Innovation, Sciences et Développement économique Canada
Office de la Propriété Intellectuelle du Canada

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 3181733 A1 2021/12/09

(21) 3 181 733

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2021/06/04
(87) Date publication PCT/PCT Publication Date: 2021/12/09
(85) Entrée phase nationale/National Entry: 2022/12/06
(86) N° demande PCT/PCT Application No.: US 2021/036031
(87) N° publication PCT/PCT Publication No.: 2021/248087
(30) Priorité/Priority: 2020/06/06 (US63/035,692)

(51) Cl.Int./Int.Cl. *C12N 9/10* (2006.01), *C12N 15/09* (2006.01), *C12N 15/52* (2006.01)

(71) Demandeur/Applicant:
CB THERAPEUTICS, INC., US

(72) Inventeurs/Inventors:
PEIFFER, LAURA FLATAUER, US;
VOGAN, JACOB MICHAEL, US;
WADE, JAMES LEE, US;
YACOUB, TYRONE JACOB, US;
TANG, KIRSTEN, US;
BURNETT, RACHEL NADINE, US

(74) Agent: GOWLING WLG (CANADA) LLP

(54) Titre : ENZYMES ET PROTEINES REGULATRICES DANS LE METABOLISME DE LA TRYPTAMINE
(54) Title: ENZYMES AND REGULATORY PROTEINS IN TRYPTAMINE METABOLISM

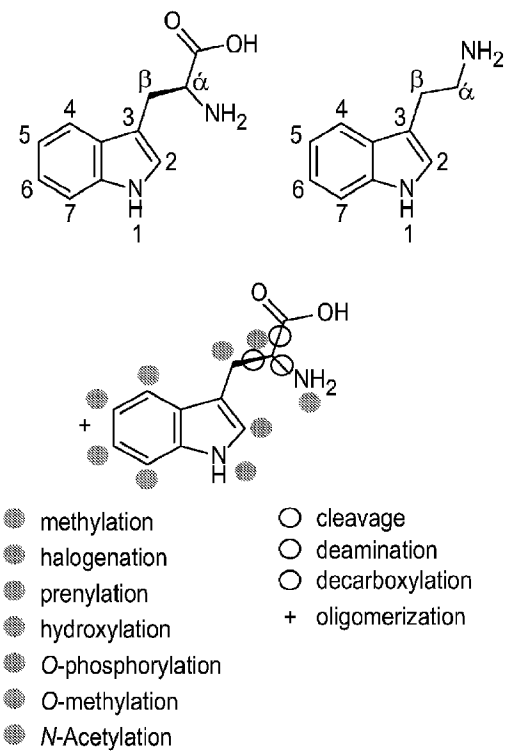


FIG. 1

(57) Abrégé/Abstract:
Provided are non-naturally occurring nucleic acids comprising a sequence encoding an enzyme or regulatory protein in tryptamine metabolism. Also provided are a recombinant microorganisms expressing the enzyme or regulatory protein. Methods of expressing the enzyme or regulatory protein are additionally provided.

Date Submitted: 2022/12/06

CA App. No.: 3181733

Abstract:

Provided are non-naturally occurring nucleic acids comprising a sequence encoding an enzyme or regulatory protein in tryptamine metabolism. Also provided are a recombinant microorganisms expressing the enzyme or regulatory protein. Methods of expressing the enzyme or regulatory protein are additionally provided.

ENZYMES AND REGULATORY PROTEINS IN TRYPTAMINE METABOLISM

CROSS-REFERENCE TO RELATED APPLICATION

This application claims the benefit of U.S. Provisional Application No. 63/035,692, filed June 6, 2020, and incorporated by reference herein in its entirety.

INCORPORATION-BY-REFERENCE OF MATERIAL SUBMITTED ON A COMPACT DISC

The Sequence Listing, which is a part of the present disclosure, includes a computer readable form and a written sequence listing comprising nucleotide and/or amino acid sequences of the present invention. The sequence listing information recorded in computer readable form is identical to the written sequence listing. The subject matter of the Sequence Listing is incorporated herein by reference in its entirety.

BACKGROUND OF THE INVENTION

(1) Field of the Invention

The present invention generally relates to the production of substituted indoles, e.g. N-methyl-L-tryptophan (NMTP), N,N-dimethyl-L-tryptophan (DMTP), and N,N,N-trimethyl-L-tryptophan (TMTP), and related tryptamines, e.g. N-methyltryptamine (NMT), N,N-dimethyltryptamine (DMT), and N,N,N-trimethyltryptamine (TMT), in a modified heterologous microorganism.

(2) Description of the related art

Mental health problems, which may also be referred to as mental illness or psychiatric disorder, are behavioral or mental patterns which impair the functioning of individuals across the world. Such mental health disorders include: personality disorders, anxiety disorders, major depressions, and various addictions. Indolic and tryptamine-based compounds similar in structure to the endogenous neurotransmitter serotonin have been increasingly evaluated for treating mental health problems. In contrast to anxiolytic medicines, usage of substituted indoles and methylated tryptamines, such as N,N-dimethyltryptamine does not lead to physical dependence.

The chemical synthesis of hydroxy, methoxy, phosphorylated, prenylated, and halogenated substituted tryptamines and indoles typically involve tedious techniques of organic chemistry.

Often, reproducibility is elusive and the solvents used during the syntheses of substituted tryptamines are environmentally toxic. Decarboxylations and selective methylations can be difficult to obtain via the techniques of organic chemistry. Further, the yields and purity of the intermediates for obtaining the target molecules can be low, where, for example, the starting molecule is L-tryptophan and the target molecule is N,N-dimethyltryptophan (DMTP), bufotenine, 5-MeO-dimethyltryptamine (5-MeO-DMT), 7-dimethylallyltryptophan, psilocybin, aeruginascin, among others.

The present invention provides for producing substituted tryptamines and indoles in recombinant microorganisms, providing for a more environmentally benign and higher yielding processes for production of those compounds.

BRIEF SUMMARY OF THE INVENTION

In some embodiments, provided is a non-naturally occurring nucleic acid comprising a sequence encoding an enzyme or regulatory protein in tryptamine metabolism, where the enzyme or regulatory protein is an N-methyltransferase (INMT, PsiM, TrpM), a tryptophan decarboxylase (AADC), a tryptophan hydroxylase (TPH), a tryptamine 4' hydroxylase (T4H), a tryptamine 5' hydroxylase (T5H), a truncated cytochrome p450 reductase (T4H-CPR, T5H-CPR), an hydroxytryptamine O-methyltransferase (IOMT or CaffOMT), an N-acetyltransferase (NAT), a deacetylase (DAC), a hydroxyl tryptamine kinase (PsiK), a tryptophan synthase (TrpS), a toluene monooxygenase (TMO), an aminotransferase/methyltransferase fusion (ATMT), a phosphatase, an oxidase, a dimethylallyltryptophan synthase (DMAT), an isopentenyl-diphosphate isomerase (IDI1), a tryptophan halogenase (TrpHalo), an aspartate oxidase/quinolinic acid synthase fusion (AOQS), a tryptophan importer (TAT2), a methionine importer (MUP1), or a SAMe importer (SAM3).

Also provided is an expression cassette comprising any of the above nucleic acids with a promoter functional in a recombinant microorganism.

Additionally provided is a recombinant microorganism comprising the above expression cassette, that expresses the enzyme or regulatory protein encoded therein.

Further provided is a non-naturally occurring enzyme or regulatory protein comprising an amino acid sequence encoded by any of the above-identified nucleic acids.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

FIG. 1 depicts the chemical structures of tryptophan and tryptamine, including various modifications which are performed by the enzymes disclosed within.

FIG. 2 depicts various substituted indole compounds in the tryptamine and tryptophan pathways utilized in the present invention. Panel A depicts the indole ring structure with positional numbering, and tryptophan and tryptamine. Panel B depicts examples of hydroxy modified tryptophan and tryptamine. Panel C depicts the 5-hydroxy indole ring structure with positional numbering, and examples of modified 5-hydroxy tryptamines. Panel D depicts the 4-hydroxy indole ring structure with positional numbering, and examples of modified 4-hydroxy tryptamines.

FIG. 3 depicts biosynthetic pathways utilized herein. Panel A depicts the biosynthetic pathways to tryptophan and genetic manipulations to increase tryptophan flux toward modified indoles and tryptamines. Panel B depicts the biosynthetic pathways to the methyl donor, SAMe and genetic manipulations to increase SAMe flux toward modified indoles and tryptamines.

FIG. 4 depicts enzymatic reactions utilized herein. Panel A depicts SAMe usage by INMT for methyltransferase activity. Panel B depicts BH4 usage by TPH for hydroxylase activity. Panel C depicts SAMe usage by INMT for methyltransferase activity on hydroxy tryptamine. Panel D depicts SAMe usage by IOMT (or CaffOMT) for methyltransferase activity. Panel E depicts NAD(P)H usage by T5H for hydroxylase activity. Panel F depicts acetyl-CoA usage by NAT for acetylation activity.

FIG. 5 depicts routes of modification of tryptamine by combinatorial usage of INMT, T5H, and IOMT enzymes.

FIG. 6 depicts routes of modification of tryptophan by combinatorial usage of TrpM, TPH, and IOMT enzymes (Panel A) and example branch points where modified tryptophan becomes modified tryptamine via use of the AADC enzyme (Panel B).

FIG. 7 depicts (A) routes of modification of serotonin by combinatorial usage of INMT and IOMT enzymes; (B) conversion of 5-HTP to serotonin by the AADC enzyme; (C) conversion of serotonin to N-acetylserotonin by the NAT enzyme, and N-acetylserotonin conversion to melatonin via the IOMT enzyme; (D) conversion of serotonin .to 5-MT by the IOMT enzyme, and 5-MT conversion to melatonin via the NAT enzyme; and (E) conversion of melatonin to 5-MeO-tryptamine by the DAC enzyme, and subsequent N-methylation by INMT to generate compounds such as 5-MeO-DMT.

FIG. 8 depicts (A) halogenation of tryptophan and tryptamine on the indole ring by the TrpHalo enzyme; (B) example route to halogenated DMT via combinatorial use of TrpHalo,

AADC, and INMT enzymes; (C) prenylation of tryptophan and tryptamine on the indole ring by the DMAT-IDI1 fusion enzyme; and (D) example route to prenylated DMT via combinatorial use of DMAT-IDI1, AADC, and INMT enzymes.

FIG. 9 depicts (A) a modified host organism expressing gene combinations with TPH, AADC, and TrpM enzymes to convert tryptophan into various hydroxy tryptamines; (B) a modified host organism expressing gene combinations with TPH, AADC, TrpM, and IOMT enzymes to convert tryptophan into various methoxy tryptamines; (C) a modified host organism expressing gene combinations with AADC, T5H, and INMT enzymes to convert tryptophan into various hydroxy tryptamines; and (D) a modified host organism expressing gene combinations with AADC, T5H, INMT, and IOMT enzymes to convert tryptophan into various methoxy tryptamines.

FIG. 10 depicts (A) a modified host organism which can generate various hydroxy tryptamines through bioconversion of serotonin provided exogenously or generated within the host organism; and (B) a modified host organism which can generate various methoxy tryptamines through bioconversion of melatonin provided exogenously or generated within the host organism.

FIG. 11 depicts (A) a scaffolded biosynthesis pathway of colocalized AADC, T5H-CPR fusion, IOMT, and NAT enzymes for conversion of tryptophan to melatonin; and (B) a modified host organism expressing the biosynthesis pathway from FIG. 11A to convert tryptophan to melatonin and related products.

FIG. 12 depicts (A) a scaffolded biosynthesis pathway of colocalized AADC, T4H-CPR fusion, PsiK, and PsiM enzymes for conversion of tryptophan to psilocybin related products; and (B) a modified host organism expressing the biosynthesis pathway from FIG. 12A to convert tryptophan to psilocybin and related products.

FIG. 13 depicts (A) example routes to halogenated, prenylated, and N-methylated alpha-methyl-tryptamine (AMT); and (B) a modified host organism expressing gene combinations to modify exogenously provided AMT to generate alpha-methylated-tryptamine variants.

FIG. 14 depicts (A) a heterologous tryptophan synthase (TrpS) route to combine synthetically modified indole with serine or threonine to generate indole modified tryptophan or indole modified beta-methyl tryptophan; and (B) a host organism expressing gene combinations to generate variants of indole modified tryptophan or indole modified beta-methyl tryptophan.

FIG. 15 depicts (A) the ATMT fusion enzyme converted tryptophan to beta-methyl tryptophan; and (B) a host organism expressing the ATMT fusion enzyme with gene combinations to generate beta-methyl tryptophan variants.

FIG. 16 depicts (A) the conversion of phosphorylated tryptamines to the corresponding hydroxy tryptamines by dephosphorylation; and (B) the oxidation of example hydroxy tryptamines which can catalyze polymerization.

FIG. 17 depicts HPLC chromatograms and UV-vis spectral matching of fermentation derived tryptamine via expression of the AADC enzyme.

FIG. 18 depicts HPLC chromatograms of fermentation derived methylated tryptamine via expression of the TrpM enzyme.

FIG. 19 depicts HPLC chromatograms of fermentation derived 4-OH tryptamine with improvements in yield via an optimal T4H-CPR fusion.

FIG. 20 depicts HPLC chromatograms of fermentation derived 5-OH-NMT via bioconversion of exogenous serotonin.

FIG. 21 depicts (A) a biosynthetic route to serotonin and 5-OH-NMT with a T5H enzyme or with a T5H-CPR fusion enzyme; and HPLC chromatograms of fermentation derived serotonin and 5-OH-NMT with improvements in yield via an optimal T5H-CPR fusion.

FIG. 22 depicts HPLC chromatograms of fermentation derived serotonin and melatonin.

FIG. 23 depicts HPLC chromatograms of fermentation derived 5-OH NMT and bufotenine.

FIG. 24 depicts HPLC chromatograms of fermentation derived psilocybin.

FIG. 25 depicts a synthetic route to methylate various tryptamines.

FIG. 26 depicts HPLC chromatograms and UV-vis spectral matching of fermentation derived DMT.

DETAILED DESCRIPTION OF THE INVENTION

Abbreviations and Definitions

To facilitate understanding of the invention, a number of terms and abbreviations as used herein are defined below as follows:

Conservative amino acid substitutions: As used herein, when referring to mutations in a protein, "conservative amino acid substitutions" are those in which at least one amino acid of the polypeptide encoded by the nucleic acid sequence is substituted with another amino acid having similar characteristics. Examples of conservative amino acid substitutions are ser for ala, thr, or

cys; lys for arg; gln for asn, his, or lys; his for asn; glu for asp or lys; asn for his or gln; asp for glu; pro for gly; leu for ile, phe, met, or val; val for ile or leu; ile for leu, met, or val; arg for lys; met for phe; tyr for phe or trp; thr for ser; trp for tyr; and phe for tyr.

Functional variant: The term "functional variant," as used herein, refers to a recombinant enzyme such as an INMTenzyme that comprises a nucleotide and/or amino acid sequence that is altered by one or more nucleotides and/or amino acids compared to the nucleotide and/or amino acid sequences of the parent protein and that is still capable of performing an enzymatic function (e.g., synthesis of DMT) of the parent enzyme. In other words, the modifications in the amino acid and/or nucleotide sequence of the parent enzyme may cause desirable changes in reaction parameters without altering fundamental enzymatic function encoded by the nucleotide sequence or containing the amino acid sequence. The functional variant may have conservative change including nucleotide and amino acid substitutions, additions and deletions. These modifications can be introduced by standard techniques known in the art, such as site-directed mutagenesis and random PCR-mediated mutagenesis, and may comprise natural as well as non-natural nucleotides and amino acids. Also envisioned is the use of amino acid analogs, e.g. amino acids not DNA or RNA encoded in biological systems, and labels such as fluorescent dyes, radioactive elements, electron dense agents, or any other protein modification, now known or later discovered.

Recombinant nucleic acid and recombinant protein: As used herein, a recombinant nucleic acid or protein is a nucleic acid or protein produced by recombinant DNA technology, e.g., as described in Green and Sambrook (2012).

Polypeptide, protein, and peptide: The terms “polypeptide,” “protein,” and “peptide” are used herein interchangeably to refer to amino acid chains in which the amino acid residues are linked by peptide bonds or modified peptide bonds. The amino acid chains can be of any length of greater than two amino acids. Unless otherwise specified, the terms “polypeptide,” “protein,” and “peptide” also encompass various modified forms thereof. Such modified forms may be naturally occurring modified forms or chemically modified forms. Examples of modified forms include, but are not limited to, glycosylated forms, phosphorylated forms, myristoylated forms, palmitoylated forms, ribosylated forms, acetylated forms, and the like. Modifications also include intra-molecular crosslinking and covalent attachment of various moieties such as lipids, flavin, biotin, polyethylene glycol or derivatives thereof, and the like. In addition, modifications may also include protein cyclization, branching of the amino acid chain, and cross-linking of the

protein. Further, amino acids other than the conventional twenty amino acids encoded by genes may also be included in a polypeptide.

The term "protein" or "polypeptide" may also encompass a "purified" polypeptide that is substantially separated from other polypeptides in a cell or organism in which the polypeptide naturally occurs (e.g., 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99%, 100% free of contaminants).

Primer, probe and oligonucleotide: The terms "primer," "probe," and "oligonucleotide" may be used herein interchangeably to refer to a relatively short nucleic acid fragment or sequence. They can be DNA, RNA, or a hybrid thereof, or chemically modified analogs or derivatives thereof. Typically, they are single-stranded. However, they can also be double-stranded having two complementing strands that can be separated apart by denaturation. In certain aspects, they are of a length of from about 8 nucleotides to about 200 nucleotides. In other aspects, they are from about 12 nucleotides to about 100 nucleotides. In additional aspects, they are about 18 to about 50 nucleotides. They can be labeled with detectable markers or modified in any conventional manners for various molecular biological applications.

Vector: As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is an episome, i.e., a nucleic acid capable of extra-chromosomal replication. Various vectors are those capable of autonomous replication and/expression of nucleic acids to which they are linked. Vectors capable of directing the expression of genes to which they are operatively linked are referred to herein as "expression vectors."

Linker: The term "linker" refers to a short amino acid sequence that separates multiple domains of a polypeptide. In some embodiments, the linker prohibits energetically or structurally unfavorable interactions between the discrete domains.

Codon optimized: As used herein, a recombinant gene is "codon optimized" when its nucleotide sequence is modified to accommodate codon bias of the host organism to improve gene expression and increase translational efficiency of the gene.

Expression cassette: As used herein, an "expression cassette" is a nucleic acid that comprises a gene and a regulatory sequence operatively coupled to the gene such that the promoter drives the expression of the gene in a cell. An example is a gene for an enzyme with a promoter functional in yeast, where the promoter is situated such that the promoter drives the expression of the enzyme in a yeast cell.

Compounds and abbreviations in use of or contained within systems and methods herein are provided in Table 1.

Table 1

Compounds
tryptamine
L-tryptophan
L-methionine
bufotenin
4-hydroxy-tryptamine
norbaeocystin
norpsilocin
baeocystin
psilocybin
psilocin
aeruginascin
NMT = N-methyltryptamine
DMT = N,N-dimethyltryptamine
TMT = N,N,N-trimethyltryptamine
NMTP = N-methyltryptophan or L-Abrine
DMTP = N,N-dimethyltryptophan
TMTP = N,N,N-trimethyltryptophan or Hypaphorine or Lenticin
5-HTP = 5-hydroxytryptophan
SAMe = S-Adenosyl-L-methionine
SAH = S-Adenosyl-L-homocysteine
DMC = dimethylcarbonate
DMAPP = dimethylallyl diphosphate
DMSO = dimethyl sulfoxide
5-HT = 5-hydroxytryptamine or Serotonin
NAS = N-acetylserotonin or Normelatonin
NA-MeO-T = N-acetyl-5-methoxy-tryptamine or Melatonin

5-MT = 5-methoxy-tryptamine or Mexamine
5-MeO-NMT = 5-methoxy-N-methytryptamine
5-MeO-DMT = 5-methoxy-N,N-Dimethyltryptamine
5-MeO-TMT = 5-methoxy-N,N,N-trimethyltryptamine
5-HO-NMT = 5-hydroxy-methyltryptamine
5-HO-DMT = 5-hydroxy-dimethyltryptamine or Bufotenine
5-HO-TMT = 5-hydroxy-trimethyltryptamine or Bufotenidine
NMT = N-methyltryptamine
DMT = N,N-dimethyltryptamine
TMT = N,N,N-trimethyltryptamine
NMTP = N-methyltryptophan or L-abrine
DMTP = N,N-dimethyltryptophan
TMTP = N,N,N-trimethyltryptophan or Hypaphorine or Lenticin
5-HO-NMTP = 5-hydroxy-methyltryptophan
5-HO-DMTP = 5-hydroxy-dimethyltryptophan
5-HO-TMTP = 5-hydroxy-methyltryptophan
5-MeO-NMTP = 5-methoxy-methyltryptophan
5-MeO-DMTP = 5-methoxy-dimethyltryptophan
5-MeO-TMTP = 5-methoxy-methyltryptophan
BH4 = Tetrahydrobiopterin
BH2 = Dihydrobiopterin
NADPH = Reduced nicotinamide adenine dinucleotide phosphate
NADP+ = Nicotinamide adenine dinucleotide phosphate
Acetyl-CoA = Acetyl coenzyme A
β-Methyltryptophan (β-mTrp)
N-acetyl-4-hydroxy-tryptamine
N-acetyl-psilocybin
N-acetyl-psilocin

Enzymes and regulatory proteins, and abbreviations, in use of or contained within systems and methods herein are provided in Table 2.

Table 2

Enzymes and Regulatory Proteins
INMT = Indolethylamine-N-methyltransferase; tryptamine N-methyltransferase
IOMT = indole-O-methyltransferase; hydroxytryptamine O-methyltransferase

CaffOMT = caffeic acid-O-methyltransferase
T5H = tryptamine 5’ hydroxylase
TrpM = tryptophan N-methyltransferase
PsiM = psilocybin synthase
AADC = Aromatic amino acid decarboxylase; tryptophan decarboxylase
TPH – tryptophan hydroxylase
T4H = tryptamine 4’ hydroxylase
T4H-CPR, T5H-CPR = chimeras with cytochrome p450 reductase
NAT = N-acetyltransferase
DAC = deacetylase
BH4syn = Tetrahydrobiopterin synthesis
BH4reg = Tetrahydrobiopterin regeneration
PsiK = hydroxy tryptamine kinase
TrpS = tryptophan synthase
TMO = toluene monooxygenase
ΛTMT = aminotransferase/methyltransferase fusion
oxidase = multi-copper oxidase
DMATS = dimethylallyltryptophan synthase
IDI1 = isopentenyl-diphosphate isomerase
TrpHalo = tryptophan halogenase
T5H-IOMT fusion polypeptide
AOQS = aspartate oxidase/quinolinic acid synthase fusion
TAT2 = tryptophan importer
MUP1 = methionine importer
SAM3 = SAMe importer
FEX1 = fluoride exporter

The present invention is directed to biosynthetic production of molecules that are analogs of indoles, tryptophans, and tryptamines, which can also serve as precursors to larger tryptamine

alkaloids, such as tryptamines and tryptophans modified by hydroxylation, halogenation, methylation, phosphorylation, prenylation, and halogenation in recombinant organisms.

FIG. 1 shows the chemical structures of tryptophan (top left) and tryptamine (top right), along with enzyme modifications at specific reaction sites of the tryptophan molecule. Tryptophan is the precursor to a wide array of complex natural products. The electron-rich indole of tryptophan is a weak base. These properties allow for enhanced reactivity as a substrate for numerous enzymes that perform but are not limited to the following activities: methylation, halogenation, prenylation, hydroxylation, isonitrile synthesis, nitration, O-phosphorylation, O-methylation, O-acetylation, N-acetylation, glycosylation, sulfation, cleavage, deamination, decarboxylation, and oligomerization of the molecule. This diverse array of indole intermediates provides a way to tune psychedelic effects. For example, 5-MeO-DMT is reported to be more potent than DMT in neural rodent studies (Lima da Cruz, Rafael Vitor, et al.).

FIG. 2 shows examples of various substituted indole compounds in the tryptamine and tryptophan pathways utilized in the present invention. Panel A depicts the indole ring structure with positional numbering, and tryptophan and tryptamine. Examples of 5-hydroxy modified tryptophan and tryptamine compounds are shown in Panel B; Panel C shows examples of modified 5-hydroxy tryptamines. Additionally, Panel D shows examples of modified 4-hydroxy tryptamines.

By engineering various enzymes and regulatory proteins into a microorganism, tryptophan, tryptamine and other substituted indoles can be modified into a large array of useful compounds, which can be harvested from cultures of the microorganisms.

As depicted in FIG. 3, the *de novo* biosynthesis pathway of L-tryptophan and SAMe are utilized as directing molecules in the systems and methods herein. The directing molecules lead to target molecules of the substituted indoles and tryptamine pathways, when on-pathway. In the systems and methods herein, glycolysis leads to chorismate via the shikimate pathway; glutamate biosynthesis pathway leads to L-glutamine via L-glutamate; and L-serine biosynthesis pathway leads to L-serine via 3-phospho-L-serine (i.e., dephosphorylation). Chorismate, glutamine, and L-serine are combined to form L-tryptophan as a directing molecule to be steered on-pathway for yielding substituted indoles and tryptamine pathways. In the systems and methods herein, L-methionine is a direct precursor leading to SAMe, when combined with ATP in the presence of Sam2 and Adk1 enzymes. A conversion cycle for yielding SAMe as a directing molecule also

involves the formation of S-adenyl-L-homocysteine; S-ribosyl-L-homocysteine; 4-5-dihydroxy-2,3-pentanedione; and homocysteine.

Nucleic acids

Thus, in some embodiments, provided is a non-naturally occurring nucleic acid comprising a sequence encoding an enzyme or regulatory protein in tryptamine metabolism, where the enzyme or regulatory protein is an N-methyltransferase (INMT, PsiM, TrpM), a tryptophan decarboxylase (AADC), a tryptophan hydroxylase (TPH), a tryptamine 4' hydroxylase (T4H), a tryptamine 5' hydroxylase (T5H), a truncated cytochrome p450 reductase (T4H-CPR, T5H-CPR), an hydroxytryptamine O-methyltransferase (IOMT or CaffOMT), an N-acetyltransferase (NAT), a deacetylase (DAC), a hydroxyl tryptamine kinase (PsiK), a tryptophan synthase (TrpS), a toluene monooxygenase (TMO), an aminotransferase/methyltransferase fusion (ATMT), a phosphatase, an oxidase, a dimethylallyltryptophan synthase (DMAT or DMATS), an isopentenyl-diphosphate isomerase (IDI1), a tryptophan halogenase (TrpHalo), an aspartate oxidase/quinolinic acid synthase fusion (AOQS), a tryptophan importer (TAT2), a methionine importer (MUP1), or a SAMe importer (SAM3).

These enzymes and regulatory proteins are further characterized as follows.

Indolethylamine N-methyltransferase (INMT) catalyzes the alkylation (i.e., adding a methyl (CH_3) group) of the primary amine on a tryptamine substrate. The methylation reaction uses up the methyl donor cofactor, SAMe (see FIG. 4, Panels A and C). As an example of INMT activity, INMT can act on serotonin to create 5-OH-DMT (bufotenine) or tryptamine to create DMT (FIG. 4, Panels A and C; FIG. 10, Panel A).

Indole-O-methyltransferase (IOMTor CaffOMT) catalyzes the alkylation of the primary amine on the 5-hydroxy moiety on an indole ring. The methylation reaction uses up the methyl donor cofactor, SAMe (FIG. 4, Panel D). As an example of IOMT activity, IOMT can act on bufotenine (5-OH-DMT) to create 5-MeO-DMT, or N-acetylserotonin to create melatonin (FIG. 4, Panel D).

Tryptamine 5' hydroxylase (T5H) is a p450 tryptamine hydroxylase which prefers hydroxylation at the 5' position of the indole ring, such as generating serotonin from tryptamine (FIG. 4, Panel E), in conjunction with the cofactors NAD(P)H, FMN, and FAD+. P450s such as the T5Hs are generally membrane-associated, with the N-termini imparting an effect on the efficiency of the p450 enzymatic function, including a p450's interaction with an associated CPR, which assists with electron transfer.

FIG. 5 shows a matrix of various compounds that can be made with INMT, IOMT and T5H.

Tryptophan methyltransferase (TrpM) catalyzes the alkylation of the primary amine of L-tryptophan to produce N-methyltryptophan (NMTP, also called L-abrine), the mono-methylated product; N,N-dimethyltryptophan (DMTP), the di-methylated product; and N,N,N-trimethyltryptophan (TMTP), the tri-methylated product. See FIG. 6, Panel A.

Psilocybin synthase (PsiM) is an N-methyltransferase that prefers a substituted tryptamine, such as the phosphorylated tryptamine, norbaeocystin. Novel chimeric PsiMs, were generated to remove potentially deleterious regulatory regions of the enzymes by swapping PsiM domains with the related small rRNA methyltransferases from Ascomycota, the phylum of *S. cerevisiae*.

Aromatic amino acid decarboxylase or tryptophan decarboxylase (AADC) catalyzes the decarboxylation of an aliphatic carboxylic acid (i.e., releases carbon dioxide) from compounds such as L-tryptophan to create tryptamine, 5-HTP to create serotonin; 5-OH-DMTP to create bufotenine; and 5-MeO-DMTP to create 5-MeO-DMT, as depicted in FIG. 6, Panel B.

Tryptophan hydroxylase (TPH), adds a hydroxy group to the 5-carbon of L-tryptophan. The L-tryptophan hydroxylase can catalyze the OH addition to the 5-carbon with the cofactor BH4 and oxygen (Biotechnol J. 2016 May;11(5):717-24) (FIG. 6, Panel A). BH4 is synthesized and regenerated in the cell with the BH4syn and BH4reg heterologous enzymes described herein. The BH4syn genes are enzymes that function as a GTP hydroxylase I, a 6-pyruvoyl-tetrahydropterin synthase, and a sepiapterin reductase to generate the BH4 cofactor necessary for TPH enzyme function. The BH4reg genes are enzymes that function as a 4a-hydroxytetrahydropterin dehydratase and a 6-pyruvoyl-tetrahydropterin synthase to regenerate the BH4 cofactor after conversion to HTHB by the TPH enzyme. As an example of TPH activity, TPH can act on L-tryptophan to generate 5-hydroxy-L-tryptophan (5-HTP), and 5-HTP can then be acted on by an AADC to generate serotonin.

Tryptamine 4' hydroxylase (T4H) is a p450 tryptamine hydroxylase which prefers hydroxylation at the 4' position of the indole ring, in conjunction with the cofactors NAD(P)H, FMN, and FAD+. When derived from psychedelic mushrooms, these are also called PsiH. The T4H enzyme can convert tryptamine to 4-OH-tryptamine, which is a part of the psilocybin pathway. P450s such as the T4Hs are generally membrane-associated, with the N-termini imparting an effect on the efficiency of the p450 enzymatic function, including a p450's interaction with an associated CPR, which assists with electron transfer.

From psychedelic a mushroom derived PsiH and CPR, we generated chimeric p450s and CPRs to better match a heterologous host **(SEQ ID NO:162, 179-180, and 451, 468-469)**, where the N termini of a yeast p450 and CPR replaced the N terminus. Due to the enhancing action of CPRs on p450 enzymatic activity, we determined an optimal fusion between T4H and T4H_CPR, where the T4H_CPR listed are truncated at the N termini and replaced with a linker region. In some embodiments, the T4H nucleic acids have, at the 3' end, an optimized nucleic acid encoding a T4H_CPR, e.g., having **SEQ ID NOs:171-180**, joining the sequences together to form a fusion polypeptide, e.g., having the amino acid sequence of **SEQ ID NOs:460-469** fused at the C terminus of the enzyme polypeptide, generating recombinant T4H-CPR fusion polypeptides.

Similar to the T4H CPR fusions, we generated T5H CPR fusions to enhance the hydroxylation activity. In those embodiments, the T5H nucleic acids have, at the 3' end, an optimized nucleic acid encoding a T5H-CPR, e.g., having **SEQ ID NOs:181-192**, joining the sequences together to form a fusion polypeptide, e.g., having the amino acid sequence of **SEQ ID NOs:470-481** fused at the C terminus of the enzyme polypeptide, generating recombinant T5H-CPR fusion polypeptides.

Examples of the utilization of the T4H-CPR and T5H-CPR in recombinant cells are shown in FIG. 9, Panels C and D; FIGS. 11 and 12; FIG. 13, Panel B; FIG. 14, Panel B; and FIG. 15, Panel B.

Localizing O-methyltransferase activity to hydroxylation can be beneficial for generating methoxytryptamines, such as 5-MT, 5-MeO-DMT, and melatonin. In some embodiments, the T5H nucleic acids have, at the 3' end, an optimized nucleic acid encoding an IOMT e.g., having **SEQ ID NOs:99-130**, joining the sequences together to form a fusion polypeptide, e.g., having the amino acid sequence of **SEQ ID NOs:388-419** fused at the C terminus of the enzyme polypeptide, generating recombinant T5H-IOMT fusion polypeptides.

In all of the fusions described herein, the N-terminal coding sequence has any STOP codon removed, if present, before fusion to a C-terminal coding sequence. If the N-terminal coding sequence does not have a START (ATG) codon, a START codon is added.

N-acetyltransferase (NAT) adds an acetyl group from acetyl-CoA to the terminal amino group of e.g., a tryptamine such as serotonin (FIG. 4, Panel F; FIG. 7, Panels C and D). As an example of NAT activity, NAT can act on serotonin to generate N-acetylserotonin, which in turn can be acted on by an IOMT to generate melatonin (FIG. 11).

Deacetylase (DAC) removes an acetyl group from the terminal amino group of a tryptamine such as melatonin. As an example of DAC activity, DAC can act on melatonin to create 5-MeO-tryptamine, which in turn can be acted on by an INMT to generate 5-MeO-DMT (FIG. 10, Panel B).

Hydroxy tryptamine kinase (PsiK) phosphorylates a hydroxy-indole, in conjunction with ATP. For example, PsiK can act on 4-OH tryptamine to generate norbaeocystin as part of the psilocybin pathway. PsiKs are found in certain mushrooms and parasitic fungi. For psychedelic mushroom derived PsiKs, we generated chimeric PsiKs based on yeast choline kinase to better match a heterologous host.

Non-natural tryptamine analogs can be created with the addition of a synthetic precursor to the fermentation of a recombinant host expressing enzymes capable of utilizing the substrate. For example, the addition of an alpha-methylated amino acid such as alpha-methyl tryptophan to a fermentation where an organism expresses a an indole-N-methyltransferase (INMT) leads to the generation of alpha-methylated DMT (e.g., FIG. 13).

For certain indole ring modifications, such as non-natural indoles, bacterial tryptophan synthases (TrpS) can be used to combine an indole with L-serine or L-threonine to create variants of tryptophan and beta-methyl tryptophan, respectively (FIG. 14, Panel A). While previous groups have made use of the flexibility of versions of bacterial tryptophan synthases to generate exotic tryptamines (De novo Biosynthesis of "Non-Natural" Thaxtomin Phytotoxins. Angew Chem Int Ed Engl. 2018 Jun 4;57(23):6830-6833), efficient bioproduction is limited by the toxic nature of indole. In one embodiment, TrpS is expressed as a modified secreted fusion polypeptide version of the *Salmonella* tryptophan synthase that is able to combine indole or a modified indole with L-serine or L-threonine in the extracellular space, allowing indole conversion away from the cell host. In some embodiments, a multidrug efflux exporter such as mdtEF (accessions: P37636, P37637) can be coexpressed with TrpS with exogenous indole, to enable the host cell to export indole and continue bioproduction of tryptophan and tryptamine analogs.

Alternatively to T4H, T5H, and TPH enzymes, hydroxylation of the indole ring of tryptamines and related indole-like compounds can be carried out by complexes known as toluene-monooxygenases (TMO) typically found in bacteria within the genus *Pseudomonas*. The polypeptides that form this complex can be expressed in a modified host as an alternative to P450-based hydroxylation for compounds such as psilocybin and aeruginascin, whose biosynthetic pathway involves 4'OH hydroxylation. Other non-P450 monooxygenases from genuses of

Pseudomonas and *Burkeholderia* can be optimized and expressed in a modified host for hydroxylation of different indole positions, such as the 3' carbon of the indole ring. TMO complexes are made up of several subunits. For efficient expression of TMOs in a recombinant heterologous host, we generated fusion polypeptide pairs of the four core subunits.

Beta-methylated tryptamine analogs are created by combined expression of a recombinant aminotransferase-methyltransferase (ATMT) fusion polypeptide and an aromatic amino acid decarboxylase (AADC) (FIG. 15). In nature, organisms which produce beta-methyl tryptophan typically express the aminotransferase (AT) and the methyltransferase (MT) as separate genes. Recombinant ATMT genes herein encode both domains as a single polypeptide. Combinatorial expression of ATMTs and other tryptamine modifying genes can be used to create compounds such as beta-methylated DMT and beta-methylated psilocybin.

In some embodiments, recombinant phosphatases and oxidases are used to generate hydroxylated tryptamine dimers such as one psilocin or bufotenine molecule conjugated to another psilocin or bufotenine molecule (FIG. 16). When certain psychedelic mushrooms which contain compounds such as psilocybin are damaged and cellular compartments compromised, phosphatases and oxidases, such as laccases or laccase-like multi-copper oxidases, can then come in contact with tryptamine substrate to dephosphorylate and catalyze hydroxy tryptamine polymerization. Similar polymerization which leads to 'blueing' can occur when psilocybin comes into contact with mitochondria. (Levine, Walter G), In some embodiments, the phosphatase is a recombinant alkaline phosphatase, which dephosphorylates phosphorylated tryptamines and tryptophans (FIG. 16, Panel A), such as psilocybin to psilocin. In some embodiments, the oxidase is a non-laccase member of the multi-copper oxidase superfamily, which creates hydroxy tryptamine radicals which catalyze polymerization (FIG. 16, Panel B). This dimer example and oligomerization of hydroxylated tryptamines can generate a blue color, lending the effect to colorimetric readout for compound production. Dimer variants and other oligomerized tryptamines can be separated from each other through chromatographic methods for purification. Efficient heterologous expression of certain oxidases such as laccases presents several challenges, such as N and C termini processing which may fail in a heterologous host. In some embodiments, to improve heterologous oxidase expression to biosynthetically produce tryptamine dimers and oligomers, we engineered chimeric oxidase yeast oxidase. Example includes SEQ ID NO:274,563 In one embodiment, the oxidases are also coexpressed with the yeast t-SNARE, SSO2, to improve protein expression, processing, and secretion for active enzyme SEQ ID NO:170,459.

Dimethylallyl tryptophan synthase (DMATS or DMAT) generates prenylated tryptophans and tryptamines. DMATS is a prenyltransferase that prefers the dimethylallyl diphosphate (DMAPP) prenyl donor to prenylate tryptophan and tryptamine compounds.

Localizing DMAPP generation to the DMATS enzyme can be beneficial for generating prenylated tryptophans, such as 7-dimethylallyltryptophan. In yeast, IDI1 is the enzyme which generates DMAPP as part of the mevalonate pathway. In some embodiments, the DMATS nucleic acids have, at the 3' end, an optimized nucleic acid encoding IDI1 e.g., having SEQ ID NO:67, joining the sequences together to form a fusion polypeptide, e.g., having the amino acid sequence of SEQ ID NO:356 fused at the C terminus of the enzyme polypeptide, generating recombinant DMATS-IDI1 fusion polypeptides (FIG. 8, Panel C).

Tryptophan halogenase (TrpHalo) is a flavin-associated halogenase that adds fluorine (F), chlorine (Cl), bromine (Br), and/or iodine (I) to various indoles and biogenic amines (FIG. 8, Panel A). In some embodiments, TrpHalo nucleic acids have, at the 5'end, a nucleic acid encoding an vacuolar localization tag to localize TrpHalo to a yeast vacuole, where Cl ions are stored, e.g., having SEQ ID NOs:287-289 , joining the sequences together to form a fusion polypeptide, e.g., having the amino acid sequence of SEQ ID NOs:576-578 fused at the N terminus of the enzyme polypeptide, generating recombinant fusion polypeptides.

In other embodiments, TrpHalo nucleic acids have, at the 5'end, a nucleic acid encoding a secretion tag with or without a 6xHIS tag for purification, e.g., having SEQ ID NO:1, joining the sequences together to form a fusion polypeptide, e.g., having the amino acid sequence of SEQ ID NO:290 fused at the N terminus of the enzyme polypeptide, generating recombinant fusion polypeptides. In one embodiment, TrpHalo is also coexpressed with the yeast fluoride exporter, Fex1, SEQ ID NO:66,355, to limit halide toxicity on the heterologous host.

To improve the yield of tryptophan and tryptamine variants discussed herein, modifying the heterologous host which expresses these genes and enzymes, in various combinatorial ways, to prevent tryptophan and tryptamine compound degradation is beneficial. Replacing the yeast pathway which degrades certain tryptamine and tryptophan compounds for *de novo* NAD+ production, which is an important source of cofactors for cell viability, with an alternative route to NAD+ production can preserve tryptophan as a precursor and increase product yields. In some embodiments, a new *de novo* pathway is expressed in a heterologous host, where the pathway is composed of a fusion protein containing the two enzymatic functions required to convert the amino

acid aspartate into quinolinic acid (AOQS), SEQ ID NO: 26-27,315-316, which replaces the endogenous use of tryptophan for generating quinolinic acid in the pathway for NAD+.

In some embodiments, the nucleic acids have, at the 5' end, a nucleic acid encoding codon optimized cofolding peptides to create a fusion protein, e.g., having SEQ ID NOs:256-269, joining the sequences together to form a fusion polypeptide, e.g., having the amino acid sequence of SEQ ID NOs:554-558 fused at the N terminus of the enzyme polypeptide, generating recombinant fusion polypeptides.

In some embodiments, the nucleic acids have, at the 5' end, a nucleic acid encoding a secretion signal, creating a secreted protein, e.g., having SEQ ID NOs:282-286, joining the sequences together to form a fusion polypeptide, e.g., having the amino acid sequence of SEQ ID NOs:571-575 fused at the N terminus of the enzyme polypeptide, generating recombinant fusion polypeptides.

In some embodiments, the nucleic acids have, at the 5' or 3' end, an optimized nucleic acid encoding a localization scaffold composed of multiple domains where proteins tagged with affibodies can bind and colocalize together (for example, FIG. 11, Panel A; FIG. 12, Panel A), creating a protein scaffold fusion, e.g., having SEQ ID NO:281, joining the sequences together to form a fusion polypeptide, e.g., having the amino acid sequence of SEQ ID NO:570 fused at the N or C terminus of the enzyme polypeptide, generating recombinant fusion polypeptides.

In some embodiments, the nucleic acids have, at the 5' or 3' end, an optimized nucleic acid encoding an affibody tag that can bind one of the domains of the localization scaffold, thereby colocalizing multiple enzymes and creating protein scaffold fusion, e.g., having SEQ ID NOs:259-264, joining the sequences together to form a fusion polypeptide, e.g., having the amino acid sequence of SEQ ID NOs:548-553 fused at the N or C terminus of the enzyme polypeptide, generating recombinant fusion polypeptides.

The initial substrates for DMTP, DMT, and related compound production are L-tryptophan and S-Adenosyl-L-methionine (SAMe). The initial substrate can be produced endogenously in a recombinant host as described and/or provided exogenously to a fermentation involving a recombinant host, whereby the host uptakes the starting substrates to feed into the biosynthetic pathway for indoles and tryptamines. The recombinant hosts herein described that are expressing all, one, or multiple combinations of the engineered INMT, AADC, TPH, T4H, T5H, T4H-CPR, T5H-CPR, IOMT, NAT, DAC, PsiK, TrpS, TMO, ATMT, DMATS, IDI1, and TrpHalo genes can produce tryptamine, NMTP, DMTP, TMTP, NMT, DMT, TMT, psilocybin, bufotenine, 5-MeO

DMT, 4-bromo-tryptamine, 4-dimethylallyl tryptamine, alpha-methylated DMTP, beta-methylated DMTP, melatonin, etc.

As depicted in FIGS. 4 and 9, the engineered INMT, IOMT and TrpM and INMT enzymes require a methyl donor in the form of SAMe to act on substrates in the biosynthetic pathway for substituted indoles and tryptamines such as DMTP, DMT, intermediates, and analogs. The methyltransferase activity of TrpM and INMT subsequently convert the methyl donor cofactor SAMe to SAH. Methylations can occur successively with multiple rounds of methyl donor usage. For instance, TrpM can methylate L-tryptophan to produce NMTP and continue to methylate NMTP to DMTP, and then TMTP (FIG. 6, Panel A). Similarly, an INMT can methylate tryptamine to produce NMT, and then continue to methylate NMT to DMT, and then TMT (FIG. 5).

The methylation occurs selectively at the primary amine of L-tryptophan and tryptamine in the presence of TrpM and INMT enzymes. The nitrogen in the heterocycle and hydroxyl group in the carboxylic acid of L-tryptophan are also sites of alkylation, as SAMe is a highly reactive methylating agent. The TrpM enzyme directs methylation such that di-methylation of the primary amine occurs. Using the traditional techniques of organic chemistry where robust methylating agents, such as methyl iodide, trimethyl sulfonium iodide, and dimethyl sulfate, are employed, a mixture of products is formed. The mixture of products may include: mono, di, and tri-methylation of the amine; O-methylation of the carboxylic acid (i.e., the methyl ester), and N-methylation of the indole ring. Separation of these products are tedious and reduces the yield of a desired product. Additionally, SAMe has a primary amine group which may readily undergo intramolecular methylation at the amine. The systems and methods herein in the recombinant host with TrpM and INMT enzymes maintain the structure of SAMe without methylation of the amine of the SAMe prior to methylating the amine of L-tryptophan and tryptamine.

Heterologous pathway enzymes that are expressed to produce substituted indole and tryptamine compounds such as DMTP and DMT use L-tryptophan as a directing molecule. Tryptophan production in cells is normally tightly regulated. Tryptophan accumulation in a recombinant host is increased by: (a) overexpressing feedback-resistant versions of the endogenous tryptophan-producing enzymes; (b) knocking out off-pathway tryptophan-consuming genes and enzymes; and (c) overexpressing a recombinant L-tryptophan transporter. This allows for exogenous tryptophan to be fed to the cells and transported in the recombinant host. These modifications, genes, and methods are disclosed in U.S. Patent Publication 2021/0147888, incorporated by reference.

On-pathway genes and enzymes can be overexpressed for L-tryptophan accumulation. The immediate precursors for L-tryptophan include chorismate, L-serine, and L-glutamine. To increase the on-pathway flux to L-tryptophan and the substituted indole and tryptamine pathway, off-pathway genes which consume L-tryptophan are deleted. The genes that encode the enzymes, Pdc5 and Aro10 are deleted to reduce pathway flux through the pathways that produce aromatic alcohols. The gene encoding the Aro7 enzyme is deleted to reduce production of tyrosine and phenylalanine from L-tryptophan. The genes that encode the enzymes Pdz1 and Pdz2 are also deleted to reduce pathway flux through the pABA production pathway. The gene encoding the enzyme Bna2 is deleted to reduce consumption of L-tryptophan by the kynurenine pathway.

In some embodiments, a recombinant host is modified to increase the accumulation of the methyl donor, SAMe, which is used by the recombinant TrpM and INMT enzymes to methylate indole and tryptamine molecules, such as L-tryptophan and NMT. SAMe accumulation in the recombinant host cell is increased by: (a) overexpressing enzymes to promote conversion of L-methionine to SAMe; (b) deleting off-pathway genes which encode for enzymes that deplete SAMe for unwanted side products; and (c) overexpressing a permease. This enables exogenous L-methionine to be fed to and transported into the cells.

The TrpM and INMT methyltransferase reactions consume one equivalent of adenosine triphosphate (ATP) and of SAMe. SAMe is a robust methyl donor synthesized from methionine and ATP via the L-methionine adenosyltransferase enzyme, Sam2. In various embodiments, Sam2 is overexpressed in a recombinant host to increase the conversion of L-methionine to SAMe. In other embodiments, to support the increased pathway flux and generate more ATP, the adenylate kinase enzyme, Adk1, is overexpressed. In additional embodiments, to increase the uptake of exogenous L-methionine fed into the SAMe pathway, recombinant Mup1 is overexpressed, which is a methionine transporter. SAMe is a precursor molecule for spermidine production and glycogen biosynthesis. To keep SAMe levels high in the pathways of the recombinant host and decrease off-pathway usage of SAMe, the SPE2 gene can be deleted in the recombinant host, thereby blocking the conversion of SAMe to spermidine. Glycogen biosynthesis consumes ATP, which is required for the conversion of L-methionine to SAMe. The gene encoding the enzyme Glc3 can be deleted in the recombinant host, thereby reducing production of glycogen, maintaining higher levels of ATP in the host cell, and increasing on-pathway flux of SAMe for methyltransferase activity.

As depicted in FIG. 5, the engineered INMT, T5H, and IOMT enzymes act on tryptamine substrates to generate hydroxy and methoxy tryptamine analogs such as serotonin, bufotenine (5-

OH-DMT) and 5-MeO-DMT. The initial substrates for this series of reactions includes compounds such as tryptamine and serotonin, which can be produced within a modified cell or added exogenously, in addition to L-tryptophan and S-Adenosyl-L-methionine (SAMe). The initial substrate can be produced endogenously in a recombinant host as described and/or provided exogenously to a fermentation involving a recombinant host, whereby the host uptakes the starting substrates to feed into the biosynthetic pathway for indoles and tryptamines.

As depicted in FIG. 3, de novo biosynthesis pathway of L-tryptophan and SAMe utilize L-tryptophan and SAMe as directing molecules in the systems and methods herein. The directing molecules lead to target molecules of substituted indoles and tryptamine pathways, when on-pathway. In the systems and methods herein, glycolysis leads to chorismate via the shikimate pathway; glutamate biosynthesis pathway leads to L-glutamine via L-glutamate; and L-serine biosynthesis pathway leads to L-serine via 3-phospho-L-serine (i.e., dephosphorylation). Chorismate, glutamine, and L-serine are combined to form L-tryptophan as a directing molecule to be steered on-pathway for yielding substituted indoles and tryptamine pathways. In the systems and methods herein, L-methionine is a direct precursor leading to SAMe, when combined with ATP in the presence of Sam2 and Adk1 enzymes. A conversion cycle for yielding SAMe as a directing molecule also involves the formation of S-adenyl-L-homocysteine; S-ribosyl-L-homocysteine; 4-5-dihydroxy-2,3-pentanedione; and homocysteine.

Heterologous pathway enzymes that are expressed to produce substituted indole and tryptamine compounds such as DMTP and DMT use L-tryptophan as a directing molecule. Tryptophan production in cells is normally tightly regulated. Tryptophan accumulation in a recombinant host is increased by: (a) overexpressing feedback-resistant versions of the endogenous tryptophan-producing enzymes; (b) knocking out off-pathway tryptophan-consuming genes and enzymes; and (c) overexpressing a recombinant L-tryptophan transporter. This allows for exogenous tryptophan to be fed to the cells and transported in the recombinant host. See also U.S. Patent Publication 2021/0147888.

On-pathway genes and enzymes can be overexpressed for L-tryptophan accumulation. The immediate precursors for L-tryptophan include chorismate, L-serine, and L-glutamine. To increase the on-pathway flux to L-tryptophan and the substituted indole and tryptamine pathway, off-pathway genes which consume L-tryptophan may be deleted. In some embodiments, the genes that encode the enzymes Pdc5 and Aro10 are deleted to reduce pathway flux through the pathways that produce aromatic alcohols. In other embodiments, the gene encoding the Aro7 enzyme is deleted

to reduce production of tyrosine and phenylalanine from L-tryptophan. In additional embodiments, the genes that encode the enzymes Pdz1 and Pdz2 are also deleted to reduce pathway flux through the pABA production pathway. In further embodiments, the gene encoding the enzyme Bna2 is deleted to reduce consumption of L-tryptophan by the kynurenine pathway.

In some embodiments, the nucleic acids described herewith encode a polypeptide or oligopeptide having an amino acid sequence that is naturally occurring. In other embodiments, the nucleic acids encode a polypeptide or oligopeptide having an amino acid sequence that is not naturally occurring. The encoded polypeptides or oligopeptides that are not naturally occurring can vary from a naturally occurring polypeptide or oligopeptide, or portion thereof, by a small amount (e.g., one conservative amino acid substitution or a histidine tag) or extensively (e.g., further comprising a fusion peptide, a substituted or added domain from another protein, a scaffold, etc.).

The nucleic acids can be derived from a naturally occurring gene from any source, e.g., any microorganism, protist, plant, or animal.

In some embodiments, the gene for the enzyme or regulatory protein is derived from a bacterium. It is envisioned that an enzyme or regulatory protein derived from any bacterium now known or later discovered can be utilized in the present invention. For example, the bacterium can be from phylum Abditibacteriota, including class Abditibacteria, including order Abditibacteriales; phylum Abyssubacteria or Acidobacteria, including class Acidobacteriia, Blastocatellia, Holophagae, Thermoanaerobaculia, or Vicinamibacteria, including order Acidobacteriales, Bryobacterales, Blastocatellales, Acanthopleuribacterales, Holophagales, Thermotomaculales, Thermoanaerobaculales, or Vicinamibacteraceae; phylum Actinobacteria, including class Acidimicrobiia, Actinobacteria, Actinomarinidae, Coriobacteriia, Nitriliruptoria, Rubrobacteria, or Thermoleophilia, including orders Acidimicrobiales, Acidothermales, Actinomycetales, Actinopolysporales, Bifidobacteriales, Nanopelagicales, Catenulisporales, Corunebacteriales, Cryptosporangiales, Frankiales, Geodermatophilales, Glycomycetales, Jiangellales, Micrococcales, Micromonosporales, Nakamurellales, Propionibacteriales, Pseudonocardiales, Sporichthyales, Streptomycetales, Streptosporangiales, Actinomarinales, Coriobacteriales, Eggerthellales, Egibacterales, Egicoccales, Euzebyales, Nitriliruptorales, Gaiellales, Rubrobacterales, Solirubrobacterales, or Thermoleophilales; phylum Aquificae, including class Aquificae, including order Aquificales or Desulfurobacteriales; phylum Armatimonadetes, including class Armatimonadia, including order Armatimonadales,

Capsulimonadales, Chthonomonadetes, Chthonomonadales, Fimbriimonadia, or Fimbriimonadales; phylum Aureabacteria or Bacteroidetes, including class Armatimonadia, Bacteroidia, Chitinophagia, Cytophagia, Flavobacteria, Saprospiria or Sphingobacteriia, including order Bacteroidales, Marinilabiliales, Chitinophagales, Cytophagales, Flavobacteriales, Saprospirales, or Sphingopacteriales; phylum Balneolaeota, Caldiserica, Calditrichaeota, or Chlamydiae, including class Balneolia, Caldisericia, Calditrichae, or Chlamydia, including order Balneolales, Caldisericales, Calditrichales, Anoxychlamydiales, Chlamydiales, or Parachlamydiales; phylum Chlorobi or Chloroflexi, including class Chlorobia, Anaerolineae, Ardenticatenia, Caldilineae, Thermofonsia, Chloroflexia, Dehalococcoidia, Ktedonobacteria, Tepidiformia, Thermoflexia, Thermomicrobia, or Sphaerobacteridae, including order Chlorobiales, Anaerolineales, Ardenticatenales, Caldilineales, Chloroflexales, Herpetosiphonales, Kallotenuales, Dehalococcoidales, Dehalogenimonas, Ktedonobacterales, Thermogemmatisporales, Tepidiformales, Thermoflexales, Thermomicrobiales, or Sphaerobacterales; phylum Chrysiogenetes, Cloacimonetes, Coprothermobacterota, Cryosericota, or Cyanobacteria, including class Chrysiogenetes, Coprothermobacteria, Gloeobacteria, or Oscillatoriophycideae, including order Chrysiogenales, Coprothermobacterales, Chroococcidiopsidales, Gloeoemargaritales, Nostocales, Pleurocapsales, Spirulinales, Synechococcales, Gloeobacterales, Chroococcales, or Oscillatoriales; phyla: Eferribacteres, Deinococcus-thermus, Dictyoglomi, Dormibacteraeota, Elusimicrobia, Eremiobacteraeota, Fermentibacteria, or Fibrobacteres, including class Deferribacteres, Deinococci, Dictyoglomia, Elusimicrobia, Endomicrobia, Chitinispirillia, Chitinivibrionia, or Fibrobacteria, including order Deferribacterales, Deinococcales, Thermales, Dictyoglomales, Elusimicrobiales, Endomicrobiales, Chitinspirillales, Chitinvibrionales, Fibrobacterales, or Fibromonadales; phylum Firmicutes, Fusobacteria, Gemmatimonadetes, or Hydrogenedentes, including class Bacilli, Clostridia, Erysipelotrichia, Limnochordia, Negativicutes, Thermolithobacteria, Tissierellia, Fusobacteriia, Gemmatimonadetes, Longimicrobia, including order Bacillales, Lactobacillales, Borkfalkiales, Clostridiales, Halanaerobiales, Natranaerobiales, Thermoanaerobacterales, Erysipelotrichales, Limnochordales, Acidaminococcales, Selenomonadales, Veillonellales, Thermolithobacterales, Tissierellales, Fusobacteriales, Gemmatimonadales, or Longimicrobia; phylum Hydrogenedentes, Ignavibacteriae, Kapabacteria, Kiritimatiellaeota, Krumholzibacteriota, Kryptonia, Latescibacteria, LCP-89, Lentisphaerae, Margulisbacteria, Marinimicrobia, Melainabacteria, Nitrospinae, or Omnitrophica, including class

Ignavibacteria, Kiritimatiellae, Krumholzibacteria, Lentisphaeria, Oligosphaeria, or Nitrospinae, including order Ignavibacteriales, Kiritimatiellales, Krumholzibacteriales, Lentisphaerales, Victivallales, Oligosphaerales, or Nitrospinia; phylum Omnitrophica or Planctomycetes, including class Brocadiae, Phycisphaerae, Planctomycetia, or Phycisphaerales, including order Sedimentisphaerales, Tepidisphaerales, Gemmatales, Isosphaerales, Pirellulales, or Planctomycetales; phylum Proteobacteria including class Acidithiobacillia, Alphaproteobacteria, Betaproteobacteria, Lambdaproteobacteria, Muproteobacteria, Deltaproteobacteria, Epsilonproteobacteria, Gammaproteobacteria, Hydrogenophilalia, Oligoflexia, or Zetaproteobacteria, including order Acidithiobacillales, Caulobacterales, Emcibacterales, Holosporales, Iodidimonadales, Kiloniellales, Kopriimonadales, Kordiimonadales, Magnetococcales, Micropepsales, Minwuiales, Parvularculales, Pelagibacterales, Rhizobiales, Rhodobacterales, Rhodospirillales, Rhodothalassiales, Rickettsiales, Sneathiellales, Sphingomonadales, Burkholderiales, Ferritrophicales, Ferrovales, Neisseriales, Nitrosomonadales, Procabacteriales, Rhodocyclales, Bradymonadales, Acidulodesulfobacterales, Desulfarculales, Desulfobacterales, Desulfovibrionales, Desulfurellales, Desulfuromonadales, Myxococcales, Syntrophobacterales, Campylobacterales, Nautiliales, Acidiferrobacterales, Aeromonadales, Alteromonadales, Arenicellales, Cardiobacteriales, Cellvibrionales, Chromatiales, Enterobacterales, Immundisolibacterales, Legionellales, Methylococcales, Nevskiales, Oceanospirillales, Orbales, Pasteurellales Pseudomonadales, Salinisphaerales, Thiotrichales, Vibrionales, Xanthomonadales, Hydrogenophilales, Bacteriovoracales, Bdellovibrionales, Oligoflexales, Silvanigrellales, or Mariprofundales; phylum Rhodothermaeota, Saganbacteria, Sericytochromatia, Spirochaetes, Synergistetes, Tectomicrobia, or Tenericutes, including class Rhodothermia, Spirochaetia, Synergistia, Izimaplasma, or Mollicutes, including order Rhodothermales, Brachyspirales, Brevinematales, Leptospirales, Spirochaetales, Synergistales, Acholeplasmatales, Anaeroplasmatales, Entomoplasmatales, or Mycoplasmatales; phylum Thermodesulfobacteria, Thermotogae, Verrucomicrobia, or Zixibacteria, including class Thermodesulfobacteria, Thermotogae, Methylacidiphilae, Opitutae, Spartobacteria, or Verrucomicrobiae, including order Thermodesulfobacteriales, Kosmotogales, Mesoaciditogales, Petrotogales, Thermotogales, Methylacidiphilales, Opitutales, Puniceicoccales, Xiphinematobacter, Chthoniobacterales, Terrimicrobium, or Verrucomicrobiales.

In other embodiments, the gene for the enzyme or regulatory protein is derived from an archaeon. It is envisioned that an enzyme or regulatory protein derived from any archaeon now

known or later discovered can be utilized in the present invention. For example, the archaeon can be from phylum Euryarchaeota, including class Archaeoglobi, Hadesarchaea, Halobacteria, Methanobacteria, Methanococci, Methanofastidiosa, Methanomicrobia, Methanopyri, Nanohaloarchaea, Theionarchaea, Thermococci, or Thermoplasmata, including order Archaeoglobales, Hadesarchaeales, Halobacteriales, Methanobacteriales, Methanococcales, Methanocellales, Methanomicrobiales, Methanophagales, Methanosarcinales, Methanopyrales, Thermococcales, Methanomassiliicoccales, Thermoplasmatales, or Nanoarchaeales; DPANN superphylum, including subphyla Aenigmarcheota, Altiarchaeota, Diapherotrites, Micrarchaeota, Nanoarchaeota, Pacearchaeota, Parvarchaeota, or Woesearchaeota; TACK superphylum, including subphylum Korarchaeota, Crenarchaeota, Aigarchaeota, Geoarchaeota, Thaumarchaeota, or Bathyarchaeota; Asgard superphylum including subphylium Odinarchaeota, Thorarchaeota, Lokiarchaeota, Helarchaeota, or Heimdallarchaeota.

In additional embodiments, the gene for the enzyme or regulatory protein is derived from a fungus. It is envisioned that an enzyme or regulatory protein derived from any fungus now known or later discovered can be utilized in the present invention. This includes but is not limited to the phyla Chytridiomycota, Basidiomycota, Ascomycota, Blastocladiomycota, Ascomycota, Microsporidia, Basidiomycota, Glomeromycota, Symbiomycota, and Neocallimastigomycota. For example, the fungus can be from the phylum Ascomycota, including classes and orders Pezizomycotina, Arthoniomycetes, Coniocybomycetes, Dothideomycetes, Eurotiomycetes, Geoglossomycetes, Laboulbeniomycetes, Lecanoromycetes, Leotiomycetes, Lichinomycetes, Orbiliomycetes, Pezizomycetes, Sordariomycetes, Xylonomycetes, Lahmiales, Itchiclahmadion, Triblidiales, Saccharomycotina, Saccharomycetes, Taphrinomycotina, Archaeorhizomyces, Neolectomycetes, Pneumocystidomycetes, Schizosaccharomycetes, Taphrinomycetes; phylum Basidiomycota including subphyla or classes Pucciniomycotina, Ustilaginomycotina, Wallemiomycetes, and Entorrhizomycetes; subphylum Agaricomycotina including classes Tremellomycetes, Dacrymycetes, and Agaricomycetes; phylum Symbiomycota, including class Entorrhizomycota; subphylum Ustilaginomycotina including classes Ustilaginomycetes and Exobasidiomycetes; phylum Glomeromycota including classes Archaeosporomycetes, Glomeromycetes, and Paraglomeromycetes; subphylum Pucciniomycotina including orders and classes: Pucciniomycotina, Cystobasidiomycetes, Agaricostilbomycetes, Microbotryomycetes, Atractiellomycetes, Classiculomycetes, Mixiomycetes, and Cryptomycocolacomycetes; subphylum incertae sedis Mucoromyceta including orders Calcarisporiellomycota and

Mucoromycota; phylum Mortierellomyceta including class Mortierellomycota; subphylum incertae sedis Entomophthoromycotina including order Entomophthorales; phylum Zoopagomyceta including classes Basidiobolomycota, Entomophthoromycota, Kickxellomycota, and Zoopagomycotina; subphylum incertae sedis Mucoromycotina including orders Mucorales, Endogonales, and Mortierellales; phylum Neocallimastigomycota including class Neocallimastigomycetes; phylum Blastocladiomycota including classes Physodermatomycetes and Blastocladiomycetes; phylum Rozellomyceta including classes Rozellomycota and Microsporidia; phylum Aphelidiomyceta including class Aphelidiomycota; Chytridiomyceta including classes Chytridiomycetes and Monoblepharidomycetes; and phylum Oomycota including classes or orders Leptomitales, Myzocytiopsidales, Olpidiopsidales, Peronosporales, Pythiales, Rhipidiales, Salilagenidiales, Saprolegniales, Sclerosporales, Anisolpidiales, Lagenismatales, Rozellopsidales, and Haptoglossales.

In additional embodiments, the gene for the enzyme or regulatory protein is derived from the organism below. This includes but is not limited to: Acanthurus tractus, Aplysina aerophoba, Bos Taurus, Bufo bufo, Bufotes viridis, Chrysochloris asiatica, Fukomys damarensis, Homo sapiens, Rattus norvegicus, Rhinella marina, Rhinella spinulosa, Schistosoma mansoni, Xenopus laevis, Xenopus tropicalis, Acacia koa, Arabidopsis thaliana, Brassica oleracea, Citrus sinensis, Hordeum vulgare, Juglans cinereal, Lophophora williamsii, Nymphaea colorata, Oryza sativa, Ricinus communis, Solanum lycopersicum, Sorghum bicolor, Theobroma cacao, and Triticum aestivum.

In some embodiments, the nucleic acids are codon optimized to improve expression, e.g., using techniques as disclosed in US Patent No. 10,435,727. More specifically, optimized nucleotide sequences are generated based on a number of considerations: (1) For each amino acid of the recombinant polypeptide to be expressed, a codon (triplet of nucleotide bases) is selected based on the frequency of each codon in the Saccharomyces cerevisiae genome; the codon can be chosen to be the most frequent codon or can be selected probabilistically based on the frequencies of all possible codons. (2) In order to prevent DNA cleavage due to a restriction enzyme, certain restriction sites are removed by changing codons that cover those sites. (3) To prevent low-complexity regions, long repeats (sequences of any single base longer than five bases) are modified. (2) and (3) are performed recursively to ensure that codon modification does not lead to additional undesirable sequences. (4) A ribosome binding site is added to the N-terminus. (5) A stop codon is added. (6) A localization signal is removed or replaced.

In some of the above embodiments, the nucleic acid provided herein comprises the sequence of any one of SEQ ID NOs:1-289.

In various embodiments, the nucleic acids further comprise additional nucleic acids encoding amino acids that are not part of the included enzymes or regulatory proteins herein. In some of these embodiments, the additional sequences encode additional amino acids present when the nucleic acid is translated, encoding, for example, a cofolding peptide, as previously discussed, or an additional protein domain, with or without a linker sequence, creating a fusion protein. Other examples are localization sequences, i.e., signals directing the localization of the folded protein to a specific subcellular compartment or membrane. Additional nonlimiting examples are an affibody tag, a localization scaffold, a vacuolar localization tag, a secretion signal, and a 6xhis tag.

In some embodiments, the nucleic acid comprises additional nucleotide sequences that are not translated. Nonlimiting examples include promoters, terminators, barcodes, Kozak sequences, targeting sequences, and enhancer elements. Particularly useful here are promoters that are functional in yeast.

Expression of a gene encoding an enzyme or regulatory protein is determined by the promoter controlling the gene. In order for a gene to be expressed, a promoter must be present within 1,000 nucleotides upstream of the gene. A gene is generally cloned under the control of a desired promoter. The promoter regulates the amount of enzyme expressed in the cell and also the timing of expression, or expression in response to external factors such as sugar source.

Any promoter now known or later discovered can be utilized to drive the expression of the enzymes and regulatory proteins described herein. See e.g. http://parts.igem.org/Yeast for a listing of various yeast promoters. Exemplary promoters listed in Table 3 below drive strong expression, constant gene expression, medium or weak gene expression, or inducible gene expression. Inducible or repressible gene expression is dependent on the presence or absence of a certain molecule. For example, the *GAL1, GAL7,* and *GAL10* promoters are activated by the presence of the sugar galactose and repressed by the presence of the sugar glucose. The *HO* promoter is active and drives gene expression only in the presence of the alpha factor peptide. The *HXT1* promoter is activated by the presence of glucose while the *ADH2* promoter is repressed by the presence of glucose.

Table 3: Exemplary yeast promoters

Strong constitutive promoters	Medium and weak constitutive promoters	Inducible/repressible promoters
TEF1	*STE2*	*GAL1*
PGK1	*TPI1*	*GAL7*
PGI1	*PYK1*	*GAL10*
TDH3		*HO*
		HXT1
		ADH2

In various embodiments, the nucleic acid is in an expression cassette, e.g., a yeast expression cassette. Any yeast expression cassette capable of expressing the enzyme in a yeast cell can be utilized.

Additional regulatory elements can also be present in the expression cassette, including restriction enzyme cleavage sites, antibiotic resistance genes, integration sites, auxotrophic selection markers, origins of replication, and degrons.

The expression cassette can be present in a vector that, when transformed into a host cell, either integrates into chromosomal DNA or remains episomal in the host cell. Such vectors are well-known in the art. See e.g. http://parts.igem.org/Yeast for a listing of various yeast vectors.

A nonlimiting example of a yeast vector is a yeast episomal plasmid (YEp) that contains the pBluescript II SK(+) phagemid backbone, an auxotrophic selectable marker, yeast and bacterial origins of replication and multiple cloning sites enabling gene cloning under a suitable promoter (see Table 3). Other exemplary vectors include pRS series plasmids.

Host cells

The present invention is also directed to genetically engineered host cells that comprise the above-described nucleic acids. Such cells may be, e.g., any species of filamentous fungus, including but not limited to any species of *Aspergillus*, which have been genetically altered to produce precursor molecules, intermediate molecules, or cannabinoid molecules. Host cells may also be any species of bacteria, including but not limited to *Escherichia*, *Corynebacterium*, *Caulobacter*, *Pseudomonas*, *Streptomyces*, *Bacillus*, or *Lactobacillus*.

In some embodiments, the genetically engineered host cell is a yeast cell, which may comprise any of the above-described expression cassettes, and capable of expressing the recombinant enzyme encoded therein.

Any yeast cell capable of being genetically engineered can be utilized in these embodiments. Nonlimiting examples of such yeast cells include species of *Saccharomyces*, *Candida*, *Pichia*, *Schizosaccharomyces*, *Scheffersomyces*, *Blakeslea*, *Rhodotorula*, or *Yarrowia*.

These cells can achieve gene expression controlled by inducible promoter systems; natural or induced mutagenesis, recombination, and/or shuffling of genes, pathways, and whole cells performed sequentially or in cycles; overexpression and/or deletion of single or multiple genes and reducing or eliminating parasitic side pathways that reduce precursor concentration.

The host cells of the recombinant organism may also be engineered to produce any or all precursor molecules necessary for the biosynthesis of substituted indoles, tryptophans and tryptamines.

Construction of *Saccharomyces cerevisiae* strains expressing the enzymes and regulatory proteins provided herein is carried out via expression of a gene which encodes for the enzyme. The gene encoding the enzyme can be cloned into vectors with the proper regulatory elements for gene expression (e.g. promoter, terminator) and the derived plasmid can be confirmed by DNA sequencing. As an alternative to expression from an episomal plasmid, the gene encoding the enzyme may be inserted into the recombinant host genome. Integration may be achieved by a single or double cross-over insertion event of a plasmid, or by nuclease-based genome editing methods, as are known in the art e.g. CRISPR, TALEN and ZFR. Strains with the integrated gene can be screened by rescue of auxotrophy and genome sequencing. See, e.g., Green and Sambrook (2012).

FIGS. 9-15 provide nonlimiting examples of host cells utilizing the nucleic acids provided herein.

In some embodiments, the recombinant microorganism expresses TPH, TrpM, and AADC, where the recombinant microorganism produces at least one hydroxy substituted tryptamine compound, e.g., bufotenine, 5-OH-NMT, or 5-OH-TMT (FIG. 9, Panel A).

In other embodiments, the recombinant microorganism expresses TPH, TrpM, AADC, and IOMT, where the recombinant microorganism produces at least one methoxy substituted tryptamine compound, e.g., 5-MeO-NMT, 5-MeO-DMT, or 5-MeO-TMT (FIG. 9, Panel B).

In additional embodiments, the recombinant microorganism expresses AADC, T5H or T5H-CPR and INMT, where the recombinant microorganism produces at least one hydroxy substituted tryptamine compound, e.g., bufotenine, 5-OH-NMT, or 5-OH-TMT (FIG. 9, Panel C).

In further embodiments, the recombinant microorganism expresses AADC, T5H or T5H-CPR, INMT, and IOMT, where the recombinant microorganism produces at least one methoxy substituted tryptamine compound, e.g., 5-MeO-NMT, 5-MeO-DMT, or 5-MeO-TMT (FIG. 9, Panel D).

In other embodiments, the recombinant microorganism expresses TrpM and TPH, where the recombinant microorganism produces at least one hydroxy substituted tryptophan compound, e.g., 5-HTP, 5-OH-NMTP, 5-OH-DMTP or 5-OH-TMTP.

In additional embodiments, the recombinant microorganism expresses TrpM, TPH and IOMT, where the recombinant microorganism produces at least one methoxy substituted tryptophan compound, e.g., 5-MeO-NMTP, 5-MeO-DMTP or 5-MeO-TMTP.

In further embodiments, the recombinant microorganism expresses INMT and T5H, where the recombinant microorganism produces at least one hydroxy substituted tryptamine compound, e.g., bufotenine, 5-OH-NMT, or 5-OH-TMT.

In other embodiments, the recombinant microorganism expresses INMT, T5H and IOMT, where the recombinant microorganism produces at least one methoxy substituted tryptamine compound, e.g., 5-MeO-NMT, 5-MeO-DMT, or 5-MeO-TMT.

In additional embodiments, the recombinant microorganism expresses INMT, where the recombinant microorganism produces at least one hydroxy substituted tryptophan compound, e.g., 5-OH-NMTP, 5-OH-DMTP or 5-OH-TMTP.

In further embodiments, the recombinant microorganism expresses INMT and IOMT, where the recombinant microorganism produces at least one methoxy substituted tryptophan compound, e.g., 5-MeO-NMTP, 5-MeO-DMTP or 5-MeO-TMTP.

In other embodiments, the recombinant microorganism expresses INMT and AADC, where the recombinant microorganism produces at least one hydroxy substituted tryptamine compound, e.g., bufotenine, 5-OH-NMT, or 5-OH-TMT.

In additional embodiments, the recombinant microorganism expresses INMT, AADC and IOMT, where the recombinant microorganism produces at least one methoxy substituted tryptamine compound, e.g., 5-MeO-NMT, 5-MeO-DMT, or 5-MeO-TMT.

In further embodiments, the recombinant microorganism expresses INMT, where the recombinant microorganism produces at least one hydroxy substituted tryptamine compound, e.g., bufotenine, 5-OH-NMT, or 5-OH-TMT.

In other embodiments, the recombinant microorganism expresses INMT and IOMT, where the recombinant microorganism produces at least one methoxy substituted tryptamine compound, e.g., 5-MeO-NMT, 5-MeO-DMT, or 5-MeO-TMT.

In additional embodiments, the recombinant microorganism expresses INMT, where the recombinant microorganism produces at least one methoxy substituted tryptamine compound, e.g., 5-MeO-NMT, 5-MeO-DMT, or 5-MeO-TMT.

As depicted in FIG. 11, in some embodiments, the recombinant microorganism expresses AADC, IOMT, T5H or T5H-CPR, and NAT, where the recombinant microorganism produces a compound in the melatonin pathway, e.g., serotonin or melatonin. In some of these embodiments, the enzymes are on a scaffold to facilitate pathway throughput.

As depicted in FIG. 12, in some embodiments, the recombinant microorganism expresses AADC, T4H or T4H-CPR, PsiK and INMT (PsiM), where the recombinant microorganism produces a compound in the psilocybin pathway, e.g., baeocystin, psilocybin or aeruginascin. In some of these embodiments, the enzymes are on a scaffold to facilitate pathway throughput.

In accordance with the present invention, a recombinant host may also be modified to increase the accumulation of the methyl donor, SAMe, which is used by the recombinant TrpM and INMT enzymes to methylate indole and tryptamine molecules such as L-tryptophan and NMT. SAMe accumulation in the recombinant host cell may be increased by: (d) overexpressing enzymes to promote conversion of L-methionine to SAMe; (e) deleting off-pathway genes that encode for enzymes that deplete SAMe for unwanted side products; and (f) overexpressing a permease, which enables exogenous L-methionine to be fed to and transported into the cells.

The TrpM and INMT methyltransferase reactions consume one equivalent of adenosine triphosphate (ATP) and of SAMe. SAMe is a robust methyl donor synthesized from methionine and ATP via the L-methionine adenosyltransferase enzyme, Sam2. Sam2 may be overexpressed in a recombinant host to increase the conversion of L-methionine to SAMe. To support the increased pathway flux and generate more ATP, the adenylate kinase enzyme, Adk1, may also be overexpressed. To increase the uptake of exogenous L-methionine to feed into the SAMe pathway, recombinant Mup1, which is a methionine transporter, may be overexpressed.

SAMe is a precursor molecule for spermidine production and glycogen biosynthesis. To keep SAMe levels high in the pathways of the recombinant host and decrease off-pathway usage of SAMe, the SPE2 gene may be deleted in the recombinant host, thereby blocking the conversion of SAMe to spermidine. Glycogen biosynthesis consumes ATP, which is required for the conversion of L-methionine to SAMe. The gene encoding the enzyme Glc3 may be deleted in the recombinant host, thereby reducing production of glycogen, maintaining higher levels of ATP in the host cell, and increasing on-pathway flux of SAMe for methyltransferase activity. FIG. 10 depicts a recombinant host modified to express the enzymes enabling uptake and biosynthesis of indole and tryptamine precursors and the enzymes to create tryptamine, DMTP, DMT, and related substituted indole and tryptamine compounds.

Recombinant enzymes and regulatory proteins

The present invention is also directed to a non-naturally occurring enzyme or regulatory protein comprising an amino acid sequence encoded by any of the nucleic acids described above. In some embodiments, the amino acid sequence is 85%, 90%, 95%, 98%, or 100% identical to any one of SEQ ID NO:290-578. In these embodiments, the enzyme or regulatory protein can be isolated *in vitro* and used *in vitro* to provide enzyme activity. Alternatively, as discussed above, the enzyme can be expressed in a recombinant organism, e.g., a microorganism or a plant. In some of these embodiments, the recombinant microorganism is a bacterium, for example an *E. coli*. In other embodiments, the recombinant microorganism is a yeast cell, e.g., a species *of Saccharomyces* (for example *S. cerevisiae*), *Candida, Pichia, Schizosaccharomyces, Schеffersomyces, Blakeslea, Rhodotorula, Aspergillus* or *Yarrowia*.

Methods

The systems and methods herein include: (i) growing modified recombinant host cells and thereby yielding a recombinant host organism; (ii) expressing engineered indole and tryptamine biosynthesis genes and enzymes in the recombinant host organism; (iii) producing or synthesizing substituted indoles and tryptamines in the recombinant host organism; (iv) fermenting the recombinant host organism; and (v) isolating the substituted indoles and tryptamines from the recombinant host organism. Endogenous pathways of the recombinant host can be modified by the systems and methods herein to produce high purity substituted indoles and tryptamines.

To produce the desired substituted indole, the nucleic acid encoding the enzymes and/or regulatory proteins are introduced into a host cell using standard cell (e.g., yeast) transformation techniques (Green and Sambrook, 2012). Cells are subjected to fermentation under conditions that activate the promoter controlling the synthesis of the enzyme and/or regulatory protein. The broth may be subsequently subjected to HPLC analysis to determine the presence or yield of the desired substituted indole, as in FIGS. 17-24 and 26.

In various embodiments, the host cells are provided with various feedstocks to drive production of the desired substituted indole, e.g., glucose, fructose, sucrose, ethanol, fatty acids, glycerol, molasses, corn steep liquor, dairy, fish waste, etc. for example as discussed in US Patent Application 17/078636.

In some embodiments, for recombinant enzyme purification, the gene encoding the enzyme and/or regulatory protein is cloned into an expression vector such as the pET expression vectors from Novagen, transformed into a protease deficient strain of *E. coli* such as BL21 and expressed by induction with IPTG. The protein of interest may be tagged with an affinity tag to facilitate purification, e.g. hexahistidine, GST, calmodulin, TAP, AP, CAT, HA, FLAG, MBP etc. Coexpression of a bacterial chaperone such as dnaK, GroES/GroEL or SecY may help facilitate protein folding. See Green and Sambrook (2012).

Any of the enzymes and/or regulatory proteins described above can also be produced in transgenic plants, using techniques known in the art (see, e.g., Keshavareddy et al., 2018). In these embodiments, the above-described nucleic acid encoding the enzyme and/or regulatory protein further comprises a promoter functional in a plant. In various embodiments, the nucleic acid is in a plant expression cassette. Any plant capable of being transformed with the nucleic acid can be utilized here. In some embodiments, the plant is a tobacco or cannabis.

Preferred embodiments are described in the following examples. Other embodiments within the scope of the claims herein will be apparent to one skilled in the art from consideration of the specification or practice of the invention as disclosed herein. It is intended that the specification, together with the examples, be considered exemplary only, with the scope and spirit of the invention being indicated by the claims, which follow the examples.

In the examples below, genetically engineered host cells may be any species of yeast herein, including but not limited to any species of *Saccharomyces*, *Candida*, *Schizosaccharomyces*, *Yarrowia*, etc.,which have been genetically altered to produce precursor

molecules, intermediate molecules, and psilocybin molecules. Additionally, genetically engineered host cells may be any species of filamentous fungus, including but not limited to any species of *Aspergillus*, which have been genetically altered to produce precursor molecules such as L-tryptophan and substituted indole and tryptamine molecules. Some of the species of yeast herein for the recombinant host organism include but are not limited to: *Schizosaccharomyces cerevisiae, Schizosaccharomyces japonicus, Schizosaccharomyces pombe, Schizosaccharomyces cryophilus, Saccharomyces cerevisiae, Kluyveromyces lactis, Kluyveromyces dobzhanskii,* and *Yarrowia lipolytica*.

In these examples, the gene sequences from gene source organisms are codon optimized to improve expression using techniques disclosed in U.S. Patent 10,435,727.

DNA sequences are synthesized and cloned using techniques known in the art. Gene expression can be controlled by inducible or constitutive promoter systems using the appropriate expression vectors. Genes are transformed into an organism using standard yeast or fungus transformation methods to generate modified host strains (i.e., the recombinant host organism). The modified strains express genes for: (i) producing L-tryptophan, SAMe and precursor molecules to L-tryptophan and SAMe; (ii) increasing an output of L-tryptophan molecules and precursor molecules to L-tryptophan and SAMe molecules; (iii) increasing the import of exogenous L-tryptophan, L-methionine, SAMe and TMG into the host strain; and (iv) the genes for biosynthetic pathways that generate DMT, DMTP, bufotenine, 5-MeO-DMT and all intermediate indole and tryptamine compounds synthesized and described herein. In the presence or absence of exogenous L-tryptophan, L-methionine, SAMe, TMG, 5-HTP, melatonin, and serotonin, fermentations are run to determine if the cell will convert the fed precursors into tryptamine, serotonin, methylated versions of serotonin, melatonin, or methylated versions of melatonin. The L-tryptophan, SAMe, hydroxylation, decarboxylation, and methylation pathway genes herein can be integrated into the genome of the cell or maintained as an episomal plasmid. Samples are: (i) prepared and extracted using a combination of fermentation, dissolution, and purification steps; and (ii) analyzed by HPLC for the presence of directing molecules (e.g., SAMe and L-tryptophan), precursor molecules, intermediate molecules, and target molecules such as bufotenine and 5-MeO-DMT.

Using the systems and methods herein, the genes which can be expressed to encode for a corresponding enzyme or other type of proteins include but are not limited to: ENO2, TAL1, ARO1, ADK1, MUP1, SAM2, MHT1, SAM4, SAM3, TAT2, AADC, TRPM, INMT, TPH, genes

encoding enzymes for the BH4 biosynthesis pathway, genes encoding enzymes for the BH4 regeneration pathway, T5H, IOMT, caffOMT, NAT, DAC, T4H, PsiK, oxidase, phosphatase, TrpHalo, DMAT, T4H-CPR, T5H-CPR, TrpS, and ATMT. For example, the AADC gene is expressed, or overexpressed, to encode for the aromatic amino decarboxylase enzyme; the TRPM gene is expressed to encode for the TrpM enzyme; and so forth. Gene sequences can be determined using standard techniques known in the art, e.g., the techniques disclosed in U.S. Patent 10,671,632.

EXAMPLES

Example 1 - Construction of Saccharomyces cerevisiae platform strains with elevated indole and tryptamine precursors.

The construction of *Saccharomyces cerevisiae* platform strains with elevated metabolic flux towards L-tryptophan is carried out by overexpressing five optimized enzymes in or upstream of the shikimate pathway to make the aromatic compound intermediate, chorismate, and one optimized enzyme in the tryptophan pathway to make L-tryptophan. Further, tryptophan levels in the cell are enhanced with the expression of TAT2, a tryptophan importer, and L-tryptophan supplementation in the media up to 1% mass to volume. Finally, five enzymes are deleted in the cell to decrease off-pathway consumption of the L-tryptophan. The genetically modified host described herein can be the same host used for production of psilocybin and DMT as both production pathways use the precursor, L-tryptophan. A specific description of the strain with elevated L-tryptophan is disclosed in U.S. Patent Publication 2021/0147888.

Example 2 - Construction of *Saccharomyces cerevisiae* platform strains with synthesis of methyl donor

Construction of *Saccharomyces cerevisiae* platform strains with elevated SAMe production is carried out via expression of SAM2, a SAMe synthetase gene. The SAM2 gene is cloned from *Saccharomyces cerevisiae* using techniques known in the art. The gene can be cloned into vectors with the proper regulatory elements for gene expression (e.g. promoter, terminator) and the derived plasmid can be confirmed by DNA sequencing. As an alternative to expression from an episomal plasmid, the SAM2 gene is inserted into the recombinant host genome. Integration is achieved by a single cross-over insertion event of the plasmid. Strains with the integrated gene can be screened by rescue of auxotrophy and genome sequencing.

Example 3 – Construction of *Saccharomyces cerevisiae* platform strains with elevated methyl donor production

Construction of *Saccharomyces cerevisiae* platform strains with elevated SAMe production via expression of the ADK1, adenylate kinase gene. The ADK1 gene is cloned from *Saccharomyces cerevisiae* using techniques known in the art. The gene can be cloned into vectors with the proper regulatory elements for gene expression (e.g. promoter, terminator) and the derived plasmid can be confirmed by DNA sequencing. As an alternative to expression from an episomal plasmid, the ADK1 gene is inserted into the recombinant host genome. Integration is achieved by a single cross-over insertion event of the plasmid. Strains with the integrated gene can be screened by rescue of auxotrophy and genome sequencing.

Further SAM accumulation for methyl donor availability is achieved herein by engineering the homocysteine to methionine side of the methylation pathway. SAH is generated after methylation of serotonin and other intermediates to produce bufotenine and other compounds described herein. SAH is recycled back to methionine after methyl donation by TMG (trimethylglycine) or betaine. TMG is fed to the cells up to 1% (v/v) in the growth media. Two *Saccharomyces cerevisiae* genes, MHT1 and SAM4 encode the enzymes, Mht1 and Sam4, that are responsible for homocysteine re-methylation using TMG as a methyl donor. MHT1 and SAM4 are overexpressed from a high copy vector with a strong promoter.

Example 4 – Construction of *Saccharomyces cerevisiae* platform strains with enhanced uptake of methyl donor precursors.

Construction of *Saccharomyces cerevisiae* platform strains with elevated SAMe production is carried out via expression of MUP1, the methionine permease gene. The MUP1 gene is cloned from *Saccharomyces cerevisiae* using techniques known in the art. The gene can be cloned into vectors with the proper regulatory elements for gene expression (e.g. promoter, terminator) and the derived plasmid can be confirmed by DNA sequencing. As an alternative to expression from an episomal plasmid, the MUP1 gene is inserted into the recombinant host genome. Integration is achieved by a single cross-over insertion event of the plasmid. Strains with the integrated gene can be screened by rescue of auxotrophy and genome sequencing.

Example 5 – Construction of *Saccharomyces cerevisiae* platform strains with enhanced uptake of methyl donors.

Herein we describe a strategy to increase the SAM accumulation by increasing transport of exogenous SAM into the cell. SAM levels are increased by overexpressing the gene, SAM3. SAM3 encodes for the Sam3 protein, the predominant *Saccharomyces cerevisiae* transporter that is responsible for SAM import. SAM3 is expressed from a high-copy vector with a strong promoter and media is supplemented with 0.5 – 1.0 mM SAMe.

Example 6 – Construction of *Saccharomyces cerevisiae* platform strains with decreased off-pathway flux of methyl donors

Construction of *Saccharomyces cerevisiae* platform strains with elevated metabolic flux towards SAMe is carried out via deletion of SPE2 to reduce SAMe decarboxylation. Deletion of SPE2 is performed by replacement of the SPE2 gene with the URA3 cassette in the recombinant host. The SPE2 URA3 knockout fragment, carrying the marker cassette, URA3, and homologous sequence to the targeted gene, SPE2, can be generated by bipartite PCR amplification. The PCR product is transformed into a recombinant host and transformants can be selected on synthetic URA drop-out media. Further verification of the modification in said strain can be carried out by genome sequencing, then analyzed by the techniques disclosed in U.S. Patent 10,671,632.

Example 7 – Construction of *Saccharomyces cerevisiae* platform strains with decreased off-pathway flux of methyl donor precursors

Saccharomyces cerevisiae platform strains are constructed with elevated metabolic flux towards SAMe via deletion of GLC3 to reduce ATP consumption. Deletion of GLC3 is performed by replacement of the GLC3 gene with the URA3 cassette in the recombinant host. The GLC3 URA3 knockout fragment, carrying the marker cassette, URA3, and homologous sequence to the targeted gene, GLC3, can be generated by bipartite PCR amplification. The PCR product is transformed into a recombinant host and transformants can be selected on synthetic URA drop-out media. Further verification of the modification in said strain can be carried out by genome sequencing and analyzed by the techniques disclosed in U.S. Patent 10,671,632.

Example 8 – Construction of *Saccharomyces cerevisiae* platform strains with increased Tryptophan accumulation

Saccharomyces cerevisiae platform strains with accumulation of tryptophan are generated by deletion of BNA2. Bna2 is an enzyme necessary for *de novo* NAD+ production from

tryptophan. Deletion of BNA2 is performed by replacement of the BNA2 gene with the URA3 cassette in the recombinant host. The BNA2 URA3 knockout fragment, carrying the marker cassette, URA3, and homologous sequence to the targeted gene, BNA2, can be generated by bipartite PCR amplification. The PCR product is transformed into a recombinant host and transformants can be selected on synthetic URA drop-out media. Further verification of the modification in said strain can be carried out by genome sequencing and analyzed by the techniques disclosed in U.S. Patent 10,671,632.

Example 9 – Expression of recombinant L-tryptophan methyltransferases in a modified host organism

Construction of *Saccharomyces cerevisiae* NMTP, DMTP, and TMTP production strains is carried out via expression of the TrpM methyltransferase gene. The optimized TrpM gene is synthesized using DNA synthesis techniques known in the art. The optimized gene can be cloned into vectors with the proper regulatory elements for gene expression (e.g. promoter, terminator) and the derived plasmid can be confirmed by DNA sequencing. As an alternative to expression from an episomal plasmid, the optimized TrpM gene is inserted into the recombinant host genome. Integration is achieved by a single cross-over insertion event of the plasmid. Strains with the integrated gene can be screened by rescue of auxotrophy and genome sequencing.

Example 10 – Expression of recombinant aromatic amino acid decarboxylases in a modified host organism

Construction of *Saccharomyces cerevisiae* tryptamine production strains is carried out via expression of the AADC gene which encodes the enzyme that converts L-tryptophan to tryptamine. AACD also encodes the enzyme that converts 5HTP to serotonin. This specific conversion may be carried out by the same enzyme encoded by the AADC gene that converts L-tryptophan to tryptamine. It also may be carried out by the gene product of a novel AADC described herein. The optimized AADC gene is synthesized using DNA synthesis techniques known in the art. The optimized gene can be cloned into vectors with the proper regulatory elements for gene expression (e.g. promoter, terminator) and the derived plasmid can be confirmed by DNA sequencing. As an alternative to expression from an episomal plasmid, the optimized AADC gene is inserted into the recombinant host genome. Integration is achieved by a single

cross-over insertion event of the plasmid. Strains with the integrated gene can be screened by rescue of auxotrophy and genome sequencing.

Example 11 – Expression of recombinant L-tryptophan hydroxylases in a modified host organism

Construction of the *Saccharomyces cerevisiae* 5-HTP production strains is carried out via expression of the gene that encodes tryptophan hydroxylase. 5-HTP is a precursor compound for production of serotonin and variants described herein. Tryptophan hydroxylase activity is dependent on the availability of the BH4 cofactor. The optimized TPH, BH4 biosynthesis and BH4 regeneration genes are synthesized using DNA synthesis techniques known in the art. The optimized genes can be cloned into vectors with the proper regulatory elements for gene expression (e.g. promoter, terminator) and the derived plasmid can be confirmed by DNA sequencing. As an alternative to expression from an episomal plasmid, the optimized TPH, BH4 biosynthesis and BH4 regeneration genes are inserted into the recombinant host genome. Integration is achieved by a single cross-over insertion event of the plasmid. Strains with the integrated gene can be screened by rescue of auxotrophy and genome sequencing.

Example 12 – Expression of recombinant tryptamine 5-hydroxylases in a modified host organism

Construction of the *Saccharomyces cerevisiae* serotonin production strains is carried out via expression of the gene that encodes tryptamine 5-hydroxylase (T5H). 5-HT or serotonin is a precursor compound for production of bufotenine and variants described herein. T5H activity is dependent on the availability of the intermediate indole compound, tryptamine, production of which is disclosed in U.S. Patent Publication 2021/0147888 and further described herein.

T5H, as a cytochrome p450-containing monooxygenase, is also dependent on the cytochrome p450 reductase enzyme (CPR) for full activity. The CPR facilitates electron transfer from the NAD(P)H. The optimized T5H and CPR genes are synthesized using DNA synthesis techniques known in the art. The optimized genes can be cloned into vectors with the proper regulatory elements for gene expression (e.g. promoter, terminator) and the derived plasmid can be confirmed by DNA sequencing. As an alternative to expression from an episomal plasmid, the optimized T5H and CPR genes are inserted into the recombinant host genome. Integration is

achieved by a single cross-over insertion event of the plasmid. Strains with the integrated gene can be screened by rescue of auxotrophy and genome sequencing.

Example 13 – Expression of recombinant indolethylamine-N-methyltransferase (INMT) in a modified host organism

Construction of *Saccharomyces cerevisiae* DMT production strains is carried out via expression of the INMT gene which encodes the enzyme that methylates tryptamine to DMT. INMT also encodes the enzyme that converts serotonin to bufotenine. Finally, INMT encodes the enzyme that converts 5-MeO-tryptamine to 5-Meo-DMT. These unique conversions may be carried out by the same enzyme encoded by the INMT gene that converts tryptamine to DMT. It also may be carried out by the gene product of a novel INMT described herein. The optimized INMT gene is synthesized using DNA synthesis techniques known in the art. The optimized gene can be cloned into vectors with the proper regulatory elements for gene expression (e.g. promoter, terminator) and the derived plasmid can be confirmed by DNA sequencing. As an alternative to expression from an episomal plasmid, the optimized INMT gene is inserted into the recombinant host genome. Integration is achieved by a single cross-over insertion event of the plasmid. Strains with the integrated gene can be screened by rescue of auxotrophy and genome sequencing.

Example 14 – Expression of 5-hydroxyindole-O-methyltransferase (IOMT) or caffeic acid-O-methyltransferase (CaffOMT) in a modified host organism

Construction of *Saccharomyces cerevisiae* 5-MeO-DMT production strains is carried out via expression of the IOMT gene which encodes the enzyme that methylates the 5-OH in bufotenine, an intermediate derived from the INMT conversion of serotonin, described herein. The IOMT gene also encodes for the enzyme that converts serotonin to 5-MeO-tryptamine in the first intermediate to make melatonin. The IOMT enzyme also methylates the 5-OH of N-acetyl-serotonin to generate melatonin as an intermediate to make 5-MeO-tryptamine and further, 5-MeO-DMT. Alternatively, the enzyme that converts serotonin to 5-MeO-tryptamine can be carried out with a CaffOMT enzyme, an enzyme shared with the phenylpropanoid biosynthesis pathway. This same CaffOMT enzyme can also methylate N-acetyl-serotonin to generate melatonin. The optimized IOMT or CaffOMT gene is synthesized using DNA synthesis techniques known in the art. The optimized gene can be cloned into vectors with the proper regulatory elements for gene expression (e.g. promoter, terminator) and the derived plasmid can be confirmed by DNA

sequencing. As an alternative to expression from an episomal plasmid, the optimized IOMT or CaffOMT gene is inserted into the recombinant host genome. Integration is achieved by a single cross-over insertion event of the plasmid. Strains with the integrated gene can be screened by rescue of auxotrophy and genome sequencing.

Example 15 – Expression of recombinant N-acetyl transferase (NAT) and melatonin deacetylase (NAD) in a modified host organism

Construction of *Saccharomyces cerevisiae* 5-MeO-DMT production strains can alternatively be carried out via expression of the two more enzymes, NAT and NAD. NAT acetylation of serotonin produces the intermediate N-acetyl-serotonin or NAS. NAS is converted to melatonin with the IOMT (or CaffOMT) enzyme described herein. DAC deacetylases melatonin to 5-MeO-tryptamine which is converted to 5-MeO-DMT via the INMT enzyme described herein. The optimized NAT and DAC genes are synthesized using DNA synthesis techniques known in the art. The optimized gene can be cloned into vectors with the proper regulatory elements for gene expression (e.g. promoter, terminator) and the derived plasmid can be confirmed by DNA sequencing. As an alternative to expression from an episomal plasmid, the optimized NAT and DAC genes are inserted into the recombinant host genome. Integration is achieved by a single cross-over insertion event of the plasmid. Strains with the integrated gene(s) can be screened by rescue of auxotrophy and genome sequencing.

Example 16 – Construction of *Saccharomyces cerevisiae* platform strains with accumulation of serotonin

Serotonin is the precursor molecule for both bufotenine and 5-MeO-DMT. Construction of a *Saccharomyces cerevisiae* serotonin strain is carried out by expression of AADC and TPH or AADC and T5H genes described herein for the enzymatic conversion of L-tryptophan to serotonin. Exogenous serotonin is also fed to the strains to increase precursor levels at concentrations of 0.5 mM to 2 mM. Exogenous 5-HTP with expression of the AADC gene is fed to the cells as a mechanism to increase the serotonin precursor.

In order to accumulate serotonin in the cell and prevent off pathway conversion of serotonin to unwanted products, the endogenous Saccharomyces cerevisiae gene, PAA1 (YDR071C) is deleted. PAA1 is a polyamine acetyltransferase that would acetylate serotonin and use up valuable acetyl-CoA.

Example 17 – Method of Growth

Modified host cells that yield substituted indoles and tryptamine compounds, such as a bufotenine-producing strain herein, express engineered bufotenine biosynthesis genes and enzymes. More specifically, the bufotenine-producing strain herein is grown in a minimal, complete culture media containing yeast nitrogen base, amino acids, vitamins, ammonium sulfate, and a carbon source of glucose and galactose. The recombinant host cells are grown in 24-well plates or shake flasks in a volume range of 2 mL to 100 mL of media starting from an inoculation density of OD_{600nm}=1. Exogenous serotonin, melatonin, tryptamine, 5HTP, SAMe and TMG can be added to media to supplement the precursor pool for final compound production or support methyl donor accumulation.

Example 18 - Conversion of melatonin to 5-methoxy-tryptamine using a bio-based enzyme factory

Herein we describe a strategy for 5-methoxy-tryptamine (5-MT) production by recombinant expression and secretion of the melatonin deacetylase, DAC in BL21(DE3)pLysS *E. coli*. The DAC enzyme is cloned into a high-copy vector with key features that allow 1) tight induction by the lactose analog, **β**-D-thiogalactoside (IPTG) 2) an N-terminal secretory signal peptide [MKKTAIAIAVALAGFATVAQA (**SEQ ID NO:286,575)**] and 3) C-terminal fusion to a HIS tag for purification. *E. coli* cells harboring the NAD-expression vector are grown in M9 minimal media with 1% glucose for 18h at 37 °C and shaking at 300 rpm. Concentrated cell culture is diluted to an OD_{600} =1 in fresh M9 minimal media with 1% glucose and 0.2 mM IPTG. After a 3h induction at 18 °C and 300 rpm shaking, melatonin is added to the media at a final concentration of 1-2 mM. Cells are grown at room temperature for 48h, shaking at 300 rpm. Media is collected at 24h and 48h and analyzed by HPLC as described herein.

Alternatively, we describe a strategy for production of the compound, 5-methoxy-tryptamine (5-MT) by recombinant expression and secretion of the melatonin deacetylase, DAC in *Saccharomyces cerevisiae*. The DAC enzyme is cloned into a high-copy vector with key features that allow 1) tight induction by the sugar, galactose 2) an N-terminal alpha factor secretion leader sequence, [MEGVSLEKREAEA (SEQ ID NO:574] and 3) c-terminal fusion to a HIS tag for purification. *Saccharomyces cerevisiae* cells harboring the DAC-expression vector are grown in CM minimal media with 2% glucose for 18h at 30 degrees C and shaking at 300 rpm. Concentrated

cell culture is diluted to an OD_{600} =1 in fresh CM minimal media with 2% galactose. After 24h of induction at 30 degrees C and 300 rpm shaking, melatonin is added to the media at a final concentration of 1-2 mM. Cells are grown at 30 °C and 300 rpm shaking for 48h. Media is collected at 24h and 48h and analyzed by HPLC as described herein.

Alternatively, we describe a strategy for production of the compound, 5-methoxy-tryptamine (5-MT) by recombinant expression and secretion of the melatonin deacetylase (DAC) in *Komagataella phaffii* (*Pichia pastoris*). The DAC enzyme is cloned into a high-copy vector with key features that allow 1) induction by methanol with the AOX1 promoter and 2) a secretion signal consisting of the α-factor pro region. *K. phaffii* cells harboring the DAC enzyme are inoculated into 5 mL of YPD in a 15-mL culture tube. After a day of incubation at 30 °C with shaking at 220 rpm, an aliquot of the culture is diluted to an OD_{600}=0.2 in 5 mL of BMG (buffered minimal glycerol media) in a 15-mL culture tube. This tube is incubated under the same conditions as before. The following day, the culture is centrifuged at 3000 rpm (2000×*g*) for 5 min and resuspended in 25 mL BMM (buffered minimal methanol media) to attain an OD_{600} = 1.0. 25 mL of this culture is placed in a 250-mL baffled flask, and during this induction phase, the cells are incubated at 25 °C with shaking at 150 rpm to reduce loss of methanol. After 1 day of induction, an additional dose of 125 μL methanol is added (yielding a final concentration of 0.5%), melatonin is added to the media at a final concentration of 1-2 mM, and the incubation is continued for another day. After 48 h of induction, media is collected at at 24h and 48h and analyzed by HPLC as described herein.

Example 19 - Conversion of melatonin to 5-methoxy-NMT, 5-methoxy-DMT, and 5-methoxy-TMT using a bio-based enzyme factory

Herein we describe a strategy for 5-MeO-NMT, 5-MeO-DMT, and 5-MeO-TMT production by recombinant expression and secretion of the indolethylamine-*N*-methyltransferase (INMT) in in BL21(DE3)pLysS *E. coli*. The INMT enzyme is cloned into a high-copy vector with key features that allow 1) tight induction by the lactose analog, **β**-D-thiogalactoside (IPTG) 2) an N-terminal secretory signal peptide [MKKTAIAIAVALAGFATVAQA (SEQ ID NO:574)] and 3) C-terminal fusion to a HIS tag for purification. *E. coli* cells harboring the INMT-expression vector are grown in M9 minimal media with 1% glucose for 18h at 37 °C and shaking at 300 rpm. Concentrated cell culture is diluted to an OD_{600} =1 in fresh M9 minimal media with 1% glucose and 0.2 mM IPTG. After a 3h induction at 18 °C and 300 rpm shaking, melatonin is added to the

media at a final concentration of 1-2 mM and SAMe is added to the media at a final concentration of 1-2 mM. Cells are grown at room temperature for 48h, shaking at 300 rpm. Media is collected at 24h and 48h and analyzed by HPLC as described herein.

Alternatively, we describe a strategy for production of the compounds, 5-MeO-NMT, 5-MeO-DMT, and 5-MeO-TMT by recombinant expression and secretion of the indolethylamine-*N*-methyltransferase (INMT) in *Saccharomyces cerevisiae*. The INMT enzyme is cloned into a high-copy vector with key features that allow 1) tight induction by the sugar, galactose 2) an N-terminal alpha factor secretion leader sequence, [MEGVSLEKREAEA (SEQ ID NO:574)] and 3) c-terminal fusion to a HIS tag for purification. *Saccharomyces cerevisiae* cells harboring the INMT-expression vector are grown in CM minimal media with 2% glucose for 18h at 30 °C and shaking at 300 rpm. Concentrated cell culture is diluted to an OD_{600} =1 in fresh CM minimal media with 2% galactose. After 24h of induction at 30 °C and 300 rpm shaking, melatonin is added to the media at a final concentration of 1-2 mM and SAMe is added to the media at a final concentration of 1-2 mM. Cells are grown at 30 °C and 300 rpm shaking for 48h. Media is collected at 24h and 48h and analyzed by HPLC as described herein.

Alternatively, we describe a strategy for production of the compounds, 5-MeO-NMT, 5-MeO-DMT, and 5-MeO-TMT by recombinant expression and secretion of the indolethylamine-*N*-methyltransferase (INMT) in *Komagataella phaffii*. The INMT enzyme is cloned into a high-copy vector with key features that allow 1) induction by methanol with the AOX1 promoter and 2) a secretion signal consisting of the α-factor pro region. *K. phaffii* cells harboring the DAC enzyme are inoculated into 5 mL of YPD in a 15-mL culture tube. After a day of incubation at 30 °C with shaking at 220 rpm, an aliquot of the culture is diluted to an OD_{600}=0.2 in 5 mL of BMG (buffered minimal glycerol media) in a 15-mL culture tube. This tube is incubated under the same conditions as before. The following day, the culture is centrifuged at 3000 rpm (2000×*g*) for 5 min and resuspended in 25 mL BMM (buffered minimal methanol media) to attain an OD_{600} = 1.0. 25 mL of this culture is placed in a 250-mL baffled flask, and during this induction phase, the cells are incubated at 25 °C with shaking at 150 rpm to reduce loss of methanol. After 1 day of induction, an additional dose of 125 μL methanol is added (yielding a final concentration of 0.5%), melatonin is added to the media at a final concentration of 1-2 mM, SAMe is added to the media at a final concentration of 1-2 mM, and the incubation is continued for another day. After 48 h of induction, media is collected at 24h and 48h and analyzed by HPLC as described herein.

Example 20 - Conversion of tryptamine to NMT, DMT, and TMT using a bio-based enzyme factory

Herein we describe a strategy for NMT, DMT, and TMT production by recombinant expression and secretion of the indolethylamine-*N*-methyltransferase (INMT) in BL21(DE3)pLysS *E. coli*. The INMT enzyme is cloned into a high-copy vector with key features that allow 1) tight induction by the lactose analog, **β**-D-thiogalactoside (IPTG) 2) an N-terminal secretory signal peptide [MKKTAIAIAVALAGFATVAQA (SEQ ID NO:574)] and 3) C-terminal fusion to a HIS tag for purification. *E. coli* cells harboring the INMT-expression vector are grown in M9 minimal media with 1% glucose for 18h at 37 °C and shaking at 300 rpm. Concentrated cell culture is diluted to an OD_{600} =1 in fresh M9 minimal media with 1% glucose and 0.2 mM IPTG. After a 3h induction at 18 °C and 300 rpm shaking, tryptamine is added to the media at a final concentration of 1-2 mM and SAMe is added to the media at a final concentration of 1-2 mM. Cells are grown at room temperature for 48h, shaking at 300 rpm. Media is collected at 24h and 48h and analyzed by HPLC as described herein.

Alternatively, we describe a strategy for production of the compounds, NMT, DMT, and TMT by recombinant expression and secretion of the indolethylamine-*N*-methyltransferase (INMT) in *Saccharomyces cerevisiae*. The INMT enzyme is cloned into a high-copy vector with key features that allow 1) tight induction by the sugar, galactose 2) an N-terminal alpha factor secretion leader sequence, [MEGVSLEKREAEA (SEQ ID NO:574)] and 3) c-terminal fusion to a HIS tag for purification. *Saccharomyces cerevisiae* cells harboring the INMT-expression vector are grown in CM minimal media with 2% glucose for 18h at 30 °C and shaking at 300 rpm. Concentrated cell culture is diluted to an OD_{600} =1 in fresh CM minimal media with 2% galactose. After 24h of induction at 30 °C and 300 rpm shaking, tryptamine is added to the media at a final concentration of 1-2 mM and SAMe is added to the media at a final concentration of 1-2 mM. Cells are grown at 30 °C and 300 rpm shaking for 48h. Media is collected at 24h and 48h and analyzed by HPLC as described herein.

Alternatively, we describe a strategy for production of the compounds, NMT, DMT, and TMT by recombinant expression and secretion of the indolethylamine-*N*-methyltransferase (INMT) in *Komagataella phaffii*. The INMT enzyme is cloned into a high-copy vector with key features that allow 1) induction by methanol with the AOX1 promoter and 2) a secretion signal consisting of the α-factor pro region. *K. phaffii* cells harboring the DAC enzyme are inoculated into 5 mL of YPD in a 15-mL culture tube. After a day of incubation at 30 °C with shaking at 220

rpm, an aliquot of the culture is diluted to an OD_{600}=0.2 in 5 mL of BMG (buffered minimal glycerol media) in a 15-mL culture tube. This tube is incubated under the same conditions as before. The following day, the culture is centrifuged at 3000 rpm (2000×*g*) for 5 min and resuspended in 25 mL BMM (buffered minimal methanol media) to attain an OD_{600} = 1.0. 25 mL of this culture is placed in a 250-mL baffled flask, and during this induction phase, the cells are incubated at 25 °C with shaking at 150 rpm to reduce loss of methanol. After 1 day of induction, an additional dose of 125 µL methanol is added (yielding a final concentration of 0.5%), tryptamine is added to the media at a final concentration of 1-2 mM, SAMe is added to the media at a final concentration of 1-2 mM, and the incubation is continued for another day. After 48 h of induction, media is collected at at 24h and 48h and analyzed by HPLC as described herein.

Example 21 - Conversion of serotonin to 5-OH-NMT, 5-OH-DMT, and 5-OH-TMT using a bio-based enzyme factory

Herein we describe a strategy for 5-OH-NMT, 5-OH-DMT, and 5-OH-TMT production by recombinant expression and secretion of the indolethylamine-*N*-methyltransferase (INMT) in BL21(DE3)pLysS *E. coli*. The INMT enzyme is cloned into a high-copy vector with key features that allow 1) tight induction by the lactose analog, **β**-D-thiogalactoside (IPTG) 2) an N-terminal secretory signal peptide [MKKTAIAIAVALAGFATVAQA (SEQ ID NO:574)] and 3) C-terminal fusion to a HIS tag for purification. *E. coli* cells harboring the INMT-expression vector are grown in M9 minimal media with 1% glucose for 18h at 37 °C and shaking at 300 rpm. Concentrated cell culture is diluted to an OD_{600} =1 in fresh M9 minimal media with 1% glucose and 0.2 mM IPTG. After a 3h induction at 18 °C and 300 rpm shaking, serotonin is added to the media at a final concentration of 5 mM and SAMe is added to the media at a final concentration of 1-2 mM. Cells are grown at room temperature for 48h, shaking at 300 rpm. Media is collected at 24h and 48h and analyzed by HPLC as described herein.

Alternatively, we describe a strategy for production of the compounds, 5-OH-NMT, 5-OH-DMT, and 5-OH-TMT by recombinant expression and secretion of the indolethylamine-*N*-methyltransferase (INMT) in *Saccharomyces cerevisiae*. The INMT enzyme is cloned into a high-copy vector with key features that allow 1) tight induction by the sugar, galactose 2) an N-terminal alpha factor secretion leader sequence, [MEGVSLEKREAEA (SEQ ID NO:574)] and 3) c-terminal fusion to a HIS tag for purification. *Saccharomyces cerevisiae* cells harboring the INMT-expression vector are grown in CM minimal media with 2% glucose for 18h at 30 °C and shaking

at 300 rpm. Concentrated cell culture is diluted to an OD_{600} =1 in fresh CM minimal media with 2% galactose. After 24h of induction at 30 °C and 300 rpm shaking, serotonin is added to the media at a final concentration of 5 mM and SAMe is added to the media at a final concentration of 1-2 mM. Cells are grown at 30 °C and 300 rpm shaking for 48h. Media is collected at 24h and 48h and analyzed by HPLC as described herein.

Alternatively, we describe a strategy for production of the compounds, 5-OH-NMT, 5-OH-DMT, and 5-OH-TMT by recombinant expression and secretion of the indolethylamine-*N*-methyltransferase (INMT) in *Komagataella phaffii*. The INMT enzyme is cloned into a high-copy vector with key features that allow 1) induction by methanol with the AOX1 promoter and 2) a secretion signal consisting of the α-factor pro region. *K. phaffii* cells harboring the DAC enzyme are inoculated into 5 mL of YPD in a 15-mL culture tube. After a day of incubation at 30 °C with shaking at 220 rpm, an aliquot of the culture is diluted to an OD_{600}=0.2 in 5 mL of BMG (buffered minimal glycerol media) in a 15-mL culture tube. This tube is incubated under the same conditions as before. The following day, the culture is centrifuged at 3000 rpm (2000×*g*) for 5 min and resuspended in 25 mL BMM (buffered minimal methanol media) to attain an OD_{600} = 1.0. 25 mL of this culture is placed in a 250-mL baffled flask, and during this induction phase, the cells are incubated at 25 °C with shaking at 150 rpm to reduce loss of methanol. After 1 day of induction, an additional dose of 125 μL methanol is added (yielding a final concentration of 0.5%), serotonin is added to the media at a final concentration of 5 mM, SAMe is added to the media at a final concentration of 1-2 mM and the incubation is continued for another day. After 48 h of induction, media is collected at 24h and 48h and analyzed by HPLC as herein.

Example 22 - Purification of recombinant INMT enzyme to use for *in vitro* reactions

The INMT enzyme is cloned into a high-copy vector with key features that allow 1) tight induction by the lactose analog, **β**-D-thiogalactoside (IPTG) 2) an N-terminal secretory signal peptide [MKKTAIAIAVALAGFATVAQA (SEQ ID NO:574] and 3) C-terminal fusion to a HIS tag for purification. *E. coli* cells harboring the INMT-expression vector are grown in M9 minimal media with 1% glucose for 18h at 37 °C and shaking at 300 rpm. Concentrated cell culture is diluted to an OD_{600} =1 in fresh M9 minimal media with 1% glucose and 0.2 mM IPTG and grown for 48h.

The supernatant containing the recombinant proteins is equilibrated in binding buffer (50 mM sodium phosphate, 0.5 M NaCl, 20 mM imidazole, 10% glycerol, 10 mM 2-mercaptoethanol,

1 mM PMSF, Complete EDTA-free (1 tablet/100 ml), 20 mM 1-phenyl-2- thio-urea; pH 7.4) and centrifuged at 2,500g for 5 min to remove insoluble matter. Then the supernatant is filtered through a 0.45 μm filter (Millipore, MA, USA) and applied onto a HisTrap HP column (GE Healthcare Bioscience). The recombinant proteins are eluted with a step gradient of imidazole (concentrations of 5, 20, 40 and 300 mM). Fractions are analyzed by SDS-PAGE and stored at -80 °C before use.

Example 23 - *In vitro* reactions with purified INMT enzyme or INMT lysate

Purified INMT protein is resuspended in activity buffer [100 mM sodium phosphate buffer, pH 6.55, PMSF (1mM), EDTA-free protease inhibitor] cocktail at working concentration (Roche, Meylan, France) for use in *in vitro* assays. 0.1 mg/mL of INMT protein is added to a tube with a final volume of 600 uL per sample and added to 100 mM sodium phosphate buffer (pH 7.5), 2 mM tryptamine, serotonin, or melatonin, 2 mM S-adenosylmethionine, and 5 mM $MgCl_2$.

Alternatively, 0.1 mg/mL BSA protein-equivalent of INMT lysate is used in the same reaction. INMT lysate is derived from *E. coli* cells harboring the INMT-expression vector. They are grown in M9 minimal media with 1% glucose for 18h at 37 °C and shaking at 300 rpm. Concentrated cell culture is diluted to an OD_{600} =1 in fresh M9 minimal media with 1% glucose and 0.2 mM IPTG and grown for 48h. Cell pellets are resuspended in 100 mM sodium phosphate buffer at pH 7.5 and lysed using sonication. After lysis, samples are pelleted by centrifugation (16,000g, 4 °C, 20 min) and supernatant containing INMT is harvested.

Example 24 – Method of Growth

Modified host cells that yield substituted indoles and tryptamine compounds, such as the DMTP-producing strain herein, express engineered DMTP biosynthesis genes and enzymes. More specifically, the DMTP-producing strain herein is grown in a minimal, complete culture media containing yeast nitrogen base, amino acids, vitamins, ammonium sulfate, and a carbon source of glucose and galactose. The recombinant host cells are grown in 24-well plates or shake flasks in a volume range of 2 mL to 100 mL of media starting from an inoculation density of OD600nm=1. Exogenous L-tryptophan and L-methionine up to 1% can be added to media to supplement the precursor pool for DMTP production. Exogenous L-tryptophan can be taken up by strains expressing the TAT2 L-tryptophan importer protein. Exogenous L-methionine can be taken up by strains expressing the MUP1 L-methionine permease protein. The strains herein can be harvested

during a fermentation period ranging from 12 hours onward from the start of pathway enzyme induction.

Example 25 – Detection of Isolated Product

To identify fermentation-derived tryptamine, DMTP, NMT, DMT, and all other products of a recombinant host expressing an engineered biosynthetic pathway for substituted indoles (see Fig. 11), an Agilent 1100 series liquid chromatography (LC) system equipped with a HILIC column (Primesep 100, SIELC, Wheeling, IL USA) is used. A gradient is used of mobile phase A (ultraviolet (UV) grade H_2O+0.2% TFA) and mobile phase B (UV grade acetonitrile+0.2% TFA). Column temperature is set at 40 °C. Compound absorbance is measured at 270 nm using a diode array detector (DAD) and spectral analysis from 200nm to 400nm wavelengths. A secondary wavelength of 315 nm is used to selectively detect 4-hydroxy and 4-methoxy substituted indoles. A 0.1 milligram (mg)/milliliter (mL) analytical standard is made from certified reference material for each of the substituted indoles (Cayman Chemical Company, USA). Each sample is prepared by diluting fermentation biomass from a recombinant host expressing the engineered biosynthesis pathway 1:1 in 100% ethanol and filtered in 0.2 um nanofilter vials. The retention time and UV-visible absorption spectrum (i.e., spectral fingerprints) of the samples are compared to the analytical standard retention time and UV-visible spectra (i.e. spectral fingerprint) when identifying the substituted indole compounds. For example, FIG. 17 depicts the detection of tryptamine isolated from a fermentation with a recombinant host expressing enzymes for L-tryptophan to tryptamine conversion. Detection and isolation is depicted by retention time matching of fermentation derived tryptamine with a tryptamine analytical standard, along with a matching UV-vis spectral fingerprint (i.e. spectral fingerprint) of the fermentation derived tryptamine with the tryptamine analytical standard. This also corroborates that the recombinant host is able to successfully convert L-tryptophan to tryptamine, which further validates that the systems and methods herein direct molecules into tryptamine pathways.

As another example, Fig. 18 depicts the production, detection, and isolation of the substituted indole, DMTP, from a fermentation of a modified recombinant host expressing the DMTP pathway. The retention time and UV-vis spectral absorption (i.e. spectral fingerprint) of the DMTP isolated from fermentation is identical to the retention time and UV-vis spectral absorption (i.e. spectral fingerprint) of the DMTP analytical standard. FIG. 18 also depicts a negative control fermentation from a host strain not expressing the TrpM enzyme or the DMTP

pathway, and this strain does not produce DMTP. The modified host strain expressing the TrpM and DMTP producing pathway, highlighted in FIG. 18, is able to produce DMTP. FIG. 26 depicts the production, isolation, and identification of the dimethylated tryptamine, DMT, derived from a fermentation of a recombinant host expressing the pathway for substituted indoles and tryptamines. The fermentation derived DMT is identified by matching retention times with the DMT analytical standard. Spectral library identification of the fermentation derived DMT matches the UV-vis absorption spectrum (i.e. spectral fingerprint) of the DMT analytical standard.

Example 26 - Synthetic preparation of substituted indoles from recombinant host products

In some instances, it may be preferable, for reasons of either cost or product quality, to utilize recombinant host pathways to accomplish the first part of a substituted indole synthesis and complete the remaining steps synthetically. The tryptamine, as obtained from the recombinant organism, is of a particular grade such that methylations with robust methylating agents selectively leads to mono- or di-methylation. One of ordinary skill in the art would appreciate this as improvement when a primary amine subjected to robust methylating agents as a mixture of alkylation products are not obtained, while obviating the need for tedious chromatography.

One example would be the production of tryptamine via fermentation of a recombinant host organism, followed by N,N-methylation via methylation chemistry to yield DMT. In one embodiment, the reaction of tryptamine would proceed with a 30-fold molar excess of dimethyl carbonate (DMC) under an inert atmosphere, utilizing a Y-type zeolite catalyst (see Fig. 25. This reaction is carried out at 190°C for 6 hours in a pressurized reactor vessel or autoclave; another embodiment would utilize a microwave oven for 15-60 minutes. The DMT product is recovered from the volatile DMC reactant via distillation. Another embodiment of the combined biosynthetic and chemical synthesis route is the production of tryptamine via recombinant host organism, followed by its reaction with DMC in the presence of the catalyst: 1,8-Diazabicyclo[5.4.0]undec-7-one (DBU). This catalyst can be used in a thermally heated reactor system at 90 °C for 6-24 hours or used in a pressurized microwave reactor system for less than one hour. Another embodiment of the combined biosynthetic and chemical synthesis route is the production of tryptamine via recombinant host organism, followed by its methylation to DMT using dimethylsulfoxide (DMSO). The catalysts for this system is acetic acid, and the reaction is carried out in a thermally heated reactor at 150 °C for 6-15 hours.

References

Bishop et al., 2006, J Virol. 80:8225-8235.

Blei et al., 2018, Chembiochem. 19:2160-2166.

Boyce et al., 2019, Fungal Ecol. 41:147-164.

Buller et al., 2015, Proc Natl Acad Sci U S A. 112:14599-604.

Byeon Y and Back K, 2016, Appl Microbiol Biotechnol. 100:6683-6691.

Fan, Aili, and Shu-Ming Li, 2014, Tetrahedron Letters 55: 5199-5202.

Francis et al., 2017, Chembiochem. 18:382-386.

Fricke et al., 2017, Angewandte Chemie International Edition 56:12352-12355.

Fujiwara et al., 2010, J Biol Chem. 285:11308-13.

Germann et al., 2016, Biotechnol J. 11:717-24.

Green and Sambrook, 2012, Molecular Cloning: A Laboratory Manual (Fourth Edition), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Hemmi et al., 2001, Biochemistry 40: 3512-3524.

Horita, A., and L. J. Weber, 1961, Proceedings of the Society for Experimental Biology and Medicine 106: 32-34.

Huang et al., 2018, AJSH 40:1-7.

Kong et al., 2016, Sci Rep. 6:20273.

Lenz, C., et al., 2019, Angewandte Chemie (International ed. in English) 59.4 (2019): 1450-1454.

Levine, WG, 1967, Nature 215: 1292-1293.

Lima da Cruz et al., 2018, Frontiers in molecular neuroscience 11: 312.

Nemoto et al., 1995, European Journal of Biochemistry, 233:1-8.

Keshavareddy et al., 2018, Int. J. Curr. Microbiol. App. Sci. 7:2656-2668.

Ruohonen, et al., 1997, *Yeast* 13: 337-351.

Winn et al., 2018, Angew Chem Int Ed Engl. 57:6830-6833.

Zehner, et al., 2005, Chemistry & biology 12: 445-452.

Zou et al., 2013, Angew Chem Int Ed Engl. 52:12951-5.

US Patent No. 10,435,727.

US Patent No. 10,671,632.

US Patent Application 16/553103.

US Patent Application 16/553120.

US Patent Application 17/068636.
US Provisional Patent Application 63/035692.
US Provisional Patent Application 63/053539.
US Provisional Patent Application 63/141486.
US Provisional Patent Application 63/164126.
US Patent Application Publication 2020/0063170.
US Patent Application Publication 2020/0063171.
U.S. Patent Application Publication 2021/0147888.
US Patent Application Publication 20210108238A1.
Finnish Patent Application Publication FI20185254A1.
http://parts.igem.org/Yeast.

In view of the above, it will be seen that several objectives of the invention are achieved and other advantages attained.

As various changes could be made in the above methods and compositions without departing from the scope of the invention, it is intended that all matter contained in the above description and shown in the accompanying drawings shall be interpreted as illustrative and not in a limiting sense.

All references cited in this specification, including but not limited to patent publications and non-patent literature, and references cited therein, are hereby incorporated by reference. The discussion of the references herein is intended merely to summarize the assertions made by the authors and no admission is made that any reference constitutes prior art. Applicants reserve the right to challenge the accuracy and pertinence of the cited references.

As used herein, in particular embodiments, the terms "about" or "approximately" when preceding a numerical value indicates the value plus or minus a range of 10%. Where a range of values is provided, it is understood that each intervening value, to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between the upper and lower limit of that range and any other stated or intervening value in that stated range is encompassed within the disclosure. That the upper and lower limits of these smaller ranges can independently be included in the smaller ranges is also encompassed within the disclosure, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the disclosure.

The indefinite articles "a" and "an," as used herein in the specification and in the embodiments, unless clearly indicated to the contrary, should be understood to mean "at least one."

The phrase "and/or," as used herein in the specification and in the embodiments, should be understood to mean "either or both" of the elements so conjoined, i.e., elements that are conjunctively present in some cases and disjunctively present in other cases. Multiple elements listed with "and/or" should be construed in the same fashion, i.e., "one or more" of the elements so conjoined. Other elements can optionally be present other than the elements specifically identified by the "and/or" clause, whether related or unrelated to those elements specifically identified. Thus, as a non-limiting example, a reference to "A and/or B", when used in conjunction with open-ended language such as "comprising" can refer, in one embodiment, to A only (optionally including elements other than B); in another embodiment, to B only (optionally including elements other than A); in yet another embodiment, to both A and B (optionally including other elements); etc.

As used herein in the specification and in the embodiments, "or" should be understood to have the same meaning as "and/or" as defined above. For example, when separating items in a list, "or" or "and/or" shall be interpreted as being inclusive, i.e., the inclusion of at least one, but also including more than one, of a number or list of elements, and, optionally, additional unlisted items. Only terms clearly indicated to the contrary, such as "only one of" or "exactly one of," or, when used in the embodiments, "consisting of," will refer to the inclusion of exactly one element of a number or list of elements. In general, the term "or" as used herein shall only be interpreted as indicating exclusive alternatives (i.e. "one or the other but not both") when preceded by terms of exclusivity, such as "either," "one of," "only one of," or "exactly one of." "Consisting essentially of," when used in the embodiments, shall have its ordinary meaning as used in the field of patent law.

As used herein in the specification and in the embodiments, the phrase "at least one," in reference to a list of one or more elements, should be understood to mean at least one element selected from any one or more of the elements in the list of elements, but not necessarily including at least one of each and every element specifically listed within the list of elements and not excluding any combinations of elements in the list of elements. This definition also allows that elements can optionally be present other than the elements specifically identified within the list of elements to which the phrase "at least one" refers, whether related or unrelated to those elements specifically identified. Thus, as a non-limiting example, "at least one of A and B" (or,

equivalently, "at least one of A or B," or, equivalently "at least one of A and/or B") can refer, in one embodiment, to at least one, optionally including more than one, A, with no B present (and optionally including elements other than B); in another embodiment, to at least one, optionally including more than one, B, with no A present (and optionally including elements other than A); in yet another embodiment, to at least one, optionally including more than one, A, and at least one, optionally including more than one, B (and optionally including other elements); etc.

SEQUENCE LISTING

Seq. ID No: 1
>6xHIS
CACCACCACCACCATCAC

Seq. ID No: 2
>AADC_1
ATGGAAACGGTGAATAAATCTTGTTGCTCCTTAGCGACACCGCATCTTCCTCTTATGTCACCGCACCTACTACACTC
TCACAGGGACTCTCCCGTTCTTAAAATAATCCACTCTATTGTGCTGACTGTGCAGAATAACCATTCTTGCTTGCAGG
CGTTTCTTGGGATTGGGACAGATAATGTAATCCTAGTAAAGACTAATGATAGGGGTAAAATGATACCAGAAGACCTG
GACCATAAGATCCAGAAAGTCAAGTCCGAAGGTTCAGTTCCATTTCTGGTTTCAACTACCTGTGGGACAACCGTGTT
CGGCGCGTTTGATCCATTAGAGGGCATTGCAGATGTTTGCGAACGTCATTCCCTGTGGCTGCATGTTGACGCTGCAT
GGGGAGGTTCAGCCCTTTTAAGCAGCCGTCACAGACATTTGCTTAAAGGCATCGAGAGGGCTGATAGTGTGACATGG
AATCCTCATAAACTATTGGGTGTTGGACTGCAGTGTTCCGCCTTTTTGCTGCGTGATACGACTCAGTTACTAGAGAG
GTGTCACGCTGCAAACGCAACGTATTTATTCCAGACCGATAAATTCTATAACCTGCAGTATGATACTGGTGATAAAA
GCATTCAATGTGGAAGAAGGGTAGATTGCTTAAAGCTGTGGCTAATGTGGAAAGCCTTGGGATCGAAGGGTTTGGAA
ACCAGAGTCGATAGAGTATTAGACCACACTAGGTACCTTGTAGAGGAAATGAAGAGGAGGGAGGGATTTAGACTGAT
AATGGAGCCAGAATTTGTCAATTTATGCTTTTGGTACGTGCCTCCCTCTCTTAGGAACAAAGAGAATAGTCCAGATT
TTTGGACCAGGTTGGGTAAAGTCGCCCCAGTCATAAAAGAGCGTATGATGAAGAAAGGGTCAATGATGGTGGGCTAC
CAGCCTCACGGGAACATGGTTAACTTTTTCAGGCAGATAGTCGTGAATCCAGAAGTCACCAAAGAGGACCTAGATTT
CTTTCTTGATGAGATAGAGCGTTTAGCAGAGGACCTT

Seq. ID No: 3
>AADC_2
ATGTGGGGTTGCGGGAACGGTGATTGTATCCACGTTTTACTGCTTATCAGCCATACGAGTCCACCACCCCTTAGTCC
GCATCTGTTGCACTCCCACAGGGACTCCCCGGTTCTGAAAATCATCCACAGTATCGTCCTTACAGTCCAGAACAATC
ATTCCTGCCTACAGGGCCATGTGCCATTTTATGTTAGTGCCACTGCGGGAACAACGGTGTATGGAGCTTTCGATCCT
TTCGTCAAGATTGCGGATATATGCCAGAAACACGGACTTTGGATGCTGCACGTAGACGCGGCTTGGGGTGGGGGTTT
GCTTATGTCTCGAAAACACAGGCACAAGATGAACGGGATAGAAAGGGCGGACAGCGTGACTTGGAACCCACATAAGA
TGATGGGAGTCCTTTTGCAATGTAGTGCAATCCTTCTGAAGGAAAAGGGTATACTACAGGGTTGTAATCAAATGTGC
GCAGGGTACTTATTTCAGCAGGACAAGCAATATGATATTTCTTACGATACCGGCGACAAAGCCATCCAGTGTGGCAG
ACACGTTGATATTTTCAAATTTTGGCTGATGTGGAAAGCAAAAGGGACAGTTGGCTTCGAGCAACAAATTAACAAGT
GCTTAGAGTTGTCCGAGTACCTATACTCTAAGATTTGCAACCGTGAAGATTTTGAGATGGTTTTTAAGGGAGAAGTG
TCTTTACACCGTTTGGAAGAGGGAATTAAG

Seq. ID No: 4
>AADC_3
ATGGAAACAGTGAACAAATCTTGTTGTTCATTAGCAACGCCGCACCTTCCTCTTATGAGTCCACATTTGCTACACTC
TCATAGAGACTCCCCAGTTCTTAAAATTATACATTCTATCGTCTTGACAGTGCAAAACAACCATTCCTGCTTACAAG
GTCACGTCCCGTTCTATGTGAGCGCCACGGCGGGCACCACAGTTTACGGAGCCTTTGATCCATTCGTTAAGATCGCT
GACATATGCCAGAAGCACGGCTTGTGGATGCATGTAGATGCAGCATGGGGAGGCGGTCTTTTACTGAGCAAAAAGCA
CAGGACGAAATTAAGTGGGATCGAGAGGGCAAATAGTGTGACTTGGAATCCTCACAAGATGATGGGGGTGCCACTTG
TAAAAGATAACATGGACCTTCTGAAGAGATGCCACTCCGCCGAAGCCAGTTACTTGTTTCAGCAGGATAAGTTCTAT
GACGTAAGATACGACACGGGAGACAAATCCATCCAGTGTAGTAGGCGTGCGGATGCATTCAAGTTTTGGATGATGTG
GAAGGCGCTGGGGACACTTGGGCTAGAGGAAAGGGTTAATCGTGCCCTGGCCCTGTCTAAGTACCTAGCCAAGGAGA
TCAAGAAAAGAGACGGATTTGAACTGATTTGGGAGCCAGAATATGCTAACATCTGTTTTTGGTACATCCCACCGAGT
TTACGTAAGATGGAAAATAAAGGTCCTGAGTACTGGCAAAGGTTAAACCAGGTCGCACCAACAATAAAAGAGCGTAT
GATGAAGAAGGGTTCCATGATGATTGGATATCAGCCCCACCGTGACAAGGTGAACTTTTTCCGTCACATCATTATTT
CACCGCAGGTCAGTCGTGAGGATATGGACTTTGTCTTAGATGAGATTGACTTACTTGGCCGTGACTTA

Seq. ID No: 5

>AADC_4
ATGAATGCCTCAGAATTTCGTAGGAGGGGGAAGGAGATGGTCGATTACGTAGCAAATTACATGGAAGGGATTGAGGG
GCGTCAGGTATACCCAGACGTGGAACCTGGGTATTTAAGGCCGCTTATTCCAGCTGCAGCCCCACAGGAGCCCGATA
CCTTCGAAGACATCATAAATGACGTTGAGAAAATCATAATGCCTGGCGTCACTCATTGGCATTCTCCATACTTTTTC
GCCTACTTTCCCACTGCTTCTTCCTACCCAGCCATGCTTGCAGACATGTTATGTGGCGCAATAGGATGTATTGGGTT
CTCCTGGGCGGCATCCCCCGCCTGCACGGAGCTAGAGACAGTTATGATGGATTGGTTGGGTAAAATGCTTGAGCTGC
CCAAAGCGTTCTTGAACGAGAAGGCAGGAGAAGGGGGTGGCGTGATCCAAGGCTCTGCTAGTGAAGCCACCCTAGTC
GCCTTGCTGGCAGCGAGAACTAAGGTGATACACAGACTTCAAGCCGCCTCTCCCGAGTTAACTCAGGCTGCAATTAT
GGAGAAATTGGTAGCCTATTCTAGCGACCAGGCTCACTCCAGCGTAGAACGTGCCGGTCTGATTGGGGGCGTGAAGT
TGAAAGCTATTCCGAGCGATGGGAATTTTGCCATGAGAGCGTCCGCGCTTCAAGAGGCCTTGGAACGTGACAAGGCT
GCCGGATTAATCCCGTTCTTTATGGTCGCTACGCTTGGTACTACCACGTGTTGTTCTTTTGATAACCTACTGGAAGT
GGGTCCTATCTGCAATAAGGAGGACATTTGGCTTCATGTAGACGCTGCCTATGCAGGTTCTGCTTTCATTTGTCCCG
AATTTAGGCACCTGCTTAATGGCGTGGAGTTCGCGGATAGTTTTAATTTTAACCCTCATAAATGGTTGCTGGTGAAT
TTTGACTGTAGCGCCATGTGGGTCAAGAAAAGAACGGATCTAACTGGAGCCTTTAGACTTGACCCCACGTACCTTAA
ACACAGCCATCAGGATTCCGGGCTAATAACCGATTATCGTCACTGGCAAATCCCGTTGGGAAGACGTTTTCGTAGCC
TGAAGATGTGGTTTGTGTTTAGAATGTATGGGGTGAAAGGGCTACAGGCGTACATACGTAAGCATGTCCAACTTTCT
CACGAGTTTGAGTCTCTTGTACGTCAAGACCCGAGATTTGAGATTTGCGTTGAGGTAATACTAGGTCTAGTCTGTTT
CAGGCTTAAAGGCTCAAACAAGGTCAACGAAGCGTTATTGCAAAGAATAAATAGTGCTAAGAAGATCCACCTTGTTC
CCTGCCATCTACGTGATAAGTTTGTCTTACGTTTTGCGATATGTAGTAGGACGGTAGAAAGCGCACACGTTCAGCGT
GCGTGGGAGCACATAAAGGAGTTAGCCGCTGATGTGTTGAGAGCTGAACGTGAG

Seq. ID No: 6
>AADC_5
ATGGCCAGTGGGTATCCAGGGGCGGGAGCCCAACAGCCGCCTGCGGCTCCTGCCAGTGGATCAGGTTCCCCCGTCTC
CATGCCGTACTACGCTTCCGAATTAGCAAGGGCGAAGCACGAGGATGACATGAAAATGCCGGAACATGGGATAGAAC
CTAGGCACTGTCTTAGACGTATCGAGGATTACCATCTATTGGACTTCAGTGAAAGGTTAAACACGAGTTCATACGTC
AATGTAGTCTTTGAACCAGAAGAGGAGACGGTTGCTAACATGGGGTTGAAAGTCAATTTAGCTGATCAAACAGTATA
TCCCGAAAGTTTCCGTATGCACAATGACACAGTCAATATGATCGCTAAGTTATGGAATTGCCCTAAGCCTGCCGACT
TTGATGAATATGGATGTTATGCCGGAGCAGGCACCGTTGGTTCCACCGAAGCGTGTCTGCTTGGAGGATTGGCTCTT
AAATTCAGATGGCGTAAATGGTACGCCGCCAAACACGGAATGGATCAGAACAAAGTAAGAGGTGTGTACCCGAACCT
TGTTATCACTACGATGTTTCAAGCCGCTTGGGAAAAACTTTTCAAATATATGGACATTGAACCCAGATTCGTGACCC
CATCCTGGAAGACATTCACTATGGACCCATCCGGCCTTGAAAAAGTGGTAGACGACAAAACTATAGGTGTGGTGTGT
ATTATGGGAAATCACTATGGTGGTCAATACGATCCAGTGTGGGAAGTGAACGACGTGCTTGAGAAGATTAACAAGGA
AAAGGGACTTCAGGTAGGTATCCATGTGGACGGAGCGTCTGGCGGATTCATAGCCCCCTTCCAAGAGGGATTACCGG
CATGGGATTTTAGACTAAAGAATGTGTTAAGTATAAGCGCGTCTGGACACAAATTCGGAAACTCCTGTTGTGGAACA
GGATGGATTATTTGGAGAGAAAGGAAAGGACTTTCTGACACTGTCGCCATCAATGTATCTTACCTTGGAGGGTCCGC
AGACAGCTATACGTTAAACTTCAGCAGACCGGCGCAAGGTGTATATGTCCAG

Seq. ID No: 7
>AADC_6
ATGGCGGGCAGCATCGACGTGATGCTTAGGGACCTGAATCTTGACACCGTTGAATCTCTTCCAGCAGACTTTGATCC
GACCACAGTAATAAATGATATACTACCACCAGTTGTAAACGACGCCAAACTGGTTAATGGCATTGAGAAACCCAGAG
AAATAGTCCTGGGCAGAAATGTACATACTTCCAGTTTGGAGGTAACAGAGCCCGACGCCGATGATGAGGTTACCGGG
GAGAGAGAGGCGTACATGGCGTCAGTACTAGCTCGTTATCGTAAGTCTCTGCTAGAGCGTACAAAACACCATTTAGG
TTACCCATACAATTTGGATTTTGACTATGGGGCTTTGTCCCAGCTTCAGCACTTTTCAATCAATAATCTTGGAGATC
CATTTATCGAGAGCAACTATGGCGTACATAGCCGTCAGTTTGAGGTAGGGGTCTTGGACTGGTTTGCAAGACTATGG
GAGCTGGAACGTAACGAATATTGGGGGTATATTACGAATTGTGGGACAGAGGGTAACTTGCATGGCATATTAGTCGG
CAGGGAGGTCTTTCCCGACGGTATCTTGTATGCAAGCAGTGAATCCCACTACAGTATATTCAAAGCCGCGAGGATGT
ATAGAATGGATTGCGAAAAGGTCAATACGCTTATCAGTGGAGAAATCGACTGTGAGGATTTCAAGGCAAAGTTATCT
CTGCACAAGGATAAGCCCGCGATCATAAACGTCAACATTGGGACAACAGTAAAGGGCGCGGTGGATGATCTGGATCT
TGTTATTAAAACCCTGGAGGAAAGTGGCTTTAGTCACGACCGTTTCTACATCCATTGTGATGGGGCGCTTTTTGGGC

TAATGATGCCTTTCGTAAAGCTGGCACCCAAGGTAAGTTTCAAGAAACCCATTGGTAGTGTGAGTGTGTCTGGGCAT
AAGTTTGTCGGCTGTCCTATGCCGTGTGGTGTACAAATTACCAGACTTGAGCATATCAATGCTCTTAGCCGTAATGT
AGAGTACCTGGCGTCTAGGGATGCTACCATTATGGGCTCAAGAAACGGTCATGCACCCTTGTTTCTTTGGTATACTT
TAAATCGTAAAGGTTACAGGGGATTCCAGAAAGAGGTACAAAAGTGTCTTCGTAACGCTCATTATCTTAAAGGTAGG
CTAACTGAAGCAGGGATTGGCGCGATGCTAAACGAGCTAAGTAGCACTGTGGTTTTTGAACGTCCTCAGGATGAAGA
GTTCACTCGTAAATGGCAATTGGCGTGTCAAGGTAATATTGCGCATGTAGTCGTCATGCCGAACATTAATATTGATA
AGTTGGACCATTTTGTCAATGAATTAGTAGAGAGAAGGGCTGTATGGTACGAGAATGGGAAGTTGAAGTCTCCATGT
GTTGCCTCCGAAATAGGCAACTCTAACTGCTTATGTGCGTTGCATAAG

Seq. ID No: 8
>AADC_7
ATGGTGAGGGCCGTTGAAAAGCCAGTACAAGCCATAGTAAATGCCGCGTTTAGAGGGAAAGATGCATATCATGTTTT
CAGAACAACAGTTCTAGCGGCCGTGCTGTTAAGACTGTGGAGACACTTGAGACGTGTGATGGCTCATGAAGGTTTGA
AAGCATACTTTATGTCCTTAGTCGCACCGCATCTGAAGAAACTTCCGTACGTCCAAAACAAGTTGAAGAAGGAGATG
GATAAAACTATGACGAAAATGAGAAATACCTTTCGTAAGGAGGTGACAGATCCAAGGACATCATTACCTGTTGATGG
GATTCCTGAGAAACAGATCTTGGAACTGATTCAACATCGTAAAGAGTTAGATACTAAGGAATGGACAAAGGGTATGA
CCACTGGCTGTGTTTACCACGGTGGTCAGGACCATTATGACTTTGTAGGACAGATCTTTGCATCATGGGGGTTCGCT
AATCCATTGCACCCCACAACCTTTCCAAGCCTACGTCAGATGGATAGCGAGGTCGTACAGATGGTTATCAACATGTA
CCACGGGGATTCAGAGTGCTGCGGGGCGTTCACTACTGGAGGAACGGAGAGCATTCTGATGGCGATGAAGGCCTATC
GTGATTGGGGTAAGGCAGAAAAGGGTATTACCGATCCCAATATCGTTATCTGTAACACCGCCCATGCTGCTTTTGAT
AAGGCCGGCAAGTACTTTAACATCTTCGTAAAACACGCGAGAACGAACTCTGAAATGGAGATTGATTTGGGCCATCT
ACGTAGCCTTATCGACTCCAATACCGTGGCCATCGTAGGCTCAGCCTGCCAATTTTCTCATGGAACCGTGGACCCGA
TCCAGGAAATGGCAAAGATTGCCATGAAGCGTAGAGTTGGCCTACACGTTGATTGTTGCCTAGGCGGTTTTCTAGTT
CCGTTCATGGAAAAAGCTGGATTCCAACTTCCGCCCTTTGACTTCAGAGTCAAGGGGGTAACGTCCATATCCTGTGA
TCCCCACAAATATGGGTTTGCCCCAAAGGGCTCCTCTGTTGTAATGTTTAGTAACCGTCACCTGAGGCACTACATGT
ATTGTTTCTTAACAGAATGGAGCGGAGGAATATACGCGACTGCAACGATGACTGGTTCCAGAGCAGGTGGCCCGGTG
GCCGCCACATGGGCGAGCATGTGCAAATTTGGGGAGAAAGGCTATATTGAGACAACTAAACAGATCGTCGGGGCCAC
TAAGAAAATTGCGGCTGGTATAGCAGAAATCGAAGGTTTACGTGTGGTTGGCAGACCGGATGTATGTGTCGTGGCCT
TCACCTGTACTGAGGGTTCCGGAATGAATTGCTACGCCGTTGGGGACTGTATGCATCAGGACTTTCATTGGGAGTTG
CAATCTTGTCAGAATCCGGCTTGTGTTCACTTGGCTTTAACCCTGCCGACCAGTAGAAATGCTGATAAATTCGTAGC
TGATTTGCGTCAGGCTGTGGAGGCGGTGAGGAGCGATAAGGACGGCAAATTCGCATCCACGGCGGGCATGTATGGGA
CCGCAGCTTCTTTGCCAGCCGCGTTTTTCGAAGACGGCGCAGCAGCATACCTTGATGCGATGTGTGAAGCCATCCCA
GCGGGAGACGCTCTACTGCCGGAGGAACCTGCCACGAAGGAATCCCCTGCAGCTGCCGGGGCACCGGCTCAAGCAAC
AGGGGGTGCT

Seq. ID No: 9
>AADC_8
ATGAATGCGAGCGAGTTTAGGCGTCGTGGGAAGGAAATGGTTGACTACGTAATCACGAATTATCTGGAACAGATAGA
GTTACGTCAGGTCTACCCTTCCGTCGAGCCGGGGTATCTTAGACCTATGATTCCTGATTCCGCGCCGGAAGAAGGTG
AGACGTATGAAGATATCATGAAAGACATCGAGAGAGTCATAATGCCGGGAGTTACTCATTGGAATAGCCCGTACTTC
TTTGCGTACTTCCCTGCGGCAACAAGTTACCCTGCTATGTTGGCCGACATGTTATGCGGCAGCCTGGGGTGCATCGG
GTTCTCATGGGCCGCTTCACCCGCCTGCACAGAGTTGGAAACCGTAATGCTTGACTGGTTGGGCAAAACAATCGGCC
TACCAGAACAGTTTCTGGCTGGTACCAACGGAGAGGGTGGCGGTGTAATCCAAGGTACTGCCAGCGAAGCGACCTTG
ATGGCACTATTGGCGGCCCGTACTAAGGTTACAAGAAGGTTACAAGCTGAGAACCCTGACCTTTCTGAAGCAGAGAT
CATTAGCCGTATGGTGGCATATTCTAGCGACCAGGCCCATTCCAGCGTGGAGAGAGCTGGGCTGATCAGTGGCGTTA
GAATGAAAAAGATCCCATCTGATGAAAACTTCACTGCTAGAGGTGAAGCACTAAAGAAAGCACTTGAGGAGGACAAA
GCTGAAGGTTTCATACCAGTCTTTCTTTGCGCGACACTGGGCACAACAACTTCTTGCGCGTTCGACAATCTTATGGA
ACTGGGTCCGATATGTAATGCGGAGAACATGTGGTTGCATATCGACGCCGCCTATGCTGGGAGTGCTTTCATCTGCC
CGGAGAACAGATATTTAATGAAGGGTGTCGAGTTCGCGGACTCCTTCAATTTCAACCCACATAAGTGGTTGTTAGTC
AACTTCGACTGCAGTGCTTTCTGGGTAAAGAAAAGGTCCGATCTTATATGTGCATTTAAGATAGACCCCGTCTACTT
ACAACACGACCAGCAGGAGAGCGGTTTAGTAACGGATTACAGACACTGGCAGATTCCTCTAGGAAGAAGGTTCCGTA

GTCTGAAACTGTGGTTTGTTCTGAGGATGTATGGCGTAAAGGGACTTCAAGCTCATATACGTAAGCATATTAGACTA
GCACAGGAATTTCACGAATTTGTTAAGAACGATGATCGTTTCGAGATATGTGCACCCGTCATATTGGGTCTGGTGTG
TTTTTGCTTGAAGGGGTCAAATACGCTAAATAAGTCCTTGCTTCAGAAAATTAACACTTTGAAGAAAATACATCTTG
TACCTTCTTGTCTAGGCGATAAATTCATCTTGAGATTCGCGGTTTGCGCTAGAACACTAGAGTCAAACCACATTGTG
TTCGCCTGGAAGCACATTGAGGAATTGGCGACTGAGGTTTTAAAGGAGGGAGAGAAACAA

Seq. ID No: 10
>AADC_9
ATGCAGAACTGCAATCAGATGCACGCCAGTTATCTTTTCCAACAAGACAAGCATTATGATCTGTCCTATGATACGGG
TGATAAGGCGCTACAGTGCGGGAGGCATGTGGATATCTTCAAACTTTGGTTAATGTGGCGTGCCAAGGGCACGACGG
GTTTTGAGGTCCAGATAGACAAATGCTTGGAACTTGCTGAGTATTTGTATGACAAAATAAAGAACCGTGAGGGTTAT
GAGCTAGTGATAGAGGGGAAGCCCCAACATACTAACGTCTGTTTCTGGTACATACCCCCATCACTTAGACACATGGA
GGATAATGAGGAAAGGATGGCTCGTTTGGTTAAAGTCGCTCCAGTAATAAAGGCTCGTATGATGGAATACGGAACTA
CAATGGTTAGCTACCAACCGCTGGGTGACAAGGTGAATTTCTTTCGTATGGTTATATCCAATCCAGCCGCAACCCAT
CAAGATATAGATTTTCTAATTGATGAAATTGAGCGTTTAGGTCAAGACCTG

Seq. ID No: 11
>AADC_10
ATGTGGGGATGCGGTAACGGAGATTGCATACACGTTTTGTTATTGATCTCCCACACGTCTCCACCTCCCCTGTCACC
CCATCTACTACATTCTCACCGTGATAGCCCGGTATTGAAGATAATCCACAGTATCGTACTGACGGTGCAGAATAACC
ATTCTTGCTTACAGGGGCATGTCCCTTTCTACGTGTCCGCAACTGCAGGGACAACTGTCTACGGCGCATTTGATCCC
TTCGTTAAAATCGCTGACATTTGTCAGAAGCACGGGTTATGGATGCACGTAGACGCGGCGTGGGGAGGAGGCTTGCT
TCTTAGCAAAAAGCACAGGACTAAACTGTCAGGTATAGAGCGTGCCAATTCCGTAACATGGAACCCACATAAGATGA
TGGGAGTCCCTCTTTTCCAGTGCTCCGCCTTTTTATTACGTGATACGACTCAACTGCTAGAGCGTTGTCATGCTGCG
AACGCAACATACCTTTTCCAAACCGATAAATTCTACAATCTACAATACGATACGGGTGACAAAAGCATCCAGTGCGG
GAGAAGAGTGGACTGTCTGAAACTTTGGTTAATGTGGAAGGCCTTGGGCTCCAAAGGTCTTGAGAGGAGAGTTGACA
GAGTCTTAGACCATACCAGATACCTGGTCGAGGAGATGAAAAACAGAGAGGGCTTTAGGCTAATTATGGAGCCCGAG
TTCGTGAATCTGTGCTTTTGGTACGTTCCCCCGAGCCTGCGTAATAAGGAGAACTCACCGGACTTTTGGACAAGGCT
GGGT

Seq. ID No: 12
>AADC_11
ATGGGGTCTTTGGGAACAAATCCGACGAGCTTTAGCGCCTTCCCTGATGACAAGGCCGCATTTGAACCATTAAACCC
AGAAGATGTTAGAGCTTATCTTCACAAGGCGGTCGATTTTATCTCAGACTACTACACGAACGTTGAAAGTATGCCTG
TTCTGCCGAACGTGAAGCCGGGATATCTGCAAGACGAGCTGACGGCTTCTCCACCCACTCACAGTGCCCCATTCGAC
GTTACAATGAAGGAACTAAGGACCAGCGTTGTACCAGGCATGACACACTGGGCGTCACCAAACTTCTTTGCATTTTT
CCCATCAACCAATTCAGCGGCGGCAATAGCCGGTGATCTAATAGCTTCTGCGATGAATACGGTAGGCTTCACATGGC
AGGCGAGTCCCGCAGCAACGGAGATGGAAGTTTTGGCCCTGGACTGGTTAGCACAGCTACTACATCTTCCGACGACT
TTTATGAATAGAACCTCCACTGGCAGGGGTACAGGCGGAGGAGTCATATTGGGTACGACGTCAGAGGCCATGTTAGT
TACGCTTGTCGCAGCTAGGGACGCTGCTCTGAGGAGAAGCGGCAGCGTGGGAGTGAGCGACATTCCTAGACTGGCAG
TATATGCTGCGGATCAAACACATTCCACCTTCTTTAAAGCGTGCCGTTTAGCAGGATTTGACCCAGCTAACATTAGG
AGTATCCCCACTGGACCTGAGACTAATTACGGCTTAGACCCCGCTAAACTATTGGAAGTGATGCAGGCTGACGCGGA
TGCGGGGCTAGTCCCGACTTACGTATGCGCCACCGTCGGTACTACATCTTCCAATGCTGTGGACCCGGTTGGCGCGG
TTGCCGATGTCGCCGCCATGTTTAATGCCTGGGTACATGTAGATGCCGCTTACGCCGGCAGCGCGTGTATATGTCCT
GAGTTCCGTCACCATCTTGACGGCGTAGAGAGAGTCGATAGCATCTCAATGTCCCCTCACAAATGGCTTTTAACCTG
CCTAGATTGCACTTGTCTGTATGTTAGGGATGCTCACAGGTTGAGTGACTCTCTGGAGACCAACCCCGAGTACTTGA
AGAACGACGTTACCGACAGTGGTGAGGTTACCGACCTAAAAGATATGCAGGTTGGGGTTGGTAGAAGGTTCCGTGGA
CTGAAGCTGTGGATGGTTATGAGAACCTATGGCACTGCGAAACTACAGGAGCACATTCGTTCCGACGTCGCAATGGC
CAAAATGTTCGAGGATAGCGTTAGAGCTGATAATCGTTTTGAGGTCGTCGTACCGAGGAACTTTGCCCTTGTATGCT
TCCGTATTAAGGCTAGAGGAGATATGACTGAGGAAGACGCGGACGAAGTAAACAGGTTACTGATGGAGAATTTAAAC

AAGACGGGTAAAGCATACTTAGCTCACACTGTGGTCGGGGACCGTTTTGTTTTACGTTTTGCGGTTGGGAGTAGTCT
GCAAGAAGAGAGACATGTGAGGTCAGCCTGGGATTTGATCAAGAAGACGACTTCATCTATTATGGAT

Seq. ID No: 13
>AADC_12
ATGGACCCTCTAAAAGCCGTTGAAATGGTTGACGAAAATACTATCTGCGTAGCAGCCATCTTGGGCTCCACTCTGAC
CGGGGAGTTTGAGAACGTAAAATTATTAAATGAGCTATTAACTAAAAAGAACAAGGATACAGGGTGGGACACGCCTA
TACACGTTGACGCAGCTTCAGGCGGCTTCATAGCTCCCTTTCTATACCCGGACCTAGAATGGGATTTTCGTCTTCCC
CTTGTGAAGAGTATCAACGTTTCTGGACACAAATATGGCTTAGTCTACCCCGGCGTGGGATGGGTAGTATGGCGTAG
CAAAGGGGATCTACCTGACGAACTGATATTTCATATCAATTATCTGGGAAGCGATCAGCCCACTTTTACCTTGAATT
TCAGTAAAGGTAACAATATCAGCACGCACGCTTACAAGAAGCCCAAGTGCCGTTTCTACTTCCTAACCAATCAAATC
TTATCTAGGCGTTATTTATTACACCGTCTGAAAAGTGGAACTTCTTTGAATAGTCTTCTA

