT] In some examples, the engineered host cell may modify the expression of the enzyme tetrahydroprotoberberine-*N*-methyltransferase. Tetrahydroprotoberberine-*N*-methyltransferase is encoded by the TNMT gene. In some examples, tetrahydroprotoberberine-*N*-methyltransferase catalyzes the reaction of canadine $\rightarrow$ *N*-methylcanadine.

[00349] In other examples, tetrahydropprotoberberine-*N*-methyltransferase catalyzes the reaction of stylophine → *cis-N*-methylstylophine. The engineered host cell may be modified to include constitutive expression of the TNMT gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the TNMT gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the TNMT gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the TNMT gene within the engineered host cell. In some cases, the TNMT gene may be codon optimized for expression in *Saccharomyces cerevisiae*. The TNMT gene may be derived from *Papaver somniferum*, *Eschscholzia californica*, *Papaver bracteatum*, *Argemone mexicana*, or another species. In some examples, the TNMT gene may be 100% similar to the naturally occurring gene. In examples, the TNMT gene may be 81% similar to the naturally occurring gene.

[00350] [CFS] In some examples, the engineered host cell may modify the expression of the enzyme cheilanthifoline synthase. Cheilanthifoline synthase is encoded by the CFS gene. In examples, cheilanthifoline synthase catalyzes the reaction of scoulerine → cheilanthifoline. An engineered host cell may be modified to include constitutive expression of the CFS gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the CFS gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the CFS gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the CFS gene within the engineered host cell. The CFS gene may be derived from *P. somniferum*, *E. californica*, *A. mexicana*, or another species. In some examples, the CFS gene may be 77%, 78%, or 79% similar to the naturally occurring gene. Additionally, the CFS gene may be codon optimized for expression in *Saccharomyces cerevisiae*.

[00351] In some examples, the engineered host cell may modify the expression of the enzyme stylophine synthase. Stylophine synthase is encoded by the STS gene. In examples, stylophine synthase catalyzes the reaction of cheilanthifoline → stylophine. An engineered host cell may be modified to include constitutive expression of the STS gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the STS gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the STS gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the STS gene within the engineered host cell. The STS gene may be derived from *P. somniferum*, *E. californica*, *A. mexicana*, or another species. In some examples, the STS gene may be 76%, 78%, or 79% similar to the naturally occurring gene. Additionally, the STS gene may be codon optimized for expression in *Saccharomyces cerevisiae*.

[00352] [STS] In some examples, the engineered host cell may modify the expression of the enzyme stylophine synthase. Stylophine synthase is encoded by the STS gene. In examples, stylophine synthase catalyzes the reaction of cheilanthifoline → stylophine, among other reactions. An engineered host cell may be modified to include constitutive expression of the STS gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the STS gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the STS gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promotor element for the overexpression of the STS gene within the engineered host cell. The STS gene may be derived from *P. somniferum*, *E. californica*, *A. mexicana*, or another species. In some examples, the STS gene may be 76%, 78%, or 79% similar to the naturally occurring gene. Additionally, the STS gene may be codon optimized for expression in *Saccharomyces cerevisiae*.

[00353] [MSH] In some examples, the engineered host cell may modify the expression of the enzyme cis-*N*-methylstylophine 14-hydroxylase. Cis-*N*-methylstylophine 14-hydroxylase is encoded by the MSH gene. In examples, cis-*N*-methylstylophine 14-hydroxylase catalyzes the reaction of cis-*N*-methylstylophine → protopine. An engineered host cell may be modified to include constitutive expression of the MSH gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the MSH gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the MSH gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promotor element for the overexpression of the MSH gene within the engineered host cell. The MSH gene may be derived from *P. somniferum* or another species. In some examples, the MSH gene may be 79% similar to the naturally occurring gene. Additionally, the MSH gene may be codon optimized for expression in *Saccharomyces cerevisiae*.

[00354] [P6H] In some examples, the engineered host cell may modify the expression of the enzyme protopine-6-hydroxylase. Protopine-6-hydroxylase is encoded by the P6H gene. In examples, protopine-6-hydroxylase catalyzes the reaction of Protopine → 6-hydroxyprotopine. An engineered host cell may be modified to include constitutive expression of the P6H gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the P6H gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the P6H gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promotor element for the overexpression of the CFS gene within the engineered host cell. The P6H gene may be derived from *P. somniferum*, *E. californica*, or another species. In some examples, the P6H gene may be 79% similar to the naturally occurring gene. Additionally, the P6H gene may be codon optimized for expression in *Saccharomyces cerevisiae*.

[00355] [DBOX] In some examples, the engineered host cell may modify the expression of the enzyme dihydrobenzophenanthridine oxidase. Dihydrobenzophenanthridine oxidase is encoded by the DBOX gene. In examples, dihydrobenzophenanthridine oxidase catalyzes the reaction of dihydrosanguinarine → sanguinarine. An engineered host cell may be modified to include constitutive expression of the DBOX gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the DBOX gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the DBOX gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promotor element for the overexpression of the DBOX gene within the engineered host cell. The DBOX gene may be derived from *P. somniferum* or another species. In some examples, the DBOX gene may be 100% similar to the naturally occurring gene. Additionally, the DBOX gene may be codon optimized for expression in *Saccharomyces cerevisiae*.

[00356] [morA] In some examples, the engineered host cell may modify the expression of the enzyme morphine dehydrogenase. Morphine dehydrogenase is encoded by the morA gene. In some examples, morphine dehydrogenase catalyzes the reaction of morphine → morphinone, as referenced in **FIG. 15**. In other examples, morphine dehydrogenase catalyzes the reaction of codeinone → codeine, also as referenced in **FIG. 15**. **FIG. 15** illustrates a biosynthetic scheme for production of semi-synthetic opioids, in accordance with embodiments of the invention. In particular, **FIG. 15** illustrates extended transformations of thebaine in yeast by incorporating morA, morphine dehydrogenase; and morB, morphine reductase.

[00357] The engineered host cell may be modified to include constitutive expression of the morA gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the morA gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the morA gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the morA gene within the engineered host cell. In some cases, the morA gene may be codon optimized for expression in *Saccharomyces cerevisiae*. The morA gene may be derived from *Pseudomonas putida* or another species. In some examples, the morA gene may be 73.7% similar to the naturally occurring gene.

[00358] [morB] In some examples, the engineered host cell may modify the expression of the enzyme morphinone reductase. Morphinone reductase is encoded by the morB gene. In some examples, morphinone reductase catalyzes the reaction of codeinone → hydrocodone, as referenced in **FIG. 15**. In other examples, morphinone reductase catalyzes the reaction of morphinone → hydromorphone, also as referenced in **FIG. 15**. In other examples, morphinone reductase catalyzes the reaction 14-hydroxycodeinone → oxycodone. The engineered host cell may be modified to include constitutive expression of the morB gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the morB gene in the engineered host

cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the *morB* gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the *morB* gene within the engineered host cell. In some cases, the *morB* gene may be codon optimized for expression in *Saccharomyces cerevisiae*. The *morB* gene may be derived from *Pseudomonas putida* or another species. In some examples, the *morB* gene may be 67.2% similar to the naturally occurring gene.

[00359] [CYP80A1] In some examples, the engineered host cell may express the enzyme berbaminine synthase. Berbaminine synthase is encoded by the gene for cytochrome P450 enzyme 80A1 (CYP80A1). In some examples, CYP80A1 catalyzes the reaction *(S)*-*N*-methylecclaurine + *(R)*-*N*-methylecclaurine → berbaminine, as referenced in **FIG. 17**. In other examples, CYP80A1 catalyzes the reaction *(R)*-*N*-methylecclaurine + *(R)*-*N*-methylecclaurine → guattegaumerine, as referenced in **FIG. 17**. In other examples, CYP80A1 catalyzes the reaction *(R)*-*N*-methylecclaurine + *(S)*-ecclaurine → 2'-norberbaminine. The engineered host cell may be modified to include constitutive expression of the CYP80A1 gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the CYP80A1 gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the CYP80A1 gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the CYP80A1 gene within the engineered host cell. In some cases, the CYP80A1 gene may be codon optimized for expression in *Saccharomyces cerevisiae*. The CYP80A1 gene may be derived from *Berberis stolonifera* or another species. In some examples, the CYP80A1 gene may be 76% similar to the naturally occurring gene.

[00360] [PODA] In some example, the engineered host cell may express the enzyme protopine *O*-dealkylase. Protopine *O*-dealkylase is encoded by the gene PODA. In some examples, PODA catalyzes the *O,O*-demethylation of protoberberines and protopines such as canadine, stylophine, berberine, cryptopine, allocryptopine, and protopine. In some examples, PODA catalyzes the *O*-demethylation of BIAs including tetrahydropapaverine, tetrahydropalmatine, and cryptopine. The engineered host cell may be modified to include constitutive expression of the PODA gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the PODA gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the PODA gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the PODA gene within the engineered host cell. In some cases, the PODA gene may be codon optimized for expression in *Saccharomyces cerevisiae*. The PODA gene may be derived from *Papaver somniferum* or other species. In some examples, the PODA gene may be 70-100% similar to the naturally occurring gene.

[00361] [RNMT] In some examples, the engineered host cell may modify the expression of the enzyme Reticuline N-methyltransferase. Reticuline N-methyltransferase is encoded by the RNMT gene. In some examples, Reticuline N-methyltransferase may catalyze reactions such as reticuline→tembetarine, among other reactions.

[00362] [P70MT] In some examples, the engineered host cell may modify the expression of the enzyme Papaverine 7-O-demethylase. Papaverine 7-O-demethylase is encoded by the P70MT gene. In some examples, Papaverine 7-O-demethylase may catalyze reactions such as papaverine→pacodine, among other reactions.

[00363] [3ODM] In some examples, the engineered host cell may modify the expression of the enzyme 3-O-demethylase. 3-O-demethylase is encoded by the 3ODM gene. In some examples, 3-O-demethylase may catalyze reactions such as oxycodone→oxymorphone; hydrocodone→hydromorphone; dihydrocodeine→dihydromorphone; 14-hydroxycodone→14-hydroxymorphone; codeinone→morphinone; and 14-hydroxycodeinone→14-hydroxymorphinone, among other reactions.

[00364] [NDM] In some examples, the engineered host cell may modify the expression of the enzyme *N*-demethylase. *N*-demethylase is encoded by the NDM gene. In some examples, *N*-demethylase may catalyze reactions, such as Codeine→Norcodeine; Morphine→Normorphine; Oxycodone→Noroxycodone; Oxymorphone→Noroxymorphone; Thebaine→Northebaine; Oripavine→Nororipavine; Hydrocodone→Norhydrocodone; Hydromorphone→Norhydromorphone; Dihydrocodeine→Nordihydrocodeine; Dihydromorphone→Nordihydromorphone; 14-hydroxycodone→Nor-14-hydroxycodone; 14-hydroxymorphone→Nor-14-hydroxymorphone; Codeinone→Norcodeinone; Morphinone→Normorphinone; 14-hydroxycodeinone→Nor-14-hydroxycodeinone; and 14-hydroxymorphinone→Nor-14-hydroxymorphinone, among other reactions.

[00365] [NMT] In some examples, the engineered host cell may modify the expression of the enzyme *N*-methyltransferase. *N*-methyltransferase is encoded by the NMT gene. In some examples, *N*-methyltransferase may catalyze reactions, such as Norcodeine→codeine; Normorphine→morphine; Noroxycodone→oxycodone; Noroxymorphone→noroxymorphone; Northebaine→thebaine; Nororipavine→oripavine; Norhydrocodone→hydrocodone; Norhydromorphone→Hydromorphone; Nordihydrocodeine→Dihydrocodeine; Nordihydromorphone→Dihydromorphone; Nor-14-hydroxycodone→14-hydroxycodone; Nor-14-hydroxymorphone→14-hydroxymorphone; Norcodeinone→Codeinone; Normorphinone→Morphinone; Nor-14-hydroxy-codeinone→14-hydroxycodeinone; Nor-14-hydroxy-morphinone→14-hydroxymorphinone.

[00366] [NAT] In some examples, the engineered host cell may modify the expression of the enzyme *N*-allyltransferase. *N*-allyltransferase is encoded by the NAT gene. In some examples, *N*-allyltransferase may catalyze reactions, such as Norcodeine→N-allyl-norcodeine; Normorphine→N-allyl-normorphine; Noroxycodone→N-allyl-noroxycodone; Noroxymorphone→N-allyl-nornoroxymorphone; Northebaine→N-allyl-northebaine; Nororipavine→N-allyl-nororipavine; Norhydrocodone→N-allyl-

norhydrocodone; Norhydromorphone→ N-allyl-norhydromorphone; Nordihydrocodeine→ N-allyl-nordihydrocodeine; Nordihydromorphone→ N-allyl-nordihydromorphone; Nor-14-hydroxycodone→ N-allyl-nor-14-hydroxycodone; Nor-14-hydroxymorphone→ N-allyl-nor-14-hydroxymorphone; Norcodeineone→ N-allyl-norcodeineone; Normorphinone→ N-allyl-normorphinone; Nor-14-hydroxycodeinone→ N-allyl-nor-14-hydroxycodeinone; Nor-14-hydroxy-morphinone→ N-allyl-nor-14-hydroxymorphinone, among other reactions.

[00367] [CPMT] In some examples, the engineered host cell may modify the expression of the enzyme *N*-cyclopropylmethyltransferase. *N*-cyclopropylmethyltransferase is encoded by the CPMT gene. In some examples, *N*-cyclopropylmethyltransferase may catalyze reactions, such as Norcodeine→N(cyclopropylmethyl)norcodeine; Normorphine→N(cyclopropylmethyl) normorphine; Noroxycodone→N(cyclopropylmethyl) noroxycodone; Noroxymorphone→N(cyclopropylmethyl) nornoroxymorphone; Northebaine→N(cyclopropylmethyl) northebaine; Nororipavine→N(cyclopropylmethyl) nororipavine; Norhydrocodone→N(cyclopropylmethyl) norhydrocodone; Norhydromorphone→ N(cyclopropylmethyl)norhydromorphone; Nordihydrocodeine→ N(cyclopropylmethyl)nordihydrocodeine; Nordihydromorphone→ N(cyclopropylmethyl)nordihydromorphone; Nor-14-hydroxycodone→ N(cyclopropylmethyl)nor-14-hydroxycodone; Nor-14-hydroxymorphone→ N(cyclopropylmethyl)nor-14-hydroxymorphone; Norcodeineone→ N(cyclopropylmethyl)norcodeineone; Normorphinone→ N(cyclopropylmethyl)normorphinone; Nor-14-hydroxy-codeinone→ N(cyclopropylmethyl)nor-14-hydroxycodeinone; and Nor-14-hydroxy-morphinone→ N(cyclopropylmethyl)nor-14-hydroxymorphinone, among other reactions.

[00368] [BM3] In some examples, the engineered host cell may express the enzyme BM3. BM3 is a *Bacillus megaterium* cytochrome P450 involved in fatty acid monooxygenation in its native host. In some cases BM3 *N*-demethylates an opioid to produce a nor-opioid, as referenced in **FIG. 9**. In some cases the host cell is modified to express BM3 in addition to other heterologous enzymes for the production of a nal-opioid or nor-opioid, as referenced in **FIG. 10**. The engineered host cell may be modified to include constitutive expression of the BM3 gene in the engineered host cell. Additionally or alternatively, the engineered host cell may be modified to synthetically regulate the expression of the BM3 gene in the engineered host cell. In examples, the engineered host cell may be modified to incorporate a copy, copies, or additional copies, of the BM3 gene. Additionally or alternatively, the engineered host cell may be modified to incorporate the introduction of a strong promoter element for the overexpression of the BM3 gene within the engineered host cell. BM3 has several advantages as a biosynthetic enzyme including that it is soluble, comes with a fused reductase partner protein, and can readily be engineered to accept new substrates. Additionally, **Table 7** illustrates variants of BM3 *N*-demethylase.

[00369] Examples of the aforementioned genes can be expressed from a number of different platforms in the host cell, including plasmid (2μ, ARS/CEN), YAC, or genome. In addition, examples of the

aforementioned gene sequences can either be native or codon optimized for expression in the desired heterologous host (e.g., *Saccharomyces cerevisiae*).

EXAMPLES

[00370] The following examples are given for the purpose of illustrating various embodiments of the invention and are not meant to limit the invention in any fashion. Where indicated, expression constructs are understood to incorporate a suitable promoter, gene, and terminator, even if the exact terminator sequence used is not specified. The present examples, along with the methods described herein are presently representative of preferred embodiments, are exemplary, and are not intended as limitations on the scope of the invention. Changes therein and other uses which are encompassed within the spirit of the invention as defined by the scope of the claims will occur to those skilled in the art.

Example 1. Intracellular activity of separately expressed DRS and DRR is substantially higher than the parental, fused enzyme.

[00371] The amino acid sequence of DRS-DRR from *P. bracteatum* (Farrow et al. 2015, GenBank: AKO60179.1) was codon-optimized (see, e.g., SEQ. ID 16) for expression in *S. cerevisiae* (codon optimization algorithm from IDT), and cloned into vectors with different promoter strengths to allow for protein engineering and enzyme improvement. The resulting plasmids, pDW10 and pDW18 (see FIG. 11), harbored DRS-DRR under control of the *S. cerevisiae* TDH3 promoter for strong expression and the CYC1 promoter for weak expression, respectively, and these were used as baseline constructs for future enzyme engineering efforts. The screening strain in this example was a derivative of YA106, an engineered *S. cerevisiae* Cen.PK strain expressing enzymes in the BIA pathway up to but not including DRS-DRR. Due to limitations in the high throughput differentiation of (*S*)- versus (*R*)-reticuline using LCMS methods, a strain expressing *P. bracteatum* SalSyn at high levels was constructed to screen for improved DRS-DRR enzymes based on the amount of salutaridine produced. SalSyn activity was empirically determined to be substantially higher than plasmid-based DRS-DRR activity, which allowed for increases in the conversion of (*S*)- to (*R*)-reticuline to be detected by the amount of salutaridine produced. Thus, improvements to DRS-DRR (or DRS alone) by protein engineering were measured by the relative concentration of salutaridine in the culture medium. To ensure that plasmids did not integrate onto the chromosome upon transformation into the screening strain, the TRP1 locus was knocked out by integration of URA3 to yield DW24, the reporter strain used in this and subsequent examples.

[00372] For propagation of yeast strains harboring engineered DRS-DRR (or separate DRS and DRR) enzymes, the reporter strain was transformed with expression plasmids using standard molecular biology techniques, and single colonies of yeast were isolated from solid agar medium plates under selective

conditions (such as synthetic complete 2% dextrose without tryptophan). Colonies were inoculated into liquid culture medium and grown for 2 days at 30°C. Cultures were then subcultured into fresh medium of the same composition, or in some cases into synthetic complete liquid medium containing 8% maltodextrin. To release monosaccharide from the maltodextrin polymer, amyloglucosidase from *A. niger* (Sigma) was added at a concentration of approximately 3 U/L. Yeast strains were grown for an additional 3 or 4 days at 30°C, cultures were separated by centrifugation, and salutaridine concentration was measured directly in the supernatant by LC-MS.

[00373] Plasmids and Strains

Plasmid/Strain	Genotype
pDW10	Spec ^R , TRP, P _{TDH3} -PbDRS-DRR-T _{CYC1}
pDW18	Spec ^R , TRP, P _{CYC1} -PbDRS-DRR-T _{CYC1}
pDW21	Spec ^R , TRP, P _{CYC1} -PbDRS-T _{ADH1} -P _{TEF1} -PbDRR-T _{CYC1}
pJL29	Spec ^R , TRP, P _{TDH3} -PbDRS-T _{ADH1} -P _{TEF1} -PbDRR-T _{CYC1}
pJL32	Spec ^R , TRP, P _{CYC1} -PbDRS-T _{ADH1} -P _{TDH3} -PbDRR-T _{CYC1}
pJL35	Spec ^R , TRP, P _{TDH3} -PbDRS-T _{ADH1} -P _{CYC1} -PbDRR-T _{CYC1}
YA106	<i>S. cerevisiae</i> Cen.PK, BIA pathway = NCS, CNMT, 6OMT, CYP80B1, CPR, 4OMT (complete genotype in Galanie et al. 2015)
YA511	<i>S. cerevisiae</i> Cen.PK, BIA pathway = NCS, CNMT, 6OMT, CYP80B1, CPR, 4OMT, DRR, SalSyn, SalR, SalAT (complete genotype in Galanie et al. 2015)
DW6	YA106, PbSalSyn(LEU+)
DW24	YA106, PbSalSyn(LEU+), ΔTRP(URA3+)

[00374] Unlike the morphinan alkaloid-producing poppy plants, which likely have highly specialized cell types for the synthesis of different metabolites in the BIA pathway (Onoyovwe et al., 2013), *S. cerevisiae* has just a single cell type, and the heterologous production of valuable compounds from this pathway requires that all enzymes, cofactors, and cellular machinery must work together in the same cell. To examine how the activities encoded in the wild-type fusion enzyme DRS-DRR function as separate enzymes in a heterologous system such as yeast, the cytochrome P450 (DRS) and the aldo-keto reductase (DRR) domains of the fused enzyme were split into two separate enzymes and expressed under different promoters. Using pDW18 as a template, the dehydroreticuline synthase (DRS) domain and the dehydroreticuline reductase (DRR) domain from DRS-DRR were cloned separately under the CYC1 and TEF1 promoters to yield plasmid pDW21, using standard molecular biology techniques. The pDW21 plasmid and the control plasmid harboring the parental fused DRS-DRR enzyme (pDW18) were separately transformed into the reporter strain. The strains were cultured and salutaridine production was measured directly in the supernatant via LC-MS analysis. **FIG. 18** shows that in independent transformants, the levels of salutaridine (gray bars) produced from transformants harboring pDW21 are

nearly 4 times higher than those from transformants harboring the control pDW18. The strains harboring pDW21 additionally exhibit a modest reduction in total reticuline (both S- and R-), and substantially less dehydroreticuline, in comparison to the control.

Example 2. Total BIA levels are increased in strains expressing DRS and DRR separately in comparison to the parental, fused enzyme.

[00375] Detailed LC-MS analysis was carried out to examine the effect of the independent expression of DRS and DRR enzymes derived from DRS-DRR on the levels of multiple BIA pathway intermediates. For BIA intermediate detection, a plasmid harboring the separate DRS enzyme from DRS-DRR under the control of the TDH3 promoter and the separate DRR enzyme under the control of the TEF1 promoter (pJL29) was constructed for comparison to the parental, fused enzyme expressed from the pDW10 plasmid. Based on promoter strength, this plasmid provided the closest approximation of equivalent expression between the separated and fused enzymes. The plasmids were separately transformed into the *S. cerevisiae* reporter strain described in **Example 1**. The strains were cultured as described in **Example 1** and BIA metabolites were measured directly in the supernatant via LC-MS analysis. As is shown in **FIG. 19**, the levels of nearly all BIA pathway intermediates are increased in the yeast strain harboring pJL29 in comparison to the strain harboring the control plasmid pDW10. Upstream pathway intermediates, in particular coclaurine and N-methyl-coclaurine, were increased in the strain expressing the separate DRS and DRR enzymes relative to the strain expressing the fused DRS-DRR enzyme, which indicates greater overall pathway capacity in cells that express DRS and DRR separately. The levels of reticuline in this experiment were approximately the same, although **Example 3** indicates that the majority of reticuline in this experiment was likely in the R- configuration. It is possible that the efficient conversion of S- to R-reticuline, and subsequent conversion to salutaridine, allow upstream enzymes to function better due to decreased product inhibition (from S-reticuline, for example) of their activity. Similar to the results shown in **FIG. 18**, cells expressing separate DRS and DRR enzymes produced greater amounts of salutaridine in comparison to cells expressing the fused parent DRS-DRR enzyme.

Example 3. Chiral Separation - Strains expressing DRS and DRR separately produce substantially more R-reticuline than the parental, fused enzyme.

[00376] Two yeast BIA pathway strains, derivatives of YA106 described in **Example 1**, expressing either the parental, fused DRS-DRR or separate DRS and DRR enzymes, were analyzed for their ability to produce S- and R-reticuline. Reticuline was concentrated from yeast media by pelleting 12.5 mL yeast culture and adding 30 mg XAD-4 resin per mL supernatant, incubating on a rotator overnight at room temperature, and eluting with 100 μ L methanol per mL supernatant. Samples were dried and resuspended

in 100 μ L water. The samples were fractionated by reverse-phase HPLC (Pursuit XRs-C18, 50 mm $\times$ 10 mm, 5 μ m) with isocratic 15% methanol.

[00377] Fractions were pooled, dried, and resuspended in isopropanol. Samples were separated on a chiral column (Phenomenex Lux cellulose- 1, 150 mm $\times$ 2 mm, 3 μ m) with isocratic 72% N-hexane, 28% isopropanol, 0.1% diethylamine with a flow rate of 0.3 mL/min and MS detection was performed with an Agilent 6320 Ion Trap mass spectrometer. The retention time of reticuline peaks was compared to that of authentic (S)-reticuline and (R)-reticuline standards. The strain that expressed separate DRS and DRR enzymes produced more than 87% of total reticuline as (R)-reticuline, whereas the strain that expressed the fused, parental enzyme produced approximately 50% of each stereoisomer as shown in **FIG. 20**.

Example 4. Construction and analysis of a structural model of DRS.

[00378] To better understand how the separate DRS enzyme folds and behaves mechanistically, and to build a framework for protein engineering efforts, structural homology models of DRS were constructed using the freely available SWISS-MODEL algorithm (<https://swissmodel.expasy.org/>). Two structural models, based on different CYP17 enzymes (PDBs 4NKX and 4R20), were constructed and analyzed in detail. The structural models based on these two different CYPs were nearly identical matches to each other, indicating broad structural homology between these types of cytochrome P450 molecules. Further, both models mapped the conserved EXXR motif and C-terminal cysteine residue (common to all P450s) of DRS in the identical location on the parent structure, suggesting the core structures of these molecules to be similar. **FIGs. 22 and 23** show the structural model of DRS based on 4NKX. The residual N-terminal ER anchor domain can be seen in the bottom left of the figure, as well as the approximate location of the heme binding region (modeled from 4R20), which is close to the enzyme active site. The face of the model shown is the proposed general binding region for cytochrome P450 reductase (CPR) where electron transfer occurs, and the back side of the model as shown in **FIGs. 22 and 23** are proposed to be proximal to the microsomal membrane. The homology model of DRS allows for potential functional roles to be assigned to any protein engineering improvements, based upon the position of residues and the amino acid changes that are identified.

Example 5. Intracellular DRS activity is rate limiting in comparison to DRR activity.

[00379] Additional experiments were performed on the separate expression of DRS and DRR enzymes to determine which activity was limiting in a heterologous cell. To better understand which enzyme activity, that of either DRS or DRR, was rate-limiting, a series of vectors expressing both enzymes under different promoter strengths was constructed. Using pDW21 as a parental construct, the TDH3 promoter, which is stronger than the original TEF1 promoter, was inserted and used to drive the expression of DRR, resulting in plasmid pJL32. Also using pDW21 as a parental construct, the TDH3 promoter was inserted

and used to drive DRS, resulting in plasmid pJL29. Using pJL29 as a parental construct, the CYC1 promoter, which is substantially weaker than TEF1, was inserted and used to drive DRR, resulting in plasmid pJL35. These plasmids were separately transformed in to the reporter yeast strain described in **Example 1**. The strains were cultured as described in **Example 1** and BIA metabolites were measured directly in the supernatant via LC-MS analysis. **FIG. 21** shows that when DRS expression is low (CYC1 promoter), and DRR expression is high (either TDH3 or TEF1), around 10 μ M of salutaridine is produced from the yeast strains. However, when DRS expression is high (TDH3 promoter), it may not matter whether DRR expression is low (CYC1) or at intermediate levels (TEF1), as both constructs result in around 20 μ M of salutaridine produced from the yeast strains, nearly double that observed from strains in which DRS is expressed at low levels. The results indicate that relative to DRR, the intracellular activity of DRS is limiting, and that additional protein engineering efforts can be applied to the DRS enzyme to further improve its activity.

Example 6. Mutagenesis and identification of improved DRS variants.

[00380] We performed mutagenesis-based enzyme engineering strategies to screen DRS libraries and identify improved DRS enzyme variants. The DRS open reading frame in pDW21 was subjected to error-prone PCR using the GeneMorph II (Agilent) random mutagenesis kit, according to the manufacturer's recommended protocol. Full-length, mutagenized PCR products of DRS were then either used as primers to directly reamplify and regenerate pDW21 plasmids, or recombined directly with linearized pDW21 using the Gibson Assembly kit (NEB), according to the manufacturer's recommended protocol. Resulting plasmids were then transformed into *E. coli*, and mutagenized pools of purified plasmids were purified and subsequently transformed into the *S. cerevisiae* reporter strain described in **Example 1**. Individual colonies were picked and cultured, and salutaridine levels were measured directly in the supernatant via LC-MS analysis. Approximately 5,000 independent variants were screened by LC-MS for increased production of salutaridine.

[00381] In comparison to control strains expressing wild type, non-mutagenized, parental DRS (DW24 harboring pDW21), strains expressing DRS variants that showed improved production of salutaridine were chosen for secondary screening and further study. The relative performance of each variant was determined by dividing salutaridine produced by the average salutaridine concentration from six independent replicates of the parental DRS expressed in the reporter strain. 25 variants that showed a fold change of 1.2X or higher than parental were pursued further. DNA shuffling of the improved variants and further screening was then used to isolate mutations that had a positive effect on activity from those that did not. For DNA shuffling, the NExT method (Mueller et al. 2005) was used as described, except DNA was digested with Endonuclease IV (NEB) and Phusion polymerase (NEB) was used for DNA amplification after reassembly. Shuffled variants were cloned into the pDW21 backbone and tested for salutaridine production in DW24 as described above.

[00382] Shuffled variants with improved activity over the parental sequence of DRS were sequenced (Quintara Biosciences) and mutations that showed the largest contribution to improved performance were identified. **Table 13** lists all shuffle variants that displayed improved performance. All mutations present in **Table 13** potentially result in increased activity of DRS relative to the parental sequence, but mutations that were identified in multiple improved shuffle variants or across multiple shuffle pools were determined most likely to result in direct improvements to DRS performance. The six sites satisfying these conditions were at positions 16, 17, 69, 110, 155, and 293.

[00383] When mapped onto the structural homology model of DRS based on CYP17A1 (PDB number 4NKX), positions 16, 17, and 69 are all located in the N-terminal ER localization domain that is responsible for tethering this class of P450 enzymes to the microsomal membrane. Without being bound by theory, it is possible that changes at these residues affect not only how DRS interacts with and is embedded into the ER, but also the overall activity of DRS, since membrane localization and orientation are critical for enzyme function. Position 110 is a surface-exposed residue on the homology model, and is located on the face of the enzyme that is proposed to interact with cytochrome P450 reductase (CPR), an essential redox donor for P450 activity. Changes at this position may positively influence how DRS interacts with CPR, by facilitating electron transfer between the two enzymes, for example. Position 155 is also a surface-exposed residue that may play a role in how DRS interacts with CPR, and changes at this position could positively affect activity. Position 293 is another surface-exposed residue that is on the opposite face of the molecule from position 110, and is located near channels in the structural model that could allow for substrate or product accessibility. Residue changes at this position could influence how substrate binds or how product is released. Additionally, this face of the enzyme is proximal to the ER membrane surface, and changes here could impact how the enzyme interfaces with the microsome.

Example 7. The amino acid positions at which PbDRS-DRR can be truncated to form separate DRS and DRR enzymes.

[00384] An alignment of the primary amino acid sequence of PbDRS-DRR versus dehydroreticuline synthase (DRS) and dehydroreticuline reductase (DRR) from *P. rhoeas* was generated using the Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) as shown in **FIG. 24**. Based on the alignment with DRR from *P. rhoeas*, we identified a truncation point at which to separate PbDRS-DRR into DRS and DRR enzymes, where a conserved methionine residue at position M569 is found (Seq. ID 16). This residue corresponds to position 1 of Seq. ID 18. The black arrow, between residues D568 and M569, represents the site at which PbDRS-DRR was truncated. The separate DRS enzyme based on PbDRS-DRR was designed to end at position D568. The dashed arrow points to a region of PbDRS-DRR that is not conserved with or homologous to either DRS or DRR from *P. rhoeas*. Truncations after each of these non-conservative residues, the sequence starting at K557 and ending at D568 within the black box, were

generated, and the activity of each successive truncation of DRS was assayed in a vector backbone identical to pDW21 (with DRR under the control of the TEF1 promoter). These plasmids were separately transformed in to the reporter yeast strain described in **Example 1**. The strains were cultured as described in **Example 1** and salutaridine was measured directly in the supernatant via LC-MS analysis.

Example 8. Identification of DRS variants that improve production of the morphinan alkaloid thebaine.

[00385] Mutations in 30 unique residues were identified in the DRS enzyme that improved epimerase activity in yeast and resulted in higher titers of salutaridine, a promorphinan alkaloid (see **Example 6**). The impact of several of these mutations on DRS activity and thebaine production were verified by cloning selected point mutations in DRS into plasmids. The point mutations selected for analysis were V16A, A17T, A50T, H69Q, V70I, M110I, and R160R (codon change). These plasmids were transformed into the screening strain yA511. yA511 is engineered with a complete biosynthetic pathway to thebaine, but lacks a functional DRS enzyme. Transformed colonies were picked in to pre-culture growth media and grown overnight as described in **Example 1**. Cells were backdiluted in to production media. After two days of production growth, supernatants from the cultures were isolated and measured using LC-MS/MS analysis. Cells expressing the DRS mutant variants accumulated up to 1.4 fold more thebaine in the culture medium compared to the control strain expressing wild type (WT) DRS (**FIG. 25**).

Example 9: Platform yeast strains engineered to produce (S)-reticuline from glucose and simple nitrogen sources

[00386] A platform yeast strain that produces the key branch point BIA intermediate (S)-reticuline from L-tyrosine was constructed (**FIG. 12**). Specifically, four multi-gene expression constructs were integrated into the genome of a yeast strain. The composition of the four constructs is indicated in **FIG. 26**. Each construct is comprised of 4 or 5 genes expressed from yeast promoters. Genes are positioned at each locus as complete expression cassettes comprising a promoter, gene open reading frame, and terminator as specified in the annotations above the schematic. The schematic shows the orientation of each expression cassette by the direction of the arrow representing a given gene. Selectable markers are italicized in the annotation and represented by grey arrows in the schematic. Each selection marker is flanked by loxP sites to allow removal of the marker from the locus. Additionally, each construct has a selectable marker flanked by loxP sites so that it can be removed by Cre recombinase.

[00387] In the first integration construct, four heterologous genes from *Rattus norvegicus* are integrated into the YBR197C locus together with a G418 selection marker (*KanMX*). *RnPTPS*, *RnSepR*, *RnPCD*,

and *RnQDHPR* are required to synthesize and regenerate tetrahydrobiopterin (BH₄) from the yeast endogenous folate synthesis pathway as indicated in **FIG. 1**. Each gene is codon optimized for expression in yeast.

[00388] In the second integration construct, four heterologous genes are integrated into the *HIS3* locus together with the *HIS5* selection marker. *Rattus norvegicus* tyrosine hydroxylase (*RnTyrH*) converts tyrosine to L-DOPA using the cosubstrate BH₄ generated by the preceding integration construct. The *RnTyrH* gene can be any of the wild-type or improved mutants which confer enhanced activity (e.g., W166Y, R37E, and R38E). A second *Rattus norvegicus* gene, *RnDHFR*, encodes an enzyme that reduces dihydrobiopterin (an oxidation product of BH₄) to BH₄, in this way increasing the availability of this cosubstrate. Also included in the third construct is *PpDODC* from *Pseudomonas putida*, an enzyme that converts L-DOPA to dopamine. The fourth enzyme is CjNCS from *Coptis japonica*, which condenses 4-HPA and dopamine to make norcoclaurine. Each gene is codon optimized for expression in yeast.

[00389] In the third integration construct, five heterologous genes from plants and the *LEU2* selection marker are integrated into the locus YDR514C. *Ps6OMT*, *Ps4'OMT*, and *PsCNMT* are methyltransferases from *Papaver somniferum* and are expressed as native plant nucleotide sequences. A fourth *P. somniferum* gene, *yPsCPRv2*, is codon optimized for yeast and encodes a reductase that supports the activity of a cytochrome P450 from *Eschscholzia californica*, *EcCYP80A1*. The enzymes encoded in this construct perform two *O*-methylations, an *N*-methylation, and a hydroxylation to produce reticuline from the norcoclaurine produced by the preceding integration construct. Each gene is codon optimized for expression in yeast.

[00390] In the final integration construct, additional copies of *Saccharomyces cerevisiae* endogenous genes *ARO4*^{Q166K}, *ARO7*^{T226I}, *TYR1*, and *ARO10* are integrated into the *ARO4* locus together with a hygromycin resistance selection marker. *ARO4*^{Q166K} and *ARO7*^{T226I} are feedback-resistant mutants of *ARO4* and *ARO10* which each encode a single base pair substitution relative to the wild-type sequence. *TYR1* and *ARO10* are identical to the native yeast genes, but are expressed behind strong promoters. Aro4p and Aro7p are enzymes in the biosynthesis of aromatic amino acids including tyrosine. Removing feedback inhibition from these enzymes results in upregulation of endogenous tyrosine biosynthesis. Overexpression of Tyr1p upregulates tyrosine biosynthesis and thus production of tyrosine. Overexpression of Aro10p increases the production of 4-HPA.

[00391] Platform yeast strains can be constructed with any number of the four expression cassettes. Specifically, platform yeast strains were constructed with integration constructs 1-4 and integration constructs 1-3. In the latter strain in which the tyrosine over-production construct (construct 4) is excluded, additional tyrosine may be supplied in the culture medium to support the biosynthesis of reticuline. Additional genetic modifications may be incorporated into the platform strains to support production of downstream BIAs and increased flux to BIA biosynthesis.

[00392] The yeast strains were grown in synthetic complete media with the appropriated amino acid drop out solution at 30 °C. BIA metabolites in the media supernatant were analyzed after 48 and 96 hours of growth by LC-MS/MS analysis.

Example 10: Platform yeast strains engineered to produce thebaine from glucose and simple nitrogen sources

[00393] Yeast strains can be engineered for the production of the morphinan alkaloid thebaine from early precursors such as tyrosine. As an example, the platform yeast strains described in **Example 9** can be further engineered to produce the morphinan alkaloid products from L-tyrosine (**FIG. 14**).

[00394] The platform yeast strain producing (*S*)-reticuline from L-tyrosine (see description in **Example 9**) was further engineered to incorporate an engineered split epimerase DRS and DRR, an engineered salutaridine synthase, salutaridine reductase, salutaridinol acetyltransferase, and thebaine synthase to convert the biosynthesized (*S*)-reticuline to the first morphinan alkaloid thebaine (**FIG. 14**). Three expression cassettes ($P_{TDH3}\text{-}yEcCFS^{1-26}\text{-}yPbSS^{33-504}$, $P_{TPII}\text{-}yPbSalR$, $P_{TEFI}\text{-}yPsSalAT$) were assembled into an integration construct with a *bleR* selective marker and integrated into the locus TRP1 in the platform yeast strain. An additional three expression cassettes ($P_{TDH3}\text{-}yPbDRS$, $P_{TEFI}\text{-}yPbDRR$, $P_{PGK1}\text{-}yPsTS$) were assembled into an integration construct with a URA3 selective marker and integrated into the locus YPL250CΔ in the platform yeast strain. The composition of the two constructs is indicated in **FIG. 27**.

[00395] The yeast strains harboring the integrated cassettes were grown in synthetic complete media with the appropriated drop out solution at 30 °C. After 96 hours of growth, the media was analyzed for BIA metabolites by LC-MS/MS analysis.

Example 11: Production of thebaine from glucose and simple nitrogen sources via engineered yeast strains

[00396] Yeast strains were engineered as described in **Examples 9** and **10** to produce the pentacyclic morphinan alkaloid thebaine directly from simple sugars (e.g., glucose) and nitrogen sources present in standard growth media. Specifically, a CEN.PK strain of *Saccharomyces cerevisiae* was engineered to express the following heterologous enzymes via integration into the yeast chromosome: TyrH, DODC, PTPS, SepR, PCD, QDHPR, NCS, 6OMT, CNMT, CYP80B1, CPR, 4OMT, DRS, DRR, SalSyn, SalR, SalAT, and TS. In this example, the SalSyn enzyme is engineered to have its leader sequence replaced with 83 amino acids from the N-terminus of *Eschscholzia californica* chelanthifoline synthase (EcCFS). Additional modifications were made to the strain to increase BIA precursor accumulation, including:

overexpression of ARO10, overexpression of TYR1, expression of a feedback resistant ARO4 (ARO4^{Q166K}), and expression of a feedback resistant ARO7 (ARO7^{T226I}). Separate engineered yeast strains were made as described, harboring different variants of enzymes encoding thebaine synthase activity (TS), including SEQ ID NOs. 35 (i.e., TS1), 37 (i.e., TS2), and a variant of 35 with a N-terminal truncation of the first 22 amino acids (i.e., tTS1), and no thebaine synthase enzyme (YA397). The sequences of the enzyme variants are provided in **Table 2**.

[00397] The described yeast strains were inoculated into 2 ml of synthetic complete media (yeast nitrogen base and amino acids) with 2% glucose and grown for approximately 4 hours at 30°C. Then, 10 uL of each culture was transferred to 400 uL of fresh media in a 96-well plate in replicates of 4 and grown for an additional 48 hours at 30°C. The production media contains 1x yeast nitrogen broth and amino acids, 20 mM ascorbic acid, 300 mg/L tyrosine, 40 g/L maltodextrin, and 2 units/L amylase. The amylase is used to mimic a fed-batch process and gradually releases glucose from maltodextrin polymer so that the yeast can use it as a carbon source. The cells were separated from the media by centrifugation, and thebaine concentration was measured directly in the supernatant by LC-MS/MS analysis. All engineered yeast strains produced thebaine from glucose and simple nitrogen sources present in the growth media (**FIGs. 29 and 30**). Strains harboring a thebaine synthase activity produced higher levels of thebaine relative to strains not harboring this activity under the described fermentation conditions.

Example 12: Yeast strains engineered to produce downstream morphinan alkaloids from glucose and simple nitrogen sources

[00398] Yeast strains can be engineered for the production of the downstream morphinan alkaloids from early precursors such as tyrosine. As an example, the platform yeast strains described in **Example 10** can be further engineered to produce the downstream morphinan alkaloid products from L-tyrosine (**FIG. 14**).

[00399] The platform yeast strain producing thebaine from L-tyrosine (see description in **Example 10**) was further engineered to incorporate thebaine 6-O-demethylase, codeinone reductase, and codeinone-O-demethylase to convert the biosynthesized thebaine to the downstream morphinan alkaloids including morphine (**FIG. 14**). Three expression cassettes (P_{ADHI} - $T6ODM$ - T_{ADHI} , P_{HXT7} - COR - T_{PGK1} , P_{TEF1} - $CODM$ - T_{CYC1}) were directly assembled with a $TRP1$ selective marker and integrated into the $trp1$ locus in the thebaine platform yeast strain (Thodey et al., 2014).

[00400] The yeast strains harboring the integrated cassettes were grown in synthetic complete media with the appropriated drop out solution at 30 °C. After 96 hours of growth, the media was analyzed for BIA metabolites by LC-MS/MS analysis.

Example 13: Yeast strains engineered to produce semi-synthetic opioids from glucose and simple nitrogen sources

[00401] Yeast strains can be engineered for the production of the downstream morphinan alkaloids from early precursors such as tyrosine. As an example, the yeast strains described in **Examples 9 and 10** can be further engineered to produce the semi-synthetic opioid products from L-tyrosine (**FIG. 15**).

[00402] The yeast strains producing downstream morphinan alkaloids from L-tyrosine (see description in **Example 10**) were further engineered to incorporate morphine dehydrogenase and morphinone reductase to convert the biosynthesized thebaine to the downstream morphinan alkaloids including morphine (**FIG. 15**). Two expression cassettes (P_{GPD} -*morA*- T_{CYC1} , P_{PGK1} -*morB*- T_{PHO5}) were directly assembled with a *KanMX* selective marker and integrated into the HO locus in the downstream morphinan alkaloids producing yeast strains (Thodey et al., 2014).

[00403] The yeast strains harboring the integrated cassettes were grown in synthetic complete media with the appropriated drop out solution at 30 °C. After 96 hours of growth, the media was analyzed for BIA metabolites by LC-MS/MS analysis.

