

METHODS FOR THE PRODUCTION OF PSILOCYBIN AND INTERMEDIATES OR SIDE PRODUCTS

FIELD

[0001] The general inventive concepts relate to the field of medical therapeutics and more particularly to methods for the production of psilocybin and intermediates or side products, and methods for the production of norbaeocystin.

CROSS-REFERENCE TO RELATED APPLICATIONS

[0002] The instant application is entitled to priority under 35 U.S.C. §119(e) to U.S. Provisional Application No. 62/926,875, filed October 28, 2019 and to U.S. Provisional Application No. 62/990,633, filed March 17, 2020, each of which is hereby incorporated by reference in its entirety.

SEQUENCE LISTING

[0003] The instant application contains a Sequence Listing which is submitted in ASCII format via EFS-Web and is hereby incorporated by reference in its entirety. The ASCII copy, created on September 17, 2020, is named 315691-00002_Sequence_Listing and is 39,654 bytes in size.

BACKGROUND

[0004] Because of its potential for treatment for a number of anxiety and mental-health related conditions, interest in psilocybin is significant. However, due to roadblocks in routing methods of obtaining drug targets (synthesis and/or extraction from a known biological source), large amounts are not currently available.

[0005] Psilocybin (4-phosphoryloxy-*N,N*-dimethyltryptamine) has gained attention in pharmaceutical markets as a result of recent clinical studies. The efficacy of psilocybin has been demonstrated for the treatment of anxiety in terminal cancer patients and alleviating the symptoms of post-traumatic stress disorder (PTSD). Most recently, the FDA has approved the first Phase IIb clinical trial for the use of psilocybin as a treatment for depression that is not well

controlled with currently available interventions such as antidepressants and cognitive behavioral therapies.

[0006] Psilocybin was first purified from the *Psilocybe mexicana* mushroom by the Swiss chemist, Albert Hoffmann, in 1958. The first reports of the complete chemical synthesis of psilocybin were published in 1959; however, large-scale synthesis methods were not developed until the early 2000's by Shirota and colleagues at the National Institute of Sciences in Tokyo. Despite significant improvements over early synthetic routes, current methods remain tedious and costly, involving numerous intermediate separation and purification steps resulting in an overall yield of 49% from 4-hydroxyindole, incurring an estimated cost of \$2 USD per milligram for pharmaceutical-grade psilocybin.

[0007] Much of the interest in psilocybin is due to its biosynthetic precursors—norbaeocystin and baeocystin. These compounds have structural similarity to the neurotransmitter serotonin and sparked the interest of researchers who were curious to understand the mechanism behind their hallucinogenic properties. After being named a Schedule I compound in the US with implementation of the Controlled Substance Act of 1970, research efforts involving psilocybin were abandoned for other less regulated bioactive molecules; however, experts in the field have suggested a reclassification to schedule IV would be appropriate if a psilocybin-containing medicine were to be approved in the future.

[0008] Clinical trials with psilocybin as a medication for individuals struggling with treatment-resistant depression are ongoing.

[0009] There remains a need for methods for the production of psilocybin and intermediates or side products thereof.

SUMMARY

[0010] The general inventive concepts relate to and contemplate methods and compositions for producing psilocybin or an intermediate or a side product thereof.

[0011] Provided is a method for the production of psilocybin or an intermediate or a side product thereof comprising contacting a prokaryotic host cell with one or more expression vectors, wherein each expression vector comprises a psilocybin production gene selected from the group consisting of *psiD*, *psiK* and *psiM* and combinations thereof; and culturing the host cell. In certain embodiments, the prokaryotic host cell is selected from the group consisting of *Escherichia coli*, *Corynebacterium glutamicum*, *Vibrio natriegens*, *Bacillus subtilis*, *Bacillus megaterium*, *Escherichia coli* Nissle 1917, *Clostridium acetobutylicum*, *Streptomyces coelicolor*, *Lactococcus lactis*, *Pseudomonas putida*, *Streptomyces clavuligerus*, and *Streptomyces venezuelae*.

[0012] In some embodiments, the intermediate or side product of psilocybin is norbaeocystin, baeocystin, 4-hydroxytryptophan, 4-hydroxytryptamine, aeruginascin, psilocin, norpsilocin, or 4-hydroxy-N,N,N-trimethyltryptamonium (4-OH-TMT). In some embodiments the intermediate of psilocybin is norbaeocystin, baeocystin, 4-hydroxytryptophan, or 4-hydroxytryptamine. In some embodiments, the side product of psilocybin is aeruginascin, psilocin, norpsilocin, or 4-hydroxy-N,N,N-trimethyltryptamonium (4-OH-TMT).

[0013] Also provided is a recombinant prokaryotic cell comprising one or more expression vectors, wherein each expression vector comprises a psilocybin production gene selected from the group consisting of *psiD*, *psiK* and *psiM* and combinations thereof. Provided is a vector for introducing at least one gene associated with psilocybin production; the gene may be selected from: *psiD*, *psiK*, and *psiM* and combinations thereof. Also provided is a transfection kit comprising an expression vector as described herein.

[0014] Provided is a method for the production of norbaeocystin comprising contacting a prokaryotic host cell with one or more expression vectors, wherein each expression vector comprises a psilocybin production gene selected from the group consisting of *psiD*, *psiK* and combinations thereof; and culturing the host cell. In certain embodiments, none of the expression vectors comprises *psiM*. In certain embodiments, the prokaryotic host cell is selected from the group consisting of *Escherichia coli*, *Corynebacterium glutamicum*, *Vibrio natriegens*, *Bacillus subtilis*, *Bacillus megaterium*, *Escherichia coli* Nissle 1917, *Clostridium*

acetobutylicum, *Streptomyces coelicolor*, *Lactococcus lactis*, *Pseudomonas putida*, *Streptomyces clavuligerus*, and *Streptomyces venezuelae*.

[0015] Also provided is a recombinant prokaryotic cell comprising one or more expression vectors, wherein each expression vector comprises a psilocybin production gene selected from the group consisting of *psiD*, *psiK*, and combinations thereof. Provided is a vector for introducing at least one gene associated with psilocybin production; the gene may be selected from: *psiD*, *psiK*, and combinations thereof. Also provided is a transfection kit comprising an expression vector as described herein.

DESCRIPTION OF THE FIGURES

[0016] FIGs. 1A-1D show a summary of library configurations and biosynthesis pathway. FIG. 1A shows a copy number library consisting of three plasmids with high (H), medium (M), and low (L) copy number. FIG. 1B shows a Pseudooperon library consisting of a promoter in front of each gene with a single terminator on the high copy number plasmid. FIG. 1C shows a basic operon library consisting of one promoter in front of all three genes on the high copy number plasmid. FIG. 1D shows Psilocybin biosynthesis pathway consisting of three heterologous enzymes, *PsiD*, *PsiK*, and *PsiM*, and highlighting the media supplements in yellow of serine and methionine (as described in the Examples). *PsiD*: L-tryptophan decarboxylase; *PsiK*: kinase; *PsiM*: S-adenosyl-L-methionine (SAM)-dependent N-methyltransferase; TrpB: tryptophan synthase beta subunit; Ser: serine; Met: methionine; 1:4-hydroxyindole; 2: 4-hydroxytryptophan; 3: 4-hydroxytryptamine; 4: norbaeocystin; 5: baecystin; 6: psilocybin.

[0017] FIGs. 2A-2D show a summary of genetic strategies for increasing production. FIG. 2A: Defined copy number library screening. The biosynthesis genes *psiD*, *psiK*, and *psiM* were expressed at either a high (H), medium (M), or low (L) copy number as indicated in the Examples. FIG. 2B: Pseudooperon library screening. The library provided very few mutant constructs with enhanced ability to produce psilocybin over levels previously achieved in the defined copy number library. FIG. 2C: Basic operon library screening. Significant enhancement of library performance was observed under the pseudooperon library. FIG. 2D: Additional screening of top mutants from operon library. The top 10 mutants from the operon library study were subjected to recloning and rescreening under standard conditions. Operon library clones

#13 and #15 (FIG. 2D) demonstrated a large reduction in product titer and were identified as false positives in the original screen. Operon clone #16 (pPsilo16, purple) was selected for further study. All combinations were screened in 48-well plates under standard screening conditions and quantified using HPLC analysis. Error bars represent ± 1 standard deviation from the mean of replicate samples. *Psilocybin not detected.

[0018] FIGs. 3A-3C show a summary of fermentation conditions optimization studies. FIG. 3A shows induction point and temperature screening. The timing of IPTG induction was varied from 1 to 5 hours post inoculation. The data suggest reduced sensitivity to induction point but high sensitivity to production phase temperature with increased production occurring at 37 °C. FIG. 3B shows media, carbon source, and inducer concentration screening. A significant preference was shown for AMM with glucose as the carbon source. FIG. 3C shows effects of media supplementation on psilocybin titer. High sensitivity was observed for changes in the supplement concentration for 4-hydroxyindole, serine, and methionine. Error bars represent ± 1 standard deviation from the mean of replicate samples.

[0019] FIGs 4A-4B show the screening evaluation and bioreactor scale up. FIG. 4A shows a comparison of intermediate and final product titers at various stages of optimization. Stage 1—Initial proof-of-concept All-High control, Stage 2—pPsilo16 post genetic optimization, Stage 3—pPsilo16 post genetic and fermentation optimization. Each additional screening stage further improved final production titer, mainly through reduction of intermediate product buildup. 4OH Ind: 4-hydroxyindole, 4OH-Trp: 4-hydroxytryptophan, 4OH Trm: 4-hydroxytryptamine. FIG. 4B shows fed-batch bioreactor scale up. Through careful monitoring of 4-hydroxyindole feed rate, the concentration of all intermediate products could be kept low resulting in improved growth and psilocybin titers. Error bars represent ± 1 standard deviation from the mean of replicate samples.

[0020] FIG. 5 is a graph showing norbaeocystin production from initial library screen in 48-well plates.

[0021] FIG. 6 is a graph showing norbaeocystin production after additional 4-hydroxyindole exposure to evaluate production in a non-substrate limited environment.

[0022] FIGs. 7A-7D show HPLC standard curves used for metabolite quantification: 4-hydroxyindole (FIG. 7A), 5-hydroxytryptophan (FIG. 7B), 5-hydroxytryptamine (FIG. 7C), psilocybin (FIG. 7D).

[0023] FIG. 8 shows an example chromatogram (280 nm) for HPLC method (1 mL/min) with retention times listed. The data was obtained from a sample of cell-free broth supernatant from an optimized psilocybin production host selected to have major peaks for all relevant metabolites.

[0024] FIG. 9 shows an example chromatogram (280 nm) for LC-MS method (0.25 mL/min) with retention times, MS and MS/MS fragmentation shown. The data was obtained from a sample of cell-free broth supernatant from an optimized psilocybin production host selected to have major peaks for all relevant metabolites.

[0025] FIG. 10 shows 4-hydroxytryptophan analysis in copy number library. 4-hydroxytryptophan was quantified based on the standard curve of 5-hydroxytryptophan due to limited commercial availability and high cost of the authentic standard. Error bars represent ± 1 standard deviation from the mean of triplicate samples.

[0026] FIG. 11 shows 4-hydroxytryptophan analysis in pseudooperon library. Variants are presented in order of decreasing psilocybin production to enable comparison with FIG. 2B. 4-hydroxytryptophan was quantified based on the standard curve of 5-hydroxytryptophan due to limited commercial availability and high cost of the authentic standard.

[0027] FIG. 12 shows 4-hydroxytryptophan analysis in basic operon library. Variants are presented in order of decreasing psilocybin production to enable comparison with FIG. 2C. 4-hydroxytryptophan was quantified based on the standard curve of 5-hydroxytryptophan due to limited commercial availability and high cost of the authentic standard.

[0028] FIG. 13 shows induction sensitivity of pPsilo16 at 37 °C from 0 to 6 hours. Error bars represent ± 1 standard deviation from the mean of duplicate samples.

[0029] FIG. 14 shows induction point sensitivity analysis for pPsilo16 growing in AMM – Glucose at different inducer concentrations. Error bars represent ± 1 deviation from the mean of duplicate samples.

[0030] FIG. 15 shows induction point sensitivity analysis for pPsilo16 growing in AMM – Glycerol at different inducer concentrations. Error bars represent ± 1 deviation from the mean of duplicate samples.

[0031] FIG. 16 shows induction point sensitivity analysis for pPsilo16 growing in LB at different inducer concentrations. Error bars represent ± 1 deviation from the mean of duplicate samples.

[0032] FIGs. 17A-17D show data for fed-batch bioreactor study. FIG. 17A: Measurement of dissolved oxygen (DO), pH, temperature, and agitation rate. FIG. 17B: Total cumulative glucose and ammonium phosphate dibasic fed. OD600 is also shown for reference. FIG. 17C: Total cumulative 4-hydroxyindole fed and 4-hydroxyindole feed rate for the bioreactor scale-up study. The feed rate represents the derivative for the cumulative amount fed. FIG. 17D: Total cumulative 4-hydroxyindole fed compared to psilocybin production for the bioreactor scale-up study. Transient product molar yield shows a maximum molar yield of 60% at roughly 48 hours and a final molar yield of 38% at the end of the scale-up study.

[0033] FIG. 18 shows data for a fed-batch bioreactor study for the high-level production of norbaeocystin in *E. coli*. Transient data for the target product, norbaeocystin, as well as intermediate product, 4-hydroxytryptophan, and starting substrate, 4-hydroxyindole is shown for the 38-hour fermentation process. 4-hydroxyindole was provided continuously using a syringe pump as to limit the accumulation of 4-hydroxytryptophan during the fermentation.

[0034] FIG. 19 shows a full mass spectrum of norbaeocystin produced via *E. coli* fermentation. This data was taken using a Thermo Scientific Orbitrap XL Mass Spectrometer in positive ion mode. The measured mass is in agreement with the actual mass of norbaeocystin to 5 significant figures, further confirming the identity of norbaeocystin in the fermentation broth.

DETAILED DESCRIPTION

[0035] While the general inventive concepts are susceptible of embodiment in many forms, there are shown in the drawings, and will be described herein in detail, specific embodiments thereof with the understanding that the present disclosure is to be considered an exemplification of the principles of the general inventive concepts. Accordingly, the general inventive concepts are not intended to be limited to the specific embodiments illustrated herein.