Seq. ID No: 14
>AADC_13
ATGCAGCCAGGGTATTTATCTCGTATGCTGCCTGATTCAGCTCCCAATCACCCCGAATCTCTTGAGGATATTTTCAA
TGACATCTCAGCAAAGATACTTCCCGGCGTCACACATTGGCAGAGCCCGAATTACTTCGCATACTTCCCCTCTAATT
CCAGTATTGCTGGGTTTCTGGGTGAAATGTTGTCCGCTGGTCTGAACATAGTAGGTTTTTCCTGGATAACTAGCCCA
GCGGCTACAGAATTAGAAATGATCGTTCTGGACTGGCTTGCGAAATTACTAAAACTTCCTGATGACTTTCTTTCCGG
TGGCACGTCT

Seq. ID No: 15
>AADC_14
ATGGTAGTAGACTACAAGGACTGGCAGATCCCATTAGGACGTCGTTTTCGTAGCTTAAAATTATGGATGGTACTGAG
ACTTTACGGAATTGAAAATTTACAATGCTACATCAGGAATCACATCAAGTTAGCCCAGCAGTTCGAGGTTCTAGTAG
CGCAAGACCTAAGGTTCGAGATCGTATCTCCAAGAATTTTTAGCTTGGTTTGCTTCAGGTTACTACCTTCCCAAAAC
TGCAAGGATCACGGGAATGAGCTAAACCACCACCTGTTAGATACTGTCAACAGCACGGGGAAAGTCTTCCTTAGTCA
TACTGTTCTATCAGGGAAATATATCTTGAGATTCGCTGTTGGCGCTCCCCTTACGGAGGAAAGGCACGTAACGGCGG
CATGGAAGGTATTACAGGATGAGGCGTCAGCGTTGTTACAGTCACTG

Seq. ID No: 16
>AADC_15
ATGGGCTCTCTTGATATAAAGCAGGAATCAAGCCCTCTAATGACAAATCCGTTAGATAGCGAGGAGTTCAGGCGTCA
AGGATACATGGTGATAGACTTTCTAGCCGAGTATTATAAGAACATACAGAAATTTCCGGTTCGTTCTCAAGTGGAAC
CTGGGTATTTAAGGAAAAGGCTTCCAGAATCTGCCCCTTACGAGCCGGAAAGTATAGAGCGTATTCTTAAAGACGTA
CATGACGACATAGTTCCTGGTCTTACTCATTGGCAATCACCGAACTATTACGCATATTTTCCGTCCAGCGGTAGCAC
AGCGGGTCTATTGGGTGAAACCCTGGCGGCCGGGTTCAATGTGGTCGGCTTCAACTGGATCTCCAGTCCTGCATCCA
CTGAACTAGAGAGTATCGTAATGGACTGGTTAGCTGAAATGTTGAACCTACCAAAGTCCTTTACATTTAGCGGAGAC
GGTGGCGGGGTGATGATGGGTACCACGTGTGAGGCAATTCTGACAACTATTACAGCTGCACGTGACAGGATACTGGA
CAGGATCGGACGTGAGCACATTAATAAGCTGGTCGTTTATGGCTCTGACCAAACGCACTGTTCTTTCTTTAAGTCAG
CTAAGATCGCTGGTATTCTACCGAATAATTTTAGACAAGTTAAAACTAGCAGAGTCAACGCGTTTAGTATGAGACCC
GACGCGCTTCGTGCCGCGATACAGGCGGACGCAGATGCAGGCCTGGTCCCGTTTTTCTTATGTACCACTGTTGGTAC
GACCTCAACAGCGGCTGTAGACCCGGTAGCGTTACTTTGCGAAGTAACAAAAGACTACGGGATGTGGGTGCACATTG
ATGCAGCGTACGCGGGCAATGCCTGCATTTGTCCAGAGTTCCGTCATATGATCAATGGGGTTGAGAATGCCGACTCA
TTCTCCTTTAACGCGCATAAGTGGTTCTTGACTACTCTTGATTGCTGTTGCTTATGGGTTAAAGATCCAAGCTCACT
GGTAAGGTGCCTTAGCACTAATCCCGAATACCTAAAGAATAAGGCCACTGACACCCAACAAGTTGTTGATTATAAGG
ATTGGCAAATCACACTATCAAGGCGTTTCAGATCTTTA

Seq. ID No: 17
>AADC_16

ATGGACGGGCAAATGCTAAAGCCAATGGACGCTGAACAACTGAGGGAGTATGGTCACCAGATGGTGGACTTTGTAGC
GGACTACTACAAAACTATTGAGAGCTTCCCCGTGCTAAGCCAAGTTCAACCTGGATACTTACGTGAGCTAGTCCCTA
CAAACCCTCCAACGCACCCGGAGTCCCTTCAGGATGTTCTGGACGATGTTAAAGCGAAGATATTACAAGGCGTAACC
CATTGGCAAAGCCCGGGATATTTCGCCTATTTTCCGTGCAACTCTAGTACCGCGGGATTCTTAGGAGAGATGCTGTC
CGCCGGGATCAATATCGTAGGATTTTCCTGGATGTCCTGCCCTGCGGCTACTGAGCTTGAAGTAATCGTCCTTGACT
GGTTAAGCAAACTG

Seq. ID No: 18
>AADC_17
ATGGTGTTGAGGTTATACGGCCAGGAAGGTCTTCAAAGCTACATAAGAAACCATATTGCTTTAGCGAAACAGTTCGA
GGAGTTGGTCATTCAGGATTCCAGATTCGAGATGGTCACCCCACGTAGGTTTAGTCTTGTGTGCTTTCGTTTACTGC
CCAGAAGCACAGACGAGAACCAGGCGGATAAGCTAAATAGGGCACTGCTAGATGCGGTGAATCTTACGGGTGCGATA
TTTATCTCTCACACAGTCTTGTCAGGGATGTACGTGTTACGTCTTGCGGTGGGGGCACCGTTGACTGAAGAGAGGCA
TGTGATAGCCGCGTGGAAGGTCTTGCAAGAGAAGGCAACGGCATTACTGGAGGGTAATGTAGCGCAAGAACCTAACG
GTCACGCGCAATTATCCAATGGTGTCGTCGAGCTTGATGGTGCGCTAGATAATGGTGTCACCACTGAACAGCATGGA
CATGTGGACGATCTTCCTATTAAATCTAAA

Seq. ID No: 19
>AADC_18
ATGAGCCTTTCCCGTCACATAGACGCGGAGCGTTTAATCGAGCAAATCAAGGAACACCCACATAAGAAACATTCTTC
TGCGGAGTCCCGTAGGGGTTTGCACAGGTCTTGGAGACGTGATGAGGATGCTGCGGACTTACCAAAGTACACCTTAC
CCAAACATGGAATTAATTCAAAAGCCGCTTACCAACTGTTACACGATGAAACTGCTCTGGACGGAAATCCTTTGCTT
AACCTGGCGAGCTTTGTGCACACCTGGATGCCGGAAGATGCGGACAAGTTAATAATGGAGAATATAAACAAAAATAT
AGTGGACATGGATGAATACCCTGCCGCCTCCCTTATCCACAACAGATGTATATCAATGCTGGCGGATTTATGGAAAG
CCCCTAAAGAAGGTAAAGTAATCGGCACGGCAACGGCAGGATCATCTGAGGCTATTATGCTAGGCGGTCTGGCTTTG
AAGAAGAGATGGCAGGAAGCGAGAAAGGCTGCTGGTAAGGATTATTTCCATCCGAACATTGTGTTCGGGAGCAATGC
TCAGGTAGCCTTGGAAAAATTTGCGCGTTATTTCGATGTCGAGACGAGACTTGTGCCGGTCAAAGAGGAGAACGGCT
TTGTTATGAATCCTCATGATGCCATACCTTATATCGACGAGAATACGATAGGAGTGATTGTGATACTGGGAAGTACT
TACACGGGGCACTTCGAGGACGTCAAATTGATGAGTGATCTGTTAGATGACTTAGAGAAAAGAACTGGGCTAGACGT
AAAAATACACGTCGATGGCGCATCAGGCGCATTCATAGCACCTTTCGCTTACCCGCACCTGAAATGGAGCTTTGACG
TACCTAGGGTTGTATCAATTAACACAAGTGGCCACAAGTTCGGGTTAGTTTATGCGGGCTTAGGGTGGGTACTGTGG
CGTGACGAGTCATTTCTACATCGTGACTTGGTTTTTGAGTTGCATTACTTGGGCTCAACCGAGTATAGTTTCACTCT
AAATTTCTCCAAGCCAGCGGCTCCAGTCATAGCGCAAATGTTTAACTTTCTAAACTTGGGTTTTGAGGGATACAAGA
AGATTGCATACAAGGATATGAGAAACGCAAGGATGCTTAGCAGAGCCCTGGAGTCTACGACGTATTTTAAGGTATTC
AGTAATATTCATGTCCCTCGTAACTCAGACAGTGCTCACGTTTCCAGCAACAACAAAGACGACCCGGAGACTTACCA
TGCAGGCTTGCCCGTTGTCGCATTTAGGCTGTCTGACGAATTTAATCAGAACTACCCGAACGTTCGTCAAGTGTGGA
TTCAGACCCTGCTGCGTACGAAAGGTTGGATTGTGCCCAATTATAACGCACCCCTTGGGGCTGAGAATATAGAGATT
TTGAGAATCGTTGTCAGGGAAACACTTAGTGAGGATCTTATTGAAAGACTGATCGTTGACATTGTCGCCGTTACCGA
AAGTTTGACTACCGAGGAAGGTAACGTATTCGCCGGTATCACGGCGTCTGGCGCGGTGGTTAAGCCCGATCTGGACG
AAGCCCGTCCAGACAGTAGTAATTTTAACGATTCAGGTGACGGCGAAGCGCAGGGCCAGACGGGTTATTCACGTCAG
TGC

Seq. ID No: 20
>AADC_19
ATGGCTCTATCTAAACATGTTAATACGGATAAATTGATTCGTGATAGCAGAGATAAAAAGTCACCAAAGGAAAAGGC
CCATCACACGGCCACTGCTCATCAGGAGGCAACTTACTCCTACGGGGATCGTTATGTCACTAACCCCGTGCCTAAAT
ACAACATAGCGAGTAAGGGAATTAGTGCCGACGCCGCATATAGACTTATTCACGACGAACTAGCCCTAGACGGGTCT
ACTGTGCTTAACCTTGCCTCATTTGTTCACACTTGGATGCCGCCCCAAGGCGAACAGCTAGTACATGAAAATATTGC
GAAGAACCTTATTGATAGCGATGAGTATCCTGCCACGCAGATCATTCATACTAGGTGCGTCAGTATTTTGGCTGATT
TATGGCATGCACCGTCCGCTAAGCAGGCGGTAGGCACTGCAACTACGGGCTCCTCAGAGGCGATACAGTTAGGCGGT
CTGGCGATGAAAAAGATGTGGCAAGCGCGTATGAAAGCAGCTGGTAAAAACATCCACGAACCGGGTCCCAATATAGT

CATGGGAGCTAACGCACAAGTGGCCCTTGAAAAGTTCGCGAGATATTTTGATGTCGAATGTAGATTAGTCCCGGTAA
GTGTAGAATCAAAGTATCGTCTTGATCCCAAGAAAGCTATGGATTTTGTAGATGAAAATACGATCGGAATATTCATC
ATTCTGGGTTCAACCTACACTGGTCACTACGAGCCTGTGAAAGAGATGAGTGATCTGTTGGATGAGTACGAGAAACG
TACAGGAATCTACGTTCCCATACACGTTGACGGCGCGTCTGGAGGGTTCGTAGCGCCTTTTGTCCACCCCAAGATGG
TTTGGGACTTTAAATTGCCAAGAGTAGTAAGTATAAATACATCTGGGCACAAATTTGGCCTTTCATACGTCGGCGTT
GGGTGGGTGGTCTGGCGTGATAAAGCTCATCTGCCCAAGGACCTAATCTTTGAGTTGCATTACCTAGGCTCCGTCGA
GTACAGCTTCAGTCTGAATTTCTCTCGCCCCGCTGCACCTATCTTGGCCCAATACTTCAATTTAGTCCATCTTGGGT
TCGAAGGTTACAGGTCAGTCGGTTTAGCAGATATGAAAAATGCCAGAGAGCTTAGCCGTGCTTTAGAAAAGACTGGA
TATTACACTGTATTGTCCGATATCCATAGAGCAGTAGGCGCAAAGGACCCGCACGGGATTGACGACGCTGATATAGA
GGCCTACGAGCCCGGGCTGCCCGTTGTGGCCTTCCGTTTTAGCGATAACTTTAAAGAGAAACACCCTGAAATCCAAC
AGAAATGGATACAGACCCTGTTGAGGGCGAAAGGGTGGATAGTACCAAACTACGAGTTACCACCTAGCCTTGAACAG
ATCGAGATACTTAGAGTCGTGGTTAGAGAAAACGTAACTGAAGTGCTTATCGACAAGCTGATAGACGATATCGTCGA
GATAACCGAACAACTTGCTGATTCAAGTTCCTCAATGCACTCTCTGAATAATTTGGGACATATACAGAGGCCCAAGA
AACACGAACACCCAGAAAGTAATTTAAAGGAAGGGGAGGGGTCCGATTACTCAGGAACATACGCTCGTCCGTGT

Seq. ID No: 21
>AADC_20
ATGGCCTTAAACGCAGTAAGCGCAGCTAGAGGTAGCGCAAGACAGTATATTTCCACGTTCTTAACGCTAGACAACGC
TAAGAGCGGATTATTTTATTATGTATTATTGGTTCAAGCCATTAAAGTGAAACGTCACCTGCGTGCCAGAGGTATCA
GTGCCTCCCTAAAAGAATTATACACATGGATTTCTCAGCAGATCATACGTTTGCTTCTGAGGCTGCCGGCCACACGT
AAGAAAGTTGCTAGTCAGATGGATCAAGCAAAGTTGGATATCGAGAACAGACTGGTCCCCAAAGGGGCCAATGTTAC
CAGGCATTTAAGTCTGCCGTCTGAAGGCAAGTCCTTGGAATGGATTACACAGGAGATGGACAAGATGGATACCGAGC
TTGGTGGCACAAGTGACGCGTGGAGACAGGGGAAGCTATCCGGGGCTGTATATCATGGAGGTGACGAGCTTGCGAAA
ATTATCGTCGCGGCCTATAGTAGATATTGCGTCTCTAACCCACTTCATCCGGATGTTTTTCCAGCTGTACGTAAAAT
GGAAGCTGAGATCGTTGCTATGTGTCTAAAGATGTATAGAGGGCCTGAAGGTGCTGCCGGCGCTATGACGAGCGGGG
GAACCGAATCAATCGTTATGTCCGTTAAAACCCACAGAGACTGGGCAAGGAGTGTAAAGGGAATTAAAGAACCCGAG
ATGGTTGTACCCGTTTCAGCCCATGCCGCCTTTGATAAGGCGGCGGCGTATCTTGGGATCAAGCTACATTCTATACC
AGTTGATTCCTATACAAGACAGGTTAACATAAAGCACGTGAAAAGAGCTATCAACTCAAATACTATCATGATAGTTG
GCTCCTGCATCGGGTTCCCTGACGGTAATCAAGACGATATTGAAGCCCTTGGTGCATTGGCCAAGAAGTACAACATA
GGCCTGCATGTTGATTGCTGCTTAGGCAGTTTCATTGTACCGTTTCTTGAGCCAGCGGGCCTAGCAAAAGGCGATAA
CAAAGGAAGATACAAGTTGACTCCTTTCGACTTCACGGTAGATGGCGTCACCGCAATCTCCTGCGATACCCATAAGT
ATGGGTTTGCTCCAAAGGGCACGAGTGTAATAATGTACCGTAGTGCGGAACTGCGTAGATTCCAATACTACGTTAAC
CCGATCTGGCCAGGGGGCGTCTACGCATCTCCATCTCTTAGTGGGTCTCGTCCCGGTGCACTGATAGCAGGCTGTTG
GGCGGTGATGCAATATATGGGTACAGAGGGGTACCTTAGCAGTTGCCGTGATATAGTAATTGCTACTAGAAAGATAG
CTGACGCGATTACAGATGATATCCCTGAGTTATACGTCCTGGGTAATCCTCCGGCTTCAGTCGTAGCATTCGGGTCA
CGTAATCCAACTGTAGACCCTCTTGAAGTCGGAGATGGCATGCGTAAGAGGGGGTGGCATCTGAACGGCTTGAGTTC
ACCAAAATCTGTCCATATTGCTTGTACCCGTCTGACTTTACCCGTCGTGGACCAGTTCATTGCAGACCTAAAGGATT
GCGTCAGGGAAGCAAAGGTGGCACCTAGTGGGAAAGGTACGATGGTTTCCGTGTACGGGCTAGGCAACAGCTCTGCG
GTGGGACCCGACATGGTTTCCCAACTTGCCAGTGCCTTTCTGGACGCCCTGTACAAAGCA

Seq. ID No: 22
>AADC_21
ATGGAACTAAAGACTGCGGCCAATGAGATCTGTAACAGTAATCAGATGTGTCAAGGTGAGAACTATTCACAGAACAT
GCTATTGAGAGATGGGCTAATTGATTTGAAGAACCAGATCAAAGAAGGAGATGAGGGTCTAGGTCATGGCTTTAGCG
ACTTCTCAAACGTATTTGCATCAGACCTACTACCAGCGCGTAACGGCGAAAAATACACAGAGGGATTCCTATTAGAG
GTATTCAATATACTATTCAGTTATATCAGGAAGACCTTTGACAGAAAATCCAAGGTATTGGATTTCCACCACCCCCA
CCAGCTGTTGGAGGGACTGGAAGGGTTTAACTTGGAATTGAGCGATCAGCCAGAACCACTGGAGCAAATTCTTGGTG
ACTGTAGAGATACGTTAAAATACGGGGTTAAGACGGCACATCCGAGGTACTTTAATCAATTAAGCTCCGGCCTTGAC
ATGGTGGGCCTTGCAGGGGAATGGCTTACGGGTGCAGCAAATACCAATATGTTCACCTACGAGATCGCTCCCGTCTT
CATCATAATGGAGGGACTACTTATAAAGAAGATGCATGAGTTGGTAGGGTGGGGAGAGTTGGAGGCGGATGGGATAT
TCTCTCCAGGGGGCACCATTTCAAATCTGTACAGTGTTTTAGTAGCCAGATACAAATTTTTCCCTATAGTCAAGCTA

AAAGGTATGGCTGCGTTACCCAGAATCGTCCTGTTCACTTCTGAGCACTCACATTACTCTTTTCAGAAGGCAAGTGC
AACTCTTGGAATTGGAATAGAGAATGTTATCGCTGTCAAATGCGATGAGAGGGGTAAGATGATACCATCCGATCTGG
ATGAGAAGATTATGGCACAGAAAGAGAAA

Seq. ID No: 23
>AADC_22
ATGTGGAAAGCTAAGGGAACAAGAGGTTTCGAGCTACAGATAGACAGTTGTTTGGAGAACGCTGAGTACCTGTACAA
GAAACTAAAAAGCAGGAACGGCTTTGAGCTTGTTTTCCCCGACGAGCCGGAGCACACCAATGTCTGCTTCTGGTATA
TTCCCCCTAGTTTGAAGGGTATGCCCAGGGACAAGGAATGGAATACGAAACTTCATAAGGTAGCGGCCAAAATTAAG
GCTCAAATGATGGAAGAGGGAACGGTCATGGTGTCATACCAGCCGCTTAAAAATAAGCCAAACTTCTTTAGGATGGT
GTTCTCAAACCCAGCTTCCAAGAAAAGTGACATAGACTTTTTGCTAGATGAGATAGAGAGGTTAGGAAGTGACCTTG
AATTT

Seq. ID No: 24
>AADC_23
ATGTTTGGATCTCAACATCAGATGGATGTTGCGGCTTTAGATAGACAGCTGAAAGAGGACAAAGAGTCCGGGAAATT
GCCACTACTTCTGGTTGCAAATGCGGGAACACCGGGTGCAGGACACACGGACAAGTTGGCCAGGTTGAAAGAACTTT
GTCATCAGTACAACATCTGGCTTCATGTTGAGGGAGTCAATTTAGCGACCTTGGCCCTTGGGTATGTAAGCGCATCT
GTGCTGGCGGCAACGAAATGCGATAGCATGACCTTGACACTTGGACCTTGGTTAGGTCTACCAGCCGTACCGGCGGT
TACCCTATACAGACACGAAGATCCGAGTCTTAGCCTGGCTGCAGGGTTGACAACATCTCAACCCGTGGAAAAGTTGA
GAGCCTTGCCGCTATGGCTGAGTCTGCAGTACCTAGGCCATAACGGGATAGTTGAGAGAATTAAACACGCCTCTCAA
TTATCCCAAAGGCTATTGGAAAATCTTAAAGATGTGACATCCATTAAAACATCCGTGGAGCCCGATGGCAATTCTCC
GGTGGTCGTGTTTAAGTTTTTCTATGACGGACCAGGCTCAGGGTCCACCATTAATCTTAACACGATTGAACGTGAGT
CCGACGCAATGAATCAGTGGTTGGGCGAACAGTTGGCCGCCCTTATCCCAAGTTGCGCTGTAGATACTGTTGAGTTG
GAAGACGAAGGTGTGTGTGTGAGATTTAACCCTATGATGACATCTGCAGTACTTGGCACAACCATAGAGGACGTAGA
TCAGCTGGTTGAATGCGTCAAGGTTAAAATTCCCATAATACACAATACACTACAGCTAAAGGAGGAGTTCAGGCTTG
AAGTAGAAAGGATTGCGGGCCTAACATATGTTGTTGATTATTCTTGGGCGGGGTTGGGCGTTTTACGTTATGATCAC
GTTTCAGAAGAGCTAGATGGGTCCAGACGTGAAGCCGAATTAGAGAAGATTAATGCTTCCCTATTAAAGAAGTTGAA
TGAGCTTGAAAGCGACTTGTCATTTAGTTCCGGACCCGAGTTTGGTGCCGAGAAAAATTGTGTCTACATAGGTATGG
CAACAGAGGACGTAGATGTTAGTGAATTGGTTGAAACAATTGCTGTCATGGGCAGGGAGATAGAGGAAAATAGCAAA
CTGTTAGAGAACATGACTGAAGTAGTTCGTAAGGGCATCCTAGAAGCTGAGGTGCAACTACAGAAAGCAAACGAAGA
AAGACTTCTGGAGGAAGGCGTACTAAGACAGATCCCTTTGGTTGGCAGCGTGTTAAACTGGCTTTCTCCTGTCCAGG
CTACTCCTAAGGGCCGTACGTTCAATCTGACCGCCGGTAGTTTGGAGTCCACAGAAATTACTTACGCTAGTAAGGCA
CAAGCTAATGGAACTTCACCCCCACCAACACCAAGCCTAGGGCATGCCAAGAGACATCCTGGCCAGAAGCTGTTTAA
GCGTTTGAGTAGAAACAGTGACGCAATGTCCGAGACTAGCTCAGTTTCTCACCTAGAGGAAGTGGAAAATCTGGAGG
CGAGTCCCACCCCCGAGCCACAACCTGGACATCCTACCGAACCTCCTGTTCCATCCGTTGAATCAAACTCTGAGGAG
CCACATGAGGCGGAAGCGTTAGATACTAAAACTGTAGAATCTGAAAGTCTACGT

Seq. ID No: 25
>ADK1
ATGAGTTCTAGTGAGTCTATCCGTATGGTTCTGATCGGGCCTCCCGGTGCTGGTAAGGGAACGCAAGCACCCAATTT
GCAGGAAAGATTTCATGCTGCTCATTTGGCTACGGGTGACATGCTTAGGAGTCAGATAGCCAAGGGCACTCAGTTGG
GTTTAGAGGCTAAGAAAATAATGGACCAAGGCGGCTTAGTGAGTGACGATATTATGGTCAATATGATTAAGGATGAA
CTGACTAACAATCCAGCATGTAAAAATGGCTTCATTCTGGATGGATTCCCAAGAACTATCCCTCAGGCCGAGAAATT
AGACCAAATGCTTAAAGAGCAGGGGACACCCCTAGAAAAAGCTATTGAACTAAAAGTTGATGATGAATTGTTGGTTG
CGAGGATAACTGGTAGGCTTATCCACCCTGCGTCTGGCAGAAGTTACCACAAGATCTTCAACCCGCCGAAGGAAGAT
ATGAAGGATGACGTAACAGGGGAGGCCCTAGTTCAGAGATCTGACGATAACGCGGATGCTCTGAAAAAGCGTCTAGC
CGCATATCATGCGCAGACTGAGCCGATCGTGGACTTCTACAAGAAAACGGGCATTTGGGCGGGCGTAGACGCCTCTC
AACCACCAGCCACAGTATGGGCAGACATCCTAAATAAGCTGGGTAAGGAT

Seq. ID No: 26

>AOQS_1
ATGTCTAATGCAGCCATAAGATCCTCCCGTGCCGTATCCGTATCATCCAGCACGAAGTATTACGACTTCACGGTAAT
AGGTAGTGGTGTTGCGGGGTTAAGGTATGCCCTTGAGGTAGCCAAACAGGGTACTGTGGCAGTGATAACAAAAGACG
AGCCACACGAAAGCAATACAAACTATGCACAGGGAGGAGTAAGTGCAGTGCTTTGCCCGTTGGATTCAGTTGAGAGT
CACATGAGAGACACGATGGTGGCAGGTGCGCACTTATGTGACGAAGAAACCGTCAGGGTCGTGTGTACTGAGGGGCC
CGAGCGTATCCGTGAGCTAATAGCTATGGGAGCATCATTCGATCACGGTGAGGACGGGAACCTACACCTAGCGAGGG
AGGGCGGCCACTCCCACTGTCGTATTGTCCATGCCGCTGACATGACGGGTAGGGAAATCGAACGTGCTCTGCTTGAG
GCAGTTTTGAATGACCCAAACATTTCTGTGTTCAAACACCATTTTGCAATAGACTTACTAACATCTCAGGACGGACT
AAATACTGTCTGCCACGGGGTTGACACCTTAAACATCAAGACTAACGAAGTAGTTAGATTTATCTCCAAGGTGACCT
TACTAGCTTCTGGAGGTGCAGGTCATATCTACCCGTCAACAACGAACCCCCTGGTAGCGACCGGAGATGGTATGGCA
ATGGCCCATAGGGCGCAAGCCGTCATTTCAAATATGGAATTTGTCCAGTTTCATCCCACTGCACTGGCTGATGAAGG
CTTGCCGATCAAATTGCAAACCGCGAGGGAGAACGCGTTCCTAATAACTGAAGCCGTGCGTGGCGACGGCGGGATTC
TGTATAATCTGGGCATGGAAAGGTTTATGCCCGTCTATGATGAGAGAGCTGAGCTTGCTCCTAGGGATGTGGTCGCG
AGGTCAATAGACGACCAATTGAAGAAACGTAATGAAAAGTATGTACTACTAGATATCAGTCACAAACCGAGGGAAAA
GATCCTGGCGCATTTTCCCAACATCGCGAGCGAGTGTCTTAAACATGGATTGGATATTACTCGTCAGCCTATCCCAG
TAGTGCCGGCGGCTCATTATATGTGCGGCGGTGTGCGTGCAGGCCTACAGGGTGAGACCAATGTGCTTGGTCTTTTT
GTAGCAGGAGAAGTCGCATGTACAGGCCTTCACGGTGCGAATAGGCTGGCTTCCAATTCATTACTTGAAGCCCTTGT
CTTCGCTCGTCGTGCAGTACAACCCAGCACGGAATTAATGAAGCGTACACGTTTGGACGTGTGTGCATCTGAAAAGT
GGACTAGGCCAGTGGTGGCCACCGCGCGTCTGTTAGGGGATGAAGTGATTGCCAAGATCATCGCTTTAACAAAGGAG
GTCAGACGTGAATTACAAGAGGTTATGTGGAAGTACGTGGGGATAGTAAGGTCTACTATTAGACTGACAACGGCTGA
GAGGAAAATTGCGGAGTTGGAAGCTAAGTGGGAGACGTTTCTGTTCGAGCACGGGTGGGAACAGACGGTTGTTGCCT
TGGAAGCCTGCGAGATGCGTAACCTATTCTGTTGCGCAAAACTTGTAGTGTCATCTGCGCTTGCCAGACACGAAAGC
AGGGGGCTGCATTATATGACGGATTTCCCGTTTGTCGAAGAGAGCAAGCGTATTCCAACCATAATTTTACCGAGCAG
TCCAACTACTGCCAGCTGGTCAAGTAGGAGATTGCAAAATATTTCTTCAAGCAGCTTGATAGATTGTGGGAGTGGCG
AAGGACGTGGCTCTCTACTGACTTGCGGTGACGTTGAAGAGAATCCAGGACCTAGCAGCTCCTCTTCTAGTCAAACA
ACAGAGTTAGTCCCCTATAAACTGCAGCGTCTTGTTAAAGAATTTAAATCATTGACAGAACCGATTGACAGGCTGAA
ATGGGTGTTACATTATGCGTCCCTTTTGCCACAGATGCCTGAGTCCTCAAAAACCGAATCAAATAGGGTGATGGGCT
GTACAGCGAGGGTTTGGCTGGATGCCGAATTAGGCCAGGATGGCAAGATGAGATTTTGTGCCGACAGTGATTCAGAC
GTGTCTAAGGGGATGTGTAGCTGTTTAATACAGGTTCTGGACGAGGCAAGCCCTGTCGAGGTAATGGAGCTTAAAAC
TGAGGACTTGGCAGAGTTAAATGTCGGTCTGCTTGGCGGAGAAAGGTCACGTGTTAACACTTGGTATAACGTACTTG
TCAGCATGCAAAAGAAAACTAGAAGGTTAGTTGCAGAGCGTGAAGGGAAGGTGCCGTCCTTTGAGCCATTTCCTTCG
CTAGTACTAACCGCCCACGGAATCGAGGCAAAGGGGTCCTTCGCCCAGGCGCAGGCAAAGTATTTATTCCCTGAGGA
AAGTAGAGTCGAAGAGCTAGTTAATGTATTAAAGGAGAAGAAGATCGGTGTCGTGGCACACTTTTATATGGACCCAG
AAGTGCAGGGTGTTTTGACAGCTGCACAGAAACATTGGCCCCATATCTCAATCTCAGACTCTTTAGTCATGGCTGAT
TCCGCGGTGACTATGGCAAAGGCTGGATGCCAGTTTATAACTGTCCTTGGAGTCGATTTCATGTCAGAAAACGTGAG
GGCCATCTTGGACCAAGCCGGTTTTGAGAAAGTTGGAGTCTATAGGATGTCTGACGAAACAATTGGTTGCTCATTGG
CTGATGCAGCTTCAGCCCCAGCTTACCTTAATTATCTGGAGGCGGCGTCCAGGAGCCCACCGTCTCTACATGTTGTT
TACATAAACACTTCCTTGGAGACGAAAGCATTCGCGCATGAGCTGGTTCCTACAATAACGTGCACATCCTCAAACGT
AGTACAGACTATTCTGCAGGCTTTTGCTCAGATGCCGGAACTGACAGTATGGTACGGACCCGATTCATACATGGGCG
CGAACATTGTCAAATTGTTTCAACAGATGACACTGATGACAAATGAGGAAATAGCAAACATACATCCCAAGCATTCA
CTGGACTCAATAAAAAGTTTATTACCGAGGCTGCACTACTTCCAAGAAGGTACATGTATAGTGCATCATCTTTTCGG
GCATGAGGTTGTAGAAAGGATAAAGTATATGTATTGCGATGCCTTCCTGACAGCCCACCTAGAGGTGCCCGGTGAAA
TGTTCTCCCTAGCGATGGAGGCAAAGAAAAGGGAAATGGGCGTCGTCGGTTCCACGCAAAACATACTAGACTTTATC
AAGCAGAAAGTCCAGGAAGCGGTAGACCGTAACGTCGATGACCATTTGCAGTTTGTTCTAGGTACTGAATCAGGAAT
GGTGACGTCCATCGTTGCAGTAATTAGAAGCCTATTAGGGAGCAGCGCCAACTCCAAACTTAAAGTGGAGGTAGTGT
TTCCAGTATCCTCCGATTCTATGACAAAGACTTCTTCTGACTCCAGCAATAGTATAAAAGTCGGAGATGTGGCCCTA
CCGGTGGTACCAGGGGTGGCGGGTGGCGAAGGGTGCTCCATACACGGAGGCTGTGCAAGTTGCCCCTACATGAAGAT
GAACTCCCTATCTAGCCTTCTGAAAGTCTGTCATAAACTGCCTGATTTGGAAAATGTGTACGGTGGGTTTATCGCGG
AGAGGTTCAAGAGACAGACGCCACAGGGAAAACTAATAGCAGATGTGGGCTGTGAGCCAATTCTACATATGAGGCAC
TTCCAGGCAAATAAGGAACTTCCTGATAAACTGGTTCATCAGGTCTTATCATGCGAGTCCAAGCGT

Seq. ID No: 27
>AOQS_2
ATGAATACCCTACCTGAACACTCATGTGACGTACTAATTATCGGCAGTGGAGCGGCGGGACTGAGCCTGGCACTTAG
GTTGGCGGACCAGCATCAGGTCATTGTACTGTCTAAGGGACCGGTAACAGAAGGTTCAACTTTCTATGCGCAAGGTG
GCATTGCCGCGGTGTTCGATGAGACGGATTCTATAGACTCCCATGTTGAAGATACCCTAATAGCCGGTGCCGGCATC
TGCGATAGGCACGCCGTTGAGTTTGTGGCCAGCAATGCCAGAAGCTGTGTACAGTGGCTGATTGACCAAGGGGTGCT
GTTTGATACCCACATTCAGCCTAATGGCGAAGAATCTTACCACCTGACACGTGAAGGTGGCCACTCCCACAGGCGTA
TATTGCACGCCGCTGACGCCACGGGCCGTGAAGTGGAGACAACTCTGGTCAGTAAGGCGCTAAATCACCCGAACATC
CGTGTGTTGGAACGTAGCAATGCGGTGGACCTTATAGTCTCTGACAAAATCGGGCTTCCCGGGACTAGACGTGTCGT
GGGCGCCTGGGTCTGGAACCGTAATAAAGAGACGGTAGAAACATGCCACGCCAAAGCTGTGGTTCTGGCTACAGGTG
GCGCATCAAAAGTATACCAATACACGACTAACCCCGATATTTCTAGCGGTGACGGCATAGCAATGGCTTGGCGTGCA
GGCTGCAGGGTAGCGAATTTAGAATTTAATCAATTTCATCCAACAGCGTTATATCATCCTCAAGCAAGAAATTTTCT
GCTAACGGAGGCGTTGAGAGGTGAAGGTGCCTATTTAAAAAGACCTGATGGAACGAGATTCATGCCCGACTTCGATG
AGCGTGGCGAATTAGCCCCAAGGGACATCGTCGCACGTGCTATTGATCATGAGATGAAAAGATTGGGTGCTGACTGT
ATGTTCTTAGATATTAGCCATAAGCCAGCTGATTTTATCAGACAACACTTTCCAATGATTTATGAGAAGCTATTAGG
CTTAGGTATTGACCTGACCCAAGAACCTGTCCCAATAGTGCCCGCTGCTCATTATACCTGCGGTGGCGTCATGGTGG
ACGATCACGGCCGTACGGACGTTGAAGGCTTGTATGCAATTGGGGAGGTGAGTTACACCGGGTTACATGGTGCTAAT
CGTATGGCATCCAACAGCCTGTTGGAATGTCTAGTATACGGCTGGTCAGCAGCTGAAGACATCACGCGTAGAATGCC
GTACGCGCACGATATCAGTACATTACCGCCGTGGGACGAGTCTAGGGTCGAAAACCCTGATGAAAGAGTTGTAATTC
AACATAATTGGCATGAACTAAGGCTTTTTATGTGGGACTATGTGGGTATAGTAAGGACCACAAAACGTTTGGAACGT
GCTCTTCGTCGTATAACCATGCTACAGCAGGAAATAGACGAGTATTACGCTCACTTTAGAGTCTCAAATAATCTTCT
TGAGTTGAGGAATTTGGTCCAAGTCGCGGAGCTAATCGTGAGGTGCGCAATGATGAGAAAGGAGTCCAGAGGGCTGC
ATTTCACTTTAGACTATCCCGAACTTCTGACACACTCAGGCCCCTCTATATTGTCCCCAGGAAACCACTATATTAAT
AGAGGTAGTGGTGAAGGCAGGGGATCTCTTTTGACTTGCGGAGACGTAGAAGAGAATCCTGGACCCGGAAGTGTGAT
GTTCGACCCCGATACCGCGATTTATCCGTTTCCTCCCAAGCCCACACCACTATCTATTGACGAAAAAGCATACTACA
GGGAGAAGATTAAGCGTTTGTTAAAGGAACGTAACGCGGTGATGGTGGCCCATTATTATACCGACCCGGAAATACAA
CAACTTGCAGAAGAGACAGGGGGCTGTATCAGTGATTCACTAGAAATGGCAAGGTTTGGAGCCAAACACCCGGCCAG
TACTTTACTGGTGGCTGGAGTACGTTTCATGGGGGAAACCGCAAAGATACTATCCCCAGAAAAGACCATCTTAATGC
CAACCTTGCAAGCGGAGTGTAGCCTGGATCTTGGCTGCCCCGTTGAGGAGTTCAACGCATTTTGTGATGCCCATCCC
GATCGTACAGTCGTAGTCTACGCCAACACCAGTGCAGCCGTCAAAGCTAGGGCTGATTGGGTGGTAACGTCAAGTAT
CGCGGTGGAGTTAATCGATCATTTAGACAGTCTAGGAGAGAAAATTATTTGGGCCCCAGATAAACACCTGGGAAGGT
ACGTCCAAAAGCAGACGGGTGGGGACATACTATGCTGGCAGGGTGCTTGTATCGTGCACGATGAGTTCAAGACGCAA
GCATTGACAAGGCTACAAGAAGAGTACCCAGACGCAGCGATCTTAGTACACCCGGAATCACCGCAGGCGATTGTGGA
CATGGCGGATGCTGTCGGCAGCACCTCCCAATTAATTGCGGCGGCAAAGACTTTACCTCACCAGAGATTAATAGTGG
CTACAGACAGAGGCATATTTTACAAGATGCAGCAGGCCGTACCAGACAAAGAATTATTAGAGGCGCCGACGGCTGGA
GAGGGCGCGACCTGTAGGAGTTGTGCACATTGTCCTTGGATGGCTATGAATGGCCTTCAAGCAATCGCAGAAGCCCT
GGAACAAGAAGGCTCCAACCATGAAGTCCACGTGGATGAGAGACTGAGAGAACGTGCCCTGGTCCCTTTAAATAGGA
TGTTAGACTTTGCTGCAACCTTGAGAGGA

Seq. ID No: 28
>ATMT_1
ATGACACCGGCTGCCGGTAAGACCTTCAACACTTCCATCGCTGGCGCCGACGACCTGATTAGACTTCATCTATCAGA
GTCCCCTCATGGCGCCAGTAAGGCAGCTCTTCAAGCCGCCGAACGTGAACTAGCCCGTGTCAATGTGTATCCAGATC
CGGAACGTCAGGAATTGGTCCGTGCTCTAGCCGCCCATTGGGGAGTCGGACCGGAGCACATTGCGGTAGCAAATGGA
TCAGACGAGTTGGTACTGGCGACAGCCCTAACTCTAGGTGATCGTAACCTTCCGGGTTTGGTCACGGATGGGACGTT
CCCCGGCTATCGTGCGTGTCTTGAACTGCTAGGGCGTGGGTGCACCGCGGTCCCGCCTGACGGGACAGCCGTCGATG
TGGCAGGATTCGCAGCCAGATTGCCCGGACATGGTATCGGCTACCTGTGTAACCCTCATAACCCATCAGGAGCAGCA
CTAACAAGACAAGAACTGGCGGCGCTGGTTGAAGTGTCCGGGCGTAGTGGTGTCCCTCTTGTATTTGACGAGGCCTA
CATGGAATTTGCGGGGCCGGATGTGCCACAGACTAGGGATTTAACAGCGGCCGGAGACGCACCGGTAGTCGCGCTGA
GAACGTTCAGCAAAGCCTATGGCTTGGCGGCACTAAGAGTTGGATATGCCGTGGGTAGGCCTGACTTGATAGCCGGT
TTAAGAGGAACGCTTAGAGCACTGCCGTTCAGTGTTAACCGTTTGGCTCAAGCAGCGGCCATAGCAGCACTGGGAGA

CCCTGATTTTGTTGACGGAGTTAGGAGAAGCACTGCTGAAAGGCGTAGGTGGTTTGTAGGCGAGTTGGACAGGCGTG
GCAGGGCTCATTTGCCGTCCGTAACGAACTTTGTCGCCGTTGCAGCTAGAGACTGTGCGCGTGCACAGGACAGACTG
GCTGCTGACTTCGGTATCTTGGTCAGAAACGCGGGCTTATTCGGGTTCCCTGGCTACCTGAGGACTTCACTAGGCGA
GAAGAAGGACTTAGAGAGGTTTTTGGACGCTCTGGATGAGATTGAACAAAACCCCGGTGGCGGATCAGGTGAAGGGC
GTGGCTCTCTACTTACGTGTGGTGACGTAGAAGAGAACCCGGGTCCCATGACTGCCCCGTTAAGCCGTGATGGTTTA
CGTGCAATGGGTGAGTCCGTCTTTCGTCCAGCGGAGTGGCAAGGCGCGGCGCATACGCCGTTGGACGCGGATACAGC
TTTCAACGGCTTTATCAGTACACATGTTGTTTTTGCCCTTGAACAACTAGGTTTGTTCGCCTGGTTTGATGAAAGCG
ATAGGTTAGATGTTCCTCAGTACTGCTGGCGTCGTAAGTTAGATGAACGTGTTTTTCGTCAACTTGTGAGTGCCGCC
GAGGCATTCGGTTATCTTGATGTTCATGATGACTTGGTGACCCCAACGCCAGCTTGGTCCGAACTGAGGAGAAAAAT
TGGCTTCTTCACCTGGGGCGTAGGCGGATACCACGATGTATTTGCCAACGCGGCTTCAATAGCTCGTGGGGAGAGGG
CCTTCGGTAAAGACGTGCTACGTGATGAGGCTATGGTGGCCCTAGGCTCAGCACAAGCGGACATGGCGCTGATGAGA
GACCTGCTGGATGAGCAGATTGCTGCGTTGGACTTTTCCGTAATAGCCGATCTTGGTTCAGGTATAAGCGAACGTGT
CTGTCGTCTGGTGAAGTCAAGACCGGGAGCCAGAGGTCTTGGGGTCGATATATCAGCGTCTGCGACCGCATTGGCAG
CTGGGACGGTTGAACGTCACGAACTAGCTGATAGAGTTCAACCTATTTGCGCAGACGTCTTAGATGTTCTGTTTCAC
GGCAGAAGGATAGAGGGTGCGGACCAGGTCGATGTAGCAATGTCATTCATGTTCCTGCACGACTTGCTAGTGGACCC
TACGACCAGAACTGACGTGATCCCAGCACTTAGGAAGGCATTTCCTAGGGCCCATACGTTCTTGCTGGCTGATACGA
CAGTTCGTCCTAGGGACGAGAAGGATACATTGCCCGTATTCTCAAGCGGTTTTGAGTTAGCACATGCGTTAATGGGC
GTCCCAATTTATACGAGGGAGGAATATGAGAATCTTTTCCATGAGGGAGGTCTTCATCTGCGTAGGACCGTTCCGTT
TGGAGCGCCTCATACATATTTATTCGTTCTTGAAGCTCAG

Seq. ID No: 29
>ATMT_2
ATGCAGGCATTGCCCGTTAAGGGTGACACTGTAAGCCGTCCTCCCACTGTCCACAGCCTGCACCATGAACATGAAAG
GGCTGATGGAATGTTGAGACTGCACTGTAACGAAAATCCTTACGGTCCGCCTTCCGGTGTTATCGCTTCTGTAACTA
AAGAACTGGAGGGAAGATGTAGCACGTATCCTGACTCCGAAGTCACTGCTCTGCGTGAGGCTTTAGCTGGGCAGGTA
GGGGTTGGTACCGACATGGTAGCTGTGGGTAACGGGGCTGATGAATTAGTCCTACTTATTACGTTAGCCTCTGCGGG
TCCTGGGGATACTGTGGTTGTGACCGAATCTACGTTTCCAGGTTATGCAGCTTCAGCTGCAGTGGCAGGCGCGACTG
TAAGAGGGGTGCCTCTACACCGTGACCGTGTATCTGCGACAGCGTTGGTAGAAGCTGTCGATGATGGAGCGAGGTTA
GTCTTTGTATGTAACCCTCATAACCCAACGGGTACCGTGCTTAGCCCCGCAGCAGTCGAAGAAATCTTGAGGGCATG
TGAGAGGACTGGTGCGGTGCCTGTTTTTGACGAAGCATACATTGAATTTGCTGGTCCCGGGTTTGATCATGCATTGG
ATGCAGTCAGAGCCGGGAGACGTCTACTGGTACTGAGAACTTTTTCAAAGGCCTGGGGACTGGCAGCACTGAGGGCG
GGGTACGCCGTCGGTCCAGCCGATCTTGTCGCGGGTATAATGGAGGCCAGGAGACCCCTGCCATTTTCCGTAAATAG
GTTGGCGCAACAGGCTGCACTTGCCGCGCTAGGATCTCCCGATCATATTGCAGAGGTCTACGAAAGGACAACACGTG
AGAGGGAACGTCTGTGTCGTGCTCTAACTGGCCTGGGGGTCGCATATGTGCCGTCCGTTACAAACTTTGTAATGGTG
AAGACTCCTGGAAACTCCACGAGATTTGCTTCTCGTTTGGCTGATGAGCATGGAATCTTAGTAAGGGATCTAGCACC
CTTCGGTTATCCTGGTCACGTGAGAGTAAGCGTGGGTACAGCTGAGGACACTGATCAATTCTGCGCAGCCCTAGGAA
GCCTTCTTGCTTCTCCGCGTTCTCATGCCGCCACCGGGCATGGTTTGGGGGCATCCTCTGGTGCTGGCGGAGCGGGA
AACGCTGCCATCAGGTCCGCGAGAGATGTTCTGCCGGTACCAACGCTTGATCCCGTCGCTCCTCAAGACCTTTTCAA
CGGTTATGTCGGGGCTCATGCAGTATTCGCGCTAACTCGTTTGGGCGTATGGGATAGATTGGCGGAAGGTTCTGAGC
CAACTGTAGATGCCTTAGCGGTTCAGGCAGGAACGGATGCTACAGGCCTGATGCCACTACTGCGTGTTGCCGCCTTA
CTGGGCTACGTGAGCTTAACTGACGGTTCCGCCCCTGCTGTACGTTTGACTGAGTCTGGGAGGGAACTTGTCCGTAT
GCGTGGCTTTTTCACGTGGGGCGTCGGAGGATATCACGAAGTTCTTAGATCGTTGCCGGCATTGGCCAGGGGGACCT
CCGTATTCGAGCAGGACGTCGATCGTGATGGAGGGATGGTAGCCGTGGGGTCTGGCGAGGTCGGACGTGAAATGATG
CTTCCTTTGGAACAAGAAGTCCTTGCAACAGTTGACTTCAGGACTGTGGCCGACTTGGGATGTGGAGATGCCACTAG
GTTGTTGAGGTTGTGCGATGGACACCCACACAGGAGGGGGACAGGCATCGAAATTAACCAAGGCGCTTGCGTACAGG
CTAATAAGCGTGTTGCAGACGCGGGGCTGGCTGATAGAGTGGACATTGTGCATGGTGATGCTCTTGACCTTTCAGGA
CGTACGTTTCCGGAAGTCGATCTAGTCACTTCCTTCCTGATGATGCACGATCTATTTGACGCGACAGGTGACCCCGT
GGGAGTCATGAGGACCTTGAGGGAGGTCTTCCCGAGAGCGCGTCACTTCCTTATAGGGGATACAGTAGCCCAGGACT
GGGAAGAGAGACGTGAGGGTTTGCCAATGTTCTCAGTGGGCTTTGAATTGGTTCACGCGTTTATGGACACTCCTATC
ATGAATAGGGGCACCTACGAGGACGCTTTCGCGGGCGCCGGTCTTCGTGTAGCCAGAAGGGAGCCACTAGGGGCTCC
TAGTACGTGGCTTTGGCTATTGTCTACAGAA

Seq. ID No: 30
>ATMT_3
ATGAGGCGTCGTTGGGCCGTTACGGCCTCTGCATCCTGGTGGGGGACCGCCTGTGAACTTCATGCCAGTGCGTCTGC
CGCATATACGCCACCGTGCCACTCTCCGGGTACAGGAGGAAGGGGTACCGAGTCTGGCCCTATGACGGCACCTGTCC
GTCAGGAGACGAGAAACTATAACGCATCTGTACCATCCGCTGACGACCTTGTCCGTTTACATCTTAGCGAGAGTCCG
TACGGTGCGTCTCCCGCGGCAGTGGCTGCGGTCACAGGGGAGTTAGAGAGAATAAATAGGTATCCAGCTCCGGGCAG
AGAGGGTTTGGTCCAGGCTTTGGCAAGACACTGGGAACTACCCGAGGAGCATATAGCTGTCGCCAATGGGAGCGACG
AGTTAGTTCTAGCCACGGCGCTAACTTTAGGTGACCCCGGGTCACCTGGTCTAGTCACCGCGGGGACCTTCCCTGGG
TACCTTGCAGCCCTAGAAAGAATAGGACGTGGTGCGGTACAGGTTCCCTTAGCAGGATCTGGAACCGACACTGCAGC
CTTTGCAGACCGTCTTCCAGGTTGCGGAATTGGTTACGTGTGTAATCCACACAACCCTTGCGGGTCCGCATTGACCC
ATGATGAACTACACCGTTTAGTCGCCGCGGCGCGTGACTCTGGGACCCCATTGGTGTTCGACGAGGCATACCATGAA
TTTGGGCCCCCGGCGCAGCCCCAAGCACGTACCCACCTAAGGGAGGATACTCCAGTCTTAGCTCTAAGGACGTTCAG
CAAGGCATACGGCCTGGCCGCACTTAGGATAGGGTACGCGTTGGGTCCTGCCGATCTGATTGCTGAGGTGAGGAGAA
CATTAACTGTGCTTCCTTTCAGCGTGAATAGAGCAGCGCAGGCTGCCGCTCTTGCTGCATTAGATGATCAGGAGTTC
TTAGGTAGCGTGAGAAGGGATTCAGCCGCCAGACGTCAGTGGTTTTGCGCTGAGCTGGAGCGTAGGGGTTATAGATA
CCTTCCGTCTGTTACGAATTTTGTGGCGGTAGAAGTTGCTGCCTCTGCAGAGGCACAAGACGTGCTGGCAAGGGATC
ATGGTATTTTGGTGAGAGATACCGGAATGTTTGGATTCCCCGGTCACCTGAGAGTGAGCTTGGGCTCAGTGGAAGAA
CTACGTGGATTCCTGGACGCCCTGGATAGAGTTACGGCTGGCAGTAGGGGTGGGGGCTCTGGAGAGGGAAGAGGTTC
CCTTCTAACTTGCGGGGATGTAGAAGAAAACCCAGGACCGATGACCGGGCCAGTTAGTACTTCAGCTCCAAGTCGTT
GGCCAAGAACTTGGAGGCCTAACAGGCTGGAACCCACCTCACGTGGTGGCCAACCAGGTCATGCAGCGAGGAGGAGT
CCTGCAGCGGGACGTAGGAGGAGAAGGGCTTCAGAAGCCCGTCCGCCCCCTAGCGGAAGACAACCGGCTGTGAGGAC
GGAAAGGTGTGAGAGAGTGAGCCCGTTAAATACGCTGCCAAGTGAATGGCAGGGCCAGGCACCTACTCCACTGAACC
CAGACACTGCATTCAATGGGTACATCTGTGCTAACGTATTGCATGGACTTGAGCGTCTTGGAGTATTCGAACTTTTG
AGGGATGAAAAGAGCTTAGATATGGATCGTTTCTGCGAGACGAATGGTCTCGATTCTGCGGTTTTTAGGGCGCTTGT
CGGAGCTGCAGAGTCATTCGGCTATCTTGACGTCAGAGGAGCCCAAGTAAGAGCCACGTCTGTGGGTGAGGATGTTG
CTAGATACTTGGGTTTCTTTACGTGGGGTGTTGGGGGATACCACGATATTTTTGCAAGTGCTGCTCCTGTCGCCAGG
GGCGAGAGGAGGTTTGGCGTTGATTTGCACAGGGATGAGGGGATGGTCGCCCTGGGAAGTGCCCAGGCTGATACCGC
GTTAATGAGACACATCCTTGACGAGGAAATAGCAGGTATAGACTTCAGAACCCTGGTCGATCTTGGAGCCGGGGTAT
CCGAGCGTGTGAGTAGACTCGTGAAGGCACGTCCTGGTACTCGTGGGATCGGGATCGATATATCACGTCCAGCCACT
GAACTTGCCAGAGACACCGTCGCAGGATACGGCCTGGCGGGAACGGTCGAGCCTGTGTGCGCCGACGTCTTGGACAT
ATTGTTCAATGGTCAAGAGATTGACGGTGGGGATGCAGCAGACGTTGTAATGTCATTTATGTTCCTACATGACTTAT
TAGCTGCGCCCGAAAGAAGGGAAGAAGTAGTACCTAGATTACGTAAGGCTTTTCCGAGGGCCCACACGTTCTTGTTG
GCTGACACTACGATAAGGCCACGTAACGAGGAAGGAGACGGCAGGTTGCCTGTCTTCTCATCCGGGTTTGAACTTGC
TCATGCCCTTATGGGTGTTCCTCTGCACACCAGAGAAGAATACGAAGAACTATTTGAACGTGGAGGTATGAAGCTAC
GTAGAAGTGTACCTTTCGGGGCGCCTCATACCTACCTATTCGTCCTGGAGGCGAGT

Seq. ID No: 31
>ATMT_4
ATGACAAACGACCCCTCCCCAAGAGATGCACGTGACGAACTGCCTGTGAGAGACGAGCTGCGTGGTCAATCCCCTTA
TGGGGCGCCGCAATTAGACGTTCCGGTCAGACTAAATACTAACGAGAATCCGTATCCTTTACCGGAGGCCCTGGTCG
AAAGGATCGCGGAGAGAGTCAGGGAAGCCGCAAGGAGCTTGAATAGGTATCCCGATAGGGATGCCGTCGAACTAAGG
ACTGAACTGGCACGTTATTTGACGAGAACAGCTGGGCATGAAGTTACCGCCGCGCATGTTTGGGCAGCGAACGGATC
AAATGAGGTCCTTCAACAGCTACTACAGACTTTCGGGGGCCCAGGTCGTACTGCTATAGGCTTTGAACCATCATACT
CAATGCATGCACTGATATCAAGATCCACTGGCACAGGATGGATCTCCGGGCCAAGAAATGACGATTTTACGATCGAT
GTGGATGCGGCTAGAGCAGCTATCGCCGAACATAGACCAGAGGTCGTGTTTATTACCTCCCCCAACAACCCGACAGG
TACTGCAGTTAGAGCAGAAACAGTACTAGCGTTATATGAGGCAGCACAGGCAGCAAGGCCGAGCATTGTAGTGGTGG
ATGAAGCATACGGGGAGTTCAGCCATCATCCCAGTCTACTGCCCCTTATAGAAGGCAGGCGTCATTTAGTTTTGTCC
AGAACAATGTCTAAGGCGTTTGGAGCTGCTGGTTTGAGGTTAGGATATCTGGCTGCCGATCCAGCCGTGGTTGATGC
TGTTCAATTAGTGCGTTTACCTTATCATTTATCCTCTGTTACGCAAGCTACCGCACTTGCAGCCTTGGAACATACTG
ATACCTTGTTAGGCTATGTCGCCCAGCTTAAAGGTGAAAGAGACAGGTTAGTTGCTGAGCTGAGAGCCATCGGGTAC

GAGGTAACGGAAAGCGACGCGAACTTCGTACAATTTGGCCGTTTTGATGATTCACATGCTGTCTGGCGTCAGATATT
GGATAGAGGTGTTCTGGTCAGGGACAACGGCGTACCTGGGTGGTTGAGGGTAACTGCTGGTACCCCAGAGGAGAACG
ATGCCTTCTTAGACGCCGTTAGAGAGCTTAAAAAGGAGCACGATGCCGGCGGTGGAAGTGGGGAAGGGCGTGGGTCC
CTGCTTACTTGCGGTGACGTAGAAGAGAACCCCGGCCCATCTAGTTCATCTTCCACACGTACTGATTTTGCCCAGTC
TGCAGTGGCCAGCATCTTTACAGGTGCGATAGCGTCACACGCTGCGGTTTTAGCAGATGATCTGGGATTGTTTGATG
CACTAGCTAAAGGCAAGCTGAGGAACAGAGACTTAGACCGTTCCCCTTGGTTGAGAAATCGTATCAGGATTAGCGGT
GCCCTAGAGGCTCTGTGCAGGGTCGGAGCCGTCCAGAGATGCACAGACGGCTACGAGCTGACCGACGTAGGGACGGA
ACTGGCTGGCCAGGTTCCAGTATTCCGTTTGTGGCTGGGAGGTTACGCTTCAGTGCTGGCTGGGCAAATTTCAATCG
GCGCTGACCCTGCGACGGGTGTCCACGGGGGAATCGTGGCCGAATCAAGCGGTGCTATCGGGGCACGTTACCTAGAT
GAGACGATAGTCAACTTGTTAGAGTCACTGAGGCCAGAGGGTAGGATATGCGATATTGGATGTGGAACCGGGGCTAG
ACTGTTAAGAGTCTGTCGTAGAGTCAATCAACCCGGTATTGGATACGATTTGTCAGCCAAGGCTGTGGAAGCAGCAA
GGGAGACTGTTGATGAAGCCAGAAGGATTGGTGTAGATATCGACGTGAGGCAGGGAGACGCCACTGCGCTTACTCAA
GATCATCCTGATGTGGACATAGTGACCCAGGCGTTTATGACTCACCACATAGCTCCTGACGAATACTGCGCGGCCGT
TTTGAGGAGTTATCGTTCCCGTTTTCCCAGAGCTAGGTATTTGGTTATTTTCGACACTGTTCCGTCTCAGGACTCCG
AGGAACCCGAAATCTTCGCCCCCGGCTTCGACTATATTCATGCACTACAAAATATGGAGCCGAGGTCACGAGGGGCG
GCTAGGAGAATGTTCACGGAAGCGGGTTATATTTGCCGTGAGGAAGTTGAGTTGGCTGTTCCCAACAGTTACGCGTG
GGTTCTGGAGATGCGTGATAGAGAAGGCCCGGCTTCC

Seq. ID No: 32
>ATMT_5
ATGACCAATGACCCTTCCCCACGTGATGCCCGTGACGAATTGCCGGTGCGTGACGAATTGAGGGGACAAAGTCCATA
TGGTGCACCACAGCTAGATGTTCCGGTTCGTCTTAATACTAACGAGAATCCTTACCCACTTCCTGAGGCATTAGTAG
AAAGGATCGCTGAGAGAGTAAGGGAAGCAGCAAGGTCCTTGAACAGGTACCCAGATAGAGACGCCGTAGAATTAAGA
ACCGAACTAGCGAGGTACCTGACAAGAACCGCAGGCCATGAGGTTACCGCTGCACACGTATGGGCCGCAAATGGCTC
CAATGAAGTACTTCAGCAGCTTCTGCAGACTTTCGGAGGGCCTGGCAGAACTGCTATTGGGTTCGAACCGTCTTACT
CAATGCACGCTCTGATTTCTCGTAGCACAGGGACCGGATGGATAAGCGGACCAAGAAACGACGACTTTACGATCGAC
GTTGATGCCGCCAGAGCTGCCATTGCCGAGCACAGGCCCGAAGTGGTTTTCATCACCTCCCCCAACAATCCAACCGG
TACGGCAGTGCGTGCCGAGACAGTTCTGGCGCTATATGAAGCGGCGCAAGCGGCCCGTCCCTCCATTGTGGTAGTGG
ATGAAGCATACGGCGAGTTCAGCCATCATCCTTCACTGTTACCATTGATAGAAGGTAGGAGGCACCTTGTCCTATCC
CGTACTATGTCCAAGGCCTTTGGGGCAGCCGGTTTAAGATTGGGGTATCTGGCCGCCGACCCAGCAGTGGTCGATGC
GGTACAATTAGTCAGGCTTCCCTATCATTTAAGCAGTGTCACCCAAGCGACTGCCCTGGCAGCGTTGGAGCACACGG
ATACGCTACTAGGATATGTAGCCCAGCTGAAGGGCGAGAGAGATCGTTTGGTTGCCGAATTAAGAGCGATAGGTTAT
GAAGTCACCGAGAGTGACGCAAACTTTGTGCAGTTTGGCCGTTTTGATGATAGCCATGCAGTTTGGAGGCAGATCCT
GGACAGGGGAGTGCTGGTGAGGGACAATGGTGTGCCAGGTTGGTTAAGAGTAACAGCCGGCACACCGGAGGAAAACG
ACGCATTCTTAGACGCGGTTAGAGAGCTTAAAAAGGAGCATGATGCGGGCGGGGGATCAGGAGAGGGGAGAGGTAGC
TTGCTGACTTGCGGTGACGTCGAGGAGAACCCGGGGCCTGCTCAGGCGGCACCCACAACGGTGACAGAGGTGTTCAA
TCACGCGATAACTGCATCCGCTATCAGTGCAGCTTGGGAGATGGGGGCCTTTGACGCTCTGAGGGTTAGTGAGAGGC
TAGACGCTGATGAGTTTGCTGCGCGTGAAGGACTTGACACTAGATCCACACATGAACTTTTTCGTGCTCTGGCTGCT
GCTGACATCGTATCAAGAGACGGCGCGCAGATTCGTCGTGGTCCGAATTTCGCTGAAGCTGATAGATGCAAAAGCCT
GTTTCACTGGATGACAAGAGGGTGCGGGGAATTATTTAGCACACTGCCTGCGCTAGTTCGTGAGAAAAATAGAGTGG
GTTCCTTTTACCGTAGAGATGCCGCAGCCATCTCAGTTGCATGCCGTGAAATTAACGCGGAATGGTGGGACCCTGTT
TTCTGGCCAGTCGTCTCTGGTCTTGATTTCACGAGTGTTGCAGACCTTGGATGCGGTTCTGGTGAACGTTTGATTCG
TTTAGCTAGGACGGGACCGGAGGTGATGGCCCTAGGTATTGACTTTGCCGCTGGGGCGATCGAAGTTGCTACGGCGG
CTGTTGCTGAGGCTGGTTTGTCAGACCGTATCTCATTAGTCCAGGGCGATGCCACAGCTTTAGAGCCTCGTCCCGAA
TTTGCTGGTGTAGATTTACTAACTTGTTTCATGATGGGGCATGATTTCTGGCCGAGAGCAGAGGCAGTGGCTAGTCT
ACGTCGTATAAGGGAGGTTTTTCCGGATTTGAAACACTTCCTTCTTGCAGATGCAACGAGGACGACGAGTTATCCTG
ATACGGATATGCCTGTTTTTAGCATGGCATTCGAACTTGCGCACGCAGTGATGGGAGACTACTTACCTACCTTAGAA
GAGTGGAGACCCGTGTTTGAAGAAGCCGGTTGGAGATGTGAGGGGGAGCATCCTATCAGCGTACCTGCTGATAGCGT
AATGTTCCACTTGGTACCTAAT

Seq. ID No: 33

>ATMT_6
ATGACAAATGATCCTTCACCGAGGGACGCGCGTGACGAGCTACCTGTACGTGACGAATTGAGAGGTCAATCACCCTA
CGGCGCTCCTCAGCTAGATGTTCCAGTTCGTTTGAATACGAACGAAAACCCGTACCCACTACCCGAAGCTCTGGTGG
AGCGTATTGCTGAGAGAGTGCGTGAGGCCGCGAGAAGTTTAAACAGGTACCCTGACAGGGACGCCGTTGAGCTTAGA
ACAGAATTGGCTAGATACCTAACTCGTACCGCGGGTCATGAAGTTACAGCAGCACATGTCTGGGCCGCCAACGGGTC
TAATGAAGTTCTGCAGCAACTACTGCAGACGTTCGGCGGACCGGGCAGGACAGCTATCGGCTTTGAACCTTCCTATA
GCATGCATGCGCTGATATCAAGATCTACCGGGACGGGGTGGATCTCCGGACCCAGGAACGACGATTTCACGATAGAT
GTAGATGCAGCTCGTGCTGCCATCGCCGAGCATAGACCAGAAGTCGTATTTATTACCAGTCCTAATAACCCAACCGG
GACGGCGGTACGTGCTGAAACTGTACTTGCTCTGTATGAGGCGGCTCAAGCAGCAAGGCCTTCCATTGTTGTTGTGG
ATGAAGCGTACGGCGAATTTTCACATCATCCGTCTCTGCTACCGCTGATAGAGGGTCGTCGTCACCTGGTATTGAGT
AGGACGATGTCAAAGGCCTTTGGGGCGGCTGGACTGAGACTGGGCTACCTAGCCGCAGATCCTGCGGTCGTCGATGC
AGTTCAATTGGTAAGGTTACCCTACCACCTATCTAGCGTCACACAGGCTACCGCTCTAGCAGCATTAGAACATACTG
ATACACTGTTAGGCTATGTGGCTCAGCTAAAAGGGGAACGTGATAGATTAGTCGCAGAGCTTAGAGCGATCGGATAT
GAGGTGACCGAGAGTGATGCGAACTTTGTGCAGTTTGGGAGATTTGACGACAGTCACGCTGTCTGGAGGCAAATTTT
GGACAGGGGAGTTTTAGTTAGAGATAACGGGGTACCCGGGTGGCTGAGGGTAACTGCTGGGACACCAGAAGAAAATG
ATGCTTTTCTGGACGCAGTCCGTGAGCTGAAGAAAGAGCACGACGCAGGAGGAGGATCTGGAGAGGGTAGAGGCTCC
CTATTGACATGCGGCGACGTCGAGGAAAATCCGGGCCCATCCACTGAGGTTAGCGAAGCCCAGGCAAGGAGGGCGGT
AGCGGACATTTTTAATAGCACTTTGGCGTCTTCCGCTATTGGCGCCGCTTGGGAGCTAGGGGCACTGGACGAACTGA
GAGAAAATGGGAAACTAGACGTGAGCGATTTTGCGGTTCGTCACGATCTGCACGAACCTGCTGTAGTAGGGATGTTT
ACGGCGCTGGCAAGCGTTGGCATTGTGCGTCGTGAAGGCGCCACCGTTGTTGTAGGTCCTTATTTCGACGAGGCAAA
TCACCACCGTTCCTTATTTCATTGGCTTAACCAGGGCTCTGGAGAACTGTTCAGAAGGATGCCACAAGTGCTGCCCA
ATGAGAATCGTACTGGAAAATTCTACCAACGTGACGCCGGAGCCATCAGTTACGCCTGTCGTGAGATCTCTGAACGT
TATTTTGACCCAGCATTCTGGGCGGCAGTAGATGGTCTTGGCTACACCCCAACTACAGTCGCAGACCTTGGCTCAGG
CTCAGGGGAGAGGCTGATTCAGATTGCGAGAAGATTCCCGGGGGTGAGGGGTTTGGGGGTAGATATCGCCGATGGAG
CCATCGCTATGGCCGAAAAAGAGGTAGCCGCTAAAGGGTTCGGTGATCAGATCTCATTTGTCAGAGGCGATGCGCGT
ACCATCGACCAAGTATCTGCCAGAGGTGAATTTGCCGAAGTGGATTTGTTAACATGCTTTATGATGGGGCATGACTT
TTGGCCCCGTGAGAATTGCGTACAGACCTTACGCAAATTGCGTGCAGCGTTCCCGAATGTGAGGAGGTTTTTACTGG
GAGACGCAACAAGGACGGTCGGGATACCGGATAGAGAGCTGCCGGTTTTTACTCTGGGTTTTGAATTTGGTCATGAC
ATGATGGGGGTCTACCTACCCACCTTGGACGAATGGGATGGAGTATTTGAGGAGGGAGGCTGGCGTTGTGTAAAGAA
GCATGCAATCGATAGCTTATCTGTGAGTGTTGTTTTCGAACTAGAG

Seq. ID No: 34
>BH4reg_1
ATGGCAGCGTCAGGGGAAGCCCGTAGGGTGCTAGTATATGGTGGTAGGGGAGCCCTGGGTTCCCGTTGCGTACAGGC
TTTTAGAGCGAGAAACTGGTGGGTGGCTAGTATCGATGTGGTCGAGAACGAAGAAGCTAGTGCTTCAGTGATAGTTA
AGATGACAGATTCATTTACTGAGCAGGCTGATCAGGTTACTGCTGAAGTTGGGAAACTTTTAGGGGACCAGAAGGTA
GACGCAATCCTTTGCGTCGCAGGTGGGTGGGCTGGCGGGAATGCGAAAAGCAAATCTCTGTTTAAAAATTGCGATTT
AATGTGGAAGCAATCAATCTGGACCAGCACCATTTCTTCTCATCTAGCTACTAAACACCTAAAAGAAGGCGGGCTGT
TAACATTGGCAGGAGCAAAAGCCGCCCTTGACGGTACCCCAGGAATGATTGGATATGGTATGGCTAAAGGAGCTGTT
CACCAGCTATGTCAAAGCCTGGCTGGTAAGAACTCTGGAATGCCGTCCGGCGCAGCGGCGATAGCAGTATTGCCCGT
AACCCTGGATACTCCTATGAATAGGAAGTCCATGCCTGAAGCCGATTTCTCTAGTTGGACTCCGCTGGAATTTCTAG
TAGAAACTTTCCATGATTGGATAACCGGAAATAAGCGTCCTAATAGTGGTAGTCTAATACAAGTCGTCACCACGGAT
GGGAAAACCGAGCTGACCCCAGCGTACTTT

Seq. ID No: 35
>BH4reg_2
ATGACCGCTCTTACCCAAGCTCATTGTGAAGCCTGCCGTGCGGATGCACCTCACGTTAGCGATGAGGAGCTTCCAGT
TTTGTTAAGGCAAATACCTGATTGGAATATAGAGGTCCGTGATGGGATAATGCAACTTGAGAAGGTGTATCTTTTTA
AGAATTTCAAACATGCCTTAGCCTTCACGAATGCAGTGGGGGAAATATCAGAGGCCGAGGGGCACCACCCGGGCCTT
CTTACTGAATGGGGTAAAGTTACTGTGACTTGGTGGAGTCACTCTATCAAAGGTCTGCACAGGAATGATTTTATTAT
GGCAGCCAGAACGGATGAGGTCGCGAAGACAGCAGAAGGGCGTAAA

Seq. ID No: 36
>BH4syn_1
ATGGAAGGGGGTCGTCTTGGCTGTGCTGTATGCGTACTAACTGGGGCCAGCAGAGGATTCGGCAGGGCCTTGGCTCC
ACAATTAGCTGGACTACTGTCCCCGGGATCTGTGCTGTTGCTTTCTGCGCGTTCCGACTCCATGCTAAGACAGTTGA
AAGAGGAGTTATGTACCCAGCAACCCGGTTTACAGGTAGTTTTGGCTGCCGCCGACCTTGGTACTGAGTCCGGCGTG
CAACAACTACTTTCTGCAGTCAGAGAATTACCCCGTCCAGAAAGACTTCAGAGGCTGCTGTTGATAAACAATGCGGG
CACACTGGGTGATGTGTCAAAGGGATTTCTAAATATCAATGATCTAGCAGAGGTGAATAACTATTGGGCCTTAAATT
TAACCAGCATGCTTTGTTTGACGACAGGGACATTAAACGCATTTTCTAACAGCCCCGGACTTTCAAAGACAGTAGTC
AATATTTCTTCACTGTGTGCGCTGCAGCCCTTCAAGGGGTGGGGACTGTACTGTGCTGGTAAAGCAGCTCGTGACAT
GCTGTACCAGGTGCTGGCCGTTGAAGAACCAAGTGTTAGAGTTTTAAGCTACGCCCCTGGACCACTAGACACGAACA
TGCAACAGTTGGCCCGTGAAACCTCTATGGACCCAGAGCTGAGATCCAGGCTGCAAAAACTTAACTCTGAGGGAGAG
CTAGTTGATTGTGGAACTAGCGCACAAAAACTATTGTCCTTGTTACAAAGGGATACCTTCCAAAGTGGAGCGCACGT
CGATTTCTACGACATT

Seq. ID No: 37
>BH4syn_2
ATGCACTCCCCGAGCCTGTCAGCGGAAGAAAATCTTAAAGTCTTTGGAAAATGTAATAACCCTAACGGCCACGGTCA
CAACTACAAAGTTGTGGTAACTATACACGGCGAGGAGGCAATCATGAAGCCTTTGGACCATAAAAATTTGGACCTGG
ACGTACCTTATTTCGCTGACGTAGTCTCTACAACCGAAAATGTAGCAGTATACATCTGGGAAAATTTGCAGAGGCTA
CTACCGGTGGGAGCACTTTACAAAGTTAAAGTATATGAGACAGACAATAATATCGTAGTCTATAAGGGTGAG

Seq. ID No: 38
>BH4syn_3
ATGCCATCATTGTCCAAGGAGGCTGCCTTGGTTCACGAGGCTTTGGTCGCAAGGGGGTTGGAAACGCCGTTGCGTCC
TCCGGTTCACGAGATGGACAACGAGACTAGGAAATCACTAATCGCTGGTCACATGACTGAAATCATGCAATTGCTTA
ATCTTGACCTTGCTGATGATTCACTGATGGAGACTCCACACAGAATTGCGAAGATGTATGTGGACGAAATATTTTCC
GGTCTTGACTATGCCAACTTCCCGAAAATTACCTTGATTGAAAACAAAATGAAAGTGGACGAGATGGTGACCGTTAG
AGACATCACCCTTACCAGCACTTGCGAACACCACTTTGTGACTATTGATGGAAAAGCCACCGTGGCGTATATCCCAA
AGGACAGTGTCATCGGGTTGTCAAAAATCAATAGAATAGTCCAGTTCTTTGCACAAAGGCCACAAGTTCAGGAACGT
CTTACTCAACAAATCTTAATCGCCCTTCAAACCCTATTGGGGACCAATAATGTCGCTGTGAGCATCGATGCGGTACA
CTATTGTGTTAAGGCCCGTGGTATTCGTGATGCGACGAGCGCCACCACGACGACCTCACTGGGAGGACTGTTTAAGA
GTTCACAAAATACTCGTCACGAGTTTTTGAGAGCCGTTAGGCACCATAAC

Seq. ID No: 39
>BH4syn_4
ATGCACCACCATCATCATCATACGTCCAGTACACCTGTAAGAACCGCTTATGTCACGCGTATAGAACATTTTAGCGC
GGCCCATAGGTTGAATAGTGTCCACTTGAGTCCAGCTGAGAACGTAAAACTGTTTGGCAAGTGCAATCATACCTCCG
GGCATGGTCATAACTACAAAGTAGAAGTAACGATTAAAGGACAGATCAACCCACAGAGCGGGATGGTGATCAACATA
ACCGACTTAAAGAAGACCTTGCAGGTGGCCGTTATGGACCCGTGCGACCATAGAAACCTGGACATTGATGTGCCATA
TTTCGAGTCCCGTCCATCAACCACTGAGAACTTAGCAGTTTTTCTCTGGGAAAACATAAAGTCCCACTTACCTCCTA
GTGATGCCTACGACCTATACGAAATTAAATTACATGAAACCGACAAAAACGTGGTGGTCTACAGAGGCGAG

Seq. ID No: 40
>BH4syn_5
ATGCATCATCACCATCACCATAGCTCCAAGGAGCATCACTTGGTCATAATCAACGGTGTCAATAGAGGGTTTGGGCA
TTCAGTTGCCTTAGACTACATTAGGCATTCTGGGGCGCATGCCGTCAGCTTCGTGCTTGTAGGGCGTACGCAACACT
CTTTAGAACAAGTGCTAACTGAGTTGCATGAAGCAGCATCCCATGCTGGAGTGGTCTTCAAGGGCGTCGTTGTTAGT
GAGGTCGATTTAGCACATTTAAATAGCTTAGACTCTAATCTAGCCCGTATTCAGTCTGCCGCAGCAGACCTAAGGGA
TGAGGCGGCGCAGAGCACTAGAACAATAACGAAGAGCGTGTTGTTTAACAACGCCGGTTCACTAGGGGACCTGTCAA
AGACTGTAAAAGAGTTCACCTGGCAAGAGGCAAGGTCATATCTGGACTTCAACGTGGTCTCACTTGTGGGACTATGC

TCTATGTTTCTTAAAGACACGTTAGAAGCGTTTCCAAAAGAACAATATCCCGACCACAGAACTGTCGTTGTCTCCAT
TTCTTCCCTGTTGGCTGTCCAGGCGTTCCCAAATTGGGGGCTATACGCTGCTGGGAAAGCGGCCCGTGATAGATTGT
TAGGCGTCATAGCGTTAGAGGAAGCCGCCAATAATGTTAAGACTTTAAATTACGCGCCGGGTCCCTTAGACAATGAA
ATGCAGGCCGATGTCAGACGTACTCTTGGCGATAAAGAGCAGTTAAAAATTTACGATGACATGCATAAATCCGGCTC
TCTAGTTAAGATGGAAGACTCAAGTAGAAAACTTATTCATTTACTAAAAGCTGATACGTTCACGTCTGGCGGGCATA
TAGACTTTTACGACGAG

Seq. ID No: 41
>DAC_1
ATGGTTGATGCAGACATAGCCCTGAATTGGGCGGGAGGACTTCATGTATGTATCGTGAGACCTCCGGGACATCATGC
AGAGCCCGGCGCTGCATGCGGTTTTTGTTTCTTCAACAATGTCGCTCTAGCTGCGAGATACGCTCAGAGCTTGCAGA
GTCCTTCCGACCCACCCCTGAGAGTCATGATTTTAGATTGGGATATTCATCACGGTAATGGCACCCAACATATCTTC
CAGGATGACGCTAGTGTGCTATATGTATCACTTCATCGTTATGACGACGGCACATTTTTCCCGTCCTCAGAAGACGC
TGCGCATGATAAAGTTGGAAGCGGACCGGGCGAGGGTTTTAATGTAAACATCCCCTGGAATGGGGGTAAAATGGGGG
ACGTGGAGTACTTGCTGGCTTTTCACAGAATAGTGATGCCTATTGCTTACGAATTTAACCCTCAACTGGTATTGGTG
TCCGCGGGCTTTGACGCGGCTAGAGGTGATCCTTTAGGGGGCTGCCGTGTCTCACCTGAAGGGTATGCACACATGAC
CCATTTGCTAATGGGTTTGGCGGGAGGAAAAGTAGTCGTTGTACTTGAGGGTGGGTACAATCTAACATCAATAAGCG
AGTCAATGAGCATGTGTACGAGAACACTATTGGGAGATCCTTTGCCTTTCATATCAGATTTGCATGCCCCTAGACCA
GCCGCGTTGAGAGCGATCTCTAGTGTCCTGGGGGTCCATCAGAAGTATTGGAGAAGTTTGTGCATAAACGTAGGGCC
GCCA

Seq. ID No: 42
>DAC_2
ATGAAAACACATCCACACCCGGAACGTCCAGATCGTTTGCAGGCCATAGCAGCGTCCTTGGCCACTGCCGGCATATT
TCCCGGAAGATGCTACCCGATCCCGGCGAGAGAGATTACCAAAGAAGAATTGCAGATGGTTCACTCACTGGAGCACA
TTGAAACTGTCGAGCTTACGGGCCAAATTCTATATTCTTATTTTACTCCTGATGGCACGAATCCACATAATAGGCTG
AAGCTGGATAACAGGAAGTTGGCGGGCATCCTATCACAGAGGATGTTTGTAATTCTACCATGTGGAGGACTGGGAGT
TGATAGCGATACTATCTGGAATGACCTTCATTCCTCCAACGCTGCCAGATGGGCCGCCGGTTCAGTTATTGACTTAG
CTTTCAAAGTGGTCACCAGAGAATTGAAAAACGGCTTCGCTCTAGTTCGTCCGCCCGGGCATCATGCAGATCCATCC
ACCGCAATGGGGTTTTGCTTCTTCAACTCAGTGGCTATTGCAGCTAAACAGCTTCAGCAGAAACTAAACGTGAGGAA
AATTCTGATTGTGGATTGGGACGTGCACCACGGGAATGGTACGCAGCGTGTATTTTATCGTGACCCTAACGTGCTTT
ACATATCCTTACATAGGCATGACGACGGAAACTTCTTCCCAGGGTCAGGAGCAGCAGATGAAGTAGGTGCCAATAGC
GGGGAAGGCTTCAACGTTAATGTTGCATGGGCAGGCGGCCTAGACCCGCCTATGGGTGATGCGGAATATCTAGCAGC
GTTTAGGACAGTTGTCATGCCGATAGCCCACGAGTTCGCCCCCGATGTAGTACTTGTGAGTGCGGGGTTCGACGCTG
CGGAGGGT

Seq. ID No: 43
>DAC_3
ATGATGGCAACAGAACCGATAGCGAGTGGGAGCGGCACAATGGACATCGACAGCGAGAAGACGCCAAGCACGAGCCA
GGCCAATCCAATGGCAGATACCTTTCAGACCCGTGAAGCGGTACTAGGATTAGGTGAAGTAGTGGAGCATGTAGGCG
GGCGTTGGGTGGCGGAGCAAGAGTGGATTAGAAGTCCTGAGAGGAAAATGGCATACACTCAAGGTACTAAAAAGAAG
GTGTGCTATTACTACGATGGAGACGTCGGCAATTATTACTATGGACAGGGTCATCCTATGAAGCCCCACAGGATCAG
GATGACCCATAACCTGCTGTTAAATTACGGATTATACAGAAAAATGGAGATTTACCGTCCCCACAAGGCCAACGCTG
AGGAGATGACTAAATATCACTCTGACGACTACATAAAGTTCCTAAGAAGTATTCGTCCAGATAATATGTCAGAGTAT
AGCAAGCAGATGCAGCGTTTTAACGTTGGTGAAGACTGCCCAGTATTCGACGGTCTATTTGAGTTTTGTCAGCTGTC
CACGGGCGGGAGTGTAGCGTCAAGCGTCAAATTGAACAAACAGCAAACGGACATAGCCGTGAACTGGGCCGGCGGAT
TACACCACGCGAAGAAGTCAGAAGCGTCTGGATTCTGTTACGTAAACGACATCGTACTAGCGATCCTTGAGCTTCTA
AAGTATCACCAGAGAGTTTTATATATTGATATAGACATCCATCACGGCGATGGGGTGGAAGAAGCGTTTTATACGAC
GGATCGTGTAATGACGGTATCATTTCACAAATATGGGGAGTATTTCCCAGGAACAGGTGACTTGAGAGACATTGGTG
CGGGTAAGGGTAAATACTACGCGGTCAACTATCCCCTTAGAGACGGCATAGACGACGAGTCC

Seq. ID No: 44
>DAC_4
ATGGACGCGGGTACGCGTAGAGTGGATGATGCGGCCGTTCCAAGTACGGGTCCATCCGCATCCTTACTAAGATCAGC
GAATATGTTAAGTGCCGCGTTCGGTCTAACAGCATCTCTATACTCCAGATTACGTGGGGTGTGCTCCAGTAGAAGAG
CCCTATCCACGTCAGCTCGTACGTCTGAGGCGGCCGGAGTAGGGGCGAAACCCGGAGTCGCTGCCGCATTAACCGTT
CCCTCAACTGGGCCCTCTGCTTCCGAAGCATCACCAGCTGCTCTATTAAGGATTCAAGTCGCGGAAGAGTGGGCGCG
TGCCAGTGGACTATTGGATAGGGAAGACTGCCAGGTTGGTCTGGCATTCGATGAAGCCATGCATTTACATAGTGGAC
CAGCTGGTCATCCTGAAAGACCAGCCAGGACCAAAGAAATACTGGCGCAGCTACACGCGTCAGGTTTAGTGAGAGCT
TGTGCCCAAGTCCCTAGTAGGGAAGCGACTGAAGAAGAATTACTATTGGTCCATGACGCCAGACATGTAGAGAGGGT
GTTACGTCATGAAGCGGCTGGTCATAAAAAGGCTAAGGCTTTCAGTTTCCCTTTCGGCCCCGATACATATGTTTGCG
AGCATACAGCGAGATGCGCTCGTCTAGCTGTAGGTTGCTTATTGTCTCTTGTTGACGCTTCTTTAGATCCAGCAAGT
CCGGTTCGTACAGGGATGGCGGTAGTTAGACCGCCCGGGCACCACGCAACCAGTGATCGTGCGTCCGGATTCTGCTT
GTTTAATAATGTCGCTGTTGCCGCCAGACACTTGCAGAGACGTCATGGTTTAAAAAGAGTTGCAATTGTTGACTGGG
ATGTTCATCACGGTAACGGTACGAACGATCTGTTTACCGAAGACCCAAATATACTGTTTTTCTCCGTTCACCGTTTT
GATAATCACGGGTTTTTCCCCGGGAGTGGTTTTCTTGAGGATGTTGGTCATGCGCAAGCTCGTGGTTACACGGTGAA
CGTCCCCCTGGAGAAAGGCTATGGGGATCTTGATATCGTTCATGTCGTAAAATACGTGCTTTGCCCCGTCCTAGAGA
GATTCAAGCCCGATGCGATCCTGGTCTCCGCTGGGTTCGATGCCGTAAAGGGCGATCCACTAGGCGAGTGCAGAGTG
TCTCCGGAGGCTTTTGGCTGGATGACGCGTTGTCTACATCGTTTAGCTCAGCGTTATTGCGACGGCAGGTTGTTCCT
GGTACTAGAGGGCGGGTATAATCCTGACATGATAGCACAGTGTTGCATCGAATGCGTTCAGTCCCTAGTGGCAGAGG
CTGCTGGATTAAGGGGCCCTTGGCCCGAGTTCCCTGCTGTGGGAGTTCCGTTGGCAGAAGGAGCGCAGCTGAGTGCA
CCCAGCTCAGCCCCTACCTCTGCCCCCGGGACGCCAACTTCAACAAGCCCAGCCAGCTCCCCAGCCCTGTCCGCCGC
AGCACCACCTTTAGCTTCCCCCGGTAGCACACCAACTTCCTCTCCTTGTTTAAGGCCAAGTGGGGGTGAAGCACCGC
CTAGATCACCTCCGTCTGCCAGCGCGTCAGCGGGAGGAGGCGCAAGGCAAAGGGCGCGTGCTCCAAGTTCTAAGACA
GTCAGGGCGGTAAGACAGCTGACAGAGATACATCACCTGCTACCTTTGGAGCTGCCTGTTGCGCCGCGTCCAGGGGA
CGGTCCCGGCGCTGCAAATAAGTCCGCCAGAAAGAATGAACGTAGGAGACTTGGGAGAGGTAGGAGGGGACCGGAAG
AGGAGGGAGCCAGTAGTGACAGTTCAGGCTGGGCCATCGCTTGTGGTTTAAGTGATGCTGAACCCTGGCCGAGTCCA
CAGGCGTCCCCTGTGGCTTCGCTTTCCCAAGGAGCTTCTTCCTTACCCACGTTAGAATTACCGCCCGCGTTTCCAAG
TTTAGACGGTGTTGGAAGTACGGCCGGGAACTCTTACCTGGGTACGTCAGGGAACGTGGGAATCGACGCGGCCGGCC
ATAGTGCATCTTCATGGCTAGGTTCTCCGACTACAGCAGCGACCGCAGTCGCCCCTCCCGCGAGGGGCGACAGGAAG
GTGAAACGTCGT

Seq. ID No: 45
>DAC_5
ATGGTGGATGCCGACATCGCGCTTAATTGGGCGGGAGGACTGCATGGTCATCACGCAGGGCGTGGATGTAGCGAAGG
ATTCTGTTTGTTGAACAACGTAGCAGTGGCCGCTGCCTACGCACGTAGTGCACATCCCGAACAGGTCAGAAGAGTTT
TGGTGCTAGACTGGGACGTTCATCATGGCCAGGGCACTCAAGAAATCTTTTGGAGGGACCCCGGAGTGTTGTATGCG
AGCGTCCATAGGGATGGTGGGGAGGGATTCTATCCGGGAACCGGGGCCGCTGAGCAAGTTGGAGACGGAGCGGGGAG
GGGCTTTACCGTGAACGTACCTCTTCCGACGGGCTATGGTGACGGCTGTCTTTGGGCTGCTTGCGCCGAAGTTTTGT
TACCAGCGGCAAGAAGATTCAGGCCGGATCTTATACTTGTATCCGCAGGATTCGACGCGGTTGCAGGAGATCCTTTG
GGCGGGTGCCGTTGTACTGCCCGTGGCTTCGGAGCACTGACGGGTGAATTAAGAAAGCTGGCTGGGGAACTGTGTAG
CGGTAGACTGCTTCTAGCCTTGGAAGGAGGCTACGATTTAAGGACGTTGAGAGCTTGCGTTGGGGAAGTGTGCCAAG
CACTAGCAGCTCCGGAACCGGCAGAAGGTGGAGCC

Seq. ID No: 46
>DAC_6
ATGAGGAATAGGTCAAGTGGGTTCTGTCTTGTAAACAATGTCGCTGTGGCAGCCGAATATGCAAGAGATCGTTACCC
TGAAGTGGAGCGTGTTCTAATCTTCGACTGGGACGTTCATCACGGCCAGGGAACGCAGCAGATCTTCGAGCAGTCCC
CAGATGTACTGGTAATCTCAGTGCACAGGCACGATGGTCACAGCTTCTATCCTGCAACTGGAAGTGCAGGGGAAGTA
GGCTCTGGGCCGGGAAGAGGGTATTCTGTCAATGTGGCTCTTCCTGCAGGTTATGGCGGGGCTGCACTTTGGACAGC
TTGTGCCCATGTCCTTCTGCCTGCGGCAAGAAATTTTCAACCCCAACTTATTCTAGTCTCCGCTGGCTTCGATGCAG
CGGCGAGTGATCCACTAGGGGGATGTTTCGTCGAACCCCGTGTCTTTGGTGCATTAACAGCGGAATTAAGGAGGCTG

GCTGCAGAGGTTGCGGAGGGGAGACTTATCCTGGCCTTAGAGGGTGGCTACAACCCGGAGGTGCTTGCAGATTGTGT
TGATGAGGTCGCCGCCGCGTTGGTTGCAGATGCCTCAAGTTCCGGTGTAGAAGCATTCGCCGAGGCCCCATCATGGC
TCGCAGGGTCTGCATGCTTCGGGGCGATAAGAAGGACTTGTGAGGCACATAGGATGGCTCCCTTGAGATTACCTCTT
CCGTCATCAAGAATCGACAGAAGAAGAGCTGCTGCTCGTCAGGCGGAAGCGTTGTCTAGCCCGTCTAGCAGAGACGC
TGGAGATACGGGAGGCGGCGAGGTATCTGCACACGGCGCCACAACAACGGTTACCACATCAGCAAATTTGGGAGCTG
GTACTCTAGCGGCGAGGCCATCCAGTATGGTGACAGGCGAGGGCAGGAGAGCCAATGGCCAGCTAGTGGACGTTCTG
GGTATCGCACTTGCTGGTAAGCCTTCAGCATCTCCGTGGCCGGAGGCTCAGCGTACACAGGGATCTGCCCCCGGTAC
CCCGGCGCCCGCGACAGGGGGCGCGTTACCACCAGCTGAAACGGCCGAATCACCTGGGTCCGTCGCAAGCGGAGCCG
CGGTTGCCCAGGGTCCAGTAGAATGTCAAGCGGCGGCACGTCAAGCTGGAGAGTGCCCGGGGCAAGCCCCGGCTCCG
GCAGGAGCGGGAGCAGCGCCTGGAGGAAGGGGTGTCGAAGCGGCAGCCGCGCAGCACGGACAGGATTTGGCTCCAGC
GGCAGGACCCGGAGCCGCCGCATTAGTGGAACTTCAAACGGGGGAGTTGGTCGTAAGAATCGCACCCCTTCCAAGGC
CAAAAGACGTTGTGGTTTCCGCAGAGGAGCTGTGGGTTTGGCATGATCAAGGTGGGCCGCTGGGCGTGCAGAGGTGG
AGGTTTGAAGGCGTACGTGCCGAGAACTCCGGAGCACTCAGATGTGCCGAATTTAGGTCCAAAAGGCACGAGTTGAC
CGTAAGACTAAGGTTGGGT

Seq. ID No: 47
>DAC_7
ATGGTGAGAAGTAGTCAGGCTACAACATTTAGTTCTTCCCCCTACTTTGCGGATAGGGGGATTCACGGTACCGCACA
GCCGATCACTGAAAGAATAAACCCGCGTAAGTGTACTTTTCACATATTGCCACCAGGCCTTGGGGGATGGTTGTATC
TTTTTCCAAAAATGAGTAAAACAGTTGCTTATTTTTATGATCCGGACGTCGGCAACTTTCACTACGGCACAGGACAT
CCGATGAAGCCGCACAGGCTGGCGCTTACACATAGCTTGGTTCTGCACTATGGGTTATACAAGAAGATGATTGTCTT
CAAACCATACCAGGCAAGTCAGCATGACATGTGCAGATTCCACTCAGAAGATTATATTGACTTCCTGCAACGTGTAA
GCCCAAATAATATGCAGGGATTCACGAAGAGCTTAAACGCGTTCAACGTAGGAGACGACTGCCCTGTGTTTCCGGGG
TTGTTTGAGTTTTGCAGTCGTTATACCGGCGCATCTTTGCAGGGAGCCACGCAACTAAACAACAAAATTTGCGATAT
AGCGATTAACTGGGCAGGGGGCCTGCATCATGCGAAGAAGTTTGAGGCCTCAGGTTTCTGTTATGTGAATGATATAG
TCATCGGGATTCTTGAGTTGTTGAAATACCACCCTAGAGTCCTATACATTGATATTGACATTCATCATGGAGACGGT
GTTCAAGAGGCATTCTACCTAACTGACAGAGTAATGACCGTAAGTTTCCACAAATACGGAAATTACTTTTTCCCAGG
AACAGGAGATATGTACGAGGTCGGCGCAGAATCAGGGAGATACTACTGTCTTAATGTTCCTCTTAGGGATGGAATTG
ATGATCAGAGCTATCGTCACTTGTTTCAGCCAGTAATCAAACAGGTCGTTGATTTCTATCAGCCAACTTGCATTGTT
CTTCAGTGCGGGGCGGATTCTCTGGGATGTGACCGTTTGGGTTGCTTCAACTTATCCATACGTGGGCATGGGGATTG
TGTACAATATGTTAAATCTTTCAATATACCACTGCTAGTATTGGGCGGTGGTGGGTACACAGTTAGGAACGTCGCCC
GTTGCTGGACATACGAGACGTCACTTCTAGTGGACGAAACGATATCTGAGGAGCTACCATACAGTGAGTACTTCGAG
TATTTTGCTCCGGATTTTACGCTACATCCAGACGTTAGTACAAGGATTGAGAACCAGAATACGCGTCAATATCTTGA
CCAAATTAGGCAAACAATCTTCGAGAATCTGAAAATGTTGAATCACGCCCCCTCCGTTCAGATCCATGACGTGCCGA
GCGATTTGCTTTCCTACGACAGGACAGATGAGCCGGACCCCGAAGAACGTGGGGCCGAGGATAATTATACGAGACCA
GAGGCATCTAACGAGTTTTACGATGGAGACCATGATAATGATAAGGAAAGTGATGTGGAGATC

Seq. ID No: 48
>DAC_8
ATGTGGGATGTTCATCATGGTCAGGGGATACAGTACATTTTTGAAGATGACCCTAGCGTGCTATATTTTAGTTGGCA
TAGGTATGAGCACGGTTCCTTCTGGCCCAATCTTTCAGAAAGCGATTACGATAGCATAGGGAAGGGTAGAGGGACGG
GTTTTAACATCAACCTACCCTGGAATCAGACTGGGATGGGTAATGCGGATTACGTGGCCGCCTTTTTCCATGTATTG
TTGCCACTTGCCTTTGAGTTCAACCCAGAGTTAGTCCTTGTTTCTGCCGGATATGATTCTGGGATCGGTGACCCCGA
GGGACATATGAGAGCTACACCTGAGTGTTTTTCTCACCTAACTCACATGTTGATGCACCTTGCAGGGGGCAAGCTGT
GCATGATATTAGAAGGTGGCTACCACCTACGTAGTTTGTCCGAGAGCGTAAGTATGACCGTAAGGACTTTGCTACGT
GATCCAGTGCCCAGGCTTTCAGGTGAAATGACGCCCTGCTATTCAGCCCTTGAGTCTATACAGAATACGCGTCATGC
TCACTCACCATATTGGAAATGCCTTCTTCACGACGAAACTAGACTGGTGGAAGAGATTTCTACGAAAGGACTGAAGG
CACCGGGGCCGTTGCATGTTGATGCAAGCGTGGTTGATGAATTTCTTGAAAACCACATGAAGAAGATCTTGCACCCT
ACTCCCCCAATTACAACTATGGTCGTCGCAAGCGTGGAGAATACGCTTAACTTACCCGCGGGTGTTCAGTTAGAGGA
GTCCACTGTCACGCCAGAGCAGGCCAGACATGCAATCTCCGTATTCAACCCAGATGAGTTAAATGAGAACGTGCTTA
ACTCAGTTTCAAAGATGCTTCCCGCCCTAGAGAAACTCGTCAAT

Seq. ID No: 49
>DAC_9
ATGAAAACCCATCCCCATCCTGAACGTCCGGATCGTCTTCAAGCGATTGCCGCAAGTTTGGCTACAGCAGGAATATT
CCCAGGGAGATGTTACCCTATCCCGGCAAGAGAGATCACTAAGGAAGAGTTACAGATGGTACATAGTTTGGAGCATA
TAGAGACGGTCGAATTGACAGGTCAGATTCTATATAGCTATTTCACGCCGGACACCTATGCGAATCAACACTCAGCT
CACGCTGCGAGGCTTGCGGCGGGTTTGTGCGCTGACCTAGCTAAAGAGGTCTTTAGTGGCAGGGCTAAGAACGGCTT
CGCATTGGTTAGACCGCCAGGGCACCATGCTGGTGTCAGGCAGGCAATGGGATTTTGCCTACATAACAATGCGGCGG
TAGCGGCGCTTGCCGCCCAAGTGGCGGGAGCCAAGAAGATCTTGATAGTAGACTGGGACGTTCATCATGGGAACGGC
ACACAGGAAATATTCGAGCAAAATAAATCTGTTTTATACATATCGTTACACAGGCACGAGGGCGGAAAGTTCTACCC
GGGTACAGGGGCGGCGCACGAATTTGGTACTATGGGAGCCGAGGGCTATTGCGTGAATATTCCTTGGAGTCGTGGTG
GCGTCGGCGACAATGATTATATATTTGCGTTCCAGCATGTTGTCTTGCCTATAGCGTCCGATTTCGCTCCTGATTTC
ACGATAATATCCGCCGGCTTTGACGCTGCAAGGGGTGATCCACTAGGGTGCTGCGACGTGACTCCCGCTGGTTACGC
TCAGATGACGCATATGTTAAACATTTTAAGCGGGGGCAAACTACTTGTAATTTTGGAGGGTGGATACAACCTACGTT
CTATCTCAAGTTCAGCAACAGCAGTAATCAAGGTGTTGCTGGGTGAATCTCCCGGCTGCAATCCCAAAAATTTCCTT
CCGTCCAAGGCGGGTGTTCAAACTGTCCTAGAGGTCTTGAAGATACAAATGAACTTCTGGCCGGCTCTAGGCTCAAT
TTATAGTGACCTACAAACGCAATGGGGAATGTACTGCATGAAGACCAAAAAGAAGCAAATAAAGAAAAGACAACGTG
CGGCGGCGCCGCTTTGGTGGAAGTGGGGGCAGAAAAGTTTCTTATATCACTTGCTTAATGGTCATCTGCATGTAAAA
AGCAAGGGCTGC

Seq. ID No: 50
>DAC_10
ATGTTTCTAGTCAGACACCACTTATATAAATGGAAATCACATTTTCGTATCGATGCCGACGGTAAATTCGTCGAAGA
TCAGTTCTTTCCAAAGAACTTGAAAAGCGGCAGACGTTTCTTACGTAGCATAGGCGCTTCTATCACGTGCTCTAATG
GCATTGGAAAAGATCCATATATTTTGAGCAATGAGAAAATTTCAGACGCTAGGCTAATATACGCAGTCGCGCCAGCG
ATGGGTCATAATCAGGAAAGCCACCCTGAGAGTCACTTTAGGGTTCCTGCCATAGTCAACGCCTTAGAGAAGATGGA
AATGACGCCTAAATTCAGGGGTTCCGAGATTATTGAGTTGCAATCTTTTAAGCCCGCTTTAGTTGATGATATAGCGT
CTGTCCATGCTAGAGCATATGTATCCGGTTTGGAAAAAGCGATGGACCAGGCGAGCCAACAGGGTATAATATTTATT
GATGGCTCAGGTCCAACATACGCAACGGCTACGACATTTCACGAGTCTCTTGTGGCGGCAGGGGCGGGGATAGCATT
GGTCGATTCTGTGGTTGCGGCTTCTAAAAACCATTTAGACCCGCCAGTGGGTTTTGCGCTAATAAGGCCCCCGGGCC
ACCACGCCATACCACTGGGCCCGATGGGGTTCTGTGTGTTTGGGAATGTCGCTATTGCCGCAAGATATGCTCAGCGT
GCCCACGGGCTAAAAAGGGTTTTTATAATTGATTTCGATGTGCACCACGGCAACGGAACCAATGATGCGTTTTATGA
TGACCCAGATATATTTTTCCTATCCACTCACCAAGACGGGTCTTATCCCGGCACGGGGAAGGTGGATGAAGTAGGAA
GGGGCGATGGAGAGGGTACCACACTGAACCTGCCATTGCCCGGGGGTAGTGGCGACATAGCTATGACAACAGTATTC
GACGAAGTCATTGCTCCATGCGCTCAACGTTTTAAACCCGACATAATTTTGGTTTCTGCAGGCTATGATGGGCATGT
ATTAGATCCATTAGCAAGCCTTCAATTTACAACTGGCACGTACTACATGTTAGCCAGTAAAATTAAGTTGCTGGCGA
AAGATTTATGCGGCGGTAGGTGTGTATTCTTTCTGGAGGGCGGTTATAACCTGGAGTCTCTATCCTATAGCGTAGCG
GATAGCTTCAGGGCATTTTTAGGAGAACAGTCCTTGGCATCAGAGTTTGATAACCCAGCGATCTTGTACGAAGAGCC
ATCTACGAAAGTAAAACAAGCTATCCAGAGGGTTAAGCATATTCACTCCTTA

Seq. ID No: 51
>DAC_11
ATGGTGGACGCTGACATCGCGTTGAACTGGGCCGGAGGATTGCACCACGCCAAGAAATCTGAGGCGAGTGGGTTTTG
TTATGTGAATGACATCGTCTTAGGAATCCTAGAACTACTGAAAGTACACAGGAGAGTCCTATACGTTGACATAGATG
TGCACCACGGTGACGGCGTGGAAGAGGCATTCTATGCTACAGATAGAGTAATGACTGTTTCTTTTCACAAGTTCGGC
GATTTCTTCCCCGGAACGGGCCATATTAAGGATACCGGGTGGGGCCCCGGAAAGAACTACGCTCTTAATGTCCCGCT
TAACGATGGAATGGACGATGAATCATTTAGGGGACTGTTTAGACCCATCATACAGGGGGTGATGGAAGTATACCAGC
CGGACGCTGTCGTATTACAATGCGGCGCTGATTCTCTATCCGGGGATAGGCTAGGTTGTTTTAACCTTTCAGTCAAA
GGACATGCTGATTGTCTTCGTTTCTTGAGGTCTTTTAACGTCCCACTAATGGTTTTAGGCGGGGGTGGTTATACCAT
CAGGAACGTCGCTAGATGCTGGTGTTACGAGACAGCTGTAGCCGTCGGAGTAGAACCAGATAACAAGCTGCCGTATA
ACGAATACTATGAATATTTCGGCCCCGATTATACCCTTCATATTGAACCATGTAATATGGAAAATCTTAATACCCCC

AAGGATATGGAGAAGATAAGGAACATGCTGCTGGAACAGCTAAGCAGGATACCTCATGTGCCGAGCGTTCCATTCCA
AACCACGCCGCCAACAACCCAGGTCCCCGAAGAGGCCGAGGAGGATATGGACAGAAGGCCAAAATGCAGAATCTGGA
ATGGGGAGGACTACGATTCTGATCCAGACGAAGATGAGAAGCCTAGACATACGGAACCGAATAGCGAGCTTCGTGAT
GTAGTGGACGAAATGGACGAGGACAAACGTGAAGAGCACCCACCCTCA

Seq. ID No: 52
>DAC_12
ATGTGTTCAGACGCAAATGGGAAGGTTGGTAACATTTCTGTGATGAGCACTGAGGGAATATCACAGGTGGAGAGTAA
AAAGGCTCGTCTTAATGGCTTACTTACACTAGAAGATATCTATAATCTGCCAGACGAGCTAGATGATGATGAAGACG
ACAGTGACTGGGAGCCTTTACTTGAACCTCTTGCGGTACGTAAATGGTTTTGCACCAACTGCACGATGGTCAATTTT
GATGGATTTGACTTCTGCGAGACGTGCGAGGAGCATAAGGAAAGTGGCATCCTGAAACAAGGCTTCTTCGCCTCTCC
TGCCTTACAAGGCACAAGAAGCACTCAAATCGAGTCTGAGGTTATCGAGAGATATACTGAGTCTATATGTGACATCT
CAGCTTCTGCGCTTAGTACCGTGGTCGGTTTCGACGAGAGGATGCTACTACATAGCGAAGTGGTCTTAAAACCACAT
CCTCACCCACATCCAGAAAGACCGGATAGACTGAGAGCAATTGCGGCATCACTGTCAACCGCAGGGATCTTCCCTGG
CAAATGCCATCCAATCGCCGCGAGAGAAATCACCCAAGAGGAACTGTTGAAAGTTCACTCACTTGAGCATGTCGAGG
CAGTTGAAGTTACACGTCAGATGCTAAGCTCATACTTTACCCCAGATACGTACGCTAACGAACACAGTGCGCAAGCC
GCGCGTTTGGCGGCCGGCCTTTGTGCAGATCTAGCATCTGAAATCTATTCCGGAAGGGCCAAAAATGGATTCGCCCT
GATTAGGCCCCCTGGACACCACGCGGGTGTACACCAGAGCATGGGTTTCTGTCTTCACAACAACGCGGCGGTTGCCG
CATTAGCCGCACAAGTAGCTGGTGCCAAAAAGGTCCTGATTGTGGATTGGGATGTCCATCACGGTAACGGCACACAA
GAGATTTTTGAAAGGAACAAGAGCGTTTTATATGTCAGTCTGCACAGGCACGAAGCGGGCAAATTCTATCCAGGAAC
CGGAGCCGCCCATGAAGTTGGTACTATGGGTGCTGAAGGTTATTGCGTTAACGTACCTTGGTCCAGGGGGAGGGTTG
GCGACAACGACTACATCTTCGCGTTCCAAAATGTTGTGATCCCAATTGCACACGAGTTTTCCCCAGACTTCATTATC
ATATCTGCCGGTTTTGACGCAGCTCGTGGAGATCCTTTGGGTGGATGCGATGTGACCCCGGCTGGCTATGCCTGTAT
GACTCATATGCTATCTGCGCTCGCTGGAGGGAAAATGTTAGTTATCCTTGAAGGTGGGTATAACTTACGTAGCATCA
GTTCATCCGCCACGGCGGTGATTAAGGTATTACTAGGTGAAAAGCCGAAGTGCCAGTTTGAGAACATAGAGCCTTCC
GCAAGCGGGCTTCAGGCTCTTCTTGAGGTCCTGAAGGTACAGACTAACTTCTGGCCCTGTCTGAGTTCTAAACTGAC
ACAACTACAAAGCTGCTGGGAAGCATACTTGTCTGGTCGTAAAAAGCAGAAGAAACGTAGGTTTCGTACTGTTGCTC
CGCCACCAATATGGTGGGCCTGGGGACGTAAAAGGTTCCTATACTTTTTAAGATGTCAAAGGTTCAGGATGAAACCA

Seq. ID No: 53
>DAC_13
ATGGCTGGAGCAGAAGAGCTGCACGTGTTTTGGGAAGAAGGTATGTTAAAACATGAGACGGGACGTGGTGTTTTTGA
CACTGGGAGCGATCCGGGTTTCCTGGACGTGTTAGAAAAACACCCAGAGAATGCAGATAGAGTTAGGAACATGGTAA
GCATTTTAAAAAGGGGGCCAATCGCCCCCTTCGTGAGTTGGCACCAGGGAAGACCGGCTAGTCTGCCCGAGCTTCTT
AGTTTCCATTCCAGCGAATACATAGAGGAGCTAGAAGAGGCGGACAGGGCGGGAGGTAAGATGATGTGTTGTGGGAC
TTTCTTAAACCCAGGCTCCTGGAACGCTGCACTATTGGCGGCCGGAACTACACTGTCTGCGGTAAAGTATATTTTGG
ATGGCCACGGTAAGATAGCGTATGCCCTTGTCCGTCCCCCTGGTCACCACGCCCAGCCGACCCAGGCCGACGGATAC
TGTTTCTTGAACAACGCTGGCTTGGCAGTGCAGCTTGCCCTTGACGAGGGTTGCCGTAAAGTGGCCGTTATCGACAT
TGATGTCCATTACGGTAATGGCACGGCTGAGGGATTCTACTGTTCTAACAAAGTTTTGACGATTAGCTTACACATGA
ACCACGGTAGCTGGGGTCCTAGTCACCGTCAAAGTGGCACGCATGATGAGCTTGGGGATGGTGATGGGTTTGGATAT
AACATGAACATTCCCTTGCCCAATGGAAGTGGCGATAGAGCATATGAATATGCAATGCAGGAGTTAGTGGTGCCTGC
TGTCCAAAAATTCGGCCCGGACATGATTGTACTGGTAGTCGGTCAAGATTCCAGTGCCTTCGACCCTAATGGGCGTC
AGTGTCTAACTATGGACGGTTACAGACAGGTTGCAAGGATTGTAAGGGGCTTGGCAGATATGCATTGCAAGGGCAAG
TTGTTGGTCGTACAGGAGGGTGGGTATCATATTACGTACGCCGCCTACTGCCTACACGCGACTCTTGAGGGAGCTTT
AAACCTTCCTTCCCCACTATTGAGCGACCCGATAGCCTACTATCCAGAGGATGAAGGATTCGCAGTGAAGGTGATTG
ATGCGATGAAGGAACATTATAAGAGTAACGTCCCTTTTCTGAAAGAGATCAAC

Seq. ID No: 54
>DAC_14
ATGGGCTTTTGTATCTTTGGTAACATAGCGATCGCAGCCCGTTACGCCCAAAGAGTCCACGGATTGAAAAGGGTTTT
CATCATCGACTTCGACGTTCATCACGGGAATGGCACCCAGGATGTCTTTTACGAGGACCCTGACATCTTCTTTTTGT

CCACGCACAAAGAAGGCTCCTATCCAGGAACAGGGAAGATTCACGAAGTAGGGTGCGGTCCCGGTGAGGGAACTACT
TTAAACCTGCCATTGCCCGGAGGGACTGGTGATGTCGCAATGAGAACAGTGTTTGACGAAGTTATAGTCCCTTGCGC
TCAGAGATTTAAACCAGATATTATTTTGGTGTCAGCGGGTTACGACGCGCACTTTCTTGATCCTCTGGCGAATTTTC
AGTTCAAAACGGCCACATATTATACTTTGGCTGCAAACATTAAACAATTAGCAAAAGAACTGTGCGGGGGCAGATGC
GTGTTCTTCCTTGAGGGAGGTTATAATCTTAAATCACTATCCTACAGTGTTGCGGATTCCTTTAGAGCCTTCTTGGG
AGAGCCGTCATGCGCCTCCGACGTTGATCCCACCTTTTTATACGATGAGCCGTCAACTAAGATTGAGCAAGCAATCG
ATAAAGTGAAGGCCATTCACTCCTTA

Seq. ID No: 55
>DAC_15
ATGGAACAATTGTGGGTTCCATCATTACCGATTTTAGGTGGGAGAATCTTGCCTATGCTGAGGCACTATTGTGGTTT
TGGTAGTCACCATCCGCTAACATGGAGATCCTTACAGATTACTGGCCGTAAACAAAAACACAACGGCTGCTGGATAG
CCTATTGCCTGCCAAGCCATAATGGGACTTCAATAAGCGACACAAACGGGGTGAGGAAGGATTTAGCATTACCTGAC
AATCTTTTAAGAGATGCGCACATCCTATATTGCACTTCTCCCGCTATGGGTCATAATAAAGAAGCACACCCGGAGAC
GAATAAAAGAGTCCCGGCAATAGTGGACGCACTTGAGAAGCTAGAGCTGACATCTAAACACAGAGGTTCACAGGTTT
TAGAAATTCAGGACTTCCAGCCAGCATCCCTTGACGATATAGCATTGGTCCACAGTAGATCATATATAACGGGACTG
GAGAAGGCTATGTCTCGTGCCTCTGACGAAGGACTGATCTTCATTGAGGGTACGGGGCCGACATACGCGACACAGAC
TACCTTTCAGGAGTGCTTATTGTCTGCTGGTGCTGGCATAACGTTAGTGGACTCTGTTGTCGCTGCTTCTAAGTTAG
GCCCGAAGCCCCCACTTGGCTTTGCGTTAGTTAGGCCGCCCGGTCACCACGCCGTTCCTGAGGGCCCGATGGGATTC
TGCGTCTTTGGTAATATAGCTGTAGCTGCAAGGTACGCGCAAAACCAACACGGCTTGAAGAGAGTCATGATAATAGA
CTTTGATGTACATCACGGGAATGGCACCTGTGACGCCTTCTACGAAGATCCCGACATTTTCTTTCTTAGTACTCACC
AACTAGGCTCCTACCCTGGCACGGGGAAAATACACCAGGTCGGCCAAGGAAACGGAGAGGGGACCACCTTAAACCTT
CCATTACCTGGGGGAAGTGGCGATTACGCAATGAGATGTGCGTTTGACGAAGTCATCGCACCCGCAGCCCAGAGGTT
CAAACCCGATATCATCCTAGTCAGCGCAGGATATGACGCTCACGCTCTGGACCCTCTGGCAGGCCTACAATTCACAA
CTGGCACCTTTTACATGCTGGCAGCACGGATTCGTGAGGTCGCTGCCGAACTATGCGGAGGTAGATGCGTTTTCTTC
TTAGAGGGTGGATATAATTTGGAAAGTCTGAGCAGTAGTGTGGCCGACACCTTCAGAGCGTTTCTGGGGGAGCCTAG
CCTAGCCGCTCGTTTCGATGATCCAGCAATGTTATATGAAGAGCCGACACGTAAAATCAGGGAGGCTATCGACAAGG
CTAAGCACCTACACAGTCTA

Seq. ID No: 56
>DAC_16
ATGATGGCCACCGAACCGATTGCCTCCGGTTCAGGTACAATGGACATAGATTCAGAGAAAACGCCGTCTACCAGTCA
AGCTAACCCTATGGCAGATACGTTCCAAACAAGAAGACCAAGGGCTTCTTCCCTGCCACTTCAACCATCTAATTTAA
AGGTAGGCTATATTTATAGCAGCGAGATGATGAATCACTTCTGCCCTGGAGGGCATCCTGAACAACCTTTGCGTATT
CAACAGATCTGGGCGACAATCGTCAATGAGCAACTGCATAAAAGAATGAAGTGGATGCCGATCAGAGAAGTCAAGAA
AGGTGAAGCACTGTTGGTCCATAGCGAGGATCACTGGAACAAGGTCATTGCAATTCAGTACCTTACCGATCAGCAAC
GTGCAGATTCCGTTGATTATTACGAGCAAATGTCACTTTACGTGATGTCAGGCACGACTAGGAGTGCGCTACTAAGT
TGTGGAGGGGTAGTCGAGGCATGTTTAGCTGTTGCAAGGAACGAGTTGAAGAAGACCTTTGCCATTGTTCGTCCTCC
AGGACACCACGCTGAACCCGATGAACATATGGGGTTTTGTTTCTTCAATAATGTAGCCGTGGCGGCGCGTGTTGTCC
AACAGAGGACTAAGTTAAAGAAGATACTAATACTAGATTGGGACGTTCATCACGGCAATGGTACACAGAGAGCCTTC
AACGATGACCCATCCGTTCTGTACATCAGTCTTCATAGGTACGAGCAAGGGACATTTTACCCATGCGGTCCCTTCGG
TTCTCTTACTTCTTGTGGTGAAGGGCCGGGGACCGGTTTTTCCGTAAATGTGCCCTGGCCGTGTGCGGGCATGGGAG
ATGCCGAGTACATATATGCGTTTCAAAAGGTTATACTGCCCATAGCGACCGAGTTTGCCCCGGAATTGGTTATAATT
TCCGCTGGGTTCGATGCTGCTGCGGGGGATGAGCTGGGTGAGTGCCTTGTGTCCCCCGCGGGGTATGCTCACATGAC
TCACATGTTAGCGGGCCTTGCAGGCGGCAGGATGGTGGTGGCGCTAGAAGGCGGTTATAACCTTGATAGCATTTCCC
AGTCCGCCTTGGCAGTTACTAAGGTGCTTCTAGGAGAGCCGCCGGACGAGCTGCCACCTTTGAAGGCCAATGAGGAA
GGTACCGAAACCGTTTGGTTGGTAGCTAGAGAGCAATCCAAATATTGGAAGTCAGTGGACCCCAAAGCATGCGAACC
TCAGGCTGATGTTGAGCCTATTAGCTTTAGTGTCCCTGAGATCCTAAAAGCCCACCGTCAACATTACTTGTACACCA
AACATGATATGATGCAAGTCCCTATGATGACGCCAGAATTAGAAGAAAAGTTTAGTAGCCAAATCATGTGTACTTCC
GACATTTTCGAATCAAAGACCTTAGTAATCTTTGTTCATGAATTTGGGAACTTGAGGTTAGAGTTAGAGTCTAGCAC
AACATGTGATGTACATCTGGAGAGATCCTATCTAATAGACTTCTCCAAGGAGCTAGTGGGCTGGGTTAAATCAGAGG

GGTACAGCTTGTTAGACGCGAATCTGTACCCCAAACCGTCCACAACACCAACGCCCAATTTAAGGCATAAGACGATG
GAGGAAGTCGGAAGAGATGTGCTTGTCTATCTTTGGGATAACTACGTCCAGCTAAGCGGCGCGGAAAGGGTTATCTT
AATAGGGCACGGCCCTGGGTGTAAGCCCCTTGTGGATTTACTAAATAGAAGAACAACCTCTGTGACGAAGAGCGCCA
AAGCGATTATACAGGTGGTCGGGTCTCAGCGTATGCCGAGTTACCCGTCTGATGTGGACGATGCTAGGCCCTGGTAC
CAGAAATCCTCACTTGTTATCGTGCCTCAGTCTCATCCCGTGATGGGCCCTCATATTAAACCAAAGGACATTCGTAG
GCATGGTGTCATGGTACCTATCGACGAGACTAGACAGATCAAGTTAATTACACGTGCCCTGCCTGCGATTAAGCAGT
TTGTACAGGAGACTTTATCAAGTTTTCCACTAGCAAACAGGACGAATCGTCCA

Seq. ID No: 57
>DAC_17
ATGTCAAAGCGTAAAGTTGCGTACTTTTATGATCCAGATGTCGGCGCTTACACATATGGTTGGAGTCATCTAATGAA
GCCTCATAGAATGCGTATCACTCACGAATTAGCAACCGCTTACGGGATGCTAGATAAAATGCATGTTCTTCGTCCCA
AGAGGGCGACCCCCGAAGCTATGACAGCGTTCCACACTGATGAATACGTCCAGTTTCTTCATTCAGTCACACCTGAA
ACGGCTGATAAGTTGACAGGTCAGAAAACTAGATTTCTGGTTGGTGACGACAATCCGGCTTTTGAGGGTGTTTTCGA
ATTTTGTTCCATTTCAGCCGGAGGATCAATTGGAGCGGCGGAGAGGATAGCTTCTGGCGCTGCTGATATAGCTATCA
ATTGGGCAGGGGGTTTGCATCACGCGAAGAAGAGAGAGGCGGCCGGGTTCTGCTATATAAACGATATAGTCTTGGGC
ATCCTAGAGTTATTAAGGACATACCCGAGAGTTCTATATATCGACATTGATTGCCACCACGGTGATGGGGTCGAAGA
AGCCTTCTACACTACGGACCGTGTCATGACGTGTTCCTTCCATAAGTTTGGGGAGTTCTTTCCTGGTACCGGCACAC
AGGAGGACACCGGGACCGGAAAGGGGAAAGGCTACAGCGTCAATGTGCCGCTTAAAGATGGCATTCAAGATGAGTCA
TTTAAGTCCGTGTTTGACCCAGTGATCTCTAAGATCCTAGAAGTATTCCAGCCATCTGCCGTTGTGTTACAATGCGG
AGCCGACAGTCTTGCCGGCGACAAGTTGGGGTGTCTGAACCTTACCATGCAAGGTCATGCTCATTGCGTCCAGTTTT
TGAGAAAATCTAACATTCCATTAATCCTACTAGGTGGTGGAGGATATACGGTTAAGAATGTAGCAAGGGCTTGGACA
TATGAAACCGCATGCGCAATCGGAATAGAAAACGAAATAGACTTGAACATGCCGTGGTCTCAATACTTCGAGTGGTT
CGGCCCGACCTACCGTCTGGAAGTTCCAGAAAACAACATGGAAGATATGAACGTTAAAGAAGGCACTTTGGATCATG
TTAGAACGACTGCACTAGCCCAGCTGCAGCAGCTTGCGTCACGTTGTGCACCTTCTGTGCAAATGCAGGACGTCCCG
CGTACATCCTTAGGCGGCCACTTGGGCTTCAAACGTGACAAGAGAGAGCATCGTGATGAACTGGACGAGAGATTGGC
CCAGCATACTCGTTATTTGTACGACCTTCAAGAGTCCGAGTCAGAAAGCGAGGATACGGAAAGCTCAGACAGTGATG
CTTCCTCTGTGTCTTTCGTAAATAACTGGAGAAGAGCGCCCCATAGAGCTAACAGCCTTCCCCGTATCTTGTCTGGT
AGGCATTCCAGCAATCCCCCTGGACACATAAGTGCCTCCGAAAGAAGAAGAATGAGCATCGTAACCGGCAAGTACTT
CGACATCCCCATACACGAGAGCGGTTATAACCACTACGAATATGGCGCAGCCCCCACTAAATCATCAAAAAGGATCT
TTTTCCAATCAGGTCTTGACATTTACAACGATGACAATGATTTCGAGGGCATTATCAATGCTAGAACCTCAGTTAGC
AACGGATTTGGTAACGGAATACACGATCTTCACGGCTTAATGGAACGTGGAGGCAGAAGTTTGAACGAAAACCTGGA
AGACGGGGACGATGAGGTGGAAGGTGAGGAATACGAGGATGACGCAGCCATGTCAGATAGT

Seq. ID No: 58
>DAC_18
ATGGAAGAACACTTCTGGGATGTCCTATACAAGGACAAATACTCAAAGCTGTTGAGCAAGGCTCGTGATTTCTTGGA
CGATACCGGCGGACCGGGAGACGATGTCTTAGTCTTTATCTCTTGTGGGATGGACGCGTGTGAACATGAATATGAAA
GTATGTCAAGACATAACAGGAAGGTCCCTGCGAGTTTCTACCATCGTTTCGCAAGAGACGCATGTGCGTTCTCTGAC
CGTTACGCGGGAGGGAGACTAATTTCTGTCCTTGAAGGGGGATACAGTGATAGGGCATTGATCTCTGGAGCAATGGC
ACATCTAAGCGGATTGGTAGATACACCTGATGGTATTCAGGTCGATGAGCAGTGGTGGAACATACCGAACCTTGTGA
AATTGGAAGCTGCCACCAAAAAGCGTAGAGGTGGACGTCCTAGCCTGCCTGCGAAAGGCTCAGTGGAGCCGTGGATA
GAAAGGACGTTATCCATCTTTTCCAGCATAGATGGTAGTGCGAGCACAACCTCATCAAGGAGTACGTTTATACCACC
GTCATCTAGGACATTAAGAGATCGTACAAAGGGGAGGGAGGCTATGCCTAAATCCCCGCCAGCTTCCAGTGCTTCAA
CCAAGCCTGTGTCCAGATCAAAAGTAAAGCCGGGTGCAAATATCAAGAGCGGAGATGAATCTTTCGCTTCAACAGGA
AGTAGTCCGCTTACCTCCCCTTCTCCCAGTAGCTCTGAAGACGAAGCACCGCCCATTAAACGTTTG

Seq. ID No: 59
>DMAT_1
ATGACCATAATAAATTCCCGTATAATTGACATTAGACAGAGTACCTTCGAAGAGAGTATCCCCGACCAAGTAACGGC
CGGTTTATCTACCACCCCCAAGACCCTGCCCGCACTGCTGTTTTACTCAGGCGAGGGTATTCGTCATTGGATTGAGC

ATAGTACCGCAGCCGATTTCTACCCGAGACACGAGGAGCTGAGGATATTACGTGCGAGGGCTGCGGAAATGGTCGAT
AGTATAGCCAACAATAGCGTGGTAGTCGATTTGGGTTCTGCTTCCTTAGATAAAGTGCTTCCACTACTTGAGGCCTT
AGAAGCCAGTAAAAAGAATATTACCTTCTATGCTTTGGATCTTAGTTTCAGTGAATTACAAAGTACCCTGCAGAGCT
TGCCATACGAACAATTCAAGTTTGTGAAGATAGGGGCGCTACACGGGACCTTCGAAGATGGTGTTCAATGGTTAAAG
GATACACCTGGCGTACAGGACAGGCCTCATTGCCTATTATTATTTGGGTTGACTGTTGGGAACTACAGTAGACCAAA
TGCCGCGAAGTTTTTGCAAAACATTGCCTCCAACGCACTGGCGGCGTCACCAGTCCAATCTTCAATTCTGTTGAGCC
TGGACTCATGTAAGATGCCAACCAAGGTGTTGCGTGCATATACAGCCGAAGGGGTAGTTCCGTTTGCATTGGCCAGT
TTAGACTACGGGAATACGCTTTTTGCTCCTAACAAAATGGGGGAGAAGGTCTTCCAGCCTAGTGATTGGTATTTTCT
GTCTGAGTGGAACTACATGCTTGGCAGGCATGAGGCCTCCTTGATCACAAAAGGGAAAGAGGTGCGTCTTGGTGGGC
CCTTGAATGATATCGTGATAGAGAAACACGAGAAAATTCGTTTTGGATGTTCCTATAAATACGATACCGATGAAAGA
CAAGTTTTATTCGGATCTGCAGGTCTGACTGACGTAAAAGAATGGTCTGTCGAAGGTTGTGACGTGAGCTTCTATCA
ACTACAGATGTGTCCTAAC

Seq. ID No: 60
>DMAT_2
ATGACCATATCCGCCCCACCAATCATAGATATTAGACAAGCGGGCTTAGAGTCATCCATACCAGATCAGGTGGTTGA
GGGATTAACAAAAGAAGTAAAAACCCTGCCAGCACTGTTGTTTTACTCCACAAAAGGTATACAACACTGGAACCGTC
ACAGCCATGCGGCTGACTTTTATCCGAGGCACGAAGAACTGTGTATCCTAAAAGCAGAGGCCTCAAAAATGGCCGCG
AGTATCGCCCAAGACTCTTTAGTTATAGACATGGGCAGTGCAAGCATGGATAAAGTGATTCTGTTACTTGAGGCGCT
TGAGGAACAAAAGAAGTCAATTACATATTGCGCGCTTGACCTTTCTTATTCAGAGCTGGCCTCCAATTTCCAAGCTA
TCCCAGTAGACAGATTTCACTATGTTAGGTTTGCTGCTTTGCACGGCACATTTGATGACGGACTTCATTGGCTGCAA
AACGCGCCAGATATCAGAAATAGGCCCAGATGCATACTGCTATTTGGTCTGACAATCGGTAATTTCTCCAGAGATAA
CGCCGCAAGTTTCTTGAGGAATATTGCTCAGTCTGCGCTTTCTACGAGTCCTACACAGTCGTCAATCATAGTCAGTC
TGGATAGCTGCAAACTACCAACTAAGATACTTCGTGCATATACGGCAGACGGTGTTGTGCCATTTGCGTTGGCCTCT
TTGTCATATGCCAACAGTTTATTTCATCCCAAAGGAGATAGGAAAATCTTCAACGAAGAGGATTGGTACTTTCACAG
TGAATGGAACCATGCTCTTGGTAGACATGAGGCGTCTCTGATCACACAGAGCAAAGATATTCAACTTGGGGCACCTT
TAGAGACGGTTATTGTCCGTAGAGACGAAAAGATCCGTTTTGGATGTAGCTATAAATACGACAAGGCAGAAAGAGAT
CAGTTGTTTCACAGTGCAGGACTGGAAGATGCAGCGGTCTGGACTGCTCCGGACTGTGATGTTGCGTTCTACCAACT
AAGGCTAAGGCTAAAT

Seq. ID No: 61
>DMAT_3
ATGTCTAAACCGAATGTACTTGATATCAGACTGGCCACGTTTGAAGACTCTATCGTGGATCTGGTGATTAACGGTCT
GCGTAAGCAACCTAAGACCCTTCCGGCATTACTATTCTACGCAAATGAAGGGCTGAAGCACTGGAATCACCATAGTC
ATCAACCCGAATTTTACCCCCGTCATCAGGAGGTACAAATATTGAAAAAGAAGGCCCAAGAAATGGCAGCTTCCATT
CCAATGAATAGTGTCGTTGTGGACCTAGGATCTGCCTCACTAGACAAGGTCATCCATCTTTTAGAGGCGCTAGAAGT
ACAGAAGAAGAATATCTCATACTATGCCCTGGATGTATCAGCGAGTCAGCTGGAGTCTACACTAGCTGCAATACCGA
CCCAGAACTTCCGTCACGTGAGATTCGCGGGGTTACACGGAACCTTCGATGACGGGCTTCACTGGCTAAAGGAAGCA
CCGGAGGCTAGAGACGTGCCGCACACAGTATTGCTATTCGGTCTAACAATAGGGAACTTCTCCAGGCCTAACGCAGC
AGCATTCTTATCAAATATCGGTCAACATGCATTTCAGGGTAAGTCAGGCGATCAATGCAGTATTCTGATGTCTCTGG
ACAGCTGCAAAGTCCCAACACAGGTCCTAAGAGCCTATACTTGTGAAGGCGTTGTTCCTTTCGCCTTGCAAAGTCTA
ACTTATGCAAATGGCCTATTTTCAGAGAAAAATAAAACCCAAGCCAGCGGCGATGTTCAACACAAGGTCTTCAACCT
GGATGAGTGGTATTACCTATCAGAATGGAATTTTGTTCTGGGCAGACACGAGGCTTCTTTAATACCGCGTTCTAAGG
ACATTAAGCTATTACCGCCGCTTGACGGAATATTAGTCTCAAAGGACGAAAAAGTCAGGTTTGGATGTAGTTATAAA
TATGACCAAGAAGAACGTATGGAACTGTTTGCCGCCGCCGGAGTCAAGAATGAAGTAACATGGAGCGATGAAGGTTG
CGATGTAGCATTCTACCAATTAAAGCTATCA

Seq. ID No: 62
>DMAT_4
ATGGGTTCCATAAATCCGCCACAAATTCTGGATATCAGGAGATCAAAATTTGAGGAATCTATTCCGAAGCAAGTGGA
GGCCGGGCTGCTAAGTAGTCCTAAGACTCTTCCAGCATTACTATTTTACAGCACGGAAGGAATACAGCATTGGAATC

GTTATAGTCATGCTTCAGATTTCTACCCTAGACACGAGGAGATTCAGATACTGAAAGATAAGGCAACGGACATGGCG
GCTAGTATAGCTGACGGATCAGTCGTGGTAGACCTTGGTTCCGCCAGTCTTGATAAAGTTATTCATCTACTGGAGGC
GCTGGAAGCAGCCCAGAAGAAGGTTACTTACTATGCCCTAGACTTATCATTCTCCGAGCTGACGAGCACCTTGCAGG
CCATACCTACCGATCAATTCGTTCATGTCCAATTTTCAGCACTTCATGGGACATTTGACGATGGCTTGCAGTGGTTG
AAGGAGACTCTGGTTATAAGAGACCAGCCCCATTGCTTGTTATTATTTGGACTTACTATTGGCAATTTTAGCAGACC
CAACGCCGCGAAGTTTTTACACAACATAGCGTCTCATGCTCTAGTGGAGAGTCCAAGTCAATCAAGTATTCTTTTAA
CCCTGGACTCCTGCAAGGTCCCGACAAAAGTAATCCGTGCATACACGGCTGAGGGCGTCGTCCCTTTCGCATTGGAG
TCATTGAAGTATGGGAATACTCTATTCCAACAGGATGCCGGAGAGAATGTGTTCGATCCCGAGGACTGGTATTTTCT
GAGTGAATGGAATTATGTCTTAGGAAGACATGAAGCATCATTGGTACCTAGGTCCAAGGACATTAAACTTGGTAGGC
CACTGGATAAGATAGTTGTAGGGAAGCACGAGAAGGTCAGATTCGGTTGTTCTTATAAATACGATTCAGAAGAGAGA
AAAGAACTATTCGGAACAGCCGGGTTAAGGGATGTAAAGTCTTGGTCAAAAGAAGGGTGCGATGTCGCATTCTATCA
ATTAAAATGCTGCCCCAAT

Seq. ID No: 63
>DMAT_5
ATGCCGGCACTGCCTGTCATAGATATCAGATCCAATCACGTTGAGGACTCCCTACCCGAACAAATAATAAAAGGCCT
TACCTCACAACCAAAGACTCTGCCGCCGCTACTTTTCTATTCAAATGAGGGCCTTGAACATTGGAATCATCATAGTC
GTCAGCCCGACTTCTATCCAAGAAGGCAGGAAATCGAGATATTAAAACAAGGAGGAAACGATATTGCAAGATCCATT
GCGCCCTCAAGTGTCATTCTGGATCTGGGATCAGCAAACCTGGAGAAGGTGGGCTATCTGCTGGAAGCTCTGGAGGC
GCAGGAGAAAGACGTCCTATATTTCGCCTTAGACATCTCAGCGCCGCAACTAGCGACTACGCTAAAAGAAATACCAA
GTAGCAACTTTCGTCACGTCCGTTTTGCCGGATTGCATGGTACGTTTGAGGATGGTCTGAGATGGATAAATGAGACC
CCTGAGATTAGGGATTTACCGCACTGCGTACTTCTGCTGGGGCTTACGATCGGAAATTTTTCAAGGCAAAACGCCGC
TGCTTTCTTGCAAAACATAGCAAATCATGCCCTTACTGGTGCTTCAAAGAACAAATCAAGCATATTATTAAGCCTAG
ATAGTTGCAAAGTACCCACAAAGGTCACTAGAGCATATACGTCAGACGGGGTAGTTCCATTCGCTCTACAGGCGTTG
ACGTATGCCAAGGCGTTATTGTGCGATAGGATCGACAATGGGATAGATGAGAAGGTTTTATCATGTAACTTGAGGCC
CGAACATTGGCACTATCTTTCCGAGTGGAACTTTGCCTTAGGGAGGCATGAAGCCAGCTTAATTCCAAGGTTCGGAG
ATGTGTGCCTGGGCTCCATGTTACAAGATATCATTGTAAAGAAAGAAGAGAAAGTCCGTTTTGCCTGTAGCTACAAG
TATGATGCTAAGGAGAGGCAAAAGCTGTTCCTGGACTCTGGAGTAGATCAAGGGATGGTCTGGACAAACGAAGGATG
CGATGTGGCAATCTACGAGTTGAAGTTAGCA

Seq. ID No: 64
>DMAT_6
ATGCTTTACAAGCCCAAAGTTTTGGACATAAGGTCAGGATCAGTAGAAGATAGCCTACGTCATTCCGTGATGGACGG
TATTCGTGAAGACCCTAGAACCTTGCCTACCTTAATTCTATACGGTCCGGAGGGTTTACAACATTGGGACGATCACT
CACATGCTCCCGACTACTACTTACGTCACGAAGAGTTGCATATACTAAGAAGCAGAGCTTATGAAATGGCGGAGACT
ATCGCGGATAATACTGCAATGGTTGATCTGGGCTCCGCGCAGGTAAGTAGATTCCACGAGTCACCGTGCCTACTGGC
GCCGACTTTATCTCTCGACAAAGCAGCTCTTCTGCTAGACGCCCTTGAGGTTCAGGCTAAGAATGTCACATATTATG
CGTTGGACCTTGACCATGCGGAATTACAGAAAACGCTGTGTAGACTACCTTTAGGGAAGTACAAACATGTTCAATGC
GTTGGCTTACAGGGAACCTTCGAGGACGGACTAGAGTGGATAAAGAACGATCCGGAGCAGTCTAGGCGTCCGCACTG
TCTACTTTTCCTGGGTTCAACTATTGGTAATTTCTCAAGAGAAAATGCTGCCCGTTTTATAAGATCAATGGCAAGTT
CTGCCTTTCTAAGTGAGTCAGCTAAATCTTCTATTATCTTGTCTATAGATTCCTGCAAATTGCCCACTAAGGTTCTT
AGAGCGTATAACTCAGAAGGTGTTGTTCCCTTCGCGATGGCCGGCTTGAAGCATGCGAGCGCGATCTTGTGTGAAGC
CGCGTGTAGGCAAGAGGACGCAGTTACCGAAACTTTTCTGCCGGACGACTGGTATTACTTATCACATTACAACCACG
TATTGGGTCGTCATGAGGCCAGCTTCACACCAAGGAACAGAGATATTCAATTGGGTTCTCCACTAGAGGACGTCGTA
ATAAGACTTGGAGAGACAATCCGTTTCGGCTATTCACATAAGTATGATTTTGCCGAGATTGAGCAACTTTTTCGTGA
GGCCGGGGTCGCTGCTGTTAATAGTTGGGGAGCGGTTGGCTGCGACCTTTCATTCTATCAGCTGGGAACGGCG

Seq. ID No: 65
>DMAT_7
ATGGCAGCACCTTCAGTTATTGACATCCGTTCTCACCTGGTGGAAGATAGTTTACCAGATCAGGTGGTGAAAGGGTT
AGGGTCCGATCCCAAGACCCTTCCAGCTTTGTTGTTCTATTCAAACGAGGGCTTGGAGTACTGGAATCATCACGCAC

GTCAACCCGATTTTTATCCAAGGCACCAAGAAATTGAAATTTTAAAGCGTAAGGGTGATGAGATCGCCAGGAGCGTT
GCGCCTAATTCCGTCATTCTGGATTTGGGATCTGCTAACCTTGAAAAAGTCACTTATCTACTAGAGGCACTGGAAGC
GCAAGCAAAGAACGTAACTTATTTTGCCCTTGATTTATCAGCTCCTCAACTGATGAGTACCCTGAAGGCGATTCCCA
CCACTAAATTTCGTCATGTCAGATTCGCAGGATTACATGGCACGTTCGTCGATGGATTAAGATGGATATCAGAGACT
CCTGATATACGTGACCTACCTCATTGTGTTCTTCTGTTCGGCCTTACTATTGGGAACTTCAGTAGGCCGAACGCAGC
CACCTTTCTAAGAAACATTGCGAGTCAGGCACTTCGAGGGGCGAGCGAGGACAAGTCTTCCATTTTCCTATCATTGG
ATTCCTGCAAAGTGCCAACCCAAATCCTGAGGGCCTACACATCCGATGGTGTGGTTCCATTCGCTCTTCAGAGTCTG
GCTTACGCTAAAACTCTATTCTGTGAACAGACGCAAAACGATTTTAATGAGAAACCTTCAAGTTGTCATTTAAACCC
CGATGACTGGCATTACCATAGTGAATGGAACTTTGTACTGGGGAGACACGAAGCCTCTCTTATTCCGAGACTAAATG
ACATACATTTGGGGCCTCTTCTACATGATATCGTAGTGAAAAAGGACGAGAAAGTCCGTTTTGGATGCTCTTATAAA
TACGATGATCTGGAAAGGGATAAATTGTTCGTTGATGCTGGAGTAAAAGACGAGATGGCTTGGACCAATGAAGGATG
CGACATCGCTATTTACGAACTGAAGTCCATG

Seq. ID No: 66
>FEX1
ATGATCTTCAACCCAGTGATATCAAATCACAAGCTGTCACACTATATACACGTCTTCTGCACGTTCACAACCTTCTG
CATTTTGGGGACGGAAACAAGGCAAGCCATAACCGCGCTAAGTACATATACTCCAGCTTTCGTTACTGCTCCGACAG
TGCTATGGAGTAATTGCTCTAGTTGTATGTTAATGGGCATAATGCAGTCATTGAACGCGTACACTTGGATGAAGGAC
CATCAAGTTCTTTTCCTGGGTGTCACTACCGGTTATTGTGGGGCCTTGTCATCTTTTTCCTCCATGCTGTTAGAGAT
GTTCGAGCATAGCACCAACCTGACAAATGGGAACATTGCGAATCATACTAAGTTACCTAATCGTGCCTATGGGATAA
TGGAATTTTTATCTGTCCTGCTGGTGCATTTAATGGTTTCTATGGGAAGCCTGATCTTTGGTCGTCAGTTGGGAAAG
GAAGTAATCGTCGCTTACGGCTCATCCAGTTTTTCTAAACCTTACACACCACCGTCTGACACAGTAAAAGAGAACGC
CGGAGATGTAGATACACAAGAGATGGAGAAGAATATATTAGAGTTTAAGTTCAAAACGCCGGCGCCCTTCTTCAAGA
AGTTCTTCGACATAGTCGATAAACTAGCTTATGCCCTAGCTTTTCCACTGATCATATTGTTCGTAGTCCTATGTGCG
TATTACGAAAACTACTCCCGTGGAAAGTGGACACTACCGTGCTTGTTCGGCATATTTGCTGGTTTTCTTCGTTACTG
GTTAGCAGAAATGTTTAATAAAACCAATAAGAAATTCCCTCTGGGGACTTTCCTGGCCAATGTTTTTGCGACCTTGC
TGATTGGTATCTTCACGATGGTGCAAAGAGGAAAGAAGCACTTTTCCACTGACGTACCCATCGTTAATTCCTTGAAC
AGCTGTCACATAGTGAGTGCATTGATAAGTGGGTTCTGTGGGACTCTGTCTACGATATCTACTTTCATTAACGAGGG
CTATAAATTGAGTTTCATCAACATGCTGATTTATTACACTGTAAGTATAGCAATATCCTATTGTCTTCTAGTTATTA
CCTTGGGGTCATACGCGTGGACAAGAGGATTAACTAACCCCATTTGC

Seq. ID No: 67
>IDI1_for_fusion
GGCGGGTCCGGTGGTTCAAGTGGCAGCGGTGGCTCCTCATCAACCGCGGACAACAACAGTATGCCGCATGGCGCCGT
CTCAAGTTATGCGAAACTGGTACAAAATCAGACACCGGAGGATATCCTGGAGGAATTTCCCGAAATCATCCCTCTGC
AGCAGAGACCTAACACCAGATCCAGTGAAACGTCCAACGACGAGTCTGGTGAAACGTGTTTCTCTGGTCACGATGAA
GAGCAAATCAAACTTATGAACGAGAACTGCATAGTGCTAGATTGGGATGATAATGCTATAGGGGCCGGGACCAAGAA
GGTGTGTCATCTTATGGAAAACATCGAGAAGGGCTTATTACACCGTGCGTTTAGTGTATTTATATTCAACGAGCAGG
GCGAGTTGTTATTACAACAAAGGGCTACGGAGAAGATCACGTTTCCAGACTTGTGGACCAACACTTGTTGTAGCCAT
CCACTGTGCATTGATGATGAACTTGGTCTCAAGGGAAAGCTAGATGACAAAATCAAAGGTGCTATCACCGCTGCCGT
GAGAAAGTTGGACCACGAGTTAGGGATTCCTGAGGACGAGACAAAAACTCGTGGAAAATTTCATTTTCTAAATCGTA
TTCATTTCATGGCACCTTCCAACGAACCGTGGGGCGAGCATGAAATTGACTATATCCTGTTCTATAAGATAAATGCG
AAAGAAAATTTAACTGTCAATCCGAATGTCAATGAGGTGAGAGATTTTAAGTGGGTGTCCCCTAATGACTTGAAGAC
GATGTTTGCTGATCCTTCATATAAATTTACCCCTTGCTTCAAAATAATATGCGAGAACTATTTATTCAACTGGTGGG
AACAGCTAGACGACTTGTCTGAGGTCGAGAATGACAGGCAAATTCACAGGATGCTG

Seq. ID No: 68
>INMT_1
ATGGCTGCTCCTCACACCTCTCAACAAGATTATATCGACAACTTTAACGCCAGAGACTATCTGCAAACAAGCTACAC
CCCGGGGAAAGGTATTCTTTTTGGTGAGTGGATTGAATTTGCTACGCAGAATCTTCACGAAACCTTTACGACAGGTG
GTGTAAGAGGTGATACGTTGCTTGATTTCGGAACTGGGCCAACAATCTACCAGTTAATATCAGCCTGCGAAGTCTTC

GACAAAATCATCGTAAGCGACTTTCTGGAACAAAATAGGGCGGAGTTCAGGAAATGGCTAAATAAGGACCCTGACGC
CTTCGACTGGACACCCATAATCAAAGGAGTGTGCGAACTGGAAGGTAATAGGGAAGACTGGGAGAAGAAAGCCACAA
AACTGCGTTCCAAGGTTAAAGAGGTTCTAAAATGCGATGCCCTAAAAAGAAACCCTTACGACCCCATTGTCGTGCCA
CCAGTAGATTGTTTGTTAAGCTGTTTATGCTTGGAGGCACCCTGTAAAGACATTAAAAGTTACTGCGAAGTCCTGAA
GAATTTTCAAAGTCTGATAAAGCCCGGTGGGCACCTTTTAATCCTTTCAGGTCTGAATGCCACATTCTATTATGTAG
GTAAAACTTATTTCTCAAGCATGACAACAAAGAAGGAGGAGCTAGAAATGGCCTTCAAGGAAGCGGGATACATCATA
AAGAAGGCAGTCTATACGCCGAGAGCGGATAAATCCAAAATAGACGTTGCCGATTACGAGGGCCATTACTTTATCCA
TGCCCACAAGCCAAAG

Seq. ID No: 69
>INMT_2
ATGGCTGCACCCCATACCTCTCAACAAGACTATATTGATAATTTCAATGCGCGTGACTACCTGCAAACAAGCTACAC
CCCAGGAAAAGGGATCTTATTCGGTGAATGGATAGAGTTCGCGACGCAAAACCTGCATGAAACTTTTACCACCGGCG
GGGTTAGAGGCGATACCCTTCTGGATTTTGGGACGGGACCGACCATATATCAGTTAATATCTGCGTGCGAGGTCTTC
GATAAGATAATAGTGAGTGACTTATTGGAACAGAATAGAACCGAGTTCCAAAAATGGCTTAATAAAGACCCTGACGC
ATTTGATTGGACTCCTATTATTAAGGGAGTTTGTGAGTTAGAAGGGAATCGTGAAAATTCTGAAAAGAAGGCTGAGA
AACTGCGTTCAAAAGTGAAACAGGTACTAAAGTGCGACGCTCTAAAAAGGAATCCCTACGATCCAGTGATCGTTCCA
CCTGCAGACTGTCTATTAGTCTGTTTGTGTTTAGAAATTCCTTGTAAGGACATGAAATCCTACTGTAACGTGCTGAA
AAATTTCAAAGATCTGTTGAAACCAGAAGGCCAAATACTGATATTAGGAACTTTAAACGGAACTTATTACCACGCAG
GAAAGAAGAGATTTAGCCTTTTATCTAGCAAGAAAGAGGATTTAGAGATGGCTTTTAAAGAAGCCGGCTATATAATT
GAAAAGGCCGTCTATACGCTACGTGCCGATAAGTCTAACATCGACGTTGCCGACTATGAAGGTCATTACTTTATACA
TGCTCATAAGCCCAAA

Seq. ID No: 70
>INMT_3
ATGTCTGATTTTACGAACACCAGAGAGTACGAGGAGCAGTTTGATCCGCGTTTGTACCTGGAGACGTACTTTCACCT
TGGTTCAGGGTCTTTAGCCGACGATTTCTTACGTTTCGTTCTGGACAATTTCAACAAAACATTCAAGTCCGGGGCTG
TCAAGGGAAGCACATTAATTGACATTGGGACTGCCCCCTCAATTTATCAGTTATTATCTGCGTGCGAGTCTTTTGAT
GACATAATTGTGACTTGGCACACGAACCGTGAACTAAAGGAGCTACAAAAGTGGCTAAACAGCGAGGCAGATGCGTT
CGATTGGTCATCAATTGTTAAGCATGTCTGCGAAATTGAGGGAAACAGGATGGCTCAGAAAGAGAAAGAAGAGAAAC
TAAAGGGTAAAATTAAACAGGTTCTAATGTGCGACGTATCCAAAAGTAACCCCTTGAGCCCGCATGAAGTCCCTAAA
GCTGATTGCCTGCTGACAACCGTCTGCCTAGAGGCCGCTTGTAAAAATTATGAAAGTTACGGCACCGCGCTTAAAAA
CCTATCCAATTTGTTAAAGCCAAAGGGTCACCTACTTATGGCAGGTGATCTTGGAGCGAACTATTACGAGGTAGGAA
GTAACAAAGTATTCTCCCTGCCCGTCAATGAAAAATTCTTAAAGAAGGTCATTAGTGAGTCCGGCTATGAGATAATC
CAACTTGTCTCATTCGGGAAGCCCGAGAATGCCGATTTTGAAACCAGCGACTATGAGGGTTTTTATTTTGTCCATGC
CCAGAAAGTG

Seq. ID No: 71
>INMT_4
ATGGACTGCTTGATTTCTTGCCTATGTTTAGAAGCCCCCTGTAAAGATTTGGAAGATTTCACGAATACGTTAAAGAA
GTTTAAGGAATTGTTAAAACCCGGAGGACACATAATTATACAGTCAGTGCTGAATTGCTCCCTTTACTTCGTCGGGA
ACAAATCATTTTCATGCCTTAGCATAACTAAGGATGAACTGGAGCAGGCATTCAAGGAGGCCGGGTATGAAATTGTA
AAGTTAAAGGTTGTCCCACGTTCTGAGAAAATTTGGGCTAACGTGTCCGATCACTCCGAGTACTATTATATACATGC
TCGTAAACCACAG

Seq. ID No: 72
>INMT_5
ATGAGCGACTTCACAGGAAAGAATGAATATCAGACCTTCTTCAACCCTAAGGCTTACTTAGAGAGCTACTATCAGTT
AGGGTCCGGCAGTATGGGTGATGAGTATCTTCAATTTGTTCTAAAGGAGCTGGCAGAGACGTTCAATCCAGGAAAGG
TCAAGGGGGACACTCTAATTGACATTGGAACCGGACCTACTATATATCAATTACTGTCTGCGTGTGAGGCGTTTAAA
AATATAATTGTCTCCGATTTTACTGATAAAAATCGTGAAGAATTTAATGTGTGGCTGAAAAATCAACCTGGAGCGTT

TGATTGGAGCCCAGTTGTTAAACATGTTTGCCGTCTGGAGGGAGATAGAATCCCTTGGGAACAAAAGGAAGAAAGAC
TTCGTAAAACGATCAAACAGGTACTAAAATGCGATGTGTTTAATATAAATCCAATTGACCCGGTTACCATTCCACAA
GTTGATTGCCTGCTATCATGTCTGTGTTTGGAAGGTGCTTGCAAAGATTTTGAATCTTATATTACTGCCCTTAAAAA
TATGACGACACTATTAAAAATCGGCGGTTATCTGGTAATGACCGGGGACCTTGGAAATACGTATTATATGGTAGGGG
ATGTGAAGTTTAGCGGGCTGAACTTAAATGAGAACTTTCTTAGAGAAGCAATAACTGGAGCGGGCTACGTTATAGAG
TCATTCCAACAATCTAAAAAGACCGAAGATAGTGTAGAAGATAAGGCGGATTTTACAGCTTACTACGTGATAGTAGC
GAGAAAAGAGAGGAATGTG

Seq. ID No: 73
>INMT_6
ATGGAGTCAGGGTTTACGAGTAAAGATACATACTTGAGTCATTTCAACCCGAGAGACTACCTTGAAAAATACTACAA
GTTCGGTTCAAGACACAGCGCGGAATCCCAGATTCTGAAGCACCTATTGAAGAATTTGTTTAAAATATTTTGCCTAG
ACGGCGTAAAAGGCGATTTATTAATTGATATAGGGTCTGGCCCCACAATTTACCAACTACTGTCTGCGTGTGAATCA
TTCAAGGAAATAGTCGTGACAGACTATTCAGATCAAAACTTACAGGAGCTTGAGAAGTGGCTGAAGAAAGAGCCAGA
GGCTTTTGACTGGAGTCCAGTAGTCACCTATGTTTGCGACCTTGAGGGCAACAGGGTTAAAGGGCCGGAGAAAGAAG
AAAAATTAAGGCAGGCCGTAAAGCAGGTATTGAAATGCGATGTGACCCAGTCCCAACCTCTGGGCGCAGTACCTCTG
CCGCCCGCTGACTGTGTCTTGTCTACCCTGTGTCTGGACGCGGCATGTCCTGACCTACCTACTTACTGTAGGGCATT
AAGGAATCTAGGCTCTCTATTGAAGCCGGGTGGCTTTCTAGTAATCATGGACGCTTTAAAGAGTTCCTACTACATGA
TTGGTGAGCAGAAGTTTAGTTCACTGCCGTTGGGCCGTGAGGCGGTGGAGGCGGCAGTAAAAGAAGCGGGCTACACG
ATTGAATGGTTCGAAGTTATCTCCCAGTCATATTCTAGCACGATGGCGAACAACGAGGGGTTGTTTTCATTGGTGGC
CAGGAAACTGAGCAGACCTTTG

Seq. ID No: 74
>INMT_7
ATGAAGGGTGGTTTCACAGGTGGAGATGAGTATCAAAAGCACTTTTTGCCCAGGGACTATTTGGCTACGTATTACTC
CTTTGACGGGTCCCCTTCCCCGGAGGCTGAAATGCTGAAATTTAACCTAGAGTGCCTACACAAAACTTTTGGTCCTG
GAGGGCTTCAGGGTGACACTTTAATAGATATAGGAAGCGGACCCACCATTTACCAAGTGCTTGCAGCATGTGACTCA
TTTCAAGATATTACTTTAAGCGATTTTACCGATAGAAACAGAGAAGAGCTTGAAAAATGGTTGAAAAAGGAACCTGG
AGCATATGACTGGACACCAGCGGTGAAGTTCGCTTGTGAATTAGAAGGCAATAGCGGACGTTGGGAGGAAAAAGAGG
AAAAACTAAGAGCTGCCGTTAAGCGTGTGCTAAAGTGCGATGTGCATTTGGGAAACCCCTTAGCACCAGCCGTATTG
CCGCTTGCAGACTGTGTGTTAACCCTATTGGCAATGGAGTGCGCTTGCTGTTCACTGGACGCATACAGGGCCGCGCT
ATGCAACTTAGCGTCTCTGCTAAAACCCGGAGGTCATCTTGTAACAACAGTGACCCTGAGACTGCCGTCCTATATGG
TAGGCAAGCGTGAATTTAGTTGCGTGGCTCTGGAAAAGGGTGAAGTCGAGCAAGCCGTATTGGACGCAGGATTCGAC
ATCGAGCAGCTACTGCATAGCCCGCAATCTTACTCAGTCACAAATGCAGCGAATAACGGAGTTTGCTGTATCGTCGC
ACGTAAGAAGCCAGGACCT

