



## (51) International Patent Classification:

**C07D 209/16** (2006.01)      **C12N 15/53** (2006.01)  
**A61K 31/4045** (2006.01)      **C12N 15/54** (2006.01)  
**A61K 31/675** (2006.01)      **C12N 15/60** (2006.01)  
**C07F 9/572** (2006.01)      **C12N 9/02** (2006.01)  
**C12P 17/10** (2006.01)      **C12N 9/10** (2006.01)  
**C12N 9/00** (2006.01)      **C12N 9/12** (2006.01)  
**C12N 15/52** (2006.01)      **C12N 9/88** (2006.01)

## (21) International Application Number:

PCT/CA2021/051209

## (22) International Filing Date:

01 September 2021 (01.09.2021)

## (25) Filing Language:

English

## (26) Publication Language:

English

## (30) Priority Data:

63/073,104      01 September 2020 (01.09.2020)      US

## (71) Applicant: MAGICMED INDUSTRIES INC. [CA/CA];

3655-36th Street NW, Calgary, Alberta T2L 1Y8 (CA).

## (72) Inventors: HAGEL, Jillian M.; 2210-6651 Ranchview

Drive NW, Calgary, Alberta T3G 1P3 (CA). FACCHINI,

Peter J.; 155 Ranchridge Drive NW, Calgary, Alberta T3G 1W1 (CA).

## (74) Agent: BERESKIN &amp; PARR LLP/S.E.N.C.R.L., S.R.L.;

Scotia Plaza, 40 King Street West, 40th Floor, Toronto, Ontario M5H 3Y2 (CA).

## (81) Designated States (unless otherwise indicated, for every

kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

## (84) Designated States (unless otherwise indicated, for every

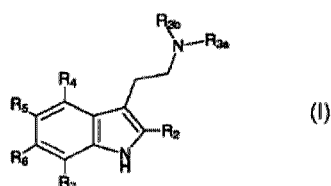
kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,

TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

## Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))
- in black and white; the international application as filed contained color or greyscale and is available for download from PATENTSCOPE

## (54) Title: HALOGENATED PSILOCYBIN DERIVATIVES AND METHODS OF USING



(57) **Abstract:** Disclosed are halogenated psilocybin derivative compounds of formula (I), and pharmaceutical and recreational drug formulations containing the same. The compounds may be produced using a biosynthetic system which comprises cells comprising a psilocybin biosynthetic enzyme complement.

**TITLE: HALOGENATED PSILOCYBIN DERIVATIVES AND METHODS OF  
USING**

**RELATED APPLICATION**

[001] This application claims the benefit of United States Provisional Application No. 63/073,104 filed September 1, 2020; the entire contents of Patent  
5 Application No. 63/073,104 are hereby incorporated by reference.

**FIELD OF THE DISCLOSURE**

[002] The compositions and methods disclosed herein relate to a chemical compound known as psilocybin. Furthermore, the compositions and methods  
10 disclosed herein relate in particular to halogenated forms of psilocybin.

**BACKGROUND OF THE DISCLOSURE**

[003] The following paragraphs are provided by way of background to the present disclosure. They are not however an admission that anything discussed  
15 therein is prior art or part of the knowledge of a person of skill in the art.

[004] The biochemical pathways in the cells of living organisms may be classified as being part of primary metabolism, or as being part of secondary metabolism. Pathways that are part of a cell's primary metabolism are involved in catabolism for energy production or in anabolism for building block production for  
20 the cell. Secondary metabolites, on the other hand, are produced by the cell without having an obvious anabolic or catabolic function. It has long been recognized that secondary metabolites can be useful in many respects, including as therapeutic compounds.

[005] Psilocybin, for example, is a secondary metabolite that is naturally  
25 produced by certain mushrooms which taxonomically can be classified as belonging the Basidiomycota division of the fungi kingdom. Mushroom species which can produce psilocybin include species belonging to the genus *Psilocybe*, such as *Psilocybe azurescens*, *Psilocybe semilanceata*, *Psilocybe serbica*, *Psilocybe mexicana*, and *Psilocybe cyanescens*, for example. The interest of the  
30 art in psilocybin is well established. Thus, for example, psilocybin is a psychoactive compound and is therefore used as a recreational drug. Furthermore, psilocybin is used as a research tool in behavioral and neuro-imaging studies in psychotic disorders, and has been evaluated for its clinical potential in the treatment of

mental health conditions (Daniel, J. *et al.* Mental Health Clin/, 2017;7(1): 24-28), including to treat anxiety in terminal cancer patients (Grob, C. *et al.* Arch. Gen. Psychiatry, 2011, 68(1) 71-78) and to alleviate symptoms of treatment-resistant depression (Cathart-Harris, R.L. *et al.* Lancet Psychiatry, 2016, 3: 619-627).

5 **[006]** Although the toxicity of psilocybin is low, adverse side effects, including, for example, panic attacks, paranoia and psychotic states, sometimes together or individually referred to as “a bad trip”, are not infrequently experienced by recreational psilocybin users.

**[007]** There exists therefore a need in the art for improved psilocybin  
10 compounds.

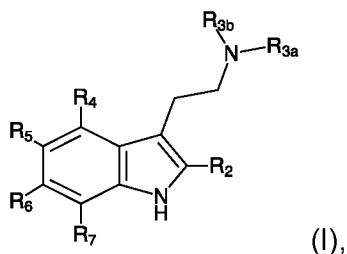
### **SUMMARY OF THE DISCLOSURE**

**[008]** The following paragraphs are intended to introduce the reader to the more detailed description, not to define or limit the claimed subject matter of the  
15 present disclosure.

**[009]** In one aspect, the present disclosure relates to psilocybin and derivative compounds.

**[0010]** In another aspect, the present disclosure relates to halogenated psilocybin derivative compounds and methods of making and using these  
20 compounds.

**[0011]** Accordingly, in one aspect, the present disclosure provides, in accordance with the teachings herein, in at least one embodiment, a chemical compound or salts thereof having the formula (I):



25 wherein, at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> is a halogen atom, wherein each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group, wherein R<sub>4</sub> when it is not halogenated is a phosphate group, hydrogen atom or an alkyl or O-alkyl group, and wherein R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group.

- [0012]** In at least one embodiment, in an aspect, the halogen atom can be a bromine atom.
- [0013]** In at least one embodiment, in an aspect, the halogen atom can be a chlorine atom.
- 5 **[0014]** In at least one embodiment, in an aspect, the halogen atom can be a fluorine atom.
- [0015]** In at least one embodiment, in an aspect, the halogen atom can be an iodine atom.
- 10 **[0016]** In at least one embodiment, in an aspect, R<sub>2</sub> can be a halogen atom and R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can each be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.
- [0017]** In at least one embodiment, in an aspect, R<sub>4</sub> can be a halogen atom and R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can each be a hydrogen atom or an alkyl group or O-alkyl group.
- 15 **[0018]** In at least one embodiment, in an aspect, R<sub>5</sub> can be a halogen atom and R<sub>2</sub>, R<sub>6</sub> and R<sub>7</sub> can each be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.
- 20 **[0019]** In at least one embodiment, in an aspect, R<sub>6</sub> can be a halogen atom and R<sub>2</sub>, R<sub>5</sub> and R<sub>7</sub> can each be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.
- [0020]** In at least one embodiment, in an aspect, R<sub>7</sub> can be a halogen atom and R<sub>2</sub>, R<sub>5</sub> and R<sub>6</sub> can each be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.
- 25 **[0021]** In at least one embodiment, in an aspect, at least two of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> can be a halogen atom.
- 30 **[0022]** In at least one embodiment, in an aspect, the at least two of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> that are halogen atoms can be at least two identical halogen atoms.



**[0023]** In at least one embodiment, in an aspect, the at least of two of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> that are halogen atoms can at least two non-identical halogen atoms.

5 **[0024]** In at least one embodiment, in an aspect, R<sub>2</sub> and R<sub>4</sub> can be a halogen atom, and R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group.

10 **[0025]** In at least one embodiment, in an aspect, R<sub>2</sub> and R<sub>5</sub> can be a halogen atom, R<sub>6</sub> and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a hydrogen atom, a phosphate group or an alkyl group or O-alkyl group.

**[0026]** In at least one embodiment, in an aspect, R<sub>2</sub> and R<sub>6</sub> can be a halogen atom, R<sub>5</sub> and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a hydrogen atom, a phosphate group or an alkyl group or O-alkyl group.

15 **[0027]** In at least one embodiment, in an aspect, R<sub>2</sub> and R<sub>7</sub> can be a halogen atom, R<sub>5</sub> and R<sub>6</sub> can be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a hydrogen atom, a phosphate group or an alkyl group or O-alkyl group.

20 **[0028]** In at least one embodiment, in an aspect, R<sub>4</sub> and R<sub>5</sub> can be a halogen atom, and R<sub>2</sub>, R<sub>6</sub> and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group.

**[0029]** In at least one embodiment, in an aspect, R<sub>4</sub> and R<sub>6</sub> can be a halogen atom, and R<sub>2</sub>, R<sub>5</sub> and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group.

25 **[0030]** In at least one embodiment, in an aspect, R<sub>4</sub> and R<sub>7</sub> can be a halogen atom, and R<sub>2</sub>, R<sub>5</sub> and R<sub>6</sub> can be a hydrogen atom or an alkyl group or O-alkyl group.

30 **[0031]** In at least one embodiment, in an aspect, R<sub>5</sub> and R<sub>6</sub> can be a halogen atom, R<sub>2</sub> and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a hydrogen atom, a phosphate group or an alkyl group or O-alkyl group.

**[0032]** In at least one embodiment, in an aspect, R<sub>5</sub> and R<sub>7</sub> can be a halogen atom, R<sub>2</sub> and R<sub>6</sub> can be a hydrogen atom or an alkyl group or O-alkyl

group, and R<sub>4</sub> can be a hydrogen atom, a phosphate group or an alkyl group or O-alkyl group.

**[0033]** In at least one embodiment, in an aspect, R<sub>6</sub> and R<sub>7</sub> can be a halogen atom, R<sub>2</sub> and R<sub>5</sub> can be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a hydrogen atom, a phosphate group or an alkyl group or O-alkyl group.

**[0034]** In at least one embodiment, in an aspect, R<sub>4</sub> when it is not halogenated can be a hydrogen atom.

**[0035]** In at least one embodiment, in an aspect, R<sub>4</sub> when it is not halogenated can be an alkyl group.

**[0036]** In at least one embodiment, in an aspect, R<sub>4</sub> when it is not halogenated can be an O-alkyl group.

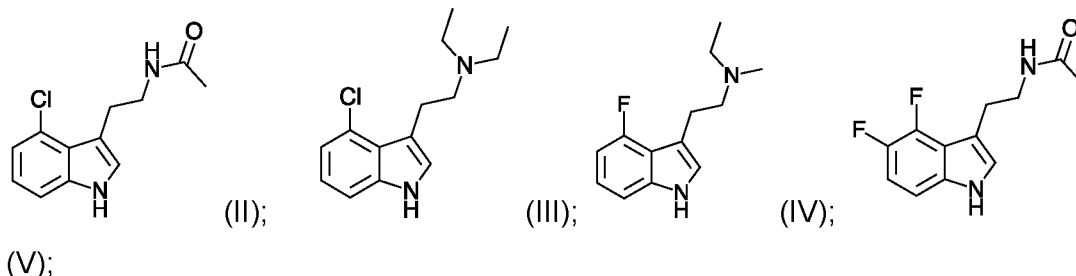
**[0037]** In at least one embodiment, in an aspect, R<sub>4</sub> when it is not halogenated can be a phosphate group.

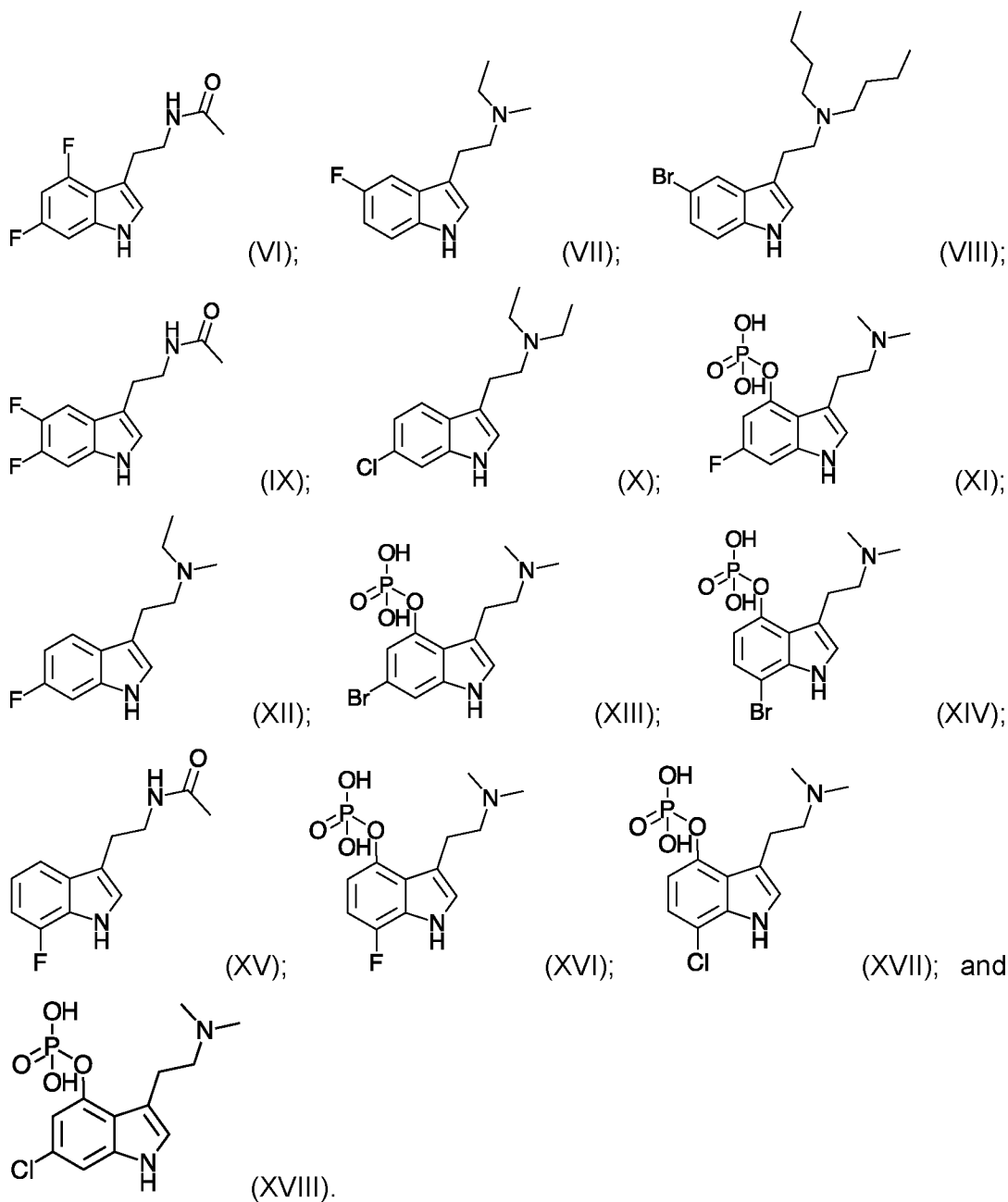
**[0038]** In at least one embodiment, in an aspect, three, four or all five of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> can be a halogen atom.

**[0039]** In at least one embodiment, in an aspect, the three, four, or five of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> that are halogen atoms can all be identical halogen atoms.

**[0040]** In at least one embodiment, in an aspect, the three, four or five of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> that are halogen atoms, can include at least two non-identical halogen atoms.

**[0041]** In at least one embodiment, in an aspect, the chemical compound can be selected from the group consisting of compounds having formulas (II); (III); (IV); (V); (VI); (VII); (VIII); (IX); (X); (XI); (XII); (XIII); (XIV); (XV); (XVI); and (XVII):

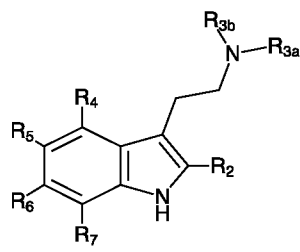




**[0042]** In at least one embodiment, the chemical compound is at least about 95% (w/w) pure.

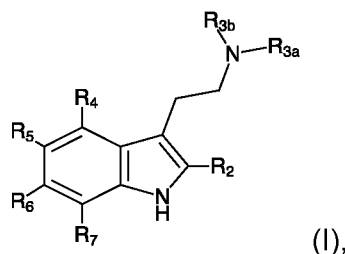
**[0043]** In another aspect, the present disclosure relates to pharmaceutical and recreational drug formulations comprising halogenated psilocybin derivatives. Accordingly, in one aspect, the present disclosure provides in at least one aspect, a pharmaceutical or recreational drug formulation comprising an effective amount of a chemical compound or salts thereof having the formula (I):

10



wherein, at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> is a halogen atom, wherein each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group, wherein R<sub>4</sub> when it is not halogenated is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group, and wherein R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group, together with a pharmaceutically acceptable excipient, diluent or carrier.

**[0044]** In another aspect, the present disclosure relates to methods of treatment of psychiatric disorders. Accordingly, the present disclosure further provide a method for treating a psychiatric disorder, the method comprising administering to a subject in need thereof a pharmaceutical formulation comprising a chemical compound or salts thereof having the formula (I):



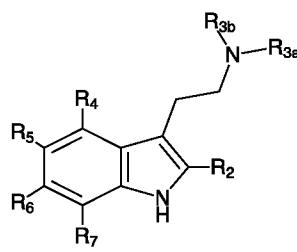
wherein, at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> is a halogen atom, wherein each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group, wherein R<sub>4</sub> when it is not halogenated is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group, and wherein R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group, wherein the pharmaceutical formulation is administered in an effective amount to treat the psychiatric disorder in the subject.

**[0045]** In at least one embodiment, in an aspect the disorder can be a 5-HT<sub>2A</sub> receptor mediated disorder.

**[0046]** In at least one embodiment, in an aspect, a dose can be administered of about 0.001 mg to about 5,000 mg.

**[0047]** In another aspect, the present disclosure relates to methods of making halogenated psilocybin derivatives. Accordingly, in one aspect, the present disclosure provides in at least one aspect, a method of making a halogenated psilocybin derivative the method comprising:

- 5 (a) contacting a halogenated psilocybin precursor compound with a host cell comprising a psilocybin biosynthetic enzyme complement; and
- (b) growing the host cell to produce a halogenated psilocybin derivative or salts thereof having the formula (I):



10

wherein, at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> is a halogen atom, wherein each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group, wherein R<sub>4</sub> when it is not halogenated is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group, and wherein R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group.

15

**[0048]** In at least one embodiment, in an aspect, the halogenated psilocybin precursor compound can be contacted with the host cell by including the psilocybin precursor compound in a growth medium for the host cell.

**[0049]** In at least one embodiment, in an aspect, the halogenated psilocybin precursor compound can be formed by contacting the host cell with a non-halogenated psilocybin precursor compound and a halogen, the host cell further comprising a halogenase capable of halogenating the non-halogenated psilocybin compound and forming the halogenated psilocybin precursor compound.

20

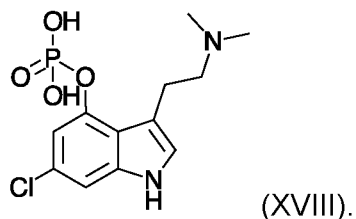
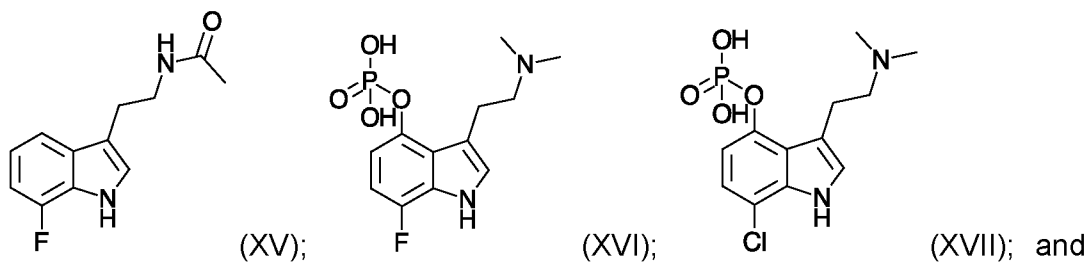
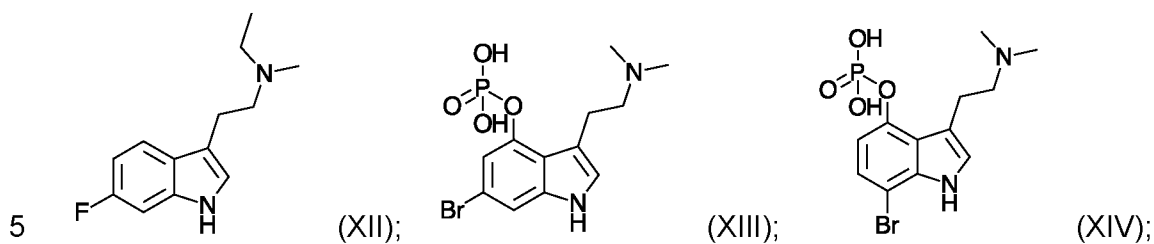
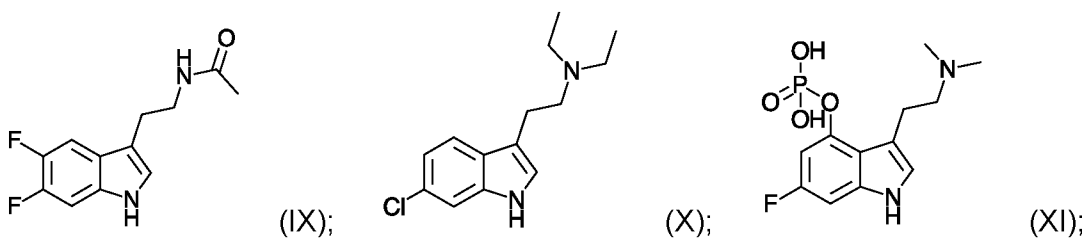
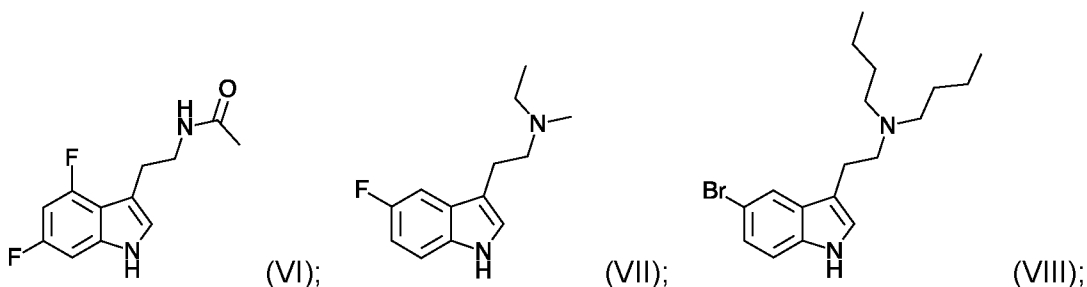
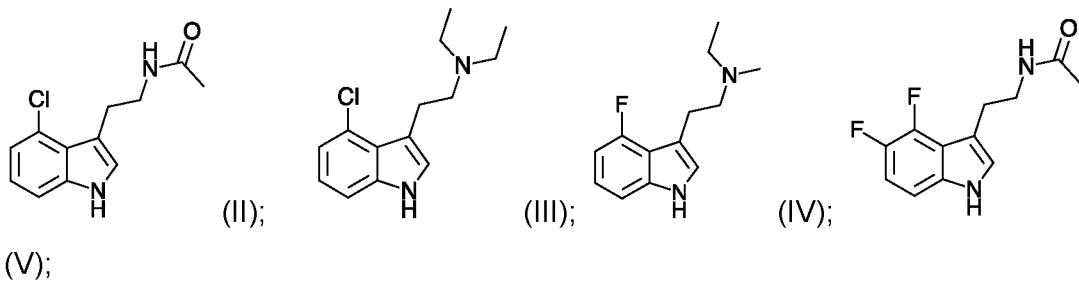
**[0050]** In at least one embodiment, in an aspect, R<sub>2</sub> can be a halogen atom and R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can each be a hydrogen atom, and R<sub>4</sub> can be a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.

25

**[0051]** In at least one embodiment, in an aspect, R<sub>4</sub> can be a halogen atom and R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can each be a hydrogen atom or an alkyl group or O-alkyl group.

- [0052]** In at least one embodiment, in an aspect, R<sub>5</sub> can be a halogen atom and R<sub>2</sub>, R<sub>6</sub> and R<sub>7</sub> can each be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.
- 5 **[0053]** In at least one embodiment, in an aspect, R<sub>6</sub> can be a halogen atom and R<sub>2</sub>, R<sub>5</sub> and R<sub>7</sub> can each be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.
- 10 **[0054]** In at least one embodiment, in an aspect, R<sub>7</sub> can be a halogen atom and R<sub>2</sub>, R<sub>5</sub> and R<sub>6</sub> can each be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.
- [0055]** In at least one embodiment, in an aspect, at least two of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> can be a halogen atom.
- 15 **[0056]** In at least one embodiment, in an aspect, the at least of two of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> that are halogen atoms can be two identical halogen atoms.
- [0057]** In at least one embodiment, in an aspect, the at least of two of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> that are halogen atoms can be two non-identical halogen atoms.
- 20 **[0058]** In at least one embodiment, in an aspect, R<sub>2</sub> and R<sub>4</sub> can be a halogen atom, and R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group.
- [0059]** In at least one embodiment, in an aspect, R<sub>2</sub> and R<sub>5</sub> can be a halogen atom, R<sub>6</sub> and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a hydrogen atom, an alkyl group or O-alkyl group or a phosphate group.
- 25 **[0060]** In at least one embodiment, in an aspect, R<sub>2</sub> and R<sub>6</sub> can be a halogen atom, R<sub>5</sub> and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a hydrogen atom, an alkyl group or O-alkyl group or a phosphate group.
- 30 **[0061]** In at least one embodiment, in an aspect, R<sub>2</sub> and R<sub>7</sub> can be a halogen atom, R<sub>5</sub> and R<sub>6</sub> can be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a hydrogen atom, an alkyl group or O-alkyl group or a phosphate group.

- [0062]** In at least one embodiment, in an aspect, R<sub>4</sub> and R<sub>5</sub> can be a halogen atom, and R<sub>2</sub>, R<sub>6</sub> and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group.
- [0063]** In at least one embodiment, in an aspect, R<sub>4</sub> and R<sub>6</sub> can be a  
5 halogen atom, and R<sub>2</sub>, R<sub>5</sub> and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group.
- [0064]** In at least one embodiment, in an aspect, R<sub>4</sub> and R<sub>7</sub> can be a halogen atom, and R<sub>2</sub>, R<sub>5</sub> and R<sub>6</sub> can be a hydrogen atom or an alkyl group or O-alkyl group.
- 10 **[0065]** In at least one embodiment, in an aspect, R<sub>5</sub> and R<sub>6</sub> can be a halogen atom, R<sub>2</sub> and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a hydrogen atom, an alkyl group or O-alkyl group or a phosphate group.
- [0066]** In at least one embodiment, in an aspect, R<sub>5</sub> and R<sub>7</sub> can be a  
15 halogen atom, R<sub>2</sub> and R<sub>6</sub> can be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a hydrogen atom, an alkyl group or O-alkyl group or a phosphate group.
- [0067]** In at least one embodiment, in an aspect, R<sub>6</sub> and R<sub>7</sub> can be a  
20 halogen atom, R<sub>2</sub> and R<sub>5</sub> can be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a hydrogen atom, an alkyl group or O-alkyl group or a phosphate group.
- [0068]** In at least one embodiment, in an aspect, R<sub>4</sub> when it is not halogenated can be a hydrogen atom.
- [0069]** In at least one embodiment, in an aspect, R<sub>4</sub> when it is not  
25 halogenated can be an alkyl group.
- [0070]** In at least one embodiment, in an aspect, R<sub>4</sub> when it is not halogenated can be an O-alkyl group.
- [0071]** In at least one embodiment, in an aspect, R<sub>4</sub> when it is not halogenated can be a phosphate group.
- 30 **[0072]** In at least one embodiment, in an aspect, the halogenated psilocybin derivative can be a derivative selected from the group consisting of derivatives having formulas (II); (III); (IV); (V); (VI); (VII); (VIII); (IX); (X); (XI); (XII); (XIII); (XIV); (XV); (XVI); and (XVII):



**[0073]** In at least one embodiment, in an aspect, the halogenated psilocybin precursor compound can be selected from the group of halogenated compounds consisting of halogenated tryptophan, halogenated tryptamine, halogenated 4-



hydroxytryptamine, halogenated 4-hydroxy-indole, halogenated 4-hydroxytryptophan, halogenated norbaeocystin, and halogenated baeocystin.

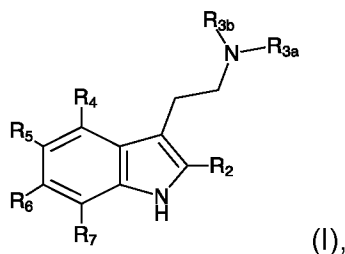
**[0074]** In at least one embodiment, in an aspect, the psilocybin biosynthetic enzyme complement can comprise at least one enzyme encoded by a nucleic acid selected from:

- (a) SEQ.ID NO: 1, SEQ.ID NO: 3, SEQ.ID NO: 5, SEQ.ID NO: 7, SEQ.ID NO: 9, SEQ.ID NO 11, SEQ.ID NO: 36, SEQ.ID NO: 38, SEQ.ID NO: 40, SEQ. ID NO: 42, SEQ.ID NO: 44, SEQ.ID NO: 46, SEQ.ID NO: 48, SEQ.ID NO: 50, SEQ.ID NO: 52, SEQ.ID NO: 54, SEQ. ID NO: 56, SEQ.ID NO: 58, SEQ.ID NO: 60, SEQ.ID NO: 64, SEQ.ID NO: 66, SEQ.ID NO: 68, SEQ. ID NO: 70, and SEQ.ID NO: 72;
- (b) a nucleic acid sequence that is substantially identical to any one of the nucleic acid sequences of (a);
- (c) a nucleic acid sequence that is substantially identical to any one of the nucleic acid sequences of (a) but for the degeneration of the genetic code;
- (d) a nucleic acid sequence that is complementary to any one of the nucleic acid sequences of (a);
- (e) a nucleic acid sequence encoding a polypeptide having any one of the amino acid sequences set forth in SEQ.ID NO: 2, SEQ.ID NO: 4, SEQ.ID NO: 6, SEQ.ID NO: 8, SEQ.ID NO: 10, SEQ.ID NO 12, SEQ.ID NO: 37, SEQ.ID NO: 39, SEQ.ID NO: 41, SEQ. ID NO: 43, SEQ.ID NO: 45, SEQ.ID NO: 47, SEQ.ID NO: 49, SEQ.ID NO: 51, SEQ.ID NO: 53, SEQ.ID NO: 55, SEQ. ID NO: 57, SEQ.ID NO: 59, SEQ.ID NO: 61, SEQ.ID NO: 65, SEQ.ID NO: 67, SEQ.ID NO: 69, SEQ. ID NO: 71, and SEQ.ID NO: 73;
- (f) a nucleic acid sequence that encodes a functional variant of any one of the amino acid sequences set forth in SEQ.ID NO: 2, SEQ.ID NO: 4, SEQ.ID NO: 6, SEQ.ID NO: 8, SEQ.ID NO: 10, SEQ.ID NO 12, SEQ.ID NO: 37, SEQ.ID NO: 39, SEQ.ID NO: 41, SEQ. ID NO: 43, SEQ.ID NO: 45, SEQ.ID NO: 47, SEQ.ID NO: 49, SEQ.ID NO: 51, SEQ.ID NO: 53, SEQ.ID NO: 55, SEQ. ID NO: 57, SEQ.ID NO: 59, SEQ.ID NO: 61, SEQ.ID NO: 65, SEQ.ID NO: 67, SEQ.ID NO: 69, SEQ. ID NO: 71, and SEQ.ID NO: 73; and
- (g) a nucleic acid sequence that hybridizes under stringent conditions to any one of the nucleic acid sequences set forth in (a), (b), (c), (d), (e) or (f).

**[0075]** In at least one embodiment, in an aspect, the halogenase can be encoded by a nucleic acid selected from:

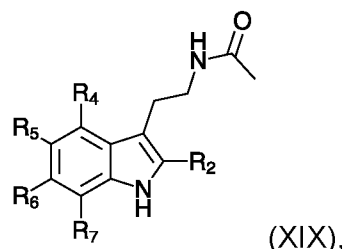
- (a) SEQ.ID NO: 13, SEQ.ID NO: 15, SEQ. ID NO: 17, SEQ.ID NO: 19, SEQ. ID NO: 21, SEQ.ID NO: 23, SEQ.ID NO: 25, and SEQ.ID NO: 27;
- 5 (b) a nucleic acid sequence that is substantially identical to any one of the nucleic acid sequences of (a);
- (c) a nucleic acid sequence that is substantially identical to any one of the nucleic acid sequences of (a) but for the degeneration of the genetic code;
- 10 (d) a nucleic acid sequence that is complementary to any one of the nucleic acid sequences of (a);
- (e) a nucleic acid sequence encoding a polypeptide having any one of the amino acid sequences set forth in SEQ.ID NO: 14, SEQ.ID NO: 16, SEQ. ID NO: 18, SEQ.ID NO: 20, SEQ. ID NO: 22, SEQ.ID NO: 24, SEQ.ID
- 15 NO: 26, SEQ.ID NO: 28, SEQ.ID NO: 29, SEQ.ID NO: 30, SEQ.ID NO: 31, SEQ.ID NO: 32. SEQ.ID NO: 33, SEQ.ID NO: 34, and SEQ.ID NO: 35;
- (f) a nucleic acid sequence that encodes a functional variant of any one of the amino acid sequences set forth in SEQ.ID NO: 14, SEQ.ID NO: 16, SEQ. ID NO: 18, SEQ.ID NO: 20, SEQ. ID NO: 22, SEQ.ID NO: 24, SEQ.ID
- 20 NO: 26, SEQ.ID NO: 28, SEQ.ID NO: 29, SEQ.ID NO: 30, SEQ.ID NO: 31, SEQ.ID NO: 32. SEQ.ID NO: 33, SEQ.ID NO: 34, and SEQ.ID NO: 35; and
- (g) a nucleic acid sequence that hybridizes under stringent conditions to any one of the nucleic acid sequences set forth in (a), (b), (c), (d), (e) or (f).

**[0076]** In at least one embodiment, in an aspect, the host cell can further  
 25 comprise an acetyl transferase capable of acetylating R<sub>3a</sub> or R<sub>3b</sub>, wherein the cell comprises a compound having chemical formula (I):



wherein, at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> is a halogen atom, wherein each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or an alkyl group or

O-alkyl group, wherein  $R_4$  when it is not halogenated is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group, and wherein at least one of  $R_{3A}$  and  $R_{3B}$  is a hydrogen atom, to form a compound having the chemical formula (XIX):



wherein, at least one of  $R_2$ ,  $R_4$ ,  $R_5$ ,  $R_6$  or  $R_7$  is a halogen atom, wherein each non-halogenated  $R_2$ ,  $R_5$ ,  $R_6$ , or  $R_7$  is a hydrogen atom or an alkyl group or O-alkyl group, wherein  $R_4$  when it is not halogenated is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.

10 **[0077]** In at least one embodiment, the acetylase can be encoded by a nucleic acid selected from:

- (a) SEQ.ID NO: 62;
- (b) a nucleic acid sequence that is substantially identical the nucleic acid sequence of (a);
- 15 (c) a nucleic acid sequence that is substantially identical to the nucleic acid sequence of (a) but for the degeneration of the genetic code;
- (d) a nucleic acid sequence that is complementary to the nucleic acid sequence of (a);
- (e) a nucleic acid sequence encoding a polypeptide having an amino acid sequences set forth in SEQ.ID NO: 63;
- 20 (f) a nucleic acid sequence that encodes a functional variant of the amino acid sequence set forth in SEQ.ID NO: 63; and
- (g) a nucleic acid sequence that hybridizes under stringent conditions to any one of the nucleic acid sequences set forth in (a), (b), (c), (d), (e) or (f).

25 **[0078]** In at least one embodiment, in an aspect,

- (a) the halogenated psilocybin precursor compound can be a halogenated indole, and the a psilocybin biosynthetic enzyme complement can comprise
- (i) a nucleic acid encoding TrpB selected from SEQ.ID NO: 12, SEQ.ID NO: 65, SEQ.ID NO: 67, SEQ.ID NO: 69, SEQ.ID NO: 71, and SEQ.ID NO: 73,

or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; (ii) a nucleic acid encoding PsiD selected from SEQ.ID NO: 2, SEQ.ID NO: 10, SEQ.ID NO: 37, SEQ.ID NO: 39, SEQ.ID NO: 41, SEQ.ID NO: 43, SEQ.ID NO: 45, SEQ.ID NO: 47, SEQ.ID NO: 49, SEQ.ID NO: 51, SEQ.ID NO: 53, SEQ.ID NO: 55, and SEQ.ID NO: 57, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; (iii) a nucleic acid encoding PsiK encoded by SEQ.ID NO: 6, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; and (iv) a nucleic acid encoding PsiM selected from SEQ.ID NO: 8, SEQ.ID NO: 59 and SEQ.ID NO: 61 or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; or

(b) the halogenated psilocybin precursor compound can be a halogenated tryptophan, and the psilocybin biosynthetic enzyme complement can comprise (i) a nucleic acid encoding PsiD selected from SEQ.ID NO: 2, SEQ.ID NO: 10, SEQ.ID NO: 37, SEQ.ID NO: 39, SEQ.ID NO: 41, SEQ.ID NO: 43, SEQ.ID NO: 45, SEQ.ID NO: 47, SEQ.ID NO: 49, SEQ.ID NO: 51, SEQ.ID NO: 53, SEQ.ID NO: 55, and SEQ.ID NO: 57, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; (ii) a nucleic acid encoding PsiK encoded by SEQ.ID NO: 6, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; and (iii) a nucleic acid encoding PsiM selected from SEQ.ID NO: 8, SEQ.ID NO: 59 and SEQ.ID NO: 61 or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; or

(c) the halogenated psilocybin precursor compound can be a halogenated tryptamine, and the psilocybin biosynthetic enzyme complement can comprise (i) a nucleic acid encoding PsiK encoded by SEQ.ID NO: 6, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; and (ii) and a nucleic acid encoding PsiM selected from SEQ.ID NO: 8, SEQ.ID NO: 59 and SEQ.ID NO: 61 or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; or

(d) the halogenated psilocybin precursor compound can be a halogenated norbaeocystin or a halogenated baecocystin and the psilocybin biosynthetic enzyme complement can comprise a nucleic acid encoding PsiM selected

from SEQ.ID NO: 8, SEQ.ID NO: 59 and SEQ.ID NO: 61 or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; or (e) the halogenated psilocybin precursor compound can be a halogenated tryptophan, and the psilocybin biosynthetic enzyme complement can  
5 comprise (i) a nucleic acid encoding PsiD selected from SEQ.ID NO: 2, SEQ.ID NO: 10, SEQ.ID NO: 37, SEQ.ID NO: 39, SEQ.ID NO: 41, SEQ.ID NO: 43, SEQ.ID NO: 45, SEQ.ID NO: 47, SEQ.ID NO: 49, SEQ.ID NO: 51, SEQ.ID NO: 53, SEQ.ID NO: 55, and SEQ.ID NO: 57, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto;  
10 (ii) a nucleic acid encoding PsiH encoded by SEQ.ID NO: 4, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; (iii) a nucleic acid encoding PsiK encoded by SEQ.ID NO: 6, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; and (iv) and a nucleic acid encoding PsiM selected from  
15 SEQ.ID NO: 8, SEQ.ID NO: 59 and SEQ.ID NO: 61 or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; or (f) the halogenated psilocybin precursor compound can be, a halogenated tryptamine, and psilocybin biosynthetic enzyme complement can comprise (i) a nucleic acid encoding PsiH encoded by SEQ.ID NO: 4, or a nucleic  
20 acid sequence encoding a polypeptide sequence substantially similar thereto; (ii) a nucleic acid encoding PsiK encoded by SEQ.ID NO: 6, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; and (iii) and a nucleic acid encoding PsiM selected from SEQ.ID NO: 8, SEQ.ID NO: 59 and SEQ.ID NO: 61 or a nucleic  
25 sequence encoding a polypeptide sequence substantially similar thereto.

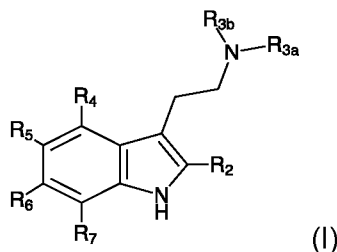
**[0079]** In at least one embodiment, in an aspect, the method can further include a step comprising isolating the halogenated psilocybin derivative.

**[0080]** In at least one embodiment, in an aspect, the host cell can be a microbial cell.

30 **[0081]** In at least one embodiment, in an aspect, the host cell can be a bacterial cell or a yeast cell.

**[0082]** In another aspect, the present disclosure provides, in at least one embodiment, a method for modulating a 5-HT<sub>2A</sub> receptor, the method comprising

contacting a 5-HT<sub>2A</sub> receptor with a chemical compound or salts thereof having the formula (I):

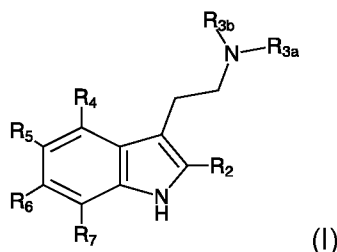


- 5 wherein, at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> is a halogen atom, wherein each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group, wherein R<sub>4</sub> when it is not halogenated is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group, and wherein R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group under reaction conditions
- 10 sufficient to thereby modulate receptor activity.

**[0083]** In at least one embodiment, in an aspect, the reaction conditions can be *in vitro* reaction conditions.

**[0084]** In at least one embodiment, in an aspect, the reaction conditions can be *in vivo* reaction conditions.

- 15 **[0085]** In another aspect, the present disclosure provides, in at least one embodiment, a use of a chemical compound or salts thereof having the formula (I):

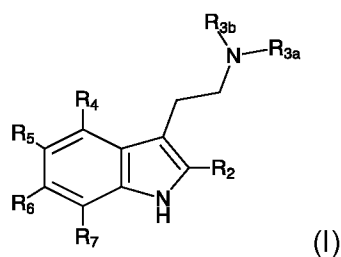


- 20 wherein, at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> is a halogen atom, wherein each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group, wherein R<sub>4</sub> when it is not halogenated is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group, and wherein R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group in the manufacture of a
- 25 pharmaceutical or recreational drug formulation.

**[0086]** In at least one embodiment, in an aspect, the manufacture of a pharmaceutical or recreational drug formulation can comprise formulating the chemical compound with an excipient, diluent or carrier.

**[0087]** In at least one embodiment, the manufacture of a pharmaceutical or recreational drug formulation can comprise derivatizing the chemical compound having the formula (I) by substituting the halogen atom with another atom or chemical group.

**[0088]** In another aspect, the present disclosure provides, in at least one embodiment, a use of a chemical compound or salts thereof having the formula (I):



wherein, at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> is a halogen atom, wherein each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group, wherein R<sub>4</sub> when it is not halogenated is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group, and wherein R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group together with a diluent, carrier or excipient, as a pharmaceutical or recreational drug formulation.

**[0089]** Other features and advantages will become apparent from the following detailed description. It should be understood, however, that the detailed description, while indicating preferred implementations of the disclosure, are given by way of illustration only, since various changes and modifications within the spirit and scope of the disclosure will become apparent to those of skill in the art from the detailed description.

## **BRIEF DESCRIPTION OF THE DRAWINGS**

**[0090]** The disclosure is in the hereinafter provided paragraphs described, by way of example, in relation to the attached figures. The figures provided herein are provided for a better understanding of the example embodiments and to show

more clearly how the various embodiments may be carried into effect. The figures are not intended to limit the present disclosure.

**[0091]** FIG. 1 depicts the chemical structure of psilocybin.

**[0092]** FIG. 2 depicts a certain prototype structure of a psilocybin precursor compound, namely an indole. Certain carbon and nitrogen atoms may be referred herein by reference to their position within the indole structure, *i.e.* N<sub>1</sub>, C<sub>2</sub>, C<sub>3</sub> *etc.* The pertinent atom numbering is shown.

**[0093]** FIGS. 3A, 3B, 3C, 3D, 3E, 3F and 3G depict the chemical structures of example psilocybin precursor compounds, notably L-tryptophan (FIG. 3A), tryptamine (FIG. 3B), 4-hydroxytryptamine (FIG. 3C), L-4-hydroxytryptophan (FIG. 3D), 4-hydroxyindole (FIG. 3E), norbaeocystin (FIG. 3F), and baeocystin (FIG. 3G).

**[0094]** FIGS. 4A, 4B, 4C, 4D, 4E, 4F, 4G, 4H, 4I, 4J, 4K, 4L, 4M, 4N, 4O, 4P, 4Q and 4R depict the chemical structures of certain example halogenated psilocybin derivative compounds, notably a 2-chloro-4-phospho psilocybin derivative (FIG. 4A); a 2-bromo psilocybin derivative (FIG. 4B); a 4-iodo psilocybin derivative (FIG. 4C); a 4-fluoro psilocybin derivative (FIG. 4D); a 4-phospho-5-fluoro psilocybin derivative (FIG. 4E); a 5-chloro psilocybin derivative (FIG. 4F); a 4-phospho-6-bromo psilocybin derivative (FIG. 4G); a 6-chloro psilocybin derivative (FIG. 4H) a 4-phospho-7-fluoro psilocybin derivative (FIG. 4I); a 7-iodo psilocybin derivative (FIG. 4J); a 2-chloro-4-methyl psilocybin derivative (FIG. 4K); a 4-ethyl-5-bromo psilocybin derivative (FIG. 4L); a 4-methyl-6-fluoro psilocybin derivative (FIG. 4M); a 4-propyl-7-iodo psilocybin derivative (FIG. 4N); a 2-chloro-4-O-methyl psilocybin derivative (FIG. 4K); a 4-O-ethyl-5-bromo psilocybin derivative (FIG. 4L); a 4-O-methyl-6-fluoro psilocybin derivative (FIG. 4M); and a 4-O-propyl-7-iodo psilocybin derivative (FIG. 4N). It is noted that in each of FIGS. 4A, 4B, 4C, 4D, 4E, 4F, 4G, 4H, 4I, 4J, 4K, 4L, 4M, 4N, 4O, 4P, 4Q and 4R R<sub>3a</sub> and R<sub>3b</sub> can be a hydrogen atom, an alkyl group, an aryl group, or an acyl group.

**[0095]** FIGS. 5A, 5B, 5C, 5D, 5E, 5F, 5G, 5H, 5I and 5J depict the chemical structures of certain example halogenated psilocybin derivative compounds, notably a 2,4-di-chloro psilocybin derivative (FIG. 5A); a 2-bromo-5-fluoro-7-methyl psilocybin derivative (FIG. 5B); a 2-chloro-O-methyl-6-bromo psilocybin derivative (FIG. 5C); a 2-fluoro-4-phospho-7-iodo psilocybin derivative (FIG. 5D); a 4,5-di-bromo psilocybin derivative (FIG. 5E); a 4,6-di-chloro psilocybin derivative



(**FIG. 5F**); a 4-chloro-5-fluoro psilocybin derivative (**FIG. 5G**); a 4-phospho-5-fluoro-6-bromo psilocybin derivative (**FIG. 5H**); a 4-phospho-5,7-di-fluoro psilocybin derivative (**FIG. 5I**); and a 6,7-di-iodo psilocybin derivative (**FIG. 5J**). It is noted that in each of **FIGS. 5A, 5B, 5C, 5D, 5E, 5F, 5G, 5H, 5I** and **5J**  $R_{3a}$  and  $R_{3b}$  can be a hydrogen atom, an alkyl group, an aryl group, or an acyl group.

**[0096]** **FIGS. 6A, 6B, 6C, 6D, 6E, and 6F** depict the chemical structures of certain example halogenated psilocybin derivative compounds, notably a 2,4,5-trichloro psilocybin derivative (**FIG. 6A**); a 2-bromo-5,6-di-fluoro psilocybin derivative (**FIG. 6B**); a 2-chloro-5-bromo-7-fluoro psilocybin derivative (**FIG. 6C**); a 4,5-dibromo-6-chloro psilocybin derivative (**FIG. 6D**); a 4,5,7-tri-bromo psilocybin derivative (**FIG. 6E**); and a 4-phospho-5-fluoro-6-chloro-7-iodo psilocybin derivative (**FIG. 6F**). It is noted that in each of **FIGS. 6A, 6B, 6C, 6D, 6E, and 6F**  $R_{3a}$  and  $R_{3b}$  can be a hydrogen atom, an alkyl group, an aryl group, or an acyl group.

**[0097]** **FIGS. 7A, 7B, 7C, 7D, and 7E** depict the chemical structures of certain example halogenated psilocybin derivative compounds, notably a 2,4,5,6-tetra-chloro psilocybin derivative (**FIG. 7A**); a 4,7-di-chloro-5,6-di-fluoro psilocybin derivative (**FIG. 7B**); a 2-chloro-4-phospho-5-fluoro-6-chloro-7-iodo psilocybin derivative (**FIG. 7C**); a 2,4,7-tri-bromo-6-chloro psilocybin derivative (**FIG. 7D**); and a 2-fluoro-4,5,7-tri-bromo psilocybin derivative (**FIG. 7E**). It is noted that in each of **FIGS. 7A, 7B, 7C, 7D, and 7E**  $R_{3a}$  and  $R_{3b}$  can be a hydrogen atom, an alkyl group, an aryl group, or an acyl group.

**[0098]** **FIGS. 8A and 8B** depict example biosynthesis pathway relating to the synthesis of a halogenated psilocybin derivatives.

**[0099]** **FIGS. 9A and 9B** depict example chemical synthesis processes for the synthesis of certain example halogenated psilocybin derivatives, notably the compound denoted as **410B01** (**FIG. 9A**) and corresponding with a halogenated psilocybin derivative compound having the chemical formula (IV) set forth herein, and the compound denoted as **1** (**FIG. 9B**) and corresponding with a halogenated psilocybin derivative compound having the chemical formula (VIII) set forth herein.

**[00100]** **FIG. 10** depicts a representation of mass spectrometry data in the form of a chromatogram, notably a chromatogram obtained in the performance of an experiment to synthesize an example halogenated psilocybin derivative compound having the chemical formula (XI) set forth herein.

**[00101]** **FIG. 11** depicts a representation of mass spectrometry data in the form of a chromatogram, notably a chromatogram obtained in the performance of an experiment to synthesize an example halogenated psilocybin derivative compound having the chemical formula (XIV) set forth herein.

5 **[00102]** **FIG. 12** depicts a representation of mass spectrometry data in the form of a chromatogram, notably a chromatogram obtained in the performance of an experiment to synthesize an example halogenated psilocybin derivative compound having the chemical formula (XVI) set forth herein.

**[00103]** **FIG. 13** depicts a representation of mass spectrometry data in the  
10 form of a chromatogram, notably a chromatogram obtained in the performance of an experiment to synthesize an example halogenated psilocybin derivative compound having the chemical formula (XVII) set forth herein.

**[00104]** **FIG. 14** depicts a graph obtained in the performance of an experimental assay to evaluate the efficacy of an example halogenated psilocybin derivative, notably a cell viability assay involving an example halogenated  
15 psilocybin derivative compound having the chemical formula (XV) set forth herein.

**[00105]** **FIGS. 15A** and **15B** depict graphs obtained in the performance of an experimental assay to evaluate the efficacy of an example halogenated psilocybin derivative, notably a 5-HT<sub>2a</sub> receptor modulation assay involving +5-HT<sub>2A</sub> cells and an example halogenated psilocybin derivative compound having the chemical formula (XV) set forth herein (**FIG. 15B**) and a control (psilocin) (**FIG. 15A**).

**[00106]** **FIGS. 16A** and **16B** depict graphs obtained in the performance of an experimental assay to evaluate the efficacy of an example halogenated psilocybin derivative, notably a 5-HT<sub>2a</sub> receptor modulation assay involving -5-HT<sub>2A</sub> cells and an example halogenated psilocybin derivative compound having the chemical formula (XV) set forth herein (**FIG. 16B**) and a control (psilocin) (**FIG. 16A**).

**[00107]** **FIG. 17** depicts a representation of mass spectrometry data in the  
30 form of a chromatogram, notably a chromatogram obtained in the performance of an experiment to synthesize an example halogenated psilocybin derivative compound having the chemical formula (II) set forth herein.

**[00108]** **FIG. 18** depicts a graph obtained in the performance of an experimental assay to evaluate the efficacy of an example halogenated psilocybin

derivative, notably a cell viability assay involving an example halogenated psilocybin derivative compound having the chemical formula (IX) set forth herein.

**[00109]**      **FIGS. 19A and 19B** depict graphs obtained in the performance of an experimental assay to evaluate the efficacy of an example halogenated psilocybin derivative, notably a 5-HT<sub>2a</sub> receptor modulation assay involving +5-HT<sub>2A</sub> cells (**FIG. 19A**) and -5-HT<sub>2A</sub> cells (**FIG. 19B**) and an example halogenated psilocybin derivative compound having the chemical formula (IX) set forth herein.

**[00110]**      **FIG. 20** depicts a representation of mass spectrometry data in the form of a chromatogram, notably a chromatogram obtained in the performance of an experiment to synthesize an example halogenated psilocybin derivative compound having the chemical formula (VI) set forth herein.

**[00111]**      **FIG. 21** depicts a graph obtained in the performance of an experimental assay to evaluate the efficacy of an example halogenated psilocybin derivative, notably a cell viability assay involving an example halogenated psilocybin derivative compound having the chemical formula (V) set forth herein.

**[00112]**      **FIGS. 22A and 22B** depict graphs obtained in the performance of an experimental assay to evaluate the efficacy of an example halogenated psilocybin derivative, notably a 5-HT<sub>2a</sub> receptor modulation assay involving +5-HT<sub>2A</sub> cells (**FIG. 22A**) and -5-HT<sub>2A</sub> cells (**FIG. 22B**), and an example halogenated psilocybin derivative compound having the chemical formula (V) set forth herein.

**[00113]**      **FIG. 23** depicts a representation of mass spectrometry data in the form of a chromatogram, notably a chromatogram obtained in the performance of an experiment to synthesize an example halogenated psilocybin derivative compound having the chemical formula (XIII) set forth herein.

**[00114]**      **FIG. 24** depicts a graph obtained in the performance of an experimental assay to evaluate the efficacy of an example halogenated psilocybin derivative, notably a cell viability assay involving an example halogenated psilocybin derivative compound having the chemical formula (X) set forth herein.

**[00115]**      **FIGS. 25A and 25B** depict graphs obtained in the performance of an experimental assay to evaluate the efficacy of an example halogenated psilocybin derivative, notably a 5-HT<sub>2a</sub> receptor modulation assay involving +5-HT<sub>2A</sub> cells (**FIG. 25A**) and -5-HT<sub>2A</sub> cells (**FIG. 25B**), and an example halogenated psilocybin derivative compound having the chemical formula (X) set forth herein.

**[00116]** **FIG. 26** depicts a graph obtained in the performance of an experimental assay to evaluate the efficacy of an example halogenated psilocybin derivative, notably a cell viability assay involving an example halogenated psilocybin derivative compound having the chemical formula (IV) set forth herein.

5 **[00117]** **FIGS. 27A** and **27B** depict graphs obtained in the performance of an experimental assay to evaluate the efficacy of an example halogenated psilocybin derivative, notably a 5-HT<sub>2a</sub> receptor modulation assay involving +5-HT<sub>2A</sub> cells (**FIG. 27A**) and -5-HT<sub>2A</sub> cells (**FIG. 27B**), and an example halogenated psilocybin derivative compound having the chemical formula (IV) set forth herein.

10 **[00118]** **FIG. 28** depicts a graph obtained in the performance of an experimental assay to evaluate the efficacy of an example halogenated psilocybin derivative, notably a cell viability assay involving an example halogenated psilocybin derivative compound having the chemical formula (VIII) set forth herein.

**[00119]** **FIGS. 29A** and **29B** depict graphs obtained in the performance of an experimental assay to evaluate the efficacy of an example halogenated psilocybin derivative, notably a 5-HT<sub>2a</sub> receptor modulation assay involving +5-HT<sub>2A</sub> cells (**FIG. 29A**) and -5-HT<sub>2A</sub> cells (**FIG. 29B**), and an example halogenated psilocybin derivative compound having the chemical formula (VIII) set forth herein.

**[00120]** **FIG. 30** depicts a representation of mass spectrometry data in the form of a chromatogram, notably a chromatogram obtained in the performance of an experiment to synthesize an example halogenated psilocybin derivative compound having the chemical formula (XVIII) set forth herein.

**[00121]** The figures together with the following detailed description make apparent to those skilled in the art how the disclosure may be implemented in practice.

#### **DETAILED DESCRIPTION**

**[00122]** Various compositions, systems or processes will be described below to provide an example of an embodiment of each claimed subject matter.

30 No embodiment described below limits any claimed subject matter and any claimed subject matter may cover processes, compositions or systems that differ from those described below. The claimed subject matter is not limited to compositions, processes or systems having all of the features of any one composition, system or process described below or to features common to multiple

or all of the compositions, systems or processes described below. It is possible that a composition, system or process described below is not an embodiment of any claimed subject matter. Any subject matter disclosed in a composition, system or process described below that is not claimed in this document may be the subject matter of another protective instrument, for example, a continuing patent application, and the applicant(s), inventor(s) or owner(s) do not intend to abandon, disclaim or dedicate to the public any such subject matter by its disclosure in this document.

**[00123]** As used herein and in the claims, the singular forms, such “a”, “an” and “the” include the plural reference and vice versa unless the context clearly indicates otherwise. Throughout this specification, unless otherwise indicated, “comprise,” “comprises” and “comprising” are used inclusively rather than exclusively, so that a stated integer or group of integers may include one or more other non-stated integers or groups of integers.

**[00124]** Various compositions, systems or processes will be described below to provide an example of an embodiment of each claimed subject matter. No embodiment described below limits any claimed subject matter and any claimed subject matter may cover processes, compositions or systems that differ from those described below. The claimed subject matter is not limited to compositions, processes or systems having all of the features of any one composition, system or process described below or to features common to multiple or all of the compositions, systems or processes described below. It is possible that a composition, system or process described below is not an embodiment of any claimed subject matter. Any subject matter disclosed in a composition, system or process described below that is not claimed in this document may be the subject matter of another protective instrument, for example, a continuing patent application, and the applicant(s), inventor(s) or owner(s) do not intend to abandon, disclaim or dedicate to the public any such subject matter by its disclosure in this document.

**[00125]** When ranges are used herein for physical properties, such as molecular weight, or chemical properties, such as chemical formulae, all combinations and sub-combinations of ranges and specific embodiments therein are intended to be included. Other than in the operating examples, or where otherwise indicated, all numbers expressing quantities of ingredients or reaction

conditions used herein should be understood as modified in all instances by the term "about." The term "about" when referring to a number or a numerical range means that the number or numerical range referred to is an approximation within experimental variability (or within statistical experimental error), and thus the number or numerical range may vary between 1% and 15% of the stated number or numerical range, as will be readily recognized by context. Furthermore any range of values described herein is intended to specifically include the limiting values of the range, and any intermediate value or sub-range within the given range, and all such intermediate values and sub-ranges are individually and specifically disclosed (e.g. a range of 1 to 5 includes 1, 1.5, 2, 2.75, 3, 3.90, 4, and 5). Similarly, other terms of degree such as "substantially" and "approximately" as used herein mean a reasonable amount of deviation of the modified term such that the end result is not significantly changed. These terms of degree should be construed as including a deviation of the modified term if this deviation would not negate the meaning of the term it modifies.

**[00126]** Unless otherwise defined, scientific and technical terms used in connection with the formulations described herein shall have the meanings that are commonly understood by those of ordinary skill in the art. The terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention, which is defined solely by the claims.

**[00127]** All publications, patents and patent applications are herein incorporated by reference in their entirety to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference in its entirety.

#### **Terms and definitions**

**[00128]** The term "psilocybin", as used herein, refers to a chemical compound having the structure set forth in **FIG. 1**, and further includes salts, thereof, such as a sodium salt, a potassium salt, *etc.*

**[00129]** The term "psilocybin precursor compound", as used herein, refers to a chemical compound that may serve as a precursor compound in the synthesis or biosynthesis of psilocybin, and includes compounds comprising an indole prototype structure, including, for example, tryptophan, tryptamine, 4-hydroxyindole, 4-hydroxytryptophan, 4-hydroxytryptamine, norbaeocystin, and

baeocystin, and further includes halogenated derivatives and salts of any of the foregoing.

**[00130]** The term “indole prototype structure”, as used herein, refers to the chemical structure shown in **FIG. 2**. It is noted that specific carbon atoms and a nitrogen atom in the indole prototype structure are numbered. Reference may be made to these carbon and nitrogen numbers herein, for example C<sub>2</sub>, C<sub>4</sub>, N<sub>1</sub>, and so forth. Furthermore, reference may be made to chemical groups attached to the indole prototype structure in accordance with the same numbering, for example R<sub>4</sub> and R<sub>6</sub> reference chemical groups attached to the C<sub>4</sub> and C<sub>6</sub> atom, respectively. In addition, R<sub>3A</sub> and R<sub>3B</sub>, in this respect, reference chemical groups extending from the ethyl-amino group extending in turn from the C<sub>3</sub> atom of the prototype indole structure.

**[00131]** The term “tryptophan”, as used herein, refers to a chemical compound having the structure set forth in **FIG. 3A** and further includes its D-enantiomeric form (not shown).

**[00132]** The term “tryptamine”, as used herein, refers to a chemical compound having the structure set forth in **FIG. 3B**.

**[00133]** The term “4-hydroxytryptamine”, as used herein, refers to a chemical compound having the structure set forth in **FIG. 3C**.

**[00134]** The term “4-hydroxytryptophan”, as used herein, refers to a chemical compound having the structure set forth in **FIG. 3D** and further includes its D-enantiomeric form (not shown).

**[00135]** The term “4-hydroxyindole”, as used herein, refers to a chemical compound having the structure set forth in **FIG. 3E**.

**[00136]** The term “norbaeocystin”, as used herein, refers to a chemical compound having the structure set forth in **FIG. 3F**.

**[00137]** The term “baeocystin”, as used herein, refers to a chemical compound having the structure set forth in **FIG. 3G**.

**[00138]** The term “phosphate group”, as used herein, is a molecule containing one atom of phosphorus, covalently bound to four oxygen atoms (three single bonds and one double bond). Of the four oxygen atoms one oxygen atom may be a hydroxy group, and one of the non-hydroxylated oxygen atom may be chemically bonded to another entity.

**[00139]** The terms “hydroxy group”, and “hydroxy”, as used herein, refers to a molecule containing one atom of oxygen bonded to one atom of hydrogen, and having the formula -OH. A hydroxy group through its oxygen atom may be chemically bonded to another entity.

5 **[00140]** The terms “halogen”, “halogenated” and “halo-”, as used herein, refer to the class of chemical elements consisting of fluorine (F), chlorine (Cl), bromine (Br), and iodine (I). Accordingly, halogenated compounds can refer to “fluorinated”, “chlorinated”, “brominated”, or “iodinated” compounds.

**[00141]** The term “halogenated psilocybin derivative”, as used herein, refers to a psilocybin derivative in which a halogen been bonded to psilocybin, or a derivative thereof. Reference may be made to specific carbon atoms of the psilocybin derivative which may be halogenated, for example, 7-halo-psilocybin refers to a halogenated psilocybin in which carbon atom number 7 (as identified in the indole prototype structure) is halogenated, or, similarly, 2-bromo-psilocybin refers to a brominated psilocybin in which carbon atom number 2 (as identified in the indole prototype structure) is brominated, and so on. The term further includes salts of halogenated psilocybin derivatives, such as sodium salt, potassium salt, or a chlorine salt, for example. Salts may be formed by virtue of a charged atom or group included in the psilocybin derivative, notably a charged nitrogen atom, and/or a charged phosphate group.

**[00142]** The term “halogenated psilocybin precursor compound”, as used herein, refers to a psilocybin precursor compound possessing a halogen. Reference may be made to specific carbon atoms of the psilocybin precursor compound which may be halogenated, for example, 6-halo-baeocystin refers to a halogenated baeocystin in which carbon atom number 6 (as identified in the indole prototype structure) is halogenated, or, similarly, 2-chloro-4-hydroxytryptamine refers to a chlorinated 4-hydroxytryptamine in which carbon atom number 2 (as identified in the indole prototype structure) is chlorinated. Furthermore, halogenated psilocybin precursor compounds include, without limitation, halogenated tryptophan, e.g. 7-bromo tryptophan, 6-chloro tryptophan, 5,6-difluoro tryptophan, halogenated 4-hydroxy indole e.g. 4-hydroxy-7-bromo indole, 4-hydroxy-5-chloro indole, 6,7-di-fluoro-indole

**[00143]** The term “alkyl”, as used herein, refers to a straight and/or branched chain, saturated alkyl radical containing from one to “p” carbon atoms (“C<sub>1</sub>-C<sub>p</sub>-



alkyl”) and includes, depending on the identity of “p”, methyl, ethyl, propyl, isopropyl, n-butyl, s-butyl, isobutyl, t-butyl, 2,2-dimethylbutyl, n-pentyl, 2-methylpentyl, 3-methylpentyl, 4-methylpentyl, n-hexyl and the like, where the variable p is an integer representing the largest number of carbon atoms in the alkyl radical. Alkyl groups further include hydrocarbon groups arranged in a chain having the chemical formula  $-C_nH_{2n+1}$ , including, without limitation, methyl groups ( $-CH_3$ ), ethyl groups ( $-C_2H_5$ ), propyl groups ( $-C_3H_7$ ), and butyl groups ( $-C_4H_9$ ).

**[00144]** The term “O-alkyl”, as used herein, refers to a hydrocarbon group arranged in a chain having the chemical formula  $-O-C_nH_{2n+1}$ . O-alkyl groups include, without limitation, O-methyl groups ( $-O-CH_3$ ), O-ethyl groups ( $-O-C_2H_5$ ), O-propyl groups ( $-O-C_3H_7$ ) and O-butyl groups ( $-O-C_4H_9$ ).

**[00145]** The term “acyl”, as used herein, refers to a carbon atom double bonded to an oxygen and single bonded to an alkyl group. The carbon atom further can be bonded to another entity. An acyl group can be described by the chemical formula:  $-C(=O)-C_nH_{2n+1}$ .

**[00146]** The term “aryl”, as used herein, refers to a monocyclic, bicyclic or tricyclic aromatic ring system containing, depending on the number of atoms in the rings, for example, from 6 to 14 carbon atoms ( $C_6-C_{14}$ -aryl) or from 6 to 10 carbons ( $C_6-C_{10}$ -aryl), and at least 1 aromatic ring and includes phenyl, naphthyl, anthracenyl, 1,2-dihydronaphthyl, 1,2,3,4-tetrahydronaphthyl, fluorenyl, phenanthrenyl, biphenylenyl, indanyl, indenyl and the like.

**[00147]** The term “5-HT<sub>2A</sub> receptor”, as used herein, refers to a subclass of a family of receptors for the neurotransmitter and peripheral signal mediator serotonin. 5-HT<sub>2A</sub> receptors can mediate a plurality of central and peripheral physiologic functions of serotonin. Central nervous system effects can include mediation of hallucinogenic effects of hallucinogenic compounds.

**[00148]** The term “modulating 5-HT<sub>2A</sub> receptors”, as used herein, refers to the ability of a compound disclosed herein to alter the function of 5-HT<sub>2A</sub> receptors. A 5-HT<sub>2A</sub> receptor modulator may activate the activity of a 5-HT<sub>2A</sub> receptor, may activate or inhibit the activity of a 5-HT<sub>2A</sub> receptor depending on the concentration of the compound exposed to the 5-HT<sub>2A</sub> receptor, or may inhibit the activity of a 5-HT<sub>2A</sub> receptor. Such activation or inhibition may be contingent on the occurrence of a specific event, such as activation of a signal transduction pathway, and/or maybe manifest only in particular cell types. The term “modulating 5-HT<sub>2A</sub>

receptors,” also refers to altering the function of a 5-HT<sub>2A</sub> receptor by increasing or decreasing the probability that a complex forms between a 5-HT<sub>2A</sub> receptor and a natural binding partner to form a multimer. A 5-HT<sub>2A</sub> receptor modulator may increase the probability that such a complex forms between the 5-HT<sub>2A</sub> receptor and the natural binding partner, may increase or decrease the probability that a complex forms between the 5-HT<sub>2A</sub> receptor and the natural binding partner depending on the concentration of the compound exposed to the 5-HT<sub>2A</sub> receptor, and or may decrease the probability that a complex forms between the 5-HT<sub>2A</sub> receptor and the natural binding partner. Furthermore, the term includes allosteric modulation of the receptor 5-HT<sub>2A</sub>, *i.e.* modulation of the 5-HT<sub>2A</sub> receptor through interaction with the 5-HT<sub>2A</sub> receptor that is topographically different than the orthosteric site recognized by the cell’s endogenous agonist, such modulation further including positive allosteric modulation (PAM), negative allosteric modulation (NAM) and silent allosteric modulation (SAM).

**[00149]** The term “5-HT<sub>2A</sub> receptor-mediated disorder”, as used herein, refers to a disorder that is characterized by abnormal 5-HT<sub>2A</sub> receptor activity. A 5-HT<sub>2A</sub> receptor-mediated disorder may be completely or partially mediated by modulating 5-HT<sub>2A</sub> receptors. In particular, a 5-HT<sub>2A</sub> receptor-mediated disorder is one in which modulation of 5-HT<sub>2A</sub> receptors results in some effect on the underlying disorder *e.g.*, administration of a 5-HT<sub>2A</sub> receptor modulator results in some improvement in at least some of the subjects being treated.

**[00150]** The term “pharmaceutical formulation”, as used herein, refers to a preparation in a form which allows an active ingredient, including a psychoactive ingredient, contained therein to provide effective treatment, and which does not contain any other ingredients which cause excessive toxicity, an allergic response, irritation, or other adverse response commensurate with a reasonable risk/benefit ratio. The pharmaceutical formulation may contain other pharmaceutical ingredients such as excipients, carriers, diluents, or auxiliary agents.

**[00151]** The term “recreational drug formulation”, as used herein, refers to a preparation in a form which allows a psychoactive ingredient contained therein to be effective for administration as a recreational drug, and which does not contain any other ingredients which cause excessive toxicity, an allergic response, irritation, or other adverse response commensurate with a reasonable risk/benefit

ratio. The recreational drug formulation may contain other ingredients such as excipients, carriers, diluents, or auxiliary agents.

**[00152]** The term “effective for administration as a recreational drug”, as used herein, refers to a preparation in a form which allows a subject to voluntarily induce a psychoactive effect for non-medical purposes upon administration, generally in the form of self-administration. The effect may include an altered state of consciousness, satisfaction, pleasure, euphoria, perceptual distortion, or hallucination.

**[00153]** The term “effective amount”, as used herein, refers to an amount of an active agent, pharmaceutical formulation or recreational drug formulation, sufficient to induce a desired biological or therapeutic effect, including a prophylactic effect, and further including a psychoactive effect. Such effect can include an effect with respect to the signs, symptoms or causes of a disorder, or disease or any other desired alteration of a biological system. The effective amount can vary depending, for example, on the health condition, injury stage, disorder stage, or disease stage, weight, or sex of a subject being treated, timing of the administration, manner of the administration, age of the subject, and the like, all of which can be determined by those of skill in the art.

**[00154]** The terms “treating” and “treatment”, and the like, as used herein, are intended to mean obtaining a desirable physiological, pharmacological, or biological effect, and includes prophylactic and therapeutic treatment. The effect may result in the inhibition, attenuation, amelioration, or reversal of a sign, symptom or cause of a disorder, or disease, attributable to the disorder, or disease, which includes mental and psychiatric diseases and disorders. Clinical evidence of the prevention or treatment may vary with the disorder, or disease, the subject and the selected treatment.

**[00155]** The term “pharmaceutically acceptable”, as used herein, refers to materials, including excipients, carriers, diluents, or auxiliary agents, that are compatible with other materials in a pharmaceutical or recreational drug formulation and within the scope of reasonable medical judgement suitable for use in contact with a subject without excessive toxicity, allergic response, irritation, or other adverse response commensurate with a reasonable risk/benefit ratio.

**[00156]** The term “psilocybin biosynthetic enzyme complement”, as used herein, refers to one or more polypeptides which alone or together are capable of

facilitating the chemical conversion of a psilocybin precursor compound and form another psilocybin precursor compound or psilocybin or a halogenated form thereof. A psilocybin biosynthetic enzyme complement can include, for example, PsiD, PsiH, PsiK, PsiM, Psi-ncAAAD and TrpB.

5 **[00157]** The term “PsiD” as used herein, refers to any and all enzymes comprising a sequence of amino acid residues which is (i) substantially identical to the amino acid sequences constituting any PsiD polypeptide set forth herein, including, for example, SEQ.ID NO: 2, SEQ.ID NO: 37, SEQ.ID NO: 39, SEQ.ID NO: 41, SEQ.ID NO: 43, SEQ.ID NO: 45, SEQ.ID NO: 47, SEQ.ID NO: 49, SEQ.ID  
10 NO: 51, SEQ.ID NO: 53, SEQ.ID NO: 55, and SEQ.ID NO: 57 or (ii) encoded by a nucleic acid sequence capable of hybridizing under at least moderately stringent conditions to any nucleic acid sequence encoding any PsiD set forth herein, but for the use of synonymous codons.

**[00158]** The term “PsiH” as used herein, refers to any and all enzymes  
15 comprising a sequence of amino acid residues which is (i) substantially identical to the amino acid sequences constituting any PsiH polypeptide set forth herein, including, for example, SEQ.ID NO: 4, or (ii) encoded by a nucleic acid sequence capable of hybridizing under at least moderately stringent conditions to any nucleic acid sequence encoding any PsiH set forth herein, but for the use of synonymous  
20 codons.

**[00159]** The term “PsiK” as used herein, refers to any and all enzymes comprising a sequence of amino acid residues which is (i) substantially identical to the amino acid sequences constituting any PsiK polypeptide set forth herein, including, for example, SEQ.ID NO: 6, or (ii) encoded by a nucleic acid sequence  
25 capable of hybridizing under at least moderately stringent conditions to any nucleic acid sequence encoding any PsiK set forth herein, but for the use of synonymous codons.

**[00160]** The term “PsiM” as used herein, refers to any and all enzymes comprising a sequence of amino acid residues which is (i) substantially identical  
30 to the amino acid sequences constituting any PsiM polypeptide set forth herein, including, for example, SEQ.ID NO: 8, and SEQ.ID NO: 59 or (ii) encoded by a nucleic acid sequence capable of hybridizing under at least moderately stringent conditions to any nucleic acid sequence encoding any PsiM set forth herein, but for the use of synonymous codons.

**[00161]** The term “Psi-ncAAAD” as used herein, refers to any and all enzymes comprising a sequence of amino acid residues which is (i) substantially identical to the amino acid sequences constituting any Psi-ncAAAD polypeptide set forth herein, including, for example, SEQ.ID NO: 10, or (ii) encoded by a nucleic acid sequence capable of hybridizing under at least moderately stringent conditions to any nucleic acid sequence encoding any Psi-ncAAAD set forth herein, but for the use of synonymous codons.

**[00162]** The term “TrpB” as used herein, refers to any and all enzymes comprising a sequence of amino acid residues which is (i) substantially identical to the amino acid sequences constituting any TrpB polypeptide set forth herein, including, for example, SEQ.ID NO: 12, SEQ.ID NO: 65, SEQ.ID NO: 67, SEQ.ID NO: 69, SEQ.ID NO: 71, and SEQ.ID NO: 73, or (ii) encoded by a nucleic acid sequence capable of hybridizing under at least moderately stringent conditions to any nucleic acid sequence encoding any TrpB set forth herein, but for the use of synonymous codons.

**[00163]** The term “halogenase”, as used herein, refers to any and all enzymes comprising a sequence of amino acid residues which is (i) substantially identical to the amino acid sequences constituting any halogenase polypeptide set forth herein, including, for example, SEQ.ID NO: 14, or (ii) encoded by a nucleic acid sequence capable of hybridizing under at least moderately stringent conditions to any nucleic acid sequence encoding any halogenase set forth herein, but for the use of synonymous codons.

**[00164]** The term “acetyl transferase”, as used herein, refers to any and all enzymes comprising a sequence of amino acid residues which is (i) substantially identical to the amino acid sequences constituting any acetyl transferase polypeptide set forth herein, including, for example, SEQ.ID NO: 63, or (ii) encoded by a nucleic acid sequence capable of hybridizing under at least moderately stringent conditions to any nucleic acid sequence encoding any acetyl transferase set forth herein, but for the use of synonymous codons.

**[00165]** The terms “nucleic acid sequence encoding PsiD”, and “nucleic acid sequence encoding a PsiD polypeptide”, as may be used interchangeably herein, refer to any and all nucleic acid sequences encoding a PsiD polypeptide, including, for example, SEQ.ID NO: 1, SEQ.ID NO: 36, SEQ.ID NO: 38, SEQ.ID NO: 40, SEQ.ID NO: 42, SEQ.ID NO: 44, SEQ.ID NO: 46, SEQ.ID NO: 48, SEQ.ID NO:

50, SEQ.ID NO: 52, SEQ.ID NO: 54, and SEQ.ID NO: 56. Nucleic acid sequences encoding a PsiD polypeptide further include any and all nucleic acid sequences which (i) encode polypeptides that are substantially identical to the PsiD polypeptide sequences set forth herein; or (ii) hybridize to any PsiD nucleic acid sequences set forth herein under at least moderately stringent hybridization conditions or which would hybridize thereto under at least moderately stringent conditions but for the use of synonymous codons.

**[00166]** The terms “nucleic acid sequence encoding PsiH”, and “nucleic acid sequence encoding a PsiH polypeptide”, as may be used interchangeably herein, refer to any and all nucleic acid sequences encoding a PsiH polypeptide, including, for example, SEQ.ID NO: 3. Nucleic acid sequences encoding a PsiH polypeptide further include any and all nucleic acid sequences which (i) encode polypeptides that are substantially identical to the PsiH polypeptide sequences set forth herein; or (ii) hybridize to any PsiH nucleic acid sequences set forth herein under at least moderately stringent hybridization conditions or which would hybridize thereto under at least moderately stringent conditions but for the use of synonymous codons.

**[00167]** The terms “nucleic acid sequence encoding PsiK”, and “nucleic acid sequence encoding a PsiK polypeptide”, as may be used interchangeably herein, refer to any and all nucleic acid sequences encoding a PsiK polypeptide, including, for example, SEQ.ID NO: 5. Nucleic acid sequences encoding a PsiK polypeptide further include any and all nucleic acid sequences which (i) encode polypeptides that are substantially identical to the PsiK polypeptide sequences set forth herein; or (ii) hybridize to any PsiK nucleic acid sequences set forth herein under at least moderately stringent hybridization conditions or which would hybridize thereto under at least moderately stringent conditions but for the use of synonymous codons.

**[00168]** The terms “nucleic acid sequence encoding PsiM”, and “nucleic acid sequence encoding a PsiM polypeptide”, as may be used interchangeably herein, refer to any and all nucleic acid sequences encoding a PsiD polypeptide, including, for example, SEQ.ID NO: 7 SEQ.ID NO: 58 and SEQ.ID NO: 61. Nucleic acid sequences encoding a PsiM polypeptide further include any and all nucleic acid sequences which (i) encode polypeptides that are substantially identical to the PsiM polypeptide sequences set forth herein; or (ii) hybridize to any PsiM nucleic

acid sequences set forth herein under at least moderately stringent hybridization conditions or which would hybridize thereto under at least moderately stringent conditions but for the use of synonymous codons.

**[00169]** The terms “nucleic acid sequence encoding Psi-ncAAAD”, and  
5 “nucleic acid sequence encoding a Psi-ncAAAD polypeptide”, as may be used interchangeably herein, refer to any and all nucleic acid sequences encoding a PsiD polypeptide, including, for example, SEQ.ID NO: 9. Nucleic acid sequences encoding a Psi-ncAAAD polypeptide further include any and all nucleic acid sequences which (i) encode polypeptides that are substantially identical to the Psi-ncAAAD polypeptide sequences set forth herein; or (ii) hybridize to any Psi-ncAAAD nucleic acid sequences set forth herein under at least moderately stringent hybridization conditions or which would hybridize thereto under at least moderately stringent conditions but for the use of synonymous codons.

**[00170]** The terms “nucleic acid sequence encoding TrpB”, and “nucleic acid  
15 sequence encoding a TrpB polypeptide”, as may be used interchangeably herein, refer to any and all nucleic acid sequences encoding a TrpB polypeptide, including, for example, SEQ.ID NO: 11, SEQ.ID NO: 64, SEQ.ID NO: 66, SEQ.ID NO: 68, SEQ.ID NO: 70, and SEQ.ID NO: 72. Nucleic acid sequences encoding a TrpB polypeptide further include any and all nucleic acid sequences which (i) encode  
20 polypeptides that are substantially identical to the TrpB polypeptide sequences set forth herein; or (ii) hybridize to any TrpB nucleic acid sequences set forth herein under at least moderately stringent hybridization conditions or which would hybridize thereto under at least moderately stringent conditions but for the use of synonymous codons.

**[00171]** The terms “nucleic acid sequence encoding a halogenase”, and  
25 “nucleic acid sequence encoding a halogenase polypeptide”, as may be used interchangeably herein, refer to any and all nucleic acid sequences encoding a halogenase polypeptide, including, for example, SEQ.ID NO: 13. Nucleic acid sequences encoding a halogenase polypeptide further include any and all nucleic acid sequences which (i) encode polypeptides that are substantially identical to the halogenase polypeptide sequences set forth herein; or (ii) hybridize to any halogenase nucleic acid sequences set forth herein under at least moderately stringent hybridization conditions or which would hybridize thereto under at least moderately stringent conditions but for the use of synonymous codons.

**[00172]** The terms “nucleic acid sequence encoding an acetyl transferase”, and “nucleic acid sequence encoding an acetyl transferase polypeptide”, as may be used interchangeably herein, refer to any and all nucleic acid sequences encoding an acetyl transferase polypeptide, including, for example, SEQ.ID NO: 5 62. Nucleic acid sequences encoding an acetyl transferase polypeptide further include any and all nucleic acid sequences which (i) encode polypeptides that are substantially identical to the acetyl transferase polypeptide sequences set forth herein; or (ii) hybridize to any acetyl transferase nucleic acid sequences set forth herein under at least moderately stringent hybridization conditions or which would 10 hybridize thereto under at least moderately stringent conditions but for the use of synonymous codons.

**[00173]** The terms “nucleic acid”, or “nucleic acid sequence”, as used herein, refer to a sequence of nucleoside or nucleotide monomers, consisting of naturally occurring bases, sugars and intersugar (backbone) linkages. The term also 15 includes modified or substituted sequences comprising non-naturally occurring monomers or portions thereof. The nucleic acids of the present disclosure may be deoxyribonucleic acids (DNA) or ribonucleic acids (RNA) and may include naturally occurring bases including adenine, guanine, cytosine, thymidine and uracil. The nucleic acids may also contain modified bases. Examples of such 20 modified bases include aza and deaza adenine, guanine, cytosine, thymidine and uracil, and xanthine and hypoxanthine. A sequence of nucleotide or nucleoside monomers may be referred to as a polynucleotide sequence, nucleic acid sequence, a nucleotide sequence or a nucleoside sequence.

**[00174]** The term “polypeptide”, as used herein in conjunction with a 25 reference SEQ.ID NO, refers to any and all polypeptides comprising a sequence of amino acid residues which is (i) substantially identical to the amino acid sequence constituting the polypeptide having such reference SEQ.ID NO, or (ii) encoded by a nucleic acid sequence capable of hybridizing under at least moderately stringent conditions to any nucleic acid sequence encoding the 30 polypeptide having such reference SEQ.ID NO, but for the use of synonymous codons. A sequence of amino acid residues may be referred to as an amino acid sequence, or polypeptide sequence.

**[00175]** The term “nucleic acid sequence encoding a polypeptide”, as used herein in conjunction with a reference SEQ.ID NO, refers to any and all nucleic



acid sequences encoding a polypeptide having such reference SEQ.ID NO. Nucleic acid sequences encoding a polypeptide, in conjunction with a reference SEQ.ID NO, further include any and all nucleic acid sequences which (i) encode polypeptides that are substantially identical to the polypeptide having such reference SEQ.ID NO; or (ii) hybridize to any nucleic acid sequences encoding polypeptides having such reference SEQ.ID NO under at least moderately stringent hybridization conditions or which would hybridize thereto under at least moderately stringent conditions but for the use of synonymous codons.

**[00176]** By the term “substantially identical” it is meant that two amino acid

sequences preferably are at least 70% identical, and more preferably are at least 85% identical and most preferably at least 95% identical, for example 96%, 97%, 98% or 99% identical. In order to determine the percentage of identity between two amino acid sequences the amino acid sequences of such two sequences are aligned, using for example the alignment method of Needleman and Wunsch (J. Mol. Biol., 1970, 48: 443), as revised by Smith and Waterman (Adv. Appl. Math., 1981, 2: 482) so that the highest order match is obtained between the two sequences and the number of identical amino acids is determined between the two sequences. Methods to calculate the percentage identity between two amino acid sequences are generally art recognized and include, for example, those described by Carillo and Lipton (SIAM J. Applied Math., 1988, 48:1073) and those described in Computational Molecular Biology, Lesk, e.d. Oxford University Press, New York, 1988, Biocomputing: Informatics and Genomics Projects. Generally, computer programs will be employed for such calculations. Computer programs that may be used in this regard include, but are not limited to, GCG (Devereux *et al.*, Nucleic Acids Res., 1984, 12: 387) BLASTP, BLASTN and FASTA (Altschul *et al.*, J. Mol. Biol., 1990:215:403). A particularly preferred method for determining the percentage identity between two polypeptides involves the Clustal W algorithm (Thompson, J D, Higgins, D G and Gibson T J, 1994, Nucleic Acid Res 22(22): 4673-4680 together with the BLOSUM 62 scoring matrix (Henikoff S & Henikoff, J G, 1992, Proc. Natl. Acad. Sci. USA 89: 10915-10919 using a gap opening penalty of 10 and a gap extension penalty of 0.1, so that the highest order match obtained between two sequences wherein at least 50% of the total length of one of the two sequences is involved in the alignment.

**[00177]** By “at least moderately stringent hybridization conditions” it is meant that conditions are selected which promote selective hybridization between two complementary nucleic acid molecules in solution. Hybridization may occur to all or a portion of a nucleic acid sequence molecule. The hybridizing portion is typically at least 15 (e.g. 20, 25, 30, 40 or 50) nucleotides in length. Those skilled in the art will recognize that the stability of a nucleic acid duplex, or hybrids, is determined by the  $T_m$ , which in sodium containing buffers is a function of the sodium ion concentration and temperature ( $T_m = 81.5^\circ \text{C} - 16.6 (\text{Log} 10 [\text{Na}^+]) + 0.41 (\% (\text{G} + \text{C}) - 600 / l)$ , or similar equation). Accordingly, the parameters in the wash conditions that determine hybrid stability are sodium ion concentration and temperature. In order to identify molecules that are similar, but not identical, to a known nucleic acid molecule a 1% mismatch may be assumed to result in about a  $1^\circ \text{C}$ . decrease in  $T_m$ , for example if nucleic acid molecules are sought that have a >95% identity, the final wash temperature will be reduced by about  $5^\circ \text{C}$ . Based on these considerations those skilled in the art will be able to readily select appropriate hybridization conditions. In preferred embodiments, stringent hybridization conditions are selected. By way of example the following conditions may be employed to achieve stringent hybridization: hybridization at  $5\times$  sodium chloride/sodium citrate (SSC)/ $5\times$ Denhardt's solution/1.0% SDS at  $T_m$  (based on the above equation)  $-5^\circ \text{C}$ , followed by a wash of  $0.2\times$ SSC/0.1% SDS at  $60^\circ \text{C}$ . Moderately stringent hybridization conditions include a washing step in  $3\times$ SSC at  $42^\circ \text{C}$ . It is understood however that equivalent stringencies may be achieved using alternative buffers, salts and temperatures. Additional guidance regarding hybridization conditions may be found in: Current Protocols in Molecular Biology, John Wiley & Sons, N.Y., 1989, 6.3.1.-6.3.6 and in: Sambrook *et al.*, Molecular Cloning, a Laboratory Manual, Cold Spring Harbor Laboratory Press, 1989, Vol. 3.

**[00178]** The term “functional variant”, as used herein in reference to polynucleotides or polypeptides, refers to polynucleotides or polypeptides capable of performing the same function as a noted reference polynucleotide or polypeptide. Thus, for example, a functional variant of the polypeptide set forth in SEQ.ID NO: 2, refers to a polypeptide capable of performing the same function as the polypeptide set forth in SEQ.ID NO: 2. Functional variants include modified a polypeptide wherein, relative to a noted reference polypeptide, the modification includes a substitution, deletion or addition of one or more amino acids. In some

embodiments, substitutions are those that result in a replacement of one amino acid with an amino acid having similar characteristics. Such substitutions include, without limitation (i) glutamic acid and aspartic acid; (ii) alanine, serine, and threonine; (iii) isoleucine, leucine and valine, (iv) asparagine and glutamine, and  
5 (v) tryptophan, tyrosine and phenylalanine. Functional variants further include polypeptides having retained or exhibiting an enhanced psilocybin biosynthetic bioactivity.

**[00179]** The term “chimeric”, as used herein in the context of nucleic acids, refers to at least two linked nucleic acids which are not naturally linked. Chimeric  
10 nucleic acids include linked nucleic acids of different natural origins. For example, a nucleic acid constituting a microbial promoter linked to a nucleic acid encoding a plant polypeptide is considered chimeric. Chimeric nucleic acids also may comprise nucleic acids of the same natural origin, provided they are not naturally linked. For example a nucleic acid constituting a promoter obtained from a  
15 particular cell-type may be linked to a nucleic acid encoding a polypeptide obtained from that same cell-type, but not normally linked to the nucleic acid constituting the promoter. Chimeric nucleic acids also include nucleic acids comprising any naturally occurring nucleic acids linked to any non-naturally occurring nucleic acids.

**[00180]** The terms “substantially pure” and “isolated”, as used herein, as may be used interchangeably herein describe a compound, *e.g.*, a secondary metabolite, psilocybin or a psilocybin derivative, polynucleotide or a polypeptide, which has been separated from components that naturally accompany it. Typically, a compound is substantially pure when at least 60%, more preferably at least 75%,  
20 more preferably at least 90%, 95%, 96%, 97%, or 98%, and most preferably at least 99% of the total material (by volume, by wet or dry weight, or by mole percent or mole fraction) in a sample is the compound of interest. Purity can be measured by any appropriate method, *e.g.*, in the case of polypeptides, by chromatography, gel electrophoresis or HPLC analysis.

**[00181]** The term “recovered” as used herein in association with an enzyme, protein, a secondary metabolite or a chemical compound, refers to a more or less pure form of the enzyme, protein, secondary metabolite, or chemical compound.

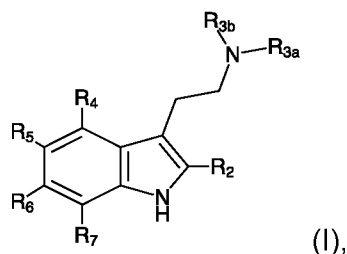
**General Implementation**

**[00182]** As hereinbefore mentioned, the present disclosure relates to psilocybin derivatives. In particular, the present disclosure provides novel halogenated psilocybin derivatives. In general, the herein provided compositions exhibit functional properties which deviate from the functional properties of psilocybin. Thus, for example, the psilocybin derivatives, can exhibit pharmacological properties which deviate from psilocybin. The halogenated psilocybin derivatives in this respect are useful in the formulation of pharmaceutical and recreation drug formulations. Furthermore, the halogenated psilocybin derivatives of the present disclosure may be used as a feedstock material for deriving further psilocybin derivatives. In one embodiment, the halogenated psilocybin derivatives of the present disclosure can conveniently be biosynthetically produced. The practice of this method avoids the extraction of psilocybin from mushrooms and the performance of subsequent chemical reactions to achieve halogenation. Furthermore, the growth of mushrooms can be avoided thus limiting the dependence on climate and weather, and potential legal and social challenges associated with the cultivation of mushrooms containing psychoactive compounds. The method can efficiently yield substantial quantities of halogenated psilocybin derivatives.

**[00183]** In what follows selected embodiments are described with reference to the drawings.

**[00184]** Initially various halogenated psilocybin derivatives will be described. Thereafter methods example methods of using and making the halogenated psilocybin derivatives will be described

**[00185]** In at least one aspect, and in at least one example embodiment, the present disclosure provides a chemical compound or salts thereof having the formula (I):



wherein, at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> is a halogen atom, wherein each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom, wherein R<sub>4</sub> when it is not halogenated is a phosphate group or a hydrogen atom, and wherein R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom, an alkyl group, an aryl group or an acyl group.

- 5 **[00186]** It is noted that in reference to the indole prototype structure shown in **FIG. 2** carbon atoms C<sub>2</sub>, C<sub>4</sub>, C<sub>5</sub>, C<sub>6</sub> and C<sub>7</sub> are bonded to R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub>, respectively. Furthermore, R<sub>3A</sub> and R<sub>3B</sub> reference chemical groups extending from the ethyl-amino group extending in turn from carbon atom C<sub>3</sub> of the prototype indole structure.
- 10 **[00187]** Referring to the chemical compound having formula (I), In one embodiment, one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a halogen atom. Thus, in one embodiment, R<sub>2</sub> can be a bromine, chlorine, fluorine or iodine atom, each of R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub>, when it is not a halogen atom, can be a phosphate group, a hydrogen atom or an
- 15 alkyl group or O-alkyl group (see: the example halogenated psilocybin derivatives shown in **FIG. 4A** (R<sub>2</sub> is chlorine; R<sub>4</sub> is a phosphate group; R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> are each a hydrogen atom; R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group); **FIG. 4B** (R<sub>2</sub> is bromine; R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> are each a hydrogen atom; R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group); **FIG. 4K** (R<sub>2</sub> is chlorine; R<sub>4</sub> is a methyl group; R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> are each a
- 20 hydrogen atom; R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group); and **FIG. 4O** (R<sub>2</sub> is chlorine; R<sub>4</sub> is an O-methyl group; R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> are each a hydrogen atom; R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).
- 25 **[00188]** Continuing to refer to the chemical compound having formula (I), in one embodiment, R<sub>4</sub> can be a bromine, chlorine, fluorine or iodine atom, and each of R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivatives shown in **FIG. 4C** (R<sub>4</sub> is iodine; R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> are each a hydrogen atom; R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen
- 30 atom, an alkyl group, an aryl group, or an acyl group); **FIG. 4D** (R<sub>4</sub> is fluorine; R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> are each a hydrogen atom; R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group))
- [00189]** Continuing to refer to the chemical compound having formula (I), in one embodiment, R<sub>5</sub> can be a bromine, chlorine, fluorine or iodine atom, and each

of R<sub>2</sub>, R<sub>6</sub> and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivatives shown in **FIG. 4E** (R<sub>5</sub> is fluorine; R<sub>4</sub> is a phosphate group; R<sub>2</sub>, R<sub>6</sub> and R<sub>7</sub> are each a hydrogen atom; R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group); **FIG. 4F** (R<sub>5</sub> is chlorine; R<sub>2</sub>, R<sub>4</sub>, R<sub>6</sub> and R<sub>7</sub> are each a hydrogen atom; R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group); **FIG. 4L** (R<sub>5</sub> is bromine; R<sub>4</sub> is an ethyl group; R<sub>2</sub>, R<sub>6</sub> and R<sub>7</sub> are each a hydrogen atom; R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group); and **FIG. 4P** (R<sub>5</sub> is bromine; R<sub>2</sub>, R<sub>6</sub> and R<sub>7</sub> are each a hydrogen atom; R<sub>4</sub> is an O-ethyl group; R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).

**[00190]** Continuing to refer to the chemical compound having formula (I), in one embodiment, R<sub>6</sub> can be a bromine, chlorine, fluorine or iodine atom, and each of R<sub>2</sub>, R<sub>5</sub> and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivatives shown in **FIG. 4G** (R<sub>6</sub> is bromine; R<sub>4</sub> is a phosphate group; R<sub>2</sub>, R<sub>5</sub> and R<sub>7</sub> are each a hydrogen atom; R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group); **FIG. 4H** (R<sub>6</sub> is chlorine; R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub> and R<sub>7</sub> are each a hydrogen atom; R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group); **FIG. 4M** (R<sub>6</sub> is fluorine; R<sub>4</sub> is a methyl group; R<sub>2</sub>, R<sub>5</sub> and R<sub>7</sub> are each a hydrogen atom; R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group); and **FIG. 4Q** (R<sub>6</sub> is fluorine; R<sub>2</sub>, R<sub>5</sub> and R<sub>7</sub> are each a hydrogen atom; R<sub>4</sub> is an O-methyl group; R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).

**[00191]** Continuing to refer to the chemical compound having formula (I), in one embodiment, R<sub>7</sub> can be a bromine, chlorine, fluorine or iodine atom, and each of R<sub>2</sub>, R<sub>5</sub> and R<sub>6</sub> can be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivatives shown in **FIG. 4I** (R<sub>7</sub> is fluorine; R<sub>4</sub> is a phosphate group; R<sub>2</sub>, R<sub>5</sub> and R<sub>6</sub> are each a hydrogen atom; R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group); **FIG. 4J** (R<sub>7</sub> is iodine; R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub> and R<sub>6</sub> are each a hydrogen atom; R<sub>3a</sub> and R<sub>3b</sub> are a

hydrogen atom, an alkyl group, an aryl group, or an acyl group); **FIG. 4N** ( $R_7$  is iodine;  $R_4$  is a propyl group;  $R_2$ ,  $R_5$  and  $R_6$  are each a hydrogen atom  $R_{3a}$  and  $R_{3b}$  are a hydrogen atom, an alkyl group, an aryl group, or an acyl group); and **FIG. 4R** ( $R_7$  is iodine;  $R_2$ ,  $R_5$  and  $R_6$  are each a hydrogen atom;  $R_4$  is an O-propyl group;  $R_{3a}$  and  $R_{3b}$  are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).

5 **[00192]** Still, continuing to refer to the chemical compound having formula (I), in one embodiment, two of  $R_2$ ,  $R_4$ ,  $R_5$ ,  $R_6$  and  $R_7$  can be a halogen atom, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom, and wherein each non-halogenated  $R_2$ ,  $R_5$ ,  $R_6$  and  $R_7$  is a  
10 hydrogen atom or an alkyl group or O-alkyl group, and wherein  $R_4$ , when it is not a halogen atom, is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.

**[00193]** Thus, continuing to refer to the chemical compound having formula (I), in one embodiment,  $R_2$  and  $R_4$  can be halogen atoms, wherein the halogen  
15 atoms are independently selected from a bromine, chlorine fluorine or iodine atom, and  $R_5$ ,  $R_6$  and  $R_7$  can be a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivative shown in **FIG. 5A** ( $R_2$  is a chlorine atom;  $R_4$  is a chlorine atom;  $R_5$ ,  $R_6$  and  $R_7$  are each a hydrogen atom; and  $R_{3a}$  and  $R_{3b}$  are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).

20 **[00194]** Continuing to refer to the chemical compound having formula (I), in one embodiment,  $R_2$  and  $R_5$  can be halogen atoms, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom, and  $R_6$  and  $R_7$  can be a hydrogen atom or an alkyl group or O-alkyl group and  $R_4$  can be a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group (see:  
25 the example halogenated psilocybin derivative shown in **FIG. 5B** ( $R_2$  is a bromine atom;  $R_5$  is a fluorine atom;  $R_4$ , and  $R_6$  are each a hydrogen atom and  $R_7$  is a methyl group; and  $R_{3a}$  and  $R_{3b}$  are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).

**[00195]** Continuing to refer to the chemical compound having formula (I), in  
30 one embodiment,  $R_2$  and  $R_6$  can be halogen atoms, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom, and  $R_5$  and  $R_7$  can be a hydrogen atom or an alkyl group or O-alkyl group, and  $R_4$  can be a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivative shown in **FIG. 5C** ( $R_2$  is a chlorine

atom; R<sub>6</sub> is a bromine atom; R<sub>5</sub> and R<sub>7</sub> are each a hydrogen atom; R<sub>4</sub> is an O-methyl group; and R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).

**[00196]** Continuing to refer to the chemical compound having formula (I), in one embodiment, R<sub>2</sub> and R<sub>7</sub> can be halogen atoms, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom, and R<sub>5</sub> and R<sub>6</sub> can be a hydrogen atom or an alkyl group or O-alkyl group and R<sub>4</sub> can be a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivative shown in **FIG. 5D** (R<sub>2</sub> is a fluorine atom; R<sub>7</sub> is an iodine atom; R<sub>4</sub> is a phosphate, R<sub>5</sub> and R<sub>6</sub> are each a hydrogen atom; and R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).

**[00197]** Continuing to refer to the chemical compound having formula (I), in one embodiment, R<sub>4</sub> and R<sub>5</sub> can be halogen atoms, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom, and R<sub>2</sub>, R<sub>6</sub> and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivative shown in **FIG. 5E** (R<sub>4</sub> and R<sub>5</sub> are each bromine atoms; R<sub>2</sub>, R<sub>6</sub> and R<sub>7</sub> are each a hydrogen atom; and R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).

**[00198]** Continuing to refer to the chemical compound having formula (I), in one embodiment, R<sub>4</sub> and R<sub>6</sub> can be halogen atoms, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom, and R<sub>2</sub>, R<sub>5</sub> and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivative shown in **FIG. 5F** (R<sub>4</sub> and R<sub>6</sub> are each chlorine atoms; R<sub>2</sub>, R<sub>5</sub> and R<sub>7</sub> are each a hydrogen atom; and R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).

**[00199]** Continuing to refer to the chemical compound having formula (I), in one embodiment, R<sub>4</sub> and R<sub>7</sub> can be halogen atoms, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom, and R<sub>2</sub>, R<sub>5</sub> and R<sub>6</sub> can be a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivative shown in **FIG. 5G** (R<sub>4</sub> is a chlorine atom; R<sub>7</sub> is a fluorine atom; R<sub>2</sub>, R<sub>5</sub> and R<sub>6</sub> are each a hydrogen atom; and R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).



- [00200]** Continuing to refer to the chemical compound having formula (I), in one embodiment, R<sub>5</sub> and R<sub>6</sub> can be halogen atoms, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom, and R<sub>2</sub> and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group and R<sub>4</sub> can be a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivative shown in **FIG. 5H** (R<sub>5</sub> is a fluorine atom; R<sub>6</sub> is a bromine atom; R<sub>4</sub> is a phosphate group, R<sub>2</sub> and R<sub>7</sub> are each a hydrogen atom; and R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).
- [00201]** Continuing to refer to the chemical compound having formula (I), in one embodiment, R<sub>5</sub> and R<sub>7</sub> can be halogen atoms, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom, and R<sub>2</sub> and R<sub>6</sub> can be a hydrogen atom or an alkyl group or O-alkyl group and R<sub>4</sub> can be a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivative shown in **FIG. 5I** (R<sub>5</sub> and R<sub>7</sub> are each fluorine atoms; R<sub>4</sub> is a phosphate group, R<sub>2</sub> and R<sub>6</sub> are each a hydrogen atom; and R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).
- [00202]** Continuing to refer to the chemical compound having formula (I), in one embodiment, R<sub>6</sub> and R<sub>7</sub> can be halogen atoms, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom, and R<sub>2</sub> and R<sub>5</sub> can be a hydrogen atom or an alkyl group or O-alkyl group and R<sub>4</sub> can be a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivative shown in **FIG. 5J** (R<sub>6</sub> and R<sub>7</sub> are each iodine atoms; R<sub>2</sub>, R<sub>4</sub> and R<sub>5</sub> are each a hydrogen atom; and R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).
- [00203]** In one embodiment, two of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a halogen atom, and the halogen atoms are identical, the remaining of R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, and R<sub>7</sub> being a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> when it is not a halogen atom, being a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group (see: e.g. the halogenated psilocybin derivatives shown in **FIGS. 5A, 5E, 5F, 5I and 5J**).
- [00204]** In one embodiment, two of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a halogen atom, and the halogen atoms are non-identical, the remaining of R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, and

R<sub>7</sub> being a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> when it is not a halogen atom, being a phosphate, a hydrogen atom or an alkyl group or O-alkyl group (see: e.g. the halogenated psilocybin derivatives shown in **FIGS. 5B, 5C, 5D, 5G and 5H**).

5 **[00205]** Still continuing to refer to the chemical compound having formula (I), In one further embodiment, three of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a halogen atom, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom, and wherein the non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group, and wherein R<sub>4</sub>, when it is not a halogen atom, is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.

**[00206]** Thus, continuing to refer to the chemical compound having formula (I), in one embodiment R<sub>2</sub>, R<sub>4</sub>, and R<sub>5</sub> can be halogen atom, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom, and R<sub>6</sub> and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivative shown in **FIG. 6A** (R<sub>2</sub>, R<sub>4</sub> and R<sub>5</sub> are each chlorine atoms; R<sub>6</sub> and R<sub>7</sub> are each a hydrogen atom; and R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).

**[00207]** Continuing to refer to the chemical compound having formula (I), in one embodiment R<sub>2</sub>, R<sub>5</sub>, and R<sub>6</sub> can be halogen atom, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivative shown in **FIG. 6B** (R<sub>2</sub> is a bromine atom, R<sub>5</sub> and R<sub>6</sub> are each fluorine atoms; R<sub>4</sub> and R<sub>7</sub> are each a hydrogen atom; and R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).

**[00208]** Continuing to refer to the chemical compound having formula (I), in one embodiment R<sub>2</sub>, R<sub>5</sub>, and R<sub>7</sub> can be halogen atom, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom and R<sub>6</sub> can be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivative shown in **FIG. 6C** (R<sub>2</sub> is a chlorine atom, R<sub>5</sub> is a bromine atom, and R<sub>7</sub> is a fluorine atoms; R<sub>4</sub> and R<sub>6</sub> are each a

hydrogen atom; and R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).

**[00209]** Continuing to refer to the chemical compound having formula (I), in one embodiment R<sub>4</sub>, R<sub>5</sub>, and R<sub>6</sub> can be halogen atom, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom, and R<sub>2</sub> and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivative shown in **FIG. 6D** (R<sub>4</sub> and R<sub>5</sub> are each a bromine atom, R<sub>6</sub> is a chlorine atom; R<sub>2</sub> and R<sub>7</sub> are each a hydrogen atom; and R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).

**[00210]** Continuing to refer to the chemical compound having formula (I), in one embodiment R<sub>4</sub>, R<sub>5</sub>, and R<sub>7</sub> can be halogen atom, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom, and R<sub>2</sub> and R<sub>6</sub> can be a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivative shown in **FIG. 6E** (R<sub>4</sub>, R<sub>5</sub> and R<sub>7</sub> are each a bromine atom; R<sub>2</sub> and R<sub>6</sub> are each a hydrogen atom; and R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).

**[00211]** Continuing to refer to the chemical compound having formula (I), in one embodiment R<sub>5</sub>, R<sub>6</sub>, and R<sub>7</sub> can be halogen atom, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom and R<sub>2</sub> can be a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> can be a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivative shown in **FIG. 6F** (R<sub>5</sub> is a fluorine atom, R<sub>6</sub> is a chlorine atom, and R<sub>7</sub> is an iodine atom; R<sub>4</sub> is a phosphate group; R<sub>2</sub> is a hydrogen atom; and R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).

**[00212]** In one embodiment, three of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a halogen atom, wherein at least two halogen atoms are identical, the remaining R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, and R<sub>7</sub> being a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> when it is not a halogen atom, being a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group (see: e.g. the halogenated psilocybin derivatives shown in **FIGS. 6A, 6B and 6E**).

**[00213]** In one embodiment, three of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a halogen atom, and at least two halogen atoms are non-identical, the remaining R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, and R<sub>7</sub> being a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> when it

is not a halogen atom, being a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group (see: e.g. the halogenated psilocybin derivatives shown in **FIGS. 6B, 6C, 6D and 6F**).

5 **[00214]** In one embodiment, three of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a halogen atom, and all three halogen atoms are identical, the remaining R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> being a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> when it is not a halogen atom, being a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group (see: e.g. the halogenated psilocybin derivatives shown in **FIGS. 6A and 6E**).

10 **[00215]** Still continuing to refer to the chemical compound having formula (I), in one embodiment, four of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a halogen atom, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom, and wherein the non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group, and wherein R<sub>4</sub>, when it is not a halogen, is a  
15 phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.

**[00216]** Thus, continuing to refer to the chemical compound having formula (I), in one embodiment R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub> and R<sub>6</sub> can be halogen atom, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom, and R<sub>7</sub> can be a hydrogen atom or an alkyl group or O-alkyl group  
20 (see: the example halogenated psilocybin derivative shown in **FIG. 7A** (R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub> and R<sub>6</sub> are each chlorine atoms; R<sub>7</sub> is a hydrogen atom; and R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).

**[00217]** Continuing to refer to the chemical compound having formula (I), in one embodiment R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be halogen atom, wherein the halogen  
25 atoms are independently selected from a bromine, chlorine fluorine or iodine atom and R<sub>2</sub> can be a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivative shown in **FIG. 7B** (R<sub>4</sub> and R<sub>7</sub> are chlorine atoms, R<sub>5</sub> and R<sub>6</sub> are each fluorine atoms; R<sub>2</sub> is a hydrogen atom; and R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).

30 **[00218]** Continuing to refer to the chemical compound having formula (I), in one embodiment R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be halogen atom, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom and R<sub>4</sub> can be a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivative shown in **FIG. 7C** (R<sub>2</sub>

and are R<sub>6</sub> are each chlorine atoms, R<sub>5</sub> is a fluorine atom, R<sub>7</sub> is an iodine atoms, R<sub>4</sub> is a phosphate group; and R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).

**[00219]** Continuing to refer to the chemical compound having formula (I), in one embodiment R<sub>2</sub>, R<sub>4</sub>, R<sub>6</sub> and R<sub>7</sub> can be halogen atom, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom and R<sub>5</sub> can be a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivative shown in **FIG. 7D** (R<sub>2</sub>, R<sub>4</sub> and are R<sub>7</sub> are each bromine atoms, R<sub>6</sub> is a chlorine atom, R<sub>5</sub> is a hydrogen atom; and R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).

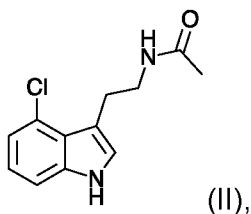
**[00220]** In one embodiment R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub> and R<sub>7</sub> can be halogen atom, wherein the halogen atoms are independently selected from a bromine, chlorine fluorine or iodine atom and R<sub>6</sub> can be a hydrogen atom or an alkyl group or O-alkyl group (see: the example halogenated psilocybin derivative shown in **FIG. 7E** (R<sub>4</sub>, R<sub>5</sub> and are R<sub>7</sub> are each bromine atoms, R<sub>2</sub> is a fluorine atom, R<sub>6</sub> is a hydrogen atom; and R<sub>3a</sub> and R<sub>3b</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group)).

**[00221]** In one embodiment, four of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a halogen atom, wherein at least two halogen atoms are identical, the remaining R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, and R<sub>7</sub> being a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> when it is not a halogen atom, being a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group (see: e.g. the halogenated psilocybin derivatives shown in **FIGS. 7A, 7B, 7C, 7D and 7E**).

**[00222]** In one embodiment, four of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a halogen atom, and at least two halogen atoms are non-identical, the remaining R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, and R<sub>7</sub> being a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> when it is not a halogen atom, being a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group (see: e.g. the halogenated psilocybin derivatives shown in **FIGS. 7B, 7C, 7D and 7E**).

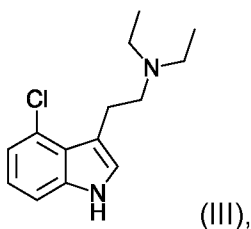
**[00223]** In one embodiment, four of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a halogen atom, and all four halogen atoms are identical, the remaining R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> being a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> when it is not a halogen atom, being a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group (see: e.g. the halogenated psilocybin derivative shown in **FIGS. 7A**).

- [00224]** Still continuing to refer to the chemical compound having formula (I), In one embodiment, all five of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a halogen atom, the halogen selected from bromine, chlorine, fluorine or iodine.
- [00225]** In one embodiment, all five of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a halogen atom, and at least two halogen atoms are identical.
- [00226]** In one embodiment, all five of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a halogen atom, and at least three halogen atoms are identical.
- [00227]** In one embodiment, all five of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a halogen atom, and at least four halogen atoms are identical.
- [00228]** In one embodiment, all five of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a halogen atom, and all five halogen atoms are identical.
- [00229]** In one embodiment, all five of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a halogen atom, and the chemical compound comprises at least two non-identical halogen atoms.
- [00230]** In one embodiment, all five of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a halogen atom, and the chemical compound comprises at least three non-identical halogen atoms.
- [00231]** In one embodiment, all five of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> can be a halogen atom, and four halogen atoms are non-identical.
- [00232]** Still continuing to refer to the chemical compound having formula (I), it is noted that, in a further aspect hereof, R<sub>3A</sub> and R<sub>3B</sub> can be a hydrogen atom, an alkyl group, an aryl group or an acyl group. Thus, for example, R<sub>3A</sub> and R<sub>3B</sub> can each be a hydrogen atom, or R<sub>3A</sub> and R<sub>3B</sub> can each be an alkyl group, such as a methyl group, ethyl group, propyl group, or longer chain alkyl group, or R<sub>3A</sub> and R<sub>3B</sub> can each be an aryl group, such as a phenyl group or a naphthyl group, or R<sub>3A</sub> and R<sub>3B</sub> can each be an acyl group, such as an acetyl group. Furthermore, one of R<sub>3A</sub> and R<sub>3B</sub> can be a hydrogen atom, and one of R<sub>3A</sub> and R<sub>3B</sub> can be an alkyl group, and aryl group, or an acyl group. Furthermore, R<sub>3A</sub> and R<sub>3B</sub> can be an aryl group and an alkyl group, an aryl group and an acyl group, or an alkyl group and an acyl group.
- [00233]** Furthermore, in one embodiment, a psilocybin derivative according to the present disclosure can be a chemical compound having the formula (II):



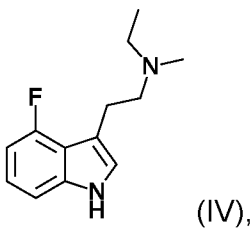
or, a chemical compound having the formula (II), wherein the chlorine is substituted for a bromine, fluorine or iodine.

**[00234]** Furthermore, in one embodiment, a psilocybin derivative according to the present disclosure can be a chemical compound having the formula (III):



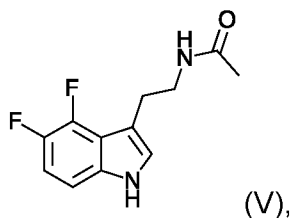
or, a chemical compound having the formula (III), wherein the chlorine is substituted for a bromine, fluorine or iodine.

**[00235]** Furthermore, in one embodiment, a psilocybin derivative according to the present disclosure can be a chemical compound having the formula (IV):



or, a chemical compound having the formula (IV), wherein the fluorine is substituted for a bromine, chlorine or iodine.

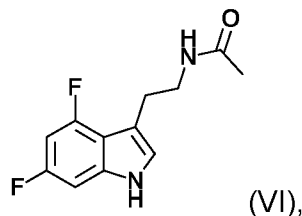
**[00236]** Furthermore, in one embodiment, a psilocybin derivative according to the present disclosure can be a chemical compound having the formula (V):



or, a chemical compound having the formula (V), wherein one or both of the fluorines are substituted for a bromine, chlorine or iodine.

**[00237]** Furthermore, in one embodiment, a psilocybin derivative according to the present disclosure can be a chemical compound having the formula (VI):

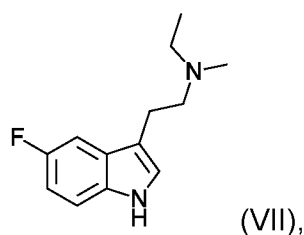
5



or, a chemical compound having the formula (VI), wherein one or both of the fluorines are substituted for a bromine, chlorine or iodine.

**[00238]** Furthermore, in one embodiment, a psilocybin derivative according to the present disclosure can be a chemical compound having the formula (VII):

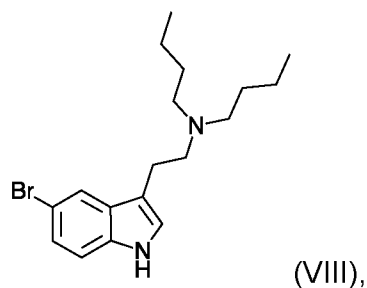
10



or, a chemical compound having the formula (VII), wherein the fluorine is substituted for a bromine, chlorine or iodine.

**[00239]** Furthermore, in one embodiment, a psilocybin derivative according to the present disclosure can be a chemical compound having the formula (VIII):

15

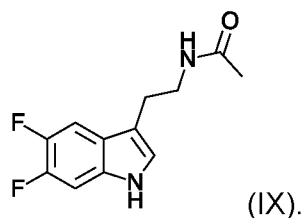


or, a chemical compound having the formula (VIII), wherein the bromine is substituted for a, chlorine, fluorine or iodine.

20

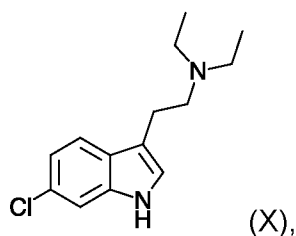
**[00240]** Furthermore, in one embodiment, a psilocybin derivative according to the present disclosure can be a chemical compound having the formula (IX):





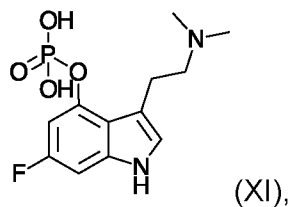
or, a chemical compound having the formula (IX), wherein one or both of the fluorines are substituted for a bromine, chlorine or iodine.

**[00241]** Furthermore, in one embodiment, a psilocybin derivative according to the present disclosure can be a chemical compound having the formula (X):



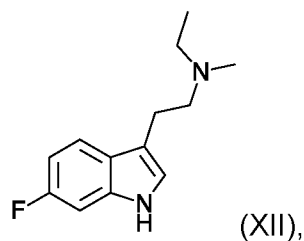
or, a chemical compound having the formula (X), wherein the chlorine is substituted for a bromine, chlorine or iodine.

**[00242]** Furthermore, in one embodiment, a psilocybin derivative according to the present disclosure can be a chemical compound having the formula (XI):



or, a chemical compound having the formula (XI), wherein the fluorine is substituted for a bromine, chlorine or iodine.

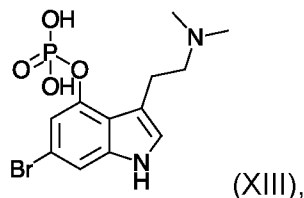
**[00243]** Furthermore, in one embodiment, a psilocybin derivative according to the present disclosure can be a chemical compound having the formula (XII):



or, a chemical compound having the formula (XII), wherein the fluorine is substituted for a bromine, chlorine or iodine.

**[00244]** Furthermore, in one embodiment, a psilocybin derivative according to the present disclosure can be a chemical compound having the formula (XIII):

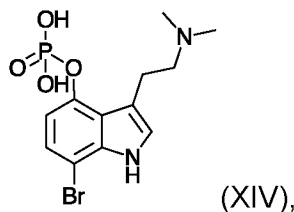
5



or, a chemical compound having the formula (XIII), wherein the bromine is substituted for a chlorine, fluorine or iodine.

**[00245]** Furthermore, in one embodiment, a psilocybin derivative according to the present disclosure can be a chemical compound having the formula (XIV):

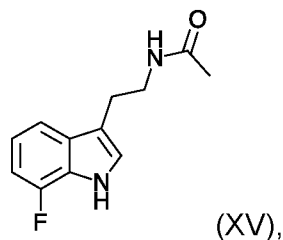
10



or, a chemical compound having the formula (XIV), wherein the bromine is substituted for a chlorine, fluorine or iodine.

**[00246]** Furthermore, in one embodiment, a psilocybin derivative according to the present disclosure can be a chemical compound having the formula (XV):

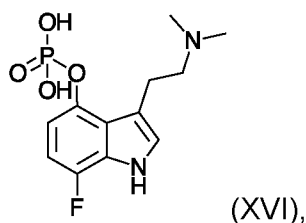
15



or, a chemical compound having the formula (XV), wherein the fluorine is substituted for a bromine, chlorine or iodine.

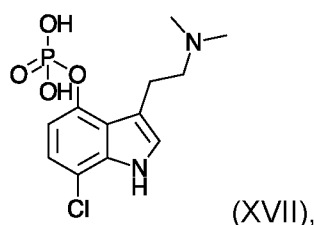
20

**[00247]** Furthermore, in one embodiment, a psilocybin derivative according to the present disclosure can be a chemical compound having the formula (XVI):



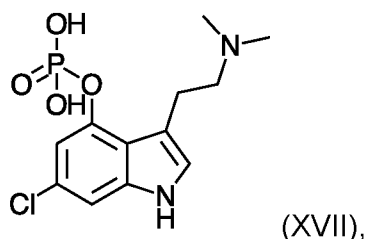
or, a chemical compound having the formula (XV), wherein the fluorine is substituted for a bromine, chlorine or iodine.

**[00248]** Furthermore, in one embodiment, a psilocybin derivative according  
5 to the present disclosure can be a chemical compound having the formula (XVII):



or, a chemical compound having the formula (XVII), wherein the fluorine is substituted for a bromine, chlorine or iodine.

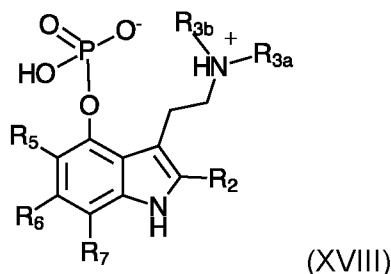
10 **[00249]** Furthermore, in one embodiment, a psilocybin derivative according to the present disclosure can be a chemical compound having the formula (XVII):



or, a chemical compound having the formula (XVII), wherein the fluorine is  
15 substituted for a bromine, chlorine or iodine.

**[00250]** Furthermore, it is noted that the halogenated psilocybin derivatives of the present disclosure include salts thereof. Thus, the nitrogen atom of the ethyl-amino group extending in turn from the C<sub>3</sub> atom may be protonated, and the positive charge may be balanced by, for example, chloride or sulfate ions, to  
20 thereby form a chloride salt or a sulfate salt. Furthermore, in compounds wherein R<sub>4</sub> is a phosphate group, the phosphate group may be de-protonated, and the negative charge may be balanced by, for example, sodium ions or potassium ions, to thereby form a sodium salt or a potassium salt.

**[00251]** Furthermore, it is noted that when R<sub>4</sub> is a phosphate group, the term halogenated psilocybin derivative also includes compounds having the formula (XVIII):

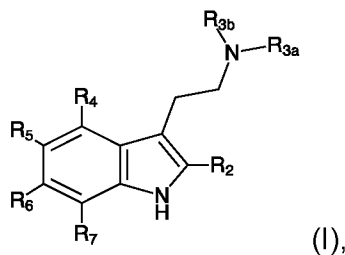


5

wherein, at least one of R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a halogen atom or an alkyl group or O-alkyl group, and wherein any R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> which are not a halogen is a hydrogen atom, and wherein R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom, an alkyl group, and aryl group or an acyl group. The term further includes salts, including pharmaceutically acceptable salts, of halogenated psilocybins having the formula (XVIII), such as a sodium salt, a potassium salt *etc.*

10

**[00252]** Thus, to briefly recap, the present disclosure provides halogenated psilocybin derivatives. The disclosure provides, in particular, a chemical compound or salts thereof having formula (I):

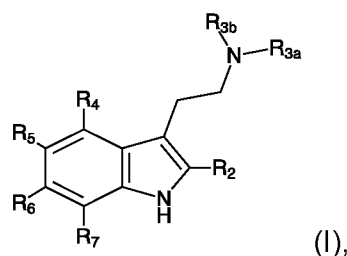


15

wherein in an aspect, at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a halogen atom. In an aspect, in formula (I), each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group. In a further aspect, in formula (I), R<sub>4</sub> when it is not halogenated is a hydrogen atom, an alkyl group or O-alkyl group, a hydroxy group, or a phosphate group. Yet in a further aspect, R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group.

20

**[00253]** In one embodiment of the disclosure, a chemical compound or salts thereof having formula (I) is included:



wherein at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a halogen atom, and wherein each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group, wherein R<sub>4</sub> when it is not halogenated is a hydrogen atom, alkyl group or O-alkyl group, a hydroxy group, or a phosphate group, and wherein R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group.

**[00254]** In one embodiment, at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a halogen atom, and wherein each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or a (C<sub>1</sub>-C<sub>20</sub>)-alkyl group or (C<sub>1</sub>-C<sub>20</sub>)-O-alkyl group. In another embodiment, each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom, a methyl group, ethyl group, a propyl group, an O-methyl group, an O-ethyl group, or an O-propyl group.

**[00255]** In another embodiment, each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or a (C<sub>1</sub>-C<sub>10</sub>)-alkyl group or (C<sub>1</sub>-C<sub>10</sub>)-O-alkyl group. In another embodiment, each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom, a methyl group, ethyl group, a propyl group, an O-methyl group, an O-ethyl group, or an O-propyl group.

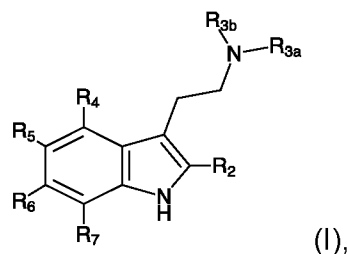
**[00256]** In another embodiment, each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or a (C<sub>1</sub>-C<sub>6</sub>)-alkyl group or (C<sub>1</sub>-C<sub>6</sub>)-O-alkyl group. In another embodiment, each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom, a methyl group, ethyl group, a propyl group, an O-methyl group, an O-ethyl group, or an O-propyl group.

**[00257]** In another embodiment, when R<sub>4</sub> is not halogenated, R<sub>4</sub> is a hydrogen atom, a (C<sub>1</sub>-C<sub>20</sub>)-alkyl group or (C<sub>1</sub>-C<sub>20</sub>)-O-alkyl group, a hydroxy group, or a phosphate group. In another embodiment, when R<sub>4</sub> is not halogenated, R<sub>4</sub> is a hydrogen atom, a (C<sub>1</sub>-C<sub>10</sub>)-alkyl group or (C<sub>1</sub>-C<sub>10</sub>)-O-alkyl group, a hydroxy group, or a phosphate group. In another embodiment, when R<sub>4</sub> is not halogenated, R<sub>4</sub> is a hydrogen atom, a (C<sub>1</sub>-C<sub>6</sub>)-alkyl group or (C<sub>1</sub>-C<sub>6</sub>)-O-alkyl group, a hydroxy group, or a phosphate group. In another embodiment, when R<sub>4</sub> is not halogenated,

R<sub>4</sub> is a hydrogen atom, a methyl group, an ethyl group, a propyl group, a phosphate group, an O-methyl group, an O-ethyl group, or an O-propyl group.

**[00258]** In another embodiment, R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom, a (C<sub>1</sub>-C<sub>20</sub>)-alkyl group, a (C<sub>6</sub>-C<sub>14</sub>)-aryl group, or a -C(=O)(C<sub>1</sub>-C<sub>20</sub>)-alkyl group. In another embodiment, R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom, a (C<sub>1</sub>-C<sub>10</sub>)-alkyl group, a (C<sub>6</sub>-C<sub>10</sub>)-aryl group, or a -C(=O)(C<sub>1</sub>-C<sub>10</sub>)-alkyl group. In another embodiment, R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom, a (C<sub>1</sub>-C<sub>6</sub>)-alkyl group, a phenyl group, or a -C(=O)(C<sub>1</sub>-C<sub>6</sub>)-alkyl group. In another embodiment, R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom, a methyl group, an ethyl group, a propyl group, a phenyl group, -C(=O)-CH<sub>3</sub>, -C(=O)-CH<sub>2</sub>CH<sub>3</sub>, or -C(=O)-CH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>.

**[00259]** In one embodiment of the disclosure, a chemical compound or salts thereof having formula (I) is included:



wherein

R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, and R<sub>7</sub> are independently or simultaneously H, an alkyl group or O-alkyl group or a halogen atom, R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group; and R<sub>4</sub> is hydrogen atom, alkyl group or O-alkyl group, a halogen, a hydroxy group, or a phosphate group; wherein at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub>, and R<sub>7</sub> is a halogen.

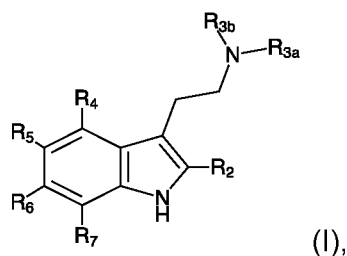
**[00260]** In one embodiment, R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, and R<sub>7</sub> are independently or simultaneously H, (C<sub>1</sub>-C<sub>20</sub>)-alkyl group or (C<sub>1</sub>-C<sub>20</sub>)-O-alkyl group or a halogen atom. In one embodiment, R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, and R<sub>7</sub> are independently or simultaneously H, (C<sub>1</sub>-C<sub>10</sub>)-alkyl group or (C<sub>1</sub>-C<sub>10</sub>)-O-alkyl group or a halogen atom. In one embodiment, R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, and R<sub>7</sub> are independently or simultaneously H, (C<sub>1</sub>-C<sub>6</sub>)-alkyl group or (C<sub>1</sub>-C<sub>6</sub>)-O-alkyl group or a halogen atom. In one embodiment, R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, and R<sub>7</sub> are independently or simultaneously H, methyl, ethyl, propyl, O-methyl, O-ethyl, O-propyl, or a halogen atom.

**[00261]** In one embodiment, R<sub>4</sub> is H, (C<sub>1</sub>-C<sub>20</sub>)-alkyl group or (C<sub>1</sub>-C<sub>20</sub>)-O-alkyl group, a halogen atom or a phosphate group. In one embodiment, R<sub>4</sub> is H, (C<sub>1</sub>-

C<sub>10</sub>)-alkyl group or (C<sub>1</sub>-C<sub>10</sub>)-O-alkyl group, a halogen atom or a phosphate group. In one embodiment, R<sub>4</sub> is H, (C<sub>1</sub>-C<sub>6</sub>)-alkyl group or (C<sub>1</sub>-C<sub>6</sub>)-O-alkyl group, a halogen atom, a hydroxy group, or a phosphate group. In one embodiment, R<sub>4</sub> is H, methyl, ethyl, propyl, O-methyl, O-ethyl, O-propyl, a halogen atom, a hydroxy group, or a phosphate group.