Example 14: Microbial strains engineered to produce O-demethylated opioid compounds from glucose and simple nitrogen sources

[00404] Enzymes listed in **Table 4** that displayed O-demethylase activity on morphinan alkaloids, were incorporated into a microbial strain (either *Saccharomyces cerevisiae* or *Escherichia coli*) which biosynthesizes morphinan alkaloids de novo (as described in **Example 10**). The complete BIA biosynthetic pathway uses L-tyrosine produced by the host cell and/or supplemented in the culture medium. Two molecules of tyrosine are modified and condensed to form the first benzylisoquinoline structure, which may be either norcoclaurine or norlaudanosoline. The benzylisoquinoline is further modified to form (S)-reticuline and then stereochemically inverted by the activity of an engineered split epimerase enzyme to yield (R)-reticuline. (R)-reticuline undergoes a carbon-carbon coupling reaction to form the first promorphinan, salutaridine, and is further modified before undergoing an oxygen-carbon coupling reaction catalyzed by a thebaine synthase to arrive at the first morphinan alkaloid structure, thebaine (see **FIG. 14**). **Table 3** lists enzymes and activities in the complete pathway.

[00405] **FIG. 10** illustrates a biosynthesis scheme in a microbial cell, in accordance with embodiments of the invention. Tyrosine produced endogenously by the cell and/or supplied in the culture medium is converted to oxycodone (broken arrows represent multiple enzymatic steps). The oxycodone is then 3-O-demethylated to oxymorphone and N-demethylated to noroxymorphone. Finally, an N-methyltransferase

accepts allyl and cyclopropylmethyl carbon moieties from SAM analogues to produce naloxone and naltrexone, respectively.

[00406] To detect *O*-demethylase activity in strains producing morphinan alkaloid molecules, cells expressing candidate enzymes, either from plasmid vectors or chromosomally-integrated cassettes, were propagated by fermentation and cell supernatants were collected to analyze the total opioid profile (as described above). *O*-demethylation of opioid molecules in strains harboring the complete BIA pathway was detected by LC-MS (as described above). Specifically, the conversion of oxycodone to oxymorphone was detected. To detect *O*-demethylation activity via biocatalysis, strains were cultured in selective medium and then lysed by glass bead disruption. Cell lysates were supplied exogenously with opioid substrates (see **FIG. 6**), and other cofactors necessary for enzyme function. *O*-demethylation of opioid molecules was detected by LC-MS.

Example 15: Microbial strains engineered to produce N-demethylated opioid compounds from glucose and simple nitrogen sources

[00407] Enzymes listed in **Table 5**, that displayed *N*-demethylase activity on morphinan alkaloids, were incorporated into a microbial strain (either *Saccharomyces cerevisiae* or *Escherichia coli*) which biosynthesizes morphinan alkaloids de novo (as described in **Example 10**). The complete BIA biosynthetic pathway uses L-tyrosine produced by the host cell and/or supplemented in the culture medium. Two molecules of tyrosine are modified and condensed to form the first benzyloquinoline structure which may be either norcoclaurine or norlaudanosoline. The benzyloquinoline is further modified to form (*S*)-reticuline and then stereochemically inverted by the activity of an engineered split epimerase enzyme to yield (*R*)-reticuline. (*R*)-reticuline undergoes a carbon-carbon coupling reaction to form the first promorphinan, salutaridine, and is further modified before undergoing an oxygen-carbon coupling reaction catalyzed by a thebaine synthase to arrive at the first morphinan alkaloid structure, thebaine (see **FIG. 14**). **Table 3** lists enzymes and activities in the complete pathway.

To detect *N*-demethylase activity in strains producing morphinan alkaloid molecules, cells expressing candidate enzymes, either from plasmid vectors or chromosomally-integrated cassettes, were propagated by fermentation and cell supernatants were collected to analyze the total opioid profile (as described above). *N*-demethylation of opioid molecules in strains harboring the complete BIA pathway was detected by LC-MS (as described above). Specifically, the conversion of oxymorphone to noroxymorphone was detected. To detect *N*-demethylation activity via biocatalysis, strains were cultured in selective medium and then lysed by glass bead disruption. Cell lysates were supplied exogenously with opioid substrates (see **FIG. 7**), and other cofactors necessary for enzyme function. *N*-demethylation of opioid molecules was detected by LC-MS.

Example 16: Microbial strains engineered to produce nal-opioid compounds from glucose and simple nitrogen sources

[00408] Enzymes listed in **Table 6**, that displayed *N*-methylase activity on morphinan alkaloids, were incorporated into a microbial strain (either *Saccharomyces cerevisiae* or *Escherichia coli*) which biosynthesizes morphinan alkaloids de novo (as described in **Example 10**). **FIG. 10** shows an example of the complete reaction scheme from the precursor molecule thebaine to the final nal-opioid compounds naloxone and naltrexone. These strains additionally express enzymes from **Examples 14 and 15** and **Table 3**, that are responsible for generating nor-opioid compounds from the complete BIA pathway. *N*-methylase enzymes were also expressed in a microbial strain (either Cen.PK2 for *S. cerevisiae* or BL21 for *E. coli*, for example) lacking the biosynthetic pathway, to generate a strain that is capable of biocatalysis of several different exogenously-supplied substrate molecules. The complete BIA biosynthetic pathway uses tyrosine produced by the host cell and/or supplemented in the culture medium. Two molecules of tyrosine are modified and condensed to form the first benzyloquinoline structure which may be either norcoclaurine or norlaudanoline. The benzyloquinoline is further modified to form (*S*)-reticuline and then stereochemically inverted by the activity of an engineered split epimerase enzyme to yield (*R*)-reticuline. (*R*)-reticuline undergoes a carbon-carbon coupling reaction to form the first promorphinan, salutaridine, and is further modified before undergoing an oxygen-carbon coupling reaction catalyzed by a thebaine synthase to arrive at the first morphinan alkaloid structure, thebaine (see **FIG. 14**). **Table 3** lists enzymes and activities in the complete pathway.

[00409] To detect *N*-modifying activity in strains with the complete BIA pathway to nor-opioids (see **FIG. 10**), cells expressing candidate enzymes were propagated by fermentation (as described above) and incubated with SAM or SAM analogs, such as those listed in **FIG. 8**. Enzymatic modification of nor-opioid or other BIA molecules in strains harboring the complete BIA pathway was detected in supernatants by LC-MS (as described above). To detect *N*-modifying activity via biocatalysis, strains were cultured in selective medium and then lysed by glass bead disruption. Cell lysates were supplied exogenously with SAM or SAM analogs, and other cofactors necessary for enzyme function. Specifically, the conversion of noroxymorphone to naloxone and naltrexone (using the SAM analogs allyl-SAM or cyclopropane-SAM, as shown in **FIG. 8**) was detected. Modification of nor-opioid or other BIA molecules was detected by LC-MS. To detect *N*-modifying activity by biocatalysis in a strain that does not have the complete BIA pathway, Cen.PK2 strains expressing the described heterologous enzymes were grown in selective medium and lysed by glass bead disruption. Cell lysates were supplied exogenously with SAM or SAM analogs, cofactors necessary for enzyme function, and nor-opioid molecules such as those listed in **FIG. 8** and **Table 3**. Modification of these compounds was detected by LC-MS.

Table 3. List of enzymes

Enzyme	Abbrev	Catalyzed Reactions	Source organisms	Genbank #
Transketolase	TKL1	fructose-6-phosphate + glyceraldehyde-3-phosphate $\leftrightarrow$ xylulose-5-phosphate + erythrose-4-phosphate (EC 2.2.1.1)	<i>Saccharomyces cerevisiae</i>	NP_015399.1
Glucose-6-phosphate dehydrogenase	ZWF1	glucose-6-phosphate $\rightarrow$ 6-phosphogluconolactone (EC 1.1.1.49)	<i>Saccharomyces cerevisiae</i>	CAA96146.1
Prephenate dehydrogenase	TYR1	prephenate + NADP ⁺ $\rightarrow$ 4-hydroxyphenylpyruvate + CO ₂ + NADPH (EC 1.3.1.13)	<i>Saccharomyces cerevisiae</i>	CAA85127.1
3-deoxy-d-arabinose-heptulosonate-7-phosphate synthase	ARO4, DAHP synthase	erythrose-4-phosphate + PEP $\rightarrow$ DAHP (EC 2.5.1.54)	<i>Saccharomyces cerevisiae</i>	CAA85212.1
Chorismate mutase	ARO7	chorismate $\rightarrow$ prephenate (EC 5.4.99.5)	<i>Saccharomyces cerevisiae</i>	NP_015385.1
Phenylpyruvate decarboxylase	ARO10	hydroxyphenylpyruvate $\rightarrow$ 4HPA (EC 4.1.1.80)	<i>Saccharomyces cerevisiae</i>	NP_010668.3
Alcohol dehydrogenase	ADH2-7, SFA1	4HPA $\rightarrow$ tyrosol (EC 1.1.1.90)	<i>Saccharomyces cerevisiae</i>	NP_014032.1, AAT93007.1, NP_011258.2, NP_009703.3, NP_014051.3, NP_010030.1, NP_010113.1
Aldehyde oxidase	ALD2-6	4HPA $\rightarrow$ hydroxyphenylacetic acid (EC 1.2.1.39)	<i>Saccharomyces cerevisiae</i>	NP_013893.1, NP_013892.1, NP_015019.1, NP_010996.2, NP_015264.1
Aryl-alcohol dehydrogenase	AAD4, 6, 10, 14-16	aromatic aldehyde + NAD ⁺ $\rightarrow$ aromatic alcohol + NADH (EC 1.1.1.90)	<i>Saccharomyces cerevisiae</i>	GAX67600, GAX72034, AAT93180, AAS56234, NP_012689, GAX69843, NP_014477, AAS56127
Aromatic aminotransferase	ARO9	hydroxyphenylpyruvate + L-alanine $\leftrightarrow$ tyrosine + pyruvate (EC 2.6.1.58)	<i>Saccharomyces cerevisiae</i>	AEC14313.1
Aromatic aminotransferase	ARO8	hydroxyphenylpyruvate + glutamate $\leftrightarrow$ tyrosine + alpha-ketoglutarate (EC 2.6.1.5)	<i>Saccharomyces cerevisiae</i>	K2V11027.1
Tyrosinase	TYR	tyrosine $\rightarrow$ L-DOPA, L-DOPA $\rightarrow$ dopaquinone (EC 1.14.18.1)	<i>Ralstonia solanacearum</i> , <i>Agaricus bisporus</i>	NP_518458.1, AJ223816,
Tyrosine hydroxylase	TyrH	tyrosine $\rightarrow$ L-DOPA (EC 1.14.16.2)	<i>Homo sapiens</i> , <i>Rattus norvegicus</i> , <i>Mus musculus</i>	NM_012740, NM_000240
L-DOPA decarboxylase	DODC	L-DOPA $\rightarrow$ dopamine (EC 4.1.1.28)	<i>Pseudomonas putida</i> , <i>Rattus norvegicus</i>	AE015451.1, NP_00125778.2.1
Tyrosine/DOPA decarboxylase	TYDC	L-DOPA $\rightarrow$ dopamine (EC 4.1.1.28) tyrosine $\rightarrow$ tyramine (EC 4.1.1.25)	<i>Papaver somniferum</i>	AAA97535
Monoamine oxidase	MAO	dopamine $\rightarrow$ 3,4-DHPA (EC 1.4.3.4)	<i>E. coli</i> , <i>Homo sapiens</i> , <i>Micrococcus luteus</i>	J03792, D2367, AB010716.1
Norcoclaurine synthase	NCS	4HPA + dopamine $\rightarrow$ S-norcoclaurine (EC 4.2.1.78) 3,4-DHPA + dopamine $\rightarrow$ S-norlaudanosoline	<i>Coptis japonica</i> , <i>Papaver somniferum</i> , <i>Papaver bracteatum</i> , <i>Thalictrum flavum</i> , <i>Corydalis saxicola</i>	BAF45337.1, AB267399.2, ACl45396.1, AC090258.1, ACO90247.1, AEB71889.1

Norcoclaurine 6-O-methyltransferase	6OMT	Norcoclaurine → coclaurine Norlaudanoline → 3'-hydroxycoclaurine EC 2.1.1.128	<i>P. somniferum</i> <i>T. flavum</i> <i>Coptis japonica</i> *	AY268894 AY610507 D29811
Coclaurine-N-methyltransferase	CNMT	Coclaurine → N-methylcoclaurine 3'-hydroxycoclaurine → 3'-hydroxy-N-methylcoclaurine EC 2.1.1.140	<i>P. somniferum</i> <i>T. flavum</i> <i>Coptis japonica</i> *	AY217336 AY610508 AB061863
4'-O-methyltransferase	4'OMT	3'-hydroxy-N-methylcoclaurine → Reticuline EC 2.1.1.116	<i>P. somniferum</i> <i>T. flavum</i> <i>Coptis japonica</i> *	AY217333, AY217334 AY610510 D29812
Cytochrome P450 80B1	CYP80B1	N-methylcoclaurine → 3'-hydroxy-N-methylcoclaurine (EC 1.14.13.71)	<i>P. somniferum</i> , <i>E. californica</i> , <i>T. flavum</i>	AAF61400.1 AAC39453.1 AAU20767.1
GTP cyclohydrolase	FOL2	GTP → dihydroneopterin triphosphate (EC 3.5.4.16)	<i>Saccharomyces cerevisiae</i> , <i>Homo sapiens</i> , <i>Mus musculus</i>	CAA97297.1, NP_00101919 5.1, NP_032128.1
6-pyruvoyl tetrahydrobiopterin (PTP) synthase	PTPS	dihydroneopterin triphosphate → PTP (EC 4.2.3.12)	<i>Rattus norvegicus</i> , <i>Homo sapiens</i> , <i>Mus musculus</i>	AAH59140.1, BAA04224.1, AAH29013.1
Sepiapterin reductase	SepR	PTP → BH4 (EC 1.1.1.153)	<i>Rattus norvegicus</i> , <i>Homo sapiens</i> , <i>Mus musculus</i>	NP_062054.1, NP_003115.1, NP_035597.2
4a-hydroxytetrahydrobiopterin (pterin-4 α -carbinolamine) dehydratase	PCD	4a-hydroxytetrahydrobiopterin → H ₂ O + quinoid dihydropteridine (EC 4.2.1.96)	<i>Rattus norvegicus</i> , <i>Homo sapiens</i> , <i>Mus musculus</i>	NP_00100760 2.1, AAB25581.1, NP_079549.1
Quinoid dihydropteridine reductase	QDHPR	quinoid dihydropteridine → BH4 (EC 1.5.1.34)	<i>Rattus norvegicus</i> , <i>Homo sapiens</i> , <i>Mus musculus</i>	AAH72536.1, NP_000311.2, AAH02107.1
Dihydrofolate reductase	DHFR	7,8-Dihydrobiopterin → 5,6,7,8-Tetrahydrobiopterin (BH4) EC 1.5.1.3	<i>Rattus norvegicus</i> , <i>Homo sapiens</i>	AF318150.1
1-benzylisoquinoline alkaloid epimerase (cytochrome P450 82Y2-codeinone reductase; dehydroreticuline synthase-dehydroreticuline reductase)	DRS-DRR (CYP-COR)	(S)-reticuline → (R)-reticuline (S)-1-benzylisoquinoline → (R)-1-benzylisoquinoline EC 1.5.1.27	<i>Papaver bracteatum</i> , <i>Papaver somniferum</i> , <i>Papaver setigerum</i> , <i>Chelidonium majus</i>	PODKI7.1, AKO60175.1, AKO60180.1, AKO60179.1, AKO60175.1
(R)-reticuline, NADPH: oxygen oxidoreductase (C-C phenol-coupling), also known as salutaridine synthase	SalSyn	(R)-reticuline + NADPH + H ⁺ + O ₂ → salutaridine + NADP ⁺ + 2 H ₂ O EC 1.14.21.4	<i>Papaver somniferum</i> , <i>Papaver spp</i> <i>Chelidonium majus</i>	EF451150 (Ref PMID 22424601)
salutaridinol: NADP ⁺ 7-oxidoreductase, also known as salutaridine reductase	SalR	salutaridinol + NADP ⁺ ↔ salutaridine + NADPH + H ⁺ EC 1.1.1.248	<i>Papaver somniferum</i> , <i>Papaver bracteatum</i> , <i>Papaver spp</i> <i>Chelidonium majus</i>	DQ316261, EF184229 (Ref PMID 22424601)
acetyl-CoA: salutaridinol 7-O-acetyltransferase	SalAT	acetyl-CoA + salutaridinol → CoA + 7-O-acetylsalutaridinol EC 2.3.1.150	<i>Papaver somniferum</i> , <i>Papaver bracteatum</i> , <i>Papaver orientale</i> , <i>Papaver spp</i>	AF339913, FJ200355, FJ200358, FJ200356, JQ659008
thebaine synthase	TS	7-O-acetylsalutaridinol → thebaine + acetate	<i>Papaver somniferum</i> , <i>Papaver bracteatum</i> , <i>Papaver orientale</i> , <i>Papaver spp</i>	AWQ63979, AWQ63980
Thebaine 6-O demethylase	T6ODM	thebaine → neopinone EC 1.14.11.31	<i>Papaver somniferum</i> , <i>Papaver spp</i> .	GQ500139.1

Codeinone reductase	COR	codeinone → codeine EC 1.1.1.247, neopinone → neopine	<i>Papaver somniferum</i> , <i>Papaver spp.</i>	AF108432.1 AF108433.1 AF108434.1 AF108435.1
Codeine O-demethylase	CODM	codeine → morphine EC 1.14.11.32, neopine → neomorphine	<i>Papaver somniferum</i> , <i>Papaver spp.</i>	GQ500141.1
Morphine dehydrogenase	morA	morphine → morphinone EC 1.1.1.218, codeinone → codeine EC 1.1.1.247	<i>Pseudomonas putida</i>	M94775.1
Morphinone reductase	morB	codeinone → hydrocodone morphinone → hydromorphone EC 1.3.1.-	<i>Pseudomonas putida</i>	U37350.1
NADPH:hemoprotein oxidoreductase, also known as cytochrome P450 reductase	ATR1, CPR	NADPH + H ⁺ + n oxidized hemoprotein → NADP ⁺ + n reduced hemoprotein EC 1.6.2.4	<i>Arabidopsis thaliana</i> , <i>E. californica</i> , <i>P. somniferum</i> , <i>H. sapiens</i> , <i>S. cerevisiae</i> , <i>P. bracteatum</i> , <i>Papaver spp.</i> , all plants	NM118585, CAB58576.1, CAB58575.1, AAC05021 .1, AAC05022.1 many others (Ref PMID 19931192)
Cytochrome P450, family 2, subfamily D, polypeptide 6	CYP2D6	Promiscuous oxidase, can perform (R)-reticuline + NADPH + H ⁺ + O ₂ → salutaridine + NADP ⁺ + 2 H ₂ O among other reactions EC 1.14.14.1	<i>Homo sapiens</i>	BC067432
S-adenosyl-L-methionine:(S)- scoulerine 9-O- methyltransferase	S9OMT	S-adenosyl-L-methionine + (S)-scoulerine → S-adenosyl-L-homocysteine + (S)- tetrahydrocolumbamine EC 2.1.1.117	<i>Thalictrum flavum</i> subsp. <i>glaucum</i> , <i>Papaver somniferum</i> , <i>Coptis japonica</i> , <i>Coptis chinensis</i> , <i>Thalictrum spp.</i> , <i>Coptis spp.</i> , <i>Papaver spp.</i>	AY610512, D29809, EU980450, JN185323
Tetrahydroprotoberberine- N-methyltransferase	TNMT	Stylopine → cis-N-methylstylopine EC 2.1.1.122 Canadine → N-methylcanadine	<i>P. somniferum</i> <i>E. californica</i> <i>P. bracteatum</i> <i>A. mexicana</i>	DQ028579 EU882977 EU882994 HQ316698
Cheilanthisfoline synthase	CFS	Cheilanthisfoline → stylopine EC 1.14.21.1	<i>P. somniferum</i> <i>E. californica</i> <i>P. bracteatum</i> <i>A. mexicana</i>	GU325749 AB434654 EF451152
Stylopine synthase	STS	Stylopine → cis-N-methylstylopine EC 2.1.1.122	<i>P. somniferum</i> <i>E. californica</i> <i>P. bracteatum</i> <i>A. mexicana</i>	GU325750 AB126257 EF451151
Cis-N-methylstylopine 14- hydroxylase	MSH	Cis-N-methylstylopine → protopine EC 1.14.13.37	<i>P. somniferum</i> <i>E. californica</i> <i>P. bracteatum</i> <i>A. mexicana</i>	KC154003

Protopine-6-hydroxylase	P6H	Protopine → 6-hydroxyprotopine EC 1.14.13.55	<i>E. californica</i> <i>P. somniferum</i> <i>P. bracteatum</i> <i>A. mexicana</i>	AB598834 AGC92397
Dihydrobenzophenanthridine oxidase	DBOX	Dihydrosanguinarine → sanguinarine EC 1.5.3.12	<i>P. somniferum</i> <i>E. californica</i> <i>P. bracteatum</i> <i>A. mexicana</i>	AGL44336, AGL44335, AGL44334
(S)-tetrahydroprotoberberine oxidase	STOX	(S)-tetrahydroberberine + 2 O ₂ → berberine + 2 H ₂ O ₂ EC 1.3.3.8	<i>Berberis wilsonae</i> , <i>Coptis japonica</i> , <i>Berberis spp.</i> , <i>Coptis spp.</i>	HQ116697, AB564543
(S)-tetrahydrocolumbamine, NADPH:oxygen oxidoreductase (methylenedioxy-bridge-forming), also known as (S)-canadine synthase	CAS	(S)-tetrahydrocolumbamine + NADPH + H ⁺ + O ₂ → (S)-canadine + NADP ⁺ + 2 H ₂ O EC 1.14.21.5	<i>Thalictrum flavum subsp. glaucum</i> , <i>Coptis japonica</i> , <i>Thalictrum spp.</i> , <i>Coptis spp.</i>	AY610513, AB026122, AB374407, AB374408
(S)-reticuline:oxygen oxidoreductase (methylene-bridge-forming), also known as berberine bridge enzyme	BBE	(S)-reticuline + O ₂ → (S)-scoulerine + H ₂ O ₂ EC 1.21.3.3	<i>Papaver somniferum</i> , <i>Argemone mexicana</i> , <i>Eschscholzia californica</i> , <i>Berberis stolonifera</i> , <i>Thalictrum flavum subsp. glaucum</i> , <i>Coptis japonica</i> , <i>Papaver spp.</i> , <i>Eschscholzia spp.</i> , <i>Berberis spp.</i> , <i>Thalictrum spp.</i>	AF025430, EU881889, EU881890, S65550 AF005655, AF049347, AY610511, AB747097
Berbamunine synthase	CYP80A1	(S)-N-methylcoclaurine + (R)-N-methylcoclaurine → berbamunine EC 1.14.21.3 1-benzylisoquinoline alkaloid + 1-benzylisoquinoline alkaloid → bis-benzylisoquinoline alkaloid	<i>Berberis stolonifera</i> , <i>Capsicum chinense</i> , <i>Quercus suber</i>	AAC48987, PHU28278, POF05239
Protopine O-dealkylase	PODA	O,O-demethylenation of canadine, stylophine and berberine	<i>P. somniferum</i> , <i>Papaver spp.</i>	GQ500140.1
Reticuline N-methyltransferase	RNMT	reticuline → tembetarine	<i>Papaver somniferum</i> , <i>Papaver spp.</i>	KX369612.1
Papaverine 7-O-demethylase	P7OMT	papaverine → pacodine	<i>Papaver somniferum</i> , <i>Papaver spp.</i>	KT159979.1
3-O-demethylase	3ODM	oxycodone → oxymorphone hydrocodone → hydromorphone dihydrocodeine → dihydromorphone 14-hydroxycodone → 14-hydroxymorphone codeinone → morphinone 14-hydroxycodeinone → 14-hydroxymorphinone	<i>Papaver somniferum</i> , <i>Papaver bracteatum</i> , <i>Papaver rhoeas</i> , <i>Papaver spp.</i>	

N-demethylase (in some cases the NDM activity is encoded in BM3 or an engineered variant of BM3)	NDM	Codeine→Norcodeine Morphine→Normorphine Oxycodone→Noroxycodone Oxymorphone→Noroxymorphone Thebaine→Northebaine Oripavine→Nororipavine Hydrocodone→Norhydrocodone Hydromorphone→Norhydromorphone Dihydrocodeine→Nordihydrocodeine Dihydromorphine→Nordihydromorphine 14-hydroxycodone→Nor-14-hydroxycodone 14-hydroxymorphine→Nor-14-hydroxymorphine Codeinone→Norcodeinone Morphinone→Normorphinone 14-hydroxycodeinone→Nor-14-hydroxycodeinone 14-hydroxymorphinone→Nor-14-hydroxymorphinone	<i>Bacillus megaterium</i> , <i>Homo sapiens</i> , <i>Papaver somniferum</i> , <i>Papaver</i> spp., <i>Chelidonium majus</i> , <i>Stylophorum diphyllum</i> , <i>Nigella sativa</i> , <i>Hydrastis canadensis</i> , <i>Glaucium flavum</i> , <i>Eschscholzia californica</i> , <i>Menispermum canadense</i> , <i>Papaver bracteatum</i>	
N-methyltransferase	NMT	Norcodeine→codeine Normorphine→morphine Noroxycodone→oxycodone Noroxymorphone→noroxymorphone Northebaine→thebaine Nororipavine→oripavine Norhydrocodone→hydrocodone Norhydromorphone→Hydromorphone Nordihydrocodeine→ Dihydrocodeine Nordihydromorphine→ Dihydromorphine Nor-14-hydroxycodone→ 14-hydroxycodone Nor-14-hydroxymorphine→ 14-hydroxymorphine Norcodeineone→ Codeineone Normorphinone→ Morphinone Nor-14-hydroxy-codeinone→ 14-hydroxycodeinone Nor-14-hydroxy-morphinone→ 14-hydroxymorphinone	<i>Papaver</i> spp., <i>Chelidonium majus</i> , <i>Thalictrum flavum</i> , <i>Coptis japonica</i> , <i>Papaver somniferum</i> , <i>Eschscholzia californica</i> , <i>Papaver bracteatum</i> , <i>Argemone mexicana</i> , <i>Glaucium flavum</i> , <i>Sanguinaria canadensis</i> , <i>Corydalis chelanthifolia</i> , <i>Nigella sativa</i> , <i>Jeffersonia diphylla</i> , <i>Berberis thunbergii</i> , <i>Mahonia aquifolium</i> , <i>Menispermum canadense</i> , <i>Tinospora cordifolia</i> , <i>Cissampelos mucronata</i> , <i>Cocculus trilobus</i>	

N-allyltransferase	NAT	Norcodeine→N-allyl-norcodeine Normorphine→N-allyl-normorphine Noroxycodone→N-allyl-noroxycodone Noroxymorphine→N-allyl-noroxymorphine Northebaine→N-allyl-northebaine Nororipavine→N-allyl-nororipavine Norhydrocodone→N-allyl-norhydrocodone Norhydromorphine→ N-allyl-norhydromorphine Nordihydrocodeine→ N-allyl-nordihydrocodeine Nordihydromorphine→ N-allyl-nordihydromorphine Nor-14-hydroxycodeine→ N-allyl-nor-14-hydroxycodeine Nor-14-hydroxymorphine→ N-allyl-nor-14-hydroxymorphine Norcodeineone→ N-allyl-norcodeineone Normorphinone→ N-allyl-normorphinone Nor-14-hydroxy-codeinone→ N-allyl-nor-14-hydroxycodeinone Nor-14-hydroxy-morphinone→ N-allyl-nor-14-hydroxymorphinone	<i>Papaver spp.</i>, <i>Chelidonium majus</i>, <i>Thalictrum flavum</i>, <i>Coptis japonica</i>, <i>Papaver somniferum</i>, <i>Eschscholzia californica</i>, <i>Papaver bracteatum</i>, <i>Argemone mexicana</i>, <i>Glaucium flavum</i>, <i>Sanguinaria canadensis</i>, <i>Corydalis chelanthifolia</i>, <i>Nigella sativa</i>, <i>Jeffersonia diphylla</i>, <i>Berberis thunbergii</i>, <i>Mahonia aquifolium</i>, <i>Menispermum canadense</i>, <i>Tinospora cordifolia</i>, <i>Cissampelos mucronata</i>, <i>Cocculus trilobus</i>	
--------------------	-----	--	---	--

N-cyclopropylmethyltransferase	CPMT	Norcodeine→N(cyclopropylmethyl)norcodeine Normorphine→N(cyclopropylmethyl) normorphine Noroxycodone→N(cyclopropylmethyl) noroxycodone Noroxymorphine→N(cyclopropylmethyl) normoroxymorphine Northebaine→N(cyclopropylmethyl) northebaine Nororipavine→N(cyclopropylmethyl) nororipavine Norhydrocodone→N(cyclopropylmethyl) norhydrocodone Norhydromorphine→ N(cyclopropylmethyl)norhydromorphine Nordihydrocodeine→ N(cyclopropylmethyl)nordihydrocodeine Nordihydromorphine→ N(cyclopropylmethyl)nordihydromorphine Nor-14-hydroxycodeine→ N(cyclopropylmethyl)nor-14-hydroxycodeine Nor-14-hydroxymorphine→ N(cyclopropylmethyl)nor-14-hydroxymorphine Norcodeineone→ N(cyclopropylmethyl)norcodeineone Normorphinone→ N(cyclopropylmethyl)normorphinone Nor-14-hydroxy-codeinone→ N(cyclopropylmethyl)nor-14-hydroxycodeinone Nor-14-hydroxy-morphinone→ N(cyclopropylmethyl)nor-14-hydroxymorphinone	<i>Papaver spp.,</i> <i>Chelidonium majus,</i> <i>Thalictrum flavum,</i> <i>Coptis japonica, Papaver</i> <i>somniferum,</i> <i>Eschscholzia californica,</i> <i>Papaver bracteatum,</i> <i>Argemone mexicana,</i> <i>Glaucium flavum,</i> <i>Sanguinaria canadensis,</i> <i>Corydalis chelanthifolia,</i> <i>Nigella sativa,</i> <i>Jeffersonia diphylla,</i> <i>Berberis thunbergii,</i> <i>Mahonia aquifolium,</i> <i>Menispermum</i> <i>canadense, Tinospora</i> <i>cordifolia, Cissampelos</i> <i>mucronata, Cocculus</i> <i>trilobus</i>	
--------------------------------	------	--	---	--

Table 4. *O*-demethylase candidate enzymes

Name	Sequence
T6ODM	MEKAKLMKLGNGMEIPSVQELAKLTAEIPSRVVCANENLLPMGASVINDHETIPVIDIENLLSPEPIHGKLELD RLHFACKEWGFFQVNVHGVDSLVDSVKSEIQGFFNLSMDEKTKYEQEDGDVEGFGQGFIESEDQTLWDADIFMMF TLPLHLRKPFLFSKLPVPLRETIESYSSEMKKLSMVLFNKMEKALQVQAAEIKGMSEVFDGTQAMRMNYYPPCPQPN LAIGLTSHSDFGGLTILLQINEVEGLQIKREGTWISVKPLPNAFVVNVGDILEIMTNGIYHVSVDHRAVVNSTNERLSIATF HDPSLESVIGPISSLITPETPALFKSGSTYGDVLEECKTRKLDGKSFLDSMRI
CODM	METPILIKLGNGLSIPSVQELAKLTAEIPSRVCTNENLLMTMGASEIDDETVPVIDLQNLSPSPVVGKLELDKLSHAC KEWGGFFQLVNHGVDALLMDNIKSEIKGFFNLPMNEKTKYQGDGDFEGFGQPYIESEDQRLDWTVEFMSLSPLHLR KPHLFPELPLPFRETLESYLSKMKKLSVVFEMLEKSLQLVEIKGMTDLFEDGLQTMRMNYYPPCPPELVGLTSHS DFSGLTILLQINEVEGLQIRKEERWISIKPLPDAFIVNVGDILEIMTNGIYRSVEHRAVVNSTKERLSIATFHDSKLESEIG PISSLVTPETPALFKRGYEDILKENLSRKLKGKSLDYMIRM
PsP7ODM	MEKAKLMKLGNGLSIPSVQELAEITFAEVPSPRYVCTNENLLMTMGASEIDDETVPVIDLQNLSPSPVVGKLELDKLSHAC LHYSCKEWGFFQLVNHGVDALLVDHVKSEIHSFFNLPLNEKTKYQGRDGDVEGFGQAFLVSENQKLDWADMFFINTL PLHLRKPFLFPNLPLPRETIESYSSEMKKLSMVLFEMMGKAIEVIDIKEAITEMFEDGMQSMRMNYYPPCPQPERVI GITPHSDFDGLTILLQINEVEGLQIRKEDKWISIKPLPDAFIVNVGDILEIMTNGVHRSVDHVRGVNSTKERLSIATFHHS PKLELEIGPISSLIRPETPAVFKSAGRFEDLLKEGLSRKLKGKSLDCMRM
PsoDIOX1	MEKAKLMKLGNGMEIPSVQELAKLTAEIPSRVVCANENLLPMGASVINDHETIPVIDIENLLSPEPIHGKLELDRLHF ACKEWGFFQVNVHGVDSLVDSVKSEIQGFFNLSMDEKTKYEQEDGDVEGFGQGFIESEDQTLWDADIFMMFTLPLH LRKPHLFSKLPVPLRETIESYSSEMKKLSMVLFNKMEKALQVQAAEIKGMSEVFDGTQAMRMNYYPPCPQPNLAIGL TSHSDFGGLTILLQINEVEGLQIKREGTWISVKPLPNAFVVNVGDILEIMTNGIYHVSVD
PsoDIOX2	METAKLMKLGNGMSIPSVQELAKLTAEIPSRVICTVENLQLPVGASVIDDHETVPVIDIENLISSEPVTEKLELDRLHS ACKEWGFFQVNVHGVDTSLVDNVKSDIQGFFNLSMNEKIKYGGQKGDGVEGFGQAFVASEDQTLWDADIFMILTLP LRKPHLFSKLPVPLRETIESYSSEMKKLSMVLFEKMEKALQVQAVEIKEISEVFKDMTQVMRMNYYPPCPQPELAIGL PHSDFGGLTILLQINEVEGLQIKNEGRWISVKPLPNAFVVNVGDILEIMTNGMYRSVDHRAVVNSTKERLSIATFHDP NLESEIGPISSLITPNTPALFRSGSTYGELVEEFHSRKLKGKSLDSMRM
PbrDIOX2	METPKSIKLGGSLLVPSVQELAQSFAPVARYVRDDLEPLTDLSGVSMIDQTIPTVIDLQKLQSPVPIIRELESEKLSHAC KEWGGFFQVNVHGVDDILLVEKTKSEIKDFNLPMDKCKFWQEEGDIQFGGQAFVQSEDDQKLDWADIFLMTPLRHT RNPRLFPKLPPLRNTMDSYSSKLSKLASTLIEMMGKALHMETSVLAELFEDGRQTMRIYPPCPQPKDVIGLTPHS DGGGLTILLQINEVDGLQIRKEKIWIPIKPLPNAFVVNIGNILEIMTNGIYRSVEHRATIHSTKERLSVAAFHNPKVGV EIPVSMITPESPALFRTIYDDYGGKYFSRKLKGKSLDFMRIGEGDEENKAT
PbrDIOX3	METPKLIKLGGSLLVPSVLELTKQSPAEPVARYIRNDLEPMTDLSSASLTDQTIPTVIDLQNLSPPELELEKLSHSGCKE WGGFFQVNVHGVDDILLVEKVKSEIQGFFNLPIDEKNKFWQEEGDLEGYGKAFVHSEDEKLDWADMFFILTQPPQMRK PRVFPKLPRLRETIESYSLELSKLGTLTLLDMGKALQIETGVMSSELFEDGRQTMRIYPPCPQPEHVIGLTPHSDGG ALTILLQLNQVDGLQIRKEEIVWPIKPLPNAFVVNIGDILEIMSNVYRSVEHRATINSSKERLSVAIFQSPKHGTEIGPIL SMITPEAPALFKTIPYEDYLRKFFSRKLKGKSFVDSMRIGESDEDNNTA
PbrDIOX4	METQKQNFAGSLVPSVQELAKQSEVQPDYRIRSDQSSNTNISCPSMTDQIPVIDLQSLSPDPIIGELELERLHSHAC KEWGGFFQVNVHGVNDLLVEKVKSEIQGFFNLPMDKCKFWQEEGDFEGFGQAFVSEDDQKLDWGDVFFILTQPPQ MRKPRLPKLPPLPRKTIESYSLETNKLSTLLEMEKALKIETGVMTELFEGGIQRMRTYPPCPQPKHVIGLTPHS DPDALTILLQLNEVDGLQIRKEKIWIPIKPLSNAFVVNIGDILEIMSNIGYRSVEHRATVNSTKERLSVATFHSPRKDTE IGPILITPETPALFRTSGFEDYFRKFFAHLNKGKSLSSIRIGETDEGNAT
PbrDIOX5	MEAPKLIMLGGSFLVPSVQELAKQSLAEVPRVYVRDDQDTLGNINITPMSMIDQSIPTVIDLEKLSPEPIVGELELERL HSACKEWGFFQVNVHGVDSLLVEKVKSEIEGFFNLPMDKCKFWQEEGDIEGFGQVVFHSDQKLDWGDVFFILTQPPQ LPRHTRKPLRFPNLPLPLRQTIESYSSELSKLVTLVDLMGKALQMESGVLTELFENGIQRMRMNYYPPCPQPEQVIGL TPHSDVGGTLILLQLNEVDGLQIKKDKVWVPIKPLANAFVVNVGDILEIMSNIGYRSVEHRATINSTKERLSIATFHNP RADREIGPIPSMISPETPALFKTTGYEYFKKFFSRKLEGKSFLDSLRIREGDEHCGRLDVKGPCN
PbrDIOX6	MEIPNPIKIGSSLLVPSVQELAKQSFAPVARYIRNDVDPLITKLSVSLIDQTVPTVIDLQKLSPEPIV GELELERLHSHACKEWGFFQVNVHGVNDLLVEKVKSEIQGFFNLPMEKCKFWQEEGDFEGFGQMFMVQSEE QKLDWGDVFFILTQPPQHMRKPRLPFLKLPPLRETIESYSLEIKLGLTIKLMKALQIDAGVMAELFED GIHTMRMNYYPPCPQPEHVIGLTPHSDGGGLTILLQLNEVDGLQIRRENIWVPIKPLPNAFVVNIGDILE ILSNIGYRSVEHRSTVNATKERLSVATFQNPQKQESVIGPNMITPERPALFRKIVKYDMKKLSRKLKGK SFLDSLRIEGGDERP
PbrDIOX8	METLKTVPKGGSLFIPNGQELAKQSLAEVYVGNDDQDTMLLIGQTIPTVIDLQKLSPEPITGDMELDKLSH ACKEWGFFQVNVHGVDDILLVEKVKSEVHDFNIPMDKCKFWQEEGDIEGFGQVVFHSDQKLDWGDVFFILTQPPQ MVTLPKHMRRKPRFLKLPPLRETIESYSLSKLGVTLELMGKALQMEDRIMSELFDGGRQTMRMNYY PPCPQPEQVIGLTPHSDPGGLTILLENEVNGLRKENIWIPIPLPNAFIVNIGDILEIMSNIGYHVSVE HRATINSTKERLSVAMFNSPKVDTEIGPIHSMITPETPALFRTIGYDEYKIFFSRKLKGKSLLESMSKI
PbrDIOX10	MEAPKLIMLGGSFLVPSVQELAKQSLAEVPRVYVRDDQDTLGNINITPMSMIDQSIPTVIDLEKLSPEP IVGELELERLHSHACKEWGFFQVNVHGVDSLLVEKVKSEIEGFFELPVDEKCKFWQEEGDIEGFGQVVFH EDQKLDWADMFMMLTLPNMRKPRLPNLPLPLRQTIDSYSSELSKLVTLVDLMGKALQMESGVLTELF ENGIQRMRMNYYPPCPQPEQVIGLTPHSDVGGTLILLQLNEVDGLQIKKDKIWIPIKPLRNFVNVGDA LEIMSNIGYRSVEHRATINSTKERLSIATFHNPADREIGPIPSMISPETPALFKTTGYEYFKKFFSRK LEGKSFLDSLRIEGDEHCGRLXVKGXCN
PbrDIOX11	METPKLMKLGGSFLVPSVQELAKQSLAEVPRVYVRDDRMVGNINVTMPSMIDQSIPTVIDLEKLSPLD IVGELELERLHSHACKEWGFFQVNVHGVDSLLVEKVKSEIEGFFELPMDEKCKFWQEEGDAEGFAQFFVQS EDQKLDYSGDMFFMLNLPQHMRKPRFLKLPPLRETIESYSLSKLGVTLELMGKALQMEDRIMSEL

	FDDGRQTMRMNYYPPCPQPEQVIGLTPHSDPGGLTILLELNEVNGLIRKENIWWPIHPLPNAFIVNIGDI LEIMSNIGIYHSEHRATINSTKERLSVAMFNSPKVDTEIGPIHSMITPETPALFRITIGYDEYLLKIFFSRK LDGKSLLSMMKI
PbrDIOX13	METPKLRDFGSFLPVPSVQELAKQVLTPEIPRYIRTDLEALNKLSCASNTDQTVPIIDMQCLLSAEPEME LEKLHSACKEWGFFRVVNHGVDNLESVKSEIESFLNLPVNAKNKYGQKQDDQGFGSRFVLSSEEQKLDWG DFFYMVTRPLYLRKPHLFPPLPLRETIESYSSEVSKLAMALFEMMGKALKIETGVMTEIFEGGMQAMR MNYPPCPRPDLVIGLNAHSDFGGLTILLQLNEVEGLEIRNKGEWVSVKPLANAFVNVGDMVIELTNGI YHSEHRATINSSKERLSVATFHYPKLETGIGPLPCMITPKTPALFGRIERYELLRKYYARKLNGKSTL DCMRIGNGFEDDNTA
PbrDIOX18	MEAPKLIMLGGSLFVPSVQELAKQSLAEVPARYVRDDQDTLGNINITPMSMIDQSIPIVIDLEKLSP IVGELELERLHSACKEWGFFQVNVHGVDSLLVEKVKSEIEGFFELPVDEKKKFWQEEGDIEGFGQIFVHS EDQKLDWADMFMFMTLPPNMRKPRLFPNLPPLRQTIDSYSSELSKLVLTLVDLMGKALQMESGVLTELF ENGIQMRMNYPPCPQPEQVIGLTPHSEVGGTILLQLNEVDGLQIRKEIWWPIKPLSNFIVNIGDI LEIMSNIGIYHSEHRATVNSTKERLSVATFHSRPRKDTEIGPILITPETPALFRITSGFEDYFRKFFAHKLN GKSFLSSIRIGETDEGNAT
PbrDIOX19	MSMIDQSIPIVIDLEKLSPPIVGELELERLHSACKEWGFFQVNVHGVDSLLVEKVKSEIEGFFELPVDE KKKFWQEEGDIEGFGQIFVHSEDQKLDWADMFMFMTLPPNMRKPRLFPNLPPLRQTIDSYSSELSKLVL TLVDLMGKALQMESGVLTELFENGIQMRMNYPPCPQPEQVIGLTPHSDVGGTLILLQLNEVDGLQIRK EKIWWPIKPLSNFIVNIGDILEIMSNIGIYHSEHRATINSTKERLSVAMFNSPKVDTEIGPIHSMITPE TPALFRITIGYDEYLLKIFFSRKLDGKSLLSMMKI
PbrDIOX21	METPKLVKSSGSLFLSTSVQELAKQSLPEVPARYIRTNLEPLSNVSGDSQSVPIIDLQKLLSSEPIGE LELDKLHSACKEWGFFQVNVHGVDSLLVEKVKSEIEGFFELPVDEKKKFWQEEGDIEGFGQIFVHS LDWGDTFGMFTLPIHMRNPRLFPPLPLRETIESYSLSVRLALALIGLMEKALKIKTSAMSELFEDGG QAMRMNYPPCPQPEHIGLTPHSDAGGLTILLQLNEVDGLQIKDKIWWPIKPLSNFIVNIGDILEIM TNGIYHSEHRATINSSKERLSVAAFHSPKGDTLIGPMVSLITPETPALFRITIGYQDYMKKFSRKLDDGK SLVNSMRIGEGDEDK
PbrDIOX- ZSNV- 2004018	METPTLMKLGNSVPSVQELAKATLAEIPSRYICTDENLLTMGASTTDNETVPVIDLQNLSPPEVIGMLELDRHSA CKEWGFFQLVNHGVDALLVDNEVQGFNLPMDCKTYGQKDGDEGFGQFFVISEDQKLDWADVFMSTLPLHSRK PHLFPPLPLRETIESYSSEMKKLSMVLDFMMGKALQVVEIKGITELFEDGAQQIRMNYPPCPQPELVFGLTSHSD FDGLTILLQLGEVEGLQIKKEERWISIKPLDAFIVNVGDILEIMTNGIYHSEHRATINSTKERLSVAMFNSPKVDTEIGPIHSMITPE PISSLITPETPALFRKGVFEDLLKEMFLRKLDDGKSFLDCMRM
PrhDIOX- MTX- 2001522	GNGLSVPSVQELAKQTLAEIPSRYICTDENPLITGASVDDDETVPVINLQNLSPPEVIGKLELDKLHSACKEWGFFQV NHGVNDSLVDSVKSEIEGFFNLPANEKLYGQKDGDEGFGQHFVISEDQKLDWADVFMVTLPLVRLRKPPLFP LPLRDTLDSYSSELNKLMSVLEMMEKALKLVECKGITDFFEDGFGQMRMNYPPCPRPPELVFGLTSHSDFGGLTILL QLNDVEGLQIKKEERWISIKPLPNAFIVNIGDILEIMSNIGIYHSEHRATINSTKERLSVAMFNSPKVDTEIGPIHSMITPE ETPALFRKGVFEDLLKEMFLRKLDDGKSFLDCMRM
PseDIOX- JSVC- 2005842	LMKLANGMSVPIVQELAKLTVEIPSRYICTDGNLLTMGASVIDYETVPVIDLQNLQSREPVEKLELDRHSA FFQLNHGVDASLMDNVRSEIRGFFNLPISDKMKYQKDGDEGFGQHFVISEDQKLDWDAFMMFTLPLHSRNPRL TPEFPQPLRETIESYSSEMKKLSVLLFELMEKALQVKGITMFEDGLQIRSNYPPCPRPPELAIGLTSHSDFDGLTILL QLNEVEGLQIKKEERWISIKPLPNAFIVNVGDVLEVMNTNGIYHSEHRATINSTKERLSVAMFNSPKVDTEIGPIHSMITPE ETPALFRKGRFKDLLENLSTKLDGKSFLDCMRM
CYP2D6	MGLEALVPLAVIVAIFLLVLDLHRRQRWAARYSPGPLPLPGLGNLLHVDQNTPYCFDQLRRRFGDVFSQLAWTP VVVLNGLAAVREALVTHGEDTADRPPVPITQILGFGPRSQGVFLARYGPAWREQRFRFSVSTLRNLGLGKKSLEQWVT EEAACLCFAFANHSRPPFRPNGLLDKAVSNVIALTCGRREFYDDPRFLRLDLAQEGLKEESGFLREVNLNAVPLLHI PALAGKVLRFQKAFLTQLDELLTEHRMTWDPAQPPRDLTEAFLAEMEKAKGNPESSFNENLRIVVADLFSAGMVT TSTTLAWGLLLMILHPDVQRRVQGEIDDVIGQVRRPEMGDQAHMPYTTAVIHEVQRFQDIVPLGVTHMTSRDIEVQG FRIPKGTTLITNLSSVLKDEAVWEKPFRRHPEHFLDAQGHFVKPEAFLPFSAGRRACLGELARMELFFFTSLLQHFS FSVPTGQPRPSHHGVFAFLVTPSPYELCAVPR