[0036] It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting.

[0037] The articles “a” and “an” are used herein to refer to one or more than one (i.e., to at least one) of the grammatical object of the article. By way of example, “a cell” means one cell or more than one cell.

[0038] “About” as used herein when referring to a measurable value such as an amount, a temporal duration, and the like, is meant to encompass variations of $\pm 5\%$, preferably $\pm 1\%$, and still more preferably $\pm 0.1\%$ from the specified value, as such variations are appropriate to perform the disclosed methods.

[0039] As used herein, the term “prokaryotic host cell” means a prokaryotic cell that is susceptible to transformation, transfection, transduction, or the like, with a nucleic acid construct or expression vector comprising a polynucleotide. The term “prokaryotic host cell” encompasses any progeny that is not identical due to mutations that occur during replication.

[0040] As used herein, the term “recombinant cell” or “recombinant host” means a cell or host cell that has been genetically modified or altered to comprise a nucleic acid sequence that is not native to the cell or host cell. In some embodiments the genetic modification comprises integrating the polynucleotide in the genome of the host cell. In further embodiments the polynucleotide is exogenous in the host cell.

[0041] As used herein, the term “intermediate” of psilocybin means an intermediate in the production or biosynthesis of psilocybin, e.g., norbaeocystin, baeocystin, 4-hydroxytryptophan, 4-hydroxytryptamine.

[0042] As used herein, the term “side product” of psilocybin means a side product in the production or biosynthesis of psilocybin, e.g., aeruginascin, psilocin, norpsilocin, or 4-hydroxy-N,N,N-trimethyltryptamonium (4-OH-TMT).

[0043] The materials, compositions, and methods described herein are intended to be used to provide novel routes for the production of psilocybin and intermediates or side products, and methods for the production of norbaeocystin.

[0044] Despite advances in the chemical synthesis of psilocybin, current methodologies struggle to provide sufficient material in a cost-effective manner. New advancements fueled Applicant's interest in developing a more cost-effective and easily manipulated host for the biosynthetic production of psilocybin.

[0045] Utilizing the recently identified gene sequences from *P. cubensis* encoding an L-tryptophan decarboxylase (PsiD), a kinase (PsiK), and an S-adenosyl-L-methionine (SAM)-dependent N-methyltransferase (PsiM), together with the promiscuity of the native *Escherichia coli* tryptophan synthase (TrpAB), the biosynthesis pathway capable of psilocybin production from 4-hydroxyindole, was expressed in the prokaryotic model organism *E. coli* BL21 starTM (DE3) (FIG. 1D).

[0046] There is an unmet need for large scale production and isolation of psilocybin. To address these limitations, a series of 3 parallel genetic screening methods were utilized, including: (1) a defined three-level copy number library, (2) a random 5-member operon library, and (3) a random 125-member pseudooperon library. After transcriptional optimization methods were employed, the best strain, pPsilo16, underwent a thorough review and revision of fermentation conditions, resulting in the production of $\sim 139 \pm 2.7$ mg/L of psilocybin from 4-hydroxyindole. Upon further work, a fed-batch bioreactor scale-up resulted in the production of ~ 1160 mg/L of psilocybin, the highest titer reported to date from a recombinant host. Accordingly, the general inventive concepts relate to a novel production pathway and new cell line according to this procedure.

I. Methods, vectors, host cells and kits for the production of psilocybin or an intermediate or a side product thereof

Methods

[0047] Provided herein are the first known methods of *in vivo* psilocybin production using a prokaryotic host. Furthermore, the general inventive concepts are based, in part, on the surprising synergy between increased production through genetic and fermentation means to quickly identify key process parameters required to enable successful scale-up studies culminating in gram scale production of a high-value chemical product.

[0048] Provided is a method for the production of psilocybin or an intermediate or a side product thereof. The method comprises contacting a host cell with at least one psilocybin production gene selected from: *psiD*, *psiK*, *psiM*, and combinations thereof to form a recombinant cell; culturing the recombinant cell; and obtaining the psilocybin. In certain embodiments, the host cell is a prokaryotic cell. In certain exemplary embodiments, the host cell is an *E. coli* cell.

[0049] Provided is a method for the production of psilocybin or an intermediate or a side product thereof comprising contacting a prokaryotic host cell with one or more expression vectors, wherein each expression vector comprises a psilocybin production gene selected from the group consisting of *psiD*, *psiK* and *psiM* and combinations thereof; and culturing the host cell. In certain embodiments, the prokaryotic host cell is selected from the group consisting of *Escherichia coli*, *Corynebacterium glutamicum*, *Vibrio natriegens*, *Bacillus subtilis*, *Bacillus megaterium*, *Escherichia coli* Nissle 1917, *Clostridium acetobutylicum*, *Streptomyces coelicolor*, *Lactococcus lactis*, *Pseudomonas putida*, *Streptomyces clavuligerus*, and *Streptomyces venezuelae*.

[0050] In certain embodiments, the *psiD* comprises the amino acid sequence of SEQ ID NO: 8 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the *psiD* comprises the amino acid sequence of Genbank accession number KY984101.1 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the *psiD* is encoded by a nucleotide sequence comprising SEQ ID NO: 5 or

a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto.

[0051] In certain embodiments, the psiK comprises the amino acid sequence of SEQ ID NO: 9 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the psiK comprises the amino acid sequence of Genbank accession number KY984099.1 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the psiK is encoded by a nucleotide sequence comprising SEQ ID NO: 6 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto.

[0052] In certain embodiments, the psiM comprises the amino acid sequence of SEQ ID NO: 10 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the psiM comprises the amino acid sequence of Genbank accession number KY984100.1 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the psiM is encoded by a nucleotide sequence comprising SEQ ID NO: 7 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto.

[0053] In certain embodiments, the prokaryotic cell is contacted with an expression vector comprising a psiD gene, a psiK gene and a psiM gene all under control of a single promoter in operon configuration. In certain embodiments, the prokaryotic cell is contacted with an expression vector comprising a psiD gene, a psiK gene and a psiM gene, wherein each gene is under control of a separate promoter in pseudooperon configuration. In certain embodiments, each gene is in monocistronic configuration, wherein each gene has a promoter and a terminator. Any configuration or arrangement of promoters and terminators is envisaged.

[0054] In some embodiments, the promoter is selected from the group consisting of G6 mutant T7, H9 mutant T7, H10 mutant T7, C4 mutant T7, consensus T7, Lac, Lac UV5, tac, trc, GAP, and xylA promoter.

[0055] It is envisaged that any intermediate or side product of psilocybin may be produced by any of the methods described herein. In some embodiments, the intermediate or side product of psilocybin is norbaeocystin, baeocystin, 4-hydroxytryptophan, 4-hydroxytryptamine, aeruginascin, psilocin, norpsilocin, or 4-hydroxy-N,N,N-trimethyltryptammonium (4-OH-TMT). In some embodiments the intermediate of psilocybin is norbaeocystin, baeocystin, 4-hydroxytryptophan, or 4-hydroxytryptamine. In some embodiments, the side product of psilocybin is aeruginascin, psilocin, norpsilocin, or 4-hydroxy-N,N,N-trimethyltryptammonium (4-OH-TMT).

[0056] In certain embodiments, the host cell is cultured with a supplement independently selected from the group consisting of 4-hydroxyindole, serine, methionine, 4-hydroxytryptophan, 4-hydroxytryptamine, and combinations thereof. In certain exemplary embodiments, the supplement is fed continuously to the host cell. In further embodiments, the host cell is grown in an actively growing culture. Continuous feeding is accomplished by using a series of syringe and/or peristaltic pumps whose outlet flow is directly connected to the bioreactor. The set point of these supplement addition pumps is adjusted in response to real-time measurement of cell biomass and specific metabolic levels using UV-vis absorption and HPLC analysis, respectively. The fed-batch fermentation process is focused on maximizing production of target metabolites through harnessing the ability of an actively growing and replicating cell culture to regenerate key co-factors and precursors which are critical to the biosynthesis of target metabolites. This process notably does not involve the centrifugal concentration and reconstitution of cell biomass to artificially higher cell density and/or into production media that was not used to build the initial biomass. The production process involves the inoculation of the reactor from an overnight preculture at low optical density, followed by exponential phase growth entering into a fed-batch phase of production, culminating in a high cell density culture.

[0057] The psilocybin and intermediate or side products are found extracellularly in the fermentation broth. In certain embodiments, the psilocybin and intermediate or side products are

isolated. These target products can be collected through drying the fermentation broth after centrifugation to remove the cell biomass. The resulting dry product can be extracted to further purify the target compounds. Alternatively, the products can be extracted from the liquid cell culture broth using a solvent which is immiscible with water and partitions psilocybin or any of the intermediate or side products into the organic phase. Furthermore, contaminants from the fermentation broth can be removed through extraction leaving the psilocybin and/or intermediate or side products in the aqueous phase for collection after drying or crystallization procedures.

[0058] In certain embodiments, the methods described herein result in a titer of psilocybin of about 0.5 to about 50 g/L. In some embodiments, the methods described herein result in a titer of psilocybin of about 0.5 to about 10 g/L. In yet further embodiments, the methods described herein result in a titer of psilocybin of about 0.5 to about 2 g/L. In certain embodiments, the methods described herein result in a titer of psilocybin of about 1.0 to about 1.2 g/L. In further embodiments, the methods described herein result in a titer of psilocybin of about 1.16 g/L.

[0059] In certain embodiments, the methods described herein result in a molar yield of psilocybin of about 10% to about 100%. In some embodiments, the methods described herein result in a molar yield of psilocybin of about 20% to about 80%. In yet further embodiments, the methods described herein result in a molar yield of psilocybin of about 30% to about 70%. In certain embodiments, the methods described herein result in a molar yield of psilocybin of about 40% to about 60%. In further embodiments, the methods described herein result in a molar yield of psilocybin of about 50%.

Recombinant prokaryotic cells for the production of psilocybin or an intermediate or a side product thereof

[0060] Provided is a recombinant prokaryotic cell comprising one or more expression vectors, wherein each expression vector comprises a psilocybin production gene selected from the group consisting of psiD, psiK and psiM and combinations thereof.

[0061] In certain embodiments, the recombinant prokaryotic cell is selected from the group consisting of *Escherichia coli*, *Corynebacterium glutamicum*, *Vibrio natriegens*, *Bacillus subtilis*, *Bacillus megaterium*, *Escherichia coli* Nissle 1917, *Clostridium acetobutylicum*,

Streptomyces coelicolor, *Lactococcus lactis*, *Pseudomonas putida*, *Streptomyces clavuligerus*, and *Streptomyces venezuelae*.

[0062] In certain embodiments, the psiD comprises the amino acid sequence of SEQ ID NO: 8 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the psiD comprises the amino acid sequence of Genbank accession number KY984101.1 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the psiD is encoded by a nucleotide sequence comprising SEQ ID NO: 5 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto.

[0063] In certain embodiments, the psiK comprises the amino acid sequence of SEQ ID NO: 9 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the psiK comprises the amino acid sequence of Genbank accession number KY984099.1 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the psiK is encoded by a nucleotide sequence comprising SEQ ID NO: 6 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto.

[0064] In certain embodiments, the psiM comprises the amino acid sequence of SEQ ID NO: 10 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the psiM comprises the amino acid sequence of Genbank accession number KY984100.1 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the psiM is encoded by a nucleotide sequence comprising SEQ ID NO: 7 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto.

[0065] In certain embodiments, the prokaryotic cell is contacted with an expression vector comprising a *psiD* gene, a *psiK* gene and a *psiM* gene all under control of a single promoter in operon configuration. In certain embodiments, the prokaryotic cell is contacted with an expression vector comprising a *psiD* gene, a *psiK* gene and a *psiM* gene, wherein each gene is under control of a separate promoter in pseudooperon configuration. In certain embodiments, each gene is in monocistronic configuration, wherein each gene has a promoter and a terminator. Any configuration or arrangement of promoters and terminators is envisaged.

[0066] In some embodiments, the promoter is selected from the group consisting of G6 mutant T7, H9 mutant T7, H10 mutant T7, C4 mutant T7, consensus T7, Lac, Lac UV5, tac, trc, GAP, and *xylA* promoter.

Expression vectors

[0067] Provided is a vector for introducing at least one gene associated with psilocybin production; the gene may be selected from: *psiD*, *psiK*, and *psiM* and combinations thereof.

[0068] In certain embodiments, the *psiD* comprises the amino acid sequence of SEQ ID NO: 8 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the *psiD* comprises the amino acid sequence of Genbank accession number KY984101.1 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the *psiD* is encoded by a nucleotide sequence comprising SEQ ID NO: 5 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto.

[0069] In certain embodiments, the *psiK* comprises the amino acid sequence of SEQ ID NO: 9 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the *psiK* comprises the amino acid sequence of Genbank accession number KY984099.1 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In

certain embodiments, the psiK is encoded by a nucleotide sequence comprising SEQ ID NO: 6 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto.

[0070] In certain embodiments, the psiM comprises the amino acid sequence of SEQ ID NO: 10 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the psiM comprises the amino acid sequence of Genbank accession number KY984100.1 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the psiM is encoded by a nucleotide sequence comprising SEQ ID NO: 7 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto.

[0071] In certain embodiments, the expression vector comprises a psiD gene, a psiK gene and a psiM gene all under control of a single promoter in operon configuration. In certain embodiments, the expression vector comprises a psiD gene, a psiK gene and a psiM gene, wherein each gene is under control of a separate promoter in pseudooperon configuration. In certain embodiments, each gene is in monocistronic configuration, wherein each gene has a promoter and a terminator. Any configuration or arrangement of promoters and terminators is envisaged.

[0072] In certain embodiments, the expression vector comprises the nucleic acid sequence of SEQ ID NO: 22 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the expression vector is pPsilo16 or a vector having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto.

[0073] In some embodiments, the promoter is selected from the group consisting of G6 mutant T7, H9 mutant T7, H10 mutant T7, C4 mutant T7, consensus T7, Lac, Lac UV5, tac, trc, GAP, and xylA promoter.

Kits

[0074] Provided is a transfection kit comprising an expression vector as described herein. Such a kit may comprise a carrying means being compartmentalized to receive in close confinement one or more container means such as, e.g., vials or test tubes. Each of such container means comprises components or a mixture of components needed to perform a transfection. Such kits may include, for example, one or more components selected from vectors, cells, reagents, lipid-aggregate forming compounds, transfection enhancers, or biologically active molecules.