**[00262]** In another embodiment, R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom independently or simultaneously a (C<sub>1</sub>-C<sub>20</sub>)-alkyl group, a (C<sub>6</sub>-C<sub>14</sub>)-aryl group, or a -C(=O)(C<sub>1</sub>-C<sub>20</sub>)-alkyl group. In another embodiment, R<sub>3A</sub> and R<sub>3B</sub> are independently or simultaneously a hydrogen atom, a (C<sub>1</sub>-C<sub>10</sub>)-alkyl group, a (C<sub>6</sub>-C<sub>10</sub>)-aryl group, or a -C(=O)(C<sub>1</sub>-C<sub>10</sub>)-alkyl group or O-alkyl group. In another embodiment, R<sub>3A</sub> and R<sub>3B</sub> are independently or simultaneously a hydrogen atom, a (C<sub>1</sub>-C<sub>6</sub>)-alkyl group, a phenyl group, or a -C(=O)(C<sub>1</sub>-C<sub>6</sub>)-alkyl group. In another embodiment, R<sub>3A</sub> and R<sub>3B</sub> are independently or simultaneously a hydrogen atom, a methyl group, an ethyl group, a propyl group, a butyl group, a phenyl group, -C(=O)-CH<sub>3</sub>, -C(=O)-CH<sub>2</sub>CH<sub>3</sub>, or -C(=O)-CH<sub>2</sub>CH<sub>2</sub>CH<sub>3</sub>.

**[00263]** The halogenated psilocybin derivatives of the present disclosure may be used to prepare a pharmaceutical or recreational drug formulation. Thus in one embodiment, the present disclosure further provides in another aspect, pharmaceutical and recreational drug formulations comprising halogenated psilocybin derivatives. Accordingly, in one aspect, the present disclosure provides in a further embodiment a pharmaceutical or recreational drug formulation comprising a chemical compound or salts thereof having the formula (I):



wherein, at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> is a halogen atom, wherein each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group, wherein R<sub>4</sub> when it is not halogenated is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group, and wherein R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group, together with a diluent, carrier

or excipient. The pharmaceutical or recreational drug formulations may be prepared as liquids, tablets, capsules, microcapsules, nanocapsules, trans-dermal patches, gels, foams, oils, aerosols, nanoparticulates, powders, creams, emulsions, micellar systems, films, sprays, ovules, infusions, teas, decoctions, suppositories, *etc.* and include a pharmaceutically acceptable salt or solvate of the halogenated psilocybin compound together with an excipient. The term “excipient” as used herein means any ingredient other than the chemical compound of the disclosure. As will readily be appreciated by those of skill in art, the selection of excipient may depend on factors such as the particular mode of administration, the effect of the excipient on solubility of the chemical compounds of the present disclosure and methods for their preparation will be readily apparent to those skilled in the art. Such compositions and methods for their preparation may be found, for example, in “Remington's Pharmaceutical Sciences”, 22nd Edition (Pharmaceutical Press and Philadelphia College of Pharmacy at the University of the Sciences, 2012).

**[00264]** The dose when using the compounds of the present disclosure can vary within wide limits, and as is customary and is known to those of skill in the art, the dose can be tailored to the individual conditions in each individual case. The dose depends, for example, on the nature and severity of the illness to be treated, on the condition of the patient, on the compound employed or on whether an acute or chronic disease state is treated or prophylaxis is conducted, on the mode of delivery of the compound, or on whether further active compounds are administered in addition to the compounds of the present disclosure. Representative doses of the present disclosure include, but are not limited to, about 0.001 mg to about 5000 mg, about 0.001 mg to about 2500 mg, about 0.001 mg to about 1000 mg, about 0.001 mg to about 500 mg, about 0.001 mg to about 250 mg, about 0.001 mg to about 100 mg, about 0.001 mg to about 50 mg, and about 0.001 mg to about 25 mg. Representative doses of the present disclosure include, but are not limited to, about 0.0001 to about 1,000 mg, about 10 to about 160 mg, about 10 mg, about 20 mg, about 40 mg, about 80 mg, or about 160 mg. Multiple doses may be administered during the day, especially when relatively large amounts are deemed to be needed, for example 2, 3 or 4, doses. Depending on the subject and as deemed appropriate from the patient's physician or care



giver it may be necessary to deviate upward or downward from the doses described herein.

**[00265]** The pharmaceutical and drug formulations comprising the halogenated psilocybin derivatives of the present disclosure may be administered orally. Oral administration may involve swallowing, so that the compound enters the gastrointestinal tract, or buccal or sublingual administration may be employed by which the compound enters the blood stream directly from the mouth. Formulations suitable for oral administration include both solid and liquid formulations.

**[00266]** Solid formulations include tablets, capsules (containing particulates, liquids, microcapsules, or powders), lozenges (including liquid-filled lozenges), chews, multi- and nano-particulates, gels, solid solutions, liposomal preparations, microencapsulated preparations, creams, films, ovules, suppositories and sprays.

**[00267]** Liquid formulations include suspensions, solutions, syrups and elixirs. Such formulations may be employed as fillers in soft or hard capsules and typically comprise a carrier, for example, water, ethanol, polyethylene glycol, propylene glycol, methylcellulose, or a suitable oil, and one or more emulsifying agents and/or suspending agents. Liquid formulations may also be prepared by the reconstitution of a solid, for example, from a sachet.

**[00268]** Binders are generally used to impart cohesive qualities to a tablet formulation. Suitable binders include microcrystalline cellulose, gelatin, sugars, polyethylene glycol, natural and synthetic gums, polyvinylpyrrolidone, pregelatinised starch, hydroxypropyl cellulose and hydroxypropyl methylcellulose.

**[00269]** Tablets may also contain diluents, such as lactose (monohydrate, spray-dried monohydrate, anhydrous and the like), mannitol, xylitol, dextrose, sucrose, sorbitol, microcrystalline cellulose, starch and dibasic calcium phosphate dihydrate.

**[00270]** Tablets may also optionally comprise surface active agents, such as sodium lauryl sulfate and polysorbate 80. When present, surface active agents may comprise from 0.2% (w/w) to 5% (w/w) of the tablet.

**[00271]** Tablets may further contain lubricants such as magnesium stearate, calcium stearate, zinc stearate, sodium stearyl fumarate, and mixtures of magnesium stearate with sodium lauryl sulphate. Lubricants generally comprise from 0.25% (w/w) to 10% (w/w), from 0.5% (w/w) to 3% (w/w) of the tablet.

**[00272]** In addition to the halogenated psilocybin derivative, tablets may contain a disintegrant. Examples of disintegrants include sodium starch glycolate, sodium carboxymethyl cellulose, calcium carboxymethyl cellulose, croscarmellose sodium, crospovidone, polyvinylpyrrolidone, methyl cellulose, microcrystalline cellulose, lower alkyl-substituted hydroxypropyl cellulose, starch, pregelatinised starch and sodium alginate. Generally, the disintegrant will comprise from 1 % (w/w) to 25% (w/w) or from 5% (w/w) to 20% (w/w) of the dosage form.

**[00273]** Other possible auxiliary ingredients include anti-oxidants, colourants, flavouring agents, preservatives and taste-masking agents.

10 **[00274]** For tablet dosage forms, depending on the desired effective amount of the chemical compound, the chemical compound of the present disclosure may make up from 1% (w/w) to 80 % (w/w) of the dosage form, more typically from 5% (w/w) to 60% (w/w) of the dosage form.

**[00275]** Exemplary tablets contain up to about 80% (w/w) of the chemical compound, from about 10% (w/w) to about 90% (w/w) binder, from about 0% (w/w) to about 85% (w/w) diluent, from about 2% (w/w) to about 10% (w/w) disintegrant, and from about 0.25% (w/w) to about 10% (w/w) lubricant.

**[00276]** The formulation of tablets is discussed in "Pharmaceutical Dosage Forms: Tablets", Vol. 1 – Vol. 3, by CRC Press (2008).

20 **[00277]** The pharmaceutical and recreational drug formulations comprising the halogenated psilocybin derivatives of the present disclosure may also be administered directly into the blood stream, into muscle, or into an internal organ. Thus, the pharmaceutical and recreational drug formulations can be administered parenterally (for example, by subcutaneous, intravenous, intraarterial, intrathecal, intraventricular, intracranial, intramuscular, or intraperitoneal injection). Parenteral formulations are typically aqueous solutions which may contain excipients such as salts, carbohydrates and buffering agents (in one embodiment, to a pH of from 3 to 9), but, for some applications, they may be more suitably formulated as a sterile non-aqueous solution or as a dried form to be used in conjunction with a suitable vehicle such as sterile water.

30 **[00278]** Formulations comprising the halogenated psilocybin derivatives of the present disclosure for parenteral administration may be formulated to be immediate and/or modified release. Modified release formulations include delayed-, sustained-, pulsed-, controlled-, targeted and programmed release.

Thus the chemical compounds of the disclosure may be formulated as a solid, semi-solid, or thixotropic liquid for administration as an implanted depot providing modified release of the active compound. Examples of such formulations include drug-coated stents and poly(dl-lactic-coglycolic)acid (PGLA) microspheres.

5   **[00279]**       The pharmaceutical or recreational drug formulations of the present disclosure also may be administered topically to the skin or mucosa, *i.e.* dermally or transdermally. Example pharmaceutical and recreational drug formulations for this purpose include gels, hydrogels, lotions, solutions, creams, ointments, dusting  
10   powders, cosmetics, oils, eye drops, dressings, foams, films, skin patches, wafers, implants, sponges, fibres, bandages and microemulsions. Liposomes may also be used. Example carriers include alcohol, water, mineral oil, liquid petrolatum, white petrolatum, glycerin, polyethylene glycol and propylene glycol. Penetration enhancers may be incorporate (see: for example, Finnin, B. and Morgan, T.M., 1999 J. Pharm. Sci, 88 (10), 955-958).

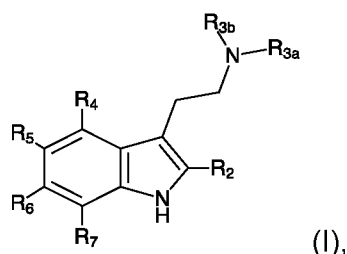
15   **[00280]**       Other means of topical administration include delivery by electroporation, iontophoresis, phonophoresis, sonophoresis and microneedle or needle-free (*e.g.*, Powderject™, Bioject™, *etc.*) injection.

**[00281]**       Pharmaceutical and recreational drug formulations for inhalation or insufflation include solutions and suspensions in pharmaceutically acceptable,  
20   aqueous or organic solvents, or mixtures thereof, and powders. The liquid or solid pharmaceutical compositions can contain suitable pharmaceutically acceptable excipients. In some embodiments, the pharmaceutical compositions are administered by the oral or nasal respiratory route for local or systemic effect. Pharmaceutical compositions in pharmaceutically acceptable solvents can be  
25   nebulized by use of inert gases. Nebulized solutions can be inhaled directly from the nebulizing device or the nebulizing device can be attached to a face mask tent, or intermittent positive pressure breathing machine. Solution, suspension, or powder pharmaceutical compositions can be administered, *e.g.*, orally or nasally, from devices that deliver the formulation in an appropriate manner.

30   **[00282]**       In further embodiments, in which the halogenated psilocybin compounds of present disclosure are used as a recreational drug, the compounds may be included in compositions such as food or food product, a beverage, a food seasoning, a personal care product, such as a cosmetic, perfume or bath oil, or oils (both for topical administration as massage oil, or to be burned or aerosolized).

The chemical compounds of the present disclosure may also be included in a “vape” product, which may also include other drugs, such as nicotine, and flavorings.

**[00283]** The pharmaceutical formulations comprising the chemical compounds of the present disclosure may be used to treat a subject, and in particular to treat a psychiatric disorder in a subject. Accordingly, the present disclosure includes in a further embodiment, a method for treating a psychiatric disorder, the method comprising administering to a subject in need thereof a pharmaceutical formulation comprising a chemical compound or salts thereof having the formula (I):



wherein, at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> is a halogen atom, wherein each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group, wherein R<sub>4</sub> when it is not halogenated is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group, and wherein R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom, or an alkyl group, an aryl group, or an acyl group, wherein the pharmaceutical formulation is administered in an effective amount to treat the psychiatric disorder in the subject.

**[00284]** Psychiatric disorders that may be treated include, for example, neurodevelopmental disorders such as intellectual disability, global development delay, communication disorders, autism spectrum disorder, and attention-deficit hyperactivity disorder (ADHD); bipolar and related disorders, such as mania, and depressive episodes; anxiety disorder, such as generalized anxiety disorder (GAD), agoraphobia, social anxiety disorder, specific phobias (natural events, medical, animal, situational, for example), panic disorder, and separation anxiety disorder; stress disorders, such as acute stress disorder, adjustment disorders, post-traumatic stress disorder (PTSD), and reactive attachment disorder; dissociative disorders, such as dissociative amnesia, dissociative identity disorder, and depersonalization/derealization disorder; somatoform disorders, such as

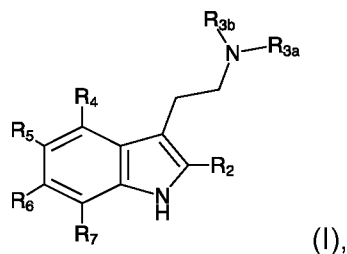
somatic symptom disorders, illness anxiety disorder, conversion disorder, and factitious disorder; eating disorders, such as anorexia nervosa, bulimia nervosa, rumination disorder, pica, and binge-eating disorder; sleep disorders, such as narcolepsy, insomnia disorder, hypersomnolence, breathing-related sleep disorders, parasomnias, and restless legs syndrome; disruptive disorders, such as kleptomania, pyromania, intermittent explosive disorder, conduct disorder, and oppositional defiant disorder; depressive disorders, such as disruptive mood dysregulation disorder, major depressive disorder, persistent depressive disorder (dysthymia), premenstrual dysphoric disorder, substance/medication-induced depressive disorder, postpartum depression, and depressive disorder caused by another medical condition, for example, psychiatric and existential distress within life-threatening cancer situations (ACS Pharmacol Transl Sci 4: 553-562; J Psychiatr Res 137: 273); substance-related disorders, such as alcohol-related disorders, cannabis related disorders, inhalant-use related disorders, stimulant use disorders, and tobacco use disorders; neurocognitive disorders, such as delirium; schizophrenia; compulsive disorders, such as obsessive compulsive disorders (OCD), body dysmorphic disorder, hoarding disorder, trichotillomania disorder, excoriation disorder, substance/medication induced obsessive-compulsive disorder, and obsessive-compulsive disorder related to another medical condition; and personality disorders, such as antisocial personality disorder, avoidant personality disorder, borderline personality disorder, dependent personality disorder, histrionic personality disorder, narcissistic personality disorder, obsessive-compulsive personality disorder, paranoid personality disorder, schizoid personality disorder, schizotypal personality disorder.

25 **[00285]** In an aspect, the compounds of the present disclosure may be used to be contacted with a 5-HT<sub>2A</sub> receptor to thereby modulate the 5-HT<sub>2A</sub> receptor. Such contacting includes bringing a compound of the present disclosure and 5-HT<sub>2A</sub> receptor together under *in vitro* conditions, for example, by introducing the compounds in a sample containing a 5-HT<sub>2A</sub> receptor, for example, a sample containing purified 5-HT<sub>2A</sub> receptors, or a sample containing cells comprising 5-HT<sub>2A</sub> receptors. *In vitro* conditions further include the conditions described in Example 3 hereof. Contacting further includes bringing a compound of the present disclosure and 5-HT<sub>2A</sub> receptor together under *in vivo* conditions. Such *in vivo*

conditions include the administration to an animal or human subject, for example, of a pharmaceutically effective amount of the compound of the present disclosure, when the compound is formulated together with a pharmaceutically active carrier, diluent or excipient, as hereinbefore described, to thereby treat the subject. Upon  
 5 having contacted the 5-HT<sub>2A</sub> receptor, the compound may activate the 5-HT<sub>2A</sub> receptor or inhibit the 5-HT<sub>2A</sub> receptor.

**[00286]** Thus, in a further aspect, the condition that may be treated in accordance herewith can be any 5-HT<sub>2A</sub> receptor mediated disorder. Such disorders include, but are not limited to schizophrenia, psychotic disorder,  
 10 attention deficit hyperactivity disorder, autism, and bipolar disorder.

**[00287]** The chemical compounds of the present disclosure may also be used as a feedstock material for other psilocybin derivatives. Thus in one embodiment, the chemical compounds of the present disclosure may be in used manufacture of a pharmaceutical or recreational drug formulation, wherein the  
 15 manufacture may comprise derivatizing a chemical compound or salts thereof having the formula (I):



wherein, at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> is a halogen atom, wherein each non-  
 20 halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group, wherein R<sub>4</sub> when it is not halogenated is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group, and wherein R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group.

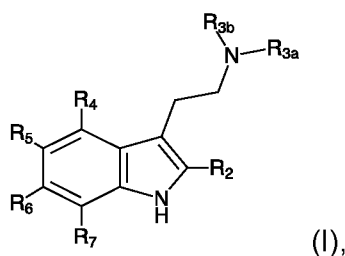
**[00288]** In order to use the compound having formula (I) as a feedstock, one  
 25 or more halogen atoms may be replaced by any atoms or groups, for example hydrocarbon groups. Those of skill in the art will be generally familiar with methods that may be used to substitute halogen groups. In this respect, guidance may be found in Schnepel C. *et al.* (2017) Chem. Eur. J. 23:12064-12086; Durak L.J. *et*

*al.* (2016) ACS Catal. 6: 1451; Runguphan W. *et al.* (2013) Org Lett 15: 2850; Corr M.J. *et al.* (2017) Chem. Sci. 8: 2039; and Roy A.D. *et al.* Chem. Comm. 4831.

**[00289]** Turning now to methods of making the halogenated psilocybin derivatives, it is noted that the psilocybin compounds of the present disclosure may be prepared in any suitable manner, including any organic chemical synthesis methods, biosynthetic methods, or a combination thereof. Synthesis generally may involve selecting a psilocybin precursor compound, and modifying the psilocybin precursor compound to form a halogenated psilocybin derivative. In this respect, it is noted that a non-halogenated psilocybin derivative may be selected and modified to form psilocybin, which subsequently may be halogenated, or, alternatively, a halogenated psilocybin derivative may be selected to subsequently form a halogenated psilocybin. Suitable psilocybin precursor compounds include compounds comprising an indole prototype structure (see: **FIG. 2**), including, for example tryptophan, tryptamine, 4-hydroxyindole, 4-hydroxytryptophan, 4-hydroxytryptamine, norbaeocystin and baeocystin, and halogenated forms thereof, notably, with respect to the indole prototype structure, 2-, 5-, 6-, 7-halogenated forms thereof. The halogenated psilocybin precursor compounds may be provided in a more or less chemically pure form, for example, in the form of a psilocybin precursor preparation having a purity of at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or at least 99.9%. The psilocybin precursor compounds may be chemically synthesized, or obtained from a fine chemical manufacturer.

**[00290]** In one embodiment of the present disclosure the halogenated psilocybin derivatives may be formed biosynthetically. Accordingly, the present disclosure further includes in one embodiment, a method of making a halogenated psilocybin derivative the method comprising:

- (a) contacting a halogenated psilocybin precursor compound with a host cell comprising a psilocybin biosynthetic enzyme complement; and
- (b) growing the host cell to produce a halogenated psilocybin compound or salts thereof having the formula (I):



wherein, at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> is a halogen atom, wherein each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group, wherein R<sub>4</sub> when it is not halogenated is a phosphate group, hydrogen atom or an alkyl group or O-alkyl group, and wherein R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom, or an alkyl group, an aryl group, or an acyl group of the chemical compound.

**[00291]** Implementation of the foregoing example embodiment initially involves providing halogenated psilocybin precursor compounds and host cells having a psilocybin biosynthetic enzyme complement. Accordingly, next, exemplary halogenated psilocybin precursor compounds and example host cells that may be selected and used in accordance with the present disclosure will be described. Thereafter, example methodologies and techniques will be described to contact and use the halogenated psilocybin precursor compounds and cells to produce example halogenated psilocybin compounds.

**[00292]** A variety of halogenated psilocybin precursor compounds may be selected, prepared and used. In some embodiments, for example, the halogenated psilocybin precursor compound is a compound comprising a halogenated indole prototype structure. Examples of such compounds are halogenated tryptophan, halogenated tryptamine, halogenated 4-hydroxyindole, halogenated 4-hydroxytryptophan, halogenated 4-hydroxytryptamine, halogenated norbaeocystin, and halogenated baeocystin, and in particular the 2-halogenated, 5-halogenated, 6-halogenated and 7-halogenated forms (with respect to the indole prototype structure) of any of the foregoing.

**[00293]** Turning now to the host cells that can be used in accordance with the present disclosure, it is initially noted that a variety of host cells may be selected in accordance with the present disclosure, including microorganism host cells, plant host cells, and animal host cells.



**[00294]** In accordance herewith the host cell includes a psilocybin biosynthetic enzyme complement. Such cells can be obtained in at least two ways. First, in some embodiments, host cells may be selected in which a psilocybin biosynthetic enzyme complement is naturally present. Generally cells naturally  
5 producing psilocybin for example, cells of fungal species belonging to the genus *psilocybe*, are suitable in this respect. Second, in some embodiments, a host cell that not naturally produces psilocybin may be modulated to produce a psilocybin biosynthetic enzyme complement. Thus, for example, a nucleic acid sequence encoding a psilocybin biosynthetic enzyme complement may be introduced into a  
10 host cell, and upon cell growth the host cells can make the psilocybin biosynthetic enzyme complement.

**[00295]** Typically a nucleic acid sequence encoding one or more enzymes constituting a psilocybin biosynthetic enzyme complement further includes one or more additional nucleic acid sequences, for example, a nucleic acid sequences  
15 controlling expression of the one or more enzymes, and these one or more additional nucleic acid sequences together with the nucleic acid sequence encoding the one or more enzymes can be said to form a chimeric nucleic acid sequence.

**[00296]** A host cell which upon cultivation expresses the chimeric nucleic acid can be selected and used in accordance with the present disclosure. Suitable  
20 host cells in this respect include, for example, microbial cells, such as bacterial cells, yeast cells, for example, and algal cells or plant cells. A variety of techniques and methodologies to manipulate host cells to introduce nucleic acid sequences in cells and attain expression exists and are well known to the skilled artisan.  
25 These methods include, for example, cation based methods, for example, lithium ion or calcium ion based methods, electroporation, biolistics, and glass beads based methods. As will be known to those of skill in the art, depending on the host cell selected, the methodology to introduce nucleic acid material in the host cell may vary, and, furthermore, methodologies may be optimized for uptake of nucleic  
30 acid material by the host cell, for example, by comparing uptake of nucleic acid material using different conditions. Detailed guidance can be found, for example, in Sambrook *et al.*, Molecular Cloning, a Laboratory Manual, Cold Spring Harbor Laboratory Press, 2012, Fourth Ed. It is noted that the chimeric nucleic acid is a

non-naturally occurring chimeric nucleic acid sequence and can be said to be heterologous to the host cell.

**[00297]** In some embodiments, the one or more enzymes constituting a psilocybin enzyme complement can be selected from by a nucleic acid sequence

5 selected from the nucleic acid sequences consisting of:

- (a) SEQ.ID NO: 1, SEQ.ID NO: 3, SEQ.ID NO: 5, SEQ.ID NO: 7, SEQ.ID NO: 9, SEQ.ID NO 11, SEQ.ID NO: 36, SEQ.ID NO: 38, SEQ.ID NO: 40, SEQ. ID NO: 42, SEQ.ID NO: 44, SEQ.ID NO: 46, SEQ.ID NO: 48, SEQ.ID NO: 50, SEQ.ID NO: 52, SEQ.ID NO: 54, SEQ. ID NO: 56, SEQ.ID NO: 58, 10 SEQ.ID NO: 60, SEQ.ID NO: 64, SEQ.ID NO: 66, SEQ.ID NO: 68, SEQ. ID NO: 70, and SEQ.ID NO: 72;
- (b) a nucleic acid sequence that is substantially identical to any one of the nucleic acid sequences of (a);
- (c) a nucleic acid sequence that is substantially identical to any one of 15 the nucleic acid sequences of (a) but for the degeneration of the genetic code;
- (d) a nucleic acid sequence that is complementary to any one of the nucleic acid sequences of (a);
- (e) a nucleic acid sequence encoding a polypeptide having any one of 20 the amino acid sequences set forth in SEQ.ID NO: 2, SEQ.ID NO: 4, SEQ.ID NO: 6, SEQ.ID NO: 8, SEQ.ID NO: 10, SEQ.ID NO 12, SEQ.ID NO: 37, SEQ.ID NO: 39, SEQ.ID NO: 41, SEQ. ID NO: 43, SEQ.ID NO: 45, SEQ.ID NO: 47, SEQ.ID NO: 49, SEQ.ID NO: 51, SEQ.ID NO: 53, SEQ.ID NO: 55, SEQ. ID NO: 57, SEQ.ID NO: 59, SEQ.ID NO: 61, SEQ.ID NO: 63, 25 SEQ.ID NO: 65, SEQ.ID NO: 67, SEQ.ID NO: 69, SEQ. ID NO: 71, and SEQ.ID NO: 73;
- (f) a nucleic acid sequence that encodes a functional variant of any one of the amino acid sequences set forth in SEQ.ID NO: 2, SEQ.ID NO: 4, SEQ.ID NO: 6, SEQ.ID NO: 8, SEQ.ID NO: 10, SEQ.ID NO 12, SEQ.ID NO: 37, SEQ.ID NO: 39, SEQ.ID NO: 41, SEQ. ID NO: 43, SEQ.ID NO: 45, 30 SEQ.ID NO: 47, SEQ.ID NO: 49, SEQ.ID NO: 51, SEQ.ID NO: 53, SEQ.ID NO: 55, SEQ. ID NO: 57, SEQ.ID NO: 59, SEQ.ID NO: 61, SEQ.ID NO: 63, SEQ.ID NO: 65, SEQ.ID NO: 67, SEQ.ID NO: 69, SEQ. ID NO: 71, and SEQ.ID NO: 73; and

(g) a nucleic acid sequence that hybridizes under stringent conditions to any one of the nucleic acid sequences set forth in (a), (b), (c), (d), (e) or (f).

**[00298]** Thus any of the nucleic acid sequence set forth in (a), (b), (c), (d), (e), (f) or (g) may be selected and introduced into a host cell. In general, however  
5 the nucleic acid sequence is selected in conjunction with the selected psilocybin precursor compound, as hereinafter further discussed in reference to **FIGS. 8A** and **8B**.

**[00299]** One example host cell that conveniently may be used is *Escherichia coli*. The preparation of the *E. coli* vectors may be accomplished using commonly  
10 known techniques such as restriction digestion, ligation, gel electrophoresis, DNA sequencing, the polymerase chain reaction (PCR) and other methodologies. A wide variety of cloning vectors is available to perform the necessary steps required to prepare a recombinant expression vector. Among the vectors with a replication system functional in *E. coli*, are vectors such as pBR322, the pUC series of  
15 vectors, the M13 mp series of vectors, pBluescript etc. Suitable promoter sequences for use in *E. coli* include, for example, the T7 promoter, the T5 promoter, tryptophan (trp) promoter, lactose (lac) promoter, tryptophan/lactose (tac) promoter, lipoprotein (lpp) promoter, and  $\lambda$  phage PL promoter. Typically, cloning vectors contain a marker, for example, an antibiotic resistance marker,  
20 such as ampicillin or kanamycin resistance marker, allowing selection of transformed cells. Nucleic acid sequences may be introduced in these vectors, and the vectors may be introduced in *E. coli* by preparing competent cells, electroporation or using other well known methodologies to a person of skill in the art. *E. coli* may be grown in an appropriate medium, such as Luria-Broth medium  
25 and harvested. Recombinant expression vectors may readily be recovered from cells upon harvesting and lysing of the cells.

**[00300]** Another example host cell that may be conveniently used is a yeast cell. Example yeast host cells that can be used are yeast cells belonging to the genus *Candida*, *Kluyveromyces*, *Saccharomyces*, *Schizosaccharomyces*, *Pichia*,  
30 *Hansenula*, and *Yarrowia*. In specific example embodiments, the yeast cell can be a *Saccharomyces cerevisiae* cell, a *Yarrowia lipolytica* cell, or *Pichia pastoris* cell.

**[00301]** A number of vectors exist for the expression of recombinant proteins in yeast host cells. Examples of vectors that may be used in yeast host cells include, for example, Yip type vectors, YEp type vectors, YRp type vectors, YCp

type vectors, pGPD-2, pAO815, pGAPZ, pGAPZ $\alpha$ , pHIL-D2, pHIL-S1, pPIC3.5K, pPIC9K, pPICZ, pPICZ $\alpha$ , pPIC3K, pHWO10, pPUZZLE and 2  $\mu$ m plasmids. Such vectors are known to the art and are, for example, described in Cregg *et al.*, Mol Biotechnol. (2000) 16(1): 23-52. Suitable promoter sequences for use in yeast host cells are also known and described, for example, in Mattanovich *et al.*, Methods Mol. Biol., 2012, 824:329-58, and in Romanos *et al.*, 1992, Yeast 8: 423-488. Examples of suitable promoters for use in yeast host cells include promoters of glycolytic enzymes, like triosephosphate isomerase (TPI), phosphoglycerate kinase (PGK), glyceraldehyde-3-phosphate dehydrogenase (GAPDH or GAP) and variants thereof, lactase (LAC) and galactosidase (GAL), *P. pastoris* glucose-6-phosphate isomerase promoter (PPGI), the 3-phosphoglycerate kinase promoter (PPGK), the glycerol aldehyde phosphate dehydrogenase promoter (PGAP), translation elongation factor promoter (PTEF), *S. cerevisiae* enolase (ENO-1), *S. cerevisiae* galactokinase (GAL1), *S. cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH1, ADH2/GAP), *S. cerevisiae* triose phosphate isomerase (TPI), *S. cerevisiae* metallothionein (CUP1), and *S. cerevisiae* 3-phosphoglycerate kinase (PGK), and the maltase gene promoter (MAL). Marker genes suitable for use in yeast host cells are also known to the art. Thus, antibiotic resistance markers, such as ampicillin resistance markers, can be used in yeast, as well as marker genes providing genetic functions for essential nutrients, for example, leucine (LEU2), tryptophan (TRP1 and TRP2), uracil (URA3, URA5, URA6), histidine (HIS3), and the like. Methods for introducing vectors into yeast host cells can, for example, be found in S. Kawai *et al.*, 2010, Bioeng. Bugs 1(6): 395-403.

**[00302]** Further, guidance with respect to the preparation of expression vectors and introduction thereof into host cells, including in *E. coli* cells, yeast cells, and other host cells, may be found in, for example: Sambrook *et al.*, Molecular Cloning, a Laboratory Manual, Cold Spring Harbor Laboratory Press, 2012, Fourth Ed.

**[00303]** Thus, to briefly recap, a host cell comprising a chimeric nucleic acid comprising (i) a nucleic acid sequence controlling expression in a host cell and (ii) a nucleic acid sequence encoding a psilocybin biosynthetic enzyme complement, can be prepared in accordance with the present disclosure.

**[00304]** In accordance herewith, host cells are grown to multiply and to express a chimeric nucleic acid. Expression of the chimeric nucleic acid results in the biosynthetic production in the host cell of a psilocybin biosynthetic enzyme complement. Growth media and growth conditions can vary depending on the host cell that is selected, as will be readily appreciated to those of ordinary skill in the art. Growth media typically contain a carbon source, one or several nitrogen sources, essential salts including salts of potassium, sodium, magnesium, phosphate and sulphate, trace metals, water soluble vitamins, and process aids including but not limited to antifoam agents, protease inhibitors, stabilizers, ligands and inducers. Example carbon sources are e.g. mono- or disaccharides. Example nitrogen sources are, e.g. ammonia, urea, amino acids, yeast extract, corn steep liquor and fully or partially hydrolyzed proteins. Example trace metals are e.g. Fe, Zn, Mn, Cu, Mo and H<sub>3</sub>BO<sub>3</sub>. Example water soluble vitamins are e.g. biotin, pantothenate, niacin, thiamine, p-aminobenzoic acid, choline, pyridoxine, folic acid, riboflavin and ascorbic acid. Further, specific example media include liquid culture media for the growth of yeast cells and bacterial cells including, Luria-Bertani (LB) broth for bacterial cell cultivation, and yeast extract peptone dextrose (YEPD or YPD), for yeast cell cultivation. Further media and growth conditions can be found in Sambrook *et al.*, Molecular Cloning, a Laboratory Manual, Cold Spring Harbor Laboratory Press, 2012, Fourth Ed.

**[00305]** In order for the host cells to produce the halogenated psilocybin compound, the cells are provided with a psilocybin precursor compound. Thus in accordance herewith, host cells may be contacted with a psilocybin precursor compound. In some embodiments, a psilocybin precursor compound can be exogenously supplied, for example, by including a psilocybin precursor compound in the growth medium of the host cells, and growing the host cells in a medium including the psilocybin precursor compound.

**[00306]** Referring next to **FIGS. 8A** and **8B**, shown therein is an example natural biosynthetic pathway showing the conversion of example psilocybin precursor compounds to form psilocybin (**FIG. 8A**). Thus, as can be appreciated from **FIG. 8A**, various psilocybin precursor compounds may be selected and prepared in halogenated form, in conjunction with a psilocybin biosynthetic enzyme complement.

**[00307]** Thus, by way of example, referring to **FIG. 8B**, a chlorinated psilocybin precursor compound, for example, 6-chloro-4-hydroxy indole may be selected and contacted with a host cell comprising a psilocybin biosynthetic enzyme complement comprising PsiD, (e.g. SEQ.ID NO: 2) PsiH (e.g. SEQ.ID NO: 4), PsiK (e.g. SEQ.ID NO: 6) and PsiM (e.g. SEQ.ID NO: 8), and upon growth of the cells 6-chloro-psilocybin can be formed. Or by way of another example, 6-chloro-tryptophan may be selected and contacted with a host cell comprising a psilocybin biosynthetic enzyme complement comprising TrpB (SEQ.ID NO: 12) PsiD, (e.g. SEQ.ID NO: 2), PsiK (e.g. SEQ.ID NO: 6) and PsiM (e.g. SEQ.ID NO: 8), and upon growth of the cells 6-chloro-psilocybin can be formed. It will be clear to those of skill in the art that by referring to **FIGS. 8A** and **8B**, other halogenated psilocybin precursors may be selected (e.g. 2-chloro tryptophan, 4-chloro-tryptophan, 5-chloro-tryptophan, 7-chlorotryptophan; 6-bromo-tryptophan, 6-fluoro-tryptophan, 6-iodotryptophan; and 6-chlorotryptamine, 6-chloro-4-hydroxytryptophan, 6-hydroxy-4-chloro-tryptamine, 6-chloro-norbaeocystin or 6-chloro-baeocystin), and contacted with a host cell comprising an appropriate psilocybin biosynthetic enzyme complement to produce a halogenated psilocybin derivative.

**[00308]** Thus, in one further example embodiment, a halogenated indole, such as 6-chloro-4-hydroxy indole, for example (see: **FIG. 8B**), may be selected and contacted with a host cell comprising a psilocybin biosynthetic enzyme complement comprising (i) a nucleic acid encoding TrpB selected from SEQ.ID NO: 12, SEQ.ID NO: 65, SEQ.ID NO: 67, SEQ.ID NO: 69, SEQ.ID NO: 71, and SEQ.ID NO: 73, or a nucleic acid sequence encoding a polypeptide sequence substantially similar to any of the aforementioned sequences; (ii) a nucleic acid encoding PsiD selected from SEQ.ID NO: 2, SEQ.ID NO: 10, SEQ.ID NO: 37, SEQ.ID NO: 39, SEQ.ID NO: 41, SEQ.ID NO: 43, SEQ.ID NO: 45, SEQ.ID NO: 47, SEQ.ID NO: 49, SEQ.ID NO: 51, SEQ.ID NO: 53, SEQ.ID NO: 55, and SEQ.ID NO: 57, or a nucleic acid sequence encoding a polypeptide sequence substantially similar to any of the aforementioned sequences; (iii) a nucleic acid encoding PsiK encoded by SEQ.ID NO: 6, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; and (iv) a nucleic acid encoding PsiM selected from SEQ.ID NO: 8, SEQ.ID NO: 59 and SEQ.ID NO: 61 or a nucleic acid

sequence encoding a polypeptide sequence substantially similar to any of the aforementioned sequences to produce a halogenated psilocybin derivative.

**[00309]** Thus, in one further example embodiment, a halogenated tryptophan, such as 6-chloro-4-hydroxy tryptophan, for example (see: **FIG. 8B**),  
5 may be selected and contacted with a host cell comprising a psilocybin biosynthetic enzyme complement comprising (i) a nucleic acid encoding PsiD selected from SEQ.ID NO: 2, SEQ.ID NO: 10, SEQ.ID NO: 37, SEQ.ID NO: 39, SEQ.ID NO: 41, SEQ.ID NO: 43, SEQ.ID NO: 45, SEQ.ID NO: 47, SEQ.ID NO: 49, SEQ.ID NO: 51, SEQ.ID NO: 53, SEQ.ID NO: 55, and SEQ.ID NO: 57, or a  
10 nucleic acid sequence encoding a polypeptide sequence substantially similar to any of the aforementioned sequences; (ii) a nucleic acid encoding PsiK encoded by SEQ.ID NO: 6, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; and (iii) a nucleic acid encoding PsiM selected from SEQ.ID NO: 8, SEQ.ID NO: 59 and SEQ.ID NO: 61 or a nucleic acid sequence  
15 encoding a polypeptide sequence substantially similar to any of the aforementioned sequences to produce a halogenated psilocybin derivative.

**[00310]** Thus, in one further example embodiment, a halogenated tryptamine, such as 6-chloro-4-hydroxy tryptamine, for example (see: **FIG. 8B**), may be selected and contacted with a host cell comprising a psilocybin  
20 biosynthetic enzyme complement comprising (i) a nucleic acid encoding PsiK encoded by SEQ.ID NO: 6, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; and (ii) and a nucleic acid encoding PsiM selected from SEQ.ID NO: 8, SEQ.ID NO: 59 and SEQ.ID NO: 61 or a nucleic acid sequence encoding a polypeptide sequence substantially similar to any of the  
25 aforementioned sequences to produce a halogenated psilocybin derivative.

**[00311]** Thus, in one further example embodiment, a halogenated norbaeocystin, such as 6-chloro-norbaeocystin, or a halogenated baecocystin may, such as 6-chlorobaecocystin, for example (see: **FIG. 8B**), may be selected and contacted with a host cell comprising a psilocybin biosynthetic enzyme  
30 complement comprising a nucleic acid encoding PsiM selected from SEQ.ID NO: 8, SEQ.ID NO: 59 and SEQ.ID NO: 61 or a nucleic acid sequence encoding a polypeptide sequence substantially similar to any of the aforementioned sequences to produce a halogenated psilocybin derivative.

**[00312]** Thus, in one further example embodiment, a halogenated tryptophan, such as 6-chloro-4-hydroxy tryptophan, for example (see: **FIG. 8B**), may be selected and contacted with a host cell comprising a psilocybin biosynthetic enzyme complement comprising (i) a nucleic acid encoding PsiD selected from SEQ.ID NO: 2, SEQ.ID NO: 10, SEQ.ID NO: 37, SEQ.ID NO: 39, SEQ.ID NO: 41, SEQ.ID NO: 43, SEQ.ID NO: 45, SEQ.ID NO: 47, SEQ.ID NO: 49, SEQ.ID NO: 51, SEQ.ID NO: 53, SEQ.ID NO: 55, and SEQ.ID NO: 57, or a nucleic acid sequence encoding a polypeptide sequence substantially similar to any of the aforementioned sequences; (ii) a nucleic acid encoding PsiH encoded by SEQ.ID NO: 4, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; (iii) a nucleic acid encoding PsiK encoded by SEQ.ID NO: 6, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; and (iv) and a nucleic acid encoding PsiM selected from SEQ.ID NO: 8, SEQ.ID NO: 59 and SEQ.ID NO: 61 or a nucleic acid sequence encoding a polypeptide sequence substantially similar to any of the aforementioned sequences to produce a halogenated psilocybin derivative.

**[00313]** Thus, in one further example embodiment, a halogenated tryptamine, such as 6-chloro-tryptamine, for example (see: **FIG. 8B**), may be selected and contacted with a host cell comprising a psilocybin biosynthetic enzyme complement comprising (i) a nucleic acid encoding PsiH encoded by SEQ.ID NO: 4, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; (ii) a nucleic acid encoding PsiK encoded by SEQ.ID NO: 6, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; and (iii) and a nucleic acid encoding PsiM selected from SEQ.ID NO: 8, SEQ.ID NO: 59 and SEQ.ID NO: 61 or a nucleic acid sequence encoding a polypeptide sequence substantially similar to any of the aforementioned sequences to produce a halogenated psilocybin derivative.

**[00314]** In some embodiments, the psilocybin can be a halogenated psilocybin precursor compound which is exogenously supplied to a host cell, for example by inclusion in the host cell's growth medium. Thus, for example, referring to **FIG. 8**, it will be understood that in accordance herewith, for example, 6-chloro-tryptophan or 6-chloro-4-hydroxy indole, may be included in the growth medium of a host cell comprising a psilocybin biosynthetic enzyme complement.



**[00315]** It is noted that in some embodiments, the host cells may in addition to the psilocybin biosynthetic enzyme complement include a halogenase. In such embodiments the cell may be contacted with a non-halogenated psilocybin precursor compound, and halogenation may occur *in vivo* in the host cells. Thus, referring again to **FIG. 8**, by way of example, non-halogenated tryptophan may be selected and contacted with a host cell comprising a psilocybin biosynthetic enzyme complement comprising PsiD, (e.g. SEQ.ID NO: 2) PsiH (e.g. SEQ.ID NO: 4), PsiK (e.g. SEQ.ID NO: 6) and PsiM (e.g. SEQ.ID NO: 8), the host cells further including a halogenase, and upon growth of the host cells in the presence of a halogen, for example in the form of a halogen salt, sodium chloride (NaCl) or sodium bromide (NaBr) halogenated psilocybin derivatives can be formed.

**[00316]** Suitable halogenases that may be used in accordance herewith include, for example, a PyrH halogenase of *Streptomyces rugosporus* (SEQ.ID NO: 13; SEQ.ID NO: 14), which may, for example, be contacted with non-halogenated tryptophan to form 5-halo-tryptophan; a KtZR halogenase of *Kutzneria* sp. 744 (SEQ.ID NO: 15; SEQ.ID NO: 16), which may be contacted with non-halogenated tryptophan to form 6-halo-tryptophan; a SstH halogenase of *Streptomyces toxytricini* (SEQ.ID NO: 17; SEQ.ID NO: 18), which may be contacted with non-halogenated tryptophan to form 6-halo-tryptophan; a Thal halogenase of *Streptomyces albogriseolus* (SEQ.ID NO: 19; SEQ.ID NO: 20), which may be contacted with non-halogenated tryptophan to form 6-halo-tryptophan; a BorH halogenase (see: SEQ.ID NO: 21; SEQ.ID NO: 22), which may be contacted with non-halogenated tryptophan to form 6-halo-tryptophan; a KtzQ halogenase of *Kutzneria* sp. 744 (SEQ.ID NO: 23; SEQ.ID NO: 24), which may be contacted with non-halogenated tryptophan to form 7-halo-tryptophan; a PrnA halogenase of *Pseudomonas fluorescens* (SEQ.ID NO: 25; SEQ.ID NO: 26), which may be contacted with non-halogenated tryptophan and other psilocybin precursor compounds to form 7-halo-tryptophan or 2-halo-tryptophan and other 7-halo or 2-halo psilocybin precursor compounds; a RebH halogenase of *Lentzea aerocolonigenes* (SEQ.ID NO: 27; SEQ.ID NO: 28), which may be contacted with non-halogenated tryptophan to form 7-halo-tryptophan; and a SatH halogenase of *Streptomyces albus* (SEQ.ID NO: 29), which may be contacted with non-halogenated tryptophan to form 6-halo-tryptophan.

**[00317]** It is noted that halogenases comprising amino acid sequences substantially similar to any of the aforementioned halogenase sequences may be prepared and used to modulate the substrate specificity and/or the enzyme's catalytic efficiency. Thus, for example, halogenases comprising amino acid sequences substantially similar to SEQ.ID NO: 14; SEQ.ID NO: 16; SEQ.ID NO: 18; SEQ.ID NO: 20; SEQ.ID NO: 22; SEQ.ID NO: 24; SEQ.ID NO: 26; SEQ.ID NO: 28, and SEQ.ID NO: 29, may be prepared and used to obtain 2-halo, 5-halo, 6-halo or 7-halo tryptophan, or halogenases comprising amino acid sequences substantially similar to SEQ.ID NO: 14; SEQ.ID NO: 16; SEQ.ID NO: 18; SEQ.ID NO: 20; SEQ.ID NO: 22; SEQ.ID NO: 24; SEQ.ID NO: 26; SEQ.ID NO: 28, and SEQ.ID NO: 29 may be prepared and used to halogenate other non-halogenated psilocybin precursor compounds, such as tryptamine, 4-hydroxyindole, 4-hydroxytryptophan, 4-hydroxytryptamine, norbaeocystin, and baeocystin, for example, to form halogenated forms thereof.

**[00318]** Thus, for example, the native polypeptide sequence of a RebH halogenase *Lentzea aerocolonigenes* (SEQ.ID NO: 27; SEQ.ID NO: 28) may be modulated to thereby modulate the substrate specificity. Thus, for example, a modulated *Lentzea aerocolonigenes* RebH halogenase, which may be referred to as a RebH-Y455W mutant, set forth in SEQ.ID NO: 30 prefers tryptamine over tryptophan as a substrate and when contacted with tryptamine forms 7-halo-tryptamine.

**[00319]** By way of a further example, a modulated *Lentzea aerocolonigenes* RebH halogenase, which may be referred to as a RebH-N470S mutant, set forth in SEQ.ID NO: 31 prefers tryptamine over tryptophan as a substrate and when contacted with tryptamine forms 7-halo-tryptamine.

**[00320]** By way of a further example, a modulated *Lentzea aerocolonigenes* RebH halogenase, which may be referred to as a RebH-8F mutant, set forth in SEQ.ID NO: 32 prefers tryptamine over tryptophan as a substrate and when contacted with tryptamine forms 6-halo-tryptamine.

**[00321]** By way of a further example, a modulated *Lentzea aerocolonigenes* RebH halogenase, which may be referred to as a RebH-10F mutant, set forth in SEQ.ID NO: 33 prefers tryptamine over tryptophan as a substrate and when contacted with tryptamine forms 5-halo-tryptamine.

**[00322]** By way of a further example, a modulated *Lentzea aerocolonigenes* RebH halogenase, which may be referred to as a RebH-3SS mutant, set forth in SEQ.ID NO: 34 has an altered substrate acceptance profile and converts tricyclic tryptoline derivatives.

- 5 **[00323]** By way of yet a further example, a modulated *Lentzea aerocolonigenes* RebH halogenase, which may be referred to as a RebH-3SS mutant set forth in SEQ.ID NO: 35 has an altered substrate acceptance profile and converts a broad range of large indoles and carbazoles.

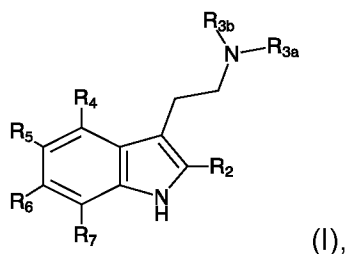
**[00324]** The foregoing examples of halogenases exhibiting modulated  
 10 substrate specificity and/or catalytic efficiency are known to the art and are documented in, for example, Andorfer M.C., *et al.* (2016) Chem. Sci. 7: 3720; Andorfer M.C. *et al.* (2017) Catalysis 7: 1897; Glenn W.S., *et al.* (2011) J. Am. Chem. Soc. 133: 19346; and Payne J.T. *et al.* (2015) Angew. Chem. Int. Ed. 54: 4226.

- 15 **[00325]** Accordingly, in one embodiment the halogenase can be encoded by a nucleic acid selected from:

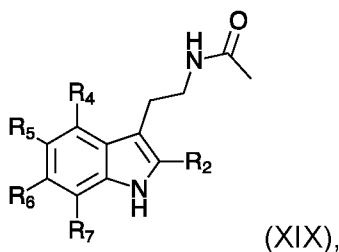
- (a) SEQ.ID NO: 13, SEQ.ID NO: 15, SEQ. ID NO: 17, SEQ.ID NO: 19, SEQ. ID NO: 21, SEQ.ID NO: 23, SEQ.ID NO: 25, and SEQ.ID NO: 27;
- (b) a nucleic acid sequence that is substantially identical to any one of  
 20 the nucleic acid sequences of (a);
- (c) a nucleic acid sequence that is substantially identical to any one of the nucleic acid sequences of (a) but for the degeneration of the genetic code;
- (d) a nucleic acid sequence that is complementary to any one of the  
 25 nucleic acid sequences of (a);
- (e) a nucleic acid sequence encoding a polypeptide having any one of the amino acid sequences set forth in SEQ.ID NO: 14, SEQ.ID NO: 16, SEQ. ID NO: 18, SEQ.ID NO: 20, SEQ. ID NO: 22, SEQ.ID NO: 24, SEQ.ID NO: 26, SEQ.ID NO: 28, SEQ.ID NO: 29, SEQ.ID NO: 30, SEQ.ID NO: 31, SEQ.ID NO: 32. SEQ.ID NO: 33, SEQ.ID NO: 34, and SEQ.ID NO: 35;  
 30
- (f) a nucleic acid sequence that encodes a functional variant of any one of the amino acid sequences set forth in SEQ.ID NO: 14, SEQ.ID NO: 16, SEQ. ID NO: 18, SEQ.ID NO: 20, SEQ. ID NO: 22, SEQ.ID NO: 24, SEQ.ID

NO: 26, SEQ.ID NO: 28, SEQ.ID NO: 29, SEQ.ID NO: 30, SEQ.ID NO: 31, SEQ.ID NO: 32. SEQ.ID NO: 33, SEQ.ID NO: 34, and SEQ.ID NO: 35; and  
 (g) a nucleic acid sequence that hybridizes under stringent conditions to any one of the nucleic acid sequences set forth in (a), (b), (c), (d), (e) or (f).

- 5 **[00326]** In one further embodiment, the host cell can comprise an acetyl transferase capable of acetylating R<sub>3a</sub> or R<sub>3b</sub>, wherein the cell comprises a compound having chemical formula (I):



- 10 wherein, at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> is a halogen atom, wherein each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group, wherein R<sub>4</sub> when it is not halogenated is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group, and wherein at least one of R<sub>3A</sub> and R<sub>3B</sub> is a hydrogen atom, to form a compound having the chemical formula (XIX):



- 15 wherein, at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> is a halogen atom, wherein each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group, wherein R<sub>4</sub> when it is not halogenated is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.

- 20 **[00327]** In some embodiments, R<sub>3a</sub> and R<sub>3b</sub> are each a hydrogen atom.

- [00328]** Suitable acyl transferases that may be used in accordance herewith include a PsmF acetyl transferase obtained from *Streptomyces griseofuscus*, including a PsmF N-acetyltransferase comprising polypeptide sequence SEQ.ID NO: 63 set forth herein, and further including a PsmF polypeptide encoded by a  
 25 nucleic acid selected from:

- (a) SEQ.ID NO: 62;
- (b) a nucleic acid sequence that is substantially identical the nucleic acid sequence of (a);
- (c) a nucleic acid sequence that is substantially identical to the nucleic acid sequence of (a) but for the degeneration of the genetic code;
- (d) a nucleic acid sequence that is complementary to the nucleic acid sequence of (a);
- (e) a nucleic acid sequence encoding a polypeptide having an amino acid sequences set forth in SEQ.ID NO: 63;
- (f) a nucleic acid sequence that encodes a functional variant of the amino acid sequence set forth in SEQ.ID NO: 63; and
- (g) a nucleic acid sequence that hybridizes under stringent conditions to any one of the nucleic acid sequences set forth in (a), (b), (c), (d), (e) or (f).

**[00329]** It will be clear to those of skill in the art that a significant variety of different halogenated and non-halogenated psilocybin precursor compounds may be selected. **FIGS. 8A** and **8B** in this respect provide guidance and allow a person of skill in the art to select appropriate psilocybin precursor compounds and a matching a psilocybin biosynthetic enzyme complement.

**[00330]** Upon production by the host cells of the halogenated psilocybin compounds in accordance with the methods of the present disclosure, the halogenated psilocybin compounds may be extracted from the host cell suspension, and separated from other constituents within the host cell suspension, such as media constituents and cellular debris. Separation techniques will be known to those of skill in the art and include, for example, solvent extraction (e.g. butane, chloroform, ethanol), column chromatography based techniques, high-performance liquid chromatography (HPLC), for example, and/or countercurrent separation (CCS) based systems. The recovered halogenated psilocybin compounds may be obtained in a more or less pure form, for example, a preparation of halogenated psilocybin compounds of at least about 60% (w/w), about 70% (w/w), about 80% (w/w), about 90% (w/w), about 95% (w/w), about 96% (w/w), about 97% (w/w), about 98% (w/w) or about 99% (w/w) purity may be obtained. Thus, in this manner, halogenated psilocybin derivatives in more or less pure form may be prepared.

**[00331]** Similarly, other methods of making the halogenated psilocybin compounds that may be used in accordance herewith may yield preparations of halogenated compounds of at least about 60% (w/w), about 70% (w/w), about 80% (w/w), about 90% (w/w), about 95% (w/w), about 96% (w/w), about 97% (w/w),  
5 about 98% (w/w), or about 99% (w/w) purity.

**[00332]** It will now be clear from the foregoing that novel halogenated psilocybin derivatives are disclosed herein. The halogenated psilocybin compounds may be formulated for use as a pharmaceutical drug or recreational drug., The halogenated psilocybin compounds may also be used as a feedstock  
10 to produce other psilocybin derivatives.

### **SUMMARY OF SEQUENCES**

**[00333]** SEQ.ID NO: 1 sets forth a *Psilocybe cubensis* nucleic acid sequence  
15 encoding a PsiD polypeptide.

**[00334]** SEQ.ID NO: 2 sets forth a deduced amino acid sequence of a *Psilocybe cubensis* PsiD polypeptide.

**[00335]** SEQ.ID NO: 3 sets forth a *Psilocybe cubensis* nucleic acid sequence encoding a PsiH polypeptide.

20 **[00336]** SEQ.ID NO: 4 sets forth a deduced amino acid sequence of a *Psilocybe cubensis* PsiH polypeptide.

**[00337]** SEQ.ID NO: 5 sets forth a *Psilocybe cubensis* nucleic acid sequence encoding a PsiK polypeptide.

25 **[00338]** SEQ.ID NO: 6 sets forth a deduced amino acid sequence of a *Psilocybe cubensis* PsiK polypeptide.

**[00339]** SEQ.ID NO: 7 sets forth a *Psilocybe cubensis* nucleic acid sequence encoding a PsiM polypeptide.

**[00340]** SEQ.ID NO: 8 sets forth a deduced amino acid sequence of a *Psilocybe cubensis* PsiM polypeptide.

30 **[00341]** SEQ.ID NO: 9 sets forth a *Psilocybe cubensis* nucleic acid sequence encoding a Psi-ncAAAD polypeptide.

**[00342]** SEQ.ID NO: 10 sets forth a deduced amino acid sequence of a *Psilocybe cubensis* Psi-ncAAAD polypeptide.

- [00343]** SEQ.ID NO: 11 sets forth a *Psilocybe cubensis* nucleic acid sequence encoding a TrpB polypeptide.
- [00344]** SEQ.ID NO: 12 sets forth a deduced amino acid sequence of a *Psilocybe cubensis* TrpB polypeptide.
- 5 **[00345]** SEQ.ID NO: 13 sets forth a *Streptomyces rugosporus* nucleic acid sequence encoding a PyrH halogenase polypeptide.
- [00346]** SEQ.ID NO: 14 sets forth a deduced amino acid sequence of a *Streptomyces rugosporus* PyrH halogenase polypeptide.
- [00347]** SEQ.ID NO: 15 sets forth a *Kutzneria* sp. 744 nucleic acid sequence  
10 encoding a KtzR halogenase polypeptide.
- [00348]** SEQ.ID NO: 16 sets forth a deduced amino acid sequence of a *Kutzneria* sp. 744 KtzR halogenase polypeptide.
- [00349]** SEQ.ID NO: 17 sets forth a *Streptomyces toxytrinitici* nucleic acid sequence encoding a SttH halogenase polypeptide.
- 15 **[00350]** SEQ.ID NO: 18 sets forth a deduced amino acid sequence of a *Streptomyces toxytrinitici* SttH halogenase polypeptide.
- [00351]** SEQ.ID NO: 19 sets forth a *Streptomyces albogriseolus* nucleic acid sequence encoding a Thal halogenase polypeptide.
- [00352]** SEQ.ID NO: 20 sets forth a deduced amino acid sequence of a  
20 *Streptomyces albogriseolus* Thal halogenase polypeptide.
- [00353]** SEQ.ID NO: 21 sets forth a nucleic acid sequence encoding a BorH halogenase polypeptide of an unidentified soil bacterium.
- [00354]** SEQ.ID NO: 22 sets forth a deduced amino acid sequence of a BorH halogenase polypeptide of an unidentified soil bacterium.
- 25 **[00355]** SEQ.ID NO: 23 sets forth a *Kutzneria* sp. 744 nucleic acid sequence encoding a KtzQ halogenase polypeptide.
- [00356]** SEQ.ID NO: 24 sets forth a deduced amino acid sequence of a *Kutzneria* sp. 744 KtzQ halogenase polypeptide.
- [00357]** SEQ.ID NO: 25 sets forth a *Pseudomonas fluorescens* nucleic acid  
30 sequence encoding a PrnA halogenase polypeptide.
- [00358]** SEQ.ID NO: 26 sets forth a deduced amino acid sequence of a *Pseudomonas fluorescens* PrnA halogenase polypeptide.
- [00359]** SEQ.ID NO: 27 sets forth a nucleic acid sequence encoding a *Lentzea areocolonigenes* RebH halogenase polypeptide.

- [00360]** SEQ.ID NO: 28 sets forth a deduced amino acid sequence of a *Lentzea aerocoligines* RebH halogenase polypeptide.
- [00361]** SEQ.ID NO: 29 sets forth a deduced amino acid sequence of a *Streptomyces albus* SatH halogenase polypeptide.
- 5 **[00362]** SEQ.ID NO: 30 sets forth a deduced amino acid sequence of a mutated *Lentzea aerocoligines* RebH halogenase polypeptide.
- [00363]** SEQ.ID NO: 31 sets forth a deduced amino acid sequence of another mutated *Lentzea aerocoligines* RebH halogenase polypeptide.
- [00364]** SEQ.ID NO: 32 sets forth a deduced amino acid sequence of another  
10 mutated *Lentzea aerocoligines* RebH halogenase polypeptide.
- [00365]** SEQ.ID NO: 33 sets forth a deduced amino acid sequence of another mutated *Lentzea aerocoligines* RebH halogenase polypeptide.
- [00366]** SEQ.ID NO: 34 sets forth a deduced amino acid sequence of another mutated *Lentzea aerocoligines* RebH halogenase polypeptide.
- 15 **[00367]** SEQ.ID NO: 35 sets forth a deduced amino acid sequence of another mutated *Lentzea aerocoligines* RebH halogenase polypeptide.
- [00368]** SEQ.ID NO: 36 sets forth a nucleic acid sequence encoding a *Bacillus atrophaeus* BaTDC tryptophan decarboxylase polypeptide.
- [00369]** SEQ.ID NO: 37 sets forth a deduced amino acid sequence of a  
20 *Bacillus atrophaeus* BaTDC tryptophan decarboxylase polypeptide.
- [00370]** SEQ.ID NO: 38 sets forth a nucleic acid sequence encoding a *Ruminococcus gnavus* RgnTDC tryptophan decarboxylase polypeptide.
- [00371]** SEQ.ID NO: 39 sets forth a deduced amino acid sequence of a *Ruminococcus gnavus* RgnTDC tryptophan decarboxylase polypeptide.
- 25 **[00372]** SEQ.ID NO: 40 sets forth a nucleic acid sequence encoding a *Streptomyces griseofuscus* SgTDC tryptophan decarboxylase polypeptide.
- [00373]** SEQ.ID NO: 41 sets forth a deduced amino acid sequence of a *Streptomyces griseofuscus* SgTDC tryptophan decarboxylase polypeptide.
- [00374]** SEQ.ID NO: 42 sets forth a nucleic acid sequence encoding a  
30 *Ceriporiopsis subvermispora* CsTDC tryptophan decarboxylase polypeptide.
- [00375]** SEQ.ID NO: 43 sets forth a deduced amino acid sequence of a *Ceriporiopsis subvermispora* CsTDC tryptophan decarboxylase polypeptide.
- [00376]** SEQ.ID NO: 44 sets forth a nucleic acid sequence encoding a *Clostridium sporogenes* ClostSporTDC tryptophan decarboxylase polypeptide.



- [00377]** SEQ.ID NO: 45 sets forth a deduced amino acid sequence of a *Clostridium sporogenes* ClostSporTDC tryptophan decarboxylase polypeptide.
- [00378]** SEQ.ID NO: 46 sets forth a nucleic acid sequence encoding a *Homo sapiens* HsDDC DOPA decarboxylase polypeptide.
- 5 **[00379]** SEQ.ID NO: 47 sets forth a deduced amino acid sequence of a *Homo sapiens* HsDDC DOPA decarboxylase polypeptide.
- [00380]** SEQ.ID NO: 48 sets forth a nucleic acid sequence encoding an *Ophiorrhiza pumila* OpTDC tryptophan decarboxylase polypeptide.
- [00381]** SEQ.ID NO: 49 sets forth a deduced amino acid sequence of an
- 10 *Ophiorrhiza pumila* OpTDC tryptophan decarboxylase polypeptide.
- [00382]** SEQ.ID NO: 50 sets forth a nucleic acid sequence encoding an *Oryza sativa* OsTDC1 tryptophan decarboxylase polypeptide.
- [00383]** SEQ.ID NO: 51 sets forth a deduced amino acid sequence of an *Oryza sativa* OsTDC1 tryptophan decarboxylase polypeptide.
- 15 **[00384]** SEQ.ID NO: 52 sets forth a nucleic acid sequence encoding another *Oryza sativa* OsTDC2 tryptophan decarboxylase polypeptide.
- [00385]** SEQ.ID NO: 53 sets forth a deduced amino acid sequence of another *Oryza sativa* OsTDC2 tryptophan decarboxylase polypeptide.
- [00386]** SEQ.ID NO: 54 sets forth a nucleic acid sequence encoding a
- 20 *Xenorhabdus doucetiae* XdTDC decarboxylase polypeptide.
- [00387]** SEQ.ID NO: 55 sets forth a deduced amino acid sequence of a *Xenorhabdus doucetiae* XdTDC decarboxylase polypeptide.
- [00388]** SEQ.ID NO: 56 sets forth a nucleic acid sequence encoding a *Camptotheca acuminata* CaTDC tryptophan decarboxylase polypeptide.
- 25 **[00389]** SEQ.ID NO: 57 sets forth a deduced amino acid sequence of a *Camptotheca acuminata* CaTDC tryptophan decarboxylase polypeptide.
- [00390]** SEQ.ID NO: 58 sets forth a nucleic acid sequence encoding an *Oryctolagus cuniculus* OcINMT indolethylamine N-methyltransferase polypeptide.
- [00391]** SEQ.ID NO: 59 sets forth a deduced amino acid sequence of an
- 30 *Oryctolagus cuniculus* OcINMT indolethylamine N-methyltransferase polypeptide.
- [00392]** SEQ.ID NO: 60 sets forth a nucleic acid sequence encoding a *Hordeum vulgare* HvNMT N-methyltransferase polypeptide.
- [00393]** SEQ.ID NO: 61 sets forth a deduced amino acid sequence of a *Hordeum vulgare* HvNMT N-methyltransferase polypeptide.

- [00394]** SEQ.ID NO: 62 sets forth a nucleic acid sequence encoding a *Streptomyces griseofuscus* PsmF N-acetyltransferase polypeptide.
- [00395]** SEQ.ID NO: 63 sets forth a deduced amino acid sequence of a *Streptomyces griseofuscus* PsmF N-acetyltransferase polypeptide.
- 5 **[00396]** SEQ.ID NO: 64 sets forth a nucleic acid sequence encoding a mutated *Thermotoga maritima* TmTrpB-2F3 tryptophan synthase subunit B polypeptide.
- [00397]** SEQ.ID NO: 65 sets forth a deduced amino acid sequence of a mutated *Thermotoga maritima* TmTrpB-2F3 tryptophan synthase subunit B
- 10 polypeptide.
- [00398]** SEQ.ID NO: 66 sets forth a nucleic acid sequence encoding a mutated *Pyrococcus furiosus* PfTrpB-0A9 tryptophan synthase subunit B polypeptide.
- [00399]** SEQ.ID NO: 67 sets forth a deduced amino acid sequence of a
- 15 mutated *Pyrococcus furiosus* PfTrpB-0A9 tryptophan synthase subunit B polypeptide.
- [00400]** SEQ.ID NO: 68 sets forth a nucleic acid sequence encoding another mutated *Pyrococcus furiosus* PfTrpB-2A6 tryptophan synthase subunit B polypeptide.
- 20 **[00401]** SEQ.ID NO: 69 sets forth a deduced amino acid sequence of a mutated *Pyrococcus furiosus* PfTrpB-2A6 tryptophan synthase subunit B polypeptide.
- [00402]** SEQ.ID NO: 70 sets forth a nucleic acid sequence encoding another mutated *Pyrococcus furiosus* PfTrpB-7E6 tryptophan synthase subunit B
- 25 polypeptide.
- [00403]** SEQ.ID NO: 71 sets forth a deduced amino acid sequence of a mutated *Pyrococcus furiosus* PfTrpB-7E6 tryptophan synthase subunit B polypeptide.
- [00404]** SEQ.ID NO: 72 sets forth a nucleic acid sequence encoding another
- 30 mutated *Pyrococcus furiosus* PfTrpB-Quat tryptophan synthase subunit B polypeptide.
- [00405]** SEQ.ID NO: 73 sets forth a deduced amino acid sequence of a mutated *Pyrococcus furiosus* PfTrpB-Quat tryptophan synthase subunit B polypeptide.

- [00406]** SEQ.ID NO: 74 sets forth a nucleic acid sequence used as integration cassette XII-4::TADH1-PsiH-HA-PPGK1-PTDH3-CPR-c-myc-TCYC1 for homologous recombination-based genome editing of *Saccharomyces cerevisiae* strain Sc-1.
- 5 **[00407]** SEQ.ID NO: 75 sets forth a nucleic acid sequence used as integration cassette XII-5::TADH1-PsiK-V5-PPGK1-PTDH3-PsiM-FLAG-TCYC1 for homologous recombination-based genome editing of *Saccharomyces cerevisiae* strain Sc-1.
- [00408]** SEQ.ID NO: 76 sets forth a nucleic acid sequence of modified pMM1  
 10 vector pMM1-PTDH3-ClostrSporTDC-His-TCYC1 for engineering of *Saccharomyces cerevisiae* strain Sc-1.
- [00409]** SEQ.ID NO: 77 sets forth a nucleic acid sequence for *Saccharomyces cerevisiae* promoter PGK1.
- [00410]** SEQ.ID NO: 78 sets forth a nucleic acid sequence for  
 15 *Saccharomyces cerevisiae* promoter TDH3.
- [00411]** SEQ.ID NO: 79 sets forth a nucleic acid sequence for *Saccharomyces cerevisiae* promoter CLN1.
- [00412]** SEQ.ID NO: 80 sets forth a nucleic acid sequence for *Saccharomyces cerevisiae* promoter UGA1.
- 20 **[00413]** SEQ.ID NO: 81 sets forth a nucleic acid sequence encoding a synthetic HA epitope tag polypeptide
- [00414]** SEQ.ID NO: 82 sets forth deduced amino acid sequence of a synthetic HA epitope tag polypeptide
- [00415]** SEQ.ID NO: 83 sets forth a nucleic acid sequence encoding a  
 25 synthetic c-MYC epitope tag polypeptide
- [00416]** SEQ.ID NO: 84 sets forth deduced amino acid sequence of a synthetic c-MYC epitope tag polypeptide
- [00417]** SEQ.ID NO: 85 sets forth a nucleic acid sequence encoding a synthetic FLAG epitope tag polypeptide
- 30 **[00418]** SEQ.ID NO: 86 sets forth deduced amino acid sequence of a synthetic FLAG epitope tag polypeptide
- [00419]** SEQ.ID NO: 87 sets forth a nucleic acid sequence encoding a synthetic V5 epitope tag polypeptide

- [00420]** SEQ.ID NO: 88 sets forth deduced amino acid sequence of a synthetic V5 epitope tag polypeptide
- [00421]** SEQ.ID NO: 89 sets forth a nucleic acid sequence encoding a synthetic HIS epitope tag polypeptide
- 5 **[00422]** SEQ.ID NO: 90 sets forth deduced amino acid sequence of a synthetic HIS epitope tag polypeptide
- [00423]** SEQ.ID NO: 91 sets forth a nucleic acid sequence of pMM1 vector
- [00424]** SEQ.ID NO: 92 sets forth a nucleic acid sequence of pCDM4 vector
- [00425]** SEQ.ID NO: 93 sets forth a nucleic acid sequence of pETM6 vector
- 10 **[00426]** SEQ.ID NO: 94 sets forth a nucleic acid sequence of pET28a(+)  
vector
- [00427]** SEQ.ID NO: 95 sets forth a nucleic acid sequence encoding a *Psilocybe cubensis* PcCPR cytochrome P450 reductase polypeptide.
- [00428]** SEQ.ID NO: 96 sets forth a deduced amino acid sequence of a  
15 *Psilocybe cubensis* PcCPR cytochrome P450 reductase polypeptide.
- [00429]** SEQ.ID NO: 97 sets forth a nucleic acid sequence of pETM6-H10 vector.
- [00430]** SEQ.ID NO: 98 sets forth a nucleic acid sequence of pRSM3 vector.

## 20 **SEQUENCE LISTING**

- [00431]** SEQ.ID NO: 1
- 25 ATGCAGGTGATACCCGCGTGCAACTCGGCAGCAATAAGATCACTATGTCCTACTCCCGA  
GTCTTTTAGAAACATGGGATGGCTCTCTGTCTCAGCGATGCGGTCTACAGCGAGTTCATAG  
GAGAGTTGGCTACCCGCGCTTCCAATCGAAATTACTCCAACGAGTTCGGCCTCATGCAA  
CCTATCCAGGAATTCAAGGCTTTCATTGAAAGCGACCCGGTGGTGCACCAAGAATTTAT  
TGACATGTTTCGAGGGCATTTCAGGACTCTCCAAGGAATTATCAGGAACTATGTAATATGT  
TCAACGATATCTTTCGCAAAGCTCCCGTCTACGGAGACCTTGGCCCTCCCGTTTATATG  
30 ATTATGGCCAAATTAATGAACACCCGAGCGGGCTTCTCTGCATTCACGAGACAAAGGTT  
GAACCTTCACCTTCAAAAACTTTTCGATACCTGGGGATTGTTCCCTGTCTTCGAAAGATT  
CTCGAAATGTTCTTGTGGCCGACCAGTTCGACGACAGACATTGCGGCTGGTTGAACGAG  
CGGGCCTTGTCTGCTATGGTTAAACATTACAATGGACGCGCATTGATGAAGTCTTCCT  
CTGCGATAAAAATGCCCCATACTACGGCTTCAACTCTTACGACGACTTCTTTAATCGCA  
35 GATTTTCGAAACCGAGATATCGACCGACCTGTAGTCGGTGGAGTTAACAACACCACCCTC  
ATTTCTGCTGCTTGCGAATCACTTTCCTACAACGCTCTTATGACGTCCAGTCTCTCGA  
CACTTTAGTTTTCAAAGGAGAGACTTATTCGCTTAAGCATTTGCTGAATAATGACCCTT  
TCACCCACAAATTCGAGCATGGGAGTATTCTACAAGGATTCTTGAACGTCACCGCTTAC  
CACCGATGGCACGCACCCGTCAATGGGACAATCGTCAAAATCATCAACGTTCCAGGTAC  
40 CTACTTTGCGCAAGCCCCGAGCAGATTGGCGACCTATCCCGGATAACGATTACGACC  
CACCTCCTTACCTTAAGTCTCTTGTCTACTTCTCTAATATTGCCGCAAGGCAAATTATG

TTTATTGAAGCCGACAACAAGGAAATTGGCCTCATTTTCCTTGTGTTTCATCGGCATGAC  
 CGAAATCTCGACATGTGAAGCCACGGTGTCCGAAGGTCAACACGTCAATCGTGGCGATG  
 ACTTGGAATGTTCCATTTTCGGTGGTTCTTCGTTTCGCGCTTGGTCTGAGGAAGGATTGC  
 AGGGCAGAGATCGTTGAAAAGTTCACCGAACCCGGAACAGTGATCAGAATCAACGAAGT  
 5 CGTCGCTGCTCTAAAGGCTTAG

**[00432]** SEQ.ID NO: 2

MQVIPACNSAAIRSLCPTPESFRNMGWLSVSDAVYSEFIGELATRASNRNYSNEFGLMQ  
 10 PIQEFKAFIESDPVVHQEFIDMFEGIQDSPRNYQELCNMFNDIFRKAPVYGDLPVPVYM  
 IMAKLMNTRAGFSAFTRQRLNLHFKKLFDTWGLFLSSKDSRNVLVADQFDDRHCWGLNE  
 RALSAMVKHYNGRAFDEVFLCDKNAPYYGFNSYDDFFNRRFRNRDIDRPVVGGVNNTTL  
 ISAACESLSYNVSYDVQSLDTLVFKGETYSLKHLNNDPFTPQFEHGSILQGFLNVTAY  
 HRWHAPVNGTIVKIIINVPGTYFAQAPSTIGDPIPDNDYDPPPYLKSLVYFSNIAARQIM  
 15 FIEADNKEIGLIFLVFIGMTEISTCEATVSEGQHVNRGDDLGMFHFGGSSFALGLRKDC  
 RAEIVEKFTEPGTVIRINEVVAALKA

**[00433]** SEQ.ID NO: 3

ATGATCGCTGTACTATTCTCCTTCGTCATTGCAGGATGCATATACTACATCGTTTCTCG  
 TAGAGTGAGGCGGTGCGCTTGCCACCAGGGCCGCTGGCATTCCATTCCCTTCATTG  
 GGAACATGTTTGATATGCCTGAAGAATCTCCATGGTTAACATTTCTACAATGGGGACGG  
 GATTACAGTCTGTCTTGCCGCGTTGACTTCTAATATATGAACAGCTAATATATTGTCAG  
 ACACCGATATTCTCTACGTGGATGCTGGAGGGACAGAAATGGTTATTCTTAACACGTTG  
 25 GAGACCATTACCGATCTATTAGAAAAGCGAGGGTCCATTTATTCTGGCCGGTGAGCTGA  
 TGTGAGTTTTTTGCAATTGAATTTGTGGTCACACGTTTCCAGACTTGAGAGTACAATG  
 GTCAACGAACCTTATGGGGTGGGAGTTTGAAGTTAGGGTTCATCACATACGGCGACAGGTG  
 GCGCGAAGAAAGGCGCATGTTCCGCAAGGAGTTCAGTGAGAAGGGCATCAAGCAATTC  
 GCCATGCTCAAGTGAAAGCTGCCCATCAGCTTGTCCAACAGCTTACCAAAACGCCAGAC  
 30 CGCTGGGCACAACATATTCGCCAGTAAGTACTACTTGAGGAAAATAGCGTACGCTTCGC  
 TGACCGGTCCGTACATCAAAGTCAGATAGCGGCAATGTCACTGGATATTGGTTATGGAA  
 TTGATCTTGAGAAGACGACCCCTTGCTGGAAGCGACCCATTTGGCTAATGAAGGCCCTC  
 GCCATAGCATCAGTGCCGGGCAAATTTTGGGTGATTCGTTCCCTTCTCGTGAGCATCC  
 TTCTTCTATGTAGGAAGGGAAGGAGTCTAACAAGTGTTAGTAAAATACCTTCTGCTTG  
 35 GTTCCCAGGTGCTGTCTTCAAGCGCAAAGCGAAGGTCTGGCGAGAAGCCGCCGACCATA  
 TGTTTGACATGCCTTATGAAACTATGAGGAAATTAGCAGTTAGTCAAATGCGTTCTCCC  
 CGTATTTTTTTCAATACTCTAACTTCAGCTCACAGCCTCAAGGATTGACTCGTCCGTCTG  
 ATGCTTCAGCTCGTCTGCAAGCCATGGATCTCAACGGTGACCTTGAGCATCAAGAACAC  
 GTAATCAAGAACACAGCCGCAGAGGTTAATGTGCGTAAGTCAAAGCGTCCGTGCGCAA  
 40 TTCAAATTCAGGCGCTAAAGTGGGTCTTCTACCAAGGTGGAGGCGATACTGTAAGGA  
 TTTCTCAATCGTTAGAGTATAAGTGTCTAATGCAGTACATACTCCACCAACCAGACTG  
 TCTCTGCTATGTCTGCGTTCATCTTGCCCATGGTGAAGTACCCTGAGGTCCAGCGAAAG  
 GTTCAAGCGGAGCTTGATGCTCTGACCAATAACGGCCAAATTCCTGACTATGACGAAGA  
 AGATGACTCCTTGCCATACCTCACCGCATGTATCAAGGAGCTTTTCCGGTGGAATCAAA  
 45 TCGCACCCCTCGCTATACCGCACAAATTAATGAAGGACGACGTGTACCGCGGGTATCTG  
 ATTCCCAAGAACACTCTAGTCTTCGCAAAACCTGGTGAGGCTGTCCATTCAATTCCTAG  
 TACATCCGTTGCCCCACTAATAGCATCTTGATAACAGGGCAGTATTAACGATCCAGAA  
 GTCTATCCAGATCCCTCTGTGTTCCGCCCAGAAAGATATCTTGGTCTTGACGGGAAGCC  
 TGATAACACTGTACGCGACCCACGTAAAGCGGCATTTGGCTATGGACGACGAAATTTGGT  
 50 AAGTGCGCTTTTCAAGACCCCCCTTCCGTTGACTAGTGCCATGCGCGCATACAATATCG  
 CTATTGATCTGATATAACTTCCCTGCGGCATTTATTTTGGCATTCCTTTAGTCCCGGAA

5 TTCATCTAGCGCAGTCGACGGTTTGGATTGCAGGGGCAACCCTCTTATCAGCGTTCAAT  
 ATCGAGCGACCTGTCGATCAGAATGGGAAGCCCATTGACATACCGGCTGATTTTACTAC  
 AGGATTCTTCAGGTAGCTAATTCCGTCTTTGTGTGCATAATACCCCTAACGACGCACG  
 TTTACCTTTTTGTAAAGACACCCAGTGCCTTTCCAGTGCAGGTTTGTTCCTCGAACAGA  
 GCAAGTCTCACAGTCGGTATCCGGACCCTGA

**[00434]** SEQ.ID NO: 4

10 MIAVLFSFVIAGCIYYIVSRRVRRSRLPPGPPGIPIPFIGNMFDMPPEESPWLTFLOWGR  
 DYNTDILYVDAGGTEMVILNTLETITDLLEKRSIYSGRLESTMVNELMGWEFDLGFIT  
 YGDRWREERRMFAKEFSEKGIKQFRHAQVKAHQVQQLTKTPDRWAQHHRHQAAMSL  
 DIGYGIDLAEDDPWLEATHLANEGLAIASVPGKFWVDSFPSLKYLPAWFFGAVFKRKAK  
 VWREAADHMDMPYETMRKLAPQGLTRPSYASARLQAMDNLNGDLEHQEHVIKNTAAEVN  
 15 VGGGDTTVSAMSAFILAMVKYPEVQRKVQAEALDALTNNQIPDYDEEDDSLPLYLTACIK  
 ELFRWNQIAPLAIIPHKLMKDDVYRGYLIIPKNTLVFANTWAVLNDPEVYPDPSPVFRPERY  
 LGPDGKPDNTVRDPRKAAFYGRNCPGIHLAQSTVWIAGATLLSAFNIERPVDQNGKP  
 IDIPADFTTGFFRHPVPFQCRFVPRTEQVSQSVSGP

**[00435]** SEQ.ID NO: 5

20 ATGGCGTTCGATCTCAAGACTGAAGACGGCCTCATCACATATCTCACTAAACATCTTTC  
 TTTGGACGTCGACACGAGCGGAGTGAAGCGCCTTAGCGGAGGCTTTGTCAATGTAACCT  
 GGCGCATTAAGCTCAATGCTCCTTATCAAGGTCATACGAGCATCATCCTGAAGCATGCT  
 CAGCCGCACATGTCTACGGATGAGGATTTTAAGATAGGTGTAGAACGTTTCGTTTACGA  
 25 ATACCAGGCTATCAAGCTCATGATGGCCAATCGGGAGGTTCTGGGAGGCGTGGATGGCA  
 TAGTTTCTGTGCCAGAAGGCCTGAACTACGACTTAGAGAATAATGCATTGATCATGCAA  
 GATGTCGGGAAGATGAAGACCCTTTTAGATTATGTCACCGCCAAACCGCCACTTGCGAC  
 GGATATAGCCCGCCTTGTTGGGACAGAAATTGGGGGGTTCGTTGCCAGACTCCATAACA  
 TAGGCCGCGAGAGGCGAGACGATCCTGAGTTCAAATTCCTTCTCTGGAAATATTGTGCGA  
 30 AGGACGACTTCAGACCAGCTGTATCAAACCATCATACCCAACGCAGCGAAATATGGCGT  
 CGATGACCCCTTGCTGCCTACTGTGGTTAAGGACCTTGTGGACGATGTCATGCACAGCG  
 AAGAGACCCTTGTATGGCGGACCTGTGGAGTGGAAATATTCTTCTCCAGTTGGAGGAG  
 GGAAACCCATCGAAGCTGCAGAAGATATATATCCTGGATTGGGAACTTTGCAAGTACGG  
 CCCAGCGTCGTTGGACCTGGGCTATTTCTTGGGTGACTGCTATTTGATATCCCGCTTTC  
 35 AAGACGAGCAGGTCGGTACGACGATGCGGCAAGCCTACTTGCAAAGCTATGCGCGTACG  
 AGCAAGCATTTCGATCAACTACGCCAAAGTCACTGCAGGTATTGCTGCTCATATTGTGAT  
 GTGGACCGACTTTATGCAGTGGGGGAGCGAGGAAGAAAGGATAAATTTTGTGAAAAAGG  
 GGGTAGCTGCCTTTCACGACGCCAGGGGCAACAACGACAATGGGGAAATTACGTCTACC  
 TTACTGAAGGAATCATCCACTGCGTAA

**[00436]** SEQ.ID NO: 6

45 MHIRNPYRTPIDYQALSEAFPPLKPFVSVNADGTSSVDLTIPQAQRAFTAALLHRDFGL  
 TMTIPEDRLCPTVPNRLNYVLWIEDIFNYTNKTLGLSDDRPIKGV DIGTGASAIYPLA  
 CARFKAWSMVGTEVERKCIDTARLNVVANLQDRLSILETSIDGPILVPIFEATEEY  
 EFTMCNPPFYDGAADMQTSDAAGFGFGVVGAPHS GTVIEMSTEGGESAFVAQMVRSLK  
 LRTRCRWYTSNLGKLKSLKEIVGLLKELEISNYAINEYVQGSTRRYAVAWSFTDIQLPE  
 ELSRPSNPELSSLF

50

**[00437]** SEQ.ID NO: 7

5 ATGCATATCAGAAATCCTTACCGTACACCAATTGACTATCAAGCACTTTTCAGAGGCCTT  
 CCCTCCCCTCAAGCCATTTGTGTCTGTCAATGCAGATGGTACCAGTTCTGTTGACCTCA  
 CTATCCCAGAAGCCCAGAGGGCGTTCACGGCCGCTCTTCTTCATCGTGAAGTTTCGGGCTC  
 ACCATGACCATAACCAGAAGACCGTCTGTGCCAACAGTCCCCAATAGGTTGAACTACGT  
 TCTGTGGATTGAAGATATTTTCAACTACACGAACAAAACCCTCGGCCTGTGGATGACC  
 10 GTCCTATTAAAGGCGTTGATATTGGTACAGGAGCCTCCGCAATTTATCCTATGCTTGCC  
 TGTGCTCGGTTCAAGGCATGGTCTATGGTTGGAACAGAGGTCGAGAGGAAGTGCATTGA  
 CACGGCCCGCTCAATGTCGTCGCGAACAATCTCCAAGACCGTCTCTCGATATTAGAGA  
 CATCCATTGATGGTCCTATTCTCGTCCCATTTTCGAGGCGACTGAAGAATACGAATAC  
 GAGTTTACTATGTGTAACCCTCCATTCTACGACGGTGTGCGGATATGCAGACTTCGGA  
 TGCTGCCAAAGGATTGGATTGGCGTGGGCGCTCCCATTCTGGAACAGTCATCGAAA  
 15 TGTCGACTGAGGGAGGTGAATCGGCTTTCGTCGCTCAGATGGTCCGTGAGAGCTTGAAG  
 CTTGCAACACGATGCAGATGGTACACGAGTAACTGGGAAAGCTGAAATCCTTGAAAGA  
 AATAGTGGGGCTGCTGAAAGAACTTGAGATAAGCAACTATGCCATTAACGAATACGTTT  
 AGGGGTCCACACGTCGTTATGCCGTTGCGTGGTCTTTCACTGATATTCAACTGCCTGAG  
 GAGCTTTCTCGTCCCTCTAACCCCGAGCTCAGCTCTCTTTTCTAG

20 **[00438]** SEQ.ID NO: 8

25 MHIRNPYRTPIDYQALSEAFPPLKPFVSVNADGTSSVDLTIPAEQRAFTAALLHRDFGL  
 TMTIPEDRLCPTVPNRLNYVLWIEDIFNYTNKTLGLSDDRPIKGVDIGTGASAIYPMLA  
 CARFKAWSMVGTEVERKCIDTARLNVVANNLQDRLSILETSIDGPILVPIFEATEEY  
 EFTMCNPPFYDGAADMQTSDAAGKGFVGVGAPHSGTVIEMSTEGGESAFVQMVRESLK  
 LRTRCRWYTSNLGKLKSLKEIVGLLKELEISNYAINEYVQGSTRRYAVAWSFTDIQLPE  
 ELSRPSNPELSSLF

30 **[00439]** SEQ.ID NO: 9

30 ATGCCTTCCAGTCACCCTCACATTACTCATCGCTATCGGGTTCCTTCGAGTGACGACCA  
 TGAACGTATATCTGCTCTGTTCTTGGGTCCCAAAGCAGAAAATGCCGCATTTCTCCAGC  
 AATGGTTGACCACGGTCGTCGCACAGCAAAAGGTGCCCGCGATGCATACTTCCCGGAT  
 GACAATGCTTTTATTACTACAGACATGCAAACTTCCCCCGCCTTGTCTCAGACTACTAA  
 35 AGTAATCGCCTCCAATCTCACCGAATTATTGACTGCACTCGGTGAAAGGTGATTCTCTT  
 TCTTCTCACCTCGGTACAGCGGCCATATGTCTGTGGACCAAAGTCTACCTGCCATTCTC  
 GGATTCTTATCGACCACATTTTATAATCCTAACAATGTTGCCTTCGAGGCTAGTCCATT  
 CACGACCTCATCGAGGAAGAAGTTGGCTTGCAACTCTCTGAAATGCTGGGTTATAATC  
 GGCTAAATAACACCGAGAAACCTCTCGCTGGGGACATATTGCATCAGGTGGAAGTGT  
 40 GCAAACTTGGAAGCGATGTGGGCGGCGGAAACCTCAAGTTTACCCTCTCTCACTCCG  
 TGATGCTTCAGCCGAAGGCGCAGAGATGGAATTCATTCGTGACACATTCTCCGTCAAAA  
 CCTGTGTGGTGACAAAAAATTATTAAAGGATTGCAGCCCATGGGAACTCCTCAATTG  
 CATGTTTCTACTATCTTAGACATGCCCGACCGTCTGCACGACGAGTACAATATTTACC  
 TCAGTTCCTCGAAAAGGTTATGCGAAAGTATATCATCCAGTCTACCAACAAAGACACGT  
 45 TGATGCAGCGTTGGGGACTTACCCAACAACCTGTGCTTTTATCCCCGAGCACAAACCAT  
 TATTCTGGCCAAAGGCTGCAGCTGTGCTCGGTATTGGCTCAGACAACCTTCGCAACGT  
 CCCAGTAGACATCCAAGCCCACATGGACATAAACGAACTCGATCGTATGTTAAAAATTT  
 GCTTGGACGAGGAGACGCCAGTATATCAAGTAGTTGCTGTTATCGGTACCACCGAAGAG  
 GCGGGTGTGATCGCATTACGGAGATCCTGAAGCTGCGCCAAAAGTATGAAGCTTTGGG  
 50 GCTGTCTTTGCCATCCATGCAGATGCTGCTTGGGGAGGCTATTTTGCAACCATGCTAC  
 CCAAAGATACATTGGGTGAAACCGGACTAGGCTTCCCAAAGAGGACACTACCTCGGGC

TTTGTCCCTCACGTCCGGTCTGCGCGAGGAGAGCGCGTTACAACCTCAGCCATATAAAGTA  
 TGCCGATTCTATTACTATCGACCCGCACAAGGCAGGCTATGTTTCCTTACCCCGCTGGGG  
 CACTCTGTTATCGCGACGGAAGAATGAGGTACCTGCTTACATGGTCCGCGCCCTACCTT  
 GCCCAAGGCAACGAGGGCCAAAGTATCGGAATATACGGGATCGAAGGAAGCAAACCTGG  
 5 TGCAGCAGCATCCGCGGTATTTCATGGCGCACGAAACCATTGGCCTGACTCCTTCTGGAT  
 ACGGGAACCTTCTTGGCCAGGCAATGTTTACATGTGCGCGATACGCTGCTCACTGGTCT  
 GCAATGTCAACGGATACTACCAGTTTCACTGTCAACCCCGTTCAATCCTATCCCTGCTGA  
 CATCGACCCCAACGCTGACCCCGCAAAGGTCGAAGAGCAAAAACAGTTCATCAGAGATC  
 GTATCTTGTTCAAATCGAACGAGGAAATATACAACGATTCTGAGGCTATGGAACCTCTTG  
 10 CACCAACTTGGGTCCGATCTCAATATCAACGTTTTTCGCATGCAACTTCCGCGACCGCGA  
 TAATAATCTCAACACCGACGTCGAGGAAGCCAACTGGCTCAATAACCGTATTTTCCAAC  
 GCTTTTCTGTTACAAGTGCTGAGGAGAACCCATTGGAAACGCCATTCTTCTCAGCTCA  
 ACTACATTGAAACAATCCGAATACGGCGTCTGCGCAACCGAAGTAAAGAGACGCATGGG  
 ACTTGTTGGTGACCAGGATGTTATAGTCTGAGGAACGTCGTTATGTCTCCATTTACTA  
 15 CAACGAACGACTTTGTGGGAACCTCTGGCAAACACCTTCCAAAAGATCGTTGAGGAGGAG  
 GTCGAGTATGCACGGATCCGCAACGATATGAAACCTAGCATTACACCTTCCTTCTTCA  
 TGGTTTCAGGAGAGCAATACTATCTTGTCCACACCCCAACGATCCATATGGCCAGCGGCC  
 GTCGCCAAATCATCCTTTTCTAGTAAATGTTGAAGGCCAAGTTCGGCAGGCGATACATGCC  
 CATGAAAGAGTTGAAGCAGTGATTGTACATAACACTGTGCCCCCTCCGCCTTGACGAAAT  
 20 CGTTGACGGAGGATCATTTGACGGCATACTCACCATCGGAAAGAGGAAAAGTAGTTTCA  
 AAGTGAAGATTTCAAACATTAAAGTAGTCAAGAAGCGCTCTCTGATGACTGAGGACCTG  
 GAATCTGCGTACCCATCGTTGATGCCATTCTATTTCTACGGGACTCAAGGACACGCTCA  
 TCTCGACCATGTCACTACTGTCGTTCTAACATCCATCTGAGTGCTGGCGAAATACAGT  
 ACAAATTCGACGACGAGGTGTCAAGCGAGGACCTCGCCAAGGGCCTCATTGTTGTTGCT  
 25 GAGAACGTACACGAGGCATCCATGCAGCCCTTCCCGCTCATGAAAGATTTCAAGATCAC  
 CAACCAATTCTTCTTCTCAGCTCCGGGCAAATACTCCGCGTCAAAGTGACAGAGATCCAT  
 ACCCGGCATCGACAATGGATCCCATCCCTCTCCACGACATCAAGAACCAGCCCGTCGTG  
 ACACAAGGCACCATCACGCTCGTCGGAAATATTTACGTCGATTCTGATGCGCTCAACGT  
 CGCTTCCGAGCCTACTGCCGACGAAGACGCGGCGCATGTTTCTCACGCTCGCAACATGT  
 30 ACGGCGAGATGACCGCTGGAACGATCAAAGGCTGGCAAACGCTGTTTCGTCAATTTCCAC  
 AACAAATTGGAGACTGTTGCTCCGACGAAGTAG

**[00440]** SEQ.ID NO: 10

35 MPSSHPHITHRYRVPSSDDHERISALFLGPKAENAAFLQQWLTTVVAQQKAARDAYFPD  
 DNAFITTTDMQTS PAF AQTTKVIASNLTELLTALGERSIPFFSPRYS GHMSVDQSLPAIL  
 GFLSTTFYNPNNVAFEAS PFTT LIEEEVGLQLSEMLGYNRLNNT EKPLAWGH IASGGTV  
 ANLEAMWAARNLKFYPLSLRDASAEGAEMEFIRD TFSVKTCVGD KKLKDCSPWELLNL  
 HVSTILDMPDR LHDEYNISPQFLEKVMRKYIIQSTNKDTLMQRWGLTQQPVVLS PSTNH  
 40 YSWPKAAAVLGIGSDNLRNVPVDIQA HMDINELDRMLKICLDEETPVYQV VAVIGTTEE  
 GGVDRITEILKLRQKY EALGLSFAIHADA AWGGYFATMLPKDTLGRNRTRL PKEDTTSG  
 FVPHVGLREESALQLSHIKYADSITIDPHKAGYVPYPAGALCYRDGRMRYLLTWSAPYL  
 AQNEGQSIGIYGIEGSKPGAAASAVFMAHETIGLTPSGYGNLLGQAMFTCRRYAAHWS  
 AMSTDTTSTFTVTFPNPIPADIDPNADPAKVEEQKFIRDRI LFKSNEEIYNDSEAMELL  
 45 HQLGSDLNINVFACNFRDRDNNLNTDVEE ANWLN NRIFQRFSVTS AEENPLETPFFLSS  
 TTLKQSEYGV CATEVKRRMGLVGDQDVIVLRNVVMS PFTTTNDFVGT LANTFQKIVEEE  
 VEYARIRNDMKPSIHTFLLHGSGEQYYLVHTPTIHMASGRRQIILSVNVEGQVRQAIHA  
 HERVEAVIVHNTVPLRLDEIVDGG SFDGILTIGKRKTSFKVKISNIKVVKKRSLMTEDL  
 ESAYPSLMPFYFYGTQGH AHL DHVITVVPNIHLSAGEIQYKFDDEVSS EDLAKGLIVVA  
 50 ENVHEASMQPFPLMKDFKITNQFFFSSGQILRVKVYRDPYPASTMDPIPLHDIKNQPVV



TQGTITLVGNIYVDS DALNVASEPTADEDAAHVPHARNMYGEMTAGTIKGWQNAVRHFH  
NKLETVAPTK

**[00441]** SEQ.ID NO: 11

5 ATGGAGGCTATCAAAAAGGTTTTTGTGAGAACAAAAAGGCGGAGGGCATTCTGTGTTGGT  
GACCTTTGTTACTGCAGGATATCCTCGTCCCGAAGATACTGTTCCCATCTTGCTGGCCA  
TGGAGGCCGGTGGTGCTGATATCATCGAGCTTGGTATGCCATTTTCAGACCCAATTGCA  
10 GATGGTCTGTCTATCCAGGAAACGAACACAATCGCCGTTGCAAACCAGGTAGATTATAC  
CACTGTTCTCGGACAACCTTCGGAAGCCCGCAAACAAGGGCTCAAGGCACCCGTTCTTC  
TGATGGGATATTATAACCCCATATTGGCTTACGGAGAAGACAGATCTATTCAAGATGCG  
GCTGAAGCTGGAGCCAATGGGTTTATTATGGTCGACCTTCCACCCGAGGAGGCTGTGCG  
TTTTTCGAGAGAAATGTATCAAATCCAACCTCTCATATGTTCTCTAATTGCACCCTCAA  
15 CGACTCTGTGCGGTATAAAGTTCCTCTCAACAATTGCAGACACGTTCTATGTGCTG  
TCTAAAATGGGAACCACCGGATCCTCAGAGAAGGTTGCCATGAATAACGCCCTTCCCAC  
CATCATCGATCGTATTTCGCGAGTACGCTGAAGTTCCTTTAGCAGTCGGATTTGGAGTCG  
CCACTCGGGCTCACTTCAACTACGTCGCCGATTCCGGTGCTGATGGTGTGCTTATTGGC  
ACCAAACTCGTTAACGTTATTAAAGAGTCACCGCAAGGGGAAGCACCCAAAAATGTTGA  
20 GGCATACTGCCGTGAGATGAGCCAAAAGGGAGAAACAAATCGCGTCAAATCTCCACCAA  
CTGCCCGTGCTGCCAGCTCCGAATCAATTCCTGTTGTTGTTCTTCTGTTCTCCCCGCA  
CGTTTTCGGAGAATTTCGGAGGACAATACGTTCCCGAAGCTCTTGTCGATTGTCTGGTTGA  
ACTAGAAGAAGCTCACAAATCTGCCATGGCTGATCCTGAATTCCAGAAGGAACATAAAT  
CGCATGCCGGATATGCAAATCGTCCTTCACAAATATACCTCGCCGAAAATCTCACCAAG  
25 GATGCTGGGGGTGCAAATATTTGGTTGAAACGTGAAGATTTGAACCACACAGGTTCCCA  
CAAAATCAATAACGCTTTGGGACAAATTCTGCTTGCCCGGAGAATCGGAAAGACCAGAA  
TTATCGCAGAAACAGGTGCCGCGCAGCATGGTGTGCAACAGCGACTGTTTGCGCTAAG  
TTTGGAATGGAATGTGTTATCTACATGGGCGCAGAAGATGTGCGACGGCAAGCTCTAAA  
TGTATTTCAGGATTGAGATGCTAGGAGCAAAAGTTGTTCTGTTACTTCAGGATCATGCA  
30 CATTGAAGGACGCTGTAAACGAGGCCTTCCGTGACTGGGTGACAAACCTTTCTACGACG  
CATAAAGGTCATTGGTCAAGAGATCAAGGCTCAGATGTTGGCCGCCCGCGGCAAACTTC  
CTGATGTGCTGCTCGCTTGTGTTGGTGGAGGAAGCAATGCTATCGGTACGTTCTATGAT  
TTTATTGGCGACAAGAGTGTACGTCTAGTTGGGGTGGAAGCAGGAGGAGAAGGTATTGA  
CGGAGACCGACATAGCGCCACACTTTTCGATGGGGCAACCGGGAGTACTTCACGGTGTTA  
35 GAACATATATTCTACAAGACAAGGCCGGTCAAATCATCGAGACGCACTCAATCAGCGCT  
GGATTGGATTATCCCGCGCTTGGACCAGAACATGCTTGGCTAAAGGACTCTAAAAGAGC  
AGAATATGTTGTGCGCCACAGACGAAGAAGCACTTCGCGGTTTCCGTATGCTAACACAAA  
GGGAGGGAATTATTCCTGCCCTTGAATCTTCCCATGCGATCTGGGAGGCTGTCAGGATT  
GCCCCGACCATGTGCAAGGACCAGGATCTTGTTGTGTGTTTGTCTGGCCGAGGTGATAA  
40 AGACGTTGAGCAAATTTCTCAACTTCTTCCCAAGTGGCGGATATTCTAGACTGGCATG  
TTTCTTCCCATGCCGTTGGACACACAACAAAATTCTAA

**[00442]** SEQ.ID NO: 12

45 MEAIKKVFENKKAEGIPVLVTFVTAGYPRPEDTVPIILLAMEAGGADIIELGMPFSDPIA  
DGPVIQETNTIAVANQVDYTTVLGQLREARKQGLKAPVLLMGYYNPILAYGEDRSIQDA  
AEAGANGFIMVDLPPEEAVAFREKCIKSNLSYVPLIAPSTTSLRIKFLSTIADTFIYV  
SKMGTTGSSEKVMNNALPTIIDRIREYAEVPLAVGFGVATRAHFNYVADSGADGVVIG  
TKLVNVIKESPOGEAPKNVEAYCREMSQKGETNRVKSPTARAASSESI PVVPSVLPA  
50 RFGEFGGQYVPEALVDCLVELEEAKHSAMADPEFQKELQSHAGYANRPSQIYLAENLTK  
DAGGANIWLKREDLNHTGSHKINNALGQILLARRIGKTRIIAETGAGQHGVATATVCAK

5 FGMECVIYMGAE DVRRQALNVFRIEMLGAKVVPVTS GSCTLKDAVNEAFRDWVTNLSTT  
 HYLVG SVIGPHPFPTIVRDFQKVIGQEIKAQMLAARGKLPDVVACVGGGSNAIGTFYD  
 FIGDKSVRLVGVEAGGEGIDGDRHSATLSMGQPGVLHGVRTYILQDKAGQIIETHSISA  
 GLDYPGVGPEHAWLKDSKRAEYVVATDEEALRGFRMLTQREGIIPALESSHAIWEAVRI  
 ARTMSKDQDLVVCLSGRGDKDVEQISQLLPKWADILDWHVSSHAVGHTTKF

**[00443]** SEQ.ID NO: 13

10 ATGGAAAGGCGGAAGCGTGAGCGTCTTGCGTCACTCGGCCGACCGACCAAAAAGGAGCT  
 CCGCATGATCCGATCTGTGGTGATCGTGGTGGTGGCACGGCGGGCTGGATGACCGCCT  
 CCTACCTCAAGGCCGCTTCGACGACCGCATCGACGTAACGCTCGTGGAGTCAGGGAAC  
 GTCAGGCGGATCGGGTTCGGCGAAGCGACCTTCAGCACGGTCCGCCACTTCTTCGACTA  
 CCTGGGCTTCGACGAGCGCGAGTGGCTGCCCCGCTGCGCCGGCGGCTACAAGCTCGGCA  
 TCCGCTTCGAGAACTGGAGCGAGCCGGGCGAGTACTTCTACCACCGTTTCGAGCGCCTG  
 15 CGCGTCGTCGACGGCTTCAACATGGCCGAGTGGTGGCTCGCGGTCGGCGACCGCAGGAC  
 GTCCTTCAGCGAGGCCTGCTATCTCACGCACCGGCTGTGCGAGGCCAAGCGGGCGCCCC  
 GCATGCTCGACGGCTCGCTCTTCGCCTCCAGGTGGACGAGTCGCTCGGCCGCTCGACC  
 CTGGCCGAGCAGCGCGCCAGTTCCTCGTACGCCTACCACTTCGACGCCGACGAGGTGCG  
 CCGCTACCTGTTCGGAGTACGCCATCGCCCGCGGCGTCCGCCACGTGGTTCGACGACGTGC  
 20 AGCACGTTCGGCCAGGACGAGCGCGGCTGGATCAGCGGCGTCCACACCAAGCAGCACGGC  
 GAGATCAGCGGCGACCTGTTCTGTCGACTGCACCGGCTTCGCGGCCTGCTCATCAACCA  
 GACGCTGGGCGGCAGGTTCCAGTCCTTCTCCGACGTGCTGCCCAACAACCGGGCGGTGCG  
 CGCTGCGCGTCCCGCGGGAGAACGACGAGGACATGCGGCCGTACACGACGGCGACCGCG  
 ATGAGCGCCGGCTGGATGTGGACGATCCCGCTGTTCAAGCGCGACGGCAACGGCTACGT  
 25 CTACTCCGACGAGTTCATCTCGCCGGAGGAGGCCGAGCGCGAGCTGCGGTCCACCGTCG  
 CCCCCGGCCGCGACGACCTGGAGGCCAACCACATCCAGATGCGGATCGGCAGGAACGAG  
 CGCACCTGGATCAACAACCTGCGTCGCGCTCGGCCTGTCCGCCGCTTCGTCGAGCCGCT  
 GGAGTCGACCGGCATCTTCTTCATCCAGCACGCCATCGAGCAGCTCGTGAAGCACTTCC  
 CCGGCGAGCGCTGGGACCCGGTGTGATCAGCGCGTACAACGAGCGCATGGCGCACATG  
 30 GTCGACGGGGTCAAGGAGTTCCTCGTCTTCCACTACAAGGGCGCCCAGCGCGAGGACAC  
 CCCCTACTGGAAGGCCGCCAAGACCAGGGCCATGCCCGACGGCCTCGCCCCGAAGCTGG  
 AGCTGTCCGCTCCACCTGCTGGACGAGCAGACGATCTACCCCTACTACCACGGCTTC  
 GAGACCTATTCTGTGGATCACCATGAACCTCGGTCTCGGCATCGTGCCCGAGCGGCCGCG  
 TCCCGCGCTCCTGCACATGGACCCGGCGCCCGCGCTGGCCGAGTTCGAACGGCTCAGGC  
 35 GCGAGGGCGACGAGCTGATCGCCGCCCTGCCAGCTGCTACGAGTACCTCGCCAGCATC  
 CAATGA

**[00444]** SEQ.ID NO: 14

40 MERRKRERLGS LGRPTKKELRMIRS VVIVGGGTAGWMTAS YLKAAFDDRIDVTLVESGN  
 VRRIGVGEATFSTVRHFFDYLG LDEREWLPRCAGGYKLGI RFENWSEPGEYFYHPFERL  
 RVVDGFNMAEWLAVGDRRTSFSEACYLTHRLCEAKRAPRMLDGS LFA SQVDESLGRST  
 LAEQRAQFPYAYHFD ADEVARYLSEYAIARGVRHVVDVQHVGQDERGWISGVHTKQHG  
 EISGDLFVDCTGFRGLLINQTLGGRFQSFS DVL PNNRAVALRVPRENDEDMRPYTTATA  
 45 MSAGWMWTIPLFKRDGN GYVYSDEFISPEEAERELRSTVAPGRDDLEANHIQMRIGRNE  
 RTWINNCVAVGLSAAFVEPLESTGIFFIQHAIEQLVKHFPGERWDPVLISAYNERMAHM  
 VDG VKEFLVLHYKGAQREDTPYWKAAKTRAMPDGLARKLELSASHLLDEQTIYPYHGF  
 ETYSWITMNLGLGIVPERPRPALLHMDPAPALAEFERLRREGDELIAALPSCYEYLASI  
 Q

50

**[00445]** SEQ.ID NO: 15

5 ATGACCGCCGCCTACCTGAAGACCGCGTTCGGCGACCGGCTTTCGATCACCGTCGTGGA  
 ATCGAGCAGAATCGGGACCATCGGCGTCGGTGAGGCGACCTTCAGCGACATCCAGCACT  
 TCTTCCAGTTCCTCAACCTCCGGGAACAGGACTGGATGCCGGCCTGCAACGCCACCTAC  
 AAGCTGGGCATCCGCTTCGAGAACTGGCGGCATGTCGGCCATCATTTCTACCAGCCGTT  
 CGAGCAGATCCGCCCCGGTCTACGGCTTTCGGCTGACCGACTGGTGGCTGCACGACGCGC  
 CGACCGACCGCTTCGACACCGACTGCTTCGTTCATGCCAACCTGTGCGAGGCCGGGCGC  
 AGCCCCGCGCCACCTCGACGGCAGCTGGCCGACGAGGACTTCGTGGAGGAGGGCGACGA  
 10 GCTCGCCAACCGCACCATGTCCGAGCACCAGGGCAAGTCGCAGTTCCTCGTACGCCTACC  
 ACTTCGAGGCGGCGCTGCTGGCCAAGTTCCTCACCGGGTACGCGGTGGACCGCGGCGTC  
 GAGCACGTCGTCGACGACGTCTGGACGTGCGGCTGGACCAGCGCGGGTGGATCGAGCA  
 CGTGGTCAACGCCGAGCAGGCGAGATCCACGGCGACCTGTTTCGTGCACTGCACCGGTT  
 TCCGGGGACTGCTGCTGAACAAGGCGCTCGGTGTGCCGTTTCGTGTCCTATCAGGACACT  
 15 CTGCCCCAACGACAGCGCGGTTCGCCCTCCAGGTCCCGCTGGACATGCAGCGGCGCGGCAT  
 CGTGCCGAACACCACGGCCACCGCGCGGGAGGCCGGCTGGATCTGGACCATCCCGCTGT  
 TCGGTTCGGTTCGGCACCGGGTACGTCTACGCCAAGGACTACCTCTCGCCCCGAGGAGGCC  
 GAGCGCACGCTGCGCGAGTTCGTTCGGCCCCGGCGGCGCGGACGTGGAGGCCAACACAT  
 CCGCATGCGGATCGGCCGAGCCAGGAGTCTTGGCGGAACAACTGCGTGGCCATCGGCC  
 20 TGTCCAGCGGCTTCGTTCGAGCCGCTGGAGTCGACCGGGATCTTCTTCATCCACCACGCC  
 ATCGAGCAGCTGGTGAAGCACTTCCCGGCGGCGGACTGGAACCCGAAGTCCCGGGACAT  
 GTACAACTCCGCGGTTCGCGCACGTGATGGACGGCATCCGCGAGTTCCTGGTGATCCACT  
 ACCGGGGCGCGGCGCGGGCCGACAACCAGTACTGGCGCGACACCAAGACCCGCCCCGCTG  
 CCGGACGGCCTGGCCGAGCGCATCGAGTGTGGCAGACGCAGCTGCCGGACACCGAGAC  
 25 GATCTACCCGTACTACCACGGGCTGCCGCCCTACTCCTACATGTGCATCCTGATGGGCG  
 GCGGGGCCATCCGGACGCCGGCCTCGGCCGCGTTGGCCCTGACCAGCAGGGGGCCGCG  
 CAGAAGGAGTTCCGCCCGCTCCGGGACCGGGCCGCGCAGCTGCGCGACACGCTGCCCAG  
 CCACTACGAGTACCTCGCGCGGATGCGTGGTCTGGATGTCTGA

30 **[00446]** SEQ.ID NO: 16

35 MTAAYLKTAFGDRLSITVVESSRIGTIGVGEATFSDIQHFFQFLNLREQDWMPACNATY  
 KLGRFENNRHVGHFYQPFEQIRPVYGFPLTDWWLHDAPTDRFDTDCFVMPNLCEAGR  
 SPRHLDGTLADEDFVEEGDELANRTMSEHQKSGFPYAYHFEEALLAKFLTGYAVDRGV  
 EHVVDVLDVRLDQRGWIEHVVTAEHGEIHGDLFVDCTGFRGLLLNKALGVPFVSYQDT  
 LPNDSAVALQVPLDMQRRGIVPNTTATAREAGWIWTIPLFGRVGTGYVYAKDYLSPEEA  
 ERTLRETFVGPAAADVEANHIRMIRIGRSQESWRNNCVAIGLSSGFVEPLESTGIFFIHHA  
 IEQLVKHFPAADWNPKSRDMYNSAVAHVMDGIREFLVIHYRGAARADNQYWRDTKTRPL  
 PDGLAERIECWQTQLPDTETIYPYHGLPPYSYMCILMGGGAIRTPASAALALTDQGAA  
 40 QKEFAAVRDRAAQLRDTLP SHYEYLARMRGLDV

**[00447]** SEQ.ID NO: 17

45 ATGAACACGAGAAATCCGGACAAGGTGCTCATCGTCGGCGGCGGGCACGGCGGGATGGAT  
 GACCGCCTCGTACCTGAAGAAGGCGTTCGGCGAGCGCGTGTCCGTGACCCCTCGTCAAT  
 CGGGCACCATCGGCACCGTTCGGCGTCGGCGAGGCGACCTTCAGCGACATCCGGCACTTC  
 TTCGAATTCTTCGACCTGCGGGAGGAGGAATGGATGCCGGCCTGCAACGCCACGTACAA  
 GCTGGCCGTCCGCTTCAGGACTGGCAGCGCCCCGGCCACCACTTCTACCACCCGTTTCG  
 AGCAGATGCGCTCCGTAGACGGGTTCCCCCTGACCGACTGGTGGCTGCAGAACGGGCCG  
 50 ACCGACCGGTTTCGACAGGGACTGCTTCGTGATGGCGTCGCTGTGCGACGCGGGGCGCAG  
 CCCGCGCTATCTCAACGGTTCCTGCTGCAGCAGGAGTTCGACGAGCGCGCGGAGGAGC

CGGCGGGGCTCACCATGTCCGAGCACACAGGGCAAGACGCAGTTCCCCTACGCCTACCAC  
 TTCGAGGCGGCCCTGCTCGCCGAATTCTTTCTGGGTACTCCAAGGACCGCGGCGTGAA  
 GCACGTGGTCGACGAGGTGCTGGAGGTGAAGCTCGACGACCGCGGATGGATCAGCCACG  
 TTGTCACCAAGGAGCACGGCGACATCGGCGGCGACCTGTTCTGTCGACTGCACCGGTTTC  
 5 CGCGGCGTACTGCTGAACCAGGCGCTGGGTGTGCCCTTCGTCTCCTACCAGGACACCCT  
 GCCCAATGACAGCGCGGTGCGCCCTCCAGGTGCCGCTCGACATGGAGGCGCGGGGCATCC  
 CGCCGTACACCCGGGCGCCACCGCCAAGGAGGCCGGGTGGATCTGGACGATCCCGCTGATC  
 GGCCGGATCGGCACCGGGTACGTGTACGCCAAGGACTACTGCTCGCCGGAGGAGGCCGA  
 GCGCACCCCTGCGCGAGTTCTGTCGGTCCGGAGGCCCGACGTGAGGCCAACCACATCC  
 10 GCATGCGGATCGGGCGCAGCGAGCAGTCTTGAAGAACAACCTGCGTCGCGATCGGGCTC  
 TCCAGTGGCTTCGTGCAACCCCTCGAGTCGACCGGCATCTTCTTCATCCACCACGCCAT  
 CGAGCAGTTGGTGAAGCACTTCCCGGCGGGCGACTGGCACCCGCGAGCTGCGCGCCGGCT  
 ACAACTCGGCCGTGCGCAACGTATGGACGGGGTCCGGGAGTTCCTCGTCTGCACTAC  
 CTGGGCGCCGCGCGCAACGACACCCGGTACTGGAAGGACACCAAGACCCGGGCGGTTC  
 15 GGATGCCCTGGCGGAGCGCATCGAGCGCTGGAAGGTGCAGCTGCCGACTCCGAGAACG  
 TCTTCCCTACTACCACGGCCTGCCGCCGTACTCGTACATGGCCATCCTGCTCGGGACC  
 GGCGCGATCGGGCTGCGACCGTCCCGGGCGCTCGCACTGGCCGACCCGGCGGCGGCGGA  
 GAAGGAGTTACCCGCGATCCGGGACAGGGCGCGGTTCTGGTGCACACCTGCCGAGCC  
 AGTACGAGTACTTCGCCGCGATGGGGCAGCGTGTCTGA

20

**[00448]** SEQ.ID NO: 18

MNTRNPDKVIVGGGTAGWMTASYLKKAFGERSVTLVESGTIGTVGVGEATFSDIRHF  
 FEFLDLREEEWMPACNATYKLAVRFQDWQRPQHFFYHPFEQMRSDVGFPLTDWWLQNGP  
 25 TDRFDRDCFVMASLCDAGRSPRYLNGSLQQEFDERAEPAGLTMSEHQKTQFPYAYH  
 FEAALLAEFLSGYSKDRGVKHVVDEVLEVKLDDRGWISHVVTKEHGDIGGDLFVDCTGF  
 RGVLLNQALGVFPVSYQDTLPNDSAVALQVPLDMEARGIPPYTRATAKEAGWIWTIPLI  
 GRIGTGYVYAKDYCSPEEAERTLREFVGPAAADVEANHIRMRIGRSEQSWKNNCVAIGL  
 SSGFVEPLESTGIFFIHHAIEQLVKHFAPGDWHPQLRAGYNSAVANVMDGVREFVLVHY  
 30 LGAARNDRYWKDTKTRAVPDALAERIERWKVQLPDSENVFPYYHGLPPYSYMAILLGT  
 GAIGLRPSPALALADPAAAEKEFTAIRDRAFLVDTLPSQYEFYFAAMGQRV

**[00449]** SEQ.ID NO: 19

ATGGACAATCGAATCAAGACAGTCGTGATCCTGGGCGGCGGCACCGCCGGCTGGATGAC  
 TGCGGCGTATCTCGGCAAGGCGCTGCAGAACACGGTGAAGATCGTGGTCTGGAGGCGC  
 CCACCATCCCGCGGATCGGTGTGGGCGAAGCCACCGTCCCCAACCTGCAGCGCGCGTTC  
 TTCGACTACCTCGGCATCCCCGAGGAGGAGTGATGCGCGAGTGCAACGCCAGTTACAA  
 GATGGCGGTGAAGTTCATCAACTGGCGCACACCGGGGAGGGCAGCCCGGATCCCCGGA  
 40 CGCTCGACGACGGTCACACCGACACCTTCCACCACCCCTTCGGACTGCTGCCGTCCGCC  
 GACCAGATACCGCTCTCCCACTACTGGGCGGCCAAGCGGCTCCAGGGCGAGACCGACGA  
 GAACTTCGACGAGGCGTGTTCGCCGACACCGCGATCATGAACGCCAAGAAGGCGCCCC  
 GCTTCCTCGACATGCGGCGGGCGACCAACTACGCCTGGCACTTCGACGCCAGCAAGGTC  
 GCCGCGTTCCTGCGCAACTTCGCCGTACCAAGCAGGCGGTGAGCACGTGAGGACGA  
 45 GATGACCGAGGTCTCACCAGCAACGGGGCTTCATCACCGCCCTGCGCACCAAGTCGG  
 GACGGATCCTGCAGGGTGACCTCTTCGTGACTGCTCCGGCTTCGCGGGGCTGCTGATC  
 AACAAGGCCATGGAGGAGCCGTTCATCGACATGAGCGACCACCTCCTGTGCAACAGCGC  
 GGTGCGCACGGCCGTACCGCACGACGACGAGAAGAACGGTGTGAGCCGTACACCTCCT  
 CGATCGCCATGGAGGCCGGCTGGACCTGGAAGATTCCCATGCTGGGCGGGTTCGGCAGC  
 50 GGTACGTCTACTCCGACCACTTCGCCACCCAGGACGAGGCCACCCTGGCCTTCTCGAA  
 GCTGTGGGGCCTCGACCCGGACAACACCGAGTTCAACCACGTCCGCTTCGGGTTCGGC

GCAACCGCAGGGCCTGGGTGCGCAACTGCGTCAGCGTCGGCCTCGCGTCGTGCTTCGTG  
 GAGCCGCTGGAGTCCAGCGGCATCTACTTCATCTACGCGGCATCCACATGCTGGCCAA  
 GCACTTCCCCGACAAGACCTTCGACAAGGTGCTCGTCGACCGGTTCAACCGGGAGATCG  
 AGGAGATGTTTCGACGACACCCGGGACTTCCTCCAGGCGCACTACTACTTCTCGCCGCGC  
 5 GTCGACACGCCGTTCTGGCGGGCGAACAAGGAAGTGAAGCTGGCCGACTCCATCAAGGA  
 CAAGGTCGAGACGTACCGGGCCGGGCTGCCGGTCAACCTGCCGGTCACCGACGAGGGGA  
 CCTACTACGGCAACTTCGAGGCCGAGTTCGGAACTTCTGGACCAACGGCAGCTACTAC  
 TGCATCTTCGCCGGCCTCGGCCTGATGCCGCGCAACCCGCTGCCGCGCTCGCCTACAA  
 GCCGAGTCGATCGCGGAGGCCGAGCTGCTGTTCCGCCGACGTCAAGCGCAAGGGGGACA  
 10 CCCTGGTCGAGTCGCTGCCGTCGACCTATGACCTGCTGCGTCAGTTGCACGGTGCGTCG  
 TGA

**[00450]** SEQ.ID NO: 20

MDNRIKTVVILGGGTAGWMTAAYLGKALQNTVKIVVLEAPTIPRIGVGEATVPNLQRAF  
 FDYLGIPPEEWMRECNASYKMAVKFINWRTPGEGSPDPRTLDDGHTDTFHHFGLLP  
 DQIPLSHYWAARKRLQGETDENFDEACFADTAIMNAKKAPRFLDMRRATNYAWHFDASKV  
 AAFLRNFAVTKQAVEHVEDEMTEVLTDERGFITALRTKSGRILQGDLFVDCSGFRGLLI  
 NKAMEEPFIDMSDHLLCNSAVATAVPHDDEKNGVEPYTSSIAMEAGWTWKI PMLGRFGS  
 20 GHVYSDHFATQDEATLAFSKLWGLDPDNTEFNHVRFRVGRNRRRAWVRNCVSVGLASCFV  
 EPLESSGIYFIYAAIHMLAKHFPDKTFDKVLVDRFNREIEEMFDDTRDFLQAHYYFS  
 PRVDTPFWRANKELKLADSIKDKVETIRAGLPVNLVPTDEGTYYGNFEAEFRNFWTNGS  
 SYCIFAGLGLMPRNPLPALAYKPQSI AEAEELLFADV KRKGDTLVESLPSTYDLLRQLHGAS

**[00451]** SEQ.ID NO: 21

ATGGACAACCGCATCAACCGGATCGTCATCCTCGGCGGCGGGACCGCGGGATGGATGAC  
 CGCCTCCTACCTGGCGAAGGCGCTCGGTGACACCGTCACGATCACCTGCTGGAAGCAC  
 CGGCGATCGGCCGGATCGGGGTGGCGAGGCCACCGTGCCCAACCTGCAGCGCGTGTTC  
 30 TTCGACTTCCTCGGCCTCCGCGAGGAGGAGTGGATGCCGAGTGCAACGCCGCGTTC  
 GACGGCCGTGAAGTTCATCAACTGGCGTACCCCCGGGCCCGGCGAGGCCAAGGCCCGCA  
 CGATCGACGGCCGGCCGGACCACTTCTACCACCGTTCGGGCTGCTGCCCCGAGCACGGC  
 CAGGTGCCGCTGTTCGCACTACTGGGCCTACAACCGGGCGGCCGGCACCACCGACGAGCC  
 GTTCGACTACGCTGTTTCGCCGAGACCGCCGCCATGGACGCGGTCCGCGCGCCCAAGT  
 35 GGCTCGACGGGCGGCCCGCCACCCGGTATGCCTGGCACTTCGACGCCACCTGGTCGCC  
 GAGTTCCTGCGCCGGCACGCCACCGAGCGGTCAACGTGGAGCACGTGCAGGGTGAGAT  
 GCAGCAGGTGCTGCGCGACGAGCGGGGCTTCATCACCGCCCTGCGCACGGTCGAGGGCC  
 GGGACCTCGAGGGTGACCTGTTTCGACTGCTCCGGCTTCCGCGGCCTGCTGATCAAC  
 AAGGCCATGGAGGAGCCCTTCATCGACATGAACGACCAGCTGCTGTGCAACCGCGCGGT  
 40 GGCCACGGCGATCAAGCACGACGACGACGCGCACGGGGTCGAGCCGTACACCTCGGCGA  
 TCGCCATGCGGTCCGGCTGGAGCTGGAAGATCCCGATGCTGGGCCGCTTCGGGACCGGC  
 TACGTCTACTCCAGCCGGTTCGCCGAGAAGGACGAGGCCACCTCGACTTCTGCCGGAT  
 GTGGGGCCTGGACCCGGAGAACACCCCGCTCAACCAGGTGGCCTTCCGGGTCCGCCGCA  
 ACCGCCGGGCCTGGGTGAAGAACTGCGTGAGCATCGGCCTGGCGTCTGCTTCCTGGAG  
 45 CCACTGGAGTCCACCGGCATCTACTTCATCACCGCGGCCATCTACCAGCTGACTCAGCA  
 CTTCCCGGACCGCACCTTCGCGCTCGCGTTGAGCGACGCGTTCAACCACGAGATCGAAG  
 CCATGTTTCGACGACACCCGGGACTTCATCCAGGCGCACTTCTACGTGTCGCCGCGCACC  
 GACACGCCGTTCTGGAAGGCGAACAAGGACCTGCACCTGCCCGAGCAGATGCGCGAGAA  
 GATCGCGATGTACAAGGCCGGCCTGCCGATCAACGCGCCGGTCACCGACGAGTCGACCT  
 50 ACTACGGCAGGTTCGAGGCCGAGTTCGCCAACTTCTGGACCAACGGCAGCTACTACTGC  
 ATCTTCGCGGGCCTCGGGCTGCGCCCGGACAACCCGCTGCCGATGCTGCGACACCGTCC

GGAGCAGGTCCGCGAGGCGCAGGCGCTGTTCCGCCGGGGTGAAGGACAAGCAGCGGGAGT  
TGGTGGAGACGCTGCCGAGCAACCTGGAGTTCCTGCGCAGCCTGCACGGCAAGTAG

**[00452]** SEQ.ID NO: 22

5 MDNRINRIVILGGGTAGWMTASYLAKALGDTVTTITLEAPAIGRIGVGEATVPNLQRVF  
FDLGLREEEWMPECNAAFKTAVKFINWRTPGPGEAKARTIDGRPDHFYHPFGLLPEHG  
QVPLSHYWAYNRAAGTTDEPFDYACFAETAAMDRAVRAPKWLDGRPATRYAWHFD AHLVA  
EFLRRHATERLNVEHVQGEMQQVLRDERGFITALRTVEGRDLEGLDFIDCSGFRGLLIN  
10 KAMEEPFIDMNDQLLCNRAVATAIKHDDDAHGVEPYTSAIAMRSGWSWKIPMLGRFGTG  
YVYSSRFAEKDEATLDFCRMWGLDPENTPLNQVAFRVGRNRRRAWVKNVCSIGLASC FLE  
PLESTGIYFITAAYQLTQHFPDRTFALALS DAFNHEIEAMFDDTRDFIQAHFYVSPRT  
DTPFWKANKDLHLPEQMREKIAMYKAGLPINAPVTDESTYYGRFEAEFRNFWTNGSYC  
IFAGLGLRPDNLPLMLRHRPEQVREAQALFAGVKDKQRELVETLPSNLEFLRSLHGK

**[00453]** SEQ.ID NO: 23

ATGGATGACAATCGAATTCGGAGCATCCTTGTCCTTGGCGGCGGCACGGCCGGCTGGAT  
GTCCGCCTGCTACCTGAGCAAGGCGCTCGGGCCCCGGCGTCGAGGTCACCGTGCTCGAGG  
20 CGCCCTCCATCTCGCGCATCCGGGTGCGCGAGGCCACCATTCCCAACCTGCACAAGGTC  
TTCTTCGACTTCCTGGGCATCGCCGAGGACGAGTGGATGCGGGAGTGCAACGCCAGCTA  
CAAGGCCGCGGTCCGGTTCGTCAACTGGCGGACGCCGGGCGACGGCCAGGCCACGCCGC  
GGCGGCGTCCGGACGGCCGCCCCGACCACTTCGACCACCTGTTGCGCCAGCTGCCCGAG  
CACGAGAACCTGCCGCTGTCGCGAGTACTGGGCGCACCGGCGCCTCAACGGCCTGACCGA  
25 CGAACCCCTTCGACCGCTCCTGCTACGTGCAGCCCGAGCTGCTGGACCGCAAGCTCTCGC  
CGAGGTTGATGGACGGCACGAACTGGCCAGCTACGCCTGGCACTTCGACGCCGACCTG  
GTGGCCGACTTCCTCTGCCGGTTCGCCGTGCAGAAAGCTGAACGTGACCCACGTCCAGGA  
CGTGTTACGCATGCCGACCTCGACCAGCGCGGCCACATCACGGCCGTCAACACCGAGT  
CCGGCCGCACGCTGGCCGCCGACCTGTTTCATCGACTGCAGCGGCTTCCGCAGCGTGCTC  
30 ATGGGCAAGGTCATGCAGGAGCCGTTTCTGGACATGAGCAAGCACCTGCTCAACGACCG  
CGCGGTGGCGCTGATGCTCCCGCACGACGACGAGAAGGTGGGCATCGAGCCGTACACCT  
CGTCGCTGGCCATGCGGTGCGGCTGGTTCGTGGAAGATCCCGCTGCTCGGCCGGTTCGGC  
TCCGGCTACGTCTACTCCAGCCAGTTCACCTCCCAGGACGAGGCGGCCGAGGAGCTCTG  
CCGCATGTGGGACGTGACCCGGCGGAGCAGACGTTCAACAACGTCCGGTTCGGGGTCG  
35 GCCGCAGCCGCCGGGCTGGGTGCGCAACTGCGTCGCCATCGGCGTGTCCGCCATGTTT  
GTGGAGCCGCTGGAGTCGACCGGCCTGTACTTCAGTTACGCCTCGCTCTACCAGCTGGT  
GAAGCACTTCCCGGACAAGCGGTTCCGGCCGATCCTGGCCGACCGGTTCAACCGCGAGG  
TGCGGACCATGTACGACGACACCCGCGACTTCCTCCAGGCGCACTTCAGCCTGTCCCG  
CGTGACGACTCCGAGTTCTGGCGGGCCTGCAAGGAGCTGCCGTTTCGCGGACGGGTTTCG  
40 CGAGAAGGTCGAGATGTACAGGGCCGGGCTGCCGGTCGAACTGCCGGTCACCATCGACG  
ACGGGCACTACTACGGCAATTTGAGGCCGAGTTCGCAACTTCTGGACCAACTCGAAC  
TACTACTGCATCTTCGCCGGGCTCGGTTTCTGCCCCGAGCATCCGCTGCCGGTGCTCGA  
ATTCCGCCCGGAGGCCGTGATCGCGCGGAGCCGGTGTTCGCCGCGGTGCGCCGGCGCA  
CGGAGGAGCTGGTCGCCACCGCCCCGACCATGCAGGCCTACCTGCGGCGCCTGCACCAG  
45 GGCACGTAG

**[00454]** SEQ.ID NO: 24

50 MDDNRIRSILVLGGGTAGWMSACYLSKALGPGVEVTVLEAPSI SRIRVGEATIPNLHKV  
FFDFLGIAEDEWMRECNASYKAAVRVFNWRTPGDGQATPRRRPDGRPDHFDHLFGQLPE  
HENLPLSQYWAHRRNLGLTDEPFDRSCYVQPELLDRKLS PRLMDGTKLASYAWHFDADL

