



УКРАЇНА

(19) UA

(11) 86918

(13) C2

(51) МПК (2009)

C12N 5/14

C12N 9/10

C12N 15/54

C12N 15/82

A01H 5/00

МІНІСТЕРСТВО ОСВІТИ  
І НАУКИ УКРАЇНИ

ДЕРЖАВНИЙ ДЕПАРТАМЕНТ  
ІНТЕЛЕКТУАЛЬНОЇ  
ВЛАСНОСТІ

## ОПИС ДО ПАТЕНТУ НА ВИНАХІД

**(54) ІЗОЛЬОВАНИЙ АБО РЕКОМБІНАНТНИЙ ПОЛІНУКЛЕОТИД, ЩО КОДУЄ ПОЛІПЕПТИД, ЯКИЙ МАЄ АКТИВНІСТЬ ГЛІФОСАТ-N-АЦЕТИЛТРАНСФЕРАЗИ (GAT)**

1

2

(21) 2003055027

(22) 29.10.2001

(24) 10.06.2009

(86) PCT/US01/46227, 29.10.2001

(31) 60/244,385

(32) 30.10.2000

(33) US

(46) 10.06.2009, Бюл.№ 11, 2009 р.

(72) КАСЛ ЛІНДА А., САЙХЛ ДЕН, ГАЙВЕР ЛО-  
РАЙН ДЖ., МІНШУЛЛ ДЖЕРЕМІ, АЙВІ КРІСТІНА,  
ЧЕН ЯНГ ХОНГ, ДАК НІКОЛАС Б., МАККАТЧЕН  
БІЛЛІ Ф., КЕМБЛ РОДЖЕР, ПАТТЕН ФІЛЛІП А.

(73) ВЕРДІА, ІНК., ПАЄНІЕ ХАЙ-БРЕД ІНТЕР-  
НЕШНЛ, ІНК., Е.І. ДЮ ПОН ДЕ НЕМУР ЕНД КОМ-  
ПАНІ

(56) DATABASE EMBL 'Online! 1 July 1997 (1997-  
07-01) Kunst et al.: "YITI protein" Database accession  
no. 006744, XP002222224.

WO A 0009727, 24.02.2000.

(57) 1. Ізольований або рекомбінантний полінуклеотид, що кодує поліпептид, який має активність гліфосат-N-ацетилтрансферази, що включає нуклеотидну послідовність, що кодує послідовність амінокислот, яка має щонайменше 60 % ідентичності послідовності по всій довжині послідовності, вибраної з групи, що включає SEQ. ID. No.: 300, SEQ. ID. No.: 445 і SEQ. ID. No.: 457.

2. Ізольований або рекомбінантний полінуклеотид, що кодує поліпептид, який має активність гліфосат-N-ацетилтрансферази, що включає нуклеотидну послідовність, яка кодує щонайменше 140 суміжних амінокислот амінокислотної послідовності, вибраної з групи, що включає SEQ. ID. No.: 300, SEQ. ID. No.: 445 і SEQ. ID. No.: 457.

3. Полінуклеотид за п. 2, де

(а) батьківський кодон був замінений синонімічним кодоном, що переважно застосовується в рослинах порівняно з батьківським кодоном; і/або

(б) згаданий полінуклеотид додатково включає нуклеотидну послідовність, що кодує N-кінцевий хлоропласт-транзитний пептид.

4. Ізольований або рекомбінантний полінуклеотид, що кодує поліпептид, який має активність гліфосат-N-ацетилтрансферази, що включає нуклеотидну послідовність, комплемент якої гібридується в жорстких умовах по суті по всій довжині з нуклеотидною послідовністю, що кодує амінокислотну послідовність, вибрану з групи, яка включає SEQ. ID. No.: 300, SEQ. ID. No.: 445 і SEQ. ID. No.: 457, де до жорстких умов гібридизації належить 50%-й формамід з 1 мг гепарину при 42°C, при проведенні гібридизації протягом ночі і промивка 0,2 x SSC при 65°C протягом 15 хвилин.

5. Полінуклеотид за п. 4, де

(с) батьківський кодон був замінений синонімічним кодоном, що переважно застосовується в рослинах порівняно з батьківським кодоном; і/або

(д) згаданий полінуклеотид додатково включає нуклеотидну послідовність, що кодує N-кінцевий хлоропласт-транзитний пептид.

6. Ізольований або рекомбінантний полінуклеотид за п. 1, де вказаний полінуклеотид включає нуклеотидну послідовність, що кодує амінокислотну послідовність з SEQ. ID. Nos.: 6-10 і 263-514.

7. Ізольований або рекомбінантний полінуклеотид за п. 1, де вказана нуклеотидна послідовність кодує поліпептид, в якому щонайменше 80 % положень не виходять за рамки наступних обмежень:

(а) в положенні 2 амінокислотний залишок являє собою I або L;

(б) в положенні 3 амінокислотний залишок являє собою E або D;

(с) в положенні 4 амінокислотний залишок являє собою V, A або I;

(д) в положенні 5 амінокислотний залишок являє собою K, R або N;

(е) в положенні 6 амінокислотний залишок являє собою P або L;

(ф) в положенні 8 амінокислотний залишок являє собою N, S або T;

(г) в положенні 10 амінокислотний залишок являє собою E або G;

(13) C2

(11) 86918

(19) UA

(h) в положенні 11 амінокислотний залишок являє собою D або E;  
(i) в положенні 12 амінокислотний залишок являє собою T або A;  
(j) в положенні 14 амінокислотний залишок являє собою E або K;  
(k) в положенні 15 амінокислотний залишок являє собою I або L;  
(l) в положенні 17 амінокислотний залишок являє собою H або Q;  
(m) в положенні 18 амінокислотний залишок являє собою R, C або K;  
(n) в положенні 19 амінокислотний залишок являє собою I або V;  
(o) в положенні 24 амінокислотний залишок являє собою Q або R;  
(p) в положенні 26 амінокислотний залишок являє собою L або I;  
(q) в положенні 27 амінокислотний залишок являє собою E або D;  
(r) в положенні 28 амінокислотний залишок являє собою A або V;  
(s) в положенні 30 амінокислотний залишок являє собою K, M або R;  
(t) в положенні 31 амінокислотний залишок являє собою Y або F;  
(u) в положенні 32 амінокислотний залишок являє собою E або G;  
(v) в положенні 33 амінокислотний залишок являє собою T, A або S;  
(w) в положенні 35 амінокислотний залишок являє собою L, S або M;  
(x) в положенні 37 амінокислотний залишок являє собою R, G, E або Q;  
(y) в положенні 38 амінокислотний залишок являє собою G або S;  
(z) в положенні 39 амінокислотний залишок являє собою T, A або S;  
(aa) в положенні 40 амінокислотний залишок являє собою F, L або S;  
(ab) в положенні 45 амінокислотний залишок являє собою Y або F;  
(ac) в положенні 47 амінокислотний залишок являє собою R, Q або G;  
(ad) в положенні 48 амінокислотний залишок являє собою G або D;  
(ae) в положенні 49 амінокислотний залишок являє собою K, R, E або Q;  
(af) в положенні 51 амінокислотний залишок являє собою I або V;  
(ag) в положенні 52 амінокислотний залишок являє собою S, C або G;  
(ah) в положенні 53 амінокислотний залишок являє собою I або T;  
(ai) в положенні 54 амінокислотний залишок являє собою A або V;  
(aj) в положенні 57 амінокислотний залишок являє собою H або N;  
(ak) в положенні 58 амінокислотний залишок являє собою Q, K, N або P;  
(al) в положенні 59 амінокислотний залишок являє собою A або S;  
(am) в положенні 60 амінокислотний залишок являє собою E, K, G, V або D;  
(an) в положенні 61 амінокислотний залишок являє собою H або Q;

(ao) в положенні 62 амінокислотний залишок являє собою P, S або T;  
(ap) в положенні 63 амінокислотний залишок являє собою E, G або D;  
(aq) в положенні 65 амінокислотний залишок являє собою E, D, V або Q;  
(ar) в положенні 67 амінокислотний залишок являє собою Q, E, R, L, H або K;  
(as) в положенні 68 амінокислотний залишок являє собою K, R, E або N;  
(at) в положенні 69 амінокислотний залишок являє собою Q або P;  
(au) в положенні 79 амінокислотний залишок являє собою E або D;  
(av) в положенні 80 амінокислотний залишок являє собою G або E;  
(aw) в положенні 81 амінокислотний залишок являє собою Y, N або F;  
(ax) в положенні 82 амінокислотний залишок являє собою R або H;  
(ay) в положенні 83 амінокислотний залишок являє собою E, G або D;  
(az) в положенні 84 амінокислотний залишок являє собою Q, R або L;  
(ba) в положенні 86 амінокислотний залишок являє собою A або V;  
(bb) в положенні 89 амінокислотний залишок являє собою T або S;  
(bc) в положенні 90 амінокислотний залишок являє собою L або I;  
(bd) в положенні 91 амінокислотний залишок являє собою I або V;  
(be) в положенні 92 амінокислотний залишок являє собою R або K;  
(bf) в положенні 93 амінокислотний залишок являє собою H, Y або Q;  
(bg) в положенні 96 амінокислотний залишок являє собою E, A або Q;  
(bh) в положенні 97 амінокислотний залишок являє собою L або I;  
(bi) в положенні 100 амінокислотний залишок являє собою K, R, N або E;  
(bj) в положенні 101 амінокислотний залишок являє собою K або R;  
(bk) в положенні 103 амінокислотний залишок являє собою A або V;  
(bl) в положенні 104 амінокислотний залишок являє собою D або N;  
(bm) в положенні 105 амінокислотний залишок являє собою L або M;  
(bn) в положенні 106 амінокислотний залишок являє собою L або I;  
(bo) в положенні 112 амінокислотний залишок являє собою T або I;  
(bp) в положенні 113 амінокислотний залишок являє собою S, T або F;  
(bq) в положенні 114 амінокислотний залишок являє собою A або V;  
(br) в положенні 115 амінокислотний залишок являє собою S, R або A;  
(bs) в положенні 119 амінокислотний залишок являє собою K, E або R;  
(bt) в положенні 120 амінокислотний залишок являє собою K або R;  
(bu) в положенні 123 амінокислотний залишок являє собою F або L;

(bv) в положенні 124 амінокислотний залишок являє собою S або R;  
 (bw) в положенні 125 амінокислотний залишок являє собою E, K, G або D;  
 (bx) в положенні 126 амінокислотний залишок являє собою Q або H;  
 (by) в положенні 128 амінокислотний залишок являє собою E, G або K;  
 (bz) в положенні 129 амінокислотний залишок являє собою V, I або A;  
 (ca) в положенні 130 амінокислотний залишок являє собою Y, H, F або C;  
 (cb) в положенні 131 амінокислотний залишок являє собою D, G, N або E;  
 (cc) в положенні 132 амінокислотний залишок являє собою I, T, A, M, V або L;  
 (cd) в положенні 135 амінокислотний залишок являє собою V, T, A або I;  
 (ce) в положенні 138 амінокислотний залишок являє собою H або Y;  
 (cf) в положенні 139 амінокислотний залишок являє собою I або V;  
 (cg) в положенні 140 амінокислотний залишок являє собою L або S;  
 (ch) в положенні 142 амінокислотний залишок являє собою Y або H;  
 (ci) в положенні 143 амінокислотний залишок являє собою K, T або E;  
 (cj) в положенні 144 амінокислотний залишок являє собою K, E або R;  
 (ck) в положенні 145 амінокислотний залишок являє собою L або I;  
 (cl) в положенні 146 амінокислотний залишок являє собою T або A;  
 (cm) в положенні 9, 76, 94 і 110 амінокислотний залишок являє собою A;  
 (cn) в положеннях 29 і 108 амінокислотний залишок являє собою C;  
 (co) в положенні 34 амінокислотний залишок являє собою D;  
 (cp) в положенні 95 амінокислотний залишок являє собою E;  
 (cq) в положенні 56 амінокислотний залишок являє собою F;  
 (cr) в положеннях 43, 44, 66, 74, 87, 102, 116, 122, 127 і 136 амінокислотний залишок являє собою G;  
 (cs) в положенні 41 амінокислотний залишок являє собою H;  
 (ct) в положенні 7 амінокислотний залишок являє собою I;  
 (cu) в положенні 85 амінокислотний залишок являє собою K;  
 (cv) в положеннях 20, 36, 42, 50, 72, 78, 98 і 121 амінокислотний залишок являє собою L;  
 (cw) в положеннях 1, 75 і 141 амінокислотний залишок являє собою M;  
 (cx) в положеннях 23, 64 і 109 амінокислотний залишок являє собою N;  
 (cy) в положеннях 22, 25, 133, 134 і 137 амінокислотний залишок являє собою P;  
 (cz) в положенні 71 амінокислотний залишок являє собою Q;  
 (da) в положеннях 16, 21, 73, 99 і 111 амінокислотний залишок являє собою R;  
 (db) в положеннях 55 і 88 амінокислотний залишок являє собою S;

(dc) в положенні 77 амінокислотний залишок являє собою T;  
 (dd) в положенні 107 амінокислотний залишок являє собою W; i  
 (de) в положеннях 13, 46, 70, 117 і 118 амінокислотний залишок являє собою Y.  
 8. Ізольований або рекомбінантний полінуклеотид за п. 1, де згаданий полінуклеотид включає нуклеотидну послідовність, що кодує послідовність амінокислот, яка має щонайменше 70 % ідентичності послідовності по всій довжині послідовності, вибраної з групи, що включає SEQ. ID. No.: 300, SEQ. ID. No.: 445 і SEQ. ID. No.: 457.  
 9. Ізольований або рекомбінантний полінуклеотид за п. 1, де згаданий полінуклеотид включає нуклеотидну послідовність, що кодує послідовність амінокислот, яка має щонайменше 80 % ідентичності послідовності по всій довжині послідовності, вибраної з групи, що включає SEQ. ID. No.: 300, SEQ. ID. No.: 445 і SEQ. ID. No.: 457.  
 10. Ізольований або рекомбінантний полінуклеотид за п. 1,  
 (i) в якому поліпептид каталізує ацетилювання гліфосату з  $k_{\text{кат}}/K_m$  щонайменше  $10 \text{ mM}^{-1}\text{хв.}^{-1}$  до гліфосату; i/або  
 (ii) поліпептид каталізує ацетилювання амінометилфосфонової кислоти; i/або  
 (iii) щонайменше 80 % положень поліпептиду узгоджуються з такими обмеженнями:  
 (a) в положенні 9, 76, 94 і 110 амінокислотний залишок являє собою A;  
 (b) в положеннях 29 і 108 амінокислотний залишок являє собою C;  
 (c) в положенні 34 амінокислотний залишок являє собою D;  
 (d) в положенні 95 амінокислотний залишок являє собою E;  
 (e) в положенні 56 амінокислотний залишок являє собою F;  
 (f) в положеннях 43, 44, 66, 74, 87, 102, 116, 122, 127 і 136 амінокислотний залишок являє собою G;  
 (g) в положенні 41 амінокислотний залишок являє собою H;  
 (h) в положенні 7 амінокислотний залишок являє собою I;  
 (i) в положенні 85 амінокислотний залишок являє собою K;  
 (j) в положеннях 20, 36, 42, 50, 72, 78, 98 і 121 амінокислотний залишок являє собою L;  
 (k) в положеннях 1, 75 і 141 амінокислотний залишок являє собою M;  
 (l) в положеннях 23, 64 і 109 амінокислотний залишок являє собою N;  
 (m) в положеннях 22, 25, 133, 134 і 137 амінокислотний залишок являє собою P;  
 (n) в положенні 71 амінокислотний залишок являє собою Q;  
 (o) в положеннях 16, 21, 73, 99 і 111 амінокислотний залишок являє собою R;  
 (p) в положеннях 55 і 88 амінокислотний залишок являє собою S;  
 (q) в положенні 77 амінокислотний залишок являє собою T;  
 (r) в положенні 107 амінокислотний залишок являє собою W; i

(s) в положеннях 13, 46, 70, 117 і 118 амінокислотний залишок являє собою Y.

11. Ізольований або рекомбінантний полінуклеотид за п. 1, в якому поліпептид включає амінокислотну послідовність, вибрану з SEQ. ID. No.: 300, SEQ. ID. No.: 445 і SEQ. ID. No.: 457.

12. Ізольований або рекомбінантний полінуклеотид за п. 11, що включає нуклеотидну послідовність, вибрану з SEQ. ID. No.: 48, SEQ. ID. No.: 193 або SEQ. ID. No.: 205.

13. Ізольований або рекомбінантний полінуклеотид за будь-яким з пп. 1, 6, 7, 8, 9, 10, 11 або 12, де

(a) батьківський кодон був замінений синонімічним кодоном, що переважно застосовується в рослинах порівняно з батьківським кодоном; і/або

(b) згаданий полінуклеотид додатково включає нуклеотидну послідовність, що кодує N-кінцевий хлоропласт-транзитний пептид.

14. Нуклеїновокислотна конструкція, що включає полінуклеотид за будь-яким з пп. 1-13, де згадана конструкція включає промотор, оперативно зв'язаний зі згаданим полінуклеотидом, при цьому промотор є гетерологічним відносно полінуклеотиду і здатен ефективно викликати достатню експресію закодованого поліпептиду для посилення толерантності до гліфосату рослинної клітини, трансформованої нуклеїновокислотою конструкцією.

15. Нуклеїновокислотна конструкція за п. 14, яка додатково включає другу полінуклеотидну послідовність, що кодує другий поліпептид, який надає клітині або організму, що експресує другий поліпептид на ефективному рівні, фенотипову ознаку, яка може бути виявлена; і/або де конструкція включає Т-ДНК-послідовність; і/або де полінуклеотид є оперативно зв'язаним з регуляторною послідовністю; і/або де конструкція являє собою вектор трансформації рослин.

16. Клітина-хозяїн, що включає щонайменше один полінуклеотид за будь-яким з пп. 1-13 або щонайменше одну конструкцію за п. 14 або 15, де полінуклеотид, що кодує активність гліфосат-N-ацетилтрансферази, є гетерологічним до клітини.

17. Клітина-хозяїн за п. 16, де клітина є рослинною клітиною.

18. Клітина-хозяїн за п. 17, що додатково містить метаболічний продукт гліфосату, який є N-ацетилгліфосатом, де зазначений метаболічний продукт утворений поліпептидом, що містить амінокислотну послідовність, яка має щонайменше 80 % послідовностей, ідентичних до амінокислотної послідовності, що знаходиться у SEQ. ID. No.: 300, 445 або 457.

19. Трансгенна рослина або насіння, вироблене нею, або культура тканини трансгенної рослини, що включає клітину за п. 17, де рослина або культура тканини рослини експресує поліпептид з активністю гліфосат-N-ацетилтрансферази.

20. Трансгенна рослина, насіння або культура тканини трансгенної рослини за п. 19, при цьому трансгенна рослина або культура тканини рослини являє собою культурну рослину, вибрану серед родів Eleusine, Lolium, Bambusa, Brassica, Dactylis, Sorghum, Pennisetum, Zea, Oryza, Triticum, Secale,

Avena, Hordeum, Saccharum, Coix, Glycine і Gossypium.

21. Трансгенна рослина, насіння або культура тканини трансгенної рослини за п. 19 або 20, де рослина або культура тканини рослини виявляє підвищену стійкість до гліфосату порівняно з дикою рослиною того ж виду, лінією або культурним сортом.

22. Ізольований або рекомбінантний поліпептид, що має активність гліфосат-N-ацетилтрансферази, де згаданий поліпептид включає амінокислотну послідовність, яка має щонайменше 60 % ідентичності послідовності по всій довжині послідовності, вибраної з групи, що включає SEQ. ID. No.: 300, SEQ. ID. No.: 445 і SEQ. ID. No.: 457.

23. Ізольований або рекомбінантний поліпептид за п. 22, де згаданий поліпептид включає амінокислотну послідовність, яка має щонайменше 70 % ідентичності послідовності по всій довжині послідовності, вибраної з групи, що включає SEQ. ID. No.: 300, SEQ. ID. No.: 445 і SEQ. ID. No.: 457.

24. Ізольований або рекомбінантний поліпептид за п. 22, де згаданий поліпептид включає амінокислотну послідовність, яка має щонайменше 80 % ідентичності послідовності по всій довжині послідовності, вибраної з групи, що включає SEQ. ID. No.: 300, SEQ. ID. No.: 445 і SEQ. ID. No.: 457.

25. Ізольований або рекомбінантний поліпептид за п. 22, де поліпептид каталізує ацетилювання гліфосату з  $K_{cat}/K_m$  щонайменше  $10\text{mM}^{-1}\text{хв}^{-1}$  до гліфосату; і/або

(i) де поліпептид каталізує ацетилювання амінометилфосфонові кислоти; і/або

(ii) де щонайменше 80 % положень поліпептиду не порушують наступні рестрикції:

(a) в положенні 9, 76, 94 і 110 амінокислотний залишок являє собою A;

(b) в положеннях 29 і 108 амінокислотний залишок являє собою C;

(c) в положенні 34 амінокислотний залишок являє собою D;

(d) в положенні 95 амінокислотний залишок являє собою E;

(e) в положенні 56 амінокислотний залишок являє собою F;

(f) в положеннях 43, 44, 66, 74, 87, 102, 116, 122, 127 і 136 амінокислотний залишок являє собою G;

(g) в положенні 41 амінокислотний залишок являє собою H;

(h) в положенні 7 амінокислотний залишок являє собою I;

(i) в положенні 85 амінокислотний залишок являє собою K;

(j) в положеннях 20, 36, 42, 50, 72, 78, 98 і 121 амінокислотний залишок являє собою L;

(k) в положеннях 1, 75 і 141 амінокислотний залишок являє собою M;

(l) в положеннях 23, 64 і 109 амінокислотний залишок являє собою N;

(m) в положеннях 22, 25, 133, 134 і 137 амінокислотний залишок являє собою P;

(n) в положенні 71 амінокислотний залишок являє собою Q;

(o) в положеннях 16, 21, 73, 99 і 111 амінокислотний залишок являє собою R;



(р) в положеннях 55 і 88 амінокислотний залишок являє собою S;  
 (q) в положенні 77 амінокислотний залишок являє собою T;  
 (r) в положенні 107 амінокислотний залишок являє собою W; i  
 (s) в положеннях 13, 46, 70, 117 і 118 амінокислотний залишок являє собою Y.  
 26. Ізольований або рекомбінантний поліпептид за п. 22, де поліпептид включає амінокислотну послідовність, вибрану з групи, що включає SEQ. ID. No.: 300, SEQ. ID. No.: 445 і SEQ. ID. No.: 457.  
 27. Ізольований або рекомбінантний поліпептид, що має активність гліфосат-N-ацетилтрансферази, де згаданий поліпептид включає щонайменше 140 суміжних амінокислот амінокислотної послідовності, вибраної з групи, що включає SEQ. ID. No.: 300, SEQ. ID. No.: 445 і SEQ. ID. No.: 457.  
 28. Ізольований або рекомбінантний поліпептид за п. 27, який додатково містить N-кінцевий хлоропласт-транзитний пептид; та/або додатково містить послідовність секреції або послідовність локалізації.  
 29. Ізольований або рекомбінантний поліпептид, що має активність гліфосат-N-ацетилтрансферази, який кодується нуклеотидною послідовністю, яка гібридується в жорстких умовах по суті по всій довжині з комплементом нуклеотидної послідовності, що кодує амінокислотну послідовність, вибрану з групи, що включає SEQ. ID. No.: 300, SEQ. ID. No.: 445 і SEQ. ID. No.: 457, де до жорстких умов гібридизації належить 50%-й формамід з 1 мг гепарину при 42°C, при проведенні гібридизації протягом ночі і промивка 0,2 x SSC при 65°C протягом 15 хвилин.  
 30. Ізольований або рекомбінантний поліпептид за п. 29, який додатково містить N-кінцевий хлоропласт-транзитний пептид; та/або додатково містить послідовність секреції або послідовність локалізації.  
 31. Ізольований або рекомбінантний поліпептид за п. 22, де згаданий поліпептид має:  
 (a)  $K_m$  до гліфосату щонайменше близько 2 мМ або менше;  $K_m$  до ацетил-CoA щонайменше близько 200 мкМ або менше; і  $k_{кат}$  дорівнює щонайменше близько 6/хвилину; або  
 (b) щонайменше 80 % положень поліпептиду узгоджуються з такими обмеженнями:  
 (a) в положенні 2 амінокислотний залишок являє собою I або L;  
 (b) в положенні 3 амінокислотний залишок являє собою E або D;  
 (c) в положенні 4 амінокислотний залишок являє собою V, A або I;  
 (d) в положенні 5 амінокислотний залишок являє собою K, R або N;  
 (e) в положенні 6 амінокислотний залишок являє собою P або L;  
 (f) в положенні 8 амінокислотний залишок являє собою N, S або T;  
 (g) в положенні 10 амінокислотний залишок являє собою E або G;  
 (h) в положенні 11 амінокислотний залишок являє собою D або E;  
 (i) в положенні 12 амінокислотний залишок являє собою T або A;

(j) в положенні 14 амінокислотний залишок являє собою E або K;  
 (k) в положенні 15 амінокислотний залишок являє собою I або L;  
 (l) в положенні 17 амінокислотний залишок являє собою H або Q;  
 (m) в положенні 18 амінокислотний залишок являє собою R, C або K;  
 (n) в положенні 19 амінокислотний залишок являє собою I або V;  
 (o) в положенні 24 амінокислотний залишок являє собою Q або R;  
 (p) в положенні 26 амінокислотний залишок являє собою L або I;  
 (q) в положенні 27 амінокислотний залишок являє собою E або D;  
 (r) в положенні 28 амінокислотний залишок являє собою A або V;  
 (s) в положенні 30 амінокислотний залишок являє собою K, M або R;  
 (t) в положенні 31 амінокислотний залишок являє собою Y або F;  
 (u) в положенні 32 амінокислотний залишок являє собою E або G;  
 (v) в положенні 33 амінокислотний залишок являє собою T, A або S;  
 (w) в положенні 35 амінокислотний залишок являє собою L, S або M;  
 (x) в положенні 37 амінокислотний залишок являє собою R, G, E або Q;  
 (y) в положенні 38 амінокислотний залишок являє собою G або S;  
 (z) в положенні 39 амінокислотний залишок являє собою T, A або S;  
 (aa) в положенні 40 амінокислотний залишок являє собою F, L або S;  
 (ab) в положенні 45 амінокислотний залишок являє собою Y або F;  
 (ac) в положенні 47 амінокислотний залишок являє собою R, Q або G;  
 (ad) в положенні 48 амінокислотний залишок являє собою G або D;  
 (ae) в положенні 49 амінокислотний залишок являє собою K, R, E або Q;  
 (af) в положенні 51 амінокислотний залишок являє собою I або V;  
 (ag) в положенні 52 амінокислотний залишок являє собою S, C або G;  
 (ah) в положенні 53 амінокислотний залишок являє собою I або T;  
 (ai) в положенні 54 амінокислотний залишок являє собою A або V;  
 (aj) в положенні 57 амінокислотний залишок являє собою H або N;  
 (ak) в положенні 58 амінокислотний залишок являє собою Q, K, N або P;  
 (al) в положенні 59 амінокислотний залишок являє собою A або S;  
 (am) в положенні 60 амінокислотний залишок являє собою E, K, G, V або D;  
 (an) в положенні 61 амінокислотний залишок являє собою H або Q;  
 (ao) в положенні 62 амінокислотний залишок являє собою P, S або T;  
 (ap) в положенні 63 амінокислотний залишок являє собою E, G або D;

(aq) в положенні 65 амінокислотний залишок являє собою E, D, V або Q;  
 (ar) в положенні 67 амінокислотний залишок являє собою Q, E, R, L, H або K;  
 (as) в положенні 68 амінокислотний залишок являє собою K, R, E або N;  
 (at) в положенні 69 амінокислотний залишок являє собою Q або P;  
 (au) в положенні 79 амінокислотний залишок являє собою E або D;  
 (av) в положенні 80 амінокислотний залишок являє собою G або E;  
 (aw) в положенні 81 амінокислотний залишок являє собою Y, N або F;  
 (ax) в положенні 82 амінокислотний залишок являє собою R або H;  
 (ay) в положенні 83 амінокислотний залишок являє собою E, G або D;  
 (az) в положенні 84 амінокислотний залишок являє собою Q, R або L;  
 (ba) в положенні 86 амінокислотний залишок являє собою A або V;  
 (bb) в положенні 89 амінокислотний залишок являє собою T або S;  
 (bc) в положенні 90 амінокислотний залишок являє собою L або I;  
 (bd) в положенні 91 амінокислотний залишок являє собою I або V;  
 (be) в положенні 92 амінокислотний залишок являє собою R або K;  
 (bf) в положенні 93 амінокислотний залишок являє собою H, Y або Q;  
 (bg) в положенні 96 амінокислотний залишок являє собою E, A або Q;  
 (bh) в положенні 97 амінокислотний залишок являє собою L або I;  
 (bi) в положенні 100 амінокислотний залишок являє собою K, R, N або E;  
 (bj) в положенні 101 амінокислотний залишок являє собою K або R;  
 (bk) в положенні 103 амінокислотний залишок являє собою A або V;  
 (bl) в положенні 104 амінокислотний залишок являє собою D або N;  
 (bm) в положенні 105 амінокислотний залишок являє собою L або M;  
 (bn) в положенні 106 амінокислотний залишок являє собою L або I;  
 (bo) в положенні 112 амінокислотний залишок являє собою T або I;  
 (bp) в положенні 113 амінокислотний залишок являє собою S, T або F;  
 (bq) в положенні 114 амінокислотний залишок являє собою A або V;  
 (br) в положенні 115 амінокислотний залишок являє собою S, R або A;  
 (bs) в положенні 119 амінокислотний залишок являє собою K, E або R;  
 (bt) в положенні 120 амінокислотний залишок являє собою K або R;  
 (bu) в положенні 123 амінокислотний залишок являє собою F або L;  
 (bv) в положенні 124 амінокислотний залишок являє собою S або R;  
 (bw) в положенні 125 амінокислотний залишок являє собою E, K, G або D;

(bx) в положенні 126 амінокислотний залишок являє собою Q або H;  
 (by) в положенні 128 амінокислотний залишок являє собою E, G або K;  
 (bz) в положенні 129 амінокислотний залишок являє собою V, I або A;  
 (ca) в положенні 130 амінокислотний залишок являє собою Y, H, F або C;  
 (cb) в положенні 131 амінокислотний залишок являє собою D, G, N або E;  
 (cc) в положенні 132 амінокислотний залишок являє собою I, T, A, M, V або L;  
 (cd) в положенні 135 амінокислотний залишок являє собою V, T, A або I;  
 (ce) в положенні 138 амінокислотний залишок являє собою H або Y;  
 (cf) в положенні 139 амінокислотний залишок являє собою I або V;  
 (cg) в положенні 140 амінокислотний залишок являє собою L або S;  
 (ch) в положенні 142 амінокислотний залишок являє собою Y або H;  
 (ci) в положенні 143 амінокислотний залишок являє собою K, T або E;  
 (cj) в положенні 144 амінокислотний залишок являє собою K, E або R;  
 (ck) в положенні 145 амінокислотний залишок являє собою L або I; i  
 (cl) в положенні 146 амінокислотний залишок являє собою T або A.  
 32. Ізольований або рекомбінантний поліпептид за будь-яким з пп. 22-28 або 31, що додатково включає N-кінцевий хлоропласт-транзитний пептид; і/або додатково включає послідовність, яка відповідає за секрецію, або послідовність, яка відповідає за локалізацію.  
 33. Спосіб одержання поліпептиду, що має активність гліфосат-N-ацетилтрансферази, при цьому спосіб включає культивування клітини за п. 16 або 17 або рослини, насіння або культури тканини рослини за пп. 19, 20 або 21.  
 34. Спосіб одержання трансгенної рослини, стійкої до гліфосату, її насіння або рослинної клітини, який включає:  
 (a) трансформацію рослини або рослинної клітини полінуклеотидом за будь-яким з пп. 1-13 або тим, що включається у конструкцію за п. 14 або 15; і  
 (b) можливо, регенерацію трансгенної рослини з трансформованої рослинної клітини.  
 35. Спосіб за п. 34, що додатково включає вирощування трансформованої рослини або рослинної клітини при такій концентрації гліфосату, яка інгібує ріст дикої рослини того ж виду, причому ця концентрація не інгібує ріст трансформованої рослини,  
 де згадане вирощування відбувається при концентраціях гліфосату, що підвищуються, і/або де згадане вирощування відбувається при концентрації гліфосату, що є летальною для дикої рослини або рослинної клітини того ж виду.  
 36. Спосіб за будь-яким з пп. 34-35, який додатково включає розмноження згаданої трансгенної рослини шляхом схрещування згаданої трансгенної рослини з другою рослиною, так, щоб щонайменше частина потомства схрещування показувала толерантність до гліфосату.

37. Спосіб селективної боротьби з бур'яном на полі з рослинною культурою, що включає:

(а) засів поля насінням або рослинами, які є толерантними до гліфосату в результаті трансформації їх полінуклеотидом за будь-яким з пп. 1-13 чи конструкції за пп. 14-15; і

(б) нанесення на культуру і бур'ян на полі достатньої кількості гліфосату для боротьби з бур'яном без суттєвого впливу на культуру.

38. Трансгенна рослина або культура тканини трансгенної рослини, що має підвищену толерантність до гліфосату, де рослина або культура тканини рослини експресує поліпептид з активністю гліфосат-N-ацетилтрансферази, де поліпептид з активністю гліфосат-N-ацетилтрансферази експресується з полінуклеотиду за будь-яким з пп. 1-13 і

(а) щонайменше один поліпептид, який надає толерантність до гліфосату за допомогою додаткового механізму, і/або

(б) щонайменше один поліпептид, який надає толерантність до додаткового гербіциду.

39. Трансгенна рослина або культура тканини трансгенної рослини за п. 38, де

(а) щонайменше один поліпептид, який надає толерантність до гліфосату за допомогою додаткового механізму, являє собою толерантну до гліфосату 5-енолпірувілшкімат-3-фосфат-синтазу або толерантну до гліфосату гліфосат-оксидоредуктазу, і/або

(б) щонайменше один поліпептид, який надає толерантність до додаткового гербіциду, являє собою мутовану гідроксифенілпіруватдіоксигеназу, сульфонамід-толерантну ацетолактатсинтазу, сульфонамід-толерантну ацетогідроксикислотну синтазу, імідазолінон-толерантну ацетолактатсинтазу, імідазолінон-толерантну ацетогідроксикислотну синтазу, фосфінотрицин-ацетилтрансферазу або мутовану протопорфіриноген-оксидазу.

40. Спосіб боротьби з бур'яном на полі з рослинною культурою, який включає:

(а) засів поля насінням або рослинами культури за будь-яким з пп. 38-39 і

(б) нанесення на культуру і бур'ян на полі ефективною кількістю гліфосату, достатньої для інгібування росту бур'яну на полі без суттєвого впливу на культуру, і

(с) можливо, нанесення на культуру і бур'ян на полі одночасно або по чергово гліфосату і, можливо, додаткового гербіциду.

41. Спосіб за п. 40, де додатковий гербіцид, що наноситься, вибирають з групи, яка включає інгібітор гідроксифенілпіруватдіоксигенази, сульфонамід, імідазолінон, біалафос, фосфінотрицин, азафенідин, бутафенацил, сульфосат, глүфосінат та протокс-інгібітор.

42. Спосіб за п. 41, де згаданий додатковий гербіцид наноситься одночасно або по черзі.

43. Спосіб оцінювання активності GAT-поліпептида у рослинній тканині, що включає обробку рослини гліфосатом та аналізування рослинної тканини зазначеної рослини на присутність N-ацетилгліфосату, де зазначений GAT-поліпептид містить амінокислотну послідовність, яка має щонайменше 80 % послідовностей, ідентичних до амінокислотної послідовності, що знаходиться у SEQ. ID. No.: 300, 445 або 457.

44. Спосіб визначення присутності GAT-поліпептиду у рослинній тканині, що включає аналізування рослинної тканини на присутність N-ацетилгліфосату, де зазначений GAT-поліпептид містить амінокислотну послідовність, яка має щонайменше 80 % послідовностей, ідентичних до амінокислотної послідовності, що знаходиться у SEQ. ID. No.: 300, 445 або 457.

45. Спосіб за п. 44, де зазначений спосіб включає аналізування рослинної тканини, що використовується в імуноаналізі.

46. Спосіб визначення присутності полінуклеотиду, що кодує GAT-поліпептид, який включає аналізування рослинної тканини з використанням ПЦР-ампліфікації, де зазначений полінуклеотид, що кодує GAT-поліпептид, містить амінокислотну послідовність, яка має щонайменше 80 % послідовностей, ідентичних до амінокислотної послідовності, що знаходиться у SEQ. ID. No.: 300, 445 або 457.

47. Спосіб визначення, чи надає GAT-поліпептид трансгенним рослинам стійкості до гліфосату, що включає етапи: трансформування рослини GAT-полінуклеотидом, який кодує GAT-поліпептид, що містить амінокислотну послідовність, яка має щонайменше 80% послідовностей, ідентичних до амінокислотної послідовності, що знаходиться у SEQ. ID. No.: 300, 445 або 457; обробка трансформованої рослини гліфосатом; та визначення, чи рослина уражена або вбита обробкою гліфосатом.

Ця заявка підтверджує пріоритет і користь тимчасової патентної заявки США сер. №60/244,385, зареєстрованої 30 жовтня 2000р., опис якої включений в цю заявку у вигляді посилання у всій його повноті для будь-якого призначення.

Частина опису цього патентного документу містить матеріали, що підлягають захисту авторських прав. Власник авторських прав не має заперечень щодо копіювання будь-якого з патентних документів або опису патенту після його реєстрації в бюро патентів і торгових марок, але у протилежному випадку залишає за собою всі авторські права.

Селективність-сільськогосподарських культур до специфічних гербіцидів може надаватися шляхом конструювання генів з отриманням культур, які кодують відповідні гербіцид-метаболізуючі ферменти. В деяких випадках ці ферменти і нуклеїнові кислоти, їх кодуючі, утворюються в рослині. В інших випадках вони походять з інших організмів, таких як мікроби. Див., наприклад, Padgett et al. (1996) *New weed control opportunities: Development of soybeans with a Round UP Ready™ gene* in *Herbicide-Resistant Crops* (Duke, ed.), pp.54-84, CRC Press, Boca Raton; і Vasil (1996) *"Phosphinothricin-resistant crops"* in *Herbicide-Resistant Crops* (Duke, ed.),

pp.85-91. Дійсно, були сконструйовані трансгенні рослини, що експресують різноманітні гени толерантності/метаболізації гербіцидів з різноманітних організмів. Наприклад, у велику кількість рослин вводять синтазу ацетогідроксикислоти, що, як з'ясувалося, надає рослинам, які експресують цей фермент, стійкість до багатьох типів гербіцидів {див., наприклад, Hattori et al. (1995) Mol Gen Genet 246:419. Інші гени, що надають толерантність до гербіцидів, включають: ген, кодуючий химерний протеїн оксидоредуктази цитохрому шкура P450A1 і NADPH-цитохрому дріжджів P450 (Shiota et al. (1994) Plant Physiol Plant Physiol 106:17), гени глутатіон-редуктази і супероксид-дисмутази (Aono et al. (1995) Plant Cell Physiol 36:1687 і гени різних фосфотрансфераз (Datta et al. (1992) Plant Mol Biol 20:619.

Одним з гербіцидів, який є предметом широкого дослідження в цьому напрямку, є N-фосфометилглїцин, загальновідомий як гліфосат. Гліфосат являє собою гербіцид, який користується найвищим попиту у всьому світі, і, за прогнозами, об'єм його продажу у 2003 році досягне 5 млрд. доларів. Цей гербіцид широкого спектру дії, який знищує як листяні, так і трав'яні рослини. Вдалий шлях надання трансгенним рослинам стійкості до гліфосату на комерційному рівні складається у введенні модифікованого гена 5-енолпірувілшкімат-3-фосфат-синтази *Agrobacterium* CP4 (званої тут EPSP-синтазою або EPSPS). Трансген таргетований на хлоропласт, де він зберігає здатність синтезувати EPSP з фосфоенолпірувінової кислоти (PEP) і шкімат-3-фосфату в присутності гліфосату. У протилежність цьому, нативна EPSP-синтаза інгібується гліфосатом. Без цього трансгена рослини, обприскані гліфосатом, швидко гинуть внаслідок інгібування EPSP-синтази, яка зупиняє низхідний шлях, необхідний для біосинтезу ароматичних амінокислот, гормонів і вітамінів, гліфосат-стійкі CP4-трансгенні соєві рослини випускає, наприклад, компанія Monsanto під торговою маркою "Round UP Ready™".

В навколишньому середовищі розкладення гліфосату здійснюється головним чином через метаболізм ґрунтової мікрофлори. Первинний метаболіт гліфосату в ґрунті був ідентифікований як амінометилфосфонові кислота (AMPA), яка зрештою перетворюється на амоній, фосфат і двооксид вуглецю. Запропонована схема метаболізму, яка описує розкладення гліфосату в ґрунті через AMPA, показана на Фіг.8. Альтернативний шлях метаболізму для розщеплення гліфосату певними ґрунтовими бактеріями, тобто, саркозиневий шлях, відбувається через початкове розщеплення зв'язку C-P з утворенням неорганічного фосфату і саркозину, як показано на Фіг.9.

Іншу вдалу комбінацію гербіциду і трансгенної рослини являє собою глүфосінат (фосфінотрицин) і LibertyLink™, що випускається, наприклад, компанією Aventis. Глүфосінат також є гербіцидом широкого спектра дії. Він таргетований на фермент глутамат-синтазу хлоропласта. Стійкі рослини несуть ген *bar* з *Streptomyces hygroscopicus* і отримують стійкість завдяки N-

ацетилюючої активності *bar*, який модифікує і детоксифікує глүфосінат.

Фермент, здатний ацетилювати первинний амоні AMPA, описаний в заявці PCT №WO00/29596. Цей фермент не описаний там як здатний ацетилювати сполуку з вторинним аміном (наприклад, гліфосатом).

Хоча існує безліч стратегій надання стійкості до гербіцидів, як вказано вище, але додаткові підходи могли б мати велику комерційну цінність. Цей винахід стосується, наприклад, нових полінуклеотидів і поліпептидів для надання толерантності до гербіцидів, а також багатьох інших корисних особливостей, як буде зрозуміло з подальшого опису.

Задача цього винаходу складається в створенні способів і реагентів для надання організму, такому як рослина, стійкості до гліфосату. Ця та інші задачі винаходу вирішуються за допомогою одного або більше варіантів здійснення винаходу, описаних нижче.

Один з варіантів здійснення винаходу стосується нових поліпептидів, що називаються тут GAT-поліпептидами. GAT-поліпептиди відрізняються їх структурною подібністю один до одного, наприклад, подібністю послідовностей GAT-поліпептидів, коли вони вирівняні один з одним. Деякі GAT-поліпептиди мають активність гліфосат-N-ацетил-трансферази, тобто, здатність каталізувати ацетилювання гліфосату. Деякі GAT-поліпептиди також здатні каталізувати ацетилювання аналогів гліфосату і/або метаболітів гліфосату, наприклад, амінометилфосфонові кислоти.

Також пропонуються нові полінуклеотиди, що називаються тут GAT-полінуклеотидами. GAT-полінуклеотиди відрізняються їх здатністю кодувати GAT-поліпептиди. В деяких варіантах здійснення винаходу GAT-полінуклеотид конструюють для кращої експресії рослиною шляхом заміни одного або більше батьківських кодонів на синонімічний кодон, який переважно застосовується в рослинах порівняно з батьківським кодоном. В інших варіантах здійснення GAT-полінуклеотид модифікують шляхом введення нуклеотидної послідовності, що кодує N-кінцевий хлоропласт-транзитний пептид.

GAT-поліпептиди, GAT-полінуклеотиди і активність гліфосат-N-ацетил-трансферази детальніше описані нижче. Винахід також включає певні фрагменти GAT-поліпептидів і GAT-полінуклеотидів, описані нижче.

Винахід включає ненативні варіанти поліпептидів і полінуклеотидів, описаних тут, в яких одна або більше амінокислот закованого поліпептиду були мутовані.

Винахід також стосується нуклеїновокислотної конструкції, що включає полінуклеотид за винаходом. Ця конструкція може являти собою вектор, такий як рослинний вектор трансформації. В деяких аспектах вектор за винаходом буде включати T-ДНК послідовність. Конструкція може включати (необов'язково) регуляторну послідовність (наприклад, промотор), оперативно зв'язану з GAT-полінуклеотидом, де промотор є гетерологічним до полінуклеотиду і здатен ефективно ви-

кликати достатню експресію закодованого поліпептиду для посилення толерантності до гліфосату рослинних клітин, трансформованих нуклеїновокислотною конструкцією.

В деяких аспектах винаходу GAT-полінуклеотид функціонує як маркер, що селектується, наприклад, у рослинах, бактеріях, актиноміцетах, водоростях, дріжджах чи інших грибах. Наприклад, організм, що був трансформований вектором, включаючим GAT-полінуклеотидний маркер, що селектується, може бути селектований на основі його здатності рости в присутності гліфосату. GAT-маркерний ген може бути застосований для селекції або скринінгу на трансформовані клітини, що експресують цей ген.

Винахід також стосується векторів з комбінованими ознаками, тобто, векторів, які кодують GAT і які також включають другу полінуклеотидну послідовність, що кодує другий поліпептид, який надає клітині або організму, експресуючому другий поліпептид на ефективному рівні, фенотипову ознаку, що може бути виявлена. Фенотипова ознака, що може бути виявлена, може функціонувати як маркер, що селектується, наприклад, шляхом надання стійкості до гербіцидів, стійкості до шкідників або утворення деякого різновиду видимого маркера.

В одному з варіантів здійснення винахід стосується композиції, що включає один або більше полінуклеотидів за винаходом.

Композиції, що містять один або більше GAT-полінуклеотидів або закодованих поліпептидів, є рисою цього винаходу. В деяких випадках ці композиції являють собою бібліотеки нуклеїнових кислот, які містять, наприклад, щонайменше три або більше нуклеїнових кислот. Композиції, вироблені шляхом ферментативного розщеплення нуклеїнових кислот за винаходом рестрикційною ендонуклеазою, ДНК-азою або РНК-азою, або іншим способом фрагментування нуклеїнових кислот, наприклад, механічним розрізанням, хімічним розщепленням і т.д., також є рисою цього винаходу, як і композиції, утворені шляхом інкубації нуклеїнової кислоти за винаходом з дезоксирибонуклеотидтрифосфатами і полімеразою нуклеїнових кислот, такою як термостабільна полімераза нуклеїнових кислот.

Клітини, трансдуковані вектором за винаходом або такі, що включають нуклеїнову кислоту за винаходом іншим способом, є аспектом винаходу. У переважному варіанті здійснення винаходу клітини експресують поліпептид, закодований нуклеїновою кислотою.

В деяких варіантах здійснення винаходу клітини, що включають нуклеїнові кислоти за винаходом, є рослинними клітинами. Трансгенні рослини, трансгенні рослинні клітини і трансгенні рослинні експлантати, що включають нуклеїнові кислоти за винаходом, також є рисою цього винаходу. В деяких варіантах здійснення трансгенні рослини, трансгенні рослинні клітини і трансгенні рослинні експлантати експресують екзогенний поліпептид з активністю гліфосат-N-ацетилтрансферази, закодований нуклеїновою кислотою за винаходом. Винахід також стосується

ся трансгенного насіння, що виробляється трансгенними рослинами за винаходом.

Винахід також стосується трансгенних рослин або трансгенних рослинних експлантатів, що мають підвищену толерантність до гліфосату завдяки експресії поліпептиду з активністю гліфосат-N-ацетилтрансферази і поліпептиду, який надає толерантність до гліфосату за допомогою іншого механізму, такого як гліфосат-толерантна 5-енолпірувілшкімат-3-фосфат-синтаза і/або гліфосат-толерантна гліфосат-оксидо-редуктаза. В іншому варіанті здійснення винахід стосується трансгенних рослин або трансгенних рослинних експлантатів, що мають підвищену толерантність до гліфосату, а також толерантність до додаткового гербіциду завдяки експресії поліпептиду з активністю гліфосат-IM-ацетилтрансферази, поліпептиду, який надає толерантність до гліфосату за допомогою іншого механізму, такого як гліфосат-толерантна 5-енолпірувілшкімат-3-фосфат-синтаза і/або гліфосат-толерантна гліфосат-оксидо-редуктаза, і поліпептиду, який надає толерантність до додаткового гербіциду, такого як мутувана гідроксифенілпіруватдіоксигеназа, сульфонамід-толерантна ацетолактатсинтаза, сульфонамід-толерантна синтаза ацетогідроксикислоти, імідазолінон-толерантна ацетолактатсинтаза, імідазолінон-толерантна синтаза ацетогідроксикислоти, фосфінотрицин-ацетилтрансфераза і мутувана протопорфіриноген-оксидаза.

Винахід також стосується трансгенних рослин або трансгенних рослинних експлантатів, що мають підвищену толерантність до гліфосату, а також толерантність до додаткового гербіциду завдяки експресії поліпептиду з активністю гліфосат-N-ацетилтрансферази і поліпептиду, який надає толерантність до додаткового гербіциду, такого як мутувана гідроксифенілпіруватдіоксигеназа, сульфонамід-толерантна ацетолактатсинтаза, сульфонамід-толерантна синтаза ацетогідроксикислоти, імідазолінон-толерантна ацетолактатсинтаза, імідазолінон-толерантна синтаза ацетогідроксикислоти, фосфінотрицин-ацетилтрансфераза і мутувана протопорфіриноген-оксидаза.

Способи отримання поліпептидів за винаходом шляхом введення нуклеїнових кислот, що їх кодують, в клітини, а потім експресії і виділення їх з клітин або з середовища культури, є рисою цього винаходу. У переважних варіантах здійснення клітини, що експресують поліпептиди за винаходом, є трансгенними рослинними клітинами.

Поліпептиди, які специфічно зв'язуються з поліклональною антисироваткою, яка реагує на антиген, що походить з SEQ ID NOS: 6-10 і 263-514, але не з спорідненою природною послідовністю, наприклад, пептидом, представленим послідовністю з GenBank під інвентарним номером CAA70664, а також антитіла, які продукуються шляхом введення антигена, що походить з SEQ ID NOS 6-10 і 263-514, і/або які специфічно зв'язуються з такими антигенами і які специфічно не зв'язуються з природним поліпептидом, що відповідає інвентарному номеру GenBank CAA70664, усі є рисами цього винаходу.

Інший аспект винаходу стосується способів диверсифікації полінуклеотидів для отримання нових GAT-полінуклеотидів і полі пептидів шляхом рекомбінації або мутації нуклеїнових кислот за винаходом *in vitro* або *in vivo*. В одному з варіантів здійснення винаходу рекомбінація забезпечує щонайменше одну бібліотеку рекомбінантних GAT-полінуклеотидів. Бібліотеки, вироблені таким чином, є варіантами здійснення цього винаходу, як і клітини, що складають бібліотеки. Крім того, способи вироблення модифікованого GAT-полінуклеотиду шляхом мутації нуклеїнової кислоти за винаходом є варіантами здійснення винаходу. Рекомбінантні і мутантні GAT-полінуклеотиди і поліпептиди, вироблені способами за винаходом, також є варіантами здійснення винаходу.

В деяких аспектах цього винаходу диверсифікація здійснюється шляхом застосування рекурсивної рекомбінації, яка може бути проведена *in vitro*, *in vivo*, *in silico* або шляхом комбінації цих способів. Деякі приклади способів диверсифікації, описані нижче більш детально, являють собою способи родового шафлінгу та синтетичні способи шафлінгу.

Винахід стосується способів вироблення гліфосат-стійкої трансгенної рослини або рослинної клітини, який включає трансформацію рослини або рослинної клітини полінуклеотидом, що кодує гліфосат-N-ацетилтрансферазу і, можливо, регенерацію трансгенної рослини з трансформованої рослинної клітини. В деяких аспектах полінуклеотид являє собою GAT-полінуклеотид, можливо, GAT-полінуклеотид, що походить з бактеріального джерела. В деяких аспектах винаходу спосіб може включати вирощування трансформованої рослини або рослинної клітини при такій концентрації гліфосату, яка інгібує ріст дикої рослини того ж виду, але не інгібує росту трансформованої рослини. Спосіб може включати вирощування трансформованої рослини або рослинної клітини, або потомства цієї рослини або рослинної клітини при концентраціях гліфосату, що підвищуються і/або при такій концентрації гліфосату, яка є летальною для дикої рослини або рослинної клітини того ж виду.

Гліфосат-стійка трансгенна рослина, вироблена способом за винаходом, може бути розмножена, наприклад, шляхом схрещування її з другою рослиною, так щоб щонайменше частина потомства після схрещування показувала толерантність до гліфосату.

Винахід також стосується способів селективної боротьби з бур'яном на полі з рослинною культурою, які включають засадження поля насінням або рослинами культури, які є гліфосат-толерантними внаслідок трансформації їх геном, кодуючим гліфосат-N-ацетилтрансферазу, і нанесення на культуру і бур'ян на цьому полі достатньої кількості гліфосату для боротьби з бур'яном без суттєвого впливу на культуру.

Винахід також стосується способів боротьби з бур'яном на полі і запобігання виникненню гліфосат-стійкого бур'яну на полі з рослинною культурою, які включають засадження поля насінням або рослинами культури, які є гліфосат-

толерантними внаслідок трансформації їх геном, кодуючим гліфосат-N-ацетилтрансферазу, і геном, кодуючим поліпептид, який надає толерантність до гліфосату за допомогою іншого механізму, такий як гліфосат-толерантна 5-енолпірувілшикімат-3-фосфат-синтаза і/або гліфосат-толерантна гліфосат-оксидо-редуктаза, і нанесення на культуру і бур'ян на цьому полі достатньої кількості гліфосату для боротьби з бур'яном без суттєвого впливу на культуру.

В наступному варіанті здійснення винахід стосується способів боротьби з бур'яном на полі і запобігання виникненню гербіцид-стійкого бур'яну на полі з рослинною культурою, які включають засадження поля насінням або рослинами культури, які є гліфосат-толерантними внаслідок трансформації їх геном, кодуючим гліфосат-N-ацетилтрансферазу, геном, кодуючим поліпептид, який надає толерантність до гліфосату за допомогою іншого механізму, такий як гліфосат-толерантна 5-енолпірувілшикімат-3-фосфат-синтаза і/або гліфосат-толерантна гліфосат-оксидо-редуктаза, і геном, кодуючим поліпептид, який надає толерантність до додаткового гербіциду, такий як мутована гідроксифенілпіруватдіоксигеназа, сульфонамід-толерантна ацетолактатсинтаза, сульфонамід-толерантна синтаза ацетогідроксикислоти, імідазолінон-толерантна ацетолактатсинтаза, імідазолінон-толерантна синтаза ацетогідроксикислоти, фосфінотрицин-ацетилтрансфераза і мутована протопорфіриноген-оксидаза, і нанесення на культуру і бур'ян на цьому полі достатньої кількості гліфосату і додаткового гербіциду, такого як інгібітор гідроксифенілпіруватдіоксигенази, сульфонамід, імідазолінон, біалафос, фосфінотрицин, азафенідин, бутафенацил, сульфосат, глуфосінат та Protox-інгібітор для боротьби з бур'яном без суттєвого впливу на культуру.

Винахід також стосується способів боротьби з бур'яном на полі і запобігання виникненню гербіцид-стійкого бур'яну на полі з рослинною культурою, які включають засадження поля насінням або рослинами культури, які є гліфосат-толерантними внаслідок трансформації їх геном, кодуючим гліфосат-N-ацетилтрансферазу, і геном, кодуючим поліпептид, який надає толерантність до додаткового гербіциду, такий як мутована гідроксифенілпіруватдіоксигеназа, сульфонамід-толерантна ацетолактатсинтаза, сульфонамід-толерантна синтаза ацетогідроксикислоти, імідазолінон-толерантна ацетолактатсинтаза, імідазолінон-толерантна синтаза ацетогідроксикислоти, фосфінотрицин-ацетилтрансфераза і мутована протопорфіриноген-оксидаза, і нанесення на культуру і бур'ян на цьому полі достатньої кількості гліфосату і додаткового гербіциду, такого як інгібітор гідроксифенілпіруватдіоксигенази, сульфонамід, імідазолінон, біалафос, фосфінотрицин, азафенідин, бутафенацил, сульфосат, глуфосінат та протокс-інгібітор для боротьби з бур'яном без суттєвого впливу на культуру.

Винахід також стосується способів отримання генетично трансформованої рослини, що є толерантною до гліфосату, які включають інсертуван-

ня в геном рослинної клітини рекомбінантної двоспиральної молекули ДНК, що містить: (i) промотор, який функціонує в рослинних клітинах, викликаючи вироблення РНК-послідовності; (ii) структурну ДНК-послідовність, яка викликає вироблення РНК-послідовності, що кодує GAT; і (iii) 3'-нетрансльовану область, яка функціонує в рослинних клітинах, викликаючи приєднання відрізка з поліаденілових нуклеотидів до 3'-закінчення РНК-послідовності; при цьому промотор є гетерологічним до структурної ДНК-послідовності і адаптованим таким чином, щоб викликати достатню експресію закодованого поліпептиду для підвищення толерантності до гліфосату рослинної клітини, трансформованої молекулою ДНК; отримання трансформованої рослинної клітини; і регенерацію з трансформованої рослинної клітини генетично трансформованої рослини, що має підвищену толерантність до гліфосату.

Винахід також стосується способів отримання культури, які включають вирощування рослини культури, яка є гліфосат-толерантною внаслідок трансформації її геном, кодуючим гліфосат-N-ацетилтрансферазу, в таких умовах, щоб рослина культури продукувала культуру; і збирання врожаю культури, отриманої з рослини культури. Ці способи часто включають нанесення гліфосату на рослину культури з концентрацією, ефективною для боротьби з бур'яном. Приклади рослин культур включають бавовну, кукурудзу і сою.

Винахід також стосується комп'ютерів, носіїв, що зчитуються комп'ютерами, і інтегрованих систем, включаючи бази даних, складені з записів послідовностей, що включають ланцюги символів, відповідні до SEQ ID NOS: 1-514. Такі інтегровані системи можуть (необов'язково) включати один або більше наборів команд щодо селекції, вирівнювання, трансляції, зворотної трансляції або порівняння одного або більше ланцюгів символів, відповідних до SEQ ID NOS: 1-514, один з одним і/або з будь-якою додатковою нуклеїновокислотою або амінокислотою послідовністю.

Фіг.1 показує N-ацетилювання гліфосату, каталізоване гліфосат-N-ацетилтрансферазою ("GAT").

Фіг.2 ілюструє мас-спектроскопічне виявлення N-ацетилгліфосату, виробленого типовою культурою *Bacillus*, яка експресує нативну GAT-активність.

Фіг.3 є таблицею, що ілюструє відносну ідентичність між GAT-послідовностями, ізольованими з різних штамів бактерій, і yifl з *Bacillus subtilis*.

Фіг.4 являє собою карту плазмід рMAXY2120 для експресії і очищення GAT-ферменту з культур *E.coli*.

Фіг.5 являє собою результати мас-спектрометрії, що показують підвищення продукції N-ацетилгліфосату з часом в типовій реакційній суміші GAT-ферменту.

Фіг.6 являє собою графік кінетичних даних GAT-ферменту, з яких було розраховано  $K_m=2,9\text{mM}$  для гліфосату.

Фіг.7 являє собою графік кінетичних даних, отриманих з даних Фіг.6, з яких було розраховано  $K_m=2\text{mM}$  для ацетил-CoA.

Фіг.8 являє собою схему, що описує розщеплення гліфосату в ґрунті через AMPA-шлях.

Фіг.9 являє собою схему, що описує саркозинний шлях розщеплення гліфосату.

Фіг.10 показує матрицю BLOSUM62.

Фіг.11 являє собою карту плазмід рMAXY2190.

Фіг.12 показує конструкцію Т-ДНК з маркером *gat*, що селектується.

Фіг.13 показує вектор експресії дріжджів з маркером *gat*, що селектується.

Цей винахід стосується нового класу ферментів, що проявляють активність N-ацетилтрансферази. В одному з аспектів винахід стосується нового класу ферментів, що здатні ацетилювати гліфосат і аналоги гліфосату, наприклад, ферментів, що мають активність гліфосат-N-ацетилтрансферази (GAT). Такі ферменти відрізняються здатністю ацетилювати вторинний амін сполуки. В деяких аспектах винаходу сполука є гербіцидом, таким як гліфосат, схематично показаний на Фіг.1. Сполука також може бути аналогом гліфосату або метаболічним продуктом розщеплення гліфосату, як, наприклад, амінометилфосфонова кислота. Хоча ацетилювання гліфосату є головним каталітичним етапом в одному з метаболічних шляхів катаболізму гліфосату, ферментативне ацетилювання гліфосату природними, ізольованими або рекомбінантними ферментами раніш не було описано. Таким чином, нуклеїнові кислоти і поліпептиди за винаходом створюють новий біохімічний шлях надання стійкості до гербіцидів.

В одному з аспектів винахід стосується нових генів, що кодують GAT-поліпептиди. Ізольовані і рекомбінантні GAT-поліпептиди, які відповідають природним поліпептидам, а також рекомбінантні і сконструйовані, наприклад, диверсифіковані, GAT-поліпептиди є рисою цього винаходу. Прикладами GAT-поліпептидів є послідовності SEQ ID NOS: 1-5 і 11-262. Специфічні послідовності GAT-поліпептидів і GAT-поліпептидів наведені лише у якості прикладів для ілюстрації винаходу і не обмежують класу GAT-поліпептидів і GAT-поліпептидів, описаного і/або заявленого тут.

Винахід також стосується способів вироблення і селекції диверсифікованих бібліотек для отримання додаткових GAT-поліпептидів, включаючи поліпептиди, що кодують GAT-поліпептиди з покращеними і/або підвищеними характеристиками, наприклад, зміненою  $K_m$  до гліфосату, підвищеною швидкістю каталізу, підвищеною стабільністю і т.д., на основі селекції поліпептидного компонента бібліотеки на нові або покращені активності, описані тут. Такі поліпептиди особливо корисні для застосування у виробленні гліфосат-стійких трансгенних рослин.

GAT-поліпептиди за винаходом показують нову ферментативну активність. Зокрема, ферментативне ацетилювання синтетичного гербіциду гліфосату до цього винаходу не було описано. Таким чином, описані тут поліпептиди, зокрема, приклади SEQ ID NOS: 6-10 і 263-514, визначають новий біохімічний шлях детоксифікації гліфо-

сату, що є функціональним *in vivo*, наприклад, в рослинах.

Таким чином, нуклеїнові кислоти і поліпептиди за винаходом мають значну корисність для вироблення гліфосат-стійких рослин шляхом утворення нових нуклеїнових кислот, поліпептидів і біохімічних шляхів для конструювання гербіцидної селективності в трансгенних рослинах.

Перед детальним описом цього винаходу слід пояснити, що цей винахід не обмежений окремими композиціями або біологічними системами, які, звичайно, можуть варіюватися. Слід також зазначити, що використана тут термінологія призначена лише для опису окремих варіантів здійснення і не є обмежуючою. Використані в цьому описі і в формулі винаходу форми однини включають множинні об'єкти, якщо в контексті не вказано інше. Так, наприклад, посилання на "прилад" включає комбінацію двох або більше таких приладів, посилання на "конструкцію злиття генів" включає суміші конструкцій і т.п.

Якщо не вказано інше, всі технічні і наукові терміни, використані тут, мають загальноприйнятні у відповідній області значення, знайомі фахівцю середньої кваліфікації. Хоча на практиці для випробування цього винаходу можуть бути використані будь-які способи й матеріали, подібні або еквівалентні описаним тут, нижче описані деякі специфічні приклади відповідних матеріалів і способів.

В описі і формулі цього винаходу буде використана така термінологія відповідно до визначень, наведених нижче.

Вираз "гліфосат" в цьому винаході включає будь-яку гербіцидно ефективну форму N-фосфометилгліцину (в тому числі будь-яку його сіль) та інші форми, які утворюють аніон гліфосату в рослинах. Вираз "аналог гліфосату" стосується будь-якого структурного аналога гліфосату, який здатен інгібувати EPSPS у такій мірі, щоб аналог гліфосату був гербіцидно ефективним.

Вираз "активність гліфосат-N-ацетилтрансферази" або "GAT-активність" тут стосується здатності каталізувати ацетилювання вторинної аміногрупи гліфосату, як показано, наприклад, на Фіг.1. "Гліфосат-N-ацетилтрансфераза" або "GAT" означає фермент, який каталізує ацетилювання аміногрупи гліфосату, аналога гліфосату і/або первинного метаболіту гліфосату (тобто, AMPA або саркозину). В деяких переважних варіантах здійснення винаходу GAT здатна переносити ацетильну групу з ацетил-CoA (ацетил-коферменту А) до вторинної аміногрупи гліфосату і первинного аміну AMPA. Типові GAT, описані тут, є активними в діапазоні pH 5-9, з оптимальною активністю в діапазоні pH 6,5-8,0. Активність може бути оцінена з використанням різних кінетичних параметрів, добре відомих у відповідній області, наприклад,  $K_{кат}$ ,  $K_m$  і  $k_{кат}/K_m$ . Ці кінетичні параметри можуть бути визначені, як описано нижче в Прикладі 7.

Вирази "полінуклеотид", "нуклеотидна послідовність" і "нуклеїнова кислота" тут означають полімер з нуклеотидів (А, С, Т, U, G і т.д. або природні або штучні нуклеотидні аналоги), напри-

клад, ДНК або РНК, або їх представлення, таке як ланцюг символів і т.п., у залежності від контексту. Даний полінуклеотид або комплементарний полінуклеотид може бути визначений з будь-якої встановленої нуклеотидної послідовності.

Подібним чином, "амінокислотна послідовність" являє собою полімер з амінокислот (протеїн, поліпептид і т.д.) або ланцюг символів, що відображає амінокислотний полімер, у залежності від контексту. Вирази "протеїн", "поліпептид" і "пептид" використовуються тут взаємозамінно.

Полінуклеотид, поліпептид або інший компонент є "ізолюваним", якщо він частково або повністю відокремлений від компонентів, з якими він є звичайно асоційованим (інші протеїни, нуклеїнові кислоти, клітини, синтетичні реагенти і т.д.). Нуклеїнова кислота або поліпептид є "рекомбінантним", якщо він є штучним або сконструйованим, або похідним з штучного протеїну або нуклеїнової кислоти. Наприклад, полінуклеотид, вставлений в вектор або будь-яку іншу гетерологічну локацію, наприклад, в геном рекомбінантного організму, так, що він не асоційований з нуклеотидними послідовностями, які звичайно фланкують полінуклеотид в його природному стані, є рекомбінантним полінуклеотидом. Протеїн, що експресується *in vivo* або *in vitro* з рекомбінантного полінуклеотиду, є прикладом рекомбінантного поліпептиду. Подібним чином, полінуклеотидна послідовність, яка не зустрічається в природі, наприклад, варіант природного гена, є рекомбінантною.

Вирази "поліпептид гліфосат-N-ацетилтрансферази" і "GAT-поліпептид" використовуються тут взаємозамінно для будь-якого поліпептиду з сімейства нових поліпептидів, що пропонуються тут.

Вирази "полінуклеотид гліфосат-N-ацетилтрансферази" і "GAT-полінуклеотид" використовуються тут взаємозамінно для полінуклеотиду, що кодує GAT-поліпептид.

"Підпослідовність" або "фрагмент" являє собою будь-яку частину цілої послідовності.

Нумерація амінокислотного або нуклеотидного полімера відповідає нумерації вибраного амінокислотного полімера або нуклеїнової кислоти, коли положення даного мономерного компонента (амінокислотного залишку, вбудованого нуклеотиду і т.д.) полімера відповідає такому самому положенню залишку у вибраному еталонному поліпептиді або полінуклеотиді.

Вектор являє собою композицію для полегшення трансдукції клітини вибраною нуклеїновою кислотою або експресії нуклеїнової кислоти в клітині. Вектори включають, наприклад, плазмідні, косміди, віруси, YACs (дріжджові штучні хромосоми), бактерії, полілізин, вектори інтеграції хромосом, епісомні вектори і т.д.

"По суті повна довжина полінуклеотидної або амінокислотної послідовності" означає щонайменше ~70%, звичайно щонайменше ~80%, або типово ~90% послідовності або більше.

Вираз "антитіло" тут означає протеїн, що містить один або більше поліпептидів, по суті або частково закодованих генами імуноглобулінів або фрагментами генів імуноглобулінів. Гени імуногло-



лобулінів, що впізнаються, включають гени константних областей каппа, лямбда, альфа, гамма, дельта, епсилон і мію, а також міради генів варіабельних областей імуноглобулінів. Легкі ланцюги класифікуються як каппа або лямбда. Важкі ланцюги класифікуються як гамма, мію, альфа, дельта або епсилон, які, в свою чергу, визначають класи імуноглобулінів, IgG, IgM, IgA, IgD і IgE, відповідно. Типовий структурний елемент імуноглобуліну (антитіла) включає тетрамер. Кожний тетрамер складається з двох ідентичних пар поліпептидних ланцюгів, де кожна пара має один "легкий" (близько 25кДа) і один "важкий" ланцюг (близько 50-70кДа). N-закінчення кожного ланцюга визначає варіабельну область розміром близько 100-110 або більше амінокислот, у першу чергу відповідальних за впізнавання антигенів. Вирази "варіабельний легкий ланцюг" (VL) і "варіабельний важкий ланцюг" (VH) стосуються цих легких і важких ланцюгів, відповідно. Антитіла існують як повні імуноглобуліни або як декілька добре визначених фрагментів, отриманих шляхом ферментаційного розщеплення різними пептидазами. Так, наприклад, пепсин розщеплює антитіло нижче дисульфідних зв'язків у шарнірній області, утворюючи фрагмент F(ab)'2, димер Fab, що сам є легким ланцюгом, приєднаним до VH-CH1 дисульфідним зв'язком. F(ab)'2 може бути відновлений в м'яких умовах з розривом дисульфідного зв'язку в шарнірній області, що перетворює димер F(ab)'2 на мономер Fab'. Мономер Fab' по суті являє собою мономер Fab з частиною шарнірної області (більш детальний опис інших фрагментів антитіл див. у *Fundamental Immunology*, 4<sup>th</sup> Edition, W.E. Paul (ed.), Raven Press, N.Y. (1998)). Хоча різні фрагменти антитіл визначені на основі ферментаційного розщеплення повного антитіла, фахівцю буде зрозуміло, що такі Fab'-фрагменти можуть бути синтезовані *de novo* хімічним шляхом або з використанням технології рекомбінантних ДНК. Таким чином, вираз "антитіло" тут також включає фрагменти антитіл, отримані шляхом модифікації повних антитіл або синтезовані *de novo* з використанням технології рекомбінантних ДНК. Антитіла включають одностанові антитіла, в тому числі одностанові Fv (sFv)-антитіла, в яких варіабельний важкий і варіабельний легкий ланцюг з'єднані між собою (безпосередньо або пептидним лінкером) з утворенням суцільного поліпептиду.

"Хлоропласт-транзитний пептид" являє собою амінокислотну послідовність, яка трансклюється разом з протеїном і направляє протеїн у хлоропласт або інші типи пластид, присутні у клітині, в якій протеїн виробляється. Вираз "хлоропласт-транзитна послідовність" стосується нуклеотидної послідовності, яка кодує хлоропласт-транзитний пептид.

"Сигнальний пептид" являє собою амінокислотну послідовність, яка трансклюється разом з протеїном і направляє протеїн у секреторну систему (Chrispeels, J.J., (1991) *Ann. Rev. Plant Phys. Plant Mol. Biol.* 42:21-53). Якщо протеїн повинен бути направлений у вакуоль, може бути додатково використаний вакуолярний таргетинг-сигнал (вище), або, якщо в ендоплазматичну сітку, - сиг-

нал утримання в ендоплазматичній сітці (вище). Якщо протеїн повинен бути направлений в ядро, будь-який присутній сигнальний пептид повинен бути видалений, а замість нього введений сигнал ядерної локалізації (Raikhel, N. (1992) *Plant Phys.* 100: 1627-1632).

Вирази "диверсифікація" і "різноманітність" стосовно полінуклеотидів означають вироблення множини модифікованих форм батьківського полінуклеотиду, або множини батьківських полінуклеотидів. У випадку, коли полінуклеотид кодує поліпептид, різноманітність в нуклеотидній послідовності полінуклеотиду може привести до різноманітності у відповідному закодованому поліпептиді, наприклад, до різноманітного пулу полінуклеотидів, кодуєчих множину поліпептидних варіантів. У деяких варіантах здійснення виноходу ця різноманітність послідовностей використовується для скринінгу/селекції бібліотек диверсифікованих полінуклеотидів на варіанти з бажаними функціональними атрибутами, наприклад, полінуклеотид, що кодує GAT-поліпептид з покращеними функціональними характеристиками.

Вираз "кодуючий" означає здатність нуклеотидної послідовності кодувати одну або більше амінокислот. Це поняття не потребує стартового або стоп-кодона. Амінокислотна послідовність може бути закодована в одній з шести різних рамок зчитування, що утворюються полінуклеотидною послідовністю і її комплементом.