Table 5. *N*-demethylase candidate enzymes

Name	Sequence
BM3	MTIKEMPQPKTFGELKNLPLNTDKPVQALMKIADDELGEIFKFEAPGRVTRYLSSQRLIKEACDESFRDKNLSQAAKF ARDFAGDGLVTSWTHEKNWKKAHNILLPSFSQQAMKGYHAMMVDIAVQLVQKWERLNADHEHIEVSEDMTRLTLD IGLCGFNYRFNSFYRDQPHPHISMVRAADEVMNKLQRANPDDPAYDENKRQFQEDIKVMNDLVDKIADRARGEQ SDDLTLQMLNGKDPETGEPLDDGNIRYQITFLIAGHETTSGLLSFALYFLVKNPHVLQKVAEEAARVLVDPPVSYKQV KQLKYVGMVLNEALRLWPTAPAFSLYAKEDTVLGGEYPLEKGDEVMVLIPQLHRDKTVWGDDVEEFRPERFENPSAI PQHAFKPPFNGQRACIGQQFALHEATVLGMMLKHFDDEDHTNYELDIKETLTLKPKGFVVAKSKKIPLGGIPSPST EQSAKKVRKKAENAHNTPLLVLYGSNMGTAEGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHP PDNAKQFVDWLDQASADEVKGVRSVFGCGDKNWATTYQKVPFIDETLAAKGAENIADRGEADASDDFEGTYEE WREHMWSDVAAYFNLDIENSEDNKSTLSLQFVDSAADMPLAKMHGAFSTNVVASKELQPPGSARSTRHLEIPLKE ASYQEGDHLGVIPRNYEGIVNRYTARFGLDASQIRLEAEEEEKLAHLPLAKTVSVEELLQYVELQDPVTRTQLRAMAA KTVGPPHKVLEALEKQAYKEQVLAKRLTMLELLEKYPACEMKFSEFIALLPSIRPRYSSISSPRVDEKQASITVSVVS GEAWSGYGEYKGIASNYLAELQEGDTITCFISTPQSEFTLPKDPETPLIMVGPVGTGAPFRGFVQARKQLKEQGQSLGE AHLVYFGCRSPHEDYLYQEELENAQSEGITLHTAFSRMPNPKTYVQHVMEQDGGKLIELLDQGAHFYICGDSQMAP AVEATLMKSYADVHVQVSEADARLWLQLEEKGRYAKDVWAG
CYP3A4-1	MALIPDLAMETWLLAVSLVLLYLYGTHSHGLFKKLGIPGPTPLPFLGNILSYHKGFCMFDMECHKKYGVWGFYDQ QQPVLAITDPMIKTVLVKECYSVFTNRRPFGVGMKSAISAEDEEWKRLRSLLSPTFTSGKLKEMVPIIAQYGDVL VRNLRREAETGKPVTLKDVFGAYSMDVITSTSGVFNIDSLNPNQDPFVENTKLLRDFDLPFFLSITVFFLIPILEVL NICVFPREVTNFRKSVKRMKESRLEDQKHRVDFLQLMIDSQNSKETESHKALSDLELVAQSIIIFAGYETTSSVLSFI MYELATHPDVQQLQEEIDAVLPNKAPPTYDTVLQMEYLDMMVNETLRLFPAMRLERVCKKDVEINGMFIPKGVV MIPSYALHRDPKYWTEPEKFLPERFSKKNKDNDIPYIYTPFGSGPRNCIGMRFALNMKALIRVLQNFSEFKPCKETQI PLKLSLGLLQPEKPVVLKVESRDGTVSGA
CYP3A4-2	MALIPDLAMETWLLAVSLVLLYLYGTHSHGLFKKLGIPGPTPLPFLGNILSYHKGFCMFDMECHKKYGVWGFYDQ QQPVLAITDPMIKTVLVKECYSVFTNRRPFGVGMKSAISAEDEEWKRLRSLLSPTFTSGKLKEMVPIIAQYGDVL VRNLRREAETGKPVTLKDVFGAYSMDVITSTSGVFNIDSLNPNQDPFVENTKLLRDFDLPFFLSIIFFLIPILEVLNI CVFPREVTNFRKSVKRMKESRLEDQKHRVDFLQLMIDSQNSKETESHKALSDLELVAQSIIIFAGYETTSSVLSFIM YELATHPDVQQLQEEIDAVLPNKAPPTYDTVLQMEYLDMMVNETLRLFPAMRLERVCKKDVEINGMFIPKGVV MIPSYALHRDPKYWTEPEKFLPERFSKKNKDNDIPYIYTPFGSGPRNCIGMRFALNMKALIRVLQNFSEFKPCKETQIP LKLSLGLLQPEKPVVLKVESRDGTVSGA
McaCYP82-4	MIMMFIDYSSWLPQTLQLSILLAVSLVIFINLFLTRRSYSSKSHNTIHPKAAGALPVIGHLYTLF RGLSAGVPLRQLDAMADRYGPAFIIHLGVYPTLVVTCRELAKECFTTNDQTFATRPSTCAGKYIGYNYA FFGFAPYGPYWREARKIATVELLSNYRLDSLHVREAEGVRNVDLYALHASSTNKNMMKIDMKQWFD QVTNLVILMMVVGKRCVTTGGNEEEVRVVKVLEHFFKHLGTLVSDVVPYVEWMDLDGNIGRMKSTAKEL DCILGRWLEEHRRERSDFMDAMLAMVEGKIPIYDSDTVIKAICLNLLNAGSDTLGITMTWALSLLN RHVLKVKDELDVHVGNRQVEELDVKNLVYLHAVVKETLRLFPAPLGPVHEAMEDCTVGFGFHVAKGTR LVVNVWKLHRDPSVWSDPLAFKPERFLDNTVDVRGQHFQQLLPFGSGRRGCPGITFALQVAHLTLARLLH GFEWDTPDGAPVDMSEVSVLTAKKNPVEVLTFRPLPAEYVTQN
NsaCYP82-4	MLSIHDSMTVFLQLQAICGIFGFIHTWVTRWSSNKMKAPEVAGAWPVIGHLHLLGGGRPLYQLLGD SDKYGPAFTLRMGIQKALVSSWEVAKELTTNDRALATRPSSAGGKYMGNYNALIPFSPYGPYWRDMRK IATLELLSNHRLEELKHVREMEINTCISDMYKLCQVEDGVEIKPISVDLSQWFADLTNNVVMITGKRY IGSTDAGDMNEIRHQAALVKFMRLLRISLVDVFPVLQWVINYGGFKGVKSTARDIDSVLENWLQEHQR KRLSPDFNGNHDFIDVMISTLEGTEFSDYDHTIHKAIISMAMVVGTDTTTTLIIWAILLNNPNAMKK VQEELEIIVGKERNVDGSDIQHLVYLQAVVKETLRLYPVPLSYMHQAMEDCVIGSYNIQAGTRVLFNLW KLHRDSSVWSDPLEFRPERFLTSHVDVVRGQHFELIPFGSGRRSCPGISFALQVIHLTIARLFHGFNL TPGNSSVDMSEISGATLSKVTPLEVLVTPRLSSKLYN
HcaCYP82-10	MDSLLQLQIIGALALIFTYKLLKVICRSPMTDGMPEPPGAWPIIGHLHLLGGQDPIARTLGVMTDKY GPILKRLGVHTGLVSVSNWELAKECFTTNDRLASRPMGAAGKYLGYNYAIFGLAPHGPVWSEVRKIVLR ELLSNQSLEKLKHVRSEINTCLKNLFSLNNGNTPIKVDMMQWFERPMFNVTMMIAGKRYFSMENDNEA MNFRRKATEFMYLTGVFVSDALPYLEWLDLQGHVSAMKRTAKELDIHVKGWLEEHRAKLLGETKNEDD FVDVLLTILPEDLKDNTYIHDRDTIATATLALFLAASDTTATLTWALSILNPNPDLKRAQDELDDKH VGKEKLKVESDIINLVYLQAIKETLRLYPAAPLLPHEAMEDCTVGGYHVPKGTRIFVNIWKLQRDPRV WFDPNFRPERFLTTHANVDFKGQHFYIYFSSGRRVCPGITFTSTQIMHLTLAHLHFEFNIPTTKSNAG VDMTESLGITMPKATPLEVLTPLRLPSNLNYQYRD
EcaCYP82-7	MNLIFFQFLQFQVLVGLSVLLAFSYLWVSKNPKNKFKGKGALLAPQAAGAWPIVGHLPQLVGPCKPL FRILGAMADNYGPIFMLRFGVHTPVVSSWEMTKECFTTNDRLASRPSNAASQYLIYEVYALFGPSLYG SSYWRDARKIATLELLSHRLELLKHVPYTEIDTCIKQLHRLWTKNNKNQNNPELKVMNQFFTDLTMTNV ILKLVVGKRRFFNVDDAADHEKEEARKIQGTIFEFFKLTGEGYSAGALPLLNWLDLNGQKRAMKRTAKKMD SIAEKLLDEHRQKRLSKEGVKGTDHNDFMDVLLSILDADQGDYSHHPFNYSRDHVIKATTLMSLLSSMS ISVLSWALSLLNRRHVLKKAQDELDMNKGDRQVEEGDIKNLVYLQAIKVTFRMYPANPLLPHEAI EDCKIGGFNVPACTRVVVNAWKLQHDPRVWSNPSEFKPERFLNDQAAKVVDRGQNFYELPFGSGRRVCP GISPSLQTIHMSLARLVQAFELGTPSNERIDMTEGSLTMPKTTPLHVLLNPLPLPLYE
GfiCYP82-8	MELINSLEIQPITISILALLTVSILLYKIIWNHGSRKNNKSNKNNRKTSSAGVVEIPGAWPIIGHLHLF NGSEQMFHKLGLSLADQYGPAPFFIRFGSRKYVVVSNWELVKTCFTAQSQIFVSRPMLAMNIFFPKDSL SYIQHGDHWRELKISSTKLLSSHRVETQKHLIASEVDYCFKQLYKLSNNGEFTLVRLNTWCEDMALNVH

	VRMIAGMKNYVAAPGSGEYGGQARRYRKALEEALDLLNQFTITDVPVWLGWLDHFRDVPVGRMKRCGAELD SIFATWVEEHRVKRASGKGGDVEPDFIDLWESMEQLPGNDPATVIKLMCKEHIFNGSGTSSLTAWILS LIMNNPYVIKKAREELEKHVGNHRQVEESDLPNLLYQAIKEGMRLYTPGPFIDRNTTDEYINGVHIP AGTCLYVNLWKIHRDPNVYEDPLEFKPERFLKNNSDLDLKGQNYQLLPFGAGRRCIPGVSALPLMYLTV SRLIHGFDMLPKGVEKADMTAHGGVINQRAYPLEVLLKPRLTFQQA
SdiCYP82-3	MTIGALALLSFIYFLRVSVIKRTKYTNTAVTATNKLLENDEDEANHSKRVPVAPPEVAGAWPILGHLPLQVVG LKQPLFRVLGDMADKYGPIFIVRFGMYPTLVVSSWEMAKECFTTNDRLASRPASASGKYLTNYAMFGF TNGPYWREIRKISMLELLSHRRVELLKHPSTEIDSSIKQLYHLWVENQNNKQGDHVKVDMSQLLRDL TLNIVLKL VVGKRLFNNDMDHEQDEAARKLQKTMVELIKVAGASVADALPFLGWLDVDGLKRTMKRIA KEIDVIAERWLQEHQRKLTSDKGGSNNIQGGGGDNDFMDVMSILDDDSNFFINYNRDTVIKATSLTM ILAGSDTTTSLTWALTLLATNPGALRKAQDELDTKVGRDRQVDERDIKNLVYLQAIVKETLRMPYAAPL AIPHEATQDCIVGGYHVTAGTRVWVNLWKLQRDPHAWPNPSEFRPERFLAVENDCKQQTCDGEANMDF RGQHFEYMPFGSGRRMCPGINFAIQIHHMTLARLLHSFELRVPEEEVIDMAEDSGLTISKVTPLELLTLP RLPLPLYI
SdiCYP82-6	FCQFQGVIGILLAFLLTFLYLWRASITGLRTKPKHNDKFKVTKAAPEADGAWPIVGHFAQFIGRPLFRIL GDMADKYGSIFMVRFGMYPTLVVSSWEMAKECFTTNDRLASRPASAAGKYLTDFAMLSFSFYGPYWRE IRKISMLELLSHRRVELLKHPSTEIDSSIKQLYHLWVENQNNKQGDHVKVDMSQLLRDLTLNIVLKL VVGKRLFNNDMDHEQDEAARKLQKTMVELIKVAGASVADALPFLGWLDVDGLKRTMKRIAKEIDVIAE RWLQEHQRKLTSDKGGSNNIQGGGGDNDFMDVMSILDDDSNFFINYNRDTVIKATSLTMLAGSDTT TSLTWALTLLATYPLCALRKAQDELDTKVGRDRQVDERDIKNLVYLQAIVKETLRMPYAAPLAIPHEAT QDCIVGGYHVTAGTRVWVNLWKLQRDPHAWPNPSEFRPERFLAVENDCKQQTCDGEANMDFRGQHFEY MPFGSGRRMCPGINFAIQIHHMTLARLLHSFELRVPEEEVIDMAEDSGLTISKVTPLELLTLPRLPLPLY I
CmaCYP82-6	MDLFIFFSRFQYIVGLLAFLLTFFYYLWRVSITGTRIKTNNMNGTNNMAPEAAGAWPIVGHLPQLVGPQ PLFKILGDMADKYGSIFMVRFGMHPTLVVSSWEMAKECFTTNDKFLASRPTSAGGKYLTDFAMFGFSFY GPYWREIRKISTLELLSHRRVELLKHPSTEIGGSIKQLYKLWMETQNNKQQRDDHVKVDMSQLVFGYLT LNTVLKL VVGKGLFNNDMDNHEQEERKLTHTVLEFFKLAGVSVADALPFLGWLDVDGQKRSMKRIAKE MDLIAERWLQEHQRKLTSDKGGSNNIQGGGGDNDFMDVMSILDDDSNFFINYNRDTVIKATSLNLILAASDTSV SLTWVLSLLVTNPGALKKQVQDELDTKVGRNRHVEERDIEKLVYLQATVKETLRMPYAGPLSVPHEATQDC TVGGYQVTAGTRLVNVWKLQRDPVWPNPSEFKPERFLPDGCEVGCGEANMDFRGQHFEYIPFGSGRR MCPGIDFAIQIHHMTLACLLHAFEFQVPSLDRHVPVAVIDMSEGSGLTMPKVTPLEVLLNPRPLPLYE L
EcaCYP82-5	MEKPILLQLQPGILGLLALMCFLLYYVIVKLSLSTRNCNLVRHPPEAAGSWPIVGHLPQLVSGKPLFRVL GDMADKFGPIFMVRFGVHPTLVVSSWEMAKECFTSNDKFLASRPPSAASIYMAIDHAMLGSSYGPYWRE IRKISTLHLLSHRRLELLKHVPHLEIHNFIKGLYGIWKDHQKQQQPTARDDQDSVMLEMSQLFGYLTN IVLSLVGKRVKNYHADGHLDDGEEAGQKQLHQTITDFFKLSGVSVASDALPFLGLFDLDGQKKMKRV AKEMDFVAERWLQDKKSSLLSSKSNKQNEAGEGDVDDFMDVLMSTLPDDDDSFFTKYSRDTVIANSL SMVAGSDTTSVSLTWALSLLNNIQLRKAQDELDTKVGRDRHVEEKIDIDLNVYLQAIVKETLRMPYAG PLSVPHEAIEDCNVGGYHIKTGTRLLVNIWKLQRDPVWSNPSEFRPERFLDNQSNGLTLLDFRGQHFEYI PFGSGRRMCPGVNLATPILHMTLARLLQSFDLTPSSSPVDMTEGSGLTMPKVTPLVLLTLPRLPLPLYD Y
PbrCYP82-5	MDVAIHDHHLQPFVSIAGLLALLSFFYCIWVFHPRHKSNDERKLSPPSPPEVAGAWPIVGHLPQ LIGSTPLFKILADMSNKYGPIMVRFGMYPTLVVSSWEMSKECFTTNDRLFATRPSSAAGKYLTALFAF SVYGPYWREIRKISTIHLLSLRRELLKHGRYLEIDCKMKRLEFYWMEHHKNIHSTTSVKNMSQVFAE LSLNVVLKIIVGKTLFIKNGNEDYTKEEEEGQKLHKTILKFMELAGVSVASDVLPFLGWLDVDGQKKQMK RVYKEMNLIAASKWLGEHRERKRLQIHKRGAARGSNYDDGNDFMDVMSILDEENDLFFGYSDTVIKS TCLQLIVAASDTSAMTWALSLLTNPVNLQKAQDELDTKVGRDRHIEHDIECLVYLQAIVKETLRLY PPAPLSLPHEAMEDCTVGGYQVQAGTRLVNVWKLQRDPVWSNPSEFKPERFLPQSDGGFGGEEARMDF RGQHFEYTPFGSGRRICPGIDFQLTVHMLARLLQAFDFNTAGGLVIDMVEGPGLTMPKVTPLEVHLNP RLPVTLY
PbrCYP82-6	MQVDWPNILQKYPIITCSLLTLLSFFYYIWSITKPSRNSKTKLPPPEVAGSWPIVGHLPQLVGSTPLFK ILANMSDKYGPIMVRFGMHPTLVVSSWEMSKECFTTNDKFLASRPPSASAKYLGYNAMFVSDYGPYW REIRKISTLQLLTHKRDLKNIPLYEINSCVKTLYTRWAKTQSQIKQNVGGAADDVFKVDMTEMFGHLN LNVVLRVLVVGKPIFIQKDNADEDYTKDGHNEELGQKLHKTIEFFELAGASVADVLPLYGWLDVDGQK KRMKKIAMEMDLFAQKWLEHRQKGINHDNENDFMAVLISVLGEGKDDHIFGYSRDTVIAA TDTTLVSLTWALSLLLTNPVLSKAQDELDTVVGKERNVEDRDVNLVYLQAIKETLRLYPPSPPLAVPH EAIENCNVGGYEVKARTRLVNLWKIHRDPVWSNPSEFKPERFLPKLDGGTGEASKLDFKGQDFVYTPF GSGRRMCPGINFASQTLHMTLARLLHAFDFDIESNGLVIDMTEGSGLTMPKVTPLVHLRPLPATLY
McaCYP82-4	MIMMFIDYSSWLPQTLLQSILLAVSLVIFINLFLTRRSYSSKSHTHNIHPPKAAGALPVIGHLYLTF RGLSAGVPLYRQLDAMADRYGPAFIHGLGVYPTLVVTCRELAKECFTTNDQTFATRPSTCAGKYIGYNYA PFGFAPGYWREARKIATVELLSNYRLDSLHVREAEGVRNVDLYALHASSSTNNKQNMKIDMKQWFD QVTLNVLMMVVGKRCVTTGGNEEVVRVVKLHEFFKHLGTLVSVDVVPYVWMDLDCNIGRMKSTAKEL DCILGRWLEEHRRERSDFMDAMLAMVEGIKIPYYSDTVIAICLNLLNAGSDTLGITMTWALSLLNN RHVLKVKDEL DVHVGKNRQVEELDKNLVYLHVVKETLRFPAPLGPVPEAMEDCVVGGFHVAKGTR LVNVWKLHRDPVWSDPLAFKPERFLDNNTVDVRGQHFQLLPFGSGRRGCPGITFALQVAHLTLARLLH GFEWDTDPDGPVDMSEVSVLTAKKNPVEVLFTRPLPAEVYTQN
NsaCYP82-4	MLSIHDSMTVFLQLQAICGIFGFIITWWTRWSSNKMKAPEVAGAWPIVGHLLGGGRPLYQLLGDM

	SDKYGPAFTLRMGIQKALVVSSWEVAKCLTTNDRALATRPSSAGGKYMGYNNALIPFSPYGPYWRDMRK IATLELLSNHRLEELKHVREMEINTCISDMYKLCQVEDGVEIKPISVDLSQWFAFLTFFNVVMMITGKRY IGSTDAGDMNEIRHFQAALVKFMRLRLISLLVDVFPVLQWINYGGFKGVMKSTARDIDSVLLENWLQEHQR KRLSPDFNGNHDVIDVMISTLEGTEFSDYDHTIKAISMAMVVGTDTTTTLIIWAISLLNNPNAMKK VQEELEIHVGKERNVDGSDIQLVYLQAVVKETLRLYPVPLSVMHQAMEDCVIGSYNIQAGTRVLFNLW KLHRDSSVWSDPLEFRPERFLTSHVDVVRGQHFEIPFGSGRRSCPGISFALQVIHLTIARLFHGFNL TPGNSSVDMSEISGATLSKVTPLEVLTPRLSSKLYN
HcaCYP82-10	MDSELLQIIGALAAIFTYKLLKVICRSPMTDGMFAPEPPGAWPHIHLHLGGQDPIARTLGVMDBKY GPILKRLGVHTGLVVSNEWELAKECFTTNDRLASRPMGAAGKYLGYNYAIFGLAPHGPYWSEVRKIVLR ELLSNQSLKELKHVRSEINTCLKNLPSLNNGNTPIKVDMMQWFERPMFNVVTMMIAGKRYFSMENDNEA MNFVKVATEFMYLTGVFVVSADALPYLEWLDLQGHVSAMKRTAKELDIHVKGWLEEHRRAKLLGETKNEDD FVDVLLTILPEDLKDNQTYIHDRDTIHKATALALFLAASDTTATITLWALSILNNPDVLRKRAQDELDKH VGKEKLVKESDIHNLVYLQAIKETLRLYPAAPLLPHEAMEDCTVGGYHVPKGRTRIFVNIWKLQDRPRV WFDPNFEFRPERFLTTHANVDKQGHFEYIPFSSGRRVCPGITFSTQIMHLLAHLLHEFNIVTPTKSNAG VDMTESLGITMPKATPLEVLLTPRLPSNLYNQYRD
EcaCYP82-7	MNLLIFFQFLLQFQVLVGLSVLLAFSYLVVSKNPKINKFKGKGALLAPQAAGAWPIVGHLPQLVGPKPL FRILGAMADNYGPIFMLRFGVHTVVVSSWEMTKECFTTNDRLASRPSNAASQYLIYEVYALFGFSLYG SSYWRDARKIATLELLSHRRLELLKHVPYTEIDTCIKQLHRLWTKNNKNQNNPELVKEMNQFFDILTMMNV ILKLTVGKRFFNVDDAADHEKEEARKIQGTIFEFFKLTGEGSVSAGALPLLNWLDLNGQKRAMKRTAKKMD SIAEKLLDEHRQKRLSKEGVKGTHDNDMDVLLSILDADQGDYSHHPFNYSRDHVIKATLLSMILSSMS ISVSLSWALSLLNNRHVLKKAQDELDMMNVGKDRQVEEGDIKNLVYLQAIKETFRMPANPLLPHEAI EDCKIGGFNVPAGTRVVVNAWKLQHDPRVWSNPSEFKPERFLNDQAAKVVDVRGQNFYELPFGSGRRVCP GISFSLQTIHMSLARLVQAFELGTPSNERIDMTEGSLTMPKTTPLHVLLNPRLPLPLYE
GHCYP82-8	MELINSLEIQTITISILALLTVSILLYKIHWNHGSRKNNKSNKNNRKTSSSAGVVEIPGAWPHIHLHLF NGSEQMFHKLGLSLADQYGPAPFFIRFGSRKYVVSNEWELVKTCTAQSQIFVSRPMLAMNIFFPKDSL SYIQHGDHWRRLRKISSTKLLSSHRVETQKHLIASVVDYCFKQLYKLSNNGEFTLVRLNTWCEDMALNVH VRMIAGMKNYVAAPGSGEYGGQARRYRKALEEALDLLNQFTTDDVVPWLGWLDHFRDVVGRMKRCGAELD SIFATWVEEHRVKRASGKGGDVEPDFIDLCWESMEQLPGNDPATVIKLMCKEHIFNGSGTSSLTAWILS LIMNNPYVKKAREELEKHVGNHRQVEESDLPNLLYIQAIEGEMRLYTPGPFIDRNTTEDYEINGVHIP AGTCLYVNLWKIHRDPNVYEDPLEFKPERFLKNNSDLDLKGQNYQLPFGAGRRICPGVSLALPLMYLTV SRLIHGFDMLKPKGVEKADMTAHGGVINQRAYPLEVLLKPRLTQQA
SdiCYP82-3	MTIGALALLSFIYFLRVSVIKRTKYTNATVATNKLLENDEANHSKRVPVAPPEVAGAWPIVGHLPQLVG LKQPLFRVLGDMADKYGPIFIVRFGMYPTLVSSWEMAKECFTTNDRLASRPASAGKYLTNYAMFGF TNGPYWREIRKISMLELLSHRRVELLKHVPSTEIDSSIKQLYHLWVENQNNQKQGDHVKVQVMSQLLRLD TLNIVLKLTVGKRLFNNDMDHEQDEAARKLQKTMVELIKVAGASVSDALPFLGWLDVLDGLKRTMKRIA KEIDVIAERWLQEHQRKLTSTNDKGGSNNIQGGGDNDFMDVMSILDDDSNFFINYNRDTVIKATSLTM ILAGSDTTTSLTWALTLLATNPGALRKAQDELDTKVGRDRQVDERDIKNLVYLQAIKETLRMYPAPL AIPHEATQDCIVGGYHVTAGTRVWVNLWKLQDRPHAWPNPSEFRPERFLAVENDCKQGTCDGEAANMDF RGQHFEYMPFGSGRRMCPGINFAIQIHHMTLARLLHSFELRVPEEEVIDMAEDSGLTISKVTPLELLLT RLPLPLYI
SdiCYP82-6	FCQFQGIIVGILLAFITFLYYLWRASITGLRTKPKHNDFKVTKAAPEADGAWPIVGHFAQFIGRPLFRIL GDMADKYGSIFMVRFGMYPTLVSSWEMAKECFTTNDRLASRPASAGKYLTIDFAMLSFSFYGPYWRE IRKISMLELLSHRRVELLKHVPSTEIDSSIKQLYHLWVENQNNQKQGDHVKVQVMSQLLRLDTLNIVLKL VYGKRLFNNDMDHEQDEAARKLQKTMVELIKVAGASVSDALPFLGWLDVLDGLKRTMKRIAEIDVIAE RWLQEHQRKLTSTNDKGGSNNIQGGGDNDFMDVMSILDDDSNFFINYNRDTVIKATSLTMILAGSDTT TSLTWALTLLATYPLCALRKAQDELDTKVGRDRQVDERDIKNLVYLQAIKETLRMYPAPLAIPEAT QDCIVGGYHVTAGTRVWVNLWKLQDRPHAWPNPSEFRPERFLAVENDCKQGTCDGEAANMDFRGQHFEY MPFGSGRRMCPGINFAIQIHHMTLARLLHSFELRVPEEEVIDMAEDSGLTISKVTPLELLLTTPRLPLPLY I
CmaCYP82-6	MDLFIFFSRFYIVGLLAFLTFYLYLWRSITGTRIKTQNNIMNGTNMMAPEAAGAWPIVGHLPQLVGPQ PLFKILGDMADKYGSIFMVRFGMHPTLVSSWEMAKECFTTNDKFLASRPTSAGGKYLTIDFAMFGFSFY GPYWREIRKISTLELLSHRRVELLKHVPYTEIGGSIKQLYKLWMETQNNQKQRDDHVKVQVMSQVFGYLT LNTVLKLTVGKGLFNNDMDNHEQEGRKLHETVLEFFKLAVSVASDALPFLGWLDVLDGQKRSMKRIAKE MDLIAERWLQEHQRKLTSTNNKASSGHDDFMSVLLSILDDDSNFFINYNRDTVIKATSLNLILAASDTTSV SLTWVLSLLVTNPGALKKVQDELDTKVGRNRHVEERDIEKLVLQATVKETLRMYPAGPLSPHEATQDC TVGGYQVTAGTRLVVNVWKLQDRPRVWPNPSEFRPERFLPDGCEVGCGEAANMDFRGQHFEYIPFGSGRR MCPGIDFAIQIHHMTLACLLHAFEFQVPSSLDKHLVPAVIDMSESGLTMPKVTPLEVLLNPRLPLPLYE L
EcaCYP82-5	MEKPILLQQLPGILGLLALMCFLYYVIVKVSLESTRNCNQLVRHPPEAAGSWPIVGHLPQLVSGKPLFRVL GDMADKFGPIFMRFGVHPTLVSSWEMAKECFTSNDKFLASRPPSAASIYMAVDHAMLGFSSYGPYWRE IRKISTLHLLSHRRLELLKHVPHEIHNFIFKGLYGIWKDHQKQQQPTARDDQDSVLMESQLFGYLTNL IVLSLVGKRVCNHYHADGHLDDGEEAGQQQLHQTTIDFFKLSGVSVASDALPFLGLFDLDGQKKIMKRV AKEMDFVAERWLQDKKSSLLSSKSNKQNEAGEGDDVDFMDVLMSTLPDDDDSFFTKYSRDTVIKANS SMVYAGSDTTVSLSLWALSLLNNIQLRKAQDELDTKVGRDRHVEEKDIDNLVYLQAIKETLRMYPAG PLSVPHEAIEDCNVGGYHIKTGTRLLVNIWKLQDRPRVWSNPSEFRPERFLDNQSNGLTLDFRGQHFEYI PFGSGRRMCPGVNLATPILHMTLARLLQSFDLTTPSSSPVDMTEGSLTMPKVTPLEVLLNPRLPLPLYD Y

PbrCYP82-5	MDVAIIVDHHYLQPFVSIAGLLALLSFFYCIWVFIRPRIKSNLDERKLSPPSPPEVAGAWPIVGHLPQ LIGSTPLFKILADMSNKYGPIFMVRFGMYPTLVVSSWEMSKECFTTNDRLFATRPPSAAGKYLTALFAF SVYGPYWREIRKISTIHLLSLRRLELLKHGRYLEIDCKMKRLFEYWMEHHKNIISTSSVKVNMSQVFAE LSLNVLKIIVGKTLFIKNGNEDYTKEEEEGQKLHKTILKFMELAGVSVASDVLPFLGWLDVDGQKKQMK RVYKEMNLIASKWLGEHRERKRLQIIKRGAAAGSNYDDGNDFMDVLSILDEENDDLFFGYSRDTVKS TCLQLIVAASDTTSLAMTWALSLLLTPNVLQKAQDELDTKVGRDRIIEEHIECLVYLQAIVKETLRLY PPAPLSLPHEAMEDCTVGGYQVKAGTRLVNLWKLQRDPRVWSNPLEFKPERFLPQSDGGFGGEEARMDF RGQHFEYTPFGSGRRICPGIDFFLQTVHMLARLLQAFDFNTAGGLVIDMVEGPGLTMPKVTPLVHLNP RLPVTLY
PbrCYP82-6	MQVDWPNILQKYYPITCSLLTLLSFYYIWVSITKPSRNSKTKLPPPEVAGSWPIVGHLPQLVGSTPLFK ILANMSDKYGPIMVRFGMHPTLVVSSWEMSKECFTTNDKFLASRPPSASAKYLYGDNAMFVPSDYGPYW REIRKISTLQELLTHKRDLKKNIPYLEINSCVKTYTRWAKTQSQIKQNVGGAADDFVKVDMTEMEGHLN LNVVRLRVGKPIFIQKDNADEDYTKDGHNKEELGQKLHKTIEFFELAGASVASDVLPYLGWLDVDGQK KRMKKIAMEMDLFAQKWLEEHQKGINHDNENDFMAVLISVLGEGKDDHIFGYSRDTVIKATCLTLIVAA TDTTLVSLTWALSLLLTPNVLSKAQDELDTVVGKERNVEDRDVNLVYLQAIKETLRLYPPSPPLAVPH EAIENCNVGGYEVKARTRLLVNLWKIHRDPRVWSNPLEFKPERFLPKLDGGTGEASKLDFKGQDFVYTPF GSGRRMCPGINFASQTLHMTLARLLHAFDFDIESNGLVIDMTEGSGLTMPKVTPLQVHLRPRIPATLY
PbrCYP82-7	MMDLAMFIDQYFSLAKIAGLLALLSFFYYLWISTLWSPRNPKLSSVSPPEVAGAWPIVGHLPQLLGSRL FKILADMSDNYGPIMVRFGMHPTLVVSSWEMAKECFTTNDRLFAGRPSGAANKYLTALFGFSTYGPYW REIRKIATLHLLSHRRLELLKHVPDLEVTNCMKHLHRRWIDSQNIQNDAAAGSVKVDMMGRVFGELTLN VVLKLVAGKSIFFKNDNTRQYDSKDGHNKEEEGKKLHKTIDFYSLAGASVASDVLPFLGWLDVDGQKK RMKRVAKMDMFIAAKWLEEHQKQKQTVLSSATLGSSNHDDAKDFMDVLSILDGENDDLFFGYSRDTV IKTTCLQLIAAAADTTSVTMTWALALLITNPILRKAQDELDTKVGKDRNIEERDINDLVYLQAIKETL RMYPAGPLNVPHEAIADCNIGGYEVRASTRLLVNLWKMHRDPRVWSNPSEFKPERFLPQLDGGSGGEAAN LDFRGQDFEYLPFSAGRRMCPGIDFSLQTLHMTLARLLHGFDFNDSAGIIMDEEGSGLTMPKLTPLLEI YLCPRLPALY

Table 6. *N*-methyltransferase and *N*-modifying candidate enzymes

Name	Sequence
TfCNMT	MAVEGKQVAPKKAHVELLKKLEGLVPDDEIKLIRIQLGRRLQWGCKSTYEEQIAQLVNLTHSLRQMKIATEVETLDDQMYEVPIDFLKIMNGS NLKGSCCYFKNDSTTLDEAEIAMLDLYCERAQIKDQGSVLDLGGCGGALTLYVAQKYKNSRVTAVTNSVSQKEFIEEESRKRNLNSNEVLLADITT HKMPDITYDRILVVELFEHMKNYELLRRKKEWMAKDGLLFEHCHKTFAYHYEPLDEDDWFTYVFPAGTMIIPSAFFLYFQDDVSVVNHWT LSGKHFSRTNEEWLKRDLANVELIKPMFVTITGQCRQEAAMKILNYWRGFCLSGMEFMFGYNNGEEMASHVLFKKK
CjCNMT	MAVEAKQTAKAAIVELLKQLEGLVPYDDIKQLIRRELARRLQWGYKPTYEEQIAEIQLNTHSLRQMKIATEVETLDSQLYEPIEFKIMNGSNLK GSCCYFKEDSTTLDEAEIAMLDLYCERAQIKDQGSVLDLGGCGGALTLYVAQKYKNCRTAVTNSVSQKEFIEEESRRRNLNVEVKLADITTHE MAETYDRILVIELFEHMKNYELLRKISEWISKDGLLFEHCHKTFAYHYEPLDDDDWFTYVFPAGTMIIPSAFFLYFQDDVSVVNHWTLSGK HFSRTNEEWLKRDLANLDVIKPMFETLMGNEEAVKLINYNWRGFCLSGMEFMFGYNNGEEMASHVLFKKK
PsCNMT	MQLKAKEELLRNMEGLIPDQEIQLIRVELEKRLQWGYKETHEEQQLSLDLVHSLKGMKMATMENLDKLYEAPMEFLKIQHGSNMKQSA GYTDESTTLDEAEIAMLDLYMERAQIKDQGSVLDLGGCGGALTLYVAQKYKNCRTAVTNSVSQKEFIEEESRRRNLNVEVKLADITTHE FDRIFAVELIEHMKNYQLLKKISEWMDGGLFVEHVCHKTAYHYEPLDAEDWYTNIFPAGTLTLSSASMLLYFQDDVSVVNHWTLSGKH SRSHHEEWLKNMDKNIVEFKEIMRSITKTEKEAKLLNFWRIFCMCGAELFGYKNGEEMWMLTHLLFKKK
PsTNMT	MGSIDEVKKESAGETLGRLLKGEIKDEELKKIKFQFEKRLQWGYKSSHQEQLSFNLDFIKSLKMMEMSGEIMTNKETELPSEFLEAVFGKTVK QSMCYFTHESATIDEAEAAHELYCERAQIKDQGTVDLIGCGQGGLVLYIAQKYKNCHVTGLTNSKAQVNYLLKQAEKGLTNVDAILADVTQY ESDKTYDRLLMIEAIEHMKNLQFLMKKLSWTMTESLLFVDHVCHKTFAHFFAEDDDWYSGFIFPPGCATILAAANSLLYFQDDVSVVDHVV VNGMHMARSVDIWRKALDKNMEAAKEILLPLGGSHETVNGVVTIHTFCMGGYEQFSMNNGDEWMVAQLLFKKK
EcTNMT	MGSAGEIMGRMLMKEIEDEELKKIRHQWDRRIEWGYKPTHEKQLAFNLDFIKLKEVMVMSGEIDTMNKETELPTAFLEAVFGKTVKQSCC YFKDENSTIDEAEAAHELYCERAQIKDQGTVDLIGCGQGGLVLYIAQKYKNCHVTGLTNSKAQVNYLLKQAEKGLTNVDAILADVTQY YDRILVVEITIEHMKNYQLFMMKLSWTMTESLLFVDHVCHKTFAHFFAEDDDWYSGFIFPPGCATILAAANSLLYFQDDVSVVDHVV MARSVEAWRKLDETIEAAREILEPGLSGKEAVNQTHTIHTFCMGGYEQFSYNNGEEMWMLTQLFKKK
PsRNMT	MSTTMETTKISQDDDLWKNMELGQISDEEVRRLMKIGIEKRIKWTGKPTQEQQLAQLLDFNKSRLGMMKMATIDLTENHKIETPESFNQIG GKESAGLFTDETTTMEANTKMDLYCERAGLKDGHITLDLGGCAGLLVHLAKYKSKITGITNTSSHKEYILKQCKNLNLSNVEIILADVTK VDIESTDRVFIGLIEHMKNFELFKISKWMDGGLLLEHLCHKSFSDDHWEPLSEDDWYAKNFFPSGTLVIPSATCLLYFQEDVTVIDHWILS GNNFARSNEVLKRIDGKIEVKDIFMSFYGIGREEAVKLINWWRLCITANELFKYNNGEEWLISQLLFKKLMTCI
TfPNMT	METRQTKKEAVANLIKRIEHEVSDIEIRGMMKIQVQKRLKWGYKPTHEEQQLAQLVTFQAQSLKGMEMAEFVDTLDAELYEIPFLHIMCGKT LKFSPGYFKDESTTLDESEVYMMMDLYCERAQIKDQGSVLDLGGCGGALTLYVAQKYKNCRTAVTNSVSQKEFIMDQCKKLDLSNVEILEDVTKF ETEITYDRIFAVALIEHMKNYELFLKKVSTWIAQYGLLFEHCHKTFAYHYEPLDEDDWYTEYIFPSGTLVMSSSSILLYFQEDVSVVNHWTLSG KHPSLGFKQWLKRLDDNIDEVKEIFESFYGSKEKAMKFTYWRVFCIAHSQMYSTNNNGEEMWMLSQVLFKKK
PbrTNMT1	MGSIDEVKKESAGETLGRLLKGEIKDEELKKIKFQFEKRLQWGYKSSHQEQLSFNLDFIKSLKMMEMSGEIMTNKETELPSEFLEAVFGKTVK QSMCYFKHESATIDEAEAAHELYCERAQIKDQGTVDLIGCGQGGLVLYIARKYKCHVTGLTNSKAQVNYLLKQAEKGLTNVDAILADVTQY SDKTYDRLLMIEAIEHMKNLQFLMKKLSWTMTESLLFVDHVCHKTFAHFFAEDDDWYSGFIFPPGCATILAAANSLLYFQDDVSVVDHVV NGMHMARSVDIWRKALDKNMEAAKEILLPLGGSHAVNGVVTIHTFCMGGYEQFSMNNGDEWMVAQLLFKKK
PbrTNMT2	MGSIEEVKKEAETLGRLLRGEINDEELKKIKYQLEKRLQWGYKSSHQEQLSFNLDFINSKKMMGMSGQVEAFTNEVYELPTECFEAYGKSM KLSGICYFKHESSTIDEAEAAHELYCERAQIKDQGTVDLIGCGQGGLVLYIAQKYKNCRTAVTNSVSQKEFIMDQCKKLDLSNVEILEDVTKF EMEDKFDRIFAVALIEHMKNYELLLRRISKWMDGGLFIEHVCHKTAYHYEPLDEDDWYTEYIFPAGTLTLSSASMLLYFQDDVSVVNHWTLSG KHYSRSHHEEWLKRIDGNMDAVKEIMKSITKTEEEAVKLINFWRIFCMCGAELFGYKNGEEMWMLSHVLFKKKQLLQCC
AmeNMT1	MCLFFAEKMGMAEANNQQQLKEDLLKNMELGLPDEEIKRLIRVQLEKRLNWGYKSTHEQQLSLLHLVHSLKMKIATEMENLDKLYEA PFSVQIQHGSTIKESSGLFKDESTTLDEAEIAMLDLYTKRAKIEDGQSVLDLGGCGGALTLYVAQKFNKYVTGITSSVEQKDFIEGRCKELKLSNV KVILADITTYETEEKYNRIFAVELIEHMKNYELLRKISEWMDGGLFIEHVCHKTAYHYEPLDEEDWYTNIFPAGTLTLSSATLLLYFQDDVAV VDQWTLSGKHYSRSHHEEWLKRIDGNIEVEKIMKSITKSEEAALLNFWRIFCMCGAELFGYKNGEEMWMLTHILFKKK
GfinNMT1	MDLMATSKQVKKKEELLKNMELGLVPDEEIRRLIRIEKRLKWGYKPTHEEQQLAQLDLVHSLKMKIATEMESLDKLYEAPFVQIKHGSTI KESSSYFKDESMTLDEAEIAMLDLYVERAQIEDGQSVLDLGGCGGALTLYVAQKYKNCRTAVTNSVSQKEFIEGCKELNLSNVKVLADVTSH EMEDKFDRIFAVALIEHMKNYELLLRRISKWMDGGLFIEHVCHKTAYHYEPLDEDDWYTEYIFPAGTLTLSSASMLLYFQDDVSVVNHWTLSG KHYSRSHHEEWLKRIDGNMDAVKEIMKSITKTEEEAVKLINFWRIFCMCGAELFGYKNGEEMWMLSHVLFKKKQLLQCC
EcaNMT1	MVDLKEKEELLKSMELGLVPDEDIRKHRSQLEKRLKWGYKPNHEEQQLAQLLDVIHSLKMKISKEYESFDRLRYEAPDFHKLGLTHLKESCS YYKDESTTLDEAEGAMLDLYTQAKIEDGQSVLDLGGCGGALTLYVAQKYKNCRTAVTNSVSQKEFIEEESRRRNLNVEVKLADITTHE FDRIFAVELIEHMKNYELLRKISKWMDGGLFIEHVCHKTAYHYEPLDEEDWYTEYIFPAGTLTLSSASMLLYFQDDVSVVHSSNLGKHYSRSH GEWLKNIDANIDEVKGIMRSITKTEEEAVRLNFWRIFCMCGIELFGYNNGEEMWMLSHILLKKK
EcaNMT2	MAADLVVKKWNNKKEIDEMELGLVGDEEIREURNDLEKRLKWGYKSNHEEQQLAQLLHFVHSLRGMKIAADEVESFNKYVEAPFSFNKIQLG SSLKESSCYKHDESTTLDEGEIAMMELYTEKAQIKDQGSVLDLGGCGGALTLYVAQKYKNCRTAVTNSVSQKEFIEEESRRRNLNVEVKLADITTHE THEMERFDRIFAVALIEHMKNYELLLRRISKWMDGGLFIEHVCHKTAYHYEPLDEDDWYTEYIFPAGTLTLSSASMLLYFQDDVSVVNHWTLSG GKHFSRTHHEEWLKNIDAKIDEVKEILKSVTKTEEEVRLTNFWRIFCMFGYVEMFGYNEGEEMWMLSQVLFKKK
CmaNMT4	MASGKVVLLKRLDSGLVSDLELRRVIRFELERRLKWGYKPTHEEQQLAELLNLAHATQMEIATKIDLTNSTMYEVPNSFLEIQLGSTLKESCLYF KDESTTVDEAEIAMMDLYLERAQIKDQGLILDGCGGALTLYVAQKYKNCRTAVTNSVSQKEFIEEESRRRNLNVEVKLADITTHE FAIGLIEHMKNYELLRKFSAWMKDQGLLFIHLCHKTGYHNEPLDEDDWYTYFFPAGTLTFIPSSFLLYFQDDVSVVNHWTLSGKHFSRSNE EWLKRMDNKIDEVKEIYKAAASETKDDDIIMKLIRLWRLSISAAEMFGYKNGEEMWMLSQVLFKKK
EcNMT3	MASLVEEGSFVNKESVKEVSELVKRLKNGLVSDLELRLKMRVELEKRLWGYKSTHEEQQLSLLDLAHSMMKMEIAMEIDALNSTVYEVPLS FLQIHGTTIKESCLYKDESTTVDEAEIAMMDLYLERAQIKDQGSVLDLGGCGGALTLYVAQKYKNCRTAVTNSVSQKEFIEEESRRRNLNVEVKLADITTHE ADICTTEIENVFDRIIAIGLIEHMKNYELLKKFSKWMQDGLLFIHLCHKTGYHNEPLDEDDWYTYFFPAGTLTFIPSSFLLYFQDDVSVVDH WTLNGKHFSRSHHEEWLKRMDKMDKIDVQKIFRSNLKSENEVTKTIGWRFLSMSAAEMFGYKNGEEMWMLSQVLFKKK