5 VADFLCRFAVQKLNVTQVDFTHADLDQRGHITAVNTESGRTLADLFDICSGFRSVL  
 MGKVMQEPFLDMSKHLNDRAVALMLPHDDEKVGIEPYTSSLAMRSGWSWKIPLLGRFG  
 SGYVYSSQFTSQDEAAEELCRMWDVDPAEQTFNNVRFRVGRSRRAWVRNCVAIGVSAMF  
 VEPLESTGLYFSYASLYQLVKHFDPKRFRIPLADRFNREVATMYDDTRDFLQAHFSLSP  
 RDDSEFWRACKELPFADGFAEKVEMYRAGLPVELPVTIDDGHYYGNFEAEFRNFWTNSN  
 YYCIFAGLGLPEHPLPVLEFRPEAVDRAEPVFAAVRRRTEELVATAPTMQAYLRRLHQ  
 GT

10 **[00455]** SEQ.ID NO: 25

ATGAACAAGCCGATCAAGAATATCGTCATCGTGGGCGGCGGTACTGCGGGCTGGATGGC  
 CGCCTCGTACCTCGTCCGGGCCCCTCCAACAGCAGGCGAACATTACGCTCATCGAATCTG  
 CGGCGATCCCTCGGATCGGCGTGGGCGAAGCGACCATCCCAAGTTTGCAGAAGGTGTTT  
 15 TTCGATTTCTCGGGATACCGGAGCGGGAATGGATGCCCCAAGTGAACGGCGCGTTCAA  
 GGCCGCGATCAAGTTCGTGAATTGGAGAAAGTCTCCCGACCCCTCGCGCGACGATCACT  
 TCTACCATTTGTTCTGGCAACGTGCCGAACGTGCGACGGCGTGCCGCTTACCCACTACTGG  
 CTGCGCAAGCGCGAACAGGGCTTCCAGCAGCCGATGGAGTACGCGTGCTACCCGCAGCC  
 CGGGGCACTCGACGGCAAGCTGGCACCGTGCTGTCGACGGCACCCGCCAGATGTCCC  
 ACGCGTGGCACTTCGACGCGCACCTGGTGGCCGACTTCTTGAAGCGCTGGGCCGTCGAG  
 20 CGCGGGGTGAACCGCGTGGTCGATGAGGTGGTGGACGTTTCGCTGAACAACCGCGGCTA  
 CATCTCCAACCTGCTCACCAAGGAGGGGCGGACGCTGGAGGCGGACCTGTTTCATCGACT  
 GCTCCGGCATGCGGGGGCTCCTGATCAATCAGGCGCTGAAGGAACCCTTCATCGACATG  
 TCCGACTACCTGCTGTGCGACAGCGCGGTGCGCCAGCGCCGTGCCAACGACGACGCGCG  
 CGATGGGGTCGAGCCGTACACCTCCTCGATCGCCATGAACTCGGGATGGACCTGGAAGA  
 25 TTCCGATGCTGGGCCGTTTCGGCAGCGGCTACGTCTTCTCGAGCCATTTACCTCGCGC  
 GACCAGGCCACCGCCGACTTCTCAAACCTCTGGGGCCTCTCGGACAATCAGCCGCTCAA  
 CCAGATCAAGTTCGGGGTCGGGCGCAACAAGCGGGCGTGGGTCAACAACCTGCGTCTCGA  
 TCGGGCTGTCGTGCTGCTTTCTGGAGCCCCTGGAATCGACGGGGATCTACTTCATCTAC  
 GCGGCGCTTTACCAGCTCGTGAAGCACTTCCCCGACACCTCGTTCGACCCGCGGCTGAG  
 30 CGACGCTTTCAACGCCGAGATCGTCCACATGTTTCGACGACTGCCGGGATTTCTGTTCAAAG  
 CGCACTATTTACACGCTCGCGCGATGACACGCCGTTCTGGCTCGCGAACCAGGCGACGAC  
 CTGCGGCTCTCGGACGCCATCAAAGAGAAGGTTTCAGCGCTACAAGGCGGGGCTGCCGCT  
 GACCACCACGTCGTTTCGACGATTCACGTAACGAGACCTTCGACTACGAATTCAAGA  
 ATTTCTGGTTGAACGGCAACTACTACTGCATCTTTGCCGGCTTGGGCATGCTGCCCGAC  
 35 CGGTCGCTGCCGCTGTTGCAGCACCGACCGGAGTCGATCGAGAAAGCCGAGGCGATGTT  
 CGCCAGCATCCGGCGCGAGGCCGAGCGTCTGCGCACACGCTGCCGACAACTACGACT  
 ACCTGCGGTGCTGCGTGACGGCGACGCGGGGCTGTGCGCGGGCCAGCGTGGGCCGAAG  
 CTCGCAGCGCAGGAAAGCCTGTAG

40 **[00456]** SEQ.ID NO: 26

MNKPIKNIVIVGGGTAGWMAASYLVRLQQQANITLIESAAIPRIGVGEATIPSLQKVF  
 FDFLGIPEREWMPQVNGAFKAAIKFVNWRKSPDPSRDDHFFYHLFGNVPNCDGVPLTHYW  
 LRKREQGFQQPMEYACYPPQGALDGKLAPCLSDGTRQMSHAWHFD AHLVADFLKRWAVE  
 45 RGVNRVVDVVDVRLNNRGYISNLLTKEGRTLEADLFDICSGMRGLLINQALKEPFIDM  
 SDYLLCDSAVASAVPNDDARDGVEPYTSSIAMNSGWTWKIPMLGRFGSGYVFSSHFTSR  
 DQATADFLKLWGLSDNQNPLNQIKFRVGRNKRKRAWVNNCVSIGLSSCFLEPLESTGIYFIY  
 AALYQLVKHFDPDTSFDPRLSDAFNAEIVHMFDDCRDFVQAHYFTTSRDDTPFWLANRHD  
 LRLSDAIKEKVQRYKAGLPLTTTSFDDSTYYETFDYEFKNFWLNGNYCYCIFAGLGLMLPD  
 50 RSLPLLQHRPESIEKAEAMFASIRREAERLRTSLPTNYDYLRSLRDGDAGLSRGQRGPK  
 LAAQESL

**[00457]** SEQ.ID NO: 27

ATGTCCGGCAAGATTGACAAGATCCTCATCGTCGGCGGGCGGCACCGCCGGATGGATGGC  
 CGCGTCCTATCTCGGCAAGGCCCTGCAGGGCACCGCGGACATCACACTGCTGCAGGCAC  
 5 CCGACATCCCGACGCTCGGGGTCGGCGAGGCCACGATCCCCAATCTGCAGACGGCGTTC  
 TTCGACTTCCTCGGAATCCCGAGGACGAGTGGATGCGGGAGTGCAACGCGAGCTACAA  
 GGTCGCCATCAAGTTCATCAACTGGCGCACCGCGGGCGAGGGGACGTCGAGGCCCGCG  
 AGCTCGACGGAGGGCCCGACCACTTCTACCACTCCTTCGGTCTGCTCAAGTACCACGAG  
 CAGATTCCGCTGTCGCACTACTGGTTCGACCGTTCGTACCGGGGAAGACCGTCGAGCC  
 10 GTTCGACTACGCCTGCTACAAGGAACCCGTCATCCTCGACGCCAACAGGTCACCGCGCA  
 GGCTCGACGGTTCGAAGGTGACGAACACGCGTGGCACTTCGACGCGCACCTCGTCGCC  
 GACTTCCTGCGCCGGTTCGCCACCGAGAAGCTCGGCGTGCGCCACGTCGAGGACCGCGT  
 CGAGCACGTCCAGCGCGACGCCAACGGCAACATCGAGTCGGTTCGCACGGCAACGGGGC  
 GTGTCTTCGATGCCGACCTCTTCGTGCTGCTCGGGCTTCGCGGGCTGCTGATCAAC  
 15 AAGGCGATGGAGGAGCCCTTCTCGACATGAGCGATCACCTGCTCAACGACAGCGCCGT  
 CGCCACCCAGGTGCCGCACGACGACGACGCGAACGGTGTGGAACCGTTCACCTCGGCGA  
 TCGCCATGAAGTCGGGCTGGACGTGGAAGATCCCGATGCTCGGCAGGTTTCGGCACCGGG  
 TACGTCTACTCGAGCCGGTTCGCCACCGAGGACGAGGCGGTGCGCGAGTTCGCGAGAT  
 GTGGCACCTCGACCCGGAGACCCAGCCCTCAACAGGATCCGGTTCGGGTTCGGCCGCA  
 20 ACCGGCGCGCGTGGGTTCGGCAACTGCGTCAGCATCGGCACGTCGTGCTGCTTCGTGGAA  
 CCACTGGAGTCGACGGGCATCTACTTCGTCTACGCCGCGCTGTACCAGCTGGTGAAGCA  
 CTTCCCCGACAAGAGCCTCAACCCCGTGTGACCGCCAGGTTCAACCGCGAGATCGAGA  
 CGATGTTTCGACGACACGCGCGACTTCATCCAGGCGCACTTCTACTTCTCGCCGCGCACG  
 GACACCCCGTTCCTGGAGGGCCAACAAGGAGCTGCGCCTGGCGGACGGCATGCAGGAGAA  
 25 GATCGACATGTACCGCGCGGGCATGGCGATCAACGCGCCCGCGTCCGACGACGCCCAGC  
 TCTACTACGGCAACTTCGAGGAGGAGTTCGCAACTTCTGGAACAACAGCAACTACTAC  
 TGGCTGCTGGCCGGCCTCGGTCTGGTGCCGACGCACCCTCACCACGCCTGGCGCACAT  
 GCCACAGGCGACGGAGTCGGTGGACGAGGTCTTCGGCGCCGTCAAGGACCGGCAGCGGA  
 ACCTGCTCGAGACCTGCCGAGCCTCCACGAGTTCCTGAGGCAACAGCACGGCCGCTGA  
 30

**[00458]** SEQ.ID NO: 28

MSGKIDKILIVGGGTAGWMAASYLGKALQGTADITLLQAPDIPTLGVGEATIPNLQTAF  
 FDFLGIPEDEWMRECNASYKVAIKFINWRTAGEGTSEARELDGGPDHFYHSFGLLKYHE  
 35 QIPLSHYWFDRSYRGKTVEPFYACYKEPVILDANRSPRRLDGSKVTNYAWHFD AHLVA  
 DFLRRFATEKLGVVRHVEDRVEHVQRDANGNIESVRTATGRVFDADLFVDCSGFRGLLIN  
 KAMEEPFLDMSDHLND SAVATQVPHDDDANGVEPFTSAIAMKSGWTWKI PMLGRFGTG  
 YVYSSRFATEDEAVREFCEMWHLD PETQPLNRIRFRVGRNRRRAWVGNVCSIGTSSCFVE  
 PLESTGIYFYVYAALYQLVKHFDPKSLNPVLTARFNREIETMFDDTRDFIQAHFYFSPRT  
 40 DTPFWRANKELRLADGMQEKIDMYRAGMAINAPASDDAQLYYGNFEEEFNFWNNSNY  
 CVLAGLGLVPDAPSPRLAHMPQATESVDEVFGAVKDRQRNLLETLP SLHEFLRQHGHR

**[00459]** SEQ.ID NO: 29

MLNNVVIIVGGGTAGWMTASYFKAAFGERINITLVESGNIGAVGVGEATFSDIRHFFDFL  
 GLKEKDWPACNATYKLAVRFENWREEGHYFYHPFEQMRSVNGFPLADWWLREKPTDRF  
 DKDCFVMPSIIDAGLSPRHLDGSLIDQPFVEGAEE MQGLTMSEHQGKTQFPYAYQFEAA  
 LLAKYLTTYSVERGVKHIVDDVQQVNLDERGWIRNVTTAEHGEIGGDLFIDCTGFRGLL  
 LNKALEEPFISFQDTLPNDSAVALQVPMDMERRGILPCTTATAQDAGWIWTIPLMGRVG  
 50 TGYVYAKDYISP EEAERTLREFVGPAAADAEANHIKMRVGRSENSWVKNCVAIGLSSGF  
 VEPLESTGIFFIHHAIEQLAKNFPGEDWNPVQRNLYNDSVAHVMDGVREFLV LHYVAAK



RS D T Q Y W R D T K T R K I P D A L A E R I E K W K T Q L P D T E T V F P Y Y H G L P P Y S Y H C V L L G M G G I D  
V K P S P A L A L G D S A A A E R E F A E I R E K T R H L T S V L P K A Y D Y F S R M R

**[00460]** SEQ.ID NO: 30

5 MSGKIDKILIVGGGTAGWMAASYLGKALQGTADITLLQAPDIPTLGVGEATIPNLQTAF  
FDFLGIPEDWWMRECNASYKVAIKFINWRTAGEGTSEARELDGGPDHFYHSFGLLKYHE  
QIPLSHYWFDRSYRGKTVEPFYACYKEPVILDANRSPRRLDGSKVTNYAWHFD AHLVA  
DFLRRFATEKLGVRHVEDRVEHVQORDANGNIESVRTATGRVFDADLFVDCSGFRGLLIN  
KAMEEPFLDMSDHLND SAVATQVPHDDDANGVEPFTSAIAMKSGWTWKI PMLGRFGTG  
10 YVYSSRFATEDEAVREFCEMWHLD PETQPLNRIRFRVGRNRRRAWVGNVCV SIGTSSCFVE  
PLESTGIYFVYAALYQLVKHFDPKSLNPVLTARFNREIETMFDDTRDFIQAHFYFSPRT  
DTPFWRANKELRLADGMQEKIDMYRAGMAINAPASDDAQLYWGNFEEEFNRNFWNNSNY  
CVLAGLGLVPDAPSPRLAHMPQATESVDEVFGAVKDRQRNLLETLP SLHEFLRQQHGR

15 **[00461]** SEQ.ID NO: 31

MSGKIDKILIVGGGTAGWMAASYLGKALQGTADITLLQAPDIPTLGVGEATIPNLQTAF  
FDFLGIPEDWWMRECNASYKVAIKFINWRTAGEGTSEARELDGGPDHFYHSFGLLKYHE  
QIPLSHYWFDRSYRGKTVEPFYACYKEPVILDANRSPRRLDGSKVTNYAWHFD AHLVA  
DFLRRFATEKLGVRHVEDRVEHVQORDANGNIESVRTATGRVFDADLFVDCSGFRGLLIN  
20 KAMEEPFLDMSDHLND SAVATQVPHDDDANGVEPFTSAIAMKSGWTWKI PMLGRFGTG  
YVYSSRFATEDEAVREFCEMWHLD PETQPLNRIRFRVGRNRRRAWVGNVCV SIGTSSCFVE  
PLESTGIYFVYAALYQLVKHFDPKSLNPVLTARFNREIETMFDDTRDFIQAHFYFSPRT  
DTPFWRANKELRLADGMQEKIDMYRAGMAINAPASDDAQLYGNFEEEFNRNFWNNSSY  
CVLAGLGLVPDAPSPRLAHMPQATESVDEVFGAVKDRQRNLLETLP SLHEFLRQQHGR

25

**[00462]** SEQ.ID NO: 32

MSGKIDKILIVGGGTAGWMAASYLGKALQGTADITLLQAPDIPTLGVGEATMPNLQTAF  
FDFLGIPEDWWMRECNASYKVAIKFINWRTAGEGTSEARELDGGPDHFYHFPGLLKYHE  
QIPLSHYWFDRLYRGKTVEPFYACYKEPVILDANRSPRRLDGSKVTSYAWHFD AHLVA  
30 DFLRRFATEKLGVRHVEDRVEHVQORDANGNIESVRTATGRVFDADLFVDCSGFRGLLIN  
KAMEEPFLDMSDHLND SAVATQVPHDDDANGVEPFTSAIAMKSGWTWKI PMLGRFGTG  
YVYSSRFATEDEAVREFCEMWHLD PETQPLNRIRFRVGRNRRRAWVGNVCV SIGTSSCFVE  
PLESTGIYFVYAALYQLVKHFDPKSFNPVLTARFNREIETMFDDTRDFIQAHFYFSPRT  
DTPFWRANKELRLADGMQEKIDMYRAGMAINAPAPDDAQLYWGNFEEEFNRNLWNNSSY  
35 CVLAGLGLVPDAPSPRLAHMPRATESVDEVFGAVKDQQRNLLETLP SLHEFLRQQHGR

**[00463]** SEQ.ID NO: 33

MSGKIDKILIVGGGTAGWMAASYLGKALQGTADITLLQAPDIPTLGVGEATHPNLQTAF  
FDFLGIPEDWWMRECNASYKVAIKFINWRTAGEGTSEARELDGGPDHFYHSFGLLKYHE  
40 QIPLSHYWFDRSYRGKTVEPFYACYKEPVILDANRSPRRLDGSKVTNYAWHFD AHLVA  
DFLRRFATEKLGVRHVEDRVEHVQORDANGNIESVRTATGRVFDADLFVDCSGFRGLLIN  
KAMEEPFLDMSDHLND SAVATQVPHDDDANGVEPFTSAIAMKSGWTWKI PMLGRFGTG  
YVYSSRFATEDEAVREFCEMWHLD PETQPLNRIRFRVGRNRRRAWVGNVCV SIGTSSCFVE  
PLESTGIYFVYAALYQLVKHFDPKSFNPVLTARFNREIETMFDDTRDFIQAHFYFSPRT  
45 DTPFWRANKELRLADGMQEKIDMYRAGMAINAPASDDAQLYGNFEEEFNRNFWNNSSY  
CVLAGLGLVPDAPSPRLAHMPRATESVDEVFGAVKDQQRNLLETLP SLHEFLRQQHGR

**[00464]** SEQ.ID NO: 34

MSGKIDKILIVGGGTAGWMAASYLGKALQGTADITLLQAPDIPTLGVGEATIPNLQTAF  
 FDFLGIPEDWMMRECNASYKVAIKFINWRTAGEGTSEARELDGGPDHIFYHSFSLLYHE  
 QIPLSHYWFDRSYRGKTVEPFYACYKEPVILDANRSPRRLDGSKVTNYAWHFD AHLVA  
 5 DFLRRFATEKLGVRHVEDRVEHVQRDANGNIESVRTATGRVFDADLFVDCSGFRGLLIN  
 KAMEEPFLDMSDHLND SAVATQVPHDDDANGVEPFTSAIAMKSGWTWKIPMLGRFGTG  
 YVYSSRFATEDEAVREFCEMWHLDPETQPLNRIRFRVGRNRRRAWVGNVCSIGTSSCFVE  
 PLESTGIYFVYAALYQLVKHFDPKSLNPVLTARFNREIETMFDDTRDFIQAHFYFSPRT  
 10 DTPFWRANKELRLADGMQEKIDMYRAGMAINAPASDDAQLYYGNFEEEFNRNFWNNSYY  
 CVLAGLGLVPDAPSPRLAHMPQATESVDEVFGAVKDRQRNLLETLP SLHEFLRQQHGR

**[00465]** SEQ.ID NO: 35

MSGKIDKILIVGGGTAGWMAASYLGKALQGTADITLLQAPDIPTLGVGEATIPNLQTAF  
 FDFLGIPEDWMMRECNASYKVAIKFINWRTAGEGTSEARELDGGPDHIFYHSFGLLYHE  
 15 QIPLSHYWFDRSYRGKTVEPFYACYKEPVILDANRSPRRLDGSKVTNYAWHFD AHLVA  
 DFLRRFATEKLGVRHVEDRVEHVQRDANGNIESVRTATGRVFDADLFVDCSGFRGLLIN  
 KAMEEPFLDMSDHLND SAVATQVPHDDDANGVEPFTSAIAMKSGWTWKIPMLGRFGTG  
 YVYSSRFATEDEAVREFCEMWHLDPETQPLNRIRFRVGRNRRRAWVGNVCSIGTSSCFVE  
 PLESTGIYFVYAALYQLVKHFDPKSLNPVLTARFNREIETMFDDTRDFIQAHFYFSPRT  
 20 DTPFWRANKELRLADGMQEKIDMYRAGMVINAPASDDAQLYYGNFEEEFNRNFWNNSYY  
 CVLAGLGLVPDAPSPRLAHMPQATESVDEVFGAVKDRQRNLLETLP SLHEFLRQQHGR

**[00466]** SEQ.ID NO: 36

ATGATGTCTGAAAATTTGCAATTGTCAGCTGAAGAAATGAGACAATTGGGTTACCAAGC  
 25 AGTTGATTTGATCATCGATCACATGAACCATTTGAAGTCTAAGCCAGTTTCAGAAACAA  
 TCGATTCTGATATCTTGAGAAATAAGTTGACTGAATCTATCCCAGAAAATGGTTCAGAT  
 CCAAAGGAATTGTTGCATTTCTTGAACAGAAACGTTTTTAATCAAATTACACATGTTGA  
 TCATCCACATTTCTTGGCTTTTGTTCAGGTCCAAAATAATTACGTTGGTGTGTTGTCAG  
 ATTTCTTGGCTTCTGGTTTTTAATGTTTTTCCAACGTCATGGATTGCTGGTGCAGGTGCT  
 30 GAACAAATCGAATTGACTACAATTAATTGGTTGAAATCTATGTTGGGTTTTCCAGATTC  
 AGCTGAAGGTTTTATTGTTTTCTGGTGGTTCAATGGCAAATTTGACAGCTTTGACTGTTG  
 CAAGACAGGCTAAGTTGAACAACGATATCGAAAATGCTGTTGTTTACTTCTCTGATCAA  
 ACACATTTCTCAGTTGATAGAGCATTGAAGTTTTAGGTTTTAAACATCATCAAATCTG  
 TAGAATCGAAACAGATGAACATTTGAGAATCTCTGTTTCAGCTTTGAAGAAACAAATTA  
 35 AAGAAGATAGAACTAAGGGTAAAAAGCCATTCTGTGTTATTGCAAATGCTGGTACTACA  
 AATTGTGGTGCTGTTGATTCTTTGAACGAATTAGCAGATTTGTGTAACGATGAAGATGT  
 TTGGTTGCATGCTGATGGTTCTTATGGTGCTCCAGCTATCTTGTCTGAAAAGGGTTTCAG  
 CTATGTTGCAAGGTATTCATAGAGCAGATTCTTTGACTTTAGATCCACATAAGTGGTTG  
 TTCCAACCATACGATGTTGGTTGTGTTTTGATCAGAACTCTCAATATTTGTCAAAGAC  
 40 TTTTAGAATGATGCCAGAATACATCAAGGATTCAGAACTAACGTTGAAGGTGAAATTA  
 ATTTCCGGTGAATGTGGTATCGAATTGTCAAGAAGATTCAGAGCTTTGAAGGTTTGGTTG  
 TCTTTTAAAGTTTTCCGGTGTGCTGCTTTTAGACAAGCAATCGATCATGGTATCATGTT  
 AGCAGAAACAGTTGAAGCATTTTTGGGTAAAGCAAAAGATTGGGAAGTTGTTACACCAG  
 CTCAATTGGGTATCGTTACTTTTAGATACATTCATCTGAATTGGCATCAACAGATACT  
 45 ATTAATGAAATTAATAAGAAATGGTTAAGGAAATCACACATAGAGGTTTTGCTATGTT  
 ATCTACTACAGAATTGAAGGAAAAGGTTGTTATTAGATTGTGTTCAATTAATCCAAGAA  
 CTACAACCTGAAGAAATGTTGCAAATCATGATGAAGATTAAAGCATTGGCTGAAGAAGTT  
 TCTATTTTCATACCCATGTGTTGCTGAATAA

**[00467] SEQ.ID NO: 37**

MMSENLQLSAEEMRQLGYQAVDLIIDHNMHLKSKPVSETIDSDILRNKLTESIPENGSD  
 PKELLHFLNRNVFNQITHVDHPHFLAFVPGPNYVGVVADFLASGFNVFPTAWIAGAGA  
 EQIELTTINWLKSMGLGFPDSAEGLFVSGGSMANLTALTVARQAKLNNDIENAVVYFSDQ  
 5 THFSVDRALKVLGFKHHQICRIETDEHLRISVSALKKQIKEDRTKGKKPFCVIANAGTT  
 NCGAVDSLNELADLCNDEDVWLHADGSYGAPAILSEKGSAMLQGIHRADSLTLDPHKWL  
 FQPYDVGCVLIRNSQYLSKTFRMMPEYIKDSETNVEGEINFGECGIELSRRFRALKVWL  
 SFKVFGVAAFRQAIDHGIMLAEQVEAFLGKAKDWEVVTPAQLGIVTFRYIPSELASTDT  
 10 INEINKKLVKEITHRGFAMLSTTELKEKVIRLCSINPRTTTEEMLQIMMKIKALAEV  
 SISYPCVAE

**[00468] SEQ.ID NO: 38**

ATGTCTCAAGTGATCAAAAAAAGAGAAACACTTTTCATGATCGGTACCGAATATATATT  
 GAACTCTACCCAATTAGAAGAGGCTATTAAATCTTTCGTCCATGATTTCTGTGCCGAAA  
 15 AGCACGAAATTCACGACCAACCAGTTGTCGTTGAAGCTAAGGAACACCAAGAAGATAAG  
 ATTAAGCAAATCAAAATTCAGAAAAGGGTCGTCCAGTCAATGAAGTTGTTTCTGAAAT  
 GATGAACGAAGTCTACAGATACAGAGGTGATGCCAACCATCCAAGATTCTTCTCTTTTCG  
 TTCCAGGTCCAGCTTCCTCTGTCTCCTGGTTGGGTGACATCATGACTTCTGCTTACAAC  
 ATCCACGCTGGTGGTTCTAAATTGGCCCCAATGGTTAACTGTATCGAACAAAGAGTCTT  
 20 GAAATGGCTCGCAAAGCAAGTTGGTTTACCGAAAAACCCTGGTGGTGTTCGTTTCCG  
 GTGGTTCTATGGCTAACATTACTGCTTTGACAGCCGCTAGAGACAACAAGCTTACCGAT  
 ATTAAGTTGCACTTGGGTACTGCTTACATTTCCGACCAAACCCACTCTTCAGTTGCTAA  
 GGGTTTACGTATCATTTGGTATCACTGACTCCAGAATCAGAAGAATTCACAACTAACTCCC  
 ACTTCCAAATGGACACCACTAAGTTGGAAGAAGCTATCGAAACCGACAAGAAGTCCGGT  
 25 TACATCCCATTTGTTGTTATCGGTACTGCCGGTACCACCAACACTGGTTCCATTGACCC  
 ATTGACTGAAATTTCCGCTTTGTGTAAGAAGCATGACATGTGGTTCCACATTGATGGTG  
 CCTACGGTGCTTCAGTTTTGTTGTCTCCAAAGTACAAGTCTTTGTTAACAGGTACCGGT  
 TTGGCTGACTCTATCTCTTGGGATGCTCACAAGTGGTTATTTCAAACCTACGGTTGTGC  
 TATGGTTTTGGTCAAGGATATTAGAACTTGTTCATTCTTCCACGTCAATCCAGAAT  
 30 ACTTGAAGGATCTGGAAAACGACATCGATAACGTTAACACCTGGGACATCGGCATGGAA  
 TTGACCAGACCAGCCAGAGGTTTGAAGTTGTGGTTGACTTTGCAAGTTCTAGGTTCTGA  
 TCTTATTGGTTCTGCCATTGAACACGGTTTCCAATTGGCTGTCTGGGCTGAAGAAGCCT  
 TGAATCCAAAGAAGGACTGGGAAATTGTCTCTCCTGCTCAAATGGCTATGATCAACTTC  
 AGATATGCTCCAAAGGACTTGACCAAGGAAGAACAAGATATCTGAACGAAAAGATCTC  
 35 TCACCGTATTTTGAATCCGGTTACGCTGCTATCTTCACTACCGTCCTAAACGGTAAGA  
 CTGTCTTGAGAATCTGCGCTATTCACCCAGAAGCTACTCAAGAAGACATGCAACACACC  
 ATTGATTTGTTGGACCAATACGGTAGAGAAATCTACACTGAAATGAAGAAGGCTTAG

**[00469] SEQ.ID NO: 39**

40 MSQVIKKKRNTFMIGTEYIILNSTQLEEAIKSFVHDFCAEKHEIHDQPVVVEAKEHQEDK  
 IKQIKIPEKGRPVNEVVSEMMNEVYRYRGDANHPRFFSFVPGPASSVSWLGDIMTSAYN  
 IHAGGSKLAPMVNCEIQEVLKWLAKQVGFTENPGGVFVSGGSMANITALTAARDNKLTD  
 INLHLGTAYISDQTHSSVAKGLRIIGITDSRIRRIPTNSHFQMDTTKLEEIETDKKSG  
 YIPFVVIGTAGTTNTGSIDPLTEISALCKKHDMWFHIDGAYGASVLLSPKYKSLLTGTG  
 45 LADSISWDAHKWLFQTYGCAMVLVKDIRNLFHSFHVNP EYLKDLENDIDNVNTWDIGME  
 LTRPARGKLWLTLQVLGSDLIGSAIEHGFQLAVWAEALNPKKDWEIVSPAQMAMINF  
 RYAPKDLTKEEQDILNEKISHRILESGYAAI FTTVLNGKTVLRICAIHPEATQEDMQHT  
 IDLLDQYGREIYTEMKKA

**[00470] SEQ.ID NO: 40**

ATGAAGCCAGCTGAAGCTAAACCACCTCACATGGACCATGATACTTTCAGATCCTTGGG  
TCACCAAGCTATTGACTGGATTGCTGATTACTGGCAACACTTGGCTGACAGACCAGTTG  
CTCCATTGGTTGAACCAGGTACTGTTAGAGCCCAATTGCCTACCACTCCACCAGAGTAC  
5 GGTGAAGACTTCCCAGCTTTGTTATCTGACTTGGAAAGAATTGTCTTGCCAGGCTTGT  
GCACTGGCAACACCCAAGATTCTTCGGTTACTTCCCAGCTAACGCTTCCGGTCCAGCCG  
TTTTAGCTGAATTATTATCTGCTGGTCTAGGTGTCCAAGGTATGAACTGGAACACTTCT  
CCAGCTTGTACCGAAGTTGAACAACAAATGTTGGACTGGTTCGTTTCATTTGTTAGGGTT  
GCCAGAACACTTCAGAGGTGGTGGTGGTATCCAAGACACTGCCTCCTCCGCCGTTTTGG  
10 TCGCTCTATTGACTGCTTTGCACAGAGCTTCTGCTGGTAGAACCAGAGATCACGGTACT  
GGTGAATGTGGTTACAGGGTTTACCTAACCCTGAACTCACTCCGCTGCTAGAAAGGC  
TGCCGTCATCACCGGTTTGGGTTTGCCTGCTGTTTGTGAAGTCGCAACTGATGCTGACG  
GTGCTATGGATGCCGCTGACTTGGCCGCTCGTATTCAAGCCGATTTGGCTGCCGGTTTG  
ACCCCATTGATGGTTGTTGCTACCCGTGGTACCACCTCTGATTTGTCTTCGATCCATT  
15 GGAAGACATCGGTCCAGTCTGTAGAAGACATGGTGTCTGGTTACACGTTGACGCTGCTT  
ATGCTGGTGGTGGTGCAGTGTGCGATGAATTGCATTGGGTAAATGATGGTGTAGATAC  
GCTGATTCTTACTGTACTAACCACACAAGTGGTTGTTGACCAACTTCGACTGTGATTT  
ATTATGGGTTGCCGACCCAGAAGCTTTGGTCAACGCTTTGTCCGTTTTGCCTGCTTACT  
TGAGAAACCCAGCTTCTGAATCAGGTAGAGTTACTGACTACAGACACTGGCAAGTTCCA  
20 TTGGGCAGAAGATTACAGAGCCTTGAAGTTGTGGTCTGTCTTAAGATGGTACGGTGCCGA  
AGGTTTGGAGATCTCACATCAGAACTGGTGTGACACACGCCAGATTGTTCCGCTGACCTGG  
TCAGAGAAGATGACCGTTTACCTTGTGGCGCCAGCTACTTTGGGTTTGGTCACTTTC  
CGTCACGCTGGTTCTGACACCGACAACAGAACTTGTGTCACGCTATCAACGCCGAAGG  
TACAACTTTTTAACTCATAGTGAAGAAGACGGTACCTTTTTCTTGAGATTTGCTACCG  
25 GTGGTACCCACACTGAAGACCATCACGTCCACGAAGCTTGGAGAGCTGTCCAAATGCC  
GTCCAAACGCCGTCCCAAGAGCCCAACATGCCACCGATGGTACTGCCGATGCTTTGAA  
CGGTTAA

**[00471] SEQ.ID NO: 41**

MKPAAEAKPPHMDHDTFRSLGHQAIDWIADYWQHLADRPVAPLVEPGTVRAQLPTTPPEY  
GEDFPALLSDLERIVLPGLLHWQHPRFFGYFPANASGPAVLAEILLSAGLGVQGMNWNNTS  
PACTEVEQQMLDWFVHLLGLPEHFRGGGVIQDTASSAVLVALLTALHRASAGRTRDHGT  
GECGYRVYLTAETHSAARKAAVITGLGLRAVCEVATDADGAMDAADLAARIQADLAAGL  
TPLMVVATRGTSSDLSFDPLEDIGPVCRRHGVWLHVDAAYAGVAAVCDELHWVNDGVRY  
35 ADSYCTNPHKWLTLNFDCLLWVADPEALVNALS VLPAYLRNPASESGRVTDYRHWQVP  
LGRFRFRALKLWSVLRWYGAEGLRSHIRTGVRHARLFADLVREDDRFTLLAPATLGLVTF  
RHAGSDTDNRNLLHAINAEGTTFLTHSEKNGTFFLR FATGGTHTEDHHVHEAWRAVQNA  
VQNAVPR AQHATDGTADALNG

**40 [00472] SEQ.ID NO: 42**

ATGGATATTGAAGCTTTTAGAAAAGCAGGTTACCAAGCTATCGATAGAATTTGTGACTA  
CTACTACTCTTTGCAAAATAGACCAGTTGTTCCATCAGTTCAACCAGGTTATTTGTTAG  
ATGCATTACCAGATTCTCCACCAGAACAAAGGTGAAGATTTCACTGTTATTGCTGATGAT  
TACCAAAAAGTACATCTTGCCAGGTTTGACACATTGGCAACATCCATCTTTCTTTGCATA  
45 CTTTCCAACAGCTTGACTTTTGAAGGTATTTTGGGTGACTTGTATTCTACATCAACTG  
CAAATCCAGGTTTTAATTGGTTGGCTTCTCCAGCATGTACTGAATTAGAAATGGTTGTT  
ATGGATTGGTCTGCTAAGTTGTTAGGTTTGTGAGAACATTTCTTGCATTCTTCAGGTAA  
AGGTGGTGGTGTATTCAAACACTACAGCATCAGAATTGGCTTTAGTTGTTGTTGCTG

CAAGAGAAAGATACTTGAGAATTCATCCAGATGCTAAGGCAGATGAATTGGTTATATAT  
 ACTACAACTCAAACACATTCTTTGGGTGTTAAAGCTGGTTTAGTTTTTGGTATGGAATG  
 TAGAGCATTTGGAAGTTAAAGCTGAAGATGCTTACGCATTAAGAGGTGCAACTTTGAAGT  
 CAGCTTTAGAAGAAGATGAAAAGAGAGGTAAAAGACCTTTTATTTTGGTTGCAACAGTT  
 5 GGTACAACTTCTTCAGGTGCTATTGATAGATTGGATGAAATTGGTCAAGTTTCTGAAGA  
 TTATCCATCATTATGGATTTCATGTTGATGCTGCATGGGCAGGTGTTACTTTGGCTTGTC  
 CAGAATACAGAGGTACAGCTCAATTAGAAAACATCAACGCTTATGCAACATCTTTCGGT  
 ACTAACTTCCATAAGTGGGGTTTGGTTAACTTTGATGCTGCATTGTTGTGGGTAAAGGA  
 TAGAAAGGATTTGACAGATGCTTTAGATGTTACTCCAGAATTCTTGAGAACAAAACAAG  
 10 GTGACGCTGGTGCAGTTGTTGATTTTAGAAATTGGCATTGTTGGGTTTGGGTAGAAGATTC  
 AGATCTTTGAAGGTTTGGTTCGTTTTGAGATCATAACGGTGTGGAAGGTTTGAAGAACTA  
 CATCAGACAAGGTATTAATGAATGAACATTTTACTTCTTTAATTAGAGCTTCTTTGG  
 ATTTTTTCATTAGTTACAGCACCATCATTTCGCTTTGACAGTTTTTAGATTGACTCCAGCT  
 GGTGCATCTTTGACTGGTTCAGAATTGAACGAATTAAACAGAGCTTTCTACGCAAGATT  
 15 ATCTTCAAGACATGATATCATGTTGACACAACTGTTTTGAACGGTGTGTTCTGTATCA  
 GATTTGCAGTTGGTGCTGCAAGAACTCAACAAGAACATATTGATACAGCTTGGGATTTG  
 TTACAACAAGAAGGTGCTGTTGCAGTTCAAGAATACATGCAAAAGATTTCTAAGGATGT  
 TCAATAA

20 **[00473]** SEQ.ID NO: 43

MDIEAFRKAGYQAIDRICDYYSLQNRPVVPSVQPGYLLDALPDSPPEQGEDFTVIADD  
 YQKYILPGLTHWQHPSFFAYFPTACTFEGILGDLYSTSTANPGFNWLASPACTELEMVV  
 MDWSAKLLGLSEHFLHSSGKGGGVIQTTASELALVVVAARERYLRIHPDAKADELVIY  
 TTTQTHSLGVKAGLVFGMECRALEVKAEDAYALRGATLKSALEEDEKRGKRPFILVATV  
 25 GTTSSGAIDRLDEIGQVSEDYPSLWIHVDAAWAGVTLACPEYRGTAQLENINAYATSFG  
 TNFHKWGLVNFDAALLWVKDRKDLTDALDVTPEFLRTKQGDAGAVVDFRNWHLGLGRRF  
 RSLKVVFLRSYGVGEFRNYIRQGIKLNEHFTSLIRASLDFSLVTAPSFAITVFRLLTPA  
 GASLTGSELNELNRAFYARLSSRHDIMLTQTVLNGVFCIRFAVGAARTQQEHIDTAWDL  
 LQQEGAVAVQEYMQKISKDVQ

30

**[00474]** SEQ.ID NO: 44

ATGAAGTTCTGGAGAAAGTACACACAACAAGAAATGGATGAAAAGATTACTGAATCTTT  
 GGAAAAGACTTTGAACTACGATAACACTAAGACAATCGGTATTCCAGGTACTAAGTTGG  
 ATGATACAGTTTTCTATGATGATCATTCTTTCGTTAAGCATTACCCATACTTGAGAACT  
 35 TTTATTCAAACCCAAACCATATCGGTTGTCATACTTATGATAAGGCTGATATCTTGTT  
 CGGTGGTACATTTCGATATCGAAAGAGAATTAATCCAATTGTTAGCAATCGATGTTTTGA  
 ACGGTAACGATGAAGAATTTGATGGTTACGTTACTCAAGGTGGTACAGAAGCTAACATC  
 CAAGCAATGTGGGTTTACAGAACTACTTCAAGAAAGAAAGAAAGGCTAAGCATGAAGA  
 AATCGCTATCATCACTTCAGCAGATACACATTACTCTGCATACAAAGGTTGAGATTTGT  
 40 TGAACATCGATATTATTAAGGTTCCAGTTGATTTTTATTCAAGAAAATTCAAGAAAAT  
 ACATTGGATTCAATTGTTAAAGAAGCTAAAGAAATTGGTAAAAAGTACTTCATCGTTAT  
 CTCTAACATGGGTACTACAATGTTTGGTTCAGTTGATGATCCAGATTTGTACGCTAACA  
 TCTTCGATAAGTACAATTTGGAATACAAAATTCATGTTGATGGTGCATTTGGTGGTTTT  
 ATATATCCAATTGATAATAAGGAATGTAAACTGATTTCTCTAATAAGAAGCTTTCTTC  
 45 AATCACATTAGATGGTCATAAGATGTTGCAAGCTCCATACGGTACTGGTATCTTCGTTT  
 CAAGAAAGAATTTGATCCATAACACTTTGACAAAGGAAGCAACTTACATCGAAAATTTG  
 GATGTTACATTGTCTGGTTCAAGATCTGGTTCAAATGCTGTTGCAATTTGGATGGTTTT  
 AGCTTCTTATGGTCCATACGGTTGGATGGAAAAGATTAATAAGTTGAGAAATAGAATA  
 AATGGTTGTGTAAGCAATTGAACGATATGAGAATTAATATTACAAAGAAGATTCAATG

AATATTGTTACAATTGAAGAACAATATGTTAATAAGGAAATCGCTGAAAAGTACTTTTT  
 AGTTCCAGAAGTTCATAACCCAACTAACAACTGGTACAAGATCGTTGTTATGGAACATG  
 TTGAATTGGATATCTTGAACCTTTTGGTTTACGATTTGAGAAAGTTTAATAAGGAACAT  
 TTGAAGGCAATGTAA

5

**[00475] SEQ.ID NO: 45**

MKFWRKYTQQEMDEKITESLEKTLNYDNTKTI GIPGTKLDDTVFYDDHSFVKHSPYLRT  
 FIQNPNHIGCHTYDKADILFGGTFDIERELIQLLAIDVLNGNDEEFDGYVTQGGTEANI  
 QAMWVYRNYFKKERKAKHEEIAIITSADTHYSAYKGSDDLNIIDIIKVPVDFYSRKIQEN  
 10 TLDSIVKEAKEIGKKYFIVISNMGTTMFGSVDDPDLYANIFDKYNLEYKIHVDGAFGGF  
 IYPIDNKECKTDFSNKNVSSITLDGHKMLQAPYGTGIFVSRKNLIHNTLTKEATYIENL  
 DVTLSGSRSGSNAVAIWMVLASYGPYGWMEKINKLRNRTKWLCQLNDMRIKYYKEDSM  
 NIVTIEEQYVNKEIAEKYFLVPEVHNPTNNWYKIVVMEHVELDILNSLVYDLRKFNKEH  
 LKAM

15

**[00476] SEQ.ID NO: 46**

ATGAATGCATCTGAATTCAGAAGAAGAGGTAAAGAAATGGTTGATTACATGGCTAACTA  
 CATGGAAGGTATTGAAGGTAGACAAGTTTATCCAGATGTTGAACCAGGTACTTGAGAC  
 CATTAAATCCAGCTGCAGCTCCACAAGAACCAGATACATTTGAAGATATTATTAACGAT  
 20 GTTGAAAAGATTATCATGCCAGGTGTTACACATTGGCATTACCATATTTCTTTGCATA  
 CTTTCCAAGTCTTCTTCATATCCAGCAATGTTGGCTGATATGTTATGTGGTGCAATTG  
 GTTGTATTGGTTTTTCTTGGGCAGCTTCACCAGCTTGTACAGAATTGGAACTGTTATG  
 ATGGATTGGTTAGGTAAAATGTTGGAATTACCAAAGGCATTTTTGAACGAAAAGGCTGG  
 TGAAGGTGGTGGTGTATTCAAGGTTCTGCATCAGAAGCTACATTAGTTGCTTTGTTAG  
 25 CAGCTAGAACTAAAGTTATTCATAGATTGCAAGCAGCTTCTCCAGAATTAACACAAGCA  
 GCTATCATGGAAAAATTGGTTGCATATTCTTCAGATCAAGCTCATTCTTCAGTTGAAAG  
 AGCTGGTTTGATCGGTGGTGTAAAGTTGAAGGCAATCCCATCTGATGGTAATTTTGCTA  
 TGAGAGCATCAGCTTTGCAAGAAGCATTAGAAAGAGATAAAGCAGCTGGTTTGATTCCA  
 TTTTTTCATGGTTGCTACTTTAGGTACTACAAGTTGTTGTTCTTTCGATAAATTTGTTAGA  
 30 AGTTGGTCCAATTTGTAATAAGGAAGATATTTGGTTGCATGTTGATGCAGCTTACGCAG  
 GTTCAGCTTTTATTTGTCCAGAATTCAGACATTTGTTGAACGGTGTGAATTCGCTGAT  
 TCTTTTAATTTCAATCCACATAAATGGTTGTTAGTTAACTTCGATTGTTTCAGCAATGTG  
 GGTAAAGAAAAGAACAGATTTGACTGGTGCTTTTAGATTAGATCCAACATATTTGAAGC  
 ATTCTCATCAAGATTCAGGTTTGATCACTGATTACAGACATTGGCAAATTCATTGGGT  
 35 AGAAGATTGAGATCTTTGAAGATGTGGTTCGTTTTTTAGAATGTACGGTGTAAAGGGTTT  
 GCAAGCTTACATCAGAAAGCATGTTCAATTGTCTCATGAATTCGAATCATTAGTTAGAC  
 AAGATCCAAGATTCGAAATCTGTGTTGAAGTTATTTTGGGTTTAGTTTGTTTCAGATTG  
 AAGGGTTCTAATAAGGTTAACGAAGCATTGTTGCAAAGAATTAATTCAGCTAAGAAAAT  
 TCATTTGGTTCCATGTCATTTGCGTGATAAGTTTCGTTTTGAGATTCGCTATTTGTTCAA  
 40 GAACTGTTGAATCAGCACATGTTCAAAGAGCTTGGGAACATATCAAGGAATTGGCAGCT  
 GATGTTTTGAGAGCTGAAAGAGAATAA

**[00477] SEQ.ID NO: 47**

MNASEFRRRGKEMVDYMANYMEGIEGRQVYPDVEPGYLRPLIPAAAPQEPDTFEDIIND  
 45 VEKIIMPGVTHWHSPYFFAYFPTASSYPAMLADMLCGAIGCIGFSWAASPACTELETVM  
 MDWLGKMLELPKAFLEKAGEGGGVIQGSASEATLVALLAARTKVIHRLQAASPELTQA  
 AIMEKLVAYSSDQAHSSVERAGLIGGVKLKAIPSDGNFAMRASALQEALERDKAAGLIP  
 FFMVATLGTTCSSFDNLLEVGPICNKEDIWLHVDAAYAGSAFICPEFRHLLNGVEFAD

SFNFNPHKWLLVNFDCSAMWVKKRTDLTGAFRLDPTYLKHS HQDSGLITDYRHWQIPLG  
RRFRSLKMWFVFRMYGVKGLQAYIRKHVQLSHEFESLVRQDPRFEICVEVILGLVCFRL  
KGSNKVNEALLQRINS AKKIHLVPCHLRDKFVLRFAICSRTVESAHVQRAW EHIKELAA  
DVLRAERE

5

**[00478] SEQ.ID NO: 48**

ATGGGTTCTATCTCAGAAAAC TGTGATGATTCTATTT CATTAGCTGCACCTTTTAGACC  
ATTGGAACCAGAAGAATT CAGAAAGCAAGCTCATGTTATGGTTGATTT CATCGCAGATT  
ACTACAAGAACATTGAAAATTATCCAGTTTGT CACAAGTTGAACCAGGTTATTTGAAA  
10 AATAGATTGCCAGAACTGCTCCACATTTGCCAGAACTTT CGAAACAATCTTGAAGGA  
TATTAAGAAAGATATCGTTCCTGGTATGACTAATTGGTTGT CACCAAATTTCTTGTCTT  
ATTTTCCAGCAACTGTTTCTTCAGCTGCATTTGTTGGTGAAATGTTGTGTACAGGTTT  
AATTCTGTTGGTTTAAATTGGTTAGCTTCACCAGCATCTACAGAATTGGAAATGGTTGT  
TATTGATTGGTTAGCTAACATGTTGAAGTTGCCAAAGTCTTTTATGTTTCATGGTACTG  
15 GTGGTGGTGTATTCAAGGTACTACATCTGAAGCTATCTTGTGTACATTGATTGCTGCA  
AGAGATGGTGCTTTGGAAAAGATTGGTATGGAAAACGTTGGTAAATTAGTTGTTTACGG  
TTCAGATCAAAC TATTCTTTCTTTCAAAGACTTGTAAGGTTGCTGGTATCTTCCCAT  
GTAACATCAAGTTGATCCCAACTACAAGAGAAGATAACTTCTCTATGTCACCAATCGCA  
TTGAGAGAACAAATTGAAGCTGATGTTGCAGATGGTTTAGTTCCAATTTTCTTGTGTAC  
20 TACAGTTGGTACTACATCAACTGCTGCAATTGATCCAGTTTCTGAAGTTGCTAAGGTTG  
CAAACGATTTCAATATTTGGGTT CATGTTGATGCTGCTTATGCTGGTTCTGCATGTATT  
TGTCCAGAATTCAGACAATACTTAGATGGTATCGAATTGGTTGATTCTTTTTTCATTGTC  
TCCACATAAGTGGTTGTTGTGTTTCTTGGATTGTTGTTGTTTGTGGTTGAAGAAACCAC  
ATTTGATGGTTAAGGCTTTGTCAACTAACC CAGAATATTTGAGAAATAAGAGATCTGAA  
25 TTCGATGGTGTGTTGATTTC AAGGATTGGCAAATCGGTACAGGTAGAAGATTCAAAGC  
TTTGAGATTATGGTTGGTTATGAGATCATACGGTGTGAAAATTTGAAGAGACATATCT  
TGTCTGATGTTCAAATGGCTAAGATGTT CGAAGGTTAGTTAAATCTGATCCAAGATTC  
GAAATCATCGTTCCAAGAGCTTTCGCATTAGTTTGTTCAGATTGAACCTGGTAAAGG  
TTATGATGATGAAATCGATAAGGAAATCTTGAATAAGGAATTGTTAGATTTGATTAATT  
30 CAACAGGTAGAGCTTACATGACTCATACAAAAGCAGGTGGTATCTATATGTTGAGATTT  
GCTGTTGGTACTACATTGACTGAAGAACATCATGTTTACGCTGCATGGGAATTGATTAA  
AGAATGTACAGATGCATCTTTGACTAAAACAAATATTATTGAATAA

**[00479] SEQ.ID NO: 49**

MGSISENCDD SISLAAPFRPLEPEEFRKQAHVMVDFIADYYKNIENYPVLSQVEPGYLK  
35 NRLPETAPHL PESFETILKDIKKDIVPGMTNWLSPNFFAYFPATVSSAA FVGEMLC TGF  
NSVGFNWLAS PASTELEMVVIDWLANMLKLPKSFMFHGTGGGVIQ GTTSEAILCTLIAA  
RDGALEKIGMENVGKLVVYGS DQTHSFFQKTCKVAGIFPCNIKLIPPTREDNFSMSPIA  
LREQIEADVADGLVPIFLCTTVGTTSTAAIDPVSEVAKVANDFN IWHVDAAYAGSACI  
40 CPEFRQYLDGIELVDSFSLSPHKWLLCFLDCCCLWLKKPHLMVKALSTNPEYLRNKRSE  
FDGVVDFKDWQIGTGRRFKALRLWLMRSYGVENLKRHILSDVQMAKMFEGLVKSDPRF  
EIIIVPRAFALVCFRLNPGKGYDDEIDKEILNKELLDLINSTGRAYMTHTKAGGIYMLRF  
AVGTTLTEEHHVYAAWELIKECTDASLTKNII E

45 **[00480] SEQ.ID NO: 50**

ATGGGTT CATTAGATACAAATCCAAC TGTCTTTTCTGCATTTCCAGCTGGTGAAGGTGA  
AACTTTTCAACCATG AACGCAGATGATGTTAGATCATATTTGCATAAGGCTGTTGATT  
TCATCTCTGATTACTACAAATCAGTTGAATCTATGCCAGTTT TGCCAAACGTTAAGCCA

5 GGTTATTTGCAAGATGAATTAAGAGCATCACCACCAACTTACTCTGCTCCATTCGATGT  
 TACAATGAAGGAATTGAGATCTTCAGTTGTTCCCTGGTATGACTCATTGGGCATCTCCAA  
 ATTTCTTTGCTTTCTTTCCATCAACAAATCTGCTGCAGCTATTGCTGGTGAAGTTGATT  
 GCATCAGCTATGAATACAGTTGGTTTTACTTGGCAAGCTTCTCCAGCAGCTACTGAAAT  
 10 GGAAGTTTTGGCATTAGATTGGTTGGCTCAAATGTTGAATTTGCCAACTTCTTTTATGA  
 ACAGAACTGGTGAAGGTAGAGGTACTGGTGGTGGTGTATTTTGGGTACTACATCTGAA  
 GCTATGTTGGTTACATTAGTTGCAGCTAGAGATGCAGCTTTGAGAAGATCTGGTTCAGA  
 TGGTGGTGCAGGTTTGCATAGATTAGCTGTTTATGCAGCTGATCAAACACATTCAACTT  
 TCTTTAAGGCATGTAGATTGGCTGGTTTTGATCCAGCAAATATTAGATCTATTCCAACA  
 15 GGTGCTGAAACTGATTATGGTTTAGATCCAGCAAGATTGTTAGAAGCTATGCAAGCAGA  
 TGCTGATGCAGGTTTAGTTCCAACATACGTTTGTGCAACTGTTGGTACTACATCTTCAA  
 ATGCTGTTGATCCAGTTGGTGTGCTGTTGCAGATGTTGCAGCTAGATTTGCAGCTTGGGT  
 CATGTTGATGCAGCTTACGCAGGTTAGCTTGTATTTGTCCAGAATTACAGACATCATTT  
 GGATGGTGTGAAAGAGTTGATTCTATCTCAATGTCTCCACATAAGTGGTTGATGACAT  
 20 GTTTAGATTGTACTTGTGTGACGTTAGAGATACACATAGATTGACTGGTCTTTAGAA  
 ACTAACCAGAATACTTGA AAAATCATGCTTCAGATTCTGGTGAAGTTACAGATTTGAA  
 AGATATGCAAGTTGGTGTGTTGGTAGAAGATTCAGAGGTTTGAAGTTGTGGATGGTTATGA  
 GAACTTATGGTGTGCTAAGTTGCAAGAACATATCAGATCAGATGTTGCTATGGCAAAA  
 GTTTTTGAAGATTTGGTTAGAGGTGACGATAGATTCTGAAGTTGTTGTTCCAAGAACTT  
 25 CGCTTTGTTTTGTTTCAGAATTAGAGCTGGTGCAGGTGCAGCTGCAGCTACAGAAGAAG  
 ATGCAGATGAAGCTAACAGAGAATTGATGGAAAGATTGAATAAGACTGGTAAAGCATAAC  
 GTTGCTCATAAGTTGTTGGTGGTAGATTGTTTTGAGATTTGCAGTTGGTTCTTCATT  
 ACAAGAAGAACATCATGTTAGATCTGCTTGGGAATTGATTAAGAAAACACTACAAGTAAA  
 TGATGAACTAA

25

**[00481] SEQ.ID NO: 51**

30 MGSLDTNPTAFSAFPAGEGETFQPLNADDVRSYLHKAVDFISDYKSVESMPVLPNVKP  
 GYLQDELRASPTYSAPFDVTMKELRSSVVPGMTHWASPNFFAFFPSTNSAAAIAGDLI  
 ASAMNTVGFTWQASPAATEMEVLALDWLAQMLNLPSTFMNRTGEGRGTTGGGVILGTTSE  
 35 AMLVTLVAARDAALRRSGSDGVAGLHRLAVYAADQTHSTFFKACRLAGFDPANIRSIPT  
 GAETDYGLDPPARLLEAMQADADAGLVPTYVCATVGTSSNAVDPPVAVADVAARFAAWV  
 HVDAAYAGSACICPEFRHHLDGVERVDSISMSPHKWLMTCLDCTCLYVRDTHRLTGSL  
 TNPEYLNKHNASDSGEVTDLKDMQVGVGRFRGLKLWVMRTYGVAKLQEHIRSDVAMAK  
 VFEDLVRGDDRFEVVPRNFALVCFRIRAGAGAAAATEEDADEANRELMERLNKTGKAY  
 40 VAHTVVGGRFVLRFAVGSSSLQEEHHVRSAWELIKKTTTEMMN

35

**[00482] SEQ.ID NO: 52**

40 ATGGAAGGTGTTGGTGGTGGTGGTGGTGAAGAATGGTTGAGACCAATGGATGCAGA  
 ACAATTAAGAGAATGTGGTCATAGAATGGTTGATTTGTTGCGATTACTACAAGTCAA  
 TCGAAGCTTTTCCAGTTTTGTCTCAAGTTCAACCAGGTTATTTGAAAGAAGTTTTGCCA  
 GATTCAGCTCCAAGACAACCAGATACTTTAGATTCTTTGTTGCGATGATATCCAACAAAA  
 GATTATCCCAGGTGTTACACATTGGCAATCACCAAAACACTTCGCATACTACCCATCTA  
 45 ATTCTTCTACTGCTGGTTTTCTTGGGTGAAATGTTGTCTGCTGCTTTTAATATCGTTGGT  
 TTTTCATGGATCACTTCTCCAGCTGCAACAGAATTGGAAGTTATTGTTTTAGATTGGTT  
 CGCAAAGATGTTGCAATTGCCATCACAATCTTATCTACTGCTTTGGGTGGTGGTGTTA  
 TTCAAGGTACAGCTTCAGAAGCAGTTTTGGTTGCATTGTTAGCTGCAAGAGATAGAGCT  
 TTGAAGAAACATGGTAAACATTCTTTAGAAAAATTGGTTGTTTACGCATCAGATCAAAC

45



TCATTCTGCTTTACAAAAAGCATGTCAAATCGCTGGTATTTTCTCTGAAAACGTTAGAG  
 TTGTTATTGCTGATTGTAATAAGAACTACGCTGTTGCACCAGAAGCAGTTTCAGAAGCT  
 TTGTCTATCGATTTGTCTTCAGGTTTGATCCCATTTTTCATTTGTGCAACAGTTGGTAC  
 TACATCTTCATCTGCTGTTGATCCATTACCAGAATTGGGTCAAATCGCTAAGTCAAACG  
 5 ATATGTGGTTCCATATCGATGCTGCTTATGCTGGTCTGCATGTATTTGTCCAGAATAC  
 AGACATCATTTGAATGGTGTGAAGAAGCAGATTCTTTTAATATGAATGCTCATAAGTG  
 GTTCTTGACTAACTTCGATTGTTTCATTGTTGTGGGTAAAGGATAGATCATTTTTGATCC  
 AATCATTATCTACAAATCCAGAATTCTTGAAAAATAAGGCTTCACAAGCAAACCTCTGTT  
 GTTGATTTCAAGGATTGGCAAATCCCATTTGGGTAGAAGATTCAGATCATTGAAGTTGTG  
 10 GATGGTTTTGAGATTGTACGGTGTGATAATTTGCAATCTTACATCAGAAAGCATATCC  
 ATTTGGCTGAACATTTCTGAACAATTGTTGTTGTGATTCAGATTCAAGATTCTGAAGTTGTTACT  
 CCAAGAACATTTTCATTGGTTTGTTCAGATTAGTTCCACCAACTTCTGATCATGAAAA  
 CGGTAGAAAGTTGAACTACGATATGATGGATGGTGTAAATTCATCTGGTAAAATTTTCT  
 TGTCACATACAGTTTTGTCTGGTAAATTCGTTTTGAGATTTGCTGTTGGTGCACCATTA  
 15 ACTGAAGAAAGACATGTTGATGCTGCATGGAAGTTGTTGAGAGATGAAGCTACAAAGGT  
 TTTAGGTAAAATGGTTTAA

**[00483] SEQ.ID NO: 53**

MEGVGGGGGGEWLRPMDAEQLRECGHRMVDFVADYYKSIEAFPVLSQVQPGYLKEVLP  
 20 DSAPRQPDTLDSLFDIQQKIIPGVTHWQSPNYFAYYPSNSSTAGFLGEMLSAAFNIVG  
 FSWITSPAATELEVIVLDWFAKMLQLPSQFLSTALGGGVIQGTASEAVLVALLAARDRA  
 LKKHGHKSLEKLVVYASDQTHSALQKACQIAGIFSENVRVVIADCNKNYAVAPEAVSEA  
 LSIDLSSGLIPFFICATVGTSSSAVDPLPELGQIAKSNDMWFHIDAAYAGSACICPEY  
 RHHLNGVEEADSFNMNAHKWFLTNFDCSLLWVKDRSFLIQSLSTNPEFLKNKASQANSV  
 25 VDFKDWQIPLGRRFRSLKLWMLRLYGVDNLQSYIRKHIHLAEHFEQLLLSDSRFEVVT  
 PRTFSLVCFRLVPPTSDHENGRLNLYMDMGVNSSGKIFLSHTVLSGKFVLRFAVGAPL  
 TEERHVDAAWKLLRDEATKVLGKMV

**[00484] SEQ.ID NO: 54**

ATGACAAAGCCATTCATCCAAATTTGAACGTTGATGCATTGTTTTTAGGTCCAAAGTC  
 TGAACACGCTGTTTTCTTTAGAGAAATGATGGATTATGCAGTTTCTGAACACATGCATT  
 GGAGATCAGTTTTTCATCCAGAAGATTCTACATTGGTTACTTCAGTTGATAGATACGAA  
 CATAACTACAGAGAAACATTGTACAGAACTGAAGGTATTTTGAATCAATTATCTGCTAA  
 GTTGAAAAATTCTTCAATCCCATTTTTCTCTCCAAGATACTTAGGTCATATTTCTTCAG  
 35 ATACTTTGATGGTTTCTAATTTGGCTTATGTTATGGCTATGATGTACAACCCAAACAAC  
 TGTTCTTATGAAGCTTCACCAACTACAACGAATTGGAATTAGAATCTGGTTTGGATTT  
 GTGTAGAATGTTTCGGTTACAATCCACAACAATCTTGGGTCATATTACATCAGGTGGTA  
 CTGTTGCAAATTACGAAGGTTTGTGGGTTGCTAGAAATTTGAAGACATTGTCTTTCGCA  
 ATCTCACAACATCCAAAGGCTAAGGATTTGATCGAACATAAGACTCCAAAGCAATTGAA  
 40 AAATATGCCAACATCTGAATTGTTAGATTGATCACTGAATTACAAAAGAGAGGTTTGT  
 TCGAAGAAATCAGAGATATGACATGTAGAGGTACTGGTGTAAAGGCTCAAATTTTAGGT  
 AAATTGTTAGTTCCACAATCTAAGCATTATTCATGGATGAAAGCTGCAGATATCTTCGG  
 TATCGGTCAAGAAAACATCATCCCATTTGCCAGTTAATGATCATTACCAAACAGATGTTG  
 CTCAAATGAGAAAGATCACTTTCGGTTTGATTGAACAAGGTGAACCAATTTTAGCAATG  
 45 ATTGCTGTTGTTGGTACAACGAAGTTGGTGAATCGATAGAATCGATGAAGTTATTGC  
 TTTAAGAAAAGAATGTGAAGAAAGATACGGTGAATCTTTCTACATCCATGTTGATGCTG  
 CTTATGCTGGTTACGCTTGTTCATGTTGTTGAACGAACAGGGTGAATTCGTTGAATAC  
 GATGAATTGGCAAACATACCATAGATCTGCTGGTATTATGCCAGAAAATATTTTCATGGCC  
 AAAACCAGAAGTTTACCAATCTTTTAAAGCTTTGAAAGATGTTGATTCAATTACAGTTG

ATCCACATAAGGTTGGTTTTATTCAATATGCTGCAGGTGCAATTTGTATGAAGGATAAG  
 AGAATTTTGAATTAATTTCTTCACATGCTGCATACGTTTTTGAAGGTCATGCTGATAA  
 GACATCTACTAAAAATAGAGGTATTTTGGGTTCTTCAATTATGGAAGGTTCTAAATCAG  
 GTGCAACAGCTGCAGCTTTGTGGGCAGCTCATAGATTGTTACCATTGAACATCAACGGT  
 5 TACGGTAAAGTTATTGCAGCTGGTATCGTACTGCACAATCTTTGTTGAATAAGTTGGC  
 TAATTTGCCACCAATCGAAATCGGTAAACATAGATTGGAAGTTCATATTTTCCATCAC  
 CAGATTTCCACATGATCAACTTCACTTTTAAAGAAGTTGGTAACGAATCTTTGTCAAAC  
 CATAACGCTTTGAATAAGAGATTGTATGAATTATGTTCTTACTCAGCAGGTAGAGCTTA  
 CACAAATGATTTCTTGACATCTTCAACTATTTTAGATTACAAGGAATACGGTGACACTC  
 10 CTTGGCATCATGCTCAAAAATGTGGTTTTTCAAGATCAGAATGGGAAAAAGTTAGAAAC  
 ATCTATGTTTTGAGAGCAGCTGTTATGACATACTGTTTGAGAGATGAAGAACATTTGCGA  
 AGAATTCTGGGAACAATTGCAAGCTATCTTCGTTAAGAAATTGACTCAAATCGTTGATG  
 AAGAAGATAAGAAAGCAAGATTGAGATTAGAACCAGATGTTCTTTTATTGGTTAA

15 **[00485]** SEQ.ID NO: 55

MTKPFHPNLNVDALFLGPKSENAVFFREMMDYAVSEHMHWRSGFHPEDSTLVTSDRYE  
 HNYRETLRTEGILNQLSAKLKNSSIPFFSPRYLGHISSDTLMVSNLAYVMAMMYNPN  
 CSYEASPTTTELELESGLDLCRMFGYNPQQSWGHITSGGTVANIEGLWVARNLKTLSFA  
 ISQHPKAKDLIEHKTPKQLKNMPTSELLDLITELQKRGLFEEIRDMTCRGTGVKAQILG  
 20 KLLVPQSKHYSWMKAADIFGIGQENIIPLPVNDHYQTDVAQMRKITFGLIEQGEPI  
 IAVVGTTEVGAIIDRIDEVIALRKECEERYGESFYIHVDAAYAGYACSMMLNEQGEFVEY  
 DELANYHRSAGIMPENISWPKPEVYQSFKALKDVSITVDPHKVGFQYAAAGAIKMDK  
 RILELISSHAAYVFEHADKTSTKNRGILGSSIMEGSKSGATAAALWAAHRLPLNING  
 YGKVIAAGIVTAQSLNKLNLPPIEIGKHRFEVHIFPSPDFHMINFTFKEVGNESLSN  
 25 HNALNKRLYELCSYSAGRAYTNDFLTSTILDYKEYGDTPWHAQKCGFSRSEWEKVRN  
 IYVLRAAVMTYCLRDEEHFEEFWEQLQAI FVKKLTQIVDEEDKKARLRLEPDVPFIG

**[00486]** SEQ.ID NO: 56

ATGGGTTCTTTGGATTCAAACACTACGATACTGAATCTCCAGCTTCAGTTGGTCAATTCAA  
 30 TCCATTAGATCCAGAAGAATTCAGAAAGCAAGCTCATTGTATCGTTGATTTTCATCGCAG  
 ATTACTACAAGAACATTGAATCTTATCCAGTTTTGTGACAAGTTGATCCAGGTTACAGA  
 CATTCAAGATTGGGTAAAAATGCTCCATACAGATCTGAACCATTCGAATCAATCTTGAA  
 GGATGTTCAAAGGATATCATCCCTGGTATGACTCATTGGATGTCTCAAATTTCTTTG  
 CTCATTTTCCAGCAACTGTTTCTTCAGCTGCATTTGTTGGTGAAATGTTGTGTACATGT  
 35 TTCAACTCTGTTGGTTTTAATTGGTTAGCATCACCAGCTGCAACAGAATTGGAAATGGT  
 TGTTATTGATTGGTTAGCTAACATGTTGAAGTTGCCAAAGTCTTTTATGTTTTTCAGGTA  
 CTGGTGGTGGTGGTTTTGCAAGGTACTACATCTGAAGCTATCTTGTGTACATTGATTGCT  
 GCATCACCATGCAATTCGAAATCGTTGGTGTAAAGACTTCTACATCATTCGTTGTTTA  
 TGGTTCTGATCAAACCTATTCAACATACGCTAAAGCATGTAAATTAGCAGGTATTTTGC  
 40 CATGTAACATCAGATCTATTCCAACACTACAGCTGATTCAAACCTCTCTGTTTCACCATTG  
 TTGTTGAGAAGAGCTATTGAAGCAGATAAAGCTGCAGGTATGGTTCCATTGTATATTTG  
 TGCTACAGTTGGTACTACATCTACTACAGCAATTGATCCATTATCTTCATTGGCTGATG  
 TTGCAAATGATTATGGTGTGTTGGTTTCATGTTGATGCTGCATACGCTGGTTCGTCATGT  
 ATTTGTCCAGAATTCAGACATTATTTGGATGGTATCGAAAGAGCAGATTCTTTGTCAAT  
 45 GTCTCCACATAAGTGGTTGTTGTCTTACTTAGATTGTTGTTGTTTGTGGGTTAAGTCAC  
 CATCTTTATTGGTTAAGGCTTTATCAACTGATCCAGAATATTTGAAAAATCAACCATCA  
 GAATCTAAGTCAGTTGTTGATTACAAGGATTGGCAAGTTGGTACAGGTAGAAGATTCAA  
 AGCTTTGAGATTGTGGTTCGTTATGAGATCTTATGGTGTGCAAATTTGCAATCACATA  
 TCAGAACTGATGTTCAAATGGCTAAGATGTTGCAAGGTTTCGTTAAGTCTGATCCAAGA

TTCGAAATTTTGGTTCCAAGAGTTTTCTCTTTGGTTTGTTCAGATTGAACCCAATCTC  
 TGGTTTCAAGATCCAAGTGGTACAGAAGCTTTGAACAGAAAGTTATTGGATTGGGTAAAT  
 CTACAGGTAGAGTTTATATGACTCATACAAAAGTTGGTGGTATCTATATGTTGAGATTT  
 GCTGTTGGTGCAACTTTAACAGAAAAGAGACATGTTTCTTCAGCTTGGAAGTTGATTAA  
 5 AGAAGGTGCAGATGTTTTGTTGAAGGAAGATTAA

**[00487] SEQ.ID NO: 57**

MGS LDSNYDTES PASVGQFNPLDPEEFRKQAHCI VDFIADYYKNIESYPVLSQVDPGYR  
 HSRLGKNAPYRSEPFESILKDVQKDIIPGMTHWMSPNFFAHFPATVSSAA FVGEMLCCTC  
 10 FNSVGFNWLASPAATELEMVVIDWLANMLKLPKSFMSGTGGGV LQGTTSEAILCTLIA  
 ASPMHFEIVGVKTSTSFVVYGS DQTHSTYAKACKLAGILPCNIRSIPTTADSNFSVSPL  
 LLRRAIEADKAAGMVPLYICATVGTTSTTAIDPLSSLADVANDYGVWFHVDAA YAGSAC  
 ICPEFRHYLDGIERADSLSLSPHKWLLSYLDCCCLWVKSPSLLVKALSTDPEY LKNQPS  
 ESKSVVDYKDWQVGTGRRFKALRLWFVMSYGVANLQSHIRTDVQMAKMFEGFVKSDPR  
 15 FEILVPRVFSLVCFRLNPISGSDPTGTEALNRKLLDWNSTGRVYMTHTKVGGIYMLRF  
 AVGATLTEKRHVSSAWKLIKEGADVLLKED

**[00488] SEQ.ID NO: 58**

ATGGAAGGTGGTTTTACTGGTGGTGACGAATACCAAAAAGCATTTCTTGCCAAGAGATTA  
 20 CTTAAACACATACTACTCATTCCAATCTGGTCCATCACCAGAAGCTGAAATGTTGAAGT  
 TTAATTTGGAATGTTTGCATAAGACTTTTCGGTCCAGGTGGTTTGCAAGGTGACACTTTA  
 ATTGATATTGGTTCTGGTCCAACAATCTATCAAGTTTTGGCTGCATGTGAATCTTTTAA  
 AGATATCACTTTGTCAGATTTTACAGATAGAAATAGAGAAGAATTGGCTAAATGGTTGA  
 AGAAAGAACCAGGTGCATACGATTGGACTCCAGCTTTGAAGTTCGCATGTGAATTAGAA  
 25 GGCAACTCTGGTAGATGGCAAGAAAAAGCTGAAAAGTTGAGAGCAACAGTTAAGAGAGT  
 TTTGAAGTGTGATGCTAATTTGTCAAATCCATTAACCTCAGTTGTTTTGCCACCAGCAG  
 ATTGTGTTTTGACATTGTTAGCTATGGAATGTGCATGTTGTTCTTTGGATGCTTATAGA  
 GCTGCATTGAGAAATTTGGCATCATTGTTAAAACCAGGTGGTCATTTGGTTACTACAGT  
 TACTTTGCAATTATCTTCATACATGGTTGGTGAAAGAGAATTTTCTTGTGTTGCTTTGG  
 30 AAAAAGAAGAAGTTGAACAAGCTGTTTTAGATGCAGGTTTCGATATCGAACAATTGTTG  
 TACTCTCCACAATCTTACTCAGCTTCTACAGCACCAAATAGAGGTGTTTGTCTTGGT  
 TGCTAGAAAGAAACCAGGTT CATAA

**[00489] SEQ.ID NO: 59**

35 MEGGFTGGDEYQKHFLPRDYLNNTYYSFQSGPSPEAEMLKFNLECLHKTFGPGGLQGD TL  
 IDIGSGPTIYQVLAACESFKDITLSDFTDRNREELAKWLKKEPGAYDWT PALKFACEL E  
 GNSGRWQEKA EKL RATVKRVLKCDANLSNPLTPVVLPPADCVLTL LAMECACCSLDAYR  
 AALRNLASLLKPGGHLVTTVT LQLSSYMVG EREFSCVALEKEEVEQAVLDAGFDIEQLL  
 YSPQSYSASTAPNRGVCFLVARKKPGS  
 40

**[00490] SEQ.ID NO: 60**

ATGGATAAGATCTCTGCACCATTTTTCTCTGGTACTTCACCAGCTGCAGCTTCAGTTGC  
 TGGTGTGATGAAGATGATAGATTGTGTTTCCAAGCTCAAGAATTGATGTTGCGCATA CA  
 ACATCTCTATGGTTTTGAGAGCAGCTATTCAATTAGGTTTGTTAGATGCTTTGTCAGCA  
 45 GCTGGTGGTAAAGCATTGACTCCAAACGAATTAGTTGAAAACGTTGAAACATCTTCAA  
 TAAGGCTGAAGCAGCTGCAGCTGTTGATAGAATTTTGAGATATTTGTCTTGTTCACG  
 TTGTTACTTGTTCTTCAGAAGCAGCTGGTCCAGATGGTACATTAGTTAGAAGATACACT

ACAGGTCCATTGTGTAGATGGTTAACTAAAGATAGAGGTGACGGTACATTATCTCCATT  
 TGCTGTTTTTGTGTTGATCCAGATCATTTGTTTCCTTGGCATCATATTGCAGAAGCTG  
 TTA CTGCTGGTGGTCCATCAGCATTCGAAAGAACACAAAAGTGGCCATACTACGAATAC  
 ATGGGTAAAAATCAAAGATTGGGTACTTTGTTTCGATAACGCAATGGCTCAACATTCTGT  
 5 TATTTTGGTTACAAAGATGTTAGAAAGATTCAAAGGTTTTGATGGTGTCAAAGATTAG  
 TTGATGTTGGTGGTGGTACTGGTTCTACATTGGGTATGATCACTTCAAAGTACAAGCAC  
 ATGACAGGTATTAATTACGATTGGCCACATGTTATTGCTCAAGGTTTGCCATTACCAGG  
 TGTGTAACATGTTGCAGGTGACATGTACGAATCTATTCCAAGTGGTACGCTGTTTTGT  
 TACAATGGATTACATTGATGTTGAACGATGATGAATTCGTTAAGATCTTGTCAAATTGT  
 10 CATAATGCTTTGCCAAAAGATGGTAAAGTTATTGTTGTTGATGGTATTTTACCAGAAAA  
 CCCAGATTCTTCATTGACTGCAAGAGATGCTTTTACATTAGATATCATCATGTTTCGTTT  
 TGTTTAAAGGTGCAAAGCAAAGAACTGAAAAGGAATTCGCAAGATTAGCTAAGCAAGCA  
 GGTTTTACTGGTGGTATTAAGAAAACCTTATTTTTCTTTAATTTTACGCTTGGAAAT  
 CACTAAATAA

15

**[00491] SEQ.ID NO: 61**

MDKISAPFFSGTSPAAASVAGVDEDDRLCFQAQELMFAYNISMVLRRAIQLGLLDALSA  
 AGGKALTPNELVENVETSSNKAEEAAVDRILRYLSCFNVVTCSSAAGPDGTLVRRYT  
 TGPLCRWLTKDRGDGTLSPFAVFVVDPDHLFPWHHIAEAVTAGGPSAFERTQKWPLYEY  
 20 MGKNQRLGTLFDNAMAQHSVILVTKMLERFKGFDGVQRLVDVGGGTGSTLGMITSKYKH  
 MTGINYDLPHVIAQGLPLPGVEHVAGDMYESIPTGDAVLLQWITLMLNDDEFVKILSNC  
 HNALPKDGKIVVDGILPENPDSSLTARDAFTLDIIMFVLFKGAKQRTEKEFARLAKQA  
 GFTGGIKKTYIFFNFYALEFTK

**25 [00492] SEQ.ID NO: 62**

ATGAACACCTTCAGAACAGCCACTGCCAGAGACATACCTGATGTAGCAGCAACTCTTAC  
 GGAAGCCTTCGCAACTGATCCACCCACGCAGTGGGTGTTCCCCGACGGTACTGCCGCCG  
 TCAGCAGGTTCTTTACACATGTTGCAGATAGGGTTCACACGGCCGGTGGTATTGTTGAG  
 CTACTACCAGACAGAGCCGCCATGATTGCATTGCCACCACACGTGAGGCTGCCAGGAGA  
 30 AGCTGCCGACGGAAGGCAGGCGGAAATTCAGAGAAGGCTGGCAGACAGGCACCCGCTGA  
 CACCTCACTACTACCTGCTGTTTTACGGAGTTAGAACGGCACACCAGGGTTCGGGATTG  
 GCGGAAGAATGCTGGCCAGATTAAGTAGCAGAGCTGATAGGGACAGGGTGGGTACATA  
 TACTGAGGCATCCACCTGGCGTGGCGCTAGACTGATGCTGAGACATGGATTCCATGCTA  
 CAAGGCCACTAAGATTGCCAGATGGACCCAGCATGTTCCACTTGGAGAGATCCAATC  
 35 CATGATCATTCTGATTAG

**[00493] SEQ.ID NO: 63**

MNTFRTATARDIPDVAATLTEAFATDPPTQWVFDPGTAAVSRFFTHVADRVHTAGGIVE  
 LLPDRAAMIALPPHVRLPGEAADGRQAEIQRRLADRHPLTPHYLLFYGVRTAHQGSGL  
 40 GGRMLARLTSRADRDVGTYTEASTWRGARLMLRHGFHATRPLRLPDGPSMFPLWRDPI  
 HDHSD

**[00494] SEQ.ID NO: 64**

ATGAAAGGATATTTTCGGACCATACGGTGGCCAGTACGTACCAGAAATATTAATGGGTGC  
 45 CTTAGAGGAGTTAGAGGCAGCATACGAGGAGATTATGAAGGATGAGAGCTTCTGGAAGG  
 AGTTCAACGATCTACTGAGGGATTACGCAGGCAGACCAACGCCATTGTACTTTGCCAGG  
 AGATTGTCTGAGAAGTACGGCGCCCGTGTTTACTTGAAGCGTGAGGATCTGCTGCACAC

TGGAGCACACAAGATAAATAACGCTATCGGACAGGTTTTATTGGCCAAATTAATGGGGA  
 AGACACGTATCATAGCCGAGACGGGAGCTGGGCAGCATGGAGTCGCTACTGCTACCGCT  
 GCTGCCCTGTTTCGGAATGGAATGTGTGATCTACATGGGTGAAGAGGACACAATCAGACA  
 GAAGTTGAACGTGGAGCGTATGAAATTATTAGGGGCTAAAGTTGTCCCTGTAAAGTCTG  
 5 GCAGTAGGACCTTGAAGGATGCGATAGACGAGGCTTTGAGAGACTGGATTACTAATTTA  
 CAGACAACATATTATGTTATCGGATCTGTTGTTGGTCCCCACCCTTACCCAATTATCGT  
 AAGGAATTTCCAGAAGGTTATCGGTGAGGAGACCAAGAAGCAAATACCAGAAAAGGAAG  
 GTCGTTTGCCAGACTATATAGTTGCCTGCGTAGGCGGCGGTAGCAATGCCGCAGGTATA  
 TTTTACCCATTTCATAGACTCTGGAGTAAAGCTGATAGGTGTTGAGGCAGGTGGCGAGGG  
 10 ATTGGAGACAGGTAAACACGCAGCCTCGTTATTAAAGGGTAAAATTGGCTATTTACATG  
 GATCGAAGACCTTTGTTCTACAAGATGACTGGGGTCAAGTCCAAGTGAGCCATTCGGTG  
 TCAGCTGGTCTTGACTATTGAGGAGTAGGACCTGAGCATGCTTATTGGAGAGAGACAGG  
 GAAGGTTCTGTACGACGCAGTGACTGACGAAGAGGCTTTGGACGCATTTATAGAGTTAT  
 CAAGACTAGAGGGCATTATACCCGCTTTAGAGTCATCGCATGCTCTAGCATATTTGAAG  
 15 AAGATAAATATAAAAAGGTAAGGTTGTGGTGGTCAACCTATCAGGGAGAGGGGATAAAGA  
 CCTGGAGTCAGTCTTAAACCATCCATACGTGAGAGAAAGAATTAGATGA

**[00495] SEQ.ID NO: 65**

MKGYFGPYGGQYVPEILMGALEEELEAAAYEIMKDESFWKEFNDLLRDYAGRPTPLYFAR  
 20 RLSEKYGARVYLKREDLLHTGAHKINNAIGQVLLAKLMGKTRIIAETGAGQHGVATATA  
 AALFGMECVIYMGEEDTIRQKLNVERMKLLGAKVVPVKSGSRTLKDAIDEALRDWITNL  
 QTTYVIGSVVGPHPYPPIIVRNFQKVI GEETKKQIPEKEGRLPDYIVACVGGGSNAAGI  
 FYPFIDSGVKLIGVEAGGEGLTGKHAASLLKGKIGYLHGSKTFVLQDDWGQVQVSHSV  
 SAGLDYSGVGPEHAYWRETGKVLVDVTDDEALDAFIELSRLEGII PALESSHALAYLK  
 25 KINIKGKVVVVNLSGRGDKDLESVLNHPYVRERIR

**[00496] SEQ.ID NO: 66**

ATGTGGTTCCGGTGAGTTTGGTGGACAATATGTGCCAGAGACTTTAGTGGGTCCCTCTTAA  
 GGAATTGGAAAAGGCATATAAAAAGGTTCAAGGACGATGAGGAGTTCAACAGGCAACTAA  
 30 ACTATTATTTGAAGACATGGGCCGGTAGACCAACGCCCTTGATTATGCTAAGAGGTTA  
 ACTGAAAAGATTGGCGGCGCGAAAGTGTATCTGAAAAGAGAAGACCTAGTTCATGGTGG  
 AGCACACAAGACAAATAATGCCATTGGACAAGCACTATTGGCAAAGCTAATGGGTAAAA  
 CTAGATTGATAGCTGAGACAGGAGCGGGTCAACATGGGGTTCGCGACAGCGATGGCTGGT  
 GCACTACTGGGGATGAAGGTAGATATTTACATGGGTGCTGAGGACGTTGAGCGTCAGAA  
 35 ACTAAATGTCTTCAGGATGAAGCTATTAGGTGCCAATGTTATACCTGTAAATTCTGGCT  
 CAAGAACACTAAAGGACGCCTTCGACGAGGCTCTTAGAGACTGGGTTGCCACTTTTCGAG  
 TATACTCATTACTTGATCGGTTGAGTGGTTGGACCACATCCATACCCAACCATCGTTAG  
 GGACTTTTCAGAGCGTGATTGGTAGAGAGGCTAAGGCACAGATCTTAGAAGCAGAGGGAC  
 AGCTACCTGACGTCATAGTTGCCTGCGTGGCGGGTGGCTCTAACGCAATGGGTATATTC  
 40 TATCCATTCGTTAATGACAAGAAGGTTAAATTAGTAGGAGTCGAAGCTGGCGGAAAGGG  
 GTTAGAGTCGGGTAAACACTCAGCAAGCTTAAATGCAGGACAGGTAGGGGTGTCCACG  
 GCATGTTGTCGTATTTCTTGCAAGACGAGGAAGGTGAGATAAAGCCAAGTCATTCAATT  
 GCTCCAGGCCTTGACCACCCCGGTGTTGGTCCAGAGCACGCTTACTTAAAGAAGATTCA  
 AAGGGCCGAGTACGTCGCTGTAACAGACGAAGAGGCATTGAAAGCTTTCCATGAGCTAT  
 45 CCAGAACTGAGGGGATTATACCCGCCCTTGAGTCTGCCCATGCTGTGGCGTACGCCATG  
 AAGTTAGCTAAAGAGATGTCCCGTGACGAAATCATCATTGTAAATCTATCAGGGAGAGG  
 AGACAAGGATTTGGACATTGTATTGAAGGCAAGCGGAAATGTTTGA

**[00497]** SEQ.ID NO: 67

MWFGEFGGQYVPETLVGPLKELEKAYKRFKDDEEFNRQLNYYLKTWAGRPTPLYAKRL  
 TEKIGGAKVYLKREDLVHGGAHKTNNAIGQALLAKLMGKTRLIAETGAGQHG VATAMAG  
 ALLGMKVDIYMGAEDVERQKLN VFRMKLLGANVIPVNSGSRTLKDAFDEALRDWVATFE  
 5 YTHYLIGSVVGPHPYPTIVRDFQSVIGREAKAQILEAEGQLPDVIVACVGGGSNAMGIF  
 YPFVNDKKVCLVGVEAGGKGLES GKHSASLNAGQVGVSHGMLS YFLQDEEGQIKPSHSI  
 APGLDHPGVGPEHAYLKKI QRAEYVAVTDEEALKA FHLSRTEGII PALESAHAVAYAM  
 KLAKEMSRDEII IVNL SGRGDKDL DIVLKASGNV

10 **[00498]** SEQ.ID NO: 68

ATGTGGTTTGGGGAATTCGGTGGTCAATACGTGCCAGAAACCTTAGTTGGGCCTTTAAA  
 GGAGCTTGAGAAAGCTTACAAAAGATTTAAGGACGACGAGGAGTTTAATAGGCAATTAA  
 ACTATTACTTGAAGACTTGGGCAGGTAGACCAACACCACTTTATTACGCGAAGAGACTT  
 ACAGAGAAGATCGGTGGCGCTAAGGTTTATTTGAAGCGTGAGGATTTGGTTCATGGTGG  
 15 AGCACATAAGACAAACAATGCCATCGGGCAAGCCTTGTTGGCCAAATTGATGGGTAAGA  
 CTAGGTTGATAGCTGGCACAGGTGCCGGCCAGCATGGCGTTGCAACAGCAATGGCTGGT  
 GCTCTTTTAGGGATGAAGGTGGATATCTATATGGGCGCCGAGGACGTTGAAAGACAAAA  
 GCTTAACGTATT CAGAATGAAGTTATTAGGTGCGAATGTTATTCCAGTGAATTCAGGTT  
 CCAGGACGCTGAAGGACGCAATAGACGAAGCCTTGAGAGACTGGGTTGCTACTTTTGAG  
 20 TACACACATTATTTGATTGGTTCGGTGGTGGACCCACCCGTACCCTACAATAGTTAG  
 AGATTTTCAATCTGTTATCGGTAGAGAGGCGAAGGCTCAAATTCAGAGGCCGAGGGAC  
 AGCTACCCGATGTTATAGTCGCCTGCGTTGGAGGAGGATCAAATGCGATGGGAATCTTC  
 TACCCTTTTGTAATGATAAGAAAGTTAAGTTAGTCGGAGTTGAGGCCGGTGGTAAAGG  
 GCTAGAAAGCGGCAAACACTCGGCATCTCTTAATGCCGGTCAAGTAGGAGTAAGTCACG  
 25 GGATGTTATCTTACTTCTTGCAGGATGAGGAAGGTCAAATTAAGCCATCTCATTCTATA  
 GCCCCAGGATTAGACTATCCCGGTGTTGGGCCTGAGCACGCTTATCTGAAGAAGATACA  
 GCGTGCTGAGTACGTTGCAGTAAGTGAAGAGGCTTTGAAGGCTTTTCACGAGTTAT  
 CAAGAACTGAGGGAATAATTCTGCCTTGGAATCGGCACATGCGGTTGCCTACGCCATG  
 AAGCTTGCTAAGGAAATGTCAAGAGACGAAATCATTATAGTGAATTTATCGGGAAGGGG  
 30 CGATAAGGACTTGGACATTGTTTTAAAAGCAAGTGGCAATGTATAA

**[00499]** SEQ.ID NO: 69

MWFGEFGGQYVPETLVGPLKELEKAYKRFKDDEEFNRQLNYYLKTWAGRPTPLYAKRL  
 TEKIGGAKVYLKREDLVHGGAHKTNNAIGQALLAKLMGKTRLIAETGAGQHG VATAMAG  
 35 ALLGMKVDIYMGAEDVERQKLN VFRMKLLGANVIPVNSGSRTLKDAIDEALRDWVATFE  
 YTHYLIGSVVGPHPYPTIVRDFQSVIGREAKAQIPEAEGQLPDVIVACVGGGSNAMGIF  
 YPFVNDKKVCLVGVEAGGKGLES GKHSASLNAGQVGVSHGMLS YFLQDEEGQIKPSHSI  
 APGLDYPGVGPEHAYLKKI QRAEYVAVTDEEALKA FHLSRTEGII PALESAHAVAYAM  
 KLAKEMSRDEII IVNL SGRGDKDL DIVLKASGNV  
 40

**[00500]** SEQ.ID NO: 70

ATGTGGTTTGGTGAGTTTGGTGGTCAGTATGTGCCTGAGACGCTAGTTGGGCCACTGAA  
 GGAGTTAGAGAAAGCGTACAAGAGATTCAAAGATGACGAGGAGTTTAACCGTCAGTTGA  
 ATTATTATCTTAAGACGTGGGCTGGTAGGCCGACCCCTCTATACTACGCTAAGCGTTTA  
 45 ACCGAAAAGATAGGTGGCGCCAAGATCTATCTGAAGAGAGAGGACCTAGTCCACGGCGG  
 GGCTCACAAAATAACAACGCAATTGGTCAGGCACCATTAGCAAAGTTAATGGGCAAAA  
 CACGTCTAATTGCCGAAACTGGTGCTGGTCAACATGGAGTTGCTACGGCAATGGCGGGC

GCACTGTTAGGTATGAAAGTGGATATCTATATGGGAGCAGAAGACGTGGAGAGACAGAA  
 GATGAACGTGTTCCGTATGAACTATTGGGAGCTAATGTTATCCCGGTTAACTCTGGCT  
 CAAGAACTGCAAAGGACGCCATAAACGAAGCTTTGAGGGATTGGGAAGCCACATTCGAG  
 TACACACACTACCTTATAGGCTCAGTCGTAGGCCCGCATCCCTACCCCACCATAGTTTCG  
 5 TGA CTTC CAGAGT GTTATTGGCCGTGAGGCCAAGGCACAGATCTTGGAAGCTGAAGGTC  
 AATTGCCTGACGTAATAGTTGCATGTGTGGGTGGCGGTTCTAACGCTATGGGTATATTC  
 TACCCCTTCGTGAACGATAAGAAAGTAAAGCTAGTTGGAGTGGAAGCAGGTGGCAAAGG  
 GTTGAGTCGGGCAAGCACTCAGCCTCATTAACGCAGGTCAGGTAGGCGTTCTGCACG  
 GTATGTTATCCTATTTCTTACAGGACGAGGAAGGCCAAATCAAACCATCGCACTCAATC  
 10 GCACCTGGTTTATAGATTACCCAGGAGTCGGTCCCGAGCACGCATATTTAAAGAAAATCCA  
 GAGAGCAGAGTACGTCACGGTGACGGATGAGGAAGCCCTTAAAGCGTTTCACGAGCTGT  
 CGAGGACCGAGGGCATAATACCTGCTCTTGAAAGTGCCACGCAGTTGCCTATGCTATG  
 AAGTTGGCCAAGGAGATGAGTAGAGACGAGATTATCATAGTAAATCTGTCAGGCAGAGG  
 TGACAAGGATCTTGACATAGTCCTTAAGGCGAGCGGTAATGTTTTGGAGTAA  
 15

**[00501] SEQ.ID NO: 71**

MWFGEFGGQYVPETLVGPLKELEKAYKRFKDDEEFNRQLNYYLKTWAGRPTPLYAKRL  
 TEKIGGAKIYLKREDLVHGGAHKTNNAIGQAPLAKLMGKTRLIAETGAGQHG VATAMAG  
 ALLGMKVDIYMGAEDVERQKMNVFRMKLLGANVIPVNSGSRTAKDAINEALRDWEATFE  
 20 YTHYLIGSVVGPHPYPTIVRDFQSVIGREAKAQILEAEGQLPDVIVACVGGGSNAMGIF  
 YPFVNDKKVKLVGVEAGGKGLESGKHSASLNAGQVGV LHGMLS YFLQDEEGQIKPSHSI  
 APGLDYPGVGPEHAYLKKIQRAEYVTVTDEEALKA FHEL SRTEGI I PALESAHAVAYAM  
 KLAKEMSRDEIIIVNLSGRGDKDLDIVLKASGNVLE

**25 [00502] SEQ.ID NO: 72**

ATGTGGTTCGGAGAGTTCGGCGGGCAGTACGTCCCAGAAACGCTAGTCGGCCCACTGAA  
 GGAGCTGGAGAAAGCATACAAAAGATTCAAGGATGACGAGGAGTTCAACAGGCAATTAA  
 ACTATTACCTAAAGACCTGGGCTGGTAGGCCACCCCTTTATACTATGCAAAACGTTTA  
 ACGGAGAAAATCGGTGGCGCAAAAGTGTACTTGAAGAGAGAAGACTTAGTCCACGGTGG  
 30 TGCTCATAAACTAACAACGCTATAGGGCAGGCACCTTTTAGCTAAGTTGATGGGTAAAA  
 CGAGGTTGATTGCGGGTACCGGTGCCGGTCAACATGGCGTTGCTACGGCAATGGCGGGA  
 GCCCTATTGGGCATGAAGGTTGACATTTATATGGGCGCTGAGGACGTGGAACGTCAGAA  
 ACTGAACGTCTATCGTATGAAATTGCTTGGTGCCAACGTTATACCAGTCAATTCGGGCT  
 CGCGTACGGTTAAAGACGCTTTCGATGAGGCACTATGTGATAGAGTTGCCACTTTTGAA  
 35 TACACCCACTATCTTATTGGTACTGTCTGGGGTCCACATCCGTATCCAACCTATTGTAAG  
 GGACTTCCAGACTGTTATCGGAAGAGAGGCAAAGGCCCAAATATTAGAGGCTGAGGGAA  
 GGTTACCCGACGCCATTGTTGCATGCGTAGGCGGTGGTAGTAACGCAATGGGTATTTTC  
 TATCCCTTCGTTAACGATAAGAAGGTTAAGTTAGTGGGTGTGAGGCCGGTGGTAAGGG  
 GCTTGAGTCCGGTAAACATTCCGCATCTTTGAATGCTGGGCAAGTAGGTGTCAGTCATG  
 40 GAATGTTATCATACTTTCTACAGGACGAGGAAGGACAGATAAAACCTTCTCATTCTATA  
 GCGCCAGGTCTTGACCACCCTGGCGTTGGACCAGAGCATGCTTATTTGAAGAAGATCCA  
 GAGAGCTGAATATGTAGCTGTTACGGATGAGGAAGCTCTAAAGGCGTTCCACGAGTTAA  
 GTAGGACTGAGGGAATTATACCGGCATTAGAGTCTGCCCACGCTGTAGCTTACGCAATG  
 AAGCTGGCCAAAGAGATGAGCCGTGACGAGATCATCATCGTAAACTTGTCTGGCAGAGG  
 45 CGATAAGGACCTAGACATCGTATTAAAGGCCTCGGGGAACGTCCTAGAGTAG

**[00503] SEQ.ID NO: 73**

MWFGFEGGQYVPETLVGPLKELEKAYKRFKDDEEFNRQLNYYLKTWAGRPTPLYYAKRL  
 TEKIGGAKVYLLKREDLVHGGAHKTNNNAIGQALLAKLMGKTRLIAGTGAGQHG VATAMAG  
 ALLGMKVDIYMGAEDVERQKLNVMKLLGANVIPVNSGSRTVKDAFDEALCDRVATFE  
 5 YTHYLIGTVWGPHPYPTIVRDFQTVIGREAKAQILEAEGRLPDAIVACVGGGSNAMGIF  
 YPFVNDKKVCLVGVEAGGKGLESGKHSASLNAGQVGVSHGMLSIFLQDEEGQIKPSHSI  
 APGLDHPGVGPEHAYLKKIQRAEYVAVTDEEALKAFFHELSRTEGII PALESAHAVAYAM  
 KLAKEMSRDEIIIVNLSGRGDKDLDIVLKASGNVLE

**10 [00504] SEQ.ID NO: 74**

GTATCCGGCTGTTCCCTTCATAGCCCTTTCAATGAACGTTGCAGCCCTTTGAAGATTGGC  
 CATTTTGTGAGGACTCGAGCCTGACAGTTGGACCAACGCAACTTTAATTTTTTGTGAAA  
 GAATCTTCGAAGCACTCATACTGGCGATCTTCACGCCCTCCTGCTATTACAAAAGCTGT  
 GTTTTTACAAGAATCAAATTAAGTTAGCAAGATATTATACAACATTATTGATAATTTCA  
 15 ATATCGTGTTTCGTACCTGATGACGTATCTGTGCATTGATAAGGCCCGCATGTTTTCAGA  
 AAGCAGAGCGGAACGATTCCAAATTAGTGGCCTTGTGCTTTGCATGTCAATTGTGTTAC  
 CTTGAGCTCGTGGATTTGTTTTATCAATACACAGTCTACAGTCAAGAATTTTTTTTTATC  
 AAATTTTGCCTTCGAGCGTATAAAATAGCCGCTGTAGCTACTTAAGTTCCTGTTGAGCG  
 ATAGTTTTTTTCCATCACACGTACTATGGCAATTAAGTCCTCAGCGAGCTCGCATGGAA  
 20 TGCCTGCGATGAGCGACCTCATGCTATACCTGAGAAAGCAACCTGACCTACAGGAAAGA  
 GTTACTCAAGAATAAGAATTTTCGTTTTTAAACCTAAGAGTCACTTTAAAAATTTGTATA  
 CACTTATTTTTTTTTATAAATTATTTAATAATAAAAAATCATAAATCATAAGAAATTCGCT  
 TATTTAGAAGTGTCAACAACGTATCTACCAACGGAATGCGTGCGATTGAGGCGTAGTCT  
 GGAACGTCGTATGGGTATGGACCAGAGACGGATTGAGAACTTGTTGCGTTCTTGGGAC  
 25 GAATCTACATTGGAATGGAACCTGGGTGCCTGAAGAAACAGTGGTGAAATCAGCTGGGA  
 TGTCAATTGGCTTACCGTTTTTGGTCAACTGGTCTTTCAATGTTGAAAGCAGACAACAAA  
 GTAGCACCGGCAATCCAAACAGTAGATTGAGCTAGGTGAATACCTGGGCAGTTTCTTCT  
 ACCGTAACCGAAAGCAGCCTTTCTTGGGTCTCTAACAGTGTGTCTGGTTTACCATCAG  
 GACCCAAGTATCTTTCTGGACGGAACCGGATGGATCTGGATAGACTTCTGGGTCAATC  
 30 AAAACTGCCCAGGTGTTAGCAAAAACCAATGTGTTCTTTGGAATCAAATAACCTCTGTA  
 AACATCATCCTTCATCAATTTATGAGGGATGGCTAATGGAGCAATTTGGTTCCATCTGA  
 ATAATTCCTTGATACAAGCGGTCAAATATGGGAGTGAGTCGTCTTCCTCATCGTAGTCA  
 GGGATTTGACCGTTGTTGGTCAAAGCATCCAATTCAGCTTGAACCTTTCTTTCGACTTC  
 TGGGTATTTAACCATGGCCAATATGAAAGCGGACATAGCAGAGACGGTAGTGTGCCAC  
 35 CACCAACGTTAACTTCAGCAGCAGTGTTCTTGATAACATGTTCTTGGTGTTCTAAATCA  
 CCGTTCAAGTCCATAGCTTGTAATCTAGCAGAAGCGTAAGATGGTCTGGTTAGACCTTG  
 TGGAGCCAACCTTTCTCATAGTTTCGTATGGCATGTCAACCATGTGGTCAGCGGCTTCTC  
 TCCAGACCTTGGCCTTACGCTTGAAGACGGCACCTGGGAACCAAGCTGGCAAGTACTTC  
 AAAGATGGGAAGGAGTCAACCCAGAACTTACCTGGAACACTAGCAATGGCCAAACCTTC  
 40 GTTGGCTAAGTGGGTAGCTTCTAACCATGGGTCTCTTCGGCTAGATCGATACCGTAAC  
 CAATGTCCAAAGACATAGCAGCGATTTGATGTCTGATGTGTTGAGCCCACCGGTCCGGA  
 GTCTTGGTCAATTGTTGAACCAATTGGTGAGCAGCCTTGACTTGAGCGTGTCTGAATTG  
 CTTGATACCCTTTTCAGAGAATTCCTTAGCAACATTCCTTCTTCTCTCTCCAACGGT  
 CACCGTAAGTGATAAAAACCCAGATCAAATTCCTCAACCCATCAATTCATTGACCATAGTG  
 45 GATTCCAATCTACCGGAGTAGATGGAACCACGTTTTTCCAAGAGATCCGTGATAGTTTC  
 CAAGGTGTTTTAAATGACCATTTCGGTACCACCAGCATCTACGTACAAGATATCAGTGT  
 TGTAGTCACGACCCCATTTGCAAGAAAGTCAACCATGGAGATTCTTCTGGCATGTGCAAC  
 ATGTTACCAATAAATGGGATTGGAATTCCTGGTGGACAGGTGGCAATCTGGATCTACG  
 AACTCTTCTGGAGACGATGTAGTAAATACAACCAGCAATGACGAAAGAGAACAAGACAG



CAATCATTGTTTTATATTTGTTGTAAAAAGTAGATAATTACTTCCTTGATGATCTGTAA  
AAAAGAGAAAAAGAAAGCATCTAAGAACTTGAAAACTACGAATTAGAAAAGACCAAAT  
ATGTATTTCTTGCAATTGACCAATTTATGCAAGTTTATATATATGTAAATGTAAGTTTCA  
CGAGGTTCTACTAACTAAACCACCCCTTGGTTAGAAGAAAAGAGTGTGTGAGAACAG  
5 GCTGTTGTTGTACACGATTTCGGACAATTCTGTTTGAAAGAGAGAGAGTAACAGTACGA  
TCGAACGAACCTTTGCTCTGGAGATCACAGTGGGCATCATAGCATGTGGTACTAAACCCT  
TTCCCGCCATTCCAGAACCTTCGATTGCTTGTTACAAAACCTGTGAGCCGTCGCTAGGA  
CCTTGTTGTGTGACGAAATTGGAAGCTGCAATCAATAGGAAGACAGGAAGTCGAGCGTG  
TCTGGGTTTTTTTTCAGTTTTGTTCTTTTTGCAAACAACAGTTTATTCCTGGCATCCACTA  
10 AATATAATGGAGCCCGCTTTTTAAGCTGGCATCCAGAAAAAAAAGAATCCCAGCACCA  
AAATATTGTTTTCTTCACCAACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACA  
GAGAACAGGGGCACAAACAGGCAAAAAACGGGCACAACCTCAATGGAGTGATGCAACCT  
GCCTGGAGTAAATGATGACACAAGGCAATTGACCCACGCATGTATCTATCTCATTTTTCT  
TACACCTTCTATTACCTTCTGCTCTCTCTGATTTGGAAAAAGCTGAAAAAAAAGGTTGA  
15 AACCAGTTCCTGAAATTATTCCCCTACTTGACTAATAAGTATATAAAGACGGTAGGTA  
TTGATTGTAATTCTGTAAATCTATTTCTTAAACTTCTTAAATTCTACTTTTATAGTTAG  
TCTTTTTTTTTAGTTTTTAAACACCAAGAACTTAGTTTCGAATAAACACACATAAACAAA  
CAAATGGCTTCTAGTTCTTCCGATGTCTTCGTTTTGGGTCTAGGTGTTGTTTTGGCTG  
CCTTGATATATCTTCAGAGACCAATTATTCGCTGCTTCTAAGCCAAAGGTGGCTCCAGTT  
20 TCCACTACGAAGCCTGCCAACGGTTCGGCTAACCCAAGAGACTTCATCGCCAAGATGAA  
ACAAGGTAAGAAGAGAATCGTAATCTTCTACGGTTCCTCAAACCTGGTACCGCTGAAGAAT  
ATGCTATTCGTTTGGCTAAGGAAGCTAAGCAAAAGTTCGGTCTAGCCTCCTTGTTTTGT  
GATCCAGAAGAATACGATTTTGAAAAGTTGGACCAATTGCCAGAAGATTCTATTGCTTT  
CTTCGTCGTTGCTACCTATGGTGAAGGTGAACCTACAGACAACGCTGTCCAATTGTTGC  
25 AAAACTTGCAAGATGAAAGCTTCGAATTCTCCTCTGGTGAGAGAAAGTTGTCAGGTTTG  
AAGTACGTTGTTTTTGGTCTGGGTAAACAAGACCTACGAACATTACAACCTCATTGGGAG  
AACTGTTGACGCTCAATTGGCCAAGATGGGTGCTATCAGAATCGGTGAAAGAGGTGAAG  
GTGATGATGACAAGTCCATGGAAGAAGACTACTTGGAATGGAAGGATGGTATGTGGGAA  
GCGTTTGCCACTGCTATGGGTGTTGAAGAAGGTCAAGGTGGTGACTCCGCTGATTTTCGT  
30 CGTTTCCGAATTGGAATCTCACCCACCAGAAAAGGTTTACCAAGGTGAATTTTCTGCTA  
GAGCTTTAACCAAAAACCAAGGGTATTACGACGCTAAGAATCCTTTTGCTGCTCCAATT  
GCGGTTGCTAGAGAATTGTTCCAATCTGTTGTGCTAGAACTGTGTCCACGTGCAATT  
CAACATTGAAGGCTCTGGTATCACCTATCAACACGGTGACCACGTTGGTTTGTGGCCAT  
TGAATCCAGATGTTGAAGTCGAACGGTTGTTGTGTGTTTTAGGTTTAGCTGAAAAGAGA  
35 GATGCTGTCATCTCCATTGAATCCTTAGACCCGGCTTTGGCTAAGGTTCCATTCCCAGT  
CCCAACTACTTACGGTGCTGTGTTGAGACACTACATTGACATCTCTGCTGTCGCCGGTA  
GACAAATCTTGGGTACTTTGTCCAAATTGCTCCAAACCCAGAAAGCTGAAGCTTTCTTG  
AGAACTTGAACACTAACAAGGAAGAATACCACAACGTCGTCGCTAACGGTTGTTTGAA  
ATTGGGTGAAATTTTGCAAATCGCTACCGGTAACGACATTACTGTCCCACCAACTACTG  
40 CCAACACCACCAAATGGCCAATTCCATTGACATCATTGTTTCTGCCATCCCAAGATTG  
CAACCAAGATACTACTCTATCTCTTCTTCCCCAAAAATTTCATCCAAACACCATCCACGC  
TACCGTTGTTGTGCTCAAATACGAAAACGTTCCAACCGAACCAATCCCAAGAAAGTGGG  
TTTACGGTGTGCGTAGTAACCTTCTTGTGAATTTAAAGTACGCTGTTAACAAGGAACCA  
GTTCCATACATCACTCAAAATGGCGAACAAAGAGTCGGTGTCCCGGAATACTTGATTGC  
45 TGGTCCACGTGGTTCTTACAAGACTGAATCTTTCTACAAGGCTCCAATCCATGTTAGAC  
GTTCTACTTTCCGTTTGCCAACCAACCCAAAGTCTCCAGTCATCATGATTGGTCCAGGT  
ACTGGTGTGCCCCATTACAGAGGCTTCGTTCAAGAAAGAGTTGCCTTGCCAGAAAGATC  
CATCGAAAAGAACGGTCCTGACTCTTTGGCTGACTGGGGTTCGTATTTCTTGTCTACG  
GTTGTAGAAGATCCGACGAAGACTTCTTGTACAAGGACGAATGGCCACAATACGAAGCT  
50 GAGTTGAAGGGTAAGTTCAAGTTGCACTGTGCTTTCTCCAGACAAAACCTACAAGCCAGA  
CGGTTCTAAGATTTACGTCCAAGATTTGATCTGGGAAGACAGAGAACACATTGCCGATG

CCATCTTAAACGGTAAGGGTTACGTCTACATCTGCGGTGAAGCTAAGTCCATGTCTAAA  
CAAGTTGAAGAAGTTCTAGCCAAGATCTTGGGCGAAGCCAAAGGTGGTTCGGGTCCAGT  
TGAAGGTGTTGCTGAAGTCAAGTTACTGAAGGAACGGTCCAGATTGATGTTGGATGTCT  
GGTCTGAACAAAAGTTAATTTCTGAAGAAGATTTGGAATGAATCGCGTGCATTCATCCG  
5 CTCTAACCGAAAAGGAAGGAGTTAGACAACCTGAAGTCTAGGTCCCTATTTATTTTTTT  
ATAGTTATGTTAGTATTAAGAACGTTATTTATATTTCAAATTTTTCTTTTTTTCTGTA  
CAGACGCGTGTACGCATGTAACATTATACTGAAAACCTTGCTTGAGAAGGTTTTGGGAC  
GCTCGAAGATCGCGTCCCAATTCGCCCTATAGTGAGTCGTATTACGCGCGCTCACTGGC  
CGTCGTTTTTACAACGTCGTGACTGGGAAAACCTGGCGTTACCCCTGCAGGACTAGTGC  
10 TGAGGCATTAATAGTTTACTCAATTCTTGAAGCCAATTTGTACAATCCCCATTAGAGT  
CAAATAAAAGGATGCCTCACGGAGGTATGTTACCCGCGCTATTTACATGGCTCATTGA  
ATTAGAGGTGGAATTTGGTGTACCCTCCCCTCCTCATCTGATGAAGTAGTGATCCGACA  
ATTCTTAAAGTTGTAGACATTACTTTTACCACCAACTAAGTTGTATTTATATTGCTAC  
CCTTATCCTTTTATATCTAACTAGCGCTCATAAGGTGGGGCAATACTAAAACGTGTGT  
15 CTTATTCAACTCATTAATAACGTGGCAGTACGTACCCTATTAGAAACAATAGGAAACAG  
CAGAGTCGGAAGAAGCCAAATGCCAGATTTGAAGTCCAAAACCTTGTCAAGCCAATCTT  
TGGGAGCGGCTATTCCCTCCAGAAATTGTGTACCAAATACTTACATACCAGTTTAGGGAT  
TTGTTAAGAAATGACCATCCAGGTACGGCAGAAA

20 **[00505]** SEQ.ID NO: 75

CAATCTGGCGGCTTGAGTTCTCAACATGTTTTATTTTTTACTTATATTGCTGGTAGGGT  
AAAAAATATAACTCCTAGGAATAGGTTGTCTATATGTTTTTGTCTTGCTTCTATAATT  
GTAACAAACAAGGAAAGGGAAAATACTGGGTGTAAAAGCCATTGAGTCAAGTTAGGTCA  
TCCCTTTTATACAAAATTTTTCAATTTTTTTTCCAAGATTCTTGACGATTAATTATTT  
25 TTTTTTTGCGTCTACAGCGTGATGAAAATTTGCGCTGCTGCAAGATGAGCGGGAACGG  
GCGAAATGTGCACGCGCACAACTTACGAAACGCGGATGAGTCACTGACAGCCACCGCAG  
AGGTTCTGACTCCTACTGAGCTCTATTGGAGGTGGCAGAACCGGTACCGGAGGAGGCCG  
CTATAACCGGTTTGAATTTATTGTACAGTGTCACATCAGCATTAAGTCTCAGCGAGC  
TCGCATGGAATGCGTGCGATGAGCGACCTCATGCTATACCTGAGAAAGCAACCTGACCT  
30 ACAGGAAAGAGTTACTCAAGAATAAGAATTTTCGTTTTTAAAACCTAAGAGTCACTTTAA  
AATTTGTATACACTTATTTTTTTTTTATAACTTATTTAATAATAAAAATCATAAATCATAA  
GAAATTCGCTTATTTAGAAGTGTCAACAACGTATCTACCAACGGAATGCGTGCGATTCA  
GGTGGAGTCCAAACCCAACAAAGGATTTGGAATTGGCTTACCGGCAGTGGAGGATTCTT  
TTAGCAACGTGGAAGTGATTTACCGTTGTGCTTGTACCTCTAGCATCGTGGAAAGCA  
35 GCAACACCCTTCTTAACAAAGTTGATTCTTTCTTCTCAGAACCCCATTCATGAAGTC  
AGTCCACATAACAATGTGAGCGGCGATACCAGCGGTAACCTTGGCGTAGTTGATGGAAT  
GCTTGGAAGTACGGGCGTAAGATTGCAAGTAAGCTTGTCTCATGGTTGTACCAACTTGT  
TCGTCTTGAATCTGCTAATCAAGTAACAGTCACCCAAGAAGTAACCCAAATCCAATGA  
AGCTGGACCGTACTTACACAATCCCAGTCTAAGATGTAGATCTTTTGCAACTTAGATG  
40 GGTTACCTTCTTCAAGTTGCAACAAGATGTTCCCAGACCACAAGTCAGCCATGACCAA  
GTTTCTTCGGAGTGCATAACATCGTCAACTAGATCCTTGACAACAGTTGGCAACAATGG  
ATCATCGACGCCGTATTTAGCGGCGTTTGGGATAATAGTTTGGTACAATTGGTCAGAGG  
TGTTTCTACCGACAATGTTACCAGAGAAGAACTTGAATTCTGGGTCTCTCTTCTTTCT  
CTACCTATGTTGTGCAATCTGGCGACGAAACCACCAATCTCGGTACCAACCAATCTAGC  
45 AATATCGGTAGCCAAAGGTGGCTTAGCAGTAACGTAGTCTAATAAGGTCTTCATTTTAC  
CGACATCTTGCATAATCAAAGCATTGTTTTCCAAGTCATAGTTGAGACCTTCTGGAACA  
GAGACAATACCATCAACACCACCCAAAACCTTCTCTGTTAGCCATCATCAACTTGATAGC  
TTGGTATTTCGTAGACAGAACGTTCAACACCGATTTTGAATCTTCATCAGTAGACATGT  
GTGGTTGAGCGTGCTTCAAATGATAGAAGTGTGACCTTGGTATGGAGCGTTCAATTTA  
50 ATACGCCAGGTGACGTTAACGAAACCACCGGATAATCTCTTGACACCAGAAAGTGTCAAC

ATCTAAAGACAAATGCTTGGTCAAATAGGTGATTAAACCGTCTTCAGTCTTCAAGTCAA  
AGGCCATTGTTTTATATTTGTTGTAAAAAGTAGATAATTACTTCCTTGATGATCTGTAA  
AAAAGAGAAAAAGAAAGCATCTAAGAACTTGAAAACTACGAATTAGAAAAGACCAAAT  
ATGTATTTCTTGCAATTGACCAATTTATGCAAGTTTATATATATGTAAATGTAAGTTTCA  
5 CGAGGTTCTACTAAACTAAACCACCCCTTGGTTAGAAGAAAAGAGTGTGTGAGAACAG  
GCTGTTGTTGTCACACGATTTCGGACAATTCTGTTTGAAAGAGAGAGAGTAACAGTACGA  
TCGAACGAACCTTGTCTGGAGATCACAGTGGGCATCATAGCATGTGGTACTAAACCCT  
TTCCCGCCATTCCAGAACCTTCGATTGCTTGTTACAAAACCTGTGAGCCGTCGCTAGGA  
CCTTGTTGTGTGACGAAATTGGAAGCTGCAATCAATAGGAAGACAGGAAGTCGAGCGTG  
10 TCTGGGTTTTTTTTCAGTTTTGTTCTTTTTGCAAACAACAGTTTTATTCTGGCATCCACTA  
AATATAATGGAGCCCGCTTTTTAAGCTGGCATCCAGAAAAAAAAGAATCCCAGCACCA  
AAATATTGTTTTCTTCACCAACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACA  
GAGAACAGGGGCACAAACAGGCAAAAAACGGGCACAACCTCAATGGAGTGATGCAACCT  
GCCTGGAGTAAATGATGACACAAGGCAATTGACCCACGCATGTATCTATCTCATTTTCT  
15 TACACCTTCTATTACCTTCTGCTCTCTCTGATTTGGAAAAAGCTGAAAAAAAAGGTTGA  
AACCAGTTCCTGAAATTATTCCCCTACTTGACTAATAAGTATATAAAGACGGTAGGTA  
TTGATTGTAATTCTGTAAATCTATTTCTTAAACTTCTTAAATTCTACTTTTATAGTTAG  
TCTTTTTTTTTAGTTTTTAAACACCAAGAAGCTTAGTTTCGAATAAACACACATAAACAAA  
CAAATGCATATCAGAAACCCATATAGAAGTCCAATTGACTACCAAGCTTTGTCTGAAG  
20 CTTTCCACCATTGAAGCCATTGTTTTCCGTTAACGCTGATGGTACCTCCTCAGTTGAC  
TTGACCATTCAGAAAGCCAAAGAGCTTTTACCGCTGCCCTTTTGCACAGAGACTTCGG  
CTTGACTATGACTATCCCAGAAAGATCGTTTGTGTCCAACCGTTCCAAACAGATTGAACT  
ACGTTTTGTGGATTGAAGACATTTTCAACTACACCAACAAGACTTTGGGTTTATCTGAC  
GACCGTCCAATCAAGGGTGTTGATATCGGTACCGGTGCTTCTGCCATTTACCCAATGTT  
25 GGCTTGCGCCAGATTCAAGGCTTGGTCCATGGTTGGTACTGAAGTTGAAAGAAAGTGTA  
TCGACACTGCTAGATTAAACGTTGTTGCTAACAACCTTGCAAGATCGTCTATCCATCTTG  
GAAACCTCTATTGACGGTCCAATTTTAGTCCCAATTTTCGAAGCTACCGAAGAATACGA  
ATACGAATTCACCATGTGTAACCCACCTTTCTACGATGGTGCCGCTGACATGCAAACCTA  
GCGATGCTGCAAAGGGTTTTGTTTTCGGTGTGCGTGCTCCACACTCTGGTACAGTCATC  
30 GAAATGTCTACTGAAGGTGGTGAATCCGCTTTTCGTGGCTCAAATGGTTAGAGAATCTCT  
TAAGTTGAGAACCAGATGTAGATGGTACACTTCTAACTTAGGTAAGTTGAAATCTTTGA  
AGGAAATTGTCGGTTTTGTTGAAGGAATTAGAAATCTCTAATTACGCCATCAACGAATAT  
GTCCAAGGTTCCACTAGAAGATACGCTGTGCTTGGAGTTTCACTGATATCCAATTGCC  
AGAAGAATTGTCCAGACCATCTAATCCTGAATTGTCCTCTTTGTTGCGACTACAAGGATG  
35 ACGATGACAAATGAATCGCGTGCATTATCCGCTCTAACCGAAAAGGAAGGAGTTAGAC  
AACCTGAAGTCTAGGTCCCTATTTATTTTTTTTATAGTTATGTTAGTATTAAGAAGCTTA  
TTTATATTTCAAATTTTTCTTTTTTTTTCTGTACAGACGCGTGACGCATGTAACATTAT  
ACTGAAAACCTTGCTTGAGAAGGTTTTGGGACGCTCGAAGATCGCGTACCCAATTCGCC  
CTATAGTGAGTCGTATTACGCGCGCTCACTGGCCGTCGTTTTTACAACGTCGTGACTGGG  
40 AAAACCTTGCGTTACCCCTGCAGGACTAGTGCTGAGGCATTAATGCAACTCAGAAGTT  
TGACAGCAAGCAAGTTCATCATTCGAAGTACGCTTATTGTTTTAGTTTCAGTGACAGCGA  
ACTGCCGTACTCGATGCTTTATTTCTCACGGTAGAGCGGAAGAACAGATAGGGGCAGCG  
TGAGAAGAGTTAGAAAGTAAATTTTTATCACGTCTGAAGTATTCTTATTCATAGGAAAT  
TTTGCAAGGTTTTTTAGCTCAATAACGGGCTAAGTTATATAAGGTGTTACGCGATTTT  
45 CTTGTTATGTATACCTCTTCTCTGAGGAATGGTACTACTGTCTGATGTAGGCTCCTTA  
AATTGGTGGGCAAGAATAACTTATCGATATTTTGTATATTGGTCTTGGAGTTCACCACG  
TAATGCCTGTTTAAAGACCATCAGTTAACTCTAGTATTATTTGGTCTTGGCTACTGGCCG  
TTTGCTATTATTCAAGTCTTTTTGTGCCTTCCCGTCGGGTAAGGGAGTTATTTAGGGATA  
CAGAATCTAACGAAAACCTAAATCTCAATGATTAACCTCTATTTAATCCTTTTTTGAAAGG  
50 CAAAAGAGGTCCCTTGTTCACTTACAACGTTCTTAGCCAAATTCGCTTATCACTTACTA

CTTCACGATATACAGAAGTAAAAACATATAAAAAGATGTCTGTTTGTGTTAGCCATCACA  
AAAGGTATCGCAG

**[00506]** SEQ.ID NO: 76

5 TGTGTTTATATTTGTTGTAAAAAGTAGATAATTACTTCCTTGATGATCTGTAAAAAGAG  
AAAAAGAAAGCATCTAAGAACTTGAAAACTACGAATTAGAAAAGACCAAATATGTATT  
TCTTGCAATTGACCAATTTATGCAAGTTTATATATATGTAAATGTAAGTTTCACGAGGTT  
CTACTAACTAAACCACCCCCCTTGGTTAGAAGAAAAGAGTGTGTGAGAACAGGCTGTTG  
TTGTCACACGATTTCGACAATTCTGTTTGAAAGAGAGAGAGTAACAGTACGATCGAACG  
10 AACTTTGCTCTGGAGATCACAGTGGGCATCATAGCATGTGGTACTAAACCCTTTCCCGC  
CATTCCAGAACCTTCGATTGCTTGTTACAAAACCTGTGAGCCGTCGCTAGGACCTTGTT  
GTGTGACGAAATTGGAAGCTGCAATCAATAGGAAGACAGGAAGTCGAGCGTGTCTGGGT  
TTTTTTCAGTTTTGTTCTTTTTGCAAACAACAGTTTATTCCTGGCATCCACTAAATATAA  
TGGAGCCCGCTTTTTAAGCTGGCATCCAGAAAAAAAAGAATCCCAGCACCAAAATATT  
15 GTTTTCTTCACCAACCATCAGTTCATAGGTCCATTCTCTTAGCGCAACTACAGAGAACA  
GGGGCACAAACAGGCAAAAAACGGGCACAACCTCAATGGAGTGATGCAACCTGCCTGGA  
GTAAATGATGACACAAGGCAATTGACCCACGCATGTATCTATCTCATTTTTCTTACACCT  
TCTATTACCTTCTGCTCTCTCTGATTTGAAAAAGCTGAAAAAAAGGTTGAAACCACT  
TCCCTGAAATTATTCCCCTACTTGACTAATAAGTATATAAAGACGGTAGGTATTGATTG  
20 TAATTCTGTAAATCTATTTCTTAACTTCTTAAATTCTACTTTTATAGTTAGTCTTTTT  
TTTAGTTTTTAAACACCAAGAAGCTTAGTTTCGAATAAACACACATAAACAAACAAAGGA  
TCCATGATGAAGTTCTGGAGAAAGTACACACAACAAGAAATGGATGAAAAGATTACTGA  
ATCTTTGAAAAGACTTTGAACTACGATAACACTAAGACAATCGGTATTCCAGGTACTA  
AGTTGGATGATACAGTTTTCTATGATGATCATTCTTTCGTTAAGCATTACCATACTTG  
25 AGAACTTTTATTCAAACCCAAACCATATCGGTTGTCATACTTATGATAAGGCTGATAT  
CTTGTTTCGGTGGTACATTTCGATATCGAAAGAGAATTAATCCAATTGTTAGCAATCGATG  
TTTTGAACGGTAACGATGAAGAATTTGATGGTTACGTTACTCAAGGTGGTACAGAAGCT  
AACATCCAAGCAATGTGGGTTTACAGAACTACTTCAAGAAAGAAAGAAAGGCTAAGCA  
TGAAGAAATCGCTATCATCACTTCAGCAGATACACATTACTCTGCATACAAAGGTTTCA  
30 ATTTGTTGAACATCGATATTATTAAGGTTCCAGTTGATTTTTATTCAAGAAAAATTCAA  
GAAAATACATTGGATTCAATTGTTAAAGAAGCTAAAGAAATTGGTAAAAAGTACTTCAT  
CGTTATCTCTAACATGGGTACTACAATGTTTGGTTTCAGTTGATGATCCAGATTTGTACG  
CTAACATCTTCGATAAGTACAATTTGGAATACAAAATTCATGTTGATGGTGCATTTGGT  
GGTTTTATATATCCAATTGATAATAAGGAATGTAAAACCTGATTTCTCTAATAAGAAGCT  
35 TTCTTCAATCACATTAGATGGTCATAAGATGTTGCAAGCTCCATACGGTACTGGTATCT  
TCGTTTTCAAGAAAGAATTTGATCCATAACACTTTGACAAAGGAAGCAACTTACATCGAA  
AATTTGGATGTTACATTGTCTGGTTCAAGATCTGGTTCAAATGCTGTTGCAATTTGGAT  
GGTTTTAGCTTCTTATGGTCCATACGGTTGGATGGAAAAGATTAATAAGTTGAGAAATA  
GAACTAAATGGTTGTGTAAGCAATTGAACGATATGAGAATTAAATATTACAAAGAAGAT  
40 TCAATGAATATTGTTACAATTGAAGAACAATATGTTAATAAGGAAATCGCTGAAAAGTA  
CTTTTTAGTTCCAGAAGTTTATAACCCAACTAACAACTGGTACAAGATCGTTGTTATGG  
AACATGTTGAATTGGATATCTTGAACCTTTGGTTTACGATTTGAGAAAGTTTAATAAG  
GAACATTTGAAGGCAATGCATCATCATCATCATTAACCGCGGCTAGCTAAGATCCG  
CTCTAACCGAAAAGGAAGGAGTTAGACAACCTGAAGTCTAGGTCCCTATTTATTTTTTT  
45 ATAGTTATGTTAGTATTAAGAAGCTTATTTATATTTCAAATTTTTCTTTTTTTCTGTA  
CAGACGCGTGTACGCATGTAACATTATACTGAAAACCTTGCTTGAGAAGGTTTTGGGAC  
GCTCGAAGATCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTTGCGT  
ATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTGCTTCGGCTGCG  
GCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATA  
50 ACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCC

GCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACG  
CTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAACCAGGCGTTTCCCCCTG  
GAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCC  
TTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTC  
5 GGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCCGACC  
GCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCG  
CCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTAC  
AGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCT  
GCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAA  
10 CAAACCACCGCTGGTAGCGGTGGTTTTTTTTTGTGTTGCAAGCAGCAGATTACGCGCAGAAA  
AAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACG  
AAAACCTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATC  
CTTTTAAATTAAAAATGAAGTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTC  
TGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTT  
15 CATCCATAGTTGCCTGACTCCCCGTCTGTAGATAACTACGATACGGGAGGGCTTACCA  
TCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACC GGCTCCAGATTTATC  
AGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCAGAAAGTGGTCCTGCAACTTTATCCG  
CCTCCATCCAGTCTATTAATTGTTGCCGGAAGCTAGAGTAAGTAGTTCGCCAGTTAAT  
AGTTTGC GCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCGTTTGG  
20 TATGGCTTCATTACGCTCCGGTCCCCAACGATCAAGGCGAGTTACATGATCCCCCATGT  
TGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCC  
GCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTATGCCATC  
CGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTA  
TGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGC  
25 AGAACTTTAAAGTGCTCATCATTGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGAT  
CTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAG  
CATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCA  
AAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATA  
TTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTT  
30 AGAAAAATAAACAAATAGGGGTTCGCGGCACATTTCCCCGAAAAGTGCCACCTGAACGA  
AGCATCTGTGCTTCATTTTGTAGAACAAAAATGCAACGCGAGAGCGCTAATTTTTCAAA  
CAAAGAATCTGAGCTGCATTTTTTACAGAACAGAAATGCAACGCGAAAGCGCTATTTTAC  
CAACGAAGAATCTGTGCTTCATTTTTGTAAAACAAAAATGCAACGCGAGAGCGCTAATT  
TTTTCAAACAAAGAATCTGAGCTGCATTTTTTACAGAACAGAAATGCAACGCGAGAGCGCT  
35 ATTTTACCAACAAAGAATCTATACTTCTTTTTTGTCTACAAAAATGCATCCCGAGAGC  
GCTATTTTTCTAACAAAGCATCTTAGATTACTTTTTTCTCCTTTGTGCGCTCTATAAT  
GCAGTCTCTTGATAACTTTTTGCACTGTAGGTCCGTTAAGGTTAGAAGAAGGCTACTTT  
GGTGTCTATTTTCTCTTCCATAAAAAAAGCCTGACTCCACTTCCCGCGTTTACTGATTA  
CTAGCGAAGCTGCGGGTGCATTTTTTCAAGATAAAGGCATCCCCGATTATATTCTATAC  
40 CGATGTGGATTGCGCATACTTTGTGAACAGAAAGTGATAGCGTTGATGATTCTTCATTG  
GTCAGAAAATTATGAACGGTTTCTTCTATTTTGTCTCTATATACTACGTATAGGAAATG  
TTTACATTTTCGTATTGTTTTCGATTCACTCTATGAATAGTTCTTACTACAATTTTTTT  
GTCTAAAGAGTAATACTAGAGATAAACATAAAAAATGTAGAGGTCGAGTTTAGATGCAA  
GTTCAAGGAGCGAAAGGTGGATGGGTAGGTTATATAGGGATATAGCACAGAGATATATA  
45 GCAAAGAGATACTTTTGAGCAATGTTTGTGGAAGCGGTATTCGCAATATTTTAGTAGCT  
CGTTACAGTCCGGTGCGTTTTTGGTTTTTGAAGTGCGTCTTCAGAGCGCTTTTGGTT  
TTCAAAAGCGCTCTGAAGTTCCTATACTTTCTAGAGAATAGGAACTTCGGAATAGGAAC  
TTCAAAGCGTTTCCGAAAACGAGCGCTTCGAAAATGCAACGCGAGCTGCGCACATACA  
GCTCACTGTTACGTCGCACCTATATCTGCGTGTTGCCTGTATATATATATACATGAGA  
50 AGAACGGCATAGTGCGTGTTTATGCTTAAATGCGTACTTATATGCGTCTATTTATGTAG  
GATGAAAGGTAGTCTAGTACCTCCTGTGATATTATCCCATTCATGCGGGGTATCGTAT

