

(12) STANDARD PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. **AU 2008335089 B2**

(54) Title
Secretion of fatty acids by photosynthetic microorganisms

(51) International Patent Classification(s)
C12P 7/64 (2006.01)

(21) Application No: **2008335089**

(22) Date of Filing: **2008.12.11**

(87) WIPO No: **WO09/076559**

(30) Priority Data

(31) Number
61/007,333

(32) Date
2007.12.11

(33) Country
US

(43) Publication Date: **2009.06.18**

(44) Accepted Journal Date: **2014.06.26**

(71) Applicant(s)
Synthetic Genomics, Inc.

(72) Inventor(s)
Liu, Bo;Chen, You;Roessler, Paul Gordon;Dodge, Corey Neal

(74) Agent / Attorney
FB Rice, L 14 90 Collins ST, Melbourne, VIC, 3000

(56) Related Art
WO 2007/136762 A2 (LS9, INC.) 29 November 2007
CHO, H. ET AL. "Defective export of a periplasmic enzyme disrupts regulation of fatty acid synthesis", The Journal of Biological Chemistry, 1995, 270, 4216-4219

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
18 June 2009 (18.06.2009)

PCT



(10) International Publication Number
WO 2009/076559 A8

(51) International Patent Classification:
C12P 7/64 (2006.01)

(21) International Application Number:
PCT/US2008/086485

(22) International Filing Date:
11 December 2008 (11.12.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/007,333 11 December 2007 (11.12.2007) US

(71) Applicant (for all designated States except US): SYN-
THETIC GENOMICS, INC. [US/US]; 11149 N. Torrey
Pines Rd, Suite 100, La Jolla, CA 92037 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): ROESSLER, Paul,
Gordon [US/US]; 11149 N. Torrey Pines Rd, Suite 100,
La Jolla, CA 92037 (US). CHEN, You [CN/US]; 11149
N. Torrey Pines Rd, Suite 100, La Jolla, CA 92037 (US).
LIU, Bo [CN/US]; 11149 N. Torrey Pines Rd, Suite 100,
La Jolla, CA 92037 (US). DODGE, Corey, Neal
[US/US]; 11149 N. Torrey Pines Rd, Suite 100, La Jolla,
CA 92037 (US).

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,

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

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

Published:

— with international search report (Art. 21(3))

(48) Date of publication of this corrected version:
20 May 2010

(15) Information about Correction:
see Notice of 20 May 2010

(54) Title: SECRETION OF FATTY ACIDS BY PHOTOSYNTHETIC MICROORGANISMS

(57) Abstract: Recombinant photosynthetic microorganisms that convert inorganic carbon to secreted fatty acids are described. Methods to recover the secreted fatty acids from the culture medium without the need for cell harvesting are also described.



WO 2009/076559 A8

SECRETION OF FATTY ACIDS BY PHOTOSYNTHETIC MICROORGANISMS

Cross-Reference to Related Applications

[0001] This application claims benefit of provisional application 61/007,333 filed 11 December 2007. The contents of this application are incorporated herein by reference.

Technical Field

[0002] This invention relates to photosynthetic microorganisms that convert inorganic carbon to fatty acids and secrete them into the culture medium, methods of production of fatty acids using such organisms, and uses thereof. The fatty acids may be used directly or may be further modified to alternate forms such as esters, reduced forms such as alcohols, or hydrocarbons, for applications in different industries, including fuels and chemicals.

Background Art

[0003] Photosynthetic microorganisms, including eukaryotic algae and cyanobacteria, contain various lipids, including polar lipids and neutral lipids. Polar lipids (*e.g.*, phospholipids, glycolipids, sulfolipids) are typically present in structural membranes whereas neutral lipids (*e.g.*, triacylglycerols, wax esters) accumulate in cytoplasmic oil bodies or oil globules. A substantial research effort has been devoted to the development of methods to produce lipid-based fuels and chemicals from photosynthetic microorganisms. Typically, eukaryotic microalgae are grown under nutrient-replete conditions until a certain cell density is achieved, after which the cells are subjected to growth under nutrient-deficient conditions, which often leads to the accumulation of neutral lipids. The cells are then harvested by various means (*e.g.*, settling, which can be facilitated by the addition of flocculants, followed by centrifugation), dried, and then the lipids are extracted from the cells by the use of various non-polar solvents. Harvesting of the cells and extraction of the lipids are cost-intensive steps. It would be desirable to obtain lipids from photosynthetic microorganisms without the requirement for cell harvesting and extraction.

[0004] PCT publication numbers WO2007/136762 and WO2008/119082 describe the production of biofuel components using microorganisms. These documents disclose the production by these organisms of fatty acid derivatives which are, apparently, short and long chain alcohols, hydrocarbons, fatty alcohols and esters including waxes, fatty acid esters or

fatty esters. To the extent that fatty acid production is described, it is proposed as an intermediate to these derivatives, and the fatty acids are therefore not secreted. Further, there is no disclosure of converting inorganic carbon directly to secreted fatty acids using a photosynthetic organism grown in a culture medium containing inorganic carbon as the primary carbon source. The present invention takes advantage of the efficiency of photosynthetic organisms in secreting fatty acids into the medium in order to recover these valuable compounds.

[0005] The invention includes the expression of heterologous acyl-ACP thioesterase (TE) genes in photosynthetic microbes. Many of these genes, along with their use to alter lipid metabolism in oilseeds, have been described previously. Genes encoding the proteins that catalyze various steps in the synthesis and further metabolism of fatty acids have also been extensively described.

[0006] The two functional classes of plant acyl-ACP thioesterases (unsaturated fatty acid-recognizing Fat A versus saturated fatty acid-recognizing FatB) can be clustered based on amino acid sequence alignments as well as function. FatAs show marked preference for 18:1-ACP with minor activity towards 18:0- and 16:0-ACPs, and FatBs hydrolyze primarily saturated acyl-ACPs with chain lengths that vary between 8–16 carbons. Several studies have focused on engineering plant thioesterases with perfected or altered substrate specificities as a strategy for tailoring specialty seed oils.

[0007] As shown in Figure 1, fatty acid synthetase catalyzes a repeating cycle wherein malonyl-acyl carrier protein (ACP) is condensed with a substrate, initially acetyl-CoA, to form acetoacetyl-ACP, liberating CO₂. The acetoacetyl-ACP is then reduced, dehydrated, and reduced further to butyryl-ACP which can then itself be condensed with malonyl-ACP, and the cycle repeated, adding a 2-carbon unit at each turn. The production of free fatty acids would therefore be enhanced by a thioesterase that would liberate the fatty acid itself from ACP, breaking the cycle. That is, the acyl-ACP is prevented from reentering the cycle. Production of the fatty acid would also be encouraged by enhancing the levels of fatty acid synthetase and inhibiting any enzymes which result in degradation or further metabolism of the fatty acid.

[0008] Figure 2 presents a more detailed description of the sequential formation of acyl-ACPs of longer and longer chains. As shown, the thioesterase enzymes listed in Figure 2 liberate the fatty acid from the ACP thioester.

[0009] Taking advantage of this principle, Dehesh, K., *et al.*, *The Plant Journal* (1996) 9:167-172, describe “Production of high levels of octanoic (8:0) and decanoic (10:0) fatty

acids in transgenic canola by overexpression of ChFatB2, a thioesterase cDNA from *Cuphea hookeriana*." Dehesh, K., *et al*, *Plant Physiology* (1996) 110:203-210, and report "Two novel thioesterases are key determinants of the bimodal distribution of acyl chain length of *Cuphea palustris* seed oil."

[0010] Voelker, T., *et al*, *Science* (1992) 251:12-1'4, describe "Fatty acid biosynthesis redirected to medium chains in transgenic oilseed plants." Voelker, T., and Davies, M., *Journal of Bacteriology* (1994) 176:7320-7327, describe "Alteration of the specificity and regulation of fatty acid synthesis of *Escherichia coli* by expression of a plant medium-chain acyl-acyl carrier protein thioesterase."

Any discussion of documents, acts, materials, devices, articles or the like which has been included in the present specification is solely for the purpose of providing a context for the present invention. It is not to be taken as an admission that any or all of these matters form part of the prior art base or were common general knowledge in the field relevant to the present invention as it existed before the priority date of each claim of this application.

Disclosure of the Invention

[0011] The present invention is directed to the production of recombinant photosynthetic microorganisms that are able to secrete fatty acids derived from inorganic carbon into the culture medium. Methods to remove the secreted fatty acids from the culture medium without the need for cell harvesting are also provided. It is anticipated that these improvements will lead to lower costs for producing lipid-based fuels and chemicals from photosynthetic microorganisms. In addition, this invention enables the production of fatty acids of defined chain length, thus allowing their use in the formulation of a variety of different products, including fuels and chemicals.

[0012] Carbon dioxide (which, along with carbonic acid, bicarbonate and/or carbonate define the term "inorganic carbon") is converted in the photosynthetic process to organic compounds. The inorganic carbon source includes any way of delivering inorganic carbon, optionally in admixture with any other combination of compounds which do not serve as the primary carbon feedstock, but only as a mixture or carrier (for example, emissions from biofuel (*e.g.*, ethanol) plants, power plants, petroleum-based refineries, as well as atmospheric and subterranean sources).

[0013] One embodiment of the invention relates to a culture of recombinant photosynthetic microorganisms, said organisms comprising at least one recombinant expression vector encoding at least one exogenous acyl-ACP thioesterase, wherein the at least one exogenous acyl-ACP thioesterase preferentially liberates fatty acid chains

containing 6 to 20 carbons from these ACP thioesters. The fatty acids are formed from inorganic carbon as their carbon source and the culture contains substantially only inorganic carbon as a carbon source. The presence of the exogenous thioesterase will increase the secretion levels of desired fatty acids by at least 2-4 fold.

[0013A] In one embodiment, the invention provides a cell culture of a recombinant cyanobacterium modified to contain a nucleic acid molecule comprising at least one recombinant expression system that produces at least one exogenous acyl-ACP thioesterase,

wherein said acyl-ACP thioesterase preferentially liberates a fatty acid chain that contains 6-20 carbons, and

wherein the culture medium provides inorganic carbon as substantially the sole carbon source and

wherein said cyanobacterium secretes the fatty acid liberated by the acyl-ACP thioesterase into the culture medium.

[0014] Specifically, in one embodiment, the invention is directed to a cell culture of a recombinant photo synthetic microorganism where the microorganism has been modified to contain a nucleic acid molecule comprising at least one recombinant expression system that produces at least one exogenous acyl-ACP thioesterase, wherein said acyl-ACP thioesterase preferentially liberates a fatty acid chain that contains 6-20 carbons, and wherein the culture medium provides inorganic carbon as substantially the sole carbon source and wherein said microorganism secretes the fatty acid liberated by the acyl-ACP thioesterase into the medium. In alternative embodiments, the thioesterase preferentially liberates a fatty acid chain that contains 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 carbons.

[0015] In other aspects, the invention is directed to a method to produce fatty acids of desired chain lengths by incubating these cultures and recovering these secreted fatty acids from the cultures. In one embodiment, the recovery employs solid particulate adsorbents to harvest the secreted fatty acids. The fatty acids thus recovered can be further modified synthetically or used directly as components of biofuels or chemicals.

[0015A] In one embodiment, the invention provides a method to convert inorganic carbon to fatty acids, said method comprising:

incubating the culture of any one of claims 1 to 10 such that the recombinant cyanobacterium therein secretes the fatty acid into the culture medium; and

recovering the secreted fatty acids from the culture medium.

Accordingly, the present invention provides a cell culture of a recombinant photosynthetic microorganism, said microorganism modified to contain a nucleic acid

molecule comprising at least one recombinant expression system that produces at least one exogenous acyl-ACP thioesterase, wherein said acyl-ACP thioesterase preferentially liberates a fatty acid chain that contains 6-20 carbons, and wherein the culture medium provides inorganic carbon as substantially the sole carbon source and wherein said microorganism secretes the fatty acid liberated by the acyl-ACP thioesterase into the culture medium.

The present invention also provides a method to convert inorganic carbon to fatty acids, said method comprising: incubating the culture of the invention such that the recombinant photosynthetic microorganism therein secretes the fatty acid into the culture medium; and recovering the secreted fatty acids from the culture medium.

Throughout this specification the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps.

Brief Description of the Drawings

[0016] Figure 1 is a diagram of the pathway of fatty acid synthesis as is known in the art.

[0017] Figure 2 is a more detailed diagram of the synthesis of fatty acids of multiple chain lengths as is known in the art.

[0018] Figure 3 is an enzymatic overview of fatty acid biosynthesis identifying enzymatic classes for the production of various chain length fatty acids.

[0019] Figure 4 is a schematic diagram of a recovery system for fatty acids from the medium.

[0020] Figure 5 shows an experimental system based on the principles in Figure 4.

[0021] Figure 6 shows representative acyl-ACP thioesterase from a variety of organisms.

Modes of Carrying Out the Invention

[0022] The present invention provides photosynthetic microorganisms that secrete fatty acids into the culture medium, along with methods to adsorb the fatty acids from the culture medium and collect them for processing into fuels and chemicals. The invention thereby

eliminates or greatly reduces the need to harvest and extract the cells, resulting in substantially reduced production costs.

[0023] Figure 2 is an overview of one aspect of the invention. As shown in Figure 2, carbon dioxide is converted to acetyl-CoA using the multiple steps in the photosynthetic process. The acetyl-CoA is then converted to malonyl-CoA by the action of acetyl-CoA carboxylase. The malonyl-CoA is then converted to malonyl-ACP by the action of malonyl-CoA:ACP transacylase which, upon progressive action of fatty acid synthetase, results in successive additions of two carbon units. In one embodiment of the invention, the process is essentially halted at carbon chain lengths of 6 or 8 or 10 or 12 or 14 or 16 or 18 carbons by supplying the appropriate thioesterase (shown in Figure 2 as FatB). To the extent that further conversions to longer chain fatty acids occur in this embodiment, the cell biomass can be harvested as well. The secreted fatty acids can be converted to various other forms including, for example, methyl esters, alkanes, alkenes, alpha-olefins and fatty alcohols.

Thioesterases (Acyl-ACP TEs)

[0024] In order to effect secretion of the free fatty acids, the organism is provided at least one expression system for at least one thioesterase that operates preferentially to liberate fatty acids of the desired length. Many genes encoding such thioesterases are available in the art. Some of these are subjects of U.S. patents as follows:

[0025] Examples include U.S. Patent 5298421, entitled "Plant medium-chain-preferring acyl-ACP thioesterases and related methods," which describes the isolation of an acyl-ACP thioesterase and the gene that encodes it from the immature seeds of *Umbellularia californica*. Other sources for such thioesterases and their encoding genes include U.S. Patent 5,304,481, entitled "Plant thioesterase having preferential hydrolase activity toward C12 acyl-ACP substrate," U.S. Patent 5,344,771, entitled "Plant thioesterases," U.S. Patent 5,455,167, entitled "Medium-chain thioesterases in plants," U.S. Patent 5,512,482, entitled "Plant thioesterases," U.S. Patent 5,530,186, entitled "Nucleotide sequences of soybean acyl-ACP thioesterase genes," U.S. Patent 5,639,790, entitled "Plant medium-chain thioesterases," U.S. Patent 5,667,997, entitled "C8 and C10 medium-chain thioesterases in plants," U.S. Patent 5,723,761, entitled "Plant acyl-ACP thioesterase sequences," U.S. Patent 5,807,893, entitled "Plant thioesterases and use for modification of fatty acid composition in plant seed oils," U.S. Patent 5,850,022, entitled "Production of myristate in plant cells," U.S. Patent 5,910,631, entitled "Middle chain-specific thioesterase genes from *Cuphea lanceolata*," U.S. Patent 5,945,585, entitled "Specific for palmitoyl, stearoyl and

oleoyl-ACP thioesters nucleic acid fragments encoding acyl-ACP thioesterase enzymes and the use of these fragments in altering plant oil composition,” U.S. Patent 5,955,329, entitled “Engineering plant thioesterases for altered substrate specificity,” U.S. Patent 5,955,650, entitled “Nucleotide sequences of canola and soybean palmitoyl-ACP thioesterase genes and their use in the regulation of fatty acid content of the oils of soybean and canola plants,” and U.S. Patent 6,331,664, entitled “Acyl-ACP thioesterase nucleic acids from maize and methods of altering palmitic acid levels in transgenic plants therewith.”

[0026] Others are described in the open literature as follows:

[0027] Dörmann, P. *et al.*, *Planta* (1993) 189:425-432, describe “Characterization of two acyl-acyl carrier protein thioesterases from developing *Cuphea* seeds specific for medium-chain and oleoyl-acyl carrier protein.” Dörmann, P., *et al.*, *Biochimica Biophysica Acta* (1994) 1212:134-136, describe “Cloning and expression in *Escherichia coli* of a cDNA coding for the oleoyl-acyl carrier protein thioesterase from coriander (*Coriandrum sativum* L.).” Filichkin, S., *et al.*, *European Journal of Lipid Science and Technology* (2006) 108:979-990, describe “New FATB thioesterases from a high-laurate *Cuphea* species: Functional and complementation analyses.” Jones, A., *et al.*, *Plant Cell* (1995) 7:359-371, describe “Palmitoyl-acyl carrier protein (ACP) thioesterase and the evolutionary origin of plant acyl-ACP thioesterases.” Knutzon, D. S., *et al.*, *Plant Physiology* (1992) 100:1751-1758, describe “Isolation and characterization of two safflower oleoyl-acyl carrier protein thioesterase cDNA clones.” Slabaugh, M., *et al.*, *The Plant Journal* (1998) 13:611-620, describe “Condensing enzymes from *Cuphea wrightii* associated with medium chain fatty acid biosynthesis.”

[0028] Additional genes, not previously isolated, that encode these acyl-ACP TEs can be isolated from plants that naturally contain large amounts of medium-chain fatty acids in their seed oil, including certain plants in the *Lauraceae*, *Lythraceae*, *Rutaceae*, *Ulmaceae*, and *Vochysiaceae* families. Typically, the fatty acids produced by the seeds of these plants are esterified to glycerol and retained inside the cells. The seeds containing the products can then be harvested and processed to isolate the fatty acids. Other sources of these enzymes, such as bacteria may also be used.

[0029] The known acyl-ACP TEs from plants can be divided into two main classes, based on their amino acid sequences and their specificity for acyl-ACPs of differing chain lengths and degrees of unsaturation. The “FatA” type of plant acyl-ACP TE has preferential activity on oleoyl-ACP, thereby releasing oleic acid, an 18-carbon fatty acid with a single double bond nine carbons distal to the carboxyl group. The “FatB” type of plant acyl-ACP

TE has preferential activity on saturated acyl-ACPs, and can have broad or narrow chain length specificities. For example, FatB enzymes from different species of *Cuphea* have been shown to release fatty acids ranging from eight carbons in length to sixteen carbons in length from the corresponding acyl-ACPs. Listed below in Table 1 are several plant acyl-ACP TEs along with their substrate preferences. (Fatty acids are designated by standard shorthand notation, wherein the number preceding the colon represents the acyl chain length and the number after the colon represents the number of double bonds in the acyl chain.)

Table 1
Plant Acyl-ACP Thioesterase

<i>Garcinia mangostana</i> FatA	18:1 and 18:0
<i>Carthamus tinctorius</i> FatA	18:1
<i>Coriandrum sativum</i> FatA	18:1
<i>Cuphea hookeriana</i> FatB1	16:0
<i>Cuphea hookeriana</i> FatB2	8:0 and 10:0
<i>Cuphea wrightii</i> FatB1	12:0 to 16:0
<i>Cuphea palustris</i> FatB1	8:0 and 10:0
<i>Cuphea palustris</i> FatB2	14:0 and 16:0
<i>Cuphea calophylla</i> FatB1	12:0 to 16:0
<i>Umbellularia californica</i> FatB1	12:0
<i>Ulmus americana</i> FatB1	8:0 and 10:0

[0030] The enzymes listed in Table 1 are exemplary and many additional genes encoding acyl-ACP TEs can be isolated and used in this invention, including but not limited to genes such as those that encode the following acyl-ACP TEs (referred to by GenPept Accession Numbers):

CAA52069.1, CAA52070.1, CAA54060.1, CAA85387.1, CAA85388.1, CAB60830.1, CAC19933.1, CAC19934.1, CAC39106.1, CAC80370.1, CAC80371.1, CAD32683.1, CAL50570.1, CAN60643.1, CAN81819.1, CAO17726.1, CAO42218.1, CAO65585.1, CAO68322.1, AAA33019.1, AAA33020.1, AAB51523.1, AAB51524.1, AAB51525.1, AAB71729.1, AAB71730.1, AAB71731.1, AAB88824.1, AAC49001.1, AAC49002.1, AAC49179.1, AAC49180.1, AAC49269.1, AAC49783.1, AAC49784.1, AAC72881.1, AAC72882.1, AAC72883.1, AAD01982.1, AAD28187.1, AAD33870.1, AAD42220.2, AAG35064.1, AAG43857.1, AAG43858.1, AAG43859.1, AAG43860.1, AAG43861.1, AAL15645.1, AAL77443.1, AAL77445.1, AAL79361.1, AAM09524.1, AAN17328.1, AAQ08202.1, AAQ08223.1, AAX51636.1, AAX51637.1, ABB71579.1, ABB71581.1, ABC47311.1, ABD83939.1, ABE01139.1, ABH11710.1, ABI18986.1, ABI20759.1, ABI20760.1, ABL85052.1, ABU96744.1, EAY74210.1, EAY86874.1, EAY86877.1, EAY86884.1, EAY99617.1, EAZ01545.1, EAZ09668.1, EAZ12044.1, EAZ23982.1, EAZ37535.1, EAZ45287.1, NP_001047567.1, NP_001056776.1, NP_001057985.1, NP_001063601.1, NP_001068400.1, NP_172327.1, NP_189147.1, NP_193041.1,

XP_001415703.1, Q39473, Q39513, Q41635, Q42712, Q9SQI3, NP_189147.1, AAC49002, CAA52070.1, CAA52069.1, 193041.1, CAC39106, CAO17726, AAC72883, AAA33020, AAL79361, AAQ08223.1, AAB51523, AAL77443, AAA33019, AAG35064, and AAL77445.

Additional sources of acyl-ACP TEs that are useful in the present invention include: *Arabidopsis thaliana* (At); *Bradyrhizobium japonicum* (Bj); *Brassica napus* (Bn); *Cinnamomum camphorum* (Cc); *Capsicum chinense* (Cch); *Cuphea hookeriana* (Ch); *Cuphea lanceolata* (Cl); *Cuphea palustris* (Cp); *Coriandrum sativum* (Cs); *Carthamus tinctorius* (Ct); *Cuphea wrightii* (Cw); *Elaeis guineensis* (Eg); *Gossypium hirsutum* (Gh); *Garcinia mangostana* (Gm); *Helianthus annuus* (Ha); *Iris germanica* (Ig); *Iris tectorum* (It); *Myristica fragrans* (Mf); *Triticum aestivum* (Ta); *Ulmus Americana* (Ua); and *Umbellularia californica* (Uc). Exemplary TEs are shown in Figure 6 with corresponding NCBI accession numbers.

[0031] In one embodiment, the present invention contemplates the specific production of an individual length of medium-chain fatty acid, for example, predominantly producing C8 fatty acids in one culture of recombinant photosynthetic microorganisms. In another embodiment, the present invention contemplates the production of a combination of two or more different length fatty acids, for example, both C8 and C10 fatty acids in one culture of recombinant photosynthetic microorganisms.

[0032] Illustrated below are manipulations of these art-known genes to construct suitable expression systems that result in production of effective amounts of the thioesterases in selected recombinant photosynthetic organisms. In such constructions, it may be desirable to remove the portion of the gene that encodes the plastid transit peptide region, as this region is inappropriate in prokaryotes. Alternatively, if expression is to take place in eukaryotic cells, the appropriate plastid transit peptide encoding region to the host organism may be substituted. Preferred codons may also be employed, depending on the host.

Other Modifications

[0033] In addition to providing an expression system for one or more appropriate acyl-ACP TE genes, further alterations in the photosynthetic host may be made. For example, the host may be modified to include an expression system for a heterologous gene that encodes a β -ketoacyl synthase (KAS) that preferentially produces acyl-ACPs having

medium chain lengths. Such KAS enzymes have been described from several plants, including various species of *Cuphea*. See Dehesh, K., *et al.*, *The Plant Journal* (1998) 15:383-390, describe "KAS IV: a 3-ketoacyl-ACP synthase from *Cuphea* sp. is a medium chain specific condensing enzyme."; Slabaugh, M., *et al.*, *The Plant Journal* (1998) 13:611-620), and would serve to increase the availability of acyl-ACP molecules of the proper length for recognition and cleavage by the heterologous medium-chain acyl-ACP TE. Another example is that the photosynthetic host cell containing a heterologous acyl-ACP TE gene may be further modified to include an expression system for a heterologous gene that encodes a multifunctional acetyl-CoA carboxylase or a set of heterologous genes that encode the various subunits of a multi-subunit type of acetyl-CoA carboxylase. Other heterologous genes that encode additional enzymes or components of the fatty acid biosynthesis pathway could also be introduced and expressed in acyl-ACP TE-containing host cells.

[0034] The photosynthetic microorganism may also be modified such that one or more genes that encode beta-oxidation pathway enzymes have been inactivated or downregulated, or the enzymes themselves may be inhibited. This would prevent the degradation of fatty acids released from acyl-ACPs, thus enhancing the yield of secreted fatty acids. In cases where the desired products are medium-chain fatty acids, the inactivation or downregulation of genes that encode acyl-CoA synthetase and/or acyl-CoA oxidase enzymes that preferentially use these chain lengths as substrates would be beneficial. Mutations in the genes encoding medium-chain-specific acyl-CoA synthetase and/or medium-chain-specific acyl-CoA oxidase enzymes such that the activity of the enzymes is diminished would also be effective in increasing the yield of secreted fatty acids. An additional modification inactivates or down-regulates the acyl-ACP synthetase gene or inactivates the gene or protein. Mutations in the genes can be introduced either by recombinant or non-recombinant methods. These enzymes and their genes are well known, and may be targeted specifically by disruption, deletion, generation of antisense sequences, generation of ribozymes or other recombinant approaches known to the practitioner. Inactivation of the genes can also be accomplished by random mutation techniques such as UV, and the resulting cells screened for successful mutants. The proteins themselves can be inhibited by intracellular generation of appropriate antibodies or intracellular generation of peptide inhibitors.

[0035] The photosynthetic microorganism may also be modified such that one or more genes that encode storage carbohydrate or polyhydroxyalkanoate (PHA) biosynthesis pathway enzymes have been inactivated or down-regulated, or the enzymes themselves may

be inhibited. Examples include enzymes involved in glycogen, starch, or chrysolaminarin synthesis, including glucan synthases and branching enzymes. Other examples include enzymes involved in PHA biosynthesis such as acetoacetyl-CoA synthase and PHA synthase.

Expression Systems

[0036] Expression of heterologous genes in cyanobacteria and eukaryotic algae is enabled by the introduction of appropriate expression vectors. For transformation of cyanobacteria, a variety of promoters that function in cyanobacteria can be utilized, including, but not limited to the lac, tac, and trc promoters and derivatives that are inducible by the addition of isopropyl β -D-1-thiogalactopyranoside (IPTG), promoters that are naturally associated with transposon- or bacterial chromosome-borne antibiotic resistance genes (neomycin phosphotransferase, chloramphenicol acetyltransferase, spectinomycin adenylyltransferase, etc.), promoters associated with various heterologous bacterial and native cyanobacterial genes, promoters from viruses and phages, and synthetic promoters. Promoters isolated from cyanobacteria that have been used successfully include the following:

- secA (secretion; controlled by the redox state of the cell)
- rbc (Rubisco operon)
- psaAB (PS I reaction center proteins; light regulated)
- psbA (D1 protein of PSII; light-inducible)

[0037] Likewise, a wide variety of transcriptional terminators can be used for expression vector construction. Examples of possible terminators include, but are not limited to, psbA, psaAB, rbc, secA, and T7 coat protein.

[0038] Expression vectors are introduced into the cyanobacterial strains by standard methods, including, but not limited to, natural DNA uptake, conjugation, electroporation, particle bombardment, and abrasion with glass beads, SiC fibers, or other particles. The vectors can be: 1) targeted for integration into the cyanobacterial chromosome by including flanking sequences that enable homologous recombination into the chromosome, 2) targeted for integration into endogenous cyanobacterial plasmids by including flanking sequences that enable homologous recombination into the endogenous plasmids, or 3) designed such that the expression vectors replicate within the chosen host.

[0039] For transformation of green algae, a variety of gene promoters and terminators that function in green algae can be utilized, including, but not limited to promoters and terminators from *Chlamydomonas* and other algae, promoters and terminators from viruses, and synthetic promoters and terminators.

[0040] Expression vectors are introduced into the green algal strains by standard methods, including, but not limited to, electroporation, particle bombardment, and abrasion with glass beads, SiC fibers, or other particles. The vectors can be 1) targeted for site-specific integration into the green algal chloroplast chromosome by including flanking sequences that enable homologous recombination into the chromosome, or 2) targeted for integration into the cellular (nucleus-localized) chromosome.