Seq. ID No: 75
>INMT_8
ATGAGCGATATAGATGACGGAGCCTTAGCTTCCGCCCAAGCAATCGTCGATGGCAACCGTTTAGCGGGGCAAATTGA
GCTTCGTCAGCAGCCTGACCCTGACAGAGTTTTCGCGGGTGTTCTTAGACAGGGCGAGGCGGTTGCTTTTTGTGTAT
GTAACCCGCCGTTTCATGAATCACTAGAACACGCCAGGCGTGCCGCTGGTGCTAAGTGGCAACGTCTAGGTAGAGCG
GTTCAAGGAAAGGAGATGAACTATCAAGGGTCCCCGGCGGAACTATGTTGTAACGGTGGAGAAGTGGGTTTCGTGAC
GAGGATGGCTGAAGAAAGTGCCCAACCGCGTCGTCAGAGGGCTTGTGTCTGGTTCAGTGCTATGCTTTCACGTGAAA
GTTCTATCGCCCCTGTCCGTGAGCGTTTGGGTGAGCTGGGCGCGAGAAGGAGAGCCTGGGAGTTACGTCAAGGGAGA
ACAACGAAGTGGGTAGTTGCTTGGACCTTCTACCCGAGGGGAGAACGTGACCAACGTCTACGTGAAATGGCGCAGAG
GCGTGCAGATCCTGAAGCACGTGCTGAAGCGGGTGCAGAAGCTGCGACCGCGAGAGATGTAGGCGCAGGAGGGGACG
GGGCCGATGGCGTGGGTGGGAGTTTAGTCCGTCGTTCAGCAGGCGCTGGAGGATCAGCCGCA

Seq. ID No: 76
>INMT_9

ATGGATTTCACGGGAGGCGAAATCTATCAATCCAGTTTCGACCCGAAAGCCTATCTTGCTTCCTTCTGCTCACTTGG
CTCCGGCAGGGACGACATACTAATGTTCAGACTTAAAAAGTGCTTTGAGACCTTTGGACCCGGTGGGTTGCGTGGTG
ACGTTTTAGTTGACATAGGAACTGGTCCAGCCATATACCACTTGTTGAGTGCTTGCGAGTCCTTCCCATACATCATT
GCTACTGATTTTACCGATAACAATCGTCAAGAACTGGAAAAGTGGTTGAGGCGTGAACCTGGCACGTTTGATTGGCT
GGAAACAGTAAAAATTGTCTGTGACTTAGAAGGGGACTCCAGGGACGATTGGGTGGAGAAGGAGGACAAACTAAGGT
CAAGGATACAAAAAGTGTTAAAGTGCGACGTTACGAAGACAAACCCCCTAGATCCAACTGTAATTCCACCTGCAGAC
TGTCTAATCACGGCATTGTGCCTTGAGACTGCGTGTACGGATATTGATACCTATTTCTGTTCTCTAAGAAATATAAC
CACTCTACTGAAACCTGGTGGGCATCTTGTTTTGATCGGAGTTCTGGGTAACAGTTTCTATAAAGTGGGTGAGAAAA
AGTTCTACTGCCTTTCCCTAGACGAGCAAACCGTAAGGAACGCGGTCATAGACGCGGGGTACTCTATAAAAGACCTG
GAACTATATTATTTACCTAATCCCGCCTCTTGTGCCCACATCACCGATACCTATGCAAATATCTTCCTTGTGGCGCA
GAAGAACGAGACC

Seq. ID No: 77
>INMT_10
ATGGAAATCGTGTCTACCTCATATAACCACATCTACGACAACTTTGATGCAAGGAAGTACCTGGACAGATACTATGG
GTTGGCATCTGAAACCCAAGAAATTGAGGAAGAGAGTGTTTTTCTTTTAACCTTTCTTAGTAACGTATTTTCTAGCG
GGAGGGTGAAGGGACATTCTTTTATTGAGATTGGTGTAGGGCCGTCCATACATTCTATACTTTCAGCATGTGAGGTG
TTTGAAAAGATTTACCTGACTGATTACTCCCAAGGAAATTTGAACGAAATCGAGAAGTGGTTGAATAGCGAAAATGA
TGCATTCGATTGGACTCCCTATATAAGATTCGTGTGCGACTTAGAGAATAATGGTTCTACTCCTAAGGGCAAGAAAG
AGAAGTTACGTAGAGCTGTCAGTTTAATGAAGTGTGACGTAAACTTGTCCAACCCGCTACACCCCCATTCACTTCCA
CTTACTGATTGCCTACTGACGGCGTCATGCCTTAGTGCTACCTGCAAAACCTTCACAGATTTTAAGATGTCTCTTAA
AATCATCGTAAGCCTAATCAAACCAGGCGGACATCTAATCCTTATTGACTATCTGAGGGCGAGTTATTACTGGGTTG
GAGAAGTCAAGCTGCCCATTTTGTCATTGGACGAACATGTTGTGAGGGAAGCTGTCGTGGAAAGCGGCTGTAAGATA
GAAGAGTTCAAATGGTTTAAGGAGTTTCACATGCCTGACGAGTTGTCAGACTGTAAAACTGTGTTCTCCTTACTAGC
TCAGAAGTTG

Seq. ID No: 78
>INMT_11
ATGGATTCTAGCAACTACAAACTATACCACGTGCACGAGTTTAATTCAAGGTCCTTCCTAGACAATTACTTCTCAGA
TGGTCCGCAAATGACTTTCGTGGACGACACACTTGTCTTTCCAATAGAAAACTTAAAGAAGACCTTCGCGGAGGGCC
ACATAAAGGGTGATGTGATGATAGACCTTTCAATAGGTGCGATGGTACACCATCTTTACGCTGCATGTGAATTTTTC
AAAGATATAATAGTCTTGAAGGCATCAGATAGATGCATCATGGAACTTAAAAGGTGGGTTGGAACAAGGACTGGGGC
ATTTTACTGGGGACATGCCACGAAACTACACGCTGACACAGAAGGGAACAGTGAACTACTGCAGGATAAAGAAGAGA
AGGTTAGATCAGCGATCCAACACGTGGTGAAGTGCGACGTAACCAAGGAATTGATGACCGATCCGATCGTTCTTCCG
CCAGCCGACTGCATCATCTCTGCATGGCTACTGGATGCAATCAGTTCAAATCAAGATGACTTCATAACGTACTTGAG
AAGGTTCATAAAGTTATTGAAGCCAGGCGGGCACTTGATCCTTATCGGAGCGTTGGAACAAACCAGTTATTCAGTAG
GCAACGAGAAATACCAGTTCCTTACGTATAATGAGGACTTTGCGAGGAAAGCTCTTATAGCTGAAGGCTTGGTCATC
GACGACTGCAAAATCAAGAAAAGAACGGCAAAGTCTGATTTAGCGGATTATAAATCCATATTGTATTTAGTTTCACA
CAAAAAG

Seq. ID No: 79
>INMT_12
ATGGACCCGTGTCTTAATCTGTACTATCCATCTCATGAGGTAAACGCCAAGAGATTGTTACATGAATACTTTTCCCA
GAATGTCCCCTACTCTATATTCAAAGAATCAACTATAAATATCATGAAATGCTGTTACAAGGCTTTTAGTTCAGGTT
TGCTTAGCGGTACAACCCTTATAGACATTAGTGTCGGCCCGTCCATAGTACATCTTCTATCTGTGTGTGAGTTTGTG
GAGGAAATAAGTATCCTAAAAGTCAATGACGCCTCAATTAGGGAGCTGGAACTTTGGAAAAACAAAGATCCTGAAAC
GTTCGACTGGACTCATACTTTGAAGCTGTTTATGGAATTGAAAGGTACTAGCCGTGACGGATGGAAAGACGCTCAAG
AAATGCTGCGTAGGAAAGTTAAACACATAGTCAAATGCGATTTCTCTAAATCCAATCTGACTAAGCCCTTCGCCTTA
CCTAGGGCTGACTGCGTTACTTGCATTTGGGGTCTTGAGACTATCAGCAGAGATCACGACGAGTGGAAAACTACCCT
TAGGAAAATTAGTGATCTGGTTAAACTTGGCGGTCATGTGTTGATTCATGCTGATATTAACGCTTCATACTTTAAAA

TTGGTGAGGACAAATATCACTTGTTTAACTTCGATGATGCGTTTCTTAGGAAAACTCTTACTGACGGTGGGTTCGCT
ATTGTTCATTACGAAAATCTGGAACGTGAGGCATGCACCGATTGTTTGGACCATTCCCACAAG

Seq. ID No: 80
>INMT_13
ATGGAATTAAAGCGTTGGGTCGATACCCGTACAGGGGCATTTGATTGGAGTCACGCTGCTAAATTGCACGTGGATAC
CGAGGGTAATTCCGACGAGCTTCAAGAGAAGAACGAGAAAGTCAAAAGTGCAATTCAGCATGTTGTGAAATGTGACT
TAGAAAAAGAAAACATGACACATCCTATTGTCTTGCCACCTGCCGATTGCATAATCAGTTTTGCCCTGCTTGATGTA
ATATCAAAAGATAAAGACGATTACATTAAATACTTGAGAAAGTTTTCAAAGTTATTGAAGCCTGGTGGGCATTTAAT
TTTAATAGGCGACCTAGATACTACGTATATTACTGTGGGTAAGCACAAAGTTCACTATCTGACATACGATGAGGAGT
TCGTGAGGAATGCCTTAGCTGGTGAGGGATTCGTCATAGACTGTTGTAAAGTCAAAGAAAGAACTGTCGAGTCTGAC
CTATGCGATTACAAAGGTATGATCTTCATAGTCGCCCACAAGGAAAAG

Seq. ID No: 81
>INMT_14
ATGGAGCTGAAGAGATGGGTGGACACCAGGACTGGAGCGTTTGACTGGTCTCATGCGGCAAAGTTACACGTAGATAC
TGAGGGAAATAGCGATGAGCTTCAAGAGAAGAACGAAAAAGTCAAGAGTGCTATACAGCACGTAGTGAAATGTGACT
TGGAGAAGGAAAATATGACACATCCGATTGTACTACCGCCCGCTGACTGCATTATCTCCTTCGGGTTCCTGGACGTT
GTCTGCAAGGACCAAGAAGACTATATCAGATATTTGAGGAAATTCAGTAGATTACTGAAACCTGGGGGCCATCTGAT
ACTAATAGGGGGTGTTGACGCTACTTATTTTACCGTCGGGAAGGAGAAGCACCATTTCTTCACATACGACGAAGCCT
TTGTGAGGAAGGCATTAGAGGGGGAGGGATTTGTCATCGATGACTGCAAAGTAAAGAAGCGTACAGCAGTAAGCGAT
TTTACCGATTATAAGGGGAGTATATTCATCGCCGCACACAAAGAACAC

Seq. ID No: 82
>INMT_15
ATGAGCGACTTCACCAATACTTCAGAATACGAAGAGCAATTTGACCCGAGGTTATACTTGGAAACATACTTCCACTT
GGGATCTGGGAGCCTGGCGGATGATTTTCTAAGATTCGTTTTAGACAATTTTAATAAAACTTTTAAATCTGGGGCGG
TGAAGGGATCTACATTAATCGATATAGGCACTGCACCATCTATTTATCAGCTTTTAAGTGCGTGTGAATCATTTGAT
GATATAACAGTTACATGGCACACTAACAGAGAGTTGAAAGAATTACAGAAGTGGCTTAACAATGAAGCTGACGCTTT
TGACTGGAGCAGTATTGTGAAGCATGTGTGTGAGATTGAAGGTAACAGAATGGGTCAGAAGGAAAAAGAAGAAAAGC
TGAAGGGCAAAATTAAACAGGTGCTGATGTGCGATGTTAGCAAGTCAAATCCTCTGAGCCCTCATGAGGTTCCGAAG
GCTGATTGTCTGCTTACGACCGTTTGTTTAGAAGCGGCCTGTAAGAATTATGAATCTTATGGTACTGCCCTTAAAAA
CCTATCCAATCTATTGAAGCCAAAAGGTCACCTTCTGATGGCCGGCGATCTAGGCGCTAATTACTACGAAGTTGGGT
CCAACAAGGTATTTAGTCTTCCCGTTAATGAAAAATTCTTAAAGAAGGTCATCTCAGAGTCCGGGTATGAGATTATA
CAGCTAGTCAGCTTCGGAAAACCCGAGAACGCTGACTTTGACACTTCAGATTATGAAGGTTTCTATTTCGTTCACGC
GCAGAAAGTC

Seq. ID No: 83
>INMT_16
ATGGCGTTACAAGAGAGGCAAGAACCTGACGTTTATCAGGAAAACTTTGAGCCAACCTCATATTTGGAATACTACCG
TATGAATCAAGATCCAGTAGGGGACGAGGTATTGCATTTCCTTTTAAAGCACTATAACGCGACGTTCAAACCGGGAG
GGTTAGAAGGGAAACTACTTATTGACATAGGTTCTGGTCCCACCATATATCAATTTTTATCCGCCTGCGAGTCCTTT
CAAGAGATTATTGCTACGGATTATACAGATAAGAACTTACAAGAATTGGAGAAGTGGCTGAAGAAGATGCCTGGCGC
ATTTGATTGGTCTCCTGTCGTCAAGTATGTTTGTGAGCTAGAAGGGAATAGAGATAAGTGGGCTGAAAAAGAAGAAC
GTGTGAGGAGGGCTGTCACCCAAGTGCTGAAGTGTGATGTGTTAAAAGAAAGGCCACTGGAGCCCGCGGTGCTTCCC
CCAGCGGATGGTCTGATTTCAAGCCTATGCCTTGAAGCAGCCTGCCCCACCCCTCAAGCGTGCAGAGATGCTCTGAG
GCATCTTAGGACACTGCTTCGTCCCGGAGGGCATCTTGTTCTAAGCGGAGGCTTCGAGACGACATTTTTCATGGTGG
GAGACAAGCGTTTCTCAACGTTACCATTAAACGAAAAATTCTTGAGGGAGGCACTGCAAGAGGCAGGCTTTATAATC
GAGAAGCTGGAAAAGGTAACCAGGGCCGCGGAAACCCACCTAGATAATCGTTCTGATTACACAGGTCTATTCTTCCT
GGTTGCGCGTCGTGGAGAC

Seq. ID No: 84
>INMT_17
ATGGACAAAATTTCTGCCCCTTTCTTTAGTGGCACATCACCAGCAGCAGCTTCAGTAGCAGGCGTTGATGAGGACGA
CAGGTTATGTTTCCAAGCCCAGGAGCTAATGTTTGCCTACAATATATCTATGGTCTTGAGAGCAGCAATCCAGTTAG
GGTTACTGGACGCACTGTCAGCCGCTGGTGGGAAAGCATTAACGCCCAACGAGCTTGTGGAAAACGTTGAGACATCT
AGTAACAAAGCCGAAGCAGCGGCAGCAGTGGACCGAATCCTGAGATATCTATCATGCTTCAACGTGGTTACCTGCAG
CTCAGAGGCTGCAGGTCCAGATGGAACTCTTGTTAGGCGTTACACCACGGGGCCCCTTTGCAGGTGGCTTACTAAAG
ACAGAGGGGATGGGACCCTGTCTCCTTTCGCTGTCTTTGTCGTAGACCCGGACCACCTGTTCCCTTGGCATCACATT
GCCGAGGCGGTCACCGCTGGGGGTCCATCTGCATTCGAAAGAACACAAAAATGGCCTTACTACGAATATATGGGGAA
AAATCAAAGACTAGGTACCCTTTTCGATAATGCTATGGCACAGCATTCTGTCATACTTGTTACTAAGATGCTAGAAA
GGTTCAAAGGATTTGACGGTGTGCAAAGACTGGTTGATGTGGGCGGAGGGACGGGTTCAACCTTGGGAATGATAACG
AGTAAATATAAACACATGACCGGAATAAACTATGATCTTCCCCACGTTATTGCGCAGGGGCTACCCTTGCCGGGAGT
AGAACATGTAGCAGGAGATATGTACGAGTCTATACCAACCGGCGATGCAGTCCTTCTTCAGTGGATAACACTAATGC
TTAATGACGATGAGTTCGTAAAGATCCTATCCAACTGTCACAATGCATTGCCCAAGGACGGGAAAGTCATAGTCGTT
GATGGGATCTTACCCGAGAATCCTGACAGCTCTTTAACGGCTCGTGACGCCTTTACGTTAGATATCATTATGTTTGT
TCTTTTCAAGGGAGCGAAACAGCGTACCGAGAAGGAGTTCGCGCGTCTGGCGAAGCAAGCCGGGTTTACGGGAGGGA
TTAAGAAGACTTATATATTCTTTAACTTCTATGCATTAGAATTTACAAAG

Seq. ID No: 85
>INMT_18
ATGGACGCTAACAAGAGATACCACGGTCCGCCGGTGTTGCTGGGGGTGGTAAGGGACTCAGAGAAGTTTGACTTCTG
TATGTGTAATCCTCCGTTCTTTGAGACAATGGAAGAGGCTGGCCTTAACCCCAAGACTAGCTGCGGAGGTACCCCAG
AGGAGATGATTTGTCCAGGCGGGGAAAAAGCCTTTATCACCCGTATAATAGAAGACAGTGCTGTTTTAAACCAATCA
TTCAGGTGGTACACGAGCATGGTTGGAAGGAAGAGTAACCTGAAGAGTCTTATATCTAAACTGAGGGAAGTCGGTGT
GACTATCGTTAAGACAACGGAGTTTGTACAAGGGCAGACATGCCGTTGGGGTCTAGCCTGGTCCTTTGTTCCTCCCG
TGAGGAAGATAGTTTCACCGCATGTGGCCGAGAAGAACATCATTAGCTTCATGCTAGAGGTGTGGGTCCCCGGATTT
AGTATATGTAGGGTTGGCGACGATCTTGTTCCGAAGAGTAAGTCTCCTCATCTGTCTCCCATACTTGGAACGAAGAA
C

Seq. ID No: 86
>INMT_19
ATGGAAGAAGCAGGCTTGAACCCTAAGACAAGTTGCGGCGGCACGCCCGAGGAAATGATCTGTCCTGGTGGAGAAAA
GGCCTTCATTACACGTATAATCGAAGATAGCGCGGTCTTAAATCAGTCCTTCAGATGGTATACCTCAATGGTTGGAA
GAAAGAGCAATCTTAAATCTTTAATCTCCAAGTTAAGAGAAGTTGGGGTCACAATCGTGAAAACCACGGAATTTGTA
CAAGGCCAAACGTGTCGTTGGGGTCTGGCGTGGAGCTTTGTGCCCCCAGTCAGGAAGATTGTAAGTCCCCATGTTGC
GGAGAAAAATATAATCAGTTTTATGTTGGAGGGATTGCAACGTCAATTTTCCGCAATCCACGTGTTGCAATCTATTG
AATCTTTCTTCAGAACTTGCGGGGCAAGCAGCGAGCTAAATGCAAGTTCATTCACTGTAGATATAACTGCTACAAAT
GATCATTGCAAGGCGATTTTGAATAATGAGCTTCAAAGCATAGATGAAGCCACTTCATGTGAACATGTCCCGGAGAC
CTCTAATAGCAGTTCTTCCCTACATCCGCACTCTAATGGTCTTGGGTTTAGGATCAGCGTGTACCAGCAAATTCCGG
GGACGTTGTTAGTTAAGGGCTCATTACAGCACAAAAACAACCCAGTGAGCGGGGCCTTTTCATTGATAATTCAGAGG
CTTGAAGAGGACCTTAAATACAAGTTTTGCAGA

Seq. ID No: 87
>INMT_20
ATGAATAGGAGCAATTATATTCATTGGATTGAGGACTTGTTAGCGTCTGACATAACCGAAAAGAACGAAGCGAACGG
AGGCAAGGTAAGGGGGTTTGACATTGGAACGGGAGCGAACTGCATATATCCACTGTTGGGCGCTTCCCTGCTGGGTT
GGAGCTTTGTAGGATCAGATGTCACTGAGGTAGCCCTAGATTGGGCTGAACAGAATGTAAGGAGTAACCCACATATC
TCAGAGCTGATAGAAATCAGGAGGGTGGATGTCGATGATCCTGCGAGTTCCTCAGGGACCGTAGAGTCATCCGGTGG
CAGTAGAATGGAAGACTCTTCACAGGGTCAGTGTGACGTAGTAGAGTTGGCATCCCTTGAAATGAAGGAGTTCTGTG
ACGTTGGCGTTACCTGTAAGGGAGGAACTGATAAGAACCAAAGGAGATACGATGAAGCAAAGCACTCAAACGTAGCT
AAGGGCTACCAAGGTCCTCCCATTTTGCTGGGGGTAGTCAAAGAGGGAGAAAAATTCGATTTCTGTATGTGT

Seq. ID No: 88
>INMT_21
ATGGAAGAGGCTGGATTAAATCCTAAAACCTGCTGTGGGGGCACACCGGAGGAAATGGTTTGCCAGGGCGGAGAGAG
GGCTTTTATTTCCAGGATTATAGAAGATTCAGCCACGCTAAAACAAAGTTTCAGGTGGTACACGTCAATGGTGGGTA
GAAAGAGCAACTTGAAGTTTTTAATGTCAAAATTGCGTGAAGTGGGCGTAACCATTGTTAAGACGACCGAGTTCGTA
CAAGGACAGACCTGTAGGTGGGGTTTGGCATGGTCTTTTATGCCAACCGCGAAGAGATCCGTGCCATCTCATGTCGC
TGAGAAGAGAAATCTGTCTTTTATGTTAGAAGGACTGCACCGTCAGACCTCTGCATTCAACGTCTTACAATCAATGG
AGTCCTTTTTCTCTCATTTCGGAGCACTTTGTAAGAGTAACCCTTCTAGCTTCACTGTTGACGTGTCAGTATCATCT
GACCACTGCGACGCAATCTTGAAATCTGACGTGGAGAAACTGGACGAGGCATCATCCCACTCCTGCGTCGCAGAGAG
TCCGGGTTCTGCGTCAAGTTACGACCCTATGGTCGTGTCATTCCGTTTGTCCGTGTTTCAACAAATTCCTGGTACCT
TGTTGGTCAGGGGGTCTCTACAACAGCGTGATTCACCATTATCTGGGGCATTTCTGTCTGTATTCCAGCAGTTAGAA
AAGTTCTTGAAACACAAGTTCTGTAGAGAGAGAGGACTTCAGTTTAACCAACGT

Seq. ID No: 89
>INMT_22
ATGGCCACGGAAATAGATGACGAGTCATACGAATCAGCTAGGAGAAACATTAGCAATAATAATATGCAGTCCCGCAT
TCACGTGGAGAAGGCGAGTCCCGACCAATCCATTTTGTTCCCCCTGGAAGATGACAGAACGTTTGAGTTCACCATGT
GCAACCCGCCATTTTACGGTAGCGCCGCTGAGGTGGTACAGTCCGCCGAAGCTAAGGAATTTCCTCCCAATGCGGTT
TGTACAGGAGCGGACATTGAAATGATTTACCCCCACGGAGGAGAAGAAGGTTTCGTCATGAAGATCCTGGATGAATC
CGAGAGGTTCATGACTAGGTGTAAATGGTACACCAGTATGTTAGGCAAGATGTCCAGCGTCGCTACTATTGTTGAAG
TTTTGAGACAAAGATCCATTACGAATTATGCCGTTACGGAGTTTGTACAGGGTCAAACCAGGAGATGGGCTATCGCT
TGGTCATTCGCAGACACAAGGCTGCCAGATACTATGGCAAGAATACAGAGCATCAGCCCGAAACACGCGCTATACCC
GTGCATGCCACCCAAAAATACTTTGGTTCAGGCTTTCCCGGGACCTGCGACCCACCTTGTCAGCACGAAACTAATTG
AAACTTTGCACGGCATAGAAGGGGTATCCTACACGACAACTAGCCTCAATTCTTTCTTCGTTGAAGCCAGACAAAAT
ACTTGGTCTCGCAGCGCCAGGAGGTCTCGTGCTAATAAGAATAGTTCGAAAAAGCCCGACCCATCCAGTTTAGACGC
GGATGACATTCTGTCCGGGAGTCAACCTGCACTTACCTGCTCATGCAGAGTGCTGGCGGATACCGCACATGCGGACC
CGGTGAACGTTGTCGAAAACCAATGGATTTTCGGTAATGACAGAGCCTTATTCGAATCCTTTGTGGGTCATGTTTCC
CGTAAGGTAGGAATGGGCCTACGTGACGTTAAA

Seq. ID No: 90
>INMT_23
ATGCTTCTGGAGAGCTACAAAACTTTCGAGCCCGCTAACTATTTGCAAGAATACTACTCTACTGTAGACTTAGAGAA
TCGTTCATTGTTAGCCTTTTTCGCAGAGGCATATAAAGGCATCGACCCTAACAGTGTTATGCTAGAGTTTAGCGGAG
GGCCATCTCTTTATTCTCTAATCACCGCGGCGGCTCACGTTAAGGAAATACATTTCAGCGATTTTCTGGAACGTAAC
GTGGAAGAGATAAAATTGTGGAAACGTTTCCGTCACCGTTCCTACATTTGGATTAACTTCTTTAAGGAAGCGTTAAT
GGCCGAGGGGCTATCAGAGGTGTCTACTGATGATATTCTTGAGCGTGAGGAACTATTAAGTAAAAAGCTATCAGATT
TCCTTCTATGCGATGCGTTCAATAGGCATCCTTTGGGTCAGAGATGCTACCAAAGATACGACGTCGTTGCAGCAAAT
TTCGTAGCCGAATCCATAACCCCATCCTTGAAAACTTGGGAAGAAGTGGTCAATAATATATGTAGTACTTTGAAGCC
GTCTGGAACCTTAATAATGACCGCGATCCAAGGGGCTAGTTTCTACTGTGTCGAAAATCACAGATACCCCGCGATTG
CCGTTACGCCAGAAGATGTCATCAGGGTTTTATCATATCAAGGCTTCGACGTAGATAACTTATTAATGCGTCATATA
CCGGCGGAAATCACAGACATCTCAGCGAAGGACTACAAGGGCTATCAAGGCATGTTATTCGTCAAGGCCACGAGA

Seq. ID No: 91
>INMT_24
ATGGAAAGTGGATTCACTTCCAAAGACACCTATTTGTCTCATTTCAATCCCAGGGACTACTTGGAGAAATACTATTC
ATTTGGTAGCAGACACTGTGCAGAGAACGAAATTCTACGTCACTTGCTGAAAAACCTTTTTAAGATATTCTGCTTAG
ATGGCGTTAAGGGAGAGCTGCTGATAGACATAGGAAGTGGCCCAACCATTTATCAATTGTTGTCCGCTTGCGAATCA
TTCACTGAAATAATCGTAACAGACTACACAGATCAAAATCTATGGGAACTACAGAAGTGGCTTAAAAAGGAGCCCGG
TGCTTTTGATTGGAGTCCTGTAGTTACGTACGTGTGCGACCTTGAAGGCAATCGTACTAAAGGGCCGGAAAAGGAAG
AAAAGCTGAGGAGAGCCATTAAACAGGTCCTTAAATGCGACGTGTCCCAAAGTCAACCCCTTGGAGGGGTCTCTCTT

CCACCCGCAGATTGTCTATTAAGTACACTATGCCTAGACGCTGCCTGCCCAGATTTACCGGCTTACAGGACCGCCTT
ACGTAATTTAGGATCTCTACTTAAACCGGGCGGGTTCTTAGTCATGGTTGATGCACTCAAGAGTTCTTACTACATGA
TCGGGGAACAGAAGTTTTCCAGTCTTAGTTTGGATAGAGAAGCAGTGAGGGATGCTGTGGAAGAGGCTGGTTACACA
ATCGAACAGTTCGAGGTCATATCACAATCCTATAGTTCTACAACTAGCAACAACGAAGGCTTATTCTCCCTGGTCGG
CCGTAAGCCCGTGGGGAGCGAG

Seq. ID No: 92
>INMT_25
ATGGAGATCGTCAGCACTTCATACAACCATGTGTATGAAAACTTTGATGCAAGAAAATATCTTGACAGATACTACGG
CATTGCGCCCGAAGCCGAGAAAATAGATGAGGAATCCGTGTTCTTGTTAACCTTCCTATCCAATGTATTCTCCTCAG
GGCGTGTCAAAGGTCACTCTTTCATAGAGATTGGGGTCGGGCCTAGTATCCACTCCATCCTTTCTGCTTGTGAAGCG
TTCGAAAAATTGTACTTGACTGACTACTTCCAGGGAAACCTGGATGAGATTAAAAAGTGGTTAAACTCTGAGAACGA
CGCATTCGACTGGACGCCATACATTAGATTTGTTTGTGATTTGGAGAATAATGGGAGTACACCGAGGGAGAAGAAGG
AGAAAATCCGTCGTTGCGTATCCCTAATGAAGTGTGACGTGAACCTAAGTAACCCGCTTCATCCCCACAGTCTTCCT
TTAACGGATTGTCTTCTGACCGCTTGCTGCCTTACCTCTACTTGCAAGACGTTTACTGATTTTAAAATGAGCTTAAA
GACGATTGTCTCCCTAATAAAGCCCGGCGGCCATCTAATTCTTATAGACTATCTGCGTGCTAGTTACTATTGGGTAG
GCGAAGCCAAATTGCCGTTACTGAGCTTGGACGAGCATGGAGTCAGGGAGGCCGTGGAAGAATCTGGATGTAAGATC
GAAGAATTTCAATGGTTCAAGGAGTTCCACATGCCTGACGAGGTCAGTGATTGCAAGACGGTGTTTATACTATTAGC
TCAAAAACTA

Seq. ID No: 93
>INMT_26
ATGAGGAACCTGCACGAGACTTTTGGCCCCGGCGGCGTCAAGGGAGATATTCTAATTGACTTCGGGGCTGGTCCGAC
GATATATCAGTTGTTATCAGCGTGCGAGGTATTTAACACAATTATTACCTCAGACTTCTTAGAACAAAATCGTGAAC
AATTGAAAAAGTGGTTAAGGAAAGATCCAGATGCACTAGATTGGAGTAATTTCGCCAAATACGTTTGCGAATTGGAG
GGTAAGTCCGACAATTGGGAGAAGAAGGAAGAGACGCTGAGACGTAAGGTCACTAAAGTTCTGAAGTGTGACGCCTT
AGCAGAGAAGCCTTATGACCCGGTCCCAATGCCTGAGGCGGACTGCCTGATATCTTGCCTATGCTTAGAAGTGGCCT
GCAAAGACCTTGAGGATTTC

Seq. ID No: 94
>INMT_27
ATGGAGACTCCTTTTACTTCACAGCAAACATACGTCGATGAGTTCAAAGCATCAGATTACTTTAAAACGTATTACGT
TGCAGAGGGCGGCATTGCCAATGAAGAATGGACCGATTTTGCACTAAGGACTCTGCATGAGACTTTTACGAAGGGAG
GTGTGAAGGGCGAAACGCTTATAGATTTCGGCGCTGGCCCTACTATCTATCATCTACTGTCAGCGTGCGAAGTGTTT
GATAAAATCATCACTAGCGACTACCTTGAGCAGAACAGAGCCGAGCTTGAAAAGTGGCTGAAAAAGGACCCGAGTGC
CTTTGATTGGACACCAATAATCAAGTTCGTCTGCGAGCTAGAAGGAAACAGGAATTATGAGAAAAAGGCTGAGAAGT
TGAGGAATAAGGTAAAAGAAGTTCTAAAATGCGATGCACTTAAAAGAAATCCGTTTGACCCTATTGTGTTGCAGCCG
GCGGACTGTTTACTAACTTGTCTATGCCTTGAAGCACCATGTGAGGATATGAAGAGTTACTTTAACGTGCTAAAGAA
TTTCAAGGATCTGATCAAACCAGGCGGACATTTGGTTATCCTATCTGTTCTTGACGCGACGTTCTACTATGTTGGAG
ACAAATACTTCAGCAGCATGACAACAAGGAAAGAGGAACTGGAACAAGCCTTGAAAGAAGCCGGTTTTGAGATTGAA
AAAGCCGTTTATACCACAAGGAAGGACAGATCACAGATGGATATAGCGGATTATCAGGGTTTTTACTACATACATGC
GCGTAACCCCAAG

Seq. ID No: 95
>INMT_28
ATGGAGGGGAGCTTTACCGGAGGTGAGGAGTACCAAAAATACTTCCAGCCGCGTGACTACCTTACTACATATTATAA
CTTCGATGGTTCCCCTACCCCAGAAGCGGAAATGTTGAAATTCAATTTGGAATGTTTACACAAGACCTTCGGACCTG
GTGGTTTAAGGGGAGATACCCTAATAGACATTGGGTCAGGACCAACGATTTACCAAGTCCTTGCCGCTTGCGAGAGC
TTTCGTGACATTACTCTTAGCGATTTTACAGACAGGAACAGAGAGGAGCTGGAAAAATGGCTGAAGAAAGAGCCTGA
GGCTTACGACTGGTCATCCGTTGTAAAATTCGCCTGTGAGCTTGAGGGTGATAGCGGGCGTTGGCAAGAGAAAGAGA
AGAAGTTGAGATCAGTGGTTAAGAGGGTCTTGAAATGCGATGCAAATTTGGCGTCTCCACTTGCACCTGCAGCCCTT

CCACCCGCCGACTGCGTTCTTACGTTACTGGCGATGGAGTGCGCCTGTTGTTCACTGGACGCATACAGGGCTGCATT
ATGCAATCTGGCCTCACTTCTAAAGCCCGGAGGTCACCTAGTGACCACGGTGACCCTAGGCATTTCATCCTATATGG
TTGGCAAGAGGGAATTTTCCTGTGTAGTCCTTGAGAAGGAGGGTGTGGAACAAGCTGTCTTAGACGCGGGTTTTGAC
ATCCAGCAGTTTCTACACATCCCTAAATGTTACTCTGCAACTATAGCCGCGAACAATGGTGTCTGCTTTATCGTGGC
TAGAAAGAAGCCTGCCCCC

Seq. ID No: 96
>INMT_29
ATGGAGGGGTCCTTTACGGGACCCGATGAGTATCAAAAATACTTCTCACCTAAGGATTACCTTGATACTTACTATAG
TTTCGAACATGGACCATCACCCGAGACCGAAATGATAAAATTTTCTCTTCAATTCCTTCATAAAGTCTTCGGGCCAG
GTGGCATCCGTGGTGAGACCCTGATAGATGTCGGAAGCGGCCCGACGATATATCAGGTCCTGGCGGCCTGTGAGGCG
TTTTCTGACATAACTCTTTCCGATTTCACGGACCGTAATAGGGAAGAACTACAAAAGTGGCTAAGAAAGGATGCCGG
AGCTTTTGATTGGACTCCCGTGTTAAAATTCGCTTGCGAACTTGAGGGCAACAGCTCACACTGGCAGGAAAAAGCAG
AAAAGCTAAGGGCAACTGTCAAGCGTGTATTGAAATGTGACGTTAACCTGGGTAAGCCTCTTGCGCCAGTCGAGTTG
CCAGCGGCGGATTGTGTCCTTACGCTTCTTGCAATGGAGTGTGCCTGCTGTTCTCTGGCCGCATATAGGGCGGCGCT
TTGTAATCTAGGATCTTTGCTGAAACCAGGAGGTCATCTAGTAACATCCATTACACTTCAAATATCCAGTTACATGG
TTGGAAAACACCAGTTCTCTTGTTTATATATAACAAAAGAAGAAGTGGAGAGGGCAATCCTGGACGCAGGTTTCGAT
ATTGAACAGTTACTTCACAGTGAGCAGAGTTACAGTGCTACGATCGCTCCAAATAAGGGTATTTGTTTCATTGTGGC
TAGGAAGCGTTCCGGTCCG

Seq. ID No: 97
>INMT_30
ATGGACGCCCAGCTAACTCAACTGAGGAACGCTGATGTAAGCTGGGCGGCTTTCGACCCAATCGCTTACGTTGACCA
CAACTATAGGGATCTACAAGCGGAGGATGCTGAAATTCTTCATCTTGTTAGGGATCACTTCGGTGACCACTTTAGGA
AACAGGGTGGGGGACCAGTCTCAGGAATAGATGTCGGTGCAGGCGCCAACCTTTACCCCGCTTTGGCAATGATGCCC
TGGTGTGAAGAGATAACCCTGTTTGAGCGTAGTCCGGCCAACGTACGTTATCTAAAGAGTCAAGTGGATAGCTATGA
TGCCAACTGGGATCAATTCTGGGACGCCTTGTGTGCCCATGAAGCATATAATAGCCTGGGTACTGACCCAAGAGAGC
GTTTCGGGAAAGTAGTCTGGGTAGAACAAGGTGACCTTTTCGACCTAGCGAGATACGAAAGGAGGTGGTCAATGGGC
ACTATGTTTTTCGTAGCGGAGTCAATGACTACATCATATCAGGAATTTATGTTAGGAGTCGAGAGATTCATGCGCGC
GCTAAGCCCTGGAGCACCCTTTGCGGCTGCGTTTATGGAACATAGTAAAGGTTACCATGCGGGGGAGCATTTCTTCC
CGGCGTGCGACGTGGGTGAGAGCGAGGTCAGAGCGTCTCTTGAGGGGTTTGCCGGAGACTTTAAAGTGCAGAGGTTG
GAATCCGCAGCGCAGCTTCGTGATGGATACAGTGGTATGATCGTGGCATAT

Seq. ID No: 98
>INMT_31
ATGTCAGATTTTACAAACGCCAGTGAGTACGAAAAGCAGTTCGACCCGAGACTATACCTTGAAACCTATTTTCATTT
AGGTTCAGGGTCATTAGCAGACGATTTCCTTAGGTTCACTTTGGGCAATTTTCATAAGACCTTCACAGAGGGCGAGG
TGAAAGGCACGACATTAATTGACATTGGAACAGCCCCTTCAATCTATCAATTGTTGTCAGCCTGTGAATACTTTCAG
GATATCACCGTAACGTGGTATACCAATAGAGAACTACAAGAGCTACAGAAATGGTTAAATAAGGACCCCGGCGCCTT
TGACTGGTCTTCCACGGTAAAACACGTTTGGGAGCTAGAGGGTAAACGTGGGATGTTGGAGGAGAAAGAGGAGAAAC
TTCGTGGTATGATTAGACAAGTGCTTTTATGTGATGTGTCAAAGAAGAACCCTTTAGAACCCGTTACTTTGCCGAAA
GCGGACTGTCTAATATCAACCGTGTGTTTGGAAGCGGCGTGCCGTAACTACGACTCTTATCGTACAGCCCTAAAGAA
TCTAAGTACGCTTCTAAAGCCAGGTGGACACCTTTTGCTTGCGGGGGACTTAGGAGCCAATTACTACGAAGTTGGTT
CTAATAAGGTTTTCAGTCTTCCCGTCAACGAGACTTTTCTAAGGAAGGCCGTCAATGAGAGCGGTTATGTCATCAAC
AAACTTGTGTCCTTCGGAAAGCCTGAAGACGCAGGGTACGACACTTCCGATTATGAAGGGTTTTACTTTATTCACGC
TCAAAAATGT

Seq. ID No: 99
>IOMT_1
ATGAGTAGTAAGTTAGACAATCAAAATATTACCGCTAATGAAGAAGAGGAGGCTTTCCACCAGGCCATGCAGCTGGC
CATGTCCACTATTTTACCGATGGTCCTGAAAGCCGCAATAGACCTAGACCTGTTAGAGATTATAGCGAAAGCTGGAC

CGGCGGGATGCAAACTTTCTCCTATAGAGATCGCTAGTCACCTTCCCACCAAAAATCCAGATGCGAGCAGCATCATC
GACAGGATTTTAAGAGTTCTGGCCTCACATTCTATCCTGACATGTGATCTTGCTACCAACGAGGACGGTCACGTACA
ACGTTTGTATGGTCTGGCCCCGATTGCGAAATATTTCTTACACAATGATGATGGGATAAGCCTAATCCCCACTTTGA
CCATTTCAACGGATAAATACCTACTTGGTGCGTGGTATCATCTACGTGAGGCAACGTTGGAAGGCGGGGCAATCCCA
CTGGTCAAGGCGTATGGAATGGACTTATTTGAGCTGGCGGCGAAGAACGACGAAATAAGTGGGAAGTTCAACAACAC
AATGGGTAACCAAACAGCGATTATCATGAAAAAGGTGTTGGAAATTTATAAGGGATTTGAAGGCATTAACCAGTTGG
TGGACGTGGGAGGCGGACTTGGTATCAACTTGAAATTGATCGTATCCAAGTACCCTCAGATCAAAGGTATAAACTTC
GACCTACCGCACGTAGTAAAGGACGCTCCTCATTTCTTAGGCGTTGACCATGTGGGTGGTGACATGTTCATCGAGGT
TCCACAGGGAGAGGTAATATTCATGAAGTGGATTCTGCACGACTGGGGAGACGATAGATGTCTAAAGCTGTTGAAAA
ATTGTTATAATGCTCTACCAAAATTCGGAAAAGTAGTTGTGGTAGAGCTAGTGGTGCCTGAATCTCCCATGACAGAC
ATCGTCACAAAGAATACGTTAACGTTGGATGCAGGCCTATTCATTGTGGTTCCAGGTGCCAAGGAAAGGACAAAAGA
GGAATACGAGGCACTAGCCAAGAAGGCGGGCTTCTCTACGTTTAGATTGGTATGTAGGGCCTATTCTTATTGGGTGA
TGGAGTTCCACAAAAACGTGATAGTG

Seq. ID No: 100
>IOMT_2
ATGGGAAGCCAAGCCGAGGTTGGGAAGGCGATGACGGAGGAAGAGGCTTGCGAATTTGCAATGCAGCTAGTCTCTTC
CAGTATCTTGCCTATGACTCTTAAAGCAGCTCTTGAGCTTGAGCTTCTTGAAATTATGGCAACGGCTGGTGAAGGAG
CCCAACTGACTCCAGCGGAGATTGCTGCTCAACTTCCTACCTCTAATCCAGACGCCCCTATAATGCTAGACAGGATG
TTAAGACTTCTTGCCTGCCACTCTGTATTAACGGCATCAACTTATACCGACGACGATGGGAAGGTCAGAAGACGTTA
TGGCTTGGCGCCTGTATGTAAGTTTTTAGTCAGGAATCAGGATGGTGTTTCTACGGCCGCGCTATCCTTGGTGAACC
AGGATAAGGTGACGATGGAGTCTTGGTATTACTTAAAAGACGCCGTACTGGAGGGCGGTATTCCATTTAACCGTGCT
CATGGGATGACGGCTTTCGACTATCCGGGGACTGATCCTAGGTTCAATAGGGTGTTTAACCAGGGAATGTCTAACCA
CTCAACCCTGACCATGAAGAAAATCCTAGAGACTTATACTGGCTTTCGTGGCCTTCACTCCTTAGTCGATGTAGGCG
GTGGCATCGGAGCTATTCTTTCCCTAATCGTCGCCAAATTCCCTCATATCAAGGGGATTAATTTCGACCTACCACAC
GTTATAGACGATGCGCCACAATTCCCAGGTGTTGAACACGTGGGAGGCGACATGTTTGCATCCGTGCCAACAGCCGA
GGCCATTTTATTAAAGCTAATTCTACATGATTGGGGTGACGAACATTGTGTGAAATTGTTGAAAAACTGCTGCAAAG
CGTTACCGGAAGATGGAAAGGTAGTCGTAGTGGAGGCCATTCTACCAGAAGGGATTGATCACTCTTATGCGTCAGCT
TGCGTTTATCAAGTCGATATGATTATGTTAGTCACAAACCCCGGGGGAAAGGAAAGGACCTTGAAAGAGTTTGAAGA
ACTTGCCAAGGCCGGAGGGTTCGCTGGCATAAGGCCAATTTGCTGCGTCTATGGAAGCTGGGTGATGGAATTTTACA
AGAAGATG

Seq. ID No: 101
>IOMT_3
ATGGGATCAACAGCAGAAACTCAACTGACCCCCGTACAGGTCACCGATGACGAGGCTGCCTTGTTTGCAATGCAATT
AGCATCCGCATCCGTCCTACCAATGGCGTTAAAAAGTGCACTGGAGCTGGACCTGTTAGAGATCATGGCGAAAAATG
GGTCTCCTATGAGCCCGACTGAAATTGCTTCCAAGCTACCTACTAAAAACCCCGAAGCGCCTGTAATGTTGGATAGG
ATACTGAGGCTATTGACAAGCTATTCCGTCCTTACCTGCTCAAATAGGAAACTTTCTGGTGATGGCGTTGAAAGGAT
ATACGGATTGGGACCGGTCTGTAAATATTTGACGAAGAATGAAGACGGAGTGAGTATCGCGGCGTTATGCCTGATGA
ATCAAGACAAAGTGCTGATGGAATCCTGGTACCACTTAAAAGATGCAATACTGGATGGCGGCATCCCGTTTAACAAG
GCATATGGAATGTCCGCGTTTGAGTACCACGGGACAGACCCACGTTTCAACAAGGTCTTCAACAATGGCATGTCAAA
TCATTCAACGATCACAATGAAAAAGATACTAGAGACCTACAAGGGCTTTGAGGGGTTGACTTCATTGGTGGATGTTG
GTGGTGGAATTGGGGCGACGTTAAAAATGATTGTGAGTAAATACCCTAACCTAAAAGGGATCAACTTCGACTTGCCT
CACGTCATTGAAGATGCACCGAGTCATCCTGGCATCGAGCACGTAGGAGGGGATATGTTTGTTAGCGTCCCGAAGGG
AGATGCTATTTTTATGAAATGGATATGTCACGATTGGAGCGACGAACATTGCGTTAAGTTCCTGAAGAATTGCTACG
AATCACTTCCTGAGGACGGCAAAGTGATCCTTGCCGAATGCATTCTACCAGAAACTCCTGACTCCAGCTTATCTACT
AAACAAGTTGTACACGTAGACTGTATAATGTTGGCTCATAACCCCGGGGGTAAGGAAAGGACCGAAAAGGAATTTGA
GGCTTTGGCAAAGGCATCAGGATTCAAGGGAATTAAAGTTGTTTGCGACGCATTCGGTGTGAATTTAATCGAGCTAT
TGAAGAAACTA

Seq. ID No: 102

>IOMT_4
ATGGGCTCCACAGCTGCAGATATGGCGGCTAGTGCCGACGAGGAGGCTTGCATGTACGCTTTGCAGCTGGTCTCCAG
CTCTATATTGCCCATGACATTGAAGAATGCCATAGAGTTGGGCTTATTAGAAACTTTGGTCGCAGCCGGGGGAAAGT
TATTAACCCCAGCAGAGGTTGCGGCAAAACTGCCCTCAACCGCAAACCCTGCCGCTGCCGATATGGTTGACCGTATG
CTAAGATTGTTAGCCTCCTACAACGTTGTTTCTTGCACTATGGAAGAGGGGAAGGACGGCAGACTAAGCAGGAGATA
CCGTGCGGCGCCCGTTTGCAAATTCTTGACTCCCAACGAGGATGGTGTAAGCATGGCAGCCCTAGCTCTAATGAACC
AGGATAAAGTATTGATGGAGTCTTGGTACTACCTTAAAGACGCAGTCTTAGATGGTGGTATCCCGTTTAATAAGGCT
TATGGCATGTCCGCGTTTGAGTACCACGGAACGGACCCTCGTTTCAACAGAGTGTTTAACGAGGGTATGAAAAATCA
CTCCATTATTATAACAAAGAAACTGCTAGAAGTGTATAAGGGCTTTGAGGGATTGGGAACTATCGTTGATGTTGGGG
GTGGAGTCGGAGCCACTGTTGGAGCGATAACAGCCGCTTATCCAGCGATTAAAGGCATCAACTTTGATCTTCCGCAC
GTCATCTCTGAGGCACAGCCGTTCCCAGGAGTGACACACGTTGGCGGGGACATGTTCCAAAAAGTGCCTTCAGGGGA
CGCGATTCTGATGAAATGGATACTTCACGATTGGAGCGATGAGCACTGTGCTACCCTGCTGAAGAACTGTTATGACG
CACTGCCCGCACATGGAAAGGTTGTACTGGTGGAGTGTATATTACCGGTTAATCCTGAGGCGACGCCTAAGGCGCAG
GGGGTTTTCCATGTAGACATGATAATGCTGGCTCATAATCCGGGAGGCCGTGAGCGTTACGAGAGAGAGTTTGAAGC
ACTGGCAAAGGGAGCCGGGTTCAAAGCCATAAAAACAACCTACATCTACGCCAATGCCTTTGCGATCGAATTTACTA
AG

Seq. ID No: 103
>IOMT_5
ATGGGCTCAGCTGGGGAAACCCAGATTACCCCGACTCATGTGAATGACGAAGAGGCCAACCTTTTCGCTATGCAGTT
AGCAAGCGCGAGTGTGTTGCCGATGATCTTAAAATCTGCGTTGGAACTGGACCTTTTAGAGATAATCGCGAAGGCCG
GGCCCAATGCGCAGCTATCTTCTAGTGACATCGCTTCACAGCTGCCTACCAAGAATCCAGACGCGGCAGTGATGCTG
GACCGTATGATGAGGCTGTTAGCTTGCTACAACGTATTGAGTTCTTCTTTGAGAACACTTCCTGATGGAAAAATCGA
ACGTCTATACGGTTTGGCTCCAGTTGCGAAATACCTGGTAAAGACGGAAGACGGTGTGTCTATCGCACCACTGAGTT
TAATGAATCAGGACAAAGTGCTGATGGAAAGTTGGTACTACCTAACCGAGGCGGTGTTGGAAGGCGGCATCCCATTT
AATAAGGCGCATGGTATGACCTCCTTCGAGTACCACGGAAAAGACGCTAGATTCAATAAGGTGTTCAATAAGGGGAT
GGCTGACCATAGTACGATTACAATGAAGAAGATTTTGGAGACATACACAGGGTTCGAGGGACTAAAGTCTTTGGTCG
ATGTGGGAGGCGGCACTGGTGCAGTTATTAGCATGATTGTAAGTAAATATCCGTCCATCAAAGGGTTCAACTTCGAC
CTACCACATGTGATAGAGGAAGCACCGTCATATCCCGGCGTCGAACACGTAGGTGGGGATATGTTTGTGTCAGTTCC
GAAGGCTGATGCTGTTTTCATGAAGTGGATCTGCCATGACTGGAGCGACGAGCATTGTGTAAAATTTCTGAAAAATT
GCTATGATGCTTTACCTGAAAACGGAAAGGTGATCGTGGCAGAGTGCATCCTTCCAGTTGCCCCAGACTCTAGCCTA
GCAACAAAGGGCGTCGTCCATATCGATGTCATAATGTTAGCCCACAATCCTGGAGGGAAGGAGAGAACCGAGAAAGA
GTTCGAAGCTCTAGCAAAAGGTGCGGGGTTCCAAGGGTTCAGGGTATGTTGCTCTGCTTTCAATTCATATATTATCG
AATTTTTGAAGAAACCG

Seq. ID No: 104
>IOMT_6
ATGGGATCTACTGCGGAGACCCAAATCACTCCAGTCCAGGTGACCGATGATGAAGCTGCTTTGTTCGCCATGCAGTT
GGCGTCCGCATCTGTTCTACCTATGGTCTTGAAGTCTGCGTTGGATCTTGATCTGTTAGAAATAATGGCGAAGAACT
CTTCACCGATGTCCCCGTCCGAAATCGCGTCTAAGCTACAAACCAAAAACCCCGAGGCACCAGTCATGTTGGATCGT
ATCCTACGTCTATTAACCTCCTACAGCATTCTTACATGCTCAAATCGTACGATTCTGGGTGGTGACAGTGTGGAGAG
AATTTATGGTCTGGGACCCGTCTGCAAATATCTTACAAAGAACGAAGACGGTGTCTCCATAGCCGCTTTATGCCTAA
TGAATCAGGATAAGGTACTTATGGAATCATGGTACCACCTTAAAGACGCTGTCTTGGATGGCGGCATACCTTTCAAC
AAAGCCTATGGCATGTCAGCCTTTGAGTATCACGGCAAAGATCTGAGGTTTAATACAGTTTTTAACAATGGAATGAG
CAATCACAGCACGATCACTATGAAGAAGATTCTGGAGACATACAAGGGCTTCGAGGGTCTAACTAGCTTGGTTGATG
TTGGGGGCGGAATAGGTGCGACTTTAAAAATGATAGTGTCCAAGTACCCGGACCTTAAAGGCATCAACTTTGACTTG
CCGCATGTGATAGAAGAAGCGACCAGTCATCCTGGGATTGACCACGTAGGTGGCGACATGTTCGTTAGTGTGCCAAA
AGGAGATGCCATATTTATGAAGTGGATATGCCATGACTGGTCCGACGAACACTGTGTTAAATTCCTTAAAAATTGCT
ACGAAGCGTTGCCCGAGGACGGAAAAGTTATCTTGGCAGAGTGCATTCTGCCCGAAACACCCGACTCATCTCTTTCC
ACAAAGCAAGTCGTCCACGTAGACTGTATAATGTTGGCCCACAATCCAGGGGGAAAAGAAAGGACAGAAAAAGAGTT

TGAAGCGTTGGCAAAGGGCTCAGGTTTTAAGGGAATCAACGTGGCATGTAATGCATTTGGCGTGTATGTGATAGAGC
TGTTGAAGAAAATG

Seq. ID No: 105
>IOMT_7
ATGGAAATGATAAACTTCATGCATCATATGGATTCAACGTGGAATCTTTGTGGTAAAGATGTAGTGCAAGCCTTTGA
CTTTTCTGAGTTCCACACGGTTTATGATCTAGGTGGCTGTTCAGGCGGCCTAGCGAAGCAGTTCGTCTCCACGTATA
ACGATTCCACGGTAACGATTATGGATCTACCAAAGGTAGTACAAACTGCCAAGAAGTACTTCGTGACGGATCAAGAA
CAACAGATTCACTTCATCGAGGGCGACCTTTTCAATGACCCCATCCCTGAGGCGGACCTGTTCATTATGGCCAGGAT
AATCCATGACTGGACAGAAGAGAAATGCCTAGAGCTGCTGAGGAAAATTTACCAGTCCTGTCGTCCAGGTGGTGGTG
TGCTGTTGGTCGAAGTTCTTTTGAATGAGGATAAGTCTGGACCCCTAATGAGCCAATTATTCAGCTTGAACATGCTA
GTTCAAACAGAGGGCAGAGAGCGTACACCAAGTGAGTACACAAAACTTCTGACCGACTCAGGCTTTCGTGACATTCA
AGTGAAGATTACTGGGAAAATATACGATGCA

Seq. ID No: 106
>IOMT_8
ATGGAACGTCTACTTGATGCCTGTGTGGGCTTAAAACTTTTGAAAGTTGAGCTAAAGTCCAACAAGGGTTACTACAG
TAACACAGATGTCTCTACCATGTACCTTGTAAAATCTAGCCCCCGTACTTTATACTATATGATAATGTTCTACTCCA
AGACTACTTACATGTGTTACAACTTCTTACCTCAAGCAGTCAGGGAGGGTCAATGTCAGTATGAGAGAGCCTTCGGC
ATATCCTCTAAAGACCTATTTGAGGCTCTATACCGTTCTGAAGAGGATACTTTAGCATTCATGTATTTCATGAACAG
CACATGGTCTATTTGTGGAAAGTATGTGGTCCAGGCGTTCGACCTGAGTGAATTTCACACAATTTATGACTTAGGCG
GCTGTACGGGCGCTCTAGCTAAGCAGTTAGTCAGTACGTATAAAGAGAGCACCGTCACTATTATGGACATGCCAAAC
ATCGTTCAAGCCGCCAAAAAGCACTTTGTTACTGACAAGGAGCAGCAAATTCATTTTCTTGAAGGCGACTTCTTTAA
TGATCCAATACCAGAAGCT

Seq. ID No: 107
>IOMT_9
ATGATCCCATTCAATAAGGCTTACGGTATGACCGCCTTTGAGTACCACGGAAAGGACGATAGATTTAACAAGGTCTT
TAACGCTGGCATGTTTAATCACTCAACGATGACCATGAAGAAGATCCTGGACATCTACGACGGCTTCAACAACCTAA
CGACGCTAGTAGACGTGGGAGGAGGAACTGGGGCTTCTCTAAATATGATCGTGAGTAAGCATCCCAGCGTAAAAGGT
ATTAATTTCGACCTTCCACACGTCATTCAAGATGCCACTACTTATCCGGGCATAGAGCACGTAGGAGGTGACATGTT
TGAATCTGTTCCAAAAGGTGATGCCATCTTTATGAAGTGGATCTGTCACGATTGGTCAGATGCCCACTGCTTAAAAT
TTCTTAAAAATTGCTACAAGGCGCTACCAGACAATGGTAAAGTCATTGTAGCC

Seq. ID No: 108
>IOMT_10
ATGGCACAAGCCGCGGCGGAGGCCGAAGGCATCACCCCGGTCATGGACTTACTATTCGCAGCACAAGGGTCCTCTGC
CTTGCTAGTCTGCGCTAGGTTAGGCCTATTCGATTACATCTCCAGCCAGGGTGAAGAAGGGGTGAGCTGCAAGCAAC
TGGCGTCCAGGGCGCAGTGGTCCACACGTGCTGCAAGTGCCGTTATGGTTAGCCTTGCTGCGAGTGGAATACTAGCT
GTAAAACCTAGTTCAGCCGGCGCTCAACACTGCTTTGAGCATTCCTACACACTTACGCCTAGGGCCCAAAGGTTCCT
GGTTACCGAGAAACCTGGGAGCATGAGCGCGTATACCGAGATACACTGGGAAGCATCCCCGGAACTACTTCTGAAAA
AGGCCGCTGAGACTGAGGACGAAAAACGTAACTTTATGCTAGAGACAGGGGGAGGAGCCCCCTCTGAGGTATTCCTA
GCGGCGATGCAGGGCCAGAGTAGTTACGCTGCGATGGTGTTGACAAGGTTAGTAGACTTATCGGACACCAGGACATT
CGTCGATGTAGGCGGTGGAAGTGGCACTTTCGCAATAGAAGCCTGCAAGGCGACCCCCAACCTTCAAGGAGTTGTGT
ACGACCTAGCCGGGGCATGTCCTACCACCGACGGTTTCATTGCAAGGGCAGGTATGGCTGAGAGGGTGAAGACACAT
GCGGGGAACATGTTCGAGGATGAGAGGTTCCCTGCGGCTGACTGTTATGCATTCGGGAATGTGCTTCATGATTGGTC
AGATCAAGATAATAGCAAGTTACTTAGGAAAGCATTTGAGAGCCTTCCTGCCCAAGGTAAAGTGCTTCTACTAGAGA
TGCTGGTGGAAGAAGACGTAGTGAGTACCTCACCCTCTGCGGCCGGACTAAATTTATGCATGGTGACTAACGAGCTG
GGGAGACAGTTTAAAGCAAGTGAACTACGTGCAATGTTGCTAGAGGCAGGGTTTGCAGGGGCCGAGGTCGTTTCATC
ACCTCTGACTCCTTACTCCCTAGTCGTGGGAACGAAAGGTGAGGCGAATCCTGTGGCGAGCAAACCCGAGGCCGCTG
CGGCTGCGGAATCAGAGAGCATTACTCCGCTGATGGATGTGTTGTTTAGCGCCCAACACAGTGCTGTACTGATAGTT

TGCTCTAGGTTAGGCGTGTTTGATTTTGTTGGCGCGCAGGGCGAATCAGGTGCTTCTTGTGCTCAGGTGGCGGCACA
TGCAAAATGGACCACCCGTGCGGCTTCAGCGATGCTTGTTAGCCTAGCTTGCAGCGGCCTGCTGGAGCCTACACCGG
GAAGCGCAGCGGCCCAACACTGCTTCGAACATAGCTATAGGCTTACTCCTCTAGCTAGGAGGTTCTTGGTAGCTGGT
CAGCCAGGACAACTGAGTGCTTACACTGAAATATTTTGGGGCGCTAGTCCCAAACAATTATTAGAAAAGGCGTCCGC
TTCACTGGGCGAGTGGGGCGAGGGAAATTTTATGTTAGACGCCGAGGGCGGAGCCCCTTCAGAGGTATTTCTAGCAG
CCATGCAGGCGCAGTCAACCTACGCCGCAATGGTTCTTACACGTTTAGTCGATCTTTCTGATGTGCGTACATTCGTT
GATGTCGGTGGAGGGTCAGGTACTTTGGCCATCGAGGCCTGTAGAGCCGCTCCAGGCCTACAGGGTGTCGTGTATGA
TCTGGCAGGTGCGTGTCCCGTTACCGACGGTTTCATTGCCCGTGCTGGAATGGCAGAACGTGTGAAGACTCACGCCG
GCAACATGTTTGCTGATGAAAGATTCCCCGCCGCTGATTGTTATGCATTTGGTAACGTGCTGCACGACTGGTCAGAC
CAAGACGATGGAAAATTACTAAGAAAGGCGTTCGAAAGTTTGCCTGCGAACGGTAAGGTCTTGTTATTGGAAATGCT
ATTAGCCGAGGACGTGGAGTCTTCAACGAGGAGTGCAACAGGACTAAACATTGTAATGGTGACCAACGAGCAGGGGC
GTCAGTTCAAGGGCTCCGAGCTTGAGGCTATGTTACGTGCTGCTGGTTTTGCTGCGACAGAAGTGGTTCGTTCACCA
CTAACTCCCTACGCATTGGTGGTTGGTACTAAAGGC

Seq. ID No: 109
>IOMT_11
ATGAGCAGGACATCATGGGACGAAGGCGAGGATGTCGATCTGGATTCAGTTGCTTATGGTTTCATGGCAAGCCAAGC
ATTGTTTACTGGTCTGGAGCTTGGCATTTTTGATCACATCGCTGCAGCGGGAGCTGGTGGCTTGTCAGCGGCTGGCA
TCGGGAAGGCTTGTGGGATAGAAGCGCCCAGGGTTCAAACGCTGCTTACTTCATTGGTCGCAGTTAAGTGCCTGAAG
CGTGACGCATCTGCCATGTACACCTTGAGTCCAAATACGGCACAATATATGGTAACTAGCAGTAGACATTTTTATGG
CGACTATTTAAGGTATCAAATAGGGCGTCAGTTTTACCACCGTATGGGGGCCTTGCCTGAGGTGATGACGTCTGGTA
AAGCCCCGAGTTACGCATCATGGTTTAGTGATCCTGAAGTAGCTAGAACTTACACACAAGCACAGCATAACGGTTCC
GTTGCCACAGCCAAGTATTTGATAAAGAAGAAATTGCAACTAGGGGGCATATCAGCGATGTTAGACGTCGGCGGGGG
TTCAGGGGCGTTTTCCTATGTCTTTACTCAGGCCACACCGGGTTTGCACAGTAAGGTACTGGAACTGCCAGAGGTCT
GCCGTACCGGTGAGGGTATCCGTGAGAAGCAACCCGAGGATGTCCGTAGCAGGGTTAGCTTTGTCGAGTTAGATGCG
AGTAGCCCGACGTGGCCGGTTGATGACTCTGCGTTTGACGTCGTGCTAATGTCCTATATATCAGGCTCAGTTCCCGA
GCCGATCATTGGATCATTATACGCCAACGCTATGAAGGCTTTAAGGCCCGGAGGACGTCTTCTTGTGCATGACTTTA
TGGTCAACGACTCACTAGACGGACCGGCGTTAGGGGCTCTGTGGGGATTGCAGCACGTCACAGTCAACGCGGACGGT
TTGGGACTGTGCCCAAAGGAGGTTATCGCGCGTATGGGAGCAGCAGGCTTTGACACAAGTAAATGTGAAGCAATGGA
GATGATTCACGGGATGACTAAGCTGATAGTCGGACACAAGGGC

Seq. ID No: 110
>IOMT_12
ATGTGCTCTAGCAAGGAACTAGATTTTCCGCATATCTTGATTGATTACCAGCACGGCTTTCTAGTATCCAAGACGAT
ATTCACGGCGTGTGAACTAGGCGTATTCGACCTATTACACGAGGTACAAGAGCCCGTGCCCGCTGCGACTATTGCAT
CTAGGCTTAGCACATCTGAAGATGGAATGGAGAGATTGTTAGACGCTTGCGTAGGCTTAAAGCTACTGAAAGTCTAC
CTGAAGAATAATAAGGGCTACTACTCAAATACTGACGTTTCCACTATTTATCTAGTTAAGAGTTCACCTAAAACCTT
ACATTACATGATGATATATTACTCCAAAATAACGTATATGTGCTGGCATTTCCTACCACAGGCCGTTAGAGAAGGTA
AAAGACAATATGAGAGGGCGCTTGGCACAACGAGTAACGATCTTTTCGAGATTGTATACAGGTCTGAAGAAGAAATG
ACGACATTCATGCATTTCATGGATTCTACTTGGAATCTATGCGGGAAAGACATAGTACAGGCGTTTGATTTGAGCGA
GTTCCACACGGTCTACGATTTGGGCGGCTGCTCTGGGAGCTTAGCGAAGCAGCTAGTATCCACGTACAAGGAATCAA
CCGTAACAATTATGGACCTACCGAAAGTGGTTCAAGCGGCTAAAAAGCACTTTGTCACTGACAAGGAACAGCAAATC
CACTTTTTAGAAGGTGATTTCTTCAATGATCCCTTACCGGAGGCCGACCTATTTATCGTCGCAAGGATAATACACGA
TTGGACGGAGGAAACTTGCATTAAGCTGTTAAAGAAAATGTACCACAGCTGCCGTCCAGGCGGAGGCGTAGTGATTG
TTGAGCTGTTACTAAACGAGGATAAGTCTGGGCCTGTTATTTCTCAAGTTTACTCCCTATATATGTTGGTACAGGCA
GAGGGAAAAGAGAGGACACCCAGCGAATACACTAAACTACTAACGGATAGCGGGTTCAAGGACATAAAAGTCAAAGC
GACTGAAAAACTTTTCGGTGCGATCCTGGGACGTAAG

Seq. ID No: 111
>IOMT_13

ATGTGCAGCCAAGAGGGAGAAGGGTATAGCTTACTGAAGGAATATGCAAATGGTTTTATGGTAAGTCAAGTGCTTTT
CGCCGCCTGCGAACTTGGCGTGTTTGAACTATTAGCTGAAGCCCTTGAACCGCTGGATAGTGCAGCTGTTAGCTCGC
ATTTGGGATCTAGTCCCCAGGGCACCGAGCTTCTGCTGAATACGTGTGTAAGTCTGAAATTACTGCAAGCCGATGTG
AGAGGGGGTAAAGCAGTATATGCGAACACAGAATTGGCCAGCACATACCTTGTGAGAGGATCTCCCAGGAGTCAGAG
GGATATGTTGTTATACGCGGGGCGTACAGCTTACGTTTGTTGGCGTCACCTTGCAGAGGCCGTGAGAGAGGGCAGAA
ATCAATACCTTAAAGCATTTGGCATCCCCTCAGAGGAACTTTTCTCAGCGATTTATAGATCCGAAGATGAGAGATTG
CAGTTTATGCAGGGGCTGCAAGATGTCTGGCGTTTGGAAGGCGCTACCGTACTAGCAGCGTTCGACCTGTCACCATT
CCCCCTAATCTGCGATTTGGGCGGAGGATCAGGCGCACTTGCAAAGGCTTGTGTCTCCCTTTATCCCGGTTGCCGTG
CCATAGTCTTCGATATTCCGGGAGTTGTCCAGATTGCCAAACGTCATTTCTCTGCCAGCGAAGACGAGAGAATATCC
TTCCACGAAGGCGACTTCTTCAAGGACGCCTTACCGGAAGCTGACCTGTATATTCTGGCCAGGGTTCTTCATGACTG
GACAGACGCGAAGTGCAGCCATCTGCTACAGAGAGTATACAGAGCATGCAGGACTGGTGGGGGTATATTAGTTATAG
AGTCCCTGCTGGATACTGACGGGAGAGGACCATTGACCACTCTACTTTATAGTCTTAACATGTTAGTACAAACCGAG
GGCAGAGAGAGGACGCCCGCCGAATACAGGGCTCTGTTGGGCCCTGCCGGATTCAGGGACGTGCGTTGCAGAAGGAC
GGGCGGCACCTATGACGCAGTTCTGGCCAGGAAA

Seq. ID No: 112
>IOMT_14
ATGGGGTATGCGGCACCCCAAGCGCGTCAGAGTGACAAACAGATCTTCGATATTTACTTCGGCTTTCTTCACAGCTA
CGCGCTTCTATTTGCTGATGAGGTCGGCCTATTCGACCTGCTAAGATGTGAAGCATTGACGTTGGATCAGGTCAGTA
TGGCAACATCCCTTCCATCCCGTAGTAGTCAAGCATTACTTTCATTATGCGCGTCTTTGGGACTATTAGAGAAAAGA
GGAGAGAGATTCGCCCTGTCTGCCCTTACGGAAGGGTTCTTAGTTCGTGAAGCGGAGACTTCATTCTGCGGAGTTTT
AGCGTCAGCACGTGGTCAGGCTGCCGCCTTCAGTTACGATTTCTTCAAGGCGTCATTGCTGAAAGGGGAATCACAAT
TATTCGGTGGACGTGACCTTTTCGATAATAACGCACAGGACCCTGAGCATTGCGAGATATTTACCAGGGCGATGCAC
TCCAAATCTAAAGGGCCAGCGCAAGCGTGGGTGGAAAAGATAGACTTGAGCGCACATGCATGCCTTCTTGATGTGGG
TGGAGGGTCTGGCGTCCATGCTATTTCCGCACTAGCGAGATGGCCTAATCTGAACGCTGTGGTGTTCGACTTGCCGC
CGGTGTGCGCCATCGCCGATACCTTCATCGAAAGGTATCAGATGACAGCAAGAGCACAGACTCATGGAGGGGATATC
TGGTACACGGATTATCCTTTCGCGGATGCACATTTCTACAGCGATATTTTTCACGATTGGCCCCTGGAAAGGTGTCG
TTTTCTAGCGAGGAAATCTTTTGATGCCTTGCCCAGTGGAGGCAGGATTATTTTACATGAGATGCTATTTAACGCAC
AAAAGACAGGTCCAAGGAATGTAGCCGCTTACAATGCGAACATGCTTCTTTGGACTCAGGGACAGCAGCTTTCAGAG
CCTGAGGCGGCAGACTTGTTGCAAGCGGCCGGCTTTGTCGAGATATTGGCCTTTCCGACAGGATATGGGGATTGGTC
ATTGGTTACGGGGGTGAAGCCC

Seq. ID No: 113
>IOMT_15
ATGGGCTCAATAGATGCCCAAATGGCAGCTGTAGAAGAGGAGTCTTGCATATACGCTATGCAATTAGCCTACACTGT
CGTCTTGCCCATGACGCTGAAAAATGCCATTGAATTGGGAATGTTAGAGATTTTAATGGGGGCCGGGGGCAAAATGC
TGTCAGCCTCTGAGGTGGCCGCACAGTTACCGTCAACAACTACTAATCCAGATGCGCCAGCAATGGTAGATAGGATG
TTACACTTACTGGCATCCTATAAGGTTGTCAGTTGCGAGGTTGAGGAGGGGACTCATTCCCGTAGATACGGCCCTGC
CCCCGTTTGCAAATGGTTTACCAGTAACAAGGATGGCGACGGAGCGTCCTTGGCGGCAATGTTACTTTTAACCAACG
AGAAGGTGTTATTGGAAAGTCTAAACCACTTGAAAGACGCCGTCTTGGATGGCGGACACCCATTTCTGAAGGCTCAT
GGGATGACGGTCTACGAGTACAACAAAACCGACGCAAGGATGAAAAGAGTTTTCAGCCAAGCCATGAATAATTATTC
TACGATCATAAATAGAAAATTGGTTGAAATGTACATGGGGTTTCATGACATCGCCTTTCTTGTTGACGTCGGGGGAG
GAGTAGGGACAACGATTAGGGCTATCACTAGCAAGTACCCCCACATAAAAGGTATTAACTTCGACCTTCCGCACGTC
ATTGCTGATGCGCCGCAGTGTCCGGGGGTACAACACGTCGCTGGTGATATGTTTAGGAATGTGCCAAGTGGAGACGC
GATTATTCTAAAGTGGATGCTACATAATTGGACGGACGAACATTGTACTACCCTGTTACGTAACTGCTACGACGCCC
TTCCACCCCACGGTAAGGTTTTTATTGTCGAAAATATTCTGCCGTTGAAACCAGACGCAACGAGCAGAGGTCAGCAA
ACGAGCCTGTCTGATATGATCATGTTGATGCACACGCCGGCCGGCAGGGAAAGAAGCCAGAGGGAGTTCCAGGAGCT
GGGCAAAGCCGCGGGGTTCACGGGCTTCAAAACAACTTACATATATGGAAACAGTTGGGTTATTGAACTTACAACG

Seq. ID No: 114
>IOMT_16

ATGAGTTTCGACACCCAGCATGCATTGCAACCATATTGGGACTTGGCAGTGGCTCCGGTTCAAGCGGACGGTCTTGC
GGCGGCACTAGAGCTAGGTATCTTTGAGGTACTTGCGACTCCCCACACACCAGCACAGCTGGCAGATGTTCTTAGTT
TACATGGGCCACACACGGCGCTGCTACTTGAGCTGTTATGGTCTATGCAAGTCCTAGAGCGTGATGGAGCAGACGCC
GACACCGACGCCAACGCTCTGAGATATCGTTGTACAGCTACGACGTTGCAATACTTCTGCCGTGACGCCGTAGCTTT
CTGCGGTGACGCTTGGCTATACCGTTTACATGCTCTGCGTCACTTCGCAACTCAATTAAACACTCTTGTACGTGACG
GAGGCAAGGTGACCCCGTATTCAACGGCCAGTGGAGTTAACTGGGCGGCTGCCGCCCAACAACAGATCGGCCAAGAG
CAAAGAGCCGTAACGATGCGTGCTGCGCTTTGCGTTATGCAACGTGTAGCTCCATTTGCGGATGGCAACACGCCGTT
GAGACTGTTAGATGCGGGAGGAGGGCCTGGGTGGGTTGCCATAGCCCTTGCGCAGGCACATGCGGGTGTTCACGGGT
GCGTCTTTGACTGGCCCGAGACGGTGGCAGTGGCAGCAGCGAATATAGCGCATGCTCAGCTGTCCGACAGGCTAGAG
ACTCTAGGCGGAGACCTTGATTCCGACGACATCGGTGGCGGATATGATCTTATTTGGTGTTCCAGTGTTCTTCACTT
TGTTCCAGACATGGCCGCTGCATTAAGGAAAATGCAGGCAGCACTGAAGCCGGGAGGCGTGCTAGTATGCATACAGG
CGGAGATTGCTGCAGCACCAGGAGACGCAGCTAGGGTGTTACCGTACTATCTGCCGATGAGAATGTTAGGAAGGACG
GTAACGAGACACGGCGAGTTGGCACAGTTATTGAGAGATACGGGGTGGCGTCAGGTAGAGCAATATGGGGCGTCAGA
CTTCCCAATGGCTCCTGTACAGGTCCTGATTGCTCGTGCG

Seq. ID No: 115
>IOMT_17
ATGCAATTGGCTTCTGCTTCCGTATTACCGATGGTATTGAAAAGCGCGATAGAACTGGACCTGTTAGATATTATTGC
AAAAGCAGGCCCTGGGGCATATCTATCCCCAAGCGAAGTCGCCTCTCAATTGCCCACTAGCAACCCGGACGCGCCAG
TGATGCTGGATAGGATTTTGAGACTACTTGCTTCATACTCTGTTCTAACATACAGCCTGAGAACGTTACCTGATGGT
AGGGTGGAAAGGCTATACGGGGTTGGTCCTGTCTGTAAGTTCCTGACTAAGAACGAGGATGGTGTCTCTATAGCCGC
TCTTTGCCTGATGAACCAAGACAAAGTATTGATGGAGTCTTGGTATTACCTGAAGGATGCTGTGCTGGAGGGTGGTA
TTCCCTTCAACAAAGCGCACGGTATGACTAGCTTTGAATATCACGGCAAGGATCTGAGATTCAACAAAGTCTTTAAT
AAAGGTATGTCTGACCACAGCACGATAACCATGAAAAAGATACTTGAGACCTATAAAGGCTTCGAAGATCTAACATC
ATTGGTGGATGTAGGTGGCGGAACCGGAGCAGTATTATCTACGATTGTCAGCAAATACCCATCTATACGTGGGATTA
ATTTTGACCTTCCACACGTTATCGAAGACGCTCCCTCTTACCCTGGTGTAGACCACGTCGGAGGTGATATGTTTGTC
TCTGTACCAAAAGGCGATGCAATATTTATGAAGTGGATTTGCCACGACTGGTCAGATGAGCACTGCTTAAAATTTCT
TAAAAACTGCTACGAAGCATTGCCCGACAATGGCAAAGTTATCGTCGCCGAATGTATACTACCCGTCGCGCCGGACA
CGTCTCTTGCCGCAAAAGGTGTGATCCATATCGACGTCATTATGCTGGCACACAATCCAGGCGGGAAAGAAAGAACA
GAAAAGGAGTTTGAAGCCCTGGCAAAAGGAGCAGGTTTTCAGGGATTTAGGGTCATGTGCTGTGCGTTTAACACTTA
CATAATGGAATTTATCAAAAAGTTG

Seq. ID No: 116
>IOMT_18
ATGCTGAATCACACAACAATGGTCATTAAGAAGATTCTTGAATGTTACAAGGGGTTTGAAACCCTAAAGCAGTTGGT
GGACGTCGGTGGGGGACTTGGCGTGGCACTAAATCTAATAACCTCCAAGTACCCACACATCAAGGGAATAAACTTTG
ATCTTCCGCATGTGGTGCAGCACGCCCCCTCCTATCCAGGTGTGGAGCACGTGGGCGGAGATATGTTCAAGAGCGTA
CCCAAGGCAGATGCAATCTTTATGAAATGGATCTTGCATGACTGGAGTGACGAGCACTGCGTGAAACTGTTGAAAAA
CTGCTATGCTGCGATCCCAAACGATGGAAACGTAATAGTCGTGGACGCAGTTTTGCCGAAAATGCCAGAAGTAAGTA
CGAGCATGCGTTGCACGTCTCAACTTGACGTCCTAATGCTGACCCAAAATCCTGGAGGTAAAGAAAGAACGGAAGAG
GAGTTTATGGCGCTGGCTACCAAAGCGGGTTTTAAAGGCATTAGGTATCAGGAGTGTTTCGTAAATACTTTCTGGCT
TATGGAATTTTTCAAG

Seq. ID No: 117
>IOMT_19
ATGGAGAGAAAGGAGGAGGTCGCGTTACTGAAAGGGCAGGCGGAAATCTGGCAGCATCTTTTCGCGTTCGCAGACAG
TATGGCGCTGAAATGCGCTGTAGAGCTGAGACTAGCAGACATAATCCATTCTCACGGTGTACCTATAACACTTAGCC
AAATAGCCTCCGCCATCGACTCACCTTCACCCGACATCGCGTATTTGTCACGTATTATGCGTTCCCTTGTGTATAAG
AAAATTTTCACTGAGCATCATCCCAGTGACGGTGGTGAAACAGTGCTGTACGGTCCAACCCATACGTCTAGGTGGCT
ATTACACGACGCCGAGCTGACTCTGGCCCCCTTCGTTCTAATGGAGAATAATCAATGGCAACTTGCCCCTTGGCACT
TCCTAAGCCAGTGTGTGAAAGAGGGAGGAATAGCCTTTAAAAAGGCGCATGGGTTTGAAATGTGGGATTTCGCCGCG

CGTAATCCGGAGTTCAATAAGATCTTTAATGATGCAATGGCCTGTACTACGAAGATTCTTATGGGAGTTTTACTAGC
GGAATACAAAGACGGCTTTGGGTCAATTGGCTCTCTTGTAGATGTGGGAGGTGGGACCGGGGAAATGATCGCTGAGA
TTATAAAACAGCATCCTCACATTAAGGGTATGAAT

Seq. ID No: 118
>IOMT_20
ATGGGCAGTGCGAGTGGTAGCGCAGAGCGTACACAGATGGGCGAGGACGAGGCTTGTAGTTTTGCTATGACAATAAC
ATCTGGCAGTGTCCCTCCAATGGTCTTGAAAGCAGTTATAGAACTGGACGTGCTAGAGATTATTAAAAGAGCTGGAC
CAGGGGCACACCTATCTCCTGCTGAAATTGCAGCCCAACTACCCACCACTAACCCCGGCGCAGCCGCTATGTTAGAC
AGAATGCTAAGACTTCTTGCCAGCTATGACGTCCTTAGTTATTCTCTGCATACTCTTCCTGACGGTCGTGTCGAACG
TCTATACGGTCTAGCGCCCGTCTGTCAGTTTCTGACTAATAACGAGGACGGGGTGACTCTATCTGCACTATCCTTAA
TGAATCAGGATAAGGTTCTGATGGAGTCTTGGTACCATCTGAAGGATGCTGTCCTTGATGGGGGCATCCCCTTTAAT
AAGGCTTACGGCATGACTGCCTTTGAGTATCACGGAACGGACCCGAGATTTAACAAGGTGTTCAACAATGGAATGAG
CAACCATTCCACAATTACAATGAAGAAACTACTAGAAAACTATAAAGGTTTTGAAGGCGTGTCAACATTAGTGGATG
TCGGCGGAGGTACTGGGGCAACCTTAAATATGATAATTTCTAAACATCCAACTATTAAGGGTATCAACTTCGACCTT
CCGCACGTTATAGAAGACGCCCCAACGTACCCCGGCGTTGAGCATATCGGCGGCGACATGTTCGTCTCAGTACCAAA
AGGTGATGCTATCTTCATGAAATGGATTTGCCATGACTGGTCCGATGAACATTGTCTTAGGTTTTTAAAGAATTGTT
ATGCGGCCCTTGCAGATCACGGAAAAGTAATTGTTTGTGAGTACATTCTACCCGTCGCACCGGAGACTAATCACGCT
GCGCGTACCGTTTTCCACGTAGACGCGATCATGCTTGCCCATAACCCTGGTGGGAAGGAGAGAACCGAGCAAGAGTT
CGAATCTTTAGCTAAGGGTGCCGGATTTGAGGGCTTTCGTGTAGCCTTCTTTTTC

Seq. ID No: 119
>IOMT_21
ATGGCACTTAACCCACCACACCAAAACAACGTCATGGAGAAAGAAGATCTATGCTCCTTTGCTTTGAGCATTGCGAC
CTCCAGTAGTCTTAGTATGGTTTTAAAAGCAATTATTGAATTAGATATCATAGGTATAATCAATAGAGCTGGTCCCG
GTGCTCACTTGAGCCCCGCACAGATAGCAGCCCAATTGCCGACTAAAGATCCGGGCGCGACAGCTTCTATGTTGGAC
CGTATGTTGAGGGTACTAGCCAATAACTCTATTTTATCATGCAGTTTACGTGCGCTACCCAATGATGGACCCATCGA
AAGGTTGTATGGTTTGGCCCCGGTCTGCCAGTTCTTTACGAAACCGGAGGATTTCGGTCCAATGGTATTATTTTCAC
AAGACAAGGTTTACACAGACACCTGGCATCACCTGAAGGATGCCGTCTTGGATGGTGGAAGTGCGTTTAAGAAGGCC
CACGGAACGACACTGTTCGAATACCTTGGGACAGATATGAGATTTTCCAAGGTTTTTAATGACGCGATGAGTAGTTC
CTCAACCATAACAATGAAGAAGATGCTGGAAAACTATAATGGCTTCGACGGACTGTCTACGTTGGTGGATGTTGGAG
GTGGAACCGGCGAGACTTTGAACATGATAATAGCTAAATATCCTACCATTAGGGGTATCAACTTCGACCTACCACAT
GTAATCAACGATGCCCCCAATTATGATGGCGTTGAACATGTTGTAGGTGACATGTTTGTATCTGTACCGAAAGGTGA
TGCTATATTCATGAAGTGGATATGTCATGATTGGTCTGACAAACTGTGTTTGAAGCTACTTAAAAACTGTTATACTG
CATTACCCAATCATGGGAAAGTTATTGTCTGCGAATGCATCTTGCCCGTGGCCCCTGAAACGTCACATAGTGCTAGG
GTTGCGTCCAATTTGGATATGCATATGCTTGCGTACTGTAGGGGCGGGAAGGAGCGTACCGAACAGGAGTTCGAAGC
ACTTGCTAAAGGGGCTGGTTTTGAGAGCTTTCGTGTGGTCTGCAGCGCCTATGATCTTAAATTATACATGTGC

Seq. ID No: 120
>IOMT_22
ATGGCAGAGATACCGACCAGTAGCAATCCTAGTGATGACCCTGAAACGCAGAAATTGAATGGCAATGAAGAGGACTA
TGATCATCATCACGATGAAGATCCCGAATCAGACGACGAAAATTACGAATACGCATTACAAATCGCAGAGATGTTAC
CATTTCCAATGGTAATGCACACAGCCATAGAGCTAGACCTGCTTGGTATTATTGCGACTGCGGGGCCAGACAGGCAA
TTGAGTGCAGCCGAGATAGCTGCAGCATTGCCCGCAGCCGGAAACCCAGACGCACCAGCCATGCTAGACAGGATGTT
ATACTTGTTAGCTACTTATAGTGTTGTAACTTGCACGGCGGTTGATGGAGGAGCCTCCGGGGGAGTCGTCCGTAAAT
ACGGACTGGCACCCGTTGCTAAATATTTCGTGTCTAATAAGGATGGAGTCAGCCTAGGTGCGGTAATCTCTCTTAAT
CAAGACCAGGCGGTACTTGCATCTTGGAGTAAACTTAAAGAAGCGGTGCTAGAGGGAGGTATTCCCTTCAATAAAGT
CCACGGGATGGACGCCTTCGAGTACCAAGGTACTAACCCGAGATTCAATGAAATCTTCAACAAGGCCATGTACGATC
AATCCACATACATAATCAAAAAGATTGTGCGTCGTTACAAGGGGTTTGAAAACATCCAGCGTCTGGTGGACGTTGGT
GGTGGCCTAGGCCATACCTTACGTGTAATTACAAGCAACTATCCGTCTATCAAAGGGATTAACTTCGACCTACCTCA
CGTCATTCAACATGCACCGACCATTCCGGGGGTTGAACACGTTGGCGGGGATATGTTCGAGTCAATCCCGCACGGTG

ACGCCATTTTTATGAAGTGTATTCTTCATGATTGGTCAGACGAACATTGCTTAAAAACGTTAAAGAACTGTTACAAA
GCTCTTCCGAGGAAGGGAAAGGTCATAGTTGTACAGATGAACATGATTGAGGAGCCCCAGACTACCCCTCTAGCCAA
AGCGATCAGTCAAATGGACCTATGGATGATGACACAAAACCCCGGCGGGAAGGAAAGGACCAGGAGAGAGTTTCAGG
CTCTTGCAGAAGCCGCAGGATTCGCTGAGTTCAACCCGGTCTGTCATGTGGCTGGGTTCTGGGTTATGGAGTTTCTG
AAA

Seq. ID No: 121
>IOMT_23
ATGTCACCCATAGACCTGGCGAATGAACTGCAGACGCTAGTGACAAGTACCTACTCCGGTGACGTAACCGACCCCTT
CAAACTTTATAAAGCTAAACATAGCATTAGCGACCTGTGTCTAAGTCTATTGAGGGCAGTTCAGGGGCCGGAAGAAT
ATACTGCCATATTAGCCGAGAGCTGCCAGGAATCCAGTGCCCTTAACGTAGTGGCCTCTCTGGGAGTTGCCGACCAT
ATTGCCGAGAGCCCTAACGGAGAACTGACTCTACAGGAGTTAAGCGAAAAAGTGAAAGCGGATGAGAAGTATTTAAG
CGTTGTTTTAAGCTCATTGGTGTACCACGGTTATTTCAAAGAAGTTGGAGGCTTCGGATCTCAAGTCTATGCGAACA
ATGATTTCAGCTCTTTGCTACTATCCGAGGAAACGAATGCGAAGGGCGGCAAGAGTATGAAGGACGCGATCGGTTTG
AGCGCGGATGACGGGGCGAAAGCTACTACGAGATTGTTAGATGCGGCAACCGGCAAGGCAAAGGGGGAGGCAAAGAC
CGCCGCGAATATAGCATTCGACTTTTCTGAATCACTTTTCCAATGGATGGCCAGTCCTGGCAATGAATGGAGGGGTA
AACGTACGGCCAAGGCTATGGTTCAATTACATGGGATGGCCAATGGAGGAATCGGAGAAGACTACCCTTGGGAGAAA
CTGGCGACGCCTATTATCGACATTGGTGGTGGGATAGGATCATTTCAAGGCATGTTGTTGGCTTTACCAAAGAACAA
AGAGTTAACTTTTACAATCTTTGACATTGAGAAAACGGTTGAGCATGCCAAGAAAGTCTGGGCCGGTAAGCCTCAAT
GGATGCAAGATAAGGTAAGTTTTATCGCTGGCGATTTCATGAAATCCTCCCCAAATGATAGTAAAATACCAACACCG
GCTCAGGGTGCCGGCACCTACGTGATTAGACACGTACTACATGACTGGGATGATGCCCAAGTAGTCACAATATTAAA
ACACGTCAGGAACGCAATGCTTGGGAGTCCGGCAAGCACACCGCCTAAGTTACTGCTTGTGGAGATGATGTTAAACG
AAACATCTTCTAGGTTCACCAGGACCACGTCCTTACAACTGTTAAGCCTAAACGGGGGTATAACGAGGACTGAGGTT
CAATTCAGGAGGTTGATCAAAGAGGCTGGGTTCACGGTTGATAGTGTAACGGAAGTGAGGGGTGTTGACCTGGTGGT
GGAGTTATCCCCTGCGAGCTTA

Seq. ID No: 122
>IOMT_24
ATGCCTTCTACCACAATTTCCCAATTGGTGGGTCTAATACAACAGTCCGTCATGGCGTTAGAGAAGCTATGCTTGGA
GAATAGGACGAGTTTGCCTGACCTAGACGCATTCCACTTCGATCAGTCCTCCGAAACCTTCAGGAGCCTGCCCGGTG
CCGCACAGGACGCTAAGATAGCAGTAGCCGCGTGCATGCAGCTAATCGCGATTCTAAGTCCCCCAACCGATACAGTT
TATAGAGCAGCTCTTGGGGGTCACCTTTCTTTTGCGACCAGAACATGTCTGGAGGCTAACATTACAGAGATTCTTAG
GGAAGCGGGACCCGAGGGATTACATATAAATGATATAGCTTCTAAATGCGGGCTAGACCCGTCCAAATTGGGAAGAG
TGATTAGGTATTTGGTTATTCATCATATATATAGAGAAGTCAAGCCCGACGTGTTTACTAATAACAGAACATCTTCA
ACAATGGATACTGGTAAGCCACTAGACAAACTTATCTCAGAACCGGATAGAAAGTATGACGACACTGGGTTTCCGGC
ATTGATTAGTCATTTTATGGACGTGGACCAGAAGTGCGGAGCCGTGGGCTGGGACGTATTAAAGGACCCAGTTCTAG
GCCATTCATGTGATCTTACAGAAACGATATTCAGTAGGGCTTTCAACACAAAGTCAAAGTATTGGGACTTCTTTGAC
CATCCCGAAAACCATTATATGCGTAGGCGTTTCGACTATGCTATGAAGGGACTGGGAGCGATCGAAGATCACGATAT
GGTGCTACATGCTTTCTCATGGGAGGATCTTGACAAAGGATCTGTTATAGTAGATGTTGGCGGAGGTATTGGAACGG
CCATGCTACCCCTGGCCAGGAAATATCCTAATTTCGATATCGTCATCCAAGACCTTCCGATAGTAATCGAGGAAGGC
ACTAAATTCTGGAGTCAAAATTTACCAGACGCCGTTGCCAATGGAAACATTAAATTGCATGCACATAACTTCTTTGA
CGAGCAACCTATTAAGAACGCATCTGTATTTTACTTACGACACGTATTGCACGACTGGCCAATGCCCGACATGGTCA
AAATATTGAGGCGTCTTCGTGACGTTGCCGCCGCGAACACGACGCTGATTATATTAGACTACATCTTACCCTATTCC
TGTAAGATGTTTGCTGACAAGGACGCCGTTTCAATCGCCTCCGCTCGTTATTATAGCGAAGCGCCGGAACCACTGTT
ACCAAACTATACCCATAAGAATGTCATAAGCGATAGTGATATGTATGTTTTTCAGATGATGTTCCACTACAATTCAC
AAGAGCATACCTATCTAAGCCTTAAATCCTTACTAGACGCTTCCGGCTGGCGTTTAGTCCGTCTGCGTGCGATCGAT
CCCAGGAATGACTATTTTCAATCAATTGAGTGCAAGATTCTAGCC

Seq. ID No: 123
>IOMT_25

ATGGCCCAGCCCATGATGCTGGCCCTGGCAAAGTTAATATCCGACAGTGTTGCAAAAGTAGACCAGCTATGCATTGA
ACAGGGTGTCATTTTCCCAAGTCTGGACGACCCTTTTACGACGGAAAGTGAGTCCATTAAGTTACACCCAGATGTTG
CAGAGGCCTCAAACTATATTATATCCGCGGCGGCTCAGCTAATCGCAATACTGAGACCTGTGCCCGTCACCTTATCT
ACGAGTGCTATCCATGTGCATGTTTCCTCAGCTCTACGTGTGGTTGTGGATTCTAATGTCGTCGAGATTCTACGTGA
GGCTGGGCCTCAGGGACTGCATGTGAAGAAGATTAGTGAGAAAAATGGCGTGGAAGCAGGTAAATTAGGAAGATTGT
TAAGGCTTCTTGCGTCCGGCCATATGTTCAAAGAGATCACTCCTGATGTGTTCGCGACAAACAGAATTTCAAGTGCT
CTGGATACCGGCAAGCCTTATGAGGAGCTAGTCAAAAATCCGGGCGAGAAATTAATCGGGACAAATGGGATCGCTGC
CTATATATCAAGATCAACAGATGAGTCCGTCAAGAGCAGTGGGTTTCTGTATGAGGCTCTGACATATAGCTCAAGCG
AGAAAGTACCCCTTCCTCCGTCACCTTTTAACCTAGCGTTTAACACGGAATTGCATATCTTTTCCTGGCTTGCACAA
AAAGGCAATGAACATCGTTTGCAAAGGTTCGGAATTGCGTTTGACGGCTTTGACAAGATGTTGCCCGTTAACGGTGT
GACCAAAGGCTATAGGTGGGGCTCCTTGCCAAAGGGTTCTATTGTCGTGGACGTGGGCGGTGGTGTTGGAAGCGAAT
CAATGAAGATCGCCAAGACATTTCCAGATCTGAAGGTTATAATACAGGATGCTGAAGGCGTCGTAGCAAATGGAGTA
AAATTTTATGAGACCCGTTTCCCAGAAGGGCTATCCTCCGGCCAGGTTACGTTTCAAGCACACGATTTCTTCACGCC
GAATCCTGTAACCAACGCGAGGGTTTTCTTCATGAGGTTTGTGCTGCACGATTGGCCTGACGCCACCTGTGTCAAGA
TACTTAAAAACCTGAGGGCGGCTGCCGCGCCTGATACAGAACTTATCATAAACGAGTGTCTAATCCAGTACGCGTGC
AGTACCGAGTCAGAAATTTCCAAGTCAATTCCCGGTGGTAGGTTCAAACCCCCACCTTCCCCGTTGCTGCCAAATTT
GGGCTATGCACGTATTTTTCATTATCTTATTGATTTACAGATGGCGATAGTTGCGCATGGAGTCGAGAGGACTGTTG
AACAATATGCGAGTATCCTTCAGAAAAGTGGATGGAAGCTGAAAGAAGTTCTGAGGATGCCTGAGTCAGCCTATAGC
TTACACAAGCTGGTAGCCGTCCCCCAGCCTGAG

Seq. ID No: 124
>IOMT_26
ATGACGAGGCTAACCGACTCTTTAGGGATGCTAAGAAGCAAACTTGTACCTCCACAGGCTACAATGCTACAGCTATT
AACGGGATATCGTGTTTCCCAAGGGATCTATGTTGTTGCAAAGCTGGGCATCGCCGATCTTTTAGCAACAGGTTCCA
AGACCAGTCAGGACCTAGCGGCGATAACTAACGTTCACGCTCCAAGCCTGTATCGTCTTATGCGTTCATTGGCCAGC
CTGGGAATCTTTACAGAAACTGAGAATGGGAGGTTTGAGCTAACTCCTCTGGCCGCTACCTTGAGGTCTGACCATCC
AAACTCCGTTCACGATGCCGCTATCATGTTTCTGGAAGACTGGCATTGGCAGGCTTGGGGCAATTTCTTTGACTGTG
TAAAGACTGGGGAAACAGCTCTTGAAAAGACATTTGGAACCAGCAATGTGTTTGACTATTTTGAGACACAAAACCCC
GAAGCTGGGCAGCATTTTGATAACGCCATGACGAATACTAGCGTGATGACAAATCAGGCGCTACCTACCGCTTACAA
TTTCGGAGCCTTCAAAACCTTGGTGGATGTTGGCGGCGGTCAGGGGTCCTTCTTATCGGCGCTGTTTCACCAGTGGG
ATCACCTGCATGGCATACTTTTTGATCTACCACCAGTGATTGAAAGCGCAGAGCAGCAAAACTTATTGTCTGGGTTT
GAAAAGCGTACTACACTTGCTGCAGGTGACTTTTTCAAGGCCGTTCCCGACGGTGCGGATGCTTATCTGCTGAAGAC
CATCATACACGACTGGGACGATGCCAGCGCGATAGCAATTTTAAAGACATGCAGGCGTGCCATGAATCATGATAGTA
AGTTACTATTAGTCGAGCTGATAGTGCCGTCCGGGAATGCGCCCTCCCTTAGTAAAATTTTAGATTTAGAGATGTTG
GCAGTTTTTGGTGGGGTCGAAAGAACGGAAGCCGAGTATCGTTCCCTGTTGCTATCAGCAGGCCTTAAATTGACTAG
GATCTACGATAGCCCGTGCCCCTGGAGTGTGATTGAAGCTATTCCCGTT

Seq. ID No: 125
>IOMT_27
ATGTCCATGCCACCGGCCCACAGTCGTTTATACAGCAGGTCCTTCTTATCTATGCTTCCCGACGCAATCACGCCATT
TCCGTATTTGCCACCTGATGCGACGGATACCAGACCTCTGTTAGCCGAACTGGAAGCCTTGTTAGAAATAATAAATA
GCTCTGCTAGATTGGCCATAACTGAGTACAAAAAGCACGGGAATAACGTACCAACGATTTATAGTACAGAATTTCAT
CCTTTAGATTTCGCCACGGACACCGTCGCGCTGAAGAAAGCGATCAGGTTGCTTGAAGACGCATGTCAACAACTTTG
TGCATCATTGGCTCCGCCACAGCACACATTAGCGAACGTGTCCCGTGTGCACCATAGACAGTACGTTACACAGCTAA
CCACGCACGACATCTTGGAAAAATATCCTTCAGGCTCTCATATCAGAGAGCTATCACAGACAGTGGGCTTGGAAAAG
GGCAAATTGGCAAGGATTTTAAGGGTTTTTGCTTTCAAGGGATGCTTCATAGAAGTCGATACCGATGTCTTTGCTAG
TAACAGGTTAAGTTTGATAATGAAATCCAGCAATGACTGTGGATGTCTAACTTGTATTCACGCGCAAGACGTATCAC
AGGGCGCTGGTGTCCTTTACGAAACGTTGACGGAACCTGAATACGCAATGAGCTACGAGCCCGACAAAGCTCCAATG
ATTTATGTCCTTAAACGTAAGGGCTTAAAGGGATCGTTCTTTGATTGGATGAAAGCCGATGCCAAAAGGAGGGAAAA
TTACCATTATGCCATGATCGCTTTAGGCCCCGTTATGGGTTCTCTTAGCATCCTGCACCACTACCCCTGGAATGACG
TGGCCACGGTATGTGATGTAGGAGCCAGCGTAGGTAGTGTTTCCATCCCCCTTAGTAAAGCTCATCCTCACCTTAAA

ATCACCGACCAAGATCTGCCTGAAGTGTTAGAAGCCGCCAGATCCGTATGGGAAAAGGAGGCATTCGAAGCGTTGCG
TGAAAAGAGGGTAGAATTTCTGACGCTTGACTTCTTTAAGGAGGCGCCCGTCCCAGGCAAAGATGTATACTACCTAA
GGCACATCATCCATGATTGGCCAGATGCCGAAGCAGCTGTTATCCTTAGGAACATAAGCAAAGCAATGGAACCGCAT
TCAAGATTGCTTATCCACAATTATGTAATCGCTGGAGCCAACCGTAGACCAGATGAAGAACAACGTGCGCCCGAGCC
AATGCTACCTAATTTTGGGGCGGGGGATAGTAGAAAGTACAGGCAAGACCTAAATATGTGGATTCTGCATAACGCTA
AAGAGAGGACGGTCGATGATCAAATCACACTAGCC

Seq. ID No: 126
>IOMT_28
ATGGCCCCTGGCCGTGAGGGGGAACTGGATAGAGATTTTAGAGTGCTAATGTCTCTAGCGCATGGGTTCATGGTATC
ACAGGTCCTATTCGCCGCATTAGACTTAGGCATCTTCGATCTGGCTGCCCAGGGCCCTGTCGCTGCCGAGGCAGTTG
CGCAAACGGGTGGCTGGTCACCAAGGGGGACACAGCTGCTTATGGATGCTTGCACTCGTCTAGGGCTACTTAGGGGT
GCGGGCGATGGTTCTTACACTAACAGTGCGTTAAGTAGTACCTTTCTGGTGAGCGGCAGCCCGCAATCACAAAGATG
CATGCTACTTTATCTTGCAGGCACAACTTACGGGTGTTGGGCGCACTTAGCTGCTGGGGTCAGAGAGGGAAGGAATC
AGTACAGTAGGGCGGTCGGTATTTCAGCTGAGGACCCCTTTTCTGCTATTTATCGTTCTGAGCCAGAGAGACTACTG
TTTATGAGGGGCCTGCAGGAAACGTGGTCCCTGTGCGGAGGTCGTGTCCTAACCGCGTTTGACCTGAGCAGATTTAG
AGTGATATGCGACCTGGGAGGGGGTTCAGGGGCACTTGCTCAGGAGGCCGCTAGGTTATATCCAGGCTCATCAGTAT
GTGTATTTGATTTACCTGATGTAATCGCTGCCGCAAGGACCCATTTCCTAAGCCCAGGGGCCAGGCCGTCCGTGCGT
TTTGTGGCCGGTGACTTTTTCCGTTCCCGTCTGCCTAGAGCAGACTTATTTATATTAGCAAGGGTTCTGCATGACTG
GGCTGATGGGGCGTGCGTCGAGCTATTGGGCCGTCTACACAGAGCGTGCCGTCCCGGTGGGGCGCTTCTGCTAGTGG
AAGCAGTATTAGCTAAGGGAGGCGCCGGTCCGTTGAGGTCACTACTGCTATCTTTAAATATGATGTTGCAAGCGGAA
GGATGGGAGCGTCAGGCCTCCGACTATAGAAATTTGGCGACGAGGGCTGGTTTTCCACGTTTACAACTACGTAGACC
TGGCGGTCCATATCATGCTATGTTGGCACGTCGTGGGCCCAGGCCCGGCATTATTACGGGAGTTGGCTCTAACACCA
CTGGAACTGGGTCATTCGTAACCGGTATTAGGCGTGACGTGCCGGGCGCGAGGTCCGATGCAGCAGGCACTGGCTCT
GGAACTGGGAATACGGGGTCTGGGATTATGCTACAAGGTGAAACTTTAGAATCAGAAGTATCAGCACCGCAGGCAGG
TTCCGACGTTGGCGGCGCTGGTAATGAACCCCGTAGCGGTACACTGAAGCAGGGTGATTGGAAA

Seq. ID No: 127
>IOMT_29
ATGGAGGTTGTACCATCTTGGTTCAAGGAAACTCTTGATAAAAGTCAATTTTCTGCGCCATATGAATATGCAGTTGA
GACAGCAAAACAAAAAGCACTGGAAGTTGCTAGGAGAATGCATGTTAAACACCTAAAGACCCCGGACATCGTAATCG
GAGCAGACACCATCGTCACGCTAGAGGGGGCCATATTAGAAAAGCCGTTTGATAAACAGGATGCTTACAATATGCTT
AGTAGACTAAGTGGGAAGGAGCACAGTGTTTTCACCGGAGTGGTAATCGTCCATTGTCGTTCCAAGGAAGAGAACCA
TTTAGAGACGGATATTATCGACTTTTACGAGGAAACGAAGGTTAAGTTCGCAGACCTGTCCGAGGACTTGTTATGGG
AGTACATCGATAGCGGTGAGCCAATGGATAAGGCGGGCGGCTACGGTATTCAATCTCTTGGCGGAATGTTAGTGGAA
TCAGTGCATGGCGACTTCCTAAACGTAGTCGGCTTCCCCTTAAATCACTTCTGTCGTAAGCTCACTGAAATATACTA
CCCGCCCCCAAAACAAGCCATATGCCGTGTAAAGCACGATTCTATTCCATATGTAGAATCATTTGAGAATCTGAGTG
ATGTCGAGACGGATTGCACCTCTACTAGCAAAGCATGTGAAGCGAAGAAAGCCGTGCAGGACGGCGTATGCAAGGCG
GATGGTTCAGGTTCAGCTGTTCTTCAAAATGGAATCGAAGAGAGACCCGTCCATTGTGCCCAGCAATTAAGTAAAAT
TACACAGCTGCTTGACGGATTTAAAGCCTCCCAGACTCTATTCGCAGCATCTAAGTTGAAAGTTTTTGACAAATTAA
AAGATAAGGGGGCGCTTAAAGCGATGGAGATTGCGGAAAAGATTAATGCGTCTGTACACGGAACGGAAAGACTACTA
GACGCCTGCGTGGCTTTAGGCCTACTAGAGAAAACCCACCAGGTTTATTCCAATACGGAACTTGCTAACACGTATTT
AGTGTCAGATGGAGCGTTCTCCATACATGAATATATTACCTACTCAAGCGATCATTTATGGTCCCACTTCACTCATT
TAGATTCTGCCGTCGTAGAGGGCGGGGGACAACATCAGACGGCCGTTAAAAAGGCGTGCGATAATAGAAATGGGTCT
GAGGTAAAGGAGAGATTTATGAGAGCCATGCATTGCATGCTTAAAATAACGGCCAGAGATTTGGTCACCGCATTTGA
TCTGTCAAAGTATTCAAGCGCCTGTGACTTGGGCGGATGTACGGGCGCCCTAGCACACGAACTGGTCTGGACTTACC
CTGAAATGAAAGTCAATGTATTTGACCTACCTGAAGTAATCAAGCATACCAGTCAATTCCAGCCTGAAAGTTTTGAC
TCAAGTAGAGTCACCTTTAGTTCCGGGAACTTTATGGAAGATACACTTCCAGAAGCAGATCTGTATATTTTGTCAAG
GGTGCTACACGACCTGCCGGAGGGCAAACTTAATCACCTGTTAAAGAAAGTTAGTGAGGCATGTTGTCCTGGTAGAA
GTGCATTGCTGGTTGCCGAGATTGTTTTGGATGAGGACAAGAAAGAAAGTAGGGGCTTACTTCAAAGCCTTAGTATG

GGAGAAGGGAAACAGAGAAGCGGGACTGAATACAAGAAGTTGCTAGAAAACCACGGATTCAATAGCGTCCAGATCAA
AAGTACCGGAAATCTATTAGATGCGATTCTTGCAATTAAGACCAGT

Seq. ID No: 128
>IOMT_30
ATGGACACAGTAAAGAACCTACAAGCGAGTAACGTCCCCTCCTCACTGTCTCAGGAGGATGAGGAAGTGTTCACTTC
AGGCCTGCATGTATGTAGTTCCGAAGTCTTTTCTCACGCGCTGTCCAATTGCATACAACTTGGCCTTTTTGATATAA
TCGCTGAAGCCGGGCCCAGTGCATACCTGACCGCTACTGAAATCACTGCACAATTGCCGACTAAGAACCCCGATGCC
GTATCAATGATAGATAGGATGTTAAGGCTTTTTAGCTGTCACTCACTATTAAATAGCTCTCTAAAGACCGTAGCAGA
TGATGTCGTCGAGACTAGGTATGGTTTGTCCCCGATCGGCCATCTGTTCGTGAGAAAGAAGGACGGCGTAACAATGG
CGGCCTGTTTTACCGACTACAAAGCGTGGACCGAAGCCTGGCTGCATTTAAAAGACGCCATATTAGAGGGTGGGAAC
CCTTATGAAAAAGCGCACGGGGTCCCTATCTACGAACATATCTCTTCCGACACTGAAAGTGTCAAGGGCTTCAGTCA
AGCGATGGACTCCATCTCATCATTCATTATGAAGAAGGTCCTTGAAAATTATTCTGGCTTTAAAGGTCTAGGGTCCC
TTGTGGATGTAGGGGGAGGGTCCGGCTTCGCGCTAAACATGATAACGTCCGAGTACCCCTCAATCTCCTGCATTAAT
TTCGACCTACCGCACGTCGTTCAAGAAGCACCTTATCATCCAGGTGTGAAGCACGTCGGGGGAGACATGTTTCTTGA
CATCCCTTCTGCTGATGCAATAATGATCAAGGAGGTGTTGCACAATTGGGGTAATGAAGATTGTGTCAAGGTCTTGA
AGAATTGTTACGAGGCGTTGCCTAAAGGAGGCAAGGTGATTGTGGTAAGTCATGTTATGCCGGAGGTAGTCGGCAGC
AGCAACGCAGCAGCCAAGTACGTGTGCCAATTAGACGTTATGATGCTACTGTTTGGCGGGGGTAAAGAGCGTACCGA
GAAGGAGTTTAAAGCCCTGGGAAAGGCTGCAGGGTTTAGTGGGTTTCAGCTAATATGTTTCGCAGCTTATAATGCCG
TTGCCGTGATGGAGTTTTACAAG

Seq. ID No: 129
>IOMT_31
ATGGCCGAGGACGTAGCCGCTGTCGCAGACGAAGAGGCATGTATGTACGCCATGCAGCTTGCCTCTAGTAGTATCCT
GCCTATGACCCTTAAAAACGCATTAGAATTGGGATTACTGGAGGTGCTTCAAAAAGACGCCGGGAAGGCATTAGCCG
CTGAAGAAGTGGTTGCTAGACTTCCCGTTGCACCGACCAATCCGGCCGCGGCGGACATGGTCGATCGTATGCTGCGT
CTACTGGCATCTTACGATGTCGTTAAATGCCAAATGGAAGATAAAGACGGTAAGTACGAACGTAGGTATAGCGCGGC
GCCCGTGGGTAAGTGGTTAACGCCTAACGAAGATGGCGTCAGCATGGCCGCATTGGCACTAATGAACCAAGACAAAG
TTTTAATGGAGAGCTGGTATTATCTGAAGGATGCTGTTCTCGACGGAGGGATACCGTTTAACAAAGCCTACGGTATG
ACGGCCTTTGAATATCACGGTACGGACCCAAGATTTAACAGGGTGTTCAACGAAGGTATGAAAAACCACTCTGTCAT
TATTACCAAAAAGTTGTTAGAGTTCTACACGGGCTTTGACGAGAGCGTCTCCACCCTAGTAGACGTCGGAGGTGGCA
TAGGCGCCACATTGCACGCAATCACTTCCCACCACAGCCATATCAGGGGTGTCAACTTTGATCTTCCTCACGTGATT
TCTGAGGCGCCACCGTTTCCTGGAGTCCAGCATGTTGGGGGTGATATGTTTAAATCTGTACCCGCAGGCGATGCGAT
CCTTATGAAATGGATTCTGCATGACTGGAGCGATGCTCATTGCGCTACCCTGCTGAAAAACTGCTACGACGCCTTAC
CAGAGAAAGGCGGGAAGGTAATTGTTGTCGAATGCGTACTTCCCGTAACGACCGATGCAGTCCCGAAGGCCCAGGGA
GTGTTCCACGTTGACATGATCATGTTGGCTCATAATCCAGGGGGCCGTGAAAGATACGAGAGAGAGTTCAGAGACTT
AGCCAAAGCTGCAGGTTTCTCCGGCTTTAAGGCAACGTACATATACGCAAACGCATGGGCAATAGAATTTATAAAA

Seq. ID No: 130
>IOMT_32
ATGACCTCACTGCAGGACCTGGATTATCCACAACAGTTACTTGAATATAAAGATGGATTCCTAGTCAGCAAGACAAT
GTTCACCGCGTGTGAACTGGGGATCTTCGATCTTCTACATAAGTCTGATGAAGCCCTGAGTGCTTTAACAATCTCCA
GCCACCTGGGGACCAGTGCCGACGGAACTGATAGACTTCTGTCCGCCTGTGTTGGTCTAAAGTTGTTAAAGGTCGAA
ATGAAGAACAATGAGGCATTCTTCTCAAATACTGATGTGTCCGATGTTTACCTGGTGCAATTAAGTCCTAGAAGTTT
GTATCATATGATGATGTATTACTCCCAAACGCTGTATAAATGTTGGCATTTCTTACCCGATGCAATAAGAGAGGGAA
AGAGCCAATACGAGAGGGCCTTTGGCGTTTCCTCTGGCGACATCTTTAAGGCTTTATATAGGTCAGAGGAGGAAATG
GTTACCTTTATGCACCATATGGATAGTGTTTGGAACATCTGTGGAAAGGACATCATCGCGGCTTTTGACCTGTCTTC
CTTTAACGAGGTTTGTGACCTTGGAGGATGCAGCGGGGGACTAGCTAAACAACTTTTGAGTATATACCCGTCCTCCA
GCGTGACAATTCTGGATCTACCCGAAGTTGTCCAGACCGCTAAGAAGCATTTTATAACGGATGCAGATTGCAACATT
GCGTTCTTACAAGGTAACTTTTTCAATGACCCAATTCCTGAGGCAGATTTATATATAATGGCCCGTATTATCCACGA
CTGGACCCAAGAAAAGTGCTTACAGTTACTTAATAAGATCTACAAAAGCTGCAGACCCGGAGGCGGAGTACTTTTGG

TGGAAGTACTATTGAACGAAGATAGGAGCGGCCCTCTGACCAGTCAGCTTTACAGTTTGAATATGCTTGTTCAGACT
GAAGGTAGGGAAAGAAGCCCGTGTGAGTATACGAAATTGTTGGCCCACTCCGGGTTCAGGGACATCCAAGTAAAGGC
GACGGGCAAGATTTATGATGCCATTCTAGGAAGGAAA

Seq. ID No: 131
>MUP1
ATGTCAGAAGGCAGAACGTTCTTGAGTCAACTGAATGTGTTTAATAAAGAAAACTACCAGTTTTCAAGCTCAACAAC
GAAGAAAGAGGTATCCAATTCAACGGTGGACGCAGACAACGGAGCCTCAGATTTCGAGGCGGGACAGCAATTTGCGA
CCGAATTGGATCAAGGAGAGAAACAACTTGGCATATTGAGCTGTATCGGTTTGATCTGCAATAGAATGTTGGGTACT
GGGGTATTCGCAGTATCTTCTACCATCTATACATTATGCGGTAGTGTAGGACTGGCTCTTATTATGTGGGCGGTCGG
CGCAATAATCGCAATTAGCGGATTGTATGTCTACATGGAGTTTGGAACAGCGATTCCCAAGAATGGCGGTGAGAAGA
ATTACCTAGAGGCTATTTTCCGTAAACCCAAATTTTTCATCACATGTATGTACGCAGCGTATATCTTCTTTCTGGGT
TGGGCAGCCGGTAACAGTATTAACACTGCTATCATGTTTTTGACTGCTGCGGACACGGAAGTCACCAAATGGAACCA
GAGGGGGATTGGAGTTGCTGTGGTGTTTTTCGCTTTCCTGATCAACAGTCTTAATGTTAAAATTGGACTTTATCTGC
AGAATATACTTGGAATTTTCAAAATCGGTATTGTACTTTTCATCTCCATAACTGGTTGGGTGGCGTTGGGAGGCGGG
CTGAAGGATGGATATCAATCTCATAATTTTAGAAATGCCTTTGAAGGAACGGAGACCGCTACTGCCTATGGGATCGT
CAACGCATTGTATTCTGTTATTTGGTCTTTTGTTGGATACTCTAATGTCAATTACGCTTTAGGAGAGGTAAAGAACC
CTGTAAGGACGTTAAAAATAGCCGGTCCCACGTCAATGGTGTTTTTGGCTATCATCTATATATTTGTAAACATAGCT
TACTTCGCAGTGGTACCCAAGGATAAGCTAATAAGTAGTAAACTAATCCTTGCAGCCGATTTCTTCGATATTGTCTT
TGGCGGTCAAGCTAAGCGTGCTGCAGCAGCGTTGGTTGGTTTGAGTGCGTTGGGCAACGTCTTGAGTGTAATCTTTT
CACAGGGCCGTATTATACAACAGCTAGGGAGGGAGGGCGTACTACCCTTTAGCAACTTCTTCGCAAGTTCAAAACCC
TTCAACTCTCCAATGGTCGGCCTTTTCCAGCACTTCATTGTCTGCACAGTAACTATACTGGCCCCACCACCCGGCGA
CGCTTATCTGCTGGTCCAAAATCTAATTAGTTACCCTATGAACATTATCAACTTCGCTATCTCCGCAGGGTTGCTGT
GGATCTATTGGCAGAGGAGGCAAGGTAAAATAGAGTGGAACCCACCTATCAAAGCTGGGGTTTTTGTAACCGGCTTT
TTCACCCTTTCAAACTTATACTTGATCATAGCGCCCTATGTCCCTCCGAGTAATGGCGAATCTGTGTACAGTAGTAT
GCCCTATTGGATACATTGCGTCATAGCCTGGGGCATCTTCTTTTTCGGGGGTGTATATTATGTGGTATGGGCTCAGT
TGTTGCCCCGTTGGGGGCATTACAAGCTCGTCTCAAAAGATGTTTTGGGTGAGGACGGTTTCTGGAGGGTGAAAATC
GCAAAAGTATATGATGACACCATCGGTGACGTGGATACGCAGGAGGATGGTGTAATCGAAACCAATATAATTGAGCA
CTATAAGAGTGAACAGGAGAAGTCTCTA

Seq. ID No: 132
>NAT_1
ATGGCTCCTATCGAAGAAGAGGAGCCGTTACCTGAGGAGCTAGTCCTGCTTGAGAGGACTCTGGCCGACGGTTCAAC
GGAGCAGATCATATTCAGCAGTGCTGGCGATGTGAACGTCTACGACCTACAGGCTCTTTGCGATAAAGTTGGGTGGC
CGAGGAGACCCTTGACCAAAATAGCAGCTAGTTTGAGGAACTCCTACTTAGTAGCTACGCTGCATAGCGTTACAATG
CCTTCCAAAGCCGAAGGCGAGGAAAGGAAACAGCTAATAGGCATGGCAAGAGCAACTTCCGATCATGCTTTTAACGC
CACTATCTGGGACGTATTGGTTGATCCATCCTATCAGGGCCAAGGTTTAGGTAAAGCTCTGATGGAAAAAGTAATAA
GGACGTTGCTGCAAAGGGATATCAGCAACATTACTTTGTTCGCAGATAACAAGGTCGTTGATTTCTATAAAAACTTA
GGGTTCGAAGCCGACCCCCAAGGTATTAAAGGTATGTTTTGGTATCCGCGTTTT

Seq. ID No: 133
>NAT_2
ATGAGTACACCGAGTGTTCACTGTCTAAAGCCCTCCCCCTTGCACCTTCCGTCAGGCATTCCTGGTTCTCCAGGTCG
TCAGAGAAGACACACTCTGCCCGCGAACGAATTTAGATGCCTTACGCCTGAAGACGCGGCTGGAGTTTTCGAAATCG
AACGTGAGGCTTTCATATCCGTGTCTGGTAACTGTCCCCTGAACCTGGACGAAGTCCAACACTTTCTGACTCTGTGT
CCCGAATTGTCACTTGGATGGTTTGTGGAAGGGCGTTTGGTAGCGTTCATCATAGGATCTCTATGGGATGAAGAGAG
ACTAACTCAAGAGAGTCTTGCACTTCATAGGCCGCGTGGACATTCCGCCCATCTACATGCGTTGGCCGTACATAGGT
CATTCCGTCAGCAGGGGAAGGGCAGTGTTCTTCTGTGGAGGTATCTACATCACGTCGGGGCACAACCAGCCGTGAGG
CGTGCTGTGCTGATGTGCGAAGATGCACTGGTCCCGTTTTACCAGAGATTCGGTTTCCACCCGGCTGGTCCTTGCGC
TATTGTCGTTGGGAGTTTAACGTTTACTGAGATGCACTGTTCTCTGAGGGGGCATGCTGCACTACGTCGTAATAGCG
ACAGG

Seq. ID No: 134
>NAT_3
ATGACAAGTGATGTCGGCGCTGACGAACATGCTACAACAGAGGCGGGCGGCGGGCGTTTACAGGCCGGCGGACATTC
TTCTGCTGAAGAGGCTTCCGAAAGATGCCCGCCTGCTGCTGCGCCGCCGAGTGGTATGAAAGGAGCGGCTGATTGCG
GGCCACAAGATTCATCAGCAAGAGATGACGTGTCTTTTATTCCCTATAAAGATGAAACTGACATGCCAGGAATCGTG
GAGCTTATCGAGAAAGATTTAAGTGAACCTTATAGCATATTTACTTACCGTTATTTCATAAATAACTGGCCAGAGCT
TTGTTTCTTAACTATGCGTGGTGATAGTTGCGTCGGCGCAATCGTGTGCAAATTAGATGTCCATAGGTGTAGGAACA
CGAATAGAGGCTACATAGCGATGTTGGCAGTTGAAAAGGGATTACGTGGTAAGGGAATAGGAAGCACTCTTGTCCGT
CTGTGCCTAGATAAAATGAGGGAAATGGGAGCGGACGAATGTGTACTTGAGACGGAGGTCACTAACAAGGGCGCACT
GGGGTTGTACCGTAATATGGGTTTCGTAAAGGAAAAGAGGTTACACAAATATTACCTGAATGGCAATGATGCCTTTA
GATTGAAATTCTTGTTCAAGCTGCCCGAAGGCTTCGATAGAGGAGAGGGTTGTCTAGGTCCTTTATGTGAGGTACCT
CCAGTGACGACT

Seq. ID No: 135
>NAT_4
ATGGTGTCCATACGTCCTGCCACAGTGGACGACTTACTTGCGATGCAAGCATGTAATCTATGCTGCCTGCCAGAGAA
TTATCAAATGAAATATTACTTCTATCACATGCTGAGTTGGCCCCAGTTGCTTTATGTTGCGGAAGACTACGGAAAGA
AGATTGTAGGCTACGTGTTAGCGAAAATGGAAGAAGACAGTAGCGAGGTCCACGGTCATATCACATCCCTGGCCGTG
CTTAGAAGTCATAGGAAACTTGGTTTAGCGTCTAAGCTGATGCGTGCGGCGATGGCTGCGATGGAAGAAACATTCGG
AGCAGAACATGTCTCCCTGCACGTGAGAGTTACTAACCGTGCCGCCTTCACATTGTATTCAGAAACGCTAGGGTTCG
AAATTAACGACGTTGAACATAAATATTATGCCGATAAGGAAGATGCTTATGATATGCGTAAGATGTTTGAAACAGGG
CTGAAGAAGCAAGAGGCCGGTAAGCAAAAGAAAAAGGAAAAGGAGAAGGAGAAAGAAAAGGAGAAGGAAAAAGAAAA
GGAGAAAGAAAAGGAAAAGGAGAAAGGTAAGGACAGCCAGCCGGTGGAACAGCAGGGTGGTGCTGCCGGTGCTGACA
AGGAAGCGCAACGTTCCAAAAGAGCGCGTTCAGGCGACAGGAAAAGGAACGGCCGTCCCCGTAGGAGAAGTGGATCT
GGG

Seq. ID No: 136
>NAT_5
ATGCTTCCACGTCCGCCGGTGGGCGCGGCCAAGGAGGGGCACCTTACGTTATTTTACCGTGAGTTAAGATGGCTTTG
CCCTGGAACTCGTTTTTATTTCGTAGTAAGGGACCCCGCCGAAAACGTAAGATCAATAGCAGACAGACTGGCCTTAG
GACCAGAAGGCCTTAGGAGACCTCCAAGGATAGTAGCTAGAGCAGACTTAGGATGGCGTGAGGTACTAAACATGTCA
TACGCAGGAGTGAGGGAAGAATCAGCGTTGGGGACTCTGGTTGGCAGGTGGAACTTAATGGCGAGATTGTACCTAGA
CGCGCCGAAGGGCGCGATGGCCCTAGTGCGTTATGAAGATTTGGTTGCGGAAGCAACTTGGGAAGCTGAAGTCAGAC
GTGTAGCGGCAGCCGAAACCTTAGACCTAAGGGAGAGAGTCCTTTGGCCTGGTCGTCCTGACCTTTGTACTTTGCCG
GGGGATGAATCAGCTCTGCATTTCGGGGCAGTAGCTGCTGGCAAAGTTCTAGGCGTCATTTCTGTGTTTCTTTCTCC
GGAACCAGGTGGTAGGGCCCAATTCAGGAAATTTGCTGTAGACCCTGAGGTCCAAGGGAGAGGGCTGGGTCGTAGAC
TACTGGAGCAGGCGGTTGCAGCAGCTAGAGAAGCAGGGGCCGGTAGCTTATTTTGTCACGCAAGAGCAGACCAACAG
GGGTTCTACGAGAGGCGTGGATTGCATGTGGTCGGTGAACCATTCGAGAAATACGGTGGCAAGCCCTACGTCGAGAT
GGAAGTGCCGTTCCAA

Seq. ID No: 137
>NAT_6
ATGAAAGGCTCAAGAATAGAGCTAGGTGACGTCACACCACATAATATCAAACAGTTAAAAAGGCTAAATCAGGTCAT
CTTCCCCGTTAGTTACAACGATAAGTTCTACAAAGATGTGTTGGAAGTAGGGGAGCTGGCGAAACTAGCGTATTTTA
ACGACATAGCAGTAGGCGCTGTGTGTTGCAGAGTCGATCATTCCCAAAATCAGAAGAGGCTATACATTATGACTCTA
GGCTGCCTTGCGCCATACAGGCGTCTGGGTATAGGAACTAAGATGCTAAACCATGTTTTGAACATCTGCGAAAAGGA
TGGTACTTTCGATAACATATATTTGCATGTCCAAATAAGCAATGAGTCCGCTATTGACTTTTATCGTAAGTTTGGTT
TCGAGATTATCGAGACGAAGAAGAATTACTATAAACGTATAGAGCCGGCGGACGCGCATGTCTTACAGAAAAATCTA
AAAGTAAGTTCCCCTGCTCCCAATGCCGACGTCCAAAAATCCGAAAAT

Seq. ID No: 138
>NAT_7
ATGTCTACTCCGTCCATTCATTGCCTGAAACCTTCACCGTTACACCTACCATCTGGCATTCCCGGTAGTCCCGGTAG
ACAAAGGAGACACACGCTACCGGCCAACGAATTTAGATGCTTAACCCCGGAAGATGCGGCTGGAGTGTTTGAGATAG
AAAGAGAGGCTTTTATATCCGTATCAGGGAACTGCCCATTGAACCTAGATGAGGTAAGACACTTCTTGACCCTTTGT
CCCGAGTTGAGTTTAGGGTGGTTTGTAGAAGGTAGGCTAGTGGCGTTCATTATAGGGTCACTGTGGGATGAGGAAAG
ACTTACCCAGGAGTCTCTGACATTGCACCGTCCCGGCGGGCGTACAGCGCACCTACATGCCTTGGCGGTCCATCATA
GCTTCAGGCAACAGGGCAAGGGCTCTGTCCTACTATGGAGATATTTGCAGCATGCTGGTGGCCAGCCAGCGGTTAGA
AGGGCTGTACTGATGTGTGAGGATGCATTGGTGCCGTTCTATCAAAGATTCGGATTCCATCCGGCAGGTCCTTGCGC
CGTTGTTGTGGGTAGCCTAACTTTCACTGAAATGCATTGTTCTTTGAGGGGGCATGCTGCCCTAAGACGTAATTCTG
ACCGT

Seq. ID No: 139
>NAT_8
ATGTCTTCTGGTGGCGTTATCGTAGACTTACATAGGAACTCTACCAATTGGGCGAAAGTAGTAGACGACATAGTTAA
GTTAGAGAGGAAGATTTTCCCCAAACACGAAAGTCTTGCAAGGTCTTTCGATGAAGAATTAGGTAAGAAGAATACAG
GACTAATATATATGGAAGTAGACGGTGAGGTTGTTGGTTATGCTATGTACTCTTGGCCTAGCTCAATGTACGCTTGT
GTAACTAAGTTAGCTGTCAAGGAGAACTGTCGTCGTCAGGGTCATGGGGAAGCCCTGCTAAAAGCTGCAATTAAGAA
ATGTAGGACCAGAAACGTTCATCGTATCAGTTTACACGTGGACCCTCTAAGAAACCCGGCTATTTCTTTATACAAAA
AGTTCGGCTTTCAGGTGGATAATCTTATTGATGGGTACTACAGCTCCGATAGGAATGCCTACCGTATGTATCTGGAC
TTCGACGCTGAT

Seq. ID No: 140
>NAT_9
ATGGATGAACGTGTGGTGGTTGAGTTAAAGAAGTCTTTAGCTGACTACCCGAAAGTCCTAGAGGAACTTGTGAGAAT
AGAGAAGAAGGTGTTTCCTAAGCATGAGTCCTTGAGTAGATCCTTCGACGAGGAGCTAGGTAAAAAGAATAGTGGGC
TTCTGTATATTTGCTCAAATGGCGAAGTTGCCGGTTATGTGATGTATAGCTGGCCTAGTGCATTGCTGGCGGTAATT
ACTAAATTGGCTGTAAAAGAAAAGTATCGTCGTCAAGGCTATGGCGAAGCCTTATTGCGTGCAGCGATTCAAAAATG
TAAGACTAGGAACATCCAAAGGATCTCATTGCACGTGGACCCTTCTAGGACACCTGCAGCGAACTTATACAAGAAAC
TTGGCTTCCGTATTGACTCTCTAGTTGAAAAGTATTACGCGGCAGACCGTGACGCCTATAGAATGTACCTTGACTTC
GACGCCGACGTC

Seq. ID No: 141
>NAT_10
ATGATGGAGGGTGCCCAAGAGGACGAAGAAACGGAAGAAAAAGCTGAGTTCGATGCCAGCGAGATTGAGTACGTGAG
TTACGGAGGGGAGCACCACCTTCCTCTAATCATGTGTCTAGTTGACCATGAACTGTCAGAACCGTATAGCATCTTTA
CCTATAGATACTTCGTTTATTTGTGGCCGCAACTTTGCTTCCTGGCCTTTCACAAAGGAAGGTGCGTAGGAACCGTG
GTTTGTAAGATGGGTGACCACAGGCATACCTTCAGGGGCTACATTGCTATGTTAGTTGTAATCAAACCTTACCGTGG
AAGGGGCATAGCTACGGAGTTGGTTACTCGTGCCATCAAGGTAATGATGGAGTCAGGGTGCGATGAAGTTACATTGG
AAGCTGAAGTGACGAACAACGGGGCGCTCGCACTGTATGGCAGATTAGGATTCATAAGGGCCAAGAGGCTTTTCAGG
TATTACCTGAATGGTGTCGATGCGTTCAGACTGAAATTGCTGTTCCCCCGTAGTGAAATGCACCCCAGTTTACACCT
TCTAGCGGACCAAGACGGACATGATGACCAGATTGCAATGGAGGGGGAAGCT

Seq. ID No: 142
>NAT_11
ATGAAGCAAGTTGGGATTAGCCTAGATGCAGTTAGAGAAAAGAATCTAATGCAGTTGAAAAAGCTAAATGTGGTACT
ATTTCCTGTGCGTTATAACGATAAGTATTACGCCGATGCTTTAGCGTCAGGTGAGTTTACAAAACTGGCGTACTATT
CCGATATTTGTGTTGGCGCTATAGCATGTAGGCTTGAGAAGAAAGACCCCGGTGCCGTACGTGTGTACATAATGACA
CTTGGCGTTTTAGCTCCCTATAGGGGTCTTGGTATAGGAACTGAGCTTCTTAATCACGTACTAGAGCAGTGCTCCAA
GCAAAACATTAGTGAAATCTATCTTCACGTCCAGACAAACAACGATGATGCGATAAACTTTTACAAGAAATTTGGTT

TTGAGGTAACGGAGACGATTCAGAATTATTATACCAATATAACACCTCCCGATTGTTATGTCGTAAGTAAAAGGCTG
GAAGCTCAACCAAAGAAG

Seq. ID No: 143
>NAT_12
ATGAATATACGTGTCGCAAAGGTTGAGGACCTTATGGGAATGCAAGCATGTAACTTGCAAAATTTACCCGAAAATTA
TATGATGAAATTTTGGATGTATCACAGCATGACCTGGCCACAAATTTCATTTGTCGCCGAAGATCATAAGGGGCGTA
TTGTGGGGTACGTACTAGCAAAGATAGAAGATCCGTCTGAAGAGGGTACGACTGAAGAGATTCATGGCCATGTTAAT
TCAATATCCGTGCTTAGGTCCTACAGGCGTCTAGGCTTAGCGAAGAAACTGATGCTTCTAAGCCAAGAGGCAATGTC
TAGTATTTACAAGGCTTCCTACGTCTCTCTTCATGTCCGTAAGTCAAATAAGGCGGCCATTGCCCTGTATAAGGACA
CCCTTGGCTTTGAAGTGGCTAAGGTGGAGAAGAAATACTACGGGGACGGTGAAGATGCGTTATCAATGAGACTAAGT
CTGAAGAACCCT

Seq. ID No: 144
>NAT_13
ATGAGTGACTTCCAAGTTGCTCCACTTACTGCAAGGGAACTTGCTAGGGTCAGGGATTTGCACGCAAAACTGCTGCC
GGTGCAGTATCCAGTAAGTTTCTTTATCCACCTATTAGTGATACCCAGTAGGGCCTGTTACGTCGCCTATAGTCATG
GGAGTCCAGTGGGCTTCATTAGTGCTGCACTACACAATCCTACCAGATGCTTCATATCTGGGGATTCTGAAGTGAGC
CCCCGTTTGGAGATTTTAACCCTTGGGGTGCTACCCGCTTTTCAACACAGAGGTTTGGCCCGTAGACTAATCATGAG
TCTTGTAAACGCCTTCAAGCAGGACCCGGCTACCCCTATTTTGATATACGCAAACGTATCAACCACAAACACTCGTG
CGTTACAATTCTATGAAAGAATGGGCATCTTGGTTTCTTCCGACATAATAACTAATTTGTACAGAACTTTATCTTAC
GGGTCTCGTGATGCTTACCTAGTGGTAGGAGCCCTT

Seq. ID No: 145
>NAT_14
ATGTTGTCTATACATCCTCTGAAACCAGAGGCGCTGCACCTTCCTCTGGGCACATCAGAGTTCCTTGGGTGTCAACG
TCGTCACACATTACCCGCGAGCGAGTTTCGTTGTCTAACGCCGGAAGATGCCACTTCAGCATTCGAAATAGAGAGAG
AAGCATTCATAAGCGTCAGCGGCACATGCCCGCTGCACTTAGATGAAATTCGTCATTTCTTGACCTTATGTCCCGAA
CTTAGTCTTGGATGGTTCGAAGAAGGATGTCTGGTGGCTTTCATTATCGGATCTCTATGGGACAAGGAAAGACTAAC
ACAGGAATCTCTAACTCTTCACAGGCCCGGTGGGAGAACTGCACATCTTCACGTGTTAGCTGTACACAGAACTTTTA
GACAGCAAGGAAAGGGCTCAGTGTTACTTTGGAGATACTTACATCATCTAGGGTCTCAACCTGCGGTTCGTAGAGCG
GTCTTGATGTGCGAGAATGCGCTAGTGCCATTCTATGAGAAGTTTGGTTTCCAGGCAATGGGACCCTGTGCTATCAC
TATGGGTAGTTTAACCTTTACAGAGTTACAATGCTCACTTCGTTGTCATACATTTCTAAGGAGAAACAGTGGATGT

Seq. ID No: 146
>NAT_15
ATGGCTGATGCCCCTTCCGGCCCCTCTGTCCTTTCTCACTATCCTGGGGCCGGTTTAGCCTTGCCACCCGGTGACGA
GCAAGAAGATGGTGAAGAAGAGGAAGAGGGGCGTTACGAACCAAGACGTGGCCACCATCACCATCGTCGTCACCACC
AACAACAACAGTTGAATGGGCTAATATCACCTGATCTGAGGCACATTAAAGCCCTTAAAAGTAAATTACCACCTCCT
CCGCACGATGAAAGGACCGGCGCACCGAATGGACTAGAGAGACTGCAAGACCTTGAGGAAGAGGAAGCCGTCTTGGC
TTCTAGGATGGGTGCCTGTTCCCTTCACCCTGGTGATGGCAGCATAAGGTACGTACGTTACGAGTCAGAATTACAGA
TGCCAGACATAATGCGTTTGATTACAAAAGATTTGAGTGAACCTTATAGCATCTACACTTATAGATATTTCATACAC
AACTGGCCCCAACTGTGTTTTCTAGCTATGGTAGAAGAGGAATGTGTGGGAGCCATAGTATGCAAATTAGACATGCA
CAAGAAAATGTTTAGACGTGGTTATATCGCAATGTTGGCCGTAGACTCAAAATATAGAAGGAAGGGTATTGGAACGA
ACCTAGTGAAGAAGGCAATTTATGCTATGGTGGAAGGCGACTGCGATGAGGTCGTGTTAGAAACGGAGATAACAAAT
AAGTCAGCCTTAAAGCTGTACGAGAACCTAGGATTTGTCAGAGACAAAAGGTTGTTTAGATACTACCTAAACGGTGT
GGATGCGCTTAGGCTTAAACTATGGTTACGT

Seq. ID No: 147
>NAT_16

ATGGATGCCGCAATGCCAACAGAAATTTCTTTTCGTCAGCCAACTCCTGACGACGCGGCCAGATGTTTCGAAATTGA
AACTAGCGCATACGAGGGTGATGAGGCTGCTACGCTTGAGAAGATCGCAACAAGGATCGCTTTATATCCGGAGGGCT
TCGTCATCCTGGAGGCGGACGGAAAAATCGCAGGGTTCATAAATTCCGGTTGTGCGTTCGAGGTGGTGATGAGTGAT
GAAGAATTTAAGGAGTTGGTCGGTCACGACCCTGCTGCACCTAATGCTGTCATCATGTCCGTAGTGGTAGATCCTGC
AGAGCAAGGCAAGGGCTATTCTAAGCTACTAATGCAACATTTTATAGCCAGGATGAAAGCTATGGACAAAAAGACAA
TCCATTTAATGTGTAAAGAGGCCCACGTCCCGCTATACGCTAGAATGGGCTACCGTTACACACGTCCATCTGCTTCC
GATCATGGAGGCATGGCTTGGCACGAGATGGTGATGGAATTG

Seq. ID No: 148
>NAT_17
ATGGAAGGCTTACACAGCGAATGGGAAGTAGGCGCAGAGTTAAAGGCTCTGGGTGCAGTTCCGAAACCTTTCATCGG
AAGTCATGTATCAGGTAAACTGATTCAGCGTTTAAAGCAAGACCTAAGGCAGAGCTGGGATCGTGGGCAATCCCAGG
CCAGACCGACTTGCACCTTGCCTCAACCTTTGCCTGCGCCTCTAGGCAGTTCCGTGCCGAGCGCGTCCGCCCAAACG
CAAGTCAGTAGGCTTGTTCCAGTAGCCCCGCCACAACCAGATCCTGCCATGAGTGTCTTAAACGCAGTACCCTTTAT
GAGACCCATCCACCTTAGAAGCCCGAGACAGCAGCGTAGGCACACATTGCCGGCCTCCGAGTTTCGTTGCTTATCTC
CGGAAGATGCTGTTAGCGTGTTTGAAATTGAACGTGAGGCTTTCATTTCAGTGAGTGGGGATTGTCCTCTGCACCTA
AACGAAGTAAGACACTTTCTAACCCTGTGCCCGGAGTTAAGCCTTGGATGGTTCGAGGAAGGGAGGCTAGTCGCTTT
CATCATTGGTAGTCTGTGGAACCAAGACCGTCTTTCTCAAGATGCATTAACTTTGCATAAGGCCGAGGGTAGTAGTG
TCCATATACACGTTCTTGCAGTTCACAGAACATTTAGACAGCAGGGTAAGGGGAGCATCTTACTATGGCGTTACCTA
CAGTACCTAAGATGTTTGCCTTTTGCCAGAAGAGCGGTACTGATGTGTGAGGACTTTCTGGTTCCATTCTACAGCAA
GTGTGGGTTCAAGGCAGTAGGTCCGTGTGATATAACCGTTGGTCCATTGACATTCATTGAGATGCAGTGTCCAGTAC
AGGGCCACGCATTTATGCGTAGGAATAGCGGATGC

Seq. ID No: 149
>PSIK_1
ATGCAGGCTAACAGACCTATATCCGACCAAGATCAAGACCAGTTTAAATTGAATCTTACCACTGCAGATGGAACTAG
GTCTTATCTGGAGAAGCATTTGAGCCTGAACGTGGAGGCGGTTGAACGTTTATCCGGTGGGTTCATAAACTTTGTTT
GGAGGGCTAAGTTAGGAACCCCCTACGAAGGTCAGAACTCAATTGTTGTAAAGCATGCACCCCCTTTCACAGCAATG
GACTCATCTCTTAATGTCGCGGTAGAGAGGTTGAAATTCGAGTATGACTCTCTTAAAATGATAGGTAGCGAGCCAAG
TATTGCCGGCGAGGATGCCCTTATATCAGTCCCCTCTGTGTACCACCATGATAACATTAAGCATGTGTTGATCATGC
AGGATGTTGGAACCATGTCAACCCTGAGAGATTTCATGGGAGCATCCCCTCCTCCGCCAACTGACATGGCTGCACTG
ATTGGGTGCCAACTAGCCACGTTCATTGCAGGACTGCACAACTGGGGCAGAAACAACGAATCTGCTCGTGCTGGTCT
ATCAGCAAATGCTTATGGGAGAACGGTCATGGACCTATGTGGATACCAGACAGTGGTGCCTAACGCTACGGCAAGTG
GGATTCTGGACCCGCTGTTATCAACTGCAATGGCAGCTTTAGCCGAGCGTGATAAGACCTCCGAAGAAACAGCCATC
ATGGGAGACTTCTGGGCCCTTAATGTACTTGTTGATATTGACATGTCAGCATCAGGAGAAAAAGCCCTTAAAAACAT
ATGGATCGTCGATTGGGAAGCATGTAGATATGGCTCACCAGCGGTGGACGTAGCGACCTTCGCAGGAGACTGTTATT
TAATTTCACGTATTCACAATGAAACCGCAACCGACGCGATGCGTCGTAACTTTCTTGGCACGTATGTCGCATTAGCT
AAGGTGGACCCTATGGAAGTTGTGATCGGAATGGGCACGATGTGGATTATGTGGACTAAGTACCAAGAGGATATAGG
GGAAGCCGAAAAGAGGGAGAGGGTCGCAAAAGGCGTCGAATACATCCACAAAGGATGGGAACGTTCCAGAGAGTGGC
TGCCTGTCTCATTAGCGCAGGAGTTGATCGCA

Seq. ID No: 150
>PSIK_2
ATGGACCTGACCACCGGAGACGGCGTTCGTGTATACTTGACGGCTCACATGACGCTAAAGGTTGAGAGCACAGAAAG
ATTGTCCGGTGGTTACTGTAACTTCGTATGGAGGGCTAAGTTAAAGACACCTTATGAGGGACAAAATAGTGTAATCG
TGAAATACGCGGCCCCCTTCACTTCCTGGGATCAGACTATTGAATTGGGCGTTGAAAGACTAGCATTTGAGTGCATG
TCACTAAAAATGATTACGTCCGAAACTCCACTGCTTGAGGAAAATGGGTTAGTTGCAGTACCTACAGTATATCACTA
CGATTCCACAGCCAATGTGTTGGTCATGCAAGATATTGGTTCCATTGCGACTTTGCATGGGTTTTTGCGTTCCAATA
CTCCCCCAACGGTGCCGATGGCCGCACTAATCGGGGCAAAGTTGGCCGCATTCATTGCGGGGGTGCACAATTGGGGA
AGGAACAACCTACCAGCCCACACCAGATTAAGCGCCAATACTGTCGGCCGTACTGCAATGAAGAAATTGTGTTACGA
AACGATCGTTCCAAAGGCGGCAAAGTCCGGGGTGGTCGATCCCCTGCTTCCGATGGTAGTGGCCGCTCTATCTGAAG

AGGCGATGACGAATGACGAAACCTTGGTAATGGGAGACTTTTGGACTGCGAACGTACTTATTGACGTTCAAGAGTCA
CATACCGGGGAGCAGGTCCTGAAGAAACTGTGGGTGATTGACTGGGAATCATGCCGTTATGGTAACCCCGCGACCGA
CATCGCGAGCTTCGCGGGAGACTCTTACCTTGTGTCCCGTTTCCAGGACCACGGTTTAGGTGAAGCCTTGAGACACA
GTTTTCTTGAAACCTATGCGGCGTTAGCAAAAGTCGATCCACTACGTGTTGCTTTGGGTCTGGGCGCGCACTGGATA
ATGTGGACCGATGATCTAGGGCAAGGTGGAGAAGCGGAAACCCGTGAATGCGTCGATAAAGGCTTAGAGTATATTCA
ACGTGCGTGGGATCAAAGCGCCGAGTGGGTCTCATTGTCTTTGGCAAAAGAGTTGGTAGTCTTG

Seq. ID No: 151
>PSIK_3
ATGGCAAACGAGAACCCTGACTTACTTACAGTCGCGGGCGTTCTAAGGTTTCTAGCTCCAACTCCTTTCGCTAGTGA
TGAAGTTCATCCTCTTTCCGGGGGCAATTGTAACTTTGTGTATAGGATTCACTTGAGGACACCGTACAATAATATAT
CTACTCTGGTCTTAAAGCACGCGGAGCCTTACGTCGCTGCGTCCGCTCACAGGATGCCCCTTGCCGTGGAGCGTCAG
AACACGGAAGTCACTGCCATGAACGCAGTCAAGGCAATTCTATCATCCGATGCGGTAGTTATAGTTCCTACAATCCA
TCATTTCGACGATGTCGCACATGTAGTTATCATGGATGATTGTGGAGTCGGAGCCGTCACCCTAAAGCAATTGATGT
TGAAAAACCCACCTCCTGTTTCCGTGGCAAAAGCGTTGGGGGCAGGTCTAGGAGAGTTTTTAAGTCGTCTACATGTG
TGGGGGAGGGACCCGCAGACGTCTAATCATGTCAGTTTCGATCAAAACCAACAGGGAAGGACAATAAGCGGCTATGT
TACATATGGTAGGTTAGTCAGCACCCTTACTGGTAAAGATAATATCCCGGCGTTATCCGACCCTCCGCTAGACATAG
CGCAAAGTAAATTGGATACGATCAGTGCACTGTCTAGTGAGAAAATCCATGCCATAAACACTTCACATCAGACCCTG
ACTATGGGTGACTTCTGGCCTGGTAATATTATGGTTAGGCTGAATCCTGCGGGTGATAGTTTGGAAAGGGCTTATGT
GTTGGATTGGGAGGTTGCTAAGCCTGGCGTGGCCGGTTTAGATATCGGCCAGTTCTGCGCTGAAATGCACAGTCTAA
GAAGGTTCAGTCCTGCCTGTGATGCTTCAGCTACCACCGTTCTGGATGCTTTTCTGAAGACCTACAGGGACGCAGCC
GGAGTCGATGTAGGGGTAGCCAAAGATGCGATGGTCCATGTCGGAGCGCATTTAGTAGCGTGGACGCCGAGAGTGCC
GTGGGGGAGTAAAGAAAGAACTCGTGAAGTGGTTGAAGAAGGCGTTGGATATTTAGTTGAGGGATACGCAGCGACAC
AAGAATGGCTACGTGGCTCTCTTGTGGGTAGGCTAGTA

Seq. ID No: 152
>PSIK_4
ATGGAAATCGAATGGTGTGACCTAGACCTAACTTCAGAAAGTAGACCTCCTACCCATAAATATACTTATTTCGCCAC
TGCCCTTATGCCCTTTGATTTAACTACGAGGGATGGCGTCCGTATGTACTTGACCGCATACCTTGCTTTAGACGTGA
TGTCTGTAGAAAGGTTGTCTGGCGGGTATTGCAATTTCTCCTGGCGTGCGAAGTTGGAAAGCCCTTACGAAGGCCAA
ATCTCAATCGTCGTAAAACACGCGGCGCCATTCACCTCTTGGGATAGAAACACAGAACTTGGTGTTGAACGTCTTGC
TTTCGAATACAAGGCATTAAAGATATTAAACAGTGAACCGTCAGTCATCGCCAAGAACTCACTGGTCGCTGTACCGG
CTGTTTACCACTACGACCCGACCGCCAACGCATTGATCATGCAAGACGTAGGTTCAATACCAACCTTGCATGCATTA
CTAAGGAATAATGCATTACCACCAGTCCCAATGGCGGAGAAGATCAGTAATGAATTGGCGGCGTTTATTGCAGGCAT
TCATAATTGGGGAAGAAACAACCAAGAAGCAAGAGCTAATTTGAGTCAGAATTTGGTTGGCAGGACAGCGATAAGGA
AGTTGTGTTACGAAACATTGGTGCCAAAAGCGGAGAAGTCTGGTGTTGATGACCCTCTATTGCAACAAGTTGCTGCG
GCGTTATCCGAGGAGGTCATGAACTCCGAGGAAACGCTGGTGATGGGTGATTTCTGGACTGCAAATGTGATGGTTGA
CATACAAGAAACCGGGGCAGGAGTAAGGAGCCTAAGAAAGATATGGGTTATAGACTGGGAGGGCTGTAGGTACGGCT
CACCTGCGGCAGACATAGCTAGTTTCGCTGGGGACTCTTACCTTGTCGCAAGGTTTCACCATCACGATCTTGGAGAG
ACTTTAAGGCATAGCTTCCTAGAGACCTATGCAGGTCTAGCAAAAGTAGATCCTTTCAGGGTCGCGTTAGGTTCTGG
TGCACACTGGATAATGTGGACTGATGATCTTTCTGAACAAGAAGAAGGTGAAATCCGTGAATGCGTAGACAAAGGGG
TGGAATACATTCATCGTGCATGGGAGCAATCCACTAAATGGATTAGCTTGTCACTAGCCAAGGAATTAGTCACA

Seq. ID No: 153
>PSIM_1
ATGCATTTCAGCTACGATTTTACTATGTGTAACCCGCCGTTTTATGGTGACTACGCTGAGCTAACGAGGCTAAGAGA
ATCCAAGTTGAAGGGACCGTTCGGGGGCGCTCATGAGGGAGTAAGTACTGAACTATTTACTGCTGGAGGCGAAATAC
ATGAGTTCATAGCAAGTGAACTGAACCAGGGCAAAACGCTGCGTTGGGTGGTAGGATGGACGTTCCATAAGGATTTA
TTCGATAAAAAGGTTATCGAGCCATTACATATTTTAATAAAATTGAAGCCCCTGAGCTGTAACACCTTGGAGCTACT
AGATTCCGACCACGAACTTGCTACGCCCTACAGCAAGAAGCGTTGTCTTAGGCGTATCTCCTTTACGCCCAAGGGGA
GGCCACCTACGAAACAAACTATAAGTGTAATGCACCCCAATGAGGCTTCGTGGGCATCTCTGCTAAATCACCTTCAA

GCATTGGACATTTCCGTGACCGTAGGCAACCACTTTCTAACAGCGGAAGTGAAAGACCCAACGTGGACAAGAGCGTG
GAGACGTAGCTCTAAGGTAACTTCTAAGATCACCCCTTTTAGTTTCTCTGGACAGTTTTCCGATCCACCAGAGAGGT
TGTTAGTATTACAGCTTCTTGTCGATGAGAGTCAGACCAGCGAAGACATTCTACTGTCATTCCAGAGCTTG

Seq. ID No: 154
>PSIM_2
ATGCACCCTCGTAATCCTTACAGAACCCCTCCGGATTACGCAGCTTTGGCGAGGAGCTTCCCGGAACTGAAGCCTTA
CGTAAGTAGGAACGCCAACGGCACAGTTTCCGTCGATTATCAAGATGAAGCCGCCCTAAGGTGCCTAACCAGAGCCC
TACTGTATAGAGATTTTGGACTATCCGTGGATCTTCCCAAGGATAGGTTGTGTCCCACCGTTCCAAACAGGCTAAAC
TATATTCTTTGGATTGAGGATATTTTGAACGTCTCATCTTTAAGCAGACTACAAAGCAACTCTGAGGCAACTGTGAG
AGGTCTTGACATAGGAACCGGCGCCAGCGCGATCTACCCGTTGCTAGGCTGCAGGGTTAGTCCAAGATGGGATTTCT
ATGCCACGGACATAGATGCCCAGAGCCTTGCCCATGCCAGGGCCAACATAACCCGTAATGGGCTACAGGGAAGGATT
AACCTGGTTGCTGCGGACCCCAAGGGTTCTATATTTGGACCATTGGAGAGCAAACACGATACGACTTTTGAGTTCAC
CATGTGCAATCCCCCATTCTATTCATCAGAGGAGGACATCGCCCAGAGCGCAGCCGTTAAGAAACTTGCTTCTAACG
CCGTCTGTACAGGGGCAGCAGTCGAAATGATCACTCCTGGAGGTGAGGCTGCGTTTGTGGTTAGGATGGTAAGAGAG
AGCCTTGCTCTAAAAATGAGGTGTCAATGGTACACGTCCATGCTAGGAAAAATGAGTAGTCTGACAGAAATAGTGGG
GCTGTTAAGGGAAAATAGTATTGATAACTACGCTATCACGGAATTTGTCCAAGGGAAAACAAGGCGTTGGGCGATTG
CATGGTCATTCGGCCACGTAAGGCTACTTGATAGTCTGGCGAGACTTTCAGGTGGGCCACTTCAGTCATTAATGCCA
ACTAGGAACACATGTCGTAGGTCTTTCGCCGTTCCTCGTATGGTCTTACATAAGCATCTAATTTTGGTGCTTGACGG
CATAGAGGGAACATCCCAGACTCCCATGTCAATCCCAGTAGGCGCTGGGGACGCCGACGGGTTGTACGGGCTACAGA
TCTCAGCTAGTAGAGACACTTGGTCCAGAGCGGCGAGGAGAAAAAGGCAACACGGTGCAATGGATATATCATTAGAT
AATGATGAAGTAGGGATGAAGTGCCTGATAAAGGTGTTATCCGTAGAGGAAGCAAGAGAAGGAGCAGAGGCCGTAGT
GCTAGAGTGTACCTGGGTGTATGGGCATGAAAGGGCGTTATTTGAGAGTTTTTGGGGCCACGTATGCCGTAAGATTG
GGGAAGCAAACGGG