II. Methods, vectors, host cells and kits for the production of norbaeocystin**Methods**

[0075] Provided is a method for the production of norbaeocystin comprising contacting a prokaryotic host cell with one or more expression vectors, wherein each expression vector comprises a psilocybin production gene selected from the group consisting of psiD, psiK and combinations thereof; and culturing the host cell. In certain embodiments, none of the expression vectors comprises psiM.

[0076] In certain embodiments, the psiD comprises the amino acid sequence of SEQ ID NO: 8 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the psiD comprises the amino acid sequence of Genbank accession number KY984101.1 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the psiD is encoded by a nucleotide sequence comprising SEQ ID NO: 5 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto.

[0077] In certain embodiments, the psiK comprises the amino acid sequence of SEQ ID NO: 9 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the psiK comprises the amino acid sequence of Genbank accession number

KY984099.1 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the psiK is encoded by a nucleotide sequence comprising SEQ ID NO: 6 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto.

[0078] In certain embodiments, the recombinant prokaryotic cell is selected from the group consisting of *Escherichia coli*, *Corynebacterium glutamicum*, *Vibrio natriegens*, *Bacillus subtilis*, *Bacillus megaterium*, *Escherichia coli* Nissle 1917, *Clostridium acetobutylicum*, *Streptomyces coelicolor*, *Lactococcus lactis*, *Pseudomonas putida*, *Streptomyces clavuligerus*, and *Streptomyces venezuelae*.

[0079] In certain embodiments, the prokaryotic cell is contacted with an expression vector comprising a psilocybin production gene selected from the group consisting of a psiD gene, a psiK gene, and combinations thereof, all under control of a single promoter in operon configuration. In certain embodiments, the prokaryotic cell is contacted with an expression vector comprising a psiD gene and a psiK gene, wherein each gene is under control of a separate promoter in pseudooperon configuration. In certain embodiments, each gene is in monocistronic configuration, wherein each gene has a promoter and a terminator. Any configuration or arrangement of promoters and terminators is envisaged. In certain embodiments, none of the expression vectors comprises a psiM gene.

[0080] In some embodiments, the promoter is selected from the group consisting of G6 mutant T7, H9 mutant T7, H10 mutant T7, C4 mutant T7, consensus T7, Lac, Lac UV5, tac, trc, GAP, and xylA promoter.

[0081] In certain embodiments, the host cell is cultured with a supplement independently selected from the group consisting of 4-hydroxyindole, serine, methionine, 4-hydroxytryptophan, 4-hydroxytryptamine, and combinations thereof. In certain exemplary embodiments, the supplement is fed continuously to the host cell. In further embodiments, the host cell is grown in an actively growing culture. Continuous feeding is accomplished by using a series of syringe and/or peristaltic pumps whose outlet flow is directly connected to the bioreactor. The set point

of these supplement addition pumps is adjusted in response to real-time measurement of cell biomass and specific metabolic levels using UV-vis absorption and HPLC analysis, respectively. The fed-batch fermentation process is focused on maximizing production of target metabolites through harnessing the ability of an actively growing and replicating cell culture to regenerate key co-factors and precursors which are critical to the biosynthesis of target metabolites. This process notably does not involve the centrifugal concentration and reconstitution of cell biomass to artificially higher cell density and/or into production media that was not used to build the initial biomass. The production process involves the inoculation of the reactor from an overnight preculture at low optical density, followed by exponential phase growth entering into a fed-batch phase of production, culminating in a high cell density culture.

[0082] The norbaeocystin is found extracellularly in the fermentation broth. In certain embodiments, the norbaeocystin is isolated. Norbaeocystin can be collected through drying the fermentation broth after centrifugation to remove the cell biomass. The resulting dry product can be extracted to further purify the norbaeocystin. Alternatively, the norbaeocystin can be extracted from the liquid cell culture broth using a solvent which is immiscible with water and partitions norbaeocystin into the organic phase. Furthermore, contaminants from the fermentation broth can be removed through extraction leaving the norbaeocystin in the aqueous phase for collection after drying or crystallization procedures.

[0083] In certain embodiments, the methods described herein result in a titer of norbaeocystin of about 0.1 to about 50 g/L. In some embodiments, the methods described herein result in a titer of norbaeocystin of about 0.1 to about 10 g/L. In yet further embodiments, the methods described herein result in a titer of norbaeocystin of about 0.1 to about 2 g/L. In certain embodiments, the methods described herein result in a titer of norbaeocystin of about 0.1 to about 1.0 g/L. In further embodiments, the methods described herein result in a titer of norbaeocystin of about 0.4 to about 0.8 g/L. In further embodiments, the methods described herein result in a titer of norbaeocystin of about 0.7 g/L.

[0084] In certain embodiments, the methods described herein result in a molar yield of norbaeocystin of about 10% to about 100%. In some embodiments, the methods described herein result in a molar yield of norbaeocystin of about 20% to about 80%. In yet further

embodiments, the methods described herein result in a molar yield of norbaeocystin of about 30% to about 70%. In certain embodiments, the methods described herein result in a molar yield of norbaeocystin of about 40% to about 60%. In further embodiments, the methods described herein result in a molar yield of norbaeocystin of about 50%.

Recombinant prokaryotic cells for the production of norbaeocystin

[0085] Provided is a recombinant prokaryotic cell comprising one or more expression vectors, wherein each expression vector comprises a psilocybin production gene selected from the group consisting of psiD, psiK, and combinations thereof. In certain embodiments, none of the expression vectors comprises psiM.

[0086] In certain embodiments, the recombinant prokaryotic cell is selected from the group consisting of *Escherichia coli*, *Corynebacterium glutamicum*, *Vibrio natriegens*, *Bacillus subtilis*, *Bacillus megaterium*, *Escherichia coli* Nissle 1917, *Clostridium acetobutylicum*, *Streptomyces coelicolor*, *Lactococcus lactis*, *Pseudomonas putida*, *Streptomyces clavuligerus*, and *Streptomyces venezuelae*.

[0087] In certain embodiments, the psiD comprises the amino acid sequence of SEQ ID NO: 8 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the psiD comprises the amino acid sequence of Genbank accession number KY984101.1 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the psiD is encoded by a nucleotide sequence comprising SEQ ID NO: 5 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto.

[0088] In certain embodiments, the psiK comprises the amino acid sequence of SEQ ID NO: 9 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the psiK comprises the amino acid sequence of Genbank accession number KY984099.1 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least

95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the *psiK* is encoded by a nucleotide sequence comprising SEQ ID NO: 6 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto.

[0089] In certain embodiments, the prokaryotic cell is contacted with an expression vector comprising a *psiD* gene and a *psiK* gene all under control of a single promoter in operon configuration. In certain embodiments, the prokaryotic cell is contacted with an expression vector comprising a *psiD* gene and a *psiK* gene, wherein each gene is under control of a separate promoter in pseudooperon configuration. In certain embodiments, each gene is in monocistronic configuration, wherein each gene has a promoter and a terminator. Any configuration or arrangement of promoters and terminators is envisaged. In certain embodiments, none of the expression vectors comprises a *psiM* gene.

[0090] In some embodiments, the promoter is selected from the group consisting of G6 mutant T7, H9 mutant T7, H10 mutant T7, C4 mutant T7, consensus T7, Lac, Lac UV5, tac, trc, GAP, and *xylA* promoter.

Expression vectors

[0091] Provided is a vector for introducing at least one gene associated with psilocybin production; the gene may be selected from: *psiD*, *psiK*, and combinations thereof.

[0092] In certain embodiments, the *psiD* comprises the amino acid sequence of SEQ ID NO: 8 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the *psiD* comprises the amino acid sequence of Genbank accession number KY984101.1 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the *psiD* is encoded by a nucleotide sequence comprising SEQ ID NO: 5 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto.

[0093] In certain embodiments, the psiK comprises the amino acid sequence of SEQ ID NO: 9 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the psiK comprises the amino acid sequence of Genbank accession number KY984099.1 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the psiK is encoded by a nucleotide sequence comprising SEQ ID NO: 6 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto.

[0094] In certain embodiments, the prokaryotic cell is contacted with an expression vector comprising a psiD gene and a psiK gene all under control of a single promoter in operon configuration. In certain embodiments, the prokaryotic cell is contacted with an expression vector comprising a psiD gene and a psiK gene, wherein each gene is under control of a separate promoter in pseudooperon configuration. In certain embodiments, each gene is in monocistronic configuration, wherein each gene has a promoter and a terminator. Any configuration or arrangement of promoters and terminators is envisaged. In certain embodiments, none of the expression vectors comprises a psiM gene.

[0095] In some embodiments, the promoter is selected from the group consisting of G6 mutant T7, H9 mutant T7, H10 mutant T7, C4 mutant T7, consensus T7, Lac, Lac UV5, tac, trc, GAP, and xylA promoter.

[0096] In certain embodiments, the expression vector comprises the nucleic acid sequence of SEQ ID NO: 23 or a sequence having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto. In certain embodiments, the expression vector is pETM6-C4-psiDK or a vector having at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity thereto.

Kits

[0097] Provided is a transfection kit comprising an expression vector as described herein. Such a kit may comprise a carrying means being compartmentalized to receive in close confinement one or more container means such as, e.g., vials or test tubes. Each of such container means comprises components or a mixture of components needed to perform a transfection. Such kits may include, for example, one or more components selected from vectors, cells, reagents, lipid-aggregate forming compounds, transfection enhancers, or biologically active molecules

EXAMPLES

[0098] The following examples describe various compositions and methods for genetic modification of cells to aid in the production of psilocybin, according to the general inventive concepts.

Example 1**Materials and Methods***Bacterial strains, vectors, and media*

[0099] *E. coli* DH5 α was used to propagate all plasmids, while BL21 starTM (DE3) was used as the host for all chemical production experiments. Plasmid transformations were completed using standard electro and chemical competency protocols as specified. Unless noted otherwise, Andrew's Magic Media (AMM) was used for both overnight growth and production media, while Luria Broth (LB) was used for plasmid propagation during cloning. The antibiotics ampicillin (80 μ g/mL), chloramphenicol (25 μ g/mL), and streptomycin (50 μ g/mL) were added at their respective concentrations to the culture media when using pETM6, pACM4, and pCDM4-derived vectors, respectively. The exogenous pathway genes encoding the enzymes PsiD, PsiK, and PsiM contained on plasmids pJF24, pJF23, and pFB13, respectively, were obtained from the Hoffmeister group of Friedrich-Schiller University, in Jena, Germany.

[0100] Plasmid construction: The original ePathBrick expression vectors, #4, #5, and #6 (Table 2) were modified through two rounds of site directed mutagenesis with primers 1 through 4 (Table 3) to result in the corresponding ‘SDM2x’ series of vectors: #7, #8, and #9 (Table 2). This mutagenesis was performed to swap the positions of the isocaudomer restriction enzyme pair *Xma*II/*Xba*I in the vector. This change allows for the monocistronic and pseudooperon pathway configurations to be constructed more cost efficiently by avoiding the use of the costly *Xma*II restriction enzyme. This series of vectors was then used to construct the vectors used in the defined copy number library study #10 - #27 (Table 2).

[0101] Plasmids #1 - #3 containing psiD, psiK, and psiM, respectively, were restriction enzyme digested with *Nde*I and *Hind*III, gel extracted, and ligated into the pETM6-SDM2x (#7, Table 2) plasmid backbone, resulting in plasmids #10, #11, and #12 (Table 2). All multigene expression plasmids were constructed in pseudooperon configuration using a modified version of the previously published ePathBrick methods as described above, while all transcriptional libraries were constructed using standard ePathOptimize methods.

[0102] Standard screening conditions: Standard screening was performed in 2 mL working volume cultures in 48-well plates at 37 °C. AMM supplemented with serine (1 g/L), 4-hydroxyindole (350 mg/L), and appropriate antibiotics were used unless otherwise noted. Overnight cultures were grown from either an agar plate or freezer stock culture in AMM with appropriate antibiotics and supplements for 14-16 hours in a shaking 37 °C incubator. Induction with 1 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) occurred four hours after inoculation, unless otherwise noted. Cultures were then sampled 24 hours post inoculation and subjected to HPLC analysis as described in analytical methods below.

[0103] Library construction: The defined copy number library was constructed using plasmid #7 (High), #8 (Medium), and #9 (Low). The pathway genes were modulated in either the high, medium, and low copy number vectors, as shown in FIG. 2A. The BL21 starTM (DE3) production host was transformed with the appropriate plasmids such that each strain had all three vectors, even if some were empty, to enable the same antibiotic resistance burden to be present in all defined library members (FIG. 1A). In the cases where multiple genes were present at a

single expression level, the plasmids were constructed in pseudooperon configuration as described above.

[0104] Random promoter libraries were assembled using standard ePathOptimize methods with the five original mutant T7 promoters: G6, H9, H10, C4, and consensus. Random libraries were built in pseudooperon (FIG. 1B) and basic operon (FIG. 1C) forms, maintaining a sufficient number of colonies at each cloning step as to not limit library size.

[0105] *Fermentation Optimization:* Once a genetically superior production strain, pPsilo16 (#28, Table 2) was identified, fermentation conditions were optimized to further enhance psilocybin production. The effect of varying induction timing was first investigated under standard screening conditions, then further evaluated under other conditions that have been shown to affect cellular growth rate and subsequently optimal induction timing including: 1. base media identity (AMM, LB), 2. media carbon source (glucose, glycerol), 3. production temperature (30 °C, 37 °C, 40 °C, 42 °C), 4. inducer concentration (1 mM, 0.5 mM, 0.1 mM), 5. concentration of media supplements: serine and methionine (0 g/L, 1 g/L, 5 g/L), and 6. concentration of 4-hydroxyindole substrate (150 mg/L, 350 mg/L, 500 mg/L). All screening was completed in 48-well plates under standard screening conditions unless otherwise noted.