GfinNMT5	MGSNETNGELTKEMVDPDLLKRLESGVLADEELRKLIRFELERRLLKWGYKPTHEQQLAELLKLAHSTKQMKIATETDSLNSTMYEVPPIFLQLQF GSAIKESCCYFKDESTTLD AEAVAMMDLYLERTQIKDGGQSILDLGCGLGALAFHIVQKYPNCNVLAITNSVQKEFIEEKCKIRKVENVKVSLADIC TLEMKTTFDRIFAIGLLEHMKNYQLLLKKFSNWMKQDGLLFIHCHKTLAYHYEPLDEDDWYTEYFFPAGTLTISSSFLLYFQDDVSIVNHWLSL SGKHFSRSNEEWLKRMDMKIDEVKEILEAAFENKDHDITKLINHWRFILAINATEMFGYNNGEEWMMVSQVLFKKK
ScaNMT1	MASDHEVSNKELKKKEVITELLKRLESGVLSDEELRGLIRFELERRLLKWGYKPTHEQQLAQLNLAHSMKQMKIATEIDALNSTMYEVPPIFLQJ QLGSTLKESCCYFKDESTTVD AEAIAMMDLYLERAQIKDGGQSILDLGCGLGALAFHIAQKYTNCNITAITNSVRQKEFIEEKCKILNVSNVKVSLADI CTLEMEATFDRIFAIGLLEHMKNYELLLKKFSEWMKQDGLLFIHCHKTLAYHYEPLDEDDWYTEYFFPAGTLTISSSFLLYFQDDVSIVVDHWL SGKHFSRSNEEWLKRMDKIDEVKEIFESVSDSKDDVTRKLNHWRFCCISSAEMFGYNNGEEWMMISQVLFKKK
CchNMT3	MIKKSIMAFSDHHHEVVKNHSSKEMIADLLKRLEAGLVPDEEMRNLFERFELERRLLQWGYKSIHQEQLSQLLKAHSTKEMTIVAEMDALNSS MYELPISFLQJQLGSNLKQSSLYFKDELTTVD AEAVAIMDLYLERAQIEDGQSILDLGCGLGAFSFHVARKYTNCNITAVTNSLTQKEFIEKKSKILNI QNVKVFADVTTEMETTFDRVFAIGLIEHMKNYELFLKLSKWMKQDGLLFIHCHKTLAYHYKPIDEDDWFTNLLYPNGTVISSSLLLYFQD DVSVDHWLSLGGKHFSRASEESLKRMDAKMDKEMKEFESITDSKEEAMKLNQWRIFCISCAEMFGYNNGEEWMTSHFLFKKKL
CchNMT6	MGSSTASDHEMVMIMENDSKNKQVJADLLKRLVGGVLPDEEMRNLMFRFELEKRLKWGYKSTHQQLSQLNLVELNKGIKIAPIEMDALNSA MYEVPPIYKLMLGSTLKQSCLYFKDESTTLD AEAIEMMDLYLERADIQDGGQSILDLGCGLGGLGFHIAQKYSCNITALTNSLTQKEFIEEKCKTLN IPNVKVLADVTVEIETTFDRIFAIGLVEHMMENYELFLRKLKSKWMKQDGLLFIHCHKTLAYHYKPIDEDDWYSNLLYPTGTLSASFLLYFQDD LSVVDHWLSLGGKHFSRATEEWLKMIDANMDKIREIYESVTSKEEATRSINQWRIFCISCAEMFGYNDGEEWMMISHFLFKKKKQIE
CchNMT1	MATSDQEVKTSKMEMIADLLKRLEAGLVPDDEIRSLRVELERRLLKWGYKSTHQEQQLDQLNLAHSIKMKIASTEMDGLTSTMYEVPISLVQJQ LGSHLKESCLYFKDELTTVD AEAIAMMDLYLERAQIKDGGQSILDLGCGLGAVSFHIAQKYTCNITAVTNSVRQKEFIEEKSKTLNPNVKVLLADI TTLEMEHTFDRIFAISLIEHMENYELLRLKLESEWMKQDGLLFIHCHKTLSYHFEPMDDEDDWYTNNLLFPAGTLTLVSASFLLYFQDDLSVNVQW VMSGKHFSRANEWLKNMDAKMDKEMREIFESITDSEEVVKLINHWRFICISSAEMFAYNDGEEWMMNSHVLFKKKKQIQ
CchNMT2	MAGSGANKEMIADLLKRLEAGLVPDDEIRSLRVELERRLLKWGYKSTHQEQLEQLLSLAHSIRKMKIATEMDALNSTMYEVPISFMQIVFGSTLK ESCLYFKDEATTVNEAEIAMMDLYLERAQIKDGGQSILDLGCGMGSLCFHIAKYTCNITAVTNSVSQKEFIEEKSKTLNPNVKVLADITTEMD DTYDCLFAIGLIEHMKNYELLRLKLSNWMKQDGLLFIHCHKTLAYHYEPIDEDDWYTNNLLFPAGTLTLVSASFLLYFQDDLSLVDHWSMSGKH FSRTNKEWLKNIDGKMDKIREIVKSITDSEEEVVKLINHWRLCINSSEMFGFNDGEEWMMNSHVLFKKKKQJ
ScaNMT2	MEMIADLLKRLEAGLVPDDEIRSLRVELERRLLKWGYKSTHQEQQLDQLNLAHSIKMKIASTEMDGLTSTMYEVPISLVQJQLGSHLKESCLYFK DETTTVD AEAIAMMDLYLERAQIKDGGQSILDLGCGLGSVCFHIAKYTCNITAVTNSVSQKEFIEEKSKTLNPNVKVLLADITTEMDDTFDCFL AIGLIEHMENYELLRLKLSWWMKQDGLLFIHCHKTLSYHFEPMDDEDDWYTNNLLFPAGTLTLVSASFLLYFQDDLSLVDHWSMSGKHFSRTN KEWLKNIDGKMDKIREIVKSITDSEEEVVKLINHWRLCINSSEMFGFNDGEEWMMNSHVLFKKKKQJ
PbrNMT2	MCTTMDTTKISQQDLWKNMELGLSDEEVRRLMKIEKRIKWGTQKPTQEQQLAQLDFNKSRLRGMKIMATEVHALENHKKIYEPDSFNQJIG GKESAGLTFDEATTTIEEANTKMMMDLYCERAGLKDGGQTLIDIGCGAGLLVLHIAKKYKNCKITGVTNTSWHKEHILEQCKNLNLSNVEILADVT TVDIERTFDRVFIHCHKTLAYHYEPIDEDDWYTNNLLFPAGTLTLVSASFLLYFQDDLSLVDHWSMSGKHFSRTNKEWLKNIDGKMDKIREIVKSITDSEEEVVKLINHWRLCINSSEMFGFNDGEEWMMNSHVLFKKKKQJ
PbrNMT1	MVKGDQFQTTTMEETKISQENDLWTNMEGLIPDEEVRRLMKIEKRIEWGMKPTQHQQLAQLLDFTKSLRGMKIMATELDKLDLSKYETPH SFNQIVNGSTLKESGLYTDVTTTMEASIKMMMDLYCERANIKDGGQTLIDIGCGPGPLVLHIAKKYSNCKITGVTNFAFSQREYILEECKKLSNVE IILADVTSLDLETTFDRVFIHCHKTLAYHYEPIDEDDWYTNNLLFPAGTLTLVSASFLLYFQDDLSLVDHWSMSGKHFSRTNKEWLKNIDGKMDKIREIVKSITDSEEEVVKLINHWRLCINSSEMFGFNDGEEWMMNSHVLFKKKKQJ
EcaNMT4	MALEQEDSMSPERNEGVAULIKRMELGLVNDDEIRRLMRIQIENRLKWGYKPTHDQQLAQLLHFINSLKEMKIMATEMDSLDSQVYESPNF QQIMCGRSMKESAGLFMDVTTVEEAHIRMMDLYCCKATFEDGQKILDLGCGHGSVVLHVAQKYKGCQVTGVTNNSAQKQYILEQCKKLDL SNVEIILADVTTEMEEEKFDRVFIHCHKTLAYHYEPIDEDDWYTNNLLFPAGTLTLVSASFLLYFQDDLSLVDHWSMSGKHFSRTNKEWLKNIDGKMDKIREIVKSITDSEEEVVKLINHWRLCINSSEMFGFNDGEEWMMNSHVLFKKKKQJ
CmaNMT5	METGKNQNMKTIDDLWNQMMGLVDPKIEIRRLMKIELKRLDWGYRPTHQQQLSLLDFAKGLCNYCWTALRCMKMSAEFTDLSKVY ETPKSFQIMCGTTIKESGLFMNESTLDDAQISMLDLYFDKAKIDGGQSILDLGCGHGALILYLAQKYQNCNITGVTNLSQKEFIEVECKKLLGL SNVEIILADVTTEMEEEKFDRVFIHCHKTLAYHYEPIDEDDWYTNNLLFPAGTLTLVSASFLLYFQDDLSLVDHWSMSGKHFSRTNKEWLKNIDGKMDKIREIVKSITDSEEEVVKLINHWRLCINSSEMFGFNDGEEWMMNSHVLFKKKKQJ
GfinNMT4	MTMEANNAKKEAIENLWEQMMMGVLPDHEITRLMKSELQKRLNWGYKPTHQQLSLLDFAKSLRRMEMS LDFDNLELDTKMYETPESFQLIMSGTTLKESSGLFTDETATLDQTQIRMMMDLYLEKAKIDGGQSILDLGC GHGALILHVAQKYRNCNVTGVTNSIAQKEFIFKQCKKGLSNVEMVLADVTCKCEMKAFTDHFVIGLIEH MKNFELFLRKVSEWMKSDGLLFMEHYCHKSFAQWEPMDDDLFSSKYVFPFGSAIHPASFLLYFQDDLT VVDHWLTLGNGHFARTHQEWLKRIDSQSDIEKGFESFYGISKEEAVKLNHWRFVCLFGVEMFGYNNGEE WMMISHLFFKKK
CchNMT5	MEVVATSSARNPKKEIVDLWKRMEGLIPDEEIRDLMKIGLEKRLKWGYKPTHEQQLSQLLHFAKSLRSM KMASEMETLDDQMYETPTAFQQLMCGSTIKESAGFFKDESTTLD AEAIKMDLYCEKARIEDGQKILDLG CGHGAVMLHIAQKYKNCNVTGVTNSISQQQFIVQRSKELNLSNVNMLADVTMLEMDATYDRIFIGLIE HMKNFELFLRKISKWITKEGLLFLHYCHKTFAQYCEPVEDDWDYNNMFIFPGTLLPILPASFLLYFQDDL IVVDRWTLNGNHARTQEEWLKRIDANVDGVKQMFESVCDGNKEEAVKLMNFWRFICISGAEMLAYNNGE EWMISHYLKRRN
NsNMT2	MEATQITKKQGVAEILKRIENGQVPDEEITRMMKIQIKRLKLGKYSTHEQQLAQLLHFVHSLQKMEMAE EVDTLDSLEYEIPFLHIMCGKALKFSFGYFKDESTTLD ESEVNMDLYCERAGIEDGQTLIDIGCGHG SLTLHVAKKYRGCKVTGITSVSQKDFIMEECKKLNLSNVEIILEDVTKFETGTTDRIFAVALIEHMKN YELFLKKYSAWMAQDGLLVEHCHKVFAYKYEPIDDDWYTEYIFPTGLVMSSSLLYFQEDVSVN HWLTLGKHPSLGFKQWLKRIDNIDEIKFESFYGSKEKATKFITYWRVFCIAHSEMYATNGGEEWMLS QVLFKRR

ScaNMT5	MGGVADLLKKMELGLVPEEEIRRLMRIIEKRLEWGYKPTHAEQLDHLTNFIQCLRGMKMADEIDALDAK MYEIPLPFMQITICGSLKFSPGYFKDESTTLDSEIHMMDLYCERAEVKDGHSLDLGCGHGGFVLHVAQ KYKNSIVTGVNSVAEKEFIMTQCKKLCLSNVEIADVTKEPETTYDRVFAIALIEHMKNYELVLEKL SKWVAQDGLFVEHHCHKVFPYKYEPLEDDWYTEYIFPGGTIVLPSASILLYFQKDVSVVNHWSLNGKH PARGFKEWLKRLDENMDAVKAIFEPFYGSKEEAMKWITYWRVFCITHSEMYAYNNGEEWMLSQVLFKRK
JdiNMT1	MSKGAKLVERMELGLVSDDEVRLMRILIEKRLKWGYKPTHEEQTLTYLNFIOGLKGMKIAEEIDALDA KMYEIPAFMQILCGYSLKFSPGFFDESTTLDSEITIMMDLYCERAAVQDQGSILDLGCGHGGFVLHVA QKYKNCVTVGTNSVSETEYIMEQCKKLGLSNVEIADVTKEPEVTYDRVFAIALIEHMKNYELVLQK LSKWVAQDGLFVDHCHCHKVFPYKYEPIDEDDWTQYIFPGGTIVLPSASILLYFQEDVSVNHWTLSGN HPARGFKEWLKRLDDNMDEKAIFEPFYGSKEEAMKWITYWRVFCITHSEMYAYNGGEEWMISQVLFKRK
BthNMT1	MEVKQAGKEGVTLLVKRMELGLVPEEEIRRLMRIQKRLDWGYKPTHEEQLAHLTKFIQNRGMKMADE EIDALDAKMYEIPLPFQITICGSLKFSPGYFKDESTTLDSEITIMMDLYCERAAVQDQGSILDLGCGHG GFLVHLAQKYRNSVTVGTNSVSETEYIEQCKKLGLSNVEIADVTKEPEVTYDRVFAIALIEHMKN YALVLNKISKWVAQDGLFVEHHCHKVFPYKYEPLEDDWYTNIFPGGTIVLPSASILLYFQEDVTVLN HWSLSGKHPSRGFIEWLKRLDENIDVIMGIFEPFYGSKEEATKWINYWRVFCITHSEMYAYNGGEEWMLS QVLLKRK
MaqNMT3	MELGLVPEKEIRRLMRIQKRLDWGYKPTHEEQLAHLTKFIQNRGMKMADEIDALDAKMYEIPLPFQ TICGSLKFSPGYFKDESTTLDSEITIMMDLYCERAAVQDQGSILDLGCGHGGFVLHVAQKYRNSIVTV TNSVSETEYIEQCKKLGLSNVEIADVTKEPEVTYDRVFAIALIEHMKNYALVLNKISKWVAQDGL FVEHHCHKVFPYKYEPLEDDWYTNIFPGGTIVLPSASILLYFQEDVTVLNHWSLSGKHPSRGFIEWLK RLDENIDVIMGIFEPFYGSKEEATKWINYWRVFCITHSEMYAYNGGEEWMLSQVLLKRK
McaNMT4	MDKANERELKRAELFKLEDDLVTYDEIKQVMRELAKRLEWGYKPTHQQLAHLDFAHLEGMKIANE VETLASEVYETPLPFIEVLGPAKXSSCLFEDESTTLEQAEIAMLDLYFERAQIRXGMSVLDLGCXGS VGLHIARKYKNCXVTCITNSISQKQYIENQCKLYLNSNVKILADIVADTDDTDFVVLVIGVIEHMKNY ALLNKISKWMAKDGLFVEHLCHKTFPYHFEPLDEDDWYSNFWPTGTLTMPVSFLLYFQADVSLNH WILSGKNFSRTXEEFLKRIDANVDAIKDGLKPSLSEGVAKLISYWRGFLTGMEMFGYNGGEEWMVSVQ LFRNK
TcoNMT3	MEDNNNLLQEEEMNVVELLQRPGLVPEDEKIRKLRLQLKRLKWGYKPTHEAQLSHLFQFIHSLPSLNM ESEDENPKSWLYETPTSFLQLLYGDCIKESDTYYKEDTATLEEAVINMLEYICERARITEGLSVLDLGC YGALTILHVAQKYKSCVTVGTSSISQKQYIMECKKLNLNVEIADVATIEIEAASYDRIFALGIFEH VNDYKFLGLKLSKWMKQDGLFVEYLCHKTFPYQNKPLDDGDKWYNEYVFPSSGLIIPASILLYFQNDV SVVRQWTQGGQHSARTFEELLKRIDGNIDKIEFIESYGSKEADAVRFINYWRVFLITGVEMFSYNDGEE WMGAHFLFKKKFIMQE
CmuNMT4	MEVKQSGDELRSRAELLERPELGLVPEDEIRRLAKARLEKRLKWGYKPTHEQQLSLLQFVESLPSLN MASEDDSPKAWLYETPTSFLQLLYGDIKESGSYYKDESTTLEEAMIHNMNLCERANIEGQSVVDLGC GYGAFILHVAQKYKTCRVGTITSSISQKHYIMEQCKKLNLNVEIADVATIKLDTAFDRVFAAGMFEH VNDYKFLRLKITNWMKPDGLFVEHLCKNTFPYQNKPLDDGDNWGEYVFPSSGLIIPASILLYFQEDVS IVNHWTFSGKHAANKFEELLKRIDAKIDAKIRFNECYGSKDSIRFINYWRVFLITAAEMFGYNGGEEWM GVHLLFKKK
CtrNMT2	GLKSSVAELLERPELGLVPEDEIRKLTKRLAKRLEWGYKPTHEQQLSHLLRFIHLPSLNMASEDDSPK AWLYETPTSFLQLLYGDIKESGTYYKDESTTLEEAIHNMNLCERARIEGQSVVDLGCYGAFTLHV AQKYKSCSVTGTSSISQKDYIMEQCKKLNLNVEIADVATIKMNTTFDRVFAAGMFEHINDYKFLR RISNWMKHDGLFVEHLCKNTFAYQNKPLDDGDDWFNEYVFPSSAGLIIPASILLYFQEDVSVHWTFS GKHAAYKFEELLERIDAKIEAIEFIECYGSKDAIRFINYWRVFLITAAEMFAYRDGEEWMGSHVLFK KK
CmuNMT5	MEAKQHESNNNIDEELKNRVNIGEQEERPGFEDEIRRLAKAQLAKRLKWGYKPTHEQQLSHLLQFLQSL PSLNMASEDESSKAWLYETPTSFLQLLYGDIKESGSYYKHESSTFEESMIHNMNLCERANIEGQSVN DLGCGYGAFTLHVAQKYKSCSVTGTITSSITQKHHIMECKKLNLNVEIADVATIELGTAFDRVFAFG MFEEINDYKILRKISNWMKPDGLFVEHLCKNTLAYQNKPLDDQDWYEEYFPGGTIVLPSASILLYFQ DDLVSYYHWYNGKHGARSFEKMLERTDANIDTIKDMFTEFYGSKEAKAIFINYWRVFLITAAEMFAYND GEEWMCSQLLFKKK
CmuNMT8	MEHKIEDIRKLSRVEEQLEPGLVKDEDIKTLAKAKLEKRLKWGYKPTAEQLSNLLQFAQSLPSLK MENVDQDQSSQWLYGVPSEFLQIYGGIIMSGSYDEDESTTLEESMIKDMDSCEKANVKEGHSVLDI GCGYGSIIHIAKKYRTCNVTGITNFVEQKQYIMECKKLNLNVEIADVATIKMNTTFDRVFAAGMFEH LEEINDYKFLKSVSDWMKPDGLLLVTHFCHKTFAYQNNKALDDEWHNEYVFPSSGLIIPASILLYFQ EDLSVSHWATNGTHTGRTCKLVERIDANIEKIEFSEFYGSKEADIRMINYWRVFLITGAEMYTKD GEEWMDVYVLFKKK

Table 7. Variants of BM3 *N*-demethylase

BM3 variant	Genotype
8F11	L437A
4H9	L181A, T260A, L437A
8C7	L75A, L181A
4H5	L75A, M177A, L181A
7A1	L75A, M177A, L181A, T260A
BM3 variant	Amino Acid Sequence
8F11	MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQRLIKEACDESFRDKN LSQALKFARDFAGDGLVTSWTHEKNWKKAHNILLPSFSQQAMKGYHAMMVDIAVQLVQKWERLNADE HIEVSEDMTRLTLDTIGLCGFNYRFNSFYRDQPHPFHSMVRALDEV MNKLQRANPDDPAYDENKRQFQE DIKVMNDLVDKIIADRKARGEQSDDLTLQMLNGKDPETGEPLDDGNIRYQIITFLIAGHETTSGLLSFALYF LVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPTAPAFSLYAKEDTVLGGEYPLE KGDEVMLVLPQLHRDKTVWGDDVEEFRPERFENPSAIPQHAFKPFNGQQRACIGQQFALHEATLVLGMM LKHDFDFEDHTNYELDIKETATLKPFGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLGYG NMGTAEGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPPDNAKQFVDWLDQASAD EVKGVRYSVFGCGDKNWATTYQKVPFIDETLAAKGAENIADRGEADASDDFEGTYEWEHMMWSDVA AYFNLDIENSEDNKSTLSLQFVDSAADMPLAKMHGAFSTNVVASKELQQPGSARSTRHLEIELPKEASYQE GDHLGVIPRNYEGIVNRVTARFGLDASQQIRLEAEEEEKLAHLPLAKTVSVEELLQYVELQDPVTRTQLRAM AAKTVCPPHKVELEALLEKQAYKEQVLAKRLTMLELLEKYPACEMKPFSEFIALLPSIRPRYSSISSSPRVDE KQASITVSVVSGEAWSGYGEYKGIASNYLAELQEGDTITCFISTPQSEFTLPKDPETPLIMVPGTGVAPFRG FVQARKQLKEQGQSLGEAHLFYGCRSPHEDYLYQEELENAQSEGIITLHTAFSRMPNQPKTYVQHVMEQD GKKLIELLDQGAHFYICGDSQMAPAVEATLMKSYADVHQVSEADARLWLQQLLEEKGRYAKDVWAG
4H9	MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQRLIKEACDESFRDKN LSQALKFARDFAGDGLVTSWTHEKNWKKAHNILLPSFSQQAMKGYHAMMVDIAVQLVQKWERLNADE HIEVSEDMTRLTLDTIGLCGFNYRFNSFYRDQPHPFHSMVRAADEVMNKLQRANPDDPAYDENKRQFQE DIKVMNDLVDKIIADRKARGEQSDDLTLQMLNGKDPETGEPLDDGNIRYQIIAFLIAGHETTSGLLSFALYF LVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPTAPAFSLYAKEDTVLGGEYPLE KGDEVMLVLPQLHRDKTVWGDDVEEFRPERFENPSAIPQHAFKPFNGQQRACIGQQFALHEATLVLGMM LKHDFDFEDHTNYELDIKETATLKPFGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLGYG NMGTAEGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPPDNAKQFVDWLDQASAD EVKGVRYSVFGCGDKNWATTYQKVPFIDETLAAKGAENIADRGEADASDDFEGTYEWEHMMWSDVA AYFNLDIENSEDNKSTLSLQFVDSAADMPLAKMHGAFSTNVVASKELQQPGSARSTRHLEIELPKEASYQE GDHLGVIPRNYEGIVNRVTARFGLDASQQIRLEAEEEEKLAHLPLAKTVSVEELLQYVELQDPVTRTQLRAM AAKTVCPPHKVELEALLEKQAYKEQVLAKRLTMLELLEKYPACEMKPFSEFIALLPSIRPRYSSISSSPRVDE KQASITVSVVSGEAWSGYGEYKGIASNYLAELQEGDTITCFISTPQSEFTLPKDPETPLIMVPGTGVAPFRG FVQARKQLKEQGQSLGEAHLFYGCRSPHEDYLYQEELENAQSEGIITLHTAFSRMPNQPKTYVQHVMEQD GKKLIELLDQGAHFYICGDSQMAPAVEATLMKSYADVHQVSEADARLWLQQLLEEKGRYAKDVWAG
8C7	MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQRLIKEACDESFRDKN LSQAAKFARDFAGDGLVTSWTHEKNWKKAHNILLPSFSQQAMKGYHAMMVDIAVQLVQKWERLNADE HIEVSEDMTRLTLDTIGLCGFNYRFNSFYRDQPHPFHSMVRAADEVMNKLQRANPDDPAYDENKRQFQE DIKVMNDLVDKIIADRKARGEQSDDLTLQMLNGKDPETGEPLDDGNIRYQIITFLIAGHETTSGLLSFALYF LVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPTAPAFSLYAKEDTVLGGEYPLE KGDEVMLVLPQLHRDKTVWGDDVEEFRPERFENPSAIPQHAFKPFNGQQRACIGQQFALHEATLVLGMM LKHDFDFEDHTNYELDIKETLTLKPFGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLGYG NMGTAEGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNGHPPDNAKQFVDWLDQASAD EVKGVRYSVFGCGDKNWATTYQKVPFIDETLAAKGAENIADRGEADASDDFEGTYEWEHMMWSDVA AYFNLDIENSEDNKSTLSLQFVDSAADMPLAKMHGAFSTNVVASKELQQPGSARSTRHLEIELPKEASYQE GDHLGVIPRNYEGIVNRVTARFGLDASQQIRLEAEEEEKLAHLPLAKTVSVEELLQYVELQDPVTRTQLRAM AAKTVCPPHKVELEALLEKQAYKEQVLAKRLTMLELLEKYPACEMKPFSEFIALLPSIRPRYSSISSSPRVDE KQASITVSVVSGEAWSGYGEYKGIASNYLAELQEGDTITCFISTPQSEFTLPKDPETPLIMVPGTGVAPFRG FVQARKQLKEQGQSLGEAHLFYGCRSPHEDYLYQEELENAQSEGIITLHTAFSRMPNQPKTYVQHVMEQD GKKLIELLDQGAHFYICGDSQMAPAVEATLMKSYADVHQVSEADARLWLQQLLEEKGRYAKDVWAG
4H5	MTIKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLSSQRLIKEACDESFRDKN LSQAAKFARDFAGDGLVTSWTHEKNWKKAHNILLPSFSQQAMKGYHAMMVDIAVQLVQKWERLNADE HIEVSEDMTRLTLDTIGLCGFNYRFNSFYRDQPHPFHSAVRAADEVMNKLQRANPDDPAYDENKRQFQE

	DIKVMNDLVDKIIADRKARGEQSDLLTQMLNGKDPETGEPLDDGNIRYQIITFLIAGHETTSGLLSFALYF LVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPTAPAFSLYAKEDTVLGGEYPLE KGDEVMVLIPQLHRDKTVWGDDVEEFRPERFENPSAIPQHAFKPFNGQQRACIGQQFALHEATLVLGMM LKHDFDFEDHTNYELDIKETLTLPKPGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLGYS NMGTAEGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNHPPDNAKQFVDWLDQASAD EVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAAKGAENIADRGEADASDDFEGTYEEWREHMWSDVA AYFNLDIENSEDNKSTLSLQFVDSAADMPLAKMHGAFSTNNVASKELQQPGSARSTRHLEIELPKEASYQE GDHLGVIPRNYEGIVNRVTARFGLDASQQIRLEAEEEEKLAHLPLAKTVSVEELLQYVELQDPVTRTQLRAM AAKTVCPPHKVELEALLEKQAYKEQVLAKRLTMLELLEKYPACEMKPFSEFIALLPSIRPRYSSISSSPRVDE KQASITVSVVSGEAWSGYGEYKGIASNYLAELQEGDTITCFISTPQSEFTLPKDPETPLIMVPGTGVAPFRG FVQARKQLKEQQQSLGEAHLFYGCRSPHEDYLYQEELNAQSEGIITLHTAFSRMPNPQPKTYVQHVMEQD GKKLIELLDQGAHFYICGDSQMAPAVEATLMKSYADVHQVSEADARLWLQQLLEEKGRYAKDVWAG
7A1	MTIKEMPQKTFGELKNLPLENTDKPVQALMKIADDELGEIFKFEAPGRVTRYLSSQRLIKEACDESFRDKN LSQAAKFARDFAGDGLVTSWTHEKNWKAHNILLPSFSQQAMKGYHAMMVDAIVQLVQKWERLNADE HIEVSEDMTRLTLDITGLCGFNRYFNSFYRDQPHPFHSAVRAADEVMNKLQRANPDDPAYDENKRQFQE DIKVMNDLVDKIIADRKARGEQSDLLTQMLNGKDPETGEPLDDGNIRYQIITFLIAGHETTSGLLSFALYF LVKNPHVLQKVAEEAARVLVDPVPSYKQVKQLKYVGMVLNEALRLWPTAPAFSLYAKEDTVLGGEYPLE KGDEVMVLIPQLHRDKTVWGDDVEEFRPERFENPSAIPQHAFKPFNGQQRACIGQQFALHEATLVLGMM LKHDFDFEDHTNYELDIKETLTLPKPGFVVKAKSKKIPLGGIPSPSTEQSAKKVRKKAENAHNTPLLVLGYS NMGTAEGTARDLADIAMSKGFAPQVATLDSHAGNLPREGAVLIVTASYNHPPDNAKQFVDWLDQASAD EVKGVRYSVFGCGDKNWATTYQKVPAFIDETLAAKGAENIADRGEADASDDFEGTYEEWREHMWSDVA AYFNLDIENSEDNKSTLSLQFVDSAADMPLAKMHGAFSTNNVASKELQQPGSARSTRHLEIELPKEASYQE GDHLGVIPRNYEGIVNRVTARFGLDASQQIRLEAEEEEKLAHLPLAKTVSVEELLQYVELQDPVTRTQLRAM AAKTVCPPHKVELEALLEKQAYKEQVLAKRLTMLELLEKYPACEMKPFSEFIALLPSIRPRYSSISSSPRVDE KQASITVSVVSGEAWSGYGEYKGIASNYLAELQEGDTITCFISTPQSEFTLPKDPETPLIMVPGTGVAPFRG FVQARKQLKEQQQSLGEAHLFYGCRSPHEDYLYQEELNAQSEGIITLHTAFSRMPNPQPKTYVQHVMEQD GKKLIELLDQGAHFYICGDSQMAPAVEATLMKSYADVHQVSEADARLWLQQLLEEKGRYAKDVWAG
BM3 variant	Nucleotide Sequence
8F11	ATGACCATCAAAGAAATGCCACAACCTAAGACTTTCGGTGAATTGAAGAAATTGCGCTTTGTTGAACACCG ATAAGCCAGTTCAAGCTTTGATGAAGATTGCTGATGAATTGGGTGAAATCTTCAAGTTTGAAGCTCCAG GTAGAGTCACTAGATACTTGTCTCATCTCAAAGATTGATCAAAGAAGCCTGCGACGAATCCAGATTTGATA AGAATTTGTCTCAAGCTTTGAAGTTCGCTAGAGATTTTGTGCTGGTGATGGTTTGGTTACTTCTTGGACTCA CGAAAAGAATTGGAAGAAGGCCATAACATTTTGTGCCATCTTCTCACAACAAGCCATGAAGGGTTAT CATGCTATGATGGTTGATATCGCCGTTCAATTGGTTCAAAAGTGGGAAAGATTGAACGCCGATGAACAT ATCGAAGTCTCTGAAGATATGACCAGATTGACCTTGGATACCATTGGTTTGTGTGGTTTCAACTACAGAT TCAACTCCTTCTACAGAGATCAACCACATCCATTTCATCATCTCTATGGTTAGAGCTTTGGATGAAGTCAT GAACAAATTGCAAAGAGCTAATCCAGACGATCCAGCTTATGACGAAAACAAGAGACAATTCCAAGAAGA TATCAAGGTCATGAACGATTTGGTCGATAAGATTATCGCTGATAGAAAGGCTAGAGGTGAACAATCTGA TGATTTGTTGACCCAAATGTTGAACGGTAAGGATCCAGAACTGGTGAACCATTGGATGATGGTAACAT CAGATACCAAATTATCACCTTCTTGATTGCTGGTCACGAACTACATCTGGTTTGTGTGCTTTTGCCTTG TACTTTTGGTTAAGAACCACACGCTTTCGAAAAGGTTGCTGAAGAAGCTGCAAGAGTTTGGTTGATC CAGTTCCATCTTACAAGCAAGTCAAGCAATTGAAGTACGTTGGTATGGTTTGAACGAAGCTTTGAGATT GTGGCCAACTGCTCCAGCTTTTTCATTATACGCTAAAGAAGATACCGTCTTGGGTGGTGAATATCCATTG GAAAAAGGTGATGAAGTTATGGTCTTGATCCCACAATTGCATAGAGATAAGACTGTTTGGGTGATGAT GTCGAAGAATTGAGACCAGAAAGATTGCAAAAACCATCTGCTATTCCACAACATGCTTTTAAGCCATTTG GTAACGGTCAAAGAGCTTGCATTGGTCAACAATTGCGTTTACATGAAGCTACCTTGGTTTGGGTATGAT GTTGAAACACTTCGACTTCGAAGATCACACCAACTACGAATTGGATATCAAAGAAACCGCTACCTTGAAG CCAAAGGGTTTGTGTTAAGGCTAAGTCCAAAAAGATTCCATTGGGTGGTATTCCATCTCCATCTACTG AACAATCCGCTAAGAAGGTTAGAAAGAAAGCTGAAAACGCTCATAACACACCTTTGTTGGTCTTGTACGG TTCTAATATGGGTACTGCTGAAGGTACAGCAAGAGATTGCGCAGATATTGCTATGTCTAAAGGTTTCGCT CCACAAGTTGCTACTTTGGATTCTCATGCTGGTAATTTGCCAAGAGAAGGTGCTGTTTTGATAGTTACTG CTTCTTACAATGGTCACCCACCAGATAATGCTAAGCAATTCGTTGATTGGTTGGATCAAGCTTCAAGTGA TGAAGTAAAAGGTGTTAGATACTCTGTTTTCGGTTGCGGTGACAAAAATTTGGGCTACTACTTACAAA GGTTCCAGCCTTTATTGACGAACTTTGGCTGCTAAGGTGCTGAAAAACATTGCTGACAGAGGTGAAGCT GATGCCCTCCGACGACTTCGAAGGTACTTACGAAGAATGGAGAGAACACATGTGGTCTGACGTTGCTGCTT ACTTCAACTTGGACATCGAAAACCTGTAAGACAACAAGTCCACTTTGTCTTTGCAATTCGTTGACTCCGC TGCTGACATGCCATTGGCTAAGATGCACGGTGCTTTCTCTACCAACGTCGTTGCCCTCAAGGAATTGCAA CAACCAGGTTCTGCTAGATCTACTAGACACTTGAAATCGAATTGCCAAAGGAAGCTTCTACCAAGAAG GTGACCACTTGGGCGTTATTCCAAGAACTACGAAGGTATCGTCAACAGAGTTACTGCTAGATTCCGTTT

	GGATGCTTCTCAACAAATCAGATTAGAAGCTGAAGAAGAAAAGTTGGCTCACTTGCCATTAGCTAAGAC TGTCTCCGTTGAAGAATTGTTGCAATACGTGCAATTGCAAGACCCAGTTACCAGAACCCAAATTGAGAGCC ATGGCTGCCAAGACCGTCTGTCCACCACACAAGGTTGAATTGGAAGCCTTGTTGGAAAAGCAAGCCTACA AGGAACAAGTTTGGCTAAGAGATTGACCATGTTGGAATTGTTGGAAAAGTACCCAGCCTGCCAAATGA AGTTCTCTGAATTTATCGCCTTGTTGCCATCTATCAGACCACGTTACTACTCTATTTCTTCTCTCCACGT GTTGACGAAAAGCAAGCTTCTATTACTGTTTCCGTTGTCTCCGGTGAAGCTTGGTCCGGTTACGGTGAAT ACAAGGTTATTGCTTCTAACTACTTGGCTGAATTGCAAGAAGGTGACACCATTACTTGTTCATCTCTAC TCCACAATCCGAATTTACTTTGCCAAAGGACCCAGAACTCCATTGATCATGGTTGGTCCAGGTACTGGT GTCGCTCCATTGAGAGGTTTCGTTCAAGCTAGAAAACAATTGAAGGAACAAGGTCAATCTTTGGGTGAA GCTCACTTGTACTTCGGTTGTAGATCTCCACACGAAGACTACTTATACCAAGAAGAATTGGAAAACGCTC AATCCGAAGGTATTATCACTTTGCACACCGCTTTCTCCAGAATGCCAAACCAACCAAGACTTACGTCCA ACACGTTATGGAACAAGACGGTAAGAAGTTGATTGAATTGTTGGACCAAGGTGCTCACTTCTACATTTGT GGTGATGGTTCTCAAATGGCTCCAGCCGTTGAAGCCACTTTGATGAAGTCTTACGCTGATGTTACCAAG TTTCCGAAGCCGATGCTAGATTATGGTTGCAACAATTGGAAGAAAAGGTGCTTACGCTAAGGATGTCT GGGCCGGTTGA
4H9	ATGACCATCAAAGAAATGCCACAACCTAAGACTTTCCGGTGAATTGAAGAATTTGCCTTTGTTGAACCCG ATAAGCCAGTTCAAGCTTTGATGAAGATTGCTGATGAATTGGGTGAAATCTTCAAGTTTGAAGCTCCAG GTAGAGTCACTAGATACTTGTCTCTCAAAGATTGATCAAAGAAGCCTGCGACGAATCCAGATTTGATA AGAATTTGTCTCAAGCTTTGAAGTTGCTAGAGATTTTGTGGTGATGGTTTGGTTACTTCTTGGACTCA CGAAAAGAATTGGAAGAAGGCCATAACATTTTGTGGCATCTTTCTCACAACAAGCCATGAAGGGTTAT CATGCTATGATGGTTGATATCGCCGTTCAATTGGTTCAAAGTGGGAAAGATTGAACGCCGATGAACAT ATCGAAGTCTCTGAAGATATGACCAGATTGACCTTGGATACCATTGGTTTGTGTGGTTTCACTACAGAT TCAACTCCTTCTACAGAGATCAACCACATCCATTCACTCTCTATGGTTAGAGTGCAGATGAAGTCAT GAACAAATTGCAAAGAGCTAATCCAGACGATCCAGCTTATGACGAAAACAAGAGACAATTCCAAGAAGA TATCAAGGTCATGAACGATTTGGTCGATAAGATTATCGCTGATAGAAAGGCTAGAGGTGAACAATCTGA TGATTTGTTGACCCAAATGTTGAACGGTAAGGATCCAGAACTGGTGAACCATTTGGATGATGGTAACAT CAGATACCAAATTTATCGCTTTCTTGATTGCTGGTCACGAACTACATCTGGTTTGTGTCTTTTGCCTTG TACTTTTGGTTAAGAACCACACGCTTTGCAAAAGGTTGCTGAAGAAGCTGCAAGAGTTTGGTTGATC CAGTCCCATCTTACAAGCAAGTCAAGCAATTGAAGTACGTTGGTATGGTTTGAACGAAGTTTGGTTGATC GTGGCCAACTGCTCCAGCTTTTTCATTATACGCTAAGAAAGATACCGTCTTGGGTGGTGAATATCCATTG GAAAAAGGTGATGAAGTTATGGTCTTGATCCCAACAATTGCATAGAGATAAGACTGTTTGGGGTGATGAT GTCGAAGAATTGAGACCAGAAAGATTGCAAAAACCCATCTGCTATTCCACAACATGCTTTTAAGCCATTTG GTAACGGTCAAAGAGCTTGCATTGGTCAACAATTGCTTTACATGAAGCTACCTTGGTTTGGGTATGAT GTTGAAACACTTCGACTTCGAAGATCACACCAACTACGAATTGGATATCAAAGAAACCGCTACCTTGAAG CCAAAGGGTTTGTGTTAAGGCTAAGTCCAAAAAGATTCCATTGGGTGGTATTCCATCTCCATCTACTG AACAATCCGCTAAGAAGGTTAGAAAGAAAGCTGAAAACGCTCATAACACACCTTTGTTGGTCTTGTACGG TTCTAATATGGGTACTGCTGAAGGTACAGCAAGAGATTTGGCAGATATTGCTATGTCTAAAGGTTTCGCT CCACAAGTTGCTACTTTGGATTCTCATGCTGGTAATTTGCCAAGAGAAGGTGCTGTTTTGATAGTTACTG CTTCTTACAATGGTCACCCACCAGATAATGCTAAGCAATTCGTTGATTGGTTGGATCAAGCTTCAGCTGA TGAAGTAAAAGGTGTTAGATACTCTGTTTTCGGTTGCGGTGACAAAAATTTGGGCTACTACTTATCAAAA GGTTCCAGCCTTTATTGACGAACTTTGGCTGCTAAGGTGCTGAAAACATTGCTGACAGAGGTGAAGCT GATGCCCTCCGACGACTTCGAAGGTACTTACGAAGATGGAGAGAACACATGTGGTCTGACGTTGCTGCTT ACTTCAACTTTGGACATCGAAAACCTGGAAGACAACAAGTCCACTTTGTCTTTGCAATTCGTTGACTCCGC TGCTGACATGCCATTGGCTAAGATGCACGGTGCTTTCTCTACCAACGTCGTTGCCCTCAAGGAATTGCAA CAACCAGGTTCTGCTAGATCTACTAGACACTTGGAAATCGAATTGCCAAAGGAAGCTTCTACCAAGAAG GTGACCACTTGGGCGTTATTCCAAGAACTACGAAGGTATCGTCAACAGAGTTACTGCTAGATTCCGTTT GGATGCTTCTCAACAAATCAGATTAGAAGCTGAAGAAGAAAAGTTGGCTCACTTGCCATTAGCTAAGAC TGTCTCCGTTGAAGAATTGTTGCAATACGTGCAATTGCAAGACCCAGTTACCAGAACCCAAATTGAGAGCC ATGGCTGCCAAGACCGTCTGTCCACCACACAAGGTTGAATTGGAAGCCTTGTTGGAAAAGCAAGCCTACA AGGAACAAGTTTGGCTAAGAGATTGACCATGTTGGAATTGTTGGAAAAGTACCCAGCCTGCCAAATGA AGTTCTCTGAATTTATCGCCTTGTTGCCATCTATCAGACCACGTTACTACTCTATTTCTTCTCTCCACGT GTTGACGAAAAGCAAGCTTCTATTACTGTTTCCGTTGTCTCCGGTGAAGCTTGGTCCGGTTACGGTGAAT ACAAGGTTATTGCTTCTAACTACTTGGCTGAATTGCAAGAAGGTGACACCATTACTTGTTCATCTCTAC TCCACAATCCGAATTTACTTTGCCAAAGGACCCAGAACTCCATTGATCATGGTTGCTCCAGGTACTGGT GTCGCTCCATTGAGAGGTTTCGTTCAAGCTAGAAAACAATTGAAGGAACAAGGTCAATCTTTGGGTGAA GCTCACTTGTACTTCGGTTGTAGATCTCCACACGAAGACTACTTATACCAAGAAGAATTGGAAAACGCTC AATCCGAAGGTATTATCACTTTGCACACCGCTTTCTCCAGAATGCCAAACCAACCAAGACTTACGTCCA ACACGTTATGGAACAAGACGGTAAGAAGTTGATTGAATTGTTGGACCAAGGTGCTCACTTCTACATTTGT GGTGATGGTTCTCAAATGGCTCCAGCCGTTGAAGCCACTTTGATGAAGTCTTACGCTGATGTTACCAAG TTTCCGAAGCCGATGCTAGATTATGGTTGCAACAATTGGAAGAAAAGGTGCTTACGCTAAGGATGTCT