Seq. ID No: 155
>PSIM_3
ATGCACGCGAGAAGCATATTCAACCCTAATTCAGCTCAGTTTCAGGCACGTCTGACATTTTCTGAGTTGTCTAACGA
GTTTCCCAAGTTGAAACCATTCTTGAAGTATAAGAGGTCCAGGAAGCAGAATGAAGCGGACCCTTTATCATCCCAGT
GTACGTTTATTGATTTCAAAGATCCAGTCGCTACGCGTGCCTATAATGAGGTGTTACTAAAGAAGTACTTTGACCTT
AGCTTGGAATTTCTGCCTGGCTCATTATGTCCGGCGGTTCCCAATAGATTAAATTACACCTTGTGGCTAGAGGACGT
CTTGAACGTCTTCCCAGGGAGTATGGGGGCAAATAACCAACGTGATGAGCTAAGAGGACTGGACATAGGAACTGGTT
CATCATGCATATATCCCCTACTTATATGTAGGACGCACCCGAACTGGAGGATGGCGGGGTCAGACATCAATCCCTCT
TCAATAGAGATCGCCAAGAAGAACGTGCAGGAGAATAGGCTGCTGGATCGTATACAGCTTTTCTTAACAACCGATAA
GAGGGATAGCGTCCTAGAGGGGCAAATATTTCAGACCCACTTGTTCTTCAATTCAAAGAAATGCCTGCAAGATGAAA
AACCTGCCAGATTCTGCTATGACTTCACGATGTGCAATCCACCATTTTATTCCGACGTTGAAGACCTTAATAACAGT
AGACAAGCGAAAACCACGACCATATTAGGTGGTGGACATGAAGGGGTGTCGTCAGAATTGTTCACTACAGGTGGAGA
ATTACTGTTCTTATCGCAGATGGTGGAGGAAAGTTTCTTGTATAAAGACAAGGTGGGGCTTGTTAGTTCATATTTCT
TTGTGCTTAAATGCATAGTGGAGATTCTTCGTCTTCTGGGACAGCATAAGATACAGGAAACTATTGCGTCTAAGCTA
ATACAAGGCAAAACCATAAGATGGGTAATAGGATGGACGTTCCACAAAGACCTTTTTGATCTGAAGCATCCCTCCTG
TAACACTCTGAAGGTGCAGGATATACATTGCACGAACGAAGAAGCCCCTGCGTCTAAGAAAATATGTCTAGGGAAAA
TTCCCCTTAACTCTAAGGAACACATTCCGTCTAAGCAGCTTATATCTTCATCCCAGTCAAGCGACATTGGCTGGACC
AGACTGAAAAGCAGACTTGGGGACCTTAGAATCGAGTTCTCTCTTGAAAAGTTATCCCTGATTGGAAAAGTTGTCTA
CCCTACCTGGACGAGGGCTTGGAGGAGGAATGGAAAAGCCAAAAGTAAACCAACCCCTTTTTCCTTCTCCGCGCAGC
AGTCAGATGCAAGTGAGACAACAATACAATTGGAACTGTTAGCGTCCGAGATAGAGGAGCCCGACGAAAATACTCTA
GTCAGCTTTCAAAGTTTGTGTAATCACCTGCGTAGCTACTTGAAAGAC

Seq. ID No: 156
>PSIM_4
ATGAGCGATATAGATGCTCAAAGTCTTGTTTACGCCCGTGCGAACGTGGCTCGTAATGCACTGGAAGGCAGGATTGC
TGTCGTGACAGCCGAACCTGAAGGCTCCATCTTTGGGCCGATTGAAGCGGAGAAGGAGATACAGTTCGACTTTACTA

TGTGCAATCCTCCCTTTTACGCTTCTGCGGAAGATATCGCTCAGTCAGCAGCAACGAAGGAATTGGGTCCAAATGCC
GTTTGCACCGGTGCGGCGGTTGAAATGATAACGTCTGGGGGCGAAGGGGCGTTTGTTGCCAGAATGATTGATAACTA
TGCTATAACTGAGTTCGTCCAAGGGCAAACTAGGCGTTGGGCAATAGCATGGAGCTTTGGTCACGATAGGTTACCAG
ATAGCCTAGCCAGGCTATCATCCGGCCCCCTACAATCATTGATGCCCACGAGAAATACATGCCGTCGTTCTTTCACC
TTTGCGAGGATGAACTTGCTTAGTAGGCTAGAGCAAGTGCTAAATAATATAGAGGGGTTGTCCCACAGCAATATGTC
TCCTAGCGAAGATAGAGGAAGCGGGGGAAGGCCATCTAGCTTACTTGTATCCGTTGCGAGGGATACCTGGTCCAGGG
CGGCTAGGAGAAAGAAGCAGCGTGGCAGCATGGATACATCACTAGATAACGATACATCAGGACTTATCTGTAGCGTC
AAGGTATTATTCGACGAGGAGGGCAGAGAAGGAAGCGAAATAGCATCATTAGAATGTACTTGGATTCATGGGAGGGA
ACGTGCTTTATTTGAATCATTTTGGGGACATATCTGTAGAAAAGTAGGGGAAGTTTCTGGT

Seq. ID No: 157
>PSIM_5
GAGGGACTATCCCCCTTCCTGCTTATGCACCCAAGAAATCAATACTGTAAAAAGAAGCCGGATTTCGCAGATCTAGC
TAAAAGCCACCCTCCGTTAAGAGAACATTTAAAGTGGAAAACTGAGGACTACGCGACTATTGATTTCAAATCTCCAT
CCGCACAGAAGGAGTTGACACGTGCATTACTGAAACAAGACTTCCATCTGGACGTTGATATGCCTGTTGACAAATTG
GTCCCAACAGTTCCCCAGAAATTAAATTACATTCACTGGATTGAAGACCTGCTGAGCGGCGGCCGTTCAGATTCTAT
CCCGCGTGGAGAAGGGATCAGGGGCATAGATATAGGAACCGGGCCGGCGTGCATCTACCCGCTACTGGCGACTTCAC
TTAATAAATGGACCTTCGTCGCGACTGATATCGACGCTGTTTCACTAGAGTACGCTGTTAAGAATGTGAGCCGTAAT
GATATGGAAGGTCGTATCAGGGTGAAAGGGGTGGACCCAGATACGTTACTTGTCGGTGTGGTCCGTGACGAACAGTT
TGACTTCTGCATGTGTAACCCGCCATTTTATGGGATCGATGAAGACCATCACGACAACCAGAGGCCCCCACCTCCCT
ATAGCAGCTGCAGTGCCCAGGCACACGAAGTCCGTGTGCAGGGAGGTGAGGTTGGCTTTGTAAGTAGAATGGTTGAG
GAAAGCCTGCTGTTGCCCTCCCGTGTGCGTTGGTTTACCTCTATGGTGGGCAAGAAAGGATCTCTTAAAAGTCTAAG
GGCACTGCTGAGGAAGAGGGAGGTGCCCACGGTAACCACGACGGAGTTCGTGCAGGGTGTAACAAAAAGGTGGGCTG
TTGCATGGAGTTTTACCGAACAAGTACCATGCATCCCTAGCCACTCACTTCCGTGTACGGTGCCATTACTGGGTTCT
ACCAGTAGCGCAGAAGGACGTGCATACGCTGAACAATGGTTAGAAAGGGTCCTTAACCATATGGAGGTAACATTCAC
AAAGAAGGATCAGGATGGGTACACTTGCACCGCAGAACGTGCTACTTGGGCTAATCAGAGGAGGAAGAGGAGGCTTA
TGCAGCGTCCGATGATGTCCCCCGAGGCTGCAGCAAAGAGATCCTGCGGAGGAAGCGATAATACGAGTGAAGGAGTC
CCTAGGAATGACTCAGACACCTTAGTCAGCGCCGGACACCTAAGTCCGAAGGCTGACAGCCTAGAAAGGAACGCTAG
TTCCGATCTTGCGGCTCAGCTAAGTGCCTTGACACCACCGTACCACGTAACCTTCTGGTGCGGAGTCCAACCAAGCG
TACCTCCCAGCACTAACAAAGCGGAGTTAGAACTAAAGATGGTACTGATTGACGGCGGAAGCGGTACTCAGCCGCTA
CAGCCCATTGCTCAATACATGAAGAATAACTGGAGTGCCACAGATTCAAGACCAACATCTGATCGTAGTAGCCAA

Seq. ID No: 158
>PSIM_6
TGCATCTACCCACTGCTAGGGGCCACGATGAACGGATGGTATTTCTTGGCGACTGAAGTGGACGACATTTGCTTCGA
TTATGCCACAAAGAACGTGGAACAAAACAACTTATCTGACTTAATTAAAGTAGTTAAAGTCCCCCAGAAAACTTTGC
TTATGGACGCATTGAAGGAGGAAACCGAAATAGTCTATGACTTCTGTATGTGCAACCCGCCCTTTTTCGCGAATCAG
TTGGAAGCCAAGGGGGTCAATTCCCGTAATTCCCGTAGGCCTCCTCCGTCTTCAGTAAACACAGGTGGAGTCACTGA
GATAATGGCAGAAGGGGGTGAATTGGAATTTGTAAAAAGAATTATTCACGACAGTCTACAGTTAAAGAAGCGTTTAA
GGTGGTATAGTTGCATGTTAGGAAAGAAGTGTTCTCTTGCACCTCTGAAAGAGGAGTTAAGAAAACAGGGGGTCCCT
AAGGTAACCCATACTGAGTTCTGCCAAGGGCGTACCATGAGGTGGGCCCTTGCATGGAGCTTTTATGATGACGTTAT
AGTGCCCTCACCACCTAATAAAAAGAGGAAACTGGAAAGGGCAAGGAAGCCGCTATCTTTCACTTTACCGGAAGCCG
GACTGAAGGAATTACAGAGTAAAGCACTTGCTCTTGGAGGCACTGCATGTAGCCCGGTGGACCGTGTAGCAGCACTT
TTGGAGAAAACTCTTACCGACCTTAGAGTCCTACATAAACGTGTTCCGTGTAGGAAGCAAGAACAAAGCCTATTTCT
AACTGCCGTTGAGAATACATGGATACATGGTCGTCAGAAAAGGCGTGAACAAAGTCGTCAGCTTAGAGAGCTGCCCC
GTGCGCCACCTTGCGCGGGTACATCCTCCCAAACTACAGTTGCTACGGCCGACTCTGTGAAAACCCCAGCATCACAG
ACTCAATCCGCATCTACTCAAAATTCCAACTCCCAAGACGATAGCAGTCAGAACAAGCGTGCGAGTGCGCAGGAATT
AGCCGGTCAACAACCCACGGATAAGGCCGGCAGTAGCGCATCTTCTGATGAAATATCAATTAAAGTGCTGCATAACT
CTACGGGAGAGCAAAAGGAGGTCACGGAAAATCTTTCAAGCGAGGCAGTAGACATGGAGTTCAGTACAAGCACAGAA
GCAGTCCAGGAGACGGGTTCTAAAGAGGCTCCCTCTGCTGAAAGCGAGCCTCCCAGCAAGAGACCCCTTTCTCCCGG
TACGGTCGAACAATTTCTATTTAAATGTTTACTAAACGTAATGCTTGAGGAAAGTGACGTCATGATCGAAATGCATT

GGGTGGAGGGTCAGAACAAGGATCTTATGAACCAATTATGCACTTACCTGAAAAACACTCTACTGAAGTCTGTCGCT
AAATCC

Seq. ID No: 159
>PSIM_7
ATGGGCTCAAAGAAGAGGCGTCGTAGACGTGAAGAGCGTCCGACAATTCATCCGAAGAATAAGTATAGTGAAAACCC
GCCCGACTTTGCGCTACTAGCCTCCTTGTACCCCAGCTTCGAACCTTTTGTTTTCTATTCTCGCGACGGCCGTCCCC
GTATAGATTGGACTGACTTCAATGCGACGCGTGAGTTAACAAGAGTTTTGCTGCTGCACGATCACGGCCTTAATTGG
TGGATACCTGACGGTCAATTATGTCCCACGGTCCCGAACCGTAGTAACTACATACATTGGATCGAGGACCTTCTGAG
TAGTAATATAATCCCGACCACTTCGAGAAATGGTGATAAGGTAAAAGGTTTCGATATCGGTACGGGTGCGAATTGTA
TTTATCCGTTGTTGGGAGCTTCTCTACTTGGTTGGAGCTTCGTAGGGAGCGACATGACCGACGTGGCGTTAGAGTGG
GCAGAGAAGAACGTAAAAAGCAACCCGCATATATCTGAACTGATCGAGATTCGTAAAGTTGACAATTCTGAATCTAC
CCCTAGTATTCAAGAATCCTTAACTGGGAAATCAGTGCAAGACGAGTCTAATATGGACATGTCCGGCCACATGGACG
AGGAGGCCGAACCAAGTTCAAGTTCTTCTTTTAACCTTCCAGCGGGCGCTCAAAGTTCCTATCACGGACCTCCGGTT
CTGGTGGGTGTTGTAAGAGATGGCGAACAGTTCGACTTCTGTATTTGTAATCCGCCCTTCTTTGAAAGTATGGAAGA
GGCTGGACTAAACCCCAAAACATCTTGCGGAGGAACTCCTGAAGAAATGGTATGTAGCGGTGGCGAAAGGGCTTTCA
TAACCAGAATAATAGAAGACAGTGTTGCTCTGAAACAAACGTTCAGGTGGTACACCAGTATGGTCGGCAGAAAAAGT
AATTTAAAATTTCTTATCTCAAAATTGAGGAAGGTCGGTGTCACAATAGTCAAAACAACGGAATTTGTCCAGGGTCA
GACGTGTCGTTGGGGTTTGGCATGGAGTTTCGTCCCGCCTGCTCGTAAAATCATCAGTCCGCATGTTGCAGAGAAAA
AGAATCTATCCTTTATGCTTGAATGTACGTTGATTAACAGGAGCTTATATCAGATGATCAATGTGACTCAAAGC

Seq. ID No: 160
>PSIM_8
ATGCATCCTCGTAACCCTTATAGGCAATTGCTTGACTTCGCCAGTTTGGCAGAGGCATACGAGCCACTAAAGCCCCA
CTTGAAACCTACCAGATCTCCCACCGCAGGCGGTTTATCATACACCATAGATTTTAAGAACTCCGAGTCTCAGAGGC
AACTGACAAAGGCAATCCTGTATAGGGATTTCGGGCTGCGTATTGCGCTGCCAGATCACCGTCTATGTCCGCCTGTG
CCAAACTCTCGTTTAAATTACATTTTGTGGTTACAAGACATCATTAAAGCGCACGACGAGTATATGGATAGGCCCGC
CTCTTGCATATGCGGATTAGATATAGGCACTGGAGCAAGCGCCATCTACCTTTTGCTTGGGTGTAGAGTAGAGCCCA
GTTTTAGGTTCATAGGCACAGAACTAGACGATATATCATTCAGCTACGCCACTCAAAACGTCGAGTCTAACGGCCTG
TCTGACAGGATACACCTGATAAAAACGACAAGCAACGATCCCATCCTGCTACCGTTCGATCTAAATCCGGCATGGTC
TTGCGATTTTACGATGTGCAATCCCCCTTTCTACGAGAGCGAAGAGGAAATGGCCCGTAGTGCGCAGGCGAAAGAAT
TAGCTCCAAATGCGGTTTGTACGGGGGCTCAAGTGGAGATGGTTACCCCCGGTGGAGAACTAGCGTTTGTATCACAA
ATTGTCAAGGAATCTTTAAAGTACACCACCCGTTGTAGGTGGTATACTTCAATGTTAGGTAAATTGAGCAGCTTAAC
AAAGTTAGTAGGACTGCTTAGAGAGTACGCGATCTCCAATTATGCTATTACGGAGTTCGTCCAGGGGCAGACAAGGA
GATGGGCTGTGGCTTGGAGCTTCGGCGAGACACACTTGCCGGATTCTGTGGCCAGGATTAGTAACCCTACGCTGCAA
CCCTTGCTGCCGGAGAGGAATACGAGTAGGCACGTGATCAATATCAGTCTTCCTCCGTTTAGCACAAGGACAGTGAA
ATCAAAACAGTCTATCAAGGCGCTTTCTGAGGTTCTTAGTCAGATCAAAGATGTGACTGTGCAGAGATTATATCAAG
TGGAACACCTAGAACCCACCGAAGAGGAAGAGGAGGACAAGAGTCTGTATAGATTGCTTGTTTACGCGAAGCAGAAT
ATGTGGAGCAGGTCTGCTCGTCGTCAAAGAGGAAGGGAGACTGGCCATAAAGCCAACGATAAAGGATGCGCTGTAGG
GGGTCCTTTAACTTCAATTCCGGCTACTTTAGATGGATTACTTTGCGGGATAGAGATAAAGGCACCGTTGATAAAGC
AAGAGCAACAGGATGTAGAGATGGAATTTGTATTTCAGTGGGTTCACGGGCAAGATAGAAGTATGTTTGAATCCTTC
GTGAACCATGTAACTAGAAAGATGAAATGTAATATAGTACTGGAC

Seq. ID No: 161
>PSIM_9
ATGGCTCTGAATAAATCCATGCACCCGCGTAACAGATATAAAGATAAGCCGCCAGATTTTGCGTTCCTAGCATCCAA
GTACCCGGAATTTAAGCAGCACGTTGACGTTGGTTTAAGTGGTAAGGTAGGACTGAATTTTAAGGACCCAGGCGCGG
TCAGGGCTCTCACGTGTACACTTCTAAAGGAGGATTTTGGGCTGACTATCGATATCCCCCTTGAGCGTCTTATTCCC
ACCGTTCCATTGCGTCTGAACTACATACATTGGGTAGAGGACCTGATAAATTTTCATGATAGCGACAAGACTACCGT
GAGAAGAGGTATAGATATAGGCACAGGCGCAAGTTGTATCTACCCCCTTTTAGGAGCAACATTAAATGGCTGGTATT
TCCTAGCGACCGAGGTAGATGATATTTGCTACAATTACGCTAAGAAGAACGTGGAACAGAACCATCTAGCCGATCTT

ATTAAGGTTGTCAAAGTCCCGCAAAAGACTCTACTTATGGACGCCCTGAAGGAAGAGTCAGGAATAATATATGACTT
TTGTATGTGCAATCCACCATTCTTTGCAAATCAAATGGAGGCTCAGGGCGTCAATTCCAGGAACCCCAGACGTCCTC
CTCCGAGCAGCGTAAATACGGGAGGGATTACCGAAATAATGGCAGAAGGTGGAGAACTTGAATTTGTGAAGCGTATC
ATTCACGATTCACTACAACTGAAGAAGAGGCTTCGTTGGTATTCCTGTATGCTGGGGAAAAAGTGTTCTCTAGCGCC
ATTGAAAGAAGAGTTGAGAATACAGGGGGTGCCAAAGGTAGCCCACACCGAATTTTACCAAGGCCGTACTATGCGTT
GGGCGCTGGCTTGGAGTTTCTATGATGACGTAACGATCCCTAACCCTCCGTCTAAAAAGCGCAAACTAGAGAAACCA
AGGAAGCCGATGATGTTCTCTGTTTTAGAAACTACCGTGAAAATGTTGATGGACAAGTTCGATTGTAGCGTAGACAG
TGAGCACGTCTCCGTCGTAACGGATTGTTTGAAGAAGATCCTTACCGACTTAAAAGTCCAACATAAACCGGTTCCAT
GCGGTAATGGGGAGGAGTCTCTATTTCTAACTGCGATCGAGAATAGTTGGGTTCATATAAGACGTAAAAAGCGAGAT
AGGATGAGGCAACTTAGAGAATTACCCAGGGCACCTGATGAAAACTTTCTCCTAGTTCAAAAGGACGAAAGGCAGGC
TGAGGACGAGGAAACGACTGAAAAGACAGTGTCAAGCAGTGAGAAATCCGTCAGTACATCAGGCATTGATGAAGCTG
CCGCCCTACCCCCAAATCCGGAGGATTCCATCTCCGAAAGTATGGGGGAAGACAGTCGTCAGCTGCCCGAGGAAGTT
AAAGATACATCCGCACTAGGGCAGATAACCGACGTCGATGAGCACCAAAATACAATGGAGGCGTCTCAACCTTGTTC
CTCAAATAGTGCCTTCCTGTTTAAGTGCCTAGTCAATGTTAAAAAGGAAGCTACCAATGTTTTAGTTGAAATGCACT
GGGTGGAAGGTCATAATCGTGATCTGATGAACCAACTGTGCACTTACCTAAGGAATCAGATTTGTAAAATTGCAACG
TCA

Seq. ID No: 162
>PsiHchimera_1
ATGTTCTGCCGTGGCCTGTTGAGTTTAATGGCAATCATCATCGTGTATTTTATAGCCCAAAAGAGGAGAAGAGCCAG
ATTACCACCCGGTCCTCGTGGCCTGCCACTGATCGGAAATTTGCACCAGGCACCGAAAGAGGCGGTGTGGCTAACCT
TCCATAAATGGGTGAAGGAATATGGCAACCTTGTTTCTGTAAATTTTGGGGGCACGGAGATGGTGATACTTAATACG
TTAGAGACGATTACAGATTTGCTAGAGAAGAGGGGATCAATTTACTCTGGCAGGCTGGAATCAACTATGGTCAACGA
ACTGATGGGCTGGGAGTTTGATCTTGGCTTTATTACATATGGAGATAGGTGGCGTGAGGAAAGACGTATGTTTGCAA
AGGAGTTCAGTGAAAAGGGAATTAAACAGTTTCGTCATGCCCAAGTTAAGGCCGCTCACCAATTAGTACAACAGTTA
ACTAAGACACCAGATCGTTGGGCACAGCACATAAGACATCAAATCGCTGCTATGAGCCTTGACATTGGCTATGGCAT
TGACTTAGCTGAGGATGATCCGTGGCTGGAGGCGACGCACTTGGCCAACGAGGGCCTAGCAATTGCCAGCGTTCCAG
GGAAATTCTGGGTTGATTCTTTTCCAAGCCTGAAATATTTACCTGCTTGGTTCCCTGGAGCTGTTTTCAAAAGAAAA
GCGAAGGTTTGGAGGGAAGCGGCCGATCACATGGTCGATATGCCCTACGAGACTATGAGAAAGTTGGCTCCTCAGGG
GCTTACGCGTCCGAGTTACGCTAGTGCAAGATTACAAGCGATGGATCTTAATGGCGATTTAGAGCACCAGGAGCACG
TAATCAAGAACACGGCGGCGGAAGTAAACGTTGGCGGAGGTGATACTACTGTGTCCGCAATGTCCGCCTTCATCTTG
GCAATGGTTAAATACCCTGAAGTCCAAAGAAAAGTTCAAGCAGAATTGGATGCTTTAACTAACAATGGGCAAATCCC
GGATTATGATGAAGAAGACGACAGTTTACCCTATCTAACGGCGTGTATTAAAGAACTGTTTAGGTGGAATCAAATAG
CCCCGCTGGCAATTCCACATAAACTTATGAAAGACGACGTGTATAGAGGCTACCTTATCCCCAAAAATACGCTAGTC
TTCGCAAACACATGGGCTGTTCTAAACGATCCCGAAGTCTATCCAGATCCTAGTGTTTTTAGACCAGAACGTTACTT
AGGACCAGACGGAAAACCAGACAACACAGTACGTGATCCTCGTAAGGCAGCATTTGGCTATGGCAGACGTAATTGCC
CGGGCATACATCTTGCGCAGAGTACCGTGTGGATAGCTGGGGCCACCCTTTTGTCAGCCTTTAATATTGAAAGGCCC
GTCGATCAGAACGGTAAACCCATAGACATACCTGCTGACTTCACGACCGGGTTTTTCCGTCATCCAGTTCCCTTCCA
ATGCCGTTTCGTACCTAGAACGGAACAGGTATCTCAAAGCGTAAGCGGACCT

Seq. ID No: 163
>PsiKchimera_1
ATGAAAACCAAATTTTGTACCGGCGGCGAAGCCGAGCCATCACCATTAGGGTTACTATTGAGCTGTGGTAGTGGCCT
TGTCCCTCGTGGTTCTCCACAACCGCCAGCTGACGAGCAACCCGAGCCGCGTACCAGAAGGAGAGCATATCTATGGT
GCAAGGAGTTCTTGCCGGGGGCATGGCGTGGACTTAGGGAAGACGAGTTTCATATCAGTGTTATCCGTGGCGGTTTA
TCCAATATGTTATTCCAATGTTCCTTGCCCGACACGACTGCTACCTTAGGCGACGAACCAAGAAAGGTGCTTCTGAG
ATTATACGGCGCTATCCTGCAGATGAGGAGTTGTAATAAGGAAGGATCTGAACAGGCTCAGAAAGAGAATGAGTTTC
AGGGAGCGGAAGCGATGGTCTTAGAGAGCGTAATGTTCGCCATCCTGGCAGAGCGTTCTCTAGGTCCAAAGTTGTAT
GGGATCTTTCCACAAGGCCGTCTGGAGCAGTTCATAAAGATGAAAACGTTGTTAGACTATGTTACTGCAAAGCCGCC
TTTAGCGACCGACATTGCTAGATTGGTGGGCACGGAAATAGGTGGTTTCGTGGCACGTCTTCACAACATCGGACGTG
AGAGACGTGATGACCCTGAGTTTAAATTTTTCTCAGGTAATATCGTAGGTAGAACAACGAGTGACCAGCTGTATCAA

ACTATCATTCCTAATGCGGCAAAGTATGGGGTCGATGATCCATTGTTACCTACCGTCGTAAAGGATCTGGTTGACGA
CGTAATGCATTCAGAGGAGACTTTGGTCATGGCGGACTTGTGGAGCGGTAACATTTTACTGCAGTTAGAGGAAGGTA
ACCCCTCCAAGCTGCAGAAAATCTACATTCTGGACTGGGAACTATGTAAATACGGTCCTGCGAGTTTGGACCTTGGT
TACTTCCTTGGGGATTGCTACTTGATATCTAGGTTCCAAGATGAGCAAGTCGGGACTACAATGCGTCAGGCGTACTT
ACAGTCTTATGCTAGGACCAGCAAGCACTCCATCAACTACGCAAAGGTTACGGCAGGCATAGCAGCCCACATCGTTA
TGTGGACTGATTTTATGCAATGGGGGAGTGAGGAGGAAAGGATCAACTTCGTCAAAAAGGGGGTGGCGGCATTCCAC
GACGCTAGAGGGAATAACGACAACGGAGAAATCACTTCGACCCTGCTTAAAGAGTCATCAACAGCG

Seq. ID No: 164
>PsiKchimera_2
ATGGCGTTCGACCTAAAGACCGAAGATGGTCTAATCACATATCTGACAAAGCACCTGTCTTTAGATGTAGATACCAG
CGGAGTTAAACGTCTATCCGGAGGGTTTGTTAACGTCACCTGGAGAATAAAACTTAATGCACCGTACCAAGGACACA
CGAGTATCATCCTAAAACACGCACAACCCCATATGTCAACGGATGAAGACTTTAAGATCGGAGTTGAGCGTAGCGTC
TACGAATATCAAGCGATTAAGTTGATGATGGCAAACCAGGAGGTGTTGGGTGGAGGCGACAGCCGTGTTAGCGTCCC
CGAAGGCTTTCACTACGACGTTGAAAATAACGCATTAATCATGCAGGATGTTGGAACTATGAAAACTCTGTTAGACT
ATGCAACTGCAAAACCGCCACTTTCAACGGAAATCGCGAGTCTAGTGGGGACCGAAATCGGGGCCTTTATCGCGCGT
CTTCACAACCTAGGTCGTAAACGTCGTGACCAACCGGCTTTCAAATTCTTCTCCGGAAATATTGTAGGCCGTACAAC
AGCCGATCAGCTTTATCAAACCATTATTCCAAATGCGGCCAAGTACGGTATTAACGACCCACTACTACCCACTGTTG
TAAAAGATCTTGTCGAAGAGGTTATGAACAGTGAGGAAACCCTGATTATGGCAGATTTATGGAGCGGGAACATTCTA
CTGCAATTAGAGGAAGGGAACCCTTCAGAGTTGAAAAAGATCTGGCTTGTTGACTGGGAATTGTGTAAGTACGGACC
AGCCAGCCTAGACATGGGGTACTTCCTAGGAGACTGCTACCTTATTGCCAGATTTCAGGACGAGTTGGTAGGTACGA
CTATGAGAAAAGCCTATCTAAAATCTTATGCGAGAACCGCTTCAGACACTATAAATTACTCAAAGGTCACGGCAAGC
ATCGGGGCGCATCTAGTTATGTGGACCGATTTTATGAAGTGGGGAAATGACGAAGAAAGGGAGGAA

Seq. ID No: 165
>PsiMchimera_1
ATGGATAGCGCAGGAAACATTTACCGTCACAAAGTAGACTTCACCGCGCTTGCGCTTCAGGACCCTGCCTTCAAGGA
GACGCTATCAGCAAAGGGACGTTTAGACTTCAGCAATCCAGATGCGGTTAGACAACTTACTGTGAGTCTGTTGAGGA
GGGATTTCGGGTTGGAGGTTGAGCTTCCAGACGACAGGCTTTGCCCTCCCGTGCCAAACAGGCTTAACTACATTTTG
TGGTTGCAAGACTTGATAGATTGTACAGGGGATGATTACCACGAGGGCTTTAATGCTGACCGTGATGTGGTGGGACT
GGACATAGGGACGGGGTCTAGCGCCATTTATCCTATGTTGGCTTGCGCAAGATTCAAGGCTTGGAGCATGGTCGGAA
CAGAGGTAGAAAGAAAATGTATAGATACGGCCAGACTAAATGTCGTCGCCAATAACTTGCAAGACAGACTGTCCATT
TTGGAGACTTCGATTGATGGGCCCATCCTAGTACCCATTTTTGAGGCCACTGAAGAGTATGAGTACGAGTTCACCAT
GTGTAACCCGCCCTTTTACGACGGGGCTGCCGATATGCAGACATCTGACGCGGCGAAAGGGTTTGGCTTTGGGGTGG
GGGCTCCTCATTCCGGTACAGTGATTGAGATGAGTACGGAAGGGGGCGAGAGCGCATTTGTAGCACAAATGGTTCGT
GAGAGTCTGAAGCTGAGGACTAGGTGCAGGTGGTACACCTCAAATCTAGGCAAGCTGAAAAGCCTTAAAGAGATCGT
AGGGCTTTTAAAAGAGTTGGAAATATCAAACTATGCTATCAACGAATATGTCCAAGGAAGTACAAGAAGATATGCGG
TGGCTTGGTCTTTTACAGACATTCAACTACCCGAAGAACTTTCAAGGCCATCTAATCCTGAACTTTCAAGCCTTTTT

Seq. ID No: 166
>PsiMchimera_2
ATGAGCGCCACCACCAACATATACAAGGAAGACATTGACTTTATAACTTTAGGGAGGGAAGATAGCGACTTTGGAAA
GTTACTAAACTCAAACGGGCAATTAGATTTCTCTGACCCCAAATCAGTACAACAGTTGACTAAGTCTCTGTTAAAGA
GGGATTTCGGCCTTAAATTGATTCTTCCCGATGACAGGTTATGCCCACCTGTTCCCAACAGGTTAAATTACGTACTA
TGGATAGAAGATATTTTCAACTACACAAACAAAACCCTAGGCCTGAGTGACGACAGACCAATCAAAGGAGTTGATAT
CGGAACAGGTGCTTCTGCTATCTATCCTATGCTGGCCTGTGCCCGTTTCAAGGCTTGGAGTATGGTTGGTACGGAGG
TCGAGAGGAAGTGTATTGATACGGCTAGATTAAACGTTGTCGCAAACAATTTACAAGACAGATTATCCATTCTGGAG
ACTTCCATAGATGGACCTATTCTTGTGCCTATTTTTGAGGCCACCGAGGAATATGAGTATGAGTTCACAATGTGTAA
CCCGCCGTTTTATGACGGCGCGGCGGACATGCAAACGAGCGACGCAGCAAAGGGGTTCGGTTTTGGAGTTGGCGCTC
CACATAGTGGGACCGTTATCGAAATGTCAACAGAAGGCGGCGAGTCCGCTTTCGTGGCTCAAATGGTGAGGGAAAGC
CTAAAGTTGAGGACCCGTTGTCGTTGGTATACTTCTAATCTTGGCAAGTTAAAGTCTTTGAAGGAAATTGTAGGTCT

ATTAAAGGAGCTAGAAATTTCCAACTACGCGATCAATGAGTACGTGCAGGGCTCCACAAGAAGGTATGCCGTCGCTT
GGTCCTTTACGGATATCCAATTACCAGAAGAGTTGTCACGTCCGTCCAACCCAGAGCTGTCAAGCCTATTT

Seq. ID No: 167
>PsiMchimera_3
ATGGCTCAAAACTCCACCATCTATGAAGATGAAGTAGATTTTGCCACACTAGCATTGCAAGATTCCGAGTTCGCGAA
GATTCTGAAATCTAACGGACAGTTGGATTTTAGTAACCCCGAATCCGTGCAGCAGTTGACAAAGTCATTGCTAAAAA
GGGATTTCAAGCTGAAACTTTCTCTGCCTCCGGATAGACTTTGCCCGCCCGTTCCTAACAGGCTTAACTACATTATA
TGGATTCAGAATCTTCTTGACACTACGTCTGATTCATATAACGACAAGTACGATCCCGAGCGTGAGGTTCTTGGATT
AGACATCGGCACGGGCGCCAGTGCAATTTACCCCATGCTTGCTTGCGCTCGTTTCAAAGCGTGGTCAATGGTCGGGA
CCGAAGTTGAGAGAAAGTGCATCGATACTGCCAGATTGAATGTGGTAGCGAACAATTTACAAGATAGGCTAAGCATT
TTGGAGACATCCATTGATGGTCCGATTTTAGTACCCATTTTTGAGGCGACAGAGGAGTACGAGTACGAATTTACGAT
GTGCAATCCGCCATTCTACGACGGGGCAGCAGACATGCAGACCTCTGATGCTGCCAAAGGATTTGGCTTCGGTGTTG
GCGCCCCACACAGTGGTACGGTGATAGAAATGAGCACTGAAGGAGGGGAGAGCGCATTTGTCGCTCAAATGGTCAGG
GAGTCACTAAAACTGAGGACCAGATGTAGATGGTATACGTCAAACTTGGGAAAATTGAAATCTTTGAAAGAAATTGT
GGGTCTGCTTAAAGAATTAGAAATTTCTAACTATGCTATAAACGAATATGTTCAGGGTAGCACGAGAAGGTATGCCG
TGGCCTGGTCATTCACGGATATTCAATTACCCGAAGAGCTTTCACGTCCCTCTAATCCCGAACTGTCCAGCTTGTTC

Seq. ID No: 168
>SAM2
ATGTCTAAATCCAAGACTTTCTTATTTACTAGCGAGAGCGTAGGAGAAGGGCACCCAGATAAAATCTGTGACCAGGT
ATCTGATGCAATTCTGGATGCCTGCCTAGAACAAGACCCCTTCTCTAAGGTAGCCTGCGAAACCGCCGCAAAGACTG
GAATGATTATGGTCTTCGGCGAAATTACTACAAAGGCAAGACTAGATTATCAACAAATCGTTAGAGATACTATCAAA
AAGATCGGGTATGATGACTCCGCGAAAGGATTTGATTATAAGACGTGTAACGTACTTGTAGCAATTGAGCAGCAAAG
TCCAGACATTGCTCAGGGGCTGCATTACGAGAAGTCTCTTGAGGATTTAGGCGCAGGCGATCAGGGCATTATGTTCG
GGTATGCTACCGACGAGACGCCGGAGGGCCTTCCGTTAACGATATTGCTGGCGCATAAGCTAAACATGGCTATGGCC
GACGCACGTCGTGATGGGTCACTACCGTGGCTGAGGCCAGACACCAAAACTCAGGTGACCGTCGAATATGAAGATGA
CAACGGTAGATGGGTTCCGAAGCGTATTGACACGGTGGTTATCTCCGCCCAACACGCAGATGAGATCTCTACTGCGG
ATCTAAGGACTCAATTGCAGAAGGACATTGTCGAAAAGGTAATTCCGAAAGATATGCTAGATGAGAACACTAAGTAT
TTTATCCAGCCAAGCGGGAGATTTGTAATAGGTGGGCCGCAGGGTGATGCCGGCTTAACAGGTAGGAAAATAATAGT
CGATGCCTACGGCGGAGCCTCAAGCGTGGGTGGTGGAGCATTTTCTGGGAAGGACTACTCAAAAGTTGATAGATCTG
CAGCGTATGCAGCGAGGTGGGTAGCCAAGTCCCTAGTGGCGGCCGGATTGTGCAAGAGAGTTCAAGTGCAATTTTCC
TACGCCATTGGTATTGCAGAACCCTTGAGCTTGCATGTGGATACCTACGGCACCGCGACAAAGAGTGATGACGAAAT
AATAGAAATTATTAAGAAGAATTTTGATCTGAGACCTGGCGTTCTGGTCAAGGAGTTAGACCTGGCCAGGCCAATAT
ATCTACCGACCGCGTCATATGGGCACTTCACAAACCAGGAGTATTCATGGGAAAAGCCCAAGAAACTTGAGTTC

Seq. ID No: 169
>SAM3
ATGGACATCCTGAAAAGAGGTAACGAGAGTGACAAGTTTACTAAGATAGAGACCGAGTCAACCACGATTCCGAACGA
CTCCGATAGAAGTGGATCTCTTATTAGACGTATGAAAGATAGTTTCAAGCAGAGCAACCTGCATGTAATCCCGGAGG
ACCTTGAAAACAGCGAACAAACTGAGCAGGAGAAGATCCAATGGAAGTTGGCTAGTCAACCTTACCAGAAAGTGTTA
TCACAACGTCACCTGACCATGATTGCAATAGGCGGCACTTTAGGTACTGGACTGTTTATCGGCCTAGGTTATTCTCT
AGCTTCCGGTCCTGCCGCGCTTCTTATTGGCTTTTTGCTGGTTGGCACCAGCATGTTTTGTGTTGTCCAATCCGCTG
CGGAGTTGTCCTGCCAATTCCCCGTCAGTGGCAGCTATGCGACCCATGTCTCTAGGTTCATTGATGAATCCGTGGGC
TTTACGGTAGCAACCAATTACGCTCTTGCGTGGTTGATATCCTTTCCCTCCGAACTGATCGGGTGTGCCTTAACCAT
TAGCTATTGGAACCAGACCGTAAATCCGGCTGTGTGGGTGGCAATTTTCTACGTATTCATAATGGTGCTGAACCTTT
TCGGAGTTAGGGGTTTTGCAGAGACAGAGTTTGCTCTGTCTATCATAAAAGTAATTGCGATATTCATATTCATCATT
ATTGGGATTGTATTGATAGCAGGTGGTGGTCCTAATAGCACTGGCTATATAGGCGCGAAGTATTGGCACGACCCAGG
TGCGTTCGCTAAACCGGTATTCAAGAATCTATGCAACACATTCGTGAGTGCTGCGTTTAGCTTCGGTGGTTCAGAAC
TGGTATTATTGACCTCAACAGAAAGCAAGAATATCAGCGCAATAAGCCGTGCGGCTAAAGGCACTTTTTGGAGGATA
GCGATATTTTATATCACCACGGTTGTGATTATTGGGTGCTTAGTCCCTTACAACGATCCCAGACTACTTTCTGGGTC

TAATAGCGAAGACGTCTCCGCCTCTCCTTTCGTCATAGCTCTGTCTAACACAGGCTCAATGGGAGCCAAGGTCAGTA
ATTTTATGAATGTCGTGATTTTAGTAGCAGTGGTGTCAGTATGCAATTCATGCGTCTATGCGAGCAGTAGGTTAATT
CAAGCCCTGGGCGCCAGCGGGCAATTGCCTTCTGTGTGTTCTTATATGGATAGGAAAGGGAGACCGCTTGTTGGTAT
CGGCATAAGTGGAGCCTTTGGTTTACTTGGATTTTTGGTGGCGAGCAAAAAGGAAGATGAGGTGTTTACGTGGCTAT
TTGCCTTATGCTCTATCAGCTCCTTCTTTACCTGGTTCTGCATCTGCATGTCCCAGATCAGGTTCAGAATGGCGTTA
AAAGCTCAAGGGAGGAGTAACGATGAAATAGCCTATAAAAGTATACTAGGTGTATATGGGGGTATCTTGGGGTGCGT
ATTGAACGCTCTTCTAATTGCGGGGGAAATCTATGTGTCTGCGGCACCCGTCGGATCTCCCTCTTCTGCTGAGGCAT
TTTTCGAATATTGCCTGAGTATTCCCATAATGATCGTAGTTTATTTTGCCCATCGTTTTTATAGGAGGGATTGGAAA
CACTTCTATATCAAGAGGTCAGAAATCGACCTTGACACGGGGTGTAGCGTCGAAAATCTTGAGTTATTCAAAGCCCA
GAAAGAAGCTGAGGAACAGCTAATTGCATCTAAGCCGTTCTACTACAAAATTTACAGGTTCTGGTGT

Seq. ID No: 170
>SSO2
ATGTCTAACGCGAATCCTTACGAGAACAATAATCCATACGCTGAAAATTACGAAATGCAGGAAGACCTGAACAATGC
CCCAACGGGGCACAGCGACGGCTCAGATGACTTCGTCGCATTTATGAATAAGATTAACTCAATCAATGCTAATCTAT
CAAGATACGAGAATATTATTAACCAAATTGACGCGCAACATAAAGACTTACTTACCCAGGTCAGCGAAGAGCAGGAA
ATGGAGCTGCGTCGTAGTCTCGACGACTACATAAGCCAAGCAACAGATTTGCAGTACCAACTAAAGGCAGACATAAA
GGACGCGCAAAGAGACGGACTACACGACTCAAATAAACAGGCTCAGGCGGAGAACTGTAGGCAAAAATTTCTGAAGT
TAATTCAAGACTACAGAATTATTGATTCTAACTATAAGGAGGAATCAAAAGAGCAAGCAAAGAGACAGTACACGATT
ATCCAACCAGAAGCAACAGATGAAGAGGTTGAAGCAGCCATCAACGATGTAAACGGGCAGCAAATATTCTCTCAAGC
ATTGCTAAATGCTAACAGGCGTGGCGAAGCCAAGACCGCCTTAGCTGAAGTCCAAGCTCGTCATCAAGAGTTATTAA
AGTTGGAGAAAACTATGGCTGAATTGACACAGTTATTTAACGACATGGAAGAGTTGGTCATCGAACAGCAGGAGAAT
GTGGATGTCATTGACAAGAACGTCGAGGATGCCCAGCAAGATGTGGAGCAGGGGGTCGGGCATACAAATAAAGCTGT
AAAATCTGCGAGGAAGGCAAGAAAGAACAAAATCCGTTGTCTAATAATTTGCTTTATCATATTTGCTATAGTCGTAG
TAGTTGTTGTTGTCCCCTCTGTTGTCGAGACGAGGAAA

Seq. ID No: 171
>T4H-CPR_1
TCTAGCTCTAGTGACGTGTTCGTACTGGGGTTAGGAGTCGTACTAGCAGCTCTTTACATATTTCGTGATCAGCTTTT
CGCTGCCTCCAAGCCAAAGGTGGCACCGGTGAGCACAACGAAACCGGCTAATGGGAGCGCAAATCCTAGAGATTTTA
TCGCCAAGATGAAACAGGGAAAGAAACGTATTGTTATTTTCTACGGCTCTCAGACAGGCACCGCGGAAGAATATGCA
ATACGTTTAGCAAAGGAGGCAAAACAAAAATTTGGACTAGCATCCCTAGTATGTGACCCAGAAGAATATGACTTCGA
GAAGCTAGATCAGTTGCCTGAGGACTCTATAGCATTCTTCGTAGTAGCCACATATGGAGAGGGTGAACCGACAGATA
ATGCGGTTCAACTATTACAAAACTTACAGGACGAATCATTCGAGTTCTCCTCAGGGGAGCGTAAGTTGTCAGGACTT
AAATATGTTGTCTTCGGCTTGGGCAATAAGACCTACGAGCATTATAATCTAATAGGGCGTACTGTAGACGCGCAGCT
TGCTAAAATGGGAGCGATCAGAATCGGAGAAAGGGGTGAAGGAGACGACGACAAATCTATGGAGGAGGACTACTTGG
AATGGAAGGATGGAATGTGGGAGGCGTTCGCGACAGCGATGGGTGTCGAGGAGGGGCAGGGAGGTGATAGCGCCGAC
TTCGTGGTAAGCGAATTAGAGTCACATCCACCTGAAAAGGTCTATCAAGGCGAGTTCTCTGCGAGGGCTCTAACTAA
GACTAAAGGCATTCATGACGCGAAGAACCCTTTCGCGGCTCCCATCGCAGTAGCGCGTGAACTGTTCCAGAGCGTAG
TAGACAGGAACTGCGTGCATGTGGAATTTAACATAGAAGGGAGTGGAATAACCTACCAGCACGGCGATCATGTCGGC
TTGTGGCCGTTGAATCCTGACGTAGAGGTGGAGCGTTTATTATGCGTTCTGGGTTTAGCGGAGAAAAGAGACGCAGT
GATCTCCATTGAGTCCCTGGACCCAGCCCTTGCGAAAGTGCCATTCCCTGTTCCAACGACCTATGGAGCGGTCCTTA
GACATTACATTGACATATCTGCCGTCGCCGGAAGGCAGATTCTTGGAACTCTATCAAAATTCGCACCAACTCCAGAG
GCTGAGGCGTTTTTAAGGAACTTAAATACGAACAAGGAGGAATACCACAATGTCGTCGCGAACGGGTGTTTAAAACT
GGGCGAAATCCTGCAGATAGCGACTGGTAATGACATCACAGTACCGCCTACGACCGCCAATACCACTAAGTGGCCGA
TACCCTTTGATATAATCGTATCCGCTATTCCCAGACTACAGCCTCGTTACTACAGTATCAGTTCTAGTCCAAAAATC
CATCCGAACACAATACATGCGACGGTCGTGGTACTGAAGTATGAAAATGTTCCCACTGAACCTATCCCAAGAAAGTG
GGTGTATGGTGTAGGGTCAAATTTCTTACTTAACCTGAAGTATGCAGTAAATAAGGAACCGGTTCCATACATAACGC
AAAACGGGGAACAGAGGGTCGGAGTTCCCGAGTACCTGATCGCAGGCCCTAGAGGCTCCTATAAGACTGAGAGCTTT
TACAAAGCTCCAATCCACGTGCGTAGGTCAACATTTAGGTTGCCAACAAACCCCAAGAGTCCCGTAATCATGATCGG
GCCAGGGACCGGAGTAGCACCCTTTCGTGGATTTGTGCAGGAGCGTGTTGCATTGGCAAGAAGGTCCATAGAGAAAA

ACGGCCCAGACTCTTTGGCGGATTGGGGCAGAATTTCACTTTTCTATGGATGTCGTCGTAGTGACGAGGACTTCTTA
TACAAGGACGAATGGCCTCAGTATGAAGCCGAACTAAAGGGAAAGTTTAAATTACACTGTGCTTTTTCCCGTCAAAA
TTACAAGCCTGACGGGTCAAAAATCTATGTGCAGGATCTGATATGGGAAGACAGAGAACACATTGCAGATGCAATAC
TAAATGGCAAAGGATATGTGTATATTTGCGGGGAGGCAAAGTCAATGTCTAAGCAGGTCGAGGAAGTCCTTGCGAAG
ATTCTTGGAGAAGCTAAGGGTGGTTCAGGTCCCGTAGAAGGCGTTGCTGAGGTAAAGCTATTGAAGGAGAGGTCTAG
GCTTATGCTGGATGTATGGTCA

Seq. ID No: 172
>T4H-CPR_2
TCATCTAGCTCTGACGTTCTGATTTTGGGCCTTGGCGTTGCGCTGGCTGCTCTTTACCTGTTCAGGGATCAGCTGTT
CGCTGCTTCTAAGCCAAAGGCTATACCGCTTACGAACAAACTGGCCGGTCTTGACAACGAAGGGAACCCGCGTGATT
TCATAGCAAAGATGAAGGCAGGAAAGAAGCGTTTGGTCATATTTTACGGCTCTCAAACAGGAACCGCTGAAGAGTAC
GCGATCAGGCTGGCCAAAGAAGCCAAATCTAAGTTCGGTTTGACCTCCCTTGTGTGCGATCCTGAAGAGTACGATTT
CGAGAATCTGGATCAGTTGCCGGAAGAGTGCGCCGTATTTTTCGTGATGGCTACGTACGGGGAAGGTGAGCCGACGG
ATAATGCGGTTCAGCTTATGCAGAATCTAGCCGACGAATCCTTCGAGTTCTCAGGCGGAGAAAGGAAATTGGAGGGT
TTGAAATATGTAATATTCGCTTTAGGAAACAAAACTTATGAGCACTACAATTTGATTGGTCGTAAAGTTGACACTCT
TCTGACTGACATGGGTGGGGTCCGTTGCGGGGAACTAGGGGAAGGCGACGATGACAAGAGTATGGAAGAGGATTATT
TGGAGTGGAAAGACGCGATGTGGGAGGACTTCGCGAGGAAGATGGGCGTCGAAGAGGGCCAGGGTGGTGATTCTGCT
GATTTTGCAGTCAGCGAATTAGACACTCACGTGCCCGAGAAGGTGTATTTAGGAGAGTTGAGTGCTAGGGCATTGAC
GAAGACGAAAGGGATACATGATGCCAAGAATCCGTATCCGGCGCCCATCGTAGCAAGTCGTGAATTGTTCCAGCAAG
GTGGCGACAGGAATTGTGTTCATGTTGAACTATCCATTGAAGGCTCAGGAATTACCTATCAGCACGGCGATCACGTC
GGAGTATGGCCCACGAATCCTGAGGTGGAAGTCAATAGGCTGCTTTGCGCCCTGGGTTTATGGGAGAAGAAGGACCA
GGTTATAGGGATTGAATCCCTTGACCCAGCCTTGGCAAAAGTCCCATTTCCTGTCCCAACTACGTACGCCACTGTGC
TGAGGAATTACATTGACATCAGCGCAGTTACGGGGCGTCAAATCCTAGGCCACTTGAGTAAATACGCGCCAGCTCCC
GACGTCGAGGAATTTCTGAAGGGTTTGTCTACAAACAAGGAACAATACGGTGCTACAGTAGCCAATGGTTGCTTGAA
GTTAGGCGAGGTCCTACAACTTGCCGCGGGCAATGATCTGAAGGCTATACCAACGACTGAGAATACTACAGCGTGGA
GTATCCCCTTCGACGTTATCGTGTCCGCCATTCCCCGTCTTCAACCACGTTATTACTCAATTTCATCATCCCCAAAG
CTAAACCCCACCTCAATTCACGTAACCGCTGTTGTGTTGAAGTATCAATCAGTCGCGAGCGAAAAACTGCCTGCTAA
GTGGGTCTACGGCGTAGGTAGCAACTTCTTACTAAATTTAAAATACGCTGCTAATGGTGAACCAGCGCCATTCGTTA
CTACCAATGGCTCAGCCGATCCTGCATCAGTTTATTATCCCACTTACGCAATCGAGGGCCCACGTGGCGCGTATAAA
CAGGAAACTATATACAAATCCCCTATACACGTGAGAAGAAGTACATTCCGTCTGCCCACGAACCCTAAATCACCGGT
GATTATGATCGGTCCGGGTACCGGTGTGGCGCCTTTCAGGGGGTTCGTACAAGAGAGAGTTGCCTTAGCAAGGAGAA
CGATTGAGAAAAATGGAGCCGACGCGTTGGCTGACTGGGGCCGTATATCTCTTTTCTATGGGTGTCGTAAATCAACA
GAGGATTTCTTATATAAAGAAGAGTGGCCGCAGTATACAGAGGAGTTAAAAGGGAAATTTAACATGCATAGTGCCTT
TAGTAGAGAAGCACCTTACAAGGCTGACGGAAGCAAGATTTATGTCCAGGACCTGATCTGGGAAGATCGTGCCAATG
TGAGCGATGCGATACTGAATGGCAAGGGATATATCTACATATGCGGTGATGCGAAATCAATGGCGAAACAGGTGGAA
GACACGCTTGCTAAGATACTAGGAGAAGCGAAGGGTGGTACCGCGGAAGTCGAGGGAGCGGCGGAAATGAAGCTGTT
AAAAGAAAGGTCTAGGCTAATGTTAGATGTTTGGTCA

Seq. ID No: 173
>T4H-CPR_3
TCTAGTTCTTCCTCAGGCGCCGGGGCAGATAGCGACGAAAACCCTCGTGATTTCATTGCCAAAATGAAAGCAGGAAA
GAAGAGGTTGGTCATCTTCTACGGCAGCCAAACAGGCACTGCCGAAGAGTATGCGATCCGTTTGGCGAAAGAGGCAA
AATCCAAGTTTGGACTAACGTCTCTAGTGTGCGACCCAGAAGAATATGATTTTGAAAACCTTGACCAGTTGCCTGAG
GATTGTGCCGTGTTCTTTGTAATGGCCACTTATGGAGAAGGCGAGCCTACCGACAATGCGGTGCAGCTTATGCAGAA
CTTGCAAGACGAGAGCTTTGAGTTCTCTAATGGCGAGAGAAAGTTGGAAGGGTTGAAATACGTAGTATTCGCATTGG
GTAACAAAACCTACGAGCATTATAACTTAATAGGAAGAAAAGTGGACACAATATTGGGGGAAATGGGGGCGGTGAGG
TGCGGAGAGAGAGGCGAAGGAGATGACGACAAATCAATGGAGGAAGATTATTTAGAGTGGAAGGACGCTATGTGGGA
AGACTTTGCTCGTAAGATGGGGGTTGAAGAGGGACAGGGCGGAGATAGCGCCGATTTTGCTGTTTCAGAGCTTGAGT
CCCATGCTCCAGAGAAAGTATACCTAGGTGAATTGTCTGCGCGTGCCTTGACCAAGACTAAAGGCATTCACGACGCT
AAAAATCCATACCCCGCCCCTATCGTCGAGAGTAGAGAACTATTCCAAGTTGGTGGGGACAGGAATTGTGTACATGT

AGAGCTAGGCATTGAGGGATCAGGGATTACCTACCAGCACGGAGACCATGTCGGTGTGTGGCCTACAAATCCTGAAG
TTGAGGTGACACGTCTATTATGCGCGCTGGGGTTGTGGGAGAAAAAGGATCAAGTTATTGGAATAGAGAGCTTAGAC
CCTGCGTTGGCAAAGGTGCCATTCCCCGTACCTACAACATACATAACTGTGCTACGTAATTATATAGACATTTCTGC
GGTCACTGGGCGTCAAATCTTAGGGCATCTAAGTAAGTTTGCGCCGAGTCCAGATGCAGAGGCGTTTCTGAAATCCT
TATCTACTAACAAGGAACAGTATGGCGCCATTGTAGCGAACGGATGTCTAAAACTTGGAGAAGTTCTGCAACTAGCC
GCTGGGAATGATTTGAAAGCCGTTCCGAACGCTGAAAACACAACGAAATGGACAATCCCGTTTGACGTTATAGTAAG
TGCTATTCCAAGATTACAGCCTAGATACTATAGTATAAGTTCTTCTCCGAAACTTAATCCTACGACAATACACGTCA
CCGCCGTAGTCCTAAAGTACGAATCAGTGGCCTCTGAGAAAGTTCCGGCAAAGTGGGTATATGGCGTAGGGTCAAAC
TTTCTATTGAACCTGAAGTATGCCGCGAATGGCGACGCAGCACCATTTGTCACAGCCAACGGTTCAGCCGATCCGGC
CTCCGTTTATGCGCCGACCTATGCCATAGAAGGACCCAGAGGGGCTTATAAGCAGGAAACCATTTACAAGTCTCCTA
TTCATGTAAGGAGAAGTACCTTTAGATTGCCGACAAACCCTAAGTCCCCGGTGATCATGATAGGTCCCGGCACTGGC
GTGGCACCATTCCGTGGCTTTGTACAGGAAAGGGTTGCACTAGCCAGGCGTACGATTGAGAAGAATGGTCCAGATGC
GCTTGCAGACTGGGGCAGAATAACATTATTTTACGGCTGTAGAAAGTCTACCGAAGATTTCCTATATAAAGATGAAT
GGCCTCAATACACCGAAGAATTGAAAGGGAAATTTACCATGCATAGTGCGTTCTCTAGGGAACCCCCGTACAAGGCG
GACGGGAGTAAAATTTATGTACAGGACCTTATCTGGGAAGATCGTGAGAAAGTCGCGGATGCTATACTTAATGGCAA
GGGGTATGTTTATATATGCGGAGACGCAAAATCAATGGCAAAGCAGGTGGAAGACACCCTAGCCAAGATACTTGGAG
AGAGTAAGGGCGGTTCAGCAGAGGTAGAAGGTGCAGCAGAAATGAAACTACTGAAGGAGAGGAGTAGACTAATGTTA
GACGTCTGGTCT

Seq. ID No: 174
>T4H-CPR_4
TCATCAAGTAGCAGCTCCAAGTTGAGCGATGGCGACGAGAATCCTAGAGACTTCATCGCAAAAATGAAGAATGGGAA
GAAACGTCTTGTCATATTCTACGGGTCCCAGACGGGGACTGCAGAAGAATATGCTATAAGACTAGCCAAAGAAGCCA
AAAGCAAATTTGGTCTTACCTCCTTAGTATGTGATCCCGAGGAGTATGACTTCGAGAATCTGGACCAGTTACCGGAC
GACTGCGCAGCGTTTTTCGTCGTAGCAACGTATGGGGAAGGTGAACCTACCGACAATGCTGTACAACTAATGCAGAA
CCTACAAGACGAATCATTCGAATTTTCTGGCGGTGAAAGAAAGTTGGAAGGATTGAAGTACGTCGTATTTGCCCTGG
GCAATAAGACTTATGAACACTATAATGTTATTGGGCGTATAGTAGACACGGAGCTTGCAAAAATGGGCGCTATACGT
TGTGGGGAGAGGGGTGAGGGTGATGACGACAAGTCTATGGAAGAGGATTACTTAGAGTGGAAAGATGGAATGTGGGA
GGAATTTGCCCGTATCATGGGAGTTGAAGAGGGTCAGGGAGGTGACACCCCAGACTTCAAAGTGACAGAGCTGCAAT
CACATCCTAGTGAGAAGGTATATTTGGGCGAGCTTTCCGCGAGAGCACTGACAAAGACTAAGGGGATACACGACGCA
AAGAATCCATACCCGGCCCCCATCCTAAAATCTAGGGAATTGTTCCAGAAACAGGGGGAGAGGAACTGCGTCCACTT
GGAGCTGGGTATTGACGGGAGTGGAATCACATATCAACACGGAGATCACGTCGGCGTGTGGCCAAGCAATCCAGAGG
TTGAGGTTAACCGTCTACTATGCGCGCTGGGCCTTTGGGACAAACGTGACCATGTCATCGGTATCGAAAGTCTTGAT
CCAGCCCTTGCTAAGGTACCATTCCCCGTTCCCACAACTTACTCCACTGTTTTAAGAAACTATATCGACATATCAGC
CGTTGCAGGAAGACAAATTTTGGGAAATCTGGCTCGTTTTTCACCCTCCCCAGATGCTGAAGGGTTTATGAGAAGTT
TGAATACGGATAAAGAGCAATACGGTAGGATCATAGCAAACGGATGTCTGAAACTAGGGGAAGTTTTGCAACTTGCG
GCGGGAAACGACATAAAAGCAGTACCAACGTTAGAAAATACTACTGCGTGGCCGATCCCATTCGATGTAATCGTTAG
CGCCATTCCTCGTTTACAACCTAGATATTTCAGCATAAGTAGTTCTCCCAAGCTACATCCTACAGCAATCCATGTTA
CTGCCGTAGTGCTAAAGTATCAGAGTGTCGCTTCAGATAAGGTGCCACCGAAATGGGTGTACGGTGTTGGTTCAAAC
TTTATCCTTAACCTAAAATACGCCGCTTGCGGCGAAACAGCTCCGCTTATTGCACAGAATGGAAGCGCCGATCCTGC
TCACACACCCTTTCCCCTATACGCTATAGAAGGTCCACGTGGGGCATACAAGCAGGAAATGATCTATAAAAGCCCAA
TCCATGTGAGAAGGAGTACATTTAGACTACCTACTAACCCTAAATCCCCTGTAATCATGGTTGGACCAGGTACCGGA
GTCGCGCCGTTTCGTGGATTCGTTCAGGAACGTATCGCTCTAGCCCGTCGTACAATCGAGAAGAATGGTCCGGATGC
ACTAGCTGATTGGGGTAGGATATCTTTATTTTATGGGTGCAGAAAAAGCAACGAGGACTTCCTATATAATGAAGAGT
GGCCACAATACATCGACGAATTAAAAGGGAAGTTTACATTGCATACAGCTTTTTCAAGGGAGCCCCCGTATAAGCCG
GATGGAAGTAAGATTTATGTGCAGGATCTATTATGGGATGACAGATCCAAGGTCGCAGACGCTATCATTAATGGAAA
GGGCTACATCTATATATGCGGTGATGCAAAGTCGATGGCAAAGTCAGTGGAGGACGTGCTGGCGAAGATATTAGGGG
AGGCTAAGGGCGGCACTATGGAAGTCGAAGGTGCGGCAGAACTTAAATTATTAAAAGAAAGAAGTCGTCTTATGTTG
GATGTTTGGTCT

Seq. ID No: 175

>T4H_1
ATGAAGACTAGGACTTCCAAACATCCTCCAGGCCCACGTGGGCTGCCCCTGATTGGTAATCTACTAGACATGCCCGC
ATCATACGAGTGGTTACAATATAGGAAGTGGTCTGAGGAATTTAAGTCCGACATCATTTACCTGAATATCTTAGGCA
CACAAATTGTAGTGACTAACACCCTTGAATCCACACTGGACTTACTTGAAAAGAGATCCTCTAAGTACTCCGGGAGG
CATTCCTTTCAACTGCCGAATAACTGCGCTATGGGGTGGGCGTGGAATCTGGCTCTAATGTCTTATGGGGACGAGTG
GAGGGCCCACCGTAGGCTTGCCGCTAGAGGTTTCGATGCACAGGCGATGCCGAAATTTAACCACGCATTTACGCGTA
ATACTCGTGGCCTGCTTAGGCGTCTTTTAGAGTCACCAGAGGCCTGGAATGAGCATGTAAGGCACGAGGTCGGGTCA
ATGATAATTGAAATTACCTACGGGTTGGACGTGCTTTCTAAGAATGATCCCTTTATTGAGTCCGCCGACAAAGGACT
AGCGACATTGGCCTTAGCAGTAGTACCGGGTGCCTTTCTGGTTGACACTCTACCAATTCTTAAACACATACCTTCAT
GGTTTCCAGGGGCTGGCTTCAAACGTAAGGCTAAGGAGTGGAAAAGATACGCCGATGAGGTTTTAGAAGCTCCTTAT
AAGGCTCTAAAGGAAGAGATGGCAAGTGGAGCCGCGAAGCCTTCTTTTGTTCAAAGGTGTCTGCAGGACATGGACCC
AAATATCGATACAACGAACCAAGAACGTGTCATAAAGAACACAGCGGCAGAGATGTACGTGGCGGGCGCCGACACTT
CTGCCAGCTTTATTGCAACGTTCGTGCTTGCTATGATACAGTACCCTCAGGTACAGCGTAGAGCACAAGCAGAGTTG
GACTCTGTCTTAGGCCCAGATAGACTGCCAACCTTCGGCGATATGCCTAGCCTTCCGTATTTGTCCGCGATAACTAA
AGAGTGTTTCCGTTGGGAAGTAATCACCCCCATAAGTATTCCACACATGCTTACAGAGGACGACGAGTATCGTGGTT
GGTTCCTACCGTCAGGAACGGTTGTAATTCCAAATAGCTGGGCTATAATGAATGACCCGACAGTATACCCAGACCCC
TCCGTGTTCAATCCTGAAAGATTCCTGAAGGATGGTAAAATAGACCTTGAAGTGCAGGACCCGCAGTTAGCGGCTTT
TGGTTATGGGCGTAGAATTTGTCCCGGCATGAGGGTCGCCAATGCTTTTACGTGGTTAAGTGCTGGATCAATCTTGG
CCTCATTCAACATTTCAAAACCGGCAGCGAAAGACGGGACCCCCATCGAGCTTGATGTTAAGTATCGTAGCTCCAGT
ATTCGTCATCCAGAGGCCTTTGACTGTCTGTTTAAGCCCCGTTCAGAGAATACTAGAGACATGATCGTAAGCGCAGC
CGCG

Seq. ID No: 176
>T4H_2
ATGAGTAAGAGGTCAAAACACCCACCTGGCCCCAGAGGCTTGCCGTTGATAGGGAACTTACTGGATATGCCTACGAA
TGATGAGTGGTTACAATATCGTAAATGGAGTCAAGAGTTCAAGTCCGACATAATATACCTTAATGTTTGCGGTACTC
AAATCGTGGTTACGAATACCCTTGAGTCTACGCTTGATTTGCTAGAAAGAAGATCTAGTAAGTATAGCGGCAGAATG
GGTCTTGAATGGGCTTTCATCTTGATGCCTTATGGAGATGAGTGGAGAGCGCACAGGAGACTAGCAGCTAAAGGTTT
CGACGCAAAGGCTATCCCAAAGTTTAATCCGACCTTCACGAGAAACGCGCAGGATCTACTGAGAAGGCTTTTAGAAT
CTCCAGAGGCTTGGCATGAGCACGTTCGTCATCAAGTCGGAGCGATGATAATAGAAGTATCTTACGGATTAGATGTA
TTGCACAAAAACGATCCCTTTATAGAGAGCGCGGATAAAGCAGCCGTGACGTTCGCAATGGCCATCAAACCCGGAGC
TTTCCTTGTCAATACCGTGCCCATTTTAAAATACGTTCCTAGTTGGTTCCCCGGTGCAGGATTTCAGAGAAAAGCTA
AGGAGTGGAAAAGGTATAACGACGCAGTGCTAGAAGCACCATTTAAGGCTTTGAAAGAAGAGATTACCAATGGGGCG
GCAAGACCATCTTTTGCCCAACAATGCCTGCAGAATATGGACCCGAACATTGACACGGCATATCAGGAACGTGTGAT
AAAGGACACAAGTGCTGCGATGTATGGGGGAGGATCAGACACGTCAGTTTCATTTTTGGCTACATTCGTCCTGGCTA
TGTTACAGTACCCCTCTGTCCAACGTCGGGCTCAGGTTGAGCTTGACTCCGTGCTAGGACGTGATCGTTTGCCAACA
TTCGACGACATGCCGGACCTTCCCTATCTAGCAGCCGTTATGAAAGAGTGCCACAGATGGGAAATAGTTTTGCCGCT
TGCGATACCTCACATGCTTACCGCTGATGACGAGTACAGAGGATGGTTCCTTTTATCCGGCACGTTAGTTATCCCAA
ATAGCTGGGCAATACTAAACGATCCTACGGTTTATCCCGATCCGTCTACTTTCAATCCTGAACGTTTCTTAAAAGAC
GGGAAGATTGACCCCAATGTACAAGATCCCGAGCTTGCGGCTTTTGGATACGGCAGAAGGACTTGTCCCGGCCGTAG
GATAACCAACGCTTTTACCTGGCTTTCCGCGGGATATATACTAGCCAGCTTCAATATAGAAAATGCAGTTGGTAATG
ATGGTATGCCCATCGAGCCCAAAGTTAAGTACCGTTCTGAAACAATCAGACACCCAGATACTTTTGAGTGCGTCTTC
ACCCCAAGGTCAGATGATACCAGAGATATGATCGGTAGCGCGTACACA

Seq. ID No: 177
>T4H_3
ATGGGCAGGTGGCCCATAATTGGGAATCTATTAGACATGCCGCAGAAATCCCCTTGGCTTACTTATGCCAAATGGAG
TGAAGATTGCGACTCTGACATAATACACTTGAATGTCCTGGGTACGAGTATAGTCGTTTTATCTTCTTTAGAGGCAA
TAAGTACGTTACTAGAGGGTAAGGCGGTCGATTTTAGTGATAGGCCTAAATCTACGATGATGTCCGAGCTTATGGGG
TGGGAAAGGGGCTTCGCGTTTATGCCCTATGGCCAGTTATGGAGATCCCACAGGAAAGCATTCCATCAGGAGTTCAG
CCCCCAGGTGGCGCACAGAAATCACCCGAAGTTAATTAAGGCGACACATAACTTACTGCGTCTGCTTCTGAACACCC

CACAGCACTGGCATGGGCATATTCGTAGACAAGCCGGGGCATCAATTATGGATATAGCATACGGTATTGAAGTACTA
CCTGAAAACGATCCGTACTTAGATATAGCAGAGGCAGCCGTGAAGGCATTCAACGATGCCAGCGTGCCTGGCGCATT
TCTAGTCGATTCTATACCCCTACTTAAACATGTTCCGGCATGGGTCCCCGGGGCGGGCTTTCAGCTTAAAGCAAAGG
AGGGTAGACAGGCACTAGAAAACCTAATTGACTCACCTTACAACGCGATGAAAAAGGATCTTGCTGGCGGAAAAGCT
AAGAGCAGCTACACGTCTAGGTCCTTGGCAGCAATGGATGCGACTGGTGTGATTGAAGAAAATGAGACGATAATAAG
GGAGACAGCGGCAATGGTGTATTTAGGGGGAAGCGACTCCACCCCGAGCACGACATCCGTGTTCATACTGGCCATGT
TAGCCCATCCCGAGGTGCAGAGAAAAGCTCATGCGGAGCTGGACAGCGTGATCGGGAAAGCACAGCTTCCTACCTTT
AAAGACCGTGGCTCACTACCATATGTTACGGCTGTGGCAAAAGAAGTCCTGAGGTGGGAACCTGTAGCGCCGCTGGC
GGTTCCTCGTAAGGTTCGTGTGGACTCCGAGTACAAGGGGTACAGGATACCCAAGGGCTCTATAGTCTTTCAAAACT
CTTGGGCCCTTCTGCATGACGAGAAGACTTACCCTAACCCGTTAGCGTTCAACCCCGAGAGATTCCTGAAAGACGGT
CAATTAGACCCTAACGTCCAGGACCCAGATGTTGTGGCCTTCGGGTATGGAAGGCGTTCTTGTCCGGGTAAAACGAT
GGGCTATGACTCCGTATGGCTGAACGTGGCAAGTATTCTAGCCGCCTTCGACATTAAGAAGGTTGCGAACCCGGATT
CCACGAACGTGGAACCAAAGTTTGAACCGTTTGGAATAACCGTT

Seq. ID No: 178
>T4H_4
ATGTATCTATTTAAGGCATACCTACGTCCTTCTCGTAGGTTACCCCCAGGGCCAAGGGGGTGGCCCCTAATAGGAAA
CTTGCTAGACATGCCGACCTCCGACGAATGGGTAAGATACGCCCAATGGGTCCGTGAGTTTAAGTCTGACGTTATCC
ACTTAGAAGTGTGTGGGACGCACATCGTTATTCTTAATAGTGTAGAAAGCGCAGTTGATCTTTTAGAGAAGCGTTCA
TCCCTGTACTCATCTCGTCCGCCAACGCCTATGATGAGTGACCTGATGGGGTGGTCTTGGAATACAGCTATGCTTCC
GTACAACGACGAGTGGAGAGCGCAAAGGCGTCATTTCCACGGTGAGTTCGACGGTAGAGCAATAGGAAAGCATTATC
CTCCGATCATTAGAAGCACGCACGATTTGTTGCAAAGATTGTTGGACACCCCAGAGCAATGGCAGAGTCATATAAGG
CACTTGGTCGGTGCAACGATATTGGATGTCGCATACGGAATCGAAGTTCTTCCAGCCGACGACCCATATGTACGTAC
AGCCGAAGCTGCCTTTGCCTCCGTATCTGAGGCTATGGTGCCAGGCGCTTTCCTTGTCGATGTACTGCCAATTTTGA
AGCACATGCCGAGTTGGATGCCAGGGGCAGGCTTCAAAAGAAAGGCAGTGGCGTGGAAGAAATTGGCAGACGCAGTT
TTTGATGCACCCTTCGCGGCGATGAAACAAGCAATGGCAGCGGGTACTGCGAAGTCTTCATTTGGCAGTAGAAGTCT
GAGGGATATAGACATTAAGGGAAACGTACAGAGCCAAGAGTTCTCTATCCAGGCCGCGGCTGGCACCATGTATAATG
CCGGGTCAGACACGACAGTTGCGCTTTTAGAAACGTTCATGCTTGCTATGGTCCTACATCCCGAAGTCCAAACCAAG
GCGCAAGCCGAAATGGACTTAGTCCTAGGGAGAAGCAACTTGCCCACATTCGCTGACCAGGAGAGTTTGAGCTACTT
AGCAGCTGTAATGCAAGAAGTTTTTAGGTGGCAAGTCGTCGCGCCCTTCGGCGTGCCGCACATGTCAACCGCAGACG
ATGAATACCGTGGTTATTTCATACCGGAAGGTACCATTGTAATTCCGAACGCCCATCAGATGCTGAACGATGAAGAT
GTTTATCCAGAGCCGTCCAAATTCAAACCCGAAAGATTTTTAAAGGATGGCAAGCTAGACTTATCAGTGCGTAGCCC
CCTAATTGCAGCATTCGGCTTCGGTAGAAGGATTTGTCCTGGCAGAGCCTTAGGAGAGAACTCCGCCTGGTTGGCAG
CTGGGTCCATCCTGACTATGTTCAATTTGTCTAAAGCGACGGACCACAATGGTGTCACGATAGAACCCTCCGGCAGA
TATACATCAGGTCTGGTTAGACACCCCGAGACGTTTAAATGTCAAATTACTCCCAGATCCAATGAGCCGAGAAGAGA
ACTGGCAGGGGAGATTGAATTGATCACGGGCAGGATTCAAGAGTCTGAAGAAGCG

Seq. ID No: 179
>T4H_CPR_chimera_1
TCCAGTAGCTCAAGCGGTTCAGTGGCCTACTTTACGAAAGGCACTTATTGGGCAGTCCCCAAGGACCCATATGCATC
CTCCTATGGCGCAGCTAATGGGGCGAAGGCGGGAAAGACAAGAGACATCATAGAGAAAATGGAGGAGACCGGGAAAA
ACTGTGTAATCTTCTACGGCTCTCAAACCGGGACAGCTGAAGATTACGCTAGTAGATTGGCAAAAGAAGGGAGTCAA
AGGTTCGGACTAAAGACTATGGTCGCGGATTTAGAAGAATATGACTACGAGAATCTGGACAAATGGCCAGAGGACAA
AGTAGCATTCTTTGTACTTGCTACATACGGCGAGGGGGAGCCTACGGATAACGCTGTACAATTACTGCAAAATTTGC
AGGATGAGAGCTTTGAGTTTTCTTCCGGTGAACGTAAGTTAAGTGGGCTGAAATACGTTGTTTTTGGCCTTGGTAAT
AAAACATATGAGCACTACAACTTGATCGGCCGTACAGTTGATGCTCAACTAGCAAAGATGGGGGCGATCAGAATAGG
AGAAAGAGGGGAGGGCGATGACGATAAGTCTATGGAGGAAGATTATCTTGAGTGGAAGGACGGTATGTGGGAAGCGT
TTGCCACGGCGATGGGAGTTGAAGAAGGCCAAGGAGGTGATTCCGCCGACTTTGTTGTTAGCGAGCTGGAATCACAT
CCACCTGAGAAAGTCTATCAAGGCGAATTTAGCGCGAGAGCCTTAACTAAAACAAAGGGGATTCACGATGCGAAAAA
TCCGTTTGCTGCACCGATCGCAGTTGCACGTGAGTTGTTTCAAAGCGTAGTGGACCGTAATTGTGTCCACGTAGAGT
TCAATATCGAGGGAAGCGGGATCACGTATCAACACGGGGACCATGTCGGTCTATGGCCTCTGAATCCGGATGTAGAG

GTAGAAAGATTATTATGCGTCCTAGGTCTAGCCGAAAAGAGAGACGCAGTTATAAGCATTGAGAGCCTAGACCCAGC
TTTAGCTAAGGTCCCCTTTCCAGTACCTACCACCTACGGTGCAGTTCTACGTCACTACATCGACATATCCGCCGTTG
CTGGCCGTCAGATCCTTGGCACTTTATCAAAATTCGCGCCGACTCCGGAAGCCGAAGCCTTTCTAAGGAATCTAAAT
ACCAACAAAGAAGAGTACCATAATGTCGTAGCCAACGGCTGCTTGAAGTTGGGTGAGATCCTGCAGATCGCGACTGG
CAATGACATAACCGTTCCCCCGACCACGGCCAACACGACAAAGTGGCCCATACCTTTTGACATTATAGTAAGCGCGA
TACCAAGGTTGCAGCCAAGATATTATTCCATTAGTAGTTCCCCGAAAATCCACCCTAACACGATACATGCAACTGTC
GTCGTATTAAAGTACGAAAACGTGCCGACAGAGCCCATTCCCAGGAAGTGGGTTTATGGGGTGGGGTCCAACTTTCT
ATTAAACCTGAAGTATGCTGTGAACAAGGAGCCAGTACCCTACATCACGCAGAATGGTGAACAGCGTGTCGGAGTCC
CAGAATATCTAATAGCGGGCCCTAGGGGGAGTTACAAAACAGAGTCCTTCTACAAGGCCCCCATACATGTGAGAAGA
TCTACCTTCAGGCTACCCACCAATCCGAAAAGTCCAGTTATCATGATTGGACCGGGTACGGGCGTTGCACCGTTTCG
TGGCTTCGTACAGGAGAGGGTCGCACTGGCTCGTCGTTCCATTGAAAAGAATGGGCCGGACAGCCTGGCTGACTGGG
GACGTATTTCCTTGTTTTACGGTTGCAGACGTAGCGATGAAGATTTCCTTTACAAAGATGAGTGGCCACAGTATGAG
GCCGAGCTTAAAGGCAAGTTCAAGCTACACTGTGCTTTCAGCCGTCAGAACTATAAACCCGATGGCAGTAAAATTTA
TGTGCAAGACCTGATATGGGAAGATCGTGAGCATATCGCAGATGCGATACTTAACGGTAAGGGTTATGTATACATCT
GTGGAGAAGCAAAGTCAATGAGCAAACAGGTAGAAGAGGTTCTGGCCAAAATTCTGGGTGAAGCCAAAGGGGGCTCT
GGTCCGGTAGAAGGAGTGGCCGAAGTCAAACTGTTGAAGGAGCGTAGCAGACTAATGCTAGACGTCTGGAGC

Seq. ID No: 180
>T4H_CPR_chimera_2
TCTTCATCATCCAGTTCAGGGACGATAGCTTATTTTACCAAAGGAACCTACTGGGGTATTGTCAAAGATCCATATGC
TCCTAACTATCCTCCTGCAAATGGGAATAAGCCCGCCAAAACAAGAAATATTGTTGAGAAAATGGACGAAAGCAACA
AGAATTGTGTTGTATTCTATGGTTCCCAGACTGGAACTGCCGAAGATTACGCGAGTCGTTTAGCCAAGGAAGGGAAA
TCACGTTTCGGGCTGGAAACGATGGTAGCTGATCTTGAGGACTATGACTTCGATAACTTGGACACACTAGGAGATGA
CAAAGTTGCAATTTTTGTGTTGGCAACCTATGGTGAGGGGGAGCCCACAGATAACGCGGTTCAATTGCTACAGAATT
TACAAGACGAGAGCTTTGAGTTTTCATCTGGCGAGAGGAAGCTAAGTGGATTGAAATATGTGGTCTTTGGACTAGGA
AACAAGACATATGAACACTACAATTTGATAGGAAGAACCGTTGATGCTCAGCTTGCTAAGATGGGTGCTATACGTAT
CGGAGAAAGAGGGGAAGGAGACGACGACAAATCTATGGAGGAGGACTATCTTGAATGGAAAGATGGGATGTGGGAGG
CATTTGCCACGGCAATGGGGGTGGAAGAAGGCCAGGGTGGTGATTCCGCTGACTTTGTTGTGAGCGAATTGGAAAGT
CATCCTCCTGAGAAAGTTTATCAAGGGGAATTTAGCGCGCGTGCTCTGACAAAGACTAAAGGAATACATGACGCAAA
GAACCCCTTCGCCGCTCCCATAGCGGTCGCCAGGGAGCTGTTCCAGTCTGTCGTTGACAGAAATTGCGTGCACGTAG
AGTTCAACATTGAAGGTAGTGGTATTACATACCAACATGGTGATCATGTCGGACTTTGGCCTCTTAACCCGGATGTA
GAAGTTGAACGTCTGCTTTGTGTGTTAGGCCTGGCCGAAAAACGTGACGCCGTTATAAGCATTGAGTCTCTGGACCC
TGCATTAGCAAAGGTCCCGTTCCCTGTTCCTACGACATACGGTGCTGTCCTTCGTCATTACATTGATATCAGTGCTG
TAGCAGGAAGGCAAATTCTAGGGACCCTAAGTAAGTTTGCACCAACCCCGGAAGCAGAGGCATTTTTACGTAACTTA
AATACCAATAAAGAAGAATACCACAACGTAGTAGCGAACGGATGTCTAAAACTTGGAGAGATTTTACAAATTGCTAC
CGGCAACGACATTACGGTACCGCCGACCACGGCCAATACAACTAAGTGGCCGATACCATTTGACATCATAGTGTCAG
CCATACCGCGTTTGCAACCACGTTATTACTCTATCTCAAGTAGCCCAAAGATCCACCCCAACACAATTCATGCAACC
GTGGTGGTTTTGAAGTATGAGAACGTCCCAACGGAGCCTATCCCCAGGAAGTGGGTATACGGAGTGGGTTCTAACTT
TTTGTTAAACTTAAAATACGCAGTAAATAAAGAGCCTGTCCCGTACATAACACAGAATGGGGAACAGCGTGTTGGCG
TGCCTGAGTATTTGATCGCTGGACCCAGAGGCTCTTATAAAACTGAAAGTTTTTACAAGGCCCCGATCCACGTGAGG
AGGAGCACCTTCAGATTACCCACCAATCCTAAGTCCCCAGTCATAATGATCGGGCCTGGAACAGGCGTGGCCCCCTT
TAGGGGCTTTGTTCAAGAGAGAGTAGCATTGGCGCGTCGTTCCATAGAAAAGAACGGCCCAGATTCACTGGCCGACT
GGGGCAGGATTAGTTTGTTTTACGGGTGCCGTCGTTCTGATGAAGATTTCTTGTACAAAGATGAGTGGCCGCAGTAC
GAAGCGGAATTGAAAGGCAAATTTAAATTACACTGCGCCTTTAGTCGTCAGAACTACAAGCCGGACGGCAGCAAGAT
TTATGTCCAGGATCTTATATGGGAGGACAGAGAACACATCGCAGATGCGATACTGAATGGTAAGGGCTACGTTTACA
TATGTGGAGAAGCCAAATCCATGTCTAAACAGGTTGAAGAGGTCCTAGCCAAAATACTAGGGGAGGCAAAAGGGGGT
AGTGGCCCCGTAGAGGGGGTGGCTGAAGTTAAGCTGTTAAAGGAGAGGAGTAGACTAATGTTAGATGTGTGGAGC

Seq. ID No: 181
>T5H-CPR_1

AGCAGTAGCTCCTCCTCAGGCGGCTTACTTGCCTTTTTATACTTATTTAGGGGTACGTTATTTGCATCAGGAAAAGC
GAGCGATGCCGGGTCAAAATTAGCAGGAGGGTCTGATCTGGACAGCAGTGCTGACGCCGCTGCCAACGACTTTGTTA
CTAAACTGACGTCCCAAAATAAACGTATAGCTATTTTCTACGGAAGCCAGACGGGAACAGCAGAAGAGTATGCGACA
AAGATCGCTAAAGAAGCGAAAGCCCGTTTTGGAACGTCATCCCTTGTGTGTGACCCCGAGGAGTATGAATTTGAAAA
ATTAGACCAGTTGCCTTCCGATTGTGTGGCCTGTTTTGTCATGGCGACCTACGGAGAGGGCGAACCGACCGACAACG
CAGTGGGTCTTATGGAGTTCTTAGATGGAGAAGATGTGCAGTTCAGCAATGGGTCTAGCCTGGATAATCTTAACTAC
GTTATCTTTGGCCTAGGAAATAGAACTTACGAGCACTATAACGCAATCGCACGTAAACTAGACGCCCGTCTTGAATC
TCTAGGGGCGAAGAGGATAGGCGAGAGGGGCGAGGGAGATGACGATAAGAGTATGGAGGAGGATTATTTAGCTTGGA
AAGATGGAATGTTTGAAGCCTTGGCAAGCTCATTAGGTTTCGAAGAGGGCGGAGGGGGCGACGTGGCCGATTTTAAG
GTACGTGAGGTAGCAGATCACCCGGAAGATAAAGTGTATAGGGGTGAACTTTCTGCGAGAGCCTTGCTAGGCACGAA
GGGCATTCATGATGCCAAGAACCCCTACAACGCAGTAGTAAAGGAGGCTAGGGAACTATTCGTCGAAGGGACAGCTG
ACAGGACGTGTGTCCATGTTGAGTTCGACATCGAGGGTTCCGGAATCTCCTACCAACACGGTGATCATATAGCCGTA
TGGGCTCATAATCCTGAACAGGAAGTTGAAAGAGCATTAGCCGTCCTTGGTTTGCTGGGCAAACGTGATACGGTGAT
AGATGTAGAAAGTCTCGATCCCACCTTGGCGAAGGTTCCTTTTCCGGTTCCAACTACGTACGAGGCCGTCTTTAGAC
ATTACCTGGATATTTGTGCACACGCGAGCAGACAAACTCTGAACAACTTCGCAAAGTATGCGCCCACCCCCGAAGCC
CGTGCTAAGCTGGAGAAAGCCTGCGGAGACAAGGCAGCGTTCCAAGAGGCAATAGGCCATAGATGTCTGAAAACTTT
TGAGGCATTACAATTGATTGTCGGCGATGATCTAGGTGGTGACTCCGTTGCAAAAGCTACTGCCTGGGAGATACCAT
TTGATAGAGTAATCTCTGACCTTCCCAGAGTCGGACCAAGGTTTTATTCAATATCATCTTCACCTAAGATGCACCCC
AAGACGGTGCACATCACGGCTGTCGTGTTGAGATATAGGCCAGAGGCAGCAGGGCAAGACTCCCCGTACGTTCATGG
CTTAGCAACGAACTTTATATCCGCTATAAAGATGGCTAAAAATAATGAGCAGCCTAGTGGACCGGATGATCCGAGAT
TCGGAACTCCTGGGTATGACCTAGCAGGCCCCAGGGGTGCCTACACAAAAGAGAGTCTATTTAGGGCCCCTATCCAT
ATAAGACGTAGTAATTTCCGTCTACCCACGTCACCCAAAATCCCAGTAATCATGGTAGGTCCTGGCACGGGGGTCGC
ACCCTTTCGTTCATTCGTGCAGGAGCGTGTCTGCAGCGCACAGAAGACCTGTGATAAAGTAAATCAATCCCCGGCCG
AGGCCCTTCAGGACTGGGGTAATATCTGGCTATTCTATGGTTGTAGAAGGAGCAACGAAGACTTCCTGTATAAGGAT
GAGTGGCCCGAGTACGCGAGCAAGCTGGGAGGGAAATTCCAAATGGAGACAGCCGTCAGTAGAGAGAAGTTCAAGCC
AGACGGCAGTAAACTTTATGTGCAGGACTTAATATGGGAGAGGCGTAAAGAACTGGCACAAGATATACTAGATAAGA
AAGCGTATATTTACATCTGTGGTGAGGCGAAAGGGATGGCCCACGATGTCGAAGAAATGTTTGGTAGGGTGCTTGAG
GAGGCTAAAGGTTCCGCTGAAGCTGGCAGACGTGAACTTAAATTACTGAAGGAGAGGTCCCGTTTACTTTTAGACGT
TTGGTCT

Seq. ID No: 182
>T5H-CPR_2
TCCTCCAGCTCTTCATCACTGTTCTCAACTACTGACGTGATTCTATTCAGCCTTATTGTGGGTGTCATGACATATTG
GTTTCTGTTTCGTAAGAAGAAGGAGGAAGTGCCAGAATTTACAAAAATACAGACAACAACGAGTTCAGTGAAAGATC
GTTCTTTCGTCGAAAAGATGAAGAAGACAGGTCGTAATATCATTGTGTTTTATGGAAGTCAGACCGGGACCGCAGAG
GAGTTTGCCAACCGTCTGTCCAAGGACGCGCACAGATATGGAATGCGTGGTATGGCTGCCGATCCTGAAGAGTACGA
TCTAGCGGACCTGTCATCACTGCCTGAGATAGAAAAAGCCTTGGCAATCTTTTGTATGGCGACTTATGGTGAGGGCG
ACCCCACGGACAATGCCCAGGACTTTTACGACTGGCTTCAGGAGACAGACGTCGATCTTAGCGGCGTTAAATATGCA
GTCTTCGCCTTGGGAAACAAGACATACGAGCATTTCAACGCGATGGGAAAATACGTAGACAAAAGGCTAGAGCAACT
TGGGGCACAGAGAATATTTGATCTTGGATTGGGAGATGATGACGGAAATTTGGAGGAAGACTTCATTACATGGAGGG
AGCAGTTTTGGCCGGCTGTGTGCGAGCACTTTGGGGTCGAAGCCACCGGTGAGGAAAGCAGTATAAGGCAGTATGAA
CTTATGGTGCATACCGATATGGACATGGCCAAAGTATATACGGGAGAGATGGGTCGTCTAAAGAGTTATGAAAATCA
GAAGCCGCCCTTTGATGCCAAGAACCCTTTCCTGGCCGTTGTCACGACCAATAGAAAGCTAAATCAAGGCACGGAGA
GGCACCTTATGCACCTGGAGCTTGATATTTCAGATAGCAAGATACGTTACGAAAGCGGGGACCATGTCGCCGTCTAT
CCTGCAAACGACTCCGCACTGGTTAACCAGTTAGGAGAGATATTGGGCGCTGATCTTGATATAATAATGTCCCTAAA
CAACCTGGACGAAGAAAGCAACAAAAAGCATCCGTTTCCTTGCCCAACCAGCTACAGGACTGCGTTAACGTATTATT
TGGATATCACTAACCCTCCTAGAACCAACGTGTTATACGAATTAGCACAATACGCGAGCGAGCCGACCGAACACGAG
CAACTGAGGAAAATGGCTAGTAGCTCAGGGGAAGGAAAGGAGTTATATCTGAGATGGGTCCTAGAGGCAAGGCGTCA
TATCTTAGCCATTCTACAGGATTATCCGTCCCTTAGGCCTCCCATAGACCATTTATGTGAGTTATTGCCCAGATTGC
AAGCGAGATATTACTCCATCGCAAGTTCATCAAAAGTTCATCCTAATTCTGTGCACATCTGTGCGGTAGCGGTGGAA
TATGAAACTAAAACGGGGAGGATTAATAAGGGTGTCGCTACGAGCTGGCTAAGGGCTAAAGAGCCGGCTGGCGAAAA

TGGCGGACGTGCATTGGTCCCAATGTACGTAAGAAAGTCCCAGTTCAGACTACCTTTCAAGGCGACGACGCCTGTCA
TAATGGTAGGCCCAGGTACCGGTGTGGCTCCATTTATCGGGTTCATTCAGGAAAGAGCGTGGCTGAGGCAACAGGGA
AAGGAGGTCGGTGAGACACTATTGTACTACGGCTGCAGGAGGAGTGACGAGGATTACCTTTACAGGGAAGAGTTGGC
AGGCTTTCATAAAGATGGTGCCCTTACACAGCTTAACGTCGCCTTCAGTAGAGAACAGCCGCAAAAGGTTTATGTTC
AACACTTACTTAAAAAGGACAAAGAACACTTGTGGAAACTAATACATGAGGGTGGGGCGCATATTTATGTGTGTGGA
GATGCGAGGAATATGGCTCGTGATGTGCAAAATACGTTTTATGACATAGTGGCTGAGCAGGGGGCGATGGAACATGC
CCAAGCCGTGGATTACGTCAAGAAGTTAATGACCAAAGGCCGTTACAGCCTAGACGTTTGGTCT

Seq. ID No: 183
>T5H-CPR_3
AGTAGCTCCTCATCATCCGAGGCAGTAGCCGAGGAGGTTTCCCTGTTTTCTATGACTGATATGATCCTATTCTCCTT
GATTGTTGGTTTGTTGACGTACTGGTTTCTTTTTCGTAAGAAGAAAGAAGAGGTTCCGGAATTTACTAAGATTCAAA
CTTTGACCAGTAGCGTCCGTGAGTCATCTTTCGTCGAAAAGATGAAAAAGACAGGTAGAAATATCATAGTATTCTAC
GGCAGTCAAACCGGTACAGCGGAGGAGTTCGCTAACAGATTATCTAAGGATGCCCACCGTTATGGCATGAGAGGAAT
GTCCGCTGATCCTGAGGAGTACGATCTGGCTGACCTAAGCTCCCTACCCGAAATTGATAATGCACTGGTGGTTTTCT
GTATGGCTACATACGGGGAGGGAGACCCGACGGACAACGCTCAAGACTTCTACGACTGGCTGCAGGAAACAGACGTT
GACCTGTCTGGTGTAAAATTCGCGGTGTTTGGCCTGGGGAATAAGACTTATGAGCACTTCAACGCGATGGGAAAATA
TGTTGATAAAAGACTTGAGCAACTTGGAGCACAGAGGATTTTCGAACTTGGATTGGGAGATGACGACGGTAACTTAG
AGGAGGATTTTATCACCTGGCGTGAACAATTCTGGCCCGCAGTGTGTGAACATTTCGGGGTGGAGGCGACCGGTGAG
GAAAGCTCAATCAGGCAATACGAGTTGGTGGTTCATACCGATATAGACGCCGCCAAAGTTTACATGGGAGAGATGGG
CAGACTGAAGTCTTACGAGAATCAAAAGCCCCCTTTCGACGCCAAGAACCCATTTTTAGCGGCAGTCACTACTAATA
GGAAACTGAACCAAGGCACCGAACGTCACCTTATGCATCTGGAACTGGATATCAGCGATAGCAAGATCAGGTATGAA
TCAGGAGATCATGTGGCGGTCTATCCTGCCAACGATTCAGCTCTTGTCAACCAGTTGGGAAAAATTTTGGGTGCTGA
CTTAGACGTTGTAATGTCACTTAACAATTTAGATGAGGAGAGCAACAAAAAGCACCCTTTCCCCTGCCCCACAAGCT
ACAGGACCGCATTAACTTATTACCTAGACATAACGAACCCTCCACGTACTAACGTCTTATATGAATTGGCACAATAC
GCCTCCGAACCGTCCGAACAGGAGTTACTGCGTAAGATGGCTAGTAGTTCAGGAGAAGGTAAGGAATTATACCTAAG
TTGGGTTGTAGAGGCAAGGAGACACATCCTTGCCATATTGCAAGATTGTCCATCTCTGAGGCCACCAATCGACCACC
TTTGCGAGTTGCTACCTAGACTTCAGGCGAGGTACTATAGTATAGCATCTTCCAGCAAAGTGCATCCCAATTCAGTG
CATATTTGCGCCGTTGTCGTTGAATACGAAACAAAGGCAGGAAGGATCAATAAGGGCGTCGCGACAAATTGGCTGCG
TGCGAAGGAGCCTGCGGGGGAGAACGGTGGACGTGCTCTAGTTCCGATGTTTGTTAGAAAATCTCAATTCCGTTTGC
CTTTTAAAGCTACCACCCCCGTAATAATGGTAGGTCCAGGAACCGGTGTAGCACCATTTATTGGTTTCATACAAGAA
AGAGCTTGGCTAAGGCAGCAAGGAAAGGAAGTAGGCGAGACCTTGTTATATTATGGTTGTCGTCGTTCAGATGAGGA
TTACCTGTACAGAGAGGAGCTGGCTCAATTCCATCGTGATGGGGCTCTGACCCAATTAAATGTGGCATTCAGCCGTG
AACAATCTCACAAAGTGTATGTTCAACATTTGCTAAAACAGGATAGAGAACATCTATGGAAATTAATAGAAGGGGGT
GCTCACATATATGTCTGCGGGGATGCCAGAAATATGGCGCGTGACGTGCAGAACACGTTTTATGACATTGTGGCCGA
GCTAGGCGCGATGGAGCACGCCCAAGCCGTCGATTACATCAAAAAGTTGATGACGAAGGGCAGGTACTCATTAGATG
TGTGGAGT

Seq. ID No: 184
>T5H-CPR_4
TCCTCTAGCAGTAGCTCTGCCGCGGCTGCAGACGGCGACGGTGGGCAATCTCGTAGACTTCTGGCTTTACTGGCTAC
CAGTCTAGCAGTATTAGTGGGATGCGGCGTTGCGTTATTGTTCAGGAGAAGTAGTAGTGGTGCAGCTCCACTAGCTA
GGCAGGCCGCTGCAGCCAAACCCCTTGCAGCTAAAAAGGACCAGGAACCAGATCCAGATGACGGCAGGCAAAGGGTC
GCCTTGTTCTTCGGCACCCAGACCGGGACAGCGGAGGGTTTTGCGAAGGCACTAGCCGAGGAAGCCAAGGCGAGGTA
CGATAAAGCAGTGTTTAAAGTCTTGGATCTGGACGATTACGCAGCCGAGGATGAAGAATATGAAGAAAAGTTGAAAA
AGGAGAACATCGCGTTTTTCTTTTTGGCAACGTATGGGGATGGAGAACCAACAGACAACGCAGCGAGATTTTACAAG
TGGTTCAGCGAAGGGAATGAAAGGGGTGAGTGGTTGTCTAACCTGCAGTACGGGGTATTTGCATTGGGGAATAGACA
GTACGAACATTTTAATAAGGTCGGAAAGGAAGTGGACCAACTACTGGCAGAACAGGGCGGGAAAAGAATTGTACCTG
TTGGATTGGGTGATGACGACCAATGTATAGAAGATGACTTCAACGCGTGGAAAGAGTTACTGTGGCCTGAACTGGAT
AAACTTTTGAGGGTAGAAGACAACTCATCCGCAGCTCAAAGTCCCTATACTGCGGCGATTCCTCAATATAGAATAGT
GCTAACAAAACCCGAAGACGCAACACATATTAATAAGAGCTTTAGTCTGTCAAATGGACACGTGGTGTACGACTCAC

AACACCCGTGCAGGGCCAACGTGGCGGTAAGACGTGAGTTACACACGCCGGCAAGTGACAGGAGCTGCATTCATCTT
GAGTTCGACATCGCCGGGACCTCACTTACGTACGAGACGGGAGATCACGTTGGAGTATATGCCGAGAACAGTACTGA
GACGGTGGAGGAGGCGGAAAAACTACTTGATTACAGCCCAGATACTTATTTCAGCATTTATGCAGATCAAGAGGACG
GTACCCCATTGTTTGGTGGTAGTCTGCCACCCCCGTTTCCGCCATGTACAGTAAGGGTGGCTTTAGCAAGATATGCT
GACCTATTAAATTCCCCCAAGAAGAGTGTATTATTAGCCCTGGCCGCTCATGCATCCGACCCCAAGGAAGCTGAGAG
GCTGCGTCACCTGGCATCTCCGGCTGGAAAGAAGGAGTATTCTCAGTGGATTATAGCGTCTCAAAGGTCTCTGTTAG
AGGTAATATCAGAGTTTCCCTCCGCTAAGCCACCACTGGGAGTTTTCTTCGCAGCGATTGCCCCGAGGTTACAACCG
CGTTATTACTCCATTAGCAGCAGTCCCAGGATGGCGCCCACAAGAATACATGTCACCTGTTCATTGGTGCACGGCCA
ATCTCCCACCGGTAGAATCCACAAAGGCGTATGCAGTACATGGATGAAAAACAGTACGCCCAGCGAGGAAGAGTCCG
AGGAGTGTTCCTGGGCACCAATCTTTGTGAGACAATCCAATTTCAAGCTACCCGCCGATCCCACTGTCCCGATCATT
ATGGTGGGGCCTGGAACAGGTCTAGCTCCATTCAGGGGTTTTCTACAGGAAAGACTGGCTTTAAAAGAAACGGGAGT
TGAGTTGGGGCGTGCAATTCTATTTTTCGGATGTCGTAATCGTCAAATGGATTTTATTTACGAGGATGAGCTGAACA
ATTTTACGGAAAGTGGCGCGCTTTCCGAGCTGGTAGTTGCCTTTTCTCGTGAAGGTCCCACCAAGGAATATGTTCAA
CACAAGATGGCAGAAAAAGCGGCAGATTTATGGTCAATAGTATCTCAAGGGGGTTATGTCTATGTGTGCGGGGACGC
TAAGGGAATGGCACGTGACGTGCACAGAGCATTGCACACCATAGTACAGGAACAAGTCACACAAAGGACCTCCAACT
TTGGGCTTTGGAAGTTCCGTCTTGTTTCACTAAAT

Seq. ID No: 185
>T5H-CPR_5
AGTTCATCTTCATCATCTGCGGCAGCCGCAGGCGGCGATCCACTAGCCGCGCTGGCTGCCACCGCCGCAGCGTTAGT
GGCAGGAGTAGTAATACTAGCTGTTTGGTTCAGGTCAGGTGGAGGTGCGCCTCCGAAAGCAGCTGCTCCACCACCTC
GTCCACCTCCGGTTAAGATTGAAGCCGACGCAGATGCTGATGATGGCAGGAAGAGGGTGACCGTGTTCTTTGGCACT
CAGACCGGGACCGCAGAAGGCTTTGCCAAAGCGATGGCCGAAGAAGCTAGGGCCCGTTACGAAAAGGCTGTGTTCAA
AGTCGTTGACCTGGATGACTATGCGGCAGAGGATGAGGAATATGAGGAAAAACTTCGTAAAGAAACCATCGTCTTAC
TGTTCCTGGCGACGTATGGGGACGGTGAGCCTACCGACAATGCGGCGCGTTTCTATAAGTGGTTTACAGAAGGTAAA
GAAAAAGAGGTTTGGCTAAAAGACCTAAAATATGCGGTATTTGGATTGGGTAACAGACAATACGAGCACTTTAATAA
GGTCGCCAAGGTAGTCGATGAGCTATTGGAGGAACAAGGAGGGAAAAGACTTGTTCCCGTGGGTTTAGGAGATGATG
ACCAGTGCATAGAGGATGATTTCACGGCTTGGAAGGAGCAGGTCTGGCCAGAACTTGATCAGCTTTTAAGAGACGAA
GACGACACAACGGGTGCATCTACACCTTACACTGCGGCCATACCCGAATATCGTATCGTTTTCATTGATAAGAGCGA
TGTATCATTTCAAGACAAGAGCTGGTCCCTTGCAAACGGTTCAGGGGTCATAGATATTCATCATCCTGTAAGGAGTA
ACGTCGCAGTCCGTAAAGAGTTGCACAAGCCGGCTTCTGATAGGTCTTGTATCCACTTGGAGTTCGACATTTCAGGC
ACCGGCCTTGTGTATGAGACCGGCGATCATGTCGGGGTGTATAGCGAGAATGCCATCGAAACAGTTGAGCAAGCGGA
AAAGCTACTAGACCTTTCTCCAGATACGTTCTTTAGCGTCCATGCAGATGCAGAGGACGGATCTCCCAGAAAAGGAG
GGGGATCTTTAGCCCCACCCTTCCCGTCTCCATGTACTCTTAGAACCGCCCTATTGAGGTATGCTGACCTGCTGAAC
AGTCCTAAGAAAGCTGCATTAGTAGCTCTGGCAGCGCATGCGTCCGATCTAGCTGAAGCAGAAAGGCTACGTTTCTT
GGCTTCCCCAGCCGGAAAAGATGAGTATTCACAATGGGTGGTAGCGTCCCAGAGGTCCCTGTTGGAAGTTATGGCCG
CGTTTCCCAGTGCAAAGCCGCCGCTGGGAGTGTTCTTTGCAGCTGTCGCCCCCAGACTTCAACCACGTTACTACTCC
ATTTCATCCTCTCCGAAAATGGCTCCCTCCCGTATCCACGTCACCTGCGCGCTAGTGTACGGCCCGACTCCTACAGG
CCGTATACACCAGGGCGTTTGTAGCACGTGGATGAAGAATGCAATACCAAGCGAGTACAGTGAGGAGTGTTCCTGGG
CACCCATATACGTTAGACAAAGTAATTTTAAACTACCCGCGGACCCGACGACCCCAATTATTATGATAGGACCCGGT
ACAGGGCTTGCTCCATTTCGTGGCTTTCTACAAGAACGTTTGGCTTTAAAACAGTCTGGAGTTGAGCTGGGGAACTC
TGTGTTATTTTTCGGGTGTAGAAACCGTAATATGGATTATATTTACGAAGATGAACTACAGAACTTTATTCAGGAGG
GCGCTTTGAGCGAGTTGATCGTCGCTTTTAGTAGGGAAGGCCCAGCGAAGGAATATGTCCAACACAAAATGACTGAA
AAGGCAACGGAAATCTGGAATATAGTCTCTCAAGGTGGTTATATTTACGTTTGCGGGGATGCAAAGGGCATGGCTAG
GGACGTCCACCGTGCTTTGCACACTATAGTACAAGAACAAGGCTCTCTTGATAGCAGTAAGACAGAGTCATATGTAA
AATCTCTTCAAATGGACGGCCGTTATTTACGTGACGTCTGG

Seq. ID No: 186
>T5H-CPR_6
TCTTCCTCCTCCTCAAGCGCCGCTGCATATCTGTTTAGGGACCAAATTTTTCGTTCTAGCAGTCCGAAAGTAGTAGT
ACCCGCCCCTTCCAAACTGGCGAACGGCCACGGAAATCCTCGTAACTTTGTTAGTAAGATGAAGGAAGGGAAGAAGA

GGATTGTAATCTTTTATGGGAGCCAGACTGGTACCGCGGAGGAGTATGCGATTAGGATCGCTAAAGAAGCTAAAACG
AAGTTCGGTCTTACCTCACTGGTTTGCGACCCTGAGGAATACGATTTCGAGAACCTGGACCAAGTACCAGAAGACTG
CTGCGTATTCTTTGTGATGGCCACATACGGAGAGGGGGAACCGACGGACAATGCTGTTCAACTGATGCAAAACTTAG
AGGATGAGTCCTTCGAGTTTAGTAATGGATCACACAGGTTGGACGGTTTGAAATATGTGGTGTTTGCGCTGGGCAAC
AAAACGTACGAGCACTATAACGCTATTGGACGTAAGGTAGACACATTGTTAACAGACATGGGGGCGACAAAAATCGG
GGAACGTGGAGAAGGAGACGACGATAAGTCAATGGAAGAGGATTATTTAGAGTGGAAAGATGGAATGTGGAAGGCGT
TTTCTGAAGCGATGGGTGTAGAGGAAGGGCAGGGCGGGGATACTCCAGATTTTGCTGTTACTGAGCTTGATAGTCAC
CCGCCAGAGAAAGTATATCTAGGCGAGCTTAGTGCCAGGGCCTTAACTAGGACTAAAGGCATTTATGACGGTAAAAA
TCCATACCCCTCCGCCGTAAAACATAGTAGGGAACTTTTTCAGGCTGGTGCAGAGAGGAACTGCGTACACGCGGAAC
TAGATATTGAGGGCTCTGGCATTACATATCAGCACGGTGATCATGTAGGAGTGTGGCCGAGCAACCCCGATGTTGAG
GTCGATCGTATGCTATACGTGTTAGGTCTATATGGCAAGAAAGACGCCGTGATAAATATAGATTCCCTAGACCCTGC
GCTGGCGAAAGTACCCTTCCCCGTACCGACTACATATGCCACGGTTCTGAGACACTACATCGACATATGTGCTGTGG
CTGGGCGTCAGATGTTGGGGGTCCTTAGCAAGTTCGCACCGCATCCGAAAGCCGAGGCTTTCCTGAAATCATTGAAT
AGCGATAAGGAAGAGTACTCAAACATAGTAACGAACGGGTGTTTTAAACTGGGTGAAGTTCTGCAGTTAGCGGCAGG
CGACGATATTAAACTATGTCCCACCCCAGACAACACCACAGCTTGGGCAATACCCTTCGATATCATAGTATCCTCTA
TACCGAGACTACAACCGCGTTTCTACAGTATTAGTTCCTCTCCCAAATTATACCCCAATGCAATCCATCTAACAGCT
GTAGTGCTGAAGTACGATAGTATCCCAAATAGGCTGGTGGAGTCTCGTTTTGTATATGGCGTGGCCACAAATTTTCT
ATTGAATGTGAAGTACGCAGCTAATGGTGAGACGGCTCCATTCATCGCCGAACCAGTAATATCTGAGCCGGCACATG
TCTCACTTCCAAAATACGCCATAGAAGGACCTAGAGGAGCCCATATCGAGGACAATATTTATAAGATACCGATACAT
GTCCGTAGATCCACTTTTAGGCTACCTGCTAATCCGAAGATTCCAGTAATCATGGTCGGACCGGGAACAGGCGTCGC
GCCCTTTAGAGGGTTCGTGCAGGAGAGAGTTGCACTTGCTAAACGTAGCATTGAGAAAAATGGGCCAGACGCCCTTG
CCGATTGGGGCAGCATTACACTGTTTTACGGTTGTAGGAAATCCAATGAAGACTTTTTATATAAGGAAGAGTGGCCT
CAATATGCGGAAGAGTTAAAAGGTAAATTTAAGATGCATTGCGCTTTCAGTAGAGAGCCTCCTTACAAGCCTGACGG
TAGCAAAATATATGTTCAAGACCTAATCTGGGAAGAAAGAGAGACGATAGCAAAGGCAATTCTTGAAGGAAAGGCGT
ATGTTTATATCTGCGGTGATGCAAAAGCCATGAGTCGTGCAGTCGAAGACACCCTGGCCAGGATTCTGGGAGAAGCA
AAAGGTGGCAATGCAGAAGTGGAAGGTGCAGCAGAGATGAAGATCTTGAAGGAACGTAGTAGACTACTGCTTGACGT
ATGGTCT

Seq. ID No: 187
>T5H-CPR_7
TCAAGCTCTTCCAGCAGCTCACTGTTTTCTACTACTGACATGGTACTGTTCTCTTTAATTGTCGGAGTCCTGACCTA
TTGGTTCATTTTCAGAAAGAAGAAGGAGGAGATTCCCGAGTTCAGTAAAATACAGACTACCGCCCCACCTGTCAAGG
AGAGCAGCTTTGTGGAGAAAATGAAGAAAACCGGCAGGAACATTATAGTCTTTTACGGCTCTCAGACCGGCACGGCG
GAGGAGTTCGCAAACAGACTGTCCAAAGACGCCCATCGTTATGGTATGCGTGGCATGAGTGCGGACCCAGAAGAATA
CGACCTTGCGGATTTATCATCCTTGCCGGAAATTGATAAGTCACTTGTGGTCTTTTGCATGGCAACATACGGAGAGG
GTGACCCAACGGACAACGCGCAAGACTTTTATGACTGGCTTCAAGAGACGGATGTAGACTTAACTGGCGTTAAATTT
GCCGTTTTTGGACTAGGTAATAAGACTTACGAACATTTCAACGCAATGGGCAAGTATGTTGATCAGCGTCTTGAGCA
GCTGGGGGCCCAACGTATTTTCGAATTGGGCTTGGGAGACGATGATGGCAACTTAGAAGAGGACTTTATTACTTGGC
GTGAGCAGTTTTGGCCTGCAGTGTGCGAGTTCTTCGGGGTGGAAGCTACCGGCGAGGAATCATCTATCCGTCAGTAT
GAATTAGTGGTCCACGAGGACATGGATGTTGCTAAAGTGTATACCGGCGAAATGGGCAGGCTAAAATCTTACGAAAA
TCAGAAGCCACCCTTCGACGCCAAGAATCCATTTCTAGCAGCGGTCACTGCCAATAGAAAACTTAATCAGGGGACTG
AGAGACATCTAATGCATTTAGAACTAGACATCAGTGACTCAAAAATAAGATATGAATCCGGCGACCACGTTGCAGTG
TACCCTGCCAACGATTCCGCGCTAGTAAACCAGATAGGTGAGATCTTAGGTGCGGACCTAGACGTAATCATGAGTTT
GAATAACCTAGATGAAGAGTCTAATAAAAAGCACCCCTTCCCTTGTCCTACAACGTATAGGACGGCCCTTACATACT
ACCTAGACATCACAAACCCGCCAAGAACTAACGTGCTTTATGAGTTAGCTCAATACGCCTCAGAACCTTCTGAGCAA
GAGCACCTGCATAAAATGGCAAGTTCCTCCGGTGAAGGTAAAGAGCTTTATTTATCATGGGTAGTAGAAGCACGTAG
ACACATATTAGCCATTCTGCAAGACTACCCATCTTTGCGTCCTCCCATCGACCATTTGTGCGAACTTTTGCCCCGTC
TGCAGGCACGTTATTATTCCATAGCCAGCAGTTCAAAGGTCCATCCAAATTCCGTACACATATGCGCTGTGGCCGTG
GAATATGAGGCGAAATCCGGGCGTGTTAATAAGGGTGTCGCGACAAGTTGGCTACGTGCCAAGGAGCCCGCGGGTGA
AAATGGGGGCAGGGCCCTAGTTCCTATGTTTGTAAGGAAGTCACAGTTTAGATTACCATTTAAAAGTACTACCCCAG
TAATCATGGTTGGCCCCGGGACCGGGATCGCGCCTTTCATGGGATTCATTCAGGAAAGAGCTTGGCTACGTGAGCAA

GGAAAGGAGGTTGGCGAGACGTTATTGTACTATGGATGTAGGAGGAGTGACGAAGACTACCTATACCGTGAAGAGCT
GGCGAGGTTTCACAAAGATGGTGCATTGACTCAGCTTAATGTGGCATTTTCCCGTGAGCAGGCCCACAAGGTCTACG
TCCAACATCTGCTTAAACGTGATCGTGAGCATTTATGGAAACTGATTCACGAAGGCGGTGCTCATATCTATGTTTGT
GGTGATGCTAGGAATATGGCCAAAGATGTTCAAAATACATTTTATGATATAGTGGCTGAATTTGGGCCGATGGAGCA
TACTCAAGCCGTTGATTACGTGAAAAAGTTGATGACAAAGGGCAGGTACAGCCTGGATGTATGGTCA

Seq. ID No: 188
>T5H-CPR_8
TCATCTAGCTCATCATCTGGGGGCTCTCCCATGTCCGATTCCGTTGTCGTAATCATCACGACGAGCTTCGCAGTCAT
AATCGGCTTACTGGTGTTTCTTTGGAAACGTTCTTCAGACAGGAGTAAGGAGGTTACTCCGCTTGTTGTACCAAAAA
GTCTTTCCGTGAAAGACGAAGAGGACGAAGCTGAAACACTTGCGGGCAAAACAAAAGTCACAATTTTCTATGGTACT
CAGACAGGCACGGCCGAAGGTTTTGCTAAGGCATTGGCGGAAGAGATAAAGGCTAGGTATGAGAAAGCAGCAGTCAA
GGTTGTGGACTTAGATGATTACGCTATGGATGATGACCAGTACGAAGAGAAATTAAAGAAGGAGACTCTAACGTTTT
TCATGGTAGCAACTTATGGAGACGGTGAGCCAACAGACAATGCTGCAAGGTTTTACAAATGGTTCACAGAAGAGCAC
GAGCGTGGGGTGTGGCTACAGCAGTTGACTTATGGGATCTTCGGGTTAGGAAATAGACAATACGAACATTTTAATAA
GATAGCCAAAGTACTAGACGAGCAATTGAATGAGCAAGGCGCCAAGCGTTTAATTCCAGTAGGACTGGGAGACGATG
ATCAGTGTATTGAGGACGACTTCACAGCGTGGAGGGAACTTCTATGGCCCGAGCTAGACAATTTACTACGTGACGAG
GATGACGTCAACGGGGCAAGCACCCCGTACACAGCGGCAATACCGGAATATCGTGTCGTAATCCATGACGCCAGTGC
TACGAGCTGCGAAGATAAAAGTGTGCTAGAGAATGGGAACACGTCCATCGACATCCACCACCCTTGCCGTGTAAACG
TAGCCGTTCAAAAGGAACTGCACAAGCCTGAAAGCGACAGGAGCTGTATACATTTGGAATTTGACATTAGCGGTACA
GGTATAATTTACGAGACTGGTGATCACGTGGGCGTGTACGCGGAGAATTTTGAGGAGAACGTTGAAGAAGCAGGGAA
ACTTCTTGGCCAGCCTTTGGACCTATTGTTTAGTATTCACGCGGACAACGAGGATGGCGCCCCTTTAGGATCAAGTC
TTGCACCACCTTTCCCCGGCCCCTGTACACTACGTACTGCACTAAGCCATTACGCAGATCTGTTGAACCCACCCCGT
AAGGCGGCACTAATAGCTTTAGCGGCCCACGCCTCAGAACCCTCTGAGGCAGAGAGACTTAAATACCTGAGTTCACC
CGAAGGGAAGGACGAGTATAGTCAATGGATCGTTGGCAGCCAAAGGTCACTGCTGGAAGTGATGGCCGAATTTCCCT
CTGCGCGTCCTCCTCTGGGCGTTTTCTTTGCTGCTATCGCGCCCAGGTTACAACCTAGGTACTACTCTATTTCTAGT
TCTCCTAGGTTTGCACTGTCCAGAGTACATGTGACCTGTGCGCTAGTCTACGGTCCAACTCCAACGGGCAGAATCCA
CAAGGGGGTATGTAGCACATGGATGAAGAATGCGGTGCCGCTAGAGAAGTCACACGACAGTAGTTGGGCGCCCGTGT
TTATTCGTACCTCTAACTTTAAACTTCCAACAGACCCTTCAATACCTATCATTATGGTAGGTCCTGGGACCGGACTT
GCGCCCTTTAGAGGTTTCCTGCAGGAGAGGATGGCGCTTAAAGAGGACGGTGCACAATTGGGGCCAGCTTTGCTGTT
CTTCGGTTGCCGTAATAGGAGAATGGACTTTATCTACGAGGACGAGCTTAACTATTTTGTGGAACAGGGGGTAATTT
CCGAATTAATCGTAGCCTTTTCTAGGGAAGGGCCGCAAAAAGAGTACGTCCAGCATAAGATGATGGACAAAGCAGCC
CAAATCTGGTCATTGATTTCTGAGCGTGGGTACATATATGTGTGTGGCGATGCAAAGGGGATGGCGAGAGATGTTCA
TAGGACCCTACATACGATCGTCCAAGAACAAGGTAATTTGGACAGTTCCAAAACGGAGAGTATGGTCAAGAAATTGC
AAATGGACGGTAGATACCTTAGAGACGTATGG

Seq. ID No: 189
>T5H-CPR_9
TCTTCCTCAAGTAGCTCCACTTCTTTCCACAAGTTAAAAAGAATTTTACACAAACACTTGCAGCGTAGTCATTCCAT
CGGAGCAGAATGTAAACCCCAAAGGTCCAATCACGAGGATCTATTAGCGGTGATGAACAGGAGTTCAATTAAGGTTT
CTATATTTTATGGTAGCCAGACAGGTACGGCGAAGAAGTTTGCAATCAACCTTGGGCATCACCTGCATAACTGTGGT
GTACGTAACTTGGTTATGGACCTGAGACAGACGAATATGGAAATCTTAGTGAACTTATCTATGCTTGATAATTGCGT
AGCATTATTTGTTGTAGCCACTTACGGAGAAGGAGAACCGACGGATAGTGCTAGGCAATTTATGGACAACTTAAAGA
ACTCCTATCAAAAGTTGGATAATCTTCGTTTTGCAGTGTTTGGATTGGGAAATAGTATGTACACGTATTTCAATGCA
GTGGGAAAGTCTATCGACCGTTTGTTGATTCAACACGGAGGAAAGCGTTTGCAAACGTTAACACTGGGCGATGAAGT
AAACGAGTTGGAGTCTACATTCCTAAACTGGCGTTCCCACCTTACCAGTTTATTAATCGATTTCTTTGACTTGAACG
ATCACGATAGAAACTACTTAAATAAACAGTATAAAAGGATGTACTCACTGAAGAGATTTAACTGGAATGTGCCCCTT
GTTTCCCACTTCGTAAACATGTTTATCAACAAGGCACACGTTAAGGAGACCCTTCCTTACGAAAATGACAATTACTT
CTACGCTTCTGTCGCAGTTAATCAAGAGCTTTATCACAAATCCTCGCGATCATGTCGTCATATCGAGCTAGATGTCT
CCGCCTCTCAGCTTAGGTATAAGACTGGGGACCATATTGCTATTTTTGCGTCAAACCCATTGGACCTTGTAGAGAAA
ATAGGAGATCTGCTGAACATCGACCTTAACGAAATGATCTCTTTGGACGCTGTCGATCCAGACTCACTAACAAAGCA

CCCCTTTCCTTGCCCGTGCACGTATAGACACGCATTTATGCATTTTGTCGATATAACGGGTCCTCCGGGTAAGTCCC
TGTTAAGTGCATGTCTGGATAGCGTCACGAACCCCGAGGAGTCCCAGTTTGTGCAATTGCTGATCTCTGACAGTGAG
GACGGCAAGAAACTGTACAGTAAATGGATATTGGAAGACCATAGGGGATTAGTCGATGTACTTCAGGATCTAAAGAG
CTTCAGGCCGCCGGCAGACCTTCTATTGGAATTATTAAATCCGCTGAAACCGAGACTATATTCTATTAGTTCCAGTA
GTCTTGTGCATACGAATCGCATTCATATTACAGCCTCTATTGTCAAGTATAAGACGAACTCTGGTAGAATCTTTAAA
GGGCTGGCCACGAACTGGTTAAAATCCTTGCAGTCAACAAACACGGAGCGTCACCTAAAGATCCCAGTTGCCATACA
TACGTCCAACTTCAATTTACCGCGTTCCAGGACCATTCCAGTGATAATGATAGCAAGCGGTACCGGGTTGGCTCCAT
TTAGAGCCTTTATCCAGGAGAGACTCAAGGTGGCGCACGACAAAGTCGGAAAGACCGGACAGATGGTATTATTCTTT
GGGTGCAGGCACGAGAATAAAGATTTCATTTATTCAGACGAGCTTAAACAAGCGTGCTCTACCGGCCTACTTGAAAT
GTTCACGGCATTCTCCAGGGATTGTCTGGATGGCAATAAGGTCTATGTACAGCATAAAGTTTTGGAGATGGGCAATA
TGGTATGGAAGTTATTAGACGAATGTTACGCCTACATATATGTGTGCGGTGACGCCGCAGGAATGGTTCGTGATGTA
CATCTTTGTTTAATCGAACTTGTTGTTCAGAGGAGTAACCTTACGCGTGAAGCGGCCACGTCCTACGTTTTAAACTT
GCGTAAACAAGGTAGGTATCGTACCGATGTTTGGAAA

Seq. ID No: 190
>T5H-CPR_10
TCAAGCTCCTCTTCCTCCGGGGGCAAAATATTTGATAAATTGAACTCTTCTCTCGATTCTGGGGACAGCACTAGCCC
TGCATCACTGACGGCTTTATTAATGGAAAATAAGGATCTTATGATGATTCTGACAACTAGCGTTGCCGTCCTAATAG
GGTGCGCGGTTGTCCTGATGTGGAGAAGATCCAGCACATCCGCGCGTAAGGTGGTAGAACTTCCCAAACTTGTAGTT
CCCAAGTCTGTTGTAGAGCCTGAAGAAATCGACGATGGCAAGAAAAAGATAGCGATCTTTTTCGGGACTCAGACCGG
CACGGCTGAGGGTTTTGCTAAGGCATTAGCAGAGGAAGCCAAGGCAAGATATGAAAAAGCAATATTTAAAGTCATTG
ACATGGATGACTACGCGGCTGATGACGAAGAATATGAAGAGAAGCTCAAAAAGGAGAAATTAGCATTTTTCTTTCTG
GCGACCTATGGAGACGGCGAGCCCACAGATAATGCGGCCAGGTTCTACAAGTGGTTCGAGGAGGGCAAGGAGAGAGG
AGACTGCTTCAAAAACCTACAATACGGTGTCTTTGGCCTAGGCAACCGTCAATATGAGCATTTTAATAAGATCGCGA
AGGTGGTCGATGAACTGCTTGCTGAACAAGGCGGGCAGCGTCTAGTCCCAGTAGGATTAGGGGATGATGATCAGTGT
ATAGAGGACGATTTTGCTGCTTGGCGTGAGTTGGTCTGGCCCGAGCTAGATAAGTTGCTGCTAGACGGTGATGATGC
CACAGCAACGACCCCCTATACTGCTGCAGTGCTAGAGTACAGGGTGGTCACGTACGATAAGAGCAACTTCGATAACG
ACTTGACTAACACGAACGGCCATGCGAACGGCCACGTCATTGTTGATGCCCAACACCCAGTTCGTGCCAATGTTGCG
GTTCGTAAGGAGCTACACACCCCCGCGTCTGATAGGTCCTGCACCCACTTGGAGTTTGACATAAGTTGTACAGGACT
TACTTATGAGACTGGAGATCATGTTGGCGTGTACTGCGAAAATTTTGTGGAAACGGTTGAAGAGGCGGAAAGACTTC
TGAACATATCACCTGACACATTCTTCTCCATACACACCGACAAGGAGGATGGAACGCCTTTGGGAGGGAGTAGTTTA
CCTTCTCCATTTCCACCCTGTACACTAAGAACGGCGTTGACTAGATATGCAGACGTCTTGAGTTCACCAAAGAAGTC
CTCCCTGCTTGCCCTGGCGGCCTGTTCATCAGATCCCAATGAGGCAGATAGACTGAGATACTTGGCCTCTCCCGCGG
GAAAAGAGGAGTATGCGCAATGGATAGTTGCCAGTCAGAGAAGTTTGTTGGAGGTTATGGCGGAATTTCCGTCAGCT
AAACCGTCCATAGGGGTGTTTTTCGCAAGCGTTGCACCTAGATTACAGCCGCGTTTCTACTCAATCTCTAGTTCTCC
GAGAATGGCTGCATCCCGTATACACGTAACTTGCGCTTTAGTATACGATAAGATGCCGACGGGACGTATCCATAAAG
GGGTTTGCAGCACCTGGATGAAAAACGCCATACCGTTGGAAGAAAGTCTAAGTTGTAGCACCGCCCCGATCTTTGTT
AGACAGTCAAATTTTAAACTTCCAGCCGACAACAAGGTTCCTATCATTATGATTGGTCCTGGCACGGGATTGGCGCC
GTTCAGAGGCTTTTTGCAGGAGAGGATGGCTCTGAAGGAAGAAGGTGCGGACCTGGGGCCCGCCGTGTTGTTCTTTG
GGTGCCGTAACAGACAAATGGATTACATATATCAGGATGAATTAGATAACTTCCTTGAAGCCGGGGCGTTAAGTAAT
CTAGTCGTCGCTTTTTCAAGAGAAGGACCCAACAAGGAATACGTCCAACACAAGATGACACAAAAGGCAGATGATAT
TTGGAATATGATTTCTCAAGGGGGTTACGTTTATGTCTGCGGTGACGCAAAAGGCATGGCCAGAGACGTACACAGAA
CACTTCATACTATCGCGCAGGATCAGGGGAGCCTAGACTCCTCTAAGGCAGAATCATTCGTCAAGAATCTGCAGACA
ACCGGTAGGTACCTAAGGGACGTGTGG

Seq. ID No: 191
>T5H-CPR_11
AGTTCATCTTCTTCCAGCGGGGGAGACGGTGCCGAAGGCAGGGCTCTGGTAGCCACTCTAGCTGCGGCGGTGCTTGG
AGCGGCTTTATTTGTCTTATGGCGTAGGGCCGCTGCAGGCAAGAAGAGGAAAAGGGAGGCCGCGGCTGCTGCAGTGG
CTGAGGCGACAGAGGTCAAAGCAAGAGCGGCCAAAGGTGGCGAAGATGAGAAAGCCGCCGATGACGGACGTAAGAAG
GTAACAGTTTTCTTCGGTACGCAAACTGGAACAGCCGAAGGTTTTGCGAAGGCCCTGGCAGAGGAGGCTAAAGCGCG

TTACGACAAGGCAATATTTAAGGTTGTAGACTTAGATGATTACGCTGCCGAAGATGAGGAATACGAAGAGAAATTGA
AGAAGGAAAAGCTGGCTCTGTTCTTTGTCGCTACGTACGGGGATGGGGAACCTACTGACAATGCCGCAAGATTTTAT
AAATGGTTCACGGAAGGTAATGAAAGGGGTGTCTGGCTGAACGATTTTGAGTATGCGGTGTTCGGTCTAGGCAACCG
TCAATACGAGCATTTCAACAAAGTAGCGAAAGTTGTTGATGAAATTCTTACAGAACAAGGGGGCAAGAGACTGGTAC
CTGTCGGACTTGGTGACGATGATCAATGCATCGAGGATGATTTTAATGCATGGAAAGAGGCGCTGTGGCCCGAGCTG
GATAGACTTTTGCGTGATGAGAACGATGCCAGCACGGGGACAACTTACACGGCAGCAATCCCGGAGTACCGTGTCGA
GTTTATAAAGCCTGAAGAGGCGGCTCACCTGGAGAGAAACTTCAGCCTTGCTAACGGTCATGCCGTGCATGATGCTC
AGCATCCGTGCCAGGCTAACGTGGCGGTTAGACGTGAGTTACATACGCCGGCGTCTGACAGGTCATGCACTCATTTG
GAGTTCGACATTGCAGGCACCGGACTTACCTACGAAACGGGGGACCACGTTGGAGTGTACACAGAGAATTGCCCGGA
GGTGGTTGAGGAAGCCGAGAGGTTGTTGGGTTACTCCCCGGACACTTTCTTCACAATTCACGCGGATAAAGAGGATG
GTACTCCCTTGTCCGGCTCATCCCTAGCGCCACCATTTCCGTCTCCCATCACGGTAAGGAACGCCTTAGCCAGATAT
GCCGATTTACTGAACTCTCCCAAAAAGACCAGCCTGGTAGCCCTAGCGACATACGCCAGTGATCCCGCTGAGGCGGA
CAGGTTAAGGTTCCTGGCTTCTGCAGCGGGGAAAGATGAATATGCTCAATGGGTCGTGGCGTCACAAAGAAGTTTAC
TAGAGGTCATGGCGGAATTTCCATCTGCCAAACCGCCACTTGGCGTGTTTTTCGCCGCCGTGGCGCCAAGGTTACAA
CCCAGATATTATTCAATTTCCTCATCCCCATCGATGGCTGCCACGAGGATTCACGTAACTTGTGCCTTGGTTCATGA
AACCACTCCAGCCGGGCGTGTACACAAGGGAGTTTGCTCAACATGGATCAAAAATGCAGTCCCCTCTGAGGAGAGCA
AAGATTGTAGCTGGGCGCCGATTTTCGTAAGGCAAAGCAATTTCAAATTACCGGCCGATCCTTCAGTTCCCATAATA
ATGATCGGGCCGGGAACAGGCCTTGCACCGTTTCGTGGCTTTTTGCAGGAACGTCTGGCCCAAAAGGAGTCTGGAGC
AGAATTGGGCCCAAGCGTCTTTTTCTTTGGTTGTAGAAATTCAAAAATGGATTTCATCTATGAAGATGAACTTAACA
ATTTCTTGGAACAGGGTGCCCTATCTGAACTTGTTTTAGCGTTCTCCCGTCAAGGGCCAACTAAGGAGTATGTTCAG
CACAAAATGGCTCAGAAGGCGTCTGAAATATGGGACATGATTAGTCAGGGGGCGTACATTTACGTTTGCGGGGATGC
GAAGGGCATGGCCAGAGACGTGCACAGAGTCCTACATACAATCGTCCAAGAACAAGGATCACTGGACAGTTCCAAGG
CCGAGAGCTTCGTGAAGAACCTGCAGATGGAAGGACGTTATCTACGTGACGTGTGG

Seq. ID No: 192
>T5H-CPR_12
TCCTCATCTTCTTCCTCAGTAAGAGAATCTAGTTTCATAGAGAAAATGAAAAAGACCGGTAAGAATATCGTAGTGTT
CTACGGATCACAAACTGGTACCGGTGAAGAGTTTGCGAATAGACTTGCAAAAGACGCTCACCGTTATGGAATGAGAG
GGATGGCTGCTGATCCCGAAGAGTTCGAAATGACGGACTTGTCCAGATTAACAGAGATCGAAAATGCACTAGCAGTT
TTCTGCATGGCTACCTACGGTGAAGGTGACCCCACTGATAATGCGCAAGACTTCTATGATTGGCTGCAGGAGACAGA
TATCGATCTGGCCGGGCTAAAGTATGCCGTGTTTGGCCTGGGAAACAAAACGTACGAACATTTTAACGCTATGGGCA
AGTATGTGGATAAGAGATTAGAGGAGTTAGGGGCAGAAAGAATTTTTGAGTTAGGGATGGGGGACGATGACGGTAAC
TTGGAAGAAGACTTTATCACGTGGAGAGAACAGTTCTGGCCGGCGGTTTGTGAACACTTTGGCGTCGAAGCGACAGG
AGAAGACAGTAGCATTAGACAATACGAATTGGTCGTACATACGGATGAAAATATGAACAAGGTATATACAGGCGAAA
TGGGGAGACTAAAGAGCTACGAGACCCAGAAACCTCCTTTTGACGCCAAGAACCCCTTCCTTGCGAACGCCACTGTC
AACAGAAAGCTGAATGAAGGGGGAGATAGACATTTTATGCATCTTGAACTAGATATTACGGGCTCAAAGATTCGTTA
TGAGAGTGGCGATCACGTCGCAGTCTATCCTGCTAACGACGCGGCACTGGTTAATAAATTAGGAGAAATTCTTGGAG
CCGACTTAGAGACAGTTATTTCTCTGAACAATCTTGATGAAGAGTCTAATAAGAAACATCCCTTCCCTTGTCCGACT
ACCTACAGGACGGCACTAACGTACTATCTGGATATCACGAATCCCCCGCGTACAAATGTTTTATACGAGTTGGCACA
ATACGCCACCGATTCAAAGGAGCAAGAGAACCTACGTAAAATGGCGTCATCTGCGCAAGATGGTAAAGCACTGTACC
TGTCTTGGGTGGTAGAATCCAGGCGTAACATATTAGCTATATTAGAAGACATACCATCACTGAGACCGCCTCTGGAC
CACTTGTGTGAACTTTTACCCAGGTTACAGGCGAGATATTACTCTATAGCAAGCTCCAGCAAAGTCCACCCTAATAG
TATTCACGTTTGCGCAGTCCTGGTGGAATATGAAACCAAAACTGGCCGTGAAAACAAAGGCGTCGCCACAAATTGGC
TGAAGAATAAGCAACCTAGCGATAACGGTCATAAATCCTCAGTACCAATGTTTGTAAGAAAGAGTCAATTTAGGTTA
CCATTTAAGCCATCCACACCCGTTATAATGATCGGTCCGGGTACCGGCATTGCGCCCTTTATGGGATTTATACAAGA
ACGTGAATGGTTAAAGCAGCAAGGTAAAGACGTCGGCGAGACGGTACTATACTATGGTTGCAGGCATGAACATGAGG
ATTTCTTGTACATAAACGAACTGAAAAGGTATCACAAAGAGGGAGTTCTAACGCAGTTGAACGTCGCGTTCTCTAGG
GATCAAGCACACAAAGTGTACGTACAACACTTGCTAAAGAACAATAAGGAAATGGTTTGGAAGTTAATCCACGAGGA
TAACGCTCATATTTATGTCTGCGGTGACGCCAGGAACATGGCCAGAGACGTACAAAATATCTTTTATGACATCGTTG
AGGAATACGGCAAGCTAGATCACGCCCAAGCCGTAGACTACATAAAGAAATTAATGACTAAGGGGAGGTATTCACAG
GATGTGTGGTCT

Seq. ID No: 193
>T5H_1
ATGCTACCGATCGTGGATCACCTACTTGACGTTTTAAACTTGGAACGTACACCCTTTCGTACATATGCTGTGACGGC
ATTACTGCTTCTTTTCGTAGGCATCATTGCCAGAGCCCTACTTAAAATGATGCTATTTATTCAGGAGTACAGTGCTA
ATTCAAAGCGTTTGAGGTGTTTCCCCGAGCCCCCGAATAGGTCCTGGATCTTGGGGCACCTGGGTCTTTTTGCCCCG
AACGAGGAGGGAATGACAGAATTTAGTAAACAGGTGAGCAAATTTACCTACTACATGAAGACATGGATGGGACCAGT
AATACCGCTGATTAGTCTAATACATCCCGATACTATTAAACCAGTAGTTGCGGCACCAGCTTCTATCGCGCCTAAAG
ACGCACTTTTCTACGGATTTTTAGAACCGTGGCTGGGCGACGGTCTACTTCTAAGCAGGGGCGAGAAGTGGGTTCGT
CATAGACGTCTACTGACGCCCGCTTTCCATTTCGACATCCTGAAGCATTATGTCAAAATATTTAACCAGTCCACAGA
TATTATGCATGCAAAGTGGAGACGTTTGTGTACAAAGGGGCCTGTCTTCCTTGATATGTTCGAGCACATAAGTCTAA
TGACCCTGGACTCCTTATTAAAATGCACGTTCTCCTATGATTCTGACTGTCAGGAAAAGCCTTCAGATTATATCGCG
GCGATATACGACTTGTCAGAGCTTATAGTGGAAAGAGAGCAATGCCCGCCTCATCACTTCGATTTCATCTATAGATT
TTCTTCTAATGGGAGGAAGTTCCAGAGGGCCTGTCGTATCGTCCACGAGTTCACTGCAAATGTCGTCCAGCAGAGAA
AGAAAGCGTTGCAAGAGAAAGGAGCGGAAAACTGGATAAGATCCAAAAAGGGCAAAACTCAAGACTTTATTGACATC
CTTCTTTTAAGTAAGGATGAAGACGGTAATACTCTGTCTGACCAGGAGATGCGTGATGAAGTCGATACGTTCATGTT
CGAAGGGCACGATACTACCGCTTCCGGTCTGTCATGGATACTTTATAACTTGGCATCACATCCCGAATATCAGGAGA
AATGTAGGGAAGAGGTGACACAACTGCTGAAAGGTGAGTCTACACACCTTGAGTGGGATGACTTATCCCTTCTTCCG
TTCACCACTATGTGCATAAAGGAAAGTCTAAGGCTGCACCCGCCGGTAACGGCAGTGAGTAGAAGATGCACTGAGGA
CATCGCAATGCCAGACGGGAAGGTAATTCCGAAAGGCAATATCTCTCTTATATCTATTTACGGTACTCACCACAATC
CAGCGGTTTGGCCTAACCCGGAAGTCTACGATCCTTACCGTTTCGATCCGAGTAGCACAGATGAAAGATCCTCCCAC
GCCTTCGTTCCTTTCAGCGCGGGCCCTAGAAACTGTATCGGGCAAAATTTCGCTATGGCAGAGATGAAAGTCGTCCT
AGCCCTAACTCTTCTTAACTTTAAGGTTGCGTTAGACCCAAACAGGGTGGTCCGTCGTAAGCCGGAGTTAGTTCTTA
GAGCGGAAGGTGGACTATGGTTGCAAGTCGAAGCCCTTAAAAGCAAGTCA

Seq. ID No: 194
>T5H_2
ATGGAGCTGTTGGGCTTGGTCTCTTGGCTTCTGCTTCTTCTACTGACTTTGGTGGTTATCTGCTTTCTACTGTACTG
TGGCTACATCCATTACCAGCACATGAAATATGACCACATACCCGGACCGCCCAGAGAGTCATTCTTGTTTGGACACG
GTTCAGCTATCTGGAAAGTGATGAGAAAGAACCAATTAGTTTATGATTTATTTCTAAATTGGGTCGAAACATATGGG
CCTGTTATACGTATAAATGCACTGCATAAAGTAACCATCGTTAGCGTCTCACCAGAGAGTGTCAAGGAAGTCTTAAT
GTCTCCAAAATACCGTAAGGACTGGTTCTACGACCATTTGCACAGTCTATTCGGAGTCCGTCTAATGGGAAATGGTT
TAGTGACAGATAGGGACAATGATCACTGGTATAAGCAGAGAAGAATCATGGACCCAGCATTCAGCAGAACGTACTTG
ATTGGGCTACTAGGGCCGTTTAACGAGAAGGCTGAAGAACTGATGGAGCGTTTAGCAGAGGAAGCTGATGGGAGGAG
TCATGTAGTTATGCACGCCATGATGAGTAGAGTGACTTTGGACGTAATTAGTAAAGTGGCCTTTGGAATGGAGATGA
ACTCCTTGAAGGATGATGGGACGCCACTTCCCAGGGCGATCAGTCTTGTAATGAGGGCTCTTGTGGAGATGAGAAAT
CCATTCATCAGGTACAGCAGAGAAAAACAAGCATTCATACGTGATGTTCAAGAGTCTGCCCGTTTATTGCGTAAAAC
TGGTCGTGAATGCATAGAGCGTCGTCAAAAGGCGATACAGGACGGTGAGGAGATACCTGTTGACATATTGACCCAAA
TCTTGAAAGGCGCCGCGTTGGAAGGGGACTGTGACATGGAAGATCTGCTGGATAATTTTGTGACGTTCTTTATTGCC
GGCCAGGAAACAACCGCGAACCAACTTGCGTTTACTATAATGGAGTTAGCGCGTAACCCCGAGATTCTGGAAAAGGC
TCAGGCAGAAGTAGACGAAGTCATTGGAGTCAAGAGGGATATAGAATACGACGACCTAGGCAAGTTGCAATACTTAT
CCCAGGTCCTGAAAGAGTCTCTTAGGCTGTATCCGACTGCTCCTGGGACAAGTAGGGCAATAGAGGAAGAGACGATA
ATAGAGGGTTTTCGTATTCCCCCTAAAGTCCCGTTAATGTTCAACAGTTACATCATGGGGAGGATGCAACAGTTTTA
TCCCGATCCCCTTACATTCAACCCGGATAGATTCCATCCAGATGCTCCCAAACCGTACTATAGCTACTTCCCGTTTT
CCTTGGGTCCGAGAAGTTGTATCGGACAGGTCTTCGCACAAATGGAAGCTAAAGTCATCATGGCCAAACTATTGCAA
AGATTTCAATTCGAGCTAGTAGAGGGGCAGAGCTTCGGTATCATGGATACCGCCAGCCTTCGTCCGGAGGGCGGGGT
TATTTGTCGTCTTACGATACGTACGAATCCGGGAAAGGCCAAGAAAGACGAT

Seq. ID No: 195
>T5H_3

ATGAGCAGGCCTCAAGTCCCAAAGGGGCTAAAGAACCCTCCAGGCCCGTGGGGCTGGCCCTTGATCGGCCATATGCT
AACCTTAGGGAAAAATCCACACTTAGCTCTTAGTAGAATGTCACAACAATATGGCGACGTTTTACAGATCCGTATTG
GGAGCACACCGGTAGTTGTACTTAGTGGCCTGGATACCATAAGGCAGGCGTTGGTTAGACAAGGAGACGACTTTAAG
GGGAGACCTGATCTGTACACGTTTACTTTAATTTCCAATGGACAATCAATGAGTTTTAGCCCCGATTCCGGTCCTGT
ATGGGCCGCCAGGAGACGTCTTGCTCAAAATGGACTAAAGTCTTTTTCAATTGCTAGTGATCCCGCCTCTTCTACGT
CTTGTTATCTGGAGGAGCATGTTTCAAAAGAAGCCGAGGTTTTGATCTCAACTCTTCAAGAATTGATGGCCGGACCA
GGACATTTCAACCCGTATAGGTACGTGGTAGTATCTGTAACGAACGTCATTTGTGCGATATGCTTCGGGAGGAGGTA
CGATCATAACCACCAAGAACTACTATCTTTAGTTAATTTGAACAATAATTTTGGAGAAGTTGTCGGGTCTGGGAACC
CGGCGGACTTTATACCTATTCTAAGATATCTACCGAATCCCTCTCTTAATGCCTTCAAAGACTTGAACGAGAAATTC
TATTCATTTATGCAGAAGATGGTGAAAGAGCACTATAAGACGTTCGAGAAGGGGCACATTAGAGACATTACAGATTC
TCTAATAGAGCACTGCCAAGAGAAGCAGCTGGATGAGAATGCAAACGTTCAACTATCTGACGAGAAGATTATTAATA
TCGTGTTAGACCTATTTGGCGCTGGCTTTGATACGGTCACCACAGCTATTTCTTGGTCCCTGATGTATTTAGTGATG
AACCCTAGAGTTCAGAGGAAGATACAAGAAGAGCTGGATACCGTAATCGGGAGATCCAGACGTCCGCGTCTGAGTGA
CAGAAGTCATTTACCCTACATGGAGGCCTTTATCTTAGAGACTTTTCGTCACAGCTCCTTAGTGCCTTTCACTATCC
CACACAGTACGACGAGGGATACATCTTTGAAAGGTTTCTACATCCCAAAAGGGAGATGCGTATTTGTCAATCAGTGG
CAAATAAACCACGACCAGAAACTATGGGTTAATCCAAGCGAATTTCTGCCTGAGCGTTTCCTTACTCCCGATGGCGC
AATAGATAAAGTGTTATCCGAGAAGGTCATCATCTTCGGTATGGGGAAAAGAAAGTGTATTGGGGAAACAGTAGCGA
GGTGGGAAGTATTCCTTTTTCTGGCCATTTTATTACAGCGTGTAGAATTTAGTGTACCACTGGGGGTTAAGGTAGAT
ATGACTCCCATATATGGGCTTACGATGAAACACGCGTGCTGTGAACACTTCCAAATGCAGTTGCGTTCC

Seq. ID No: 196
>T5H_4
ATGCCGACTCCTGGAGGCAGATTGGTTGCCTTCCTGCAGAGGAGGGGCAAGCTGGCAGGTAGTCTTGCTGTGATACT
ACTGTTAATACTGAAGAGATTGAGGGACGCTCCAAGAAAGGTGAGGTGGTTGAGAGGACCGCCACTACTTGGGGTCG
TGCTAAAAGTTTTTCAGGGACTGCGTGAGCACGCCTTGCTAGATATGTATGACAGGTGGCATCAAAGGCTTGGACCC
ACTTTTGCATATTGTGCCCCCGGAAAAATGGTTGTAGCAACAATCGATCCCAAAAACATAGAGCACGTTTTGAAGAC
GAAATTTGATAATTACGTCAAGGGGCACGTGTTTGCCGAACCGTTTACGGATCTACTGGGCGATGGTATATTTAACG
CAGACGGGGAGATGTGGCATAGACAAAGGAAAACAGCCAGCAGGATGTTTACCAAGCGTCAATTCGAGACACATATA
TGGAAAGCAATAGAAGCGAACACGGCCAAAGTAGGTAGGATATTGGAACGTTCCGAGGGAACCCTTGATATGTTCAA
TCTTATGAACAGGTTTACCCTAGATACGATCGGGCGTATTGGATTTAGCAAAGATATTGGTTCTTTAGAGGACCCGT
CATCACCCTTCCTGAGATCCTTCGACAGAGCACAACAAATTCTGATCTTACGCTTTTGGACTAATCCCGCTTGGAAA
GTTCTAAGGTGGTTAGGGGTGGGCTGGGAACGTGAATTGAAGGAACACCTAGGTAGACTAGATGGGTATGCACGTGG
CATAGTCAGGGAACTTAGGCAAAAAGCGGAAGCTGGACAGGATGATTCCTTTGTAGGGTTGTTCATGAAGGAAGAGC
AAGCGGCTCCTGCAGCTAGATCCCCCGAGCTACAAGAGAAGTTTATGAGGGACTTAGTACTGAATTTCTTGATAGCA
GGCAGGGACACTACTGCACAATGTATTAGTTGGACTCTTTTCGAACTGACTCAGCACCCTGCCGTTGCGGCGAAAGC
AAGGCAGGAGGTTTTAGATGTATGCGGCGAGGGCCCCGTCACTTTCGAGCATCTAAAATCCTTACAATACGTGAGAG
CCATTTTAGACGAAGGCCTGAGACTTCACCCGTCTGTTCCTTACGACGGCAAACTATGTCTAGGCAAGGACACGTTG
CCTGACGGCACTGTCGTGCCGGCCGGCTGCATAATTCAATACATCCCTTACGCTCAGGGCCGTTGTAAAGACATATG
GGGTGAGGACGCTTGTTCCTTCAGACCTGAAAGGTGGCTAGAGATGCCCCGTAGACCATCCTCTTTCGCCTTTGCGG
CCTTCAATGCTGGTCCCAGGGAGTGTCTTGGACGTAGGCTTGCGGAAGCGGAGATGGCGGCTCTTGTTTCCACAGTC
GTTCGTGATTTCGACATGCGTTTAGAGGTAGAGCCATCTAGTGTGCGTTATGATGCGCAGTTAACATTAGGGATGTG
TGGGTTGCCCGTTAGCGTGAGGCGTTGTCGTAGGGCCTATGGTGTTGCTGAACCACTTGCCGGAGCG

Seq. ID No: 197
>T5H_5
ATGCTACCGTTGAGACATAAGATGCTTACGGGAGAGGCCGAACCCTGTCTAGTCTCAAAGACTGCAGAGACTGACGC
TGAGTGGACCAGGGACGCGTTTGGTATGGGGCAATACACAGCCGGACGTTGTGACCACTTATTATCCTGGGTGGTCT
TCTTACTTCTGGCCCCTGTACTATTGATAGTGTGGTTACCTCTATCTTGTATCTGCTGTGCCTCTCCGGTATTATTA
GTCCAGAGATTCGCAGGATGGGTTCTTTCCGGGTGTTTGGCAAGGACATATTTGGGAGTCCTGTTGATCAGATTGTG
CGGGAAGTGCGACCTAATCCTTACAGGTATGCATTTTATTAGGACGGGAAGTCAAAGGTTCTGGATGGATACACTTG
ACCCTCAGGATTGGGCGTATCACAACGAGACGTATGGGAGAAACATCATCCTATGGGCGAATCTGCGTGTAGGATCA

TACAAGCAAGTTAGGGATATTGTTCTTAACCCCGCGCGTAAAAGAACAAGGGCGCTGGACGGCTGGATATCAGGGTT
CGCGAGGCACTACCCCAATCTGCCCGTGTTTTTCAACACTGGCTCAAACATGCATACTACGTTCAGGCAGATCTTTT
TCGCCAACTTCACCAAAACGGATTTCGTACTTAGGGCGTTGGAAGACGAAGGAGCGGGCTTGGCGAAAATGGCGGCG
CCTATTTTGCAAAGGTGGCTTGCAGGGAGCTTTCGTGAAAGCAAGTCAGGTGAGGGGAATTTATATATGGTGGAACC
AGTGGCACCTCTGATTTTGTTTCTGCTGTTCGAAGTAGAAGTGGAGTCCATTCCGCCCGAATTGTTGACCGCCTTCT
CAGACGTAGTAACTGTCGGTGCGTCATACTTTTTGCTGCCCCCACACTCACCGTATTGGCTACTGAGTGGAAAGGTG
AAAGCTATTGCTCTGCTTAAAGATTTCCTACTAGAACACTGTAACGCGGCGAGACCCGAGTCACTGAAGGGCCGTGC
TGTCGATTGGAGGAGCTTGGCGGCACAAATGCCCGCTTTTCTTCCTAAGGACGAGTGTCGTTGTCCATGTAGTGGGA
CCCCGGCTGTAGATCCCGTCGATGCGTATTTGGAAGTCATATCAGTAATGGTTTGCGTTGCGGGTGTGACTGGCACT
ACTAATGGGTTCACATCAGTCATAAGAAAATTCGCCGACGTACCTGTAGGACCTACAAAGTCCCGTTGGCCTAGTGC
GCCTGTACAATGGCGTCCAGACGCTGATGATATGGTGAGACTTTACCGTAGGGACCCACTGGGTTTTATCTTGGAGG
CCCTTCGTCTTGGTACACCTGTTGCAGGGACCCACCAGGTCTTAGAGGAGGAATTAACTTGTCCTTTCCTGCACAAG
GAGACAACGTTCCCAAAGGGGACGGTCGTATGCGCAAATCTAAATGCCTGCCACACTGATCCCGAGGAGTGGGGATC
AGATGCTCTGGAATTTAGACCAGGAAGGGCTGCAAGGAACCGTTACCTGATGTGGAACGGACCCTTCGGTGAGGCCG
CGCCGAGGCAGTGTCCCGGCGAACAGGTCGCGGCACATTGCATCAAAGTTTCTATCGATGCGTTCTTAGACATGCAC
AAGCCTCAA

Seq. ID No: 198
>T5H_6
ATGGCTACGTCTATCCTATCCCTGAGCCTTATGGACTTGTTATATTGGGGTGCGTGTTTGTGCGTACTAAGTGTTTT
GTACAAGATCTCAGCCCTATACCTTAGACAAAAGAATTTTGAAAGGGTTTTCTCAGCATTTCCCGGACCGAAGAGGC
ACTGGTTGTATGGCAACGCTCACGAGTTCAAGCAAGACGGAACAGACTTAGATATCCTGAACGGGTACGCAAAGCAG
TTTGATTGCGCCTTTCCATTGTGGCTTGGAAACTTTTTCGCCTCCCTAGCTATATACCACCCAGACTATATAAAAGC
TATCCTTAGTCGTCAAGACCCCAAGGATAACTTCGTTTATCACTTCATCACCCCGTGGATCGGCAAAGGACTACTTG
TCCTAAGTGGCCAAAAATGGTATCAACATAGGAAACTTTTAACTCCCGGGTTTCACTATGACGTCTTAAAACCTTAT
GTTGGCGTTATGAGTGATTGTGTTAACGTTATGTTGGATAAGTGGGAGCGTCTAGTCCCCGACAAGAAGCCAGTGGA
ACTGTTTCACTATATTTCCCTAATGACGCTAGATACTATAATGAAATGCGCCTTCAGTTACCAAAGCAACTGTCAAA
ACGACTCCGAAAATGAGTACATTAAGGCAGTTTATGAGTTGTCCTACTTGGTGGATCACCGTACTAGATGCCCTCCA
TACCACAATGACTTTATTTTCTACTGGAGCCCGCACGGATTCAGACTGCGTCGTGCTTTGAAGACGGCTCATCAACA
CACGGAGAAGGTAATAAAGTTGAGAAAAGAATCTCTTAAACAGGAGACGGAGCTTGAGAAGATAAAACAAAAGAGGA
GGCTTGATTTCTTGGATATTCTGTTGTGTGCGAGAGACGAGAACGGACAAGGCCTTAGTGATGAAGATTTGCGTGCG
GAGGTCGATACCTTTATGTTCGAAGGGCATGACACTACTGCTTCTGGGGTCTCCTGGACGTTTTATTGCTTGGCAAA
GAACCCAGAACATCAGGAAAAATGTCGTGAGGAAATACGTCAGGTCCTAGGCGACCGTAGAACCGTAGAGTGGGAGG
ACTTGTCTAAGTTACCCTACACAACCATGTGCATCAAGGAGTCCATGCGTTTATATCCTCCGGTACCTGAGGTGGCA
AGAGAATTAAAAGAGCCCATTACTTTCTGTGATGGCAGGAGCGTGCCAAAAGGTAGTATCGTTTTCTTATGTATCTA
TGCTATCAATAGATGCCCTGGCATCTGGGAAGATCCTGAGGTTTTCGATCCGTTACGTTTTAGCCCCGAGAACTCTT
CAACGAGACATTCCCATGCTTTTCTACCTTTTTCTGCGGGAGGCCGTAACTGCATAGGTCAGAATTTCGCGATGAAC
GAGATGAAGATAGCCACTGCACTGACCCTACAGAGGTTCGAATTGCAATTGGAGACCAAAAGAGAGCCAGTCAAACG
TGCTCAGTTAGTGTTAAGAAGCATGAATGGAATCTATATCAACCTTAAAAAGATTCATAGCGACAAGACTAAGATTA
TC

Seq. ID No: 199
>T5H_7
ATGGGTTTATGGACCTTTATGACAGGTGCCTTAATATTATTACTAATTCTAGTGGTACTATGTTTTCTGCTGTACTG
CGGTTACATATATTACATGCATATGAAATATGACCATATCCCCGGTCCTCCGAGAGATTCCTTCTTCTTTGGTCACA
GTCCGACCATTATGAAATTAATGAGAAATAACGTGATTATGTACGATACGTTTTTAGAATGGGTTAAGACATACGGA
CCGGTCGTTCGTGTAAATCTATCATGTAGTACCATTGTTTTTGTCATCAGTCAAGAAGCCGTGAAAGAGTTTCTGAT
GTCTCCCAAATACACGAAAGATAACTTTTACGAATGTGTAGAAACACTTTTCGGTGTCAGGTATATGGGTAAGGGCC
TTTTGACTGACAGGGACTACGAGCACTGGCACAAGCAGCGTAGGATCATGGACCCTGCATTTAGTCGTAATTATCTG
ATAGGTCTAATGGGGACATTCAACGAAATTGCAGAAGACCTGGTGGACATCTTGGGCGACAAGGCAGACGGCAAGTG
TCAGGTTGGTATGCATGATATGATGGGAAGAGTCACGTTAGATATTATCGCCAAAGTAGCATTTGGAATGGAGCTTA