[0106] *Scale-up study:* In order to demonstrate the scalability of our selected production host and process, a scale-up study was performed in an Eppendorf BioFlo120 bioreactor with 1.5 L working volume. The cylindrical vessel was mixed by a direct drive shaft containing two Rushton-type impellers positioned equidistance under the liquid surface. The overnight culture of pPsilo16 was grown for 14 hours at 37 °C in AMM supplemented with serine (5 g/L), methionine (5 g/L), and appropriate antibiotics. The bioreactor was inoculated at 2% v/v to an initial OD₆₀₀ of approximately 0.09. The bioreactor was initially filled with AMM media (1.5 L) supplemented with 150 mg/L 4-hydroxyindole, 5 g/L serine, and 5 g/L methionine. Temperature was held constant at 37 °C with a heat jacket and recirculating cooling water, pH was automatically controlled at 6.5 with the addition of 10 M NaOH, and dissolved oxygen (DO) was maintained at 20% of saturation through agitation cascade control (250 - 1000 rpm). Full oxygen saturation was defined under the conditions of 37 °C, pH 7.0, 250 rpm agitation, and 3 lpm of standard air. The zero-oxygen set point was achieved by a nitrogen gas flush. Samples were

collected periodically for measurement of OD₆₀₀ and metabolite analysis. The bioreactor was induced with 1 mM IPTG 4 hours post inoculation. Once the initial 20 g/L of glucose was exhausted, as identified by a DO spike, separate feed streams of 500 g/L glucose and 90 g/L (NH₄)₂HPO₄ were fed at a flow rate ranging from 2.0 to 4.0 mL/L/hr (FIG. 17B). Beginning 12 hours post inoculation, a continuous supply of 4-hydroxyindole was supplied by external syringe pump to the bioreactor. The feed rate of 4-hydroxyindole was manually varied from 11 to 53 mg/L/hr according to the observed buildup of the key pathway intermediate 4-hydroxytryptophan (FIG. 17C). The concentration of psilocybin and all intermediate compounds were immediately analyzed via HPLC on an approximate 45-minute delay and were used as feedback into the feeding strategy described above.

[0107] Analytical Methods: Samples were prepared by adding an equal volume of 100% ethanol or 100% deionized water and fermentation broth, vortexed briefly, and then centrifuged at 12000 x g for 10 minutes. 2 µL of the resulting supernatant was then injected for HPLC or LC-MS analysis. Analysis was performed on a Thermo Scientific Ultimate 3000 High-Performance Liquid Chromatography (HPLC) system equipped with Diode Array Detector (DAD) and Refractive Index Detector (RID). Authentic standards were purchased for glucose (Sigma), psilocybin (Cerilliant), and 4-hydroxyindole (BioSynth). Standards for baeocystin, norbaeocystin, 4-hydroxytryptamine, and 4-hydroxytryptophan were quantified using a standard for a similar analog due to limited commercial availability and extremely high cost, approx. \$2000 USD for 1 mg of the authentic standard. Baeocystin and norbaeocystin were quantified on the psilocybin standard curve, while 4-hydroxytryptamine and 4-hydroxytryptophan were quantified on the standard curves of 5-hydroxytryptamine (Alfa Aesar, Haverhill Massachusetts) and 5-hydroxytryptophan (Alfa Aesar, Haverhill Massachusetts), respectively (FIG. 7). No significant intracellular accumulation of target metabolites was observed upon analysis with and without cell lysis. Transport across the cell membrane was assumed to be passive, however, specific investigation into this phenomenon was not undertaken for this work.

[0108] Glucose analysis was performed using an Aminex HPX-87H column maintained at 30 °C followed by a refractive index detector (RID) held at 35 °C. The mobile phase was 5 mM H₂SO₄ in water at a flow rate of 0.6 mL/min. Glucose was quantified using a standard curve with a retention time of 8.8 min.

[0109] UV absorbance at 280 nm was used to quantify all aromatic compounds. Analysis was performed using an Agilent ZORBAX Eclipse XDB-C18 analytical column (3.0 mm x 250 mm, 5 μ m) with mobile phases of acetonitrile (A) and water (B) both containing 0.1% formic acid at a flow rate of 1 mL/min: 0 min, 5% A; 0.43 min, 5% A; 5.15 min, 19% A; 6.44 min, 100 % A; 7.73 min 100% A; 7.73 min, 5% A; 9.87 min, 5% A. This method resulted in the following observed retention times: psilocybin (2.2 min), baeocystin (1.9 min), norbaeocystin (1.7 min), 4-hydroxytryptamine (3.4 min), 4-hydroxytryptophan (3.6 min), and 4-hydroxyindole (6.6 min). High Resolution Liquid Chromatography Mass Spectrometry (LC-MS) and Mass Spectrometry-Mass Spectrometry (LC-MS/MS) data were measured on a Thermo Scientific LTQ Orbitrap XL mass spectrometer equipped with an Ion Max ESI source using the same mobile phases and column described above. The flow rate was adjusted to 0.250 mL/min resulting in a method with the following gradient: 0 min, 5% A; 1 min, 5% A; 24 min, 19% A; 30 min, 100 % A; 36 min 100% A; 36 min, 5% A; 46 min, 5% A. This method resulted in the following observed retention times: psilocybin (8.7 min), baeocystin (7.6 min), norbaeocystin (6.4 min), 4-hydroxytryptamine (13.3 min), 4-hydroxytryptophan (14.2 min), and 4-hydroxyindole (27 min). The Orbitrap was operated in positive mode using direct infusion from a syringe at 5 μ L/min for optimization of tuning parameters and for external calibration. A 5-hydroxytryptamine sample was prepared at ~0.1 mg/ml (570 μ M) in 50% ethanol / 50% water for tuning. External calibration was performed using the Pierce LTQ ESI Positive Ion Calibration Solution, allowing for a less than 5 ppm mass accuracy.

[0110] Mass spectrometry parameters in positive mode were spray voltage 3.5 kV, capillary temperature 275 $^{\circ}$ C, capillary voltage 23 V and tube lens voltage 80 V (optimized by tuning on 5-hydroxytryptamine), nitrogen sheath, auxiliary, and sweep gas were 15, 30, 1 a.u., full scan mode (m/z 100-500) at a resolution of 60,000 and an AGC target of 1e6.

[0111] LC-MS/MS data was collected in the data-dependent acquisition mode, where the full MS scan was followed by fragmentation of the three most abundant peaks by higher energy collisional dissociation (HCD). Data was collected in the Orbitrap with a minimum m/z of 50 at 30,000 resolution, AGC target of 1e5, and intensity threshold of 200K using normalized collision energy of 40, default charge state of 1, activation time of 30 ms, and maximum injection times of

200 msec for both MS and MS/MS scans. All data were processed using Xcalibur/Qual Browser 2.1.0 SP1 build (Thermo Scientific). MS/MS fragmentation data can be found in FIG. 9.

Results

[0112] Psilocybin production genes (*psiD*, *psiK*, and *psiM*) from *P. cubensis* were heterologously expressed in *E. coli* using the strong T7 promoter system. Induction with IPTG allowed for the production of 2.19 ± 0.02 mg/L psilocybin. To confirm compound identities, culture media from the psilocybin production host was subjected to liquid chromatography-mass spectroscopy analysis on a Thermo Orbitrap XL LC-MS system. Psilocybin, as well as all precursor and intermediate compounds in the biosynthetic pathway, were identified with better than 5 ppm mass accuracy. The sample was then subjected to additional MS/MS fragmentation analysis to further support structural identification of all indole derived intermediates and final products. In each case, fragmentation products for the deamination, dephosphorylation (if applicable), and loss of both functional groups were observed, confirming the identification of psilocybin, and its intermediates: 4-hydroxytryptophan, 4-hydroxytryptamine, norbaeocystin, and baeocystin, with better than 5 ppm mass accuracy. Additionally, expected retention times and order of elution were consistent with previously published efforts. The overexpression of the native tryptophan synthase (TrpAB) was also performed in an attempt to push flux through the heterologous production pathway. The native expression level was determined to be sufficient to maintain the necessary pathway flux, as supported by the buildup of 4-hydroxytryptophan in nearly all fermentation studies performed.

[0113] *Defined Copy Number Library*: A defined 27-member copy number library consisting of the 3 heterologous biosynthesis genes (*psiD*, *psiK*, and *psiM*) each expressed on 3 different copy number plasmids was constructed and screened in 48-well plates as shown in FIG. 2A. Each member of the library contained each of the three genes spread across a low (pACM4-SDM2X), medium (pCDM4-SDM2x), or high (pETM6-SDM2x) copy number plasmid (FIG.1A). This library screen realized minor improvements over the original All-High construct (2.19 ± 0.02 mg/L), where final titers of 4.0 ± 0.2 mg/L were achieved with the combination of *psiK* expressed from the pETM6-SDM2x vector, *psiD* expressed from the pCDM4-SDM2x vector, and *psiM* expressed from the pACM4-SDM2X vector in the BL21 starTM (DE3) expression host.

[0114] FIGs. 2A-2D show a summary of genetic strategies for increasing production.

[0115] *Pseudooperon Library*: The pseudooperon library were constructed having a different mutant promoter in front of each of the three enzyme encoding sequences, psiD, psiK, and psiM, while having a single terminator at the end of the 3-gene synthetic operon (FIG. 1B). This configuration resulted in a widely variable transcriptional landscape in which each promoter resulted in a distinct mRNA capable of encoding translation of either 1, 2, or 3 of the pathway enzymes. In this configuration, a possible library size of 125 pathway configurations existed, and 231 random colonies were screened. The large majority of variants demonstrated low (30%) or no production (65%); however, a small population of mutants demonstrated significant improvements in production (FIG. 2B) over the previous defined library screen (FIG. 2A). Additional analysis of the HPLC data revealed a significant accumulation of the intermediate, 4-hydroxytryptophan, suggesting that poor functional activity from the PsiD enzyme led to the underperformance of the majority of the members in the pseudooperon library.

[0116] *Basic Operon Library*: In the operon configuration, the three-gene pathway was expressed from a single high-copy plasmid under the control of a single promoter and terminator where each gene has an identical ribosome binding site (RBS) (FIG. 1C). The promoter sequence was randomized to one of five mutant T7 promoters (G6, H9, H10, C4, Consensus) using the ePathOptimize approach, resulting in a library that contained 5 potential promoter combinations (FIG. 2C). After screening nearly 50X the library size, the top 10 variants were selected for further screening. These top variants were re-cloned into an empty plasmid backbone and transformed to eliminate the possibility of spurious plasmid or strain mutations (FIG. 2D). Mutant #16 (pPsilo16) was selected for further investigation due to its top production and high reproducibility across multiple fermentations. The sequencing results revealed that pPsilo16 contained the H10 mutant promoter which has been previously characterized as a medium strength promoter, with between 40% and 70% of the effective expression strength of the consensus T7 sequence. The top mutants from the basic operon screen showed a 17-fold improvement in titer over the best performing mutants from the defined copy number library study.

[0117] *Fermentation Conditions:* After identifying pPsilo16 as the best strain with respect to the highest psilocybin production, low buildup of intermediate products, and consistent reproducibility, the strain underwent a series of experiments to determine the best fermentation conditions for the production of psilocybin. All genetic optimization experiments were conducted under standard conditions (as described in the Materials and Methods) determined from initial screening. Many studies in the metabolic engineering literature have demonstrated high sensitivity to variations in induction point for pathways controlled by the T7-lac inducible promoter. Additionally, induction timing can have a large impact on overall cell growth and can lead to difficulties achieving reproducible production upon scale-up. Upon evaluation of induction sensitivity for pPsilo16, it was found that the cells demonstrate low sensitivity to induction point, with the maximum production achieved with induction 3 to 4 hours post inoculation (FIG.3A). No psilocybin production was observed in the non-induced controls.

[0118] FIGs. 3A-3C show a summary of fermentation conditions screening studies.

[0119] Next, base media, carbon source identity, and inducer concentration was evaluated. Since these variables can affect cellular growth rate and corresponding optimal induction points, each of these variables was evaluated across a range of induction points from 1 to 6 hours. As demonstrated in FIG. 3B, psilocybin production was very sensitive to both media and carbon source selection ($p < 0.05$). When production was attempted in a rich undefined media such as LB, a dark colored insoluble product was observed along with low psilocybin production. Similarly, low production was also observed when grown on glycerol, however no colored products were observed. pPsilo16 demonstrated moderate sensitivity to IPTG concentration, with higher final concentrations of 0.5 and 1.0 mM over a range of induction time conditions ($p < 0.05$) (FIG. 3B). This trend is likely influenced by the initial library screening, which was performed at 1.0 mM IPTG.

[0120] Production temperatures of 30 °C, 37 °C, 40 °C, and 42 °C were also evaluated for their effect on psilocybin production (FIG. 3A). In an attempt to minimize the effect on changing induction points, all fermentations were started at 37 °C through the growth phase of the fermentation before being shifted to the production temperature 1 hour prior to induction. A

significant preference ($p < 0.05$) was seen for maintaining an isothermal fermentation temperature of 37 °C throughout both growth and production phases (FIG. 3A).

[0121] The fermentation screening was completed by evaluating the effects of the targeted media supplements: 4-hydroxyindole, serine, and methionine (FIG. 3C). Each media supplement was provided at high, medium, and low levels: 4-hydroxyindole (150, 350, and 500 mg/L), serine and methionine (0, 1, and 5 g/L). At high concentrations of 4-hydroxyindole, the cells demonstrated noticeable growth decline due to presumed cellular toxicity leading to reduced productivity. Serine addition showed minimal effects on psilocybin production, however, the addition of methionine in the presence of greater than 350 mg/L of 4-hydroxyindole resulted in a significant enhancement of psilocybin titer ($p < 0.05$). Under the identified optimal screening conditions, psilocybin was produced at 139 ± 2.7 mg/L, which represents a 63-fold improvement through the synergistic efforts of genetic and fermentation optimization.

[0122] *Scale-up*: After identification of preferred production conditions for pPsilo16 strain, a fed-batch scale up study was completed as described in the Materials and Methods. This study resulted in the production of 1.16 g/L of psilocybin which represents an 8.3-fold improvement over the top conditions screening case in 48-well plates and a 528-fold improvement over the original construct. Precursor and intermediate product titers remained low throughout the fermentation enabling the culture to achieve a final OD₆₀₀ of 35 (FIG. 4B). Pathway intermediate concentrations were maintained at a low level through the use of a HPLC informed feeding strategy which enabled the substrate feed rate to be tailored to specific pathway bottleneck flux within 45 min of sampling. This led to an oscillatory concentration profile for the key pathway intermediate 4-hydroxytryptophan (FIG. 4B). The initial 20 g/L of glucose was completely consumed from the media after 25 hours and was then externally fed such that the culture maintained robust growth with low residual sugar content in the media to maximize product yield on glucose.