[0041] For transformation of diatoms, a variety of gene promoters that function in diatoms can be utilized in these expression vectors, including, but not limited to: 1) promoters from *Thalassiosira* and other heterokont algae, promoters from viruses, and synthetic promoters. Promoters from *Thalassiosira pseudonana* that would be suitable for use in expression vectors include an alpha-tubulin promoter (SEQ ID NO:1), a beta-tubulin promoter (SEQ ID NO:2), and an actin promoter (SEQ ID NO:3). Promoters from *Phaeodactylum tricornutum* that would be suitable for use in expression vectors include an alpha-tubulin promoter (SEQ ID NO:4), a beta-tubulin promoter (SEQ ID NO:5), and an actin promoter (SEQ ID NO:6). These sequences are deduced from the genomic sequences of the relevant organisms available in public databases and are merely exemplary of the wide variety of promoters that can be used. The terminators associated with these and other genes, or particular heterologous genes can be used to stop transcription and provide the appropriate signal for polyadenylation and can be derived in a similar manner or are known in the art.

Expression vectors are introduced into the diatom strains by standard methods, including, but not limited to, electroporation, particle bombardment, and abrasion with glass beads, SiC fibers, or other particles. The vectors can be 1) targeted for site-specific integration into the diatom chloroplast chromosome by including flanking sequences that enable homologous recombination into the chromosome, or 2) targeted for integration into the cellular (nucleus-localized) chromosome.

Host Organisms

[0042] The host cells used to prepare the cultures of the invention include any photosynthetic organism which is able to convert inorganic carbon into a substrate that is in

turn converted to fatty acid derivatives. These organisms include prokaryotes as well as eukaryotic organisms such as algae and diatoms.

[0043] Host organisms include eukaryotic algae and cyanobacteria (blue-green algae). Representative algae include green algae (chlorophytes), red algae, diatoms, prasinophytes, glaucophytes, chlorarachniophytes, euglenophytes, chromophytes, and dinoflagellates. A number of cyanobacterial species are known and have been manipulated using molecular biological techniques, including the unicellular cyanobacteria *Synechocystis* sp. PCC6803 and *Synechococcus elongates* PCC7942, whose genomes have been completely sequenced.

[0044] The following genera of cyanobacteria may be used: one group includes

<i>Chamaesiphon</i>	<i>Dactylococcopsis</i>	<i>Prochlorococcus</i>
<i>Chroococcus</i>	<i>Gloeobacter</i>	<i>Prochloron</i>
<i>Cyanobacterium</i>	<i>Gloeocapsa</i>	<i>Synechococcus</i>
<i>Cyanobium</i>	<i>Gloeotheca</i>	<i>Synechocystis</i>
<i>Cyanothece</i>	<i>Microcystis</i>	

[0045] Another group includes

<i>Cyanocystis</i>	<i>Xenococcus</i>	<i>Pleurocapsa</i>
<i>Dermocarpella</i>	<i>Chroococcidiopsis</i>	
<i>Stanieria</i>	<i>Myxosarcina</i>	

[0046] Still another group includes

<i>Arthrospira</i>	<i>Limnothrix</i>	<i>Pseudanabaena</i>
<i>Borzia</i>	<i>Lyngbya</i>	<i>Spirulina</i>
<i>Crinalium</i>	<i>Microcoleus</i>	<i>Starria</i>
<i>Geitlerinema</i>	<i>Oscillatoria</i>	<i>Symploca</i>
<i>Halospirulina</i>	<i>Planktothrix</i>	<i>Trichodesmium</i>
<i>Leptolyngbya</i>	<i>Prochlorothrix</i>	<i>Tychonema</i>

[0047] Still another group includes

<i>Anabaena</i>	<i>Cyanospira</i>	<i>Nostoc</i>
<i>Anabaenopsis</i>	<i>Cylindrospermopsis</i>	<i>Rivularia</i>
<i>Aphanizomenon</i>	<i>Cylindrospermum</i>	<i>Scytonema</i>
<i>Calothrix</i>	<i>Nodularia</i>	<i>Tolypothrix</i>

[0048] And another group includes

<i>Chlorogloeopsis</i>	<i>Geitleria</i>	<i>Nostochopsis</i>
<i>Fischerella</i>	<i>Iyengariella</i>	<i>Stigonema</i>

[0049] In addition, various algae, including diatoms and green algae can be employed.

[0050] Desirable qualities of the host strain include high potential growth rate and lipid productivity at 25-50°C, high light intensity tolerance, growth in brackish or saline water, *i.e.*, in wide range of water types, resistance to growth inhibition by high O₂ concentrations, filamentous morphology to aid harvesting by screens; resistance to predation, ability to be flocculated (by chemicals or 'on-demand autoflocculation'), excellent inorganic carbon

uptake characteristics, virus or cyanophage-resistance, tolerance to free fatty acids or other compounds associated with the invention method, and ability to undergo metabolic engineering.

[0051] Metabolic engineering is facilitated by the ability to take up DNA by electroporation or conjugation, lack of a restriction system and efficient homologous recombination in the event gene replacement or gene knockouts are required.

Fatty Acid Adsorption, Removal, and Recovery

[0052] The fatty acids secreted into the culture medium by the recombinant photosynthetic microorganisms described above can be recovered in a variety of ways. A straightforward isolation method by partition using immiscible solvents may be employed. In one embodiment, particulate adsorbents can be employed. These may be lipophilic particulates or ion exchange resins, depending on the design of the recovery method. They may be circulating in the separated medium and then collected, or the medium may be passed over a fixed bed column, for example, a chromatographic column containing these particulates. The fatty acids are then eluted from the particulate adsorbents by the use of an appropriate solvent. Evaporation of the solvent, followed by further processing of the isolated fatty acids and lipids can then be carried out to yield chemicals and fuels that can be used for a variety of commercial purposes.

[0053] The particulate adsorbents may have average diameters ranging from 0.5 mm to 30 mm which can be manufactured from various materials including, but not limited to, polyethylene and derivatives, polystyrene and derivatives, polyamide and derivatives, polyester and derivatives, polyurethane and derivatives, polyacrylates and derivatives, silicone and derivatives, and polysaccharide and derivatives. Certain glass and ceramic materials can also be used as the solid support component of the fat adsorbing objects. The surfaces of the particulate adsorbents may be modified so that they are better able to bind fatty acids and lipids. An example of such modification is the introduction of ether-linked alkyl groups having various chain lengths, preferably 8-30 carbons. In another example, acyl chains of various lengths can be attached to the surface of the fat adsorbing objects via ester, thioester, or amide linkages.

[0054] In one embodiment, the particulate adsorbents are coated with inorganic compounds known to bind fatty acids and lipids. Examples of such compounds include but are not limited to aluminum hydroxide, graphite, anthracite, and silica.

[0055] The particles used may also be magnetized or otherwise derivatized to facilitate recovery. For instance the particles may be coupled to one member of a binding pair and the adsorbed to a substrate containing the relevant binding partner.

[0056] The fatty acids may be eluted from the particulate adsorbents by the use of an appropriate solvent such as hexane or ethanol. The particulate adsorbents may be reused by returning them to the culture medium or used in a regenerated column. The solvent containing the dissolved fatty acids is then evaporated, leaving the fatty acids in a purified state for further conversion to chemicals and fuels. The particulate adsorbents can be designed to be neutrally buoyant or positively buoyant to enhance circulation in the culture medium. A continuous cycle of fatty acid removal and recovery can be implemented by utilizing the steps outlined above. The recovered fatty acids may be converted to alternative organic compounds, used directly, or mixed with other components. Chemical methods for such conversions are well understood in the art, and developments of biological methods for such conversions are also contemplated

[0057] The present invention further contemplates a variety of compositions comprising the fatty acids produced by the recombinant photosynthetic microorganisms described herein, and uses thereof. The composition may comprise the fatty acids themselves, or further derivatives of the fatty acids, such as alcohols, alkanes, and alkenes which can be generated from the fatty acids produced by the microorganisms by any methods that are known in the art, as well as by development of biological methods of conversion. For examples, fatty acids may be converted to alkenes by catalytic hydrogenation and catalytic dehydration.

[0058] The composition may serve, for example, as a biocrude. The biocrude can be processed through refineries that will convert the composition compounds to various petroleum and petrochemical replacements, including alkanes, olefins and aromatics through processes including hydrotreatment, decarboxylation, isomerization and catalytic cracking and reforming. The biocrude can be also converted to ester-based fuels, such as fatty acid methyl ester (commercially known as biodiesel), through established chemical processes including transesterification and esterification.

[0059] In addition, one of skill in the art could contemplate a variety of other uses for the fatty acids of the present invention, and derivatives thereof, that are well known in the art, for example, the production of chemicals, soaps, surfactants, detergents, lubricants, nutraceuticals, pharmaceuticals, cosmetics, *etc.* For example, derivatives of the fatty acids

of the present invention include C8 chemicals, such as octanol, used in the manufacture of esters for cosmetics and flavors as well as for various medical applications, and octane, used primarily as a co-monomer in production of polyethylene. Derivatives of the fatty acids of the present invention may also include C10 chemicals, such as decanol, used in the manufacture of plasticizers, surfactants and solvents, and decene, used in the manufacture of lubricants.

[0060] Biocrudes are biologically produced compounds or a mix of different biologically produced compounds that are used as a feedstock for refineries in replacement of, or in complement to, crude oil or other forms of petroleum. In general, but not necessarily, these feedstocks have been pre-processed through biological, chemical, mechanical or thermal processes in order to be in a liquid state that is adequate for introduction in a petroleum refinery.

[0061] The fatty acids of the present invention can be a biocrude, and further processed to a biofuel composition. The biofuel can then perform as a finished fuel or a fuel additive.

[0062] “Finished fuel” is defined as a chemical compound or a mix of chemical compounds (produced through chemical, thermochemical or biological routes) that is in an adequate chemical and physical state to be used directly as a neat fuel or fuel additive in an engine. In many cases, but not always, the suitability of a finished fuel for use in an engine application is determined by a specification which describes the necessary physical and chemical properties that need to be met. Some examples of engines are: internal combustion engine, gas turbine, steam turbine, external combustion engine, and steam boiler. Some examples of finished fuels include: diesel fuel to be used in a compression-ignited (diesel) internal combustion engine, jet fuel to be used in an aviation turbine, fuel oil to be used in a boiler to generate steam or in an external combustion engine, ethanol to be used in a flex-fuel engine. Examples of fuel specifications are ASTM standards, mainly used in the US, and the EN standards, mainly used in Europe.

[0063]

“Fuel additive” refers to a compound or composition that is used in combination with another fuel for a variety of reasons, which include but are not limited to complying with mandates on the use of biofuels, reducing the consumption of fossil fuel-derived products or enhancing the performance of a fuel or engine. For example, fuel additives can be used to alter the freezing/gelling point, cloud point, lubricity, viscosity, oxidative stability, ignition quality, octane level, and flash point. Additives can further function as antioxidants,

demulsifiers, oxygenates, thermal stability improvers, cetane improvers, stabilizers, cold flow improvers, combustion improvers, anti-foams, anti-haze additives, icing inhibitors, injector cleanliness additives, smoke suppressants, drag reducing additives, metal deactivators, dispersants, detergents, demulsifiers, dyes, markers, static dissipaters, biocides, and/or corrosion inhibitors.

The following examples are offered to illustrate but not to limit the invention.

Example 1

Secretion of Fatty Acids by Strains Derived from the Unicellular Photoautotrophic Cyanobacterium *Synechococcus elongatus* PCC 7942

[0064] The *Cuphea hookeriana* FatB2 gene encoding an acyl-ACP thioesterase (ChFatB2) enzyme was modified for optimized expression in *Synechococcus elongatus* PCC 7942. First, the portion of the gene that encodes the plastid transit peptide region of the native ChFatB2 protein was removed. The remainder of the coding region was then codon-optimized using the “Gene Designer” software program (version 1.1.4.1) provided by DNA2.0, Inc. The nucleotide sequence of this derivative of the ChFatB2 gene (hereafter ChFatB2-7942) is provided as SEQ ID NO:7. The protein sequence encoded by this gene is provided in SEQ ID NO:8.

[0065] Two different versions of the *trc* promoter, *trc* (Egon, A., *et al.*, *Gene* (1983) 25:167-178) and “enhanced *trc*” (hereafter *trcE*, from pTrcHis A, Invitrogen) were used to drive the expression of ChFatB2-7942 in *S. elongatus* PCC 7942. The *trc* promoter is repressed by the Lac repressor protein encoded by the *lacIq* gene and can be induced by the addition of isopropyl β -D-1-thiogalactopyranoside (IPTG). The *trcE* promoter is a derivative of *trc* designed to facilitate expression of eukaryotic proteins in *E. coli* and is also inducible by IPTG.

[0066] The fusion fragments of ChFatB2-7942 operably linked to *trc* or *trcE*, together with the *lacIq* gene, were cloned into the shuttle vector pAM2314 (Mackey, S. R., *et al.*, *Methods Mol. Biol.* (2007) 362:115-129), which enables transformation of *S. elongatus* PCC 7942 via double homologous recombination-mediated integration into the “NS1” site of the chromosome. The constructed plasmid containing the *trcE*::ChFatB2-7942 expression cassette and *lacIq* gene is designated pSGI-YC01. SEQ ID NO:9 represents the sequence between and including the NS1 recombination sites of pSGI-YC01. The

constructed plasmid containing the *trc::ChFatB2-7942* expression cassette and *lacIq* gene is designated pSGI-YC09. SEQ ID NO:10 represents the sequence between and including the NS1 recombination sites of pSGI-YC09.

[0067] Each of the plasmids pSGI-YC01 and pSGI-YC09, along with the control vector pAM2314, were introduced into wild-type *S. elongatus* PCC 7942 cells as described by Golden and Sherman (*J. Bacteriol.* (1984) 158:36-42). Both recombinant and control strains were pre-cultivated in 100 mL of BG-11 medium supplied with spectinomycin (5 mg/L) to late-log phase ($OD_{730nm} = 1.0$) on a rotary shaker (150 rpm) at 30°C with constant illumination ($60 \mu E m^{-2} sec^{-1}$). Cultures were then subcultured at initial $OD_{730nm} = 0.4-0.5$ in BG-11 and cultivated overnight to $OD_{730nm} = 0.7-0.9$. For time-course study, 60 mL aliquots of the culture were transferred into 250-mL flasks and induced by adding IPTG (final conc. = 1 mM) if applicable. Cultures were sampled 0, 48, 96, and 168 hours after IPTG induction and then filtered through Whatman® GF/F filters using a Millipore vacuum filter manifold. Filtrates were collected in screw top culture tubes for gas chromatographic (GC) analysis.

[0068] Free fatty acids (FFAs) were separated from filtered cell cultures using liquid-liquid extraction. Five mL of the filtrate were mixed with 125 μL of 1 M H_3PO_3 and 0.25 mL of 5 M NaCl, followed by addition of 2 mL of hexane and thorough mixing. For GC-FID analyze, a 0.2 μl sample of the hexane was injected using a 40:1 split ratio onto a DB-FFAP column (J&W Scientific, 15 m x 250 μm x 0.25 μm), with a temperature profile starting at 150°C for 0.5 min, then heating at 15°C/min to 230°C and holding for 7.1 min (1.1 mL/min He).

[0069] GC analysis results indicating the levels of medium-chain FFAs (8:0 and 10:0) in cultures containing various *Synechococcus elongatus* strains 168 hours after IPTG induction are shown in Table 1-1.

Table 1-1
Medium-chain fatty acid secretion in various strains of *S. elongatus*

Strain	Parent Strain	Plasmid Added	Transgenes	Fatty Acids (mg/L)	
				8:0	10:0
SGC-YC2-5	PCC 7942	pAM2314	none	ND	ND
SGC-YC1-2	PCC 7942	pSGI-YC01	<i>trcE::ChFatB2-7942</i>	1.5	3.5
SGC-YC14-4	PCC 7942	pSGI-YC09	<i>trc::ChFatB2-7942</i>	5.1	10.1

Note: ND represents “not detected” (<1 mg/L).

Example 2

Secretion of Fatty Acids by Strains Derived from the Unicellular Photoheterotrophic Cyanobacterium *Synechocystis* sp. PCC 6803

[0070] The *trcE::ChFatB2-7942* and *trc::ChFatB2-7942* fusion fragments, together with the *lacIq* gene, were cloned into the shuttle vector pSGI-YC03 (SEQ ID NO:11), which enables transformation of *Synechocystis* sp. PCC 6803 via double homologous recombination-mediated integration into the “RS1” site of the chromosome (Williams, Methods Enzymol. (1988) 167:766–778). The constructed plasmid containing the *trcE::ChFatB2-7942* expression cassette and *lacIq* gene is designated pSGI-YC08. SEQ ID NO:12 represents the sequence between and including the RS1 recombination sites of pSGI-YC08. The constructed plasmid containing the *trc::ChFatB2-7942* expression cassette and *lacIq* gene is designated pSGI-YC14. SEQ ID NO:13 represents the sequence between and including the RS1 recombination sites of pSGI-YC14.

[0071] Each of the plasmids pSGI-YC08, pSGI-YC14, and the control vector pSGI-YC03, was introduced into wild-type *Synechocystis* PCC 6803 cells, as described by Zang, X. *et al.*, *J Microbiol.* (2007) 45:241-245. Both recombinant and control strains were pre-cultivated in 100 mL of BG-11 medium supplied with kanamycin (10 mg/L) to late-log phase ($OD_{730nm} = 1.0$) on a rotary shaker (150 rpm) at 30°C with constant illumination ($60 \mu E \cdot m^{-2} \cdot sec^{-1}$). Cultures were then subcultured at initial $OD_{730nm} = 0.4-0.5$ in BG-11 and cultivated overnight to $OD_{730nm} = 0.7-0.9$. For time-course studies, 60-mL aliquots of the culture were transferred into 250-mL flasks and induced by adding IPTG (final conc. = 1 mM) when applicable. Cultures were sampled 0, 72, and 144 hours after IPTG induction and then filtered through Whatman® GF/B filters using a Millipore vacuum filter manifold. Filtrates were collected in screw top culture tubes for gas chromatographic (GC) analysis. Free fatty acids (FFA) were separated from the filtered culture supernatant solutions by liquid-liquid extraction. For each sample, 2 mL filtered culture was extracted with a mixture of 50 μ l phosphoric acid (1 M), 100 μ l NaCl (5 M) and 2 mL hexane. A 0.2 μ l sample was injected using a 40:1 split ratio on to a DB-FFAP column (J&W Scientific, 15 m x 250 μ m x 0.25 μ m), with a temperature profile starting at 150°C for 0.5 min, then heating at 15°C/min to 230°C and holding for 7.1 min (1.1 mL/min He).

[0072] GC analysis results indicating the levels of medium-chain FFAs (8:0 and 10:0) in cultures 144 hours after IPTG induction are shown in Table 2-1.

Table 2-1

Medium-chain fatty acid secretion in various strains of *Synechocystis*.

Strain	Parent Strain	Plasmid Added	Transgenes	Fatty Acids (mg/L)	
				8:0	10:0
SGC-YC9-8	PCC 6803	pSGI-YC03	none	ND	ND
SGC-YC10-5	PCC 6803	pSGI-YC08	trcE::ChFatB2-7942	61.3	52.7
SGC-YC16-2	PCC 6803	pSGI-YC14	trc::ChFatB2-7942	2.7	5.8

Note: ND represents “not detected” (<1 mg/L).

Example 3Secretion of Fatty Acids by Strains Derived from the Filamentous Cyanobacterium*Anabaena variabilis* ATCC 29413

[0073] The trc::ChFatB2-7942 and trcE::ChFatB2-7942 fusion fragments, together with the lacIq gene, were PCR amplified using primers RS3-3F (SEQ ID NO:14) and 4YC-rmBter-3 (SEQ ID NO:15) from pSGI-YC14 and pSGI-YC08, respectively, and then cloned into the shuttle vector pEL17, which enables transformation of *A. variabilis* ATCC 29413 via double homologous recombination-mediated integration into the nifU1 locus of the chromosome (Lyons and Thiel, J. Bacteriol. (1995) 177:1570-1575). The constructed plasmids are designated pSGI-YC69 and pSGI-YC70 for trc::ChFatB2-7942 and trcE::ChFatB2-7942, respectively.

[0074] Each of the plasmids pSGI-YC69, pSGI-YC70, along with the control vector pEL17, are introduced into wild-type *A. variabilis* ATCC 29413 cells via tri-parental conjugation, as described by Elhai and Wolk (*Methods Enzymol.* (1988) 167:747-754). Both recombinant and control strains are pre-cultivated in 100 mL of BG-11 medium supplied with 5 mM NH₄Cl and spectinomycin (3 mg/L) to late-log phase (OD_{730nm} = 1.0) on a rotary shaker (150 rpm) at 30°C with constant illumination (60 μE·m⁻²·sec⁻¹). Cultures are then subcultured at initial OD_{730nm} = 0.4-0.5 in BG-11 and cultivated overnight to OD_{730nm} = 0.7-0.9. For time-course studies, 60-mL aliquots of the culture are transferred into 250 mL flasks and induced by adding IPTG (final conc. = 1 mM) if applicable. Cultures are sampled every 72 hours and then filtered through Whatman® GF/F filters using a Millipore vacuum filter manifold. Filtrates are collected in screw top culture tubes for gas chromatographic (GC) analysis as described in Example 1.

Example 4

Secretion of Fatty Acids in Strains Derived from *Synechococcus elongatus* PCC 7942 Containing an Inactivated Acyl-ACP Synthetase Gene

[0075] A putative acyl-ACP synthetase gene in *S. elongatus* PCC 7942, *synpcc7942_0918* (Cyanobase gene designation), was disrupted via replacing of an internal 422-bp portion of its coding region with a 1,741-bp DNA sequence carrying the chloramphenicol resistance marker gene, *cat* (which encodes chloramphenicol acetyltransferase). Primer pairs 918-15 (SEQ ID NO:16) / 918-13 (SEQ ID NO:17) and 918-25 (SEQ ID NO:18) / 918-23 (SEQ ID NO:19) were used to amplify two DNA fragments corresponding to a 5' portion (1-480 bp) and a 3' portion (903-1521 bp) of the coding region of *synpcc7942_0918*, respectively. The *cat* fragment was amplified from plasmid pAM1573 (Mackey *et al.*, Methods Mol. Biol. 362:115-29) using PCR with primers NS21-3Cm (SEQ ID NO:20) and ter-3Cm (SEQ ID NO:21), which overlap primers 918-13 and 918-25, respectively. The recombinant chimeric PCR technique was then used to amplify the complete disruption cassette with the three aforementioned PCR fragments, as well as primers 918-15 and 918-23. The resulting 2,840-bp blunt-end PCR fragment (SEQ ID NO:22) was then ligated into pUC19 (Yanisch-Perron *et al.*, Gene 33:103-119), which has been digested with both *HindIII* and *EcoRI* to remove the multiple cloning sites and subsequently blunted with T4 DNA polymerase, to yield plasmid pSGI-YC04.

[0076] Plasmid pSGI-YC04 was introduced into *S. elongatus* strain SGC-YC1-2, which harbors a copy of *trcE::ChFatB2-7942* integrated into NS1 (see Example 1). The resulting strain was designated SGC-YC4-7. Fatty acid production assays and GC analyses were performed as described in Example 1. The results of GC analyses indicating the levels of FFAs in cultures of various *S. elongatus* strains 168 hours after IPTG induction are shown in Table 4-1. It is possible that inactivation of the acyl-ACP synthetase gene has a larger impact on secretion of long-chain fatty acids than on secretion of medium-chain fatty acids.

Table 4-1

Medium-chain fatty acid secretion in various strains of *S. elongatus*.

Strain	Parent Strain	Plasmid Added	Transgenes	Deletions	Fatty Acids (mg/L)			
					8:0	10:0	16:0	16:1
SGC-YC2-5	PCC 7942	pAM2314	none	none	ND	ND	ND	1.4
SGC-YC1-2	PCC 7942	pSGI-YC01	trcE::ChFatB2-7942	none	1.4	4.2	ND	1.6
SGC-YC4-7	SGC-YC1-2	pSGI-YC04	trcE::ChFatB2-7942	<i>synpcc7942_0918</i>	1.0	3.1	1.1	3.9

Note: ND represents “not detected” (<1 mg/L).

Example 5Secretion of Fatty Acids in Strains Derived from *Synechocystis* sp. PCC6803 Containing an Inactivated Acyl-ACP Synthetase Gene

[0077] A ~1.7-kbp DNA fragment spanning an area upstream and into the coding region of the acyl-ACP synthetase-encoding gene, *slr1609* (Cyanobase gene designation), from *Synechocystis* sp. PCC 6803 was amplified from genomic DNA using PCR with primers NB001 (SEQ ID NO:23) and NB002 (SEQ ID NO:24). This fragment was cloned into the pCR2.1 vector (Invitrogen) to yield plasmid pSGI-NB3 and subsequently cut with the restriction enzyme MfeI. A chloramphenicol resistance marker cassette containing the *cat* gene and associated regulatory control sequences was amplified from plasmid pAM1573 (Andersson, *et al.*, *Methods Enzymol.* (2000) 305:527–542) to contain flanking MfeI restriction sites using PCR with primers NB010 (SEQ ID NO:25) and NB011 (SEQ ID NO:26). The *cat* gene expression cassette was then inserted into the MfeI site of pSGI-NB3 to yield pSGI-NB5 (SEQ ID NO:27).

[0078] The pSGI-NB5 vector was transformed into trcE::ChFatB2-7942-containing *Synechocystis* strain SGC-YC10-5 (see Example 1) according to Zang *et al.*, *J Microbiology* (2007) 45:241-245. Insertion of the chloramphenicol resistance marker into the *Slr1609* gene through homologous recombination was verified by PCR screening of insert and insertion site. The resulting strain was designated SGC-NB10-4, which was tested in liquid BG-11 medium for fatty acid secretion. All liquid medium growth conditions used a rotary shaker (150 rpm) at 30°C with constant illumination ($60 \mu\text{E}\cdot\text{m}^{-2}\cdot\text{sec}^{-1}$). Cultures were inoculated in 25 mL of BG-11 medium containing chloramphenicol and/or kanamycin (5 $\mu\text{g}/\text{mL}$) accordingly and grown to a sufficient density (minimal $\text{OD}_{730\text{nm}} = 1.6\text{-}2$). Cultures were then used to inoculate 100 mL BG-11 medium in 250 mL polycarbonate flasks to

OD_{730nm} = 0.4-0.5 and incubated overnight. 45 mL of overnight culture at OD_{730nm} = 0.7-0.9 were added to new 250 mL flasks, inducing with 1 mM IPTG or using as uninduced controls. 5 mL samples were taken at 0, 72 and 144 hours post induction and processed as described in Example 2.

[0079] Free fatty acids (FFA) were separated from the filtered culture supernatant solutions by liquid-liquid extraction for GC/FID (flame ionization detector) analysis. For each sample, 2 mL filtered culture was extracted with a mixture of 50 μ l phosphoric acid (1 M), 100 μ l NaCl (5 M) and 2 mL hexane. A 0.2 μ l sample was injected using a 40:1 split ratio on to a DB-FFAP column (J&W Scientific, 15 m x 250 μ m x 0.25 μ m), with a temperature profile starting at 150°C for 0.5 min, then heating at 15°C/min to 230°C and holding for 7.1 min (1.1 mL/min He).

[0080] GC results indicating secreted levels of free fatty acids after 144 hours are shown in Table 5-1.

Table 5.1

Medium-chain fatty acid secretion in various strains of *Synechocystis*.

Strain	Parent Strain	Plasmid Added	Transgenes	Deletions	Fatty Acids (mg/L)	
					8:0	10:0
SGC-YC10-5	PCC 6803	pSGI-YC08	trcE::ChFatB2-7942	none	58.3	67.7
SGC-NB10-4	SGC-YC10-5	pSGI-NB5	trcE::ChFatB2-7942	<i>slr1609</i>	57.7	73.7

Note: ND represents “not detected” (<1 mg/L).

Example 6

Expression of *Cuphea lanceolata* Kas-IV and *Helianthus annuus* Kas-III genes in *Synechocystis* sp.

[0081] A DNA fragment comprising a functional operon was synthesized such that it contained the following elements in the given order: the trc promoter, the *Cuphea lanceolata* 3-ketoacyl-acyl carrier protein synthase IV gene (ClKas-IV, GenBank Accession No. CAC59946) codon-optimized for expression in *Synechococcus elongatus* PCC 7942, and the rps14 terminator (SEQ ID NO:28) from *Synechococcus* sp. WH8102. The nucleotide sequence of this entire functional operon, along with various flanking restriction enzyme recognition sites, is provided in SEQ ID NO:29.

[0082] Another DNA fragment comprising a functional operon was synthesized such that it contained the following elements in the given order: the trc promoter, the *Helianthus annuus* 3-ketoacyl-acyl carrier protein synthase III gene (HaKas-III, GenBank Accession No. ABP93352) codon-optimized for expression in both *Synechococcus elongatus* PCC 7942 and *Synechocystis* sp. PCC 6803, and rps14 terminator from *Synechococcus* sp. WH8102. The nucleotide sequence of this functional operon, along with various flanking restriction enzyme recognition sites, is provided in SEQ ID NO:30.