ACTCCTTACACGACGATCAGACGCCGTTCACCCGTGCAATCACTACAGTTATGAGGGGAATGGTAGAAACGAGAAAC
CCTTTGGCCCGTTATATCCCTGGTAAACAGGCACTTATACGTGATATCAAGGAGTCACTGAAATTTTTACGTAAAAC
GGGGAGGGAGTGTATTCTACAGAGAAGAAAGGCAATCCAAGACGGAGAAGACATACCAAGAGATATCCTGACACAAA
TCCTAAAAGGCGCTGAGACCGAGGGAGACTGCTCCTTGGAAAACCTGATAGACAACTTTGTAACTTTCTTCATTGCC
GGCCAGGAGACGACAGCCAATCAATTGAGCTTCGCAGTTATGGAGTTAGGTAGACACCCTGAGATTTTGACACGTGT
GCAGGCGGAAGTAGATGAGGTATTAGGTAGCAAGAGGGATATTGAATACGAAGACCTAGGTAAGTTGCAGTATCTGT
CACAGGTATTGAAGGAGACCTTGAGACTTTATCCCATCGCGCCCGGGACCTCCCGTGCCCTAGAGAAAGAAATGGTT
ATTGAGGGCGTTCGTGTTCCACCCGGGACTACGTTAATGTTTAATGCATATATAATGGGAAGGATGGAGAAATACTA
CCATGACCCTCTAGTCTTTAATCCCGATAGGTTTCATCCAGATGCCCCGAAAATATCTTATGCTTATCTACCATTCA
GCCTAGGACCCAGATCTTGTATTGGTCAAGTTTTTGCACAGATGGAGGCGAAAGTAGTTATGGCAAAACTATTACAG
AGGTTTGAGTTCGAGTTAGTCGAGGGACAGAGCTTCAGAATATTGGACACGGGTACACTTCGTCCCCTTGATGGTGT
CATCTGTAGGCTAAGACCGAGAGCAGAGCATAAGTCCAGGAAG

Seq. ID No: 200
>T5H_8
ATGTGGACAATCCTTCTTTCTACCATCAACATAACGCTGGCAACAGCTCTAATGTTATCCTTCATAATAATCTACCT
TCTATATATCCAAAATTCAACAAAACTGCCCCCTGGCCCCACCTCTTGGCCTTTAATTGGCTATACTAGCTGCCTTG
GCACTGATGCGTTCAGGAAGATTCAAGATCTTAACAAGATATATGGCGACATAGTCAGCTTCCAAGTATTAGGGAAG
ACTATAATTATATTATACAATTATGATTTAATCCATGAAGCTGCCAATGGCAATAGGTCTAAGGTGGGGCGTTACAC
CATGACCGTGAATGATTTGCTGGCGGAAAATAGCGGCATCTCTAATTATGACACGCAGAAGGCACTTGAAATGCGTA
AAGCGTTCGTTAGGCTTGTCCACAACAATATCAAGACTACGGAGGAACATGAGGGCAACAAGCTACAACCGTTTATT
AGCCAAAATATAATAAACGCACAGATCAATAAGCTGATTAGGCAGCTGAGAATCAGGCAGGGCAAGCCGGTTAACGT
TTTACAGCTTATGAGGTGCACCGTCTGGAGAATCATTTGGAATTTGATATTTGGCAAGGAGTGCCAACTGACTGATA
AGCAAATCTCAGACACACTGGACGATATTTCCAGCAATAACCTGCAGAACCAGCTATTTCAAATCAGGCAGTTGCTG
CCAAGGTTTTGTGTTAACATATTCAAGCACAGCCAGTTCGCCCGTAAGCTATTCGAGATCGAGGAAATAATCTACAA
GTATAAAACTGTGAGACAATTAATTGATAATAATGTTGGAGAGATGCATAACTCCGATAGCTTGCTGGGTCAATTAA
TAAACGATTTGAAACTGAATCTAACAAAGAATGACATCTCCCGTCTAAGTTTTGAATTTATGGCAGCTGGAACCGAT
ACTACCTCTCTAACCTTGACGTGGGCATGTGACTATTTAGCAAGAGCCCCGCCAAAGGAGTCACTGAAATTATCCTC
CGACCTGATTGATATGATCCATAGATGGGCCTCAGTTGTGCCTCTTTCTTTGCCTCACATAGTTCGTGAAAGTTTCA
AGCTAAAGAACTATTATATCCCGAAATCAAGCATATTAATATATAACTTGTACGCCGTCCACAATTCCCAAATTAAG
AAGCTAATCAATACGGAGCAAAATTCTGATGAGATTCAGGAATCCGATAAGCCAATCCCCTTCTCTCTTGGTAGTAG
ATCATGTCCGGGGGCTCGTATAGCAAATTTATTGATCGAACAAATTCTGACAGCCATTAACCAGGAGTTCCTTATTC
AAAATATAACGCAGAGCCCGTTTGAGACCATTTCCCCAGGAAACCAGGAAAGTTTAACGCCTTTCGGGATTACTAGG
ACTCCACATAAAAGTATGTATATATTCGTAACGAAACTAAATGGAAATCGTAGGACTTCAATT

Seq. ID No: 201
>T5H_9
ATGTCCCAGTTGCTATCGAGCTTGATAGAGCTACCTACTCAGACGCTGGTGCTTGCAACAGCAGTGGCTGTAGGGGC
AGCCGCCTTATTAGTCCACGCGTATCTTTTCGACGCCGTTGGGAAGCATGGAAACCTACCGCCGGGCCCACCAGTTG
ATTCACTATTTTCTGGACATAGAATACCGTCAACGCATCCTTGGAGGTACCTGGAGAAGCTAACAGAAGAATATGGC
GATATTTTTACCCTGAGGATTGGTAGATCACCTTTATTTGTCCTTGGAAGAGCTAGTTCAGCTCATCGTATTCTGGA
GAAACAAAGTGCGCTGAGCAGCTCCAGACCTCGTCTAGTTCTTGCCGGTGAACTATTGAGTAATAATAAAAGAATAT
TACTTATGCCTTATGGAGACCAGTGGAGGCTGTATAGAAAAGCCATGCACGAAACATTGAACGACACGGTGGCGAAA
CAATATGAACCTATCCAAGAGAGGGAGGCTAGAATTGCGACGTTACACCTAGGAAGGCTGGGTCAGGCCGATGGCGG
TGGAGGTGACTTCCAGCGTGTCTTGCACAGGTACGCGGCCAGCGTAATAATGCAAGTTACGTATGATTACCAGGTTC
AGACTCTGGACGATCCCCTGGTGAGGTCAGTGGCCCAGAGGGGTCACGCGTTGGCGATGTGCATCCGTCCTGGCGCT
TCTGTTCTTGACAGGTATCCACTTTTAGAGCACGTTCCCACGTGGTTGAATCCCTGGAAACAGGAAGGTTTGAGGCT
AAGAAAACTAGAGCAAGAGCTATATCTAGGCCAAGTGATCAAGGTCAGAGAAAGAATGGAGAGAGGGGAGTGTGCCC
CGTGCTTTGTTTCTAAAATGACTGAGCGTCAGCAAGAACTAGGGCTTACCGATCTGGATGTTGCGGGGATGAGTGGG
AGTCTGTTCGGTGCTGGGTCAGATACAACGGCGAGCGCACTTTCCATCTTCGTGATGGCTGTGTGCAGATATCCCGC
AGTGTTAGCCAGACTTCACGAGGAATTAGACAGGGTGGTAAGCCGTGATCGTATGCCCACCTTTGATGACATTCCGC

AAATGCCCTATGTGCGTGCCACAGTGCAAGAAGTACTAAGATGGAGGCCAGTCTCTGCCGGTGGTTTCCAACACAGC
CTGACAGCGGATGTAGAGTACAAAGGCTACGTTTTGCCTAAGGGGTCCACGGTGGTGGGCCCCCATTGGTCAATCAG
CAGGGACGAGCACGAATATCCTGAACACGATGTCTTTAAACCTGAAAGGTTTTTACAAAGCGGGGGAGCGGAGGCTA
ACGGCACTTCAGCCCAGGACGAAGTAAAAGGGACTTGGTTCGCGCCCGCACGTGGCTCTGTCGCATTTGGATTTGGG
AGAAGAGTTTGTCCAGGGTTAAACGTTGCCATGCGTTCTTTACACATTAACATTGCTTGTATGGCATGGGCTTTTGA
CATTGCCCAGCCCGATGGCAGACCGGAGAGGGTCGATACGTTTGCGTTCAACTCTGCAGCGAACTCACATCCTCTAC
CATTTGACGCTACCTTCACGTACAGGGACCCCGCGAGGAAAGGCGTAGTCGAGGAAGAGAATATTGCTACAGGGGAG
TTAGACAGGATCGCTGCATCAAGAGGTGCAACT

Seq. ID No: 202
>T5H_10
ATGTTAGAAGCATTGTCCAGTCTAGCTACAGCACTGTGGGCTGCTTTGAGACCTGATACCGTGCTTCTAGGGACACT
GGCTTTTCTGCTATTTGTTGACTTCTTAAAAAGAAGGCATCCTAAAAACTATCCACCAGGGCCCCCTGGTCTACCGT
TCGTGGGAAACCTGTTCCAACTAGATCCCGAAAAGGTGCCATTAGTCCTACACCAGTTTGTGAAAAAGTATGGGAAT
GTCTTCTCCTTAGACTTTGGAACCGTCCCATCTGTGCTAATAACCGGCTTGCCGTTGATAAAAGAAGTGCTAGTTCA
TCAAGGACAAATCTTCTCTAACAGGCCTATAGTCCCGCTACAAGAACATATCATTAATAACAAAGGCTTAATTATGT
CATCAGGCCAATTATGGAAAGAACAACGTCGTTTTGCATTAACGACACTGAGAAATTTTGGTTTAGGGAAAAAGTCA
CTTGAAGAGAGAATCCAAGAAGAGGCCAGCTATTTAATCCAGACAATCAGAGAAGAAAACGGACAGCCTTTTGATCC
TCATCTGACGATTAATAATGCGGTTTCAAACATTATTTGTAGTATTACCTTCGGGGAAAGGTTTGACTACCAAGATG
ACCAATTTCAGGAGCTGCTAAGGATGTTGGATGAGATACTAAACCTTCAGACATCCATGTGCTGTCAACTTTACAAT
GTCTTTCCGAGAATAATGAACTTCCTGCCAGGACCACATCAGGCGCTTTTCAGTAACATGGAGAAAATGAAGATGTT
TGTCGCAAGAATGATCGAGAATCACAAGAGAGACTGGAATCCTGCTGAGGCCCGTGATTTCATTGACGCCTATCTAC
AGGAGATTGAAAAGCACAAGGGCGACGCCACATCCAGTTTTCAAGAAGAGAATCTAATTTATAATACACTTGATCTG
TTTCTTGCGGGGACCGAGACTACGTCAACTAGCCTACGTTGGGGGCTACTGTTCATGGCTTTGAACCCGGAAATACA
AGAAAAAGTTCAAGCAGAAATCGATCGTGTCCTGGGCCAGTCCCAACAGCCGTCAATGGCTGCCAGAGAAAGCATGC
CCTATACTAACGCCGTAATACACGAAGTATTGAGGATGGGCAACATAATCCCCCTAAACGTCCCCAGAGAGGTTGCC
GTAGATACTACCCTGGCGGGATACCACTTACCCAAAGGTACGATGGTTATGACTAATTTAACTGCGCTTCACCGTGA
CCCCACAGAATGGGCTACACCCGATACTTTTAATCCCGAGCATTTTCTTGAGAATGGGCAGTTCAAGAAGAGAGAAT
CTTTCCTGCCCTTCTCCATAGGGAAAAGAATGTGCCTTGGAGAACAACTGGCAAGGACCGAACTATTCATCTTTTTC
ACAAGTTTACTGCAGAAATTCACTTTTCGTCCGCCTGAAAACGAGCAGTTGAGCTTAAAATTTAGAGTATCTCTTAC
CCTAGCCCCGGTAAGCCACAGGCTGTGTGCGGTCCCTCGTGGG

Seq. ID No: 203
>T5H_11
ATGAAGACCCCGCCCCAGTCCTCATGCCCTTTTCATGCGGTTGGACGTCCTCCCACACCCCCTAGATCATCTGCGGG
CAGATGGCCCCCAGGACCAGAGTCAGGTCTGACGGGTTGGGGGTTATTAAAGCTGATGAGCAGAGATTTAATGGGAA
CGCTTGCAGGGTGGCAAAGGGAGTTTGGCGATCTTGTGCATGTCAGGACCTGGCCTGAGCATCAAGTCATAGTGTCC
GATCCACAACTAGCGAGAGAACTTCTAGTCAATCAAGCCGACGCATTACAAAGGTGGGAGCGTGCGCTTACAGTATA
TCGTAGGGTTCACGGACACAGCGTCCTGATAGCCGAGGGCCAGGCCTGGAGGGAGAAGAGACAAGCCCTTCAGCCCG
ACTTTACAAGGAAATCAGTACAGGCCTTTAGTCCCAGTATCGTAGAAGCAGCGCGTAGGGCGTTCGAACAGTGGCCT
GCGAGGCATGCTGCTTGGCCAATTGAATCCGAGCTTACCTCAGTTACGATGGAGGTGATACTTCGTATGATGTTTAG
CAGTGGAGTGGGTTCTGAAGCCCAGCAAGCTGAGGAAGCTGTACATACCCTAATGGTGGCATCAACCGAAGAGCTTT
GGAGACCCGCCAGCTTGCCTGATTGGGTTCCCTGGCAAAGGAAAAGAAGGCGTGCGAGACTACTAATGAACGGTCTA
ATTGAAAGGCACTTGCAGGCGCGACTAGCAATGCCTCAGGATGCTTGGCCCGAGGACTTGTTGAGTCGTTTACTTCG
TCTTCACTTGCAACAGCCCCAATCCTGGCCACTACAGGCTGTGAGGGATGAGTGCAAAACAGCATTCTTGGCTGGGC
ACGAAACTGTCGCAACATCACTTACATGGTGGGCTTGGTGCATGGCGTCTCATCCGGAAATCCAGGAGAGGGCTCGT
GAAGAGGCCCTAGCGGCCTTGTCTGGTGGAGGACAAGCAGACCCAGCTGCCCTACAATACGTCAATCAAACATTGCT
AGAAACTATGCGTCTATACCCTGCCGTCCCCTTACTTATGAGTAGAAGAGCACTGAAACCAGTGACTCTGGGCGACT
GGACTTTTCCTGCGAAGACTGTCTTCATGGTTCCAATGCAGCTTATGCAGCATGACGAGCGTTGGTTCCCGGAACCG
AGATCTTATCGTCCAGAAAGATTTGGGCCTGATGCTGCCCGTCCACAACAGGGAGCGTATTTGCCATTCGGAGGGGG
CCCTCGTGTGTGTTTAGGACAACATCTAGCTATGGCGGAAATGGCGTTGGTGGCAGCACAGCTGCTTCTTAGGTATA

GGCTGTCTGCGCCAGAAGGCGCGGAACCCCCTAGACCCGTTTTCCACGTCTCACAAAGACCAAGCCAGCCACTGACT
CTAGGTATTGCCCGTATT

Seq. ID No: 204
>T5H_12
ATGAAACTTGCAGGTAAAAGGTTCAGGCTACCCCCTGGGCCATCAGGAGCGCCTATTGTCGGCAACTGGCTTCAAGT
GGGGGATGATCTAAATCACAGAAATCTAATGGGTCTAGCTAAGAGATTTGGAGAAGTGTTCCTTTTGAGGATGGGTG
TGAGGAATCTAGTTGTCGTGTCTTCACCAGAACTTGCTAAAGAGGTTCTGCACACCCAAGGAGTTGAGTTTGGTTCA
AGGACGAGGAATGTAGTGTTTGATATTTTTACGGGCAAAGGACAAGATATGGTCTTCACCGTTTACGGGGACCACTG
GCGTAAGATGCGTCGTATAATGACGGTGCCGTTCTTTACCAATAAGGTAGTTGCGCAAAATAGAGTCGGTTGGGAAG
AGGAAGCCAGACTTGTGGTGGAAGATTTGAGGGCGGACCCTGCAGCAGCAACCAAAGGGGTAGTAGTCAGAAGGCGT
TTGCAACTTATGATGTACAACGATATGTTTAGGATCATGTTCGACAGAAGGTTTGAAACCGTGGCAGACCCCCTGTT
CAACCAACTTAAAGCACTCAATGCGGAGAGGTCCATCCTTAGTCAGAGCTTCGATTACAACTATGGAGATTTTATTC
CGGTGCTGCGTCCCTTTCTGAGACGTTACTTAAACAGGTGTACTAATTTAAAAACGAAGAGGATGAAGGTCTTCGAA
GATCACTTTGTCCAACAGAGAAAAGAAGCCCTAGAGAAAACTGGCGAGATCAAGTGTGCAATGGATCACATATTGGA
AGCAGAGCGTAAAGGTGAAATTAATCATGATAATGTCCTTTATATTGTGGAAAACATCAATGTTGCGGCGATCGAGA
CGACACTATGGTCTATTGAGTGGGGTCTAGCAGAGCTGGTCAATCACCCCGAAATACAGCAGAAACTTAGAGAGGAA
ATTGTAGCGGTGCTAGGTCCAGGTACACCAGTAACGGAGCCCGATTTGGAACGTCTGCCTTATTTGCAGTCCGTAGT
AAAAGAAACACTGAGGTTACGTATGGCTATTCCTTTACTGGTCCCTCACATGAACTTGAGCGATGCGAAATTAGCCG
GGTATGATATTCCGGCGGAATCAAAGATCCTGGTGAATGCCTGGTTCTTGGCAAACGACCCAAAGAGGTGGGTCAGA
GCGGATGAGTTTAGACCCGAAAGATTCTTGGAGGAGGAAAAGAGCGTGGAAGCACATGGGAACGATTTCCGTTTTGT
CCCCTTCGGCGTTGGCAGAAGGAGCTGTCCAGGCATAATATTAGCCTTGCCGATTATTGGTATTACCCTGGGTAGGT
TGGTACAGAATTTTGAGCTGTTACCGCCACCTGGCCAGGACAAAATTGATACTACGGAAAAGCCGGGTCAATTCTCC
AACCAGATCTTAAAACACGCAACAATAGTATGCAAGCCGCTAGAGGCT

Seq. ID No: 205
>T5H_13
ATGCACACGGATACCCCTGACACGACCGCCGATCAACCGCTTAGGCGTATAAAAGATTTGCCAGGACCTAGACCTTT
GCCCCTTATAGGCAACGGTCATCAGATAAAACCCCAGAGGATTCACCAGCATGTAGAGAGATGGAGTCTTCAGTACG
GGCCCCTAATGAGGATGTACTTCGGAGCGACTCCCATTTTGGTCGTCGCAGATCACGAAATGGTAGGTGCTGTCTTA
CGTGATAGACCTGACGGGTTTCGTCGTCCCAGTATAAGTGCTACGATTTCAAATGAGATGGGAGGCATACCGGGCTT
ATTCCTGGCAGAAGGAGCAGACTGGCGTAATCAGAGACGTATGGTTATGGCAGGCTTCGCGCCCACGGCAATAAAAG
CGTATTTTCCTGCACTGGTCGCGGTAGCTCTAAGGCTAAGAAGACGTTGGCAAGCAGCAGCATCCGCAAGAAAGGCT
ATCGATCTTGAGAGCGACCTGAAGAGGTACACAGTTGATATAATAGCTGGATTGGCCTTCGGGTCAGATGTTAATAC
CCTGGAGTCTGGAGAAGATGTGATTCAAAGACATCTTGATGACATCCTTCCGGCAGTCGCACGTCGTTCTTTGGCAT
TAGTACCCTATTGGCGTTACGTGAAACTTCCGGCCGATCGTAGATTAGATAGGTCAGTGGCCGTGCTAAGAACTGCC
GTTCAGGACCTTATTGGGCAGGCTCGTCAGAGAATGCTTGATAACCCGGCCAGACGTGAGAGGCCCCCGAATCTTTT
GGAGGCTATGATAGCTGCCGCGGACCAGAGCGGTAGTGGAGTTACCGATTTGAACGTTGCTGGGAACGTCACCAATA
TGTTATTAGCTGGAGAGGACACTACTGCGAACACAATATCCTGGATGATCTACCTTCTGCAGCGTCACCCTCATACC
TTGCAAAAAGCTAGAGACGAAGTTAGGAGAAACGCCCCTGACGCAGCCAGATTTACGATCGAGCAACTAGATTCTTT
GGACTACTTAGGCGCCTGTGCGAACGAGGCCATGAGACTTAAACCTGTTGCGCCTTATTTGCCTCTGGAGGCTCTAA
GGGATACAGTAATTGGAGATGTTGCGGTGCCTGCGGGTACCATGATCTGGTGTGTGTTGAGACATGACAGTGTAGCA
GAAAAGCACTTCCCCGATCCACTGTTGTTCGACCCACAGAGATGGCTGCAGGCTGATGGAAAGCCGAACTCCGATAA
GCGTGTTACAATGCCCTTCGGTGCCGGACTAAGAACTTGTCCAGGCAGATATCTGGCTTTACTAGAGATCAAAATCG
CAATGGCTATGTTGCTGGGAAGTTTTGATATTGCCGGGGTGGACACGCCGGACGGCAAGGAGGCTCAGGAGCTTATG
GGGTTTGTGATGTCTCCAGTTGGATTATCTTTGAGATTAGAG

Seq. ID No: 206
>T5H_14
ATGCTGATGAAAACCTTAATGGCATCACTTCAATGGTTGAAGGAATCCTTCCAGCCGTTTATGTTACTGTTTGCTTC
CATTTTTCTTGCCGTGCTTCTAAAGTTCTTCTTCAAGGAGAAGAGCCGTAAAAGGTCAAACTTACCTCCGAGCCCAC

CAAAGTTACCGATTATTGGGAACTTGCATCAGCTTGGCAACATGCCTCACCTTTCCCTGCATAACTTGGCAAAGAAA
TATGGGTCAATTATCTTTCTTCAACTGGGGGAAATTCCCACCGTGGTAGTCAGTTCAGCTAGGCTTGCGAAAGAGGT
AATGAAAACTCATGACCTTGCTCTGAGTTCTAGGCCGCAAATTTTCTCAGCGAAGCATCTTTTCTACAACTGCACAG
ATGTTGTATTCAGTCCGTACGGCGCTTACTGGCGTCACATTAGGAAGATCTGTATCCTGGAGTTACTGTCAGTTAAG
AGGGTCCAGAGTTACTCTTTTGTGAGGGAGGAAGAAGTGGCTCGTTTAGTCAGGAGAGTCGCAGAGTTTTACCCAGG
CACTACGGACTTGACTAAGATACTGGGATTATACGCAAATGACGTCCTTTGCCGTGTAGCCTTTGGCAGAGATTTCA
GTGGGGGCGGAGAGTATGATCAGCACGGGTTCCAGAAAATGCTGGAGGAGTACCAGGAATTGTTGGGCGGTTTCTCT
CTAGGTGACTTTTTCCCCAGTATGGAGTTCGTGCACAGTTTAACCGGGATGAAGAGTCGTTTGCAAGACACGTTTCG
TAGGTTCGACCAGCTATTCGACCTTTTCCTGACGGAACACAGAGACCCAAAGAGAGAGACAGAAGAACACAAAGACC
TGGTGGACGTGTTATTAGACTTACAAAAGAATGCCTACGATGAAATGCCGCTGACAACGGACAATATCAAAGCTATA
ATATTGGACATGTTTGCGGCAGGGACCGATACTACGTTTATCACTTTAGATTGGGGGATGACTGAACTAATCATGAA
CCCCGAGGTAATGGAGCGTGCCCAGGCCGAGGTCAGGTCTGTGGTCGGTGACCGTGAGGTTGTATTACAGAGTGATC
TGCCTCAGCTACATTACATTAAGGCCGTCATCAAAGAGATCTTTAGACTACATCCCCCGGCACCCGTTTTGGTGCCC
AGGGAATCTATGGAAGACGTTAGTATAGATGGTTATAATATCCCTAGCAAAACCAGGTTCTTCGTTAATGCGTGGGC
TATTGGGCGGGACCCAGAGTCCTGGGAAAACCCCAACGCTTTCGAACCAGAAAGATTTATGGACAGTACAATAGATT
TCAAAGGTCAGCACTTTGAACTTATTCCCTTTGGTGCAGGTAGGCGTAGTTGCCCCGCAATCGCCTTCGGCGAAGCG
ACTATCGAATTAGCCCTAGCGCAACTTTTGCATAGCTTTGACTGGGAGCTGCCTCCAGGTACGACGCCCAAAGATCT
TGACATGTCAGAAGTGTTTGGAATCACTATGCACCGTATTGCCCATTTAATAGTCATCGCGAAACCGAGATTTCCCG
TCGGCCAAAACAAA

Seq. ID No: 207
>T5H_15
ATGCCCAAACAGAAGAAACGTTTGCCGCCAGGGCCTCCAACACTACCAATCATTGGGAACATGCACCAGCTGGGAGA
ACTTGCGCACAAATCTCTGTCCGAGCTGTCAAAGAAGTATGGCCCTATTATGCTTTTAAAAATTGGTAGTAAAACAA
TAATCAACATAAGTAGCGCAGAAGCGGCGAGACAGGTATTGAAAGTGCATGATTTGGATTGTTGCTCAAGACCGGTA
AGCTCCACAGCTGGCAGACTAACGTACAATTTTAAAGATATTGTATTTGCACCGTATGGCGATTACTGGAGGGAGAT
GAGGAAAATCTGCGCACTAGAGCTGTTGAGCGTGGCGAGAGTGCAAAGCTATCGTTTCATTAGAGAGGAAGAAGTCG
CAAGTCTAGTAAATTCTATCTCCCAGAGTGCATCTTCTGCCACCCCGGTTGATCTGTCTGAGAAGATGTTAGCATTG
ACAGTGAACATACTTTGCAGGACTGCGTTCGGTAAAAGTTTCAGAGGATCAGGACTTGACAATGGTAAATTAAGAGA
GGTCGTCCACGAGGCTGAAGTCATGTTCGCGAGCTTTTCTGCGACCGAATTTTTCCCCTACGTAGGTTGGATTATTG
ATCGTCTAAGCGGACGTATAAGAAGATTAGAGAAAATTTTTAGAGGGCTAGACGATTTCCTGCAACAGGCTATTGAT
TTGCACTTAAAACCTAAAAAGACCGAACAAGACCATGAAGATCTAATTGACGTACTATTGAAGATAGAGCGTGACCA
ACAAACCAATACGGGGGCACCCCCGTTCAACAAGGACAACATCAAGGCCATACTTTTTGATATGTTCTTGGGTGGAT
CAAATACAGCCGCAGTGACCATGTTGTGGGCTATGGCTGAACTAGCAAGGAACCCCAGAGCTATGAAGAAAGCTCAA
GATGAGGTACGTAATGTCGTAGGCAATCGTGGAAAAGTCACTGAAAGTGACATCACACATTTGCATTACCTGAAGAT
GACCATCAAAGAGACGTTTAGGTTACACCCGCCCGCAGCGATACTTCTTCCAAGGCAAACTATGGCCGAGGTCAAGA
TCGGGGGCTATGACATCGGCCCGAACAGCCTTTTACAGGTTAATGCGTGGGCTCTAGGTAGGGACCCGGAGTACTGG
ATGAACCCCGAAGAGTTCTACCCAGAGAGATTCGTTGACTCTAGCATAGATTACAAGGGACAACATTTCGAATTATT
ACCTTTTGGTTCAGGTCGTAGGGGCTGCCCGGGTATGCACATGGGAACAACAACGGTCGAGTTAGCATTGGCTAACT
TGCTGTATTGCTTCGACTGGAAGTTGCCCAGTGGTCTTAAAGAGGAGGATATTAATATGGATGAGTCTACGGGTCCG
GGGCTGACCCAAAAGAGAACGACACTGAAGTTAGTGCCGGTGAAGTTGTTT

Seq. ID No: 208
>T5H_16
ATGAAGCTACTGCTAGATAGGACCAGAACTAATGGGTACCTTCCCCCTAGCCCTCCAAAGCTACCCATCATTGGTAA
TTTACATCAATTAGGCAAAATGCCGCACATCAGCCTGTGTGAGCGTGCTCAGAAGTTGGGACCGATTATGTTCTTGC
AACTTGGAGAGGTACCGACTGTAGTCATCTCCTCAGCGGCTATGGCAAAGGAGGTGATGAAAACGCATGATTTAGCT
TTCTCAAGTAGACCGCAGTTATATTCCGCCAAGTGGTTGTTCTACAATTGCACTAATATCGTATTTTCACCTTATGG
AGCGTATTGGAGACACGTTAGAAAGATTTGCATTTTAGAGCTGCTGTCCACTAAAAGGGTCCAGTCTTATGGGTTTA
TTCGTCAAGAAGAGGTTTCAAGACTACTTCATAGAATAGCTGATTCTTGCAGCAAGCCAATAAATCTAAGCAAATTG
TTGGGCCTTTACGCGAATGATGTGCTATGCCGTGCTGTCCTGGGCAGAAATTTTAGTGAAGGCGGGGATTATGATAT

GCATGGGTTCCAGTCTATGTTGAAAGAGTACCAAGAGTTGTTGGGTGGCTTCAGTATAGGAGATTTCTTTCCAAGTA
AAGAATTTGTCCATTTATTAACAGGCCACAAAAGGAGGTTACAAAATACGTTTAAAAGATTCGATAACTTCTTTCAA
CAGGTTGTGCGTGAGCATTTAGATCCGGAAAGAAACTACGAAGGTGAAAAAGATATCTTAGATGTCTTGTTAGATAT
CCAAAAGAATGGAAGTAGCGAAATGCCCTTAACACTGGATAATGTTAAAGCCATTCTGTTGGATATGTTTGCAGCCG
GAACAGACACATCATTCATCGTCCTGGACTGGGGAATGACCGAACTAATCATGAACCCAAAGGTTATGAAAAAGGCG
CAAGCAGAAATAAGGAGAGTAGTAGGTGAGAGGCAAGTCGTCCTTGAAAATGATCTGCCGCAACTACATTATTTAAA
AGCAGTCATTAAAGAAATATTCAGGCTGCATCCGCCTGTCCCGGTACTGGTCCCGCGTGAATCTATACAAGATGTGA
CCATTGAGGGCTACAATATCCCAGCAAAGACAAGGGTGTTTATCAACGTGTGGGCAATAGGCCGTGATCCAGAAAGC
TGGAAGAACCCAGAGACGTTCGATCCGGAACGTTTCGTTGGTAGCACAATCGACTTTAAAGGCCAGGACTTCGAACT
GCTGCCCTTTGGGGCTGGCAGGAGGGGATGTCCAGGAATTACATTCGGAGCAGTAACTGTTGAATTGGCCTTAGCGC
AACTGTTGCATAGCTTCGATTGGAAATTGCCTTTAGGGGTTGAAGCAAAGGACCTAGATCTAACGGAGGCATTTGGA
ATTTCCATGCCTAAGACGAGTGACTTGATCGTCGTCGCCAAGCCGTGTTTCGCA

Seq. ID No: 209
>T5H_17
ATGAGAATGGACGGAAATTCAACTACGATGTTCCCTCTTCTGATAACGGTCATAATGCTTTTGGCCTCTGTGTTATT
TTATATATTCAACAGGTGGACTCATAGGTATAGTAAGTCAGGAATACTGCCTCCTAGTCCGCCTAAGTTACCCCTGT
TAGGTCACCTGCACTTGCTTTCAGATCAACCGCACGTTGCCCTGTCAAGGTTAGCTCAAAAGTATGGCCCAATTATG
TACTTAGAGCTGGGACAAGTGCCGACTGTGGTTGTTAGCTCAGCGTCCTTGGCACGTGAAGTCCTTAAAACCCACGA
CCATGTATTTTGTAATCGTCCGCAAACTATAGCCGCACAGTACATATCATTCGGGTGTTCAGACGTGACTTTTTCTC
CTTATGGGCCGTATTGGAGACAGGTGAGAAAAATATGCGTCACAGAATTATTGACCCTGAGGAGGGTTAACTCTTTT
CAGTTAATACGTGAAGAGGAAACTAATAGGTTGCTTACGGCCGTCGGAGCGCACTCAGGGAGTGAGGTGAACTTAAC
GAAATTATTTTTCAATTTGGCCAATGATACGCTATGCAGAGCCGCATTCGGAACGAGATTCATGAGCGAGTCTACCC
AGTTAGAAAGGCAAAGGGAGGGCAAAAGACTTGAGGATATTCTGATTGAGACCGTAAAACTGTTAAGTGGATTCTAT
GTGGGCGAGTTCTTTCCCAGGTGGGGGTGGATTAACAGCGTCAGCGGTTTCAAACGTAGACTTGAGAGAAACTTAGC
AGATTTGCGTTCAGTAGGTGATGAGATAATCCAAGAACACATTAAGAAGAGAGGACGTGGTAACGAAGAAGAGGATT
TTGTAGACGTATTGTTACGTGTCCAACGTCAACAAGACTTGCAAGTCCCAATCACTGATGACAACGTAACGGCACTT
GTTATGGATCTTTTCGTGGCGGGGACAGATACCACATCCTCCACTTTGGAGTGGACGATGACCGAAATGGCCAGGCA
CCCGGAGGTGATGAAAAAGGCTCAAGCAGAAGTGAGATCTATGTCACCCGAAGGGGGCACGTTGGACGAGTCCCATT
TAAGACATTTGCACTATCTTAAAGCTGTTATTAAAGAGGCACTTAGATTACATCCACCAATACCTCTACTTCTGCCC
CGTGAAAGCATGGATAAGTGTGCCATCGACGGTTATGAAATCCCTGCTAAGACAAGAGTACTTATAAATAACTTCGC
ACTTGGAAGGGACCCCGATTCTTGGGACGACCCTCTTCGTTATAACCCTGCCCGTTTTATGGGTGGCGATGAACATA
AAATCGATTTCAAAGGGGAAGATTTCAGATTTGTTCCGTTTGGCGGTGGTAGACGTGGGTGTCCCGGCTACTCCCTG
GGACTTGCTACCGTTGAATTGACCCTAGCCCGTTTACTATACCATTTTGACTGGAAGTTGCCGCCTGGAGTAGAGGC
AGAGAAAATTGATTTGACAGAAATTTTCGGCCTAGCTACTAGGAAGAAAACCCCATTACTGCTTATTCCGACTGCCA
GGAAAGCTCCACCGCATGAA

Seq. ID No: 210
>T5H_18
ATGGAATTAACTATGGCTTCTACCATGTCTCTGGCATTGCTAGTTTTAAGCGCCGCATATGTCTTAGTCGCTCTAAG
AAGATCCCGTAGTAGTTCCTCAAAACCTCGTAGACTTCCACCATCACCTCCCGGTTGGCCGGTTATTGGACACCTGC
ATTTGATGAGTGGGATGCCACACCACGCGCTTGCGGAATTAGCAAGAACCATGAGGGCTCCACTATTTAGAATGCGT
TTAGGTTCCGTGCCAGCTGTTGTAATATCAAAACCCGACTTGGCTCGTGCTGCACTTACCACAAACGATGCGGCATT
GGCGTCCCGTCCGCATCTGCTTTCAGGACAGTTTCTGAGCTTCGGCTGCAGCGATGTAACTTTCGCACCGGCAGGAC
CGTATCATAGAATGGCTCGTAGGGTGGTTGTAAGCGAGTTACTAAGCGCACGTCGTGTGGCAACATATGGGGCGGTG
AGAGTTAAAGAGCTTCGTAGGCTATTGGCCCATTTGACGAAAAATACAAGCCCTGCTAAGCCTGTAGATTTATCAGA
GTGTTTTCTGAATTTAGCCAATGACGTGCTATGTAGAGTAGCTTTTGGTAGACGTTTCCCACACGGAGAAGGGGACA
AGCTGGGAGCCGTACTGGCTGAAGCGCAAGATCTATTCGCAGGCTTCACTATAGGGGATTTCTTCCCGGAATTGGAG
CCTGTAGCTAGTACGGTTACTGGGCTACGTCGTAGATTAAAGAAATGCCTAGCCGATCTTAGGGAGGCGTGTGACGT
GATAGTAGACGAGCACATTTCCGGTAACAGGCAAAGAATCCCAGGTGACCGTGACGAGGACTTCGTTGATGTATTAC
TACGTGTCCAAAAGAGCCCCGACCTGGAAGTACCACTTACGGACGACAATCTGAAGGCCTTAGTGTTAGATATGTTC

GTGGCAGGTACGGATACCACTTTCGCAACACTGGAGTGGGTCATGACAGAGCTTGTACGTCATCCGAGAATCCTAAA
GAAAGCTCAGGAGGAAGTTCGTAGAGTGGTCGGCGACTCTGGTCGTGTTGAGGAAAGCCATTTAGGAGAACTGCATT
ACATGAGGGCCATTATTAAAGAAACTTTTAGACTTCACCCAGCGGTTCCCCTATTGGTACCGAGGGAGTCTGTTGCA
CCATGTACTTTGGGTGGGTATGACATACCTGCAAGGACGAGAGTATTCATCAACACGTTTGCGATGGGGAGGGACCC
GGAAATTTGGGATAATCCTCTTGAGTATTCTCCAGAAAGATTTGAGTCTGCCGGCGGGGGAGGCGAGATTGATCTCA
AGGACCCCGATTACAAACTTCTACCATTCGGCGGCGGAAGACGTGGCTGTCCAGGATATACGTTCGCCCTTGCGACC
GTTCAGGTTTCATTGGCTAGTTTATTATATCACTTTGAATGGGCACTACCGGCAGGGGTAAGAGCCGAGGACGTCAA
CCTTGACGAAACTTTCGGTCTAGCAACTAGAAAGAAAGAGCCCCTGTTTGTTGCGGTACGTAAGAGTGATGCTTACG
AGTTCAAAGGAGAGGAGCTAAGTGAAGTC

Seq. ID No: 211
>T5H_19
ATGCCCCTAAGCGACAGCACTATCTCCCTTCTACTGGCCGTGTTGCCGATTAGCGGGATCATTTTCGCCCTATATAA
CCAGTACCAGATTTGGCTTAAATCTCCGATAAGAGGGCTGCCCTATCCCCCTGGACCACCCTTGTTGTTGGGGAACG
CGAACAGGGCTGTACAGTCTAGGCCCTGGTTAACGTATACAGAGTGGGCCAAGCAGTACGGAGACATAATATATGTC
AATATATACGGAGAACACACGGTCATTTTGAATAACTTAGAGGATGTTATGGAACTGTTCGAACAGAGATCAAGGGT
TTATTCTTCCAGACAAAACAACCCATACATCGAGCTGATGGGATGGCAGTTTAACGCTGGGCTTCTGCCCTATGGTG
ATCTGTGGAGGAGACATAGGAAGTTACTACAGCAGTGTTTCAGAAGAAAGATTTCCACACAGTACGAACCCATACAG
ATAGCTAAGACCCACAACCTACTGAATGATTTACTGCAAACACCTAGCGACTTCATAGAGCATATCAAACGTAACAG
TTCTGCCATGATAATGAGTATCCTATACGGTCAAGACATCTCTGACGAAATGTCTGCCCAATTCGTTTCTGTGGCTG
AAGAAAGTGTAAAGGCTCTGGGGAAGTGCCTGCGTCCCGGCACCTACCTTGTATCTTACATCCCAATGTTGCGTTAC
CTGCCAGCCTGGTTTCCCGGAGCAGAATTTCAACGTCAAGCTGCTGAGGTGAAAAAGCTAACTACCAAAATGAAAGA
TGAGCCAATTGATTTCGTAGGGAAGGGTCTGTTACACGGCACAGCATCAGCATCTCTTGTGGCAGACCTATTGGAAA
ACTGCTATGTGCAAAGAGAATACGACGTAATTAAAGACGTAGCCGCGACCGTTTTTGCAGCGGGGGCAGATACCTCT
GTGGCAGCATTGGAGTCCTTTTTCCTAGCTATGAGTCTTTTCCCGGAAGCTCAGAAAAAGGCCCAGGCTGAGATGGA
CAGAGTAATTGGTAATAAAAGATTACCAACGACCGATGACAGGCCACTGCTACCTTATCTGGAAGCCGTATACCGTG
AGCTGATGAGGTGGGCTCCTGTAGTGCCTCTAAACGCGGCGCACACCACAATCGCGGACGACATCTACAAAGGCTAT
TACATACCGAAGGGCACTGCAGTCTACGCTAACACGTGGGCCTTGACGCGTAATGAGGAAAAGTACCCCAATCCAGA
CATATTTAATCCCGATAGGTTCTTTACCGAGACCGGAGAGCTAAATGACGATGATACGGTATTGACTTTCGGGTTTG
GAAGGCGTATATGTCCGGGCAGGCACATGGCATCTACGACGGTTTGGTTAACTATTGCTTCAGTTTTAAGTAATTTT
GATATAAAGGGGAAAGGCACAAATACGAAAGATCAGAAATTTACTAGCATCGGGGAAATGTTTACCGACAACTTCAT
TTCCAGGCCGGTGCCATTTGAATGCGACATTGTGCCGAGAAAGAACGCAGCCCTGCTGGCAAGTAAA

Seq. ID No: 212
>T5H_20
ATGGCGTTCGAAACGACGAACGGTATCCTATTAGCTGCCAGTCTATTCGCGGGGGTGGTTCTATATCTGCAAAAGAG
AAAGAGATACACCCTACCGTATCCTCCGGGGCCAAAGAAGCACTTCCTTTTGGGCAACTTACTAGATGTTCCAACCA
CGTTTGCATGGAAGAGATACGCAGAGTGGGGGAAGACTTTTGATTCTGATGTGCTACACTTGAGCGTTGCAGGGAGC
CATTTCATCATACTTAACTCCTTTAAGGCCGCCAACGATCTGTTCGAAAAACGTAGCTCTATTTATTCTTCAAGGGC
ACAGATGATAATGTTTAGCGAACTGATCGGTTGGGATTGGCTGATGAGCGGCATGGTCTATGGGGAGCCTTGGCGTG
AAAGGAGGAAGGCGTTCCAACAATATTTTCACGTAGGCAATGCCCACTTATATGAACCCGTTCAGATGCAAGCAGTG
AGAAAGATGCTACCACGTCTGTTAAAAGAACCTGAAGACTTTTTAAGCATCACGAGGCATGCGCTTGGCAGTATGGC
ACTAACCCTAGCCTACGGCCTAGACATACAAGAAAAGAATGATCCCTATTTAAGAGTCTCTGAAGCTGCAGTCAAGT
CTATAGGTGAAGTCGCAATTCCAGGCGCTTTTCTTGTGGACATGATCCCGGCGCTGAAATATGTACCTGAATTTTTC
CCTGGAGCAGGTTTTAAGAAAAAGGCACGTATTTGGAGAAAAGTCCAGGAAAACATGCGTGAGATACCTTTCGCTGC
TACACTGAAGAATATAGCGAGTGGTAGCGCCAAGGTCAGTTTTACAAGTACATGCCTAGAAAACCTTGATGAATCAA
GGGACGTAGACCACCAACGTACTATTATTAAAGATACGGCCGGAAATATGTTTGCCGCTGCGACTGATACAACTATT
AGTGCTATCCATACCTTCTTTGTTGCGATGTTGTGCTTTCCCGAGGTCCAGAAAAAGGCTCAACAAGAAATTGACAG
GGTATTACAAGGTAGGTTACCTGAGTTTAGTGATGAGGCCGACTTACCGTACTTATCTGCATTGGTTAAAGAAACGC
TGAGGTGGGAGCCTAGCACCCCTATCGGTGTGCCGCACTACAGTTCTGAAGATGATGTCTATAACGGATATCACATA
CCGAAAGGAAGTCTAGTTATCGGAAATGCGTGGGCAATGCTTCACAATGAGGAGGATTACCCGGAGCCTTCACTTTT

CAAGCCAGAAAGGTTCATCAAGGACGGTAAATTGAATCCCAACGTGAGGGACCCGGCCGAAATGGCCTTCGGGTTCG
GGCGTCGTCTGTGTCCGGGGAACCACATTGCCATTTCTGCGCTATGGTTAACCGCTGCAACAGTACTAGCTACCTTT
AACATCACCGAGGCAATAGATGACGATGGTAGACCGATCAAGCCATGTGTGGAATATGAATCTGCGTTAATATGCCA
CCCCCTTCCTTTTAAATGCACAATCAAACCTAGATCTAAGGAATGCACAATGTTAATTCAAGCGGCAGCTGACAGCT
AT

Seq. ID No: 213
>T5H_21
ATGATTATCGATTCTAGTAACTCCGAAGGCAATAGTGAAGGCCAGTACACCATCGATGGACCTAAGGCCAAAGGACT
GAGAAGGATGTTTAGAATCTTCCACTTAATCTTACAGCCGACTAAGTATATGGAGTCTTCCGTACAGAGGTACGGTA
GTATGTTTCAGATAGGAAGCGAAGGAGCATCACCATTAGTATACGTAGGAGAACCAGAAGTTGTGAAAGAGATTTTC
GCTTTGGACGGAGATCAAGTCGTGACAGGGCAGGGTAACGGAGTGCTAGAGACTATGGTTGGCAAGCACAGTATTCT
TTTACTTGATGGCGACCCTCACCGTCAACAAAGAAAGTTGTTAATGCCTCCATTCCACGGTGAGCAGCTAAGGGCAT
ACGCACATTTGATTTGCGATATTACCAGGCAGATCAGTGCACAGTGGCAGCCGGGACAAACCATCGTGGCTAGACCG
CCCATTCAGAATCTTACTCTAGGAGTGATCCTACAGGCAGTGTTTGGGGTCCCCTCTGGCGAAAGGTTGTCCCGTCT
TCAACAACTTATGTCCACGCTGTTAGACTCTTTTGCCTACCCCATATCAGCATCATTTTTATTCTTCCCGGCGTTAC
AAAAAGACTTAGGAGAGTGGTCCCCGTGGGGTAAGTTTATTAGGCTGAGAGAGGAAGTCCGTAGTCTTATTTATGCG
GAAATTAGGGATAGAAGGCAGCAGTTAGAGAGAAGCGCAATTGAGCAAGACGAAAAACTTGGAGAGAAGTTAGGCGA
GAAGACAGATATATTAACCTTGCTACTACAAGCGCGTGACGAGGACGGAGGGGCGATGTCTGATGCTGAGCTTCACG
ATGAGATAGTAACCCTTTTGCTTGCCGGCCATGAAACGACAGCCAGCGCAATAGTGTGGATGCTATACTGGATACAC
TACCTTCCTGAAGTGCAGCAGAAGTTGCGTGCCGAGCTGGATGCGTTGGGACCGGACCCTGATCCTATGGCTATAGC
ACAGCTTCCCTACCTGACCGCCGTGTGCCAAGAAGCACTGAGGATATATCCTATCACTCCAACCACATTTATTAGAC
GTTTAAGGGAGCCTATGACTCTTGCAGGTTACAGATTTAAAGCGGGTACGGCACTTATGCCGGCGACGTACATCATC
CATCAAAGACCCGATCTATACCCCGAGCCTAAACAATTTAGGCCAGAGAGGTTCCTAGAAAGACAGTTTGCTCCGCA
TGAGTTTTTACCATTCGGCGGGGGCCATAGATATTGCATCGGAAGTGCCTTGGCCATGATGGAGTTGAAGTTATCAA
TTGCAACGCTGCTTGCGGATTTCGAACTGGCATTACTGCACTCCAGGCCGTTGCTGCCAGCTAGGAGAGGACTAACA
ATGGCTCCACCAGCAGCAATGAAGTTAAGGATCAAGGCGAGAAAAACCAACAAAGCT

Seq. ID No: 214
>T5H_22
ATGCCCGCGCCTAAAACTGCACCGAGCACACTGCCATTACCACCAGGAAGGTTAGGATTGCCGTGGATTGGTGAAAC
ACTGTCATTTTTGAGGGACCCAAACTTTGCGACAAAAAGACAGGCACAGTATGGTAGTTTGTTCAAATCTAGGATCA
TTGGCCAGCCGACCGTATTCTTTTGCGGACCAGAGGCGAATGCTTTCCTTCTGTCCAGCCACGCGGACTGCTTCTCC
TGGAGAGATGGCTGGCCAGGCACATTTCAAGAACTTCTTGGTGAGTCCTTATTTTTACAAGAGGGTGAGACTCATTT
ACGTAACCGTAGACTGCTGATGCCAGCCTTTCACGGAAAAGCGTTGGCAAGTTATTTCAGCACTATGGTTGCCCTTT
CCGATAGCTATCTGGCCCGTTGGGAGAAGAAACAGCAATTGACATGGTTCCTAGAGTTCAAAAAGTTTACGTTTGAA
GTAGCCTCTGTTCTATTGGTCGGCTCAGCCCCAGGTCACGACGAGACCGACAACACTATCGGCACGGCTGAAAGTGC
GGAGACGGAGGCCCAAATCGCGCAGCTGGCTTCCTGGTTTGCAGATCTAACGAACGGCTTATTCACCCTTCCTATCC
GTTGGGGTCCCACGACCTATAGGAAGGCCCTAAGAGGCAGGGATAGATTGCTGAGTTATATAGAACAGGAGATTACT
AAAAGGAGACAATTATTGGCACGTTTACAAACCGATCCAACCGCAGCTCTACCGACGGATGTTTTAACCTTACTTCT
TCAAACGGAAGATGATGAGGGAAACAGACTTTCTGAGGCAGAAATCAAGGTTCAGACATTGCTAATGCTGTTTGCGG
GGCACGAGACGACGACCTCTATGCTAACTTCATTAGTTATGTCTCTTGCCCAAAACCCGGACGTCCTAGCCAAAGCA
CGTGCGGAGCAACAAGCGTTCCCTGCAGAGTCAGCCCTAACGTTTGAACAGATACAACAAATGCCATATTTAGATCA
GATTCTAAAAGAGGTCGAACGTCAGTATCCTCCCGTCGGGGGAGGATTCAGGAGGGTCATCAAGCCGTTCAACTTTA
ATGGATACCATGTCCCTGCAGGTTGGCTGGCATTATACAGGATTGATGCGGCGCACAAGGACGAAAGATGTTATACG
AACCCGTCAGACTTTGACCCTGATAGATTCAGCCCTGAGCGTGCTGAACAAAAACGTTATGACTACTCACTTGTTGG
GTTCGGTGGGGGCCCACGTGTGTGCCTAGGCATGGCGTTTGCCAAGCTCGAAATGAAAATAATGGCTGCACAATTGC
TTAGACGTTATCACTGGCAACTAGACGCTGATCAAGATCTGACTATGAACCCTGTACCTAGTTTGAGGCCTGCGGAC
GGTTTGAAGGTCCGTTTCAGTAAGCTAAGTTTCACAGCC

Seq. ID No: 215

>T5H_23
ATGTTGGATATGCCCTCCGTGAAACCTTGGCTAACCTTCTCAGACTGGGCCTCCAAGTTCGGTGATATAAGTCATTT
AGAAATATTTGGTCAACACATTGTGGTCCTTAACAGTGCTAAAACTGCTGTTGAAATGTTAGATAGAAAATCTAGCA
TCTACTCCGATCGTCCGGTATTACCAATGGGTGGGGAATTAGTAGGCTGGAGAAATACCCTGGTCCTATTACCGTAT
GGAGATAACTTCAGGGAATATAGAAGGAATTTTCACAGAGTGATTGGAAGCCGTGCGGCTATGAGCGTTTACCACGC
GATTGAAGAGGAGGAAACTCATAAGTTTCTACAAAGAGTGCTGACGAAACCGGCTGATTTGAGTGCTCATGTAAGGA
CAACAGCCGGTGCAATTATTCTAAGGATTAGTCATGGGTATCATATTCAAGAAGACGGGGACCCATTTGTCAGTTTG
GCGGATACAGCTGTGGATCAATTCTCCCGTAGTACGGCTACAGGAGCGTTCATGGTTGACCTGATCCCGGCCCTAGC
TTACGTACCAGAATGGTTTCCTGGAGCTAGTTTCCAAAGAAAAGCGCGTGAGTGGAGGGCCACATTACACGAGATGG
TCAATCAGCCATATAAGTTCGTGCAAGACCAGATGGCTGCCGGCATAGCCCCTAAGAGCTTTACCAGTAACCTGCTG
GAGGGACGTACATTAACGGAAGAGGAGGAGCATATTATCAAATGGTCAGGTGCGTCTCTGTACAGCGGGGGTGCCGA
CACTACGGTTAGTGCCATTTATGGATTTTTCCTGGCCATGACGCTTTATCCCGAGGCGCAGAAGAAAGCCCAGGCCG
AGATCGACGCAGTAGTCGGGAGTGATCGTCTACCCACATTTGCAGACAGGGAATCTCTACCTTACGCTGAAGCATTG
GTCAAAGAAGTTCTTAGATGGTGTCCCGTCGTCCCGATTGTAGTTCCGCACCGTGTTACCGCAGATGATATACATAA
TGGATATTATATTCCCAAGGGTACCTTAGTACTAGCTAATGCTTGGTATATGTTGAGGGACCCGTCAATATATCCCG
ATCCAATGAACTTTAACCCCGATAGGTTCCTTCCGTCAGGCGGGAAGGAGCCCCCAACAGACCCGCGTGATATTTGT
TTTGGTTTTGGACGTAGAATCTGCCCTGGAATGCACTTGGCGGACGCTAGTGTTTGGTTATCCGCCGTAATGTCTCT
GGCGGTGTTTAATGTCTCAAAGGTGGTCGAAAACGGTGTGGAGATCACCCCTGAAGTCGATCCTAGTTCCGGAACTA
TCAGTCACCCGAAGCCCTTCAAATGTTCTATTAAGCCAAGATCTGCTAAAGCACTAGAGCTAATTCAGCAAACGCCA
CATTAC

Seq. ID No: 216
>T5H_24
ATGCACTTACCTCCAGGTCCAAGGCCGCTTCCATTCCTTGGGAATTTGCTTCAGATGAACAGACGTGGATTGTTGAG
GAGCTTTATGCAGTTGCAAGAGAAGTACGGCGATGTATTTACAGTGCACCTTGGTCCTAGGCCAGTCGTAATTCTAT
GCGGGACCGATACAATTAGAGAAGCCTTGGTGGATCAGGCCGAAGCGTTCTCCGGAAGAGGTACCGTTGCTGTTCTG
CATCCCGTCGTCCAAGGCTACGGAGTCATTTTTGCAAATGGCGAGAGATGGAAAATCCTACGTCGTTTTAGTCTAGT
TACTATGAGAAATTTCGGAATGGGAAAGAGGAGCGTCGAGGAAAGGATAAAAGAGGAGGCACAATGCTTGGTTGAAG
AGCTAAAGAAATACAAAGGCGCGCTACTAAACCCAACTTCCATCTTTCAGAGTATTGCTGCAAACATAATCTGTTCC
ATAGTTTTTGGGGAAAGATTTGATTACAAAGACCATCAATTTTTAAGGTTACTGGACTTGATTTATCAGACATTTTC
TCTTATGGGCTCATTGAGTAGTCAAGTATTCGAGCTATTTTCCGGTTTTCTGAAGTATTTCCCTGGGGTGCACAAAC
AGATTAGTAAAAACTTGCAGGAGATTTTAAATTATATAGATCATTCCGTTGAGAAGCACCGTGCCACATTGGACCCA
AATACTCCCAGGGACTTCATCGACACATACCTACTTCACATGGAAAAGGAAAAATCTAACCATCACACTGAATTTCA
CCATCAAAACTTGGTAATTTCCGTCCTGTCCCTGTTCTTTGCAGGCACCGAGACCACTTCCACGACTTTACGTTACT
CATTTTTAATCATGCTGAAATATCCACACGTCGCAGAAAAGGTTCAAAAGGAGATCGACCAAGTGATATCTAGTCAT
AGGTTGCCGACGTTGGATGACAGGATTAAGATGCCCTATACCGACGCAGTCATCCATGAAATTCAGCGTTTTGCAGA
CCTAGCGCCAATAGGGTTACCGCACAGGGTTACAAAAGATACAATGTTCCGTGGATATCTGCTACCTAAAAATACCG
AGGTATATCCGATTTTATCCTCCGCACTTCATGATCCGAGATACTTCGACCATCCTGACACTTTTAACCCGGAACAC
TTCTTAGACGCTAATGGAACGTTGAAAAAGTCCGAGGCATTCCTGCCGTTCAGTACAGGAAAGAGAACTTGTCTGGG
TGAGGGAATAGCACGTAACGAGTTGTTCATTTTCTTTACAGCCCTACTTCAGAATTTTAGCCTTGCCAGCCCTGTCG
CACCCGAGGACATTGATCTTACTCCAATAAACTCCGGGGCGGGTAAAATTCCAAGTCCATACCAAATCAATTTTCTA
AGTCGTTGTGTCGGA

Seq. ID No: 217
>T5H_25
ATGTATCTAATACCCGACTTTAGTAAAGAGACCTGGATACTACTTATCATCTTACTTGCGCTGTTGGCTTATTATGG
TATTTGGCCCTATAGGCTTTTCAAGAAATATGGGATACCTGGCCCTAAACCTTTACCTTTCTTTGGTACCTTTCTGG
AGAACAGGAATGGAGTGTTTGAATTTGATATGGAATGCTTCAAAAAGTTCGGTAAAGTATGGGGGTTTTACGACGGC
CGTCAGCCCGTATTGGCCATTATGGACCCCGTTATTATCAAGGCCATTCTGGTTAAAGAATGCTATACAGTGTTTAC
GAATCGTAGGAATTTCGGGCTAAACGGTCCTCTTAACTCAGCAGTGTCTATTGCAGCAGACGATCAGTGGAAACGTA
TTAGAACAGTGCTATCACCGACTTTTACTTCCGGGAAGCTGAAACAGATGTTTCCAATCATTAAGCAATATGGCGAC

TTGCTAGTCAAAAATATACAGAAAAAGGTAGATAACAAGGAGTTCATAGATATGAAAAATATATTTGGTTCATACAG
TATGGATATCGTTTTAAGCACTAGCTTTTCAGTAAACGTAGACTCACTAAATAACCCTAACGACCCATTCGTCACCA
ACGGGAGGAATCTGTTTACTTTCTCTTTCTTAAATCCCCTATTTCTGACGACGCTACTGTGTCCTTTCCTGATTCCA
ATACTGGACAAGCTGAATTTCTGCTTTCTGCCCATTTCAGTTCTAAACTTCTTTCAGGATGCCATTACATCTATCAA
GAAGAATAGACAAAAAGGCATTCACAAGGATAGAGTAGACTTCCTTCAGTTGATGGTTGACGCACAAGCCAATGATT
CAAAAGGGGGCGCAGACCACGGCTACAAGGAATTAACGGATACTGAGATCATGGCTCAGGGGTTAATTTTCATAATC
GCGGGGTATGAGACTACGAGCACAACTCTAATGTTTCTAGCGTACCACTTGGCGACCCACCCTGACGTGCAGACCAA
ATTACAGGAAGAAATTGATATCATACTTCCCAACAAAGCGCCACCTACATATGAGGCACTTATGCAAATGGAGTATT
TGGACATGGTATTGTACGAAAACCTGAGGTTGTATCCGGCAGCAGGTCGTATTGAGAGAGTTTGTAAAGCAACGACA
GAAATTAACGGAGTTACCATTCCGAAAGGTGTTGTGACAGTGATTCCAGCTTTCGTCTTGCATCGAGATCCCGAGTT
ATGGCCCGAACCGGACGAGTTCCGTCCCGAAAGATTTTCCAAGGAAAATAGAGAAACGCAAGACCCATATACTTTTC
TTCCATTCGGGGCGGGTCCAAGGAACTGTATAGGGATGAGGTTTGCACTTATTAACATGAAGTCAGTAATTACATTG
TTGCTTCAGAACTTTTCCTTCCGTACTTGCAAGGACACGCCAATCCCGCTACAAATAGATACTAGGGGATTCTTGAA
AACAACTAAGCCTGTCATCCTTAACTTGGTCCCCAGGGAAGCTCAGAAAACCGAGAAA

Seq. ID No: 218
>T5H_26
ATGTATGATACGTTCCTTGAATGGATCGAAAAGTATGGACCTGTAGTACGTGTCAACTCATCTCACTCAACCTTCGT
CATTGTGATTTCCCCCGAGGGAGTTAAGGAATTTCTTATGTCACCCAAGTATACCAAGGATAATTTCTATGAACGTA
TCGAAACCCTGTTTGGCGCCAGATTCTTAGGTAAGGGACTTGTGACCGATAGGGACTACGATCACTGGCACAAACAA
CGTAGGATGATGGACCCCGCGTTCAGCAGGACGTATTTAATAGGTTTAATGGGGACCTTTAATGAGACTGCCGAGGA
CCTTATGGACGTGCTAGGTGACAAAGCGGACGGTAAGTGCCAGGTGGGGATGCATGACATGTTATCTCGTGTGACTC
TTGATGTTATTGCGAAGGCCGCTTTCGGGATGGAGCTGAACAGTTTACATGACGATCAAACACCTTTCACAAGAGCG
ATAAGTACCGTAATGAAAGGGATGGTAGAGACCAGAAACCCTCTAGCACGTTACATACCGGGGAAACAAGCGTTTAT
ACGTGAGGTGAAAGAAAGTATCAAACTGCTGAGGGAGACTGGGAGAGAATGTATCCTTCAAAGGAGGAAGGAGATTC
AAGATGGAGAGGATATACCGATGGATATATTAACGCAAATCTTGAAAGGGGCCGAGATCGAGGACGGATGCAGTTTA
GAAGATTTGATCGATAACTTCGTTACGTTCTTTGTCGCGGGACAGGAAACGACAGCGAATCAATTGAGTTTCGCGGT
TATGGAATTAGCCAGGAACCCAGAGATTCTTACCAGAGTACAAACGGAAGTGGATGAGGTACTTGGATCTAAGCGTG
ATATCGAATATGAAGATTTGGGTAAGCTGCAGTATTTAAGCCAGGTTTTGAAAGAAACGCTTCGTTTATATCCGATT
GCTCCAGGAACATCCAGAGCCCTGGAGAAGGAGACCGTTATTGAAGGTGTAAGGGTTCCACCCGGCACTACACTTAT
GTTCAACAGCTATATAATGGGACGTATGGAGAAGTATTACCACGATCCTTTTATCTTCAATCCAGACAGGTTCCATC
CGGATGCGCCCAAGCCAAGCTGCGCTTATTTCCCATTTTCCCTTGGACCACGTTCCTGTATTGGGCAAGTTTTTGCT
AGGATGGAAGCAAAGGTCGTCATGGCCAAGCTATTACAGCGTTTCGAGTTTGAGTTGGTTGAAGGACAGTCTTTTAG
AATCATGGATACGGGTACCCTCAGGCCGATGGATGGGGTTATATGCAGACTTAGACCGCGTGCAGAGAGAAAATCCA
GGAAG

Seq. ID No: 219
>T5H_27
ATGGCCGCTCGTCCAAAGCCGGCAACCCCACCCAGCCCTCCAGCGTTACCGGTTATCGGTCATTTGCATCTTCTAAC
AGATATGCCACATCACACGTTTGCTGACCTTTCCAACAGCCTTGGTCCACTGATTTACCTTAGATTGGGACAGGTCC
CGACCATAGTGATTCATTCCGCACACCTAGCAAAATTAGTTCTTAGGACACACGATCACGCTTTTGCTAACCGTCCT
CAATTGATCTCCGCTCAATACCTGTCTTTTGGTTGTTCCGATGTGACTTTTTCAAGTTACGGAGCGTACTGGAGGCA
GGCTAGGAAAATATGCGTCACGGAGTTACTAAGTGCTAAAAGGGTACATTCTTTCAGATTGGTTAGAAAAGAGGAGG
TGGATAGACTGTTGGATGCAGTGCTAACGAGCAGTGGAAAGGAGGTAGACATGTCCCAAATGCTGTTTTGTTTGGCG
AATGATGTACTATGTAAGGTTGCGTTCGGACGTCGTTTTATGGCAGAAAAGGACGGGAAGGGGAAGAACCTTGGATC
TGTACTAATGGAAACCCAAGCGTTATTTGCTGGATTCTGCTTAGGAGACTTCTTTCCCAAGTGGGAATGGGTTAATA
GCATGTCAGGTTATAGGAAGAGACTGTTGAAGAATCTAAAGGACTTGAAGGAAGTCTGTGACGAGATAATTGAAGAG
CATTTAAAGAAGAAGAAAAAGAAAAACGGCACAGAAAACGCCGATGACGACGATGACTATAACGAGAAGGAGGACTT
TGTAGATGTTTTGCTTAGAGTGCAGAAAAGGGAAGACCTTGAAGTACCCATCACTGACGACAACTTGAAGGCGCTGG
TCCTAGACATGTTCGTCGCTGGCACTGACACCAGCAGTGCTACTTTAGAGTGGGTGTTCACAGAGCTTGCGAGACAC
CCCCGTGTTATGAAGAAAGCTCAAGAAGAGGTTCGTATGATTGCTTCCGGAAATGGTAAGGTCGATGAGTCAGACCT

ACAGCATCTTCACTACATGAAGGCGGTAATAAAGGAGACGATGAGGCTACACCCGCCAGTGCCCCTTTTAGTCCCAA
GGGAGTCAATGGAAAAATGCGCCCTAGACGGGTATGAAATCCCTGCCAAGACACGTGTCCTGATTAACACTTACGCT
ATCGGTCGTGACCCGAAATCATGGGAAAATCCCCTAGACTATGACCCAGAAAGGTTCATGGAGGACGACATTGACTT
TAAGGATCAAGATTTCAGATTCCTTCCTTTTGGCGGAGGGCGTAGAGGCTGCCCTGGTTATTCCTTCGGACTTGCGA
CAATTGAGATAACACTAGCCCGTCTTCTATATCACTTCGATTGGGCTTTACCGCATGGAGTAGAGGCGGACGATGTG
GACTTGAGCGAGGTGTTCGGACTAGCAACTCGTAAAAAGACTGCTCTGGTTTTAGTGCCGACGGCAAACAAGGATTT
CCAGTTTAGAGGGCATGACTTC

Seq. ID No: 220
>T5H_28
ATGGGGAAAAACAAAGTGCCGCCAGGACCGATCGGACTGCCGTTCATCGGTAACCTGCATCAGTTCGATACTCTAGC
ACCACATATCTACTTTTGGGAGCTATCCAAAAAGTACGGAAAAATCTTCTCATTTAAGTTGACGAGTAACGTGCCTA
TAATAGTGGTTAGCAGTGCAAAATTGGCGAAAGAAGTGCTGAAGACACAGGATCTAGTGTTTTGCAGCAGACCATCA
CTGGTCGGACAGCAAAAGCTGTCATACAATGGGCATGACATAGGCTTCGCACCATATAACGATTATTGGAGGGAAAT
GAGGAAAATTTGCGTACTTCACCTATTCTCTCTGAAGAAGGTGCAGCTATTTTCCCCCATTCGTGAGGATGAGGTTA
GCCGTATGATCAAGAAAATCTATCAACAAGCAGTGAACAGTCAAGTCACTAACCTATCCAACCTAATGATCAGCTTA
AATTCTACTATAATATGTCGTGTAGCGTTCGGAGTAAGATTCGATGAAGAAGCCCATGAGAGAAAGAGATTCAACTA
CATATTGGCAGAAGCACAGGCGATGTTTGCGGGTTTCTTCATGTCAGATTTCTTCCCATCACTTAGCTGGATAGATA
AGTTGACAGGGATGATCGATAGGCTAGAAAAGAATTTTAAAGACCTTGACGAATTTTATGAGGAACTTATAGAACAG
CACTACAATCCTAACAGGCCGAAAAGTATGGAGGGAGACTTTATTGATATTCTGCTACAGCTCAAAAAGGACCAGCT
AACTCCTATAGACTTGAGCCTGGAGGATATTAAAGGAATCTTGATGAATGTACTATTGGCAGGCTCTGACACCTCTA
GTAGTGTCATAATTTGGGCGATGACTATATTAATCAAGAACCCCAAGGCAATGAAGAAGGTGCAAGAGGAAATTCGT
AACCTAATCGGGAATAAAGGGATAGTCAATGAGGATGACATCCAAAACATGCACTATCTAAAGGCCGTCATCAAAGA
AACTCTTAGACTATTTCCGCCAGCACCTCTACTTATTCCTCGTGAGTCAATGAAGATCTCTACTCTGGAGGGGTACG
AGTTTCAGCCAAGAACTATCGTATACGTCAATGCATGGGCGATCGCTAGGGACCCGGAAATATGGGAAAACCCAGAA
GAGTTCATGCCAGAACGTTTTCTGAATAGTAACATTGATTTTAAGGGACAAGACTACGAACTAATACCATTTGGCGC
AGGGAGGAGAGGATGTCCGGGGTTGGCTTTGGGCGTAGCGTCAGTGGAACTAGCTCTATCTAATCTACTTTATGCGT
TCGACTGGGAGTTACCCTATGGCTTGAAGAAGGAGGACATAGACATTAATGGTAAACCGGGGATCACAGTTAACAAG
AAGAACGATCTGTGTCTTATCCCCAAGAAGTACTTT

Seq. ID No: 221
>T5H_29
ATGAAGTTAACGGGGAAGCGTTATAGATTACCACCAGGACCTGCCGGCGCTCCCGTGGTTGGGAATTGGCTTCAAGT
GGGTGACGACCTGAATCACAGAAACTTAATGTCTTTGGCGAAGCGTTTCGGTGATATTTTCCTGCTGCGTATGGGAG
TTAGGAACTTAGTGGTGGTGTCAACGCCCGAACTTGCAAAGGAGGTTCTACATACCCAAGGGGTAGAATTTGGGTCT
CGAACACGTAATGTTGTGTTTGACATTTTTACAGGCAAAGGACAAGACATGGTTTTCACTGTATATGGCGACCATTG
GAGAAAGATGAGGAGGATAATGACGGTACCTTTCTTCACAAATAAAGTAGTAGCACAGAACAGGGTTGGTTGGGAGG
AAGAAGCGAGGCTTGTAGTGGAGGATGTAAGAAAAGACCCCCGTGCAGCCGCCGAGGGAGTTGTAATCCGTAGACGT
TTACAATTAATGATGTACAATGACATGTTTAGGATAATGTTCGACACTAGATTTGAATCAGAGCAGGACCCCTTGTT
TAATAAGTTAAAAGCATTGAATGCGGAGCGTTCACGTCTTTCCCAATCTTTTGAATACAATTACGGAGACTTCATAC
CTGTCCTGAGACCTTTTCTGAGAGGGTACCTAAATCGTTGTCATGATTTGAAAACCCGTAGAATGAAGGTTTTCGAG
GATAACTTCGTACAGGAACGTAAGAAAGTAATGGCGCAAACCGGAGAGATCCGTTGCGCTATGGATCACATTCTTGA
AGCGGAGAGGAAAGGCGAGATAAACCACGATAATGTTCTTTATATTGTTGAGAACATCAATGTGGCTGCAATAGAAA
CGACACTATGGTCTATCGAGTGGGGCATTGCGGAGCTAGTTAACCATCCCGCGATTCAAAGTAAACTACGTGAAGAA
ATGGATAGTGTACTGGGCGCCGGAGTACCCGTGACCGAGCCTGACTTAGAAAGACTACCCTATCTACAAGCCATTGT
TAAAGAAACACTTCGTCTTCGTATGGCGATACCTCTGTTAGTCCCCCATATGAACCTAAATGATGGGAAACTAGCTG
GATATGACATCCCCGCGGAATCAAAGATATTAGTTAACGCCTGGTTCCTTGCAAATGACCCAAAGAGATGGGTCAGA
CCTGACGAGTTCAGACCAGAGCGATTTTTAGAGGAGGAAAAGACAGTAGAGGCGCATGGGAATGATTTTAGATTCGT
ACCATTCGGGGTAGGAAGAAGGTCATGCCCAGGTATCATTTTAGCTTTACCAATCATCGGCATCACGCTAGGCAGAT
TGGTCCAAAACTTCCAGCTGTTGCCCCCGCCTGGGCAGGATAAGATAGATACGACAGAAAAGCCGGGTCAGTTTTCT
AACCAAATCGCCAAACATGCTACAATTGTGTGTAAGCCCCTAGAAGCG

Seq. ID No: 222
>T5H_30
ATGAAACCGCGTGGCGCTAAATATCCTAATAGTCTTCCATGTTTGCCCTTCATCGGCTCACTGTTGCACCTTGCATC
ACACCTGGCTCCTCATATCTTATTTAATAAACTACAAGAAAAATATGGGTCACTATATTCATTCAAAATGGGATCTC
ACTACATCGTAATAGTAAATCATCATGAACATGCTAAGGAGGTATTACTGAAGAAGGGTAAAACATTTGGTGGTCGT
CCTAGAGCTGTGACCACAGACTTGCTAACTAGGAATGCTAAGGATATTGCTTTTGCGGACTATAGCCCCACGTGGAA
GTTCCATCGTAAACTTGTACACGCAGCACTGTCTATGTTTGGTGAGGGGACTGTAGCGATAGAGAAAATCATTTCAC
GTGAAGCCGCCTCCCTTTGTCAAACTCTTATTACATTTCAAGGGAGCCCCTTAGATATGGCCCCGGAGCTGACCCGT
GCGGTGACGAACGTAGTATGCGCACTTTGTTTCAACGCGAGATACAAAAGATGTGATCCTGAATTTGAAGAGATGCT
TGCATATTCCAAAGGAATCGTTGACACGGTCGCTAAAGATAGTCTGGTAGACATTTTTCCCTGGTTGCAAATATTTC
CCAATAAAGACCTGGAGATTCTGAAGAGATCAGTGGCCATTAGAGACAAACTATTACAGAAAAAGCTGAAGGAACAT
AAGGAGGCATTTTGTGGGGAAGAGGTGAATGACCTGCTGGACGCGTTATTAAAAGCAAAGCTGTCAATGGAAAACAA
CAACTCAAATATATCCCAGGAAGTCGGGCTTACAGACGACCATCTTTTGATGACCGTGGGTGACATATTTGGCGCGG
GAGTTGAGACAACCACCACAGTACTAAAGTGGGCTGTAGCGTATCTACTACACTATCCGAAGGTCCAAGCTAAGATT
CAAGAAGAATTAGATGTCAAGGTAGGTTTCGGGAGACATCCGGTTTTGTCCGATCGTCGTATCCTGCCCTACTTGGA
CGCGACGATTAGCGAAGTCCTAAGGATAAGACCAGTTGCCCCTCTACTGATCCCCCACGTCGCATTACACGAAAGCA
GTATAGGGGAATATACCATCCCACAGGATGCGAGAGTTGTAATAAATCTGTGGAGCCTGCATCACGATCCTAACGAA
TGGGAGAACCCCGAGGAGTTCATTCCTGACCGTTTTCTTGACGAAAATGGGAACCATCTTTACACTCCTAGCCAAAG
CTACCTACCCTTCGGTGCAGGCATTCGTGTTTGTTTGGGGGAGGCATTGGCTAAGATGGAGATTTTTCTATTTCTGT
CATGGATCTTACAGAGGTTCACGCTAGAAGTCCCTGCCGGGGATTCCCTGCCGGATCTTGACGGGAAGTTTGGCGTG
GTGTTGCAAGTGAAGAAATTTCGTGTAACAGCGAAGCTGCGTGAAGTGTGGAAGAACATCGACCTTACAACA

Seq. ID No: 223
>TAT2
ATGACGGAGGATTTCATTTCTTCTGTTAAAAGATCCAACGAAGAATTAAAAGAGAGAAAATCCAATTTTGGATTTGT
GGAGTACAAATCAAAACAGCTGACGAGCAGTAGCTCACACAATAGCAATAGCAGCCATCACGATGATGACAACCAAC
ACGGGAAAAGGAATATATTTCAAAGGTGTGTTGACAGTTTCAAAAGTCCCTTAGACGGATCTTTTGACACATCAAAT
CTAAAACGTACGCTAAAGCCCAGACATCTTATCATGATCGCTATAGGTGGGAGTATAGGAACTGGTCTGTTCGTCGG
GTCAGGTAAAGCCATTGCCGAAGGTGGTCCATTAGGGGTGGTCATTGGCTGGGCCATCGCTGGCAGTCAGATCATTG
GTACCATCCACGGACTAGGCGAAATAACCGTCAGATTCCCCGTGGTCGGCGCGTTCGCGAATTACGGAACTAGGTTC
TTAGATCCATCAATATCATTTGTCGTGTCTACTATCTACGTACTTCAATGGTTCTTCGTACTACCGCTGGAGATCAT
CGCAGCGGCTATGACAGTGCAGTACTGGAACTCATCAATTGATCCGGTCATCTGGGTTGCGATCTTCTATGCCGTCA
TAGTCAGCATAAATTTGTTTGGTGTACGTGGATTTGGAGAGGCAGAATTTGCATTCTCTACGATAAAGGCCATCACG
GTGTGCGGTTTCATAATACTATGCGTGGTACTTATTTGCGGCGGGGGTCCTGATCACGAATTTATAGGGGCCAAGTA
CTGGCACGATCCAGGCTGTCTAGCAAATGGGTTTCCGGGCGTACTTTCCGTACTGGTCGTTGCCTCATATAGCTTGG
GCGGTATAGAGATGACCTGCTTAGCGTCAGGGGAGACGGACCCCAAGGGATTACCGTCTGCCATAAAGCAGGTGTTC
TGGAGAATACTGTTCTTCTTTCTGATTTCTCTTACGCTTGTTGGATTTCTTGTCCCGTACACGAATCAGAATTTACT
GGGAGGGTCCAGCGTCGATAACAGCCCCTTCGTGATCGCAATAAAGCTGCACCACATCAAAGCCCTTCCCTCAATCG
TCAACGCGGTAATCTTGATCAGTGTTCTGAGTGTAGGTAACAGTTGTATCTTTGCGTCCAGTAGAACATTGTGTTCA
ATGGCCCACCAAGGCTTGATTCCCTGGTGGTTCGGTTACATAGATAGAGCAGGCAGACCTCTAGTGGGAATTATGGC
AAACAGTTTGTTTGGACTTCTAGCTTTCTTAGTGAAGTCAGGCTCAATGAGCGAAGTATTCAACTGGTTAATGGCGA
TAGCGGGGCTGGCCACTTGCATAGTCTGGTTATCTATAAATCTTTCCCATATACGTTTTAGGTTGGCAATGAAAGCG
CAAGGCAAATCCCTAGATGAGCTAGAGTTTGTAAGTGCAGTTGGTATATGGGGTTCCGCTTATTCTGCGTTGATAAA
CTGCTTGATATTGATTGCGCAATTCTATTGTTCTCTGTGGCCGATAGGTGGATGGACGTCAGGAAAAGAAAGAGCCA
AAATCTTTTTCCAAAACTATCTGTGCGCCCTAATTATGCTGTTCATTTTCATTGTTCATAAAATCTATTACAAATGC
CAAACTGGTAAGTGGTGGGGTGTCAAGGCTTTGAAGGACATTGACCTGGAAACAGACAGAAAAGACATTGATATAGA
AATTGTGAAGCAAGAGATCGCCGAAAAGAAAATGTACCTAGATAGCAGACCCTGGTATGTAAGACAGTTTCATTTCT
GGTGC

Seq. ID No: 224

>TMO_1
ATGAGCACGTTAGCAGATCAGGCATTGCATAACAACAATGTCGGCCCGATTATTCGTGCAGGGGACCTAGTCGAACC
TGTCATCGAAACTGCCGAAATCGACAATCCCGGGAAGGAAATAACGGTCGAAGACCGTCGTGCCTACGTAAGGATTG
CGGCAGAAGGGGAGCTAATATTGACTCGTAAGACTCTTGAGGAACAGCTAGGTAGACCGTTTAATATGCAGGAACTA
GAGATAAATCTAGCCTCATTCGCTGGGCAAATACAGGCAGATGAAGACCAAATCAGATTTTATTTCGATAAGACAAT
GGGCGGAGGGAGTGGTGAGGGAAGAGGCAGCCTTCTTACTTGTGGAGACGTGGAAGAAAACCCCGGGCCCATGTTTA
ACATACAATCTGACGATCTATTACATCACTTTGAAGCTGATAGTAATGATACACTATTGTCCGCCGCGCTTCGTGCT
GAACTTGTATTCCCCTACGAATGCAACTCAGGTGGTTGCGGCGCTTGTAAAATCGAACTACTTGAAGGAGAAGTATC
CAATTTATGGCCGGACGCACCTGGGTTAGCCGCCAGAGAGCTACGTAAGAATAGGTTTCTGGCCTGCCAGTGTAAGC
CGCTTTCTGATCTTAAAATTAAAGTTATCAATAGAGCCGAGGGCCGTGCTTCACATCCCCCGAAGAGGTTCTCTACG
AGAGTTGTAAGCAAAAGGTTTTTAAGTGACGAAATGTTTGAACTACGTCTGGAGGCCGAACAGAAGGTCGTTTTCTC
CCCCGGACAGTATTTTATGGTGGACGTTCCAGAGTTAGGCACCAGGGCATATAGCGCCGCGAACCCAGTTGATGGTA
ATACACTAACGCTAATCGTAAAAGCCGTGCCGAACGGAAAGGTGAGCTGCGCTCTAGCTAATGAGACGATCGAGACA
CTACAACTTGACGGACCTTATGGTTTGAGCGTGCTGAAGACTGCTGATGAAACGCAGTCCGTATTTATCGCGGGAGG
GAGTGGTATAGCTCCTATGGTCTCTATGGTCAATACTTTAATAGCACAAGGCTACGAGAAGCCTATTACGGTCTTCT
ACGGTTCTCGTTTGGAGGCGGAACTAGAGGCAGCCGAGACATTGTTCGGCTGGAAGGAGAACCTAAAATTGATTAAC
GTGAGCAGTTCCGTCGTTGGTAATTCCGAGAAAAAGTATCCCACCGGGTACGTGCATGAGATCATACCGGAATACAT
GGAAGGCCTGTTAGGTGCGGAGTTCTACCTGTGTGGTCCACCACAAATGATCAACTCTGTCCAGAAACTGTTGATGA
TTGAAAATAAGGTCCCATTCGAGGCGATTCATTTTGATAGGTTCTTC

Seq. ID No: 225
>TMO_2
ATGGCCATGCATCCGCGTAAAGATTGGTATGAGCTGACGCGTGCGACCAATTGGACACCTTCATACGTGACGGAAGA
ACAACTGTTTCCAGAACGTATGAGCGGTCACATGGGTATTCCACTTGAGAAATGGGAGTCCTATGACGAGCCGTACA
AGACTTCGTACCCGGAGTATGTTTCCATACAAAGAGAGAAGGACGCGGGTGCCTATTCTGTTAAAGCTGCCCTGGAA
CGTGCGAAGATTTATGAAAATAGTGATCCCGGTTGGATTAGTACATTGAAGTCCCACTATGGCGCTATTGCCGTGGG
TGAATACGCAGCGGTTACTGGTGAGGGTAGAATGGCTAGATTCAGTAAAGCCCCTGGCAATAGGAACATGGCTACTT
TCGGTATGATGGACGAGCTGAGGCATGGTCAATTACAACTATTTTTCCCCCACGAATACTGCAAAAAGGACAGGCAA
TTCGACTGGGCCTGGAGAGCTTACCATTCTAACGAATGGGCCGCCATCGCGGCCAAACACTTTTTCGACGATATCAT
CACAGGCAGGGACGCTATATCTGTCGCGATCATGCTGACATTTTCATTCGAAACGGGTTTCACAAACATGCAGTTTC
TAGGACTTGCTGCTGATGCAGCCGAGGCGGGAGACTACACGTTTGCGAACCTTATATCAAGCATTCAAACCGATGAA
TCCAGACATGCACAGCAGGGTGGTCCCGCATTACAACTGCTAATTGAAAACGGTAAGCGTGAGGAAGCGCAGAAGAA
AGTCGATATGGCAATTTGGAGGGCCTGGAGATTATTCGCAGTGCTGACCGGTCCGGTAATGGACTATTATACGCCCT
TGGAGGATCGTTCCCAAAGTTTCAAGGAGTTCATGTATGAATGGATCATTGGCCAGTTCGAACGTAGCCTGATTGAC
TTGGGTTTGGATAAACCCTGGTACTGGGATTTGTTCTTAAAGGACATCGACGAGCTACACCACAGTTATCACATGGG
AGTTTGGTACTGGCGTACCACCGCCTGGTGGAACCCTGCTGCAGGGGTAACTCCGGAGGAGAGGGATTGGTTGGAAG
AAAAATATCCGGGATGGAATAAGCGTTGGGGCAGGTGCTGGGACGTGATAACCGAGAATGTCCTAAATGACAGAATG
GATTTAGTGTCTCCAGAGACATTACCCTCTGTCTGCAACATGTCTCAGATTCCATTAGTTGGAGTACCGGGCGACGA
CTGGAACATAGAGGTCTTTAGTCTGGAACATAACGGCAGATTGTATCACTTCGGAAGTGAGGTAGACAGATGGGTTT
TCCAGCAAGATCCGGTCCAATATCAAAATCATATGAACATCGTTGATAGATTCCTGGCGGGCCAAATTCAACCCATG
ACATTAGAAGGTGCGCTTAAATACATGGGGTTCCAGTCCATCGAGGAAATGGGGAAGGACGCTCACGATTTCGCATG
GGCAGATAAGTGCAAACCGGCAATGAAGAAGTCAGCAGGTGGGGGCAGCGGAGAAGGCAGGGGAAGCCTGTTAACCT
GTGGCGACGTAGAGGAGAATCCCGGTCCTATGTCATTCGAAAAGATTTGCTCTCTGGATGATATCTGGGTTGGTGAA
ATGGAGACGTTCGAAACTTCAGATGGAACAGAGGTTTTGATCGTGAACTCTGAGGAACACGGCGTTAAAGCATACCA
GGCTATGTGTCCTCACCAGGAGATTCTTCTAAGTGAGGGGAGTTACGAGGGAGGTGTCATAACGTGTAGAGCGCACT
TATGGACATTCAATGACGGCACGGGCCACGGCATTAACCCCGATGACTGCTGTTTAGCGGAGTACCCTGTGGAGGTT
AAGGGTGATGACATTTATGTCAGTACCAAAGGTATACTTCCCAATAAGGCACACAGC

Seq. ID No: 226
>TMO_3

ATGTCTTTTACCAAGGTTTGTTCTGTCGGAGATATCTGGGAGGGGGAAATGGAGCCATTTACTGTTGATGGTCACGA
GATATTACTTGTAGGTGTGGAAGGGGGAGGGATCAAAGCCTTCCAGGGTATCTGTCCTCACCAAGATATAGCGTTGT
CAGAGGGGAAATTTGATGGGAAGAAACTTATCTGCCGTGCACACCTATGGCAATTCGATGCCAGCAACGGAAAAGGT
ATCAACCCAGACGATTGCGCCCTAGCCGAATACCCTGTCAAGGTTGACGGCGACGACGTCTACGTACAGACTGCTGG
GGTGGAAGCATTATTTGCACATAGCGGTGGTGGCTCAGGCGAAGGGAGGGGCTCCCTGCTAACGTGCGGAGATGTCG
AGGAAAACCCTGGTCCGATGGCCTTGCTGAATAGAATGGACTGGTATGACCTTGCGCGTACAACTAACTGGAGTCCA
AAATATGTAACGGAGTCTGAATTGTTTCCGCCCGAGCTGTCAGGAGACCACGGAATACCTATGGAGAAGTGGGAAAC
TTATGATGAACCCTACAAGCAGACTTATCCAGAGTATGTTAAAGTACAAAGGGAAAAGGATGCGGGCGCCTATAGTG
TTAAGGCAGCCCTGGAAAGATCCCAGATCTATGAGCGTTCTGACCCGGGGTGGCTAACGGTTATGAAGCAGCACTAT
GGCGCGATTGCACTTGGAGAGTACGCCGCTTCATCAGCAGAGGCCAGGATGATGAGATTTAGTAAAGCTCCAGGTAT
GAGAAACATGGCCACACTAGGGAGCATGGATGAAATACGTCACGGACAGATACAGCTATATTTTCCCCATGAGCATG
TTTCCAAGGATCGTCAGTTTGATTGGGCAGCTAAGGCTTTCCACACCAACGAATGGGCAGCGATAGCTGCGCGTCAC
TTTTTCGACGATATAATGATGACCAGAGATGCTATAAGCGTAGCAATTATGCTTACATTCTCATTCGAGACGGGGTT
TACGAATATGCAGTTTCTTGGCTTGGCAGCAGACGCGGCGGAGGCTGGAGATCACACATTTGCAAGTTTGATATCCT
CCGTACAAACAGATGAGAGCAGGCATGCCCAGATTGGAGGCCCTACGCTGCAAATCCTTATTGAGAACGGCAAGAAG
GCTGAAGCCCAAAAGAAGGTGGATATCGCTTTCTGGCGTGCCTGGAGACTGTTTTCCGTCCTGACTGGTCCAGTCAT
GGATTACTATACGCCGCTAGAACACCGTAAACAGAGCTTTAAGGAATTTATGCAAGAATGGATAGTCGCGCAGTTCG
AGAGAGCACTTAGCGATCTAGGGCTGGATAAACCCTGGTATTGGGATACGTTCTTACAGCAGCTTGACCAACAGCAT
CACGGGATGCACCTGGGGGTCTGGTATTGGAGACCAACGGTTTGGTGGAACCCTGCTGCTGGCGTTACACCAGCAGA
AAGGGATTGGCTGGAAGAAAAATACCCCGGCTGGAACGATACTTGGGGACAATGCTGGGACGTGATTATTGACAATT
TAGTTGACGGCAACATAGCTCAAACATATCCGGAGACATTGCCGATTGTCTGCAACATGTGCAATCTGCCTATAAAC
TGCACGCCCGGCAATGGTTGGGCCGTGCAAGATTACCCGCTGGAATACAACGGGCGTCTATACCACTTTGGGAGTGA
GCCTGATCGTTGGTGTTTTGAACAGGAACCTGAGCGTTATGCGGGCCACATGACCCTTGTCGATCGTTTTCTGGCTG
GTCTAGTGCAACCAATGGACCTGGGCGGGGCCCTGGCCTATATGGGGCTGGCGCCAGGGGAAATAGGGGACGATGCA
CACGGGTATAGCTGGGTCGATATTTATAAGAAAATGCGTATGAAGAAGGCAAGC

Seq. ID No: 227
>TMO_4
ATGTCAGTTGCATCAAGCGCGCAAGCGTATCATAACAACATGGTCGGACCCGTTATGAGAGCAGGCGACTTAGCGTT
AGCAGTAATTGAAGCGGCCAGGGTAGACAATCCAGGCAAGGAGGTTTTTGTTGATGATAAACGTGCTTATGTACGTA
TTCACACCGAGCAAGAAATGATATTGAGGAGAGAAACGATAGAGGAGGAACTTGGTAGACCATTTAAGATGAATGAT
CTTGAGGTAGATTTAAGTAGTTTTGCGGGGCAAATCGAATCTTTGGATGACGCCGTTAGATTTTACTTCACAAAGAA
GTTGGGTGGGGGATCTGGCGAAGGGAGGGGTAGCCTATTGACATGTGGGGACGTGGAGGAAAATCCGGGACCATCAT
CAAACCCAATAATCCATAATCAAAAGGACGGCTCAAGATTCGCACAAAGGGAGGGGGATACAATACTTCGTGCAGCT
CTGAGGGCTGGTGTTGGGCTGTCCTACGAGTGTAACTCCGGTGGCTGTGGAGGGTGCAAGTTCGAACTACTGGAGGG
AGAAGTGGACACGTTGTGGCCGGATGCTCCCGGGCTATCCGACAAAGACAGAAGGAGAGGGCGTCACCTAGCTTGCC
AGTGTCGTGCTAGGGGGCCTGTCAGTATAAAAGCCGCGACGGGAGCCGAGTACGTGCCTAAGGTCGTTCCACAAAGA
CAAACGGCTAGACTGGTTGGGAGCACTGATATTACGCACGACCTTAGAGAGTTTAGGTTCAGATCCGCGGCAGGAGC
CAGTTTCTTGCCCGGGCAGTTCGCCATGCTAGATCTACCCGGGCTTGCGAGTGCTAGGGCCTATTCTATGAGTAACA
CTGCCAATGATGATGGGGAGTGGCATTTCCAAGTGAGGAGAGTCCCGCATGGGCAAGGGACTCATGTATTGTTTGAC
AGGCTGGGGGTCGGAGACGAAATCGGGTTGGACGGACCTTACGGAGTAGCCTGGCTAAGAACTGGTGCTCCCAGAGA
TATAGTATGCGTAGCGGGTGGGAGTGGACTGGCTCCTATGGTCTCCATCGCAAGAGGAGCCGCCGCCGCCGGCATGC
TAAAGGATAGGAAACTGTATTTCTTTTATGGCGCAAGAACTCCTAGGGACGTTTGTGGAGCGGAGATGCTTGCACAG
CTGGATGGCTTCGGGGAACGTATCATATATCTTCCGGTGGTTTCTCTTCCGGGTGGCGAGGGTGAATGGCAGGGGGA
GACTGGTTACGTACACGATGCAGTAGCTAGGACACTACCCGGCTCATTGGCGGGTTTCGAGTTTTATTTTGCTGGGC
CGCCGCCGATGACGCAAGCCCTGCAAGAGATGCTGATGGTAGGACACAGGGTTCCTTTTGAGCAAATTCATTTTGAT
CGGTTTTTC

Seq. ID No: 228
>TPH_1

ATGCCGTCCAGACTGAATAAGGATGAATATCAGTTTTACATCGATCTTGATAATAAGTCCACACCAGCCTTGAATGA
AATTGTTAAATGTTTACGTCTTGACATTGGCGCGACAGTGCATGAGTTGAGCAGAGATAAGAAGAAGGATGCGGTGC
CGTGGTTTCCGAAAACAATACAAGACTTGGACAAATTCGCGAACCAGATCCTTTCCTACGGCGCTGAACTTGATTCT
GATCACCCTGGCTTCACTGATCCCGTCTATAGAGCTAGAAGGAAAGAATTTGCTGACATTGCTTTTCATTATAAGCA
CGGCCAGCCAATCCCTTGCGTAACGTACACAGAGGAAGAGAAGAAGACTTGGGGCACAGTATTCAAAGAGTTAAAGT
TGCTTTATCCGACTCACGCATGCTACGAACATAATCACGTTTTTCCTTTACTAGAGAAATATTGCGGTTATAACGAG
AACAACATTCCTCAATTAGAGGACGTCTCTAAATTTCTTCAGACCTGTACGGGATTTCGTCTGAGGCCGGTTGCAGG
CCTTCTGAGTTCCAGGGACTTCTTAGCTGGCCTAGCTTTCCGTGTATTCCACTCAACTCAATACATTCGTCACTGGA
GTAAGCCAATGTATACGCCTGAACCGGACATCTGTCACGAGCTATTGGGCCACGCTCCGCTTTTTGCAGATCCAAGT
TTCGCTCAATTCTCACAAGAAATCGGCCTGGCGTCACTGGGCGCTCCAGATGAATACATAGAGCGTCTAGCGACTCT
TTACTGGTTCACAGTAGAGTTCGGATTATGTAAGCAGGATGACAAAATTAAGGCCTATGGAGCCGGACTTCTTTCAT
CTTTCGGCGAGCTTCAGTATTGCCTTACAGATAAGCCTGAATTGAAGCCCTTTGAGCCAGAAAAGACGTCCCTTCAG
AAGTACCCGATTACAGAATTTCAACCTGTGTACTTCATAGCCGAGTCATTCGAAGACGCTAAAGAGAAAATGAGAAA
ATTCGCCACTACTATTCCTAGACCGTTCAGTGTGCGTTATAACCCGTATACTCAAAGTATCGAGGTCTTAGACAATG
TCCAGCAGTTAAAGAATTTAGCGGATTGTATAAACTCAGAGATCGGAACCCTGTGTTGCGCG

Seq. ID No: 229
>TPH_2
ATGATCGAGGATAATAAAGAAAATAAAGATCACAGTCTGGAACGTGGCAGGGCTTCATTGATATTCTCTTTAAAGAA
TGAGGTAGGCGGCTTGATAAAGGCGCTTAAAATATTCCAAGAGAAACATGTTAACCTATTGCATATCGAGAGTAGAA
AAAGTAAAAGGCGTAACTCCGAGTTCGAGATATTCGTAGATTGTGATATTAACAGGGAACAGCTGAATGACATTTTC
CATCTTCTAAAGAGCCACACTAACGTACTGTCTGTAAACCTACCGGACAATTTTACCCTAAAGGAAGATGGTATGGA
AACAGTCCCTTGGTTTCCCAAAAAGATAAGTGATTTGGACCACTGCGCTAACAGAGTGCTAATGTATGGAAGCGAGT
TAGATGCCGACCATCCTGGATTTAAGGACAACGTTTATAGGAAGAGGAGGAAGTATTTCGCTGATTTGGCGATGAAT
TATAAACATGGCGATCCAATTCCCAAGGTTGAGTTCACTGAAGAGGAGATTAAAACTTGGGGAACCGTTTTTCAAGA
ATTAAATAAGTTGTATCCGACACACGCATGTAGAGAGTATCTAAAGAACCTTCCTCTTCTGAGTAAGTATTGTGGGT
ATAGGGAGGACAATATCCCGCAACTGGAGGATGTTAGCAACTTCTTGAAAGAAAGGACCGGATTCAGTATCCGTCCA
GTGGCAGGCTATCTATCACCTAGAGATTTTCTAAGTGGGTTAGCATTCAGGGTATTCCATTGCACACAATACGTCCG
TCACAGTAGTGATCCGTTCTATACACCAGAGCCCGATACATGCCACGAGCTTTTAGGCCACGTTCCTTTACTTGCGG
AGCCGTCCTTTGCACAATTCTCTCAGGAGATCGGACTTGCATCCCTTGGCGCTTCCGAAGAAGCAGTTCAGAAACTT
GCCACTTGCTACTTCTTTACTGTAGAGTTTGGCTTGTGCAAGCAGGATGGGCAACTGAGAGTGTTTGGGGCGGGATT
GCTTTCCAGCATCAGCGAGTTGAAGCATGCGTTAAGCGGACATGCAAAAGTCAAGCCATTCGATCCAAAAATCACAT
GCAAGCAGGAGTGCTTGATCACAACTTTCCAGGACGTTTATTTCGTGTCTGAATCTTTTGAAGACGCCAAGGAGAAG
ATGAGGGAGTTCACGAAAACTATTAAGCGTCCTTTCGGGGTAAAATATAATCCGTACACAAGATCCATTCAGATTCT
GAAAGACACAAAATCAATCACGTCTGCGATGAACGAACTTCAACACGACTTGGATGTAGTTTCAGATGCCCTAGCTA
AGGTTTCCAGGAAGCCTTCTATA

Seq. ID No: 230
>TPH_3
ATGTCCGGTCTAGCACTTGACCGTAGTTCACAGCCACACGAGGTGAGAACACTGGAGGTCAACGAGTTGGACCCGAA
AGTTTTCGCGGTAGTCGAAGTAAGGAAGGACGAGCCAGGAGTACTGGGCGATGTGTTGAAAGTATTTACAGAGTCTA
GCATAAACATCACTAACATTGAGTCCAGGTTCAAGAGCTTCGCAAGGGATGGGCCAGCATTTCATATTGATTTTGAA
GGTGAAGCCAGGGAGCACAGAGTACAACGTGTGCTAAGGGATGTAAAATCTGTTCCAGGCGTCTCTCAGGTGACAGT
AATGGAAGAGAGGGAAGTCCCCTGGTTCCCAATTAATATTAGGGACTTAGATCTAACCACGGATACTCTTGACGGAG
GAACAGCCTTGATCAATGAGGATCATCCTGGCTTCAACGACCTAGCGTACAGGCAGAGAAGAGAGGAGATCGTTACC
GCTGCAAAGGAACATAGGCACGGAGACAGAATTGCCAGGGTACAGTACTTGGAACATGAAGTCGAGACTTGGAGGGC
AGTCTACGAACAACTTCGTGAATGTCATTCAAGGTGGGCTTGTACTGAGTACCTAGAGATGTTGCCGCAAATGGAGA
GGTTCTGCGGTTATGCACCCGGGAACATACCGCAACTAGCAGATATATCTGACTTTTTACAACAGAGAACTGGGTTT
ACCTTGCGTCCTATTACAGGTCTATTATCTGCTAGAGATTTTCTGAACGCACTTGCATTCAGAGTTTTCTACAGCAC
CCAATACATCAGGCACCACGGAAATCCCTTTTACACACCAGAACCGGATATTTGTCATGAGCTGATGGGACATGTAC
CTCTTTTTGCGAATGCCGCTTTTGCCGATTTTAGTCAGGAAATAGGGTTGGCATCCTTGGCTGCCAGCGACGATGAT

ATAGCCCGTCTTGCGGCAGTATATTGGTTCACAGTTGAATTTGGGCTGGTCAGACAGGGAGGAGAGGTGAAAGCGTA
CGGGGCTGGACTGCTGTCATCCTTTGGTGAGATGGAGTGGTCTTGTTCCAGAGAACCCAGCACGACCTGTAGAGAAA
TGGGTAGCGTAGCGGAGCTGCAAGCTCCTTCCATAGTCCCACTGGACCCGACCCAAGCAGGCAAACAGGCTTACCCT
ATAACCACGTATCAACCCTTGTACTTTTGTGCCGAGTCAATGCAGGACGCTAAGGCCAAGATCTCCCAGTTTTGCGA
TACTCTGACAAGGCCCTTCTTCCCTCAATATGATCCGTTAACGCAAAACATCCGTGTTACTAAAGCAGTACGTCGTG
CGAGGCGTATCTCTACTGTAGAAATGCAGATGGCAAAGCAACTTGATTATTTCGAGAAGCAA

Seq. ID No: 231
>TPH_4
ATGGCAGTTCCCTGGTTTCCGAAAACTATACAAGACTTAGACAAGTTCGCCAATCAAATCCTAAGTTATGGTGCCGA
ATTAGATTCAGACCACCCTGGGTTCACGGACCCTGTATACAGAACGCGTAGAAAGGAGTTTGCTGACATAGCGTTTC
ATTATAAGCATGGCCAACCAATTCCACGTGTAACGTATACAGAGGAGGAGAAGAAGACCTGGGGTACGGTGTTCAAG
GAGTTAAAATTGCTATACCCTACACACGCGTGTTATGAACACAATCATGTATTTCCACTTCTTGAAAAGTACTGCGG
GTACAATGAAAACAACATACCCCAGCTTGAGGACGTTTCAAACTTCTTGCAGACGTGTACTGGGTTTAGGCTTAGGC
CCGTGGCGGGTTTACTAAGTTCCAGGGATTTTCTGGCTGGACTGGCGTTTCGTGTATTTCACAGCACACAATATATC
AGACATTGGTCTAAACCGATGTATACGCCGGAGCCGGATATTTGTCACGAGCTTCTAGGTCATGCACCTTTATTTGC
TGACCCTAGCTTCGCACAGTTCAGTCAAGAAATTGGCCTGGCTTCACTAGGCGCTCCAGATGAGTATATTGAGAGAC
TTGCCACATTATACTGGTTTACCATTGAGTTTGGCTTGTGTAAACAAGACGACAAGATAAAAGCCTATGGGGCGGGG
TTATTATCTAGTTTCGGTGAGCTACAGTATTGTCTAACGGACAAACCGGACTTAAAGCCGTTTGAGCCAGAAAAGAC
ATCATTACAGAAATATCCCATAACCGAGTTCCAGCCTGTCTATTTCATAGCGGAGAGCTTTGAAGATGCCAAAGAGA
AAGTTCGTAAATTCGCAACGACGATTCCGAGGCCCTTCAGCGTACGTTATAATCCTTATACACAGAGTATTGAGGTG
TTGGATAATGTCCAACAGCTTAAAAATCTAGCCGATTGCATAAACTCCGAAATTGGTATCTTGTGTTGTGCACTGCG
TAAATTGGAG

Seq. ID No: 232
>TPH_5
ATGTTGATTTCCTTTACGCTTAACCTGGTCCACCAAAAGAAGAACAGCGAGTTTGAGATTTTCCTAGACTGCGACAG
TAACAGGGAACAACTAAACGAGATATTCCAGCTGCTAAGACCCCACGTCAATCTGATAACCATGAATCCGCAAGAGG
ACTTTAGCGTGGAAGAAGATGACATGGAGTCAGTACCCTGGTTTCCCATCAAAATTAGCGATTTAGATAAAAGCGCA
AACAGGGTGCTTATGTACGGCAGTGACTTAGACGCGGACCATCCTGGCTTCAAGGATAACGTATATAGAAGAAGACG
TAAATATTTTGCCGACGTCGCGATGAATTATAAATACGGTGACCCAATCCCACACATCGAGTTCACTGAAGAGGAGG
TGAAAACATGGGGGACTGTGTTCCGTGAATTAAATAAGCTACACCAAACCCACGCATGCCGTGAATACCTAAAGAAT
TTACCTTTGTTAGTCAAGCACTGTGGGTATAGGGAAGACAACATTCCACAACTAGAGGACGTTTCCAGATTTTTGAA
AGAACGTAGTGGCTTTACCATCCGTCCGGTCGCCGGCTACTTATCCCCGAGGGATTTCTTAGCTGGCCTAGCCTTTA
GAGTATTTCACTGCACTCAATATGTTCGTCACTCATCAGATCCGTTATACACGCCGGAGCCCGACACCTGCCATGAG
TTGCTAGGCCATGTACCTTTGTTAGCTGAGCCCAGTTTTGCACAGTTTAGTCAGGAGATAGGTTTGGCCAGCTTGGG
CGCCTCTGATGAAGCTGTCCAGAAGTTGGCCACCTGCTATTTCTTTACGGTGGAGTTCGGTCTGTGTAAACAAGAAG
GGAAACTAAAGGTGTACGGCGCTGGGCTTCTGTCTAGTATCTCAGAGTTGAAACACAGCCTGTCAGGTAATGCAAAT
GTGAAACCGTTCGATCCAATGGTCACTTGCTCACAAGAATGCATCATCACAAGTTTTCAGGAAGTTTATTTTTACTC
TGAATCATTTGAGGAAGCTAAAGAGAAGATGAGAGAATTTGCCAAGACAATAAAACGTCCGTTTGGTCTGAAATACA
ATCCGTACACGCAGAGCGTACAAATGTTGAAGGATACCCAGTCTATCACTACGCTAGTATCTGAACTTCGTCACGAG
CTGGATATCATCAGCGATGCTCTTAATAAAATGAACAAACAGCTAGGTGTC

Seq. ID No: 233
>TPH_6
ATGCATTCCCCCGAACCTGATTGCTGCCACGAATTATTAGGGCACGTCCCGATGTTAGCTGATAAAACTTTTGCCCA
GTTTTCTCAGGATATAGGATTGGCATCCCTAGGTGTCACGGACGAGGAAATAGAAAAGTTGAGCACACTATATTGGT
TTACGGTCGAGTTTGGTCTATGCAAGCAAGACGGTGAGGTTAAAGCATACGGAGCAGGTTTACTGTCCTCATATGGC
GAGTTACTACACGCGCTATCCGACAAACCTGAGGTAAGACCTTTCGATCCAGACGAGGCTGCGATTCAGCCGTACCA
AGATCAAAATTATCAACCAGTTTACTTTGTTTCCGAGTCTTTTACCGACGCCAAGGAGAAATTGAGAAATTACGCCA

GCAGGATTAAGCGTCCCTTCGCGGCGAGGTATGACCCATATACAGTCTCCATCGAGGTGTTGGACTCCCCTGGTCAA
ATTCAAAGTTCTTTGGAGGAACTTAAAGACGAGTTGCAGACACTGACTACAGCGTTAAATATACTATCT

Seq. ID No: 234
>TPH_7
ATGATGATATCTACTGAATCTGACCTAAGGAGGCAGTTGGACGAAAACGTACGTTCCGAGGCCGATGAGTCCACTAA
GGAGGAATGTCCATATATTAACGCAGTACAATCCCATCACCAGAATGTCCAGGAAATGTCCATAATCATATCGTTAG
TTAAGAATATGAATGACATGAAGTCTATCATCTCCATTTTCACAGATCGTAATATCAATATACTTCATATTGAAAGT
AGGCTGGGCAGGCTGAACATGAAGAAACATACTGAGAAGTCCGAATTTGAGCCCCTAGAGTTATTAGTTCACGTAGA
GGTACCCTGCATAGAAGTCGAAAGGTTGTTGGAAGAACTTAAATCATTTAGTTCTTACCGTATCGTCCAGAACCCGT
TGATGAATTTACCTGAGGCGAAGAATCCAACCTTGGACGACAAGGTGCCGTGGTTTCCCAGGCACATAAGTGACCTT
GATAAGGTGAGCAATAGTGTTCTTATGTACGGGAAAGAGTTAGATGCCGATCATCCCGGATTCAAAGATAAGGAGTA
CAGGAAAAGGAGGATGATGTTTGCCGACATAGCTCTAAACTATAAGTGGGGCCAACAGATTCCCATTGTTGAGTATA
CTGAAATAGAGAAGACAACTTGGGGAAGAATTTATAGGGAATTAACCAGACTATATAAGACGTCCGCATGTCACGAG
TTTCAGAAGAATCTAGGGTTACTACAAGATAAGGCGGGGTACAACGAATTTGACCTGCCGCAACTACAGGTAGTCTC
CGACTTTTTGAAGGCCAGGACTGGATTTTGTCTTCGTCCCGTTGCGGGGTATCTTAGTGCAAGGGATTTCCTTAGCG
GACTAGCCTTTAGGGTCTTTTACTGTACTCAGTATATTCGTCACCAGGCCGATCCGTTCTACACACCCGAACCAGAT
TGCTGCCATGAGCTTCTTGGACACGTTCCGATGCTGGCGGACCCCAAGTTTGCTCGTTTCTCACAAGAAATAGGGCT
TGCGTCACTAGGTACCAGTGATGAAGAAATTAAGAAATTAGCAACTTGTTATTTCTTTACGATAGAGTTCGGCTTGT
GCAGACAGGACAACCAGCTAAAAGCGTACGGTGCAGGTTTGCTTAGTTCCGTCGCAGAATTACAGCACGCACTTTCC
GACAAAGCGGTTATCAAACCGTTTATTCCTATGAAGGTAATAAATGAGGAATGTTTGGTTACAACATTCCAAAATGG
CTATTTCGAGACCAGTTCATTTGAAGATGCTACTCGTCAAATGAGGGAATTTGTCAGAACGATCAAGCGTCCTTTCG
ATGTCCACTATAATCCGTACACTCAATCTATTGAAATAATCAAAACCCCAAAATCTGTAGCAAAATTAGTACAAGAT
TTACAATTTGAGCTAACCGCAATTAATGAGTCCTTGCTGAAAATGAACAAGGAGATTCGTAGTCAACAGTTTACCAC
AAACAAGATTGTGACTGAAAATCGTAGTTCC

Seq. ID No: 235
>TrpHalo_1
ATGTCTACAGCTTCAAAGAACATAGATATTACAAGATTCCCCAAGAAATACGATGCGGCCACGAAAGACTCTGATTT
TTATGATGTTGTAATCGTAGGTGCTGGCCCAGGTGGCAGTACAACTGCTTACTACTTGGCAAAAGAAGGGAAAAAGG
TTTTATTGCTTGAGAAGAAAAAGTTCCCCAGAGATAAGATATGCGGTGACGCTATATGCAAACTAGCCATAGAGATG
CTTATGGATATGGGTGTCTACGAGGGATTAGTTAGGGAGAAGAAAGCTCGTGTAGCTCACAACGGAGGATTAGTTTC
TCCCTCCGGCCTTAGTTTCATCGGGAATACGTACCTAAAGCCGGGGGAGATCCCTGCTGCAGCTGCGTGCAAACGTA
TGGTCCTGGACGAAGCCATCGCTAAGGCCGCTATTGGGGCCGGAGCCGAACTAAAAGAAAATTCCCCAGTGACAGAC
GCTGTCTTCGATAGTAGTACCGGGCTGTGGACGATTAGCATTGAGGGCAGTGATGTGAAACATATGGGCCGTGTGTT
AGTATGTGCAGACGGTGCTCCTAGTAAACTAGCTACGCAATTAGGCATCGTAAAACAGGCGCCCCAGGGGGTGTGTA
GCAGAGCGTATATTAAGGAAGGTACACATAGATTCAGAGCGGACGGCGTCGTATTTTATCCGAGAAATATCTTACCA
GCATATGCTGCCCTTTTCAGGCACATAGACGACACTGTCGCGTACTGCACCTATATATTACCCTTTAACCCAAAAGT
TACGACAGACGACCTGTCATATTGGCATCATCGTCTATTAGAGGAGGACCCTTCTATTTCTCAAGCAGTCGGTAAAA
ATGCTGATATGGAGAGAATGAAAGCCTGGGGTCTGCGTATGGGAGGTGAACCCGTGACGTATGGTAACCATGTTTTG
GTAGTTGGAGACGCAGCTGGAATGATCGATCCACTAACAGGGGAAGGTATACACCACGCTATGGATGGCGGTCGTAT
CGCCGCACATTTCCTATGCGAAGCCATAGCGGTTGGGAATTTTGACAAGGAAGTCATGAAGGAATACCAGAACAGGT
GGCTTTACACGTTCGGAAATGACTACAAATGGAGCCAGGCCATATGCCACTTTTTATATCGTTTCCCTATCTTCATC
GATGCTACTGCCGCCGCCGCACAAAGGCGTGGGAACAACTTTCTTGCTCTTTACGCAGACATTATGACCGGGAGGAT
ACCTAAGGCCAATATTTTCCGTCCCGATATTAGTTTGCCCATAGCATTTGAAGTCTTGGTACTATTGTGGAAGATGA
TGTTCACTGGGGGCGGAGGTAATAATAAAATGAAGTCACAG

Seq. ID No: 236
>TrpHalo_2
ATGAGTACGGCGTCTAAAAATATCGACATCACCCGTTTTCCGAAGAAGTATGACGCGGCGACCAAGGACAGTGACTT
CTATGACGTCGTCATCGTCGGAGCAGGGCCAGGGGGATCAACGACGGCTTATTATCTAGCGAAAGAGGGAAAGAAGG

TACTACTTTTGGAAAAGAAAAAGTTTCCGAGGGACAAAATTTGTGGAGATGCTATATGCAAGTTAGCGATCGAGATG
CTTATGGATATGGGCGTCTACGAGGGTCTGGTAAGGGAGAAAAAGGCAAGAGTTGCACATAACGGGGGTCTAGTTAG
TCCAAGTGGTTTAAGTTTCATAGGCAACACGTACCTTAAACCAGGCGAGATTCCTGCCGCCGCAGCGTGCAAAAGAA
TGGTATTGGACGAGGCTATAGCAAAGGCAGCTATAGGGGCAGGGGCTGAACTAAAGGAAAATAGCCCTGTGACTGAC
GCCGTGTTCGATTCTAGTACCGGTCTGTGGACAATCTCTATCGAAGGTTCAGACGTAAAACATATGGGCAGGGTCCT
AGTTTGTGCGGACGGAGCACCCAGCAGACTTGCGATGCAATTAGGCATCGTCAAGGGGACACCTAAATGCGTTTGCT
CCAGAGCGTACATAAAAGGCGGGACCCACCGTTTTAAAGAGGACGGAATGGTATTCTATGTACCTTCTATCCTACCG
GGGTACGTGGCTTTGTTGCGTCATATCGACGACCAACTTACATACTGCACCTACATTCTACCAGGGAACCCAAGAGC
AACTACGAAGGATTTGTCTTATTGGCATCATAGACTTCTTGAGGAAGACCCAAACATAAGCCAGGCAGTCGGGAAGA
ACGCCGAACTTGAGAAGATGAAAGCCTGGGACTTACGTGTGGGCGGTGAACCAGTAACGTACGGTAATCATGTCCTT
GTGGTCGGTGACGCGGCCGGTATGATAGATCCATTGACAGGTGAAGGCATCCACCACGCAATGGATGGGGGCCGTAT
TGCCGCTCACTTTTTGTGTGAAGCTATTGCAGTGGGTAATTTCGATAAGGAAGTTATGAAAGAGTACCAGAATAGAT
GGTTGAAGGCATTCGGAAATGACTTCAGATGGTCTCAAGCTATCGGCAATTTTCTTTACCGTTATCCGATTTTCATC
GACGCAACCGCCGCTGTCGCCGAGAAGAAAGGAGACAGGTTCTTAGCAAGGTGGGCAGATATTATGGCAGGAAGAAT
ACCAAAGATTAGTGTCCTTAGGCCGCAGTTTTTGCTGGCTGTAGGCTTCCAAGCTCTCCTACTATTCTATAAGAAGA
TTTTCAAAGGGGGTTATGGAAAGAAGACAAAGATTCTA

Seq. ID No: 237
>TrpHalo_3
ATGAGTAGCTTAATCGCGCCAAAAGTAGATACCATAGATATTACAAGGTTCCCCAAAAAGTACGACCCTGCAGCGGA
GGACTCCGATTTCTACGACGTAGTAATAGTCGGAGCGGGTCCCGGTGGCAGCACGACTGCGTATTACCTAGCGAAGA
AGGGTAAGAAAGTGTTACTACTAGAAAAGAAGAAATTTCCTAGAGACAAGATCTGTGGTGACGCAATCTGTAAAACG
GCAATTGAAATTTTAATGGATATGGGGGTCTACGGCGGACTGATCAGGGAACAGAAGGCTTACATGATTGACTACGG
GGGACTTGTGTCCCCTTCCGGTTTGAGCTTTGTCGGCCACACCCATGAGCTTTTCGGTGAAATACCAGGAGCTGTTG
TGTGCAAGAGAGTTGTGTTGGACAAGGTAATATCAAGAACGGCACAATCAGCTGGAGCTGAATTGTTGGAGAACTCC
CCTGTAACAGATGCAGTTTTCGATTCCAGCACTGGTCTTTGGACTATTTCTATTGAGGGCTCAGACGTGAAGCATAT
GGGTAGAGTACTGGTATGCGCTGACGGAGCACCTTCCAGATTGGCCATGCAACTAGGAATTGTCAAGGGGACACCAA
AGTGTGTTTGTAGTCGTGCCTATATTAAAGGCGGCACCCATAGGTTCAAAGAAGACGGCATGGTGTTCTATGTCCCT
TCAATACTTCCCGGATATGTCGCCCTTCTTCGTCATATAGATGACCAGCTGACTTACTGTACGTATATCTTACCTGG
AAATCCCAGAGCTACCACTAAGGACCTTTCTTACTGGCACCATAGGCTGCTAGAGGAGGACCCCAATATTAGCCAAG
CAGTTGGTAAAAATGCCGAATTAGAGAAAATGAAGGCCTGGGACCTAAGAGTGGGCGGAGAACCCGTCACGTACGGC
AATCATGTACTGGTAGTAGGTGACGCTGCGGGCATGATCGATCCATTAACCGGAGAGGGAATCCACCATGCTATGGA
CGGAGGCCGTATCGCGGCTCACTTTCTTTGCGAGGCTATAGCTGTCGGCAACTTTGACAAAGAAGTCATGAAGGAGT
ATCAAAATCGTTGGTTAAAGGCATTTGGTAATGATTTTAGATGGTCCCAAGCGATTGGGAACTTCCTATATAGATAC
CCGATATTTATAGATGCCACTGCGGCCGTGGCTGAGAAGAAGGGTGACAGATTCCTGGCCCGTTGGGCTGACATCAT
GGCGGGTAGGATTCCCAAAATATCCGTCTTGAGACCGCAATTCCTGTTGGCAGTAGGATTCCAAGCCCTTCTGCTTT
TCTATAAGAAGATTTTTAAAGGGGGCTACGGTAAGAAAACTAAGATCCTT

Seq. ID No: 238
>TrpHalo_4
ATGTCTGGAAAGATTGACAAGATTTTGATTGTGGGAGGTGGAACCGCGGGATGGATGGCTGCAAGTTATCTAGGGAA
GGCGTTACAGGGCACTGCGGACATTACGTTATTACAGGCTCCAGATATCCCGACACTAGGGGTAGGAGAAGCAACCA
TACCTAATTTACAGACCGCGTTCTTTGATTTCTTGGGAATACCCGAGGATGAGTGGATGAGGGAATGCAATGCCAGT
TACAAAGTTGCGATAAAGTTCATCAACTGGCGTACGGCGGGCGAAGGTACCTCTGAGGCGAGGGAATTAGATGGCGG
GCCGGACCATTTTTACCACAGCTTTGGCCTTCTAAAGTACCATGAGCAGATTCCTCTGTCACACTACTGGTTCGACC
GTTCCTACAGGGGCAAGACGGTTGAGCCGTTCGACTACGCGTGCTACAAGGAGCCGGTTATTTTAGATGCTAACAGA
AGCCCAAGGAGATTGGACGGCTCCAAGGTCACTAACTATGCGTGGCATTTTGATGCGCACCTAGTCGCTGATTTTCT
GAGAAGATTTGCAACTGAGAAGTTAGGGGTAAGGCATGTGGAAGATAGAGTTGAACACGTGCAGCGTGATGCCAATG
GGAATATTGAGTCCGTGAGAACGGCTACCGGAAGAGTATTTGACGCTGACTTGTTTGTAGACTGCAGTGGATTTAGG
GGGTTACTAATTAATAAGGCAATGGAAGAACCCTTTTTAGACATGTCCGACCATTTGTTGAATGATTCAGCCGTAGC
GACTCAGGTACCACACGACGATGATGCTAACGGAGTGGAACCTTTTACCAGTGCAATAGCCATGAAGTCAGGTTGGA

CATGGAAAATTCCCATGCTGGGCAGATTCGGGACAGGTTATGTGTATTCTTCCAGGTTTGCCACTGAGGACGAGGCA
GTGCGTGAATTTTGCGAGATGTGGCATTTAGATCCAGAGACGCAGCCTTTGAATAGGATTAGGTTTCGTGTCGGGCG
TAACCGTAGGGCTTGGGTCGGTAATTGTGTGTCCATTGGAACTTCTTCTTGCTTCGTAGAGCCACTTGAGTCAACTG
GAATTTACTTCGTGTATGCCGCACTATACCAGCTTGTCAAGCACTTCCCTGACAAGAGTCTGAACCCTGTTCTAACT
GCTCGTTTCAACCGTGAAATTGAGACAATGTTCGATGATACCCGTGACTTCATTCAGGCACACTTTTACTTCAGTCC
ACGTACCGATACACCGTTCTGGCGTGCTAACAAAGAGCTGAGACTTGCAGATGGGATGCAGGAAAAGATCGATATGT
ATAGGGCTGGCATGGCTATCAATGCGCCTGCCAGTGACGATGCTCAGCTTTACTATGGAAATTTTGAAGAGGAGTTC
CGTAACTTCTGGAACAACTCAAATTACTACTGTGTTTTAGCGGGATTAGGTCTTGTTCCCGATGCCCCTAGCCCCAG
GCTGGCCCACATGCCTCAAGCAACGGAATCAGTGGACGAGGTGTTTGGCGCCGTTAAGGACAGACAACGTAACTTGC
TTGAGACCCTACCATCACTGCATGAGTTCCTTAGGCAGCAGCATGGGAGG

Seq. ID No: 239
>TrpHalo_5
ATGGACGAAATAGACGACCCCCGTATAAGGAGTGTAGTGATAGTGGGAGGCGGTACCGCGGGGTGGATGACCGCCGC
CGCTTTAGTTCAGCACTTTAGAACCGCGCCTCTAAAAATAACGGTGGTGGAGAGTTCCGATATTGGGACTATTGGTG
TAGGTGAAGCTACGATTCCGACCATTCGTAGATTCTATGGCCAACTAGGACTTAGGGACGACGATGTGATGAGAGCT
ACGCAAGCGACATGTAAATTGGGCATCAGGTTCTTAGATTGGAGTGGACCTGGGTCCGACTTTATTCATCCCTTTGG
TCTTTATGGCCAAGATGTTAAGGGCATTGGCTTCCACCACTATTGGCTTAAACAAAGGCGTGCCGGAGATGCAGCGC
CGTTGGCTGCGTACAGCTTAGGAGCGGCACTGGCTGCCGGTGGCAAATTTACGTTGCCCAGTCCTCATCCACCCTCT
CAACTGTCTGTATTTGATTGGGCATTGCACCTTGACGCAGGCCTATTCGCCCAGCACCTAAGAGCGTACGCAGAAGC
CGGTGGCTGTGCCAGAATTGATGCCCGTATAAGATCAGTTGAGCTTAGACCAGAAGACGGCTTTGTCAGAGCCCTTA
CATTGGACGACGGAAGAGAGGTAGAGGGGGACTTGTTCGTGGACTGTAGCGGGTTCAAAGGACTTGTTATCGGTGAG
GCGCTGGGAGTTGGGTTCGAAGACTGGGGGAGGTGGTTACCTTGCGACGCAGCGTATGCTGTACAGTCTGAAAACAG
GCCAGGCGATGCGCCAGCTCCGTTCACACGTGTTACAGCTAGAAGCGCCGGATGGCAGTGGGGCATACCGCTTAGAC
ATAGGGCGGGTAACGGGTTAGTATTTTCAAGTGCCCATCTGTCAGACGACCAAGCCCTTGCTGAGTTGATGCCGCAT
TTATTAGGAGACCCCCTTACCGAGCCGAGGCGTATACCTTTCCGTCCCGGCCGTAGATCTCAAGCGTGGGCAAAGAA
TTGTGTGGCTATTGGCCTTTCATCCGGCTTCTTAGAGCCCCTGGAGTCCACTAGCATCGCTCTGATCGAGACAGGAA
TTGAACGTTTGAAAGCCCTGTTCCCTGACCGTAGGTTCGCGCAACCTATCCTAGACGAATTTAATGATCAAACGGCG
AGAGAGATGGAGAGGGTTCGTGACTTTATCATACTACATTATAAACTTAATAGGAGAACTGACACCGACTTCTGGAG
AGACTGCAGGGAAATGCCCGTTCCCGAAACATTGGAGAGAAAGATAGCCCTATGGACGGCAAGAGGACAGTTTGTAC
GTTATCGTTGGGAAATGTTCCATCCAGCAAGCTGGTTAGCGATCTATGATGGTTTCGGCCTTTATCCCGATCACCAT
GATCCAGCGGTGGATGCTATGGACCCCGCCTATTTAGCTAGGTCATTAGCGGAAATGAGAGCTAACATTGCGGACTT
AGTTGCGCGTACCCCCGAACATGCTCAGTTCCTTGCTGGGTTGGACCCGGCGGCTTCGGCCGCT

Seq. ID No: 240
>TrpHalo_6
ATGATACGTAGTGTCGTGATTGTAGGAGGAGGAACGGCCGGTTGGATGACGGCATCTTATCTAAAAGCTGCCTTCGA
TGACCGTATCGACGTAACTCTGGTAGAATCCGGTAACGTGCGTCGTATTGGAGTCGGGGAAGCTACTTTCTCAACAG
TAAGACACTTTTTCGATTATCTGGGGTTGGATGAGCGTGAGTGGTTGCCACGTTGTGCTGGTGGATATAAATTAGGT
ATACGTTTTGAAAACTGGTCCGAACCAGGGGAATACTTTTATCACCCATTCGAGAGGCTAAGGGTAGTCGATGGTTT
TAATATGGCAGAATGGTGGCTTGCGGTGGGAGATCGTCGTACTTCTTTCAGTGAGGCTTGCTACTTAACACATAGGC
TTTGTGAGGCCAAGAGGGCTCCAAGGATGCTAGACGGGAGTCTTTTCGCATCTCAAGTAGATGAGTCATTAGGGAGA
AGCACACTAGCAGAACAGAGAGCCCAATTTCCCTATGCTTATCATTTCGATGCCGACGAGGTGGCGAGATACCTATC
AGAATATGCAATAGCGAGGGGCGTGAGGCACGTAGTGGACGACGTACAACATGTAGGTCAGGATGAAAGGGGATGGA
TCAGTGGTGTTCATACTAAGCAGCACGGAGAGATTAGTGGAGATCTTTTTGTAGACTGTACAGGGTTCAGAGGCCTA
CTTATAAACCAAACTCTGGGTGGTAGATTTCAGAGTTTTTCCGATGTGCTTCCTAACAATAGGGCTGTTGCTCTGAG
GGTGCCAAGAGAGAACGACGAGGACATGCGTCCTTACACAACGGCTACCGCAATGTCTGCCGGTTGGATGTGGACTA
TACCCCTTTTCAAGAGGGATGGAAACGGTTACGTGTATTCAGACGAGTTCATTTCTCCTGAAGAAGCCGAGCGTGAG
TTGAGATCAACAGTGGCTCCCGGCAGAGATGACCTGGAGGCAAATCATATCCAGATGCGTATCGGGAGGAACGAGCG
TACGTGGATTAATAACTGCGTGGCCGTCGGTTTGTCAGCAGCGTTTGTAGAGCCATTGGAGTCAACCGGAATCTTTT
TCATCCAACACGCAATCGAGCAGCTGGTTAAGCACTTCCCCGGAGAGAGATGGGACCCTGTCCTAATAAGTGCATAC

AATGAACGTATGGCACACATGGTTGACGGGGTGAAGGAATTTTTAGTATTACACTACAAGGGTGCCCAGCGTGAAGA
CACTCCTTACTGGAAAGCAGCCAAAACCAGGGCGATGCCGGACGGACTGGCCAGGAAACTAGAGCTTAGCGCGTCTC
ACTTGCTAGACGAGCAGACTATATACCCCTACTATCATGGATTCGAAACCTACAGCTGGATCACCATGAACTTAGGG
TTAGGCATAGTGCCCGAGAGACCTCGTCCAGCTCTTTTGCACATGGACCCTGCTCCCGCCTTAGCAGAATTTGAACG
TCTGCGTAGAGAAGGAGACGAGCTAATTGCAGCCCTTCCATCATGCTACGAGTATCTGGCTAGTATTCAA

Seq. ID No: 241
>TrpHalo_7
ATGTTAGAATCTATTGTGGTCGTAGGAGGTGGTACCAGTGGATGGATGACCGCTTCATATCTGAGTGCTGCCTTTGG
TGAGCGTATTTCTGTAACAGTTGTAGAATCAGCGAGGGTCGGGACGATAGGGGTTGGCGAAGCTACCTTCAGCACAG
TAAGACACTTTTTCGAATACCTGGGGTTAAGTGAGGAAACATGGATGCCAGCTTGCAATGCTACGTATAAACTAGGG
ATCAGATTTGAAAACTGGAGAGCGCCAGGGCACCATTTTTATCATCCTTTTGAAAGACAGAGGGTAGTGGACGGGTT
CACCCTGCCAGACTGGTGGCTGGCCGATGGCGGGGCCACCGAGAGATTCGATAAGGAATGCTTCCTGGTTGGCACAC
TGTGTGACACTATGAGATCCCCTAGACACATGGACGGTGCTTTATTTGAGGGCGATCTAACGGACAGGCCCGCTGGA
CGTTCAACGCTGGCCGAGCAAGGAACGCAGTTTCCTTATGCATATCATTTCGATGCGGCGTTGCTTGCCGACTTTTT
ACGTGATTACGCTGTGGCGAGAGGGGTCTTGCATGTTGTGGACGATGTGGTTCACGTGGCTAGGGACGAAAGGGGAT
GGATCTCTCATGTTGCCACAAGGGGATCTGGAGACCTGGCGGGCGATTTGTTTGTCGATTGTACTGGATTCAGAGGA
TTACTTATTAACGACGCTTTGGACGAGCCGTTTGAGTCTTATCAAGACACCCTACCGAATGATTCCGCAGTGGCTCT
GCGTGTACCAGTTGACATGGAGCGTGAAGGCTTGAGGCCATGTACCACATCAACCGCTCAAGCGGCGGGGTGGATAT
GGACTATTCCACTTTTTGGACGTGTTGGAACCGGCTATGTCTACGCGAGGGATTACTGCACACCTGAGGAAGCTGAA
CGTACGCTGCGTCGTTTCGTAGGCCCGGCTGCCGACGACTTAGAGGCCAATCACATTCGTATGAGAATTGGAAGGAG
TAGACGTTCATGGGTCAACAATTGCGTCGCAGTGGGTTTAAGCAGCGGGTTCGTAGAACCGTTAGAAAGTACCGGTA
TCTTCTTTATTCAGCACGCAATAGAACAACTTGTTAAACACTTCCCCGATGCGGACTGGGACCCCGCTCTGCGTTCC
GCCTACAACACCTTGGTTAACAGGTGCATGGACGGAGTAAGGGAGTTCTTGGTCTTGCACTATTACGGGGCAGCGAG
AGCTGACAATGAATATTGGAGAGATACAAAGACGAGGAAAATCCCCGACTCTTTGGCGGAGAGGGTGGAGCAGTGGC
GTACCAAATTGCCCCATCCAGAGAGTGTGTACCCTCACTATCATGGTTTTGAAGCGTACTCCTACGTCTGTATGGTG
CTTGGACTGGGTGGTATTCCTCTAAAACCGAGCCCTGCCCTACGTATGCTTGACCCCTCCGCAGCCCAGCGTGAGTT
CCGTCTGTTGGCGACTCAAGCAGAGGATCTTAGGCGTACCCTACCGTCTCAATATGCTTACTTCGCACAGTTTAGA

Seq. ID No: 242
>TrpHalo_8
ATGAATAAACCTATAAAGAATATAGTGATCGTCGGAGGCGGAACCGCAGGCTGGATGGCAGCGTCATATTTAGTAAG
GGCCCTTCAGCAGCAGGCCAATATAACCTTAATTGAATCCGCTGCTATACCTCGTATCGGTGTCGGGGAGGCCACCA
TACCCAGCTTACAGAAGGTCTTCTTCGACTTCCTTGGTATTCCCGAGAGAGAGTGGATGCCTCAGGTTAATGGAGCG
TTCAAGGCAGCTATCAAGTTTGTTAATTGGAGGAAAAGCCCGGACCCTTCCAGGGATGACCATTTCTACCATCTATT
CGGCAATGTTCCGAATTGCGATGGCGTGCCCCTAACGCACTACTGGTTGCGTAAGAGGGAGCAGGGATTTCAGCAAC
CGATGGAGTACGCTTGCTATCCCCAGCCTGGGGCGTTAGACGGCAAATTAGCGCCATGCCTTTCCGATGGCACACGT
CAAATGAGTCACGCTTGGCATTTCGACGCTCACTTGGTAGCTGATTTTCTGAAGAGGTGGGCCGTTGAGAGGGGCGT
CAATAGAGTAGTTGATGAGGTCGTCGATGTCAGATTGAATAACCGTGGCTACATAAGTAATTTACTGACCAAAGAGG
GGAGAACTTTGGAAGCGGACTTATTCATTGATTGCTCCGGGATGCGTGGACTGCTGATTAACCAAGCGTTAAAAGAA
CCATTTATAGATATGTCCGATTACTTACTGTGTGATTCTGCCGTTGCGAGTGCCGTACCCAATGACGACGCCCGTGA
CGGGGTTGAGCCCTATACGTCTAGTATCGCGATGAATAGCGGCTGGACGTGGAAGATACCAATGTTAGGTCGTTTCG
GTAGTGGATACGTCTTTAGCTCACATTTCACGAGTCGTGACCAAGCAACAGCGGATTTCTTAAAGCTGTGGGGACTA
TCTGACAACCAACCCTTGAACCAGATTAAATTCCGTGTCGGGCGTAATAAGAGAGCCTGGGTAAATAATTGCGTCTC
AATCGGGCTATCAAGTTGTTTCCTGGAACCGCTAGAATCTACTGGGATTTACTTTATATACGCAGCATTATATCAAC
TGGTCAAACACTTCCCGGATACGTCTTTTGATCCACGTTTATCAGATGCCTTCAACGCCGAGATCGTACATATGTTC
GACGATTGCCGTGACTTTGTCCAAGCGCATTATTTCACAACTTCGAGAGACGACACCCCGTTCTGGTTAGCAAACAG
GCATGATCTAAGACTTAGCGATGCGATCAAAGAAAAAGTCCAGCGTTACAAAGCTGGGCTTCCTCTTACCACAACGA
GTTTCGACGACTCTACCTATTATGAAACGTTTGACTACGAATTTAAAAATTTTTGGTTAAACGGCAACTATTATTGT
ATTTTTGCGGGTCTTGGGATGTTGCCGGATCGTAGCCTTCCCCTTCTGCAACACCGTCCCGAGAGTATAGAAAAAGC

AGAGGCTATGTTTGCGTCCATCCGTAGGGAGGCAGAAAGATTGAGGACCAGTCTACCTACAAACTATGACTACTTGA
GATCCCTACGTGATGGTGACGCGGGATTAAGCAGAGGCCAAAGGGGACCGAAACTTGCCGCACAAGAGTCTTTG

Seq. ID No: 243
>TrpM_1
ATGAGTCCCGTCGCACTTTCTCCTAAAAGGGTTGACATAGTAGATATTAGAGGGAACGACATGCAGTATAGTTTAGT
CAACGAAATTCATAAAGGTCTAAATCCTCCTAACGGTACCAGGAGAAGCCTTCCCACAATGCTATTATACGATAGTG
AAGGGCTTAAACTATTTGAGAAAATTACTTACGTCGATGAGTACTATCTGACTAACGCAGAGATTGAGGTCTTGGAA
AAGCATTCCCGTAGGTTAGTGGAGAAAATTCCGAGTAACGCGCAACTGCTAGAGTTGGGCTCCGGTAATCTTAGGAA
AATTGAGATCCTTCTGAGAGAGTTTGAGCGTGTAGGTAAGCCCGTCGATTACTACGCATTGGATCTGTCACTGAGTG
AATTAGAGCGTACCTTCAGCAATGTCTCATTAGAAGAATATAAGTCTGTAGGATTTCATGGCTTGCACGGTACATAT
GACGATGCACACACTTGGCTTTCTGACCCGAAAAATCGTGAGAGGCCCACTGTCGTTTTGTCAATGGGGTCTAGTCT
AGGGAACTTTTCTCCCCCTGATGCAGCTGCCTTCCTTGCTGGGTTTGCAACCCTTTTGAAACCCTCCGACTTCATGG
TCATTGGTTTAGATGCCTGTGAGGACCCTGATCGTGTATATAAGGCTTACAACGACAGTGCCGGTATAACGAGAAAG
TTTTATGAGAACGGCCTTGCAAATGCAAATAAAACTCTGGGTCACGAGGTATTTAGACCTGACGAGTGGGAAGTGGT
AACGGAATATGATGCGGTCAATGGAAGGCATCAAGTATTCTATGTCCCTACCAAGGACGTAAGCGTGGGTGACGTAT
TGCTAAGGCGTGGCGAGAAGATTATCTTCGCTGAGGCATTTAAGTACGGATGCCAAGCACGTGAAAAGTTGTGGCAC
GACGCTGGACTGATTGAAGCAGCCGAGTTCGGCTCCGGGAGTGAGGACTATAGGACTTATATATAA

Seq. ID No: 244
>TrpM_10
ATGCTAGGACCAGTACCTTCTCCTAGTCCAGTCCCAATACCGCCTGGCTCACGTCCTGGGGCAAGCCCCGGACTGGA
GGCTACCATCCCCATCATAGACATACGTAGCACAGCTCACTCCGTTACAGTAGCAGCTCTTGAGGACGGTATCCGTG
CTAACGTATTAAGCGGGTTTACAAAGCCTTACAATGAAAAAGAGCTACCAAATCTTTTATTGTATAACGAGGAAGGA
TTAAGATTATTTGAACAAATTACCTACCAGCCAGATTACTACTTGACCAGATTGGAAATCGATATTTTATCAAGGCA
CGCCCATCAGATTGCGAACTCAGTACCCGACGGTGCCATCTTATTGGAGTTAGGGGCAGGTGCACTTAGAAAAACGG
CGCTAATTCTGGACGCGTTGGAGGCGCAGGGCAAGGACGTTACATACTTTGCTTTAGACCTAGATAAACCGGAGCTG
CTACGTACATTGGCCGAAGTCAAAGGGAGATATACGCACGTGAGCCTGGCTGGACTATGGGGGACGTACGATGATGG
CTGCACGTGGCTTAAACAAGTGAAAGATAGACCTCGTATTATTCTATGGTTGGGTTCCTCAGTAGGAAACATGAGCC
GTAAAGAAGCTGGCCAATTCATCAGAACCTTCGGCGATATTCTGGCGCCCAGGGACAGGTTCATTGTTGCTATAGAC
AGTAAGAACCATAAGTTGAACGACATTCGTGCCGCTTACGACGACCGTGCGGGCGTTACCCGTAGATTCGCATTGAA
CGCGTTGGGAAACATTAATGACCTGTTCAATGCGGATGTTGTTGACGTCAGTTCTTTTGATTACAACCCGTACTATA
ATGAAGTACAGGGCAGAAACGAAGCGTATTTCAGGTGCCTTAAAGATACGCAAGTCCGTATCCCCTCCGAGACACCT
ATATTGGTACACGAGGGTGAGTATATCAGATTCGCGTTTAGTCATAAATATGATAGAGTTGAAAGGCAAGTTCTGTG
GACAGCCGCAGGGGCATACCCCGTTCAGGAATGGATGTCACAGGACGGTGACTATGCACTGACAATGCTTAGTTGGT
CAAGCTAA

Seq. ID No: 245
>TrpM_11
ATGACCTACAGTATAGTTGATATTCGTAAAACCGATACTTGTTTAAAGAATAGCATAATCAATGGCATTAACCAGTC
CACTAAATCAATCCCTGCTATTGTACTATACGACGAACTTGGCCTGCAATACTATGAAAAAGTGACCTATTTAAAGG
AATATTATTTAACGGAAGCCGAGATTGACATCCTGAAGAACAAGGCTGACCAGATCTCTGACTATATTCCGGAAGGC
AGCTCCCTGATAGAGCTTGGGTCTGGGGCGCTTAGAAAGACAAGATTACTACTAGATTCTATCGAGAAGCAGAAAAA
GAAGGTTATCTACTATGCCCTGGATTTAATGGAGGGCGAATTGAAGCGTACGCTTTCCTCATTGGGCAAGTTTCAAT
ATGTAAAACTTGTTGGCTTGTGGGGTGTTTACGAAGATGGCATTGATTATGCGTCTAATCTGCCAGGGGATTCACAC
AAAACTATTCTATGGATGGGAAGCAGTATCGGCAACTTTAATCGTGACGAGGCTGCTAATTTCGTTAAGACGATCCA
GGACAAAGCTATGAACCCAGGAGATCTATTCCTGATTGGGATAGACAGACGTAAAAACCCGGACAAGATAACAGCCG
CCTACAACGATCCAAAGGGAATCAACGCAAAATTCATAATGAACGGTTTGAACCACGTAAATGCAATATTCGACCAA
CCTATTTTCGATTCTAATAACTTCGAACATGTAACTATGTATAACGATGATGTTGGAAGGCATGAGGCTTACTGCAA
GGTAAAGAATGATACAACTCTTGAGTTTAAAGAATCCAAAGATAACCCGAAAACAATAATCAAACTTAATAAGAATG
AGTTAATTAACATAGGTTACTCCCATAAGTATAATAAGGCGGAAACAGATGCTCTATTTGACTTCTCACTACTAAGC

TACATGGAGTCTTGGACCGACTCACAGAGCCTATATGACCTGCATTTGGTTTATAAGAGTCCCTTCCACTTCACACG
TAAGTTTGATAGCCACAAATAA

Seq. ID No: 246
>TrpM_12
ATGTCTAAAGACGTGCAAGTTCTTGATATACGTGCAAGCCCCCAGTCTAAGGGCTCAATTCCTAATCTGCGTACAGC
GATATTGGATGGCTTGCAGAAGGCGCCCGGTATGAGAACTCTACCTAGTGAAATTCTATATGACGACAGGGGCCTTA
AAATCTATAATGACTGTATCAGATCTTGGTCTGAGTGGTATTATCCCATCTCTGCCGAAACGGAAATCCTTGAGATC
AACGGTAAAGACATAGCCAGGGTGTTCAGTACAAGTGATCGTGGTGAGGCGGTGTTAATTGAGCTGGGTGCAGGGAG
CCTGGATAAGACCAGCAAGATTCTAGTATCCCTTTCTGAAACCGTTCAAAATGTATCCGATAGCCAGCCTCCGATAA
CTTATTACGCACTTGACCTGGAGCGTAGTGAACTGCAACGTACACTTAGTGAACTACAAAAGAACATCGGGGAGAAG
ATAGCGGGGAAAATTGCGACGAAAGGGATGTGGGGTACTTATGACGATGGAATCCGTTCTGTTGAGAATAACGAACT
ACACCTGGACGCAGCAGTGCCGGTTCACTTTTTGTTCCTTGGTGGTACTATCGGCAATTTCTCAAAAGGAGAGGGAG
ATGTGACCTTTCTTAGGAATCTGCCCTTGAATGCACAGAGAGGAGATACGATTCTTTTGGGCATTGACCGTGAGAAA
TCCAAAGAGATTATCGAGAGAGCGTATAATTTTCCCGCCGCCCGTGAGTGGATTATGAACGGTCTAAACGTCTCAGG
GCACTTACTTTCTGGTGACAAAGACCTGTTCCAATTAGATAACTGGGACCGTTATGCGATGTACGACGAAAAACTGG
GCAGATTAGAAGCGGGATATCGTAGCAAGATCGATCAAATCATCGAAGTGACGGCGAATTACAGCATACCTTTTAAG
AAGGATGAATCAGTGATGGCCATTTTCTCCAATAAGTACACCGACGATGAACTTAATTTCCTTATTAGCAAGGCCAA
TCTAAAGACAATCAACAGTTGGGTCGATCACAAGGCCTTGTACTATATTTTCTCACTGAGGAAGGTCTGA

Seq. ID No: 247
>TrpM_13
ATGCCGCGTATACAGGTGCTAGATATAAGAGGCAGTAAGGAGAGCGTGGGCAGTACGCCACACTTACGGGCTGCGAT
ACTTGAAGGGTTACTGAAACCACCCGGTTCCCGTACCCTGCCGAGCGAGACTCTATATGATGAAGTCGGACTGAAGA
TGTACAATGATGGCATGAAGGCCTGGGCTGAATGGTATTATCCCGTAGAAGCCGAACGTCAAATCTTGGAAAGATAT
GGGCGTGATATTGCAAAGCTGTTCACAACTTCTGCCAAGGGGAAAGCGGTACTTATAGAACTGGGAGCCGGGTCACT
TGACAAAACCAGTCAGGTGCTGCTAAGCGCGGCAGAAATCACACGTACCACGGGACCCATGAACAATATAGCATATT
ACGCATTGGATTTGGAGCGAGGGGAGCTGGAGAGAACCATAGGTCGTTTACAGGAAGTCATAGGCGATCAAATAGCA
GGTAAAATCAGTACAGCCGGTATGTGGGGAACCTATGATGACGGGATTAGGGTAATTGAGAAAAACGAACTTGAGCT
AGAACCGGACATCCCTGTGCACATTCTGTTTCTCGGGGGCACTATCGGAAATTTCTCTAAGCAGGACGGAGACGTCG
CCTTTCTTAAATCCCTGCCACTGGACCATAAAAGGGGAGATACGCTTTTAGTAGGGATGGATAGACACAAGAGCGCA
GACGCCATCGAGAGGTCCTATGGATTCGCCGCGGCGAAGGATTGGATCATGAACGGCCTTAAAGTATCAGGTAGAGT
TCTTACCGGGGACGAAGGTTTATTTGAAATAGGGAACTGGGAGAGATACGCAAAATACAATGAAGAACTGGGCAGGT
ATGAAGCAGGATATAAAAGTCAGAAAGAGCACGCCTTAAAAATCAGTGAGGGAGTAGACATCACCTTCTTAAAAGAT
GAGGTAGTGTTGGTTATGTTTTCTAATAAGTACACCGACGCGGAAATGGACAGCGTAGTCGATTCTGCGGGCCTAGT
TAAAAACGGCTCCTGGATGGATGAAAAAGCTCAGTATTGCCTGTTATCTCTGAGAGCCAACAACGGCCCCGTCTAG

Seq. ID No: 248
>TrpM_14
ATGAGCCAGATAGAAGTTCTGGACATTCGTGGAAGCAAGGAAGCTACAGGTTCTACACCACACTTGAGGGCGGAAAT
TTTGCAAGGTCTTTCAAAGTCACCTGGGCATCGTACCATTCCCGGTGAGACCTTATTTGATGAGACCGGGTTGAAGA
TGTATGACGAGGGGATGAAGACTTGGAGAAAGTGGTATTACCCGTTCGAGGCTGAGAAGGAAATCCTTGAAGTTAGG
GGTCTGGAGATAGCCAAATTGTTGAAAACCAGTTCCAAAGGGGAAGCCGTACTAATTGAATTAGGCGCTGGGTCTTT
AGAGAAAACTTCACAAATACTTTTATCTGCCGCTCAGATAGCCGAAACAGCCGATAACAGCACTACTAACCCCATTA
CATACTACGCACTAGACCTGGAACATAGAGAGCTGGAGAGAACACTGGCGGCCTTACAAGACGCTATCGGTCCGAGA
ATCGCTGGTAAAATCACAACCAAAGGGATGTGGGGGACCTACGAAGACGGAATTAGAGTCGTGGAAAGGAATGATCT
AAAATTTCCTTCAGATGTACCTTTACACATCCTATTTCTAGGAGGAACCATTGGCAACTTTTCTAAAGCTGATGGCG
ATATTGCATTTTTGAAGAGCCTGCCTTTGAATCGTAAGAGGGGAGATACATTGCTATTGGGAGTGGACAGAGCTAAG
GCGGTTGAACTAATTGAAAGAGCCTACGGCTTCGCGGCTGCTACAGGTTGGATAATGAACGGCTTAAAGGTAAGTGG
AAGGGTTCTAACCGGAGATGAAGAGCTGTTCGAGAGCGGGAATTGGGAAAGGTACTCCAAGTACAACGAGGAACTAG
GAAGATATGAAGCAGGATATAAGAGCAGGAAGGATCAAACAATAAAGGTTGCTAAGGATGTAGACATCGTTTTTAGT

AAGGATGAAGTCATTCTAGTTACTTACTCAAATAAGTATACTGATGCTGAAATAAAGACGGTCTTTGACGGTGCCGG
ACTGGAAATTGTGGAATCTTGGATGGACAAAAAGGCACAGTATTGCTTGTTTTTACTAAAAGCCTGA

Seq. ID No: 249
>TrpM_2
ATGACGTTGAGTCTTGCTAATTACCTAGCGGCGGATAGCGCCGCCGAGGCTTTAAGGCGTGACGTGAGGACCGGTCT
GACGGCGACACCAAAGAGTCTTCCTCCCAAATGGTTCTACGACGCTGTAGGTTCCGACCTTTTTGACCAAATAACTA
GACTACCCGAATATTACCCTACCCGTACTGAAGCACAAATTTTACGTACCCGTTCAGCGGAAATAATCGCGGCTGCC
GGGGCGGATACGTTGGTCGAGTTGGGCTCAGGGACTTCTGAGAAGACTCGTATGCTGTTGGACGCAATGAGAGATGC
GGATCTATTGAGGAGGTTCATTCCCTTCGACGTCGATGCTGGGGTATTAAGATCTGCGGGAGCGGCCATTGGGGCGG
AGTATCCAGGTATAGAAATTGATGCAGTCTGTGGGGATTTCGAAGAACATTTGGGGAAGATCCCAAGAGTAGGTAGA
AGGTTAGTTGTGTTTTTAGGTAGTACGATCGGTAACCTGACACCACAGCCGAGGGCTGAATTTTTAGCCACACTAGC
AGATACTTTACAACCAGGAGACTCCCTGTTGTTAGGAACGGATCTTGTAAAGGATACTGGGAGACTTGTAAGAGCAT
ACGATGACGCGGCTGGCGTAACAGCTGCGTTTAATAGGAACGTACTTGCAGTCGTTAATAGGGAACTGTCCGCCGAC
TTTGATCTTGATGCGTTCGAACATATAGCTAAGTGGAACGATGACGAGGAGAGAATCGAAGTTTGGCTACGTGCCAG
AACAGCCCAGCATGTCAGAATACCTGCACTAGACCTGGAAATAGATTTCGCAGCCGGGGAGCAGATGCTGACCGCAG
TAAGCTGTAAGTTTCGTCCCGATAGTGTTGCCGCTGAACTAGCCGAAGCAGGTTTGAGACAAACACATTGGTGGACC
GATCCAGCGGGCGACTTCGGACTGAGTCTAGCTGTGAGATAG

Seq. ID No: 250
>TrpM_3
ATGACACTTAGCCTTGCAAATTACCTTGCGGCCGATTCAGCGGCCGAAGCACTAAGGAGGGACGTCAGAGCCGGATT
AACTGCTGCCCCGAAATCTCTTCCACCTAAGTGGTTCTATGACGCCGTAGGAAGCGATCTTTTCGACCAGATTACCA
GATTGCCCGAGTACTATCCAACAAGAACAGAAGCTCAGATACTAAGGACTAGGAGCGCGGAAATAATAGCGGCGGCT
GGAGCAGACACACTTGTAGAACTAGGTAGCGGTACTTCCGAGAAGACACGTATGCTACTAGATGCAATGAGAGACGC
TGAGTTATTGCGTAGGTTCATTCCGTTCGACGTGGACGCGGGAGTGTTAAGGAGTGCTGGGGCGGCAATTGGTGCGG
AGTATCCTGGCATCGAGATAGACGCTGTTTGCGGAGATTTTGAGGAACACCTGGGTAAGATTCCCCATGTCGGTCGT
CGTTTGGTGGTCTTCTTAGGGAGCACCATAGGCAATCTGACGCCTGCGCCCAGAGCTGAATTTCTAAGTACGTTAGC
CGATACCCTTCAACCAGGGGATAGCTTGTTATTGGGTACTGACCTAGTCAAGGATACCGGACGTTTGGTTAGAGCTT
ACGATGATGCGGCAGGAGTGACGGCCGCTTTTAATAGAAATGTACTTGCTGTGGTTAACAGGGAACTATCCGCCGAC
TTTGATCTGGACGCGTTTGAGCATGTCGCCAAATGGAACAGCGATGAAGAGAGGATTGAAGTCTGGTTGAGGGCGAG
GACTGCTCAACACGTGAGGGTTGCAGCTCTGGATCTGGAAGTGGATTTTGCGGCTGGAGAGGAGATGCTGACCGCCG
TTTCCTGTAAATTTAGACCTGAGAATGTCGTAGCCGAGCTTGCGGAGGCAGGCTTGAGACAGACTCATTGGTGGACA
GATCCAGCGGGAGACTTTGGGTTGAGCTTGGCCGTTCGTTAG

Seq. ID No: 251
>TrpM_4
ATGAGGGTCAGTGGGGCAAACCATCTGGGCGAGGATGCCGGGCACCTGGCATTAAGGCGTGATGTATATAGTGGGCT
TCAGAAAACACCTAAAAGTCTACCACCCAAGTGGTTTTACGATACCGTTGGCTCAGAACTATTTGATCAGATCACGA
GATTACCGGAGTATTATCCCACGAGAGCCGAGGCCGAAATATTGCGTGCGAGGAGCGCAGAGGTAGCTTCAGCTTGC
AGGGCAGATACTCTAGTCGAGCTTGGGTCAGGCACGTCAGAAAAGACCAGGATGCTGCTAGATGCGCTAAGGCATCG
TGGGTCCCTGAGGAGATTCGTACCTTTCGATGTAGATGCTTCCGTGCTTTCCGCCACTGCGACTGCGATTCAAAGAG
AATATTCCGGCGTGGAGATCAACGCCGTCTGTGGGGATTTTGAAGAACATCTAACCGAGATACCCAGGGGCGGGAGG
CGTCTATTTGTGTTTTTGGGCTCCACGATTGGTAACCTGACACCGGGTCCCCGTGCCCAGTTTCTAACTGCCCTGGC
AGGCGTCATGCGTCCAGGGGACTCCTTGTTATTAGGCACAGATTTAGTGAAAGATGCTGCAAGGTTAGTGAGAGCTT
ATGATGATCCAGGAGGAGTGACGGCGCAGTTTAACAGAAATGTCCTGGCCGTCATAAACCGTGAGCTAGAAGCAGAT
TTTGACGTAGACGCATTTCAACACGTAGCCCGTTGGAACAGCGCCGAGGAAAGGATTGAGGTCTGGCTTAGGGCGGA
CGGAAGACAACGTGTGAGGGTGGGTGCTTTAGATTTGACAGTAGATTTTGATGCGGGTGAAGAGATGTTAACTGCTG
TGTCATGTAAGTTTCGTCCACAGGCAGTTGGTGCAGAGCTTGCAGCAGCCGGCTTACATAGGATTAGGTGGTGGACC
GATGAAGCCGGTGACTTCGGTCTAAGCCTCGCCGCAAAATGA