[0123] The production of psilocybin and all pathway intermediates were confirmed through the use of high mass accuracy LC-MS (FIG. 9). HPLC analysis of fermentation broth from strains containing incomplete pathways (i.e. psiDM and psiDK) was consistent with the conclusions of

previous studies aimed at identifying the order of specific biosynthetic steps in the synthesis pathway.

[0124] Multiple genetic screening methods were utilized in parallel to identify a genetically superior mutant. Starting with the copy-number based approach, a 27-member library of 3 pathway genes, each at 3 discrete copy numbers (FIG. 2A) was constructed. Of the three genetic optimization screens presented, this method was the most tedious to construct, requiring each plasmid to be independently cloned and verified prior to screening. This defined library approach also yielded the lowest product titers with the best mutants demonstrating small but statistically significant ($p < 0.05$) improvements over the All-High initial construct. Without wishing to be bound by theory, the limited titer improvement from this approach may be due to the increased metabolic burden associated with selection for and propagation of three independent plasmids.

[0125] Subsequent screening of two independent single-plasmid transcriptionally-varied promoter libraries with pathway genes in basic operon (FIG. 2C) and pseudooperon (FIG. 2B) configuration yielded considerably improved results over the initial copy number library. In each case, the library was screened using a medium-throughput HPLC-based screen. Each of these transcriptionally varied libraries were constructed using the high copy pETM6 plasmid vector. This enabled a wide range of expression levels to be screened, resulting in greater coverage of the psilocybin transcriptional landscape. The pseudooperon library screen demonstrated that a large majority of mutants (~95%) showed low or no psilocybin production. The reason for this widespread underperformance is unknown; however, it does motivate the use of random libraries coupled with variant screening for the identification of genetically superior mutants as the current predictive power of *a priori* pathway design is still lacking for most applications. Surprisingly, the simplistic basic-operon pathway design yielded the highest titer psilocybin production in this study. This coupled with the smallest library size of only 5 mutants, enabled rapid screening of several times the theoretical library size, resulting in high confidence of complete coverage of the full transcriptional landscape. Upon recloning and rescreening the top mutants from the operon library screen, several false positives were identified as shown in FIG. 2D. The source of error for these false positive mutants was not investigated as the false positive rate was at an acceptable level for the study design.

[0126] Additional increases in titer and yield were achieved through careful optimization of fermentation conditions (FIGs. 3A-3C). The genetically superior strain, pPsilo16, demonstrated low sensitivity to induction timing as compared to that of other amino acid derived high-value products; however, this could also be due to the supplementation of both 4-hydroxyindole and serine to the fermentation media, reducing the requirement for high flux through amino acid metabolism. Therefore, all additional fermentation optimization experiments were performed under a range of induction times. Little variation from the induction optimum of 4 hours post inoculation was observed, strengthening the observation of reduced sensitivity to induction timing.

[0127] The psilocybin production host demonstrated high sensitivity to media composition, carbon source identity, fermentation temperature, and inducer concentration (FIGs. 3A-3B). In each case, this preferred level was similar to that of the standard screening conditions. This is likely not a coincidence, as some basic initial screening was performed to identify conditions under which our proof-of-principle strain best performed. Furthermore, the initial genetic screening studies were performed under standard screening conditions, which also self-selects for mutants with top performance under the test conditions.

[0128] The largest gains in the fermentation optimization aspect of this study were achieved through the media supplementation studies (FIG. 3C). In this study, the concentrations of 4-hydroxyindole, serine, and methionine were varied. These supplements were selected specifically for their direct effect on the psilocybin production pathway (FIG. 1D). 4-hydroxyindole and serine are condensed by TrpAB in the first dedicated step of the pathway to form the intermediate 4-hydroxytryptophan. Although *E. coli* can produce serine and indole naturally, it lacks the ability to express the P450 hydroxylase that oxidizes indole into 4-hydroxyindole. Additionally, with the high fluxes through our engineered pathway, it was hypothesized that the cellular supply of serine would be quickly depleted, requiring additional supplementation to not limit pathway flux. Finally, methionine was supplemented to enhance intercellular pools of the activated methyl donor, SAM. The final two biosynthetic steps are both catalyzed by the SAM-dependent methyltransferase, PsiM. Previous studies with SAM-dependent methylations in *E. coli* have documented SAM-limited flux to final products.

[0129] Analysis of intermediate product concentrations was performed to evaluate the success of each study. A comparison is presented (FIG. 4A) between the initial proof-of-principle 'All-High' strain (Stage 1) and the top production strain, both post genetic optimization (Stage 2) and post genetic and fermentation optimization (Stage 3). Each additional optimization stage resulted in further enhanced psilocybin titers, accomplished through a reduction in intermediate product concentrations, and generally enhanced flux towards the final product.

[0130] The information gained from the genetic and fermentation optimization studies was applied in a scale-up study for the production of psilocybin in a fed-batch bioreactor. In this study, many of the optimization parameters such as temperature, inducer concentration, and induction timing were applied as previously optimized. Information from the supplement addition studies was used but applied with modification from the 2 mL batch studies. In the fed-batch studies, both serine and methionine were supplemented at the high level of 5 g/L to account for higher cellular demand due to enhanced cell growth. Furthermore, in the small-scale studies a growth deficit was observed at higher concentrations of 4-hydroxyindole and 4-hydroxytryptophan. To counter this, a low amount of 4-hydroxyindole (150 mg/L) was added initially to the media, while a low-flow syringe pump, containing a 40 mg/mL 4-hydroxyindole solution, was connected for slow external supplementation. To determine the optimal feed rate, the pathway flux through the bottleneck point, PsiD, was estimated through frequent HPLC analysis of the fermentation broth. As 4-hydroxytryptophan titers fell, the flux of 4-hydroxyindole was increased to meet the high flux demand, and vice-versa. This strategy resulted in an oscillatory concentration profile for 4-hydroxytryptophan and maintained all intermediates at low levels, enabling robust and extended growth and psilocybin production (FIG. 4B).

[0131] In small batch fermentation studies, the work presented above resulted in a similar titer of psilocybin to that presented previously in the *A. nidulans* host. This indicates that both bacterial and fungal hosts show potential as production platforms for this important chemical. However, upon scale-up to a fed batch reactor our bacterial host demonstrated greatly enhanced psilocybin production resulting in a 10-fold enhancement over previously published results.

[0132] Provided is the first example of effective psilocybin production in a prokaryotic organism and the highest psilocybin titer to date from a recombinant host from any kingdom. This was accomplished through the combination of increased genetic and fermentation production in small scale, coupled with a scaled-up fed-batch study utilizing a unique HPLC informed substrate feeding strategy. The fed-batch study resulted in a psilocybin titer of 1.16 g/L with maximum and final molar yields from the 4-hydroxyindole substrate of 0.60 and 0.38 mol/mol, respectively (FIG. 17D).

Example 2: Production of Norbaeocystin

Materials and Methods

[0133] A transcriptional library comprised of five IPTG-inducible T7 promoter mutants of varied strength (G6, H9, H10, C4, and consensus) were used to construct two independently pooled libraries capable of norbaeocystin production: pETM6-xx5-psiDK (operon form, 5 member) and pETM6-xx5-psiD-xx5-psiDK (pseudooperon form, 25 members). These libraries were constructed using standard molecular cloning and ePathOptimize techniques analogous to those used for the construction of the psilocybin production plasmid libraries discussed above. The plasmid DNA libraries were then transformed into the production host strain BL21 StarTM (DE3) and screened in a medium throughput fermentation assay in 48-well plates. Andrew's Magic Media (AMM) supplemented with 20 g/L glucose, 350 mg/L of 4-hydroxyindole, and 1 g/L of serine was used as the microbial growth media and the fermentation screening and HPLC sample preparation was performed as described elsewhere herein. Andrew's Magic Media (AMM) is rich semi-defined media containing: 3.5 g/L KH₂PO₄, 5.0 g/L K₂HPO₄, 3.5g/L (NH₄)₂HPO₄, 2g/L casamino acids, 100mL of 10x MOPS Mix, 1mL of 1M MgSO₄, 0.1mL of 1M CaCl₂, 1mL of 0.5g/L thiamine HCL, supplemented with 20g/L glucose). 10x MOPS Mix consisted of 83.72g/L MOPS, 7.17g/L Tricine, 28mg/L FeSO₄·7H₂O, 29.2g/L NaCl, 5.1g/L NH₄Cl, 1.1g/L MgCl₂, 0.48g/L K₂SO₄, 0.2mL Micronutrient Stock. Micronutrient Stock consisted of 0.18g/L (NH₄)₆Mo₇O₂₄, 1.24g/L H₃BO₃, 0.12g/L CuSO₄, 0.8g/L MnCl₂, 0.14g/L ZnSO₄.

[0134] All norbaeocystin titers were quantified using a psilocybin standard curve due to the lack of a commercially available analytical standard.

[0135] Upon identification of top mutants from both the operon and pseudooperon libraries, the plasmids were purified, retransformed in the plasmid storage strain, DH5 α . A single DH5 α colony was grown overnight, plasmid was purified, retransformed into BL21 StarTM (DE3) for additional screening, and sequenced to identify the mutant promoters controlling transcription of the exogenous pathway genes, psiD and psiK. The retransformed production strains were subjected to additional screening identical to that of the initial screen and with an additional 350 mg/L of 4-hydroxyindole added approximately 24 hours after inoculation. Final samples for HPLC analysis were taken 48 hours post inoculation.

Results

[0136] The initial screening resulted in a range of production levels in both the operon and pseudooperon libraries. 47 random mutants from the operon and 143 random mutants from pseudooperon library were screened. This represents 9.4x and 5.7x their respective library sizes. The top mutants from both libraries demonstrated complete consumption of the 4-hydroxyindole, no endpoint buildup of the 4-hydroxytryptophan, and produced approximately 400 mg/L of norbaeocystin (FIG. 5). This is a significant observation as the production of norbaeocystin in the top mutants is roughly 400% higher than the production of psilocybin from the optimized pPsilo16 mutant under similar conditions. Without wishing to be bound by theory, this indicates that regeneration of the methyl donor, SAM, is likely limiting in the psilocybin production case and supports the need for further studies targeted at alleviating this bottleneck.

[0137] Seven mutants from this initial screen at a variety of production levels were selected for additional testing and sequencing (Table 1). The sequencing results revealed an interesting trend of the top producing strains having the exogenous pathway controlled by the strong mutant promoter, C4, in both top producing mutants deriving from the operon and pseudooperon libraries. The data also supports a trend of reduced strength promoters leading to reduced norbaeocystin production. This is in contrast with the similarly performed psilocybin production work which resulted in the best performance from the medium strength, H10, mutant promoter.

Additionally, production of another tryptophan derived compound, violacein, found weakened promoters to produce significantly more product than strong promoters. Taken together, this data supports the discovery of a non-obvious and interesting solution for the biological production of norbaeocystin.

[0138] Table 1

Strain #	Original Library	Plasmid Name	Norbaeocystin Titer	Promoter(s) Strength
O-H1	Operon	pETM6-C4-psiDK	415 mg/L	High
P3-D4	Pseudooperon	pETM6-C4-psiD-C4-psiK	398 mg/L	High
O-B1	Operon	pETM6-H9-psiDK	192 mg/L	Medium/Low
P2-E1	Pseudooperon	pETM6-T7-psiD-H10-psiK	152 mg/L	Medium/High
O-F1	Operon	pETM6-G6-psiDK	32 mg/L	Low

[0139] The select mutants were additionally screened after plasmid retransformation to confirm their norbaeocystin production capability. Additionally, all selected mutants were also given additional 4-hydroxyindole to further evaluate their production in a non-substrate limited environment (FIG. 6, right). Upon rescreening, the mutants maintained their high titer production with the top mutant, O-H1, showing production just under 400 mg/L. Adding an additional 350 mg/L of the 4-hydroxyindole substrate approximately 24 hours after inoculation resulted in a significant ($p < 0.05$) enhancement in overall titer for the best producing mutant, O-H1, to over 0.5 g/L of norbaeocystin in a 48-well plate fermentation assay.

Example 3: Optimization of Production of Norbaeocystin

Materials and Methods

[0140] The top norbaeocystin production strain identified from the library screens, O-H1 (Table 1), was subjected to scaleup screening in a 1.5-L working volume bioreactor controlled by the Eppendorf BioFLO120 system. This bioreactor system was operated as described above for the psilocybin scale up study (Example 1).

[0141] Norbaeocystin was quantified as described above using a psilocybin standard curve due to the lack of a commercially available analytical standard. Norbaeocystin identity was verified using an accurate mass OrbitrapXL spectrometer (FIG. 19). The measured mass resulted in an acceptable error of 6.2 ppm.

Results

[0142] The concentration of psilocybin and other key intermediates were tracked over the course of the fed-batch bioreactor study. The results of this HPLC analysis are shown in FIG. 18. The figure shows that the intermediate product, 4-hydroxytryptophan, and the 4-hydroxyindole substrate were maintained below inhibitory levels throughout the fermentation. This was achieved by using an HPLC-informed feeding strategy coupled with frequent sampling and analysis. This study resulted in the production of 700 mg/L of norbaeocystin over 32 hours. This is the first reported example of norbaeocystin production from a prokaryotic host.