Вираз "штучний варіант" тут означає поліпептид, що має GAT-активність, закодований модифікованим GAT-полінуклеотидом, наприклад, модифікованою формою будь-якої з SEQ ID NOS: 1-5 і 11-262, або природним GAT-полінуклеотидом, ізольованим з організму. Модифікований полінуклеотид, з якого шляхом експресії в придатному хазяїні виробляють штучний варіант, отримують модифікують GAT-полінуклеотиду через втручання людини.

Вираз "нуклеїновокислотна конструкція" або "полінуклеотидна конструкція" означає молекулу нуклеїнової кислоти, одно- або двоспиральну, яка є ізольованою з природного гена або була модифікована таким чином, що вона містить сегменти нуклеїнових кислот у порядку, який в природі не існує. Вираз "нуклеїновокислотна конструкція" є синонімом до виразу "касета експресії", якщо нуклеїновокислотна конструкція містить контрольні послідовності, потрібні для експресії кодуючої послідовності за цим винаходом.

Вираз "контрольні послідовності" тут включає всі компоненти, що є необхідними або переважними для експресії поліпептиду за цим винаходом. Кожна контрольна послідовність може бути нативною або чужорідною до нуклеотидної послідовності, що кодує поліпептид. Такі контрольні послідовності включають, але не обмежуються, лідерну послідовність, поліаденілаційну послідовність, пропептидну послідовність, промотор, послідовність сигнального пептиду і термінатор транскрипції. Як мінімум, контрольні послідовності включають промотор і транскрипційні і трансляційні стоп-сигнали. Контрольні послідовності можуть бути забезпечені лінкерами для введення

специфічних рестрикційних сайтів, полегшуючих лігування контрольних послідовностей з кодуючою областю нуклеотидної послідовності, яка кодує поліпептид.

Вираз "оперативно зв'язаний" тут означає конфігурацію, в якій контрольна послідовність розміщена у відповідному положенні відносно кодуючої послідовності ДНК-послідовності, так що контрольна послідовність наставляє експресію поліпептиду.

Вираз "кодуюча послідовність" тут включає нуклеотидну послідовність, яка безпосередньо визначає амінокислотну послідовність її протеїнового продукту. Межі кодуючої послідовності загалом визначаються відкритою рамкою зчитування, яка звичайно починається стартовим кодоном АТГ. Типова кодуюча послідовність включає ДНК, кДНК і/або рекомбінантну нуклеотидну послідовність.

В цьому контексті вираз "експресія" включає будь-який етап вироблення поліпептиду, включаючи, але не тільки, транскрипцію, трансляцію, пост-трансляційну модифікацію і секрецію.

В цьому контексті вираз "вектор експресії" включає молекулу ДНК, лінійну або циклічну, яка містить сегмент, кодує поліпептид за винаходом, і яка є оперативно зв'язаною з додатковими сегментами, що забезпечують її транскрипцію.

Вираз "клітина-хазяїн" тут включає будь-який тип клітин, схильний до трансформації нуклеїновокислотною конструкцією.

Вираз "рослина" включає цілі рослини, паросткові вегетативні органи/структури (наприклад, листя, стебла і бульби), корені, квітки і квіткові органи/структури (наприклад, приквітники, чашолистки, пелюстки, тичинки, плодолистки, пильовики і насінні зачатки), насіння (включаючи зародки, ендосперм і насінневу оболонку) і фрукти (зрілу зав'язь), рослинну тканину (наприклад, васкулярну тканину, покривну тканину і т.п.) і клітини (наприклад, замикаючі клітини, яйцеклітини, трихоми і т.п.) та їх потомство. Клас рослин, які можуть бути використані в способі за винаходом, загалом включає широкий клас вищих та нижчих рослин, піддатливих технологіям трансформації, включаючи покритонасінні (односім'ядольні і двосім'ядольні рослини), голонасінні, папороті і багатоклітинні водорості. Вираз включає рослини різних рівнів плідності, в тому числі анеуплоїдні, поліплоїдні, диплоїдні, гаплоїдні та гемізиготні.

Вираз "гетерологічний" тут описує таке взаємовідношення між двома або більше елементами, яке показує, що ці елементи звичайно не зустрічаються в природі у близькій спорідненості один до одного. Так, наприклад, поліпептидна послідовність "гетерологічна" організму або другій поліпептидній послідовності, якщо вона походить з іншого виду, або, якщо вона походить з того ж виду, то вона є модифікованою порівняно з оригінальною формою. Наприклад, вираз "промотор, оперативно зв'язаний з гетерологічною кодуючою послідовністю" стосується кодуючої послідовності з виду, відмінного від того, до якого належить промотор, або, якщо з того ж виду, - кодуючої послідовності, яка в природі не асоційована з цим промотором (наприклад, гене-

тично сконструйована кодуюча послідовність або аллель з іншого екотипу чи різновиду). Прикладом гетерологічного поліпептиду є поліпептид, експресований з рекомбінантного поліпептиду в трансгенному організмі. Гетерологічні поліпептиди і поліпептиди є формами рекомбінантних молекул.

Крім того, далі визначено або описано іншим способом багато додаткових термінів.

Гліфосат-N-ацетилтрансферази

В одному з аспектів винахід стосується нового сімейства ізоляованих або рекомбінантних ферментів, що називаються тут " гліфосат-N-ацетилтрансферазами", "GATs" або "GAT-ферментами". GATs є ферментами, що мають GAT-активність, переважно достатню активність для надання деякого ступеню толерантності до гліфосату трансгенній рослині, сконструйованій зі здатністю експресувати GAT. Деякі приклади GATs включають GAT-поліпептиди, детальніше описані нижче.

Звичайно, GAT-опосереднена толерантність до гліфосату є комплексною функцією GAT-активності, рівнів експресії GAT в транс генній рослині, конкретної рослини, природи і режиму нанесення гербіциду і т.д. Досвідчений фахівець може без надмірного експериментування визначити рівень GAT-активності, необхідний для надання толерантності до гліфосату в конкретній ситуації.

GAT-активність може бути оцінена з використанням звичайних кінетичних параметрів  $K_{кат}$ ,  $K_m$  і  $K_{кат}/K_m$ .  $K_{кат}$  можна представити як міру швидкості ацетилювання, зокрема, при високих концентраціях субстрату,  $K_m$  є мірою афінності GAT до її субстратів (наприклад, ацетил-CoA і гліфосату), а  $K_{кат}/K_m$  є мірою каталітичної ефективності, яка приймає до уваги і афінність до субстрату, і каталітичну швидкість; цей параметр особливо важливий в ситуації, коли концентрація субстрату, щонайменше частково, є фактором, обмежуючим швидкість. Загалом, GAT з вищими параметрами  $K_{кат}$  або  $K_{кат}/K_m$  є більш ефективним каталізатором, ніж інша GAT з нижчими параметрами  $K_{кат}$  або  $K_{кат}/K_m$ . GAT з нижчим  $K_m$  є більш ефективним каталізатором, ніж інша GAT з вищим  $K_m$ . Таким чином, для визначення ефективності однієї GAT порівняно з іншою можна порівняти кінетичні параметри для цих двох ферментів. Відносна важливість  $K_{кат}$ ,  $K_m$  і  $K_{кат}/K_m$  буде змінюватися в залежності від очікуваної ситуації, в якій GAT буде функціонувати, наприклад, очікуваної ефективної концентрації гліфосату відносно  $K_m$  до гліфосату. GAT-активність також може бути оцінена на основі будь-якої з ряду функціональних характеристик, наприклад, стабільності, схильності до інгібування або активації іншими молекулами і т.д.

Поліпептид гліфосат-N-ацетилтрансферази

В одному з аспектів винахід стосується нового сімейства ізоляованих або рекомбінантних поліпептидів, що називаються тут "поліпептидами гліфосат-N-ацетилтрансферази" або "GAT-поліпептидами". GAT-поліпептиди відрізняються їх структурною подібністю до нового сімейства GATs. Багато, але не всі GAT-поліпептиди являють собою GATs. Відмінність складається в тому,

що GATs визначаються на основі функції, а GAT-поліпептиди визначаються на основі структури. Одна підгрупа GAT-поліпептидів включає ті GAT-поліпептиди, які мають GAT-активність, переважно на рівні, який забезпечить надання стійкості до гліфосату трансгенній рослині, яка експресує цей протеїн на ефективному рівні. Деякі GAT-поліпептиди, переважні для надання толерантності до гліфосату, мають  $k_{cat}$  щонайменше  $1 \text{ хв}^{-1}$ , більш переважно  $10 \text{ хв}^{-1}$ ,  $100 \text{ хв}^{-1}$  або  $1000 \text{ хв}^{-1}$ . Інші GAT-поліпептиди, переважні для надання толерантності до гліфосату, мають  $K_m$  не більше ніж  $100 \text{ мМ}$ , більш переважно не більше ніж  $10 \text{ мМ}$ ,  $1 \text{ мМ}$ , or  $0,1 \text{ мМ}$ . Інші GAT-поліпептиди, переважні для надання толерантності до гліфосату, мають  $k_{cat}/K_m$  щонайменше  $1 \text{ мМ}^{-1} \text{ хв}^{-1}$  або більше, переважно щонайменше  $10 \text{ мМ}^{-1} \text{ хв}^{-1}$ ,  $100 \text{ мМ}^{-1} \text{ хв}^{-1}$ ,  $1000 \text{ мМ}^{-1} \text{ хв}^{-1}$  або  $10000 \text{ мМ}^{-1} \text{ хв}^{-1}$ .

Зразки GAT-поліпептидів з різних бактеріальних штамів були ізолювані і описані. Один із зразків мономерного GAT-поліпептиду, що був ізолюваний і описаний, має радіус молекули близько  $17 \text{ кДа}$ . Зразок GAT-ферменту, ізолюваний зі штаму *B. licheniformis*, SEQ ID NO:7, показує  $K_m$  до гліфосату близько  $2,9 \text{ мМ}$ , а  $K_m$  до ацетил-CoA - близько  $2 \text{ мМ}$ , при цьому  $k_{cat}$  дорівнює  $6/\text{хвилину}$ .

Вираз "GAT-поліпептид" стосується будь-якого поліпептиду, що включає амінокислотну послідовність, яка може бути оптимально вирівняна з амінокислотою послідовністю, вибраною з групи, що включає SEQ ID NOS: 6-10 і 263-514, з отриманням показника подібності щонайменше 430 при використанні матриці BLOSUM62, штрафу за існування пропусків 11 і штрафу за розширення пропусків 1. Деякі аспекти винаходу стосуються GAT-поліпептидів, що включають амінокислотну послідовність, яка може бути оптимально вирівняна з амінокислотою послідовністю, вибраною з групи, що включає SEQ ID NOS: 6-10 і 263-514, з отриманням показника подібності щонайменше 440, 445, 450, 455, 460, 465, 470, 475, 480, 485, 490, 495, 500, 505, 510, 515, 520, 525, 530, 535, 540, 545, 550, 555, 560, 565, 570, 575, 580, 585, 590, 595, 600, 605, 610, 615, 620, 625, 630, 635, 640, 645, 650, 655, 660, 665, 670, 675, 680, 685, 690, 695, 700, 705, 710, 715, 720, 725, 730, 735, 740, 745, 750, 755 або 760 при використанні матриці BLOSUM62, штрафу за існування пропусків 11 і штрафу за розширення пропусків 1.

В одному з аспектів винахід стосується GAT-поліпептиду, що включає амінокислотну послідовність, яка може бути оптимально вирівняна з SEQ ID NO. 457 з отриманням показника подібності щонайменше 430 при використанні матриці BLOSUM62, штрафу за існування пропусків 11 і штрафу за розширення пропусків 1. Деякі аспекти винаходу стосуються GAT-поліпептидів, що включають амінокислотну послідовність, яка може бути оптимально вирівняна з SEQ ID NO. 457 з отриманням показника подібності щонайменше 440, 445, 450, 455, 460, 465, 470, 475, 480, 485, 490, 495, 500, 505, 510, 515, 520, 525, 530, 535, 540, 545, 550, 555, 560, 565, 570, 575, 580, 585, 590, 595, 600, 605, 610, 615, 620, 625, 630, 635,

640, 645, 650, 655, 660, 665, 670, 675, 680, 685, 690, 695, 700, 705, 710, 715, 720, 725, 730, 735, 740, 745, 750, 755 або 760 при використанні матриці BLOSUM62, штрафу за існування пропусків 11 і штрафу за розширення пропусків 1.

В одному з аспектів винахід стосується GAT-поліпептиду, що включає амінокислотну послідовність, яка може бути оптимально вирівняна з SEQ ID NO. 445 з отриманням показника подібності щонайменше 430 при використанні матриці BLOSUM62, штрафу за існування пропусків 11 і штрафу за розширення пропусків 1. Деякі аспекти винаходу стосуються GAT-поліпептидів, що включають амінокислотну послідовність, яка може бути оптимально вирівняна з SEQ ID NO. 445 з отриманням показника подібності щонайменше 440, 445, 450, 455, 460, 465, 470, 475, 480, 485, 490, 495, 500, 505, 510, 515, 520, 525, 530, 535, 540, 545, 550, 555, 560, 565, 570, 575, 580, 585, 590, 595, 600, 605, 610, 615, 620, 625, 630, 635, 640, 645, 650, 655, 660, 665, 670, 675, 680, 685, 690, 695, 700, 705, 710, 715, 720, 725, 730, 735, 740, 745, 750, 755 або 760 при використанні матриці BLOSUM62, штрафу за існування пропусків 11 і штрафу за розширення пропусків 1.

В одному з аспектів винахід стосується GAT-поліпептиду, що включає амінокислотну послідовність, яка може бути оптимально вирівняна з SEQ ID NO:300 з отриманням показника подібності щонайменше 430 при використанні матриці BLOSUM62, штрафу за існування пропусків 11 і штрафу за розширення пропусків 1. Деякі аспекти винаходу стосуються GAT-поліпептидів, що включають амінокислотну послідовність, яка може бути оптимально вирівняна з SEQ ID NO:300 з отриманням показника подібності щонайменше 440, 445, 450, 455, 460, 465, 470, 475, 480, 485, 490, 495, 500, 505, 510, 515, 520, 525, 530, 535, 540, 545, 550, 555, 560, 565, 570, 575, 580, 585, 590, 595, 600, 605, 610, 615, 620, 625, 630, 635, 640, 645, 650, 655, 660, 665, 670, 675, 680, 685, 690, 695, 700, 705, 710, 715, 720, 725, 730, 735, 740, 745, 750, 755 або 760 при використанні матриці BLOSUM62, штрафу за існування пропусків 11 і штрафу за розширення пропусків 1.

Дві послідовності є "оптимально вирівняними", якщо вони вирівняні для підрахунку подібності з використанням вибраної матриці амінокислотного заміщення (наприклад, BLOSUM62), штрафу за існування пропусків і штрафу за розширення пропусків таким чином, щоб досягти найвищого показника, можливого для цієї пари послідовностей. Матриці амінокислотного заміщення та їх застосування для кількісної оцінки подібності між двома послідовностями добре відомі у відповідній області і описані, наприклад, у Dayhoff et al. (1978) "A model of evolutionary change in proteins." In "Atlas of Protein Sequence and Structure", Vol.5, Suppl. 3 (ed. M.O. Dayhoff), pp.345-352. Natl. Biomed. Res. Pound., Washington, DC and Henikoff et al. (1992) Proc. Natl. Acad. Sci. USA 89:10915-10919. Матрицю BLOSUM62 (Fig.10) часто використовують у якості стандартної матриці заміщення для підрахунку в протоколах вирівнювання послідовностей, таких як Gapped BLAST 2.0. Штраф за існування

пропусків накладається за введення одного амінокислотного пропуску в одну з послідовностей, що вирівнюються, а штраф за розширення пропусків накладається на кожне додаткове порожнє амінокислотне положення, введене у вже існуючий пропуск. Вирівнювання визначається амінокислотними положеннями в кожній послідовності, в яких вирівнювання починається і закінчується, і, можливо, ісерцією пропуску або множини пропусків в одній чи обох послідовностях, для того щоб досягти найвищого можливого показника. Хоча оптимальне вирівнювання і підрахунок можуть бути виконані вручну, цей процес полегшується шляхом використання комп'ютерного алгоритму вирівнювання, наприклад, Gapped BLAST 2.0, описаного у Altschul et al, (1997) *Nucleic Acids Res.* 25:3389-3402 і доступного на вебсайті Національного центру інформації по біотехнології (<http://www.ncbi.nlm.nih.gov>). Оптимальні вирівнювання, включаючи множинні вирівнювання, можуть бути отримані з використанням, наприклад, PSI-BLAST, доступного на <http://www.ncbi.nlm.nih.gov> і описаного у Altschul et al, (1997) *Nucleic Acids Res.* 25:3389-3402.

В амінокислотній послідовності, що є оптимально вирівняною з еталонною послідовністю, амінокислотний залишок "відповідає" тому положенню в еталонній послідовності, з яким залишок паруються при вирівнюванні. "Положення" позначаються номерами, які послідовно ідентифікують кожну амінокислоту в еталонній послідовності на основі її положення відносно N-закінчення. Наприклад, в SEQ ID NO:300 положення 1 є M, положення 2 є I, положення 3 є E, і т.д. Коли тестова послідовність є оптимально вирівняною з SEQ ID NO:300, той залишок в тестовій послідовності, який вирівнюється з E в положенні 3, називають "відповідаючим положенню 3" SEQ ID NO:300. Внаслідок делецій, інсерцій, усічень, злиттів і т.д., які можуть прийматися до уваги при визначенні оптимального вирівнювання, загалом номер амінокислотного залишку в тестовій послідовності, визначений простим підрахунком від N-закінчення, не обов'язково буде співпадати з номером його відповідного положення в еталонній послідовності. Наприклад, у випадку, коли є делеція у вирівняній тестовій послідовності, то в ній не буде амінокислоти, що відповідає положенню в еталонній послідовності на сайті делеції. Якщо є інсерція у вирівняній тестовій послідовності, ця інсерція не буде відповідати жодному амінокислотному положенню в еталонній послідовності. У випадку усічень або злиттів можуть бути присутні амінокислотні відрізки в еталонній або тестовій послідовності, які не відповідають жодній амінокислоті у відповідній послідовності.

Вираз "GAT-поліпептид" також стосується будь-якого поліпептиду, що включає амінокислотну послідовність, яка має щонайменше 40% ідентичності послідовності з амінокислотою послідовністю, вибраною з групи, що включає SEQ ID NOS: 6-10 і 263-514. Деякі аспекти винаходу стосуються GAT-поліпептидів, що включають амінокислотну послідовність, яка має щонайменше 60%, 70%, 80%, 90%, 92%, 95%, 96%, 97%, 98% або 99% ідентичності послідовності з амінокисло-

тню послідовністю, вибраною з групи, що включає SEQ ID NOS: 6-10 і 263-514.

Один з аспектів винаходу стосується GAT-поліпептиду, що включає амінокислотну послідовність, яка має щонайменше 40% ідентичності послідовності з SEQ ID NO. 457. Деякі аспекти винаходу стосуються GAT-поліпептидів, що включають амінокислотну послідовність, яка має щонайменше 60%, 70%, 80%, 90%, 92%, 95%, 96%, 97%, 98% або 99% ідентичності послідовності з SEQ ID NO. 457.

Один з аспектів винаходу стосується GAT-поліпептиду, що включає амінокислотну послідовність, яка має щонайменше 40% ідентичності послідовності з SEQ ID NO. 445. Деякі аспекти винаходу стосуються GAT-поліпептидів, що включають амінокислотну послідовність, яка має щонайменше 60%, 70%, 80%, 90%, 92%, 95%, 96%, 97%, 98% або 99% ідентичності послідовності з SEQ ID NO. 445.

Один з аспектів винаходу стосується GAT-поліпептиду, що включає амінокислотну послідовність, яка має щонайменше 40% ідентичності послідовності з SEQ ID NO. 300. Деякі аспекти винаходу стосуються GAT-поліпептидів, що включають амінокислотну послідовність, яка має щонайменше 60%, 70%, 80%, 90%, 92%, 95%, 96%, 97%, 98% або 99% ідентичності послідовності з SEQ ID NO. 300.

Вираз "GAT-поліпептид" також стосується будь-якого поліпептиду, що включає амінокислотну послідовність, яка має щонайменше 40% ідентичності послідовності з залишками 1-96 амінокислотної послідовності, вибраної з групи, що включає SEQ ID NOS: 6-10 і 263-514. Деякі аспекти винаходу стосуються поліпептидів, що включають амінокислотну послідовність, яка має щонайменше 60%, 70%, 80%, 90%, 92%, 95%, 96%, 97%, 98% або 99% ідентичності послідовності з залишками 1-96 амінокислотної послідовності, вибраної з групи, що включає SEQ ID NOS: 6-10 і 263-514.

Один з аспектів винаходу стосується GAT-поліпептиду, що включає амінокислотну послідовність, яка має щонайменше 40% ідентичності послідовності з залишками 1-96 SEQ ID NO. 457. Деякі аспекти винаходу стосуються GAT-поліпептидів, що включають амінокислотну послідовність, яка має щонайменше 60%, 70%, 80%, 90%, 92%, 95%, 96%, 97%, 98% або 99% ідентичності послідовності з залишками 1-96 SEQ ID NO. 457.

Один з аспектів винаходу стосується GAT-поліпептиду, що включає амінокислотну послідовність, яка має щонайменше 40% ідентичності послідовності з залишками 1-96 SEQ ID NO. 445. Деякі аспекти винаходу стосуються GAT-поліпептидів, що включають амінокислотну послідовність, яка має щонайменше 60%, 70%, 80%, 90%, 92%, 95%, 96%, 97%, 98% або 99% ідентичності послідовності з залишками 1-96 SEQ ID NO. 445.

Один з аспектів винаходу стосується GAT-поліпептиду, що включає амінокислотну послідовність, яка має щонайменше 40% ідентичності послідовності з залишками 1-96 SEQ ID NO. 300.

Деякі аспекти винаходу стосуються GAT-поліпептидів, що включають амінокислотну послідовність, яка має щонайменше 60%, 70%, 80%, 90%, 92%, 95%, 96%, 97%, 98% або 99% ідентичності послідовності з залишками 1-96 SEQ ID NO. 300.

Вираз "GAT-поліпептид" також стосується будь-якого поліпептиду, що включає амінокислотну послідовність, яка має щонайменше 40% ідентичності послідовності з залишками 51-146 амінокислотної послідовності, вибраної з групи, що включає SEQ ID NOS: 6-10 і 263-514. Деякі аспекти винаходу стосуються поліпептидів, що включають амінокислотну послідовність, яка має щонайменше 60%, 70%, 80%, 90%, 92%, 95%, 96%, 97%, 98% або 99% ідентичності послідовності з залишками 51-146 амінокислотної послідовності, вибраної з групи, що включає SEQ ID NOS: 6-10 і 263-514.

Один з аспектів винаходу стосується GAT-поліпептиду, що включає амінокислотну послідовність, яка має щонайменше 40% ідентичності послідовності з залишками 51-146 SEQ ID NO. 457. Деякі аспекти винаходу стосуються GAT-поліпептидів, що включають амінокислотну послідовність, яка має щонайменше 60%, 70%, 80%, 90%, 92%, 95%, 96%, 97%, 98% або 99% ідентичності послідовності з залишками 51-146 SEQ ID NO. 457.

Один з аспектів винаходу стосується GAT-поліпептиду, що включає амінокислотну послідовність, яка має щонайменше 40% ідентичності послідовності з залишками 51-146 SEQ ID NO. 445. Деякі аспекти винаходу стосуються GAT-поліпептидів, що включають амінокислотну послідовність, яка має щонайменше 60%, 70%, 80%, 90%, 92%, 95%, 96%, 97%, 98% або 99% ідентичності послідовності з залишками 51-146 SEQ ID NO. 445.

Один з аспектів винаходу стосується GAT-поліпептиду, що включає амінокислотну послідовність, яка має щонайменше 40% ідентичності послідовності з залишками 51-146 SEQ ID NO. 300. Деякі аспекти винаходу стосуються GAT-поліпептидів, що включають амінокислотну послідовність, яка має щонайменше 60%, 70%, 80%, 90%, 92%, 95%, 96%, 97%, 98% або 99% ідентичності послідовності з залишками 51-146 SEQ ID NO. 300.

Вираз "ідентичність" або "процентна ідентичність" стосовно окремої пари вирівняних амінокислотних послідовностей тут означає процентну ідентичність амінокислотних послідовностей, отриману аналізом ClustalW (версія W 1.8, що поставляється European Bioinformatics Institute, Cambridge, UK), яка включає підрахунок числа ідентичних збігів у вирівнюванні і ділення цього числа ідентичних збігів на більшу величину з (i) довжини вирівняних послідовностей, і (ii) 96, і використання таких стандартних параметрів ClustalW для досягнення повільних/точних вирівнювань за парами: штраф за відкриття пропуску - 10; штраф за розширення пропуску - 0,10; матриця маси протеїну - серії Gonnet; матриця маси ДНК - IUB; вирівнювання за парами з перемиканням "повільне/швидке" - SLOW or FULL Alignment.

В іншому аспекті винахід стосується ізольованого або рекомбінантного поліпептиду, що включає щонайменше 20, або, в альтернативі, 50, 75, 100, 125 або 140 суміжних амінокислот з амінокислотної послідовності, вибраної з групи, що включає SEQ ID NOS: 6-10 і 263-514.

В наступному аспекті винахід стосується ізольованого або рекомбінантного поліпептиду, що включає щонайменше 20, або, в альтернативі, 50, 100 або 140 суміжних амінокислот з SEQ ID NO: 457.

В наступному аспекті винахід стосується ізольованого або рекомбінантного поліпептиду, що включає щонайменше 20, або, в альтернативі, 50, 100 або 140 суміжних амінокислот з SEQ ID NO: 445.

В наступному аспекті винахід стосується ізольованого або рекомбінантного поліпептиду, що включає щонайменше 20, або, в альтернативі, 50, 100 або 140 суміжних амінокислот з SEQ ID NO: 300.

В наступному аспекті винахід стосується поліпептиду, що включає амінокислотну послідовність, вибрану з групи, що включає SEQ ID NOS: 6-10 і 263-514.

Деякі переважні GAT-поліпептиди за винаходом відрізняються таким чином. При оптимальному вирівнюванні з еталонною амінокислотною послідовністю, вибраною з групи, що включає SEQ ID NOS: 6-10 і 263-514, щонайменше 90% амінокислотних залишків в поліпептиді, що відповідають вказаним далі положенням, узгоджуються з такими обмеженнями: (a) в положеннях 2, 4, 15, 19, 26, 28, 31, 45, 51, 54, 86, 90, 91, 97, 103, 105, 106, 114, 123, 129, 139 і/або 145 амінокислотний залишок являє собою B1; і (b) в положеннях 3, 5, 8, 10, 11, 14, 17, 18, 24, 27, 32, 37, 38, 47, 48, 49, 52, 57, 58, 61, 62, 63, 68, 69, 79, 80, 82, 83, 89, 92, 100, 101, 104, 119, 120, 124, 125, 126, 128, 131, 143 і/або 144 амінокислотний залишок являє собою B2; при цьому B1 є амінокислотою, вибраною з групи, що включає A, I, L, M, F, W, Y і V, а B2 є амінокислотою, вибраною з групи, що включає R, N, D, C, Q, E, G, H, K, P, S і T. При використанні літер для позначення амінокислоти або амінокислотного залишку однолітерні позначення A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W і Y відповідають їх стандартним значенням, що звичайно використовуються у біотехнології і наведені тут в Таблиці 2.

Деякі переважні GAT-поліпептиди за винаходом відрізняються таким чином. При оптимальному вирівнюванні з еталонною амінокислотною послідовністю, вибраною з групи, що включає SEQ ID NOS: 6-10 і 263-514, щонайменше 80% амінокислотних залишків в поліпептиді, що відповідають вказаним далі положенням, узгоджуються з такими обмеженнями: (a) в положеннях 2, 4, 15, 19, 26, 28, 51, 54, 86, 90, 91, 97, 103, 105, 106, 114, 129, 139 і/або 145 амінокислотний залишок являє собою Z1; (b) в положеннях 31 і/або 45 амінокислотний залишок являє собою Z2; (c) в положеннях 8 і/або 89 амінокислотний залишок являє собою Z3; (d) в положеннях 82, 92, 101 і/або 120 амінокислотний залишок являє собою Z4; (e) в положеннях 3, 11, 27 і/або 79 амінокислотний

залишок являє собою Z5; (f) в положенні 123 амінокислотний залишок являє собою Z1 або Z2; (g) в положеннях 12, 33, 35, 39, 53, 59, 112, 132, 135, 140 i/або 146 амінокислотний залишок являє собою Z1 або Z3; (h) в положенні 30 амінокислотний залишок являє собою Z1 або Z4; (i) в положенні 6 амінокислотний залишок являє собою Z1 або Z6; (j) в положеннях 81 i/або 113 амінокислотний залишок являє собою Z2 або Z3; (k) в положеннях 138 i/або 142 амінокислотний залишок являє собою Z2 або Z4; (l) в положеннях 5, 17, 24, 57, 61, 124 i/або 126 амінокислотний залишок являє собою Z3 або Z4; (m) в положенні 104 амінокислотний залишок являє собою Z3 або Z5; (o) в положеннях 38, 52, 62 i/або 69 амінокислотний залишок являє собою Z3 або Z6; (p) в положеннях 14, 119 i/або 144 амінокислотний залишок являє собою Z4 або Z5; (q) в положенні 18 амінокислотний залишок являє собою Z4 або Z6; (r) в положеннях 10, 32, 48, 63, 80 i/або 83 амінокислотний залишок являє собою Z5 або Z6; (s) в положенні 40 амінокислотний залишок являє собою Z1, Z2 або Z3; (t) в положеннях 65 i/або 96 амінокислотний залишок являє собою Z1, Z3 або Z5; (u) в положеннях 84 i/або 115 амінокислотний залишок являє собою Z1, Z3 або Z4; (v) в положенні 93 амінокислотний залишок являє собою Z2, Z3 або Z4; (w) в положенні 130 амінокислотний залишок являє собою Z2, Z4 або Z6; (x) в положеннях 47 i/або 58 амінокислотний залишок являє собою Z3, Z4 або Z6; (y) в положеннях 49, 68, 100 i/або 143 амінокислотний залишок являє собою Z3, Z4 або Z5; (z) в положенні 131 амінокислотний залишок являє собою Z3, Z5 або Z6; (aa) в положеннях 125 i/або 128 амінокислотний залишок являє собою Z4, Z5 або Z6; (ab) в положенні 67 амінокислотний залишок являє собою Z1, Z3, Z4 або Z5; (ac) в положенні 60 амінокислотний залишок являє собою Z1, Z4, Z5 або Z6; i (ad) в положенні 37 амінокислотний залишок являє собою Z3, Z4, Z5 або Z6; при цьому Z1 є амінокислотою, вибраною з групи, що включає A, I, L, M і V; Z2 є амінокислотою, вибраною з групи, що включає F, W і Y; Z3 є амінокислотою, вибраною з групи, що включає N, Q, S і T; Z4 є амінокислотою, вибраною з групи, що включає R, H і K; Z5 є амінокислотою, вибраною з групи, що включає D і E; Z6 є амінокислотою, вибраною з групи, що включає C, G і P.

Деякі переважні GAT-поліпептиди за винаходом відрізняються таким чином. При оптимальному вирівнюванні з еталонною амінокислотою послідовністю, вибраною з групи, що включає SEQ ID NO:6-10 і 263-514, щонайменше 90% амінокислотних залишків в поліпептиді, що відповідають вказаним далі положенням, узгоджуються з такими обмеженнями: (a) в положеннях 1, 7, 9, 13, 20, 36, 42, 46, 50, 56, 64, 70, 72, 75, 76, 78, 94, 98, 107, 110, 117, 118, 121 i/або 141 амінокислотний залишок являє собою B1; i (b) в положеннях 16, 21, 22, 23, 25, 29, 34, 41, 43, 44, 55, 66, 71, 73, 74, 77, 85, 87, 88, 95, 99, 102, 108, 109, 111, 116, 122, 127, 133, 134, 136 i/або 137 амінокислотний залишок являє собою B2; при цьому B1 є амінокислотою, вибраною з групи, що включає A, I, L, M, F, W, Y і V; a B2 є амінокислотою, вибраною з

групи, що включає R, N, D, C, Q, E, G, H, K, P, S і T.

Деякі переважні GAT-поліпептиди за винаходом відрізняються таким чином. При оптимальному вирівнюванні з еталонною амінокислотою послідовністю, вибраною з групи, що включає SEQ ID NO:6-10 і 263-514, щонайменше 90% амінокислотних залишків в поліпептиді, що відповідають вказаним далі положенням, узгоджуються з такими обмеженнями: (a) в положеннях 1, 7, 9, 20, 36, 42, 50, 64, 72, 75, 76, 78, 94, 98, 110, 121 i/або 141 амінокислотний залишок являє собою Z1; (b) в положеннях 13, 46, 56, 70, 107, 117 i/або 118 амінокислотний залишок являє собою Z2; (c) в положеннях 23, 55, 71, 77, 88, i/або 109 амінокислотний залишок являє собою Z3; (d) в положеннях 16, 21, 41, 73, 85, 99 i/або 111 амінокислотний залишок являє собою Z4; (e) в положеннях 34 i/або 95 амінокислотний залишок являє собою Z5; (f) в положенні 22, 25, 29, 43, 44, 66, 74, 87, 102, 108, 116, 122, 127, 133, 134, 136 i/або 137 амінокислотний залишок являє собою Z6; при цьому Z1 є амінокислотою, вибраною з групи, що включає A, I, L, M і V; Z2 є амінокислотою, вибраною з групи, що включає F, W і Y; Z3 є амінокислотою, вибраною з групи, що включає N, Q, S і T; Z4 є амінокислотою, вибраною з групи, що включає R, H і K; Z5 є амінокислотою, вибраною з групи, що включає D і E; i Z6 є амінокислотою, вибраною з групи, що включає C, G і P.

Деякі переважні GAT-поліпептиди за винаходом відрізняються таким чином. При оптимальному вирівнюванні з еталонною амінокислотою послідовністю, вибраною з групи, що включає SEQ ID NO:6-10 і 263-514, щонайменше 80% амінокислотних залишків в поліпептиді, що відповідають вказаним далі положенням, узгоджуються з такими обмеженнями: (a) в положенні 2 амінокислотний залишок являє собою I або L; (b) в положенні 3 амінокислотний залишок являє собою E або D; (c) в положенні 4 амінокислотний залишок являє собою V, A або I; (d) в положенні 5 амінокислотний залишок являє собою K, R або N; (e) в положенні 6 амінокислотний залишок являє собою P або L; (f) в положенні 8 амінокислотний залишок являє собою N, S або T; (g) в положенні 10 амінокислотний залишок являє собою E або G; (h) в положенні 11 амінокислотний залишок являє собою D або E; (i) в положенні 12 амінокислотний залишок являє собою T або A; (j) в положенні 14 амінокислотний залишок являє собою E або K; (k) в положенні 15 амінокислотний залишок являє собою I або L; (l) в положенні 17 амінокислотний залишок являє собою H або Q; (m) в положенні 18 амінокислотний залишок являє собою R, C або K; (n) в положенні 19 амінокислотний залишок являє собою I або V; (o) в положенні 24 амінокислотний залишок являє собою Q або R; (p) в положенні 26 амінокислотний залишок являє собою L або I; (q) в положенні 27 амінокислотний залишок являє собою E або D; (r) в положенні 28 амінокислотний залишок являє собою A або V; (s) в положенні 30 амінокислотний залишок являє собою K, M або R; (t) в положенні 31 амінокислотний залишок являє собою Y або F; (u) в положенні 32 амінокислотний залишок являє

собою Е або G; (v) в положенні 33 амінокислотний залишок являє собою Т, А або S; (w) в положенні 35 амінокислотний залишок являє собою L, S або M; (x) в положенні 37 амінокислотний залишок являє собою R, G, Е або Q; (y) в положенні 38 амінокислотний залишок являє собою G або S; (z) в положенні 39 амінокислотний залишок являє собою Т, А або S; (aa) в положенні 40 амінокислотний залишок являє собою F, L або S; (ab) в положенні 45 амінокислотний залишок являє собою Y або F; (ac) в положенні 47 амінокислотний залишок являє собою R, Q або G; (ad) в положенні 48 амінокислотний залишок являє собою G або D; (ae) в положенні 49 амінокислотний залишок являє собою K, R, Е або Q; (af) в положенні 51 амінокислотний залишок являє собою I або V; (ag) в положенні 52 амінокислотний залишок являє собою S, C або G; (ah) в положенні 53 амінокислотний залишок являє собою I або T; (ai) в положенні 54 амінокислотний залишок являє собою А або V; (aj) в положенні 57 амінокислотний залишок являє собою H або N; (ak) в положенні 58 амінокислотний залишок являє собою Q, K, N або P; (al) в положенні 59 амінокислотний залишок являє собою А або S; (am) в положенні 60 амінокислотний залишок являє собою E, K, G, V або D; (an) в положенні 61 амінокислотний залишок являє собою H або Q; (ao) в положенні 62 амінокислотний залишок являє собою P, S або T; (ap) в положенні 63 амінокислотний залишок являє собою E, G або D; (aq) в положенні 65 амінокислотний залишок являє собою E, D, V або Q; (ar) в положенні 67 амінокислотний залишок являє собою Q, E, R, L, H або K; (as) в положенні 68 амінокислотний залишок являє собою K, R, Е або N; (at) в положенні 69 амінокислотний залишок являє собою Q або P; (au) в положенні 79 амінокислотний залишок являє собою E або D; (av) в положенні 80 амінокислотний залишок являє собою G або E; (aw) в положенні 81 амінокислотний залишок являє собою Y, N або F; (ax) в положенні 82 амінокислотний залишок являє собою R або H; (ay) в положенні 83 амінокислотний залишок являє собою E, G або D; (az) в положенні 84 амінокислотний залишок являє собою Q, R або L; (ba) в положенні 86 амінокислотний залишок являє собою А або V; (bb) в положенні 89 амінокислотний залишок являє собою Т або S; (be) в положенні 90 амінокислотний залишок являє собою L або I; (bd) в положенні 91 амінокислотний залишок являє собою I або V; (be) в положенні 92 амінокислотний залишок являє собою R або K; (bf) в положенні 93 амінокислотний залишок являє собою H, Y або Q; (bg) в положенні 96 амінокислотний залишок являє собою E, А або Q; (bh) в положенні 97 амінокислотний залишок являє собою L або I; (bi) в положенні 100 амінокислотний залишок являє собою K, R, N або E; (bj) в положенні 101 амінокислотний залишок являє собою K або R; (bk) в положенні 103 амінокислотний залишок являє собою А або V; (bl) в положенні 104 амінокислотний залишок являє собою D або N; (bm) в положенні 105 амінокислотний залишок являє собою L або M; (bn) в положенні 106 амінокислотний залишок являє собою L або I; (bo) в положенні 112 амінокислотний залишок

являє собою Т або I; (bp) в положенні 113 амінокислотний залишок являє собою S, Т або F; (bq) в положенні 114 амінокислотний залишок являє собою А або V; (br) в положенні 115 амінокислотний залишок являє собою S, R або A; (bs) в положенні 119 амінокислотний залишок являє собою K, Е або R; (bt) в положенні 120 амінокислотний залишок являє собою K або R; (bu) в положенні 123 амінокислотний залишок являє собою F або L; (bv) в положенні 124 амінокислотний залишок являє собою S або R; (bw) в положенні 125 амінокислотний залишок являє собою E, K, G або D; (bx) в положенні 126 амінокислотний залишок являє собою Q або H; (by) в положенні 128 амінокислотний залишок являє собою E, G або K; (bz) в положенні 129 амінокислотний залишок являє собою V, I або A; (ca) в положенні 130 амінокислотний залишок являє собою Y, H, F або C; (cb) в положенні 131 амінокислотний залишок являє собою D, G, N або E; (cc) в положенні 132 амінокислотний залишок являє собою I, T, A, M, V або L; (cd) в положенні 135 амінокислотний залишок являє собою V, T, A або I; (ce) в положенні 138 амінокислотний залишок являє собою H або Y; (cf) в положенні 139 амінокислотний залишок являє собою I або V; (eg) в положенні 140 амінокислотний залишок являє собою L або S; (ch) в положенні 142 амінокислотний залишок являє собою Y або H; (ci) в положенні 143 амінокислотний залишок являє собою K, Т або E; (cj) в положенні 144 амінокислотний залишок являє собою K, Е або R; (ck) в положенні 145 амінокислотний залишок являє собою L або I; (cl) в положенні 146 амінокислотний залишок являє собою Т або A.

Деякі переважні GAT-поліпептиди за винаходом відрізняються таким чином. При оптимальному вирівнюванні з еталонною амінокислотою послідовністю, вибраною з групи, що включає SEQ ID NO:6-10 і 263-514, щонайменше 80% амінокислотних залишків в поліпептиді, що відповідають вказаним далі положенням, узгоджуються з такими обмеженнями: (a) в положенні 9, 76, 94 і 110 амінокислотний залишок являє собою A; (b) в положеннях 29 і 108 амінокислотний залишок являє собою C; (c) в положенні 34 амінокислотний залишок являє собою D; (d) в положенні 95 амінокислотний залишок являє собою E; (e) в положенні 56 амінокислотний залишок являє собою F; (f) в положеннях 43, 44, 66, 74, 87, 102, 116, 122, 127 і 136 амінокислотний залишок являє собою G; (g) в положенні 41 амінокислотний залишок являє собою H; (h) в положенні 7 амінокислотний залишок являє собою I; (i) в положенні 85 амінокислотний залишок являє собою K; (j) в положеннях 20, 36, 42, 50, 72, 78, 98 і 121 амінокислотний залишок являє собою L; (k) в положеннях 1, 75 і 141 амінокислотний залишок являє собою M; (l) в положеннях 23, 64 і 109 амінокислотний залишок являє собою N; (m) в положеннях 22, 25, 133, 134 і 137 амінокислотний залишок являє собою P; (n) в положенні 71 амінокислотний залишок являє собою Q; (o) в положеннях 16, 21, 73, 99 і 111 амінокислотний залишок являє собою R; (p) в положеннях 55 і 88 амінокислотний залишок являє собою S; (q) в

положенні 77 амінокислотний залишок являє собою T; (r) в положенні 107 амінокислотний залишок являє собою W; i (s) в положеннях 13, 46, 70, 117 і 118 амінокислотний залишок являє собою Y.

Деякі переважні GAT-поліпептиди за винаходом відрізняються таким чином. При оптимальному вирівнюванні з еталонною амінокислотою послідовністю, вибраною з групи, що включає SEQ ID NO:6-10 і 263-514, амінокислотний залишок в поліпептиді, що відповідає положенню 28, являє собою V або A. Валін в положенні 28 загалом корелює зі зниженим параметром  $K_m$ , у той час як аланін у цьому положенні загалом корелює з підвищеним параметром  $K_{кат}$ . Інші переважні GAT-поліпептиди відрізняються тим, що вони містять I27 (тобто, I в положенні 27), M30, S35, R37, S39, G48, K49, N57, Q58, P62, Q65, Q67, K68, E83, S89, A96, E96, R101, T112, A114, K119, K120, E128, V129, D131, T131, V134, R144, 1145 або T146, або будь-яку їх комбінацію.

Деякі переважні GAT-поліпептиди за винаходом включають амінокислотну послідовність, вибрану з групи, що включає SEQ ID NO:6-10 і 263-514.

Винахід також стосується переважних GAT-поліпептидів, які відрізняються комбінацією вищезгаданих обмежень у положеннях амінокислотних залишків.

Крім того, винахід стосується GAT-поліпептидів, що кодують переважні GAT-поліпептиди, описані вище, і їх комплементарні нуклеотидні послідовності.

Деякі аспекти винаходу окремо стосуються підгрупи будь-якої з вищеприписаних категорій GAT-поліпептидів, яка має GAT-активність, описану тут. Ці GAT-поліпептиди є переважними, наприклад, для застосування у якості засобів для надання рослині стійкості до гліфосату. Приклади бажаних рівнів GAT-активності описані тут.

В одному з аспектів GAT-поліпептиди включають амінокислотну послідовність, що кодується рекомбінантною або ізольованою формою природних нуклеїнових кислот, ізольованих з природного джерела, наприклад, бактеріального штаму. Поліпептиди дикого типу, що кодують такі GAT-поліпептиди, можуть бути піддані специфічному скринінгу за допомогою стандартних методик, відомих в цій області. Наприклад, поліпептиди, визначені послідовностями SEQ ID NO:6 - SEQ ID NO:10, були відкриті шляхом експресійного клонування послідовностей з штамів *Bacillus*, які проявляють GAT-активність, як детальніше описано нижче.

Винахід також включає ізольовані або рекомбінантні поліпептиди, що кодуються ізольованим або рекомбінантним поліпептидом, який включає нуклеотидну послідовність, яка гібридується в жорстких умовах по суті по всій довжині нуклеотидної послідовності, вибраної з групи, що включає SEQ ID NOS: 1-5 і 11-262, їх компоненти і нуклеотидні послідовності, які кодують амінокислотну послідовність, вибрану з групи, що включає SEQ ID NOS: 6-10 і 263-514, включаючи їх компоненти.

Винахід також включає будь-який поліпептид, що має GAT-активність, який кодується фрагмен-

том будь-якого з GAT-кодуючих поліпептидів, описаних тут.

Винахід також стосується фрагментів GAT-поліпептидів, які можуть бути сплайсовані один з іншим з утворенням функціонального GAT-поліпептиду. Сплайсінг може бути здійснений *in vitro* або *in vivo* і може включати цис- або транс- (тобто, внутрішньомолекулярний або міжмолекулярний) сплайсінг. Ці фрагменти самі можуть, але не обов'язково, мати GAT-активність. Наприклад, два або більше сегментів GAT-поліпептиду можуть бути розділені інтеїнами; видалення інтеїнової послідовності шляхом цис-сплайсінгу приводить до утворення функціонального GAT-поліпептиду. В іншому прикладі зашифрований GAT-поліпептид може експресуватися у вигляді двох або більше окремих фрагментів; транс-сплайсінг цих сегментів приводить до відновлення функціонального GAT-поліпептиду. Різні аспекти цис- і транс-сплайсінгу, кодування генів і введення проміжних послідовностей детальніше описані в патентних заявках США №№09/517,933 і 09/710,686, обидві з яких повністю включені в цю заявку у вигляді посилання.

Загалом, винахід включає будь-який поліпептид, закодований модифікованим GAT-поліпептидом, отриманим шляхом мутації, рекурсивної рекомбінації послідовностей і/або диверсифікації поліпептидних послідовностей, описаних тут. В деяких аспектах винаходу GAT-поліпептид є модифікованим шляхом одного або багатьох амінокислотних заміщень, делецій, інсерцій або комбінації одного або більше з цих типів модифікацій. Заміщення можуть бути консервативними або неконсервативними, можуть змінювати або не змінювати функцію, а також можуть додавати нову функцію. Інсерції і делеції можуть бути значними, як у випадку усичення значного фрагмента послідовності або злиття з додатковою послідовністю, всередині або у N- чи C-закінчення. В деяких варіантах здійснення винаходу GAT-поліпептид є частиною злитого протеїну, що включає функціональний додаток, наприклад, сигнал секреції, хлоропласт-транзитний пептид, таг для очищення або будь-яку з безлічі інших функціональних груп, що відомі досвідченим фахівцям і детально описані в цьому документі.

Поліпептиди за винаходом можуть містити одну або більше модифікованих амінокислот. Присутність модифікованих амінокислот може бути сприятливою, наприклад, для (а) підвищення полуперіоду існування поліпептиду *in vivo*, (b) зниження або підвищення антигенності поліпептиду, (с) підвищення стабільності зберігання поліпептиду. Амінокислоти модифікують, наприклад, ко-трансляційно або пост-трансляційно під час рекомбінантного вироблення (наприклад, N-зчепленої глікозилації в N-X-S/T-мотивах під час експресії в клітинах ссавців) або модифікують синтетичними способами.

Необмежуючі приклади модифікованої амінокислоти включають глікозильовану амінокислоту, сульфатовану амінокислоту, пренілатовану (наприклад, фарнезильовану, геранілгеранільовану) амінокислоту, ацетильовану амінокислоту, аци-



льовану амінокислоту, PEG-ильовану амінокислоту, біотинильовану амінокислоту, карбоксильовану амінокислоту, фосфорильовану амінокислоту і т.п. Посилань для отримання довідок про модифікацію амінокислот достатньо в літературі. Типові протоколи описані у Walker (1998) Protein Protocols on CD-ROM Human Press, Towata, NJ.

Рекомбінантні способи вироблення і ізоляції поліпептидів за винаходом описані в цьому документі. На додаток до рекомбінантного вироблення поліпептиди можуть вироблятися прямим пептидним синтезом з використанням твердофазних методик (наприклад, Stewart et al. (1969) Solid-Phase Peptide Synthesis. WH Freeman Co, San Francisco; Merrifield J (1963) J. Am. Chem. Soc. 85:2149-2154). Пептидний синтез може виконуватися з використанням ручних або автоматизованих методик. Автоматизований синтез може бути здійснений, наприклад, за допомогою Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer, Foster City, Calif.) згідно з інструкціями виробника. Наприклад, послідовності можуть бути хімічно синтезовані окремо, а потім з'єднані хімічними способами з отриманням повного GAT-поліпептиду. Пептиди також можуть бути вибрані з різноманітних джерел.

В іншому аспекті винаходу GAT-поліпептид за винаходом застосовують для вироблення антитіл, наприклад, для діагностичного застосування, зокрема, для визначення активності, розподілення і експресії GAT-поліпептидів, наприклад, у різних тканинах трансгенної рослини.

GAT-гомологічні поліпептиди для індукції антитіл не потребують біологічної активності; проте поліпептид або олігопептид повинен бути антигенним. Пептиди, що використовуються для індукції специфічних антитіл, можуть мати амінокислотну послідовність, яка включає щонайменше 10 амінокислот, переважно щонайменше 15 або 20 амінокислот. Короткі відрізки GAT-поліпептиду можуть бути злиті з іншим протеїном типу keyhole limpet hemocyanin (протеїн гемоціаніну молюсків) і антитілом, що виробляється проти химерної молекули.

Способи вироблення поліклональних і моноклональних антитіл відомі фахівцям, і багато антитіл випускаються серійно. Див., наприклад, Coligan (1991) Current Protocols in Immunology Wiley/Green, NY; and Harlow and Lane (1989) Antibodies: A Laboratory Manual Cold Spring Harbor Press, NY; Stites et al. (eds.) Basic and Clinical Immunology (4<sup>th</sup> ed.) Lange Medical Publications, Los Altos, CA, і посилання, наведені там; Goding (1986) Monoclonal Antibodies: Principles and Practice (2d ed.) Academic Press, New York, NY; and Kohler and Milstein (1975) Nature 256: 495-497. Інші придатні методики для отримання антитіл включають селекцію бібліотек рекомбінантних антитіл в фагових або подібних векторах. Див. Huse et al. (1989) Science 246: 1275-1281; and Ward, et al. (1989) Nature 341: 544-546. Специфічні моноклональні і поліклональні антитіла і антисироватка у більшості випадків будуть зв'язуватися з  $K_D$  щонайменше  $\sim 0,1 \mu\text{M}$ , переважно щонайменше  $\sim 0,01 \mu\text{M}$  або краще, і найбільш

типово і переважно - щонайменше  $\sim 0,001 \mu\text{M}$  або краще.

Додаткові деталі методик вироблення і конструювання антитіл можна знайти у Borrebaeck (ed) (1995) Antibody Engineering. 2<sup>nd</sup> Edition Freeman and Company, NY (Borrebaeck); McCafferty et al. (1996) Antibody Engineering. A Practical Approach IRL at Oxford Press, Oxford, England (McCafferty), and Paul (1995) Antibody Engineering Protocols Humana Press, Towata, NJ (Paul).

#### Варіації послідовностей

GAT-поліпептиди за винаходом включають консервативно модифіковані варіації послідовностей, описаних тут як SEQ ID NOS: 6-10 і 263-514. Такі консервативно модифіковані варіації включають заміщення, приєднання або делеції, які змінюють, додають або делетують одну амінокислоту або невеликий процент амінокислот (типово менш ніж  $\sim 5\%$ , більш типово менш ніж  $\sim 4\%$ ,  $2\%$  або  $1\%$ ) в будь-якій з SEQ ID NOS: 6-10 і 263-514.

Наприклад, консервативно модифікована варіація (наприклад, делеція) 146-амінокислотного поліпептиду, ідентифікованого тут як SEQ ID NO:6, матиме довжину щонайменше 140 амінокислот, переважно щонайменше 141 амінокислоту, більш переважно щонайменше 144 амінокислоти і ще більш переважно - 146 амінокислот, що відповідає делеції менш ніж  $\sim 5\%$ ,  $4\%$ ,  $2\%$  або  $\sim 1\%$ , або менше, нуклеотидної послідовності.

Інший приклад консервативно модифікованої варіації (наприклад, "консервативно заміщеної варіації") поліпептиду, ідентифікованого тут як SEQ ID NO:6, матиме "консервативні заміщення", відповідні до шести груп заміщення, представлених в Табл.2 (нижче), в загальній кількості до  $\sim 7$  залишків (тобто, менше  $5\%$ ) 146-амінокислотного поліпептиду.

Гомологи GAT-поліпептидної послідовності за винаходом, включаючи консервативно заміщені послідовності, можуть бути присутні у вигляді частини більших поліпептидних послідовностей, як це зустрічається в GAT-поліпептиді, в GAT-злитті з сигнальною послідовністю, наприклад, таргетинг-послідовністю хлоропластів, або після додання одного або більше доменів для очищення протеїну (наприклад, полі-гіс-сегментів, FLAG-таг-сегментів і т.д.). В останньому випадку додаткові функціональні домени майже або зовсім не впливають на активність GAT-частини протеїну, так же як у випадку, коли додаткові домени можуть бути видалені за допомогою етапів пост-синтезу, таких як обробка протеазою.

Визначення поліпептидів за імунореактивністю

Оскільки поліпептиди за винаходом утворюють новий клас ферментів з визначеною активністю, тобто, активністю ацетилювання гліфосату, ці поліпептиди також забезпечують нові структурні особливості, які можуть розпізнаватися, наприклад, методами імунологічного аналізу. Вироблення антисироватки, яка специфічно зв'язується з поліпептидами за винаходом, а та-

кож поліпептидів, що зв'язуються цією антисироваткою, є рисою цього винаходу.

Винахід включає GAT-поліпептиди, які специфічно зв'язуються або є специфічно імунореактивними з антитілом або антисироваткою, виробленою проти імуногена, що включає амінокислотну послідовність, вибрану з однієї або більше з SEQ ID NO:6 - SEQ ID NO:10. Для усунення крос-реактивності з іншими GAT-гомологами антитіло або антисироватку віднімають доступними спорідненими протеїнами, такими як протеїни або пептиди, що відповідають інвентарним номерам GenBank, існуючим на момент реєстрації цієї заявки, наприклад, CAA70664, Z99109 і Y09476. У випадку, коли інвентарний номер відповідає нуклеїновій кислоті, виробляється поліпептид, що кодується цією нуклеїновою кислотою, який використовується для віднімання антитіла/антисироватки. Фіг.3 показує відносну ідентичність між типовими GAT-поліпептидами і найбільш спорідненою послідовністю з GenBank, Yitl. Функція нативної Yitl ще має бути визначена, але з'ясовано, що цей фермент має помітну GAT-активність.

В одному з типових форматів імуноаналізу використовує поліклональну антисироватку, що була отримана проти одного або більше поліпептидів, які включають одну або більше послідовностей, відповідаючих одній або більше з SEQ ID NOS: 6-10 і 263-514, або її значній частині (тобто, щонайменше ~30% повної довжини послідовності). Повний набір потенціальних поліпептидних імуногенів, виведених з SEQ ID NOS: 6-10 і 263-514, нижче називаються "імуногенними поліпептидами". Результируюча антисироватка може бути вибрана таким чином, щоб вона мала низьку крос-реактивність проти інших споріднених послідовностей, і будь-яка така крос-реактивність усувається шляхом імуноабсорбції однією або більше спорідненими послідовностями перед використанням поліклональної антисироватки в імуноаналізі.

Для отримання антисироватки для використання її в імуноаналізі один або більше імуногенних поліпептидів отримують і очищують, як описано тут. Наприклад, може бути отриманий рекомбінантний протеїн в бактеріальній клітинній лінії. Штам мишей, виведений від близьких родичів (використаний в цьому аналізі тому, що результати є більш відтворюваними завдяки фактичній генетичній ідентичності мишей), імунізували імуногенним протеїном (протеїнами) в комбінації зі стандартним ад'ювантом, таким як ад'ювант Фреунда, відповідно до стандартного протоколу імунізації мишей (див., наприклад, в Harlow and Lane (1989) *Antibodies: A Laboratory Manual*. Cold Spring Harbor Publications, New York, стандартний опис вироблення антитіл, формати імуноаналізу і умови, які можуть використовуватися для визначення специфічної імунореактивності). В альтернативі один або більше синтетичних або рекомбінантних поліпептидів, отриманих з послідовностей, описаних тут, кон'югують з протеїном-носієм і використовують у якості імуногену.

Поліклональні сироватки збирають і титрують проти імуногенного поліпептиду в імуноаналізі, наприклад, твердофазному імуноаналізі з одним або більше імуногенних поліпептидів, імобілізованих на твердій підкладці. Поліклональні антисироватки з титром  $10^6$  або більше збирають, об'єднують і віднімають спорідненими протеїнами, наприклад, ідентифікованими в GenBank, як описано вище, для отримання віднятих об'єднаних титрованих поліклональних антисироваток.

Відняті об'єднані титровані поліклональні антисироватки тестують на крос-реактивність зі спорідненими поліпептидами. Переважно щонайменше два з імуногенних GATs використовують в цьому досліді, переважно в поєднанні з щонайменше двома спорідненими поліпептидами, для ідентифікації антитіл, які специфічно зв'язуються імуногенним протеїном (протеїнами).

В цьому порівняльному аналізі визначалися умови вибіркового зв'язування для віднятих титрованих поліклональних антисироваток, що забезпечують щонайменше в ~5-10 вище відношення сигнал/шум для зв'язування титрованих поліклональних антисироваток з імуногенними поліпептидами, ніж для зв'язування зі спорідненими поліпептидами. Тобто, точність реакції зв'язування регулюється додаванням неспецифічних конкурентів, таких як альбумін або нежирне сухе молоко, або шляхом регулювання сольових умов, температури і т.п. Ці умови зв'язування використовуються в наступних аналізах для визначення, чи буде тестовий поліпептид специфічно зв'язуватися об'єднаними віднятими поліклональними антисироватками. Зокрема, тестові поліпептиди, що показують щонайменше в ~2-5 вище відношення сигнал/шум, ніж контрольні поліпептиди в умовах вибіркового зв'язування, і щонайменше половинне відношення сигнал/шум порівняно з імуногенним поліпептидом (поліпептидами), мають значну структурну подібність з імуногенним поліпептидом порівняно з відомим GAT, і тому є поліпептидами за винаходом.

В іншому прикладі імуноаналізи в форматі конкурентного зв'язування використовуються для виявлення тестового поліпептиду. Наприклад, як було зазначено вище, крос-реактивні антитіла видаляють з об'єднаної суміші антисироваток шляхом імуноабсорбції контрольними GAT-поліпептидами. Потім імуногенний поліпептид (поліпептиди) імобілізують на твердій основі, яку піддають дії віднятої об'єднаної антисироватки. Тестові протеїни додають в аналітичну суміш для конкурентного зв'язування з об'єднаною віднятою антисироваткою. Здатність тестового протеїну (протеїнів) зв'язуватися з об'єднаною віднятою антисироваткою порівняно з імобілізованим протеїном (протеїнами) порівнюють зі здатністю імуногенного поліпептиду (поліпептидів), що додаються в аналітичну суміш, конкурентно зв'язуватися (імуногенні поліпептиди ефективно конкурують з імуногенними імуногенними поліпептидами за зв'язування з об'єднаною антисироваткою). Процентну крос-реактивність для тестових протеїнів розраховують з використанням стандартних розрахунків.

В паралельному аналізі здатність контрольних протеїнів конкурувати за зв'язування з об'єднаною віднятою антисироваткою можна визначати порівняно зі здатністю імуногенного поліпептиду (поліпептидів) конкурувати за зв'язування з антисироваткою. Процентну крос-реактивність для контрольних поліпептидів розраховують з використанням стандартних розрахунків. Якщо процентна крос-реактивність для тестових поліпептидів щонайменше в 5-10 разів вище, кажуть, що тестові поліпептиди специфічно зв'язуються з об'єднаною віднятою антисироваткою.

Загалом, імуноабсорбована і об'єднана антисироватка може бути використана в імуноаналізі на конкурентне зв'язування, як описано тут, для порівняння будь-якого тестового поліпептиду з імуногенним поліпептидом (поліпептидами). Щоб здійснити це порівняння, два поліпептиди аналізують окремо в широкому діапазоні концентрацій, і кількість кожного поліпептиду, необхідну для інгібування 50% зв'язування віднятої антисироватки з імобілізованим протеїном, визначають за допомогою стандартних методик. Якщо необхідна кількість тестового поліпептиду менше подвійної необхідної кількості імуногенного поліпептиду, тоді кажуть, що тестовий поліпептид специфічно зв'язується з антитілом, виробленим до імуногенного протеїну, за умови, що ця кількість щонайменше в 5-10 разів вище, ніж для контрольного поліпептиду.

У якості остаточного визначення специфічності (необов'язкового) відняту антисироватку повністю імуносорбують імуногенним поліпептидом (поліпептидами) (але не контрольними поліпептидами), доки зв'язування результуючої імуногенної поліпептидної віднятої об'єднаної антисироватки з імуногенним поліпептидом (поліпептидами), використовуваним в імуносорбції, майже або зовсім не буде виявлятися. Тоді цю повністю імуносорбовану антисироватку тестують на реактивність з тестовим поліпептидом. Якщо спостерігається мінімум або відсутність реактивності (тобто, не більше подвійного відношення сигнал/шум порівняно з отриманим при зв'язуванні повністю імуносорбованої антисироватки з імуногенним поліпептидом), це означає, що тестовий поліпептид специфічно зв'язується антисироваткою, витягнутою імуногенним поліпептидом.

Поліпептиди гліфосат-N-ацетилтрансферази

В одному з аспектів винахід стосується нового сімейства ізолюваних або рекомбінантних поліпептидів, що називаються тут "поліпептидами гліфосат-N-ацетилтрансферази" або "GAT-поліпептидами". GAT-поліпептидні послідовності відрізняються здатністю кодувати GAT-поліпептид. Загалом, винахід включає будь-яку нуклеотидну послідовність, яка кодує будь-який з нових GAT-поліпептидів, описаних тут. В деяких аспектах винаходу переважним є GAT-поліпептид, який кодує GAT-поліпептид з GAT-активністю.

В одному з аспектів GAT-поліпептиди включають рекомбінантні або ізолювані форми

природних нуклеїнових кислот, ізолюваних з організму, наприклад, бактеріального штаму. Типові GAT-поліпептиди, наприклад, SEQ ID NO:1 - SEQ ID NO:5, були відкриті шляхом експресійного клонування послідовностей з штамів *Bacillus*, що показують GAT-активність. У короткому викладі, колекцію близько 500 штамів *Bacillus* і *Pseudomonas* піддали скринінгу на нативну здатність N-ацетилувати гліфосат. Штами вирощували в LB протягом ночі, зібрали шляхом центрифугування, пермеабілізували в розведеному толуолі, а потім промили і пересушували в реакційній суміші, що містила буфер, 5мМ гліфосату і 200мМ ацетил-CoA. Клітини інкубували в цій реакційній суміші від 1 до 48 годин, додавши за цей час в реакцію рівний об'єм метанолу. Потім клітини пелетизували шляхом центрифугування і відфільтрували супернатант для проведення аналізу методом мас-спектрометрії батьківських іонів. Продукт реакції був позитивно ідентифікований як N-ацетилгліфосат шляхом порівняння профілю мас-спектрометрії реакційної суміші з еталоном N-ацетилгліфосату, як показано на Фіг.2. На виявлення продукту могло впливати включення обох субстратів (ацетилкоферменту A та гліфосату), що було усунено шляхом теплової денатурації бактеріальних клітин.

Потім окремі GAT-поліпептиди клонували з ідентифікованих штамів шляхом функціонального скринінгу. Отримали геномну ДНК і частково розщепили ферментом *Sau3A1*. Фрагменти ~4Kb клонували в вектор експресії *E.coli* і трансформували в електрокомпетентні *E.coli*. Окремі клони, що показували GAT-активність, ідентифікували мас-спектрометрією після реакції, описаної вище, за тим винятком, що толуол замінили пермеабілізацією в PMBS. Геномні фрагменти секвенували і ідентифікували очікувану GAT-поліпептид-кодуючу відкриту рамку зчитування. Ідентичність GAT-гена підтвердили експресією відкритої рамки зчитування в *E. coli* і виявленням високих рівнів N-ацетилгліфосату, вироблених в реакційних сумішах.

В іншому аспекті винаходу GAT-поліпептиди виробляються шляхом диверсифікації, наприклад, рекомбінації і/або мутації одного або більше природних, ізолюваних або рекомбінантних GAT-поліпептидів. Як детально описано в цьому документі, часто буває можливо генерувати диверсифіковані GAT-поліпептиди, що кодують GAT-поліпептиди з високими функціональними показниками, наприклад, з підвищеною каталітичною функцією, підвищеною стабільністю, більш високим рівнем експресії, ніж GAT-поліпептиди, що використовуються у якості субстрату або предка в процесі диверсифікації.

Поліпептиди за винаходом мають численні призначення, наприклад, для рекомбінантного виробництва (тобто, експресії) GAT-поліпептидів за винаходом; у якості трансгенів (наприклад, для надання стійкості до гербіцидів трансгенним рослинам); у якості маркерів, що селектуються, для трансформації і підтримання плазмід; у якості імуногенів; у якості діагностич-

них зондів на присутність комплементарних або частково комплементарних нуклеїнових кислот (у тому числі для виявлення природних GAT-кодуючих нуклеїнових кислот); у якості субстратів для подальшого вироблення різноманітності, наприклад, рекомбінаційних реакцій або мутаційних реакцій для отримання нових і/або покращених GAT-гомолів і т.п.

Слід зазначити, що деякі специфічні, суттєві та надійні корисності GAT-полінуклеотидів не потребують, щоб полінуклеотид кодував поліпептид зі значною GAT-активністю. Наприклад, GAT-полінуклеотиди, які не кодують активних ферментів, можуть бути цінними джерелами батьківських полінуклеотидів для використання в диверсифікаційних процедурах для отримання варіантів GAT-полінуклеотидів, або не-GAT-полінуклеотидів, з бажаними функціональними властивостями (наприклад, високими  $K_{кат}$  або  $K_{кат}/K_m$ , низькими  $K_m$ , високою стабільністю при дії тепла або інших навколишніх факторів, високими швидкостями транскрипції або трансляції, стійкістю до протеолітичного розщеплення, зниженою антигенністю і т.д.). Наприклад, нуклеотидні послідовності, що кодують варіанти протеаз з мінімумом або відсутністю помітної активності, були застосовані у якості батьківських полінуклеотидів в дослідях з ДНК-шафлінгом для вироблення потомства, кодуючого високоактивні протеази (Ness et al. (1999) Nature Biotechnology 17:893-96).

Полінуклеотидні послідовності, вироблені способами генерації різноманітності або способами рекурсивної рекомбінації послідовностей ("RSR") (наприклад, ДНК-шафлінгом), є рисою цього винаходу. Способи мутації і рекомбінації з використанням нуклеїнових кислот, описаних тут, є рисою цього винаходу. Наприклад, один із способів за винаходом включає рекурсивну рекомбінацію однієї або більше нуклеотидних послідовностей за винаходом, описаних вище або нижче, з одним або більше додатковими нуклеотидами. Етапи рекомбінації можуть виконуватися *in vivo*, *ex vivo*, *in silico* або *in vitro*. Вищезгадана генерація різноманітності або рекурсивна рекомбінація послідовностей виробляє щонайменше одну бібліотеку рекомбінантних модифікованих GAT-полінуклеотидів. Поліпептиди, що кодуються членами цієї бібліотеки, включені у винахід.

Винахід також включає застосування полінуклеотидів, також званих тут олігонуклеотидами, які звичайно мають щонайменше 12 основ, переважно щонайменше 15, більш переважно щонайменше 20, 30 або 50 або більше основ, які гібридизуються в жорстких або високо жорстких умовах з послідовністю GAT-полінуклеотиду. Ці полінуклеотиди можуть бути використані у якості зондів, праймерів, смислових і антисмислових агентів і т.п., відповідно до способів, згаданих тут.

За винаходом, GAT-полінуклеотиди, включаючи нуклеотидні послідовності, що кодують GAT-поліпептиди, фрагменти GAT-поліпептидів, споріднені злиті протеїни або їх функціональні еквіваленти, використовуються в рекомбінантних молекулах ДНК, які направляють експресію GAT-поліпептидів у відповідних клітинах-хазяях, таких

як бактеріальні або рослинні клітини. Внаслідок властивості їм дегенерації генетичного коду, інші нуклеїновокислотні послідовності, що кодують по суті ту саму або функціонально еквівалентну амінокислотну послідовність, також можуть бути використані для клонування і експресії GAT-полінуклеотидів.

Винахід стосується GAT-полінуклеотидів, кодуючих продукти транскрипції і/або трансляції, з яких шляхом сплайсінгу зрештою отримують функціональні GAT-поліпептиди. Сплайсінг може виконуватися *in vitro* або *in vivo* і може включати цис- або транс-сплайсінг. Субстратом для сплайсінгу можуть бути полінуклеотиди (наприклад, РНК-транскрипти) або поліпептиди. Прикладом цис-сплайсінгу полінуклеотиду є випадок, коли інтрон, вставлений в кодуючу послідовність, видаляють, а дві фланкуючі екзонні області піддають сплайсінгу з отриманням GAT-поліпептид-кодуючої послідовності. Прикладом транс-сплайсінгу полінуклеотиду є випадок, коли GAT-полінуклеотид кодують шляхом розділення кодуючої послідовності на два або більше фрагментів, які можуть бути окремо транскрибовані і потім піддані сплайсінгу з утворенням повної GAT-кодуючої послідовності. Використання сплайсінг-енхансерної послідовності (яка може бути введена в конструкцію за винаходом) може полегшити цис- або транс-сплайсінг. Цис- і транс-сплайсінг поліпептидів детально описаний в цьому документі. Більш детальний опис цис- і транс-сплайсінгу можна знайти в патентних заявках США №№09/517,933 і 09/710,686.

Таким чином, деякі з GAT-полінуклеотидів безпосередньо не кодують повний GAT-поліпептид, але кодують фрагмент або фрагменти GAT-поліпептиду. Ці GAT-полінуклеотиди можуть бути використані для експресії функціонального GAT-поліпептиду за допомогою механізму, що включає сплайсінг, при цьому сплайсінг може відбуватися на рівні полінуклеотиду (наприклад, інтрона/екзона) і/або поліпептиду (інтеїну/екстеїну). Це може бути корисно, наприклад, для контролю експресії GAT-активності, оскільки функціональний GAT-поліпептид буде експресуватися, тільки якщо всі потрібні фрагменти будуть експресуватися в навколишньому середовищі, дозволяючи процеси сплайсінгу для вироблення функціонального продукту. В іншому прикладі введення однієї або більше інсерційних послідовностей в GAT-полінуклеотид може полегшити рекомбінацію з низькогомологічним полінуклеотидом; застосування інтрона або інтеїна для інсерційної послідовності полегшує видалення проміжної послідовності для відновлення функції закодованого варіанту.

Досвідченому фахівцеві буде зрозуміло, що може бути сприятливою модифікація кодуючої послідовності для посилення її експресії в конкретному хазяїні. Генетичний код, який має 64 можливі кодони, є надлишковим, і більшість організмів переважно використовують певну підмножину цих кодонів. Кодони, що використовуються кожним видом найчастіше, називаються оптимальними кодонами, а ті, що не дуже часто використовуються, класифікуються як рідкі кодони або

кодони низького застосування (див., наприклад, Zhang SP et al. (1991) *Gene* 105:61-72). Кодони можуть бути замінені для відображення переважного застосування кодонів хазяїном; цей процес іноді називають "оптимізацією кодонів" або "контролем зміщення кодонів даного виду".

Оптимізована кодуюча послідовність, що містить кодони, переважні для конкретного прокариотичного або еукаріотичного хазяїна (див. також Murray, E. et al. (1989) *Nuc. Acids Res.* 17:477-508), може бути отримана, наприклад, для збільшення швидкості трансляції або для вироблення рекомбінантних транскриптів РНК, що мають бажані властивості, такі як довший напівперіод існування порівняно з транскриптами, отриманими з неоптимізованої послідовності. Трансляційні стоп-кодони також можуть бути модифіковані для відображення переважних кодонів хазяїна. Наприклад, переважними стоп-кодонами для *S. cerevisiae* і ссавців є UAA і UGA, відповідно. Переважним стоп-кодоном для односім'ядольних рослин є UGA, у той час як комахи і *E. coli* віддають перевагу застосуванню UAA у якості стоп-кодона (Dolphin ME et al. (1996) *Nuc. Acids Res.* 24: 216-218). Методика оптимізації нуклеотидної послідовності для експресії в рослині описана, наприклад, в патенті США №6,015,891 і посиланнях, наведених там.

Один з варіантів здійснення винаходу включає GAT-полінуклеотид, що містить оптимальні кодони для експресії у відповідному хазяїні, наприклад, трансгенному рослинному хазяїні. Це особливо бажано, коли GAT-полінуклеотид бактеріального походження вводять в трансгенну рослину, наприклад, для надання рослині стійкості до гліфосату.