Seq. ID No: 252
>TrpM_5
ATGACGCTGACATTATCAAATTATCTTGCTGCAGATTCTGCCGCCACTGCTCTAAGGCGTGACGTGCATGAGGGGTT
GACCCAGTCTCCGAAGATGCTTCCGCCGAAGTGGTTTTATGATAGCGTTGGGTCAGATCTTTTCGACCAAATTACCA
GACTACCTGAGTACTACCCTACAAGAACGGAAGCGCAGATACTTACCCATCGTTCTCCAGAAATTGTCGCGGCAGCG
GGGGCCGATACACTTGTGGAATTAGGATCAGGAACATCTGAGAAGACCAGAATGCTATTAGATGCGATGAGAGACGG
GGGTCAGCTAAGACGTTTTATTCCGTTCGACGTTGACGCGGGCGTGCTAAGAGCAGCAGGGGCGGCTATAGGACAAG
AATATCCAGGGATTGAAATTGACGCGGTGTGTGGGGATTTTGAGGAACATTTAGGGAAAATACCTGCGGTCGGCAGG
AGACTGGTGGCTTTCTTAGGGTCTACGATCGGTAACCTAACCCCAGGTCCACGTGCGGACTTTCTGGCAAGTCTTGC
CGAGACTCTTCAGCCCGGCGATTCCGTGTTGCTTGGGACGGACCTGGTTAAGGACACTGGTCGTCTAGTCTCTGCAT
ATGACGATTCCGCAGGAGTTACAGCAGCGTTCAATAGAAATGTCCTTTCTGTAGTTAACAGAGAATTGGACGCTGAT
TTTGACCTAGATGCTTTCGCACATGTCGCTAAATGGAACGCGGAAGAAGAGCGTATCGAGGTATGGCTTAGGGCCGA
CGCACCCCAGCAAGTAAGGATAGCTGGGCTAGACTTAGACGTTGCTTTCGGGGCTGGCGAGGAGATGTTGACGGCAG
TTAGTTGTAAGTTCCGTGCTGATGGTGTCGCGGATGAATTAGCGAAGGCCGGGCTGAGGCAAACACACTGGTGGACC
GACGAGGCGGGTGACTTTGGTCTAAGTCTTGCGGTCAAGTAG

Seq. ID No: 253
>TrpM_6
ATGCTTGAAGCGACGTCTACCCAAAATCTGGTGAGCTTCCAAATACCCATCGTTGATATTAGGACACCGTCTTGTTT
AGAGGAAACTATTAGGAAGAAAGTAGTATCCGGTTTAGCCCGTCCTTATAATAAGAAATCTATTCCGGACTTGCTAC
TGTATAATGAGACAGGCTTACGTCTTTTTGAGGACTTGACATACCAGCCGGATTACTATCTAACCGGACTGGAAATC
GAAATATTGTCAAAGCATTCTCTTCAGATAGCCGACTCAATCCCGGTAGGGTCCCTGATAATGGAACTAGGTGCTGG
CGCCCTTCGTAAAACGGCCCTAATACTAGACGCCTTAGAGGCACAGAAAAAGGAAGTTGCTTATTTAGCACTGGACC
TGGACCGTCCTGAGTTGGTGAGAACTCTTGGACAGTTGAACGGGAAATACACACATGTGAAGCTGGGCGGGTTATGG
GGTACGTATGATGATGGACGTCGTTGGCTTAGTGAAAACACCAGTGATTCCCCCAGAACAATACTGTGGCTTGGAAG
CTCTATTGGTAACGTCAAAAGAGATGACGCGGGGGACTTTATTCGTTCCTTTGGTGACGTCCTTAGTTCTAAAGATC
GTTTTGTGGTGGCAATAGATTCTAGGTATCACGAGGTTGATACCATTTGTAGGGCGTATAACGATAGAGAGGGTTTC
GCTGAAAGATTTTGCTTGAACGGAATAGACTCATTTAACCAACTATTCGGCCGTGCTATCATTGACATCTCTTGCGC
GAAATACCGTACAGTTTATAACGAGGTCAAAGGGAGACACGAGGTGTACTACAGATGCACCCACGATTTCGAAATCA
GGCTACCTGGAGATTATCCACCCACCTTCTTGTACGAGGGAGAACTTATCCTGTTGGCGCATTCATATAAATATGCG
GCGGTTGAGAGAGAGACGCTATGGCTGCGTGCAGGAGCTAGACCGGAGAAAGAATGGATGACGGACGGGAGCTATAC
CGTGACAATGTTGTCATGGCCATAA

Seq. ID No: 254
>TrpM_7
ATGTCTCCATCTACGGTTAACAAAATCGCTTCTAGCCCGGTTTTCGACATTCGTTCCGACGAAACGAAAGGCTTTGC
TAAGGCTCCGATAGAAGATGAGTTAGCAGGGCTACAAGCAGTATACAACGAGAAAACGCTGCCAAATGTACTTCTAT
ATGATGCCAAGGGGCTGCAACTTTTCGAAAAGATAACATACACCAATGACTACTACTTAACTGGACTGGAGATGGAC
CTTCTTGGAGAACACGCCGATGAGATGGCAGAATGGATAAAGGATGGAGCGGCTCTGGTCGAACTTGGTGCGGGCGC
ACTTAGGAAAACCGCTATCTTACTGGATGCCATTGAACGTCAAGGTAAGAGGATAACCTTTTATGCGTTGGACCTAG
ACCACTCAGAGTTGACCAGGACATTGGCGGAGCTAGAGGGTAGGTATAGGCACATAACACTATGTGGCTTGTGGGGA
ACGTATGATGATGGCAGGGCCTGGCTTGCATCTACAAATGAAGAGCAGCGTGTCCTATTGTGGTTGGGTAGCAGTAT
TGGCAATCTGTCAAGACAAGAGGCCAAGGATTTTCTGCATTCTTTCGGACGTGCCTTGAGACCTGGAATCGACAAGT
TTATAGTAGCAATGGACAGTAAGTATAACGCTGTCTCTTCCATGACCAGGGCATATAACGACTCAGAAGGAGTAACA
GCAAGTTTCGCGTTAAACTTACTTGACGCATTCAATGCTAAAGTCGGATTCAAAGCACTGCCCCCATCATCCTTTTG
CTATAGTCCCTTCTTTAATCAGGCACAGGGGAGAAACGAGGCATATCTAAGGGCACGTCATGGTGTGAGGTTCGAGG
TTAACGGCATAGCAGTCGAAGTGAGAGATGAGGAGTTGATTCGTTTCGCGTACTCCCATAAATACGATAACGCTGAG
CGTGACTTGCTTTGGAGAGCTGCTGAAGCAAATGTAGAACAAGAGTGGCTACATTCACCCCAGTCAGGTAGGGCACG
TTATTCCATAAGTCTGCTGTCCTTCCGTGATTAG

Seq. ID No: 255

>TrpM_8
ATGACCCTGTCTCTATCAAACCATCTTCCAGCCAACAGTGCAGCAAGGGTACTGAGGAGAGATGTCCTGGATGGTTT
AACGCAAACGCCCAAAGCCTTGCCGCCTAAGTGGTTTTACGACTCAGTCGGAAGTGATCTTTTCGACCAGATCACAA
GACTTCCAGAGTATTACCCGACTAGAACAGAAGCACAGATCTTGAGGACCCGTAGTGCGGAGATTGCCGAGGCTTCT
GGAGCCGATACATTGGTAGAGTTAGGGAGTGGGACAAGCGAAAAGACGAGGATGCTTCTGGATGCCTTGAGAGATAA
CGGTACTCTTAGAAGATTTATACCTTTTGATGTGGACGCTGGGGTACTAAATGCAGCCGGAGCGGCCATCCAGAAGG
AATATCCAGGCGTAGAGGTTGATGCAGTATGCGGCGACTTCGAGGAGCACCTTGGCGAAATCCCTCGTGTCGGGAGG
CGTCTAATCGCGTTCTTGGGGTCCACTATTGGAAACCTTACACCCCAACCGAGAGCCCGTTTTCTGACGGCCTTAGC
GCAAACCATGCGTCCAGGGGACAGCCTGCTTCTTGGAACTGACCTTGTCAAAGACACAGAGAGATTAGTAAGGGCAT
ACGACGATTCAGCGGGAGTTACCGCGCGTTTTAACAGGAATGTACTGGCGGTAATCAACAGGGAACTGGATGCGGAT
TTTGATCTAGCCGCTTTCGATCATGTCGCTAGGTTCAATGCAGCCGAGGAACGTATTGAGGTGTGGCTGCGTGCACG
TGGGGCGCAAAGGGTGTATGTCAGAGAACTTGACTTGACGGTAGACTTTGCCGATGGAGAAGAAATGTTGACCGCGG
TGAGTTGTAAATTCAGGCCCGATGGAGTTGCAGCTGAGTTAGCCGCAGCGGGTCTGCGTAGAACTCACTGGTGGACC
GATCCTGCTGGAGATTTCGGCCTTAGTTTAAGCACGAAATAA

Seq. ID No: 256
>TrpM_9
ATGACCATTTCCATAGCCAACTATCTGGCTGCCGACTCAGCAGCTACAGCCCTGCGTAGAGACGTGAGAGAGGGATT
GGCAGGGACCCCTAAGTCTCTACCGCCCAAATGGTTTTATGACTCAGTTGGCTCCGACCTTTTTGACCAGATCACCA
GATTGCCGGAATACTATCCTACTAGGGCAGAGGCCCAGATACTAAGAACACACGCGGTTGACGTAGCTGCGGCCTCT
GGCGCTGACACGTTAGTTGAGCTTGGCAGCGGGACCAGTGAGAAGACGCGTTTATTACTTGATGCTTTGCATCGTGC
TGACAGTTTGAGGCGTTTCATACCGTTCGATGTAGATGCGAGTATCCTGCAGTCCGCCGGCGCTGCTATCAGTCAGG
AGTACCCAGATGTCGAAATTGAAGCTGTCTGTGGCGACTTTGAGGAACACCTGGGAAAAATACCTCTTCAGGGTAGA
AGGCTGGTTGTTTTCCTGGGATCTACTATTGGCAACCTAACTTCCGGACCCAGAGCAACTTTCCTATCCGCATTGGC
AGATTCCTTGCAACCGGGTGACACTCTATTGCTGGGTACCGATCTAGTGAAGGACGTAGATAGGTTGAAGAGAGCCT
ACGATGACGCAGCGGGAGTGACGGCAAGGTTCAATAAGAACGTCCTTACTGTAGTCAATAGGGAGTTGGGCGCGGAC
TTCGACCTGGACGCATTTGAGCACGTCTGCAAGTGGAACGCGGATGAAGAACGTATCGAGGTTTGGTTGAGGGCTAA
TACTCTGCAACGTGTCCATATATCAGGACTGGAACTAGATGTAGAATATGCTGCGGGGGAAGAGATGTTGACAGCAG
TCTCTTGTAAGTTCCGTCCTGAGGGGATCGCGGCCGAATTAGCCGCAGTCGGATTAAATCGTACCCACTGGTGGACG
GATGACGCGGGAGATTTCGGCCTTAGCTTAGCAGTAAAGTAA

Seq. ID No: 257
>TrpS_1
ATGACGACGTTATTAAATCCCTACTTTGGTGAGTTTGGAGGTATGTACGTACCTCAGATTTTGATGCCTGCTCTAAA
TCAACTTGAGGAGGCATTTGTGTCCGCACAAAAAGATCCGGAGTTCCAGGCTCAATTTGCCGATCTGCTAAAGAATT
ATGCAGGGAGACCGACGGCGCTTACTAAGTGCCAGAACATCACAGCCGGGACCAGGACCACACTATACCTTAAAAGA
GAAGACCTGCTGCATGGTGGTGCCCATAAAACGAACCAGGTTTTAGGCCAGGCTCTTTTAGCCAAACGTATGGGAAA
AAGCGAAATCATAGCAGAAACGGGAGCTGGGCAGCATGGTGTAGCCTCTGCCTTAGCCTCCGCTCTATTAGGCCTGA
AGTGCCGTATCTACATGGGGGCAAAAGATGTGGAGAGGCAGAGTCCCAATGTTTTTCGTATGCGTCTGATGGGAGCG
GAAGTCATCCCAGTACACTCTGGGTCAGCTACGGTAAAAGACGCGTGCAACGAGGCGTTGCGTGACTGGAGTGGGTC
ATACGAGACAGCCCATTATATGCTTGGGACAGCAGCTGGCCCGCACCCGTACCCTACGATAGTAAGGGAATTTCAAA
GGATGATTGGTGAAGAGACAAAAGCTCAGATATTAGATAAAGAGGGAAGGCTGCCTGATGCGGTGATCGCATGCGTT
GGAGGGGGTAGTAATGCCATTGGCATGTTTGCTGACTTTATCAATGACACAAGCGTTGGCCTAATTGGTGTCGAACC
GGGCGGGCATGGCATCGAGACTGGAGAACATGGTGCCCCACTGAAACACGGCAGAGTTGGGATATACTTTGGGATGA
AGGCACCTATGATGCAGACGGCCGACGGTCAAATAGAAGAGTCATACAGCATAAGTGCGGGGCTAGATTTTCCAAGT
GTTGGTCCACAACATGCCTATCTGAACTCCATCGGGAGAGCGGATTATGTTAGCATTACTGACGACGAAGCACTAGA
GGCCTTTAAGACCCTGTGCAGGCATGAGGGCATAATTCCGGCATTGGAAAGTAGTCACGCTCTAGCGCACGCGTTAA
AGATGATGAGAGAACAACCCGAGAAAGAACAATTATTAGTAGTGAATCTTTCAGGCAGGGGCGATAAGGACATATTC
ACAGTACATGACATCTTGAAGGCTAGAGGGGAAATA

Seq. ID No: 258

>TrpS_2
ATGTGGTTTGGAGAATTTGGAGGCCAATATGTCCTAGAAACACTAATAGGACCACTAAAGGAGTTAGAGAAAGCATA
TAAGCGTTTTAAAGACGACGAGGAGTTCAATCGTCAGCTTAACTATTACCTGAAGACCTGGGCAGGAAGGCCAACGC
CATTGTATTATGCTAAGAGACTTACCGAGAAGATTGGCGGCGCTAAGGTCTACTTAAAGAGGGAGGACCTGGTCCAC
GGGGGCGCACACAAAACAAATAACGCCATAGGTCAGGCCCTACTAGCTAAGTTTATGGGAAAAACTAGGTTAATTGC
CGAAACAGGAGCAGGACAGCACGGCGTTGCCACAGCGATGGCTGGGGCTTTGCTTGGAATGAAAGTAGATATATACA
TGGGAGCGGAAGACGTTGAACGTCAGAAAATGAATGTTTTTAGGATGAAGCTATTGGGCGCTAATGTCATACCAGTT
AACAGCGGTTCCAGGACACTGAAGGACGCAATCAATGAAGCATTACGTGATTGGGTGGCGACGTTTGAGTACACACA
TTATCTGATAGGCAGCGTCGTTGGCCCTCATCCCTATCCGACAATAGTCAGGGATTTTCAATCTGTAATAGGTAGGG
AGGCCAAAGCTCAAATACTAGAAGCTGAAGGGCAATTGCCTGATGTCATCGTTGCTTGTGTCGGGGGAGGCAGTAAC
GCTATGGGGATATTTTATCCATTTGTTAACGACAAAAAGGTTAAATTGGTAGGTGTCGAGGCAGGGGGAAAAGGTCT
GGAAAGCGGTAAGCATAGTGCCAGTCTAAATGCTGGGCAGGTAGGTGTTTCACACGGCATGTTGTCTTATTTTCTCC
AAGATGAGGAAGGCCAAATCAAGCCGTCTCATAGCATAGCACCAGGGTTAGACTACCCGGGGGTCGGGCCCGAACAC
GCGTATCTAAAGAAAATCCAGAGGGCGGAGTACGTTGCGGTAACAGACGAAGAAGCGTTAAAGGCCTTTCACGAGCT
ATCTAGGACAGAAGGTATTATTCCAGCCCTAGAATCCGCGCACGCGGTGGCATACGCTATGAAGTTAGCCAAGGAAA
TGAGTCGTGACGAAATTATTATAGTGAACCTATCTGGGAGGGGCGACAAGGATCTGGACATAGTATTAAAAGTGTCA
GGGAACGTA

Seq. ID No: 259
>affibody_tag_1
ATGGTGGACAACAAGTTCAACAAGGAGACGATACAGGCATCACAAGAAATAAGGCTGTTACCTAATCTTAACGGTCG
TCAAAAACTGGCGTTCATTCACTCACTGTTAGACGACCCCAGCCAGTCCGCGAACCTGCTGGCGGAAGCTAAAAAGT
TAAATGATGCACAAGCTCCGAAGAATGCCGCAATCCGTTCTTCCTCTGCTTCATCTGGAGGGAGCGGGGGCAGCTCA
AGTAGC

Seq. ID No: 260
>affibody_tag_2
AACGCAGCAATTAGAAGCAGTAGTGCTAGTAGTGGTGGATCAGGAGGTTCCTCATCCTCCGTGGATAATAAGTTTAA
CAAAGAAACAATACAAGCATCACAGGAGATTAGACTATTGCCAAACCTTAATGGGAGACAGAAACTGGCGTTCATTC
ACAGTCTATTAGACGATCCCAGCCAATCCGCCAATCTACTTGCAGAGGCCAAGAAACTGAACGATGCACAGGCACCG
AAA

Seq. ID No: 261
>affibody_tag_3
AACGCCGCAATTAGGTCATCCTCAGCATCTAGCGGCGGGTCAGGCGGGTCATCTTCTTCAGGCGTGGATAACAAGTT
TAACAAGGAATTGGGCTGGGCGACATGGGAAATTTTTAACCTACCAAATCTGAATGGAGTCCAGGTAAAGGCGTTTA
TCGATTCCCTGAGGGACGATCCCTCCCAGTCCGCGAATCTACTAGCGGAGGCAAAGAAATTGAATGACGCTCAAGCG
CCGAAG

Seq. ID No: 262
>affibody_tag_4
ATGGTGGACAACAAATTCAACAAGGAGCTGGGTTGGGCCACCTGGGAAATCTTCAATCTTCCAAACCTTAACGGCGT
ACAAGTAAAGGCTTTCATCGACTCTCTAAGAGACGATCCGTCACAGAGTGCTAATCTGTTGGCCGAAGCGAAGAAGC
TGAATGACGCTCAGGCTCCCAAAGGCAACGCGGCCATAAGATCAAGCTCAGCTTCCTCTGGAGGTAGCGGCGGTTCC
AGTTCTTCA

Seq. ID No: 263
>affibody_tag_5
ATGGTAGACAACAAGTTCAATAAAGAGATGAGAAATGCTTACTGGGAGATCGCTCTGCTACCAAACTTAAATAATCA
ACAAAAGAGAGCATTTATTAGAAGTTTATATGACGACCCATCCCAGTCCGCGAACCTTCTTGCGGAGGCAAAGAAAT

TGAACGATGCTCAGGCTCCGAAGTCAAGTAACGCAGCAATACGTTCTTCCAGTGCAAGCTCTGGGGGAAGTGGGGGA
AGTTCATCCTCC

Seq. ID No: 264
>affibody_tag_6
AATGCCGCCATCAGGTCTTCATCAGCCTCAAGTGGCGGCTCTGGAGGGTCCAGCTCCTCTGGGGTCGATAATAAGTT
TAATAAAGAGATGAGGAACGCATATTGGGAAATTGCCTTACTGCCGAATCTGAATAATCAGCAGAAAAGGGCTTTCA
TCAGATCCCTGTACGATGACCCGTCCCAGAGCGCGAATCTTTTGGCAGAAGCGAAAAAGCTGAACGACGCACAGGCA
CCAAAG

Seq. ID No: 265
>cofold_1
ATGAAAATCGAGGAGGGGAAGCTAGTTATCTGGATCAACGGGGATAAGGGCTATAATGGCCTAGCCGAAGTAGGAAA
GAAGTTTGAGAAAGACACCGGTATAAAAGTGACAGTGGAGCACCCTGATAAGTTGGAGGAGAAGTTCCCTCAGGTGG
CTGCTACTGGGGATGGTCCAGATATTATCTTCTGGGCTCATGACCGTTTCGGCGGATACGCCCAAAGTGGATTGCTT
GCTGAAATAACGCCAGATAAAGCATTCCAGGACAAATTGTACCCATTCACCTGGGACGCGGTGCGTTACAATGGAAA
ACTGATCGCTTACCCTATCGCGGTAGAAGCTCTTTCACTTATATATAACAAAGATCTGTTGCCTAACCCTCCAAAGA
CCTGGGAAGAGATTCCAGCATTAGACAAAGAATTGAAAGCAAAAGGCAAGAGTGCTCTGATGTTTAATTTACAAGAG
CCGTATTTTACCTGGCCCTTGATCGCCGCGGACGGGGGTTACGCTTTTAAGTATGAAAACGGTAAGTACGACATCAA
GGATGTTGGAGTAGATAACGCCGGAGCAAAAGCGGGTCTTACCTTTTTAGTAGATCTGATCAAGAATAAGCATATGA
ACGCTGATACGGATTATAGCATCGCAGAAGCGGCATTTAATAAGGGTGAGACAGCGATGACAATAAACGGCCCCTGG
GCATGGTCTAATATCGATACTTCTAAGGTGAATTACGGAGTGACTGTCTTACCTACTTTCAAAGGGCAACCTAGCAA
GCCGTTTGTTGGAGTCCTGTCTGCGGGGATTAACGCAGCATCCCCTAATAAGGAGTTGGCTAAAGAATTTTTGGAAA
ATTATCTGCTTACCGATGAGGGTTTGGAGGCGGTTAATAAAGATAAACCCCTGGGGGCGGTGGCTCTTAAATCATAC
GAAGAAGAGCTGGCGAAAGATCCAAGAATAGCTGCTACAATGGAGAATGCACAGAAGGGTGAGATAATGCCGAATAT
TCCACAAATGTCTGCCTTTTGGTACGCTGTCAGGACTGCGGTCATTAATGCTGCCTCAGGAAGGCAAACGGTGGACG
AGGCGTTAAAAGACGCCCAGACAAACTCAAGTAGTAATAATAACAACAATAACAACAACAATAATTTAGGTATTGAA
GGTAGG

Seq. ID No: 266
>cofold_2
ATGGTGTCTAAGGGTGAGGAGTTGTTCACCGGAGTTGTCCCAATTTTAGTGGAATTAGACGGAGACGTCAACGGACA
CAAATTCAGCGTCTCAGGAGAGGGGGAGGGAGACGCGACTTACGGAAAACTTACCCTAAAACTGATATGTACGACGG
GAAAACTTCCAGTACCCTGGCCTACTCTAGTTACAACTTTAGGGTATGGCCTACAGTGTTTCGCGCGTTATCCAGAC
CACATGAAACAACACGACTTTTTCAAGTCAGCTATGCCGGAGGGCTACGTTCAAGAACGTACCATTTTCTTCAAAGA
TGATGGAAATTACAAAACAAGAGCGGAGGTCAAATTTGAGGGAGATACCCTAGTAAATCGTATAGAACTGAAAGGGA
TAGACTTTAAAGAAGATGGAAACATCCTAGGCCATAAACTTGAGTATAATTACAATTCCCACAATGTGTACATTACG
GCTGATAAACAGAAAAACGGGATCAAGGCCAATTTCAAGATTCGTCATAACATCGAGGATGGAGGGGTGCAATTAGC
AGACCACTACCAACAGAATACTCCTATTGGTGACGGACCTGTCCTTTTACCGGACAATCACTATTTGAGTTACCAGT
CTGCTCTGTCCAAAGACCCGAATGAGAAACGTGACCACATGGTGTTATTGGAGTTCGTGACAGCAGCTGGAATCACG
CTTGGTATGGATGAGTTGTATAAA

Seq. ID No: 267
>cofold_3
ATGGCGATGTTCTGCACGTTCTTTGAAAAACATCACCGTAAATGGGATATACTACTAGAGAAAAGTACCGGGGTCAT
GGAAGCGATGAAAGTGACCAGCGAGGAGAAGGAGCAGCTTTCTACGGCGATTGATAGGATGAATGAAGGCCTGGACG
CTTTCATTCAATTATATAATGAATCTGAAATAGATGAGCCCTTAATCCAACTGGATGACGATACCGCCGAGTTAATG
AAACAAGCTAGGGATATGTATGGCCAGGAAAAGCTAAATGAGAAACTAAACACCATAATCAAACAGATTTTATCTAT
ATCCGTTTCAGAAGAAGGAGAGAAGGAAGGTTCAGGCTCTGGA

Seq. ID No: 268

>cofold_4
ATGTACCTGTTAGGAATCGGGCTGATTCTTGCACTAATTGCCTGTAAACAAAACGTGTCCAGTTTAGATGAGAAAAA
TTCAGTGAGTGTTGACTTACCTGGAGAAATGAAGGTCCTAGTATCTAAGGAGAAAAATAAGGATGGAAAGTACGATT
TGATAGCAACGGTTGATAAATTGGAACTAAAAGGTACTTCTGACAAAAACAATGGTAGTGGGGTTTTAGAAGGTGTG
AAAGCGGACAAAAGCAAGGTTAAACTGACGATCAGTGATGACGGATCAGGT

Seq. ID No: 269
>cofold_5
ATGGCAGACAGGGATAGGTCTGGGATATACGGAGGGGCTCACGCAACCTATGGCCAGCAGCAGCAGCAAGGTGGAGG
TGGTAGGCCGATGGGGGAACAAGTCAAAGGAATGCTTCACGACAAAGGGCCGACGGCTAGTCAGGCTCTGACGGTCG
CGACACTTTTTCCCCTGGGCGGTTTATTGCTAGTGCTTAGTGGACTGGCCCTAACGGCGTCAGTGGTGGGATTGGCT
GTAGCTACCCCAGTTTTCTTGATATTTAGTCCTGTTCTAGTACCAGCCGCCCTTTTGATTGGAACAGCGGTGATGGG
CTTCCTGACGAGCGGCGCTTTAGGCTTAGGAGGCTTATCTAGTCTTACTTGTCTTGCAAACACGGCACGTCAGGCTT
TTCAAAGAACGCCTGATTACGTAGAGGAGGCTCATAGAAGAATGGCGGAGGCTGCCGCGCATGCCGGACATAAAACG
GCTCAAGCGGGCCAGGCTATTCAGGGCCGTGCTCAGGAAGCTGGGGCCGGAGGTGGAGCGGGA

Seq. ID No: 270
>oxidase_1
ATGAAGATCTTAATCTTGGGCATCTTTCTGTTTCTATGTAGCACTCCAGCTTGGGCCAAGGAGAAGCACTATTACAT
CGGGATTATAGAGACGACGTGGGATTATGCCAGCGACCACGGCGAGAAGAAACTTATTAGCGTGGACACTGAGCACT
CTAACATCTATCTTCAGAATGGCCCAGATAGGATAGGACGTTTATATAAAAAGGCATTATACCTTCAATATACTGAC
GAAACTTTCAGAACGACGATTGAAAAGCCTGTGTGGCTAGGGTTCCTGGGGCCTATCATTAAGGCGGAGACGGGAGA
TAAGGTCTACGTACACCTAAAGAACCTTGCTAGTCGTCCTTACACATTTCACTCTCATGGTATTACATACTACAAGG
AGCATGAGGGTGCTATTTATCCAGACAACACCACAGATTTCCAAAGAGCCGATGACAAAGTTTACCCAGGGGAACAG
TACACCTATATGCTACTGGCAACGGAGGAACAAAGTCCCGGGGAGGGCGACGGGAACTGCGTAACAAGAATCTACCA
TAGCCATATCGACGCACCAAAGGATATAGCAAGCGGCCTAATCGGGCCTCTGATTATTTGCAAGAAGGACTCCCTGG
ATAAAGAGAAAGAGAAGCATATTGATCGTGAGTTTGTGGTGATGTTCTCAGTAGTTGACGAAAATTTCAGCTGGTAT
CTGGAGGACAACATTAAGACATATTGTAGCGAACCGGAGAAGGTAGATAAAGATAATGAAGATTTTCAAGAGTCTAA
TAGGATGTATAGTGTAAACGGTTATACCTTTGGCTCACTTCCAGGGCTGAGTATGTGCGCAGAGGATAGAGTTAAAT
GGTACCTGTTTGGCATGGGGAACGAAGTAGACGTACACGCAGCATTCTTTCATGGACAGGCGCTAACTAACAAAAAC
TATAGAATCGATACAATAAATCTATTCCCAGCTACGCTGTTTGACGCCTACATGGTTGCTCAAAACCCCGGAGAATG
GATGCTATCTTGTCAGAATTTGAATCACCTTAAAGCCGGTTTGCAAGCATTCTTTCAGGTACAAGAGTGTAATAAGT
CCTCATCTAAAGATAACATTAGAGGAAAGCACGTCCGTCATTACTACATCGCGGCGGAGGAAATAATTTGGAATTAT
GCACCCAGCGGCATAGACATCTTTACTAAAGAGAACTTGACTGCCCCTGGGTCTGACAGTGCCGTGTTCTTCGAGCA
AGGTACCACCCGTATTGGTGGATCTTATAAGAAACTTGTCTATAGGGAATATACCGACGCTAGTTTCACTAACCGTA
AAGAAAGAGGCCCTGAGGAAGAGCACCTAGGAATCTTAGGGCCGGTCATTTGGGCGGAAGTGGGGGACACAATACGT
GTTACCTTCCATAATAAAGGTGCTTATCCATTGTCCATCGAGCCAATCGGCGTAAGATTTAACAAGAATAACGAGGG
GACTTATTACTCTCCTAACTATAACCCTCAGTCAAGATCTGTGCCCCCAAGCGCCTCACACGTTGCTCCAACTGAAA
CTTTTACTTACGAATGGACGGTACCCAAAGAGGTCGGGCCCACTAATGCTGACCCTGTATGCCTAGCTAAGATGTAC
TATTCCGCTGTAGATCCTACCAAAGATATCTTTACGGGGTTAATAGGACCCATGAAAATCTGTAAGAAAGGTAGCCT
ACACGCCAATGGCAGACAAAAGGACGTGGATAAGGAGTTTTATCTTTTCCCTACAGTTTTTGACGAAAATGAAAGCC
TTCTATTGGAGGATAATATCCGTATGTTCACTACCGCGCCCGACCAAGTTGATAAAGAGGACGAAGACTTTCAAGAG
AGTAATAAGATGCACAGTATGAACGGATTCATGTACGGCAATCAGCCGGGCCTTACCATGTGCAAGGGTGATTCAGT
AGTGTGGTATCTATTCAGCGCAGGCAACGAGGCAGACGTGCATGGAATTTATTTTTCAGGAAACACTTATCTTTGGA
GAGGCGAAAGACGTGACACAGCAAATCTTTTCCCACAGACGAGTCTAACGCTACACATGTGGCCAGATACGGAAGGA
ACTTTCAATGTGGAGTGTCTTACGACCGATCATTACACCGGGGGAATGAAGCAGAAGTACACCGTCAACCAGTGCAG
GAGACAAAGCGAGGATTCAACCTTCTACCTTGGAGAACGTACATACTACATTGCAGCTGTCGAGGTGGAGTGGGATT
ATTCTCCTCAGCGTGAATGGGAAAAAGAACTTCATCACCTGCAAGAGCAAAATGTCTCTAACGCGTTTCTCGACAAA
GGTGAGTTCTACATAGGATCAAAGTATAAGAAAGTGGTTTACCGTCAATATACAGACAGTACATTTAGAGTCCCCGT
GGAAAGAAAGGCAGAGGAAGAGCACCTGGGCATACTAGGCCCTCAGTTGCACGCGGACGTCGGGGACAAAGTAAAGA
TCATATTTAAGAACATGGCTACTAGGCCGTATAGCATACATGCACACGGAGTACAAACCGAGAGTTCCACAGTCACA

CCAACGTTGCCGGGCGAAACTCTGACTTATGTGTGGAAAATTCCTGAAAGATCAGGTGCTGGAACGGAAGACAGTGC
TTGCATTCCCTGGGCGTACTATTCAACAGTCGATCAGGTCAAAGACCTTTACAGTGGTTTAATAGGGCCTCTAATAG
TCTGTAGAAGACCTTACCTTAAAGTCTTTAATCCTCGTAGGAAACTTGAGTTTGCGTTATTATTCCTGGTCTTCGAC
GAGAACGAAAGCTGGTACTTAGACGACAACATTAAAACGTACTCCGACCACCCAGAGAAAGTTAATAAGGACGACGA
GGAATTTATAGAATCTAACAAGATGCACGCTATCAACGGTAGAATGTTTGGGAATCTTCAGGGTTTGACGATGCACG
TGGGAGATGAGGTCAACTGGTATTTGATGGGTATGGGAAACGAAATAGATCTGCATACCGTCCATTTTCATGGCCAT
AGTTTCCAGTACAAACATAGGGGAGTCTACTCAAGCGACGTATTTGATATTTTCCCGGGTACCTACCAAACTCTGGA
AATGTTTCCCCGTACGCCGGGTATTTGGCTGTTACATTGTCATGTAACGGATCACATTCATGCAGGGATGGAAACGA
CCTATACAGTATTGCAAAACGAGGATACAAAAAGCGGA

Seq. ID No: 271
>oxidase_2
ATGGGACTGAACTCCGCGATACCGTCTTTAGCGATTTTGGCACTGTCCGTAGGATCTTATGCCGCAATCGGGCCAGT
ATCTGACCTACACATTGTAAACAAAGATTTGGCACCCGACGGCGTGCAGAGGCCCACAGTATTAGCTGGCGGCACGT
TCCCCGGAACGTTAATAACTGGACAGAAGGGTGATAATTTTCAACTTAACGTTATCGATGACCTTACCGATGACCGT
ATGCTAACCCCTACGAGCATCCATTGGCATGGCTTTTTCCAGAAAGGTACTGCCTGGGCAGACGGTCCAGCTTTTGT
TACGCAATGCCCGATTATTGCGGACAATAGTTTCTTATACGATTTTGATGTGCCAGACCAGGCGGGAACTTTCTGGT
ACCACTCCCATTTATCAACACAATACTGCGACGGTCTACGTGGGGCATTTGTCGTATATGACCCGAACGATCCCCAT
AAAGATCTTTATGATGTCGATGATGAAAGTACAGTCATCACACTAGCAGATTGGTACCACGTCTTAGCGCAAACAGT
GGTAGGCGCTGCTACGCCAGACAGCACATTGATTAATGGCCTAGGTAGGTCCCAAACTGGTCCCGCAGACGCAGAGT
TAGCCGTAATCAGCGTAGAGCACAATAAGAGGTATAGATTCAGACTTGTGTCTATATCTTGCGACCCAAACTTCACC
TTCTCTATAGATGGTCACAACATGACCGTAATCGAGGTCGATGGGGTAAATACCCGTCCTTTAACCGTGGATTCCAT
TCAAATTTTCGCCGGCCAAAGATATTCTTTTGTCTTGAATGCAAATCAGCCGGACGACAACTACTGGATTAGAGCCA
TGCCGAATATAGGTAGGAATACGACGACCCTAGACGGTAAGAACGCGGCTATACTAAGATATAAAAATGCTTCCGTA
GAAGAGCCCAAGACAGTTGGCGGTCCAGCTCAGAGCCCATTGAATGAAGCCGACCTTCGTCCGCTTGTGCCAGCACC
CGTGCCGGGAAATGCGGTCCCGGGAGGTGCAGATATTAATCACAGATTGAACCTGACATTTTCCAATGGGCTGTTCT
CCATAAATAACGCAAGTTTCACAAACCCATCCGTACCGGCATTACTGCAGATCCTGAGCGGTGCACAAAACGCGCAG
GACCTTCTGCCAACAGGTTCCTACATTGGGTTAGAGCTGGGGAAGGTGGTAGAACTTGTCATCCCGCCACTAGCGGT
AGGGGGTCCTCATCCATTCCACTTACACGGGCATAACTTCTGGGTTGTCAGATCAGCCGGGAGCGATGAATATAACT
TCGACGATGCTATACTTAGGGATGTGGTTTCCATTGGTGCGGGCACGGACGAGGTGACCATTAGGTTTGTCACCGAC
AACCCGGGGCCCTGGTTCTTACACTGTCACATAGATTGGCATCTGGAAGCGGGACTAGCTATCGTATTCGCTGAAGG
TATAAACCAGACTGCCGCTGCAAACCCGACCCCTCAAGCATGGGACGAACTATGCCCTAAATACAACGGGCTTAGCG
CGAGCCAAAAGGTTAAACCGAAAAAGGGTACCGCTATT

Seq. ID No: 272
>oxidase_3
ATGAGTAGGTTCCAGTCTCTGCTATCTTTCGTACTTGTAAGTTTGGCGGCCGTTGCCAATGCTGCTATAGGCCCTGT
TGCCGATTTAACTTTGACAAATGCTGCAGTGAGTCCGGACGGATTTAGCAGAGAGGCTGTCGTCGTTAATGGAATTA
CCCCAGCCCCTCTAATTGCTGGTCAGAAAGGCGATAGATTCCAATTAAATGTCATTGATAACTTAACAAACCATACG
ATGCTAAAAACCACATCAATCCACTGGCATGGCTTCTTCCAGCACGGCACGAACTGGGCCGATGGCGTGTCTTTCGT
GAACCAATGCCCAATCGCCAGCGGCCACTCCTTTTTGTACGATTTTCAAGTCCCGGACCAGGCAGGCACATTTTGGT
ATCATAGTCACCTGAGTACGCAGTACTGTGATGGTTTAAGGGGTCCGTTTGTGGTCTACGACCCGAATGATCCCCAA
GCGTCCTTATATGATATCGACAACGACGACACAGTAATCACGTTGGCAGATTGGTACCATGTTGCAGCCAAGTTGGG
CCCAAGGTTTCCTCTAGGAGCAGACGCCACCCTAATTAATGGACTAGGCCGTTCACCCGGGACAACAACAGCCGACC
TAGCCGTCATTAAAGTGACCCAAGGCAAGAGATATAGATTTCGTTTAGTTTCTCTTAGCTGTGACCCTAACCACACC
TTTTCTATTGATGGCCATACGATGACTGTGATCGAAGCTGACAGTGTAAACACCCAACCACTTGAAGTAGATTCAAT
TCAGATCTTTGCGGCTCAGCGTTACAGTTTCGTTTTAGATGCGTCCCAACCAGTGGATAACTACTGGATTAGGGCAA
ACCCAGCATTTGGGAATGTGGGTTTTGCAGGTGGAATTAACTCTGCGATTCTGAGATACGATGGAGCCCCCGAAGTT
GAGCCGACCACAACCCAGACCACCTCCACCAAACCGCTTAATGAAGCTGACCTACATCCACTGACACCTATGCCCGT
TCCGGGTAGACCTGAGGCCGGGGGCGTAGACAAGCCACTTAATATGGTATTTAACTTTAATGGAACTAATTTCTTCA
TTAATAATCATTCCTTCGTACCCCCGAGCGTACCCGTTCTGCTTCAGATTCTATCCGGCGCGCAGGCTGCACAAGAC

CTTGTCCCAGACGGGAGCGTGTACGTCTTACCGTCAAATTCAAGCATAGAAATATCTTTTCCAGCCACCGCCAATGC
CCCGGGCACGCCCCATCCGTTTCACCTGCATGGCCACACTTTCGCCGTCGTCAGGTCTGCCGGATCTAGTGAATATA
ACTATGACAACCCAATCTTCAGGGATGTTGTCAGCACGGGTCAGCCAGGCGATAACGTGACTATAAGGTTCCAGACT
AATAATCCCGGTCCCTGGTTTTTGCATTGCCATATAGACTTTCACCTAGAGGCCGGATTTGCAGTTGTTCTTGCGGA
GGATACCCCTGATACGGCAGCAGTGAATCCAGTGCCACAGTCCTGGAGCGACTTATGCCCAATCTATGACGCGCTAG
ACCCTTCTGACTTG

Seq. ID No: 273
>oxidase_4
ATGAAATTCTTGTTACTATCCGCACTGCTTTTCCTGCATTCTTCATTGGCTTGGACTAGGGAGAAGCACTACTACAT
CGGGATTACCGAGGCGGTATGGGATTACGCGAGCGGAAGCGAAGAAAAAGAGTTAATTTCCGTAGACACAGAACAAT
CCAATTTCTACCTTAGGAATGGCCCTGACAGGATCGGTCGTAAATACAAGAAAGCCTTGTACAGCGAGTACACAGAC
GGCACATTTACAAAAACCATAGACAAACCAGCCTGGTTGGGATTCCTGGGCCCCGTAATAAAGGCCGAAGTAGGGGA
TAAGGTTTCCGTACACGTGAAGAACTTCGCCTCACGTCCTTACACATTTCATGCTCATGGTGTTACTTATACCAAAG
CAAACGAGGGTGCTATTTACCCTGACAACACAACAGACTTCCAGAGAGCAGACGATAAGTTATTTCCGGGTCAGCAA
TATCTATACGTTCTTCGTGCGAATGAACCAAGTCCTGGCGAAGGGGATTCCAATTGCGTGACGAGGATATACCACAG
TCATGTTGACGCTCCAAAGGACATCGCTTCTGGACTGATCGGCCCCTTGATTCTATGTAAGAAGGGCAGTTTGCACA
AGGAAAAGGAAGAGAACATTGATCAAGAGTTCGTGTTGATGTTTAGCGTGGTCGATGAGAACTTGTCATGGTACTTG
GAAGACAACATCAAAACATTTTGTTCCGAACCAGAAAAGGTTGATAAAGACAACGAGGATTTTCAGGAATCTAACAG
GATGTATTCTATAAACGGTTACACATTCGGAAGTTTACCTGGACTAAGCATGTGCGCCGAAGACCGTGTAAAATGGT
ATCTTTTTGGTATGGGAAACGAGGTGGATGTTCATAGCGAACTGTTCCACGGACAAGCCCTTACCTCCAAGAATTAC
CATACAGACATAATTAACCTGTTTCCCGCCACATTAATTGACGTGTCAATGGTGGCACAGAACCCCGGCGTATGGAT
GCTATCCTGCCAGAATCTTAACCACTTAAAAGCTGGATTGCAGGCATTCTTCCAAGTCCGTGATTGTAACAAACCAT
CCCCTGACGATGACATCCAAGACCGTCATGTCAGGCACTATTACATAGCCGCGGAGGAAACGATCTGGGATTACGCG
CCATCAGGGACAGATACGTTTACAGGCGAAAATTTTACCTCCTTAGGATCAGATTCGAGAGTGTTTTTCGAACAAGG
AGCTACTAGAATCGGCGGATCGTACAAGAAGTTGGTCTACAGGGAGTATACGGATGATTCCTTTACTAACAGGAAGG
AGCGTGGCCCAGATGAGGAACATTTGGGGATATTAGGGCCGGTGATCTGGGCGGAAGTCGGCGATATCATTAGGGTT
ACCTTTCACAATAAGGGTCAGTTTCCGCTAAGTATACAACCTATGGGAGTGCGTTTCACTAAAGAAAATGAGGGAAC
TTATTATGGGCCAGATGGAAGGAGCAGCAAACAAGCGAGCCATGTCGCCCCGAAAGAAACTTTCACTTACGAATGGA
CTGTACCTAAAGAAATGGGACCGACCTACGCCGATCCCGTTTGTTTGTCCAAAATGTATTACAGTGGAGTAGATCTA
ACTAAAGATATATTTACAGGCCTTATTGGCCCTATGAAGATATGCAAGAAAGGAAGTTTATTGGCTGACGGTAGACA
GAAAGACGTAGATAAAGAGTTCTACCTTTTCGCGACGGTCTTTGACGAAAATGAGAGTCTTCTACTGGACGACAATA
TTAGAATGTTTACTACCGCGCCTGAAAACGTAGATAAGGAAGACGAGGACTTCCAGGAATCAAACAAAATGCATTCT
ATGAATGGCTTTATGTATGGCAACTTGCCCGGGCTAAACATGTGCCTTGGTGAATCTATCGTTTGGTATCTTTTCTC
TGCGGGCAATGAAGCCGATGTTCATGGCATATACTTTTCAGGCAACACTTATTTGTCAAAAGGAGAGAGACGTGACA
CGGCCAATCTGTTCCCTCACAAGTCCTTAACGCTTCTGATGACACCTGACACAGAGGGGAGCTTTGATGTCGAATGC
CTGACAACTGACCATTATACAGGAGGCATGAAGCAAAAGTACACGGTCAACCAATGCAAAGGACAATTCGAAGACGT
CACCCTGTACCAGGGTGAACGTACGTACTACATAGCGGCTGTAGAAGTAGAATGGGACTATAGTCCCAGCAGGGATT
GGGAGATGGAGCTTCACCATCTGCAGGAGCAAAACGTGAGTAACGCATTCTTAGATAAGGAAGAATTTTTCATCGGT
TCAAAATATAAGAAAGTGGTCTACAGAGAATTTACAGACAGCACATTCAGAGAGCAGGTTAAACGTCGTGCGGAGGA
AGAGCACTTGGGTATGTTGGGGCCCTTAATCCACGCAGACGTAGGTGCTAAGGTCAAGGTCGTCTTTAAGAATATGG
CTACGAGGCCGTATTCAATACACGCGCATGGGGTCAAAACCAAGAGTTCCACCGTAGCACCAACACTACCGGGTGAA
GTAAGAACCTACATATGGCAAATTCCCGAAAGAAGTGGTGCAGGTACCGAGGATTCACCCTGCATCCCGTGGGCTTA
CTATAGCACGGTAGATCGTGTAAAAGACTTATATAGCGGCCTGATAGGGCCGTTAATAGTATGCAGAAAATCATACG
TTAAAGTCTTCAATCCTAAGAAGAAGATGGAATTTTCTCTGTTATTTCTGGTGTTCGATGAGAACGAGTCATGGTAC
CTTGATGATAATATTAATACATATCCAGATCATCCTGAGAAGGATAATAAGGATAATGAAGAGTTTATTGAGAGTAA
TAAGATGCACGCAATCAATGGGAAAATGTTCGGAAACCTTCAGGGGCTGACCATGCACGTAGGTGACGAAGTAAACT
GGTACGTCATGGCTATGGGCAATGAGATCGATTTACACACAGTTCACTTTCATGGCCACTCTTTTCAGTATAAGCAC
AGAGGGATTCATAGTAGTGATGTATTCGACTTCTTTCCTGGAACGTATCAAACCCTAGAGATGTTCCCTCAAACGCC
CGGAACGTGGTTGCTACATTGCCACGTTACCGATCACATACACGCTGGTATGGTTACTACGTACACTGTGCTTCCAA
ACCAGGAAACCAAATCCGGG

Seq. ID No: 274
>oxidase_5
ATGAACTTCGTCACGGCGCTGCCTCTGATTGCGCAGTTGATAGGAACAGCAAGGGCGGCTATAGGCCCGGTCACAAA
CCTTTTGGTGAAAAACGCGGACATACCCCCGGACGGATTCACTAGGGCTGCGGTTGTCGCTAATAATCAGTTTCCTG
GTCCTGTTATCAGGGCCACTAAAGGAGATACGCTTTCCCTAAACGTGGTAAATCAGCTAACTGATGCCACCATGCTT
ATGGGGACTTCAATACATTGGCACGGATTCCACCAGAAGGGTACCTCTTGGGCCGACGGAGTTGTAGGCGTTACCCA
ATGTCCGATAGCTCCCGGACATAGCTTCTTGTATCAATTCCCTACAGCCAATCAAGCTGGGACTTTCTGGTACCACT
CTCATTATTCCACACAATACTGCGATGGTCTAAGGGGAGCGTTAATAGTTTATGACCCAACTGACCCATACAGGACC
TGGTATGACATTGACGACGAGTCTACCATAATTACACTAGCGGATTGGTACCATAAGGCTGCCCCTCTGCAAACGTT
AAGAACCGCGAAAGAGGATAGCGTACTGATCAACGGACAAGGCAGAGTCCCAGGGGATAAGACCACCGACAGTACAC
CGTTGTCCGTGATTAATATAATTCCTCAGAAGAGATATAGGTTTAGGCTGATTTCAATATCATGCGATCCAGCATTT
TCTTTTTCAATAGACGGGCATAGCATGACTGTCATCGAGGCCGATTCACAAAGCGTCCAACCGTTGACGGTCAACGA
GATAACTATTTTCGCGGGACAGCGTTATTCTTTCATCCTGTACGCGAACAACCCGGTTGGTAATTACTGGATTCGTT
CACAGCCTACATACCCTGACGACGGTATTCAAGGGTATGCAGGGGGTATCAACTCCGCAATTTTGAGGTATTCAGGC
GCTCCTGCCGTCAACCCAACTACGAAGAAGGCATCCATTACCATCCCGCTGGTAGAGGCCGATCTAAGACCATTATA
TAGTCCCGCCGCTCCTGGATTGCCTTCACCCGGTGCGGCAGATGTGAATATCAAACTGGATATAAGTTATAATTCTC
CGTCTGAGACATTTTTCGTAAACAACTCCACATTTCCTGAAGTTCCTGTGCCTGTCCTTTTGCAAATCCTGAGTGGC
GCGCAAAGCGCCAACGATTTGTTGCCCGCTGGCAGCGTGTATACCCTTCCGCCCAATAAAGTAATCGAGATTAGTAT
GCCAGGGGGCCGTCCTGGGAGTCCGCATCCTATGCATCTACATGGTCATGACTTCAGTGTGGTAAGATCTGCAGGAT
CAAATAGGTATAATTATGCCAACCCCGTTAGGAGAGACGTAGTGAATATTGGGATGGAGGACACTGATAATGTCACG
ATCAGGTTCCGTGTTTGCAGTCATACTTATTTGAGTTTACATTGTCATATCGATTTCCATTTGGAAGACGGACAGTC
TGGGACCCTAGTTCCACCTCTTCCACACAGATTGCCGCCACGTGGAAGGATCAGGTGTAGACTACATAGAGGCATTT
TAGTAAGAGGCAGGCTTGGACCTGACCTTCAG

Seq. ID No: 275
>phosphatase_1
ATGCAGGGGCCGTGGGTTCTTTTACTGTTGGGCCTGAGACTACAACTATCTCTGGGCATTATCCCTGTAGAAGAGGA
GAATCCAGATTTCTGGAACCGTCAGGCTGCCGAGGCCCTGGGTGCAGCTAAAAAGCTGCAACCCGCCCAGACGGCAG
CTAAGAATTTGATCATTTTCTTGGGAGACGGCATGGGTGTTTCTACTGTAACAGCGGCTAGGATTCTAAAGGGGCAA
AAGAAAGACAAACTTGGACCAGAAACGTTCCTTGCAATGGATCGTTTTCCTTATGTGGCGCTGAGCAAGACTTACTC
AGTGGACAAGCATGTCCCTGATTCCGGAGCAACCGCAACAGCTTATCTTTGTGGAGTCAAAGGTAATTTTCAAACCA
TAGGGTTGAGTGCGGCAGCAAGATTTAATCAATGCAACACAACAAGGGGGAACGAAGTGATTAGTGTAATGAATAGG
GCCAAGAAGGCTGGCAAGTCCGTAGGGGTCGTTACGACCACTAGAGTTCAACACGCGAGCCCCGCTGGGGCGTATGC
ACACACTGTAAACAGAAATTGGTATAGCGACGCCGATGTGCCTGCTTCTGCAAGACAAGAGGGATGTCAGGACATCG
CTACGCAATTGATAAGTAACATGGATATTGATGTTATTTTGGGCGGAGGCAGAAAATACATGTTCCCTATGGGAACG
CCGGACCCTGAATATCCTGACGACTATTCACAAGGGGGAACTAGGTTAGACGGAAAGAATTTGGTTCAGGAGTGGTT
GGCCAAACACCAAGGCGCACGTTACGTGTGGAATCGTACGGAGTTATTACAGGCATCTCTTGACCCATCTGTAACGC
ATCTGATGGGCCTTTTCGAGCCCGGTGACATGAAATACGAGATCCATCGTGATTCTACTTTGGACCCAAGTCTTATG
GAAATGACTGAAGCAGCTTTATTACTACTATCACGTAATCCGAGGGGATTCTTTCTTTTTGTTGAGGGAGGAAGAAT
CGACCACGGGCACCATGAGTCTCGTGCCTATAGGGCGTTAACCGAGACCATTATGTTCGACGATGCAATCGAGCGTG
CCGGGCAACTGACCTCTGAGGAAGATACTCTGTCCCTTGTAACCGCGGATCATTCTCACGTATTCTCATTTGGCGGA
TATCCTTTACGTGGCAGTTCTATCTTTGGGTTAGCTCCAGGTAAGGCGCGTGATAGAAAAGCCTACACAGTGCTTTT
GTACGGGAATGGCCCCGGTTATGTTTTGAAAGACGGCGCAAGACCAGACGTTACCGAATCCGAAAGCGGTAGTCCAG
AGTATAGGCAGCAAAGCGCAGTTCCTTTGGATGGGGAGACTCACGCAGGGGAAGATGTTGCTGTGTTCGCGCGTGGA
CCGCAGGCTCACCTTGTGCATGGCGTGCAGGAGCAAACCTTTATTGCCCATGTGATGGCTTTTGCAGCGTGTTTAGA
GCCCTATACTGCCTGTGACTTGGCTCCACGTGCGGGCACAACAGATGCTGCACACCCCGGCCCCTCTGTTGTACCGG
CTCTGCTTCCGCTACTTGCTGGCACGTTACTGCTATTGGGAACAGCTACCGCACCT

Seq. ID No: 276
>phosphatase_2

ATGCAAGGCGCTTGTGTGCTGTTGCTTCTTGGCTTGCACTTGCAACTATCACTTGGTCTAGTTCCTGTCGAAGAAGA
AGACCCAGCGTTCTGGAATAGGCAGGCCGCGCAAGCATTAGACGTCGCCAAGAAGCTGCAACCTATCCAAACGGCCG
CTAAAAACGTCATCCTGTTCCTTGGCGACGGGATGGGGGTCCCTACAGTCACAGCCACACGTATACTTAAAGGCCAA
ATGAATGGTAAATTGGGGCCGGAAACTCCACTGGCTATGGACCAGTTCCCCTACGTCGCATTATCCAAGACGTACAA
TGTAGACCGTCAGGTACCCGACTCCGCAGGAACCGCCACGGCTTATCTATGTGGTGTAAAAGGCAACTACAGAACGA
TTGGGGTCTCTGCTGCCGCAAGATACAATCAATGTAAAACAACTAGGGGTAATGAAGTCACGAGTGTCATGAACAGA
GCTAAAAAGGCGGGGAAGAGCGTAGGCGTGGTCACAACAACGAGAGTACAACATGCTTCTCCCGCGGGAGCCTACGC
ACACACGGTTAATAGGAACTGGTACTCTGATGCAGACTTGCCGGCCGACGCGCAAATGAACGGTTGTCAAGATATAG
CAGCGCAGCTAGTAAACAATATGGACATCGATGTAATATTAGGGGGTGGCAGGAAATATATGTTCCCAGTCGGGACG
CCAGATCCGGAATATCCCGACGACGCTAGTGTAAATGGCGTCAGAAAGAGGAAGCAGAATTTAGTTCAAGCCTGGCA
GGCTAAACATCAAGGCGCTCAGTACGTATGGAATAGGACCGCACTGCTTCAAGCTGCAGATGATTCCAGCGTGACAC
ATCTGATGGGCCTGTTTGAACCCGCAGACATGAAGTACAATGTACAGCAGGACCATACTAAAGATCCAACTCTACAG
GAAATGACTGAGGTCGCATTGAGAGTCGTTAGTAGAAACCCCAGGGGCTTTTACCTATTTGTCGAAGGCGGGAGAAT
CGATCACGGTCACCACGACGATAAAGCGTACATGGCCTTAACTGAAGCCGGCATGTTCGACAATGCGATAGCCAAAG
CCAACGAACTTACGAGTGAACTAGACACGTTAATTCTAGTAACTGCGGACCACAGTCACGTCTTTTCATTCGGCGGC
TATACTCTAAGGGGCACATCCATATTTGGGCTTGCACCGTCTAAGGCTCTCGATTCTAAGTCATACACCAGCATATT
GTACGGCAACGGACCTGGCTATGCATTAGGGGGTGGATCACGTCCCGACGTCAATGATTCAACCAGTGAAGATCCGT
CTTACCAGCAACAGGCCGCCGTGCCTCAAGCCAGCGAGACACACGGTGGCGAGGACGTTGCTGTGTTTGCTAGGGGA
CCTCAGGCACATTTGGTACACGGCGTCGAAGAAGAAACCTTTGTAGCCCACATAATGGCGTTCGCCGGATGTGTTGA
GCCTTATACCGATTGTAACCTGCCAGCTCCGACAACGGCAACTTCTATTCCAGACGCGGCACATCTTGCAGCCTCTC
CCCCACCTCTTGCCTTGTTAGCAGGTGCTATGCTGTTACTGCTTGCTCCTACGCTTTAT

Seq. ID No: 277
>phosphatase_3
ATGCAGGGTCCTTGGGTATTATTGTTGCTTGGATTGCGTCTGCAGCTATCCCTTGGTGTAATCCCAGCCGAAGAAGA
AAACCCAGCTTTCTGGAATAGGCAAGCAGCGGAGGCCCTGGACGCCGCCAAAAAGCTGCAGCCTATTCAGAAGGTAG
CCAAAAATTTGATATTGTTCTTGGGAGACGGTCTGGGTGTACCAACCGTAACAGCTACCCGTATCCTAAAGGGTCAG
AAGAATGGCAAGTTAGGACCTGAAACACCGTTAGCTATGGACAGATTCCCCTACCTTGCCCTGAGCAAGACTTACAA
CGTGGATCGTCAGGTACCTGATTCCGCTGCAACAGCAACCGCGTATCTTTGTGGAGTCAAGGCCAATTTTCAAACAA
TCGGACTATCTGCAGCCGCCAGATTTAACCAATGCAACACAACTAGAGGAAACGAGGTTATTTCTGTGATGAATCGT
GCAAAGCAAGCTGGAAAGAGTGTCGGGGTGGTAACAACGACGCGTGTTCAACATGCTTCTCCTGCGGGAACATATGC
GCACACTGTTAACAGGAATTGGTACTCAGACGCGGATATGCCCGCGTCAGCCAGACAGGAGGGCTGCCAAGACATCG
CCACTCAATTAATATCCAACATGGATATCGACGTCATTCTAGGAGGCGGTAGGAAGTATATGTTCCCTATGGGGACA
CCAGACCCTGAATACCCCGCTGACGCCTCTCAGAATGGCATAAGACTTGACGGGAAAAATCTAGTACAAGAATGGTT
AGCAAAGCATCAGGGCGCGTGGTATGTCTGGAATAGAACGGAGTTGATGCAGGCAAGCCTGGACCAGTCAGTTACGC
ACCTGATGGGGCTTTTTGAACCCGGCGATACAAAGTACGAGATCCACCGCGATCCCACACTTGACCCTTCTTTAATG
GAAATGACCGAAGCGGCACTGCGTTTGCTGTCCAGAAACCCCAGAGGCTTTTACCTTTTTGTGGAGGGTGGTAGAAT
AGATCACGGACACCACGAGGGCGTGGCATACCAAGCTCTAACTGAGGCGGTAATGTTTGATGACGCAATTGAAAGGG
CCGGACAACTGACCTCCGAAGAAGATACTTTGACCTTAGTCACCGCCGACCATTCCCATGTCTTTTCCTTTGGTGGT
TATACCCTGAGAGGTTCCTCTATATTCGGACTAGCTCCCTCTAAGGCGCAAGACAGCAAAGCATATACGAGTATCCT
GTACGGGAATGGACCGGGCTACGTATTTAACTCAGGTGTTAGGCCCGACGTAAATGAAAGTGAGTCCGGCTCCCCAG
ACTATCAGCAGCAGGCGGCTGTTCCACTTTCATCAGAAACCCACGGAGGTGAGGACGTAGCGGTGTTTGCCAGGGGC
CCCCAGGCTCATTTAGTTCACGGAGTTCAAGAGCAGTCATTCGTTGCTCACGTGATGGCATTTGCCGCATGTTTGGA
GCCCTACACTGCCTGTGATTTAGCTCCTCCCGCTTGCACAACAGACGCGGCACATCCTGTCGCAGCTTCCCTGCCAT
TATTAGCTGGAACTTTACTTCTTCTAGGAGCGTCAGCCGCTCCT

Seq. ID No: 278
>phosphatase_4
ATGCAAGGCGCCTGGGTATTACTTTTATTGGGTTTCCGTTTGCAGCTGTCTCTGAGCGTAATCCCTGTAGAAGAAGA
AAATCCGGCGTTTTGGACGCAAAAAGCTGCTGACGCGTTGAACGTGGCGAAGAAGTTACAGCCAATTCAGACAAGTG
CCAAGAATCTAATAATATTTTTAGGAGATGGTATGGGTGTTGCAACAGTAACGGCCACGAGGATCTTAAAGGGTCAG

CTGGAAGGGAATTTGGGTCCCGAAACACCGCTGGCGATGGACCACTTCCCGTATATGGCTCTTTCCAAGACTTACTC
AGTTGATAGACAGGTACCGGATAGCGCATCTACAGCTACTGCTTATCTGTGCGGCGTAAAGACAAACTATAAGACGA
TTGGGGTTAGCGCAGCTGCGAGGTTCGATCAATGTAACACAACGTTTGGCAATGAGGTTCTTTCTGTAATGTACCGT
GCAAAGAAGGCGGGTAAATCCGTGGGCGTGGGTGATCATACGAGGGTGCAGCACGCGAGTCCGGCTGGCACATACGT
GCACACCGTTACGTCTAATTGGTATGGGGATGCAGACATGCCAGCCCTGCCGTTGCAAGAGGGTTGCAAGGACATAG
CGACGCAGTTGATTAGTAACATGGATATAAACGTAATACTTGGTGGGGGCAGAAAATATATGTTTCCGGCTGGTACA
CCCGATCCTGAGTATCCAAATGATGTCAATGAAACAGGCACGCGTCTTGATGGCAAGAATTTAGTTCAAGAGTGGCT
ATCCAAACACCAGGGTAGCCAGTACGTTTGGAACAGGCAAGAGTTAATTCAAAAGTCCTTGGACCCTAGTGTTACCT
ATTTGATGGGTCTATTTGAGCCAGTAGACACGAAGTTCGAAATACAACGTGACCCTCTGATGGACCCCAGTTTGAAG
GATATGACAGAGGCTGCGCTTCACGTGCTTTCTAGGAACCCTAAAGGATTCTACCTATTCGTAGAGGGTGGCCGTAT
AGATAGAGGGCACCACTTGGGTACCGCCTATCTGGCACTAACTGAGGCCGTCATGTTTGATTCTGCGATAGAAAGAG
CATCACTACAAGCGTCTGAGCAAGACACCCTAACTATCGTGACGGCAGACCACAGCCACGTCTTTTCCTTCGGGGGT
TATACATTAAGGGGAACCTCCATCTTTGGGCTTGCCCCGTTGAATGCTTTGGATGGAAAGCCCTACACATCTATTTT
ATACGGGAACGGGCCGGGGTATGTGGGGACAGGGGAAAGACCCAACGTTACCGACGCCGAAAGCCACGATCCTTCCT
ACCAACAACAAGCAGCGGTCCCTGTTAAAAGTGAGACAACTGTGGGAAAGGATGTGGCAATCTTTGCGCGTGGACCT
CAGGCACACTTACTGCATGGCGTACAGGAGCAGAATTATATAGCACACGTTATGGCTTTTGCTGGGTGCTTGGAACC
CTACACGGATTGCGGCCTAGCTCCACCTGCTGACGAGAATAGACCTACGACACCTGTACAGAACTCCACGACAACGA
CTACGACTACTACCACAACCACTACAACCACCACTACTACTAGAGTCCAAAACTCAGCTTCCTCACTTGGGCCAGCA
ACCGCCCCTCTTGCTTGGCATTACTGGCCGCGTCGT

Seq. ID No: 279
>phosphatase_5
ATGTCAGGATCTAGCGTTACAGGCGGGGGTGCAAGCCTACCTGCGGAATTGTACAAGGGGTCCGCTGACTCAATCTT
GCCAGCCAATTTCTCTTATGCTGTTACGGGTAGCGGTACGGGTAAAAATGCGTTCCTGACGAACAACTCTAGCTTGT
TCGGTACAACCGGCACAGTTCATTACGCGGGGTCCGATTCTGTTCTGTCCGGCAGCGAATTGACCACCTACAATTCA
AACTACAATGGCACCTATGGTCCCTTAATCCAGATACCTAGTGTGGCGACTTCTGTTACCGTGCCGTACCGTAAAGA
TGGTAACACTACGCTTAATTTGACATCAGCACAATTGTGTGACGCATTTAGTGGAGCTAAGACAACTTGGGGTCAAC
TACTTGGCACAACTGACTCTACACCCATTAGGATAGTCTATAGGACAGGGTCATCAGGGACTACCGAATTGTTTACT
AGACACTTAAACTCTATTTGCCCCACAAGGTTTGCGACAAACAGCACCTTCACCAATGCGAGATTACCCGCCGGAGG
GACTCTACCTAGTAATTGGGTTGGTGTAGCAGCTACGTCTACAGTAGTGAGCACGGTGAAAGCTACGAACGGCAGTC
TGGGTTATGTATCTCCTGACGCAGTTAACATAAATTCCAATGCGGAGGTTAGTAGGGTCAATGGAAACTTACCTACA
CAGGCCAATGTGTCAACCGCGTTGGGGTCCGTGGCACCCCCGGCCAATGCGGCAGATCGTGCAGATCCATCAAAGTG
GGTCCCAGTATTCACAAATCCTAGTGCCGGTTATTCAATCGTGGGATATACTAATTTCGTGTTTGGCCAGTGCTATA
AAGACGCCTCTGTCTCCACGGACGTTAGAGCCTTTATCAATAAGCACTATGGAGGCACAACCACCAATGCAGCGGTC
GCGGCGCATGGCTTCATTCCGTTGACACCAGCCTGGAAATCCGCTATAGTATCCGCCTTTTATACTGGTACGTCTGA
AAATTTAGCGATAGGCAATACAAACGTATGTAACACTAAGGGAAGGCCC

Seq. ID No: 280
>phosphatase_6
ATGAACTTCGTGACAGCCTTACCTCTTATCGCCCAATTGATAGGGACAGCTAGAGCGGCTATCGGACCGGTAACTAA
CCTACTGGTCAAGAATGCGGACATTCCGCCCGATGGTTTTACACGTGCTGCGGTGGTTGCTAACAATCAGTTTCCTG
GTCCGCTAATCACGGGTAATCAGGGGATTAATTTTCAGATCAACGTAGTGGCCAGATTAAATAACGAGTGTAGTTCC
CGTCCTCTGTGGTATCATTGGCATGGTTTCTTCCAGAAGGGAACGAATTGGGCTGACGGTCCGGCATTTGTAAACCA
ATGTCCAATTTCTACTGGAAATTCATTTCTTTATGACTTCACGGCTGCCGACCAGGCTGGTACGTTCTGGTACCATT
CCCACCTAAGTACCCAATATTGTGACGGTCTGAGAGGGCCGATGGTCGTATATGATCCAAACGATCCCCACGCAAGT
CTATATGACGTAGACGACGAAAGCACCGTAATAACCCTTAGTGACTGGTACCACACTGCCGCCCGTTTAGGCTCTGC
CTTTCCGCTAGGTCCGGACTCAGTGCTAATCAATGGTCTAGGTCGTTTTGCAGGAGGGGATGGAAATGCTGAACTAG
CGGTAATCCCAGTAACTCAGGGCAAGCGTTATAGATTTCGTCTGATTAGCTTATCTTGTGACCCTAACTTCGTCTTT
AGCATCGATAAACACAACATGACCGTTATCGAGGCGGACGCCGTCAGTCACGAGCCTGTGACGGTAGATAGCATACA
TATCTACGCGGGACAGAGATACTCATTTGTTCTATCCGCTCATCGTGACATAGACAACTACTGGATACGTGCTCTTC
CTTCAGGCGGCACTGTAAATTTCGTCGGGGGCGTGAACTCAGCCTTAATAAGATATGATGGTGCAGCGGAAGTTGAA

CCAGTTACCAACACTACGATGAGTATTGCCCCACTTGTTGAAACTGACCTTGTACCACTGGACAGTCCAGCGGCTCC
CGGCGAGGCCAGCATAGGAGGAGTGGACTATGCACTTAGTCTTGTTCCATCTTTCGTTAGTCGTACCTTGTTTTGCG
TCCGTTCAATAGCAGATGACCTTAGGTTACGTACGGGAAGAACCATATTAAGCGGTTCAACGGAGCTACCCTCCTCC
CATCCCCCTTGCAGAGTATATACTTTGCCCTCTAATGCTACGATAGAATTGTCCTTTCCAATCACCGCCACAAACGC
GCCGGGGGCTCCCCATCCCTTCCACTTGCATGGACACGTTTTTAGTGTTGTTAGATCAGCGGGGTCTAGTGAGTACA
ATTACGCAAACCCTCCACGTAGAGATGTGGTCAATACAGGTACAGCCGGTGACAATGTAACTATTCGTTTTAGGGTA
TGTTCTCACACTTACTTGTCATTGCATTGTCATATTGACTTTCACCTGGAGGACGGTCAAAGTGGGACGTTAGTACC
ACCTCTGCCTCACCGTCTGCCACCGAGGGGGAGGATTCGTTGTCGTTTACACAGAGGTATTTTAGTCAGAGGAAGAT
TAGGTCCTGACCTACAA

Seq. ID No: 281
>scaffold
GGGGGCGCTAGTGCGTCCGGTGTGGATAATAAATTTAATAAAGAGAGAGTCATTGCTATTGGTGAAATTATGAGATT
ACCGAATTTGAATAGCTTGCAAGTAGTCGCCTTCATCAATTCCCTTCGTGACGATCCTTCTCAAAGTGCCAACCTTC
TGGCGGAGGCGAAGAAGTTAAATGACGCCCAGGCCCCTAAAGGAGGGAGTTCTGCTTCCAGTGCGGGCGGATCATCA
GTAGATAATAAGTTTAACAAGGAGGCGCAAACCGCCGGGGTTGAAATAATGGAGTTGCCAAACTTGAATACACGTCA
GCTATTGGCTTTCATTCAAAGTTTACGTGATGATCCCTCTCAGAGCGCGAATTTATTAGCGGAAGCTAAGAAATTAA
ACGATGCTCAGGCCCCAAAGACAAGTGGAGGCTCCAGCGCGTCATCTGCAGGTGGATCAAGCGTGGATAATAAGTTC
AACAAAGAGAGGAGGATGGCGGCCTACGAGATAATCGATCTGCCTAATCTAAATTGGTTTCAATTAGAGGCGTTCAT
CACGAGCCTAAGCGATGACCCGAGCCAAAGCGCAAATTTGCTTGCCGAAGCGAAGAAACTTAACGATGCGCAGGCAC
CCAAGGGATCAGGCTCTAACGCCGCAATTCGTTCTAGCGGGAGTGGGTCCGTAGATAACAAATTTAACAAGGAAAGA
GTAATTGCTATTGGAGAAATCATGAGGTTGCCTAACCTTAACTCCCTTCAGGTAGTGGCATTTATAAACAGTTTACG
TGATGACCCTTCACAGAGCGCAAACCTACTTGCAGAAGCCAAGAAATTAAATGATGCACAGGCGCCCAAAGGAGGGT
CAAGTGCGTCCTCTGCAGGAGGGAGTAGTGTTGATAATAAATTTAATAAAGAAGCCCAAACAGCGGGTGTTGAAATT
ATGGAACTACCGAACTTAAACACCCGTCAATTACTGGCATTTATTCAGAGCCTGAGAGATGATCCATCTCAATCCGC
TAACCTCCTAGCTGAGGCGAAGAAACTTAATGATGCCCAGGCGCCAAAAACCTCAGGTTCAGGTTCAGCAAACGCTG
CTATTCGTAGTGCAGGATCTGGAAGCGTAGATAATAAATTCAACAAAGAACGTCGTATGGCTGCTTATGAGATCATC
GACTTACCTAATCTAAACTGGTTTCAACTTGAGGCTTTCATCACTTCATTATCTGATGATCCATCCCAGAGCGCCAA
CCTACTAGCGGAAGCGAAGAAGCTAAACGACGCTCAAGCTCCTAAAGGTTCAGGGGCGTCTGGTTCTGGTGCGGGCG
GCAGTTCCGTAGATAACAAATTTAACAAAGAAAGAGTTATCGCCATAGGGGAAATTATGCGTCTGCCGAATCTGAAT
AGTCTGCAAGTCGTCGCATTTATAAACTCTTTACGTGATGATCCCAGCCAATCTGCGAATTTACTGGCAGAGGCTAA
GAAGCTAAACGATGCGCAAGCCCCGAAGGGCTCTGGCGCGTCTGGTAGTGGGGCAGGTGGGAGCAGCGTAGACAACA
AGTTTAACAAGGAGGCCCAGACTGCAGGAGTCGAAATCATGGAATTACCTAATTTGAACACCAGACAGCTGCTGGCG
TTTATACAATCTCTTAGAGATGATCCTAGCCAGTCCGCTAATTTACTAGCCGAGGCCAAAAAGTTGAATGACGCACA
GGCACCGAAGACAAGTGGTTCCGGGTCCGCTAACGCAGCGATAAGGTCCGCGGGTTCCGGATCTGTGGACAATAAGT
TTAATAAAGAGCGTAGGATGGCCGCGTACGAAATAATCGATCTTCCTAATCTAAATTGGTTTCAGTTAGAGGCCTTT
ATTACTAGCCTTTCCGATGACCCCAGTCAGTCAGCGAACCTATTAGCGGAGGCCAAAAAGCTGAACGACGCGCAGGC
ACCTAAG

Seq. ID No: 282
>sec_1
ATGCAGCTATTAAGGTGTTTTTCTATTTTCTCAGTTATAGCCTCCGTGTTGGCCCAAGAGCTAACTACCATCTGCGA
GCAGATCCCATCTCCTACGCTAGAATCAACCCCATATAGTTTATCAACGACCACAATCTTGGCTAATGGCAAA

Seq. ID No: 283
>sec_2
ATGCTATCTTTGAAAACGCTGCTGTGTACGCTGTTAACGGTTTCATCAGTTCTAGCCACACCCGTACCCGCTCGTGA
TCCATCCTCTATTCAATTTGTCCATGAAGAAAATAAGAAGCGTTATTATGATTACGATCATGGAAGCCTTGGTGAA

Seq. ID No: 284
>sec_3

ATGAAACTTCAGTCACTGCTGGTAAGCGCGGCAGTCCTGACCAGCTTAACGGAGAACGTGAACGCATGGAGCCCAAA
TAACTCTTACGTTCCCGCCAATGTCACTTGTGATGACGATATTAACCTTGTGAGGGAGGCAAGTGGATTAAGCGACA
ATGAGACTGAATGGTTGAAGAAGCGTGATGCGTACACCAAAGAG

Seq. ID No: 285
>sec_4
ATGGAGGGCGTGTCACTAGAAAAACGAGAGGCTGAGGCA

Seq. ID No: 286
>sec_5
ATGAAAAAGACAGCGATTGCTATAGCGGTAGCCTTGGCGGGCTTTGCCACAGTCGCTCAGGCC

Seq. ID No: 287
>vac_1
ATGTTCTCTTTGAAAGCTCTACTACCCTTGGCTTTGTTGTTGGTAAGCGCAAATCAAGTTGCAGCGAAGGTTCATAA
GGCTAAAATATATAAACACGAACTGTCA

Seq. ID No: 288
>vac_2
ATGACGAAGAACTTTATTGTTACACTGAAAAAGAACACTCCAGATGTCGAGGCGAAAAAGTTCCTGGATAGTGTACA
TCACGCAGGGGGTTCCATAGTACATAAATTTGACATAATTAAGGGCTATACAATAAAAGTTCCAGATGTGTTGCACC
TAAACAAACTAAAAGAGAAACATAATGACGTGATAGAGAACGTGGAAGAGGACAAGGAGGTCCACACCAAT

Seq. ID No: 289
>vac_3
ATGGAAGAGCAGAGGGAAATCCTTGAGCAATTAAAGAAGACATTACAAATGCTAACTGTGGAGCCTAGCAAGAACAA
CCAAATCGCAAATGAGGAGAAAGAGAAAAAGGAAAACGAGAACTCCTGGTGCATTCTTGAACACAATTATGAAGACA
TCGCGCAGGAGTTTATTGATTTTATCTATAAGAATCCCACCACCTACCATGTCGTTTCATTCTTCGCAGAATTACTT
GACAAGCATAATTTCAAGTACCTATCTGAGAAGAGCAACTGGCAGGACAGTATCGGCGAAGATGGCGGT

Seq. ID No: 290
>6xHIS
HHHHHH

Seq. ID No: 291
>AADC_1
METVNKSCCSLATPHLPLMSPHLLHSHRDSPVLKIIHSIVLTVQNNHSCLQAFLGIGTDNVILVKTNDRGKMIPEDL
DHKIQKVKSEGSVPFLVSTTCGTTVFGAFDPLEGIADVCERHSLWLHVDAAWGGSALLSSRHRHLLKGIERADSVTW
NPHKLLGVGLQCSAFLLRDTTQLLERCHAANATYLFQTDKFYNLQYDTGDKSIQCGRRVDCLKLWLMWKALGSKGLE
TRVDRVLDHTRYLVEEMKRREGFRLIMEPEFVNLCFWYVPPSLRNKENSPDFWTRLGKVAPVIKERMMKKGSMMVGY
QPHGNMVNFFRQIVVNPEVTKEDLDFFLDEIERLAEDL

Seq. ID No: 292
>AADC_2
MWGCGNGDCIHVLLLISHTSPPPLSPHLLHSHRDSPVLKIIHSIVLTVQNNHSCLQGHVPFYVSATAGTTVYGAFDP
FVKIADICQKHGLWMLHVDAAWGGGLLMSRKHRHKMNGIERADSVTWNPHKMMGVLLQCSAILLKEKGILQGCNQMC
AGYLFQQDKQYDISYDTGDKAIQCGRHVDIFKFWLMWKAKGTVGFEQQINKCLELSEYLYSKICNREDFEMVFKGEV
SLHRLEEGIK

Seq. ID No: 293
>AADC_3

METVNKSCCSLATPHLPLMSPHLLHSHRDSPVLKIIHSIVLTVQNNHSCLQGHVPFYVSATAGTTVYGAFDPFVKIA
DICQKHGLWMHVDAAWGGGLLLSKKHRTKLSGIERANSVTWNPHKMMGVPLVKDNMDLLKRCHSAEASYLFQQDKFY
DVRYDTGDKSIQCSRRADAFKFWMMWKALGTLGLEERVNRALALSKYLAKEIKKRDGFELIWEPEYANICFWYIPPS
LRKMENKGPEYWQRLNQVAPTIKERMMKKGSMMIGYQPHRDKVNFFRHIIISPQVSREDMDFVLDEIDLLGRDL

Seq. ID No: 294
>AADC_4
MNASEFRRRGKEMVDYVANYMEGIEGRQVYPDVEPGYLRPLIPAAAPQEPDTFEDIINDVEKIIMPGVTHWHSPYFF
AYFPTASSYPAMLADMLCGAIGCIGFSWAASPACTELETVMMDWLGKMLELPKAFLNEKAGEGGGVIQGSASEATLV
ALLAARTKVIHRLQAASPELTQAAIMEKLVAYSSDQAHSSVERAGLIGGVKLKAIPSDGNFAMRASALQEALERDKA
AGLIPFFMVATLGTTTCCSFDNLLEVGPICNKEDIWLHVDAAYAGSAFICPEFRHLLNGVEFADSFNFNPHKWLLVN
FDCSAMWVKKRTDLTGAFRLDPTYLKHSHQDSGLITDYRHWQIPLGRRFRSLKMWFVFRMYGVKGLQAYIRKHVQLS
HEFESLVRQDPRFEICVEVILGLVCFRLKGSNKVNEALLQRINSAKKIHLVPCHLRDKFVLRFAICSRTVESAHVQR
AWEHIKELAADVLRAERE

Seq. ID No: 295
>AADC_5
MASGYPGAGAQQPPAAPASGSGSPVSMPYYASELARAKHEDDMKMPEHGIEPRHCLRRIEDYHLLDFSERLNTSSYV
NVVFEPEEETVANMGLKVNLADQTVYPESFRMHNDTVNMIAKLWNCPKPADFDEYGCYAGAGTVGSTEACLLGGLAL
KFRWRKWYAAKHGMDQNKVRGVYPNLVITTMFQAAWEKLFKYMDIEPRFVTPSWKTFTMDPSGLEKVVDDKTIGVVC
IMGNHYGGQYDPVWEVNDVLEKINKEKGLQVGIHVDGASGGFIAPFQEGLPAWDFRLKNVLSISASGHKFGNSCCGT
GWIIWRERKGLSDTVAINVSYLGGSADSYTLNFSRPAQGVYVQ

Seq. ID No: 296
>AADC_6
MAGSIDVMLRDLNLDTVESLPADFDPTTVINDILPPVVNDAKLVNGIEKPREIVLGRNVHTSSLEVTEPDADDEVTG
EREAYMASVLARYRKSLLERTKHHLGYPYNLDFDYGALSQLQHFSINNLGDPFIESNYGVHSRQFEVGVLDWFARLW
ELERNEYWGYITNCGTEGNLHGILVGREVFPDGILYASSESHYSIFKAARMYRMDCEKVNTLISGEIDCEDFKAKLS
LHKDKPAIINVNIGTTVKGAVDDLDLVIKTLEESGFSHDRFYIHCDGALFGLMMPFVKLAPKVSFKKPIGSVSVSGH
KFVGCPMPCGVQITRLEHINALSRNVEYLASRDATIMGSRNGHAPLFLWYTLNRKGYRGFQKEVQKCLRNAHYLKGR
LTEAGIGAMLNELSSTVVFERPQDEEFTRKWQLACQGNIAHVVVMPNINIDKLDHFVNELVERRAVWYENGKLKSPC
VASEIGNSNCLCALHK

Seq. ID No: 297
>AADC_7
MVRAVEKPVQAIVNAAFRGKDAYHVFRTTVLAAVLLRLWRHLRRVMAHEGLKAYFMSLVAPHLKKLPYVQNKLKKEM
DKTMTKMRNTFRKEVTDPRTSLPVDGIPEKQILELIQHRKELDTKEWTKGMTTGCVYHGGQDHYDFVGQIFASWGFA
NPLHPTTFPSLRQMDSEVVQMVINMYHGDSECCGAFTTGGTESILMAMKAYRDWGKAEKGITDPNIVICNTAHAAFD
KAGKYFNIFVKHARTNSEMEIDLGHLRSLIDSNTVAIVGSACQFSHGTVDPIQEMAKIAMKRRVGLHVDCCLGGFLV
PFMEKAGFQLPPFDFRVKGVTSISCDPHKYGFAPKGSSVVMFSNRHLRHYMYCFLTEWSGGIYATATMTGSRAGGPV
AATWASMCKFGEKGYIETTKQIVGATKKIAAGIAEIEGLRVVGRPDVCVVAFTCTEGSGMNCYAVGDCMHQDFHWEL
QSCQNPACVHLALTLPTSRNADKFVADLRQAVEAVRSDKDGKFASTAGMYGTAASLPAAFFEDGAAAYLDAMCEAIP
AGDALLPEEPATKESPAAAGAPAQATGGA

Seq. ID No: 298
>AADC_8
MNASEFRRRGKEMVDYVITNYLEQIELRQVYPSVEPGYLRPMIPDSAPEEGETYEDIMKDIERVIMPGVTHWNSPYF
FAYFPAATSYPAMLADMLCGSLGCIGFSWAASPACTELETVMLDWLGKTIGLPEQFLAGTNGEGGGVIQGTASEATL
MALLAARTKVTRRLQAENPDLSEAEIISRMVAYSSDQAHSSVERAGLISGVRMKKIPSDENFTARGEALKKALEEDK
AEGFIPVFLCATLGTTTSCAFDNLMELGPICNAENMWLHIDAAYAGSAFICPENRYLMKGVEFADSFNFNPHKWLLV
NFDCSAFWVKKRSDLICAFKIDPVYLQHDQQESGLVTDYRHWQIPLGRRFRSLKLWFVLRMYGVKGLQAHIRKHIRL

AQEFHEFVKNDDRFEICAPVILGLVCFCLKGSNTLNKSLLQKINTLKKIHLVPSCLGDKFILRFAVCARTLESNHIV
FAWKHIEELATEVLKEGEKQ

Seq. ID No: 299
>AADC_9
MQNCNQMHASYLFQQDKHYDLSYDTGDKALQCGRHVDIFKLWLMWRAKGTTGFEVQIDKCLELAEYLYDKIKNREGY
ELVIEGKPQHTNVCFWYIPPSLRHMEDNEERMARLVKVAPVIKARMMEYGTTMVSYQPLGDKVNFFRMVISNPAATH
QDIDFLIDEIERLGQDL

Seq. ID No: 300
>AADC_10
MWGCGNGDCIHVLLLISHTSPPPLSPHLLHSHRDSPVLKIIHSIVLTVQNNHSCLQGHVPFYVSATAGTTVYGAFDP
FVKIADICQKHGLWMHVDAAWGGGLLLSKKHRTKLSGIERANSVTWNPHKMMGVPLFQCSAFLLRDTTQLLERCHAA
NATYLFQTDKFYNLQYDTGDKSIQCGRRVDCLKLWLMWKALGSKGLERRVDRVLDHTRYLVEEMKNREGFRLIMEPE
FVNLCFWYVPPSLRNKENSPDFWTRLG

Seq. ID No: 301
>AADC_11
MGSLGTNPTSFSAFPDDKAAFEPLNPEDVRAYLHKAVDFISDYYTNVESMPVLPNVKPGYLQDELTASPPTHSAPFD
VTMKELRTSVVPGMTHWASPNFFAFFPSTNSAAAIAGDLIASAMNTVGFTWQASPAATEMEVLALDWLAQLLHLPTT
FMNRTSTGRGTGGGVILGTTSEAMLVTLVAARDAALRRSGSVGVSDIPRLAVYAADQTHSTFFKACRLAGFDPANIR
SIPTGPETNYGLDPAKLLEVMQADADAGLVPTYVCATVGTTSSNAVDPVGAVADVAAMFNAWVHVDAAYAGSACICP
EFRHHLDGVERVDSISMSPHKWLLTCLDCTCLYVRDAHRLSDSLETNPEYLKNDVTDSGEVTDLKDMQVGVGRRFRG
LKLWMVMRTYGTAKLQEHIRSDVAMAKMFEDSVRADNRFEVVVPRNFALVCFRIKARGDMTEEDADEVNRLLMENLN
KTGKAYLAHTVVGDRFVLRFAVGSSLQEERHVRSAWDLIKKTTSSIMD

Seq. ID No: 302
>AADC_12
MDPLKAVEMVDENTICVAAILGSTLTGEFENVKLLNELLTKKNKDTGWDTPIHVDAASGGFIAPFLYPDLEWDFRLP
LVKSINVSGHKYGLVYPGVGWVVWRSKGDLPDELIFHINYLGSDQPTFTLNFSKGNNISTHAYKKPKCRFYFLTNQI
LSRRYLLHRLKSGTSLNSLL

Seq. ID No: 303
>AADC_13
MQPGYLSRMLPDSAPNHPESLEDIFNDISAKILPGVTHWQSPNYFAYFPSNSSIAGFLGEMLSAGLNIVGFSWITSP
AATELEMIVLDWLAKLLKLPDDFLSGGTS

Seq. ID No: 304
>AADC_14
MVVDYKDWQIPLGRRFRSLKLWMVLRLYGIENLQCYIRNHIKLAQQFEVLVAQDLRFEIVSPRIFSLVCFRLLPSQN
CKDHGNELNHHLLDTVNSTGKVFLSHTVLSGKYILRFAVGAPLTEERHVTAAWKVLQDEASALLQSL

Seq. ID No: 305
>AADC_15
MGSLDIKQESSPLMTNPLDSEEFRRQGYMVIDFLAEYYKNIQKFPVRSQVEPGYLRKRLPESAPYEPESIERILKDV
HDDIVPGLTHWQSPNYYAYFPSSGSTAGLLGETLAAGFNVVGFNWISSPASTELESIVMDWLAEMLNLPKSFTFSGD
GGGVMMGTTCEAILTTITAARDRILDRIGREHINKLVVYGSDQTHCSFFKSAKIAGILPNNFRQVKTSRVNAFSMRP
DALRAAIQADADAGLVPFFLCTTVGTTSTAAVDPVALLCEVTKDYGMWVHIDAAYAGNACICPEFRHMINGVENADS
FSFNAHKWFLTTLDCCCLWVKDPSSLVRCLSTNPEYLKNKATDTQQVVDYKDWQITLSRRFRSL

Seq. ID No: 306

>AADC_16
MDGQMLKPMDAEQLREYGHQMVDFVADYYKTIESFPVLSQVQPGYLRELVPTNPPTHPESLQDVLDDVKAKILQGVT
HWQSPGYFAYFPCNSSTAGFLGEMLSAGINIVGFSWMSCPAATELEVIVLDWLSKL

Seq. ID No: 307
>AADC_17
MVLRLYGQEGLQSYIRNHIALAKQFEELVIQDSRFEMVTPRRFSLVCFRLLPRSTDENQADKLNRALLDAVNLTGAI
FISHTVLSGMYVLRLAVGAPLTEERHVIAAWKVLQEKATALLEGNVAQEPNGHAQLSNGVVELDGALDNGVTTEQHG
HVDDLPIKSK

Seq. ID No: 308
>AADC_18
MSLSRHIDAERLIEQIKEHPHKKHSSAESRRGLHRSWRRDEDAADLPKYTLPKHGINSKAAYQLLHDETALDGNPLL
NLASFVHTWMPEDADKLIMENINKNIVDMDEYPAASLIHNRCISMLADLWKAPKEGKVIGTATAGSSEAIMLGGLAL
KKRWQEARKAAGKDYFHPNIVFGSNAQVALEKFARYFDVETRLVPVKEENGFVMNPHDAIPYIDENTIGVIVILGST
YTGHFEDVKLMSDLLDDLEKRTGLDVKIHVDGASGGAFIAPFAYPHLKWSFDVPRVVSINTSGHKFGLVYAGLGWVLW
RDESFLHRDLVFELHYLGSTEYSFTLNFSKPAAPVIAQMFNFLNLGFEGYKKIAYKDMRNARMLSRALESTTYFKVF
SNIHVPRNSDSAHVSSNNKDDPETYHAGLPVVAFRLSDEFNQNYPNVRQVWIQTLLRTKGWIVPNYNAPLGAENIEI
LRIVVRETLSEDLIERLIVDIVAVTESLTTEEGNVFAGITASGAVVKPDLDEARPDSSNFNDSGDGEAQGQTGYSRQ
C

Seq. ID No: 309
>AADC_19
MALSKHVNTDKLIRDSRDKKSPKEKAHHTATAHQEATYSYGDRYVTNPVPKYNIASKGISADAAYRLIHDELALDGS
TVLNLASFVHTWMPPQGEQLVHENIAKNLIDSDEYPATQIIHTRCVSILADLWHAPSAKQAVGTATTGSSEAIQLGG
LAMKKMWQARMKAAGKNIHEPGPNIVMGANAQVALEKFARYFDVECRLVPVSVESKYRLDPKKAMDFVDENTIGIFI
ILGSTYTGHYEPVKEMSDLLDEYEKRTGIYVPIHVDGASGGFVAPFVHPKMVWDFKLPRVVSINTSGHKFGLSYVGV
GWVVWRDKAHLPKDLIFELHYLGSVEYSFSLNFSRPAAPILAQYFNLVHLGFEGYRSVGLADMKNARELSRALEKTG
YYTVLSDIHRAVGAKDPHGIDDADIEAYEPGLPVVAFRFSDNFKEKHPEIQQKWIQTLLRAKGWIVPNYELPPSLEQ
IEILRVVVRENVTEVLIDKLIDDIVEITEQLADSSSSMHSLNNLGHIQRPKKHEHPESNLKEGEGSDYSGTYARPC

Seq. ID No: 310
>AADC_20
MALNAVSAARGSARQYISTFLTLDNAKSGLFYYVLLVQAIKVKRHLRARGISASLKELYTWISQQIIRLLLRLPATR
KKVASQMDQAKLDIENRLVPKGANVTRHLSLPSEGKSLEWITQEMDKMDTELGGTSDAWRQGKLSGAVYHGGDELAK
IIVAAYSRYCVSNPLHPDVFPAVRKMEAEIVAMCLKMYRGPEGAAGAMTSGGTESIVMSVKTHRDWARSVKGIKEPE
MVVPVSAHAAFDKAAAYLGIKLHSIPVDSYTRQVNIKHVKRAINSNTIMIVGSCIGFPDGNQDDIEALGALAKKYNI
GLHVDCCLGSFIVPFLEPAGLAKGDNKGRYKLTPFDFTVDGVTAISCDTHKYGFAPKGTSVIMYRSAELRRFQYYVN
PIWPGGVYASPSLSGSRPGALIAGCWAVMQYMGTEGYLSSCRDIVIATRKIADAITDDIPELYVLGNPPASVVAFGS
RNPTVDPLEVGDGMRKRGWHLNGLSSPKSVHIACTRLTLPVVDQFIADLKDCVREAKVAPSGKGTMVSVYGLGNSSA
VGPDMVSQLASAFLDALYKA

Seq. ID No: 311
>AADC_21
MELKTAANEICNSNQMCQGENYSQNMLLRDGLIDLKNQIKEGDEGLGHGFSDFSNVFASDLLPARNGEKYTEGFLLE
VFNILFSYIRKTFDRKSKVLDFHHPHQLLEGLEGFNLELSDQPEPLEQILGDCRDTLKYGVKTAHPRYFNQLSSGLD
MVGLAGEWLTGAANTNMFTYEIAPVFIIMEGLLIKKMHELVGWGELEADGIFSPGGTISNLYSVLVARYKFFPIVKL
KGMAALPRIVLFTSEHSHYSFQKASATLGIGIENVIAVKCDERGKMIPSDLDEKIMAQKEK

Seq. ID No: 312
>AADC_22

MWKAKGTRGFELQIDSCLENAEYLYKKLKSRNGFELVFPDEPEHTNVCFWYIPPSLKGMPRDKEWNTKLHKVAAKIK
AQMMEEGTVMVSYQPLKNKPNFFRMVFSNPASKKSDIDFLLDEIERLGSDLEF

Seq. ID No: 313
>AADC_23
MFGSQHQMDVAALDRQLKEDKESGKLPLLLVANAGTPGAGHTDKLARLKELCHQYNIWLHVEGVNLATLALGYVSAS
VLAATKCDSMTLTLGPWLGLPAVPAVTLYRHEDPSLSLAAGLTTSQPVEKLRALPLWLSLQYLGHNGIVERIKHASQ
LSQRLLENLKDVTSIKTSVEPDGNSPVVVFKFFYDGPGSGSTINLNTIERESDAMNQWLGEQLAALIPSCAVDTVEL
EDEGVCVRFNPMMTSAVLGTTIEDVDQLVECVKVKIPIIHNTLQLKEEFRLEVERIAGLTYVVDYSWAGLGVLRYDH
VSEELDGSRREAELEKINASLLKKLNELESDLSFSSGPEFGAEKNCVYIGMATEDVDVSELVETIAVMGREIEENSK
LLENMTEVVRKGILEAEVQLQKANEERLLEEGVLRQIPLVGSVLNWLSPVQATPKGRTFNLTAGSLESTEITYASKA
QANGTSPPPTPSLGHAKRHPGQKLFKRLSRNSDAMSETSSVSHLEEVENLEASPTPEPQPGHPTEPPVPSVESNSEE
PHEAEALDTKTVESESLR

Seq. ID No: 314
>ADK1
MSSSESIRMVLIGPPGAGKGTQAPNLQERFHAAHLATGDMLRSQIAKGTQLGLEAKKIMDQGGLVSDDIMVNMIKDE
LTNNPACKNGFILDGFPRTIPQAEKLDQMLKEQGTPLEKAIELKVDDELLVARITGRLIHPASGRSYHKIFNPPKED
MKDDVTGEALVQRSDDNADALKKRLAAYHAQTEPIVDFYKKTGIWAGVDASQPPATVWADILNKLGKD

Seq. ID No: 315
>AOQS_1
MSNAAIRSSRAVSVSSSTKYYDFTVIGSGVAGLRYALEVAKQGTVAVITKDEPHESNTNYAQGGVSAVLCPLDSVES
HMRDTMVAGAHLCDEETVRVVCTEGPERIRELIAMGASFDHGEDGNLHLAREGGHSHCRIVHAADMTGREIERALLE
AVLNDPNISVFKHHFAIDLLTSQDGLNTVCHGVDTLNIKTNEVVRFISKVTLLASGGAGHIYPSTTNPLVATGDGMA
MAHRAQAVISNMEFVQFHPTALADEGLPIKLQTARENAFLITEAVRGDGGILYNLGMERFMPVYDERAELAPRDVVA
RSIDDQLKKRNEKYVLLDISHKPREKILAHFPNIASECLKHGLDITRQPIPVVPAAHYMCGGVRAGLQGETNVLGLF
VAGEVACTGLHGANRLASNSLLEALVFARRAVQPSTELMKRTRLDVCASEKWTRPVVATARLLGDEVIAKIIALTKE
VRRELQEVMWKYVGIVRSTIRLTTAERKIAELEAKWETFLFEHGWEQTVVALEACEMRNLFCCAKLVVSSALARHES
RGLHYMTDFPFVEESKRIPTIILPSSPTTASWSSRRLQNISSSSLIDCGSGEGRGSLLTCGDVEENPGPSSSSSSQT
TELVPYKLQRLVKEFKSLTEPIDRLKWVLHYASLLPQMPESSKTESNRVMGCTARVWLDAELGQDGKMRFCADSDSD
VSKGMCSCLIQVLDEASPVEVMELKTEDLAELNVGLLGGERSRVNTWYNVLVSMQKKTRRLVAEREGKVPSFEPFPS
LVLTAHGIEAKGSFAQAQAKYLFPEESRVEELVNVLKEKKIGVVAHFYMDPEVQGVLTAAQKHWPHISISDSLVMAD
SAVTMAKAGCQFITVLGVDFMSENVRAILDQAGFEKVGVYRMSDETIGCSLADAASAPAYLNYLEAASRSPPSLHVV
YINTSLETKAFAHELVPTITCTSSNVVQTILQAFAQMPELTVWYGPDSYMGANIVKLFQQMTLMTNEEIANIHPKHS
LDSIKSLLPRLHYFQEGTCIVHHLFGHEVVERIKYMYCDAFLTAHLEVPGEMFSLAMEAKKREMGVVGSTQNILDFI
KQKVQEAVDRNVDDHLQFVLGTESGMVTSIVAVIRSLLGSSANSKLKVEVVFPVSSDSMTKTSSDSSNSIKVGDVAL
PVVPGVAGGEGCSIHGGCASCPYMKMNSLSSLLKVCHKLPDLENVYGGFIAERFKRQTPQGKLIADVGCEPILHMRH
FQANKELPDKLVHQVLSCESKR

Seq. ID No: 316
>AOQS_2
MNTLPEHSCDVLIIGSGAAGLSLALRLADQHQVIVLSKGPVTEGSTFYAQGGIAAVFDETDSIDSHVEDTLIAGAGI
CDRHAVEFVASNARSCVQWLIDQGVLFDTHIQPNGEESYHLTREGGHSHRRILHAADATGREVETTLVSKALNHPNI
RVLERSNAVDLIVSDKIGLPGTRRVVGAWVWNRNKETVETCHAKAVVLATGGASKVYQYTTNPDISSGDGIAMAWRA
GCRVANLEFNQFHPTALYHPQARNFLLTEALRGEGAYLKRPDGTRFMPDFDERGELAPRDIVARAIDHEMKRLGADC
MFLDISHKPADFIRQHFPMIYEKLLGLGIDLTQEPVPIVPAAHYTCGGVMVDDHGRTDVEGLYAIGEVSYTGLHGAN
RMASNSLLECLVYGWSAAEDITRRMPYAHDISTLPPWDESRVENPDERVVIQHNWHELRLFMWDYVGIVRTTKRLER
ALRRITMLQQEIDEYYAHFRVSNNLLELRNLVQVAELIVRCAMMRKESRGLHFTLDYPELLTHSGPSILSPGNHYIN
RGSGEGRGSLLTCGDVEENPGPGSVMFDPDTAIYPFPPKPTPLSIDEKAYYREKIKRLLKERNAVMVAHYYTDPEIQ
QLAEETGGCISDSLEMARFGAKHPASTLLVAGVRFMGETAKILSPEKTILMPTLQAECSLDLGCPVEEFNAFCDAHP

DRTVVVYANTSAAVKARADWVVTSSIAVELIDHLDSLGEKIIWAPDKHLGRYVQKQTGGDILCWQGACIVHDEFKTQ
ALTRLQEEYPDAAILVHPESPQAIVDMADAVGSTSQLIAAAKTLPHQRLIVATDRGIFYKMQQAVPDKELLEAPTAG
EGATCRSCAHCPWMAMNGLQAIAEALEQEGSNHEVHVDERLRERALVPLNRMLDFAATLRG

Seq. ID No: 317
>ATMT_1
MTPAAGKTFNTSIAGADDLIRLHLSESPHGASKAALQAAERELARVNVYPDPERQELVRALAAHWGVGPEHIAVANG
SDELVLATALTLGDRNLPGLVTDGTFPGYRACLELLGRGCTAVPPDGTAVDVAGFAARLPGHGIGYLCNPHNPSGAA
LTRQELAALVEVSGRSGVPLVFDEAYMEFAGPDVPQTRDLTAAGDAPVVALRTFSKAYGLAALRVGYAVGRPDLIAG
LRGTLRALPFSVNRLAQAAAIAALGDPDFVDGVRRSTAERRRWFVGELDRRGRAHLPSVTNFVAVAARDCARAQDRL
AADFGILVRNAGLFGFPGYLRTSLGEKKDLERFLDALDEIEQNPGGGSGEGRGSLLTCGDVEENPGPMTAPLSRDGL
RAMGESVFRPAEWQGAAHTPLDADTAFNGFISTHVVFALEQLGLFAWFDESDRLDVPQYCWRRKLDERVFRQLVSAA
EAFGYLDVHDDLVTPTPAWSELRRKIGFFTWGVGGYHDVFANAASIARGERAFGKDVLRDEAMVALGSAQADMALMR
DLLDEQIAALDFSVIADLGSGISERVCRLVKSRPGARGLGVDISASATALAAGTVERHELADRVQPICADVLDVLFH
GRRIEGADQVDVAMSFMFLHDLLVDPTTRTDVIPALRKAFPRAHTFLLADTTVRPRDEKDTLPVFSSGFELAHALMG
VPIYTREEYENLFHEGGLHLRRTVPFGAPHTYLFVLEAQ

Seq. ID No: 318
>ATMT_2
MQALPVKGDTVSRPPTVHSLHHEHERADGMLRLHCNENPYGPPSGVIASVTKELEGRCSTYPDSEVTALREALAGQV
GVGTDMVAVGNGADELVLLITLASAGPGDTVVVTESTFPGYAASAAVAGATVRGVPLHRDRVSATALVEAVDDGARL
VFVCNPHNPTGTVLSPAAVEEILRACERTGAVPVFDEAYIEFAGPGFDHALDAVRAGRRLLVLRTFSKAWGLAALRA
GYAVGPADLVAGIMEARRPLPFSVNRLAQQAALAALGSPDHIAEVYERTTRERERLCRALTGLGVAYVPSVTNFVMV
KTPGNSTRFASRLADEHGILVRDLAPFGYPGHVRVSVGTAEDTDQFCAALGSLLASPRSHAATGHGLGASSGAGGAG
NAAIRSARDVLPVPTLDPVAPQDLFNGYVGAHAVFALTRLGVWDRLAEGSEPTVDALAVQAGTDATGLMPLLRVAAL
LGYVSLTDGSAPAVRLTESGRELVRMRGFFTWGVGGYHEVLRSLPALARGTSVFEQDVDRDGGMVAVGSGEVGREMM
LPLEQEVLATVDFRTVADLGCGDATRLLRLCDGHPHRRGTGIEINQGACVQANKRVADAGLADRVDIVHGDALDLSG
RTFPEVDLVTSFLMMHDLFDATGDPVGVMRTLREVFPRARHFLIGDTVAQDWEERREGLPMFSVGFELVHAFMDTPI
MNRGTYEDAFAGAGLRVARREPLGAPSTWLWLLSTE

Seq. ID No: 319
>ATMT_3
MRRRWAVTASASWWGTACELHASASAAYTPPCHSPGTGGRGTESGPMTAPVRQETRNYNASVPSADDLVRLHLSESP
YGASPAAVAAVTGELERINRYPAPGREGLVQALARHWELPEEHIAVANGSDELVLATALTLGDPGSPGLVTAGTFPG
YLAALERIGRGAVQVPLAGSGTDTAAFADRLPGCGIGYVCNPHNPCGSALTHDELHRLVAAARDSGTPLVFDEAYHE
FGPPAQPQARTHLREDTPVLALRTFSKAYGLAALRIGYALGPADLIAEVRRTLTVLPFSVNRAAQAAALAALDDQEF
LGSVRRDSAARRQWFCAELERRGYRYLPSVTNFVAVEVAASAEAQDVLARDHGILVRDTGMFGFPGHLRVSLGSVEE
LRGFLDALDRVTAGSRGGGSGEGRGSLLTCGDVEENPGPMTGPVSTSAPSRWPRTWRPNRLEPTSRGGQPGHAARRS
PAAGRRRRRASEARPPPSGRQPAVRTERCERVSPLNTLPSEWQGQAPTPLNPDTAFNGYICANVLHGLERLGVFELL
RDEKSLDMDRFCETNGLDSAVFRALVGAAESFGYLDVRGAQVRATSVGEDVARYLGFFTWGVGGYHDIFASAAPVAR
GERRFGVDLHRDEGMVALGSAQADTALMRHILDEEIAGIDFRTLVDLGAGVSERVSRLVKARPGTRGIGIDISRPAT
ELARDTVAGYGLAGTVEPVCADVLDILFNGQEIDGGDAADVVMSFMFLHDLLAAPERREEVVPRLRKAFPRAHTFLL
ADTTIRPRNEEGDGRLPVFSSGFELAHALMGVPLHTREEYEELFERGGMKLRRSVPFGAPHTYLFVLEAS

Seq. ID No: 320
>ATMT_4
MTNDPSPRDARDELPVRDELRGQSPYGAPQLDVPVRLNTNENPYPLPEALVERIAERVREAARSLNRYPDRDAVELR
TELARYLTRTAGHEVTAAHVWAANGSNEVLQQLLQTFGGPGRTAIGFEPSYSMHALISRSTGTGWISGPRNDDFTID
VDAARAAIAEHRPEVVFITSPNNPTGTAVRAETVLALYEAAQAARPSIVVVDEAYGEFSHHPSLLPLIEGRRHLVLS
RTMSKAFGAAGLRLGYLAADPAVVDAVQLVRLPYHLSSVTQATALAALEHTDTLLGYVAQLKGERDRLVAELRAIGY
EVTESDANFVQFGRFDDSHAVWRQILDRGVLVRDNGVPGWLRVTAGTPEENDAFLDAVRELKKEHDAGGGSGEGRGS

LLTCGDVEENPGPSSSSSTRTDFAQSAVASIFTGAIASHAAVLADDLGLFDALAKGKLRNRDLDRSPWLRNRIRISG
ALEALCRVGAVQRCTDGYELTDVGTELAGQVPVFRLWLGGYASVLAGQISIGADPATGVHGGIVAESSGAIGARYLD
ETIVNLLESLRPEGRICDIGCGTGARLLRVCRRVNQPGIGYDLSAKAVEAARETVDEARRIGVDIDVRQGDATALTQ
DHPDVDIVTQAFMTHHIAPDEYCAAVLRSYRSRFPRARYLVIFDTVPSQDSEEPEIFAPGFDYIHALQNMEPRSRGA
ARRMFTEAGYICREEVELAVPNSYAWVLEMRDREGPAS

Seq. ID No: 321
>ATMT_5
MTNDPSPRDARDELPVRDELRGQSPYGAPQLDVPVRLNTNENPYPLPEALVERIAERVREAARSLNRYPDRDAVELR
TELARYLTRTAGHEVTAAHVWAANGSNEVLQQLLQTFGGPGRTAIGFEPSYSMHALISRSTGTGWISGPRNDDFTID
VDAARAAIAEHRPEVVFITSPNNPTGTAVRAETVLALYEAAQAARPSIVVVDEAYGEFSHHPSLLPLIEGRRHLVLS
RTMSKAFGAAGLRLGYLAADPAVVDAVQLVRLPYHLSSVTQATALAALEHTDTLLGYVAQLKGERDRLVAELRAIGY
EVTESDANFVQFGRFDDSHAVWRQILDRGVLVRDNGVPGWLRVTAGTPEENDAFLDAVRELKKEHDAGGGSGEGRGS
LLTCGDVEENPGPAQAAPTTVTEVFNHAITASAISAAWEMGAFDALRVSERLDADEFAAREGLDTRSTHELFRALAA
ADIVSRDGAQIRRGPNFAEADRCKSLFHWMTRGCGELFSTLPALVREKNRVGSFYRRDAAAISVACREINAEWWDPV
FWPVVSGLDFTSVADLGCGSGERLIRLARTGPEVMALGIDFAAGAIEVATAAVAEAGLSDRISLVQGDATALEPRPE
FAGVDLLTCFMMGHDFWPRAEAVASLRRIREVFPDLKHFLLADATRTTSYPDTDMPVFSMAFELAHAVMGDYLPTLE
EWRPVFEEAGWRCEGEHPISVPADSVMFHLVPN

Seq. ID No: 322
>ATMT_6
MTNDPSPRDARDELPVRDELRGQSPYGAPQLDVPVRLNTNENPYPLPEALVERIAERVREAARSLNRYPDRDAVELR
TELARYLTRTAGHEVTAAHVWAANGSNEVLQQLLQTFGGPGRTAIGFEPSYSMHALISRSTGTGWISGPRNDDFTID
VDAARAAIAEHRPEVVFITSPNNPTGTAVRAETVLALYEAAQAARPSIVVVDEAYGEFSHHPSLLPLIEGRRHLVLS
RTMSKAFGAAGLRLGYLAADPAVVDAVQLVRLPYHLSSVTQATALAALEHTDTLLGYVAQLKGERDRLVAELRAIGY
EVTESDANFVQFGRFDDSHAVWRQILDRGVLVRDNGVPGWLRVTAGTPEENDAFLDAVRELKKEHDAGGGSGEGRGS
LLTCGDVEENPGPSTEVSEAQARRAVADIFNSTLASSAIGAAWELGALDELRENGKLDVSDFAVRHDLHEPAVVGMF
TALASVGIVRREGATVVVGPYFDEANHHRSLFHWLNQGSGELFRRMPQVLPNENRTGKFYQRDAGAISYACREISER
YFDPAFWAAVDGLGYTPTTVADLGSGSGERLIQIARRFPGVRGLGVDIADGAIAMAEKEVAAKGFGDQISFVRGDAR
TIDQVSARGEFAEVDLLTCFMMGHDFWPRENCVQTLRKLRAAFPNVRRFLLGDATRTVGIPDRELPVFTLGFEFGHD
MMGVYLPTLDEWDGVFEEGGWRCVKKHAIDSLSVSVVFELE

Seq. ID No: 323
>BH4reg_1
MAASGEARRVLVYGGRGALGSRCVQAFRARNWWVASIDVVENEEASASVIVKMTDSFTEQADQVTAEVGKLLGDQKV
DAILCVAGGWAGGNAKSKSLFKNCDLMWKQSIWTSTISSHLATKHLKEGGLLTLAGAKAALDGTPGMIGYGMAKGAV
HQLCQSLAGKNSGMPSGAAAIAVLPVTLDTPMNRKSMPEADFSSWTPLEFLVETFHDWITGNKRPNSGSLIQVVTTD
GKTELTPAYF

Seq. ID No: 324
>BH4reg_2
MTALTQAHCEACRADAPHVSDEELPVLLRQIPDWNIEVRDGIMQLEKVYLFKNFKHALAFTNAVGEISEAEGHHPGL
LTEWGKVTVTWWSHSIKGLHRNDFIMAARTDEVAKTAEGRK

Seq. ID No: 325
>BH4syn_1
MEGGRLGCAVCVLTGASRGFGRALAPQLAGLLSPGSVLLLSARSDSMLRQLKEELCTQQPGLQVVLAAADLGTESGV
QQLLSAVRELPRPERLQRLLLINNAGTLGDVSKGFLNINDLAEVNNYWALNLTSMLCLTTGTLNAFSNSPGLSKTVV
NISSLCALQPFKGWGLYCAGKAARDMLYQVLAVEEPSVRVLSYAPGPLDTNMQQLARETSMDPELRSRLQKLNSEGE
LVDCGTSAQKLLSLLQRDTFQSGAHVDFYDI

Seq. ID No: 326
>BH4syn_2
MHSPSLSAEENLKVFGKCNNPNGHGHNYKVVVTIHGEEAIMKPLDHKNLDLDVPYFADVVSTTENVAVYIWENLQRL
LPVGALYKVKVYETDNNIVVYKGE

Seq. ID No: 327
>BH4syn_3
MPSLSKEAALVHEALVARGLETPLRPPVHEMDNETRKSLIAGHMTEIMQLLNLDLADDSLMETPHRIAKMYVDEIFS
GLDYANFPKITLIENKMKVDEMVTVRDITLTSTCEHHFVTIDGKATVAYIPKDSVIGLSKINRIVQFFAQRPQVQER
LTQQILIALQTLLGTNNVAVSIDAVHYCVKARGIRDATSATTTTSLGGLFKSSQNTRHEFLRAVRHHN

Seq. ID No: 328
>BH4syn_4
MHHHHHHTSSTPVRTAYVTRIEHFSAAHRLNSVHLSPAENVKLFGKCNHTSGHGHNYKVEVTIKGQINPQSGMVINI
TDLKKTLQVAVMDPCDHRNLDIDVPYFESRPSTTENLAVFLWENIKSHLPPSDAYDLYEIKLHETDKNVVVYRGE

Seq. ID No: 329
>BH4syn_5
MHHHHHHSSKEHHLVIINGVNRGFGHSVALDYIRHSGAHAVSFVLVGRTQHSLEQVLTELHEAASHAGVVFKGVVVS
EVDLAHLNSLDSNLARIQSAAADLRDEAAQSTRTITKSVLFNNAGSLGDLSKTVKEFTWQEARSYLDFNVVSLVGLC
SMFLKDTLEAFPKEQYPDHRTVVVSISSLLAVQAFPNWGLYAAGKAARDRLLGVIALEEAANNVKTLNYAPGPLDNE
MQADVRRTLGDKEQLKIYDDMHKSGSLVKMEDSSRKLIHLLKADTFTSGGHIDFYDE

Seq. ID No: 330
>DAC_1
MVDADIALNWAGGLHVCIVRPPGHHAEPGAACGFCFFNNVALAARYAQSLQSPSDPPLRVMILDWDIHHGNGTQHIF
QDDASVLYVSLHRYDDGTFFPSSEDAAHDKVGSGPGEGFNVNIPWNGGKMGDVEYLLAFHRIVMPIAYEFNPQLVLV
SAGFDAARGDPLGGCRVSPEGYAHMTHLLMGLAGGKVVVVLEGGYNLTSISESMSMCTRTLLGDPLPFISDLHAPRP
AALRAISSVLGVHQKYWRSLCINVGPP

Seq. ID No: 331
>DAC_2
MKTHPHPERPDRLQAIAASLATAGIFPGRCYPIPAREITKEELQMVHSLEHIETVELTGQILYSYFTPDGTNPHNRL
KLDNRKLAGILSQRMFVILPCGGLGVDSDTIWNDLHSSNAARWAAGSVIDLAFKVVTRELKNGFALVRPPGHHADPS
TAMGFCFFNSVAIAAKQLQQKLNVRKILIVDWDVHHGNGTQRVFYRDPNVLYISLHRHDDGNFFPGSGAADEVGANS
GEGFNVNVAWAGGLDPPMGDAEYLAAFRTVVMPIAHEFAPDVVLVSAGFDAAEG

Seq. ID No: 332
>DAC_3
MMATEPIASGSGTMDIDSEKTPSTSQANPMADTFQTREAVLGLGEVVEHVGGRWVAEQEWIRSPERKMAYTQGTKKK
VCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKANAEEMTKYHSDDYIKFLRSIRPDNMSEY
SKQMQRFNVGEDCPVFDGLFEFCQLSTGGSVASSVKLNKQQTDIAVNWAGGLHHAKKSEASGFCYVNDIVLAILELL
KYHQRVLYIDIDIHHGDGVEEAFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNYPLRDGIDDES

Seq. ID No: 333
>DAC_4
MDAGTRRVDDAAVPSTGPSASLLRSANMLSAAFGLTASLYSRLRGVCSSRRALSTSARTSEAAGVGAKPGVAAALTV
PSTGPSASEASPAALLRIQVAEEWARASGLLDREDCQVGLAFDEAMHLHSGPAGHPERPARTKEILAQLHASGLVRA
CAQVPSREATEEELLLVHDARHVERVLRHEAAGHKKAKAFSFPFGPDTYVCEHTARCARLAVGCLLSLVDASLDPAS
PVRTGMAVVRPPGHHATSDRASGFCLFNNVAVAARHLQRRHGLKRVAIVDWDVHHGNGTNDLFTEDPNILFFSVHRF
DNHGFFPGSGFLEDVGHAQARGYTVNVPLEKGYGDLDIVHVVKYVLCPVLERFKPDAILVSAGFDAVKGDPLGECRV

SPEAFGWMTRCLHRLAQRYCDGRLFLVLEGGYNPDMIAQCCIECVQSLVAEAAGLRGPWPEFPAVGVPLAEGAQLSA
PSSAPTSAPGTPTSTSPASSPALSAAAPPLASPGSTPTSSPCLRPSGGEAPPRSPPSASASAGGGARQRARAPSSKT
VRAVRQLTEIHHLLPLELPVAPRPGDGPGAANKSARKNERRRLGRGRRGPEEEGASSDSSGWAIACGLSDAEPWPSP
QASPVASLSQGASSLPTLELPPAFPSLDGVGSTAGNSYLGTSGNVGIDAAGHSASSWLGSPTTAATAVAPPARGDRK
VKRR

Seq. ID No: 334
>DAC_5
MVDADIALNWAGGLHGHHAGRGCSEGFCLLNNVAVAAAYARSAHPEQVRRVLVLDWDVHHGQGTQEIFWRDPGVLYA
SVHRDGGEGFYPGTGAAEQVGDGAGRGFTVNVPLPTGYGDGCLWAACAEVLLPAARRFRPDLILVSAGFDAVAGDPL
GGCRCTARGFGALTGELRKLAGELCSGRLLLALEGGYDLRTLRACVGEVCQALAAPEPAEGGA

Seq. ID No: 335
>DAC_6
MRNRSSGFCLVNNVAVAAEYARDRYPEVERVLIFDWDVHHGQGTQQIFEQSPDVLVISVHRHDGHSFYPATGSAGEV
GSGPGRGYSVNVALPAGYGGAALWTACAHVLLPAARNFQPQLILVSAGFDAAASDPLGGCFVEPRVFGALTAELRRL
AAEVAEGRLILALEGGYNPEVLADCVDEVAAALVADASSSGVEAFAEAPSWLAGSACFGAIRRTCEAHRMAPLRLPL
PSSRIDRRRAAARQAEALSSPSSRDAGDTGGGEVSAHGATTTVTTSANLGAGTLAARPSSMVTGEGRRANGQLVDVL
GIALAGKPSASPWPEAQRTQGSAPGTPAPATGGALPPAETAESPGSVASGAAVAQGPVECQAAARQAGECPGQAPAP
AGAGAAPGGRGVEAAAAQHGQDLAPAAGPGAAALVELQTGELVVRIAPLPRPKDVVVSAEELWVWHDQGGPLGVQRW
RFEGVRAENSGALRCAEFRSKRHELTVRLRLG

Seq. ID No: 336
>DAC_7
MVRSSQATTFSSSPYFADRGIHGTAQPITERINPRKCTFHILPPGLGGWLYLFPKMSKTVAYFYDPDVGNFHYGTGH
PMKPHRLALTHSLVLHYGLYKKMIVFKPYQASQHDMCRFHSEDYIDFLQRVSPNNMQGFTKSLNAFNVGDDCPVFPG
LFEFCSRYTGASLQGATQLNNKICDIAINWAGGLHHAKKFEASGFCYVNDIVIGILELLKYHPRVLYIDIDIHHGDG
VQEAFYLTDRVMTVSFHKYGNYFFPGTGDMYEVGAESGRYYCLNVPLRDGIDDQSYRHLFQPVIKQVVDFYQPTCIV
LQCGADSLGCDRLGCFNLSIRGHGDCVQYVKSFNIPLLVLGGGGYTVRNVARCWTYETSLLVDETISEELPYSEYFE
YFAPDFTLHPDVSTRIENQNTRQYLDQIRQTIFENLKMLNHAPSVQIHDVPSDLLSYDRTDEPDPEERGAEDNYTRP
EASNEFYDGDHDNDKESDVEI

Seq. ID No: 337
>DAC_8
MWDVHHGQGIQYIFEDDPSVLYFSWHRYEHGSFWPNLSESDYDSIGKGRGTGFNINLPWNQTGMGNADYVAAFFHVL
LPLAFEFNPELVLVSAGYDSGIGDPEGHMRATPECFSHLTHMLMHLAGGKLCMILEGGYHLRSLSESVSMTVRTLLR
DPVPRLSGEMTPCYSALESIQNTRHAHSPYWKCLLHDETRLVEEISTKGLKAPGPLHVDASVVDEFLENHMKKILHP
TPPITTMVVASVENTLNLPAGVQLEESTVTPEQARHAISVFNPDELNENVLNSVSKMLPALEKLVN

Seq. ID No: 338
>DAC_9
MKTHPHPERPDRLQAIAASLATAGIFPGRCYPIPAREITKEELQMVHSLEHIETVELTGQILYSYFTPDTYANQHSA
HAARLAAGLCADLAKEVFSGRAKNGFALVRPPGHHAGVRQAMGFCLHNNAAVAALAAQVAGAKKILIVDWDVHHGNG
TQEIFEQNKSVLYISLHRHEGGKFYPGTGAAHEFGTMGAEGYCVNIPWSRGGVGDNDYIFAFQHVVLPIASDFAPDF
TIISAGFDAARGDPLGCCDVTPAGYAQMTHMLNILSGGKLLVILEGGYNLRSISSSATAVIKVLLGESPGCNPKNFL
PSKAGVQTVLEVLKIQMNFWPALGSIYSDLQTQWGMYCMKTKKKQIKKRQRAAAPLWWKWGQKSFLYHLLNGHLHVK
SKGC

Seq. ID No: 339
>DAC_10

MFLVRHHLYKWKSHFRIDADGKFVEDQFFPKNLKSGRRFLRSIGASITCSNGIGKDPYILSNEKISDARLIYAVAPA
MGHNQESHPESHFRVPAIVNALEKMEMTPKFRGSEIIELQSFKPALVDDIASVHARAYVSGLEKAMDQASQQGIIFI
DGSGPTYATATTFHESLVAAGAGIALVDSVVAASKNHLDPPVGFALIRPPGHHAIPLGPMGFCVFGNVAIAARYAQR
AHGLKRVFIIDFDVHHGNGTNDAFYDDPDIFFLSTHQDGSYPGTGKVDEVGRGDGEGTTLNLPLPGGSGDIAMTTVF
DEVIAPCAQRFKPDIILVSAGYDGHVLDPLASLQFTTGTYYMLASKIKLLAKDLCGGRCVFFLEGGYNLESLSYSVA
DSFRAFLGEQSLASEFDNPAILYEEPSTKVKQAIQRVKHIHSL

Seq. ID No: 340
>DAC_11
MVDADIALNWAGGLHHAKKSEASGFCYVNDIVLGILELLKVHRRVLYVDIDVHHGDGVEEAFYATDRVMTVSFHKFG
DFFPGTGHIKDTGWGPGKNYALNVPLNDGMDDESFRGLFRPIIQGVMEVYQPDAVVLQCGADSLSGDRLGCFNLSVK
GHADCLRFLRSFNVPLMVLGGGGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYYEYFGPDYTLHIEPCNMENLNTP
KDMEKIRNMLLEQLSRIPHVPSVPFQTTPPTTQVPEEAEEDMDRRPKCRIWNGEDYDSDPDEDEKPRHTEPNSELRD
VVDEMDEDKREEHPPS

Seq. ID No: 341
>DAC_12
MCSDANGKVGNISVMSTEGISQVESKKARLNGLLTLEDIYNLPDELDDDEDDSDWEPLLEPLAVRKWFCTNCTMVNF
DGFDFCETCEEHKESGILKQGFFASPALQGTRSTQIESEVIERYTESICDISASALSTVVGFDERMLLHSEVVLKPH
PHPHPERPDRLRAIAASLSTAGIFPGKCHPIAAREITQEELLKVHSLEHVEAVEVTRQMLSSYFTPDTYANEHSAQA
ARLAAGLCADLASEIYSGRAKNGFALIRPPGHHAGVHQSMGFCLHNNAAVAALAAQVAGAKKVLIVDWDVHHGNGTQ
EIFERNKSVLYVSLHRHEAGKFYPGTGAAHEVGTMGAEGYCVNVPWSRGRVGDNDYIFAFQNVVIPIAHEFSPDFII
ISAGFDAARGDPLGGCDVTPAGYACMTHMLSALAGGKMLVILEGGYNLRSISSSATAVIKVLLGEKPKCQFENIEPS
ASGLQALLEVLKVQTNFWPCLSSKLTQLQSCWEAYLSGRKKQKKRRFRTVAPPPIWWAWGRKRFLYFLRCQRFRMKP

Seq. ID No: 342
>DAC_13
MAGAEELHVFWEEGMLKHETGRGVFDTGSDPGFLDVLEKHPENADRVRNMVSILKRGPIAPFVSWHQGRPASLPELL
SFHSSEYIEELEEADRAGGKMMCCGTFLNPGSWNAALLAAGTTLSAVKYILDGHGKIAYALVRPPGHHAQPTQADGY
CFLNNAGLAVQLALDEGCRKVAVIDIDVHYGNGTAEGFYCSNKVLTISLHMNHGSWGPSHRQSGTHDELGDGDGFGY
NMNIPLPNGSGDRAYEYAMQELVVPAVQKFGPDMIVLVVGQDSSAFDPNGRQCLTMDGYRQVARIVRGLADMHCKGK
LLVVQEGGYHITYAAYCLHATLEGALNLPSPLLSDPIAYYPEDEGFAVKVIDAMKEHYKSNVPFLKEIN

Seq. ID No: 343
>DAC_14
MGFCIFGNIAIAARYAQRVHGLKRVFIIDFDVHHGNGTQDVFYEDPDIFFLSTHKEGSYPGTGKIHEVGCGPGEGTT
LNLPLPGGTGDVAMRTVFDEVIVPCAQRFKPDIILVSAGYDAHFLDPLANFQFKTATYYTLAANIKQLAKELCGGRC
VFFLEGGYNLKSLSYSVADSFRAFLGEPSCASDVDPTFLYDEPSTKIEQAIDKVKAIHSL

Seq. ID No: 344
>DAC_15
MEQLWVPSLPILGGRILPMLRHYCGFGSHHPLTWRSLQITGRKQKHNGCWIAYCLPSHNGTSISDTNGVRKDLALPD
NLLRDAHILYCTSPAMGHNKEAHPETNKRVPAIVDALEKLELTSKHRGSQVLEIQDFQPASLDDIALVHSRSYITGL
EKAMSRASDEGLIFIEGTGPTYATQTTFQECLLSAGAGITLVDSVVAASKLGPKPPLGFALVRPPGHHAVPEGPMGF
CVFGNIAVAARYAQNQHGLKRVMIIDFDVHHGNGTCDAFYEDPDIFFLSTHQLGSYPGTGKIHQVGQGNGEGTTLNL
PLPGGSGDYAMRCAFDEVIAPAAQRFKPDIILVSAGYDAHALDPLAGLQFTTGTFYMLAARIREVAAELCGGRCVFF
LEGGYNLESLSSSVADTFRAFLGEPSLAARFDDPAMLYEEPTRKIREAIDKAKHLHSL

Seq. ID No: 345
>DAC_16

MMATEPIASGSGTMDIDSEKTPSTSQANPMADTFQTRRPRASSLPLQPSNLKVGYIYSSEMMNHFCPGGHPEQPLRI
QQIWATIVNEQLHKRMKWMPIREVKKGEALLVHSEDHWNKVIAIQYLTDQQRADSVDYYEQMSLYVMSGTTRSALLS
CGGVVEACLAVARNELKKTFAIVRPPGHHAEPDEHMGFCFFNNVAVAARVVQQRTKLKKILILDWDVHHGNGTQRAF
NDDPSVLYISLHRYEQGTFYPCGPFGSLTSCGEGPGTGFSVNVPWPCAGMGDAEYIYAFQKVILPIATEFAPELVII
SAGFDAAAGDELGECLVSPAGYAHMTHMLAGLAGGRMVVALEGGYNLDSISQSALAVTKVLLGEPPDELPPLKANEE
GTETVWLVAREQSKYWKSVDPKACEPQADVEPISFSVPEILKAHRQHYLYTKHDMMQVPMMTPELEEKFSSQIMCTS
DIFESKTLVIFVHEFGNLRLELESSTTCDVHLERSYLIDFSKELVGWVKSEGYSLLDANLYPKPSTTPTPNLRHKTM
EEVGRDVLVYLWDNYVQLSGAERVILIGHGPGCKPLVDLLNRRTTSVTKSAKAIIQVVGSQRMPSYPSDVDDARPWY
QKSSLVIVPQSHPVMGPHIKPKDIRRHGVMVPIDETRQIKLITRALPAIKQFVQETLSSFPLANRTNRP

Seq. ID No: 346
>DAC_17
MSKRKVAYFYDPDVGAYTYGWSHLMKPHRMRITHELATAYGMLDKMHVLRPKRATPEAMTAFHTDEYVQFLHSVTPE
TADKLTGQKTRFLVGDDNPAFEGVFEFCSISAGGSIGAAERIASGAADIAINWAGGLHHAKKREAAGFCYINDIVLG
ILELLRTYPRVLYIDIDCHHGDGVEEAFYTTDRVMTCSFHKFGEFFPGTGTQEDTGTGKGKGYSVNVPLKDGIQDES
FKSVFDPVISKILEVFQPSAVVLQCGADSLAGDKLGCLNLTMQGHAHCVQFLRKSNIPLILLGGGGYTVKNVARAWT
YETACAIGIENEIDLNMPWSQYFEWFGPTYRLEVPENNMEDMNVKEGTLDHVRTTALAQLQQLASRCAPSVQMQDVP
RTSLGGHLGFKRDKREHRDELDERLAQHTRYLYDLQESESESEDTESSDSDASSVSFVNNWRRAPHRANSLPRILSG
RHSSNPPGHISASERRRMSIVTGKYFDIPIHESGYNHYEYGAAPTKSSKRIFFQSGLDIYNDDNDFEGIINARTSVS
NGFGNGIHDLHGLMERGGRSLNENLEDGDDEVEGEEYEDDAAMSDS

Seq. ID No: 347
>DAC_18
MEEHFWDVLYKDKYSKLLSKARDFLDDTGGPGDDVLVFISCGMDACEHEYESMSRHNRKVPASFYHRFARDACAFSD
RYAGGRLISVLEGGYSDRALISGAMAHLSGLVDTPDGIQVDEQWWNIPNLVKLEAATKKRRGGRPSLPAKGSVEPWI
ERTLSIFSSIDGSASTTSSRSTFIPPSSRTLRDRTKGREAMPKSPPASSASTKPVSRSKVKPGANIKSGDESFASTG
SSPLTSPSPSSSEDEAPPIKRL

Seq. ID No: 348
>DMAT_1
MTIINSRIIDIRQSTFEESIPDQVTAGLSTTPKTLPALLFYSGEGIRHWIEHSTAADFYPRHEELRILRARAAEMVD
SIANNSVVVDLGSASLDKVLPLLEALEASKKNITFYALDLSFSELQSTLQSLPYEQFKFVKIGALHGTFEDGVQWLK
DTPGVQDRPHCLLLFGLTVGNYSRPNAAKFLQNIASNALAASPVQSSILLSLDSCKMPTKVLRAYTAEGVVPFALAS
LDYGNTLFAPNKMGEKVFQPSDWYFLSEWNYMLGRHEASLITKGKEVRLGGPLNDIVIEKHEKIRFGCSYKYDTDER
QVLFGSAGLTDVKEWSVEGCDVSFYQLQMCPN

Seq. ID No: 349
>DMAT_2
MTISAPPIIDIRQAGLESSIPDQVVEGLTKEVKTLPALLFYSTKGIQHWNRHSHAADFYPRHEELCILKAEASKMAA
SIAQDSLVIDMGSASMDKVILLLEALEEQKKSITYCALDLSYSELASNFQAIPVDRFHYVRFAALHGTFDDGLHWLQ
NAPDIRNRPRCILLFGLTIGNFSRDNAASFLRNIAQSALSTSPTQSSIIVSLDSCKLPTKILRAYTADGVVPFALAS
LSYANSLFHPKGDRKIFNEEDWYFHSEWNHALGRHEASLITQSKDIQLGAPLETVIVRRDEKIRFGCSYKYDKAERD
QLFHSAGLEDAAVWTAPDCDVAFYQLRLRLN

Seq. ID No: 350
>DMAT_3
MSKPNVLDIRLATFEDSIVDLVINGLRKQPKTLPALLFYANEGLKHWNHHSHQPEFYPRHQEVQILKKKAQEMAASI
PMNSVVVDLGSASLDKVIHLLEALEVQKKNISYYALDVSASQLESTLAAIPTQNFRHVRFAGLHGTFDDGLHWLKEA
PEARDVPHTVLLFGLTIGNFSRPNAAAFLSNIGQHAFQGKSGDQCSILMSLDSCKVPTQVLRAYTCEGVVPFALQSL
TYANGLFSEKNKTQASGDVQHKVFNLDEWYYLSEWNFVLGRHEASLIPRSKDIKLLPPLDGILVSKDEKVRFGCSYK
YDQEERMELFAAAGVKNEVTWSDEGCDVAFYQLKLS

Seq. ID No: 351
>DMAT_4
MGSINPPQILDIRRSKFEESIPKQVEAGLLSSPKTLPALLFYSTEGIQHWNRYSHASDFYPRHEEIQILKDKATDMA
ASIADGSVVVDLGSASLDKVIHLLEALEAAQKKVTYYALDLSFSELTSTLQAIPTDQFVHVQFSALHGTFDDGLQWL
KETLVIRDQPHCLLLFGLTIGNFSRPNAAKFLHNIASHALVESPSQSSILLTLDSCKVPTKVIRAYTAEGVVPFALE
SLKYGNTLFQQDAGENVFDPEDWYFLSEWNYVLGRHEASLVPRSKDIKLGRPLDKIVVGKHEKVRFGCSYKYDSEER
KELFGTAGLRDVKSWSKEGCDVAFYQLKCCPN

Seq. ID No: 352
>DMAT_5
MPALPVIDIRSNHVEDSLPEQIIKGLTSQPKTLPPLLFYSNEGLEHWNHHSRQPDFYPRRQEIEILKQGGNDIARSI
APSSVILDLGSANLEKVGYLLEALEAQEKDVLYFALDISAPQLATTLKEIPSSNFRHVRFAGLHGTFEDGLRWINET
PEIRDLPHCVLLLGLTIGNFSRQNAAAFLQNIANHALTGASKNKSSILLSLDSCKVPTKVTRAYTSDGVVPFALQAL
TYAKALLCDRIDNGIDEKVLSCNLRPEHWHYLSEWNFALGRHEASLIPRFGDVCLGSMLQDIIVKKEEKVRFACSYK
YDAKERQKLFLDSGVDQGMVWTNEGCDVAIYELKLA

Seq. ID No: 353
>DMAT_6
MLYKPKVLDIRSGSVEDSLRHSVMDGIREDPRTLPTLILYGPEGLQHWDDHSHAPDYYLRHEELHILRSRAYEMAET
IADNTAMVDLGSAQVSRFHESPCLLAPTLSLDKAALLLDALEVQAKNVTYYALDLDHAELQKTLCRLPLGKYKHVQC
VGLQGTFEDGLEWIKNDPEQSRRPHCLLFLGSTIGNFSRENAARFIRSMASSAFLSESAKSSIILSIDSCKLPTKVL
RAYNSEGVVPFAMAGLKHASAILCEAACRQEDAVTETFLPDDWYYLSHYNHVLGRHEASFTPRNRDIQLGSPLEDVV
IRLGETIRFGYSHKYDFAEIEQLFREAGVAAVNSWGAVGCDLSFYQLGTA

Seq. ID No: 354
>DMAT_7
MAAPSVIDIRSHLVEDSLPDQVVKGLGSDPKTLPALLFYSNEGLEYWNHHARQPDFYPRHQEIEILKRKGDEIARSV
APNSVILDLGSANLEKVTYLLEALEAQAKNVTYFALDLSAPQLMSTLKAIPTTKFRHVRFAGLHGTFVDGLRWISET
PDIRDLPHCVLLFGLTIGNFSRPNAATFLRNIASQALRGASEDKSSIFLSLDSCKVPTQILRAYTSDGVVPFALQSL
AYAKTLFCEQTQNDFNEKPSSCHLNPDDWHYHSEWNFVLGRHEASLIPRLNDIHLGPLLHDIVVKKDEKVRFGCSYK
YDDLERDKLFVDAGVKDEMAWTNEGCDIAIYELKSM

Seq. ID No: 355
>FEX1
MIFNPVISNHKLSHYIHVFCTFTTFCILGTETRQAITALSTYTPAFVTAPTVLWSNCSSCMLMGIMQSLNAYTWMKD
HQVLFLGVTTGYCGALSSFSSMLLEMFEHSTNLTNGNIANHTKLPNRAYGIMEFLSVLLVHLMVSMGSLIFGRQLGK
EVIVAYGSSSFSKPYTPPSDTVKENAGDVDTQEMEKNILEFKFKTPAPFFKKFFDIVDKLAYALAFPLIILFVVLCA
YYENYSRGKWTLPCLFGIFAGFLRYWLAEMFNKTNKKFPLGTFLANVFATLLIGIFTMVQRGKKHFSTDVPIVNSLN
SCHIVSALISGFCGTLSTISTFINEGYKLSFINMLIYYTVSIAISYCLLVITLGSYAWTRGLTNPIC

Seq. ID No: 356
>IDI1_for_fusion
GGSGGSSGSGGSSSTADNNSMPHGAVSSYAKLVQNQTPEDILEEFPEIIPLQQRPNTRSSETSNDESGETCFSGHDE
EQIKLMNENCIVLDWDDNAIGAGTKKVCHLMENIEKGLLHRAFSVFIFNEQGELLLQQRATEKITFPDLWTNTCCSH
PLCIDDELGLKGKLDDKIKGAITAAVRKLDHELGIPEDETKTRGKFHFLNRIHFMAPSNEPWGEHEIDYILFYKINA
KENLTVNPNVNEVRDFKWVSPNDLKTMFADPSYKFTPCFKIICENYLFNWWEQLDDLSEVENDRQIHRML

Seq. ID No: 357
>INMT_1

MAAPHTSQQDYIDNFNARDYLQTSYTPGKGILFGEWIEFATQNLHETFTTGGVRGDTLLDFGTGPTIYQLISACEVF
DKIIVSDFLEQNRAEFRKWLNKDPDAFDWTPIIKGVCELEGNREDWEKKATKLRSKVKEVLKCDALKRNPYDPIVVP
PVDCLLSCLCLEAPCKDIKSYCEVLKNFQSLIKPGGHLLILSGLNATFYYVGKTYFSSMTTKKEELEMAFKEAGYII
KKAVYTPRADKSKIDVADYEGHYFIHAHKPK

Seq. ID No: 358
>INMT_2
MAAPHTSQQDYIDNFNARDYLQTSYTPGKGILFGEWIEFATQNLHETFTTGGVRGDTLLDFGTGPTIYQLISACEVF
DKIIVSDLLEQNRTEFQKWLNKDPDAFDWTPIIKGVCELEGNRENSEKKAEKLRSKVKQVLKCDALKRNPYDPVIVP
PADCLLVCLCLEIPCKDMKSYCNVLKNFKDLLKPEGQILILGTLNGTYYHAGKKRFSLLSSKKEDLEMAFKEAGYII
EKAVYTLRADKSNIDVADYEGHYFIHAHKPK

Seq. ID No: 359
>INMT_3
MSDFTNTREYEEQFDPRLYLETYFHLGSGSLADDFLRFVLDNFNKTFKSGAVKGSTLIDIGTAPSIYQLLSACESFD
DIIVTWHTNRELKELQKWLNSEADAFDWSSIVKHVCEIEGNRMAQKEKEEKLKGKIKQVLMCDVSKSNPLSPHEVPK
ADCLLTTVCLEAACKNYESYGTALKNLSNLLKPKGHLLMAGDLGANYYEVGSNKVFSLPVNEKFLKKVISESGYEII
QLVSFGKPENADFETSDYEGFYFVHAQKV

Seq. ID No: 360
>INMT_4
MDCLISCLCLEAPCKDLEDFTNTLKKFKELLKPGGHIIIQSVLNCSLYFVGNKSFSCLSITKDELEQAFKEAGYEIV
KLKVVPRSEKIWANVSDHSEYYYIHARKPQ

Seq. ID No: 361
>INMT_5
MSDFTGKNEYQTFFNPKAYLESYYQLGSGSMGDEYLQFVLKELAETFNPGKVKGDTLIDIGTGPTIYQLLSACEAFK
NIIVSDFTDKNREEFNVWLKNQPGAFDWSPVVKHVCRLEGDRIPWEQKEERLRKTIKQVLKCDVFNINPIDPVTIPQ
VDCLLSCLCLEGACKDFESYITALKNMTTLLKIGGYLVMTGDLGNTYYMVGDVKFSGLNLNENFLREAITGAGYVIE
SFQQSKKTEDSVEDKADFTAYYVIVARKERNV

Seq. ID No: 362
>INMT_6
MESGFTSKDTYLSHFNPRDYLEKYYKFGSRHSAESQILKHLLKNLFKIFCLDGVKGDLLIDIGSGPTIYQLLSACES
FKEIVVTDYSDQNLQELEKWLKKEPEAFDWSPVVTYVCDLEGNRVKGPEKEEKLRQAVKQVLKCDVTQSQPLGAVPL
PPADCVLSTLCLDAACPDLPTYCRALRNLGSLLKPGGFLVIMDALKSSYYMIGEQKFSSLPLGREAVEAAVKEAGYT
IEWFEVISQSYSSTMANNEGLFSLVARKLSRPL

Seq. ID No: 363
>INMT_7
MKGGFTGGDEYQKHFLPRDYLATYYSFDGSPSPEAEMLKFNLECLHKTFGPGGLQGDTLIDIGSGPTIYQVLAACDS
FQDITLSDFTDRNREELEKWLKKEPGAYDWTPAVKFACELEGNSGRWEEKEEKLRAAVKRVLKCDVHLGNPLAPAVL
PLADCVLTLLAMECACCSLDAYRAALCNLASLLKPGGHLVTTVTLRLPSYMVGKREFSCVALEKGEVEQAVLDAGFD
IEQLLHSPQSYSVTNAANNGVCCIVARKKPGP

Seq. ID No: 364
>INMT_8
MSDIDDGALASAQAIVDGNRLAGQIELRQQPDPDRVFAGVLRQGEAVAFCVCNPPFHESLEHARRAAGAKWQRLGRA
VQGKEMNYQGSPAELCCNGGEVGFVTRMAEESAQPRRQRACVWFSAMLSRESSIAPVRERLGELGARRRAWELRQGR
TTKWVVAWTFYPRGERDQRLREMAQRRADPEARAEAGAEAATARDVGAGGDGADGVGGSLVRRSAGAGGSAA

Seq. ID No: 365
>INMT_9
MDFTGGEIYQSSFDPKAYLASFCSLGSGRDDILMFRLKKCFETFGPGGLRGDVLVDIGTGPAIYHLLSACESFPYII
ATDFTDNNRQELEKWLRREPGTFDWLETVKIVCDLEGDSRDDWVEKEDKLRSRIQKVLKCDVTKTNPLDPTVIPPAD
CLITALCLETACTDIDTYFCSLRNITTLLKPGGHLVLIGVLGNSFYKVGEKKFYCLSLDEQTVRNAVIDAGYSIKDL
ELYYLPNPASCAHITDTYANIFLVAQKNET

Seq. ID No: 366
>INMT_10
MEIVSTSYNHIYDNFDARKYLDRYYGLASETQEIEEESVFLLTFLSNVFSSGRVKGHSFIEIGVGPSIHSILSACEV
FEKIYLTDYSQGNLNEIEKWLNSENDAFDWTPYIRFVCDLENNGSTPKGKKEKLRRAVSLMKCDVNLSNPLHPHSLP
LTDCLLTASCLSATCKTFTDFKMSLKIIVSLIKPGGHLILIDYLRASYYWVGEVKLPILSLDEHVVREAVVESGCKI
EEFKWFKEFHMPDELSDCKTVFSLLAQKL

Seq. ID No: 367
>INMT_11
MDSSNYKLYHVHEFNSRSFLDNYFSDGPQMTFVDDTLVFPIENLKKTFAEGHIKGDVMIDLSIGAMVHHLYAACEFF
KDIIVLKASDRCIMELKRWVGTRTGAFYWGHATKLHADTEGNSELLQDKEEKVRSAIQHVVKCDVTKELMTDPIVLP
PADCIISAWLLDAISSNQDDFITYLRRFIKLLKPGGHLILIGALEQTSYSVGNEKYQFLTYNEDFARKALIAEGLVI
DDCKIKKRTAKSDLADYKSILYLVSHKK

Seq. ID No: 368
>INMT_12
MDPCLNLYYPSHEVNAKRLLHEYFSQNVPYSIFKESTINIMKCCYKAFSSGLLSGTTLIDISVGPSIVHLLSVCEFV
EEISILKVNDASIRELELWKNKDPETFDWTHTLKLFMELKGTSRDGWKDAQEMLRRKVKHIVKCDFSKSNLTKPFAL
PRADCVTCIWGLETISRDHDEWKTTLRKISDLVKLGGHVLIHADINASYFKIGEDKYHLFNFDDAFLRKTLTDGGFA
IVHYENLEREACTDCLDHSHK

Seq. ID No: 369
>INMT_13
MELKRWVDTRTGAFDWSHAAKLHVDTEGNSDELQEKNEKVKSAIQHVVKCDLEKENMTHPIVLPPADCIISFALLDV
ISKDKDDYIKYLRKFSKLLKPGGHLILIGDLDTTYITVGKHKVHYLTYDEEFVRNALAGEGFVIDCCKVKERTVESD
LCDYKGMIFIVAHKEK

Seq. ID No: 370
>INMT_14
MELKRWVDTRTGAFDWSHAAKLHVDTEGNSDELQEKNEKVKSAIQHVVKCDLEKENMTHPIVLPPADCIISFGFLDV
VCKDQEDYIRYLRKFSRLLKPGGHLILIGGVDATYFTVGKEKHHFFTYDEAFVRKALEGEGFVIDDCKVKKRTAVSD
FTDYKGSIFIAAHKEH

Seq. ID No: 371
>INMT_15
MSDFTNTSEYEEQFDPRLYLETYFHLGSGSLADDFLRFVLDNFNKTFKSGAVKGSTLIDIGTAPSIYQLLSACESFD
DITVTWHTNRELKELQKWLNNEADAFDWSSIVKHVCEIEGNRMGQKEKEEKLKGKIKQVLMCDVSKSNPLSPHEVPK
ADCLLTTVCLEAACKNYESYGTALKNLSNLLKPKGHLLMAGDLGANYYEVGSNKVFSLPVNEKFLKKVISESGYEII
QLVSFGKPENADFDTSDYEGFYFVHAQKV

Seq. ID No: 372
>INMT_16
MALQERQEPDVYQENFEPTSYLEYYRMNQDPVGDEVLHFLLKHYNATFKPGGLEGKLLIDIGSGPTIYQFLSACESF
QEIIATDYTDKNLQELEKWLKKMPGAFDWSPVVKYVCELEGNRDKWAEKEERVRRAVTQVLKCDVLKERPLEPAVLP

PADGLISSLCLEAACPTPQACRDALRHLRTLLRPGGHLVLSGGFETTFFMVGDKRFSTLPLNEKFLREALQEAGFII
EKLEKVTRAAETHLDNRSDYTGLFFLVARRGD

Seq. ID No: 373
>INMT_17
MDKISAPFFSGTSPAAASVAGVDEDDRLCFQAQELMFAYNISMVLRAAIQLGLLDALSAAGGKALTPNELVENVETS
SNKAEAAAAVDRILRYLSCFNVVTCSSEAAGPDGTLVRRYTTGPLCRWLTKDRGDGTLSPFAVFVVDPDHLFPWHHI
AEAVTAGGPSAFERTQKWPYYEYMGKNQRLGTLFDNAMAQHSVILVTKMLERFKGFDGVQRLVDVGGGTGSTLGMIT
SKYKHMTGINYDLPHVIAQGLPLPGVEHVAGDMYESIPTGDAVLLQWITLMLNDDEFVKILSNCHNALPKDGKVIVV
DGILPENPDSSLTARDAFTLDIIMFVLFKGAKQRTEKEFARLAKQAGFTGGIKKTYIFFNFYALEFTK

Seq. ID No: 374
>INMT_18
MDANKRYHGPPVLLGVVRDSEKFDFCMCNPPFFETMEEAGLNPKTSCGGTPEEMICPGGEKAFITRIIEDSAVLNQS
FRWYTSMVGRKSNLKSLISKLREVGVTIVKTTEFVQGQTCRWGLAWSFVPPVRKIVSPHVAEKNIISFMLEVWVPGF
SICRVGDDLVPKSKSPHLSPILGTKN

Seq. ID No: 375
>INMT_19
MEEAGLNPKTSCGGTPEEMICPGGEKAFITRIIEDSAVLNQSFRWYTSMVGRKSNLKSLISKLREVGVTIVKTTEFV
QGQTCRWGLAWSFVPPVRKIVSPHVAEKNIISFMLEGLQRQFSAIHVLQSIESFFRTCGASSELNASSFTVDITATN
DHCKAILNNELQSIDEATSCEHVPETSNSSSSLHPHSNGLGFRISVYQQIPGTLLVKGSLQHKNNPVSGAFSLIIQR
LEEDLKYKFCR

Seq. ID No: 376
>INMT_20
MNRSNYIHWIEDLLASDITEKNEANGGKVRGFDIGTGANCIYPLLGASLLGWSFVGSDVTEVALDWAEQNVRSNPHI
SELIEIRRVDVDDPASSSGTVESSGGSRMEDSSQGQCDVVELASLEMKEFCDVGVTCKGGTDKNQRRYDEAKHSNVA
KGYQGPPILLGVVKEGEKFDFCMC

Seq. ID No: 377
>INMT_21
MEEAGLNPKTCCGGTPEEMVCQGGERAFISRIIEDSATLKQSFRWYTSMVGRKSNLKFLMSKLREVGVTIVKTTEFV
QGQTCRWGLAWSFMPTAKRSVPSHVAEKRNLSFMLEGLHRQTSAFNVLQSMESFFSHFGALCKSNPSSFTVDVSVSS
DHCDAILKSDVEKLDEASSHSCVAESPGSASSYDPMVVSFRLSVFQQIPGTLLVRGSLQQRDSPLSGAFLSVFQQLE
KFLKHKFCRERGLQFNQR

Seq. ID No: 378
>INMT_22
MATEIDDESYESARRNISNNNMQSRIHVEKASPDQSILFPLEDDRTFEFTMCNPPFYGSAAEVVQSAEAKEFPPNAV
CTGADIEMIYPHGGEEGFVMKILDESERFMTRCKWYTSMLGKMSSVATIVEVLRQRSITNYAVTEFVQGQTRRWAIA
WSFADTRLPDTMARIQSISPKHALYPCMPPKNTLVQAFPGPATHLVSTKLIETLHGIEGVSYTTTSLNSFFVEARQN
TWSRSARRSRANKNSSKKPDPSSLDADDILSGSQPALTCSCRVLADTAHADPVNVVENQWIFGNDRALFESFVGHVS
RKVGMGLRDVK

Seq. ID No: 379
>INMT_23
MLLESYKTFEPANYLQEYYSTVDLENRSLLAFFAEAYKGIDPNSVMLEFSGGPSLYSLITAAAHVKEIHFSDFLERN
VEEIKLWKRFRHRSYIWINFFKEALMAEGLSEVSTDDILEREELLSKKLSDFLLCDAFNRHPLGQRCYQRYDVVAAN
FVAESITPSLKTWEEVVNNICSTLKPSGTLIMTAIQGASFYCVENHRYPAIAVTPEDVIRVLSYQGFDVDNLLMRHI
PAEITDISAKDYKGYQGMLFVKATR

Seq. ID No: 380
>INMT_24
MESGFTSKDTYLSHFNPRDYLEKYYSFGSRHCAENEILRHLLKNLFKIFCLDGVKGELLIDIGSGPTIYQLLSACES
FTEIIVTDYTDQNLWELQKWLKKEPGAFDWSPVVTYVCDLEGNRTKGPEKEEKLRRAIKQVLKCDVSQSQPLGGVSL
PPADCLLSTLCLDAACPDLPAYRTALRNLGSLLKPGGFLVMVDALKSSYYMIGEQKFSSLSLDREAVRDAVEEAGYT
IEQFEVISQSYSSTTSNNEGLFSLVGRKPVGSE

Seq. ID No: 381
>INMT_25
MEIVSTSYNHVYENFDARKYLDRYYGIAPEAEKIDEESVFLLTFLSNVFSSGRVKGHSFIEIGVGPSIHSILSACEA
FEKLYLTDYFQGNLDEIKKWLNSENDAFDWTPYIRFVCDLENNGSTPREKKEKIRRCVSLMKCDVNLSNPLHPHSLP
LTDCLLTACCLTSTCKTFTDFKMSLKTIVSLIKPGGHLILIDYLRASYYWVGEAKLPLLSLDEHGVREAVEESGCKI
EEFQWFKEFHMPDEVSDCKTVFILLAQKL

Seq. ID No: 382
>INMT_26
MRNLHETFGPGGVKGDILIDFGAGPTIYQLLSACEVFNTIITSDFLEQNREQLKKWLRKDPDALDWSNFAKYVCELE
GKSDNWEKKEETLRRKVTKVLKCDALAEKPYDPVPMPEADCLISCLCLEVACKDLEDF

Seq. ID No: 383
>INMT_27
METPFTSQQTYVDEFKASDYFKTYYVAEGGIANEEWTDFALRTLHETFTKGGVKGETLIDFGAGPTIYHLLSACEVF
DKIITSDYLEQNRAELEKWLKKDPSAFDWTPIIKFVCELEGNRNYEKKAEKLRNKVKEVLKCDALKRNPFDPIVLQP
ADCLLTCLCLEAPCEDMKSYFNVLKNFKDLIKPGGHLVILSVLDATFYYVGDKYFSSMTTRKEELEQALKEAGFEIE
KAVYTTRKDRSQMDIADYQGFYYIHARNPK

Seq. ID No: 384
>INMT_28
MEGSFTGGEEYQKYFQPRDYLTTYYNFDGSPTPEAEMLKFNLECLHKTFGPGGLRGDTLIDIGSGPTIYQVLAACES
FRDITLSDFTDRNREELEKWLKKEPEAYDWSSVVKFACELEGDSGRWQEKEKKLRSVVKRVLKCDANLASPLAPAAL
PPADCVLTLLAMECACCSLDAYRAALCNLASLLKPGGHLVTTVTLGISSYMVGKREFSCVVLEKEGVEQAVLDAGFD
IQQFLHIPKCYSATIAANNGVCFIVARKKPAP

Seq. ID No: 385
>INMT_29
MEGSFTGPDEYQKYFSPKDYLDTYYSFEHGPSPETEMIKFSLQFLHKVFGPGGIRGETLIDVGSGPTIYQVLAACEA
FSDITLSDFTDRNREELQKWLRKDAGAFDWTPVLKFACELEGNSSHWQEKAEKLRATVKRVLKCDVNLGKPLAPVEL
PAADCVLTLLAMECACCSLAAYRAALCNLGSLLKPGGHLVTSITLQISSYMVGKHQFSCLYITKEEVERAILDAGFD
IEQLLHSEQSYSATIAPNKGICFIVARKRSGP