GCTTCCTTCAGCACTACCCTTTAGCTGTTCTATATGCTGCCACTCCTCAATTGGATTAG  
TCTCATCCTTCAATGCTATCATTTTCCTTTGATATTGGATCATACTAAGAAACCATTATT  
ATCATGACATTAACTATAAAAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTT  
CGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTC  
5 TGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGG  
TGTCGGGGCTGGCTTAACCTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATAT  
TTCACACCGCATAGATCCGTCGAGTTCAAGAGAAAAAAGAAAAAGCAAAAAGAAAA  
AAGGAAAGCGCGCCTCGTTTCAGAATGACACGTATAGAATGATGCATTACCTTGTCTCT  
TCAGTATCATACTGTTTCGTATACATACTTACTGACATTTCATAGGTATACATATATACAC  
10 ATGTATATATATCGTATGCTGCAGCTTTAAATAATCGGTGTCACTACATAAGAACACCT  
TTGGTGGAGGGAACATCGTTGGTACCATTGGGCGAGGTGGCTTCTCTTATGGCAACCGC  
AAGAGCCTTGAACGCACTCTCACTACGGTGATGATCATTCTTGCCTCGCAGACAATCAA  
CGTGGAGGGTAATTCTGCTAGCCTCTGCAAAGCTTTCAAGAAAATGCGGGATCATCTCG  
CAAGAGAGATCTCCTACTTTCTCCCTTTGCAAACCAAGTTCGACAACTGCGTACGGCCT  
15 GTTCGAAAGATCTACCACCGCTCTGGAAAGTGCCTCATCCAAAGGCGCAAATCCTGATC  
CAAACCTTTTTACTCCACGCACGGCCCCCTAGGGCCTCTTTAAAAGCTTGACCGAGAGCA  
ATCCCGCAGTCTTCAGTGGTGTGATGGTCGTCTATGTGTAAGTCACCAATGCACTCAAC  
GATTAGCGACCAGCCGGAATGCTTGGCCAGAGCATGTATCATATGGTCCAGAAACCTTA  
TACCTGTGTGGACGTTAATCACTTGCATTTGTGTGGCCTGTTCTGCTACTGCTTCTGCC  
20 TCTTTTTCTGGGAAGATCGAGTGCTCTATCGCTAGGGGACCACCCTTTAAAGAGATCGC  
AATCTGAATCTTGGTTTCATTTGTAATACGCTTTACTAGGGCTTTCTGCTCTGTCATCT  
TTGCCTTCGTTTATCTTGCCTGCTCATTTTTTAGTATATTCTTCGAAGAAATCACATTA  
CTTTATATAATGTATAATTCATTATGTGATAATGCCAATCGCTAAGAAAAAAGAGT  
CATCCGCTAGGTGGAAAAAATGAAAATCATTACCGAGGCATAAAAAATATAGA  
25 GTGTACTAGAGGAGGCCAAGAGTAATAGAAAAAGAAAATTGCGGGAAAGGACTGTGTTA  
TGACTTCCCTGACTAATGCCGTGTTCAAACGATACCTGGCAGTGACTCCTAGCGCTCAC  
CAAGCTCTTAAACGGGAATTTATGGTGCATCTCAGTACACGCGCCAGATCTGTTTAG  
CTTGCCTCGTCCCCGCCGGGTACCCGGCCAGCGACATGGAGGCCAGAAATACCCTCCT  
TGACAGTCTTGACGTGCGCAGCTCAGGGGCATGATGTGACTGTGCGCCGTACATTTAGC  
30 CCATACATCCCCATGTATAATCATTTGCATCCATACATTTTGATGGCCGCACGGCGCGA  
AGCAAAATACGGCTCCTCGCTGCAGACCTGCGAGCAGGGAAACGCTCCCCCTCACAGA  
CGCGTTGAATTGTCCCCACGCCGCGCCCCCTGTAGAGAAATATAAAAGGTTAGGATTTGC  
CACTGAGGTTCTTCTTTTATATACTTCCCTTTTAAATCTTGCTAGGATACAGTTCTCAC  
ATCACATCCGAACATAAACAACCATGGGTAAGGAAAAGACTCACGTTTCGAGGCCGCGA  
35 TTAAATTCGAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGG  
GCAATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTC  
TGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAAC  
TGGCTGACGGAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGA  
TGCATGGTTACTCACCCTGCGATCCCCGGCAAACAGCATTCAGGTATTAGAAGAAT  
40 ATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCTGCGCCGGTTGCAT  
TCGATTCCTGTTTGTAAATTGTCCTTTTAAACAGCGATCGCGTATTTTCGTCTCGCTCAGGC  
GCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATG  
GCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAGCTTTTGCCATTCTCACCGGAT  
TCAGTCGTCACTCATGGTGATTTCTCACTTGATAACCTTATTTTGACGAGGGGAAAT  
45 AATAGGTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCA  
TCCTATGGAAGTGCCTCGGTGAGTTTCTCCTTCATTACAGAAACGGCTTTTCAAAAA  
TATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTT  
TTTCTAATCAGTACTGACAATAAAAAGATTCTTGTTTTCAAGAACTTGTCAATTTGTATA  
GTTTTTTTATATTGTAGTTGTTCTATTTTAAATCAAATGTTAGCGTGATTTATATTTTTT  
50 TTCGCTCGACATCATCTGCCAGATGCGAAGTTAAGTGCAGAGAAAGTAATATCATGC  
GTCAATCGTATGTGAATGCTGGTCGCTATACTGCTGTCGATTTCGATACTAACGCCGCCA

TCCAGTGTGCAATTCGCCATTTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGG  
 CCTCTTCGCTATTACGCCAGCTGAATTGGAGCGACCTCATGCTATACCTGAGAAAGCAA  
 CCTGACCTACAGGAAAGAGTTACTCAAGAATAAGAATTTTCGTTTTAAACCTAAGAGT  
 CACTTTAAATTTGTATACACTTATTTTTTTTATAACTTATTTAATAATAAAAATCATA  
 5 AATCATAAGAAATTCGCTTATTTAGAAGTGTCAACAACGTATCTACCAACGATTTGACC  
 CTTTTCCATCTTTTCGTAAATTTCTGGCAAGGTAGACAAGCCGACAACCTTGATTGGAG  
 ACTTGACCAAACCTCTGGCGAAGAATTGTTAATTAAGAGCTCAGATCTTTTGCGGCCGC

**[00507] SEQ.ID NO: 77**

10 GTTTGCAAAAAGAACAAAACCTGAAAAACCCAGACACGCTCGACTTCCTGTCTTCCTAT  
 TGATTGCAGCTTCCAATTTTCGTACACAACAAGGTCCTAGCGACGGCTCACAGGTTTTG  
 TAACAAGCAATCGAAGGTTCTGGAATGGCGGGAAAGGGTTTAGTACCACATGCTATGAT  
 GCCCACTGTGATCTCCAGAGCAAAGTTCGTTTCGATCGTACTGTTACTCTCTCTCTTTCA  
 AACAGAAATGTCCGAATCGTGTGACAACAACAGCCTGTTCTCACACACTCTTTTCTTCT  
 15 AACCAAGGGGGTGGTTTTAGTTTTAGTAGAACCTCGTGAACTTACATTTACATATATATA  
 AACTTGCATAAATTGGTCAATGCAAGAAATACATATTTGGTCTTTTCTAATTCGTAGTT  
 TTTCAAGTTCTTAGATGCTTTCTTTTTCTCTTTTTTACAGATCATCAAGGAAGTAATTA  
 TCTACTTTTTACAACAAATATAAAACA

**20 [00508] SEQ.ID NO: 78**

ACAGTTTATTCCTGGCATCCACTAAATATAATGGAGCCCGCTTTTTTAAGCTGGCATCCA  
 GAAAAAAAAAGAATCCCAGCACCAAAATATTGTTTTCTTCACCAACCATCAGTTCATAG  
 GTCCATTCTCTTAGCGCAACTACAGAGAACAGGGGCACAAACAGGCCAAAAACGGGCAC  
 AACCTCAATGGAGTGATGCAACCTGCCTGGAGTAAATGATGACACAAGGCAATTGACCC  
 25 ACGCATGTATCTATCTCATTTTCTTACACCTTCTATTACCTTCTGCTCTCTCTGATTTG  
 GAAAAAGCTGAAAAAAAAGGTTGAAACCAGTTCCCTGAAATTATTCCCCTACTTGACTA  
 ATAAGTATATAAAGACGGTAGGTATTGATTGTAATTCTGTAAATCTATTTCTTAACTT  
 CTTAAATTCTACTTTTATAGTTAGTCTTTTTTTTTTAGTTTTTAAACACCAAGAACTTAGT  
 TTCGAATAAACACACATAAACAAACAAA

30

**[00509] SEQ.ID NO: 79**

GCTGGATTGAGCTGAATGGTGCCAGGTCGAGGCTGGGAGGGAGACTAACTCGAAAGTGA  
 CGAAGACTCGAAAATTAAAAAAAAGATACTGCAGAAGGCAAGATTGAGAATGGAGTAA  
 AGGCAGCGTGGGTCCCCTGTGGAAACCGCAGTTTTCTGCGCCAAGTGGTACCGGTGCG  
 35 AGTGCAGCAATTAATCTCTCGATATTTTCTTAGTATCTCTTTTTTATATAAGAATATATT  
 TTGGAATTGGTAATGCTTATCTTCAATAGTTTCTTAGTTGAATGCACACTTAAGAGCAA  
 ATTGGCCAAGGAGTTCTTCGTTTCGCTTTAATTTATTTCTGTTATTGTCAATTTATTC  
 ATCCCATCTCCCAGGATAGAAGAAATTAGTGTAATTTTGCTGACAATACATTTTAACG  
 ACGATAACAATAATAGCAATTAAATAAAATAGCACTACCACCACTCCACTGCTCGTTAG  
 40 CTATTTCTGTAAAATAAATAAAAAGATC

**[00510] SEQ.ID NO: 80**

ATTCGCGCTATCTCGATTTCTACCTATATAGTTAATCTCTGTACAAAAACAATCTTTCC  
 AACTATCCATTAATCATAGTATATTATCAGCGTCGGCGATTTTACCACGCTTGACAAAA  
 45 GCCGCGGGCGGGATTCTGTGGGTAGTGGCACCAGGCAATTAATCTAATCAAAGGCGCTT  
 GAAGGAAGAGATAGATAATAGAACAAAGCAATCGCCGCTTTGGACGGCAAATATGTTTA  
 TCCATTGGTGCGGTGATTGGATATGATTTGTCTCCAGTAGTATAAGCAAGCGCCAGATC  
 TGTTTACTGTAAAATTAAAGTGAGTAATCTCGCGGGATGTAATGATTTAAGGGAATCTGG

TTCAGGTTTTACATATATTTGTATATAAGGCCATTTGTAATTTCAATAGTTTTAGGAT  
TTTTCCCTTCTCCCAAATACTCACTTACTGTGTTACATTACAGAAAGAACAGACAAGAA  
ACCGTCAATAAGAAATATAACTAAGAACA

5 **[00511]** SEQ.ID NO: 81  
TACCCATACGACGTTCCAGACTACGCC

**[00512]** SEQ.ID NO: 82  
YPYDVDPDYA

10 **[00513]** SEQ.ID NO: 83  
GAACAAAAGTTAATTTCTGAAGAAGATTTGGAA

**[00514]** SEQ.ID NO: 84  
15 EQKLISEEDL

**[00515]** SEQ.ID NO: 85  
GACTACAAGGATGACGATGACAAA

20 **[00516]** SEQ.ID NO: 86  
DYKDDDDK

**[00517]** SEQ.ID NO: 87  
GGTAAGCCAATTCCAAATCCTTTGTTGGGTTTGGACTCCACC

25 **[00518]** SEQ.ID NO: 88  
GKPIPNPLLGLDST

**[00519]** SEQ.ID NO: 89  
30 CATCATCATCATCATCAT

**[00520]** SEQ.ID NO: 90  
HHHHHH

**[00521]** SEQ.ID NO: 91  
35 CAGGTGGCACTTTTCGGGGAAAATGTGCGCGGAACCCCTATTTGTTTATTTTCTAAATA  
CATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTG  
AAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGG  
CATTTTGCCTTCCTGTTTTTGTCTACCCAGAAACGCTGGTGAAAGTAAAGATGCTGAA  
GATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCT



TGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCTAT  
 GTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTGCGCGCATACAC  
 TATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGG  
 CATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCA  
 5 ACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATG  
 GGGGATCATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAA  
 CGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACCTATTAA  
 CTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGAT  
 AAAGTTGCAGGACCCTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAA  
 10 ATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTA  
 AGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGA  
 AATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACCTGTCAGACCA  
 AGTTTACTCATATATACTTTAGATTGATTTAAACTTCATTTTTTAATTTAAAGGATCT  
 AGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTC  
 15 CACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCT  
 GCGCGTAATCTGCTGCTTGCAAACAAAAAACACCGCTACCAGCGGTGGTTTGGTTTGC  
 CGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATA  
 CCAAATACTGTTCTTCTAGTG TAGCCGTAGTTAGGCCACCCTTCAAGAACTCTGTAGC  
 ACCGCTACATACTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATA  
 20 AGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCTG  
 GGCTGAACGGGGGGTTCGTGCACACAGCCAGCTTGGAGCGAACGACCTACACCGAACT  
 GAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGG  
 ACAGGTATCCGGTAAGCGGCAGGGTCCGAACAGGAGAGCGCACGAGGGAGCTTCCAGGG  
 GGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCG  
 25 ATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGA AAAACGCCAGCAACGCGGCCT  
 TTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCC  
 CCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAG  
 CCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCA  
 AACCGCCTCTCCCCGCGCGTTGGCCGATTCAATTAATGCAGCTGGATCTTCGAGCGTCCC  
 30 AAAACCTTCTCAAGCAAGGTTTTTCAGTATAATGTTACATGCGTACACGCGTCTGTACAG  
 AAAAAAAGAAAAATTTGAAATATAAATAACGTTCTTAATACTAACATAACTATAAAAA  
 AATAAATAGGGACCTAGACTTCAGGTTGTCTAACTCCTTCCTTTTCGGTTAGAGCGGAT  
 CTTAGCTAGCCGCGGTACCAAGCTGGTGGATCCTTTGTTTGTGTTTATGTGTGTTTATTTCG  
 AAATAAGTTCTTGGTGTGTTTTAAACTAAAAAAGACTAACTATAAAAGTAGAATTTA  
 35 AGAAGTTTAAGAAATAGATTTACAGAATTACAATCAATACCTACCGTCTTTATATACTT  
 ATTAGTCAAGTAGGGGAATAATTTAGGGAAGTGGTTTCAACCTTTTTTTTTCAGCTTTT  
 TCCAAATCAGAGAGAGCAGAAGGTAATAGAAGGTGTAAGAAAATGAGATAGATACATGC  
 GTGGGTCAATTGCCTTGTGTATCATTTACTCCAGGCAGGTTGCATCACTCCATTGAGG  
 TTGTGCCCCTTTTTTGCCTGTTTGTGCCCCTGTTCTCTGTAGTTGCGCTAAGAGAATGG  
 40 ACCTATGAACTGATGGTTGGTGAAGAAAAACAATATTTTGGTGCTGGGATTCTTTTTTTT  
 TCTGGATGCCAGCTTAAAAAGCGGGCTCCATTATATTTAGTGGATGCCAGGAATAAACT  
 GTTGTGTTGCAAAAAGAACAAAACCTGAAAAAACCAGACACGCTCGACTTCCTGTCTTCC  
 TATTGATTGCAGCTTCCAATTTCTGTACACAACAAGGTCCTAGCGACGGCTCACAGGTT  
 TTGTAACAAGCAATCGAAGGTTCTGGAATGGCGGGAAAGGGTTAGTACCACATGCTAT  
 45 GATGCCCCTGTGATCTCCAGAGCAAAGTTCGTTTCGATCGTACTGTTACTCTCTCTCTT  
 TCAAACAGAATTGTCCGAATCGTGTGACAACAACAGCCTGTTCTCACACACTCTTTTCT  
 TCTAACCAAGGGGGTGGTTTAGTTTAGTAGAACCTCGTGAACTTACATTTACATATAT  
 ATAACTTGCATAAATTGGTCAATGCAAGAAATACATATTTGGTCTTTTCTAATTCGTA  
 GTTTTTCAAGTTCTTAGATGCTTTCTTTTCTCTTTTACAGATCATCAAGGAAGTAA  
 50 TTATCTACTTTTTACAACAAATATAAAACAGCGGCCGAAAAGATCTGAGCTCTTAATT  
 AACAATCTTCGCCAGAGGTTTGGTCAAGTCTCCAATCAAGGTTGTCGGCTTGTCTACC

TTGCCAGAAATTTACGAAAAGATGGAAAAGGGTCAAATCGTTGGTAGATACGTTGTTGA  
 CACTTCTAAATAAGCGAATTTCTTATGATTTATGATTTTTATTATTAATAAGTTATAA  
 AAAAAATAAGTGTATACAAATTTTAAAGTGAAGTCTTAGGTTTTAAACGAAAATTCTTA  
 TTCTTGAGTAACTCTTTCCTGTAGGTGAGGTTGCTTCTCAGGTATAGCATGAGGTGCG  
 5 TCCAATTGAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTG  
 CGCAGCCTGAATGGCGAATTCGACACTGGATGGCGGCGTTAGTATCGAATCGACAGCAG  
 TATAGCGACCAGCATTCACATACGATTGACGCATGATATTACTTTCTGCGCACTTAACT  
 TCGCATCTGGGCAGATGATGTGAGGCGGAAAAAATATAAATCACGCTAACATTTGAT  
 TAAAAATAGAACAACTACAATATAAAAAAACTATACAAATGACAAGTTCTTGAAAACAAG  
 10 AATCTTTTTTATTGTCAGTACTGATTAGAAAACTCATCGAGCATCAAATGAACTGCAA  
 TTTATTCATATCAGGATTATCAATACCATATTTTTTGAAAAAGCCGTTTCTGTAATGAAG  
 GAGAAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGCGATT  
 CCGACTCGTCCAACATCAATACAACCTATTAATTTCCCCTCGTCAAAAATAAGGTTATC  
 AAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGCTTATGCA  
 15 TTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACTCGCA  
 TCAACCAAACCGTTATTTCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGATCGCT  
 GTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCCAGCG  
 CATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGTTTTG  
 CCGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCTTGAT  
 20 GGTGCGAAGAGGCATAAATTCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTAACAT  
 CATTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACCTCTGGCGCATCGGGCTTCCCA  
 TACAATCGATAGATTGTGCGACCTGATTGCCCGACATTATCGCGAGCCCATTTATACCC  
 ATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCTCGAAACGTGAGTCTTTTCT  
 TACCCATGGTTGTTTATGTTGCGATGTGATGTGAGAACTGTATCCTAGCAAGATTTTAA  
 25 AAGGAAGTATATGAAAGAAGAACCTCAGTGGCAAATCCTAACCTTTTATATTTCTCTAC  
 AGGGGCGCGCGTGGGGACAATTC AACCGCTCTGTGAGGGGAGCGTTTCCCTGCTCGCA  
 GGTCTGCAGCGAGGAGCCGTAATTTTTGCTTCGCGCCGTGCGGCCATCAAAATGTATGG  
 ATGCAAATGATTATACATGGGGATGTATGGGCTAAATGTACGGGCGACAGTCACATCAT  
 GCCCCTGAGCTGCGCACGTCAAGACTGTCAAGGAGGGTATTCTGGGCCTCCATGTCGCT  
 30 GGCCGGGTGACCCGGCGGGGACGAGGCAAGCTAAACAGATCTGGCGCGTGTACTGAGAG  
 TGCACCATAAATTTCCGTTTTAAGAGCTTGGTGAGCGCTAGGAGTCACTGCCAGGTATC  
 GTTTGAACACGGCATTAGTCAAGGAAGTCATAACACAGTCCTTTCCCGCAATTTTCTTT  
 TTCTATTACTCTTGCCCTCCTCTAGTACACTCTATATTTTTTTTATGCCTCGGTAATGAT  
 TTTTATTTTTTTTTTTTCCACCTAGCGGATGACTCTTTTTTTTTTCTTAGCGATTGGCATT  
 35 ATCACATAATGAATTATACATTATATAAAGTAATGTGATTTCTTGAAGAATATACTAA  
 AAAATGAGCAGGCAAGATAAACGAAGGCAAAGATGACAGAGCAGAAAGCCCTAGTAAAG  
 CGTATTACAAATGAAACCAAGATTGAGATTGCGATCTCTTTAAAGGGTGGTCCCCTAGC  
 GATAGAGCACTCGATCTTCCCAGAAAAAGAGGCAGAAGCAGTAGCAGAACAGGCCACAC  
 AATCGCAAGTGATTAACGTCCACACAGGTATAGGGTTTCTGGACCATATGATACATGCT  
 40 CTGGCCAAGCATTCCGGCTGGTTCGCTAATCGTTGAGTGCATTGGTGACTTACACATAGA  
 CGACCATCACACCACTGAAGACTGCGGGATTGCTCTCGGTCAAGCTTTTAAAGAGGCC  
 TAGGGGCCGTGCGTGAGTAAAAAGGTTTGGATCAGGATTTGCGCCTTTGGATGAGGCA  
 CTTTCCAGAGCGGTGGTAGATCTTTCGAACAGGCCGTACGCAGTTGTGCAACTTGGTTT  
 GCAAAGGGAGAAAGTAGGAGATCTCTCTTGCAGATGATCCCGCATTTTCTTGAAAGCT  
 45 TTGCAGAGGCTAGCAGAATTACCTCCACGTTGATTGTCTGCGAGGCAAGAATGATCAT  
 CACCGTAGTGAGAGTGCGTTCAAGGCTCTTGCAGTTGCCATAAGAGAAGCCACCTCGCC  
 CAATGGTACCAACGATGTTCCCTCCACCAAGGTGTTCTTATGTAGTGACACCGATTAT  
 TTAAAGCTGCAGCATACGATATATATACATGTGTATATATGTATACCTATGAATGTCAG  
 TAAGTATGTATACGAACAGTATGATACTGAAGATGACAAGGTAATGCATCATTCTATAC  
 50 GTGTCATTCTGAACGAGGCGCGCTTTCTTTTTTCTTTTTGCTTTTTCTTTTTTTTTCT  
 CTTGAACTCGACGGATCTATGCGGTGTGAAATATGGTGCACCTCAGTACAATCTGCTC

5 TGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGAC  
GGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGC  
ATGTGTCAGAGGTTTTACCGTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGAT  
ACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGTATGATCCAATAT  
CAAAGGAAATGATAGCATTGAAGGATGAGACTAATCCAATTGAGGAGTGGCAGCATATA  
10 GAACAGCTAAAGGGTAGTGCTGAAGGAAGCATACGATACCCCGCATGGAATGGGATAAT  
ATCACAGGAGGTACTAGACTACCTTTCATCCTACATAAATAGACGCATATAAGTACGCA  
TTTAAGCATAAACACGCACTATGCCGTTCTTCTCATGTATATATATACAGGCAACAC  
GCAGATATAGGTGCGACGTGAACAGTGAGCTGTATGTGCGCAGCTCGCGTTGCATTTTC  
GGAAGCGCTCGTTTTTCGGAACGCTTTGAAGTTCCTATTCCGAAGTTCCTATTCTCTAG  
AAAGTATAGGAACCTTCAGAGCGCTTTTGAAAACCAAAAGCGCTCTGAAGACGCACTTTC  
AAAAAACCAAAACGCACCGGACTGTAACGAGCTACTAAAATATTGCGAATACCGCTTC  
CACAAACATTGCTCAAAAGTATCTCTTTGCTATATATCTCTGTGCTATATCCCTATATA  
ACCTACCCATCCACCTTTCGCTCCTTGAACCTTGCATCTAAACTCGACCTCTACATTTTT  
15 TATGTTTATCTCTAGTATTACTCTTTAGACAAAAAATTGTAGTAAGAACTATTCATAG  
AGTGAATCGAAAACAATACGAAAATGTAAACATTTCTTATACGTAGTATATAGAGACAA  
AATAGAAGAAACCGTTTCATAATTTTCTGACCAATGAAGAATCATCAACGCTATCACTTT  
CTGTTTACAAAGTATGCGCAATCCACATCGGTATAGAATATAATCGGGGATGCCTTTAT  
CTTGAAAAAATGCACCCGCAGCTTCGCTAGTAATCAGTAAACGCGGGAAGTGGAGTCAG  
20 GCTTTTTTTTATGGAAGAGAAAAATAGACACCAAGTAGCCTTCTTCTAACCTTAACGGAC  
CTACAGTGCAAAAAGTTATCAAGAGACTGCATTATAGAGCGCACAAAGGAGAAAAAAG  
TAATCTAAGATGCTTTGTTAGAAAAATAGCGCTCTCGGGATGCATTTTTGTAGAACAAA  
AAAGAAGTATAGATTCTTTGTTGGTAAAAATAGCGCTCTCGCGTTGCATTTCTGTTCTGT  
AAAAATGCAGCTCAGATTCTTTGTTTGA AAAAATTAGCGCTCTCGCGTTGCATTTTTGTT  
25 TTACAAAAATGAAGCACAGATTCTTCGTTGGTAAAAATAGCGCTTTTCGCGTTGCATTTCT  
GTTCTGTAAAAATGCAGCTCAGATTCTTTGTTTGA AAAAATTAGCGCTCTCGCGTTGCAT  
TTTTGTTCTACAAAAATGAAGCACAGATGCTTCGTT

**[00522] SEQ.ID NO: 92**

30 GCTAACAGCGCGATTTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCCAGTCGCGT  
ACCGTCTTCATGGGAGAAAAATAACTGTTGATGGGTGTCTGGTCAGAGACATCAAGAA  
ATAACGCCGGAACATTAGTGCAAGGAGCTTCCACAGCAATGGCATCCTGGTCATCCAGC  
GGATAGTTAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTT  
ACAGGCTTCGACGCGCTTCGTTCTACCATCGACACCACCGCTGGCAGCCAGTTGAT  
35 CGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCGTGCAAGGCCAGACTGGAG  
GTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGTTGGG  
AATGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTTCGCAGAAACGT  
GGCTGGCCTGGTTACCCACGCGGGAAACGGTCTGATAAGAGACACCGGCATACTCTGCG  
ACATCGTATAACGTTACTGGTTTCACATTCACCACCCTGAATTGACTCTCTTCCGGGCG  
40 CTATCATGCCATACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCCGGGATCTCGACGC  
TCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCAGTAGTAGGTTGAGGCCGTTGAG  
CACCGCCGCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCA  
CGGGGCTTGCCACCATAACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCC  
CGATCTTCCCATCGGTGATGTGCGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCC  
45 GGTGATGCCGGCCACGATGCGTCCGGCGTAGCCTAGGATCGAGATCGATCTCGATCCCG  
CGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCTCTAGAA  
ATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGGCAGATCTCAATTGGATATCG  
GCCGGCCACGCGATCGCTGACGTCGGTACCCTCGAGTCTGGTAAAGAAACCGCTGCTGC  
GAAATTTGAACGCCAGCACATGGACTCGTCTACTAGTCGCAGCTTAATTAACCTAAACT  
50 GCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAG

GGGTTTTTTGCTAGCGAAAGGAGGAGTCGACACTGCTTCCGGTAGTCAATAAACCGGTA  
AACCAGCAATAGACATAAGCGGCTATTTAACGACCTGCCCTGAACCGACGACCGGGTC  
ATCGTGGCCGGATCTTGCGGGCCCTCGGCTTGAACGAATTGTTAGACATTATTTGCCGA  
CTACCTTGGTGATCTCGCCTTTCACGTAGTGGACAAATTCTTCCAAGTATGACGGC  
5 GAGGCCAAGCGATCTTCTTCTTGTCGAAGATAAGCCTGTCTAGCTTCAAGTATGACGGG  
CTGATACTGGGCCGGCAGGCGCTCCATTGCCAGTCGGCAGCGACATCCTTCGGCGCGA  
TTTTGCGGTTACTGCGCTGTACCAAATGCGGGACAACGTAAGCACTACATTTGCTCA  
TCGCCAGCCCAGTCGGGCGGCGAGTTCCATAGCGTTAAGGTTTCATTTAGCGCCTCAAA  
TAGATCCTGTTTCAGGAACCGGATCAAAGAGTTCTCCGCCGCTGGACCTACCAAGGCAA  
10 CGCTATGTTCTCTTGCTTTTGTGAGCAAGATAGCCAGATCAATGTCGATCGTGGCTGGC  
TCGAAGATACCTGCAAGAATGTCATTGCGCTGCCATTCTCCAAATTGCAGTTCGCGCTT  
AGCTGGATAACGCCACGGAATGATGTGCTGCTGCACAACAATGGTGACTTCTACAGCGC  
GGAGAATCTCGCTCTCTCCAGGGGAAGCCGAAGTTTCCAAAAGGTCGTTGATCAAAGCT  
CGCCGCGTTGTTTCATCAAGCCTTACGGTCACCGTAACCAGCAAATCAATATCACTGTG  
15 TGGCTTCAGGCCGCCATCCACTGCGGAGCCGTACAAATGTACGGCCAGCAACGTCGGTT  
CGAGATGGCGCTCGATGACGCCAACTACCTCTGATAGTTGAGTCGATACTTCGGCGATC  
ACCGCTTCCCTCATACTCTTCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTG  
TCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGCCAGCTCAC  
TCGGTCGCTACGCTCCGGGCGTGAGACTGCGGCGGGCGCTGCGGACACATACAAAGTTA  
20 CCCACAGATTCCGTGGATAAGCAGGGGACTAACATGTGAGGCAAAACAGCAGGGCCGCG  
CCGGTGGCGTTTTTCCATAGGCTCCGCCCTCTGCCAGAGTTCACATAAACAGACGCTT  
TTCCGGTGCATCTGTGGGAGCCGTGAGGCTCAACCATGAATCTGACAGTACGGGCGAAA  
CCCGACAGGACTTAAAGATCCCCACCGTTTCCGGCGGGTCGCTCCCTCTTGCGCTCTCC  
TGTTCCGACCCTGCCGTTTACCGGATACCTGTTCCGCCTTTCTCCCTTACGGGAAGTGT  
25 GGCGCTTTCTCATAGCTCACACACTGGTATCTCGGCTCGGTGTAGGTGCTTCGCTCCAA  
GCTGGGCTGTAAGCAAGAACTCCCCGTTTCAGCCCGACTGCTGCGCCTTATCCGGTAACT  
GTTCACTTGAGTCCAACCCGGAAAAGCACGGTAAACGCCACTGGCAGCAGCCATTGGT  
AACTGGGAGTTTCGAGAGGATTGTTTAGCTAAACACGCGGTTGCTCTTGAAGTGTGCG  
CCAAAGTCCGGCTACACTGGAAGGACAGATTTGGTTGCTGTGCTCTGCGAAAGCCAGTT  
30 ACCACGGTTAAGCAGTTCCCCAACTGACTTAACCTTCGATCAAACCACCTCCCCAGGTG  
GTTTTTTCGTTTACAGGGCAAAAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCC  
TTTGATCTTTTCTACTGAACCGCTCTAGATTTTCAGTGCAATTTATCTCTTCAAATGTAG  
CACCTGAAGTCAGCCCCATACGATATAAGTTGTAATTCTCATGTTAGTCATGCCCCGCG  
CCCACCGGAAGGAGCTGACTGGGTGTAAGGCTCTCAAGGGCATCGGTTCGAGATCCCGGT  
35 GCCTAATGAGTGAGCTAACTTACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTC  
GGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTT  
TGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTACCAGTGAGACGGGCAACAGCTGATT  
GCCCTTACCAGCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGCCCCA  
GCAGGCGAAAATCCTGTTTGATGGTGGTTAACGGCGGGATATAACATGAGCTGTCTTCG  
40 GTATCGTCGTATCCCACTACCGAGATGTCCGCACCAACGCGCAGCCGGACTCGGTAAAT  
GGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATCGCAGTGGGAACGA  
TGCCCTCATTGAGCATTTGTCATGGTTTGTGAAAACCGGACATGGCACTCCAGTCGCCT  
TCCCCTTCGCTATCGGCTGAATTTGATTGCGAGTGAGATATTTATGCCAGCCAGCCAG  
ACGCAGACGCGCCGAGACAGAACTTAATGGGCCC

45

**[00523]** SEQ.ID NO: 93

GGGGAATTGTGAGCGGATAACAATTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAA  
GGAGATATACATATGGCAGATCTCAATTGGATATCGGCCGGCCACGCGATCGCTGACGT  
CGGTACCCTCGAGTCTGGTAAAGAAACCGCTGCTGCGAAATTTGAACGCCAGCACATGG  
50 ACTCGTCTACTAGTCGAGCTTAATTAACCTAAACTGCTGCCACCGCTGAGCAATAACT

AGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTAGCGAAAGGAG  
GAGTCGACTATATCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCG  
GCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGC  
TCCTTTTCGCTTTCTTCCCTTCTTTCTCGCCACGTTTCGCCGGCTTTCCCCGTCAAGCTC  
5 TAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCAAA  
AAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCG  
CCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAACTGGAACAA  
CACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTTCGGCC  
TATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAAATATT  
10 AACGTTTACAATTTCTGGCGGCACGATGGCATGAGATTATCAAAAAGGATCTTCACCTA  
GATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAACTT  
GGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTT  
CGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTT  
ACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATT  
15 TATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTA  
TCCGCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCCGCCAGT  
TAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCTGT  
TTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCC  
ATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCTCCGATCGTTGTGAGAAGTAAGTT  
20 GGCCGCGAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCTATGC  
CATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAG  
TGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACA  
TAGCAGAACTTTAAAAGTGCTCATCATTGGAACGTTCTTCGGGGCGAAAACCTCTCAA  
GGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCT  
25 TCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGC  
CGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTTC  
AATCATGATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATG  
TATTTAGAAAAATAAACAAATAGGTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTC  
CACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCT  
30 GCGCGTAATCTGCTGCTTGCAAACAAAAAACACCGCTACCGAGCGGTGGTTTGTTCG  
CGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATA  
CCAAATACTGTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGC  
ACCGCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATA  
AGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCTG  
35 GGCTGAACGGGGGGTTCGTGCACACAGCCAGCTTGGAGCGAACGACCTACACCGAACT  
GAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGG  
ACAGGTATCCGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGG  
GGAAACGCCTGGTATCTTTATAGTCCTGTGCGGTTTCGCCACCTCTGACTTGAGCGTCG  
ATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCT  
40 TTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCC  
CCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCGAG  
CCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGT  
ATTTTCTCCTTACGCATCTGTGCGGTATTTACACCGCATATATGGTGCACCTCTCAGTA  
CAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATACTCCGCTATCGCTACGTGACT  
45 GGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGT  
CTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCA  
GAGGTTTTACCGTCATCACCGAAACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGT  
GGTCGTGAAGCGATTACAGATGTCTGCCTGTTTCATCCGCGTCCAGCTCGTTGAGTTTC  
TCCAGAAGCGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTTC  
50 CTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGATTCTGTTTCATGGGGGTAATGATAC  
CGATGAAACGAGAGAGGATGCTCACGATACGGGTTACTGATGATGAACATGCCCGGTTA

CTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGATGCGGCGGGACCAGAGAAAAAT  
CACTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCC  
AGCAGCATCCTGCGATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTT  
TCCAGACTTTACGAAACACGGAAACCGAAGACCATTATGTTGTTGCTCAGGTCGCAGA  
5 CGTTTTGTCAGCAGCAGTCGCTTCACGTTTCGCTCGCGTATCGGTGATTCACTCTGCTAAC  
CAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCTCAACGACAGGAGCACGATCATGCTA  
GTCATGCCCCGCGCCACCGGAAGGAGCTGACTGGGTGAAGGCTCTCAAGGGCATCGG  
TCGAGATCCCGGTGCCTAATGAGTGAGCTAACTTACATTAATTGCGTTGCGCTCACTGC  
CCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCG  
10 GGGAGAGGCGGTTTTCGCTATTGGGCGCCAGGGTGGTTTTTTCTTTTACCAGTGAGACGG  
GCAACAGCTGATTGCCCTTCACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACG  
CTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGTTAACGGCGGGATATAACA  
TGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGATGTCCGCACCAACGCGCAGCC  
CGGACTCGGTAATGGCGCGCATTCGCGCCAGCGCCATCTGATCGTTGGCAACCAGCATC  
15 GCAGTGGGAACGATGCCCTCATTCAGCATTTGCATGGTTTGTGAAAACCGGACATGGC  
ACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGATTGCGAGTGAGATATTTAT  
GCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGCCCCGCTAACAGCGCG  
ATTTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCCAGTCGCGTACCGTCTTCATG  
GGAGAAAATAATACTGTTGATGGGTGTCTGGTCAGAGACATCAAGAAATAACGCCGGAA  
20 CATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCGGATAGTTAATG  
ATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGAC  
GCCGCTTCGTTCTACCATCGACACCACCGCTGGCACCCAGTTGATCGGCGCGAGATT  
TAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAGACTGGAGGTGGCAACGCCA  
ATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAATGTAATTCAG  
25 CTCCGCCATCGCCGCTTCCACTTTTTTCCCGCGTTTTTCGCAGAAACGTGGCTGGCCTGGT  
TCACCACGCGGGAAACGGTCTGATAAGAGACACCGGCATACTCTGCGACATCGTATAAC  
GTTACTGGTTTCACATTACCAACCCTGAATTGACTCTCTTCCGGGCGCTATCATGCCAT  
ACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCCGGGATCTCGACGCTCTCCCTTATGC  
GACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCGCCGC  
30 AAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCCTGCCA  
CCATACCCACGCGGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCCA  
TCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGC  
CACGATGCGTCCGGCGTAGCCTAGGATCGAGATCGATCTCGATCCCGCGAAATTAATAC  
GACTCACTATA

35

**[00524] SEQ.ID NO: 94**

ATCCGGATATAGTTCCCTCCTTTCAGCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCA  
AGGGGTTATGCTAGTTATTGCTCAGCGGTGGCAGCAGCCAACTCAGCTTCCTTTCGGGC  
TTTGTTAGCAGCCGGATCTCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT  
40 TGTCGACGGAGCTCGAATTCGGATCCGCGACCCATTGCTGTCCACCAGTCATGCTAGC  
CATATGGCTGCCGCGCGGCACCAGGCCGCTGCTGTGATGATGATGATGATGGCTGCTGC  
CCATGGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGGGAATTGTTATC  
CGCTCACAATTCCTTATAGTGAGTCGTATTAATTTTCGCGGGATCGAGATCTCGATCCT  
CTACGCCGGACGCATCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCT  
45 ATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCT  
TGTTTTCGGCGTGGGTATGGTGGCAGGCCCCGTGGCCGGGGGACTGTTGGGCGCCATCTC  
CTTGATGCACCATTCCTTGCGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGGGCT  
GCTTCCTAATGCAGGAGTCGCATAAGGGAGAGCGTCGAGATCCCGGACACCATCGAATG  
GCGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGGTGG  
50 TGAATGTGAAACCAGTAACGTTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAG

ACCGTTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGT  
GGAAGCGGCGATGGCGGAGCTGAATTACATTCCCAACCGCGTGGCACAACAACCTGGCGG  
GCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCCGTCG  
CAAATTGTGCGGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTC  
5 GATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGC  
AACGCGTCAGTGGGCTGATCATTAATATCCGCTGGATGACCAGGATGCCATTGCTGTG  
GAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGACACCCAT  
CAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATCTGGTCG  
CATTGGGTCACCAGCAAATCGCGCTGTTAGCGGGCCATTAAAGTTCTGTCTCGGCGCGT  
10 CTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGA  
ACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAATGCTGAATG  
AGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAATG  
CGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCGGATATCTCGGTAGTGGGATACGA  
CGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAACCACCATCAAACAGGATTTTC  
15 GCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTG  
AAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAAGAAAAACCACCTGGCGCCCCAA  
TACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCAATTAATGCAGCTGGCACGACAGG  
TTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTAAGTTAGCTCACTCA  
TTAGGCACCGGGATCTCGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAGCTCCTTCC  
20 GGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGACTGTCTTCTTTATCATGCAA  
CTCGTAGGACAGGTGCCGGCAGCGCTCTGGGTCAATTTTCGGCGAGGACCGCTTTCGCTG  
GAGCGCGACGATGATCGGCCTGTCGCTTGCGGTATTGGAATCTTGACGCCCCTCGCTC  
AAGCCTTCGTCACTGGTCCCGCCACCAAACGTTTTCGGCGAGAAGCAGGCCATTATCGCC  
GGCATGGCGGCCCCACGGGTGCGCATGATCGTGCTCCTGTGCTTGAGGACCCGGCTAGG  
25 CTGGCGGGGTTGCCTTACTGGTTAGCAGAATGAATCACCGATACGCGAGCGAACGTGAA  
GCGACTGCTGCTGCAAAACGTCTGCGACCTGAGCAACAACATGAATGGTCTTCGGTTTC  
CGTGTTTCGTAAAGTCTGGAAACGCGGAAGTCAGCGCCCTGCACCATTATGTTCCGGAT  
CTGCATCGCAGGATGCTGCTGGCTACCCTGTGGAACACCTACATCTGTATTAACGAAGC  
GCTGGCATTGACCCTGAGTGATTTTTCTCTGGTCCCGCCGCATCCATACCGCCAGTTGT  
30 TTACCCTCACAACGTTCCAGTAACCGGGCATGTTCAATCATCAGTAACCCGTATCGTGAG  
CATCCTCTCTCGTTTTATCGGTATCATTACCCCCATGAACAGAAATCCCCCTTACACGG  
AGGCATCAGTGACCAAACAGGAAAAAACCGCCCTTAACATGGCCCCGCTTTATCAGAAGC  
CAGACATTAACGCTTCTGGAGAAACTCAACGAGCTGGACGCGGATGAACAGGCAGACAT  
CTGTGAATCGCTTACGACCACGCTGATGAGCTTTACCGCAGCTGCCTCGCGCGTTTTCG  
35 GTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTG  
TAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTG  
TCGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGATACTGGCTTAACTA  
TGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATATGCGGTGTGAAATACCGCA  
CAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTCTCGCTCACTGACT  
40 CGCTGCGCTCGGTGCTTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATA  
CGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCA  
AAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCC  
CTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTA  
TAAAGATACCAGGCGTTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCT  
45 GCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATA  
GCTCACGCTGTAGGTATCTCAGTTCGGTGATGGTTCGCTCCAAGCTGGGCTGTGTG  
CACGAACCCCCCGTTACGCCCAGCGCTGCGCCTTATCCGGTAACATCGTCTTGAGTC  
CAACCCGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCA  
GAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTAC  
50 ACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAG  
AGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTT

GCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCT  
ACGGGGTCTGACGCTCAGTGGAACGAAACTCACGTTAAGGGATTTTGGTCATGAACAA  
TAAACTGTCTGCTTACATAAACAGTAATACAAGGGTGTTATGAGCCATATTCAACGG  
GAAACGTCTTGCTCTAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTA  
5 TAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGATTGTATGGGA  
AGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTT  
ACAGATGAGATGGTCAGACTAACTGGCTGACGGAATTTATGCCTCTTCCGACCATCAA  
GCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCCTGCGATCCCCGGGAAAA  
CAGCATTCCAGGTATTAGAAGAATATCCTGATTGAGGTGAAAAATATTGTTGATGCGCTG  
10 GCAGTGTTCTGCGCCGGTTGCATTCGATTCCTGTTTGTAAATTGTCCTTTTAACAGCGA  
TCGCGTATTTCTGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGA  
GTGATTTTGTATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCAT  
AACTTTTGGCATTCTCACCAGGATTCAGTCGTCACTCATGGTGATTTCTCACTTGATAA  
CCTTATTTTGGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCG  
15 CAGACCGATACCGAGGATCTTGCCATCCTATGGAAGTGCCTCGGTGAGTTTTCTCCTTCA  
TTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCA  
GTTTCATTTGATGCTCGATGAGTTTTTCTAAGAATTAATTCATGAGCGGATACATATTT  
GAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCC  
ACCTGAAATTGTAAACGTTAATATTTTGTAAATTCGCGTTAAATTTTTGTAAATCA  
20 GCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAGAATAG  
ACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGAACAAGAGTCCACTATTAAAGAACGT  
GGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAAC  
CATCACCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCTT  
AAAGGGAGCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGA  
25 AGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTACGCTGC  
GCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCA

**[00525] SEQ.ID NO: 95**

ATGGCTTCTAGTTCTTCCGATGTCTTCGTTTTGGGTCTAGGTGTTGTTTTGGCTGCCTT  
30 GTATATCTTCAGAGACCAATTATTCGCTGCTTCTAAGCCAAAGGTGGCTCCAGTTTCCA  
CTACGAAGCCTGCCAACGGTTCCGCTAACCCAAGAGACTTCATCGCCAAGATGAAACAA  
GGTAAGAAGAGAATCGTAATCTTCTACGGTTCTCAAACCTGGTACCGCTGAAGAATATGC  
TATTCGTTTGGCTAAGGAAGCTAAGCAAAAGTTTCGGTCTAGCCTCCTTGGTTTGTGATC  
CAGAAGAATACGATTTTGAAAAGTTGGACCAATTGCCAGAAGATTCTATTGCTTTCTTC  
35 GTCGTTGCTACCTATGGTGAAGGTGAACCTACAGACAACGCTGTCCAATTGTTGCAAAA  
CTTGCAAGATGAAAGCTTCGAATTCTCCTCTGGTGAGAGAAAGTTGTCAGGTTTGAAGT  
ACGTTGTTTTTGGTCTGGGTAACAAGACCTACGAACATTACAACCTCATTGGGAGAACT  
GTTGACGCTCAATTGGCCAAGATGGGTGCTATCAGAATCGGTGAAAGAGGTGAAGGTGA  
TGATGACAAGTCCATGGAAGAAGACTACTTGAATGGAAGGATGGTATGTGGGAAGCGT  
40 TTGCCACTGCTATGGGTGTTGAAGAAGGTCAAGGTGGTGAAGTCCGCTGATTTCTGCTGCTT  
TCCGAATTGGAATCTCACCACCAAGAAAAGGTTTACCAAGGTGAATTTTCTGCTAGAGC  
TTTAACCAAAACCAAGGTATTACGACGCTAAGAATCCTTTTGGTCTGCTCCAATTGCGG  
TTGCTAGAGAATTGTTCCAATCTGTTGTGATAGAACTGTGTCCACGTCGAATTCAAC  
ATTGAAGGCTCTGGTATCACCTATCAACACGGTGACCACGTTGGTTTGTGGCCATTGAA  
45 TCCAGATGTTGAAGTCGAACGGTTGTTGTGTGTTTTAGGTTTAGCTGAAAAGAGAGATG  
CTGTCATCTCCATTGAATCCTTAGACCCGGCTTTGGCTAAGGTTCCATTCCCAGTCCCA  
ACTACTTACGGTGCTGTGTTGAGACACTACATTGACATCTCTGCTGTGCGCGGTAGACA  
AATCTTGGGTACTTTGTCCAATTCGCTCCAACCCAGAAAGCTGAAGCTTTCTTGAGAA  
ACTTGAACACTAACAAGGAAGAATACCACAACGTCGTCGCTAACGGTTGTTTGAATTTG  
50 GGTGAAATTTTGCAAATCGCTACCGGTAACGACATTACTGTCCCACTACTGCCAA



CACCACCAAATGGCCAATTCCATTGACATCATTGTTTCTGCCATCCCAAGATTGCAAC  
 CAAGATACTACTCTATCTCTTCTTCCCCAAAAATTCATCCAAACACCATCCACGCTACC  
 GTTGTGTGCTCAAATACGAAAACGTTCCAACCGAACCAATCCCAAGAAAGTGGGTTTA  
 CGGTGTCGGTAGTAACCTTCTTGTGAATTTAAAGTACGCTGTTAACAAGGAACAGTTC  
 5 CATACTACTCAAAATGGCGAACAAAGAGTCGGTGTCCCGGAATACTTGATTGCTGGT  
 CCACGTGGTTCTTACAAGACTGAATCTTTCTACAAGGCTCCAATCCATGTTAGACGTTT  
 TACTTTCCGTTTGCCAACCAACCCAAAGTCTCCAGTCATCATGATTGGTCCAGGTACTG  
 GTGTCGCCCCATTGAGAGGCTTCGTTCAAGAAAGAGTTGCCTTGGCCAGAAGATCCATC  
 10 GAAAAGAACGGTCCTGACTCTTTGGCTGACTGGGGTCGTATTTCTTGTCTACGGTTG  
 TAGAAGATCCGACGAAGACTTCTTGTACAAGGACGAATGGCCACAATACGAAGCTGAGT  
 TGAAGGGTAAGTTCAAGTTGCACTGTGCTTTCTCCAGACAAAACCTACAAGCCAGACGGT  
 TCTAAGATTTACGTCCAAGATTTGATCTGGGAAGACAGAGAACACATTGCCGATGCCAT  
 CTTAAACGGTAAGGGTTACGTCTACATCTGCGGTGAAGCTAAGTCCATGTCTAAACAAG  
 TTGAAGAAGTTCTAGCCAAGATCTTGGGCGAAGCCAAAGGTGGTTCGGTCCAGTTGAA  
 15 GGTGTTGCTGAAGTCAAGTTACTGAAGGAACGGTCCAGATTGATGTTGGATGTCTGGTC  
 TTGA

**[00526] SEQ.ID NO: 96**

MASSSSDVFLGLGVVLAALYIFRDQLFAASKPKVAPVSTTKPANGSANPRDFIAKMKQ  
 20 GKKRIVIFYGSQTGTAEYAIRLAKEAKQKFGLASLVCDPPEYDFEKLQLPEDSIAFF  
 VVATYGEGETDNAVQLLQNLQDESFEFSSGERKLSGLKYVVFGLGNKTYEHYNLIGRT  
 VDAQLAKMGAIRIGERGEEDDKSMEEDYLEWKDGMWEAFATAMGVEEGQGGDSADFFV  
 SELESHPEKVYQGEFSARALTKTKGIHDAKNPFAAPIAVARELFQSVVDRNCVHVEFN  
 IEGSGITYQHGDHVLWPLNPDVEVERLLCVLGLAEKRDAVISIESLDPALAKVPFPVP  
 25 TTYGAVLRHYIDISAVAGRQILGTLSKFAPTPEAEAFRLNLNTNKEEYHNVVANGCLKL  
 GEILQIATGNDITVPPTTANTTKWPIPFDIIVSAIPRLQPRYSSISSPKIHPNTIHAT  
 VVVLKYENVPTPIPRKWVYGVGSNFLNLKYAVNKEPVPIYITQNGEQRVGVPEYLIAG  
 PRGSYKTESFYKAPIHVRSTFRLPTNPKSPVIMIGPGTGVAFFRGFVQERVALARRSI  
 EKNPDSLADWGRISLFYGCRRSDEDFLYKDEWPQYEAELKGKFKLHCAFSRQNYKPDG  
 30 SKIYVQDLIWEDREHIADAILNGKYVYICGEAKSMSKQVEEVLAAILGEAKGSGPVE  
 GVAEVKLLKERSRLMLDVWS

**[00527] SEQ.ID NO: 97**

GAAGAATTGTGAGCGGATAACAATCCCCCTCTAGAAATAATTTTGTTTAACTTTAAGAA  
 35 GGAGATATACATATGGCAGATCTCAATTGGATATCGGCCGGCCACGCGATCGCTGACGT  
 CGGTACCCCTCGAGTCTGGTAAAGAAACCGCTGCTGCGAAATTTGAACGCCAGCACATGG  
 ACTCGTCTACTAGTCGCAGCTTAATTAACCTAAACTGCTGCCACCGCTGAGCAATAACT  
 AGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTTGCTAGCGAAAGGAG  
 GAGTCGACTATATCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCG  
 40 GCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGC  
 TCCTTTTCGCTTTCTTCCCTTCTTTCTCGCCACGTTGCGCCGGCTTTCCCGTCAAGCTC  
 TAAATCGGGGGCTCCCTTTAGGGTTCGATTTAGTGCTTTACGGCACCTCGACCCAAA  
 AACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCG  
 CCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAA  
 45 CACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTTCGGCC  
 TATTGGTTAAAAAATGAGCTGATTTAAACAAAATTTAACGCGAATTTTAAACAAAATATT  
 AACGTTTACAATTTCTGGCGGCACGATGGCATGAGATTATCAAAAAGGATCTTCACCTA  
 GATCCTTTTAAATTAATAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAACTT  
 GGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTT  
 50 CGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTT

ACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATT  
TATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTA  
TCCGCCTCCATCCAGTCTATTAATTGTTGCCGGAAGCTAGAGTAAGTAGTTCGCCAGT  
TAATAGTTTGCACAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCGT  
5 TTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCC  
ATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTGAGAAGTAAGTT  
GGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGC  
CATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAG  
TGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACA  
10 TAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACCTCTCAA  
GGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCT  
TCAGCATCTTTTACTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGC  
CGCAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTC  
AATCATGATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATG  
15 TATTTAGAAAAATAAACAAATAGGTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTC  
CACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTTCT  
GCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTCG  
CGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACCTGGCTTCAGCAGAGCGCAGATA  
CCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGC  
20 ACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATA  
AGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTTCG  
GGCTGAACGGGGGGTTCGTGCACACAGCCAGCTTGGAGCGAACGACCTACACCGAACT  
GAGATACCTACAGCGTGAGCTATGAGAAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGG  
ACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGG  
25 GGAAACGCCTGGTATCTTTATAGTCCTGTGCGGGTTTCGCCACCTCTGACTTGAGCGTCG  
ATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGAAAAACGCCAGCAACGCGGCCT  
TTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCTGCGTTATCC  
CCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAG  
CCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGT  
30 ATTTTCTCCTTACGCATCTGTGCGGTATTTACACCGCATATATGGTGCACTCTCAGTA  
CAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATACACTCCGCTATCGCTACGTGACT  
GGGTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGT  
CTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCA  
GAGGTTTTACCGTCATACCGAAACGCGCGAGGCAGCTGCGGTAAAGCTCATCAGCGT  
35 GGTGCTGAAGCGATTACAGATGTCTGCCTGTTTCATCCGCGTCCAGCTCGTTGAGTTTC  
TCCAGAAGCGTTAATGTCTGGCTTCTGATAAAGCGGGCCATGTTAAGGGCGGTTTTTTC  
CTGTTTGGTCACTGATGCCTCCGTGTAAGGGGGATTCTGTTTCATGGGGGTAATGATAC  
CGATGAAACGAGAGAGGATGCTCACGATACGGGTTACTGATGATGAACATGCCCGGTTA  
CTGGAACGTTGTGAGGGTAAACAACTGGCGGTATGGATGCGGCGGGACCAGAGAAAAAT  
40 CACTCAGGGTCAATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCC  
AGCAGCATCCTGCGATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGCGTT  
TCCAGACTTTACGAAACACGGAAACCGAAGACCATTATGTTGTTGCTCAGGTCGCAGA  
CGTTTTTGACGACGAGTCGCTTACGTTTCGCTCGCGTATCGGTGATTATCTGCTAAC  
CAGTAAGGCAACCCCGCCAGCCTAGCCGGTCTCAACGACAGGAGCACGATCATGCTA  
45 GTCATGCCCCGCGCCACCGGAAGGAGCTGACTGGGTGAAGGCTCTCAAGGGCATCGG  
TCGAGATCCCGGTGCCTAATGAGTGAGCTAACTTACATTAATTGCGTTGCGCTCACTGC  
CCGCTTTCCAGTCGGGAAACCTGTGCTGCCAGCTGCATTAATGAATCGGCCAACGCGCG  
GGGAGAGGCGGTTTGGCTATTGGGCGCCAGGGTGGTTTTTTCTTTTACCAGTGAGACGG  
GCAACAGCTGATTGCCCTTACCGCCTGGCCCTGAGAGAGTTGCAGCAAGCGGTCCACG  
50 CTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGTTAACGGCGGGATATAACA  
TGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGATGTCCGCACCAACGCGCAGCC

CGGACTCGGTAATGGCGCGCATTGCGCCCAGCGCCATCTGATCGTTGGCAACCAGCATC  
GCAGTGGGAACGATGCCCTCATTGAGCATTTGCATGGTTTGTGAAAACCGGACATGGC  
ACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGATTGCGAGTGAGATATTTAT  
GCCAGCCAGCCAGACGCGAGACGCGCCGAGACAGAACTTAATGGGCCCCGCTAACAGCGCG  
5 ATTTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCCAGTCGCGTACCGTCTTCATG  
GGAGAAAATAATACTGTTGATGGGTGTCTGGTCAGAGACATCAAGAAATAACGCCGGAA  
CATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAGCGGATAGTTAATG  
ATCAGCCCCTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTTTACAGGCTTCGAC  
GCCGCTTCGTTCTACCATCGACACCACCAGCTGGCACCCAGTTGATCGGCGCGAGATT  
10 TAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAGACTGGAGGTGGCAACGCCA  
ATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGGGAATGTAATTCAG  
CTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTTCGCAGAAACGTGGCTGGCCTGGT  
TCACCACGCGGGAAACGGTCTGATAAGAGACACCGGCATACTCTGCGACATCGTATAAC  
GTTACTGGTTTCACATTACCAACCCTGAATTGACTCTCTTCCGGGCGCTATCATGCCAT  
15 ACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCCGGGATCTCGACGCTCTCCCTTATGC  
GACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCGCCGC  
AAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCCTGCCA  
CCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCCA  
TCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGC  
20 CACGATGCGTCCGGCGTAGCCTAGGATCGAGATCGATCTCGATCCCGCGAAATTAATAC  
GACTCACTACG

**[00528] SEQ.ID NO: 98**

CGCTAACAGCGCGATTTGCTGGTGACCCAATGCGACCAGATGCTCCACGCCCAGTCGCG  
25 TACCGTCTTCATGGGAGAAAATAATACTGTTGATGGGTGTCTGGTCAGAGACATCAAGA  
AATAACGCCGGAACATTAGTGCAGGCAGCTTCCACAGCAATGGCATCCTGGTCATCCAG  
CGGATAGTTAATGATCAGCCCACTGACGCGTTGCGCGAGAAGATTGTGCACCGCCGCTT  
TACAGGCTTCGACGCCGCTTCGTTCTACCATCGACACCACCACGCTGGCACCCAGTTGA  
TCGGCGCGAGATTTAATCGCCGCGACAATTTGCGACGGCGCGTGCAGGGCCAGACTGGA  
30 GGTGGCAACGCCAATCAGCAACGACTGTTTGCCCGCCAGTTGTTGTGCCACGCGGTTGG  
GAATGTAATTCAGCTCCGCCATCGCCGCTTCCACTTTTTCCCGCGTTTTTCGCAGAAACG  
TGGCTGGCCTGGTTCCACCACGCGGGAACGGTCTGATAAGAGACACCGGCATACTCTGC  
GACATCGTATAACGTTACTGGTTTACATTACCAACCCTGAATTGACTCTCTTCCGGGC  
GCTATCATGCCATACCGCGAAAGGTTTTGCGCCATTCGATGGTGTCCGGGATCTCGACG  
35 CTCTCCCTTATGCGACTCCTGCATTAGGAAGCAGCCAGTAGTAGGTTGAGGCCGTTGA  
GCACCGCCGCCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGCC  
ACGGGGCCTGCCACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGC  
CCGATCTTCCCCATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGC  
CGGTGATGCCGGCCACGATGCGTCCGGCGTAGCCTAGGATCGAGATCGATCTCGATCCC  
40 GCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTTCCCTCTAGA  
AATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGGCAGATCTCAATTGGATATC  
GGCCGGCCACGCGATCGCTGACGTCGGTACCCTCGAGTCTGGTAAAGAAACCGCTGCTG  
CGAAATTTGAACGCCAGCACATGGACTCGTCTACTAGTCGCAGCTTAATTAACCTAAAC  
TGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGA  
45 GGGGTTTTTTGCTAGCGAAAGGAGGAGTCGACAAGCTGACGACCGGGTCTCCGCAAGTG  
GCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCA  
AATATGTATCCGCTCATGAATTAATTCTTAGAAAACTCATCGAGCATCAAATGAAACT  
GCAATTTATTATATCAGGATTATCAATACCATATTTTTTGAAGGCGGTTTCTGTAAT  
GAAGGAGAAACTCACCGAGGCAGTTCCATAGGATGGCAAGATCCTGGTATCGGTCTGC  
50 GATTCCGACTCGTCCAACATCAATACAACCTATTAATTTCCCTCGTCAAAAATAAGGT

TATCAAGTGAGAAATCACCATGAGTGACGACTGAATCCGGTGAGAATGGCAAAAGTTTA  
 TGCATTTCTTTCCAGACTTGTTCAACAGGCCAGCCATTACGCTCGTCATCAAAATCACT  
 CGCATCAACCAACCGTTATTTCATTCGTGATTGCGCCTGAGCGAGACGAAATACGCGGT  
 CGCTGTTAAAAGGACAATTACAAACAGGAATCGAATGCAACCGGCGCAGGAACACTGCC  
 5 AGCGCATCAACAATATTTTCACCTGAATCAGGATATTCTTCTAATACCTGGAATGCTGT  
 TTTCCCGGGGATCGCAGTGGTGAGTAACCATGCATCATCAGGAGTACGGATAAAATGCT  
 TGATGGTCGGAAGAGGCATAAAATCCGTCAGCCAGTTTAGTCTGACCATCTCATCTGTA  
 ACATCATTTGGCAACGCTACCTTTGCCATGTTTCAGAAACAACCTCTGGCGCATCGGGCTT  
 CCCATACAATCGATAGATTGTCGCACCTGATTGCCCCGACATTATCGCGAGCCCATTTAT  
 10 ACCCATATAAATCAGCATCCATGTTGGAATTTAATCGCGGCCTAGAGCAAGACGTTTCC  
 CGTTGAATATGGCTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTA  
 TTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGCATGC  
 AGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTACGCTCGGTCTGTTGACTGCGGCGA  
 GCGGTGTGAGCTCACTCAAAAGCGGTAATACGGTTATCCACAGAATCAGGGGATAAAGC  
 15 CGGAAAGAACATGTGAGCAAAAAGCAAAGCACCAGGAAGAAGCCAACGCCGAGGCGTTT  
 TTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGCCAGAGGTG  
 GCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGC  
 GCTCTCCTGTTCCGACCCTGCCGTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGA  
 AGCGTGGCGCTTTCTCATAGCTCACGCTGTTGGTATCTCAGTTCGGTGTAGGTCGTTTCG  
 20 CTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTACGCCCCGACCGCTGCGCCTTATCCG  
 GTAACATATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCC  
 ATTGGTAACTGATTTAGAGGACTTTGTCTTGAAGTTATGCACCTGTTAAGGCTAAACTG  
 AAAGAACAGATTTTGGTGAGTGCGGTCTCCAACCCACTTACCTTGTTCAAAGAGTTG  
 GTAGCTCAGCGAACCTTGAGAAAACCACCGTTGGTAGCGGTGGTTTTTCTTTATTTATG  
 25 AGATGATGAATCAATCGGTCTATCAAGTCAACGAACAGCTATTCCGTTACTCTAGATTT  
 CAGTGCAATTTATCTCTTCAAATGTAGCACCTGAAGTCAGCCCCATACGATATAAGTTG  
 TAATTCTCATGTTAGTCATGCCCCGCGCCACCGGAAGGAGCTGACTGGGTGTAAGGCT  
 CTAAGGGCATCGGTGAGATCCCGGTGCCTAATGAGTGAGCTAACTTACATTAATTGC  
 GTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAA  
 30 TCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTT  
 CACCAGTGAGACGGGCAACAGCTGATTGCCCTTACCGCCTGGCCCTGAGAGAGTTGCA  
 GCAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAATCCTGTTTGATGGTGGTTAAC  
 GGCGGGATATAACATGAGCTGTCTTCGGTATCGTCGTATCCCACTACCGAGATGTCCGC  
 ACCAACGCGCAGCCCGGACTCGGTAATGGCGCGCATTCGCGCCAGCGCCATCTGATCGT  
 35 TGGCAACCAGCATCGCAGTGGGAACGATGCCCTCATTGAGCATTTGCATGGTTTGTGTA  
 AAACCGGACATGGCACTCCAGTCGCCTTCCCGTTCCGCTATCGGCTGAATTTGATTGCG  
 AGTGAGATATTTATGCCAGCCAGCCAGACGCAGACGCGCCGAGACAGAACTTAATGGGC  
 C

40

**[00529]** Hereinafter are provided examples of specific implementations for  
 performing the methods of the present disclosure, as well as implementations  
 representing the compositions of the present disclosure. The examples are  
 provided for illustrative purposes only, and are not intended to limit the scope of  
 45 the present disclosure in any way.

## **EXAMPLES**

### **Example 1 – Process for making a first halogenated psilocybin derivative from halogenated tryptophan feedstock.**

5

**[00530]** Yeast (*Saccharomyces cerevisiae*) was genetically engineered to enable bioconversion of commercially obtained, halogenated, simple indole, tryptophan, or tryptamine feedstock to generate final product. The parent yeast (*Saccharomyces cerevisiae*) strain was CEN.PK with genotype *Mat $\alpha$* ; *ura3-52*; *trp1-289*; *leu2-3,112*; *his3 $\Delta$ 1*; *MAL2-8C*; *SUC2*. The parent strain was engineered to include three enzymatic steps through chromosomal homologous recombination of integration cassettes as described previously (Dastmalchi *et al.*, 2019, Nat. Chem Biol. 15: 384-390; Chen *et al.*, 2018, Nat. Chem Biol. 14: 738-743). The three steps were comprised of PsiH (SEQ.ID NO: 4), PsiK (SEQ.ID NO: 6) and PsiM (SEQ.ID NO: 8), encoded by SEQ.ID NO: 3, SEQ.ID NO: 5 and SEQ.ID NO: 7, respectively, with addition of in-frame, C-terminal HA (SEQ.ID NO: 81, SEQ.ID NO: 82), V5 (SEQ.ID NO: 87, SEQ.ID NO: 88), and FLAG (SEQ.ID NO: 85, SEQ.ID NO: 86) epitope tags, respectively. DNA (SEQ.ID NO: 95) encoding a fifth enzyme, PcCPR (SEQ.ID NO: 96) was also integrated following PCR-based, C-terminal, in-frame addition of epitope tag c-MYC (SEQ.ID NO: 83, SEQ.ID NO: 84) to support functionality of PsiH. Integration cassettes were built using yeast promoter sequences amplified from *S. cerevisiae* genomic DNA as described (Dastmalchi *et al.* 2019; Chen *et al.* 2018) enabling constitutive gene expression. Amplified promoters included *PGK1* (SEQ.ID NO: 77), *TDH3* (SEQ.ID NO: 78), *CLN1* (SEQ.ID NO: 79), and *UGA1* (SEQ.ID NO: 80). Two integration cassettes were assembled: the first (SEQ.ID NO: 74) harbored tagged PsiH and PcCPR, and the second (SEQ.ID NO: 75) harbored tagged PsiK and PsiM. Successive genomic integration of these cassettes was performed as described previously (Chen *et al.* 2018). Following stable integration of these two cassettes, the strain was further manipulated by transformation with a yeast episomal vector encoding a promiscuous decarboxylase, ClostSporTDC (pMM1-pTDH3-ClostSporTDC-His-tCYC1) (SEQ.ID NO: 76) enabling efficient decarboxylation of halogenated tryptophan intermediates and/or feedstocks. For this Example, heterologous expression of a non-native or engineered TrpB gene

was not necessary, as endogenous tryptophan synthase activity proved sufficient. The final engineered strain was called Sc-1. For scaled-up production of halogenated product, culturing was performed as follows. Seed cultures were inoculated in SD-drop-out medium overnight. The overnight culture was then

5 divided into two flasks containing 500 ml each of SD-drop-out medium containing 2% (w/v) glucose, 0.3%(w/v)  $\text{KH}_2\text{PO}_4$ , 0.05% (w/v)  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ , 0.5% (w/v)  $(\text{NH}_4)_2\text{SO}_4$  plus 500  $\mu\text{M}$  6-fluorotryptophan (ThermoFisher Scientific) for conversion by Sc-1. Yeast cultures were grown for 48 h. Cultures were then centrifuged (10,000g x 5 minutes) to remove cellular content, and culture broth

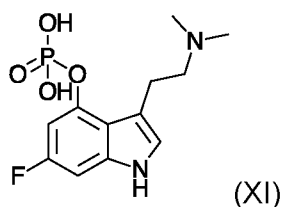
10 containing secreted halogenated product was stored at  $-80^\circ\text{C}$  until further processing. To determine successful biosynthesis of 6-fluoro-psilocybin (compound with formula (XI)), culture broth was analyzed using high-resolution LC-HESI-LTQ-Orbitrap-XL MS (Thermo Fisher Scientific), employing a modified version of a method described previously (Chang *et al.*, 2015, Plant Physiol. 169:

15 1127-1140), with the exception that liquid chromatography was carried out using an UltiMate 3000 HPLC (Thermo Fisher Scientific) equipped with a Poroshell 120 SB-C18 column (Agilent Technologies) instead of an Accela HPLC system (Thermo Fisher Scientific) equipped with a Zorbax C18 column (Agilent Technologies). Briefly, 10 microliters of sample was injected at a flow rate of 0.5

20 mL/min and a gradient of solvent A (water with 0.1 % of formic Acid) and solvent B (ACN with 0.1% formic Acid) as follows: 100% to 0% (v/v) solvent A over 5 min; isocratic at 0% (v/v) for 1 min; 0% to 100% (v/v) over 0.1 min; and isocratic at 100% (v/v) for 1.9 min. Total run time was 8 minutes. Heated ESI source and interface conditions were operated in positive ion mode as follows: vaporizer

25 temperature,  $400^\circ\text{C}$ ; source voltage, 3 kV; sheath gas, 60 au, auxiliary gas, 20 au; capillary temperature,  $380^\circ\text{C}$ ; capillary voltage, 6 V; tube lens, 45 V. Instrumentation was performed as a single, HR scan event using Orbitrap detection of  $m/z$  in the range of 100-500  $m/z$ . Ion injection time was 300 ms with scan time of 1 s. External and internal calibration procedures ensured <2 ppm

30 error to facilitate elemental formulae predictions. Singly protonated product with exact  $m/z$  and expected elemental formula matching 6-fluoro-psilocybin having chemical formula (XI):

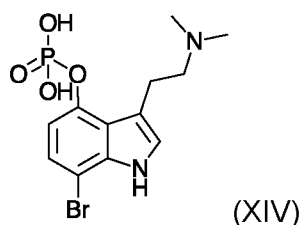


eluted at 1.9 minutes (EIC, see: **FIG. 10**).

**5 Example 2 – Process for making a second halogenated psilocybin derivative from halogenated tryptophan feedstock.**

**[00531]** The same yeast strain (Sc-1) and procedures described in Example 1 were used to biosynthesize a halogenated psilocybin derivative with chemical formula (XIV), with the following exception: in place of 6-fluorotryptophan, 500  $\mu$ M 7-bromotryptophan (Combi-Blocks, [www.combi-blocks.com](http://www.combi-blocks.com)) was supplied as feedstock for bioconversion. Singly protonated product with exact  $m/z$  and expected elemental formula matching 7-bromo-psilocybin having chemical formula (XIV):

15



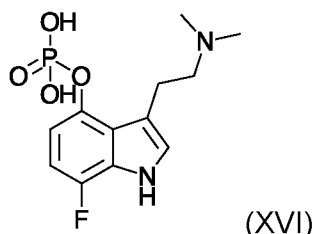
eluted at 2.8 minutes (EIC, see: **FIG. 11**).

**20 Example 3 – Process for making a third halogenated psilocybin derivative from halogenated indole feedstock.**

**[00532]** The same yeast strain (Sc-1) and procedures described in Example 1 were used to biosynthesize a halogenated psilocybin derivative with chemical formula (XVI), with the following exception: in place of 6-fluorotryptophan, 500  $\mu$ M 7-fluoroindole (Combi-Blocks, [www.combi-blocks.com](http://www.combi-blocks.com)) was supplied as feedstock for bioconversion. Singly protonated product with exact  $m/z$  and expected

25

elemental formula matching 7-fluoro-psilocybin compound having chemical formula (XVI):



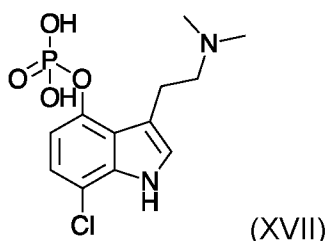
5

eluted at 2.3 minutes (EIC, see: **FIG. 12**).

**Example 4 – Process for making a fourth halogenated psilocybin derivative from halogenated indole feedstock.**

10

**[00533]** The same yeast strain (Sc-1) and procedures described in Example 1 were used to biosynthesize a halogenated psilocybin derivative with formula (XVII), with the following exception: in place of 6-fluorotryptophan, 500  $\mu$ M 7-chloroindole (Combi-Blocks, [www.combi-blocks.com](http://www.combi-blocks.com)) was supplied as feedstock for bioconversion. Singly protonated product with exact  $m/z$  and expected elemental formula matching 7-chloro-psilocybin having formula (XVII):



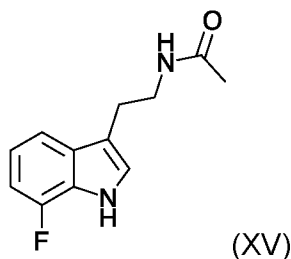
eluted at 2.7 minutes (EIC, see: **FIG. 13**).

**Example 5 – Process for making a fifth halogenated psilocybin derivative from halogenated indole feedstock.**

**[00534]** *E. coli* strain Ec-1 was constructed as follows. For plasmid cloning, Top10 or XL1-blue strains were used depending on antibiotic markers. Standard LB media was used for culturing. For gene expression and feeding experiments, the parent host strain employed was BL21 (DE3). For this Example, heterologous



expression of a non-native or engineered TrpB gene was not necessary, as endogenous tryptophan synthase activity proved sufficient. An in frame, C-terminally His-tagged (SEQ.ID NO: 89, SEQ.ID NO: 90) BaTDC (SEQ.ID NO: 36, SEQ.ID NO: 37) was generated by PCR, followed by cloning into the *NdeI* and *XhoI* sites of the plasmid pCDM4 (SEQ.ID NO: 92) where the expression of the gene was driven by a T7 promoter. In a second plasmid, an in-frame, C-terminally His-tagged (SEQ.ID NO: 89, SEQ.ID NO: 90) PsmF (SEQ.ID NO: 62, SEQ.ID NO: 63) was cloned into the *NdeI* and *XhoI* sites of pET28a(+) (SEQ.ID NO: 94) where the expression of the gene was also driven by a T7 promoter. The *E. coli* host strain BL21 (DE3) was transformed with both plasmids and selected for using streptomycin and kanamycin antibiotics. This strain (Ec-1, F<sup>-</sup> *ompT hsdSB* (rB<sup>-</sup>, mB<sup>-</sup>) *gal dcm* (DE3), pCDM4-BaTDC-His, pET28a(+)-PsmF-His) enabled production of *N*-acetylated products from modified indole feedstock. Further, the use of promiscuous BaTDC (SEQ.ID NO: 36, SEQ.ID NO: 37) allowed decarboxylation of a wide variety of halogenated tryptophan intermediates. Scaled-up culturing of engineered *E. coli* was conducted as follows: seed cultures were inoculated in AMM (Jones *et al.*, 2015, Sci. Rep. 5: 11301) medium overnight. The overnight culture was then divided into two flasks containing 500 mL each of AMM medium additionally containing 0.5% (w/v) serine, 1M IPTG, 50ug/L kanamycin and streptomycin, and 100 mg/L 7-fluoroindole (Combi-Blocks, [www.combi-blocks.com](http://www.combi-blocks.com)) for conversion by the *E. coli* strain. Cultures were grown for 24 h. Cultures were then centrifuged (10,000g x 5 minutes) to remove cellular content, and culture broth containing secreted derivative was stored at -80C until further processing. The *N*-(2-(7-difluoro-1*H*-indol-3-yl)ethyl)acetamide product having chemical formula (XV):



**[00535]** contained in 1 L of *E. coli* culture was extracted by ethyl acetate (3×300 ml). The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, followed by concentration under reduced pressure. The residue was purified by column chromatography on

silica gel using ethyl acetate – hexane (50 → 80 %) as eluent to give the compound as a light yellow solid (22 mg). Following purification, high-resolution MS (HRMS), <sup>1</sup>H NMR, and selective <sup>13</sup>C NMR were performed to assess purity, estimate total quantity, and confirm molecular structure. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ = 1.96 (s, 3H), 2.97 (t, *J* = 6.8 Hz, 2H), 3.59 (dt, *J* = 6.7, 6.1 Hz, 2H), 5.7 (br, s, 1H), 6.91 (dd, *J* = 11.2, 7.8 Hz, 1H), 7.03 (m, 2H), 7.35 (d, *J* = 7.9 Hz, 1H), 8.6 (br, s, 1H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>): δ = 23.3, 25.3, 39.8, 106.9 (d, *J*<sub>C,F</sub> = 16.1 Hz), 113.7, 114.4 (d, *J*<sub>C,F</sub> = 3.4 Hz), 119.6 (d, *J*<sub>C,F</sub> = 6.1 Hz), 122.8, 124.7, 131.1 (d, *J*<sub>C,F</sub> = 5.3 Hz), 149.7 (d, *J*<sub>C,F</sub> = 244.2 Hz), 170.3. HRMS (ESI) *m/z*: calcd. for C<sub>12</sub>H<sub>13</sub>FN<sub>2</sub>O [M+H]<sup>+</sup> 221.1090, found 221.1087. The compound having chemical formula (XV) was 95% (w/w) pure. It is noted that the acetyl group of the compound having chemical formula (XV) is included therein by virtue of the acetyl transferase having SEQ.ID NO: 63 in *Escherichia coli* strain Ec-1.

**[00536]** Cell lines for pharmacology assays. CHO-K1/Galpha15 (GenScript, M00257) (-5-HT<sub>2A</sub>) and CHO-K1/5-HT<sub>2A</sub> (GenScript, M00250) (+5-HT<sub>2A</sub>) cells lines were used in both toxicology/growth inhibition (MTT) and calcium release assays. Briefly, CHO-K1/Galpha15 is a control cell line that constitutively expresses Galpha15 which is a promiscuous Gq protein. It is engineered as a host cell, allowing transfected receptor(s) to signal through the Gq signal transduction pathway and mobilize intracellular calcium from the endoplasmic reticulum (ER). These control cells lack any transgene encoding 5-HT<sub>2A</sub> receptors, thus preventing calcium mobilization in response to 5-HT<sub>2A</sub> activation. Conversely, CHO-K1/5-HT<sub>2A</sub> cells stably express 5-HT<sub>2A</sub> receptor in the CHO-K1 host background. This design enables Gq-11 expressed in CHO-K1 cells to mobilize intracellular calcium changes when 5-HT<sub>2A</sub> receptors are activated by ligands.

**[00537]** Cell lines were maintained in Ham's F12 media plus 10% FBS in the presence of 100 ug/ml hygromycin for CHO-K1/Ga15 or 400 ug/ml G418 for CHO-K1/5-HT<sub>2A</sub> unless indicated otherwise for specific assays. Cell maintenance was carried out as recommended by the cell supplier. Briefly, vials with cells were removed from the liquid nitrogen and thawed quickly in 37°C water bath. Just before cells were completely thawed, vial exteriors were decontaminated with 70% ethanol spray. Cell suspension was then retrieved from the vial and added to warm (37°C), 'complete' (non-dropout) growth media, and centrifuged at 1,000 rpm for 5

minutes. The supernatant was discarded and the cell pellet was then resuspended in another 10 ml of complete growth media, and added to a 10 cm cell culture dish (Greiner Bio-One #664160). The media was changed every third day until the cells reached ~90% confluence. The ~90% confluent cells were then split 10:1, and used either for maintenance or pharmacological study.

**[00538]**      *Assessment of cell viability upon treatment of glycosylated psilocybin derivatives.* To establish suitable ligand concentrations for the calcium release assays, MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium) assays were first performed. Results of these assays were conducted using both control ligands (e.g. psilocybin, psilocin, DMT) and novel derivatives, in part as a pre-screen for any remarkable toxic effects on cell cultures up to concentrations of 1 mM. A known cellular toxin (Triton X-100, Pyrgiotakis G. *et al.*, 2009, Ann. Biomed. Eng. 37: 1464-1473) was included as a general marker of toxicity. Modified Chinese Hamster Ovary cells (CHO-K1/Ga15) were cultured using standard procedures using the manufacture's protocols (Genscript, M00257). Briefly, cells were cultured in Ham's F12 medium supplemented with 10% fetal bovine serum and 100 mg/ml Hygromycin B, and grown at 37°C in the presence of 5% CO<sub>2</sub>. To test the various compounds with the cell line, cells were seeded in a clear 96-well culture plate at 10,000 cells per well. After allowing cells to attach and grow for 24 hours, assay compounds were added at 1 µM, 10 µM, 100 µM, and 1 mM final concentrations. Methanol concentrations used are 0.001, 0.01, 0.1, and 1%. Triton concentrations used are 0.0001, 0.001, 0.01 and 0.1%. Cells were incubated with compounds for 48 hours before accessing cell viability with the MTT assay following the manufacture's protocol (MTT Cell Growth Assay Kit; Millipore Sigma, CT02). MTT reagent was added to cells and allowed to incubate for 4 hours before solubilization with isopropanol plus 0.04 N HCl. Absorbance readings were performed at 570 nm with the reference at 630 nm on a SpectraMax iD3 plate reader. Non-treated cells were assigned 100% viability. Results of the cell viability assays are shown in **FIG. 14**. Bar graphs show the mean +/- SD (n=3). Significance (P<0.0001), as indicated by (\*\*\*) was determined using 2-way ANOVA with Dunnett's multiple comparisons test. The results using compound with formula (XV) are indicated as "(XV)" on the x-axis.

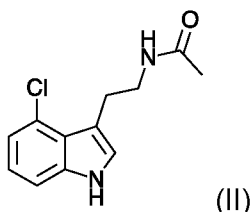
**[00539]**      *Increase in cytosolic calcium concentration by 5-HT<sub>2A</sub> activation.* Changes in intracellular calcium concentration due to the treatment with assay

compounds was measured using Fluo-8 dye (Abcam, #ab112129) according to the manufacturer's instructions. Briefly, CHO-K1 cells stably expressing 5-HT<sub>2A</sub> (Genscript # M00250) (+5-HT<sub>2A</sub>) or lacking 5-HT<sub>2A</sub> (Genscript, M00257) (-5-HT<sub>2A</sub>) were seeded on black walled clear bottom 96-well plates (Thermo Scientific #NUNC165305), allowing 70,000 cells/well in 100 ul media (HAM's F12, GIBCO #11765-047) with 1% FBS (Thermo Scientific #12483020). Cultures were maintained in a humidified incubator at 37°C and 5% CO<sub>2</sub>. Fluo-8 dye was loaded into the cultures for 30 min at 37°C, followed by 30 min additional incubation at room temperature. Next, different dilutions of novel molecules and controls were prepared in serum-free culture media and added to the cells. Fluorescence (ex 490 nm / em 525 nm) obtained after the addition of molecules was expressed relative to values obtained before addition of the molecules (relative Fluo-8 fluorescence = F<sub>max</sub>/ F<sub>0</sub>, where F<sub>max</sub>=maximum fluorescence and F<sub>0</sub>=baseline fluorescence). Fluorescence intensities were measured using a Spectramax ID3 plate reader (www.moleculardevices.com). Relative fluorescence (RFU) at increasing concentrations of compound was determined, illustrating concentration-dependent calcium flux. Psilocin, a known agonist with binding activity at 5-HT<sub>2A</sub> (Rickli A. *et al.*, 2016, *Europ. Neuropsychopharmacol.*, 26: 1326-1337) was used as a positive control to establish assay functionality. The Example compound (XV) was then evaluated. Results are shown in **FIGS. 15A – 15B** and **16A – 16B**. The results using different concentrations of psilocin and compound with formula (XV) (indicated as "(XV)") in +5-HT<sub>2A</sub> cell cultures are shown in **FIG. 15A** and **FIG. 15B**, respectively. The results using different concentrations of psilocin and compound with formula (XV) (indicated as "(XV)") in -5-HT<sub>2A</sub> cell cultures are shown in **FIG. 16A** and **FIG. 16B**, respectively. Results for negative control treatment with vehicle containing neither psilocin nor derivative are shown in **FIGS. 15A – 15B** and **16A – 16B**. Error bars show +/- SD about the mean (n=3).

**Example 6 – Process for making a sixth halogenated psilocybin derivative from halogenated indole feedstock.**

**[00540]** *Escherichia coli* strain Ec-1 was used to biosynthesize halogenated tryptamine derivative having chemical formula (II) from halogenated indole

feedstock. The construction of Ec-1 is described in Example 5. Scaled-up culturing and material storage of engineered *E. coli* was conducted as described in Example 5, except that 4-chloroindole (Combi-Blocks, [www.combi-blocks.com](http://www.combi-blocks.com)) was used in place of 7-fluoroindole. To assess product, high-resolution LC-HESI-LTQ-Orbitrap-XL MS analysis was conducted as described in Example 1. Singly protonated product with exact  $m/z$  and expected elemental formula matching 4-chloro-N-acetyltryptamine having chemical formula (II):

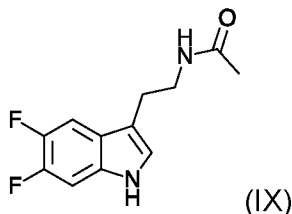


Eluted at 4.1 minutes (EIC, see: **FIG. 17**). It is noted that the acetyl group of the compound having chemical formula (II) is included therein by virtue of the acetyl transferase having SEQ.ID NO: 63 in *Escherichia coli* strain Ec-1.

**Example 7 – Process for making a seventh halogenated psilocybin derivative from halogenated indole feedstock.**

15

**[00541]** *Escherichia coli* strain Ec-1 was used to biosynthesize halogenated tryptamine derivative having chemical formula (IX) from halogenated indole feedstock. The construction of Ec-1 is described in Example 5. Scaled-up culturing and material storage of engineered *E. coli* was conducted as described in Example 5, except that 5,6-difluoroindole (Combi-Blocks, [www.combi-blocks.com](http://www.combi-blocks.com)) was used in place of 7-fluoroindole. N-(2-(5,6-difluoro-1H-indol-3-yl)ethyl)acetamide product having Chemical formula (IX):



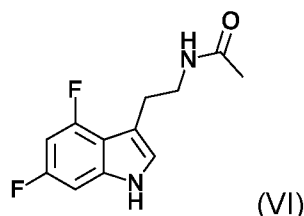
**[00542]** contained in 0.55 L of *E. coli* culture was extracted by ethyl acetate (3×300 ml). The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, followed by concentration under reduced pressure. The residue was purified by column chromatography on

silica gel using ethyl acetate – hexane (50 → 80 %) as eluent to give the compound as a light yellow solid (11 mg). Following purification, high-resolution MS (HRMS), <sup>1</sup>H NMR, and selective <sup>13</sup>C NMR were performed to assess purity, estimate total quantity, and confirm molecular structure. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ = 1.70-1.79 (m, 1H), 1.99-2.06 (m, 1H), 2.29 (s, 6H), 2.33-2.38 (m, 1H), 2.40-2.48 (m, 1H), 3.22 (t, 1H, *J* = 8.0 Hz), 3.28-3.36 (m, 1H), 3.68 (t, 1H, *J* = 8.4 Hz), 3.94 (br, 1H), 6.65 (d, 1H, *J* = 8.4 Hz), 6.74 (t, 1H, *J* = 7.6 Hz), 7.05 (t, 1H, *J* = 7.6 Hz), 7.11 (d, 1H, *J* = 7.6 Hz). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>): δ = 23.4, 25.2, 39.7, 99.1 (d, *J*<sub>C, F</sub> = 21.7 Hz), 105.4 ((d, *J*<sub>C, F</sub> = 20.0 Hz), 123.3 (d, *J*<sub>C, F</sub> = 3.7 Hz), 150.6 (dd, *J*<sub>C, F</sub> = 237.4, 17.2 Hz), 148.9 (dd, *J*<sub>C, F</sub> = 238.3, 18.1 Hz), 170.1. HRMS (ESI) calcd. for C<sub>12</sub>H<sub>12</sub>F<sub>2</sub>N<sub>2</sub>O 239.0996 [M+H]<sup>+</sup>, found 239.0989. The compound having chemical formula (IX) was 95% (w/w) pure. It is noted that the acetyl group of the compound having chemical formula (IX) is included therein by virtue of the acetyl transferase having SEQ.ID NO: 63 in *Escherichia coli* strain Ec-1.

**[00543]** Efficacy testing was carried out as described in Example 5. MTT assay results are shown in **FIG. 18**, and calcium mobility assay results are shown in **FIG. 19A** (+5-HT<sub>2A</sub> cells) and **FIG. 19B** (-5-HT<sub>2A</sub> cells). In both cases, the example compound having the chemical formula (IX) is shown as “(IX)”.

## **Example 8 – Process for making an eighth halogenated psilocybin derivative from halogenated indole feedstock.**

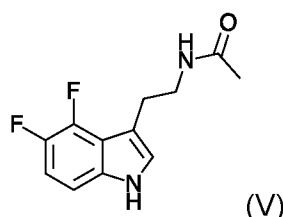
**[00544]** *Escherichia coli* strain Ec-1 was used to biosynthesize halogenated tryptamine derivative having chemical formula (VI) from halogenated indole feedstock. The construction of Ec-1 is described in Example 5. Scaled-up culturing and material storage of engineered *E. coli* was conducted as described in Example 5, except that 4,6-difluoroindole (Combi-Blocks, [www.combi-blocks.com](http://www.combi-blocks.com)) was used in place of 7-fluoroindole. To assess product, high-resolution LC-HESI-LTQ-Orbitrap-XL MS analysis was conducted as described in Example 1. Singly protonated product with exact *m/z* and expected elemental formula matching 4,6-difluoro-N-acetyltryptamine having chemical formula (VI):



eluted at 5.1 minutes (EIC, see: **FIG. 20**).

**Example 9 – Process for making a ninth halogenated psilocybin derivative from halogenated indole feedstock.**

**[00545]** *Escherichia coli* strain Ec-1 was used to biosynthesize halogenated tryptamine derivative having chemical formula (V) from halogenated indole feedstock. The construction of Ec-1 is described in Example 5. Scaled-up culturing and material storage of engineered *E. coli* was conducted as described in Example 5, except that 4,5-difluoroindole (Combi-Blocks, [www.combi-blocks.com](http://www.combi-blocks.com)) was used in place of 7-fluoroindole. *N*-(2-(4,5-difluoro-1*H*-indol-3-yl)ethyl)acetamide product having chemical formula (V):



**[00546]** contained in 1 L of *E. coli* culture was extracted by ethyl acetate (3×600 ml). The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub>, followed by concentration under reduced pressure. The residue was purified by flash chromatography on silica gel using ethyl acetate – hexane (50 → 80 %) as eluent, followed by crystallization from EtOAc/Hexane to give the compound as a white solid (20 mg). Following purification, high-resolution MS (HRMS), <sup>1</sup>H NMR, and selective <sup>13</sup>C NMR were performed to assess purity, estimate total quantity, and confirm molecular structure. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ = 1.96 (s, 3H), 3.04 (t, *J* = 6.9 Hz, 2H), 3.60 (dt, *J* = 6.6, 6.4 Hz, 2H), 5.64 (br, s, 1H), 7.01 (m, 3H), 8.41 (br, s, 1H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>): δ = 23.4, 26.3, 40.5 (d, *J*<sub>C,F</sub> = 2.0 Hz), 106.6 (dd, *J*<sub>C,F</sub> = 7.7, 4.2 Hz), 111.7 (d, *J*<sub>C,F</sub> = 21.7 Hz), 124.0, 143.9 (dd, *J*<sub>C,F</sub> = 246.4, 14.5 Hz), 143.8 (dd, *J*<sub>C,F</sub> = 246.4, 14.5 Hz), 170.2. HRMS (ESI) *m/z*: calcd. for C<sub>12</sub>H<sub>12</sub>F<sub>2</sub>N<sub>2</sub>O [M+H]<sup>+</sup> 239.0996, found 239.0989. The compound having chemical

formula (V) was 95% (w/w) pure. It is noted that the acetyl group of the compound having chemical formula (II) is included therein by virtue of the acetyl transferase having SEQ.ID NO: 63 in *Escherichia coli* strain Ec-1.

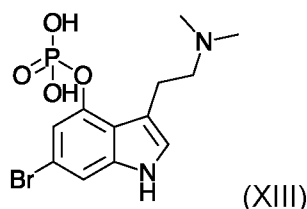
5 [00547] Efficacy testing was carried out as described in Example 5. MTT assay results are shown in **FIG. 21**, and calcium mobility assay results are shown in **FIGS. 22A** and **22B**. In both cases, the example compound having the chemical formula (V) is shown as “(V)”.

10 **Example 10 – Process for making a tenth halogenated psilocybin derivative from halogenated indole feedstock.**

[00548] *E. coli* strain Ec-2 was constructed as follows. For plasmid cloning, Top10 or XL1-blue strains were used depending on antibiotic markers. Standard LB media was used for culturing. For gene expression and feeding experiments, 15 the parent host strain employed was BL21 (DE3). For this Example, heterologous expression of a non-native or engineered TrpB gene was not necessary, as endogenous tryptophan synthase activity proved sufficient. From plasmid pCDM4 (SEQ.ID NO: 92), the plasmid pCDM4-PsiD-FLAG was created by inserting an in-frame, C-terminally FLAG-tagged (SEQ.ID NO: 85) PsiD gene (SEQ.ID NO: 1) into 20 the *NdeI/XhoI* site of pCDM4. The plasmid pETM6-H10-PsiK-V5-PsiM-FLAG was created by first cloning the in-frame, C-terminally V5-tagged (SEQ.ID NO: 87) PsiK (SEQ.ID NO: 5) into the *NdeI/XhoI* site of pETM6-H10 (SEQ.ID NO: 97) to create pETM6-H10-PsiK-V5. This intermediate plasmid was digested with *SpeI* and *Sall*, and in-frame, C-terminally FLAG tagged (SEQ.ID NO: 85) PsiM (SEQ.ID NO: 7) 25 was cloned into the site with *XbaI* and *Sall*, nullifying the *SpeI* restriction site. In this setup, the T7 polymerase was able to drive the expression of the polycistronic DNA containing both PsiK and PsiM. The two target plasmids pCDM4-PsiD-FLAG and pETM6-H10-PsiK-V5-PsiM-FLAG were transformed into BL21 (DE3) cells, and antibiotics ampicillin plus streptomycin were used to select for the correct 30 clones containing both plasmids. Scaled-up culturing of engineered *E.coli* was conducted as follows: seed cultures were inoculated in AMM (Jones *et al.* 2015, Sci Rep. 5: 11301) medium overnight. The overnight culture was then divided into two flasks containing 500 mL each of AMM medium additionally containing 0.5% (w/v) serine, 1M IPTG, 50ug/L ampicillin and streptomycines, and 100 mg/L 4-



hydroxy-6-bromoindole (BLDPharm, www.bldpharm.com) for conversion by Ec-2. Cultures were grown for 24 h. Cultures were then centrifuged (10,000g x 5 minutes) to remove cellular content, and culture broth containing secreted derivative was stored at -80C until further processing. To assess product, high-resolution LC-HESI-LTQ-Orbitrap-XL MS analysis was conducted as described in Example 1. Singly protonated product with exact  $m/z$  and expected elemental formula matching 6-bromo-psilocybin having chemical formula (XIII):

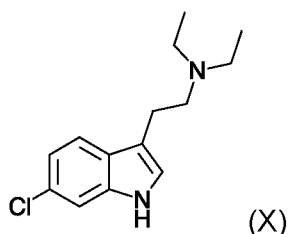


eluted at 2.8 minutes (EIC, see: **FIG. 23**).

**Example 11 – Process for making an eleventh halogenated psilocybin derivative from halogenated indole feedstock.**

**[00549]** In a flame-dried round bottom flask under argon, 6-chloroindole (**compound 1**, see: **FIG. 9A**) (210 mg, 1.39 mmol, 1.00 eq) was dissolved in diethyl ether (4.0 mL) and cooled to 0 °C. To this solution, oxalyl chloride (369 mg, 2.91 mmol, 2.10 eq) in diethyl ether (2.0 mL) was added dropwise over a period of 10 minutes. The reaction mixture was allowed to stir at room temperature for 4 hours. Additional ether (5.0 mL) was added to the reaction mixture which was then cooled to 0 °C and subsequently vacuum-filtered to yield a bright yellow solid (**compound 2**) (211 mg, 0.872 mmol, 63%). **Compound 2** (see: **FIG. 9A**) was washed with cold ether (10 mL) and allowed to briefly dry before being resuspended in diethyl ether (3.0 mL) and cooled to 0 °C. Diethylamine (510 mg, 6.97 mmol, 8.00 eq) was diluted with diethyl ether (3.0 mL) and added dropwise to the reaction flask, immediately precipitating out a white solid. This mixture was allowed to stir for 10 minutes at 0 °C, warmed to room temperature, and stirred for an additional 1.5 hours. The visible precipitate was vacuum filtered to yield a solid which was washed with cold diethyl ether (10 mL). This solid was suspended in

deionized water (50 mL), stirred for 10 minutes, vacuum-filtered, and allowed to dry overnight, yielding **compound 3** (see: **FIG. 9A**) as a white powder (124 mg, 0.445 mmol, 51%). No further purification was required. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) = 9.56 (s, 1H), 8.24 (d, *J* = 8.5 Hz, 1H), 7.74 (d, *J* = 3.0 Hz, 1H), 7.36 (dd, *J* = 1.8, 0.6 Hz, 1H), 7.28 (d, *J* = 1.8 Hz, 1H), 3.54 (q, *J* = 7.1 Hz, 2H), 3.36 (q, *J* = 7.1 Hz, 2H), 1.27 (t, *J* = 7.2 Hz, 3H), 1.19 (t, *J* = 7.1 Hz, 3H). A THF solution of lithium aluminum hydride (2.20 mL, 2.22 mmol, 5.00 eq, 1.0 M) was added to a 3-neck, flame-dried, round-bottom flask under argon. Anhydrous dioxane (3.0 mL) was added to the reaction mixture. In a separate vial, **compound 3** (124 mg, 0.445 mmol) was dissolved in a THF (2.0 mL)/dioxane (1.0 mL) solution and added dropwise to the reaction vessel. This mixture was then allowed to reflux at 95 °C for 1.5 hours. The reaction mixture was cooled to 0 °C and quenched through dropwise addition of a water (1.0 mL)/THF (2.0 mL) solution, yielding a flocculent precipitate. This precipitate was removed through vacuum-filtration and the filter-cake was washed with diethyl ether (20 mL). The filtrate was collected, dried with MgSO<sub>4</sub>, and solvent removed to yield a pale-yellow oil. Purification through column chromatography on silica gel (2% MeOH/DCM + 1% NEt<sub>3</sub>) yielded compound **410B01** (see: **FIG. 9A**), corresponding to compound with formula (X):



20

as a clear, colourless oil (34 mg, 0.136 mmol, 30%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) = 8.37 (s, 1H), 7.49 (d, *J* = 8.5 Hz, 1H), 7.30 (d, *J* = 1.8 Hz, 1H), 7.07 (dd, *J* = 8.5, 1.8 Hz, 1H), 6.98 (dd, *J* = 2.5, 1.1 Hz, 1H), 2.90 (high order m, 2H), 2.79 (high order m, 2H), 2.74 – 2.62 (q, *J* = 7.1 Hz, 4H), 1.15 – 1.04 (t, *J* = 7.1 Hz, 6H). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>): δ (ppm) = 136.6, 127.8, 126.2, 122.2, 119.9, 119.7, 114.8, 111.0, 70.5, 53.5, 46.9, 22.8, 11.7. The compound having chemical formula (IX) was 95% (w/w) pure.

25

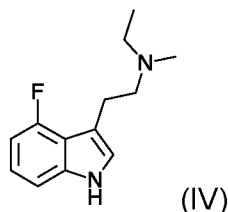
**[00550]** Efficacy testing was carried out as described in Example 5. MTT assay results are shown in **FIG. 24**, and calcium mobility assay results are shown

in **FIGS. 25A** and **25B**. In both cases, the example compound having chemical formula (X) is shown as "(X)".

**Example 12 – Process for making a twelfth halogenated psilocybin derivative from halogenated indole feedstock.**

**[00551]** A similar process as illustrated in **FIG. 9A** (Example 11) was followed. In a flame-dried round bottom flask under argon, 4-fluoroindole (**compound 1**) (200 mg, 1.48 mmol, 1.00 eq) was dissolved in diethyl ether (4.0 mL) and cooled to 0 °C. To this solution, oxalyl chloride (394 mg, 3.11 mmol, 2.10 eq) in diethyl ether (2.0 mL) was added dropwise over a period of 10 minutes. The reaction mixture was allowed to stir at room temperature for 4 hours. Additional ether (5.0 mL) was added to the reaction mixture which was then cooled to 0 °C and subsequently vacuum-filtered to yield a bright yellow solid (**compound 2**) (264 mg, 1.17 mmol, 79%). **Compound 2** was washed with cold ether (10 mL) and allowed to briefly dry before being resuspended in diethyl ether (3.0 mL) and cooled to 0 °C. N-Methylethylamine (574 mg, 9.72 mmol, 8.00 eq) was diluted with diethyl ether (3.0 mL) and added dropwise to the reaction flask. This mixture was allowed to stir for 10 minutes at 0 °C, warmed to room temperature, and stirred for an additional 1.5 hours, yielding a suspension of light-pink gooey solid. The diethyl ether was decanted, and the solid dissolved in dichloromethane (10 mL) and washed with water (3 x 10 mL). The organic phase was dried with MgSO<sub>4</sub> and solvent removed to yield a pale-pink glassy solid (128 mg, 0.516 mmol, 43%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) = 10.72 (s, 1H), 7.72 (d, *J* = 3.8 Hz, 1H), 7.10 – 7.00 (m, 2H), 6.84 (ddd, *J* = 11.1, 7.2, 1.6 Hz, 1H), 3.55 (q, *J* = 7.2 Hz, 1H), 3.35 (q, *J* = 7.1 Hz, 1H), 3.05 (s, 2H), 2.98 (s, 1H), 1.23 (t, *J* = 7.2 Hz, 1H), 1.15 (t, *J* = 7.1 Hz, 2H). A THF solution of lithium aluminum hydride (2.58 mL, 2.58 mmol, 5.00 eq, 1.0 M) was added to a 3-neck, flame-dried, round-bottom flask under argon. Anhydrous dioxane (3.0 mL) was added to the reaction mixture. In a separate vial, **compound 3** (128 mg, 0.516 mmol) was dissolved in an anhydrous THF (2.0 mL)/dioxane (1.0 mL) solution and added dropwise to the reaction vessel. This mixture was then allowed to reflux at 95 °C for 2.5 hours. The reaction mixture was cooled to 0 °C and quenched through dropwise addition of a water (1 mL)/THF (2 mL) solution, yielding a flocculent precipitate. This precipitate was removed

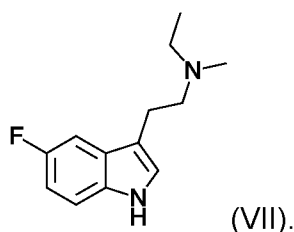
through vacuum-filtration and the filter-cake was washed with diethyl ether (20 mL). The filtrate was collected, dried with  $\text{MgSO}_4$ , and solvent removed to yield a pale-orange oil. Purification through column chromatography on silica gel (1%  $\text{MeOH}/\text{CH}_2\text{Cl}_2$  + 1%  $\text{NEt}_3$ ) yielded compound **456B01**, corresponding to  
 5 compound having chemical formula (IV):



as clear, colourless oily crystal (86 mg, 0.390 mmol, 76%).  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 8.68 (s, 1H), 7.12 – 6.99 (m, 2H), 6.91 (d,  $J$  = 2.3 Hz, 1H), 6.73 (ddd,  $J$  = 11.2, 6.9, 1.6 Hz, 1H), 3.04 (high order m, 2H), 2.75 (high order m, 2H),  
 10 2.57 (q,  $J$  = 7.2 Hz, 2H), 2.38 (s, 3H), 1.12 (t,  $J$  = 7.2 Hz, 3H).  $^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 157.2 (d,  $J_{\text{C-F}}$  = 245.6 Hz), 139.2 (d,  $J_{\text{C-F}}$  = 11.8 Hz), 122.2 (d,  $J_{\text{C-F}}$  = 8.2 Hz), 121.9, 107.2 (d,  $J_{\text{C-F}}$  = 3.6 Hz), 104.2 (d,  $J_{\text{C-F}}$  = 19.7 Hz), 58.5, 51.2, 41.5, 24.1, 12.1. The compound having chemical formula (IV) was 95% (w/w) pure.  
**[00552]** Efficacy testing was carried out as described in Example 5. MTT  
 15 assay results are shown in **FIG. 26**, and calcium mobility assay results are shown in **FIG. 27A** and **FIG. 27B**. In both cases, the example compound having the chemical formula (IV) is shown as “(IV)”.

**Example 13 – Process for making a thirteenth halogenated psilocybin  
 20 derivative from halogenated indole feedstock.**

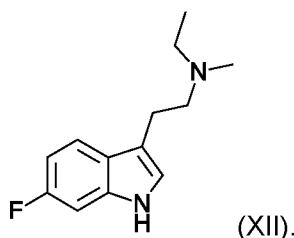
**[00553]** A thirteenth halogenated psilocybin derivative compound was synthesized according to the method of Example 11, except using 5-fluoroindole and N-methylethylamine as starting materials. Analysis indicated the isolation of a  
 25 compound having chemical formula (VII):



<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) = 8.09 (s, 1H), 7.29 – 7.20 (m, 2H), 7.06 (d, *J* = 2.4 Hz, 1H), 6.96 – 6.87 (m, 1H), 2.93 – 2.86 (m, 2H), 2.72 – 2.64 (m, 2H), 2.53 (q, *J* = 7.2 Hz, 2H), 2.35 (s, 3H), 1.11 (t, *J* = 7.2 Hz, 3H).

**Example 14 – Process for making a fourteenth halogenated psilocybin derivative from halogenated indole feedstock.**

10 **[00554]** A fourteenth halogenated psilocybin derivative was synthesized according to the method of Example 11, except using 6-fluoroindole and N-methylethylamine as starting materials. Analysis indicated the isolation of a compound having chemical formula (XII):

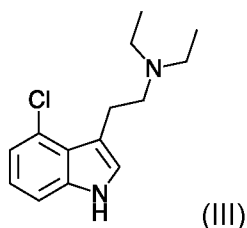


15 <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) = 8.05 (s, 1H), 7.50 (ddt, *J* = 8.7, 5.3, 0.6 Hz, 1H), 7.02 (ddd, *J* = 9.7, 2.3, 0.5 Hz, 1H), 7.00 – 6.97 (m, 1H), 6.88 (ddd, *J* = 9.6, 8.6, 2.3 Hz, 1H), 2.96 – 2.89 (m, 2H), 2.74 – 2.66 (m, 2H), 2.53 (q, *J* = 7.2 Hz, 2H), 2.35 (s, 3H), 1.11 (t, *J* = 7.2 Hz, 3H).

20

**Example 15 – Process for making a fifteenth halogenated psilocybin derivative from halogenated indole feedstock.**

25 **[00555]** A fifteenth halogenated psilocybin derivative was synthesized according to the method of Example 11, except using 4-chloroindole as starting material. Analysis indicated the isolation of a compound having chemical formula (III):

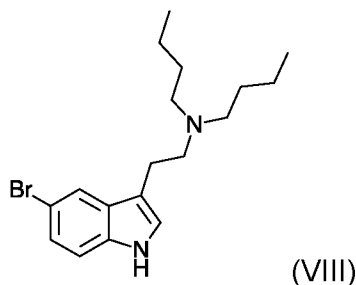


<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ (ppm) = 8.28 (s, 1H), 7.26 – 7.19 (m, 1H), 7.08 – 7.01 (m, 3H), 3.18 – 3.10 (m, 2H), 2.86 – 2.77 (m, 2H), 2.67 (q, *J* = 7.1 Hz, 4H),  
 5 1.09 (t, *J* = 7.1 Hz, 6H).

**Example 16 – Process for making a sixteenth halogenated psilocybin derivative from halogenated indole feedstock.**

10 **[00556]** The initial reaction to generate **compound 1.1** (see: **FIG. 9B**) followed a previously reported procedure with modifications (Campos *et al.* 2004, Org. Lett 6: 79-82). To a solution of 4-bromophenylhydrazine hydrochloride (300 mg, 1.34 mmol) in a mixture of *N,N*-dimethylacetamide (1.9 mL) and 4% H<sub>2</sub>SO<sub>4</sub> (1.9 mL), was added 2,3-dihydrofuran (0.10 mL, 1.33 mmol). The reaction mixture  
 15 was heated at 100°C for 16 hours. Once the mixture was cooled to room temperature, ethyl acetate (15 mL) was added, and the organic solution was washed with water (2 × 6 mL), and dried over anhydrous MgSO<sub>4</sub>. The organic solvent was removed *in vacuo*. The product (**compound 1.1**, (see: **FIG. 9B**)) was purified by flash chromatography on silica gel (eluted with a gradient of hexanes-EtOAc, 100:0 to 0:100) to afford an orange oil. Yield: 37%. The characterization of  
 20 this first product agreed with previously reported procedure (Go *et al.* 2010, J. Med. Chem. 53: 6838-6850). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.13 (bs, 1H), 7.74 (d, *J* = 1.7 Hz, 1H), 7.28 (dd, *J* = 8.6, 1.9 Hz, 1H), 7.23 (d, *J* = 8.6 Hz, 1H), 7.08 (d, *J* = 2.3 Hz, 1H), 3.89 (t, *J* = 6.3 Hz, 2H), 2.98 (t, *J* = 6.4 Hz, 2H). The second reaction  
 25 (see: **FIG. 9B**) followed a previously reported procedure (Lombardo *et al.* 2013, Angew. Chem. Int. Ed. 52: 12910-12914). Imidazole (73 mg, 1.1 mmol), triphenyl phosphine (140 mg, 0.54 mmol), and iodine (140 mg, 0.54 mmol) were added to a solution of **compound 1.1** (see: **FIG. 9B**) (119 mg, 0.49 mmol) in anhydrous dichloromethane (5.0 mL) cooled in an ice-water bath. The reaction was stirred at  
 30 room temperature for 6 hours. The mixture was quenched with saturated sodium

thiosulfate (10 mL), extracted in dichloromethane (3 × 10 mL), and the combined organic solution was dried over anhydrous MgSO<sub>4</sub>, and concentrated under reduced pressure. The crude product was purified by flash chromatography on silica gel (eluted with a gradient of hexanes-EtOAc, 100:0 to 50:50) to afford an orange oil. Yield of **compound 1.2** (see: **FIG. 9B**): 74%. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.07 (s, 1H), 7.71–7.69 (m, 1H), 7.29 (dd, *J* = 8.6, 1.7 Hz, 1H), 7.25 (d, *J* = 8.6 Hz, 1H), 7.10 (d, *J* = 2.4 Hz, 1H), 3.44–3.39 (m, 2H), 3.30 (t, *J* = 7.0 Hz, 2H). To a solution of **compound 1.2** (129 mg, 0.37 mmol) in anhydrous DMF (1.0 mL) was added di-*n*-butylamine (0.25 mL, 1.48 mmol). The reaction mixture was heated at 60°C for 16 hours. The reaction was quenched with saturated sodium bicarbonate solution, extracted in ethyl acetate (3 × 20 mL), and the combined organic solution was washed with water (2 × 40 mL), brine (2 × 40 mL), dried over anhydrous MgSO<sub>4</sub> and concentrated under reduced pressure. The crude mixture was purified by flash column chromatography on silica gel (eluted with a gradient of hexanes-EtOAc, 100:0 to 0:100) to afford an orange oil. Yield of **compound 1**, with structure as indicated by compound (VIII):



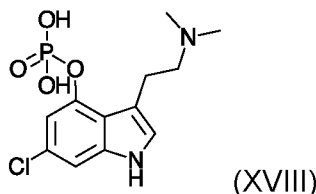
50%. <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 8.16 (s, 1H), 7.72 (s, 1H), 7.28 – 7.22 (m, 1H), 7.22–7.18 (m, 1H), 7.03–6.99 (m, 1H), 2.89–2.82 (m, 2H), 2.81–2.73 (m, 2H), 2.54 (t, *J* = 7.6 Hz, 4H), 1.52–1.43 (m, 5H), 1.33 (h, *J* = 7.0 Hz, 4H), 0.93 (t, *J* = 7.4 Hz, 6H) (Figure X, NMR). <sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>) δ 134.95, 129.58, 124.81, 122.83, 121.65, 114.83, 112.62, 112.57, 54.75, 54.01, 29.42, 22.87, 20.95, 14.27. HRMS (ESI, positive) *m/z* for C<sub>18</sub>H<sub>28</sub>BrN<sub>2</sub> [M+H]<sup>+</sup> calcd. 351.1430, found 351.1427. The compound having chemical formula (VIII) was 95% (w/w) pure.

**[00557]** Efficacy testing was carried out as described in Example 5. MTT assay results are shown in **FIG. 28**, and calcium mobility assay results are shown

in **FIGS. 29A** and **29B**. In both cases, the example compound having the chemical formula (VIII) is shown as “(VIII)”.

**Example 17 – Process for making a seventeenth halogenated psilocybin derivative from halogenated indole feedstock.**

**[00558]** *E. coli* strain Ec-3 was constructed as follows. For plasmid cloning, Top10 or XL1-blue strains were used depending on antibiotic markers. Standard LB media was used for culturing. For gene expression and feeding experiments, the parent host strain employed was BL21 (DE3). For this Example, heterologous expression of a non-native or engineered TrpB gene was not necessary, as endogenous tryptophan synthase activity proved sufficient. Plasmids pCDM4-PsiD-FLAG and pETM6-H10-PsiK-V5-PsiM-FLAG are described in Example 10. A third plasmid, pRSM3-Thal-HIS, was created by cloning an in-frame, C-terminally HIS-tagged (SEQ.ID NO: 89) Thal gene (SEQ.ID NO: 19), encoding a halogenase having SEQ.ID NO: 20, into the *Nde*//*Xho*I site of pRSM3 (SEQ.ID NO: 98). The three target plasmids pCDM4-PsiD-FLAG, pETM6-H10-PsiK-V5-PsiM-FLAG, and pRSM3-Thal-HIS were transformed into BL21 (DE3) cells and antibiotics ampicillin plus streptomycin plus kanamycin were used to select for the correct clones containing all three plasmids. Scaled-up culturing of engineered *E.coli* was conducted as described in Example 10. To assess product, high-resolution LC-HESI-LTQ-Orbitrap-XL MS analysis was conducted as described in Example 1. Singly protonated product with exact *m/z* and expected elemental formula matching 6-chloro-psilocybin having chemical formula (XVIII):

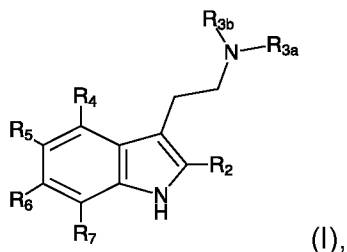


eluted at 2.3 minutes (EIC, see: **FIG. 30**).



**CLAIMS**

1. a chemical compound or salts thereof having the formula (I):



- 5 wherein, at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> is a halogen atom, wherein each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group, wherein R<sub>4</sub> when it is not halogenated is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group, and wherein R<sub>3A</sub> and R<sub>3B</sub> are a hydrogen atom, an alkyl group, an aryl group, or an acyl group.
- 10
2. A chemical compound according to claim 1, wherein the halogen atom is a bromine atom.
3. A chemical compound according to claim 1, wherein the halogen atom is a
- 15 chlorine atom.
4. A chemical compound according to claim 1, wherein the halogen atom is a fluorine atom.
- 20 5. A chemical compound according to claim 1, wherein the halogen atom is an iodine atom.
6. A chemical compound according to any one of claims 1 to 5, wherein R<sub>2</sub> is a halogen atom and R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> each are a hydrogen atom or an alkyl group or
- 25 O-alkyl group, and R<sub>4</sub> is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.
7. A chemical compound according to any one of claims 1 to 5, wherein R<sub>4</sub> is a halogen atom and R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> each are a hydrogen atom or an alkyl group
- 30 or O-alkyl group.

**8.** A chemical compound according to any one of claims **1** to **5**, wherein R<sub>5</sub> is a halogen atom and R<sub>2</sub>, R<sub>6</sub> and R<sub>7</sub> each are a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.

**9.** A chemical compound according to any one of claims **1** to **5**, wherein R<sub>6</sub> is a halogen atom or an alkyl group or O-alkyl group and R<sub>2</sub>, R<sub>5</sub> and R<sub>7</sub> each are a hydrogen atom, and R<sub>4</sub> is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.

**10.** A chemical compound according to any one of claims **1** to **5**, wherein R<sub>7</sub> is a halogen atom and R<sub>2</sub>, R<sub>5</sub> and R<sub>6</sub> each are a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.

**11.** A chemical compound according to any one of claims **1** to **5**, wherein at least two of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> are a halogen atom.

**12.** A chemical compound according to any one of claims **1** to **11**, wherein the at least two of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> that are halogen atoms are at least two identical halogen atoms.

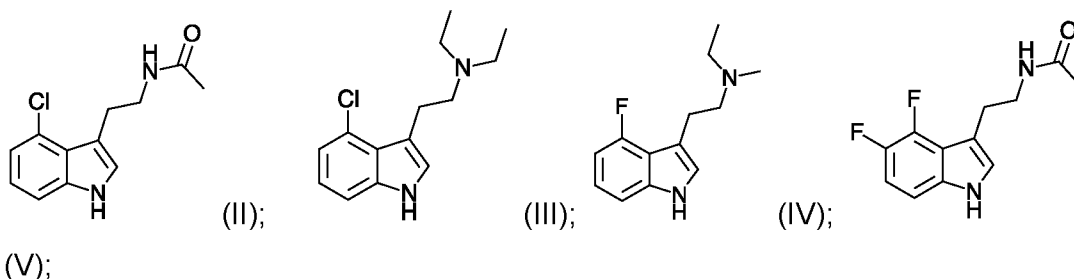
**13.** A chemical compound according to any one of claims **1** to **11**, wherein the at least two of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> that are halogen atoms are at least two non-identical halogen atoms.

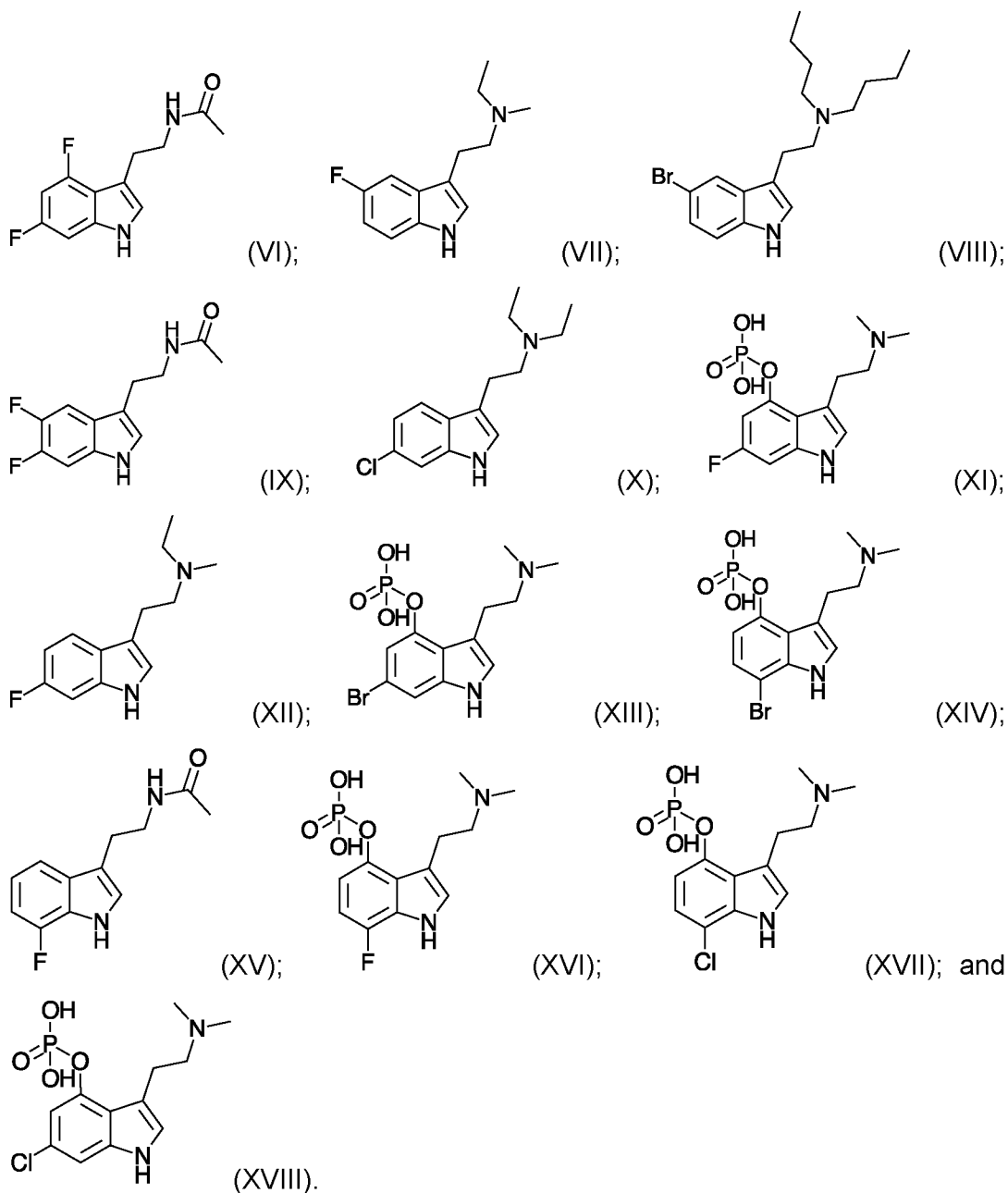
**14.** A chemical compound according to any one of claims **1** to **5**, wherein R<sub>2</sub> and R<sub>4</sub> are a halogen atom, and R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> are a hydrogen atom or an alkyl group or O-alkyl group.

**15.** A chemical compound according to any one of claims **1** to **5**, wherein R<sub>2</sub> and R<sub>5</sub> are a halogen atom, R<sub>6</sub> and R<sub>7</sub> are a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> is a hydrogen atom, or an alkyl group or O-alkyl group or a phosphate group.

16. A chemical compound according to any one of claims 1 to 5, wherein R<sub>2</sub> and R<sub>6</sub> are a halogen atom, R<sub>5</sub> and R<sub>7</sub> are a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> is a hydrogen atom, an alkyl group or O-alkyl group or a phosphate group.
- 5 17. A chemical compound according to any one of claims 1 to 5, wherein R<sub>2</sub> and R<sub>7</sub> are a halogen atom, R<sub>5</sub> and R<sub>6</sub> are a hydrogen atom, an alkyl group or O-alkyl group, and R<sub>4</sub> is a hydrogen atom, an alkyl group or O-alkyl group or a phosphate group.
- 10 18. A chemical compound according to any one of claims 1 to 5, wherein R<sub>4</sub> and R<sub>5</sub> are a halogen atom, and R<sub>2</sub>, R<sub>6</sub> and R<sub>7</sub> are a hydrogen atom or an alkyl group or O-alkyl group.
- 15 19. A chemical compound according to any one of claims 1 to 5, wherein R<sub>4</sub> and R<sub>6</sub> are a halogen atom, and R<sub>2</sub>, R<sub>5</sub> and R<sub>7</sub> are a hydrogen atom or an alkyl group or O-alkyl group.
- 20 20. A chemical compound according to any one of claims 1 to 5, wherein R<sub>4</sub> and R<sub>7</sub> are a halogen atom, and R<sub>2</sub>, R<sub>5</sub> and R<sub>6</sub> are a hydrogen atom or an alkyl group or O-alkyl group.
- 25 21. A chemical compound according to any one of claims 1 to 5, wherein R<sub>5</sub> and R<sub>6</sub> are a halogen atom, R<sub>2</sub> and R<sub>7</sub> are a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> is a hydrogen atom, an alkyl group or O-alkyl group or a phosphate group.
- 30 22. A chemical compound according to any one of claims 1 to 5, wherein R<sub>5</sub> and R<sub>7</sub> are a halogen atom, R<sub>2</sub> and R<sub>6</sub> are a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> is a hydrogen atom, an alkyl group or O-alkyl group or a phosphate group.
- 35 23. A chemical compound according to any one of claims 1 to 5, wherein R<sub>6</sub> and R<sub>7</sub> are a halogen atom, R<sub>2</sub> and R<sub>5</sub> are a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> is a hydrogen atom, an alkyl group or O-alkyl group or a phosphate group.

- 24.** A chemical compound according to any one of claims **1** to **23**, wherein  $R_4$  when it is not halogenated is a hydrogen atom or an alkyl group or O-alkyl group.
- 5 **25.** A chemical compound according to any one of claims **1** to **23**, wherein  $R_4$  when it is not halogenated is an alkyl group.
- 26.** A chemical compound according to any one of claims **1** to **23**, wherein  $R_4$  when it is not halogenated is an O-alkyl group.
- 10 **27.** A chemical compound according to any one of claims **1** to **23**, wherein  $R_4$  when it is not halogenated is a phosphate group.
- 28.** A chemical compound according to any one of claims **1** to **5**, wherein three,  
15 four or all five of  $R_2$ ,  $R_4$ ,  $R_5$ ,  $R_6$  or  $R_7$  are a halogen atom.
- 29.** A chemical compound according to claim **28**, wherein the three, four, or five of  $R_2$ ,  $R_4$ ,  $R_5$ ,  $R_6$  or  $R_7$  that are halogen atoms are all identical halogen atoms.
- 20 **30.** A chemical compound according to claim **28**, wherein the three, four or five of  $R_2$ ,  $R_4$ ,  $R_5$ ,  $R_6$  or  $R_7$  that are halogen atoms include at least two non-identical halogen atoms.
- 31.** A chemical compound according to claim **1**, wherein the chemical  
25 compound is selected from the group consisting of compounds having formulas (II); (III); (IV); (V); (VI); (VII); (VIII); (IX); (X); (XI); (XII); (XIII); (XIV); (XV); (XVI); and (XVII):





**32.** A chemical compound according to any one of claims **1** to **31** wherein the compound is at least about 95% pure.

10 **33.** A pharmaceutical or recreational drug formulation comprising an effective amount of the chemical compound according to claim **1**, together with a pharmaceutically acceptable excipient, diluent or carrier.

**34.** A method for treating a psychiatric disorder, the method comprising administering to a subject in need thereof a pharmaceutical formulation comprising a chemical compound according to claim **1**, wherein the pharmaceutical formulation is administered in an effective amount to treat the psychiatric disorder  
5 in the subject.

**35.** A method according to claim **34**, wherein the disorder is a 5-HT<sub>2A</sub> receptor mediated disorder.

10 **36.** A method according to claim **34** or **35**, wherein a dose is administered of about 0.001 mg to about 5,000 mg.

**37.** A method of making a halogenated psilocybin derivative of claim **1**, the method comprising:

- 15 (a) contacting a halogenated psilocybin precursor compound with a host cell comprising a psilocybin biosynthetic enzyme complement; and  
(b) growing the host cell to produce the halogenated psilocybin derivative.

20 **38.** A method according to claim **37**, wherein the halogenated psilocybin precursor compound is contacted with the host cell by including the psilocybin precursor compound in a growth medium for the host cell.

**39.** A method according to claim **38**, wherein the halogenated psilocybin  
25 precursor compound is formed by contacting the host cell with a non-halogenated psilocybin precursor compound and a halogen, the host cell further comprising a halogenase capable of halogenating the non-halogenated psilocybin compound and forming the halogenated psilocybin precursor compound.

30 **40.** A method according to claim **38** or **39**, wherein R<sub>2</sub> is a halogen atom and R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> each are a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.

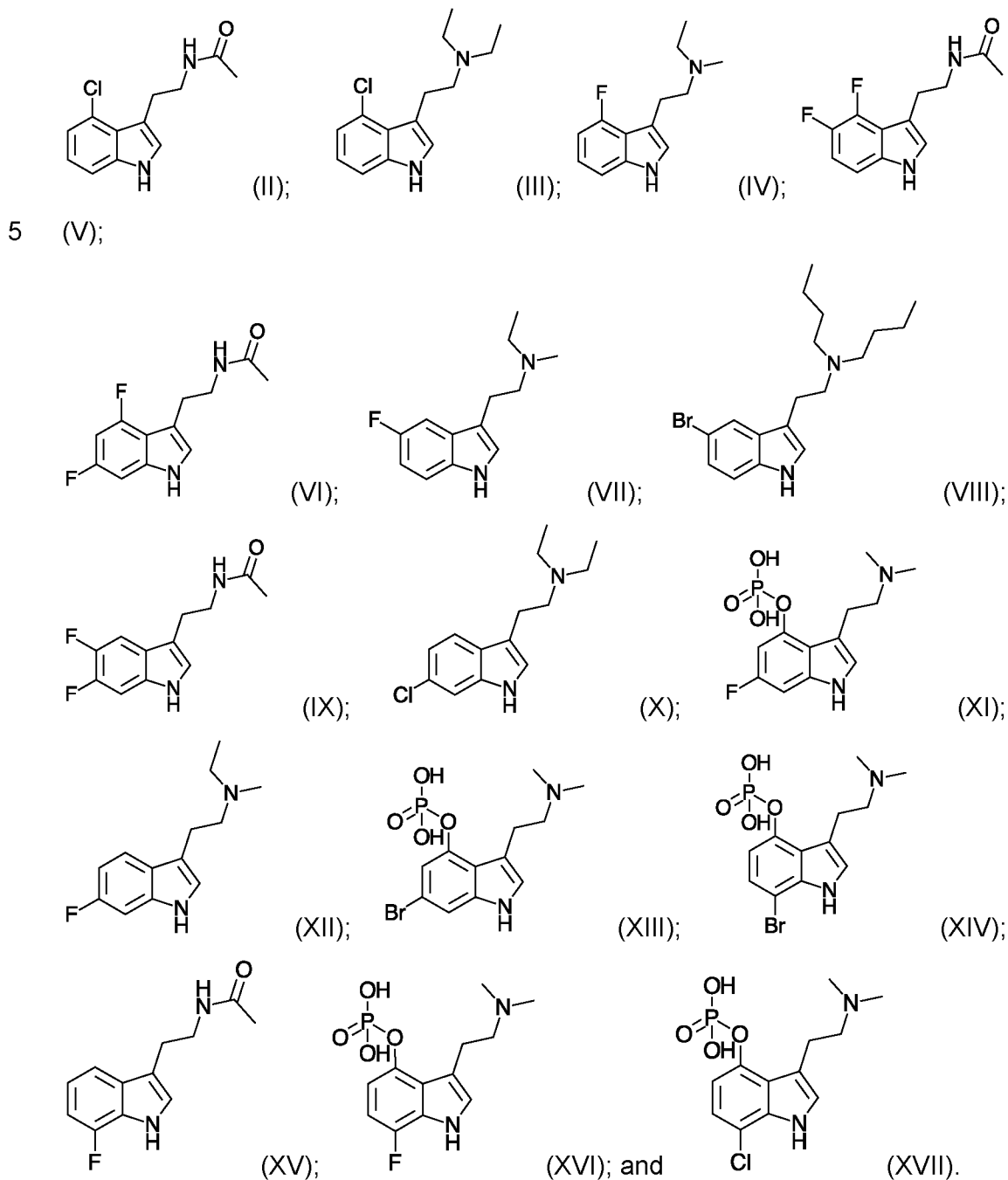
**41.** A method according to claim **38** or **39**, wherein R<sub>4</sub> is a halogen atom and  
35 R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> are a hydrogen atom or an alkyl group or O-alkyl group.