Полінуклеотидні послідовності за винаходом можуть конструюватися таким чином, щоб змінити GAT-полінуклеотид для різних цілей, включаючи, але не обмежуючись, зміни, які модифікують клонування, обробку і/або експресію генного продукту. Наприклад, зміни можуть бути введені з використанням технологій, добре відомих в біотехнології, таких як сайт-направлений мутагенез, для вставлення нових рестрикційних сайтів, зміни шаблонів глікозиляції, зміни переважних кодонів, введення сайтів сплайсінгу і т.д.

Як детально описано в цьому документі, полінуклеотиди за винаходом включають послідовності, які кодують нові GAT-поліпептиди і послідовності, комплементарні до кодуючих послідовностей, і нові фрагменти кодуючої послідовності і їх комплементи. Полінуклеотиди можуть мати форму РНК або ДНК і включають мРНК, кРНК, синтетичні РНК і ДНК, геномні ДНК і кДНК. Полінуклеотиди можуть бути двоспіральними або односпіральними, і у випадку односпіральних молекул можуть являти собою кодуючу спіраль або некодуючу (антисмыслову, комплементарну) спіраль. Полінуклеотиди можуть включати кодуючу послідовність GAT-поліпептиду (i) в ізоляції, (ii) в комбінації з додатковою кодуючою послідовністю, так щоб кодувати, наприклад, злитий протеїн, препро-протеїн, препро-протеїн і т.п., (iii) в комбінації з некодуючими послідовностями, такими як інтрони або інтеїни, контрольними

елементами, такими як промотор, енхансер, термінуючий елемент, або 5'- і/або 3'-нетрансльованими областями, ефективними для експресії кодуючої послідовності в придатному хазяїні, і/або (iv) у середовищі вектора або хазяїна, в якому GAT-полінуклеотид є гетерологічним геном. Послідовності можуть також знаходитися у комбінації з типовими композиціями нуклеїнових кислот, що включають, наприклад, носії, буфери, ад'юванти, наповнювачі і т.п.

Полінуклеотиди і олігонуклеотиди за винаходом можуть бути отримані за стандартними твердофазними методиками, відповідно до відомих способів синтезу. Звичайно синтезують окремі фрагменти розміром до -100 основ, потім їх з'єднують (наприклад, способами ферментативного або хімічного лігування, або полімераза-опосередненими способами) з утворенням по суті будь-якої бажаної безперервної послідовності. Наприклад, полінуклеотиди і олігонуклеотиди за винаходом можуть бути отримані хімічним синтезом з використанням класичного фосфорамідитного способу, описаного Beaucage et al. (1981) *Tetrahedron Letters* 22:1859-69, або способу, описаного Matthes et al. (1984) *EMBO J.* 3: 801-05, наприклад, як це звичайно використовується в автоматизованих способах синтезу. Відповідно до фосфорамідитного способу, олігонуклеотиди синтезують, наприклад, в автоматичному ДНК-синтезаторі, очищують, відпалюють, лігують і клонують у відповідні вектори.

Крім того, по суті будь-яку нуклеїнову кислоту можна замовити в одному з багатьох комерційних джерел, таких як The Midland Certified Reagent Company (mcrc@oligos.com), The Great American Gene Company (<http://www.genco.com>), ExpressGen Inc. ([www.expressgen.com](http://www.expressgen.com)), Operon Technologies Inc. (Alameda, CA) та інших. Подібним чином, пептиди і антитіла можуть бути замовлені в будь-якому з багатьох джерел, таких як PeptideGenic (pkim@ccnet.com), HTI Bio-products, Inc. (<http://www.htibio.com>), BMA Biomedicals Ltd (U.K.), Bio.Synthesis, Inc., та інших.

Полінуклеотиди також можуть бути синтезовані за відомими технологіями, описаними в технічній літературі. Див. наприклад, Carruthers et al. *Cold Spring Harbor Symp. Quant. Biol.* 47:411-418 (1982), and Adams et al., *J. Am. Chem. Soc.* 105:661 (1983). Тоді двоспіральні фрагменти ДНК можуть бути отримані синтезом комплементарної спіралі і спільного відпалу спіралей у відповідних умовах, або шляхом приєднання комплементарної спіралі з використанням ДНК-полімерази з відповідною праймерною послідовністю.

Основні документи, що описують молекулярно-біологічні технології, корисні для цього винаходу, в тому числі мутагенез, включають Berger and Kimmel, *Guide to Molecular Cloning Techniques. Methods in Enzymology*, volume 152 Academic Press, Inc., San Diego, CA ("Berger"); Sambrook et al., *Molecular Cloning - A Laboratory Manual* (2nd Ed.), volumes 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989 ("Sambrook"); and *Current Protocols in Molecular Biology*, F.M. Ausubel et al., eds. Current Protocols, a joint venture between Greene Publishing

Associates, Inc. and John Wiley & Sons, Inc., (supplemented through 2000) ("Ausubel"). Приклади технологій, достатні для створення уяви про способи ампліфікації *in vitro*, що включають полімеразну ланцюгову реакцію (ПЛР), лігазну ланцюгову реакцію (ЛЛР), Q $\beta$ -репліказну ампліфікацію та інші технології, опосереднені РНК-полімеразою (наприклад, NASBA), можуть бути знайдені в Berger, Sambrook, and Ausubel, а також Mullis et al., (1987) U.S. Patent No. 4,683,202; PCR Protocols A Guide to Methods and Applications (Innis et al., eds.) Academic Press Inc. San Diego, CA (1990); Amheim & Levinson (October 1, 1990) Chemical and Engineering News 36-47; The Journal Of NIH Research (1991) 3:81-94; Kwoh et al. (1989) Proc. Natl. Acad. Sci. USA 86:1173; Guatelli et al. (1990) Proc. Natl. Acad. Sci. USA 87:1874; Lomell et al. (1989) J.Clin. Chem. 35:1826; Landegren et al., (1988) Science 241:1077-1080; Van Brunt (1990) Biotechnology 8:291-294; Wu and Wallace, (1989) Gene 4:560; Bardnager et al. (1990) Gene 89:117, and Sooknanan and Malek (1995)

Biotechnology 13:563-564. Вдосконалені способи клонування *in vitro* ампліфікованих нуклеїнових кислот описані в Wallace et al., патент США №5,426,039. Вдосконалені способи ампліфікації великих нуклеїнових кислот шляхом ПЛР описані в Cheng et al. (1994) Nature 369:684-685 і наведених там посиланнях, де отримують ПЛР-реплікони розміром до 40kb. Фахівцеві буде зрозуміло, що по суті будь-яка РНК може бути перетворена в двоспіральну ДНК, придатну для рестрикційної ферментації, ПЛР-розширення і секвенування з використанням зворотної транскриптази і полімерази. Див. Ausubel, Sambrook and Berger, усі вище.

Варіації послідовностей

Фахівцям буде зрозуміло, що завдяки дегенерації генетичного коду може вироблятися безліч нуклеотидних послідовностей, що кодують GAT-поліпептиди за винаходом, деякі з яких мають суттєву ідентичність з нуклеїновокислотними послідовностями, детально описаними тут.

Таблиця 1

Таблиця кодонів

Амінокислоти			Кодон			
Аланін	Ala	A	GCA	GCC	GCG	GCU
Цистеїн	Cys	C	UGC	UGU		
Аспарагінова кислота	Asp	D	GAC	GAU		
Глутамінова кислота	Glu	E	GAA	GAG		
Фенілаланін	Phe	F	UUC	UUU		
Гліцин	Gly	G	GGA	GGC	GGG	GGU
Гістидин	His	H	CAC	CAU		
Ізолейцин	Ile	I	AUA	AUC	AUU	
Лізин	Lys	K	AAA	AAG		
Лейцин	Leu	L	UUA	UUG	CUA	CUC CUG CUU
Метіонін	Met	M	AUG			
Аспарагін	Asn	N	AAC	AAU		
Пролін	Pro	P	CCA	CCC	CCG	CCU
Глутамін	Gln	Q	CAA	CAG		
Аргінін	Arg	R	AGA	AGG	CGA	CGC CGG CGU
Серин	Ser	S	AGC	AGU	UCA	UCC UCG UCU
Треонін	Thr	T	ACA	ACC	ACG	ACU
Валін	Val	V	GUA	GUC	GUG	GUU
Триптофан	Trp	W	UGG			
Тирозин	Tyr	Y	UAC	UAU		

Наприклад, таблиця кодонів (Табл.1) показує, що всі кодони AGA, AGG, CGA, CGC, CGG і CGU кодують амінокислоту аргінін. Таким чином, у кожному положенні в нуклеїнових кислотах за винаходом, де кодоном визначається аргінін, цей кодон може бути змінений на будь-який з відповідних кодонів, описаних вище, без зміни закодованого поліпептиду. Зрозуміло, що U в РНК-послідовності відповідає T в ДНК-послідовності.

При використанні, у якості прикладу, нуклеїновокислотної послідовності, що відповідає нуклеотидам 1-15 з SEQ ID NO:1, ATG ATT GAA GTC AAA, німа варіація цієї послідовності включає AGT ATC GAG GTG AAG, при цьому обидві послідовності кодують амінокислотну послідов-

ність MD5VK, що відповідає амінокислотам 1-5 SEQ ID NO:6.

Такі "німі варіації" належать до одного виду "консервативно модифікованих варіацій", описаних нижче. Фахівцеві буде зрозуміло, що кожний кодон в нуклеїновій кислоті (крім AUG, який звичайно є єдиним кодоном для метіоніну) може бути модифікований за стандартними методиками, так щоб кодувати функціонально ідентичний поліпептид. Відповідним чином, кожна німа варіація нуклеїнової кислоти, яка кодує поліпептид, є можливою у будь-якій описаній послідовності. Винахід стосується кожної можливої варіації нуклеїновокислотної послідовності, що кодує поліпептид за винаходом, який може бути отриманий шляхом вибору комбінацій на основі можливих

варіантів кодонів. Ці комбінації здійснюються у відповідності до стандартного триплетного генетичного коду (наприклад, як вказано в Табл.1), застосованого до нуклеїновокислотної послідовності, яка кодує GAT-гомологічний поліпептид за винаходом. Всі такі варіації кожної нуклеїнової кислоти специфічно передбачені і описані тут шляхом розгляду послідовності в комбінації з генетичним кодом. Будь-який варіант може бути вироблений як зазначено тут.

Група з двох або більше різних кодонів, які при трансляції в одному й тому самому контексті кодують одну й ту саму амінокислоту, тут називаються "синонімічними кодонами." Як описано тут, в деяких аспектах винаходу GAT-полінуклеотид є сконструйованим для оптимізованого застосування кодонів в бажаному організмі-хазяїні, наприклад, рослинному хазяїні. Вираз "оптимізований" або "оптимальний" не обмежується найкращою можливою комбінацією кодонів, а просто показує, що кодуюча послідовність в цілому має покращене застосування кодонів порівняно з поліпептидом-попередником, з якого вона була виведена. Таким чином, в одному з аспектів винахід стосується способу вироблення варіанту GAT-полінуклеотиду шляхом заміни щонайменше одного батьківського кодона в нуклео-

тидній послідовності синонімічним кодоном, який переважно використовується в бажаному організмі-хазяїні, наприклад, рослині, порівняно з батьківським кодоном.

Вираз "консервативно модифіковані варіації" або просто "консервативні варіації" конкретної нуклеїновокислотної послідовності стосується тих нуклеїнових кислот, які кодують ідентичні або по суті ідентичні амінокислотні послідовності, або, якщо нуклеїнова кислота не кодує амінокислотної послідовності, стосується по суті ідентичних послідовностей. Фахівцеві буде зрозуміло, що індивідуальні заміщення, делеції або приєднання, які змінюють, додають або делетують одну амінокислоту або малий процент амінокислот (звичайно менше 5%, більш звичайно менше 4%, 2% або 1%, або менше) в послідовність, що кодується, є "консервативно модифікованими варіаціями", при цьому зміни складаються в делеції амінокислоти, приєднанні амінокислоти або заміщенні амінокислоти хімічно подібною амінокислотою.

Таблиці консервативних заміщень, які включають функціонально подібні амінокислоти, добре відомі у біотехнології. Таблиця 2 показує шість груп, що включають амінокислоти, які є "консервативними заміщеннями" одна до одної.

Таблиця 2

## Групи консервативних заміщень

1	Аланін (A)	Серин (S)	Треонін (T)
2	Аспарагінова кислота (D)	Глутамінова кислота (E)	
3	Аспарагін (N)	Глутамін (Q)	
4	Аргінін (R)	Лізин (K)	
5	Ізолейцин (I)	Лейцин (L)	Метіонін (M) Валін (V)
6	Фенілаланін (F)	Тирозин (Y)	Триптофан (W)

Таким чином, "консервативно заміщені варіації" перелічених поліпептидних послідовностей за винаходом включають заміщення малого процента, звичайно менше 5%, більш звичайно менше 2% і часто менше 1% амінокислот поліпептидної послідовності консервативно вибраними амінокислотами тієї самої групи консервативного заміщення.

Наприклад, консервативно заміщена варіація поліпептиду, ідентифікованого тут як SEQ ID NO:6, буде містити "консервативні заміщення", відповідні до шести груп, визначених вище, в загальній кількості до 7 залишків (тобто, менше 5% амінокислот) 146-амінокислотного поліпептиду.

В іншому прикладі, якщо чотири консервативні заміщення були локалізовані в області, що відповідає амінокислотам 21-30 SEQ ID NO:6, приклади консервативно заміщених варіацій цієї області, RPN QPL EAC M, включають:

KPQ QPV ESC M і

KPN NPL DAC V і т.п., відповідно до консервативних заміщень, перелічених в Табл.2 (у наведеному вище прикладі консервативні заміщення підкреслені). Перелік протеїнових послідовностей, у поєднанні з наведеною вище таблицею заміщень, надає тут експрес-перелік всіх консервативно заміщених протеїнів.

Нарешті, приєднання послідовностей, які не змінюють закодованої активності молекули нуклеїнової кислоти, наприклад, приєднання нефункціональної або некодуєчої послідовності, є консервативною варіацією базової нуклеїнової кислоти.

Фахівцеві буде зрозуміло, що численні консервативні варіації нуклеїновокислотних конструкцій, описаних тут, в результаті утворюють функціонально ідентичну конструкцію. Наприклад, як описано вище, завдяки дегенерації генетичного коду, "німі заміщення" (тобто, заміщення в нуклеїновокислотній послідовності, які не приводять до змін в закодованому поліпептиді) є властивою рисою кожної нуклеїновокислотної послідовності, яка кодує амінокислоту. Подібним чином, "консервативні амінокислотні заміщення" в одній або декількох амінокислотах в амінокислотній послідовності, що заміщені іншими амінокислотами з високо подібними властивостями, також легко ідентифікуються як високо подібні до конструкції, що описується. Такі консервативні варіації кожної описаної тут послідовності є рисою цього винаходу.

Неконсервативними модифікаціями конкретної нуклеїнової кислоти є такі заміщення будь-якої амінокислоти, які не характеризуються як

консервативні заміщення. Наприклад, це будь-яке заміщення, яке перетинає межі шести груп, зазначених в Табл.2. Вони включають заміщення основних або кислотних амінокислот на нейтральні амінокислоти (наприклад, Asp, Glu, Asn або Gin на Val, Ile, Leu або Met), ароматичних амінокислот на основні або кислотні амінокислоти (наприклад, Phe, Tyr або Trp на Asp, Asn, Glu або Gin) або будь-яке інше заміщення, яке не замінює амінокислоту подібною амінокислотою.

#### Гібридизація нуклеїнових кислот

Нуклеїнові кислоти "гібридизуються", коли вони асоціюються, звичайно в розчині. Нуклеїнові кислоти гібридизуються завдяки різноманітним, докладно описаним фізико-хімічним силам, таким як водневі зв'язки, виключення розчинника, накладення основ і т.п. Детальну інформацію стосовно гібридизації нуклеїнових кислот можна знайти в Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology- Hybridization with Nucleic Acid Probes*, part I, chapter 2, "Overview of principles of hybridization and the strategy of nucleic acid probe assays", (Elsevier, New York), а також в Ausubel, вище, Hames and Higgins (1995) *Gene Probes 1*, IRL Press at Oxford University Press, Oxford, England (Hames and Higgins 1) and Hames and Higgins (1995) *Gene Probes 2*, IRL Press at Oxford University Press, Oxford, England (Hames and Higgins 2), де наданий детальний опис синтезу, введення міток, виявлення і кількісної оцінки ДНК і РНК, включаючи олігонуклеотиди.

"Жорсткі умови гібридизації-промивки" в контексті дослідів з гібридизацією нуклеїнових кислот, таких як саузерн- і нозерн-гібридизація, є послідовність-залежними і є різними при різних параметрах навколишнього середовища. Детальну інформацію стосовно гібридизації нуклеїнових кислот можна знайти в Tijssen (1993), вище, і в Hames and Higgins 1 and Hames and Higgins 2, вище.

Загалом, у цьому винаході "високо жорсткі" умови гібридизації і промивки вибирають приблизно на  $<5^{\circ}\text{C}$  нижче температури плавлення ( $T_m$ ) для специфічної послідовності при визначеній іонній силі і рН (як зазначено нижче, вираз "високо жорсткі умови" може також використовуватися у порівняльному сенсі).  $T_m$  являє собою температуру (при визначеній іонній силі і рН), при якій 50% тестової послідовності гібридується з ідеально співпадаючим зондом. Високо жорсткі умови підбирають рівними  $T_m$  для конкретного зонда.

$T_m$  нуклеїновокислотного дуплекса означає температуру, при якій дуплекс денатурується на 50% в даних умовах, і вона являє собою безпосередню міру стабільності нуклеїновокислотного гібрида. Так,  $T_m$  відповідає температурі середньої точки переходу від спіралі до випадкового клубка; вона залежить від довжини, нуклеотидного складу і іонної сили для довгих нуклеотидних відрізків.

Після гібридизації негібридизований нуклеїновокислотний матеріал може бути видалений серією промивок, жорсткість яких може регулюватися в залежності від бажаних результатів. Умови промивки низької жорсткості (тобто, з ви-

користанням вищої концентрації солі і нижчої температури) підвищують чутливість, але можуть сприяти виробленню сигналів неспецифічної гібридизації і високих фонових сигналів. Умови більш високої жорсткості (тобто, з використанням нижчої концентрації солі і вищої температури, ближчої до температури гібридизації) знижують фоновий сигнал, звичайно залишаючи лише специфічний сигнал. Див. публікацію Rapley, R. and Walker, J.M. eds., *Molecular Biomethods Handbook* (Humana Press, Inc. 1998) (тут далі "Rapley and Walker"), що включена в цю заявку у вигляді посилання у всій її повноті для всіх призначень.

$T_m$  ДНК-ДНК-дуплекса може бути оцінена за допомогою рівняння 1 таким чином:

$$T_r (^{\circ}\text{C}) = 81,5^{\circ}\text{C} + 16,6 (\log_{10}M) + 0,41 (\%G+C) - 0,72 (\%f) - 500/n,$$

де  $M$  - молярність моновалентних катіонів (звичайно  $\text{Na}^+$ ),  $(\%G+C)$  - процент нуклеотидів гуанозину (G) і цистозину (C),  $(\%f)$  - процент формаліну і  $n$  - кількість нуклеотидних основ (тобто, довжина) гібриду. Див. Rapley and Walker, вище.

$T_m$  РНК-ДНК-дуплексу може бути оцінена за допомогою рівняння 2 таким чином:

$$T_m (^{\circ}\text{C}) = 79,8^{\circ}\text{C} + 18,5 (\log_{10}M) + 0,58 (\%G+C) - 11,8 (\%G+C)^2 - 0,56 (\%f) - 820/n,$$

де  $M$  - молярність моновалентних катіонів (звичайно  $\text{Na}^+$ ),  $(\%G+C)$  - процент нуклеотидів гуанозину (G) і цистозину (C),  $(\%f)$  - процент формаміду і  $n$  - кількість нуклеотидних основ (тобто, довжина) гібриду. Там же.

Рівняння 1 і 2 звичайно точні лише для гібридних дуплексів, довгих ніж 100-200 нуклеотидів. Там же.

$T_m$  нуклеїновокислотних послідовностей, коротших ніж 50 нуклеотидів, може бути розрахована таким чином:

$$T_m (^{\circ}\text{C}) = 4(G+C) + 2(A+T),$$

де  $A$  (аденін),  $C$ ,  $T$  (тимін) і  $G$  є числами відповідних нуклеотидів.

Приклад жорстких гібридизаційних умов для гібридизації комплементарних нуклеїнових кислот, що містять більше ніж 100 комплементарних залишків на фільтрі в методі саузерн- або нозерн-блоттингу, включає 50%-й формамід з 1мг гепарину при  $42^{\circ}\text{C}$ , при проведенні гібридизації протягом ночі. Приклад жорстких умов промивки включає промивку  $0,2\times\text{SSC}$  при  $65^{\circ}\text{C}$  протягом 15 хвилин (опис буферу SSC owe. Sambrook, вище). Часто перед промивкою високої жорсткості застосовують промивку низької жорсткості для усунення фонового сигналу зонда. Приклад умов низької жорсткості включає  $2\times\text{SSC}$  при  $40^{\circ}\text{C}$  15 хвилин.

Загалом, відношення сигнал/шум у 2,5-5 разів вище, ніж для неспорідненого зонда в конкретному досліді з гібридизацією означає виявлення специфічної гібридизації. Виявлення щонайменше жорсткої гібридизації між двома послідовностями в контексті цього винаходу показує відносно велику структурну подібність або гомологічність, наприклад, до нуклеїнових кислот за винаходом, що вказані в наведених тут переліках послідовностей.

Як було зазначено, "високо жорсткі" умови вибирають приблизно на  $\leq 5^{\circ}\text{C}$  нижче температу-



ри плавлення ( $T_m$ ) для конкретної послідовності при визначеній іонній силі і рН. Цільові послідовності, близько споріднені або ідентичні нуклеотидній послідовності, що представляє інтерес (наприклад, "зонд"), можуть бути ідентифіковані в високо жорстких умовах. Умови нижчої жорсткості придатні для послідовностей, що є менш комплементарними. Див., наприклад, Rapley and Walker, вище.

Порівняльна гібридизація може використовуватися для ідентифікації нуклеїнових кислот за винаходом, і цей спосіб порівняльної гібридизації є переважним способом розрізнення нуклеїнових кислот за винаходом. Виявлення високо жорсткої гібридизації між двома нуклеотидними послідовностями в контексті цього винаходу означає відносно високу структурну подібність/гомологічність, наприклад, до нуклеїнових кислот, наведених тут в переліку послідовностей. Високо жорстка гібридизація між двома нуклеотидними послідовностями демонструє таку ступінь подібності або гомологічності структури, складу нуклеотидних основ, розміщення або порядку, що є більшою, ніж виявлено жорсткими умовами гібридизації. Зокрема, виявлення високо жорсткої гібридизації в контексті цього винаходу означає високу структурну подібність або структурну гомологічність (наприклад, нуклеотидної структури, складу нуклеотидних основ, розміщення або порядку), наприклад, до нуклеїнових кислот, наведених тут в переліках послідовностей. Наприклад, бажано ідентифікувати тестові нуклеїнові кислоти, які гібридизуються із зразковими нуклеїновими кислотами, наведеними тут, в жорстких умовах.

Таким чином, одною мірою жорсткої гібридизації є здатність гібридизуватися з однією з перелічених нуклеїнових кислот (наприклад, нуклеїновокислотними послідовностями SEQ ID NO:1 - SEQ ID NO:5 і SEQ ID NO:11 - SEQ ID NO:262 і комплементарними до них полінуклеотидними послідовностями) в високо жорстких умовах (або дуже жорстких умовах, або умовах гібридизації ультра-високої жорсткості, або умовах гібридизації ультра-ультра-високої жорсткості). Жорсткі умови гібридизації (а також високо жорсткі умови, умови гібридизації ультра-високої жорсткості, або умови гібридизації ультра-ультра-високої жорсткості) і умови промивки можуть бути легко визначені емпіричним шляхом для будь-якої тестової нуклеїнової кислоти. Наприклад, для визначення умов високо жорсткої гібридизації і промивки умови гібридизації і промивки поступово посилюють (наприклад, шляхом підвищення температури, зниження концентрації солі, підвищення концентрації детергента і/або підвищення концентрації органічних розчинників, таких як формалін, на етапах гібридизації або промивки), доки не буде задовільнений набір критеріїв. Наприклад, умови гібридизації і промивки поступово посилюють, доки зонд, що включає одну або більше нуклеїновокислотних послідовностей, вибраних з SEQ ID NO:1 - SEQ ID NO:5 і SEQ ID NO:11 - SEQ ID NO:262 і комплементарних до них полінуклеотидних послідовностей, не буде зв'язуватися з ідеально співпадаючою комплементарною

ною ціллю (знов-таки, нуклеїновою кислотою, що включає одну або більше нуклеїновокислотних послідовностей, вибраних з SEQ ID NO:1 - SEQ ID NO:5 і SEQ ID NO:11 - SEQ ID NO:262 і комплементарних до них полінуклеотидних послідовностей), з відношенням сигнал/шум щонайменше в 2,5, можливо, 5 або більше разів вище, ніж спостерігається для гібридизації зонда з неспівпадаючою ціллю. В цьому випадку неспівпадаюча ціль являє собою нуклеїнову кислоту, що відповідає нуклеїновій кислоті (яка не входить в наведений тут перелік послідовностей), присутній в загальнодоступній базі даних, такий як GenBank™, на момент реєстрації цієї заявки. Такі послідовності можуть бути ідентифіковані в GenBank. Приклади включають інвентарні номери Z99109 і Y09476. Додаткові такі послідовності можуть бути ідентифіковані, наприклад, в GenBank, фахівцем середньої кваліфікації.

Кажуть, що тестова нуклеїнова кислота специфічно гібридизується з зондовою нуклеїновою кислотою, якщо вона гібридизується з зондом щонайменше наполовину так добре, як з ідеально співпадаючою комплементарною ціллю, тобто, щонайменше з половинним відношенням сигнал/шум порівняно з гібридизацією зонда з ціллю в умовах, в яких ідеально співпадаючий зонд зв'язується з ідеально співпадаючою комплементарною ціллю з відношенням сигнал/шум щонайменше в ~2-10, іноді 20, 50 або більше разів, ніж спостерігається для гібридизації з будь-яким з неспівпадаючих полінуклеотидів з інвентарними номерами Z99109 і Y09476.

Умови гібридизації і промивки ультра-високої жорсткості включають такі, в яких жорсткість умов гібридизації і промивки посилюють, доки відношення сигнал/шум для зв'язування зонда з ідеально співпадаючою комплементарною цільовою нуклеїновою кислотою не стане щонайменше в 10 разів вище, ніж спостерігається для гібридизації з будь-якою з неспівпадаючих цільових нуклеїнових кислот з інвентарними номерами GenBank Z99109 і Y09476. Цільову нуклеїнову кислоту, яка гібридизується з зондом в таких умовах, щонайменше з половинним відношенням сигнал/шум порівняно з тим, що спостерігається для ідеально співпадаючої комплементарної цільової нуклеїнової кислоти, називають такою, що зв'язується з зондом в умовах гібридизації ультра-високої жорсткості.

Подібним чином, навіть вищі рівні жорсткості можуть бути визначені шляхом поступового посилення умов гібридизації і/або промивки у відповідному досліді з гібридизацією. Наприклад, жорсткість умов гібридизації і промивки посилюють, доки відношення сигнал/шум для зв'язування зонда з ідеально співпадаючою комплементарною цільовою нуклеїновою кислотою не стане щонайменше в 10, 20, 50, 100 або 500 або більше разів вище, ніж спостерігається для гібридизації з будь-якою з неспівпадаючих цільових нуклеїнових кислот з інвентарними номерами GenBank Z99109 і Y09476. Цільову нуклеїнову кислоту, яка гібридизується з зондом в таких умовах, щонайменше з половинним відношенням сигнал/шум порівняно з тим, що спостерігається для ідеально

співпадаючої комплементарної цільової нуклеїнової кислоти, називають такою, що зв'язується з зондом в умовах гібридизації ультра-ультра-високої жорсткості.

Цільові нуклеїнові кислоти, які гібридизуються з нуклеїновими кислотами, представленими послідовностями SEQ ID NO:1 - SEQ ID NO:5 і SEQ ID NO:11 - SEQ ID NO:262 в умовах високої, ультра-високої і ультра-ультра-високої жорсткості, є рисою цього винаходу. Приклади таких нуклеїнових кислот включають нуклеїнові кислоти з кількома німими або консервативними нуклеїновокислотними заміщеннями у порівнянні з даною нуклеїновокислотою послідовністю.

Нуклеїнові кислоти, які не гібридизуються одна з одною в жорстких умовах, все одно є по суті ідентичними, якщо поліпептиди, які вони кодують, є по суті ідентичними. Це відбувається, наприклад, якщо копію нуклеїнової кислоти створюють з використанням максимальної дегенерації кодонів, дозволеної генетичним кодом, або якщо вироблені антисироватки або антисироватка проти одної або більше з послідовностей SEQ ID NO:6 - SEQ ID NO:10 і SEQ ID NO:263 - SEQ ID NO:514, що були відняті з використанням поліпептидів, закодованих відомими нуклеотидними послідовностями, включаючи інвентарний номер GenBank CAA70664. Додаткові подробиці імунологічної ідентифікації поліпептидів за винаходом описані нижче. Крім того, для розрізнення дуплексів з послідовностями, що складаються менш ніж з 100 нуклеотидів, може бути використана відома фахівцям процедура гібридизації TMACI. Див., наприклад, публікацію Sorg, U. et al. 1 Nucleic Acids Res. (Sept. 11, 1991) 19 (17), включену в цю заявку у вигляді посилання у всій її повноті для будь-якого призначення.

В одному з аспектів винахід стосується нуклеїнової кислоти, яка включає унікальну підпослідовність в нуклеїновій кислоті, вибраній з SEQ ID NO:1 - SEQ ID NO:5 і SEQ ID NO:11 - SEQ ID NO:262. Унікальна підпослідовність є унікальною порівняно з нуклеїновою кислотою, відповідною до інвентарних номерів GenBank Z99109 і Y09476. Такі унікальні підпослідовності можуть бути визначені шляхом вирівнювання будь-якої з SEQ ID NO:1 - SEQ ID NO:5 і SEQ ID NO:11 - SEQ ID NO:262 з повним набором нуклеїнових кислот, представлених інвентарними номерами GenBank Z99109, Y09476 або іншими спорідненими послідовностями в загальнодоступних базах даних на момент реєстрації цієї заявки. Вирівнювання може виконуватися з використанням алгоритму BLAST, встановленого на стандартні параметри. Будь-яка унікальна підпослідовність є корисною, наприклад, у якості зонда для ідентифікації нуклеїнових кислот за винаходом.

Подібним чином, винахід стосується поліпептиду, який включає унікальну підпослідовність в поліпептиді, вибраному з SEQ ID NO:6 - SEQ ID NO:10 і SEQ ID NO:263 - SEQ ID NO:514. У цьому випадку унікальна підпослідовність є унікальною порівняно з поліпептидом, відповідним до інвентарного номеру GenBank CAA70664. Поліпептид також вирівнюють з послідовностями, представленими інвентарним номером GenBank

CAA70664. Слід зазначити, що, коли послідовність відповідає нетрансльованій послідовності, такий як псевдоген, відповідний поліпептид виробляють просто шляхом трансляції in silico нуклеїновокислотної послідовності в амінокислотну послідовність, де рамку зчитування вибирають таким чином, щоб вона відповідала рамці зчитування гомологічних GAT-полінуклеотидів.

Винахід також стосується цільових нуклеїнових кислот, які гібридизуються в жорстких умовах з унікальним кодуючим олігонуклеотидом, що кодує унікальну підпослідовність в поліпептиді, вибраному з SEQ ID NO:6 - SEQ ID NO:10 і SEQ ID NO:263 - SEQ ID NO:514, де унікальна підпослідовність є унікальною порівняно з поліпептидом, відповідним до будь-якого з контрольних поліпептидів. Унікальні послідовності визначають, як описано вище.

В одному з прикладів жорсткі умови вибирають таким чином, щоб олігонуклеотид, ідеально комплементарний до кодуючого олігонуклеотиду, гібридизувався з кодуючим олігонуклеотидом з відношенням сигнал/шум щонайменше в 2,5-10 разів вище, переважно щонайменше в 5-10 разів вище, ніж при гібридизації ідеально комплементарного олігонуклеотиду з контрольною нуклеїновою кислотою, відповідною до будь-якого з контрольних поліпептидів. Умови можуть вибиратися таким чином, щоб в конкретному досліді спостерігалися вищі відношення сигнал/шум, наприклад, в 15, 20, 30, 50 разів або більше. В цьому прикладі цільова нуклеїнова кислота гібридується з унікальним кодуючим олігонуклеотидом щонайменше з 2-кратним відношенням сигнал/шум порівняно з гібридизацією контрольної нуклеїнової кислоти з кодуючим олігонуклеотидом. Знов-таки, можуть бути вибрані вищі відношення сигнал/шум, наприклад, приблизно в 2,5, 5, 10, 20, 30, 50 разів або більше.

Конкретний сигнал буде залежати від мітки, що використовується у відповідному досліді, наприклад, флуоресцентної мітки, радіоактивної мітки і т.п.

Вектори, промотори і системи експресії

Цей винахід також стосується рекомбінантних конструкцій, що включають одну або більше з нуклеїновокислотних послідовностей у діапазоні, описаному вище. Конструкції включають вектор, такий як плазмід, космід, фаг, вірус, штучна бактеріальна хромосома (BAC), штучна хромосома дріжджів (YAC) і т.п., в яку вставлена нуклеїновокислотна послідовність за винаходом, в прямій або зворотній орієнтації. В переважному аспекті цього варіанту здійснення конструкція додатково включає регуляторні послідовності, в тому числі, наприклад, промотор, оперативно зв'язаний з послідовністю. Великі кількості придатних векторів і промоторів відомі фахівцям і серійно випускаються.

Основні документи, що описують молекулярно-біологічні технології, корисні для цього винаходу, в тому числі застосування векторів, промоторів і багато інших релевантних питань, включають Berger and Kimmel, Guide to Molecular Cloning Techniques. Methods in Enzymology, volume 152 Academic Press, Inc., San Diego, CA

("Berger"); Sambrook et al., *Molecular Cloning - A Laboratory Manual* (2nd Ed.), volumes 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989 ("Sambrook"); and *Current Protocols in Molecular Biology*. F.M. Ausubel et al., eds. *Current Protocols*, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (supplemented through 1999) ("Ausubel"). Приклади протоколів, достатні для створення уяви про способи ампліфікації *in vitro*, включають полімеразну ланцюгову реакцію (ПЛР), лігазну ланцюгову реакцію (ЛЛР), Q $\beta$ -репліказну ампліфікацію та інші технології, опосереднені РНК-полімеразою (наприклад, NASBA), наприклад, для вироблення гомологічних нуклеїнових кислот за винаходом, можуть бути знайдені в Berger, Sambrook, and Ausubel, а також Mullis et al., (1987) U.S. Patent No. 4,683,202; *PCR Protocols A Guide to Methods and Applications* (Innis et al., eds.) Academic Press Inc. San Diego, CA (1990); Amheim & Levinson (October 1, 1990) *Chemical and Engineering News* 36:47; *The Journal Of NIH Research* (1991) 3:81-94; Kwoh et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:1173; Guatelli et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:1874; Lomell et al. (1989) *J. Clin. Chem.* 35:1826; Landegren et al., (1988) *Science* 241:1077-1080; Van Brunt (1990) *Biotechnology* 8:291-294; Wu and Wallace, (1989) *Gene* 4:560; Bardnager et al. (1990) *Gene* 89:117, and Sooknanan and Malek (1995) *Biotechnology* 13:563-564. Вдосконалені способи клонування *in vitro* ампліфікованих нуклеїнових кислот описані в Wallace et al., патент США №5,426,039. Вдосконалені способи ампліфікації великих нуклеїнових кислот шляхом ПЛР описані у Cheng et al. (1994) *Nature* 369:684-685 і наведених там посиланнях, де отримують ПЛР-реплікони розміром до 40kb. Фахівцям буде зрозуміло, що по суті будь-яка РНК може бути перетворена в двоспиральну ДНК, придатну для рестрикційної ферментації, ПЛР-розширення і секвенування з використанням зворотної транскриптази і полімерази. Див., наприклад, Ausubel, Sambrook and Berger, усі вище.

Винахід також стосується сконструйованих клітин-хазяїв, трансдукованими (трансформованими або трансфектованими) вектором за винаходом (наприклад, вектором клонування за винаходом або вектором експресії за винаходом, а також вироблення поліпептидів за винаходом за допомогою рекомбінантних технологій. Вектором може бути, наприклад, плазміда, вірусна частка, фаг і т.д. Сконструйовані клітини-хазяї можуть культивуватися в звичайних живильних середовищах, модифікованих відповідним чином для активації промоторів, селекції трансформантів або ампліфікації GAT-гомологічного гена. Умови культивування, такі як температура, рН і т.п., такі ж, як звичайно використовуються для клітини-хазяїна, вибраного для експресії, добре відомі фахівцям і можуть бути знайдені в посиланнях, наведених тут, наприклад, Sambrook, Ausubel and Berger, а також, наприклад, Freshney (1994) *Culture of Animal Cells, a Manual of Basic Technique*, third edition, Wiley-Liss, New York, і посилання, наведені там.

GAT-поліпептиди за винаходом можуть вироблятися в нетваринних клітинах, таких як рослинні, дріжджові, грибові, бактеріальні і т.п. На додаток до Sambrook, Berger and Ausubel, деталі щодо нетваринних клітинних культур можуть бути знайдені в Payne et al. (1992) *Plant Cell and Tissue Culture in Liquid Systems* John Wiley & Sons, Inc. New York, NY; Gamborg and Phillips (eds) (1995) *Plant Cell, Tissue and Organ Culture: Fundamental Methods* Springer Lab Manual, Springer-Verlag (Berlin Heidelberg New York) and Atlas and Parks (eds) *The Handbook of Microbiological Media* (1993) CRC Press, Boca Raton, FL.

Полінуклеотиди за винаходом можуть бути вбудовані в будь-який з різноманітних векторів експресії, придатних для експресії поліпептиду. Придатні вектори включають хромосомні, нехромосомні і синтетичні ДНК-послідовності, наприклад, похідні SV40; бактеріальні плазмідні; фагові ДНК; бакуловіруси; дріжджові плазмідні; вектори, отримані з комбінацій плазмідних і фагових ДНК, вірусних ДНК, таких як вірус коров'ячої віспи, аденовірус, вірус вітряної віспи, псевдосказу, адено-асоційований вірус, ретровіруси та багато інших. Може бути використаний будь-який вектор, що трансдукує генетичний матеріал в клітину, і, якщо бажана реплікація, такий, що реплікується і є життєздатним у відповідній клітині-хазяїні.

Полінуклеотид за винаходом, вбудований в вектор експресії, є оперативно зв'язаним з відповідною транскрипційною контрольною послідовністю (промотором), що направляє синтез мРНК. Приклади таких транскрипційних контрольних послідовностей, найбільш придатних для застосування в трансгенних рослинах, включають промотори мозаїчного вірусу кольорової капусти (CaMV), мозаїчного вірусу норицника (FMV) і вірусу облямування жилок суниці (SVBV), описані в тимчасовій заявці США No. 60/245,354. Інші промотори, відомі як контролюючі експресію генів в прокариотичних або еукаріотичних клітинах або їх вірусах і які можуть бути застосовані в деяких варіантах здійснення винаходу, включають промотор SV40, промотор *E.coli lac* або *trp*, фаговий промотор лямбда-Р<sub>L</sub>. Вектор експресії може містити сайт зв'язування рибосом для ініціації трансляції, а також термінатор трансляції. Вектор також може включати відповідні послідовності для ампліфікації експресії, наприклад, енхансер. Крім того, вектори експресії за винаходом містять один або більше генів маркерів, що селектуються, для створення фенотипової ознаки для селекції трансформованих клітин-хазяїв, наприклад, дигідрофолат-редуктазу або стійкості до неомицину для еукаріотичної клітинної культури, або стійкості до тетрацикліну або ампіциліну в *E.coli*.

Вектори за винаходом можуть бути застосовані для трансформації відповідного хазяїна, яка дозволить хазяїну експресувати протеїн або поліпептид за винаходом. Приклади відповідних хазяїв експресії включають: бактеріальні клітини, такі як *E.coli*, *B. subtilis*, *Streptomyces* і *Salmonella typhimurium*; грибові клітини, такі як *Saccharomyces cerevisiae*, *Pichia pastoris* і *Neurospora crassa*; клітини комах, такі як *Drosophila* і *Spodoptera frugiperda*; клітини ссав-

ців, такі як CHO, COS, BHK, HEK 293 або меляноми Bowes; або рослинні клітини або експлантати і т.д. Зрозуміло, що не всі клітини або клітинні лінії повинні мати здатність виробляти повністю функціональні GAT-поліпептиди; наприклад, можуть вироблятися антигенні фрагменти GAT-поліпептиду. Винахід не обмежений переліченими клітинами-хазяями.

В бактеріальних системах кількість векторів експресії може вибиратися в залежності від призначення GAT-поліпептиду. Наприклад, якщо потрібні великі кількості GAT-поліпептиду або його фрагментів для серійного випуску або для індукції антитіл, можуть бути бажані вектори, які забезпечують високі рівні експресії злитих протеїнів, що легко очищаються. Такі вектори включають, але не обмежуються, багатфункціональні вектори клонування і експресії *E. coli*, такі як BLUESCRIPT (Stratagene), в яких GAT-поліпептид-кодуюча послідовність може бути лігвана в вектор в одній рамці з послідовностями для аміно-термінального Met і наступними 7 залишками бета-галактозидази, для отримання гібридного протеїну; вектори pIN (Van Heeke & Schuster (1989) *J Biol Chem* 264:5503-5509); вектори pET (Novagen, Madison WI) і т.п.

Подібним чином, в дріжджах *Saccharomyces cerevisiae* кількість векторів, що містять конститутивні промотори або такі, що індукуються, такі як альфа-фактор, алкоголь-оксидаза і PGH, можуть бути використані для вироблення GAT-поліпептидів за винаходом. Див. Ausubel et al. (вище) і Grant et al. (1987; *Methods in Enzymology* 153:516-544).

В клітинах-хазяях ссавців можуть бути застосовані різноманітні системи експресії, включаючи системи на основі вірусів. У випадках, коли у якості вектора експресії використовується аденовірус, кодуюча послідовність, наприклад, GAT-поліпептиду, може бути лігвана в транскрипційно-трансляційний комплекс аденовірусу, що складається з пізнього промотора і троїстої лідерної послідовності. Інсерція кодуючої області GAT-поліпептиду в заміні області вірусного генома E1 і E3 створить життєздатний вірус, що здатен експресувати GAT в інфікованій клітині-хазяїні (Logan and Shenk (1984) *Proc Natl Acad Sci USA* 81:3655-3659). Крім того, енхансери транскрипції, такі як енхансер вірусу саркоми Рауса (RSV), можуть бути використані для підвищення експресії в клітинах-хазяях ссавців.

Подібним чином, в рослинних клітинах експресія може збуджуватися трансгеном, вбудованим в рослинну хромосому, або цитоплазматичним шляхом, епісомною або вірусною нуклеїновою кислотою. У випадку стабільно вбудованих трансгенів часто буває бажано отримати послідовності, здатні збуджувати конститутивну або індуковану експресію GAT-поліпептидів за винаходом, наприклад, з використанням вірусних регуляторних послідовностей, таких як CaMV, або регуляторних послідовностей рослинного походження. Описані численні регуляторні послідовності рослинного походження, включаючи послідовності, які направляють експресію тканино-специфічним чином, наприклад, TobRB7, па-

татин B33, промотори гена GRP, промотор *rbcs*-3A і т.п. В альтернативі, високі рівні експресії можуть бути досягнуті екзогенними послідовностями тимчасової експресії з рослинного вірусного вектора, наприклад, TMV, BMV і т.д. Звичайно переважними є трансгенні рослини, конститутивно експресуючі GAT-поліпептид за винаходом, і регуляторні послідовності, селектовані таким чином, щоб забезпечити конститутивну стабільну експресію GAT-поліпептиду.

В деяких варіантах здійснення винаходу отримують GAT-поліпептидну конструкцію, придатну для трансформації рослинних клітин. Наприклад, бажаний GAT-поліпептид може бути вбудований в рекомбінантну касету експресії для полегшення введення гена в рослину і подальшої експресії закодованого поліпептиду. Касета експресії звичайно містить GAT-поліпептид або його функціональний фрагмент, оперативно зв'язаний з промоторною послідовністю та іншими регуляторними послідовностями ініціації транскрипції і трансляції, які будуть направляти експресію послідовності в цільових тканинах (наприклад, у всій рослині, листі, насінні) трансформованої рослини.

Наприклад, може бути використаний сильно або слабо конститутивний промотор, який буде направляти експресію GAT-поліпептиду у всіх тканинах рослини. Такі промотори є активними у більшості навколишніх умов і станів розвитку або диференціації клітин. Приклади конститутивних промоторів включають 1'- або 2'-промотор, отриманий з T-ДНК *Agrobacterium tumefaciens*, та інші області ініціації транскрипції з різних рослинних генів, відомі фахівцям. Після ознайомлення з цим описом фахівцеві буде зрозуміло, що у випадках, коли надмірна експресія GAT-поліпептиду є згубною для рослини або небажаною з інших причин, для низьких рівнів експресії може бути використаний слабо конститутивний промотор. У випадках, коли високі рівні експресії не шкідливі для рослини, може бути використаний сильний промотор, наприклад, т-РНК або інший промотор *pol III*, або сильний промотор *pol II*, такий як промотор мозаїчного вірусу кольорової капусти.

В альтернативі, рослинний промотор може знаходитися під контролем навколишнього середовища. Такі промотори називаються тут промоторами, що індукуються. Приклади умов навколишнього середовища, які можуть здійснювати транскрипцію за допомогою промоторів, що індукуються, включають атаку патогенів, анаеробні умови або присутність світла.

Промотори, що використовуються в цьому винаході, можуть бути "тканино-специфічними" і, як такі, можуть знаходитися під еволюційним контролем, внаслідок того, що поліпептид експресується лише у певних тканинах, таких як листя і насіння. У варіантах здійснення винаходу, в яких в конструкцію вбудовуються одна або більше нуклеїнових кислотних послідовностей, ендегенних до рослинної системи, ендегенні промотори (або їх варіанти) з цих генів можуть бути використані для управління експресією генів в трансфектованій рослині. Тканино-специфічні промотори можуть бути також використані для

управління експресією гетерологічних полінуклеотидів.

Загалом, конкретний промотор, що використовується в касеті експресії в рослинах, залежить від призначення. Придатними є будь-які з численних промоторів, що направляють транскрипцію в рослинних клітинах. Промотор може бути конститутивним або таким, що індукується. На додаток до промоторів, описаних вище, промотори бактеріального походження, що діють в рослинах, включають промотор октопін-синтази, промотор нопалін-синтази та інші промотори, отримані з нативних Ті-плазмід (див. Herrera-Estrella et al., (1983) *Nature* 303:209-213). Вірусні промотори включають РНК-промотори 35S і 19S мозаїчного вірусу кольорової капусти (Odell et al. (1985) *Nature* 313:810-812). Інші рослинні промотори включають малий субодиничний промотор рибулоза-1,3-бісфосфат-карбоксилази і фазеоліновий промотор. Також може бути використана промоторна послідовність з гена E8 та інших генів. Ізоляція і послідовність промотора E8 детально описані у Deikman and Fischer (1988) *EMEA* 7:3315-3327.

Для ідентифікації промоторів-кандидатів 5'-частини геномного клону аналізують на послідовності, характерні для промоторних послідовностей. Наприклад, елементи промоторної послідовності включають консенсусну послідовність TATA box (TATAAT), яка звичайно знаходиться на 20-30 пар основ вище від початкового сайту транскрипції. В рослинах ще вище від TATA box, в положеннях 80-100, звичайно знаходиться промоторний елемент з серією аденінів, оточуючих тринуклеотид G (або T), як описано у Messing et al. (1983) *Genetic Engineering in Plants*. Kosage, et al. (eds.), pp.221-227.

При отриманні полінуклеотидних конструкцій, наприклад, векторів, за винаходом можуть також використовуватися послідовності, відмінні від промотора і приєднаного полінуклеотиду. Якщо бажана нормальна експресія поліпептиду, може бути включена поліаденілаційна область у 3'-закінчення GAT-кодуючої області. Поліаденілаційна область може бути отримана, наприклад, з різноманітних рослинних генів або з Т-ДНК.

Конструкція може також включати маркерний ген, який надає рослинним клітинам фенотип, що селектується. Наприклад, маркер може кодувати біоцидну толерантність, зокрема, толерантність до антибіотиків, таку як толерантність до хлорсульфурону або фосфіотрицину (активний інгредієнт в гербіцидах біалафос і Basta).

Специфічні сигнали ініціації можуть сприяти ефективній трансляції GAT-полінуклеотид-кодуючої послідовності за винаходом. Ці сигнали можуть включати, наприклад, кодон ініціації ATG і суміжні послідовності. У випадках, коли GAT-полінуклеотид-кодуюча послідовність, її кодон ініціації і вищерозташовані послідовності інсертовані у відповідний вектор експресії, додаткові сигнали контролю трансляції можуть бути не потрібні. Проте у випадках, коли інсертована лише кодуюча послідовність (наприклад, кодуюча послідовність зрілого протеїну) або її частина, необхідна присутність екзогенних сигналів контролю

трансляції, що включають кодон ініціації. Крім того, кодон ініціації повинен знаходитися у належній рамці зчитування, щоб гарантувати транскрипцію цілої вставки. Екзогенні транскрипційні елементи і кодони ініціації можуть бути різного походження, як природного, так і синтетичного. Ефективність експресії може бути посилена шляхом включення енхансерів, відповідних до клітинної системи, що використовується (Scharf D et al. (1994) *Results Probl Cell Differ* 20:125-62; Bittner et al. (1987) *Methods in Enzymol* 153:516-544).

Послідовності секреції/локалізації

Полінуклеотиди за винаходом можуть бути також злиті, наприклад, в рамці з нуклеїновими кислотами, кодуючими послідовність секреції/локалізації, щоб націлити експресію поліпептиду в бажаний відсік клітини, мембрану або органелу клітини ссавців, або направити секрецію поліпептиду в периплазматичний простір або в середовище клітинної культури. Такі послідовності відомі фахівцям і включають лідерні пептиди секреції, таргетинг-послідовності органел (наприклад, послідовності ядерної локалізації, сигнали ER-утримання, послідовності транзиту в мітохондрії, послідовності транзиту в хлоропласти), послідовності мембранної локалізації/якірні послідовності (наприклад, стоп-трансферні послідовності, GPI-якірні послідовності) і т.п.

У переважному варіанті здійснення винаходу полінуклеотид за винаходом зливають в рамці з N-кінцевою послідовністю транзиту в хлоропласти (або пептидною хлоропласт-транзитною послідовністю), отриманою з гена, кодуючого поліпептид, що звичайно таргетований на хлоропласт. Такі послідовності звичайно багаті на серин і треонін, бідні на аспартат, глутамат і тирозин і у більшості випадків мають центральний домен, багатий на позитивно заряджені амінокислоти.

Хазяї експресії

У наступному варіанті здійснення винахід стосується клітин-хазяїв, що містять вищеописані конструкції. Клітина-хазяїн може бути еукаріотичною клітиною, такою як клітина ссавців, дріжджова клітина або рослинна клітина, або клітина-хазяїн може бути прокаріотичною клітиною, такою як бактеріальна клітина. Введення конструкції в клітину-хазяїна може здійснюватися шляхом кальцієво-фосфатної трансфекції, DEAE-Dextran-опосередкованої трансфекції, електропорації або іншими традиційними способами (Davis, L, Dibner, M., and Battey, I. (1986) *Basic Methods in Molecular Biology*).

Штам клітин-хазяїв може бути вибраний за їхньою здатністю модулювати експресію інсертованих послідовностей або обробляти експресований протеїн бажаним чином. Такі модифікації протеїну включають, але не обмежуються, ацетилювання, карбоксилювання, глікозилювання, фосфорилювання, ліпидування і ацилювання. Пост-трансляційна обробка, яка розщеплює "пре" або "препро" форму протеїну, також може бути важлива для коректної інсерції, складання і/або функціонування. Пізні клітини-хазяї, такі як *E.coli*, *Bacillus* sp., дріжджові клітини або клітини ссавців, такі як CHO, HeLa, BHK, MDCK, 293, WI38 і т.д., мають специфічний клітинний апарат і типові

механізми, наприклад, пост-трансляційної діяльності, і можуть бути вибрані для забезпечення бажаних модифікацій і обробки введеного чужорідного протеїну.

Для довготривалого, високопродуктивного вироблення рекомбінантних протеїнів можуть бути застосовані стабільні системи експресії. Наприклад, рослинні клітини, експлантати або тканини, такі як паростки, листяні диски, які стабільно експресують поліпептид за винаходом, трансдукують за допомогою векторів експресії, що містять вірусні джерела реплікації або ендогенні елементи експресії і ген маркера, що селектується. Після введення вектора клітинам дозволяють рости протягом періоду, визначеного для даного клітинного типу, наприклад, 1 або більше годин для бактеріальних клітин, 1-4 доби для рослинних клітин, 2-4 тижні для деяких рослинних експлантатів, в збагаченому середовищі до переносу в селекційне середовище. Задача маркера, що селектується, - надання стійкості до селекції, і його присутність забезпечує можливість росту і регенерації клітин, які успішно експресують введені послідовності. Наприклад, трансгенні рослини, експресуючі поліпептиди за винаходом, можуть бути безпосередньо селектовані на стійкість до гербіциду гліфосату. Стійкі зародки, виведені з стабільно трансформованих експлантатів, можуть бути розмножені, наприклад, з використанням методик тканинних культур, придатних для даного клітинного типу.

Клітини-хазяї, трансформовані нуклеотидною послідовністю, кодуною поліпептид за винаходом, можуть бути культивовані в умовах, придатних для експресії і регенерації закодованого протеїну з культури клітин. Протеїн або його фрагмент, вироблений рекомбінантною клітиною, може бути секретований, мембрано-зв'язаний або може міститися внутрішньоклітинно, в залежності від послідовності і/або вектора, що використовується. Фахівцям буде зрозуміло, що вектори експресії, які містять GAT-полінуклеотиди за винаходом, можуть бути сконструйовані з сигнальними послідовностями, які направляють секрецію зрілих поліпептидів через мембрану прокаріотичних або еукаріотичних клітин.

Додаткові поліпептидні послідовності

Полінуклеотиди за винаходом можуть також містити кодуєчу послідовність, зливу в рамці з маркерною послідовністю, яка, наприклад, полегшує очищення закодованого поліпептиду. Такі домени полегшення очищення включають, але не обмежуються, металохелатуючі пептиди, такі як гістидино-триптофанові модулі, які дозволяють очищення на іммобілізованих металах, послідовність, яка зв'язує глутатіон (наприклад, GST), гемаглютинін (HA)-таг (що відповідає епітопу, отриманому з протеїну гемаглютиніну інфлуенци; Wilson et al. (1984) Cell 37:767), мальтоза-зв'язуючі протеїнові послідовності, FLAG-епітоп, що використовується в системі очищення FLAGS extension/affinity (Immunex Corp, Seattle, WA) і т.д. Включення поліпептидної лінкерної послідовності, що розщеплюється протеазою, між доменом очищення і GAT-гомологічною послідовністю корисно для полегшення очистки. Один з векторів

експресії, придатний для застосування в композиціях і способах, описаних тут, забезпечує експресію злитого протеїну, що включає поліпептид за винаходом, злитий з полігістидиновою областю, відділеною сайтом ентерокиназного розщеплення. Гістидинові залишки полегшують очищення на IMIAC (афінна хроматографія на іммобілізованому іоні металу, описана у Porath et al. (1992) Protein Expression and Purification 3:263-281), в той час як сайт ентерокиназного розщеплення забезпечує засіб відділення GAT-гомологічного поліпептиду від злитого протеїну. Вектори pGEX (Promega; Madison, WI) також можуть бути застосовані для експресії чужорідних поліпептидів у якості злитих протеїнів з глутатіон-S-трансферазою (GST). Загалом, такі злиті протеїни є розчинними і легко можуть бути очищені з лізованих клітин шляхом адсорбції на ліганд-агарозному бісері (наприклад, глутатіон-агароза у випадку GST-злиття) з наступним елюванням у присутності вільного ліганду.

Вироблення і регенерація поліпептиду

Після трансдукції придатного штаму клітин-хазяїв і вирощування штаму до відповідної щільності клітин вибраний промотор індукують відповідним засобом (наприклад, температурним зсувом або хімічною індукцією) і культивують клітини протягом додаткового періоду. Клітини звичайно збирають шляхом центрифугування, руйнують фізичними або хімічними засобами і результируючий необроблений екстракт витримують для подальшого очищення. Мікробні клітини, використані в експресії протеїнів, можуть руйнуватися будь-яким зручним способом, включаючи цикли заморожування - розморожування, ультразвукову руйнацію, механічну руйнацію або використання лізуючих агентів, або іншими способами, відомими фахівцям.

Як зазначено вище, багато публікацій присвячено культивуванню і виробленню різноманітних клітин, включаючи клітини бактеріального, рослинного, тваринного (зокрема, клітини ссавців) і архебактеріального походження. Див., наприклад, Sambrook, Ausubel and Berger, (yci вище), а також Freshney (1994) Culture of Animal Cells, a Manual of Basic Technique, third edition, Wiley-Liss, New York, і посилання, наведені там; Doyle and Griffiths (1997) Mammalian Cell Culture: Essential Techniques John Wiley and Sons, NY; Humason (1979) Animal Tissue Techniques, fourth edition W.H. Freeman and Company; і Ricciardelli, et al., (1989) In vitro Cell Dev. Biol. 25:1016-1024. Література стосовно культивування і регенерації рослинних клітин: Payne et al. (1992) Plant Cell and Tissue Culture in Liquid Systems John Wiley & Sons, Inc. New York, NY; Gamborg and Phillips (eds) (1995) Plant Cell. Tissue and Organ Culture: Fundamental Methods Springer Lab Manual, Springer-Verlag (Berlin Heidelberg New York); Jones, ed. (1984) Plant Gene Transfer and Expression Protocols. Humana Press, Totowa, New Jersey and Plant Molecular Biolav (1993) R.R.D.Croy, Ed. Bios Scientific Publishers, Oxford, U.K. ISBN 0 12 198370 6. Середовища клітинних культур загалом описані в Atlas and Parks (eds) The Handbook of Microbiological Media (1993) CRC

Press, Boca Raton, FL. Додаткову інформацію щодо клітинних культур можна знайти в комерційній літературі, такий як Life Science Research Cell Culture Catalogue (1998) from Sigma-Aldrich, Inc (St Louis, MO) ("Sigma-LSRCCC") і, наприклад, The Plant Culture Catalogue і додатки до нього (1997), також від Sigma-Aldrich, Inc (St Louis, MO) ("Sigma-PPCS"). Додаткові подробиці стосовно трансформації рослинних клітин і вироблення трансгенних рослин наведені нижче.

Поліпептиди за винаходом можуть бути регенеровані і очищені з рекомбінантних клітинних культур будь-яким з численних способів, добре відомих у відповідній області техніки, включаючи осадження сульфатом амонію або етанолом, кислотну екстракцію, аніоно- або катіонообмінну хроматографію, фосфоцелюлозну хроматографію, хроматографію гідрофобної взаємодії, афінну хроматографію (наприклад, з використанням будь-якої з систем мічення, згаданих тут), гідроксилпатитну хроматографію і лектинову хроматографію. Якщо це бажано, можуть бути використані етапи перескладання протеїну для довершення конфігурації зрілого протеїну. Нарешті, у кінцевих етапах очищення може бути використана рідинна хроматографія високого тиску (HPLC). На додаток до посилань, наведених вище, у відповідній області відомо багато інших способів очищення, включаючи, наприклад, описані у Sandana (1997) Bioseparation of Proteins. Academic Press, Inc.; Bollag et al. (1996) Protein Methods. 2<sup>nd</sup> Edition Wiley-Uss, NY; Walker (1996) The Protein Protocols Handbook Humana Press, NJ, Harris and Angal (1990) Protein Purification Applications: A Practical Approach IRL Press at Oxford, Oxford, England; Harris and Angal Protein Purification Methods: A Practical Approach IRL Press at Oxford, Oxford, England; Scopes (1993) Protein Purification: Principles and Practice 3<sup>rd</sup> Edition Springer Verlag, NY; Janson and Ryden (1998) Protein Purification: Principles. High Resolution Methods and Applications. Second Edition Wiley-VCH, NY; Walker (1998) Protein Protocols on CD-ROM Humana Press, NJ.

У деяких випадках бажано продукувати GAT-поліпептид за винаходом у великому об'ємі для промислових і/або комерційних призначень. У таких випадках застосовують процедури ферментації в масі. У короткому викладі, GAT-полінуклеотид, наприклад, полінуклеотид, що містить будь-яку з SEQ ID NOS: 1-5 і 11-262 або інші нуклеїнові кислоти, кодує GAT-поліпептиди за винаходом, можуть бути клоновані в вектор експресії. Наприклад, в патенті США №5,955,310, Widner et al. "METHODS FOR PRODUCING A POLYPEPTIDE IN A BACILLUS CELL" описаний вектор з тандемними промоторами і стабілізуючими послідовностями, оперативно зв'язаними з поліпептид-кодуючою послідовністю. Після інсертування у вектор полінуклеотиду, що представляє інтерес, вектор трансформують в бактеріальний штам клітин-хазяїв, наприклад, *Bacillus subtilis* PL18011IE (amyE, apr, npr, spollE::Tn917). Введення вектора експресії в клітину *Bacillus* може бути здійснено, наприклад, шляхом трансформації протопластів (див., наприклад, Chang and

Cohen (1979) Molecular General Genetics 168:111), шляхом використання компетентних клітин (див., наприклад, Young and Spizizin (1961) Journal of Bacteriology 81:823, або Dubnau and Davidoff-Abelson (1971) Journal of Molecular Biology 56:209), шляхом електропорації (див., наприклад, Shigekawa and Dower (1988) Biotechniques 6:742), або шляхом кон'югації (див., наприклад, Koehler and Thome (1987) Journal of Bacteriology 169:5271), також див. Ausubel, Sambrook and Berger, усі вище.

Трансформовані клітини культивують у живильному середовищі, придатному для вироблення поліпептиду, з використанням способів, відомих в біотехнології. Наприклад, клітина може бути культивована шляхом культивування у вібраційній колбі, маломасштабної або крупномасштабної ферментації (включаючи безперервну, періодичну, періодичну з підживленням або твердофазну ферментацію) в лабораторних або промислових ферментерах у придатному середовищі і в умовах, що дозволяють експресію і ізоляцію поліпептиду. Культивування відбувається у придатному живильному середовищі, що включає джерела вуглецю і азоту і неорганічні солі, з використанням процедур, відомих у біотехнології. Придатні середовища серійно випускаються виробниками або можуть бути приготовані згідно з опублікованими композиціями (наприклад, в каталогах American Type Culture Collection). Секретований поліпептид може бути регенерований безпосередньо з середовища.

Результуючий поліпептид може бути ізолюваний способами, відовими у біотехнології. Наприклад, поліпептид може бути ізолюваний з живильного середовища за звичайними методиками, що включають, але не обмежуються, центрифугування, фільтрацію, екстракцію, сушку розпиленням, випарювання або осадження. Ізолюваний поліпептид може бути додатково очищений за допомогою численних методик, відомих в технології, що включають, але не обмежуються, хроматографію (наприклад, іонообмінну, афінну, гідрофобну хроматографію, хроматофокусування і хроматографію з виключенням за розміром часток), електрофоретичні методики (наприклад, препаративне ізоелектричне фокусування), диференційну розчинність (наприклад, осадження сульфатом амонію) або екстракцію (див., наприклад, Bollag et al. (1996) Protein Methods. 2<sup>nd</sup> Edition Wiley-Liss, NY; Walker (1996) The Protein Protocols Handbook Humana Press, NJ; Bollag et al. (1996) Protein Methods. 2<sup>nd</sup> Edition Wiley-Liss, NY; Walker (1996) The Protein Protocols Handbook Humana Press, NJ).

Також можуть бути застосовані безклітинні системи транскрипції/трансляції для вироблення поліпептидів з використанням ДНК або РНК за винаходом. Кілька таких систем серійно випускаються. Загальну інформацію щодо протоколів *in vitro* транскрипції і трансляції можна знайти у Tymms (1995) *In vitro* Transcription and Translation Protocols: Methods in Molecular Biology Volume 37, Garland Publishing, NY.

Субстрати і формати для рекомбінації послідовностей

Полінуклеотиди за винаходом можуть бути використані у якості субстратів для багатьох процедур вироблення різноманітності, наприклад, реакцій мутації, рекомбінації і рекурсивної рекомбінації, на додаток до їх використання в стандартних способах клонування, описаних, наприклад, у Ausubel, Berger and Sambrook, тобто, для вироблення додаткових GAT-полінуклеотидів і поліпептидів з бажаними властивостями. Численні протоколи вироблення різноманітності створені і описані у відповідній області. Ці методи можуть застосовуватися окремо і/або в комбінації для вироблення одного або більше варіантів полінуклеотиду або набору полінуклеотидів, а також варіантів закодованих протеїнів. Окремо і спільно, ці методи забезпечують ефективні, загальнодоступні способи генерування диверсифікованих полінуклеотидів і наборів полінуклеотидів (включаючи, наприклад, бібліотеки полінуклеотидів), корисних, наприклад, для конструювання або швидкої еволюції полінуклеотидів, протеїнів, шляхів, клітин і/або організмів з новими і/або покращеними характеристиками. Конкретний спосіб зміни послідовності може забезпечувати, наприклад, одиночні заміщення нуклеотидів, множинні заміщення нуклеотидів, а також інсерції або делеції областей нуклеїновокислотної послідовності.

Хоча у подальшому описі для ясності зроблені розмежування і класифікації, але фахівцям буде зрозуміло, що технології часто не є взаємно виключними. Дійсно, різні методи можуть використовуватися окремо або в комбінації, паралельно або послідовно, для отримання різноманітних варіантів послідовностей.

Результатом будь-якої з процедур вироблення різноманітності, описаних тут, може бути вироблення одного або більше полінуклеотидів, які можуть бути піддані селекції або скринінгу на полінуклеотиди, які кодують протеїни з бажаними властивостями або такі протеїни, що надають бажані властивості. Після диверсифікації одним або більше з описаних тут способів, або іншими, відомими фахівцям, будь-які полінуклеотиди, що виробляються, можуть бути селектовані на бажану активність або властивість, наприклад, змінену  $K_m$  до гліфосату, змінену  $K_m$  до ацетил-CoA, використання альтернативних кофакторів (наприклад, пропіоніл-CoA), підвищену  $k_{cat}$  і т.д. Це може включати ідентифікацію будь-якої активності, яка може бути виявлена, наприклад, в автоматизованому або автоматичному форматі будь-якою аналітичною методикою, відомою у відповідній області. Наприклад, гомологи GAT з підвищеною специфічною активністю можуть бути виявлені шляхом аналізу перетворення гліфосату в N-ацетилгліфосат, наприклад, мас-спектрометрією. В альтернативі, покращена здатність надавати стійкість до гліфосату може бути досліджена шляхом вирощування бактерій, трансформованих нуклеїновою кислотою за винаходом, на агарі, який містить концентрації гліфосату, що поступово підвищуються, або шляхом обприскування трансгенних рослин, включаючих нуклеїнову кислоту за винаходом, гліфосатом. Можуть бути оцінені різні споріднені (або навіть неспоріднені)

властивості, послідовно або паралельно, на розсуд дослідника. Додаткові подробиці стосовно рекомбінації і селекції на толерантність до гербіциду можна знайти, наприклад, в "DNA SHUFFLING TO PRODUCE HERBICIDE RESISTANT CROPS" (USSN 09/373,333), зареєстрований 12 серпня 1999.

Описи багатьох методик вироблення різноманітності, включаючи родовий шафлінг і способи вироблення модифікованих нуклеїновокислотних послідовностей, кодуючих численні домени ферментів, можна знайти в таких публікаціях і посиланнях, наведених там: Soong, N. et al. (2000) "Molecular breeding of viruses" *Nat Genet* 25(4):436-39; Stemmer, et al. (1999) "Molecular breeding of viruses for targeting and other clinical properties" *Tumor Targeting* 4:1-4; Ness et al. (1999) "DNA Shuffling of subgenomic sequences of subtilisin" *Nature Biotechnology* 17:893-896; Chang et al. (1999) "Evolution of a cytokine using DNA family shuffling" *Nature Biotechnology* 17:793-797; Minshull and Stemmer (1999) "Protein evolution by molecular breeding" *Current Opinion in Chemical Biology* 3:284-290; Christians et al. (1999) "Directed evolution of thymidine kinase for AZT phosphorylation using DNA family shuffling" *Nature Biotechnology* 17:259-264; Cramer et al. (1998) "DNA shuffling of a family of genes from diverse species accelerates directed evolution" *Nature* 391:288-291; Cramer et al. (1997) "Molecular evolution of an arsenate detoxification pathway by DNA shuffling", *Nature Biotechnology* 15:436-438; Zhang et al. (1997) "Directed evolution of an effective fucosidase from a galactosidase by DNA shuffling and screening" *Proc. Natl. Acad. Sci. USA* 94:4504-4509; Patten et al. (1997) "Applications of DNA Shuffling to Pharmaceuticals and Vaccines" *Current Opinion in Biotechnology* 8:724-733; Cramer et al. (1996) "Construction and evolution of antibody-phage libraries by DNA shuffling" *Nature Medicine* 2:100-103; Cramer et al. (1996) "Improved green fluorescent protein by molecular evolution using DNA shuffling" *Nature Biotechnology* 14:315-319; Gates et al. (1996) "Affinity selective isolation of ligands from peptide libraries through display on a lac repressor 'headpiece dimer'" *Journal of Molecular Biology* 255:373-386; Stemmer (1996) "Sexual PCR and Assembly PCR" In: *The Encyclopedia of Molecular Biology*. VCH Publishers, New York, pp.447-457; Cramer and Stemmer (1995) "Combinatorial multiple cassette mutagenesis creates all the permutations of mutant and wildtype cassettes" *BioTechniques* 18:194-195; Stemmer et al., (1995) "Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides" *Gene*, 164:49-53; Stemmer (1995) "The Evolution of Molecular Computation" *Science* 270:1510; Stemmer (1995) "Searching Sequence Space" *Bio/Technology* 13:549-553; Stemmer (1994) "Rapid evolution of a protein in vitro by DNA shuffling" *Nature* 370:389-391; Stemmer (1994) "DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution." *Proc. Natl. Acad. Sci. USA* 91:10747-10751.



Мутаційні способи вироблення різноманітності включають, наприклад, сайт-направлений мутагенез (Ling et al. (1997) "Approaches to DNA mutagenesis: an overview" *Anal Biochem.* 254(2): 157-178; Dale et al. (1996) "Oligonucleotide-directed random mutagenesis using the phosphorothioate method" *Methods Mol. Biol.* 57:369-374; Smith (1985) "In vitro mutagenesis" *Ann. Rev. Genet.* 19:423-462; Botstein & Shortle (1985) "Strategies and applications of in vitro mutagenesis" *Science* 229:1193-1201; Carter (1986) "Site-directed mutagenesis" *Biochem. J.* 237:1-7; Kunkel (1987) "The efficiency of oligonucleotide directed mutagenesis" in *Nucleic Acids & Molecular Biology* (Eckstein, F. and Lilley, D.M.J. eds., Springer Verlag, Berlin)); мутагенез з використанням матриць, що містять урацил (Kunkel (1985) "Rapid and efficient site-specific mutagenesis without phenotypic selection" *Proc. Natl. Acad. Sci. USA* 82:488-492; Kunkel et al. (1987) "Rapid and efficient site-specific mutagenesis without phenotypic selection" *Methods in Enzymol.* 154:367-382; Bass et al. (1988) "Mutant Trp repressors with new DNA-binding specificities" *Science* 242:240-245); олігонуклеотид-направлений мутагенез (Methods in *Enzymol.* 100:468-500 (1983); *Methods in Enzymol.* 154: 329-350 (1987); Zoller & Smith (1982) "Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any DNA fragment" *Nucleic Acids Res.* 10:6487-6500; Zoller & Smith (1983) "Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors" *Methods in Enzymol.* 100:468-500; Zoller & Smith (1987) "Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template" *Methods in Enzymol.* 154:329-350); фосфоротіоат-модифікований мутагенез ДНК (Taylor et al. (1985) "The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA" *Nucl. Acids Res.* 13: 8749-8764; Taylor et al. (1985) "The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA" *Nucl. Acids Res.* 13: 8765-8787 (1985); Nakamaye & Eckstein (1986) "Inhibition of restriction endonuclease Nci I cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis" *Nucl. Acids Res.* 14: 9679-9698; Sayers et al. (1988) "Y-T Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis" *Nucl. Acids Res.* 16:791-802; Sayers et al. (1988) "Strand specific cleavage of phosphorothioate-containing DNA by reaction with restriction endonucleases in the presence of ethidium bromide" *Nucl. Acids Res.* 16: 803-814); мутагенез з використанням дуплексної ДНК з пропуском (Kramer et al. (1984) "The gapped duplex DNA approach to oligonucleotide-directed mutation construction" *Nucl. Acids Res.* 12: 9441-9456; Kramer & Pritz (1987) *Methods in Enzymol.* "Oligonucleotide-directed construction of mutations via gapped duplex DNA" 154:350-367; Kramer et al. (1988) "Improved enzymatic in vitro reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations"

*Nucl. Acids Res.* 16: 7207; Fritz et al. (1988) "Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure without enzymatic reactions in vitro" *Nucl. Acids Res.* 16: 6987-6999).