[0143] Table 2: Plasmid and Strain List

Number	Strain or vector	Relevant properties	Reference
S1	<i>Escherichia coli</i> DH5a	F-, ϕ 80d lacZ Δ M15, Δ (lacZYA-argF)U169, recA1, endZ1, hsdR17(rk ⁻ , mk ⁺), phoA, supE44 λ ⁻ , thi ⁻¹ , gyrA96, relA1	Novagen
S2	<i>E. coli</i> BL21 Star TM (DE3)	F- ompT gal dcm rne131 lon hsdS _B (rB-mB-) λ (DE3)	Invitrogen

1	pJF24	pET28a containing tryptophan decarboxylase from <i>Psilocybe cubensis</i> (PcPsiD), Kan ^R	(Fricke et al., 2017)
2	pJF23	pET28a containing Kinase from <i>Psilocybe cubensis</i> (PcPsiK), Kan ^R	(Fricke et al., 2017)
3	pFB13	pET28a containing SAM-dependant methyl transferase from <i>Psilocybe cubensis</i> (PcPsiM), Kan ^R	(Fricke et al., 2017)
4	pETM6	ColE1(pBR322), Amp ^R	(Xu et al., 2012)
5	pACM4	P15A(pACYC184), Cm ^R	(Xu et al., 2012)
6	pCDM4	CloDF13, Str ^R	(Xu et al., 2012)
7	pETM6-SDM2x	#4 with XbaI and XmaJI sites switched	This Study
8	pACM4-SDM2x	#5 with XbaI and XmaJI sites switched	This Study
9	pCDM4-SDM2x	#6 with XbaI and XmaJI sites switched	This Study
10	pETM6-SDM2x-psiD	#7 containing psiD	This Study
11	pETM6-SDM2x-psiK	#7 containing psiK	This Study
12	pETM6-SDM2x-psiM	#7 containing psiM	This Study
13	pETM6-SDM2x-psiDK	#7 containing psiDK in pseudooperon configuration	This Study

14	pETM6-SDM2x-psiKM	#7 containing psiKM in pseudooperon configuration	This Study
15	pETM6-SDM2x-psiDKM (All-High)	#7 containing psiDKM in pseudooperon	This Study
16	pACM4-SDM2x-psiD	#8 containing psiD	This Study
17	pACM4-SDM2x-psiK	#8 containing psiK	This Study
18	pACM-SDM2x-psiM	#8 containing psiM	This Study
19	pACM4-SDM2x-psiDK	#8 containing psiDK in pseudooperon configuration	This Study
20	pACM4-SDM2x-psiKM	#8 containing psiKM in pseudooperon configuration	This Study
21	pACM4-SDM2x-psiDKM	#8 containing psiDKM in pseudooperon configuration	This Study
22	pCDM4-SDM2x-psiD	#9 containing psiD	This Study
23	pCDM4-SDM2x-psiK	#9 containing psiK	This Study
24	pCDM4-SDM2x-psiM	#9 containing psiM	This Study

25	pCDM4-SDM2x-psiDK	#9 containing psiDK in pseudooperon configuration	This Study
26	pCDM4-SDM2x-psiKM	#9 containing psiKM in pseudooperon configuration	This Study
27	pCDM4-SDM2x-psiDKM	#9 containing psiDKM in pseudooperon configuration	This Study
28	pPsilo16	pETM6-SDM2x-psiDKM in basic operon configuration with T7 mutant promoter H10 (TAATACGACTCACTACGGAAGAA [SEQ ID NO: 11]) sequence in front of psiD controlling expression of all three genes in basic operon configuration	This Study
29	pETM6-G6-mCherry	#4 containing the mCherry reporter under control of the G6 mutant T7 promoter	(Jones et al., 2015)
30	pETM6-H9-mCherry	#4 containing the mCherry reporter under control of the H9 mutant T7 promoter	(Jones et al., 2015)
31	pETM6-H10-mCherry	#4 containing the mCherry reporter under control of the H10 mutant T7 promoter	(Jones et al., 2015)

32	pETM6-mCherry	#4 containing the mCherry reporter under control of the consensus T7 promoter	(Xu et al., 2012)
33	pETM6-C4-mCherry	#4 containing the mCherry reporter under control of the C4 mutant T7 promoter	(Jones et al., 2015)

[0144] Bibliography

[0145] Fricke, J., Blei, F., Hoffmeister, D., 2017. Enzymatic synthesis of psilocybin. *Angew. Chemie Int. Ed.* 56, 12352-12355. (doi.org/10.1002/anie.201705489)

[0146] Jones, J. Andrew, Vernacchio, V.R., Lachance, D.M., Lebovich, M., Fu, L., Shirke, A.N., Schultz, V.L., Cress, B., Linhardt, R.J., Koffas, M.A.G., 2015. ePathOptimize: a combinatorial approach for transcriptional balancing of metabolic pathways. *Sci. Rep.* 5, 11301 (doi.org/10.1038/srep11301)

[0147] Xu, P., Vansiri, A., Bhan, N., Koffas, M.A.G., 2012. ePathBrick: A synthetic biology platform for engineering metabolic pathways in *E. coli*. *Biol* 1, 256–266. (doi.org/10.1021/sb300016b)

[0148] Table 3: Sequences

SEQ ID NO:	Description	Sequence
1	SDM_XbaI-AvrII-FWD	GAATTGTGAGCGGATAACAATTCCCCCTAGGAATAATTTG TTTAACTTTAAGAAG

	Primer 1	(5'-3)
2	SDM_XbaI-AvrII-REV	CTTCTTAAAGTTAAACAAAATTATTCCTAGGGGGAATTGTT ATCCGCTCACAATTC
	Primer 2	(5'-3)
3	SDM_AvrII-XbaI_FWD	CCGGCCACGATGCGTCCGGCGTAGTCTAGAATCGAGATCGA TCTCGATCCCCG
	Primer 3	(5'-3)
4	SDM_AvrII-XbaI_REV	CGGGATCGAGATCGATCTCGATTCTAGACTACGCCGGACGC ATCGTGGCCGG
	Primer 4	(5'-3)
5	PsiD (Genbank KY984101.1)	ATGCAGGTGATACCCGCGTGCAACTCGGCAGCAATAAGATC ACTATGTCCTACTCCCGAGTCTTTTAGAAACATGGGATGGCT CTCTGTCAGCGATGCGGTCTACAGCGAGTTCATAGGAGAGTT GGCTACCCGCGCTTCCAATCGAAATTAATCCAACGAGTTCGG CCTCATGCAACCTATCCAGGAATTCAAGGCTTTCATTGAAAG CGACCCGGTGGTGCACCAAGAATTTATTGACATGTTTCGAGG GCATTCAAGGACTCTCCAAGGAATTATCAGGAACTATGTAATA TGTTCAACGATATCTTTCGCAAAGCTCCCGTCTACGGAGACC TTGGCCCTCCCGTTTATATGATTATGGCCAAATTAATGAACA CCCGAGCGGGCTTCTCTGCATTACGAGACAAAGGTTGAAC CTTCACTTCAAAAACTTTTCGATACCTGGGGATTGTTCCCTG TCTTCGAAAGATTCTCGAAATGTTCTTGTGGCCGACCAGTTC GACGACAGACATTGCGGCTGGTTGAACGAGCGGGCCTTGTC TGCTATGGTTAAACATTACAATGGACGCGCATTTGATGAAGT CTTCCTCTGCGATAAAAATGCCCCATACTACGGCTTCAACTC TTACGACGACTTCTTTAATCGCAGATTTTCGAAACCGAGATAT

		CGACCGACCTGTAGTCGGTGGAGTTAACAACACCACCCTCAT TTCTGCTGCTTGCGAATCACTTTCCTACAACGTCTCTTATGAC GTCCAGTCTCTCGACACTTTAGTTTTCAAAGGAGAGACTTAT TCGCTTAAGCATTTGCTGAATAATGACCCTTTCACCCACAA TTCGAGCATGGGAGTATTCTACAAGGATTCTTGAACGTCACC GCTTACCACCGATGGCACGCACCCGTCAATGGGACAATCGT CAAAATCATCAACGTTCCAGGTACCTACTTTGCGCAAGCCCC GAGCACGATTGGCGACCCTATCCCGGATAACGATTACGACC CACCTCCTTACCTTAAGTCTCTTGTCTACTTCTCTAATATTGC CGCAAGGCAAATTATGTTTATTGAAGCCGACAACAAGGAAA TTGGCCTCATTTTCCTTGTGTTCATCGGCATGACCGAAATCTC GACATGTGAAGCCACGGTGTCCGAAGGTCAACACGTCAATC GTGGCGATGACTTGGAATGTTCCATTTCCGGTGGTTCTTCGT TCGCGCTTGGTCTGAGGAAGGATTGCAGGGCAGAGATCGTT GAAAAGTTCACCGAACCCGGAACAGTGATCAGAATCAACGA AGTCGTCGCTGCTCTAAAGGCTTAG
6	PsiK (Genbank KY984099.1)	ATGGCGTTCGATCTCAAGACTGAAGACGGCCTCATCACATAT CTCACTAAACATCTTCTTTGGACGTCGACACGAGCGGAGTG AAGCGCCTTAGCGGAGGCTTTGTCAATGTAACCTGGCGCATT AAGCTCAATGCTCCTTATCAAGGTCATACGAGCATCATCCTG AAGCATGCTCAGCCGCACATGTCTACGGATGAGGATTTTAA GATAGGTGTAGAACGTTTCGGTTTACGAATACCAGGCTATCA AGCTCATGATGGCCAATCGGGAGGTTCTGGGAGGCGTGGAT GGCATAGTTTCTGTGCCAGAAGGCCTGAACTACGACTTAGA GAATAATGCATTGATCATGCAAGATGTCGGGAAGATGAAGA CCCTTTTAGATTATGTCACCGCCAAACCGCCACTTGCGACGG ATATAGCCCGCCTTGTTGGGACAGAAATTGGGGGGTTTCGTTG CCAGACTCCATAACATAGGCCGCGAGAGGCGAGACGATCCT GAGTTCAAATTCTTCTCTGGAAATATTGTCGGAAGGACGACT

		TCAGACCAGCTGTATCAAACCATCATACCCAACGCAGCGAA ATATGGCGTCGATGACCCCTTGCTGCCTACTGTGGTTAAGGA CCTTGTGGACGATGTCATGCACAGCGAAGAGACCCTTGTCAT GGCGGACCTGTGGAGTGGAAATATTCTTCTCCAGTTGGAGG AGGGAAACCCATCGAAGCTGCAGAAGATATATATCCTGGAT TGGGAACTTTGCAAGTACGGCCCAGCGTCGTTGGACCTGGG CTATTTCTTGGGTGACTGCTATTTGATATCCCGCTTTCAAGAC GAGCAGGTCCGTACGACGATGCGGCAAGCCTACTTGCAAAG CTATGCGCGTACGAGCAAGCATTGATCAACTACGCCAAAG TCACTGCAGGTATTGCTGCTCATATTGTGATGTGGACCGACT TTATGCAGTGGGGGAGCGAGGAAGAAAGGATAAAATTTGTG AAAAAGGGGGTAGCTGCCTTTCACGACGCCAGGGGCAACAA CGACAATGGGGAAATTACGTCTACCTTACTGAAGGAATCAT CCACTGCGTAA
7	PsiM (Genbank KY984100.1)	ATGCATATCAGAAATCCTTACCGTACACCAATTGACTATCAA GCACTTTCAGAGGCCTTCCCTCCCCTCAAGCCATTTGTGTCT GTCAATGCAGATGGTACCAGTTCTGTTGACCTCACTATCCCA GAAGCCCAGAGGGCGTTCACGGCCGCTCTTCTTCATCGTGAC TTCGGGCTCACCATGACCATAACCAGAAGACCGTCTGTGCCCA ACAGTCCCCAATAGGTTGAACTACGTTCTGTGGATTGAAGAT ATTTTCAACTACACGAACAAAACCCTCGGCCTGTCCGATGAC CGTCCTATTAAAGGCGTTGATATTGGTACAGGAGCCTCCGCA ATTTATCCTATGCTTGCCTGTGCTCGGTTCAAGGCATGGTCT ATGGTTGGAACAGAGGTCGAGAGGAAGTGCATTGACACGGC CCGCCTCAATGTCGTCGCGAACAATCTCCAAGACCGTCTCTC GATATTAGAGACATCCATTGATGGTCCTATTCTCGTCCCCAT TTTCGAGGCGACTGAAGAATACGAATACGAGTTTACTATGTG TAACCCTCCATTCTACGACGGTGCTGCCGATATGCAGACTTC GGATGCTGCCAAAGGATTTGGATTTGGCGTGGGCGCTCCCCA TTCTGGAACAGTCATCGAAATGTCGACTGAGGGAGGTGAAT

		CGGCTTTCGTCGCTCAGATGGTCCGTGAGAGCTTGAAGCTTC GAACACGATGCAGATGGTACACGAGTAACTTGGGAAAGCTG AAATCCTTGAAAGAAATAGTGGGGCTGCTGAAAGAACTTGA GATAAGCAACTATGCCATTAACGAATACGTTCAAGGGGTCCA CACGTCGTTATGCCGTTGCGTGGTCTTTCAGTATATTCAACT GCCTGAGGAGCTTCTCGTCCCTCTAACCCCGAGCTCAGCTC TCTTTTCTAG
8	PsiD amino acid sequence	MQVIPACNSAAIRSLCPTPESFRNMGWLSVSDAVYSEFIGELAT RASNRNYSNEFGLMQPIQEFKAFIESDPVVHQEFIDMFEGIQDSP RNYQELCNMFNDIFRKAPVYGDLGPPVYMIMAKLMNTRAGFS AFTRQRLNLHFKKLFDTWGLFLSSKDSRNVLVADQFDDRHC WLNERALSAMVKHYNGRAFDEVFLCDKNAPYYGFNSYDDFFN RRFRNRDIDRPVVGGVNNTTLISAACESLSYNVSVDVQSLDTLV FKGETYSLKHLLNNDPFTPQFEHGSILQGFLNVTAYHRWHAPV NGTIVKIINVPPTYFAQAPSTIGDPIPDNDYDPPPYLKSLVYFSNI AARQIMFIEADNKEIGLIFLVFIGMTEISTCEATVSEGQHVNRGD DLGMFHFHGGSSFALGLRKDCRAEIVEKFTEPGTVIRINEVVAAL KA
9	PsiK amino acid sequence	MAFDLKTEDGLITYLTKHLSLDVDTSGVKRLSGGFVNVTWRIK LNAPYQGHTSIILKHAQPHMSTDEDFKIGVERSVYEQAIKLM MANREVLGGVDGIVSVPEGLNYDLENNALIMQDVGKMKTLTD YVTAKPPLATDIARLVGTEIGGFVARLHNIGRERRDDPEFKFFS GNIVGRITSDQLYQTIIPNAAKYGVDDPLLPTVVKDLVDDVMH SEETLVMADLWSGNILLQLEEGNPSKLQKIYILDWELCKYGPAS LDLGYFLGDCYLISRQDEQVGTTMRQAYLQSYARTSKHSINY AKVTAGIAAHIVMWTDFMQWGSEERINFVKKGVAAAFHDARG NNDNGEITSTLLKESSTA
10	PsiM	MHIRNPYRTPIDYQALSEAFPPLKPFVSVNADGTSSVDLTIPEAQ RAFTAALLHRDFGLTMTIPEDRLCPTVPNRLNYVLWIEDIFNYT

	amino acid sequence	NKTLGLSDDRPIKGVDIGTGASAIYPMLACARFKAWSMVGTEV ERKCIDTARLNVVANNLQDRLSILETSIDGPILVPIFEATEEY EYE FTMCNPPFYDGAADMQTS DAAKGFGFGVGAPHSGTVIEMSTE GGESAFVAQMVRESLKLRTRCRWYTSNLGKLKSLKEIVGLLKE LEISNYAINEYVQGSTRRYAVAWSFTDIQLPEELSRPSNPELSSL F
11	H10 mutant T7 promoter	TAATACGACTCACTACGGAAGAA
12	G6 mutant T7 promoter	TAATACGACTCACTATTTTCGGAA
13	H9 mutant T7 promoter	TAATACGACTCACTAATACTGAA
14	C4 mutant T7 promoter	TAATACGACTCACTATCAAGGAA
15	consensus T7 promoter	TAATACGACTCACTATAGGGGAA
16	Lac promoter	TTTACACTTTATGCTTCCGGCTCGTATGTTG
17	Lac UV5 promoter	TTTACACTTTATGCTTCCGGCTCGTATAATG
18	tac promoter	TTGACAATTAATCATCGGCTCGTATAATG
19	trc promoter	TTGACAATTAATCATCCGGCTCGTATAATG