- 42.** A method according to claim **38** or **39**, wherein R<sub>5</sub> is a halogen atom and R<sub>2</sub>, R<sub>6</sub> and R<sub>7</sub> each are a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.
- 5 **43.** A method according to claim **38** or **39**, wherein R<sub>6</sub> is a halogen atom and R<sub>2</sub>, R<sub>5</sub> and R<sub>7</sub> each are a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.
- 10 **44.** A method according to claim **38** or **39**, wherein R<sub>7</sub> is a halogen atom and R<sub>2</sub>, R<sub>5</sub> and R<sub>6</sub> each are a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.
- 45.** A method according to claim **38** or **39**, wherein at least two of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> are a halogen atom.
- 15 **46.** A method according to any one of claims **38** to **45**, wherein the at least of two of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> that are halogen atoms are two identical halogen atoms.
- 20 **47.** A method according to any one of claims **38** to **45**, wherein the at least of two of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> that are halogen atoms are two non-identical halogen atoms.
- 48.** A method according to claim **38** or **39**, wherein R<sub>2</sub> and R<sub>4</sub> are a halogen atom, and R<sub>5</sub>, R<sub>6</sub> and R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group.
- 25 **49.** A method according to claim **38** or **39**, wherein R<sub>2</sub> and R<sub>5</sub> are a halogen atom, R<sub>6</sub> and R<sub>7</sub> are a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> is a hydrogen atom, an alkyl group or O-alkyl group or a phosphate group.
- 30 **50.** A method according to claim **38** or **39** wherein R<sub>2</sub> and R<sub>6</sub> are a halogen atom, R<sub>5</sub> and R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> is a hydrogen atom, an alkyl group or O-alkyl group or a phosphate group.
- 35 **51.** A method according to claim **38** or **39**, wherein R<sub>2</sub> and R<sub>7</sub> is a halogen atom, R<sub>5</sub> and R<sub>6</sub> are a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> is a hydrogen atom, an alkyl group or O-alkyl group or a phosphate group.

- 52.** A method according to claim **38** or **39**, wherein R<sub>4</sub> and R<sub>5</sub> is a halogen atom, and R<sub>2</sub>, R<sub>6</sub> and R<sub>7</sub> are a hydrogen atom or an alkyl group or O-alkyl group.
- 5 **53.** A method according to claim **38** or **39**, wherein R<sub>4</sub> and R<sub>6</sub> are a halogen atom, and R<sub>2</sub>, R<sub>5</sub> and R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group.
- 54.** A method according to claim **38** or **39**, wherein R<sub>4</sub> and R<sub>7</sub> are a halogen atom, and R<sub>2</sub>, R<sub>5</sub> and R<sub>6</sub> is a hydrogen atom or an alkyl group or O-alkyl group.
- 10 **55.** A method according to claim **38** or **39**, wherein R<sub>5</sub> and R<sub>6</sub> are a halogen atom, R<sub>2</sub> and R<sub>7</sub> are a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> is a hydrogen atom, or an alkyl group or O-alkyl group or a phosphate group.
- 15 **56.** A method according to claim **38** or **39**, wherein R<sub>5</sub> and R<sub>7</sub> are a halogen atom, R<sub>2</sub> and R<sub>6</sub> are a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> is a hydrogen atom, an alkyl group or O-alkyl group or a phosphate group.
- 57.** A method according to claim **38** or **39**, wherein R<sub>6</sub> and R<sub>7</sub> are a halogen atom, R<sub>2</sub> and R<sub>5</sub> are a hydrogen atom or an alkyl group or O-alkyl group, and R<sub>4</sub> is a hydrogen atom, or an alkyl group or O-alkyl group or a phosphate group.
- 20 **58.** A method according to any one of claims **38** to **57**, wherein R<sub>4</sub> when it is not halogenated is a hydrogen atom.
- 25 **59.** A method according to any one of claims **38** to **57**, wherein R<sub>4</sub> when it is not halogenated is an alkyl group.
- 60.** A method according to any one of claims **38** to **57**, wherein R<sub>4</sub> when it is not halogenated is an O-alkyl group.
- 30 **61.** A method according to any one of claims **38** to **57**, wherein R<sub>4</sub> when it is not halogenated is a phosphate group.
- 35 **62.** A method according to claim **38**, wherein the halogenated psilocybin derivative is a derivative selected from the group consisting of derivatives having



formulas (II); (III); (IV); (V); (VI); (VII); (VIII); (IX); (X); (XI); (XII); (XIII); (XIV); (XV); (XVI); and (XVII):



**63.** A method according to any one of claims **38** to **62**, wherein the halogenated psilocybin precursor compound is selected from the group of halogenated compounds consisting of halogenated tryptophan, halogenated tryptamine,

halogenated 4-hydroxytryptamine, halogenated 4-hydroxy-indole, halogenated 4-hydroxytryptophan, halogenated norbaeocystin, and halogenated baecocystin.

**64.** A method according to any one of claims **38** to **62**, wherein the psilocybin  
5 biosynthetic enzyme complement comprises at least one enzyme encoded by a nucleic acid selected from:

- (a) SEQ.ID NO: 1, SEQ.ID NO: 3, SEQ.ID NO: 5, SEQ.ID NO: 7, SEQ.ID NO: 9, SEQ.ID NO 11, SEQ.ID NO: 36, SEQ.ID NO: 38, SEQ.ID NO: 40, SEQ. ID NO: 42, SEQ.ID NO: 44, SEQ.ID NO: 46, SEQ.ID NO: 48, SEQ.ID  
10 NO: 50, SEQ.ID NO: 52, SEQ.ID NO: 54, SEQ. ID NO: 56, SEQ.ID NO: 58, SEQ.ID NO: 60, SEQ.ID NO: 64, SEQ.ID NO: 66, SEQ.ID NO: 68, SEQ. ID NO: 70, and SEQ.ID NO: 72;
- (b) a nucleic acid sequence that is substantially identical to any one of the nucleic acid sequences of (a);
- 15 (c) a nucleic acid sequence that is substantially identical to any one of the nucleic acid sequences of (a) but for the degeneration of the genetic code;
- (d) a nucleic acid sequence that is complementary to any one of the nucleic acid sequences of (a);
- 20 (e) a nucleic acid sequence encoding a polypeptide having any one of the amino acid sequences set forth in SEQ.ID NO: 2, SEQ.ID NO: 4, SEQ.ID NO: 6, SEQ.ID NO: 8, SEQ.ID NO: 10, SEQ.ID NO 12, SEQ.ID NO: 37, SEQ.ID NO: 39, SEQ.ID NO: 41, SEQ. ID NO: 43, SEQ.ID NO: 45, SEQ.ID NO: 47, SEQ.ID NO: 49, SEQ.ID NO: 51, SEQ.ID NO: 53, SEQ.ID NO: 55, SEQ. ID NO: 57, SEQ.ID NO: 59, SEQ.ID NO: 61, SEQ.ID NO: 65, SEQ.ID NO: 67, SEQ.ID NO: 69, SEQ. ID NO: 71, and SEQ.ID NO: 73;
- 25 (f) a nucleic acid sequence that encodes a functional variant of any one of the amino acid sequences set forth in SEQ.ID NO: 2, SEQ.ID NO: 4, SEQ.ID NO: 6, SEQ.ID NO: 8, SEQ.ID NO: 10, SEQ.ID NO 12, SEQ.ID NO: 37, SEQ.ID NO: 39, SEQ.ID NO: 41, SEQ. ID NO: 43, SEQ.ID NO: 45, SEQ.ID NO: 47, SEQ.ID NO: 49, SEQ.ID NO: 51, SEQ.ID NO: 53, SEQ.ID NO: 55, SEQ. ID NO: 57, SEQ.ID NO: 59, SEQ.ID NO: 61, SEQ.ID NO: 65, SEQ.ID NO: 67, SEQ.ID NO: 69, SEQ. ID NO: 71, and SEQ.ID NO: 73; and
- 30

(g) a nucleic acid sequence that hybridizes under stringent conditions to any one of the nucleic acid sequences set forth in (a), (b), (c), (d), (e) or (f).

**65.** A method according to claim **39**, wherein the halogenase is encoded by a  
5 nucleic acid selected from:

(a) SEQ.ID NO: 13, SEQ.ID NO: 15, SEQ. ID NO: 17, SEQ.ID NO: 19, SEQ. ID NO: 21, SEQ.ID NO: 23, SEQ.ID NO: 25, and SEQ.ID NO: 27;

(b) a nucleic acid sequence that is substantially identical to any one of the nucleic acid sequences of (a);

10 (c) a nucleic acid sequence that is substantially identical to any one of the nucleic acid sequences of (a) but for the degeneration of the genetic code;

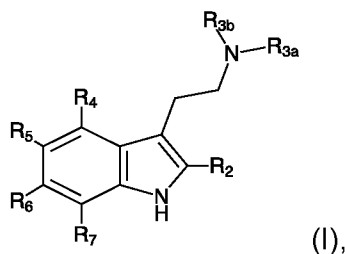
(d) a nucleic acid sequence that is complementary to any one of the nucleic acid sequences of (a);

15 (e) a nucleic acid sequence encoding a polypeptide having any one of the amino acid sequences set forth in SEQ.ID NO: 14, SEQ.ID NO: 16, SEQ. ID NO: 18, SEQ.ID NO: 20, SEQ. ID NO: 22, SEQ.ID NO: 24, SEQ.ID NO: 26, SEQ.ID NO: 28, SEQ.ID NO: 29, SEQ.ID NO: 30, SEQ.ID NO: 31, SEQ.ID NO: 32, SEQ.ID NO: 33, SEQ.ID NO: 34, and SEQ.ID NO: 35;

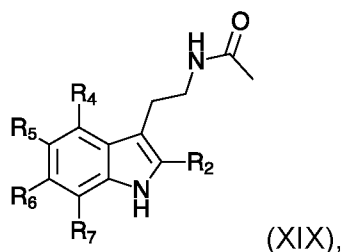
20 (f) a nucleic acid sequence that encodes a functional variant of any one of the amino acid sequences set forth in SEQ.ID NO: 14, SEQ.ID NO: 16, SEQ. ID NO: 18, SEQ.ID NO: 20, SEQ. ID NO: 22, SEQ.ID NO: 24, SEQ.ID NO: 26, SEQ.ID NO: 28, SEQ.ID NO: 29, SEQ.ID NO: 30, SEQ.ID NO: 31, SEQ.ID NO: 32, SEQ.ID NO: 33, SEQ.ID NO: 34, and SEQ.ID NO: 35; and

25 (g) a nucleic acid sequence that hybridizes under stringent conditions to any one of the nucleic acid sequences set forth in (a), (b), (c), (d), (e) or (f).

**66.** A method according to any one of claims **37** to **65**, wherein the host cell further comprises an acetyl transferase capable of acetylating R<sub>3a</sub> or R<sub>3b</sub>, wherein  
30 the cell comprises a compound having chemical formula (I):



wherein, at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> is a halogen atom, wherein each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group, wherein R<sub>4</sub> when it is not halogenated is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group, and wherein at least one of R<sub>3A</sub> and R<sub>3B</sub> is a hydrogen atom, to form a compound having the chemical formula (XIX):



wherein, at least one of R<sub>2</sub>, R<sub>4</sub>, R<sub>5</sub>, R<sub>6</sub> or R<sub>7</sub> is a halogen atom, wherein each non-halogenated R<sub>2</sub>, R<sub>5</sub>, R<sub>6</sub>, or R<sub>7</sub> is a hydrogen atom or an alkyl group or O-alkyl group, wherein R<sub>4</sub> when it is not halogenated is a phosphate group, a hydrogen atom or an alkyl group or O-alkyl group.

**67.** A method according to claim **66**, wherein, the acetylase is encoded by a nucleic acid selected from:

- (a) SEQ.ID NO: 62;
- (b) a nucleic acid sequence that is substantially identical the nucleic acid sequence of (a);
- (c) a nucleic acid sequence that is substantially identical to the nucleic acid sequence of (a) but for the degeneration of the genetic code;
- (d) a nucleic acid sequence that is complementary to the nucleic acid sequence of (a);
- (e) a nucleic acid sequence encoding a polypeptide having an amino acid sequences set forth in SEQ.ID NO: 63;

(f) a nucleic acid sequence that encodes a functional variant of the amino acid sequence set forth in SEQ.ID NO: 63; and

(g) a nucleic acid sequence that hybridizes under stringent conditions to any one of the nucleic acid sequences set forth in (a), (b), (c), (d), (e) or (f).

5 **68.** A method according to claim **37**, wherein:

(a) the halogenated psilocybin precursor compound is a halogenated indole, and the a psilocybin biosynthetic enzyme complement comprises

10 (i) a nucleic acid encoding TrpB selected from SEQ.ID NO: 12, SEQ.ID NO: 65, SEQ.ID NO: 67, SEQ.ID NO: 69, SEQ.ID NO: 71, and SEQ.ID NO: 73, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; (ii) a nucleic acid encoding PsiD selected from SEQ.ID NO: 2, SEQ.ID NO: 10, SEQ.ID NO: 37, SEQ.ID NO: 39, SEQ.ID NO: 41, SEQ.ID NO: 43, SEQ.ID NO: 45, SEQ.ID NO: 47, SEQ.ID NO: 49, SEQ.ID NO: 51, SEQ.ID NO: 53, SEQ.ID NO: 55, and  
15 SEQ.ID NO: 57, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; (iii) a nucleic acid encoding PsiK encoded by SEQ.ID NO: 6, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; and (iv) a nucleic acid encoding PsiM selected from SEQ.ID NO: 8, SEQ.ID NO: 59 and  
20 SEQ.ID NO: 61 or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; or

(b) the halogenated psilocybin precursor compound is a halogenated tryptophan, and the psilocybin biosynthetic enzyme complement comprises (i) a nucleic acid encoding PsiD selected from SEQ.ID NO:  
25 2, SEQ.ID NO: 10, SEQ.ID NO: 37, SEQ.ID NO: 39, SEQ.ID NO: 41, SEQ.ID NO: 43, SEQ.ID NO: 45, SEQ.ID NO: 47, SEQ.ID NO: 49, SEQ.ID NO: 51, SEQ.ID NO: 53, SEQ.ID NO: 55, and SEQ.ID NO: 57, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; (ii) a nucleic acid encoding PsiK encoded  
30 by SEQ.ID NO: 6, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; and (iii) a nucleic acid encoding PsiM selected from SEQ.ID NO: 8, SEQ.ID NO: 59 and SEQ.ID NO: 61 or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; or

- 5 (c) the halogenated psilocybin precursor compound is a halogenated tryptamine, and the psilocybin biosynthetic enzyme complement comprises (i) a nucleic acid encoding PsiK encoded by SEQ.ID NO: 6, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; and (ii) and a nucleic acid encoding PsiM selected from SEQ.ID NO: 8, SEQ.ID NO: 59 and SEQ.ID NO: 61 or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; or
- 10 (d) the halogenated psilocybin precursor compound is a halogenated norbaeocystin or a halogenated baecocystin and the psilocybin biosynthetic enzyme complement comprises a nucleic acid encoding PsiM selected from SEQ.ID NO: 8, SEQ.ID NO: 59 and SEQ.ID NO: 61 or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; or
- 15 (e) the halogenated psilocybin precursor compound is a halogenated tryptophan, and the psilocybin biosynthetic enzyme complement comprises (i) a nucleic acid encoding PsiD selected from SEQ.ID NO: 2, SEQ.ID NO: 10, SEQ.ID NO: 37, SEQ.ID NO: 39, SEQ.ID NO: 41, SEQ.ID NO: 43, SEQ.ID NO: 45, SEQ.ID NO: 47, SEQ.ID NO: 49, SEQ.ID NO: 51, SEQ.ID NO: 53, SEQ.ID NO: 55, and SEQ.ID NO: 57, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; (ii) a nucleic acid encoding PsiH encoded by SEQ.ID NO: 4, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; (iii) a nucleic acid encoding PsiK encoded by SEQ.ID NO: 6, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; and (iv) and a nucleic acid encoding PsiM selected from SEQ.ID NO: 8, SEQ.ID NO: 59 and SEQ.ID NO: 61 or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; or
- 20
- 25
- 30 (f) the halogenated psilocybin precursor compound is a halogenated tryptamine, and psilocybin biosynthetic enzyme complement comprises (i) a nucleic acid encoding PsiH encoded by SEQ.ID NO: 4, or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; (ii) a nucleic acid encoding PsiK encoded by SEQ.ID NO: 6, or

5 a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto; and (iii) and a nucleic acid encoding PsIM selected from SEQ.ID NO: 8, SEQ.ID NO: 59 and SEQ.ID NO: 61 or a nucleic acid sequence encoding a polypeptide sequence substantially similar thereto.

**69.** A method according to any one of claims **37** to **68**, wherein the method further includes a step comprising isolating the halogenated psilocybin derivative.

10 **70.** A method according to any one of claims **37** to **69**, wherein the host cell is a microbial cell.

**71.** A method according to any one of claims **37** to **69**, wherein the host cell is a bacterial cell or a yeast cell.

15 **72.** A method for modulating a 5-HT<sub>2A</sub> receptor, the method comprising contacting a 5-HT<sub>2A</sub> receptor with a chemical compound of claim **1** under reaction conditions sufficient to thereby modulate receptor activity.

20 **73.** A method according to claim **72**, wherein the reaction conditions are *in vitro* reaction conditions.

**74.** A method according to claim **72**, wherein the reaction conditions are *in vivo* reaction conditions.

25 **75.** A use of a chemical compound according to claim **1** in the manufacture of a pharmaceutical or recreational drug formulation.

30 **76.** A use according to claim **75**, wherein the manufacture of a pharmaceutical or recreational drug formulation comprises formulating the chemical compound with an excipient, diluent or carrier.

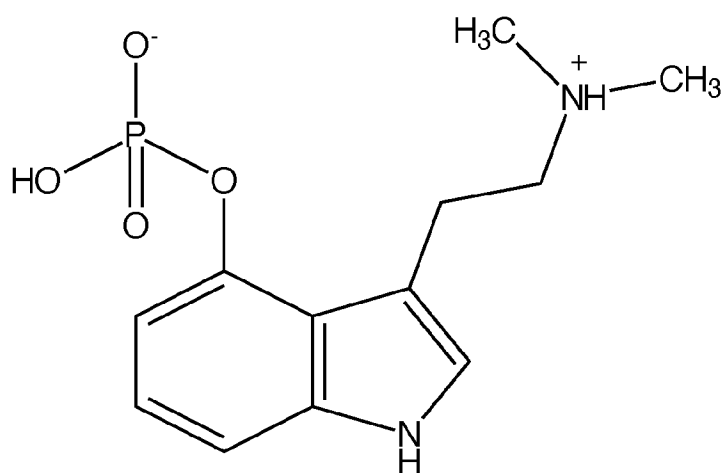
**77.** A use according to claim **75**, wherein the manufacture of a pharmaceutical or recreational drug formulation comprises derivatizing the chemical compound

having the formula (I) by replacing the halogen atom with another atom or chemical group.

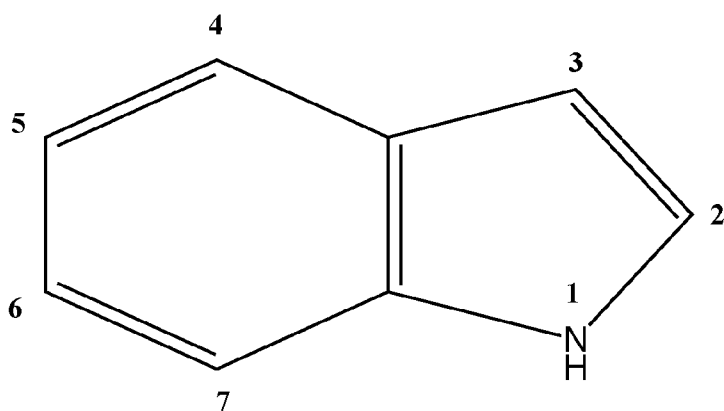
- 78.** A use of a chemical compound of claim **1**, together with a diluent, carrier or  
5 excipient, as a pharmaceutical or recreational drug formulation.

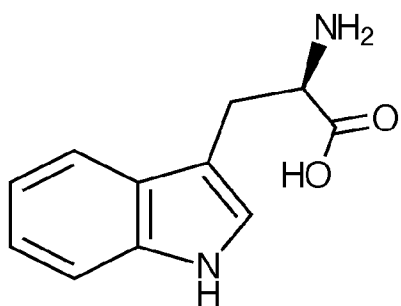
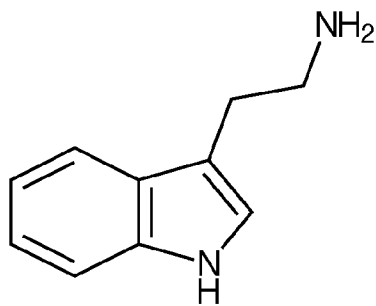
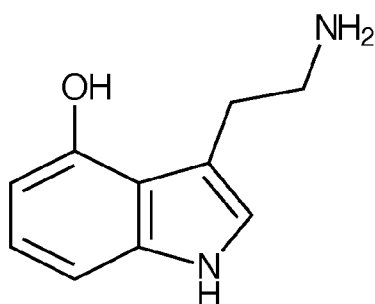
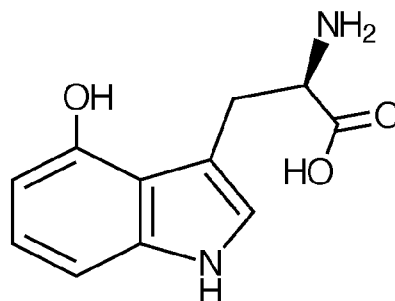
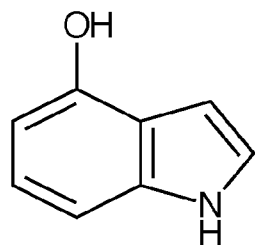
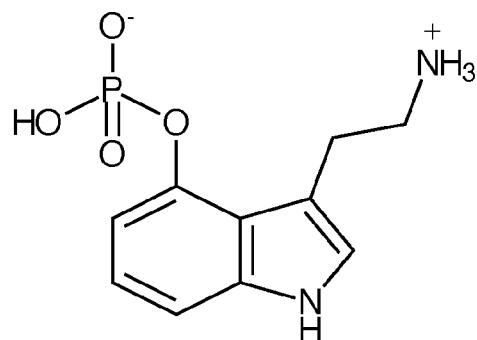
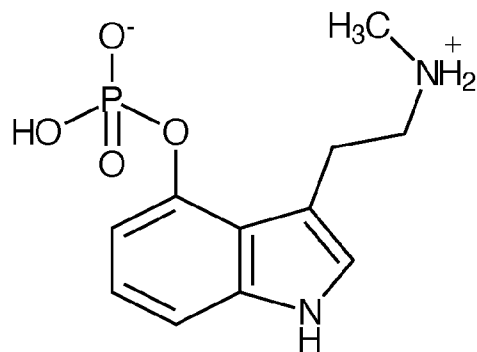


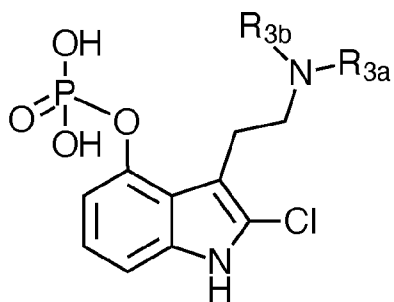
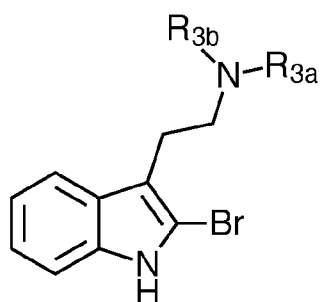
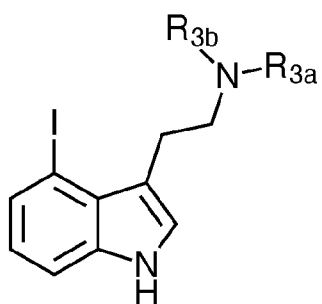
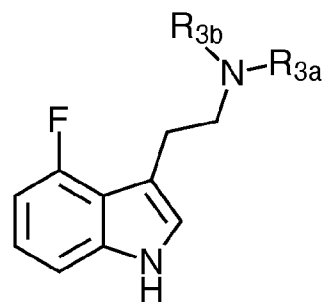
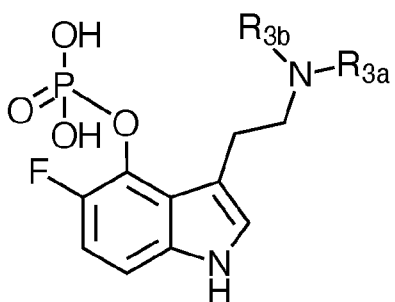
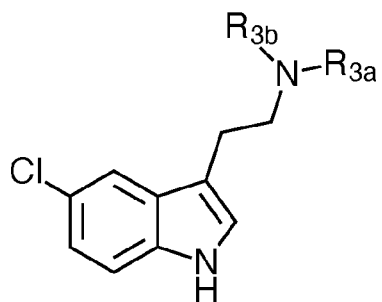
1/36

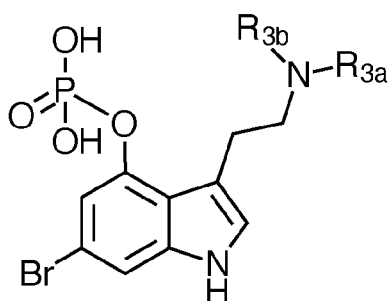
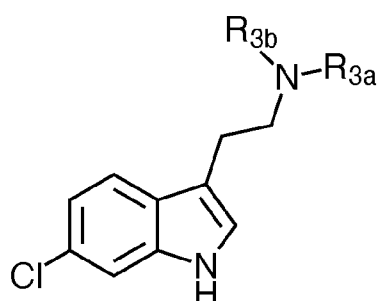
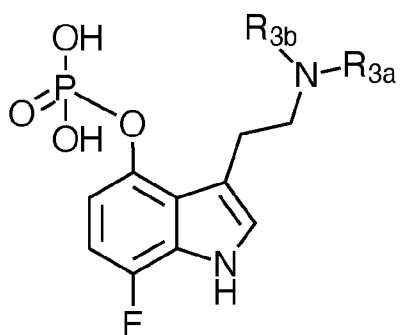
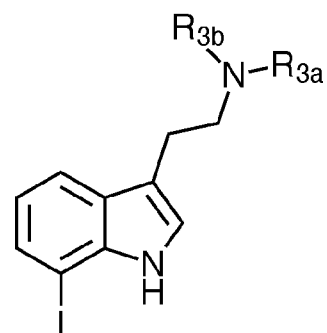
**FIG. 1**

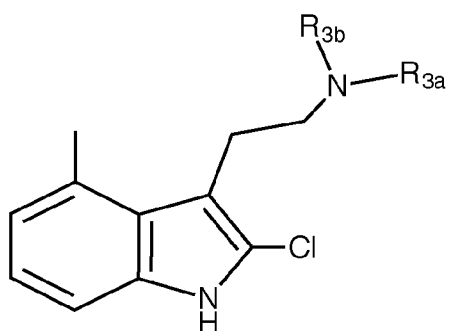
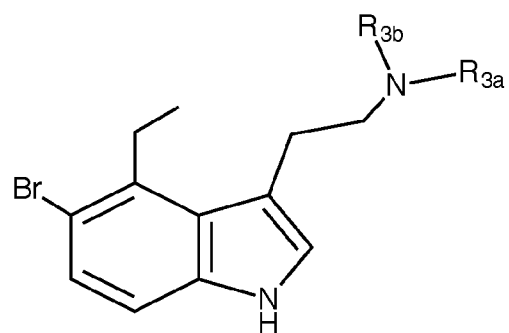
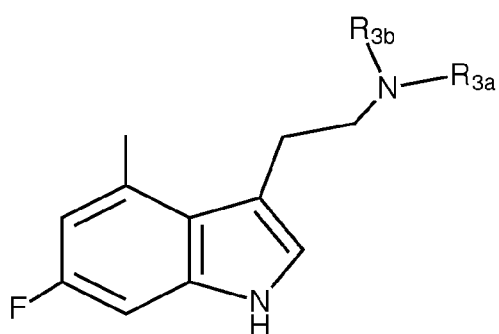
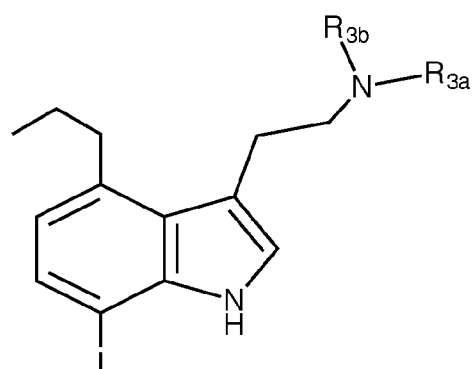
2/36

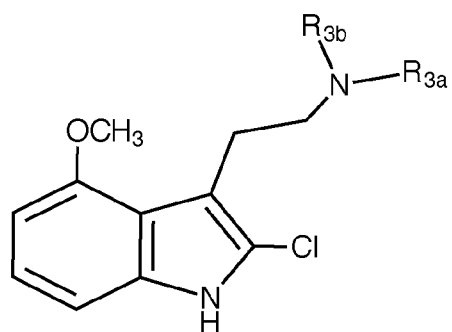
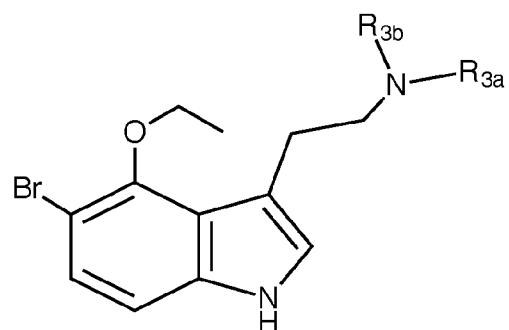
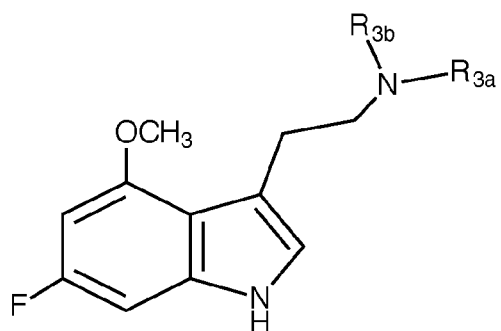
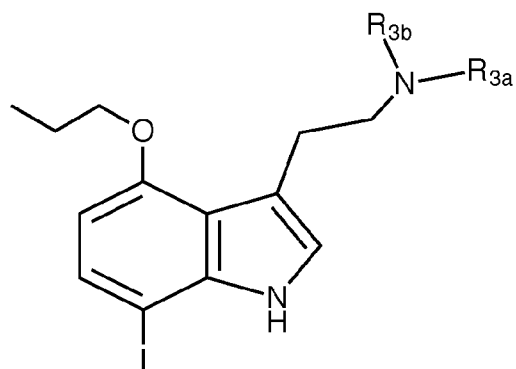
**FIG. 2**

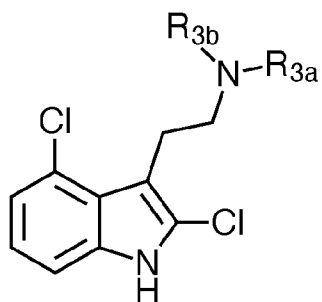
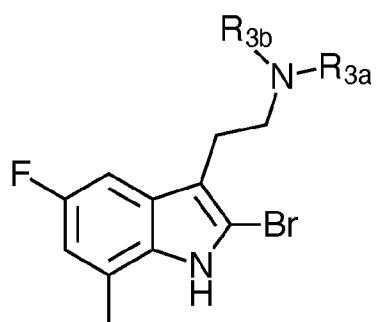
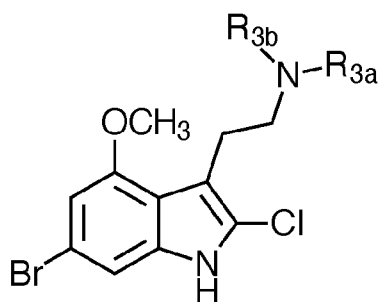
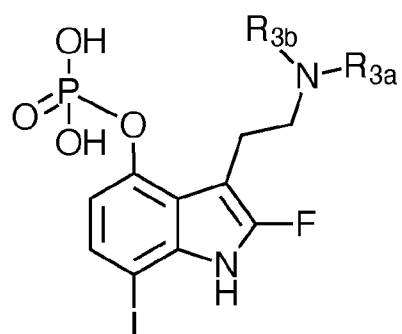
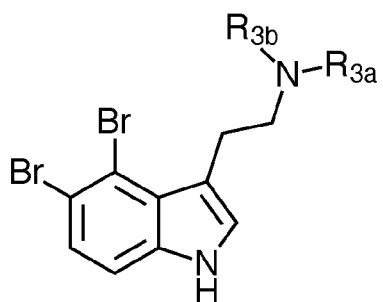
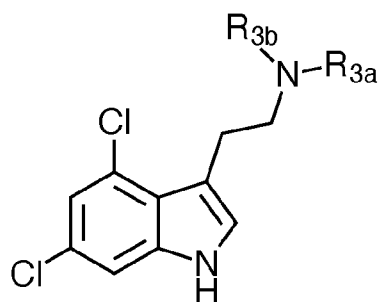
**FIG. 3A****FIG. 3B****FIG. 3C****FIG. 3D****FIG. 3E****FIG. 3F****FIG. 3G**

**FIG. 4A****FIG. 4B****FIG. 4C****FIG. 4D****FIG. 4E****FIG. 4F**

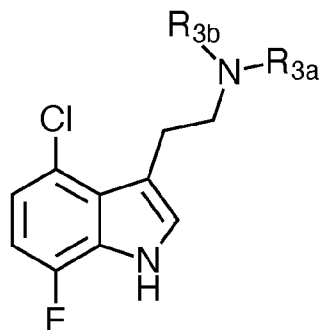
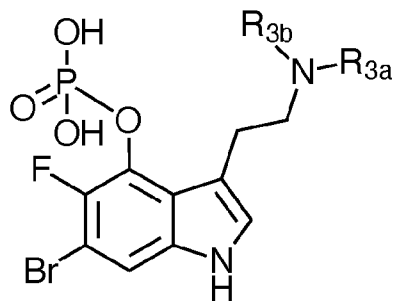
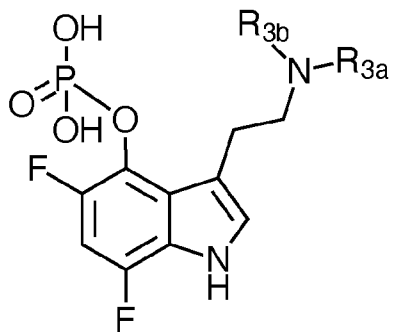
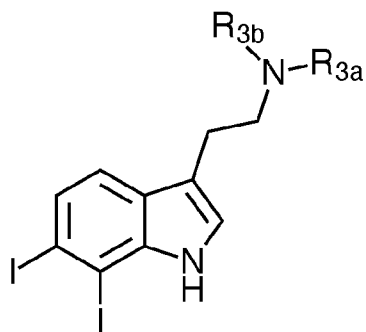
**FIG. 4G****FIG. 4H****FIG. 4I****FIG. 4J**

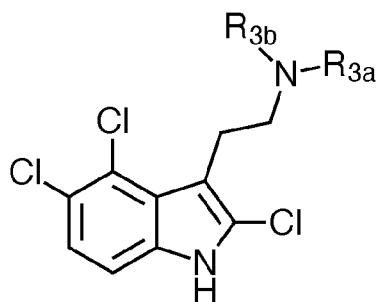
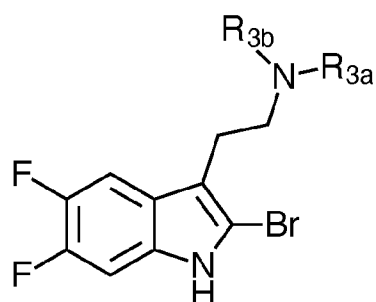
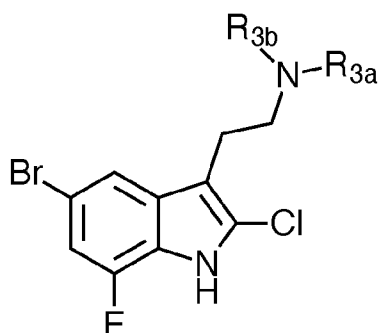
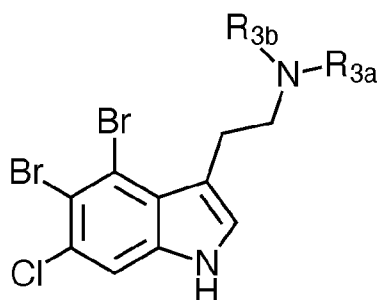
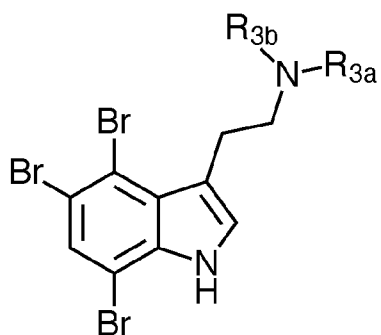
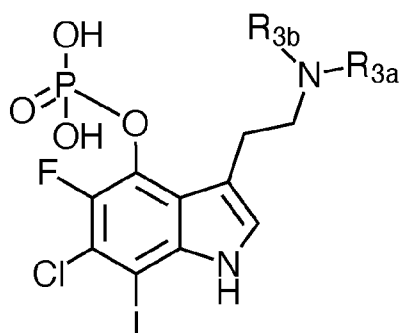
**FIG. 4K****FIG. 4L****FIG. 4M****FIG. 4N**

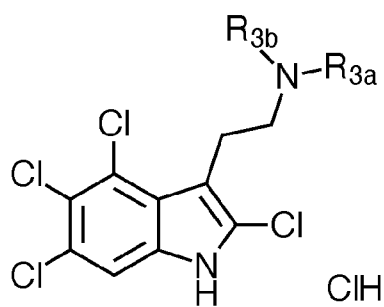
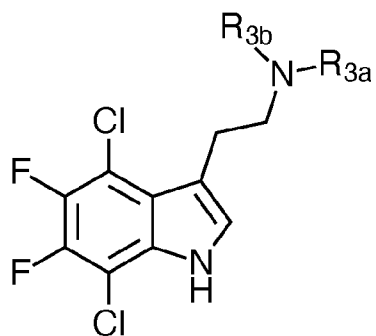
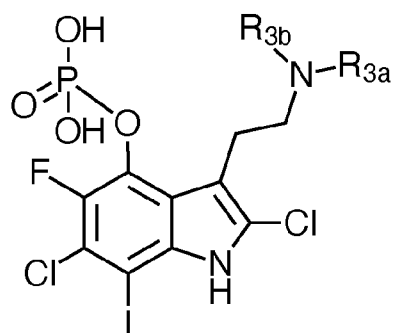
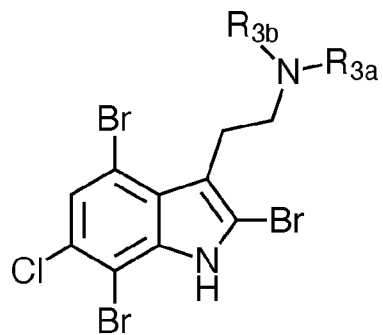
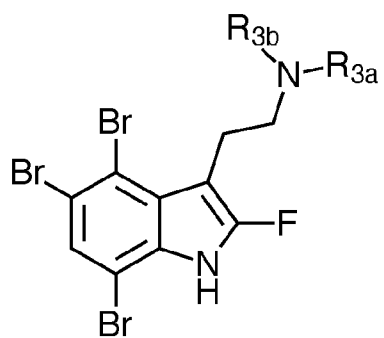
**FIG. 4O****FIG. 4P****FIG. 4Q****FIG. 4R**

**FIG. 5A****FIG. 5B****FIG. 5C****FIG. 5D****FIG. 5E****FIG. 5F**

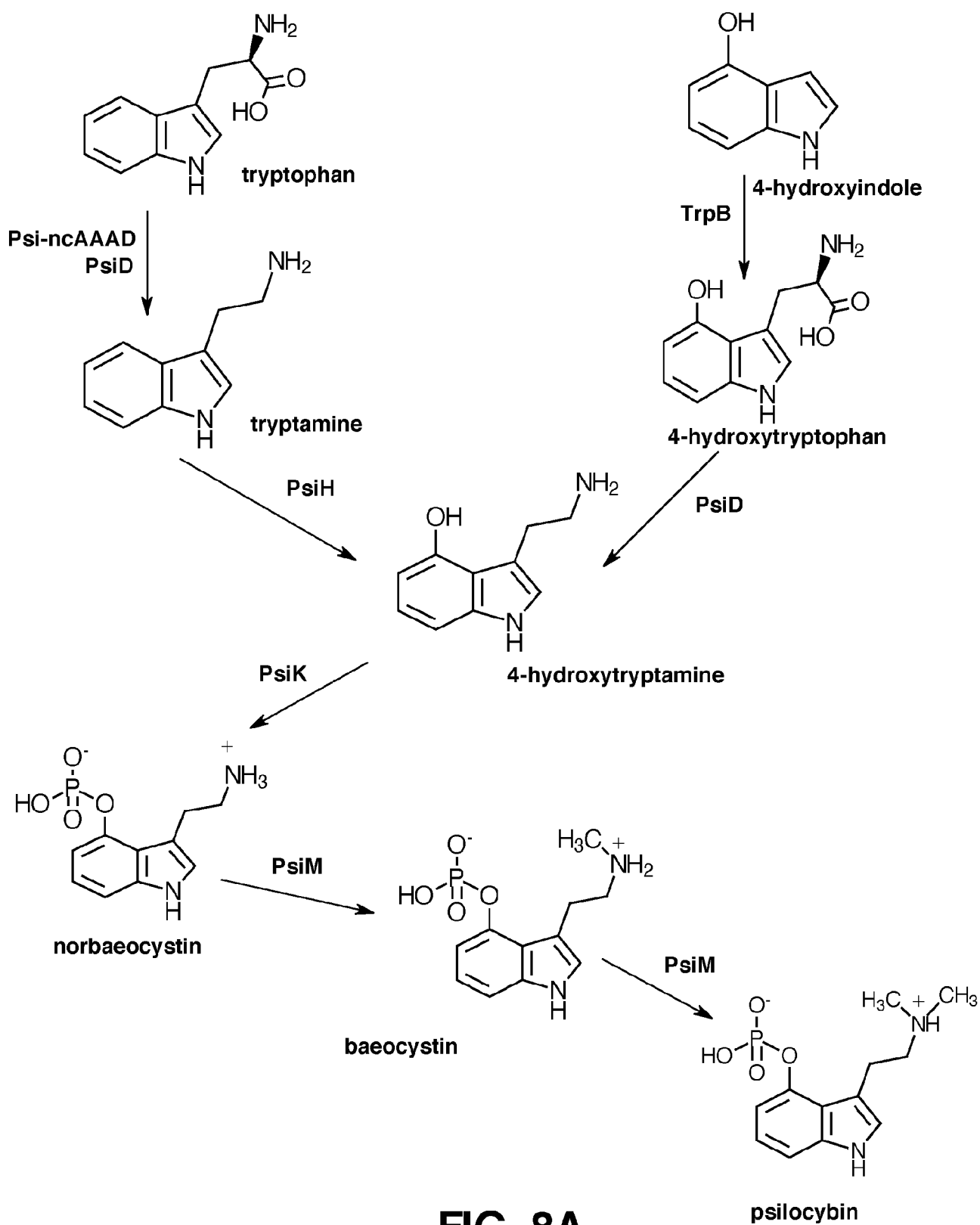


**FIG. 5G****FIG. 5H****FIG. 5I****FIG. 5J**

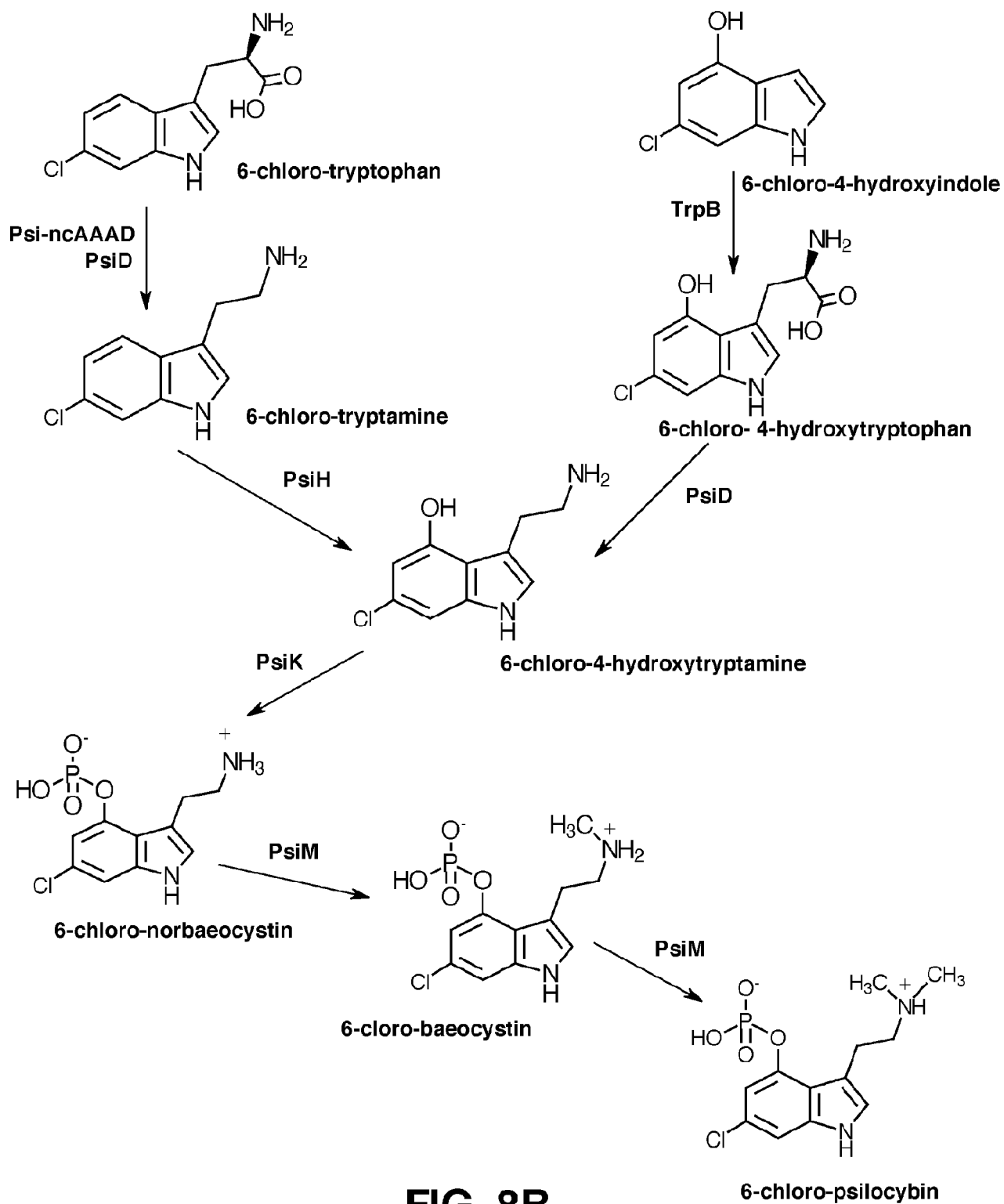
**FIG. 6A****FIG. 6B****FIG. 6C****FIG. 6D****FIG. 6E****FIG. 6F**

**FIG. 7A****FIG. 7B****FIG. 7C****FIG. 7D****FIG. 7E**

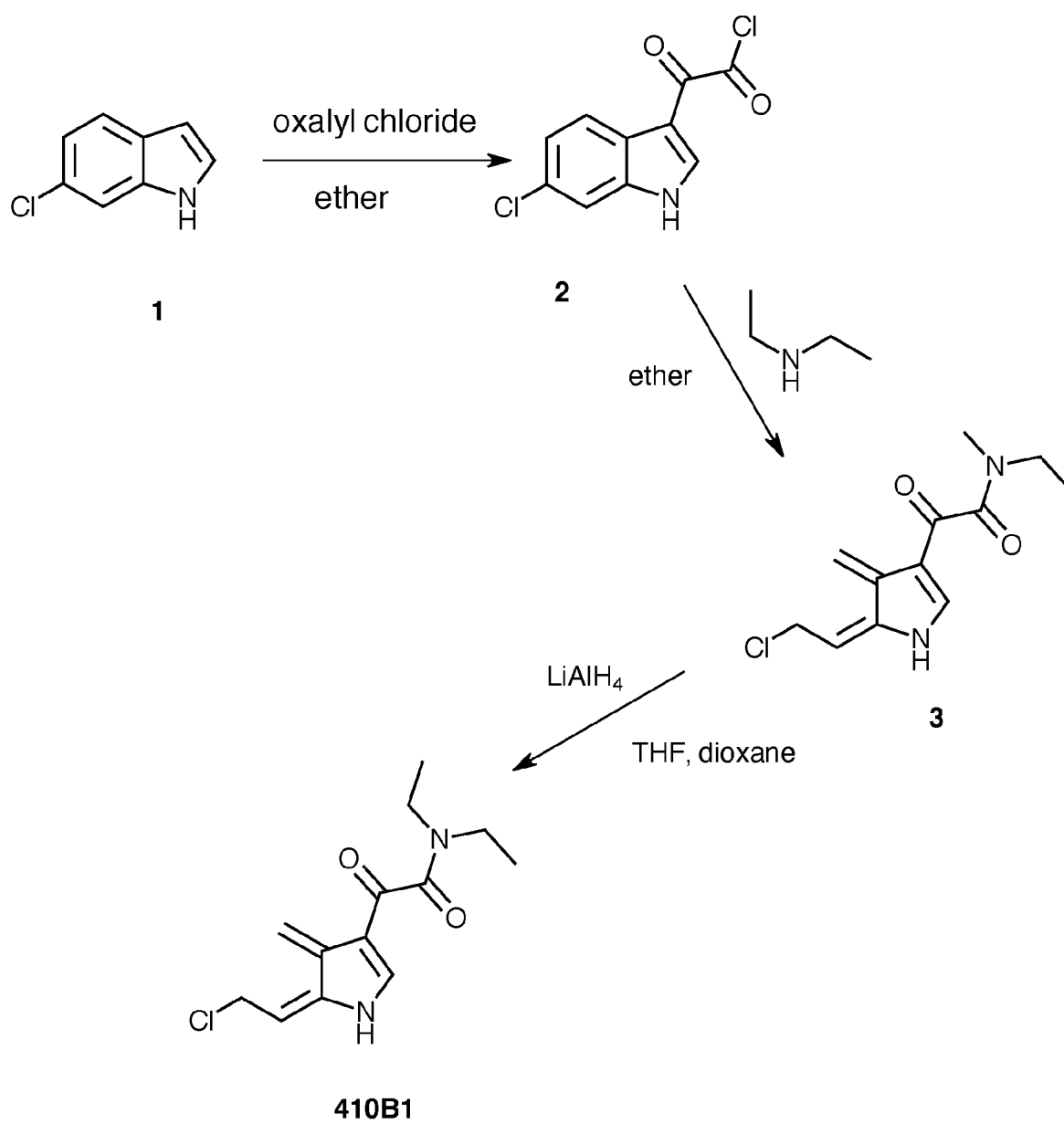
12/36



13/36



14/36

**FIG. 9A**

15/36

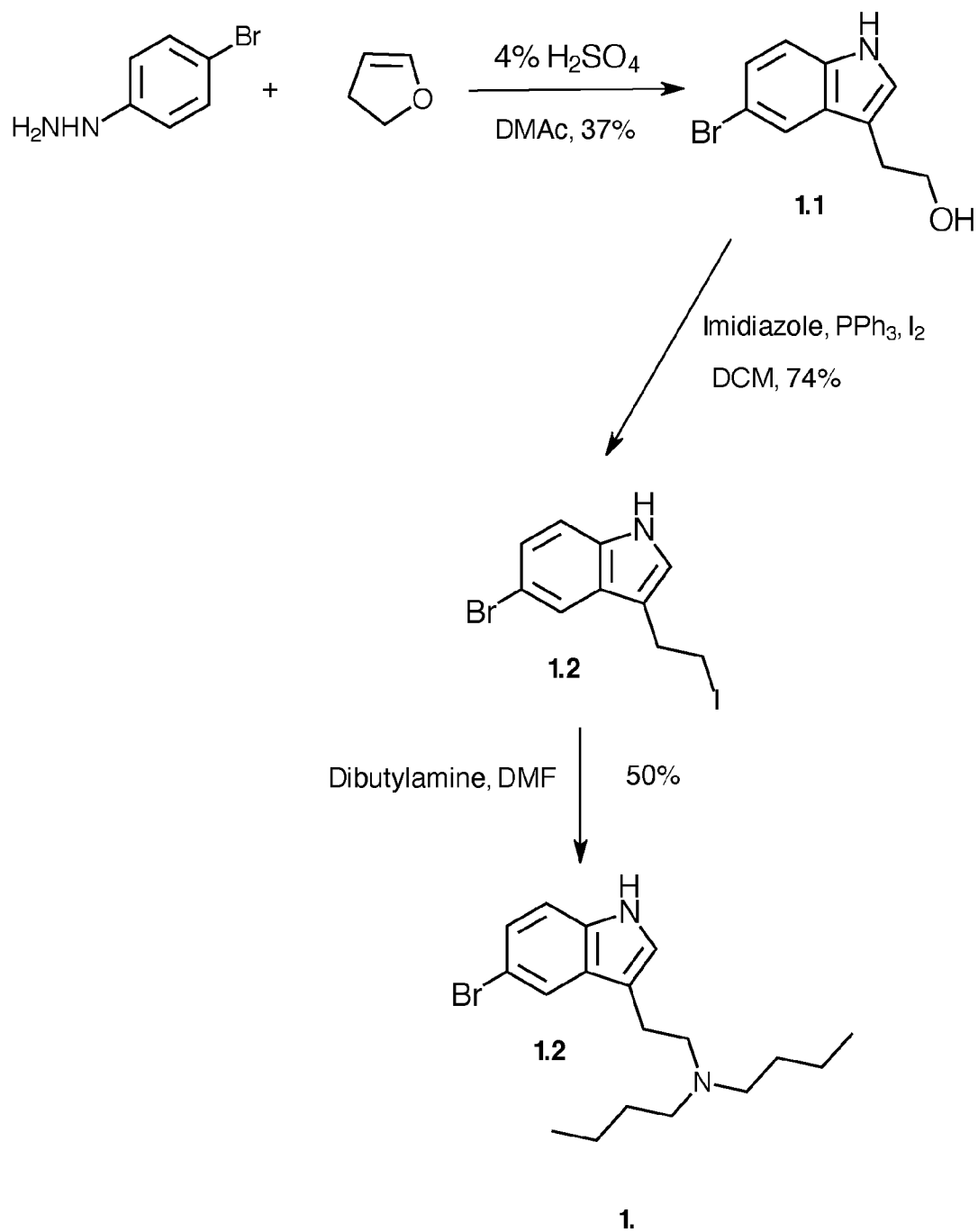
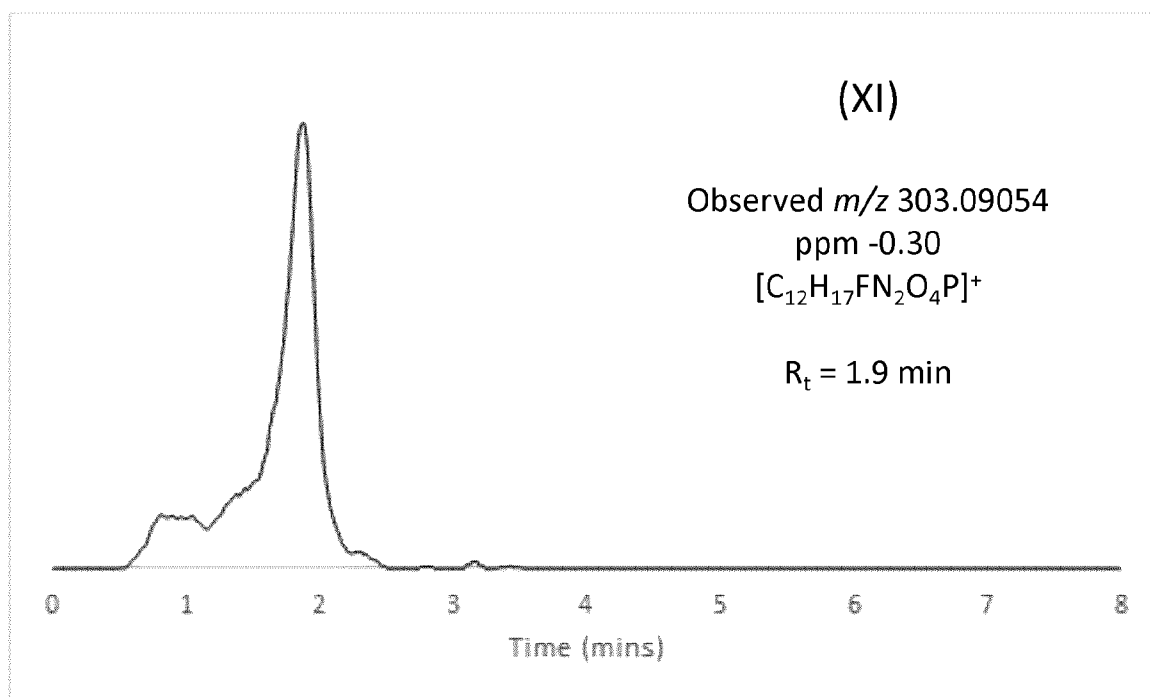


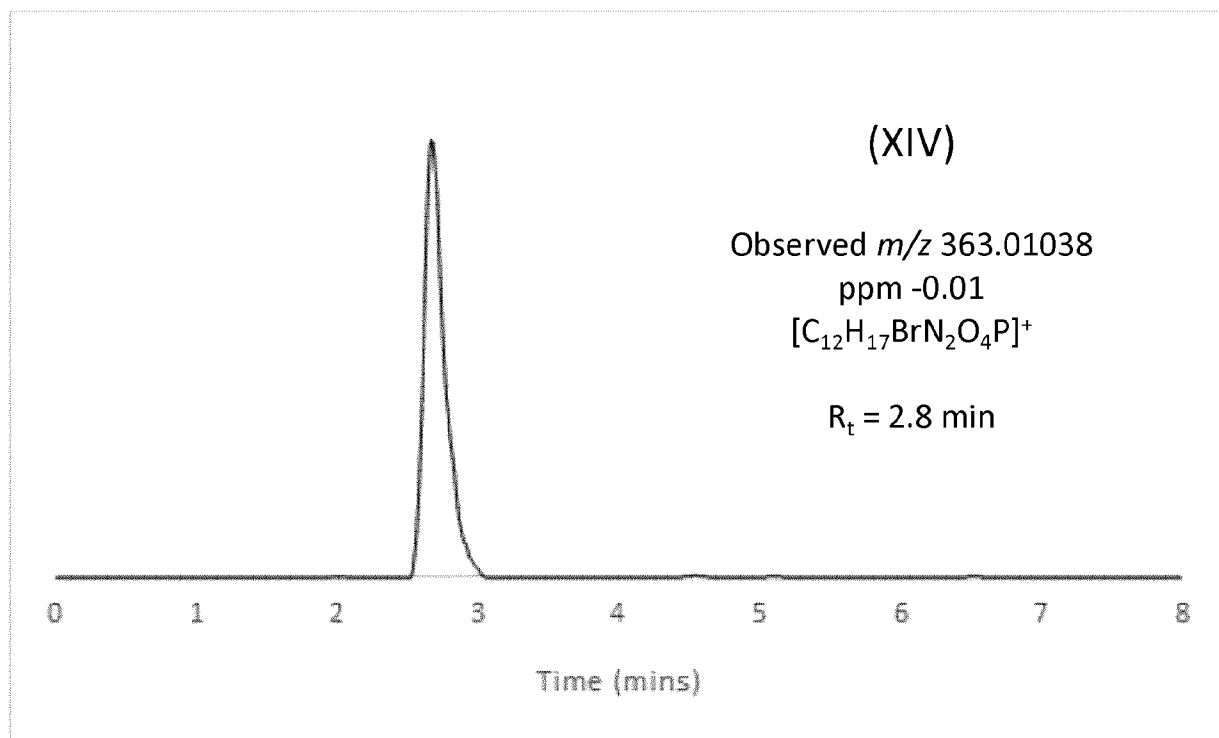
FIG. 9B

16/36

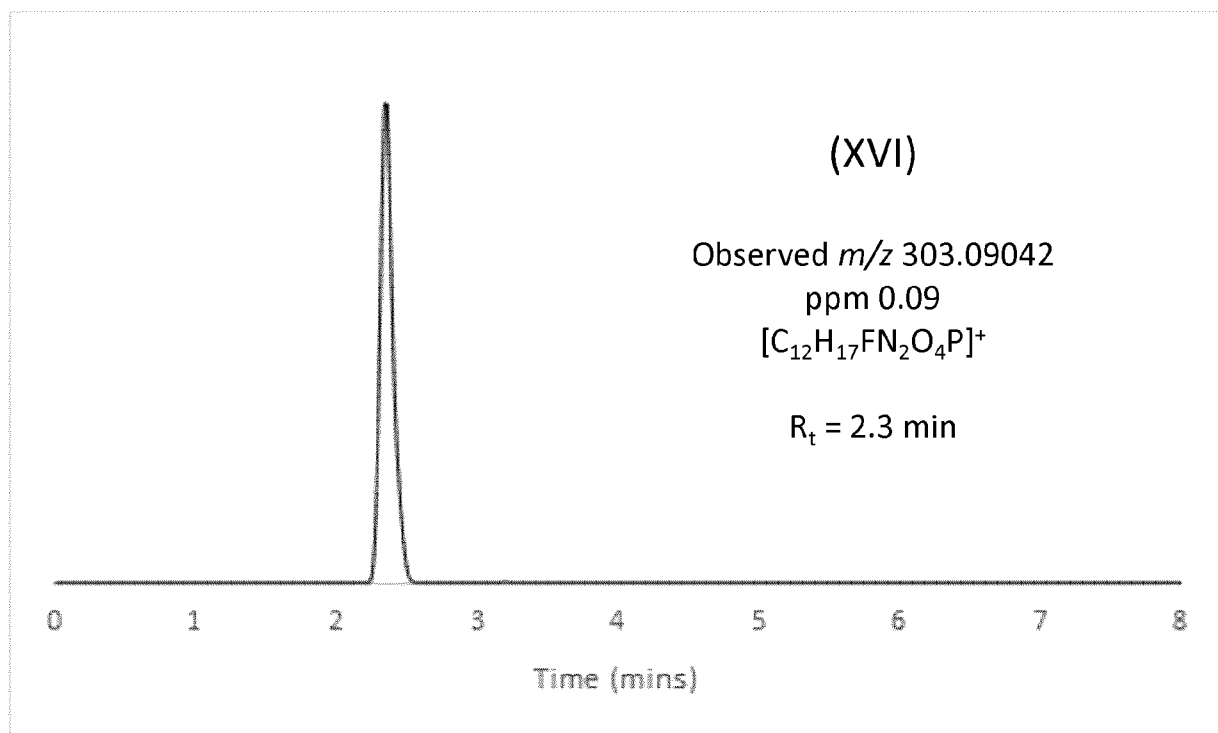
**FIG. 10**



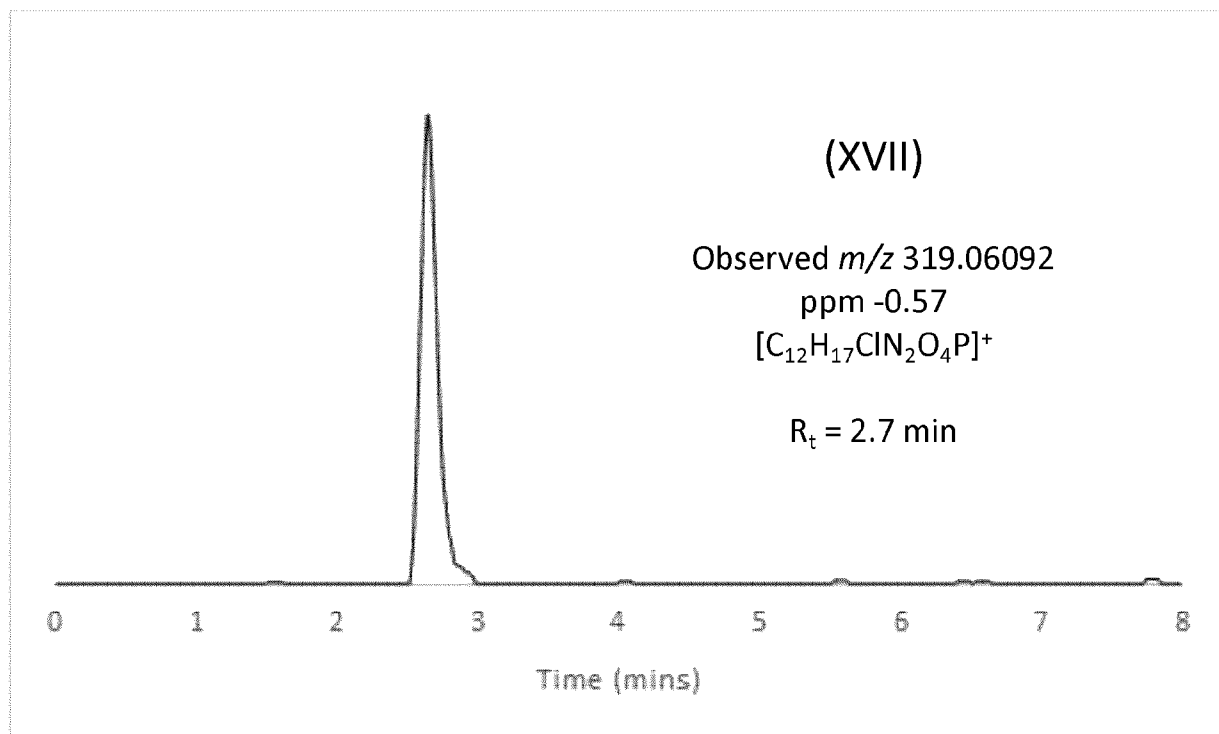
17/36

**FIG. 11**

18/36

**FIG. 12**

19/36

**FIG. 13**

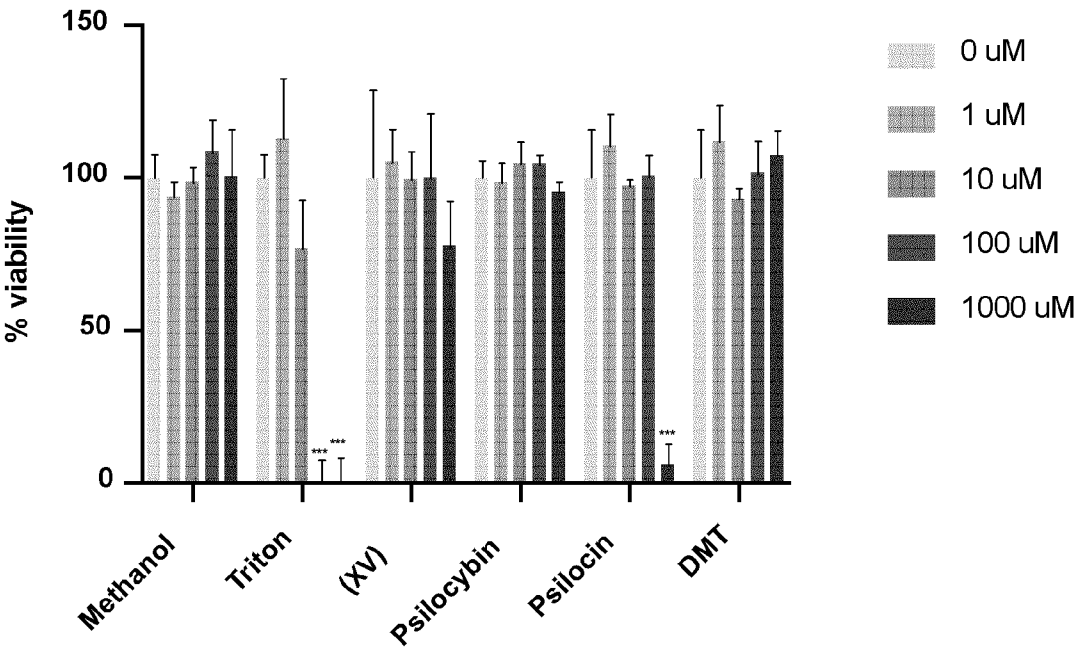


FIG. 14

21/36

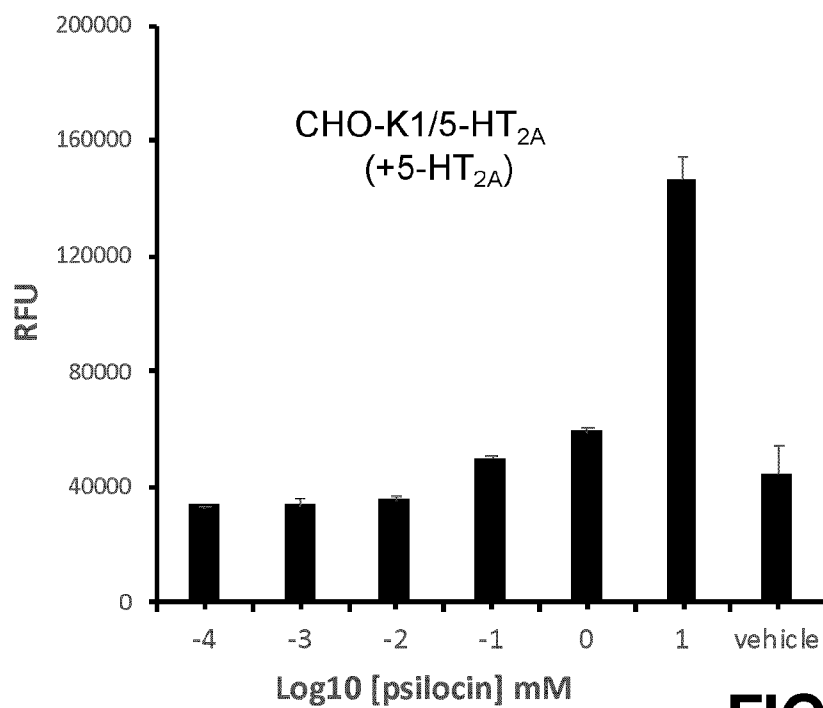


FIG. 15A

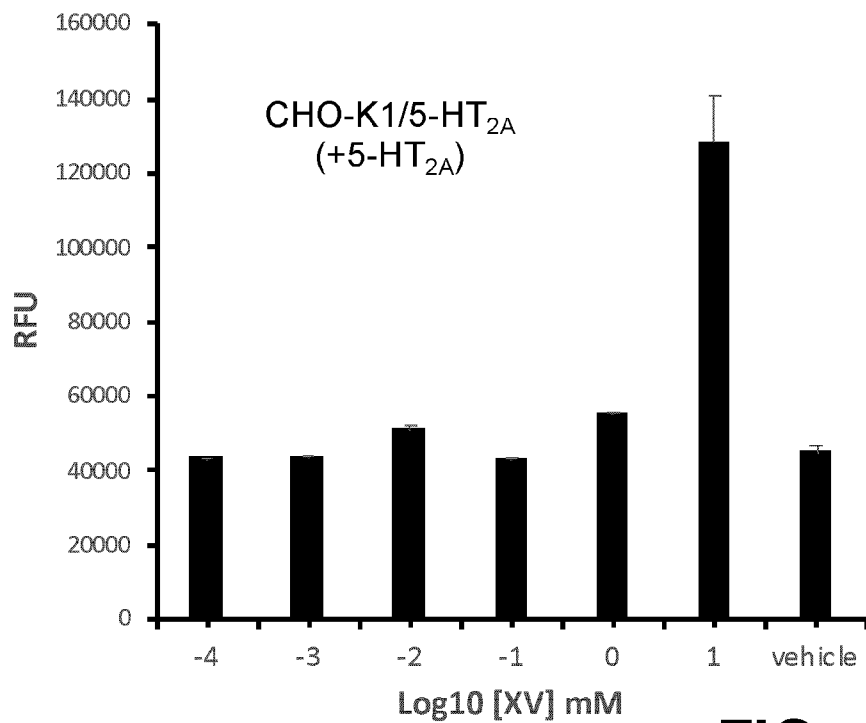
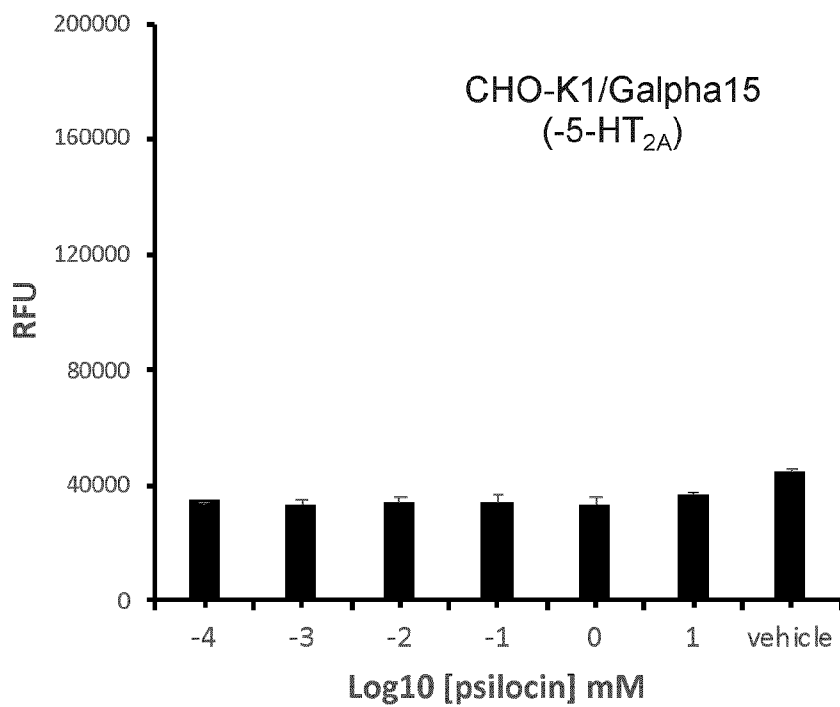
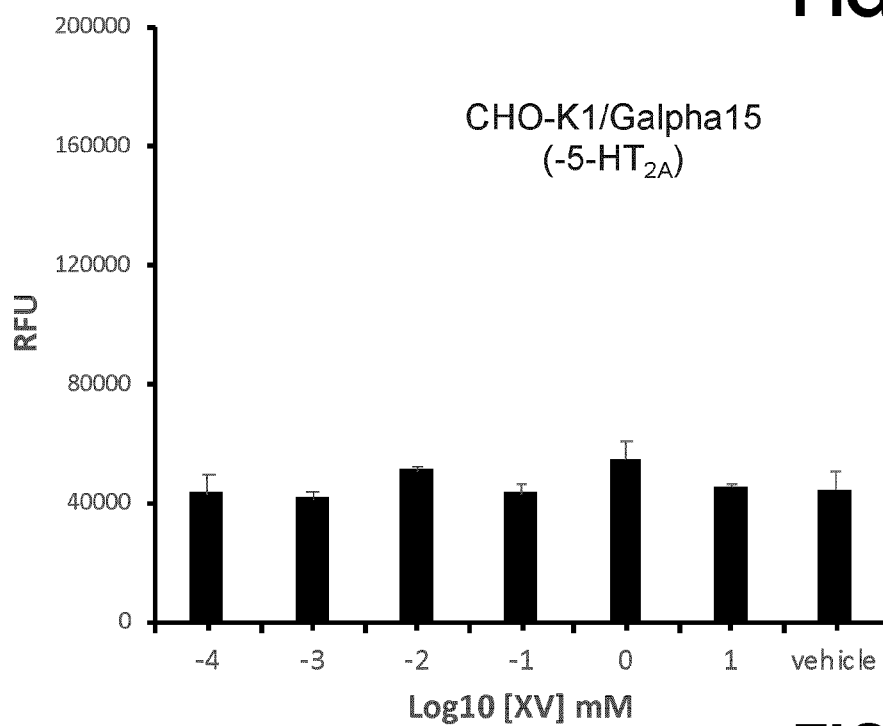
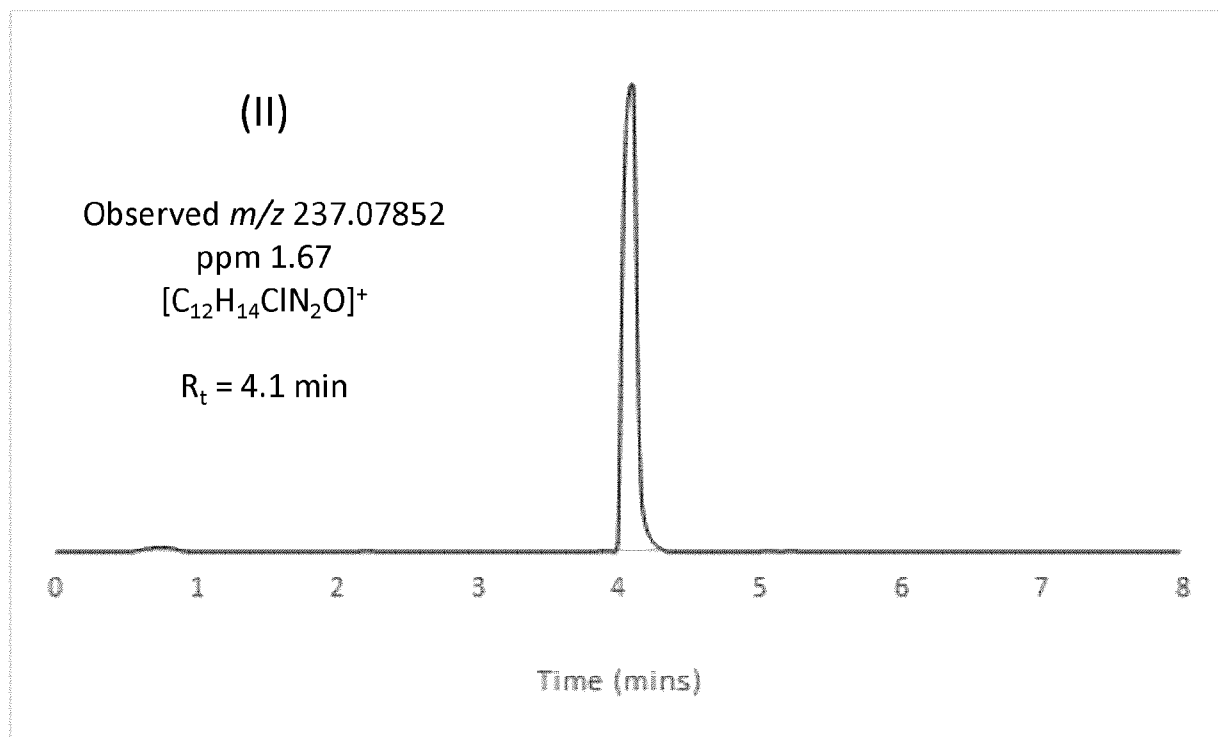


FIG. 15B

22/36

**FIG. 16A****FIG. 16B**

23/36

**FIG. 17**

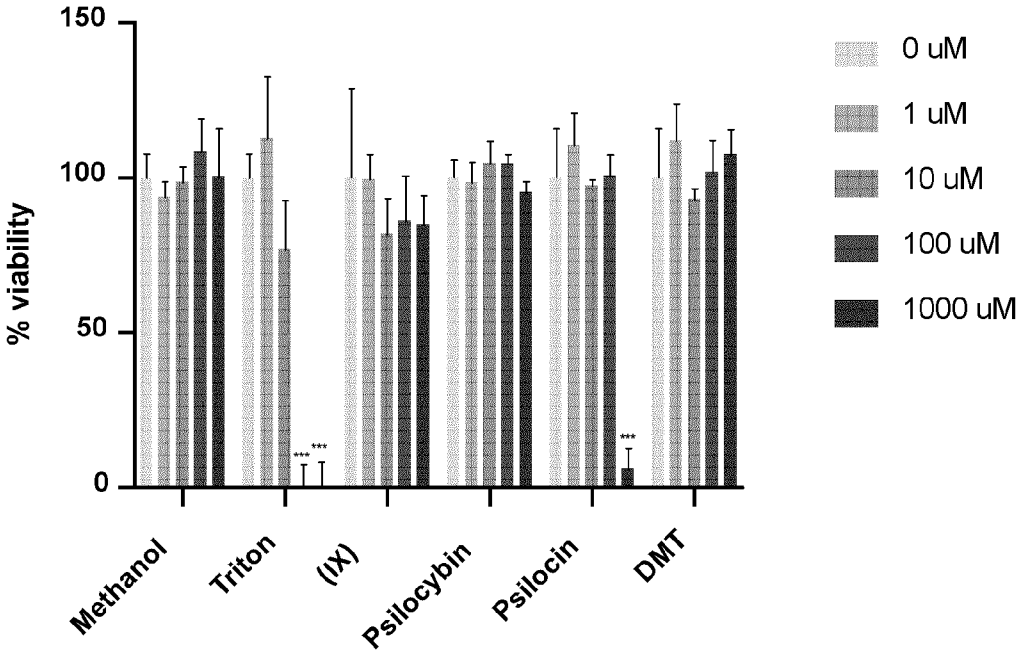


FIG. 18



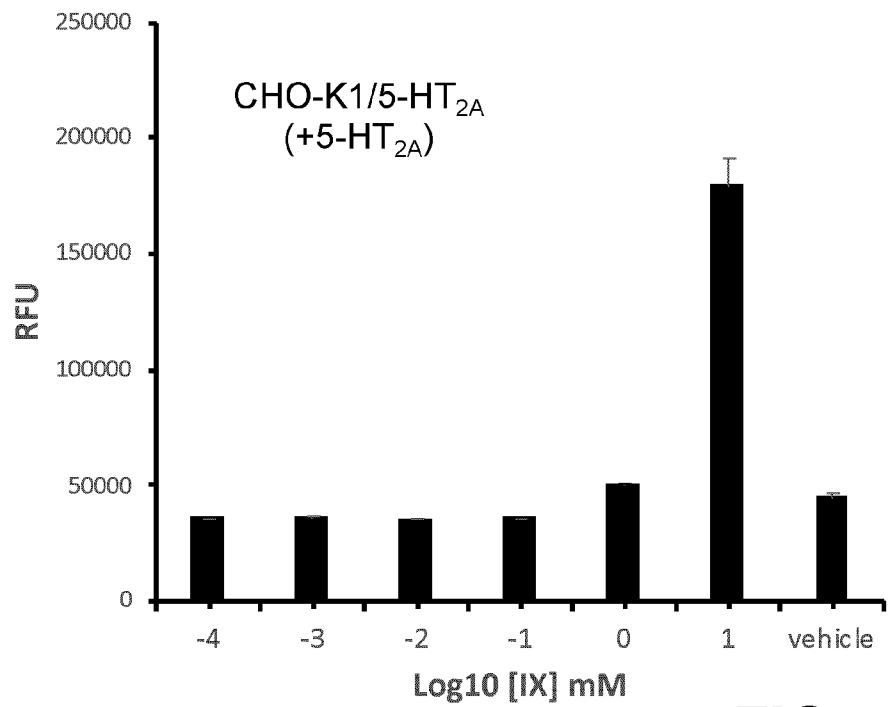


FIG. 19A

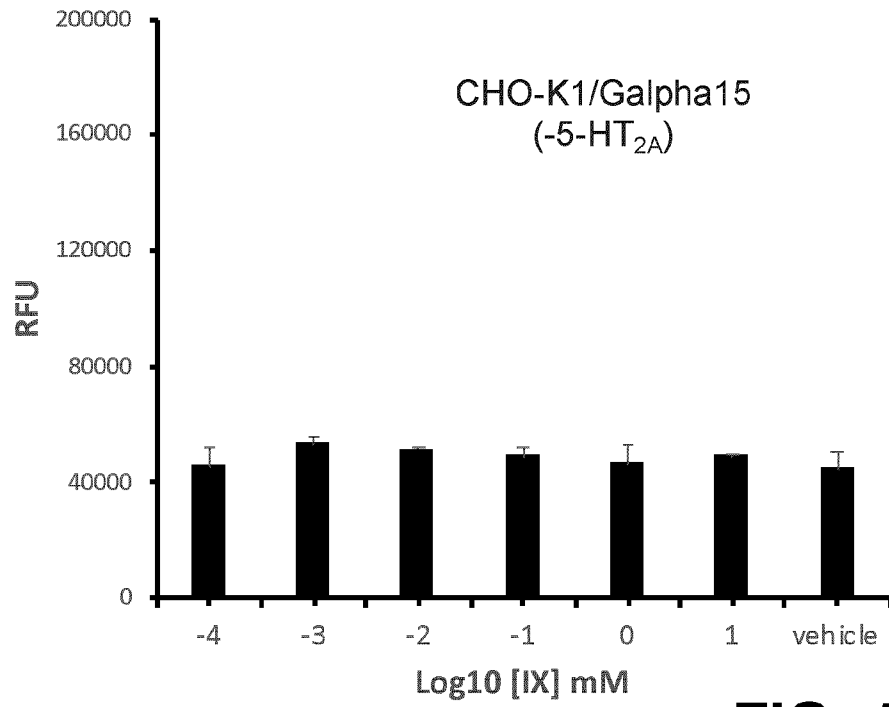
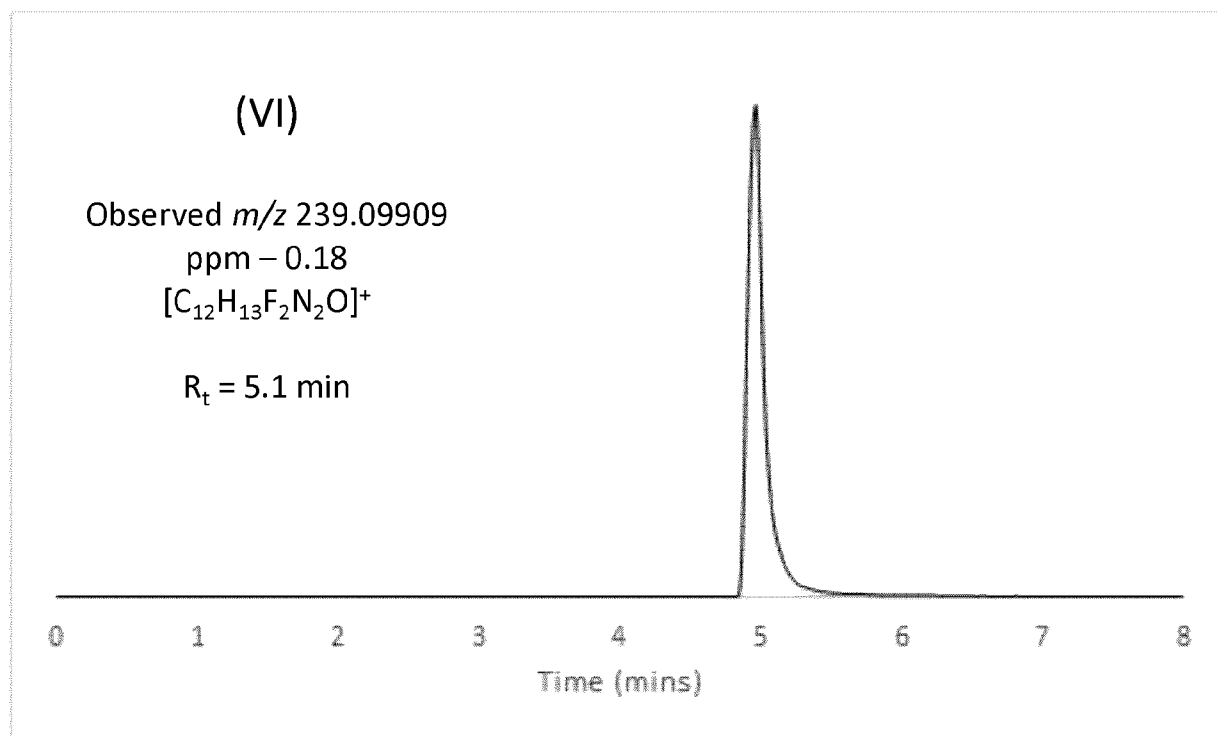


FIG. 19B

26/36

**FIG. 20**

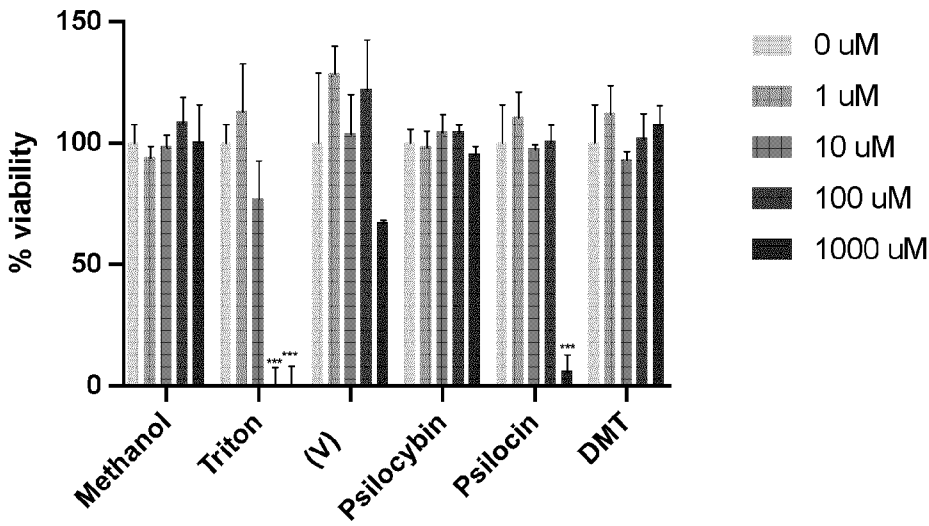


FIG. 21

28/36

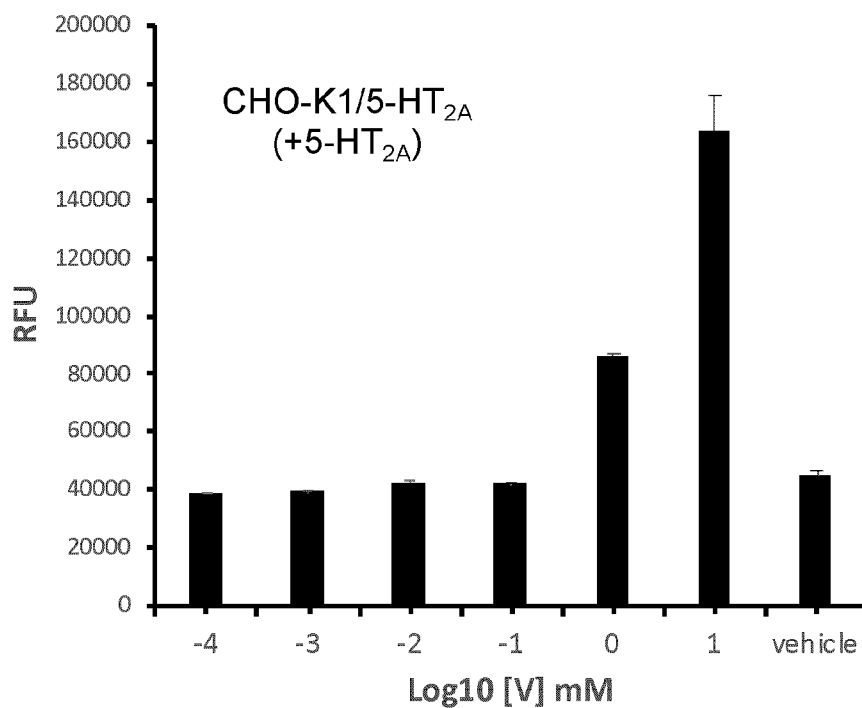


FIG. 22A

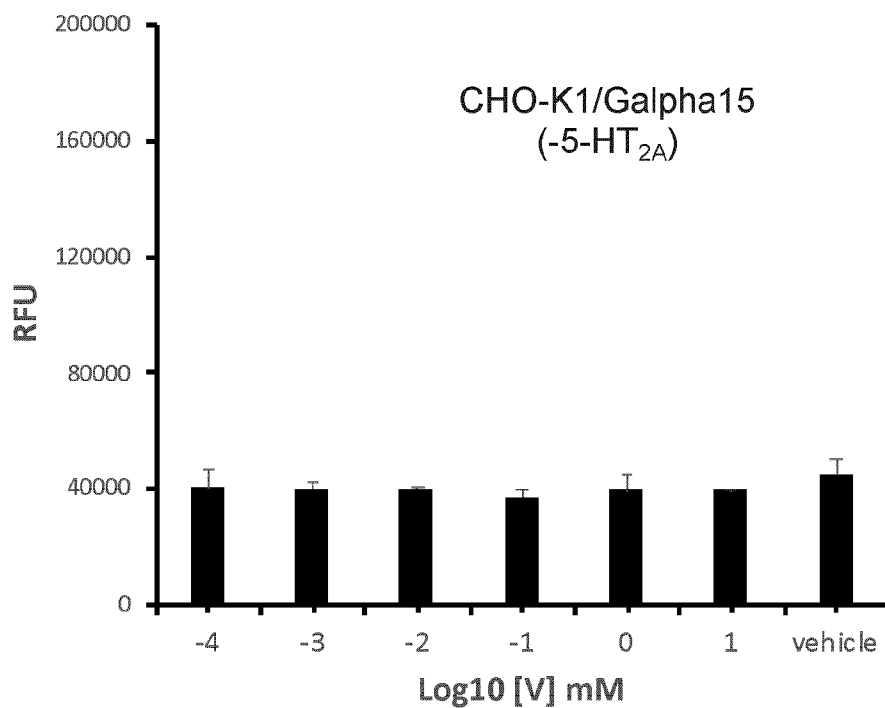
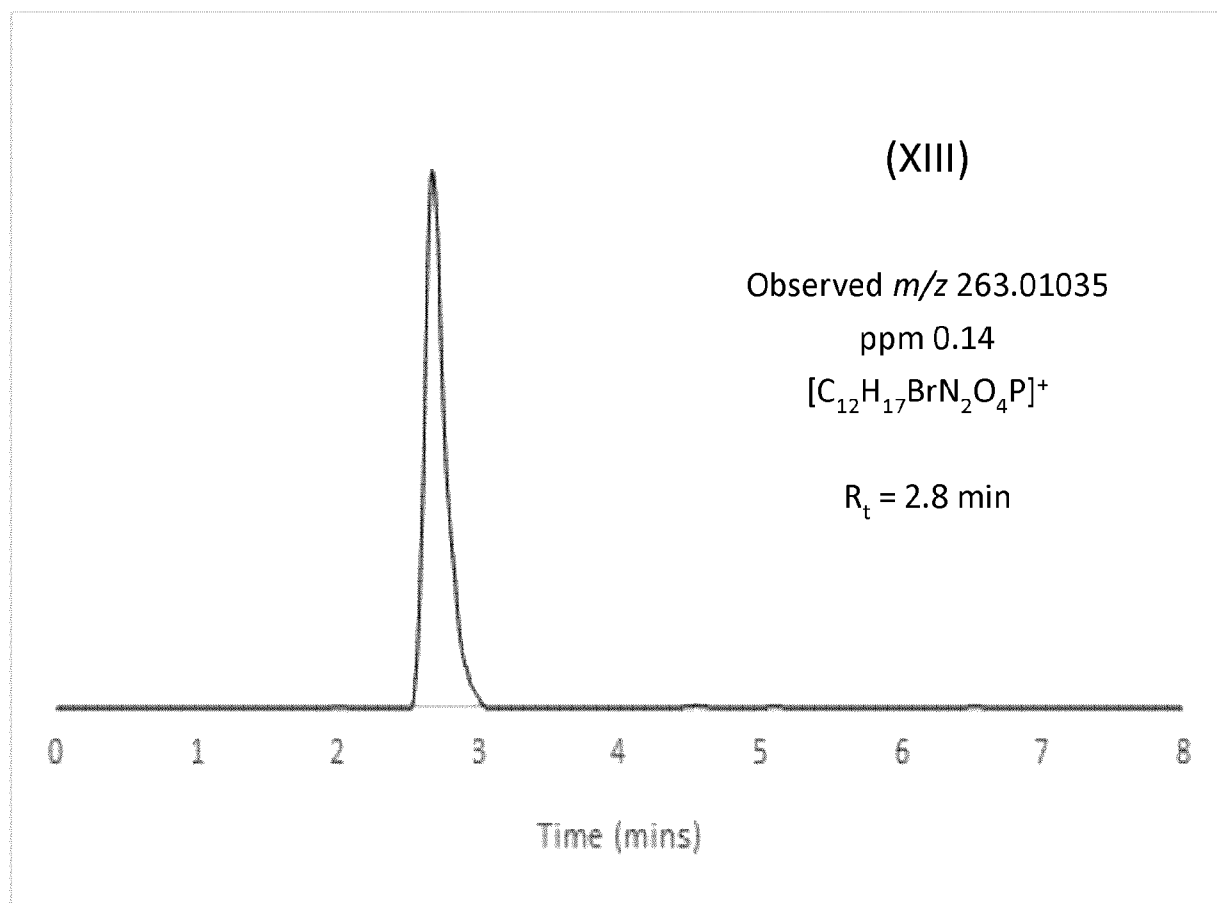


FIG. 22B

29/36

**FIG. 23**

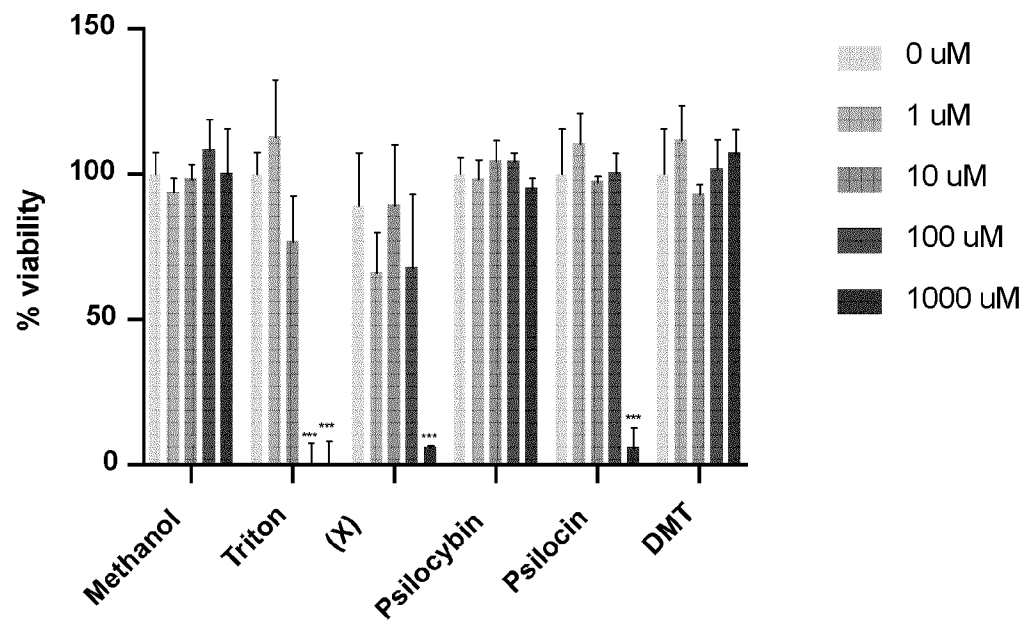


FIG. 24

31/36

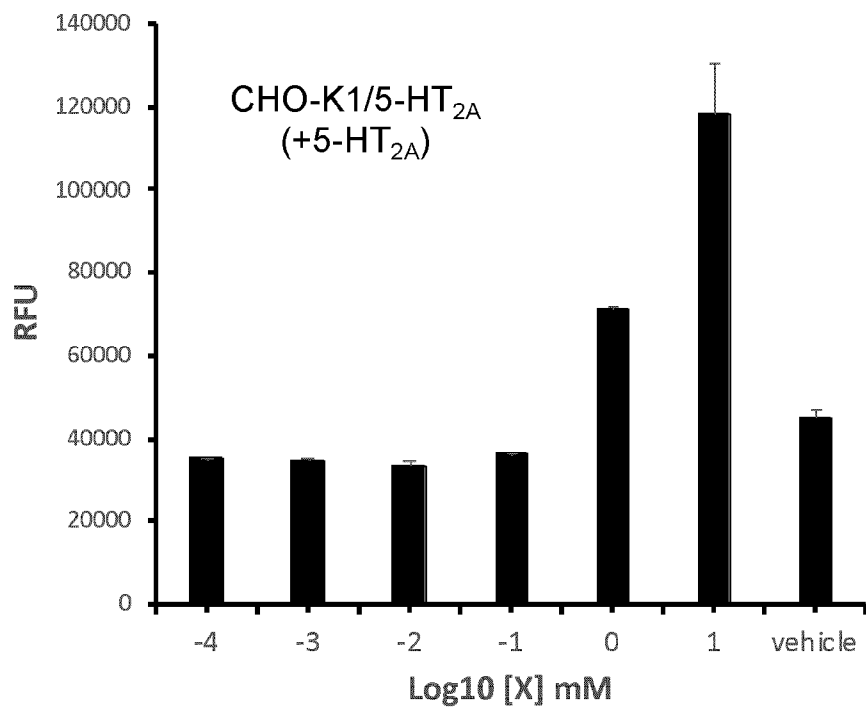


FIG. 25A

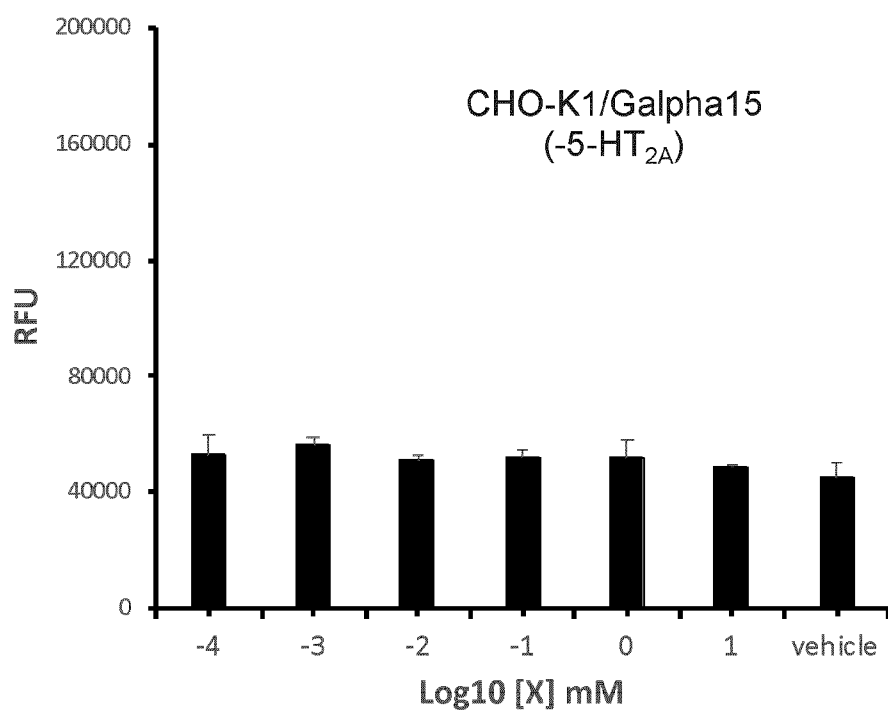


FIG. 25B

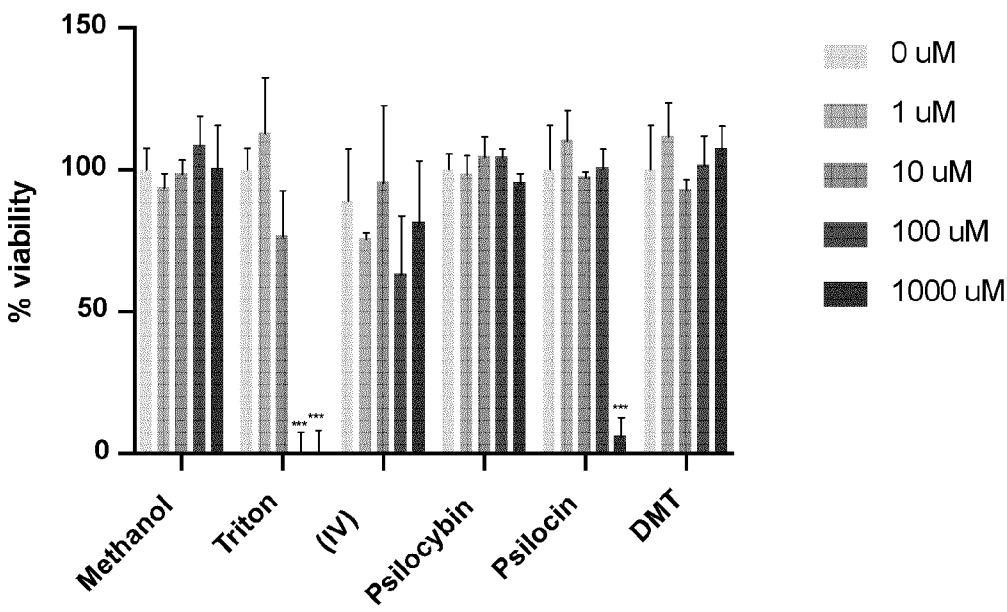


FIG. 26



33/36

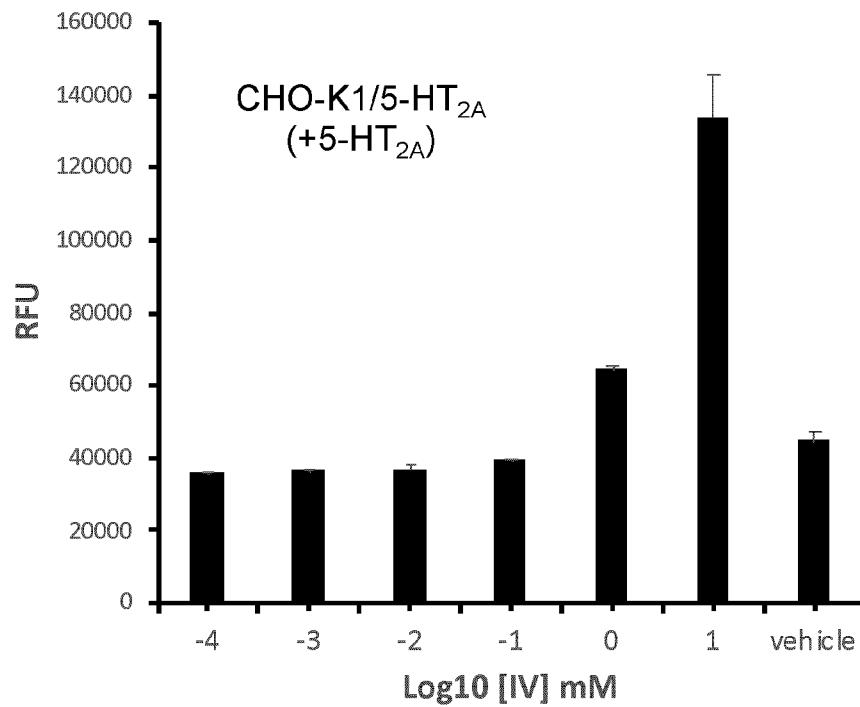


FIG. 27A

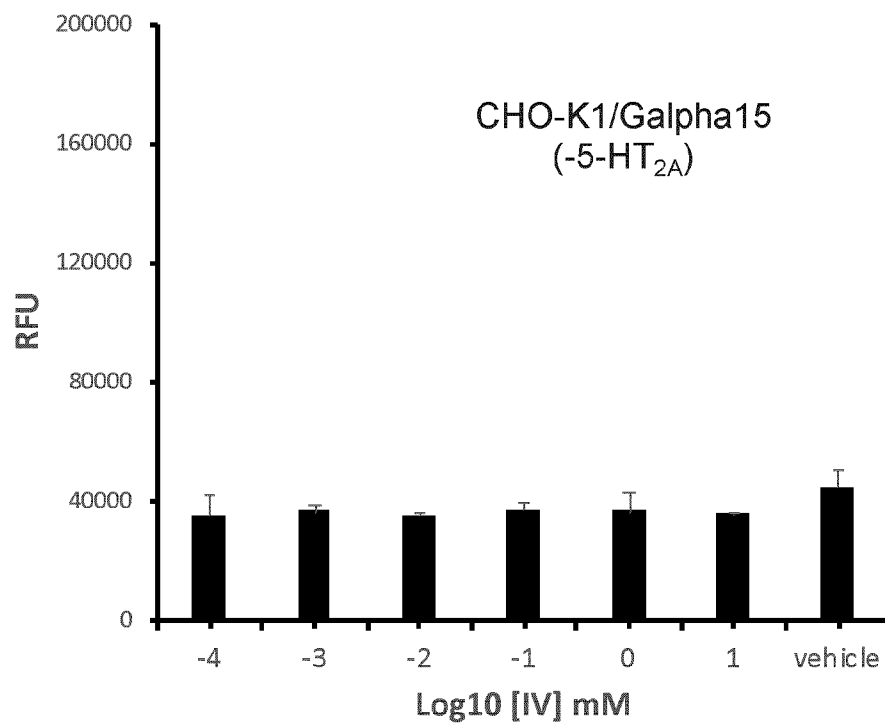


FIG. 27B

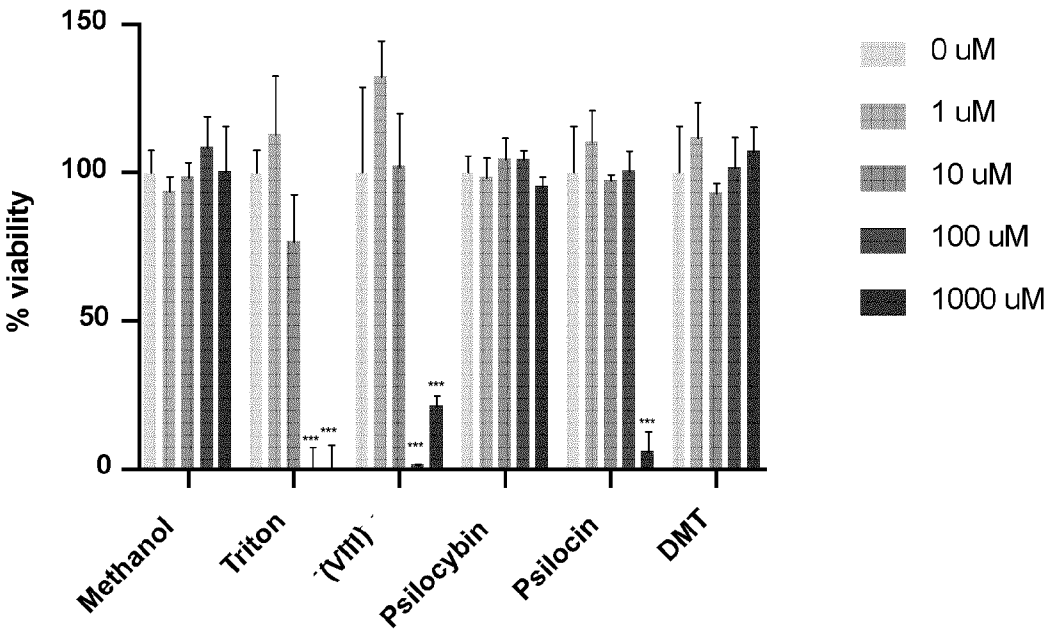


FIG. 28

35/36

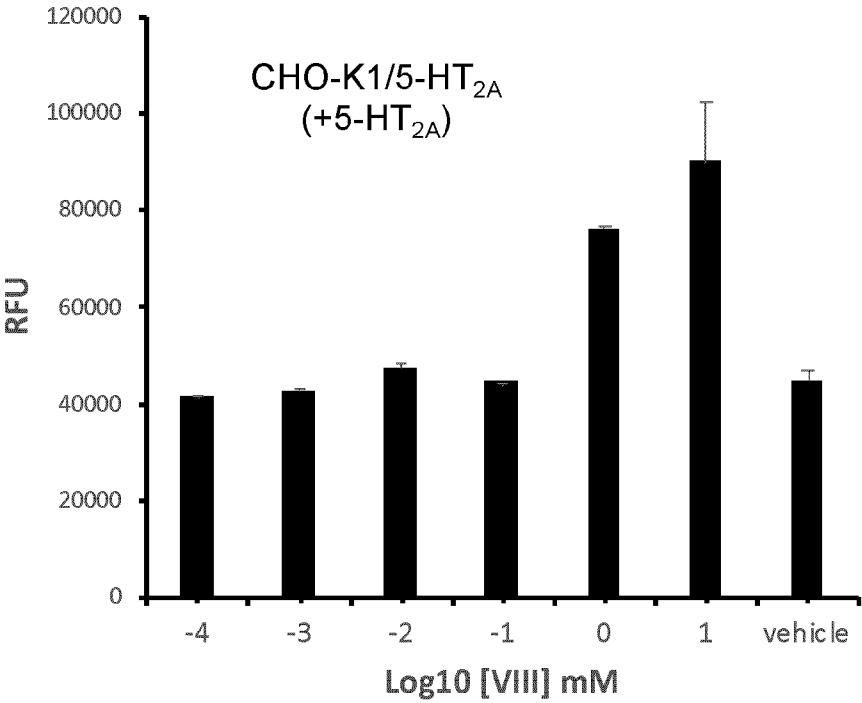


FIG. 29A

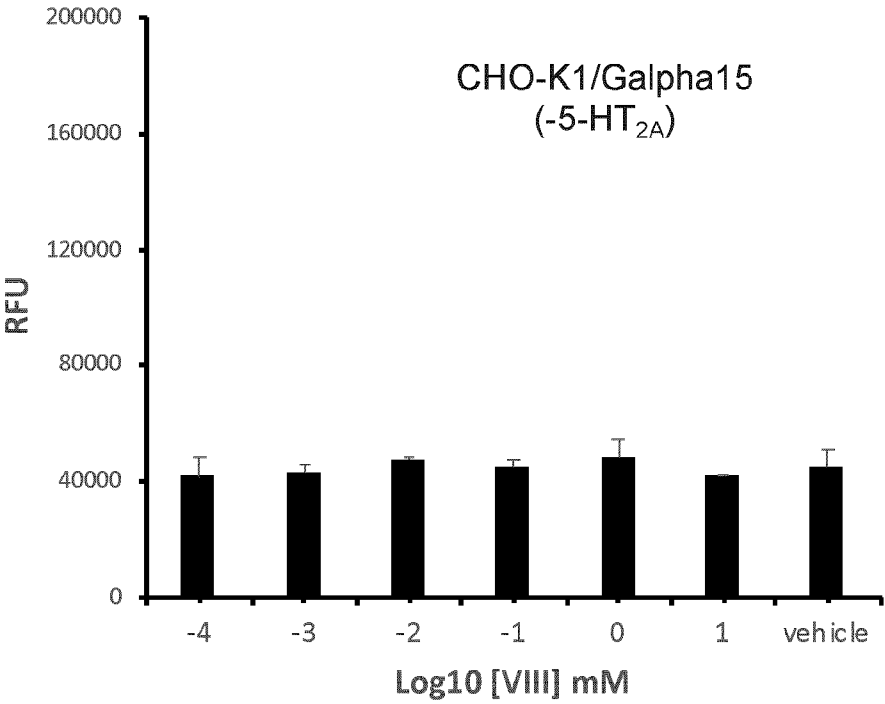
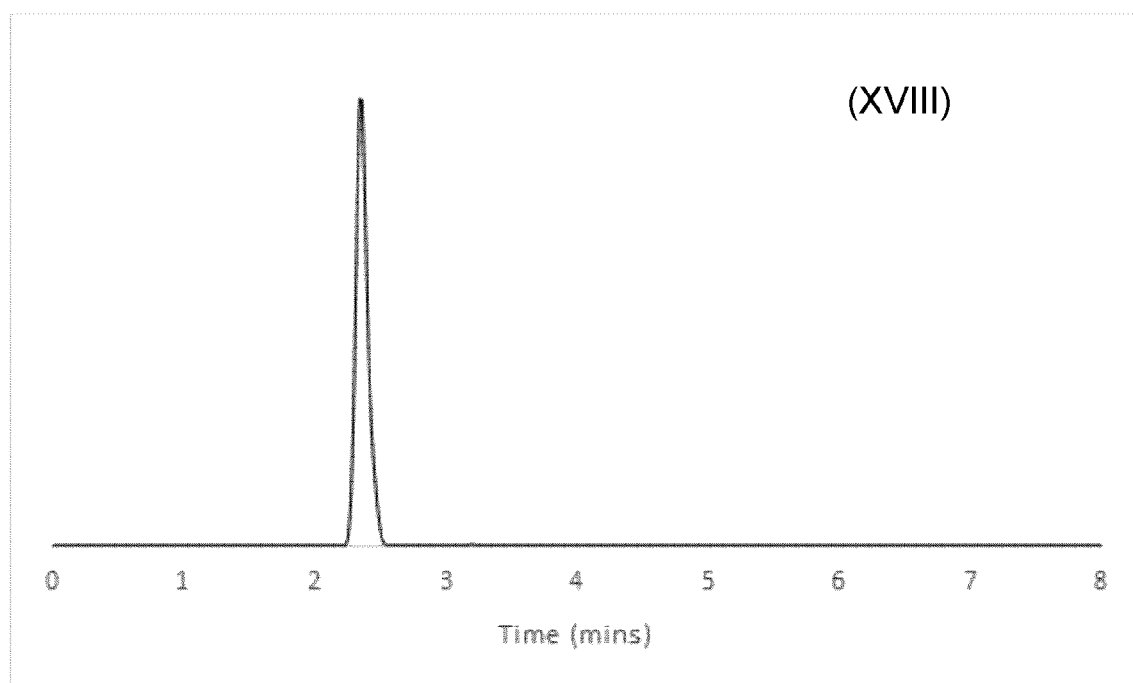


FIG. 29B

36/36

**FIG. 30**

# SEQUENCE LISTING

<110> MAGICMED INDUSTRIES INC.