	GGGCCGGTTGA
8C7	ATGACCATCAAAGAAATGCCACAACCTAAGACTTTCGGTGAATTGAAGAATTTGCCTTTGTTGAACACCG ATAAGCCAGTTCAAGCTTTGATGAAGATTGCTGATGAATTGGGTGAAATCTTCAAGTTTGAAGCTCCAG GTAGAGTCACTAGATACTTGTCTCATCAAGATTGATCAAAGAAGCCTGCGACGAATCCAGATTTGATA AGAATTTGTCTCAAGCTGCTAAGTTTCGCTAGAGATTTTGGTGGTGATGGTTTGGTTACTTCTTGGACTCA CGAAAAGAATTGGAAGAAGGCCCATTAACATTTTGTGGCATCTTCTCACAACAAGCCATGAAGGGTTAT CATGCTATGATGGTTGATATCGCCGTTCAATTGGTTCAAAAGTGGGAAAGATTGAACGCCGATGAACAT ATCGAAGTCTCTGAAGATATGACCAGATTGACCTTGGATACCATTGGTTTGTGTGGTTTCAACTACAGAT TCAACTCCTTCTACAGAGATCAACCACATCCATTTCATCATCTCTATGGTTAGAGCTGCAGATGAAGTCAT GAACAAATGCAAAGAGCTAATCCAGACGATCCAGCTTATGACGAAAACAAGAGACAATTCGAAGAAGA TATCAAGGTCATGAACGATTTGGTCGATAAGATTATCGCTGATAGAAAGGCTAGAGGTGAACAATCTGA TGATTTGTGACCCAAATGTTGAACGGTAAGGATCCAGAAACTGGTGAACCATTTGGATGATGGTAACAT CAGATACCAAATATCACCTTCTTGATTGCTGGTCACGAACTACATCTGGTTTGTGTCTTTTGCCTTG TACTTTTGGTTAAGAACCCACAGTCTTGCAAAAGGTTGCTGAAGAAGCTGCAAGAGTTTGGTTGATC CAGTTCCATCTTACAAGCAAGTCAAGCAATTGAAGTACGTTGGTATGGTTTGAACGAAGCTTTGAGATT GTGGCCAACTGCTCCAGCTTTTTCATTATACGCTAAAGAAGATACCGTCTTGGGTGGTGAATATCCATTG GAAAAAGGTGATGAAGTTATGGTCTTGATCCCACAATTGTCATAGAGATAAGACTGTTTGGGGTGATGAT GTCGAAGAATTGAGACCAGAAAGATTGCAAAACCCATCTGCTATTCCACAACATGCTTTTAAGCCATTTG GTAACGGTCAAAGAGCTTGCAATTGGTCAACAATTGCGTTTACATGAAGCTACCTTGGTTTGGGTATGAT GTTGAACACTTCGACTTCGAAGATCACACCAACTACGAATTGGATATCAAAGAAACCTTGACCTTGAAG CCAAAGGGTTTGTGTTAAGGCTAAGTCCAAAAAGATTCCATTGGGTGGTATTCCATCTCCATCTACTG AACAATCCGCTAAGAAGGTTAGAAAGAAAGTGAACACGCTCATAACACACCTTTGTTGGTCTTGTACGG TTCTAATATGGGTACTGCTGAAGGTACAGCAAGAGATTTGGCAGATATTGCTATGTCTAAAGGTTTCGCT CCACAAGTTGCTACTTTGGATTCTCATGCTGGTAATTTGCCAAGAGAAGGTGCTGTTTTGATAGTTACTG CTTCCTACAATGGTCACCCACCAGATAATGCTAAGCAATTCGTTGATTGGTTGGATCAAGCTTCAGCTGA TGAAGTAAAAGGTGTTAGATACTCTGTTTTCGGTTGCGGTGACAAAAATTGGGCTACTACTTATCAAAA GGTTCCAGCCTTTATTGACGAACTTTGGCTGCTAAAGGTGCTGAAAACATTGCTGACAGAGGTGAAGCT GATGCTCCGACGACTTCGAAGGTACTTACGAAGAATTGGAGAGAACACATGTGGCTCTGACGTTGTCTGCT ACTTCAACTTGGACATCGAAAACCTGGAAGACAACAAGTCCACTTTGTCTTTGGTACTGTTGACTCCGC TGCTGACATGCCATTGGCTAAGATGCACGGTGCTTTCTCTACCAACGTCGTTGCCCTCCAAGGAATTGCAA CAACCAGGTTCTGCTAGATCTACTAGACACTTGGAAATCGAATTGCCAAAGGAAGCTTCTACCAAGAAG GTGACCACTTGGGCGTTATTCCAAGAAACTACGAAGGTATCGTCAACAGAGTTACTGCTAGATTCCGTTT GGATGCTTCTCAACAAATCAGATTAGAAGCTGAAGAAGAAAAGTTGGTCACTTGCATTAGCTAAGAC TGTCTCCGTTGAAGAATTGTTGCAATACGTCGAATTGCAAGACCCAGTTACCAGAACCCAATTGAGAGCC ATGGGTGCCAAGACCGTCTGTCCACCACACAAGGTTGAATTGGAAGCCTTGTGGAAAAGCAAGCCTACA AGGAACAAGTTTGGCTAAGAGATTGACCATGTTGGAATTGTTGGAAAAGTACCCAGCCTGCCAAATGA AGTTCTCTGAATTTATCGCCTTGTGGCATCTATCAGACCACGTTACTACTCTATTCTTCTCTCCACGT GTTGACGAAAAGCAAGCTTCTATTACTGTTTCCGTTGTCTCCGGTGAAGCTTGGTCCGGTTACGGTGAAT ACAAGGTATTGCTTCTAACTACTTGGCTGAATTGCAAGAAGGTGACACCACTACTTGTTCATCTCTAC TCCCAATCCGAATTTACTTTGCCAAAGGACCCAGAAACTCCATTGATCATGTTGGTCCAGGTACTGGT GTCGCTCCATTCAGAGGTTTCGTTCAAGCTAGAAACAATTGAAGGAACAAGGTCAATCTTTGGGTGAA GCTCACTTGTAATCTCGTTGTAGATCTCCACACGAAGACTACTTATACCAAGAAGAATTGGAAAACGCTC AATCCGAAGGTATTATCACTTTGCACACCGCTTTCTCCAGAATGCCAAACCAACCAAGACTTACGTCCA ACACGTTATGGAACAAGACGGTAAGAAGTTGATTGAATTGTTGGACCAAGGTGCTCACTTCTACATTGT GGTGATGGTTCTCAATGGCTCCAGCCGTTGAAGCCACTTTGATGAAGTCTTACGCTGATGTTACCAAG TTTCCGAAGCCGATGCTAGATTATGGTTGCAACAATTGGAAGAAAAGGTGCTTACGCTAAGGATGTCT GGGCCGGTTGA
4H5	ATGACCATCAAAGAAATGCCACAACCTAAGACTTTCGGTGAATTGAAGAATTTGCCTTTGTTGAACACCG ATAAGCCAGTTCAAGCTTTGATGAAGATTGCTGATGAATTGGGTGAAATCTTCAAGTTTGAAGCTCCAG GTAGAGTCACTAGATACTTGTCTCATCAAGATTGATCAAAGAAGCCTGCGACGAATCCAGATTTGATA AGAATTTGTCTCAAGCTGCTAAGTTTCGCTAGAGATTTTGGTGGTGATGGTTTGGTTACTTCTTGGACTCA CGAAAAGAATTGGAAGAAGGCCCATTAACATTTTGTGGCATCTTCTCACAACAAGCCATGAAGGGTTAT CATGCTATGATGGTTGATATCGCCGTTCAATTGGTTCAAAAGTGGGAAAGATTGAACGCCGATGAACAT ATCGAAGTCTCTGAAGATATGACCAGATTGACCTTGGATACCATTGGTTTGTGTGGTTTCACTACAGAT TCAACTCCTTCTACAGAGATCAACCACATCCATTTCATCATCTCTGCTGTTAGAGCTGCAGATGAAGTCAT GAACAAATGCAAAGAGCTAATCCAGACGATCCAGCTTATGACGAAAACAAGAGACAATTCGAAGAAGA TATCAAGGTCATGAACGATTTGGTCGATAAGATTATCGCTGATAGAAAGGCTAGAGGTGAACAATCTGA TGATTTGTGACCCAAATGTTGAACGGTAAGGATCCAGAAACTGGTGAACCATTTGGATGATGGTAACAT CAGATACCAAATATCACCTTCTTGATTGCTGGTCACGAACTACATCTGGTTTGTGTCTTTTGCCTTG TACTTTTGGTTAAGAACCCACAGTCTTGCAAAAGGTTGCTGAAGAAGCTGCAAGAGTTTGGTTGATC

	CAGTTCCATCTTACAAGCAAGTCAAGCAATTGAAGTACGTTGGTATGGTTTTGAACGAAGCTTTGAGATT GTGGCCAACTGCTCCAGCTTTTTCATTATACGCTAAAGAAGATACCGTCTTGGGTGGTGAATATCCATTG GAAAAAGGTGATGAAGTTATGGTCTTGATCCCACAATTGCATAGAGATAAGACTGTTTGGGGTGATGAT GTCGAAGAATTGAGACCAGAAAGATTGCAAAAACCCATCTGCTATTCCACAACATGCTTTTAAGCCATTTG GTAACGGTCAAAGAGCTTGCAATTGGTCAACAATTCGCTTTACATGAAGCTACCTTGGTTTTGGGTATGAT GTTGAAACACTTCGACTTCGAAGATCACACCAACTACGAATTGGATATCAAAGAAACCTTGACCTTGAAG CCAAAGGGTTTTGTTGTTAAGGCTAAGTCCAAAAAGATTCCATTGGGTGGTATTCCATCTCCATCTACTG AACAATCCGCTAAGAAGGTTAGAAAGAAAGCTGAAAACGCTCATAACACACCTTTGTTGGTCTTGTACGG TTCTAATATGGGTACTGCTGAAGGTACAGCAAGAGATTTGGCAGATATTGCTATGTCTAAAGGTTTCGCT CCACAAGTTGCTACTTTGGATTCTCATGCTGGTAATTTGCCAAGAGAAGGTGCTGTTTTGATAGTTACTG CTTCTTACAATGGTCAACCCACCAGATAATGCTAAGCAATTCGTTGATTGGTTGGATCAAGCTTCAGCTGA TGAAGTAAAAGGTGTTAGATACTCTGTTTTCGGTTGCGGTGACAAAAATTGGGCTACTACTTATCAAAA GGTTCAGCCTTTATTGACGAACTTTGGCTGCTAAAGGTGCTGAAAAATTGCTGACAGAGGTGAAGCT GATGCCCTCCGACGACTTCGAAGGTACTTACGAAGAATGGAGAGAACACATGTGGTCTGACGTTGCTGCTT ACTTCAACTTGGACATCGAAAACCTCTGAAGACAACAAGTCCACTTTGTCTTTGCAATTCGTTGACTCCGC TGCTGACATGCCATTGGCTAAGATGCACGGTGCTTTCTCTACCAACGTCGTTGCCTCCAAGGAATTCGCA CAACCAGGTTCTGCTAGATCTACTAGACACTTGAAATCGAATTGCCAAAGGAAGCTTCCTACCAAGAAG GTGACCACTTGGGCGTTATTCCAAGAACTACGAAGGTATCGTCAACAGAGTTACTGCTAGATTCCGTTTT GGATGCTTCTCAACAAATCAGATTAGAAGCTGAAGAAGAAAAGTTGGCTCACTTGCCATTAGCTAAGAC TGTCTCCGTTGAAGAATTGTTGCAATACGTCGAATTGCAAGACCCAGTTACCAGAACCCAATTGAGAGCC ATGGCTGCCAAGACCGTCTGTCCACCACACAAGGTTGAATTGGAAGCCTTGTGGAAAAGCAAGCCTACA AGGAACAAGTTTTGGCTAAGAGATTGACCATGTTGGAATTGTTGGAAAAGTACCAGCCTGCGAAATGA AGTTCTCTGAATTTATCGCCTTGTGCCATCTATCAGACCACGTTACTACTCTATTTCTTCTCTCCACGT GTTGACGAAAAGCAAGCTTCTATTACTGTTTCCGTTGTCTCCGGTGAAGCTTGGTCCGGTTACGGTGAAT ACAAGGTATTGCTTCTAACTACTTTGGCTGAATTGCAAGAAGGTGACACCATTACTTGTTCATCTCTAC TCCACAATCCGAATTTACTTTGCCAAAGGACCCAGAACTCCATTGATCATGGTTGGTCCAGGTACTGGT GTCGCTCCATTGAGAGTTTCGTTCAAGCTAGAAAACAATTGAAGGAACAAGTCAATCTTTGGGTGAA GCTCACTTGTACTTCGTTGTAGATCTCCACAGAACTACTTATACCAAGAAGAATTGGAACAGCTC AATCCGAAGGTATTATCACTTTGACACCCGCTTTGCTCAGAAATGCCAAACCAACCAAGACTTACGTCCA ACACGTTATGGAACAAGACGGTAAGAAGTTGATTGAATTGTTGGACCAAGGTGCTCACTTCTACATTTGT GGTGATGGTTCTCAAATGGCTCCAGCCGTTGAAGCCACTTTGATGAAGTCTTACGCTGATGTTACCAAG TTTCCGAAGCCGATGCTAGATTATGGTTGCAACAATTGGAAGAAAAGGTGCTTACGCTAAGGATGTCT GGGCCGTTGA
7A1	ATGACCATCAAAGAAATGCCACAACCTAAGACTTTTCGGTGAATTGAAGAATTTGCTTTTGTGAACACCG ATAAGCCAGTTCAAGCTTTGATGAAGATTGCTGATGAATTGGGTGAAATCTTCAAGTTTGAAGCTCCAG GTAGAGTCACTAGATACTTGTCTCTCAAAGATTGATCAAAGAAGCCTGCCAGCAATCCAGATTTGATA AGAAATTTGTCTCAAGCTGCTAAGTTGCTAGAGATTTTGTGCTGGTGATGGTTTGGTTACTTCTTGGACTCA CGAAAAGAATTGGAAGAAGGCCCATAACATTTTGTGGCATCTTTCTCACAACAAGCCATGAAGGGTTAT CATGCTATGATGGTTGATATCGCCGTTCAATTGGTTCAAAAGTGGGAAAGATTGAACGCCGATGAACAT ATCGAAGTCTCTGAAGATATGACCAGATTGACCTTGGATACCATTGGTTTGTGTGGTTTCACTACAGAT TCAACTCCTTCTACAGAGTCAACACATCCATTCATCATCTCTGCTGTTAGAGCTGCAGATGAAGTCAT GAACAAATTTGCAAAGAGCTAATCCAGACGATCCAGCTTATGACGAAAACAAGAGACAATTCCAAGAAGA TATCAAGGTCATGAACGATTTGGTCGATAAGATTATCGCTGATAGAAAGGCTAGAGGTGAACAATCTGA TGATTTGTTGACCCAAATGTTGAACGGTAAGGATCCAGAACTGGTGAACCATTTGGATGATGGTAACAT CAGATACCAAATATCGCTTTCTTGATTGCTGGTCACGAACTACATCTGGTTTGTGTCTTTTGCTTG TACTTTTGGTTAAGAACCACACGCTTTGCAAAAGGTTGCTGAAGAAGCTGCAAGAGTTTGGTTGATC CAGTTCCATCTTACAAGCAAGTCAAGCAATTGAAGTACGTTGGTATGGTTTTGAACGAAGCTTTGAGATT GTGGCCAACTGCTCCAGCTTTTTCATTATACGCTAAAGAAGATACCGTCTTGGGTGGTGAATATCCATTG GAAAAAGGTGATGAAGTTATGGTCTTGATCCCACAATTGCATAGAGATAAGACTGTTTGGGGTGATGAT GTCGAAGAATTGAGACCAGAAAGATTGCAAAAACCCATCTGCTATTCCACAACATGCTTTTAAGCCATTTG GTAACGGTCAAAGAGCTTGCAATTGGTCAACAATTCGCTTTACATGAAGCTACCTTGGTTTTGGGTATGAT GTTGAAACACTTCGACTTCGAAGATCACACCAACTACGAATTGGATATCAAAGAAACCTTGACCTTGAAG CCAAAGGGTTTTGTTGTTAAGGCTAAGTCCAAAAAGATTCCATTGGGTGGTATTCCATCTCCATCTACTG AACAATCCGCTAAGAAGGTTAGAAAGAAAGCTGAAAACGCTCATAACACACCTTTGTTGGTCTTGTACGG TTCTAATATGGGTACTGCTGAAGGTACAGCAAGAGATTTGGCAGATATTGCTATGTCTAAAGGTTTCGCT CCACAAGTTGCTACTTTGGATTCTCATGCTGGTAATTTGCCAAGAGAAGGTGCTGTTTTGATAGTTACTG CTTCTTACAATGGTCAACCCACCAGATAATGCTAAGCAATTCGTTGATTGGTTGGATCAAGCTTCAGCTGA TGAAGTAAAAGGTGTTAGATACTCTGTTTTCGGTTGCGGTGACAAAAATTGGGCTACTACTTATCAAAA GGTTCAGCCTTTATTGACGAACTTTGGCTGCTAAAGGTGCTGAAAAATTGCTGACAGAGGTGAAGCT GATGCCCTCCGACGACTTCGAAGGTACTTACGAAGAATGGAGAGAACACATGTGGTCTGACGTTGCTGCTT

ACTTCAACTTGGACATCGAAAACCTCTGAAGACAACAAGTCCACTTTGTCTTTGCAATTTCGTTGACTCCGC TGCTGACATGCCATTGGCTAAGATGCACGGTGCTTTCTCTACCAACGTCGTTGGCTCCAAGGAATTGCAA CAACCAGGTTCTGCTAGATCTACTAGACACTTGGAAATCGAATTGCCAAAGGAAGCTTCTTACCAAGAAG GTGACCACTTTGGGCGTTATTCCAAGAACTACGAAGGTATCGTCAACAGAGTTACTGCTAGATTTCGGTTT GGATGCTTCTCAACAAATCAGATTAGAAGCTGAAGAAGAAAAGTTGGCTCACTTGGCATTAGCTAAGAC TGCTCCCGTTGAAGAATTGTTGCAATACGTCGAATTGCAAGACCCAGTTACCAGAACCCAATTGAGAGCC ATGGCTGCCAAGACCGTCTGTCCACCACACAAGGTTGAATTGGAAGCCTTGTGGAAAAGCAAGCCTACA AGGAACAAGTTTGGCTAAGAGATTGACCATGTTGGAATTGTTGGAAAAGTACCCAGCCTGCCAAATGA AGTTCTCTGAATTTATCGCCTTGTGCCATCTATCAGACCACGTTACTACTCTATTTCTTCTCTCCACGT GTTGACGAAAAGCAAGCTTCTATTACTGTTTCCGTTGTCTCCGGTGAAGCTTGGTCCGGTTACGGTGAAT ACAAGGGTATTGCTTCTAATCTACTTGGCTGAATTGCAAGAAGGTGACACCATTACTTGTTCATCTCTAC TCCACAATCCGAATTTACTTTGCCAAAGGACCCAGAAAACCTCCATTGATCATGGTTGGTCCAGGTACTGGT GTCGCTCCATTACAGAGGTTTCGTTCAAGCTAGAAAACAATTGAAGGAACAAGGTCAATCTTTGGGTGAA GCTCACTTGTACTTCGGTTGTAGATCTCCACACGAAGACTACTTATACCAAGAAGAATTGGAAAACGCTC AATCCGAAGGTATTATCACTTTGCACACCGCTTCTCCAGAATGCCAAACCAACCAAGACTTACGTCCA ACACGTTATTGGAACAAGACGGTAAGAAGTTGATTGAATTGTTGGACCAAGGTGCTCACTTCTACTTTGT GGTGATGGTTCTCAAATGGCTCCAGCCGTTGCAAGCCACTTTGATGAAGTCTTACGCTGATGTTTACCAAG TTTCCGAAGCCGATGCTAGATTATGGTTGCAACAATTGGAAGAAAAGGTGCTTACGCTAAGGATGTCT GGGCCGGTTGA

Table 8. pA24, pA25, and pA26 sequences

pA24 Sequence	cctcgccgcagttaataaagtcagtgcgaggaagcgcgtaactataacggctctaagtagcgaatcctgatcggtattttccttacgcatctgtg cggatttcacaccgcatagatcgcaagtcacaaacaataacttaataaataactactcagtaataacattttcttagcatttttgacgaattgtcatttt tgtagagctctttacaccattgtctccacactccgcttacatcaacaccaataacgcatcttaataagcgcatcaccaacatttttggcgctcagtc ccagctaacataaaatgaagcttcggggctctctgcttccaaccagtcagaaatcaggtccaatccaaaagttcacctgtcccactgtctctgaat caacaagggaataaacgaatgaggtttctgtgaagctgcactgagtagtagtgcagtcctttggaaatcagagctcttaataactggcaaacggagg aactcttggtattctgacgactcatctcatgagtggaacaaatcatttgcggtcaactttggacgatatcaatgccgtaacattgaccagagcc aaaacatctccttaagttgattacgaacacgccaaccaagtaatttcggagtgctgaactattttatagcttttacaagactgaaattttccttgaat aacgggtcaattgtctcttctattgggcacacataataaccagcaagtcagcatcggaatctagagcacattctgcggcctctgtgctctgcaagcc gcaaatcttcacaaatggaccagaactcctgtgaaattaataacagacatactccaagtcgcttctgtgcttaatacgtatactacgtgctcaatag tcaccaatgccctcctcttgccctctcttcttttttcgaccgaattaattcttaacggcaaaaaagaaaagtcggatcaagattgtacgtaaggt gacaagctatttttaataaagaatatcttcactactgccaatctggcgctcataactgcaagtacacatatattacgatgctgtcttataaagcttctctat attatatatatagtaatgtcgtgatctatggtgactctcagtaacatctgctctgatccgcatagttaagccagcccgacacccgcaacacccgctga cgccctgacgggctgtgctcctccgcatccgcttacagacaagctgtgacgctcctcgaggagctgcatgtgtagaggttttaccgctcatccgga aacgcgcgagagcaaaaggcctctgtagacgctatttttaggttaattgtcatgataataatgggttcttagacggatcgcttgcctgttaacttacacgc gcctctgtatctttatagtggaataaattgggaattactctgtgtttatttttttattgttttatttggatttttagaaagtaataaagaaggtagaagag ttacggaatgaagaaaaaaataacaaagggttaaaaaattcaacaaaaagcgtactttacatatattatttagacaagaaaagcagattaaata gatatacattcgattaacgataagtaaatgtaaaatcacaggattttcgtgtgtgcttctacacagacaaggtgaacaattcgccattaatacttgag agcaggaagagcaagataaaagtagtattgttggcgatccccctagagctttttacatcttcggaaaacaaaaatatttttcttaattcttttttact ttctatttttaatttatatttatataaaaaatttaattataatttttatagcagctgatgaaaaggaccaggtggcattttcggggaaatgtgctgc ggaaacctattgtttatttttcaataacattcaaatatgtatccgctcatgagacaataaccctgataaattgctcaataattgaaaggaagagta tgagtattcaacatttcgtgtgccttattcccttttttggcgattttgcttctctgttttctcaccagaaacgctgtgtaagtaaaagatgctgaag atcagttgggacgcgtagctagaccagcaggacagaaatgcctcgactctgctgctacccaaggttgcgggtgacgcacacgttggaacggatga aggcagcaaacccagtgacataagcctgttcggtctgaagctgtaagtagcgtatgctcagcaactggtccagaaccttgaccgaacgcga cgggtggaacggcgagtgccggttttcatggtctgtatgactgttttttgggtacagctctatgctcgggcatccaagcagcaagcgcttacgccg tgggtcgatgtttgatgttatggagcagcaacgatgttacgcagcaggcgagtcgccccaaaaacaaagtttaacattatgaggaagcgggtgatccgcg aagtatcgactcaactatcagaggttagttggcgccatcgagcgccatctcgaaacggacgttgcggcggtacatttgcagcctccgagtggtgagcg cctgaagccacacagtgatattgattgtggttacggtgacgtaaggcttgatgaacaacgcggcgagcttgatcaacgacattttgaaacttcgg cttcccctggagagcgagattctccgctgtagaagtcaccattgtgtgcacgacacattcctggtgcttatcagctaaagcgcgaactgcaa ttggagaatggcagcgcaatgacattcttcaggtatcttcgagccagccagatcgacattgatctggctatctgctgacaaaagcaagagaacatag cgttgcttgtaggtccagcgccggagggaactcttgatccggttctgaacaggatctatttgaggcgtaaatgaaccttaacgctatggaactcgc cgcccgactgggctggcgatgagcgaatgtagtgcctacgtgtgtcccgacttggtagcgcagtaaacggcgaataacgcgcgcaaggtatgctgctgc cgctgctggcgaatggagcgctgcccggcagatcagccgtctacattgtagcagagcagcttatcttggacagaagaagatcgttggcctgcgcg gcagatcagttggaagaattgtccactacgtgaaaggcgagatccaaggtgtagcggaataaacctcgagcattcaaggcgcttgattttgac gtggtttgatggcctccacgcagctgtgatagtagatgattcagttcgagtttatcattatcaataactgcccatttcaagaataacgtaataataatagta gtgattttcctaactttatttagtaaaaaattagccttttaattctgctgaacccgacatgccccaaaataggggcggggttacagaataataacatc gtaggtgtctgggtgaacagtttattcctggcatccactaaataatggagccgctttttaagctggcatccagaaaaaaaagaatcccagcaccaaa atattgtttcttcaccaaccatcgattcataggtccattctctagcgcaactacagagaacaggggcacaacacaggcaaaaacggcacaacctcaat ggagtgatgcaacctgcctggagtaaatgatgacacaaggcaattgaccacgcgatgtatctatcttcttacccttctattaccttctgctctctctg atttggaaaaagctgaaaaaaagggtgaaccagttccctgaattatccctacttgactaataagtatataaagcggtaggtattgattgtaattct gtaaatctatttctaaacttctaaattctattttatagtagtcttttttagtttttaaacaccaagaacttagtttgaataaacacacataaacaaca aaacaggccccctttcttctgtgatcatgtaattagttatgtcacgcttaccctcacgccccctccacatcgctctaaccgaaaagggaaggtag
--------------------------	--

-173-

accaataacgccatttaataagcgcatcaccaacattttctggcgctcagtcaccagctaacataaaatgaagcttcggggctctcttgcctccaac cagtcagaaatcgagttccaatcctgaatcctgtccacgtctctgaatcaaacagggaataaacgaatgaggtttctgtgaagctgactga gtagtatgttcagctctttggaaatcagagcttttaataactggcaaacagggaactcttggtattcttccacgactcatctccatgcagtgaggc tcaattcttgcggtaactttggacgatatcaatgcgtaacattgaccagagcgaataatctcttaagttgattacgaaacacgccaaccaagatt tcggagtgctgaactattttatatgctttacaagactgaaatcttctgcaataacggggtcaattgttctcttattgggcacacataataacca gcaagtcagcatcggaatctagagcacattctgcggcctctgtgctctgcaagccgcaaaccttcaccaatggaccagaactacgtgaaataaaca gacatactccaagctgcctttgtgtctaatcacgtatactcacgtgctcaatagtcaccaatgccctccctcttggccctctctttcttttcgaccgaat taattcttaactggcaaaaaaagaaagctccggatcaagattgtacgtaagggtgacaagctattttcaataaagaaatctccactactgccaactctggc gtcataactgcaaaagtacacatatattacgtagctgttctattaaatgcttctattatataatagtaattgctgtagctatgggtgcaactctcagtagacaatc tgctctgatgcgcatagttaagccagcccgacacccgccaacacccgctgacgctgctgacgggctgtctgctccggcatccgcttacagacaag ctgtgaccgtctccgggagctgcatgtgtcagagggtttcaccgtcatcaccgaaacgcgcgagacgaaaggccctcgtgatacgcctattttataggtta atgtcatgataaataaggtttcttagacggatcgttgcctgtaacttacacgcgcctcgtatcttttaagtaggaataattgggaattactctgtgttatt tattttatgtttgtattggatttagaaagtaataaagaaggtagaagattacggaatgaagaaaaaaataaacaaggttaaaaaatttcaa caaaaagcgtaactttacatatattttatgacaagaaaagcagattaaatagatatatacttgcatttaacgataagtaaaatgtaaaatcacaggattttc gtgtgtggtctttcacacagacaaggtgaaacaattcggcattaatcctgagagcaggaagagcaagataaaaagtagtattttgttggcgatccccc gagctttttacatcttcggaacaaaaactattttcttaattctttttactttctatttttaattatattatataaaaaattaaattataattttt tatagcagtgatgaaaaggaccaggtggcacttttcggggaaatgtgcgcggaacccctattgtttattttctaaatacattcaaatatgtatccgctc atgagacaataacccgtataatgtctcaataatattgaaaaggaagagtatgagtattcaacattccgtgtgccttattcccttttttcggcattttg ccttctgttttctcaccagaaacgtggtgaaagtaaaagatgctgaagatcagttgggacgctagtctagaccagccagagacagaatgcctcg acttctgctacccaaggttgcgggtgacgcaacccgtggaaacggatgaaggcacgaacccagtggaacataagcctgttcgggtcgtaagctgta gcaagtagcgtatgcgtcacgcaactggtccagaaccttgaccgaacgcagcgggtgtaacggcgagtgccgggtttctagctgtgtatgactgtttt ttgggtacagctatgcctcgggcatccaagcagcaagcgctgtacgcggtgggtgcatgtttgatgttatggagcagcaacatgttacgcagcagg cagtcgcccataaaaggttaacattatgagggaaagcgggtgatcgcgaaagtatcgactcaactatcagaggtatgtggcgccatcgagcgaatctc gaaccgacgttctggcgctacattgtacggctcgcagtggtgagggcctgaagccacacagtgatattgttgcgtgtacgggtacgctaaggct tgatgaacaacgcggcgagctttgatcaacgacctttggaaactcggcttccctggagagagcagagattctccgctgtagaagtcaccattgtt tgacgacgacatcattccgtggcggttatccagtaagcgcgaactgcaatttggagaatggcagcgcaatgacattctgcaggtatcttcagccagcc acgatcgacattgatctggctatcttctgacaaaagcaagagaacatagcgttgccttggtaggtccagcggcgagggaactcttgatccggttctga acaggatctatttgagcgctaaatgaaccttaacgctatggaactcgcgcgcccagctgggtggcgatgagcgaaatgtagtcttacgttgcctccgca tttggtacagcgagtaaccggcaaaatcgccggaaggatgtcgtcgggtgggcaatggagcgctgccggccagtatcagccgctacatttg aagctagacaggcttatttggaagaagaagatcgcttggcctcgcgcgagatcagttggaagaattgtccactacgtgaaaggcgagatcacca aggtagtcggcaataaacctcagcattcaaggcgcttattttgacgtggtttgatggcctccacgcacggttgatagtagtagactcgttaggaa caatttcgggcccctgcgtgttcttgcaggttcatctttacatttgccttctggtgataatttcagaggcaacaaggaataatagatggcaaaaagtcg tcttcaaggaaaaatccccaccatcttctgagatccctgtaacttattggcaactgaaagaatgaaaggaggaaaaatacaaaatatacagaactga aaaaaaaagtataatagagacgatatatgccaatactcacaatgttcgaatctattcttatttgagctattgtaaaataataaacatcaagaac aaacaagctcaactgtcttttcaagaacaaagaataaacacaaaaacaaaagtttttaatttaatacaaaaaacaggcccttttcttctgcatat catgtaattagttagtcacgcttacattcacgccccccacatccgcttaaccgaaaagggaaggagttagacaacctgaagctaggtccctattta ttttttatagttatgttagtattaagaacgttatttatatttcaaatttttcttttttctgtacaaacgcgtgtacgcatgtaacattatctgaaaacgttgc gagaagggtttgggacgctcgaaggcttaattgtaatcattatcattacgggtcctttccgggtatccgacaggttacggggcgcgacctcgcgggt tttcgtatttatgaaaatttccggttaaggcgtttccgttcttctcgtcataactaatgtttttatataaacctcgcgagtggaacactgaaaatac ccatggagcggcgttaaccgtcgcacaggatctaggtgaagatccttttgataatctcatgacaaaatcccttaacgtgagttttcgttccactgagcgtc agaccctgataaaaagatcaaaggatcttctgagatcctttttctgcgctaattctgctgtgcaaaaaaaaccacgcgtaccagcggtggtttg tttcgggatcaagagctaccaactcttttccgaaggtaactggcttcagcagagcgcagatacaaaatctgtccttctagttagccgtagttaggcca ccacttcaagaactctgtagcaccgcctacatacctcgtctgtaatcctgttaccagtggtgctgcccagtgaggaataagctgtcttaccgggttga ctcaagacgatagttaccgataaggcgagcgtcgggctgaacgggggttctgtgcacacagccagcttgagcga
--

Table 9. Tailoring enzymes

<i>Reaction Catalyzed</i>	<i>Enzyme</i>	<i>Species</i>
Carbon-carbon coupling	Berberine bridge enzyme (BBE)	Ps, Ec, Cj, Bs, Tf
	Salutaridine synthase (SalSyn)	Ps
	Corytuberine synthase (CorSyn)	Cj
Oxidation	Tetrahydroprotoberberine oxidase (STOX)	Cj, Am, Bw
	Dihydrobenzophenanthridine oxidase (DBOX)	Ps
	Methylstylopine hydroxylase (MSH)	Ps
	Protopine 6-hydroxylase (P6H)	Ps, Ec
Methylenedioxy bridge formation	Stylopine synthase (StySyn)	Ps, Ec, Am
	Chellanthifoline synthase (CheSyn)	Ps, Ec, Am
	Canadine synthase (CAS)	Tf, Cc
O-methylation	Norcoclaurine 6-O-methyltransferase (6OMT)	Ps, Tf, Cj, Pb
	3'-hydroxy-N-methylcoclaurine 4'-O-methyltransferase (4'OMT)	Ps, Tf, Cj, Cc
	Reticuline 7-O-methyltransferase (7OMT)	Ps, Ec
	Scoulerine 9-O-methyltransferase (9OMT)	Ps, Tf, Cj, Cc
N-methylation	Coclaurine N-methyltransferase (CNMT)	Ps, Tf, Cj
	Tetrahydroprotoberberine N-methyltransferase (TMNT)	Ps, Ec, Pb
O-demethylation	Thebaine demethylase (T6ODM)	Ps
	Codeine demethylase (CODM)	Ps, Ga
Reduction	Salutaridine reductase (SalR)	Ps, Pb, Ga
	Codeinone reductase (COR)	Ps
	Sanguinarine reductase (SanR)	Ec
Acetylation	Salutaridine acetyltransferase (SalAT)	Ps

Table 10. Comparison of impurities that may be present in concentrate of poppy straw and clarified yeast culture medium.

Impurities:		Concentrate of Poppy Straw	Clarified Yeast Culture Medium
Inorganic	Sodium	✓	✓
	Magnesium	✓	✓
	Silicon	✓	✗ (not soluble in water)
	Phosphorus	✓	✓
	Sulfur	✓	✓
	Chloride	✓	✓
	Potassium	✓	✓
	Calcium	✓	✓
	Copper	✓	✓
	Zinc	✓	✓
	Molybdenum	✓	✓ (not soluble in water)
	Iron	✓	✓
	Manganese	✓	✓
	Ammonium	✓	✓
	Boron	✓	✓
Organic	Polysaccharides (starch, cellulose, xylan)	✓	✗ (not sol. in water)
	Lipids (g-mucilage, wax, sterol, sterol succinate)	✓	✗
	Pigments (chlorophyll, anthocyanins, carotenoids)	✓	✗
	Flavonoids	✓	✗
	Phenanthroic acids	✓	✗
	Latex, gum, and wax	✓	✗
	Rubicon	✓	✗
	Shikonic acid	✓	✗
	Pseudonitroglycine	✓	✗
	Nitrobenzene	✓	✗
	Theobromine	✓	✗
Other	Pesticides	✓	✗
	Protein	✓	✗

Table 11. Distinct groups of molecules present in clarified yeast culture medium (CYCM). Unlike concentrate of poppy straw (CPS), yeast host strains may be engineered to produce molecules of a predetermined class of alkaloids (i.e., only one biosynthesis pathway per strain) such that other classes of alkaloids are not present. Therefore, the CYCM may contain molecules within a single biosynthesis pathway including a subset of molecules spanning one or two columns, whereas the CPS may contain a subset of molecules across many columns.

4-Benzylisoquinoline	Protuberberine and Pthalideisoquinoline	Morphine	Isopavine	Apoptine	Brada
Tetrahydrogambogine	Scolimine	Salsoline	Pavine	Magnofoline	Staurine
Dihydrogambogine	Diamorphiline	Salsolactone	Gambogine	Corydalline	Scopolamine
Papaverine	Silypine	Solidoline-7-O-acetate	Hydrogambogine	Apocryptine	Lignosidine
	6a,6-methylsilypine	Naltrexone	Isomorphamine	Bulbine	Fargoline
	Protocle	Codolone			Tetradolone
	Ethylgambogine	Opavine			Cine
	Sanguinarine	Morphine			Cephaeline
	Tetrahydrocolumbamine	Necoptine			Berberine
	Canadine	Necoptine			
	N-methylgambogine	Cocaine			
	Necoptine	Morphine			
	Berberine	Necoptine			
		Hydrocaine			
		Oxycaine			
		14-hydrogambogine			
		12-hydrogambogine			
		Dihydrogambogine			
		Hydrogambogine			

Table 12. Impurities that may be present in chemical synthesis preparations of compounds

Compound	Impurities
Buprenorphine	15,16-Dehydrobuprenorphine, 17,18-Dehydrobuprenorphine, 18,19-demethylbuprenorphine, 19,19'-Ethylbuprenorphine, 2,2'-Bisbuprenorphine, 3-Deshydroxybuprenorphine, 3-O-Methylbuprenorphine, 3-O-Methyl-N-cyanonorbuprenorphine, 3-O-Methyl-N-methylnorbuprenorphine, 6-O-Desmethylbuprenorphine, Buprenorphine N-oxide, N-But-3-enylnorbuprenorphine, N-But-3-enylnormethylbuprenorphine, N-Butylnorbuprenorphine, N-Methylbuprenorphine, Norbuprenorphine, Tetramethylfuran buprenorphine
Oxymorphone	1-Bromooxymorphone, 6-Beta oxymorphol, 10-Alpha-hydroxyoxymorphone, 10-Ketooxymorphone, 2,2-Bisoxymorphone, Noroxymorphone, Oxymorphone N-oxide, 10-Hydroxyoxymorphone, 4-Hydroxyoxymorphone, 8-Hydroxyoxymorphone, Hydromorphinol.
Naltrexone	10-Hydroxynaltrexone, 10-Ketonaltrexone, 14-Hydroxy-17-cyclopropylmethylnormorphinone, 2,2'-Bisnaltrexone, 3-Cyclopropylmethylnaltrexone, 3-O-Methylnaltrexone, 8-Hydroxynaltrexone, N-(3-Butenyl)-noroxymorphone, Naltrexone aldol dimer, N-Formyl-noroxymorphone
Naloxone	10-Alpha-hydroxynaloxone, 10-Beta-hydroxynaloxone, 10-Ketonaloxone, 3-O-Allylnaloxone, 7,8-Didehydronaloxone, 2,2'-Bisnaloxone, Naloxone N-oxide
Nalbuphine	Beta-epimer of nalbuphine, 2,2'-Bisnalbuphine, 6-Ketonalbuphine, 10-Ketonalbuphine, Alpha-noroxymorphol, N-(Cyclobutylcarbonyl)-alpha-noroxymorphol, N-Formyl-6-alpha-noroxymorphol.

7. The method of clause 2, wherein said oxidase component comprises a sequence that is at least 85% homologous to SEQ ID NO. 17.

8. The method of clause 2, wherein said oxidase component comprises a sequence that is at least 90% homologous to SEQ ID NO. 17.

9. The method of clause 2, wherein said oxidase component comprises a sequence that is at least 95% homologous to SEQ ID NO. 17.

10. The method of clause 6, wherein said oxidase component comprises a sequence that is encoded by a DNA sequence which comprises one or more codon substitutions, relative to a wildtype sequence, at a codon position which corresponds to an amino acid position in SEQ ID NO. 17 selected from the group consisting of position 75, 160, 191, 256, 359, 366, and 500.

11. The method of clause 2, wherein said reductase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 18.

12. The method of clause 2, wherein said reductase component comprises a sequence that is at least 85% homologous to SEQ ID NO. 18.

13. The method of clause 2, wherein said reductase component comprises a sequence that is at least 90% homologous to SEQ ID NO. 18.

14. The method of clause 2, wherein said reductase component comprises a sequence that is at least 95% homologous to SEQ ID NO. 18.

15. A method of increasing production of an alkaloid product through the epimerization of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid via an engineered split epimerase in an engineered host cell in comparison to the epimerization of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid via a fused epimerase in similar conditions, said method comprising:

contacting said (*S*)-1-benzylisoquinoline alkaloid with said engineered split epimerase in said engineered host cell, thereby increasing production of an alkaloid product,

wherein contacting said (*S*)-1-benzylisoquinoline alkaloid with said engineered split epimerase converts said (*S*)-1-benzylisoquinoline alkaloid to said (*R*)-1-benzylisoquinoline alkaloid within said engineered host cell, and wherein said engineered split epimerase produces at least 1% more of said (*R*)-1-benzylisoquinoline alkaloid than a fused epimerase under similar conditions.

16. The method of clause 15, wherein said engineered split epimerase produces at least 3% more of said (*R*)-1-benzylisoquinoline alkaloid than a wildtype fused epimerase under similar conditions.

17. The method of clause 15, wherein said engineered split epimerase produces at least 5% more of said (*R*)-1-benzylisoquinoline alkaloid than a wildtype fused epimerase under similar conditions.

18. The method of clause 15, wherein said engineered split epimerase produces at least 10% more of said (*R*)-1-benzylisoquinoline alkaloid than a wildtype fused epimerase under similar conditions.

19. The method of clause 15, wherein said engineered split epimerase produces at least 15% more of said (*R*)-1-benzylisoquinoline alkaloid than a wildtype fused epimerase under similar conditions.