Додаткові корисні способи включають репарацію точкових невідповідностей (Kramer et al. (1984) "Point Mismatch Repair" *Cell* 38:879-887), мутагенез з використанням репаративно-дефіцитних штамів-хазяїв (Carter et al. (1985) "Improved oligonucleotide site-directed mutagenesis using M13 vectors" *Nucl. Acids Res.* 13:4431-4443; and Carter (1987) "Improved oligonucleotide-directed mutagenesis using M13 vectors" *Methods in Enzymol.* 154: 382-403), делеційний мутагенез (Eghtedarzadeh & Henikoff (1986) "Use of oligonucleotides to generate large deletions" *Nucl. Acids Res.* 14: 5115), рестрикцію-селекцію і рестрикцію-селекцію і рестрикцію-очищення (Wells et al. (1986) "Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin" *Phil. Trans. R. Soc. Lond. A* 317:415-423), мутагенез шляхом тотального синтезу гена (Nambiar et al. (1984) "Total synthesis and cloning of a gene coding for the ribonuclease S protein" *Science* 223:1299-1301; Sakamar and Khorana (1988) "Total synthesis and expression of a gene for the a-subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin)" *Nucl. Acids Res.* 14:6361-6372; Wells et al. (1985) "Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites" *Gene* 34:315-323; Grundstrom et al. (1985) "Oligonucleotide-directed mutagenesis by microscale 'shot-gun' gene synthesis" *Nucl. Acids Res.* 13: 3305-3316), репарацію шляхом розриву подвійної спіралі (Mandecki (1986); Arnold (1993) "Protein engineering for unusual environments" *Current Opinion in Biotechnology* 4:450-455. "Oligonucleotide-directed double-strand break repair in plasmids of *Escherichia coli*: a method for site-specific mutagenesis" *Proc. Natl. Acad. Sci. USA*, 83:7177-7181). Додаткові подробиці стосовно багатьох з цих способів можна знайти в *Methods in Enzymology Volume 154*, де також описані корисні регулятори для усунення проблем при здійсненні різних способів мутагенезу.

Додаткові подробиці стосовно багатьох способів вироблення різноманітності можна знайти в таких патентах США, публікаціях РСТ і ЕПО: патент США №5,605,793, Stemmer (February 25, 1997), "Methods for In Vitro Recombination"; патент США №5,811,238, Stemmer et al. (September 22, 1998) "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination"; патент США №5,830,721, Stemmer et al. (November 3, 1998), "DNA Mutagenesis by Random Fragmentation and Reassembly"; патент США №5,834,252, Stemmer, et al. (November 10, 1998) "End-Complementary Polymerase Reaction"; патент США №5,837,458, Minshall, et al. (November 17, 1998), "Methods and Compositions for Cellular and Metabolic Engineering"; WO 95/22625, Stemmer and Cramer, "Mutagenesis by Random Fragmentation and Reassembly", WO 96/33207, Stemmer and Uppschütz "End Complementary Polymerase Chain Reaction"; WO 97/20078, Stemmer and Cramer

"Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination"; WO 97/35966, Minshull and Stemmer, "Methods and Compositions for Cellular and Metabolic Engineering"; WO 99/41402, Punnonen et al. "Targeting of Genetic Vaccine Vectors"; WO 99/41383, Punnonen et al. "Antigen Library Immunization"; WO 99/41369, Punnonen et al. "Genetic Vaccine Vector Engineering"; WO 99/41368, Punnonen et al. "Optimization of Immunomodulatory Properties of Genetic Vaccines"; EP 752008, Stemmer and Cramer, "DNA Mutagenesis by Random Fragmentation and Reassembly"; EP 0932670, Stemmer "Evolving Cellular DNA Uptake by Recursive Sequence Recombination"; WO 99/23107, Stemmer et al., "Modification of Virus Tropism and Host Range by Viral Genome Shuffling"; WO 99/21979, Apt et al., "Human Papillomavirus Vectors"; WO 98/31837, del Cardayre et al. "Evolution of Whole Cells and Organisms by Recursive Sequence Recombination"; WO 98/27230, Patten and Stemmer, "Methods and Compositions for Polypeptide Engineering"; WO 98/13487, Stemmer et al., "Methods for Optimization of Gene Therapy by Recursive Sequence Shuffling and Selection"; WO 00/00632, "Methods for Generating Highly Diverse Libraries"; WO 00/09679, "Methods for Obtaining in Vitro Recombined Polynucleotide Sequence Banks and Resulting Sequences"; WO 98/42832, Arnold et al., "Recombination of Polynucleotide Sequences Using Random or Defined Primers"; WO 99/29902, Arnold et al., "Method for Creating Polynucleotide and Polypeptide Sequences"; WO 98/41653, Vind, "An in Vitro Method for Construction of a DNA Library"; WO 98/41622, Borchert et al., "Method for Constructing a Library Using DNA Shuffling"; WO 98/42727, Pad and Zarling, "Sequence Alterations using Homologous Recombination"; WO 00/18906, Patten et al., "Shuffling of Codon-Altered Genes"; WO 00/04190, del Cardayre et al. "Evolution of Whole Cells and Organisms by Recursive Recombination"; WO 00/42561, Cramer et al., "Oligonucleotide Mediated Nucleic Acid Recombination"; WO 00/42559, Selifonov and Stemmer "Methods of Populating Data Structures for Use in Evolutionary Simulations"; WO 00/42560, Selifonov et al., "Methods for Making Character Strings, Polynucleotides & Polypeptides Having Desired Characteristics"; WO 01/23401, Welch et al., "Use of Codon-Variied Oligonucleotide Synthesis for Synthetic Shuffling"; PCT/US01/06775 "Single-Stranded Nucleic Acid Template-Mediated Recombination and Nucleic Acid Fragment Isolation", автор Affholter.

В деяких патентних заявках США описані додаткові подробиці стосовно багатьох способів вироблення різноманітності, включаючи "SHUFFLING OF CODON ALTERED GENES", Patten et al., зареєстрована 28 вересня 1999, (USSN 09/407,800); "EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE SEQUENCE RECOMBINATION", del Cardayre et al., зареєстрована 15 липня 1998 (USSN 09/166,188) і 15 липня 1999 (USSN 09/354,922); "OLIGONUCLEOTIDE MEDIATED NUCLEIC ACID

RECOMBINATION", Cramer et al., зареєстрована 28 вересня 1999 (USSN 09/408,392), і "OLIGONUCLEOTIDE MEDIATED NUCLEIC ACID RECOMBINATION", Cramer et al., зареєстрована 18 січня 2000 (PCT/US00/01203); "USE OF CODON-BASED OLIGONUCLEOTIDE SYNTHESIS FOR SYNTHETIC SHUFFLING", Welch et al., зареєстрована 28 січня 1999 (USSN 09/408,393); "METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES & POLYPEPTIDES HAVING DESIRED CHARACTERISTICS", Selifonov et al., зареєстрована 18 січня 2000, (PCT/US00/01202) і, наприклад, "METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES & POLYPEPTIDES HAVING DESIRED CHARACTERISTICS", Selifonov et al., зареєстрована 18 липня 2000 (USSN 09/618,579); "METHODS OF POPULATING DATA STRUCTURES FOR USE IN EVOLUTIONARY SIMULATIONS", Selifonov and Stemmer (PCT/US00/01138), зареєстрована 18 січня 2000; і "SINGLE-STRANDED NUCLEIC ACID TEMPLATE-MEDIATED RECOMBINATION AND NUCLEIC ACID FRAGMENT ISOLATION", Affholter (USSN 60/186F.482, зареєстрована 2 березня 2000).

У короткому викладі, кілька різних загальних класів способів модифікації послідовностей, такі як мутація, рекомбінація і т.д., придатні для цього винаходу і перелічені, наприклад, в наведених вище посиланнях. Тобто, зміни складових нуклеїновокислотних послідовностей у модифікованих конструкціях злиття генів, що виробляються, можуть виконуватися за будь-яким з описаних протоколів, або до з'єднання послідовностей, або після етапу з'єднання. Далі описані приклади декількох різних типів переважних форматів для вироблення різноманітності в контексті цього винаходу, включаючи, наприклад, деякі формати вироблення різноманітності на основі рекомбінації.

Нуклеїнові кислоти можуть бути рекомбіновані in vitro за допомогою будь-якої з методик, описаних вище у посиланнях, включаючи, наприклад, ферментативне розщеплення ДНК-азою нуклеїнових кислот, що підлягають рекомбінації, з наступним лігуванням і/або ПЛР-перезбиранням нуклеїнових кислот. Наприклад, може бути використаний статевий ПЛР-мутагенез, в якому після випадкової (або псевдовипадкової, або навіть не випадкової) фрагментації молекули ДНК проводиться рекомбінація (на основі подібності послідовностей) між молекулами ДНК з різними, але спорідненими ДНК-послідовностями, in vitro, з наступною фіксацією кросоверу шляхом ПЛР-розширення. Цей спосіб і багато варіантів способу описані в декількох з наведених вище посилань, наприклад, в Stemmer (1994) Proc. Natl. Acad. Sci. USA 91:10747-10751.

Подібним чином, нуклеїнові кислоти можуть бути рекурсивно рекомбіновані in vivo, наприклад, шляхом дозволу рекомбінації між нуклеїновими кислотами в клітинах. Багато таких форматів in vivo рекомбінації описані в наведених вище посиланнях. Такі формати можуть забезпечувати

безпосередню рекомбінацію між нуклеїновими кислотами, що представляють інтерес, або забезпечувати рекомбінацію між векторами, вірусами, плазмідами і т.д., які включають нуклеїнові кислоти, що представляють інтерес, а також інші формати. Деталі стосовно таких методик можна знайти в наведених вище посиланнях.

Також можуть бути використані способи рекомбінації цілого генома, в яких рекомбінуються цілі геноми клітин або інших організмів, можливо, включаючи спайкінг геномних рекомбінаційних сумішей з бажаними компонентами бібліотеки (наприклад, генами, відповідними до шляхів за винаходом). Ці способи мають багато застосувань, включаючи ті, в яких ідентичність цільового гена невідома. Подробиці стосовно цих способів можна знайти, наприклад, в WO 98/31837, del Cardayre et al. "Evolution of Whole Cells and Organisms by Recursive Sequence Recombination"; і в PCT/US99/15972, del Cardayre et al., також під назвою "Evolution of Whole Cells and Organisms by Recursive Sequence Recombination." Таким чином, будь-які з цих способів і технологій рекомбінації, рекурсивної рекомбінації і рекомбінації цілого генома, окремо або в комбінації, можуть бути застосовані для вироблення модифікованих нуклеїнових кислотних послідовностей і/або модифікованих конструкцій злиття генів за винаходом.

Також можуть бути застосовані синтетичні способи рекомбінації, в яких олігонуклеотиди, які відповідають об'єктам, що представляють інтерес, синтезують і перескладають за допомогою реакцій ПЛР або лігування. Це олігонуклеотиди, які відповідають більше ніж одній батьківській нуклеїновій кислоті, за рахунок чого виробляються нові рекомбіновані нуклеїнові кислоти. Олігонуклеотиди можуть вироблятися стандартними способами нуклеотидного приєднання, або можуть вироблятися, наприклад, способами три нуклеотидного синтезу. Подробиці стосовно цих способів можна знайти в наведених вище посиланнях, наприклад, в WO 00/42561, Crameri et al., "Oligonucleotide Mediated Nucleic Acid Recombination"; WO 01/23401, Welch et al., "Use of Codon-Varied Oligonucleotide Synthesis for Synthetic Shuffling"; WO 00/42560, Selifonov et al., "Methods for Making Character Strings, Polynucleotides and Polypeptides Having Desired Characteristics"; WO 00/42559, Selifonov and Stemmer "Methods of Populating Data Structures for Use in Evolutionary Simulations".

Можуть бути здійснені способи рекомбінації *in silico*, в яких використовуються комп'ютерні генетичні алгоритми для рекомбінації ланцюгів послідовностей, що відповідають гомологічним (або навіть негомологічним) нуклеїновим кислотам. Результуючі рекомбіновані ланцюги послідовностей можуть бути перетворені в нуклеїнові кислоти шляхом синтезу нуклеїнових кислот, які відповідають рекомбінованим послідовностям, наприклад, у взаємодії з олігонуклеотидним синтезом/технологією перезбирання генів. Цей підхід може генерувати випадкові, частково випадкові або заплановані варіанти. Багато подробиць стосовно рекомбінації *in silico*, включаючи використання генетичних алгоритмів, генетичних опера-

торів і т.п., в комп'ютерних системах, у поєднанні з генерацією відповідних нуклеїнових кислот (і/або протеїнів), а також комбінацій сконструйованих нуклеїнових кислот і/або протеїнів (наприклад, на основі селекції сайта кросовера), а також способів запланованої, псевдовипадкової або випадкової рекомбінації описані в WO 00/42560, Selifonov et al., "Methods for Making Character Strings, Polynucleotides and Polypeptides Having Desired Characteristics" і WO 00/42559, Selifonov and Stemmer "Methods of Populating Data Structures for Use in Evolutionary Simulations". Багато подробиць стосовно рекомбінації *in silico* можна знайти в цих заявках. Ця методологія загалом придатна у цьому винаході для рекомбінації нуклеїнових кислотних послідовностей і/або конструкцій злиття генів, кодуєних протеїнів, що приймають участь в різних метаболічних шляхах (наприклад, таких як каротиноїдні біосинтетичні шляхи, ектоїнові біосинтетичні шляхи, полігідроксисилканоатні біосинтетичні шляхи, біосинтетичні шляхи ароматичних полікетидів і т.п.) *in silico* і/або для вироблення відповідних нуклеїнових кислот або протеїнів.

Багато способів отримання природної різноманітності, наприклад, шляхом гібридизації різноманітних нуклеїнових кислот або фрагментів нуклеїнових кислот з односпіральними матрицями, з наступною полімеризацією і/або лігуванням для регенерації послідовностей повної довжини, можливо, з наступною деградацією матриць і регенерацією результуючих модифікованих нуклеїнових кислот, можуть бути використані подібним чином. В одному зі способів, що використовують односпіральну матрицю, фрагментарну популяцію, отриману з геномної бібліотеки (бібліотек) відпаляють з частковою, або, частіше, майже повною ссДНК або РНК, що відповідає протилежній спіралі. Потім модуль з комплексних химерних генів з цієї популяції опосереднюють нуклеаза-основним видаленням негібридизуючих кінців фрагмента, полімеризацією для заповнення пропусків між такими фрагментами і, нарешті, односпіральним лігуванням. Батьківська нуклеотидна спіраль може бути видалена шляхом ферментації (наприклад, якщо вона являє собою РНК або таку, що містить урацил), магнітного розділення в денатуруючих умовах (якщо вона мічена відповідним чином для такого розділення) та іншими доступними способами розділення/очищення. В альтернативі, батьківську спіраль можна очищати спільно з химерними спіралями і видаляти під час наступних етапів скринінгу і подальшої обробки. Додаткові подробиці стосовно цього підходу можна знайти, наприклад, в "Single-Stranded Nucleic Acid Template-Mediated Recombination and Nucleic Acid Fragment Isolation", Affholter, PCT/US01/06775.

В іншому підході односпіральні молекули перетворюють на двоспіральні ДНК (дсДНК), а молекули дсДНК зв'язують з твердою основою шляхом ліганд-опосередненого зв'язування. Після відділення незв'язаної ДНК селектовані молекули ДНК звільнюють від основи і вводять в придатну клітину-хазяїна для генерації бібліотеки збагачених послідовностей, які гібридизуються з зондом.

Бібліотека, вироблена таким чином, утворює бажаний субстрат для подальшої диверсифікації з використанням будь-якої з методик, описаних тут.

Будь-який з вищеописаних загальних форматів рекомбінації може бути застосований у повторній схемі (наприклад, один або більше циклів мутації/рекомбінації або інших способів вироблення різноманітності, можливо, з наступним використанням одного або більше способів селекції) для вироблення більш різноманітного набору рекомбінантних нуклеїнових кислот.

Також був запропонований мутагенез, що використовує способи термінування полінуклеотидного ланцюга (див., наприклад, патент США №5,965,408, "Method of DNA reassembly by interrupting synthesis", Short, і наведені там посилання), який може бути використаний в цьому винаході. В цьому підході двоспіральні ДНК, що відповідають одному або більше генам, які поділяють області подібності послідовностей, поєднують і денатурують у присутності або у відсутності праймерів, специфічних для цього гена. Потім односпіральні полінуклеотиди відпалюють і інкубують у присутності полімерази і реагента для термінування ланцюга (ультрафіолетове, гамма- або рентгенівське опромінювання; етидія бромід та інші інтеркалятори; ДНК-зв'язуючі протеїни, такі як односпіральні зв'язуючі протеїни, фактори активації транскрипції, або гістони; поліциклічні ароматичні вуглеводні; тривалентний хром або сіль тривалентного хрому; або скорочена полімеризація, опосередкована швидкими термоциклами; і т.п.), що приводить до вироблення часткових дуплексних молекул. Часткові дуплексні молекули, наприклад, що містять частково подовжені ланцюги, денатурують і відпалюють в наступних раундах реплікації або часткової реплікації, що приводить до утворення полінуклеотидів, які поділяють різні ступені подібності послідовностей і які є диверсифікованими відносно стартової популяції молекул ДНК. Продукти, або часткові пули продуктів, можуть бути ампліфіковані в одну або більше стадій. Полінуклеотиди, вироблені способом термінування ланцюга, таким як описаний вище, є зручними субстратами для будь-якого іншого описаного формату рекомбінації.

Різнноманітність також може бути вироблена в нуклеїнових кислотах або популяціях нуклеїнових кислот з використанням рекомбінаційної методики, званої "інкрементне усичення для створення гібридних ферментів" ("ITCHY"), що описана в Ostermeier et al. (1999) "A combinatorial approach to hybrid enzymes independent of DNA homology" *Nature Biotech* 17:1205. Цей підхід може бути використаний для генерування початкової бібліотеки варіантів, яка, наприклад, може служити субстратом для одного або більше способів рекомбінації *in vitro* або *in vivo*. Див. також Ostermeier et al. (1999) "Combinatorial Protein Engineering by Incremental Truncation", *Proc. Natl. Acad. Sci. USA*, 96: 3562-67; Ostermeier et al. (1999) "Incremental Truncation as a Strategy in the Engineering of Novel Biocatalysts", *Biological and Medicinal Chemistry*, 7: 2139-44.

Способи мутації, які приводять до змін окремих нуклеотидів або груп суміжних або несуміжних нуклеотидів, можуть бути успішно використані для введення нуклеотидної різноманітності в нуклеїновокіслотні послідовності і/або конструкції злиття генів за винаходом. Багато способів мутагенезу можна знайти в наведених вище посиланнях; далі можна знайти додаткові подробиці стосовно способів мутагенезу, які також можуть бути використані у цьому винаході.

Наприклад, для генерації варіантів нуклеїнових кислот може бути використана схильна до помилок ПЛР. За цією технологією, ПЛР виконують в умовах, при яких точність відтворення ДНК-полімерази низька, в результаті чого отримують високу швидкість точкових мутацій по всій довжині продукту ПЛР. Приклади таких технологій можна знайти в наведених вище посиланнях і, наприклад, в Leung et al. (1989) *Technique* 1:11-15 and Caldwel et al. (1992) *PCR Methods Applic.* 2:28-33. Подібним чином, може бути використана складальна ПЛР, спосіб, що включає збирання продукту ПЛР із суміші малих фрагментів ДНК. Велика кількість різних реакцій ПЛР може відбуватися паралельно в одній реакційній суміші, де продукти однієї реакції примують продукти іншої реакції.

Олігонуклеотид-направлений мутагенез може бути використаний для введення сайт-специфічних мутацій в нуклеїновокіслотну послідовність, що представляє інтерес. Приклади таких технологій можна знайти в наведених вище посиланнях і, наприклад, в Reidhaar-Olson et al. (1988) *Science*. 241:53-57. Подібним чином, може бути використаний касетний мутагенез, спосіб, що замінює малу область двоспіральної молекули ДНК синтетичною олігонуклеотидною касетою, яка відрізняється від нативної послідовності. Цей олігонуклеотид може містити, наприклад, повністю і/або частково рандомізовану нативну послідовність (послідовності).

Рекурсивний ансамбль-мутагенез являє собою спосіб, в якому алгоритм протеїнового мутагенезу використовується для вироблення різноманітних популяцій фенотипово споріднених мутантів, члени яких відрізняються амінокислотною послідовністю. Цей спосіб використовує механізм зворотного зв'язку для контролю послідовних раундів комбінаторного касетного мутагенезу. Приклади цього підходу можна знайти в Arkin & Youvan (1992) *Proc. Natl. Acad. Sci. USA* 89:7811-7815.

Експоненціальний ансамбль-мутагенез може бути використаний для вироблення комбінаторних бібліотек з високим процентом унікальних і функціональних мутантів. Малі групи залишків у послідовності, що представляє інтерес, рандомізують паралельно з ідентифікацією, у кожному зміненому положенні, амінокислот, які ведуть до функціональних протеїнів. Приклади таких методик можна знайти в Delegrave & Youvan (1993) *Biotechnology Research* 11:1548-1552.

*In vivo* мутагенез може бути використаний для генерації випадкових мутацій в будь-якій клонованій ДНК, що представляє інтерес, шляхом розмноження ДНК, наприклад, в штамі *E.coli*,

який несе мутації в одному або більше шляхів репарації ДНК. Ці "мутаторні" штами мають вищу швидкість випадкових мутацій, ніж предок дикого типу. Розмноження ДНК в одному з цих штамів у кінцевому рахунку буде генерувати випадкові мутації всередині ДНК. Такі методики описані в посиланнях, наведених вище.

Інші методики для введення різноманітності в геном, наприклад, бактеріальний, грибовий, тваринний або рослинний геном, можуть бути використані у поєднанні з описаними вище і/або наведеними в посиланнях способами. Наприклад, на додаток до способів, описаних вище, були запропоновані технології, які забезпечують вироблення мультимерів нуклеїнових кислот, придатних для трансформації у різні біологічні види (див., наприклад, Schellenberger, патент США №5,756,316 і наведені вище посилання). Трансформація придатного хазяїна такими мультимерами, що складаються з генів, які є дивергентними один до одного (наприклад, отриманими з природної різноманітності або шляхом застосування сайт-направленого мутагенезу, схильної до помилок ПЛР, пасажу через мутагенні бактеріальні штами і т.п.), забезпечує джерело різноманітності нуклеїнових кислот для диверсифікації ДНК, наприклад, шляхом рекомбінації *in vivo*, як описано вище.

В альтернативі, безліч мономерних полінуклеотидів, що поділяють області часткової подібності послідовностей, можуть бути трансформовані у види-хазяї і рекомбіновані *in vivo* клітиною-хазяїном. Наступні цикли ділення клітин можуть бути використані для генерації бібліотек, члени яких включають одиночну, гомогенну популяцію або пул мономерних полінуклеотидів. В альтернативі, мономерна нуклеїнова кислота може бути регенована стандартними технологіями, наприклад, ПЛР і/або клонуванням, і рекомбінована в будь-якому з форматів рекомбінації, описані вище.

Описані способи генерації багатовидових бібліотек експресії (на додаток до посилань, наведених вище, див., наприклад, Peterson et al. (1998) патент США №5,783,431 "METHODS FOR GENERATING AND SCREENING NOVEL METABOLIC PATHWAYS", і Thompson, et al. (1998) патент США №5,824,485 "METHODS FOR GENERATING AND SCREENING NOVEL METABOLIC PATHWAYS") і запропоновано їх використання для ідентифікації активностей протеїну, що представляє інтерес (на додаток до посилань, наведених вище, див. Short (1999) патент США №5,958,672 "PROTEIN ACTIVITY SCREENING OF CLONES HAVING DNA FROM UNCULTIVATED MICROORGANISMS"). Багатовидові бібліотеки експресії включають, загалом, бібліотеки, що містять кДНК або геномні послідовності з численних видів або штамів, оперативно зв'язані з відповідними регуляторними послідовностями, в касеті експресії. кДНК і/або геномні послідовності можуть бути ліговані випадковим чином для додаткового посилення різноманітності. Вектор може бути човниковим вектором, придатним для трансформації і експресії більш ніж в одному виді організму-хазяїна, наприклад, бакте-

ріальному виді, еукаріотичних клітинах. В деяких випадках бібліотеку зміщують шляхом попередньої селекції послідовностей, які кодують протеїн, що представляє інтерес, або які гібридизуються з нуклеїновою кислотою, що представляє інтерес. Будь-які такі бібліотеки можуть бути створені у якості субстратів для будь-якого з способів, описаних тут.

Описані вище методики у значній мірі були направлені на підвищення різноманітності нуклеїнових кислот і/або закодованого протеїну. Проте в багатьох випадках не вся різноманітність корисна, наприклад, функціональна, а тільки приймає участь в підвищенні фону варіантів, що повинні бути піддані скринінгу або селекції для ідентифікації кількох переважних варіантів. В деяких прикладних задачах бажано піддати попередній селекції або попередньому скринінгу бібліотеки (наприклад, ампліфіковану бібліотеку, геномну бібліотеку, бібліотеку кДНК, нормалізовану бібліотеку і т.д.) або інші субстратні нуклеїнові кислоти перед диверсифікацією, наприклад, шляхом мутагенезу на основі рекомбінації, або іншим способом змістити субстрати в сторону нуклеїнових кислот, які кодують функціональні продукти. Наприклад, у випадку конструювання антитіл можливо змістити процес вироблення різноманітності у сторону антитіл з функціональними сайтами зв'язування антигена шляхом використання рекомбінаційних подій *in vivo* до застосування будь-якого з описаних способів. Наприклад, рекомбіновані CDRs (області, що визначають комплементарність), отримані з В-клітинних кДНК-бібліотек, можуть бути ампліфіковані і зібрані в каркасні області (наприклад, Jirholt et al. (1998) "Exploiting sequence space: shuffling *in vivo* formed complementarity determining regions into a master framework" *Gene* 215: 471) перед диверсифікацією будь-яким з методів, описаних тут.

Бібліотеки можуть бути зміщені в сторону нуклеїнових кислот, які кодують протеїни з бажаними ферментативними активностями. Наприклад, після ідентифікації клона з бібліотеки, який показує специфічну активність, клон може бути мутагенізований з використанням будь-якого відомого способу введення ДНК-змін. Потім бібліотеку, яка включає мутагенізовані гомологи, піддають скринінгу на бажану активність, яка може бути тою самою або відрізнятися від початково визначеної активності. Приклад такої методики запропонований у Short (1999) патент США №5,958,672 "PRODUCTION OF ENZYMES HAVING DESIRED ACTIVITIES BY MUTAGENESIS". Бажані активності можуть бути ідентифіковані будь-яким способом, відомим в біотехнології. Наприклад, у WO 99/10539 запропоновано піддавати скринінгу генні бібліотеки комбінуючими екстрактами з генної бібліотеки, що містить компоненти, отримані з метаболічно багатих клітин, і ідентифікувати комбінації, які показують бажану активність. Також було запропоновано (наприклад, у WO 98/58085) ідентифікувати клони з бажаною активністю шляхом вставлення біоактивних субстратів у зразки бібліотеки і виявлення біоактивної флуоресценції, відповідаючої продукту з бажаною активністю, з використанням флуоресцентного аналізатора,

наприклад, приладу проточної цитометрії, CCD, флуорометру або спектрофотометру.

Бібліотеки можуть бути також зміщені в сторону нуклеїнових кислот, що мають визначені характеристики, наприклад, гібридизацією з вибраним нуклеїновокислотним зондом. Наприклад, в заявці WO 99/10539 запропонована ідентифікація полінуклеотидів, кодуючих бажану активність (наприклад, ферментативну активність, включаючи активність ліпази, естерази, протеази, глікозидази, глікозилтрансферази, фосфатази, кінази, оксигенази, пероксидази, гідролази, гідратази, нітрилази, трансамінази, амідази або ацилази), з різних послідовностей геномних ДНК таким чином. Односпіральні молекули ДНК з популяції геномної ДНК гібридизують з ліганд-кон'югованим зондом. Геномна ДНК може бути отримана з культивованого або некультивованого мікроорганізму або із зразка навколишнього середовища. В альтернативі геномна ДНК може бути отримана з багатоклітинного організму або з отриманої з нього тканини. Синтез другої спіралі може бути проведений безпосередньо від гібридизаційного зонда, що використовується для захоплення, з попереднім звільненням від середовища захоплення або без звільнення, або за допомогою інших численних методик, відомих у відповідній технології. В альтернативі, ізольована односпіральна популяція геномної ДНК може бути фрагментована без додаткового клонування і безпосередньо застосована, наприклад, у рекомбінаційному підході, який використовує односпіральні матриці, як описано вище.

"Нестохастичні" способи вироблення нуклеїнових кислот і поліпептидів запропоновані у Short "Non-Stochastic Generation of Genetic Vaccines and Enzymes" WO 00/46344. Ці способи, включаючи запропоновані способи нестохастичного полінуклеотидного перезбирання і мутагенезу сайт-направленого, також можуть бути використані у цьому винаході. Також описаний випадковий або напіввипадковий мутагенез, що використовує домішкові або вироджені олігонуклеотиди, наприклад, у Arkin and Youvan (1992) "Optimizing nucleotide mixtures to encode specific subsets of amino acids for semi-random mutagenesis" *Biotechnology* 10:297-300; Reidhaar-Olson et al. (1991) "Random mutagenesis of protein sequences using oligonucleotide cassettes" *Methods Enzymol.* 208:564-86; Lim and Sauer (1991) "The role of internal packing interactions in determining the structure and stability of a protein" *J. Mol. Biol.* 219:359-76; Breyer and Sauer (1989) "Mutational analysis of the fine specificity of binding of monoclonal antibody 51F to lambda repressor" *J. Biol. Chem.* 264:13355-60; і "Walk-Through Mutagenesis" (Crea, R; патенти США 5,830,685 і 5,798,208, а також патент EP 0527809 B1).

Зрозуміло, що будь-яка з описаних вище технологій, придатних для збагачення бібліотеки перед диверсифікацією, також можуть бути використані для скринінгу продуктів, або бібліотек продуктів, вироблених способами вироблення різноманітності. Будь-який з описаних вище способів може бути застосований рекурсивно або у

комбінації для зміни нуклеїнових кислот, наприклад, GAT-кодуючих полінуклеотидів.

Комплекти для мутагенезу, конструювання бібліотек та інших способів вироблення різноманітності також випускаються серійно. Наприклад, комплекти випускаються компаніями Stratagene (наприклад, комплект для сайт-направленого мутагенезу QuickChange™; комплект для двоспирального сайт-направленого мутагенезу Chameleon™), Bio/Can Scientific, Bio-Rad (наприклад, при застосуванні способу Kunkel, описаного вище), Boehringer Mannheim Corp., Clontech Laboratories, DNA Technologies, Epicentre Technologies (наприклад, комплект 5 prime 3 prime); Genpak Inc, Lemargo Inc, Life Technologies (Gibco BRL), New England Biolabs, Pharmacia Biotech, Promega Corp., Quantum Biotechnologies, Amersham International pic (наприклад, при застосуванні способу Eckstein, описаного вище), і Anglian Biotechnology Ltd (наприклад, при застосуванні способу Carter/Winter, описаного вище).

Наведені вище посилання надають багато мутаційних форматів, включаючи рекомбінацію, рекурсивну рекомбінацію, рекурсивну мутацію і їх комбінації, або рекомбінацію з іншими формами мутагенезу, а також багато модифікацій цих форматів. Незалежно від формату вироблення різноманітності, що використовується, нуклеїнові кислоти за винаходом можуть бути рекомбіновані (одна з одною або із спорідненими (або навіть неспорідненими) послідовностями) з отриманням різноманітного набору рекомбінаційних нуклеїнових кислот для застосування в конструкціях злиття генів і модифікованих конструкціях злиття генів за винаходом, включаючи, наприклад, набори гомологічних нуклеїнових кислот, а також відповідні поліпептиди.

Багато з методик вироблення модифікованих поліпептидів забезпечують вироблення великої кількості різноманітних варіантів батьківської послідовності або послідовностей. В деяких переважних варіантах здійснення винаходу технологія модифікації (наприклад, деякі форми шафлінгу) використовується для вироблення бібліотеки варіантів, яку потім піддають скринінгу на полінуклеотид або пул модифікованих полінуклеотидів, кодуючих деякий бажаний функціональний атрибут, наприклад, покращену GAT-активність. Приклади ферментативних активностей, що можуть бути виявлені скринінгом, включають каталітичні швидкості (звичайно визначаються на основі кінетичних констант, таких як  $k_{кат}$  і  $K_M$ ), специфічність субстрату і схильність до активації або інгібування субстратом, продуктом або іншими молекулами (наприклад, інгібіторами або активаторами).

Один з прикладів селекції на бажану ферментативну активність включає вирощування клітин-хазяїв в умовах, що інгібують ріст і/або виживання клітин, які недостатньо експресують ферментативну активність, що представляє інтерес, наприклад, GAT-активність. Застосування такого способу селекції може виключити з розгляду всі модифіковані полінуклеотиди, крім тих, які кодують бажану ферментативну активність. Наприклад, в деяких варіантах здійснення винаходу

клітини-хазяї підтримуються в умовах, що інгібують ріст або виживання клітин у відсутність достатніх рівнів GAT, наприклад, при концентрації гліфосату, що є летальною або інгібує ріст дикої рослини того ж різновиду, яка не експресує GAT-полінуклеотид. В цих умовах лише клітини-хазяї, які вміщують модифіковані нуклеїнові кислоти, кодує ферментативну активність або активності, що здатні каталізувати вироблення достатніх рівнів продукту, будуть виживати і рости. Деякі варіанти здійснення винаходу використовують багаторазові цикли скринінгу при концентраціях гліфосату або аналога гліфосату, що поступово підвищуються.

В деяких варіантах здійснення винаходу використовується мас-спектрометрія для виявлення ацетилювання гліфосату або аналога або метаболіта гліфосату. Використання мас-спектрометрії детальніше описано в Прикладах, наведених нижче.

Для зручності і високої продуктивності часто буває бажаний скринінг/селекція на бажані модифіковані нуклеїнові кислоти в мікроорганізмі, наприклад, бактерії, такої як *E.coli*. З іншого боку, у деяких випадках може бути бажаний скринінг в рослинних клітинах або рослинах, якщо кінцевою ціллю є вироблення модифікованих нуклеїнових кислот для експресії в рослинній системі.

В деяких варіантах здійснення винаходу продуктивність підвищують шляхом скринінгу пулів клітин-хазяїв, експресуючих різні модифіковані нуклеїнові кислоти, окремо або у вигляді частини конструкції злиття генів. Будь-які пули, що показують значну активність, можуть бути піддані деконволюції для ідентифікації одиночних клонів, експресуючих бажану активність.

Досвідченому фахівцеві буде зрозуміло, що відповідний спосіб аналізу, скринінгу або селекції буде залежати від бажаного організму-хазяїна і т.д. Як правило, сприятливо застосування такого аналізу, який може бути здійснений у високопродуктивному форматі.

У високопродуктивних аналізах можливо проведення скринінгу до кількох тисяч різних варіантів за один день. Наприклад, кожна а мікротитрувальної пластини може бути використана для проведення окремого аналізу або, якщо повинні бути досліджені ефекти концентрації або часу інкубації, кожні 5-10 чарунок можуть бути призначені для дослідження одного варіанту.

На додаток до рідинних підходів можливо, як зазначено вище, просто вирощувати клітини на пластинках з середовищем культури для селекції на бажану ферментативну або метаболічну функцію. Цей підхід пропонує простий і високопродуктивний спосіб скринінгу.

Також було розроблено декілька добре відомих роботизованих систем для рідкофазних хімічних методик, корисних в аналітичних системах. Ці системи включають автоматизовані робочі станції, такі як автоматизовані апарати синтезу, розроблені Takeda Chemical Industries, LTD. (Osaka, Japan), і багато роботизованих систем, що використовують маніпулятори (Zymate II, Zymark Corporation, Hopkinton, MA.; Orca, Hewlett-Packard, Palo Alto, CA), які імітують ручні операції

синтезу, що виконуються людиною. Будь-які з описаних приладів придатні для застосування у цьому винаході. Природа і реалізація модифікацій цих приладів (якщо потрібно), для того щоб вони могли працювати, як описано тут, застосовно до інтегрованої системи, будуть зрозумілі фахівцям у відповідній області техніки.

Високопродуктивні системи скринінгу серійно випускаються (див., наприклад, Zymark Corporation, Hopkinton, MA; Air Technical Industries, Mentor, OH; Beckman Instruments, Inc. Fullerton, CA; Precision Systems, Inc., Natick, MA, і т.д.). Ці системи звичайно автоматизують усі процедури, включаючи піпетування всіх зразків і реагентів, дозування рідин, регулювання часу інкубації і визначення кінцевих даних мікропластини за допомогою детектора (детекторів), відповідного до конкретного аналізу. Ці системи, що конфігуруються, забезпечують високу продуктивність і швидкий запуск, а також високий ступінь гнучкості і налаштованості для різних задач.

Виробники таких систем надають докладні протоколи для різних високопродуктивних приладів. Так, наприклад, Zymark Corp. надає технічні бюлетені, що описують системи скринінгу для виявлення модуляції транскрипції генів, зв'язування лігандів і т.п. Також розроблені мікрорідинні підходи до маніпулювання реагентами, наприклад, Caliper Technologies (Mountain View, CA).

Оптичні зображення, що спостерігаються (і, можливо, записуються) камерою або іншим записуючим пристроєм (наприклад, фотодіодом і пристроєм зберігання даних), можуть додатково оброблятися у будь-якому з описаних тут варіантів здійснення винаходу, наприклад, шляхом оцифрування зображення і/або зберігання і аналізу зображення на комп'ютері. Багато периферійного обладнання і програмного забезпечення серійно випускається для оцифрування, зберігання і аналізу оцифрованих відео- або оцифрованих оптичних зображень, наприклад, з використанням PC (машин на базі Intel x86 або Pentium-сумісних DOS™, OS2™, WINDOWS™, WINDOWS NT™ або WINDOWS 95™), комп'ютерів на базі MACINTOSH™ або UNIX (наприклад, робоча станція SUN™).

Одна з традиційних систем переносить світло з аналітичного приладу на охолоджувану камеру-прилад с зарядовим зв'язком (CCD), що широко використовується у техніці. CCD-камера включає масив елементів зображення (пікселів). Світло від зразка створює зображення на CCD. Вибирають окремі піксели, що відповідають певним областям зразка (наприклад, окремим сайтам гібридизації в масиві біологічних полімерів), з отриманням даних світлової інтенсивності для кожного положення. Множину пікселів оброблюють паралельно з підвищенням швидкості. Апарат і способи за винаходом можуть бути легко застосовані для візуалізації будь-якого зразка, наприклад, за допомогою флуоресцентної технології або мікроскопії за методом темного поля.

Інші полінуклеотидні композиції

Винахід також стосується композицій, що включають два або більше полінуклеотидів за винаходом (наприклад, у якості субстратів для

рекомбінації). Композиції можуть включати бібліотеку рекомбінантних нуклеїнових кислот, при цьому бібліотека містить щонайменше 2, 3, 5, 10, 20 або 50 або більше полінуклеотидів. Полінуклеотиди можуть бути клоновані в вектори експресії з утворенням бібліотек експресії.

Винахід також стосується композицій, отриманих шляхом ферментації одного або більше полінуклеотидів за винаходом рестрикційною ендонуклеазою, РНК-азою або ДНК-азою (наприклад, як виконується в певних рекомбінаційних форматах, описаних вище), і композицій, отриманих шляхом фрагментування або розрізання одного або більше полінуклеотидів за винаходом механічними засобами (такими як ультразвукова руйнація, завихрення і т.п.), які також можуть бути застосовані для створення субстратів для рекомбінації способами, описаними вище. Подібним чином, композиції, що містять набори олігонуклеотидів, що відповідають більше ніж одній нуклеїновій кислоті за винаходом, є корисними у якості рекомбінаційних субстратів і є рисою цього винаходу. Для зручності ці фрагментовані, розрізані або олігонуклеотид-синтезовані суміші називаються наборами фрагментованих нуклеїнових кислот. Винахід також стосується композицій, отриманих шляхом інкубації одного або більше наборів фрагментованих нуклеїнових кислот у присутності рибонуклеотид- або дезоксирибонуклеотидтрифосфатів і полімерази нуклеїнових кислот. Ця результуюча композиція утворює рекомбінаційну суміш для багатьох рекомбінаційних форматів, зазначених вище. Полімеразою нуклеїнових кислот може бути РНК-полімераза, ДНК-полімераза або РНК-направлена ДНК-полімераза (наприклад, "зворотна транскриптаза"); полімеразою може бути наприклад, термостабільна ДНК-полімераза (така як VENT, TAQ і т.п.).

#### Інтегровані системи

Цей винахід стосується комп'ютерів, комп'ютерних носіїв інформації і інтегрованих систем, що включають ланцюги символів, які містять інформацію про послідовності поліпептидів і нуклеїнових кислот, описаних тут, включаючи, наприклад, послідовності, наведені у переліку, і їх різноманітні німі заміщення і консервативні заміщення.

Наприклад, різні способи і генетичні алгоритми (GAs), відомі у біотехнології, можуть бути використані для виявлення гомологічності або подібності між різними ланцюгами символів, або можуть бути використані для виконання інших бажаних функцій, наприклад, для контролю вихідних файлів, створення базису для здійснення презентацій інформації, включаючи послідовності і т.п. Приклади включають алгоритм BLAST, описаний вище.

Таким чином, різні типи гомологічності або подібності різної точності і довжини можуть бути виявлені і впізнані в інтегрованих системах, описаних тут. Наприклад, багато способів визначення гомологічності було розроблено для порівняльного аналізу послідовностей біополімерів, для перевірки орфографії при обробці текстів і для пошуку даних в різних базах даних. Маючи розуміння двоспиральних парних комплементарних

взаємодій між 4 принциповими нуклеотидними основами в природних полінуклеотидах, можна також використовувати моделі, що імітують відповідні комплементарних гомологічних полінуклеотидних ланцюгів, у якості фундаменту вирівнювання послідовностей або інших операцій, що звичайно виконуються над ланцюгами символів, які відповідають описаним тут послідовностям (наприклад, маніпуляції при обробці текстів, конструювання фігур, що містять ланцюги символів послідовностей або підпослідовностей, вихідні таблиці і т.д.). Прикладом пакета програмного забезпечення з GAs для розрахунку подібності послідовностей є BLAST, який може бути адаптований для цього винаходу шляхом вводу ланцюга символів, що відповідає послідовностям, описаним тут.

Подібним чином, стандартні настільні системи, такі як програмне забезпечення обробки текстів (наприклад, Microsoft Word™ або Corel WordPerfect™) і програмне забезпечення баз даних (наприклад, електронні таблиці, такі як Microsoft Excel™, Corel Quattro Pro™, або програми баз даних, такі як Access™ або Paradox™) можуть бути адаптовані для цього винаходу шляхом вводу ланцюга символів, що відповідає GAT-гомологам за винаходом (нуклеїновим кислотам або протеїнам, або тим і іншим). Наприклад, інтегровані системи можуть включати згадане вище програмне забезпечення, яке використовує відповідну інформацію стосовно ланцюга символів, наприклад, у поєднанні з інтерфейсом користувача (наприклад, графічним інтерфейсом користувача в стандартній операційній системі, такий як Windows, Macintosh або LINUX) для маніпулювання ланцюгом символів. Як зазначено вище, спеціалізовані програми вирівнювання, такі як BLAST, також можуть бути вбудовані в системи за винаходом для вирівнювання нуклеїнових кислот або протеїнів (або відповідних ланцюгів символів).

Інтегровані системи для аналізу в цьому винаході звичайно включають цифровий комп'ютер з програмним забезпеченням GA для вирівнювання послідовностей, а також набори даних, що вводяться в систему програмного забезпечення, які включають будь-яку з послідовностей, описаних тут. Комп'ютер може являти собою, наприклад, PC (машину на базі Intel x86 або Pentium-сумісних DOS™, OS2™, WINDOWS™, WINDOWS NT™ або WINDOWS 95™, WINDOWS 98™, LINUX), MACINTOSH™, Power PC або UNIX (наприклад, робоча станція SUN™) або інший серійний комп'ютер, відомий фахівцям. Програмне забезпечення для вирівнювання або інших маніпуляцій з послідовностями може бути придбане або легко розроблено фахівцем у відповідній області техніки з використанням стандартних мов програмування, таких як Visualbasic, Fortran, Basic, Java і т.п.

Будь-який контролер або комп'ютер може включати монітор, який часто являє собою дисплей на основі електронно-променевої трубки (ЕПТ-дисплей), дисплей з плоским екраном (наприклад, рідкокристалічний дисплей з активною матрицею, рідкокристалічний дисплей) та інші.



Комп'ютерну схему часто вміщують в корпус, який включає багато інтегральних мікросхем, таких як мікропроцесор, пам'ять, схеми інтерфейса та інші. Корпус також може включати жорсткий диск, дисковод гнучких дисків, змінні носії високої місткості, такі як CD-ROM, що перезаписується, та інші звичайні елементи периферії. Додатково надаються пристрої вводу, такі як клавіатура або маніпулятор "миша", для вводу користувачем і для вибору користувачем послідовностей, які підлягають порівнянню або іншим маніпуляціям у відповідній комп'ютерній системі.

Звичайно комп'ютер містить відповідне програмне забезпечення для отримання інструкцій користувача, або в формі вводу користувачем в поля заданих параметрів, наприклад, в графічному інтерфейсі користувача, або в формі заздалегідь запрограмованих інструкцій, наприклад, інструкцій для множини різних специфічних операцій. Тоді програмне забезпечення конвертує ці інструкції у відповідну мову для керування напрямою течії і роботою контролера транспорту для здійснення бажаної операції.

Програмне забезпечення також може включати вихідні елементи для контролю синтезу нуклеїнової кислоти (наприклад, на основі послідовності або вирівнювання послідовностей, описаних тут) або інших операцій, що виконуються з використанням ланцюга символів, відповідаючого описаній тут послідовності. Відповідним чином, обладнання для синтезу нуклеїнових кислот може бути компонентом однієї або більше описаних тут інтегрованих систем.

У додатковому аспекті цей винахід стосується комплектів, що поєднують способи, композиції, системи і апаратуру, описані тут. Комплект за винаходом може включати один або більше з таких компонентів: (1) апарат, систему, компонент системи або компонент апарату, описані тут; (2) інструкції для здійснення способів, описаних тут, і/або для управління апаратом або компонентом апарату, описаними тут, і/або для застосування композицій, описаних тут; (3) одну або більше GAT-композицію або компонент; (4) контейнер для вмісту компонентів або композицій і (5) пакувальні матеріали.

У наступному аспекті цей винахід стосується застосування будь-якого апарату, компонента апарату, композиції або комплексу, описаних тут, для здійснення будь-якого способу або аналізу, описаних тут, і/або застосування будь-якого апарату або комплексу для здійснення будь-якого аналізу або способу, описаних тут.

Клітини та організми-хазяї

Клітини-хазяї можуть бути еукаріотичними, наприклад, еукаріотична клітина, рослинна клітина, тваринна клітина, протопласт або тканинна культура. Клітина-хазяїн може включати множини клітин, наприклад, організм. В альтернативі, клітина-хазяїн може бути прокаріотичною, включаючи, але не обмежуючись, бактерії (тобто, грам-позитивні бактерії, пурпурні бактерії, зелені сірчані бактерії, зелені несірчані бактерії, цианобактерії, спірохети, терматогали, флавобактерії і бактероїди) і архебактерії (тобто, Korarchaeota, Thermoproteus, Pyrodictum, Thermococcales, ме-

тан-продукуючі бактерії, Archaeoglobus і екстремальні галофіли). Трансгенні рослини, або рослинні клітини, що включають GAT-нуклеїнові кислоти і/або експресуючі GAT-поліпептиди за винаходом, є рисою цього винаходу. Трансформація рослинних клітин і протопластів може проводитися по суті будь-яким з різних способів, відомих фахівцям з рослинної молекулярної біології, включаючи, але не обмежуючись, способи, описані тут. Див., загалом, публікацію *Methods in Enzymology*. Vol. 153 (Recombinant DNA PartD) Wu and Grossman (eds.) 1987, Academic Press, що включена в цю заявку у вигляді посилання. Вираз "трансформація" тут означає зміну генотипу рослини-хазяїна шляхом введення нуклеїновокислотної послідовності, наприклад, "гетерологічної" або "чужорідної" нуклеїновокислотної послідовності. Гетерологічна нуклеїновокислотна послідовність не обов'язково походить з іншого джерела, але в деякій точці вона є зовнішньою до клітини, в яку її вводять.

На додаток до Berger, Ausubel і Sambrook, корисні загальні посилання щодо клонування, культивування і регенерації рослинних клітин включають Jones (ed) (1995) *Plant Gene Transfer and Expression Protocols - Methods in Molecular Biology*. Volume 49 Humana Press Towata NJ; Payne et al. (1992) *Plant Cell and Tissue Culture in Liquid Systems* John Wiley & Sons, Inc. New York, NY (Payne); Gamborg and Phillips (eds) (1995) *Plant Cell, Tissue and Organ Culture: Fundamental Methods* Springer Lab Manual, Springer-Verlag (Berlin Heidelberg New York) (Gamborg). Пізні середовища клітинних культур описані в Atlas and Parks (eds) *The Handbook of Microbiological Media* (1993) CRC Press, Boca Raton, FL (Atlas). Додаткову інформацію стосовно культур рослинних клітин можна знайти в комерційній літературі, наприклад, *Life Science Research Cell Culture Catalogue* (1998) від Sigma-Aldrich, Inc (St Louis, MO) (Sigma-LSRCCC) і, наприклад, *Plant Culture Catalogue* і додатках (1997), також від Sigma-Aldrich, Inc (St Louis, MO) (Sigma-PPCS). Додаткові подробиці стосовно культур рослинних клітин можна знайти в Cray, (ed.) (1993) *Plant Molecular Biology* Bios Scientific Publishers, Oxford, U.K.

В одному з варіантів здійснення винаходу отримують рекомбінантні вектори, що включають один або більше GAT-полінуклеотидів, придатні для трансформації рослинних клітин. ДНК-послідовність, кодуючу бажаний GAT-поліпептид, наприклад, вибраний серед SEQ ID NOS: 1-5 і 11-262, звичайно використовують для конструювання рекомбінантної касети експресії, яка може бути введена в бажану рослину. В контексті цього винаходу, касета експресії звичайно буде містити вибраний GAT-полінуклеотид, оперативно зв'язаний з промоторною послідовністю та іншими регуляторними послідовностями ініціації транскрипції і трансляції, достатніми для управління транскрипцією GAT-послідовності у відповідних тканинах (наприклад, цілій рослині, листі, корінні і т.д.) трансформованої рослини.

Наприклад, може бути використаний сильно або слабо конститутивний промотор, який буде направляти експресію GAT-поліпептиду у всіх

тканинах рослини. Такі промотори є активними у більшості навколишніх умов і станів розвитку або диференціації клітин. Приклади конститутивних промоторів включають 1'- або 2'-промотор, отриманий з Т-ДНК *Agrobacterium tumefaciens*, та інші області ініціації транскрипції з різних рослинних генів, відомих фахівцям. Після ознайомлення з цим описом фахівцеві буде зрозуміло, що у випадках, коли надмірна експресія GAT-полінуклеотиду є згубною для рослини або небажаною з інших причин, для низьких рівнів експресії може бути використаний слабо конститутивний промотор. У випадках, коли високі рівні експресії не шкідливі для рослини, може бути використаний сильний промотор, наприклад, т-РНК або інший промотор *rol III*, або сильний промотор *rol II* (наприклад, промотор мозаїчного вірусу кольорової капусти, CaMV, промотор 35S).

В альтернативі, рослинний промотор може знаходитися під контролем навколишнього середовища. Такі промотори називаються тут промоторами, що індукуються. Приклади умов навколишнього середовища, які можуть здійснювати транскрипцію за допомогою промоторів, що індукуються, включають атаку патогенів, анаеробні умови або присутність світла. В деяких випадках бажано використання промоторів, що є "тканино-специфічними" і/або знаходяться під еволюційним контролем внаслідок того, що GAT-полінуклеотид експресується лише у певних тканинах або на певних стадіях розвитку, наприклад, у листі, корінні, паростках і т.д. Ендогенні промотори генів, що мають відношення до гербіцидної толерантності, і відповідних фенотипів особливо корисні для здійснення експресії GAT-нуклеїнових кислот, наприклад, монооксигенази P450, глутатіон-S-трансферази, гомоглутатіон-S-трансферази, гліфосат-оксидази і 5-енолпірувілкімат-2-фосфат-синтази.

Тканино-специфічні промотори також можуть бути використані для управління експресією гетерологічних структурних генів, включаючи GAT-полінуклеотиди, описані тут. Так, промотори можуть бути використані в рекомбінантних касетах експресії для здійснення експресії будь-якого гена, експресія якого є бажаною в трансгенних рослинах за винаходом, наприклад, GAT і/або інших генів, що надають стійкість або толерантність до гербіцидів, генів, що впливають на інші корисні характеристики, наприклад, гетерозис. Подібним чином, енхансерні елементи, наприклад, отримані з 5'-регуляторних послідовностей або інтрона гетерологічного гена, також можуть бути використані для покращення експресії гена гетерологічної структури, такого як GAT-полінуклеотид.

Загалом, конкретний промотор для використання в касеті експресії в рослинах залежить від майбутнього призначення. Будь-які з численних промоторів, що направляють транскрипцію в рослинних клітинах, можуть бути придатними. Промотор може бути конститутивним або таким, що індукується. На додаток до промоторів, зазначених вище, промотори бактеріального походження, які працюють в рослинах, включають промотор октопін-синтази, промотор нопалін-синтази та

інші промотори, отримані з Ti-плазмід. Див. Herrera-Estrella et al. (1983) *Nature* 303:209. Вірусні промотори включають РНК-промотори CaMV 35S і 19S. Див. Odell et al. (1985) *Nature* 313:810. Інші рослинні промотори включають малий субодиничний промотор рибулоза-1,3-бісфосфат-карбоксилази і фазеоліновий промотор. Також успішно може бути використана промоторна послідовність з гена E8 (див. Deikman and Fischer (1988) *EMBO J.* 7:3315) та інших генів. Також до уваги приймаються промотори, специфічні для односім'ядольних видів (McElroy D., Bretell R.I.S. 1994. *Foreign gene expression in transgenic cereals. Trends Biotech.*, 12:62-68). В альтернативі, можуть бути ідентифіковані нові промотори з корисними характеристиками з будь-якого вірусного, бактеріального або рослинного джерела різними способами, включаючи аналіз послідовностей, захоплення енхансера або промотора і т.п., відомими у відповідній технології.

Для отримання векторів за винаходом також успішно використовуються інші послідовності, крім промотора і GAT-кодуючого гена. Якщо потрібна належна експресія поліпептиду, область поліаденілації може бути отримана з природного гена, з різних інших рослинних генів або з Т-ДНК. Також можуть бути використані сигнальні/локалізаційні пептиди, які, наприклад, полегшують транслокацію експресованого поліпептиду у внутрішні органели (наприклад, хлоропласти) або позаклітинну секрецію.

Вектор, що містить GAT-полінуклеотид, також може включати маркерний ген, який надає рослинним клітинам фенотип, що селектується. Наприклад, маркер може кодувати біоцидну толерантність, зокрема, толерантність до антибіотиків, таку як толерантність до хлорсульфурону або фосфінотрицину. Гени-репортери, які використовуються для спостереження за експресією генів і локалізацією протеїну через візуалізовані продукти реакції (наприклад, бета-глюкуронідаза, бета-галактозидаза і хлорамфенікол-ацетилтрансфераза) або шляхом прямої візуалізації самого генного продукту (наприклад, зелений флуоресцентний протеїн, GFP; Sheen et al. (1995) *The Plant Journal* 8:777), можуть бути використані наприклад, для спостереження за тимчасовою експресією гена в рослинних клітинах. Системи тимчасової експресії можуть бути використані в рослинних клітинах, наприклад, для скринінгу культур рослинних клітин на гербіцидну толерантність.

Трансформація рослин

Протопласти

Численні протоколи отримання протопластів, що трансформуються, з різних типів рослин і наступної трансформації культивованих протопластів описані в літературі і включені в цю заявку у вигляді посилання. Наприклад, див. Hashimoto et al. (1990) *Plant Physiol.* 93:857; Fowke and Constabel 9eds) (1994) *Plant Protoplasts: Saunders et al. (1993) Applications of Plant In Vitro Technology Symposium. UPM 16-18; i Lyznik et al. (1991) BioTechniques 10:295*, кожна з яких включена в цю заявку у вигляді посилання.

Хлоропласти

Хлоропласти являють собою ділянку дії деяких активностей гербіцидної толерантності, і в деяких випадках GAT-полінуклеотид зливають з хлоропласт-транзитним пептидом для полегшення транслокації генних продуктів в хлоропласти. В цих випадках може бути сприятливо трансформувати GAT-полінуклеотид в хлоропласти рослинних клітин-хазяїв. У відповідній технології існують численні способи здійснення трансформації і експресії в хлоропластах (наприклад, Daniell et al. (1998) *Nature Biotechnology* 16:346; O'Neill et al. (1993) *The Plant Journal* 3:729; Maliga (1993) *TIBTECH* 11:1). Конструкція експресії включає транскрипційну регуляторну послідовність, функціональну в рослинах, оперативно зв'язану з полінуклеотидом, кодуєчим GAT-поліпептид. Касети експресії, сконструйовані для функціонування в хлоропластах (такі як касета експресії, що включає GAT-полінуклеотид), включають послідовності, необхідні для забезпечення експресії в хлоропластах. Звичайно кодуєча послідовність фланкована двома областями гомологічності з хлоропластидним геномом для здійснення гомологічної рекомбінації з геномом хлоропласта; часто всередині фланкуючих пластидних ДНК-послідовностей також присутній ген маркера, що селектується, для полегшення селекції генетично стабільно трансформованих хлоропластів в результируючих транспластонових рослинних клітинах (див., наприклад, Maliga (1993) і Daniell (1998), а також наведені там посилання).

Загальні способи трансформації

ДНК-конструкції за винаходом можуть бути введені в геном бажаної рослини-хазяїна багатьма звичайними технологіями. Технології трансформації різноманітних видів вищих рослин добре відомі і описані в технічній і науковій літературі. Див., наприклад, Payne, Gamborg, Cray, Jones і т.д., усі вище, а також, наприклад, Weising et al. (1998) *Ann. Rev. Genet.* 22:421.

Наприклад, ДНК можуть бути введені безпосередньо в геномну ДНК рослинної клітини з використанням таких технологій, як електропорація і мікроін'єкції протопластів рослинної клітини, або ДНК-конструкції можуть бути введені безпосередньо в тканину рослини з використанням балістичних способів, таких як бомбардування частками, що містять ДНК. В альтернативі, конструкції ДНК можуть бути поєднані з придатними Т-ДНК-фланкуючими областями і введені в звичайний вектор хазяїна *Agrobacterium tumefaciens*. Функції вірулентності хазяїна *Agrobacterium* будуть направляти вставлення конструкції і суміжного маркера в ДНК рослинної клітини, якщо інфікувати рослинну клітину бактеріями.

Методи мікроін'єкцій відомі у відповідній області і детально описані в науковій і патентній літературі. Введення конструкцій ДНК з використанням поліетилеңглікольного осадження описано в Paszkowski et al. (1984) *EMBO J* 3:2717. Методи електропорації описані в Fromm et al. (1985) *Proc. Natl. Acad. Sci. USA* 82:5824. Методи балістичної трансформації описані в Klein et al. (1987) *Nature* 327:70; і Weeks et al. *Plant Physiol* 102:1077.

В деяких варіантах здійснення винаходу використовуються *Agrobacterium*-опосереднені методи трансформації для переносу GAT-послідовностей за винаходом в трансгенні рослини. *Agrobacterium*-опосереднена трансформація широко використовується для трансформації двосім'ядольних рослин, проте деякі односім'ядольні також можуть бути трансформовані *Agrobacterium*. Наприклад, *Agrobacterium*-трансформація рису описана Hiei et al. (1994) *Plant J.* 6:271; патент США №5,187,073; патент США №5,591,616; Li et al. (1991) *Science in China* 34:54; і Raineri et al. (1990) *Bio/Technology* 8:33. Також описані трансформована кукурудза, ячмінь, тритикале і спаржа шляхом *Agrobacterium*-опосередненої трансформації (Xu et al. (1990) *Chinese J Bot* 2:81).

Методи *Agrobacterium*-опосередненої трансформації використовують здатність пухлиноіндукуючої (Ti) плазмиди *A. tumefaciens* вбудовуватися в геном рослинної клітини для спільного переносу нуклеїнової кислоти, що представляє інтерес, в рослинну клітину. Звичайно виробляють вектор експресії, в якому нуклеїнова кислота, що представляє інтерес, така як GAT-полінуклеотид за винаходом, лігвана в плазмиду, що автономно реплікується, яка теж містить Т-ДНК-послідовності. Т-ДНК-послідовності звичайно фланкують нуклеїнову кислоту касети експресії, що представляє інтерес, і включають послідовності інтеграції плазмиди. На додаток до касети експресії Т-ДНК також звичайно включає маркерну послідовність, наприклад, гени стійкості до антибіотиків. Потім плазмиду з Т-ДНК і касетою експресії трансфектують в клітини *Agrobacterium*. Звичайно для ефективної трансформації рослинних клітин бактерія *A. tumefaciens* також має необхідні області *vir* в плазмиді або вбудовані в хромосому. Інформацію стосовно *Agrobacterium*-опосередненої трансформації див. у Firoozabady and Kuehnle, (1995) *Plant Cell Tissue and Organ Culture: Fundamental Methods*. Gamborg and Phillips (eds.).

Регенерація трансгенних рослин

Трансформовані рослинні клітини, отримані за методиками рослинної трансформації, включаючи описані вище, можуть бути культивовані для регенерації цілої рослини, яка має трансформований генотип (тобто, містить GAT-полінуклеотид) і, таким чином, бажаний фенотип, наприклад, набути стійкість (тобто, толерантність) до гліфосату або аналога гліфосату. Такі методи регенерації ґрунтовані на маніпулюванні певними фітогормонами в середовищі росту тканинної культури, звичайно ґрунтованому на біоцидному і/або гербіцидному маркері, що вводиться разом з бажаними нуклеотидними послідовностями. В альтернативі, може бути проведена селекція на стійкість до гліфосату, що надається GAT-полінуклеотидом за винаходом. Регенерація рослини з культивованих протопластів описана в Evans et al. (1983) *Protoplasts Isolation and Culture*. Handbook of Plant Cell Culture, pp.124-176, Macmillan Publishing Company, New York; і Binding (1985) *Regeneration of Plants*. Plant Protoplasts pp.21-73, CRC Press, Boca Raton. Pe-

генерація також може бути отримана з калуса, експлантатів, органів рослини або їх частин. Такі регенераційні методики загалом описані у Klee et al. (1987) *Ann Rev of Plant Phys* 38:467. Див. також, наприклад, Payne and Gamborg. Після трансформації *Agrobacterium* експлантати звичайно переносять в середовище селекції. Фахівцю буде зрозуміло, що середовище селекції буде залежати від маркера, що селекується, який був спільно трансфектований в експлантати. Через достатній період часу трансформанти почнуть утворювати паростки. Коли паростки досягнуть 1-2 см в довжину, вони повинні бути перенесені у відповідне середовище для коріння і паростків. У середовищі для коріння і паростків повинен підтримуватися тиск селекції.

Звичайно трансформанти будуть розвивати коріння протягом ~1-2 тижнів і формувати прорості. Коли прорості досягнуть 3-5 см заввишки, їх вміщують в стерильний ґрунт у горщики з фібри. Фахівцям буде зрозуміло, що для отримання трансформованих рослин різних видів використовуються різні процедури акліматизації. Наприклад, після розвитку коріння й паростків живці, а також соматичні зародки трансформованих рослин переносять в середовище для розвитку проростей. Опис селекції і регенерації трансформованих рослин див., наприклад, у Dodds and Roberts (1995) *Experiments in Plant Tissue Culture*. 3<sup>rd</sup> Ed., Cambridge University Press.

Також існують способи *Agrobacterium*-трансформації *Arabidopsis* з використанням вакуумної інфільтрації (Bechtold N., Ellis J. and Pelletier G., 1993, *In planta Agrobacterium mediated gene transfer by infiltration of adult Arabidopsis thaliana plants*. *CR Acad Sci Paris Life Sci* 316:1194-1199) і простого занурення квітучих рослин (Desfeux, C, Clough S.J., and Bent A.J.P., 2000, *Female reproductive tissues are the primary target of Agrobacterium-mediated transformation by the Arabidopsis floral-dip method*. *Plant Physiol*. 123:895-904). Цими способами виробляють трансгенне насіння без необхідності в тканинній культурі.

Існує багато рослин, для яких протоколи ефективної *Agrobacterium*-опосередненої трансформації ще мають бути розроблені. Наприклад, успішна тканинна трансформація разом з регенерацією трансформованої тканини для вироблення трансгенної рослини ще не описана для деяких найбільш комерційно популярних культурних сортів бавовни. Проте, підхід, що може бути використаний для цих рослин, включає стабільне введення полінуклеотиду у відповідне різноманіття рослин шляхом *Agrobacterium*-опосередненої трансформації, підтвердження працездатності, а потім - перенос трансгена в бажаний серійний штам з використанням стандартних методик статевого схрещування або зворотного схрещування. Наприклад, у випадку бавовни, *Agrobacterium* може бути використана для трансформації лінії Coker Gossypium hirsutum (наприклад, лінії Coker 310, 312, 5110 Deltapine 61 або Stoneville 213), а потім трансген може бути введений в інший, більш комерційно відповідний сорт *G. hirsutum* шляхом зворотного схрещування.

Трансгенні рослини за винаходом можуть бути проаналізовані за генотипом або за фенотипом для визначення присутності GAT-полінуклеотиду за винаходом. Аналіз генотипу може бути проведений будь-яким з численних відомих способів, включаючи ПЛР-ампліфікацію геномної ДНК і гібридизацію геномної ДНК специфічними міченими зондами. Аналіз фенотипу включає, наприклад, виживання рослин або рослинних тканин під дією вибраного гербіциду, такого як гліфосат.

По суті будь-яка рослина може бути трансформована GAT-полінуклеотидами за винаходом. Придатні рослини для трансформації і експресії нових GAT-полінуклеотидів за винаходом включають важливі в агрономії і садівництві види. Такі види включають, але не обмежуються, членів сімейств Graminae (в тому числі кукурудзу, жито, тритикале, ячмінь, просо, рис, пшеницю, овес і т.д.); Leguminosae (в тому числі горох, квасолю, сочевицю, арахіс, ямс, вігну, квасолю мунго, сою, конюшину, люцерну, люпин, вику, лотос, буркун, гліцинію і запашний горошок); Compositae (найбільше сімейство судинних рослин, що включає щонайменше 1,000 видів, в тому числі важливі комерційні культури, такі як соняшник) і Rosaciae (в тому числі малину, абрикос, мигдаль, персик, розу і т.д.), а також горіхоплідні рослини (в тому числі волоський горіх, горіх пекан, лісовий горіх і т.д.) і лісові дерева (в тому числі Pinus, Quercus, Pseudotsuga, Sequoia, Populus.etc.)

Додаткові цілі для модифікації GAT-полінуклеотидами за винаходом, а також описаними вище, включають рослини з видів Agrostis, Allium, Antirrhinum, Apium, Arachis, Asparagus, Atropa, Avena (наприклад, овес), Bambusa, Brassica, Bromus, Browallia, Camellia, Cannabis, Capsicum, Cicer, Chenopodium, Chichorium, Citrus, Coffea, Coix, Cucumis, Curcubita, Cynodon, Dactylis, Datura, Daucus, Digitalis, Dioscorea, Elaeis, Eleusine, Festuca, Fragaria, Geranium, Gossypium, Glycine, Helianthus, Heterocallis, Hevea, Hordeum (наприклад, ячмінь), Hyoscyamus, Ipomoea, Lactuca, Lens, Lilium, Linum, Lolium, Lotus, Lycopersicon, Majorana, Malus, Mangifera, Manihot, Medicago, Nemesis, Nicotiana, Onobrychis, Oryza (наприклад, рис), Panicum, Pelargonium, Pennisetum (наприклад, просо), Petunia, Pisum, Phaseolus, Phleum, Poa, Prunus, Ranunculus, Raphanus, Ribes, Ricinus, Rubus, Saccharum, Salpiglossis, Secale (наприклад, жито), Senecio, Setaria, Sinapis, Solanum, Sorghum, Stenotaphrum, Theobroma, Trifolium, Trigonella, Triticum (наприклад, пшениця), Vicia, Vigna, Vitis, Zea (наприклад, кукурудза) і Olyreae, Pharioideae та багато інших. Як зазначено вище, рослини сімейства Graminae є особливо переважними цільовими рослинами для способів за винаходом.

Популярні культурні рослини, що є цілями за цим винаходом, включають кукурудзу, рис, тритикале, жито, бавовну, сою, сорго, пшеницю, овес, ячмінь, просо, соняшник, канولا, горох, боби, сочевиця, арахіс, ямс, вігну, квасолю мунго, конюшину, люцерну, люпин, вику, лотос, буркун,

гліцинію, запашний горошок і горіхоплідні рослини (наприклад, волоський горіх, горіх пекан і т.д.).

В одному з аспектів винахід стосується способу вироблення рослинної культури шляхом вирощування культурної рослини, що є гліфосат-толерантною в результаті трансформації геном, кодуєчим гліфосат-М-ацетилтрансферазу, в таких умовах, що рослина продукує культуру, і збирання врожаю культури. Переважно, гліфосат наносять на рослину або поблизу рослини з концентрацією, ефективною для боротьби з бур'яном, не впливаючи на ріст і продукування трансгенної культури. Нанесення гліфосату може виконуватися перед посадкою, або в будь-який час після посадки аж до збору врожаю. Гліфосат може наноситися один або кілька разів. Тривалість нанесення гліфосату, кількість, що наноситься, режим нанесення та інші параметри будуть варіюватися в залежності від специфічної природи культурної рослини і середовища росту і можуть бути легко визначені фахівцем. Винахід також стосується культури, виробленої цим способом.

Винахід стосується розмноження рослини, що містить GAT-полінуклеотидний трансген. Рослина може бути, наприклад, односім'ядольною або двосім'ядольною. В одному з аспектів розмноження включає схрещування рослини, що містить GAT-полінуклеотидний трансген, з другою рослиною, так щоб щонайменше частина потомства після схрещування показувала толерантність до гліфосату.

В одному з аспектів винахід стосується способу селективної боротьби з бур'яном на полі, де вирощується культура. Спосіб включає посадку насіння культури або рослин, що є гліфосат-толерантними в результаті трансформації геном, кодуєчим GAT, наприклад, GAT-полінуклеотид, і нанесення на культуру і бур'ян достатньої кількості гліфосату для боротьби з бур'яном без суттєвого впливу на культуру. Слід зазначити, що культура не повинна бути цілком нечутливою до гербіциду, оскільки користь, отримана від інгібування бур'яну, переважає негативний вплив гліфосату або аналога гліфосату на культуру або культурну рослину.

В іншому аспекті винахід стосується застосування GAT-полінуклеотиду у якості гена маркера, що селектується. В цьому варіанті здійснення винаходу присутність GAT-полінуклеотиду в клітині або організмі надає клітині або організму помітну фенотипову ознаку стійкості до гліфосату, що дає можливість селекції на клітини або організми, які були трансформовані геном, що представляє інтерес, зв'язаним з GAT-полінуклеотидом. Так, наприклад, GAT-полінуклеотид може бути введений в нуклеїновокислотну конструкцію, наприклад, вектор, що забезпечує можливість ідентифікації хазяїна (наприклад, клітини або трансгенної рослини), який містить нуклеїновокислотну конструкцію, шляхом вирощування хазяїна в присутності гліфосату і селекції на здатність виживати і/або рости зі швидкістю, яка є помітно більшою, ніж у хазяїна, що не містить нуклеїновокислотної конструкції. GAT-полінуклеотид може бути застосований у

якості маркера, що селектується, в різноманітних хазяях, що є чутливими до гліфосату, включаючи рослини, більшість бактерій (в тому числі *E.coli*), актиноміцети, дріжджі, водорості і грибки. Одна з переваг застосування стійкості до гербіциду у якості маркера в рослинах, у протилежність звичайній стійкості до антибіотиків, складається в тому, що вона усуває тривогу деяких членів суспільства в тому відношенні, що стійкість до антибіотиків може розповсюдитися у навколишньому середовищі. Деякі дослідницькі дані з експериментів, що демонструють застосування GAT-полінуклеотиду у якості маркера, що селектується, в різноманітних системах хазяїв описані в Прикладах цього документа.