[0083] Codon optimization was performed by the use of the “Gene Designer” (version 1.1.4.1) software program provided by DNA2.0, Inc. The functional operon (expression cassette) containing the codon-modified ClKas-IV gene as represented in SEQ ID NO:29 was digested by the restriction enzymes SpeI and XbaI and inserted into plasmid pSGI-YC39 between the restriction sites SpeI and XbaI to form plasmid pSGI-BL26, which enables integration of the functional operon into the *Synechocystis* sp. PCC 6803 chromosome at the “RS2” recombination site (Aoki, *et al.*, *J. Bacteriol* (1995) 177:5606-5611). The plasmid pSGI-BL27 containing the DNA fragment represented in SEQ ID NO:30 was constructed in the same way.

[0084] Plasmid pSGI-BL43 contains the trcE promoter, the codon-optimized ClKas-IV gene, and the rps14 terminator as represented in SEQ ID NO:31 and was made by inserting a SpeI/NcoI trcE fragment from pTrcHis A (Invitrogen) into SpeI/NcoI-digested pSGI-BL26. An additional plasmid, pSGI-BL44, contains the trcE promoter, the optimized ClKas-IV gene, the *S. elongatus* PCC 7942 *kaiBC* intergenic region, the optimized HaKas-III gene, and the rps14 terminator as represented in SEQ ID NO:32 and was made by inserting a BamHI/SacI fragment (containing the *S. elongatus kaiBC* intergenic region, the HaKas-III gene, and the rps14 terminator) generated via PCR amplification into BglII/SacI-digested pSGI-BL43. The PCR primers used to generate the DNA fragment containing the *kaiBC* region, HaKas-III, and rps14 terminator are provided as SEQ ID NO:33 and SEQ ID NO:34.

[0085] Wild-type *Synechocystis* PCC 6803 cells and transgenic *Synechocystis* strain SGC-YC10-5, which contains the ChFatB2-7942 gene, were transformed with plasmids pSGI-BL26, pSGI-BL27, pSGI-BL43 and pSGI-BL44 as described by Zang, X. *et al. J. Microbiol.* (2007) 45:241-245. Both recombinant and wild-type control strains were pre-cultivated in 20 mL of BG-11 medium to mid-log phase ($OD_{730nm} = 0.7-0.9$) on a rotary shaker (150 rpm) at 30°C with constant illumination ($60 \mu E \cdot m^{-2} \cdot sec^{-1}$). Kanamycin (5

μg/mL) and/or spectinomycin (10 μg/mL) were included in recombinant cultures as appropriate. Cultures were then subcultured at initial OD_{730nm} = 0.4-0.5 in BG-11 and cultivated overnight to OD_{730nm} = 0.7-0.9. For a time-course study, 45-mL aliquots of the culture were transferred into 250 mL flasks and induced by adding IPTG (final conc. = 1 mM) when applicable. Cultures were sampled 0, 72, and 144 hours after IPTG induction and then filtered through Whatman® GF/B filters using a Millipore vacuum filter manifold. Filtrates were collected in screw top culture tubes for gas chromatographic (GC) analysis as described in Example 2.

[0086] Results indicating the levels of secreted octanoic acid and decanoic acid in culture supernatants 144 hours after culture inoculation are shown in Table 6-1. The ClKas-IV and HaKas-III genes present in the indicated strains were under the control of the trc promoter.

Table 6-1

Medium-chain fatty acid secretion in (in mg/L) various *Synechocystis sp.* strains

Strain	Parent Strain	Plasmid Added	Transgenes	Fatty Acids (mg/L)	
				8:0	10:0
PCC 6803	n/a	n/a	None	ND	ND
SGC-YC10-5	PCC 6803	pSGI-YC08	trcE-ChFatB2-7942	69.8	68.4
SGC-BL26-3	PCC 6803	pSGI-BL26	trc-ClKas-IV	ND	ND
SGC-BL26-5	SGC-YC10-5	pSGI-BL26	trcE-ChfatB2-7942 trc-ClKas-IV	69.5	71.9
SGC-BL27-1	PCC 6803	pSGI-BL27	trc-HaKas-III	ND	ND
SGC-BL27-2	SGC-YC10-5	pSGI-BL27	trcE-ChFatB2-7942 trc-HaKas-III	65.7	66.6

Note: ND represents “not detected” (<1 mg/L).

[0087] For a more optimized measurement of fatty acid secretion in these strains, the fatty acid secretion data shown in Table 6-1 were normalized to cell culture density, measured as optical density at 730 nm (OD_{730nm}); these data are presented in Table 6-2. Other experiments described in this application could be normalized in a similar fashion.

Table 6-2

Normalized medium-chain fatty acid secretion (mg/L/OD_{730nm}) in various *Synechocystis sp.* strains

Strain	Parent Strain	Plasmid Added	Transgenes	Fatty Acids	
				8:0	10:0
PCC 6803	n/a	n/a	None	ND	ND
SGC-YC10-5	PCC 6803	pSGI-YC08	trcE-ChFatB2-7942	11.7	11.4
SGC-BL26-3	PCC 6803	pSGI-BL26	trc-ClKas-IV	ND	ND
SGC-BL26-5	SGC-YC10-5	pSGI-BL26	trcE-ChfatB2-7942 trc-ClKas-IV	11.7	12.1
SGC-BL27-1	PCC 6803	pSGI-BL27	trc-HaKas-III	ND	ND
SGC-BL27-2	SGC-YC10-5	pSGI-BL27	trcE-ChFatB2-7942 trc-HaKas-III	12.2	12.3

Note: ND represents “not detected” (<1 mg/L).

[0088] Results indicating the levels of secreted octanoic acid and decanoic acid in culture supernatants of additional strains 120 hours after culture inoculation are shown in Table 6-3. The ClKas-IV and HaKas-III genes present in the indicated strains were under the control of the trcE promoter.

Table 6-3

Medium-chain fatty acid secretion (in mg/L) in various *Synechocystis* sp. strains

Strain	Parent Strain	Plasmid Added	Transgenes	Fatty Acids (mg/L)	
				8:0	10:0
SGC-YC10-5	PCC 6803	pSGI-YC08	trcE-ChFatB2-7942	34.8	43.5
SGC-BL44	PCC 6803	pSGI-BL44	trcE-ClKAS-IV+HaKAS-III	ND	ND
SGC-YC10-5-BL43	SGC-YC10-5	pSGI-BL43	trcE-ChFatB2-7942 trcE-ClKas-IV	40.0	48.1
SGC-YC10-5-BL44	SGC-YC10-5	pSGI-BL44	trcE-ChfatB2-7942 trcE-ClKAS-IV+HaKAS-III	38.5	47.1

Note: ND represents “not detected” (<1 mg/L).

[0089] For a more optimized measurement of fatty acid secretion in these strains, the fatty acid secretion data shown in Table 6-1 were normalized to cell culture density, measured as optical density at 730 nm (OD_{730nm}); these data are presented in Table 6-4.

Table 6-4

Normalized medium-chain fatty acid secretion (mg/L/OD_{730nm}) in various *Synechocystis* sp. strains

Strain	Parent Strain	Plasmid Added	Transgenes	Fatty Acids	
				8:0	10:0
SGC-YC10-5	PCC 6803	pSGI-YC08	trcE-ChFatB2-7942	6.8	8.5
SGC-BL44	PCC 6803	pSGI-BL44	trcE-ClKAS-IV+HaKAS-III	ND	ND
SGC-YC10-5-BL43	SGC-YC10-5	pSGI-BL43	trcE-ChFatB2-7942 trcE-ClKas-IV	7.4	8.9
SGC- YC10-5-BL44	SGC-YC10-5	pSGI-BL44	trcE-ChfatB2-7942 trcE-ClKAS-IV+HaKAS-III	8.3	10.2

Example 7

Introduction of a Heterologous Acyl-ACP Thioesterase Gene into a Diatom

[0090] A synthetic gene that encodes a derivative of the ChFatB2 enzyme with specificity for medium-chain (8:0-10:0) acyl-ACPs is expressed in various diatoms (*Bacillariophyceae*) by constructing and utilizing expression vectors comprising the ChFatB2 gene operably linked to gene regulatory regions (promoters and terminators) that function in diatoms. In a preferred embodiment, the gene is optimized for expression in specific diatom species and the portion of the gene that encodes the plastid transit peptide region of the native ChFatB2 protein is replaced with a plastid transit peptide that functions optimally in diatoms. The nucleotide sequence provided as SEQ ID NO:35 represents a synthetic derivative of the ChFatB2 gene that has been optimized for expression in *Thalassiosira pseudonana* and in which the native plastid transit peptide-encoding region of the gene has been replaced with the plastid transit peptide (including coupled signal sequence) associated with the gamma subunit of the coupling factor portion (CF1) of the chloroplast ATP synthase from *T. pseudonana* (JGI Identifier = jgi/Thaps3/40156/est Ext_gwp_gwl.C_chr_40019). The protein encoded by this gene, referred to hereafter as ChFatB2-Thal,) is provided in SEQ ID NO:36.

[0091] To produce an expression vector for *T. pseudonana*, the ChFatB2-Thal gene was placed between the *T. pseudonana* alpha-tubulin promoter and terminator regulatory sequences. The alpha-tubulin promoter was amplified from genomic DNA isolated from *T. pseudonana* CCMP 1335 by use of primers PR1 (SEQ ID NO:37) and PR3 (SEQ ID NO:38), whereas the alpha-tubulin terminator was amplified by use of primers PR4 (SEQ

ID NO:39) and PR8 (SEQ ID NO:40). The KpnI/BamHI fragment from the alpha-tubulin promoter amplicon, the BamHI/XbaI fragment from the alpha-tubulin terminator and the large fragment from KpnI/XbaI- cut pUC118 (Vieira and Messing, *Meth. Enzymol.* (1987) 153:3-11) were then combined to form pSGI-PR5. The NcoI/BamHI fragment from ChFatB2-Thal gene was then inserted into NcoI/BamHI-digested pSGI-PR5 to form pSGI-PR16. In addition, a codon-optimized gene that encodes the nourseothricin acetyltransferase (NAT) enzyme from *Streptomyces noursei* (SEQ ID NO:41) (Krugel, *et al.*, *Gene* (1993) 127:127-131) was synthesized and the NcoI/BamHI fragment from this NAT-encoding DNA molecule was inserted into the large NcoI/BamHI fragment from pSGI-PR5 to form pSGI-PR7, which upon introduction into *T. pseudonana* and other diatoms can provide resistance to the antibiotic nourseothricin.

[0092] pSGI-PR16 and pSGI-PR7 were co-transformed into *T. pseudonana* CCMP 1335 by means of particle bombardment essentially as described by Poulsen, *et al.*, (*J. Phycol.* (2006) 42:1059-1065). Transformed cells were selected on agar plates in the presence of 100 mg/L nourseothricin (ClonNAT, obtained from Werner BioAgents, Germany). The presence of the ChFatB2-Thal gene in cells was confirmed by the use of PCR. Transformants were grown in ASW liquid medium (Darley and Volcani, *Exp. Cell Res.* (1964) 58:334) on a rotary shaker (150 rpm) at 18°C with constant illumination (60 $\mu\text{E}\cdot\text{m}^{-2}\cdot\text{sec}^{-1}$). Samples were removed seven days after inoculation and the culture medium was tested for the presence of FFAs as described in Example 1.

[0093] Although no fatty acid secretion was detected under these particular experimental conditions, optimization of the ChFatB2-Thal gene and diatom host strain can be performed to achieve fatty acid secretion in diatoms, which are known to have relatively impervious cell walls.

Example 8

Secretion of Fatty Acids by Green Algae

[0094] A synthetic gene that encodes a derivative of the ChFatB2 enzyme with specificity for medium-chain (8:0-10:0) acyl-ACPs is expressed in green algae (*Chlorophyceae*) by constructing and utilizing expression vectors comprising the ChFatB2 gene operably linked to gene regulatory regions (promoters and terminators) that function in green algae. The gene is optimized for expression in specific green algal species and the portion of the gene that encodes the plastid transit peptide region of the native ChFatB2

protein is replaced with a plastid transit peptide that functions optimally in green algae. The nucleotide sequence provided as SEQ ID NO:42 represents a derivative of the ChFatB2 gene optimized for expression in *Chlamydomonas reinhardtii* and in which the native plastid transit peptide-encoding region of the gene has been replaced with the plastid transit peptide associated with the gamma subunit of the coupling factor portion (CF1) of the chloroplast ATP synthase from *C. reinhardtii* (GenPept Accession No. XP 001696335). The protein encoded by this gene is provided in SEQ ID NO:43.

Example 9

Secretion of Fatty Acids in Strains of *Synechocystis* sp. Containing a Disrupted 1,4-alpha-Glucan Branching Enzyme Gene

[0095] A 1.4-kbp DNA fragment spanning an area upstream and into the coding region of the 1,4-alpha-glucan branching enzyme gene (*glgB*, Cyanobase gene designation = sll0158) from *Synechocystis* sp. PCC6803 was amplified from genomic DNA using PCR with primers glgB-5 (SEQ ID NO:44) and glgB-3 (SEQ ID NO:45). This fragment was cloned into the pCR4-Topo vector (Invitrogen) to yield plasmid pSGI-BL32 and subsequently cut with the restriction enzyme *Ava*I. A spectinomycin resistance marker cassette containing the *aadA* gene and associated regulatory control sequences was digested by *Hind*III from plasmid pSGI-BL27. Both of the linear fragments were treated with the Quick Blunting™ Kit (New England Biolabs). The *aadA* gene expression cassette was then inserted into the *Ava*I site of pSGI-BL32 to yield pSGI-BL33. The portion of pSGI-BL33 that inserts into and inactivates the *glgB* gene is provided as SEQ ID NO:46).

[0096] The pSGI-BL33 vector was transformed into wild-type *Synechocystis* PCC 6803 and into *trcE::ChFatB2-7942*-containing *Synechocystis* strain SGC-YC10-5 (see Example 1) according to Zang, *et al.*, *J. Microbiology* (2007) 45:241-245. Insertion of the spectinomycin resistance marker into the Sll0158 (*glgB*) gene via homologous recombination was verified by PCR screening of insert and insertion site. Verified knockout strains were tested in liquid BG-11 medium for secretion of fatty acids. All liquid medium growth conditions used a rotary shaker (150 rpm) at 30°C with constant illumination (60 $\mu\text{E}\cdot\text{m}^{-2}\cdot\text{sec}^{-1}$). Cultures were inoculated in 25 mL of BG-11 medium containing spectinomycin (10 $\mu\text{g}/\text{mL}$) and/or kanamycin (5 $\mu\text{g}/\text{mL}$) accordingly and grown to a sufficient density (minimal $\text{OD}_{730\text{nm}} = 1.6\text{-}2$). Cultures were then used to inoculate 100 mL

BG-11 medium in 250-mL polycarbonate flasks to $OD_{730nm} = 0.4-0.5$ and incubated overnight. Forty-five mL of overnight culture at $OD_{730nm} = 0.5$ were added to new 250-mL flasks; some cultures were induced with 1 mM IPTG or used as uninduced controls. Samples (0.5 mL) were taken at 0, 72, 144, and 216 hours post induction and processed as described in Example 2.

[0097] Free fatty acids (FFA) were separated from the filtered culture supernatant solutions by liquid-liquid extraction for GC/FID analysis. For each sample, 2 mL of filtered culture were extracted with a mixture of 50 μ L phosphoric acid (1 M), 100 μ L NaCl (5 M) and 2 mL hexane. A 0.2 μ L sample was injected using a 40:1 split ratio on to a DB-FFAP column (J&W Scientific, 15 m x 250 μ m x 0.25 μ m), with a temperature profile starting at 150°C for 0.5 min, then heating at 15°C/min to 230°C and holding for 7.1 min (1.1 mL/min He).

[0098] GC results indicating secreted levels of free fatty acids after 216 hours are shown in Table 9-1.

Table 9-1
Medium-chain Fatty Acid Secretion (in mg/L) in Various *Synechocystis* sp. Strains

Strain	Parent Strain	Plasmid Added	Deletion	Transgenes	Fatty Acids	
					8:0	10:0
PCC 6803	n/a	n/a	None	None	ND	ND
SGC-BL33-1	PCC 6803	pSGI-BL33	SI10158 (<i>glgB</i>)	None	ND	ND
SGC-YC10-5	PCC 6803	pSGI-YC08	None	trcE-ChFatB2-7942	70.0	68.7
SGC-BL33-2	SGC-YC10-5	pSGI-BL33	SI10158 (<i>glgB</i>)	trcE-ChFatB2-7942	66.2	68.1

Note: ND represents “not detected” (<1 mg/L).

[0099] For a more optimized measurement of fatty acid secretion in these strains, the fatty acid secretion data shown in Table 9-1 were normalized to cell culture density, measured as optical density at 730 nm (OD_{730nm}); these data are presented in Table 9-2. Other experiments described in this application could be normalized in a similar fashion.

Table 9-2
Normalized Medium-chain Fatty Acid Secretion (mg/L/ OD_{730nm}) in Various *Synechocystis* sp. Strains

Strain	Parent Strain	Plasmid Added	Deletion	Transgenes	Fatty Acids	
					8:0	10:0
PCC 6803	n/a	n/a	None	None	ND	ND
SGC-BL33-1	PCC 6803	pSGI-BL33	SII0158 (<i>glgB</i>)	None	ND	ND
SGC-YC10-5	PCC 6803	pSGI-YC08	None	trcE-ChFatB2-7942	9.8	9.7
SGC-BL33-2	SGC-YC10-5	pSGI-BL33	SII0158 (<i>glgB</i>)	trcE-ChFatB2-7942	10.4	10.7

Note: ND represents “not detected” (<1 mg/L).

Example 10

Capture of Free Fatty Acids from Model Solutions with Hydrophobic Adsorbent Resins

[00100] A spike solution was formulated by dissolving 75 mg/L octanoic acid and 75 mg/L decanoic acid in BG-11 medium supplemented with 300 mM NaCl and adjusting the pH to 5.8. 50 mg of each of the resins listed in Table 1 were weighed into a 50 mL centrifuge tube and combined with 1.0 mL of methanol and shaken gently. The excess methanol was decanted and the resins were dried under a 25 in Hg vacuum, room temperature, overnight. 50 mL of the spike solution was then added to each of the resins and incubated with gentle shaking at 31°C for 24 hours. Following incubation, the resins were removed by filtering over a Whatman® GF/F glass fiber filter and the filtrates were analyzed for octanoic acid and decanoic acid content by gas chromatography as described in Example 2. The capacity of each resin for octanoic and decanoic acid could then be determined by the difference in the concentration of each fatty acid before and after incubation with each resin. The results are shown in Table 10-1 below.

Table 10-1

Adsorption capacities of several commercially-available adsorbents

Description	Resin type	Adsorption Capacity (mg/g)		
		Octanoic Acid	Decanoic acid	Total free fatty acids
Dowex Optipore [®] V503 (Dow Chemical)	Post cross-linked macroporous polystyrene divinyl benzene	26.3	69.8	96.0
Lewatit 1064 MD (LanXess)	Macroporous polystyrene divinyl benzene	1.1	46.7	47.8
Zeolyst CBV 28014 (Zeolyst)	Very low-alumina zeolite	17.4	74.7	92.0
Zeolyst CBV 901 (Zeolyst)	Low-alumina zeolite	5.4	64.8	70.1
Hisiv 3000 Silicalite (UOP Honeywell)	Hydrophobic silicalite	15.3	23.7	39.1
Lipidex 5000 (Packard Instrument Co.)	Alkylated sephadex gel	0.00	18.6	18.6
Norit ROW 0.8 (Fluka)	Extruded activated charcoal	40.2	71.8	112.1

[00101] Elution of free fatty acids from the hydrophobic adsorbents was also investigated. Dowex[®] Optipore[®] V503, Zeolyst CBV 28014, Zeolyst CBV 901, and Norit[®] ROW were incubated with 1.0 mL of spike solution per mg of adsorbent as described above. After the incubation period, the adsorbents were rinsed and combined with 0.1, 0.5, or 1.0 mL methanol per mg of adsorbent and shaken gently at room temperature for 4 hours. The methanol eluates and post-adsorption spikes were analyzed for free fatty acid concentration by gas chromatography. The results are listed in Table 10-2 below.

Table 10-2

Desorption of free fatty acids in methanol

mL MeOH/mg Resin	% Desorption		
	0.1 mL/mg	0.5 mL/mg	1.0 mL/mg
Dowex Optipore [®] V503	92%	84%	100%
CBV 28014	53%	76%	84%
CBV 901	78%	76%	57%
Norit [®] ROW	44%	85%	77%

[00102] The effect of pH on adsorbent capacity was studied utilizing Dowex[®] Optipore[®] V503. 40 mg of the resin were combined with 40 mL of BG-11 media spiked with 150 mg/L of octanoic and decanoic acid and adjusted to a pH of 10.0, 7.5, 4.8, or 2.8.

The pH 10 spike was buffered with 5 mM CAPS. The pH 7.5 and 2.8 spikes were buffered with 5 mM phosphate, and the pH 4.8 was buffered naturally by the dissolved fatty acids, with 5 mM NaCl added to maintain consistent conductivity. The spikes were incubated with resin as described above. Free fatty acid concentrations were measured with an enzymatic assay purchased from Zen-bio. The results are displayed in Table 10-3 below. From these results, it is clear that hydrophobic adsorption of free fatty acids is possible over a wide range of pH.

Table 10-3
Adsorption capacity of Dowex[®] Optipore[®] V503 at various pH values

pH	Adsorption Capacity (mg FFA/g resin)
10	42 ± 13
7.5	64 ± 4
4.8	172 ± 4
2.8	259 ± 1

Reported values are the mean of two experimental replicates, +/- one standard deviation.

Example 11

In Vivo Capture of Free Fatty Acids from Cultures of *Synechocystis* Strain SGC-YC10-5

[00103] *Synechocystis* sp. strain SGC-YC10-5, which contains the ChFatB2-7942 gene as described in Example 1, was cultured in BG-11 with and without Dowex[®] Optipore[®] V503 resin. 400 mL of fresh culture was induced with 5 mM IPTG and incubated at room temperature for 1 hour to allow for uptake of the inducer. The culture was then divided into four 1,000 mL baffled Erlenmeyer flasks with PTFE vent caps. To two of the flasks, approximately 400 mg of Dowex[®] Optipore[®] V503 were added. The adsorbent resin in the test flasks was recovered and exchanged for fresh resin daily for 10 days. The recovered resin was washed liberally with deionized water and eluted with 2 mL of methanol. Samples of culture medium from the test flasks and control flasks were also taken daily. The samples were measured for OD_{730nm} and filtered over a Whatman[®] GF/B glass fiber filter and analyzed for octanoic acid and decanoic acid content by gas chromatography as previously described in Example 2. The results are presented in Table 11-1.

Table 11-1

In vivo capture of free fatty acids from *Synechocystis* SGC-YC10-5 cultures

	Avg. Specific Growth Rate (d^{-1})	Average Free Fatty Acid Productivity ($\text{mg L}^{-1} \text{d}^{-1}$)
Without Dowex	0.090 ± 0.005	16 ± 0.8
With Dowex	0.090 ± 0.010	31 ± 3

Reported values are the mean of two biological duplicates +/- one standard deviation.

Example 12

Integration of CO₂ Delivery and Product Recovery as a Means for Enhancing the Efficiency and Economy of Both

[00104] Table 10-3 above reveals a clear relationship between free fatty acid adsorption capacity and pH. This relationship results from the inefficiency of extraction of the ionized form of the free fatty acids. Many potential production hosts require a pH significantly higher than the pKa of free fatty acids in order to survive and reproduce. An extreme example of this would be the alkalophilic cyanobacteria such as those belonging to the genera *Synechococcus*, *Synechocystis*, *Spirulina*, and many others, which prefer a pH between 9 and 11 for optimum growth. Figure 5 outlines an embodiment of the invention wherein this problem is solved by recycling a portion of the culture first through a vessel where it is contacted with concentrated CO₂ gas to lower the pH, then through a stationary adsorbent column wherein the protonated free fatty acids are captured.

[00105] The CO₂-enriched, free fatty acid-depleted suspension is then returned to the bulk culture. The pressure inside the gas-liquid contactor can be controlled independently to provide a constant pH in the stream exiting the adsorption column. Further, the pressure of the post-column flash vessel can be controlled so as to provide a supply of CO₂ which is titrated to the CO₂ consumption rate of the bulk culture through PID control of pH, dissolved CO₂, off-gas CO₂, or any combination of the three. The excess CO₂ can then be recycled.

[00106] In order to demonstrate proof of concept for the invention described above, an experimental system was constructed as displayed in Figure 5.

[00107] Vessel E-1 was filled with 4L of a spike solution containing 700 mg/L octanoic acid dissolved in 100 mM NaCl, pH 11.1. Column C1 was filled with 45.2 g of Dowex[®] Optipore[®] V503 polymeric resin. The resin was activated with two column volumes of

methanol, followed by a wash of three column volumes of 100 mM NaCl, pH 11.1. Liquid-gas contact vessel E2 was then filled with 200 mL of spike solution and 34.7 psia of CO₂. When the pH of the spike solution inside E-2 had decreased to between 5 and 6 (as determined by a slip of pH paper contained within E-2) peristaltic pumps P-1 and P-2 were set to the same flow rate and column loading was initiated. Valve V-2 was adjusted as needed to increase the column pressure and prevent the formation of gas bubbles.

[00108] Fractions of the flowthrough were taken at periodic intervals of 70-100 mL and assayed for octanoic acid by a commercially-available free fatty acid assay purchased from Zen-Bio. Two superficial linear flow rates were evaluated: 16.3 cm/min and 6.1 cm/min. For both flow rates, a control run was performed whereby vessel E-2 was bypassed and the column was loaded directly at a pH of 11.1. Table 12-1 below displays the results of this experiment. For both flow rates, column dynamic binding capacity was approximately 4-fold greater when CO₂ was used to lower the pH of the load.

Table 12-1

Dynamic binding capacity with and without CO₂-mediated load acidification

Flow velocity (cm/min)	Dynamic Binding Capacity (mg/g)	
	+34.7 psia CO ₂	Control (pH 11.1) 0 psia CO ₂
6.1	43.5	10.5
16.3	7.2	1.9

Example 13

Secretion of Oleic Acid by Photosynthetic Microorganisms

[00109] A synthetic gene that encodes a derivative of a FatA-type plant acyl-ACP TE enzyme with specificity for oleoyl-ACP is expressed in various photosynthetic microorganisms by constructing and utilizing expression vectors comprising a FatA gene operably linked to gene regulatory regions (promoters and terminators) that function in the host photosynthetic microorganism. The gene is optimized for expression in the host photosynthetic microorganism and the portion of the gene that encodes the plastid transit peptide region of the native FatA protein is removed for expression in cyanobacteria or replaced with a plastid transit peptide that functions effectively in the host eukaryotic photosynthetic microorganisms.

[00110] Genes that could be used for this purpose include, but are not limited to, those that encode the following acyl-ACP TEs (referred to by GenPept Accession Numbers): NP_189147.1, AAC49002, CAA52070.1, CAA52069.1, 193041.1, CAC39106, CAO17726, AAC72883, AAA33020, AAL79361, AAQ08223.1, AAB51523, AAL77443, AAA33019, AAG35064, and AAL77445.

[00111] The following is a sequence listing of all sequences referred to above.