20	GAP promoter	GCGTAATGCTTAGGCACAGGATTGATTTGTCGCAATGATTGA CACGATTCCGCTTGACGCTGCGTAAGGTTTTTGTAATTTTAC AGGCAACCTTTTATTCA
21	xylA promoter	TTGAAATAAACATTTATTTTGTATATGATGAGATAAAGTTAG TTTATTGGATAAAACAACTAACTCAATTAAGATAGTTGATGG ATAAACTT
22	pPsilo16 vector	TAATACGACTCACTACGGAAGAATTGTGAGCGGATAACAAT TCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATA TACATATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGC GGATCCATGCAGGTGATACCCGCGTGCAACTCGGCAGCAAT AAGATCACTATGTCCTACTCCCGAGTCTTTTAGAAACATGGG ATGGCTCTCTGTCAGCGATGCGGTCTACAGCGAGTTCATAGG AGAGTTGGCTACCCGCGCTTCCAATCGAAATTACTCCAACGA GTTTCGGCCTCATGCAACCTATCCAGGAATTCAAGGCTTTCAT TGAAAGCGACCCGGTGGTGCACCAAGAATTTATTGACATGTT CGAGGGCATTCAAGACTCTCCAAGGAATTATCAGGAACATAT GTAATATGTTCAACGATATCTTTCGCAAAGCTCCCGTCTACG GAGACCTTGGCCCTCCCGTTTATATGATTATGGCCAAATTAA TGAACACCCGAGCGGGCTTCTCTGCATTACGAGACAAAGG TTGAACCTTCACTTCAAAAACTTTTCGATACCTGGGGATTG TTCCTGTCTTCGAAAGATTCTCGAAATGTTCTTGTGGCCGAC CAGTTCGACGACAGACATTGCGGCTGGTTGAACGAGCGGGC CTTGTCTGCTATGGTTAAACATTACAATGGACGCGCATTGA TGAAGTCTTCCTCTGCGATAAAAATGCCCCATACTACGGCTT CAACTCTTACGACGACTTCTTTAATCGCAGATTCGAAACCG AGATATCGACCGACCTGTAGTCGGTGGAGTTAACAACACCA CCCTCATTCTGCTGCTTGCGAATCACTTTCCTACAACGTCTC TTATGACGTCCAGTCTCTCGACACTTTAGTTTTCAAAGGAGA GACTTATTCGCTTAAGCATTTGCTGAATAATGACCCTTTCAC

	CCCACAATTCGAGCATGGGAGTATTCTACAAGGATTCTTGAA CGTCACCGCTTACCACCGATGGCACGCACCCGTCAATGGGA CAATCGTCAAAATCATCAACGTTCCAGGTACCTACTTTGCGC AAGCCCCGAGCACGATTGGCGACCCTATCCCGGATAACGAT TACGACCCACCTCCTTACCTTAAGTCTCTTGTCTACTTCTCTA ATATTGCCGCAAGGCAAATTATGTTTATTGAAGCCGACAACA AGGAAATTGGCCTCATTTTCCTTGTGTTTCATCGGCATGACCG AAATCTCGACATGTGAAGCCACGGTGTCCGAAGGTCAACAC GTCAATCGTGGCGATGACTTGGGAATGTTCCATTTCTGGTGGT TCTTCGTTTCGCGCTTGGTCTGAGGAAGGATTGCAGGGCAGAG ATCGTTGAAAAGTTCACCGAACCCGGAACAGTGATCAGAAT CAACGAAGTCGTCGCTGCTCTAAAGGCTTAGAAGCTTGCGG CCGCACTCGAGTCTGGTAAAGAAACCGCTGCTGCGAAATTT GAACGCCAGCACATGGACTCGTCTACTAGAAATAATTTTGT TAACTTTAAGAAGGAGATATACATATGGCTAGCATGACTGG TGGACAGCAAATGGGTCGCGGATCCATGGCGTTCGATCTCA AGACTGAAGACGGCCTCATCACATATCTCACTAAACATCTTT CTTTGGACGTCGACACGAGCGGAGTGAAGCGCCTTAGCGGA GGCTTTGTCAATGTAACCTGGCGCATTAAAGCTCAATGCTCCT TATCAAGGTCATACGAGCATCATCCTGAAGCATGCTCAGCCG CACATGTCTACGGATGAGGATTTTAAGATAGGTGTAGAACG TTCGGTTTACGAATACCAGGCTATCAAGCTCATGATGGCCAA TCGGGAGGTTCTGGGAGGCGTGGATGGCATAAGTTTCTGTGCC AGAAGGCCTGAACTACGACTTAGAGAATAATGCATTGATCA TGCAAGATGTGCGGAAGATGAAGACCCTTTTAGATTATGTCA CCGCCAAACCGCCACTTGCGACGGATATAGCCCGCCTTGTTG GGACAGAAATTGGGGGGTTCGTTGCCAGACTCCATAACATA GGCCGCGAGAGGCGAGACGATCCTGAGTTCAAATTCTTCTCT GGAAATATTGTCGGAAGGACGACTTCAGACCAGCTGTATCA AACCATCATACCCAACGCAGCGAAATATGGCGTCGATGACC
--	---

	<p>CCTTGCTGCCTACTGTGGTTAAGGACCTTGTGGACGATGTCA TGCACAGCGAAGAGACCCTTGTTCATGGCGGACCTGTGGAGT GGAAATATTCTTCTCCAGTTGGAGGAGGGAAACCCATCGAA GCTGCAGAAGATATATATCCTGGATTGGGAACTTTGCAAGTA CGGCCAGCGTCGTTGGACCTGGGCTATTTCTTGGGTGACTG CTATTTGATATCCCGCTTTCAAGACGAGCAGGTCGGTACGAC GATGCGGCAAGCCTACTTGCAAAGCTATGCGCGTACGAGCA AGCATTTCGATCAACTACGCCAAAGTCACTGCAGGTATTGCTG CTCATATTGTGATGTGGACCGACTTTATGCAGTGGGGGAGCG AGGAAGAAAGGATAAATTTTGTGAAAAAGGGGGTAGCTGCC TTTACGACGCCAGGGGCAACAACGACAATGGGGAAATTAC GTCTACCTTACTGAAGGAATCATCCACTGCGTAAAAGCTTGC GGCCGCACTCGAGTCTGGTAAAGAAACCGCTGCTGCGAAAT TTGAACGCCAGCACATGGACTCGTCTACTAGAAATAATTTTG TTAACTTTAAGAAGGAGATATACATATGGCTAGCATGACTG GTGGACAGCAAATGGGTCGCGGATCCATGCATATCAGAAAT CCTTACCGTACACCAATTGACTATCAAGCACTTTCAGAGGCC TTCCCTCCCCTCAAGCCATTTGTGTCTGTCAATGCAGATGGT ACCAGTTCTGTTGACCTCACTATCCCAGAAGCCCAGAGGGCG TTCACGGCCGCTCTTCTTCATCGTGACTIONCGGGCTCACCATG ACCATACCAGAAGACCGTCTGTGCCCAACAGTCCCCAATAG GTTGAACTACGTTCTGTGGATTGAAGATATTTTCAACTACAC GAACAAAACCCTCGGCCTGTCCGATGACCGTCCTATTAAAG GCGTTGATATTGGTACAGGAGCCTCCGCAATTTATCCTATGC TTGCCTGTGCTCGGTTCAAGGCATGGTCTATGGTTGGAACAG AGGTCGAGAGGAAGTGCATTGACACGGCCCGCCTCAATGTC GTCGCGAACAATCTCCAAGACCGTCTCTCGATATTAGAGACA TCCATTGATGGTCCTATTCTCGTCCCCATTTTCGAGGCGACTG AAGAATACGAATACGAGTTTACTATGTGTAACCCTCCATTCT ACGACGGTGCTGCCGATATGCAGACTTCGGATGCTGCCAAA</p>
--	---

	GGATTTGGATTTGGCGTGGGCGCTCCCCATTCTGGAACAGTC ATCGAAATGTCTGACTGAGGGAGGTGAATCGGCTTTCGTCGCT CAGATGGTCCGTGAGAGCTTGAAGCTTCGAACACGATGCAG ATGGTACACGAGTAACCTGGGAAAGCTGAAATCCTTGAAAG AAATAGTGGGGCTGCTGAAAGAACTTGAGATAAGCAACTAT GCCATTAACGAATACGTTTCAGGGGTCCACACGTCGTTATGCC GTTGCGTGGTCTTTCCTGATATTCAACTGCCTGAGGAGCTT TCTCGTCCCTCTAACCCCGAGCTCAGCTCTCTTTTCTAGCTCG AGTCTGGTAAAGAAACCGCTGCTGCGAAATTTGAACGCCAG CACATGGACTCGTCTACTAGTCGCAGCTTAATTAACCTAAAC TGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGC CTCTAAACGGGTCTTGAGGGGTTTTTGTAGCGAAAGGAGG AGTCGACTATATCCGGATTGGCGAATGGGACGCGCCCTGTA GCGGCGCATTAAAGCGCGGGGGTGTGGTGGTTACGCGCAGC GTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTC GCTTTCCTCCCTTCCTTTCCTCGCCACGTTGCGCGGCTTCCCC GTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTA GTGCTTTACGGCACCTCGACCCCAAAAACTTGATTAGGGTG ATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTC GCCCTTTGACGTTGGAGTCCACGTTCTTAATAGTGGACTCT TGTTCCAACTGGAACAACACTCAACCCTATCTCGGTCTATT CTTTTGATTTATAAGGGATTTTGCCGATTTTCGGCCTATTGGTT AAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTA ACAAAATATTAACGTTTACAATTTCTGGCGGCACGATGGCAT GAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTA AAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAA CTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTA TCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACT CCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATC TGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCAC
--	--

	CGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGG GCCGAGCGCAGAAAGTGGTCCTGCAACTTTATCCGCCTCCATC CAGTCTATTAATTGTTGCCGGAAGCTAGAGTAAGTAGTTCG CCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGC ATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTACG TCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCATG TTGTGCAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTT GTCAGAAGTAAGTTGGCCGAGTGTTATCACTCATGGTTATG GCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAAG TGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGA GAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCCGGCGTCA ATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGT GCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACCTCTCAAG GATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCG TGCACCCAACCTGATCTTCAGCATCTTTTACTTTACCCAGCGTT TCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAA AGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC TTCCTTTTTCAATCATGATTGAAGCATTATCAGGGTTATTGT CTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAA CAAATAGGTCATGACCAAAATCCCTTAACGTGAGTTTTCGTT CCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTT CTTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTTGCAAAC AAAAAAACCACCGCTACCAGCGGTGGTTTGGTTGCCGGATC AAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCA GAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGT TAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACC TCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCG ATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTAC CGGATAAGGCGCAGCGGTCTGGGCTGAACGGGGGGTTCGTGC ACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAG
--	---

	ATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCG AAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGT CGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAAC GCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGAC TTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCC TATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGG CCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATC CCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGC TGATAACCGCTCGCCGACGCCGAACGACCGAGCGCAGCGAGT CAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTT CTCCTTACGCATCTGTGCGGTATTTACACCCGCATATATGGT GCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCC AGTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGC GCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGG CTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCG TCTCCGGGAGCTGCATGTGTCAGAGGTTTTACCGTCATCAC CGAAACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGTGG TCGTGAAGCGATTACAGATGTCTGCCTGTTTCATCCGCGTCC AGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTCTGGCTTCTG ATAAAGCGGGCCATGTTAAGGGCGGTTTTTTCCTGTTTGGTC ACTGATGCCTCCGTGTAAGGGGGATTTCTGTTCATGGGGGTA ATGATAACCGATGAAACGAGAGAGGATGCTCACGATACGGGT TACTGATGATGAACATGCCCCGTTACTGGAACGTTGTGAGG GTAAACAACCTGGCGGTATGGATGCGGCGGGACCAGAGAAAA ATCACTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTA GGTGTTCCACAGGGTAGCCAGCAGCATCCTGCGATGCAGAT CCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTTTCCAG ACTTTACGAAACACGGAAACCGAAGACCATTTCATGTTGTTGC TCAGGTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTTTCG CTCGCGTATCGGTGATTTCATTCTGCTAACCAGTAAGGCAACC
--	--