<120> HALOGENATED PSILOCYBIN DERIVATIVES AND METHODS OF USING

<130> 21806-P62051PC00

<150> US 63/073,104

<151> 2020-09-01

<160> 98

<170> PatentIn version 3.5

<210> 1

<211> 1320

<212> DNA

<213> Psilocybe cubensis

<400> 1

```
atgcaggatga taccgcgtg caactcggca gcaataagat cactatgtcc tactcccgag      60
tcttttagaa acatgggatg gctctctgtc agcgatgcgg tctacagcga gttcatagga      120
gagttggcta cccgcgttc caatcgaaat tactccaacg agttcggcct catgcaacct      180
atccaggaat tcaaggcttt cattgaaagc gaccgggtgg tgcaccaaga atttattgac      240
atgttcgagg gcattcagga ctctccaagg aattatcagg aactatgtaa tatgttcaac      300
gatatctttc gcaaagctcc cgtctacgga gaccttggcc ctccggttta tatgattatg      360
gccaaattaa tgaacacccg agcgggcttc tctgcattca cgagacaaag gttgaacctt      420
cacttcaaaa aacttttcga tacctgggga ttgttctgtt cttcgaaaga ttctcgaaat      480
gttcttgtgg ccgaccagtt cgacgacaga cattgcggct ggttgaacga gcgggccttg      540
tctgctatgg ttaaacatta caatggacgc gcatttgatg aagtcttcct ctgcgataaa      600
aatgccccat actacggctt caactcttac gacgacttct ttaatcgag atttcgaaac      660
cgagatatcg accgacctgt agtcgggtga gttacaaca ccacctcat ttctgctgct      720
tgcgaaatcac tttctacaa cgtctcttat gacgtccagt ctctcgacac tttagttttc      780
aaaggagaga cttattcgct taagcatttg ctgaataatg accctttcac cccacaattc      840
gagcatggga gtattctaca aggattcttg aacgtcaccg cttaccaccg atggcacgca      900
cccgatcaatg ggacaatcgt caaaatcatc aacgttccag gtacctactt tgcgcaagcc      960
ccgagcacga ttggcgaccc tatcccgatg aacgattacg acccacctcc ttaccttaag     1020
tctcttgtct acttctctaa tattgccgca aggcaaatta tgtttattga agccgacaac     1080
aaggaaattg gcctcatctt ctttgtgttc atcggcatga ccgaaatctc gacatgtgaa     1140
gccacggtgt ccgaaggcca acacgtcaat cgtggcgatg acttggaat gttccatttc     1200
ggtggttctt cgttcgcgtt tggcttgagg aaggattgca gggcagagat cgttgaaaag     1260
ttcaccgaac ccggaacagt gatcagaatc aacgaagtcg tcgctgctct aaaggcttag     1320
```

<210> 2

<211> 439

<212> PRT

<213> *Psilocybe cubensis*

<400> 2

Met Gln Val Ile Pro Ala Cys Asn Ser Ala Ala Ile Arg Ser Leu Cys  
1 5 10 15

Pro Thr Pro Glu Ser Phe Arg Asn Met Gly Trp Leu Ser Val Ser Asp  
20 25 30

Ala Val Tyr Ser Glu Phe Ile Gly Glu Leu Ala Thr Arg Ala Ser Asn  
35 40 45

Arg Asn Tyr Ser Asn Glu Phe Gly Leu Met Gln Pro Ile Gln Glu Phe  
50 55 60

Lys Ala Phe Ile Glu Ser Asp Pro Val Val His Gln Glu Phe Ile Asp  
65 70 75 80

Met Phe Glu Gly Ile Gln Asp Ser Pro Arg Asn Tyr Gln Glu Leu Cys  
85 90 95

Asn Met Phe Asn Asp Ile Phe Arg Lys Ala Pro Val Tyr Gly Asp Leu  
100 105 110

Gly Pro Pro Val Tyr Met Ile Met Ala Lys Leu Met Asn Thr Arg Ala  
115 120 125

Gly Phe Ser Ala Phe Thr Arg Gln Arg Leu Asn Leu His Phe Lys Lys  
130 135 140

Leu Phe Asp Thr Trp Gly Leu Phe Leu Ser Ser Lys Asp Ser Arg Asn  
145 150 155 160

Val Leu Val Ala Asp Gln Phe Asp Asp Arg His Cys Gly Trp Leu Asn  
165 170 175

Glu Arg Ala Leu Ser Ala Met Val Lys His Tyr Asn Gly Arg Ala Phe  
180 185 190

Asp Glu Val Phe Leu Cys Asp Lys Asn Ala Pro Tyr Tyr Gly Phe Asn  
195 200 205

Ser Tyr Asp Asp Phe Phe Asn Arg Arg Phe Arg Asn Arg Asp Ile Asp  
210 215 220

Arg Pro Val Val Gly Gly Val Asn Asn Thr Thr Leu Ile Ser Ala Ala  
225 230 235 240

Cys Glu Ser Leu Ser Tyr Asn Val Ser Tyr Asp Val Gln Ser Leu Asp  
245 250 255

Thr Leu Val Phe Lys Gly Glu Thr Tyr Ser Leu Lys His Leu Leu Asn  
 260 265 270  
 Asn Asp Pro Phe Thr Pro Gln Phe Glu His Gly Ser Ile Leu Gln Gly  
 275 280 285  
 Phe Leu Asn Val Thr Ala Tyr His Arg Trp His Ala Pro Val Asn Gly  
 290 295 300  
 Thr Ile Val Lys Ile Ile Asn Val Pro Gly Thr Tyr Phe Ala Gln Ala  
 305 310 315 320  
 Pro Ser Thr Ile Gly Asp Pro Ile Pro Asp Asn Asp Tyr Asp Pro Pro  
 325 330 335  
 Pro Tyr Leu Lys Ser Leu Val Tyr Phe Ser Asn Ile Ala Ala Arg Gln  
 340 345 350  
 Ile Met Phe Ile Glu Ala Asp Asn Lys Glu Ile Gly Leu Ile Phe Leu  
 355 360 365  
 Val Phe Ile Gly Met Thr Glu Ile Ser Thr Cys Glu Ala Thr Val Ser  
 370 375 380  
 Glu Gly Gln His Val Asn Arg Gly Asp Asp Leu Gly Met Phe His Phe  
 385 390 395 400  
 Gly Gly Ser Ser Phe Ala Leu Gly Leu Arg Lys Asp Cys Arg Ala Glu  
 405 410 415  
 Ile Val Glu Lys Phe Thr Glu Pro Gly Thr Val Ile Arg Ile Asn Glu  
 420 425 430  
 Val Val Ala Ala Leu Lys Ala  
 435

<210> 3  
 <211> 2155  
 <212> DNA  
 <213> *Psilocybe cubensis*

<400> 3  
 atgatcgctg tactattctc cttcgtcatt gcaggatgca tatactacat cgtttctcgt 60  
 agagtgaggc ggtcgcgctt gccaccaggg ccgcctggca ttcctattcc cttcattggg 120  
 aacatgtttg atatgcctga agaatctcca tggtaacat ttctacaatg gggacgggat 180  
 tacagtctgt cttgccgcgt tgacttctaa tatatgaaca gctaatatat tgtcagacac 240  
 cgatattctc tacgtggatg ctggagggac agaaatgggt attcttaaca cgttggagac 300  
 cattaccgat ctattagaaa agcgagggtc catttattct ggccgggtgag ctgatgttga 360  
 gttttttgca attgaatttg tggtcacacg tttccagact tgagagtaca atgggtcaacg 420  
 aacttatggg gtgggagttt gacttagggt tcatcacata cggcgacagg tggcgcgaag 480

|            |             |             |             |            |             |            |      |
|------------|-------------|-------------|-------------|------------|-------------|------------|------|
| aaaggcgc   | cat         | gttcgccaag  | gagttcagtg  | agaagggcat | caagcaattt  | cgccatgctc | 540  |
| aagtgaagc  | tgcccatcag  | cttgtccaac  | agcttaccaa  | aacgccagac | cgctgggcac  |            | 600  |
| aacatattcg | ccagtaagta  | ctacttgagg  | aaaatagcgt  | acgcttcgct | gaccggtccg  |            | 660  |
| tacatcaaag | tcagatagcg  | gcaatgtcac  | tggatattgg  | ttatggaatt | gatcttgtag  |            | 720  |
| aagacgaccc | ttggctggaa  | gcgacccatt  | tggctaata   | aggcctcgcc | atagcatcag  |            | 780  |
| tgccgggcaa | atthttgggtc | gattcgttcc  | cttctcgtga  | gcacccctt  | tctatgtagg  |            | 840  |
| aagggaagga | gtctaacaag  | tgtagtaaaa  | atactttcct  | gcttggttcc | cagggtgctgt |            | 900  |
| cttcaagcgc | aaagcgaagg  | tctggcgaga  | agccgccgac  | catatggttg | acatgcctta  |            | 960  |
| tgaaactatg | aggaaattag  | cagttagtca  | aatgcgttct  | ccccgtattt | tttcaatact  |            | 1020 |
| ctaacttcag | ctcacagcct  | caaggattga  | ctcgtccgtc  | gtatgcttca | gctcgtctgc  |            | 1080 |
| aagccatgga | tctcaacggt  | gaccttgagc  | atcaagaaca  | cgtaatcaag | aacacagccg  |            | 1140 |
| cagaggttaa | tgctggtaag  | tcaaaagcgt  | ccgtcggcaa  | ttcaaaattc | aggcgctaaa  |            | 1200 |
| gtgggtcttc | tcaccaaggt  | ggaggcgata  | ctgtaaggat  | ttctcaatcg | ttagagtata  |            | 1260 |
| agtgttctaa | tgtagtacat  | actccaccaa  | ccagactgtc  | tctgctatgt | ctgcgttcat  |            | 1320 |
| cttggccatg | gtgaagtacc  | ctgagggtcca | gcgaaagggt  | caagcggagc | ttgatgctct  |            | 1380 |
| gaccaataac | ggccaaattc  | ctgactatga  | cgaagaagat  | gactccttgc | catacctcac  |            | 1440 |
| cgcatgtatc | aaggagcttt  | tccggtggaa  | tcaaatcgca  | cccctcgcta | taccgcacaa  |            | 1500 |
| attaatgaag | gacgacgtgt  | accgcgggta  | tctgattccc  | aagaacactc | tagtcttcgc  |            | 1560 |
| aaacacctgg | tgaggctgtc  | cattcattcc  | tagtacatcc  | gttgccccac | taatagcatc  |            | 1620 |
| ttgataacag | ggcagtatta  | aacgatccag  | aagtctatcc  | agatccctct | gtgttccgcc  |            | 1680 |
| cagaaagata | tcttggtcct  | gacgggaagc  | ctgataacac  | tgtacgcgac | ccacgtaaag  |            | 1740 |
| cggcatttgg | ctatggacga  | cgaattgggt  | aagtgcgctt  | tcagaacccc | cccttccgtt  |            | 1800 |
| gactagtgcc | atgcgcgc    | acaatatcgc  | tattgatctg  | atataacttc | cctgcggcat  |            | 1860 |
| ttattttggc | attccttttag | tcccgggaatt | catctagcgc  | agtcgacggt | ttggattgca  |            | 1920 |
| ggggcaaccc | tcttatcagc  | gttcaatatc  | gagcgacctg  | tcgatcagaa | tggaagccc   |            | 1980 |
| attgacatac | cggtgattt   | tactacagga  | ttcttcagggt | agctaatttc | cgcttttgtg  |            | 2040 |
| tgcataatac | ccctaacgac  | gcacgtttac  | ctttttgtaa  | agacaccag  | tgcttttcca  |            | 2100 |
| gtgcaggttt | gttcctcgaa  | cagagcaagt  | ctcacagtcg  | gtatccggac | cctga       |            | 2155 |

<210> 4  
 <211> 508  
 <212> PRT  
 <213> Psilocybe cubensis

<400> 4

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ile | Ala | Val | Leu | Phe | Ser | Phe | Val | Ile | Ala | Gly | Cys | Ile | Tyr | Tyr |
| 1   |     |     |     | 5   |     |     |     | 10  |     |     |     |     |     | 15  |     |



Ile Val Ser Arg Arg Val Arg Arg Ser Arg Leu Pro Pro Gly Pro Pro  
 20 25 30

Gly Ile Pro Ile Pro Phe Ile Gly Asn Met Phe Asp Met Pro Glu Glu  
 35 40 45

Ser Pro Trp Leu Thr Phe Leu Gln Trp Gly Arg Asp Tyr Asn Thr Asp  
 50 55 60

Ile Leu Tyr Val Asp Ala Gly Gly Thr Glu Met Val Ile Leu Asn Thr  
 65 70 75 80

Leu Glu Thr Ile Thr Asp Leu Leu Glu Lys Arg Gly Ser Ile Tyr Ser  
 85 90 95

Gly Arg Leu Glu Ser Thr Met Val Asn Glu Leu Met Gly Trp Glu Phe  
 100 105 110

Asp Leu Gly Phe Ile Thr Tyr Gly Asp Arg Trp Arg Glu Glu Arg Arg  
 115 120 125

Met Phe Ala Lys Glu Phe Ser Glu Lys Gly Ile Lys Gln Phe Arg His  
 130 135 140

Ala Gln Val Lys Ala Ala His Gln Leu Val Gln Gln Leu Thr Lys Thr  
 145 150 155 160

Pro Asp Arg Trp Ala Gln His Ile Arg His Gln Ile Ala Ala Met Ser  
 165 170 175

Leu Asp Ile Gly Tyr Gly Ile Asp Leu Ala Glu Asp Asp Pro Trp Leu  
 180 185 190

Glu Ala Thr His Leu Ala Asn Glu Gly Leu Ala Ile Ala Ser Val Pro  
 195 200 205

Gly Lys Phe Trp Val Asp Ser Phe Pro Ser Leu Lys Tyr Leu Pro Ala  
 210 215 220

Trp Phe Pro Gly Ala Val Phe Lys Arg Lys Ala Lys Val Trp Arg Glu  
 225 230 235 240

Ala Ala Asp His Met Val Asp Met Pro Tyr Glu Thr Met Arg Lys Leu  
 245 250 255

Ala Pro Gln Gly Leu Thr Arg Pro Ser Tyr Ala Ser Ala Arg Leu Gln  
 260 265 270

Ala Met Asp Leu Asn Gly Asp Leu Glu His Gln Glu His Val Ile Lys  
 275 280 285

Asn Thr Ala Ala Glu Val Asn Val Gly Gly Gly Asp Thr Thr Val Ser

290

295

300

Ala Met Ser Ala Phe Ile Leu Ala Met Val Lys Tyr Pro Glu Val Gln  
305 310 315 320

Arg Lys Val Gln Ala Glu Leu Asp Ala Leu Thr Asn Asn Gly Gln Ile  
325 330 335

Pro Asp Tyr Asp Glu Glu Asp Asp Ser Leu Pro Tyr Leu Thr Ala Cys  
340 345 350

Ile Lys Glu Leu Phe Arg Trp Asn Gln Ile Ala Pro Leu Ala Ile Pro  
355 360 365

His Lys Leu Met Lys Asp Asp Val Tyr Arg Gly Tyr Leu Ile Pro Lys  
370 375 380

Asn Thr Leu Val Phe Ala Asn Thr Trp Ala Val Leu Asn Asp Pro Glu  
385 390 395 400

Val Tyr Pro Asp Pro Ser Val Phe Arg Pro Glu Arg Tyr Leu Gly Pro  
405 410 415

Asp Gly Lys Pro Asp Asn Thr Val Arg Asp Pro Arg Lys Ala Ala Phe  
420 425 430

Gly Tyr Gly Arg Arg Asn Cys Pro Gly Ile His Leu Ala Gln Ser Thr  
435 440 445

Val Trp Ile Ala Gly Ala Thr Leu Leu Ser Ala Phe Asn Ile Glu Arg  
450 455 460

Pro Val Asp Gln Asn Gly Lys Pro Ile Asp Ile Pro Ala Asp Phe Thr  
465 470 475 480

Thr Gly Phe Phe Arg His Pro Val Pro Phe Gln Cys Arg Phe Val Pro  
485 490 495

Arg Thr Glu Gln Val Ser Gln Ser Val Ser Gly Pro  
500 505

&lt;210&gt; 5

&lt;211&gt; 1089

&lt;212&gt; DNA

&lt;213&gt; Psilocybe cubensis

&lt;400&gt; 5

atggcggttcg atctcaagac tgaagacggc ctcatcacat atctcactaa acatctttct 60

ttggacgtcg acacgagcgg agtgaagcgc cttagcggag gctttgtcaa tgtaacctgg 120

cgcattaagc tcaatgctcc ttatcaaggt catagagca tcatcctgaa gcatgctcag 180

ccgcacatgt ctacggatga ggattttaag ataggtgtag aacgttcggt ttacgaatac 240

caggctatca agctcatgat ggccaatcgg gaggttctgg gaggcgtgga tggcatagtt 300  
tctgtgccag aaggcctgaa ctacgactta gagaataatg cattgatcat gcaagatgtc 360  
gggaagatga agaccctttt agattatgtc accgccaaac cgccacttgc gacggatata 420  
gcccgccttg ttgggacaga aattggggggg ttctgttgcca gactccataa cataggccgc 480  
gagaggcgag acgatcctga gttcaaattc ttctctggaa atattgtcgg aaggacgact 540  
tcagaccagc tgtatcaaac catcataccc aacgcagcga aatatggcgt cgatgacccc 600  
ttgtgccta ctgtggttaa ggaccttggtg gacgatgtca tgcacagcga agagaccctt 660  
gtcatggcgg acctgtggag tggaaatatt cttctccagt tggaggaggg aaacccatcg 720  
aagctgcaga agatatatat cctggattgg gaactttgca agtacggccc agcgtcgttg 780  
gacctgggct atttcttggg tgactgctat ttgatatccc gctttcaaga cgagcaggtc 840  
ggtacgacga tgcggcaagc ctacttgcaa agctatgcgc gtacgagcaa gcattcgatc 900  
aactacgcca aagtcactgc aggtattgct gctcatattg tgatgtggac cgactttatg 960  
cagtggggga gcgaggaaga aaggataaat tttgtgaaaa agggggtagc tgcctttcac 1020  
gacgccaggg gcaacaacga caatggggaa attacgtcta cttactgaa ggaatcatcc 1080  
actgcgtaa 1089

<210> 6  
<211> 309  
<212> PRT  
<213> Psilocybe cubensis

<400> 6

Met His Ile Arg Asn Pro Tyr Arg Thr Pro Ile Asp Tyr Gln Ala Leu  
1 5 10 15  
  
Ser Glu Ala Phe Pro Pro Leu Lys Pro Phe Val Ser Val Asn Ala Asp  
20 25 30  
  
Gly Thr Ser Ser Val Asp Leu Thr Ile Pro Glu Ala Gln Arg Ala Phe  
35 40 45  
  
Thr Ala Ala Leu Leu His Arg Asp Phe Gly Leu Thr Met Thr Ile Pro  
50 55 60  
  
Glu Asp Arg Leu Cys Pro Thr Val Pro Asn Arg Leu Asn Tyr Val Leu  
65 70 75 80  
  
Trp Ile Glu Asp Ile Phe Asn Tyr Thr Asn Lys Thr Leu Gly Leu Ser  
85 90 95  
  
Asp Asp Arg Pro Ile Lys Gly Val Asp Ile Gly Thr Gly Ala Ser Ala  
100 105 110  
  
Ile Tyr Pro Met Leu Ala Cys Ala Arg Phe Lys Ala Trp Ser Met Val  
115 120 125

Gly Thr Glu Val Glu Arg Lys Cys Ile Asp Thr Ala Arg Leu Asn Val  
130 135 140

Val Ala Asn Asn Leu Gln Asp Arg Leu Ser Ile Leu Glu Thr Ser Ile  
145 150 155 160

Asp Gly Pro Ile Leu Val Pro Ile Phe Glu Ala Thr Glu Glu Tyr Glu  
165 170 175

Tyr Glu Phe Thr Met Cys Asn Pro Pro Phe Tyr Asp Gly Ala Ala Asp  
180 185 190

Met Gln Thr Ser Asp Ala Ala Lys Gly Phe Gly Phe Gly Val Gly Ala  
195 200 205

Pro His Ser Gly Thr Val Ile Glu Met Ser Thr Glu Gly Gly Glu Ser  
210 215 220

Ala Phe Val Ala Gln Met Val Arg Glu Ser Leu Lys Leu Arg Thr Arg  
225 230 235 240

Cys Arg Trp Tyr Thr Ser Asn Leu Gly Lys Leu Lys Ser Leu Lys Glu  
245 250 255

Ile Val Gly Leu Leu Lys Glu Leu Glu Ile Ser Asn Tyr Ala Ile Asn  
260 265 270

Glu Tyr Val Gln Gly Ser Thr Arg Arg Tyr Ala Val Ala Trp Ser Phe  
275 280 285

Thr Asp Ile Gln Leu Pro Glu Glu Leu Ser Arg Pro Ser Asn Pro Glu  
290 295 300

Leu Ser Ser Leu Phe  
305

<210> 7  
<211> 930  
<212> DNA  
<213> Psilocybe cubensis

<400> 7  
atgcatatca gaaatcctta ccgtacacca attgactatc aagcactttc agaggccttc 60  
cctccccctca agccatttgt gtctgtcaat gcagatggta ccagttctgt tgacctcact 120  
atcccagaag cccagagggc gttcacggcc gctcttcttc atcgtgactt cgggctcacc 180  
atgaccatac cagaagaccg tctgtgccca acagtcccca ataggttgaa ctacgttctg 240  
tggattgaag atattttcaa ctacacgaac aaaaccctcg gcctgtcgga tgaccgtcct 300  
attaaaggcg ttgatattgg tacaggagcc tccgcaattt atcctatgct tgcctgtgct 360  
cggttcaagg catggtctat gggttgaaca gaggtcgaga ggaagtgcac tgacacggcc 420

cgctcaatg tcgtcgcgaa caatctccaa gaccgtctct cgatattaga gacatccatt 480  
gatggtccta ttctcgtccc cattttcgag gcgactgaag aatacgaata cgagtttact 540  
atgtgtaacc ctccattcta cgacgggtgct gccgatatgc agacttcgga tgctgccaaa 600  
ggatttggat ttggcgtggg cgctcccat tctggaacag tcatcgaaat gtcgactgag 660  
ggaggatgaat cggctttcgt cgctcagatg gtccgtgaga gcttgaagct tcgaacacga 720  
tgcagatggg acacgagtaa cttgggaaag ctgaaatcct tgaagaaat agtggggctg 780  
ctgaaagaac ttgagataag caactatgcc attaacgaat acgttcaggg gtccacacgt 840  
cgttatgccg ttgcgtggtc tttcactgat attcaactgc ctgaggagct ttctcgtccc 900  
tctaaccg agctcagctc tcttttctag 930

<210> 8  
<211> 309  
<212> PRT  
<213> Psilocybe cubensis

<400> 8

Met His Ile Arg Asn Pro Tyr Arg Thr Pro Ile Asp Tyr Gln Ala Leu  
1 5 10 15  
  
Ser Glu Ala Phe Pro Pro Leu Lys Pro Phe Val Ser Val Asn Ala Asp  
20 25 30  
  
Gly Thr Ser Ser Val Asp Leu Thr Ile Pro Glu Ala Gln Arg Ala Phe  
35 40 45  
  
Thr Ala Ala Leu Leu His Arg Asp Phe Gly Leu Thr Met Thr Ile Pro  
50 55 60  
  
Glu Asp Arg Leu Cys Pro Thr Val Pro Asn Arg Leu Asn Tyr Val Leu  
65 70 75 80  
  
Trp Ile Glu Asp Ile Phe Asn Tyr Thr Asn Lys Thr Leu Gly Leu Ser  
85 90 95  
  
Asp Asp Arg Pro Ile Lys Gly Val Asp Ile Gly Thr Gly Ala Ser Ala  
100 105 110  
  
Ile Tyr Pro Met Leu Ala Cys Ala Arg Phe Lys Ala Trp Ser Met Val  
115 120 125  
  
Gly Thr Glu Val Glu Arg Lys Cys Ile Asp Thr Ala Arg Leu Asn Val  
130 135 140  
  
Val Ala Asn Asn Leu Gln Asp Arg Leu Ser Ile Leu Glu Thr Ser Ile  
145 150 155 160  
  
Asp Gly Pro Ile Leu Val Pro Ile Phe Glu Ala Thr Glu Glu Tyr Glu  
165 170 175

Tyr Glu Phe Thr Met Cys Asn Pro Pro Phe Tyr Asp Gly Ala Ala Asp  
180 185 190

Met Gln Thr Ser Asp Ala Ala Lys Gly Phe Gly Phe Gly Val Gly Ala  
195 200 205

Pro His Ser Gly Thr Val Ile Glu Met Ser Thr Glu Gly Gly Glu Ser  
210 215 220

Ala Phe Val Ala Gln Met Val Arg Glu Ser Leu Lys Leu Arg Thr Arg  
225 230 235 240

Cys Arg Trp Tyr Thr Ser Asn Leu Gly Lys Leu Lys Ser Leu Lys Glu  
245 250 255

Ile Val Gly Leu Leu Lys Glu Leu Glu Ile Ser Asn Tyr Ala Ile Asn  
260 265 270

Glu Tyr Val Gln Gly Ser Thr Arg Arg Tyr Ala Val Ala Trp Ser Phe  
275 280 285

Thr Asp Ile Gln Leu Pro Glu Glu Leu Ser Arg Pro Ser Asn Pro Glu  
290 295 300

Leu Ser Ser Leu Phe  
305

<210> 9  
<211> 3042  
<212> DNA  
<213> *Psilocybe cubensis*

<400> 9  
atgccttcca gtcaccctca cattactcat cgctatcggg ttccttcgag tgacgaccat 60  
gaacgtatat ctgctctgtt cttgggtccc aaagcagaaa atgccgcatt tctccagcaa 120  
tggttgacca cggtcgtcgc acagcaaaag gctgcccgcg atgcatactt cccggatgac 180  
aatgctttta ttactacaga catgcaaact tccccgcct ttgctcagac tactaaagta 240  
atgcctcca atctaccga attattgact gcactcgggt aaaggtcgat tcctttcttc 300  
tcacctcggg acagcggcca tatgtctgtg gaccaaagtc tacctgcat tctcggattc 360  
ttatcgacca cattttataa tcctaacaat gttgccttcg aggctagtcc attcacgacc 420  
ctcatcgagg aagaagtgg cttgcaactc tctgaaatgc tgggttataa tcggctaaat 480  
aacaccgaga aacctctcgc ctggggacat attgcatcag gtggaactgt tgcaaacttg 540  
gaagcgatgt gggcggcgcg aaacctcaag ttttacctc tctactccg tgatgcttca 600  
gccgaaggcg cagagatgga attcattcgt gacacattct ccgtcaaaac ctgtgttggt 660  
gacaaaaaat tattaaagga ttgcagccca tgggaactcc tcaatttgca tgtttctact 720  
atcttagaca tgcccaccg tctgcacgac gagtacaata ttccacctca gttcctcgaa 780

|                                                                    |      |
|--------------------------------------------------------------------|------|
| aaggttatgc gaaagtatat catccagtct accaacaag acacgttgat gcagcgttgg   | 840  |
| ggacttacct aacaacctgt cgttttatcc ccgagcacia accattattc ctggccaaag  | 900  |
| gctgcagctg tgctcggtat tggctcagac aaccttcgca acgtcccagt agacatccaa  | 960  |
| gcccacatgg acataaacga actcgatcgt atgttaaaaa tttgcttgga cgaggagacg  | 1020 |
| ccagtatatc aagtagttgc tgttatcggc accaccgaag agggcgggtg cgatcgcatt  | 1080 |
| acggagatcc tgaagctgcg ccaaaagtat gaagctttgg ggctgtcttt tgccatccat  | 1140 |
| gcagatgctg cttggggagg ctattttgca accatgctac ccaaagatac attgggtcga  | 1200 |
| aaccggacta ggcttcccaa agaggacact acctcgggct ttgtccctca cgtcggctctg | 1260 |
| cgcgaggaga gcgcgttaca actcagccat ataaagtatg ccgattctat tactatcgac  | 1320 |
| ccgcacaagg caggctatgt tccttacctt gctggggcac tctgttatcg cgacggaaga  | 1380 |
| atgaggtacc tgcttacatg gtccgcgccc taccttgccc aaggcaacga gggccaaagt  | 1440 |
| atcggaatat acgggatcga aggaagcaaa cctggtgcag cagcatccgc ggtattcatg  | 1500 |
| gcgcacgaaa ccattggcct gactccttct ggatacggga accttcttgg ccaggcaatg  | 1560 |
| tttcatgctc gccgatacgc tgctcactgg tctgcaatgt caacggatac taccagtttc  | 1620 |
| actgtcacc cgttcaatcc tatccctgct gacatcgacc ccaacgctga ccccgcaaag   | 1680 |
| gtcgaagagc aaaaacagtt catcagagat cgtatcttgt tcaaatcgaa cgaggaaata  | 1740 |
| tacaacgatt ctgaggctat ggaactcttg caccaacttg ggtccgatct caatatcaac  | 1800 |
| gttttcgcat gcaacttccg cgaccgcgat aataatctca acaccgacgt cgaggaagcc  | 1860 |
| aactggctca ataaccgtat tttccaacgc ttttctgtta caagtgtgga ggagaacca   | 1920 |
| ttggaaacgc cattcttctt cagctcaact acattgaaac aatccgaata cggcgtctgc  | 1980 |
| gcaaccgaag taaagagacg catgggactt gttggtgacc aggatgttat agtcctgagg  | 2040 |
| aacgtcgtta tgtctccatt tactacaacg aacgactttg tgggaactct ggcaaacacc  | 2100 |
| ttccaaaaga tcgttgagga ggaggctcag tatgcacgga tccgcaacga tatgaaacct  | 2160 |
| agcattcaca ccttccttct tcatggttca ggagagcaat actatcttgt ccacaccca   | 2220 |
| acgatccata tggccagcgg ccgtcgcaa atcatccttt cagtaaatgt tgaaggcaa    | 2280 |
| gttcggcagg cgatacatgc ccatgaaaga gttgaagcag tgattgtaca taacactgtg  | 2340 |
| cccctccgcc ttgacgaaat cgttgacgga ggatcatttg acggcatact caccatcgga  | 2400 |
| aagaggaaaa ctagtttcaa agtgaagatt tcaaacatta aagtagtcaa gaagcgtctt  | 2460 |
| ctgatgactg aggacctgga atctgcgtac ccatcgttga tgccattcta tttctacggg  | 2520 |
| actcaaggac acgctcatct cgaccatgtc attactgtcg ttcctaacat ccatctgagt  | 2580 |
| gctggcgaaa tacagtacaa attcgacgac gaggtgtcaa gcgaggacct cgccaagggc  | 2640 |
| ctcattgttg ttgctgagaa cgtacacgag gcatccatgc agcccttccc gctcatgaaa  | 2700 |
| gatttcaaga tcaccaacca attcttcttc agctccgggc aaatactccg cgtcaaagtg  | 2760 |
| tacagagatc cataccggc atcgacaatg gatcccatcc ctctccacga catcaagaac   | 2820 |
| cagcccgtcg tgacacaagg caccatcacg ctcgctcgaa atatttacgt cgattctgat  | 2880 |

gcgctcaacg tcgcttccga gcctactgcc gacgaagacg cggcgcatgt tcctcacgct 2940  
cgcaacatgt acggcgagat gaccgctgga acgatcaaag gctggcaaaa cgctgttcgt 3000  
catttccaca acaaattgga gactgttgct ccgacgaagt ag 3042

<210> 10  
<211> 1013  
<212> PRT  
<213> Psilocybe cubensis

<400> 10

Met Pro Ser Ser His Pro His Ile Thr His Arg Tyr Arg Val Pro Ser  
1 5 10 15

Ser Asp Asp His Glu Arg Ile Ser Ala Leu Phe Leu Gly Pro Lys Ala  
20 25 30

Glu Asn Ala Ala Phe Leu Gln Gln Trp Leu Thr Thr Val Val Ala Gln  
35 40 45

Gln Lys Ala Ala Arg Asp Ala Tyr Phe Pro Asp Asp Asn Ala Phe Ile  
50 55 60

Thr Thr Asp Met Gln Thr Ser Pro Ala Phe Ala Gln Thr Thr Lys Val  
65 70 75 80

Ile Ala Ser Asn Leu Thr Glu Leu Leu Thr Ala Leu Gly Glu Arg Ser  
85 90 95

Ile Pro Phe Phe Ser Pro Arg Tyr Ser Gly His Met Ser Val Asp Gln  
100 105 110

Ser Leu Pro Ala Ile Leu Gly Phe Leu Ser Thr Thr Phe Tyr Asn Pro  
115 120 125

Asn Asn Val Ala Phe Glu Ala Ser Pro Phe Thr Thr Leu Ile Glu Glu  
130 135 140

Glu Val Gly Leu Gln Leu Ser Glu Met Leu Gly Tyr Asn Arg Leu Asn  
145 150 155 160

Asn Thr Glu Lys Pro Leu Ala Trp Gly His Ile Ala Ser Gly Gly Thr  
165 170 175

Val Ala Asn Leu Glu Ala Met Trp Ala Ala Arg Asn Leu Lys Phe Tyr  
180 185 190

Pro Leu Ser Leu Arg Asp Ala Ser Ala Glu Gly Ala Glu Met Glu Phe  
195 200 205

Ile Arg Asp Thr Phe Ser Val Lys Thr Cys Val Gly Asp Lys Lys Leu  
210 215 220



Leu Lys Asp Cys Ser Pro Trp Glu Leu Leu Asn Leu His Val Ser Thr  
 225 230 235 240

Ile Leu Asp Met Pro Asp Arg Leu His Asp Glu Tyr Asn Ile Ser Pro  
 245 250 255

Gln Phe Leu Glu Lys Val Met Arg Lys Tyr Ile Ile Gln Ser Thr Asn  
 260 265 270

Lys Asp Thr Leu Met Gln Arg Trp Gly Leu Thr Gln Gln Pro Val Val  
 275 280 285

Leu Ser Pro Ser Thr Asn His Tyr Ser Trp Pro Lys Ala Ala Ala Val  
 290 295 300

Leu Gly Ile Gly Ser Asp Asn Leu Arg Asn Val Pro Val Asp Ile Gln  
 305 310 315 320

Ala His Met Asp Ile Asn Glu Leu Asp Arg Met Leu Lys Ile Cys Leu  
 325 330 335

Asp Glu Glu Thr Pro Val Tyr Gln Val Val Ala Val Ile Gly Thr Thr  
 340 345 350

Glu Glu Gly Gly Val Asp Arg Ile Thr Glu Ile Leu Lys Leu Arg Gln  
 355 360 365

Lys Tyr Glu Ala Leu Gly Leu Ser Phe Ala Ile His Ala Asp Ala Ala  
 370 375 380

Trp Gly Gly Tyr Phe Ala Thr Met Leu Pro Lys Asp Thr Leu Gly Arg  
 385 390 395 400

Asn Arg Thr Arg Leu Pro Lys Glu Asp Thr Thr Ser Gly Phe Val Pro  
 405 410 415

His Val Gly Leu Arg Glu Glu Ser Ala Leu Gln Leu Ser His Ile Lys  
 420 425 430

Tyr Ala Asp Ser Ile Thr Ile Asp Pro His Lys Ala Gly Tyr Val Pro  
 435 440 445

Tyr Pro Ala Gly Ala Leu Cys Tyr Arg Asp Gly Arg Met Arg Tyr Leu  
 450 455 460

Leu Thr Trp Ser Ala Pro Tyr Leu Ala Gln Gly Asn Glu Gly Gln Ser  
 465 470 475 480

Ile Gly Ile Tyr Gly Ile Glu Gly Ser Lys Pro Gly Ala Ala Ala Ser  
 485 490 495

Ala Val Phe Met Ala His Glu Thr Ile Gly Leu Thr Pro Ser Gly Tyr  
500 505 510

Gly Asn Leu Leu Gly Gln Ala Met Phe Thr Cys Arg Arg Tyr Ala Ala  
515 520 525

His Trp Ser Ala Met Ser Thr Asp Thr Thr Ser Phe Thr Val Thr Pro  
530 535 540

Phe Asn Pro Ile Pro Ala Asp Ile Asp Pro Asn Ala Asp Pro Ala Lys  
545 550 555 560

Val Glu Glu Gln Lys Gln Phe Ile Arg Asp Arg Ile Leu Phe Lys Ser  
565 570 575

Asn Glu Glu Ile Tyr Asn Asp Ser Glu Ala Met Glu Leu Leu His Gln  
580 585 590

Leu Gly Ser Asp Leu Asn Ile Asn Val Phe Ala Cys Asn Phe Arg Asp  
595 600 605

Arg Asp Asn Asn Leu Asn Thr Asp Val Glu Glu Ala Asn Trp Leu Asn  
610 615 620

Asn Arg Ile Phe Gln Arg Phe Ser Val Thr Ser Ala Glu Glu Asn Pro  
625 630 635 640

Leu Glu Thr Pro Phe Phe Leu Ser Ser Thr Thr Leu Lys Gln Ser Glu  
645 650 655

Tyr Gly Val Cys Ala Thr Glu Val Lys Arg Arg Met Gly Leu Val Gly  
660 665 670

Asp Gln Asp Val Ile Val Leu Arg Asn Val Val Met Ser Pro Phe Thr  
675 680 685

Thr Thr Asn Asp Phe Val Gly Thr Leu Ala Asn Thr Phe Gln Lys Ile  
690 695 700

Val Glu Glu Glu Val Glu Tyr Ala Arg Ile Arg Asn Asp Met Lys Pro  
705 710 715 720

Ser Ile His Thr Phe Leu Leu His Gly Ser Gly Glu Gln Tyr Tyr Leu  
725 730 735

Val His Thr Pro Thr Ile His Met Ala Ser Gly Arg Arg Gln Ile Ile  
740 745 750

Leu Ser Val Asn Val Glu Gly Gln Val Arg Gln Ala Ile His Ala His  
755 760 765

Glu Arg Val Glu Ala Val Ile Val His Asn Thr Val Pro Leu Arg Leu  
770 775 780

Asp Glu Ile Val Asp Gly Gly Ser Phe Asp Gly Ile Leu Thr Ile Gly  
785 790 795 800

Lys Arg Lys Thr Ser Phe Lys Val Lys Ile Ser Asn Ile Lys Val Val  
805 810 815

Lys Lys Arg Ser Leu Met Thr Glu Asp Leu Glu Ser Ala Tyr Pro Ser  
820 825 830

Leu Met Pro Phe Tyr Phe Tyr Gly Thr Gln Gly His Ala His Leu Asp  
835 840 845

His Val Ile Thr Val Val Pro Asn Ile His Leu Ser Ala Gly Glu Ile  
850 855 860

Gln Tyr Lys Phe Asp Asp Glu Val Ser Ser Glu Asp Leu Ala Lys Gly  
865 870 875 880

Leu Ile Val Val Ala Glu Asn Val His Glu Ala Ser Met Gln Pro Phe  
885 890 895

Pro Leu Met Lys Asp Phe Lys Ile Thr Asn Gln Phe Phe Phe Ser Ser  
900 905 910

Gly Gln Ile Leu Arg Val Lys Val Tyr Arg Asp Pro Tyr Pro Ala Ser  
915 920 925

Thr Met Asp Pro Ile Pro Leu His Asp Ile Lys Asn Gln Pro Val Val  
930 935 940

Thr Gln Gly Thr Ile Thr Leu Val Gly Asn Ile Tyr Val Asp Ser Asp  
945 950 955 960

Ala Leu Asn Val Ala Ser Glu Pro Thr Ala Asp Glu Asp Ala Ala His  
965 970 975

Val Pro His Ala Arg Asn Met Tyr Gly Glu Met Thr Ala Gly Thr Ile  
980 985 990

Lys Gly Trp Gln Asn Ala Val Arg His Phe His Asn Lys Leu Glu Thr  
995 1000 1005

Val Ala Pro Thr Lys  
1010

<210> 11  
<211> 2103  
<212> DNA  
<213> Psilocybe cubensis

<400> 11

|                                                                    |      |
|--------------------------------------------------------------------|------|
| atggaggccta tcaaaaaggt ttttgagaac aaaaaggcgg agggcattcc tgtgttggtg | 60   |
| acctttgtta ctgcaggata tcctcgtccc gaagatactg ttcccatctt gctggccatg  | 120  |
| gaggccggtg gtgctgatat catcgagctt ggtatgcat tttcagaccc aattgcagat   | 180  |
| ggtcctgtca tccaggaaac gaacacaatc gccgttgcaa accaggtaga ttataccact  | 240  |
| gttctcggac aacttcggga agccccgaaa caagggtca aggcacccgt tcttctgatg   | 300  |
| ggatattata accccatatt ggcttacgga gaagacagat ctattcaaga tgcggctgaa  | 360  |
| gctggagcca atgggtttat tatggtcgac cttccacccg aggaggctgt cgcttttcga  | 420  |
| gagaaatgta tcaaatacaa cctctcatat gttcctctaa ttgcaccctc aacgactctg  | 480  |
| tcgcgtataa agttcctctc aacaattgca gacacgttca tctatgtcgt gtctaaaatg  | 540  |
| ggaaccaccg gatcctcaga gaaggttgcc atgaataacg cccttccac catcatcgat   | 600  |
| cgtattcgcg agtacgtga agttccttta gcagtcggat ttggagtcgc cactcgggct   | 660  |
| cacttcaact acgtcgccga ttccgggtgt gatggtgtcg ttattggcac caaactcgtt  | 720  |
| aacgttatta aagagtcacc gcaaggggaa gcacccaaaa atgttgaggc atactgccgt  | 780  |
| gagatgagcc aaaagggaga aacaaatcgc gtcaaattct caccaactgc ccgtgctgcc  | 840  |
| agctccgaat caattcctgt tgttgttcct tctgttctcc ccgcacgttt cggagaattc  | 900  |
| ggaggacaat acgttcccga agctcttgct gattgtctgg ttgaactaga agaagctcac  | 960  |
| aaatctgcca tggctgatcc tgaattccag aaggaactac aatcgcatgc cggatatgca  | 1020 |
| aatcgtcctt cacaaatata cctcgccgaa aatctcacca aggatgctgg gggtgcaaat  | 1080 |
| atttggttga aacgtgaaga tttgaaccac acagggtccc acaaaatcaa taacgctttg  | 1140 |
| ggacaaattc tgcttgcccg gagaatcgga aagaccagaa ttatcgcaga aacaggtgcc  | 1200 |
| ggccagcatg gtgttgcaac agcgactgtt tgcgctaagt ttggaatgga atgtgttatc  | 1260 |
| tacatgggcg cagaagatgt gcgacggcaa gctctaaatg tattcaggat tgagatgcta  | 1320 |
| ggagcaaaag ttgttcctgt tacttcagga tcatgcacat tgaaggacgc tgtaaacgag  | 1380 |
| gccttccgtg actgggtgac aaacctttct acgacgcatt atttggttgg ctctgtaatt  | 1440 |
| ggacctcatc ctttccccac cattgtccga gatttccaaa aggtcattgg tcaagagatc  | 1500 |
| aaggctcaga tgttggccgc ccgcgcaaaa cttcctgatg tcgtcgtcgc ttgtgttggt  | 1560 |
| ggaggaagca atgctatcgg tacgttctat gattttattg gcgacaagag tgtacgtcta  | 1620 |
| gttgggggtg aagcaggagg agaaggtatt gacggagacc gacatagcgc cacactttcg  | 1680 |
| atggggcaac cgggagtact tcacggtgtt agaacatata ttctacaaga caaggccggt  | 1740 |
| caaatcatcg agacgcactc aatcagcgct ggattggatt atcccggcgt tggaccagaa  | 1800 |
| catgcttggc taaaggactc taaaagagca gaatatgttg tcgccacaga cgaagaagca  | 1860 |
| cttcgcgggt tccgtatgct aacacaaagg gagggaatta ttcctgccct tgaatcttcc  | 1920 |
| catgcgatct gggaggctgt caggattgcc cgcaccatgt cgaaggacca ggatcttgtt  | 1980 |
| gtgtgtttgt ctggccgagg tgataaagac gttgagcaaa tttctcaact tcttccaag   | 2040 |

tgggcggata ttctagactg gcatgtttct tcccatgccg ttggacacac aacaaaattc 2100  
taa 2103

<210> 12  
<211> 700  
<212> PRT  
<213> *Psilocybe cubensis*

<400> 12

Met Glu Ala Ile Lys Lys Val Phe Glu Asn Lys Lys Ala Glu Gly Ile  
1 5 10 15

Pro Val Leu Val Thr Phe Val Thr Ala Gly Tyr Pro Arg Pro Glu Asp  
20 25 30

Thr Val Pro Ile Leu Leu Ala Met Glu Ala Gly Gly Ala Asp Ile Ile  
35 40 45

Glu Leu Gly Met Pro Phe Ser Asp Pro Ile Ala Asp Gly Pro Val Ile  
50 55 60

Gln Glu Thr Asn Thr Ile Ala Val Ala Asn Gln Val Asp Tyr Thr Thr  
65 70 75 80

Val Leu Gly Gln Leu Arg Glu Ala Arg Lys Gln Gly Leu Lys Ala Pro  
85 90 95

Val Leu Leu Met Gly Tyr Tyr Asn Pro Ile Leu Ala Tyr Gly Glu Asp  
100 105 110

Arg Ser Ile Gln Asp Ala Ala Glu Ala Gly Ala Asn Gly Phe Ile Met  
115 120 125

Val Asp Leu Pro Pro Glu Glu Ala Val Ala Phe Arg Glu Lys Cys Ile  
130 135 140

Lys Ser Asn Leu Ser Tyr Val Pro Leu Ile Ala Pro Ser Thr Thr Leu  
145 150 155 160

Ser Arg Ile Lys Phe Leu Ser Thr Ile Ala Asp Thr Phe Ile Tyr Val  
165 170 175

Val Ser Lys Met Gly Thr Thr Gly Ser Ser Glu Lys Val Ala Met Asn  
180 185 190

Asn Ala Leu Pro Thr Ile Ile Asp Arg Ile Arg Glu Tyr Ala Glu Val  
195 200 205

Pro Leu Ala Val Gly Phe Gly Val Ala Thr Arg Ala His Phe Asn Tyr  
210 215 220

Val Ala Asp Ser Gly Ala Asp Gly Val Val Ile Gly Thr Lys Leu Val

| 225 | 230 |     |     |     |     |     |     |     | 235 |     |     |     | 240 |     |     |  |  |  |  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|--|--|
| Asn | Val | Ile | Lys | Glu | Ser | Pro | Gln | Gly | Glu | Ala | Pro | Lys | Asn | Val | Glu |  |  |  |  |
|     |     |     |     | 245 |     |     |     |     | 250 |     |     |     |     | 255 |     |  |  |  |  |
| Ala | Tyr | Cys | Arg | Glu | Met | Ser | Gln | Lys | Gly | Glu | Thr | Asn | Arg | Val | Lys |  |  |  |  |
|     |     |     | 260 |     |     |     |     | 265 |     |     |     |     | 270 |     |     |  |  |  |  |
| Ser | Pro | Pro | Thr | Ala | Arg | Ala | Ala | Ser | Ser | Glu | Ser | Ile | Pro | Val | Val |  |  |  |  |
|     |     |     | 275 |     |     |     | 280 |     |     |     |     | 285 |     |     |     |  |  |  |  |
| Val | Pro | Ser | Val | Leu | Pro | Ala | Arg | Phe | Gly | Glu | Phe | Gly | Gly | Gln | Tyr |  |  |  |  |
|     | 290 |     |     |     |     | 295 |     |     |     |     | 300 |     |     |     |     |  |  |  |  |
| Val | Pro | Glu | Ala | Leu | Val | Asp | Cys | Leu | Val | Glu | Leu | Glu | Glu | Ala | His |  |  |  |  |
| 305 |     |     |     |     | 310 |     |     |     |     | 315 |     |     |     |     | 320 |  |  |  |  |
| Lys | Ser | Ala | Met | Ala | Asp | Pro | Glu | Phe | Gln | Lys | Glu | Leu | Gln | Ser | His |  |  |  |  |
|     |     |     |     | 325 |     |     |     |     | 330 |     |     |     |     | 335 |     |  |  |  |  |
| Ala | Gly | Tyr | Ala | Asn | Arg | Pro | Ser | Gln | Ile | Tyr | Leu | Ala | Glu | Asn | Leu |  |  |  |  |
|     |     |     | 340 |     |     |     |     | 345 |     |     |     |     | 350 |     |     |  |  |  |  |
| Thr | Lys | Asp | Ala | Gly | Gly | Ala | Asn | Ile | Trp | Leu | Lys | Arg | Glu | Asp | Leu |  |  |  |  |
|     |     | 355 |     |     |     |     | 360 |     |     |     |     | 365 |     |     |     |  |  |  |  |
| Asn | His | Thr | Gly | Ser | His | Lys | Ile | Asn | Asn | Ala | Leu | Gly | Gln | Ile | Leu |  |  |  |  |
|     | 370 |     |     |     |     | 375 |     |     |     |     | 380 |     |     |     |     |  |  |  |  |
| Leu | Ala | Arg | Arg | Ile | Gly | Lys | Thr | Arg | Ile | Ile | Ala | Glu | Thr | Gly | Ala |  |  |  |  |
| 385 |     |     |     |     | 390 |     |     |     |     | 395 |     |     |     |     | 400 |  |  |  |  |
| Gly | Gln | His | Gly | Val | Ala | Thr | Ala | Thr | Val | Cys | Ala | Lys | Phe | Gly | Met |  |  |  |  |
|     |     |     |     | 405 |     |     |     |     | 410 |     |     |     |     | 415 |     |  |  |  |  |
| Glu | Cys | Val | Ile | Tyr | Met | Gly | Ala | Glu | Asp | Val | Arg | Arg | Gln | Ala | Leu |  |  |  |  |
|     |     |     | 420 |     |     |     |     | 425 |     |     |     |     | 430 |     |     |  |  |  |  |
| Asn | Val | Phe | Arg | Ile | Glu | Met | Leu | Gly | Ala | Lys | Val | Val | Pro | Val | Thr |  |  |  |  |
|     |     | 435 |     |     |     |     | 440 |     |     |     |     | 445 |     |     |     |  |  |  |  |
| Ser | Gly | Ser | Cys | Thr | Leu | Lys | Asp | Ala | Val | Asn | Glu | Ala | Phe | Arg | Asp |  |  |  |  |
|     | 450 |     |     |     |     | 455 |     |     |     |     | 460 |     |     |     |     |  |  |  |  |
| Trp | Val | Thr | Asn | Leu | Ser | Thr | Thr | His | Tyr | Leu | Val | Gly | Ser | Val | Ile |  |  |  |  |
| 465 |     |     |     |     | 470 |     |     |     |     | 475 |     |     |     |     | 480 |  |  |  |  |
| Gly | Pro | His | Pro | Phe | Pro | Thr | Ile | Val | Arg | Asp | Phe | Gln | Lys | Val | Ile |  |  |  |  |
|     |     |     |     | 485 |     |     |     |     | 490 |     |     |     |     | 495 |     |  |  |  |  |
| Gly | Gln | Glu | Ile | Lys | Ala | Gln | Met | Leu | Ala | Ala | Arg | Gly | Lys | Leu | Pro |  |  |  |  |
|     |     |     | 500 |     |     |     |     | 505 |     |     |     |     | 510 |     |     |  |  |  |  |

Asp Val Val Val Ala Cys Val Gly Gly Gly Ser Asn Ala Ile Gly Thr  
 515 520 525  
 Phe Tyr Asp Phe Ile Gly Asp Lys Ser Val Arg Leu Val Gly Val Glu  
 530 535 540  
 Ala Gly Gly Glu Gly Ile Asp Gly Asp Arg His Ser Ala Thr Leu Ser  
 545 550 555 560  
 Met Gly Gln Pro Gly Val Leu His Gly Val Arg Thr Tyr Ile Leu Gln  
 565 570 575  
 Asp Lys Ala Gly Gln Ile Ile Glu Thr His Ser Ile Ser Ala Gly Leu  
 580 585 590  
 Asp Tyr Pro Gly Val Gly Pro Glu His Ala Trp Leu Lys Asp Ser Lys  
 595 600 605  
 Arg Ala Glu Tyr Val Val Ala Thr Asp Glu Glu Ala Leu Arg Gly Phe  
 610 615 620  
 Arg Met Leu Thr Gln Arg Glu Gly Ile Ile Pro Ala Leu Glu Ser Ser  
 625 630 635 640  
 His Ala Ile Trp Glu Ala Val Arg Ile Ala Arg Thr Met Ser Lys Asp  
 645 650 655  
 Gln Asp Leu Val Val Cys Leu Ser Gly Arg Gly Asp Lys Asp Val Glu  
 660 665 670  
 Gln Ile Ser Gln Leu Leu Pro Lys Trp Ala Asp Ile Leu Asp Trp His  
 675 680 685  
 Val Ser Ser His Ala Val Gly His Thr Thr Lys Phe  
 690 695 700

<210> 13  
 <211> 1599  
 <212> DNA  
 <213> Streptomyces rugosporus

<400> 13  
 atggaaaggc ggaagcgtga gcgtcttggc tcactcggcc gaccgaccaa aaaggagctc 60  
 cgcatgatcc gatctgtggt gatcgtgggt ggtggcacgg cgggctggat gaccgcctcc 120  
 tacctcaagg ccgccttcga cgaccgcatc gacgtaacgc tcgtggagtc agggaacgtc 180  
 aggcggatcg gggtcggcga agcgaccttc agcacggtcc gccacttctt cgactacctg 240  
 ggccctcgacg agcgcgagtg gctgccccgc tgcgccggcg gctacaagct cggcatccgc 300  
 ttcgagaact ggagcgagcc gggcgagtac ttctaccacc cgttcgagcg cctgcgcgtc 360  
 gtcgacggct tcaacatggc cgagtgttg ctcgcggtcg gcgaccgcag gacgtccttc 420

```

agcgaggcct gctatctcac gcaccggctg tgcgaggcca agcgggcgcc ccgcatgctc      480
gacggctcgc tcttcgcctc ccagggtggac gagtcgctcg gccgctcgac cctggccgag      540
cagcgcgccc agttccccta gcctaccac ttcgacgccg acgaggtcgc ccgctacctg      600
tcggagtacg ccatcgcccc ggcgctccgc cacgtggtcg acgacgtgca gcacgtcggc      660
caggacgagc gcggctggat cagcggcgct cacaccaagc agcacggcga gatcagcggc      720
gacctgttcg tcgactgcac cggcttccgc ggcctgtcta tcaaccagac gctgggcggc      780
aggttccagt ctttctccga cgtgtgccc aacaaccggg cggtcgcgct gcgcgtcccg      840
cgggagaacg acgaggacat gcggccgtac acgacggcga ccgcgatgag cgccggctgg      900
atgtggacga tcccgtgtt caagcgcgac ggcaacggct acgtctactc cgacgagttc      960
atctcgccgg aggaggccga gcgcgagctg cggtcaccg tcgccccgg ccgcgacgac     1020
ctggaggcca accacatcca gatgcggatc ggagggaac agcgcacctg gatcaacaac     1080
tgcgtcgccg tcggcctgtc cgccgccttc gtcgagccgc tggagtcgac cggcatcttc     1140
ttcatccagc acgccatcga gcagctcgtg aagcacttcc ccggcgagcg ctgggaccg      1200
gtgctgatca gcgcgtacaa cgagcgcgatg gcgcacatgg tcgacggggt caaggagttc     1260
ctcgtcctcc actacaaggg gcccagcgc gaggacaccc cctactggaa ggccgccaag     1320
accagggcca tgcccgacgg cctcgccgc aagctggagc tgtccgcctc ccacctgctg     1380
gacgagcaga cgatctaccc ctactaccac ggcttcgaga cctattcgtg gatcaccatg     1440
aacctcggtc tcggcatcgt gcccgagcgg ccgcgtcccg cgctcctgca catggaccg      1500
gcgcccgcgc tggccgagtt cgaacggctc aggcgcgagg gcgacgagct gatcgccgcc      1560
ctgcccagct gctacgagta cctcgccagc atccaatga      1599

```

```

<210> 14
<211> 532
<212> PRT
<213> Streptomyces rugosporus

```

```

<400> 14

```

```

Met Glu Arg Arg Lys Arg Glu Arg Leu Gly Ser Leu Gly Arg Pro Thr
1           5           10          15

```

```

Lys Lys Glu Leu Arg Met Ile Arg Ser Val Val Ile Val Gly Gly Gly
          20          25          30

```

```

Thr Ala Gly Trp Met Thr Ala Ser Tyr Leu Lys Ala Ala Phe Asp Asp
          35          40          45

```

```

Arg Ile Asp Val Thr Leu Val Glu Ser Gly Asn Val Arg Arg Ile Gly
          50          55          60

```

```

Val Gly Glu Ala Thr Phe Ser Thr Val Arg His Phe Phe Asp Tyr Leu
65          70          75          80

```



|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|
| Gly | Leu | Asp | Glu | Arg | Glu | Trp | Leu | Pro | Arg | Cys | Ala | Gly | Gly | Tyr | Lys |  |  |
|     |     |     | 85  |     |     |     |     |     | 90  |     |     |     |     | 95  |     |  |  |
| Leu | Gly | Ile | Arg | Phe | Glu | Asn | Trp | Ser | Glu | Pro | Gly | Glu | Tyr | Phe | Tyr |  |  |
|     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |  |  |
| His | Pro | Phe | Glu | Arg | Leu | Arg | Val | Val | Asp | Gly | Phe | Asn | Met | Ala | Glu |  |  |
|     |     | 115 |     |     |     |     | 120 |     |     |     |     | 125 |     |     |     |  |  |
| Trp | Trp | Leu | Ala | Val | Gly | Asp | Arg | Arg | Thr | Ser | Phe | Ser | Glu | Ala | Cys |  |  |
|     | 130 |     |     |     |     | 135 |     |     |     |     | 140 |     |     |     |     |  |  |
| Tyr | Leu | Thr | His | Arg | Leu | Cys | Glu | Ala | Lys | Arg | Ala | Pro | Arg | Met | Leu |  |  |
| 145 |     |     |     |     | 150 |     |     |     |     | 155 |     |     |     |     | 160 |  |  |
| Asp | Gly | Ser | Leu | Phe | Ala | Ser | Gln | Val | Asp | Glu | Ser | Leu | Gly | Arg | Ser |  |  |
|     |     |     |     | 165 |     |     |     |     | 170 |     |     |     |     | 175 |     |  |  |
| Thr | Leu | Ala | Glu | Gln | Arg | Ala | Gln | Phe | Pro | Tyr | Ala | Tyr | His | Phe | Asp |  |  |
|     |     |     | 180 |     |     |     |     | 185 |     |     |     |     | 190 |     |     |  |  |
| Ala | Asp | Glu | Val | Ala | Arg | Tyr | Leu | Ser | Glu | Tyr | Ala | Ile | Ala | Arg | Gly |  |  |
|     |     | 195 |     |     |     |     | 200 |     |     |     |     | 205 |     |     |     |  |  |
| Val | Arg | His | Val | Val | Asp | Asp | Val | Gln | His | Val | Gly | Gln | Asp | Glu | Arg |  |  |
|     | 210 |     |     |     |     | 215 |     |     |     |     | 220 |     |     |     |     |  |  |
| Gly | Trp | Ile | Ser | Gly | Val | His | Thr | Lys | Gln | His | Gly | Glu | Ile | Ser | Gly |  |  |
| 225 |     |     |     |     | 230 |     |     |     |     | 235 |     |     |     |     | 240 |  |  |
| Asp | Leu | Phe | Val | Asp | Cys | Thr | Gly | Phe | Arg | Gly | Leu | Leu | Ile | Asn | Gln |  |  |
|     |     |     |     | 245 |     |     |     |     | 250 |     |     |     |     | 255 |     |  |  |
| Thr | Leu | Gly | Gly | Arg | Phe | Gln | Ser | Phe | Ser | Asp | Val | Leu | Pro | Asn | Asn |  |  |
|     |     |     | 260 |     |     |     |     | 265 |     |     |     |     | 270 |     |     |  |  |
| Arg | Ala | Val | Ala | Leu | Arg | Val | Pro | Arg | Glu | Asn | Asp | Glu | Asp | Met | Arg |  |  |
|     |     | 275 |     |     |     |     | 280 |     |     |     |     | 285 |     |     |     |  |  |
| Pro | Tyr | Thr | Thr | Ala | Thr | Ala | Met | Ser | Ala | Gly | Trp | Met | Trp | Thr | Ile |  |  |
|     | 290 |     |     |     |     | 295 |     |     |     |     | 300 |     |     |     |     |  |  |
| Pro | Leu | Phe | Lys | Arg | Asp | Gly | Asn | Gly | Tyr | Val | Tyr | Ser | Asp | Glu | Phe |  |  |
| 305 |     |     |     |     | 310 |     |     |     |     | 315 |     |     |     |     | 320 |  |  |
| Ile | Ser | Pro | Glu | Glu | Ala | Glu | Arg | Glu | Leu | Arg | Ser | Thr | Val | Ala | Pro |  |  |
|     |     |     |     | 325 |     |     |     |     | 330 |     |     |     |     | 335 |     |  |  |
| Gly | Arg | Asp | Asp | Leu | Glu | Ala | Asn | His | Ile | Gln | Met | Arg | Ile | Gly | Arg |  |  |
|     |     |     | 340 |     |     |     |     | 345 |     |     |     |     | 350 |     |     |  |  |
| Asn | Glu | Arg | Thr | Trp | Ile | Asn | Asn | Cys | Val | Ala | Val | Gly | Leu | Ser | Ala |  |  |

355

360

365

Ala Phe Val Glu Pro Leu Glu Ser Thr Gly Ile Phe Phe Ile Gln His  
370 375 380

Ala Ile Glu Gln Leu Val Lys His Phe Pro Gly Glu Arg Trp Asp Pro  
385 390 395 400

Val Leu Ile Ser Ala Tyr Asn Glu Arg Met Ala His Met Val Asp Gly  
405 410 415

Val Lys Glu Phe Leu Val Leu His Tyr Lys Gly Ala Gln Arg Glu Asp  
420 425 430

Thr Pro Tyr Trp Lys Ala Ala Lys Thr Arg Ala Met Pro Asp Gly Leu  
435 440 445

Ala Arg Lys Leu Glu Leu Ser Ala Ser His Leu Leu Asp Glu Gln Thr  
450 455 460

Ile Tyr Pro Tyr Tyr His Gly Phe Glu Thr Tyr Ser Trp Ile Thr Met  
465 470 475 480

Asn Leu Gly Leu Gly Ile Val Pro Glu Arg Pro Arg Pro Ala Leu Leu  
485 490 495

His Met Asp Pro Ala Pro Ala Leu Ala Glu Phe Glu Arg Leu Arg Arg  
500 505 510

Glu Gly Asp Glu Leu Ile Ala Ala Leu Pro Ser Cys Tyr Glu Tyr Leu  
515 520 525

Ala Ser Ile Gln  
530

<210> 15

<211> 1518

<212> DNA

<213> Kutzneria sp. 744

<400> 15

atgaccgccg cctacctgaa gaccgcgttc ggcgaccggc tttcgatcac cgtcgtggaa 60

tcgagcagaa tcgggacat cggcgtcggg gaggcgacct tcagcgacat ccagcacttc 120

ttccagttcc tcaacctccg ggaacaggac tggatgccgg cctgcaacgc cacctacaag 180

ctgggcatcc gcttcgagaa ctggcggcat gtcggccatc atttctacca gccgttcgag 240

cagatccgcc cggctctacgg ctttccgctg accgactggg ggctgcacga cgcgccgacc 300

gaccgcttcg acaccgactg cttcgtcatg cccaacctgt gcgaggccgg gcgcagcccg 360

cgccacctcg acggcacgct ggccgacgag gacttcgtgg aggagggcga cgagctcgcc 420

aaccgcacca tgtccgagca ccagggaag tcgcagttcc cgtacgccta ccacttcgag 480

gcggcgctgc tggccaagtt cctcaccggg tacgcggtgg accgcggcgt cgagcacgtc 540  
 gtcgacgacg tcctggacgt gcggctggac cagcgcgggg ggatcgagca cgtggtcacc 600  
 gccgagcacg gcgagatcca cggcgacctg ttctgtcact gcaccggttt ccgggggactg 660  
 ctgctgaaca aggcgctcgg tgtgccgttc gtgtcctatc aggacactct gcccaacgac 720  
 agcgcggtcg ccctccaggt cccgctggac atgcagcggc gcggcatcgt gccgaacacc 780  
 acggccaccg cgcgggaggc cggctggatc tggaccatcc cgctgttcgg tcgggtcggc 840  
 accgggtacg tctacgcaa ggactacctc tcgcccaggg aggccgagcg cacgctgcgc 900  
 gagttcgtcg gcccggcggc cgcggacgtg gaggccaacc acatccgcat gcggatcggc 960  
 cgcagccagg agtcctggcg gaacaactgc gtggccatcg gcctgtccag cggcttcgtc 1020  
 gagccgctgg agtcgaccgg gatcttcttc atccaccacg ccatcgagca gctggtgaag 1080  
 cacttcccgg cggccgactg gaaccggaag tcccgggaca tgtacaactc cgcggtcgcg 1140  
 cacgtgatgg acggcatccg cgagttcctg gtgatccact accggggcgc ggcgcgggcc 1200  
 gacaaccagt actggcgcga caccaagacc cggccgctgc cggacggcct ggccgagcgc 1260  
 atcgagtgtc ggagacgca gctgccggac accgagacga tctaccgta ctaccacggg 1320  
 ctgccgccct actcctacat gtgcatcctg atgggcggcg gggccatccg gacgccggcc 1380  
 tcggccgcgt tggccctgac cgaccagggg gccgcgcaga aggagtctgc cgccgtccgg 1440  
 gaccgggccc cgagctgcg cgacacgtg cccagccact acgagtacct cgcgcggtatg 1500  
 cgtggtctgg atgtctga 1518

<210> 16  
 <211> 505  
 <212> PRT  
 <213> Kutzneria sp. 744

<400> 16

Met Thr Ala Ala Tyr Leu Lys Thr Ala Phe Gly Asp Arg Leu Ser Ile  
1 5 10 15

Thr Val Val Glu Ser Ser Arg Ile Gly Thr Ile Gly Val Gly Glu Ala  
20 25 30

Thr Phe Ser Asp Ile Gln His Phe Phe Gln Phe Leu Asn Leu Arg Glu  
35 40 45

Gln Asp Trp Met Pro Ala Cys Asn Ala Thr Tyr Lys Leu Gly Ile Arg  
50 55 60

Phe Glu Asn Trp Arg His Val Gly His His Phe Tyr Gln Pro Phe Glu  
65 70 75 80

Gln Ile Arg Pro Val Tyr Gly Phe Pro Leu Thr Asp Trp Trp Leu His  
85 90 95

Asp Ala Pro Thr Asp Arg Phe Asp Thr Asp Cys Phe Val Met Pro Asn

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |  |  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|--|--|
| 100 | 105 | 110 |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |  |  |
| Leu | Cys | Glu | Ala | Gly | Arg | Ser | Pro | Arg | His | Leu | Asp | Gly | Thr | Leu | Ala |  |  |  |  |
|     | 115 |     |     |     |     |     | 120 |     |     |     |     | 125 |     |     |     |  |  |  |  |
| Asp | Glu | Asp | Phe | Val | Glu | Glu | Gly | Asp | Glu | Leu | Ala | Asn | Arg | Thr | Met |  |  |  |  |
|     | 130 |     |     |     |     | 135 |     |     |     |     | 140 |     |     |     |     |  |  |  |  |
| Ser | Glu | His | Gln | Gly | Lys | Ser | Gln | Phe | Pro | Tyr | Ala | Tyr | His | Phe | Glu |  |  |  |  |
| 145 |     |     |     |     | 150 |     |     |     |     | 155 |     |     |     |     | 160 |  |  |  |  |
| Ala | Ala | Leu | Leu | Ala | Lys | Phe | Leu | Thr | Gly | Tyr | Ala | Val | Asp | Arg | Gly |  |  |  |  |
|     |     |     |     | 165 |     |     |     |     | 170 |     |     |     |     | 175 |     |  |  |  |  |
| Val | Glu | His | Val | Val | Asp | Asp | Val | Leu | Asp | Val | Arg | Leu | Asp | Gln | Arg |  |  |  |  |
|     |     |     | 180 |     |     |     |     | 185 |     |     |     |     | 190 |     |     |  |  |  |  |
| Gly | Trp | Ile | Glu | His | Val | Val | Thr | Ala | Glu | His | Gly | Glu | Ile | His | Gly |  |  |  |  |
|     |     | 195 |     |     |     |     | 200 |     |     |     |     | 205 |     |     |     |  |  |  |  |
| Asp | Leu | Phe | Val | Asp | Cys | Thr | Gly | Phe | Arg | Gly | Leu | Leu | Leu | Asn | Lys |  |  |  |  |
|     | 210 |     |     |     |     | 215 |     |     |     |     | 220 |     |     |     |     |  |  |  |  |
| Ala | Leu | Gly | Val | Pro | Phe | Val | Ser | Tyr | Gln | Asp | Thr | Leu | Pro | Asn | Asp |  |  |  |  |
| 225 |     |     |     |     | 230 |     |     |     |     | 235 |     |     |     |     | 240 |  |  |  |  |
| Ser | Ala | Val | Ala | Leu | Gln | Val | Pro | Leu | Asp | Met | Gln | Arg | Arg | Gly | Ile |  |  |  |  |
|     |     |     |     | 245 |     |     |     |     | 250 |     |     |     |     | 255 |     |  |  |  |  |
| Val | Pro | Asn | Thr | Thr | Ala | Thr | Ala | Arg | Glu | Ala | Gly | Trp | Ile | Trp | Thr |  |  |  |  |
|     |     |     | 260 |     |     |     |     | 265 |     |     |     |     | 270 |     |     |  |  |  |  |
| Ile | Pro | Leu | Phe | Gly | Arg | Val | Gly | Thr | Gly | Tyr | Val | Tyr | Ala | Lys | Asp |  |  |  |  |
|     | 275 |     |     |     |     |     | 280 |     |     |     |     | 285 |     |     |     |  |  |  |  |
| Tyr | Leu | Ser | Pro | Glu | Glu | Ala | Glu | Arg | Thr | Leu | Arg | Glu | Phe | Val | Gly |  |  |  |  |
|     | 290 |     |     |     |     | 295 |     |     |     |     | 300 |     |     |     |     |  |  |  |  |
| Pro | Ala | Ala | Ala | Asp | Val | Glu | Ala | Asn | His | Ile | Arg | Met | Arg | Ile | Gly |  |  |  |  |
| 305 |     |     |     |     | 310 |     |     |     |     | 315 |     |     |     |     | 320 |  |  |  |  |
| Arg | Ser | Gln | Glu | Ser | Trp | Arg | Asn | Asn | Cys | Val | Ala | Ile | Gly | Leu | Ser |  |  |  |  |
|     |     |     |     | 325 |     |     |     |     | 330 |     |     |     |     | 335 |     |  |  |  |  |
| Ser | Gly | Phe | Val | Glu | Pro | Leu | Glu | Ser | Thr | Gly | Ile | Phe | Phe | Ile | His |  |  |  |  |
|     |     |     | 340 |     |     |     |     | 345 |     |     |     |     | 350 |     |     |  |  |  |  |
| His | Ala | Ile | Glu | Gln | Leu | Val | Lys | His | Phe | Pro | Ala | Ala | Asp | Trp | Asn |  |  |  |  |
|     |     | 355 |     |     |     |     | 360 |     |     |     |     | 365 |     |     |     |  |  |  |  |
| Pro | Lys | Ser | Arg | Asp | Met | Tyr | Asn | Ser | Ala | Val | Ala | His | Val | Met | Asp |  |  |  |  |
|     | 370 |     |     |     |     | 375 |     |     |     |     | 380 |     |     |     |     |  |  |  |  |

Gly Ile Arg Glu Phe Leu Val Ile His Tyr Arg Gly Ala Ala Arg Ala  
385 390 395 400

Asp Asn Gln Tyr Trp Arg Asp Thr Lys Thr Arg Pro Leu Pro Asp Gly  
405 410 415

Leu Ala Glu Arg Ile Glu Cys Trp Gln Thr Gln Leu Pro Asp Thr Glu  
420 425 430

Thr Ile Tyr Pro Tyr Tyr His Gly Leu Pro Pro Tyr Ser Tyr Met Cys  
435 440 445

Ile Leu Met Gly Gly Gly Ala Ile Arg Thr Pro Ala Ser Ala Ala Leu  
450 455 460

Ala Leu Thr Asp Gln Gly Ala Ala Gln Lys Glu Phe Ala Ala Val Arg  
465 470 475 480

Asp Arg Ala Ala Gln Leu Arg Asp Thr Leu Pro Ser His Tyr Glu Tyr  
485 490 495

Leu Ala Arg Met Arg Gly Leu Asp Val  
500 505

<210> 17  
<211> 1572  
<212> DNA  
<213> Streptomyces toxytrinic

<400> 17  
atgaacacga gaaatccgga caaggtcgtc atcgtcggcg gcggcacggc gggatggatg 60  
accgcctcgt acctgaagaa ggcgttcggc gagcgcgtgt ccgtgaccct cgtcgaatcg 120  
ggcaccatcg gcaccgtcgg cgtcggcgag gcgaccttca gcgacatccg gcacttcttc 180  
gaattcctcg acctgcggga ggaggaatgg atgccggcct gcaacgccac gtacaagctg 240  
gccgtccgct tccaggactg gcagcgcccc ggccaccact tctaccacc gttcgagcag 300  
atgcgctccg tagacgggtt ccccttgacc gactggtggc tgcagaacgg gccgaccgac 360  
cggttcgaca gggactgctt cgtgatggcg tcgctgtgcg acgcggggcg cagcccgcgc 420  
tatctcaacg gttccctgct gcagcaggag ttcgacgagc gcgcggagga gccggcgggg 480  
ctcaccatgt ccgagcacca gggcaagacg cagttcccct acgcctacca cttcgaggcg 540  
gccctgctcg ccgaattcct ttctgggtac tccaaggacc gcggcgtgaa gcacgtggtc 600  
gacgaggtgc tggaggtgaa gctcgacgac cgcggatgga tcagccacgt tgtcaccaag 660  
gagcacggcg acatcggcgg cgacctgttc gtcgactgca ccggtttccg cggcgtactg 720  
ctgaaccagg cgctgggtgt gcccttcgtc tcctaccagg acaccctgcc caatgacagc 780  
gcggtcgccc tccagggtgcc gtcgacatg gaggcgcggg gcatcccgcc gtacacccgg 840  
gccaccgcca aggaggccgg gtggatctgg acgatccgc tgatcggccg gatcggcacc 900

|                                                                     |      |
|---------------------------------------------------------------------|------|
| gggtacgtgt acgccaagga ctactgctcg ccggaggagg ccgagcgac cctgcgcgag    | 960  |
| ttcgtcggtc cggaggccgc cgacgtcgag gccaaaccaca tccgcatgcg gatcgggctgc | 1020 |
| agcgagcagt cctggaagaa caactgcgtc gcgacgaggc tctccagtgg cttcgtcgaa   | 1080 |
| cccctcgagt cgaccggcat cttcttcac caccacgcca tcgagcagtt ggtgaagcac    | 1140 |
| ttcccgccgg gcgactggca cccgcagctg cgcgccggct acaactcggc cgtcgccaac   | 1200 |
| gtcatggacg ggggccggga gttcctcgtc ctgcactacc tgggcgccgc gcgcaacgac   | 1260 |
| acccgggtact ggaaggacac caagaccgg gcggttcgg atgccctggc ggagcgcatc    | 1320 |
| gagcgctgga aggtgcagct gcccgactcc gagaacgtct tcccctacta ccacggcctg   | 1380 |
| ccgccgtact cgtacatggc catcctgctc gggaccggcg cgatcgggct gcgaccgtcg   | 1440 |
| ccggcgctcg cactggccga cccggcgccg gcggagaagg agttcaccgc gatccgggac   | 1500 |
| agggcgcggt tcctggctga caccctgcc agccagtac agtacttcgc cgcgatgggg     | 1560 |
| cagcgtgtct ga                                                       | 1572 |

<210> 18  
 <211> 523  
 <212> PRT  
 <213> Streptomyces toxytrinic

<400> 18

|                                                                 |  |
|-----------------------------------------------------------------|--|
| Met Asn Thr Arg Asn Pro Asp Lys Val Val Ile Val Gly Gly Gly Thr |  |
| 1 5 10 15                                                       |  |

|                                                                 |  |
|-----------------------------------------------------------------|--|
| Ala Gly Trp Met Thr Ala Ser Tyr Leu Lys Lys Ala Phe Gly Glu Arg |  |
| 20 25 30                                                        |  |

|                                                                 |  |
|-----------------------------------------------------------------|--|
| Val Ser Val Thr Leu Val Glu Ser Gly Thr Ile Gly Thr Val Gly Val |  |
| 35 40 45                                                        |  |

|                                                                 |  |
|-----------------------------------------------------------------|--|
| Gly Glu Ala Thr Phe Ser Asp Ile Arg His Phe Phe Glu Phe Leu Asp |  |
| 50 55 60                                                        |  |

|                                                                 |  |
|-----------------------------------------------------------------|--|
| Leu Arg Glu Glu Glu Trp Met Pro Ala Cys Asn Ala Thr Tyr Lys Leu |  |
| 65 70 75 80                                                     |  |

|                                                                 |  |
|-----------------------------------------------------------------|--|
| Ala Val Arg Phe Gln Asp Trp Gln Arg Pro Gly His His Phe Tyr His |  |
| 85 90 95                                                        |  |

|                                                                 |  |
|-----------------------------------------------------------------|--|
| Pro Phe Glu Gln Met Arg Ser Val Asp Gly Phe Pro Leu Thr Asp Trp |  |
| 100 105 110                                                     |  |

|                                                                 |  |
|-----------------------------------------------------------------|--|
| Trp Leu Gln Asn Gly Pro Thr Asp Arg Phe Asp Arg Asp Cys Phe Val |  |
| 115 120 125                                                     |  |

|                                                                 |  |
|-----------------------------------------------------------------|--|
| Met Ala Ser Leu Cys Asp Ala Gly Arg Ser Pro Arg Tyr Leu Asn Gly |  |
| 130 135 140                                                     |  |

Ser Leu Leu Gln Gln Glu Phe Asp Glu Arg Ala Glu Glu Pro Ala Gly  
145 150 155 160

Leu Thr Met Ser Glu His Gln Gly Lys Thr Gln Phe Pro Tyr Ala Tyr  
165 170 175

His Phe Glu Ala Ala Leu Leu Ala Glu Phe Leu Ser Gly Tyr Ser Lys  
180 185 190

Asp Arg Gly Val Lys His Val Val Asp Glu Val Leu Glu Val Lys Leu  
195 200 205

Asp Asp Arg Gly Trp Ile Ser His Val Val Thr Lys Glu His Gly Asp  
210 215 220

Ile Gly Gly Asp Leu Phe Val Asp Cys Thr Gly Phe Arg Gly Val Leu  
225 230 235 240

Leu Asn Gln Ala Leu Gly Val Pro Phe Val Ser Tyr Gln Asp Thr Leu  
245 250 255

Pro Asn Asp Ser Ala Val Ala Leu Gln Val Pro Leu Asp Met Glu Ala  
260 265 270

Arg Gly Ile Pro Pro Tyr Thr Arg Ala Thr Ala Lys Glu Ala Gly Trp  
275 280 285

Ile Trp Thr Ile Pro Leu Ile Gly Arg Ile Gly Thr Gly Tyr Val Tyr  
290 295 300

Ala Lys Asp Tyr Cys Ser Pro Glu Glu Ala Glu Arg Thr Leu Arg Glu  
305 310 315 320

Phe Val Gly Pro Glu Ala Ala Asp Val Glu Ala Asn His Ile Arg Met  
325 330 335

Arg Ile Gly Arg Ser Glu Gln Ser Trp Lys Asn Asn Cys Val Ala Ile  
340 345 350

Gly Leu Ser Ser Gly Phe Val Glu Pro Leu Glu Ser Thr Gly Ile Phe  
355 360 365

Phe Ile His His Ala Ile Glu Gln Leu Val Lys His Phe Pro Ala Gly  
370 375 380

Asp Trp His Pro Gln Leu Arg Ala Gly Tyr Asn Ser Ala Val Ala Asn  
385 390 395 400

Val Met Asp Gly Val Arg Glu Phe Leu Val Leu His Tyr Leu Gly Ala  
405 410 415

Ala Arg Asn Asp Thr Arg Tyr Trp Lys Asp Thr Lys Thr Arg Ala Val

420

425

430

Pro Asp Ala Leu Ala Glu Arg Ile Glu Arg Trp Lys Val Gln Leu Pro  
 435 440 445

Asp Ser Glu Asn Val Phe Pro Tyr Tyr His Gly Leu Pro Pro Tyr Ser  
 450 455 460

Tyr Met Ala Ile Leu Leu Gly Thr Gly Ala Ile Gly Leu Arg Pro Ser  
 465 470 475 480

Pro Ala Leu Ala Leu Ala Asp Pro Ala Ala Ala Glu Lys Glu Phe Thr  
 485 490 495

Ala Ile Arg Asp Arg Ala Arg Phe Leu Val Asp Thr Leu Pro Ser Gln  
 500 505 510

Tyr Glu Tyr Phe Ala Ala Met Gly Gln Arg Val  
 515 520

<210> 19  
 <211> 1596  
 <212> DNA  
 <213> Streptomyces albogriseolus

<400> 19  
 atggacaatc gaatcaagac agtcgtgata ctgggcggcg gcaccgccgg ctggatgact 60  
 gcggcgtatc tcggcaaggc gctgcagaac acggtgaaga tcgtggctct ggaggcgccc 120  
 accatcccg c gcatcgggtgt gggcgaagcc accgtcccca acctgcagcg cgcgttcttc 180  
 gactacctcg gcatccccga ggaggagtgg atgcgcgagt gcaacgccag ttacaagatg 240  
 gcggtgaagt tcatcaactg gcgcacaccg ggggagggca gcccgatcc ccggacgctc 300  
 gacgacggtc acaccgacac cttccaccac cccttcggac tgctgccgtc cgccgaccag 360  
 ataccgctct cccactactg ggccggccaag cggctccagg gcgagaccga cgagaacttc 420  
 gacgaggcgt gtttcgccga caccgcgata atgaacgcca agaaggcgcc ccgcttcctc 480  
 gacatgcggc gggcgaccaa ctacgcctgg cacttcgacg ccagcaaggt cgccgcgttc 540  
 ctgcgcaact tcgccgtcac caagcaggcg gtcgagcacg tcgaggacga gatgaccgag 600  
 gtcctcaccg acgaacgggg cttcatcacc gccctgcgca ccaagtcggg acggatcctg 660  
 cagggtgacc ttttcgtcga ctgctccggc ttccgcgggc tgctgatcaa caaggccatg 720  
 gaggagccgt tcatcgacat gagcgaccac ctctgtgca acagcgcggt cgccacggcc 780  
 gtaccgcacg acgacgagaa gaacgggtgtc gagccgtaca cctcctcgat cgccatggag 840  
 gccggctgga cctggaagat tcccatgctg ggccgggttcg gcagcggtca cgtctactcc 900  
 gaccacttcg ccaccagga caggccacc ctggccttct cgaagctgtg gggcctcgac 960  
 ccggacaaca ccgagttcaa ccacgtccgc ttccgggtcg gccgaaccg cagggcctgg 1020  
 gtgcgcaact gcgtcagcgt cggcctcgcg tcgtgcttcg tggagccgct ggagtccagc 1080



|            |            |             |            |            |            |      |
|------------|------------|-------------|------------|------------|------------|------|
| ggcatctact | tcatctacgc | ggccatccac  | atgctggcca | agcacttccc | cgacaagacc | 1140 |
| ttcgacaagg | tgctcgtcga | ccggttcaac  | cgggagatcg | aggagatgtt | cgacgacacc | 1200 |
| cgggacttcc | tccaggcgca | ctactacttc  | tcgccgcgcg | tcgacacgcc | gttctggcgg | 1260 |
| gcgaacaagg | aactgaagct | ggccgactcc  | atcaaggaca | aggtcgagac | gtaccggggc | 1320 |
| gggctgccgg | tcaacctgcc | ggtcaccgac  | gaggggacct | actacggcaa | cttcgaggcc | 1380 |
| gagttccgga | acttctggac | caacggcagc  | tactactgca | tcttcgccgg | cctcggcctg | 1440 |
| atgccgcgca | acccgctgcc | cgcgctcgcc  | tacaagccgc | agtcgatcgc | ggaggccgag | 1500 |
| ctgctgttcg | ccgacgtcaa | gcgcaagggg  | gacaccctgg | tcgagtcgct | gccgtcgacc | 1560 |
| tatgacctgc | tgcgtcagtt | gcacgggtgcg | tcgtga     |            |            | 1596 |

```
<210> 20
<211> 531
<212> PRT
<213> Streptomyces albobriseolus
```

<400> 20

Met Asp Asn Arg Ile Lys Thr Val Val Ile Leu Gly Gly Gly Thr Ala  
1 5 10 15

Gly Trp Met Thr Ala Ala Tyr Leu Gly Lys Ala Leu Gln Asn Thr Val  
20 25 30

Lys Ile Val Val Leu Glu Ala Pro Thr Ile Pro Arg Ile Gly Val Gly  
35 40 45

Glu Ala Thr Val Pro Asn Leu Gln Arg Ala Phe Phe Asp Tyr Leu Gly  
50 55 60

Ile Pro Glu Glu Glu Trp Met Arg Glu Cys Asn Ala Ser Tyr Lys Met  
65 70 75 80

Ala Val Lys Phe Ile Asn Trp Arg Thr Pro Gly Glu Gly Ser Pro Asp  
85 90 95

Pro Arg Thr Leu Asp Asp Gly His Thr Asp Thr Phe His His Pro Phe  
100 105 110

Gly Leu Leu Pro Ser Ala Asp Gln Ile Pro Leu Ser His Tyr Trp Ala  
115 120 125

Ala Lys Arg Leu Gln Gly Glu Thr Asp Glu Asn Phe Asp Glu Ala Cys  
130 135 140

Phe Ala Asp Thr Ala Ile Met Asn Ala Lys Lys Ala Pro Arg Phe Leu  
145 150 155 160

Asp Met Arg Arg Ala Thr Asn Tyr Ala Trp His Phe Asp Ala Ser Lys  
165 170 175

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Val | Ala | Ala | Phe | Leu | Arg | Asn | Phe | Ala | Val | Thr | Lys | Gln | Ala | Val | Glu | 180 | 185 | 190 |     |
| His | Val | Glu | Asp | Glu | Met | Thr | Glu | Val | Leu | Thr | Asp | Glu | Arg | Gly | Phe | 195 | 200 | 205 |     |
| Ile | Thr | Ala | Leu | Arg | Thr | Lys | Ser | Gly | Arg | Ile | Leu | Gln | Gly | Asp | Leu | 210 | 215 | 220 |     |
| Phe | Val | Asp | Cys | Ser | Gly | Phe | Arg | Gly | Leu | Leu | Ile | Asn | Lys | Ala | Met | 225 | 230 | 235 | 240 |
| Glu | Glu | Pro | Phe | Ile | Asp | Met | Ser | Asp | His | Leu | Leu | Cys | Asn | Ser | Ala | 245 | 250 | 255 |     |
| Val | Ala | Thr | Ala | Val | Pro | His | Asp | Asp | Glu | Lys | Asn | Gly | Val | Glu | Pro | 260 | 265 | 270 |     |
| Tyr | Thr | Ser | Ser | Ile | Ala | Met | Glu | Ala | Gly | Trp | Thr | Trp | Lys | Ile | Pro | 275 | 280 | 285 |     |
| Met | Leu | Gly | Arg | Phe | Gly | Ser | Gly | His | Val | Tyr | Ser | Asp | His | Phe | Ala | 290 | 295 | 300 |     |
| Thr | Gln | Asp | Glu | Ala | Thr | Leu | Ala | Phe | Ser | Lys | Leu | Trp | Gly | Leu | Asp | 305 | 310 | 315 | 320 |
| Pro | Asp | Asn | Thr | Glu | Phe | Asn | His | Val | Arg | Phe | Arg | Val | Gly | Arg | Asn | 325 | 330 | 335 |     |
| Arg | Arg | Ala | Trp | Val | Arg | Asn | Cys | Val | Ser | Val | Gly | Leu | Ala | Ser | Cys | 340 | 345 | 350 |     |
| Phe | Val | Glu | Pro | Leu | Glu | Ser | Ser | Gly | Ile | Tyr | Phe | Ile | Tyr | Ala | Ala | 355 | 360 | 365 |     |
| Ile | His | Met | Leu | Ala | Lys | His | Phe | Pro | Asp | Lys | Thr | Phe | Asp | Lys | Val | 370 | 375 | 380 |     |
| Leu | Val | Asp | Arg | Phe | Asn | Arg | Glu | Ile | Glu | Glu | Met | Phe | Asp | Asp | Thr | 385 | 390 | 395 | 400 |
| Arg | Asp | Phe | Leu | Gln | Ala | His | Tyr | Tyr | Phe | Ser | Pro | Arg | Val | Asp | Thr | 405 | 410 | 415 |     |
| Pro | Phe | Trp | Arg | Ala | Asn | Lys | Glu | Leu | Lys | Leu | Ala | Asp | Ser | Ile | Lys | 420 | 425 | 430 |     |
| Asp | Lys | Val | Glu | Thr | Tyr | Arg | Ala | Gly | Leu | Pro | Val | Asn | Leu | Pro | Val | 435 | 440 | 445 |     |

Thr Asp Glu Gly Thr Tyr Tyr Gly Asn Phe Glu Ala Glu Phe Arg Asn  
450 455 460

Phe Trp Thr Asn Gly Ser Tyr Tyr Cys Ile Phe Ala Gly Leu Gly Leu  
465 470 475 480

Met Pro Arg Asn Pro Leu Pro Ala Leu Ala Tyr Lys Pro Gln Ser Ile  
485 490 495

Ala Glu Ala Glu Leu Leu Phe Ala Asp Val Lys Arg Lys Gly Asp Thr  
500 505 510

Leu Val Glu Ser Leu Pro Ser Thr Tyr Asp Leu Leu Arg Gln Leu His  
515 520 525

Gly Ala Ser  
530

<210> 21  
<211> 1590  
<212> DNA  
<213> BorH halogenase

<400> 21  
atggacaacc gcatcaaccg gatcgtcatc ctcggcggcg ggaccgcggg atggatgacc 60  
gcctcctacc tggcgaaggc gctcggtgac accgtcacga tcaccctgct ggaagcaccg 120  
gcgatcggcc ggatcgggggt cggcgaggcc accgtgccca acctgcagcg cgtgttcttc 180  
gacttcctcg gcctccgcga ggaggagtgg atgccggagt gcaacgccgc gttcaagacg 240  
gccgtgaagt tcatacaactg gcgtaccccc gggcccggcg aggccaaggc ccgcacgac 300  
gacggccggc cggaccactt ctaccacccg ttcgggctgc tgcccagca cggccagggtg 360  
ccgctgtcgc actactgggc ctacaaccgg gcggccggca ccaccgacga gccgttcgac 420  
tacgcctgtt tcgccgagac cgccgccatg gacgcggtcc gcgcgcccga gtggctcgac 480  
gggcgggccc ccaccggta tgcttggcac ttcgacgccc acctggctgc cgagttcctg 540  
cgccggcacg ccaccgagcg gctcaacgtg gacgacgtgc aggggtgagat gcagcagggtg 600  
ctgcgcgacg agcggggctt catcacccgc ctgcgcacgg tcgagggccg ggacctcgag 660  
ggtgacctgt tcatcgactg ctccggcttc cgcggcctgc tgatcaacaa ggccatggag 720  
gagcccttca tcgacatgaa cgaccagctg ctgtgcaacc gcgcggtggc cacggcgcac 780  
aagcacgacg acgacgcgca cggggctcgag ccgtacacct cggcgatcgc catgcggtcc 840  
ggctggagct ggaagatccc gatgctgggc cgcttcggga ccggctacgt ctactccagc 900  
cggttcgccg agaaggacga ggccaccctc gacttctgcc ggatgtgggg cctggaccgg 960  
gagaacaccc cgctcaacca ggtggccttc cgggtcggcc gcaaccgccg ggcctgggtg 1020  
aagaactgcg tgagcatcgg cctggcgctg tgcttcctgg agccactgga gtccaccggc 1080  
atctacttca tcaccgggc catctaccag ctgactcagc acttcccgga ccgcaccttc 1140  
gcgctcgcgt tgagcgacgc gttcaaccac gagatcgaag ccatgttcga cgacacccgg 1200

gacttcatcc aggcgcactt ctacgtgtcg ccgcgcaccg acacgccgtt ctggaaggcg 1260  
aacaaggacc tgcacctgcc cgagcagatg cgcgagaaga tcgcgatgta caaggccggc 1320  
ctgccgatca acgcgccggt caccgacgag tcgacctact acggcagggt cgaggccgag 1380  
ttccgcaact tctggaccaa cggcagctac tactgcatct tcgcgggcct cgggctgcgc 1440  
ccggacaacc cgctgccgat gctgcgacac cgtccggagc aggtccgcga ggcgcaggcg 1500  
ctgttcgccg gggatgaagga caagcagcgg gagttggtgg agacgctgcc gagcaacctg 1560  
gagttcctgc gcagcctgca cggcaagtag 1590

<210> 22  
<211> 529  
<212> PRT  
<213> BorH halogenase  
  
<400> 22

Met Asp Asn Arg Ile Asn Arg Ile Val Ile Leu Gly Gly Gly Thr Ala  
1 5 10 15  
  
Gly Trp Met Thr Ala Ser Tyr Leu Ala Lys Ala Leu Gly Asp Thr Val  
20 25 30  
  
Thr Ile Thr Leu Leu Glu Ala Pro Ala Ile Gly Arg Ile Gly Val Gly  
35 40 45  
  
Glu Ala Thr Val Pro Asn Leu Gln Arg Val Phe Phe Asp Phe Leu Gly  
50 55 60  
  
Leu Arg Glu Glu Glu Trp Met Pro Glu Cys Asn Ala Ala Phe Lys Thr  
65 70 75 80  
  
Ala Val Lys Phe Ile Asn Trp Arg Thr Pro Gly Pro Gly Glu Ala Lys  
85 90 95  
  
Ala Arg Thr Ile Asp Gly Arg Pro Asp His Phe Tyr His Pro Phe Gly  
100 105 110  
  
Leu Leu Pro Glu His Gly Gln Val Pro Leu Ser His Tyr Trp Ala Tyr  
115 120 125  
  
Asn Arg Ala Ala Gly Thr Thr Asp Glu Pro Phe Asp Tyr Ala Cys Phe  
130 135 140  
  
Ala Glu Thr Ala Ala Met Asp Ala Val Arg Ala Pro Lys Trp Leu Asp  
145 150 155 160  
  
Gly Arg Pro Ala Thr Arg Tyr Ala Trp His Phe Asp Ala His Leu Val  
165 170 175  
  
Ala Glu Phe Leu Arg Arg His Ala Thr Glu Arg Leu Asn Val Glu His  
180 185 190

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|
| Val | Gln | Gly | Glu | Met | Gln | Gln | Val | Leu | Arg | Asp | Glu | Arg | Gly | Phe | Ile |  |  |
|     | 195 |     |     |     |     |     | 200 |     |     |     |     | 205 |     |     |     |  |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |
| Thr | Ala | Leu | Arg | Thr | Val | Glu | Gly | Arg | Asp | Leu | Glu | Gly | Asp | Leu | Phe |  |  |
|     | 210 |     |     |     |     | 215 |     |     |     |     | 220 |     |     |     |     |  |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |
| Ile | Asp | Cys | Ser | Gly | Phe | Arg | Gly | Leu | Leu | Ile | Asn | Lys | Ala | Met | Glu |  |  |
| 225 |     |     |     |     | 230 |     |     |     |     | 235 |     |     |     |     | 240 |  |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |
| Glu | Pro | Phe | Ile | Asp | Met | Asn | Asp | Gln | Leu | Leu | Cys | Asn | Arg | Ala | Val |  |  |
|     |     |     |     | 245 |     |     |     |     | 250 |     |     |     |     | 255 |     |  |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |
| Ala | Thr | Ala | Ile | Lys | His | Asp | Asp | Asp | Ala | His | Gly | Val | Glu | Pro | Tyr |  |  |
|     |     |     | 260 |     |     |     |     | 265 |     |     |     |     | 270 |     |     |  |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |
| Thr | Ser | Ala | Ile | Ala | Met | Arg | Ser | Gly | Trp | Ser | Trp | Lys | Ile | Pro | Met |  |  |
|     | 275 |     |     |     |     |     | 280 |     |     |     |     | 285 |     |     |     |  |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |
| Leu | Gly | Arg | Phe | Gly | Thr | Gly | Tyr | Val | Tyr | Ser | Ser | Arg | Phe | Ala | Glu |  |  |
|     | 290 |     |     |     |     | 295 |     |     |     |     | 300 |     |     |     |     |  |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |
| Lys | Asp | Glu | Ala | Thr | Leu | Asp | Phe | Cys | Arg | Met | Trp | Gly | Leu | Asp | Pro |  |  |
| 305 |     |     |     |     | 310 |     |     |     |     | 315 |     |     |     |     | 320 |  |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |
| Glu | Asn | Thr | Pro | Leu | Asn | Gln | Val | Ala | Phe | Arg | Val | Gly | Arg | Asn | Arg |  |  |
|     |     |     |     | 325 |     |     |     |     | 330 |     |     |     |     | 335 |     |  |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |
| Arg | Ala | Trp | Val | Lys | Asn | Cys | Val | Ser | Ile | Gly | Leu | Ala | Ser | Cys | Phe |  |  |
|     |     |     | 340 |     |     |     |     | 345 |     |     |     |     | 350 |     |     |  |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |
| Leu | Glu | Pro | Leu | Glu | Ser | Thr | Gly | Ile | Tyr | Phe | Ile | Thr | Ala | Ala | Ile |  |  |
|     | 355 |     |     |     |     |     | 360 |     |     |     |     | 365 |     |     |     |  |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |
| Tyr | Gln | Leu | Thr | Gln | His | Phe | Pro | Asp | Arg | Thr | Phe | Ala | Leu | Ala | Leu |  |  |
|     | 370 |     |     |     |     | 375 |     |     |     |     | 380 |     |     |     |     |  |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |
| Ser | Asp | Ala | Phe | Asn | His | Glu | Ile | Glu | Ala | Met | Phe | Asp | Asp | Thr | Arg |  |  |
| 385 |     |     |     |     | 390 |     |     |     |     | 395 |     |     |     |     | 400 |  |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |
| Asp | Phe | Ile | Gln | Ala | His | Phe | Tyr | Val | Ser | Pro | Arg | Thr | Asp | Thr | Pro |  |  |
|     |     |     | 405 |     |     |     |     |     | 410 |     |     |     |     | 415 |     |  |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |
| Phe | Trp | Lys | Ala | Asn | Lys | Asp | Leu | His | Leu | Pro | Glu | Gln | Met | Arg | Glu |  |  |
|     |     |     | 420 |     |     |     |     | 425 |     |     |     |     | 430 |     |     |  |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |
| Lys | Ile | Ala | Met | Tyr | Lys | Ala | Gly | Leu | Pro | Ile | Asn | Ala | Pro | Val | Thr |  |  |
|     | 435 |     |     |     |     |     | 440 |     |     |     |     | 445 |     |     |     |  |  |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |
| Asp | Glu | Ser | Thr | Tyr | Tyr | Gly | Arg | Phe | Glu | Ala | Glu | Phe | Arg | Asn | Phe |  |  |
|     | 450 |     |     |     |     | 455 |     |     |     |     | 460 |     |     |     |     |  |  |

Trp Thr Asn Gly Ser Tyr Tyr Cys Ile Phe Ala Gly Leu Gly Leu Arg  
465 470 475 480

Pro Asp Asn Pro Leu Pro Met Leu Arg His Arg Pro Glu Gln Val Arg  
485 490 495

Glu Ala Gln Ala Leu Phe Ala Gly Val Lys Asp Lys Gln Arg Glu Leu  
500 505 510

Val Glu Thr Leu Pro Ser Asn Leu Glu Phe Leu Arg Ser Leu His Gly  
515 520 525

Lys

<210> 23  
<211> 1602  
<212> DNA  
<213> Kutzneria sp. 744

<400> 23  
atggatgaca atcgaattcg gagcatcctt gtccttggcg gcggcacggc cggctggatg 60  
tccgcctgct acctgagcaa ggcgctcggg cccggcgctcg aggtcaccgt gctcgaggcg 120  
ccctccatct cgcgcatccg ggtcggcgag gccaccattc ccaacctgca caaggtcttc 180  
ttcgacttcc tgggcatcgc cgaggacgag tggatgcggg agtgcaacgc cagctacaag 240  
gccgcggtcc ggttcgtcaa ctggcggacg ccgggcgacg gccaggccac gccgcggcgg 300  
cgtccggacg gccgccccga ccacttcgac cacctgttcg gccagctgcc cgagcacgag 360  
aacctgccgc tgtcgcagta ctgggcgcac cggcgcctca acggcctgac cgacgaaccc 420  
ttcgaccgct cctgctacgt gcagcccagag ctgctggacc gcaagctctc gccgaggttg 480  
atggacggca cgaactggc cagctacgcc tggcacttcg acgccgacct ggtggccgac 540  
ttcctctgcc ggttcgccgt gcagaagctg aacgtgacct acgtccagga cgtgttcacg 600  
catgccgacc tcgaccagcg cggccacatc acggccgtca acaccgagtc cggccgcacg 660  
ctggccgccg acctgttcac cgactgcagc ggcttccgca gcgtgctcat gggcaaggctc 720  
atgcaggagc cgttcctgga catgagcaag cacctgctca acgaccgcgc ggtggcgctg 780  
atgctccgc acgacgacga gaaggctggg atcgagccgt acacctcgtc gctggccatg 840  
cggtcgggct ggtcgtggaa gatcccgtg ctcggccggt tcggctccgg ctacgtctac 900  
tccagccagt tcacctcca ggacgaggcg gccgaggagc tctgccgcat gtgggacgtc 960  
gacccggcgg agcagacgtt caacaacgtc cggttccggg tcggccgcag ccgccgggcc 1020  
tgggtgcgca actgcgtcgc catcggcgtg tccgccatgt tcgtggagcc gctggagtcg 1080  
accggcctgt acttcagtta cgctcgctc taccagctgg tgaagcactt cccggacaag 1140  
cggttccggc cgatcctggc cgaccggttc aaccgcgagg tggcgaccat gtacgacgac 1200  
acccgcgact tcctccaggc gcacttcagc ctgtcgccgc gtgacgactc cgagttcttg 1260

|                                                                    |      |
|--------------------------------------------------------------------|------|
| cgggcctgca aggagctgcc gttcgcggac gggttcgcgc agaaggtcga gatgtacagg  | 1320 |
| gccgggctgc cggtcgaact gccggtcacc atcgacgacg ggcactacta cggcaatttc  | 1380 |
| gaggccgagt tccgcaactt ctggaccaac tcgaactact actgcatctt cgccgggctc  | 1440 |
| ggtttcctgc ccgagcatcc gctgccggtg ctcgaattcc gcccgaggagc cgtcgatcgc | 1500 |
| gcgagagccgg tgttcgccgc ggtgcgccgg cgcacggagg agctggtcgc caccgccccg | 1560 |
| accatgcagg cctacctgcg gcgcctgcac cagggcacgt ag                     | 1602 |

<210> 24  
 <211> 533  
 <212> PRT  
 <213> Kutzneria sp. 744

<400> 24

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asp | Asp | Asn | Arg | Ile | Arg | Ser | Ile | Leu | Val | Leu | Gly | Gly | Gly | Thr |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | Gly | Trp | Met | Ser | Ala | Cys | Tyr | Leu | Ser | Lys | Ala | Leu | Gly | Pro | Gly |
|     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Val | Glu | Val | Thr | Val | Leu | Glu | Ala | Pro | Ser | Ile | Ser | Arg | Ile | Arg | Val |
|     | 35  |     |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gly | Glu | Ala | Thr | Ile | Pro | Asn | Leu | His | Lys | Val | Phe | Phe | Asp | Phe | Leu |
| 50  |     |     |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gly | Ile | Ala | Glu | Asp | Glu | Trp | Met | Arg | Glu | Cys | Asn | Ala | Ser | Tyr | Lys |
| 65  |     |     |     |     | 70  |     |     |     |     | 75  |     |     |     |     | 80  |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | Ala | Val | Arg | Phe | Val | Asn | Trp | Arg | Thr | Pro | Gly | Asp | Gly | Gln | Ala |
|     |     |     | 85  |     |     |     |     | 90  |     |     |     |     |     | 95  |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Thr | Pro | Arg | Arg | Arg | Pro | Asp | Gly | Arg | Pro | Asp | His | Phe | Asp | His | Leu |
|     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Phe | Gly | Gln | Leu | Pro | Glu | His | Glu | Asn | Leu | Pro | Leu | Ser | Gln | Tyr | Trp |
|     |     | 115 |     |     |     |     | 120 |     |     |     |     | 125 |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | His | Arg | Arg | Leu | Asn | Gly | Leu | Thr | Asp | Glu | Pro | Phe | Asp | Arg | Ser |
|     | 130 |     |     |     |     | 135 |     |     |     |     | 140 |     |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Cys | Tyr | Val | Gln | Pro | Glu | Leu | Leu | Asp | Arg | Lys | Leu | Ser | Pro | Arg | Leu |
| 145 |     |     |     |     | 150 |     |     |     |     | 155 |     |     |     |     | 160 |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asp | Gly | Thr | Lys | Leu | Ala | Ser | Tyr | Ala | Trp | His | Phe | Asp | Ala | Asp |
|     |     |     |     | 165 |     |     |     |     | 170 |     |     |     |     | 175 |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Val | Ala | Asp | Phe | Leu | Cys | Arg | Phe | Ala | Val | Gln | Lys | Leu | Asn | Val |
|     |     |     | 180 |     |     |     |     | 185 |     |     |     |     |     | 190 |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Thr | His | Val | Gln | Asp | Val | Phe | Thr | His | Ala | Asp | Leu | Asp | Gln | Arg | Gly |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|

205

|            |            |            |            |            |            |            |            |            |            |            |            |            |            |            |            |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| His<br>210 | Ile        | Thr        | Ala        | Val        | Asn        | Thr<br>215 | Glu        | Ser        | Gly        | Arg        | Thr<br>220 | Leu        | Ala        | Ala        | Asp        |
| Leu<br>225 | Phe        | Ile        | Asp        | Cys        | Ser<br>230 | Gly        | Phe        | Arg        | Ser        | Val<br>235 | Leu        | Met        | Gly        | Lys        | Val<br>240 |
| Met        | Gln        | Glu        | Pro        | Phe<br>245 | Leu        | Asp        | Met        | Ser        | Lys<br>250 | His        | Leu        | Leu        | Asn        | Asp<br>255 | Arg        |
| Ala        | Val        | Ala        | Leu<br>260 | Met        | Leu        | Pro        | His        | Asp<br>265 | Asp        | Glu        | Lys        | Val        | Gly<br>270 | Ile        | Glu        |
| Pro        | Tyr        | Thr<br>275 | Ser        | Ser        | Leu        | Ala        | Met<br>280 | Arg        | Ser        | Gly        | Trp        | Ser<br>285 | Trp        | Lys        | Ile        |
| Pro        | Leu<br>290 | Leu        | Gly        | Arg        | Phe        | Gly<br>295 | Ser        | Gly        | Tyr        | Val        | Tyr<br>300 | Ser        | Ser        | Gln        | Phe        |
| Thr<br>305 | Ser        | Gln        | Asp        | Glu        | Ala<br>310 | Ala        | Glu        | Glu        | Leu        | Cys<br>315 | Arg        | Met        | Trp        | Asp        | Val<br>320 |
| Asp        | Pro        | Ala        | Glu        | Gln<br>325 | Thr        | Phe        | Asn        | Asn        | Val<br>330 | Arg        | Phe        | Arg        | Val        | Gly<br>335 | Arg        |
| Ser        | Arg        | Arg        | Ala<br>340 | Trp        | Val        | Arg        | Asn        | Cys<br>345 | Val        | Ala        | Ile        | Gly        | Val<br>350 | Ser        | Ala        |
| Met        | Phe<br>355 | Val        | Glu        | Pro        | Leu        | Glu        | Ser<br>360 | Thr        | Gly        | Leu        | Tyr        | Phe<br>365 | Ser        | Tyr        | Ala        |
| Ser<br>370 | Leu        | Tyr        | Gln        | Leu        | Val        | Lys<br>375 | His        | Phe        | Pro        | Asp        | Lys<br>380 | Arg        | Phe        | Arg        | Pro        |
| Ile<br>385 | Leu        | Ala        | Asp        | Arg        | Phe<br>390 | Asn        | Arg        | Glu        | Val        | Ala<br>395 | Thr        | Met        | Tyr        | Asp        | Asp<br>400 |
| Thr        | Arg        | Asp        | Phe<br>405 | Leu        | Gln        | Ala        | His        | Phe        | Ser<br>410 | Leu        | Ser        | Pro        | Arg        | Asp<br>415 | Asp        |
| Ser        | Glu        | Phe        | Trp<br>420 | Arg        | Ala        | Cys        | Lys        | Glu<br>425 | Leu        | Pro        | Phe        | Ala        | Asp<br>430 | Gly        | Phe        |
| Ala        | Glu        | Lys<br>435 | Val        | Glu        | Met        | Tyr        | Arg<br>440 | Ala        | Gly        | Leu        | Pro        | Val        | Glu        | Leu        | Pro        |
| Val<br>450 | Thr        | Ile        | Asp        | Asp        | Gly        | His<br>455 | Tyr        | Tyr        | Gly        | Asn        | Phe<br>460 | Glu        | Ala        | Glu        | Phe        |
| Arg<br>465 | Asn        | Phe        | Trp        | Thr        | Asn<br>470 | Ser        | Asn        | Tyr        | Tyr        | Cys<br>475 | Ile        | Phe        | Ala        | Gly        | Leu<br>480 |



Gly Phe Leu Pro Glu His Pro Leu Pro Val Leu Glu Phe Arg Pro Glu  
485 490 495

Ala Val Asp Arg Ala Glu Pro Val Phe Ala Ala Val Arg Arg Arg Thr  
500 505 510

Glu Glu Leu Val Ala Thr Ala Pro Thr Met Gln Ala Tyr Leu Arg Arg  
515 520 525

Leu His Gln Gly Thr  
530

<210> 25  
<211> 1617  
<212> DNA  
<213> Pseudomonas fluorescens

<400> 25  
atgaacaagc cgatcaagaa tatcgtcatc gtgggcggcg gtactgcggg ctggatggcc 60  
gcctcgtacc tcgtccgggc cctccaacag caggcgaaca ttacgtcatc cgaatctgcg 120  
gcgatccctc ggatcggcgt gggcgaagcg accatcccaa gtttgagaa ggtgttcttc 180  
gatttcctcg ggataccgga gcgggaatgg atgccccaa tgaacggcgc gttcaaggcc 240  
gcgatcaagt tcgtgaattg gagaaagtct cccgaccctc cgcgcgacga tcacttctac 300  
catttgttcg gcaacgtgcc gaactgcgac ggcgtgccgc ttaccacta ctggctgcgc 360  
aagcgcgaac agggcttcca gcagccgatg gactacgcgt gctaccgca gcccggggca 420  
ctcgacggca agctggcacc gtgcctgtcc gacggcacc gccagatgtc ccacgcgtgg 480  
cacttcgacg cgcacctggg gcccgacttc ttgaagcgt gggccgtcga gcgcgggggtg 540  
aaccgcgtgg tcgatgaggt ggtggacgtt cgcctgaaca accgcggcta catctccaac 600  
ctgctcacca aggaggggcg gacgtggag gcggacctgt tcatcgactg ctccggcatg 660  
cgggggctcc tgatcaatca ggcgtgaag gaacccttca tcgacatgtc cgactacctg 720  
ctgtgcgaca gcgcggctgc cagcgccgtg cccaacgacg acgcgcgca tggggtcgag 780  
ccgtacacct cctcgatcgc catgaactcg ggatggacct ggaagattcc gatgctgggc 840  
cggttcggca gcggctacgt cttctcgagc catttcacct cgcgcgacca ggccaccgcc 900  
gacttcctca aactctgggg cctctcgac aatcagccgc tcaaccagat caagttccgg 960  
gtcgggcgca acaagcgggc gtgggtcaac aactgcgtct cgatcgggct gtcgtcgtgc 1020  
tttctggagc cctggaatc gacggggatc tacttcatct acgcggcgct ttaccagctc 1080  
gtgaagcact tccccgacac ctcgttcgac ccgcggctga gcgacgttt caacgccgag 1140  
atcgtccaca tggtcgacga ctgccgggat ttcgtccaag cgcactatct caccacgtcg 1200  
cgcgatgaca cgccgttctg gtcgcgaac cggcacgacc tgcggctctc ggacgccatc 1260  
aaagagaagg ttcagcgcta caaggcgggg ctgccgtga ccaccacgtc gttcgacgat 1320  
tccacgtact acgagacctt cgactacgaa ttcaagaatt tctggttgaa cggcaactac 1380

|            |            |            |            |            |            |      |
|------------|------------|------------|------------|------------|------------|------|
| tactgcatct | ttgccggctt | gggcatgctg | cccgaccggt | cgctgccgct | gttgagcac  | 1440 |
| cgaccggagt | cgatcgagaa | agccgaggcg | atgttcgcca | gcatccggcg | cgaggccgag | 1500 |
| cgtctgcgca | ccagcctgcc | gacaaactac | gactacctgc | ggtcgctgcg | tgacggcgac | 1560 |
| gcggggctgt | cgcgcgccca | gcgtgggccg | aagctcgcag | cgcaggaaag | cctgtag    | 1617 |

|       |                                |
|-------|--------------------------------|
| <210> | 26                             |
| <211> | 538                            |
| <212> | PRT                            |
| <213> | <i>Pseudomonas fluorescens</i> |

<400> 26

Met Asn Lys Pro Ile Lys Asn Ile Val Ile Val Gly Gly Gly Thr Ala  
1 5 10 15

Gly Trp Met Ala Ala Ser Tyr Leu Val Arg Ala Leu Gln Gln Gln Ala  
20 25 30

Asn Ile Thr Leu Ile Glu Ser Ala Ala Ile Pro Arg Ile Gly Val Gly  
35 40 45

Glu Ala Thr Ile Pro Ser Leu Gln Lys Val Phe Phe Asp Phe Leu Gly  
50 55 60

Ile Pro Glu Arg Glu Trp Met Pro Gln Val Asn Gly Ala Phe Lys Ala  
65 70 75 80

Ala Ile Lys Phe Val Asn Trp Arg Lys Ser Pro Asp Pro Ser Arg Asp  
85 90 95

Asp His Phe Tyr His Leu Phe Gly Asn Val Pro Asn Cys Asp Gly Val  
100 105 110

Pro Leu Thr His Tyr Trp Leu Arg Lys Arg Glu Gln Gly Phe Gln Gln  
115 120 125

Pro Met Glu Tyr Ala Cys Tyr Pro Gln Pro Gly Ala Leu Asp Gly Lys  
130 135 140

Leu Ala Pro Cys Leu Ser Asp Gly Thr Arg Gln Met Ser His Ala Trp  
145 150 155 160

His Phe Asp Ala His Leu Val Ala Asp Phe Leu Lys Arg Trp Ala Val  
165 170 175

Glu Arg Gly Val Asn Arg Val Val Asp Glu Val Val Asp Val Arg Leu  
180 185 190

Asn Asn Arg Gly Tyr Ile Ser Asn Leu Leu Thr Lys Glu Gly Arg Thr  
195 200 205

Leu Glu Ala Asp Leu Phe Ile Asp Cys Ser Gly Met Arg Gly Leu Leu  
210 215 220

Ile Asn Gln Ala Leu Lys Glu Pro Phe Ile Asp Met Ser Asp Tyr Leu  
225 230 235 240

Leu Cys Asp Ser Ala Val Ala Ser Ala Val Pro Asn Asp Asp Ala Arg  
245 250 255

Asp Gly Val Glu Pro Tyr Thr Ser Ser Ile Ala Met Asn Ser Gly Trp  
260 265 270

Thr Trp Lys Ile Pro Met Leu Gly Arg Phe Gly Ser Gly Tyr Val Phe  
275 280 285

Ser Ser His Phe Thr Ser Arg Asp Gln Ala Thr Ala Asp Phe Leu Lys  
290 295 300

Leu Trp Gly Leu Ser Asp Asn Gln Pro Leu Asn Gln Ile Lys Phe Arg  
305 310 315 320

Val Gly Arg Asn Lys Arg Ala Trp Val Asn Asn Cys Val Ser Ile Gly  
325 330 335

Leu Ser Ser Cys Phe Leu Glu Pro Leu Glu Ser Thr Gly Ile Tyr Phe  
340 345 350

Ile Tyr Ala Ala Leu Tyr Gln Leu Val Lys His Phe Pro Asp Thr Ser  
355 360 365

Phe Asp Pro Arg Leu Ser Asp Ala Phe Asn Ala Glu Ile Val His Met  
370 375 380

Phe Asp Asp Cys Arg Asp Phe Val Gln Ala His Tyr Phe Thr Thr Ser  
385 390 395 400

Arg Asp Asp Thr Pro Phe Trp Leu Ala Asn Arg His Asp Leu Arg Leu  
405 410 415

Ser Asp Ala Ile Lys Glu Lys Val Gln Arg Tyr Lys Ala Gly Leu Pro  
420 425 430

Leu Thr Thr Thr Ser Phe Asp Asp Ser Thr Tyr Tyr Glu Thr Phe Asp  
435 440 445

Tyr Glu Phe Lys Asn Phe Trp Leu Asn Gly Asn Tyr Tyr Cys Ile Phe  
450 455 460

Ala Gly Leu Gly Met Leu Pro Asp Arg Ser Leu Pro Leu Leu Gln His  
465 470 475 480

Arg Pro Glu Ser Ile Glu Lys Ala Glu Ala Met Phe Ala Ser Ile Arg

Arg Glu Ala Glu Arg Leu Arg Thr Ser Leu Pro Thr Asn Tyr Asp Tyr  
500 505 510

Leu Arg Ser Leu Arg Asp Gly Asp Ala Gly Leu Ser Arg Gly Gln Arg  
515 520 525

Gly Pro Lys Leu Ala Ala Gln Glu Ser Leu  
530 535

<210> 27

<211> 1593

<212> DNA

<213> *Lentzea areocolonigenes*

<400> 27

```

atgtccggca agattgacaa gatcctcatc gtcggcgggc gcaccgccgg atggatggcc      60
gcgtcctatc tcggcaaggc cctgcagggc accgcggaca tcacactgct gcaggcaccc      120
gacatcccga cgctcggggg cggcgaggcc acgatcccca atctgcagac ggcgttcttc      180
gacttcctcg gaatccccga ggacgagtgg atgcgggagt gcaacgcgag ctacaaggtc      240
gccatcaagt tcatcaactg gcgcaccgcg ggcgagggga cgtccgaggc ccgcgagctc      300
gacggagggc ccgaccactt ctaccactcc ttcggtctgc tcaagtacca cgagcagatt      360
ccgctgtcgc actactgggt cgaccgttcg taccggggga agaccgtcga gccgttcgac      420
tacgcctgct acaaggaacc cgtcatcctc gacgccaaca ggtcaccgcg caggctcgac      480
ggtccaagg tgacgaacta cgcggtggc cgttcgacgc acctcgtcgc cgacttcctg      540
cgccggttcg ccaccgagaa gtcggcggtg cgccacgtcg aggaccgct cgagcacgtc      600
cagcgcgacg ccaacggcaa catcgagtcg gttcgcacgg caacggggcg tgtcttcgat      660
gccgacctct tcgtcgactg ctcgggcttc cgcgggctgc tgatcaacaa ggcgatggag      720
gagcccttcc tcgacatgag cgatcacctg ctcaacgaca gcgccgtcgc caccagggtg      780
ccgcacgacg acgacgcgaa cgggtgtgga ccgttcacct cggcgatcgc catgaagtcg      840
ggctggacgt ggaagatccc gatgctcggc aggttcggca ccgggtacgt ctactcgagc      900
cggttcgcca ccgaggacga ggcggtgcgc gagttctgcg agatgtggca cctcgaccgc      960
gagaccacgc ccctcaacag gatccggttc cgggtcggcc gcaaccggcg cgcgtgggtc     1020
ggcaactgcg tcagcatcgg cacgtcgtcg tgcttcgtgg aaccactgga gtcgacgggc     1080
atctacttcg tctacgccgc gctgtaccag ctggtgaagc acttccccga caagagcctc     1140
aaccctgtgc tgaccgccag gttcaaccgc gagatcgaga cgatgttcga cgacacgcgc     1200
gacttcatcc aggcgactt ctacttctcg ccgcgcacgg acaccccggt ctggagggcc     1260
aacaaggagc tgcgcctggc ggacggcatg caggagaaga tcgacatgta ccgcgcgggc     1320
atggcgatca acgcgcccgc gtccgacgac gccagctct actacggcaa cttcgaggag     1380
gagttccgca acttctggaa caacagcaac tactactgcg tgctggccgg cctcggtctg     1440

```

gtgcccgacg caccctcacc acgcctggcg cacatgccac aggcgacgga gtcggtggac 1500  
gaggtcttcg gcgccgtcaa ggaccggcag cggaacctgc tcgagaccct gccgagcctc 1560  
cacgagttcc tgaggcaaca gcacggccgc tga 1593

<210> 28  
<211> 530  
<212> PRT  
<213> Lentzea areocolonigenes

<400> 28

Met Ser Gly Lys Ile Asp Lys Ile Leu Ile Val Gly Gly Gly Thr Ala  
1 5 10 15

Gly Trp Met Ala Ala Ser Tyr Leu Gly Lys Ala Leu Gln Gly Thr Ala  
20 25 30

Asp Ile Thr Leu Leu Gln Ala Pro Asp Ile Pro Thr Leu Gly Val Gly  
35 40 45

Glu Ala Thr Ile Pro Asn Leu Gln Thr Ala Phe Phe Asp Phe Leu Gly  
50 55 60

Ile Pro Glu Asp Glu Trp Met Arg Glu Cys Asn Ala Ser Tyr Lys Val  
65 70 75 80

Ala Ile Lys Phe Ile Asn Trp Arg Thr Ala Gly Glu Gly Thr Ser Glu  
85 90 95

Ala Arg Glu Leu Asp Gly Gly Pro Asp His Phe Tyr His Ser Phe Gly  
100 105 110

Leu Leu Lys Tyr His Glu Gln Ile Pro Leu Ser His Tyr Trp Phe Asp  
115 120 125

Arg Ser Tyr Arg Gly Lys Thr Val Glu Pro Phe Asp Tyr Ala Cys Tyr  
130 135 140

Lys Glu Pro Val Ile Leu Asp Ala Asn Arg Ser Pro Arg Arg Leu Asp  
145 150 155 160

Gly Ser Lys Val Thr Asn Tyr Ala Trp His Phe Asp Ala His Leu Val  
165 170 175

Ala Asp Phe Leu Arg Arg Phe Ala Thr Glu Lys Leu Gly Val Arg His  
180 185 190

Val Glu Asp Arg Val Glu His Val Gln Arg Asp Ala Asn Gly Asn Ile  
195 200 205

Glu Ser Val Arg Thr Ala Thr Gly Arg Val Phe Asp Ala Asp Leu Phe  
210 215 220

Val Asp Cys Ser Gly Phe Arg Gly Leu Leu Ile Asn Lys Ala Met Glu  
225 230 235 240

Glu Pro Phe Leu Asp Met Ser Asp His Leu Leu Asn Asp Ser Ala Val  
245 250 255

Ala Thr Gln Val Pro His Asp Asp Asp Ala Asn Gly Val Glu Pro Phe  
260 265 270

Thr Ser Ala Ile Ala Met Lys Ser Gly Trp Thr Trp Lys Ile Pro Met  
275 280 285

Leu Gly Arg Phe Gly Thr Gly Tyr Val Tyr Ser Ser Arg Phe Ala Thr  
290 295 300

Glu Asp Glu Ala Val Arg Glu Phe Cys Glu Met Trp His Leu Asp Pro  
305 310 315 320

Glu Thr Gln Pro Leu Asn Arg Ile Arg Phe Arg Val Gly Arg Asn Arg  
325 330 335

Arg Ala Trp Val Gly Asn Cys Val Ser Ile Gly Thr Ser Ser Cys Phe  
340 345 350

Val Glu Pro Leu Glu Ser Thr Gly Ile Tyr Phe Val Tyr Ala Ala Leu  
355 360 365

Tyr Gln Leu Val Lys His Phe Pro Asp Lys Ser Leu Asn Pro Val Leu  
370 375 380

Thr Ala Arg Phe Asn Arg Glu Ile Glu Thr Met Phe Asp Asp Thr Arg  
385 390 395 400

Asp Phe Ile Gln Ala His Phe Tyr Phe Ser Pro Arg Thr Asp Thr Pro  
405 410 415

Phe Trp Arg Ala Asn Lys Glu Leu Arg Leu Ala Asp Gly Met Gln Glu  
420 425 430

Lys Ile Asp Met Tyr Arg Ala Gly Met Ala Ile Asn Ala Pro Ala Ser  
435 440 445

Asp Asp Ala Gln Leu Tyr Tyr Gly Asn Phe Glu Glu Glu Phe Arg Asn  
450 455 460

Phe Trp Asn Asn Ser Asn Tyr Tyr Cys Val Leu Ala Gly Leu Gly Leu  
465 470 475 480

Val Pro Asp Ala Pro Ser Pro Arg Leu Ala His Met Pro Gln Ala Thr  
485 490 495

Glu Ser Val Asp Glu Val Phe Gly Ala Val Lys Asp Arg Gln Arg Asn  
500 505 510

Leu Leu Glu Thr Leu Pro Ser Leu His Glu Phe Leu Arg Gln Gln His  
515 520 525

Gly Arg  
530

<210> 29  
<211> 516  
<212> PRT  
<213> Streptomyces albus

<400> 29

Met Leu Asn Asn Val Val Ile Val Gly Gly Gly Thr Ala Gly Trp Met  
1 5 10 15

Thr Ala Ser Tyr Phe Lys Ala Ala Phe Gly Glu Arg Ile Asn Ile Thr  
20 25 30

Leu Val Glu Ser Gly Asn Ile Gly Ala Val Gly Val Gly Glu Ala Thr  
35 40 45

Phe Ser Asp Ile Arg His Phe Phe Asp Phe Leu Gly Leu Lys Glu Lys  
50 55 60

Asp Trp Met Pro Ala Cys Asn Ala Thr Tyr Lys Leu Ala Val Arg Phe  
65 70 75 80

Glu Asn Trp Arg Glu Glu Gly His Tyr Phe Tyr His Pro Phe Glu Gln  
85 90 95

Met Arg Ser Val Asn Gly Phe Pro Leu Ala Asp Trp Trp Leu Arg Glu  
100 105 110

Lys Pro Thr Asp Arg Phe Asp Lys Asp Cys Phe Val Met Pro Ser Ile  
115 120 125

Ile Asp Ala Gly Leu Ser Pro Arg His Leu Asp Gly Ser Leu Ile Asp  
130 135 140

Gln Pro Phe Val Glu Gly Ala Glu Glu Met Gln Gly Leu Thr Met Ser  
145 150 155 160

Glu His Gln Gly Lys Thr Gln Phe Pro Tyr Ala Tyr Gln Phe Glu Ala  
165 170 175

Ala Leu Leu Ala Lys Tyr Leu Thr Thr Tyr Ser Val Glu Arg Gly Val  
180 185 190

Lys His Ile Val Asp Asp Val Gln Gln Val Asn Leu Asp Glu Arg Gly  
195 200 205

Trp Ile Arg Asn Val Thr Thr Ala Glu His Gly Glu Ile Gly Gly Asp  
 210 215 220

Leu Phe Ile Asp Cys Thr Gly Phe Arg Gly Leu Leu Leu Asn Lys Ala  
 225 230 235 240

Leu Glu Glu Pro Phe Ile Ser Phe Gln Asp Thr Leu Pro Asn Asp Ser  
 245 250 255

Ala Val Ala Leu Gln Val Pro Met Asp Met Glu Arg Arg Gly Ile Leu  
 260 265 270

Pro Cys Thr Thr Ala Thr Ala Gln Asp Ala Gly Trp Ile Trp Thr Ile  
 275 280 285

Pro Leu Met Gly Arg Val Gly Thr Gly Tyr Val Tyr Ala Lys Asp Tyr  
 290 295 300

Ile Ser Pro Glu Glu Ala Glu Arg Thr Leu Arg Glu Phe Val Gly Pro  
 305 310 315 320

Ala Ala Ala Asp Ala Glu Ala Asn His Ile Lys Met Arg Val Gly Arg  
 325 330 335

Ser Glu Asn Ser Trp Val Lys Asn Cys Val Ala Ile Gly Leu Ser Ser  
 340 345 350

Gly Phe Val Glu Pro Leu Glu Ser Thr Gly Ile Phe Phe Ile His His  
 355 360 365

Ala Ile Glu Gln Leu Ala Lys Asn Phe Pro Gly Glu Asp Trp Asn Pro  
 370 375 380

Val Gln Arg Asn Leu Tyr Asn Asp Ser Val Ala His Val Met Asp Gly  
 385 390 395 400

Val Arg Glu Phe Leu Val Leu His Tyr Val Ala Ala Lys Arg Ser Asp  
 405 410 415

Thr Gln Tyr Trp Arg Asp Thr Lys Thr Arg Lys Ile Pro Asp Ala Leu  
 420 425 430

Ala Glu Arg Ile Glu Lys Trp Lys Thr Gln Leu Pro Asp Thr Glu Thr  
 435 440 445

Val Phe Pro Tyr Tyr His Gly Leu Pro Pro Tyr Ser Tyr His Cys Val  
 450 455 460

Leu Leu Gly Met Gly Gly Ile Asp Val Lys Pro Ser Pro Ala Leu Ala  
 465 470 475 480



Leu Gly Asp Ser Ala Ala Ala Glu Arg Glu Phe Ala Glu Ile Arg Glu  
485 490 495

Lys Thr Arg His Leu Thr Ser Val Leu Pro Lys Ala Tyr Asp Tyr Phe  
500 505 510

Ser Arg Met Arg  
515

<210> 30  
<211> 530  
<212> PRT  
<213> Lentzea aerocologinenes

<400> 30

Met Ser Gly Lys Ile Asp Lys Ile Leu Ile Val Gly Gly Gly Thr Ala  
1 5 10 15

Gly Trp Met Ala Ala Ser Tyr Leu Gly Lys Ala Leu Gln Gly Thr Ala  
20 25 30

Asp Ile Thr Leu Leu Gln Ala Pro Asp Ile Pro Thr Leu Gly Val Gly  
35 40 45

Glu Ala Thr Ile Pro Asn Leu Gln Thr Ala Phe Phe Asp Phe Leu Gly  
50 55 60

Ile Pro Glu Asp Glu Trp Met Arg Glu Cys Asn Ala Ser Tyr Lys Val  
65 70 75 80

Ala Ile Lys Phe Ile Asn Trp Arg Thr Ala Gly Glu Gly Thr Ser Glu  
85 90 95

Ala Arg Glu Leu Asp Gly Gly Pro Asp His Phe Tyr His Ser Phe Gly  
100 105 110

Leu Leu Lys Tyr His Glu Gln Ile Pro Leu Ser His Tyr Trp Phe Asp  
115 120 125

Arg Ser Tyr Arg Gly Lys Thr Val Glu Pro Phe Asp Tyr Ala Cys Tyr  
130 135 140

Lys Glu Pro Val Ile Leu Asp Ala Asn Arg Ser Pro Arg Arg Leu Asp  
145 150 155 160

Gly Ser Lys Val Thr Asn Tyr Ala Trp His Phe Asp Ala His Leu Val  
165 170 175

Ala Asp Phe Leu Arg Arg Phe Ala Thr Glu Lys Leu Gly Val Arg His  
180 185 190

Val Glu Asp Arg Val Glu His Val Gln Arg Asp Ala Asn Gly Asn Ile

205

Phe Trp Asn Asn Ser Asn Tyr Tyr Cys Val Leu Ala Gly Leu Gly Leu  
465 470 475 480

Val Pro Asp Ala Pro Ser Pro Arg Leu Ala His Met Pro Gln Ala Thr  
485 490 495

Glu Ser Val Asp Glu Val Phe Gly Ala Val Lys Asp Arg Gln Arg Asn  
500 505 510

Leu Leu Glu Thr Leu Pro Ser Leu His Glu Phe Leu Arg Gln Gln His  
515 520 525

Gly Arg  
530

<210> 31  
<211> 530  
<212> PRT  
<213> Lentzea aerocologinenes

<400> 31

Met Ser Gly Lys Ile Asp Lys Ile Leu Ile Val Gly Gly Gly Thr Ala  
1 5 10 15

Gly Trp Met Ala Ala Ser Tyr Leu Gly Lys Ala Leu Gln Gly Thr Ala  
20 25 30

Asp Ile Thr Leu Leu Gln Ala Pro Asp Ile Pro Thr Leu Gly Val Gly  
35 40 45

Glu Ala Thr Ile Pro Asn Leu Gln Thr Ala Phe Phe Asp Phe Leu Gly  
50 55 60

Ile Pro Glu Asp Glu Trp Met Arg Glu Cys Asn Ala Ser Tyr Lys Val  
65 70 75 80

Ala Ile Lys Phe Ile Asn Trp Arg Thr Ala Gly Glu Gly Thr Ser Glu  
85 90 95

Ala Arg Glu Leu Asp Gly Gly Pro Asp His Phe Tyr His Ser Phe Gly  
100 105 110

Leu Leu Lys Tyr His Glu Gln Ile Pro Leu Ser His Tyr Trp Phe Asp  
115 120 125

Arg Ser Tyr Arg Gly Lys Thr Val Glu Pro Phe Asp Tyr Ala Cys Tyr  
130 135 140

Lys Glu Pro Val Ile Leu Asp Ala Asn Arg Ser Pro Arg Arg Leu Asp  
145 150 155 160

Gly Ser Lys Val Thr Asn Tyr Ala Trp His Phe Asp Ala His Leu Val  
165 170 175

Ala Asp Phe Leu Arg Arg Phe Ala Thr Glu Lys Leu Gly Val Arg His  
180 185 190

Val Glu Asp Arg Val Glu His Val Gln Arg Asp Ala Asn Gly Asn Ile  
195 200 205

Glu Ser Val Arg Thr Ala Thr Gly Arg Val Phe Asp Ala Asp Leu Phe  
210 215 220

Val Asp Cys Ser Gly Phe Arg Gly Leu Leu Ile Asn Lys Ala Met Glu  
225 230 235 240

Glu Pro Phe Leu Asp Met Ser Asp His Leu Leu Asn Asp Ser Ala Val  
245 250 255

Ala Thr Gln Val Pro His Asp Asp Asp Ala Asn Gly Val Glu Pro Phe  
260 265 270

Thr Ser Ala Ile Ala Met Lys Ser Gly Trp Thr Trp Lys Ile Pro Met  
275 280 285

Leu Gly Arg Phe Gly Thr Gly Tyr Val Tyr Ser Ser Arg Phe Ala Thr  
290 295 300

Glu Asp Glu Ala Val Arg Glu Phe Cys Glu Met Trp His Leu Asp Pro  
305 310 315 320

Glu Thr Gln Pro Leu Asn Arg Ile Arg Phe Arg Val Gly Arg Asn Arg  
325 330 335

Arg Ala Trp Val Gly Asn Cys Val Ser Ile Gly Thr Ser Ser Cys Phe  
340 345 350

Val Glu Pro Leu Glu Ser Thr Gly Ile Tyr Phe Val Tyr Ala Ala Leu  
355 360 365

Tyr Gln Leu Val Lys His Phe Pro Asp Lys Ser Leu Asn Pro Val Leu  
370 375 380

Thr Ala Arg Phe Asn Arg Glu Ile Glu Thr Met Phe Asp Asp Thr Arg  
385 390 395 400

Asp Phe Ile Gln Ala His Phe Tyr Phe Ser Pro Arg Thr Asp Thr Pro  
405 410 415

Phe Trp Arg Ala Asn Lys Glu Leu Arg Leu Ala Asp Gly Met Gln Glu  
420 425 430

Lys Ile Asp Met Tyr Arg Ala Gly Met Ala Ile Asn Ala Pro Ala Ser  
435 440 445

Asp Asp Ala Gln Leu Tyr Tyr Gly Asn Phe Glu Glu Glu Phe Arg Asn

|                                                                 |                         |     |  |     |
|-----------------------------------------------------------------|-------------------------|-----|--|-----|
| 450                                                             |                         | 455 |  | 460 |
| Phe Trp Asn Asn Ser Ser Tyr Tyr Cys Val Leu Ala Gly Leu Gly Leu |                         |     |  |     |
| 465                                                             |                         | 470 |  | 475 |
|                                                                 |                         |     |  | 480 |
| Val Pro Asp Ala Pro Ser Pro Arg Leu Ala His Met Pro Gln Ala Thr |                         |     |  |     |
|                                                                 |                         | 485 |  | 490 |
|                                                                 |                         |     |  | 495 |
| Glu Ser Val Asp Glu Val Phe Gly Ala Val Lys Asp Arg Gln Arg Asn |                         |     |  |     |
|                                                                 |                         | 500 |  | 505 |
|                                                                 |                         |     |  | 510 |
| Leu Leu Glu Thr Leu Pro Ser Leu His Glu Phe Leu Arg Gln Gln His |                         |     |  |     |
|                                                                 |                         | 515 |  | 520 |
|                                                                 |                         |     |  | 525 |
| Gly Arg                                                         |                         |     |  |     |
| 530                                                             |                         |     |  |     |
| <210>                                                           | 32                      |     |  |     |
| <211>                                                           | 530                     |     |  |     |
| <212>                                                           | PRT                     |     |  |     |
| <213>                                                           | Lentzea aerocologinenes |     |  |     |
| <400>                                                           | 32                      |     |  |     |
| Met Ser Gly Lys Ile Asp Lys Ile Leu Ile Val Gly Gly Gly Thr Ala |                         |     |  |     |
| 1                                                               |                         | 5   |  | 10  |
|                                                                 |                         |     |  | 15  |
| Gly Trp Met Ala Ala Ser Tyr Leu Gly Lys Ala Leu Gln Gly Thr Ala |                         |     |  |     |
|                                                                 |                         | 20  |  | 25  |
|                                                                 |                         |     |  | 30  |
| Asp Ile Thr Leu Leu Gln Ala Pro Asp Ile Pro Thr Leu Gly Val Gly |                         |     |  |     |
|                                                                 |                         | 35  |  | 40  |
|                                                                 |                         |     |  | 45  |
| Glu Ala Thr Met Pro Asn Leu Gln Thr Ala Phe Phe Asp Phe Leu Gly |                         |     |  |     |
|                                                                 |                         | 50  |  | 55  |
|                                                                 |                         |     |  | 60  |
| Ile Pro Glu Asp Glu Trp Met Arg Glu Cys Asn Ala Ser Tyr Lys Val |                         |     |  |     |
| 65                                                              |                         | 70  |  | 75  |
|                                                                 |                         |     |  | 80  |
| Ala Ile Lys Phe Ile Asn Trp Arg Thr Ala Gly Glu Gly Thr Ser Glu |                         |     |  |     |
|                                                                 |                         | 85  |  | 90  |
|                                                                 |                         |     |  | 95  |
| Ala Arg Glu Leu Asp Gly Gly Pro Asp His Phe Tyr His Pro Phe Gly |                         |     |  |     |
|                                                                 |                         | 100 |  | 105 |
|                                                                 |                         |     |  | 110 |
| Leu Leu Lys Tyr His Glu Gln Ile Pro Leu Ser His Tyr Trp Phe Asp |                         |     |  |     |
|                                                                 |                         | 115 |  | 120 |
|                                                                 |                         |     |  | 125 |
| Arg Leu Tyr Arg Gly Lys Thr Val Glu Pro Phe Asp Tyr Ala Cys Tyr |                         |     |  |     |
|                                                                 |                         | 130 |  | 135 |
|                                                                 |                         |     |  | 140 |
| Lys Glu Pro Val Ile Leu Asp Ala Asn Arg Ser Pro Arg Arg Leu Asp |                         |     |  |     |
| 145                                                             |                         | 150 |  | 155 |
|                                                                 |                         |     |  | 160 |

Gly Ser Lys Val Thr Ser Tyr Ala Trp His Phe Asp Ala His Leu Val  
 165 170 175

Ala Asp Phe Leu Arg Arg Phe Ala Thr Glu Lys Leu Gly Val Arg His  
 180 185 190

Val Glu Asp Arg Val Glu His Val Gln Arg Asp Ala Asn Gly Asn Ile  
 195 200 205

Glu Ser Val Arg Thr Ala Thr Gly Arg Val Phe Asp Ala Asp Leu Phe  
 210 215 220

Val Asp Cys Ser Gly Phe Arg Gly Leu Leu Ile Asn Lys Ala Met Glu  
 225 230 235 240

Glu Pro Phe Leu Asp Met Ser Asp His Leu Leu Asn Asp Ser Ala Val  
 245 250 255

Ala Thr Gln Val Pro His Asp Asp Asp Ala Asn Gly Val Glu Pro Phe  
 260 265 270

Thr Ser Ala Ile Ala Met Lys Ser Gly Trp Thr Trp Lys Ile Pro Met  
 275 280 285

Leu Gly Arg Phe Gly Thr Gly Tyr Val Tyr Ser Ser Arg Phe Ala Thr  
 290 295 300

Glu Asp Glu Ala Val Arg Glu Phe Cys Glu Met Trp His Leu Asp Pro  
 305 310 315 320

Glu Thr Gln Pro Leu Asn Arg Ile Arg Phe Arg Val Gly Arg Asn Arg  
 325 330 335

Arg Ala Trp Val Gly Asn Cys Val Ser Ile Gly Thr Ser Ser Cys Phe  
 340 345 350

Val Glu Pro Leu Glu Ser Thr Gly Ile Tyr Phe Val Tyr Ala Ala Leu  
 355 360 365

Tyr Gln Leu Val Lys His Phe Pro Asp Lys Ser Phe Asn Pro Val Leu  
 370 375 380

Thr Ala Arg Phe Asn Arg Glu Ile Glu Thr Met Phe Asp Asp Thr Arg  
 385 390 395 400

Asp Phe Ile Gln Ala His Phe Tyr Phe Ser Pro Arg Thr Asp Thr Pro  
 405 410 415

Phe Trp Arg Ala Asn Lys Glu Leu Arg Leu Ala Asp Gly Met Gln Glu  
 420 425 430

Lys Ile Asp Met Tyr Arg Ala Gly Met Ala Ile Asn Ala Pro Ala Pro  
435 440 445

Asp Asp Ala Gln Leu Tyr Trp Gly Asn Phe Glu Glu Glu Phe Arg Asn  
450 455 460

Leu Trp Asn Asn Ser Ser Tyr Tyr Cys Val Leu Ala Gly Leu Gly Leu  
465 470 475 480