20. The method of clause 15, wherein said engineered split epimerase produces at least 20% more of said (*R*)-1-benzylisoquinoline alkaloid than a wildtype fused epimerase under similar conditions.

21. The method of clause 15, wherein said engineered split epimerase produces at least 25% more of said (*R*)-1-benzylisoquinoline alkaloid than a wildtype fused epimerase under similar conditions.

22. The method of clause 15, wherein said engineered split epimerase produces at least 30% more of said (*R*)-1-benzylisoquinoline alkaloid than a wildtype fused epimerase under similar conditions.

23. The method of clause 15, wherein said engineered split epimerase comprises one or more mutations that increase activity of said engineered split epimerase.

24. The method of clause 23, wherein said engineered split epimerase comprises one or more mutations selected from the group consisting of a localization mutation, a cytochrome P450 reductase interaction mutation, and an accessibility mutation.

25. The method of clause 15, wherein said engineered split epimerase comprises an oxidase component and a reductase component.

26. The method of clause 25, wherein said oxidase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 17.

27. The method of clause 25, wherein said oxidase component comprises a sequence that is at least 85% homologous to SEQ ID NO. 17.

28. The method of clause 25, wherein said oxidase component comprises a sequence that is at least 90% homologous to SEQ ID NO. 17.

29. The method of clause 25, wherein said oxidase component comprises a sequence that is at least 95% homologous to SEQ ID NO. 17.

30. The method of clause 25, wherein said oxidase component comprises a sequence that is encoded by a DNA sequence which comprises one or more codon substitutions, relative to a wildtype sequence, at a codon position which is equivalent to an amino acid position in SEQ ID NO. 17 selected from the group consisting of position 75, 160, 191, 256, 359, 366, and 500.

31. The method of clause 25, wherein said reductase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 18.

32. The method of clause 25, wherein said reductase component comprises a sequence that is at least 85% homologous to SEQ ID NO. 18.

33. The method of clause 25, wherein said reductase component comprises a sequence that is at least 90% homologous to SEQ ID NO. 18.

34. The method of clause 25, wherein said reductase component comprises a sequence that is at least 95% homologous to SEQ ID NO. 18.

35. An engineered host cell which comprises an engineered split epimerase, wherein said engineered split epimerase increases conversion of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid relative to conversion of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid via a fused epimerase under similar conditions.

36. The engineered host cell of clause 35, wherein said engineered split epimerase increases production of said alkaloid product in comparison to an epimerization via a parent fused epimerase under similar conditions.

37. The engineered host cell of clause 35, wherein said engineered split epimerase increases production of said alkaloid product in comparison to an epimerization via a wild-type fused epimerase.

38. The engineered host cell of clause 35, wherein said engineered split epimerase comprises one or more mutations that increase activity of said engineered split epimerase.

39. The engineered host cell of clause 35, wherein said engineered split epimerase comprises one or more mutations selected from the group consisting of a localization mutation, a cytochrome P450 reductase interaction mutation, and an accessibility mutation.

40. The engineered host cell of clause 35, wherein said engineered split epimerase comprises an oxidase component and a reductase component.

41. The engineered host cell of clause 40, wherein said oxidase component comprises a sequence that is encoded by a DNA sequence which comprises one or more mutations, relative to a wildtype sequence, each mutation at a codon position which corresponds to a position selected from the group consisting of positions 16, 17, 37, 50, 64, 69, 70, 110, 128, 155, 232, 234, 244, 293, 303, 328, 354, 387, 393, 427, and 477 within SEQ ID NO. 17.

42. The engineered host cell of clause 40, wherein said oxidase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 17 and is encoded by a DNA sequence which comprises one or more codon substitutions, relative to a wildtype sequence, at a codon position which corresponds to an amino acid position in SEQ ID NO. 17 selected from the group consisting of position 75, 160, 191, 256, 359, 366, and 500.

43. The engineered host cell of clause 40, wherein said reductase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 18.

44. An engineered split epimerase that comprises one or more mutations selected from the group consisting of a localization mutation, a cytochrome P450 reductase interaction mutation, and an accessibility mutation.

45. The engineered split epimerase of clause 44, wherein said engineered split epimerase comprises an oxidase component and a reductase component.

46. The engineered split epimerase of clause 44, wherein said engineered split epimerase has increased activity compared to a wildtype fused epimerase.

47. The engineered split epimerase of clause 44, wherein said engineered split epimerase comprises one or more mutations that increase activity of said engineered split epimerase.

48. The engineered split epimerase of clause 45, wherein said oxidase component comprises a sequence that is encoded by a DNA sequence which comprises one or more mutations, relative to a wildtype sequence, each mutation at a codon position which corresponds to a position selected from the

group consisting of positions 16, 17, 37, 50, 64, 69, 70, 110, 128, 155, 232, 234, 244, 293, 303, 328, 354, 387, 393, 427, and 477 within SEQ ID NO. 17.

49. The engineered split epimerase of clause 45, wherein said oxidase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 17 and is encoded by a DNA sequence which comprises one or more codon substitutions, relative to a wildtype sequence, at a codon position which corresponds to an amino acid position in SEQ ID NO. 17 selected from the group consisting of position 75, 160, 191, 256, 359, 366, and 500.

50. The engineered split epimerase of clause 45, wherein said reductase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 18.

51. An engineered host cell which comprises an engineered split epimerase, wherein said engineered split epimerase increases conversion of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid relative to conversion of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid via a fused epimerase under similar conditions.

52. The engineered host cell of clause 51, wherein said engineered split epimerase comprises one or more mutations that increase activity of said engineered split epimerase.

53. The engineered host cell of clause 51, wherein said engineered split epimerase comprises one or more mutations selected from the group consisting of a localization mutation, a cytochrome P450 reductase interaction mutation, and an accessibility mutation.

54. The engineered host cell of clause 51, wherein said engineered split epimerase comprises an oxidase component and a reductase component.

55. The engineered host cell of clause 54, wherein said oxidase component comprises a sequence that is encoded by a DNA sequence which comprises one or more mutations, relative to a wildtype sequence, each mutation at a codon position which corresponds to a position selected from the group consisting of positions 16, 17, 37, 50, 64, 69, 70, 110, 128, 155, 232, 234, 244, 293, 303, 328, 354, 387, 393, 427, and 477 within SEQ ID NO. 17.

56. The engineered host cell of clause 54, wherein said oxidase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 17 and is encoded by a DNA sequence which comprises one or more codon substitutions, relative to a wildtype sequence, at a codon position which is equivalent to an amino acid position in SEQ ID NO. 17 selected from the group consisting of position 75, 160, 191, 256, 359, 366, and 500.

57. An engineered split epimerase that comprises one or more mutations selected from the group consisting of a localization mutation, a cytochrome P450 reductase interaction mutation, and an accessibility mutation.

58. The engineered split epimerase of clause 57, wherein said engineered split epimerase has increased activity compared to a wildtype fused epimerase.

59. The engineered split epimerase of clause 57, wherein said engineered split epimerase comprises one or more mutations that increase activity of said engineered split epimerase.

60. The engineered split epimerase of clause 57, wherein said engineered split epimerase comprises an oxidase component and a reductase component.

61. The engineered split epimerase of clause 60, wherein said oxidase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 17.

62. The engineered split epimerase of clause 60, wherein said oxidase component comprises a sequence that is at least 85% homologous to SEQ ID NO. 17.

63. The engineered split epimerase of clause 60, wherein said oxidase component comprises a sequence that is at least 90% homologous to SEQ ID NO. 17.

64. The engineered split epimerase of clause 60, wherein said oxidase component comprises a sequence that is at least 95% homologous to SEQ ID NO. 17.

65. The engineered split epimerase of clause 60, wherein said reductase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 18.

66. The engineered split epimerase of clause 60, wherein said reductase component comprises a sequence that is at least 85% homologous to SEQ ID NO. 18.

67. The engineered split epimerase of clause 60, wherein said reductase component comprises a sequence that is at least 90% homologous to SEQ ID NO. 18.

68. The engineered split epimerase of clause 60, wherein said reductase component comprises a sequence that is at least 95% homologous to SEQ ID NO. 18.

69. The engineered split epimerase of clause 60, wherein said oxidase component comprises one or more mutations, each mutation at a codon position which corresponds to a position selected from the group consisting of positions 16, 17, 37, 50, 64, 69, 70, 110, 128, 155, 232, 234, 244, 293, 303, 328, 354, 387, 393, 427, and 477 within SEQ ID NO. 17.

70. The engineered split epimerase of clause 60, wherein said oxidase component comprises a sequence that is encoded by a DNA sequence which comprises one or more codon substitutions, relative to a wildtype sequence, at a codon position which corresponds to an amino acid position in SEQ ID NO. 17 selected from the group consisting of position 75, 160, 191, 256, 359, 366 and 500.

71. A method of producing an alkaloid product through the epimerization of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid via an engineered split epimerase in an engineered host cell said method comprising:

contacting said (*S*)-1-benzylisoquinoline alkaloid with said engineered split epimerase in said engineered host cell, thereby increasing production of an alkaloid product,

wherein contacting said (*S*)-1-benzylisoquinoline alkaloid with said engineered split epimerase converts said (*S*)-1-benzylisoquinoline alkaloid to said (*R*)-1-benzylisoquinoline alkaloid within said engineered host cell, wherein said engineered split epimerase comprises an oxidase component and a reductase component, wherein said oxidase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 17, and wherein said reductase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 18.

72. The method of clause 71, wherein said engineered split epimerase comprises one or more mutations that increase activity of said engineered split epimerase.

73. The method of clause 71, wherein said engineered split epimerase comprises one or more mutations selected from the group consisting of a localization mutation, a cytochrome P450 reductase interaction mutation, and an accessibility mutation.

74. The method of clause 71, wherein said oxidase component comprises a sequence that is encoded by a DNA sequence which comprises one or more mutations, each mutation at a codon position which corresponds to a position selected from the group consisting of positions 16, 17, 37, 50, 64, 69, 70, 110, 128, 155, 232, 234, 244, 293, 303, 328, 354, 387, 393, 427, and 477 within SEQ ID NO. 17.

75. The method of clause 71, wherein said oxidase component comprises a sequence that is at least 85% homologous to SEQ ID NO. 17.

76. The method of clause 71, wherein said oxidase component comprises a sequence that is at least 90% homologous to SEQ ID NO. 17.

77. The method of clause 71, wherein said oxidase component comprises a sequence that is at least 95% homologous to SEQ ID NO. 17.

78. The method of clause 71, wherein said oxidase component comprises a sequence that is encoded by a DNA sequence which comprises one or more codon substitutions, relative to a wildtype sequence, at a codon position which corresponds to an amino acid position in SEQ ID NO. 17 selected from the group consisting of position 75, 160, 191, 256, 359, 366, and 500.

79. The method of clause 71, wherein said reductase component comprises a sequence that is at least 85% homologous to SEQ ID NO. 18.

80. The method of clause 71, wherein said reductase component comprises a sequence that is at least 90% homologous to SEQ ID NO. 18.

81. The method of clause 71, wherein said reductase component comprises a sequence that is at least 95% homologous to SEQ ID NO. 18.

82. The method of any of the previous clauses, wherein said fused epimerase is a wildtype epimerase.

83. The method of any of the previous clauses, wherein said fused epimerase is a parent epimerase of said engineered split epimerase.

84. The method of any of the previous clauses, wherein said fused epimerase is an engineered fused epimerase comprising an oxidase domain with a same sequence as said oxidase component of said engineered split epimerase.

85. The method of any of the previous clauses, wherein said fused epimerase is an engineered fused epimerase comprising a reductase domain with a same sequence as said reductase component of said engineered split epimerase.

86. The engineered host cell of any of the previous clauses, wherein said fused epimerase is a wildtype epimerase.

87. The engineered host cell of any of the previous clauses, wherein said fused epimerase is a parent epimerase of said engineered split epimerase.

88. The engineered host cell of any of the previous clauses, wherein said fused epimerase is an engineered fused epimerase comprising an oxidase domain with a same sequence as said oxidase component of said engineered split epimerase.

89. The engineered host cell of any of the previous clauses, wherein said fused epimerase is an engineered fused epimerase comprising a reductase domain with a same sequence as said reductase component of said engineered split epimerase.

90. The method of any of the previous clauses, wherein the oxidase component has a cytochrome P450 oxidase-like domain.

91. The method of any of the previous clauses, wherein the reductase component has a codeinone reductase-like domain.

92. A method of producing an alkaloid product through the epimerization of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid via an engineered fused epimerase in an engineered host cell, said method comprising:

contacting said (*S*)-1-benzylisoquinoline alkaloid with said engineered fused epimerase in said engineered host cell, wherein said engineered fused epimerase comprises one or more mutations that increase activity of said engineered fused epimerase,

wherein contacting said (*S*)-1-benzylisoquinoline alkaloid with said engineered fused epimerase converts said (*S*)-1-benzylisoquinoline alkaloid to said (*R*)-1-benzylisoquinoline alkaloid within said engineered host cell.

93. The method of clause 92, wherein said engineered fused epimerase comprises one or more mutations selected from the group consisting of a localization mutation, a cytochrome P450 reductase interaction mutation, and an accessibility mutation.

94. The method of clause 92, wherein said engineered fused epimerase comprises a sequence that is encoded by a DNA sequence which comprises one or more mutations, relative to a wildtype sequence, each mutation at a codon position which corresponds to a position that is selected from the group consisting of positions 16, 17, 37, 50, 64, 69, 70, 110, 128, 155, 232, 234, 244, 293, 303, 328, 354, 387, 393, 427, and 477 within SEQ ID NO. 16.

95. The method of clause 92, wherein said engineered fused epimerase comprises a sequence that is encoded by a DNA sequence which comprises one or more codon substitutions, relative to a wildtype sequence, at a codon position which corresponds to an amino acid position in SEQ ID NO. 16 selected from the group consisting of position 75, 160, 191, 256, 359, 366, and 500.

96. The method of clause 92, wherein said engineered fused epimerase comprises a sequence that is at least 80% homologous to SEQ ID NO. 16.

97. The method of clause 92, wherein said engineered fused epimerase comprises a sequence that is at least 85% homologous to SEQ ID NO. 16.

98. The method of clause 92, wherein said engineered fused epimerase comprises a sequence that is at least 90% homologous to SEQ ID NO. 16.

99. The method of clause 92, wherein said engineered fused epimerase comprises a sequence that is at least 95% homologous to SEQ ID NO. 16.

100. A method of increasing production of an alkaloid product through the epimerization of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid via an engineered fused epimerase in an engineered host cell in comparison to the epimerization of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid via a wildtype fused epimerase in similar conditions, said method comprising:

contacting said (*S*)-1-benzylisoquinoline alkaloid with said engineered fused epimerase in said engineered host cell, thereby increasing production of an alkaloid product,

wherein contacting said (*S*)-1-benzylisoquinoline alkaloid with said engineered fused epimerase converts said (*S*)-1-benzylisoquinoline alkaloid to said (*R*)-1-benzylisoquinoline alkaloid within said engineered host cell, and wherein said engineered fused epimerase produces at least 1% more of said (*R*)-1-benzylisoquinoline alkaloid than a wildtype fused epimerase under similar conditions.

101. The method of clause 100, wherein said engineered fused epimerase produces at least 3% more of said (*R*)-1-benzylisoquinoline alkaloid than a wildtype fused epimerase under similar conditions.

102. The method of clause 100, wherein said engineered fused epimerase produces at least 5% more of said (*R*)-1-benzylisoquinoline alkaloid than a wildtype fused epimerase under similar conditions.

103. The method of clause 100, wherein said engineered fused epimerase produces at least 10% more of said (*R*)-1-benzylisoquinoline alkaloid than a wildtype fused epimerase under similar conditions.

104. The method of clause 100, wherein said engineered fused epimerase produces at least 15% more of said (*R*)-1-benzylisoquinoline alkaloid than a wildtype fused epimerase under similar conditions.

105. The method of clause 100, wherein said engineered fused epimerase produces at least 20% more of said (*R*)-1-benzylisoquinoline alkaloid than a wildtype fused epimerase under similar conditions.

106. The method of clause 100, wherein said engineered fused epimerase produces at least 25% more of said (*R*)-1-benzylisoquinoline alkaloid than a wildtype fused epimerase under similar conditions.

107. The method of clause 100, wherein said engineered fused epimerase produces at least 30% more of said (*R*)-1-benzylisoquinoline alkaloid than a wildtype fused epimerase under similar conditions.

108. The method of clause 100, wherein said engineered fused epimerase comprises one or more mutations that increase activity of said engineered fused epimerase.

109. The method of clause 100, wherein said engineered fused epimerase comprises one or more mutations selected from the group consisting of a localization mutation, a cytochrome P450 reductase interaction mutation, and an accessibility mutation.

110. The method of clause 100, wherein said engineered fused epimerase comprises a sequence that is encoded by a DNA sequence which comprises one or more mutations, relative to a wildtype sequence, each mutation at a codon position which corresponds to a position that is selected from the group consisting of positions 16, 17, 37, 50, 64, 69, 70, 110, 128, 155, 232, 234, 244, 293, 303, 328, 354, 387, 393, 427, and 477 within SEQ ID NO. 16.

111. The method of clause 100, wherein said engineered fused epimerase comprises a sequence that is encoded by a DNA sequence which comprises one or more codon substitutions, relative to a wildtype sequence, at a codon position which corresponds to an amino acid position in SEQ ID NO. 16 selected from the group consisting of position 75, 160, 191, 256, 359, 366, and 500.

112. The method of clause 100, wherein said engineered fused epimerase comprises a sequence that is at least 80% homologous to SEQ ID NO. 16.

113. The method of clause 100, wherein said engineered fused epimerase comprises a sequence that is at least 85% homologous to SEQ ID NO. 16.

114. The method of clause 100, wherein said engineered fused epimerase comprises a sequence that is at least 90% homologous to SEQ ID NO. 16.

115. The method of clause 100, wherein said engineered fused epimerase comprises a sequence that is at least 95% homologous to SEQ ID NO. 16.

116. An engineered host cell which comprises an engineered fused epimerase, wherein said epimerase increases conversion of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid, and wherein said engineered fused epimerase produces more of said (*R*)-1-benzylisoquinoline alkaloid than a wildtype fused epimerase under similar conditions.

117. The engineered host cell of clause 116, wherein said engineered fused epimerase comprises one or more mutations that increase activity of said engineered fused epimerase.

118. The engineered host cell of clause 116, wherein said engineered fused epimerase comprises one or more mutations selected from the group consisting of a localization mutation, a cytochrome P450 reductase interaction mutation, and an accessibility mutation.

119. The engineered host cell of clause 116, wherein said engineered fused epimerase comprises a sequence that is encoded by a DNA sequence which comprises one or more mutations, relative to a wildtype sequence, each mutation at a codon position which corresponds to a position that is selected from the group consisting of positions 16, 17, 37, 50, 64, 69, 70, 110, 128, 155, 232, 234, 244, 293, 303, 328, 354, 387, 393, 427, and 477 within SEQ ID NO. 16.

120. The engineered host cell of clause 116, wherein said engineered fused epimerase comprises a sequence that is at least 80% homologous to SEQ ID NO. 16 and is encoded by a DNA sequence which comprises one or more codon substitutions, relative to a wildtype sequence, at a codon

position which corresponds to an amino acid position in SEQ ID NO. 16 selected from the group consisting of position 75, 160, 191, 256, 359, 366, and 500.

121. An engineered fused epimerase that comprises one or more mutations selected from the group consisting of a localization mutation, a cytochrome P450 reductase interaction mutation, and an accessibility mutation.

122. The engineered epimerase of clause 121, wherein said engineered fused epimerase has increased activity compared to a wildtype fused epimerase.

123. The engineered fused epimerase of clause 121, wherein said engineered fused epimerase comprises one or more mutations that increase activity of said engineered fused epimerase.

124. The engineered fused epimerase of clause 121, wherein said engineered fused epimerase comprises a sequence that is encoded by a DNA sequence which comprises one or more mutations, relative to a wildtype sequence, each mutation at a codon position which corresponds to a position that is selected from the group consisting of positions 16, 17, 37, 50, 64, 69, 70, 110, 128, 155, 232, 234, 244, 293, 303, 328, 354, 387, 393, 427, and 477 within SEQ ID NO. 16.

125. The engineered fused epimerase of clause 121, wherein said engineered fused epimerase comprises a sequence that is at least 80% homologous to SEQ ID NO. 16 and is encoded by a DNA sequence which comprises one or more codon substitutions, relative to a wildtype sequence, at a codon position which is equivalent to an amino acid position in SEQ ID NO. 16 selected from the group consisting of position 75, 160, 191, 256, 359, 366, and 500.

126. A method for forming a product stream having a benzyloquinoline alkaloid product, comprising:

(a) providing engineered host cells and a feedstock including nutrients and water to a batch reactor, which engineered host cells comprise at least one heterologous coding sequence encoding at least one engineered epimerase, wherein the engineered epimerase is an engineered split epimerase or engineered fused epimerase from any of the previous clauses;

(b) in said batch reactor, subjecting said engineered host cells to fermentation by incubating said engineered host cells for a time period of at least about 5 minutes to produce a solution comprising said benzyloquinoline alkaloid product and cellular material; and

(c) using at least one separation unit to separate said benzyloquinoline alkaloid product from said cellular material to provide said product stream comprising said benzyloquinoline alkaloid product.

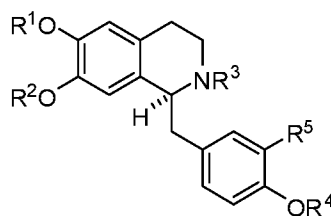
127. The method of clause 126, wherein at least one process parameter of the batch reactor is modifiable to alter a resultant benzyloquinoline alkaloid product composition.

128. The method of clause 126, wherein the at least one process parameter that is modifiable comprises at least one of dissolved oxygen, pH, stirring speed, aeration rate, and cell density.

129. A medicinal product that comprises the benzyloquinoline alkaloid product produced from any of the previous clauses.

130. The medicinal product of clause 129, wherein the medicinal product has cancer therapeutic properties.
131. The medicinal product of clause 129, wherein the medicinal product has analgesic properties.
132. The medicinal product of clause 129, wherein the medicinal product has antimalarial properties.
133. The medicinal product of clause 129, wherein the medicinal product has antitussive properties.
134. The medicinal product of clause 129, wherein the medicinal product has opioid overdose treatment properties.
135. The medicinal product of clause 129, wherein the medicinal product has anti-addiction properties.
136. The method of any of the previous clauses, wherein said product is further processed within said engineered host cell to produce a nal-opioid product.
137. The method of any of the previous clauses, wherein said product is further processed within said engineered host cell to produce a nor-opioid product.
138. The method of any of the previous clauses, wherein said product is further processed within said engineered host cell to produce a bisbenzylisoquinoline alkaloid product.
139. The engineered host cell of any of the previous clauses, wherein said engineered host cell is engineered to convert said product to a nal-opioid product.
140. The engineered host cell of any of the previous clauses, wherein said engineered host cell is engineered to convert said product to a nor-opioid product.
141. The engineered host cell of any of the previous clauses, wherein said engineered host cell is engineered to convert said product to a bisbenzylisoquinoline alkaloid product.
142. The engineered host cell of any of the previous clauses, wherein said engineered host cell is engineered to convert said product to a promorphinan alkaloid product.
143. The engineered host cell of any of the previous clauses, wherein said engineered host cell is engineered to convert said product to a morphinan alkaloid product.
144. The method of any of the previous clauses, wherein the engineered fused epimerase comprises an oxidase domain and a reductase domain.
145. The method of clause 144, wherein the oxidase domain is a cytochrome P450 oxidase-like domain.
146. The method of clause 144, wherein the reductase domain is a codeinone reductase-like domain.
147. The method of any of the previous clauses, wherein the (*S*)-1-benzylisoquinoline alkaloid is (*S*)-reticuline.

148. The method of any of the previous clauses, wherein the (*R*)-1-benzylisoquinoline alkaloid is (*R*)-reticuline.
149. The method of any of the previous clauses, further comprising:
converting the (*R*)-1-benzylisoquinoline alkaloid to a 4-ring promorphinan alkaloid.
150. The method of clause 75, wherein the 4-ring promorphinan alkaloid is salutaridine.
151. The method of any of the previous clauses, further comprising:
converting the (*R*)-1-benzylisoquinoline alkaloid to a 5-ring morphinan alkaloid.
152. The method of any of the previous clauses, further comprising:
converting the (*R*)-1-benzylisoquinoline alkaloid to a bisbenzylisoquinoline alkaloid.
153. The method of any of the previous clauses, further comprising:
converting the (*R*)-1-benzylisoquinoline alkaloid to a nal-opioid.
154. The method of any of the previous clauses, further comprising:
converting the (*R*)-1-benzylisoquinoline alkaloid to a nor-opioid.
155. The method of any of the previous clauses, further comprising:
converting the (*R*)-1-benzylisoquinoline alkaloid to a promorphinan.
156. The method of any of the previous clauses, further comprising:
converting the (*R*)-1-benzylisoquinoline alkaloid to a morphinan.
157. The method of any of the previous clauses, wherein the (*S*)-1-benzylisoquinoline alkaloid present within the engineered host cell is produced within an engineered microbial cell.
158. The method of any of the previous clauses, wherein the (*S*)-1-benzylisoquinoline alkaloid is produced within the engineered host cell by a metabolic pathway starting with L-tyrosine.
159. The method of any of the previous clauses, wherein the (*S*)-1-benzylisoquinoline alkaloid is produced within the engineered host cell by a metabolic pathway starting with a carbohydrate and nitrogen source.
160. The method of any of the previous clauses, wherein the (*S*)-1-benzylisoquinoline alkaloid is selected from the group consisting of: (*S*)-norreticuline; (*S*)-reticuline; (*S*)-tetrahydropapaverine; (*S*)-norcocclaurine; (*S*)-cocclaurine; (*S*)-*N*-methylococclaurine; (*S*)-3'-hydroxy-*N*-methylococclaurine; (*S*)-norisoorientaline; (*S*)-orientaline; (*S*)-isoorientaline; (*S*)-norprotosinomenine; (*S*)-protosinomenine; (*S*)-norlaudanoline; (*S*)-laudanoline; (*S*)-4'-*O*-methyllaudanoline; (*S*)-6-*O*-methylnorlaudanoline; and (*S*)-4'-*O*-methylnorlaudanoline.
161. The method of any one of the previous clauses wherein the (*S*)-1-benzylisoquinoline alkaloid is a compound of Formula I:



Formula I,

wherein:

R^1 , R^2 , R^3 , and R^4 are independently hydrogen or methyl; and

R^5 is hydrogen, hydroxy, or methoxy.

162. The method of any one of the above clauses, wherein said engineered host cell is an engineered non-plant cell.

163. The method of any one of the above clauses, wherein said engineered host cell is an engineered yeast cell.

164. The method of any one of the above clauses, wherein said engineered host cell is an engineered bacteria cell.

165. The engineered host cell of any of the previous clauses, wherein the engineered fused epimerase comprises an oxidase domain and a reductase domain.

166. The engineered host cell of clause 165, wherein the oxidase component has a cytochrome P450 oxidase-like domain.

167. The engineered host cell of clause 165, wherein the reductase component has a codeinone reductase-like domain.

168. The engineered host cell of any one of the above clauses, wherein said engineered host cell is an engineered non-plant cell.

169. The engineered host cell of any one of the above clauses, wherein said engineered host cell is an engineered yeast cell.

170. The engineered host cell of any one of the above clauses, wherein said engineered host cell is an engineered bacteria cell.

171. A method of increasing production of an alkaloid product through the epimerization of an (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid via an engineered epimerase in an engineered host cell, said method comprising:

contacting said (*S*)-1-benzylisoquinoline alkaloid with said engineered epimerase,

wherein contacting said (*S*)-1-benzylisoquinoline alkaloid with said engineered epimerase converts said (*S*)-1-benzylisoquinoline alkaloid to said (*R*)-1-benzylisoquinoline alkaloid.

172. The method of clause 171, wherein said engineered epimerase is an engineered split epimerase.

173. The method of clause 172, wherein said engineered split epimerase increases said production of said alkaloid product in comparison to an epimerization via a wild-type fused epimerase.

174. The method of clause 172, wherein said engineered split epimerase comprises one or more mutations that increase activity of said engineered split epimerase.

175. The method of clause 172, wherein said engineered split epimerase comprises one or more mutations selected from the group consisting of a localization mutation, a cytochrome P450 reductase interaction mutation, and an accessibility mutation.

176. The method of clause 172, wherein said engineered split epimerase comprises one or more mutations, each mutation at a position which is equivalent to a position within SEQ ID NO. 16 selected from the group consisting of positions 16, 17, 37, 50, 64, 69, 70, 110, 128, 155, 232, 234, 244, 293, 303, 328, 354, 387, 393, 427, and 477 within SEQ ID NO. 16.

177. The method of clause 172, wherein said engineered split epimerase comprises a sequence that is at least 80% homologous to SEQ ID NO. 16 and is encoded by a DNA sequence which comprises one or more codon substitutions, relative to a wildtype sequence, at a codon position which is equivalent to an amino acid position in SEQ ID NO. 16 selected from the group consisting of position 75, 160, 191, 256, 359, 366 and 500.

178. The method of clause 171, wherein said engineered epimerase is an engineered fused epimerase.

179. The method of clause 178, wherein said engineered fused epimerase comprises one or more mutations that increase activity of said engineered fused epimerase.

180. The method of clause 178, wherein said engineered fused epimerase comprises one or more mutations selected from the group consisting of a localization mutation, a cytochrome P450 reductase interaction mutation, and an accessibility mutation.

181. The method of clause 178, wherein said engineered fused epimerase comprises one or more mutations, each mutation at a position which is equivalent to a position within SEQ ID NO. 16 selected from the group consisting of positions 16, 17, 37, 50, 64, 69, 70, 110, 128, 155, 232, 234, 244, 293, 303, 328, 354, 387, 393, 427, and 477 within SEQ ID NO. 16.

182. The method of clause 178, wherein said engineered fused epimerase comprises a sequence that is at least 80% homologous to SEQ ID NO. 16 and is encoded by a DNA sequence which comprises one or more codon substitutions, relative to a wildtype sequence, at a codon position which is equivalent to an amino acid position in SEQ ID NO. 16 selected from the group consisting of position 75, 160, 191, 256, 359, 366 and 500.

183. The method of any one of the above clauses 171-182, wherein said engineered host cell is an engineered non-plant cell.

184. An engineered host cell which comprises an engineered epimerase, wherein said epimerase increases conversion of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid.

185. The engineered host cell of clause 184, wherein said engineered epimerase is an engineered split epimerase.

186. The engineered host cell of clause 185, wherein said engineered split epimerase increases production of said alkaloid product in comparison to an epimerization via a wild-type fused epimerase.

187. The engineered host cell of clause 185, wherein said engineered split epimerase comprises one or more mutations that increase activity of said engineered split epimerase.

188. The engineered host cell of clause 185, wherein said engineered split epimerase comprises one or more mutations selected from the group consisting of a localization mutation, a cytochrome P450 reductase interaction mutation, and an accessibility mutation.

189. The engineered host cell of clause 185, wherein said engineered split epimerase comprises one or more mutations, each mutation at a position which is equivalent to a position within SEQ ID NO. 16 selected from the group consisting of positions 16, 17, 37, 50, 64, 69, 70, 110, 128, 155, 232, 234, 244, 293, 303, 328, 354, 387, 393, 427, and 477 within SEQ ID NO. 16.

190. The engineered host cell of clause 185, wherein said engineered split epimerase comprises a sequence that is at least 80% homologous to SEQ ID NO. 16 and is encoded by a DNA sequence which comprises one or more codon substitutions, relative to a wildtype sequence, at a codon position which is equivalent to an amino acid position in SEQ ID NO. 16 selected from the group consisting of position 75, 160, 191, 256, 359, 366 and 500.

191. The engineered host cell of clause 184, wherein said engineered epimerase is an engineered fused epimerase.

192. The engineered host cell of clause 191, wherein said engineered fused epimerase comprises one or more mutations that increase activity of said engineered fused epimerase.

193. The engineered host cell of clause 191, wherein said engineered split epimerase comprises one or more mutations selected from the group consisting of a localization mutation, a cytochrome P450 reductase interaction mutation, and an accessibility mutation.

194. The engineered host cell of clause 191, wherein said engineered fused epimerase comprises one or more mutations, each mutation at a position which is equivalent to a position within SEQ ID NO. 16 selected from the group consisting of positions 16, 17, 37, 50, 64, 69, 70, 110, 128, 155, 232, 234, 244, 293, 303, 328, 354, 387, 393, 427, and 477 within SEQ ID NO. 16.

195. The engineered host cell of clause 191, wherein said engineered fused epimerase comprises a sequence that is at least 80% homologous to SEQ ID NO. 16 and is encoded by a DNA sequence which comprises one or more codon substitutions, relative to a wildtype sequence, at a codon position which is equivalent to an amino acid position in SEQ ID NO. 16 selected from the group consisting of position 75, 160, 191, 256, 359, 366 and 500.

196. The engineered host cell of any one of clauses 184-195, wherein said engineered host cell is an engineered non-plant cell.

197. An engineered epimerase that comprises one or more mutations selected from the group consisting of a localization mutation, a cytochrome P450 reductase interaction mutation, and an accessibility mutation.

198. The engineered epimerase of clause 197, wherein said engineered epimerase is an engineered split epimerase.

199. The engineered epimerase of clause 197, wherein said engineered split epimerase has increased activity compared to a wildtype fused epimerase.

200. The engineered epimerase of clause 197, wherein said engineered split epimerase comprises one or more mutations that increase activity of said engineered split epimerase.

201. The engineered epimerase of clause 197, wherein said engineered split epimerase comprises one or more mutations, each mutation at a position which is equivalent to a position within SEQ ID NO. 16 selected from the group consisting of positions 16, 17, 37, 50, 64, 69, 70, 110, 128, 155, 232, 234, 244, 293, 303, 328, 354, 387, 393, 427, and 477 within SEQ ID NO. 16.

202. The engineered epimerase of clause 197, wherein said engineered split epimerase comprises a sequence that is at least 80% homologous to SEQ ID NO. 16 and is encoded by a DNA sequence which comprises one or more codon substitutions, relative to a wildtype sequence, at a codon position which is equivalent to an amino acid position in SEQ ID NO. 16 selected from the group consisting of position 75, 160, 191, 256, 359, 366 and 500.

203. An engineered epimerase, wherein said engineered epimerase is an engineered fused epimerase that comprises one or more mutations selected from the group consisting of a localization mutation, a cytochrome P450 reductase interaction mutation, and an accessibility mutation.

204. The engineered epimerase of clause 203, wherein said engineered fused epimerase comprises one or more mutations at a position corresponding to SEQ ID NO. 16, said one or more mutations corresponding to a position selected from the group consisting of position 16, 17, 37, 50, 64, 69, 70, 110, 128, 155, 232, 234, 244, 293, 303, 328, 354, 387, 393, 427, and 477.

205. The engineered epimerase of clause 203, wherein said engineered fused epimerase comprises a sequence that is at least 80% homologous to SEQ ID NO. 16 and is encoded by a DNA sequence which comprises one or more codon substitutions, relative to a wildtype sequence, at a codon position which is equivalent to an amino acid position in SEQ ID NO. 16 selected from the group consisting of position 75, 160, 191, 256, 359, 366 and 500.

[00411] While preferred embodiments of the invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the invention. It should be understood that various alternatives to the embodiments of the invention described herein may be employed in practicing the invention. It is intended that the following claims define the scope of the invention and that methods and structures within the scope of these claims and their equivalents be covered thereby.

CLAIMS

WHAT IS CLAIMED IS:

1. A method of increasing production of an alkaloid product through the epimerization of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid via an engineered split epimerase in an engineered host cell in comparison to the epimerization of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid via a fused epimerase in similar conditions, said method comprising:
contacting said (*S*)-1-benzylisoquinoline alkaloid with said engineered split epimerase in said engineered host cell, thereby increasing production of an alkaloid product,
wherein contacting said (*S*)-1-benzylisoquinoline alkaloid with said engineered split epimerase converts said (*S*)-1-benzylisoquinoline alkaloid to said (*R*)-1-benzylisoquinoline alkaloid within said engineered host cell.
2. The method of claim 1, wherein said fused epimerase is a wildtype epimerase.
3. The method of claim 1, wherein said fused epimerase is a parent epimerase of said engineered split epimerase.
4. The method of claim 1, wherein said fused epimerase is an engineered fused epimerase comprising an oxidase domain with a same sequence as said oxidase component of said engineered split epimerase.
5. The method of claim 1, wherein said engineered split epimerase comprises one or more mutations that increase activity of said engineered split epimerase.
6. The method of claim 1, wherein said engineered split epimerase comprises one or more mutations selected from the group consisting of a localization mutation, a cytochrome P450 reductase interaction mutation, and an accessibility mutation.
7. The method of claim 1, wherein said engineered split epimerase comprises an oxidase component and a reductase component.
8. The method of claim 7, wherein said oxidase component comprises a sequence that is encoded by a DNA sequence which comprises one or more mutations, each mutation at a codon position which corresponds to a position selected from the group consisting of positions 16, 17, 37, 50, 64, 69, 70, 110, 128, 155, 232, 234, 244, 293, 303, 328, 354, 387, 393, 427, and 477 within SEQ ID NO. 17.
9. The method of claim 7, wherein said oxidase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 17.
10. The method of claim 7, wherein said oxidase component comprises a sequence that is at least 95% homologous to SEQ ID NO. 17.
11. The method of claim 7, wherein said oxidase component comprises a sequence that is encoded by a DNA sequence which comprises one or more codon substitutions, relative to a wildtype sequence, at a codon position which corresponds to an amino acid position in SEQ ID NO. 17 selected from the group consisting of position 75, 160, 191, 256, 359, 366, and 500.

12. The method of claim 7, wherein said reductase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 18.

13. The method of claim 7, wherein said reductase component comprises a sequence that is at least 95% homologous to SEQ ID NO. 18.

14. The method of claim 1, wherein said engineered split epimerase produces at least 1% more of said (*R*)-1-benzylisoquinoline alkaloid than said fused epimerase under similar conditions.

15. The method of claim 1, wherein said engineered split epimerase produces at least 3% more of said (*R*)-1-benzylisoquinoline alkaloid than a wildtype fused epimerase under similar conditions.

16. The method of claim 1, wherein said engineered split epimerase produces at least 5% more of said (*R*)-1-benzylisoquinoline alkaloid than a wildtype fused epimerase under similar conditions.

17. A method of producing an alkaloid product through the epimerization of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid via an engineered split epimerase in an engineered host cell said method comprising:

contacting said (*S*)-1-benzylisoquinoline alkaloid with said engineered split epimerase in said engineered host cell, thereby increasing production of an alkaloid product,

wherein contacting said (*S*)-1-benzylisoquinoline alkaloid with said engineered split epimerase converts said (*S*)-1-benzylisoquinoline alkaloid to said (*R*)-1-benzylisoquinoline alkaloid within said engineered host cell, wherein said engineered split epimerase comprises an oxidase component and a reductase component, wherein said oxidase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 17, and wherein said reductase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 18.

18. The method of claim 17, wherein said engineered split epimerase comprises one or more mutations that increase activity of said engineered split epimerase.

19. The method of claim 17, wherein said engineered split epimerase comprises one or more mutations selected from the group consisting of a localization mutation, a cytochrome P450 reductase interaction mutation, and an accessibility mutation.

20. The method of claim 17, wherein said engineered split epimerase comprises an oxidase component and a reductase component.

21. The method of claim 20, wherein said oxidase component comprises a sequence that is encoded by a DNA sequence which comprises one or more mutations, each mutation at a codon position which corresponds to a position selected from the group consisting of positions 16, 17, 37, 50, 64, 69, 70, 110, 128, 155, 232, 234, 244, 293, 303, 328, 354, 387, 393, 427, and 477 within SEQ ID NO. 17.

22. The method of claim 20, wherein said oxidase component comprises a sequence that is at least 85% homologous to SEQ ID NO. 17.

23. The method of claim 20, wherein said oxidase component comprises a sequence that is at least 90% homologous to SEQ ID NO. 17.

24. The method of claim 20, wherein said oxidase component comprises a sequence that is at least 95% homologous to SEQ ID NO. 17.

25. The method of claim 20, wherein said oxidase component comprises a sequence that is encoded by a DNA sequence which comprises one or more codon substitutions, relative to a wildtype sequence, at a codon position which corresponds to an amino acid position in SEQ ID NO. 17 selected from the group consisting of position 75, 160, 191, 256, 359, 366, and 500.

26. The method of claim 20, wherein said reductase component comprises a sequence that is at least 85% homologous to SEQ ID NO. 18.

27. The method of claim 20, wherein said reductase component comprises a sequence that is at least 90% homologous to SEQ ID NO. 18.

28. The method of claim 20, wherein said reductase component comprises a sequence that is at least 95% homologous to SEQ ID NO. 18.

29. The method of claim 7 or 20, wherein a first nucleic acid which encodes said oxidase component has a first translation initiation site and wherein a second nucleic acid which encodes said reductase component has a second translation initiation site.

30. The method of claim 29, wherein said first nucleic acid is separate from said second nucleic acid.

31. The method of claim 29, wherein said first nucleic acid is a same nucleic acid as said second nucleic acid.

32. The method of claim 7 or 20, wherein said oxidase component is encoded by a first messenger RNA and wherein said reductase component is encoded by a second messenger RNA.

33. The method of claim 7 or 20, wherein said oxidase component comprises a first molecule, and wherein said reductase component comprises a second molecule, wherein said first molecule is separate from said second molecule.

34. An engineered host cell which comprises an engineered split epimerase, wherein said engineered split epimerase increases conversion of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid relative to conversion of a (*S*)-1-benzylisoquinoline alkaloid to a (*R*)-1-benzylisoquinoline alkaloid via a fused epimerase under similar conditions.

35. The engineered host cell of claim 34, wherein said fused epimerase is a wildtype epimerase.

36. The engineered host cell of claim 34, wherein said fused epimerase is a parent epimerase of said engineered split epimerase.

37. The engineered host cell of claim 34, wherein said fused epimerase is an engineered fused epimerase comprising an oxidase domain with a same sequence as said oxidase component of said engineered split epimerase.