	CCGCCAGCCTAGCCGGGTCCTCAACGACAGGAGCACGATCA TGCTAGTCATGCCCCGCGCCACCGGAAGGAGCTGACTGGG TTGAAGGCTCTCAAGGGCATCGGTCGAGATCCCGGTGCCTA ATGAGTGAGCTAACTTACATTAATTGCGTTGCGCTCACTGCC CGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATG AATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGC GCCAGGGTGGTTTTTCTTTTACCAGTGAGACGGGCAACAGC TGATTGCCCTTACC GCCTGGCCCTGAGAGAGTTGCAGCAAG CGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTG ATGGTGGTTAACGGCGGGATATAACATGAGCTGTCTTCGGTA TCGTCGTATCCCACTACCGAGATGTCCGCACCAACGCGCAGC CCGGA CT CGGTAATGGCGCGCATTGCGCCCAGCGCCATCTG ATCGTTGGCAACCAGCATCGCAGTGGGAACGATGCCCTCATT CAGCATTTGCATGGTTTGTTGAAAACCGGACATGGCACTCCA GTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGATTGCGAGT GAGATATTTATGCCAGCCAGCCAGACGCGAGACGCGCCGAGA CAGAACTTAATGGGCCCCGCTAACAGCGCGATTTGCTGGTGA CCCAATGCGACCAGATGCTCCACGCCCAGTCGCGTACCGTCT TCATGGGAGAAAATAACTGTTGATGGGTGTCTGGTCAGA GACATCAAGAAATAACGCCGGAACATTAGTGCAGGCAGCTT CCACAGCAATGGCATCCTGGTCATCCAGCGGATAGTTAATG ATCAGCCC ACTGACGCGTTGCGCGAGAAGATTGTGCACCGC CGCTTTACAGGCTTCGACGCCGCTTCGTTCTACCATCGACAC CACCACGCTGGCACCCAGTTGATCGGCGCGAGATTTAATCGC CGCGACAATTTGCGACGGCGCGTGCAGGGCCAGACTGGAGG TGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTT GTGCCACGCGGTTGGGAATGTAATTCAGCTCCGCCATCGCCG CTTCCACTTTTCCCGCGTTTTTCGCAGAAACGTGGCTGGCCT GGTTCACCACGCGGGAAACGGTCTGATAAGAGACACCGGCA TACTCTGCGACATCGTATAACGTTACTGGTTTTACATTACCC
--	--

		ACCCTGAATTGACTCTCTTCCGGGCGCTATCATGCCATACCG CGAAAGGTTTTGCGCCATTCGATGGTGTCCGGGATCTCGACG CTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCCAGTAG TAGGTTGAGGCCGTTGAGCACCGCCGCCGCAAGGAATGGTG CATGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCACGGGG CCTGCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCC GAAGTGGCGAGCCCGATCTTCCCCATCGGTGATGTCGGCGAT ATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGG CCACGATGCGTCCGGCGTAGCCTAGGATCGAGATCGATCTC GATCCCGCGAAAT
23	pETM6-C4- psiDK	TAATACGACTCACTATCAAGGAATTGTGAGCGGATAACAAT TCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATA TACATATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGC GGATCCATGCAGGTGATACCCGCGTGCAACTCGGCAGCAAT AAGATCACTATGTCCTACTCCCGAGTCTTTTAGAAACATGGG ATGGCTCTCTGTCAGCGATGCGGTCTACAGCGAGTTCATAGG AGAGTTGGCTACCCGCGCTTCCAATCGAAATTACTCCAACGA GTTCGGCCTCATGCAACCTATCCAGGAATTCAAGGCTTTCAT TGAAAGCGACCCGGTGGTGCACCAAGAATTTATTGACATGTT CGAGGGCATTCAAGACTCTCCAAGGAATTATCAGGAACTAT GTAATATGTTCAACGATATCTTTCGCAAAGCTCCCGTCTACG GAGACCTTGGCCCTCCCGTTTATATGATTATGGCCAAATTAA TGAACACCCGAGCGGGCTTCTCTGCATTACGAGACAAAGG TTGAACCTTCACTTCAAAAACTTTTCGATACCTGGGGATTG TTCCTGTCTTCGAAAGATTCTCGAAATGTTCTTGTGGCCGAC CAGTTCGACGACAGACATTGCGGCTGGTTGAACGAGCGGGC CTTGTCTGCTATGGTTAAACATTACAATGGACGCGCATTTGA TGAAGTCTTCCTCTGCGATAAAAAATGCCCCATACTACGGCTT CAACTCTTACGACGACTTCTTTAATCGCAGATTCGAAACCG AGATATCGACCGACCTGTAGTCGGTGGAGTTAACAACACCA

	CCCTCATTCTGCTGCTTGCGAATCACTTTCCTACAACGTCTC TTATGACGTCCAGTCTCTCGACACTTTAGTTTTCAAAGGAGA GACTTATTCGCTTAAGCATTGCTGAATAATGACCCTTTCAC CCCACAATTCGAGCATGGGAGTATTCTACAAGGATTCTTGAA CGTCACCGCTTACCACCGATGGCACGCACCCGTCAATGGGA CAATCGTCAAAATCATCAACGTTCCAGGTACCTACTTTGCGC AAGCCCCGAGCACGATTGGCGACCCTATCCCGGATAACGAT TACGACCCACCTCCTTACCTTAAGTCTCTTGTCTACTTCTCTA ATATTGCCGCAAGGCAAATTATGTTTATTGAAGCCGACAACA AGGAAATTGGCCTCATTTTCCTTGTGTTTCATCGGCATGACCG AAATCTCGACATGTGAAGCCACGGTGTCCGAAGGTCAACAC GTCAATCGTGGCGATGACTTGGGAATGTTCCATTTCCGGTGGT TCTTCGTTTCGCGCTTGGTCTGAGGAAGGATTGCAGGGCAGAG ATCGTTGAAAAGTTCACCGAACCCGGAACAGTGATCAGAAT CAACGAAGTCGTCGCTGCTCTAAAGGCTTAGAAGCTTGCGG CCGCACTCGAGTCTGGTAAAGAAACCGCTGCTGCGAAATTT GAACGCCAGCACATGGACTCGTCTACTAGAAATAATTTTGTT TAACTTTAAGAAGGAGATATACATATGGCTAGCATGACTGG TGGACAGCAAATGGGTCGCGGATCCATGGCGTTCGATCTCA AGACTGAAGACGGCCTCATCACATATCTCACTAAACATCTTT CTTTGGACGTTCGACACGAGCGGAGTGAAGCGCCTTAGCGGA GGCTTTGTCAATGTAACCTGGCGCATTAAAGCTCAATGCTCCT TATCAAGGTCATACGAGCATCATCCTGAAGCATGCTCAGCCG CACATGTCTACGGATGAGGATTTTAAGATAGGTGTAGAACG TTCGGTTTACGAATACCAGGCTATCAAGCTCATGATGGCCAA TCGGGAGGTTCTGGGAGGCGTGGATGGCATAGTTTCTGTGCC AGAAGGCCTGAACTACGACTTAGAGAATAATGCATTGATCA TGCAAGATGTCGGGAAGATGAAGACCCTTTTAGATTATGTCA CCGCCAAACCGCCACTTGCGACGGATATAGCCCGCCTTGTTG GGACAGAAATTGGGGGGTTCGTTGCCAGACTCCATAACATA
--	---

	<p>GGCCGCGAGAGGCGAGACGATCCTGAGTTCAAATTCTTCTCT GGAAATATTGTCGGAAGGACGACTTCAGACCAGCTGTATCA AACCATCATACCCAACGCAGCGAAATATGGCGTCGATGACC CCTTGCTGCCTACTGTGGTTAAGGACCTTGTGGACGATGTCA TGCACAGCGAAGAGACCCTTGTTCATGGCGGACCTGTGGAGT GGAAATATTCTTCTCCAGTTGGAGGAGGGAAACCCATCGAA GCTGCAGAAGATATATATCCTGGATTGGGAACCTTTCGAAGTA CGGCCCAGCGTCGTTGGACCTGGGCTATTTCTTGGGTGACTG CTATTTGATATCCCGCTTTCAAGACGAGCAGGTCGGTACGAC GATGCGGCAAGCCTACTTGCAAAGCTATGCGCGTACGAGCA AGCATTGATCAACTACGCCAAAGTCACTGCAGGTATTGCTG CTCATATTGTGATGTGGACCGACTTTATGCAGTGGGGGAGCG AGGAAGAAAGGATAAAATTTGTGAAAAAGGGGGTAGCTGCC TTTCACGACGCCAGGGGCAACAACGACAATGGGGAAATTAC GTCTACCTTACTGAAGGAATCATCCACTGCGTAAAAGCTTGC GGCCGCACTCGAGTCTGGTAAAGAAACCGCTGCTGCGAAAT TTGAACGCCAGCACATGGACTCGTCTACTAGTCGCAGCTTAA TTAACCTAAACTGCTGCCACCGCTGAGCAATAACTAGCATAA CCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTTGCTA GCGAAAGGAGGAGTCGACTATATCCGGATTGGCGAATGGGA CGCGCCCTGTAGCGGCGCATTAAAGCGCGGCGGGTGTGGTGG TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGC CCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTTCGC CGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGG GTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAACT TGATTAGGGTGATGGTTACGTAGTGGGCCATCGCCCTGATA GACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAAT AGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATC TCGGTCTATTCTTTTGATTATAAGGGATTTTGCCGATTTTCGG CCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACG</p>
--	---

	CGAATTTTAACAAAATATTAACGTTTACAATTTCTGGCGGCA CGATGGCATGAGATTATCAAAAAGGATCTTCACCTAGATCCT TTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATA TGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGA GGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTT GCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGG CTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAG CCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCC GCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTA AGTAGTTCGCCAGTTAATAGTTTTCGCAACGTTGTTGCCATT GCTACAGGCATCGTGGTGTACGCTCGTCGTTTGGTATGGCT TCATTACAGCTCCGGTTCCTAACGATCAAGGCGAGTTACATGA TCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCT CCGATCGTTGTGAGAAGTAAGTTGGCCGAGTGTTATCACTC ATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCA TCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGC CCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAAC TTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAA ACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGTA ACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTC ACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGC CGCAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATA CTCATACTCTTCCTTTTTCAATCATGATTGAAGCATTATCAG GGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAG AAAAATAACAAATAGGTCATGACCAAAATCCCTTAACGTG AGTTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCA AAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTG CTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGT
--	--

	<p>GCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAGTGG CTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTA GCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCC TACATACCTCGCTCTGCTAATCCTGTTACCACTGGCTGCTGC CAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGAC GATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGG GGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACAC CGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCA CGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGC GGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAG GGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCC ACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGG GGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTA CGGTTCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTC CTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTT TGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGC GCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGAT GCGGTATTTTCTCCTTACGCATCTGTGCGGTATTTACACCGC ATATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCAT AGTTAAGCCAGTATACTCCGCTATCGCTACGTGACTGGGT CATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGC CCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAG CTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTAC CGTCATACCGAAACGCGCGAGGCAGCTGCGGTAAAGCTCA TCAGCGTGGTCGTGAAGCGATTACAGATGTCTGCCTGTTCA TCCGCGTCCAGCTCGTTGAGTTTCTCCAGAAGCGTTAATGTC TGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTCC TGTTTGGTCACTGATGCCTCCGTGTAAGGGGGATTTCTGTTC ATGGGGGTAATGATACCGATGAAACGAGAGAGGATGCTCAC GATACGGGTTACTGATGATGAACATGCCCGGTTACTGGAAC</p>
--	--

	GTTGTGAGGGTAAACAACCTGGCGGTATGGATGCGGCGGGAC CAGAGAAAAATCACTCAGGGTCAATGCCAGCGCTTCGTAA TACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTG CGATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTC CGCGTTTCCAGACTTTACGAAACACGGAAACCGAAGACCAT TCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGTC GCTTCACGTTTCGCTCGCGTATCGGTGATTCTCTGCTAACC AGTAAGGCAACCCCGCCAGCCTAGCCGGGTCTCAACGACA GGAGCACGATCATGCTAGTCATGCCCCGCGCCACCGGAAG GAGCTGACTGGGTGAAGGCTCTCAAGGGCATCGGTGAGA TCCCGGTGCCTAATGAGTGAGCTAACTTACATTAATTGCGTT GCGCTCACTGCCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCA GCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTT TGCGTATTGGGCGCCAGGGTGGTTTTCTTTTACCAGTGAG ACGGGCAACAGCTGATTGCCCTTCACCGCCTGGCCCTGAGA GAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGC GAAAATCCTGTTTGATGGTGGTTAACGGCGGGATATAACAT GAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGATGTCC GCACCAACGCGCAGCCCGGACTCGGTAATGGCGCGCATTGC GCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGG GAACGATGCCCTCATTAGCATTTGCATGGTTTGTTGAAAAC CGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCT GAATTTGATTGCGAGTGAGATATTTATGCCAGCCAGCCAGAC GCAGACGCGCCGAGACAGAACTTAATGGGCCCGCTAACAGC GCGATTTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCC AGTCGCGTACCGTCTTCATGGGAGAAAATAATACTGTTGATG GGTGTCTGGTCAGAGACATCAAGAAATAACGCCGGAACATT AGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCA GCGGATAGTTAATGATCAGCCCACTGACGCGTTGCGCGAGA AGATTGTGCACCGCCGCTTTACAGGCTTCGACGCCGCTTCGT
--	---

		TCTACCATCGACACCACCACGCTGGCACCCAGTTGATCGGCG CGAGATTTAATCGCCGCGACAATTTGCGACGGCGCGTGCGAG GGCCAGACTGGAGGTGGCAACGCCAATCAGCAACGACTGTT TGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAATGTAATTCA GCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTTCGCAGA AACGTGGCTGGCCTGGTTCACCACGCGGGAAACGGTCTGAT AAGAGACACCGGCATACTCTGCGACATCGTATAACGTTACT GGTTTCACATTCACCACCCTGAATTGACTCTCTTCCGGGCGC TATCATGCCATACCGCGAAAGGTTTTGCGCCATTCGATGGTG TCCGGGATCTCGACGCTCTCCCTTATGCGACTCCTGCATTAG GAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCG CCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGT CCCCCGGCCACGGGGCCTGCCACCATACCCACGCCGAAACA AGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCCATC GGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGG CGCCGGTGATGCCGGCCACGATGCGTCCGGCGTAGCCTAGG ATCGAGATCGATCTCGATCCCGCGAAAT
--	--	--

* * *

All publications and patents referred to herein are incorporated by reference. Various modifications and variations of the described subject matter will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific embodiments, it should be understood that the invention as claimed should not be unduly limited to these embodiments. Indeed, various modifications for carrying out the invention are obvious to those skilled in the art and are intended to be within the scope of the following claims.