Val Pro Asp Ala Pro Ser Pro Arg Leu Ala His Met Pro Arg Ala Thr  
485 490 495

Glu Ser Val Asp Glu Val Phe Gly Ala Val Lys Asp Gln Gln Arg Asn  
500 505 510

Leu Leu Glu Thr Leu Pro Ser Leu His Glu Phe Leu Arg Gln Gln His  
515 520 525

Gly Arg  
530

<210> 33  
<211> 530  
<212> PRT  
<213> Lentzea aerocologinenes

<400> 33

Met Ser Gly Lys Ile Asp Lys Ile Leu Ile Val Gly Gly Gly Thr Ala  
1 5 10 15

Gly Trp Met Ala Ala Ser Tyr Leu Gly Lys Ala Leu Gln Gly Thr Ala  
20 25 30

Asp Ile Thr Leu Leu Gln Ala Pro Asp Ile Pro Thr Leu Gly Val Gly  
35 40 45

Glu Ala Thr His Pro Asn Leu Gln Thr Ala Phe Phe Asp Phe Leu Gly  
50 55 60

Ile Pro Glu Asp Glu Trp Met Arg Glu Cys Asn Ala Ser Tyr Lys Val  
65 70 75 80

Ala Ile Lys Phe Ile Asn Trp Arg Thr Ala Gly Glu Gly Thr Ser Glu  
85 90 95

Ala Arg Glu Leu Asp Gly Gly Pro Asp His Phe Tyr His Ser Phe Gly  
100 105 110

Leu Leu Lys Tyr His Glu Gln Ile Pro Leu Ser His Tyr Trp Phe Asp  
115 120 125

Arg Ser Tyr Arg Gly Lys Thr Val Glu Pro Phe Asp Tyr Ala Cys Tyr  
130 135 140

Lys Glu Pro Val Ile Leu Asp Ala Asn Arg Ser Pro Arg Arg Leu Asp  
 145 150 155 160

Gly Ser Lys Val Thr Asn Tyr Ala Trp His Phe Asp Ala His Leu Val  
 165 170 175

Ala Asp Phe Leu Arg Arg Phe Ala Thr Glu Lys Leu Gly Val Arg His  
 180 185 190

Val Glu Asp Arg Val Glu His Val Gln Arg Asp Ala Asn Gly Asn Ile  
 195 200 205

Glu Ser Val Arg Thr Ala Thr Gly Arg Val Phe Asp Ala Asp Leu Phe  
 210 215 220

Val Asp Cys Ser Gly Phe Arg Gly Leu Leu Ile Asn Lys Ala Met Glu  
 225 230 235 240

Glu Pro Phe Leu Asp Met Ser Asp His Leu Leu Asn Asp Ser Ala Val  
 245 250 255

Ala Thr Gln Val Pro His Asp Asp Asp Ala Asn Gly Val Glu Pro Phe  
 260 265 270

Thr Ser Ala Ile Ala Met Lys Ser Gly Trp Thr Trp Lys Ile Pro Met  
 275 280 285

Leu Gly Arg Phe Gly Thr Gly Tyr Val Tyr Ser Ser Arg Phe Ala Thr  
 290 295 300

Glu Asp Glu Ala Val Arg Glu Phe Cys Glu Met Trp His Leu Asp Pro  
 305 310 315 320

Glu Thr Gln Pro Leu Asn Arg Ile Arg Phe Arg Val Gly Arg Asn Arg  
 325 330 335

Arg Ala Trp Val Gly Asn Cys Val Ser Ile Gly Thr Ser Ser Cys Phe  
 340 345 350

Val Glu Pro Leu Glu Ser Thr Gly Ile Tyr Phe Val Tyr Ala Ala Leu  
 355 360 365

Tyr Gln Leu Val Lys His Phe Pro Asp Lys Ser Phe Asn Pro Val Leu  
 370 375 380

Thr Ala Arg Phe Asn Arg Glu Ile Glu Thr Met Phe Asp Asp Thr Arg  
 385 390 395 400

Asp Phe Ile Gln Ala His Phe Tyr Phe Ser Pro Arg Thr Asp Thr Pro  
 405 410 415



Phe Trp Arg Ala Asn Lys Glu Leu Arg Leu Ala Asp Gly Met Gln Glu  
420 425 430

Lys Ile Asp Met Tyr Arg Ala Gly Met Ala Ile Asn Ala Pro Ala Ser  
435 440 445

Asp Asp Ala Gln Leu Tyr Tyr Gly Asn Phe Glu Glu Glu Phe Arg Asn  
450 455 460

Cys Trp Asn Asn Ser Ser Tyr Tyr Cys Val Leu Ala Gly Leu Gly Leu  
465 470 475 480

Val Pro Asp Ala Pro Ser Pro Arg Leu Ala His Met Pro Arg Ala Thr  
485 490 495

Glu Ser Val Asp Glu Val Phe Gly Ala Val Lys Asp Gln Gln Arg Asn  
500 505 510

Leu Leu Glu Thr Leu Pro Ser Leu His Glu Phe Leu Arg Gln Gln His  
515 520 525

Gly Arg  
530

<210> 34  
<211> 530  
<212> PRT  
<213> Lentzea aerocologinenes  
  
<400> 34

Met Ser Gly Lys Ile Asp Lys Ile Leu Ile Val Gly Gly Gly Thr Ala  
1 5 10 15

Gly Trp Met Ala Ala Ser Tyr Leu Gly Lys Ala Leu Gln Gly Thr Ala  
20 25 30

Asp Ile Thr Leu Leu Gln Ala Pro Asp Ile Pro Thr Leu Gly Val Gly  
35 40 45

Glu Ala Thr Ile Pro Asn Leu Gln Thr Ala Phe Phe Asp Phe Leu Gly  
50 55 60

Ile Pro Glu Asp Glu Trp Met Arg Glu Cys Asn Ala Ser Tyr Lys Val  
65 70 75 80

Ala Ile Lys Phe Ile Asn Trp Arg Thr Ala Gly Glu Gly Thr Ser Glu  
85 90 95

Ala Arg Glu Leu Asp Gly Gly Pro Asp His Phe Tyr His Ser Phe Ser  
100 105 110

Leu Leu Lys Tyr His Glu Gln Ile Pro Leu Ser His Tyr Trp Phe Asp

115

120

125

Arg Ser Tyr Arg Gly Lys Thr Val Glu Pro Phe Asp Tyr Ala Cys Tyr  
 130 135 140

Lys Glu Pro Val Ile Leu Asp Ala Asn Arg Ser Pro Arg Arg Leu Asp  
 145 150 155 160

Gly Ser Lys Val Thr Asn Tyr Ala Trp His Phe Asp Ala His Leu Val  
 165 170 175

Ala Asp Phe Leu Arg Arg Phe Ala Thr Glu Lys Leu Gly Val Arg His  
 180 185 190

Val Glu Asp Arg Val Glu His Val Gln Arg Asp Ala Asn Gly Asn Ile  
 195 200 205

Glu Ser Val Arg Thr Ala Thr Gly Arg Val Phe Asp Ala Asp Leu Phe  
 210 215 220

Val Asp Cys Ser Gly Phe Arg Gly Leu Leu Ile Asn Lys Ala Met Glu  
 225 230 235 240

Glu Pro Phe Leu Asp Met Ser Asp His Leu Leu Asn Asp Ser Ala Val  
 245 250 255

Ala Thr Gln Val Pro His Asp Asp Asp Ala Asn Gly Val Glu Pro Phe  
 260 265 270

Thr Ser Ala Ile Ala Met Lys Ser Gly Trp Thr Trp Lys Ile Pro Met  
 275 280 285

Leu Gly Arg Phe Gly Thr Gly Tyr Val Tyr Ser Ser Arg Phe Ala Thr  
 290 295 300

Glu Asp Glu Ala Val Arg Glu Phe Cys Glu Met Trp His Leu Asp Pro  
 305 310 315 320

Glu Thr Gln Pro Leu Asn Arg Ile Arg Phe Arg Val Gly Arg Asn Arg  
 325 330 335

Arg Ala Trp Val Gly Asn Cys Val Ser Ile Gly Thr Ser Ser Cys Phe  
 340 345 350

Val Glu Pro Leu Glu Ser Thr Gly Ile Tyr Phe Val Tyr Ala Ala Leu  
 355 360 365

Tyr Gln Leu Val Lys His Phe Pro Asp Lys Ser Leu Asn Pro Val Leu  
 370 375 380

Thr Ala Arg Phe Asn Arg Glu Ile Glu Thr Met Phe Asp Asp Thr Arg  
 385 390 395 400

Asp Phe Ile Gln Ala His Phe Tyr Phe Ser Pro Arg Thr Asp Thr Pro  
405 410 415

Phe Trp Arg Ala Asn Lys Glu Leu Arg Leu Ala Asp Gly Met Gln Glu  
420 425 430

Lys Ile Asp Met Tyr Arg Ala Gly Met Ala Ile Asn Ala Pro Ala Ser  
435 440 445

Asp Asp Ala Gln Leu Tyr Tyr Gly Asn Phe Glu Glu Glu Phe Arg Asn  
450 455 460

Phe Trp Asn Asn Ser Ser Tyr Tyr Cys Val Leu Ala Gly Leu Gly Leu  
465 470 475 480

Val Pro Asp Ala Pro Ser Pro Arg Leu Ala His Met Pro Gln Ala Thr  
485 490 495

Glu Ser Val Asp Glu Val Phe Gly Ala Val Lys Asp Arg Gln Arg Asn  
500 505 510

Leu Leu Glu Thr Leu Pro Ser Leu His Glu Phe Leu Arg Gln Gln His  
515 520 525

Gly Arg  
530

<210> 35  
<211> 530  
<212> PRT  
<213> Lentzea aerocologinenes

<400> 35

Met Ser Gly Lys Ile Asp Lys Ile Leu Ile Val Gly Gly Gly Thr Ala  
1 5 10 15

Gly Trp Met Ala Ala Ser Tyr Leu Gly Lys Ala Leu Gln Gly Thr Ala  
20 25 30

Asp Ile Thr Leu Leu Gln Ala Pro Asp Ile Pro Thr Leu Gly Val Gly  
35 40 45

Glu Ala Thr Ile Pro Asn Leu Gln Thr Ala Phe Phe Asp Phe Leu Gly  
50 55 60

Ile Pro Glu Asp Glu Trp Met Arg Glu Cys Asn Ala Ser Tyr Lys Val  
65 70 75 80

Ala Ile Lys Phe Ile Asn Trp Arg Thr Ala Gly Glu Gly Thr Ser Glu  
85 90 95

Ala Arg Glu Leu Asp Gly Gly Pro Asp His Phe Tyr His Ser Phe Gly  
100 105 110

Leu Leu Lys Tyr His Glu Gln Ile Pro Leu Ser His Tyr Trp Phe Asp  
115 120 125

Arg Ser Tyr Arg Gly Lys Thr Val Glu Pro Phe Asp Tyr Ala Cys Tyr  
130 135 140

Lys Glu Pro Val Ile Leu Asp Ala Asn Arg Ser Pro Arg Arg Leu Asp  
145 150 155 160

Gly Ser Lys Val Thr Asn Tyr Ala Trp His Phe Asp Ala His Leu Val  
165 170 175

Ala Asp Phe Leu Arg Arg Phe Ala Thr Glu Lys Leu Gly Val Arg His  
180 185 190

Val Glu Asp Arg Val Glu His Val Gln Arg Asp Ala Asn Gly Asn Ile  
195 200 205

Glu Ser Val Arg Thr Ala Thr Gly Arg Val Phe Asp Ala Asp Leu Phe  
210 215 220

Val Asp Cys Ser Gly Phe Arg Gly Leu Leu Ile Asn Lys Ala Met Glu  
225 230 235 240

Glu Pro Phe Leu Asp Met Ser Asp His Leu Leu Asn Asp Ser Ala Val  
245 250 255

Ala Thr Gln Val Pro His Asp Asp Asp Ala Asn Gly Val Glu Pro Phe  
260 265 270

Thr Ser Ala Ile Ala Met Lys Ser Gly Trp Thr Trp Lys Ile Pro Met  
275 280 285

Leu Gly Arg Phe Gly Thr Gly Tyr Val Tyr Ser Ser Arg Phe Ala Thr  
290 295 300

Glu Asp Glu Ala Val Arg Glu Phe Cys Glu Met Trp His Leu Asp Pro  
305 310 315 320

Glu Thr Gln Pro Leu Asn Arg Ile Arg Phe Arg Val Gly Arg Asn Arg  
325 330 335

Arg Ala Trp Val Gly Asn Cys Val Ser Ile Gly Thr Ser Ser Cys Phe  
340 345 350

Val Glu Pro Leu Glu Ser Thr Gly Ile Tyr Phe Val Tyr Ala Ala Leu  
355 360 365

Tyr Gln Leu Val Lys His Phe Pro Asp Lys Ser Leu Asn Pro Val Leu

370

375

380

Thr Ala Arg Phe Asn Arg Glu Ile Glu Thr Met Phe Asp Asp Thr Arg  
385 390 395 400

Asp Phe Ile Gln Ala His Phe Tyr Phe Ser Pro Arg Thr Asp Thr Pro  
405 410 415

Phe Trp Arg Ala Asn Lys Glu Leu Arg Leu Ala Asp Gly Met Gln Glu  
420 425 430

Lys Ile Asp Met Tyr Arg Ala Gly Met Val Ile Asn Ala Pro Ala Ser  
435 440 445

Asp Asp Ala Gln Leu Tyr Tyr Gly Asn Phe Glu Glu Glu Phe Arg Asn  
450 455 460

Phe Trp Asn Asn Ser Asn Tyr Tyr Cys Val Leu Ala Gly Leu Gly Leu  
465 470 475 480

Val Pro Asp Ala Pro Ser Pro Arg Leu Ala His Met Pro Gln Ala Thr  
485 490 495

Glu Ser Val Asp Glu Val Phe Gly Ala Val Lys Asp Arg Gln Arg Asn  
500 505 510

Leu Leu Glu Thr Leu Pro Ser Leu His Glu Phe Leu Arg Gln Gln His  
515 520 525

Gly Arg  
530

<210> 36  
<211> 1446  
<212> DNA  
<213> Bacillus atrophaeus

<400> 36  
atgatgtctg aaaatttgca attgtcagct gaagaaatga gacaattggg ttaccaagca 60  
gttgatttga tcatcgatca catgaaccat ttgaagtcta agccagtttc agaaacaatc 120  
gattctgata tcttgagaaa taagttgact gaatctatcc cagaaaatgg ttcagatcca 180  
aaggaattgt tgcatttctt gaacagaaac gtttttaatc aaattacaca tgttgatcat 240  
ccacatttct tggcttttgt tccaggtcca aataattacg ttggtgttgt tgcagatttc 300  
ttggcttctg gttttaatgt ttttccaact gcatggattg ctggtgcagg tgctgaacaa 360  
atcgaattga ctacaattaa ttggttgaaa tctatgttgg gttttccaga ttcagctgaa 420  
ggtttatttg tttctggtgg ttcaatggca aatttgacag ctttgactgt tgcaagacag 480  
gctaagtga acaacgatat cgaaaatgct gttgtttact tctctgatca aacacatttc 540  
tcagttgata gagcattgaa gggttttaggt tttaaactc atcaaactcg tagaatcgaa 600

|                                                                     |      |
|---------------------------------------------------------------------|------|
| acagatgaac atttgagaat ctctgtttca gctttgaaga aacaaattaa agaagataga   | 660  |
| actaagggtta aaaagccatt ctgtgtttatt gcaaatgctg gtactacaaa ttgtggtgct | 720  |
| gttgattcctt tgaacgaatt agcagatttg tgtaacgatg aagatgtttg gttgcatgct  | 780  |
| gatggttcctt atgggtgctcc agctatcttg tctgaaaagg gttcagctat gttgcaaggt | 840  |
| attcatagag cagattcttt gacttttagat ccacataagt ggttgttcca accatacgat  | 900  |
| gttggttggtg ttttgatcag aaactctcaa tatttgtcaa agacttttag aatgatgcca  | 960  |
| gaatacatca aggattcaga aactaacgtt gaaggtgaaa ttaatttcgg tgaatgtggt   | 1020 |
| atcgaattgt caagaagatt cagagctttg aagggttggt tgtcttttaa agttttcggt   | 1080 |
| gttgctgctt ttagacaagc aatcgatcat ggtatcatgt tagcagaaca agttgaagca   | 1140 |
| tttttgggta aagcaaaaga ttgggaagtt gttacaccag ctcaattggg tatcgttact   | 1200 |
| tttagataca ttccatctga attggcatca acagatacta ttaatgaaat taataagaaa   | 1260 |
| ttggttaagg aaatcacaca tagaggtttc gctatgttat ctactacaga attgaaggaa   | 1320 |
| aaggttggtta ttagattgtg ttcaattaat ccaagaacta caactgaaga aatgttgcaa  | 1380 |
| atcatgatga agattaaagc attggctgaa gaagtttcta tttcataccc atgtgttgct   | 1440 |
| gaataa                                                              | 1446 |

<210> 37  
 <211> 481  
 <212> PRT  
 <213> Bacillus atrophaeus  
 <400> 37

|                                                                 |  |
|-----------------------------------------------------------------|--|
| Met Met Ser Glu Asn Leu Gln Leu Ser Ala Glu Glu Met Arg Gln Leu |  |
| 1                  5                  10                  15    |  |

|                                                                 |  |
|-----------------------------------------------------------------|--|
| Gly Tyr Gln Ala Val Asp Leu Ile Ile Asp His Met Asn His Leu Lys |  |
| 20                  25                  30                      |  |

|                                                                 |  |
|-----------------------------------------------------------------|--|
| Ser Lys Pro Val Ser Glu Thr Ile Asp Ser Asp Ile Leu Arg Asn Lys |  |
| 35                  40                  45                      |  |

|                                                                 |  |
|-----------------------------------------------------------------|--|
| Leu Thr Glu Ser Ile Pro Glu Asn Gly Ser Asp Pro Lys Glu Leu Leu |  |
| 50                  55                  60                      |  |

|                                                                 |  |
|-----------------------------------------------------------------|--|
| His Phe Leu Asn Arg Asn Val Phe Asn Gln Ile Thr His Val Asp His |  |
| 65                  70                  75                  80  |  |

|                                                                 |  |
|-----------------------------------------------------------------|--|
| Pro His Phe Leu Ala Phe Val Pro Gly Pro Asn Asn Tyr Val Gly Val |  |
| 85                  90                  95                      |  |

|                                                                 |  |
|-----------------------------------------------------------------|--|
| Val Ala Asp Phe Leu Ala Ser Gly Phe Asn Val Phe Pro Thr Ala Trp |  |
| 100                  105                  110                   |  |

|                                                                 |  |
|-----------------------------------------------------------------|--|
| Ile Ala Gly Ala Gly Ala Glu Gln Ile Glu Leu Thr Thr Ile Asn Trp |  |
| 115                  120                  125                   |  |

Leu Lys Ser Met Leu Gly Phe Pro Asp Ser Ala Glu Gly Leu Phe Val  
130 135 140

Ser Gly Gly Ser Met Ala Asn Leu Thr Ala Leu Thr Val Ala Arg Gln  
145 150 155 160

Ala Lys Leu Asn Asn Asp Ile Glu Asn Ala Val Val Tyr Phe Ser Asp  
165 170 175

Gln Thr His Phe Ser Val Asp Arg Ala Leu Lys Val Leu Gly Phe Lys  
180 185 190

His His Gln Ile Cys Arg Ile Glu Thr Asp Glu His Leu Arg Ile Ser  
195 200 205

Val Ser Ala Leu Lys Lys Gln Ile Lys Glu Asp Arg Thr Lys Gly Lys  
210 215 220

Lys Pro Phe Cys Val Ile Ala Asn Ala Gly Thr Thr Asn Cys Gly Ala  
225 230 235 240

Val Asp Ser Leu Asn Glu Leu Ala Asp Leu Cys Asn Asp Glu Asp Val  
245 250 255

Trp Leu His Ala Asp Gly Ser Tyr Gly Ala Pro Ala Ile Leu Ser Glu  
260 265 270

Lys Gly Ser Ala Met Leu Gln Gly Ile His Arg Ala Asp Ser Leu Thr  
275 280 285

Leu Asp Pro His Lys Trp Leu Phe Gln Pro Tyr Asp Val Gly Cys Val  
290 295 300

Leu Ile Arg Asn Ser Gln Tyr Leu Ser Lys Thr Phe Arg Met Met Pro  
305 310 315 320

Glu Tyr Ile Lys Asp Ser Glu Thr Asn Val Glu Gly Glu Ile Asn Phe  
325 330 335

Gly Glu Cys Gly Ile Glu Leu Ser Arg Arg Phe Arg Ala Leu Lys Val  
340 345 350

Trp Leu Ser Phe Lys Val Phe Gly Val Ala Ala Phe Arg Gln Ala Ile  
355 360 365

Asp His Gly Ile Met Leu Ala Glu Gln Val Glu Ala Phe Leu Gly Lys  
370 375 380

Ala Lys Asp Trp Glu Val Val Thr Pro Ala Gln Leu Gly Ile Val Thr  
385 390 395 400

Phe Arg Tyr Ile Pro Ser Glu Leu Ala Ser Thr Asp Thr Ile Asn Glu  
405 410 415

Ile Asn Lys Lys Leu Val Lys Glu Ile Thr His Arg Gly Phe Ala Met  
420 425 430

Leu Ser Thr Thr Glu Leu Lys Glu Lys Val Val Ile Arg Leu Cys Ser  
435 440 445

Ile Asn Pro Arg Thr Thr Thr Glu Glu Met Leu Gln Ile Met Met Lys  
450 455 460

Ile Lys Ala Leu Ala Glu Glu Val Ser Ile Ser Tyr Pro Cys Val Ala  
465 470 475 480

Glu

<210> 38  
<211> 1473  
<212> DNA  
<213> Ruminococcus gnayus

<400> 38  
atgtctcaag tgatcaaaaa aaagagaaac actttcatga tcggtaccga atatatttg 60  
aactctaccc aattagaaga ggctattaaa tctttcgtcc atgatttctg tgccgaaaag 120  
cacgaaattc acgaccaacc agttgtcgtt gaagctaagg aacaccaaga agataagatt 180  
aagcaaatca aaattccaga aaagggtcgt ccagtcaatg aagttgtttc tgaaatgatg 240  
aacgaagtct acagatacag aggtgatgcc aaccatccaa gattcttctc tttcgttcca 300  
gggccagctt cctctgtctc ctggttgggt gacatcatga cttctgctta caacatccac 360  
gctggtgggt ctaaattggc cccaatgggt aactgtatcg aacaagaagt cttgaaatgg 420  
ctcgcaaagc aagttggttt caccgaaaac cctggtgggt ttttcgtttc cgggtggttct 480  
atggctaaca ttactgcttt gacagccgct agagacaaca agcttaccga tattaacttg 540  
cacttgggta ctgcttacat ttccgaccaa acccactctt cagttgctaa gggtttacgt 600  
atcattggta tcaactgact cagaatcaga agaattccaa ctaactccca cttccaaatg 660  
gacaccacta agttggaaga agctatcgaa accgacaaga agtccggta catccattt 720  
gttgttatcg gtactgccgg taccaccaac actggttcca ttgaccatt gactgaaatt 780  
tccgctttgt gtaagaagca tgacatgtgg ttccacattg atgggtgccta cgggtgcttca 840  
gttttgttgt ctcaaagta caagtctttg ttaacaggta ccggtttggc tgactctatc 900  
tcttgggatg ctcaaatg gttatttcaa acttacggtt gtgctatggt tttggtcaag 960  
gatattagaa acttgtttca ttccttccac gtcaatccag aatacttgaa ggatctggaa 1020  
aacgacatcg ataacgttaa cacctgggac atcggcattg aattgaccag accagccaga 1080  
ggtttgaagt tgtggttgac tttgcaagtt ctaggttctg atcttattgg ttctgccatt 1140  
gaacacggtt tccaattggc tgtctgggct gaagaagcct tgaatccaaa gaaggactgg 1200



|                                                                   |      |
|-------------------------------------------------------------------|------|
| gaaattgtct ctcctgctca aatggctatg atcaacttca gatatgctcc aaaggacttg | 1260 |
| accaaggaag aacaagatat cttgaacgaa aagatctctc accgtatttt ggaatccggt | 1320 |
| tacgctgcta tcttcactac cgtcctaaac ggtaagactg tcttgagaat ctgcgctatt | 1380 |
| caccagaag ctactcaaga agacatgcaa cacaccattg atttgttgga ccaatacggg  | 1440 |
| agagaaatct acactgaaat gaagaaggct tag                              | 1473 |

<210> 39  
 <211> 490  
 <212> PRT  
 <213> Ruminococcus gnavus

<400> 39

|                                                                 |     |     |     |     |
|-----------------------------------------------------------------|-----|-----|-----|-----|
| Met Ser Gln Val Ile Lys Lys Lys Arg Asn Thr Phe Met Ile Gly Thr | 1   | 5   | 10  | 15  |
| Glu Tyr Ile Leu Asn Ser Thr Gln Leu Glu Glu Ala Ile Lys Ser Phe | 20  | 25  | 30  |     |
| Val His Asp Phe Cys Ala Glu Lys His Glu Ile His Asp Gln Pro Val | 35  | 40  | 45  |     |
| Val Val Glu Ala Lys Glu His Gln Glu Asp Lys Ile Lys Gln Ile Lys | 50  | 55  | 60  |     |
| Ile Pro Glu Lys Gly Arg Pro Val Asn Glu Val Val Ser Glu Met Met | 65  | 70  | 75  | 80  |
| Asn Glu Val Tyr Arg Tyr Arg Gly Asp Ala Asn His Pro Arg Phe Phe | 85  | 90  | 95  |     |
| Ser Phe Val Pro Gly Pro Ala Ser Ser Val Ser Trp Leu Gly Asp Ile | 100 | 105 | 110 |     |
| Met Thr Ser Ala Tyr Asn Ile His Ala Gly Gly Ser Lys Leu Ala Pro | 115 | 120 | 125 |     |
| Met Val Asn Cys Ile Glu Gln Glu Val Leu Lys Trp Leu Ala Lys Gln | 130 | 135 | 140 |     |
| Val Gly Phe Thr Glu Asn Pro Gly Gly Val Phe Val Ser Gly Gly Ser | 145 | 150 | 155 | 160 |
| Met Ala Asn Ile Thr Ala Leu Thr Ala Ala Arg Asp Asn Lys Leu Thr | 165 | 170 | 175 |     |
| Asp Ile Asn Leu His Leu Gly Thr Ala Tyr Ile Ser Asp Gln Thr His | 180 | 185 | 190 |     |
| Ser Ser Val Ala Lys Gly Leu Arg Ile Ile Gly Ile Thr Asp Ser Arg | 195 | 200 | 205 |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ile | Arg | Arg | Ile | Pro | Thr | Asn | Ser | His | Phe | Gln | Met | Asp | Thr | Thr | Lys | 210 | 215 | 220 |
| Leu | Glu | Glu | Ala | Ile | Glu | Thr | Asp | Lys | Lys | Ser | Gly | Tyr | Ile | Pro | Phe | 225 | 230 | 235 |
| Val | Val | Ile | Gly | Thr | Ala | Gly | Thr | Thr | Asn | Thr | Gly | Ser | Ile | Asp | Pro | 245 | 250 | 255 |
| Leu | Thr | Glu | Ile | Ser | Ala | Leu | Cys | Lys | Lys | His | Asp | Met | Trp | Phe | His | 260 | 265 | 270 |
| Ile | Asp | Gly | Ala | Tyr | Gly | Ala | Ser | Val | Leu | Leu | Ser | Pro | Lys | Tyr | Lys | 275 | 280 | 285 |
| Ser | Leu | Leu | Thr | Gly | Thr | Gly | Leu | Ala | Asp | Ser | Ile | Ser | Trp | Asp | Ala | 290 | 295 | 300 |
| His | Lys | Trp | Leu | Phe | Gln | Thr | Tyr | Gly | Cys | Ala | Met | Val | Leu | Val | Lys | 305 | 310 | 315 |
| Asp | Ile | Arg | Asn | Leu | Phe | His | Ser | Phe | His | Val | Asn | Pro | Glu | Tyr | Leu | 325 | 330 | 335 |
| Lys | Asp | Leu | Glu | Asn | Asp | Ile | Asp | Asn | Val | Asn | Thr | Trp | Asp | Ile | Gly | 340 | 345 | 350 |
| Met | Glu | Leu | Thr | Arg | Pro | Ala | Arg | Gly | Leu | Lys | Leu | Trp | Leu | Thr | Leu | 355 | 360 | 365 |
| Gln | Val | Leu | Gly | Ser | Asp | Leu | Ile | Gly | Ser | Ala | Ile | Glu | His | Gly | Phe | 370 | 375 | 380 |
| Gln | Leu | Ala | Val | Trp | Ala | Glu | Glu | Ala | Leu | Asn | Pro | Lys | Lys | Asp | Trp | 385 | 390 | 395 |
| Glu | Ile | Val | Ser | Pro | Ala | Gln | Met | Ala | Met | Ile | Asn | Phe | Arg | Tyr | Ala | 405 | 410 | 415 |
| Pro | Lys | Asp | Leu | Thr | Lys | Glu | Glu | Gln | Asp | Ile | Leu | Asn | Glu | Lys | Ile | 420 | 425 | 430 |
| Ser | His | Arg | Ile | Leu | Glu | Ser | Gly | Tyr | Ala | Ala | Ile | Phe | Thr | Thr | Val | 435 | 440 | 445 |
| Leu | Asn | Gly | Lys | Thr | Val | Leu | Arg | Ile | Cys | Ala | Ile | His | Pro | Glu | Ala | 450 | 455 | 460 |
| Thr | Gln | Glu | Asp | Met | Gln | His | Thr | Ile | Asp | Leu | Leu | Asp | Gln | Tyr | Gly | 465 | 470 | 475 |

Arg Glu Ile Tyr Thr Glu Met Lys Lys Ala  
485 490

<210> 40  
<211> 1482  
<212> DNA  
<213> *Streptomyces griseofuscus*

<400> 40  
atgaagccag ctgaagctaa accacctcac atggaccatg atactttcag atccttgggt 60  
caccaagcta ttgactggat tgctgattac tggcaacact tggctgacag accagttgct 120  
ccattggttg aaccaggtac tgtagagcc caattgccta cactccacc agagtacggt 180  
gaagacttcc cagctttgtt atctgacttg gaaagaattg tcttgccagg cttgttgcac 240  
tggcaacacc caagattctt cggttacttc ccagctaacg cttccggtcc agccgtttta 300  
gctgaattat tatctgctgg tctaggtgtc caaggtatga actggaacac ttctccagct 360  
tgtaccgaag ttgaacaaca aatgttggac tggttcgttc atttgttagg gttgccagaa 420  
cacttcagag gtggtggtgt tatccaagac actgcctcct ccgccgtttt ggctgctcta 480  
ttgactgctt tgcacagagc ttctgctggt agaaccagag atcacggtac tggatgaatgt 540  
ggttacaggg ttacctaac cgctgaaact cactccgctg ctagaaaggc tgccgtcatc 600  
accggtttgg gtttgcgtgc tgtttgtgaa gtcgcaactg atgctgacgg tgctatggat 660  
gccgctgact tggccgctcg tattcaagcc gatttggctg ccggtttgac cccattgatg 720  
gttgttgcta cccgtggtac cacctctgat ttgtccttcg atccattgga agacatcggt 780  
ccagtctgta gaagacatgg tgtctgggta cacgttgacg ctgcttatgc tgggtgttgct 840  
gcagtgtgcg atgaattgca ttgggttaat gatggtgtta gatacgctga ttcttactgt 900  
actaaccac acaagtgggt gttgaccaac ttcgactgtg atttattatg gggtgccgac 960  
ccagaagctt tggcaacgc tttgtccgtt ttgcctgctt acttgagaaa cccagcttct 1020  
gaatcaggta gagttactga ctacagacac tggcaagttc cattgggcag aagattcaga 1080  
gccttgaagt tgtggtctgt cttaagatgg tacggtgccg aaggtttgag atctcacatc 1140  
agaactgggtg tcagacacgc cagattgttc gctgacctgg tcagagaaga tgaccgtttc 1200  
acctgttggt cgccagctac tttgggtttg gtcactttcc gtcacgctgg ttctgacacc 1260  
gacaacagaa acttgttgca cgctatcaac gccgaaggta caactttttt aactcatagt 1320  
gaaaagaacg gtaccttttt cttgagattt gctaccggtg gtaccacac tgaagaccat 1380  
cacgtccacg aagcttggag agctgtccaa aatgccgtcc aaaacgccgt cccaagagcc 1440  
caacatgcc a ccgatggtag tgccgatgct ttgaacggtt aa 1482

<210> 41  
<211> 493  
<212> PRT  
<213> *Streptomyces griseofuscus*

<400> 41

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Lys | Pro | Ala | Glu | Ala | Lys | Pro | Pro | His | Met | Asp | His | Asp | Thr | Phe | 1   | 5   | 10  | 15  |
| Arg | Ser | Leu | Gly | His | Gln | Ala | Ile | Asp | Trp | Ile | Ala | Asp | Tyr | Trp | Gln | 20  | 25  | 30  |     |
| His | Leu | Ala | Asp | Arg | Pro | Val | Ala | Pro | Leu | Val | Glu | Pro | Gly | Thr | Val | 35  | 40  | 45  |     |
| Arg | Ala | Gln | Leu | Pro | Thr | Thr | Pro | Pro | Glu | Tyr | Gly | Glu | Asp | Phe | Pro | 50  | 55  | 60  |     |
| Ala | Leu | Leu | Ser | Asp | Leu | Glu | Arg | Ile | Val | Leu | Pro | Gly | Leu | Leu | His | 65  | 70  | 75  | 80  |
| Trp | Gln | His | Pro | Arg | Phe | Phe | Gly | Tyr | Phe | Pro | Ala | Asn | Ala | Ser | Gly | 85  | 90  | 95  |     |
| Pro | Ala | Val | Leu | Ala | Glu | Leu | Leu | Ser | Ala | Gly | Leu | Gly | Val | Gln | Gly | 100 | 105 | 110 |     |
| Met | Asn | Trp | Asn | Thr | Ser | Pro | Ala | Cys | Thr | Glu | Val | Glu | Gln | Gln | Met | 115 | 120 | 125 |     |
| Leu | Asp | Trp | Phe | Val | His | Leu | Leu | Gly | Leu | Pro | Glu | His | Phe | Arg | Gly | 130 | 135 | 140 |     |
| Gly | Gly | Val | Ile | Gln | Asp | Thr | Ala | Ser | Ser | Ala | Val | Leu | Val | Ala | Leu | 145 | 150 | 155 | 160 |
| Leu | Thr | Ala | Leu | His | Arg | Ala | Ser | Ala | Gly | Arg | Thr | Arg | Asp | His | Gly | 165 | 170 | 175 |     |
| Thr | Gly | Glu | Cys | Gly | Tyr | Arg | Val | Tyr | Leu | Thr | Ala | Glu | Thr | His | Ser | 180 | 185 | 190 |     |
| Ala | Ala | Arg | Lys | Ala | Ala | Val | Ile | Thr | Gly | Leu | Gly | Leu | Arg | Ala | Val | 195 | 200 | 205 |     |
| Cys | Glu | Val | Ala | Thr | Asp | Ala | Asp | Gly | Ala | Met | Asp | Ala | Ala | Asp | Leu | 210 | 215 | 220 |     |
| Ala | Ala | Arg | Ile | Gln | Ala | Asp | Leu | Ala | Ala | Gly | Leu | Thr | Pro | Leu | Met | 225 | 230 | 235 | 240 |
| Val | Val | Ala | Thr | Arg | Gly | Thr | Thr | Ser | Asp | Leu | Ser | Phe | Asp | Pro | Leu | 245 | 250 | 255 |     |
| Glu | Asp | Ile | Gly | Pro | Val | Cys | Arg | Arg | His | Gly | Val | Trp | Leu | His | Val | 260 | 265 | 270 |     |
| Asp | Ala | Ala | Tyr | Ala | Gly | Val | Ala | Ala | Val | Cys | Asp | Glu | Leu | His | Trp |     |     |     |     |

275

280

285

Val Asn Asp Gly Val Arg Tyr Ala Asp Ser Tyr Cys Thr Asn Pro His  
290 295 300

Lys Trp Leu Leu Thr Asn Phe Asp Cys Asp Leu Leu Trp Val Ala Asp  
305 310 315 320

Pro Glu Ala Leu Val Asn Ala Leu Ser Val Leu Pro Ala Tyr Leu Arg  
325 330 335

Asn Pro Ala Ser Glu Ser Gly Arg Val Thr Asp Tyr Arg His Trp Gln  
340 345 350

Val Pro Leu Gly Arg Arg Phe Arg Ala Leu Lys Leu Trp Ser Val Leu  
355 360 365

Arg Trp Tyr Gly Ala Glu Gly Leu Arg Ser His Ile Arg Thr Gly Val  
370 375 380

Arg His Ala Arg Leu Phe Ala Asp Leu Val Arg Glu Asp Asp Arg Phe  
385 390 395 400

Thr Leu Leu Ala Pro Ala Thr Leu Gly Leu Val Thr Phe Arg His Ala  
405 410 415

Gly Ser Asp Thr Asp Asn Arg Asn Leu Leu His Ala Ile Asn Ala Glu  
420 425 430

Gly Thr Thr Phe Leu Thr His Ser Glu Lys Asn Gly Thr Phe Phe Leu  
435 440 445

Arg Phe Ala Thr Gly Gly Thr His Thr Glu Asp His His Val His Glu  
450 455 460

Ala Trp Arg Ala Val Gln Asn Ala Val Gln Asn Ala Val Pro Arg Ala  
465 470 475 480

Gln His Ala Thr Asp Gly Thr Ala Asp Ala Leu Asn Gly  
485 490

<210> 42

<211> 1482

<212> DNA

<213> Ceriporiopsis subvermispora

<400> 42

atggatattg aagcttttag aaaagcaggt taccaagcta tcgatagaat ttgtgactac 60

tactactctt tgcaaaatag accagttggt ccatcagttc aaccagggtta tttgttagat 120

gcattaccag attctccacc agaacaaggt gaagatttca ctgttattgc tgatgattac 180

caaaagtaca tcttgccagg tttgacacat tggcaacatc catctttctt tgcatacttt 240

```

ccaacagctt gtacttttga aggtattttg ggtgacttgt attctacatc aactgcaaat    300
ccaggtttta attggttggc ttctccagca tgtactgaat tagaaatggt tgttatggat    360
tggtctgcta agttgttagg tttgtcagaa catttcttgc attcttcagg taaaggtggt    420
ggtgttattc aaactacagc atcagaattg gcttttagttg ttgttgttgc tgcaagagaa    480
agatacttga gaattcatcc agatgctaag gcagatgaat tggttatata tactacaact    540
caaacacatt ctttgggtgt taaagctggt ttagtttttg gtatggaatg tagagcattg    600
gaagttaaag ctgaagatgc ttacgcatta agaggtgcaa ctttgaagtc agctttagaa    660
gaagatgaaa agagaggtaa aagacctttt attttggttg caacagttgg tacaacttct    720
tcaggtgcta ttgatagatt ggatgaaatt ggtcaagttt ctgaagatta tccatcatta    780
tggattcatg ttgatgctgc atgggcagggt gttactttgg cttgtccaga atacagaggt    840
acagctcaat tagaaaacat caacgcttat gcaacatctt tcggtactaa cttccataag    900
tggggtttgg ttaactttga tgctgcattg ttgtgggtta aggatagaaa ggatttgaca    960
gatgctttag atgttactcc agaattcttg agaacaaaac aaggtgacgc tgggtgcagtt 1020
gttgatttta gaaattggca tttgggtttg ggtagaagat tcagatcttt gaaggtttgg 1080
ttcgttttga gatcatacgg tgttgaaggt ttagaaaact acatcagaca aggtattaaa 1140
ttgaatgaac attttacttc ttaattaga gcttctttgg atttttcatt agttacagca 1200
ccatcattcg ctttgacagt ttttagattg actccagctg gtgcatcttt gactggttca 1260
gaattgaacg aattaaacag agctttctac gcaagattat cttcaagaca tgatatcatg 1320
ttgacacaaa ctgttttgaa cgggtgttttc tgtatcagat ttgcagttgg tgctgcaaga 1380
actcaacaag aacatattga tacagcttgg gatttgttac aacaagaagg tgctgttgca 1440
gttcaagaat acatgcaaaa gatttctaag gatgttcaat aa                        1482

```

```

<210> 43
<211> 493
<212> PRT
<213> Ceriporiopsis subvermispora

```

```
<400> 43
```

```

Met Asp Ile Glu Ala Phe Arg Lys Ala Gly Tyr Gln Ala Ile Asp Arg
1           5           10          15

```

```

Ile Cys Asp Tyr Tyr Tyr Ser Leu Gln Asn Arg Pro Val Val Pro Ser
          20          25          30

```

```

Val Gln Pro Gly Tyr Leu Leu Asp Ala Leu Pro Asp Ser Pro Pro Glu
          35          40          45

```

```

Gln Gly Glu Asp Phe Thr Val Ile Ala Asp Asp Tyr Gln Lys Tyr Ile
          50          55          60

```

```

Leu Pro Gly Leu Thr His Trp Gln His Pro Ser Phe Phe Ala Tyr Phe
65           70           75           80

```

Pro Thr Ala Cys Thr Phe Glu Gly Ile Leu Gly Asp Leu Tyr Ser Thr  
85 90 95

Ser Thr Ala Asn Pro Gly Phe Asn Trp Leu Ala Ser Pro Ala Cys Thr  
100 105 110

Glu Leu Glu Met Val Val Met Asp Trp Ser Ala Lys Leu Leu Gly Leu  
115 120 125

Ser Glu His Phe Leu His Ser Ser Gly Lys Gly Gly Gly Val Ile Gln  
130 135 140

Thr Thr Ala Ser Glu Leu Ala Leu Val Val Val Val Ala Ala Arg Glu  
145 150 155 160

Arg Tyr Leu Arg Ile His Pro Asp Ala Lys Ala Asp Glu Leu Val Ile  
165 170 175

Tyr Thr Thr Thr Gln Thr His Ser Leu Gly Val Lys Ala Gly Leu Val  
180 185 190

Phe Gly Met Glu Cys Arg Ala Leu Glu Val Lys Ala Glu Asp Ala Tyr  
195 200 205

Ala Leu Arg Gly Ala Thr Leu Lys Ser Ala Leu Glu Glu Asp Glu Lys  
210 215 220

Arg Gly Lys Arg Pro Phe Ile Leu Val Ala Thr Val Gly Thr Thr Ser  
225 230 235 240

Ser Gly Ala Ile Asp Arg Leu Asp Glu Ile Gly Gln Val Ser Glu Asp  
245 250 255

Tyr Pro Ser Leu Trp Ile His Val Asp Ala Ala Trp Ala Gly Val Thr  
260 265 270

Leu Ala Cys Pro Glu Tyr Arg Gly Thr Ala Gln Leu Glu Asn Ile Asn  
275 280 285

Ala Tyr Ala Thr Ser Phe Gly Thr Asn Phe His Lys Trp Gly Leu Val  
290 295 300

Asn Phe Asp Ala Ala Leu Trp Val Lys Asp Arg Lys Asp Leu Thr  
305 310 315 320

Asp Ala Leu Asp Val Thr Pro Glu Phe Leu Arg Thr Lys Gln Gly Asp  
325 330 335

Ala Gly Ala Val Val Asp Phe Arg Asn Trp His Leu Gly Leu Gly Arg  
340 345 350

Arg Phe Arg Ser Leu Lys Val Trp Phe Val Leu Arg Ser Tyr Gly Val  
355 360 365

Glu Gly Phe Arg Asn Tyr Ile Arg Gln Gly Ile Lys Leu Asn Glu His  
370 375 380

Phe Thr Ser Leu Ile Arg Ala Ser Leu Asp Phe Ser Leu Val Thr Ala  
385 390 395 400

Pro Ser Phe Ala Leu Thr Val Phe Arg Leu Thr Pro Ala Gly Ala Ser  
405 410 415

Leu Thr Gly Ser Glu Leu Asn Glu Leu Asn Arg Ala Phe Tyr Ala Arg  
420 425 430

Leu Ser Ser Arg His Asp Ile Met Leu Thr Gln Thr Val Leu Asn Gly  
435 440 445

Val Phe Cys Ile Arg Phe Ala Val Gly Ala Ala Arg Thr Gln Gln Glu  
450 455 460

His Ile Asp Thr Ala Trp Asp Leu Leu Gln Gln Glu Gly Ala Val Ala  
465 470 475 480

Val Gln Glu Tyr Met Gln Lys Ile Ser Lys Asp Val Gln  
485 490

<210> 44  
<211> 1254  
<212> DNA  
<213> Clostridium sporogenes

<400> 44  
atgaagtctt ggagaaagta cacacaacaa gaaatggatg aaaagattac tgaatctttg 60  
gaaaagactt tgaactacga taacactaag acaatcggta ttccaggtac taagttggat 120  
gatacagttt tctatgatga tcattctttc gttaagcatt caccatactt gagaactttt 180  
attcaaaacc caaacatat cggttgtcat acttatgata aggctgatat cttgttcggt 240  
ggtacattcg atatcgaaag agaattaatc caattgttag caatcgatgt tttgaacggt 300  
aacgatgaag aatttgatgg ttacgttact caaggtggta cagaagctaa catccaagca 360  
atgtgggttt acagaaacta cttcaagaaa gaaagaaagg ctaagcatga agaaatcgct 420  
atcatcactt cagcagatac acattactct gcatacaaag gttcagattt gttgaacatc 480  
gatattatta aggttccagt tgatttttat tcaagaaaaa ttcaagaaaa tacattggat 540  
tcaattgtta aagaagctaa agaaattggg aaaaagtact tcatcgttat ctctaocatg 600  
ggtactacaa tgtttggttc agttgatgat ccagatttgt acgctaacat cttcgataag 660  
tacaatttgg aatacaaaat tcatgttgat ggtgcatttg gtggttttat atatccaatt 720  
gataataagg aatgtaaaac tgatttctct aataagaacg tttcttcaat cacattagat 780  
ggtcataaga tgttgcaagc tccatacggg actggtatct tcgtttcaag aaagaatttg 840



|                                                                    |      |
|--------------------------------------------------------------------|------|
| atccataaca ctttgacaaa ggaagcaact tacatcgaaa atttggatgt tacattgtct  | 900  |
| ggttcaagat ctggttcaaa tgctgttgca atttggatgg ttttagcttc ttatgggtcca | 960  |
| tacggttgga tggaaaagat taataagttg agaaatagaa ctaaattggt gtgtaagcaa  | 1020 |
| ttgaacgata tgagaattaa atattacaaa gaagattcaa tgaatattgt tacaattgaa  | 1080 |
| gaacaatatg ttaataagga aatcgctgaa aagtactttt tagttccaga agttcataac  | 1140 |
| ccaactaaca actggtacaa gatcgttgtt atggaacatg ttgaattgga tatcttgaac  | 1200 |
| tctttggttt acgatttgag aaagtttaat aaggaacatt tgaaggcaat gtaa        | 1254 |

<210> 45  
 <211> 417  
 <212> PRT  
 <213> Clostridium sporogenes  
 <400> 45

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Lys | Phe | Trp | Arg | Lys | Tyr | Thr | Gln | Gln | Glu | Met | Asp | Glu | Lys | Ile | 1   | 5   | 10  | 15  |
| Thr | Glu | Ser | Leu | Glu | Lys | Thr | Leu | Asn | Tyr | Asp | Asn | Thr | Lys | Thr | Ile | 20  | 25  | 30  |     |
| Gly | Ile | Pro | Gly | Thr | Lys | Leu | Asp | Asp | Thr | Val | Phe | Tyr | Asp | Asp | His | 35  | 40  | 45  |     |
| Ser | Phe | Val | Lys | His | Ser | Pro | Tyr | Leu | Arg | Thr | Phe | Ile | Gln | Asn | Pro | 50  | 55  | 60  |     |
| Asn | His | Ile | Gly | Cys | His | Thr | Tyr | Asp | Lys | Ala | Asp | Ile | Leu | Phe | Gly | 65  | 70  | 75  | 80  |
| Gly | Thr | Phe | Asp | Ile | Glu | Arg | Glu | Leu | Ile | Gln | Leu | Leu | Ala | Ile | Asp | 85  | 90  | 95  |     |
| Val | Leu | Asn | Gly | Asn | Asp | Glu | Glu | Phe | Asp | Gly | Tyr | Val | Thr | Gln | Gly | 100 | 105 | 110 |     |
| Gly | Thr | Glu | Ala | Asn | Ile | Gln | Ala | Met | Trp | Val | Tyr | Arg | Asn | Tyr | Phe | 115 | 120 | 125 |     |
| Lys | Lys | Glu | Arg | Lys | Ala | Lys | His | Glu | Glu | Ile | Ala | Ile | Ile | Thr | Ser | 130 | 135 | 140 |     |
| Ala | Asp | Thr | His | Tyr | Ser | Ala | Tyr | Lys | Gly | Ser | Asp | Leu | Leu | Asn | Ile | 145 | 150 | 155 | 160 |
| Asp | Ile | Ile | Lys | Val | Pro | Val | Asp | Phe | Tyr | Ser | Arg | Lys | Ile | Gln | Glu | 165 | 170 | 175 |     |
| Asn | Thr | Leu | Asp | Ser | Ile | Val | Lys | Glu | Ala | Lys | Glu | Ile | Gly | Lys | Lys | 180 | 185 | 190 |     |

Tyr Phe Ile Val Ile Ser Asn Met Gly Thr Thr Met Phe Gly Ser Val  
195 200 205

Asp Asp Pro Asp Leu Tyr Ala Asn Ile Phe Asp Lys Tyr Asn Leu Glu  
210 215 220

Tyr Lys Ile His Val Asp Gly Ala Phe Gly Gly Phe Ile Tyr Pro Ile  
225 230 235 240

Asp Asn Lys Glu Cys Lys Thr Asp Phe Ser Asn Lys Asn Val Ser Ser  
245 250 255

Ile Thr Leu Asp Gly His Lys Met Leu Gln Ala Pro Tyr Gly Thr Gly  
260 265 270

Ile Phe Val Ser Arg Lys Asn Leu Ile His Asn Thr Leu Thr Lys Glu  
275 280 285

Ala Thr Tyr Ile Glu Asn Leu Asp Val Thr Leu Ser Gly Ser Arg Ser  
290 295 300

Gly Ser Asn Ala Val Ala Ile Trp Met Val Leu Ala Ser Tyr Gly Pro  
305 310 315 320

Tyr Gly Trp Met Glu Lys Ile Asn Lys Leu Arg Asn Arg Thr Lys Trp  
325 330 335

Leu Cys Lys Gln Leu Asn Asp Met Arg Ile Lys Tyr Tyr Lys Glu Asp  
340 345 350

Ser Met Asn Ile Val Thr Ile Glu Glu Gln Tyr Val Asn Lys Glu Ile  
355 360 365

Ala Glu Lys Tyr Phe Leu Val Pro Glu Val His Asn Pro Thr Asn Asn  
370 375 380

Trp Tyr Lys Ile Val Val Met Glu His Val Glu Leu Asp Ile Leu Asn  
385 390 395 400

Ser Leu Val Tyr Asp Leu Arg Lys Phe Asn Lys Glu His Leu Lys Ala  
405 410 415

Met

<210> 46  
<211> 1443  
<212> DNA  
<213> Homo sapiens

<400> 46  
atgaatgcat ctgaattcag aagaagaggt aaagaaatgg ttgattacat ggctaactac

|                                                                    |      |
|--------------------------------------------------------------------|------|
| atggaaggta ttgaaggtag acaagtttat ccagatgttg aaccaggtta cttgagacca  | 120  |
| ttaattccag ctgcagctcc acaagaacca gatacatttg aagatattat taacgatgtt  | 180  |
| gaaaagatta tcatgccagg tgttacacat tggcattcac catatttctt tgcatacttt  | 240  |
| ccaactgctt cttcatatcc agcaatgttg gctgatatgt tatgtggtgc aattggttgt  | 300  |
| attggttttt cttgggcagc ttcaccagct tgtacagaat tggaaactgt tatgatggat  | 360  |
| tggttaggta aaatgttggg attaccaaag gcatttttga acgaaaaggc tggatgaagg  | 420  |
| ggtggtgtta ttcaaggttc tgcacagaa gctacattag ttgctttgtt agcagctaga   | 480  |
| actaaagtta ttcataagatt gcaagcagct tctccagaat taacacaagc agctatcatg | 540  |
| gaaaaattgg ttgcatattc ttcagatcaa gctcattctt cagttgaaag agctggtttg  | 600  |
| atcggtggtg ttaagttgaa ggcaatccca tctgatggta attttgctat gagagcatca  | 660  |
| gctttgcaag aagcattaga aagagataaa gcagctggtt tgattccatt tttcatggtt  | 720  |
| gctactttag gtactacaac ttgttgttct ttcgataatt tgttagaagt tggccaatt   | 780  |
| tgtaataagg aagatatttg gttgcatgtt gatgcagctt acgcaggttc agcttttatt  | 840  |
| tgtccagaat tcagacattt gttgaacggg gttgaattcg ctgattcttt taatttcaat  | 900  |
| ccacataaat ggttgtagt taacttcgat tgttcagcaa tgtgggttaa gaaaagaaca   | 960  |
| gatttgactg gtgcttttag attagatcca acatatttga agcattctca tcaagattca  | 1020 |
| ggtttgatca ctgattacag acattggcaa attccattgg gtagaagatt cagatctttg  | 1080 |
| aagatgtggt tcgttttttag aatgtacggg gttgaagggt tgcaagctta catcagaaag | 1140 |
| catgttcaat tgtctcatga attcgaatca ttagttagac aagatccaag attcgaatc   | 1200 |
| tgtgttgaag ttattttggg tttagtttgt ttcagattga agggttctaa taaggttaac  | 1260 |
| gaagcattgt tgcaaagaat taattcagct aagaaaattc atttggttcc atgtcatttg  | 1320 |
| cgtgataagt tcgttttgag attcgctatt tgttcaagaa ctgttgaatc agcacatgtt  | 1380 |
| caaagagctt gggaacatat caaggaattg gcagctgatg ttttgagagc tgaaagagaa  | 1440 |
| taa                                                                | 1443 |

<210> 47  
 <211> 480  
 <212> PRT  
 <213> Homo sapiens

<400> 47

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ala | Ser | Glu | Phe | Arg | Arg | Arg | Gly | Lys | Glu | Met | Val | Asp | Tyr |
| 1   |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ala | Asn | Tyr | Met | Glu | Gly | Ile | Glu | Gly | Arg | Gln | Val | Tyr | Pro | Asp |
|     |     |     | 20  |     |     |     | 25  |     |     |     |     | 30  |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Val | Glu | Pro | Gly | Tyr | Leu | Arg | Pro | Leu | Ile | Pro | Ala | Ala | Ala | Pro | Gln |
|     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |     |

Glu Pro Asp Thr Phe Glu Asp Ile Ile Asn Asp Val Glu Lys Ile Ile  
50 55 60

Met Pro Gly Val Thr His Trp His Ser Pro Tyr Phe Phe Ala Tyr Phe  
65 70 75 80

Pro Thr Ala Ser Ser Tyr Pro Ala Met Leu Ala Asp Met Leu Cys Gly  
85 90 95

Ala Ile Gly Cys Ile Gly Phe Ser Trp Ala Ala Ser Pro Ala Cys Thr  
100 105 110

Glu Leu Glu Thr Val Met Met Asp Trp Leu Gly Lys Met Leu Glu Leu  
115 120 125

Pro Lys Ala Phe Leu Asn Glu Lys Ala Gly Glu Gly Gly Gly Val Ile  
130 135 140

Gln Gly Ser Ala Ser Glu Ala Thr Leu Val Ala Leu Leu Ala Ala Arg  
145 150 155 160

Thr Lys Val Ile His Arg Leu Gln Ala Ala Ser Pro Glu Leu Thr Gln  
165 170 175

Ala Ala Ile Met Glu Lys Leu Val Ala Tyr Ser Ser Asp Gln Ala His  
180 185 190

Ser Ser Val Glu Arg Ala Gly Leu Ile Gly Gly Val Lys Leu Lys Ala  
195 200 205

Ile Pro Ser Asp Gly Asn Phe Ala Met Arg Ala Ser Ala Leu Gln Glu  
210 215 220

Ala Leu Glu Arg Asp Lys Ala Ala Gly Leu Ile Pro Phe Phe Met Val  
225 230 235 240

Ala Thr Leu Gly Thr Thr Thr Cys Cys Ser Phe Asp Asn Leu Leu Glu  
245 250 255

Val Gly Pro Ile Cys Asn Lys Glu Asp Ile Trp Leu His Val Asp Ala  
260 265 270

Ala Tyr Ala Gly Ser Ala Phe Ile Cys Pro Glu Phe Arg His Leu Leu  
275 280 285

Asn Gly Val Glu Phe Ala Asp Ser Phe Asn Phe Asn Pro His Lys Trp  
290 295 300

Leu Leu Val Asn Phe Asp Cys Ser Ala Met Trp Val Lys Lys Arg Thr  
305 310 315 320

Asp Leu Thr Gly Ala Phe Arg Leu Asp Pro Thr Tyr Leu Lys His Ser

325

330

335

His Gln Asp Ser Gly Leu Ile Thr Asp Tyr Arg His Trp Gln Ile Pro  
 340 345 350

Leu Gly Arg Arg Phe Arg Ser Leu Lys Met Trp Phe Val Phe Arg Met  
 355 360 365

Tyr Gly Val Lys Gly Leu Gln Ala Tyr Ile Arg Lys His Val Gln Leu  
 370 375 380

Ser His Glu Phe Glu Ser Leu Val Arg Gln Asp Pro Arg Phe Glu Ile  
 385 390 395 400

Cys Val Glu Val Ile Leu Gly Leu Val Cys Phe Arg Leu Lys Gly Ser  
 405 410 415

Asn Lys Val Asn Glu Ala Leu Leu Gln Arg Ile Asn Ser Ala Lys Lys  
 420 425 430

Ile His Leu Val Pro Cys His Leu Arg Asp Lys Phe Val Leu Arg Phe  
 435 440 445

Ala Ile Cys Ser Arg Thr Val Glu Ser Ala His Val Gln Arg Ala Trp  
 450 455 460

Glu His Ile Lys Glu Leu Ala Ala Asp Val Leu Arg Ala Glu Arg Glu  
 465 470 475 480

&lt;210&gt; 48

&lt;211&gt; 1521

&lt;212&gt; DNA

&lt;213&gt; Ophiorrhiza pumila

&lt;400&gt; 48

atgggttcta tctcagaaaa ctgtgatgat tctatttcat tagctgcacc ttttagacca 60

ttggaaccag aagaattcag aaagcaagct catgttatgg ttgatttcat cgcagattac 120

tacaagaaca ttgaaaatta tccagttttg tcacaagttg aaccaggtta tttgaaaaat 180

agattgccag aaactgctcc acatttgcca gaatctttcg aaacaatctt gaaggatatt 240

aagaaagata tcgttcctgg tatgactaat tggttgtcac caaatttctt tgcttatttt 300

ccagcaactg tttcttcagc tgcatttggt ggtgaaatgt tgtgtacagg ttttaattct 360

gttggtttta attggtagc ttcaccagca tctacagaat tggaaatggt tgttattgat 420

tggtagcta acatgttgaa gttgcaaag tcttttatgt ttcattgtac tgggtggtggt 480

gttattcaag gtactacatc tgaagctatc ttgtgtacat tgattgctgc aagagatggt 540

gctttggaaa agattggtat ggaaaacgtt ggtaaattag ttgtttacgg ttcagatcaa 600

actcattctt tctttcaaaa gacttgtaag gttgctggta tcttcccatg taacatcaag 660

ttgatcccaa ctacaagaga agataacttc tctatgtcac caatcgcat gagagaacaa 720

attgaagctg atgttgcaga tggtttagtt ccaattttct tgtgtactac agttggtact 780  
acatcaactg ctgcaattga tccagtttct gaagttgcta aggttgcaaa cgatttcaat 840  
atttgggttc atgttgatgc tgcttatgct ggttctgcat gtatttgtcc agaattcaga 900  
caatacttag atggtatcga attggttgat tctttttcat tgtctccaca taagtggttg 960  
ttgtgtttct tggattgttg ttgtttgtgg ttgaagaaac cacatttgat ggtaaggct 1020  
ttgtcaacta acccagaata tttgagaaat aagagatctg aattcgatgg tgttgttgat 1080  
ttcaaggatt ggcaaatcgg tacaggtaga agattcaaag ctttgagatt atggttggtt 1140  
atgagatcat acggtgttga aaatttgaag agacatatct tgtctgatgt tcaaattggct 1200  
aagatgttcg aaggtttagt taaatctgat ccaagattcg aaatcatcgt tccaagagct 1260  
ttcgcattag tttgtttcag attgaaccct ggtaaagggt atgatgatga aatcgataag 1320  
gaaatcttga ataaggaatt gttagatttg attaattcaa caggtagagc ttacatgact 1380  
catacaaaag caggtggtat ctatatgttg agatttgctg ttggtactac attgactgaa 1440  
gaacatcatg tttacgctgc atgggaattg attaaagaat gtacagatgc atctttgact 1500  
aaaacaaata ttattgaata a 1521

<210> 49  
<211> 506  
<212> PRT  
<213> Ophiorrhiza pumila

<400> 49

Met Gly Ser Ile Ser Glu Asn Cys Asp Asp Ser Ile Ser Leu Ala Ala  
1 5 10 15

Pro Phe Arg Pro Leu Glu Pro Glu Glu Phe Arg Lys Gln Ala His Val  
20 25 30

Met Val Asp Phe Ile Ala Asp Tyr Tyr Lys Asn Ile Glu Asn Tyr Pro  
35 40 45

Val Leu Ser Gln Val Glu Pro Gly Tyr Leu Lys Asn Arg Leu Pro Glu  
50 55 60

Thr Ala Pro His Leu Pro Glu Ser Phe Glu Thr Ile Leu Lys Asp Ile  
65 70 75 80

Lys Lys Asp Ile Val Pro Gly Met Thr Asn Trp Leu Ser Pro Asn Phe  
85 90 95

Phe Ala Tyr Phe Pro Ala Thr Val Ser Ser Ala Ala Phe Val Gly Glu  
100 105 110

Met Leu Cys Thr Gly Phe Asn Ser Val Gly Phe Asn Trp Leu Ala Ser  
115 120 125

Pro Ala Ser Thr Glu Leu Glu Met Val Val Ile Asp Trp Leu Ala Asn

|            |            |            |            |            |            |            |            |            |            |            |            |            |            |            |            |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 130        |            |            | 135        |            |            |            |            | 140        |            |            |            |            |            |            |            |
| Met<br>145 | Leu        | Lys        | Leu        | Pro        | Lys<br>150 | Ser        | Phe        | Met        | Phe        | His<br>155 | Gly        | Thr        | Gly        | Gly        | Gly<br>160 |
| Val        | Ile        | Gln        | Gly        | Thr<br>165 | Thr        | Ser        | Glu        | Ala        | Ile<br>170 | Leu        | Cys        | Thr        | Leu        | Ile<br>175 | Ala        |
| Ala        | Arg        | Asp        | Gly<br>180 | Ala        | Leu        | Glu        | Lys        | Ile<br>185 | Gly        | Met        | Glu        | Asn        | Val<br>190 | Gly        | Lys        |
| Leu        | Val        | Val<br>195 | Tyr        | Gly        | Ser        | Asp        | Gln<br>200 | Thr        | His        | Ser        | Phe        | Phe<br>205 | Gln        | Lys        | Thr        |
| Cys        | Lys<br>210 | Val        | Ala        | Gly        | Ile        | Phe<br>215 | Pro        | Cys        | Asn        | Ile        | Lys<br>220 | Leu        | Ile        | Pro        | Thr        |
| Thr<br>225 | Arg        | Glu        | Asp        | Asn        | Phe<br>230 | Ser        | Met        | Ser        | Pro        | Ile<br>235 | Ala        | Leu        | Arg        | Glu        | Gln<br>240 |
| Ile        | Glu        | Ala        | Asp        | Val<br>245 | Ala        | Asp        | Gly        | Leu        | Val<br>250 | Pro        | Ile        | Phe        | Leu        | Cys<br>255 | Thr        |
| Thr        | Val        | Gly        | Thr<br>260 | Thr        | Ser        | Thr        | Ala        | Ala<br>265 | Ile        | Asp        | Pro        | Val        | Ser<br>270 | Glu        | Val        |
| Ala        | Lys        | Val<br>275 | Ala        | Asn        | Asp        | Phe        | Asn<br>280 | Ile        | Trp        | Val        | His        | Val<br>285 | Asp        | Ala        | Ala        |
| Tyr        | Ala<br>290 | Gly        | Ser        | Ala        | Cys        | Ile<br>295 | Cys        | Pro        | Glu        | Phe        | Arg<br>300 | Gln        | Tyr        | Leu        | Asp        |
| Gly<br>305 | Ile        | Glu        | Leu        | Val        | Asp<br>310 | Ser        | Phe        | Ser        | Leu        | Ser<br>315 | Pro        | His        | Lys        | Trp        | Leu        |
| Leu        | Cys        | Phe        | Leu        | Asp<br>325 | Cys        | Cys        | Cys        | Leu        | Trp<br>330 | Leu        | Lys        | Lys        | Pro        | His<br>335 | Leu        |
| Met        | Val        | Lys        | Ala<br>340 | Leu        | Ser        | Thr        | Asn        | Pro<br>345 | Glu        | Tyr        | Leu        | Arg        | Asn<br>350 | Lys        | Arg        |
| Ser        | Glu        | Phe<br>355 | Asp        | Gly        | Val        | Val        | Asp<br>360 | Phe        | Lys        | Asp        | Trp        | Gln<br>365 | Ile        | Gly        | Thr        |
| Gly        | Arg<br>370 | Arg        | Phe        | Lys        | Ala        | Leu        | Arg<br>375 | Leu        | Trp        | Leu        | Val<br>380 | Met        | Arg        | Ser        | Tyr        |
| Gly<br>385 | Val        | Glu        | Asn        | Leu        | Lys<br>390 | Arg        | His        | Ile        | Leu        | Ser<br>395 | Asp        | Val        | Gln        | Met        | Ala        |
| Lys        | Met        | Phe        | Glu        | Gly<br>405 | Leu        | Val        | Lys        | Ser        | Asp<br>410 | Pro        | Arg        | Phe        | Glu        | Ile<br>415 | Ile        |

Val Pro Arg Ala Phe Ala Leu Val Cys Phe Arg Leu Asn Pro Gly Lys  
 420 425 430

Gly Tyr Asp Asp Glu Ile Asp Lys Glu Ile Leu Asn Lys Glu Leu Leu  
 435 440 445

Asp Leu Ile Asn Ser Thr Gly Arg Ala Tyr Met Thr His Thr Lys Ala  
 450 455 460

Gly Gly Ile Tyr Met Leu Arg Phe Ala Val Gly Thr Thr Leu Thr Glu  
 465 470 475 480

Glu His His Val Tyr Ala Ala Trp Glu Leu Ile Lys Glu Cys Thr Asp  
 485 490 495

Ala Ser Leu Thr Lys Thr Asn Ile Ile Glu  
 500 505

<210> 50  
 <211> 1545  
 <212> DNA  
 <213> Oryza sativa

<400> 50  
 atgggttcat tagatacaaa tccaactgct ttttctgcat ttccagctgg tgaaggtgaa 60  
 acttttcaac cattgaacgc agatgatgtt agatcatatt tgcataaggc tggtgatttc 120  
 atctctgatt actacaaatc agttgaatct atgccagttt tgccaaacgt taagccaggt 180  
 tatttgcaag atgaattaag agcatcacca ccaacttact ctgctccatt cgatgttaca 240  
 atgaaggaat tgagatcttc agttgttctt ggtatgactc attgggcatc tccaaatttc 300  
 tttgctttct ttccatcaac aaattctgct gcagctattg ctggtgactt gattgcatca 360  
 gctatgaata cagttggttt tacttggcaa gcttctccag cagctactga aatggaagtt 420  
 ttggcattag attggttggc tcaaattgtt aatttgccaa cttcttttat gaacagaact 480  
 ggtgaaggta gaggtactgg tgggtggtgtt attttgggta ctacatctga agctatgttg 540  
 gttacattag ttgcagctag agatgcagct ttgagaagat ctggttcaga tgggtgttgca 600  
 ggtttgcata gattagctgt ttatgcagct gatcaaacac attcaacttt ctttaaggca 660  
 tgtagattgg ctggttttga tccagcaaatt attagatcta ttccaacagg tgctgaaact 720  
 gattatgggt tagatccagc aagattgtta gaagctatgc aagcagatgc tgatgcaggt 780  
 ttagttccaa catacgtttg tgcaactgtt ggtactacat cttcaaatgc tgttgatcca 840  
 gttggtgctg ttgcagatgt tgcagctaga tttgcagctt gggttcatgt tgatgcagct 900  
 tacgcagggt cagcttgtat ttgtccagaa ttcagacatc atttggatgg tgttgaaaga 960  
 gttgattcta tctcaatgtc tccacataag tggttgatga catgtttaga ttgtacttgt 1020  
 ttgtacgtta gagatacaca tagattgact ggttcttttag aaactaacc agaatacttg 1080  
 aaaaatcatg cttcagattc tgggtgaagtt acagatttga aagatatgca agttggtgtt 1140



ggtagaagat tcagaggttt gaagttgtgg atggttatga gaacttatgg tgttgctaag 1200  
ttgcaagaac atatcagatc agatgttgct atggcaaaag tttttgaaga tttggttaga 1260  
ggtgacgata gattcgaagt tgttgttcca agaaacttcg ctttggtttg tttcagaatt 1320  
agagctggtg caggtgcagc tgcagctaca gaagaagatg cagatgaagc taacagagaa 1380  
ttgatggaaa gattgaataa gactggtaaa gcatacgttg ctcatacagt tgttggtggt 1440  
agatttgttt tgagatttgc agttggttct tcattacaag aagaacatca tgttagatct 1500  
gcttgggaat tgattaagaa aactacaact gaaatgatga actaa 1545

<210> 51  
<211> 514  
<212> PRT  
<213> Oryza sativa

<400> 51

Met Gly Ser Leu Asp Thr Asn Pro Thr Ala Phe Ser Ala Phe Pro Ala  
1 5 10 15

Gly Glu Gly Glu Thr Phe Gln Pro Leu Asn Ala Asp Asp Val Arg Ser  
20 25 30

Tyr Leu His Lys Ala Val Asp Phe Ile Ser Asp Tyr Tyr Lys Ser Val  
35 40 45

Glu Ser Met Pro Val Leu Pro Asn Val Lys Pro Gly Tyr Leu Gln Asp  
50 55 60

Glu Leu Arg Ala Ser Pro Pro Thr Tyr Ser Ala Pro Phe Asp Val Thr  
65 70 75 80

Met Lys Glu Leu Arg Ser Ser Val Val Pro Gly Met Thr His Trp Ala  
85 90 95

Ser Pro Asn Phe Phe Ala Phe Phe Pro Ser Thr Asn Ser Ala Ala Ala  
100 105 110

Ile Ala Gly Asp Leu Ile Ala Ser Ala Met Asn Thr Val Gly Phe Thr  
115 120 125

Trp Gln Ala Ser Pro Ala Ala Thr Glu Met Glu Val Leu Ala Leu Asp  
130 135 140

Trp Leu Ala Gln Met Leu Asn Leu Pro Thr Ser Phe Met Asn Arg Thr  
145 150 155 160

Gly Glu Gly Arg Gly Thr Gly Gly Gly Val Ile Leu Gly Thr Thr Ser  
165 170 175

Glu Ala Met Leu Val Thr Leu Val Ala Ala Arg Asp Ala Ala Leu Arg  
180 185 190

Arg Ser Gly Ser Asp Gly Val Ala Gly Leu His Arg Leu Ala Val Tyr  
 195 200 205

Ala Ala Asp Gln Thr His Ser Thr Phe Phe Lys Ala Cys Arg Leu Ala  
 210 215 220

Gly Phe Asp Pro Ala Asn Ile Arg Ser Ile Pro Thr Gly Ala Glu Thr  
 225 230 235 240

Asp Tyr Gly Leu Asp Pro Ala Arg Leu Leu Glu Ala Met Gln Ala Asp  
 245 250 255

Ala Asp Ala Gly Leu Val Pro Thr Tyr Val Cys Ala Thr Val Gly Thr  
 260 265 270

Thr Ser Ser Asn Ala Val Asp Pro Val Gly Ala Val Ala Asp Val Ala  
 275 280 285

Ala Arg Phe Ala Ala Trp Val His Val Asp Ala Ala Tyr Ala Gly Ser  
 290 295 300

Ala Cys Ile Cys Pro Glu Phe Arg His His Leu Asp Gly Val Glu Arg  
 305 310 315 320

Val Asp Ser Ile Ser Met Ser Pro His Lys Trp Leu Met Thr Cys Leu  
 325 330 335

Asp Cys Thr Cys Leu Tyr Val Arg Asp Thr His Arg Leu Thr Gly Ser  
 340 345 350

Leu Glu Thr Asn Pro Glu Tyr Leu Lys Asn His Ala Ser Asp Ser Gly  
 355 360 365

Glu Val Thr Asp Leu Lys Asp Met Gln Val Gly Val Gly Arg Arg Phe  
 370 375 380

Arg Gly Leu Lys Leu Trp Met Val Met Arg Thr Tyr Gly Val Ala Lys  
 385 390 395 400

Leu Gln Glu His Ile Arg Ser Asp Val Ala Met Ala Lys Val Phe Glu  
 405 410 415

Asp Leu Val Arg Gly Asp Asp Arg Phe Glu Val Val Val Pro Arg Asn  
 420 425 430

Phe Ala Leu Val Cys Phe Arg Ile Arg Ala Gly Ala Gly Ala Ala Ala  
 435 440 445

Ala Thr Glu Glu Asp Ala Asp Glu Ala Asn Arg Glu Leu Met Glu Arg  
 450 455 460

Leu Asn Lys Thr Gly Lys Ala Tyr Val Ala His Thr Val Val Gly Gly  
465 470 475 480

Arg Phe Val Leu Arg Phe Ala Val Gly Ser Ser Leu Gln Glu Glu His  
485 490 495

His Val Arg Ser Ala Trp Glu Leu Ile Lys Lys Thr Thr Thr Glu Met  
500 505 510

Met Asn

<210> 52  
<211> 1494  
<212> DNA  
<213> Oryza sativa

<400> 52  
atggaagggtg ttggtggtgg tgggtggtggt gaagaatggt tgagaccaat ggatgcagaa 60  
caattaagag aatgtggtca tagaatggtt gatttcgttg cagattacta caagtcaatc 120  
gaagcttttc cagttttgtc tcaagttcaa ccaggttatt tgaaagaagt ttgcccagat 180  
tcagctccaa gacaaccaga tacttttagat tctttgttcg atgatatcca acaaaagatt 240  
atcccagggtg ttacacattg gcaatcacca aactacttcg catactaccc atctaattct 300  
tctactgctg gtttcttggg tgaaatgttg tctgctgctt ttaatatcgt tggtttttca 360  
tggatcactt ctccagctgc aacagaattg gaagttattg ttttagattg gttcgcaaag 420  
atgttgcaat tgccatcaca attccttatct actgctttgg gtggtggtgt tattcaaggt 480  
acagcttcag aagcagtttt gggtgcattg ttagctgcaa gagatagagc tttgaagaaa 540  
catggtaaac attcctttaga aaaattggtt gtttacgcat cagatcaaac tcattctgct 600  
ttacaaaaag catgtcaaat cgctgggtatt ttctctgaaa acgttagagt tgttattgct 660  
gattgtaata agaactacgc tgttgcacca gaagcagttt cagaagcttt gtctatcgat 720  
ttgtcttcag gtttgatccc atttttcatt tgtgcaacag ttggtactac atcttcatct 780  
gctgttgatc cattaccaga attgggtcaa atcgctaagt caaacgatat gtggttccat 840  
atcgatgctg cttatgctgg ttctgcatgt atttgtccag aatacagaca tcatttgaat 900  
ggtgttgaag aagcagattc ttttaatatg aatgctcata agtggttctt gactaacttc 960  
gattgttcat tgttgtgggt taaggataga tcatttttga tccaatcatt atctacaaat 1020  
ccagaattct tgaaaaataa ggcttcacaa gcaaactctg ttgttgattt caaggattgg 1080  
caaatcccat tgggtagaag attcagatca ttgaagttgt ggatggtttt gagattgtac 1140  
ggtgttgata atttgcaatc ttacatcaga aagcatatcc atttggctga acatttcgaa 1200  
caattgttgt tgtcagattc aagattcgaa gttgttactc caagaacatt ttcatgggtt 1260  
tgtttcagat tagttccacc aacttctgat catgaaaacg gtagaaagtt gaactacgat 1320  
atgatggatg gtgttaattc atctggtaaa attttcttgt cacatacagt tttgtctggt 1380

aaattcgttt tgagatttgc tgttggtgca ccattaactg aagaaagaca tgttgatgct 1440  
gcatggaagt tgttgagaga tgaagctaca aaggtttttag gtaaatgggt ttaa 1494

<210> 53  
<211> 497  
<212> PRT  
<213> Oryza sativa

<400> 53

Met Glu Gly Val Gly Gly Gly Gly Gly Gly Glu Glu Trp Leu Arg Pro  
1 5 10 15

Met Asp Ala Glu Gln Leu Arg Glu Cys Gly His Arg Met Val Asp Phe  
20 25 30

Val Ala Asp Tyr Tyr Lys Ser Ile Glu Ala Phe Pro Val Leu Ser Gln  
35 40 45

Val Gln Pro Gly Tyr Leu Lys Glu Val Leu Pro Asp Ser Ala Pro Arg  
50 55 60

Gln Pro Asp Thr Leu Asp Ser Leu Phe Asp Asp Ile Gln Gln Lys Ile  
65 70 75 80

Ile Pro Gly Val Thr His Trp Gln Ser Pro Asn Tyr Phe Ala Tyr Tyr  
85 90 95

Pro Ser Asn Ser Ser Thr Ala Gly Phe Leu Gly Glu Met Leu Ser Ala  
100 105 110

Ala Phe Asn Ile Val Gly Phe Ser Trp Ile Thr Ser Pro Ala Ala Thr  
115 120 125

Glu Leu Glu Val Ile Val Leu Asp Trp Phe Ala Lys Met Leu Gln Leu  
130 135 140

Pro Ser Gln Phe Leu Ser Thr Ala Leu Gly Gly Gly Val Ile Gln Gly  
145 150 155 160

Thr Ala Ser Glu Ala Val Leu Val Ala Leu Leu Ala Ala Arg Asp Arg  
165 170 175

Ala Leu Lys Lys His Gly Lys His Ser Leu Glu Lys Leu Val Val Tyr  
180 185 190

Ala Ser Asp Gln Thr His Ser Ala Leu Gln Lys Ala Cys Gln Ile Ala  
195 200 205

Gly Ile Phe Ser Glu Asn Val Arg Val Val Ile Ala Asp Cys Asn Lys  
210 215 220

Asn Tyr Ala Val Ala Pro Glu Ala Val Ser Glu Ala Leu Ser Ile Asp

|                                                                 |                                                 |     |  |     |  |     |
|-----------------------------------------------------------------|-------------------------------------------------|-----|--|-----|--|-----|
| 225                                                             |                                                 | 230 |  | 235 |  | 240 |
| Leu Ser Ser Gly                                                 | Leu Ile Pro Phe Phe Ile Cys Ala Thr Val Gly Thr | 245 |  | 250 |  | 255 |
| Thr Ser Ser Ser                                                 | Ala Val Asp Pro Leu Pro Glu Leu Gly Gln Ile Ala | 260 |  | 265 |  | 270 |
| Lys Ser Asn Asp Met Trp Phe His Ile Asp Ala Ala Tyr Ala Gly Ser |                                                 | 275 |  | 280 |  | 285 |
| Ala Cys Ile Cys Pro Glu Tyr Arg His His Leu Asn Gly Val Glu Glu |                                                 | 290 |  | 295 |  | 300 |
| Ala Asp Ser Phe Asn Met Asn Ala His Lys Trp Phe Leu Thr Asn Phe |                                                 | 305 |  | 310 |  | 315 |
| Asp Cys Ser Leu Leu Trp Val Lys Asp Arg Ser Phe Leu Ile Gln Ser |                                                 | 325 |  | 330 |  | 335 |
| Leu Ser Thr Asn Pro Glu Phe Leu Lys Asn Lys Ala Ser Gln Ala Asn |                                                 | 340 |  | 345 |  | 350 |
| Ser Val Val Asp Phe Lys Asp Trp Gln Ile Pro Leu Gly Arg Arg Phe |                                                 | 355 |  | 360 |  | 365 |
| Arg Ser Leu Lys Leu Trp Met Val Leu Arg Leu Tyr Gly Val Asp Asn |                                                 | 370 |  | 375 |  | 380 |
| Leu Gln Ser Tyr Ile Arg Lys His Ile His Leu Ala Glu His Phe Glu |                                                 | 385 |  | 390 |  | 395 |
| Gln Leu Leu Leu Ser Asp Ser Arg Phe Glu Val Val Thr Pro Arg Thr |                                                 | 405 |  | 410 |  | 415 |
| Phe Ser Leu Val Cys Phe Arg Leu Val Pro Pro Thr Ser Asp His Glu |                                                 | 420 |  | 425 |  | 430 |
| Asn Gly Arg Lys Leu Asn Tyr Asp Met Met Asp Gly Val Asn Ser Ser |                                                 | 435 |  | 440 |  | 445 |
| Gly Lys Ile Phe Leu Ser His Thr Val Leu Ser Gly Lys Phe Val Leu |                                                 | 450 |  | 455 |  | 460 |
| Arg Phe Ala Val Gly Ala Pro Leu Thr Glu Glu Arg His Val Asp Ala |                                                 | 465 |  | 470 |  | 475 |
| Ala Trp Lys Leu Leu Arg Asp Glu Ala Thr Lys Val Leu Gly Lys Met |                                                 | 485 |  | 490 |  | 495 |

Val

<210> 54  
<211> 1944  
<212> DNA  
<213> Xenorhabdus doucetiae

<400> 54  
atgacaaagc cattccatcc aaatttgaac gttgatgcat tgtttttagg tccaaagtct 60  
gaaaacgctg ttttcttttag agaaatgatg gattatgcag tttctgaaca catgcattgg 120  
agatcagggtt ttcattccaga agattctaca ttggttactt cagttgatag atacgaacat 180  
aactacagag aaacattgta cagaactgaa ggtattttga atcaattatc tgctaagttg 240  
aaaaattctt caatcccatt tttctctcca agatacttag gtcataattc ttcagatact 300  
ttgatggttt ctaatttggc ttatgttatg gctatgatgt acaaccctaa caactgttct 360  
tatgaagctt caccaactac aactgaattg gaattagaat ctggtttgga tttgtgtaga 420  
atgttcgggtt acaatccaca acaatcttgg ggtcatatta catcagggtg tactgttgca 480  
aattacgaag gtttgtgggt tgctagaaat ttgaagacat tgtctttcgc aatctcacia 540  
catccaaagg ctaaggattt gatcgaacat aagactccaa agcaattgaa aaatatgcca 600  
acatctgaat tgtttagattt gatcactgaa ttacaaaaga gaggtttgtt cgaagaaatc 660  
agagatatga catgtagagg tactgggtgtt aaggctcaaa ttttaggtaa attgttagtt 720  
ccacaatcta agcattattc atggatgaaa gctgcagata tcttcggtat cggtaagaa 780  
aacatcatcc cattgccagt taatgatcat taccaaacag atgttgctca aatgagaaag 840  
atcactttcg gtttgattga acaagggtgaa ccaatttttag caatgattgc tgttggttgg 900  
acaactgaag ttgggtgcaat cgatagaatc gatgaagtta ttgctttaag aaaagaatgt 960  
gaagaaagat acggtgaatc tttctacatc catgttgatg ctgcttatgc tggttacgct 1020  
tgttcaatgt tgttgaacga acagggtgaa ttcgttgaat acgatgaatt ggcaaactac 1080  
catagatctg ctgggtattat gccagaaaat atttcatggc caaaaccaga agtttaccaa 1140  
tcttttaaag ctttgaaaga tgttgattca attacagttg atccacataa ggttggtttt 1200  
attcaatatg ctgcagggtc aatttgtatg aaggataaga gaattttgga attaatcttct 1260  
tcacatgctg catacgtttt tgaaggtcat gctgataaga catctactaa aaatagaggt 1320  
attttgggtt cttcaattat ggaaggttct aaatcagggt caacagctgc agctttgtgg 1380  
gcagctcata gattgttacc attgaacatc aacgggttac gttaaagttat tgcagctgg 1440  
atcgttactg cacaatcttt gttgaataag ttggctaatt tgccaccaat cgaaatcgg 1500  
aaacatagat tcgaagttca tatttttcca tcaccagatt tccacatgat caacttcact 1560  
tttaaagaag ttggtaacga atctttgtca aaccataacg ctttgaataa gagattgtat 1620  
gaattatgtt cttactcagc aggtagagct tacacaaatg atttcttgac atcttcaact 1680  
attttagatt acaaggaata cgggtgacact ctttggcatc atgctcaaaa atgttggttt 1740  
tcaagatcag aatgggaaaa agttagaac atctatgttt tgagagcagc tgttatgaca 1800  
tactgtttga gagatgaaga acatttcgaa gaattctggg aacaattgca agctatcttc 1860

gttaagaaat tgactcaaat cgttgatgaa gaagataaga aagcaagatt gagattagaa 1920

ccagatgttc cttttattgg ttaa 1944

<210> 55

<211> 647

<212> PRT

<213> Xenorhabdus doucetiae

<400> 55

Met Thr Lys Pro Phe His Pro Asn Leu Asn Val Asp Ala Leu Phe Leu  
1 5 10 15

Gly Pro Lys Ser Glu Asn Ala Val Phe Phe Arg Glu Met Met Asp Tyr  
20 25 30

Ala Val Ser Glu His Met His Trp Arg Ser Gly Phe His Pro Glu Asp  
35 40 45

Ser Thr Leu Val Thr Ser Val Asp Arg Tyr Glu His Asn Tyr Arg Glu  
50 55 60

Thr Leu Tyr Arg Thr Glu Gly Ile Leu Asn Gln Leu Ser Ala Lys Leu  
65 70 75 80

Lys Asn Ser Ser Ile Pro Phe Phe Ser Pro Arg Tyr Leu Gly His Ile  
85 90 95

Ser Ser Asp Thr Leu Met Val Ser Asn Leu Ala Tyr Val Met Ala Met  
100 105 110

Met Tyr Asn Pro Asn Asn Cys Ser Tyr Glu Ala Ser Pro Thr Thr Thr  
115 120 125

Glu Leu Glu Leu Glu Ser Gly Leu Asp Leu Cys Arg Met Phe Gly Tyr  
130 135 140

Asn Pro Gln Gln Ser Trp Gly His Ile Thr Ser Gly Gly Thr Val Ala  
145 150 155 160

Asn Tyr Glu Gly Leu Trp Val Ala Arg Asn Leu Lys Thr Leu Ser Phe  
165 170 175

Ala Ile Ser Gln His Pro Lys Ala Lys Asp Leu Ile Glu His Lys Thr  
180 185 190

Pro Lys Gln Leu Lys Asn Met Pro Thr Ser Glu Leu Leu Asp Leu Ile  
195 200 205

Thr Glu Leu Gln Lys Arg Gly Leu Phe Glu Glu Ile Arg Asp Met Thr  
210 215 220

Cys Arg Gly Thr Gly Val Lys Ala Gln Ile Leu Gly Lys Leu Leu Val  
225 230 235 240

Pro Gln Ser Lys His Tyr Ser Trp Met Lys Ala Ala Asp Ile Phe Gly  
245 250 255

Ile Gly Gln Glu Asn Ile Ile Pro Leu Pro Val Asn Asp His Tyr Gln  
260 265 270

Thr Asp Val Ala Gln Met Arg Lys Ile Thr Phe Gly Leu Ile Glu Gln  
275 280 285

Gly Glu Pro Ile Leu Ala Met Ile Ala Val Val Gly Thr Thr Glu Val  
290 295 300

Gly Ala Ile Asp Arg Ile Asp Glu Val Ile Ala Leu Arg Lys Glu Cys  
305 310 315 320

Glu Glu Arg Tyr Gly Glu Ser Phe Tyr Ile His Val Asp Ala Ala Tyr  
325 330 335

Ala Gly Tyr Ala Cys Ser Met Leu Leu Asn Glu Gln Gly Glu Phe Val  
340 345 350

Glu Tyr Asp Glu Leu Ala Asn Tyr His Arg Ser Ala Gly Ile Met Pro  
355 360 365

Glu Asn Ile Ser Trp Pro Lys Pro Glu Val Tyr Gln Ser Phe Lys Ala  
370 375 380

Leu Lys Asp Val Asp Ser Ile Thr Val Asp Pro His Lys Val Gly Phe  
385 390 395 400

Ile Gln Tyr Ala Ala Gly Ala Ile Cys Met Lys Asp Lys Arg Ile Leu  
405 410 415

Glu Leu Ile Ser Ser His Ala Ala Tyr Val Phe Glu Gly His Ala Asp  
420 425 430

Lys Thr Ser Thr Lys Asn Arg Gly Ile Leu Gly Ser Ser Ile Met Glu  
435 440 445

Gly Ser Lys Ser Gly Ala Thr Ala Ala Ala Leu Trp Ala Ala His Arg  
450 455 460

Leu Leu Pro Leu Asn Ile Asn Gly Tyr Gly Lys Val Ile Ala Ala Gly  
465 470 475 480

Ile Val Thr Ala Gln Ser Leu Leu Asn Lys Leu Ala Asn Leu Pro Pro  
485 490 495

Ile Glu Ile Gly Lys His Arg Phe Glu Val His Ile Phe Pro Ser Pro



500

505

510

Asp Phe His Met Ile Asn Phe Thr Phe Lys Glu Val Gly Asn Glu Ser  
515 520 525

Leu Ser Asn His Asn Ala Leu Asn Lys Arg Leu Tyr Glu Leu Cys Ser  
530 535 540

Tyr Ser Ala Gly Arg Ala Tyr Thr Asn Asp Phe Leu Thr Ser Ser Thr  
545 550 555 560

Ile Leu Asp Tyr Lys Glu Tyr Gly Asp Thr Pro Trp His His Ala Gln  
565 570 575

Lys Cys Gly Phe Ser Arg Ser Glu Trp Glu Lys Val Arg Asn Ile Tyr  
580 585 590

Val Leu Arg Ala Ala Val Met Thr Tyr Cys Leu Arg Asp Glu Glu His  
595 600 605

Phe Glu Glu Phe Trp Glu Gln Leu Gln Ala Ile Phe Val Lys Lys Leu  
610 615 620

Thr Gln Ile Val Asp Glu Glu Asp Lys Lys Ala Arg Leu Arg Leu Glu  
625 630 635 640

Pro Asp Val Pro Phe Ile Gly  
645

<210> 56  
<211> 1509  
<212> DNA  
<213> *Camptotheca acuminata*

<400> 56  
atgggttctt tggattcaaa ctacgatact gaatctccag cttcagttgg tcaattcaat 60  
ccattagatc cagaagaatt cagaaagcaa gctcattgta tcgttgattt catcgagat 120  
tactacaaga acattgaatc ttatccagtt ttgtcacaag ttgatccagg ttacagacat 180  
tcaagattgg gtaaaaatgc tccatacaga tctgaaccat tcgaatcaat cttgaaggat 240  
gttcaaaagg atatcatccc tggatatgact cattggatgt ctccaaattt ctttgctcat 300  
tttccagcaa ctgtttcttc agctgcattt gttggtgaaa tgttgtgtac atgtttcaac 360  
tctgttggtt ttaattgggt agcatcacca gctgcaacag aattggaaat ggttggttatt 420  
gattggttag ctaacatggt gaagttgcc aagtctttta tgttttcagg tactggtggt 480  
ggtgttttgc aagggtactac atctgaagct atcttgtgta cattgattgc tgcacacca 540  
atgcatttcg aaatcggttg tgtaagact tctacatcat tcgttggtta tggttctgat 600  
caaaactcatt caacatacgc taaagcatgt aaattagcag gtattttgcc atgtaacatc 660  
agatctattc caactacgc tgattcaaac ttctctgttt caccattggt gttgagaaga 720



|                                                                 |  |     |  |     |     |
|-----------------------------------------------------------------|--|-----|--|-----|-----|
| 130                                                             |  | 135 |  | 140 |     |
| Asn Met Leu Lys Leu Pro Lys Ser Phe Met Phe Ser Gly Thr Gly Gly |  |     |  |     |     |
| 145                                                             |  | 150 |  | 155 | 160 |
| Gly Val Leu Gln Gly Thr Thr Ser Glu Ala Ile Leu Cys Thr Leu Ile |  |     |  |     |     |
|                                                                 |  | 165 |  | 170 | 175 |
| Ala Ala Ser Pro Met His Phe Glu Ile Val Gly Val Lys Thr Ser Thr |  |     |  |     |     |
|                                                                 |  | 180 |  | 185 | 190 |
| Ser Phe Val Val Tyr Gly Ser Asp Gln Thr His Ser Thr Tyr Ala Lys |  |     |  |     |     |
|                                                                 |  | 195 |  | 200 | 205 |
| Ala Cys Lys Leu Ala Gly Ile Leu Pro Cys Asn Ile Arg Ser Ile Pro |  |     |  |     |     |
|                                                                 |  | 210 |  | 215 | 220 |
| Thr Thr Ala Asp Ser Asn Phe Ser Val Ser Pro Leu Leu Leu Arg Arg |  |     |  |     |     |
| 225                                                             |  | 230 |  | 235 | 240 |
| Ala Ile Glu Ala Asp Lys Ala Ala Gly Met Val Pro Leu Tyr Ile Cys |  |     |  |     |     |
|                                                                 |  | 245 |  | 250 | 255 |
| Ala Thr Val Gly Thr Thr Ser Thr Thr Ala Ile Asp Pro Leu Ser Ser |  |     |  |     |     |
|                                                                 |  | 260 |  | 265 | 270 |
| Leu Ala Asp Val Ala Asn Asp Tyr Gly Val Trp Phe His Val Asp Ala |  |     |  |     |     |
|                                                                 |  | 275 |  | 280 | 285 |
| Ala Tyr Ala Gly Ser Ala Cys Ile Cys Pro Glu Phe Arg His Tyr Leu |  |     |  |     |     |
|                                                                 |  | 290 |  | 295 | 300 |
| Asp Gly Ile Glu Arg Ala Asp Ser Leu Ser Leu Ser Pro His Lys Trp |  |     |  |     |     |
| 305                                                             |  | 310 |  | 315 | 320 |
| Leu Leu Ser Tyr Leu Asp Cys Cys Cys Leu Trp Val Lys Ser Pro Ser |  |     |  |     |     |
|                                                                 |  | 325 |  | 330 | 335 |
| Leu Leu Val Lys Ala Leu Ser Thr Asp Pro Glu Tyr Leu Lys Asn Gln |  |     |  |     |     |
|                                                                 |  | 340 |  | 345 | 350 |
| Pro Ser Glu Ser Lys Ser Val Val Asp Tyr Lys Asp Trp Gln Val Gly |  |     |  |     |     |
|                                                                 |  | 355 |  | 360 | 365 |
| Thr Gly Arg Arg Phe Lys Ala Leu Arg Leu Trp Phe Val Met Arg Ser |  |     |  |     |     |
|                                                                 |  | 370 |  | 375 | 380 |
| Tyr Gly Val Ala Asn Leu Gln Ser His Ile Arg Thr Asp Val Gln Met |  |     |  |     |     |
| 385                                                             |  | 390 |  | 395 | 400 |
| Ala Lys Met Phe Glu Gly Phe Val Lys Ser Asp Pro Arg Phe Glu Ile |  |     |  |     |     |
|                                                                 |  | 405 |  | 410 | 415 |

Leu Val Pro Arg Val Phe Ser Leu Val Cys Phe Arg Leu Asn Pro Ile  
420 425 430

Ser Gly Ser Asp Pro Thr Gly Thr Glu Ala Leu Asn Arg Lys Leu Leu  
435 440 445

Asp Trp Val Asn Ser Thr Gly Arg Val Tyr Met Thr His Thr Lys Val  
450 455 460

Gly Gly Ile Tyr Met Leu Arg Phe Ala Val Gly Ala Thr Leu Thr Glu  
465 470 475 480

Lys Arg His Val Ser Ser Ala Trp Lys Leu Ile Lys Glu Gly Ala Asp  
485 490 495

Val Leu Leu Lys Glu Asp  
500

<210> 58  
<211> 792  
<212> DNA  
<213> *Oryctolagus cuniculus*

<400> 58  
atggaagggtg gttttactgg tggtagcgaa taccaaaagc atttcttgcc aagagattac 60  
ttaaacacat actactcatt ccaatctggt ccatcaccag aagctgaaat gttgaagttt 120  
aatttggaat gtttgcataa gactttcggg ccagggtggtt tgcaagggtga cactttaatt 180  
gatattgggtt ctgggtccaac aatctatcaa gttttggctg catgtgaatc ttttaaagat 240  
atcactttgt cagattttac agatagaaat agagaagaat tggctaaatg gttgaagaaa 300  
gaaccagggtg catacgattg gactccagct ttgaagttcg catgtgaatt agaaggcaac 360  
tctggtagat ggcaagaaaa agctgaaaag ttgagagcaa cagttaagag agttttgaag 420  
tgtgatgcta atttgtcaaa tccattaact ccagttgttt tgccaccagc agattgtgtt 480  
ttgacattgt tagctatgga atgtgcatgt tgttctttgg atgcttatag agctgcattg 540  
agaaatttgg catcattgtt aaaaccagggt ggatcatttgg ttactacagt tactttgcaa 600  
ttatcttcat acatggtttg tgaaagagaa ttttcttgtg ttgctttgga aaaagaagaa 660  
gttgaacaag ctgttttaga tgcaggtttc gatatcgaa aattgttgta ctctccacaa 720  
tcttactcag cttctacagc accaaataga ggtgtttgtt tcttggttgc tagaaagaaa 780  
ccaggttcat aa 792

<210> 59  
<211> 263  
<212> PRT  
<213> *Oryctolagus cuniculus*

<400> 59

Met Glu Gly Gly Phe Thr Gly Gly Asp Glu Tyr Gln Lys His Phe Leu

|                     |                     |                     |             |
|---------------------|---------------------|---------------------|-------------|
| 1                   | 5                   | 10                  | 15          |
| Pro Arg Asp Tyr     | Leu Asn Thr Tyr     | Tyr Ser Phe Gln Ser | Gly Pro Ser |
| 20                  | 25                  | 30                  |             |
| Pro Glu Ala Glu Met | Leu Lys Phe Asn     | Leu Glu Cys Leu His | Lys Thr     |
| 35                  | 40                  | 45                  |             |
| Phe Gly Pro Gly Gly | Leu Gln Gly Asp Thr | Leu Ile Asp Ile     | Gly Ser     |
| 50                  | 55                  | 60                  |             |
| Gly Pro Thr Ile Tyr | Gln Val Leu Ala Ala | Cys Glu Ser Phe Lys | Asp         |
| 65                  | 70                  | 75                  | 80          |
| Ile Thr Leu Ser Asp | Phe Thr Asp Arg Asn | Arg Glu Glu Leu Ala | Lys         |
| 85                  | 90                  | 95                  |             |
| Trp Leu Lys Lys Glu | Pro Gly Ala Tyr Asp | Trp Thr Pro Ala     | Leu Lys     |
| 100                 | 105                 | 110                 |             |
| Phe Ala Cys Glu Leu | Glu Gly Asn Ser Gly | Arg Trp Gln Glu Lys | Ala         |
| 115                 | 120                 | 125                 |             |
| Glu Lys Leu Arg Ala | Thr Val Lys Arg Val | Leu Lys Cys Asp Ala | Asn         |
| 130                 | 135                 | 140                 |             |
| Leu Ser Asn Pro Leu | Thr Pro Val Val Leu | Pro Pro Ala Asp Cys | Val         |
| 145                 | 150                 | 155                 | 160         |
| Leu Thr Leu Leu Ala | Met Glu Cys Ala Cys | Cys Ser Leu Asp Ala | Tyr         |
| 165                 | 170                 | 175                 |             |
| Arg Ala Ala Leu Arg | Asn Leu Ala Ser Leu | Leu Lys Pro Gly Gly | His         |
| 180                 | 185                 | 190                 |             |
| Leu Val Thr Thr Val | Thr Leu Gln Leu Ser | Ser Tyr Met Val Gly | Glu         |
| 195                 | 200                 | 205                 |             |
| Arg Glu Phe Ser Cys | Val Ala Leu Glu Lys | Glu Glu Val Glu Gln | Ala         |
| 210                 | 215                 | 220                 |             |
| Val Leu Asp Ala Gly | Phe Asp Ile Glu Gln | Leu Leu Tyr Ser Pro | Gln         |
| 225                 | 230                 | 235                 | 240         |
| Ser Tyr Ser Ala Ser | Thr Ala Pro Asn Arg | Gly Val Cys Phe Leu | Val         |
| 245                 | 250                 | 255                 |             |
| Ala Arg Lys Lys Pro | Gly Ser             |                     |             |
| 260                 |                     |                     |             |

<210> 60  
 <211> 1131

<212> DNA  
<213> Hordeum vulgare

<400> 60  
atggataaga tctctgcacc atttttctct ggtacttcac cagctgcagc ttcagttgct 60  
ggtgttgatg aagatgatag attgtgtttc caagctcaag aattgatgtt cgcatacaac 120  
atctctatgg ttttgagagc agctattcaa ttaggtttgt tagatgcttt gtcagcagct 180  
ggtggtaaag cattgactcc aaacgaatta gttgaaaacg ttgaaacatc ttcaaataag 240  
gctgaagcag ctgcagctgt tgatagaatt ttgagatatt tgtcttgttt caacgttggt 300  
acttgttctt cagaagcagc tgggccagat ggtacattag ttagaagata cactacaggt 360  
ccattgtgta gatggttaac taaagataga ggtgacggta cattatctcc atttgctggt 420  
tttgttgttg atccagatca tttgtttcct tggcatcata ttgcagaagc tgttactgct 480  
ggtgggtccat cagcattcga aagaacacaa aagtggccat actacgaata catgggtaaa 540  
aatcaaagat tgggtacttt gttcgataac gcaatggctc aacattctgt tatttttggt 600  
acaaagatgt tagaaagatt caaaggtttt gatggtgttc aaagattagt tgatgttggt 660  
ggtggtactg gttctacatt gggatatgac acttcaaagt acaagcacat gacaggtatt 720  
aattacgatt tgccacatgt tattgctcaa ggtttgccat taccaggtgt tgaacatgtt 780  
gcaggtgaca tgtacgaatc tattccaact ggtgacgctg ttttgttaca atggattaca 840  
ttgatgttga acgatgatga attcgttaag atcttgtcaa attgtcataa tgctttgcca 900  
aaagatggta aagttattgt tgttgatggt attttaccag aaaaccaga ttcttcattg 960  
actgcaagag atgctttttac attagatatc atcatgttcg ttttgtttaa aggtgcaaag 1020  
caaagaactg aaaaggaatt cgcaagatta gctaagcaag caggttttac tgggtggtatt 1080  
aagaaaactt atattttctt taatttttac gctttggaat tcactaaata a 1131

<210> 61  
<211> 376  
<212> PRT  
<213> Hordeum vulgare

<400> 61  
Met Asp Lys Ile Ser Ala Pro Phe Phe Ser Gly Thr Ser Pro Ala Ala  
1 5 10 15  
Ala Ser Val Ala Gly Val Asp Glu Asp Asp Arg Leu Cys Phe Gln Ala  
20 25 30  
Gln Glu Leu Met Phe Ala Tyr Asn Ile Ser Met Val Leu Arg Ala Ala  
35 40 45  
Ile Gln Leu Gly Leu Leu Asp Ala Leu Ser Ala Ala Gly Gly Lys Ala  
50 55 60  
Leu Thr Pro Asn Glu Leu Val Glu Asn Val Glu Thr Ser Ser Asn Lys  
65 70 75 80

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|
| Ala | Glu | Ala | Ala | Ala | Ala | Val | Asp | Arg | Ile | Leu | Arg | Tyr | Leu | Ser | Cys |  |  |
|     |     |     |     | 85  |     |     |     |     | 90  |     |     |     |     | 95  |     |  |  |
| Phe | Asn | Val | Val | Thr | Cys | Ser | Ser | Glu | Ala | Ala | Gly | Pro | Asp | Gly | Thr |  |  |
|     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |  |  |
| Leu | Val | Arg | Arg | Tyr | Thr | Thr | Gly | Pro | Leu | Cys | Arg | Trp | Leu | Thr | Lys |  |  |
|     |     | 115 |     |     |     |     | 120 |     |     |     |     | 125 |     |     |     |  |  |
| Asp | Arg | Gly | Asp | Gly | Thr | Leu | Ser | Pro | Phe | Ala | Val | Phe | Val | Val | Asp |  |  |
|     | 130 |     |     |     |     | 135 |     |     |     |     | 140 |     |     |     |     |  |  |
| Pro | Asp | His | Leu | Phe | Pro | Trp | His | His | Ile | Ala | Glu | Ala | Val | Thr | Ala |  |  |
| 145 |     |     |     |     | 150 |     |     |     |     | 155 |     |     |     |     | 160 |  |  |
| Gly | Gly | Pro | Ser | Ala | Phe | Glu | Arg | Thr | Gln | Lys | Trp | Pro | Tyr | Tyr | Glu |  |  |
|     |     |     |     | 165 |     |     |     |     | 170 |     |     |     |     | 175 |     |  |  |
| Tyr | Met | Gly | Lys | Asn | Gln | Arg | Leu | Gly | Thr | Leu | Phe | Asp | Asn | Ala | Met |  |  |
|     |     |     | 180 |     |     |     |     | 185 |     |     |     |     | 190 |     |     |  |  |
| Ala | Gln | His | Ser | Val | Ile | Leu | Val | Thr | Lys | Met | Leu | Glu | Arg | Phe | Lys |  |  |
|     |     | 195 |     |     |     |     | 200 |     |     |     |     | 205 |     |     |     |  |  |
| Gly | Phe | Asp | Gly | Val | Gln | Arg | Leu | Val | Asp | Val | Gly | Gly | Gly | Thr | Gly |  |  |
|     | 210 |     |     |     |     | 215 |     |     |     |     | 220 |     |     |     |     |  |  |
| Ser | Thr | Leu | Gly | Met | Ile | Thr | Ser | Lys | Tyr | Lys | His | Met | Thr | Gly | Ile |  |  |
| 225 |     |     |     |     | 230 |     |     |     |     | 235 |     |     |     |     | 240 |  |  |
| Asn | Tyr | Asp | Leu | Pro | His | Val | Ile | Ala | Gln | Gly | Leu | Pro | Leu | Pro | Gly |  |  |
|     |     |     | 245 |     |     |     |     |     | 250 |     |     |     |     | 255 |     |  |  |
| Val | Glu | His | Val | Ala | Gly | Asp | Met | Tyr | Glu | Ser | Ile | Pro | Thr | Gly | Asp |  |  |
|     |     |     | 260 |     |     |     |     | 265 |     |     |     |     | 270 |     |     |  |  |
| Ala | Val | Leu | Leu | Gln | Trp | Ile | Thr | Leu | Met | Leu | Asn | Asp | Asp | Glu | Phe |  |  |
|     |     | 275 |     |     |     |     | 280 |     |     |     |     | 285 |     |     |     |  |  |
| Val | Lys | Ile | Leu | Ser | Asn | Cys | His | Asn | Ala | Leu | Pro | Lys | Asp | Gly | Lys |  |  |
|     | 290 |     |     |     |     | 295 |     |     |     |     | 300 |     |     |     |     |  |  |
| Val | Ile | Val | Val | Asp | Gly | Ile | Leu | Pro | Glu | Asn | Pro | Asp | Ser | Ser | Leu |  |  |
| 305 |     |     |     |     | 310 |     |     |     |     | 315 |     |     |     |     | 320 |  |  |
| Thr | Ala | Arg | Asp | Ala | Phe | Thr | Leu | Asp | Ile | Ile | Met | Phe | Val | Leu | Phe |  |  |
|     |     |     |     | 325 |     |     |     |     | 330 |     |     |     |     | 335 |     |  |  |
| Lys | Gly | Ala | Lys | Gln | Arg | Thr | Glu | Lys | Glu | Phe | Ala | Arg | Leu | Ala | Lys |  |  |
|     |     |     | 340 |     |     |     |     | 345 |     |     |     |     | 350 |     |     |  |  |

Gln Ala Gly Phe Thr Gly Gly Ile Lys Lys Thr Tyr Ile Phe Phe Asn  
355 360 365

Phe Tyr Ala Leu Glu Phe Thr Lys  
370 375

<210> 62  
<211> 549  
<212> DNA  
<213> Streptomyces griseofuscus

<400> 62  
atgaacacct tcagaacagc cactgccaga gacataacctg atgtagcagc aactcttacg 60  
gaagccttcg caactgatcc acccacgcag tgggtgttcc ccgacggtac tgccgccgtc 120  
agcaggttct ttacacatgt tgcagatagg gttcacacgg ccggtggtat tgttgagcta 180  
ctaccagaca gagccgccat gattgcattg ccaccacacg tgaggctgcc aggagaagct 240  
gccgacggaa ggcaggcgga aattcagaga aggctggcag acaggcaccc gctgacacct 300  
cactactacc tgctgtttta cggagttaga acggcacacc agggttcggg attgggcgga 360  
agaatgctgg ccagattaac tagcagagct gatagggaca ggggtgggtac atatactgag 420  
gcatccacct ggcgtggcgc tagactgatg ctgagacatg gattccatgc tacaaggcca 480  
ctaagattgc cagatggacc cagcatgttt ccactttgga gagatccaat ccatgatcat 540  
tctgattag 549

<210> 63  
<211> 182  
<212> PRT  
<213> Streptomyces griseofuscus

<400> 63

Met Asn Thr Phe Arg Thr Ala Thr Ala Arg Asp Ile Pro Asp Val Ala  
1 5 10 15

Ala Thr Leu Thr Glu Ala Phe Ala Thr Asp Pro Pro Thr Gln Trp Val  
20 25 30

Phe Pro Asp Gly Thr Ala Ala Val Ser Arg Phe Phe Thr His Val Ala  
35 40 45

Asp Arg Val His Thr Ala Gly Gly Ile Val Glu Leu Leu Pro Asp Arg  
50 55 60

Ala Ala Met Ile Ala Leu Pro Pro His Val Arg Leu Pro Gly Glu Ala  
65 70 75 80

Ala Asp Gly Arg Gln Ala Glu Ile Gln Arg Arg Leu Ala Asp Arg His  
85 90 95

Pro Leu Thr Pro His Tyr Tyr Leu Leu Phe Tyr Gly Val Arg Thr Ala  
100 105 110



His Gln Gly Ser Gly Leu Gly Gly Arg Met Leu Ala Arg Leu Thr Ser  
 115 120 125

Arg Ala Asp Arg Asp Arg Val Gly Thr Tyr Thr Glu Ala Ser Thr Trp  
 130 135 140

Arg Gly Ala Arg Leu Met Leu Arg His Gly Phe His Ala Thr Arg Pro  
 145 150 155 160

Leu Arg Leu Pro Asp Gly Pro Ser Met Phe Pro Leu Trp Arg Asp Pro  
 165 170 175

Ile His Asp His Ser Asp  
 180

<210> 64  
 <211> 1170  
 <212> DNA  
 <213> Thermotoga maritima

<400> 64  
 atgaaaggat atttcggacc atacggtggc cagtacgtac cagaaatatt aatgggtgcc 60  
 ttagaggagt tagaggcagc atacgaggag attatgaagg atgagagctt ctggaaggag 120  
 ttcaacgata tactgaggga ttacgcaggc agaccaacgc cattgtactt tgccaggaga 180  
 ttgtctgaga agtacggcgc ccgtgtttac ttgaagcgtg aggatctgct gcacactgga 240  
 gcacacaaga taaataacgc tatcggacag gttttattgg ccaaattaat ggggaagaca 300  
 cgtatcatag ccgagacggg agctgggcag catggagtcg ctactgctac cgctgctgcc 360  
 ctgttcggaa tggaatgtgt gatctacatg ggtgaagagg acacaatcag acagaagttg 420  
 aacgtggagc gtatgaaatt attaggggct aaagtgtgcc ctgttaagtc tggcagtagg 480  
 acctgaagg atgcgataga cgaggctttg agagactgga ttactaattt acagacaaca 540  
 tattatgtta tcggatctgt tgttgggtccc cacccttacc caattatcgt aaggaatttc 600  
 cagaaggtta tcggtgagga gaccaagaag caaataccag aaaaggaagg tcgtttgcca 660  
 gactatatag ttgcctgcgt aggcggcggg agcaatgccg caggatatatt ttaccattc 720  
 atagactctg gagtaaagct gataggtgtt gaggcagggt gcgagggatt ggagacaggt 780  
 aaacacgcag cctcgttatt aaagggtaaa attggctatt tacatggatc gaagacctt 840  
 gttctacaag atgactgggg tcaagtccaa gtgagccatt cgggtgtcagc tggctctgac 900  
 tattcaggag taggacctga gcatgcttat tggagagaga cagggaaggt tctgtacgac 960  
 gcagtgactg acgaagaggc tttggacgca tttatagagt tatcaagact agagggcatt 1020  
 ataccgctt tagagtcata gcatgctcta gcatatttga agaagataaa tataaaaggt 1080  
 aaggttgttg tggtaacct atcagggaga ggggataaag acctggagtc agtcttaaac 1140  
 catccatacg tgagagaaag aattagatga 1170

<210> 65

<211> 389  
<212> PRT  
<213> Thermotoga maritima

<400> 65

Met Lys Gly Tyr Phe Gly Pro Tyr Gly Gly Gln Tyr Val Pro Glu Ile  
1 5 10 15

Leu Met Gly Ala Leu Glu Glu Leu Glu Ala Ala Tyr Glu Glu Ile Met  
20 25 30

Lys Asp Glu Ser Phe Trp Lys Glu Phe Asn Asp Leu Leu Arg Asp Tyr  
35 40 45

Ala Gly Arg Pro Thr Pro Leu Tyr Phe Ala Arg Arg Leu Ser Glu Lys  
50 55 60

Tyr Gly Ala Arg Val Tyr Leu Lys Arg Glu Asp Leu Leu His Thr Gly  
65 70 75 80

Ala His Lys Ile Asn Asn Ala Ile Gly Gln Val Leu Leu Ala Lys Leu  
85 90 95

Met Gly Lys Thr Arg Ile Ile Ala Glu Thr Gly Ala Gly Gln His Gly  
100 105 110

Val Ala Thr Ala Thr Ala Ala Ala Leu Phe Gly Met Glu Cys Val Ile  
115 120 125

Tyr Met Gly Glu Glu Asp Thr Ile Arg Gln Lys Leu Asn Val Glu Arg  
130 135 140

Met Lys Leu Leu Gly Ala Lys Val Val Pro Val Lys Ser Gly Ser Arg  
145 150 155 160

Thr Leu Lys Asp Ala Ile Asp Glu Ala Leu Arg Asp Trp Ile Thr Asn  
165 170 175

Leu Gln Thr Thr Tyr Tyr Val Ile Gly Ser Val Val Gly Pro His Pro  
180 185 190

Tyr Pro Ile Ile Val Arg Asn Phe Gln Lys Val Ile Gly Glu Glu Thr  
195 200 205

Lys Lys Gln Ile Pro Glu Lys Glu Gly Arg Leu Pro Asp Tyr Ile Val  
210 215 220

Ala Cys Val Gly Gly Gly Ser Asn Ala Ala Gly Ile Phe Tyr Pro Phe  
225 230 235 240

Ile Asp Ser Gly Val Lys Leu Ile Gly Val Glu Ala Gly Gly Glu Gly  
245 250 255

Leu Glu Thr Gly Lys His Ala Ala Ser Leu Leu Lys Gly Lys Ile Gly  
260 265 270

Tyr Leu His Gly Ser Lys Thr Phe Val Leu Gln Asp Asp Trp Gly Gln  
275 280 285

Val Gln Val Ser His Ser Val Ser Ala Gly Leu Asp Tyr Ser Gly Val  
290 295 300

Gly Pro Glu His Ala Tyr Trp Arg Glu Thr Gly Lys Val Leu Tyr Asp  
305 310 315 320

Ala Val Thr Asp Glu Glu Ala Leu Asp Ala Phe Ile Glu Leu Ser Arg  
325 330 335

Leu Glu Gly Ile Ile Pro Ala Leu Glu Ser Ser His Ala Leu Ala Tyr  
340 345 350

Leu Lys Lys Ile Asn Ile Lys Gly Lys Val Val Val Val Asn Leu Ser  
355 360 365

Gly Arg Gly Asp Lys Asp Leu Glu Ser Val Leu Asn His Pro Tyr Val  
370 375 380

Arg Glu Arg Ile Arg  
385

<210> 66  
<211> 1167  
<212> DNA  
<213> *Pyrococcus furiosus*

<400> 66  
atgtggttcg gtgagtttgg tggacaatat gtgccagaga ctttagtggg tcctcttaag 60  
gaattggaaa aggcatataa aaggttcaag gacgatgagg agttcaacag gcaactaaac 120  
tattatttga agacatgggc cggtagacca acgcccttgt attatgctaa gaggttaact 180  
gaaaagattg gcggcgcgaa agtgtatctg aaaagagaag acctagttca tgggtggagca 240  
cacaagacaa ataatgccat tggacaagca ctattggcaa agctaattggg taaaactaga 300  
ttgatagctg agacaggagc ggggtcaacat ggggtcgcga cagcgatggc tgggtgcacta 360  
ctggggatga aggtagatat ttacatgggt gctgaggacg ttgagcgtca gaaactaaat 420  
gtcttcagga tgaagctatt aggtgccaat gttataacctg taaattctgg ctcaagaaca 480  
ctaaaggacg ccttcgacga ggctcttaga gactgggttg ccactttcga gtatactcat 540  
tacttgatcg gttcagtggg tggaccacat ccatacccaa ccatcgttag ggactttcag 600  
agcgtgattg gtagagaggc taaggcacag atcttagaag cagagggaca gctacctgac 660  
gtcatagtgg cctgcgtcgg cgggtggctct aacgcaatgg gtatattcta tccattcgtt 720  
aatgacaaga aggttaaatt agtaggagtc gaagctggcg gaaaggggtt agagtcgggt 780

|                                                                   |      |
|-------------------------------------------------------------------|------|
| aaacactcag caagcttaaa tgcaggacag gtaggggtgt cccacggcat gttgtcgtat | 840  |
| ttcttgaag acgaggaagg tcagataaag ccaagtcatt caattgctcc aggccttgac  | 900  |
| caccccggtg ttggtccaga gcacgcttac ttaaagaaga ttcaaagggc cgagtacgtc | 960  |
| gctgtaacag acgaagaggc attgaaagct ttccatgagc tatccagaac tgaggggatt | 1020 |
| ataccgccc ttgagtctgc ccatgctgtg gcgtacgcca tgaagttagc taaagagatg  | 1080 |
| tcccgtagcg aaatcatcat tgtaaactta tcagggagag gagacaagga tttggacatt | 1140 |
| gtattgaagg caagcggaaa tgtttga                                     | 1167 |

<210> 67  
 <211> 388  
 <212> PRT  
 <213> *Pyrococcus furiosus*

<400> 67

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Trp | Phe | Gly | Glu | Phe | Gly | Gly | Gln | Tyr | Val | Pro | Glu | Thr | Leu | Val |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gly | Pro | Leu | Lys | Glu | Leu | Glu | Lys | Ala | Tyr | Lys | Arg | Phe | Lys | Asp | Asp |
|     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Glu | Glu | Phe | Asn | Arg | Gln | Leu | Asn | Tyr | Tyr | Leu | Lys | Thr | Trp | Ala | Gly |
|     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Arg | Pro | Thr | Pro | Leu | Tyr | Tyr | Ala | Lys | Arg | Leu | Thr | Glu | Lys | Ile | Gly |
| 50  |     |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gly | Ala | Lys | Val | Tyr | Leu | Lys | Arg | Glu | Asp | Leu | Val | His | Gly | Gly | Ala |
| 65  |     |     |     |     | 70  |     |     |     | 75  |     |     |     |     |     | 80  |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| His | Lys | Thr | Asn | Asn | Ala | Ile | Gly | Gln | Ala | Leu | Leu | Ala | Lys | Leu | Met |
|     |     |     | 85  |     |     |     |     |     | 90  |     |     |     |     | 95  |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gly | Lys | Thr | Arg | Leu | Ile | Ala | Glu | Thr | Gly | Ala | Gly | Gln | His | Gly | Val |
|     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | Thr | Ala | Met | Ala | Gly | Ala | Leu | Leu | Gly | Met | Lys | Val | Asp | Ile | Tyr |
|     |     | 115 |     |     |     |     | 120 |     |     |     |     | 125 |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Gly | Ala | Glu | Asp | Val | Glu | Arg | Gln | Lys | Leu | Asn | Val | Phe | Arg | Met |
| 130 |     |     |     |     |     | 135 |     |     |     |     | 140 |     |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Lys | Leu | Leu | Gly | Ala | Asn | Val | Ile | Pro | Val | Asn | Ser | Gly | Ser | Arg | Thr |
| 145 |     |     |     |     | 150 |     |     |     |     | 155 |     |     |     |     | 160 |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Lys | Asp | Ala | Phe | Asp | Glu | Ala | Leu | Arg | Asp | Trp | Val | Ala | Thr | Phe |
|     |     |     | 165 |     |     |     |     |     | 170 |     |     |     |     | 175 |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Glu | Tyr | Thr | His | Tyr | Leu | Ile | Gly | Ser | Val | Val | Gly | Pro | His | Pro | Tyr |
|     |     |     | 180 |     |     |     |     | 185 |     |     |     |     | 190 |     |     |

Pro Thr Ile Val Arg Asp Phe Gln Ser Val Ile Gly Arg Glu Ala Lys  
195 200 205

Ala Gln Ile Leu Glu Ala Glu Gly Gln Leu Pro Asp Val Ile Val Ala  
210 215 220

Cys Val Gly Gly Gly Ser Asn Ala Met Gly Ile Phe Tyr Pro Phe Val  
225 230 235 240

Asn Asp Lys Lys Val Lys Leu Val Gly Val Glu Ala Gly Gly Lys Gly  
245 250 255

Leu Glu Ser Gly Lys His Ser Ala Ser Leu Asn Ala Gly Gln Val Gly  
260 265 270

Val Ser His Gly Met Leu Ser Tyr Phe Leu Gln Asp Glu Glu Gly Gln  
275 280 285

Ile Lys Pro Ser His Ser Ile Ala Pro Gly Leu Asp His Pro Gly Val  
290 295 300

Gly Pro Glu His Ala Tyr Leu Lys Lys Ile Gln Arg Ala Glu Tyr Val  
305 310 315 320

Ala Val Thr Asp Glu Glu Ala Leu Lys Ala Phe His Glu Leu Ser Arg  
325 330 335

Thr Glu Gly Ile Ile Pro Ala Leu Glu Ser Ala His Ala Val Ala Tyr  
340 345 350

Ala Met Lys Leu Ala Lys Glu Met Ser Arg Asp Glu Ile Ile Ile Val  
355 360 365

Asn Leu Ser Gly Arg Gly Asp Lys Asp Leu Asp Ile Val Leu Lys Ala  
370 375 380

Ser Gly Asn Val  
385

|            |                     |            |
|------------|---------------------|------------|
| <210>      | 68                  |            |
| <211>      | 1167                |            |
| <212>      | DNA                 |            |
| <213>      | Pyrococcus furiosus |            |
| <400>      | 68                  |            |
| atgtggtttg | gggaattcgg          | tggtcaatac |
| gtgccagaaa | ccttagttgg          | gcctttaaag |
|            |                     | 60         |
| gagcttgaga | aagcttaca           | aagatttaag |
| gacgacgagg | agtttaatag          | gcaattaaac |
|            |                     | 120        |
| tattacttga | agacttgggc          | aggtagacca |
| acaccacttt | attacgcgaa          | gagacttaca |
|            |                     | 180        |
| gagaagatcg | gtggcgctaa          | ggtttatttg |
| aagcgtgagg | atttggttca          | tggtggagca |
|            |                     | 240        |
| cataagacaa | acaatgccat          | cgggcaagcc |
| ttgttgGCCA | aattgatggg          | taagactagg |
|            |                     | 300        |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| ttgatagctg gcacaggtgc cggccagcat ggcgttgcaa cagcaatggc tggtgctctt  | 360  |
| ttagggatga aggtggatat ctatatgggc gccgaggacg ttgaaagaca aaagcttaac  | 420  |
| gtattcagaa tgaagttatt aggtgcgaat gttattccag tgaattcagg ttccaggacg  | 480  |
| ctgaaggacg caatagacga agccttgaga gactgggttg ctacttttga gtacacacat  | 540  |
| tatttgattg gttccgtggt tggaccccac ccgtacccta caatagttag agattttcaa  | 600  |
| tctgttatcg gtagagaggc gaaggctcaa attccagagg ccgagggaca gctacccgat  | 660  |
| gttatagtcg cctgcgttgg aggaggatca aatgcgatgg gaatcttcta cccttttgta  | 720  |
| aatgataaga aagttaagtt agtcggagtt gaggccggtg gtaaagggt agaaagcggc   | 780  |
| aaacactcgg catctcttaa tgccggtcaa gtaggagtaa gtcacgggat gttatcttac  | 840  |
| ttcttgcagg atgaggaagg tcaaattaag ccatctcatt ctatagcccc aggattagac  | 900  |
| tatcccgggtg ttgggcctga gcacgcttat ctgaagaaga tacagcgtgc tgagtacgtt | 960  |
| gcagtaactg atgaagaggc tttgaaggct ttccacgagt tatcaagaac tgagggaata  | 1020 |
| attcctgcct tggaatcggc acatgcggtt gcctacgcca tgaagcttgc taaggaaatg  | 1080 |
| tcaagagacg aaatcattat agtgaattta tcgggaaggg gcgataagga cttggacatt  | 1140 |
| gttttaaaag caagtggcaa tgtataa                                      | 1167 |

<210> 69  
 <211> 388  
 <212> PRT  
 <213> *Pyrococcus furiosus*

<400> 69

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Trp | Phe | Gly | Glu | Phe | Gly | Gly | Gln | Tyr | Val | Pro | Glu | Thr | Leu | Val |
| 1   |     |     |     | 5   |     |     |     | 10  |     |     |     |     |     | 15  |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gly | Pro | Leu | Lys | Glu | Leu | Glu | Lys | Ala | Tyr | Lys | Arg | Phe | Lys | Asp | Asp |
|     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Glu | Glu | Phe | Asn | Arg | Gln | Leu | Asn | Tyr | Tyr | Leu | Lys | Thr | Trp | Ala | Gly |
|     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Arg | Pro | Thr | Pro | Leu | Tyr | Tyr | Ala | Lys | Arg | Leu | Thr | Glu | Lys | Ile | Gly |
|     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gly | Ala | Lys | Val | Tyr | Leu | Lys | Arg | Glu | Asp | Leu | Val | His | Gly | Gly | Ala |
| 65  |     |     |     |     | 70  |     |     |     |     | 75  |     |     |     |     | 80  |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| His | Lys | Thr | Asn | Asn | Ala | Ile | Gly | Gln | Ala | Leu | Leu | Ala | Lys | Leu | Met |
|     |     |     |     | 85  |     |     |     |     | 90  |     |     |     |     | 95  |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Gly | Lys | Thr | Arg | Leu | Ile | Ala | Gly | Thr | Gly | Ala | Gly | Gln | His | Gly | Val |
|     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | Thr | Ala | Met | Ala | Gly | Ala | Leu | Leu | Gly | Met | Lys | Val | Asp | Ile | Tyr |
|     |     | 115 |     |     |     |     | 120 |     |     |     |     |     | 125 |     |     |

Met Gly Ala Glu Asp Val Glu Arg Gln Lys Leu Asn Val Phe Arg Met  
130 135 140

Lys Leu Leu Gly Ala Asn Val Ile Pro Val Asn Ser Gly Ser Arg Thr  
145 150 155 160

Leu Lys Asp Ala Ile Asp Glu Ala Leu Arg Asp Trp Val Ala Thr Phe  
165 170 175

Glu Tyr Thr His Tyr Leu Ile Gly Ser Val Val Gly Pro His Pro Tyr  
180 185 190

Pro Thr Ile Val Arg Asp Phe Gln Ser Val Ile Gly Arg Glu Ala Lys  
195 200 205

Ala Gln Ile Pro Glu Ala Glu Gly Gln Leu Pro Asp Val Ile Val Ala  
210 215 220

Cys Val Gly Gly Gly Ser Asn Ala Met Gly Ile Phe Tyr Pro Phe Val  
225 230 235 240

Asn Asp Lys Lys Val Lys Leu Val Gly Val Glu Ala Gly Gly Lys Gly  
245 250 255

Leu Glu Ser Gly Lys His Ser Ala Ser Leu Asn Ala Gly Gln Val Gly  
260 265 270

Val Ser His Gly Met Leu Ser Tyr Phe Leu Gln Asp Glu Glu Gly Gln  
275 280 285

Ile Lys Pro Ser His Ser Ile Ala Pro Gly Leu Asp Tyr Pro Gly Val  
290 295 300

Gly Pro Glu His Ala Tyr Leu Lys Lys Ile Gln Arg Ala Glu Tyr Val  
305 310 315 320

Ala Val Thr Asp Glu Glu Ala Leu Lys Ala Phe His Glu Leu Ser Arg  
325 330 335

Thr Glu Gly Ile Ile Pro Ala Leu Glu Ser Ala His Ala Val Ala Tyr  
340 345 350

Ala Met Lys Leu Ala Lys Glu Met Ser Arg Asp Glu Ile Ile Ile Val  
355 360 365

Asn Leu Ser Gly Arg Gly Asp Lys Asp Leu Asp Ile Val Leu Lys Ala  
370 375 380

Ser Gly Asn Val  
385

<210> 70  
 <211> 1173  
 <212> DNA  
 <213> Pyrococcus furiosus

<400> 70  
 atgtggtttg gtgagtttgg tggtcagtat gtgcctgaga cgctagttgg gccactgaag 60  
 gagttagaga aagcgtacaa gagattcaaa gatgacgagg agtttaaccg tcagttgaat 120  
 tattatctta agacgtgggc tggtaggccg acccctctat actacgctaa gcgtttaacc 180  
 gaaaagatag gtggcgccaa gatctatctg aagagagagg acctagtcca cggcggggct 240  
 cacaaaacta acaacgcaat tggtcaggca ccattagcaa agttaatggg caaaacacgt 300  
 ctaattgccg aaactggtgc tggccaacat ggagttgcta cggcaatggc gggcgccactg 360  
 ttaggtatga aagtggatat ctatatggga gcagaagacg tggagagaca gaagatgaac 420  
 gtgttccgta tgaactattt gggagctaata gttatcccgg ttaactctgg ctcaagaact 480  
 gcaaaggacg ccataaacga agctttgagg gattgggaag ccacattcga gtacacacac 540  
 taccttatag gctcagtcgt agggccgcat ccctacccca ccatagttcg tgacttcag 600  
 agtgttattg gccgtgaggc caaggcacag atcttggaag ctgaagggtca attgcctgac 660  
 gtaatagttg catgtgtggg tggcggttct aacgctatgg gtatatctta ccccttcgtg 720  
 aacgataaga aagtaaagct agttggagtg gaagcagggtg gcaaagggtt ggagtcgggc 780  
 aagcactcag cctcattaaa cgcaggtcag gtaggcgttc tgcacggtat gttatcctat 840  
 ttcttacagg acgaggaagg ccaaatcaaa ccatacgact caatcgacc tggttttagat 900  
 taccaggag tcgggtccga gcacgcatat ttaaagaaaa tccagagagc agagtacgtc 960  
 acggtgacgg atgaggaagc ccttaaagcg tttcacgagc tgtcgaggac cgagggcata 1020  
 atacctgctc ttgaaagtgc ccacgcagtt gcctatgcta tgaagttggc caaggagatg 1080  
 agtagagacg agattatcat agtaaactctg tcaggcagag gtgacaagga tcttgacata 1140  
 gtccttaagg cgagcggtaa tgttttggag taa 1173