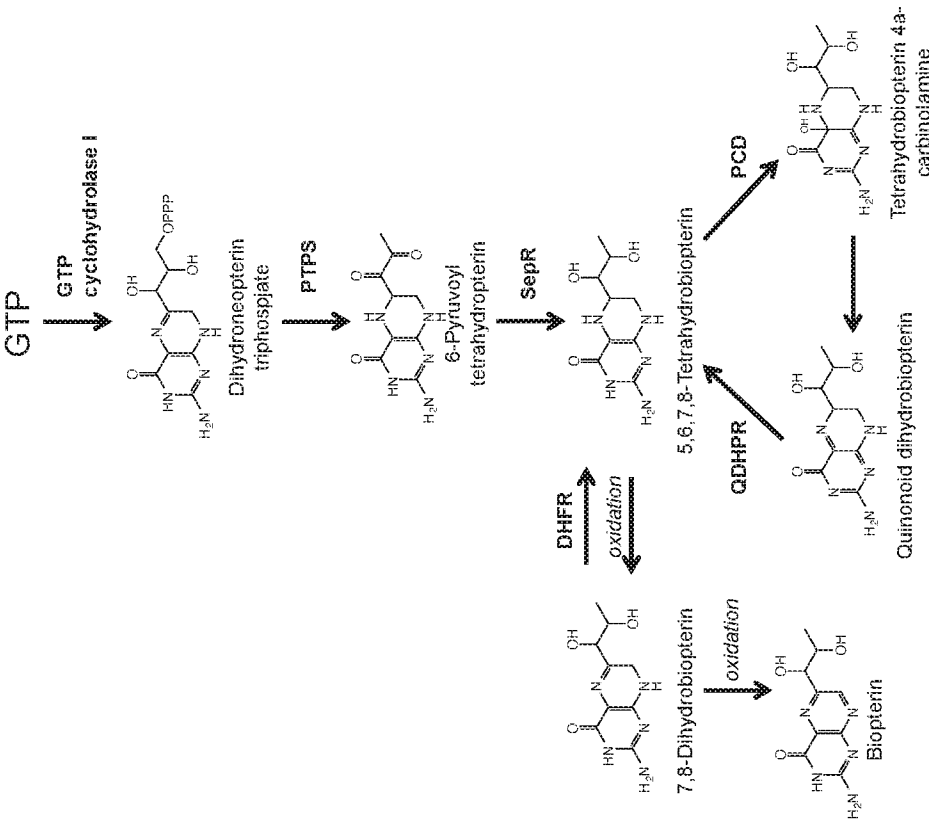


FIG. 1

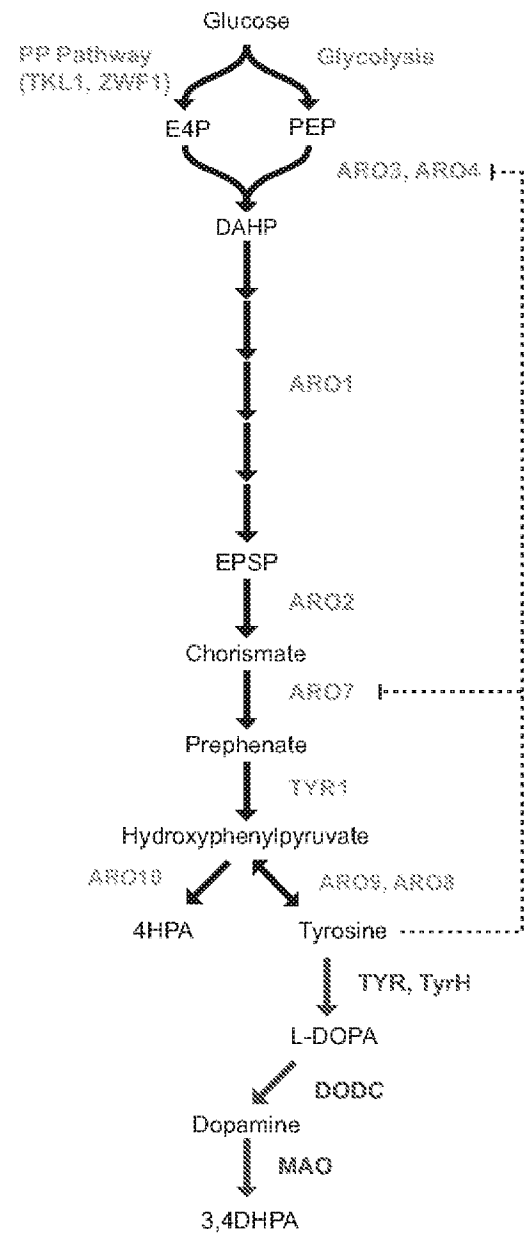


FIG. 2

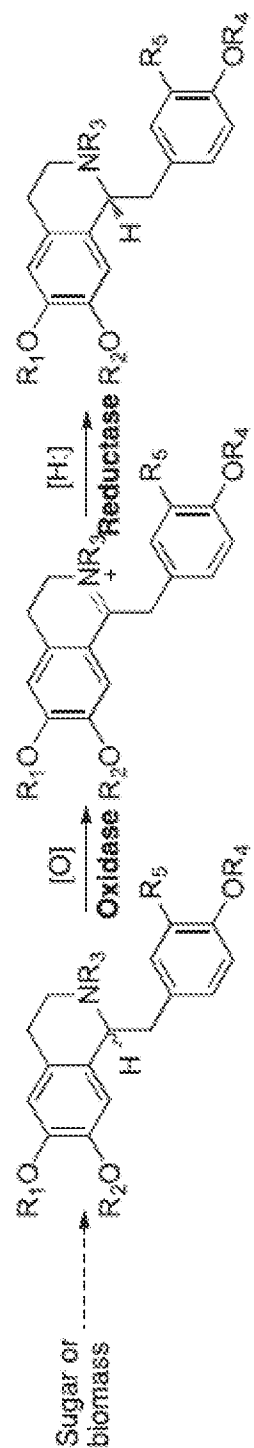


FIG. 3

4/30

SEQ ID No. 16

MELQYISYFQPTSSVALLALVSILSSVVLRKTFLNYYSSPASSTKTAVLSHQRQQSCALPISGLLH
IFMKNKGLIHVTLGNMADKYGPIFSPTGSHRTLVSSSWEMVKECFTGNNDTAFSNRPIPLAFKTIYAC
GGIDSYGLSSVPYGYWRELKVCVHNLLSNQQLKFRHLIISQVDTSFNKLVELCKNSEDNQGNYPPTT
TAAGMVRIDDWLAELSFNIGRIVCGFQSGPKTGAPSRVEQFKEAINEASYFMSTSPVSDNVPMLGWIDQ
LTGLTRNMKHCGKKLDLVESIINDHRQKRFRSRTKGGDEKDDFIDICLSIMEQPQLPGNNNPSQI
PIKSIVLDMIGGGTDTTKLTITWTLSSLNPHVLDKAKQEVDAHFRTKRRSTNDAAAADVDFDDIRNLV
YIQAIKESMRLYPASPVVERLSGEDCVVGGFHVPAGTRLWANVWKMQRDPKVWDDPLVFRPDRFLSDEQ
KMVDVVRGQNYELLPGAGRRVCPGVSFSLDLMQLVLTRLILEFEMKSPSGKVDMTATPGLMSYKVIPLDI
LLTHRRIKPCVQSAASERDMESSGPVITLGSQKMPVPLGMGTFEKVKGSERLAILKAIEVGYRYFD
TAAAYETEEVLGEAIAEALQLGLVKSRLDELFSSMLWCTDAHADRVLLALQNSLRNLKLEYVDLYMLPFP
ASLKP GKITMDIPEEDICRMDYRSVWAAAMEECQNLGFTKSGVSNFSCCKLQELMATANIPPAVNQVEMS
PAFQQKKLREYCNANNILVSAISVLGSNGTPWGSNAVIGSEVLKKIAMAKGKSAQVSMRWVVEQGASLV
VKSFSEERLRENINIFDWELTKEDHEKIGEIPQCRILSAYFLVSPNGPFKSQEELWDDEA

FIG. 4

Split Enzymes originating from SEQ ID No. 16

DRS enzyme (SEQ ID No. 17):

MELQFSYFQPTSSVALLALVSILFSVVLKTFSNYSSPASSTETAVLCHQRQQSCALPISGLLH
VFMNKNGLIHVTLGNMADKYGPFSFPTGSHRTLVSSEWEMVKECFTGNNDTAFSNRPIPLAFQTI
FYACGGIDSYGLSSVPYGYWRELKVCVHNLLSNQQLKFRHLISQVDTSFNKLVELCKNSEDNQ
GMVRMDDWLALQSFNVIGRIVCGFQSDPKTGAPSRVEQKEVINEASYFMSTSPVSDNVPMLG
WIDQLTGLTRNMKHCCKKLDLVESIHDRQKRRFSRTKGGDEKDEQDDFIDICLSIMEQPQLP
GNNSPPQIPIKSIVLDMIGGGTDTTKLTTIWTLSLLNPHVLDKAKQEVDAHFRKKRRSTDAAA
AVVDFDIRNLVYIQAIKESMRLYPASPVVERLSGEDCVGGFHVPAGTRLWANVWKMQRDPKV
WDDPLVFRPERFLSDEQKMDVVRGQNYELLPGAGRRICPGVSFSLDLMQLVLTRLILEFEMKSPS
GKVDMTATPGLMSYKVPLDILLTHRRIKSCVQLASSERD *

5/30

DRR enzyme (SEQ ID No. 18):

MESSGVPVITLSSGKVMPLVGMGTFEKVKGSERLAILKAIEVGYRYFDTAAAYETEEVLGEAIAE
ALQLGLIESRDELFISSMLWCTDAHPPDRVLLALQNSLRNLKLEYLDLYMLPFPASLKPGKITMDIPEE
DICRMDYRSVWSAMEECQNLGFTKISIGVSNFSCCKLQELMATANIPPAVNQVEMSPAFOQKKLRE
YCNANNILVSAVSILGSNGTPWGSNAVLGSEVLKQIAMAKGKSAQVSMRWVYEQQGASLVVKSFS
EERLRENLFIDWELTKEDNEKIGEIPQCRILTAYFLVSPNGPFKSQEELWDDKA*

FIG. 5

3-O-demethylase (ODM):

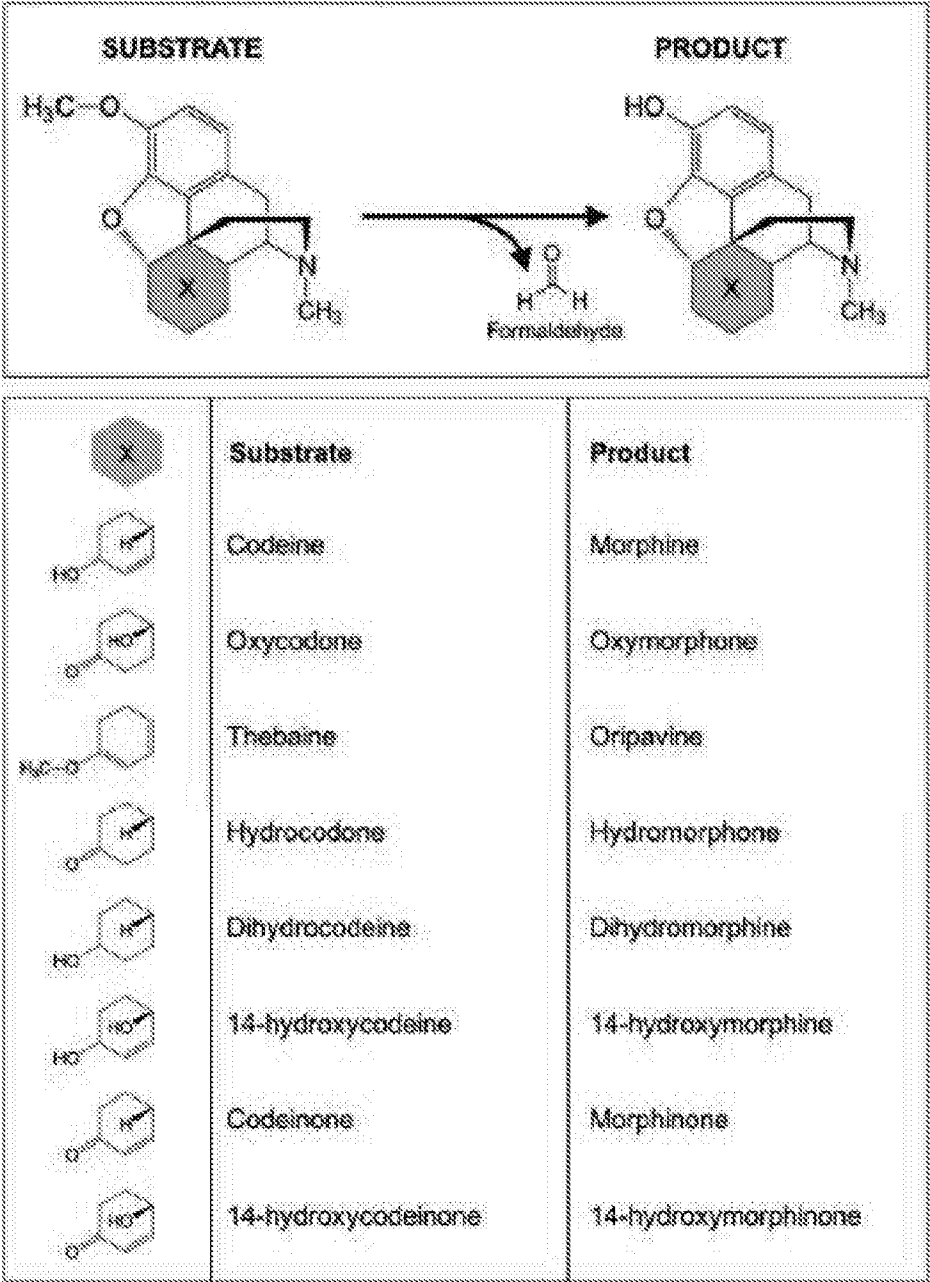


FIG. 6

N-demethylase (NDM):

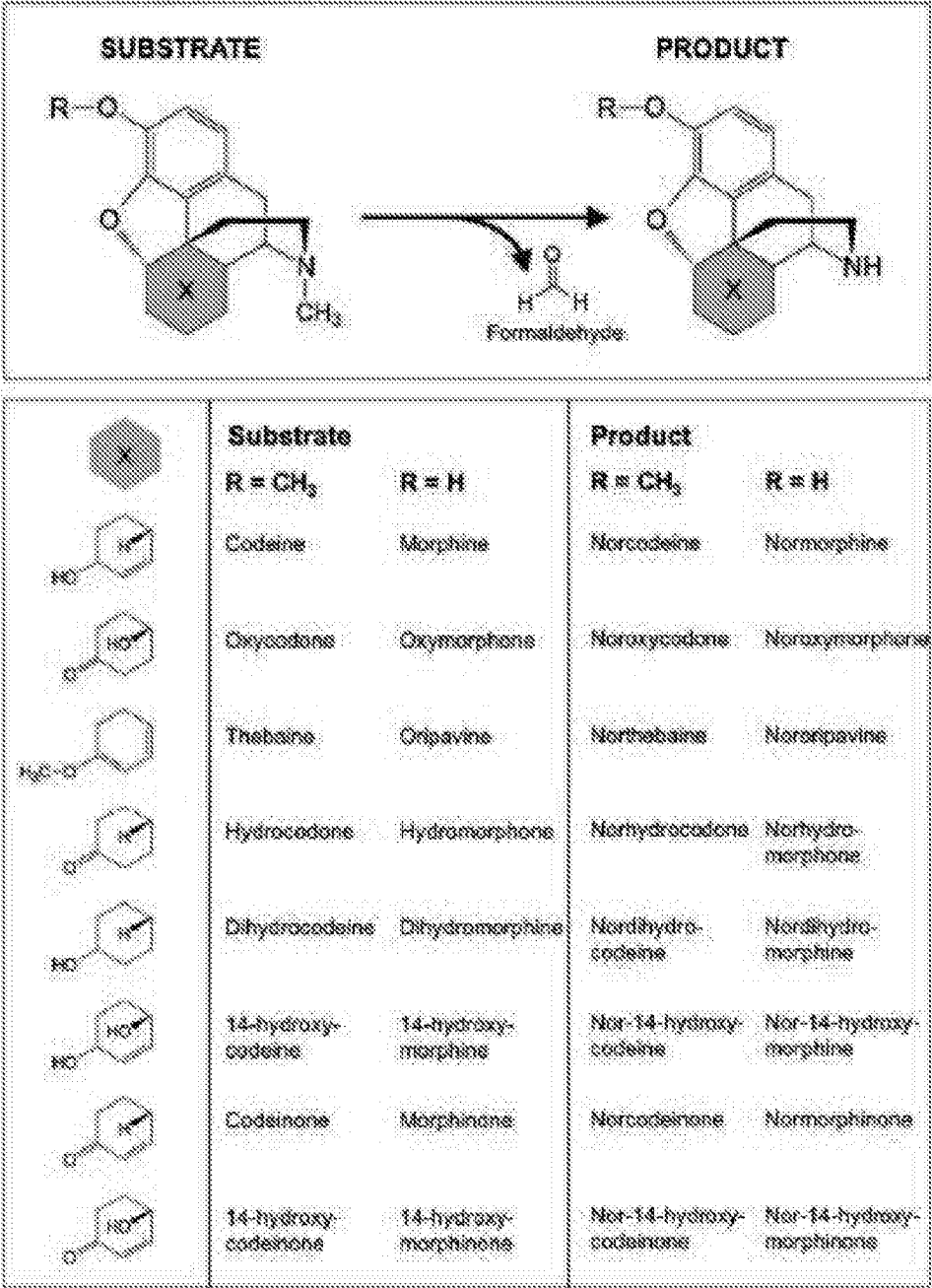


FIG. 7

8/30

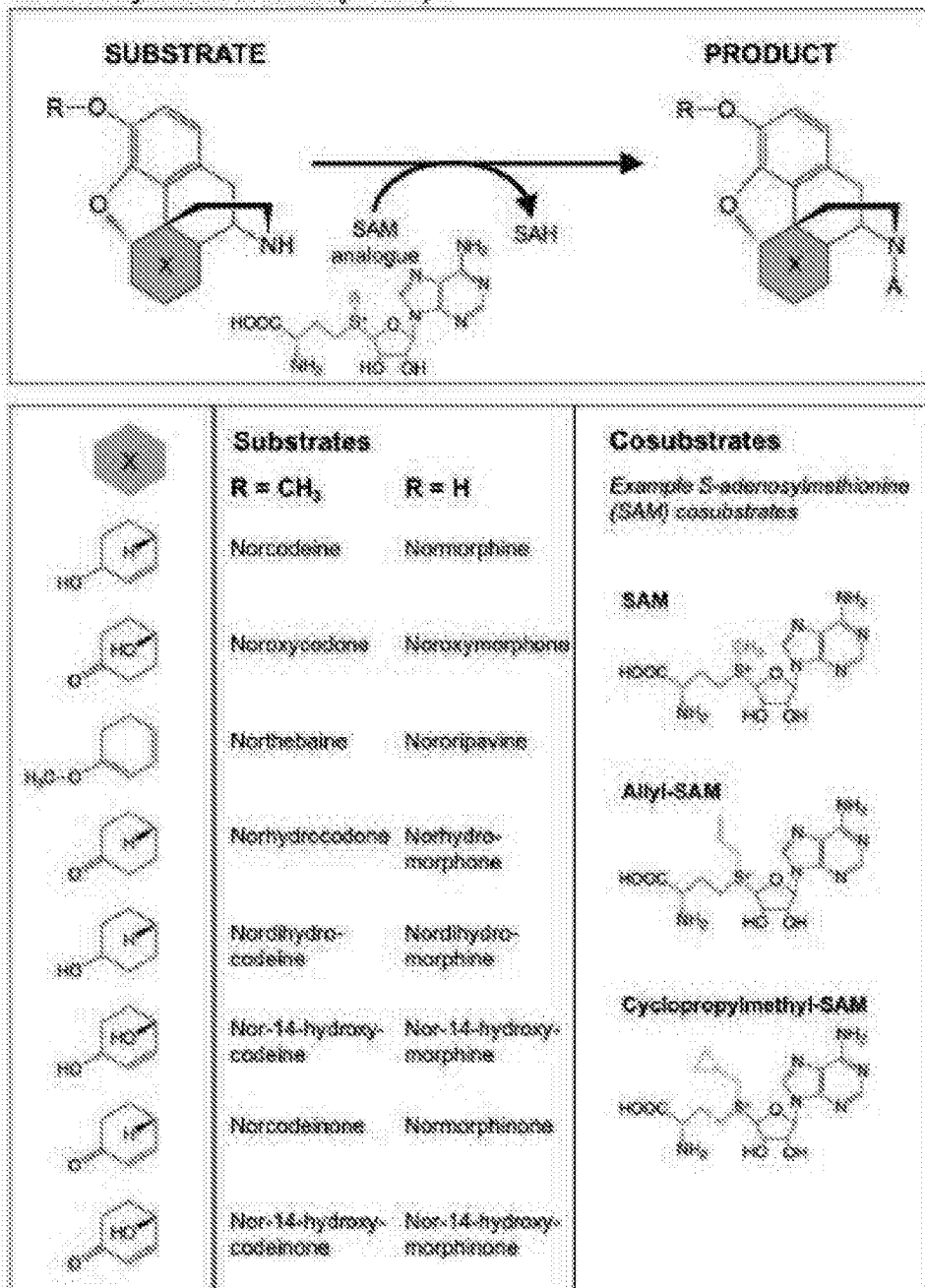
N-methyltransferase (NMT):

FIG. 8

9/30

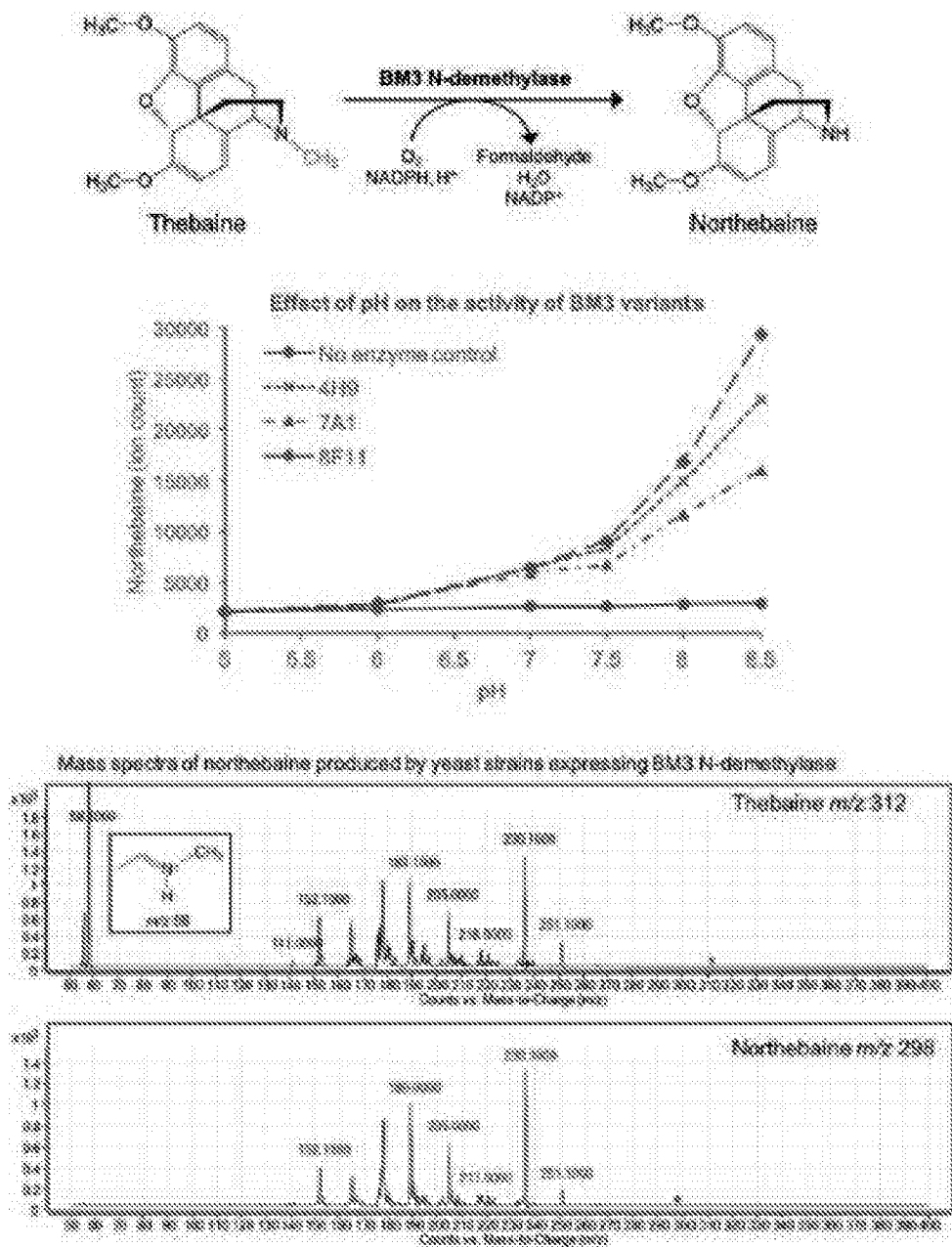


FIG. 9

10/30

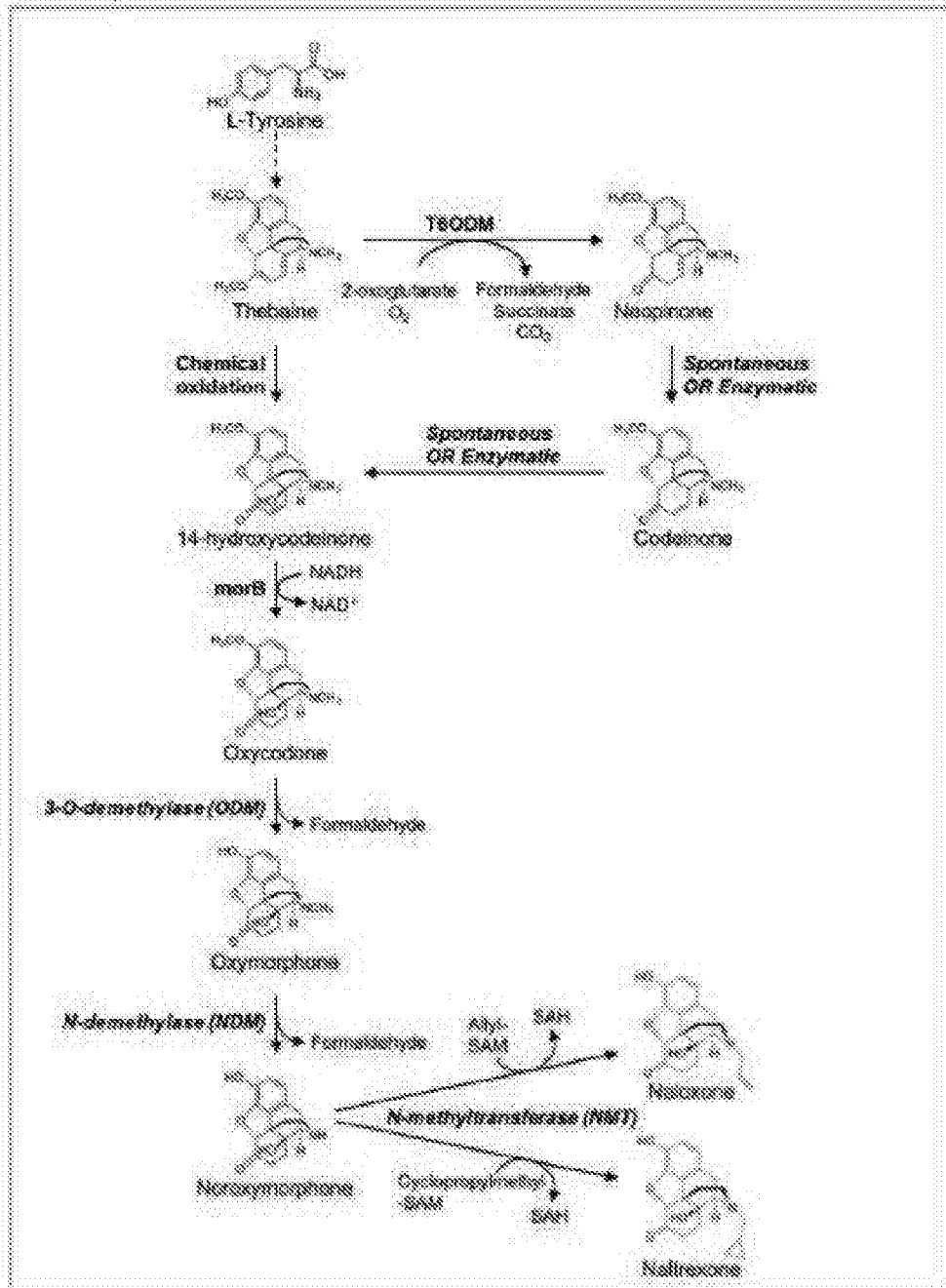
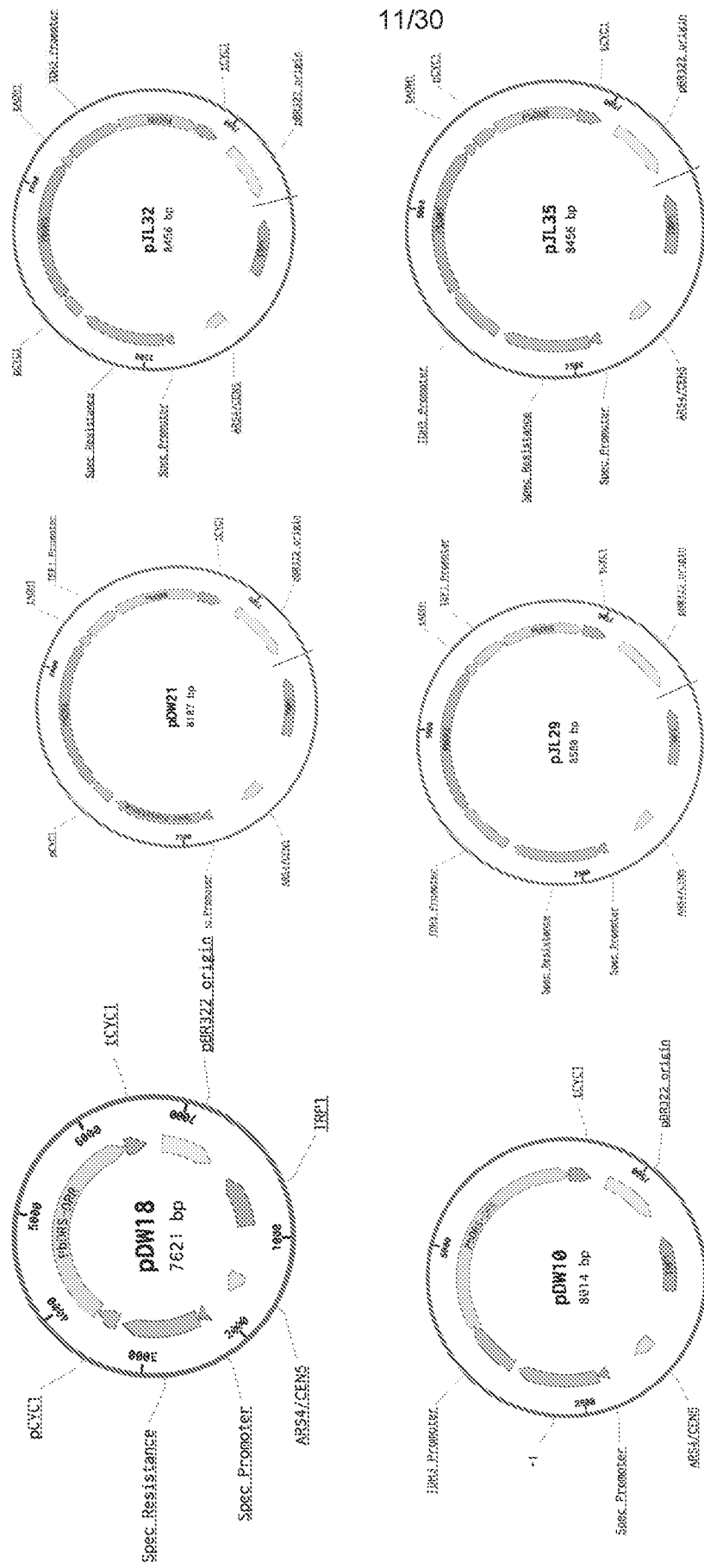
Example reaction scheme:

FIG. 10



11/30

Expression plasmids

FIG. 11

12/30

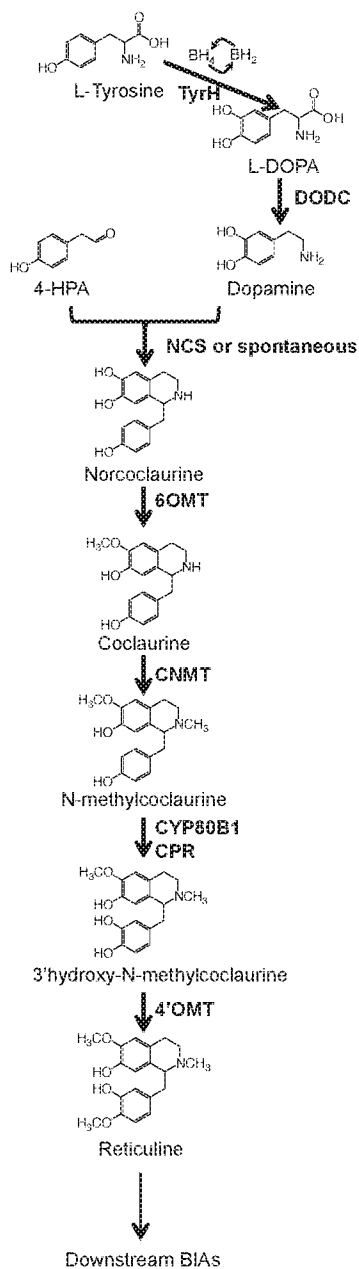


FIG. 12

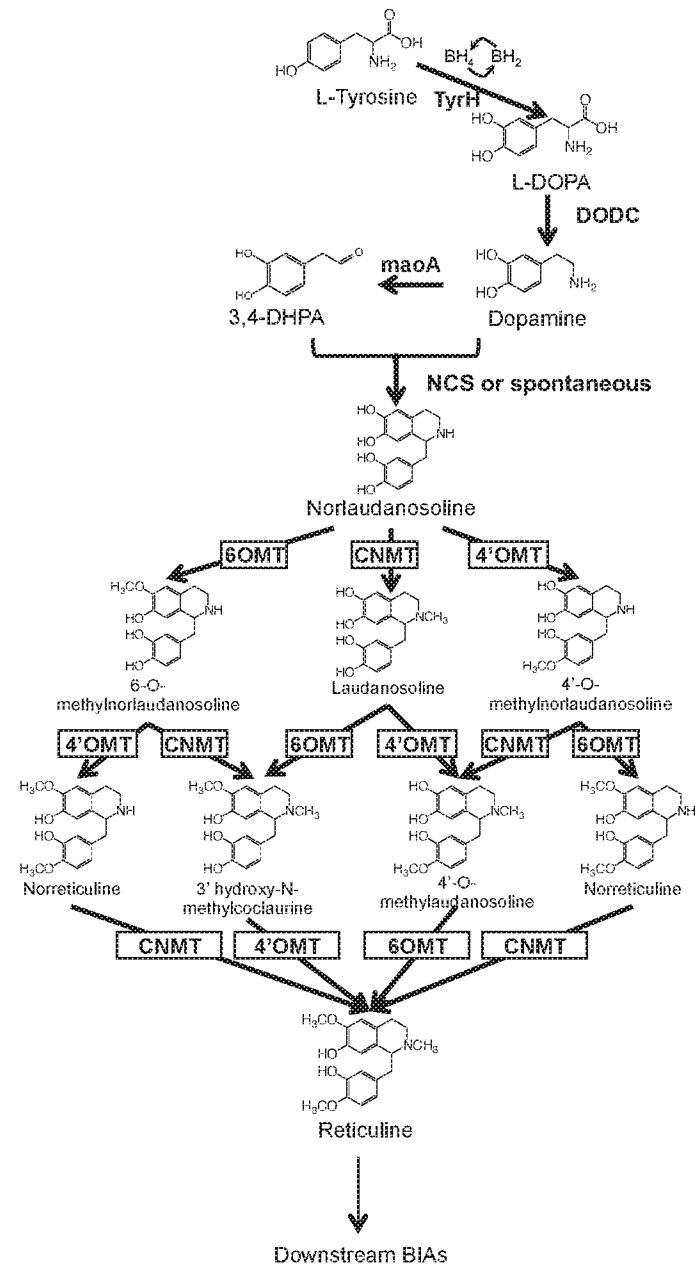


FIG. 13

14/30

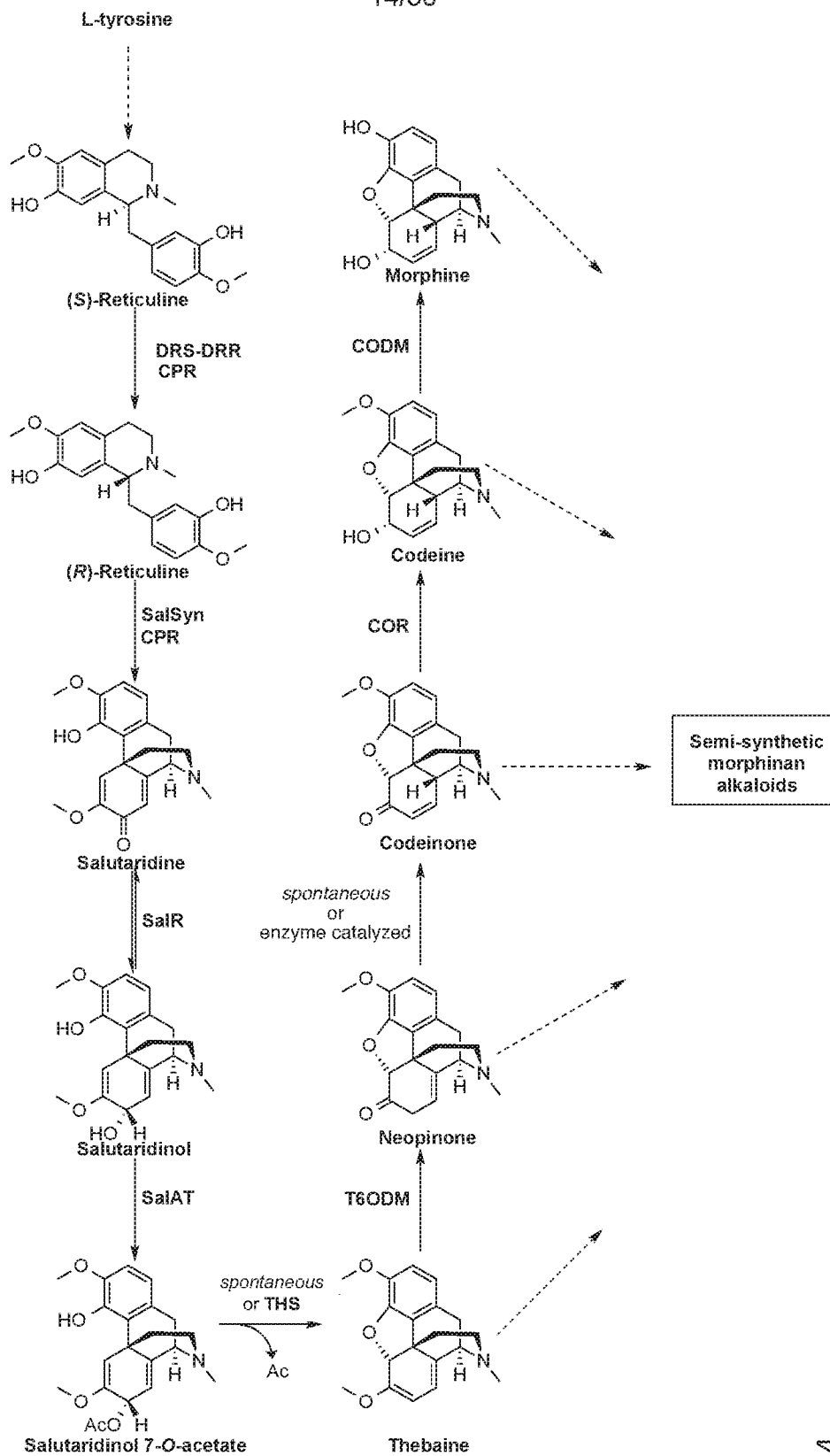
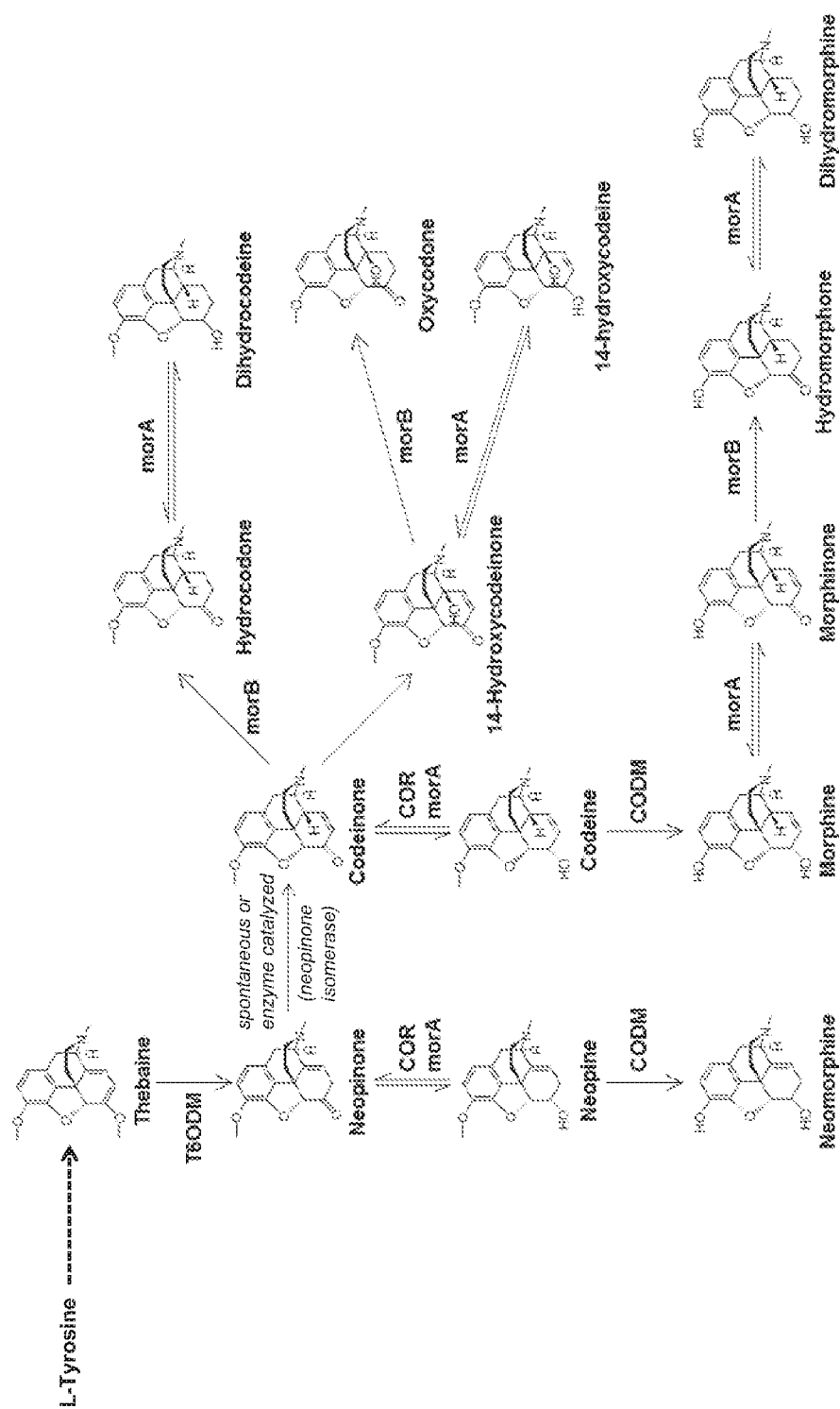