SEQ ID NO:1

ATATCGTGGAGTATATCAATGGTGGGGAGGTGTGGTGTAGTAGTTGCGAGCAAAGATGACACTTGGTAAACTGATGCGACGTGGATACTGCGACGAAGATTGGCCGTACACACGTGCGATTGGAATGAACATATGTGTTTTATTCAAACCAATTTGACTAGTTTGGAGAACCTTCACGTGTTTCGCTCTCAAACCTTTGAGACAACAGCCTCCGAATCCAAATGAATGACTTTTAAACACAAGCTAGGAGCTGGTGATATATAATATGCTGGTTGTATGAAAGAGACTAATCGTGTGAAATAATGATGGCTCGCCCTAGTGAATGCTCCTCAGAGACGCTCATTTCGTCCAAGTGTTCTGTCACCTTCTGTCA TTGTTTCTCCGAGGCCAAGGTGGTCGAGTAGGTAGATACCAGCTATTCTCTTGCTTCTTTACTTTATCTCCC TCTACCAAAAACAGCACGTTATTATCTCCTTTCCATTCCACGCAATAACAAGAGGCAATCGGTAAAGAGGCACA AACAAGAGAACAAGACCCCGGCTGCTTCTCTCGTCCGTCCGCCGCCCTAAACTTCAAGTTTACTTCAAGTT CAATCTGTTTTTTGGCGCAAAAAGCGCGTTGCTCCGCCGTCTCCGCACTTTTCAGTTCTCTGTGCTCGAGGA CTGTTATCAACTTCCAAGATCTCCATCTCTTCTCTATCTCTCCCTAACAAAGTACG

SEQ ID NO:2

GTTCAATGCCTTTGGTGTGTCGTCAATAGGCACTTCGACTTTGCTCTTGTTCCGTTATCCCAAACCTTGAACG AGCGCCACGGTCTCTCGGTTTCGGTGGTATCCCAGGACCTCTCGTAGTTGATGCAGGGTTTCAAGATCGAGATA ACTCATGTTGTCGTTTGTGTTTGTGATTACCTTGCTTCCAGCTTTCGGTCTGTAATTACAGTGACACGC TGTACTAGAAATGATGTACGTTTGTGGAATCTCTAAAATTATGAGCTATTTATGAACACAGGAGTTCTCATCA ACTTTCCATCGAAATCCGTAGGAGAATTCTAATGTCTCTTCGGACGAGAGACAGACGTATCAGGAGTCACCTG AAGGTTCCAAGATTCTATCTTCATGAGGTCTGGATATGACAGTCTGCTTCGAGGCAAGCCCTGTCACTGTGA CCTTTTCGCTCGTCAATAATTTTAGGAACGCAAGGATAGGGATTCTCCATAGTAAGGACTATTGTTTGACCCC TGAAACTTCAACCTTTACCCCAAGAATGGGGCATTCTAAGTGAAAAACGTTTGTATGTATGCCCCAATTCTC ACACAGGAATAGGTATTGAATCACGTAGAAAATGATCGTTGCGCCGCAAGCAACACACCGGCTCTCTTCCGCC GCACCTCTCTTCCAATCCAACAAACAAACGCAACC

SEQ ID NO:3

ACGCAGATAGTGTATATTTGCGTCACAGTCTCTTGTGTCATAGGAGAGGAGAACTAGAGAACAAAAAGCGTCA TGTAATAAATGTTGGATGTTGGCATGTGTCGCCAGCCAGTATCCAAAACACCGAATTGTCGAGGTTTCGTGAGCT TGCAGCACTCATGGCAACGGCTAATTTTCATATCTATGTTATCAATGTTATCTGTAACACTAATGCTAAGTAATG CGTCAACAACCTTATCTCTCCGGCTCTTCACTCCACTTCGCTGACGTCGTTTGATATTTATCTGCTCTATTAT TCAAGTTGAATCTGCAGTTGAGGCATTCTCTAACTTAGCCGAGAAATCAAGACGGTGACTTTGAATTTACAAGT ACAGTTACGCTTACACAAGATACCTTTCTCACAAAAAAGATTCCGTTGGCTCCCACTGCGCATTGCTACTTTGGT ACTATTCCCATGTGGAACGTGATTGTTGGGGAAAGAGGGAGTCTGAGTTTGTAAATGTACATTTGTTATTCCCTT CATTATCGACAACATCACTAACTCATCGTGATACAGAGAAAAACAATCTCCACTTTCTCAACAAAAGTGGCCA CAATGTGCCTCCGACACAGCCTCAAGAGCCGACCGATCGTTGCATTTTTTCACTCTCGAACACACACACACAC ACACCCACACACCACCACCTCTCTTTATCCAACC

SEQ ID NO:4

AGTCGGATTGAAAACAGCGAATGTACGCCATTCCAAAGGCGCTCAGCAAAAGGAGACATATGCACACATCCAGC GGAAGTAAGTACGACACTTGAACAAGAGCATGACCTGTCAAAGCATGCTGCCATCGTCTGCTTCTATTCC CAATGACACTTTGGTCAACACGACTTGAAAAACGGCAATCAGCAAAATAAGCGATAGACCCTGACCAACGGCAG CTTTCATCTTTTATGAACGGCAGATATTGCGATCCTCTTTTATCGATACAGCAACACGCAGAAATTTCTGTTCT CTTTCAAGACGACAAGCAGAAATTTCCGTACGCTGTCATAATTTATTGACTATGTTAGATAACACAACCTCTCAT GCGCTTTGAAAATCTGCTTACTTCACAGTAAAGAGACAAGCTCTTTCGACTGACTGCGACAGAGATGGAAAAA GGAATTTTACCGGCAATTGACAGACTGATGTGAAAACAGAGAGTAACCGTAAACAAGTACCGGTAAGTATGCGC GCAACCTTTACTTGTTCCGTTGGCGTCTGTCAATTTGATGTCACGCAGACTTGAAAAGTCGTTGCGCTCCATTGTG AAAAAATATCATGCGACAACGTTTCAAGAGCCGGCGTCAATCGGTTTGCCTTGTTTCTGATCCGCTGCTTTTT GAGCAACGACCTGCGGAGGACCACAATGATCTTTCTCTGTCGTGAGAGCTAGTTCTATTACCTGTTCAATTAC CTGCTTTCTTGTATTACTCGAAGCTCTCGTTCTTCTATC

SEQ ID NO:5

TTTTGTAATTCGCCACTACCTTTACGCAAGTAAGAACGTTTCATGCTGGAGTCGTGGACCAATCGTAAGGTATA
CGTTAGTCATACCGCGCCTGTACTATTTACGACACGAGAGAAAGCCACTGCAGTTCTGGGATGGGATCAGATGC
TTGCTCCTTTCACTGCGCTGGCAAACGTATGCTAGACACGACTCGGATCGGATATCGAAATCAAACGGCGGAG
AATGGGTTTCGGATGACTGTCCGGAGCTACCTAGGAAAAGCTTCTTTTTCGTTTCGGACCACCAAGAGGGAAGCG
CTGCCTGTACTCGTGCATAGGAAGCATCAGACGTATTTGTTTCGGATGAGATCACACCAGAACTAGCCAGGCAG
CCAGCTAGCTATTGTCTATCTACAGATTTTGAACCAAACGTGGATACTAGAAAGCATGGGATTGACTGTGACTGT
GATTTGTGTTGCACACTTTATACCTACCCTCGACCTCGTACTTTGTGTAGTAGCAAAATGTGGATTGTGCGTTG
AAATGTAGAAGGGTTTGGGGTTGACACGGGTTTCATTATATCCGGGTACTCGAAAATGACCGCAACGATACTCA
TCGATCGAGATACGGTGTACACGTAGACTACGTAGAAAACCTACGAGGAAGCAGATATGATTTTCCGGTCCGCA
GCATCCACCCAGCCAACGTCCGCAAACAACCAAACAACCTCGTCGCCCCTTGTGTTCAAGATCTGCATTCCAT
TGACAGCCTTTTCAACGAAACCGTTTCGCTCGTTTGATTCCATACGTCTTTGAATACCAACAGAAAAAT

SEQ ID NO:6

AAAGTATCAATAGCTTATTCAGATTTTTGTGATGTTAGCCTACTTGTAAGCAGCGGAGGTCTGTCTATGACGG
TGATGTGGCTGGTTTCGCTCCGGAATTAAGTTCTGGTTTTATATCTCAACATAACTAGAGATAAAGTTACAGG
CACGTTACTGTAAGTCCGCAGATTGCTAATGCTTTGCTTCGGTGTCCGTAAAGCTTATGTTACTGTTCTAGATT
AGAGTGGTATCCACGATTTTCAAACGAAAGTGACATATTGCGAATTGTGCAGTATCAGAAAAATCTCCAAAGCAG
GAGCATACATTAGTTTGGCCGTATTGCAACGAGTAGCTCTCCTGAAGATGCAAGTAATAGAGGCTGTGAGCGTG
AATAATGAATTTGCCTGTTTAGAAGCTGGGGATCACATCTCGTGCTCCCCAAAAGTCTCTCAGTAAATCAAGAA
TGTTCTTATTTTCGAAAACATTGCTATTTATTTAGTTAACCGGCTTCGTCTCCCATTTAAATAAAGATTTTCA
AAAATGACACCACCAACGTCCGCAAGATCACGATTGAGAGGATTCTTCTTTGTCCCAACCATGGATGACCTCT
CCTATTAACACGTATATGAAGTACCGCTGCTGGTACCCGGAAAAGAGAGGACATTCTTGTGGGAGAGTCATCG
ATGCGCTGCCAATCGAAAAAATGCCAAGGCGAGAAAAGCGCAGTTCGTTCTTATAATCCAATTTTGAGTTTCA
AGACATACTCGTTGCTACCTTCCCACCTTCCCAACCAACCACTCGCAACC

SEQ ID NO:7

CCATGGCGAATGGTTCTGCAGTCTCTTTGAAATCTGGAAGCTTGAATACGCAGGAGGATACTAGTTCAGTCCC
CCTCCTCGGACGTTTTTGCATCAGCTGCCCCGACTGGAGTCGCTTGCTGACCGCCATCACAAACAGTGTGTTGTC
ATCTAAACGACCGGACATGCATGATCGGAAAAGCAAGCGCCAGATATGCTCGTCGATAGTTTCGGACTCGAGT
CTACTGTGCAGGACGGCCTGGTGTTCGTCATCCTTCAGCATCCGAAGCTACGAGATTGGTACGGACCGTACC
GCTAGCATTGAAACGTTGATGAACCATCTCCAAGAAACAGTTTGAACCACTGCAAGAGCACGGGCATCCTGCT
GGATGGTTTTGGCCGCACATTGGAAATGTGCAAGCGAGACTTGATCTGGGTGGTCATTAAAAATGCAGATCAAAG
TTAATCGATACCCGGCTGGGGAGATACCGTTGAGATCAATACACGCTTTTCCGTTTGGGCAAAATGGCGATG
GGTCGCGATTGGCTGATCTCCGACTGCAACACCGGTGAGATCTTGGTCCGTGCAACGCTCTGCGTACGCGATGAT
GAATCAAAAGACGCGTCGGTTGAGTAAGCTGCCGTATGAAGTTCACCAAGAAATTGTTCCATTGTTGTTGATA
GTCCCGTTATCGAGGATTCTGACCTCAAAGTCCACAAGTTTAAAGTCAAGACTGGCGATTCCATCCAGAAGGGC
CTGACGCCAGGTTGGAACGATCTGGATGTGAACCAACACGTTAGCAACGTTAAGTATATCGGCTGGATCTTGGA
AAGTATGCCTACGGAAGTCTGGAGACGCAGGAACCTGCGAGTCTCGCTCTGGAGTACCGCCGTGAGTGTGGCC
GTGATTCCGTGCTCGAGTCCGTCAGTGGACCTAGCAAAAGTGGGTGTTGCGAGTCAATACCAACACCTC
TTGCGGCTCGAAGATGGGACCGCCATTGTGAACGGCGGACCGAATGGCGCCCCAAAAATGCCGGCGCTAACGG
GGCAATTAGTACCGGGAAAACCTCCAATGGAACAGCGTCAGCTAATGATAGGATCC

SEQ ID NO:8

MANGSAVSLKSGSLNTQEDTSSSPPPRTFLHQLPDWSRLTLAITTVFVKSkrpdmHDRKSKRPDMLVDSFGLES
TVQDGLVFRQSFIRSIEIGTDRTASIEITLMNHLQETSLNHCKSTGILLDGFGRITLEMCKRDLIWWVIKMQIKV
NRYPAWGDVEINTRFSRLKIGMRDWLISDCNTGEILVRATSAYAMNQKTRRLSKLPYEVHQEIVPLFVDS
PVIEDSDLKVHKFKVKTGDSIQKGLTPGWNLDVNHVSNVKYIGWILESMPTEVLETQELCSLALEYRRECGR
DSVLESVTAMDPSKVGVRSQYQHLLRLEDGTAIVNGATEWRPNAGANGAISTGKTSNGNSVS

SEQ ID NO:9

CGCCGGGGCTGGCAGCTTAGTCCTGCGCAATCTCTACTACATCTGCCAACCCAGTGAAATTTTGATCTTTGCTG
GCAGTAGTCGCCGAGTAGTGATGGCCGCCGAGTTGGCTATCGCTTGGTCAAGGGCGGCAGCAGCCTGCGGGTA
CCTCTGCTGGAAAAAGCGCTCCGCATGGATCTGACCAACATGATCATTGAGTTGCGCGTTTCCAATGCCTTCTC
CAAGGGCGGCATTCCCTGACTGTTGAAGGCGTTGCCAATATCAAGATTGCTGGGGAAGAACCGACCATCCACA
ACGCGATCGAGCGGCTGCTTGCAAAAACCGTAAGGAAATCGAGCAAATTGCCAAGGAGACCCTCGAAGGCAAC
TTGCGTGGTGTGTTTAGCCAGCCTCACGCCGAGCAGATCAACGAGGACAAAATTGCCTTTGCCAAAAGTCTGCT
GGAAGAGGCGGAGGATGACCTTGAGCAGCTGGGTCTAGTCTCGATACGCTGCAAGTCCAGAACATTTCCGATG
AGGTGCGTTATCTCTCGGCTAGTGACGCAAGCAGCGGGCTGATCTGCAGCGAGATGCCCCGAATTGCTGAAGCC
GATGCCAGGCTGCCTCTGCGATCCAAACGGCCGAAAATGACAAGATCACGGCCCTGCGTCGGATCGATCGCGA
TGATAGCGATCGCCCAAGCCGAGGCCGAGCGCCGATTGAGGATGCGTTGACGCGGCGCAAGCGGTGGTGGCCG

AAGCTGAAGCGGACATTGCTACCGAAGTCGCTCGTAGCCAAGCAGAACTCCCTGTGCAGCAGGAGCGGATCAAA
CAGGTGCAGCAGCAACTTCAAGCCGATGTGATCGCCCCAGCTGAGGCAGCTTGTAACGGGCGATCGCGGAAGC
GCGGGGGGCGCCGCCCGTATCGTCGAAGATGGAAAAGCTCAAGCGGAAGGGACCCAAACGGCTGGCGGAGGCTT
GGCAGACCGCTGGTGCTAATGCCCCGCGACATCTTCTGCTCCAGAAGCTCGAAATTCGAGCTCGGTACCATTTA
CGTTGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGG
TGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGC
GTGGTGAACCGAGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGATGGCGGAGCTGAATTA
CATTCCCAACCGCGTGGCACAACAACCTGGCGGGCAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGG
CCCTGCACGCGCCGTCGCAAAATGTGCGGCGGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTG
TCGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGG
GCTGATCATTAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGT
TATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGC
GTGGAGCATCTGGTCGCTTGGGTACCAGCAAATCGCGCTGTTAGCGGGGCCATTAAATTCTGTCTCGGCGCG
TCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCGACT
GGAGTGCCATGTCCGTTTTCAACAAACCATGCAATGCTGAATGAGGGCATCGTTCCTACTGCGATGCTGGTT
GCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTGGTGCGGATATCTC
GGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAACCAACATCAAAACAGGATTTTC
GCCTGCTGGGGCAAACCGCTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTG
TTGCCCGTCTCACTGGTGAAGAAAAACCACCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGC
CGATTCAATTAATGCAGCTGGCAGCAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTA
AGTTAGCGCGAATTGATCTGGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCGTCAGGCA
GCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAATCACTGCATAATTCGTGTGCTCAAGGCGCACTCCC
GTTCTGGATAATGTTTTTTCGCGCGACATCATAACGGTTCTGGCAAATATTCTGAAATGAGCTGTTGACAATTA
ATCATCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGCGCCGCTGAGAAA
AAGCGAAGCGGCACTGCTCTTTAACAATTTATCAGACAATCTGTGTGGGCACTGACCGGAATTATCGATTAAC
TTTATTATAAAAAATTAAAGAGGTATATATTAATGTATCGATTAAATAAGGAGGAATAAACCATGCGCAATGGT
TCTGCAGTCTCTTTGAAATCTGGAAGCTTGAATACGCAGGAGGATACTAGTTCCAGTCCCCCTCCTCGGACGTT
TTTGCATCAGCTGCCCCGACTGGAGTCGCTTGCTGACCGCCATCACAACAGTGTGTTGTCAAATCTAAACGACCGG
ACATGCATGATCGGAAAAGCAAGCGCCAGATATGCTCGTCGATAGTTTTCGGACTCGAGTCTACTGTGCAGGAC
GGCCTGGTGTTCCTCAATCCTTCAGCATCCGAAGCTACGAGATTGGTACGGACCGTACCCTAGCATTGAAAC
GTTGATGAACCATCTCCAAGAAACAGTTTGAACCACTGCAAGAGCACGGGCATCCTGCTGGATGGTTTTGGCC
GCACATTGGAATGTGCAAGCGAGACTTGATCTGGGTGGTCATTAAAAATGCAGATCAAAGTTAATCGATACCCG
GCCTGGGGAGATACCGTTGAGATCAATACACGCTTTTCCCGTTTGGGCAAAATTGGCATGGGTGCGGATTTGGCT
GATCTCCGACTGCAACACCGGTGAGATCTTGGTCGTGCAACGCTGTCGTACGCGATGATGAATCAAAAGACGC
GTCGGTTGAGTAAGCTGCCGTATGAAGTTACCAAGAAATTTGTTCCATTGTTGCTGATAGTCCCGTTATCGAG
GATTCTGACCTCAAAGTCCACAAGTTTAAAGTCAAGACTGGCGATTCCATCCAGAAGGGCCTGACGCCAGGTTG
GAACGATCTGGATGTGAACCAACACGTTAGCAACGTTAAGTATATCGGCTGGATCTTGGAAGATATGCCTACGG
AAGTCTTGAGACGCAGGAATCTGCAGTCTCGCTCTGGAGTACCGCCGTGAGTGTGGCCGTGATTCCTGCTC
GAGTCCGTCACTGCGATGGACCCTAGCAAAGTGGGTGTTGCGAGTCAATACCAACACCTCTTGCAGGCTCGAAGA
TGGGACCGCCATTGTGAACGGCGCGACCGAATGGCGCCCCAAAAATGCCGGCGCTAACGGGGCAATTAGTACCG
GGAAAACCTCCAATGGAAACAGCGTCAGCTAATGATAGGATCCGAGCTCGAGATCTGCAGCTGGTACCATATGG
GAATTGCAAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCTGATACAGATTAATCAGAACGCAGAA
CGGCTCTGATAAAACAGAAATTTGCCTGGCGGAGTAGCGCGGTGGTCCACCTGACCCCATGCCAATCAGAA
GTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACCTGCCAGGCATCAATATA
AACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCTGTTTTATCTGTTGTTTGTGCGGTGAACGCTCTCCTGAGTAGG
ACAAATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCCGCCATA
AACTGCCAGGCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTTGCGTTTTCTACAAACTCTTTTTGT
TTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGGGGATCCGACTAGTAGGCCTCGAGGAATTCACGCG
TACGTAGATCTCCGCGGCCCGCGATCCTCTAGTATGCTTGTAACCGTTTTGTGAAAAAATTTTTAAATAAAAA
AAGGGGACCTCTAGGGTCCCCAATTAATTAGTAATATAATCTATTAAAGGTCATTCAAAGGTCATCCACCGGA
TCAGCTTAGTAAAGCCCTCGCTAGATTTTAAATGCGGATGTTGCGATTACTTCGCCAACTATTGCGATAACAAGA
AAAAGCCAGCCTTTTCATGATATATCTCCAATTTGTGTAGGGCTTATTATGCACGCTTAAAAATAATAAAGCA
GACTTGACCTGATAGTTTGGCTGTGAGCAATTATGTGCTTAGTGCTATCTAACGCTTGAGTTAAGCCGCGCCGCG
AAGCGGCGTCGGCTTGAACGAATTGTTAGACATTATTTGCCGACTACCTTGGTGATCTCGCCTTTCACGTAGTG
GACAAATTCTTCCAATGATCTGCGCGCGAGGCCAAGCGATCTTCTTCTTGTCCAAGATAAGCCTGTCTAGCTT
CAAGTATGACGGGCTGATACTGGGCGGCGAGGCGCTCCATTGCCAGTGGCGAGCGACATCCTTCGGCGCGATT
TTGCCGGTTACTGCGCTGTACCAAAATGCGGGACAACGTAAGCACTACATTTGCTCATCGCCAGCCAGTCGGG
CGGCGAGTTCCATAGCGTTAAGGTTTTCATTTAGCGCTCAAATAGATCCTGTTGAGGAACCGGATCAAAGAGTT
CCTCCGCCGCTGGACCTACCAAGGCAACGCTATGTTCTCTTGTCTTTGTGTCAGCAAGATAGCCAGATCAATGTCG
ATCGTGGCTGGCTCGAAGATACCTGCAAGAATGTCATTGCGCTGCCATTCTCCAAATTGCAGTTTCGCGCTTAGC
TGGATAACGCCACGGAATGATGTGCTGTCACAACAATGGTGACTTCTACAGCGCGGAGAATCTCGCTCTCTC

CAGGGGAAGCCGAAGTTTCCAAAAGGTCGTTGATCAAAGCTCGCCGCGTTGTTTTCATCAAGCCTTACGGTCACC
GTAACCAGCAAATCAATATCACTGTGTGGCTTCAGGCCGCCATCCACTGCGGAGCCGTACAAATGTACGGCCAG
CAACGTCGGTTCGAGATGGCGCTCGATGACGCCAACTACCTCTGATAGTTGAGTCGATACTTCGGCGATCACC
CTTCCCTCATGATGTTTAACTTTGTTTTAGGGCGACTGCCCTGCTGCGTAACATCGTTGCTGCTCCATAACATC
AAACATCGACCCACGGCGTAACCGCGCTTGCTGCTTGGATGCCCCGAGGCATAGACTGTACCCCCAAAAAACAGTC
ATAACAAGCCATGAAAACCGCCACTGCGCCGTTACCACCGCTGCGTTCGGTCAAGGTTCTGGACCAAGTTGCGTG
AGCGCATACGCTACTTGCAATTACAGCTTACGAACCGAACAGGCTTATGTCCACTGGGTTCTGTCCTTCATCCGT
TTCCACGGTGTGCGTCACCCGGCAACCTTGGGCAGCAGCGAAGTCGAGGCATTTCTGTCTGGCTGGCGAACGA
GCGCAAGGTTTTCGGTCTCCACGCATCGTCAGGCATTGGCGGCCCTTGCTGTTCTTCTACGGCAAGGTGCTGTGCA
CGGATCTGCCCTGGCTTCAGGAGATCGGAAGACCTCGGCCGTGCGGGCGCTTGCCGGTGGTGCTGACCCCGGAT
GAAGTGGTTTCGCATCCTCGGTTTTCTGGAAGGCGAGCATCGTTTGTTCGCCAGCTTCTGTATGGAACGGGCAT
GCGGATCAGTGAGGGTTTTGCAACTGCGGGTCAAGGATCTGGATTTTCGATCACGGCACGATCATCGTGCGGGAGG
GCAAGGGCTCCAAGGATCGGGCCTTGATGTTACCCGAGAGCTTGGCACCCAGCCTGCGCGAGCAGGGGAATTGA
TCCGGTGGATGACCTTTTGAATGACCTTTAATAGATTATATTACTAATTAATTGGGGACCCCTAGAGGTCCCCTT
TTTTATTTTAAAAATTTTTTCAAAAACGGTTTACAAGCATAAAGCTCTAGAGTCGACCTGCAGGCATGCAAGC
TTCGAGTCCCCTGCTCGTCACGCTTTTACGGCACCGTGCCAGATATCGACGTGGAGTCGATCACTGTGATTGGCGA
AGGGGAAGGCAGCGCTACCCAAATCGCTAGCTTGCTGGAGAAGCTGAAACAAACCACGGGCATTGATCTGGCGA
AATCCCTACCGGTCAATCCGACTCGCCGCTGCGAAGTCCTAAGAGATAGCGATGTGACCGCGATCGCTTGTC
AAGAATCCCAGTGATCCCGAACCATAGGAAGGCAAGCTCAATGCTTGCTCGTCTTGAGGACTATCTAGATGTC
TGTGGAACGCACATTTATTGCCATCAAGCCCGATGGCGTTTACGCGGGTTTTGGTTCGGTACGATCATCGGCCGCT
TTGAGCAAAAAGGCTTCAAACCTGGTGGGCCTAAAGCAGCTGAAGCCCAGTCGCGAGCTGGCCGAACAGCACTAT
GCTGTCCACCGCGAGCGCCCTTCTTCAATGGCCTCGTCGAGTTTCATCACCTCTGGGCGGATCGTGCGCATCGT
CTTGGAAGGCGAAGGCGTTTGTGGCGGCTGCTCGCAAGTTGATCGGCGCTACCAATCCGCTGACGGCAGAACC
GCACCATCCGTGGTGATTTTGGTGTCAATATTGGCCGCAACATCATCCATGGCTCGGATGCAATCGAAACAGCA
CAACAGGAAATTGCTCTCTGTTTTAGCCAGCAGAGCTAAGTGATTGGACCCACGATTCAACCCCTGGCTGTA
CGAATAAGGTCTGCATTCTTCAGAGAGACATTGCCATGCCCTGCGCGTCTGCGATCGCCCTTCCAAGCTTGCC
CCGCTGTTTTCGGGCTGGCAGCCCTGGCGTTGGGGCTGGCGACCGCTTGCCAAGAAAGCAGCGCTCCGCCGGCTG
CCGATC

SEQ ID NO:10

CGCCGGGGCTGGCAGCTTAGTCCTGCGCAATCTCTACTACATCTGCCAACCCAGTGAAATTTTGATCTTTGCTG
GCAGTAGTCGCCGAGTAGTGATGGCCGCCGAGTTGGCTATCGCTTGGTCAAGGGCGGCAGCAGCCTGCGGGTA
CCTCTGCTGGAAAAAGCGCTCCGCATGGATCTGACCAACATGATCATTGAGTTGCGCGTTTCCAATGCCTTCTC
CAAGGGCGGATTTCCCCTGACTGTTGAAGCGTTTGCCAAATCAAGATTGCTGGGGAAGAACCAGCACTCCACA
ACGCGATCGAGCGGCTGCTTGGCAAAAACCGTAAGGAAATCGAGCAAAATTGCCAAGGAGACCCTCGAAGGCAAC
TTGCGTGGTGTTTTAGCCAGCCTCACGCCGAGCAGATCAACGAGGACAAAATTGCCTTTGCCAAAAGTCTGCT
GGAAGAGGCGGAGGATGACCTTGAGCAGCTGGGTCTAGTCTCGATACGCTGCAAGTCCAGAACATTTCCGATG
AGGTTCGGTTATCTCTCGGCTAGTGACGCAAGCAGCGGGCTGATCTGCAGCGAGATGCCGAATTGCTGAAGCC
GATGCCCAGGCTGCCTCTGCGATCCAAACGGCCGAAAAATGACAAGATCACGGCCCTGCGTCGGATCGATCGCGA
TGATAGCGATCGCCCAAGCCGAGGCCGAGCGCCGGATTAGGATGCGTTGACGCGGCGCGAAGCGGTGGTGCCG
AAGCTGAAGCGGACATTGCTACCGAAGTCGCTCGTAGCCAAGCAGAACTCCCTGTGCAGCAGGAGCGGATCAAA
CAGGTGCAGCAGCAACTTCAAGCCGATGTGATGCCCCAGCTGAGGCAGCTTGTAACGGGCGATCGCGGAAGC
GCGGGGGCGCCCGCGGTATCGTCGAAGATGGAAGCTCAAGCGGAAGGGACCCACGGCTGGCGGAGGCTT
GGCAGACCGCTGGTGCTAATGCCCCGCGACATCTTCTGCTCCAGAAGCTCGAAATTCGAGCTCGGTACCATTTA
CGTTGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGG
TGGTGAATGTGAAACCAGTAACGTTATACGATGTCGAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGC
GTGGTGAACCAGGCCAGCCAGTCTTCTGCGAAAACGCGGAAAAAGTGGAAGCGGCGATGGCGGAGCTGAATTA
CATTTCCCAACCGCGTGGCACAACAACCTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGG
CCCTGCACGCGCCGTCGCAAAATTGTGCGGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTG
TCGATGGTAGAACGAAGCGGCTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGG
GCTGATCATTAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGT
TATTTCTTGATGTCTTGACAGACACCCATCAACAGTATTATTTTCTCCATGAAGACGGTACGCACTGGGCG
GTGGGAGCATCTGGTCGATTGGGTCAACAGCAATCGCGCTGTAGCGGGCCATTAGTTCTGTCTCGGCGCG
TCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCGACT
GGAGTGCCATGTCCGGTTTTCAACAAACCATGCAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTT
GCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCGGATATCTC
GGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAACCACCATCAAACAGGATTTTC
GCCTGCTGGGGCAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTG
TTGCCCGTCTCACTGGTGAAAAGAAAAACCCCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGC
CGATTCAATTAATGCAGCTGGCAGCAGGTTTTCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTA
AGTTAGCGCAATTGATCTGGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCGTCAGGCA

GCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAAATCACTGCATAAATTCGTGTCGCTCAAGGCGCACTCCC
GTTCTGGATAATGTTTTTTCGCGCCGACATCATAACGGTTCTGGCAAATATTCTGAAATGAGCTGTTGACAATTA
ATCATCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGACCATGGCGAATG
GTTCTGCAGTCTCTTTGAAATCTGGAAGCTTGAATACGAGGAGGATACTAGTTCCAGTCCCCCTCCTCGGACG
TTTTTGCATCAGCTGCCCCGACTGGAGTGCCTTGCTGACCGCCATCACAACAGTGTTTGTCAAATCTAAACGACC
GGACATGCATGATCGGAAAAAGCAAGCGCCCCAGATATGCTCGTCGATAGTTTCGGACTCGAGTCTACTGTGCAGG
ACGGCCTGGTGTTCCTCAATCCTTCAGCATCCGAAGCTACGAGATTGGTACGGACCGTACCGCTAGCATTGAA
ACGTTGATGAACCATCTCCAAGAAACAGTTTTGAACCACTGCAAGAGCACGGGCATCCTGCTGGATGGTTTTTG
CCGCACATTGGAATGTGCAAGCGAGACTTGATCTGGGTGGTCATTAAAATGCAGATCAAAGTTAATCGATACC
CGGCTGGGGAGATACCGTTGAGATCAATACACGCTTTTCCCGTTTGGGCAAAATTTGGCATGGGTGCGGATTGG
CTGATCTCCGACTGCAACACCGGTGAGATCTTGGTCCGTGCAACGTCTGCGTACGCGATGATGAATCAAAAGAC
GCGTCGGTTGAGTAAGCTGCCGTATGAAGTTCACCAAGAAATTGTTCCATTGTTCTGTTGATAGTCCCGTTATCG
AGGATTCTGACCTCAAAGTCCACAAGTTTAAAGTCAAGACTGGCGATTCCATCCAGAAGGGCCTGACGCCAGGT
TGGAACGATCTGGATGTGAACCAACACGTTAGCAACGTTAAGTATATCGGCTGGATCTTGAAAGTATGCCTAC
GGAAGTCTTGAGACGCGAGGAATCTGCAGTCTCGCTCTGGAGTACCGCCGTGAGTGTGGCCGTGATTCCGTGC
TCGAGTCCGTCAGTCGATGGACCTAGCAAAGTGGGTGTTGCGAGTCAATACCAACACCTCTTGGCGGCTCGAA
GATGGGACCGCCATTGTGAACGGCGCGACCGAATGGCGCCCCAAAAATGCCGGCGCTAACGGGGCAATTAGTAC
CGGGAACCTCCAATGGAACAGCGTCAGCTAATGATAGGATCCGAGCTCGAGATCTGCAGCTGGTACCATAT
GGGAATTCGAAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTACGCTGATACAGATTAAATCAGAACGCAG
AAGCGGTCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACCTCAG
AAGTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAAT
AAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGTGCGGTGAACGCTCTCCTGAGTA
GGACAAATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAACGGCCCCGAGGGTGGCGGGCAGGACGCCGCCCA
TAAACTGCCAGGCATCAAATTAAGCAGAAGGCCATCTGACGGATGGCCTTTTTGCGTTTCTACAAACTCTTTT
GTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGGGGATCCGACTAGTAGGCCTCGAGGAATTCACG
CGTAGCTAGATCTCCGCGCGCCGATCCTCTAGTATGCTTGTAAACCGTTTTGTGAAAAAATTTTTAAATAA
AAAAGGGGACCTCTAGGGTCCCCAATTAATTAGTAATATAATCTATTAAAGGTCATTCAAAGGTCATCCACCG
GATCAGCTTAGTAAAGCCCTCGCTAGATTTTAAATGCGGATGTTGCGATTACTTCGCCAACTATTGCGATAACAA
GAAAAAGCCAGCCTTTTCATGATATATCTCCCAATTTGTGTAGGGCTTATTATGCACGCTTAAAAATAATAAAG
CAGACTTGACCTGATAGTTTGGCTGTGAGCAATTATGTCTTAGTGCATCTAACGCTTGAGTTAAGCCGCGCCG
CGAAGCGGCGTCCGGCTTGAACGAATTGTTAGACATTATTTGCCGACTACCTTGGTGATCTCGCCTTTCACGTAG
TGGACAAATTTCTCCAATGATCTGCGCGCGAGGCCAAGCGATCTTCTTCTTGTCCAAGATAAGCCTGTCTAGC
TTCAAGTATGACGGGCTGATACTGGGCGCGCAGGCGCTCCATTGCCAGTCGGCAGCGACATCTTCCGGCGCGA
TTTTTCCGGTTACTGCGCTGTACCAATGCGGGACAACGTAAAGCACTACATTTGCTCATCGCCAGCCCACTCG
GGCGGCGAGTTCCATAGCGTTAAGGTTTTCAATTTAGCGCCTCAAATAGATCCTGTTTCAGGAACCGGATCAAAGAG
TTCTCCGCGCTGGACCTACCAAGGCAACGCTATGTTCTCTTGTCTTTGTGTCAGCAAGATAGCCAGATCAATGT
CGATCGTGGCTGGCTCGAAGATACCTGCAAGAATGTCAATTGCGCTGCCATTCTCCAAATTCAGTTCGCGCTTA
GCTGGATAACGCCACGGAATGATGTGCTGTCGTCACAACAATGGTGACTTCTACAGCGCGGAGAATCTCGCTCTC
TCCAGGGGAAGCCGAAGTTTCCAAAAGGTCGTTGATCAAAGCTCGCCGCGTTGTTTCATCAAGCCTTACGGTCA
CCGTAACCAGCAAATCAATATCACTGTGTGGCTTCAGGCCGCCATCCACTGCGGAGCCGTACAAATGTACGGCC
AGCAACGTGGTTTCGAGATGGCGCTCGATGACGCCAACTACCTCTGATAGTTGAGTCGATACTTCGGCGATCAC
CGCTTCCCTCATGATGTTTAACTTTGTTTTAGGGCAGTGCCTGCTGCGTAACATCGTTGCTGCTCCATAACA
TCAAACCTGACGCCACGGCGTAACGCGCTTGTCTGTTGATGCCCGAGGCATAGACTGACCCCAAGCAAGACAG
TCATAACAAGCCATGAAAACCGCCACTGCGCCGTTACCAACCGCTGCGTTCCGGTCAAGGTTCTGGACCAAGTGGC
TGAGCGCATACGCTACTTGCAATTACAGCTTACGAACCGAACAGGCTTATGTCCACTGGGTTCTGTGCTTTCATCC
GTTTCCACGGTGTGCGTCACCCGGCAACCTTGGGCGAGCAGCGAAGTCGAGGCATTTCTGTCTGGCTGGCGAAC
GAGCGCAAGGTTTTCGGTCTCCACGCATCGTCAGGCATTGGCGGCCCTTGTGTTCTTCTACGGCAAGGTGCTGTG
CACGGATCTGCCCTGGCTTCAGGAGATCGGAAGACCTCGGCCGTGCGGCGCTTGCCTGGTGTGCTGACCCCGG
ATGAAGTGGTTCGCATCCTCGGTTTTTCTGGAAGGCGAGCATCGTTTGTTCGCCCAGCTTCTGTATGGAACGGGC
ATGCGGATCAGTGAGGGTTTGAACCTGCGGGTCAAGGATCTGGATTTTCGATCAGGGCAGCATCATCGTGGGGA
GGGCAAGGGCTCCAAGGATCGGGCCTTGATGTTACCCGAGAGCTTGGCACCCAGCCTGCGCGAGCAGGGGAATT
GATCCGGTGGATGACCTTTTGAATGACCTTTAATAGATTATATTACTAATTAATTGGGGACCTTAGAGGTCCCC
TTTTTTATTTTAAAAATTTTTTACAAAACGGTTTACAAGCATAAAGCTCTAGAGTCGACCTGCAGGCATGCAA
GCTTCGAGTCCCTGCTCGTCACGCTTTCAGGCACCGTGCCAGATATCGACGTGGAGTCGATCACTGTGATTGGC
GAAGGGGAAGGCAGCGCTACCCAAATCGCTAGCTTGTGGAGAAGCTGAAACAAACCACGGGCATTGATCTGGC
GAAATCCCTACCGGTCAATCCGACTCGCCCGCTGCGAAGTCTAAGAGATAGCGATGTGACCGCGATCGCTTG
TCAAGAATCCAGTGATCCCGAACCATAGGAAGGCAAGCTCAATGCTTGCCTCGTCTTGAGGACTATCTAGATG
TCTGTGGAACGCACATTTATTGCCATCAAGCCGATGGCGTTTACGCGGGGTTTTGGTCCGTACGATCATCGGCCG
CTTTGAGCAAAAAGGCTTCAAACCTGGTGGGCCTAAAGCAGCTGAAGCCCAGTCGCGAGCTGGCCGAACAGCACT
ATGCTGTCCACCGCAGCGCCCTTCTTCAATGGCCTCGTCGAGTTTATCACCTCTGGGCCGATCGTGGCGATC
GTCTTGGAAGGCGAAGGCTTGTGGCGGCTGCTCGCAAGTTGATCGGCGCTACCAATCCGCTGACGGCAGAACC

GGGCACCATCCGTGGTGATTTTGGTGTCAATATTGGCCGCAACATCATCCATGGCTCGGATGCAATCGAAACAG
CACAACAGGAAATTGCTCTCTGGTTTAGCCAGCAGAGCTAAGTGATTGGACCCCCACGATTCAACCCTGGCTG
TACGAATAAGGTCTGCATTCCCTCAGAGAGACATTGCCATGCCCGTGCTGCGATCGCCCTTCCAAGCTGCCTTG
CCCCGCTGTTTCGGGCTGGCAGCCCTGGCGTTGGGGCTGGCGACCGCTTGCCAAGAAAGCAGCGCTCCGCCGGC
TGCCGGATC

SEQ ID NO:11

CTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAG
GCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGC
CAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCAAAAAATC
GACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTTCCCCTGGAAGCTCCCTC
GTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCT
TTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGGTGTAGGTGCTTCGCTCCAAGCTGGGCTGTGTGCACGAAC
CCCCCGTTACAGCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTA
TCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAA
GTGGTGGCCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG
GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGGTTTGAAGCAG
CAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAAC
CGAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTA
AATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAG
GCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCGCTGACTCCCCGTCGTGTAGATAACTACGAT
ACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACC GGCTCCAGATTTAT
CAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCTGCAACTTTATCCGCTCCATCCAGTCT
ATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCCCGAGTTAATAGTTTGCACAACGTTGTTGCCATTGCTAC
AGGCATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTGAGTCCGGTTCCCAACGATCAAGGCGAGTTA
CATGATCCCCCATGTTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTTTGTGAGAAGTAAGTTGGCC
GCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTATGCCATCCGTAAGATGCTTTTC
TGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGT
CAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAGTGCTCATCATTGGAACGTTCTTCGGGGCGA
AACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGACCCAACTGATCTTCAGC
ATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGG
CGACACGGAAATGTTGAATACTCATACTCTTCTTTTTCAATATTATTGAAGCATTATCAGGGTTATTGTCTC
ATGAGCGGATAACATATTTGAATGTATTTAGAAAAATAAACAATAAGGGGTTCCGCGCACATTTCCCCGAAAAGT
GCCACCTGACGCTCAAGAAACCATTTATCATGACATTAACCTATAAAAAATAGGCGTATCAGAGGCCCTTTC
GTCTCGCGCTTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCAAGCTTGTCTG
TAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGTGCGGGCTGGCTTAA
CTATGCGGCATCAGAGCAGATTGTACTGAGAGTGACCATAAAATGTAAACGTTAATATTTGTTAAATTCG
CGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAA
GAATAGCCCCGAGATAGGGTTGAGTGTTGTTCCAGTTTGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAA
CGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCAAATCAAGTTTTTTGG
GGTCGAGGTGCCGTAAAGCACTAAATCGGAACCTAAAGGGAGCCCCGATTTAGAGCTTGACGGGGAAAGCCG
CGCAACGTGGCGAGAAAGGAAGGAAGGAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGCTAGCGGTAC
GCTGCGCTGAACCAACACCCGCGCGCTTAATGCGCGCTACAGGGCGCTACTGTTGTTGACGTAT
GCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAAATACCGCATCAGGCGCCATTGCGCAATTGCTGCGCA
ACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGG
CGATTAAGTTGGGTAAACGCCAGGGTTTTCCAGTCACGACGTTGTAAACGACGGCCAGTGCCAAGCTATTGCT
GAAGCGGAATCCCTGGTTAATGCCGCCGCGATGCCAATTGCATTCTCCAAGTGGGGCACATTGAACGCTTCAA
CCCGGCATTTTTAGAGCTAACC AAAATTTCTAAAACGGAAGAGTTATTGGCGATCGAAGCCCATCGCATGAGTC
CCTATTCCCAGCGGGCCAATGATGTCTCCGTGGTATTGGATTTGATGATCCATGACATTGACCTGTTGCTGGAA
TTGGTGGGTTCCGAAGTGGTTAAACTGTCCGCCAGTGGCAGTGGGCTTCTGGGTGAGGATATTTGGATTATGT
CACCCTACGTTAGGCTTCTCCTCCGGCATTTGTTGCCACCTCACC GCCAGTAAGGTACCCATCGTAAATTC
GTTCCATCGCCGCCACTGCAAAAATTCCTCACC GAAGCGGATTTTCTCAATAACGAAATTTTGATCCATCGC
CAAAACACCGCTGATTGGAGCGCGGACTATGGCCAGGTATTGTATCGCCAGGATGGTCTAATCGAAAAGGTTTA
CACCAGTAATATTGAACCTCTCCACGCTGAATTAGAACATTTTATTTCATTGTGTTAGGGGAGGTGATCAACCCT
CAGTGGGGGGGAGAACAGGCCCTCAAGGCCCTGAAGTTAGCCAGTTTAATTGAAGAAATGGCCCTGGACAGTCAG
GAATGGCATGGGGGGGAAGTTGTGACAGAATATCAAGATGCCACCCTGGCCCTCAGTGCGAGTGTTTAAATCAA
CTTAATTAATGCAATTATTGCGAGTTCAAACCTCGATAACTTTGTGAAATATTACTGTTGAATTAATCTATGACT
ATTCAATACACCCCTAGCCGATCGCCTGTTGGCTACCTCGCCGCCGATCGCCTAAATCTCAGCGCCAAGAG
TAGTTCCCTCAACACCAGTATTCTGCTCAGCAGTGACCTATTCAATCAGGAAGGGGAATTGTAACAGCCAAC
ATGGCTTTGATGGTTATATGGTACCATATGCATGCGAGCTCAGATCTACCAGGTTGTCCTTGGCGCAGCGCTTC
CCACGCTGAGAGGGTGTAGCCCGTCACGGGTAACCGATATCGTCGACAGGCCTCTAGACCCGGGCTCGAGCTAG

CAAGCTTGGCCGGATCCGGCCGGATCCGGAGTTTGTAGAAAACGCAAAAAGGCCATCCGTCAGGATGGCCTTCTG
CTTAATTTGATGCCTGGCAGTTTATGGCGGGCGTCTGCCCCGCCACCCTCCGGGCGGTGCTCCGCAACGTTCA
AATCCGCTCCCGGGCGGATTGTCTACTCAGGAGAGCGTTACCGACAAACAACAGATAAAACGAAAGGCCAG
TCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCTGGCAGTTCCCTACTCTCGCATGGGGAGACCCACACTAC
CATCGGCGCTACGGCGTTTCACTTCTGAGTTCGGCATGGGGTCAGGTGGGACCACCGCGCTACTGCCGCCAGGC
AAATTCTGTTTTATTGAGCCGTTACCCACCTACTAGCTAATCCCATCTGGGCACATCCGATGGCAAGAGGCC
GAAGGTCCCCCTCTTTGGTCTTGCGACGTTATGCGGTATTAGCTACCGTTTCCAGTAGTTATCCCCCTCCATCA
GGCAGTTTTCCAGACATTACTCACCCGTCCGCCACTCGTCAGCAAGAAGCAAGCTTAGATCGACCTGCAGGGG
GGGGGGGAAAGCCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAAATATATCATCATGAA
CAATAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCTTGC
TCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCA
ATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTTCTGAAACATGGCAAAGGTA
GCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCATC
AAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGGAAAAACAGCATTCCAGGT
ATTAGAAGAATATCCTGATTGAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCTGCGCCGGTTGCATTGCA
TTCTGTGTTGTAATTGTCTTTAACAGCGATCGCGTATTTCTGCTCGCTCAGGCGCAATCACGAATGAATAAC
GGTTTGTTGATGCGAGTGATTTTATGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAAGAAATGCA
TAAGCTTTTTGCCATTCTCACCGGATTGAGTCGTCATGCTGATGTTTCTCACTTGATAACCTTATTTTTGACG
AGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTA
TGGAAGTGCCTCGGTGAGTTTTCTCTTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGA
TATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTCTAATCAGAATTGGTTAATTGGTTGTAACACT
GGCAGAGCATTACGCTGACTTGACGGGACGGCGGCTTTGTTGAATAAATCGAATTTTGTCTGAGTTGAAGGATC
AGATCACGCATCTTCCCGACAACGCAGACCGTTCCGTGGCAAAGCAAAAGTTCAAAATCACCAACTGGTCCACC
TACAACAAAGCTCTCATCAACCGTGGCTCCCTCACTTTCTGGCTGGATGATGGGGCGATTGAGGCTGGTATGA
GTCAGCAACACCTTCTTACGAGGACAGCTCAGCGCCCCCCCCCTGCAGGTCGATGTTGTAACCCGACGG
CGGTTGCTACCAAGTAGTGACCCGCTTCGTGATGCAAAATCCGCTGACGATATTCCGGCGCATCGCTGTAATG
CCATCGAGCAGTAACGTGGCGAATTCCGTACCGGTATGGATGGCACCGATGCGGAATCCCAACAGATTGCCTTT
GACAACAATGTGGCCTGGAATAACCTGGGGGATTTGTCCACCACCACCAACGGGCCTACACTTCGGCTATTAG
CACAGACACAGTGCAGAGTGTATGCGGTTAATCTGAAAAAAACGATAACATTCCCATTGTTTTTGCCTGGC
CCATTTTTCCACCACCCTTAATCCACAGATTTTTCAGTAATGCTTAACACGGGGGAAATTGTACCCCGGTG
ATCGCCTCTTTGATTCCCAACAGTGAATACAACGAACGGCAACGGTAGTAATTACGGGCAATTTTGGTAATCG
TTTAACCCACAGGCACGGAGGGAGCGATTTATCCCGTTTCCGTAGGCACAGTGTGGACAGTACTCCTTTGAAA
TGGTGGGACCCAACGGCCCGGTGAGTGGCGGTGGGTATTACCATTGATAGTCTCAACCCCTACGTGGCCGGCAAT
GGTCCCAAAATTGTCGCCGCTAAGTTAGACCGCTTCAGTGACCTGGGGGAAGGGGCTCCCTCTGGTTAGCCAC
CAATCAAAATAACAGTGGCGGGGATTTATATGGAGACCAAGCCCAATTTCTGTTTGCGAATTTACACCAAGCGCCG
GTTTTTCCCCCGATGGCATTGCCAGTTTACTACCCACAGAATTTGAACGGTATTTTCAACTCCAAGCGGAAGAT
ATTACGGGACGGACAGTTATCCTAACCCAACTGGTGTGATTATGAAATTTCCCGGCTTTGGTCTGGTGCAGGT
GTTGGGGCTGGCGGATTTGGCCGGGGTTCAGGACAGCTATGACCTGACTTACATCGAAGATCATGACAATATT
ACGACATTATCCTCAAAGGGGACGAAGCCGCGATTCGCCAAATTAAGAGGGTTGCTTTGCCCTCCGAAGGGGAT
TATTCGGCGGTTTTATAATCCCGGTGGCCCCGGCAATGATCCAGAGAATGGTCCCCCAAATTCGTAATCATGTCA
TAGCTGTTTTCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTA
AGCCTGGGGTGCCTAATGAGTGAGCTAATCATTAAATTGCGTTGCGCTCACTGCCCCGCTTCCAGTCGGGAA
ACCTGTCTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCT

SEQ ID NO:12

ATTGCTGAAGCGGAATCCCTGGTTAATGCCGCCGCCGATGCCAATTGCATTCTCCAAGTGGGGCACATTGAACG
CTTCAACCCGGCATTTTTAGAGCTAACCAAAATTCTCAAAACGGAAGAGTTATTGGCGATCGAAGCCCATCGCA
TGAGTCCCTATTCCAGCGGGCCAATGATGTCTCCGTGGTATTGGATTGATGATCCATGACATTGACCTGTTG
CTGGAATTGGTGGGTTCCGGAAGTGTTAACTGTCCGCCAGTGGCAGTCGGGCTTCTGGGTGAGGATATTTGGA
TTATGTCACCGCTACGTTAGGCTTCTCCTCCGGCATTGTGGCCACCCTCACCGCCAGTAAGGTACCCATCGTA
AAATTCGTTCCATCGCCGCCACTGCAAAAATTCCTCACCGAAGCGGATTTTCTCAATAACGAAATTTTGATC
CATCGCAAACACCGCTGATTGGAGCGCGGACTATGGCCAGGTATTGTATCGCCAGGATGGTCTAATCGAAAA
GGTTTACACCAAGTAATATTGAACCTCTCCACGCTGAATTAAGAACATTTTATTCAATTGTGTTAGGGGAGGTGATC
AACCCTCAGTGGGGGAGAACAGGCCCTCAAGGCCCTGAAGTTAGCCAGTTTAATTGAAGAAATGGCCCTGGAC
AGTCAGGAATGGCATGGGGGGGAAGTTGTGACAGAATATCAAGATGCCACCCTGGCCCTCAGTGCAGTGTTTA
AATCAACTTAATTAATGCAATTATTGCGAGTTCAAACCTGATAACTTTGTGAAATATTACTGTTGAATTAATCT
ATGACTATTCAATACACCCCTTAGCCGATCGCCTGTGGCCTACCTCGCCGCCGATCGCCTAAATCTCAGCGC
CAAGAGTAGTTCCCTCAACACCAAGTATTCTGCTCAGCAGTGACCTATTCAATCAGGAAGGGGGGAATTGTAACAG
CCAATATGGCTTTGATGGTTATATGGTACCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGT
TAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGGTGATGGCTGCGCCCCGACACCCGCCAACACCCGCTG
ACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATG

TGTCAGAGGTTTTTACCCTCATCACCGAAACGCGCGAGGCAGCAGATCAATTTCGCGCGCGAAGGCGAAGCGGCA
TGCAATTTACGTTGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCA
ATTCAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCGAGATATGCCGGTGTCTCTTATCAGACCG
TTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGAAGCGGCGATGGCGGAG
CTGAATTACATTCCCAACCGCGTGGCACAACAACCTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTC
CAGTCTGGCCCTGCACGCGCCGTGCGAAATTGTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCG
TGGTGGTGTGATGGTAGAACGAAGCGGCGTGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGC
GTCAGTGGGCTGATCATTAACCTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAAGCTGCCTGCACTAATGT
TCCGGCGTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGC
GACTGGGCGTGGAGCATCTGGTTCGATTGGGTACCAGCAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCT
TCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGGA
AGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGA
TGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGCTGCGCGTTGGTGCG
GATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAACCACCATCAAACA
GGATTTTCGCTGTGGGGCAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCA
ATCAGCTGTTGCCGTCTCACTGGTGAAAAGAAAAACCCCTGGCGCCCAATACGCAAACCGCCTCTCCCGCG
CGGTTGGCCGATTTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAA
TTAATGTAAGTTAGCGCGAATTGATCTGGTTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGC
GTCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAAGTTCGTAATCACTGCATAATTCTGTGTGCTCAAGGC
GCACTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCATAACGGTTCTGGCAAATATCTGAAATGAGCTGTT
GACAATTAATCATCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGCGCCG
CTGAGAAAAAGCGAAGCGGCACTGCTCTTTAACAATTTATCAGACAATCTGTGTGGGCACTCGACCGGAATTAT
CGATTAACCTTTATTTAAAAATTAAGAGGTATATATTAATGTATCGATTAATAAGGAGGAATAAACCATGG
CGAATGGTTCTGCAGTCTCTTTGAAATCTGGAAGCTGAATACGCAGGAGGATACTAGTTCAGTCCCCCTCTCT
CGGACGTTTTTGCATCAGCTGCCCCGACTGGAGTCTGTTGCTGACCGCCATCACAACAGTGTGTTGTCAAACTAA
ACGACCGGACATGATGATCGGAAAAAGCAAGCGCCAGATATGCTCGTCGATAGTTTTCGGACTCGAGCTACTAG
TGCAGGACGGCCTGGTGTTCCTGCAATCCTTCAGCATCCGAAGCTACGAGATTGGTACGGACCGTACCGCTAGC
ATTGAAACGTTGATGAACCATCTCCAAGAAACAGTTTGAACCACTGCAAGAGCACGGGCATCCTGCTGGATGG
TTTTGGCCGCACATTGGAATGTGCAAGCGAGACTTGATCTGGGTGGTCAATTAATGCGATCAAAGTTAATC
GATACCCGGCTGGGGAGATACCGTTGAGATCAATACAGCTTTTTCCCGTTTGGGCAAATTTGGCATGGGTGCG
GATTGGCTGATCTCCGACTGCAACACCGGTGAGATCTTGGTCCGTGCAACGTCTGCGTACCGGATGATGAATCA
AAAGACGCGTCGGTTGAGTAAGCTGCCGTATGAAGTTACCAAGAAATTTGTTCCATTGTTGTTGATAGTCCCG
TTATCGAGGATCTGACCTCAAAGTCCACAAGTTTAAAGTCAAGACTGGCGATTCCATCCAGAAGGGCCTGACG
CCAGGTTGGAACGATCTGGATGTGAACCAACAGTTAGCAACGTTAAGTATATCGGCTGGATCTTGGAAAGTAT
GGTACGGAAGTCCCTGGAGACGCAAGAACTCTGCACTCTCGCTCTGGAGTACCGCCGTGAGTGTGGCCGTGATT
CCGTGCTCGAGTCCGTCACTGCGATGGACCCTAGCAAAGTGGGTGTTGCGAGTCAATACCAACACCTCTTGCGG
CTCGAAGATGGGACCGCCATTGTGAACGGCGCGACCGAATGGCGCCCCAAAAATGCCGGCGCTAACGGGGCAAT
TAGTACCGGGAAAACCTCCAATGGAACAGCGTCAGCTAATGATAGGATCCGAGCTCAGATCTACCAGGTTGTC
CTTGGCGCAGCGCTTCCACGCTGAGAGGGTGTAGCCCGTCACGGGTAACCGATATCGTCGACAGGCCTCTAGA
CCCGGGCTCGAGCTAGCAAGCTTGGCCGGATCCGGCCGGATCCGGAGTTTGTAGAAACGCAAAAAGGCCATCCG
TCAGGATGGCCTTCTGCTTAATTTGATGCCTGGCAGTTTATGGCGGGCGTCTGCCCGCCACCTCCGGGCCGT
TGCTTCGAAACGTTCAAATCCGCTCCCGGCGGATTTTGCTCTACTCAGGAGAGCGTTACCGGACAAACAACAGAT
AAAACGAAAGGCCAGTCTTTCGACTGAGCCTTTTCGATTTTATTTGATGCCTGGCAGTTCCCTACTCTCGCATGG
GGAGACCCACACTACCATCGGCGCTACGGCGTTTCACTTCTGAGTTTCGGCATGGGTCAGGTGGGACCAACCGC
GCTACTGCCGCCAGGCAAATTTCTGTTTTATTGAGCCGTTACCCACCTACTAGCTAATCCCATCTGGGCACATC
CGATGGCAAGAGGCCCCGAAGGTCCCCCTCTTTGGTCTTTCGACGTTATGCGGTATTAGCTACCGTTTCCAGTAG
TTATCCCCCTCCATCAGGCAGTTTCCAGACATTACTACCCGTCCGCCACTCGTCAGCAAAGAAGCAAGCTTA
GATCGACCTGCAGGGGGGGGGGGGAAAGCCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAA
AATATATCATCATGAACAATAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCA
ACGGGAAACGTCCTTGCTCGAGGCGCGGATTAATTTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTC
GCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTG
AAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAACTGGCTGAGCAATTTAT
GCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCCTGCGATCCCCGGGA
AAACAGCATTCCAGGTATTAGAAGAATATCCTGATTACAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTG
CGCCGTTGCAATTCGATTCTGTTTGTAAATTGTCCTTTTAACAGCGATCGCGTATTTTCGTCTCGCTCAGGCGCA
ATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTGTGACGAGCGTAATGGCTGGCCTGTTGAACAAG
TCTGGAAAGAAATGCATAAGCTTTTGCCATTCTACCCGATTACAGTCGTCACTCATGGTGATTTCTCACTTGAT
AACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCA
GGATCTTGCCATCCTATGGAAGTGCCTCGGTGAGTTTTCTCCTTCAATACAGAAACGGCTTTTTCAAAAATATG
GTATTGATAATCCTGATATGAATAAATTGCAGTTTCAATTGATGCTCGATGAGTTTTTCTAATCAGAATTGGTT
AATTGGTTGTAACACTGGCAGAGCATTACGCTGACTTGACGGGACGGCGGCTTTGTTGAATAAATCGAACTTTT