Селекція GAT-полінуклеотидів, що надають підвищену стійкість до гліфосату в трансгенних рослинах

Бібліотеки GAT-кодуєчих нуклеїнових кислот, диверсифіковані способами, описаними тут, можуть бути селекціоновані на здатність надавати стійкість до гліфосату трансгенним рослинам. Після одного або більше циклів диверсифікації і селекції модифіковані GAT-гени можуть бути застосовані у якості маркера селекції для полегшення вироблення і оцінки трансгенних рослин і у якості засобу надання стійкості до гербіциду експериментальним або сільськогосподарським рослинам. Наприклад, після диверсифікації будь-якої одної або більше з SEQ ID NO:1 - SEQ ID NO:5 для вироблення бібліотеки GAT-полінуклеотидів первинна функціональна оцінка може бути здійснена шляхом експресії бібліотеки GAT-кодуєчих послідовностей в *E.coli*. Експресовані GAT-поліпептиди можуть бути очищені або частково очищені, як описано вище, і піддані скринінгу на покращену кінетику за допомогою мас-спектрометрії. Після одного або більше попередніх циклів диверсифікації і селекції полінуклеотиди, кодуєчі покращені GAT-поліпептиди, клонують в рослинний вектор експресії, оперативно зв'язаний, наприклад, з сильним конститутивним промотором, таким як промотор CaMV 35S. Вектори експресії, що містять модифіковані GAT-нуклеїнові кислоти, трансформують, звичайно шляхом *Agrobacterium*-опосередкованої трансформації, в рослини-хазяї *Arabidopsis thaliana*. Наприклад, *Arabidopsis*-хазяї легко трансформуються шляхом занурення суцвіть в розчини *Agrobacterium*, після чого їм дозволяють рости і утворювати насіння. Тисячі насінин можуть бути отримані приблизно за 6 тижнів. Тоді збирають усю масу насіння від занурених рослин і пророщують у ґрунті. Таким чином можна генерувати кілька тисяч незалежно трансформованих рослин для оцінки у форматі високопродуктивної трансформації рослин (HTP). Пророщене в масі насіння обприскують гліфосатом, і сіянці, що вижили і показали стійкість до гліфосату, виживають в процесі селекції, в той час як нетрансгенні рослини і рослини, що включають менш сприятливо модифіковані GAT-нуклеїнові кислоти, пошкоджуються або гинуть після гербіцидної обробки. Можлива регенерація GAT-кодуєчих нуклеїнових кислот, що надають покращену стійкість до гліфосату, наприклад, шляхом ПЛР-ампліфікації з

використанням Т-ДНК-праймерів, фланкуючих вставки бібліотек, і застосування їх в подальших процедурах диверсифікації або для вироблення додаткових трансгенних рослин того самого або інших видів. Якщо це бажано, можуть бути проведені додаткові цикли диверсифікації і селекції з використанням концентрацій гліфосату, що підвищуються у кожній наступній селекції. Таким чином можуть бути отримані GAT-полінуклеотиди і поліпептиди, що надають стійкість до концентрацій гліфосату, корисних у польових умовах.

Стійкість до гербіцидів

Механізм стійкості до гліфосату за винаходом може бути поєднаний з іншими формами стійкості до гліфосату, відомими у відповідній області, призначеними для вироблення рослин і рослинних експлантатів з високою стійкістю до гліфосату. Наприклад, гліфосат-толерантні рослини можуть бути вироблені шляхом вставлення в геном рослини здатності продукувати вищі рівні 5-енолпірувілшкімат-3-фосфат-синтази (EPSP), як найбільш детально описано в патентах США №№6,248,876 В1; 5,627,061; 5,804,425; 5,633,435; 5,145,783; 4,971,908; 5,312,910; 5,188,642; 4,940,835; 5,866,775; 6,225,114 В1; 6,130,366; 5,310,667; 4,535,060; 4,769,061; 5,633,448; 5,510,471; Re. 36,449; RE 37,287 Е і 5,491,288, а також в міжнародних публікаціях WO 97/04103; WO 00/66746; WO 01/66704 і WO 00/66747, що включені в цю заявку у вигляді посилань у всій їх повноті для всіх призначень. Стійкість до гліфосату також властива рослинам, які експресують ген, кодуючий фермент гліфосат-оксидо-редуктазу, як описано більш детально в патентах США №№5,776,760 і 5,463,175, що включені в цю заявку у вигляді посилань у всій їх повноті для всіх призначень.

Крім того, механізм стійкості до гліфосату за винаходом може бути поєднаний з іншими формами гербіцидної стійкості для отримання рослин і рослинних експлантатів, стійких до гліфосату і одного або більше інших гербіцидів. Наприклад, гідроксифенілпіруватдіоксигенази являють собою ферменти, що каталізують реакцію, в якій пара-гідроксифенілпіруват (HPP) трансформується в гомогентизат. Молекули, які інгібують цей фермент, і які зв'язуються з ферментом, інгібуючи трансформацію HPP в гомогентизат, корисні у якості гербіцидів. Рослини, більш стійкі до певних гербіцидів, описані в патентах США №№6,245,968 В1; 6,268,549 і 6,069,115, а також в міжнародній публікації WO 99/23886, що включені в цю заявку у вигляді посилань у всій їх повноті для всіх призначень.

Гербіциди сульфонілсечовина і імідазолінон також інгібують ріст вищих рослин шляхом блокування ацетолактат-синтази (ALS) або синтази ацетогідроксикислоти (AHAS). Вироблення сульфонілсечовина- і імідазолінон-толерантних рослин описано більш детально в патентах США №№5,605,011; 5,013,659; 5,141,870; 5,767,361; 5,731,180; 5,304,732; 4,761,373; 5,331,107; 5,928,937 і 5,378,824, а також в міжнародній публікації WO 96/33270, що включені в цю заявку у вигляді посилань у всій їх повноті для всіх призначень.

Глутамін-синтетаза (GS), судячи з усього, є незамінним ферментом, необхідним для розвитку і життєдіяльності більшості рослинних клітин. Інгібітори GS є токсичними для рослинних клітин. На основі цього токсичного ефекту завдяки інгібуванню GS в рослинах були розроблені гліфосінатні гербіциди. Ці гербіциди є невибірковыми. Вони інгібують ріст всіх існуючих видів рослин, викликаючи їх загальну деструкцію. Розвиток рослин, що містять екзогенну фосфінотрицин-ацетилтрансферазу, описаний в патентах США №№5,969,213; 5,489,520; 5,550,318; 5,874,265; 5,919,675; 5,561,236; 5,648,477; 5,646,024; 6,177,616 В1 і 5,879,903, що включені в цю заявку у вигляді посилань у всій їх повноті для всіх призначень.

Протопорфіриноген-оксидаза (Protox) необхідна для вироблення хлорофілу, який є необхідним для життєдіяльності рослин. Фермент Protox служить ціллю для різних гербіцидних сполук. Ці гербіциди також інгібують ріст всіх існуючих видів рослин, викликаючи їх загальну деструкцію. Розвиток рослин, що містять змінену Protox-активність і є стійкими до цих гербіцидів, описані в патентах США №№6,288,306 В1; 6,282,837 В1 і 5,767,373, а також в міжнародній публікації WO 01/12825, що включені в цю заявку у вигляді посилань у всій їх повноті для всіх призначень.

Приклади

Наступні приклади є лише ілюстративними і не обмежуючими. Фахівцеві буде зрозуміло, що існує велика кількість некритичних параметрів, які можуть змінюватися з отриманням по суті подібних результатів.

Приклад 1: Ізоляція нових нативних GAT-полінуклеотидів

П'ять нативних GAT-полінуклеотидів (тобто, GAT-полінуклеотидів, що існують в природі в генетично немодифікованому організмі) були відкриті шляхом експресійного клонування послідовностей з штамів *Bacillus*, демонструючих GAT-активність. Їх нуклеотидні послідовності були визначені і представлені тут як SEQ ID NO:1 - SEQ ID NO:5. У короткому викладі, колекцію близько 500 штамів *Bacillus* і *Pseudomonas* піддали скринінгу на нативну здатність N-ацетилювати гліфосат. Штами вирощували в LB протягом ночі, зібрали шляхом центрифугування, пермеабілізували в розведеному толуолі і потім промили і пересушували в реакційній суміші, що містила буфер, 5мМ гліфосату і 200мкМ ацетил-CoA. Клітини інкубували в цій реакційній суміші від 1 до 48 годин, додавши за цей час в реакцію рівний об'єм метанолу. Потім клітини пелетизували шляхом центрифугування і супернатант відфільтрували перед аналізом методом мас-спектроскопії батьківських іонів. Продукт реакції був позитивно ідентифікований як N-ацетилгліфосат шляхом порівняння профілю мас-спектроскопії реакційної суміші з еталоном N-ацетилгліфосату, як показано на Фіг.2. На виявлення продукту могло впливати включення обох субстратів (ацетил-CoA та гліфосату), що було усунено шляхом теплової денатурації бактеріальних клітин.

Потім окремі GAT-полінуклеотиди клонували з ідентифікованих штамів шляхом функціонального скринінгу. Отримали геномну ДНК і частково розщепили ферментом Sau3A1. Фрагменти ~4Kb клонували в вектор експресії E.coli і трансформували в електрокомпетентні E.coli. Окремі клони, що показували GAT-активність, ідентифікували мас-спектроскопією після реакції, описаної вище, за тим винятком, що толуол замінили пермеабілізацією в PMBS. Геномні фрагменти секвенували і ідентифікували очікувану GAT-поліпептид-кодуючу відкриту рамку зчитування. Ідентичність GAT-гена підтвердили експресією відкритої рамки зчитування в E.coli і виявленням високих рівнів N-ацетил гліфосату, вироблених в реакційних сумішах.

Приклад 2: Оцінка характеристик GAT-поліпептиду, ізолюваного з штаму B6 B. licheniformis

Геномну ДНК з штаму B6 B. licheniformis очищали, частково розщепили ферментом Sau3A1 і фрагменти 1-10Kb клонували в вектор експресії E.coli. Клон з вставкою 2,5kb надав хазяїну E.coli активність гліфосат-N-ацетилтрансферази (GAT), що було виявлено мас-спектроскопією. Секвенування вставки виявило одну повну відкриту рамку зчитування з 441 пари основ. Наступне клонування цієї відкритої рамки зчитування підтвердило, що вона кодує GAT-фермент. Плазмиду pMAXY2120, показану на Фіг.4, з геном, кодуючим GAT-фермент штаму B6, трансформували в штам E.coli XL1 Blue. 10% інокуляту насиченої культури додали в бульйон Luria і інкубували культуру при 37°C протягом 1 години. Експресію GAT індукували шляхом додавання IPTG при концентрації 1мМ. Культуру інкубували ще 4 години, після чого клітини зібрали шляхом центрифугування і пелетизовані клітини заморозили при -80°C.

Лізис клітин здійснили шляхом додавання 1мл до 0,2г клітин такого буфера: 25мМ HEPES, pH 7,3, 100мМ KCl і 10% метанолу (HKM) з 0,1мМ EDTA, 1мМ DTT, 1мг/мл лізозиму курячого яйця і протеаза-інгібіторний коктейль виробництва Sigma, використаний за інструкціями виробника. Після 20 хвилин інкубації при кімнатній температурі (наприклад, 22-25°C) лізис доповнили короткою ультразвуковою руйнацією. Лізат центрифугували і супернатант знесолили пропусканням через Sephadex G25, врівноважену HKM. Часткове очищення провели шляхом афінної хроматографії на CoA Agarose (Sigma). Колонку врівноважили HKM і прояснений екстракт пропустили через колонку під гідростатичним тиском. Незв'язані протеїни видалили промивкою колонки HKM, а GAT елюювали HKM з 1мМ Coenzyme A. Ця процедура забезпечила 4-кратне очищення. На цій стадії близько 65% забарвлення протеїну, що спостерігалось на SDS-поліакриламідному гелі, завантаженому неочищеним лізатом, було зумовлено GAT, а ще 20% - хлорамфенікол-ацетилтрансферазою, закодованою вектором.

Очищення до гомогенності було досягнуто гель-фільтрацією частково очищеного протеїну через Superdex 75 (Pharmacia). В мобільній фазі, що являла собою HKM, GAT-активність елююва-

ли при об'ємі, відповідаючому молекулярному радіусу 17KDa. Цей матеріал був гомогенним на основі оцінки шляхом забарвлення Coomassie 3-м кг зразка GAT, який піддали SDS-поліакриламідному гель-електрофорезу на 12%-му акриламідному гелі товщиною 1мм. Очищення було досягнуто з 6-кратним підвищенням специфічної активності.

Видиму  $K_m$  до гліфосату визначали на реакційних сумішах, які містили ацетил-CoA, що насичується (200мкМ), різні концентрації гліфосату і 1мкМ очищеного GAT в буфері, що містив 5мМ морфоліну, відрегульованого до pH 7,7 оцтовою кислотою і 20% етиленгліколем. Початкові швидкості реакції визначали шляхом безперервного моніторингу гідролізу тіоефірного зв'язку ацетил-CoA при 235нм ( $E=3,4OD/мМ/см$ ). Спостерігалася гіперболічна кінетика насичення (Фіг.5), з якої отримали видиму  $K_m=2,9\pm0,2$  (стандартне відхилення) мМ.

Видиму  $K_m$  до AcCoA визначали на реакційних сумішах, які містили 5мМ гліфосату, різні концентрації ацетил-CoA і 0,19мкМ GAT в буфері, що містив 5мМ морфоліну, відрегульованого до pH 7,7 оцтовою кислотою і 50% метанолом. Початкові швидкості реакції визначали з використанням мас-спектрометричного виявлення N-ацетилгліфосату. В прилад многократно інжектували 5-мкл порції і отримували швидкості реакції з графіка площини інтегрованого піку від часу реакції (Фіг.6). Спостерігалася гіперболічна кінетика насичення (Фіг.7), з якої отримали видиму  $K_m=2мкМ$ . З величин  $V_{max}$ , отриманих при відомій концентрації ферменту, розраховували  $K_{cat}=6/хв$ .

Приклад 3: Мас-спектрометричний (МС) спосіб скринінгу

Зразки (5мкл) відбирають з 96-лункової мікротитрувальної пластини зі швидкістю 1 зразок кожні 26 секунд і інжектують в мас-спектрометр (Micromass Quattro LC, потрібний квадрупольний мас-спектрометр) без будь-якого розділення. Зразок інжектують в мас-спектрометр за допомогою мобільної фази вода/метанол (50:50) зі швидкістю потоку 500Uл/хв. Кожний введений зразок іонізують способом негативної електророзпилювальної іонізації (напруга голчастого електрода - 3,5KV, напруга конічного електрода - 20V, температура джерела 120°C, температура десольвації 250°C, швидкість течії газу з конуса 90 л/год., швидкість течії газу десольвації 600л/год.). Іони молекул ( $m/z$  (маса/заряд) 210), утворені під час цього процесу, вибираються першим квадруполем для виконання дисоціації, індукованої зіткненнями (CID), в другому квадруполі, де тиск встановлюють на  $5\times10^{-4}$ мбар, а енергію удару регулюють до 20Ев. Третій квадруполь встановлюють таким чином, щоб тільки дозволити одному з дочірніх іонів ( $m/z$  124), отриманих з батьківських іонів ( $m/z$  210) потрапити в детектор для запису сигналу. Перший і третій квадруполі встановлюють на одиничне розділення, в той час як фотопомножувач працює при 650В. Чисті еталони N-ацетилгліфосату використовують для порівняння, а інтегрування піків використовують для оцінки концентрацій. За допомогою цього способу

можливо виявлення менш ніж 200Nм N-ацетилгліфосату.

Приклад 4: Виявлення нативних ферментів або ферментів з низькою GAT-активністю

Нативні ферменти або ферменти з низькою GAT-активністю звичайно мають  $k_{cat}$  близько  $1\text{хв.}^{-1}$  і  $K_m$  до гліфосату 1,5-10мМ.  $K_m$  до ацетилCoA звичайно менше ніж 25мМ.

Бактеріальні культури вирощують в збагаченому середовищі в глибоких 96-лункових пластинах, і 0,5мл клітин стаціонарної фази збирають шляхом центрифугування, промивають 5мМ ацетатом морфоліну, рН 8, і пересуспендують в 0,1мл реакційної суміші, що містить 200мМ амоній-ацетил-CoA, 5мМ амоній-гліфосату і 5мкг/мл PMBS (Sigma) в 5мМ ацетаті морфоліну, рН 8. PMBS пермеабілізує клітинну мембрану, що дозволяє субстратам і продуктам переходити з клітин в буфер, не вивільняючи в'язкого вмісту клітин. Реакції проводять при 25-37°C протягом 1-48 годин. Реакції гасять рівним об'ємом 100% етанолу і всю суміш фільтрують на 0,45-мкм фільтрувальній пластині MAHV Multiscreen (Millipore). Зразки аналізують мас-спектрометрією, як описано вище, і порівнюють з синтетичними еталонами N-ацетилгліфосату.

Приклад 5: Виявлення GAT-ферментів з високою активністю

GAT-ферменти з високою активністю звичайно мають  $k_{cat}$  до 400 $\text{хв.}^{-1}$  і  $K_m$  нижче 0,1мМ гліфосату.

Гени, кодуєчі GAT-ферменти, клонують в вектори експресії E.coli, такі як pQE80 (Qiagen), і вводять в штами E.coli, наприклад, XL1 Blue (Stratagene). Культури вирощують в 150мкл збагаченого середовища (такого як LB з 50мкг/мл карбеніциліну) в мілких 96-лункових круглодонних полістирольних пластинах до пізньої логарифмічної фази росту і розбавляють 1:9 свіжим середовищем, що містить 1мМ IPTG (USB). Після 4-8-годинної індукції клітини збирають, промивають 5мМ ацетатом морфоліну, рН 6,8, і пересуспендують рівним об'ємом того ж морфолінового буферу. Реакції проводять з кількостями до 10мкл промитих клітин. При вищих рівнях активності клітини спочатку розбавляють до 1:200 і додають 5мкл до 100мкл реакційної суміші. Для вимірювання GAT-активності можна використовувати ту ж реакційну суміш, що описана для низької активності. Проте для виявлення високо-активних GAT-ферментів концентрацію гліфосату знижують до 0,15-0,5мМ, рН знижують до 6,8 і проводять реакції протягом 1 години при 37°C. Обробка продукту реакції і МС-виявлення описані тут.

Приклад 6: Очищення GAT-ферментів

Очищення ферментів здійснюють за допомогою афінної хроматографії клітинних лізатів на CoA-Agarose і гелі-фільтрації на Superdex-75. Кількості очищеного GAT-ферменту до 10мг отримують таким чином: 100мл культуру E.coli, що несе GAT-полінуклеотид в векторі pQE80, яку вирощували протягом ночі в LB з 50мкг/мл карбеніциліну, використовують для інокуляції 1л LB з 50мкг/мл карбеніциліну. Через 1 годину додають до 1мМ IPTG і вирощують культуру ще 6 годин.

Клітини збирають шляхом центрифугування. Лізис здійснюють шляхом суспендування клітин в 25мМ HEPES (рН 7,2), 100 мМ KCl, 10% метанолу (HKM), 0,1мМ EDTA, 1мМ DTT, протеаза-інгібіторному коктейлі виробництва Sigma-Aldrich і 1мг/мл лізозиму курячого яйця. Після 30 хвилин інкубації при кімнатній температурі клітини піддають короткій ультразвуковій руйнації. Уламки клітин видаляють шляхом центрифугування, а лізат пропускають крізь шар коферменту A-Agarose. Колонку промивають кількома об'ємами шару HKM, а GAT елюють в 1,5 об'ємах шару HKM, що містить 1мМ ацетил-CoA. GAT в елюаті концентрують шляхом його утримання над ультрафільтраційною мембраною Centricon YM 50. Подальше очищення здійснюють шляхом пропускання протеїну через колонку Superdex-75 серією 0,6-мл інжекцій. Пік GAT-активності елюється при об'ємі, що відповідає молекулярній масі 17КДа. Цей спосіб забезпечує очищення GAT-ферменту до гомогенності з >85% регенерацією. Подібну методику використовують для отримання до 96 шафлінг-варіантів за один раз у кількостях 0,1-0,4мг кожного. Об'єм індукованої культури знижують до 1-10мл, афінну хроматографію CoA-Agarose виконують в 0,15-мл колонках, включених в фільтрувальну пластину MAHV (Millipore), уникаючи використання хроматографії Superdex-75.

Приклад 7: Стандартний протокол для визначення  $k_{cat}$  і  $K_m$

$k_{cat}$  і  $K_m$  очищеного протеїну до гліфосату визначають з використанням безперервного спектрометричного аналізу, в якому гідроліз сульфоефірного з'язку ацетил-CoA спостерігають при 235нм. Реакції проводять при навколишній температурі (близько 23°C) в лунках 96-лункової аналітичної пластини, при цьому в кінцевому об'ємі 0,3мл присутні такі компоненти: 20мМ HEPES, рН 6,8, 10% етиленгліколю, 0,2мМ ацетил-коферменту A і різні концентрації амоній-гліфосату. При порівнянні кінетики двох GAT-ферментів обидва ферменти повинні аналізуватися в однакових умовах, наприклад, при 23°C.  $k_{cat}$  розраховують з  $V_{max}$  і концентрації ферменту, визначеної аналізом Бредфорда.  $K_m$  розраховують з початкових швидкостей реакції, отриманих для концентрацій гліфосату в діапазоні від 0,125 до 10мМ, з використанням трансформації Lineweaver-Burke рівняння Michaelis-Menten.  $k_{cat}/K_m$  визначають шляхом ділення величини, отриманої для  $k_{cat}$ , на величину, отриману для  $K_m$ .

З використанням цієї методики були отримані кінетичні параметри для ряду GAT-поліпептидів, описаних тут. Наприклад, визначені  $k_{cat}$ ,  $K_m$  і  $k_{cat}/K_m$  для GAT-поліпептиду, що відповідає SEQ ID NO:445, склали 322 $\text{хв.}^{-1}$ , 0,5мМ і 660мМ $^{-1}\text{хв.}^{-1}$ , відповідно, з використанням умов аналізу, описаних вище. Визначені  $k_{cat}$ ,  $K_m$  і  $k_{cat}/K_m$  для GAT-поліпептиду, що відповідає SEQ ID NO:447, склали 118 $\text{хв.}^{-1}$ , 0,1мМ і 1184мМ $^{-1}\text{хв.}^{-1}$ , відповідно, з використанням умов аналізу, описаних вище. Визначені  $k_{cat}$ ,  $K_m$  і  $k_{cat}/K_m$  для GAT-поліпептиду, що відповідає SEQ ID NO:300, склали 296 $\text{хв.}^{-1}$ , 0,65мМ і 465мМ $^{-1}\text{хв.}^{-1}$ , відповідно, з використан-



ням умов аналізу, описаних вище. Фахівець може використати ці величини для підтвердження того, що аналіз GAT-активності надає кінетичні параметри для GAT, придатні для порівняння з величинами, наведеними тут. Наприклад, умови, які використовуються для порівняння активності GATs, повинні забезпечувати ті самі кінетичні константи для SEQ ID NOS: 300, 445 і 457 (в ме-

жах нормальних експериментальних відхилень), що й описані тут, якщо ці умови будуть використані для порівняння тестових GAT з GAT-поліпептидами, переліченими тут. Кінетичні параметри для ряду варіантів GAT-поліпептидів були визначені за цією методикою і наведені в Табл.3, 4 і 5.

Таблиця 3. Величини  $k_{cat}$  GAT-поліпептидів

ID № послідовності	ID клону	$k_{cat}(min^{-1})$
SEQ ID NO:263	13_10F6	48.6
SEQ ID NO:264	13_12G6	52.1
SEQ ID NO:265	14_2A5	280.8
SEQ ID NO:266	14_2C1	133.4
SEQ ID NO:267	14_2F11	136.9
SEQ ID NO:268	CHIMERA	155.4
SEQ ID NO:269	10_12D7	77.3
SEQ ID NO:270	10_15F4	37.6
SEQ ID NO:271	10_17D1	176.2
SEQ ID NO:272	10_17F6	47.9
SEQ ID NO:273	10_18G9	24
SEQ ID NO:274	10_1H3	78.2
SEQ ID NO:275	10_20D10	86.2
SEQ ID NO:276	10_23F2	101.3
SEQ ID NO:277	10_2B8	108.4
SEQ ID NO:278	10_2C7	135
SEQ ID NO:279	10_3G5	87.4
SEQ ID NO:280	10_4H7	112
SEQ ID NO:281	10_6D11	62.4
SEQ ID NO:282	10_8C8	21.7
SEQ ID NO:283	11C3	2.8
SEQ ID NO:284	11G3	15.6
SEQ ID NO:285	11H3	1.2
SEQ ID NO:286	12_1F9	80.4
SEQ ID NO:287	12_2G9	151.4
SEQ ID NO:288	12_3F1	44.1
SEQ ID NO:289	12_5C10	89.6
SEQ ID NO:290	12_6A10	54.7
SEQ ID NO:291	12_6D1	49
SEQ ID NO:292	12_6F9	89.1
SEQ ID NO:293	12_6H6	90.5
SEQ ID NO:294	12_7D6	53.9
SEQ ID NO:295	12_7G11	234.5
SEQ ID NO:296	12F5	3.1
SEQ ID NO:297	12G7	2.3
SEQ ID NO:298	1_2H6	9.3
SEQ ID NO:299	13_12G12	36.1
SEQ ID NO:300	13_6D10	296.5
SEQ ID NO:301	13_7A7	117
SEQ ID NO:302	13_7B12	68.9
SEQ ID NO:303	13_7C1	48.1
SEQ ID NO:304	13_8G6	33.7
SEQ ID NO:305	13_9F6	59
SEQ ID NO:306	14_10C9	127
SEQ ID NO:307	14_10H3	105.2
SEQ ID NO:308	14_10H9	127.2

SEQ ID NO:309	14_11C2	108.7
SEQ ID NO:310	14_12D8	62.1
SEQ ID NO:311	14_12H6	91.1
SEQ ID NO:312	14_2B6	34.2
SEQ ID NO:313	14_2G11	69.4
SEQ ID NO:314	14_3B2	68.7
SEQ ID NO:315	14_4H8	198.8
SEQ ID NO:316	14_6A8	43.7
SEQ ID NO:317	14_6B10	134.7
SEQ ID NO:318	14_6D4	256
SEQ ID NO:319	14_7A11	197.2
SEQ ID NO:320	14_7A1	155.8
SEQ ID NO:321	14_7A9	245.9
SEQ ID NO:322	14_7G1	136.7
SEQ ID NO:323	14_7H9	64.4
SEQ ID NO:324	14_8F7	90.5
SEQ ID NO:325	15_10C2	69.9
SEQ ID NO:326	15_10D6	67.1
SEQ ID NO:327	15_11F9	76.4
SEQ ID NO:328	15_11H3	61.9
SEQ ID NO:329	15_12A8	77.1
SEQ ID NO:330	15_12D6	148.6
SEQ ID NO:331	15_12D8	59.7
SEQ ID NO:332	15_12D9	59.7
SEQ ID NO:333	15_3F10	48.7
SEQ ID NO:334	15_3G11	71.5
SEQ ID NO:335	15_4F11	80.3
SEQ ID NO:336	15_4H3	93.3
SEQ ID NO:337	15_6D3	85.9
SEQ ID NO:338	15_6G11	36.9
SEQ ID NO:339	15_8F6	59.6
SEQ ID NO:340	15F5	0.5
SEQ ID NO:341	16A1	10.4
SEQ ID NO:342	16H3	3.5
SEQ ID NO:343	17C12	3.2
SEQ ID NO:344	18D6	9.6
SEQ ID NO:345	19C6	2.2
SEQ ID NO:346	19D5	2.2
SEQ ID NO:347	20A12	2.8
SEQ ID NO:348	20F2	3.9
SEQ ID NO:349	2.10E+12	1.1
SEQ ID NO:350	23H11	7.1
SEQ ID NO:351	24C1	1.7
SEQ ID NO:352	24C6	2.7
SEQ ID NO:353	2.40E+08	8.9
SEQ ID NO:354	2_8C3	24.8
SEQ ID NO:355	2H3	16.1
SEQ ID NO:356	30G8	10.2
SEQ ID NO:357	3B_10C4	24.8
SEQ ID NO:358	3B_10G7	19.6
SEQ ID NO:359	3B_12B1	22.8
SEQ ID NO:360	3B_12D10	5.4
SEQ ID NO:361	3B_2E5	16.4
SEQ ID NO:362	3C_10H3	33.9
SEQ ID NO:363	3C_12H10	9.1
SEQ ID NO:364	3C_9H8	11.7
SEQ ID NO:365	4A_1B11	23.2
SEQ ID NO:366	4A_1C2	20.4

SEQ ID NO:367	4B_13E1	37.2
SEQ ID NO:368	4B_13G10	34.9
SEQ ID NO:369	4B_16E1	17
SEQ ID NO:370	4B_17A1	19.1
SEQ ID NO:371	4B_18F11	14.6
SEQ ID NO:372	4B_19C8	15.9
SEQ ID NO:373	4B_1G4	3.7
SEQ ID NO:374	4B_21C6	11.8
SEQ ID NO:375	4B_2H7	27
SEQ ID NO:376	4B_2H8	38.3
SEQ ID NO:377	4B_6D8	22.7
SEQ ID NO:378	4B_7E8	20.5
SEQ ID NO:379	4C_8C9	9
SEQ ID NO:380	4H1	1.3
SEQ ID NO:381	6_14D10	42.2
SEQ ID NO:382	6_15G7	48.4
SEQ ID NO:383	6_16A5	43.8
SEQ ID NO:384	6_16F5	35.2
SEQ ID NO:385	6_17C5	35.2
SEQ ID NO:386	6_18C7	32.2
SEQ ID NO:387	6_18D7	43
SEQ ID NO:388	6_19A10	86.8
SEQ ID NO:389	6_19B6	23.9
SEQ ID NO:390	6_19C3	23.1
SEQ ID NO:391	6_19C8	74.8
SEQ ID NO:392	6_20A7	40.4
SEQ ID NO:393	6_20A9	45.1
SEQ ID NO:394	6_20H5	19.5
SEQ ID NO:395	6_21F4	24.3
SEQ ID NO:396	6_22C9	47.4
SEQ ID NO:397	6_22D9	43.9
SEQ ID NO:398	6_22H9	17.4
SEQ ID NO:399	6_23H3	43.9
SEQ ID NO:400	6_23H7	46.2
SEQ ID NO:401	6_2H1	26.6
SEQ ID NO:402	6_3D6	41.7
SEQ ID NO:403	6_3G3	51.9
SEQ ID NO:404	6_3H2	57.2
SEQ ID NO:405	6_4A10	55
SEQ ID NO:406	6_4B1	27
SEQ ID NO:407	6_5D11	15.2
SEQ ID NO:408	6_5F11	40.1
SEQ ID NO:409	6_5G9	35.8
SEQ ID NO:410	6_6D5	55.3
SEQ ID NO:411	6_7D1	19.7
SEQ ID NO:412	6_8H3	44.7
SEQ ID NO:413	6_9G11	78.4
SEQ ID NO:414	6F1	10.1
SEQ ID NO:415	7_1C4	17.4
SEQ ID NO:416	7_2A10	14.5
SEQ ID NO:417	7_2A11	46.8
SEQ ID NO:418	7_2D7	54.9
SEQ ID NO:419	7_5C7	44.7
SEQ ID NO:420	7_9C9	65
SEQ ID NO:421	9_13F10	34.7
SEQ ID NO:422	9_13F1	31.6
SEQ ID NO:423	9_15D5	27.6
SEQ ID NO:424	9_15D8	107.3

SEQ ID NO:425	9_15H3	68.7
SEQ ID NO:426	9_18H2	25
SEQ ID NO:427	9_20F12	37.8
SEQ ID NO:428	9_21C8	28.6
SEQ ID NO:429	9_22B1	50.1
SEQ ID NO:430	9_23A10	21
SEQ ID NO:431	9_24F6	52.5
SEQ ID NO:432	9_4H10	101.3
SEQ ID NO:433	9_4H8	47.1
SEQ ID NO:434	9_8H1	74.8
SEQ ID NO:435	9_9H7	28
SEQ ID NO:436	9C6	13
SEQ ID NO:437	9H11	4
SEQ ID NO:438	0_4B10	190
SEQ ID NO:439	0_5B11	219
SEQ ID NO:440	0_5B3	143
SEQ ID NO:441	0_5B4	180
SEQ ID NO:442	0_5B8	143
SEQ ID NO:443	0_5C4	205
SEQ ID NO:444	0_5D11	224
SEQ ID NO:445	0_5D3	322
SEQ ID NO:446	0_5D7	244
SEQ ID NO:447	0_6B4	252
SEQ ID NO:448	0_6D10	111
SEQ ID NO:449	0_6D11	212
SEQ ID NO:450	0_6F2	175
SEQ ID NO:451	0_6H9	228
SEQ ID NO:452	10_4C10	69.6
SEQ ID NO:453	10_4D5	82.72
SEQ ID NO:454	10_4F2	231.04
SEQ ID NO:455	10_4F9	55.39
SEQ ID NO:456	10_4G5	176.65
SEQ ID NO:457	10_4H4	118.36
SEQ ID NO:458	11_3A11	55.66
SEQ ID NO:459	11_3B1	219.97
SEQ ID NO:460	11_3B5	194.61
SEQ ID NO:461	11_3C12	49.07
SEQ ID NO:462	11_3C3	214.02
SEQ ID NO:463	11_3C6	184.44
SEQ ID NO:464	11_3D6	55.3
SEQ ID NO:465	1_1G12	58.48
SEQ ID NO:466	1_1H1	291
SEQ ID NO:467	1_1H2	164
SEQ ID NO:468	1_1H5	94
SEQ ID NO:469	1_2A12	229
SEQ ID NO:470	1_2B6	138
SEQ ID NO:471	1_2C4	193
SEQ ID NO:472	1_2D2	124
SEQ ID NO:473	1_2D4	182
SEQ ID NO:474	1_2F8	161
SEQ ID NO:475	1_2H8	141
SEQ ID NO:476	1_3A2	181
SEQ ID NO:477	1_3D6	226
SEQ ID NO:478	1_3F3	167
SEQ ID NO:479	1_3H2	128
SEQ ID NO:480	1_4C5	254
SEQ ID NO:481	1_4D6	137
SEQ ID NO:482	1_4H1	236

SEQ ID NO:483	1_5H5	214
SEQ ID NO:484	1_6F12	209
SEQ ID NO:485	1_6H6	274
SEQ ID NO:486	3_11A10	135.41
SEQ ID NO:487	3_14F6	188.43
SEQ ID NO:488	3_15B2	104.13
SEQ ID NO:489	3_6A10	126.48
SEQ ID NO:490	3_6B1	263.08
SEQ ID NO:491	3_7F9	193.55
SEQ ID NO:492	3_8G11	99.14
SEQ ID NO:493	4_1B10	77.09
SEQ ID NO:494	5_2B3	56.75
SEQ ID NO:495	5_2D9	75.44
SEQ ID NO:496	5_2F10	54.72
SEQ ID NO:497	6_1A11	45.54
SEQ ID NO:498	6_1D5	42.92
SEQ ID NO:499	6_1F11	105.76
SEQ ID NO:500	6_1F1	69.81
SEQ ID NO:501	6_1H10	17.01
SEQ ID NO:502	6_1H4	85.91
SEQ ID NO:503	8_1F8	82.88
SEQ ID NO:504	8_1G2	67.47
SEQ ID NO:505	8_1G3	108.9
SEQ ID NO:506	8_1H7	101.24
SEQ ID NO:507	8_1H9	78.39
SEQ ID NO:508	GAT1_21F12	5.4
SEQ ID NO:509	GAT1_24G3	4.9
SEQ ID NO:510	GAT1_29G1	6.2
SEQ ID NO:511	GAT1_32G1	4.5
SEQ ID NO:512	GAT2_15G8	4.5
SEQ ID NO:513	GAT2_19H8	4.1
SEQ ID NO:514	GAT2_21F1	4.2

Таблиця 4 Величини  $K_m$  GAT поліпептиду (гліфосат )

ID № послідовності	ID клону	$K_m$ (mM)
SEQ ID NO:263	13_10F6	1.3
SEQ ID NO:264	13_12G6	1.2
SEQ ID NO:265	14_2A5	1.6
SEQ ID NO:266	14_2C1	3.1
SEQ ID NO:267	14_2F11	1.7
SEQ ID NO:268	CHIMERA	1.3
SEQ ID NO:269	10_12D7	1.8
SEQ ID NO:270	10_15F4	1
SEQ ID NO:271	10_17D1	2.2
SEQ ID NO:272	10_17F6	1.4
SEQ ID NO:273	10_18G9	1.2
SEQ ID NO:274	10_1H3	1.9
SEQ ID NO:275	10_20D10	1.6
SEQ ID NO:276	10_23F2	0.9
SEQ ID NO:277	10_2B8	1.1
SEQ ID NO:278	10_2C7	1.4
SEQ ID NO:279	10_3G5	2
SEQ ID NO:280	10_4H7	1.7
SEQ ID NO:281	10_6D11	1.2
SEQ ID NO:282	10_8C6	0.7
SEQ ID NO:283	11C3	3.1

SEQ ID NO:284	11G3	1.7
SEQ ID NO:285	11H3	1.4
SEQ ID NO:286	12_1F9	3
SEQ ID NO:287	12_2G9	1.5
SEQ ID NO:288	12_3F1	0.9
SEQ ID NO:289	12_5C10	1.5
SEQ ID NO:290	12_6A10	1.1
SEQ ID NO:291	12_6D1	1.2
SEQ ID NO:292	12_6F9	1.9
SEQ ID NO:293	12_6H6	1.6
SEQ ID NO:294	12_7D6	1.4
SEQ ID NO:295	12_7G11	2
SEQ ID NO:296	12F5	1.8
SEQ ID NO:297	12G7	3.7
SEQ ID NO:298	1_2H6	0.9
SEQ ID NO:299	13_12G12	0.69
SEQ ID NO:300	13_6D10	0.65
SEQ ID NO:301	13_7A7	0.5
SEQ ID NO:302	13_7B12	1.7
SEQ ID NO:303	13_7C1	1.5
SEQ ID NO:304	13_8G6	0.61
SEQ ID NO:305	13_9F6	1.3
SEQ ID NO:306	14_10C9	0.9
SEQ ID NO:307	14_10H3	0.6
SEQ ID NO:308	14_10H9	1.1
SEQ ID NO:309	14_11C2	1
SEQ ID NO:310	14_12D8	1
SEQ ID NO:311	14_12H6	0.9
SEQ ID NO:312	14_2B6	0.63
SEQ ID NO:313	14_2G11	1.4
SEQ ID NO:314	14_3B2	0.85
SEQ ID NO:315	14_4H8	2
SEQ ID NO:316	14_6A8	0.78
SEQ ID NO:317	14_6B10	1.4
SEQ ID NO:318	14_6D4	1
SEQ ID NO:319	14_7A11	3.7
SEQ ID NO:320	14_7A1	1.6
SEQ ID NO:321	14_7A9	3.2
SEQ ID NO:322	14_7G1	0.66
SEQ ID NO:323	14_7H9	1.3
SEQ ID NO:324	14_8F7	1.8
SEQ ID NO:325	15_10C2	0.8
SEQ ID NO:326	15_10D6	1
SEQ ID NO:327	15_11F9	1
SEQ ID NO:328	15_11H3	1
SEQ ID NO:329	15_12A8	1.6
SEQ ID NO:330	15_12D6	0.74
SEQ ID NO:331	15_12D8	1.3
SEQ ID NO:332	15_12D9	1.4
SEQ ID NO:333	15_3F10	0.9
SEQ ID NO:334	15_3G11	1.2
SEQ ID NO:335	15_4F11	0.9
SEQ ID NO:336	15_4H3	1
SEQ ID NO:337	15_6D3	1.4
SEQ ID NO:338	15_6G11	0.9
SEQ ID NO:339	15_9F6	1.1
SEQ ID NO:340	15F5	2.9
SEQ ID NO:341	16A1	2.9

SEQ ID NO:342	16H3	2.9
SEQ ID NO:343	17C12	1.4
SEQ ID NO:344	18D6	1.2
SEQ ID NO:345	19C6	1.1
SEQ ID NO:346	19D5	1.7
SEQ ID NO:347	20A12	1.1
SEQ ID NO:348	20F2	1.9
SEQ ID NO:349	2.10E+12	0.7
SEQ ID NO:350	23H11	2.2
SEQ ID NO:351	24C1	0.9
SEQ ID NO:352	24C6	1.3
SEQ ID NO:353	2.40E+08	0.9
SEQ ID NO:354	2_8C3	1.5
SEQ ID NO:355	2H3	0.9
SEQ ID NO:356	30G8	1.6
SEQ ID NO:357	3B_10C4	1.6
SEQ ID NO:358	3B_10G7	1
SEQ ID NO:359	3B_12B1	1.2
SEQ ID NO:360	3B_12D10	0.9
SEQ ID NO:361	3B_2E5	1.3
SEQ ID NO:362	3C_10H3	1.1
SEQ ID NO:363	3C_12H10	1.2
SEQ ID NO:364	3C_9H8	1
SEQ ID NO:365	4A_1B11	1.6
SEQ ID NO:366	4A_1C2	1.2
SEQ ID NO:367	4B_13E1	2
SEQ ID NO:368	4B_13G10	7.6
SEQ ID NO:369	4B_16E1	1
SEQ ID NO:370	4B_17A1	1.1
SEQ ID NO:371	4B_18F11	1.7
SEQ ID NO:372	4B_19C8	1.2
SEQ ID NO:373	4B_1G4	1
SEQ ID NO:374	4B_21C6	0.8
SEQ ID NO:375	4B_2H7	6.2
SEQ ID NO:376	4B_2H8	1.2
SEQ ID NO:377	4B_6D8	1.5
SEQ ID NO:378	4B_7E8	1.2
SEQ ID NO:379	4C_8C9	0.6
SEQ ID NO:380	4H1	1.4
SEQ ID NO:381	6_14D10	1.5
SEQ ID NO:382	6_15G7	1.3
SEQ ID NO:383	6_16A5	1.1
SEQ ID NO:384	6_16F5	1
SEQ ID NO:385	6_17C5	1.3
SEQ ID NO:386	6_18C7	1.2
SEQ ID NO:387	6_18D7	1.2
SEQ ID NO:388	6_19A10	1.9
SEQ ID NO:389	6_19B6	0.7
SEQ ID NO:390	6_19C3	1.4
SEQ ID NO:391	6_19C8	2
SEQ ID NO:392	6_20A7	1
SEQ ID NO:393	6_20A9	1.3
SEQ ID NO:394	6_20H5	0.8
SEQ ID NO:395	6_21F4	0.7
SEQ ID NO:396	6_22C9	3.2
SEQ ID NO:397	6_22D9	1.3
SEQ ID NO:398	6_22H9	1.1
SEQ ID NO:399	6_23H3	1.1

SEQ ID NO:400	6_23H7	1.2
SEQ ID NO:401	6_2H1	0.9
SEQ ID NO:402	6_3D6	1
SEQ ID NO:403	6_3G3	1
SEQ ID NO:404	6_3H2	1
SEQ ID NO:405	6_4A10	1.1
SEQ ID NO:406	6_4B1	1
SEQ ID NO:407	6_5D11	1
SEQ ID NO:408	6_5F11	1.9
SEQ ID NO:409	6_5G9	1.4
SEQ ID NO:410	6_6D5	1
SEQ ID NO:411	6_7D1	0.5
SEQ ID NO:412	6_8H3	1
SEQ ID NO:413	6_9G11	1.3
SEQ ID NO:414	6F1	1.8
SEQ ID NO:415	7_1C4	1.1
SEQ ID NO:416	7_2A10	0.8
SEQ ID NO:417	7_2A11	1.1
SEQ ID NO:418	7_2D7	1.1
SEQ ID NO:419	7_5C7	1
SEQ ID NO:420	7_9C9	1
SEQ ID NO:421	9_13F10	0.7
SEQ ID NO:422	9_13F1	1.1
SEQ ID NO:423	9_15D5	1.2
SEQ ID NO:424	9_15D8	1.1
SEQ ID NO:425	9_15H3	1.9
SEQ ID NO:426	9_18H2	1.1
SEQ ID NO:427	9_20F12	1
SEQ ID NO:428	9_21C8	1.2
SEQ ID NO:429	9_22B1	1.4
SEQ ID NO:430	9_23A10	1
SEQ ID NO:431	9_24F6	0.9
SEQ ID NO:432	9_4H10	1.5
SEQ ID NO:433	9_4H8	0.6
SEQ ID NO:434	9_8H1	1.7
SEQ ID NO:435	9_9H7	0.7
SEQ ID NO:436	9C6	2.5
SEQ ID NO:437	9H11	2.3
SEQ ID NO:438	0_4B10	0.68
SEQ ID NO:439	0_5B11	0.54
SEQ ID NO:440	0_5B3	0.39
SEQ ID NO:441	0_5B4	0.6
SEQ ID NO:442	0_5B8	0.27
SEQ ID NO:443	0_5C4	0.67
SEQ ID NO:444	0_5D11	0.67
SEQ ID NO:445	0_5D3	0.5
SEQ ID NO:446	0_5D7	1.1
SEQ ID NO:447	0_6B4	0.8
SEQ ID NO:448	0_6D10	0.1
SEQ ID NO:449	0_6D11	0.44
SEQ ID NO:450	0_6F2	0.34
SEQ ID NO:451	0_6H9	0.47
SEQ ID NO:452	10_4C10	0.1
SEQ ID NO:453	10_4D5	0.1
SEQ ID NO:454	10_4F2	0.2
SEQ ID NO:455	10_4F9	0.1
SEQ ID NO:456	10_4G5	0.58
SEQ ID NO:457	10_4H4	0.1



SEQ ID NO:458	11_3A11	0.1
SEQ ID NO:459	11_3B1	0.63
SEQ ID NO:460	11_3B5	0.26
SEQ ID NO:461	11_3C12	0.1
SEQ ID NO:462	11_3C3	0.22
SEQ ID NO:463	11_3C6	0.21
SEQ ID NO:464	11_3D6	0.1
SEQ ID NO:465	1_1G12	0.1
SEQ ID NO:466	1_1H1	1.8
SEQ ID NO:467	1_1H2	0.44
SEQ ID NO:468	1_1H5	1.5
SEQ ID NO:469	1_2A12	1.3
SEQ ID NO:470	1_2B6	0.58
SEQ ID NO:471	1_2C4	0.8
SEQ ID NO:472	1_2D2	1.2
SEQ ID NO:473	1_2D4	1.2
SEQ ID NO:474	1_2F8	1.9
SEQ ID NO:475	1_2H8	0.48
SEQ ID NO:476	1_3A2	0.8
SEQ ID NO:477	1_3D6	3.5
SEQ ID NO:478	1_3F3	1.5
SEQ ID NO:479	1_3H2	0.7
SEQ ID NO:480	1_4C5	0.93
SEQ ID NO:481	1_4D6	1.4
SEQ ID NO:482	1_4H1	1.2
SEQ ID NO:483	1_5H5	0.51
SEQ ID NO:484	1_6F12	14.7
SEQ ID NO:485	1_6H6	1.05
SEQ ID NO:486	3_11A10	0.17
SEQ ID NO:487	3_14F6	0.25
SEQ ID NO:488	3_15B2	0.1
SEQ ID NO:489	3_6A10	0.66
SEQ ID NO:490	3_6B1	0.43
SEQ ID NO:491	3_7F9	0.29
SEQ ID NO:492	3_8G11	0.1
SEQ ID NO:493	4_1B10	0.1
SEQ ID NO:494	5_2B3	0.1
SEQ ID NO:495	5_2D9	0.1
SEQ ID NO:496	5_2F10	0.1
SEQ ID NO:497	6_1A11	0.1
SEQ ID NO:498	6_1D5	0.1
SEQ ID NO:499	6_1F11	0.1
SEQ ID NO:500	6_1F1	0.1
SEQ ID NO:501	6_1H10	0.1
SEQ ID NO:502	6_1H4	0.1
SEQ ID NO:503	8_1F8	0.1
SEQ ID NO:504	8_1G2	0.1
SEQ ID NO:505	8_1G3	0.1
SEQ ID NO:506	8_1H7	0.1
SEQ ID NO:507	8_1H9	0.1
SEQ ID NO:508	GAT1_21F12	4.6
SEQ ID NO:509	GAT1_24G3	3.8
SEQ ID NO:510	GAT1_29G1	4
SEQ ID NO:511	GAT1_32G1	3.3
SEQ ID NO:512	GAT2_15G8	2.8
SEQ ID NO:513	GAT2_19H8	2.8
SEQ ID NO:514	GAT2_21F1	3

Таблиця 5 Величини  $K_{cat}/K_m$  GAT поліпептиду

ID № послідовності	ID клону	$K_{cat}/K_m (mM^{-1} min^{-1})$
SEQ ID NO:263	13_10F6	37.4
SEQ ID NO:264	13_12G6	43.4
SEQ ID NO:265	14_2A5	175.5
SEQ ID NO:266	14_2C1	43
SEQ ID NO:267	14_2F11	80.6
SEQ ID NO:268	CHIMERA	119.6
SEQ ID NO:269	10_12D7	43
SEQ ID NO:270	10_15F4	37.6
SEQ ID NO:271	10_17D1	80.1
SEQ ID NO:272	10_17F6	34.2
SEQ ID NO:273	10_18G9	20
SEQ ID NO:274	10_1H3	40.1
SEQ ID NO:275	10_20D10	53.9
SEQ ID NO:276	10_23F2	112.5
SEQ ID NO:277	10_2B8	98.5
SEQ ID NO:278	10_2C7	96.4
SEQ ID NO:279	10_3G5	43.7
SEQ ID NO:280	10_4H7	65.9
SEQ ID NO:281	10_6D11	52
SEQ ID NO:282	10_8C6	31
SEQ ID NO:283	11C3	0.9
SEQ ID NO:284	11G3	8.9
SEQ ID NO:285	11H3	0.9
SEQ ID NO:286	12_1F9	26.8
SEQ ID NO:287	12_2G9	101
SEQ ID NO:288	12_3F1	49
SEQ ID NO:289	12_5C10	59.7
SEQ ID NO:290	12_6A10	49.7
SEQ ID NO:291	12_6D1	40.8
SEQ ID NO:292	12_6F9	46.9
SEQ ID NO:293	12_6H6	56.5
SEQ ID NO:294	12_7D6	38.5
SEQ ID NO:295	12_7G11	117.2
SEQ ID NO:296	12F5	1.7
SEQ ID NO:297	12G7	0.6
SEQ ID NO:298	1_2H6	10.4
SEQ ID NO:299	13_12G12	52.4
SEQ ID NO:300	13_6D10	456.1
SEQ ID NO:301	13_7A7	234
SEQ ID NO:302	13_7B12	40.5
SEQ ID NO:303	13_7C1	32.1
SEQ ID NO:304	13_8G6	55.2
SEQ ID NO:305	13_9F6	45.3
SEQ ID NO:306	14_10C9	141.1
SEQ ID NO:307	14_10H3	175.3
SEQ ID NO:308	14_10H9	115.6
SEQ ID NO:309	14_11C2	108.7
SEQ ID NO:310	14_12D8	62.1
SEQ ID NO:311	14_12H6	101.3
SEQ ID NO:312	14_2B6	54.3
SEQ ID NO:313	14_2G11	49.6
SEQ ID NO:314	14_3B2	80.9

SEQ ID NO:315	14_4H8	99.4
SEQ ID NO:316	14_6A8	56
SEQ ID NO:317	14_6B10	96.2
SEQ ID NO:318	14_6D4	256
SEQ ID NO:319	14_7A11	53.3
SEQ ID NO:320	14_7A1	97.4
SEQ ID NO:321	14_7A9	76.9
SEQ ID NO:322	14_7G1	207.1
SEQ ID NO:323	14_7H9	49.5
SEQ ID NO:324	14_8F7	50.3
SEQ ID NO:325	15_10C2	87.3
SEQ ID NO:326	15_10D6	67.1
SEQ ID NO:327	15_11F9	76.4
SEQ ID NO:328	15_11H3	61.9
SEQ ID NO:329	15_12A8	48.2
SEQ ID NO:330	15_12D6	200.8
SEQ ID NO:331	15_12D8	45.9
SEQ ID NO:332	15_12D9	42.6
SEQ ID NO:333	15_3F10	54.1
SEQ ID NO:334	15_3G11	59.6
SEQ ID NO:335	15_4F11	89.2
SEQ ID NO:336	15_4H3	93.3
SEQ ID NO:337	15_6D3	61.3
SEQ ID NO:338	15_6G11	41
SEQ ID NO:339	15_9F6	54.2
SEQ ID NO:340	15F5	0.2
SEQ ID NO:341	16A1	3.6
SEQ ID NO:342	16H3	1.2
SEQ ID NO:343	17C12	2.3
SEQ ID NO:344	18D6	8
SEQ ID NO:345	19C6	2
SEQ ID NO:346	19D5	1.3
SEQ ID NO:347	20A12	2.5
SEQ ID NO:348	20F2	2
SEQ ID NO:349	2.10E+12	1.5
SEQ ID NO:350	23H11	3.2
SEQ ID NO:351	24C1	1.8
SEQ ID NO:352	24C6	2.1
SEQ ID NO:353	2.40E+08	9.8
SEQ ID NO:354	2_8C3	16.6
SEQ ID NO:355	2H3	17.7
SEQ ID NO:356	30G8	6.4
SEQ ID NO:357	3B_10C4	15.5
SEQ ID NO:358	3B_10G7	19.6
SEQ ID NO:359	3B_12B1	19
SEQ ID NO:360	3B_12D10	6
SEQ ID NO:361	3B_2E5	12.6
SEQ ID NO:362	3C_10H3	30.8
SEQ ID NO:363	3C_12H10	7.6
SEQ ID NO:364	3C_9H8	11.7
SEQ ID NO:365	4A_1B11	15
SEQ ID NO:366	4A_1C2	17
SEQ ID NO:367	4B_13E1	18.6
SEQ ID NO:368	4B_13G10	4.6
SEQ ID NO:369	4B_16E1	17
SEQ ID NO:370	4B_17A1	17.4
SEQ ID NO:371	4B_18F11	8.6
SEQ ID NO:372	4B_19C8	13.2

SEQ ID NO:373	4B_1G4	3.7
SEQ ID NO:374	4B_21C6	14.8
SEQ ID NO:375	4B_2H7	4.4
SEQ ID NO:376	4B_2H8	31.9
SEQ ID NO:377	4B_6D8	15.2
SEQ ID NO:378	4B_7E8	17.1
SEQ ID NO:379	4C_8C9	15.1
SEQ ID NO:380	4H1	0.9
SEQ ID NO:381	6_14D10	28.2
SEQ ID NO:382	6_15G7	37.3
SEQ ID NO:383	6_16A5	39.8
SEQ ID NO:384	6_16F5	35.2
SEQ ID NO:385	6_17C5	27.1
SEQ ID NO:386	6_18C7	26.8
SEQ ID NO:387	6_18D7	35.8
SEQ ID NO:388	6_19A10	45.7
SEQ ID NO:389	6_19B6	34.2
SEQ ID NO:390	6_19C3	16.5
SEQ ID NO:391	6_19C8	37.4
SEQ ID NO:392	6_20A7	40.4
SEQ ID NO:393	6_20A9	34.7
SEQ ID NO:394	6_20H5	24.3
SEQ ID NO:395	6_21F4	34.7
SEQ ID NO:396	6_22C9	14.8
SEQ ID NO:397	6_22D9	33.8
SEQ ID NO:398	6_22H9	15.9
SEQ ID NO:399	6_23H3	39.9
SEQ ID NO:400	6_23H7	38.5
SEQ ID NO:401	6_2H1	29.5
SEQ ID NO:402	6_3D6	41.7
SEQ ID NO:403	6_3G3	51.9
SEQ ID NO:404	6_3H2	57.2
SEQ ID NO:405	6_4A10	50
SEQ ID NO:406	6_4B1	27
SEQ ID NO:407	6_5D11	15.2
SEQ ID NO:408	6_5F11	21.1
SEQ ID NO:409	6_5G9	25.6
SEQ ID NO:410	6_6D5	55.3
SEQ ID NO:411	6_7D1	39.5
SEQ ID NO:412	6_8H3	44.7
SEQ ID NO:413	6_9G11	60.3
SEQ ID NO:414	6F1	5.6
SEQ ID NO:415	7_1C4	15.9
SEQ ID NO:416	7_2A10	18.2
SEQ ID NO:417	7_2A11	42.6
SEQ ID NO:418	7_2D7	49.9
SEQ ID NO:419	7_5C7	44.7
SEQ ID NO:420	7_9C9	65
SEQ ID NO:421	9_13F10	49.6
SEQ ID NO:422	9_13F1	28.7
SEQ ID NO:423	9_15D5	23
SEQ ID NO:424	9_15D8	97.6
SEQ ID NO:425	9_15H3	36.2
SEQ ID NO:426	9_18H2	22.7
SEQ ID NO:427	9_20F12	37.8
SEQ ID NO:428	9_21C8	23.8
SEQ ID NO:429	9_22B1	35.8
SEQ ID NO:430	9_23A10	21

SEQ ID NO:431	9_24F6	58.3
SEQ ID NO:432	9_4H10	67.5
SEQ ID NO:433	9_4H8	78.5
SEQ ID NO:434	9_8H1	44
SEQ ID NO:435	9_9H7	40
SEQ ID NO:436	9C6	5.1
SEQ ID NO:437	9H11	1.7
SEQ ID NO:438	0_4B10	279
SEQ ID NO:439	0_5B11	406
SEQ ID NO:440	0_5B3	367
SEQ ID NO:441	0_5B4	301
SEQ ID NO:442	0_5B8	522
SEQ ID NO:443	0_5C4	306
SEQ ID NO:444	0_5D11	334
SEQ ID NO:445	0_5D3	660
SEQ ID NO:446	0_5D7	222
SEQ ID NO:447	0_6B4	315
SEQ ID NO:448	0_6D10	1177
SEQ ID NO:449	0_6D11	481
SEQ ID NO:450	0_6F2	516
SEQ ID NO:451	0_6H9	486
SEQ ID NO:452	10_4C10	695.98
SEQ ID NO:453	10_4D5	827.16
SEQ ID NO:454	10_4F2	1155.19
SEQ ID NO:455	10_4F9	553.93
SEQ ID NO:456	10_4G5	304.57
SEQ ID NO:457	10_4H4	1183.6
SEQ ID NO:458	11_3A11	556.62
SEQ ID NO:459	11_3B1	349.17
SEQ ID NO:460	11_3B5	748.49
SEQ ID NO:461	11_3C12	490.67
SEQ ID NO:462	11_3C3	972.81
SEQ ID NO:463	11_3C6	878.27
SEQ ID NO:464	11_3D6	553.01
SEQ ID NO:465	1_1G12	584.79
SEQ ID NO:466	1_1H1	162
SEQ ID NO:467	1_1H2	366
SEQ ID NO:468	1_1H5	63
SEQ ID NO:469	1_2A12	176
SEQ ID NO:470	1_2B6	239
SEQ ID NO:471	1_2C4	242
SEQ ID NO:472	1_2D2	104
SEQ ID NO:473	1_2D4	152
SEQ ID NO:474	1_2F8	85
SEQ ID NO:475	1_2H8	294
SEQ ID NO:476	1_3A2	227
SEQ ID NO:477	1_3D6	64
SEQ ID NO:478	1_3F3	112
SEQ ID NO:479	1_3H2	183
SEQ ID NO:480	1_4C5	273
SEQ ID NO:481	1_4D6	98
SEQ ID NO:482	1_4H1	196
SEQ ID NO:483	1_5H5	419
SEQ ID NO:484	1_6F12	14
SEQ ID NO:485	1_6H6	259
SEQ ID NO:486	3_11A10	796.55
SEQ ID NO:487	3_14F6	753.73
SEQ ID NO:488	3_15B2	1041.32

SEQ ID NO:489	3_6A10	191.64
SEQ ID NO:490	3_6B1	611.81
SEQ ID NO:491	3_7F9	667.4
SEQ ID NO:492	3_8G11	991.44
SEQ ID NO:493	4_1B10	770.91
SEQ ID NO:494	5_2B3	567.5
SEQ ID NO:495	5_2D9	754.36
SEQ ID NO:496	5_2F10	547.22
SEQ ID NO:497	6_1A11	455.41
SEQ ID NO:498	6_1D5	429.16
SEQ ID NO:499	6_1F11	1057.6
SEQ ID NO:500	6_1F1	698.15
SEQ ID NO:501	6_1H10	170.11
SEQ ID NO:502	6_1H4	859.12
SEQ ID NO:503	8_1F8	828.78
SEQ ID NO:504	8_1G2	674.73
SEQ ID NO:505	8_1G3	1088.97
SEQ ID NO:506	8_1H7	1012.4
SEQ ID NO:507	8_1H9	783.89
SEQ ID NO:508	GAT1_21F12	1.2
SEQ ID NO:509	GAT1_24G3	1.3
SEQ ID NO:510	GAT1_29G1	1.5
SEQ ID NO:511	GAT1_32G1	1.4
SEQ ID NO:512	GAT2_15G8	1.6
SEQ ID NO:513	GAT2_19H8	1.5
SEQ ID NO:514	GAT2_21F1	1.4

$K_m$  для АсСоА вимірюють з використанням мас-спектрометричного способу з многократним відбором зразків під час реакції. АсСоА і гліфосат (амонієву сіль) вміщують у вигляді 50-кратно сконцентрованих базових розчинів в лунку мас-спектрометричної пластини для зразків. Реакції ініціюють додаванням ферменту, розбавленого відповідним чином у легкому буфері, такому як морфоліну ацетат або амонію карбонат, рН 6,8 або 7,7. Зразок многократно інжектують в прилад і розраховують початкові швидкості з графіків площини піку від часу утримання.  $K_m$  розраховують, як для гліфосату.

Приклад 8: Селекція трансформованих *E. coli*

Виведений ген *gat* (химера з нативним сайтом зв'язування рибосом в *Licheniformis* (AACTGAAGGAGGAATCTC, SEQ ID NO:515), приєднаним безпосередньо до 5'-закінчення GAT-кодуючої послідовності) клонували в вектор експресії pQE80 (Qiagen) між сайтами EcoRI і HindIII, з отриманням плазмиди pMAXY2190 (Фіг.11). В результаті цього з плазмиди був видалений домен His-таг і залишився ген В-лактамази, що надає стійкість до антибіотиків ампіциліну і карбеніциліну. pMAXY2190 піддали електропорації (BioRad

Gene Pulser) в клітини *E. coli* XL1 Blue (Stratagene). Клітини пересуспендували в збагаченому середовищі SOC і залишили для регенерації на 1 годину. Потім клітини обережно пелетизували, промили один раз мінімальним середовищем M9 з нестачею амінокислот (12,8г/л  $\text{Mg}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$ , 3,0г/л  $\text{KH}_2\text{PO}_4$ , 0,5г/л NaCl, 1,0г/л  $\text{NH}_4\text{Cl}$ , 0,4% глюкози, 2мМ  $\text{MgSO}_4$ , 0,1мМ  $\text{CaCl}_2$ , 10мг/л тіаміну, 10мг/л проліну, 30мг/л карбеніциліну) і пересуспендували в 20мл того ж середовища M9. Після вирощування протягом ночі при 37°C і 250об/хв. рівні об'єми клітин вміщували на пластини з середовищем M9 або M9 + 1мМ гліфосату. Вектор pQE80 без гена *gat* подібним чином вводили в клітини *E. coli* і вміщували на пластини для отримання окремих колоній для порівняння. Результати наведені в Табл.6 і ясно показують, що GAT-активність дозволяє селекцію і ріст трансформованих клітин *E. coli* з фоном менше 1%. Слід зазначити, що індукція IPTG була необхідна для достатньої GAT-активності, щоб дозволити ріст трансформованих клітин. Трансформацію підтверджували шляхом переізоляції pMAXY2190 з клітин *E. coli*, вирощуваних в присутності гліфосату.

Таблиця 6. Гліфосатна селекція pMAXY2190 в *E. coli*

Плазміда	Число колоній	
	M9 без гліфосату	M9 + 1 мМ гліфосату
pMAXY2190	568	512
pQE80	324	3

Приклад 9: Селекція трансформованих рослинних клітин

*Agrobacterium*-опосереднена трансформація рослинних клітин відбувається з низькою ефективністю. Щоб дозволити розмноження трансформованих клітин при одночасному інгібуванні проліферації нетрансформованих клітин, необхідний маркер, що селектується. Антибіотичні маркери

для канаміцину і гідроміцину і гербіцид-модифікуючий ген *bag*, який детоксифікує гербіцидну сполуку фосфінотрицин, є прикладами маркерів, що селектуються, які використовують в рослинах (Methods in Molecular Biology, 1995, 49:9-18). Ми демонструємо тут, що GAT-активність служить у якості ефективного маркера, що селектується, для трансформації рослин. Виведений ген *gat*

(0\_5B8) клонували між рослинним промотором (посилений вірус облямування жилок суниці) і термінатором убіхінону і ввели в Т-ДНК-область бінарного вектора pMAXY3793, придатного для трансформації рослинних клітин за допомогою *Agrobacterium tumefaciens* EHA105, як показано на Фіг.12. Скринабельний маркер GUS був присутній в Т-ДНК для можливості підтвердження трансформації. Паростки трансгенного тютюну генерували з використанням гліфосату у якості єдиного агента селекції.

Зародкова брунька *Nicotiana tabacum* L. Xanthi субкультивували в середовищі MS половинної сили з сахарозою (1,5%) і Gerlite (0,3%) під дією світла протягом 16 годин (35-42 рЕйнштейн м<sup>-2</sup>с<sup>-1</sup>, білі флуоресцентні лампи з холодним катодом) при 24°C кожні 2-3 тижні. Молоде листя зрізали з рослин через 2-3 тижні субкультуваці і розрізали на сегменти 3×3mm. *A. tumefaciens* EHA105 інокулювали в середовище і вирощували протягом ночі до щільності A600 = 1,0. Клітини пелетизували при 4000об/хв. протягом 5 хвилин і пересушували в 3 об'ємах рідкого кокультуваційного середовища, скомбінованого з Murashige і Skoog (середовище MS) (pH 5,2) з 2мг/л N<sub>6</sub>-бензиладеніну (BA), 1% глюкози і 400мкМ ацетосирінгону. Тоді шматочки листя повністю занурили в 20мл *A. tumefaciens* в

100×25-мм чашках Петрі на 30 хвилин, промкнули стерилізованим фільтрувальним папером, вмістили на тверде кокультуваційне середовище (0,3% Gerlite) і інкубували, як описано вище. Через 3 доби кокультуваці 20-30 сегментів перенесли в середовище базальної індукції паростків (BSI), що містило тверде середовище MS (pH 5,7) з 2мг/л BA, 3% сахарози, 0,3% Gerlite, 0-200мкМ гліфосату і 400мкг/мл Timentin.

Через 3 тижні паростки були виразно помітними на експлантатах, вміщених на середовище без гліфосату, незважаючи на присутність або відсутність гена *gat*. Т-ДНК-перенос з обох конструкцій був підтверджений гістохімічним GUS-забарвленням листя з регенованих паростків. Концентрації гліфосату, більші ніж 20мкМ, повністю інгібували утворення паростків на експлантатах без гена *gat*. Експлантати, інфіковані *A. tumefaciens* з конструкцією гена *gat*, регенерували паростки при концентраціях гліфосату до 200мкМ (найвищий рівень, що аналізувався). Трансформацію підтверджували гістохімічним GUS-забарвленням і фрагментарною ПЛР-ампліфікацією гена *gat* з використанням відпалу праймерів для промотора і 3'-областей. Результати наведені в Табл.7.

Таблиця 7. Регенерація паростків тютюну з селекцією гліфосатом

Перенесені гени	Концентрація гліфосату % регенерації паростків				
	0 мкМ	20 мкМ	40 мкМ	80 мкМ	200 мкМ
GUS	100	0	0	0	0
gat і GUS	100	60	30	5	3

Приклад 10: Селекція гліфосатом трансформованих дріжджових клітин

Маркери селекції для трансформації дріжджів звичайно являють собою ауксотрофні гени, що дозволяють ріст трансформованих клітин в середовищі з нестачею специфічної амінокислоти або нуклеотиду. Оскільки *Saccharomyces cerevisiae* чутливий до гліфосату, GAT також може бути використана у якості маркера, що селектується. Щоб продемонструвати це, виведений ген *gat* (0\_6D10) клонують з Т-ДНК-вектора pMAXY3793 (як показано в Прикладі 9) у вигляді фрагмента PstI-ClaI, що містить повну кодуючу область, і лігують в PstI-ClaI-розщеплений p424TEF (Gene, 1995, 156:119-122), як показано на Фіг.13. Ця плазміда містить джерело реплікації *E.coli* і ген, який надає стійкість до карбеніциліну, а також TRP1, триптофан-ауксотрофний маркер, що селектується, для трансформації дріжджів.

Конструкцію, що містить *gat*, трансформують в *E.coli* XL1 Blue (Stratagene) і вміщують на пластини в агарово середовищі LB з карбеніциліном (50мкг/мл). Плазмідну ДНК отримують і застосовують для трансформації дріжджового штаму YPH499 (Stratagene) з використанням комплексу для трансформації (Bio101). Рівні кількості трансформованих клітин вміщують на пластини в CSM-YNB-глюкозне середовище (Bio101) без усіх ароматичних амінокислот (триптофану, тирозину і

фенілаланіну) з додаванням гліфосату. Для порівняння в YPH499 також вводять p424TEF, що не містить гена *gat*, і вміщують на пластини, як описано вище. Результати показують, що GAT-активність функціонує як ефективний маркер, що селектується. Присутність вектора, що містить *gat*, в селектованих гліфосатом колоніях може бути підтверджена переізоляцією плазміди і аналізом рестрикційного розщеплення.

Хоча вищевикладений винахід був описаний для кращого пояснення в деяких деталях, фахівцям буде зрозуміло з цього опису, що численні зміни в формі і подробицях можуть бути проведені в рамках дійсної області винаходу. Наприклад, всі технології, способи, композиції, апаратура і системи, описані вище, можуть бути застосовані в різних комбінаціях. Винахід включає всі способи і реагенти, описані тут, а також всі полінуклеотиди, поліпептиди, клітини, організми, рослини, рослинні культури і т.д., що є продуктами цих нових способів і реагентів.

Всі публікації, патенти, патентні заявки або інші документи, що цитуються в цій заявці, включені в неї у вигляді посилання у всій їх повноті для будь-якого призначення у тій же мірі, як це було б, якби кожна окрема публікація, патент, патентна заявка або інший документ були окремо включені у вигляді посилання для будь-якого призначення.