Seq. ID No: 386
>INMT_30
MDAQLTQLRNADVSWAAFDPIAYVDHNYRDLQAEDAEILHLVRDHFGDHFRKQGGGPVSGIDVGAGANLYPALAMMP
WCEEITLFERSPANVRYLKSQVDSYDANWDQFWDALCAHEAYNSLGTDPRERFGKVVWVEQGDLFDLARYERRWSMG
TMFFVAESMTTSYQEFMLGVERFMRALSPGAPFAAAFMEHSKGYHAGEHFFPACDVGESEVRASLEGFAGDFKVQRL
ESAAQLRDGYSGMIVAY

Seq. ID No: 387
>INMT_31

MSDFTNASEYEKQFDPRLYLETYFHLGSGSLADDFLRFTLGNFHKTFTEGEVKGTTLIDIGTAPSIYQLLSACEYFQ
DITVTWYTNRELQELQKWLNKDPGAFDWSSTVKHVWELEGKRGMLEEKEEKLRGMIRQVLLCDVSKKNPLEPVTLPK
ADCLISTVCLEAACRNYDSYRTALKNLSTLLKPGGHLLLAGDLGANYYEVGSNKVFSLPVNETFLRKAVNESGYVIN
KLVSFGKPEDAGYDTSDYEGFYFIHAQKC

Seq. ID No: 388
>IOMT_1
MSSKLDNQNITANEEEEAFHQAMQLAMSTILPMVLKAAIDLDLLEIIAKAGPAGCKLSPIEIASHLPTKNPDASSII
DRILRVLASHSILTCDLATNEDGHVQRLYGLAPIAKYFLHNDDGISLIPTLTISTDKYLLGAWYHLREATLEGGAIP
LVKAYGMDLFELAAKNDEISGKFNNTMGNQTAIIMKKVLEIYKGFEGINQLVDVGGGLGINLKLIVSKYPQIKGINF
DLPHVVKDAPHFLGVDHVGGDMFIEVPQGEVIFMKWILHDWGDDRCLKLLKNCYNALPKFGKVVVVELVVPESPMTD
IVTKNTLTLDAGLFIVVPGAKERTKEEYEALAKKAGFSTFRLVCRAYSYWVMEFHKNVIV

Seq. ID No: 389
>IOMT_2
MGSQAEVGKAMTEEEACEFAMQLVSSSILPMTLKAALELELLEIMATAGEGAQLTPAEIAAQLPTSNPDAPIMLDRM
LRLLACHSVLTASTYTDDDGKVRRRYGLAPVCKFLVRNQDGVSTAALSLVNQDKVTMESWYYLKDAVLEGGIPFNRA
HGMTAFDYPGTDPRFNRVFNQGMSNHSTLTMKKILETYTGFRGLHSLVDVGGGIGAILSLIVAKFPHIKGINFDLPH
VIDDAPQFPGVEHVGGDMFASVPTAEAILLKLILHDWGDEHCVKLLKNCCKALPEDGKVVVVEAILPEGIDHSYASA
CVYQVDMIMLVTNPGGKERTLKEFEELAKAGGFAGIRPICCVYGSWVMEFYKKM

Seq. ID No: 390
>IOMT_3
MGSTAETQLTPVQVTDDEAALFAMQLASASVLPMALKSALELDLLEIMAKNGSPMSPTEIASKLPTKNPEAPVMLDR
ILRLLTSYSVLTCSNRKLSGDGVERIYGLGPVCKYLTKNEDGVSIAALCLMNQDKVLMESWYHLKDAILDGGIPFNK
AYGMSAFEYHGTDPRFNKVFNNGMSNHSTITMKKILETYKGFEGLTSLVDVGGGIGATLKMIVSKYPNLKGINFDLP
HVIEDAPSHPGIEHVGGDMFVSVPKGDAIFMKWICHDWSDEHCVKFLKNCYESLPEDGKVILAECILPETPDSSLST
KQVVHVDCIMLAHNPGGKERTEKEFEALAKASGFKGIKVVCDAFGVNLIELLKKL

Seq. ID No: 391
>IOMT_4
MGSTAADMAASADEEACMYALQLVSSSILPMTLKNAIELGLLETLVAAGGKLLTPAEVAAKLPSTANPAAADMVDRM
LRLLASYNVVSCTMEEGKDGRLSRRYRAAPVCKFLTPNEDGVSMAALALMNQDKVLMESWYYLKDAVLDGGIPFNKA
YGMSAFEYHGTDPRFNRVFNEGMKNHSIIITKKLLEVYKGFEGLGTIVDVGGGVGATVGAITAAYPAIKGINFDLPH
VISEAQPFPGVTHVGGDMFQKVPSGDAILMKWILHDWSDEHCATLLKNCYDALPAHGKVVLVECILPVNPEATPKAQ
GVFHVDMIMLAHNPGGRERYEREFEALAKGAGFKAIKTTYIYANAFAIEFTK

Seq. ID No: 392
>IOMT_5
MGSAGETQITPTHVNDEEANLFAMQLASASVLPMILKSALELDLLEIIAKAGPNAQLSSSDIASQLPTKNPDAAVML
DRMMRLLACYNVLSSSLRTLPDGKIERLYGLAPVAKYLVKTEDGVSIAPLSLMNQDKVLMESWYYLTEAVLEGGIPF
NKAHGMTSFEYHGKDARFNKVFNKGMADHSTITMKKILETYTGFEGLKSLVDVGGGTGAVISMIVSKYPSIKGFNFD
LPHVIEEAPSYPGVEHVGGDMFVSVPKADAVFMKWICHDWSDEHCVKFLKNCYDALPENGKVIVAECILPVAPDSSL
ATKGVVHIDVIMLAHNPGGKERTEKEFEALAKGAGFQGFRVCCSAFNSYIIEFLKKP

Seq. ID No: 393
>IOMT_6
MGSTAETQITPVQVTDDEAALFAMQLASASVLPMVLKSALDLDLLEIMAKNSSPMSPSEIASKLQTKNPEAPVMLDR
ILRLLTSYSILTCSNRTILGGDSVERIYGLGPVCKYLTKNEDGVSIAALCLMNQDKVLMESWYHLKDAVLDGGIPFN
KAYGMSAFEYHGKDLRFNTVFNNGMSNHSTITMKKILETYKGFEGLTSLVDVGGGIGATLKMIVSKYPDLKGINFDL

PHVIEEATSHPGIDHVGGDMFVSVPKGDAIFMKWICHDWSDEHCVKFLKNCYEALPEDGKVILAECILPETPDSSLS
TKQVVHVDCIMLAHNPGGKERTEKEFEALAKGSGFKGINVACNAFGVYVIELLKKM

Seq. ID No: 394
>IOMT_7
MEMINFMHHMDSTWNLCGKDVVQAFDFSEFHTVYDLGGCSGGLAKQFVSTYNDSTVTIMDLPKVVQTAKKYFVTDQE
QQIHFIEGDLFNDPIPEADLFIMARIIHDWTEEKCLELLRKIYQSCRPGGGVLLVEVLLNEDKSGPLMSQLFSLNML
VQTEGRERTPSEYTKLLTDSGFRDIQVKITGKIYDA

Seq. ID No: 395
>IOMT_8
MERLLDACVGLKLLKVELKSNKGYYSNTDVSTMYLVKSSPRTLYYMIMFYSKTTYMCYNFLPQAVREGQCQYERAFG
ISSKDLFEALYRSEEDTLAFMYFMNSTWSICGKYVVQAFDLSEFHTIYDLGGCTGALAKQLVSTYKESTVTIMDMPN
IVQAAKKHFVTDKEQQIHFLEGDFFNDPIPEA

Seq. ID No: 396
>IOMT_9
MIPFNKAYGMTAFEYHGKDDRFNKVFNAGMFNHSTMTMKKILDIYDGFNNLTTLVDVGGGTGASLNMIVSKHPSVKG
INFDLPHVIQDATTYPGIEHVGGDMFESVPKGDAIFMKWICHDWSDAHCLKFLKNCYKALPDNGKVIVA

Seq. ID No: 397
>IOMT_10
MAQAAAEAEGITPVMDLLFAAQGSSALLVCARLGLFDYISSQGEEGVSCKQLASRAQWSTRAASAVMVSLAASGILA
VKPSSAGAQHCFEHSYTLTPRAQRFLVTEKPGSMSAYTEIHWEASPELLLKKAAETEDEKRNFMLETGGGAPSEVFL
AAMQGQSSYAAMVLTRLVDLSDTRTFVDVGGGSGTFAIEACKATPNLQGVVYDLAGACPTTDGFIARAGMAERVKTH
AGNMFEDERFPAADCYAFGNVLHDWSDQDNSKLLRKAFESLPAQGKVLLLEMLVEEDVVSTSPSAAGLNLCMVTNEL
GRQFKASELRAMLLEAGFAGAEVVSSPLTPYSLVVGTKGEANPVASKPEAAAAAESESITPLMDVLFSAQHSAVLIV
CSRLGVFDFVGAQGESGASCAQVAAHAKWTTRAASAMLVSLACSGLLEPTPGSAAAQHCFEHSYRLTPLARRFLVAG
QPGQLSAYTEIFWGASPKQLLEKASASLGEWGEGNFMLDAEGGAPSEVFLAAMQAQSTYAAMVLTRLVDLSDVRTFV
DVGGGSGTLAIEACRAAPGLQGVVYDLAGACPVTDGFIARAGMAERVKTHAGNMFADERFPAADCYAFGNVLHDWSD
QDDGKLLRKAFESLPANGKVLLLEMLLAEDVESSTRSATGLNIVMVTNEQGRQFKGSELEAMLRAAGFAATEVVRSP
LTPYALVVGTKG

Seq. ID No: 398
>IOMT_11
MSRTSWDEGEDVDLDSVAYGFMASQALFTGLELGIFDHIAAAGAGGLSAAGIGKACGIEAPRVQTLLTSLVAVKCLK
RDASAMYTLSPNTAQYMVTSSRHFYGDYLRYQIGRQFYHRMGALPEVMTSGKAPSYASWFSDPEVARTYTQAQHNGS
VATAKYLIKKKLQLGGISAMLDVGGGSGAFSYVFTQATPGLHSKVLELPEVCRTGEGIREKQPEDVRSRVSFVELDA
SSPTWPVDDSAFDVVLMSYISGSVPEPIIGSLYANAMKALRPGGRLLVHDFMVNDSLDGPALGALWGLQHVTVNADG
LGLCPKEVIARMGAAGFDTSKCEAMEMIHGMTKLIVGHKG

Seq. ID No: 399
>IOMT_12
MCSSKELDFPHILIDYQHGFLVSKTIFTACELGVFDLLHEVQEPVPAATIASRLSTSEDGMERLLDACVGLKLLKVY
LKNNKGYYSNTDVSTIYLVKSSPKTLHYMMIYYSKITYMCWHFLPQAVREGKRQYERALGTTSNDLFEIVYRSEEEM
TTFMHFMDSTWNLCGKDIVQAFDLSEFHTVYDLGGCSGSLAKQLVSTYKESTVTIMDLPKVVQAAKKHFVTDKEQQI
HFLEGDFFNDPLPEADLFIVARIIHDWTEETCIKLLKKMYHSCRPGGGVVIVELLLNEDKSGPVISQVYSLYMLVQA
EGKERTPSEYTKLLTDSGFKDIKVKATEKLFGAILGRK

Seq. ID No: 400
>IOMT_13

MCSQEGEGYSLLKEYANGFMVSQVLFAACELGVFELLAEALEPLDSAAVSSHLGSSPQGTELLLNTCVSLKLLQADV
RGGKAVYANTELASTYLVRGSPRSQRDMLLYAGRTAYVCWRHLAEAVREGRNQYLKAFGIPSEELFSAIYRSEDERL
QFMQGLQDVWRLEGATVLAAFDLSPFPLICDLGGGSGALAKACVSLYPGCRAIVFDIPGVVQIAKRHFSASEDERIS
FHEGDFFKDALPEADLYILARVLHDWTDAKCSHLLQRVYRACRTGGGILVIESLLDTDGRGPLTTLLYSLNMLVQTE
GRERTPAEYRALLGPAGFRDVRCRRTGGTYDAVLARK

Seq. ID No: 401
>IOMT_14
MGYAAPQARQSDKQIFDIYFGFLHSYALLFADEVGLFDLLRCEALTLDQVSMATSLPSRSSQALLSLCASLGLLEKR
GERFALSALTEGFLVREAETSFCGVLASARGQAAAFSYDFFKASLLKGESQLFGGRDLFDNNAQDPEHCEIFTRAMH
SKSKGPAQAWVEKIDLSAHACLLDVGGGSGVHAISALARWPNLNAVVFDLPPVCAIADTFIERYQMTARAQTHGGDI
WYTDYPFADAHFYSDIFHDWPLERCRFLARKSFDALPSGGRIILHEMLFNAQKTGPRNVAAYNANMLLWTQGQQLSE
PEAADLLQAAGFVEILAFPTGYGDWSLVTGVKP

Seq. ID No: 402
>IOMT_15
MGSIDAQMAAVEEESCIYAMQLAYTVVLPMTLKNAIELGMLEILMGAGGKMLSASEVAAQLPSTTTNPDAPAMVDRM
LHLLASYKVVSCEVEEGTHSRRYGPAPVCKWFTSNKDGDGASLAAMLLLTNEKVLLESLNHLKDAVLDGGHPFLKAH
GMTVYEYNKTDARMKRVFSQAMNNYSTIINRKLVEMYMGFHDIAFLVDVGGGVGTTIRAITSKYPHIKGINFDLPHV
IADAPQCPGVQHVAGDMFRNVPSGDAIILKWMLHNWTDEHCTTLLRNCYDALPPHGKVFIVENILPLKPDATSRGQQ
TSLSDMIMLMHTPAGRERSQREFQELGKAAGFTGFKTTYIYGNSWVIELTT

Seq. ID No: 403
>IOMT_16
MSFDTQHALQPYWDLAVAPVQADGLAAALELGIFEVLATPHTPAQLADVLSLHGPHTALLLELLWSMQVLERDGADA
DTDANALRYRCTATTLQYFCRDAVAFCGDAWLYRLHALRHFATQLNTLVRDGGKVTPYSTASGVNWAAAAQQQIGQE
QRAVTMRAALCVMQRVAPFADGNTPLRLLDAGGGPGWVAIALAQAHAGVHGCVFDWPETVAVAAANIAHAQLSDRLE
TLGGDLDSDDIGGGYDLIWCSSVLHFVPDMAAALRKMQAALKPGGVLVCIQAEIAAAPGDAARVLPYYLPMRMLGRT
VTRHGELAQLLRDTGWRQVEQYGASDFPMAPVQVLIARA

Seq. ID No: 404
>IOMT_17
MQLASASVLPMVLKSAIELDLLDIIAKAGPGAYLSPSEVASQLPTSNPDAPVMLDRILRLLASYSVLTYSLRTLPDG
RVERLYGVGPVCKFLTKNEDGVSIAALCLMNQDKVLMESWYYLKDAVLEGGIPFNKAHGMTSFEYHGKDLRFNKVFN
KGMSDHSTITMKKILETYKGFEDLTSLVDVGGGTGAVLSTIVSKYPSIRGINFDLPHVIEDAPSYPGVDHVGGDMFV
SVPKGDAIFMKWICHDWSDEHCLKFLKNCYEALPDNGKVIVAECILPVAPDTSLAAKGVIHIDVIMLAHNPGGKERT
EKEFEALAKGAGFQGFRVMCCAFNTYIMEFIKKL

Seq. ID No: 405
>IOMT_18
MLNHTTMVIKKILECYKGFETLKQLVDVGGGLGVALNLITSKYPHIKGINFDLPHVVQHAPSYPGVEHVGGDMFKSV
PKADAIFMKWILHDWSDEHCVKLLKNCYAAIPNDGNVIVVDAVLPKMPEVSTSMRCTSQLDVLMLTQNPGGKERTEE
EFMALATKAGFKGIRYQECFVNTFWLMEFFK

Seq. ID No: 406
>IOMT_19
MERKEEVALLKGQAEIWQHLFAFADSMALKCAVELRLADIIHSHGVPITLSQIASAIDSPSPDIAYLSRIMRSLVYK
KIFTEHHPSDGGETVLYGPTHTSRWLLHDAELTLAPFVLMENNQWQLAPWHFLSQCVKEGGIAFKKAHGFEMWDFAA
RNPEFNKIFNDAMACTTKILMGVLLAEYKDGFGSIGSLVDVGGGTGEMIAEIIKQHPHIKGMN

Seq. ID No: 407

>IOMT_20
MGSASGSAERTQMGEDEACSFAMTITSGSVPPMVLKAVIELDVLEIIKRAGPGAHLSPAEIAAQLPTTNPGAAAMLD
RMLRLLASYDVLSYSLHTLPDGRVERLYGLAPVCQFLTNNEDGVTLSALSLMNQDKVLMESWYHLKDAVLDGGIPFN
KAYGMTAFEYHGTDPRFNKVFNNGMSNHSTITMKKLLENYKGFEGVSTLVDVGGGTGATLNMIISKHPTIKGINFDL
PHVIEDAPTYPGVEHIGGDMFVSVPKGDAIFMKWICHDWSDEHCLRFLKNCYAALADHGKVIVCEYILPVAPETNHA
ARTVFHVDAIMLAHNPGGKERTEQEFESLAKGAGFEGFRVAFFF

Seq. ID No: 408
>IOMT_21
MALNPPHQNNVMEKEDLCSFALSIATSSSLSMVLKAIIELDIIGIINRAGPGAHLSPAQIAAQLPTKDPGATASMLD
RMLRVLANNSILSCSLRALPNDGPIERLYGLAPVCQFFTKPEDFGPMVLFSQDKVYTDTWHHLKDAVLDGGSAFKKA
HGTTLFEYLGTDMRFSKVFNDAMSSSSTITMKKMLENYNGFDGLSTLVDVGGGTGETLNMIIAKYPTIRGINFDLPH
VINDAPNYDGVEHVVGDMFVSVPKGDAIFMKWICHDWSDKLCLKLLKNCYTALPNHGKVIVCECILPVAPETSHSAR
VASNLDMHMLAYCRGGKERTEQEFEALAKGAGFESFRVVCSAYDLKLYMC

Seq. ID No: 409
>IOMT_22
MAEIPTSSNPSDDPETQKLNGNEEDYDHHHDEDPESDDENYEYALQIAEMLPFPMVMHTAIELDLLGIIATAGPDRQ
LSAAEIAAALPAAGNPDAPAMLDRMLYLLATYSVVTCTAVDGGASGGVVRKYGLAPVAKYFVSNKDGVSLGAVISLN
QDQAVLASWSKLKEAVLEGGIPFNKVHGMDAFEYQGTNPRFNEIFNKAMYDQSTYIIKKIVRRYKGFENIQRLVDVG
GGLGHTLRVITSNYPSIKGINFDLPHVIQHAPTIPGVEHVGGDMFESIPHGDAIFMKCILHDWSDEHCLKTLKNCYK
ALPRKGKVIVVQMNMIEEPQTTPLAKAISQMDLWMMTQNPGGKERTRREFQALAEAAGFAEFNPVCHVAGFWVMEFL
K

Seq. ID No: 410
>IOMT_23
MSPIDLANELQTLVTSTYSGDVTDPFKLYKAKHSISDLCLSLLRAVQGPEEYTAILAESCQESSALNVVASLGVADH
IAESPNGELTLQELSEKVKADEKYLSVVLSSLVYHGYFKEVGGFGSQVYANNDFSSLLLSEETNAKGGKSMKDAIGL
SADDGAKATTRLLDAATGKAKGEAKTAANIAFDFSESLFQWMASPGNEWRGKRTAKAMVQLHGMANGGIGEDYPWEK
LATPIIDIGGGIGSFQGMLLALPKNKELTFTIFDIEKTVEHAKKVWAGKPQWMQDKVSFIAGDFMKSSPNDSKIPTP
AQGAGTYVIRHVLHDWDDAQVVTILKHVRNAMLGSPASTPPKLLLVEMMLNETSSRFTRTTSLQLLSLNGGITRTEV
QFRRLIKEAGFTVDSVTEVRGVDLVVELSPASL

Seq. ID No: 411
>IOMT_24
MPSTTISQLVGLIQQSVMALEKLCLENRTSLPDLDAFHFDQSSETFRSLPGAAQDAKIAVAACMQLIAILSPPTDTV
YRAALGGHLSFATRTCLEANITEILREAGPEGLHINDIASKCGLDPSKLGRVIRYLVIHHIYREVKPDVFTNNRTSS
TMDTGKPLDKLISEPDRKYDDTGFPALISHFMDVDQKCGAVGWDVLKDPVLGHSCDLTETIFSRAFNTKSKYWDFFD
HPENHYMRRRFDYAMKGLGAIEDHDMVLHAFSWEDLDKGSVIVDVGGGIGTAMLPLARKYPNFDIVIQDLPIVIEEG
TKFWSQNLPDAVANGNIKLHAHNFFDEQPIKNASVFYLRHVLHDWPMPDMVKILRRLRDVAAANTTLIILDYILPYS
CKMFADKDAVSIASARYYSEAPEPLLPNYTHKNVISDSDMYVFQMMFHYNSQEHTYLSLKSLLDASGWRLVRLRAID
PRNDYFQSIECKILA

Seq. ID No: 412
>IOMT_25
MAQPMMLALAKLISDSVAKVDQLCIEQGVIFPSLDDPFTTESESIKLHPDVAEASNYIISAAAQLIAILRPVPVTLS
TSAIHVHVSSALRVVVDSNVVEILREAGPQGLHVKKISEKNGVEAGKLGRLLRLLASGHMFKEITPDVFATNRISSA
LDTGKPYEELVKNPGEKLIGTNGIAAYISRSTDESVKSSGFLYEALTYSSSEKVPLPPSPFNLAFNTELHIFSWLAQ
KGNEHRLQRFGIAFDGFDKMLPVNGVTKGYRWGSLPKGSIVVDVGGGVGSESMKIAKTFPDLKVIIQDAEGVVANGV
KFYETRFPEGLSSGQVTFQAHDFFTPNPVTNARVFFMRFVLHDWPDATCVKILKNLRAAAAPDTELIINECLIQYAC

STESEISKSIPGGRFKPPPSPLLPNLGYARIFHYLIDLQMAIVAHGVERTVEQYASILQKSGWKLKEVLRMPESAYS
LHKLVAVPQPE

Seq. ID No: 413
>IOMT_26
MTRLTDSLGMLRSKLVPPQATMLQLLTGYRVSQGIYVVAKLGIADLLATGSKTSQDLAAITNVHAPSLYRLMRSLAS
LGIFTETENGRFELTPLAATLRSDHPNSVHDAAIMFLEDWHWQAWGNFFDCVKTGETALEKTFGTSNVFDYFETQNP
EAGQHFDNAMTNTSVMTNQALPTAYNFGAFKTLVDVGGGQGSFLSALFHQWDHLHGILFDLPPVIESAEQQNLLSGF
EKRTTLAAGDFFKAVPDGADAYLLKTIIHDWDDASAIAILKTCRRAMNHDSKLLLVELIVPSGNAPSLSKILDLEML
AVFGGVERTEAEYRSLLLSAGLKLTRIYDSPCPWSVIEAIPV

Seq. ID No: 414
>IOMT_27
MSMPPAHSRLYSRSFLSMLPDAITPFPYLPPDATDTRPLLAELEALLEIINSSARLAITEYKKHGNNVPTIYSTEFH
PLDFATDTVALKKAIRLLEDACQQLCASLAPPQHTLANVSRVHHRQYVTQLTTHDILEKYPSGSHIRELSQTVGLEK
GKLARILRVFAFKGCFIEVDTDVFASNRLSLIMKSSNDCGCLTCIHAQDVSQGAGVLYETLTEPEYAMSYEPDKAPM
IYVLKRKGLKGSFFDWMKADAKRRENYHYAMIALGPVMGSLSILHHYPWNDVATVCDVGASVGSVSIPLSKAHPHLK
ITDQDLPEVLEAARSVWEKEAFEALREKRVEFLTLDFFKEAPVPGKDVYYLRHIIHDWPDAEAAVILRNISKAMEPH
SRLLIHNYVIAGANRRPDEEQRAPEPMLPNFGAGDSRKYRQDLNMWILHNAKERTVDDQITLA

Seq. ID No: 415
>IOMT_28
MAPGREGELDRDFRVLMSLAHGFMVSQVLFAALDLGIFDLAAQGPVAAEAVAQTGGWSPRGTQLLMDACTRLGLLRG
AGDGSYTNSALSSTFLVSGSPQSQRCMLLYLAGTTYGCWAHLAAGVREGRNQYSRAVGISAEDPFSAIYRSEPERLL
FMRGLQETWSLCGGRVLTAFDLSRFRVICDLGGGSGALAQEAARLYPGSSVCVFDLPDVIAAARTHFLSPGARPSVR
FVAGDFFRSRLPRADLFILARVLHDWADGACVELLGRLHRACRPGGALLLVEAVLAKGGAGPLRSLLLSLNMMLQAE
GWERQASDYRNLATRAGFPRLQLRRPGGPYHAMLARRGPRPGIITGVGSNTTGTGSFVTGIRRDVPGARSDAAGTGS
GTGNTGSGIMLQGETLESEVSAPQAGSDVGGAGNEPRSGTLKQGDWK

Seq. ID No: 416
>IOMT_29
MEVVPSWFKETLDKSQFSAPYEYAVETAKQKALEVARRMHVKHLKTPDIVIGADTIVTLEGAILEKPFDKQDAYNML
SRLSGKEHSVFTGVVIVHCRSKEENHLETDIIDFYEETKVKFADLSEDLLWEYIDSGEPMDKAGGYGIQSLGGMLVE
SVHGDFLNVVGFPLNHFCRKLTEIYYPPPKQAICRVKHDSIPYVESFENLSDVETDCTSTSKACEAKKAVQDGVCKA
DGSGSAVLQNGIEERPVHCAQQLSKITQLLDGFKASQTLFAASKLKVFDKLKDKGALKAMEIAEKINASVHGTERLL
DACVALGLLEKTHQVYSNTELANTYLVSDGAFSIHEYITYSSDHLWSHFTHLDSAVVEGGGQHQTAVKKACDNRNGS
EVKERFMRAMHCMLKITARDLVTAFDLSKYSSACDLGGCTGALAHELVWTYPEMKVNVFDLPEVIKHTSQFQPESFD
SSRVTFSSGNFMEDTLPEADLYILSRVLHDLPEGKLNHLLKKVSEACCPGRSALLVAEIVLDEDKKESRGLLQSLSM
GEGKQRSGTEYKKLLENHGFNSVQIKSTGNLLDAILAIKTS

Seq. ID No: 417
>IOMT_30
MDTVKNLQASNVPSSLSQEDEEVFTSGLHVCSSEVFSHALSNCIQLGLFDIIAEAGPSAYLTATEITAQLPTKNPDA
VSMIDRMLRLFSCHSLLNSSLKTVADDVVETRYGLSPIGHLFVRKKDGVTMAACFTDYKAWTEAWLHLKDAILEGGN
PYEKAHGVPIYEHISSDTESVKGFSQAMDSISSFIMKKVLENYSGFKGLGSLVDVGGGSGFALNMITSEYPSISCIN
FDLPHVVQEAPYHPGVKHVGGDMFLDIPSADAIMIKEVLHNWGNEDCVKVLKNCYEALPKGGKVIVVSHVMPEVVGS
SNAAAKYVCQLDVMMLLFGGGKERTEKEFKALGKAAGFSGFQLICFAAYNAVAVMEFYK

Seq. ID No: 418
>IOMT_31

MAEDVAAVADEEACMYAMQLASSSILPMTLKNALELGLLEVLQKDAGKALAAEEVVARLPVAPTNPAAADMVDRMLR
LLASYDVVKCQMEDKDGKYERRYSAAPVGKWLTPNEDGVSMAALALMNQDKVLMESWYYLKDAVLDGGIPFNKAYGM
TAFEYHGTDPRFNRVFNEGMKNHSVIITKKLLEFYTGFDESVSTLVDVGGGIGATLHAITSHHSHIRGVNFDLPHVI
SEAPPFPGVQHVGGDMFKSVPAGDAILMKWILHDWSDAHCATLLKNCYDALPEKGGKVIVVECVLPVTTDAVPKAQG
VFHVDMIMLAHNPGGRERYEREFRDLAKAAGFSGFKATYIYANAWAIEFIK

Seq. ID No: 419
>IOMT_32
MTSLQDLDYPQQLLEYKDGFLVSKTMFTACELGIFDLLHKSDEALSALTISSHLGTSADGTDRLLSACVGLKLLKVE
MKNNEAFFSNTDVSDVYLVQLSPRSLYHMMMYYSQTLYKCWHFLPDAIREGKSQYERAFGVSSGDIFKALYRSEEEM
VTFMHHMDSVWNICGKDIIAAFDLSSFNEVCDLGGCSGGLAKQLLSIYPSSSVTILDLPEVVQTAKKHFITDADCNI
AFLQGNFFNDPIPEADLYIMARIIHDWTQEKCLQLLNKIYKSCRPGGGVLLVEVLLNEDRSGPLTSQLYSLNMLVQT
EGRERSPCEYTKLLAHSGFRDIQVKATGKIYDAILGRK

Seq. ID No: 420
>MUP1
MSEGRTFLSQLNVFNKENYQFSSSTTKKEVSNSTVDADNGASDFEAGQQFATELDQGEKQLGILSCIGLICNRMLGT
GVFAVSSTIYTLCGSVGLALIMWAVGAIIAISGLYVYMEFGTAIPKNGGEKNYLEAIFRKPKFFITCMYAAYIFFLG
WAAGNSINTAIMFLTAADTEVTKWNQRGIGVAVVFFAFLINSLNVKIGLYLQNILGIFKIGIVLFISITGWVALGGG
LKDGYQSHNFRNAFEGTETATAYGIVNALYSVIWSFVGYSNVNYALGEVKNPVRTLKIAGPTSMVFLAIIYIFVNIA
YFAVVPKDKLISSKLILAADFFDIVFGGQAKRAAAALVGLSALGNVLSVIFSQGRIIQQLGREGVLPFSNFFASSKP
FNSPMVGLFQHFIVCTVTILAPPPGDAYLLVQNLISYPMNIINFAISAGLLWIYWQRRQGKIEWNPPIKAGVFVTGF
FTLSNLYLIIAPYVPPSNGESVYSSMPYWIHCVIAWGIFFFGGVYYVVWAQLLPRWGHYKLVSKDVLGEDGFWRVKI
AKVYDDTIGDVDTQEDGVIETNIIEHYKSEQEKSL

Seq. ID No: 421
>NAT_1
MAPIEEEEPLPEELVLLERTLADGSTEQIIFSSAGDVNVYDLQALCDKVGWPRRPLTKIAASLRNSYLVATLHSVTM
PSKAEGEERKQLIGMARATSDHAFNATIWDVLVDPSYQGQGLGKALMEKVIRTLLQRDISNITLFADNKVVDFYKNL
GFEADPQGIKGMFWYPRF

Seq. ID No: 422
>NAT_2
MSTPSVHCLKPSPLHLPSGIPGSPGRQRRHTLPANEFRCLTPEDAAGVFEIEREAFISVSGNCPLNLDEVQHFLTLC
PELSLGWFVEGRLVAFIIGSLWDEERLTQESLALHRPRGHSAHLHALAVHRSFRQQGKGSVLLWRYLHHVGAQPAVR
RAVLMCEDALVPFYQRFGFHPAGPCAIVVGSLTFTEMHCSLRGHAALRRNSDR

Seq. ID No: 423
>NAT_3
MTSDVGADEHATTEAGGGRLQAGGHSSAEEASERCPPAAAPPSGMKGAADCGPQDSSARDDVSFIPYKDETDMPGIV
ELIEKDLSEPYSIFTYRYFINNWPELCFLTMRGDSCVGAIVCKLDVHRCRNTNRGYIAMLAVEKGLRGKGIGSTLVR
LCLDKMREMGADECVLETEVTNKGALGLYRNMGFVKEKRLHKYYLNGNDAFRLKFLFKLPEGFDRGEGCLGPLCEVP
PVTT

Seq. ID No: 424
>NAT_4
MVSIRPATVDDLLAMQACNLCCLPENYQMKYYFYHMLSWPQLLYVAEDYGKKIVGYVLAKMEEDSSEVHGHITSLAV
LRSHRKLGLASKLMRAAMAAMEETFGAEHVSLHVRVTNRAAFTLYSETLGFEINDVEHKYYADKEDAYDMRKMFETG
LKKQEAGKQKKKEKEKEKEKEKEKEKEKEKEKEKGKDSQPVEQQGGAAGADKEAQRSKRARSGDRKRNGRPRRRSGS
G

Seq. ID No: 425
>NAT_5
MLPRPPVGAAKEGHLTLFYRELRWLCPGTRFYFVVRDPAENVRSIADRLALGPEGLRRPPRIVARADLGWREVLNMS
YAGVREESALGTLVGRWNLMARLYLDAPKGAMALVRYEDLVAEATWEAEVRRVAAAETLDLRERVLWPGRPDLCTLP
GDESALHFGAVAAGKVLGVISVFLSPEPGGRAQFRKFAVDPEVQGRGLGRRLLEQAVAAAREAGAGSLFCHARADQQ
GFYERRGLHVVGEPFEKYGGKPYVEMEVPFQ

Seq. ID No: 426
>NAT_6
MKGSRIELGDVTPHNIKQLKRLNQVIFPVSYNDKFYKDVLEVGELAKLAYFNDIAVGAVCCRVDHSQNQKRLYIMTL
GCLAPYRRLGIGTKMLNHVLNICEKDGTFDNIYLHVQISNESAIDFYRKFGFEIIETKKNYYKRIEPADAHVLQKNL
KVSSPAPNADVQKSEN

Seq. ID No: 427
>NAT_7
MSTPSIHCLKPSPLHLPSGIPGSPGRQRRHTLPANEFRCLTPEDAAGVFEIEREAFISVSGNCPLNLDEVRHFLTLC
PELSLGWFVEGRLVAFIIGSLWDEERLTQESLTLHRPGGRTAHLHALAVHHSFRQQGKGSVLLWRYLQHAGGQPAVR
RAVLMCEDALVPFYQRFGFHPAGPCAVVVGSLTFTEMHCSLRGHAALRRNSDR

Seq. ID No: 428
>NAT_8
MSSGGVIVDLHRNSTNWAKVVDDIVKLERKIFPKHESLARSFDEELGKKNTGLIYMEVDGEVVGYAMYSWPSSMYAC
VTKLAVKENCRRQGHGEALLKAAIKKCRTRNVHRISLHVDPLRNPAISLYKKFGFQVDNLIDGYYSSDRNAYRMYLD
FDAD

Seq. ID No: 429
>NAT_9
MDERVVVELKKSLADYPKVLEELVRIEKKVFPKHESLSRSFDEELGKKNSGLLYICSNGEVAGYVMYSWPSALLAVI
TKLAVKEKYRRQGYGEALLRAAIQKCKTRNIQRISLHVDPSRTPAANLYKKLGFRIDSLVEKYYAADRDAYRMYLDF
DADV

Seq. ID No: 430
>NAT_10
MMEGAQEDEETEEKAEFDASEIEYVSYGGEHHLPLIMCLVDHELSEPYSIFTYRYFVYLWPQLCFLAFHKGRCVGTV
VCKMGDHRHTFRGYIAMLVVIKPYRGRGIATELVTRAIKVMMESGCDEVTLEAEVTNNGALALYGRLGFIRAKRLFR
YYLNGVDAFRLKLLFPRSEMHPSLHLLADQDGHDDQIAMEGEA

Seq. ID No: 431
>NAT_11
MKQVGISLDAVREKNLMQLKKLNVVLFPVRYNDKYYADALASGEFTKLAYYSDICVGAIACRLEKKDPGAVRVYIMT
LGVLAPYRGLGIGTELLNHVLEQCSKQNISEIYLHVQTNNDDAINFYKKFGFEVTETIQNYYTNITPPDCYVVSKRL
EAQPKK

Seq. ID No: 432
>NAT_12
MNIRVAKVEDLMGMQACNLQNLPENYMMKFWMYHSMTWPQISFVAEDHKGRIVGYVLAKIEDPSEEGTTEEIHGHVN
SISVLRSYRRLGLAKKLMLLSQEAMSSIYKASYVSLHVRKSNKAAIALYKDTLGFEVAKVEKKYYGDGEDALSMRLS
LKNP

Seq. ID No: 433
>NAT_13

MSDFQVAPLTARELARVRDLHAKLLPVQYPVSFFIHLLVIPSRACYVAYSHGSPVGFISAALHNPTRCFISGDSEVS
PRLEILTLGVLPAFQHRGLARRLIMSLVNAFKQDPATPILIYANVSTTNTRALQFYERMGILVSSDIITNLYRTLSY
GSRDAYLVVGAL

Seq. ID No: 434
>NAT_14
MLSIHPLKPEALHLPLGTSEFLGCQRRHTLPASEFRCLTPEDATSAFEIEREAFISVSGTCPLHLDEIRHFLTLCPE
LSLGWFEEGCLVAFIIGSLWDKERLTQESLTLHRPGGRTAHLHVLAVHRTFRQQGKGSVLLWRYLHHLGSQPAVRRA
VLMCENALVPFYEKFGFQAMGPCAITMGSLTFTELQCSLRCHTFLRRNSGC

Seq. ID No: 435
>NAT_15
MADAPSGPSVLSHYPGAGLALPPGDEQEDGEEEEEGRYEPRRGHHHHRRHHQQQQLNGLISPDLRHIKALKSKLPPP
PHDERTGAPNGLERLQDLEEEEAVLASRMGACSLHPGDGSIRYVRYESELQMPDIMRLITKDLSEPYSIYTYRYFIH
NWPQLCFLAMVEEECVGAIVCKLDMHKKMFRRGYIAMLAVDSKYRRKGIGTNLVKKAIYAMVEGDCDEVVLETEITN
KSALKLYENLGFVRDKRLFRYYLNGVDALRLKLWLR

Seq. ID No: 436
>NAT_16
MDAAMPTEISFRQPTPDDAARCFEIETSAYEGDEAATLEKIATRIALYPEGFVILEADGKIAGFINSGCAFEVVMSD
EEFKELVGHDPAAPNAVIMSVVVDPAEQGKGYSKLLMQHFIARMKAMDKKTIHLMCKEAHVPLYARMGYRYTRPSAS
DHGGMAWHEMVMEL

Seq. ID No: 437
>NAT_17
MEGLHSEWEVGAELKALGAVPKPFIGSHVSGKLIQRLKQDLRQSWDRGQSQARPTCTLPQPLPAPLGSSVPSASAQT
QVSRLVPVAPPQPDPAMSVLNAVPFMRPIHLRSPRQQRRHTLPASEFRCLSPEDAVSVFEIEREAFISVSGDCPLHL
NEVRHFLTLCPELSLGWFEEGRLVAFIIGSLWNQDRLSQDALTLHKAEGSSVHIHVLAVHRTFRQQGKGSILLWRYL
QYLRCLPFARRAVLMCEDFLVPFYSKCGFKAVGPCDITVGPLTFIEMQCPVQGHAFMRRNSGC

Seq. ID No: 438
>PSIK_1
MQANRPISDQDQDQFKLNLTTADGTRSYLEKHLSLNVEAVERLSGGFINFVWRAKLGTPYEGQNSIVVKHAPPFTAM
DSSLNVAVERLKFEYDSLKMIGSEPSIAGEDALISVPSVYHHDNIKHVLIMQDVGTMSTLRDFMGASPPPPTDMAAL
IGCQLATFIAGLHNWGRNNESARAGLSANAYGRTVMDLCGYQTVVPNATASGILDPLLSTAMAALAERDKTSEETAI
MGDFWALNVLVDIDMSASGEKALKNIWIVDWEACRYGSPAVDVATFAGDCYLISRIHNETATDAMRRNFLGTYVALA
KVDPMEVVIGMGTMWIMWTKYQEDIGEAEKRERVAKGVEYIHKGWERSREWLPVSLAQELIA

Seq. ID No: 439
>PSIK_2
MDLTTGDGVRVYLTAHMTLKVESTERLSGGYCNFVWRAKLKTPYEGQNSVIVKYAAPFTSWDQTIELGVERLAFECM
SLKMITSETPLLEENGLVAVPTVYHYDSTANVLVMQDIGSIATLHGFLRSNTPPTVPMAALIGAKLAAFIAGVHNWG
RNNLPAHTRLSANTVGRTAMKKLCYETIVPKAAKSGVVDPLLPMVVAALSEEAMTNDETLVMGDFWTANVLIDVQES
HTGEQVLKKLWVIDWESCRYGNPATDIASFAGDSYLVSRFQDHGLGEALRHSFLETYAALAKVDPLRVALGLGAHWI
MWTDDLGQGGEAETRECVDKGLEYIQRAWDQSAEWVSLSLAKELVVL

Seq. ID No: 440
>PSIK_3
MANENPDLLTVAGVLRFLAPTPFASDEVHPLSGGNCNFVYRIHLRTPYNNISTLVLKHAEPYVAASAHRMPLAVERQ
NTEVTAMNAVKAILSSDAVVIVPTIHHFDDVAHVVIMDDCGVGAVTLKQLMLKNPPPVSVAKALGAGLGEFLSRLHV
WGRDPQTSNHVSFDQNQQGRTISGYVTYGRLVSTLTGKDNIPALSDPPLDIAQSKLDTISALSSEKIHAINTSHQTL

TMGDFWPGNIMVRLNPAGDSLERAYVLDWEVAKPGVAGLDIGQFCAEMHSLRRFSPACDASATTVLDAFLKTYRDAA
GVDVGVAKDAMVHVGAHLVAWTPRVPWGSKERTREVVEEGVGYLVEGYAATQEWLRGSLVGRLV

Seq. ID No: 441
>PSIK_4
MEIEWCDLDLTSESRPPTHKYTYFATALMPFDLTTRDGVRMYLTAYLALDVMSVERLSGGYCNFSWRAKLESPYEGQ
ISIVVKHAAPFTSWDRNTELGVERLAFEYKALKILNSEPSVIAKNSLVAVPAVYHYDPTANALIMQDVGSIPTLHAL
LRNNALPPVPMAEKISNELAAFIAGIHNWGRNNQEARANLSQNLVGRTAIRKLCYETLVPKAEKSGVDDPLLQQVAA
ALSEEVMNSEETLVMGDFWTANVMVDIQETGAGVRSLRKIWVIDWEGCRYGSPAADIASFAGDSYLVARFHHHDLGE
TLRHSFLETYAGLAKVDPFRVALGSGAHWIMWTDDLSEQEEGEIRECVDKGVEYIHRAWEQSTKWISLSLAKELVT

Seq. ID No: 442
>PSIM_1
MHFSYDFTMCNPPFYGDYAELTRLRESKLKGPFGGAHEGVSTELFTAGGEIHEFIASELNQGKTLRWVVGWTFHKDL
FDKKVIEPLHILIKLKPLSCNTLELLDSDHELATPYSKKRCLRRISFTPKGRPPTKQTISVMHPNEASWASLLNHLQ
ALDISVTVGNHFLTAEVKDPTWTRAWRRSSKVTSKITPFSFSGQFSDPPERLLVLQLLVDESQTSEDILLSFQSL

Seq. ID No: 443
>PSIM_2
MHPRNPYRTPPDYAALARSFPELKPYVSRNANGTVSVDYQDEAALRCLTRALLYRDFGLSVDLPKDRLCPTVPNRLN
YILWIEDILNVSSLSRLQSNSEATVRGLDIGTGASAIYPLLGCRVSPRWDFYATDIDAQSLAHARANITRNGLQGRI
NLVAADPKGSIFGPLESKHDTTFEFTMCNPPFYSSEEDIAQSAAVKKLASNAVCTGAAVEMITPGGEAAFVVRMVRE
SLALKMRCQWYTSMLGKMSSLTEIVGLLRENSIDNYAITEFVQGKTRRWAIAWSFGHVRLLDSLARLSGGPLQSLMP
TRNTCRRSFAVPRMVLHKHLILVLDGIEGTSQTPMSIPVGAGDADGLYGLQISASRDTWSRAARRKRQHGAMDISLD
NDEVGMKCLIKVLSVEEAREGAEAVVLECTWVYGHERALFESFWGHVCRKIGEANG

Seq. ID No: 444
>PSIM_3
MHARSIFNPNSAQFQARLTFSELSNEFPKLKPFLKYKRSRKQNEADPLSSQCTFIDFKDPVATRAYNEVLLKKYFDL
SLEFLPGSLCPAVPNRLNYTLWLEDVLNVFPGSMGANNQRDELRGLDIGTGSSCIYPLLICRTHPNWRMAGSDINPS
SIEIAKKNVQENRLLDRIQLFLTTDKRDSVLEGQIFQTHLFFNSKKCLQDEKPARFCYDFTMCNPPFYSDVEDLNNS
RQAKTTTILGGGHEGVSSELFTTGGELLFLSQMVEESFLYKDKVGLVSSYFFVLKCIVEILRLLGQHKIQETIASKL
IQGKTIRWVIGWTFHKDLFDLKHPSCNTLKVQDIHCTNEEAPASKKICLGKIPLNSKEHIPSKQLISSSQSSDIGWT
RLKSRLGDLRIEFSLEKLSLIGKVVYPTWTRAWRRNGKAKSKPTPFSFSAQQSDASETTIQLELLASEIEEPDENTL
VSFQSLCNHLRSYLKD

Seq. ID No: 445
>PSIM_4
MSDIDAQSLVYARANVARNALEGRIAVVTAEPEGSIFGPIEAEKEIQFDFTMCNPPFYASAEDIAQSAATKELGPNA
VCTGAAVEMITSGGEGAFVARMIDNYAITEFVQGQTRRWAIAWSFGHDRLPDSLARLSSGPLQSLMPTRNTCRRSFT
FARMNLLSRLEQVLNNIEGLSHSNMSPSEDRGSGGRPSSLLVSVARDTWSRAARRKKQRGSMDTSLDNDTSGLICSV
KVLFDEEGREGSEIASLECTWIHGRERALFESFWGHICRKVGEVSG

Seq. ID No: 446
>PSIM_5
EGLSPFLLMHPRNQYCKKKPDFADLAKSHPPLREHLKWKTEDYATIDFKSPSAQKELTRALLKQDFHLDVDMPVDKL
VPTVPQKLNYIHWIEDLLSGGRSDSIPRGEGIRGIDIGTGPACIYPLLATSLNKWTFVATDIDAVSLEYAVKNVSRN
DMEGRIRVKGVDPDTLLVGVVRDEQFDFCMCNPPFYGIDEDHHDNQRPPPPYSSCSAQAHEVRVQGGEVGFVSRMVE
ESLLLPSRVRWFTSMVGKKGSLKSLRALLRKREVPTVTTTEFVQGVTKRWAVAWSFTEQVPCIPSHSLPCTVPLLGS
TSSAEGRAYAEQWLERVLNHMEVTFTKKDQDGYTCTAERATWANQRRKRRLMQRPMMSPEAAAKRSCGGSDNTSEGV

PRNDSDTLVSAGHLSPKADSLERNASSDLAAQLSALTPPYHVTFWCGVQPSVPPSTNKAELELKMVLIDGGSGTQPL
QPIAQYMKNNWSATDSRPTSDRSSQ

Seq. ID No: 447
>PSIM_6
CIYPLLGATMNGWYFLATEVDDICFDYATKNVEQNNLSDLIKVVKVPQKTLLMDALKEETEIVYDFCMCNPPFFANQ
LEAKGVNSRNSRRPPPSSVNTGGVTEIMAEGGELEFVKRIIHDSLQLKKRLRWYSCMLGKKCSLAPLKEELRKQGVP
KVTHTEFCQGRTMRWALAWSFYDDVIVPSPPNKKRKLERARKPLSFTLPEAGLKELQSKALALGGTACSPVDRVAAL
LEKTLTDLRVLHKRVPCRKQEQSLFLTAVENTWIHGRQKRREQSRQLRELPRAPPCAGTSSQTTVATADSVKTPASQ
TQSASTQNSNSQDDSSQNKRASAQELAGQQPTDKAGSSASSDEISIKVLHNSTGEQKEVTENLSSEAVDMEFSTSTE
AVQETGSKEAPSAESEPPSKRPLSPGTVEQFLFKCLLNVMLEESDVMIEMHWVEGQNKDLMNQLCTYLKNTLLKSVA
KS

Seq. ID No: 448
>PSIM_7
MGSKKRRRRREERPTIHPKNKYSENPPDFALLASLYPSFEPFVFYSRDGRPRIDWTDFNATRELTRVLLLHDHGLNW
WIPDGQLCPTVPNRSNYIHWIEDLLSSNIIPTTSRNGDKVKGFDIGTGANCIYPLLGASLLGWSFVGSDMTDVALEW
AEKNVKSNPHISELIEIRKVDNSESTPSIQESLTGKSVQDESNMDMSGHMDEEAEPSSSSSFNLPAGAQSSYHGPPV
LVGVVRDGEQFDFCICNPPFFESMEEAGLNPKTSCGGTPEEMVCSGGERAFITRIIEDSVALKQTFRWYTSMVGRKS
NLKFLISKLRKVGVTIVKTTEFVQGQTCRWGLAWSFVPPARKIISPHVAEKKNLSFMLECTLINRSLYQMINVTQS

Seq. ID No: 449
>PSIM_8
MHPRNPYRQLLDFASLAEAYEPLKPHLKPTRSPTAGGLSYTIDFKNSESQRQLTKAILYRDFGLRIALPDHRLCPPV
PNSRLNYILWLQDIIKAHDEYMDRPASCICGLDIGTGASAIYLLLGCRVEPSFRFIGTELDDISFSYATQNVESNGL
SDRIHLIKTTSNDPILLPFDLNPAWSCDFTMCNPPFYESEEEMARSAQAKELAPNAVCTGAQVEMVTPGGELAFVSQ
IVKESLKYTTRCRWYTSMLGKLSSLTKLVGLLREYAISNYAITEFVQGQTRRWAVAWSFGETHLPDSVARISNPTLQ
PLLPERNTSRHVINISLPPFSTRTVKSKQSIKALSEVLSQIKDVTVQRLYQVEHLEPTEEEEEDKSLYRLLVYAKQN
MWSRSARRQRGRETGHKANDKGCAVGGPLTSIPATLDGLLCGIEIKAPLIKQEQQDVEMEFVFQWVHGQDRSMFESF
VNHVTRKMKCNIVLD

Seq. ID No: 450
>PSIM_9
MALNKSMHPRNRYKDKPPDFAFLASKYPEFKQHVDVGLSGKVGLNFKDPGAVRALTCTLLKEDFGLTIDIPLERLIP
TVPLRLNYIHWVEDLINFHDSDKTTVRRGIDIGTGASCIYPLLGATLNGWYFLATEVDDICYNYAKKNVEQNHLADL
IKVVKVPQKTLLMDALKEESGIIYDFCMCNPPFFANQMEAQGVNSRNPRRPPPSSVNTGGITEIMAEGGELEFVKRI
IHDSLQLKKRLRWYSCMLGKKCSLAPLKEELRIQGVPKVAHTEFYQGRTMRWALAWSFYDDVTIPNPPSKKRKLEKP
RKPMMFSVLETTVKMLMDKFDCSVDSEHVSVVTDCLKKILTDLKVQHKPVPCGNGEESLFLTAIENSWVHIRRKKRD
RMRQLRELPRAPDENFLLVQKDERQAEDEETTEKTVSSSEKSVSTSGIDEAAALPPNPEDSISESMGEDSRQLPEEV
KDTSALGQITDVDEHQNTMEASQPCSSNSAFLFKCLVNVKKEATNVLVEMHWVEGHNRDLMNQLCTYLRNQICKIAT
S

Seq. ID No: 451
>PsiHchimera_1
MFCRGLLSLMAIIIVYFIAQKRRRARLPPGPRGLPLIGNLHQAPKEAVWLTFHKWVKEYGNLVSVNFGGTEMVILNT
LETITDLLEKRGSIYSGRLESTMVNELMGWEFDLGFITYGDRWREERRMFAKEFSEKGIKQFRHAQVKAAHQLVQQL
TKTPDRWAQHIRHQIAAMSLDIGYGIDLAEDDPWLEATHLANEGLAIASVPGKFWVDSFPSLKYLPAWFPGAVFKRK
AKVWREAADHMVDMPYETMRKLAPQGLTRPSYASARLQAMDLNGDLEHQEHVIKNTAAEVNVGGGDTTVSAMSAFIL
AMVKYPEVQRKVQAELDALTNNGQIPDYDEEDDSLPYLTACIKELFRWNQIAPLAIPHKLMKDDVYRGYLIPKNTLV
FANTWAVLNDPEVYPDPSVFRPERYLGPDGKPDNTVRDPRKAAFGYGRRNCPGIHLAQSTVWIAGATLLSAFNIERP
VDQNGKPIDIPADFTTGFFRHPVPFQCRFVPRTEQVSQSVSGP

Seq. ID No: 452
>PsiKchimera_1
MKTKFCTGGEAEPSPLGLLLSCGSGLVPRGSPQPPADEQPEPRTRRRAYLWCKEFLPGAWRGLREDEFHISVIRGGL
SNMLFQCSLPDTTATLGDEPRKVLLRLYGAILQMRSCNKEGSEQAQKENEFQGAEAMVLESVMFAILAERSLGPKLY
GIFPQGRLEQFIKMKTLLDYVTAKPPLATDIARLVGTEIGGFVARLHNIGRERRDDPEFKFFSGNIVGRTTSDQLYQ
TIIPNAAKYGVDDPLLPTVVKDLVDDVMHSEETLVMADLWSGNILLQLEEGNPSKLQKIYILDWELCKYGPASLDLG
YFLGDCYLISRFQDEQVGTTMRQAYLQSYARTSKHSINYAKVTAGIAAHIVMWTDFMQWGSEEERINFVKKGVAAFH
DARGNNDNGEITSTLLKESSTA

Seq. ID No: 453
>PsiKchimera_2
MAFDLKTEDGLITYLTKHLSLDVDTSGVKRLSGGFVNVTWRIKLNAPYQGHTSIILKHAQPHMSTDEDFKIGVERSV
YEYQAIKLMMANQEVLGGGDSRVSVPEGFHYDVENNALIMQDVGTMKTLLDYATAKPPLSTEIASLVGTEIGAFIAR
LHNLGRKRRDQPAFKFFSGNIVGRTTADQLYQTIIPNAAKYGINDPLLPTVVKDLVEEVMNSEETLIMADLWSGNIL
LQLEEGNPSELKKIWLVDWELCKYGPASLDMGYFLGDCYLIARFQDELVGTTMRKAYLKSYARTASDTINYSKVTAS
IGAHLVMWTDFMKWGNDEEREE

Seq. ID No: 454
>PsiMchimera_1
MDSAGNIYRHKVDFTALALQDPAFKETLSAKGRLDFSNPDAVRQLTVSLLRRDFGLEVELPDDRLCPPVPNRLNYIL
WLQDLIDCTGDDYHEGFNADRDVVGLDIGTGSSAIYPMLACARFKAWSMVGTEVERKCIDTARLNVVANNLQDRLSI
LETSIDGPILVPIFEATEEYEYEFTMCNPPFYDGAADMQTSDAAKGFGFGVGAPHSGTVIEMSTEGGESAFVAQMVR
ESLKLRTRCRWYTSNLGKLKSLKEIVGLLKELEISNYAINEYVQGSTRRYAVAWSFTDIQLPEELSRPSNPELSSLF

Seq. ID No: 455
>PsiMchimera_2
MSATTNIYKEDIDFITLGREDSDFGKLLNSNGQLDFSDPKSVQQLTKSLLKRDFGLKLILPDDRLCPPVPNRLNYVL
WIEDIFNYTNKTLGLSDDRPIKGVDIGTGASAIYPMLACARFKAWSMVGTEVERKCIDTARLNVVANNLQDRLSILE
TSIDGPILVPIFEATEEYEYEFTMCNPPFYDGAADMQTSDAAKGFGFGVGAPHSGTVIEMSTEGGESAFVAQMVRES
LKLRTRCRWYTSNLGKLKSLKEIVGLLKELEISNYAINEYVQGSTRRYAVAWSFTDIQLPEELSRPSNPELSSLF

Seq. ID No: 456
>PsiMchimera_3
MAQNSTIYEDEVDFATLALQDSEFAKILKSNGQLDFSNPESVQQLTKSLLKRDFKLKLSLPPDRLCPPVPNRLNYII
WIQNLLDTTSDSYNDKYDPEREVLGLDIGTGASAIYPMLACARFKAWSMVGTEVERKCIDTARLNVVANNLQDRLSI
LETSIDGPILVPIFEATEEYEYEFTMCNPPFYDGAADMQTSDAAKGFGFGVGAPHSGTVIEMSTEGGESAFVAQMVR
ESLKLRTRCRWYTSNLGKLKSLKEIVGLLKELEISNYAINEYVQGSTRRYAVAWSFTDIQLPEELSRPSNPELSSLF

Seq. ID No: 457
>SAM2
MSKSKTFLFTSESVGEGHPDKICDQVSDAILDACLEQDPFSKVACETAAKTGMIMVFGEITTKARLDYQQIVRDTIK
KIGYDDSAKGFDYKTCNVLVAIEQQSPDIAQGLHYEKSLEDLGAGDQGIMFGYATDETPEGLPLTILLAHKLNMAMA
DARRDGSLPWLRPDTKTQVTVEYEDDNGRWVPKRIDTVVISAQHADEISTADLRTQLQKDIVEKVIPKDMLDENTKY
FIQPSGRFVIGGPQGDAGLTGRKIIVDAYGGASSVGGGAFSGKDYSKVDRSAAYAARWVAKSLVAAGLCKRVQVQFS
YAIGIAEPLSLHVDTYGTATKSDDEIIEIIKKNFDLRPGVLVKELDLARPIYLPTASYGHFTNQEYSWEKPKKLEF

Seq. ID No: 458
>SAM3
MDILKRGNESDKFTKIETESTTIPNDSDRSGSLIRRMKDSFKQSNLHVIPEDLENSEQTEQEKIQWKLASQPYQKVL
SQRHLTMIAIGGTLGTGLFIGLGYSLASGPAALLIGFLLVGTSMFCVVQSAAELSCQFPVSGSYATHVSRFIDESVG

FTVATNYALAWLISFPSELIGCALTISYWNQTVNPAVWVAIFYVFIMVLNLFGVRGFAETEFALSIIKVIAIFIFII
IGIVLIAGGGPNSTGYIGAKYWHDPGAFAKPVFKNLCNTFVSAAFSFGGSELVLLTSTESKNISAISRAAKGTFWRI
AIFYITTVVIIGCLVPYNDPRLLSGSNSEDVSASPFVIALSNTGSMGAKVSNFMNVVILVAVVSVCNSCVYASSRLI
QALGASGQLPSVCSYMDRKGRPLVGIGISGAFGLLGFLVASKKEDEVFTWLFALCSISSFFTWFCICMSQIRFRMAL
KAQGRSNDEIAYKSILGVYGGILGCVLNALLIAGEIYVSAAPVGSPSSAEAFFEYCLSIPIMIVVYFAHRFYRRDWK
HFYIKRSEIDLDTGCSVENLELFKAQKEAEEQLIASKPFYYKIYRFWC

Seq. ID No: 459
>SSO2
MSNANPYENNNPYAENYEMQEDLNNAPTGHSDGSDDFVAFMNKINSINANLSRYENIINQIDAQHKDLLTQVSEEQE
MELRRSLDDYISQATDLQYQLKADIKDAQRDGLHDSNKQAQAENCRQKFLKLIQDYRIIDSNYKEESKEQAKRQYTI
IQPEATDEEVEAAINDVNGQQIFSQALLNANRRGEAKTALAEVQARHQELLKLEKTMAELTQLFNDMEELVIEQQEN
VDVIDKNVEDAQQDVEQGVGHTNKAVKSARKARKNKIRCLIICFIIFAIVVVVVVVPSVVETRK

Seq. ID No: 460
>T4H-CPR_1
SSSSDVFVLGLGVVLAALYIFRDQLFAASKPKVAPVSTTKPANGSANPRDFIAKMKQGKKRIVIFYGSQTGTAEEYA
IRLAKEAKQKFGLASLVCDPEEYDFEKLDQLPEDSIAFFVVATYGEGEPTDNAVQLLQNLQDESFEFSSGERKLSGL
KYVVFGLGNKTYEHYNLIGRTVDAQLAKMGAIRIGERGEGDDDKSMEEDYLEWKDGMWEAFATAMGVEEGQGGDSAD
FVVSELESHPPEKVYQGEFSARALTKTKGIHDAKNPFAAPIAVARELFQSVVDRNCVHVEFNIEGSGITYQHGDHVG
LWPLNPDVEVERLLCVLGLAEKRDAVISIESLDPALAKVPFPVPTTYGAVLRHYIDISAVAGRQILGTLSKFAPTPE
AEAFLRNLNTNKEEYHNVVANGCLKLGEILQIATGNDITVPPTTANTTKWPIPFDIIVSAIPRLQPRYYSISSSPKI
HPNTIHATVVVLKYENVPTEPIPRKWVYGVGSNFLLNLKYAVNKEPVPYITQNGEQRVGVPEYLIAGPRGSYKTESF
YKAPIHVRRSTFRLPTNPKSPVIMIGPGTGVAPFRGFVQERVALARRSIEKNGPDSLADWGRISLFYGCRRSDEDFL
YKDEWPQYEAELKGKFKLHCAFSRQNYKPDGSKIYVQDLIWEDREHIADAILNGKGYVYICGEAKSMSKQVEEVLAK
ILGEAKGGSGPVEGVAEVKLLKERSRLMLDVWS

Seq. ID No: 461
>T4H-CPR_2
SSSSDVLILGLGVALAALYLFRDQLFAASKPKAIPLTNKLAGLDNEGNPRDFIAKMKAGKKRLVIFYGSQTGTAEEY
AIRLAKEAKSKFGLTSLVCDPEEYDFENLDQLPEECAVFFVMATYGEGEPTDNAVQLMQNLADESFEFSGGERKLEG
LKYVIFALGNKTYEHYNLIGRKVDTLLTDMGGVRCGELGEGDDDKSMEEDYLEWKDAMWEDFARKMGVEEGQGGDSA
DFAVSELDTHVPEKVYLGELSARALTKTKGIHDAKNPYPAPIVASRELFQQGGDRNCVHVELSIEGSGITYQHGDHV
GVWPTNPEVEVNRLLCALGLWEKKDQVIGIESLDPALAKVPFPVPTTYATVLRNYIDISAVTGRQILGHLSKYAPAP
DVEEFLKGLSTNKEQYGATVANGCLKLGEVLQLAAGNDLKAIPTTENTTAWSIPFDVIVSAIPRLQPRYYSISSSPK
LNPTSIHVTAVVLKYQSVASEKLPAKWVYGVGSNFLLNLKYAANGEPAPFVTTNGSADPASVYYPTYAIEGPRGAYK
QETIYKSPIHVRRSTFRLPTNPKSPVIMIGPGTGVAPFRGFVQERVALARRTIEKNGADALADWGRISLFYGCRKST
EDFLYKEEWPQYTEELKGKFNMHSAFSREAPYKADGSKIYVQDLIWEDRANVSDAILNGKGYIYICGDAKSMAKQVE
DTLAKILGEAKGGTAEVEGAAEMKLLKERSRLMLDVWS

Seq. ID No: 462
>T4H-CPR_3
SSSSSGAGADSDENPRDFIAKMKAGKKRLVIFYGSQTGTAEEYAIRLAKEAKSKFGLTSLVCDPEEYDFENLDQLPE
DCAVFFVMATYGEGEPTDNAVQLMQNLQDESFEFSNGERKLEGLKYVVFALGNKTYEHYNLIGRKVDTILGEMGAVR
CGERGEGDDDKSMEEDYLEWKDAMWEDFARKMGVEEGQGGDSADFAVSELESHAPEKVYLGELSARALTKTKGIHDA
KNPYPAPIVESRELFQVGGDRNCVHVELGIEGSGITYQHGDHVGVWPTNPEVEVTRLLCALGLWEKKDQVIGIESLD
PALAKVPFPVPTTYITVLRNYIDISAVTGRQILGHLSKFAPSPDAEAFLKSLSTNKEQYGAIVANGCLKLGEVLQLA
AGNDLKAVPNAENTTKWTIPFDVIVSAIPRLQPRYYSISSSPKLNPTTIHVTAVVLKYESVASEKVPAKWVYGVGSN
FLLNLKYAANGDAAPFVTANGSADPASVYAPTYAIEGPRGAYKQETIYKSPIHVRRSTFRLPTNPKSPVIMIGPGTG
VAPFRGFVQERVALARRTIEKNGPDALADWGRITLFYGCRKSTEDFLYKDEWPQYTEELKGKFTMHSAFSREPPYKA

DGSKIYVQDLIWEDREKVADAILNGKGYVVYICGDAKSMAKQVEDTLAKILGESKGGSAEVEGAAEMKLLKERSRLML
DVWS

Seq. ID No: 463
>T4H-CPR_4
SSSSSSKLSDGDENPRDFIAKMKNGKKRLVIFYGSQTGTAEEYAIRLAKEAKSKFGLTSLVCDPEEYDFENLDQLPD
DCAAFFVVATYGEGEPTDNAVQLMQNLQDESFEFSGGERKLEGLKYVVFALGNKTYEHYNVIGRIVDTELAKMGAIR
CGERGEGDDDKSMEEDYLEWKDGMWEEFARIMGVEEGQGGDTPDFKVTELQSHPSEKVYLGELSARALTKTKGIHDA
KNPYPAPILKSRELFQKQGERNCVHLELGIDGSGITYQHGDHVGVWPSNPEVEVNRLLCALGLWDKRDHVIGIESLD
PALAKVPFPVPTTYSTVLRNYIDISAVAGRQILGNLARFSPSPDAEGFMRSLNTDKEQYGRIIANGCLKLGEVLQLA
AGNDIKAVPTLENTTAWPIPFDVIVSAIPRLQPRYFSISSSPKLHPTAIHVTAVVLKYQSVASDKVPPKWVYGVGSN
FILNLKYAACGETAPLIAQNGSADPAHTPFPLYAIEGPRGAYKQEMIYKSPIHVRRSTFRLPTNPKSPVIMVGPGTG
VAPFRGFVQERIALARRTIEKNGPDALADWGRISLFYGCRKSNEDFLYNEEWPQYIDELKGKFTLHTAFSREPPYKP
DGSKIYVQDLLWDDRSKVADAIINGKGYIYICGDAKSMAKSVEDVLAKILGEAKGGTMEVEGAAELKLLKERSRLML
DVWS

Seq. ID No: 464
>T4H_1
MKTRTSKHPPGPRGLPLIGNLLDMPASYEWLQYRKWSEEFKSDIIYLNILGTQIVVTNTLESTLDLLEKRSSKYSGR
HSFQLPNNCAMGWAWNLALMSYGDEWRAHRRLAARGFDAQAMPKFNHAFTRNTRGLLRRLLESPEAWNEHVRHEVGS
MIIEITYGLDVLSKNDPFIESADKGLATLALAVVPGAFLVDTLPILKHIPSWFPGAGFKRKAKEWKRYADEVLEAPY
KALKEEMASGAAKPSFVQRCLQDMDPNIDTTNQERVIKNTAAEMYVAGADTSASFIATFVLAMIQYPQVQRRAQAEL
DSVLGPDRLPTFGDMPSLPYLSAITKECFRWEVITPISIPHMLTEDDEYRGWFLPSGTVVIPNSWAIMNDPTVYPDP
SVFNPERFLKDGKIDLEVQDPQLAAFGYGRRICPGMRVANAFTWLSAGSILASFNISKPAAKDGTPIELDVKYRSSS
IRHPEAFDCLFKPRSENTRDMIVSAAA

Seq. ID No: 465
>T4H_2
MSKRSKHPPGPRGLPLIGNLLDMPTNDEWLQYRKWSQEFKSDIIYLNVCGTQIVVTNTLESTLDLLERRSSKYSGRM
GLEWAFILMPYGDEWRAHRRLAAKGFDAKAIPKFNPTFTRNAQDLLRRLLESPEAWHEHVRHQVGAMIIEVSYGLDV
LHKNDPFIESADKAAVTFAMAIKPGAFLVNTVPILKYVPSWFPGAGFQRKAKEWKRYNDAVLEAPFKALKEEITNGA
ARPSFAQQCLQNMDPNIDTAYQERVIKDTSAAMYGGGSDTSVSFLATFVLAMLQYPSVQRRAQVELDSVLGRDRLPT
FDDMPDLPYLAAVMKECHRWEIVLPLAIPHMLTADDEYRGWFLLSGTLVIPNSWAILNDPTVYPDPSTFNPERFLKD
GKIDPNVQDPELAAFGYGRRTCPGRRITNAFTWLSAGYILASFNIENAVGNDGMPIEPKVKYRSETIRHPDTFECVF
TPRSDDTRDMIGSAYT

Seq. ID No: 466
>T4H_3
MGRWPIIGNLLDMPQKSPWLTYAKWSEDCDSDIIHLNVLGTSIVVLSSLEAISTLLEGKAVDFSDRPKSTMMSELMG
WERGFAFMPYGQLWRSHRKAFHQEFSPQVAHRNHPKLIKATHNLLRLLLNTPQHWHGHIRRQAGASIMDIAYGIEVL
PENDPYLDIAEAAVKAFNDASVPGAFLVDSIPLLKHVPAWVPGAGFQLKAKEGRQALENLIDSPYNAMKKDLAGGKA
KSSYTSRSLAAMDATGVIEENETIIRETAAMVYLGGSDSTPSTTSVFILAMLAHPEVQRKAHAELDSVIGKAQLPTF
KDRGSLPYVTAVAKEVLRWEPVAPLAVPRKVRVDSEYKGYRIPKGSIVFQNSWALLHDEKTYPNPLAFNPERFLKDG
QLDPNVQDPDVVAFGYGRRSCPGKTMGYDSVWLNVASILAAFDIKKVANPDSTNVEPKFEPFGITV

Seq. ID No: 467
>T4H_4
MYLFKAYLRPSRRLPPGPRGWPLIGNLLDMPTSDEWVRYAQWVREFKSDVIHLEVCGTHIVILNSVESAVDLLEKRS
SLYSSRPPTPMMSDLMGWSWNTAMLPYNDEWRAQRRHFHGEFDGRAIGKHYPPIIRSTHDLLQRLLDTPEQWQSHIR
HLVGATILDVAYGIEVLPADDPYVRTAEAAFASVSEAMVPGAFLVDVLPILKHMPSWMPGAGFKRKAVAWKKLADAV
FDAPFAAMKQAMAAGTAKSSFGSRSLRDIDIKGNVQSQEFSIQAAAGTMYNAGSDTTVALLETFMLAMVLHPEVQTK

AQAEMDLVLGRSNLPTFADQESLSYLAAVMQEVFRWQVVAPFGVPHMSTADDEYRGYFIPEGTIVIPNAHQMLNDED
VYPEPSKFKPERFLKDGKLDLSVRSPLIAAFGFGRRICPGRALGENSAWLAAGSILTMFNLSKATDHNGVTIEPSGR
YTSGLVRHPETFKCQITPRSNEPRRELAGEIELITGRIQESEEA

Seq. ID No: 468
>T4H_CPR_chimera_1
SSSSSGSVAYFTKGTYWAVPKDPYASSYGAANGAKAGKTRDIIEKMEETGKNCVIFYGSQTGTAEDYASRLAKEGSQ
RFGLKTMVADLEEYDYENLDKWPEDKVAFFVLATYGEGEPTDNAVQLLQNLQDESFEFSSGERKLSGLKYVVFGLGN
KTYEHYNLIGRTVDAQLAKMGAIRIGERGEGDDDKSMEEDYLEWKDGMWEAFATAMGVEEGQGGDSADFVVSELESH
PPEKVYQGEFSARALTKTKGIHDAKNPFAAPIAVARELFQSVVDRNCVHVEFNIEGSGITYQHGDHVGLWPLNPDVE
VERLLCVLGLAEKRDAVISIESLDPALAKVPFPVPTTYGAVLRHYIDISAVAGRQILGTLSKFAPTPEAEAFLRNLN
TNKEEYHNVVANGCLKLGEILQIATGNDITVPPTTANTTKWPIPFDIIVSAIPRLQPRYYSISSSPKIHPNTIHATV
VVLKYENVPTEPIPRKWVYGVGSNFLLNLKYAVNKEPVPYITQNGEQRVGVPEYLIAGPRGSYKTESFYKAPIHVRR
STFRLPTNPKSPVIMIGPGTGVAPFRGFVQERVALARRSIEKNGPDSLADWGRISLFYGCRRSDEDFLYKDEWPQYE
AELKGKFKLHCAFSRQNYKPDGSKIYVQDLIWEDREHIADAILNGKGYVYICGEAKSMSKQVEEVLAKILGEAKGGS
GPVEGVAEVKLLKERSRLMLDVWS

Seq. ID No: 469
>T4H_CPR_chimera_2
SSSSSSGTIAYFTKGTYWGIVKDPYAPNYPPANGNKPAKTRNIVEKMDESNKNCVVFYGSQTGTAEDYASRLAKEGK
SRFGLETMVADLEDYDFDNLDTLGDDKVAIFVLATYGEGEPTDNAVQLLQNLQDESFEFSSGERKLSGLKYVVFGLG
NKTYEHYNLIGRTVDAQLAKMGAIRIGERGEGDDDKSMEEDYLEWKDGMWEAFATAMGVEEGQGGDSADFVVSELES
HPPEKVYQGEFSARALTKTKGIHDAKNPFAAPIAVARELFQSVVDRNCVHVEFNIEGSGITYQHGDHVGLWPLNPDV
EVERLLCVLGLAEKRDAVISIESLDPALAKVPFPVPTTYGAVLRHYIDISAVAGRQILGTLSKFAPTPEAEAFLRNL
NTNKEEYHNVVANGCLKLGEILQIATGNDITVPPTTANTTKWPIPFDIIVSAIPRLQPRYYSISSSPKIHPNTIHAT
VVVLKYENVPTEPIPRKWVYGVGSNFLLNLKYAVNKEPVPYITQNGEQRVGVPEYLIAGPRGSYKTESFYKAPIHVR
RSTFRLPTNPKSPVIMIGPGTGVAPFRGFVQERVALARRSIEKNGPDSLADWGRISLFYGCRRSDEDFLYKDEWPQY
EAELKGKFKLHCAFSRQNYKPDGSKIYVQDLIWEDREHIADAILNGKGYVYICGEAKSMSKQVEEVLAKILGEAKGG
SGPVEGVAEVKLLKERSRLMLDVWS

Seq. ID No: 470
>T5H-CPR_1
SSSSSSGGLLAFLYLFRGTLFASGKASDAGSKLAGGSDLDSSADAAANDFVTKLTSQNKRIAIFYGSQTGTAEEYAT
KIAKEAKARFGTSSLVCDPEEYEFEKLDQLPSDCVACFVMATYGEGEPTDNAVGLMEFLDGEDVQFSNGSSLDNLNY
VIFGLGNRTYEHYNAIARKLDARLESLGAKRIGERGEGDDDKSMEEDYLAWKDGMFEALASSLGFEEGGGGDVADFK
VREVADHPEDKVYRGELSARALLGTKGIHDAKNPYNAVVKEARELFVEGTADRTCVHVEFDIEGSGISYQHGDHIAV
WAHNPEQEVERALAVLGLLGKRDTVIDVESLDPTLAKVPFPVPTTYEAVFRHYLDICAHASRQTLNNFAKYAPTPEA
RAKLEKACGDKAAFQEAIGHRCLKTFEALQLIVGDDLGGDSVAKATAWEIPFDRVISDLPRVGPRFYSISSSPKMHP
KTVHITAVVLRYRPEAAGQDSPYVHGLATNFISAIKMAKNNEQPSGPDDPRFGTPGYDLAGPRGAYTKESLFRAPIH
IRRSNFRLPTSPKIPVIMVGPGTGVAPFRSFVQERVCSAQKTCDKVNQSPAEALQDWGNIWLFYGCRRSNEDFLYKD
EWPEYASKLGGKFQMETAVSREKFKPDGSKLYVQDLIWERRKELAQDILDKKAYIYICGEAKGMAHDVEEMFGRVLE
EAKGSAEAGRRELKLLKERSRLLLDVWS

Seq. ID No: 471
>T5H-CPR_2
SSSSSSLFSTTDVILFSLIVGVMTYWFLFRKKKEEVPEFTKIQTTTSSVKDRSFVEKMKKTGRNIIVFYGSQTGTAE
EFANRLSKDAHRYGMRGMAADPEEYDLADLSSLPEIEKALAIFCMATYGEGDPTDNAQDFYDWLQETDVDLSGVKYA
VFALGNKTYEHFNAMGKYVDKRLEQLGAQRIFDLGLGDDDGNLEEDFITWREQFWPAVCEHFGVEATGEESSIRQYE
LMVHTDMDMAKVYTGEMGRLKSYENQKPPFDAKNPFLAVVTTNRKLNQGTERHLMHLELDISDSKIRYESGDHVAVY
PANDSALVNQLGEILGADLDIIMSLNNLDEESNKKHPFPCPTSYRTALTYYLDITNPPRTNVLYELAQYASEPTEHE
QLRKMASSSGEGKELYLRWVLEARRHILAILQDYPSLRPPIDHLCELLPRLQARYYSIASSSKVHPNSVHICAVAVE

YETKTGRINKGVATSWLRAKEPAGENGGRALVPMYVRKSQFRLPFKATTPVIMVGPGTGVAPFIGFIQERAWLRQQG
KEVGETLLYYGCRRSDEDYLYREELAGFHKDGALTQLNVAFSREQPQKVYVQHLLKKDKEHLWKLIHEGGAHIYVCG
DARNMARDVQNTFYDIVAEQGAMEHAQAVDYVKKLMTKGRYSLDVWS

Seq. ID No: 472
>T5H-CPR_3
SSSSSSEAVAEEVSLFSMTDMILFSLIVGLLTYWFLFRKKKEEVPEFTKIQTLTSSVRESSFVEKMKKTGRNIIVFY
GSQTGTAEEFANRLSKDAHRYGMRGMSADPEEYDLADLSSLPEIDNALVVFCMATYGEGDPTDNAQDFYDWLQETDV
DLSGVKFAVFGLGNKTYEHFNAMGKYVDKRLEQLGAQRIFELGLGDDDGNLEEDFITWREQFWPAVCEHFGVEATGE
ESSIRQYELVVHTDIDAAKVYMGEMGRLKSYENQKPPFDAKNPFLAAVTTNRKLNQGTERHLMHLELDISDSKIRYE
SGDHVAVYPANDSALVNQLGKILGADLDVVMSLNNLDEESNKKHPFPCPTSYRTALTYYLDITNPPRTNVLYELAQY
ASEPSEQELLRKMASSSGEGKELYLSWVVEARRHILAILQDCPSLRPPIDHLCELLPRLQARYYSIASSSKVHPNSV
HICAVVVEYETKAGRINKGVATNWLRAKEPAGENGGRALVPMFVRKSQFRLPFKATTPVIMVGPGTGVAPFIGFIQE
RAWLRQQGKEVGETLLYYGCRRSDEDYLYREELAQFHRDGALTQLNVAFSREQSHKVYVQHLLKQDREHLWKLIEGG
AHIYVCGDARNMARDVQNTFYDIVAELGAMEHAQAVDYIKKLMTKGRYSLDVWS

Seq. ID No: 473
>T5H-CPR_4
SSSSSSAAAADGDGGQSRRLLALLATSLAVLVGCGVALLFRRSSSGAAPLARQAAAAKPLAAKKDQEPDPDDGRQRV
ALFFGTQTGTAEGFAKALAEEAKARYDKAVFKVLDLDDYAAEDEEYEEKLKKENIAFFFLATYGDGEPTDNAARFYK
WFSEGNERGEWLSNLQYGVFALGNRQYEHFNKVGKEVDQLLAEQGGKRIVPVGLGDDDQCIEDDFNAWKELLWPELD
KLLRVEDNSSAAQSPYTAAIPQYRIVLTKPEDATHINKSFSLSNGHVVYDSQHPCRANVAVRRELHTPASDRSCIHL
EFDIAGTSLTYETGDHVGVYAENSTETVEEAEKLLDYSPDTYFSIYADQEDGTPLFGGSLPPPFPPCTVRVALARYA
DLLNSPKKSVLLALAAHASDPKEAERLRHLASPAGKKEYSQWIIASQRSLLEVISEFPSAKPPLGVFFAAIAPRLQP
RYYSISSSPRMAPTRIHVTCSLVHGQSPTGRIHKGVCSTWMKNSTPSEEESEECSWAPIFVRQSNFKLPADPTVPII
MVGPGTGLAPFRGFLQERLALKETGVELGRAILFFGCRNRQMDFIYEDELNNFTESGALSELVVAFSREGPTKEYVQ
HKMAEKAADLWSIVSQGGYVYVCGDAKGMARDVHRALHTIVQEQVTQRTSNFGLWKFRLVSLN

Seq. ID No: 474
>T5H-CPR_5
SSSSSSAAAAGGDPLAALAATAAALVAGVVILAVWFRSGGGAPPKAAAPPPRPPPVKIEADADADDGRKRVTVFFGT
QTGTAEGFAKAMAEEARARYEKAVFKVVDLDDYAAEDEEYEEKLRKETIVLLFLATYGDGEPTDNAARFYKWFTEGK
EKEVWLKDLKYAVFGLGNRQYEHFNKVAKVVDELLEEQGGKRLVPVGLGDDDQCIEDDFTAWKEQVWPELDQLLRDE
DDTTGASTPYTAAIPEYRIVFIDKSDVSFQDKSWSLANGSGVIDIHHPVRSNVAVRKELHKPASDRSCIHLEFDISG
TGLVYETGDHVGVYSENAIETVEQAEKLLDLSPDTFFSVHADAEDGSPRKGGGSLAPPFPSPCTLRTALLRYADLLN
SPKKAALVALAAHASDLAEAERLRFLASPAGKDEYSQWVVASQRSLLEVMAAFPSAKPPLGVFFAAVAPRLQPRYYS
ISSSPKMAPSRIHVTCALVYGPTPTGRIHQGVCSTWMKNAIPSEYSEECSWAPIYVRQSNFKLPADPTTPIIMIGPG
TGLAPFRGFLQERLALKQSGVELGNSVLFFGCRNRNMDYIYEDELQNFIQEGALSELIVAFSREGPAKEYVQHKMTE
KATEIWNIVSQGGYIYVCGDAKGMARDVHRALHTIVQEQGSLDSSKTESYVKSLQMDGRYLRDVW

Seq. ID No: 475
>T5H-CPR_6
SSSSSSAAAYLFRDQIFRSSSPKVVVPAPSKLANGHGNPRNFVSKMKEGKKRIVIFYGSQTGTAEEYAIRIAKEAKT
KFGLTSLVCDPEEYDFENLDQVPEDCCVFFVMATYGEGEPTDNAVQLMQNLEDESFEFSNGSHRLDGLKYVVFALGN
KTYEHYNAIGRKVDTLLTDMGATKIGERGEGDDDKSMEEDYLEWKDGMWKAFSEAMGVEEGQGGDTPDFAVTELDSH
PPEKVYLGELSARALTRTKGIYDGKNPYPSAVKHSRELFQAGAERNCVHAELDIEGSGITYQHGDHVGVWPSNPDVE
VDRMLYVLGLYGKKDAVINIDSLDPALAKVPFPVPTTYATVLRHYIDICAVAGRQMLGVLSKFAPHPKAEAFLKSLN
SDKEEYSNIVTNGCFKLGEVLQLAAGDDIKLCPTPDNTTAWAIPFDIIVSSIPRLQPRFYSISSSPKLYPNAIHLTA
VVLKYDSIPNRLVESRFVYGVATNFLLNVKYAANGETAPFIAEPVISEPAHVSLPKYAIEGPRGAHIEDNIYKIPIH
VRRSTFRLPANPKIPVIMVGPGTGVAPFRGFVQERVALAKRSIEKNGPDALADWGSITLFYGCRKSNEDFLYKEEWP

QYAEELKGKFKMHCAFSREPPYKPDGSKIYVQDLIWEERETIAKAILEGKAYVYICGDAKAMSRAVEDTLARILGEA
KGGNAEVEGAAEMKILKERSRLLLDVWS

Seq. ID No: 476
>T5H-CPR_7
SSSSSSSLFSTTDMVLFSLIVGVLTYWFIFRKKKEEIPEFSKIQTTAPPVKESSFVEKMKKTGRNIIVFYGSQTGTA
EEFANRLSKDAHRYGMRGMSADPEEYDLADLSSLPEIDKSLVVFCMATYGEGDPTDNAQDFYDWLQETDVDLTGVKF
AVFGLGNKTYEHFNAMGKYVDQRLEQLGAQRIFELGLGDDDGNLEEDFITWREQFWPAVCEFFGVEATGEESSIRQY
ELVVHEDMDVAKVYTGEMGRLKSYENQKPPFDAKNPFLAAVTANRKLNQGTERHLMHLELDISDSKIRYESGDHVAV
YPANDSALVNQIGEILGADLDVIMSLNNLDEESNKKHPFPCPTTYRTALTYYLDITNPPRTNVLYELAQYASEPSEQ
EHLHKMASSSGEGKELYLSWVVEARRHILAILQDYPSLRPPIDHLCELLPRLQARYYSIASSSKVHPNSVHICAVAV
EYEAKSGRVNKGVATSWLRAKEPAGENGGRALVPMFVRKSQFRLPFKSTTPVIMVGPGTGIAPFMGFIQERAWLREQ
GKEVGETLLYYGCRRSDEDYLYREELARFHKDGALTQLNVAFSREQAHKVYVQHLLKRDREHLWKLIHEGGAHIYVC
GDARNMAKDVQNTFYDIVAEFGPMEHTQAVDYVKKLMTKGRYSLDVWS

Seq. ID No: 477
>T5H-CPR_8
SSSSSSGGSPMSDSVVVIITTSFAVIIGLLVFLWKRSSDRSKEVTPLVVPKSLSVKDEEDEAETLAGKTKVTIFYGT
QTGTAEGFAKALAEEIKARYEKAAVKVVDLDDYAMDDDQYEEKLKKETLTFFMVATYGDGEPTDNAARFYKWFTEEH
ERGVWLQQLTYGIFGLGNRQYEHFNKIAKVLDEQLNEQGAKRLIPVGLGDDDQCIEDDFTAWRELLWPELDNLLRDE
DDVNGASTPYTAAIPEYRVVIHDASATSCEDKSVLENGNTSIDIHHPCRVNVAVQKELHKPESDRSCIHLEFDISGT
GIIYETGDHVGVYAENFEENVEEAGKLLGQPLDLLFSIHADNEDGAPLGSSLAPPFPGPCTLRTALSHYADLLNPPR
KAALIALAAHASEPSEAERLKYLSSPEGKDEYSQWIVGSQRSLLEVMAEFPSARPPLGVFFAAIAPRLQPRYYSISS
SPRFALSRVHVTCALVYGPTPTGRIHKGVCSTWMKNAVPLEKSHDSSWAPVFIRTSNFKLPTDPSIPIIMVGPGTGL
APFRGFLQERMALKEDGAQLGPALLFFGCRNRRMDFIYEDELNYFVEQGVISELIVAFSREGPQKEYVQHKMMDKAA
QIWSLISERGYIYVCGDAKGMARDVHRTLHTIVQEQGNLDSSKTESMVKKLQMDGRYLRDVW

Seq. ID No: 478
>T5H-CPR_9
SSSSSSTSFHKLKRILHKHLQRSHSIGAECKPQRSNHEDLLAVMNRSSIKVSIFYGSQTGTAKKFAINLGHHLHNCG
VRNLVMDLRQTNMEILVNLSMLDNCVALFVVATYGEGEPTDSARQFMDNLKNSYQKLDNLRFAVFGLGNSMYTYFNA
VGKSIDRLLIQHGGKRLQTLTLGDEVNELESTFLNWRSHLTSLLIDFFDLNDHDRNYLNKQYKRMYSLKRFNWNVPL
VSHFVNMFINKAHVKETLPYENDNYFYASVAVNQELYHKSSRSCRHIELDVSASQLRYKTGDHIAIFASNPLDLVEK
IGDLLNIDLNEMISLDAVDPDSLTKHPFPCPCTYRHAFMHFVDITGPPGKSLLSACLDSVTNPEESQFVQLLISDSE
DGKKLYSKWILEDHRGLVDVLQDLKSFRPPADLLLELLNPLKPRLYSISSSSLVHTNRIHITASIVKYKTNSGRIFK
GLATNWLKSLQSTNTERHLKIPVAIHTSNFNLPRSRTIPVIMIASGTGLAPFRAFIQERLKVAHDKVGKTGQMVLFF
GCRHENKDFIYSDELKQACSTGLLEMFTAFSRDCLDGNKVYVQHKVLEMGNMVWKLLDECYAYIYVCGDAAGMVRDV
HLCLIELVVQRSNLTREAATSYVLNLRKQGRYRTDVWK

Seq. ID No: 479
>T5H-CPR_10
SSSSSSGGKIFDKLNSSLDSGDSTSPASLTALLMENKDLMMILTTSVAVLIGCAVVLMWRRSSTSARKVVELPKLVV
PKSVVEPEEIDDGKKKIAIFFGTQTGTAEGFAKALAEEAKARYEKAIFKVIDMDDYAADDEEYEEKLKKEKLAFFFL
ATYGDGEPTDNAARFYKWFEEGKERGDCFKNLQYGVFGLGNRQYEHFNKIAKVVDELLAEQGGQRLVPVGLGDDDQC
IEDDFAAWRELVWPELDKLLLDGDDATATTPYTAAVLEYRVVTYDKSNFDNDLTNTNGHANGHVIVDAQHPVRANVA
VRKELHTPASDRSCTHLEFDISCTGLTYETGDHVGVYCENFVETVEEAERLLNISPDTFFSIHTDKEDGTPLGGSSL
PSPFPPCTLRTALTRYADVLSSPKKSSLLALAACSSDPNEADRLRYLASPAGKEEYAQWIVASQRSLLEVMAEFPSA
KPSIGVFFASVAPRLQPRFYSISSSPRMAASRIHVTCALVYDKMPTGRIHKGVCSTWMKNAIPLEESLSCSTAPIFV
RQSNFKLPADNKVPIIMIGPGTGLAPFRGFLQERMALKEEGADLGPAVLFFGCRNRQMDYIYQDELDNFLEAGALSN
LVVAFSREGPNKEYVQHKMTQKADDIWNMISQGGYVYVCGDAKGMARDVHRTLHTIAQDQGSLDSSKAESFVKNLQT
TGRYLRDVW

Seq. ID No: 480
>T5H-CPR_11
SSSSSSGGDGAEGRALVATLAAAVLGAALFVLWRRAAAGKKRKREAAAAAVAEATEVKARAAKGGEDEKAADDGRKK
VTVFFGTQTGTAEGFAKALAEEAKARYDKAIFKVVDLDDYAAEDEEYEEKLKKEKLALFFVATYGDGEPTDNAARFY
KWFTEGNERGVWLNDFEYAVFGLGNRQYEHFNKVAKVVDEILTEQGGKRLVPVGLGDDDQCIEDDFNAWKEALWPEL
DRLLRDENDASTGTTYTAAIPEYRVEFIKPEEAAHLERNFSLANGHAVHDAQHPCQANVAVRRELHTPASDRSCTHL
EFDIAGTGLTYETGDHVGVYTENCPEVVEEAERLLGYSPDTFFTIHADKEDGTPLSGSSLAPPFPSPITVRNALARY
ADLLNSPKKTSLVALATYASDPAEADRLRFLASAAGKDEYAQWVVASQRSLLEVMAEFPSAKPPLGVFFAAVAPRLQ
PRYYSISSSPSMAATRIHVTCALVHETTPAGRVHKGVCSTWIKNAVPSEESKDCSWAPIFVRQSNFKLPADPSVPII
MIGPGTGLAPFRGFLQERLAQKESGAELGPSVFFFGCRNSKMDFIYEDELNNFLEQGALSELVLAFSRQGPTKEYVQ
HKMAQKASEIWDMISQGAYIYVCGDAKGMARDVHRVLHTIVQEQGSLDSSKAESFVKNLQMEGRYLRDVW

Seq. ID No: 481
>T5H-CPR_12
SSSSSSVRESSFIEKMKKTGKNIVVFYGSQTGTGEEFANRLAKDAHRYGMRGMAADPEEFEMTDLSRLTEIENALAV
FCMATYGEGDPTDNAQDFYDWLQETDIDLAGLKYAVFGLGNKTYEHFNAMGKYVDKRLEELGAERIFELGMGDDDGN
LEEDFITWREQFWPAVCEHFGVEATGEDSSIRQYELVVHTDENMNKVYTGEMGRLKSYETQKPPFDAKNPFLANATV
NRKLNEGGDRHFMHLELDITGSKIRYESGDHVAVYPANDAALVNKLGEILGADLETVISLNNLDEESNKKHPFPCPT
TYRTALTYYLDITNPPRTNVLYELAQYATDSKEQENLRKMASSAQDGKALYLSWVVESRRNILAILEDIPSLRPPLD
HLCELLPRLQARYYSIASSSKVHPNSIHVCAVLVEYETKTGRENKGVATNWLKNKQPSDNGHKSSVPMFVRKSQFRL
PFKPSTPVIMIGPGTGIAPFMGFIQEREWLKQQGKDVGETVLYYGCRHEHEDFLYINELKRYHKEGVLTQLNVAFSR
DQAHKVYVQHLLKNNKEMVWKLIHEDNAHIYVCGDARNMARDVQNIFYDIVEEYGKLDHAQAVDYIKKLMTKGRYSQ
DVWS

Seq. ID No: 482
>T5H_1
MLPIVDHLLDVLNLERTPFRTYAVTALLLLFVGIIARALLKMMLFIQEYSANSKRLRCFPEPPNRSWILGHLGLFAP
NEEGMTEFSKQVSKFTYYMKTWMGPVIPLISLIHPDTIKPVVAAPASIAPKDALFYGFLEPWLGDGLLLSRGEKWVR
HRRLLTPAFHFDILKHYVKIFNQSTDIMHAKWRRLCTKGPVFLDMFEHISLMTLDSLLKCTFSYDSDCQEKPSDYIA
AIYDLSELIVEREQCPPHHFDFIYRFSSNGRKFQRACRIVHEFTANVVQQRKKALQEKGAENWIRSKKGKTQDFIDI
LLLSKDEDGNTLSDQEMRDEVDTFMFEGHDTTASGLSWILYNLASHPEYQEKCREEVTQLLKGESTHLEWDDLSLLP
FTTMCIKESLRLHPPVTAVSRRCTEDIAMPDGKVIPKGNISLISIYGTHHNPAVWPNPEVYDPYRFDPSSTDERSSH
AFVPFSAGPRNCIGQNFAMAEMKVVLALTLLNFKVALDPNRVVRRKPELVLRAEGGLWLQVEALKSKS

Seq. ID No: 483
>T5H_2
MELLGLVSWLLLLLLTLVVICFLLYCGYIHYQHMKYDHIPGPPRESFLFGHGSAIWKVMRKNQLVYDLFLNWVETYG
PVIRINALHKVTIVSVSPESVKEVLMSPKYRKDWFYDHLHSLFGVRLMGNGLVTDRDNDHWYKQRRIMDPAFSRTYL
IGLLGPFNEKAEELMERLAEEADGRSHVVMHAMMSRVTLDVISKVAFGMEMNSLKDDGTPLPRAISLVMRALVEMRN
PFIRYSREKQAFIRDVQESARLLRKTGRECIERRQKAIQDGEEIPVDILTQILKGAALEGDCDMEDLLDNFVTFFIA
GQETTANQLAFTIMELARNPEILEKAQAEVDEVIGVKRDIEYDDLGKLQYLSQVLKESLRLYPTAPGTSRAIEEETI
IEGFRIPPKVPLMFNSYIMGRMQQFYPDPLTFNPDRFHPDAPKPYYSYFPFSLGPRSCIGQVFAQMEAKVIMAKLLQ
RFQFELVEGQSFGIMDTASLRPEGGVICRLTIRTNPGKAKKDD

Seq. ID No: 484
>T5H_3
MSRPQVPKGLKNPPGPWGWPLIGHMLTLGKNPHLALSRMSQQYGDVLQIRIGSTPVVVLSGLDTIRQALVRQGDDFK
GRPDLYTFTLISNGQSMSFSPDSGPVWAARRRLAQNGLKSFSIASDPASSTSCYLEEHVSKEAEVLISTLQELMAGP
GHFNPYRYVVVSVTNVICAICFGRRYDHNHQELLSLVNLNNNFGEVVGSGNPADFIPILRYLPNPSLNAFKDLNEKF
YSFMQKMVKEHYKTFEKGHIRDITDSLIEHCQEKQLDENANVQLSDEKIINIVLDLFGAGFDTVTTAISWSLMYLVM

NPRVQRKIQEELDTVIGRSRRPRLSDRSHLPYMEAFILETFRHSSLVPFTIPHSTTRDTSLKGFYIPKGRCVFVNQW
QINHDQKLWVNPSEFLPERFLTPDGAIDKVLSEKVIIFGMGKRKCIGETVARWEVFLFLAILLQRVEFSVPLGVKVD
MTPIYGLTMKHACCEHFQMQLRS

Seq. ID No: 485
>T5H_4
MPTPGGRLVAFLQRRGKLAGSLAVILLLILKRLRDAPRKVRWLRGPPLLGVVLKVFQGLREHALLDMYDRWHQRLGP
TFAYCAPGKMVVATIDPKNIEHVLKTKFDNYVKGHVFAEPFTDLLGDGIFNADGEMWHRQRKTASRMFTKRQFETHI
WKAIEANTAKVGRILERSEGTLDMFNLMNRFTLDTIGRIGFSKDIGSLEDPSSPFLRSFDRAQQILILRFWTNPAWK
VLRWLGVGWERELKEHLGRLDGYARGIVRELRQKAEAGQDDSFVGLFMKEEQAAPAARSPELQEKFMRDLVLNFLIA
GRDTTAQCISWTLFELTQHPAVAAKARQEVLDVCGEGPVTFEHLKSLQYVRAILDEGLRLHPSVPYDGKLCLGKDTL
PDGTVVPAGCIIQYIPYAQGRCKDIWGEDACSFRPERWLEMPRRPSSFAFAAFNAGPRECLGRRLAEAEMAALVSTV
VRDFDMRLEVEPSSVRYDAQLTLGMCGLPVSVRRCRRAYGVAEPLAGA

Seq. ID No: 486
>T5H_5
MLPLRHKMLTGEAEPCLVSKTAETDAEWTRDAFGMGQYTAGRCDHLLSWVVFLLLAPVLLIVWLPLSCICCASPVLL
VQRFAGWVLSGCLARTYLGVLLIRLCGKCDLILTGMHFIRTGSQRFWMDTLDPQDWAYHNETYGRNIILWANLRVGS
YKQVRDIVLNPARKRTRALDGWISGFARHYPNLPVFFNTGSNMHTTFRQIFFANFTKTDFVLRALEDEGAGLAKMAA
PILQRWLAGSFRESKSGEGNLYMVEPVAPLILFLLFEVEVESIPPELLTAFSDVVTVGASYFLLPPHSPYWLLSGKV
KAIALLKDFLLEHCNAARPESLKGRAVDWRSLAAQMPAFLPKDECRCPCSGTPAVDPVDAYLEVISVMVCVAGVTGT
TNGFTSVIRKFADVPVGPTKSRWPSAPVQWRPDADDMVRLYRRDPLGFILEALRLGTPVAGTHQVLEEELTCPFLHK
ETTFPKGTVVCANLNACHTDPEEWGSDALEFRPGRAARNRYLMWNGPFGEAAPRQCPGEQVAAHCIKVSIDAFLDMH
KPQ

Seq. ID No: 487
>T5H_6
MATSILSLSLMDLLYWGACLCVLSVLYKISALYLRQKNFERVFSAFPGPKRHWLYGNAHEFKQDGTDLDILNGYAKQ
FDCAFPLWLGNFFASLAIYHPDYIKAILSRQDPKDNFVYHFITPWIGKGLLVLSGQKWYQHRKLLTPGFHYDVLKPY
VGVMSDCVNVMLDKWERLVPDKKPVELFHYISLMTLDTIMKCAFSYQSNCQNDSENEYIKAVYELSYLVDHRTRCPP
YHNDFIFYWSPHGFRLRRALKTAHQHTEKVIKLRKESLKQETELEKIKQKRRLDFLDILLCARDENGQGLSDEDLRA
EVDTFMFEGHDTTASGVSWTFYCLAKNPEHQEKCREEIRQVLGDRRTVEWEDLSKLPYTTMCIKESMRLYPPVPEVA
RELKEPITFCDGRSVPKGSIVFLCIYAINRCPGIWEDPEVFDPLRFSPENSSTRHSHAFLPFSAGGRNCIGQNFAMN
EMKIATALTLQRFELQLETKREPVKRAQLVLRSMNGIYINLKKIHSDKTKII

Seq. ID No: 488
>T5H_7
MGLWTFMTGALILLLILVVLCFLLYCGYIYYMHMKYDHIPGPPRDSFFFGHSPTIMKLMRNNVIMYDTFLEWVKTYG
PVVRVNLSCSTIVFVISQEAVKEFLMSPKYTKDNFYECVETLFGVRYMGKGLLTDRDYEHWHKQRRIMDPAFSRNYL
IGLMGTFNEIAEDLVDILGDKADGKCQVGMHDMMGRVTLDIIAKVAFGMELNSLHDDQTPFTRAITTVMRGMVETRN
PLARYIPGKQALIRDIKESLKFLRKTGRECILQRRKAIQDGEDIPRDILTQILKGAETEGDCSLENLIDNFVTFFIA
GQETTANQLSFAVMELGRHPEILTRVQAEVDEVLGSKRDIEYEDLGKLQYLSQVLKETLRLYPIAPGTSRALEKEMV
IEGVRVPPGTTLMFNAYIMGRMEKYYHDPLVFNPDRFHPDAPKISYAYLPFSLGPRSCIGQVFAQMEAKVVMAKLLQ
RFEFELVEGQSFRILDTGTLRPLDGVICRLRPRAEHKSRK

Seq. ID No: 489
>T5H_8
MWTILLSTINITLATALMLSFIIIYLLYIQNSTKLPPGPTSWPLIGYTSCLGTDAFRKIQDLNKIYGDIVSFQVLGK
TIIILYNYDLIHEAANGNRSKVGRYTMTVNDLLAENSGISNYDTQKALEMRKAFVRLVHNNIKTTEEHEGNKLQPFI
SQNIINAQINKLIRQLRIRQGKPVNVLQLMRCTVWRIIWNLIFGKECQLTDKQISDTLDDISSNNLQNQLFQIRQLL
PRFCVNIFKHSQFARKLFEIEEIIYKYKTVRQLIDNNVGEMHNSDSLLGQLINDLKLNLTKNDISRLSFEFMAAGTD

TTSLTLTWACDYLARAPPKESLKLSSDLIDMIHRWASVVPLSLPHIVRESFKLKNYYIPKSSILIYNLYAVHNSQIK
KLINTEQNSDEIQESDKPIPFSLGSRSCPGARIANLLIEQILTAINQEFLIQNITQSPFETISPGNQESLTPFGITR
TPHKSMYIFVTKLNGNRRTSI

Seq. ID No: 490
>T5H_9
MSQLLSSLIELPTQTLVLATAVAVGAAALLVHAYLFDAVGKHGNLPPGPPVDSLFSGHRIPSTHPWRYLEKLTEEYG
DIFTLRIGRSPLFVLGRASSAHRILEKQSALSSSRPRLVLAGELLSNNKRILLMPYGDQWRLYRKAMHETLNDTVAK
QYEPIQEREARIATLHLGRLGQADGGGGDFQRVLHRYAASVIMQVTYDYQVQTLDDPLVRSVAQRGHALAMCIRPGA
SVLDRYPLLEHVPTWLNPWKQEGLRLRKLEQELYLGQVIKVRERMERGECAPCFVSKMTERQQELGLTDLDVAGMSG
SLFGAGSDTTASALSIFVMAVCRYPAVLARLHEELDRVVSRDRMPTFDDIPQMPYVRATVQEVLRWRPVSAGGFQHS
LTADVEYKGYVLPKGSTVVGPHWSISRDEHEYPEHDVFKPERFLQSGGAEANGTSAQDEVKGTWFAPARGSVAFGFG
RRVCPGLNVAMRSLHINIACMAWAFDIAQPDGRPERVDTFAFNSAANSHPLPFDATFTYRDPARKGVVEEENIATGE
LDRIAASRGAT

Seq. ID No: 491
>T5H_10
MLEALSSLATALWAALRPDTVLLGTLAFLLFVDFLKRRHPKNYPPGPPGLPFVGNLFQLDPEKVPLVLHQFVKKYGN
VFSLDFGTVPSVLITGLPLIKEVLVHQGQIFSNRPIVPLQEHIINNKGLIMSSGQLWKEQRRFALTTLRNFGLGKKS
LEERIQEEASYLIQTIREENGQPFDPHLTINNAVSNIICSITFGERFDYQDDQFQELLRMLDEILNLQTSMCCQLYN
VFPRIMNFLPGPHQALFSNMEKMKMFVARMIENHKRDWNPAEARDFIDAYLQEIEKHKGDATSSFQEENLIYNTLDL
FLAGTETTSTSLRWGLLFMALNPEIQEKVQAEIDRVLGQSQQPSMAARESMPYTNAVIHEVLRMGNIIPLNVPREVA
VDTTLAGYHLPKGTMVMTNLTALHRDPTEWATPDTFNPEHFLENGQFKKRESFLPFSIGKRMCLGEQLARTELFIFF
TSLLQKFTFRPPENEQLSLKFRVSLTLAPVSHRLCAVPRG

Seq. ID No: 492
>T5H_11
MKTPPQSSCPFHAVGRPPTPPRSSAGRWPPGPESGLTGWGLLKLMSRDLMGTLAGWQREFGDLVHVRTWPEHQVIVS
DPQLARELLVNQADALQRWERALTVYRRVHGHSVLIAEGQAWREKRQALQPDFTRKSVQAFSPSIVEAARRAFEQWP
ARHAAWPIESELTSVTMEVILRMMFSSGVGSEAQQAEEAVHTLMVASTEELWRPASLPDWVPWQRKRRRARLLMNGL
IERHLQARLAMPQDAWPEDLLSRLLRLHLQQPQSWPLQAVRDECKTAFLAGHETVATSLTWWAWCMASHPEIQERAR
EEALAALSGGGQADPAALQYVNQTLLETMRLYPAVPLLMSRRALKPVTLGDWTFPAKTVFMVPMQLMQHDERWFPEP
RSYRPERFGPDAARPQQGAYLPFGGGPRVCLGQHLAMAEMALVAAQLLLRYRLSAPEGAEPPRPVFHVSQRPSQPLT
LGIARI

Seq. ID No: 493
>T5H_12
MKLAGKRFRLPPGPSGAPIVGNWLQVGDDLNHRNLMGLAKRFGEVFLLRMGVRNLVVVSSPELAKEVLHTQGVEFGS
RTRNVVFDIFTGKGQDMVFTVYGDHWRKMRRIMTVPFFTNKVVAQNRVGWEEEARLVVEDLRADPAAATKGVVVRRR
LQLMMYNDMFRIMFDRRFETVADPLFNQLKALNAERSILSQSFDYNYGDFIPVLRPFLRRYLNRCTNLKTKRMKVFE
DHFVQQRKEALEKTGEIKCAMDHILEAERKGEINHDNVLYIVENINVAAIETTLWSIEWGLAELVNHPEIQQKLREE
IVAVLGPGTPVTEPDLERLPYLQSVVKETLRLRMAIPLLVPHMNLSDAKLAGYDIPAESKILVNAWFLANDPKRWVR
ADEFRPERFLEEEKSVEAHGNDFRFVPFGVGRRSCPGIILALPIIGITLGRLVQNFELLPPPGQDKIDTTEKPGQFS
NQILKHATIVCKPLEA

Seq. ID No: 494
>T5H_13
MHTDTPDTTADQPLRRIKDLPGPRPLPLIGNGHQIKPQRIHQHVERWSLQYGPLMRMYFGATPILVVADHEMVGAVL
RDRPDGFRRPSISATISNEMGGIPGLFLAEGADWRNQRRMVMAGFAPTAIKAYFPALVAVALRLRRRWQAAASARKA
IDLESDLKRYTVDIIAGLAFGSDVNTLESGEDVIQRHLDDILPAVARRSLALVPYWRYVKLPADRRLDRSVAVLRTA
VQDLIGQARQRMLDNPARRERPPNLLEAMIAAADQSGSGVTDLNVAGNVTNMLLAGEDTTANTISWMIYLLQRHPHT

LQKARDEVRRNAPDAARFTIEQLDSLDYLGACANEAMRLKPVAPYLPLEALRDTVIGDVAVPAGTMIWCVLRHDSVA
EKHFPDPLLFDPQRWLQADGKPNSDKRVTMPFGAGLRTCPGRYLALLEIKIAMAMLLGSFDIAGVDTPDGKEAQELM
GFVMSPVGLSLRLE

Seq. ID No: 495
>T5H_14
MLMKTLMASLQWLKESFQPFMLLFASIFLAVLLKFFFKEKSRKRSNLPPSPPKLPIIGNLHQLGNMPHLSLHNLAKK
YGSIIFLQLGEIPTVVVSSARLAKEVMKTHDLALSSRPQIFSAKHLFYNCTDVVFSPYGAYWRHIRKICILELLSVK
RVQSYSFVREEEVARLVRRVAEFYPGTTDLTKILGLYANDVLCRVAFGRDFSGGGEYDQHGFQKMLEEYQELLGGFS
LGDFFPSMEFVHSLTGMKSRLQDTFRRFDQLFDLFLTEHRDPKRETEEHKDLVDVLLDLQKNAYDEMPLTTDNIKAI
ILDMFAAGTDTTFITLDWGMTELIMNPEVMERAQAEVRSVVGDREVVLQSDLPQLHYIKAVIKEIFRLHPPAPVLVP
RESMEDVSIDGYNIPSKTRFFVNAWAIGRDPESWENPNAFEPERFMDSTIDFKGQHFELIPFGAGRRSCPAIAFGEA
TIELALAQLLHSFDWELPPGTTPKDLDMSEVFGITMHRIAHLIVIAKPRFPVGQNK

Seq. ID No: 496
>T5H_15
MPKQKKRLPPGPPTLPIIGNMHQLGELAHKSLSELSKKYGPIMLLKIGSKTIINISSAEAARQVLKVHDLDCCSRPV
SSTAGRLTYNFKDIVFAPYGDYWREMRKICALELLSVARVQSYRFIREEEVASLVNSISQSASSATPVDLSEKMLAL
TVNILCRTAFGKSFRGSGLDNGKLREVVHEAEVMFASFSATEFFPYVGWIIDRLSGRIRRLEKIFRGLDDFLQQAID
LHLKPKKTEQDHEDLIDVLLKIERDQQTNTGAPPFNKDNIKAILFDMFLGGSNTAAVTMLWAMAELARNPRAMKKAQ
DEVRNVVGNRGKVTESDITHLHYLKMTIKETFRLHPPAAILLPRQTMAEVKIGGYDIGPNSLLQVNAWALGRDPEYW
MNPEEFYPERFVDSSIDYKGQHFELLPFGSGRRGCPGMHMGTTTVELALANLLYCFDWKLPSGLKEEDINMDESTGP
GLTQKRTTLKLVPVKLF

Seq. ID No: 497
>T5H_16
MKLLLDRTRTNGYLPPSPPKLPIIGNLHQLGKMPHISLCERAQKLGPIMFLQLGEVPTVVISSAAMAKEVMKTHDLA
FSSRPQLYSAKWLFYNCTNIVFSPYGAYWRHVRKICILELLSTKRVQSYGFIRQEEVSRLLHRIADSCSKPINLSKL
LGLYANDVLCRAVLGRNFSEGGDYDMHGFQSMLKEYQELLGGFSIGDFFPSKEFVHLLTGHKRRLQNTFKRFDNFFQ
QVVREHLDPERNYEGEKDILDVLLDIQKNGSSEMPLTLDNVKAILLDMFAAGTDTSFIVLDWGMTELIMNPKVMKKA
QAEIRRVVGERQVVLENDLPQLHYLKAVIKEIFRLHPPVPVLVPRESIQDVTIEGYNIPAKTRVFINVWAIGRDPES
WKNPETFDPERFVGSTIDFKGQDFELLPFGAGRRGCPGITFGAVTVELALAQLLHSFDWKLPLGVEAKDLDLTEAFG
ISMPKTSDLIVVAKPCFA

Seq. ID No: 498
>T5H_17
MRMDGNSTTMFPLLITVIMLLASVLFYIFNRWTHRYSKSGILPPSPPKLPLLGHLHLLSDQPHVALSRLAQKYGPIM
YLELGQVPTVVVSSASLAREVLKTHDHVFCNRPQTIAAQYISFGCSDVTFSPYGPYWRQVRKICVTELLTLRRVNSF
QLIREEETNRLLTAVGAHSGSEVNLTKLFFNLANDTLCRAAFGTRFMSESTQLERQREGKRLEDILIETVKLLSGFY
VGEFFPRWGWINSVSGFKRRLERNLADLRSVGDEIIQEHIKKRGRGNEEEDFVDVLLRVQRQQDLQVPITDDNVTAL
VMDLFVAGTDTTSSTLEWTMTEMARHPEVMKKAQAEVRSMSPEGGTLDESHLRHLHYLKAVIKEALRLHPPIPLLLP
RESMDKCAIDGYEIPAKTRVLINNFALGRDPDSWDDPLRYNPARFMGGDEHKIDFKGEDFRFVPFGGGRRGCPGYSL
GLATVELTLARLLYHFDWKLPPGVEAEKIDLTEIFGLATRKKTPLLLIPTARKAPPHE

Seq. ID No: 499
>T5H_18
MELTMASTMSLALLVLSAAYVLVALRRSRSSSSKPRRLPPSPPGWPVIGHLHLMSGMPHHALAELARTMRAPLFRMR
LGSVPAVVISKPDLARAALTTNDAALASRPHLLSGQFLSFGCSDVTFAPAGPYHRMARRVVVSELLSARRVATYGAV
RVKELRRLLAHLTKNTSPAKPVDLSECFLNLANDVLCRVAFGRRFPHGEGDKLGAVLAEAQDLFAGFTIGDFFPELE
PVASTVTGLRRRLKKCLADLREACDVIVDEHISGNRQRIPGDRDEDFVDVLLRVQKSPDLEVPLTDDNLKALVLDMF
VAGTDTTFATLEWVMTELVRHPRILKKAQEEVRRVVGDSGRVEESHLGELHYMRAIIKETFRLHPAVPLLVPRESVA

PCTLGGYDIPARTRVFINTFAMGRDPEIWDNPLEYSPERFESAGGGGEIDLKDPDYKLLPFGGGRRGCPGYTFALAT
VQVSLASLLYHFEWALPAGVRAEDVNLDETFGLATRKKEPLFVAVRKSDAYEFKGEELSEV

Seq. ID No: 500
>T5H_19
MPLSDSTISLLLAVLPISGIIFALYNQYQIWLKSPIRGLPYPPGPPLLLGNANRAVQSRPWLTYTEWAKQYGDIIYV
NIYGEHTVILNNLEDVMELFEQRSRVYSSRQNNPYIELMGWQFNAGLLPYGDLWRRHRKLLQQCFRRKISTQYEPIQ
IAKTHNLLNDLLQTPSDFIEHIKRNSSAMIMSILYGQDISDEMSAQFVSVAEESVKALGKCLRPGTYLVSYIPMLRY
LPAWFPGAEFQRQAAEVKKLTTKMKDEPIDFVGKGLLHGTASASLVADLLENCYVQREYDVIKDVAATVFAAGADTS
VAALESFFLAMSLFPEAQKKAQAEMDRVIGNKRLPTTDDRPLLPYLEAVYRELMRWAPVVPLNAAHTTIADDIYKGY
YIPKGTAVYANTWALTRNEEKYPNPDIFNPDRFFTETGELNDDDTVLTFGFGRRICPGRHMASTTVWLTIASVLSNF
DIKGKGTNTKDQKFTSIGEMFTDNFISRPVPFECDIVPRKNAALLASK

Seq. ID No: 501
>T5H_20
MAFETTNGILLAASLFAGVVLYLQKRKRYTLPYPPGPKKHFLLGNLLDVPTTFAWKRYAEWGKTFDSDVLHLSVAGS
HFIILNSFKAANDLFEKRSSIYSSRAQMIMFSELIGWDWLMSGMVYGEPWRERRKAFQQYFHVGNAHLYEPVQMQAV
RKMLPRLLKEPEDFLSITRHALGSMALTLAYGLDIQEKNDPYLRVSEAAVKSIGEVAIPGAFLVDMIPALKYVPEFF
PGAGFKKKARIWRKVQENMREIPFAATLKNIASGSAKVSFTSTCLENLDESRDVDHQRTIIKDTAGNMFAAATDTTI
SAIHTFFVAMLCFPEVQKKAQQEIDRVLQGRLPEFSDEADLPYLSALVKETLRWEPSTPIGVPHYSSEDDVYNGYHI
PKGSLVIGNAWAMLHNEEDYPEPSLFKPERFIKDGKLNPNVRDPAEMAFGFGRRLCPGNHIAISALWLTAATVLATF
NITEAIDDDGRPIKPCVEYESALICHPLPFKCTIKPRSKECTMLIQAAADSY

Seq. ID No: 502
>T5H_21
MIIDSSNSEGNSEGQYTIDGPKAKGLRRMFRIFHLILQPTKYMESSVQRYGSMFQIGSEGASPLVYVGEPEVVKEIF
ALDGDQVVTGQGNGVLETMVGKHSILLLDGDPHRQQRKLLMPPFHGEQLRAYAHLICDITRQISAQWQPGQTIVARP
PIQNLTLGVILQAVFGVPSGERLSRLQQLMSTLLDSFAYPISASFLFFPALQKDLGEWSPWGKFIRLREEVRSLIYA
EIRDRRQQLERSAIEQDEKLGEKLGEKTDILTLLLQARDEDGGAMSDAELHDEIVTLLLAGHETTASAIVWMLYWIH
YLPEVQQKLRAELDALGPDPDPMAIAQLPYLTAVCQEALRIYPITPTTFIRRLREPMTLAGYRFKAGTALMPATYII
HQRPDLYPEPKQFRPERFLERQFAPHEFLPFGGGHRYCIGSALAMMELKLSIATLLADFELALLHSRPLLPARRGLT
MAPPAAMKLRIKARKTNKA

Seq. ID No: 503
>T5H_22
MPAPKTAPSTLPLPPGRLGLPWIGETLSFLRDPNFATKRQAQYGSLFKSRIIGQPTVFFCGPEANAFLLSSHADCFS
WRDGWPGTFQELLGESLFLQEGETHLRNRRLLMPAFHGKALASYFSTMVALSDSYLARWEKKQQLTWFLEFKKFTFE
VASVLLVGSAPGHDETDNTIGTAESAETEAQIAQLASWFADLTNGLFTLPIRWGPTTYRKALRGRDRLLSYIEQEIT
KRRQLLARLQTDPTAALPTDVLTLLLQTEDDEGNRLSEAEIKVQTLLMLFAGHETTTSMLTSLVMSLAQNPDVLAKA
RAEQQAFPAESALTFEQIQQMPYLDQILKEVERQYPPVGGGFRRVIKPFNFNGYHVPAGWLALYRIDAAHKDERCYT
NPSDFDPDRFSPERAEQKRYDYSLVGFGGGPRVCLGMAFAKLEMKIMAAQLLRRYHWQLDADQDLTMNPVPSLRPAD
GLKVRFSKLSFTA

Seq. ID No: 504
>T5H_23
MLDMPSVKPWLTFSDWASKFGDISHLEIFGQHIVVLNSAKTAVEMLDRKSSIYSDRPVLPMGGELVGWRNTLVLLPY
GDNFREYRRNFHRVIGSRAAMSVYHAIEEEETHKFLQRVLTKPADLSAHVRTTAGAIILRISHGYHIQEDGDPFVSL
ADTAVDQFSRSTATGAFMVDLIPALAYVPEWFPGASFQRKAREWRATLHEMVNQPYKFVQDQMAAGIAPKSFTSNLL
EGRTLTEEEEHIIKWSGASLYSGGADTTVSAIYGFFLAMTLYPEAQKKAQAEIDAVVGSDRLPTFADRESLPYAEAL
VKEVLRWCPVVPIVVPHRVTADDIHNGYYIPKGTLVLANAWYMLRDPSIYPDPMNFNPDRFLPSGGKEPPTDPRDIC

FGFGRRICPGMHLADASVWLSAVMSLAVFNVSKVVENGVEITPEVDPSSGTISHPKPFKCSIKPRSAKALELIQQTP
HY

Seq. ID No: 505
>T5H_24
MHLPPGPRPLPFLGNLLQMNRRGLLRSFMQLQEKYGDVFTVHLGPRPVVILCGTDTIREALVDQAEAFSGRGTVAVL
HPVVQGYGVIFANGERWKILRRFSLVTMRNFGMGKRSVEERIKEEAQCLVEELKKYKGALLNPTSIFQSIAANIICS
IVFGERFDYKDHQFLRLLDLIYQTFSLMGSLSSQVFELFSGFLKYFPGVHKQISKNLQEILNYIDHSVEKHRATLDP
NTPRDFIDTYLLHMEKEKSNHHTEFHHQNLVISVLSLFFAGTETTSTTLRYSFLIMLKYPHVAEKVQKEIDQVISSH
RLPTLDDRIKMPYTDAVIHEIQRFADLAPIGLPHRVTKDTMFRGYLLPKNTEVYPILSSALHDPRYFDHPDTFNPEH
FLDANGTLKKSEAFLPFSTGKRTCLGEGIARNELFIFFTALLQNFSLASPVAPEDIDLTPINSGAGKIPSPYQINFL
SRCVG

Seq. ID No: 506
>T5H_25
MYLIPDFSKETWILLIILLALLAYYGIWPYRLFKKYGIPGPKPLPFFGTFLENRNGVFEFDMECFKKFGKVWGFYDG
RQPVLAIMDPVIIKAILVKECYTVFTNRRNFGLNGPLNSAVSIAADDQWKRIRTVLSPTFTSGKLKQMFPIIKQYGD
LLVKNIQKKVDNKEFIDMKNIFGSYSMDIVLSTSFSVNVDSLNNPNDPFVTNGRNLFTFSFLNPLFLTTLLCPFLIP
ILDKLNFCFLPISVLNFFQDAITSIKKNRQKGIHKDRVDFLQLMVDAQANDSKGGADHGYKELTDTEIMAQGLIFII
AGYETTSTTLMFLAYHLATHPDVQTKLQEEIDIILPNKAPPTYEALMQMEYLDMVLYENLRLYPAAGRIERVCKATT
EINGVTIPKGVVTVIPAFVLHRDPELWPEPDEFRPERFSKENRETQDPYTFLPFGAGPRNCIGMRFALINMKSVITL
LLQNFSFRTCKDTPIPLQIDTRGFLKTTKPVILNLVPREAQKTEK

Seq. ID No: 507
>T5H_26
MYDTFLEWIEKYGPVVRVNSSHSTFVIVISPEGVKEFLMSPKYTKDNFYERIETLFGARFLGKGLVTDRDYDHWHKQ
RRMMDPAFSRTYLIGLMGTFNETAEDLMDVLGDKADGKCQVGMHDMLSRVTLDVIAKAAFGMELNSLHDDQTPFTRA
ISTVMKGMVETRNPLARYIPGKQAFIREVKESIKLLRETGRECILQRRKEIQDGEDIPMDILTQILKGAEIEDGCSL
EDLIDNFVTFFVAGQETTANQLSFAVMELARNPEILTRVQTEVDEVLGSKRDIEYEDLGKLQYLSQVLKETLRLYPI
APGTSRALEKETVIEGVRVPPGTTLMFNSYIMGRMEKYYHDPFIFNPDRFHPDAPKPSCAYFPFSLGPRSCIGQVFA
RMEAKVVMAKLLQRFEFELVEGQSFRIMDTGTLRPMDGVICRLRPRAERKSRK

Seq. ID No: 508
>T5H_27
MAARPKPATPPSPPALPVIGHLHLLTDMPHHTFADLSNSLGPLIYLRLGQVPTIVIHSAHLAKLVLRTHDHAFANRP
QLISAQYLSFGCSDVTFSSYGAYWRQARKICVTELLSAKRVHSFRLVRKEEVDRLLDAVLTSSGKEVDMSQMLFCLA
NDVLCKVAFGRRFMAEKDGKGKNLGSVLMETQALFAGFCLGDFFPKWEWVNSMSGYRKRLLKNLKDLKEVCDEIIEE
HLKKKKKKNGTENADDDDDYNEKEDFVDVLLRVQKREDLEVPITDDNLKALVLDMFVAGTDTSSATLEWVFTELARH
PRVMKKAQEEVRMIASGNGKVDESDLQHLHYMKAVIKETMRLHPPVPLLVPRESMEKCALDGYEIPAKTRVLINTYA
IGRDPKSWENPLDYDPERFMEDDIDFKDQDFRFLPFGGGRRGCPGYSFGLATIEITLARLLYHFDWALPHGVEADDV
DLSEVFGLATRKKTALVLVPTANKDFQFRGHDF

Seq. ID No: 509
>T5H_28
MGKNKVPPGPIGLPFIGNLHQFDTLAPHIYFWELSKKYGKIFSFKLTSNVPIIVVSSAKLAKEVLKTQDLVFCSRPS
LVGQQKLSYNGHDIGFAPYNDYWREMRKICVLHLFSLKKVQLFSPIREDEVSRMIKKIYQQAVNSQVTNLSNLMISL
NSTIICRVAFGVRFDEEAHERKRFNYILAEAQAMFAGFFMSDFFPSLSWIDKLTGMIDRLEKNFKDLDEFYEELIEQ
HYNPNRPKSMEGDFIDILLQLKKDQLTPIDLSLEDIKGILMNVLLAGSDTSSSVIIWAMTILIKNPKAMKKVQEEIR
NLIGNKGIVNEDDIQNMHYLKAVIKETLRLFPPAPLLIPRESMKISTLEGYEFQPRTIVYVNAWAIARDPEIWENPE
EFMPERFLNSNIDFKGQDYELIPFGAGRRGCPGLALGVASVELALSNLLYAFDWELPYGLKKEDIDINGKPGITVNK
KNDLCLIPKKYF

Seq. ID No: 510
>T5H_29
MKLTGKRYRLPPGPAGAPVVGNWLQVGDDLNHRNLMSLAKRFGDIFLLRMGVRNLVVVSTPELAKEVLHTQGVEFGS
RTRNVVFDIFTGKGQDMVFTVYGDHWRKMRRIMTVPFFTNKVVAQNRVGWEEEARLVVEDVRKDPRAAAEGVVIRRR
LQLMMYNDMFRIMFDTRFESEQDPLFNKLKALNAERSRLSQSFEYNYGDFIPVLRPFLRGYLNRCHDLKTRRMKVFE
DNFVQERKKVMAQTGEIRCAMDHILEAERKGEINHDNVLYIVENINVAAIETTLWSIEWGIAELVNHPAIQSKLREE
MDSVLGAGVPVTEPDLERLPYLQAIVKETLRLRMAIPLLVPHMNLNDGKLAGYDIPAESKILVNAWFLANDPKRWVR
PDEFRPERFLEEEKTVEAHGNDFRFVPFGVGRRSCPGIILALPIIGITLGRLVQNFQLLPPPGQDKIDTTEKPGQFS
NQIAKHATIVCKPLEA

Seq. ID No: 511
>T5H_30
MKPRGAKYPNSLPCLPFIGSLLHLASHLAPHILFNKLQEKYGSLYSFKMGSHYIVIVNHHEHAKEVLLKKGKTFGGR
PRAVTTDLLTRNAKDIAFADYSPTWKFHRKLVHAALSMFGEGTVAIEKIISREAASLCQTLITFQGSPLDMAPELTR
AVTNVVCALCFNARYKRCDPEFEEMLAYSKGIVDTVAKDSLVDIFPWLQIFPNKDLEILKRSVAIRDKLLQKKLKEH
KEAFCGEEVNDLLDALLKAKLSMENNNSNISQEVGLTDDHLLMTVGDIFGAGVETTTTVLKWAVAYLLHYPKVQAKI
QEELDVKVGFGRHPVLSDRRILPYLDATISEVLRIRPVAPLLIPHVALHESSIGEYTIPQDARVVINLWSLHHDPNE
WENPEEFIPDRFLDENGNHLYTPSQSYLPFGAGIRVCLGEALAKMEIFLFLSWILQRFTLEVPAGDSLPDLDGKFGV
VLQVKKFRVTAKLREVWKNIDLTT

Seq. ID No: 512
>TAT2
MTEDFISSVKRSNEELKERKSNFGFVEYKSKQLTSSSSHNSNSSHHDDDNQHGKRNIFQRCVDSFKSPLDGSFDTSN
LKRTLKPRHLIMIAIGGSIGTGLFVGSGKAIAEGGPLGVVIGWAIAGSQIIGTIHGLGEITVRFPVVGAFANYGTRF
LDPSISFVVSTIYVLQWFFVLPLEIIAAAMTVQYWNSSIDPVIWVAIFYAVIVSINLFGVRGFGEAEFAFSTIKAIT
VCGFIILCVVLICGGGPDHEFIGAKYWHDPGCLANGFPGVLSVLVVASYSLGGIEMTCLASGETDPKGLPSAIKQVF
WRILFFFLISLTLVGFLVPYTNQNLLGGSSVDNSPFVIAIKLHHIKALPSIVNAVILISVLSVGNSCIFASSRTLCS
MAHQGLIPWWFGYIDRAGRPLVGIMANSLFGLLAFLVKSGSMSEVFNWLMAIAGLATCIVWLSINLSHIRFRLAMKA
QGKSLDELEFVSAVGIWGSAYSALINCLILIAQFYCSLWPIGGWTSGKERAKIFFQNYLCALIMLFIFIVHKIYYKC
QTGKWWGVKALKDIDLETDRKDIDIEIVKQEIAEKKMYLDSRPWYVRQFHFWC

Seq. ID No: 513
>TMO_1
MSTLADQALHNNNVGPIIRAGDLVEPVIETAEIDNPGKEITVEDRRAYVRIAAEGELILTRKTLEEQLGRPFNMQEL
EINLASFAGQIQADEDQIRFYFDKTMGGGSGEGRGSLLTCGDVEENPGPMFNIQSDDLLHHFEADSNDTLLSAALRA
ELVFPYECNSGGCGACKIELLEGEVSNLWPDAPGLAARELRKNRFLACQCKPLSDLKIKVINRAEGRASHPPKRFST
RVVSKRFLSDEMFELRLEAEQKVVFSPGQYFMVDVPELGTRAYSAANPVDGNTLTLIVKAVPNGKVSCALANETIET
LQLDGPYGLSVLKTADETQSVFIAGGSGIAPMVSMVNTLIAQGYEKPITVFYGSRLEAELEAAETLFGWKENLKLIN
VSSSVVGNSEKKYPTGYVHEIIPEYMEGLLGAEFYLCGPPQMINSVQKLLMIENKVPFEAIHFDRFF

Seq. ID No: 514
>TMO_2
MAMHPRKDWYELTRATNWTPSYVTEEQLFPERMSGHMGIPLEKWESYDEPYKTSYPEYVSIQREKDAGAYSVKAALE
RAKIYENSDPGWISTLKSHYGAIAVGEYAAVTGEGRMARFSKAPGNRNMATFGMMDELRHGQLQLFFPHEYCKKDRQ
FDWAWRAYHSNEWAAIAAKHFFDDIITGRDAISVAIMLTFSFETGFTNMQFLGLAADAAEAGDYTFANLISSIQTDE
SRHAQQGGPALQLLIENGKREEAQKKVDMAIWRAWRLFAVLTGPVMDYYTPLEDRSQSFKEFMYEWIIGQFERSLID
LGLDKPWYWDLFLKDIDELHHSYHMGVWYWRTTAWWNPAAGVTPEERDWLEEKYPGWNKRWGRCWDVITENVLNDRM
DLVSPETLPSVCNMSQIPLVGVPGDDWNIEVFSLEHNGRLYHFGSEVDRWVFQQDPVQYQNHMNIVDRFLAGQIQPM
TLEGALKYMGFQSIEEMGKDAHDFAWADKCKPAMKKSAGGGSGEGRGSLLTCGDVEENPGPMSFEKICSLDDIWVGE

METFETSDGTEVLIVNSEEHGVKAYQAMCPHQEILLSEGSYEGGVITCRAHLWTFNDGTGHGINPDDCCLAEYPVEV
KGDDIYVSTKGILPNKAHS

Seq. ID No: 515
>TMO_3
MSFTKVCSVGDIWEGEMEPFTVDGHEILLVGVEGGGIKAFQGICPHQDIALSEGKFDGKKLICRAHLWQFDASNGKG
INPDDCALAEYPVKVDGDDVYVQTAGVEALFAHSGGGSGEGRGSLLTCGDVEENPGPMALLNRMDWYDLARTTNWSP
KYVTESELFPPELSGDHGIPMEKWETYDEPYKQTYPEYVKVQREKDAGAYSVKAALERSQIYERSDPGWLTVMKQHY
GAIALGEYAASSAEARMMRFSKAPGMRNMATLGSMDEIRHGQIQLYFPHEHVSKDRQFDWAAKAFHTNEWAAIAARH
FFDDIMMTRDAISVAIMLTFSFETGFTNMQFLGLAADAAEAGDHTFASLISSVQTDESRHAQIGGPTLQILIENGKK
AEAQKKVDIAFWRAWRLFSVLTGPVMDYYTPLEHRKQSFKEFMQEWIVAQFERALSDLGLDKPWYWDTFLQQLDQQH
HGMHLGVWYWRPTVWWNPAAGVTPAERDWLEEKYPGWNDTWGQCWDVIIDNLVDGNIAQTYPETLPIVCNMCNLPIN
CTPGNGWAVQDYPLEYNGRLYHFGSEPDRWCFEQEPERYAGHMTLVDRFLAGLVQPMDLGGALAYMGLAPGEIGDDA
HGYSWVDIYKKMRMKKAS

Seq. ID No: 516
>TMO_4
MSVASSAQAYHNNMVGPVMRAGDLALAVIEAARVDNPGKEVFVDDKRAYVRIHTEQEMILRRETIEEELGRPFKMND
LEVDLSSFAGQIESLDDAVRFYFTKKLGGGSGEGRGSLLTCGDVEENPGPSSNPIIHNQKDGSRFAQREGDTILRAA
LRAGVGLSYECNSGGCGGCKFELLEGEVDTLWPDAPGLSDKDRRRGRHLACQCRARGPVSIKAATGAEYVPKVVPQR
QTARLVGSTDITHDLREFRFRSAAGASFLPGQFAMLDLPGLASARAYSMSNTANDDGEWHFQVRRVPHGQGTHVLFD
RLGVGDEIGLDGPYGVAWLRTGAPRDIVCVAGGSGLAPMVSIARGAAAAGMLKDRKLYFFYGARTPRDVCGAEMLAQ
LDGFGERIIYLPVVSLPGGEGEWQGETGYVHDAVARTLPGSLAGFEFYFAGPPPMTQALQEMLMVGHRVPFEQIHFD
RFF

Seq. ID No: 517
>TPH_1
MPSRLNKDEYQFYIDLDNKSTPALNEIVKCLRLDIGATVHELSRDKKKDAVPWFPKTIQDLDKFANQILSYGAELDS
DHPGFTDPVYRARRKEFADIAFHYKHGQPIPCVTYTEEEKKTWGTVFKELKLLYPTHACYEHNHVFPLLEKYCGYNE
NNIPQLEDVSKFLQTCTGFRLRPVAGLLSSRDFLAGLAFRVFHSTQYIRHWSKPMYTPEPDICHELLGHAPLFADPS
FAQFSQEIGLASLGAPDEYIERLATLYWFTVEFGLCKQDDKIKAYGAGLLSSFGELQYCLTDKPELKPFEPEKTSLQ
KYPITEFQPVYFIAESFEDAKEKMRKFATTIPRPFSVRYNPYTQSIEVLDNVQQLKNLADCINSEIGTLCCA

Seq. ID No: 518
>TPH_2
MIEDNKENKDHSLERGRASLIFSLKNEVGGLIKALKIFQEKHVNLLHIESRKSKRRNSEFEIFVDCDINREQLNDIF
HLLKSHTNVLSVNLPDNFTLKEDGMETVPWFPKKISDLDHCANRVLMYGSELDADHPGFKDNVYRKRRKYFADLAMN
YKHGDPIPKVEFTEEEIKTWGTVFQELNKLYPTHACREYLKNLPLLSKYCGYREDNIPQLEDVSNFLKERTGFSIRP
VAGYLSPRDFLSGLAFRVFHCTQYVRHSSDPFYTPEPDTCHELLGHVPLLAEPSFAQFSQEIGLASLGASEEAVQKL
ATCYFFTVEFGLCKQDGQLRVFGAGLLSSISELKHALSGHAKVKPFDPKITCKQECLITTFQDVYFVSESFEDAKEK
MREFTKTIKRPFGVKYNPYTRSIQILKDTKSITSAMNELQHDLDVVSDALAKVSRKPSI

Seq. ID No: 519
>TPH_3
MSGLALDRSSQPHEVRTLEVNELDPKVFAVVEVRKDEPGVLGDVLKVFTESSINITNIESRFKSFARDGPAFHIDFE
GEAREHRVQRVLRDVKSVPGVSQVTVMEEREVPWFPINIRDLDLTTDTLDGGTALINEDHPGFNDLAYRQRREEIVT
AAKEHRHGDRIARVQYLEHEVETWRAVYEQLRECHSRWACTEYLEMLPQMERFCGYAPGNIPQLADISDFLQQRTGF
TLRPITGLLSARDFLNALAFRVFYSTQYIRHHGNPFYTPEPDICHELMGHVPLFANAAFADFSQEIGLASLAASDDD
IARLAAVYWFTVEFGLVRQGGEVKAYGAGLLSSFGEMEWSCSREPSTTCREMGSVAELQAPSIVPLDPTQAGKQAYP
ITTYQPLYFCAESMQDAKAKISQFCDTLTRPFFPQYDPLTQNIRVTKAVRRARRISTVEMQMAKQLDYFEKQ

Seq. ID No: 520
>TPH_4
MAVPWFPKTIQDLDKFANQILSYGAELDSDHPGFTDPVYRTRRKEFADIAFHYKHGQPIPRVTYTEEEKKTWGTVFK
ELKLLYPTHACYEHNHVFPLLEKYCGYNENNIPQLEDVSNFLQTCTGFRLRPVAGLLSSRDFLAGLAFRVFHSTQYI
RHWSKPMYTPEPDICHELLGHAPLFADPSFAQFSQEIGLASLGAPDEYIERLATLYWFTIEFGLCKQDDKIKAYGAG
LLSSFGELQYCLTDKPDLKPFEPEKTSLQKYPITEFQPVYFIAESFEDAKEKVRKFATTIPRPFSVRYNPYTQSIEV
LDNVQQLKNLADCINSEIGILCCALRKLE

Seq. ID No: 521
>TPH_5
MLISFTLNLVHQKKNSEFEIFLDCDSNREQLNEIFQLLRPHVNLITMNPQEDFSVEEDDMESVPWFPIKISDLDKSA
NRVLMYGSDLDADHPGFKDNVYRRRRKYFADVAMNYKYGDPIPHIEFTEEEVKTWGTVFRELNKLHQTHACREYLKN
LPLLVKHCGYREDNIPQLEDVSRFLKERSGFTIRPVAGYLSPRDFLAGLAFRVFHCTQYVRHSSDPLYTPEPDTCHE
LLGHVPLLAEPSFAQFSQEIGLASLGASDEAVQKLATCYFFTVEFGLCKQEGKLKVYGAGLLSSISELKHSLSGNAN
VKPFDPMVTCSQECIITSFQEVYFYSESFEEAKEKMREFAKTIKRPFGLKYNPYTQSVQMLKDTQSITTLVSELRHE
LDIISDALNKMNKQLGV

Seq. ID No: 522
>TPH_6
MHSPEPDCCHELLGHVPMLADKTFAQFSQDIGLASLGVTDEEIEKLSTLYWFTVEFGLCKQDGEVKAYGAGLLSSYG
ELLHALSDKPEVRPFDPDEAAIQPYQDQNYQPVYFVSESFTDAKEKLRNYASRIKRPFAARYDPYTVSIEVLDSPGQ
IQSSLEELKDELQTLTTALNILS

Seq. ID No: 523
>TPH_7
MMISTESDLRRQLDENVRSEADESTKEECPYINAVQSHHQNVQEMSIIISLVKNMNDMKSIISIFTDRNINILHIES
RLGRLNMKKHTEKSEFEPLELLVHVEVPCIEVERLLEELKSFSSYRIVQNPLMNLPEAKNPTLDDKVPWFPRHISDL
DKVSNSVLMYGKELDADHPGFKDKEYRKRRMMFADIALNYKWGQQIPIVEYTEIEKTTWGRIYRELTRLYKTSACHE
FQKNLGLLQDKAGYNEFDLPQLQVVSDFLKARTGFCLRPVAGYLSARDFLSGLAFRVFYCTQYIRHQADPFYTPEPD
CCHELLGHVPMLADPKFARFSQEIGLASLGTSDEEIKKLATCYFFTIEFGLCRQDNQLKAYGAGLLSSVAELQHALS
DKAVIKPFIPMKVINEECLVTTFQNGYFETSSFEDATRQMREFVRTIKRPFDVHYNPYTQSIEIIKTPKSVAKLVQD
LQFELTAINESLLKMNKEIRSQQFTTNKIVTENRSS

Seq. ID No: 524
>TrpHalo_1
MSTASKNIDITRFPKKYDAATKDSDFYDVVIVGAGPGGSTTAYYLAKEGKKVLLLEKKKFPRDKICGDAICKLAIEM
LMDMGVYEGLVREKKARVAHNGGLVSPSGLSFIGNTYLKPGEIPAAAACKRMVLDEAIAKAAIGAGAELKENSPVTD
AVFDSSTGLWTISIEGSDVKHMGRVLVCADGAPSKLATQLGIVKQAPQGVCSRAYIKEGTHRFRADGVVFYPRNILP
AYAALFRHIDDTVAYCTYILPFNPKVTTDDLSYWHHRLLEEDPSISQAVGKNADMERMKAWGLRMGGEPVTYGNHVL
VVGDAAGMIDPLTGEGIHHAMDGGRIAAHFLCEAIAVGNFDKEVMKEYQNRWLYTFGNDYKWSQAICHFLYRFPIFI
DATAAAAQRRGNNFLALYADIMTGRIPKANIFRPDISLPIAFEVLVLLWKMMFTGGGGNNKMKSQ

Seq. ID No: 525
>TrpHalo_2
MSTASKNIDITRFPKKYDAATKDSDFYDVVIVGAGPGGSTTAYYLAKEGKKVLLLEKKKFPRDKICGDAICKLAIEM
LMDMGVYEGLVREKKARVAHNGGLVSPSGLSFIGNTYLKPGEIPAAAACKRMVLDEAIAKAAIGAGAELKENSPVTD
AVFDSSTGLWTISIEGSDVKHMGRVLVCADGAPSRLAMQLGIVKGTPKCVCSRAYIKGGTHRFKEDGMVFYVPSILP
GYVALLRHIDDQLTYCTYILPGNPRATTKDLSYWHHRLLEEDPNISQAVGKNAELEKMKAWDLRVGGEPVTYGNHVL
VVGDAAGMIDPLTGEGIHHAMDGGRIAAHFLCEAIAVGNFDKEVMKEYQNRWLKAFGNDFRWSQAIGNFLYRYPIFI
DATAAVAEKKGDRFLARWADIMAGRIPKISVLRPQFLLAVGFQALLLFYKKIFKGGYGKKTKIL

Seq. ID No: 526
>TrpHalo_3
MSSLIAPKVDTIDITRFPKKYDPAAEDSDFYDVVIVGAGPGGSTTAYYLAKKGKKVLLLEKKKFPRDKICGDAICKT
AIEILMDMGVYGGLIREQKAYMIDYGGLVSPSGLSFVGHTHELFGEIPGAVVCKRVVLDKVISRTAQSAGAELLENS
PVTDAVFDSSTGLWTISIEGSDVKHMGRVLVCADGAPSRLAMQLGIVKGTPKCVCSRAYIKGGTHRFKEDGMVFYVP
SILPGYVALLRHIDDQLTYCTYILPGNPRATTKDLSYWHHRLLEEDPNISQAVGKNAELEKMKAWDLRVGGEPVTYG
NHVLVVGDAAGMIDPLTGEGIHHAMDGGRIAAHFLCEAIAVGNFDKEVMKEYQNRWLKAFGNDFRWSQAIGNFLYRY
PIFIDATAAVAEKKGDRFLARWADIMAGRIPKISVLRPQFLLAVGFQALLLFYKKIFKGGYGKKTKIL

Seq. ID No: 527
>TrpHalo_4
MSGKIDKILIVGGGTAGWMAASYLGKALQGTADITLLQAPDIPTLGVGEATIPNLQTAFFDFLGIPEDEWMRECNAS
YKVAIKFINWRTAGEGTSEARELDGGPDHFYHSFGLLKYHEQIPLSHYWFDRSYRGKTVEPFDYACYKEPVILDANR
SPRRLDGSKVTNYAWHFDAHLVADFLRRFATEKLGVRHVEDRVEHVQRDANGNIESVRTATGRVFDADLFVDCSGFR
GLLINKAMEEPFLDMSDHLLNDSAVATQVPHDDDANGVEPFTSAIAMKSGWTWKIPMLGRFGTGYVYSSRFATEDEA
VREFCEMWHLDPETQPLNRIRFRVGRNRRAWVGNCVSIGTSSCFVEPLESTGIYFVYAALYQLVKHFPDKSLNPVLT
ARFNREIETMFDDTRDFIQAHFYFSPRTDTPFWRANKELRLADGMQEKIDMYRAGMAINAPASDDAQLYYGNFEEEF
RNFWNNSNYYCVLAGLGLVPDAPSPRLAHMPQATESVDEVFGAVKDRQRNLLETLPSLHEFLRQQHGR

Seq. ID No: 528
>TrpHalo_5
MDEIDDPRIRSVVIVGGGTAGWMTAAALVQHFRTAPLKITVVESSDIGTIGVGEATIPTIRRFYGQLGLRDDDVMRA
TQATCKLGIRFLDWSGPGSDFIHPFGLYGQDVKGIGFHHYWLKQRRAGDAAPLAAYSLGAALAAGGKFTLPSPHPPS
QLSVFDWALHLDAGLFAQHLRAYAEAGGCARIDARIRSVELRPEDGFVRALTLDDGREVEGDLFVDCSGFKGLVIGE
ALGVGFEDWGRWLPCDAAYAVQSENRPGDAPAPFTRVTARSAGWQWGIPLRHRAGNGLVFSSAHLSDDQALAELMPH
LLGDPLTEPRRIPFRPGRRSQAWAKNCVAIGLSSGFLEPLESTSIALIETGIERLKALFPDRRFAQPILDEFNDQTA
REMERVRDFIILHYKLNRRTDTDFWRDCREMPVPETLERKIALWTARGQFVRYRWEMFHPASWLAIYDGFGLYPDHH
DPAVDAMDPAYLARSLAEMRANIADLVARTPEHAQFLAGLDPAASAA

Seq. ID No: 529
>TrpHalo_6
MIRSVVIVGGGTAGWMTASYLKAAFDDRIDVTLVESGNVRRIGVGEATFSTVRHFFDYLGLDEREWLPRCAGGYKLG
IRFENWSEPGEYFYHPFERLRVVDGFNMAEWWLAVGDRRTSFSEACYLTHRLCEAKRAPRMLDGSLFASQVDESLGR
STLAEQRAQFPYAYHFDADEVARYLSEYAIARGVRHVVDDVQHVGQDERGWISGVHTKQHGEISGDLFVDCTGFRGL
LINQTLGGRFQSFSDVLPNNRAVALRVPRENDEDMRPYTTATAMSAGWMWTIPLFKRDGNGYVYSDEFISPEEAERE
LRSTVAPGRDDLEANHIQMRIGRNERTWINNCVAVGLSAAFVEPLESTGIFFIQHAIEQLVKHFPGERWDPVLISAY
NERMAHMVDGVKEFLVLHYKGAQREDTPYWKAAKTRAMPDGLARKLELSASHLLDEQTIYPYYHGFETYSWITMNLG
LGIVPERPRPALLHMDPAPALAEFERLRREGDELIAALPSCYEYLASIQ

Seq. ID No: 530
>TrpHalo_7
MLESIVVVGGGTSGWMTASYLSAAFGERISVTVVESARVGTIGVGEATFSTVRHFFEYLGLSEETWMPACNATYKLG
IRFENWRAPGHHFYHPFERQRVVDGFTLPDWWLADGGATERFDKECFLVGTLCDTMRSPRHMDGALFEGDLTDRPAG
RSTLAEQGTQFPYAYHFDAALLADFLRDYAVARGVLHVVDDVVHVARDERGWISHVATRGSGDLAGDLFVDCTGFRG
LLINDALDEPFESYQDTLPNDSAVALRVPVDMEREGLRPCTTSTAQAAGWIWTIPLFGRVGTGYVYARDYCTPEEAE
RTLRRFVGPAADDLEANHIRMRIGRSRRSWVNNCVAVGLSSGFVEPLESTGIFFIQHAIEQLVKHFPDADWDPALRS
AYNTLVNRCMDGVREFLVLHYYGAARADNEYWRDTKTRKIPDSLAERVEQWRTKLPHPESVYPHYHGFEAYSYVCMV
LGLGGIPLKPSPALRMLDPSAAQREFRLLATQAEDLRRTLPSQYAYFAQFR

Seq. ID No: 531
>TrpHalo_8

MNKPIKNIVIVGGGTAGWMAASYLVRALQQQANITLIESAAIPRIGVGEATIPSLQKVFFDFLGIPEREWMPQVNGA
FKAAIKFVNWRKSPDPSRDDHFYHLFGNVPNCDGVPLTHYWLRKREQGFQQPMEYACYPQPGALDGKLAPCLSDGTR
QMSHAWHFDAHLVADFLKRWAVERGVNRVVDEVVDVRLNNRGYISNLLTKEGRTLEADLFIDCSGMRGLLINQALKE
PFIDMSDYLLCDSAVASAVPNDDARDGVEPYTSSIAMNSGWTWKIPMLGRFGSGYVFSSHFTSRDQATADFLKLWGL
SDNQPLNQIKFRVGRNKRAWVNNCVSIGLSSCFLEPLESTGIYFIYAALYQLVKHFPDTSFDPRLSDAFNAEIVHMF
DDCRDFVQAHYFTTSRDDTPFWLANRHDLRLSDAIKEKVQRYKAGLPLTTTSFDDSTYYETFDYEFKNFWLNGNYYC
IFAGLGMLPDRSLPLLQHRPESIEKAEAMFASIRREAERLRTSLPTNYDYLRSLRDGDAGLSRGQRGPKLAAQESL

Seq. ID No: 532
>TrpM_1
MSPVALSPKRVDIVDIRGNDMQYSLVNEIHKGLNPPNGTRRSLPTMLLYDSEGLKLFEKITYVDEYYLTNAEIEVLE
KHSRRLVEKIPSNAQLLELGSGNLRKIEILLREFERVGKPVDYYALDLSLSELERTFSNVSLEEYKSVGFHGLHGTY
DDAHTWLSDPKNRERPTVVLSMGSSLGNFSPPDAAAFLAGFATLLKPSDFMVIGLDACEDPDRVYKAYNDSAGITRK
FYENGLANANKTLGHEVFRPDEWEVVTEYDAVNGRHQVFYVPTKDVSVGDVLLRRGEKIIFAEAFKYGCQAREKLWH
DAGLIEAAEFGSGSEDYRTYI*

Seq. ID No: 533
>TrpM_10
MLGPVPSPSPVPIPPGSRPGASPGLEATIPIIDIRSTAHSVTVAALEDGIRANVLSGFTKPYNEKELPNLLLYNEEG
LRLFEQITYQPDYYLTRLEIDILSRHAHQIANSVPDGAILLELGAGALRKTALILDALEAQGKDVTYFALDLDKPEL
LRTLAEVKGRYTHVSLAGLWGTYDDGCTWLKQVKDRPRIILWLGSSVGNMSRKEAGQFIRTFGDILAPRDRFIVAID
SKNHKLNDIRAAYDDRAGVTRRFALNALGNINDLFNADVVDVSSFDYNPYYNEVQGRNEAYFRCLKDTQVRIPSETP
ILVHEGEYIRFAFSHKYDRVERQVLWTAAGAYPVQEWMSQDGDYALTMLSWSS*

Seq. ID No: 534
>TrpM_11
MTYSIVDIRKTDTCLKNSIINGINQSTKSIPAIVLYDELGLQYYEKVTYLKEYYLTEAEIDILKNKADQISDYIPEG
SSLIELGSGALRKTRLLLDSIEKQKKKVIYYALDLMEGELKRTLSSLGKFQYVKLVGLWGVYEDGIDYASNLPGDSH
KTILWMGSSIGNFNRDEAANFVKTIQDKAMNPGDLFLIGIDRRKNPDKITAAYNDPKGINAKFIMNGLNHVNAIFDQ
PIFDSNNFEHVTMYNDDVGRHEAYCKVKNDTTLEFKESKDNPKTIIKLNKNELINIGYSHKYNKAETDALFDFSLLS
YMESWTDSQSLYDLHLVYKSPFHFTRKFDSHK*

Seq. ID No: 535
>TrpM_12
MSKDVQVLDIRASPQSKGSIPNLRTAILDGLQKAPGMRTLPSEILYDDRGLKIYNDCIRSWSEWYYPISAETEILEI
NGKDIARVFSTSDRGEAVLIELGAGSLDKTSKILVSLSETVQNVSDSQPPITYYALDLERSELQRTLSELQKNIGEK
IAGKIATKGMWGTYDDGIRSVENNELHLDAAVPVHFLFLGGTIGNFSKGEGDVTFLRNLPLNAQRGDTILLGIDREK
SKEIIERAYNFPAAREWIMNGLNVSGHLLSGDKDLFQLDNWDRYAMYDEKLGRLEAGYRSKIDQIIEVTANYSIPFK
KDESVMAIFSNKYTDDELNFLISKANLKTINSWVDHKALYYIFSLRKV*

Seq. ID No: 536
>TrpM_13
MPRIQVLDIRGSKESVGSTPHLRAAILEGLLKPPGSRTLPSETLYDEVGLKMYNDGMKAWAEWYYPVEAERQILERY
GRDIAKLFTTSAKGKAVLIELGAGSLDKTSQVLLSAAEITRTTGPMNNIAYYALDLERGELERTIGRLQEVIGDQIA
GKISTAGMWGTYDDGIRVIEKNELELEPDIPVHILFLGGTIGNFSKQDGDVAFLKSLPLDHKRGDTLLVGMDRHKSA
DAIERSYGFAAAKDWIMNGLKVSGRVLTGDEGLFEIGNWERYAKYNEELGRYEAGYKSQKEHALKISEGVDITFLKD
EVVLVMFSNKYTDAEMDSVVDSAGLVKNGSWMDEKAQYCLLSLRANNGPV*

Seq. ID No: 537
>TrpM_14

MSQIEVLDIRGSKEATGSTPHLRAEILQGLSKSPGHRTIPGETLFDETGLKMYDEGMKTWRKWYYPFEAEKEILEVR
GLEIAKLLKTSSKGEAVLIELGAGSLEKTSQILLSAAQIAETADNSTTNPITYYALDLEHRELERTLAALQDAIGPR
IAGKITTKGMWGTYEDGIRVVERNDLKFPSDVPLHILFLGGTIGNFSKADGDIAFLKSLPLNRKRGDTLLLGVDRAK
AVELIERAYGFAAATGWIMNGLKVSGRVLTGDEELFESGNWERYSKYNEELGRYEAGYKSRKDQTIKVAKDVDIVFS
KDEVILVTYSNKYTDAEIKTVFDGAGLEIVESWMDKKAQYCLFLLKA*

Seq. ID No: 538
>TrpM_2
MTLSLANYLAADSAAEALRRDVRTGLTATPKSLPPKWFYDAVGSDLFDQITRLPEYYPTRTEAQILRTRSAEIIAAA
GADTLVELGSGTSEKTRMLLDAMRDADLLRRFIPFDVDAGVLRSAGAAIGAEYPGIEIDAVCGDFEEHLGKIPRVGR
RLVVFLGSTIGNLTPQPRAEFLATLADTLQPGDSLLLGTDLVKDTGRLVRAYDDAAGVTAAFNRNVLAVVNRELSAD
FDLDAFEHIAKWNDDEERIEVWLRARTAQHVRIPALDLEIDFAAGEQMLTAVSCKFRPDSVAAELAEAGLRQTHWWT
DPAGDFGLSLAVR*