GCTGAGTTGAAGGATCAGATCACGCATCTTCCCGACAACGCAGACCGTTCCGTGGCAAAGCAAAAGTTCAAAAT
CACCAACTGGTCCACCTACAACAAAGCTCTCATCAACCGTGGCTCCCTCACTTTCTGGCTGGATGATGGGGCGA
TTCAGGCCTGGTATGAGTCAGCAACACCTTCTTACAGAGGCAGACCTCAGCGCCCCCCCCCCCCCTGCAGGTGCGA
TCTGGTAACCCAGCGCGGTTGCTACCAAGTAGTGACCCGCTTCGTGATGCAAAATCCGCTGACGATATTCGGG
CGATCGCTGCTGAATGCCATCGAGCAGTAACGTGGCGAATTCGGTACCGGTATGGATGGCACCGATGCGGAATC
CCAACAGATTGCCTTTGACAACAATGTGGCCTGGAATAACCTGGGGGATTGTGCCACCACCACCCAACGGGCCT
ACACTTCGGCTATTAGCACAGACACAGTGCAGAGTGTATGGCGTTAATCTGGAAAAAACGATAACATTCCC
ATTGTTTTTGGCTGGCCCATTTTTCCCAACCACCTTAATCCCAACAGATTTTCAGGTAATGCTTAACACGGGGGA
AATTGTACCCCCGGTGATCGCCTCTTTGATTCCCAACAGTGAATACAACGAACGGCAAACGGTAGTAATTACGG
GCAATTTTGGTAATCGTTTAACCCAGGCACGGAGGGAGCGATTTATCCCGTTTCCGTAGGCACAGTGTGGAC
AGTACTCCTTTGGAAATGGTGGGACCCAACGGCCCGTCACTGCGGTGGGTATTACCATTGATAGTCTCAACCC
CTACGTGGCCGGCAATGGTCCCAAAATTGTGCGCGCTAAGTTAGACCGCTTCAGTGACCTGGGGGAAGGGGCTC
CCCTCTGGTTAGCCACCAATCAAAATAACAGTGGCGGGGATTTATATGGAGACCAAGCCCAATTTTCGTTTGCGA
ATTTACACCAGCGCCGGTTTTTCCCCGATGGCATTGCCAGTTTACTACCCACAGAATTTGAACGGTATTTTCA
ACTCCAAGCGGAAGATATTACGGGACGGACAGTTATCCTAACCCAAACTGGTGTGATTATGAAATTTCCGGCT
TTGGTCTGGTGCAGGTGTGGGGCTGGCGGATTTGGCCGGGGTTCAGGACAGCTATGACCTGACTTACATCGAA
GATCATGACAACTATTACGACATTATCCTCAAAGGGGACGAAGCCGCAAGTTTCGCCAAATTAAGAGGGTTGCTTT
GCCCTCCGAAGGGGATTATTCGGCGGTTTATAATCCCGTGGCCCCGGCAATGATCCAGAGAATGGTCCCCCA

SEQ ID NO:13

ATTGCTGAAGCGGAATCCCTGGTTAATGCCGCCGCGATGCCAATTGCATTCTCCAAGTGGGGCACATTGAACG
CTTCAACCCGGCATTTTTAGAGCTAACCAAAATTCTCAAACCGGAAGAGTTATTGGCGATCGAAGCCCATCGCA
TGAGTCCCTATTCCCAGCGGGCCAATGATGTCTCCGTGGTATTGGATTTGATGATCCATGACATTGACCTGTTG
CTGGAATTGGTGGGTTTCGGAAGTGGTTAACTGTCCGCCAGTGGCAGTGGGGCTTCTGGGTCAAGATATTTGGA
TTATGTCAACCGCTACGTTAGGCTTCTCCTCCGGCATTGTGGCCACCCTCACCAGTAAAGTCAACCCATCGTA
AAATTCGTTCCATCGCCGCCCACTGCAAAAATTCCCTCACCAGAAAGCGGATTTTCTCAATAACGAAATTTTGATC
CATCGCCAAACCACCGCTGATTGGAGCGCGGACTATGGCCAGGTATTGTATCGCCAGGATGGTCTAATCGAAAA
GGTTTACACCAGTAATATTGAACCTCTCCACGCTGAATTAGAACATTTTATTCAATTGTGTAGGGGAGGTGATC
AACCCTCAGTGGGGGGAGAACAGGCCCTCAAGGCCCTGAAGTTAGCCAGTTTAATTGAAGAAATGGCCCTGGAC
AGTCAGGAATGGCATGGGGGGGAAGTTGTGACAGAATATCAAGATGCCACCCTGGCCCTCAGTGCAGAGTGTTA
AATCAACTTAATTAATGCAATTATTGCGAGTTCAAACTCGATAACTTTGTGAAATATTACTGTTGAATTAATCT
ATGACTATTCAATACACCCCTTAGCCGATCGCCTGTTGGCCTACCTCGCCGCCGATCGCCTAAATCTCAGCGC
CAAGAGTAGTTCCCTCAACACCAAGTATTCTGCTCAGCAGTGACCTATTCAATCAGGAAGGGGGAATTTGAACAG
CAAACCTATGGCTTTGATGGTTATATGGTACCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGATAGT
TAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGGTTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTG
ACGCGCCCTGACGGGCTTGCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATG
TGTCAGAGGTTTTTACCCTCATCACCGAAACGCGGAGGCAGCAGATCAATTCGCGCGCGAAGGCGAAGCGGCA
TGCATTTACGTTGACACCATCGAATGGTGCAAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGAGTCA
ATTCAGGGTGGTGAATGTGAAACAGTAACGTTATACGATGTGCGCAGAGTATGCCGGTGTCTCTTATCAGACCG
TTTCCCGCGTGGTGAACCAAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAAGCGCGCATGGCGGAG
CTGAATTACATTCCCAACCGCTGGCACAACAACCTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTC
CAGTCTGGCCCTGCACGCGCCGTGCGAAATTGTGCGGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCG
TGGTGGTGTGATGTTAGAACGAAGCGCGCTCGAAGCCTGTAAAGCGGGCGGTGCACAACTCTTCGCGCAACGCG
GTCAGTGGGCTGATCATTAACCTATCCGCTGGATGACCAAGTATGCCATTGCTGTGGAAAGCTGCCTGCACTAATGT
TCCGGCGTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGC
GACTGGGCGTGGAGCATCTGGTGCATTGGGTACCAGCAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCT
TCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGGA
AGGCGACTGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGAATGAGGGCATCGTTCCCACTGCGA
TGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTGTGGTGCG
GATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAACCACCATCAAACA
GGATTTTCGCTGCTGGGGCAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCA
ATCAGCTGTTGGCCGTCTCACTGGTGAAAAGAAAAACACCCCTGGCGCCCAATACGCAACCCGCTCTCCCCGC
GCGTTGGCCGATTCAATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAA
TTAATGTAAGTTAGCGCGAATTGATCTGGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGC
GTCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAAGTTCGTAATCACTGCATAATTCTGTGCTGCTCAAGGC
GCACTCCCGTCTGGATAATGTTTTTTCGCGCCGACATCATAACGGTCTGGCAAATATCTGAAATGAGCTGTT
GACAATTAATCATCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGACCAT
GGCGAATGGTTCTGCAGTCTCTTTGAAATCTGGAAGCTTGAATACGCAGGAGGATACTAGTTCAGTCCCCCTC
CTCGGACGTTTTTGCATCAGCTGCCCCGACTGGAGTTCGCTTGCTGACCGCCATCACAACAGTGTGTGCAATCT
AAACGACCGGACATGCATGATCGGAAAAGCAAGCGCCAGATATGCTCGTCGATAGTTTCGGACTCGAGTCTAC
TGTGCAGGACGGCTGGTGTTCGTCATCTTCAGCATCCGAAGCTACGAGATTGGTACGGACCGTACCGCTA

GCATTGAAACGTTGATGAACCATCTCCAAGAAACCAGTTTGAACCACTGCAAGAGCACGGGCATCCTGCTGGAT
GGTTTTGGCCGCACATTGGAATGTGCAAGCGAGACTTGATCTGGGTGGTCATTAAAAATGCAGATCAAAGTTAA
TCGATACCCGGCCTGGGGAGATACCGTTGAGATCAATACACGCTTTTCCCGTTTGGGCAAAATTGGCATGGGTC
GCGATTGGCTGATCTCCGACTGCAACACCGGTGAGATCTTGGTCCGTGCAACGTCTGCGTACGCGATGATGAAT
CAAAAGACGCGTCGGTTGAGTAAGCTGCCGTATGAAGTTCACCAAGAAATTGTTCCATTGTTGCTTGATAGTCC
CGTTATCGAGGATTCTGACCTCAAAGTCCACAAGTTTAAAGTCAAGACTGGCGATTCCATCCAGAAGGGCCTGA
CGCCAGGTTGGAACGATCTGGATGTGAACCAACACGTTAGCAACGTTAAGTATATCGGCTGGATCTTGGAAAGT
ATGCCTACGGAAGTCCTGGAGACGCAGGAACCTCTGCAGTCTCGCTCTGGAGTACCGCCGTGAGTGTGGCCGTGA
TTCCGTGCTCGAGTCCGTCACTGCGATGGACCCTAGCAAAGTGGGTGTTGCGAGTCAATACCAACACCTCTTGC
GGCTCGAAGATGGGACCGCCATTGTGAACGGCGCGACCGAATGGCGCCCCAAAAATGCCGGCGCTAACGGGGCA
ATTAGTACCGGGAAAACCTCCAATGGAACAGCGTCAGTAATGATAGGATCCGAGCTCAGATCTACCAGGTTG
TCCTTGGCGCAGCGCTTCCCACGCTGAGAGGGTGTAGCCCGTCACGGGTAACCGATATCGTCGACAGGCCTCTA
GACCCGGGCTCGAGCTAGCAAGCTTGGCCGGATCCGGCCGGATCCGGAGTTTGTAGAAACGCAAAAAGGCCATC
CGTCAGGATGGCCTTCTGCTTAATTTGATGCCTGGCAGTTTATGGCGGGCGTCTGCCCGCCACCCTCCGGGCC
GTTGCTTCGCAACGTTCAAATCCGCTCCCGCGGATTGTGCTACTCAGGAGAGCGTTACCGACAAACAACAG
ATAAAACGAAAGGCCAGTCTTTCGACTGAGCCTTTCGTTTTATTTGATGCCGTGGCAGTTCCTACTCTCGCAT
GGGGAGACCCACACTACCATCGGCCTACGGCGTTTCACTTCTGAGTTCGGCATGGGGTCAGGTGGGACCACC
GCGCTACTGCCGCCAGGCAAATCTGTTTTATTGAGCCGTTACCCACCTACTAGCTAATCCCATCTGGGCACA
TCCGATGGCAAGAGGGCCGAAGGTCCCCCTCTTTGGTCTTGCGACGTTATGCGGTATTAGCTACCGTTTCCAGT
AGTTATCCCCCTCCATCAGGCAGTTTCCCAGACATTACTACCCGTCGCCCACTCGTCAGCAAAGAAGCAAGCT
TAGATCGACCTGCAGGGGGGGGGGGGAAAGCCACGTTGTGTCTCAAATCTCTGATGTTACATTGCACAAGATA
AAAATATATCATCATGAACAATAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATT
CAACGGGAAACGTCCTTGCTCGAGGCCGCGATTAAATTCACATGGATGCTGATTTATATGGGTATAAATGGGC
TCGCGATAATGTGGGCAATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTT
TGAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTT
ATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCCTGCGATCCCCGG
GAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTACAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTC
TGCGCCGTTGCAATTCGATTCTGTTTGTAAATTGCTCTTTAACAGCGATCGCGTATTTCTGCTCGCTCAGGCG
CAATCACGAATGAATAACGTTTGGTTGATGCGAGTGATTTTATGATGACGAGCGTAATGGCTGGCCTGTTGAACA
AGTCTGGAAAGAAATGCATAAGCTTTTGCCATTCTACCGGATTACAGTCGTCACTCATGGTGATTTCTCACTTG
ATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATAC
CAGGATCTTGCCATCCTATGGAAGTGCCTCGGTGAGTTTTCTCCTTCAATACAGAAACGGCTTTTTCAAATA
TGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTCTAATCAGAATTGG
TTAATTGGTTGTAACACTGGCAGAGCATTACGCTGACTTGACGGGACGGCGGCTTTGTTGAATAAATCAACTT
TTGCTGAGTTGAAGGATCAGATCACGCATCTTCCCGACAACGCAGACCGTTCCGTGGCAAAAGCAAAAGTTCAAA
ATCACCAACTGGTCCACCTACAACAAAGCTCTCATCAACCGTGGCTCCCTCACTTTCTGGCTGGATGATGGGGC
GATTACGGCCTGGTATGAGTCAGCAACACCTTCTTACGAGGCAGACCTCAGCGCCCCCCCCCCCCCTGCAGGTC
GATCTGGTAACCCAGCGCGGTTGCTACCAAGTAGTGACCCGCTTCGTGATGCAAAATCCGCTGACGATATTCTG
GGCGATCGCTGCTGAATGCCATCGAGCAGTAACGTGGCGAATTCGGTACCGGTATGGATGGCACCAGATGCGGAA
TCCCAACAGATTGCCTTTGACAACAATGTGGCCTGGAATAACCTGGGGGATTTGTCCACCACCACCCAACGGGC
CTACACTTCGGCTATTAGCACAGACACAGTGCAGAGTGTTTATGGCGTTAATCTGGAAAAAACGATAACATTC
CCATTGTTTTTGGCTGGCCCATTTTTCCACCACCCTTAATCCACAGATTTTCAGGTAATGCTTAACACGGGG
GAAATTGTCACCCCGGTGATCGCCTCTTTGATTCCCAACAGTGAATACAACGAACGGCAACCGGTAGTAATTAC
GGGCAATTTTGGTAATCGTTTAACCCAGGCAGGAGGAGCGATTTATCCCGTTTCCGTAGGCACAGTGTGG
ACAGTACTCCTTTGGAATGGTGGGACCCAACGGCCCGTCACTGCGGTGGGTATTACCATGATAGTCTCAAC
CCCTACGTGGCCGGCAATGGTCCCAAAATTTGTCGCCGCTAAGTTAGACCGCTTCAGTGACCTGGGGGAAGGGC
TCCCTCTGGTTAGCCACCAATCAAATAACAGTGGCGGGGATTTATATGGAGACCAAGCCCAATTTCTGTTTGC
GAATTTACACCAGCGCCGTTTTTCCCCCGATGGCATTGCCAGTTTACTACCCACAGAATTTGAACGGTATTTT
CAACTCCAAGCGGAAGATATTACGGGACGGACAGTTATCCTAACCCAAACTGGTGTGATTATGAAATTCCCGG
CTTTGGTCTGGTGCAGGTGTTGGGGCTGGCGGATTTGGCCGGGGTTTACAGACAGCTATGACCTGACTTACATCG
AAGATCATGACAATATTACGACATTATCCTCAAAGGGGACGAAGCCGCAGTTCGCCAAATTAAGAGGGTTGCT
TTGCCCTCCGAAGGGGATTATTGCGCGGTTTATAATCCCGGTGGCCCCGGCAATGATCCAGAGAATGGTCCCC
A

SEQ ID NO:14

ACCCTGGCCCTCAGTGCAG

SEQ ID NO:15

TGCTTCTTTGCTGACGAGTGG

SEQ ID NO:16

GTGACTGGAACCGCCCTCG

SEQ ID NO:17

CCATCGAGCAGTAACGTGGCCGATAGTGACGCTAAACCAGGCTG

SEQ ID NO:18

CGAGTGGCGGACGGGTGAGTCTACGAGGGCGTGCAGAAGC

SEQ ID NO:19

CACCAAGTTGCCTTCACCGAC

SEQ ID NO:20

CAGCCTGGTTTAGCGTCACTATCGGCCACGTTACTGCTCGATGG

SEQ ID NO:21

GCTTCTGCACGCCCTCGTAGACTCACCCGTCCGCCACTCG

SEQ ID NO:22

GTGACTGGAACCGCCCTCGCGCAACCCCGCGCCATTACGCCCCACGAACAGCAGCTTTTGGCCAAACTGAAAAG
CTATCGCGATATCCAAAGCTTGTCGCAAATTTGGGGACGTGCTGCCAGTCAATTTGGATCGATGCCGGCTTTGG
TTGCACCCCATGCCAAACCAGCGATCACCCCTCAGTTATCAAGAATTGGCGATTTCAGATCCAAGCGTTTGCAGCC
GGACTGCTCGCGCTGGGAGTGCCTACCTCCACAGCCGATGACTTTCCGCCTCGCTTGGCGCAGTTTGCAGGATAA
CAGCCCCCGCTGGTTGATTGCTGACCAAGGCACGTTGCTGGCAGGGGCTGCCAATGCGGTGCGCGGCGCCCAAG
CTGAAGTATCGGAGCTGCTCTACGTCTTAGAGGACAGCGGTTTCGATCGGCTTGATTGTCGAAGACGCGGCGCTG
CTGAAGAACTACAGCCTGGTTTAGCGTCACTATCGGCCACGTTACTGCTCGATGGCATTTCAGCAGCGATCGCC
CGAATATCGTCAGCGGATTTTGCATCACGAAGCGGGTCACTACTTGGTAGCAACCGCGCTGGGGTTACCAGATC
CGTCGATCATATCGTCAATTATTACCTCCACGGGGAGAGCCTGAGCAAACCTGGCCTCAGGCATTTGAGAAGCAC
ACGGTCACACTGCTTCCGGTAGTCAATAAACCGGTAAACCAGCAATAGACATAAGCGGCTATTTAACGACCCTG
CCCTGAACCGACGACCGGGTCGAATTTGCTTTTCGAATTTCTGCCATTTCATCCGCTTATTATCACTTATTCAGGC
GTAGCACCAGGCGTTTAAAGGCGACCAATAACTGCCTTAAAAAATTACGCCCGGCCCTGCCACTCATCGCAGTA
CTGTTGTAATTCATTAAAGCATTCTGCCGACATGGAAGCCATCACAAACGGCATGATGAACCTGAATCGCCAGCG
GCATCAGCACCTTGTGCGCTTGCGTATAATTTTGCCCATGGTGAAAACGGGGGCGAAGAAGTTGTCCATATTG
GCCACGTTTAAATCAAACCTGGTGAACTCACCAGGGATTGGCTGAGACGAAAAACATATTCTCAATAAACCC
TTTAGGGAAATAGGCCAGGTTTTCACCGTAACAGCCACATCTTGCGAATATATGTGTAGAACTGCCGGAAT
CGTCGTGGTATTCACTCCAGAGCGATGAAAACGTTTCAGTTTGCTCATGGAAAACGGTGTAACAAGGGTGAACA
CTATCCCATATCACCAGCTCACCCTCTTTTATTGCCATACGGAATTCCGGATGAGCATTTCATCAGGCGGGCAAG
AATGTGAATAAAGGCCGGATAAAACTTGTGCTTATTTTTCTTTACGGTCTTTAAAAAGGCCGTAATATCCAGCT
GAACGGTCTGGTTATAGGTACATTGAGCAACTGACTGAAATGCCTCAAATGTTCTTTACGATGCCATTGGGAT
ATATCAACGGTGGTATATCCAGTGATTTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAAATCTCGATAACTC
AAAAAATACGCCCGGTAGTGATCTTATTTTATTATGGTGAAAGTTGGAACCTCTTACGTGCCGATCAACGTCTC
ATTTTCGCCAAAAGTTGGCCAGGGCTTCCCGGTATCAACAGGGACACCAGGATTTATTTATCTGCGAAGTGA
TCTTCCGTCACAGGTATTTATTGAAAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTAAATGTCATGA
TAATAATGGTTTTCTTAGACGTACGGTGGCACTTTTCGGGGAATGTGCGCGGAACCCCTATTTGTTTTATTTTTT
TAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAA
GAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTC
ACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGAACTGGAT
CTCAACAGCGGTAAGATCCTTGAGAGTTTTTCGCCCCGAAGAAGCTTTTCCAATGATGAGCACTTTTAAAGTTCT
GCTATGTGGCGCGGTATTATCCCGTGTGACGGATCTAAGCTTGCTTCTTTGCTGACGAGTGGCGGACGGGTGAG
TCTACGAGGGCGTGCAGAAGCAGTTTCGCGAGCAACCGGCGAAGAAACGTGCTTGATCGATACTTCTTTGGC
TTGAGTCAACGCTATGTTTTGGCACGGCGCCGCTGGCAAGGACTGGATTTGCTGGCACTGAACCAATCCCCAGC
CCAGCGCCTCGCTGAGGGTGTCCGGATGTTGGCGCTAGCACCGTTGCATAAGCTGGGCGATCGCCTCGTCTACG
GCAAAGTACGAGAAGCCACGGGTGGCCGAATTCGGCAGGTGATCAGTGGCGGTGGCTACTGGCACTGCACCTC
GATACCTTCTTCGAAATTGTTGGTGTTGATTTGCTGGTGGGTATGGCTTGACAGAAACCTCACCAGTGCTGAC
GGGGCGACGGCCTTGGCACAACTACGGGGTTCGGCCGCTCAGCCGATTCCAGGTACGGCGATTTCGGATCGTCG
ATCCTGAAACGAAGGAAAACCGACCCAGTGGCGATCGCGGCTTGGTGCTGGCGAAAGGGCCGAAATCATGCAG
GGCTACTTCAATAAACCCGAGGCGACCGCGAAAGCGATCGATGCCGAAGGTTGGTTTGACACCGGCGACTTAGG
CTACATCGTCGGTGAAGGCAACTTGGTG

SEQ ID NO:23

CTCGAGCCCCCGTGCTATGACTAGC

SEQ ID NO:24

CTCGAGCCCCGGAACGTTTTTTGTACCCC

SEQ ID NO:25

CAATTGGTCACACGGGATAATACCGCGCC

SEQ ID NO:26

CAATTGGTCGATCATATCGTCAATTATTACCTCCAC

SEQ ID NO:27

CCCCCGTGCTATGACTAGCGGCGATCGCCATACCGGCCACGACCATTTGCATTGGATCCCCAACGGCGGCCACA
ACTTCCATGGCATTGAGATGCGGGGAATGATGTTCTAGACTCTGACGCACCAAAGCCAATTTTTGTTGATGGTT
GCAATGGGGATGACTACTGTTCACTTTGCCCCAGCGTCAATGCCTAGACCTAGCAGTACCCCCAGGGCTGTGG
TAGTGCCCCCACCACGCATTTCGCTTAGCACTAAGTAACTTTGGCGATGTTCTCGGGCTAACTGTGCGCCCCAC
TGCAAACCTGCTGAAAAAGATGCTCCACCAGGGCCAACGGTAACGCTTGCCCTGTGGAAAGACAGCGGGCGGG
TTGTCCGTCTAGATTGATGACTGGCACCCTGGGGGAATGGGTAAACCAGAGTTAAATAAATAAACCGGAGTAT
GGAGGGCATCCACCAACGCTTTGGTGATGAACACTGGGGAAACCCAGAAATGAGGGGAGGTAAGGGATAGGTT
GCCCCTGCCGTAGTTCCCTTGATTAAAAATTCCGCATCGGCGATCGCCGTCAATTTTCGATCAGCGGGGTTTT
ACCCGCCGCAGAAATGCCCGGAATTAACACAGTTTCCGTAAAGCCCAACACACAGACAAAACCGGTGGACAGT
GGCCATGGCGCTCAATCCAGGATAAAGCTTGGTCAGACTGGGTATAAACTGTCAACATATTTCTGCAAGAGTGG
GCCAATTGGGAAAATCAACCTCAAATCCATTGGAATAGCCTTTTTTCAACCGTAAAAATCCAACCTTTCTCTCT
TCCCTTCTTCTTCCATCTGATTATGGTTACGCCAATTAACCTACCATTCCATCCATTGCCTGGCGGATATCTGG
GCTATCACCGGAGAAAAATTTGCCGATATTGTGGCCCTCAACGATCGCCATAGTCATCCCCCGTAACCTTTAAC
CTATGCCCAATTGGTCACACGGGATAATACCGCGCCACATAGCAGAACTTTAAAGTGCTCATCATTGGAAAAAC
GTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCAACC
AACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTTCTGGGTGAGCAAAAACAGGAAGGCAAAAATGCCGAAA
AAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCTTTTTCAATATTATTGAAGCATTTATC
AGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACA
TTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTTATCATGACATTAACCTATAAAAAATAGGCGTAT
CACGAGGCCCTTTTCGTCTCGAATAAATACCTGTGACGGAAGATCACTTCGCAGAAATAAATAAATCCTGGTGTC
CCTGTTGATACCGGGAAGCCCTGGGCCAACTTTTGGCGAAAATGAGACGTTGATCGGCACGTAAGAGGTTCCAA
CTTTCACCATAATGAAATAAGATCACTACCGGGCGTATTTTTTGAGTTATCGAGATTTTCAGGAGCTAAGGAAG
CTAAAAATGGAGAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAG
GCATTTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTTCAGCTGGATATTACGGCCTTTTTAAAGACCGT
AAAGAAAAATAAGCACAAAGTTTATCCGGCCTTTTATTCACATCTTGGCCGCCTGATGAATGCTCATCCGGAAT
TCCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTACCCCTTGTACACCGTTTTTCATGAG
CAAACCTGAAACGTTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATATTCGCA
AGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTTATTGAGAATATGTTTTTCGTCTCAG
CCAATCCCTGGGTGAGTTTACCAGTTTTGATTTAAACGTGGCCAATATGGACAACCTCTTCGCCCCCCGTTTTTC
ACCATGGGCAAATATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTTCAGGTTTCATCATGCCGTTTTG
TGATGGCTTCCATGTTCGGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGGGCGTAAT
TTTTTTAAGGCAGTTATTGGTGCCCTTAAACGCTGGTGCTACGCCTGAATAAGTGATAATAAGCGGATGAATG
GCAGAAATTCGAAAGCAAATTCGACCCGGTCTGTCGGTTACGGGCAGGGTCGTTAAATAGCCGCTTATGTCTATT
GCTGGTTTTACCGGTTTTATTGACTACCGGAAGCAGTGTGACCGGTGTGCTTCTCAAATGCCTGAGGCGAGTTTGTCT
CAGGCTCTCCCCGTGGAGGTAATAATTGACGATATGATCGACCAATTGCGGGAAGAAATTACAGCTTTTGCCGC
TGGCCTACAGAGTTTAGGAGTTACCCCCATCAACACCTGGCCATTTTCGCCGACAACAGCCCCCGGTGGTTTTA
TCGCCGATCAAGGCAGTATGTTGGCTGGAGCCGTCAACGCCGTCCGTTCTGCCAAGCAGAGCGCCAGGAATTA
CTCTACATCCTAGAAGACAGCAACAGCCGTACTTTAATCGCAGAAAATCGGCAAACCCTAAGCAAATTTGGCCCT
AGATGGCGAAACCATTTGACCTGAAACTAATCATCCTCCTCACCGATGAAGAAGTGGCAGAGGACAGCGCCATTC
CCCAATATACTTTGCCCAGGTGATGGCCCTAGGGGCCGGCAAAATCCCCACTCCCCGTTCCCCGCCAGGAAGAA
GATTTAGCCACCCTGATCTACACCTCCGGCACCACAGGACAACCCAAAGGGGTGATGCTCAGCCACGGTAATTT
ATTGCACCAAGTACGGGAATTGGATTTCGGTGATTATTTCCCGCCCCGGCGATCAGGTGTTGAGCATTTTGGCCT
GTTGGCAACTCCCTAGAAAAGAGCGCCGAATTTTTCTTCTTTCCCGGGGCTGCACGATGAACCTACACAGCATT
CGCCATTTCAAGGGGGATGTGAAGGACATTAACCCCATCACATTGTGCGGTGTGCCCGGCTGTGGGAATCCCT
CTACGAAGGGGTACAAAAACGTTCCGGGCTAAGGGCGAATTTCTGCAGATATCCATCACACTGGCGGCCGCTCG
AGCATGCATCTAGAGGGCCCAATTCGCCCTATAGTGAGTGTGATTACAATTCAGTGGCCGTCTGTTTTACAACGT
CGTGACTGGGAAAACCCCTGGCGTTACCCAACCTAATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGCGTAA
TAGCGAAGAGGCCCCGACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAG
CGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCCG
CTCCTTTTCGCTTTCTTCCCTTCTTCTCGCCACGTTTCGCCGGCTTTCCCGTCAAGCTCTAAATCGGGGGCTC
CCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAACTTGATTAGGGTGATGGTTACAGTAG

TGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGT
TCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTGATTTATAAGGGATTTTGCCGATTTTCGGCC
TATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACCGGAATTTTAACAAAATTACAGGGCGCAAGGGCTGC
TAAAGGAAGCGGAACACGTAGAAAGCCAGTCCGCAGAAACGGTGCTGACCCCGGATGAATGTGAGCTACTGGGC
TATCTGGACAAGGGAAAAACGCAAGCGCAAAGAGAAAGCAGGTAGCTTGACGTGGGCTTACATGGCGATAGCTAG
ACTGGGCGGTTTTTATGGACAGCAAGCGAACCAGGAATTGCCAGCTGGGGCGCCCTCTGGTAAGGTTGGGAAGCCC
TGCAAAGTAAACTGGATGGCTTTCTTGCCGCCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGAC
AGGATGAGGATCGTTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGC
TATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTGAGCGCAGGGG
CGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATC
GTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGC
TATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCT
GATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGA
GCGAGCACGTACTCGGATGGAAGCCGGTCTTGTGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGC
CAGCCGAAGTGTTCGCCAGGCTCAAGGCGCGCATGCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCC
TGCTTGCCGAATATCATGGTGAAAAATGGCCGCTTTTCTGGATTATCGACTGTGGCCGGCTGGGTGTGGCGGA
CCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCC
TCGTGCTTTACGGTATCGCCGCTCCCGATTGCGAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA
ATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCCTT
CCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTA
CATCGAAGTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAAGCTTTTCCAATGATGAGCA
CTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATA
CACTATTCTCAGAATGACTTGGTTGAGTACTACCCAGTACAGAAAAGCATCTTACGGATGGCATGACAGTAAG
AGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGAC
CGAAGGAGCTAACCGCTTTTTTGACACAACATGGGGGATCATGTAATCGCCTTGATCGTTGGGAACCGGCTG
AATGAAGCCATACCAAAACGACGAGCGTGACACCAGTGCCTGTAGCAATGGCAACAGCTTGCGCAAACTATT
AACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAAGTTGCAGGAC
CACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGC
GGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGC
AACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTTGGTAAGTGTGAGACC
AAGTTTACTCATATATACTTTAGATTGATTTAAACTTCATTTTTTAATTTAAAGGATCTAGGTGAAGATCCTT
TTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGAT
CAAAGGATCTTCTTGAGATCCTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAACCACCGCTACCAG
CGGTGGTTTTGTTTGGCGGATCAAGAGCTACCAACTTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATA
CCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAAGTCTGTAGCACCAGCTACATACCT
CGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCTGTGCTTACCAGGTTGGACTCAAGAC
GATAGTTACCGGATAAGGCGCAGCGGTGCGGCTGAACGGGGGGTTCGTGCACACAGCCAGCTTGGAGCGAAGC
ACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGA
CAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATC
TTTATAGTCCTGTGCGGTTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGC
CTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTTCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTT
TCCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGACGCC
GAACGACCGAGCGCAGGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAAACCGCCTTCCCCGCG
CGTTGGCCGATTATTAATGCGAGCTGGCAGCAGAGTTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAAT
TAATGTGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAA
TTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGGTACCGAGCTCGGA
TCCACTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCCTTCTCGAG

SEQ ID NO:28

GATCCGCTGTTGACCCAACAGCATGAGTCGTTATCCAAGGGGAGCTTCGGCTCCCTTTTTTCATGCGCGGATGC
GGTGA

SEQ ID NO:29

GGATCCACTAGTCCTGAGGTGTTGACAATTAATCATCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACA
ATTTACACAGGAAACAGACCATGGCCGTGCGACTGCAACCAGCTCAAGAAGTCGCAACTAAGAAAAAGCCTGC
AATCAAACAGCGGCGCGTGGTGGTTACCGGCATGGGTGTGGTACTCCCCTCGGGCATGAACCGGATGTGTTTT
ACAACAATCTCCTGGATGGCGTGAGCGGCATTAGTGAGATCGAGAATTTTACTCGACGCAGTTTCCCACTCGC
ATTGCCGGCGAAATCAAGAGTTTACGACCCGACGGCTGGGTGCGGCCCAAATTGAGCAAACGGATGGATAAATT
GATGCTGTATCTGCTCACCGCAGGCAAGAAAGCGCTGGCCGATGCGGGCATCACGGATGATGTGATGAAAGAGC
TGGATAAACGCAAATGTGGAGTTCTGATTGGCAGTGGCATGGGCGGCATGAAGCTGTTCTACGATGCGCTCGAA
GCCCTGAAGATTTCTGATCGAAAGATGAACCCATTCTGTGTGCCTTTTGCACCACGAATATGGGTAGCGCCAT

GCTGGCTATGGATTTGGGGTGGATGGGGCCGAATTATAGTATTTCCACCGCGTGCGCAACCTCGAACTTCTGCA
TCTTGAACGCGGCTAACCACATTATCCGTGGTGAAGCAGACATGATGCTCTGCGGCGGCTCCGATGCGGTCAAT
ATCCCTATCGGTTTGGGCGGCTTTGTTGCTTGCCGCGCCTTGAGCCAACGCAATAACGACCCAACCAAGGCATC
GCGCCCGTGGGACAGCAATCGCGATGGCTTCGTTCATGGCGAGGGAGCCGGGGTGCTGCTGTTGGAGGAGCTGG
AACACGCGAAAAAGCGAGGCGCGACAATCTATGCTGAGTTCTTGGGAGGGTCCTTTACATGCGATGCCTACCAC
ATGACGGAGCCTCACCAGAGGGGCGCAGGCGTGATCTTGTGTATCGAGAAGGCAATGGCTCAGGCAGGAGTCTC
TCGCGAGGATGTTAACTACATTAATGCTCAGCAACGTCCACGCGGCTGGTGACATCAAGGAATACCAAGCTC
TCGCCCCATTGTTTCGGCCAGAACTCGGAGCTGCGGGTCAATAGTACAAAGTCCATGATCGGTTCATCTGCTGGGT
GCTGCCGGTGGCGTGAAGCTGTGACAGTCATTCAAGCCATCCGCACCGGCTGGATTACCCCTAATCTGAACCT
GGAAGACCCGGACAAGGCCGTTGACGCAAAATTCCTCGTCGGACCGGAGAAAGAAGCTCTCAACGTTAAAGTCG
GATTGAGCAATAGTTTCGGTTTGGTGGCCATAACTCTAGTATCCTGTTTGACCCCTATAATTGATAATAGATC
TGATCCGCTGTTGACCCAACAGCATGAGTCGTTATCCAAGGGGAGCTTCGGCTCCCTTTTTTCATGCGCGGATG
CGGTGAGAGCTCACGTGTCTAGA

SEQ ID NO:30

GGATCCACTAGTCTGAGGTGTTGACAATTAATCATCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACA
ATTTACACACAGGAAACAGACCATGGCAAGCCGTGTTGTTGGTAAAGGTTGTAAACTCGTTGGATGTGGTAGTGC
CGTCCCGAAGTTGGAGGTGAGTAACGACGACCTCAGTAAGATCGTGGATACTTCCGATGAATGGATTTCTGTTT
GGACGGGAATCCGCAACCGGCGGGTGATTACTGGTAAGGATAAGATGACGGGGCTGGCGGTGAGGCAGCCCAG
AAAGCCCTGGAAATGGCTGAAGTCGATGCTGACGATGTGGACTTGCTCCTGTTGTGCACCTCCACCCCAGATGA
TCTCTTTGGAAGTGCGCCGCAATCCAGGCGGCCTCGGCTGCAAAGGAAACCCTCTGGCATTGATATTACAG
CCGCTTGTAGCGGCTTCGTTCTGGGTCTGGTGAGTGCTTCCTGCTATATCCGCGGCGCGGGTTCAAGAACGTC
CTGGTTATCGGCGCGGACGCACTGAGCCGCTACGTCGATTGGACTGACCGCGGCACATGCATTCTCTTTGGTGA
CGCCGCTGGCGCTGTGTTGGTCCAGGCGTGTGAGAGCGAGGACGACGGCGTCTTCGGGTTTGATCTGCATAGCG
ATGGAGAGGGTTATGCCACCTGCATACTGGGCTCAAGGCGAACGAGGAGTTCGGGACGAACGGTTCCGTTTGTG
GATTTTCCGCCCAAGCGCAGCTACTCTTCCATCCAAATGAATGGGAAAGAAGTGTCCGTTTCGCCTGCCG
CGTCGTGCCCCAGTCTATTGAGATCGCACTCGAGAACCGGGCCTCACACGTTCTAGCATTGATTGGCTGCTGC
TCCACCAAGCAAACCAACGAATCTTGGATGCCGTCGCAACGCGTCTGGAAATTCCCGACAGACCGGTGATTAGT
AACTTGGCTAATTACGGCAATACTTCTGCCGCCAGCATTCCGTTGGCACTGGATGAAGCCGTGCGCAGCGGTAA
GGTCAAACCCGGTCAGACTATCGCAACTTCGGGGTTTGGAGCAGGCTTGACATGGGGCAGCGCATCATTGCT
GGAATTAATGATAGATCTGATCCGCTGTTGACCCAACAGCATGAGTCGTTATCCAAGGGGAGCTTCGGCTCCCT
TTTTTCATGCGCGGATGCGGTGAGAGCTCACGTGTCTAGA

SEQ ID NO:31

TGTTGACAATTAATCATCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGC
GCCGCTGAGAAAAAGCGAAGCGGCACTGCTCTTTAAACAATTTATCAGACAATCTGTGTGGGCACTCGACCGGAA
TTATCGATTAACTTTATTATTAATAAATAAGAGGTATATTAATGTATCGATTAAATAAGGAGGAATAAACC
ATGGCCGTGCACTGCAACAGCTCAAGAAGTCGCAACTAAGAAAAAGCCTGCAATCAAACAGCGGCGCGTGGT
GGTTACCGGCATGGGTGTGGTGACTCCCCTCGGGCATGAACCGGATGTGTTTTACAACAATCTCCTGGATGGCG
TGAGCGGCATTAGTGAGATCGAGAATTTTGAATCGACGCAGTTTCCCACTCGCATTGCCGGCGAAATCAAGAGT
TTCAGCACCGACGGCTGGGTGCGGCCCAAATTGAGCAAACGGATGGATAAATTGATGCTGTATCTGCTCACCGC
AGGCAAGAAAGCGCTGGCCGATGCGGGCATCACGGATGATGTGATGAAAGAGCTGGATAAACGCAAAATGTGGAG
TTCTGATTGGCAGTGGCATGGCGGCGATGAAGCTGTTCTACGATGCGCTCGAAGCCCTGAAGATTTTCGTATCGA
AAGATGAACCCATTCTGTGTCCTTTTGGCACCAGAAATATGGGTAGCGCCATGCTGGCTATGGATTGGGGTG
GATGGGGCCGAATTATAGTATTTCCACCGCGTGCGCAACCTCGAACTTCTGCATCTTGAACGCGGCTAACCACA
TTATCCGTGGTGAAGCAGACATGATGCTCTGCGGCGGCTCCGATGCGGTCAATTATCCCTATCGGTTTGGGCGGC
TTTGTGCTTGCCGCGCCTTGAGCCAACGCAATAACGACCCAACCAAGGCATCGCGCCGTGGGACAGCAATCG
CGATGGCTTCGTTCATGGGCGAGGGAGCCGGGGTGCTGCTGTTGGAGGAGCTGGAACACGCGAAAAAGCGAGGCG
CGACAATCTATGCTGAGTTCTTGGGAGGGTCCTTTACATGCGATGCCTACCACATGACGGAGCCTCACCAGAG
GGCGCAGGCGTGATCTTGTGTATCGAGAAGGCAATGGCTCAGGCAGGAGTCTCTCGCGAGGATGTTAACTACAT
TAATGCTCACGCAACGTCCACGCGGCTGGTGACATCAAGGAATACCAAGCTCTCGCCATTGTTTCGGCCAGA
ACTCGGAGCTGCGGGTCAATAGTACAAAGTCCATGATCGGTTCATCTGCTGGGTGCTGCGGTGGCGTCAAGCT
GTGACAGTCATTCAAGCCATCCGCACCGGCTGGATTACCCCTAATCTGAACCTGGAAGACCCGGACAAGGCCGT
TGACGCAAAATTCCTCGTCGGACCGGAGAAAGAAGCTCTCAACGTTAAAGTCGGATTGAGCAATAGTTTCGGTT
TTGGTGGCCATAACTCTAGTATCCTGTTTGACCCCTATAATTGATAATAGATCTGATCCGCTGTTGACCCAACA
GCATGAGTCGTTATCCAAGGGGAGCTTCGGCTCCCTTTTTTCATGCGCGGATGCGGTGA

SEQ ID NO:32

CCTGAGGTGTTGACAATTAATCATCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTACACAGG
AAACAGCGCCGCTGAGAAAAAGCGAAGCGGCACTGCTCTTTAAACAATTTATCAGACAATCTGTGTGGGCACTCG
ACCGGAATTATCGATTAACTTTATTATTAATAAATAAGAGGTATATTAATGTATCGATTAAATAAGGAGGA

ATAAACCATGGCCGTCGCACTGCAACCAGCTCAAGAACTCGCAACTAAGAAAAAGCCTGCAATCAAACAGCGGC
GCGTGGTGGTTACCGGCATGGGTGTGGTGAATCCCTCGGGCATGAACCGGATGTGTTTTACAACAATCTCCTG
GATGGCGTGAGCGGCATTAGTGAGATCGAGAATTTTACTCGACGCGATTTCCCACTCGCATTGCCGGCGAAAT
CAAGAGTTTTAGCACCAGCGCTGGGTGCGGCCCAAATTGAGCAAACGGATGGATAAATTGATGCTGTATCTGC
TCACCGCAGGCAAGAAAGCGCTGGCCGATGCGGGCATCACGGATGATGTGATGAAAGAGCTGGATAAACGCAAA
TGTGGAGTTCTGATTGGCAGTGGCATGGGCGGCATGAAGCTGTTCTACGATGCGCTCGAAGCCCTGAAGATTTT
GTATCGAAAGATGAACCCATTCTGTGTGCCTTTTGCACCACGAATATGGGTAGCGCCATGCTGGCTATGGATT
TGGGGTGGATGGGGCCGAATTATAGTATTTCCACCGCTGCGCAACCTCGAAGTTCTGCATCTTGAACGCGGCT
AACCACATTATCCGTGGTGAAGCAGACATGATGCTCTGCGGCGGCTCCGATGCGGTCAATTATCCCTATCGTTTT
GGGCGGCTTTGTTGCTTGCCGCGCCTTGAGCCAACGCAATAACGACCCAACCAAGGCATCGCGCCCGTGGGACA
GCAATCGCGATGGCTTCGTCATGGGCGAGGGAGCCGGGTGCTGCTGTTGGAGGAGCTGGAACACGCGAAAAAG
CGAGGCGCGACAATCTATGCTGAGTTCTTGGGAGGGTCCCTTACATGCGATGCCTACCACATGACGGAGCCTCA
CCCAGAGGGCGCAGGCGTGATCTTGTGTATCGAGAAGCAATGGCTCAGGCAGGAGTCTCTCGCGAGGATGTTA
ACTACATTAATGCTCACGCAACGTCCACGCCGGCTGGTGACATCAAGGAATACCAAGCTCTCGCCCATTGTTTC
GGCCAGAATCGGAGCTGCGGGTCAATAGTACAAAGTCCATGATCGGTATCTGCTGGGTGCTGCCGGTGGCGT
CGAAGCTGTGACAGTCAATCAAGCCATCCGACCGGCTGGATTACCCCTAATCTGAACCTGGAAGACCCGGACA
AGGCCGTTGACGCAAAAATTCTCGTCGGACCGGAGAAAGAACGTCTCAACGTTAAAGTCGGATTGAGCAATAGT
TTCGGTTTTGGTGGCCATAACTCTAGTATCCTGTTTGACCCCTATAATTGATAATAGATCCTGTCTGTTAACTGC
TTTGTGGTACTACCTGACTTCACCCCTCTTTTAAAGATGGCAAGCCGTGTTGTTGGTAAAGGTTGTAACTCGTT
GGATGTGGTAGTGCCGTCCCGAAGTTGGAGGTGAGTAACGACGACCTCAGTAAGATCGTGGATACTTCCGATGA
ATGGATTTCTGTTGCGACGGGAATCCGCAACCGGCGGGTGATTACTGGTAAGGATAAGATGACGGGGCTGGCGG
TCGAGGCAGCCAGAAAGCCCTGGAAATGGCTGAAGTCGATGCTGACGATGTGGACTTGCTCCTGTTGTGCACC
TCCACCCAGATGATCTCTTTGGAAGTGCGCCGCAATCCAGGCGGCACTCGGCTGCAAAAGGAAACCCCTCTGGC
ATTTGATATTACAGCCGCTTGTAGCGGCTTCGTTCTGGGTCTGGTGAGTGCTTCTGCTATATCCGCGGCGGCG
GGTTCAGAACGTCTCTGTTATCGGCGCGGACGCACTGAGCCGCTACGTCGATTGGACTGACCGCGGCACATGC
ATTCTCTTTGGTGACGCCGCTGGCGCTGTGTTGGTCCAGCGCTGTGAGAGCGAGGACGACGGCGTCTTCGGGTT
TGATCTGCATAGCGATGGAGAGGGTTATCGCCACCTGCATACTGGGATCAAGGCGAACGAGGAGTTCGGGACGA
ACGTTTCCGTTGTGGATTTTCCGCCCAAGCGCAGCAGCTACTCTTCCATCCAAATGAATGGGAAAGAAGTGTTT
CGTTTCGCTGCCGCTGCTGCCCCAGTCTATTGAGATCGCACTCGAGAACGCGGGCCTCACACGTTCTAGCAT
TGATTGGCTGCTGCTCCACCAAGCAAAACCAACGAATCTTGGATGCCGTGCAACGCGTCTGGAAATTCCCGCAG
ACCGCGTGATTAGTAACCTGGCTAATTACGGCAATACTTCTGCCGCCAGCATTCCGTTGGCACTGGATGAAGCC
GTGCGCAGCGGTAAGGTCAAACCCGGTCAGACTATCGCAACTTCGGGGTTTTGGAGCAGGCTTGACATGGGGCAG
CGCGATCATTGCTGGAATTAATGATAGATCTGATCCGCTGTTGACCCAACAGCATGAGTCGTTATCCAAGGGG
AGCTTCGGCTCCCTTTTTTCATGCGCGGATGCGG

SEQ ID NO:33

GTACGGGATCCCTGTCTGTTAACTGCTTTGTTGGTACTACCTGACTTCACCCCTCTTTTAAAGATGGCAAGCCGTGT
TGTTGGTAAAGGTTG

SEQ ID NO:34

CACGTGAGCTCTCACCGCATCCGCGCATG

SEQ ID NO:35

TCATGAAGTTCTTTGTCGTCGCCGTCTCAGCACTTGCAACTGCATCTGCTTTTACAACCAGTCTGCTCTTTT
ACCACTGTGACGAGTCTTTCGGTGAACAATGTGTTCCGACAGGAGGGAAATGCTCACAGGAACAGGAGAGCTAC
CATTGTGATGGATGGAGCTAACGGAAGTGCAGTCAGTTTGAAAAGTGGGTCAATTGAATACGACAGGAGGACACAA
GTTTCGTCGCCACCGCCCCGTACATTCTTACCAACTCCCTGATTGGAGCAGATTGCTCACTGCCATCACAAACC
GTTTTTGTAAAGTAAGCGTCCGGATATGCATGATCGTAAGTCGAAAAGGCCGGACATGCTCGTGGATAGTTT
CGGGTTGGAGAGTACCGTTACAGGATGGACTCGTGTTCCGTCAAAGCTTTTCGATCCGTTTCATATGAGATTGGAA
CTGATCGTACGGCTTCCATTGAGACTTTGATGAACCATCTTCAGGAGACTTCCCTCAACCATTTGTAAGAGTACA
GGAATTTTGTGGATGGATTTCGGACGCACACTCGAAATGTGTAAGCGCGATTGATTGTTGGGTGCTCATTAAAT
GCAGATCAAGGTTAATAGATACCCGGCTGGGGCGATACAGTAGAAATCAATACTAGGTTACGACAGACTTGGTA
AGATCGGCATGGGTGAGATTGGCTCATTAGCGACTGCAATACCGGTGAGATCCTCGTCAGGGCAACCGAGCGCC
TACGCCATGATGAATCAGAAAGACCCGAAGACTCTCGAAGCTTCCGTACGAGGTCCACCAAGAGATTGTCCCCCT
TTTTGTGCACTCCCCCGTAATTGAAGATTCCGATCTCAAGGTCCACAAATTCAAAGTTAAACGGGTGACAGCA
TCCAGAAGGGACTTACTCCTGGTTGGAACGACCTCGATGTGAACCAACATGTTTCGAACGTGAAATATATCGGC
TGGATTCTTGAGAGTATGCAACCGAGGTACTTGAGACCGAGGAATTGTGCTCGTTGGCATTGGAGTATCGTCG
TGAGTGTGGGCGAGACTCAGTCCTCGAAAGTGTAAACGCAATGGACCCAAGCAAAGTTGGTGTTCGTTACAGT
ATCAACACCTCCTCCGTCTCGAGGATGGAACAGCCATTGTGAACGGGGCCACAGAGTGGAGGCCAAAGAACGCT
GGCGCTAACGGAGCTATCTCCACAGGAAAGACCAGCAATGGTAACTCTGTGAGTTAATGATAGGATCC

SEQ ID NO:36

MKFLVVAVSALATASAFITTPSPASFTTVSSPSVNNVFGQEGNAHRNRRATIVMDGANGSAVSLKSGSLNTQEDTS
SSPPPRTFLHQLPDWSRLLTAITTVFVKSKRPDMHDRKSKRPDMLVDSFGLESTVQDGLVFRQSFIRSIEIGT
DRTASIEITLMNHLQETSLNHCKSTGILLDGFGRITLEMCKRDLIWVVIKMQIKVNRYPAWGDTVEINTRFSRLGK
IGMGRDWLISDCNTGEILVRATSAYAMMNQKTRRLSKLPYEVHQEIVPLFVDSPIEDSDLKVHKFKVKGTGDSI
QKGLTPGWNLDVNDQHVSNVKYIGWILESMPTVELETQELCSLALEYRRECGRDSVLESVTAMDPSKVGVR SQY
QHLLRLEDGTAIVNGATEWRPKNAGANGAISTGKTSNGNSVS

SEQ ID NO:37

CAGGATCCGGGGAGGTGTGGTGTAGT

SEQ ID NO:38

TAGGATCCAGTGGTGCCCATGGTACTTTGTTAGGGGAGGATAG

SEQ ID NO:39

CAGGATCCTCACTCTGTCGCGCTGTTG

SEQ ID NO:40

CATCTAGAGAGGATTGATTTCCGAGTC

SEQ ID NO:41

ATGGGCACTACTCTCGACGACACGGCTTACCGCTACCGCACCAAGTGTGCCGGGGGACGCCGAGGCCATCGAGGC
ACTGGATGGGTCTTACCACCGACACCGTCTTCCGCGTACCGGCCACCGGGGACGGCTTACCCTGCGGGAGG
TGCCGGTGGACCGCCCTGACCAAGGTGTTCCCGACGACGAGTCGGACGACGAGTCGGACGACGGGGAGGAC
GGCGACCGGACTCCCGACGTTCTGTCGCGTACGGGGACGACGGCGACCTGGCGGGCTTCGTGGTCTGTCGTA
CTCCGGCTGGAACCGCCGCTGACCGTCGAGGACATCGAGGTGCGCCCGGAGCACCGGGGGACGGGGTCGGGC
GCGCGCTGATGGGGCTCGCGACGGAGTTCGCGCGGAGCGGGGTGCCGGGCACCTCTGGCTGGAGGTACCAAC
GTCAACGACCGGCGATCCACGCGTACCGGCGGATGGGGTTACCCTCTGCGGCCTGGACACCGCCCTGTACGA
CGGCACCGCCTCGGACGGCGAGCAGGCGCTCTACATGTCCATGCCCTGCCCTAA

SEQ ID NO:42

CCATGGCCGCTATGCTCGCTCTAAGCAGGGCGCCTTCATGGGCGCAGCTCCTTTGCCCCGCCCCCAAGGGC
GTCGCCAGCCGCGGCTCCCTGCAGGTGGTGGCGGCCAACGGCAGCGCGGTGAGCCTGAAGTCGGGTTCCT
CAACACTCAGGAGGACACCTCGTCTCGCCCCGCGCACGTTCTGACACGCTGCCGACTGTTCCCGCC
TGCTGACGGCTATTACGACCGGTGTTCTGTAAGTCAAGCGCCCGACATGCACGACCGCAAGAGCAAGCGGCCT
GATATGCTGGTGGACAGCTTTGGCCTGGAGTCCACGGTGCAGGACGGCCTCGTGTTCGGGCAAGCTTCAGCAT
CCGCGACTACGAGATCGGCACGGACCGCACCGCGTTCGATCGAGACGCTCATGAACCACCTCCAGGAGACGTCGC
TCAACCCTGCAAGTCCACCGGTATCCTGCTGGACGGCTTTGGCCGACCCCTGGAGATGTGCAAGCGGGATCTG
ATCTGGGTGGTATCAAGATGCAGATCAAGGTGAACCGCTATCCCGCCTGGGGTGACACCGTCGAGATTAACAC
CCGCTTCTCGCGCCTGGGCAAGATCGGCATGGGGCGGACTGGCTGATCTCGGACTGCAACACTGGCGAGATCC
TGGTCCGGGCCACGTCGGCCTACGCCATGATGAACCAGAAGACTCGGCGGCTGAGCAAGCTGCCTTACGAGGTG
CATCAGGAGATCGTCCGCTCTTCTGTTGACAGCCCCGTGATCGAGGACAGCGATCTGAAGGTGCACAAGTTCAA
GGTCAAGACCGGCGACAGCATCCAGAAGGGCCTGACTCCCGGCTGGAACGACCTGGACGTGAACGACACGTCT
CGAACGTGAAGTACATCGGCTGGATTCTGGAGTCGATGCCCCACCGAGGTGCTGGAGACGAGGAGCTGTGCTCC
CTGGCGCTGGAGTATCGCCGCGAGTGGCGCCGCGACTCCGCTGCTGGAGTCCGTACCGCGATGGACCGCTCGAA
GGTGGGTGTCCGACCCAGTACCAACACCTGCTGCGCCTCGAGGACGGCACCGCCATTGTGAACGGCGCGACGG
AGTGGCGGCCGAAGAACGCGGGCGCTAACGGCGCCATCTCCACGGGCAAGACCTCCAACGGCAACTCGGTGAGC
TAATGATAGGATCC

SEQ ID NO:43

MAAMLASKQGAFFMGRSSFAPAPKGVASRGSQVAVANGSAVSLKSGSLNTQEDTSSPPPPRTFLHQLPDWSRL
LTAITTVFVKSKRPDMHDRKSKRPDMLVDSFGLESTVQDGLVFRQSFIRSIEIGTDRTASIEITLMNHLQETSL
NHCKSTGILLDGFGRITLEMCKRDLIWVVIKMQIKVNRYPAWGDTVEINTRFSRLGKIGMGRDWLISDCNTGEIL
VRATSAYAMMNQKTRRLSKLPYEVHQEIVPLFVDSPIEDSDLKVHKFKVKGTGDSIQKGLTPGWNLDVNDQHV
SNVKYIGWILESMPTVELETQELCSLALEYRRECGRDSVLESVTAMDPSKVGVR SQYQHLLRLEDGTAIVNGATE
WRPKNAGANGAISTGKTSNGNSVS