ID № послідовності	ID клону	Послідовність
SEQ ID NO:1	ST401 gat	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTTGCTCGGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATAAAGCCGAACATTTCAGAGCTT GAAGGCGAAGAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGATACCGTGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG GGGCAGACCTTTTATGGTGCAATGCCAGGACATCTGTG AGCGGCTACTATGAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:2	B6 gat	ATGATTGAAGTCAAACCTATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGC ACGTTTCACCTCGGCGGATATTATCGGGACAGGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATTTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCAGACCTTTTATGGTGCAACGCCAGGACATCTGTGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAAGG GGGGTCTACGATATACCGCCGATCGGACCTCATATTTT ATGTATAAGAAATTGACATAA
SEQ ID NO:3	DS3	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATAATGCCGAACATTTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGATATCTGTG AGCGGCTACTATGAAAAGCTCGGCTTCAGCGAACAAAGG CGGGATCTACGACATACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGGCATAA
SEQ ID NO:4	NHA-2	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATAATGCCGAACATTTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGATATCTGTG AGCGGCTACTATGAAAAGCTCGGCTTCAGCGAACAAAGG CGGGATCTACGACATACCGCCGATCGGACCTCATATTTT



		GATGTATAAGAAATTGGCATAA
SEQ ID NO:5	NH5-2	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGT GCGTTTCACCTCGGTGGATATTACCAGGGCAAGCTGATC AGCATCGCTTCCTTTCATAAAGCCGAACATTGAGAGCTT GAGGGCGAAGAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGATACCGTGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG GGGCAGACCTTTTATGGTGCAATGCCAGGACATCTGTG AGCGGCTACTATGAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGATCGGACCTCATATTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:6	ST401 GAT	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGAFH LGGYYRGKLISIASFHKAHSELEGEEQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYEK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:7	B6 GAT	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRDLISIASFHQAHSSELEGOKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYKK LGFSEQGGVYDIPPIGPHILMYKKLT
SEQ ID NO:8	DS3 GAT	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHNAHSELEGOKQYQLRGMAT TLEGYREQKAGSTLIRHAEELLRKKGADLLWCNARISVSG YYEKLGFSEQGGIYDIPPIGPHILMYKKLA
SEQ ID NO:9	NHA-2 GAT	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHNAHSELEGOKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARISVSGYYEKL GLSEQGGIYDIPPIGPHILMYKKLA
SEQ ID NO:10	NH5-2 GAT	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGAFH LGGYYQGKLISIASFHKAHSELEGEEQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYEK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:11	13_10F6	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGT ACGTTTCACCTCGGTGGATATTACCAGGGGCAAGCTGATC AGCATCGCTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTCGAAGGATACCGTGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGACGTCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGCTCGGACCTCATATTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:12	13_12G6	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTCGGTGGATATTACCAGGGGCAAGCTGGTC AGCATCGCTCCTTTCATCAAGCCGAACATCCAGAGCTT

		GAAGGCCAAAGACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGGG GGCAGACCTCTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGACTGGGCCCCATATTTG ATGTATAAGAAATTGACATAA
SEQ ID NO:13	14_2A5	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGAGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTAATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGTGAGCAAAAAGCGGGGAAGCAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGACGTCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACACACCGCCGGTTCGGACCTCATATTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:14	14_2C1	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGGTC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTCGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACACACCGCCGACTGGGCCCCATATTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:15	14_2F11	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTTGGTGGATATTACCGGGGCAAGCTGGTC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTCGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGCGCTTCTTCGGAAAAAGG GGCAGACCTCTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACACACCGCCGGCCGACCCCATATTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:16	CHIMERA	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATTTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGCGAGCAAAAAGCGGGCAGTACG

		CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACACACCGCCGGTCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:17	10_12D7	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGNATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGC ACGCTTCACCTCGGTGGATATTACCGGGGCAAGCTGAT CAGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGCA CGCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAG GGGGCAGACCTCTTATGGTGCAACGCCAGGACATCTGC GAGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAG GCGAAGTCTACGACATACCGCCGACCGGACCCCATATT TTGATGTATAAGAAATTGACGTAA
SEQ ID NO:18	10_15F4	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTTGCTCAGGGGT ACGTTTCACCTCGGTGGGTATTACCGGGGCAAGCTGGTC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAGG CGGGGTCTACGACATACCGCCGGTCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:19	10_17D1	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGGG CGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAGGC GAAGTCTACGACACACCGCCGGTCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:20	10_17F6	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGGTC AGCATCGCTTCCTTTCATCAAGCCGAACATTAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGCAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG GCGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG

		AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGGTTCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:21	10_18G9	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACTGATTTGCTCGGTGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGGTC AGCATCGCTTCCTTTCATCAAGCCGAACATTGAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG GGGCAGACCTCTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAAGG CGGGGTCTACGACATACCGCCGGTTCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:22	10_1H3	ATGATTGAAGTCAAACCAATAAACGCGGAGGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTATCGGGGCAAGCTGGTC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCGAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGACCGGACCCCATATTTT GATGTATAAGAAATTGACATAA
SEQ ID NO:23	10_20D10	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTTGCTCGGGGGC ACGCTTCACCTCGGTGGATATTACCGGGGCAAGCTGAT CAGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTTGAAGAGTACCGCGAGCAAAAAGCGGGCAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGCTACTATAAAAAGCTCGGCTTCAGCGAACAAAGG CGGGGTCTACGACATACCGCCGGTTCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:24	10_23F2	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATACTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGGTC AGCATCGCTTCCTTTCATCAAGCCGAACACCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTCGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG GGCAGACCTCTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACACACCGCCGGTTCGGACCTCATATTTT

		ATGTATAAGAAATTGACGTAA
SEQ ID NO:25	10_2B8	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGGG GGCAGACCTCTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACACACCGCCGGTCGGACCTCATATTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:26	10_2C7	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGAT CAGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTCGAAGGGTACCGTGAGCAAAAAGCGGGAAGCA CGCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAA GGCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGC GAGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGG GCGAAGTCTACGACACACCGCCGGTCGGACCTCATATT TTGATGTATAAGAAATTGACGTAA
SEQ ID NO:27	10_3G5	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGGTC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGACCGGACCCCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:28	10_4H7	ATGATTGAAGTCAAACCGATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGGTC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGACCGGACCCCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID	10_6D11	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA

NO:29		TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGC ACGCTTCACCTCGGTGGATATTACCGGGGCAAGCTGGT CAGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCG ACGCTTGAAGGGTACCGTGAGCAAAAAGCGGGCAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGGTCCGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:30	10_8C6	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCGA CACTCGAAGGATACCGCGAGCAAAAAGCGGGAAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAAG CGGGGTCTACGACATACCGCCGGTCCGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:31	11C3	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCAGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATTAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCGA CGCTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGTG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAAG CGGGGTCTACGATATACCGCCGATCCGACCTCATATTTT GATGTATAAGAAATTGACATAA
SEQ ID NO:32	11G3	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCGTGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGCGGATATTACCAGGGCAAGCTGAT CAGCATCGCTTCCTTTCATCAAGCCGAACATTAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCG ACGCTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGCTACTATGAAAAGCTCGGCTTCAGCGAACAAAG CGGGGTCTACGATATACCGCCGATCCGACCTCATATTTT GATGTATAAGAAATTGGCATAA
SEQ ID NO:33	11H3	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATACTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGT

		GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACACCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGCGCTTCTTCGGA AAAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCA ACTGGGCCCCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:34	12_1F9	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATACTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGA AAAAAGG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCCACGACATACCGCCGACCGGACCCCATATTT TGATGTATAAGAAATTGACGTAA
SEQ ID NO:35	12_2G9	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGT ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGGT CAGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGA AAAAAGG GGGCAGACCTCTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACACACCGCCGGTCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:36	12_3F1	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTTGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTCGAAGGATACCGTGAGCAAAAAGCGGGAAGTAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGA AAAAAGG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAAG CGGGGTCTACGACATACCGCCGGTCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:37	12_5C10	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTATCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT

		GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACGCACCGCCGACCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:38	12_6A10	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGGT CAGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAAGG CGGGGTCTACGACATACCGCCGGTTCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:39	12_6D1	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAAGG CGGGGTCTACGACATACCGCCTGTTCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:40	12_6F9	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTAAGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTCGAAGGATACCGCGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGCTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGACCGGACCCCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:41	12_6H6	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATACTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGGTC AGCATCGCCTCCTTTCACCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGTGAGCAAAAAGCGGGCAGTACG



		CTTATCCGCCATGCCGAAGCGCTTCTTCGGAAAAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGACCGGACCCCATATTTT GATGTATAAGAAATTGACATAA
SEQ ID NO:42	12_7D6	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACTGATTTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GGGGTCTACGACATACCGCCGACCGGACCCCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:43	12_7G11	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATTTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACACACCGCCGGTCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:44	12F5	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCAGGGCAAGCTGATC AGCATCGCTTCCTTTCATAAAGCCGAACATTTCAGAGCTT GAGGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGGG GGCAGACCTTTTATGGTGCAATGCCAGGACATCTGTGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GGGATCTACGACATACCGCCGATCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:45	12G7	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCAGGGCAAGCTGATC AGCATCGCTTCCTTTCATAAAGCCGAACATTTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGATACCGTGAGCAAAAAGCGGGAAGCAC ACTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCAGACCTTTTATGGTGCAACGCCAGGACATCTGTG

		AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:46	1_2H6	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGTATGTATGAAACCGATTGCTCGGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAAGG CGGGGTCTACGACATACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:47	13_12G12	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTAATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG GCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGGTCGGACCTCATATTTT GATGCATAAGAAATTGACGTAA
SEQ ID NO:48	13_6D10	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGAGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTAATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTCGAAGGGTACCGTGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG GGGCAGACCTCTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACACACCGCCGGTCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:49	13_7A7	ATGATCGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCAGGAGT GCGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGAT CAGCATCGCCTCCTTTCACCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGGGGGATGGCG ACACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGTA CGCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAG GGGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGC GAGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGG GCGAAGTCTACGACACACCGCCGGTCGGACCTCATATT

		TTGATGTATAAGAAATTGACGTAA
SEQ ID NO:50	13_7B12	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGAGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTCGAAGGATACCGCGAGCAAAAAGCGGGCAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCGGACCTTTTGTGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGACTGGGCCCCATATTTT GATGTATAAGAAGTTGACGTAA
SEQ ID NO:51	13_7C1	ATGATTGAAGTCAAACCAATAAATGCGGAAGATACGTA TGAGATCAGGCACCGCATACTCCGGCCGAATCAGCCGC TTGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAACTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGCGGGTAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCGGACCTTTTGTGGTGCAACGCCAGGACATCTGCGA GAGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAAGJC GAAGTCTACGACATACCGCCGACTGGGCCCCATATTTTG ATGTATAAAGAAATTGACGTAA
SEQ ID NO:52	13_8G6	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGAT CAGCATCGCTTCCTTTAATCAAGCCGAACATCCAGAGCT TGAAGGTCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGACGTCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAAGGC GGGGTCTACGACATACCGCCGGTCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:53	13_9F6	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATCTGCTTGGGGGC ACGTTTCACCTAGGTGGATATTACCGGGGCAAGCTGAT CAGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGTA CGCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GGGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGC GAGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGG GCGAAGTCTACGACATACCGCCGGTCGGACCTCATATTT TGATGTATAAGAAATTGACGTAA
SEQ ID	14_10C9	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA

NO:54		TGAGATCAGGCACCGCATACTCCGGCCGAATCAGCCGC TAGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCTGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGTAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG GGGCAGACCTTTTATGGTGCAACGCCAGGACGTCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACACACCGCCGGTCGGACCTCATATTTT GATGTATAAGAAGTTGACGTAA
SEQ ID NO:55	14_10H3	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGGT CAGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGCA CGCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAA GGCGCAGACCTTTTATGGTGCAACGCCAGGACATCTGC GAGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGG GCGAAGTCTACGACACACCGCCGGTCGGACCTCATATT TTGATGTATAAGAAGTTGACGTAA
SEQ ID NO:56	14_10H9	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATACTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGGTC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAAGG CGCGGACCTTTTGTGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACACACCGCCGGTCGGACCTCATATTTTG ATGTATAAGAATTGACATAA
SEQ ID NO:57	14_11C2	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGAGC ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGGT CAGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTTGAAGAGTACCGCGAGCAAAAAGCGGGCAGTAC GCTTATCCGCCATGCCGAAGCGCTTCTTCGGAAAAAGG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACACACCGCCGACCGGACCCCATATTT TGATGTATAAGAATTGACGTAA
SEQ ID NO:58	14_12D8	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTAAGTATGAAACCGATTGCTCGGGGGT

		ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGGT CAGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTTGAAGGATACCGTGAGCAAAAAGCTGGCAGTAC GCTTATCCGCCATGCCGAAGCGCTTCTTCGGAAAAAAG GCGCGGACCTTTTGTGGTGCAACGCCAGGACATCTGCG AGCGGCTACTATAAAAAGCTCGGCTTCAGGGAACAAGG CGGGGTCTACGACATACCGCCTGTTCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:59	14_12H6	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCGGACCTTTTGTGGTGCAACGCCAGGACGTCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGACTGGGCCCCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:60	14_2B6	ATGATTGAAGTCAAACCAATAAATGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTAATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTCGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGACGTCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GGGGTCTACGACATACCGCCGGTTCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:61	14_2G11	ATGATTGAAGTCAAACCAATAAATGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCAGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGGTC AGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTCGAAGGGTACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCGGACCTTTTGTGGTGCAACGCCAGGACATCTGCGA GTGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGACTGGGCCCCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:62	14_3B2	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCAGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCAGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGGTC AGCATCGCCTCCTTTCATCAGGCCGAACATCCAGAGCTT

		GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGCGGGAAGCAC GCTTATCCGCCATGCCGAAGCGCTTCTTCGGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAAGG CGGGGTCTACGACATACCGCCGGCCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:63	14_4H8	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGAGC ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGAT CAGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTCGAAGGGTACCGTGAGCAAAAAGCGGGAAGCA CGCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAA GGCGCGGACCTTTTGTGGTGCAACGCCAGGACGTCTGC GAGCGGCTACTATAAAAAGCTCGGCTTCAGCGAACAGG GCGAAGTCTACGACACACCGCCGGTCCGGACCTCATATT TTGATGTATAAGAAATTGACGTAA
SEQ ID NO:64	14_6A8	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTAGTC AGCATCGCTTCCTTTAATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCGGACCTTTTGTGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACACACCGCCGGTCCGGACCTCATGTTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:65	14_6B10	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTTGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTCGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAAGGC GGGGTCTACGACATGCCGCCGGTCCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:66	14_6D4	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGACCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGAGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTAATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG

		CTTATCCGCCATGCCGAAGCGCTTCTTCGGAAAAAGGG GGCAGACCTCTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACACACCGCCGGTCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:67	14_7A11	ATGATTGAAGTCAAACCAATAAACGCGGAGGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGGTC AGCATCGCCTCCTTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCTAAAACAGTATCAGCTGAGAGGGATGGCGAC ACTCGAAGGGTACCGTGAGCAAAAAGCGGGAAGTACG CTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGGG GGCAGACCTCTTATGGTGCAACGCCAGGACGTCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACACACCGCCGACCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:68	14_7A1	ATGATTGAAGTCAAACCAATAAACGCGGAGGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGGTC AGCATCGCCTCCTTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCTAAAACAGTATCAGCTGAGAGGGATGGCGAC ACTCGAAGGGTACCGTGAGCAAAAAGCGGGAAGTACG CTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGGG GGCAGACCTCTTATGGTGCAACGCCAGGACGTCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACACACCGCCGACCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:69	14_7A9	ATGATTGAAGTCAAACCAATAAACGCGGAGGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGT ACGTTTCACCTCGGCGGATATTACCGGGGCAAGTTGGTC AGCATCGCCTCCTTTTCATCAAGCCAAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTCGAAGGGTACCGTGAGCAAAAAGCGGGTAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCGACCTTTTATGGTGCAACGCCAGGACGTCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACACACCGCCGGTCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:70	14_7G1	ATGATTGAAGTCAAACCAATAAACGCGAGAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTAATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGTTGAGAGGGATGGCGA CACTTGAAGAGTACCGTGAGCAAAAAGCGGGAAGTACG CTTATCCGCCATGCCGAAGCGCTTCTTCGGAAAAAAGGG GGCAGACCTCTTATGGTGCAACGCCAGGACATCTGCGA

		GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACACACCGCCGGTCGGACCTCATATTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:71	14_7H9	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGGT CAGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCG ACACTTGAAGGATACCGTGAGCAAAAAGCGGGAAGCA CGCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAA GGCGCGGACCTTTTGTGGTGCAACGCCAGGACATCTGC GAGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGG GCGAAGTCTACGACATACCGCCGGTCGGACCTCATATTT TGATGTATAAGAAATTGACGTAA
SEQ ID NO:72	14_8F7	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGGT CAGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCG ACACTTGAAGAGTACCGCGAGCAAAAAGCGGGCAGTAC GCTTATCCGCCATGCTGAAGCGCTTCTTCGGAAAAAAG GCGCGGACCTTTTGTGGTGCAACGCCAGGACATCTGCA AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGACTGGGCCCCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:73	15_10C2	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGGTC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGCGGGAAGTACG CTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG GGCAGACCTCTTATGGTGCAACGCCAGGACAACGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGT GAAGTCTTCGACATACCGCCGACCGGACCCCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:74	15_10D6	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTAGGTGGATATTACCGGGGCAAGCTGGT CAGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCG ACACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGCA CGCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GGGGCAGACCTCTTATGGTGCAACGCCAGGACATCTGC GAGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGG GCGAAGTCTACGACATACCGCCGGTCGGACCTCATATTT



		TGATGTATAAGAAATTGACGTAA
SEQ ID NO:75	15_11F9	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTTGGTGGATATTACCGGGGCAAGCTGGTC AGCATCGCCTCCTTTAATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTCGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAGAAAGGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGCTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGACCGGACCCCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:76	15_11H3	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATACTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACACCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGCGCTTCTTCGGAAAAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCAACCTGGGCCCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:77	15_12A8	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGT ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGCGCTTCTTCGGAAAAAAGG GGCAGACCTCTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGACCGGACCCCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:78	15_12D6	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGTATGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGGT CAGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTTGAAGGGTACCGTGAGCAAAAAGCGGGCAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACACACCGCCGGTCGGACCTCATATTTT GATGTATAAGAAAGTTGACGTAA
SEQ ID	15_12D8	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA

NO:79		TGAGATCAGGCACCGCATACTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGGT CAGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAACTGAGAGGGGATGGCG ACACTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGACGTCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CAAAGTCTACGACATACCGCCGGTCGGACCTCATATTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:80	15_12D9	ATGATTGAAGTCAAACCAATAAACGCGGAGGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGGT ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGGT CAGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCG ACACTCGAAGAGTACCGCGAGCAAAAAGCGGGGAAGCA CGCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GGGGCAGACCTCTTATGGTGCAACGCCAGGACATCTGC GAGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGG GCGAAGTCTACGACATACCGCCGGTCGGACCTCATATTT TGATGTATAAGAAATTGACATAA
SEQ ID NO:81	15_3F10	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGGT GCGTTTCACCTTGGTGGATATTACCGGGGCAAGCTGATC AGCATCGTTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCGA CACTTGAAGGGTACCGTGAGCAAAAAGCGGGCAGCACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACACACCGCCGGCCGGACCTCATATTTT GATGTATACGAAATTGACGTAA
SEQ ID NO:82	15_3G11	ATGATTGAAGTTAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATACTCCGGCCGAATCAGCCGC TTGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGGT CAGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCG ACACTTGAAGAGTACCGCGAGCAAAAAGCGGGCAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCGGACCTTTTGTGGTGCAACGCCAGGACGTCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGGTCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:83	15_4F11	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TAAGATCAGGCACCGCATACTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC

		ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGGTC AGCATCGCTTCCTTTAATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGCGCTTCTTCGGAAGAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGACCGGACCCCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:84	15_4H3	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGGT CAGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGTA CGCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAA GGCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGC GAGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGG GCGAAGTCTACGACATACCGCCGACTGGGCCCATATT TTGATGTATAAGAAATTGACGTAA
SEQ ID NO:85	15_6D3	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATACTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACACCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGA AAAAGG GGGCAGACCTCTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGACCGGACCCCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:86	15_6G11	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGGTC AGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGA AAAAGG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CAAAGTCTACGACATACCGCCGGTCGGACCTCATATTTT GATGTATAAGAAGTTGACGTAA
SEQ ID NO:87	15_9F6	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGAT CAGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCT

		TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTCGAAGAGTACCGCGAGCAAAAAGCGGGCAGTA CGCTTATCCGCCATGCCGAAGAGCTTCTTCGGAGAAAA GGCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGC GAGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGG GCGAAGTCTACGACATACCGCCTGTCGGACCTCATATTT TGATGTATAAGAAATTGACGTAA
SEQ ID NO:88	15F5	ATGATCGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGT ACGTTTCACCTCGGTGGGTACTACCGGGGCAAGCTGAT CAGCATCGCTTCCTTTCATAAAGCCGAACATTCAGAGCT TGAGGGCGAAGAACAGTATCAGCTGAGAGGGATGGCG ACGCTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTAC GCTTATCCGCTATGCCGAAGAGCTTCTTCGAAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGTG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGATCGGACCTCATATTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:89	16A1	ATGATTGAAGTCAAACCTATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTTGCTCGGGGGT ACGCTTCACCTCGGTGGATATTACCAGGGGCAAGCTGAT CAGCATCGCTTCCTTTCATAAAGCCGAACATTCAGGGCT TGAGGGCGAAGAACAGTATCAGCTGAGAGGGATGGCG ACGCTCGAAGGGTACCGCGAGCAAAAAGCGGGCAGTA CGCTTATCCGCCATGCCGAAGAGCTTCTTCGAAAAAA GGCGCGGACCTTTTATGGTGCAATGCCAGGACATCTGT GAGCGGCTACTATGAAAAGCTCGGCTTCAGCGAACAGG GCGAAGTCTACGACATACCGCCGATCGGACCTCATATTT TGATGTATAAGAAATTGACGTAA
SEQ ID NO:90	16H3	ATGATTGACGTCAAACCTATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGC ACGTTTCACCTCGGCGGATATTACCAGGGGCAAGCTGAT CAGCATCGCTTCCTTTCATCAAGCCGAACATTCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTTGAAGGGTACCGCGAGCAAAAAGCGGGGAAGTA CGCTCATCCGCCATGCCGAAGAGCTTCTTCGAAAAAA GGGGCAGACCTTTTATGGTGCAATGCCAGGACATCTGT GAGCGGGTACTATGAAAAGCTCGGCTTCAGCGAACAGG GCGAAGTCTACGACATACCGCCGATCGGACCTCATATTT TGATGTATAAGAAATTGACGTAA
SEQ ID NO:91	17C12	ATGATTGAAGTCAAACCAATAAGCGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGTATGTATGAAACCGATTTGCTCGGGGGT GCGTTTCACCTCGGTGGATATTACCAGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGGAAGTAC

		GCTTATCCGCCATGCCGAAGAGCTTCTTCGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGTG AGCGGGTACTATGAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:92	18D6	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATAAAGCCGAACATTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCAA CGCTTGAAGGATACCGTGAGCAAAAAGCGGGAAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGAAAAAAGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGTGA GCGGCTACTATGAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGATCGGACCTCATATTTT ATGTATAAGAAATTGGCATAA
SEQ ID NO:93	19C6	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC TGCATCGCTCCTTTCATCAAGCCGAACATTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGAAAAAAGG GGCAGACCTTTTATGGTGCAATGCCAGGACATCTGTG AGAGGCTACTATGAAAAGCTCGGCTTCAGCGAACAAAGG CGGGGTCTACGATATACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGGCGTAA
SEQ ID NO:94	19D5	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACTGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCAGGGCAAGCTGATC AGCATCGCTTCCTTTCATAAAGCCGAACATTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTACG CTCATCCGCCATGCCGAAGAGCTTCTTCGAAAAAAGGG GGCAGACCTTTTATGGTGCAATGCCAGGACATCTGTGA GCGGCTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGATCGGACCTCATATTTT ATGTATAAGAAATTGACGTAA
SEQ ID NO:95	20A12	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCAGGGCAAGCTGATC AGCATCGCTTCCTTTCATAATGCCGAACATTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGGTACCGTGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGAAAAAAGG GGTAGACCTTTTATGGTGCAACGCCAGGACATCTGTG

		AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAGG CGGGATCTACGACATACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGGCATAA
SEQ ID NO:96	20F2	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTTCATCAAGCCGAACATTTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCAGACCTTTTATGGTGCAACGCCAGGACATCTGTG AGCGGCTACTATGAAAAGCTCGGCTTCAGCGAACAAGG CGAAGTCTACGACATACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:97	2.10E+12	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGT GCGTTTCACCTCGGTGGATATTACCAGGGCAAGCTGATC AGCATCGCTTCCTTTTCATCAAGCCGAACATTTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGATACCGTGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCAGACCTTTTATGGTGCAACGCCAGGACATCTGTG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAGG CGAAGTCTACGACATACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:98	23H11	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGTATGTATGAAACCGATTTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCAGGGCAAGCTGATC AGCATCGCTTCCTTTTCATAAAGCCGAACATTTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTCCGAAAAAAGG CGCGGACCTTTTATGGTGCAATGCCAGGACATCTGCGA GCGGCTACTATGAAAAGCTCGGCTTCAGCGAACAAGGC GAAGTCTACGACATACCACCGATCGGACCTCATATTTTG ATGTATAAGAAATTGGCATAA
SEQ ID NO:99	24C1	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGC ACGTTTCACCTCGGCGGATATTATCGGGACAGGCTGATC AGCATCGCTTCCTTTTCATCAAGCCGAACATTTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGTG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAGG CGAAGTCTACGACATACCGCCGATCGGACCTCATATTTT

		GATGTATAAGAACTGACGTAA
SEQ ID NO:100	24C6	ATGATTGAAGTCAAACCTATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATTGAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGGTACCGCGAGCAAAAAGCGGGGAAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGATATCTGTG AGCGGCTACTATAAAAAGCTCGGCTTCAGCGAACAAGG CGGGGTCTACGACATACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGGCATAA
SEQ ID NO:101	2.40E+08	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAGGCATGCAAGTATGAAACCGATTTGCTCGGGGGC ACGTTTCATCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATAATGCCGAACATTGAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGATACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGAAAAAAGGG GGCAGACCTTTTATGGTGCAATGCCAGGACATCTGCGA GCGGCTACTATGAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGATCGGACCTCATATTTTG ATGTATAAGAAATTGGCATAA
SEQ ID NO:102	2_8C3	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGTATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTTGCTCGGGGGT ACGTTTCACCTCGGCGGATATTATCGGGACAGGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATTGAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGAAAAAAGGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGCTACTATGAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGATCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:103	2H3	ATGATTGAAGTCAAACCGATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACAGGGCAAGCTGATC AGCACCGCTTCCTTTCATCAAGCCGGACATTGAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCGAAAAGCGGGGAAGTAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGAAAAAAGG GGGCAGACCTTTTATGGTGCAACGCCAGGATATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAGG CGGGGTCTACGATATACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID	30G8	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA

NO:104		TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTTTCAAACCGATTGCTCGGGGGTG CGTTTCACCTCGGTGGATATTACCAGGGCAAGCTGATCA GCATCGCTTCCTTTCATCAAGCCGAACATTCAGAGCTTG AAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCGAC GCTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTACGC TTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGGC GCAGACCTTTTATGGTGCAACGCCAGGACATCTGTGAG CGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGCG AAGTCTACGACATACCGCCGATCGGACCTCATATTTTGA TGTATAAGAAATTGACGTAA
SEQ ID NO:105	3B_10C4	ATGATTGAAGTCAGACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGTATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCAGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGCCTACGACATACCGCCGATCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:106	3B_10G7	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCAGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCGA CGCTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAAGG GGGGTCTACGACATACCGCCGATCGGACCCCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:107	3B_12B1	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCAGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:108	3B_12D10	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGTATTCTCCGGCCGAATCAGCCGC TGGAAGCATGTATGTACGAAACCGATTGCTCGGGGGT



		GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTTCATCCAGCCGAACATTGAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGATATCTGCGA GCGGGTACTATGAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGATCGGACCCCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:109	3B_2E5	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTTCATCAAGCCGAACATTGAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGCTACTATGAAAAGCTCGGCTTCAGCAAACAGGGC GAAGTCTACGACATACCGCCGATCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:110	3C_10H3	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGTATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTTCATCAAGCCGAACATTGAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGCGGGAAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG GGCAGACCTTTTATGGTGCAACGCCAGGATATCTGCGA GCGGCTACTATAAAAAGCTCGGCTTCAGCGAACAAAGGC GGGGTCTACGACATACCGCCGTCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:111	3C_12H10	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTTCATCAAGCCGAACATTGAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGTGGGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGCTACTATGAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGATCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:112	3C_9H8	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGTATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGCGGATATTATCAGGACAGGCTGATC AGCATCGCCTCCTTTTCATCAAGCCGAACATTGAGAGCTT

		GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGTAC GCTTATCCGCTATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGATATCTGCG AGCGGCTACTATGAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:113	4A_1B11	ATGATTGAAGTCAAACCTATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATACTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTCCTTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGCTACTATGAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGATCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:114	4A_1C2	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGC ACGTTTCACCTCGGCGGATATTATCGGGGCAAGCTGATC AGCATCGCTCCTTTTCATCAAGCCGAACATTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGAGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGATCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:115	4B_13E1	ATGATTGAAGTCAAACCTATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTCCTTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCGGACCTTTTGTGGTGCAACGCCAGGATATCTGCGA GCGGCTACTATGAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGATCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:116	4B_13G10	TTACGTCAATTTCTTATACATCAAAATATGAGGTCCGAT CGGCGGTATGTCGTAGACTTCGCCCTGTTGCTGAAGCC GAGCTTTTATAGTACCCGCTCGCAGATGTCCTGGCGTT GCACCATAAAAAGGTCCGCGCCTTTTTTCCGAAGAAGCTC TTCGGCATGGCGGATGAGCGTGCTTCCCGCTTTTTGCTC GCGGTACCCTTCAAGCGTCGCCATCCCTCTCAGCTGATA CTGTTTTTGGCCTTCAAGCTCTGAATGTTCCGCTTGATG

		AAAGGAGGCGATGCTGATCAGCTTGCCCCGGTAATATC CACCGAGGTGAAACGTGCCCCGAGCAAATCAGTTTCA TACTTGCATGCTTCCAGCGGCTGATTTCGGCCGGAGAATG CGGTGCCTGATCTCATACGTATCTTCCGCGTTTATTGGT TTGGCTTCAATCAT
SEQ ID NO:117	4B_16E1	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGT ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGAT CAGCATCGCTTCTTTTCATCAAGCCGAACATTCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCG ACACTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAGG CGGGGTCTACGATATACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:118	4B_17A1	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGC ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGAT CAGCATCGCTTCTTTTCATCAAGCCGAGCATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCG ACGCTTGAAGAGTACCGCGAGCAAAAAGCGGGCAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGCTACTATGAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGACATAA
SEQ ID NO:119	4B_18F11	ATGATTGAAGTCAATCCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTTGCTCGGGGGC ACGTCTCACCTCGGCGGATATTACCGGGGCAAGCTGAT CAGCATCGCTTCTTTTCATAATGCCGAACATTCAGAGCT TGATGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCAGACCTTTTATGGTGCAACGCCAGGACATCTGTG AGCGGCTACTATGAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:120	4B_19C8	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGC ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGAT CAGCATCGCTTCTTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCG ACGCTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGCA CGCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGC

		GAGCGGGTACTATAAAAAAGCTCGGCTTCAGCGAACAAG GCGGGGTCTACGATATACCGCCGATCGGACCTCATATTT TGATGTATAAGAAATTGGCATAA
SEQ ID NO:121	4B_1G4	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGT GCGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGAT CAGCATCGCCTCCTTTCATCAATCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACGCTTGAAGGGTACCGCGAGCTAAAAGCGGGAAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGATATCTGCG AGCGGGTACTATAAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:122	4B_21C6	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATTGAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGATATCTGCG AGCGGCTACTATAAAAAAGCTCGGCTTCAGCGAACAAAGG CGGGGTCTACGATATACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:123	4B_2H7	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATTGAGAGCTT GAAGGCCAAAAACAGTACCAGCTGAGAGGGATGGCGA CGCTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGCTACTATAAAAAAGCTCGGCTTCAGCGAACAAAGG GGGGTCTACGGCATAACCGCCGATCGGACCTCATATTTT ATGTATAAGAAATTGACATAA
SEQ ID NO:124	4B_2H8	ATGATTGAAGCCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACTGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATTGAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATAACCGCCGATCGGACCTCATATTTT

		GATGTATAAGAAATTGACGTAA
SEQ ID NO:125	4B_6D8	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATACTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGGTACCGCGAGCAAAAAGCGGGTAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACATGGC GAAGTCTACGACATAACCGCCGATCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:126	4B_7E8	ATGATTGAAGTCAAACC <sup>A</sup> ATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCA <sup>T</sup> ICTCCGGCCGAATCAGCCGC TGGAAGCATGCATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGTGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATAACCGCCGATCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:127	4C_8C9	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGCTACTATGAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATAACCGCCGATCGGACCTCATATTTTG ATGTATAAGAAATTAACATAA
SEQ ID NO:128	4H1	ATGATTGAGGTGAAACCGATTAAACGCAGAGGAGACCTA TGAAC <sup>T</sup> AAGGCATAGGATACTCAGACCACACCAGCCGA TAGAGGTTTGTATGTATGAAACCGATTACTTCGTGGTG CGTTTCACTTAGGCGGCTTTTACAGGGGCAAGCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCATCCAGAACTCC AGGGCCAGAAACAATACTCACTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGACCAGAAAGCGGGATCGAGCCT AATTAAACACGCTGAACAGATCCTTCGGAAGCGGGGGG CGGACATGCTATGGTGCAATGCGCGGACATCCGCCGCT GGCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGA GGTATTTGAAACGCCGCCAGTAGGACCTCACATCGTAA TGTATAAACGCCTCACATAA
SEQ ID	6_14D10	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA

NO:129		TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGTATGTATGAAACCGATTTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGAGGCAAGCTGATC AGCATCGCCTCCTTCCATCAAGCCGAACATTCAGAGCTT GAAGGCCATAAACAGTATCAGCTGAGAGGGATGGCGAC ACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGCACG CTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGCTACTATAAAAAGCTCGGCTTCAGCGAACAAGGC GGGGTCTACGACATACCGCCGGTTCGGACCTCATATTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:130	6_15G7	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTAAGTATGAAACCGATTTGCTCGGGGGC ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGAT CAGCATCGCTTCTTTTCATCAAGCCGAACATTCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGCA CGCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAA GGCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGC GAGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGG GCGAAGTCTACGACATACCGCCGGTTCGGACCTCATATTT TGATGTATAAGAAATTGACGTAA
SEQ ID NO:131	6_16A5	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCACCAAGCCGAACATTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGCGGGAAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAGGC GGGGTCTACGACATACCGCCGGTTCGGACCTCATATTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:132	6_16F5	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCTTTTCATCAAGCCGTACATTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGCTACTATAAAAAGCTCGGCTTCAGCGAACAAGGC GGGGTCTACGACATACCGCCGGTTCGGACCTCATATTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:133	6_17C5	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGCAAGTATGAAGCCGATTTGCTCGGGGGC

		ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAGCATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGAAACCGTGAGCAAAAAGCGGGCAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACGTACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:134	6_18C7	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAGGTATGAAACCGATTTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTATCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGGG GGCAGACCTTTTATGGTGCAACGCCAGGATATCTGCGA GCGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTTTACGACATACCGCCGGTCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:135	6_18D7	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCMCCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGTACTATAAAAAGCTCGGCTTCAGCGAACAAAGG CGGGGTCTACGACATACCGCCGGTCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:136	6_19A10	ATGATTGAAGCCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGACCGGACCCCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:137	6_19B6	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTTGCTCAGGGGT GCGTTTCACCTCGGTGGATATTATCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATTAGAGCTT

		GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTCGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGGG CGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGCTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATAACCGCCGGTCGGACCTCATATTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:138	6_19C3	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGT ACGTTTACCTCGGCGGATATTACCGGGGCAAGCTGAT CAGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATAACCGCCGATCGGACCTCATATTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:139	6_19C8	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGT ACGTTACACCTCGGTGGATATTACCGGGGCAAGCTGAT CAGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTAC GCTTATCCGCCAAGCCGAAGAGCTTCTTCGGAAAAAAG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAAG CGGGGTCTACGACATAACCGCCGGTCGGACCTCATATTT GATGTATAAGGAATTGACGTAA
SEQ ID NO:140	6_20A7	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTTGCTCAGGGGC ACGTTTACCTCGGCGGATATTACCGGGGCAAGCTGAT CAGCATCGCTTCCTTTCATCAAGCCGAACATTGAGATCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTTGAAGAGTACCGCGAGCAAAAAGCGGGGAAGCA CGCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGC GAGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGG GCGAAGTCTACGACATAACCGCCGGTCGGACCTCATATTT TGATGTATAAGAAATTGACGTAA
SEQ ID NO:141	6_20A9	ATGATTGAAGTCAAACCAATAAACGCGGGAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGC ACGTTTACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATTGAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTACG



		CTTATCCGCCATGCCGAAGAGCTTCTACGGAAAAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAGGC GGGGTCTACGACATACCGCCGGTCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:142	6_20H5	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGAT CAGCATCGCCTCCTTTCATCAAGCCGAACATTCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTTGAAGGATACCGTGAGCAAAAAGCGGGGAAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGCTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:143	6_21F4	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCGTTCTCCGGCCGAATCAGCCGC TGGAAGCATGTATGTATGAAACCGATTGCTCGGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG CGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACGTACCGCCGGTCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:144	6_22C9	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCGGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATCCAGGGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCGCGAGCAAAAAGCGGGGAAGCAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGACTTCCGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAAG AGGGGTCTACGACATACCGCCGGTCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:145	6_22D9	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGTATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCATGTATGAAACCGATTGCTCGAGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAGCATTAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCGA

		GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGGTCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:146	6_22H9	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATTTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCGA CGCTTGATGAGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGATCGGACCCCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:147	6_23H3	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGGAAGTATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAGCAACCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAGCAAGGC GGGGTCTACGACATACCGCCGGTCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:148	6_23H7	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATTTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCGA CGCTTGAAGGATACCGCGAGCAAAAAGCGGGAAGTAC GCTTATCCGCCATGCAGAAGAGATTCTTCGGAAAAAAG GCGCGGACCTCTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAGG CGGGGTCTACGACATACCGCCGGTCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:149	6_2H1	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGGCTTCTCCGGCCGAATCAGCCGC TGGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACCGTATCAGCTGAGAGGGGATGGCGA CACTTGAAGGATACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAATCTACGACATACCGCCGATCGGACCTCATATTTTG

		ATGTATAAGAAATTGACGTAA
SEQ ID NO:150	6_3D6	ATGATTGAAATCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCGAGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CTCTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAGGTCTACGACATACCGCCGGTCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:151	6_3G3	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGCTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGGTCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:152	6_3H2	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG GGCAGACCTCTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGGTCGGACCTCATATTTT GATGTATAAGAAATTGACATAA
SEQ ID NO:153	6_4A10	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAACTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGATACCGTGAGCAAAAAGCGGGAAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGCTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGGTCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID	6_4B1	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA

NO:154		TGAGATCAGGCACCGCGTACTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC GGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGCTACTATGAAAAGCTCGGCTTCAGCGGACAGGGC GAAGTCTACGACATACCGCCGATCGGACCTCATATTTTG ATGTATAAGAAATTGACATAA
SEQ ID NO:155	6_5D11	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGATCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:156	6_5F11	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTAATC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCCACGACATACCGCCGGTTCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:157	6_5G9	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTAAT CAGCATCGCTTCCTTTCATCAAGCCGAACATTCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACGCTTGAAGAGTACCGTGAGCAAAAAGCGGGCAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG GGCAGACCTTTTATGGTGCAACGCCAGGATATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAAGG CGGGGTCTACGACATACCGCCGGTTCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:158	6_6D5	ATGATTGAAGTCAAACCAATAAACGCGGAAGATGCGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACTGATTGCTCGGGGGC

		ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGAT CAGCATCGCTTCCTTTCATCAAGCCGAACATTTCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTTGAAGGGTACCGCGAGCAAAAAGCGGCGCAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCGGACCTTTTGTGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAGG CGGGGTCTACGACATACCGCCGGTCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:159	6_7D1	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATTTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAGGC GGGGTCTACGACATACCGCCGGTCGGACCTCATATTTT ATGTATAAGAAATTGACGTAA
SEQ ID NO:160	6_8H3	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGTATGTATGAAACCGATTGCTCAGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAGG CGGGGTCTACGACATACCGCCGGTCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:161	6_9G11	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGGC ACGCTTCACCTCGGTGGATATTACCGGGGCAAGCTGAT CAGCATCGCTTCCTTTCATCAAGCCGAACATTTCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACGCTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGTA CGCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAA GGCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGC GAGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAG GCGAAGTCTACGACATACCGCCGGTCGGACCTCATATTT TGATGTATAAGAAATTGACGTAA
SEQ ID NO:162	6F1	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCAGGGGGT ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGGTC TGCATCGCTTCCTTTCATAAAGCCGAACATTTCAGAGCTT

		GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGATGGATACCGCGAGCAAAAAGCGGGAAGCACG CTCATCCGCCATGCCGAAGAGCTTCTTCGAAAAAAGG CGCGGACCTTTTATGGTGCAATGCCAGGATCTGTGA GCGGCTACTATGAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGGTCGGACCTCATATTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:163	7_1C4	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCTTTTCATCAAGCCGAGCATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGAAAAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAGGC GGGGTCTACGATATACCGCCGATCGGACCTCATATTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:164	7_2A10	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACTGATTGCTCGGGGGC ACGTTTCATCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCTTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAGG CGGGGTCTACGATATACCGCCGATCGGACCTCATATTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:165	7_2A11	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCTTTTCATCAAGCCGAACATTAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGAAAAAAG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAGG CGGGGTCTACGACATACCGCCGGTCGGACCTCATATTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:166	7_2D7	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCTTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGGTACCGTGAGCAAAAAGCGGGAAGTACG

		CTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GTGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGGTCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:167	7_5C7	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTCCTTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGTGGAAGCACG CTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGC GGGGTCTACGATATACCGCCGGTCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:168	7_9C9	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAAATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTCCTTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGTAC GCTCATCCGCCATGCCGAAGAGCTTCTACGGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:169	9_13F10	ATGATTGAAGTCAAACCAATAAACGCGGAGGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTTGGTGGATATTACCGGGGCAAGCTGGTC AGCATCGCTCCTTTTCATCAAGCCGAACATTTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG GGGCAGACCTCTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGACTGGGCCCCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:170	9_13F1	ATGATTGAAGTCAAACCAATAAACGCGGAGGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTTGGTGGATATTACCGGGGCAAGCTGGTC AGCATCGCTCCTTTTCATCAAGCCGAACATTTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG GGGCAGACCTCTTATGGTGCAACGCCAGGACATCTGCG

		AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGACTGGGCCCCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:171	9_15D5	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGACGCATGCAAGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGA AAAAGGG GGCAGACCTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGGTCGGACCTCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:172	9_15D8	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATACTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGGT CAGCATCGCCTCCTTTCATCAAGCTGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTTGAAGGGTACCGTGAGCAAAAAGCGGGCAGTAC GCTTATCCGCCATGCCGAAGCGCTTCTTCGGAAGAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACACACCGCCGGTCGGACCCCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:173	9_15H3	ATGATTGAAGTCAAGCCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATATGCTCAGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCACGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGA AAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTTAGCGAACAGGG CGAAGTCTACAACACACCGCCGGTTGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:174	9_18H2	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGAT CAGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCT TGTAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTACA CTTATCCGCCATGCCGAAGAGCTTCTTCGGA AAAAGGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGGTCGGACCTCATATTTTG



SEQ ID NO:175	9_20F12	ATGTATAAGAAATTGACGTAA ATGATTGAAGTAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCTTCTCCGGCCGAATCAGCCGC TGGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCGAGCTGGTC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCGGACCTTTTGTGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAAGG GGGGTCTACGACATACCGCCGGTCGGACCTCATATTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:176	9_21C8	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGTATGTATGAAACTGATTGCTCGGGGGC ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGAT CAGCATCGCTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTCGAAGGATACCGCGAGCAAAAAGCGGGCAGTA CGCTAATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GGGGCAGACCTCTTATGGTGCAACGCCAGGACATCTGC GAGCGGGTACTATAAAAAGCTCGGCTTCAGCGATCAGG GCGAAGTCTACGACATACCGCCGGTCGGACCTCATATTT TGATGTATAAGAAATTGACGTAA
SEQ ID NO:177	9_22B1	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATAAGGCACCGCATCCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGGTC AGCATCGCTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACTTACCGCCGACCGGACCCCATATTTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:178	9_23A10	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGC ACGCTTCACCTCGGTGGATATTACCGGGGCAAGCTGGT CAGCATTGCTTCCTTTCATCAAGCCGAACATCCAGAGCT TGAGGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTTGAAGGGTACCGCGGGCAAAAAGCGGGCAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG GGGCAGACCTTTTATGGTGCAATGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAAGG CGGGGTCTACGACATACCGCCGGTCGGACCTCATATTT GATGTATAAGAAATTGACGTAA
SEQ ID	9_24F6	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA

NO:179		TGAGATCAGGCACCGCATTCTCAGGCCGAATCAGCCGC TAGAAGCATGCAAGTATGAAACCGATTGCTCAGGGGT GCGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGCGCTTCTTCGGAAAAAAGG CGCGGACCTTTTGTGGTGCAACGCCAGGACGTCTGCGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGACCGGACCCCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:180	9_4H10	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACTGATTGCTAGGGGGT ACGCTTCACCTCGGTGGATATTACCGGGGCAAGCTGAT CAGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTTGAAGGGTACCGTGAGCAAAAAGCGGGCAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG GCGCGGACCTTATATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGGTCGGACCTCATATTTT GATGTATAAGAAATTGACATAA
SEQ ID NO:181	9_4H8	ATGATTGAAGTCAAACCAATAAATGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGAGGC ACGTTTCACCTAGGTGGATATTACCGGGGCAAGCTGAT CAGCATCGCTTCCTTTAATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACACTTGAAGGGTACCGTGAGCAAAAAGCGGGCAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAA'AAGG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGGTCGGACCTCATATTTT GATGTATAAGAAATTGACATAA
SEQ ID NO:182	9_8H1	ATGATTGAAGTCAAACCAATAACCGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGT ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTAGAAGGGTACCGCGAGCAAAAAGCGGGCAGTAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG GGGCAGACCTTTTATGGTGCAACGCCAGAACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGACCGGACCCCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:183	9_9H7	ATGATTGAAGTCAAACCAATAAACGCGGAAGATGCGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGAGC

		ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGAGTACCGCGAGCAAAAAGCGGGAAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCTGTTCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:184	9C6	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGT ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC TGCATCGCCTCCTTTCATCAAGCCGAACATTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GGGCAGACCTTTTATGGTGCAATGCCAGGACATCTGTG AGAGGCTACTATGAAAAGCTCGGCTTCAGCGAACAAAG CGGGGTCTACGATATACCGGATCGGACCTCATATTTT GATGTATAAGAAATTGGCGTAA
SEQ ID NO:185	9H11	ATGATTGAAGTCAAACCAA <sup>*</sup> AAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGT TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGT ACGTTTCACCTCGGCGGATATTACCGGGGCAAGCTGAT CAGCATCGCTTCCTTTCATAAAGCCGAACATTCAGAGCT TGAGGGCGAAGAACAGTATCAGCTGAGAGGGATGGCG ACGCTTGAAGGATACCGTGAGCAAAAAGCGGGAAGCA CGCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GGGGCAGACCTTTTATGGTGCAATGCCAGGACATCTGT GAGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGG GCGAAGTCTACGACATACCGCCGATCGGACCTCATATTT TGATGTATAAGAAATTGACGTAA
SEQ ID NO:186	0_4B10	ATGATAGAAGTGAAACCGATTAAACGCAGAGGATACCTA TGAAC <sup>T</sup> AAGGCATAAAATACTCAGACCAAACAGCCGA TAGAAGCGTGTATGTATGAAAGCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGCAAACTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCG AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGATCAGAAAGCGGGATCGACTCT AATTAACACGCTGAAGAAATTCTTCGTAAGAGGGGGG CGGACATGCTTTGGTGCAATGCGCGGACAACCGCCTCA GGCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGA GATATTTGATACGCCGCCAGTAGGACCTCACATCCTGAT GTATAAAAGGCTCACATAA
SEQ ID NO:187	0_5B11	ATGATAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAAC <sup>T</sup> AAGGCATAAAATACTCAGACCAAACAGCCGA TAGAAGCGTGTATGTATGAAAGCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACGGGGGCAAACTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCG

		AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGATCAGAAAGCGGGATCGACTCT AATTAAACACGCTGAACAACTTCTTCGTAAGAGGGGGG CGGACATGCTTTGGTGCAATGCGCGGACATCCGCCTCA GGCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGA GGTATTTGAAACGCCGCCAGTAGGACCTCACATCCTGA TGTATAAAAAGATCACA
SEQ ID NO:188	0_5B3	ATGCTAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAACATAAGGCATAGAATACTCAGACCAAACAGCCGA TAGAAGCGTGTATGTATGAAACCGATTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGAACTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGATCAGAAAGCGGGATCGAGTCT AATTAAACACGCTGAACAACTTCTTCGTAAGAGGGGGG CGGACTTGCTTTGGTGCAATGCGCGGACATCCGCCTCAG GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGAG GTATTTGATACGCCGCCAGTAGGACCTCACATCCTGATG TATAAAAGGATCACA
SEQ ID NO:189	0_5B4	ATGCTAGAGGTGAAACTGATTAAACGCAGAGGATACCTA TGAACATAAGGCATAGAATACTCAGACCAAACAGCCGT TAGAAGCGTGTATGTATGAAACCGATTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCG AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTTTCTCGTGATCAGAAAGCGGGATCGAGTCT AATTAAACACGCTGAAGAAATTCTTCGTAAGAGGGGGG CGAACTTGCTTTGGTGTAATGCGCGGACATCCGCCTCAG GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGAG GTATTTGATACGCCGCCAGTAGGACCTCACATCCTGATG TATAAAAGGATCACA
SEQ ID NO:190	0_5B8	ATGATAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAACATAAGGCATAAAATACTCAGACCAAACAGCCGA TAGAAGCGTGTATGTATGAAAGCGATTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGATCAGAAAGCGGGATCGAGTCT AATTAGACACGCTGAACAAATTCTTCGTAAGAGGGGGG CGGACTTGCTTTGGTGCAATGCGCGGACATCCGCCTCAG GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGAG ATATTTGATACGCCGCCAGTAGGACCTCACATCCTGATG TATAAAAGGCTCACA
SEQ ID NO:191	0_5C4	ATGATAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAACATAAGGCATAAAATACTCAGACCAAACAGCCGT TAGAAGCGTGTATGTATGAAACCGATTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGGCCTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGAGTAT

		AATTAACACGCTGAAGAAATTCTTCGTAAGAAGGGGG CGGACTTGCTTTGGTGCAATGCGCGGACGTCCGCCTCAG GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGAG ATATTTGACACGCCGCCAGTAGGACCTCACATCCTGATG TATAAAAGGATCACA
SEQ ID NO:192	0_5D11	ATGATAGAGGTGAAACCGATTAACGCAGAGGATACCTA TGAAC TAAGGCATAGAATACTCAGACCAAACCAGCCGA TAGAAGCGTGTATGTATGAAAGCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGCAAAC TGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGACTCT AATTAGACACGCTGAACAAC TTCTTCGTAAGAGGGGGG CGGACTTGCTTTGGTGCAATGCGCGGACATCCGCCTCAG GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGAG GTATTTGATACGCCGCCAGTAGGACCTCACATCCTGATG TATAAAAGGCTCACA
SEQ ID NO:193	0_5D3	ATGCTAGAGGTGAAACCGATTAACGCAGAGGATACCTA TGAAC TAAGGCATAGAATACTCAGACCAAACCAGCCGA TAGAAGCGTGTATGTATGAAAGCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTATTACAGGGGCAAAC TGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGAACTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGAGTCT AATTAACACGCTGAAGAAATTCTTCGTAAGAGGGGGG CGGACTTGCTTTGGTGTAATGCGCGGACATCCGCCTCAG GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGAG ATATTTGAAACGCCGCCAGTAGGACCTCACATCCTGAT GTATAAAAGGATCACATAA
SEQ ID NO:194	0_5D7	ATGATAGAAGTGAAACCGATTAACGCAGAGGAGACCTA TGAAC TAAGGCATAGAATACTCAGACCAAACCAGCCGA TAGAAGCGTGTATGTATGAAACCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGCAAAC TGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGAACTC GAAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTAC CTTGGAAGGTTATCGTGATCAGAAAGCGGGATCGAGTC TAATTAGACACGCTGAACAAC TTCTTCGTAAGAAGGGG GCGAATATGCTTTGGTGTAATGCGCGGACAACCGCCTC AGGCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAG AGATATTTGATACGCCGCCAGTAGGACCTCACATCCTG ATGTATAAAAGGATCACA
SEQ ID NO:195	0_6B4	ATGCTAGAGGTGAAACCGATTAACGCAGAGGATACCTA TGAAC TAAGGCATAGAATACTCAGACCAAACCAGCCGA TAGAAGCGTGTATGTATGAAAGCGATTTACTTCGTGGTG CACTTCACTTAGGCGGCTTTTACAGGGGCAAAC TGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTTTCGTGATCAGAAAGCGGGATCGAGTCT AATTAGACACGCTGAACAAC TTCTTCGTAAGAGGGGGG CGGACTTGCTTTGGTGCAATGCGCGGACATCCGCCTCAG

		GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAAAG GTATTTGATACGCCGCCAGTAGGACCTCACATCCTGATG TATAAAAGGATCACA
SEQ ID NO:196	0_6D10	ATGCTAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAAC TAAGGCATAAAATACTCAGACCAAACCGCCGT TAGAAGTGTGTATGTATGAAACCGATTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGGCAAAC TGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGATCAGAAAGCGGGATCGAGTCT AATTAGACACGCTGAACAAATTCTTCGTAAGAGGGGGG CGGACATGCTTTGGTGCAATGCGCGGACATCCGCCTCA GGCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGA GGTATTTGAAACGCCGCCAGTAGGACCTCACATCCTGA TGTATAAAAGGCTCACA
SEQ ID NO:197	0_6D11	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGC ACGCTTCACCTCGGTGGATATTACCGGGGCAAGCTGGT CAGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGGATGGCG ACGCTTGAAGGGTACCGTGAGCAAAAAGCGGGCAGTAC GCTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGG GGGCAGACCTTTTATGGTGCAACGCCAGGACATCTGCG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGGTCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:198	0_6F2	ATGATAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAAC TAAGGCATAGAATACTCAGACCAAACCGCCGA TAGAAGCGTGTATGTATGAAAGCGATTACTTCGTGGTG CATTTCACTTAGGCGGCTATTACAGGGGGCAAAC TGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGAACTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTTTCTGTGAGCAGAAAGCGGGATCGACTCT AATTAGACACGCTGAACAAATTCTTCGTAAGAGGGGGG CGGACATGCTTTGGTGCAATGCGCGGACATCCGCCTCA GGCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGA GATATTTGATACGCCGCCAGTAGGACCTCACATCCTGAT GTATAAAAGGATCACA
SEQ ID NO:199	0_6H9	ATGATAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAAC TAAGGCATAAAATACTCAGACCAAACCGCCGA TAGAAGCGTGTATGTATGAAACCGATTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACGGGGGGCAAAC TGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCG AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGACTCT AATTAGACACGCTGAAGAAATTCTTCGTAAGAAGGGGG CGAACTTGCTTTGGTGCAATGCGCGGACATCCGCCTCAG GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGAG GTATTTGACACGCCGCCAGTAGGACCTCACATCCTGATG

		TATAAAAGGCTCACA
SEQ ID NO:200	10_4C10	ATGATAGAGGTGAAACCGATTAACGCAGAGGATACCTA TGAAC TAAGGCAT AAAA TACTCAGACCAA ACCAGCCGT TAGAAGTGTGTATGTATGAAACCGATT TACTTCGTGGTG CATTTCACTTAGGCGGCTNTTACAGGGGCAA ACTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGAACTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGATCAGAAAGCGGGATCGAGTCT AATTAAACACGCTGAACAAATTCTTCGTAAGAGGGGGG CGGACNTGCTTTGGTGCAATGCGCGGACATCCGCCTCA GGCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGA GATATTTGATACGCCGCCAGTAGGACCTCACATCCTGAT GTATAAAAGGCTCACATAA
SEQ ID NO:201	10_4D5	ATGATAGAGGTGAAACCGATTAACGCAGAGGATACCTA TGAAC TAAGGCATAGA TACTCAGACCAA ACCAGCCGA TAGAAGTGTGTATGTATGAAACCGATT TACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGCAA ACTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGACTCT AATTAGACACGCTGAACAAATTCTTCGTAAGAGGGGGG CGGACTTGCTTTGGTGCAATGCGCGGACATCCGCCTCAG GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGAG GTATTTGATACGCCGCCAGTAGGACCTCACATCCTGATG TATAAAAGGATCACATAA
SEQ ID NO:202	10_4F2	ATGCTAGAGGTGAAACCGATTAACGCAGAGGATACCTA TGAAC TAAGGCATAGA TACTCAGACCAA ACCAGCCGA TAGAAGCGTGTATGTTTGAAAGCGATT TACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGCAA ACTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGAACTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGAGTCT AATTAGACACGCTGAAGAAATTCTTCGTAAGAGGGGGG CGGACATGCTTTGGTGTAATGCGCGGACATCCGCCTCA GGCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGA GATATTTGAAACGCCGCCAGTAGGACCTCACATCCTGA TGTATAAAAGGCTCACATAA
SEQ ID NO:203	10_4F9	ATGATAGAGGTGAAACCGATTAACGCAGAGGATACCTA TGAAC TAAGGCATAGA TACTCAGACCAA ACCAGCCGA TAGAAGTGTGTATGTATGAAACCGATT TACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGCAA ACTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGAACTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTTTTCGTGAGCAGAAAGCGGGATCGAGTCT AATTAGACACGCTGAACAAATTCTTCGTAAGAGGGGGG CGGACTTGCTTTGGTGTAATGCGCGGACATCCGCCTCAG GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGAG ATATTTGATACGCCGCCAGTAGGACCTCACATCCTGATG TATAAAAGGCTCACATAA
SEQ ID	10_4G5	ATGATAGAGGTGAAACCGATTAACGCAGAGGATACCTA

NO:204		TGAACTAAGGCATAGAATACTCAGACCAAACCCAGCCGA TAGAAGCGTGTATGTTTGAAGCGATTACTTCGTGGTG CATTTCACTTAGGCGGCTATTACAGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTACCGCGATCAGAAAGCGGGATCGAGTCT AATTAGACACGCTGAACAAATTCTTCGTAAGAGGGGGG CGGACTTGCTTTGGTGTAATGCGCGGACATCCGCCTCAG GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGAG ATATTTGATACGCCGCCAGTAGGACCTCACATCCTGATG TATAAAAGGCTCACATAA
SEQ ID NO:205	10_4H4	ATGCTAGAGGTGAAACCGATTAACGCAGAGGATACCTA TGAACTAAGGCATAAAATACTCAGACCAAACCCAGCCGT TAGAAGTGTGTATGTATGAAACCGATTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGAACTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGAGTCT AATTAAACACGCTGAAGAAATTCTTCGTAAGAGGGGGG CGGACTTGCTTTGGTGCAATGCGCGGACATCCGCCTCAG GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGAG GTATTTGATACGCCGCCAGTAGGACCTCACATCCTGATG TATAAAAGGATCACATAA
SEQ ID NO:206	11_3A11	ATGATAGAAGTGAAACCGATTAACGCAGAGGATACCTA TGAACTGAGGCATAAAATACTCAGACCAAACCCAGCCGA TAGAAGTGTGTATGTATGAAAGCGATTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGCAAACCTGATTT CCATAGCGTCATTCCACCAGGCCGAGCACCCAGACCTC CAAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTAC CTTGAAGGTTATCGTGATCAGAAAGCGGGATCGAGTC TAATTAAACACGCTGAACAAATTCTTCGTAAGAGGGGGG GCGGACTTGCTTTGGTGCAATGCGCGGACATCCGCCTCA GGCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGA GGTATTTGAAACGCCGCCAGTAGGACCTCACATCCTGA TGTATAAAAGGCTCACATAA
SEQ ID NO:207	11_3B1	ATGCTAGAGGTGAAACCGATTAACGCAGAGGATACCTA TGAACTGAGGCATAGAATACTCAGACCAAACCCAGCCGA TAGAAGCGTGTATGTTTGAAGCGATTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCC AAGGCCAGAAACAGTACCAACTCCGAGGTATGGCTACC TTGGAAGGTTTTCGTGAGCAGAAAGCGGGATCGACTCT AATTAGACACGCTGAAGAAATTCTTCGTAAGAGGGGGG CGGACTTGCTTTGGTGCAATGCGCGGACATCCGCCTCAG GCTACTACAAAAAGGTTAGGCTTCAGCGAGCAGGGAGAG ATATTTGACACGCCGCCAGTAGGGCCTCACATCCTGATG TATAAAAGGCTCACATAA
SEQ ID NO:208	11_3B5	ATGATAGAGGTGAAACCGATTAACGCAGAGGATACCTA TGAACTAAGGCATAGAATACTCAGACCAAACCCAGCCGA TAGAAGCGTGTATGTTTGAAGCGATTACTTCGTGGTG



		CATTTCACTTAGGCGGCTATTACAGGGGCAAACCTGATTT CCATAGCGTCATTCCACCAGGCCGAGCACTCGGAACCTC CAAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTAC CTTGGAAGGTTATCGTGATCAGAAAGCGGGATCGAGTC TAATTAGACACGCTGAACAAATTCTTCGTAAGAGGGGG GCGGACATGCTTTGGTGCAATGCGCGGACATCCGCCTC AGGCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAG AGGTATTTGATACGCCGCCAGTAGGACCTCACATCCTG ATGTATAAAAGGATCACATAA
SEQ ID NO:209	11_3C12	ATGATAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAACCTAAGGCATAGAATACTCAGACCAAACCGCCGT TAGAAGTGTGTATGTATGAAACCGATTTACTTCGTGGTG CATTTCACTTGGGCGGCTTTTACGGGGGCAAACCTGATTT CCATAGCGTCATTCCACCAGGCCGAGCACCCAGACCTC CAAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTAC CTTGGAAGGTTATCGTGATCAGAAAGCGGGATCGAGTC TAATTAGACACGCTGAACAACTTCTTCGTAAGAGGGGG GCGGACTTGCTTTGGTGCAATGCGCGGACATCCGCCTCA GGCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGA GATATTCGAAACGCCGCCAGTAGGACCTCACATCCTGA TGTATAAAAGGATCACATAA
SEQ ID NO:210	11_3C3	ATGATAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAACCTAAGGCATAAAATACTCAGACCAAACCGCCGA TAGAAGCGTGTATGTATGAAAGCGATTTACTTCGTGGTG CACTTCACTTAGGCGGCTATTACAGGGGCAAACCTGATTT CCATAGCGTCATTCCACCAGGCCGAGCACTCAGAACTC CAAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTAC CTTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGAGTC TAATTAAACACGCTGAAGAAATTCTTCGTAAGAGGGGG GCGGACTTGCTTTGGTGCAATGCGCGGACATCCGCCTCA GGCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGA GGTATTTGACACGCCGCCAGTAGGACCTCACATCCTGAT GTATAAAAGGATCACATAA
SEQ ID NO:211	11_3C6	ATGCTAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAACCTAAGGCATAAAATACTCAGACCAAACCGCCGA TAGAAGCGTGTATGTTTGAAAGCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCG AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGACTCT AATTAGACACGCTGAAGAAATTCTTCGTAAGAGGGGGG CGGACTTGCTTTGGTGCAATGCGCGGACATCCGCCTCAG GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGAG ATATTTGATACGCCGCCAGTAGGACCTCACATCCTGATG TATAAAAGGATCACATAA
SEQ ID NO:212	11_3D6	ATGATAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAACCTAAGGCATAGAATACTCAGACCAAACCGCCGA TAGAAGTGTGTATGTATGAAACCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCC

		AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGAGTCT AATTAACACGCTGAACAAATTCTTCGTAAGAGGGGGG CGGACTTGCTTTGGTGCAATGCGCGGACATCCGCCTCAG GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGAG GTATTTGATACGCCGCCAGTAGGACCTCACATCCTGATG TATAAAAGGCTCACATAA
SEQ ID NO:213	1_1G12	ATGCTAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAACCTAAGGCATAGAATACTCAGACCAAACCGCCGA TAGAAGTGTGTATGTATGAAACCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACGGGGGCAAACGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGAACTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGATCAGAAAGCGGGATCGAGTCT AATTAACACGCTGAAGAAATTCTTCGTAAGAGGGGGG CGGACTTGCTTTGGTGTAATGCGCGGACATCCGCCTCAG GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGAG GTATTTGAAACGCCGCCAGTAGGACCTCACATCCTGAT GTATAAAAGGCTCACATAA
SEQ ID NO:214	1_1H1	ATGATAGAAGTGAAACCTATTAACGCAGAGGAGACTTA CGAACTTCGACACAAGATCCTGCGCCCTAATCAGCCGA TAGAGGCATGCATGTATGAAAGCGATCTGCTGCGGGGC TCGTTCCATTTGGGCGGGTTCTATCGTGGCAAATTGATC TCGATTGCGAGTTTCCACAAAGCTGAACACTCAGAACT GCAAGGGCAAAAGCAGTATCAATTACGAGGGATGGCG ACCCTCGAAGGATTCCGTGAGCAGAAGGCTGGCTCTTC GCTTATTAGGCACGCCGAGGAGATACTACGGAATAAAG GGGCAGATCTGCTTTGGTGTAATGCACGCACGACAGCC TCCGGTTACTATAAAAGGCTTGGTTTTAGTGAGCACGGC GAAGTTTTCGAAACCCCGCCGGTTGGGCCGCACATTCTT ATGTACAAAAGAATCACT
SEQ ID NO:215	1_1H2	ATGATAGAAGTGAAACCTATTAACGCAGAGGATACTTA CGAACTTCGACACAGGATCCTGCGCCCTAATCAGCCGTT AGAGGCATGCATGTATGAAAGCGATCTGCTGCGGGGCT CGTTCCATTTGGGCGGGTTCTATCGTGGCAAATTGATCT CGATTGCGAGTTTCCACCAAGCTGAACACTCAGAACTG GAAGGGCAAAAGCAGTATCAATTACGAGGGATGGCGA CCCTCGAAGGATTCCGTGAGCAGAAGGCTGGCTCTTCG CTTATTAGGCACGCCGAGGAGATACTACGGAATAAAG GGCAGATCTGCTTTGGTGTAATGCACGCACGACAGCCG CCGGTTACTATAAAAGGCTTGGTTTTAGTGAGCAGGGC GAAATTTTCGACACCCCGCCGGTTGGGCCGCACATTCTT ATGTACAAAAGAATCACT
SEQ ID NO:216	1_1H5	ATGATAGAAGTGAAACCTATTAACGCAGAGGATACTTA CGAAATTCGACACAGGATCCTGCGCCCTAATCAGCCGT TAGAGGCATGCATGTATGAAAGCGATCTGCTGCGGGGC TCGTTCCATTTGGGCGGGTTCTATCGTGGCAAATTGATC TCGATTGCGAGTTTCCACCAAGCTGAACACTCAGACCTG GAAGGGCAAAAGCAGTATCAATTACGAGGGATGGCGA CCCTCGAAGGATACCGTGATCAGAAGGCTGGCTCTTCG

		CTTATTAGGCACGCCGAGCAGATACTACGGAAAAGAGG GGCAGATCTGCTTTGGTGCAATGCACGCACGACAGCCG CCGGTTACTATAAAAGGCTTGGTTTTAGTGAGCAGGGC GAAGTTTTTCGACACCCCGCCGGTTGGGCCGCACATTCTT ATGTACAAAAAACTCACT
SEQ ID NO:217	1_2A12	ATGATAGAAGTGAAACCTATTAACGCAGAGGATACTTA CGAACTTCGACACAGGATCCTGCGCCCTAATCAGCCGA TAGAGGCATGCATGTATGAAAGCGATCTGCTGCGGGGC TCGTTCCATTTGGGCGGGTTCTATCGTGGCAAATTGATC TCGATTGCGAGTTTCCACCAAGCTGAACAGTCAGAACT GGAAGGGCAAAAGCAGTATCAATTACGAGGGATGGCG ACCTCGAAGGATACCGTGATCAGAAGGCTGGCTCTAC GCTTATTAAGCACGCCGAGGAGATACTACGGAAAAAAG GGGCAGATCTGCTTTGGTGCAATGCACGCACGTCAGCC GCCGGTTACTATAAAAGGCTTGGTTTTAGTGAGCAGGG CGAAATTTTCGACACCCCGCCGGTTGGGCCGCACATTCT TATGTACAAAAGACTCACT
SEQ ID NO:218	1_2B6	ATGATAGAAGTGAAACCTATTAACGCAGAGGAGACTTA CGAACTTCGACACAAGATCCTGCGCCCTAATCAGCCGTT AGAGGCATGCATGTATGAAACCGATCTGCTGCGGGGCT CGTTCCATTTGGGCGGGTTCTATCGTGGCAAATTGATCT CGATTGCGAGTTTCCACCAAGCTGAACACTCAGAACTG GAAGGGCAAAAGCAGTATCAATTACGAGGGATGGCGA CCCTCGAAGGATTCCGTGATCAGAAGGCTGGCTCTTCGC TTATTAAGCACGCCGAGGAGATACTACGGAAAAGAGGG GCAGATCTGCTTTGGTGCAATGCACGCACGTCAGCCTCC GGTTACTATAAAAAGCTTGGTTTTAGTGAGCAGGGCGA AATTTTCGAAACCCCGCCGGTTGGGCCGCACATTCTTAT GTACAAAAGACTCACT
SEQ ID NO:219	1_2C4	ATGCTAGAAGTGAAACCTATTAACGCAGAGGAGACTTA CGAACTTCGACACAAGATCCTGCGCCCTAATCAGCCGA TAGAGGCATGCATGTATGAAACCGATCTGCTGCGGGGC TCGTTCCATTTGGGCGGGTTCTATCGTGGCCAATTGATC TCGATTGCGAGTTTCCACCAAGCTGAACACTCAGACCTG CAAGGGCAAAAGCAGTATCAATTACGAGGGATGGCGAC CCTCGAAGGATACCGTGAGCAGAAGGCTGGCTCTACGC TTATTAAGCACGCCGAGGAGCTACTACGGAAAAAAGGG GCAGATCTGCTTTGGTGCAATGCACGCACGACAGCCGC CGGTTACTATAAAAAGCTTGGTTTTAGTGAGCAGGGCG AAGTTTTTCGACACCCCGCCGGTTGGGCCGCACATTCTTA TGTACAAAAAAATCACT
SEQ ID NO:220	1_2D2	ATGATAGAAGTGAAACCTATTAACGCAGAGGATACTTA CGAACTTCGACACAAGATCCTGCGCCCTAATCAGCCGTT AGAGGCATGCATGTATGAAAGCGATCTGCTGCGGAGCG CATTCCATTTGGGCGGGTTCTATCGTGGCAAATTGATCT CGATTGCGAGTTTCCACAAAGCTGAACACTCAGAACTG CAAGGGCAAAAGCAGTATCAATTACGAGGGATGGCGAC CCTCGAAGGATACCGTGATCAGAAGGCTGGCTCTTCGC TTATTAGGCACGCCGAGGAGATACTACGGAAAAGAGGG GCAGATATGCTTTGGTGCAATGCACGCACGTCAGCCGC

		CGGTTACTATAAAAGGCTTGGTTTTAGTGAGCAGGGCG AAGTTTTTCGACACCCCGCGGTTGGGCCGCACATTCTTA TGTACAAAAGAATCACTTAA
SEQ ID NO:221	1_2D4	ATGATAGAAAGTGAAACCTATTAACGCAGAGGATACTTA CGAACTTCGACACAGGATCCTGCGCCCTAATCAGCCGA TAGAGGCATGCATGTATGAAAGCGATCTGCTGCGGGGC TCGTTCCATTTGGGCGGGTTCTATCGTGGCAAATTGATC TCGATTGCGAGTTTCCACCAAGCTGAACACTCAGACCTG CAAGGGCAAAGCAGTATCAATTACGAGGGATGGCGAC CCTCGAAGGATACCGTGAGCAGAAGGCTGGCTCTTCGC TTATTAAGCACGCCGAGCAGCTACTACGGAAAAAGGG GCAGATATGCTTTGGTGTAATGCACGCACGTCAGCCGC CGGTTACTATAAAAGGCTTGGTTTTAGTGAGCAGGGCG AAATTTTCGAAACCCCGCGGTTGGGCCGCACATTCTTA TGTACAAAAGAATCACT
SEQ ID NO:222	1_2F8	ATGCTAGAAAGTGAAACCTATTAACGCAGAGGATACTTA CGAACTTCGACACAGGATCCTGCGCCCTAATCAGCCGTT AGAGGCATGCATGTATGAAACCGATCTGCTGCGGGGCT CGTTCCATTTGGGCGGGTTCTATCGTGGCAAATTGATCT CGATTGCGAGTTTCCACCAAGCTGAACATTCAAGACTG GAAGGGCAAAGCAGTATCAATTACGAGGGATGGCGA CTCTCGAAGGATACCGTGATCAGAAGGCTGGCTCTTCG CTTATTAGGCACGCCGAGGAGATACTACGGAAAAGAGG GGCAGATATGCTTTGGTGCAATGCACGCACGACAGCCG CCGGTTACTATAAAAAGCTTGGTTTTAGTGAGCAGGGC GAAATTTACGACACCCCGCGGTTGGGCCGCACATTCTT ATGTACAAAAAACTCACT
SEQ ID NO:223	1_2H8	ATGATAGAAAGTGAAACCTATTAACGCAGAGGAGACTTA CGAACTTCGACACAAGATCCTGCGCCCTAATCAGCCGTT AGAGGCATGCATGTATGAAACCGATCTGCTGCGGGGCG CGTTCCATTTGGGCGGGTTCTATCGTGGCAAATTGATCT CGATTGCGAGTTTCCACCAAGCTGACCACTCAGAACTG CAAGGGCAAAGCAGTATCAATTACGAGGGATGGCGAC CCTCGAAGGATACCGTGAGCAGAAGGCTGGCTCTACGC TTATTAGGCACGCCGAGCAGATACTACGGAAAAGAGGG GCAGATCTACTTTGGTGCAATGCACGCACGTCAGCCGC CGGTTACTATAAAAAGCTTGGTTTTAGTGAGCAGGGCG AAATTTTCGAAACCCCGCGGTTGGGCCGCACATTCTTA TGTACAAAAGACTCACTTAA
SEQ ID NO:224	1_3A2	ATGATAGAAAGTGAAACCTATTAACGCAGAGGATACTTA CGAACTTCGACACAGGATCCTGCGCCCTAATCAGCCGA TAGAGGCATGCATGTATGAAAGCGATCTGCTGCGGGGC GCGTTCCATTTGGGCGGGTTCTATCGTGGCAAATTGATC TCGATTGCGAGTTTCCACCAAGCTGAACACTCAGACCTG CAAGGGCAAAGCAGTATCAATTACGAGGGATGGCGAC CCTCGAAGGATACCGTGAGCAGAAGGCTGGCTCTTCGC TTATTAGGCACGCCGAGGAGATACTACGGAAAAAGGG GCAGATATGCTTTGGTGCAATGCACGCACGACAGCCGC CGGTTACTATAAAAAGGCTTGGTTTTAGTGAGCAGGGCG AAGTTTTTCGACACCCCGCGGTTGGGCCGCACATTCTTA