Seq. ID No: 539
>TrpM_3
MTLSLANYLAADSAAEALRRDVRAGLTAAPKSLPPKWFYDAVGSDLFDQITRLPEYYPTRTEAQILRTRSAEIIAAA
GADTLVELGSGTSEKTRMLLDAMRDAELLRRFIPFDVDAGVLRSAGAAIGAEYPGIEIDAVCGDFEEHLGKIPHVGR
RLVVFLGSTIGNLTPAPRAEFLSTLADTLQPGDSLLLGTDLVKDTGRLVRAYDDAAGVTAAFNRNVLAVVNRELSAD
FDLDAFEHVAKWNSDEERIEVWLRARTAQHVRVAALDLEVDFAAGEEMLTAVSCKFRPENVVAELAEAGLRQTHWWT
DPAGDFGLSLAVR*

Seq. ID No: 540
>TrpM_4
MRVSGANHLGEDAGHLALRRDVYSGLQKTPKSLPPKWFYDTVGSELFDQITRLPEYYPTRAEAEILRARSAEVASAC
RADTLVELGSGTSEKTRMLLDALRHRGSLRRFVPFDVDASVLSATATAIQREYSGVEINAVCGDFEEHLTEIPRGGR
RLFVFLGSTIGNLTPGPRAQFLTALAGVMRPGDSLLLGTDLVKDAARLVRAYDDPGGVTAQFNRNVLAVINRELEAD
FDVDAFQHVARWNSAEERIEVWLRADGRQRVRVGALDLTVDFDAGEEMLTAVSCKFRPQAVGAELAAAGLHRIRWWT
DEAGDFGLSLAAK*

Seq. ID No: 541
>TrpM_5
MTLTLSNYLAADSAATALRRDVHEGLTQSPKMLPPKWFYDSVGSDLFDQITRLPEYYPTRTEAQILTHRSPEIVAAA
GADTLVELGSGTSEKTRMLLDAMRDGGQLRRFIPFDVDAGVLRAAGAAIGQEYPGIEIDAVCGDFEEHLGKIPAVGR
RLVAFLGSTIGNLTPGPRADFLASLAETLQPGDSVLLGTDLVKDTGRLVSAYDDSAGVTAAFNRNVLSVVNRELDAD
FDLDAFAHVAKWNAEEERIEVWLRADAPQQVRIAGLDLDVAFGAGEEMLTAVSCKFRADGVADELAKAGLRQTHWWT
DEAGDFGLSLAVK*

Seq. ID No: 542
>TrpM_6
MLEATSTQNLVSFQIPIVDIRTPSCLEETIRKKVVSGLARPYNKKSIPDLLLYNETGLRLFEDLTYQPDYYLTGLEI
EILSKHSLQIADSIPVGSLIMELGAGALRKTALILDALEAQKKEVAYLALDLDRPELVRTLGQLNGKYTHVKLGGLW
GTYDDGRRWLSENTSDSPRTILWLGSSIGNVKRDDAGDFIRSFGDVLSSKDRFVVAIDSRYHEVDTICRAYNDREGF
AERFCLNGIDSFNQLFGRAIIDISCAKYRTVYNEVKGRHEVYYRCTHDFEIRLPGDYPPTFLYEGELILLAHSYKYA
AVERETLWLRAGARPEKEWMTDGSYTVTMLSWP*

Seq. ID No: 543
>TrpM_7
MSPSTVNKIASSPVFDIRSDETKGFAKAPIEDELAGLQAVYNEKTLPNVLLYDAKGLQLFEKITYTNDYYLTGLEMD
LLGEHADEMAEWIKDGAALVELGAGALRKTAILLDAIERQGKRITFYALDLDHSELTRTLAELEGRYRHITLCGLWG
TYDDGRAWLASTNEEQRVLLWLGSSIGNLSRQEAKDFLHSFGRALRPGIDKFIVAMDSKYNAVSSMTRAYNDSEGVT

ASFALNLLDAFNAKVGFKALPPSSFCYSPFFNQAQGRNEAYLRARHGVRFEVNGIAVEVRDEELIRFAYSHKYDNAE
RDLLWRAAEANVEQEWLHSPQSGRARYSISLLSFRD*

Seq. ID No: 544
>TrpM_8
MTLSLSNHLPANSAARVLRRDVLDGLTQTPKALPPKWFYDSVGSDLFDQITRLPEYYPTRTEAQILRTRSAEIAEAS
GADTLVELGSGTSEKTRMLLDALRDNGTLRRFIPFDVDAGVLNAAGAAIQKEYPGVEVDAVCGDFEEHLGEIPRVGR
RLIAFLGSTIGNLTPQPRARFLTALAQTMRPGDSLLLGTDLVKDTERLVRAYDDSAGVTARFNRNVLAVINRELDAD
FDLAAFDHVARFNAAEERIEVWLRARGAQRVYYVRELDLTVDFADGEEMLTAVSCKFRPDGVAAELAAAGLRRTHWWT
DPAGDFGLSLSTK*

Seq. ID No: 545
>TrpM_9
MTISIANYLAADSAATALRRDVREGLAGTPKSLPPKWFYDSVGSDLFDQITRLPEYYPTRAEAQILRTHAVDVAAAS
GADTLVELGSGTSEKTRLLLDALHRADSLRRFIPFDVDASILQSAGAAISQEYPDVEIEAVCGDFEEHLGKIPLQGR
RLVVFLGSTIGNLTSGPRATFLSALADSLQPGDTLLLGTDLVKDVDRLKRAYDDAAGVTARFNKNVLTVVNRELGAD
FDLDAFEHVCKWNADEERIEVWLRANTLQRVHISGLELDVEYAAGEEMLTAVSCKFRPEGIAAELAAVGLNRTHWWT
DDAGDFGLSLAVK*

Seq. ID No: 546
>TrpS_1
MTTLLNPYFGEFGGMYVPQILMPALNQLEEAFVSAQKDPEFQAQFADLLKNYAGRPTALTKCQNITAGTRTTLYLKR
EDLLHGGAHKTNQVLGQALLAKRMGKSEIIAETGAGQHGVASALASALLGLKCRIYMGAKDVERQSPNVFRMRLMGA
EVIPVHSGSATVKDACNEALRDWSGSYETAHYMLGTAAGPHPYPTIVREFQRMIGEETKAQILDKEGRLPDAVIACV
GGGSNAIGMFADFINDTSVGLIGVEPGGHGIETGEHGAPLKHGRVGIYFGMKAPMMQTADGQIEESYSISAGLDFPS
VGPQHAYLNSIGRADYVSITDDEALEAFKTLCRHEGIIPALESSHALAHALKMMREQPEKEQLLVVNLSGRGDKDIF
TVHDILKARGEI

Seq. ID No: 547
>TrpS_2
MWFGEFGGQYVLETLIGPLKELEKAYKRFKDDEEFNRQLNYYLKTWAGRPTPLYYAKRLTEKIGGAKVYLKREDLVH
GGAHKTNNAIGQALLAKFMGKTRLIAETGAGQHGVATAMAGALLGMKVDIYMGAEDVERQKMNVFRMKLLGANVIPV
NSGSRTLKDAINEALRDWVATFEYTHYLIGSVVGPHPYPTIVRDFQSVIGREAKAQILEAEGQLPDVIVACVGGGSN
AMGIFYPFVNDKKVKLVGVEAGGKGLESGKHSASLNAGQVGVSHGMLSYFLQDEEGQIKPSHSIAPGLDYPGVGPEH
AYLKKIQRAEYVAVTDEEALKAFHELSRTEGIIPALESAHAVAYAMKLAKEMSRDEIIIVNLSGRGDKDLDIVLKVS
GNV

Seq. ID No: 548
>affibody_tag_1
MVDNKFNKETIQASQEIRLLPNLNGRQKLAFIHSLLDDPSQSANLLAEAKKLNDAQAPKNAAIRSSSASSGGSGGSS
SS

Seq. ID No: 549
>affibody_tag_2
NAAIRSSSASSGGSGGSSSSVDNKFNKETIQASQEIRLLPNLNGRQKLAFIHSLLDDPSQSANLLAEAKKLNDAQAP
K

Seq. ID No: 550
>affibody_tag_3
NAAIRSSSASSGGSGGSSSSGVDNKFNKELGWATWEIFNLPNLNGVQVKAFIDSLRDDPSQSANLLAEAKKLNDAQA
PK

Seq. ID No: 551
>affibody_tag_4
MVDNKFNKELGWATWEIFNLPNLNGVQVKAFIDSLRDDPSQSANLLAEAKKLNDAQAPKGNAAIRSSSASSGGSGGS
SSS

Seq. ID No: 552
>affibody_tag_5
MVDNKFNKEMRNAYWEIALLPNLNNQQKRAFIRSLYDDPSQSANLLAEAKKLNDAQAPKSSNAAIRSSSASSGGSGG
SSSS

Seq. ID No: 553
>affibody_tag_6
NAAIRSSSASSGGSGGSSSSGVDNKFNKEMRNAYWEIALLPNLNNQQKRAFIRSLYDDPSQSANLLAEAKKLNDAQA
PK

Seq. ID No: 554
>cofold_1
MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLL
AEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQE
PYFTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPW
AWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSY
EEELAKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNNNNNNNNLGIE
GR

Seq. ID No: 555
>cofold_2
MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKLICTTGKLPVPWPTLVTTLGYGLQCFARYPD
HMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIT
ADKQKNGIKANFKIRHNIEDGGVQLADHYQQNTPIGDGPVLLPDNHYLSYQSALSKDPNEKRDHMVLLEFVTAAGIT
LGMDELYK

Seq. ID No: 556
>cofold_3
MAMFCTFFEKHHRKWDILLEKSTGVMEAMKVTSEEKEQLSTAIDRMNEGLDAFIQLYNESEIDEPLIQLDDDTAELM
KQARDMYGQEKLNEKLNTIIKQILSISVSEEGEKEGSGSG

Seq. ID No: 557
>cofold_4
MYLLGIGLILALIACKQNVSSLDEKNSVSVDLPGEMKVLVSKEKNKDGKYDLIATVDKLELKGTSDKNNGSGVLEGV
KADKSKVKLTISDDGSG

Seq. ID No: 558
>cofold_5
MADRDRSGIYGGAHATYGQQQQQGGGGRPMGEQVKGMLHDKGPTASQALTVATLFPLGGLLLVLSGLALTASVVGLA
VATPVFLIFSPVLVPAALLIGTAVMGFLTSGALGLGGLSSLTCLANTARQAFQRTPDYVEEAHRRMAEAAAHAGHKT
AQAGQAIQGRAQEAGAGGGAG

Seq. ID No: 559
>oxidase_1

MKILILGIFLFLCSTPAWAKEKHYYIGIIETTWDYASDHGEKKLISVDTEHSNIYLQNGPDRIGRLYKKALYLQYTD
ETFRTTIEKPVWLGFLGPIIKAETGDKVYVHLKNLASRPYTFHSHGITYYKEHEGAIYPDNTTDFQRADDKVYPGEQ
YTYMLLATEEQSPGEGDGNCVTRIYHSHIDAPKDIASGLIGPLIICKKDSLDKEKEKHIDREFVVMFSVVDENFSWY
LEDNIKTYCSEPEKVDKDNEDFQESNRMYSVNGYTFGSLPGLSMCAEDRVKWYLFGMGNEVDVHAAFFHGQALTNKN
YRIDTINLFPATLFDAYMVAQNPGEWMLSCQNLNHLKAGLQAFFQVQECNKSSSKDNIRGKHVRHYYIAAEEIIWNY
APSGIDIFTKENLTAPGSDSAVFFEQGTTRIGGSYKKLVYREYTDASFTNRKERGPEEEHLGILGPVIWAEVGDTIR
VTFHNKGAYPLSIEPIGVRFNKNNEGTYYSPNYNPQSRSVPPSASHVAPTETFTYEWTVPKEVGPTNADPVCLAKMY
YSAVDPTKDIFTGLIGPMKICKKGSLHANGRQKDVDKEFYLFPTVFDENESLLLEDNIRMFTTAPDQVDKEDEDFQE
SNKMHSMNGFMYGNQPGLTMCKGDSVVWYLFSAGNEADVHGIYFSGNTYLWRGERRDTANLFPQTSLTLHMWPDTEG
TFNVECLTTDHYTGGMKQKYTVNQCRRQSEDSTFYLGERTYYIAAVEVEWDYSPQREWEKELHHLQEQNVSNAFLDK
GEFYIGSKYKKVVYRQYTDSTFRVPVERKAEEEHLGILGPQLHADVGDKVKIIFKNMATRPYSIHAHGVQTESSTVT
PTLPGETLTYVWKIPERSGAGTEDSACIPWAYYSTVDQVKDLYSGLIGPLIVCRRPYLKVFNPRRKLEFALLFLVFD
ENESWYLDDNIKTYSDHPEKVNKDDEEFIESNKMHAINGRMFGNLQGLTMHVGDEVNWYLMGMGNEIDLHTVHFHGH
SFQYKHRGVYSSDVFDIFPGTYQTLEMFPRTPGIWLLHCHVTDHIHAGMETTYTVLQNEDTKSG

Seq. ID No: 560
>oxidase_2
MGLNSAIPSLAILALSVGSYAAIGPVSDLHIVNKDLAPDGVQRPTVLAGGTFPGTLITGQKGDNFQLNVIDDLTDDR
MLTPTSIHWHGFFQKGTAWADGPAFVTQCPIIADNSFLYDFDVPDQAGTFWYHSHLSTQYCDGLRGAFVVYDPNDPH
KDLYDVDDESTVITLADWYHVLAQTVVGAATPDSTLINGLGRSQTGPADAELAVISVEHNKRYRFRLVSISCDPNFT
FSIDGHNMTVIEVDGVNTRPLTVDSIQIFAGQRYSFVLNANQPDDNYWIRAMPNIGRNTTTLDGKNAAILRYKNASV
EEPKTVGGPAQSPLNEADLRPLVPAPVPGNAVPGGADINHRLNLTFSNGLFSINNASFTNPSVPALLQILSGAQNAQ
DLLPTGSYIGLELGKVVELVIPPLAVGGPHPFHLHGHNFWVVRSAGSDEYNFDDAILRDVVSIGAGTDEVTIRFVTD
NPGPWFLHCHIDWHLEAGLAIVFAEGINQTAAANPTPQAWDELCPKYNGLSASQKVKPKKGTAI

Seq. ID No: 561
>oxidase_3
MSRFQSLLSFVLVSLAAVANAAIGPVADLTLTNAAVSPDGFSREAVVVNGITPAPLIAGQKGDRFQLNVIDNLTNHT
MLKTTSIHWHGFFQHGTNWADGVSFVNQCPIASGHSFLYDFQVPDQAGTFWYHSHLSTQYCDGLRGPFVVYDPNDPQ
ASLYDIDNDDTVITLADWYHVAAKLGPRFPLGADATLINGLGRSPGTTTADLAVIKVTQGKRYRFRLVSLSCDPNHT
FSIDGHTMTVIEADSVNTQPLEVDSIQIFAAQRYSFVLDASQPVDNYWIRANPAFGNVGFAGGINSAILRYDGAPEV
EPTTTQTTSTKPLNEADLHPLTPMPVPGRPEAGGVDKPLNMVFNFNGTNFFINNHSFVPPSVPVLLQILSGAQAAQD
LVPDGSVYVLPSNSSIEISFPATANAPGTPHPFHLHGHTFAVVRSAGSSEYNYDNPIFRDVVSTGQPGDNVTIRFQT
NNPGPWFLHCHIDFHLEAGFAVVLAEDTPDTAAVNPVPQSWSDLCPIYDALDPSDL

Seq. ID No: 562
>oxidase_4
MKFLLLSALLFLHSSLAWTREKHYYIGITEAVWDYASGSEEKELISVDTEQSNFYLRNGPDRIGRKYKKALYSEYTD
GTFTKTIDKPAWLGFLGPVIKAEVGDKVSVHVKNFASRPYTFHAHGVTYTKANEGAIYPDNTTDFQRADDKLFPGQQ
YLYVLRANEPSPGEGDSNCVTRIYHSHVDAPKDIASGLIGPLILCKKGSLHKEKEENIDQEFVLMFSVVDENLSWYL
EDNIKTFCSEPEKVDKDNEDFQESNRMYSINGYTFGSLPGLSMCAEDRVKWYLFGMGNEVDVHSELFHGQALTSKNY
HTDIINLFPATLIDVSMVAQNPGVWMLSCQNLNHLKAGLQAFFQVRDCNKPSPDDDIQDRHVRHYYIAAEETIWDYA
PSGTDTFTGENFTSLGSDSRVFFEQGATRIGGSYKKLVYREYTDDSFTNRKERGPDEEHLGILGPVIWAEVGDIIRV
TFHNKGQFPLSIQPMGVRFTKENEGTYYGPDGRSSKQASHVAPKETFTYEWTVPKEMGPTYADPVCLSKMYYSGVDL
TKDIFTGLIGPMKICKKGSLLADGRQKDVDKEFYLFATVFDENESLLLDDNIRMFTTAPENVDKEDEDFQESNKMHS
MNGFMYGNLPGLNMCLGESIVWYLFSAGNEADVHGIYFSGNTYLSKGERRDTANLFPHKSLTLLMTPDTEGSFDVEC
LTTDHYTGGMKQKYTVNQCKGQFEDVTLYQGERTYYIAAVEVEWDYSPSRDWEMELHHLQEQNVSNAFLDKEEFFIG
SKYKKVVYREFTDSTFREQVKRRAEEEHLGMLGPLIHADVGAKVKVVFKNMATRPYSIHAHGVKTKSSTVAPTLPGE
VRTYIWQIPERSGAGTEDSPCIPWAYYSTVDRVKDLYSGLIGPLIVCRKSYVKVFNPKKKMEFSLLFLVFDENESWY
LDDNINTYPDHPEKDNKDNEEFIESNKMHAINGKMFGNLQGLTMHVGDEVNWYVMAMGNEIDLHTVHFHGHSFQYKH
RGIHSSDVFDFFPGTYQTLEMFPQTPGTWLLHCHVTDHIHAGMVTTYTVLPNQETKSG

Seq. ID No: 563
>oxidase_5
MNFVTALPLIAQLIGTARAAIGPVTNLLVKNADIPPDGFTRAAVVANNQFPGPVIRATKGDTLSLNVVNQLTDATML
MGTSIHWHGFHQKGTSWADGVVGVTQCPIAPGHSFLYQFPTANQAGTFWYHSHYSTQYCDGLRGALIVYDPTDPYRT
WYDIDDESTIITLADWYHKAAPLQTLRTAKEDSVLINGQGRVPGDKTTDSTPLSVINIIPQKRYRFRLISISCDPAF
SFSIDGHSMTVIEADSQSVQPLTVNEITIFAGQRYSFILYANNPVGNYWIRSQPTYPDDGIQGYAGGINSAILRYSG
APAVNPTTKKASITIPLVEADLRPLYSPAAPGLPSPGAADVNIKLDISYNSPSETFFVNNSTFPEVPVPVLLQILSG
AQSANDLLPAGSVYTLPPNKVIEISMPGGRPGSPHPMHLHGHDFSVVRSAGSNRYNYANPVRRDVVNIGMEDTDNVT
IRFRVCSHTYLSLHCHIDFHLEDGQSGTLVPPLPHRLPPRGRIRCRLHRGILVRGRLGPDLQ

Seq. ID No: 564
>phosphatase_1
MQGPWVLLLLGLRLQLSLGIIPVEEENPDFWNRQAAEALGAAKKLQPAQTAAKNLIIFLGDGMGVSTVTAARILKGQ
KKDKLGPETFLAMDRFPYVALSKTYSVDKHVPDSGATATAYLCGVKGNFQTIGLSAAARFNQCNTTRGNEVISVMNR
AKKAGKSVGVVTTTRVQHASPAGAYAHTVNRNWYSDADVPASARQEGCQDIATQLISNMDIDVILGGGRKYMFPMGT
PDPEYPDDYSQGGTRLDGKNLVQEWLAKHQGARYVWNRTELLQASLDPSVTHLMGLFEPGDMKYEIHRDSTLDPSLM
EMTEAALLLLSRNPRGFFLFVEGGRIDHGHHESRAYRALTETIMFDDAIERAGQLTSEEDTLSLVTADHSHVFSFGG
YPLRGSSIFGLAPGKARDRKAYTVLLYGNGPGYVLKDGARPDVTESESGSPEYRQQSAVPLDGETHAGEDVAVFARG
PQAHLVHGVQEQTFIAHVMAFAACLEPYTACDLAPRAGTTDAAHPGPSVVPALLPLLAGTLLLLGTATAP

Seq. ID No: 565
>phosphatase_2
MQGACVLLLLGLHLQLSLGLVPVEEEDPAFWNRQAAQALDVAKKLQPIQTAAKNVILFLGDGMGVPTVTATRILKGQ
MNGKLGPETPLAMDQFPYVALSKTYNVDRQVPDSAGTATAYLCGVKGNYRTIGVSAAARYNQCKTTRGNEVTSVMNR
AKKAGKSVGVVTTTRVQHASPAGAYAHTVNRNWYSDADLPADAQMNGCQDIAAQLVNNMDIDVILGGGRKYMFPVGT
PDPEYPDDASVNGVRKRKQNLVQAWQAKHQGAQYVWNRTALLQAADDSSVTHLMGLFEPADMKYNVQQDHTKDPTLQ
EMTEVALRVVSRNPRGFYLFVEGGRIDHGHHDDKAYMALTEAGMFDNAIAKANELTSELDTLILVTADHSHVFSFGG
YTLRGTSIFGLAPSKALDSKSYTSILYGNGPGYALGGGSRPDVNDSTSEDPSYQQQAAVPQASETHGGEDVAVFARG
PQAHLVHGVEEETFVAHIMAFAGCVEPYTDCNLPAPTTATSIPDAAHLAASPPPLALLAGAMLLLLAPTLY

Seq. ID No: 566
>phosphatase_3
MQGPWVLLLLGLRLQLSLGVIPAEEENPAFWNRQAAEALDAAKKLQPIQKVAKNLILFLGDGLGVPTVTATRILKGQ
KNGKLGPETPLAMDRFPYLALSKTYNVDRQVPDSAATATAYLCGVKANFQTIGLSAAARFNQCNTTRGNEVISVMNR
AKQAGKSVGVVTTTRVQHASPAGTYAHTVNRNWYSDADMPASARQEGCQDIATQLISNMDIDVILGGGRKYMFPMGT
PDPEYPADASQNGIRLDGKNLVQEWLAKHQGAWYVWNRTELMQASLDQSVTHLMGLFEPGDTKYEIHRDPTLDPSLM
EMTEAALRLLSRNPRGFYLFVEGGRIDHGHHEGVAYQALTEAVMFDDAIERAGQLTSEEDTLTLVTADHSHVFSFGG
YTLRGSSIFGLAPSKAQDSKAYTSILYGNGPGYVFNSGVRPDVNESESGSPDYQQQAAVPLSSETHGGEDVAVFARG
PQAHLVHGVQEQSFVAHVMAFAACLEPYTACDLAPPACTTDAAHPVAASLPLLAGTLLLLGASAAP

Seq. ID No: 567
>phosphatase_4
MQGAWVLLLLGFRLQLSLSVIPVEEENPAFWTQKAADALNVAKKLQPIQTSAKNLIIFLGDGMGVATVTATRILKGQ
LEGNLGPETPLAMDHFPYMALSKTYSVDRQVPDSASTATAYLCGVKTNYKTIGVSAAARFDQCNTTFGNEVLSVMYR
AKKAGKSVGVGDHTRVQHASPAGTYVHTVTSNWYGDADMPALPLQEGCKDIATQLISNMDINVILGGGRKYMFPAGT
PDPEYPNDVNETGTRLDGKNLVQEWLSKHQGSQYVWNRQELIQKSLDPSVTYLMGLFEPVDTKFEIQRDPLMDPSLK
DMTEAALHVLSRNPKGFYLFVEGGRIDRGHHLGTAYLALTEAVMFDSAIERASLQASEQDTLTIVTADHSHVFSFGG
YTLRGTSIFGLAPLNALDGKPYTSILYGNGPGYVGTGERPNVTDAESHDPSYQQQAAVPVKSETTVGKDVAIFARGP
QAHLLHGVQEQNYIAHVMAFAGCLEPYTDCGLAPPADENRPTTPVQNSTTTTTTTTTTTTTTTRVQNSASSLGPA
TAPLAWHYWPRR

Seq. ID No: 568
>phosphatase_5
MSGSSVTGGGASLPAELYKGSADSILPANFSYAVTGSGTGKNAFLTNNSSLFGTTGTVHYAGSDSVLSGSELTTYNS
NYNGTYGPLIQIPSVATSVTVPYRKDGNTTLNLTSAQLCDAFSGAKTTWGQLLGTTDSTPIRIVYRTGSSGTTELFT
RHLNSICPTRFATNSTFTNARLPAGGTLPSNWVGVAATSTVVSTVKATNGSLGYVSPDAVNINSNAEVSRVNGNLPT
QANVSTALGSVAPPANAADRADPSKWVPVFTNPSAGYSIVGYTNFVFGQCYKDASVSTDVRAFINKHYGGTTTNAAV
AAHGFIPLTPAWKSAIVSAFYTGTSENLAIGNTNVCNTKGRP

Seq. ID No: 569
>phosphatase_6
MNFVTALPLIAQLIGTARAAIGPVTNLLVKNADIPPDGFTRAAVVANNQFPGPLITGNQGINFQINVVARLNNECSS
RPLWYHWHGFFQKGTNWADGPAFVNQCPISTGNSFLYDFTAADQAGTFWYHSHLSTQYCDGLRGPMVVYDPNDPHAS
LYDVDDESTVITLSDWYHTAARLGSAFPLGPDSVLINGLGRFAGGDGNAELAVIPVTQGKRYRFRLISLSCDPNFVF
SIDKHNMTVIEADAVSHEPVTVDSIHIYAGQRYSFVLSAHRDIDNYWIRALPSGGTVNFVGGVNSALIRYDGAAEVE
PVTNTTMSIAPLVETDLVPLDSPAAPGEASIGGVDYALSLVPSFVSRTLFCVRSIADDLRLRTGRTILSGSTELPSS
HPPCRVYTLPSNATIELSFPITATNAPGAPHPFHLHGHVFSVVRSAGSSEYNYANPPRRDVVNTGTAGDNVTIRFRV
CSHTYLSLHCHIDFHLEDGQSGTLVPPLPHRLPPRGRIRCRLHRGILVRGRLGPDLQ

Seq. ID No: 570
>scaffold
GGASASGVDNKFNKERVIAIGEIMRLPNLNSLQVVAFINSLRDDPSQSANLLAEAKKLNDAQAPKGGSSASSAGGSS
VDNKFNKEAQTAGVEIMELPNLNTRQLLAFIQSLRDDPSQSANLLAEAKKLNDAQAPKTSGGSSASSAGGSSVDNKF
NKERRMAAYEIIDLPNLNWFQLEAFITSLSDDPSQSANLLAEAKKLNDAQAPKGSGSNAAIRSSGSGSVDNKFNKER
VIAIGEIMRLPNLNSLQVVAFINSLRDDPSQSANLLAEAKKLNDAQAPKGGSSASSAGGSSVDNKFNKEAQTAGVEI
MELPNLNTRQLLAFIQSLRDDPSQSANLLAEAKKLNDAQAPKTSGSGSANAAIRSAGSGSVDNKFNKERRMAAYEII
DLPNLNWFQLEAFITSLSDDPSQSANLLAEAKKLNDAQAPKGSGASGSGAGGSSVDNKFNKERVIAIGEIMRLPNLN
SLQVVAFINSLRDDPSQSANLLAEAKKLNDAQAPKGSGASGSGAGGSSVDNKFNKEAQTAGVEIMELPNLNTRQLLA
FIQSLRDDPSQSANLLAEAKKLNDAQAPKTSGSGSANAAIRSAGSGSVDNKFNKERRMAAYEIIDLPNLNWFQLEAF
ITSLSDDPSQSANLLAEAKKLNDAQAPK

Seq. ID No: 571
>sec_1
MQLLRCFSIFSVIASVLAQELTTICEQIPSPTLESTPYSLSTTTILANGK

Seq. ID No: 572
>sec_2
MLSLKTLLCTLLTVSSVLATPVPARDPSSIQFVHEENKKRYYDYDHGSLGE

Seq. ID No: 573
>sec_3
MKLQSLLVSAAVLTSLTENVNAWSPNNSYVPANVTCDDDINLVREASGLSDNETEWLKKRDAYTKE

Seq. ID No: 574
>sec_4
MEGVSLEKREAEA

Seq. ID No: 575
>sec_5
MKKTAIAIAVALAGFATVAQA

Seq. ID No: 576
>vac_1
MFSLKALLPLALLLVSANQVAAKVHKAKIYKHELS

Seq. ID No: 577
>vac_2
MTKNFIVTLKKNTPDVEAKKFLDSVHHAGGSIVHKFDIIKGYTIKVPDVLHLNKLKEKHNDVIENVEEDKEVHTN

Seq. ID No: 578
>vac_3
MEEQREILEQLKKTLQMLTVEPSKNNQIANEEKEKKENENSWCILEHNYEDIAQEFIDFIYKNPTTYHVVSFFAELL
DKHNFKYLSEKSNWQDSIGEDGG

What is claimed is:

1. A non-naturally occurring nucleic acid comprising a sequence encoding an enzyme or regulatory protein in tryptamine metabolism,

wherein the enzyme or regulatory protein is an N-methyltransferase (INMT, PsiM, TrpM), a tryptophan decarboxylase (AADC), a tryptophan hydroxylase (TPH), a tryptamine 4' hydroxylase (T4H), a tryptamine 5' hydroxylase (T5H), a truncated cytochrome p450 reductase (T4H-CPR, T5H-CPR), an hydroxytryptamine O-methyltransferase (IOMT or CaffOMT), an N-acetyltransferase (NAT), a deacetylase (DAC), a hydroxyl tryptamine kinase (PsiK), a tryptophan synthase (TrpS), a toluene monooxygenase (TMO), an aminotransferase/methyltransferase fusion (ATMT), a phosphatase, an oxidase, a dimethylallyltryptophan synthase (DMATS), an isopentenyl-diphosphate isomerase (IDI1), a tryptophan halogenase (TrpHalo), an aspartate oxidase/quinolinic acid synthase fusion (AOQS), a tryptophan importer (TAT2), a methionine importer (MUP1), or a SAMe importer (SAM3).

2. The nucleic acid of claim 1, encoding a methyltransferase or hydroxylase.

3. The nucleic acid of claim 2, wherein the methyltransferase or hydroxylase is a tryptamine N-methyltransferase (INMT), a hydroxytryptamine O-methyltransferase (IOMT), a tryptamine 5' hydroxylase (T5H) or a tryptophan N-methyltransferase (TrpM).

4. The nucleic acid of claim 1, encoding an amino acid sequence that is naturally occurring.

5. The nucleic acid of claim 1, encoding an amino acid sequence that is not naturally occurring.

6. The nucleic acid of claim 1, wherein the sequence is codon-optimized for yeast expression.

7. The nucleic acid of claim 1, further comprising nucleotides encoding amino acids that are not part of the enzyme or regulatory protein.

8. The nucleic acid of claim 7, having a 5' end, wherein the additional nucleotides are at the 5' end of the nucleic acid and encode a codon optimized cofolding peptide.

9. The nucleic acid of claim 8, wherein the codon optimized cofolding peptide comprises an amino acid sequence of any one of SEQ ID NO:554-558.

10. The nucleic acid of claim 9, wherein the codon optimized cofolding peptide is encoded by any one of SEQ ID NOs:265-269.

11. The nucleic acid of claim 7, wherein the amino acids that are not part of the enzyme or regulatory protein are an affibody tag, a localization scaffold, a vacuolar localization tag, a secretion signal, a 6xhis tag, or any combination thereof.

12. The nucleic acid of claim 1, comprising the sequence of SEQ ID NOs:1-289

13. The nucleic acid of claim 1, further comprising a promoter functional in a recombinant microorganism.

14. The nucleic acid of claim 13, wherein the recombinant microorganism is a yeast.

15. An expression cassette comprising the nucleic acid of claim 13.

16. The expression cassette of claim 15, which is a yeast expression cassette.

17. A recombinant microorganism comprising the expression cassette of claim 15, that expresses the enzyme or regulatory protein encoded therein.

18. The recombinant microorganism of claim 17, which is an E. coli.

19. The recombinant microorganism of claim 17, which is a yeast cell.

20. The yeast cell of claim 19, which is a species of Saccharomyces, Candida, Pichia, Schizosaccharomyces, Scheffersomyces, Blakeslea, Rhodotorula, Aspergillus or Yarrowia.

21. The yeast cell of claim 19, which is a Saccharomyces cerevisiae.

22. A non-naturally occurring enzyme or regulatory protein comprising an amino acid sequence encoded by the nucleic acid of any one of claims 1-12.

23. A recombinant microorganism expressing at least one enzyme or regulatory protein of claim 22.

24. The recombinant microorganism of claim 23, which is an E. coli.

25. The recombinant microorganism of claim 23, which is a yeast cell.

26. The yeast cell of claim 25, which is a species of Saccharomyces, Candida, Pichia, Schizosaccharomyces, Scheffersomyces, Blakeslea, Rhodotorula, Aspergillus or Yarrowia.

27. The yeast cell of claim 25, which is a Saccharomyces cerevisiae.

28. The recombinant microorganism of claim 23, expressing INMT, wherein the recombinant microorganism produces at least one hydroxy substituted tryptophan compound.

29. The recombinant microorganism of claim 28, wherein the at least one hydroxy substituted tryptophan compound is 5-OH-NMTP, 5-OH-DMTP or 5-OH-TMTP.

30. The recombinant microorganism of claim 23, expressing INMT, wherein the recombinant microorganism produces at least one hydroxy substituted tryptamine compound.

31. The recombinant microorganism of claim 30, wherein the at least one hydroxy substituted tryptamine compound is bufotenine, 5-OH-NMT, or 5-OH-TMT.

32. The recombinant microorganism of claim 23, expressing INMT, wherein the recombinant microorganism produces at least one methoxy substituted tryptamine compound.

33. The recombinant microorganism of claim 32, wherein the at least one methoxy substituted tryptamine compound is 5-MeO-NMT, 5-MeO-DMT, or 5-MeO-TMT.

34. The recombinant microorganism of claim 23, expressing IOMT, wherein the recombinant microorganism methylates the primary amine on the 5-hydroxy moiety on an indole ring.

35. The recombinant microorganism of claim 34, wherein the microorganism acts on (a) bufotenine to create 5-MeO-DMT, or (b) N-acetylserotonin to create melatonin.

36. The recombinant microorganism of claim 23, expressing T5H, wherein the recombinant microorganism hydroxylates at the 5' position of an indole ring.

37. The recombinant microorganism of claim 36, wherein the microorganism generates serotonin from tryptamine.

38. The recombinant microorganism of claim 36, wherein the T5H is a fusion polypeptide with a cytochrome P450 reductase (CPR).

39. The recombinant microorganism of claim 36, wherein the T5H is a fusion polypeptide with an IOMT.

40. The recombinant microorganism of claim 23, expressing TrpM, wherein the recombinant microorganism catalyzes the alkylation of the primary amine of L-tryptophan to produce NMTP, DMTP, TMTP, or any combination thereof.

41. The recombinant microorganism of claim 23, expressing PsiM, wherein the recombinant microorganism methylates norbaeocystin.

42. The recombinant microorganism of claim 41, wherein the PsiM comprises a domain from an rRNA methyltransferase from Ascomycota.

43. The recombinant microorganism of claim 23, expressing AADC, wherein the recombinant microorganism decarboxylates an aliphatic carboxylic acid.

44. The recombinant microorganism of claim 43, wherein the recombinant microorganism creates tryptamine from L-tryptophan, creates serotonin from 5-HTP, creates bufotenine from 5-OH-DMTP, creates 5-MeO-DMT from 5-MeO-DMTP, or any combination thereof.

45. The recombinant microorganism of claim 23, expressing TPH, wherein the recombinant microorganism adds a hydroxy group to the 5-carbon of L-tryptophan.

46. The recombinant microorganism of claim 23, expressing T4H, wherein the recombinant microorganism hydroxylates the 4' position of an indole ring.

47. The recombinant microorganism of claim 46, wherein the recombinant microorganism converts tryptamine to 4-OH-tryptamine.

48. The recombinant microorganism of claim 46, wherein the T4H is a chimera of sequences from T4H from different species, wherein the T4H comprises a yeast p450 N terminus.

49. The recombinant microorganism of claim 46, wherein the T4H is a chimera of a mushroom PsiH and a yeast p450 N terminus.

50. The recombinant microorganism of claim 23, expressing NAT, wherein the recombinant microorganism adds an acetyl group from acetyl-CoA to the terminal amino group of a tryptamine.

51. The recombinant microorganism of claim 50, wherein the recombinant microorganism acts on serotonin to generate N-acetylserotonin.

52. The recombinant microorganism of claim 23, expressing DAC, wherein the recombinant microorganism removes an acetyl group from the terminal amino group of a tryptamine.

53. The recombinant microorganism of claim 52, wherein the recombinant microorganism acts on melatonin to create 5-MeO-tryptamine.

54. The recombinant microorganism of claim 23, expressing PsiK, wherein the recombinant microorganism phosphorylates a hydroxy-indole.

55. The recombinant microorganism of claim 54, wherein the PsiK is a chimera of a PsiK and a yeast kinase.

56. The recombinant microorganism of claim 23, expressing TrpS, wherein the recombinant microorganism combines an indole with L-serine or L-threonine to create variants of tryptophan or beta-methyl tryptophan, respectively.

57. The recombinant microorganism of claim 56, wherein the TrpS is coexpressed with a multidrug exporter, wherein the recombinant microorganism exports indole while continuing bioproduction of tryptophan and/or tryptamine analogs.

58. The recombinant microorganism of claim 23, expressing TMO, wherein the recombinant microorganism hydroxylates the indole ring of tryptamines.

59. The recombinant microorganism of claim 58, wherein the TMO comprises four subunits that are fused into a fusion polypeptide.

60. The recombinant microorganism of claim 23, expressing a phosphatase, wherein the recombinant microorganism dephosphorylates a phosphorylated tryptamine.

61. The recombinant microorganism of claim 60, wherein the recombinant microorganism dephosphorylates psilocybin into psilosin.

62. The recombinant microorganism of claim 23, expressing an oxidate, wherein the recombinant microorganism creates a tryptamine radical which reacts with another tryptamine to form a dimer or oligomer.

63. The recombinant microorganism of claim 62, wherein the oxidase is a chimera with a yeast laccase.

64. The recombinant microorganism of claim 63, wherein the oxidase is coexpressed with a yeast t-SNARE.

65. The recombinant microorganism of claim 23, expressing a DMATS, wherein the recombinant microorganism prenylates a tryptophan and/or a tryptamine.

66. The recombinant microorganism of claim 65, wherein the DMATS is a fusion polypeptide with IDI1.

67. The recombinant microorganism of claim 23, expressing a TrpHalo, wherein the recombinant microorganism adds fluorine (F), chlorine (Cl), bromine (Br), and/or iodine (I) to an indole or biogenic amine.

68. The recombinant microorganism of claim 67, wherein the TrpHalo further comprises a secretion tag.

69. The recombinant microorganism of claim 68, wherein the TrpHalo further comprises a 6xhis tag.

70. The recombinant microorganism of claim 67, wherein the TrpHalo is coexpressed with a fluoride exporter.

71. The recombinant microorganism of claim 23, expressing more than one of the enzyme and/or regulatory protein.

72. The recombinant microorganism of claim 71, expressing TPH, TrpM, and AADC, wherein the recombinant microorganism produces at least one hydroxy substituted tryptamine compound.

73. The recombinant microorganism of claim 72, wherein the at least one hydroxy substituted tryptamine compound is bufotenine, 5-OH-NMT, or 5-OH-TMT.

74. The recombinant microorganism of claim 71, expressing TPH, TrpM, AADC, and IOMT, wherein the recombinant microorganism produces at least one methoxy substituted tryptamine compound.

75. The recombinant microorganism of claim 74, wherein the at least one methoxy substituted tryptamine compound is 5-MeO-NMT, 5-MeO-DMT, or 5-MeO-TMT.

76. The recombinant microorganism of claim 71, expressing AADC, T5H and INMT, wherein the recombinant microorganism produces at least one hydroxy substituted tryptamine compound.

77. The recombinant microorganism of claim 76, wherein the at least one hydroxy substituted tryptamine compound is bufotenine, 5-OH-NMT, or 5-OH-TMT.

78. The recombinant microorganism of claim 71, expressing AADC, T5H, INMT, and IOMT, wherein the recombinant microorganism produces at least one methoxy substituted tryptamine compound.

79. The recombinant microorganism of claim 78, wherein the at least one methoxy substituted tryptamine compound is 5-MeO-NMT, 5-MeO-DMT, or 5-MeO-TMT.

80. The recombinant microorganism of claim 71, expressing TrpM and TPH, wherein the recombinant microorganism produces at least one hydroxy substituted tryptophan compound.

81. The recombinant microorganism of claim 80, wherein the at least one hydroxy substituted tryptophan compound is 5-HTP, 5-OH-NMTP, 5-OH-DMTP or 5-OH-TMTP.

82. The recombinant microorganism of claim 71, expressing TrpM, TPH and IOMT, wherein the recombinant microorganism produces at least one methoxy substituted tryptophan compound.

83. The recombinant microorganism of claim 82, wherein the at least one methoxy substituted tryptophan compound is 5-MeO-NMTP, 5-MeO-DMTP or 5-MeO-TMTP.

84. The recombinant microorganism of claim 71, expressing INMT and T5H, wherein the recombinant microorganism produces at least one hydroxy substituted tryptamine compound.

85. The recombinant microorganism of claim 84, wherein the at least one hydroxy substituted tryptamine compound is bufotenine, 5-OH-NMT, or 5-OH-TMT.

86. The recombinant microorganism of claim 71, expressing INMT, T5H and IOMT, wherein the recombinant microorganism produces at least one methoxy substituted tryptamine compound.

87. The recombinant microorganism of claim 86, wherein the at least one methoxy substituted tryptamine compound is 5-MeO-NMT, 5-MeO-DMT, or 5-MeO-TMT.

88. The recombinant microorganism of claim 71, expressing INMT and IOMT, wherein the recombinant microorganism produces at least one methoxy substituted tryptophan compound.

89. The recombinant microorganism of claim 88, wherein the at least one methoxy substituted tryptophan compound is 5-MeO-NMTP, 5-MeO-DMTP or 5-MeO-TMTP.

90. The recombinant microorganism of claim 71, expressing INMT and AADC, wherein the recombinant microorganism produces at least one hydroxy substituted tryptamine compound.

91. The recombinant microorganism of claim 90, wherein the at least one hydroxy substituted tryptamine compound is bufotenine, 5-OH-NMT, or 5-OH-TMT.

92. The recombinant microorganism of claim 71, expressing INMT, AADC and IOMT, wherein the recombinant microorganism produces at least one methoxy substituted tryptamine compound.

93. The recombinant microorganism of claim 92, wherein the at least one methoxy substituted tryptamine compound is 5-MeO-NMT, 5-MeO-DMT, or 5-MeO-TMT.

94. The recombinant microorganism of claim 71, expressing INMT and IOMT, wherein the recombinant microorganism produces at least one methoxy substituted tryptamine compound.

95. The recombinant microorganism of claim 94, wherein the at least one methoxy substituted tryptamine compound is 5-MeO-NMT, 5-MeO-DMT, or 5-MeO-TMT.

96. The recombinant microorganism of claim 71, expressing TPH and AADC, wherein the recombinant microorganism generates serotonin from L-tryptophan through a 5-HTP intermediate.

97. The recombinant microorganism of claim 71, expressing NAT and IOMT, wherein the recombinant microorganism generates melatonin from serotonin through an N-acetylserotonin intermediate.

98. The recombinant microorganism of claim 71, expressing DAC and INMT, wherein the recombinant microorganism generates 5-MeO-DMT from melatonin through a 5-MeO-tryptamine intermediate.

99. The recombinant microorganism of claim 71, expressing ATMT and AADC, wherein ATMT is a fusion polypeptide of aminotransferase and methyltransferase, wherein the recombinant microorganism produces meta-methylated tryptamine analogs.

100. The recombinant microorganism of claim 71, wherein the recombinant microorganism does not express Pdc5, Aro10, Aro7, Pdz1, Pdz2, Bna2, SPE2, Glc3 or any combination thereof.

101. The recombinant microorganism of claim 71, wherein the recombinant microorganism overexpresses an enzyme that promotes conversion of L-methionine to SAMe, does not express an off-pathway gene that encodes for enzymes that deplete SAMe for unwanted side products, overexpresses a permease, or any combination thereof.

102. The recombinant microorganism of claim 71, wherein the recombinant microorganism overexpresses a Sam2, Adk1, Mup1, Sam3.

103. The nucleic acid of claim 1, further comprising a promoter functional in a plant.

104. A plant expression cassette comprising the nucleic acid of claim 103.

105. A recombinant plant comprising the plant expression cassette of claim 104, capable of expressing the enzyme or regulatory protein encoded therein.

106. The recombinant plant of claim 105, expressing TrpM and TPH, wherein the recombinant plant produces at least one hydroxy substituted tryptophan compound.

107. The recombinant plant of claim 106, wherein the at least one hydroxy substituted tryptophan compound is 5-HTP, 5-OH-NMTP, 5-OH-DMTP or 5-OH-TMTP.

108. The recombinant plant of claim 105, expressing TrpM, TPH and IOMT, wherein the recombinant plant produces at least one methoxy substituted tryptophan compound.

109. The recombinant plant of claim 105, which is a tobacco or *Arabadopsis* plant.

110. A method of producing a substituted indole, the method comprising

(i) growing the recombinant microorganism of any one of claims 23-102 or the recombinant plant of any one of claims 105-109;

(ii) expressing the at least one enzyme or regulatory protein in the recombinant microorganism;

(iii) producing or synthesizing substituted indoles and tryptamines in the recombinant microorganism; and

(iv) isolating the substituted indole from the recombinant microorganism.

111. The method of claim 110, wherein precursor chemicals are added to the growing recombinant microorganism or plant, wherein the precursor chemicals are utilized by the at least one enzyme or regulatory protein.

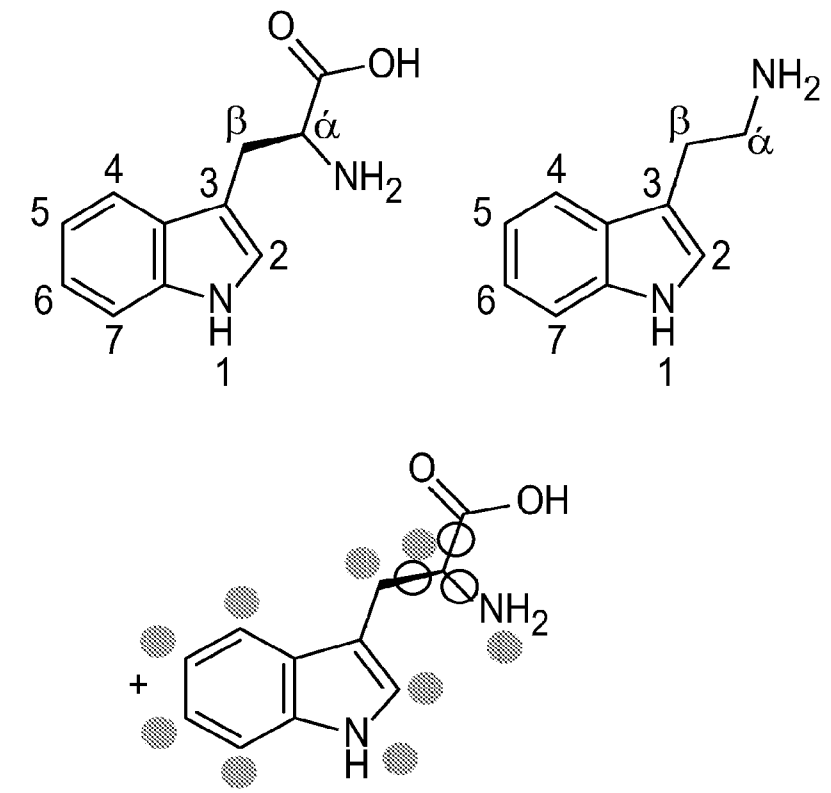

O
OH
β
α
NH2
4
3
5
2
6
N
H
7
1
NH2
β
α
4
3
5
2
6
N
H
7
1
O
OH
NH2
+
N
H
methylation
halogenation
prenylation
hydroxylation
O-phosphorylation
O-methylation
N-Acetylation
cleavage
deamination
decarboxylation
+ oligomerization

FIG. 1

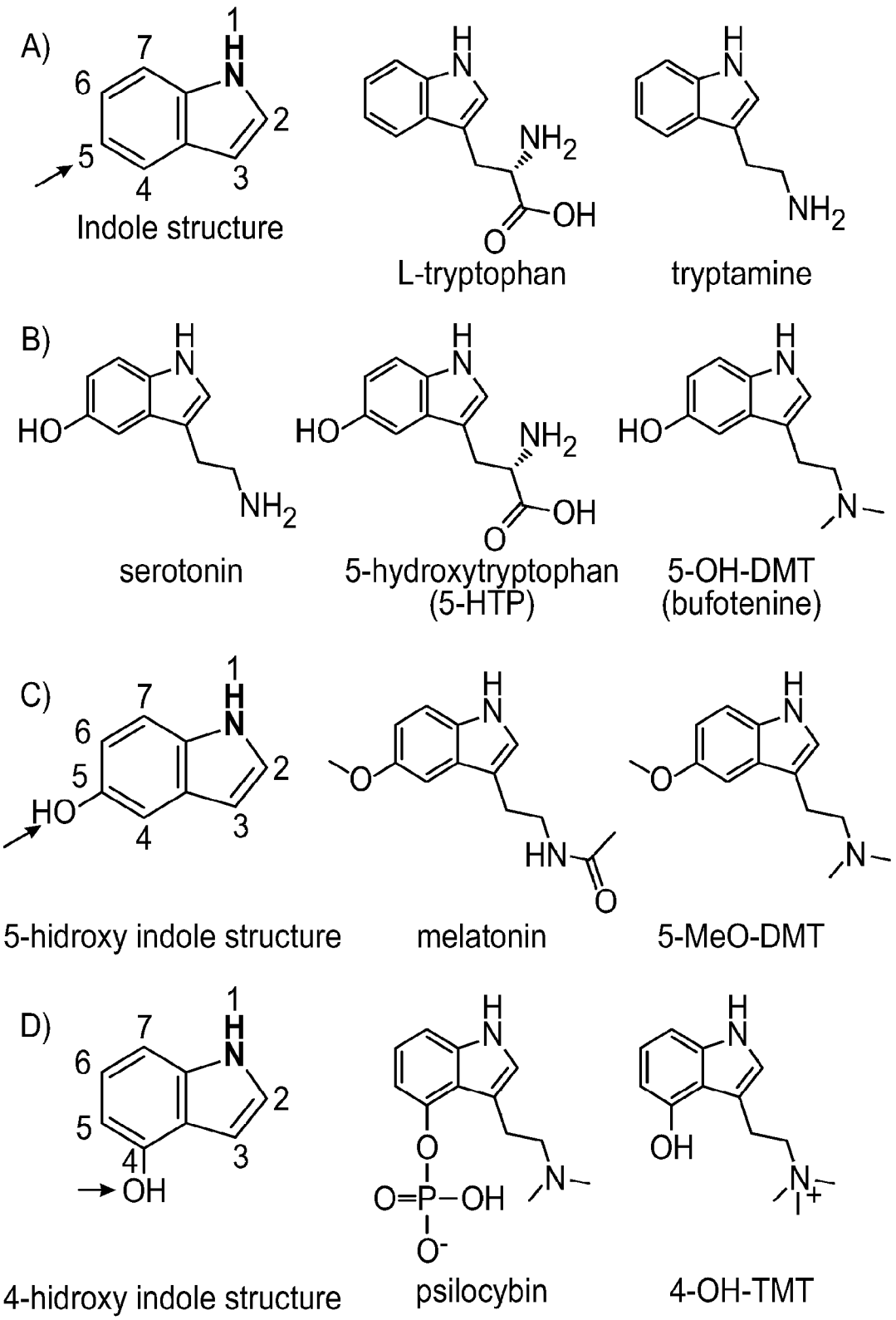

A)
Indole structure
L-tryptophan
tryptamine
B)
serotonin
5-hydroxytryptophan
(5-HTP)
5-OH-DMT
(bufotenine)
C)
5-hidroxy indole structure
melatonin
5-MeO-DMT
D)
4-hidroxy indole structure
psilocybin
4-OH-TMT

FIG. 2

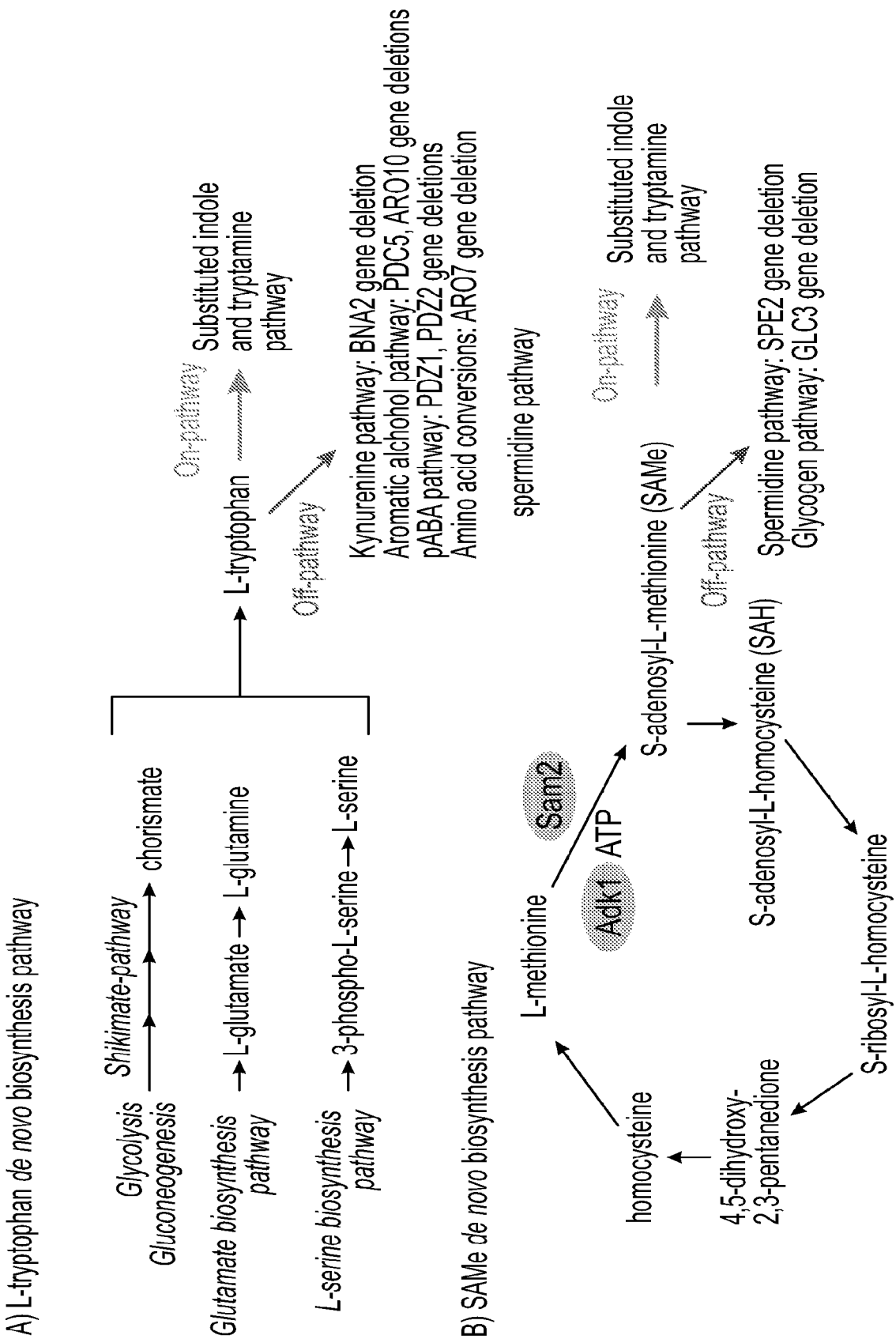

A) L-tryptophan de novo biosynthesis pathway
Glycolysis
Gluconeogenesis
Shikimate-pathway
chorismate
Glutamate biosynthesis pathway
L-glutamate
L-glutamine
L-serine biosynthesis pathway
3-phospho-L-serine
L-serine
L-tryptophan
On-pathway
Substituted indole and tryptamine pathway
Off-pathway
Kynurenine pathway: BNA2 gene deletion
Aromatic alchohol pathway: PDC5, ARO10 gene deletions
pABA pathway: PDZ1, PDZ2 gene deletions
Amino acid conversions: ARO7 gene deletion
B) SAMe de novo biosynthesis pathway
spermidine pathway
L-methionine
Sam2
Adk1
ATP
S-adenosyl-L-methionine (SAMe)
On-pathway
Substituted indole and tryptamine pathway
Off-pathway
homocysteine
S-adenosyl-L-homocysteine (SAH)
Spermidine pathway: SPE2 gene deletion
Glycogen pathway: GLC3 gene deletion
4,5-dihydroxy-2,3-pentanedione
S-ribosyl-L-homocysteine

FIG. 3

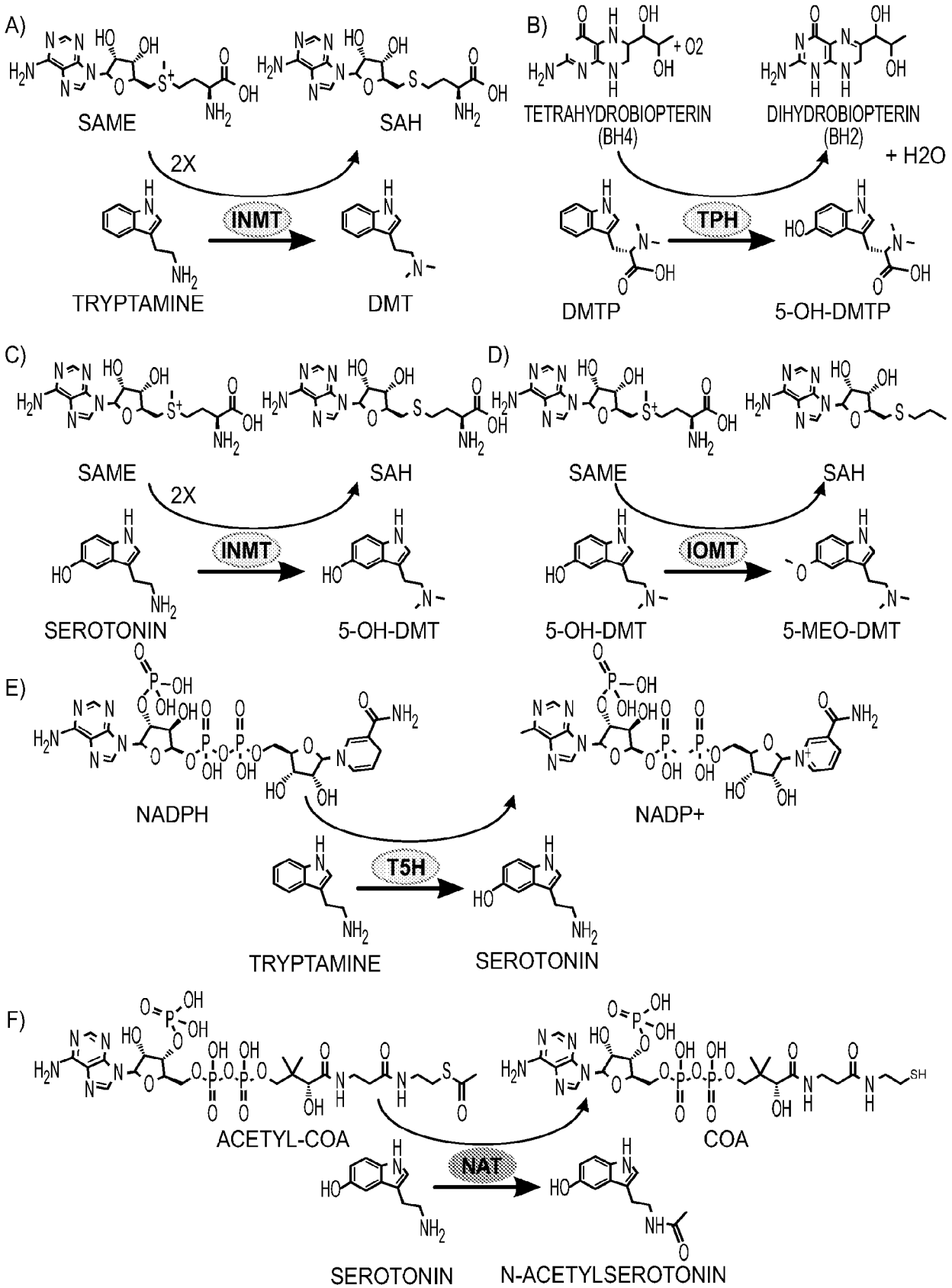

A)
SAME
SAH
2X
INMT
TRYPTAMINE
DMT
B)
TETRAHYDROBIOPTERIN (BH4)
DIHYDROBIOPTERIN (BH2)
+ O2
+ H2O
TPH
DMTP
5-OH-DMTP
C)
SAME
SAH
2X
INMT
SEROTONIN
5-OH-DMT
D)
SAME
SAH
IOMT
5-OH-DMT
5-MEO-DMT
E)
NADPH
NADP+
T5H
TRYPTAMINE
SEROTONIN
F)
ACETYL-COA
COA
NAT
SEROTONIN
N-ACETYLSEROTONIN

FIG. 4

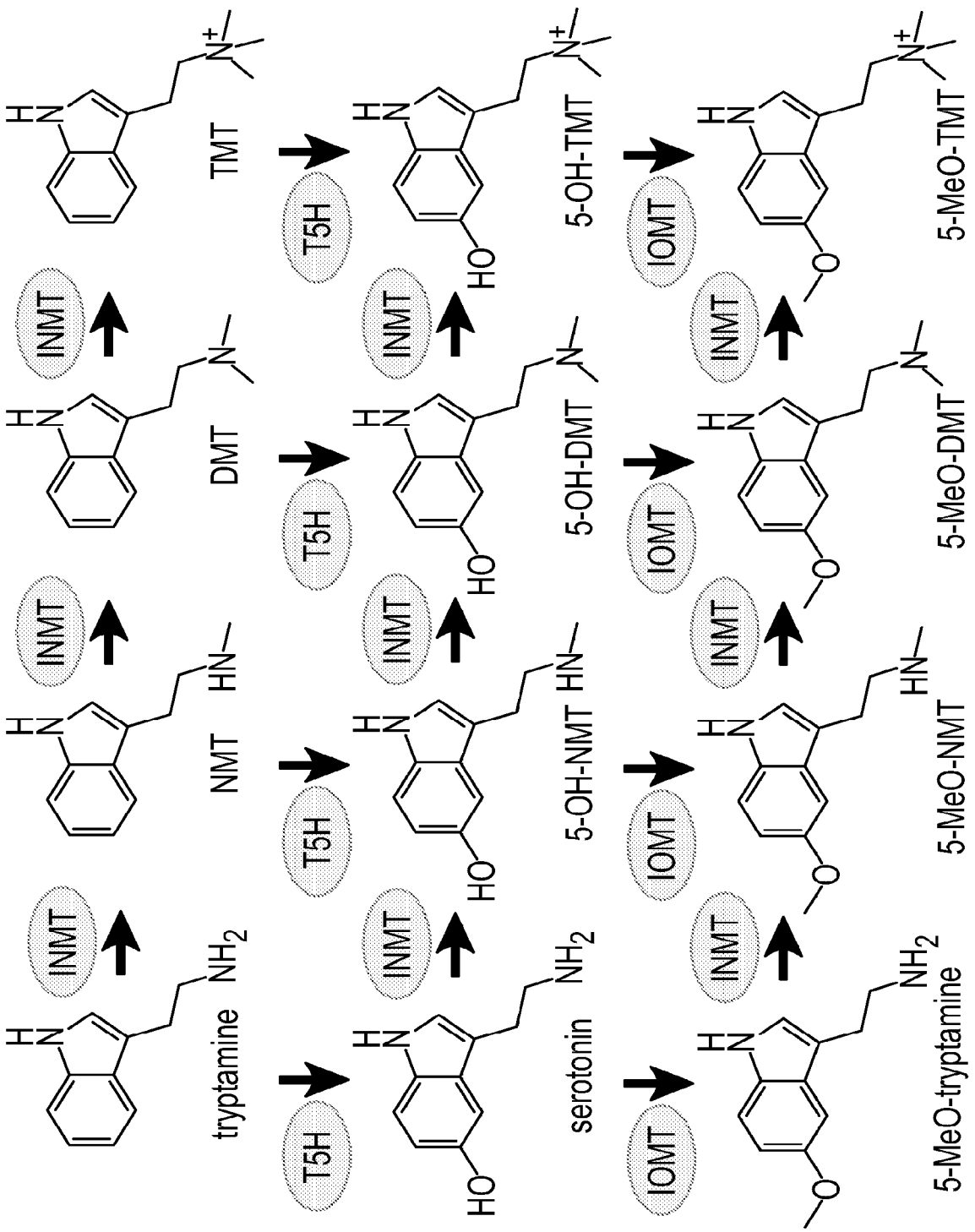

tryptamine
NMT
DMT
TMT
serotonin
5-OH-NMT
5-OH-DMT
5-OH-TMT
5-MeO-tryptamine
5-MeO-NMT
5-MeO-DMT
5-MeO-TMT
INMT
T5H
IOMT

FIG. 5

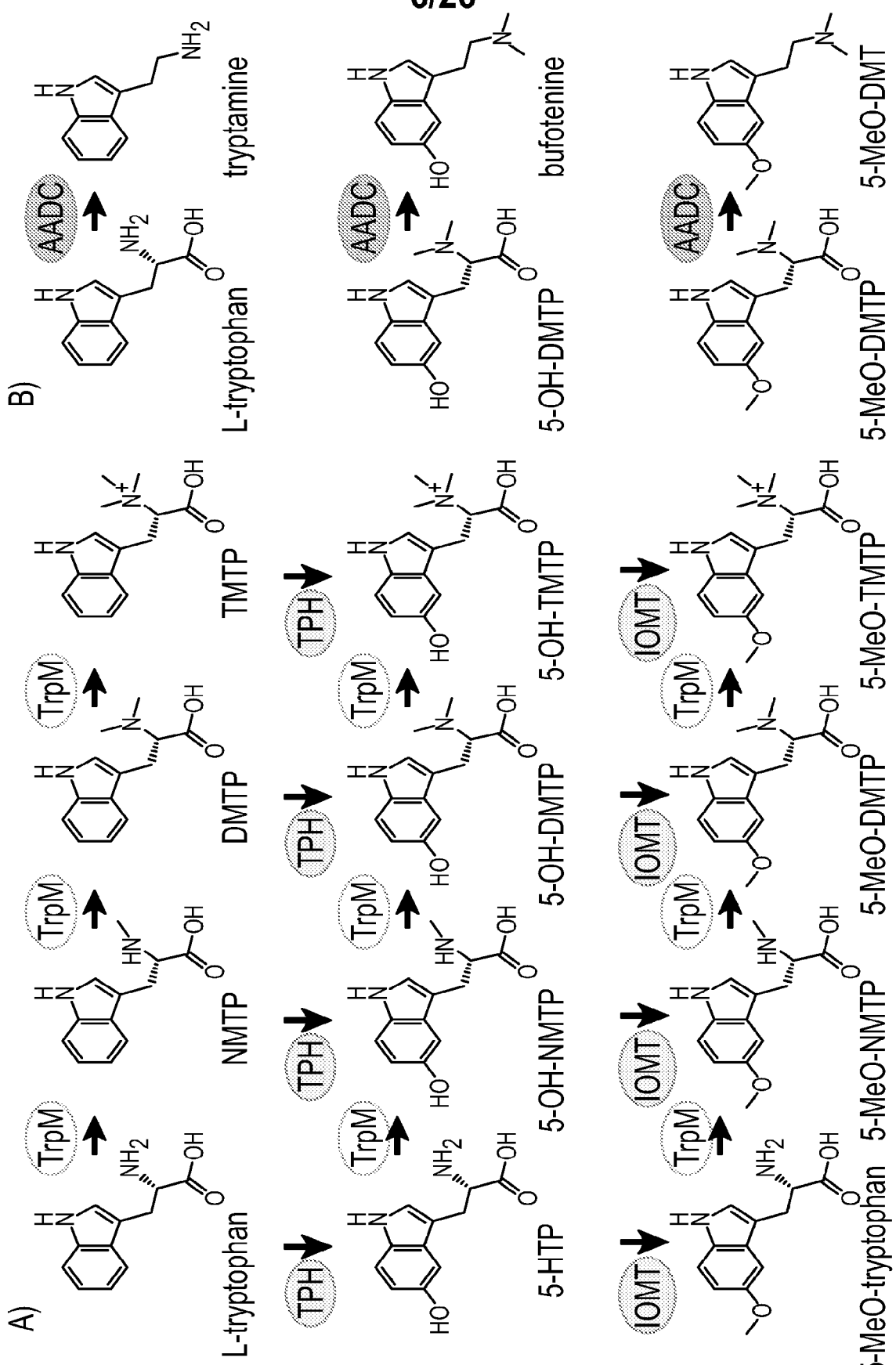

A)
L-tryptophan
TrpM
NMTP
TrpM
DMTP
TrpM
TMTP
TPH
5-HTP
TrpM
5-OH-NMTP
TrpM
5-OH-DMTP
TrpM
5-OH-TMTP
IOMT
5-MeO-tryptophan
TrpM
5-MeO-NMTP
TrpM
5-MeO-DMTP
TrpM
5-MeO-TMTP
B)
L-tryptophan
AADC
tryptamine
5-OH-DMTP
AADC
bufotenine
5-MeO-DMTP
AADC
5-MeO-DMT

FIG. 6

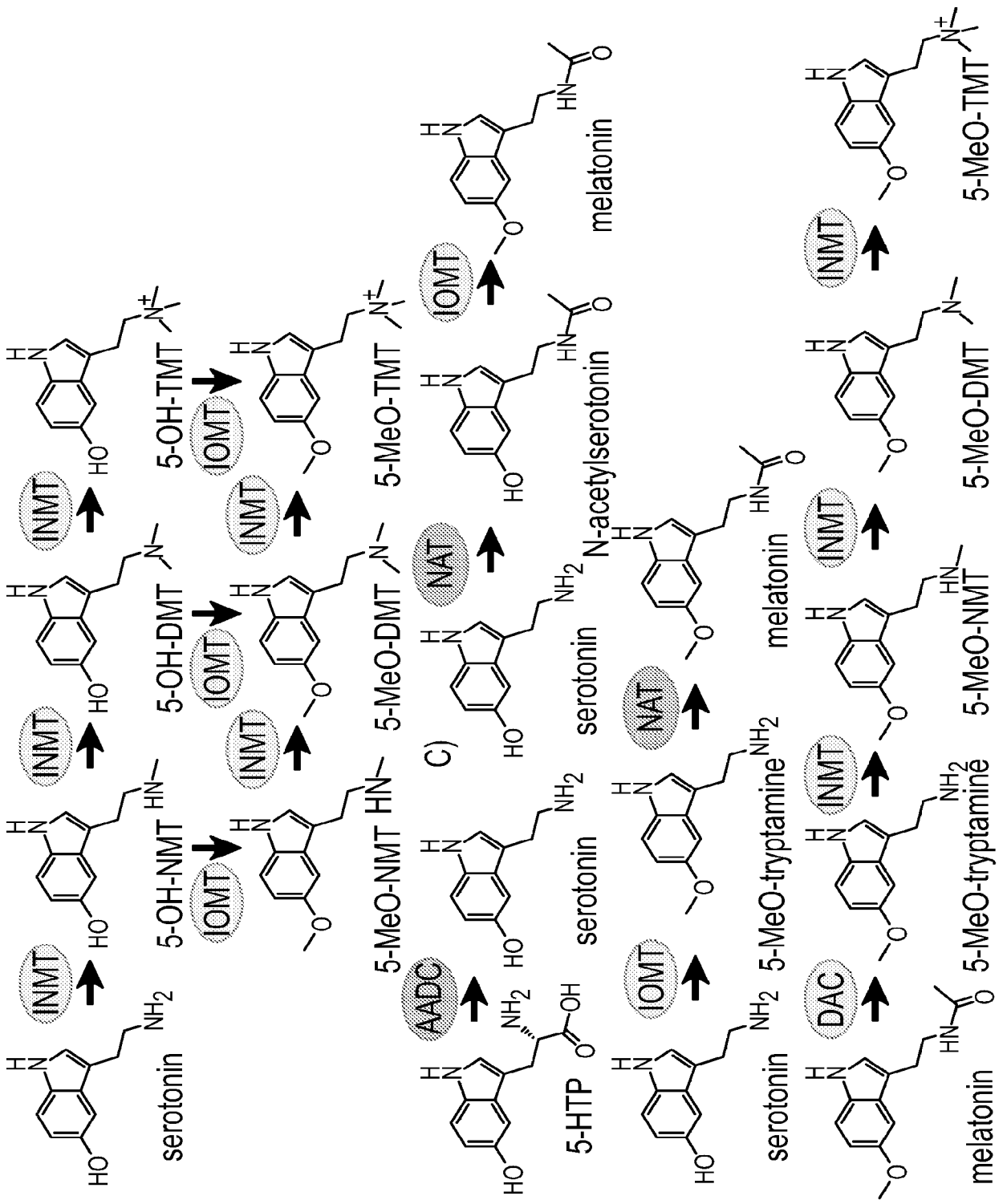

A)
serotonin
INMT
5-OH-NMT
INMT
5-OH-DMT
INMT
5-OH-TMT
IOMT
5-MeO-NMT
IOMT
5-MeO-DMT
IOMT
5-MeO-TMT
B)
5-HTP
AADC
serotonin
NAT
N-acetylserotonin
IOMT
melatonin
C)
D)
serotonin
IOMT
5-MeO-tryptamine
NAT
melatonin
E)
serotonin
DAC
5-MeO-tryptamine
INMT
5-MeO-NMT
INMT
5-MeO-DMT
INMT
5-MeO-TMT
melatonin

FIG. 7

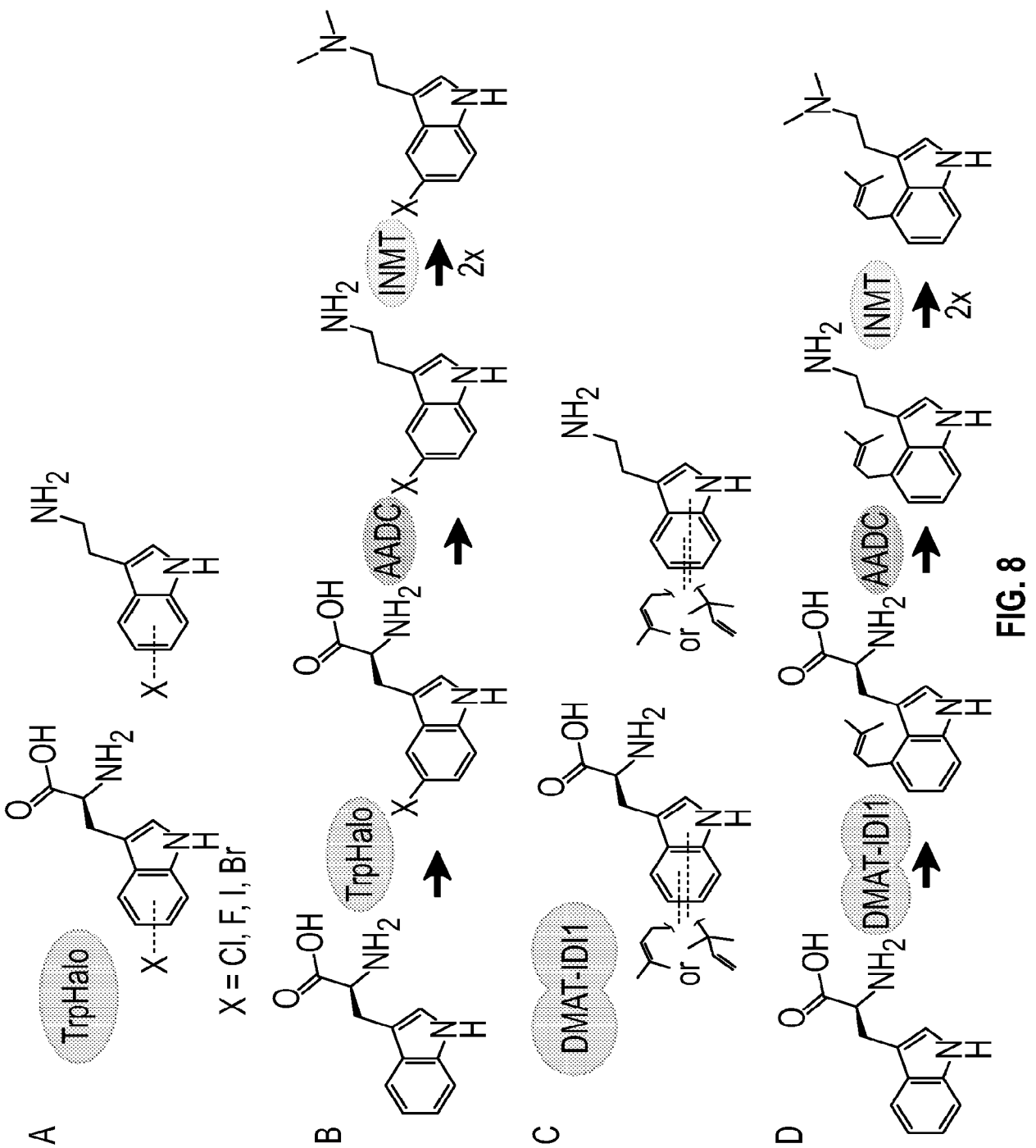

A
TrpHalo
X = Cl, F, I, Br
B
TrpHalo
AADC
INMT
2x
C
DMAT-IDI1
or
D
DMAT-IDI1
AADC
INMT
2x

FIG. 8

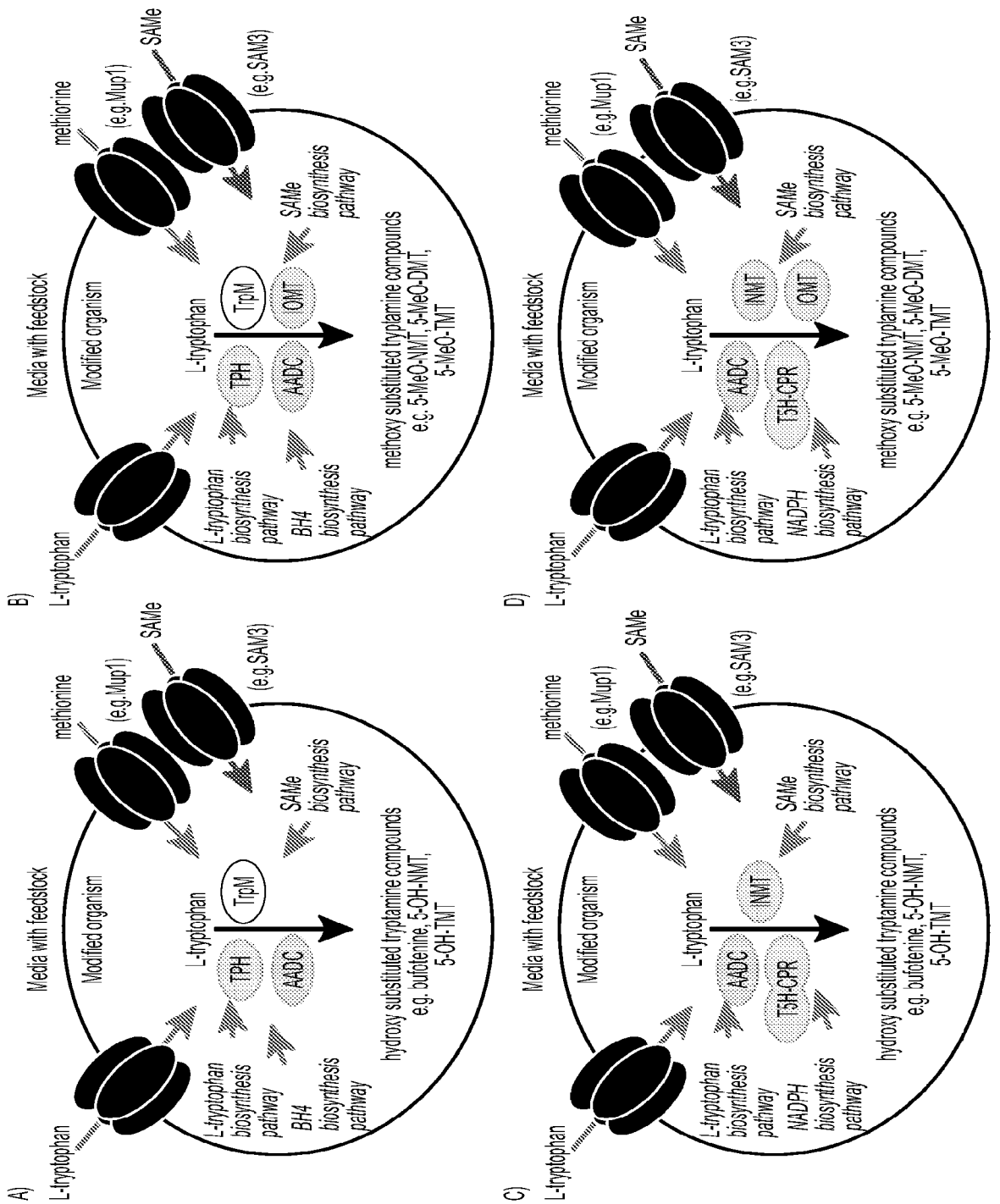

A)
L-tryptophan
Media with feedstock
Modified organism
methionine
(e.g.Mup1)
SAMe
(e.g.SAM3)
L-tryptophan
TPH
TrpM
AADC
L-tryptophan biosynthesis pathway
BH4 biosynthesis pathway
SAMe biosynthesis pathway
hydroxy substituted tryptamine compounds
e.g. bufotenine, 5-OH-NMT,
5-OH-TMT
B)
L-tryptophan
Media with feedstock
Modified organism
methionine
(e.g.Mup1)
SAMe
(e.g.SAM3)
L-tryptophan
TPH
TrpM
AADC
OMT
L-tryptophan biosynthesis pathway
BH4 biosynthesis pathway
SAMe biosynthesis pathway
methoxy substituted tryptamine compounds
e.g. 5-MeO-NMT, 5-MeO-DMT,
5-MeO-TMT
C)
L-tryptophan
Media with feedstock
Modified organism
methionine
(e.g.Mup1)
SAMe
(e.g.SAM3)
L-tryptophan
AADC
INMT
T5H-CPR
L-tryptophan biosynthesis pathway
NADPH biosynthesis pathway
SAMe biosynthesis pathway
hydroxy substituted tryptamine compounds
e.g. bufotenine, 5-OH-NMT,
5-OH-TMT
D)
L-tryptophan
Media with feedstock
Modified organism
methionine
(e.g.Mup1)
SAMe
(e.g.SAM3)
L-tryptophan
AADC
INMT
IOMT
T5H-CPR
L-tryptophan biosynthesis pathway
NADPH biosynthesis pathway
SAMe biosynthesis pathway
methoxy substituted tryptamine compounds
e.g. 5-MeO-NMT, 5-MeO-DMT,
5-MeO-TMT

FIG. 9

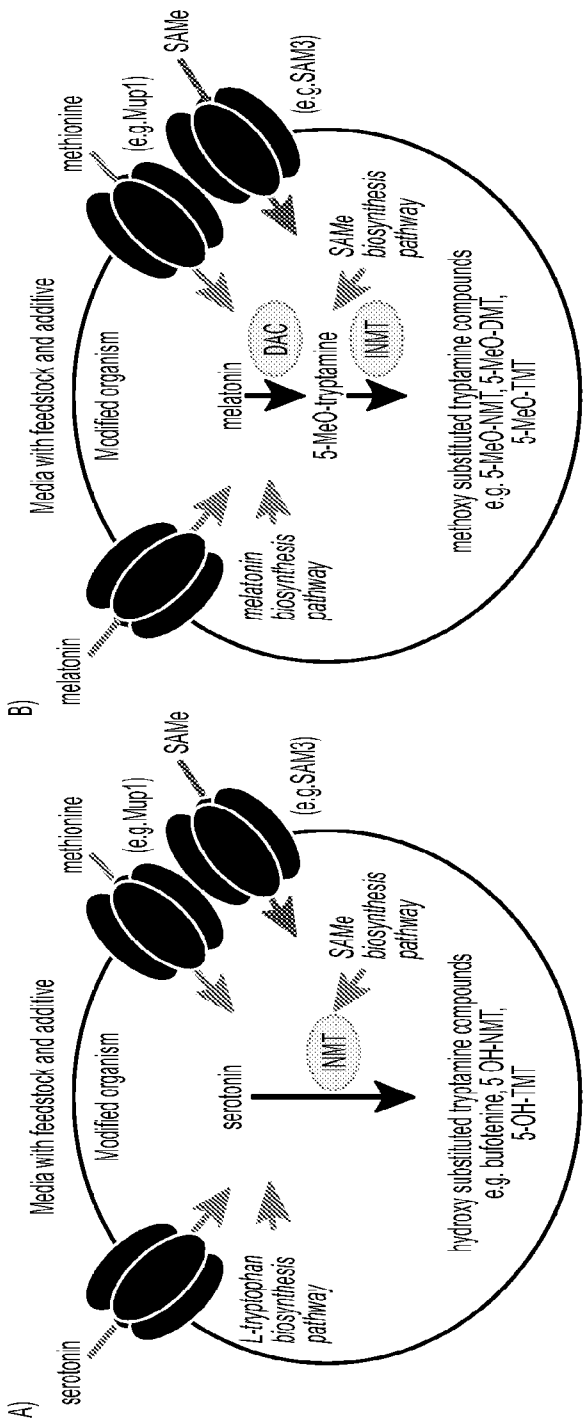

A)
serotonin
L-tryptophan biosynthesis pathway
Media with feedstock and additive
Modified organism
serotonin
NMT
hydroxy substituted tryptamine compounds
e.g. bufotenine, 5 OH-NMT,
5-OH-TMT
SAMe biosynthesis pathway
methionine
(e.g.Mup1)
SAMe
(e.g.SAM3)
B)
melatonin
melatonin biosynthesis pathway
Media with feedstock and additive
Modified organism
melatonin
DAC
5-MeO-tryptamine
NMT
methoxy substituted tryptamine compounds
e.g. 5-MeO-NMT, 5-MeO-DMT,
5-MeO-TMT
SAMe biosynthesis pathway
methionine
(e.g.Mup1)
SAMe
(e.g.SAM3)

FIG. 10

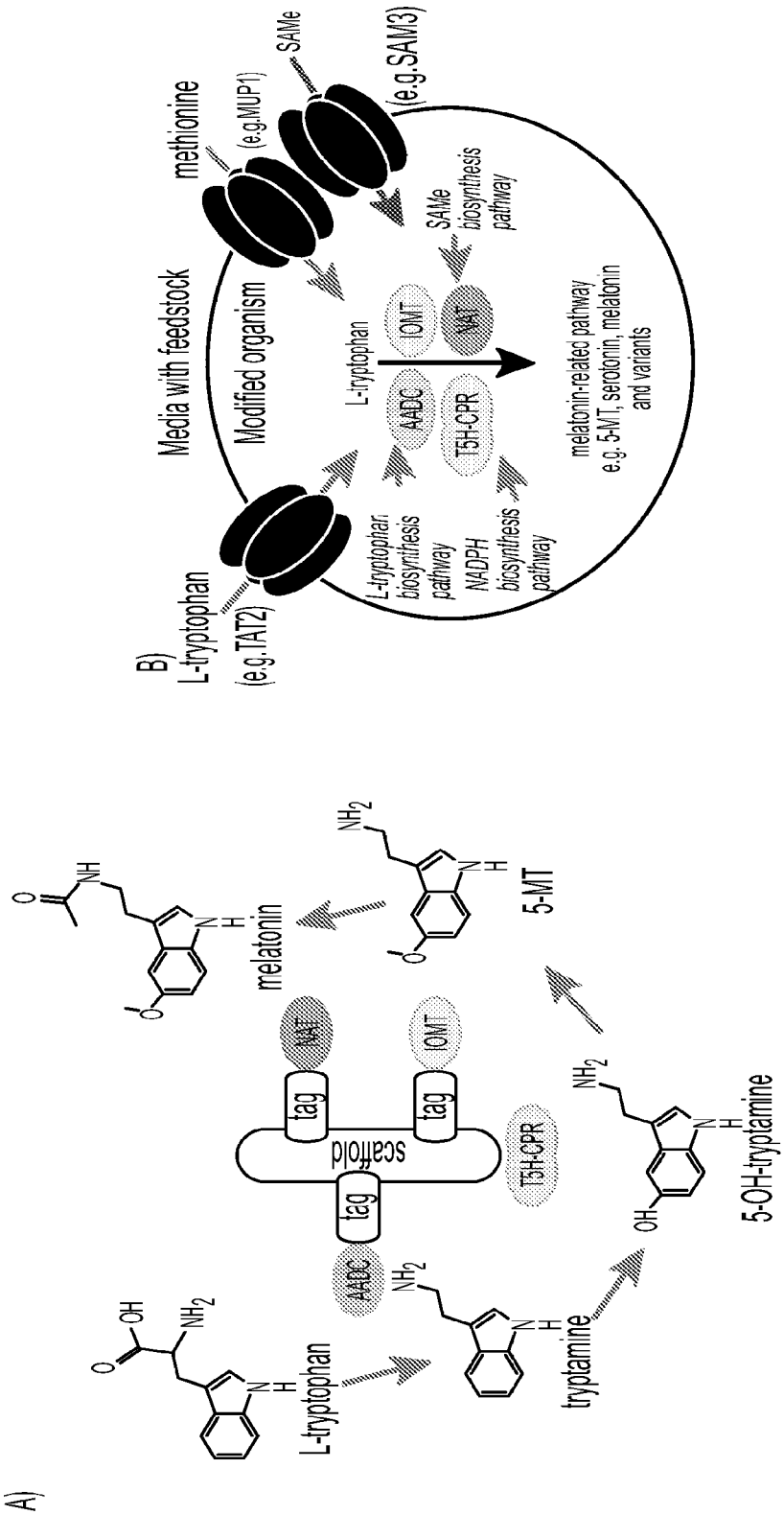

A)
L-tryptophan
tryptamine
5-OH-tryptamine
5-MT
melatonin
AADC
T5H-CPR
IOMT
NAT
tag
scaffold
B)
L-tryptophan
(e.g.TAT2)
Media with feedstock
Modified organism
methionine
(e.g.MUP1)
SAMe
(e.g.SAM3)
L-tryptophan biosynthesis pathway
NADPH biosynthesis pathway
SAMe biosynthesis pathway
melatonin-related pathway
e.g. 5-MT, serotonin, melatonin and variants

FIG. 11

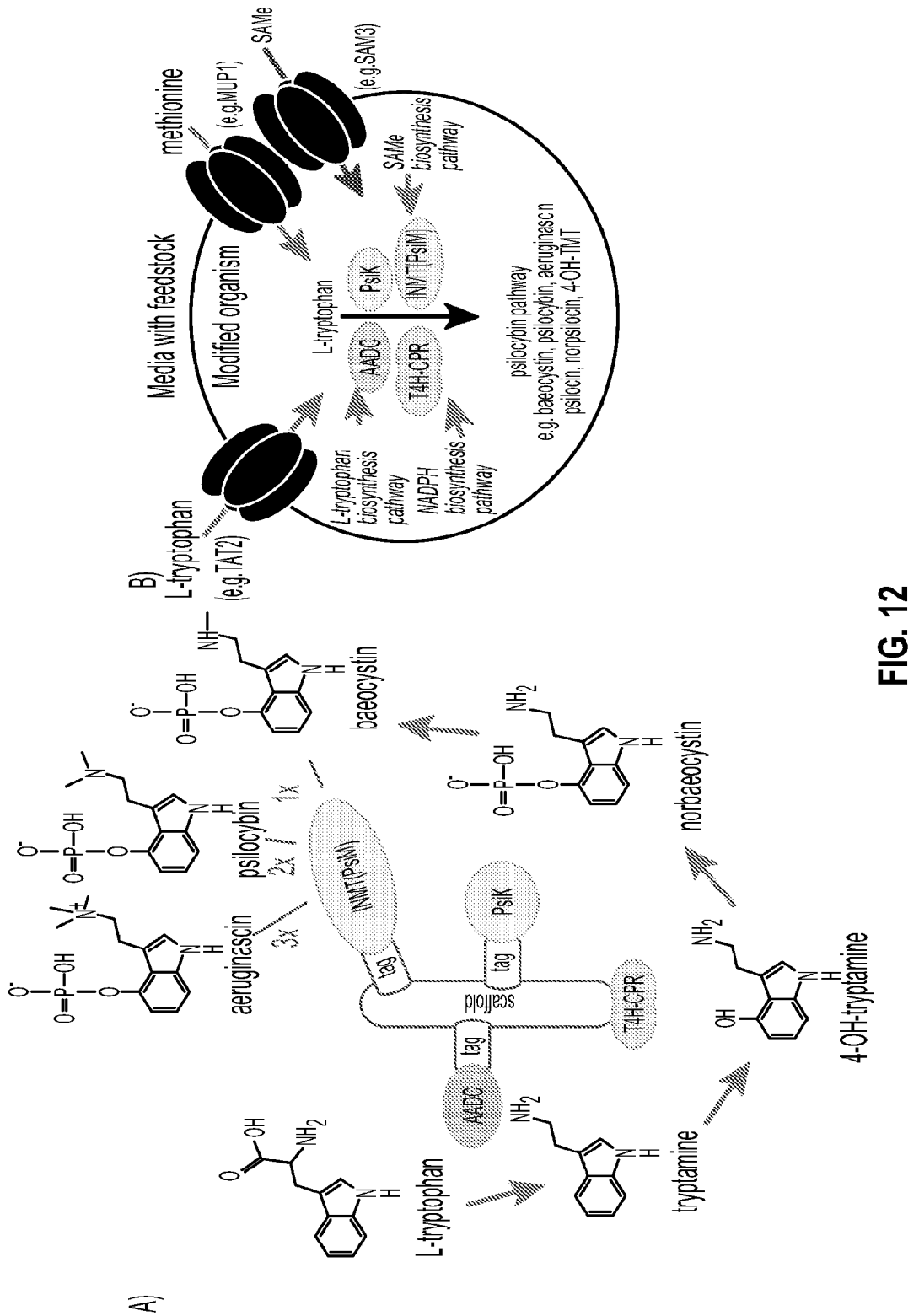

A)
L-tryptophan
AADC
tryptamine
4-OH-tryptamine
T4H-CPR
scaffold
tag
PsiK
INMTPsiM
norbaeocystin
baeocystin
psilocybin
2x
1x
aeruginascin
3x
B)
L-tryptophan
(e.g.TAT2)
Media with feedstock
Modified organism
methionine
(e.g.MUP1)
SAMe
(e.g.SAM3)
L-tryptophan biosynthesis pathway
NADPH biosynthesis pathway
SAMe biosynthesis pathway
psilocybin pathway
e.g. baeocystin, psilocybin, aeruginascin
psilocin, norpsilocin, 4-OH-TMT

FIG. 12

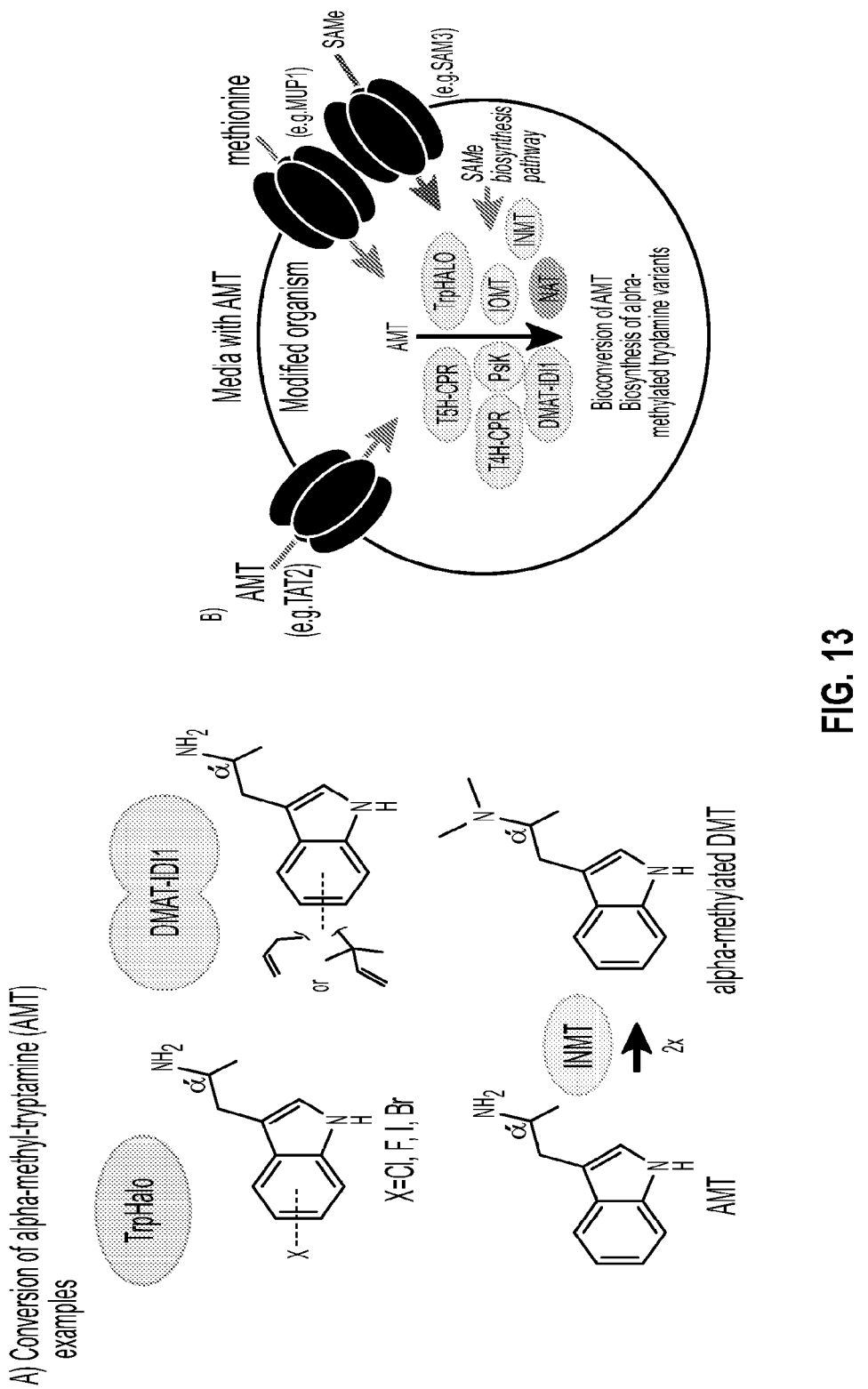
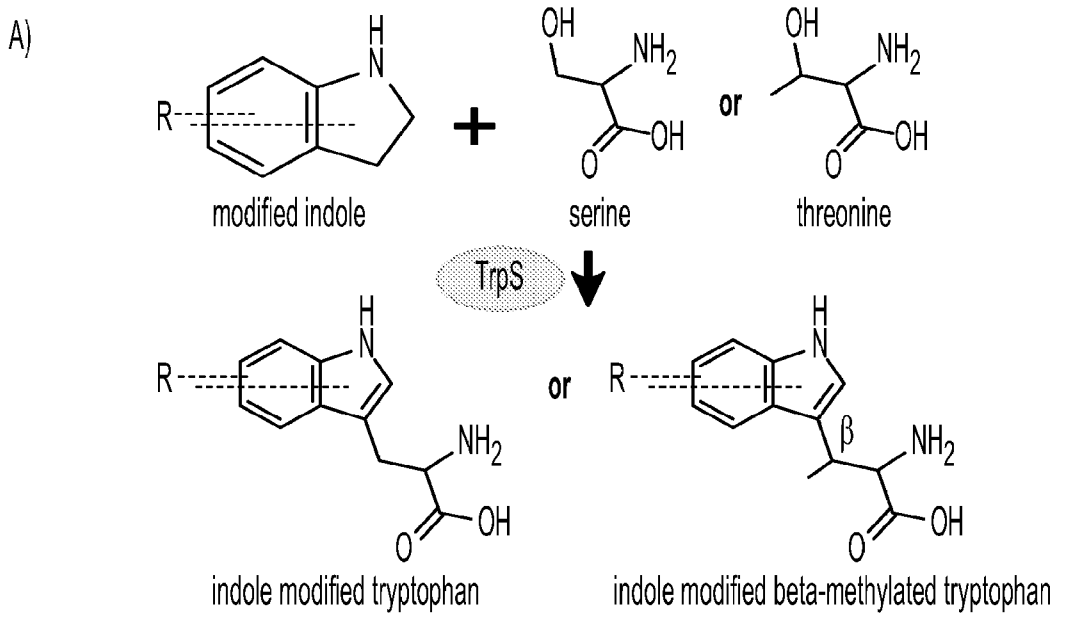

A) Conversion of alpha-methyl-tryptamine (AMT) examples
TrpHalo
DMAT-ID1
X=Cl, F, I, Br
or
INMT
2x
AMT
alpha-methylated DMT
B)
AMT
(e.g.TAT2)
Media with AMT
Modified organism
methionine
(e.g.MUP1)
SAMe
(e.g.SAM3)
SAMe biosyntheis pathway
AMT
TrpHALO
OMT
INMT
NMT
T5H-CPR
PsiK
T4H-CPR
DMAT-ID11
Bioconversion of AMT
Biosynthesis of alpha-methylated tryptamine variants

FIG. 13


A)
R
H
N
modified indole
+
OH
NH2
O
OH
serine
or
OH
NH2
O
OH
threonine
TrpS
R
H
N
NH2
O
OH
indole modified tryptophan
or
R
H
N
β
NH2
O
OH
indole modified beta-methylated tryptophan

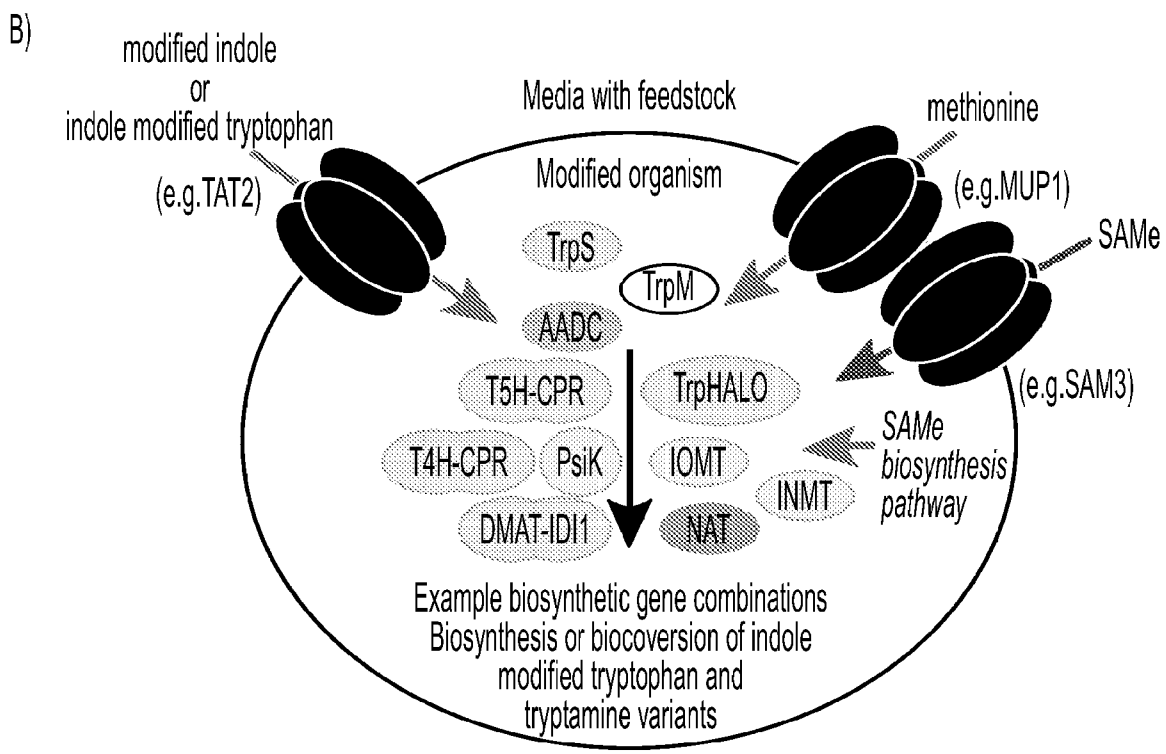
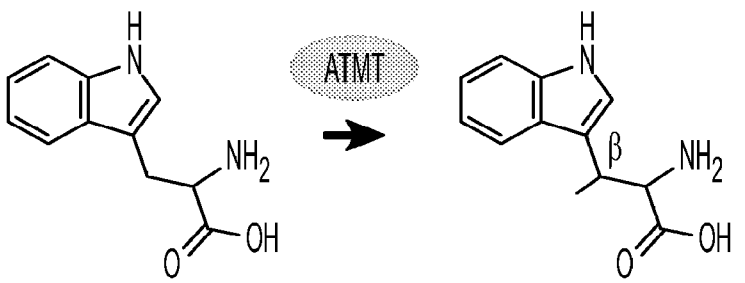

B)
modified indole
or
indole modified tryptophan
(e.g.TAT2)
Media with feedstock
Modified organism
methionine
(e.g.MUP1)
SAMe
(e.g.SAM3)
TrpS
TrpM
AADC
T5H-CPR
TrpHALO
T4H-CPR
PsiK
IOMT
INMT
DMAT-IDI1
NAT
SAMe biosynthesis pathway
Example biosynthetic gene combinations
Biosynthesis or biocoversion of indole modified tryptophan and tryptamine variants

FIG. 14

A)


H
N
NH2
OH
O
ATMT
β
tryptophan
beta-methylated tryptophan

B)

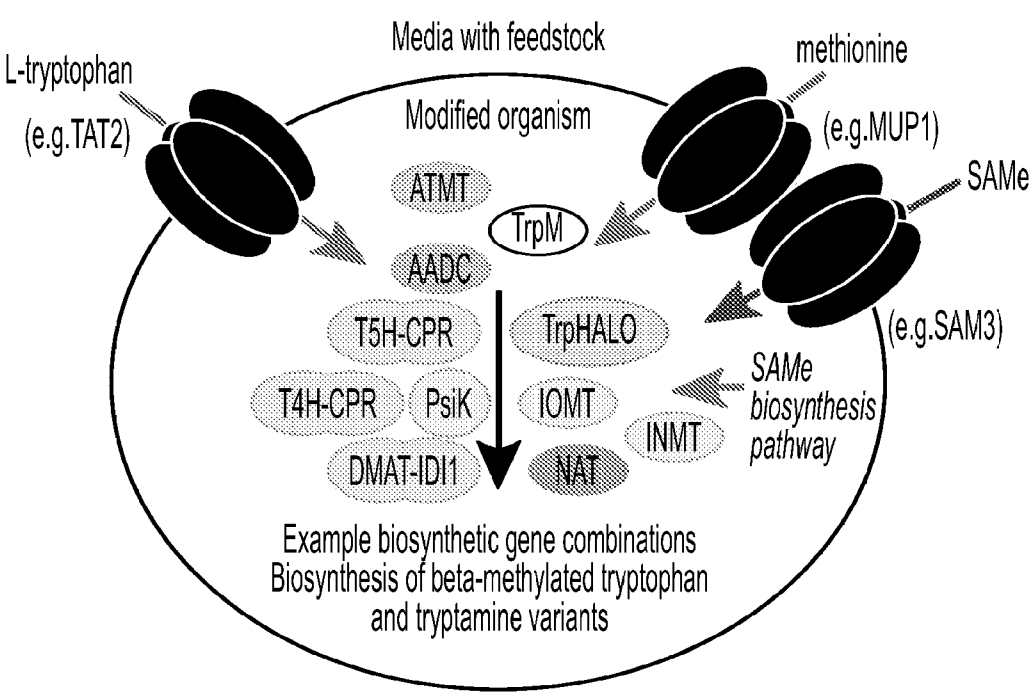

Media with feedstock
L-tryptophan
(e.g.TAT2)
Modified organism
methionine
(e.g.MUP1)
SAMe
(e.g.SAM3)
ATMT
TrpM
AADC
T5H-CPR
TrpHALO
T4H-CPR
PsiK
IOMT
INMT
DMAT-IDI1
NAT
SAMe biosynthesis pathway
Example biosynthetic gene combinations
Biosynthesis of beta-methylated tryptophan
and tryptamine variants

FIG. 15

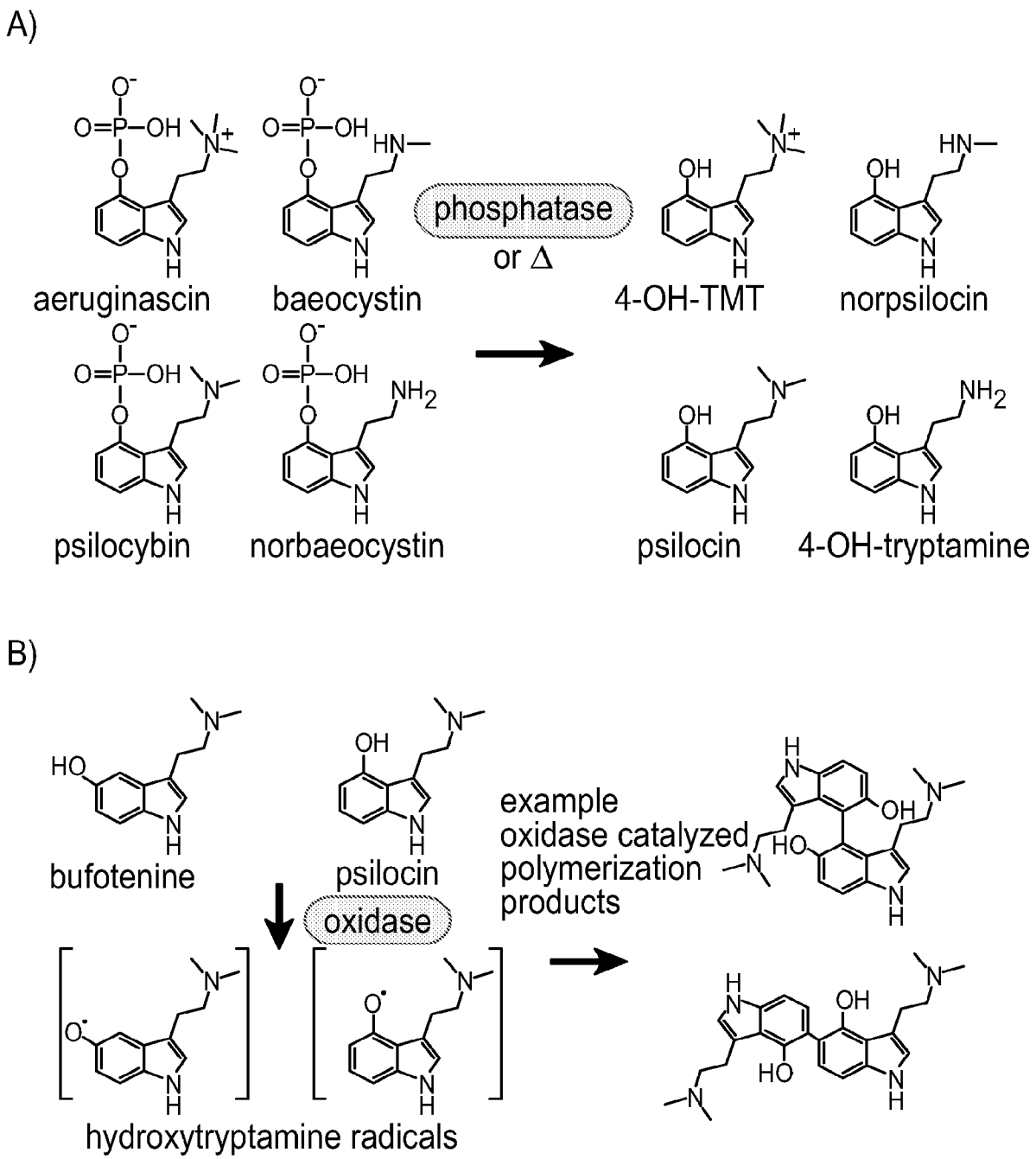

A)
aeruginascin
baeocystin
phosphatase
or Δ
4-OH-TMT
norpsilocin
psilocybin
norbaeocystin
psilocin
4-OH-tryptamine
B)
bufotenine
psilocin
oxidase
example
oxidase catalyzed
polymerization
products
hydroxytryptamine radicals

FIG. 16

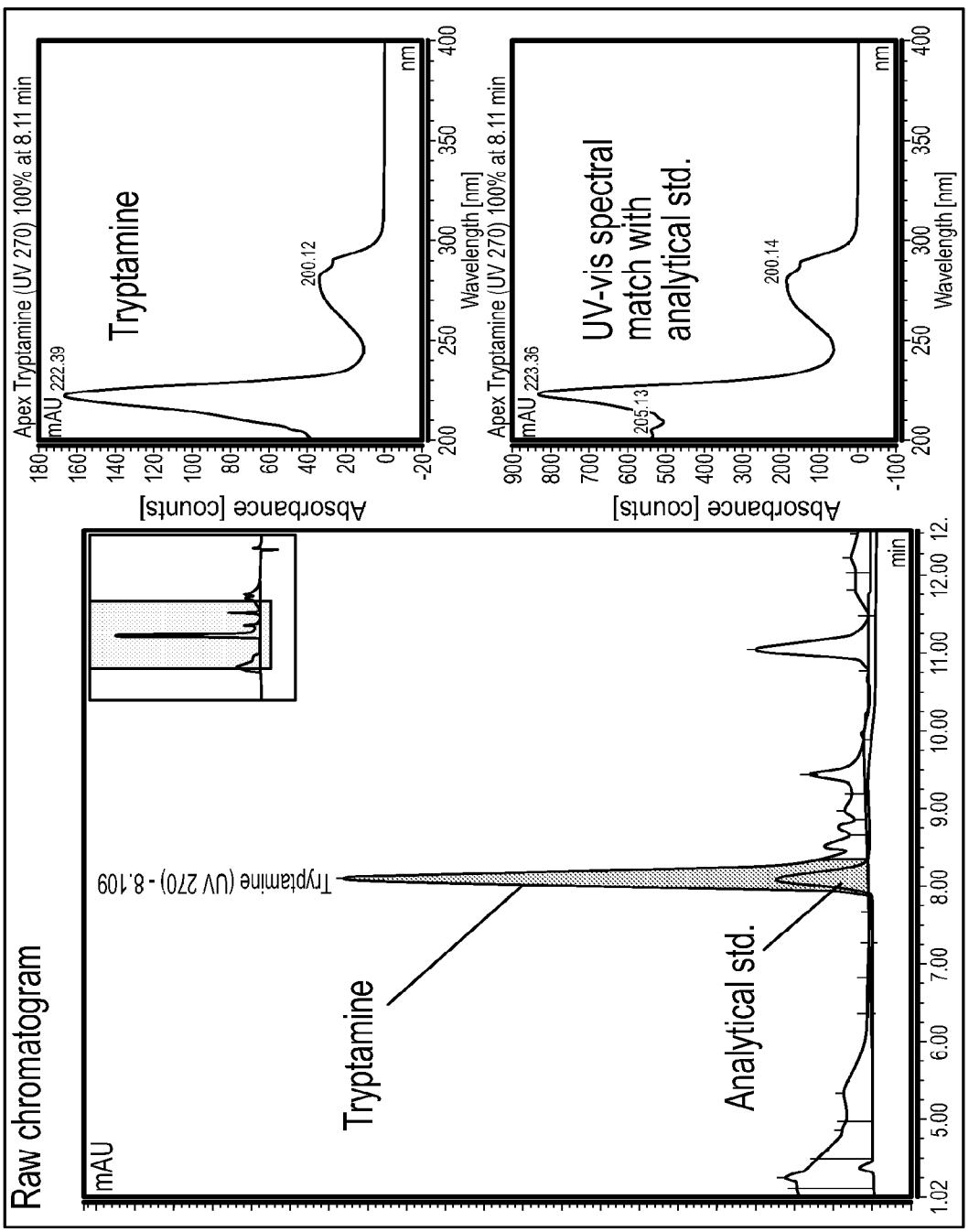

Raw chromatogram
mAU
Tryptamine (UV 270) - 8.109
Tryptamine
Analytical std.
min
1.02
5.00
6.00
7.00
8.00
9.00
10.00
11.00
12.00
12.
Absorbance [counts]
Apex Tryptamine (UV 270) 100% at 8.11 min
mAU
222.39
Tryptamine
200.12
nm
Wavelength [nm]
Absorbance [counts]
Apex Tryptamine (UV 270) 100% at 8.11 min
mAU
223.36
UV-vis spectral match with analytical std.
205.13
200.14
nm
Wavelength [nm]

FIG. 17

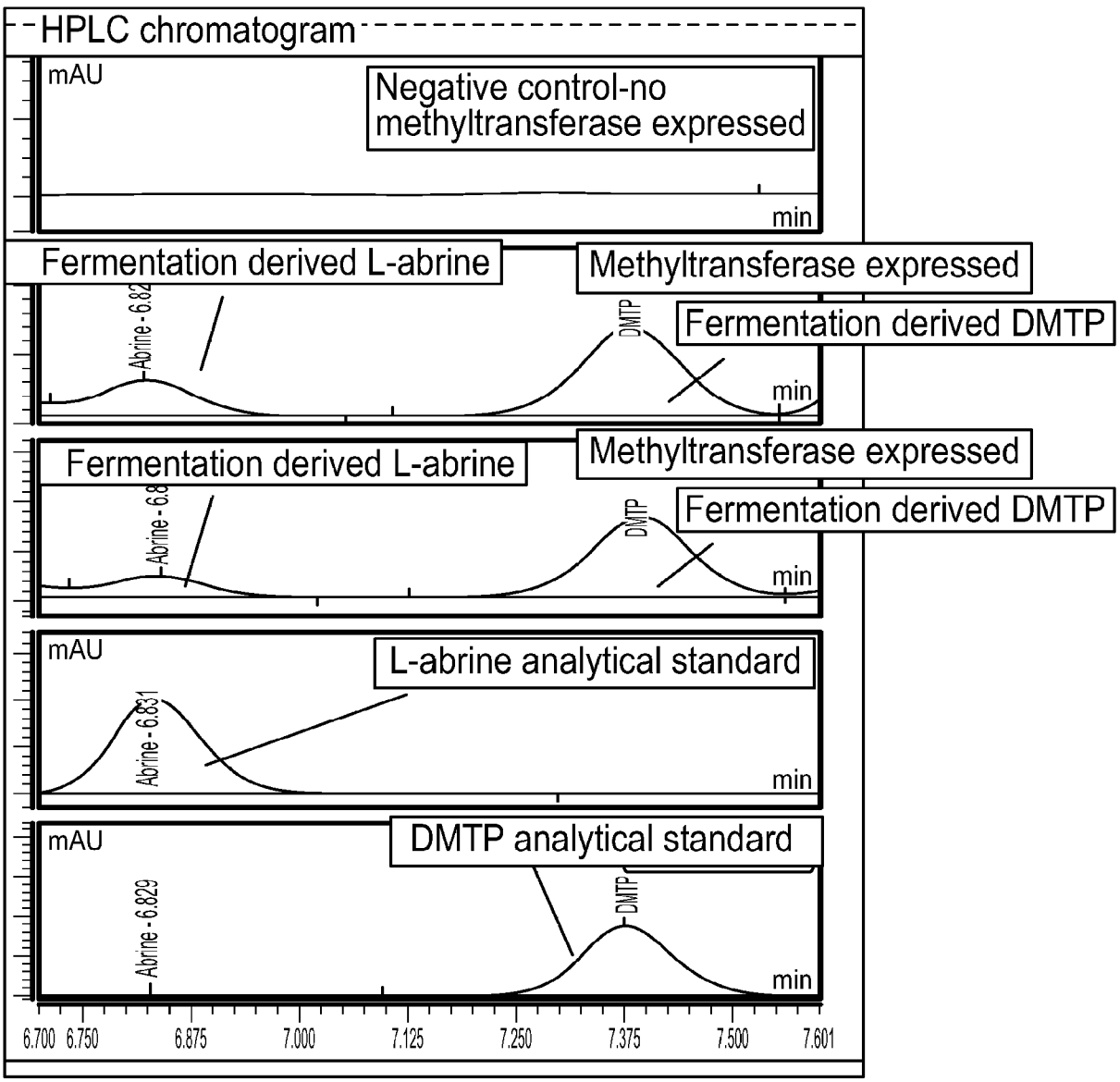

HPLC chromatogram
mAU
Negative control-no methyltransferase expressed
min
Fermentation derived L-abrine
Methyltransferase expressed
Fermentation derived DMTP
Abrine - 6.82
DMTP
min
Fermentation derived L-abrine
Methyltransferase expressed
Fermentation derived DMTP
Abrine - 6.8
DMTP
min
mAU
L-abrine analytical standard
Abrine - 6.831
min
mAU
DMTP analytical standard
Abrine - 6.829
DMTP
min
6.700 6.750 6.875 7.000 7.125 7.250 7.375 7.500 7.601

FIG. 18

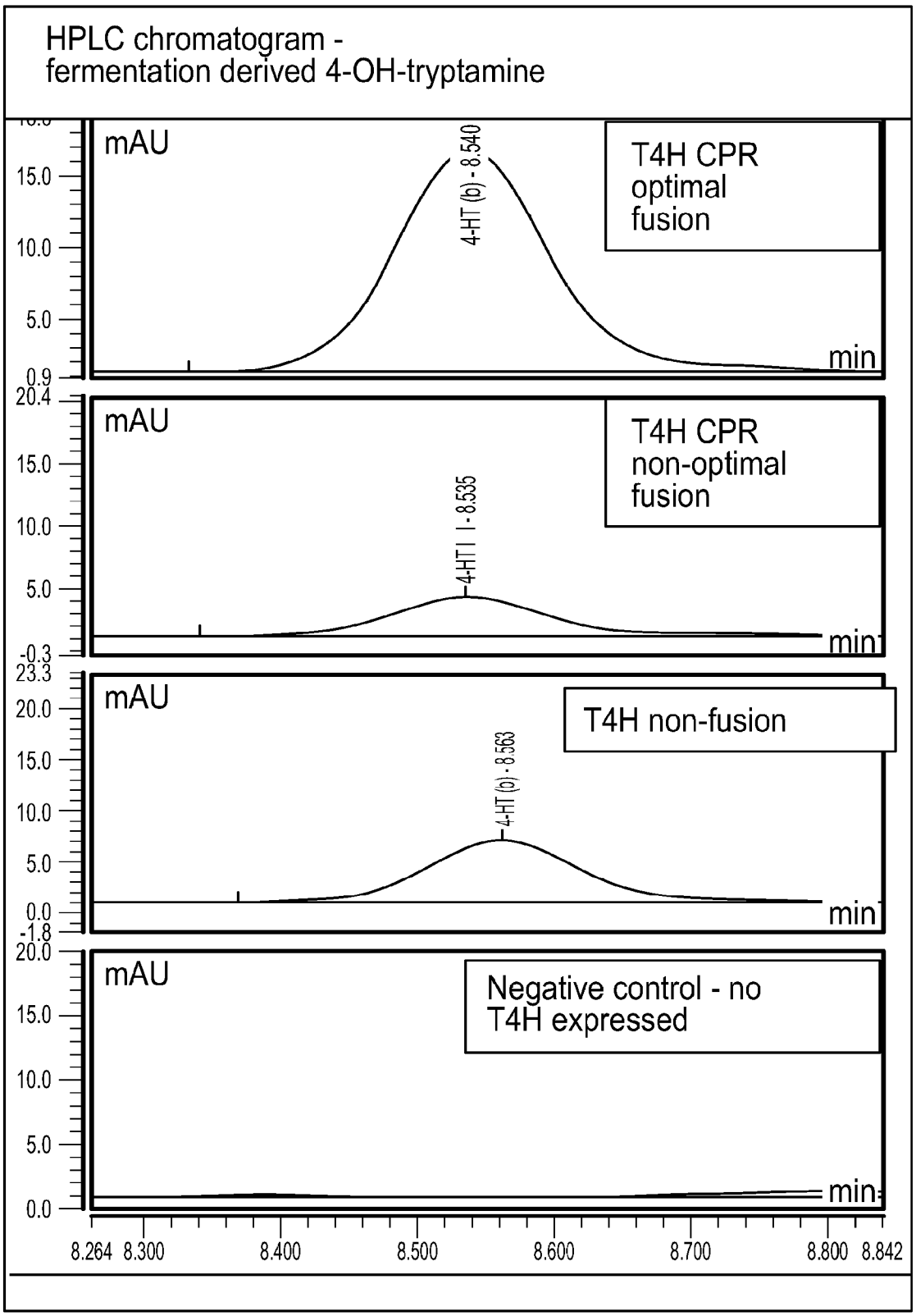

HPLC chromatogram -
fermentation derived 4-OH-tryptamine
mAU
4-HT (b) - 8.540
T4H CPR
optimal
fusion
min
mAU
4-HT I - 8.535
T4H CPR
non-optimal
fusion
min
mAU
4-HT (b) - 8.563
T4H non-fusion
min
mAU
Negative control - no
T4H expressed
min
8.264 8.300 8.400 8.500 8.600 8.700 8.800 8.842

FIG. 19

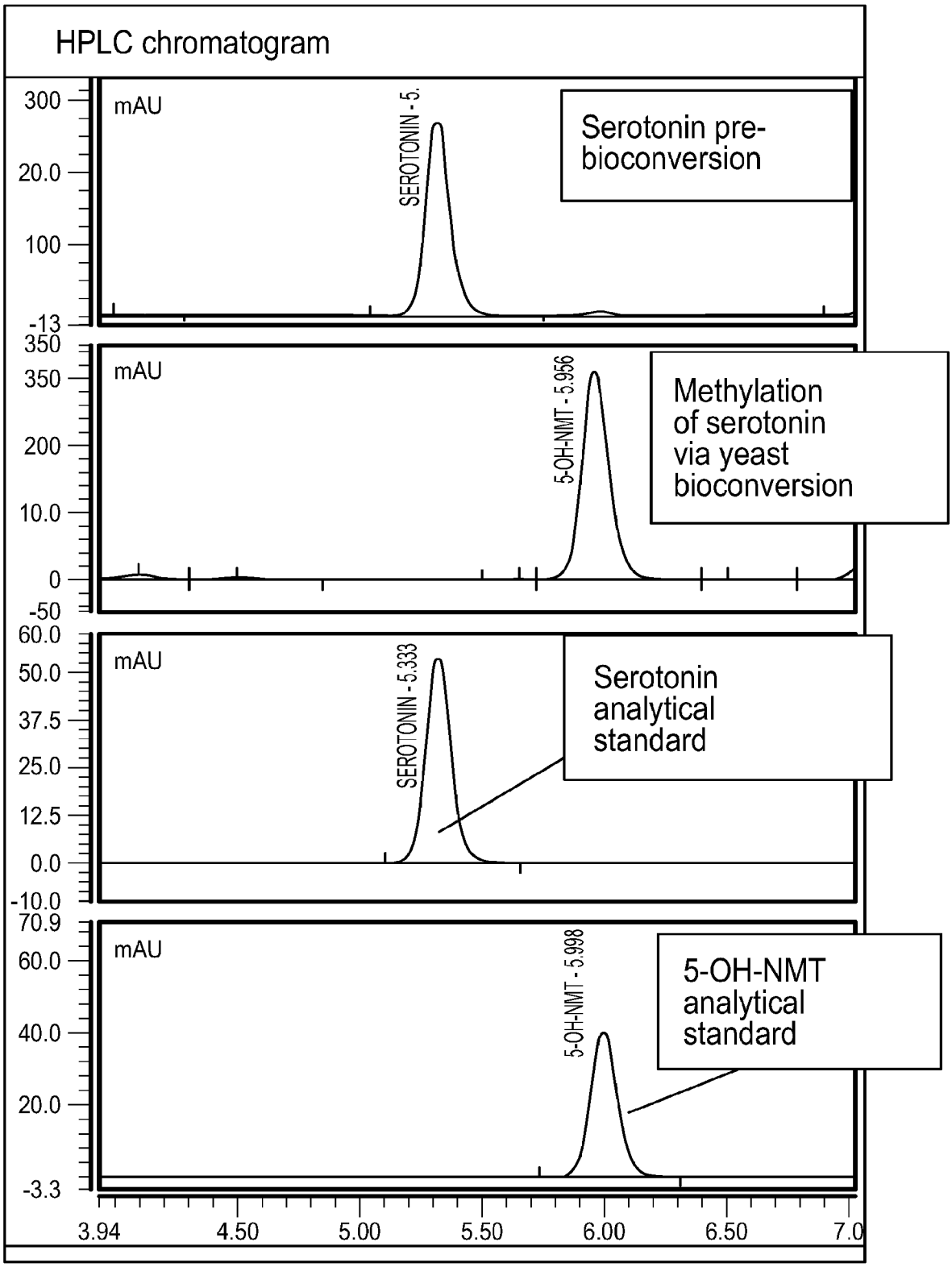

HPLC chromatogram
mAU
SEROTONIN - 5.
Serotonin pre-bioconversion
Methylation of serotonin via yeast bioconversion
5-OH-NMT - 5.956
Serotonin analytical standard
SEROTONIN - 5.333
5-OH-NMT analytical standard
5-OH-NMT - 5.998
3.94
4.50
5.00
5.50
6.00
6.50
7.0

FIG. 20

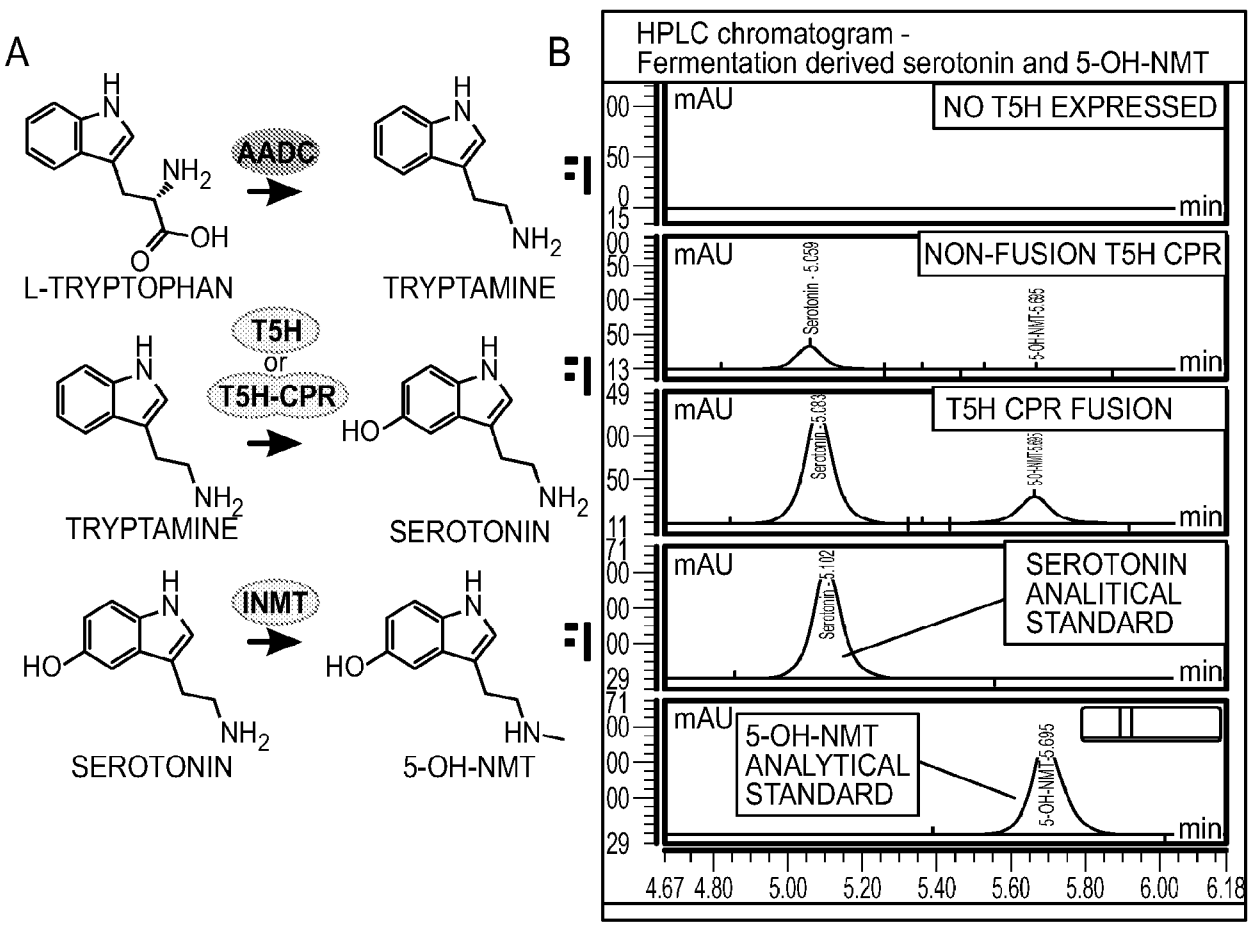

A
B
L-TRYPTOPHAN
AADC
TRYPTAMINE
T5H
or
T5H-CPR
TRYPTAMINE
SEROTONIN
INMT
SEROTONIN
5-OH-NMT
HPLC chromatogram -
Fermentation derived serotonin and 5-OH-NMT
NO T5H EXPRESSED
NON-FUSION T5H CPR
T5H CPR FUSION
SEROTONIN ANALITICAL STANDARD
5-OH-NMT ANALYTICAL STANDARD
mAU
min
4.67 4.80 5.00 5.20 5.40 5.60 5.80 6.00 6.18

FIG. 21

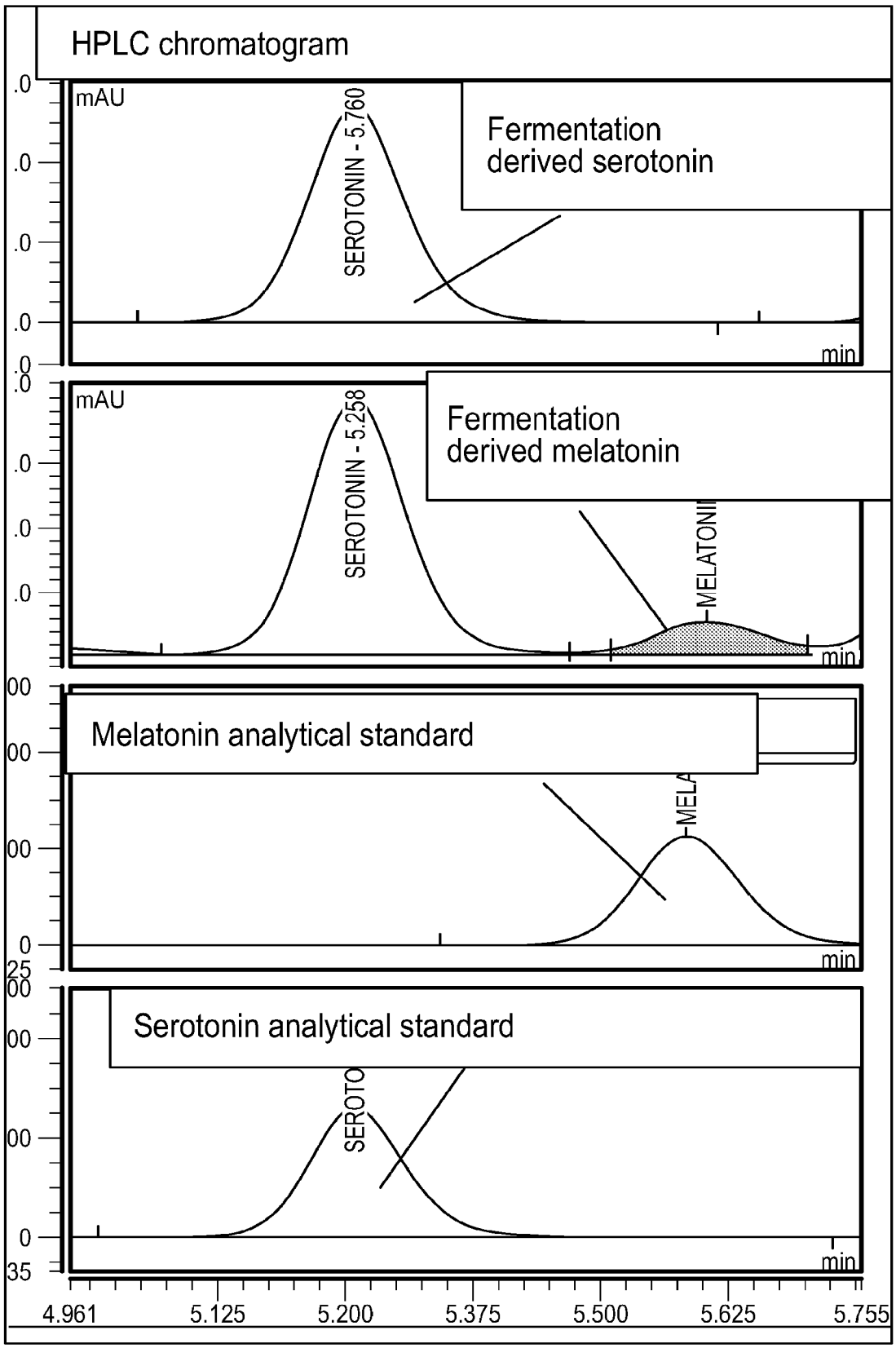

HPLC chromatogram
mAU
SEROTONIN - 5.760
Fermentation derived serotonin
min
mAU
SEROTONIN - 5.258
Fermentation derived melatonin
min
Melatonin analytical standard
min
Serotonin analytical standard
min
4.961
5.125
5.200
5.375
5.500
5.625
5.755

FIG. 22

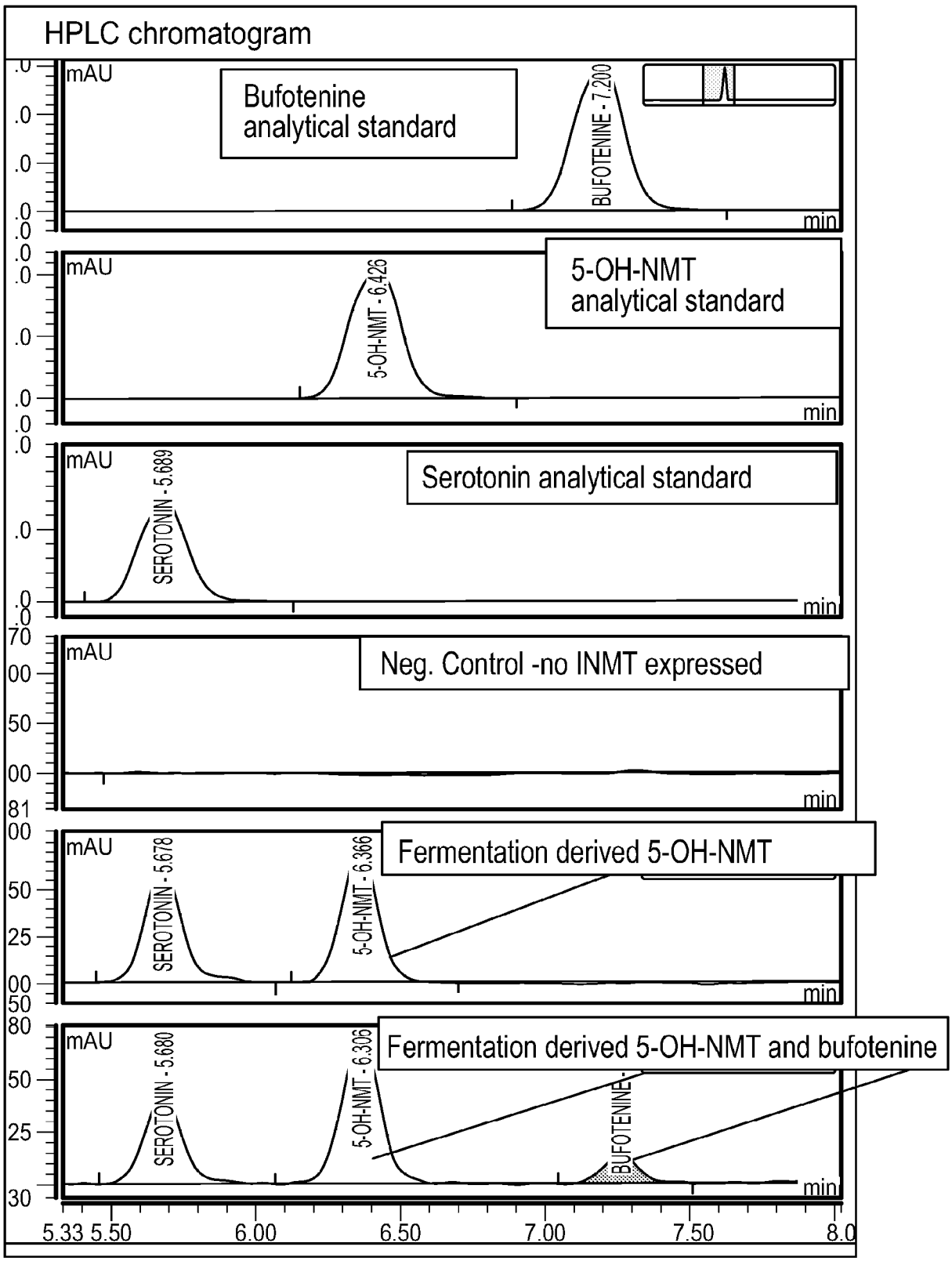

HPLC chromatogram
mAU
Bufotenine analytical standard
BUFOTENINE - 7.200
min
mAU
5-OH-NMT - 6.426
5-OH-NMT analytical standard
min
mAU
SEROTONIN - 5.689
Serotonin analytical standard
min
mAU
Neg. Control -no INMT expressed
min
mAU
SEROTONIN - 5.678
5-OH-NMT - 6.366
Fermentation derived 5-OH-NMT
min
mAU
SEROTONIN - 5.680
5-OH-NMT - 6.306
BUFOTENINE -
Fermentation derived 5-OH-NMT and bufotenine
min
5.33 5.50 6.00 6.50 7.00 7.50 8.0

FIG. 23

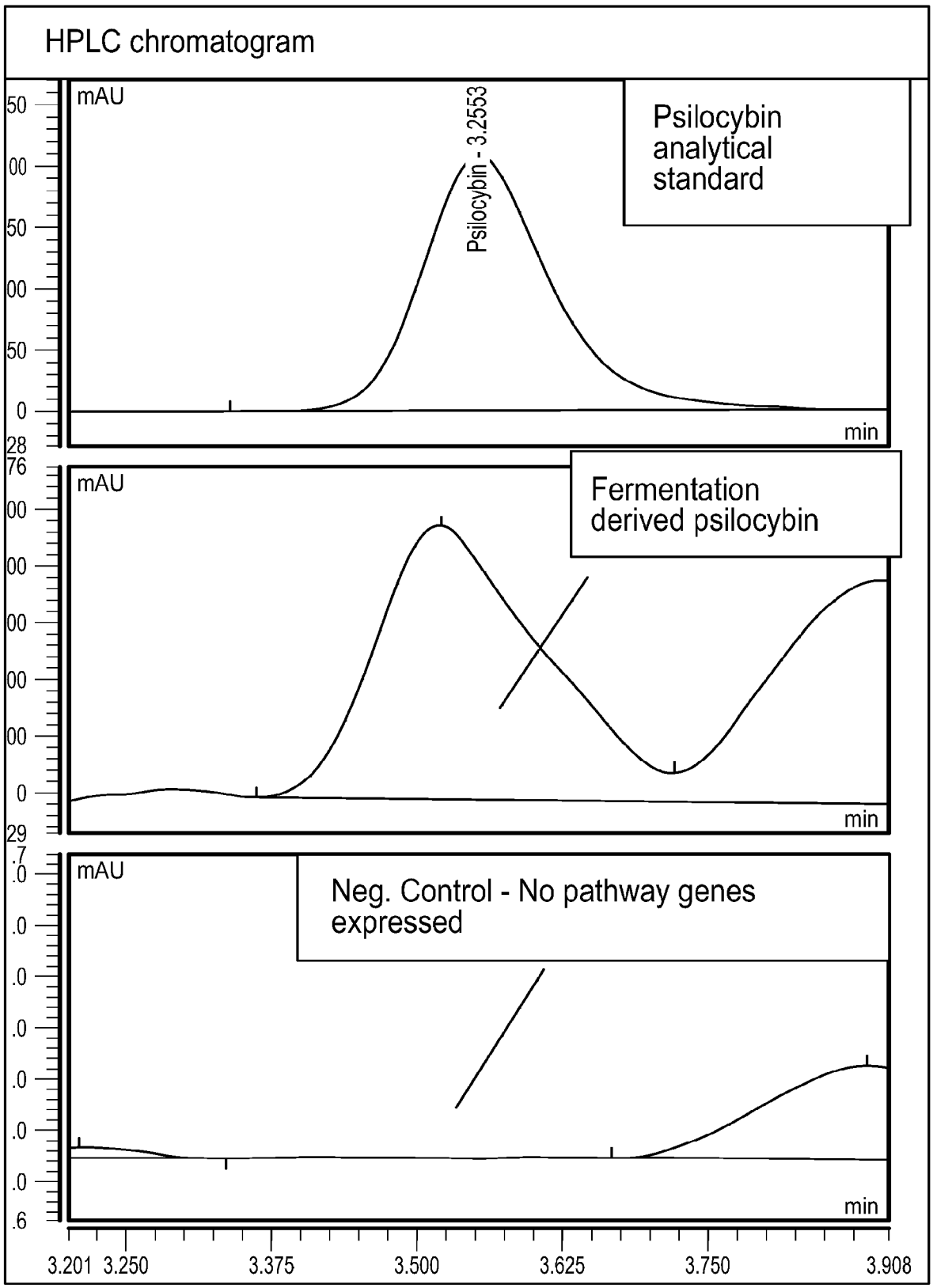

HPLC chromatogram
mAU
Psilocybin - 3.2553
Psilocybin analytical standard
min
mAU
Fermentation derived psilocybin
min
mAU
Neg. Control - No pathway genes expressed
min
3.201
3.250
3.375
3.500
3.625
3.750
3.908

FIG. 24

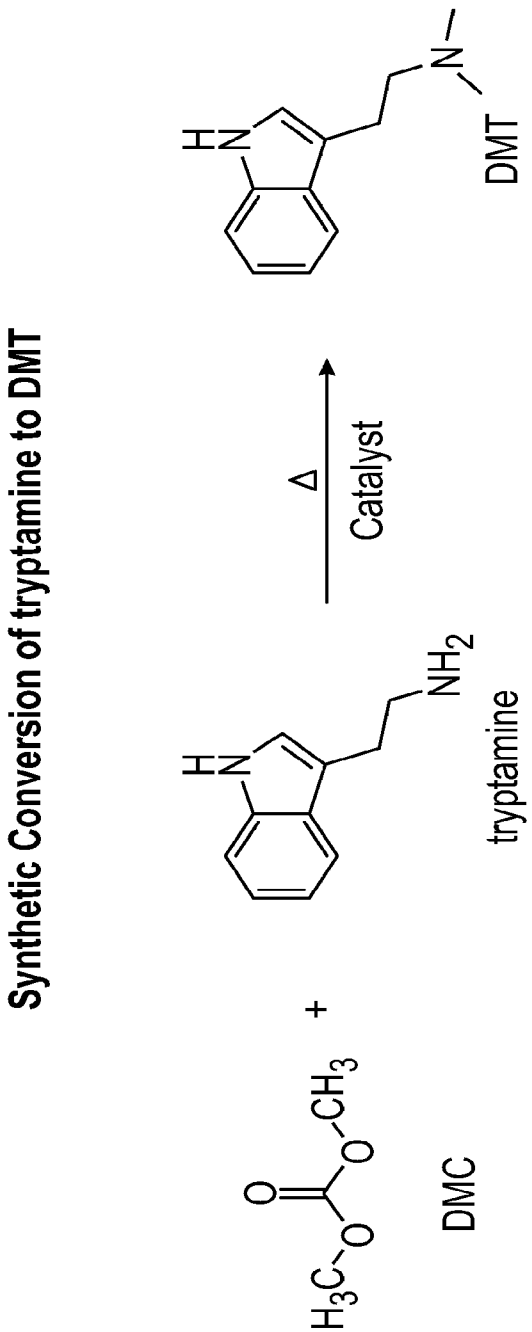

Synthetic Conversion of tryptamine to DMT
DMC
tryptamine
Catalyst
Δ
DMT

FIG. 25

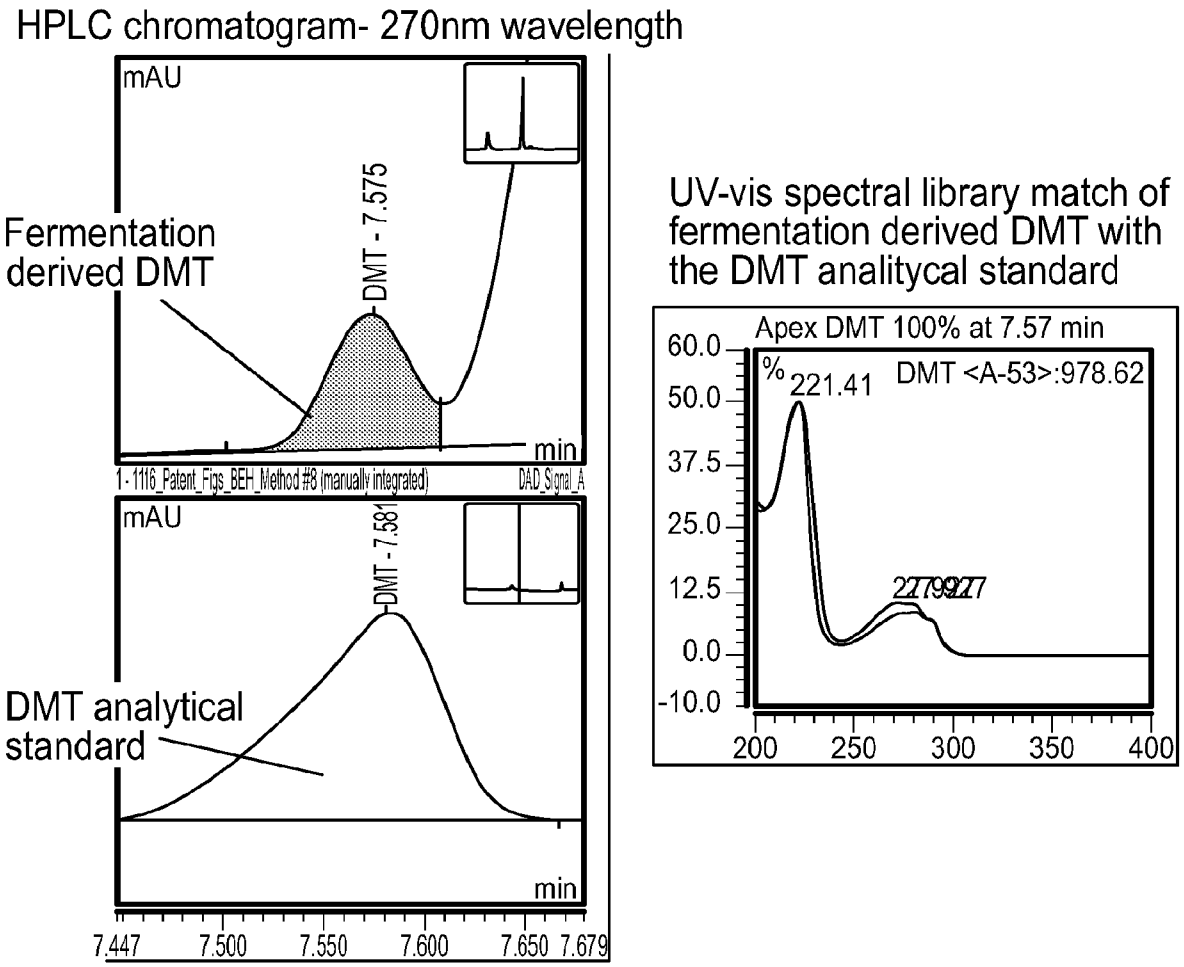

HPLC chromatogram- 270nm wavelength
mAU
Fermentation derived DMT
DMT - 7.575
min
1-1116_Patent_Figs_BEH_Method #8 (manually integrated)
DAD_Signal_A
mAU
DMT - 7.581
DMT analytical standard
min
7.447
7.500
7.550
7.600
7.650
7.679
UV-vis spectral library match of fermentation derived DMT with the DMT analitycal standard
Apex DMT 100% at 7.57 min
%
221.41
DMT <A-53>:978.62
60.0
50.0
37.5
25.0
12.5
0.0
-10.0
200
250
300
350
400

FIG. 26

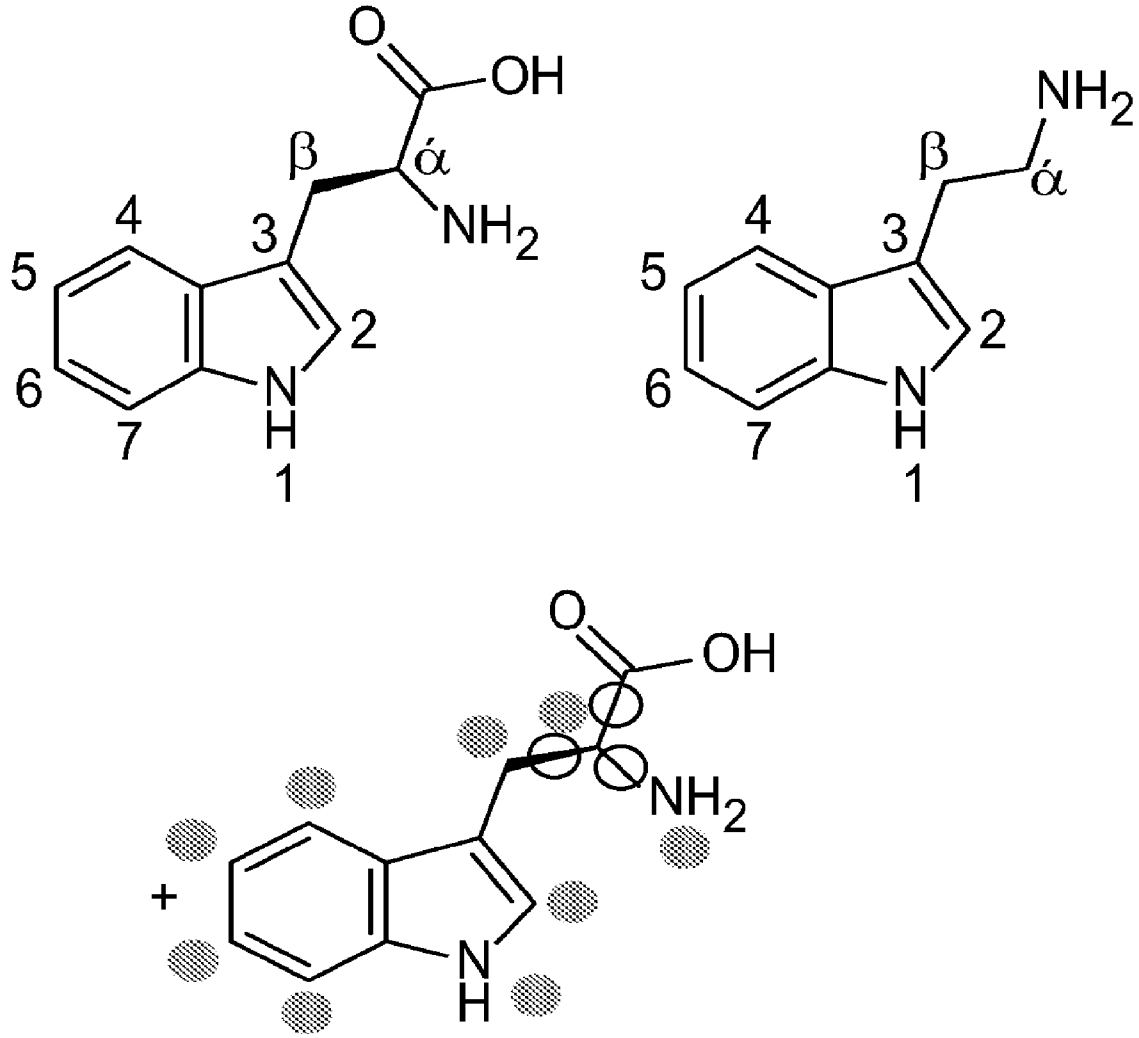

O
OH
β
α
NH2
4
3
5
2
6
N
7
H
1
NH2
β
α
4
3
5
2
6
N
7
H
1
O
OH
NH2
+
N
H
methylation
halogenation
prenylation
hydroxylation
O-phosphorylation
O-methylation
N-Acetylation
cleavage
deamination
decarboxylation
+ oligomerization

FIG. 1