SEQ ID NO:44

GGTGGAAAATGCCTATGTGTTAACG

SEQ ID NO:45

CGTAGGCAGTGTGCAACCAGGAGCC

SEQ ID NO:46

GGTGGAAAATGCCTATGTGTTAACGGATCTACAAACCAGCACCAAACTCTATTACGAACCCACGGTTTCCACT
CTCCCCAACTGCAAGACTTGGGGCCCATTGATGTGGTTTTAACCCCGTCATTGGCATCAATATCCTCGGATTC
CTGCCGGTGCTCAATGGCCAGAAAACCACCTGGAGCTTTGTGCGACTGTCCATCCCCAGGCGATCGTCCCCAC
CTCTGGAGCCGCAGAAATTGAACATATAGCGGTTTACTAACTAAAGTTTTACGTTTAGACGGCGATCTCAGTCAAT
TTCGCCAGTCCCTAATTGACGAAGGGATACAAGCTTCCCTATGGGAACCCAGGTGGGAGTGCCCTCAATGTG
CCCCAATCCACCGTTGGCTAGGTTGGAATGTTCAAATCACTGTGCGGTGTGATGCTTGATAAAATACAGTGAGCC
AGGGAAAAGTGCAAAAAGTGTATAAAGTAGGTTTAACTTGAATCAAAATCCTTTCTCCGCAGTCATAGCCAGG
AGTAGGAAGATTACCAGCGAAGCAAGTTGTCTTCCCTAGCTTTGGGCGGGCAAACCCCTTGCAGTATTGCCAA
CGTCAAAAAATCACCATAGCCGAATGACCTACACCATCAACGCTGACCAAGTCCATCAGATTGTCCATAATCTT
CACCACGATCCCTTTGAAGTGTGGGCTGCCATCCCTCGGAGCTTTATGCTTGTAAACCGTTTTGTGAAAAAA
TTTTTAAATAAAAAAGGGGACCTCTAGGGTCCCCAATTAATTAGTAATATAATCTATTAAAGGTCATTCAAAA
GGTCATCCACCGGATCAGCTTAGTAAAGCCCTCGCTAGATTTTAAATGCGGATGTTGCGATTACTTCGCCAACTA
TTGCGATAACAAGAAAAAGCCAGCCTTTTCATGATATATCTCCCAATTTGTGTAGGGCTTATTATGCACGCTTAA
AAATAATAAAAGCAGACTTGACCTGATAGTTTGGCTGTGAGCAATTATGTGCTTAGTGCATCTAACGCTTGAGT
TAAGCCGCGCCGCGAAGCGCGCTCGGCTTGAACGAATTGTTAGACATTATTTGCCGACTACCTTGGTGATCTCG
CCTTTACGTAGTGGACAAATCTTCCAAGTATCTGCGCGGAGGCCAAGCGATCTTCTTCTTGTCCAAGATA
AGCCTGTCTAGCTTCAAGTATGACGGGCTGATACTGGGCCGGCAGGCGCTCCATTGCCAGTCGGCAGCGACAT
CCTTCGGCGGATTTTGGCGTTACTGCGCTGTACCAAAATGCGGGACAACGTAAGCACTACATTTGCTCATCG
CCAGCCAGTCGGGCGGCGAGTTCCATAGCGTTAAGGTTTCATTTAGCGCCTCAAATAGATCCTGTTTCAGGAAC
CGGATCAAAGAGTTCTCCGCGCTGGACCTACCAAGGCAACGCTATGTTCTCTTGTCTTTGTGAGCAAGATAG
CCAGATCAATGTCGATCGTGGCTGGCTCGAAGATACCTGCAAGAATGTCATTGCGCTGCCATTCTCCAAATTGC
AGTTTCGCGCTTAGCTGGATAACGCCACGGAATGATGTCGTGTCGACACAACATGGTGACTTCTACAGCGCGGAG
AATCTCGCTCTCTCCAGGGGAAGCCGAAGTTTCCAAAAGGTCGTTGATCAAAGCTCGCCGCGTTGTTTCATCAA
GCCTTACGGTCACCGTAACCAGCAAATCAATATCACTGTGTGGCTTCAGGCCGCCATCCACTGCGGAGCCGTAC
AAATGTACGGCCAGCAACGTCGGTTTCGAGATGGCGCTCGATGACGCCAACTACCTCTGATAGTTGAGTCGATAC
TTCGGCGATCACCGCTTCCCTCATGATGTTTAACTTTGTTTTAGGGCGACTGCCCTGCTGCGTAACATCGTTGC
TGCTCCATAACATCAAACATCGACCCACGGCGTAACGCGCTTGTCTGGATGCCCCAGCCAGGCGATAGACTG
TACCCCAAAAAACAGTCATAACAAGCCATGAAAACCGCCACTGCGCCGTTACCACCGCTGCGTTTCGGTCAAGG
TTCTGGACCAGTTGCGTGAGCGCATACGCTACTTGCATTACAGCTTACGAACCGAACAGGCTTATGTCCACTGG
GTTCTGTCCTTCATCCGTTTCCACGGTGTGCGTCACCCGGCAACCTTGGGCAGCAGCGAAGTCGAGGCATTTCT
GTCTTGGCTGGCGAACGAGCGCAAGGTTTTCGGTCTCCACGCATCGTCAGGCATTGGCGGCCTTGTCTTCTTCT
ACGGCAAGGTGCTGTGACGGATCTGCCCTGGCTTCAGGAGATCGGAAGACCTCGGCCGTCGCGGCGCTTGCCG
GTGGTGCTGACCCCGGATGAAGTGGTTTCGCATCCTCGGTTTTCTGGAAGGCGAGCATCGTTTGTTCGCCAGCT
TCTGTATGGAACGGGCATGCGGATCAGTGAGGGTTTGCAACTGCGGGTCAAGGATCTGGATTTTCGATCACGGCA
CGATCATCGTGCGGGAGGGCAAGGGCTCCAAGGATCGGGCTGGCACCCAGCCTGCGCGAGCAGGGGAATTGAT
CCGGTGGATGACCTTTTGAATGACCTTTAATAGATTATATTACTAATTAATTGGGGACCCTAGAGGTCCCCTTT
TTATTTTTAAAAATTTTTTACAAAAACGGTTTACAAGCATAAAGCTTCGGGGACCACGGCAAGGTCAATCAATG
GGTCATTCTGTGCTATTTTACCCACGGCTGAAGCGGTAACGGTGTGCTTCCCACCGATCGCCGGGAAGTGATTA
TGACCACGGTCCACCATCCCACTTTTTTGAATGCGTGTGGAGTTGGAAGAACCAGGAAGATTATCAATTAAGA
ATTACCGAAAATGGCCACGAAAGGGTAATTTATGACCCCTATGGTTTTTAAACTCCCAAACTGACGGATTTTGA
CCTCCATGTGTTTGGGGAAGGCAACCACCACCGTATTACGAAAAACTCGGTGCTACCTGATGACGGTGGATG
GAGTTAAAGGGGTTTTTTTTGTGTGTGGGCCCCCAATGCCCGCAACGTTTTCCATTTTGGGGGATTTCAACAAC
TGGGACGGCAGATTGCACCAATGCGGAAACGCAACAACATGGTGTGGGAATTATTTATCCCTGAGTTGGGGGT
GGGCACCTTCTTATAAGTATGAGATTAAAAACTGGGAAGGGCACATCTACGAAAAGACTGACCCCTACGGTTTTT
ACCAAGAAGTACGCCCAAAACCGCTTCCATTGTGGCAGACTTGGACGGTTACCAATGGCACGACGAAGATTGG
TTGGAAGCTAGGCGCACCGATCCCCTGAGCAAACCGTTTTCCGTTTACGAACCTCATTTAGGCTCCTGGTT
GCACACTGCCTACG

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. A cell culture of a recombinant cyanobacterium modified to contain a nucleic acid molecule comprising at least one recombinant expression system that produces at
5 least one exogenous acyl-ACP thioesterase,
wherein said acyl-ACP thioesterase preferentially liberates a fatty acid chain that contains 6-20 carbons, and
wherein the culture medium provides inorganic carbon as substantially the sole carbon source and
10 wherein said cyanobacterium secretes the fatty acid liberated by the acyl-ACP thioesterase into the culture medium.
2. The culture of claim 1, wherein the at least one exogenous acyl-ACP thioesterase is a Fat B thioesterase.
15
3. The culture of claim 1, wherein the at least one exogenous acyl-ACP thioesterase is a Fat B thioesterase derived from the genus *Cuphea*.
4. The culture of claim 1, wherein the at least one exogenous acyl-ACP
20 thioesterase is ChFatB2.
5. The culture of any one of claims 1 to 4, wherein the recombinant cyanobacterium has further been modified to produce an exogenous β -ketoacyl synthase (KAS).
25
6. The culture of claim 5, wherein the exogenous KAS preferentially produces acyl-ACPs having the chain length for which the thioesterase has preferred activity.
7. The culture of any one of claims 1 to 6, wherein the recombinant
30 cyanobacterium is further modified so that one or more genes encoding beta-oxidation pathway enzymes are inactivated or downregulated, or said enzymes are inhibited.
8. The culture of any one of claims 1 to 7, wherein the recombinant cyanobacterium is further modified so that one or more genes encoding acyl-ACP
35 synthetases are inactivated or downregulated, or said synthetases are inhibited.

9. The culture of any one of claims 1 to 7, wherein the recombinant cyanobacterium is further modified so that one or more genes encoding an enzyme involved in carbohydrate biosynthesis are inactivated or downregulated, or said enzymes are inhibited.
- 5
10. The culture of claim 9, wherein the enzyme involved in carbohydrate biosynthesis is a branching enzyme.
11. A method to convert inorganic carbon to fatty acids, said method comprising:
- 10 incubating the culture of any one of claims 1 to 10 such that the recombinant cyanobacterium therein secretes the fatty acid into the culture medium; and recovering the secreted fatty acids from the culture medium.
12. The method of claim 11, wherein the fatty acids are recovered from the culture
- 15 by contacting the medium with particulate adsorbents.
13. The method of claim 12, wherein the particulate adsorbents circulate in the medium.
14. The method of claim 12, wherein the particulate adsorbents are contained in a fixed bed column.
15. The method of claim 14, wherein the pH of the medium is lowered during said contacting.
- 25
16. The method of claim 15, wherein said pH lowering process comprises adding CO₂.
17. The method of any one of claims 11 to 16, wherein the medium is recirculated
- 30 to the culture.
18. The method of claim 12, wherein the particulate adsorbents are lipophilic.
19. The method of claim 12, wherein the particulate adsorbents are ion exchange
- 35 resins.

20. A cell culture of a recombinant cyanobacterium, said cyanobacterium modified to contain a nucleic acid molecule comprising at least one recombinant expression system that produces at least one exogenous acyl-ACP thioesterase substantially as hereinbefore described with reference to any one of the Examples, excluding
- 5 comparative examples, and/or figures 2-6.

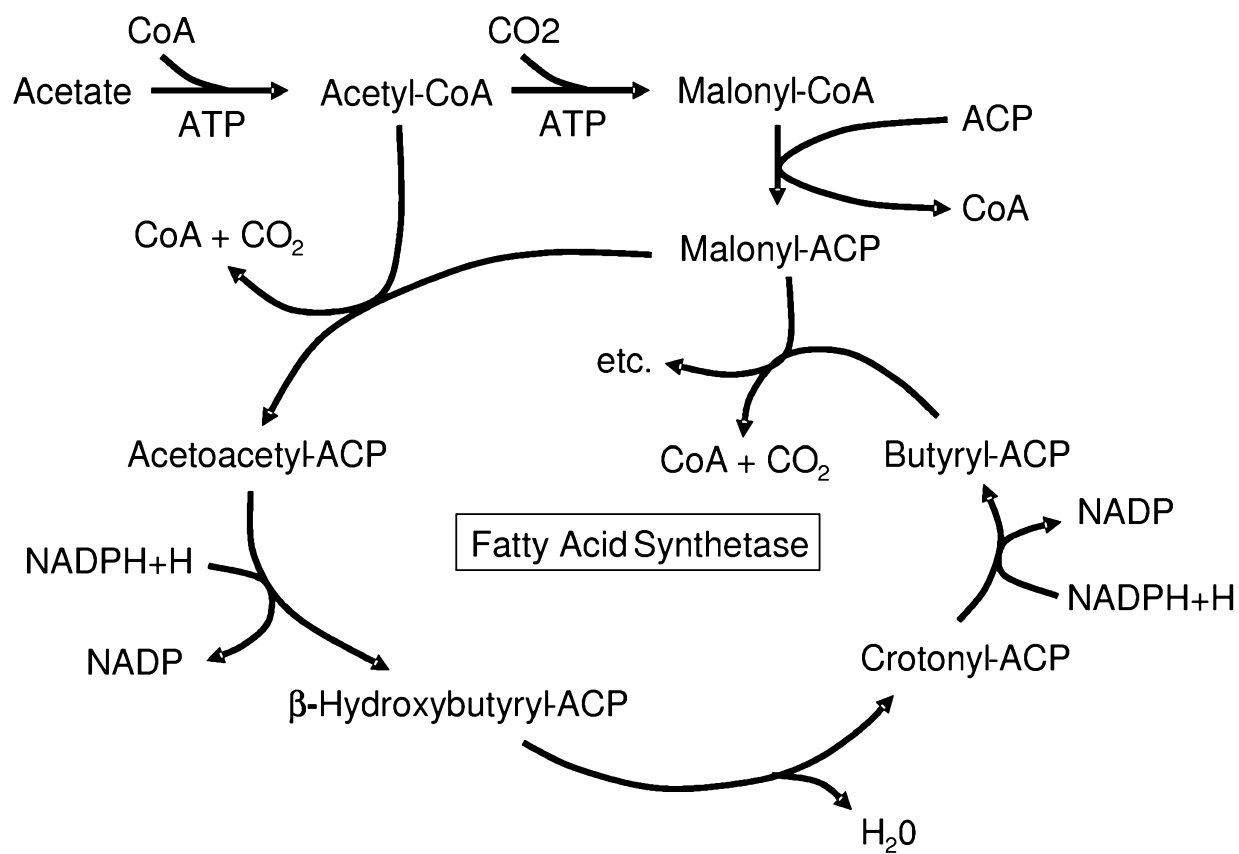


FIG. 1

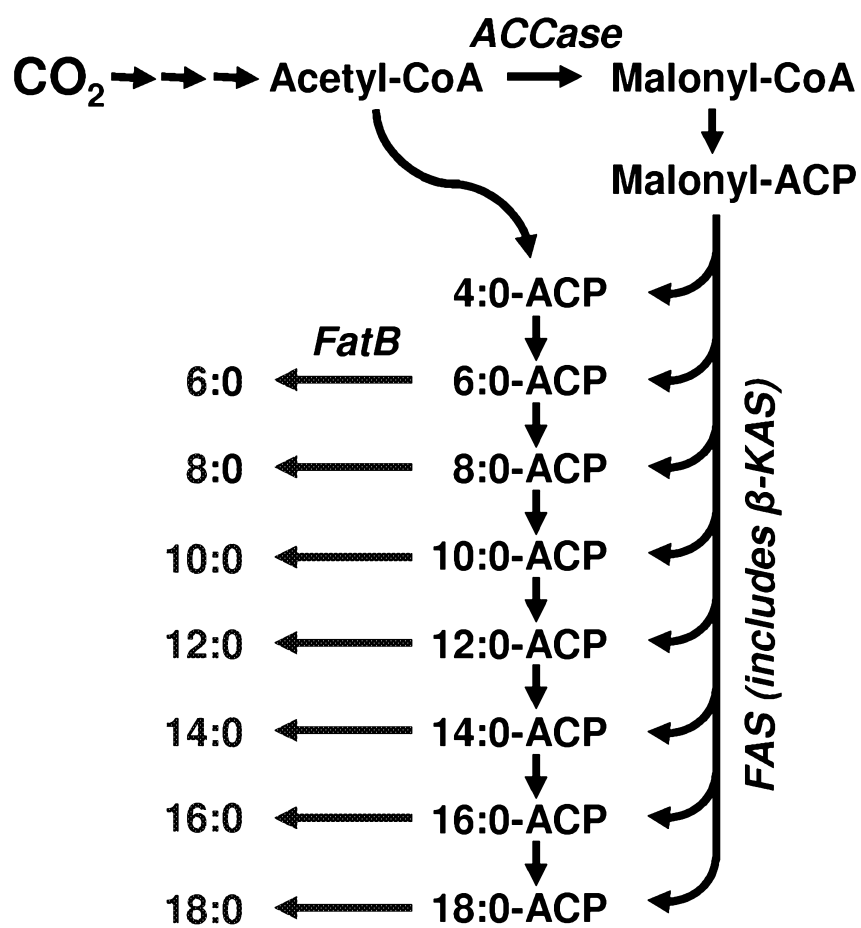


FIG. 2

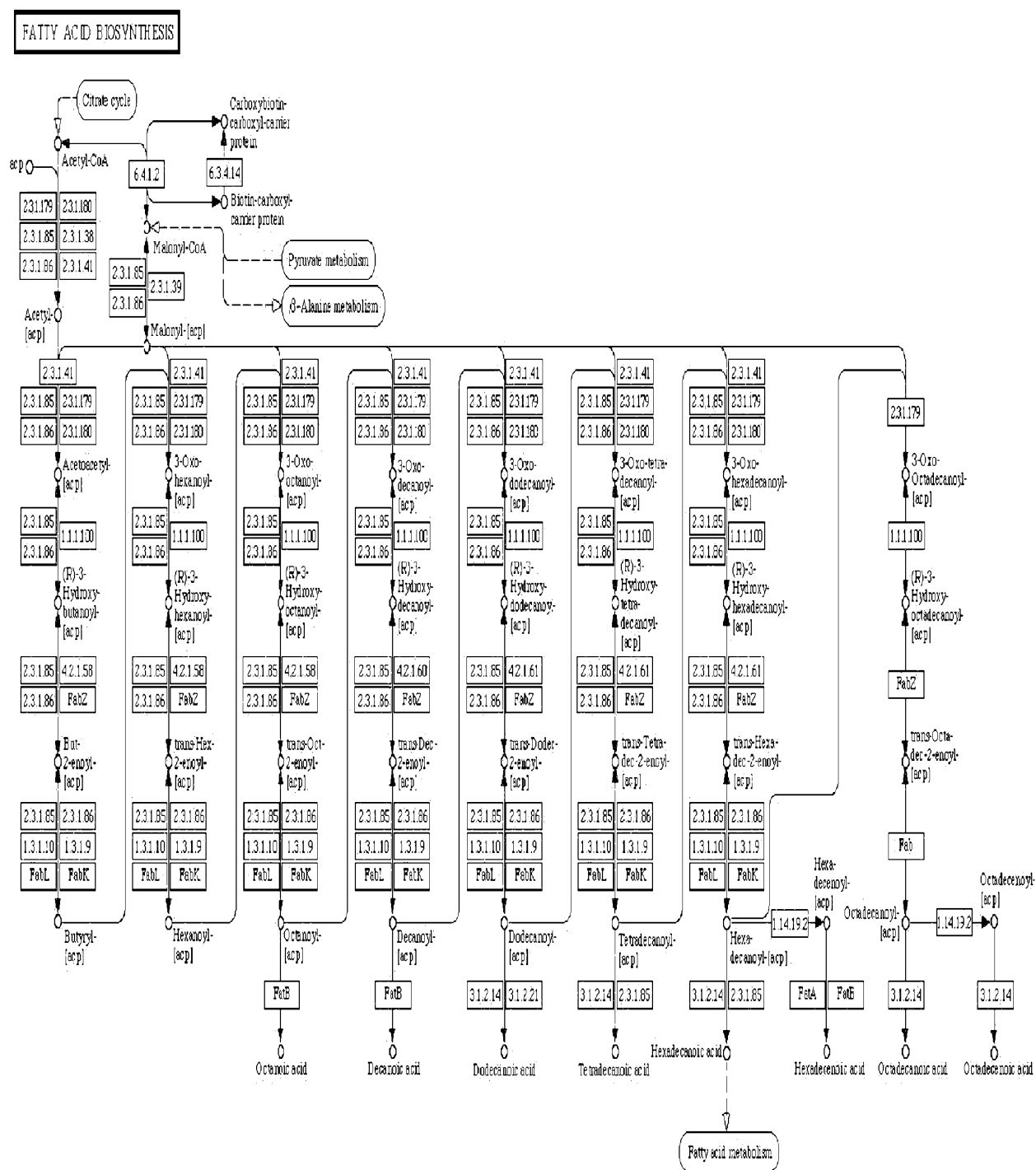


FIG. 3

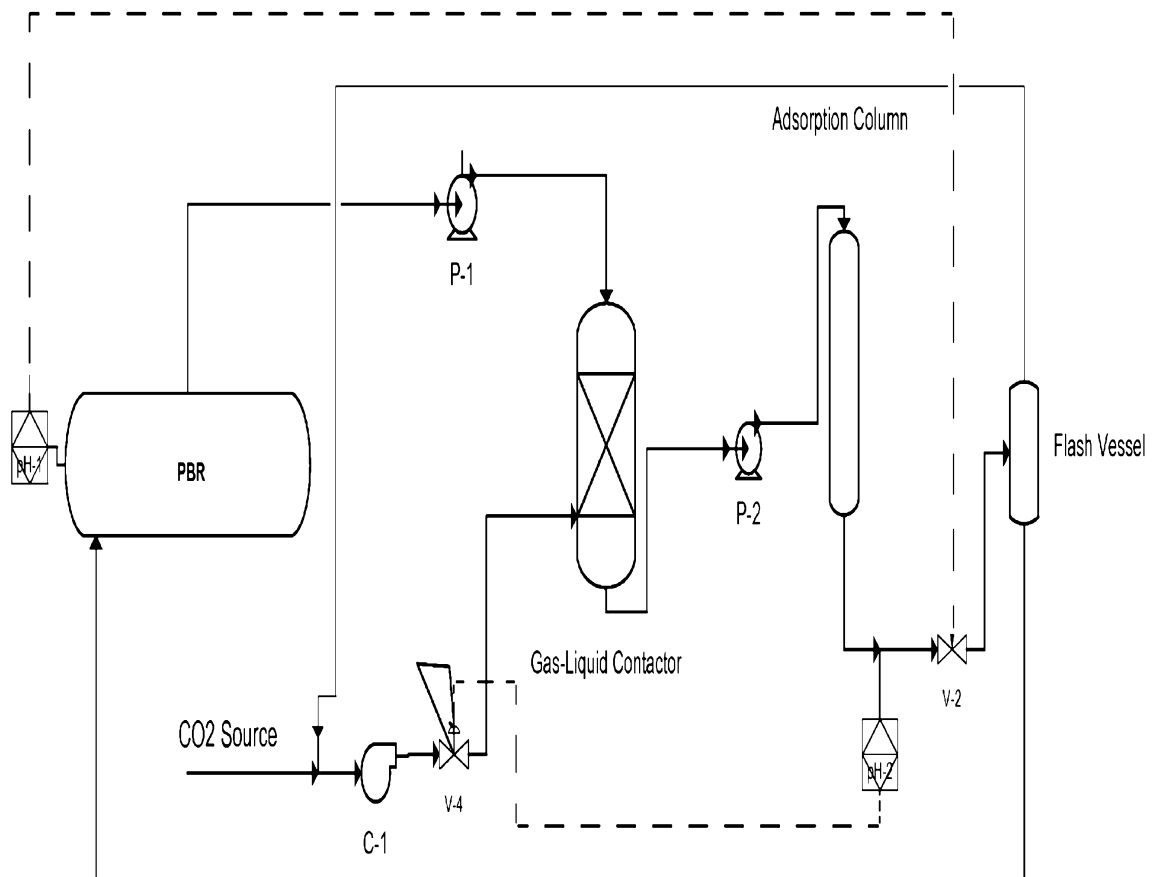
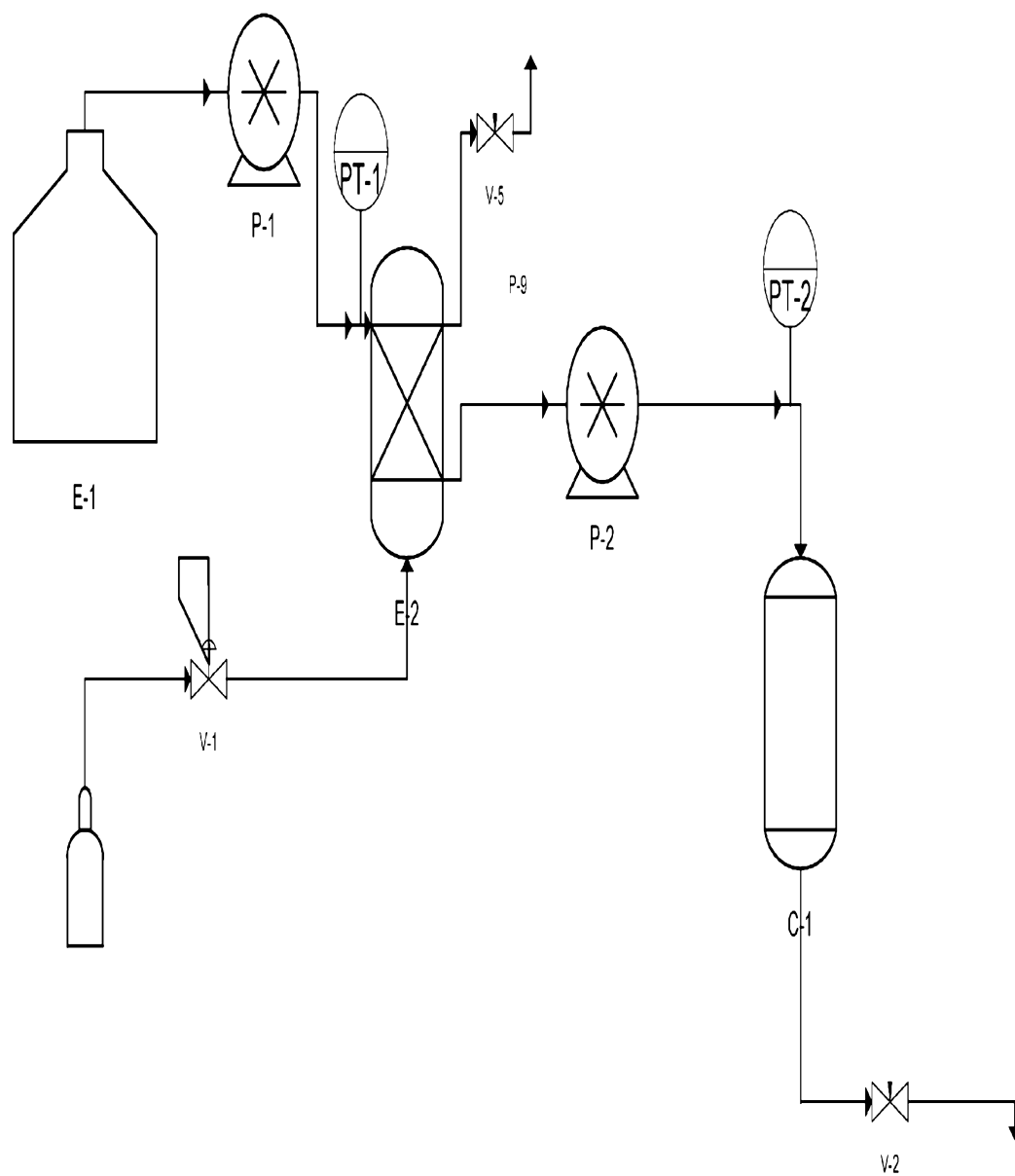
Integrated CO₂ delivery and product recovery schematic

FIG. 4



CO₂-mediated acidification of hydrophobic adsorption column load

FIG. 5

5/6

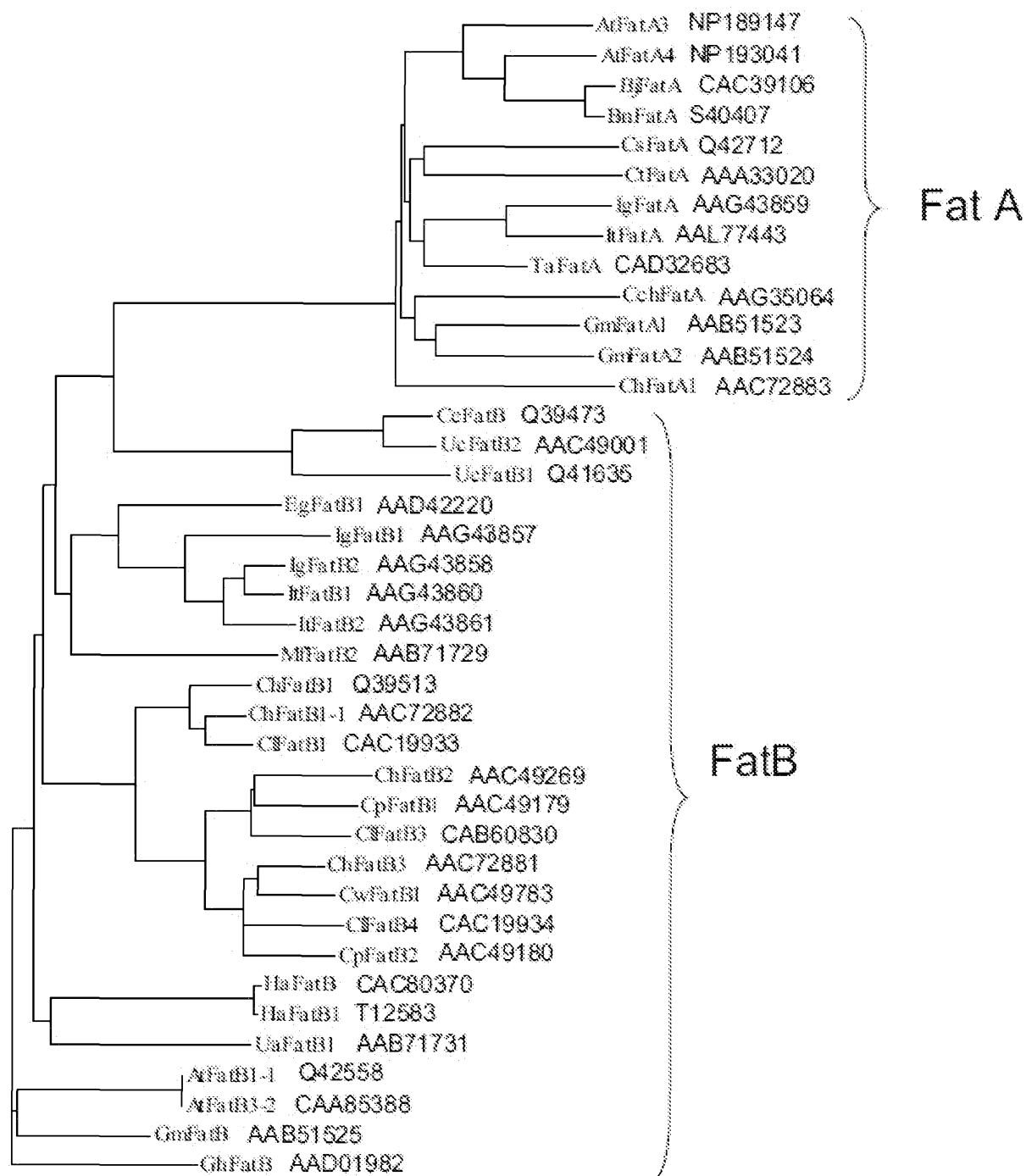


FIG. 6