FIG. 14



U.S. DEPARTMENT OF JUSTICE

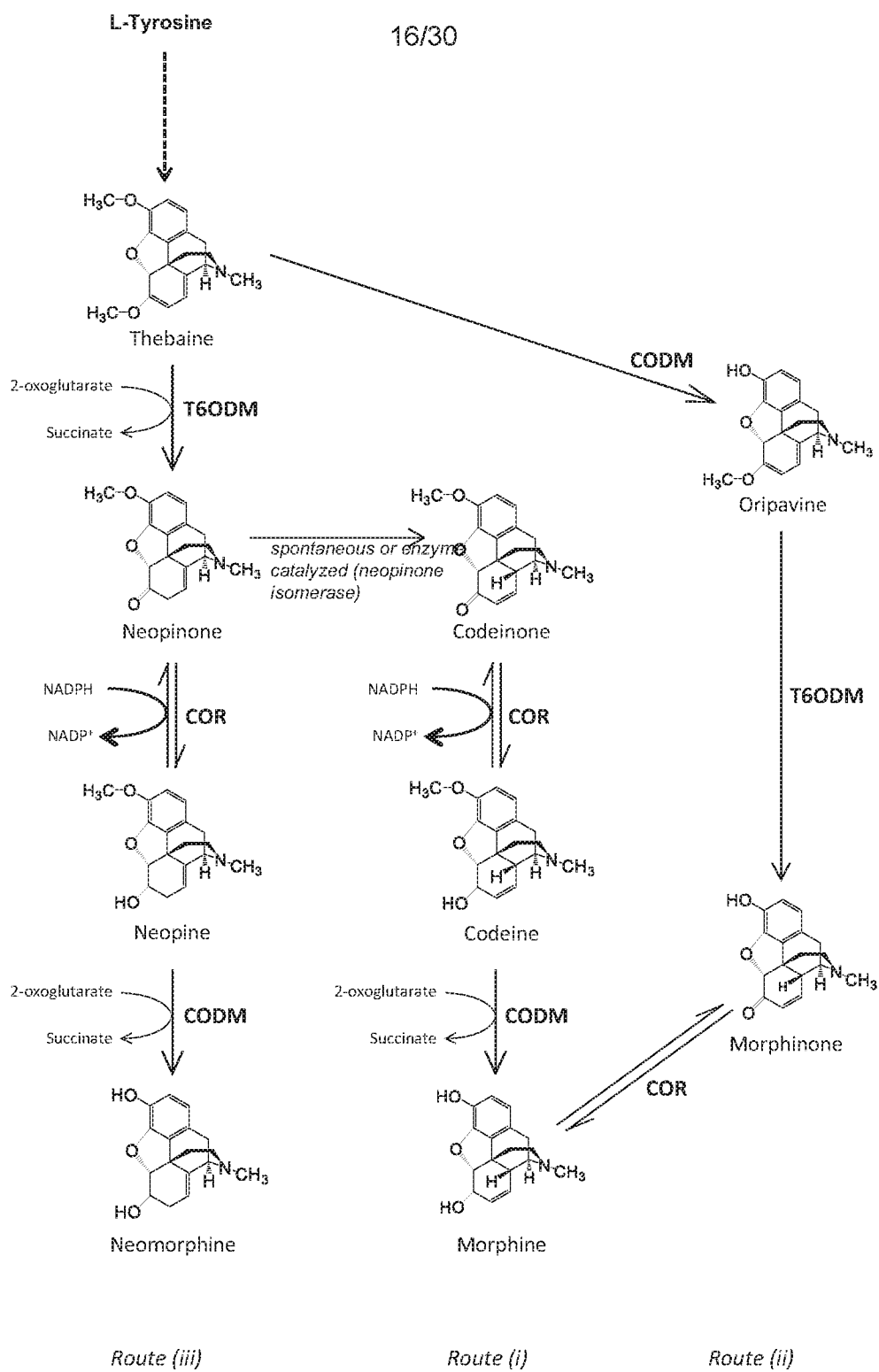


FIG. 16

17/30

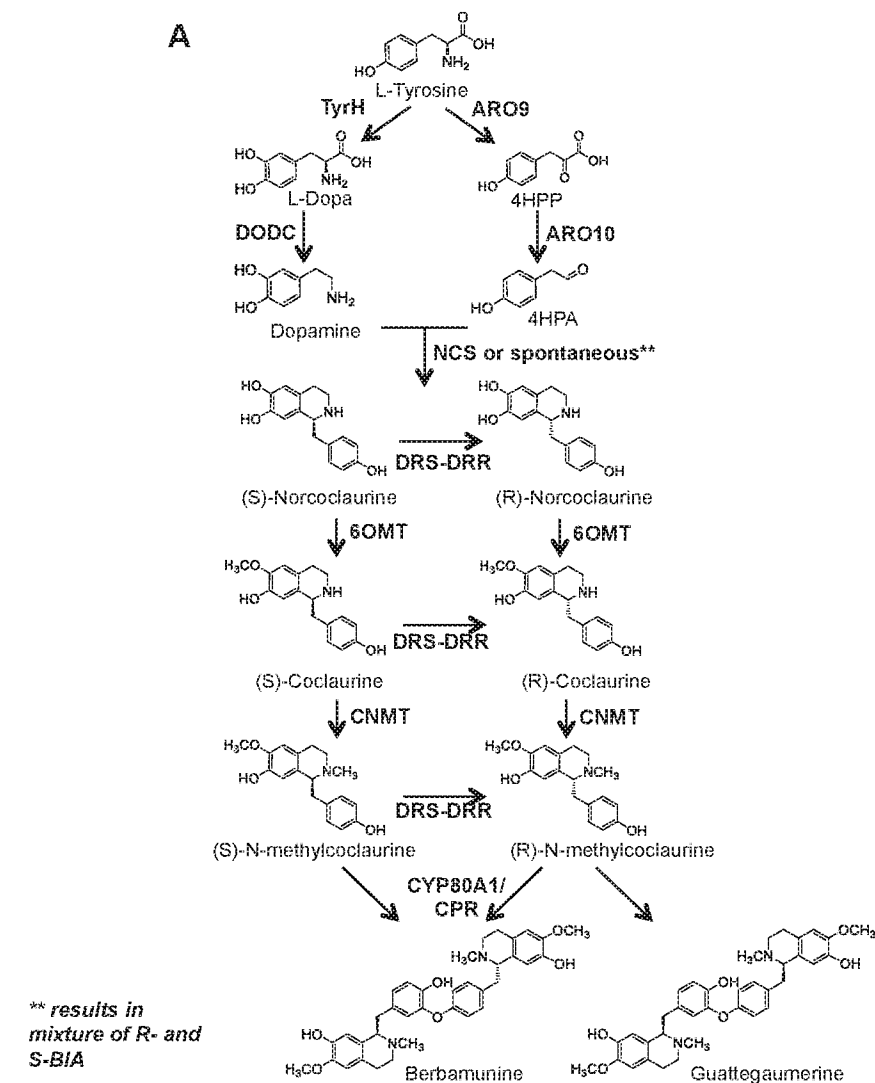
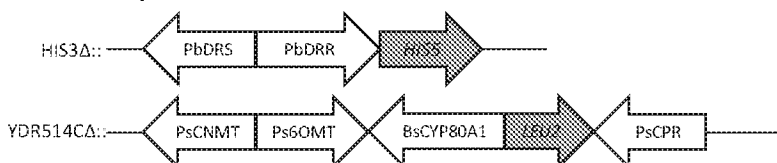
**B****Genotype:**1. HIS3 Δ::P_{GPD}-PbDRS-T_{ADH1}, P_{TP1}-PbDRR-T_{STE2}, HIS5 marker2. YDR514CΔ::P_{PK1}-PsCNMT-T_{Mfa1}, P_{PGK1}-Ps6OMT-T_{PHO5}, P_{GPD}-BsCYP80A1-T_{ADH1}, LEU2 marker, P_{TEF1}-PsCPR-T_{CYC1}**Schematic representation:**

FIG. 17

Reticuline,
salutaridine, and
dehydroreticuline
concentrations

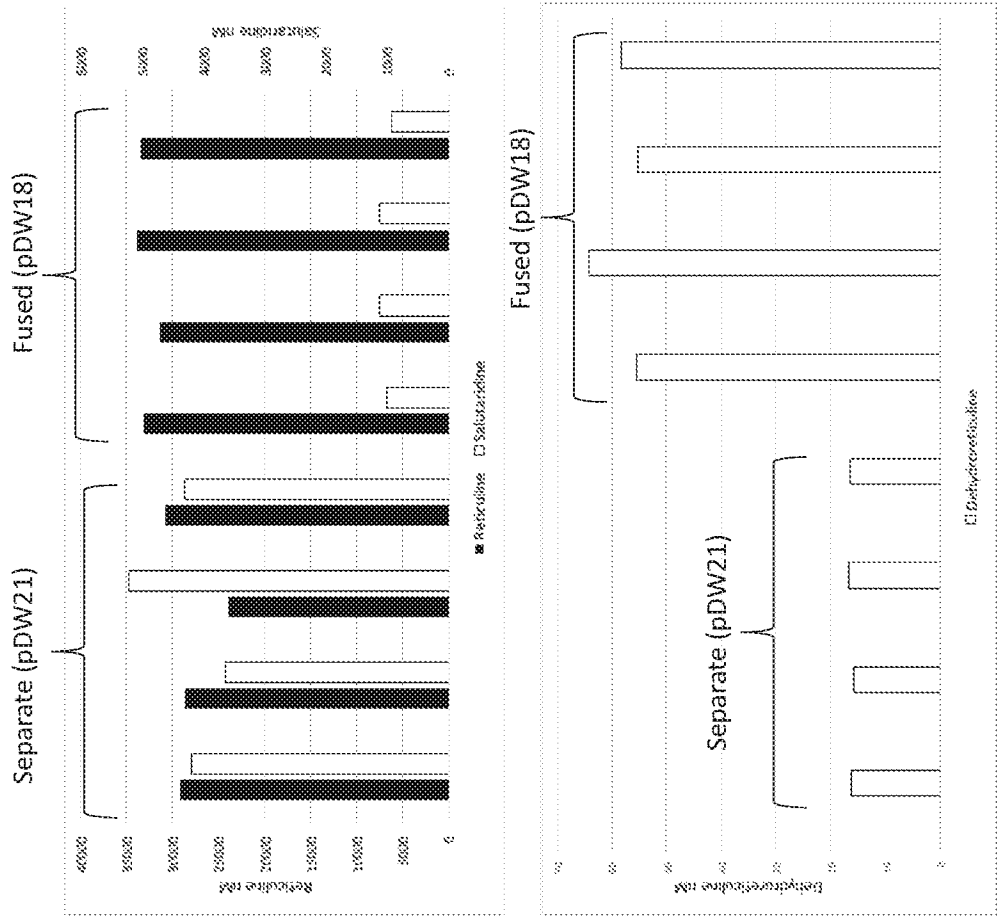
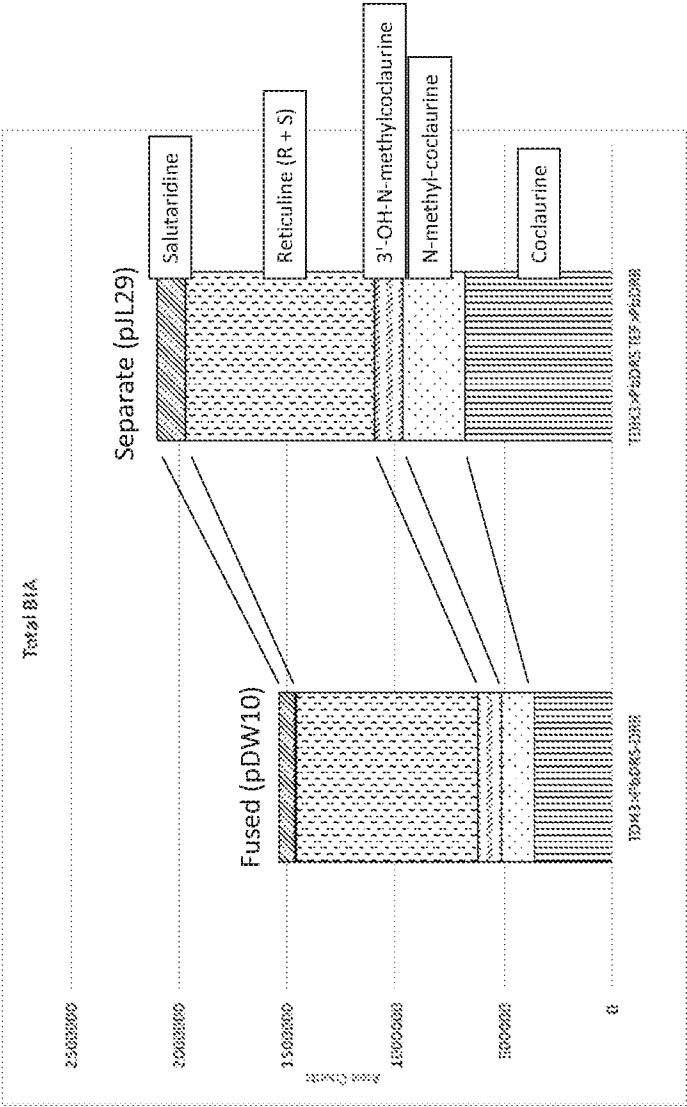
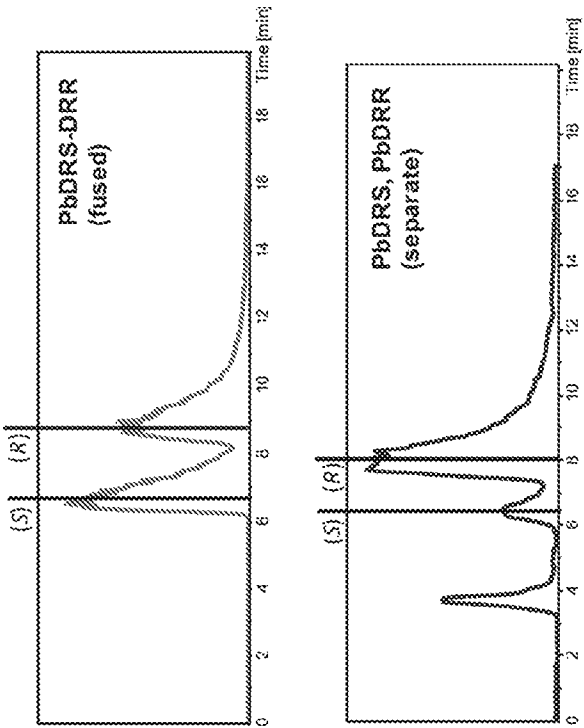


FIG. 18



Total benzylisoquinoline alkaloid (BIA) levels

FIG. 19



Chiral separation of reticuline

FIG. 20

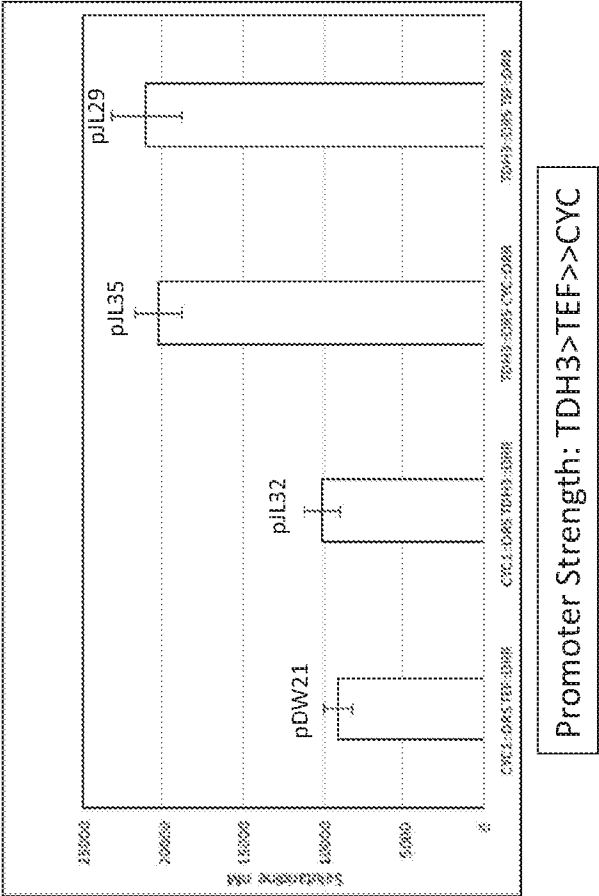
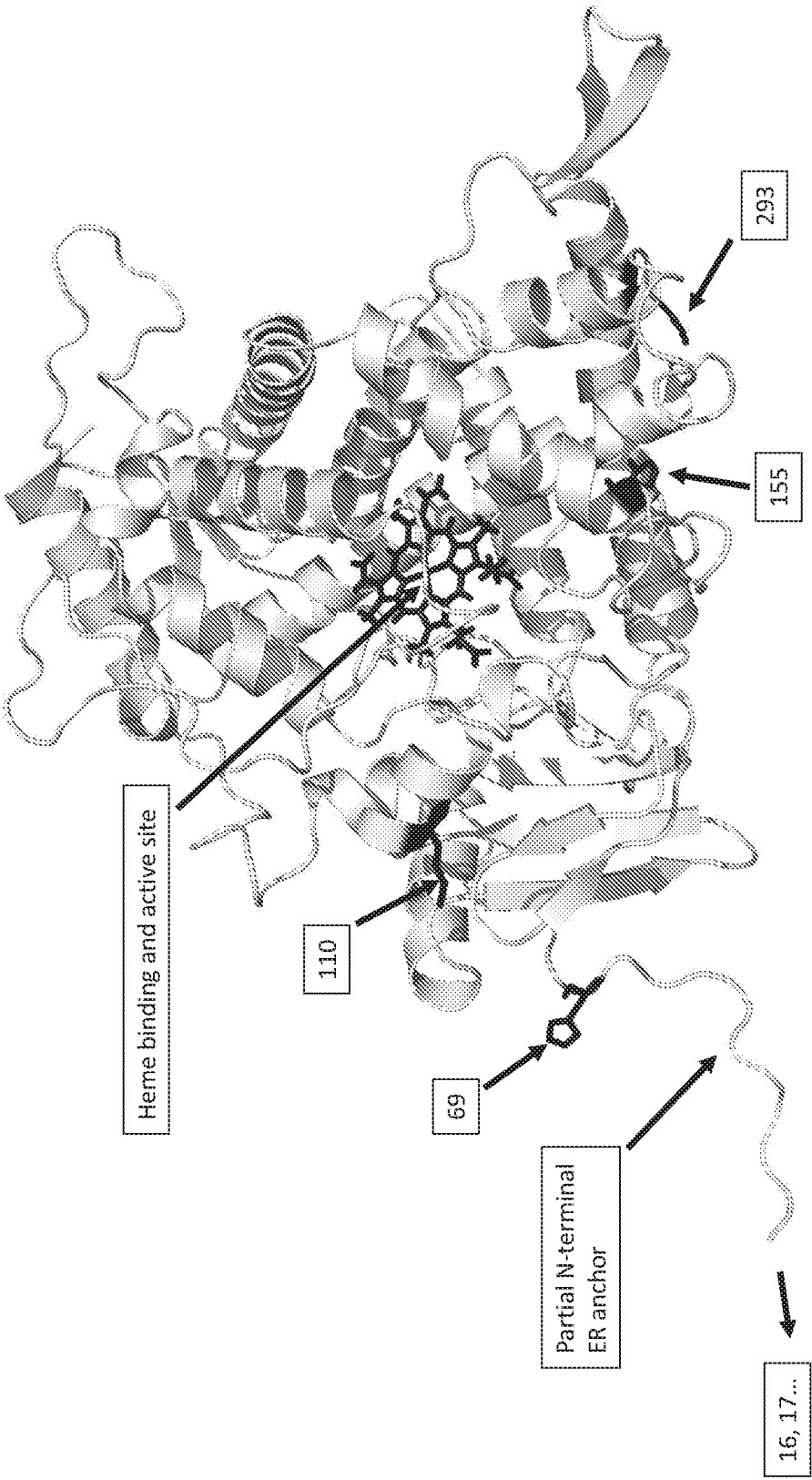
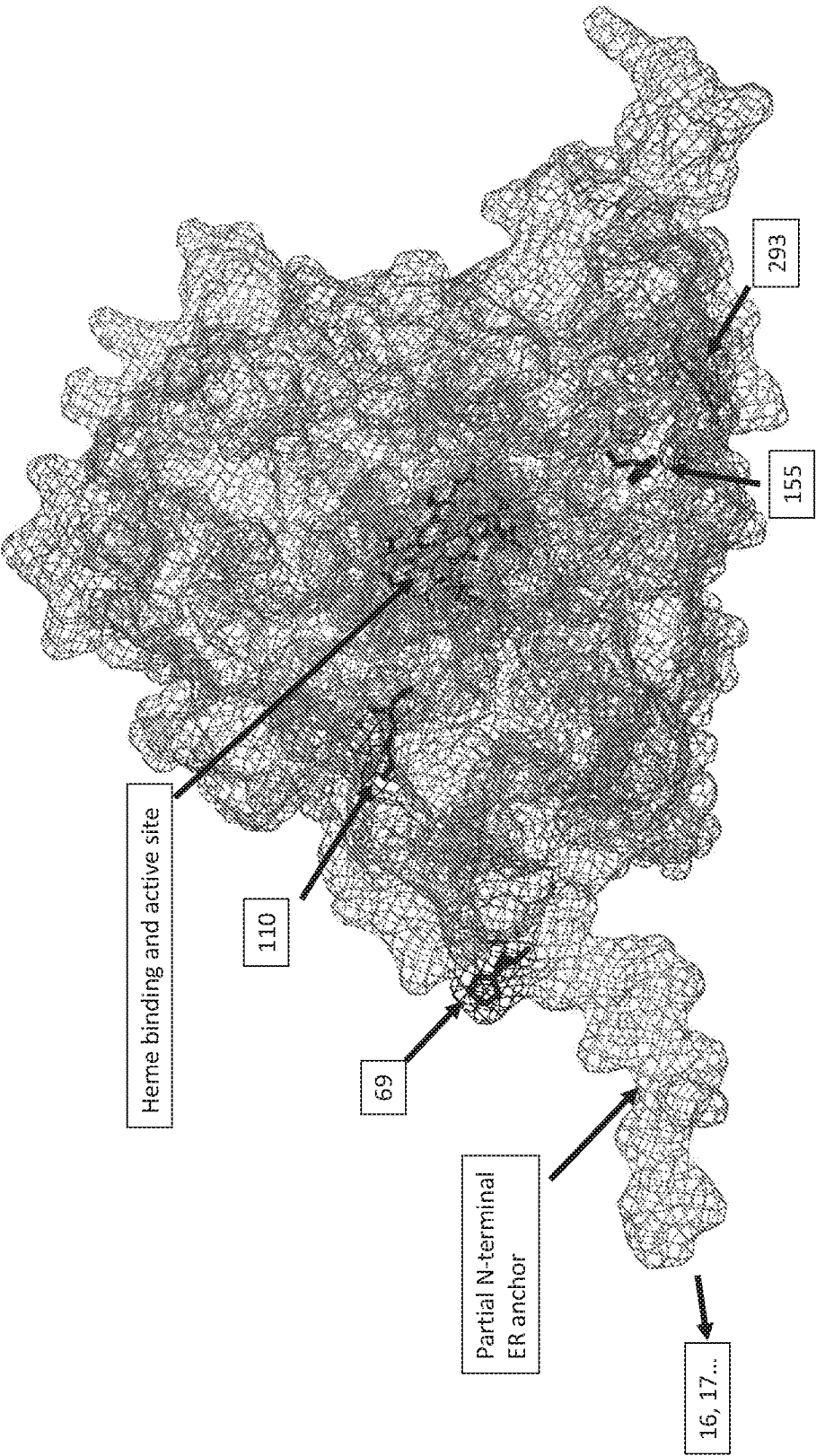


FIG. 21 Salutaridine concentrations



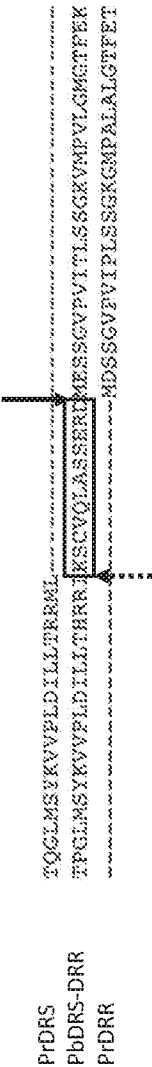
Homology model of DRS based on 4NKX

FIG. 22



Homology model of DRS based on 4NKX

FIG. 23



Alignment between PbDRS-DRR, PrDRS, and PrDRR

FIG. 24

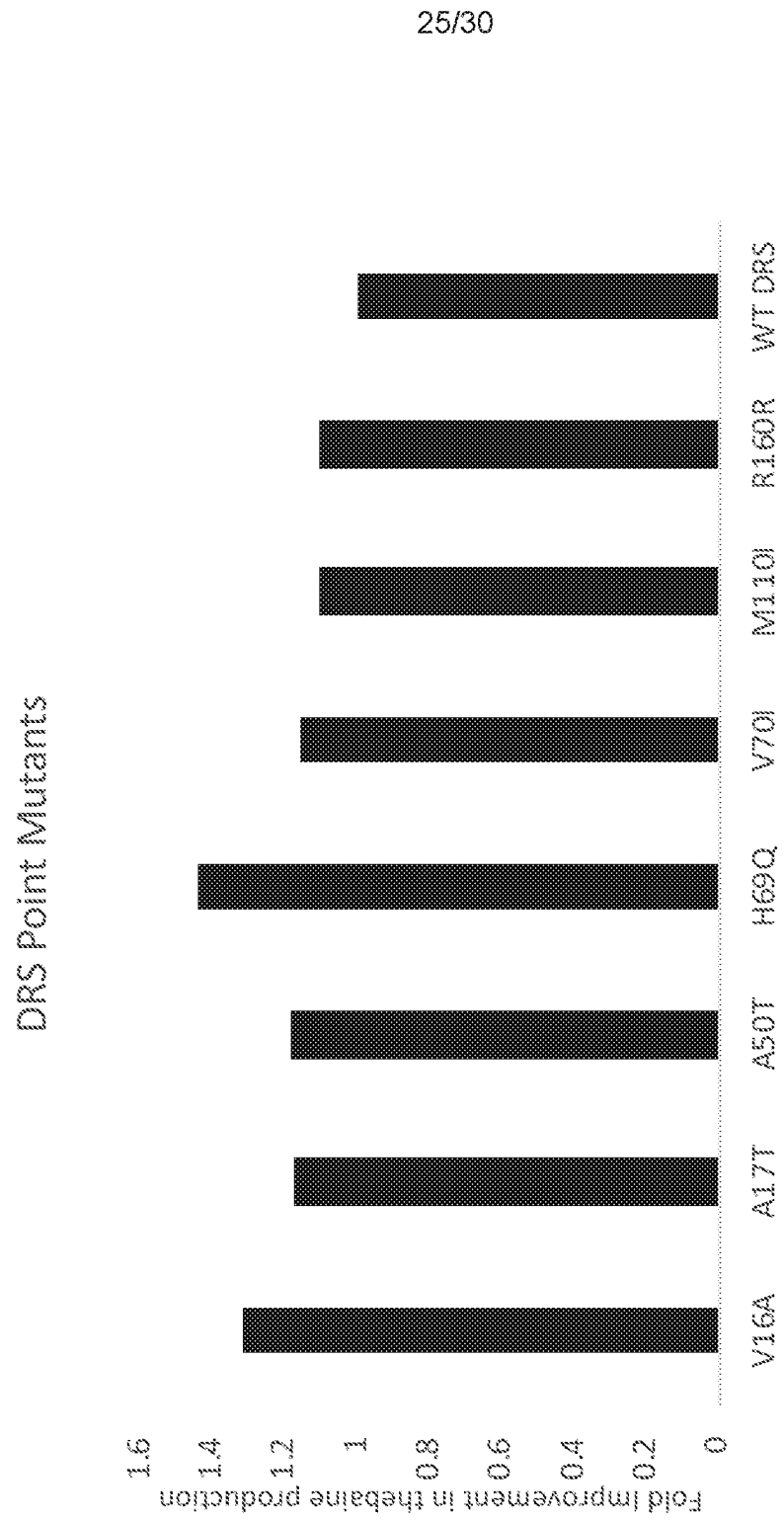


FIG. 25

- Four genetic constructs integrated into the yeast genome:
1. YBR197CA::P_{TPI1}-RnSepR-T_{STE2}, P_{TEF3}-RnPTPS-T_{CVL1}, KanMX marker, P_{GPD}-RnQDHPR-T_{ADH1}, P_{PGK1}-RnPCD-T_{PHOS}
 2. HIS3 Δ::P_{GPD}-RnTyrH-T_{ADH1}, P_{TPI1}-PpDODC-T_{STE2}, HIS5 marker, P_{TEF1}-RnDHR-T_{CVL1}, P_{PGK1}-CJNCS-T_{PHOS}
 3. YDR514CA::P_{PGK1}-PscNMT-T_{MFAL}, P_{PGK1}-PscOMT-T_{PHOS}, P_{GPD}-EcCYP80B1-T_{ADH1}, LEU2 marker, P_{TEF1}-PscPR-T_{CVL1}, P_{TPI1}-Psc4OMT-T_{STE2}
 4. ARO4Δ::P_{TEF1}-ARO4^{FBR}-T_{CVL1}, P_{GPD}-ARO7^{FBR}-T_{ADH1}, HygR marker, P_{PGK1}-TKL1-T_{PHOS}, P_{TPI1}-ARO10-T_{STE2}

Schematic of the integrated constructs indicating the orientation of each expression cassette:

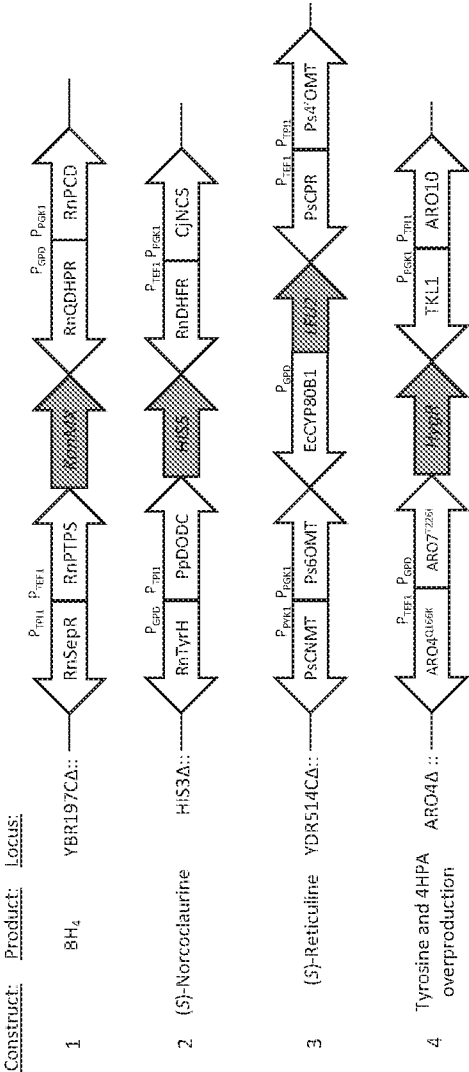


FIG. 26

Example genetic constructs:
5. YPL250CA::P_{GPD}-PbDRS-T_{ADH1}, P_{TEF1}-PbDRR-T_{CYC1}, P_{PGK1}-PstTS-T_{PHOS}, *bleR* marker
6. TRP1::P_{PGK1}-PstSalAT-T_{PHOS5}, P_{TP11}-PbSalR-T_{STE2}, P_{GPD}-EcCFS^{1.83}-PbSalSyn⁹²⁻⁵⁰⁴-T_{ADH1}, *URA3* marker
7. HOA::P_{GPD}-PstT6ODM-T_{ADH1}, P_{PGK1}-morB-T_{PHOS}

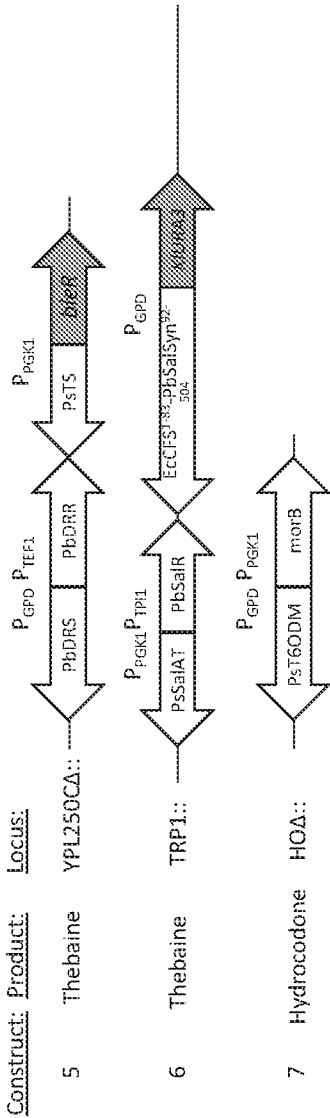


FIG. 27

28/30

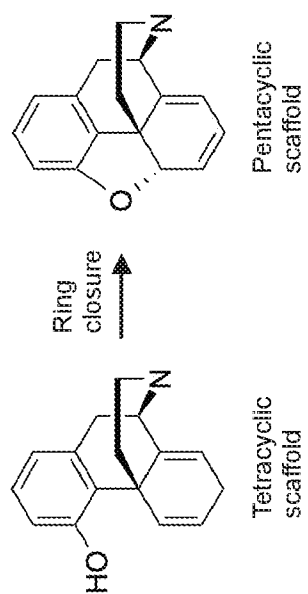


FIG. 28

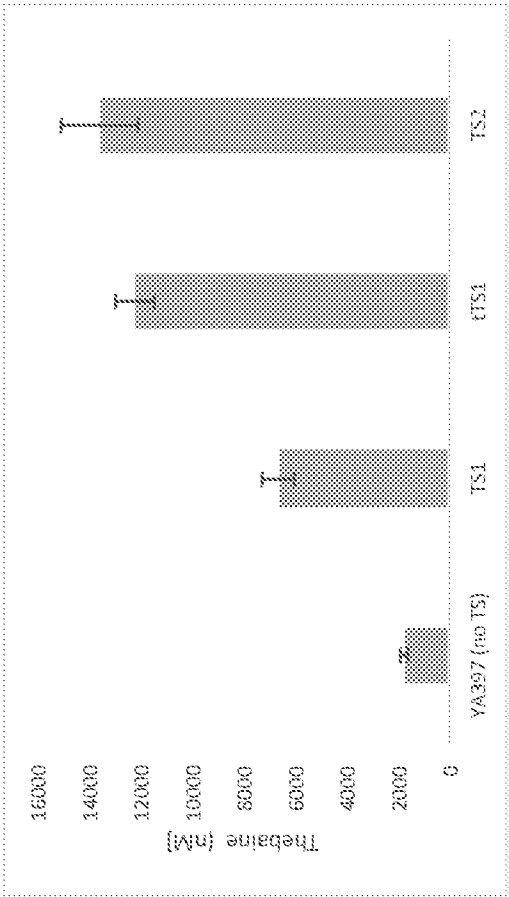


FIG. 29

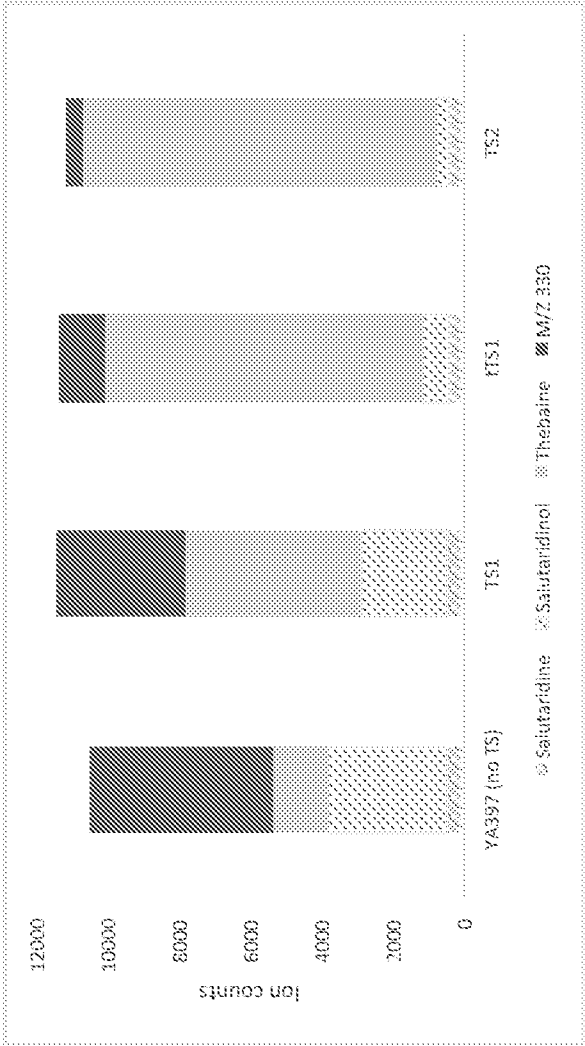


FIG. 30

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US18/45222

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. ☐ forming part of the international application as filed:
 - ☐ in the form of an Annex C/ST.25 text file.
 - ☐ on paper or in the form of an image file.
 - b. ☐ furnished together with the international application under PCT Rule 13ter.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
 - c. ☒ furnished subsequent to the international filing date for the purposes of international search only:
 - ☒ in the form of an Annex C/ST.25 text file (Rule 13ter.1(a)).
 - ☐ on paper or in the form of an image file (Rule 13ter.1(b) and Administrative Instructions, Section 713).
2. ☒ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US18/45222

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

.-***-Please see supplemental page-***-

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
Claims 1-7, 8 (in-part), 9, 10, 12-20, 21 (in-part), 22-24 and 26-37, SEQ ID NO: 17 with a mutation at position 16 (split epimerase oxidase component) and SEQ ID NO: 18 (split epimerase reductase component).

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US18/45222

A. CLASSIFICATION OF SUBJECT MATTER

IPC - C12N 9/90, 15/52, 15/63, 1/19; C12P 17/18 (2019.01)

CPC - C12N 9/90, 15/52, 15/62, 9/0004, 9/0006; C12P 17/18, 41/002; C12Y 101/01247

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

See Search History document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	WO 2016/183023 A1 (THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY) 17 November 2016; paragraphs [0007], [0062], [00101], [00103], [00104], [00148], claim 8, table 1	1-5, 7, 9-10, 12-17, 20, 22-24, 26-28, 34-37 ----- 6, 9, 18, 29/7, 29/20, 30/29/7, 30/29/20, 31/29/7, 31/29/20, 32/7, 32/20, 33/7, 33/20
Y	~ (KEENAN, C et al.) Uncovering the role of hydrophobic residues in Cytochrome P450-Cytochrome P450 reductase interactions. Biochemistry. 17 May 2011, Epub 22 April 2011, Vol. 50, No. 19, pages 3957-3967; abstract; DOI: 10.1021/bi1020748	6, 19
Y	~ (GIRVAN, HM et al.) Applications of microbial cytochrome P450 enzymes in biotechnology and synthetic biology. Current Opinion in Chemical Biology. April 2016, Epub 22 March 2016; Vol. 31, pages 136-145; page 139, 2nd column, 1st paragraph DOI: 10.1016/j.cbpa.2016.02.018	18
Y	~ (FARROW, SC et al.) Stereochemical inversion of (S)-reticuline by a cytochrome P450 fusion in opium poppy. Nature Chemical Biology. September 2015, Epub 1 July 2015, Vol. 11, No. 9, pages 728-732; page 728, 2nd column, 2nd paragraph, page 729, 1st column, 1st paragraph, DOI: 10.1038/nchembio.1879	29/7, 29/20, 30/29/7, 30/29/20, 31/29/7, 31/29/20, 32/7, 32/20, 33/7, 33/2

☒ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

13 December 2018 (13.12.2018)

Date of mailing of the international search report

03 JAN 2019

Name and mailing address of the ISA/

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450
Facsimile No. 571-273-8300

Authorized officer

Shane Thomas

PCT Helpdesk: 571-272-4300
PCT OSF: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US18/45222

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	(DORAY, B et al.) N-Terminal Deletions and His-Tag Fusions Dramatically Affect Expression of Cytochrome P450 2C2 in Bacteria. Archives of Biochemistry and Biophysics. 1 September 2001, Vol. 393, No. 1, pages 143-153; abstract	8, 21
A	(ZELASKO, S et al.) Optimizations to achieve high-level expression of cytochrome P450 proteins using Escherichia coli expression systems. Protein Expression and Purification. November 2013, Epub 20 August 2013, Vol. 9, No. 1, pages 77-87; Table 1; DOI: 10.1016/j.pep.2013.07.017	8, 21
P, X	US 2018/0163241 A1 (SMOLKE, C et al.) 14 June 2018; entire document	1-10, 12-24, 26-28, 29/7, 29/20, 30/29/7, 30/29/20, 31/29/7, 31/29/20, 32/7, 32/20, 33/7, 33/20, 34-37

-Continued from Box No. III: Observations where unity of invention is lacking-

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Groups I+, Claims 1-37, SEQ ID NO: 17 with a mutation at position 16 (split epimerase oxidase component); SEQ ID NO: 18 (split epimerase reductase component) are directed toward engineered host cells comprising an engineered split epimerase capable of catalyzing the epimerization of a (S)-1-benzylisoquinoline alkaloid to a (R)-1-benzylisoquinoline alkaloid and methods for producing said alkaloids.

The engineered cells, engineered epimerase enzymes and methods will be searched to the extent that they encompass a split epimerase enzyme encompassing SEQ ID NOs: 17 with a mutation at position 16 (first exemplary split epimerase oxidase component) and 18 (first exemplary split epimerase reductase component). Applicant is invited to elect additional epimerase polypeptide sequence(s), with specified SEQ ID NO: for each, such that the sequence of each elected species is fully specified (i.e. no optional or variable residues or substituents), to be searched. Additional epimerase sequence(s) will be searched upon the payment of additional fees. It is believed that claims 1-7, 8 (in-part), 9, 10, 12-20, 21 (in-part), 22-24 and 26-37 encompass this first named invention and thus these claims will be searched without fee to the extent that they encompass SEQ ID NO: 17 with a mutation at position 16 (split epimerase oxidase component) and SEQ ID NO: 18 (split epimerase reductase component). Applicants must specify the claims that encompass any additionally elected epimerase polypeptide sequence(s). Applicants must further indicate, if applicable, the claims which encompass the first named invention, if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched/examined. An exemplary election would be SEQ ID NO: 17 with a mutation at position 17 (epimerase oxidase component).

No technical features are shared between the polypeptide sequences of Groups I+ and, accordingly, these groups lack unity a priori.

Additionally, even if Groups I+ were considered to share the technical features of: a method of increasing production of an alkaloid product through the epimerization of a (S)-1-benzylisoquinoline alkaloid to a (R)-1-benzylisoquinoline alkaloid via an engineered split epimerase in an engineered host cell in comparison to the epimerization of a (S)-1-benzylisoquinoline alkaloid to a (R)-1-benzylisoquinoline alkaloid via a fused epimerase in similar conditions, said method comprising: contacting said (S)-1-benzylisoquinoline alkaloid with said engineered split epimerase in said engineered host cell, thereby increasing production of an alkaloid product, wherein contacting said (S)-1-benzylisoquinoline alkaloid with said engineered split epimerase converts said (S)-1-benzylisoquinoline alkaloid to said (R)-1-benzylisoquinoline alkaloid within said engineered host cell; the method of producing an alkaloid product through the epimerization of a (S)-1-benzylisoquinoline alkaloid to a (R)-1-benzylisoquinoline alkaloid via an engineered split epimerase in an engineered host cell said method further comprising: wherein said engineered split epimerase comprises an oxidase component and a reductase component, wherein said oxidase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 17, and wherein said reductase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 18; and an engineered host cell which comprises said engineered split epimerase, wherein said engineered split epimerase increases conversion of a (S)-1-benzylisoquinoline alkaloid to a (R)-1-benzylisoquinoline alkaloid relative to conversion of a (S)-1-benzylisoquinoline alkaloid to a (R)-1-benzylisoquinoline alkaloid via a fused epimerase under similar conditions; however, these shared technical features are previously disclosed by WO2016/183023 A1 to The Board of Trustees of the Leland Stanford Jr. University (hereinafter 'Stanford').

Stanford discloses a method of increasing production of an alkaloid product (increased production of BIAs (benzylisoquinoline alkaloid) in engineered host cells; paragraph [0062]) through the epimerization of a (S)-1-benzylisoquinoline alkaloid to a (R)-1-benzylisoquinoline alkaloid (abstract; paragraph [0004]) via an engineered split epimerase (paragraph [00103]) in an engineered host cell (paragraph [0004]) in comparison to the epimerization of a (S)-1-benzylisoquinoline alkaloid to a (R)-1-benzylisoquinoline alkaloid (100% conversion of (S)-1-benzylisoquinoline alkaloid to a (R)-1-benzylisoquinoline alkaloid by split epimerase; paragraph [00103]) via a fused epimerase (CYR-COR fusions for conversions; paragraph [00102]; Figure 4) in similar conditions (paragraph [00154]), said method comprising: contacting said (S)-1-benzylisoquinoline alkaloid (abstract; paragraph [0006]) with said engineered split epimerase (paragraph [00103]) in said engineered host cell (paragraph [0004]), thereby increasing production of an alkaloid product (increased production of BIAs (benzylisoquinoline alkaloid) in engineered host cells; paragraph [0062]), wherein contacting said (S)-1-benzylisoquinoline alkaloid (abstract; paragraph [0006]) with said engineered split epimerase (paragraph [00103]) converts said (S)-1-benzylisoquinoline alkaloid to said (R)-1-benzylisoquinoline alkaloid (abstract; paragraph [0004]) within said engineered host cell (paragraph [0004]); the method of producing an alkaloid product (paragraph [00105]) through the epimerization of a (S)-1-benzylisoquinoline alkaloid to a (R)-1-benzylisoquinoline alkaloid (abstract; paragraph [0004]) via an engineered split epimerase (paragraph [00103]) in an engineered host cell (paragraph [0004]) said method further comprising: wherein said engineered split epimerase comprises an oxidase component (paragraphs [0058], [0059]) and a reductase component (paragraphs [0058], [0059]), wherein said oxidase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 17 (the epimerase of SEQ ID NO: 14 is 100% identical to SEQ ID NO: 17 of the instant application; paragraph [0007]), and wherein said reductase component comprises a sequence that is at least 80% homologous to SEQ ID NO. 18 (the epimerase of SEQ ID NO: 8 is 100% identical to SEQ ID NO: 18 of the instant application; paragraph [0007]); and an engineered host cell (paragraph [0004]) which comprises said engineered split epimerase (paragraph [00103]), wherein said engineered split epimerase increases conversion of a (S)-1-benzylisoquinoline alkaloid to a (R)-1-benzylisoquinoline alkaloid (100% conversion of (S)-1-benzylisoquinoline alkaloid to a (R)-1-benzylisoquinoline alkaloid by split epimerase; paragraph [00103]) relative to conversion of a (S)-1-benzylisoquinoline alkaloid to a (R)-1-benzylisoquinoline alkaloid via a fused epimerase (CYR-COR fusions for conversions; paragraph [00102]; Figure 4) under similar conditions (paragraph [00154]).

Since none of the special technical features of the Groups I+ inventions is found in more than one of the inventions, and since all of the shared technical features are previously disclosed by the Stanford reference, unity of invention is lacking.