<210> 71  
 <211> 390  
 <212> PRT  
 <213> Pyrococcus furiosus

<400> 71

Met Trp Phe Gly Glu Phe Gly Gly Gln Tyr Val Pro Glu Thr Leu Val  
 1 5 10 15

Gly Pro Leu Lys Glu Leu Glu Lys Ala Tyr Lys Arg Phe Lys Asp Asp  
 20 25 30

Glu Glu Phe Asn Arg Gln Leu Asn Tyr Tyr Leu Lys Thr Trp Ala Gly  
 35 40 45

Arg Pro Thr Pro Leu Tyr Tyr Ala Lys Arg Leu Thr Glu Lys Ile Gly  
 50 55 60



Gly Ala Lys Ile Tyr Leu Lys Arg Glu Asp Leu Val His Gly Gly Ala  
65 70 75 80

His Lys Thr Asn Asn Ala Ile Gly Gln Ala Pro Leu Ala Lys Leu Met  
85 90 95

Gly Lys Thr Arg Leu Ile Ala Glu Thr Gly Ala Gly Gln His Gly Val  
100 105 110

Ala Thr Ala Met Ala Gly Ala Leu Leu Gly Met Lys Val Asp Ile Tyr  
115 120 125

Met Gly Ala Glu Asp Val Glu Arg Gln Lys Met Asn Val Phe Arg Met  
130 135 140

Lys Leu Leu Gly Ala Asn Val Ile Pro Val Asn Ser Gly Ser Arg Thr  
145 150 155 160

Ala Lys Asp Ala Ile Asn Glu Ala Leu Arg Asp Trp Glu Ala Thr Phe  
165 170 175

Glu Tyr Thr His Tyr Leu Ile Gly Ser Val Val Gly Pro His Pro Tyr  
180 185 190

Pro Thr Ile Val Arg Asp Phe Gln Ser Val Ile Gly Arg Glu Ala Lys  
195 200 205

Ala Gln Ile Leu Glu Ala Glu Gly Gln Leu Pro Asp Val Ile Val Ala  
210 215 220

Cys Val Gly Gly Gly Ser Asn Ala Met Gly Ile Phe Tyr Pro Phe Val  
225 230 235 240

Asn Asp Lys Lys Val Lys Leu Val Gly Val Glu Ala Gly Gly Lys Gly  
245 250 255

Leu Glu Ser Gly Lys His Ser Ala Ser Leu Asn Ala Gly Gln Val Gly  
260 265 270

Val Leu His Gly Met Leu Ser Tyr Phe Leu Gln Asp Glu Glu Gly Gln  
275 280 285

Ile Lys Pro Ser His Ser Ile Ala Pro Gly Leu Asp Tyr Pro Gly Val  
290 295 300

Gly Pro Glu His Ala Tyr Leu Lys Lys Ile Gln Arg Ala Glu Tyr Val  
305 310 315 320

Thr Val Thr Asp Glu Glu Ala Leu Lys Ala Phe His Glu Leu Ser Arg  
325 330 335

Thr Glu Gly Ile Ile Pro Ala Leu Glu Ser Ala His Ala Val Ala Tyr  
340 345 350

Ala Met Lys Leu Ala Lys Glu Met Ser Arg Asp Glu Ile Ile Ile Val  
355 360 365

Asn Leu Ser Gly Arg Gly Asp Lys Asp Leu Asp Ile Val Leu Lys Ala  
370 375 380

Ser Gly Asn Val Leu Glu  
385 390

<210> 72  
<211> 1173  
<212> DNA  
<213> Pyrococcus furiosus

<400> 72  
atgtggttcg gagagttcgg cgggcagtac gtcccagaaa cgctagtcgg cccactgaag 60  
gagctggaga aagcatataa aagattcaag gatgacgagg agttcaacag gcaattaaac 120  
tattacctaa agacctgggc tggtaggccc acccctttat actatgcaaa acgtttaacg 180  
gagaaaatcg gtggcgcaaa agtgtacttg aagagagaag acttagtcca cggtggtgct 240  
cataaaacta acaacgctat agggcaggca cttttagcta agttgatggg taaaacgagg 300  
ttgattgcgg gtaccggtgc cggccaacat ggcgttgcta cggcaatggc gggagcccta 360  
ttgggcatga aggttgacat ttatatgggc gctgaggacg tggaacgtca gaaactgaac 420  
gtctatcgta tgaattgct tggtgccaac gttataccag tcaattccgg ctgcgtacg 480  
gttaaagacg ctttcgatga ggcactatgt gatagagttg ccactttcga atacaccac 540  
tatcttattg gtactgtctg ggggccacat ccgtatccaa ctattgtaag ggacttccag 600  
actgttatcg gaagagaggc aaaggcccaa atattagagg ctgagggaag gttacccgac 660  
gccattgttg catgcgtagg cggtggtagt aacgcaatgg gtatcttcta tcccttcgtt 720  
aacgataaga aggttaagtt agtgggtgtc gaggccggtg gtaaggggct tgagtccggt 780  
aaacattccg catctttgaa tgctgggcaa gtaggtgtca gtcattggaat gttatcatac 840  
tttctacagg acgaggaagg acagataaaa ctttctcatt ctatagcgcc aggtcttgac 900  
caccctggcg ttggaccaga gcatgcttat ttgaagaaga tccagagagc tgaatatgta 960  
gctgttacgg atgaggaagc tctaaaggcg ttccacgagt taagtaggac tgagggaatt 1020  
ataccggcat tagagtctgc ccacgctgta gcttacgcaa tgaagctggc caaagagatg 1080  
agccgtgacg agatcatcat cgtaaacttg tctggcagag gcgataagga cctagacatc 1140  
gtattaaagg cctcggggaa cgtcctagag tag 1173

<210> 73  
<211> 390  
<212> PRT  
<213> Pyrococcus furiosus

<400> 73

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Trp | Phe | Gly | Glu | Phe | Gly | Gly | Gln | Tyr | Val | Pro | Glu | Thr | Leu | Val | 1   | 5   | 10  | 15  |
| Gly | Pro | Leu | Lys | Glu | Leu | Glu | Lys | Ala | Tyr | Lys | Arg | Phe | Lys | Asp | Asp | 20  | 25  | 30  |     |
| Glu | Glu | Phe | Asn | Arg | Gln | Leu | Asn | Tyr | Tyr | Leu | Lys | Thr | Trp | Ala | Gly | 35  | 40  | 45  |     |
| Arg | Pro | Thr | Pro | Leu | Tyr | Tyr | Ala | Lys | Arg | Leu | Thr | Glu | Lys | Ile | Gly | 50  | 55  | 60  |     |
| Gly | Ala | Lys | Val | Tyr | Leu | Lys | Arg | Glu | Asp | Leu | Val | His | Gly | Gly | Ala | 65  | 70  | 75  | 80  |
| His | Lys | Thr | Asn | Asn | Ala | Ile | Gly | Gln | Ala | Leu | Leu | Ala | Lys | Leu | Met | 85  | 90  | 95  |     |
| Gly | Lys | Thr | Arg | Leu | Ile | Ala | Gly | Thr | Gly | Ala | Gly | Gln | His | Gly | Val | 100 | 105 | 110 |     |
| Ala | Thr | Ala | Met | Ala | Gly | Ala | Leu | Leu | Gly | Met | Lys | Val | Asp | Ile | Tyr | 115 | 120 | 125 |     |
| Met | Gly | Ala | Glu | Asp | Val | Glu | Arg | Gln | Lys | Leu | Asn | Val | Tyr | Arg | Met | 130 | 135 | 140 |     |
| Lys | Leu | Leu | Gly | Ala | Asn | Val | Ile | Pro | Val | Asn | Ser | Gly | Ser | Arg | Thr | 145 | 150 | 155 | 160 |
| Val | Lys | Asp | Ala | Phe | Asp | Glu | Ala | Leu | Cys | Asp | Arg | Val | Ala | Thr | Phe | 165 | 170 | 175 |     |
| Glu | Tyr | Thr | His | Tyr | Leu | Ile | Gly | Thr | Val | Trp | Gly | Pro | His | Pro | Tyr | 180 | 185 | 190 |     |
| Pro | Thr | Ile | Val | Arg | Asp | Phe | Gln | Thr | Val | Ile | Gly | Arg | Glu | Ala | Lys | 195 | 200 | 205 |     |
| Ala | Gln | Ile | Leu | Glu | Ala | Glu | Gly | Arg | Leu | Pro | Asp | Ala | Ile | Val | Ala | 210 | 215 | 220 |     |
| Cys | Val | Gly | Gly | Gly | Ser | Asn | Ala | Met | Gly | Ile | Phe | Tyr | Pro | Phe | Val | 225 | 230 | 235 | 240 |
| Asn | Asp | Lys | Lys | Val | Lys | Leu | Val | Gly | Val | Glu | Ala | Gly | Gly | Lys | Gly | 245 | 250 | 255 |     |
| Leu | Glu | Ser | Gly | Lys | His | Ser | Ala | Ser | Leu | Asn | Ala | Gly | Gln | Val | Gly | 260 | 265 | 270 |     |

Val Ser His Gly Met Leu Ser Tyr Phe Leu Gln Asp Glu Glu Gly Gln  
275 280 285

Ile Lys Pro Ser His Ser Ile Ala Pro Gly Leu Asp His Pro Gly Val  
290 295 300

Gly Pro Glu His Ala Tyr Leu Lys Lys Ile Gln Arg Ala Glu Tyr Val  
305 310 315 320

Ala Val Thr Asp Glu Glu Ala Leu Lys Ala Phe His Glu Leu Ser Arg  
325 330 335

Thr Glu Gly Ile Ile Pro Ala Leu Glu Ser Ala His Ala Val Ala Tyr  
340 345 350

Ala Met Lys Leu Ala Lys Glu Met Ser Arg Asp Glu Ile Ile Ile Val  
355 360 365

Asn Leu Ser Gly Arg Gly Asp Lys Asp Leu Asp Ile Val Leu Lys Ala  
370 375 380

Ser Gly Asn Val Leu Glu  
385 390

<210> 74  
<211> 6347  
<212> DNA  
<213> *Saccharomyces cerevisiae*

<400> 74  
gtatccggct gttccttcat agccctttca atgaacgttg cagccctttg aagattggcc 60  
atthttgtcag gactcgagcc tgacagttgg accaacgcaa ctttaatttt ttgtgaaaga 120  
atcttcgaag cactcactact ggcatcttc acgccctcct gctattacaa aagctgtgtt 180  
tttacaagaa tcaaattaag ttagcaagat attatacaac attattgata atttcaatat 240  
cgtgttcgta cctgatgacg tatctgtgca ttgataaggc ccgcatggtt tcagaaagca 300  
gagcggaaacg attccaaatt agtggccttg tgctttgcat gtcaattgtg ttaccttcag 360  
ctcgtggatt tgttttatca atacacagtc tacagtcaag aatttttttt atcaaatttt 420  
gcgttcgagc gtataaaata gccgtgttag ctacttaagt tcctgttcag cgatagtttt 480  
tttccatcac acgtactatg gcaattaagt cctcagcgag ctgcgatgga atgcgtgcga 540  
tgagcgacct catgctatac ctgagaaagc aacctgacct acaggaaaga gttactcaag 600  
aataagaatt ttcgttttaa aacctaagag tcactttaaa atttgtatac acttattttt 660  
tttataactt atttaataat aaaaatcata aatcataaga aattcgctta tttagaagtg 720  
tcaacaacgt atctaccaac ggaatgcgtg cgattcaggc gtagtctgga acgtcgtatg 780  
ggatggacc agagacggat tgagaaactt gttcggttct tgggacgaat ctacattgga 840  
atggaactgg gtgcctgaag aaaccagtgg tgaaatcagc tgggatgtca attggcttac 900  
cgttttggtc aactggctctt tcaatgttga aagcagacaa caaagtagca ccggcaatcc 960

|                                                                    |      |
|--------------------------------------------------------------------|------|
| aaacagtaga ttgagctagg tgaataacctg ggcagtttct tctaccgtaa ccgaaagcag | 1020 |
| cctttcttgg gtctctaaca gtgttgtctg gtttaccatc aggaccaag tatctttctg   | 1080 |
| gacggaaaac ggatggatct ggatagactt ctgggtcatt caaaactgcc caggtgttag  | 1140 |
| caaaaaccaa tgtgttcttt ggaatcaaat aacctctgta aacatcatcc ttcatcaatt  | 1200 |
| tatgagggat ggctaattga gcaatttggg tccatctgaa taattccttg atacaagcgg  | 1260 |
| tcaaatatgg gagtgagtcg tcttcctcat cgtagtcagg gatttgaccg ttgttggtca  | 1320 |
| aagcatccaa ttcagcttga acctttcttt gcatttctgg gtatttaacc atggccaata  | 1380 |
| tgaaagcggg catagcagag acggtagtgt cgccaccacc aacgttaact tcagcagcag  | 1440 |
| tgttcttgat aacatgttct tgggtttcta aatcaccgtt caagtccata gcttgtaatc  | 1500 |
| tagcagaagc gtaagatggg ctgggttagac cttgtggagc caactttctc atagtttcgt | 1560 |
| atggcatgtc aaccatgtgg tcagcggcctt ctctccagac cttggcctta cgcttgaaga | 1620 |
| cggcacctgg gaaccaagct ggcaagtact tcaaagatgg gaaggagtca acccagaact  | 1680 |
| tacctggaac actagcaatg gccaaacctt cgttggctaa gtgggtagct tctaaccatg  | 1740 |
| ggatcatctt ggctagatcg ataccgtaac caatgtccaa agacatagca gcgatttgat  | 1800 |
| gtctgatgtg ttgagcccac cggtcgggag tcttgggtcaa ttgttgaacc aattggtgag | 1860 |
| cagccttgac ttgagcgtgt ctgaattgct tgataccctt ttcagagaat tccttagcaa  | 1920 |
| acattcttct tcttctctc caacggtcac cgtaagtgtg aaaaccaga tcaaattccc    | 1980 |
| aacctcatca ttcattgacc atagtggtatt ccaatctacc ggagtagatg gaaccacgtt | 2040 |
| ttccaagag atccgtgata gttccaagg tgtttaaaat gaccatttcg gtaccaccag    | 2100 |
| catctacgta caagatatca gtgttgtagt cagcacccca ttgcaagaaa gtcaaccatg  | 2160 |
| gagattcttc tggcatgtcg aacatgttac caataaatgg gattggaatt cctggtggac  | 2220 |
| caggtggcaa tctggatcta cgaactcttc tggagacgat gtagtaaata caaccagcaa  | 2280 |
| tgacgaaaga gaacaagaca gcaatcattg ttttatattt gttgtaaaaa gtagataatt  | 2340 |
| acttccttga tgatctgtaa aaaagagaaa aagaaagcat ctaagaactt gaaaaactac  | 2400 |
| gaattagaaa agaccaaata tgtatttctt gcattgacca atttatgcaa gtttatatat  | 2460 |
| atgtaaatgt aagtttcacg aggttctact aaactaaacc acccccttgg ttagaagaaa  | 2520 |
| agagtgtgtg agaacaggct gttgttgtca caggttcgg acaattctgt ttgaaagaga   | 2580 |
| gagagtaaca gtacgatcga acgaactttg ctctggagat cacagtgggc atcatagcat  | 2640 |
| gtggtactaa accctttccc gccattccag aaccttcgat tgcttgttac aaaacctgtg  | 2700 |
| agccgtcgct aggaccttgt tgtgtgacga aattggaagc tgcaatcaat aggaagacag  | 2760 |
| gaagtcgagc gtgtctgggt tttttcagtt ttgttctttt tgcaaacaac agttttattcc | 2820 |
| tggcatccac taaatataat ggagcccgtt ttttaagctg gcatccagaa aaaaaagaa   | 2880 |
| tcccagcacc aaaatattgt tttcttcacc aaccatcagt tcataggtcc attctcttag  | 2940 |
| cgcaactaca gagaacaggg gcacaaacag gcaaaaaacg ggcacaacct caatggagtg  | 3000 |

|                                                                     |      |
|---------------------------------------------------------------------|------|
| atgcaacctg cctggagtaa atgatgacac aaggcaattg acccacgcat gtatctatct   | 3060 |
| cattttctta caccttctat taccttctgc tctctctgat ttggaaaaag ctgaaaaaaa   | 3120 |
| aggttgaaac cagttccctg aaattattcc cctacttgac taataagtat ataaagacgg   | 3180 |
| taggtattga ttgtaattct gtaaattctat ttcttaaact tcttaaattc tacttttata  | 3240 |
| gtagtcttt tttttagttt taaaacacca agaacttagt ttcgaataaa cacacataaa    | 3300 |
| caaacaaaat ggcttctagt tcttccgatg tcttcgtttt gggcttaggt gttgttttgg   | 3360 |
| ctgccttgta tatcttcaga gaccaattat tcgctgcttc taagccaaag gtggctccag   | 3420 |
| tttccactac gaagcctgcc aacggttccg ctaacccaag agacttcac gccaaagatga   | 3480 |
| aacaaggtaa gaagagaatc gtaattcttct acggttctca aactgggtacc gctgaagaat | 3540 |
| atgctattcg tttggctaag gaagctaagc aaaagttcgg tctagcctcc ttggtttgtg   | 3600 |
| atccagaaga atacgatttt gaaaagttgg accaattgcc agaagattct attgctttct   | 3660 |
| tcgtcgttg c tacctatggt gaaggtgaac ctacagacaa cgctgtccaa ttgttgcaaa  | 3720 |
| acttgcaaga tgaaagcttc gaattctcct ctggtgagag aaagttgtca ggtttgaagt   | 3780 |
| acgttgtttt tggctcgggt aacaagacct acgaacatta caacctcatt gggagaactg   | 3840 |
| ttgacgctca attggccaag atgggtgcta tcagaatcgg tgaaagaggt gaaggtgatg   | 3900 |
| atgacaagtc catggaagaa gactacttgg aatggaagga tggtatgtgg gaagcgttg    | 3960 |
| ccactgctat ggggtgttgaa gaaggtcaag gtggtgactc cgctgatttc gtcgtttccg  | 4020 |
| aattggaatc tcaccaccca gaaaaggttt accaagggtga attttctgct agagctttaa  | 4080 |
| ccaaaaccaa gggatttcac gacgctaaga atccttttgc tgctccaatt gcggttgcta   | 4140 |
| gagaattggt ccaatctggt gtcgatagaa actgtgtcca cgtcgaattc aacattgaag   | 4200 |
| gctctggtat cacctatcaa cacggtgacc acgttggttt gtggccattg aatccagatg   | 4260 |
| ttgaagtcga acggttggtg tgtgttttag gtttagctga aaagagagat gctgtcatct   | 4320 |
| ccattgaatc ctagacccg gctttggcta aggttccatt cccagtcca actacttacg     | 4380 |
| gtgctgtgtt gagacactac attgacatct ctgctgtcgc cggtagacaa atcttgggta   | 4440 |
| ctttgtccaa attcgctcca accccagaag ctgaagcttt cttgagaaac ttgaacacta   | 4500 |
| acaaggaaga ataccacaac gtcgtcgcta acggttggtt gaaattgggt gaaattttgc   | 4560 |
| aaatcgctac cggtaacgac attactgtcc caccaactac tgccaacacc accaaatggc   | 4620 |
| caattccatt cgacatcatt gtttctgcca tccaagatt gcaaccaaga tactactcta    | 4680 |
| tctcttcttc cccaaaaatt catccaaaca ccatccacgc taccgttggt gtgctcaaat   | 4740 |
| acgaaaacgt tccaaccgaa ccaatcccaa gaaagtgggt ttacggtgtc ggtagtaact   | 4800 |
| tcttggtgaa tttaaagtac gctgttaaca aggaaccagt tccatacatc actcaaaatg   | 4860 |
| gcgaacaaag agtcggtgtc ccggaatact tgattgctgg tccacgtggt tcttacaaga   | 4920 |
| ctgaatcttt ctacaaggct ccaatccatg ttagacgttc tactttccgt ttgccaacca   | 4980 |
| acccaaagtc tccagtcac atgattgggt caggtactgg tgtcgcccc tttagaggct     | 5040 |
| tcgttcaaga aagagttgcc ttggccagaa gatccatcga aaagaacggt cctgactctt   | 5100 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| tggtgactg gggtcgtatt tccttgttct acggttgtag aagatccgac gaagacttct   | 5160 |
| tgtacaagga cgaatggcca caatacgaag ctgagttgaa gggtaagttc aagttgcact  | 5220 |
| gtgctttctc cagacaaaac tacaagccag acggttctaa gatttacgtc caagatttga  | 5280 |
| tctgggaaga cagagaacac attgccgatg ccatcttaaa cggttaagggt tacgtctaca | 5340 |
| tctgcggtga agctaagtcc atgtctaaac aagttgaaga agttctagcc aagatcttgg  | 5400 |
| gcgaagccaa aggtggttcc ggtccagttg aaggtgttgc tgaagtcaag ttactgaagg  | 5460 |
| aacggtccag attgatgttg gatgtctggt ctgaacaaaa gttaatttct gaagaagatt  | 5520 |
| tggaatgaat cgcggtgcatt catccgctct aaccgaaaag gaaggagtta gacaacctga | 5580 |
| agtctaggtc cctatttatt tttttatagt tatgttagta ttaagaacgt tatttatatt  | 5640 |
| tcaaattttt ctttttttct tgtacagacg cgtgtacgca tgtaacatta tactgaaaac  | 5700 |
| cttgcttgag aagggttttg gacgctcgaa gatcgctcc caattcgccc tatagtgagt   | 5760 |
| cgtattacgc gcgctcactg gccgtcgttt tacaacgtcg tgactgggaa aaccctggcg  | 5820 |
| ttaccctgc aggactagtg ctgaggcatt aatagtttac tcaattcttg aagccaattt   | 5880 |
| gtacaattcc ccattagagt caaataaaag gatgcctcac ggaggtatgt taccgctgct  | 5940 |
| atttcacatg gctcattgaa ttagagggtg aatttggtgt accctcccct cctcatctga  | 6000 |
| tgaagtagtg atccgacaat tcttaaaagt tgtagacatt acttttacca ccaactaagt  | 6060 |
| tgtatttata ttgctaccct tatcctttta tatctaacta gcgctcataa gggtggggca  | 6120 |
| atactaaaac tgtgttctta ttcaactcat taaatacgtg gcagtacgta ccctattaga  | 6180 |
| aacaatagga aacagcagag tcggaagaag ccaaatgcc aatttgaagt ccaaaacctt   | 6240 |
| gtcaagccaa tctttgggag cggctattcc tccagaaatt gtgtaccaa tacttacata   | 6300 |
| ccagtttagg gatttgttaa gaaatgacca tccaggtagc gcagaaa                | 6347 |

<210> 75  
 <211> 4792  
 <212> DNA  
 <213> *Saccharomyces cerevisiae*

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| <400> 75                                                           |     |
| caatctggcg gcttgagttc tcaacatgtt ttatttttta cttatattgc tggtagggta  | 60  |
| aaaaaatata actcctagga atagggtgtc tatatgtttt tgtcttgctt ctataattgt  | 120 |
| aacaacaag gaaagggaaa atactgggtg taaaagccat tgagtcaagt taggtcatcc   | 180 |
| ctttatata aaatttttca attttttttc caagattctt gtacgattaa ttattttttt   | 240 |
| tttgctcct acagcgtgat gaaaatttcg cctgctgcaa gatgagcggg aacgggcgaa   | 300 |
| atgtgcacgc gcacaactta cgaaacgcgg atgagtcact gacagccacc gcagaggttc  | 360 |
| tgactcctac tgagctctat tggagggtggc agaaccggta ccggaggagg ccgctataac | 420 |
| cggtttgaat ttattgtcac agtgtcacat cagcattaag tcctcagcga gctcgcatgg  | 480 |
| aatgcgtgcg atgagcgacc tcatgctata cctgagaaag caacctgacc tacaggaaag  | 540 |
| agttactcaa gaataagaat tttcgtttta aaacctaa gtcacttta aatttgtata     | 600 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| cacttatttt ttttataact tattttaataa taaaaatcat aaatcataag aaattcgctt | 660  |
| atttagaagt gtcaacaacg tatctaccaa cggaatgcgt gcgattcagg tggagtccaa  | 720  |
| acccaacaaa ggatttggaa ttggcttacc ggacgtggag gattccttta gcaacgtgga  | 780  |
| agtgaattca ccgttgtcgt tgttacctct agcatcgtgg aaagcagcaa cacccttctt  | 840  |
| aacaaagttg attcttttctt cttcagaacc ccattgcatg aagtcagtcc acataacaat | 900  |
| gtgagcggcg ataccagcgg taaccttggc gtagttgatg gaatgcttgg aagtacgggc  | 960  |
| gtaagattgc aagtaagctt gtctcatggt tgtaccaact tgttcgtctt ggaatctgct  | 1020 |
| aatcaagtaa cagtcaccca agaagtaacc caaatccaat gaagctggac cgtacttaca  | 1080 |
| caattcccag tctaagatgt agatcttttg caacttagat gggttacctt cttcaagttg  | 1140 |
| caacaagatg ttcccagacc acaagtcagc catgaccaa gtttcttcgg agtgcataac   | 1200 |
| atcgtcaact agatccttga caacagttgg caacaatgga tcatcgacgc cgtatttagc  | 1260 |
| ggcgtttggg ataatagttt ggtacaattg gtcagagggt gttctaccga caatgttacc  | 1320 |
| agagaagaac ttgaattctg ggtcgtctct tctttctcta cctatgttgt gcaatctggc  | 1380 |
| gacgaaacca ccaatctcgg taccaaccaa tctagcaata tcggtagcca aaggtggctt  | 1440 |
| agcagtaacg tagtctaata aggtcttcat tttaccgaca tcttgcataa tcaaagcatt  | 1500 |
| gttttccaag tcatagttga gaccttctgg aacagagaca ataccatcaa caccaccaa   | 1560 |
| aacttctctg ttagccatca tcaacttgat agcttgggtat tcgtagacag aacgttcaac | 1620 |
| accgattttg aaatcttcat cagtagacat gtgtggttga gcgtgcttca aaatgataga  | 1680 |
| agtgtgacct tggtagggag cgttcaattt aatacgccag gtgacgttaa cgaaaccacc  | 1740 |
| ggataatctc ttgacaccag aagtgtcaac atctaaagac aaatgcttgg tcaaataagg  | 1800 |
| gattaaaccg tcttcagtct tcaagtcaaa ggccattggt ttatatattgt tgtaaaaagt | 1860 |
| agataattac ttccttgatg atctgtaaaa aagagaaaaa gaaagcatct aagaacttga  | 1920 |
| aaaactacga attagaaaag accaaatatg tatttcttgc attgaccaat ttatgcaagt  | 1980 |
| ttatatatat gtaaagttaa gtttcacgag gttctactaa actaaaccac ccccttggtt  | 2040 |
| agaagaaaag agtgtgtgag aacaggctgt tgttgtcaca cgattcggac aattctgttt  | 2100 |
| gaaagagaga gagtaacagt acgatcgaac gaactttgct ctggagatca cagtgggcat  | 2160 |
| catagcatgt ggtactaaac ctttcccg cattccagaa ctttcgattg cttgttacia    | 2220 |
| aacctgtgag ccgtcgctag gaccttggtg tgtgacgaaa ttggaagctg caatcaatag  | 2280 |
| gaagacagga agtcgagcgt gtctgggttt tttcagtttt gttctttttg caaacaacag  | 2340 |
| tttattcctg gcatccacta aatataatgg agcccgttt ttaagctggc atccagaaaa   | 2400 |
| aaaaagaatc ccagcaccaa aatattgttt tcttcaccaa ccatcagttc ataggtccat  | 2460 |
| tctcttagcg caactacaga gaacaggggc acaaacaggc aaaaaacggg cacaacctca  | 2520 |
| atggagtgat gcaacctgcc tggagtaaat gatgacacaa ggcaattgac ccacgcatgt  | 2580 |
| atctatctca ttttcttaca ctttctatta ctttctgctc tctctgattt ggaaaaagct  | 2640 |



|                                                                    |      |
|--------------------------------------------------------------------|------|
| gaaaaaaaaag gttgaaacca gttccctgaa attattcccc tacttgacta ataagtatat | 2700 |
| aaagacggta ggtattgatt gtaattctgt aaatctattt cttaaacttc ttaaattcta  | 2760 |
| cttttatagt tagtcttttt tttagtttta aaacaccaag aacttagttt cgaataaaca  | 2820 |
| cacataaaca aacaaaatgc atatcagaaa cccatataga actccaattg actaccaagc  | 2880 |
| tttgtctgaa gctttccac cattgaagcc atttgtttcc gttaacgctg atggtacctc   | 2940 |
| ctcagttgac ttgaccattc cagaagccca aagagctttt accgctgccc ttttgcacag  | 3000 |
| agacttcggc ttgactatga ctatcccaga agatcgtttg tgtccaaccg ttccaaacag  | 3060 |
| attgaactac gttttgtgga ttgaagacat tttcaactac accaacaaga ctttgggttt  | 3120 |
| atctgacgac cgtccaatca aggggtgttg tatcggtacc ggtgcttctg ccatttaccc  | 3180 |
| aatgttggct tgcgccagat tcaaggcttg gtccatgggt ggtactgaag ttgaaagaaa  | 3240 |
| gtgtatcgac actgctagat taaacgttgt tgctaacaac ttgcaagatc gtctatccat  | 3300 |
| cttggaacc tctattgacg gtccaatttt agtcccaatt ttcgaagcta ccgaagaata   | 3360 |
| cgaatacgaa ttcacatgt gtaaccacc tttctacgat ggtgccgctg acatgcaaac    | 3420 |
| tagcgatgct gcaaagggtt ttggtttcgg tgtcgggtgct ccacactctg gtacagtcac | 3480 |
| cgaatgtct actgaagggt gtgaatccgc tttcgtggct caaatggta gagaatctct    | 3540 |
| taagttgaga accagatgta gatggtacac ttctaactta ggtaagttga aatctttgaa  | 3600 |
| ggaaattgtc ggtttgttga aggaattaga aatctctaata tacgccatca acgaatatgt | 3660 |
| ccaaggttcc actagaagat acgctgtcgc ttggagtttc actgatatcc aattgccaga  | 3720 |
| agaattgtcc agaccatcta atcctgaatt gtctcttttg ttcgactaca aggatgacga  | 3780 |
| tgacaaatga atcgcgtgca ttcattccgct ctaaccgaaa aggaaggagt tagacaacct | 3840 |
| gaagtctagg tccctattta tttttttata gttatgttag tattaagaac gttatttata  | 3900 |
| tttcaaattt ttcttttttt tctgtacaga cgcgtgtacg catgtaacat tatactgaaa  | 3960 |
| accttgcttg agaaggtttt gggacgctcg aagatcgctg acccaattcg ccctatagtg  | 4020 |
| agtcgtatta cgcgcgtca ctggccgtcg ttttacaacg tcgtgactgg gaaaaccctg   | 4080 |
| gcgttacccc tgcaggacta gtgctgaggc attaatgcaa ctcagaagtt tgacagcaag  | 4140 |
| caagttcatc attcgaacta gccttattgt tttagttcag tgacagcgaa ctgccgtact  | 4200 |
| cgatgcttta tttctcacgg tagagcggaa gaacagatag gggcagcgtg agaagagtta  | 4260 |
| gaaagtaaata ttttatcacg tctgaagtat tcttattcat aggaaatttt gcaaggtttt | 4320 |
| ttagctcaat aacgggctaa gttatataag gtgttcacgc gattttcttg ttatgtatac  | 4380 |
| ctcttctctg aggaatggta ctactgtcct gatgtaggct ctttaaattg gtgggcaaga  | 4440 |
| ataacttatc gatattttgt atattggtct tggagtccac cacgtaatgc ctgtttaaga  | 4500 |
| ccatcagtta actctagtat tatttggctt tggctactgg ccgtttgcta ttattcaagt  | 4560 |
| cttttgtgcc ttcccgtcgg gtaaggaggt tatttaggga tacagaatct aacgaaaact  | 4620 |
| aatctcaat gattaactct atttaactct tttttgaaag gcaaaagagg tcccttgttc   | 4680 |
| acttacaacg ttcttagcca aattcgctta tcacttacta cttcacgata tacagaagta  | 4740 |

aaaacatata aaaagatgtc tgtttgttta gccatcacia aaggtatcgc ag 4792

<210> 76

<211> 9145

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 76

tgttttatat ttgttgtaaa aagtagataa ttacttcctt gatgatctgt aaaaaagaga 60  
aaaagaaagc atctaagaac ttgaaaaact acgaattaga aaagaccaa tatgtatttc 120  
ttgcattgac caatttatgc aagtttatat atatgtaaat gtaagtttca cgaggttcta 180  
ctaaactaaa ccacccctt ggtagaaga aaagagtgtg tgagaacagg ctgttgttgt 240  
cacacgattc ggacaattct gtttgaaaga gagagagtaa cagtacgac gaacgaactt 300  
tgctctggag atcacagtgg gcatcatagc atgtgggtact aaacccttc ccgccattcc 360  
agaaccttcg attgcttggt acaaaacctg tgagccgtcg ctaggacctt gttgtgtgac 420  
gaaattggaa gctgcaatca ataggaagac aggaagtcga gcgtgtctgg gttttttcag 480  
ttttgttctt ttgcaaaca acagtttatt cctggcatcc actaaatata atggagcccg 540  
ctttttaagc tggcatccag aaaaaaaaaa aatcccagca ccaaatatt gttttcttca 600  
ccaacatca gttcataggt ccattctctt agcgcaacta cagagaacag gggcacaaac 660  
aggcaaaaaa cgggcacaaac ctcaatggag tgatgcaacc tgcctggagt aaatgatgac 720  
acaaggcaat tgaccacgc atgtatctat ctcatcttct tacaccttct attaccttct 780  
gctctctctg atttgaaaa agctgaaaa aaaggttgaa accagttccc tgaaattatt 840  
cccctacttg actaataagt atataagac ggtaggtatt gattgtaatt ctgtaaatct 900  
atttcttaaa cttcttaaat tctactttta tagttagtct tttttttagt tttaaaacac 960  
caagaactta gtttgaata aacacacata aacaacaaa ggatccatga tgaagtcttg 1020  
gagaaagtac acacaacaag aaatggatga aaagattact gaatctttgg aaaagacttt 1080  
gaactacgat aacactaaga caatcggat tccaggtact aagttggatg atacagtttt 1140  
ctatgatgat cattctttcg ttaagcattc accatacttg agaactttta ttcaaaacc 1200  
aaaccatatt ggttgtcata cttatgataa ggctgatatt ttgttcgggtg gtacattcga 1260  
tatcgaaaga gaattaatcc aattgttagc aatcgatgtt ttgaacggta acgatgaaga 1320  
atttgatggt tacgttactc aaggtgggtac agaagctaac atccaagcaa tgtgggttta 1380  
cagaaactac ttcaagaaag aaagaaaggc taagcatgaa gaaatcgcta tcatcacttc 1440  
agcagatata cattactctg catacaaagg ttcagatttg ttgaacatcg atattattaa 1500  
ggttccagtt gatttttatt caagaaaaat tcaagaaaat acattggatt caattgttaa 1560  
agaagctaaa gaaattggta aaaagtactt catcgttatc tctaactgg gtactacaat 1620  
gtttgggtca gttgatgac cagatttgta cgctaacatc ttcgataagt acaatttgga 1680  
atacaaaatt catgttgatg gtgcatttgg tggttttata tatccaattg ataataagga 1740  
atgtaaaact gatttctcta ataagaacgt ttcttcaatc acattagatg gtcataagat 1800

|                                                                    |      |
|--------------------------------------------------------------------|------|
| gttgcaagct ccatacggta ctggtatctt cgtttcaaga aagaatttga tccataacac  | 1860 |
| tttgacaaag gaagcaactt acatcgaaaa tttggatggt acattgtctg gttcaagatc  | 1920 |
| tggttcaaat gctgttgcaa tttggatggt tttagcttct tatggtccat acggttggat  | 1980 |
| ggaaaagatt aataagttga gaaatagaac taaatggttg tgtaagcaat tgaacgatat  | 2040 |
| gagaattaaa tattacaaag aagattcaat gaatattggt acaattgaag aacaatatgt  | 2100 |
| taataaggaa atcgctgaaa agtacttttt agttccagaa gttcataacc caactaacia  | 2160 |
| ctggtacaag atcgttgtta tggaacatgt tgaattggat atcttgaact ctttggttta  | 2220 |
| cgatttgaga aagtttaata aggaacattt gaaggcaatg catcatcatc atcatcatta  | 2280 |
| accgcggcta gctaagatcc gctctaaccg aaaaggaagg agttagacaa cctgaagtct  | 2340 |
| aggtccttat ttattttttt atagttatgt tagtattaag aacgttattt atatttcaaa  | 2400 |
| tttttctttt ttttctgtac agacgcgtgt acgcatgtaa cattatactg aaaaccttgc  | 2460 |
| ttgagaaggt tttgggacgc tcgaagatcc agctgcatta atgaatcggc caacgcgcgg  | 2520 |
| ggagaggcgg tttgcgtatt gggcgctctt ccgcttcctc gctcactgac tcgctgcgct  | 2580 |
| cggtcgttcg gctgcggcga gcggtatcag ctactcaaa ggcggtaata cggttatcca   | 2640 |
| cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga  | 2700 |
| accgtaaaaa ggccgcgttg ctggcgtttt tccataggct ccgccccct gacgagcatc   | 2760 |
| acaaaaatcg acgctcaagt cagagggtggc gaaacccgac aggactataa agataccagg | 2820 |
| cgtttcccc tggaagctcc ctctgtgcgt ctctgtttc gaccctgccg cttaccggat    | 2880 |
| acctgtccgc ctttctcct tcgggaagcg tggcgctttc tcatagctca cgctgtaggt   | 2940 |
| atctcagttc ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa cccccgttc   | 3000 |
| agcccgaccg ctgcgcctta tccggtaact atcgtcttga gtccaaccg gtaagacacg   | 3060 |
| acttatcgcc actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg  | 3120 |
| gtgctacaga gttcttgaag tggtagccta actacggcta cactagaaga acagtatttg  | 3180 |
| gtatctgcgc tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg  | 3240 |
| gcaaacaaac caccgctggt agcggtaggt tttttgtttg caagcagcag attacgcgca  | 3300 |
| gaaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga  | 3360 |
| acgaaaactc acgttaaggg attttggctc tgagattatc aaaaaggatc ttcacctaga  | 3420 |
| tccttttaaa ttaaaaatga agttttaaat caatctaaag tatatatgag taaacttgg   | 3480 |
| ctgacagtta ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcgtt  | 3540 |
| catccatagt tgcctgactc cccgtcgtgt agataactac gatacgggag ggcttaccat  | 3600 |
| ctggccccag tgctgcaatg ataccgcgag acccacgctc accggctcca gatttatcag  | 3660 |
| caataaacca gccagccgga agggccgagc gcagaagtgg tcctgcaact ttatccgcct  | 3720 |
| ccatccagtc tattaattgt tgccgggaag ctagagtaag tagttcgcca gttaatagtt  | 3780 |
| tgcgcaacgt tgttgccatt gctacaggca tcgtggtgtc acgctcgtcg tttggtatgg  | 3840 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| cttcattcag ctccggttcc caacgatcaa ggcgagttac atgatcccc atgttgtgca   | 3900 |
| aaaaagcggg tagctccttc ggtcctccga tcgttgtcag aagtaagttg gccgcagtgt  | 3960 |
| tatcactcat ggttatggca gactgcata attctcttac tgtcatgcc tccgtaagat    | 4020 |
| gcttttctgt gactgggtgag tactcaacca agtcattctg agaatagtgt atgcggcgac | 4080 |
| cgagttgctc ttgccggcg tcaatacggg ataataccgc gccacatagc agaactttaa   | 4140 |
| aagtgctcat cattggaaaa cgttcttcgg ggcgaaaact ctcaaggatc ttaccgctgt  | 4200 |
| tgagatccag ttcgatgtaa cccactcgtg cacccaactg atcttcagca tcttttactt  | 4260 |
| tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa  | 4320 |
| ggcgacacg gaaatgttga atactcatac tcttcctttt tcaatattat tgaagcattt   | 4380 |
| atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa aataaacaaa  | 4440 |
| taggggttcc gcgcacattt ccccgaaaag tgccacctga acgaagcatc tgtgcttcat  | 4500 |
| ttttagaac aaaaatgcaa cgcgagagcg ctaatttttc aaacaaagaa tctgagctgc   | 4560 |
| atttttacag aacagaaatg caacgcgaaa gcgctatttt accaacgaag aatctgtgct  | 4620 |
| tcatttttgt aaaacaaaaa tgcaacgcga gagcgctaata ttttcaaaca aagaatctga | 4680 |
| gctgcatttt tacagaacag aaatgcaacg cgagagcgct attttacc aaagaatct     | 4740 |
| atacttcttt tttgttctac aaaaatgcat cccgagagcg ctatttttct aacaaagcat  | 4800 |
| cttagattac tttttttctc ctttgtgcgc tctataatgc agtctcttga taactttttg  | 4860 |
| cactgtaggt ccgttaaggt tagaagaagg ctactttggt gtctattttc tcttcataa   | 4920 |
| aaaaagcctg actccacttc ccgcgtttac tgattactag cgaagctgcg ggtgcatttt  | 4980 |
| ttcaagataa aggcattccc gatttatattc tataccgatg tggattgcgc atactttgtg | 5040 |
| aacagaaagt gatagcgttg atgattcttc attggtcaga aaattatgaa cggtttcttc  | 5100 |
| tattttgtct ctatatacta cgtataggaa atgtttacat tttcgtattg ttttcgattc  | 5160 |
| actctatgaa tagttcttac tacaattttt ttgtctaaag agtaatacta gagataaaca  | 5220 |
| taaaaaatgt agaggtcgag tttagatgca agttcaagga gcgaaagggtg gatgggtagg | 5280 |
| ttatataggg atatagcaca gagatatata gcaaagagat acttttgagc aatgtttgtg  | 5340 |
| gaagcggtat tcgcaatatt ttagtagctc gttacagtcc ggtgcgtttt tggttttttg  | 5400 |
| aaagtgcgtc ttcagagcgc ttttggtttt caaagcgct ctgaagttcc tatactttct   | 5460 |
| agagaatagg aacttcggaa taggaacttc aaagcgtttc cgaaaacgag cgcttccgaa  | 5520 |
| aatgcaacgc gagctgcgca catacagctc actgttcacg tcgcacctat atctgcgtgt  | 5580 |
| tgctgtata tatatataca tgagaagaac ggcatagtgc gtgtttatgc ttaaatgcgt   | 5640 |
| acttatatgc gtctatttat gtaggatgaa aggtagtcta gtacctctg tgatattatc   | 5700 |
| ccattccatg cggggtatcg tatgcttcct tcagcactac ctttagctg ttctatatgc   | 5760 |
| tgccactcct caattggatt agtctcatcc ttcaatgcta tcatttcctt tgatattgga  | 5820 |
| tcatactaag aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg  | 5880 |
| ccctttcgtc tcgcgcgttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccg | 5940 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| gagacgggtca cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg | 6000 |
| tcagcgggtg ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta  | 6060 |
| ctgagagtgc accatatctt acaccgcata gatccgtcga gttcaagaga aaaaaaaga   | 6120 |
| aaaagcaaaa agaaaaaagg aaagcgcgcc tcgttcagaa tgacacgtat agaattgatgc | 6180 |
| attaccttgt catcttcagt atcatactgt tcgtatacat acttactgac attcataggt  | 6240 |
| atacatatat acacatgtat atatatcgta tgctgcagct ttaaataatc ggtgtcacta  | 6300 |
| cataagaaca cttttggtgg agggaaacatc gttggtacca ttgggcgagg tggcttctct | 6360 |
| tatggcaacc gcaagagcct tgaacgcact ctactacgg tgatgatcat tcttgccctcg  | 6420 |
| cagacaatca acgtggaggg taattctgct agcctctgca aagctttcaa gaaaatgcgg  | 6480 |
| gatcatctcg caagagagat ctctactttt ctccctttgc aaaccaagtt cgacaactgc  | 6540 |
| gtacggcctg ttcgaaagat ctaccaccgc tctggaaagt gcctcatcca aaggcgcaaa  | 6600 |
| tcctgatcca aaccttttta ctccacgcac ggcccctagg gcctctttaa aagcttgacc  | 6660 |
| gagagcaatc ccgcagtctt cagtgggtgtg atggctgtct atgtgtaagt caccaatgca | 6720 |
| ctcaacgatt agcgaccagc cggaatgctt ggccagagca tgtatcatat ggtccagaaa  | 6780 |
| ccctatacct gtgtggacgt taatcacttg cgattgtgtg gcctgttctg ctactgcttc  | 6840 |
| tgctcttttt tctgggaaga tcgagtgtc tatcgctagg ggaccaccct ttaaagagat   | 6900 |
| cgcaatctga atcttggttt catttgtaat acgctttact agggctttct gctctgtcat  | 6960 |
| ctttgccttc gtttatcttg cctgctcatt ttttagtata ttcttcgaag aaatcacatt  | 7020 |
| actttatata atgtataatt cattatgtga taatgccaat cgctaagaaa aaaaaagagt  | 7080 |
| catccgctag gtggaaaaaa aaaaatgaaa atcattaccg aggcataaaa aaatatagag  | 7140 |
| tgtactagag gaggccaaga gtaatagaaa aagaaaattg cgggaaagga ctgtgttatg  | 7200 |
| acttccctga ctaatgccgt gttcaaacga tacctggcag tgactcctag cgctcaccaa  | 7260 |
| gctcttaaaa cgggaattta tgggtgcactc tcagtacacg cgccagatct gtttagcttg | 7320 |
| cctcgcccc gccgggtcac ccggccagcg acatggaggc ccagaatacc ctcttgaca    | 7380 |
| gtcttgacgt gcgcagctca ggggcatgat gtgactgtcg cccgtacatt tagcccatac  | 7440 |
| atccccatgt ataattcttt gcatccatac attttgatgg ccgcacggcg cgaagcaaaa  | 7500 |
| attacggctc ctcgctgcag acctgcgagc agggaaacgc tcccctcaca gacgcgttga  | 7560 |
| attgtcccca cgccgcgccc ctgtagagaa atataaaagg ttaggatttg cactgaggt   | 7620 |
| tcttctttca tatacttcct tttaaaatct tgctaggata cagttctcac atcacatccg  | 7680 |
| aacataaaca accatgggta aggaaaagac tcacgtttcg aggccgcgat taaattccaa  | 7740 |
| catggatgct gatattatg ggtataaatg ggctcgcgat aatgtcgggc aatcaggtgc   | 7800 |
| gacaatctat cgattgtatg ggaagcccga tgcgccagag ttgtttctga aacatggcaa  | 7860 |
| aggtagcggt gccaatgatg ttacagatga gatggtcaga ctaaactggc tgacggaatt  | 7920 |
| tatgcctctt ccgaccatca agcattttat ccgtactcct gatgatgcat ggttactcac  | 7980 |

|            |            |            |            |            |             |      |
|------------|------------|------------|------------|------------|-------------|------|
| cactgcgac  | cccggcaaaa | cagcattcca | ggtattagaa | gaatatcctg | attcaggtga  | 8040 |
| aaatatgtt  | gatgcgctgg | cagtgttcct | gcgccggtt  | cattcgattc | ctgtttgtaa  | 8100 |
| ttgtccttt  | aacagcgatc | gcgtatttcg | tctcgctcag | gcgcaatcac | gaatgaataa  | 8160 |
| cggtttggt  | gatgcgagt  | attttgatga | cgagcgtaat | ggctggcctg | ttgaacaagt  | 8220 |
| ctggaaagaa | atgcataagc | ttttgccatt | ctcaccggat | tcagtcgtca | ctcatggtga  | 8280 |
| tttctcactt | gataacctta | tttttgacga | ggggaaatta | ataggttgta | ttgatgttgg  | 8340 |
| acgagtcgga | atcgcagacc | gataccagga | tcttgccatc | ctatggaact | gcctcgggtga | 8400 |
| gttttctcct | tcattacaga | aacggctttt | tcaaaaatat | ggtattgata | atcctgatata | 8460 |
| gaataaattg | cagtttcatt | tgatgctcga | tgagtttttc | taatcagtac | tgacaataaaa | 8520 |
| aagattcttg | ttttcaagaa | cttgtcattt | gtatagtttt | tttatattgt | agttgttcta  | 8580 |
| ttttaatcaa | atgttagcgt | gatttatatt | ttttttcgcc | tcgacatcat | ctgcccagat  | 8640 |
| gcgaagttaa | gtgcgcagaa | agtaatatca | tgcgtaaatc | gtatgtgaat | gctggtcgct  | 8700 |
| atactgctgt | cgattcgata | ctaacgccgc | catccagtgt | cgaattcgcc | attcaggctg  | 8760 |
| cgcaactgtt | gggaagggcg | atcgggtcgg | gcctcttcgc | tattacgcca | gctgaattgg  | 8820 |
| agcgacctca | tgctatacct | gagaaagcaa | cctgacctac | aggaaagagt | tactcaagaa  | 8880 |
| taagaatttt | cgttttaaaa | cctaagagtc | actttaaaat | ttgtatacac | ttatTTTTTT  | 8940 |
| tataacttat | ttaataataa | aatcataaaa | tcataagaaa | ttcgcttatt | tagaagtgtc  | 9000 |
| aacaacgtat | ctaccaacga | tttgaccctt | ttccatcttt | tcgtaaattt | ctggcaaggt  | 9060 |
| agacaagccg | acaaccttga | ttggagactt | gaccaaacct | ctggcgaaga | attgttaatt  | 9120 |
| aagagctcag | atcttttgcg | gccgc      |            |            |             | 9145 |

<210> 77  
 <211> 499  
 <212> DNA  
 <213> *Saccharomyces cerevisiae*

|             |            |            |
|-------------|------------|------------|
| <400>       | 77         |            |
| gtttgcaaaa  | agaacaaaac | tgaaaaaacc |
| cagacacgct  | cgacttcctg | tcttcctatt |
|             |            | 60         |
| gattgcagct  | tccaatttcg | tcacacaaca |
| aggctcctagc | gacggctcac | aggttttgta |
|             |            | 120        |
| acaagcaatc  | gaaggttctg | gaatggcggg |
| aaagggttta  | gtaccacatg | ctatgatgcc |
|             |            | 180        |
| cactgtgatc  | tccagagcaa | agttcgttcg |
| atcgtactgt  | tactctctct | ctttcaaaca |
|             |            | 240        |
| gaattgtccg  | aatcgtgtga | caacaacagc |
| ctgttctcac  | acactctttt | cttctaacca |
|             |            | 300        |
| agggggtggt  | ttagtttagt | agaacctcgt |
| gaaacttaca  | tttacaataa | tataaacttg |
|             |            | 360        |
| cataaattgg  | tcaatgcaag | aaatacatat |
| ttggctcttt  | ctaattcgta | gtttttcaag |
|             |            | 420        |
| ttcttagatg  | ctttcttttt | ctctttttta |
| cagatcatca  | aggaagtaat | tatctacttt |
|             |            | 480        |
| ttacaacaaa  | tataaaaca  |            |
|             |            | 499        |

<210> 78  
 <211> 500  
 <212> DNA

<213> Saccharomyces cerevisiae

<400> 78

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| acagtttatt cctggcatcc actaaatata atggagcccg ctttttaagc tggcatccag | 60  |
| aaaaaaaaag aatcccagca ccaaaatatt gttttcttca ccaaccatca gttcataggt | 120 |
| ccattctctt agcgcaacta cagagaacag gggcacaaac aggcaaaaaa cgggcacaac | 180 |
| ctcaatggag tgatgcaacc tgcttgaggt aaatgatgac acaaggcaat tgaccacgc  | 240 |
| atgtatctat ctcatcttct tacaccttct attaccttct gctctctctg atttgaaaa  | 300 |
| agctgaaaaa aaagggtgaa accagttccc tgaaattatt cccctacttg actaataagt | 360 |
| atataaagac ggtaggtatt gattgtaatt ctgtaaatct atttcttaaa cttcttaaat | 420 |
| tctactttta tagttagtct ttttttagt tttaaaacac caagaactta gtttcgaata  | 480 |
| aacacacata aacaaacaaa                                             | 500 |

<210> 79

<211> 500

<212> DNA

<213> Saccharomyces cerevisiae

<400> 79

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| gctggattga gctgaatggt gccaggtcga ggctgggagg gagactaact cgaaagtgac  | 60  |
| gaagactcga aaattaaaaa aaaagatact gcagaaggca agattgagaa tggagttaaag | 120 |
| gcagcgtggg tccccgtgg aaaccgcagt tttcctgcgc caagtggtag cggtgcgagt   | 180 |
| gcagcaatta atctctcgat attttcttag tatctctttt tatataagaa tatattttgg  | 240 |
| aattggtaat gcttatcttc aatagtttct tagttgaatg cacacttaag agcaaattgg  | 300 |
| ccaaggagtt cttcgttcgc tttaatat ttcctggta ttgtcaattt attcatcca      | 360 |
| tctccccagg atagaagaaa ttagtgtaat ttgctgaca atacatttta acgacgataa   | 420 |
| caataatagc aattaaataa aatagcacta ccaccactcc actgctcgtt agctatttct  | 480 |
| gtaaaataaa taaaagatc                                               | 500 |

<210> 80

<211> 501

<212> DNA

<213> Saccharomyces cerevisiae

<400> 80

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| attcgcgcta tctcgatttc tacctatata gttaatctct gtacaaaaac aatctttcca  | 60  |
| actatccatt aatcatagta tattatcagc gtcggcgatt ttaccacgct tgacaaaagc  | 120 |
| cgcgggcggg attcctgtgg gtagtggcac cggcagttaa tctaatacaa ggcgcttgaa  | 180 |
| ggaagagata gataatagaa caaagcaatc gccgctttgg acggcaaata tgtttatcca  | 240 |
| ttggtgcggt gattggatat gatttgtctc cagtagtata agcaagcgc agatctgttt   | 300 |
| actgtaaaat taagttagta atctcgcggg atgtaatgat ttaagggaat ctggttcagg  | 360 |
| ttttcacata tatttgtata taaggccatt tgtaatttca atagtttttag gatttttcct | 420 |
| tctcccaaaa tactcactta ctgtgttaca ttacagaaag aacagacaag aaaccgtcaa  | 480 |

taagaaatat aactaagaac a

501

<210> 81  
<211> 27  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Synthetic HA epitope tag polypeptide

<400> 81  
taccatacgc acgttccaga ctacgcc

27

<210> 82  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Synthetic HA epitope tag polypeptide

<400> 82

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala  
1 5

<210> 83  
<211> 33  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Synthetic c-MYC epitope tag polypeptide

<400> 83  
gaacaaaagt taatttctga agaagatttg gaa

33

<210> 84  
<211> 10  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Synthetic c-MYC epitope tag polypeptide

<400> 84

Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu  
1 5 10

<210> 85  
<211> 24  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Synthetic FLAG epitope tag polypeptide

<400> 85  
gactacaagg atgacgatga caaa

24

<210> 86  
<211> 8



<212> PRT  
<213> Artificial Sequence

<220>  
<223> Synthetic FLAG epitope tag polypeptide

<400> 86

Asp Tyr Lys Asp Asp Asp Asp Lys  
1 5

<210> 87  
<211> 42  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Synthetic V5 epitope tag polypeptide

<400> 87  
ggtaagccaa ttccaaatcc tttgttgggt ttggactcca cc

42

<210> 88  
<211> 14  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Synthetic V5 epitope tag polypeptide

<400> 88

Gly Lys Pro Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr  
1 5 10

<210> 89  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Synthetic HIS epitope tag polypeptide

<400> 89  
catcatcatc atcatcat

18

<210> 90  
<211> 6  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Synthetic HIS epitope tag polypeptide

<400> 90

His His His His His His  
1 5

<210> 91  
<211> 7882  
<212> DNA  
<213> Artificial Sequence

<220>

<223> pMM1 vector

<400> 91

|             |            |            |            |             |             |      |
|-------------|------------|------------|------------|-------------|-------------|------|
| cagggtggcac | ttttcgggga | aatgtgcgcg | gaacccttat | ttgtttat    | ttctaaatac  | 60   |
| attcaaatat  | gtatccgctc | atgagacaat | aaccctgata | aatgcttcaa  | taatattgaa  | 120  |
| aaaggaagag  | tatgagtatt | caacatttcc | gtgtcgccct | tattcccttt  | tttgcggcat  | 180  |
| tttgccttcc  | tgTTTTgtct | caccagaaa  | cgctggtgaa | agtaaaagat  | gctgaagatc  | 240  |
| agttgggtgc  | acgagtgggt | tacatcgaac | tggatctcaa | cagcggtaag  | atccttgaga  | 300  |
| gttttcgccc  | cgaagaacgt | tttccaatga | tgagcacttt | taaagttctg  | ctatgtggcg  | 360  |
| cggattatc   | ccgtattgac | gccgggcaag | agcaactcgg | tcgccgcata  | cactattctc  | 420  |
| agaatgactt  | ggttgagtac | tcaccagtca | cagaaaagca | tcttacggat  | ggcatgacag  | 480  |
| taagagaatt  | atgcagtgtc | gccataacca | tgagtataa  | actgcggcc   | aacttacttc  | 540  |
| tgacaacgat  | cggaggaccg | aaggagctaa | ccgctttttt | gcacaacatg  | ggggatcatg  | 600  |
| taactcgctt  | tgatcgttgg | gaaccggagc | tgaatgaagc | cataccaaac  | gacgagcgtg  | 660  |
| acaccacgat  | gcctgtagca | atggcaacaa | cgttgcgcaa | actattaact  | ggcgaactac  | 720  |
| ttactctagc  | ttcccggcaa | caattaatag | actggatgga | ggcggataaa  | gttgcaggac  | 780  |
| cacttctgcg  | ctcggccctt | ccggctggct | ggttttattg | tgataaatct  | ggagccgggtg | 840  |
| agcgtgggtc  | tcgcggtatc | attgcagcac | tggggccaga | tggttaagccc | tcccgtatcg  | 900  |
| tagttatcta  | cacgacgggg | agtcaggcaa | ctatggatga | acgaaataga  | cagatcgctg  | 960  |
| agatagggtc  | ctcactgatt | aagcattggg | aactgtcaga | ccaagtttac  | tcatatatac  | 1020 |
| tttagattga  | tttaaaactt | catttttaat | ttaaaaggat | ctaggtgaag  | atcctttttg  | 1080 |
| ataatctcat  | gaccaaatac | ccttaacgtg | agttttcgtt | ccactgagcg  | tcagaccccg  | 1140 |
| tagaaaagat  | caaaggatct | tcttgagatc | ctttttttct | gcgcgtaatc  | tgctgcttgc  | 1200 |
| aaacaaaaaa  | accaccgcta | ccagcgggtg | tttgtttgcc | ggatcaagag  | ctaccaactc  | 1260 |
| tttttccgaa  | ggtaactggc | ttcagcagag | cgcagatacc | aaatactgtt  | cttctagtgt  | 1320 |
| agccgtagtt  | aggccaccac | ttcaagaact | ctgtagcacc | gcctacatac  | ctcgctctgc  | 1380 |
| taatcctgtt  | accagtggct | gctgccagt  | gcgataagtc | gtgtcttacc  | gggttggact  | 1440 |
| caagacgata  | gttaccggat | aaggcgcagc | ggtcgggctg | aacggggggg  | tcgtgcacac  | 1500 |
| agcccagctt  | ggagcgaacg | acctacaccg | aactgagata | cctacagcgt  | gagctatgag  | 1560 |
| aaagcgccac  | gcttcccga  | gggagaaagg | cggacaggta | tccggtaagc  | ggcagggtcg  | 1620 |
| gaacaggaga  | gcgcacgagg | gagcttccag | ggggaaacgc | ctggatatct  | tatagtcctg  | 1680 |
| tcgggtttcg  | ccacctctga | cttgagcgtc | gattttttgt | atgctcgtca  | ggggggcgga  | 1740 |
| gcctatggaa  | aaacgccagc | aacgcggcct | ttttacgggt | cctggccttt  | tgctggcctt  | 1800 |
| ttgctcacat  | gttctttcct | gcgttatccc | ctgattctgt | ggataaccgt  | attaccgcct  | 1860 |
| ttgagtgagc  | tgataccgct | cgccgcagcc | gaacgaccga | gcgcagcgag  | tcagtgagcg  | 1920 |
| aggaagcgga  | agagcgccca | atacgcaaac | cgcctctccc | cgcgcgttgg  | ccgattcatt  | 1980 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| aatgcagctg gatcttcgag cgtcccaaaa ccttctcaag caaggttttc agtataatgt  | 2040 |
| tacatgcgta cacgcgtctg tacagaaaaa aaagaaaaat ttgaaatata aataacgttc  | 2100 |
| ttaataactaa cataactata aaaaaataaa tagggaccta gacttcaggt tgtctaactc | 2160 |
| cttccttttc ggtagagcg gatcttagct agccgcggta ccaagctggg ggatcctttg   | 2220 |
| tttgtttatg tgtgtttatt cgaaactaag ttcttggtgt tttaaaacta aaaaaagac   | 2280 |
| taactataaa agtagaattt aagaagttta agaaatagat ttacagaatt acaatcaata  | 2340 |
| cctaccgtct ttatatactt attagtcaag taggggaata atttcaggga actggtttca  | 2400 |
| accttttttt tcagcttttt ccaaatacaga gagagcagaa ggtaatagaa ggtgtaagaa | 2460 |
| aatgagatag atacatgcgt gggtaattg ccttgtgtca tcatttactc caggcagggt   | 2520 |
| gcatcactcc attgaggttg tgcccgtttt ttgcctgttt gtgcccctgt tctctgtagt  | 2580 |
| tgcgctaaga gaatggacct atgaactgat ggttggtgaa gaaaacaata ttttggtgct  | 2640 |
| gggattcttt ttttttctgg atgccagctt aaaaagcggg ctccattata tttagtggat  | 2700 |
| gccaggaata aactgttggt tgcaaaaaga acaaaactga aaaaaccag acacgctcga   | 2760 |
| cttcctgtct tcctattgat tgcagcttcc aatttcgtca cacaacaagg tcctagcgac  | 2820 |
| ggctcacagg ttttgtaaca agcaatcgaa ggttctggaa tggcgggaaa gggtttagta  | 2880 |
| ccacatgcta tgatgccac tgtgatctcc agagcaaagt tcgttcgatc gtactgttac   | 2940 |
| tctctctctt tcaaacagaa ttgtccgaat cgtgtgacaa caacagcctg ttctcacaca  | 3000 |
| ctcttttctt ctaaccaagg gggtggttta gtttagtaga acctcgtgaa acttacattt  | 3060 |
| acatatatat aaacttgcac aaattggtca atgcaagaaa tacatatattg gtcttttcta | 3120 |
| attcgtagtt tttcaagttc ttagatgctt tctttttctc ttttttacag atcatcaagg  | 3180 |
| aagtaattat ctacttttta caacaaatat aaaacagcgg ccgcaaaaaga tctgagctct | 3240 |
| taattaacaa ttcttcgcca gaggtttggt caagtctcca atcaaggttg tcggcttgtc  | 3300 |
| taccttgcca gaaatttacg aaaagatgga aaagggtcaa atcgttggta gatacgttgt  | 3360 |
| tgacacttct aaataagcga atttcttatg atttatgatt tttattatta aataagttat  | 3420 |
| aaaaaaaaata agtgtataca aattttaaag tgactcttag gttttaaaac gaaaattctt | 3480 |
| attcttgagt aactcttttc ttaggtcag gttgctttct caggatatagc atgaggctgc  | 3540 |
| tccaattcag ctggcgtaat agcgaagagg cccgcaccga tcgcccttcc caacagttgc  | 3600 |
| gcagcctgaa tggcgaattc gacactggat ggcggcggtta gtatcgaatc gacagcagta | 3660 |
| tagcgaccag cattcacata cgattgacgc atgatattac tttctgcgca cttaacttcg  | 3720 |
| catctgggca gatgatgtcg aggcgaaaaa aaatataaat cagcctaaca tttgattaaa  | 3780 |
| atagaacaac tacaatataa aaaaactata caaatgacaa gttcttgaaa acaagaatct  | 3840 |
| ttttattgtc agtactgatt agaaaaactc atcgagcatc aaatgaaact gcaatttatt  | 3900 |
| catatcagga ttatcaatac catatttttg aaaaagccgt ttctgtaatg aaggagaaaa  | 3960 |
| ctcaccgagg cagttccata ggatggcaag atcctggtat cggctctcga ttccgactcg  | 4020 |

|                                                                     |      |
|---------------------------------------------------------------------|------|
| tccaacatca atacaaccta ttaatttccc ctctgcaaaa ataaggttat caagtgagaa   | 4080 |
| atcaccatga gtgacgactg aatccggtga gaatggcaaa agcttatgca tttctttcca   | 4140 |
| gacttgttca acaggccagc cattacgctc gtcatacaaaa tcaactcgcat caaccaaacc | 4200 |
| gttattcatt cgtgattgcg cctgagcgag acgaaatacg cgatcgctgt taaaaggaca   | 4260 |
| attacaaaca ggaatcgaat gcaaccggcg caggaacact gccagcgcat caacaatatt   | 4320 |
| ttcacctgaa tcaggatatt cttctaatac ctggaatgct gttttgccgg ggatcgcagt   | 4380 |
| ggtgagtaac catgcatcat caggagtacg gataaaatgc ttgatggtcg gaagaggcat   | 4440 |
| aaattccgct agccagttta gtctgacat ctcatctgta acatcattgg caacgctacc    | 4500 |
| tttgccatgt ttcagaaaca actctggcgc atcgggcttc ccatacaatc gatagattgt   | 4560 |
| cgcacctgat tgcccagacat tatcgcgagc ccatttatac ccatataaat cagcatccat  | 4620 |
| gttggaattt aatcgcggcc tcgaaacgtg agtcttttcc ttacccatgg ttgtttatgt   | 4680 |
| tcggatgtga tgtgagaact gtatcctagc aagattttta aaggaagtat atgaaagaag   | 4740 |
| aacctcagtg gcaaatccta accttttata tttctctaca ggggcgcggc gtggggacaa   | 4800 |
| ttcaacgcgt ctgtgagggg agcgtttccc tgctcgcagg tctgcagcga ggagccgtaa   | 4860 |
| tttttgcttc gcgccgtgcg gccatcaaaa tgtatggatg caaatgatta tacatgggga   | 4920 |
| tgtatgggct aaatgtacgg gcgacagtca catcatgccc ctgagctgcg cacgtcaaga   | 4980 |
| ctgtcaagga gggatttctg ggctccatg tcgctggccg ggtgaccgg cggggacgag     | 5040 |
| gcaagctaaa cagatctggc gcgtgtactg agagtgcacc ataaattccc gttttaagag   | 5100 |
| cttggtgagc gctaggagtc actgccaggt atcgtttgaa cacggcatta gtcagggaag   | 5160 |
| tcataacaca gtcctttccc gcaattttct ttttctatta ctcttgccct cctctagtac   | 5220 |
| actctatatt ttttatgcc tcggtaatga ttttcatttt tttttttcca cctagcggat    | 5280 |
| gactcttttt ttttcttagc gattggcatt atcacataat gaattataca ttatataaag   | 5340 |
| taatgtgatt tcttgaaga atatactaaa aaatgagcag gcaagataaa cgaaggcaaa    | 5400 |
| gatgacagag cagaaagccc tagtaaagcg tattacaaat gaaaccaaga ttcagattgc   | 5460 |
| gatctcttta aagggtggtc ccctagcgat agagcactcg atcttccag aaaaagaggc    | 5520 |
| agaagcagta gcagaacagg ccacacaatc gcaagtgatt aacgtccaca caggtatagg   | 5580 |
| gtttctggac catatgatac atgctctggc caagcattcc ggctggctgc taatcgttga   | 5640 |
| gtgcattggt gacttacaca tagacgacca tcacaccact gaagactgcg ggattgctct   | 5700 |
| cggccaagct tttaaagagg ccctaggggc cgtgcgtgga gtaaaaaggt ttggatcagg   | 5760 |
| atttgcgcct ttggatgagg cactttccag agcggtggtg gatctttcga acaggccgta   | 5820 |
| cgcagttgtc gaacttggtt tgcaaaggga gaaagtagga gatctctctt gcgagatgat   | 5880 |
| cccgcathtt cttgaaagct ttgcagaggc tagcagaatt accctccacg ttgattgtct   | 5940 |
| gcgaggcaag aatgatcatc accgtagtga gagtgcgttc aaggctcttg cggttgccat   | 6000 |
| aagagaagcc acctgcacca atggtaccaa cgatgttccc tccaccaaag gtgttcttat   | 6060 |
| gtagtgacac cgattattta aagctgcagc atacgatata tatacatgtg tatatatgta   | 6120 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| tacctatgaa tgtcagtaag tatgtatacg aacagtatga tactgaagat gacaaggtaa  | 6180 |
| tgcatcattc tatacgtgtc attctgaacg aggcgcgctt tccttttttc tttttgcttt  | 6240 |
| ttcttttttt ttctcttgaa ctgcacggat ctatgcggtg tgaaatatgg tgcactctca  | 6300 |
| gtacaatctg ctctgatgcc gcatagttaa gccagcccc acacccgcca acacccgctg   | 6360 |
| acgcgccctg acgggcttgt ctgctcccgg catccgctta cagacaagct gtgaccgtct  | 6420 |
| ccgggagctg catgtgtcag aggttttcac cgtcatcacc gaaacgcgcg agacgaaagg  | 6480 |
| gcctcgtgat acgcctattt ttataggtta atgtcatgat aataatggtt tcttagtatg  | 6540 |
| atccaatatc aaaggaaatg atagcattga aggatgagac taatccaatt gaggagtggc  | 6600 |
| agcatataga acagctaaag ggtagtgctg aaggaagcat acgatacccc gcatggaatg  | 6660 |
| ggataatatc acaggaggta ctgactacc tttcatccta cataaataga cgcatataag   | 6720 |
| tacgcattta agcataaaca cgcactatgc cgttcttctc atgtatatat atatacaggc  | 6780 |
| aacacgcaga tatagggtgcg acgtgaacag tgagctgtat gtgcgcagct cgcgttgcat | 6840 |
| tttcggaagc gtcgttttc ggaaacgctt tgaagttcct attccgaagt tcctattctc   | 6900 |
| tagaaagtat aggaacttca gagcgctttt gaaaaccaa agcgctctga agacgcactt   | 6960 |
| tcaaaaaacc aaaaacgcac cggactgtaa cgagctacta aaatattgcg aataccgctt  | 7020 |
| ccacaaacat tgctcaaaag tatctctttg ctatatatct ctgtgctata tccctatata  | 7080 |
| acctacccat ccacctttcg ctcttgaac ttgcatctaa actcgacctc tacatttttt   | 7140 |
| atgtttatct ctagtattac tctttagaca aaaaaattgt agtaagaact attcatagag  | 7200 |
| tgaatcgaaa acaatacgaa aatgtaaaca tttcctatac gtagtatata gagacaaaat  | 7260 |
| agaagaaacc gttcataatt ttctgaccaa tgaagaatca tcaacgctat cactttctgt  | 7320 |
| tcacaaagta tgcgcaatcc acatcggtat agaataaat cggggatgcc tttatcttga   | 7380 |
| aaaaatgcac ccgcagcttc gctagtaatc agtaaacgcg ggaagtggag tcaggctttt  | 7440 |
| tttatggaag agaaaataga caccaaagta gccttcttct aaccttaacg gacctacagt  | 7500 |
| gcaaaaagtt atcaagagac tgcatatatag agcgcacaaa ggagaaaaaa agtaatctaa | 7560 |
| gatgctttgt tagaaaaata gcgctctcgg gatgcatttt tgtagaacia aaaagaagta  | 7620 |
| tagattcttt gttggtaaaa tagcgctctc gcgttgcat tctgttctgt aaaaatgcag   | 7680 |
| ctcagattct ttgtttgaaa aattagcgct ctgcggttg atttttgttt tacaaaaatg   | 7740 |
| aagcacagat tcttcgttgg taaaatagcg ctttcgcgtt gcatttctgt tctgtaaaaa  | 7800 |
| tgcagctcag attctttgtt tgaaaaatta gcgctctcgc gttgcatttt tgttctacaa  | 7860 |
| aatgaagcac agatgcttcg tt                                           | 7882 |

<210> 92

<211> 3810

<212> DNA

<213> Artificial Sequence

<220>

<223> pCDM4 vector

<400> 92

|                                                                   |      |
|-------------------------------------------------------------------|------|
| gctaacagcg cgatttgctg gtgacccaat gcgaccagat gctccacgcc cagtcgcgta | 60   |
| ccgtcttcat gggagaaaat aatactgttg atgggtgtct ggtagagac atcaagaaat  | 120  |
| aacgccggaa cattagtga ggcagcttcc acagcaatgg catcctggtc atccagcgga  | 180  |
| tagttaatga tcagcccact gacgcgttgc gcgagaagat tgtgcaccgc cgctttacag | 240  |
| gcttcgacgc cgcttcgttc taccatcgac accaccacgc tggcaccag ttgatcggcg  | 300  |
| cgagatttaa tcgccgcgac aatttgcgac ggcgcgtgca gggccagact ggaggtggca | 360  |
| acgccaatca gcaacgactg tttgcccgcc agttgttgtg ccacgcggtt gggaatgtaa | 420  |
| ttcagctccg ccatcgccgc ttccactttt tcccgcgttt tcgcagaaac gtggctggcc | 480  |
| tggttcacca cgcgggaaac ggtctgataa gagacaccgg catactctgc gacatcgtat | 540  |
| aacgttactg gtttcacatt caccaccctg aattgactct cttccgggcg ctatcatgcc | 600  |
| ataccgcgaa aggttttgcg ccattcgatg gtgtccggga tctcgacgct ctcccttatg | 660  |
| cgactcctgc attaggaagc agcccagtag taggttgagg ccgttgagca ccgccccgc  | 720  |
| aaggaatggt gcatgcaagg agatggcgcc caacagtccc ccggccacgg ggcctgccac | 780  |
| catacccacg ccgaaacaag cgctcatgag cccgaagtgg cgagcccgat cttcccatc  | 840  |
| ggtgatgtcg gcgatatagg gccagcaac cgcacctgtg gcgccggtga tgccggccac  | 900  |
| gatgcgtccg gcgtagccta ggatcgagat cgatctcgat cccgcgaaat taatacgact | 960  |
| cactataggg gaattgtgag cggataacaa ttcccctcta gaaataattt tgtttaactt | 1020 |
| taagaaggag atatacatat ggcagatctc aattggatat cggccggcca cgcgatcgct | 1080 |
| gacgtcggta ccctcgagtc tggtaaagaa accgctgctg cgaaatttga acgccagcac | 1140 |
| atggactcgt ctactagtcg cagcttaatt aacctaaact gctgccaccg ctgagcaata | 1200 |
| actagcataa ccccttgggg cctctaaacg ggtcttgagg ggttttttgc tagcgaaagg | 1260 |
| aggagtcgac actgcttccg gtagtcaata aaccggtaaa ccagcaatag acataagcgg | 1320 |
| ctatttaacg accctgccct gaaccgacga ccgggtcatc gtggccggat cttgcggccc | 1380 |
| ctcggcttga acgaattgtt agacattatt tgccgactac cttggtgatc tcgcctttca | 1440 |
| cgtagtggac aaattcttcc aactgatctg cgcgcgaggc caagcgatct tcttcttgtc | 1500 |
| caagataagc ctgtctagct tcaagtatga cgggctgata ctgggcccgc aggcgctcca | 1560 |
| ttgcccagtc ggcagcgaca tccttcggcg cgattttgcc ggttactgcg ctgtacaaa  | 1620 |
| tgcgggacaa cgtaagcact acatttcgct catcgccagc ccagtcgggc ggcgagttcc | 1680 |
| atagcgtaa ggtttcattt agcgcctcaa atagatcctg ttcaggaacc ggatcaaaga  | 1740 |
| gttcctccgc cgctggacct accaaggcaa cgctatgttc tcttgctttt gtcagcaaga | 1800 |
| tagccagatc aatgtcgatc gtggctggct cgaagatacc tgcaagaatg tcattgcgct | 1860 |
| gccatttctc aaattgcagt tcgcgcttag ctggataacg ccacggaatg atgtcgtcgt | 1920 |
| gcacaacaat ggtgacttct acagcgcgga gaatctcgct ctctccaggg gaagccgaag | 1980 |
| tttccaaaag gtcgttgatc aaagctcgcc gcgttgtttc atcaagcctt acggtcaccg | 2040 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| taaccagcaa atcaatatca ctgtgtggct tcaggccgcc atccactgcg gagccgtaca  | 2100 |
| aatgtacggc cagcaacgtc ggttcgagat ggcgctcgat gacgccaact acctctgata  | 2160 |
| gttgagtcga tacttcggcg atcaccgctt ccctcatact cttccttttt caatattatt  | 2220 |
| gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa  | 2280 |
| ataaacaat agccagctca ctcggtcgct acgctccggg cgtgagactg cggcgggcgc   | 2340 |
| tgcggacaca taaaaagtta cccacagatt ccgtggataa gcaggggact aacatgtgag  | 2400 |
| gcaaaacagc agggccgcgc cggtagcggtt tttccatagg ctccgccctc ctgccagagt | 2460 |
| tcacataaac agacgctttt ccggtgcatc tgtgggagcc gtgaggctca accatgaatc  | 2520 |
| tgacagtacg ggcgaaaccc gacaggactt aaagatcccc accgtttccg gcgggtcgct  | 2580 |
| ccctcttgcg ctctcctggt ccgaccctgc cgtttaccgg atacctgttc cgcctttctc  | 2640 |
| ccttacggga agtgtagcgc tttctcatag ctcacacact ggtatctcgg ctcggtgtag  | 2700 |
| gtcgttcgct ccaagctggg ctgtaagcaa gaactcccc tttagccga ctgctgcgcc    | 2760 |
| ttatccggta actgttcaact tgagtccaac ccgaaaagc acggtaaaac gccactggca  | 2820 |
| gcagccattg gtaactggga gttcgagag gatttgttta gctaaacacg cggttgctct   | 2880 |
| tgaagtgtgc gccaaagtcc ggctacactg gaaggacaga tttggttgct gtgctctgcg  | 2940 |
| aaagccagtt accacggtta agcagttccc caactgactt aaccttcgat caaaccacct  | 3000 |
| ccccaggtagg ttttttcgtt tacagggcaa aagattacgc gcagaaaaaa aggatctcaa | 3060 |
| gaagatcctt tgatcttttc tactgaaccg ctctagattt cagtgcatt tatctcttca   | 3120 |
| aatgtagcac ctgaagtcag ccccatacga tataagttgt aattctcatg ttagtcatgc  | 3180 |
| ccgcgcccc ccggaaggag ctgactgggt tgaaggctct caagggcacg ggtcgagatc   | 3240 |
| ccggtgccta atgagtgagc taacttacat taattgcgtt gcgctcactg cccgctttcc  | 3300 |
| agtcgggaaa cctgtcgtgc cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg  | 3360 |
| gtttgcgtat tgggcgccag ggtggttttt cttttcacca gtgagacggg caacagctga  | 3420 |
| ttgcccttca ccgcctggcc ctgagagagt tgcagcaagc ggtccacgct ggtttgcccc  | 3480 |
| agcaggcgaa aatcctgttt gatggtgggt aacggcggga tataacatga gctgtcttcg  | 3540 |
| gtatcgtcgt atcccactac cgagatgtcc gcaccaacgc gcagcccga ctcggtaatg   | 3600 |
| gcgcgcattg cggccagcgc catctgatc ttggcaacca gcatcgagt gggaacgatg    | 3660 |
| ccctcattca gcatttgcgt ggtttgttga aaaccggaca tggcactcca gtcgccttcc  | 3720 |
| cgttccgcta tcggctgaat ttgattgcga gtgagatatt tatgccagcc agccagacgc  | 3780 |
| agacgcgccg agacagaact taatgggccc                                   | 3810 |

<210> 93

<211> 5203

<212> DNA

<213> Artificial Sequence

<220>

<223> pETM6 vector

<400> 93

|                                                                   |      |
|-------------------------------------------------------------------|------|
| ggggaattgt gagcggataa caattcccct ctagaataa ttttgtttaa ctttaagaag  | 60   |
| gagatataca tatggcagat ctcaattgga tatcggccgg ccacgcgac gctgacgtcg  | 120  |
| gtaccctcga gtctggtaaa gaaaccgctg ctgcgaaatt tgaacgccag cacatggact | 180  |
| cgtctactag tcgcagctta attaacctaa actgctgcc cgcctgagca ataactagca  | 240  |
| taacccttgg gggcctctaa acgggtcttg aggggttttt tgctagcgaa aggaggagtc | 300  |
| gactatatcc ggattggcga atgggacgcg ccctgtagcg gcgcattaag cgcggcgggt | 360  |
| gtggtgggta cgcgacgcgt gaccgctaca cttgccagcg ccctagcgcc cgctcctttc | 420  |
| gctttcttcc cttcctttct cgccacgttc gccggctttc cccgtcaagc tctaaatcgg | 480  |
| gggctccctt tagggttccg atttagtgct ttacggcacc tcgaccccaa aaaacttgat | 540  |
| tagggtagtg gttcacgtag tgggcatcgc ccctgataga cggtttttcg ccctttgacg | 600  |
| ttggagtcca cgttctttaa tagtggactc ttgttccaaa ctggaacaac actcaaccct | 660  |
| atctcggctc attcttttga tttataaggg attttgccga tttcggccta ttggttaaaa | 720  |
| aatgagctga ttttaacaaa atttaacgcg aattttaaca aaatattaac gtttacaatt | 780  |
| tctggcggca cgatggcatg agattatcaa aaaggatctt cacctagatc cttttaaatt | 840  |
| aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc | 900  |
| aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg | 960  |
| cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct ggccccagtg | 1020 |
| ctgcaatgat accgcgagac ccacgctcac cggtccaga tttatcagca ataaaccagc  | 1080 |
| cagccggaag ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc atccagtcta | 1140 |
| ttaattgttg ccgggaagct agagtaagta gttcgccagt taatagtttg cgcaacgttg | 1200 |
| ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt tggtagggct tcattcagct | 1260 |
| ccggttccca acgatcaagg cgagttacat gatccccat gttgtgcaaa aaagcgggta  | 1320 |
| gctccttcgg tcctccgacg gttgtcagaa gtaagttggc cgcagtgta tcaactcatg  | 1380 |
| ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc ttttctgtga | 1440 |
| ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt | 1500 |
| gccccgcgtc aatacgggat aataccgcgc cacatagcag aactttaaaa gtgctcatca | 1560 |
| ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg agatccagtt | 1620 |
| cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc accagcgttt | 1680 |
| ctgggtgagc aaaaacagga aggcaaatg ccgcaaaaaa gggaataagg gcgacacgga  | 1740 |
| aatgttgaat actcactatc ttcctttttc aatcatgatt gaagcattta tcagggttat | 1800 |
| tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggtcatgac | 1860 |
| caaaatccct taacgtgagt tttcgttcca ctgagcgtca gacccgtag aaaagatcaa  | 1920 |
| aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc | 1980 |
| accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt ttccgaaggt | 2040 |



|                                                                    |      |
|--------------------------------------------------------------------|------|
| aactggcttc agcagagcgc agataccaaa tactgtcctt ctagtgtagc cgtagttagg  | 2100 |
| ccaccacttc aagaactctg tagcaccgcc tacatacctc gctctgctaa tcctgttacc  | 2160 |
| agtggctgct gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatatgt  | 2220 |
| accggataag gcgcagcggg cgggctgaac ggggggttcg tgcacacagc ccagcttgga  | 2280 |
| gcgaacgacc tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct  | 2340 |
| tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa caggagagcg  | 2400 |
| cacgagggag cttccagggg gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca  | 2460 |
| cctctgactt gagcgtcgat ttttgtgatg ctcgctcaggg gggcggagcc tatggaaaaa | 2520 |
| cgccagcaac gcggcctttt tacggttcct ggcccttttg tggccttttg ctcacatgtt  | 2580 |
| ctttcctgcg ttatcccctg attctgtgga taaccgtatt accgcctttg agtgagctga  | 2640 |
| taccgctcgc cgagccgaa cgaccgagcg cagcgagtca gtgagcgagg aagcgggaaga  | 2700 |
| gcgcctgatg cggtatTTTT tccttacgca tctgtgcggg atttcacacc gcatatatgg  | 2760 |
| tgcactctca gtacaatctg ctctgatgcc gcatagttaa gccagtatac actccgctat  | 2820 |
| cgctacgtga ctgggtcatg gctgcgcccc gacacccgcc aacacccgct gacgcgccct  | 2880 |
| gacgggcttg tctgctcccg gcatccgctt acagacaagc tgtgaccgtc tccgggagct  | 2940 |
| gcatgtgtca gaggttttca ccgtcatcac cgaaacgcgc gaggcagctg cggtaaagct  | 3000 |
| catcagcgtg gtcgtgaagc gattcacaga tgtctgcctg ttcattccgcg tccagctcgt | 3060 |
| tgagtttctc cagaagcgtt aatgtctggc ttctgataaa gcgggccatg ttaagggcgg  | 3120 |
| ttttttcctg tttggtcact gatgcctccg tgtaaggggg atttctgttc atgggggtaa  | 3180 |
| tgataccgat gaaacgagag aggatgtcga cgatacgggt tactgatgat gaacatgccc  | 3240 |
| ggttactgga acgttgtgag ggtaaacaac tggcgggtatg gatgcggcgg gaccagagaa | 3300 |
| aaatcactca gggccaatgc cagcgcttcg ttaatacaga tgtaggtgtt ccacagggta  | 3360 |
| gccagcagca tcctgcgatg cagatccgga acataatggt gcagggcgct gacttccgcg  | 3420 |
| ttccagact ttacgaaaca cggaaccga agaccattca tgttgttgct caggtcgcag    | 3480 |
| acgttttgca gcagcagtcg cttcacgttc gctcgcgtat cggtgattca ttctgctaac  | 3540 |
| cagtaaggca accccgccag cctagccggg tcctcaacga caggagcacg atcatgctag  | 3600 |
| tcatgccccg cgcccaccgg aaggagctga ctgggttgaa ggctctcaag ggcacggtc   | 3660 |
| gagatcccgg tgcctaata gaagagctaac ttacattaat tgcgttgcg tcaactgccc   | 3720 |
| ctttccagtc gggaaacctg tcgtgccagc tgcattaatg aatcgccaa cgcgcgggga   | 3780 |
| gaggcggttt gcgtattggg gccagggtg gtttttctt tcaccagtga gacgggcaac    | 3840 |
| agctgattgc cttcaccgc ctggccctga gagagttgca gcaagcggc cagctgggtt    | 3900 |
| tgccccagca ggcgaaaatc ctgtttgatg gtggttaacg gcgggatata acatgagctg  | 3960 |
| tcttcggtat cgtcgtatcc cactaccgag atgtccgcac caacgcgcag cccggactcg  | 4020 |
| gtaatggcgc gcattgcgc cagcgccatc tgatcgttgg caaccagcat cgagtgggga   | 4080 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| acgatgccct cattcagcat ttgcatggtt tgttgaaaac cggacatggc actccagtcg  | 4140 |
| ccttcccgtt ccgctatcgg ctgaatttga ttgcgagtga gatatttatg ccagccagcc  | 4200 |
| agacgcagac gcgccgagac agaacttaat gggcccgccta acagcgcgat ttgctggtga | 4260 |
| cccaatgcga ccagatgctc cacgcccagt cgcgtaccgt cttcatggga gaaaataata  | 4320 |
| ctgttgatgg gtgtctggtc agagacatca agaaataacg ccggaacatt agtgcaggca  | 4380 |
| gcttccacag caatggcatc ctgggtcatcc agcggatagt taatgatcag cccactgacg | 4440 |
| cgttgcgcgga gaagattgtg caccgccgct ttacaggctt cgacgccgct tcgttctacc | 4500 |
| atcgacacca ccacgtggc acccagttga tcggcgcgag atttaatcgc cgcgacaatt   | 4560 |
| tgcgacggcg cgtgcagggc cagactggag gtggcaacgc caatcagcaa cgactgtttg  | 4620 |
| cccgccagtt gttgtgccac gcggttggga atgtaattca gctccgcat cgccgcttcc   | 4680 |
| actttttccc gcgttttcgc agaaacgtgg ctggcctggt tcaccacgcg ggaaacggtc  | 4740 |
| tgataagaga caccggcata ctctgcgaca tcgtataacg ttactggttt cacattcacc  | 4800 |
| accctgaatt gactctcttc cgggcgctat catgccatac cgcgaaaggt tttgcgcat   | 4860 |
| tcgatggtgt ccgggatctc gacgctctcc cttatgcgac tcctgcatta ggaagcagcc  | 4920 |
| cagtagtagg ttgaggccgt tgagcaccgc cgccgcaagg aatggtgcat gcaaggagat  | 4980 |
| ggcgcccaac agtcccccg ccacggggcc tgccaccata cccacgccga aacaagcgct   | 5040 |
| catgagcccg aagtggcgag ccgatcttc cccatcggtg atgtcggcga tataggcgcc   | 5100 |
| agcaaccgca cctgtggcgc cggatgatgc ggccacgatg cgtccggcgt agcctaggat  | 5160 |
| cgagatcgat ctcgatcccg cgaaattaat acgactcact ata                    | 5203 |

<210> 94  
 <211> 5369  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> pET28a(+) vector

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| <400> 94                                                           |     |
| atccggatat agttcctcct ttcagcaaaa aaccctcaa gaccgttta gaggcccaa     | 60  |
| ggggttatgc tagttattgc tcagcggtagg cagcagccaa ctacgcttcc tttcgggctt | 120 |
| tgtagtagc cggatctcag tggtaggtggt ggtggtgctc gtagtcggcc gcaagcttgt  | 180 |
| cgacggagct cgaattcgga tccgcgacct atttgctgtc caccagtcac gtagccata   | 240 |
| tggctgccgc gcggcaccag gccgctgctg tgatgatgat gatgatggct gctgcccatg  | 300 |
| gtatatctcc ttcttaaagt taaacaaaat tatttctaga ggggaattgt tatccgctca  | 360 |
| caattcccct atagttagtc gtattaattt cgcgggatcg agatctcgat cctctacgcc  | 420 |
| ggacgcatcg tggccggcat caccggcgcc acaggtgcgg ttgctggcgc ctatatcgcc  | 480 |
| gacatcaccg atggggaaga tcgggctcgc cacttcgggc tcatgagcgc ttgtttcggc  | 540 |
| gtgggtatgg tggcaggccc cgtggccggg ggactgttgg gcgccatctc cttgcatgca  | 600 |
| ccattccttg cggcggcggt gctcaacggc ctcaacctac tactgggctg cttcctaattg | 660 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| caggagtcgc ataagggaga gcgtcgagat cccggacacc atcgaatggc gcaaaacctt  | 720  |
| tcgcggtatg gcatgatagc gcccggaaga gagggtggtga atgtgaaacc            | 780  |
| agtaacgtta tacgatgtcg cagagtatgc cggtgtctct tatcagaccg tttcccgcgt  | 840  |
| ggtgaaccag gccagccacg tttctgcgaa aacgcgggaa aaagtggaag cggcgatggc  | 900  |
| ggagctgaat tacattccca accgcgtggc acaacaactg gcggggcaaac agtcgttgct | 960  |
| gattggcggt gccacctcca gtctggccct gcacgcgccg tcgcaaattg tcgcggcgat  | 1020 |
| taaatctcgc gccgatcaac tgggtgccag cgtggtggtg tcgatggtag aacgaagcgg  | 1080 |
| cgtcgaagcc tgtaaagcgg cggcgcacaa tcttctcgcg caacgcgtca gtgggctgat  | 1140 |
| cattaactat ccgctggatg accaggatgc cattgctgtg gaagctgcct gcactaatgt  | 1200 |
| tccggcgtaa tttcttgatg tctctgacca gacacccatc aacagtatta ttttctccca  | 1260 |
| tgaagacggt acgcgactgg gcgtggagca tctggtcgca ttgggtcacc agcaaatacg  | 1320 |
| gctgttagcg ggcccattaa gttctgtctc ggcgcgtctg cgtctggctg gctggcataa  | 1380 |
| atatctcact cgcaatcaaa ttcagccgat agcggaaagg gaaggcgact ggagtgccat  | 1440 |
| gtccggtttt caacaaacca tgcaaatgct gaatgagggc atcgttccca ctgcgatgct  | 1500 |
| ggttgccaac gatcagatgg cgctgggcgc aatgcgcgcc attaccgagt ccgggctgcg  | 1560 |
| cgttggtgcg gatatctcgg tagtgggata cgacgatacc gaagacagct catgttatat  | 1620 |
| cccgcgtaa accaccatca aacaggattt tcgcctgctg gggcaaacca gcgtggaccg   | 1680 |
| cttgctgcaa ctctctcagg gccaggcggg gaagggaat cagctgttgc ccgtctcact   | 1740 |
| ggtgaaaaga aaaaccaccc tggcgcccaa tacgcaaacc gcctctcccc gcgcgttggc  | 1800 |
| cgattcatta atgcagctgg cacgacaggt ttcccgaactg gaaagcgggc agtgagcgca | 1860 |
| acgcaattaa tgtaagttag ctactcatt aggcaccggg atctcgaccg atgcccttga   | 1920 |
| gagccttcaa cccagtcagc tccttcgggt gggcgcgggg catgactatc gtcgccgcac  | 1980 |
| ttatgactgt cttctttatc atgcaactcg taggacaggt gccggcagcg ctctgggtca  | 2040 |
| ttttcggcga ggaccgcttt cgctggagcg cgacgatgat cggcctgtcg cttgcggtat  | 2100 |
| tcggaatctt gcacgccctc gctcaagcct tcgtcactgg tcccgccacc aaacgtttcg  | 2160 |
| gcgagaagca ggccattatc gccggcatgg cgggccacg ggtgcgcatg atcgtgctcc   | 2220 |
| tgctggttag gaccgggcta ggctggcggg gttgccttac tggttagcag aatgaatcac  | 2280 |
| cgatacgcga gcgaacgtga agcgactgct gctgcaaaac gtctgcgacc tgagcaacaa  | 2340 |
| catgaatggt cttcggtttc cgtgtttcgt aaagtctgga aacgcggaag tcagcgccct  | 2400 |
| gcaccattat gttccggatc tgcatcgag gatgctgctg gctaccctgt ggaacaccta   | 2460 |
| catctgtatt aacgaagcgc tggcattgac cctgagtgat ttttctctgg tcccgccgca  | 2520 |
| tccataccgc cagttgttta ccctcacaac gttccagtaa ccgggcatgt tcatcatcag  | 2580 |
| taacccgtat cgtgagcatc ctctctcgtt tcatcggtat cattaccccc atgaacagaa  | 2640 |
| atccccctta cacggaggca tcagtgacca aacaggaaaa aaccgccctt aacatggccc  | 2700 |

|                                                                     |      |
|---------------------------------------------------------------------|------|
| gctttatcag aagccagaca ttaacgcttc tggagaaact caacgagctg gacgcggatg   | 2760 |
| aacaggcaga catctgtgaa tcgcttcacg accacgctga tgagctttac cgcagctgcc   | 2820 |
| tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca | 2880 |
| cagcttgctt gtaagcggat gccgggagca gacaagcccc tcagggcgcg tcagcgggtg   | 2940 |
| ttggcgggtg tcggggcgca gccatgaccc agtcacgtag cgatagcgga gtgtatactg   | 3000 |
| gcttaactat gcggcatcag agcagattgt actgagagtg caccatatat gcggtgtgaa   | 3060 |
| ataccgcaca gatgcgtaag gagaaaatac cgcacagggc gctcttcgcg ttcctcgctc   | 3120 |
| actgactcgc tgcgctcggc cgcttcggctg cggcgagcgg tatcagctca ctcaaaggcg  | 3180 |
| gtaatacggc tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc   | 3240 |
| cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc   | 3300 |
| ccccctgacg agcatcacia aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga   | 3360 |
| ctataaagat accaggcggt tccccctgga agctccctcg tgcgctctcc tgttccgacc   | 3420 |
| ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat   | 3480 |
| agctcacgct gtaggtatct cagttcgggt taggtcgctt gctccaagct gggctgtgtg   | 3540 |
| cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc   | 3600 |
| aaccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga    | 3660 |
| gcgaggatat taggcggtgc tacagagttc ttgaagtggc ggcctaacta cggctacact   | 3720 |
| agaaggacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt   | 3780 |
| ggtagctctt gatccggcaa acaaacacc gctggtagcg gtgggttttt tgtttgcaag    | 3840 |
| cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg   | 3900 |
| tctgacgctc agtggaaacga aaactcacgt taagggattt tggatcatgaa caataaaact | 3960 |
| gtctgcttac ataaacagta atacaagggg tgttatgagc catattcaac gggaaacgtc   | 4020 |
| ttgctctagg ccgcgattaa attccaacat ggatgctgat ttatatgggt ataaatgggc   | 4080 |
| tcgcgataat gtcgggcaat caggtgcgac aatctatcga ttgtatggga agcccgatgc   | 4140 |
| gccagagttg tttctgaaac atggcaaagg tagcgttgcc aatgatgtta cagatgagat   | 4200 |
| ggtcagacta aactggctga cggaatttat gcctcttcgc accatcaagc attttatccg   | 4260 |
| tactcctgat gatgcatggt tactcaccac tgcgatcccc gggaaaacag cattccaggt   | 4320 |
| attagaagaa tatcctgatt caggtgaaaa tattgttgat gcgctggcag tgttcctgcg   | 4380 |
| ccggttgcat tcgattcctg tttgtaattg tccttttaac agcgatcgcg tatttcgtct   | 4440 |
| cgctcaggcg caatcacgaa tgaataacgg tttggttgat gcgagtgatt ttgatgacga   | 4500 |
| gcgtaatggc tggcctgttg aacaagtctg gaaagaaatg cataaacttt tgccattctc   | 4560 |
| accggattca gtcgtcactc atggtgattt ctacttgat aaccttattt ttgacgaggg    | 4620 |
| gaaattaata ggttgatttg atgttgacg agtcggaatc gcagaccgat accaggatct    | 4680 |
| tgccatccta tggaactgcc tcggtgagtt ttctccttca ttacagaaac ggctttttca   | 4740 |
| aaaatatggt attgataatc ctgatatgaa taaattgcag tttcatttga tgctcgatga   | 4800 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| gtttttctaa gaattaattc atgagcggat acatatattga atgtatttag aaaaataaac | 4860 |
| aaataggggt tccgcgcaca tttccccgaa aagtgccacc tgaaattgta aacgttaata  | 4920 |
| ttttgttaaa attcgcgtta aatttttgtt aaatcagctc attttttaac caataggccg  | 4980 |
| aaatcggcaa aatcccttat aaatcaaaag aatagaccga gatagggttg agtgttgttc  | 5040 |
| cagtttgga caagagtcca ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa   | 5100 |
| ccgtctatca gggcgatggc ccactacgtg aaccatcacc ctaatcaagt tttttggggt  | 5160 |
| cgaggtgccg taaagcacta aatcggaacc ctaaaggag cccccgattt agagcttgac   | 5220 |
| ggggaaagcc ggcgaaactg gcgagaaagg aagggaagaa agcgaaagga gcgggcgcta  | 5280 |
| gggcgctggc aagtgtagcg gtcacgtgc gcgtaaccac cacaccgcc gcgcttaatg    | 5340 |
| cgccgctaca gggcgctcc cattcgcca                                     | 5369 |

<210> 95  
 <211> 2187  
 <212> DNA  
 <213> *Psilocybe cubensis*

|                                                                    |      |
|--------------------------------------------------------------------|------|
| <400> 95                                                           |      |
| atggcttcta gttcttccga tgtcttcgtt ttgggtctag gtgttgtttt ggctgccttg  | 60   |
| tatatcttca gagaccaatt attcgctgct tctaagccaa aggtggctcc agtttccact  | 120  |
| acgaagcctg ccaacggttc cgctaaccac agagacttca tcgccaagat gaaacaaggt  | 180  |
| aagaagagaa tcgtaatctt ctacggttct caaactggta ccgctgaaga atatgctatt  | 240  |
| cgtttggtta aggaagctaa gcaaaagttc ggtctagcct ctttggtttg tgatccagaa  | 300  |
| gaatacgatt ttgaaaagtt ggaccaattg ccagaagatt ctattgcttt cttcgtcgtt  | 360  |
| gctacctatg gtgaaggatga acctacagac aacgctgtcc aattgttgca aaacttgcaa | 420  |
| gatgaaagct tcgaattctc ctctggtgag agaaagttgt caggtttgaa gtacgttggt  | 480  |
| tttggtctgg gtaacaagac ctacgaacat tacaacctca ttgggagaac tgttgacgct  | 540  |
| caattggcca agatgggtgc tatcagaatc ggtgaaagag gtgaaggatga tgatgacaag | 600  |
| tccatggaag aagactactt ggaatggaag gatggtatgt gggaagcgtt tgccactgct  | 660  |
| atgggtgttg aagaaggtca aggtggtgac tccgctgatt tcgtcgtttc cgaattggaa  | 720  |
| tctcaccac cagaaaaggt ttaccaaggt gaattttctg ctagagcttt aaccaaaacc   | 780  |
| aagggtattc acgacgctaa gaatcctttt gctgctccaa ttgcggttgc tagagaattg  | 840  |
| ttccaatctg ttgtcgatag aaactgtgtc cacgtcgaat tcaacattga aggctctggt  | 900  |
| atcacctatc aacacggtga ccacgttggg ttgtggccat tgaatccaga tgttgaagtc  | 960  |
| gaacggttgt tgtgtgtttt aggttttagct gaaaagagag atgctgtcat ctccattgaa | 1020 |
| tccttagacc cggctttggc taaggttcca ttcccagtc cactactta cggtgctgtg    | 1080 |
| ttgagacact acattgacat ctctgctgtc gccggtagac aaatcttggg tactttgtcc  | 1140 |
| aaattcgctc caacccaga agctgaagct ttcttgagaa acttgaacac taacaaggaa   | 1200 |
| gaataccaca acgtcgtcgc taacggttgt ttgaaattgg gtgaaatttt gcaaatcgct  | 1260 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| accggtaacg acattactgt cccaccaact actgccaaca ccaccaaagtg gccaatcca  | 1320 |
| ttcgacatca ttgtttctgc catcccaaga ttgcaaccaa gatactactc tatctcttct  | 1380 |
| tccccaaaaa ttcatccaaa caccatccac gctaccgttg ttgtgctcaa atacgaaaac  | 1440 |
| gttccaaccg aaccaatccc aagaaagtgg gtttacgggtg tcggtagtaa cttcttggtg | 1500 |
| aatttaaagt acgctgttaa caaggaacca gttccataca tcaactcaaaa tggcgaacaa | 1560 |
| agagtcggtg tcccggaata cttgattgct ggtccacgtg gttcttataa gactgaatct  | 1620 |
| ttctacaagg ctccaatcca tgtagacgt tctactttcc gtttgccaac caacccaaag   | 1680 |
| tctccagtca tcatgattgg tccaggtact ggtgtcgccc cattcagagg cttcgttcaa  | 1740 |
| gaaagagttg ctttggccag aagatccatc gaaaagaacg gtcctgactc tttggctgac  | 1800 |
| tggggctgta tttccttggt ctacggttgt agaagatccg acgaagactt cttgtacaag  | 1860 |
| gacgaatggc cacaatacga agctgagttg aagggttaagt tcaagttgca ctgtgctttc | 1920 |
| tccagacaaa actacaagcc agacggttct aagattttacg tccaagattt gatctgggaa | 1980 |
| gacagagaac acattgccga tgccatctta aacggtaagg gttacgtcta catctgcggt  | 2040 |
| gaagctaagt ccatgtctaa acaagttgaa gaagttctag ccaagatctt gggcgaagcc  | 2100 |
| aaaggtggtt ccggtccagt tgaagggtt gctgaagtca agttactgaa ggaacgttcc   | 2160 |
| agattgatgt tggatgtctg gtcttga                                      | 2187 |

<210> 96  
 <211> 728  
 <212> PRT  
 <213> *Psilocybe cubensis*  
  
 <400> 96

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ala | Ser | Ser | Ser | Asp | Val | Phe | Val | Leu | Gly | Leu | Gly | Val | Val |
| 1   |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Ala | Ala | Leu | Tyr | Ile | Phe | Arg | Asp | Gln | Leu | Phe | Ala | Ala | Ser | Lys |
|     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Pro | Lys | Val | Ala | Pro | Val | Ser | Thr | Thr | Lys | Pro | Ala | Asn | Gly | Ser | Ala |
|     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | Pro | Arg | Asp | Phe | Ile | Ala | Lys | Met | Lys | Gln | Gly | Lys | Lys | Arg | Ile |
| 50  |     |     |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Val | Ile | Phe | Tyr | Gly | Ser | Gln | Thr | Gly | Thr | Ala | Glu | Glu | Tyr | Ala | Ile |
| 65  |     |     |     |     | 70  |     |     |     |     | 75  |     |     |     | 80  |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Arg | Leu | Ala | Lys | Glu | Ala | Lys | Gln | Lys | Phe | Gly | Leu | Ala | Ser | Leu | Val |
|     |     |     | 85  |     |     |     |     |     | 90  |     |     |     |     | 95  |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Cys | Asp | Pro | Glu | Glu | Tyr | Asp | Phe | Glu | Lys | Leu | Asp | Gln | Leu | Pro | Glu |
|     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |

Asp Ser Ile Ala Phe Phe Val Val Ala Thr Tyr Gly Glu Gly Glu Pro  
115 120 125

Thr Asp Asn Ala Val Gln Leu Leu Gln Asn Leu Gln Asp Glu Ser Phe  
130 135 140

Glu Phe Ser Ser Gly Glu Arg Lys Leu Ser Gly Leu Lys Tyr Val Val  
145 150 155 160

Phe Gly Leu Gly Asn Lys Thr Tyr Glu His Tyr Asn Leu Ile Gly Arg  
165 170 175

Thr Val Asp Ala Gln Leu Ala Lys Met Gly Ala Ile Arg Ile Gly Glu  
180 185 190

Arg Gly Glu Gly Asp Asp Asp Lys Ser Met Glu Glu Asp Tyr Leu Glu  
195 200 205

Trp Lys Asp Gly Met Trp Glu Ala Phe Ala Thr Ala Met Gly Val Glu  
210 215 220

Glu Gly Gln Gly Gly Asp Ser Ala Asp Phe Val Val Ser Glu Leu Glu  
225 230 235 240

Ser His Pro Pro Glu Lys Val Tyr Gln Gly Glu Phe Ser Ala Arg Ala  
245 250 255

Leu Thr Lys Thr Lys Gly Ile His Asp Ala Lys Asn Pro Phe Ala Ala  
260 265 270

Pro Ile Ala Val Ala Arg Glu Leu Phe Gln Ser Val Val Asp Arg Asn  
275 280 285

Cys Val His Val Glu Phe Asn Ile Glu Gly Ser Gly Ile Thr Tyr Gln  
290 295 300

His Gly Asp His Val Gly Leu Trp Pro Leu Asn Pro Asp Val Glu Val  
305 310 315 320

Glu Arg Leu Leu Cys Val Leu Gly Leu Ala Glu Lys Arg Asp Ala Val  
325 330 335

Ile Ser Ile Glu Ser Leu Asp Pro Ala Leu Ala Lys Val Pro Phe Pro  
340 345 350

Val Pro Thr Thr Tyr Gly Ala Val Leu Arg His Tyr Ile Asp Ile Ser  
355 360 365

Ala Val Ala Gly Arg Gln Ile Leu Gly Thr Leu Ser Lys Phe Ala Pro  
370 375 380

Thr Pro Glu Ala Glu Ala Phe Leu Arg Asn Leu Asn Thr Asn Lys Glu

|                                                                 |  |     |  |     |  |     |
|-----------------------------------------------------------------|--|-----|--|-----|--|-----|
| 385                                                             |  | 390 |  | 395 |  | 400 |
| Glu Tyr His Asn Val Val Ala Asn Gly Cys Leu Lys Leu Gly Glu Ile |  |     |  |     |  |     |
|                                                                 |  | 405 |  | 410 |  | 415 |
| Leu Gln Ile Ala Thr Gly Asn Asp Ile Thr Val Pro Pro Thr Thr Ala |  |     |  |     |  |     |
|                                                                 |  | 420 |  | 425 |  | 430 |
| Asn Thr Thr Lys Trp Pro Ile Pro Phe Asp Ile Ile Val Ser Ala Ile |  |     |  |     |  |     |
|                                                                 |  | 435 |  | 440 |  | 445 |
| Pro Arg Leu Gln Pro Arg Tyr Tyr Ser Ile Ser Ser Ser Pro Lys Ile |  |     |  |     |  |     |
|                                                                 |  | 450 |  | 455 |  | 460 |
| His Pro Asn Thr Ile His Ala Thr Val Val Val Leu Lys Tyr Glu Asn |  |     |  |     |  |     |
|                                                                 |  | 465 |  | 470 |  | 475 |
|                                                                 |  |     |  |     |  | 480 |
| Val Pro Thr Glu Pro Ile Pro Arg Lys Trp Val Tyr Gly Val Gly Ser |  |     |  |     |  |     |
|                                                                 |  | 485 |  | 490 |  | 495 |
| Asn Phe Leu Leu Asn Leu Lys Tyr Ala Val Asn Lys Glu Pro Val Pro |  |     |  |     |  |     |
|                                                                 |  | 500 |  | 505 |  | 510 |
| Tyr Ile Thr Gln Asn Gly Glu Gln Arg Val Gly Val Pro Glu Tyr Leu |  |     |  |     |  |     |
|                                                                 |  | 515 |  | 520 |  | 525 |
| Ile Ala Gly Pro Arg Gly Ser Tyr Lys Thr Glu Ser Phe Tyr Lys Ala |  |     |  |     |  |     |
|                                                                 |  | 530 |  | 535 |  | 540 |
| Pro Ile His Val Arg Arg Ser Thr Phe Arg Leu Pro Thr Asn Pro Lys |  |     |  |     |  |     |
|                                                                 |  | 545 |  | 550 |  | 555 |
|                                                                 |  |     |  |     |  | 560 |
| Ser Pro Val Ile Met Ile Gly Pro Gly Thr Gly Val Ala Pro Phe Arg |  |     |  |     |  |     |
|                                                                 |  | 565 |  | 570 |  | 575 |
| Gly Phe Val Gln Glu Arg Val Ala Leu Ala Arg Arg Ser Ile Glu Lys |  |     |  |     |  |     |
|                                                                 |  | 580 |  | 585 |  | 590 |
| Asn Gly Pro Asp Ser Leu Ala Asp Trp Gly Arg Ile Ser Leu Phe Tyr |  |     |  |     |  |     |
|                                                                 |  | 595 |  | 600 |  | 605 |
| Gly Cys Arg Arg Ser Asp Glu Asp Phe Leu Tyr Lys Asp Glu Trp Pro |  |     |  |     |  |     |
|                                                                 |  | 610 |  | 615 |  | 620 |
| Gln Tyr Glu Ala Glu Leu Lys Gly Lys Phe Lys Leu His Cys Ala Phe |  |     |  |     |  |     |
|                                                                 |  | 625 |  | 630 |  | 635 |
|                                                                 |  |     |  |     |  | 640 |
| Ser Arg Gln Asn Tyr Lys Pro Asp Gly Ser Lys Ile Tyr Val Gln Asp |  |     |  |     |  |     |
|                                                                 |  | 645 |  | 650 |  | 655 |
| Leu Ile Trp Glu Asp Arg Glu His Ile Ala Asp Ala Ile Leu Asn Gly |  |     |  |     |  |     |
|                                                                 |  | 660 |  | 665 |  | 670 |



Lys Gly Tyr Val Tyr Ile Cys Gly Glu Ala Lys Ser Met Ser Lys Gln  
675 680 685

Val Glu Glu Val Leu Ala Lys Ile Leu Gly Glu Ala Lys Gly Gly Ser  
690 695 700

Gly Pro Val Glu Gly Val Ala Glu Val Lys Leu Leu Lys Glu Arg Ser  
705 710 715 720

Arg Leu Met Leu Asp Val Trp Ser  
725

<210> 97  
<211> 5203  
<212> DNA  
<213> Artificial Sequenec

<220>  
<223> pETM6-H10 vector

<400> 97  
gaagaattgt gagcggataa caattcccct ctagaaataa ttttgtttaa ctttaagaag 60  
gagatataca tatggcagat ctcaattgga tatcggccgg ccacgcgac gctgacgtcg 120  
gtaccctcga gtctggtaaa gaaaccgctg ctgcgaaatt tgaacgccag cacatggact 180  
cgtctactag tcgcagctta attaacctaa actgctgcca ccgctgagca ataactagca 240  
taacccttg gggcctctaa acgggtcttg aggggttttt tgctagcgaa aggaggagtc 300  
gactatatcc ggattggcga atgggacgcg ccctgtagcg gcgcattaa cgcggcgggt 360  
gtggtggtta cgcgcagcgt gaccgctaca cttgccagcg ccctagcgcc cgctcctttc 420  
gctttcttcc cttcctttct cgccacgttc gccggctttc cccgtcaagc tctaaatcgg 480  
gggctccctt tagggttcgc atttagtgct ttacggcacc tcgacccaa aaaacttgat 540  
tagggtgatg gttcacgtag tgggccatcg ccctgataga cggtttttcg ccctttgacg 600  
ttggagtcca cgttctttta tagtggactc ttgttcctaa ctggaacaac actcaaccct 660  
atctcggctc attcttttga tttataaggg attttgccga tttcggccta ttggttaaaa 720  
aatgagctga ttttaaaaa atttaacgcg aattttaaca aaatattaac gtttacaatt 780  
tctggcggca cgatggcatg agattatcaa aaaggatctt cacctagatc cttttaaatt 840  
aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc 900  
aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca tccatagttg 960  
cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct ggccccagtg 1020  
ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca ataaaccagc 1080  
cagccggaag ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc atccagtcta 1140  
ttaattgttg ccgggaagct agagtaagta gttcgccagt taatagtttg cgcaacgttg 1200  
ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt tggatatggct tcattcagct 1260

|            |            |            |             |            |            |      |
|------------|------------|------------|-------------|------------|------------|------|
| ccggttccca | acgatcaagg | cgagttacat | gatccccc    | gttgtgcaaa | aaagcggta  | 1320 |
| gctccttcgg | tcctccgatc | gttgtcagaa | gtaagttggc  | cgcagtgtta | tcactcatgg | 1380 |
| ttatggcagc | actgcataat | tctcttactg | tcatgccatc  | cgtaagatgc | ttttctgtga | 1440 |
| ctggtgagta | ctcaaccaag | tcattctgag | aatagtgtat  | gcggcgaccg | agttgctctt | 1500 |
| gcccggcgtc | aatacgggat | aataccgcgc | cacatagcag  | aactttaaaa | gtgctcatca | 1560 |
| ttggaaaacg | ttcttcgggg | cgaaaactct | caaggatctt  | accgctgttg | agatccagtt | 1620 |
| cgatgtaacc | cactcgtgca | cccaactgat | cttcagcatc  | ttttactttc | accagcgttt | 1680 |
| ctgggtgagc | aaaaacagga | aggcaaaatg | ccgcaaaaaa  | gggaataagg | gcgacacgga | 1740 |
| aatgttgaat | actcatactc | ttcctttttc | aatcatgatt  | gaagcattta | tcagggttat | 1800 |
| tgtctcatga | gcggatacat | atttgaatgt | atttagaaaa  | ataaaciaat | aggtcatgac | 1860 |
| caaaatccct | taacgtgagt | tttcgttcca | ctgagcgtca  | gaccccgtag | aaaagatcaa | 1920 |
| aggatcttct | tgagatcctt | tttttctgcg | cgtaatctgc  | tgcttgcaaa | caaaaaaacc | 1980 |
| accgctacca | gcggttggtt | gtttgccgga | tcaagagcta  | ccaactcttt | ttccgaaggt | 2040 |
| aactggcttc | agcagagcgc | agataccaaa | tactgtcctt  | ctagtgtagc | cgtagttagg | 2100 |
| ccaccacttc | aagaactctg | tagcaccgcc | tacatacctc  | gctctgctaa | tcctgttacc | 2160 |
| agtggctgct | gccagtggcg | ataagtcgtg | tcttaccggg  | ttggactcaa | gacgatagtt | 2220 |
| accggataag | gcgcagcggg | cgggctgaac | gggggggttc  | tgcacacagc | ccagcttga  | 2280 |
| gcgaacgacc | tacaccgaac | tgagatacct | acagcgtgag  | ctatgagaaa | gcgccacgct | 2340 |
| tcccgaaggg | agaaaggcgg | acaggtatcc | ggtaagcggc  | agggtcggaa | caggagagcg | 2400 |
| cacgaggagg | cttccagggg | gaaacgcctg | gtatctttat  | agtcctgtcg | ggtttcgcca | 2460 |
| cctctgactt | gagcgtcgat | ttttgtgatg | ctcgtcaggg  | gggcggagcc | tatggaaaaa | 2520 |
| cgccagcaac | gcggcctttt | tacggttcct | ggccttttgc  | tggccttttg | ctcacatgtt | 2580 |
| ctttcctgcg | ttatcccctg | attctgtgga | taaccgtatt  | accgcctttg | agtgagctga | 2640 |
| taccgctcgc | gcgagccgaa | cgaccgagcg | cagcgagtca  | gtgagcgagg | aagcggaaga | 2700 |
| gcgcctgatg | cggatatttc | tccttacgca | tctgtgcggt  | atttcacacc | gcatatatgg | 2760 |
| tgcactctca | gtacaatctg | ctctgatgcc | gcatagttaa  | gccagtatac | actccgctat | 2820 |
| cgctacgtga | ctgggtcatg | gctgcgcccc | gacacccgcc  | aacacccgct | gacgcgccct | 2880 |
| gacgggcttg | tctgctcccg | gcatccgctt | acagacaagc  | tgtgaccgtc | tccgggagct | 2940 |
| gcatgtgtca | gaggttttca | ccgtcatcac | cgaaacgcgc  | gaggcagctg | cggtaaagct | 3000 |
| catcagcgtg | gtcgtgaagc | gattcacaga | tgtctgcctg  | ttcatccgcg | tccagctcgt | 3060 |
| tgagtttctc | cagaagcggt | aatgtctggc | ttctgataaa  | gcgggccatg | ttaaggcg   | 3120 |
| ttttttcctg | tttggtcact | gatgcctccg | tgtaaggggg  | atttctgttc | atgggggtaa | 3180 |
| tgataccgat | gaaacgagag | aggatgctca | cgatacgggt  | tactgatgat | gaacatgccc | 3240 |
| ggttactgga | acgttgtgag | ggtaaacaac | tggcgggtatg | gatgcggcgg | gaccagagaa | 3300 |
| aaatcactca | gggtcaatgc | cagcgcttcg | ttaatacaga  | tgtaggtgtt | ccacagggta | 3360 |

|                                                                   |      |
|-------------------------------------------------------------------|------|
| gccagcagca tcctgcgatg cagatccgga acataatggt gcagggcgct gacttccgcg | 3420 |
| tttccagact ttacgaaaca cggaaccga agaccattca tgttgttgct caggtcgcag  | 3480 |
| acgttttgca gcagcagtcg cttcacgttc gctcgcgtat cggtgattca ttctgctaac | 3540 |
| cagtaaggca accccgccag cctagccggg tcctcaacga caggagcacg atcatgctag | 3600 |
| tcatgccccg cgcccaccgg aaggagctga ctgggttgaa ggctctcaag ggcatcggtc | 3660 |
| gagatcccgg tgccaatga gtgagctaac ttacattaat tgcgttgcg cactgccccg   | 3720 |
| ctttccagtc gggaaacctg tcgtgccagc tgcattaatg aatcggccaa cgcgcgggga | 3780 |
| gaggcggttt gcgtattggg cgccagggtg gtttttcttt tcaccagtga gacgggcaac | 3840 |
| agctgattgc cttcaccgc ctggccctga gagagtgcga gcaagcggtc cacgctgggt  | 3900 |
| tgccccagca ggcgaaaatc ctgtttgatg gtggttaacg gcgggatata acatgagctg | 3960 |
| tcttcggtat cgtcgtatcc cactaccgag atgtccgcac caacgcgcag cccggactcg | 4020 |
| gtaatggcgc gcattgcgc cagcgccatc tgatcgttgg caaccagcat cgcagtggga  | 4080 |
| acgatgccct cattcagcat ttgcatggtt tgttgaaaac cggacatggc actccagtcg | 4140 |
| ccttcccgtt ccgctatcgg ctgaatttga ttgcgagtga gatatttatg ccagccagcc | 4200 |
| agacgcagac gcgccgagac agaacttaat gggcccgcta acagcgcgat ttgctggtga | 4260 |
| cccaatgcga ccagatgctc cacgcccagt cgcgtaccgt cttcatggga gaaaataata | 4320 |
| ctgttgatgg gtgtctggtc agagacatca agaaataacg ccggaacatt agtgcaggca | 4380 |
| gcttccacag caatggcatc ctggatcatc agcggatagt taatgatcag cccactgacg | 4440 |
| cgttcgcgga gaagattgtg caccgccgct ttacaggctt cgacgccgct tcgttctacc | 4500 |
| atcgacacca ccacgctggc acccagttga tcggcgcgag atttaatcgc cgcgacaatt | 4560 |
| tgcgacggcg cgtgcagggc cagactggag gtggcaacgc caatcagcaa cgactgtttg | 4620 |
| cccgccagtt gttgtgccac gcggttggga atgtaattca gtcgcccat cgccgcttcc  | 4680 |
| actttttccc gcgttttcgc agaaacgtgg ctggcctggt tcaccacgcg ggaaacggtc | 4740 |
| tgataagaga caccggcata ctctgcgaca tcgtataacg ttactggttt cacattcacc | 4800 |
| accctgaatt gactctcttc cgggcgctat catgccatac cgcgaaaggt tttgcgcat  | 4860 |
| tcgatggtgt ccgggatctc gacgtcttc cttatgcgac tcctgcatta ggaagcagcc  | 4920 |
| cagtagtagg ttgaggccgt tgagcaccgc cgccgcaagg aatggtgcat gcaaggagat | 4980 |
| ggcgcccaac agtcccccg ccacggggcc tgccaccata ccacgccga aacaagcgct   | 5040 |
| catgagcccc aagtggcgag cccgatcttc cccatcggtg atgtcggcga tataggcgcc | 5100 |
| agcaaccgca cctgtggcgc cggatgatcc ggccacgatg cgtccggcgt agcctaggat | 5160 |
| cgagatcgat ctcgatcccg cgaaattaat acgactcact acg                   | 5203 |

<210> 98

<211> 3777

<212> DNA

<213> Artificial Sequence

<220>

<223> pRSM3 vector

<400> 98

|                                                                    |      |
|--------------------------------------------------------------------|------|
| cgctaacagc gcgatttgct ggtgacccaa tgcgaccaga tgctccacgc ccagtcgcgt  | 60   |
| accgtcttca tgggagaaaa taatactgtt gatgggtgtc tggtcagaga catcaagaaa  | 120  |
| taacgccgga acattagtgc aggagcttc cacagcaatg gcacacctgt catccagcgg   | 180  |
| atagttaatg atcagcccac tgacgcgttg cgcgagaaga ttgtgcaccg ccgctttaca  | 240  |
| ggcttcgacg ccgcttcgtt ctaccatcga caccaccacg ctggcaccca gttgatcggc  | 300  |
| gcgagattta atcgccgcga caatttgcga cggcgcgtgc agggccagac tggaggtggc  | 360  |
| aacgccaatc agcaacgact gtttgccgc cagttgttgt gccacgcggt tgggaatgta   | 420  |
| attcagctcc gccatcgccg cttccacttt tccccgcgtt ttcgcagaaa cgtggctggc  | 480  |
| ctggttcacc acgcgggaaa cggcttgata agagacaccg gcatactctg cgacatcgta  | 540  |
| taacgttact ggtttcacat tcaccaccct gaattgactc tcttcgggc gctatcatgc   | 600  |
| cataccgcga aaggttttgc gccattcgat ggtgtccggg atctcgacgc tctcccttat  | 660  |
| gcgactcctg cattaggaag cagcccagta gtaggttgag gccgttgagc accgccgccg  | 720  |
| caaggaatgg tgcagcaag gagatggcgc ccaacagtcc cccggccacg gggcctgcca   | 780  |
| ccataccac gccgaaacaa gcgctcatga gccgaagtg gcgagcccga tcttccccat    | 840  |
| cggatgatgtc ggcgatatag gcgccagcaa ccgcacctgt ggcgccggtg atgccggcca | 900  |
| cgatgcgtcc ggcgtagcct aggatcgaga tcgatctcga tcccgcgaaa ttaatacgac  | 960  |
| tcactatagg ggaattgtga gcggataaca attcccctct agaaataatt ttgtttaact  | 1020 |
| ttaagaagga gatatacata tggcagatct caattggata tcggccggcc acgcgatcgc  | 1080 |
| tgacgtcggc accctcgagt ctggtaaaga aaccgctgct gcgaaatttg aacgccagca  | 1140 |
| catggactcg tctactagtc gcagcttaat taacctaaac tgctgccacc gctgagcaat  | 1200 |
| aactagcata accccttggg gcctctaaac gggctttgag gggttttttg ctagcgaaag  | 1260 |
| gaggagtcga caagctgacg accgggtctc cgcaagtggc acttttcggg gaaatgtgcg  | 1320 |
| cggaaaccct atttgtttat ttttctaaat acattcaaat atgtatccgc tcatgaatta  | 1380 |
| attcttagaa aaactcatcg agcatcaaat gaaactgcaa ttatttcata tcaggattat  | 1440 |
| caataccata tttttgaaaa agccgtttct gtaatgaagg agaaaactca ccgaggcagt  | 1500 |
| tccataggat ggcaagatcc tggatcggc ctgcgattcc gactcgtcca acatcaatac   | 1560 |
| aacctattaa tttcccctcg tcaaaaataa gggttatcaag tgagaaatca ccatgagtga | 1620 |
| cgactgaatc cggtgagaat ggcaaaagtt tatgcatttc tttccagact tgttcaacag  | 1680 |
| gccagccatt acgctcgtca tcaaaatcac tcgcatcaac caaacgtta ttcattcgtg   | 1740 |
| attgcgcctg agcgagacga aatacgcggt cgctgttaaa aggacaatta caaacaggaa  | 1800 |
| tcgaatgcaa ccggcgcagg aacactgccg gcgcatcaac aatattttca cctgaatcag  | 1860 |
| gatattcttc taatacttgg aatgctgttt tccgggggat cgcagtgggtg agtaaccatg | 1920 |
| catcatcagg agtacggata aaatgcttga tggtcggaag aggcataaat tccgtcagcc  | 1980 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| agtttagtct gaccatctca tctgtaacat cattggcaac gctacctttg ccatgtttca  | 2040 |
| gaaacaactc tggcgcatcg ggcttcccat acaatcgata gattgtcgca cctgattgcc  | 2100 |
| cgacattatc gcgagcccat ttatacccat ataatcagc atccatgttg gaatttaatc   | 2160 |
| gcggcctaga gcaagacgtt tcccgttgaa tatggctcat actcttcctt tttcaatatt  | 2220 |
| attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga  | 2280 |
| aaaataaaca aataggcatg cagcgctctt ccgcttcctc gctcactgac tcgctacgct  | 2340 |
| cggtcgttcg actgcggcga gcggtgtcag ctactcaaa agcggtaata cggttatcca   | 2400 |
| cagaatcagg ggataaagcc ggaaagaaca tgtgagcaaa aagcaaagca ccggaagaag  | 2460 |
| ccaacgccgc aggcgttttt ccataggctc cgccccctg acgagcatca caaaaatcga   | 2520 |
| cgctcaagcc agaggtagcg aaacccgaca ggactataaa gataccaggc gtttccccct  | 2580 |
| ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata cctgtccgcc  | 2640 |
| tttctccctt cggaagcgtt ggcgctttct catagctcac gctgttggtt tctcagttcg  | 2700 |
| gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac cccccgttca gcccgaccgc  | 2760 |
| tgcgcttat ccggtaaacta tcgtcttgag tccaaccgg taagacacga cttatcgcca   | 2820 |
| ctggcagcag ccattggtaa ctgatttaga ggactttgtc ttgaagttat gcacctgtta  | 2880 |
| aggctaaact gaaagaacag attttggtga gtgcggctct ccaaccact taccttggtt   | 2940 |
| caaagagttg gtagctcagc gaaccttgag aaaaccaccg ttggtagcgg tggtttttct  | 3000 |
| ttatttatga gatgatgaat caatcggctt atcaagtcaa cgaacagcta ttccgttact  | 3060 |
| ctagatttca gtgcaattta tctcttcaaa tgtagcacct gaagtcagcc ccatacgata  | 3120 |
| taagttgtaa ttctcatgtt agtcatgccc cgcgcccacc ggaaggagct gactgggttg  | 3180 |
| aaggctctca agggcatcgg tcgagatccc ggtgcctaata gagtgagcta acttacatta | 3240 |
| attgcgttgc gctcactgcc cgctttccag tcgggaaacc tgctgtgcca gctgcattaa  | 3300 |
| tgaatcggcc aacgcgcggg gagaggcggg ttgcgtattg ggccgaggg tggtttttct   | 3360 |
| tttcaccagt gagacgggca acagctgatt gcccttcacc gcctggccct gagagagttg  | 3420 |
| cagcaagcgg tccacgtggg tttgccccag caggcgaaaa tcctgtttga tggtaggttaa | 3480 |
| cggcgggata taacatgagc tgtcttcggg atcgtcgtat cccactaccg agatgtccgc  | 3540 |
| accaacgcgc agcccgact cggtaatggc gcgcattgcg cccagcgcca tctgatcgtt   | 3600 |
| ggcaaccagc atcgcagtgg gaacgatgcc ctcatcagc atttgcatgg tttgttgaaa   | 3660 |
| accggacatg gcactccagt cgccttcccc ttccgctatc ggctgaattt gattgcgagt  | 3720 |
| gagatattta tgccagccag ccagacgcag acgcgccgag acagaactta atgggcc     | 3777 |