		TGTACAAAAGAATCACT
SEQ ID NO:225	1_3D6	ATGATAGAGGTGAAACCGATTAACGCAGAGGATACCTA TGAAC TAAGGCATAGAATACTCAGACCAAACAGCCGA TAGAAGTGTGTATGTATGAAACCGATTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGAGTCT AATTAACACGCTGAACAAATTCTTCGTAAGAGGGGGG CGGACTTGCTTTGGTGCAATGCGCGGACATCCGCCTCAG GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGAG GTATTTGATACGCCGCCAGTAGGACCTCATCTCTGATG TATAAAGGCTCACATAA
SEQ ID NO:226	1_3F3	ATGATAGAAGTGAAACCTATTAACGCAGAGGAGACTTA CGAACTTCGACAGAGGATCCTGCGCCCTAATCAGCCGA TAGAGGCATGCATGTATGAAAGCGATCTGCTGCGGGGC TCGTTCCATTTGGGCGGGTTCTATCGTGGCCAATTGATC TCGATTGCGAGTTTCCACCAAGCTGAACACTCAGAACT GCAAGGGCAAAAGCAGTATCAATTACGAGGGATGGCG ACCCTCGAAGGATACCGTGAGCAGAAGGCTGGCTCTAC GCTTATTAAGCACGCCGAGGAGATACTACGGAAAAAAG GGGCAGATCTGCTTTGGTGCAATGCACGCACGTCAGCC GCCGGTTACTATAAAAGGCTTGGTTTTAGTGAGCACGG CGAAATTTTCGACACCCCGCCGGTTGGGCCGCACATTCT TATGTACAAAAGAATCACT
SEQ ID NO:227	1_3H2	ATGATAGAAGTGAAACCTATTAACGCAGAGGATACTTA CGAACTTCGACACAGGATCCTGCGCCCTAATCAGCCGA TAGAGGCATGCATGTATGAAACCGATCTGCTGCGGGGC GCGTTCCATTTGGGCGGGTACTATCGTGGCCAATTGATC TCGATTGCGAGTTTCCACAAAGCTGAACACTCAGAACT GCAAGGGCAAAAGCAGTATCAATTACGAGGGATGGCG ACCCTCGAAGGATACCGTGAGCAGAAGGCTGGCTCTAC GCTTATTAAGCACGCCGAGCAGCTACTACGGGAAAAAAG GGGCAGATATGCTTTGGTGCAATGCACGCACGTCAGCC GCCGGTTACTATAAAAGGCTTGGTTTTAGTGAGCAGGG CGAAGTTTTTCGACACCCCGCCGGTTGGGCCGCACATTCT TATGTACAAAAAATCACT
SEQ ID NO:228	1_4C5	ATGATAGAAGTGAAACCTATTAACGCAGAGGATACTTA CGAACTTCGACACAAGATCCTGCGCCCTAATCAGCCGA TAGAGGCATGCATGTATGAAAGCGATCTGCTGCGGGGC TCGTTCCATTTGGGCGGGTTCTATCGTGGCAAATTGATC TCGATTGCGAGTTTCCACAAAGCTGAACACTCAGACCT GGAAGGGCAAAACCAGTATCAATTACGAGGGATGGCG ACCCTCGAAGGATACCGTGAGCAGAAGGCTGGCTCTAC GCTTATTAGGCACGCCGAGGAGATACTACGGAAAAAGAG GGGCAGATATGCTTTGGTGCAATGCACGCACGTCAGCC TCCGGTTACTATAAAAGGCTTGGTTTTAGTGAGCACGGC GAAATTTTCGACACCCCGCCGGTTGGGCCGCACATTCTT ATGTACAAAAGACTCACTTAA
SEQ ID	1_4D6	ATGCTAGAAGTGAAACCTATTAACGCAGAGGATACTTA

NO:229		CGAACTTCGACACAGGATCCTGCGCCCTAATCAGCCGA TAGAGGCATGCATGTATGAAACCGATCTGCTGCGGGGC TCGTTCCATTTGGGCGGGTTCTATCGTGGCCAATTGATC TCGATTGCGAGTTTCCACAAAGCTGAACACTCAGACCT GGAAGGGCAAAAGCAGTATCAATTACGAGGGATGGCG ACCCTCGAAGGATACCGTGAGCAGAAGGCTGGCTCTAC GCTTATTAGGCACGCCGAGCAGATACTACGGAAAAGAG GGGCAGATATGCTCTGGTGCAATGCACGCACGTCAGCC GCCGGTTACTATAAAAGGCTTGGTTTTAGTGAGCAGGG CGAAGTTTTCGAAACCCCGCCGGTTGGGCCGCACATTCT TATGTACAAAAGACTCACT
SEQ ID NO:230	1_4H1	ATGATAGAAGTGAAACCTATTAACGCAGAGGATACTTA CGAACTTCGACACAGGATCCTGCGCCCTAATCAGCCGTT AGAGGCATGCATGTATGAAACCGATCTGCTGCGGGGCT CGTTCCATTTGGGCGGGTTCTATCGTGGCAAATTGATCT CGATTGCGAGTTTCCACCAAGCTGAACACTCAGACCTG CAAGGGCAAAAGCAGTATCAATTACGAGGGATGGCGAC CCTCGAAGGATACCGTGAGCAGAAGGCTGGCTCTACGC TTATTAGGCACGCCGAGCAGCTACTACGGAAAAGAGGG GCAGATCTGCTTTGGTGCAATGCACGCACGTCAGCCTCC GGTACTATAAAAGGCTTGGTTTTAGTGAGCACGGCGA AGTTTTCGACACCCCGCCGGTTGGGCCGCACATTCTTAT GTACAAAAGACTCACT
SEQ ID NO:231	1_5H5	ATGCTAGAAGTGAAACCTATTAACGCAGAGGAGACTTA CGAACTTCGACACAAGATCCTGCGCCCTAATCAGCCGTT AGAGGCATGCATGTATGAAAGCGATCTGCTGCGGGGCT CGTTCCATTTGGGCGGGTACTATCGTGGCCAATTGATCT CGATTGCGAGTTTCCACCAAGCTGAACACTCAGAACTG GAAGGGCAAAAGCAGTATCAATTACGAGGGATGGCGA CCCTCGAAGGATTCCGTGAGCAGAAGGCTGGCTCTACG CTTATTAAGCACGCCGAGCAGATACTACGGAAAAGAGG GGCAGATATGCTTTGGTGCAATGCACGCACGTCAGCCG CCGGTTACTATAAAAAGCTTGGTTTTAGTGAGCACGGC GAAATTTTCGACACCCCGCCGGTTGGGCCGCACATTCTT ATGTACAAAAAACTCACTTAA
SEQ ID NO:232	1_6F12	ATGATAGAAGTGAAACCTATTAACGCAGAGGAGACTTA CGAACTTCGACACAGGATCCTGCGCCCTAATCAGCCGA TAGAGGCATGCATGTATGAAAGCGATCTGCTGCGGGGC TCGTTCCATTTGGGCGGGTTCTATCGTGGCAAATTGATC TCGATTGCGAGTTTCCACCAAGCTGAACACTCAGACCTA GAAGGGCAAAAGCAGTATCAATTACGAGGGATGGCGA CCCTCGAAGGATACCGTGATCAGAAGGCTGGCTCTACG CTTATTAAGCACGCCGAGGAGCTACTACGGAAAAGAGG GGCAGATATGCTTTGGTGCAATGCACGCACGTCAGCCG CCGGTTACTATAAAAGGCTTGGTTTTAGTGAGCACGGC GAAATTTACGAAACCCCGCCGGTTGGGCCGCACATTCTT ATGTACAAAAAAATCACT
SEQ ID NO:233	1_6H6	ATGATAGAAGTGAAACCTATTAACGCAGAGGATACTTA CGAACTTCGACACAAGATCCTGCGCCCTAATCAGCCGA TAGAGGCATGCATGTATGAAAGCGATCTGCTGCGGGGC

		TCGTTCCATTTGGGCGGGTTCTATCGTGGCCAATTGATC TCGATTGCGAGTTTCCACCAAGCTGAACACTCAGACCTG GAAGGGCAAAGCAGTATCAATTACGAGGGATGGCGA CCCTCGAAGGATACCGTGATCAGAAGGCTGGCTCTTCG CTTATTAAGCACGCCGAGGAGATACTACGGAAAAGAGG GGCAGATCTGCTTTGGTGCAATGCACGCACGTCAGCCG CCGGTTACTATAAAAGGCTTGGTTTTAGTGAGCAGGGC GAAATTTTCGACACCCCGCGGTTGGGCCGCACATTCTT ATGTACAAAAAATCACT
SEQ ID NO:234	3_11A10	ATGCTAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAAC TAAGGCATAGAACTCAGACCAAACCAGCCGA TAGAAGCGTGTATGTATGAAAGCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTATTACAGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGAACTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGAGTCT AGTTAAACACGCTGAAGAAATTCCTCGTAAGAGGGGGG CGGACTTGCTTTGGTGTAATGCGCGGACATCCGCCTCAG GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGAG ATATTTGAAACGCCGCCAGTAGGACCTCACATCCTGAT GTATAAAAGGATCACATAA
SEQ ID NO:235	3_14F6	ATGCTAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAAC TAAGGCATAGAACTCAGACCAAACCAGCCGA TAGAAGCGTGTATGTATGAAAGCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGAACTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGAGTCT AATTAACACGCTGAAGAAATTCCTCGTAAGAGGGGGG CGGACTTGCTTTGGTGTAATGCGCGGACGTCCGCCTCAG GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGAG ATATTTGAAACGCCGCCAGTAGGACCTCACATCCTGAT GTATAAAAGGCTCACATAA
SEQ ID NO:236	3_15B2	ATGCTAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAAC TAAGGCATAAAATACTCAGACCAAACCAGCCGT TAGAAGTGTGTATGTATGAAACCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTATTACGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGAACTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGAGTCT AATTAACACGCTGAAGAAATTCCTCGTAAGAGGGGGG CGGACTTGCTTTGGTGTAATGCGCGGACATCCGCCTCAG GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGAG ATATTTGAAACGCCGCCAGTAGGACCTCACATCCTGAT GTATAAAAGGATCACATAA
SEQ ID NO:237	3_6A10	ATGATAGAAGTGAAACCGATTAAACGCAGAGGATACCTA TGAAC TAAGGCATAGAACTCAGACCAAACCAGCCGA TAGAAGCGTGTATGTATGAAAGCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTATTACAGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGAACTCC

		AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGAGTCT AATTAAACACGCTGAAGAAATTCTTCGTAAGAGGGGGG CGGACTTGCTTTGGTGTAAATGCGCGGACATCCGCCTCAG GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGAG ATATTTGAAACGCCGCCAGTAGGACCTCACATCCTGAT GTATAAAAGGATCACATAA
SEQ ID NO:238	3_6B1	ATGCTAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAACCTAAGGCATAGAATACTCAGACCAAACCAGCCGA TAGAAGCGTGTATGTATGAAAGCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTATTACAGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCAACCAGAACTC CAAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTAC CTTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGAGTC TAATTAAACACGCTGAAGAAATTCTTCGTAAGAGGGGG GCGGACTTGCTTTGGTGTAAATGCGCGGACATCCGCCTCA GGCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGA GGTATTTGAAACGCCGCCAGTAGGACCTCACATCCTGA TGTATAAAAGGATCACATAA
SEQ ID NO:239	3_7F9	ATGCTAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAACCTAAGGCATAGAATACTCAGACCAAACCAGCCGA TAGAAGCGTGTATGTATGAAAGCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTATTACGGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGAGTCT AATTAAACACGCTGAAGAAATTCTTCGTAAGAGGGGGG CGGACTTGCTTTGGTGTAAATGCGCGGACATCCGCCTCAG GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGAG ATATTTGAAACGCCGCCAGTAGGACCTCACATCCTGAT GTATAAAAGGATCACATAA
SEQ ID NO:240	3_8G11	ATGCTAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAACCTAAGGCATAGAATACTCAGACCAAACCAGCCGA TAGAAGTGTGTATGTATGAAAGCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTATTACAGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGAACTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGAGTCT AATTAAACACGCTGAAGAAATTCTTCGTAAGAGGGGGG CGGACTTGCTTTGGTGTAAATGCGCGGACATCCGCCTCAG GCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGAG ATATTTGAAACGCCGCCAGTAGGACCTCACATCCTGAT GTATAAAAGGATCACATAA
SEQ ID NO:241	4_1B10	ATGATAGAAGTGAAACCTATTAACGCAGAGGATACCTA TGAACCTAAGGCATAGAATACTCAGACCAAACCAGCCGA TAGAAGTGTGTATGTATGAAACCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACGGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGATCAGAAAGCGGGATCGAGTCT



		AATTAGACACGCTGAACAAATTCCTTCGTAAGAGGGGGG CGGACATGCTTTGGTGCAATGCGCGGACATCCGCCTCA GGCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGA GATATTTGAAACGCCGCCAGTAGGACCTCACATCCTGA TGTATAAAAGGATCACATAA
SEQ ID NO:242	5_2B3	ATGATAGAAGTGAAACCTATTAACGCAGAGGATACCTA TGAAC TAAGGCATAGAATACTCAGACCAAACCAGCCGT TAGAAGTGTGTATGTATGAAACCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGATCAGAAAGCGGGATCGAGTCT AATTAGACACGCTGAACAAATTCCTTCGTAAGAGGGGGG CGGACATGCTTTGGTGTAATGCGCGGACATCCGCCTCA GGCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGA GATATTTGAAACGCCGCCAGTAGGACCTCACATCCTGA TGTATAAAAGGATCACATAA
SEQ ID NO:243	5_2D9	ATGCTAGANGTGAAACCGATTAACGCAGAGGATACCTA TGAAC TAAGGCATAAAAATACTCAGACCAAACCAGCCGN TAGAAGTGTGTATGTATGAAANCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGATCAGAAAGCGGGATCGAGTCT AATTAACACGCTGAACAAATTCCTTCGTGAGAGGGGGG CGGACATGCTTTGGTGCAATGCGCGGACATCCGCCTCA GGCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGA GGTATTTGACACGCCGCCAGTAGGACCTCACATCCTGAT GTATAAAAGGCTCACATAA
SEQ ID NO:244	5_2F10	ATGCTAGAAGTGAAACCTATTAACGCAGAGGATACCTA TGAAC TAAGGCATAAAAATACTCAGACCAAACCAGCCGA TAGAAGTGTGTATGTATGAAACCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGATCAGAAAGCGGGATCGAGTCT AATTAGACACGCTGAACAAATTCCTTCGTAAGAGGGGGG CGGACATGCTTTGGTGCAATGCGCGGACATCCGCCTCA GGCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGAGA GATATTTGAAACGCCGCCAGTAGGACCTCACATCCTGA TGTATAAAAGGCTCACATAA
SEQ ID NO:245	6_1A11	ATGCTAGAGGTGAAACCGATTAACGCAGAGGATACCTA TGAAC TAAGGCATAAAAATACTCAGACCAAACCAGCCGT TAGAAGTGTGTATGTATGAAACCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGGCAAACCTGATTT CCATAGCGTCATTCCACCAGGCCGAGCACTCAGACCTC CAAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTAC CTTGGAAGGTTATCGTGATCAGAAAGCGGGATCGAGTC TAATTAGACACGCTGAACAAATTCCTTCGTAAGAGGGGG GCGGACATGCTTTGGTGCAATGCGCGGACATCCGCCTC

		AGGCTACTACAGAAAGTTAGGCTTCAGCGAGCAGGGAG AGGTATTTGAAACGCCGCCAGTAGGACCTCACATCCTG ATGTATAAAAGGCTCACATAA
SEQ ID NO:246	6_1D5	ATGATAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAACCTAAGGCATAAAATACTCAGACCAAACAGCCGT TAGAAGTGTGTATGTATGAAACCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGATCAGAAAGCGGGATCGAGTCT AATTAGACACGCTGAACAAATTCTTCGTAAGAGGGGGG CGGACATGCTTTGGTGCAATGCGCGGACATCCGCCTCA GGCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGGGA GGTATTTGAAACGCCGCCAGTAGGACCTCACATCCTGA TGTATAAAAGGATCACATAA
SEQ ID NO:247	6_1F11	ATGATAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAACCTAAGGCATAAAATACTCAGACCAAACAGCCGT TAGAAGTGTGTATGTATGAAACCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGAGTCT AATTAGACACGCTGAACAAATTCTTCGTAAGAGGGGGG CGGACATGCTTTGGTGCAATGCGCGGACATCCGCCTCA GGCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGGAGA GGTATTTGAAACGCCGCCAGTAGGACCTCACATCCTGA TGTATAAAAGGCTCACATAA
SEQ ID NO:248	6_1F1	ATGATAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAACCTAAGGCATAAAATACTCAGACCAAACAGCCGT TAGAAGTGTGTATGTATGAAACCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGAGTCT AATTAGACACGCTGAACAAATTCTTCGTAAGAGGGGGG CGGACATGCTTTGGTGCAATGCGCGGACATCCGCCTCA GGCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGGAGA GGTATTTGAAACGCCGCCAGTAGGACCTCACATCCTGA TGTATAAAAGGCTCACATAA
SEQ ID NO:249	6_1H10	ATGCTAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAACCTAAGGCATAAAATACTCAGACCAAACAGCCGT TAGAAGTGTGTATGTATGAAACCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACGGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCGGACCTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGATCAGAAAGCGGGATCGAGTCT AATTAGACACGCTGAAGAAATTCTTCGTAAGAGGGGGG CGGACATGCTTTGGTGCAATGCGCGGACATCCGCCTCA GGCTACTACAAAAAGTTAGGCTTCAGCGAGCAGGGGAGA GGTATTTGACACGCCGCCAGTAGGACCTCACATCCTGAT

		GTATAAAAAGATCACATAA
SEQ ID NO:250	6_1H4	ATGCTAGAAAGTGAAACCGATTAAACGCAGAGGATACCTA TGAACATAAGGCATAAAATACTCAGACCAAACCGCCGT TAGAAGTGTGTATGTATGAAACCGATTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGATCAGAAAGCGGGATCGACTCT AATTAACACGCTGAACAAATTCTTCGTAAGAGGGGGG CGGACATGCTTTGGTGCAATGCGCGGACATCCGCCTCA GGCTACTACAAAAGTTAGGCTTCAGCGAGCAGGGAGA GGTATTTGAAACGCCGCCAGTAGGACCTCACATCCTGA TGTATAAAAGGCTCACATAA
SEQ ID NO:251	8_1F8	ATGATAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAACATAAGGCATAGAATACTCAGACCAAACCGCCGT TAGAAGTGTGTATGTATGAAACCGATTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGACCTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGAGTCT AATTAACACGCTGAAGAAATTCTTCGTAAGAGGGGGG CGGACTTGCTTTGGTGTAATGCGCGGACATCCGCCTCAG GCTACTACAAAAGTTAGGCTTCAGCGAGCAGGGAGAG ATATTTGATACGCCGCCAGTAGGACCTCACATCCTGATG TATAAAAGGATCACATAA
SEQ ID NO:252	8_1G2	ATGATAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAACATAAGGCATAGAGTACTCAGACCAAACCGCCGT TAGAAGTGTGTATGTATGAAACCGATTACTTCGTGGTG CATTTCACTTAGGCGGCTATTACAGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGAACTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGAGTCT AATTAACACGCTGAAGAAATTCTTCGTAAGAGGGGGG CGGACTTGCTTTGGTGCAATGCGCGGACATCCGCCTCAG GCTACTACAAAAGTTAGGCTTCAGCGAGCAGGGAGAG GTATTTGAGACGCCGCCAGTAGGACCTCACATCCTGAT GTATAAAAGGCTCACGTAA
SEQ ID NO:253	8_1G3	ATGCTAGAGGTGAAACCGATTAAACGCAGAGGATACTTA CGAACTAAGGCATAAAATACTCAGACCAAACCGCCGA TAGAAGTGTGTATGTATGAAACCGATTACTTCGTGGTG CATTTCACTTAGGCGGCTATTACAGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGAACTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGAGTCT AATTAGACACGCTGAAGAAATTCTTCGTAAGAGGGGGG CGGACTTGCTTTGGTGTAATGCGCGGACATCCGCCTCAG GCTACTACAAAAGTTAGGCTTCAGCGAGCAGGGAGAG ATATTTGATACGCCGCCAGTAGGACCTCACATCCTGATG TATAAAAGGATCACGTAA
SEQ ID	8_1H7	ATGCTAGAGGTGAAACCGATTAAACGCAGAGGATACCTA

NO:254		TGAACTAAGGCATAGAATACTCAGACCAAACCAGCCGA TAGAAGTGTGTATGTATGAAACCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTTTTACAGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGAACTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGAGTCT AATTAAACACGCTGAAGAAATTCCTTCGTAAGAGGGGGG CGGACATGCTTTGGTGCAATGCGCGGACATCCGCCTCA GGCTACTACAAAAGTTAGGCTTCAGCGAGCAGGGAGA GATATTTGAAACGCCGCCAGTAGGACCTCACATCCTGA TGTATAAAAGGCTCACATAA
SEQ ID NO:255	8_1H9	ATGCTAGAGGTGAAACCGATTAAACGCAGAGGATACCTA TGAACTAAGGCATAAAATACTCAGACCAAACCAGCCGT TAGAAGTGTGTATGTATGAAACCGATTTACTTCGTGGTG CATTTCACTTAGGCGGCTATTACAGGGGGCAAACCTGATTT CCATAGCTTCATTCCACCAGGCCGAGCACTCAGAACTCC AAGGCCAGAAACAGTACCAGCTCCGAGGTATGGCTACC TTGGAAGGTTATCGTGAGCAGAAAGCGGGATCGAGTCT AATTAGACACGCTGAAGAAATTCCTTCGTAAGAGGGGGG CGGACTTGCTTTGGTGTAATGCGCGGACATCCGCCTCAG GCTACTACAAAAGTTAGGCTTCAGCGAGCAGGGAGAG GTATTTGATACGCCGCCAGTAGGACCTCACATCCTGATG TATAAAAGGCTCACATAA
SEQ ID NO:256	GAT1_21F 12	ATGATTGAAGTCAAACCTATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTTGCTCGGGGGC ACGTTTCACCTCGGCGGATATTACCGGGGGCAAGCTGAT CAGCATCGCTTCCTTTTATAATGCCGAACATTCAGAGCT TGAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCG ACGCTTGAAGGATACCGTGAGCAAAAAGCGGGAAGCA CGCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAA GGCGCGGACCTTTTATGGTGCAACGCCAGGACATCTGT GAGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGG GCGAAGTCTACGACATACCGCCGATCGGACCTCATATTT TGATGTATAAGAAATTGACGTAA
SEQ ID NO:257	GAT1_24G 3	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGGCAAGCTGATC AGCATCGCCTCCTTTCATCAAGCCGAACATTCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGGG GGCAGACCTTTTATGGTGCAATGCCAGGACATTTGTGA GCGGTTACTATGAAAAGCTCGGTTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGATCGGACCTTATATTTG ATGTATTAGAAATTGACATAA
SEQ ID NO:258	GAT1_29G 1	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTTGCTCGGGGGT

		ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATTAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGCTAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG GGCAGACCTTTTATGGTGCAACGCCAGGACATCTGTGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGC GGGGTCTGCGATATACCGCCGATCGGACCTCATATTTTG ATGTATAAGAAATTGGCATAA
SEQ ID NO:259	GAT1_32G 1	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATACTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCAGACCTTTTATGGTGCAACGCCAGGACATCTGTGA GCGGCTACTATGAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTACGACATACCGCCGATCGGACCTCATATTTTG ATGTATAAGAAATTGACATAA
SEQ ID NO:260	GAT2_15G 8	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TGGAAGCATGCAAGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATAATGCCGAACATTAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CGCTTGAAGGGTACCGCGAGCAAAAAGCGGGAAGCAC GCTCATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAG GCGCAGACCTTTTATGGTGCAACGCCAGGACATCTGTG AGCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAGGG CGAAGTCTACGACATACCGCCGATCGGACCTCATATTTT GATGTATAAGAAATTGACGTAA
SEQ ID NO:261	GAT2_19H 8	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATACTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATCCAGAGCTT GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGGTACCGCGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAAGG CGCAGACCTTTTATGGTGCAACGCCAGGACATCTGTGA GCGGCTACTATGAAAAGCTCGGCTTCAGCGAACAGGGC GAAGTCTGCGACATACCGCCGATCGGACCTCATATTTTG ATGTATAAGAAATTGACATAA
SEQ ID NO:262	GAT2_21F 1	ATGATTGAAGTCAAACCAATAAACGCGGAAGATACGTA TGAGATCAGGCACCGCATTCTCCGGCCGAATCAGCCGC TTGAAGCATGTATGTATGAAACCGATTGCTCGGGGGC ACGTTTCACCTCGGTGGATATTACCGGGGCAAGCTGATC AGCATCGCTTCCTTTCATCAAGCCGAACATTAGAGCTT

		GAAGGCCAAAAACAGTATCAGCTGAGAGGGATGGCGA CACTTGAAGGATACCGTGAGCAAAAAGCGGGCAGTACG CTTATCCGCCATGCCGAAGAGCTTCTTCGGAAAAAGGG GGCAGACCTTTTATGGTGCAACGCCAGGAATCTGTGA GCGGGTACTATAAAAAGCTCGGCTTCAGCGAACAAGGC GGGGTCTACGATATACCGCCGATCGGACCTCATATTTG ATGTATAAGAAATTGACGTAA
SEQ ID NO:263	13_10F6	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLRGTFH LGGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPVGP HILMYKKLT
SEQ ID NO:264	13_12G6	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLRGAFH LGGYYRGKLVSIASFHQAEPHELEGQRQYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK KLGFSEQGEVYDIPPTGP HILMYKKLT
SEQ ID NO:265	14_2A5	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGSTFHL GGYYRGKLISIASFNQAEPHELEGQKQYQLRGMATLEGYR EQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKKL GFSEQGEVYDTPVGP HILMYKKLT
SEQ ID NO:266	14_2C1	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLRGAFH LGGYYRGKLVSIASFHQAEPHELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK KLGFSEQGEVYDTPPTGP HILMYKKLT
SEQ ID NO:267	14_2F11	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLRGAFH LGGYYRGKLVSIASFHQAEPHELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYKK KLGFSEQGEVYDTPPAGP HILMYKKLT
SEQ ID NO:268	CHIMERA	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLRGAFH LGGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDTPVGP HILMYKKLT
SEQ ID NO:269	10_12D7	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTLH LGGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLEEY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPTGP HILMYKKLT
SEQ ID NO:270	10_15F4	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLRGTFH LGGYYRGKLVSIASFHQAEPHELEGQKQYQLRGMATLEE YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK KLGFSEQGGVYDIPVGP HILMYKKLT
SEQ ID NO:271	10_17D1	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDTPVGP HILMYKKLT
SEQ ID NO:272	10_17F6	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLVSIASFHQAEPHELEGQKQYQLRGMATLEE YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK KLGFSEQGEVYDIPVGP HILMYKKLT
SEQ ID NO:273	10_18G9	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLVSIASFHQAEPHELEGQKQYQLRGMATLEE YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK

		KLGFSEQGGVYDIPPVGPHILMYKKLT
SEQ ID NO:274	10_1H3	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYK KLGFSEQGEVYDIPPTGPHILMYKKLT
SEQ ID NO:275	10_20D10	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTLH LGGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLECY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPPVGPHILMYKKLT
SEQ ID NO:276	10_23F2	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYK KLGFSEQGEVYDTPPVGPHILMYKKLT
SEQ ID NO:277	10_2B8	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLECY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDTPPVGPHILMYKKLT
SEQ ID NO:278	10_2C7	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLRGAFH LGGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDTPPVGPHILMYKKLT
SEQ ID NO:279	10_3G5	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYK KLGFSEQGEVYDIPPTGPHILMYKKLT
SEQ ID NO:280	10_4H7	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYK KLGFSEQGEVYDIPPTGPHILMYKKLT
SEQ ID NO:281	10_6D11	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTLH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYK KLGFSEQGEVYDIPPVGPHILMYKKLT
SEQ ID NO:282	10_8C6	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGAFH LGGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPPVGPHILMYKKLT
SEQ ID NO:283	11C3	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYQGKLISIASFHQAEPHELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYKK LGFSEQGGVYDIPPIGPHILMYKKLT
SEQ ID NO:284	11G3	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYQGKLISIASFHQAEPHELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYEK LGFSEQGGVYDIPPIGPHILMYKKLA
SEQ ID NO:285	11H3	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGAFH LGGYYQGKLISIASFHQAEPHELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVRGYK LGFSEQGGVYDIPPIGPHILMYKKLT
SEQ ID	12_1F9	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH

NO:286		LGGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVHDIPPTGPHILMYKKLT
SEQ ID NO:287	12_2G9	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLVSIASFHQAEPHELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK KLGFEQGEVYDTPPVGPHILMYKKLT
SEQ ID NO:288	12_3F1	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPPVGPHILMYKKLT
SEQ ID NO:289	12_5C10	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLECY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDAPPTGPHILMYKKLT
SEQ ID NO:290	12_6A10	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLVSIASFHQAEPHELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK KLGFEQGGVYDIPPVGPHILMYKKLT
SEQ ID NO:291	12_6D1	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLECY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPPVGPHILMYKKLT
SEQ ID NO:292	12_6F9	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPTGPHILMYKKLT
SEQ ID NO:293	12_6H6	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLVSIASFHQAEPHELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYKK KLGFEQGEVYDIPPTGPHILMYKKLT
SEQ ID NO:294	12_7D6	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPPTGPHILMYKKLT
SEQ ID NO:295	12_7G11	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDTPPVGPHILMYKKLT
SEQ ID NO:296	12F5	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYQGKLISIASFHKAEPHELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYKK LGFSEQGGIYDIPPIGPHILMYKKLT
SEQ ID NO:297	12G7	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYQGKLISIASFHKAEPHELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYKK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:298	1_2H6	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGAFH LGGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK



		LGFEQGGVYDIPPVGP HILMYKKLT
SEQ ID NO:299	13_12G12	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFNQAEHPELEGQKQYQLRGMATLEEY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFEQGEVYDIPPVGP HILMYKKLT
SEQ ID NO:300	13_6D10	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDSLGGTFH LGGYYRGKLISIASFNQAEHPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFEQGEVYDTPPVGP HILMYKKLT
SEQ ID NO:301	13_7A7	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLRGA FH LGGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEEY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFEQGEVYDTPPVGP HILMYKKLT
SEQ ID NO:302	13_7B12	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGSTFHL GGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEGYR EQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKKL GFSEQGEVYDIPPTGP HILMYKKLT
SEQ ID NO:303	13_7C1	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLRGA FH LGGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSARGYYKK LGFEQGEVYDIPPTGP HILMYKKLT
SEQ ID NO:304	13_8G6	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDSLGGTFH LGGYYRGKLISIASFNQAEHPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFEQGGVYDIPPVGP HILMYKKLT
SEQ ID NO:305	13_9F6	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEEY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFEQGEVYDIPPVGP HILMYKKLT
SEQ ID NO:306	14_10C9	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLRGA FH LGGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEEY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFEQGEVYDTPPVGP HILMYKKLT
SEQ ID NO:307	14_10H3	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLRGA FH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEE YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYK KLGFSEQGEVYDTPPVGP HILMYKKLT
SEQ ID NO:308	14_10H9	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLRGA FH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYK KLGFSEQGEVYDTPPVGP HILMYKKLT
SEQ ID NO:309	14_11C2	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGSTFHL GGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEEY REQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYKK LGFEQGEVYDTPPTGP HILMYKKLT
SEQ ID NO:310	14_12D8	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYK KLGFREQGGVYDIPPVGP HILMYKKLT
SEQ ID	14_12H6	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGA FH

NO:311		LGGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEEY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPTGPHILMYKKLT
SEQ ID NO:312	14_2B6	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFNQAEHPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPPVGPVPHILMYKKLT
SEQ ID NO:313	14_2G11	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLRGAFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK KLGFSSEQGEVYDIPPTGPHILMYKKLT
SEQ ID NO:314	14_3B2	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLRGAFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYKK KLGFSSEQGGVYDIPPAGPHILMYKKLT
SEQ ID NO:315	14_4H8	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGSTFHL GGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEGYR EQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKKL GFSEQGEVYDTPPVGPVPHILMYKKLT
SEQ ID NO:316	14_6A8	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLVSIASFNQAEHPELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK KLGFSSEQGEVYDTPPVGPVPHVLMYKKLT
SEQ ID NO:317	14_6B10	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDMPVGPVPHILMYKKLT
SEQ ID NO:318	14_6D4	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFNQAEHPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDTPPVGPVPHILMYKKLT
SEQ ID NO:319	14_7A11	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLRGAFH LGGYYRGKLVSIASFHQAEHPELEGKQYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK KLGFSSEQGEVYDTPPTGPHILMYKKLT
SEQ ID NO:320	14_7A1	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLRGTFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEE YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK KLGFSSEQGEVYDTPPAGPHILMYKKLT
SEQ ID NO:321	14_7A9	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLVSIASFHQAKHPELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK KLGFSSEQGEVYDTPPVGPVPHILMYKKLT
SEQ ID NO:322	14_7G1	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLRGAFH LGGYYRGKLISIASFNQAEHPELEGQKQYQLRGMATLEEY REQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDTPPVGPVPHILMYKKLT
SEQ ID NO:323	14_7H9	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK

		KLGFSEQGEVYDIPVGP HILMYKKLT
SEQ ID NO:324	14_8F7	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEE YREQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYK KLGFSEQGEVYDIPPTGP HILMYKKLT
SEQ ID NO:325	15_10C2	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLRGAFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTTASGYK KLGFSEQGEVDFDIPPTGP HILMYKKLT
SEQ ID NO:326	15_10D6	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEE YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYK KLGFSEQGEVYDIPVGP HILMYKKLT
SEQ ID NO:327	15_11F9	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLRGAFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYK KLGFSEQGEVYDIPPTGP HILMYKKLT
SEQ ID NO:328	15_11H3	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLRGAFH LGGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYK LGFSEQGEVYDIPPTGP HILMYKKLT
SEQ ID NO:329	15_12A8	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYK LGFSEQGEVYDIPPTGP HILMYKKLT
SEQ ID NO:330	15_12D6	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLRGAFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYK KLGFSEQGEVYDTPVGP HILMYKKLT
SEQ ID NO:331	15_12D8	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYK KLGFSEQGKVYDIPVGP HILMYKKLT
SEQ ID NO:332	15_12D9	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLRGTFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEE YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYK KLGFSEQGEVYDIPVGP HILMYKKLT
SEQ ID NO:333	15_3F10	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLRGAFH LGGYYRGKLISIVSFHQAEHPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYK LGFSEQGEVYDTPPAGP HILMYTKLT
SEQ ID NO:334	15_3G11	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEE YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYK KLGFSEQGEVYDIPVGP HILMYKKLT
SEQ ID NO:335	15_4F11	MIEVKPINAEDTYKIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYK KLGFSEQGEVYDIPPTGP HILMYKKLT
SEQ ID	15_4H3	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH

NO:336		LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEE YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK KLGFEQGEVYDIPPTGPHILMYKKLT
SEQ ID NO:337	15_6D3	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEEY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPTGPHILMYKKLT
SEQ ID NO:338	15_6G11	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLRGAFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEE YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK KLGFEQGEVYDIPVGPVPHILMYKKLT
SEQ ID NO:339	15_9F6	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEEY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPVGPVPHILMYKKLT
SEQ ID NO:340	15F5	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLVSIASFHKAHSELEGEEQYQLRGMATLEGY REQKAGSTLIRYAEELLRKKGADLLWCNARTSVSGYYKK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:341	16A1	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTLH LGGYYQGKLVSIASFHKAHSGLEGEQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYEK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:342	16H3	MIDVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYQGKLVSIASFHQAEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYEK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:343	17C12	MIEVKPISAEDTYEIRHRILRPNQPLEACMYETDLLGGAFH LGGYYQGKLVSIASFHQAEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYEK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:344	18D6	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLVSIASFHKAHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYEK LGFSEQGEVYDIPPIGPHILMYKKLA
SEQ ID NO:345	19C6	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLVSIASFHQAEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVRGYYEK LGFSEQGGVYDIPPIGPHILMYKKLA
SEQ ID NO:346	19D5	MIEVKPINAEDTYEIRHCILRPNQPLEACMYETDLLGGTFH LGGYYQGKLVSIASFHKAHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYKK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:347	20A12	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYQGKLVSIASFHNAEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGVDLLWCNARTSVSGYYKK LGFSEQGGIYDIPPIGPHILMYKKLA
SEQ ID NO:348	20F2	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLVSIASFHQAEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYEK

		LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:349	2.10E+12	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGAFH LGGYYQGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYKK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:350	23H11	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYQGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYEYK LGFSEQGEVYDIPPIGPHILMYKKLA
SEQ ID NO:351	24C1	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRDLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYKK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:352	24C6	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARISVSG`KKL GFSEQGGVYDIPPIGPHILMYKKLA
SEQ ID NO:353	2.40E+08	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHNAEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYEYK LGFSEQGEVYDIPPIGPHILMYKKLA
SEQ ID NO:354	2_8C3	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRDLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYEYK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:355	2H3	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYQGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY YRERKAGSTLIRHAEELLRKKGADLLWCNARISASGYEYK LGFSEQGGVYDIPPIGPHILMYKKLT
SEQ ID NO:356	30G8	MIEVKPINAEDTYEIRHRILRPNQPLEACMFETDLLGGAFH LGGYYQGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYKK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:357	3B_10C4	MIEVRPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYEYK LGFSEQGEAYDIPPIGPHILMYKKLT
SEQ ID NO:358	3B_10G7	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYEYK LGFSEQGGVYDIPPIGPHILMYKKLT
SEQ ID NO:359	3B_12B1	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYEYK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:360	3B_12D10	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGAFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARISASGYEYK GFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID	3B_2E5	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH

NO:361		LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYEYK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:362	3C_10H3	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARISASGYKKL GFSEQGGVYDIPVGPILMYKKLT
SEQ ID NO:363	3C_12H10	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY RGQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYEYK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:364	3C_9H8	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYQDRLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRYAEELLRKKGADLLWCNARISASGYEYK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:365	4A_1B11	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYEYK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:366	4A_1C2	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEEY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:367	4B_13E1	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEEY REQKAGSTLIRHAEELLRKKGADLLWCNARISASGYEYK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:368	4B_13G10	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPPIGPYILMYKKLT
SEQ ID NO:369	4B_16E1	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPPIGPHILMYKKLT
SEQ ID NO:370	4B_17A1	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEEY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYEYK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:371	4B_18F11	MIEVNPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTSH LGGYYRGKLISIASFHNAEHSLEDGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYEYK LGFSEQGEVYDIPPIGPHISMYKKLT
SEQ ID NO:372	4B_19C8	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPPIGPHILMYKKLA
SEQ ID NO:373	4B_1G4	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGAFH LGGYYRGKLISIASFHQSEHSELEGQKQYQLRGMATLEGY RELKAGSTLIRHAEELLRKKGADLLWCNARISASGYKKL

		GFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:374	4B_21C6	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGQKQYJLRGMATLEEY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKKL GFSEQGGVYDIPPIGPHILMYKKLT
SEQ ID NO:375	4B_2H7	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPPIGPHILMYKKLT
SEQ ID NO:376	4B_2H8	MIEAKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:377	4B_6D8	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEHGEVYDIPPIGPHILMYKKLT
SEQ ID NO:378	4B_7E8	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYKK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:379	4C_8C9	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLRGAFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:380	4H1	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGAFH LGGYYQGKLISIASFHQAVHSELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYK KLGFSQGGVYDIPPIGPHILMYKKLT
SEQ ID NO:381	6_14D10	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGHKQYQLRGMATLEEY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPVGPILMYKKLT
SEQ ID NO:382	6_15G7	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPVGPILMYKKLT
SEQ ID NO:383	6_16A5	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPVGPILMYKKLT
SEQ ID NO:384	6_16F5	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAVHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPVGPILMYKKLT
SEQ ID NO:385	6_17C5	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAHEHSELEGQKQYQLRGMATLEGN REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID	6_18C7	MIEVKPINAEDTYEIRHRILRPNQPLEACRYETDLLGGTFH

NO:386		LGGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARISASGYKKL LGFSEQGEVYDIPPVGPHILMYKKLT
SEQ ID NO:387	6_18D7	MIEVKPINAEDTYEIRXRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPPVGPHILMYKKLT
SEQ ID NO:388	6_19A10	MIEAKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPTGPHILMYKKLT
SEQ ID NO:389	6_19B6	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLRGAFH LGGYYRGKLISIASFHQAEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPVGPHILMYKKLT
SEQ ID NO:390	6_19C3	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:391	6_19C8	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTLH LGGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPPVGPHILMYKELT
SEQ ID NO:392	6_20A7	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLRGTFH LGGYYRGKLISIASFHQAEHSDLEGQKQYQLRGMATLEEEY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPVGPHILMYKKLT
SEQ ID NO:393	6_20A9	MIEVKPINAGDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPPVGPHILMYKKLT
SEQ ID NO:394	6_20H5	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:395	6_21F4	MIEVKPINAEDTYEIRHRVLRPNQPLEACMYETDLLGGAF HLGGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK KLGFSEQGEVYDVPPVGPHILMYKKLT
SEQ ID NO:396	6_22C9	MIEVKPINAEDTYEIRHRILRPNRPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAEHPGLEGKKQYQLRGMATLEEEY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPPVGPHILMYKKLT
SEQ ID NO:397	6_22D9	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGTFH LGGYYRGKLISIASFHQAEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPVGPHILMYKKLT
SEQ ID NO:398	6_22H9	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAEHSELEGQKQYQLRGMATLDEY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK



		LGFEQGEVYDIPPVGPILMYKKLT
SEQ ID NO:399	6_23H3	MIEVKPINAEDTYEIRHRILRPNQPLEACMYGTDLLGGTFH LGGYYRGKLISIASFHQAEOPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFEQGGVYDIPPVGPILMYKKLT
SEQ ID NO:400	6_23H7	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAHESELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKKL GFSEQGGVYDIPPVGPILMYKKLT
SEQ ID NO:401	6_2H1	MIEVKPINAEDTYEIRHRVLRPNQPLEACMYETDLLGGTF HLGGYYRGKLISIASFHQAHEPELEGQKPYQLRGMATLEG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK KLGFSEQGEIYDIPPVGPILMYKKLT
SEQ ID NO:402	6_3D6	MIEIKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH GGYYRGKLISIASFHQAHEPELEGQKQYQLRGMATLEGYR EQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKKL GFSEQGEVYDIPPVGPILMYKKLT
SEQ ID NO:403	6_3G3	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAHESELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFEQGEVYDIPPVGPILMYKKLT
SEQ ID NO:404	6_3H2	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAHEPELEGQKQYQLRGMATLECY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFEQGEVYDIPPVGPILMYKKLT
SEQ ID NO:405	6_4A10	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAHEPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFEQGEVYDIPPVGPILMYKKLT
SEQ ID NO:406	6_4B1	MIEVKPINAEDTYEIRHRVLRPNQPLEACMYETDLLGGTF HLGGYYRGKLIGIASFHQAHEPELEGQKQYQLRGMATLE GYREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGY EKLGFSGQGEVYDIPPVGPILMYKKLT
SEQ ID NO:407	6_5D11	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAHEPELEGQKQYQLRGMATLECY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFEQGEVYDIPPVGPILMYKKLT
SEQ ID NO:408	6_5F11	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAHEPELEGQKQYQLRGMATLECY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFEQGEVYDIPPVGPILMYKKLT
SEQ ID NO:409	6_5G9	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAHESELEGQKQYQLRGMATLECY REQKAGSTLIRHAEELLRKKGADLLWCNARISASGYKKL GFSEQGGVYDIPPVGPILMYKKLT
SEQ ID NO:410	6_6D5	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAHESELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFEQGGVYDIPPVGPILMYKKLT
SEQ ID	6_7D1	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLRGAFH

NO:411		LGGYYRGKLISIASFHQAEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPPVGP HILMYKKLT
SEQ ID NO:412	6_8H3	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPPVGP HILMYKKLT
SEQ ID NO:413	6_9G11	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTLH LGGYYRGKLISIASFHQAEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPVGP HILMYKKLT
SEQ ID NO:414	6F1	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLVCIASFHKAHSELEGQKQYQLRGMATLDG YREQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYE KLGFSEQGEVYDIPPVGP HILMYKKLT
SEQ ID NO:415	7_1C4	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEEEY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPPVGP HILMYKKLT
SEQ ID NO:416	7_2A10	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPPVGP HILMYKKLT
SEQ ID NO:417	7_2A11	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPPVGP HILMYKKLT
SEQ ID NO:418	7_2D7	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPVGP HILMYKKLT
SEQ ID NO:419	7_5C7	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEGY REQKVGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGGVYDIPPVGP HILMYKKLT
SEQ ID NO:420	7_9C9	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPVGP HILMYKKLT
SEQ ID NO:421	9_13F10	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLRGAFH LGGYYRGKLVSIASFHQAHEHSELEGQKQYQLRGMATLEE YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK KLGFSEQGEVYDIPPTGP HILMYKKLT
SEQ ID NO:422	9_13F1	MIEAKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLVSIASFHQAHEHTELEGQKQYQLRGMATLEE YREQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK KLGFSEQGEVYDIPPVGP HILMYKKLT
SEQ ID NO:423	9_15D5	MIEVKPINAEDTYEIRHRILRPNQPLDACKYETDLLGGTFH LGGYYRGKLISIASFHQAEHPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSASGYKK

		LGFSEQGEVYDIPPVGPHILMYKKLT
SEQ ID NO:424	9_15D8	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYK KLGFSQGEVYDTPPVGPHILMYKKLT
SEQ ID NO:425	9_15H3	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDMLRGAFH LGGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLECY HEQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYKK LGFSEQGEVYNTTPVGPHILMYKKLT
SEQ ID NO:426	9_18H2	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAEPHELVGQKQYQLRGMATLEGY REQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPVGPHILMYKKLT
SEQ ID NO:427	9_20F12	MIEVKPINAEDTYEIRHRVLRPNQPLEACMYETDLLGGTF HLGGYYRGELVSIASFHQAEHPELEGQKQYQLRGMATLE GYREQKAGSTLIRHAEALLRKKGADLLWCNARTSASGY KKLGFSEQGGVYDIPPVGPHILMYKKLT
SEQ ID NO:428	9_21C8	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYKK LGFSDQGEVYDIPPVGPHILMYKKLT
SEQ ID NO:429	9_22B1	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEG YREQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYK KLGFSQGEVYDLPTGPHILMYKKLT
SEQ ID NO:430	9_23A10	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTLH LGGYYRGKLVSIASFHQAEHPELEGQKQYQLRGMATLEG YRGQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYK KLGFSQGGVYDIPPVGPHILMYKKLT
SEQ ID NO:431	9_24F6	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLRGAFH LGGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPTGPHILMYKKLT
SEQ ID NO:432	9_4H10	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTLH LGGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPVGPHILMYKKLT
SEQ ID NO:433	9_4H8	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFNQAEPHELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPVGPHILMYKKLT
SEQ ID NO:434	9_8H1	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH GGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLEGY EQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPTGPHILMYKKLT
SEQ ID NO:435	9_9H7	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGSTFH LGGYYRGKLISIASFHQAEPHELEGQKQYQLRGMATLECY REQKAGSTLIRHAEALLRKKGADLLWCNARTSASGYKK LGFSEQGEVYDIPPVGPHILMYKKLT
SEQ ID	9C6	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH

NO:436		LGGYYQGKLISIASFHNAEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYEK LGFSEQGEVYDIPPVGPHILMYKKLA
SEQ ID NO:437	9H11	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHKAEHSELEGEEQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYKK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:438	0_4B10	MIEVKPINAEDTYELRHKILRPNQPIEACMYESDLLRGAFH LGGFYRGKLISIASFHQAHS DLEGQKQYQLRGMATLEGY RDQKAGSTLIKHAEEILRKR GADMLWCNARTTASGYYKK LGFSEQGEIFDTPPVGPHILMYKRIT
SEQ ID NO:439	0_5B11	MIEVKPINAEDTYELRHKILRPNQPIEACMYESDLLRGAFH LGGFYGGKLISIASFHQAHS DLEGQKQYQLRGMATLEGY RDQKAGSTLIKHAEEILRKR GADMLWCNARTSASGYYK KLGFSEQGEVFETPPVGPHILMYKKIT
SEQ ID NO:440	0_5B3	MIEVKPINAEDTYELRHRILRPNQPIEACMYETDLLRGAFH LGGFYRGKLISIASFHQAHS ELQGQKQYQLRGMATLEGY RDQKAGSSLIKHAEEILRKR GADLLWCNARTSASGYYKK LGFSEQGEVFDTPPVGPHILMYKRIT
SEQ ID NO:441	0_5B4	MIEVKLINAEDTYELRHRILRPNQPLEACMYETDLLRGAF HLGGFYRGKLISIASFHQAHS DLEGQKQYQLRGMATLEG FRDQKAGSSLIKHAEEILRKR GANLLWCNARTSASGYYKK LGFSEQGEVFDTPPVGPHILMYKRIT
SEQ ID NO:442	0_5B8	MIEVKPINAEDTYELRHKILRPNQPIEACMYESDLLRGAFH LGGFYRGKLISIASFHQAHS DLQGQKQYQLRGMATLEGY RDQKAGSSLIRHAEQILRKR GADLLWCNARTSASGYYKK LGFSEQGEIFDTPPVGPHILMYKRIT
SEQ ID NO:443	0_5C4	MIEVKPINAEDTYELRHKILRPNQPLEACMYETDLLRGAF HLGGFYRGKLISIASFHQAHS GLQGQKQYQLRGMATLEG YREQKAGSSIIKHAEEILRKK GADLLWCNARTSASGYYKK LGFSEQGEIFDTPPVGPHILMYKRIT
SEQ ID NO:444	0_5D11	MIEVKPINAEDTYELRHRILRPNQPIEACMYESDLLRGAFH LGGFYRGKLISIASFHQAHS DLQGQKQYQLRGMATLEGY REQKAGSTLIRHAEQILRKR GADLLWCNARTSASGYYKR LGFSEQGEVFDTPPVGPHILMYKRIT
SEQ ID NO:445	0_5D3	MIEVKPINAEDTYELRHRILRPNQPIEACMYESDLLRGAFH LGGYYRGKLISIASFHQAHS ELQGQKQYQLRGMATLEGY REQKAGSSLIKHAEEILRKR GADLLWCNARTSASGYYKKL GFSEQGEIFETPPVGPHILMYKRIT
SEQ ID NO:446	0_5D7	MIEVKPINAEDTYELRHRILRPNQPIEACMYETDLLRGAFH LGGFYRGKLISIASFHQAHS SELEGQKQYQLRGMATLEGY RDQKAGSSLIRHAEQILRKK GANMLWCNARTTASGYYK KLGFSEQGEIFDTPPVGPHILMYKRIT
SEQ ID NO:447	0_6B4	MIEVKPINAEDTYELRHRILRPNQPIEACMYESDLLRGALH LGGFYRGKLISIASFHQAHS DLQGQKQYQLRGMATLEGF RDQKAGSSLIRHAEQILRKR GADLLWCNARTSASGYYKK LGFSEQGKVFDTPPVGPHILMYKRIT
SEQ ID NO:448	0_6D10	MIEVKPINAEDTYELRHKILRPNQPLEVCMYETDLLRGAF HLGGFYRGKLISIASFHQAHS DLQGQKQYQLRGMATLEG YRDQKAGSSLIRHAEQILRKR GADMLWCNARTSASGYYK

		KLGFSEQGEVFETPPVGPHILMYKRLT
SEQ ID NO:449	0_6D11	MIEVKPINAEDTYELRHRILRPNQPIEACMYESDLLRGAFH LGGYYRGKLISIASFHQAHSDDLQGGKQYQLRGMATLEGF RDQKAGSSLIRHAEQILRKRGA DLLWCNARTSASGYKK LGFSEQGEVFETPPVGPHILMYKRLT
SEQ ID NO:450	0_6F2	MIEVKPINAEDTYELRHRILRPNQPIEACMYESDLLRGAFH LGGYYRGKLISIASFHQAHSDDLQGGKQYQLRGMATLEGF REQKAGSTLIRHAEQILRKRGA DMLWCNARTSASGYKK LGFSEQGEIFDTPPVGPHILMYKRLT
SEQ ID NO:451	0_6H9	MIEVKPINAEDTYELRHKILRPNQPIEACMYETDLLRGAFH LGGFYGGKLISIASFHQAHS DLEGKQYQLRGMATLEGY REQKAGSTLIRHAEILRKKGANLLWCNARTSASGYKKL GFSEQGEVFDTPPVGPHILMYKRLT
SEQ ID NO:452	10_4C10	MIEVKPINAEDTYELRHKILRPNQPLEVCMYETDLLRGAF HLGGXYRGKLISIASFHQAHSDDLQGGKQYQLRGMATLEG YRDQKAGSSLIKHAEQILRKRGA DXLWCNARTSASGYK KLGFSEQGEIFDTPPVGPHILMYKRLT
SEQ ID NO:453	10_4D5	MIEVKPINAEDTYELRHRILRPNQPIEVCMYETDLLRGAFH LGGFYRGKLISIASFHQAHSDDLQGGKQYQLRGMATLEGY REQKAGSTLIRHAEQILRKRGA DLLWCNARTSASGYKKL GFSEQGEVFDTPPVGPHILMYKRLT
SEQ ID NO:454	10_4F2	MLEVKPINAEDTYELRHRILRPNQPIEACMFESDLLRGAFH LGGFYRGKLISIASFHQAHSDDLQGGKQYQLRGMATLEGY REQKAGSSLIRHAEILRKRGA DMLWCNARTSASGYKK LGFSEQGEIFETPPVGPHILMYKRLT
SEQ ID NO:455	10_4F9	MIEVKPINAEDTYELRHRILRPNQPIEVCMYETDLLRGAFH LGGFYRGKLISIASFHQAHSDDLQGGKQYQLRGMATLEGF REQKAGSSLIRHAEQILRKRGA DLLWCNARTSASGYKKL GFSEQGEIFDTPPVGPHILMYKRLT
SEQ ID NO:456	10_4G5	MIEVKPINAEDTYELRHRILRPNQPIEACMFESDLLRGAFH LGGYYRGKLISIASFHQAHSDDLQGGKQYQLRGMATLEG YRDQKAGSSLIRHAEQILRKRGA DLLWCNARTSASGYK KLGFSEQGEIFDTPPVGPHILMYKRLT
SEQ ID NO:457	10_4H4	MLEVKPINAEDTYELRHKILRPNQPLEVCMYETDLLRGAF HLGGFYRGKLISIASFHQAHSDDLQGGKQYQLRGMATLEG YREQKAGSSLIKHAEILRKRGA DLLWCNARTSASGYKK LGFSEQGEVFDTPPVGPHILMYKRLT
SEQ ID NO:458	11_3A11	MIEVKPINAEDTYELRHKILRPNQPIEVCMYESDLLRGAFH LGGFYRGKLISIASFHQAHPDLQGGKQYQLRGMATLEGY RDQKAGSSLIKHAEQILRKRGA DLLWCNARTSASGYKK LGFSEQGEVFETPPVGPHILMYKRLT
SEQ ID NO:459	11_3B1	MLEVKPINAEDTYELRHRILRPNQPIEACMFETDLLRGAFH LGGFYRGKLISIASFHQAHSDDLQGGKQYQLRGMATLEGF REQKAGSTLIRHAEILRKRGA DLLWCNARTSASGYKKRL GFSEQGEIFDTPPVGPHILMYKRLT
SEQ ID NO:460	11_3B5	MIEVKPINAEDTYELRHRILRPNQPIEACMFESDLLRGAFH LGGYYRGKLISIASFHQAHSDDLQGGKQYQLRGMATLEGY RDQKAGSSLIRHAEQILRKRGA DMLWCNARTSASGYKK LGFSEQGEVFDTPPVGPHILMYKRLT
SEQ ID	11_3C12	MIEVKPINAEDTYELRHRILRPNQPLEVCMYETDLLRGAFH

NO:461		LGGFYGGKLISIASFHQAEHPDLQGQKQYQLRGMATLEGY RDQKAGSSLRHAEQLLRKRGADLLWCNARTSASGYKK LGFSEQGEIFETPPVGPHILMYKRIT
SEQ ID NO:462	11_3C3	MIEVKPINAEDTYELRHKILRPNQPIEACMYESDLLRGALH LGGFYRGKLISIASFHQAEHSELQGQKQYQLRGMATLEGY REQKAGSSLIKHAEEILRKRKADLLWCNARTSASGYKKL GFSEQGEVFDTPPVGPHILMYKRIT
SEQ ID NO:463	11_3C6	MLEVKPINAEDTYELRHKILRPNQPIEACMFESDLLRGAFH LGGFYGGKLISIASFHQAEHSDLEGQKQYQLRGMATLEGY REQKAGSTLRHAEELRKRKADLLWCNARTSASGYKKL GFSEQGEIFDTPPVGPHILMYKRIT
SEQ ID NO:464	11_3D6	MIEVKPINAEDTYELRHRILRPNQPIEVCMYETDLLRGAFH LGGFYRGKLISIASFHQAEHSDLQGQKQYQLRGMATLEGY REQKAGSSLIKHAEQILRKRKADLLWCNARTSASGYKKL GFSEQGEVFDTPPVGPHILMYKRIT
SEQ ID NO:465	1_1G12	MLEVKPINAEDTYELRHRILRPNQPIEVCMYETDLLRGAFH LGGFYGGKLISIASFHQAEHSELQGQKQYQLRGMATLEGY RDQKAGSSLIKHAEEILRKRKADLLWCNARTSASGYKKL GFSEQGEVFETPPVGPHILMYKRIT
SEQ ID NO:466	1_1H1	MIEVKPINAEDTYELRHKILRPNQPIEACMYESDLLRGSFH LGGFYRGKLISIASFHKAEHSELQGQKQYQLRGMATLEGF REQKAGSSLRHAEELRKRKADLLWCNARTTASGYKKRL GFSEHGEVFETPPVGPHILMYKRIT
SEQ ID NO:467	1_1H2	MIEVKPINAEDTYELRHRILRPNQPIEACMYESDLLRGSFH LGGFYRGKLISIASFHQAEHSELEGQKQYQLRGMATLEGF REQKAGSSLRHAEELRKRKADLLWCNARTTAAGYYKK LGFSEQGEIFDTPPVGPHILMYKRIT
SEQ ID NO:468	1_1H5	MIEVKPINAEDTYEIRHRILRPNQPIEACMYESDLLRGSFH LGGFYRGKLISIASFHQAEHSDLEGQKQYQLRGMATLEGY RDQKAGSSLRHAEQILRKRKADLLWCNARTTAAGYYKR LGFSEQGEVFDTPPVGPHILMYKKLT
SEQ ID NO:469	1_2A12	MIEVKPINAEDTYELRHRILRPNQPIEACMYESDLLRGSFH LGGFYRGKLISIASFHQAEQSELEGQKQYQLRGMATLEGY RDQKAGSTLIKHAEEILRKRKADLLWCNARTSAAGYYKR LGFSEQGEIFDTPPVGPHILMYKRIT
SEQ ID NO:470	1_2B6	MIEVKPINAEDTYELRHKILRPNQPIEACMYETDLLRGSFH LGGFYRGKLISIASFHQAEHSELEGQKQYQLRGMATLEGF RDQKAGSSLIKHAEEILRKRKADLLWCNARTSASGYKKL GFSEQGEIFETPPVGPHILMYKRIT
SEQ ID NO:471	1_2C4	MLEVKPINAEDTYELRHKILRPNQPIEACMYETDLLRGSFH LGGFYRGKLISIASFHQAEHSDLQGQKQYQLRGMATLEGY REQKAGSTLIKHAEEILRKRKADLLWCNARTTAAGYYKK LGFSEQGEVFDTPPVGPHILMYKKIT
SEQ ID NO:472	1_2D2	MIEVKPINAEDTYELRHKILRPNQPIEACMYESDLLRSFHF LGGFYRGKLISIASFHKAEHSELQGQKQYQLRGMATLEGY RDQKAGSSLRHAEELRKRKADMLWCNARTSAAGYYKR LGFSEQGEVFDTPPVGPHILMYKRIT
SEQ ID NO:473	1_2D4	MIEVKPINAEDTYELRHRILRPNQPIEACMYESDLLRGSFH LGGFYRGKLISIASFHQAEHSDLQGQKQYQLRGMATLEGY REQKAGSSLIKHAEQILRKRKADMLWCNARTSAAGYYK

		RLGFSEHGEIFETPPVGPHILMYKRIT
SEQ ID NO:474	1_2F8	MLEVKPINAEDTYELRHRILRPNQPLEACMYETDLLRGSF HLGGFYRGKLISIASFQAHEHSELEGQKQYQLRGMATLEG YRDQKAGSSLRHAEELRKRKADMLWCNARTTAAGYYK KLGFEHQGEIYDTPPVGPHILMYKKLT
SEQ ID NO:475	1_2H8	MIEVKPINAETTYELRHKILRPNQPLEACMYETDLLRGAFH LGGFYRGKLISIASFQAADHSELQGGQKQYQLRGMATLEGY REQKAGSTLIRHAEQILRKRKADLLWCNARTSAAGYYKK LGFSEHGEIFETPPVGPHILMYKRIT
SEQ ID NO:476	1_3A2	MIEVKPINAEDTYELRHRILRPNQPIEACMYESDLLRGAFH LGGFYRGKLISIASFQAHEHSDLGQKQYQLRGMATLEGY REQKAGSSLRHAEELRKKGADMLWCNARTTAAGYYKR LGFSEHQGEVFDTPPVGPHILMYKRIT
SEQ ID NO:477	1_3D6	MIEVKPINAEDTYELRHKILRPNQPIEACMYESDLLQGSFH LGGFYRGQLISIASFQAHEHSDLGQKQYQLRGMATLEGF REQKAGSTLIKHAEEILRKKGADLLWCNARTSAAGYYKK LGFSEHGEIFDTPPAGPHILMYKKLT
SEQ ID NO:478	1_3F3	MIEVKPINAETTYELRQRILRPNQPIEACMYESDLLRGSFHL GGFYRGQLISIASFQAHEHSELQGGQKQYQLRGMATLEGYR EQKAGSTLIKHAEEILRKKGADLLWCNARTSAAGYYKRL GFSEHGEIFDTPPVGPHILMYKRIT
SEQ ID NO:479	1_3H2	MIEVKPINAEDTYELRHRILRPNQPIEACMYETDLLRGAFH LGGYYRGQLISIASFQAHEHSELQGGQKQYQLRGMATLEGY REQKAGSTLIKHAEEILRKRKADMLWCNARTSAAGYYK RLGFSEHQGEVFDTPPVGPHILMYKKLT
SEQ ID NO:480	1_4C5	MIEVKPINAEDTYELRHKILRPNQPIEACMYESDLLRGSFH LGGFYRGKLISIASFQAHEHSDLEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELRKRKADMLWCNARTSASGYKKR LGFSEHGEIFDTPPVGPHILMYKRIT
SEQ ID NO:481	1_4D6	MLEVKPINAEDTYELRHRILRPNQPIEACMYETDLLRGSFH LGGFYRGQLISIASFQAHEHSDLEGQKQYQLRGMATLEGY REQKAGSTLIRHAEQILRKRKADMLWCNARTSAAGYYKR LGFSEHQGEVFETPPVGPHILMYKRIT
SEQ ID NO:482	1_4H1	MIEVKPINAEDTYELRHRILRPNQPLEACMYETDLLRGSFH LGGFYRGKLISIASFQAHEHSDLGQKQYQLRGMATLEGY REQKAGSTLIRHAEQILRKRKADLLWCNARTSASGYKKR LGFSEHGEVFDTPPVGPHILMYKRIT
SEQ ID NO:483	1_5H5	MLEVKPINAETTYELRHKILRPNQPLEACMYESDLLRGSFH LGGYYRGQLISIASFQAHEHSELEGQKQYQLRGMATLEGF REQKAGSTLIKHAEEILRKRKADMLWCNARTSAAGYYKK LGFSEHGEIFDTPPVGPHILMYKKLT
SEQ ID NO:484	1_6F12	MIEVKPINAETTYELRHRILRPNQPIEACMYESDLLRGSFHL GGFYRGKLISIASFQAHEHSDLEGQKQYQLRGMATLEGYR DQKAGSTLIKHAEEILRKRKADMLWCNARTSAAGYYKR LGFSEHGEIYETPPVGPHILMYKKIT
SEQ ID NO:485	1_6H6	MIEVKPINAEDTYELRHKILRPNQPIEACMYESDLLRGSFH LGGFYRGQLISIASFQAHEHSDLEGQKQYQLRGMATLEGY RDQKAGSSLIKHAEEILRKRKADLLWCNARTSAAGYYKR LGFSEHQGEIFDTPPVGPHILMYKKIT
SEQ ID	3_11A10	MLEVKPINAEDTYELRHRILRPNQPIEACMYESDLLRGAFH

NO:486		LGGYYRGKLISIASFHQAEHSELQGQKQYQLRGMATLEGY REQKAGSSLVKHAEELRKRADLLWCNARTSASGYKK LGFSEQGEIFETPPVGPHILMYKRIT
SEQ ID NO:487	3_14F6	MLEVKPINAEDTYELRHRILRPNQPIEACMYESDLLRGAFH LGGFYRGKLISIASFHQAEHSELQGQKQYQLRGMATLEGY REQKAGSSLIKHAEEILRKRADLLWCNARTSASGYKKL GFSEQGEIFETPPVGPHILMYKRIT
SEQ ID NO:488	3_15B2	MLEVKPINAEDTYELRHKILRPNQPLEVCMYETDLLRGAF HLGGYYGGKLISIASFHQAEHSELQGQKQYQLRGMATLE GYREQKAGSSLIKHAEEILRKRADLLWCNARTSASGYKK KLGFSEQGEIFETPPVGPHILMYKRIT
SEQ ID NO:489	3_6A10	MIEVKPINAEDTYELRHRILRPNQPIEACMYESDLLRGAFH LGGYYRGKLISIASFHQAEHSELQGQKQYQLRGMATLEGY REQKAGSSLIKHAEEILRKRADLLWCNARTSASGYKKL GFSEQGEIFETPPVGPHILMYKRIT
SEQ ID NO:490	3_6B1	MLEVKPINAEDTYELRHRILRPNQPIEACMYESDLLRGAFH LGGYYRGKLISIASFHQAEHPELQGQKQYQLRGMATLEGY REQKAGSSLIKHAEEILRKRADLLWCNARTSASGYKKL GFSEQGEVFETPPVGPHILMYKRIT
SEQ ID NO:491	3_7F9	MLEVKPINAEDTYELRHRILRPNQPIEACMYESDLLRGAFH LGGYYGGKLISIASFHQAEHSDLQGQKQYQLRGMATLEG YREQKAGSSLIKHAEEILRKRADLLWCNARTSASGYKK LGFSEQGEIFETPPVGPHILMYKRIT
SEQ ID NO:492	3_8G11	MLEVKPINAEDTYELRHRILRPNQPIEVCMYESDLLRGAFH LGGYYRGKLISIASFHQAEHSELQGQKQYQLRGMATLEGY REQKAGSSLIKHAEEILRKRADLLWCNARTSASGYKKL GFSEQGEIFETPPVGPHILMYKRIT
SEQ ID NO:493	4_1B10	MIEVKPINAEDTYELRHRILRPNQPIEVCMYETDLLRGAFH LGGFYGGKLISIASFHQAEHSDLQGQKQYQLRGMATLEGY RDQKAGSSLRHAEQILRKRADMLWCNARTSASGYKK LGFSEQGEIFETPPVGPHILMYKRIT
SEQ ID NO:494	5_2B3	MIEVKPINAEDTYELRHRILRPNQPLEVCMYETDLLRGAFH LGGFYGGKLISIASFHQAEHSDLQGQKQYQLRGMATLEGY RDQKAGSSLRHAEQILRKRADMLWCNARTSASGYKK LGFSEQGEIFETPPVGPHILMYKRIT
SEQ ID NO:495	5_2D9	MLXVKPINAEDTYELRHKILRPNQPXEVCMYEXDLLRGAF HLGGFYRGKLISIASFHQAEHSDLQGQKQYQLRGMATLEG YRDQKAGSSLIKHAEQILRERGADMLWCNARTSASGYKK KLGFSEQGEVFETPPVGPHILMYKRIT
SEQ ID NO:496	5_2F10	MLEVKPINAEDTYELRHKILRPNQPIEVCMYETDLLRGAF HLGGFYGGKLISIASFHQAEHSDLQGQKQYQLRGMATLEG YRDQKAGSSLRHAEQILRKRADMLWCNARTSASGYKK KLGFSEQGEIFETPPVGPHILMYKRIT
SEQ ID NO:497	6_1A11	MLEVKPINAEDTYELRHKILRPNQPLEVCMYETDLLRGAF HLGGFYRGKLISIASFHQAEHSDLQGQKQYQLRGMATLEG YRDQKAGSSLRHAEQILRKRADMLWCNARTSASGYR KLGFSEQGEVFETPPVGPHILMYKRIT
SEQ ID NO:498	6_1D5	MIEVKPINAEDTYELRHKILRPNQPLEVCMYETDLLRGAF HLGGFYRGKLISIASFHQAEHSDLQGQKQYQLRGMATLEG YRDQKAGSSLRHAEQILRKRADMLWCNARTSASGYKK



		KLGFSEQGEVFETPPVGPHILMYKRIT
SEQ ID NO:499	6_1F11	MIEVKPINAEDTYELRHKILRPNQPLEVCMYETDLLRGAF HLGGFYRGKLISIASFQAHSDDLQGGKQYQLRGMATLEG YREQKAGSSLIRHAEQILRKRGAADMWLCNARTSASGYK KLGFSEQGEVFETPPVGPHILMYKRIT
SEQ ID NO:500	6_1F1	MIEVKPINAEDTYELRHKILRPNQPLEVCMYETDLLRGAF HLGGFYRGKLISIASFQAHSDDLQGGKQYQLRGMATLEG YRDQKAGSSLIRHAEQILRKRGAADMWLCNARTSASGYK KLGFSEQGEVFETPPVGPHILMYKRIT
SEQ ID NO:501	6_1H10	MIEVKPINAEDTYELRHKILRPNQPLEVCMYETDLLRGAF HLGGFYGGKLISIASFQAHSDDLQGGKQYQLRGMATLEG YRDQKAGSSLIRHAEELRKRGAADMWLCNARTSASGYK KLGFSEQGEVFDTPPVGPHILMYKKIT
SEQ ID NO:502	6_1H4	MIEVKPINAEDTYELRHKILRPNQPLEVCMYETDLLRGAF HLGGFYGGKLISIASFQAHSDDLQGGKQYQLRGMATLEG YRDQKAGSTLIKHAEQILRKRGAADMWLCNARTSASGYK KLGFSEQGEVFETPPVGPHILMYKRIT
SEQ ID NO:503	8_1F8	MIEVKPINAEDTYELRHRILRPNQPLEVCMYETDLLRGAFH LGGFYRGKLISIASFQAHSDDLQGGKQYQLRGMATLEGY REQKAGSSLIKHAEEILRKRGAADLLWCNARTSASGYKKL GFSEQGEIFDTPPVGPHILMYKRIT
SEQ ID NO:504	8_1G2	MIEVKPINAEDTYELRHRVLRPNQPLEVCMYETDLLRGAF HLGGYYRGKLISIASFQAHSDDLQGGKQYQLRGMATLEG YREQKAGSSLIKHAEEILRKRGAADLLWCNARTSASGYKK LGFSEQGEVFETPPVGPHILMYKRIT
SEQ ID NO:505	8_1G3	MIEVKPINAEDTYELRHKILRPNQPIEVCMYETDLLRGAF HLGGYYRGKLISIASFQAHSDDLQGGKQYQLRGMATLEG YREQKAGSSLIRHAEELRKRGAADLLWCNARTSASGYKK LGFSEQGEIFDTPPVGPHILMYKRIT
SEQ ID NO:506	8_1H7	MIEVKPINAEDTYELRHRILRPNQPIEVCMYETDLLRGAFH LGGFYRGKLISIASFQAHSDDLQGGKQYQLRGMATLEGY REQKAGSSLIKHAEEILRKRGAADMWLCNARTSASGYKK LGFSEQGEIFETPPVGPHILMYKRIT
SEQ ID NO:507	8_1H9	MIEVKPINAEDTYELRHKILRPNQPLEVCMYETDLLRGAF HLGGYYRGKLISIASFQAHSDDLQGGKQYQLRGMATLE GYREQKAGSSLIRHAEELRKRGAADLLWCNARTSASGYK KLGFSEQGEVFDTPPVGPHILMYKRIT
SEQ ID NO:508	GAT1_21F 12	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISIASFHNAEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYKK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:509	GAT1_24G 3	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFQAHSDDLQGGKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTFVSGYYEK LGFSEQGEVYDIPPIGPYILMYEKL
SEQ ID NO:510	GAT1_29G 1	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISIASFQAHSDDLQGGKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYKK LGFSEQGGVCDIPPIGPHILMYKKLA
SEQ ID	GAT1_32G	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH

NO:511	1	LGGYYRGKLISLASFHQAEHPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYEK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:512	GAT2_15G 8	MIEVKPINAEDTYEIRHRILRPNQPLEACKYETDLLGGTFH LGGYYRGKLISLASFHNAEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYKK LGFSEQGEVYDIPPIGPHILMYKKLT
SEQ ID NO:513	GAT2_19H 8	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISLASFHQAEHPELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYEK LGFSEQGEVCDIPPIGPHILMYKKLT
SEQ ID NO:514	GAT2_21F 1	MIEVKPINAEDTYEIRHRILRPNQPLEACMYETDLLGGTFH LGGYYRGKLISLASFHQAEHSELEGQKQYQLRGMATLEGY REQKAGSTLIRHAEELLRKKGADLLWCNARTSVSGYYKK LGFSEQGGVYDIPPIGPHILMYKKLT
SEQ ID NO:515	<i>B.</i> <i>licheniform-</i> сайт зв'язування з рибосомою	AACTGAAGGAGGAATCTC

## СПИСОК ПОСЛІДОВНОСТЕЙ

<110> КАСЛ, Лінда А.  
 САЙХЛ, Ден  
 ГАЙВЕР, Лорайн Дж.  
 МІНШУЛЛ, Джеремі  
 АЙВІ, Крістіна  
 ЧЕН, Янг Хонг  
 ДАК, Ніколас Б.

<120> НОВІ ГЕНИ ГЛІФОСАТ-N-АЦЕТИЛТРАНСФЕРАЗИ (GAT)

<130> 48949-20001.40

<140> PCT/US01/46227

<141> 2001-10-29

<150> US 60/244,385

<151> 2000-10-30

<160> 515

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 441

<212> DNA

<213> *Bacillus licheniformis*

<400> 1

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcgcttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca taaagccgaa 180
cattcagagc ttgaaggcga agaacagtat cagctgagag ggatggcgac gcttgaagga 240
taccgtgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag acctttttatg gtgcaatgcc aggacatctg tgagcgggta ctatgaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441
```

<210> 2

<211> 441

<212> DNA

<213> Unknown

<220>

<223> Unidentified microorganism derived from soil sample

<400> 2

```
atgattgaag tcaaacctat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggcacgttt 120
cacctcggcg gatattatcg ggacaggctg atcagcatcg cctcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcag acctttttatg gtgcaacgcc aggacatctg tgagcgggta ctataaaaag 360
```

ctcgggttca gccaacaagg cgggggtctac gatataccgc cgatcggacc tcatattttg 420  
atgtataaga aattgacata a 441

<210> 3

<211> 441

<212> DNA

<213> Unknown

<220>

<223> Unidentified microorganism derived from soil sample

<400> 3

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
cgccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtacgttt 120  
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca taatgccgaa 180  
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaaggg 240  
taccgcgagc aaaaagcggg aagtacgctt atccgccatg ccgaagagct tcttcgaaaa 300  
aaaggcggg accttttatg gtgcaacgcc aggatatctg tgagcggcta ctatgaaaag 360  
ctcgggttca gccaacaagg cgggatctac gacataccgc cgatcggacc tcatattttg 420  
atgtataaga aattggcata a 441

<210> 4

<211> 441

<212> DNA

<213> Unknown

<220>

<223> Unidentified microorganism derived from soil sample

<400> 4

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
cgccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtacgttt 120  
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca taatgccgaa 180  
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaaggg 240  
taccgcgagc aaaaagcggg aagtacgctt atccgccatg ccgaagagct tcttcgaaaa 300  
aaaggcggg accttttatg gtgcaacgcc aggatatctg tgagcggcta ctatgaaaag 360  
ctcgggttca gccaacaagg cgggatctac gacataccgc cgatcggacc tcatattttg 420  
atgtataaga aattggcata a 441

<210> 5

<211> 441

<212> DNA

<213> Unknown

<220>

<223> Unidentified microorganism derived from soil sample

<400> 5

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
cgccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtgcgttt 120  
cacctcgggtg gatattacca gggcaagctg atcagcatcg cttcctttca taaagccgaa 180  
cattcagagc ttgagggcga agaacagtat cagctgagag ggatggcgac gcttgaagga 240  
taccgtgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcgaaaa 300  
aagggggcag accttttatg gtgcaatgcc aggacatctg tgagcggcta ctatgaaaag 360  
ctcgggttca gccaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420  
atgtataaga aattgacgta a 441

<210> 6  
 <211> 146  
 <212> PRT  
 <213> *Bacillus licheniformis*

<400> 6  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Lys Ala Glu His Ser Glu Leu  
 50 55 60  
 Glu Gly Glu Glu Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Val Ser Gly Tyr Tyr Glu Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

<210> 7  
 <211> 146  
 <212> PRT  
 <213> Unknown

<220>  
 <223> Unidentified microorganism derived from soil sample

<400> 7  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Asp  
 35 40 45  
 Arg Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Val Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly  
 115 120 125  
 Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

<210> 8  
 <211> 146  
 <212> PRT  
 <213> Unknown

<220>

<223> Unidentified microorganism derived from soil sample

<400> 8

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Asn Ala Glu His Ser Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Ile
           100          105          110
Ser Val Ser Gly Tyr Tyr Glu Lys Leu Gly Phe Ser Glu Gln Gly Gly
           115          120          125
Ile Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Ala
145

```

<210> 9  
 <211> 146  
 <212> PRT  
 <213> Unknown

<220>

<223> Unidentified microorganism derived from soil sample

<400> 9

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Asn Ala Glu His Ser Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Ile
           100          105          110
Ser Val Ser Gly Tyr Tyr Glu Lys Leu Gly Leu Ser Glu Gln Gly Gly

```

309

86918

310

115 120 125  
 Ile Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Ala  
 145

<210> 10  
 <211> 146  
 <212> PRT  
 <213> Unknown

<220>  
 <223> Unidentified microorganism derived from soil sample

<400> 10  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Ala Phe His Leu Gly Gly Tyr Tyr Gln Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Lys Ala Glu His Ser Glu Leu  
 50 55 60  
 Glu Gly Glu Glu Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Val Ser Gly Tyr Tyr Glu Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

<210> 11  
 <211> 441  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic DNA Sequence

<400> 11  
 atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
 cggccgaatc agcogctgga agcatgcaag tatgaaaccg atttgctcag gggtaoqttt 120  
 cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180  
 catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac actogaagga 240  
 taccgtgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300  
 aaaggcgcgg accttttatg gtgcaacgcc aggacgtctg cgagcgggta ctataaaaag 360  
 ctcggttca gccaacaggg cgaagtctac gacataccgc cggtcgggacc tcatatattg 420  
 atgtataaga aattgacgta a 441

<210> 12

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 12

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcag gggcgctgtt 120
cacctcggtg gatattaccg gggcaagctg gtcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aagacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag acctcttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggttca gcgaacaggg cgaagtctac gacataccgc cgactgggac ccataatttg 420
atgtataaga aattgacata a                                     441

```

&lt;210&gt; 13

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 13

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcag gagcacgttt 120
cacctcggtg gatattaccg gggcaagctg atcagcatcg cttcctttta tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgtgagc aaaaagcggg aagcacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgctg accttttatg gtgcaacgcc aggacgtctg cgagcgggta ctataaaaag 360
ctcggttca gcgaacaggg cgaagtctac gacacaccgc cggtcggacc tcataatttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 14

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 14

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcag gggcgctgtt 120
cacctcggtg gatattaccg gggcaagctg gtcagcatcg cttcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac actcgaagga 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgctg accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggttca gcgaacaggg cgaagtctac gacacaccgc cgactgggac ccataatttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 15

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence



&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 15

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcag gggcgctgtt 120
caccttggtg gatattaccg gggcaagctg gtcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac actcgaagga 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagcgtt tcttcggaaa 300
aagggggcag acctcttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacacaccgc cggccgggacc ccataattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 16

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 16

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcag gggcgctgtt 120
cacctcggtg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagga 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacacaccgc cggtcgggacc tcatatatttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 17

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;221&gt; misc\_feature

&lt;222&gt; 54

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 17

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgnattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggcacgctt 120
cacctcggtg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240
taccgcgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag acctcttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaagg cgaagtctac gacataaccg cgaccgggacc ccataattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 18

&lt;211&gt; 441

&lt;212&gt; DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 18

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cgcccggaatc agccgcttga agcatgtatg tatgaaaccg atttgctcag gggtagcttt 120
cacctcggtg ggtattaccg gggcaagctg gtcagcatcg ctccctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240
taccgcgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggttca gcgaacaagg cgggtcttac gacataccgc cggtcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441
```

<210> 19

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 19

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cgcccggaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggcacgttt 120
cacctcggtg gatattaccg gggcaagctg atcagcatcg ctccctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggcgcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggttca gcgaacaagg cgaagtctac gacacaccgc cggtcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441
```

<210> 20

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 20

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cgcccggaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggcacgttt 120
cacctcggtg gatattaccg gggcaagctg gtcagcatcg ctccctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240
taccgcgagc aaaaagcggg aagcacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggcgcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggttca gcgaacaggg cgaagtctac gacataccgc cggtcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441
```

<210> 21

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 21

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca cgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaactg atttgctcgg tggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg gtcagcatcg cttcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240
taccgcgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag acctcttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaagg cggggtctac gacataccgc cggtcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 22

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 22

```

atgattgaag tcaaaccaat aaacgcggag gatacgtatg agatcaggca cgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattatcg gggcaagctg gtcagcatcg cttcctttca tcaagccgaa 180
catccagagc ttgaaggccg aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgaccggacc ccatattttg 420
atgtataaga aattgacata a                                     441

```

&lt;210&gt; 23

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 23

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca cgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcacgctt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaagg cggggtctac gacataccgc cggtcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 24

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 24

319	86918	320	
atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg agatcaggca ccgcatactc 60
cggccgaatc	agccgcttga	agcatgtatg	tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg	gatattaccg	gggcaagctg	gtcagcatcg cttcctttca tcaagccgaa 180
caccagagc	ttgaaggcca	aaaacagtat	cagctgagag ggatggcgac actcgaagga 240
taccgtgagc	aaaaagcggg	cagtaacgtt	atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag	acctcttatg	gtgcaacgcc	aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca	gcgaacaggg	cgaagtctac	gacacaccgc cggtcggacc tcataatttg 420
atgtataaga	aattgacgta	a	441

<210> 25

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 25

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc	agccgctgga	agcatgcaag	tatgaaaccg atttgctcgg gggtagcttt 120
cacctcgggtg	gatattaccg	gggcaagctg	atcagcatcg cctcctttca tcaagccgaa 180
catccagagc	ttgaaggcca	aaaacagtat	cagctgagag ggatggcgac acttgaagag 240
taccgcgagc	aaaaagcggg	cagtaacgtt	atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag	acctcttatg	gtgcaacgcc	aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca	gcgaacaggg	cgaagtctac	gacacaccgc cggtcggacc tcataatttg 420
atgtataaga	aattgacgta	a	441

<210> 26

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 26

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc	agccgctgga	agcatgcaag	tatgaaaccg atttgctcag gggtagcttt 120
cacctcggcg	gatattaccg	gggcaagctg	atcagcatcg cctcctttca tcaagccgaa 180
catccagagc	ttgaaggcca	aaaacagtat	cagctgagag ggatggcgac actcgaaggg 240
taccgtgagc	aaaaagcggg	aagcacgctc	atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcg	accttttatg	gtgcaacgcc	aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca	gcgaacaggg	cgaagtctac	gacacaccgc cggtcggacc tcataatttg 420
atgtataaga	aattgacgta	a	441

<210> 27

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 27

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc	agccgctgga	agcatgcaag	tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg	gatattaccg	gggcaagctg	gtcagcatcg cttcctttca tcaagccgaa 180

321	86918	322
catccagagc ttgaaggcca aaaacagtat	cagctgagag ggatggcgac	acttgaaggg 240
taccgcgagc aaaaagcggg aagtacgctt	atccgccatg ccgaagagct	tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc	aggacatctg cgagcgggta	ctataaaaaag 360
ctcggcttca gcgaacaggg cgaagtctac	gacataccgc cgaccggacc	ccatatttttg 420
atgtataaga aattgacgta a		441
<210> 28		
<211> 441		
<212> DNA		
<213> Artificial Sequence		
<220>		
<223> Synthetic DNA Sequence		
<400> 28		
atgattgaag tcaaaccgat aaacgcggaa	gatacgtatg agatcaggca	ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg	tatgaaaccg atttgctcgg	gggcacgctt 120
cacctcggtg gatattaccg gggcaagctg	gtcagcatcg cttcctttca	tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat	cagctgagag ggatggcgac	acttgaaggg 240
taccgtgagc aaaaagcggg cagtacgctt	atccgccatg ccgaagagct	tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc	aggacatctg cgagcgggta	ctataaaaaag 360
ctcggcttca gcgaacaggg cgaagtctac	gacataccgc cgaccggacc	ccatatttttg 420
atgtataaga aattgacgta a		441
<210> 29		
<211> 441		
<212> DNA		
<213> Artificial Sequence		
<220>		
<223> Synthetic DNA Sequence		
<400> 29		
atgattgaag tcaaaccaat aaacgcggaa	gatacgtatg agatcaggca	ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag	tatgaaaccg atttgctcgg	gggcacgctt 120
cacctcggtg gatattaccg gggcaagctg	gtcagcatcg cttcctttca	tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat	cagctgagag ggatggcgac	gcttgaaggg 240
taccgtgagc aaaaagcggg cagtacgctt	atccgccatg ccgaagagct	tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc	aggacatctg cgagcgggta	ctataaaaaag 360
ctcggcttca gcgaacaggg cgaagtctac	gacataccgc cggcgggacc	tcataattttg 420
atgtataaga aattgacgta a		441
<210> 30		
<211> 441		
<212> DNA		
<213> Artificial Sequence		
<220>		
<223> Synthetic DNA Sequence		
<400> 30		
atgattgaag tcaaaccaat aaacgcggaa	gatacgtatg agatcaggca	ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag	tatgaaaccg atttgctcgg	gggtgcgctt 120
cacctcggtg gatattaccg gggcaagctg	atcagcatcg cctcctttca	tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat	cagctgagag ggatggcgac	actcgaagga 240
taccgcgagc aaaaagcggg aagtacgctt	atccgccatg ccgaagagct	tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc	aggacatctg cgagcgggta	ctataaaaaag 360

323

86918

324

ctcggcttca gcgaacaagg cggggtctac gacataccgc cggtcggacc tcatattttg 420  
atgtataaga aattgacgta a 441

&lt;210&gt; 31

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 31

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
cggccgaatc agccgcttga agcatgcaag tatgaaaccg atttgctcgg gggtagcttt 120  
cacctcgggtg gatattacca gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180  
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaaggg 240  
taccgcgagc aaaaagcggg aagtacgctt atccgccatg ccgaagagct tcttcggaaa 300  
aagggggcag accttttatg gtgcaacgcc aggacatctg tgagcgggta ctataaaaag 360  
ctcggcttca gcgaacaagg cggggtctac gatataccgc cgatcggacc tcatattttg 420  
atgtataaga aattgacata a 441

&lt;210&gt; 32

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 32

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
cggccgaatc agccgcttga agcgtgtatg tatgaaaccg atttgctcgg gggcacgttt 120  
cacctcggcg gatattacca gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180  
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaaggg 240  
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300  
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctatgaaaag 360  
ctcggcttca gcgaacaagg cggggtctac gatataccgc cgatcggacc tcatattttg 420  
atgtataaga aattggcata a 441

&lt;210&gt; 33

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 33

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcatactc 60  
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcag gggtagcttt 120  
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180  
caccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240  
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagcgtt tcttcggaaa 300  
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360  
ctcggcttca gcgaacaggg cgaagtctac gacataccgc caactggggc ccatattttg 420  
atgtataaga aattgacgta a 441

<210> 34  
 <211> 441  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> Synthetic DNA Sequence

<400> 34  
 atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcatactc 60  
 cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggcacgttt 120  
 cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180  
 catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240  
 taccgcgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300  
 aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360  
 ctcggcttca gcgaacaggg cgaagtccac gacataccgc cgaccggacc ccatattttg 420  
 atgtataaga aattgacgta a 441

<210> 35  
 <211> 441  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> Synthetic DNA Sequence

<400> 35  
 atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
 cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggtaacttt 120  
 cacctcggcg gatattaccg gggcaagctg gtcagcatcg cttcctttca tcaagccgaa 180  
 catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagga 240  
 taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300  
 aagggggcag acctcttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360  
 ctcggcttca gcgaacaggg cgaagtctac gacacaccgc cggtcggacc tcatattttg 420  
 atgtataaga aattgacgta a 441

<210> 36  
 <211> 441  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic DNA Sequence

<400> 36  
 atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
 cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgcttgg gggcacgttt 120  
 cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180  
 catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac actcgaagga 240  
 taccgtgagc aaaaagcggg aagtacgctc atccgccatg ccgaagagct tcttcggaaa 300  
 aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360  
 ctcggcttca gcgaacaagg cggggtctac gacataccgc cggtcggacc tcatattttg 420  
 atgtataaga aattgacgta a 441

<210> 37  
 <211> 441  
 <212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 37

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cgcccgaaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattatcg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240
taccgcgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggtttca gcgaacaggg cgaagtctac gacgcaccgc cgaccggacc tcatattttg 420
atgtataaga aattgacgta a                                     441
```

<210> 38

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 38

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cgcccgaaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg gtcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagga 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggtttca gcgaacaagg cgggggtctac gacataccgc cggtcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441
```

<210> 39

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 39

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cgcccgaaatc agccgctgga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240
taccgcgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggtttca gcgaacaagg cgggggtctac gacataccgc ctgtcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441
```

<210> 40

<211> 441

<212> DNA

<213> Artificial Sequence

<220>



## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 40

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtaag tatgaaaccg atttgctcgg gggtagcttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac actcgaagga 240
taccgcgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag acctttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgaccggacc ccatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 41

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 41

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcatactc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg gtcagcatcg cctcctttca ccaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagcgct tcttcggaaa 300
aaaggcgcgg acctttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgaccggacc ccatattttg 420
atgtataaga aattgacata a                                     441

```

&lt;210&gt; 42

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 42

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaactg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag acctttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaagg cggggtctac gacataccgc cgaccggacc ccatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 43

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 43

331

86918

332

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgcaag tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagga 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacacaccgc cggtcgggacc tcataattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 44

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 44

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtagcttt 120
cacctcgggtg gatattacca gggcaagctg atcagcatcg cttcctttca taaagccgaa 180
cattcagagc ttgagggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaatgcc aggacatctg tgagcgggta ctataaaaaag 360
ctcggcttca gcgaacaagg cgggatctac gacataccgc cgatcgggacc tcataattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 45

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 45

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgcaag tatgaaaccg atttgctcgg gggtagcttt 120
cacctcgggtg gatattacca gggcaagctg atcagcatcg cttcctttca taaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaagga 240
taccgtgagc aaaaagcggg aagcacactc atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcag accttttatg gtgcaacgcc aggacatctg tgagcgggta ctataaaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcgggacc tcataattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 46

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 46

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtagcttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180

```

```

cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgcgagc aaaaagcggg aagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag acctttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaaag 360
ctcggcttca gcgaacaagg cggggtctac gacataccgc cgatcggacc tcataattttg 420
atgtataaga aattgacgta a                                     441

```

<210> 47

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 47

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtagcttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttta tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240
taccgcgagc aaaaagcggg aagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg acctttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cggtcggacc tcataattttg 420
atgcataaga aattgacgta a                                     441

```

<210> 48

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 48

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg aggcacgctt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttta tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac actcgaaggg 240
taccgtgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag acctctttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacacaccgc cggtcggacc tcataattttg 420
atgtataaga aattgacgta a                                     441

```

<210> 49

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 49

```

atgatcgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcag gaggcgttt 120
cacctcggcg gatattaccg gggcaagctg atcagcatcg cctcctttca ccaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgaggg ggatggcgac acttgaagag 240
taccgcgagc aaaaagcggg aagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag acctttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaaag 360

```

ctcggcttca gcgaacaggg cgaagtctac gacacaccgc cggtcggacc tcatattttg 420  
atgtataaga aattgacgta a 441

<210> 50

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 50

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gagcacgttt 120  
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180  
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac actcgaagga 240  
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300  
aaaggcgcgg acctttttgtg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360  
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgactggggc ccatattttg 420  
atgtataaga agttgacgta a 441

<210> 51

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 51

atgattgaag tcaaaccaat aaatgcggaa gatacgtatg agatcaggca ccgcatactc 60  
cggccgaatc agccgcttga agcatgcaag tatgaaaccg atttgctcag gggtgcggtt 120  
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180  
catccagaac ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagga 240  
taccgtgagc aaaaagcggg tagtacgctt atccgccatg ccgaagagct tcttcggaaa 300  
aaaggcgcgg acctttttgtg gtgcaacgcc aggacatctg cgagagggta ctataaaaag 360  
ctcggcttca gcgaacaagg cgaagtctac gacataccgc cgactggggc ccatattttg 420  
atgtataaga aattgacgta a 441

<210> 52

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 52

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
cggccgaatc agccgctgga agcatgcaag tatgaaaccg attcgtcggg gggcacgttt 120  
cacctcggcg gatattaccg gggcaagctg atcagcatcg cttcctttta tcaagccgaa 180  
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagga 240  
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300  
aaaggcgcgg acctttttatg gtgcaacgcc aggacgtctg cgagcgggta ctataaaaag 360  
ctcggcttca gcgaacaagg cggggtctac gacataccgc cggtcggacc tcatattttg 420  
atgtataaga aattgacgta a 441

<210> 53  
 <211> 441  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> Synthetic DNA Sequence

<400> 53  
 atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
 cggccgaatc agccgctgga agcatgcaag tatgaaaccg atctgcttgg gggcacgttt 120  
 cacctagggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180  
 catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240  
 taccgcgagc aaaaagcggg aagtacgctt atccgccatg ccgaagagct tcttcggaaa 300  
 aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360  
 ctcggttca gcgaacaggg cgaagtctac gacataccgc cggtcggacc tcataattttg 420  
 atgtataaga aattgacgta a 441

<210> 54  
 <211> 441  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> Synthetic DNA Sequence

<400> 54  
 atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcatactc 60  
 cggccgaatc agccgctaga agcatgcaag tatgaaaccg atttgctcag gggcgctttt 120  
 cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagctgaa 180  
 catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240  
 taccgcgagc aaaaagcggg aagtacgctc atccgccatg ccgaagagct tcttcggaaa 300  
 aagggggcag accttttatg gtgcaacgcc aggacgtctg cgagcgggta ctataaaaag 360  
 ctcggttca gcgaacaggg cgaagtctac gacacaccgc cggtcggacc tcataattttg 420  
 atgtataaga agttgacgta a 441

<210> 55  
 <211> 441  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> Synthetic DNA Sequence

<400> 55  
 atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
 cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcag gggcgctttt 120  
 cacctcggcg gatattaccg gggcaagctg gtcagcatcg cctcctttca tcaagccgaa 180  
 catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240  
 taccgcgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300  
 aaaggcgcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360  
 ctcggttca gcgaacaggg cgaagtctac gacacaccgc cggtcggacc tcataattttg 420  
 atgtataaga agttgacgta a 441

<210> 56  
 <211> 441  
 <212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 56

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcatactc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcag gggtagcttt 120
cacctcgggtg gatattaccg gggcaagctg gtcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagga 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg acctttttgtg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacacaccgc cggtcgggacc tcatattttg 420
atgtataaga aattgacata a 441
```

<210> 57

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 57

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gagcacgttt 120
cacctcggcg gatattaccg gggcaagctg gtcagcatcg cttcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagcgct tcttcggaaa 300
aagggggcag acctttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacacaccgc cgaccgggacc ccatattttg 420
atgtataaga aattgacgta a 441
```

<210> 58

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 58

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtaag tatgaaaccg atttgctcgg gggtagcttt 120
cacctcggcg gatattaccg gggcaagctg gtcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagga 240
taccgtgagc aaaaagctgg cagtacgctt atccgccatg ccgaagcgct tcttcggaaa 300
aaaggcgcgg acctttttgtg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gggaacaagg cggggtctac gacataaccgc ctgtcgggacc tcatattttg 420
atgtataaga aattgacgta a 441
```

<210> 59

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 59

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggcogaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggtagcttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagcogaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttgtg gtgcaacgcc aggacgtctg cgagcgggta ctataaaaag 360
ctcggtttca gcgaacaggg cgaagtctac gacataccgc cgactgggac ccatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 60

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 60

```

atgattgaag tcaaaccaat aaatgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggcogaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggtagcttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttta tcaagcogaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac actcgaagga 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttgtg gtgcaacgcc aggacgtctg cgagcgggta ctataaaaag 360
ctcggtttca gcgaacaagg cggggtctac gacataccgc cggtcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 61

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 61

```

atgattgaag tcaaaccaat aaatgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggcogaatc agccgctgga agcatgcaag tatgaaaccg atttgctcag gggtagcttt 120
cacctcgggtg gatattaccg gggcaagctg gtcagcatcg cctcctttca tcaagcogaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac actcgaaggg 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttgtg gtgcaacgcc aggacatctg cgagtgggta ctataaaaag 360
ctcggtttca gcgaacaggg cgaagtctac gacataccgc cgactgggac ccatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 62

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 62

343	86918	344	
atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg agatcaggca ccgcattctc 60
aggccgaatc	agccgctgga	agcatgcaag	tatgaaaccg atttgctcag gggtagcttt 120
cacctcggtg	gatattaccg	gggcaagctg	gtcagcatcg cctcctttca tcaggccgaa 180
catccagagc	ttgaaggcca	aaaacagtat	cagctgagag ggatggcgac acttgaagga 240
taccgtgagc	aaaaagcggg	aagcacgctt	atccgccatg ccgaagcgtt tcttcggaaa 300
aaaggcgcgg	acctttttatg	gtgcaacgcc	aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca	gcgaacaagg	cggggtctac	gacataccgc cggccggacc tcataattttg 420
atgtataaga	aattgacgta	a	441

<210> 63

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 63

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc	agccgctgga	agcatgcaag	tatgaaaccg atttgctcgg gagcacgttt 120
cacctcggtg	gatattaccg	gggcaagctg	atcagcatcg cctcctttca tcaagccgaa 180
catccagagc	ttgaaggcca	aaaacagtat	cagctgagag ggatggcgac actcgaaggg 240
taccgtgagc	aaaaagcggg	aagcacgctc	atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg	acctttttgtg	gtgcaacgcc	aggacgtctg cgagcgggta ctataaaaag 360
ctcggcttca	gcgaacaggg	cgaagtctac	gacacaccgc cggtcggacc tcataattttg 420
atgtataaga	aattgacgta	a	441

<210> 64

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 64

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc	agccgcttga	agcatgtatg	tatgaaaccg atttgctcgg gggtagcttt 120
cacctcggtg	gatattaccg	gggcaagcta	gtcagcatcg cttcctttta tcaagccgaa 180
catccagagc	ttgaaggcca	aaaacagtat	cagctgagag ggatggcgac acttgaagga 240
taccgtgagc	aaaaagcggg	cagtaacgctt	atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg	acctttttgtg	gtgcaacgcc	aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca	gcgaacaggg	cgaagtctac	gacacaccgc cggtcggacc tcattgttttg 420
atgtataaga	aattgacgta	a	441

<210> 65

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 65

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc	agccgctgga	agcatgcaag	tatgaaaccg atttgctcgg gggtagcttt 120
caccttggtg	gatattaccg	gggcaagctg	atcagcatcg cttcctttca tcaagccgaa 180



345	86918	346	
catccagagc	ttgaaggcca	aaaacagtat	cagctgagag ggatggcgac actcgaagga 240
taccgtgagc	aaaaagcggg	cagtacgctt	atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgagg	accttttatg	gtgcaacgcc	aggacatctg cgagcgggta ctataaaaaag 360
ctcggcttca	gcgaacaagg	cggggtctac	gacatgccgc cggtcggacc tcataattttg 420
atgtataaga	agttgacgta	a	441

<210> 66

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 66

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg	agatcaggca	ccgcattctc	60
cgaccgaatc	agccgctgga	agcatgcaag	tatgaaaccg	atttgctcgg	aggcacgttt	120
cacctcggtg	gatattaccg	gggcaagctg	atcagcatcg	cttcctttta	tcaagccgaa	180
catccagagc	ttgaaggcca	aaaacagtat	cagctgagag	ggatggcgac	acttgaagga	240
taccgtgagc	aaaaagcggg	cagtacgctt	atccgccatg	ccgaagcgct	tcttcggaaa	300
aagggggcag	acctcttatg	gtgcaacgcc	aggacatctg	cgagcgggta	ctataaaaaag	360
ctcggcttca	gcgaacaggg	cgaagtctac	gacacaccgc	cggtcggacc	tcataattttg	420
atgtataaga	aattgacgta	a				441

<210> 67

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 67

atgattgaag	tcaaaccaat	aaacgcggag	gatacgtatg	agatcaggca	ccgcattctc	60
cggccgaatc	agccgctgga	agcatgcaag	tatgaaaccg	atttgctcag	gggtgcgttt	120
cacctcggtg	gatattaccg	gggcaagctg	gtcagcatcg	cctcctttca	tcaagccgaa	180
catccagagc	ttgaaggcct	aaaacagtat	cagctgagag	ggatggcgac	actcgaaggg	240
taccgtgagc	aaaaagcggg	aagtacgctc	atccgccatg	ccgaagagct	tcttcggaaa	300
aagggggcag	acctcttatg	gtgcaacgcc	aggacgtctg	cgagcgggta	ctataaaaaag	360
ctcggcttca	gcgaacaggg	cgaagtctac	gacacaccgc	cgaccggacc	tcataattttg	420
atgtataaga	aattgacgta	a				441

<210> 68

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 68

atgattgaag	tcaaaccaat	aaacgcggag	gatacgtatg	agatcaggca	ccgcattctc	60
cggccgaatc	agccgctgga	agcatgcaag	tatgaaaccg	atttgctcag	gggtgcgttt	120
cacctcggtg	gatattaccg	gggcaagctg	gtcagcatcg	cctcctttca	tcaagccgaa	180
catccagagc	ttgaaggcct	aaaacagtat	cagctgagag	ggatggcgac	actcgaaggg	240
taccgtgagc	aaaaagcggg	aagtacgctc	atccgccatg	ccgaagagct	tcttcggaaa	300
aagggggcag	acctcttatg	gtgcaacgcc	aggacgtctg	cgagcgggta	ctataaaaaag	360

347	86918	348	
ctcggcttca	gcgaacaggg	cgaagtctac	gacacaccgc
atgtataaga	aattgacgta	a	cgaccggacc
			tcataattttg
			420
			441

<210> 69

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 69

atgattgaag	tcaaaccaat	aaacgcggag	gatacgtatg	agatcaggca	cgcattctc	60
cggccgaatc	agccgctgga	agcatgcaag	tatgaaaccg	atttgctcgg	gggtacgttt	120
cacctcggcg	gatattaccg	gggcaagttg	gtcagcatcg	cctcctttca	tcaagccaaa	180
catccagagc	ttgaaggcca	aaaacagtat	cagctgagag	ggatggcgac	actcgaaggg	240
taccgtgagc	aaaaagcggg	tagtacgctt	atccgccatg	ccgaagagct	tcttcggaaa	300
aaaggcgcgg	accttttatg	gtgcaacgcc	aggacgtctg	cgagcgggta	ctataaaaaag	360
ctcggcttca	gcgaacaggg	cgaagtctac	gacacaccgc	cggtcggacc	tcataattttg	420
atgtataaga	aattgacgta	a				441

<210> 70

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 70

atgattgaag	tcaaaccaat	aaacgcagaa	gatacgtatg	agatcaggca	cgcattctc	60
cggccgaatc	agccgctgga	agcatgcaag	tatgaaaccg	atttgctcag	gggtgcgttt	120
cacctcggtg	gatattaccg	gggcaagctg	atcagcatcg	cttcctttaa	tcaagccgaa	180
catccagagc	ttgaaggcca	aaaacagtat	cagttgagag	ggatggcgac	acttgaagag	240
taccgtgagc	aaaaagcggg	aagtacgctt	atccgccatg	ccgaagcgtc	tcttcggaaa	300
aaaggggcag	acctcttatg	gtgcaacgcc	aggacatctg	cgagcgggta	ctataaaaaag	360
ctcggcttca	gcgaacaggg	cgaagtctac	gacacaccgc	cggtcggacc	tcataattttg	420
atgtataaga	aattgacgta	a				441

<210> 71

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 71

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg	agatcaggca	cgcattctc	60
cggccgaatc	agccgctgga	agcatgcaag	tatgaaaccg	atttgctcgg	gggtacgttt	120
cacctcggcg	gatattaccg	gggcaagctg	gtcagcatcg	cttcctttca	tcaagccgaa	180
catccagagc	ttgaaggcca	aaaacagtat	cagctgagag	ggatggcgac	acttgaagga	240
taccgtgagc	aaaaagcggg	aagcacgctc	atccgccatg	ccgaagagct	tcttcggaaa	300
aaaggcgcgg	accttttgtg	gtgcaacgcc	aggacatctg	cgagcgggta	ctataaaaaag	360
ctcggcttca	gcgaacaggg	cgaagtctac	gacataaccgc	cggtcggacc	tcataattttg	420
atgtataaga	aattgacgta	a				441

<210> 72  
 <211> 441  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> Synthetic DNA Sequence

<400> 72  
 atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
 cggccgaatc agccgcttga agcatgcaag tatgaaaccg atttgctcgg ggggtacgttt 120  
 cacctcggcg gatattaccg gggcaagctg gtcagcatcg cctcctttca tcaagccgaa 180  
 catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240  
 taccgcgagc aaaaagcggg cagtacgctt atccgccatg ctgaagcgtt tcttcggaaa 300  
 aaaggcggcg accttttgtg gtgcaacgcc aggacatctg caagcgggta ctataaaaag 360  
 ctgggcttca gcgaacaggg cgaagtctac gacataccgc cgactggggc ccataattttg 420  
 atgtataaga aattgacgta a 441

<210> 73  
 <211> 441  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic DNA Sequence

<400> 73  
 atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
 cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcag ggggtgcgttt 120  
 cacctcgggtg gatattaccg gggcaagctg gtcagcatcg cctcctttca tcaagccgaa 180  
 catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagga 240  
 taccgtgagc aaaaagcggg aagtacgctc atccgccatg ccgaagagct tcttcggaaa 300  
 aagggggcag acctcttatg gtgcaacgcc aggacaactg cgagcgggta ctataaaaag 360  
 ctgggcttca gcgaacaggg tgaagtcttc gacataccgc cgaccgggacc ccataattttg 420  
 atgtataaga aattgacgta a 441

<210> 74  
 <211> 441  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic DNA Sequence

<400> 74  
 atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
 cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120  
 cacctagggtg gatattaccg gggcaagctg gtcagcatcg cctcctttca tcaagccgaa 180  
 catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240  
 taccgcgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300  
 aagggggcag acctcttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360  
 ctgggcttca gcgaacaggg cgaagtctac gacataccgc cggtcgggacc tcataattttg 420  
 atgtataaga aattgacgta a 441

<210> 75  
 <211> 441  
 <212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 75

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcag ggggtgcgttt 120
caccttggtg gatattaccg gggcaagctg gtcagcatcg cctcctttaa tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac actcgaagga 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaga 300
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgaccggacc ccatattttg 420
atgtataaga aattgacgta a                                     441
```

<210> 76

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 76

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcatactc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcag ggggtgcgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
caccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagcgct tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc caactgggac ccatattttg 420
atgtataaga aattgacgta a                                     441
```

<210> 77

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 77

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcag ggggtacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagga 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagcgct tcttcggaaa 300
aagggggcag acctcttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgaccggacc ccatattttg 420
atgtataaga aattgacgta a                                     441
```

<210> 78

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 78

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgtatg tatgaaaccg atttgctcag gggtagcttt 120
cacctcggcg gatattaccg gggcaagctg gtcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacacaccgc cggtcgggac tcataatttg 420
atgtataaga agttgacgta a                                     441

```

&lt;210&gt; 79

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 79

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcatactc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggtagcttt 120
cacctcggcg gatattaccg gggcaagctg gtcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat caactgagag ggatggcgac acttgaaggg 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggacgtctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg caaagtctac gacataccgc cggtcgggac tcataatttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 80

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 80

```

atgattgaag tcaaaccaat aaacgcggag gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcag gggtagcttt 120
cacctcggcg gatattaccg gggcaagctg gtcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac actcgaagag 240
taccgcgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag acctcttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cggtcgggac tcataatttg 420
atgtataaga aattgacata a                                     441

```

&lt;210&gt; 81

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 81

355	86918	356	
atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg agatcaggca ccgcattctc 60
cgccgaatc	agccgcttga	agcatgcaag	tatgaaaccg atttgctcag gggtagcttt 120
caccttggtg	gatattaccg	gggcaagctg	atcagcatcg ttccctttca tcaagccgaa 180
catccagagc	ttgaaggcca	aaaacagtat	cagctgagag ggatggcgac acttgaaggg 240
taccgtgagc	aaaaagcggg	cagcacgctt	atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcg	acctttttatg	gtgcaacgcc	aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca	gcgaacagg	cgaagtctac	gacacaccgc cggccggacc tcatattttg 420
atgtatacga	aattgacgta	a	441

<210> 82

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 82

atgattgaag	ttaaaccaat	aaacgcggaa	gatacgtatg agatcaggca ccgcatactc 60
cgccgaatc	agccgcttga	agcatgcaag	tatgaaaccg atttgctcgg gggtagcttt 120
cacctcggtg	gatattaccg	gggcaagctg	gtcagcatcg cctcctttca tcaagccgaa 180
catccagagc	ttgaaggcca	aaaacagtat	cagctgagag ggatggcgac acttgaagag 240
taccgcgagc	aaaaagcggg	cagtacgctt	atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcg	acctttttgtg	gtgcaacgcc	aggacgtctg cgagcgggta ctataaaaag 360
ctcggcttca	gcgaacagg	cgaagtctac	gacataaccgc cggtcggacc tcatattttg 420
atgtataaga	aattgacgta	a	441

<210> 83

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 83

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtata agatcaggca ccgcatactc 60
cgccgaatc	agccgcttga	agcatgtatg	tatgaaaccg atttgctcgg gggcacgctt 120
cacctcggtg	gatattaccg	gggcaagctg	gtcagcatcg cttcctttta tcaagccgaa 180
catccagagc	ttgaaggcca	aaaacagtat	cagctgagag ggatggcgac acttgaaggg 240
taccgtgagc	aaaaagcggg	cagtacgctt	atccgccatg ccgaagcgtt tcttcggaa 300
aaaggcgcg	acctttttatg	gtgcaacgcc	aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca	gcgaacagg	cgaagtctac	gacataaccgc cgaccggacc ccatattttg 420
atgtataaga	aattgacgta	a	441

<210> 84

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 84

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg agatcaggca ccgcattctc 60
cgccgaatc	agccgcttga	agcatgcaag	tatgaaaccg atttgctcgg gggtagcttt 120
cacctcggtg	gatattaccg	gggcaagctg	gtcagcatcg cttcctttca tcaagccgaa 180

357	86918	358
catccagagc ttgaaggcca aaaacagtat	cagctgagag ggatggcgac	acttgaagag 240
taccgcgagc aaaaagcggg aagtacgctt	atccgccatg ccgaagagct	tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc	aggacatctg cgagcgggta	ctataaaaaag 360
ctcggcttca gcgaacaggg cgaagtctac	gacataccgc cgactgggac	ccatattttg 420
atgtataaga aattgacgta a		441
<210> 85		
<211> 441		
<212> DNA		
<213> Artificial Sequence		
<220>		
<223> Synthetic DNA Sequence		
<400> 85		
atgattgaag tcaaaccaat aaacgcggaa	gatacgtatg agatcaggca ccgcatactc	60
cggccgaatc agccgctgga agcatgcaag	tatgaaaccg atttgctcgg gggtagcttt	120
cacctcggtg gatattaccg gggcaagctg	atcagcatcg cctcctttca tcaagccgaa	180
caccagagc ttgaaggcca aaaacagtat	cagctgagag ggatggcgac acttgaagag	240
taccgcgagc aaaaagcggg aagcacgctc	atccgccatg ccgaagagct tcttcggaaa	300
aagggggcag acctcttatg gtgcaacgcc	aggacatctg cgagcgggta ctataaaaaag	360
ctcggcttca gcgaacaggg cgaagtctac	gacataccgc cgaccggacc ccatattttg	420
atgtataaga aattgacgta a		441
<210> 86		
<211> 441		
<212> DNA		
<213> Artificial Sequence		
<220>		
<223> Synthetic DNA Sequence		
<400> 86		
atgattgaag tcaaaccaat aaacgcggaa	gatacgtatg agatcaggca ccgcattctc	60
cggccgaatc agccgctgga agcatgcaag	tatgaaaccg atttgctcag gggtagcttt	120
cacctcggtg gatattaccg gggcaagctg	gtcagcatcg cctcctttca tcaagccgaa	180
catccagagc ttgaaggcca aaaacagtat	cagctgagag ggatggcgac acttgaagag	240
taccgcgagc aaaaagcggg aagcacgctc	atccgccatg ccgaagagct tcttcggaaa	300
aagggggcag accttttatg gtgcaacgcc	aggacatctg cgagcgggta ctataaaaaag	360
ctcggcttca gcgaacaggg caaagtctac	gacataccgc cggtcggacc tcatattttg	420
atgtataaga agttgacgta a		441
<210> 87		
<211> 441		
<212> DNA		
<213> Artificial Sequence		
<220>		
<223> Synthetic DNA Sequence		
<400> 87		
atgattgaag tcaaaccaat aaacgcggaa	gatacgtatg agatcaggca ccgcattctc	60
cggccgaatc agccgctgga agcatgcaag	tatgaaaccg atttgctcgg gggtagcttt	120
cacctcggtg gatattaccg gggcaagctg	atcagcatcg cctcctttca tcaagccgaa	180
catccagagc ttgaaggcca aaaacagtat	cagctgagag ggatggcgac actcgaagag	240
taccgcgagc aaaaagcggg cagtacgctt	atccgccatg ccgaagagct tcttcggaga	300
aaaggcgcgg accttttatg gtgcaacgcc	aggacatctg cgagcgggta ctataaaaaag	360

359

86918

360

ctcggcttca gcgaacaggg cgaagtctac gacataccgc ctgtcggacc tcatattttg 420  
atgtataaga aattgacgta a 441

&lt;210&gt; 88

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 88

atgatcgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggtagcttt 120  
cacctcgggtg ggtactaccg gggcaagctg atcagcatcg cttcctttca taaagccgaa 180  
cattcagagc ttgagggcga agaacagtat cagctgagag ggatggcgac gcttgaagga 240  
taccgtgagc aaaaagcggg cagtacgctt atccgctatg ccgaagagct tcttcgaaaa 300  
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg tgagcgggta ctataaaaag 360  
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420  
atgtataaga aattgacgta a 441

&lt;210&gt; 89

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 89

atgattgaag tcaaacctat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtagcttt 120  
cacctcgggtg gatattacca gggcaagctg atcagcatcg cttcctttca taaagccgaa 180  
cattcagggc ttgagggcga agaacagtat cagctgagag ggatggcgac gctcgaaggg 240  
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcgaaaa 300  
aaaggcgcgg accttttatg gtgcaatgcc aggacatctg tgagcgggta ctatgaaaag 360  
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420  
atgtataaga aattgacgta a 441

&lt;210&gt; 90

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 90

atgattgacg tcaaacctat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggcacgttt 120  
cacctcggcg gatattacca gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180  
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240  
taccgcgagc aaaaagcggg aagtacgctc atccgccatg ccgaagagct tcttcggaaa 300  
aagggggcag accttttatg gtgcaatgcc aggacatctg tgagcgggta ctatgaaaag 360  
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420  
atgtataaga aattgacgta a 441



&lt;210&gt; 91

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 91

```

atgattgaag tcaaaccaat aagcgcggaa gatacgtatg agatcaggca ccgcattctc 60
cgcccggaatc agccgctgga agcatgtatg tatgaaaccg atttgctcgg gggtagcttt 120
cacctcgggtg gatattacca gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgcgagc aaaaagcggg aagtacgctt atccgccatg ccgaagagct tcttcgaaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg tgagcgggta ctatgaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 92

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 92

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cgcccggaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca taaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcaac gcttgaagga 240
taccgtgagc aaaaagcggg aagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcggcta ctatgaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattggcata a                                     441

```

&lt;210&gt; 93

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 93

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cgcccggaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggtagcttt 120
cacctcgggtg gatattaccg gggcaagctg atctgcatcg cctcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaaggg 240
taccgcgagc aaaaagcggg aagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaatgcc aggacatctg tgagaggcta ctatgaaaag 360
ctcggcttca gcgaacaagg cgggtctac gatataccgc cgatcggacc tcatattttg 420
atgtataaga aattggcgta a                                     441

```

&lt;210&gt; 94

&lt;211&gt; 441

&lt;212&gt; DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 94

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ctgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtaacgttt 120
cacctcgggtg gatattacca gggcaagctg atcagcatcg cttcctttca taaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaagg 240
taccgcgagc aaaaagcggg cagtacgtc atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaatgcc aggacatctg tgagcggcta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacgta a 441
```

<210> 95

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 95

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtaacgttt 120
cacctcgggtg gatattacca gggcaagctg atcagcatcg cttcctttca taatgccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaagg 240
taccgtgagc aaaaagcggg aagcacgtc atccgccatg ccgaagagct tcttcggaaa 300
aaggggtag accttttatg gtgcaacgcc aggacatctg tgagcgggta ctataaaaag 360
ctcggcttca gcgaacaagg cgggatctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattggcata a 441
```

<210> 96

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 96

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtaacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagg 240
taccgcgagc aaaaagcggg aagtacgtt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcag accttttatg gtgcaacgcc aggacatctg tgagcggcta ctatgaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacgta a 441
```

<210> 97

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 97

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggtagcttt 120
cacctcgggtg gatattacca gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaaggga 240
taccgtgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcag accttttatg gtgcaacgcc aggacatctg tgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 98

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 98

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga ggcattgtatg tatgaaaccg atttgctcgg gggtagcttt 120
cacctcgggtg gatattacca gggcaagctg atcagcatcg cttcctttca taaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaaggga 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tctccgaaaa 300
aaaggcgcgg accttttatg gtgcaatgcc aggacatctg cgagcgggta ctatgaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccac cgatcggacc tcatattttg 420
atgtataaga aattggcata a                                     441

```

&lt;210&gt; 99

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 99

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggcacgcttt 120
cacctcgggtg gatattatcg ggacaggctg atcagcatcg cttcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaaggga 240
taccgcgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg tgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aactgacgta a                                     441

```

&lt;210&gt; 100

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 100

367

86918

368

```

atgattgaag tcaaacctat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtagcttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaaggg 240
taccgcgagc aaaaagcggg aagtacgctt atccgccatg ccgaagagct tcttcgaaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggatatctg tgagcgggcta ctataaaaag 360
ctcggcttca gcgaacaagg cggggtctac gacataccgc cgatcgggacc tcatattttg 420
atgtataaga aattggcata a 441

```

&lt;210&gt; 101

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 101

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga ggcattgcaag tatgaaaccg atttgctcgg gggcacgttt 120
catctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca taatgccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaagga 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcgaaaa 300
aagggggcag accttttatg gtgcaatgcc aggacatctg cgagcgggcta ctatgaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcgggacc tcatattttg 420
atgtataaga aattggcata a 441

```

&lt;210&gt; 102

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 102

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgtattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtagcttt 120
cacctcggcg gatattatcg ggacaggctg atcagcatcg cctcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaaggg 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggcta ctatgaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcgggacc tcatattttg 420
atgtataaga aattgacgta a 441

```

&lt;210&gt; 103

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 103

```

atgattgaag tcaaaccgat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggtagcttt 120
cacctcgggtg gatattacca gggcaagctg atcagcaccg cttcctttca tcaagccgga 180

```

369

86918

370

```

cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgcgagc gaaaagcggg aagtacgctc atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggatatctg cgagcgggta ctataaaaag 360
ctcggttca gcgaacaagg cggggtctac gatataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 104

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 104

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tttgaaaccg atttgctcgg gggtagcttt 120
cacctcgggtg gatattacca gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaaggg 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcag accttttatg gtgcaacgcc aggacatctg tgagcgggta ctataaaaag 360
ctcggttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 105

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 105

```

atgattgaag tcagaccaat aaacgcggaa gatacgtatg agatcaggca ccgtattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagga 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggttca gcgaacaggg cgaagcctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 106

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 106

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtaogttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaaggg 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360

```

371

86918

372

ctcgggttca gccaacaagg cgggggtctac gacataccgc cgatcggacc ccatattttg 420  
atgtataaga aattgacgta a 441

&lt;210&gt; 107

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 107

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
cgcccgaaatc agccgctgga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120  
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180  
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240  
taccgcgagc aaaaagcggg aagtacgctt atccgccatg ccgaagagct tcttcggaaa 300  
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360  
ctcgggttca gccaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420  
atgtataaga aattgacgta a 441

&lt;210&gt; 108

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 108

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgtattctc 60  
cgcccgaaatc agccgctgga agcatgtatg taagaaaccg atttgctcgg gggcgctttt 120  
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tccagccgaa 180  
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggga 240  
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300  
aaaggcgagg accttttatg gtgcaacgcc aggatatctg cgagcgggta ctatgaaaag 360  
ctcgggttca gccaacaggg cgaagtctac gacataccgc cgatcggacc ccatattttg 420  
atgtataaga aattgacgta a 441

&lt;210&gt; 109

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 109

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
cgcccgaaatc agccgctgga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120  
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180  
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggga 240  
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300  
aagggcgagg accttttatg gtgcaacgcc aggacatctg cgagcgggta ctatgaaaag 360  
ctcgggttca gccaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420  
atgtataaga aattgacgta a 441

&lt;210&gt; 110

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 110

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgtattctc 60
cgcccggaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagga 240
taccgtgagc aaaaagcggg aagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggatatctg cgagcggcta ctataaaaag 360
ctcggcttca gcgaacaagg cgggtcttac gacataccgc cggtcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 111

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 111

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cgcccggaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgtgggc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg cgagcggcta ctatgaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 112

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 112

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgtattctc 60
cgcccggaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcggcg gatattatca ggacaggctg atcagcatcg cctcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgcgagc aaaaagcggg aagtacgctt atccgctatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggatatctg cgagcggcta ctatgaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 113

&lt;211&gt; 441

&lt;212&gt; DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 113

```
atgattgaag tcaaacctat aaacgcggaa gatacgatg agatcaggca ccgcatactc 60
cgcccgaaac agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtagcttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaagg 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcggcta ctatgaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacgta a 441
```

<210> 114

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 114

```
atgattgaag tcaaaccaat aaacgcggaa gatacgatg agatcaggca ccgcattctc 60
cgcccgaaac agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattatcg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaagag 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacgta a 441
```

<210> 115

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 115

```
atgattgaag tcaaacctat aaacgcggaa gatacgatg agatcaggca ccgcattctc 60
cgcccgaaac agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttgtg gtgcaacgcc aggatatctg cgagcgggta ctatgaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacgta a 441
```

<210> 116

<211> 441

<212> DNA

<213> Artificial Sequence

<220>



## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 116

```

ttacgtcaat ttcttataca tcaaaatatg aggtccgata gccggtatgt cgtagacttc 60
gccctgttcg ctgaagccga gctttttata gtaccgctc gcagatgtcc tggcggtgca 120
ccataaaagg tccgcgcctt tttccgaag aagctcttcg gcattggcga tgagcgtgct 180
tcccgctttt tgctcgcggt acccttcaag cgtcgccatc cctctcagct gatactgttt 240
ttggccttca agctctgaat gttcggttg atgaaaggag gcgatgctga tcagcttgcc 300
ccggtaatat ccaccgaggt gaaacgtgcc cccgagcaaa tcagtttcat acttgcattgc 360
ttccagcggc tgattcggcc ggagaatgcg gtgcctgata tcatacgtat cttccgcgtt 420
tattggtttg gcttcaatca t

```

&lt;210&gt; 117

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 117

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggtacgttt 120
cactcggcg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaagg cggggtctac gatataccgc cgatcggacc tcataatttg 420
atgtataaga aattgacgta a

```

&lt;210&gt; 118

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 118

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggcacgttt 120
cactcggcg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgag 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaagag 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctatgaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcataatttg 420
atgtataaga aattgacata a

```

&lt;210&gt; 119

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 119

379

86918

380

```

atgattgaag tcaatccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcacgtct 120
cacctcggcg gatattaccg gggcaagctg atcagcatcg cttcctttca taatgccgaa 180
cattcagagc ttgatggcca aaaacagtat cagctgagag ggatggcgac acttgaagg 240
taccgcgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgagc accttttatg gtgcaacgcc aggacatctg tgagcggcta ctatgaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcataatttcg 420
atgtataaga aattgacgta a

```

&lt;210&gt; 120

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 120

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggcacgttt 120
cacctcggcg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaagg 240
taccgcgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaagg cggggtctac gatataccgc cgatcggacc tcataattttg 420
atgtataaga aattggcata a

```

&lt;210&gt; 121

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 121

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggcgcttt 120
cacctcggcg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaatccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaagg 240
taccgcgagc taaaagcggg aagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggatatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcataattttg 420
atgtataaga aattgacgta a

```

&lt;210&gt; 122

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 122

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggt gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180

```

381	86918	382	
cattcagagc	ttgaaggcca	aaaacagtat	cagctgagag ggatggcgac acttgaagag 240
taccgcgagc	aaaaagcggg	aagcacgctc	atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg	accttttatg	gtgcaacgcc	aggatatctg cgagcgggta ctataaaaag 360
ctcggcttca	gcgaacaagg	cggggtctac	gatataaccg cgatcggacc tcatattttg 420
atgtataaga	aattgacgta	a	441

<210> 123

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 123

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg	agatcaggca	ccgcattctc	60
cggccgaatc	agccgcttga	agcatgtatg	tatgaaaccg	atttgctcgg	gggcacgttt	120
cacctcggtg	gatattaccg	gggcaagctg	atcagcatcg	cctcctttca	tcaagccgaa	180
cattcagagc	ttgaaggcca	aaaacagtac	cagctgagag	ggatggcgac	gcttgaaggg	240
taccgcgagc	aaaaagcggg	cagtacgctt	atccgccatg	ccgaagagct	tcttcggaaa	300
aagggggcag	accttttatg	gtgcaacgcc	aggacatctg	cgagcgggta	ctataaaaag	360
ctcggcttca	gcgaacaagg	cggggtctac	ggcataaccg	cgatcggacc	tcatattttg	420
atgtataaga	aattgacata	a				441

<210> 124

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 124

atgattgaag	ccaaaccaat	aaacgcggaa	gatacgtatg	agatcaggca	ccgcattctc	60
cggccgaatc	agccgctgga	agcatgcaag	tatgaaactg	atttgctcgg	gggcacgttt	120
cacctcggtg	gatattaccg	gggcaagctg	atcagcatcg	cctcctttca	tcaagccgaa	180
cattcagagc	ttgaaggcca	aaaacagtat	cagctgagag	ggatggcgac	gcttgaaggg	240
taccgcgagc	aaaaagcggg	aagcacgctc	atccgccatg	ccgaagagct	tcttcggaaa	300
aaaggcgcgg	accttttatg	gtgcaacgcc	aggacatctg	cgagcgggta	ctataaaaag	360
ctcggcttca	gcgaacaggg	cgaagtctac	gacataaccg	cgatcggacc	tcatattttg	420
atgtataaga	aattgacgta	a				441

<210> 125

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 125

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg	agatcaggca	ccgcatactc	60
cggccgaatc	agccgctgga	agcatgcaag	tatgaaaccg	atttgctcgg	gggcacgttt	120
cacctcggtg	gatattaccg	gggcaagctg	atcagcatcg	cttcctttca	tcaagccgaa	180
cattcagagc	ttgaaggcca	aaaacagtat	cagctgagag	ggatggcgac	gcttgaaggg	240
taccgcgagc	aaaaagcggg	tagtacgctt	atccgccatg	ccgaagagct	tcttcggaaa	300
aagggggcag	accttttatg	gtgcaacgcc	aggacatctg	cgagcgggta	ctataaaaag	360

383 86918 384  
 ctcggtctca gcgaacatgg cgaagtctac gacataccgc cgatcggacc tcatattttg 420  
 atgtataaga aattgacgta a 441

<210> 126

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 126

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
 cggccgaatc agccgctgga agcatgcatg tatgaaaccg atttgctcgg gggcacgttt 120  
 caccctcggg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180  
 cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240  
 taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300  
 aagggggcag accttttatg gtgcaacgcc aggacatctg tgagcgggta ctataaaaag 360  
 ctcgggttca gcgaacaggc cgaagtctac gacataccgc cgatcggacc tcatattttg 420  
 atgtataaga aattgacgta a 441

<210> 127

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 127

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
 cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcag gggcgctttt 120  
 caccctcggg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180  
 catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagga 240  
 taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300  
 aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctatgaaaag 360  
 ctcgggttca gcgaacaggc cgaagtctac gacataccgc cgatcggacc tcatattttg 420  
 atgtataaga aattaacata a 441

<210> 128

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 128

atgattgagg tgaaaccgat taacgcagag gagacctatg aactaaggca taggatactc 60  
 agaccacacc agccgataga ggtttgatg tatgaaaccg atttacttcg tggcgctttt 120  
 cacttaggcg gcttttacag gggcaagctg atttccatag cttcattcca ccaggccgag 180  
 catccagaac tccagggcca gaaacaatac caactccgag gtatggctac cttggaagg 240  
 tatcgtgacc agaaaagcggg atcgagccta attaaacacg ctgaacagat ccttcggaag 300  
 cggggggcgg acatgctatg gtgcaatgcg cggacatccg ccgctgggta ctacaaaaag 360  
 ttaggcttca gcgagcaggg agagggtattt gaaacgcgcg cagtaggacc tcacatcgta 420  
 atgtataaac gcctcacata a 441

&lt;210&gt; 129

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 129

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcggtg gatattaccg aggcaagctg atcagcatcg cctccttcca tcaagccgaa 180
cattcagagc ttgaaggcca taaacagtat cagctgagag ggatggcgac acttgaagag 240
taccgcgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaagg cggggtctac gacataccgc cggtcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 130

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 130

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtaag tatgaaaccg atttgctcgg gggcacgttt 120
cacctcggcg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgcgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cggtcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 131

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 131

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggcacgttt 120
cacctcggtg gatattaccg gggcaagctg atcagcatcg cctcctttca ccaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagga 240
taccgtgagc aaaaagcggg aagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaagg cggggtctac gacataccgc cggtcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 132

&lt;211&gt; 441

&lt;212&gt; DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 132

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgta 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagga 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag acctttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaagg cggggtctac gacataccgc cggtcggacc tcataatttg 420
atgtataaga aattgacgta a                                     441
```

<210> 133

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 133

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgcaag tatgaagccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgag 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagga 240
aaccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag acctttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacgtaccgc cgatcggacc tcataatttg 420
atgtataaga aattgacgta a                                     441
```

<210> 134

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 134

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcagg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattatcg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaagga 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag acctttttatg gtgcaacgcc aggataatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtttac gacataccgc cggtcggacc tcataatttg 420
atgtataaga aattgacgta a                                     441
```

<210> 135

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 135

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggcm ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtaacgtt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgcgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgagg accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaagg cggggtctac gacataccgc cggtcgggacc tcatattttg 420
atgtataaga aattgacgta a

```

&lt;210&gt; 136

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 136

```

atgattgaag ccaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtaacgtt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgcgagc aaaaagcggg aagtaacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgaccgggacc ccatattttg 420
atgtataaga aattgacgta a

```

&lt;210&gt; 137

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 137

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcag gggtagcttt 120
cacctcgggtg gatattatcg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac actcgaagga 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggcgagc accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cggtcgggacc tcatattttg 420
atgtataaga aattgacgta a

```

&lt;210&gt; 138

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 138

391	86918	392	
atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc	agccgctgga	agcatgcaag	tatgaaaccg atttgctcgg gggtagcttt 120
cacctcggcg	gatattaccg	gggcaagctg	atcagcatcg cctcctttca tcaagccgaa 180
catccagagc	ttgaaggcca	aaaacagtat	cagctgagag ggatggcgac acttgaagga 240
taccgtgagc	aaaaagcggg	cagtacgctt	atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcg	accttttatg	gtgcaacgcc	aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca	gcgaacagg	cgaagtctac	gacataccgc cgatcggacc tcataattttg 420
atgtataaga	aattgacgta	a	441

<210> 139

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 139

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc	agccgctgga	agcatgcaag	tatgaaaccg atttgctcgg gggtagctta 120
cacctcgggtg	gatattaccg	gggcaagctg	atcagcatcg cctcctttca tcaagccgaa 180
catccagagc	ttgaaggcca	aaaacagtat	cagctgagag ggatggcgac acttgaagga 240
taccgtgagc	aaaaagcggg	cagtacgctt	atccgccaa g ccgaagagct tcttcggaaa 300
aagggggcag	accttttatg	gtgcaacgcc	aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca	gcgaacaagg	cggggtctac	gacataccgc cggtcggacc tcataattttg 420
atgtataaag	aattgacgta	a	441

<210> 140

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 140

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc	agccgcttga	agcatgtatg	tatgaaaccg atttgctcag gggcacgttt 120
cacctcggcg	gatattaccg	gggcaagctg	atcagcatcg cttcctttca tcaagccgaa 180
cattcagatc	ttgaaggcca	aaaacagtat	cagctgagag ggatggcgac acttgaagag 240
taccgcgagc	aaaaagcggg	aagcacgctc	atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag	accttttatg	gtgcaacgcc	aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca	gcgaacagg	cgaagtctac	gacataccgc cggtcggacc tcataattttg 420
atgtataaga	aattgacgta	a	441

<210> 141

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 141

atgattgaag	tcaaaccaat	aaacgcggga	gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc	agccgctgga	agcatgcaag	tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg	gatattaccg	gggcaagctg	atcagcatcg cctcctttca tcaagccgaa 180



393	86918	394	
cattcagagc	ttgaaggcca	aaaacagtat	cagctgagag
taccgcgagc	aaaaagcggg	cagtagcgtt	atccgccatg
aaaggcgcg	accttttatg	gtgcaacgcc	aggacatctg
ctcggcttca	gcgaacaagg	cggggtctac	gacataccgc
atgtataaga	aattgacgta	a	

<210> 142

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 142

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg	agatcaggca	ccgcattctc	60
cggccgaatc	agccgctgga	agcatgcaag	tatgaaaccg	atttgctcgg	gggcacgttt	120
cacctcggtg	gatattaccg	gggcaagctg	atcagcatcg	cctcctttca	tcaagccgaa	180
cattcagagc	ttgaaggcca	aaaacagtat	cagctgagag	ggatggcgac	acttgaagga	240
taccgtgagc	aaaaagcggg	aagtagcgtt	atccgccatg	ccgaagagct	tcttcggaaa	300
aaaggcgcg	accttttatg	gtgcaacgcc	aggacatctg	cgagcgggta	ctataaaaag	360
ctcggcttca	gcgaacagg	cgaagtctac	gacataccgc	cgatcggacc	tcatattttg	420
atgtataaga	aattgacgta	a				441

<210> 143

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 143

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg	agatcaggca	ccgcgttctc	60
cggccgaatc	agccgctgga	agcatgtatg	tatgaaaccg	atttgctcgg	gggtgcgttt	120
cacctcggtg	gatattaccg	gggcaagctg	atcagcatcg	cctcctttca	tcaagccgaa	180
catccagagc	ttgaaggcca	aaaacagtat	cagctgagag	ggatggcgac	acttgaaggg	240
taccgcgagc	aaaaagcggg	cagtagcgtt	atccgccatg	ccgaagagct	tcttcggaaa	300
aaaggcgcg	accttttatg	gtgcaacgcc	aggacatctg	cgagcgggta	ctataaaaag	360
ctcggcttca	gcgaacagg	cgaagtctac	gacgtaccgc	cggtcggacc	tcatattttg	420
atgtataaga	aattgacgta	a				441

<210> 144

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 144

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg	agatcaggca	ccgcattctc	60
cggccgaatc	ggccgcttga	agcatgtatg	tatgaaaccg	atttgctcgg	gggcacgttt	120
cacctcggtg	gatattaccg	gggcaagctg	atcagcatcg	cctcctttca	tcaagccgaa	180
catccagggc	ttgaaggcaa	aaaacagtat	cagctgagag	ggatggcgac	acttgaagag	240
taccgcgagc	aaaaagcggg	aagcacgctt	atccgccatg	ccgaagagct	tcttcggaaa	300
aaaggcgcg	accttttatg	gtgcaacgcc	aggacttccg	cgagcgggta	ctataaaaag	360

395	86918	396	
ctcggcttca	gccaacaagg	aggggtctac	gacataccgc
atgtataaga	aattgacgta	a	cggtcggacc
			tcataattttg
			420
			441

<210> 145

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 145

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg	agatcaggca	cogtattctc	60
cggccgaatc	agccgctgga	agcatgcatg	tatgaaaccg	atttgctcga	gggcacgttt	120
cacctcggtg	gatattaccg	gggcaagctg	atcagcatcg	cctcctttca	tcaagccgag	180
cattcagagc	ttgaaggcca	aaaacagtat	cagctgagag	ggatggcgac	acttgaagga	240
taccgtgagc	aaaaagcggg	cagtacgctt	atccgccatg	ccgaagagct	tcttcggaaa	300
aaaggcgcg	accttttatg	gtgcaacgcc	aggacatctg	cgagcgggta	ctataaaaaag	360
ctcggcttca	gcgaacagg	cggaagtctac	gacataccgc	cggtcggacc	tcataattttg	420
atgtataaga	aattgacgta	a				441

<210> 146

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 146

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg	agatcaggca	ccgcattctc	60
cggccgaatc	agccgcttga	agcatgtatg	tatgaaaccg	atttgctcgg	gggcacgttt	120
cacctcggtg	gatattaccg	gggcaagctg	atcagcatcg	cctcctttca	tcaagccgaa	180
cattcagagc	ttgaaggcca	aaaacagtat	cagctgagag	ggatggcgac	gcttgatgag	240
taccgcgagc	aaaaagcggg	cagtacgctt	atccgccatg	ccgaagagct	tcttcggaaa	300
aaaggcgcg	accttttatg	gtgcaacgcc	aggacatctg	cgagcgggta	ctataaaaaag	360
ctcggcttca	gcgaacagg	cggaagtctac	gacataccgc	cgatcggacc	ccataattttg	420
atgtataaga	aattgacgta	a				441

<210> 147

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 147

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg	agatcaggca	ccgcattctc	60
cggccgaatc	agccgcttga	agcatgtatg	tatggaactg	atttgctcgg	gggcacgttt	120
cacctcggtg	gatattaccg	gggcaagctg	atcagcatcg	cttcctttca	tcaagccgag	180
caaccagagc	ttgaaggcca	aaaacagtat	cagctgagag	ggatggcgac	acttgaagg	240
taccgcgagc	aaaaagcggg	cagtacgctt	atccgccatg	ccgaagagct	tcttcggaaa	300
aagggggcg	accttttatg	gtgcaacgcc	aggacatctg	cgagcgggta	ctataaaaaag	360
ctcggcttca	gcgagcaagg	cgggggtctac	gacataccgc	cggtcggacc	tcataattttg	420
atgtataaga	aattgacgta	a				441

&lt;210&gt; 148

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 148

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaagga 240
taccgcgagc aaaaagcggg aagtacgctt atccgccatg cagaagagat tcttcggaaa 300
aaaggcgcgg acctcttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaagg cggggtctac gacataccgc cggtcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 149

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 149

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcgttctc 60
cggccgaatc agccgctgga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaaccgtat cagctgagag ggatggcgac acttgaagga 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaaatctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 150

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 150

```

atgattgaaa tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtagcttt 120
cacctcgggtg gatattaccg aggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac tcttgaagga 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgagggtctac gacataccgc cggtcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 151

&lt;211&gt; 441

&lt;212&gt; DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 151

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cgcccgaaatc agccgctgga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagga 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg cgagcggcta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cggtcggacc tcatattttg 420
atgtataaga aattgacgta a 441
```

<210> 152

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 152

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cgcccgaaatc agccgctgga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240
tacgcgcagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag acctcttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cggtcggacc tcatattttg 420
atgtataaga aattgacata a 441
```

<210> 153

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 153

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cgcccgaaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaaaactg atcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaagga 240
taccgtgagc aaaaagcggg aagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg cgagcggcta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cggtcggacc tcatattttg 420
atgtataaga aattgacgta a 441
```

<210> 154

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 154

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcgtactc 60
cgcccggaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcggcatcg cttcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctatgaaaag 360
ctcggcttca gcggacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacata a                                     441

```

&lt;210&gt; 155

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 155

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cgcccggaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 156

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 156

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cgcccggaatc agccgctgga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagcta atcagcatcg cttcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240
taccgcgagc aaaaagcggg aagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaaag 360
ctcggcttca gcgaacaggg cgaagtccac gacataccgc cggtcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 157

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 157

403	86918	404	
atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc	agccgctgga	agcatgtatg	tatgaaaccg atttgctcgg gggcacgttt 120
cacctcggcg	gatattaccg	gggcaagcta	atcagcatcg cctcctttca tcaagccgaa 180
cattcagagc	ttgaaggcca	aaaacagtat	cagctgagag ggatggcgac gcttgaagag 240
taccgtgagc	aaaaagcggg	cagtacgctt	atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag	acctttttatg	gtgcaacgcc	aggatatctg cgagcgggta ctataaaaag 360
ctcggcttca	gcgaacaagg	cggggtctac	gacataccgc cggtcggacc tcataatttg 420
atgtataaga	aattgacgta	a	441

<210> 158

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 158

atgattgaag	tcaaaccaat	aaacgcggaa	gatgcgtatg agatcaggca ccgcattctc 60
cggccgaatc	agccgctgga	agcatgcaag	tatgaaactg atttgctcgg gggcacgttt 120
cacctcggcg	gatattaccg	gggcaagctg	atcagcatcg cttcctttca tcaagccgaa 180
cattcagagc	ttgaaggcca	aaaacagtat	cagctgagag ggatggcgac acttgaaggg 240
taccgcgagc	aaaaagcggg	cagtacgctt	atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg	acctttttgtg	gtgcaacgcc	aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca	gcgaacaagg	cggggtctac	gacataccgc cggtcggacc tcataatttg 420
atgtataaga	aattgacgta	a	441

<210> 159

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 159

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc	agccgcttga	agcatgtatg	tatgaaaccg atttgctcag gggcgcgttt 120
cacctcggtg	gatattaccg	gggcaagctg	atcagcatcg cttcctttca tcaagccgaa 180
cattcagagc	ttgaaggcca	aaaacagtat	cagctgagag ggatggcgac acttgaaggga 240
taccgtgagc	aaaaagcggg	cagtacgctt	atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag	acctttttatg	gtgcaacgcc	aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca	gcgaacaagg	cggggtctac	gacataccgc cggtcggacc tcataatttg 420
atgtataaga	aattgacgta	a	441

<210> 160

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 160

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc	agccgctgga	agcatgtatg	tatgaaaccg atttgctcgg gggcacgttt 120
cacctcggtg	gatattaccg	gggcaagctg	atcagcatcg cctcctttca tcaagccgaa 180

405

86918

406

```

catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaaggg 240
taccgcgagc aaaaagcggg aagtaagctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaagg cggggtctac gacataccgc cggtcggacc tcatattttg 420
atgtataaga aattgacgta a 441

```

&lt;210&gt; 161

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 161

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggcacgctt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaaggg 240
taccgcgagc aaaaagcggg aagtaagctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaagg cgaagtctac gacataccgc cggtcggacc tcatattttg 420
atgtataaga aattgacgta a 441

```

&lt;210&gt; 162

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 162

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg gtctgcacg cttcctttca taaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgatgga 240
taccgcgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcgaaaa 300
aaaggcgcgg accttttatg gtgcaatgcc aggacatctg tgagcgggta ctatgaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cggtcggacc tcatattttg 420
atgtataaga aattgacgta a 441

```

&lt;210&gt; 163

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 163

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgag 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240
taccgcgagc aaaaagcggg cagtaagctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360

```

ctcggcttca gcgaacaagg cggggtctac gatataccgc cgatcggacc tcatattttg 420  
atgtataaga aattgacgta a 441

<210> 164

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 164

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
cggccgaatc agccgctgga agcatgcaag tatgaaactg atttgctcgg gggcacgttt 120  
catctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180  
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaaggg 240  
taccgcgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300  
aaaggcgagg accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360  
ctcggcttca gcgaacaagg cggggtctac gatataccgc cgatcggacc tcatattttg 420  
atgtataaga aattgacgta a 441

<210> 165

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 165

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120  
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180  
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240  
taccgcgagc aaaaagcggg aagtacgctt atccgccatg ccgaagagct tcttcggaaa 300  
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360  
ctcggcttca gcgaacaagg cggggtctac gacataccgc cggtcggacc tcatattttg 420  
atgtataaga aattgacgta a 441

<210> 166

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 166

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggcacgttt 120  
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180  
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaaggg 240  
taccgtgagc aaaaagcggg aagtacgctc atccgccatg ccgaagagct tcttcggaaa 300  
aaaggcgagg accttttatg gtgcaacgcc aggacatctg cgagtgggta ctataaaaag 360  
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cggtcggacc tcatattttg 420  
atgtataaga aattgacgta a 441



&lt;210&gt; 167

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 167

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagga 240
taccgtgagc aaaaagtggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaagg cgggtcttac gatataccgc cggtcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 168

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 168

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg aaatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagga 240
taccgcgagc aaaaagcggg aagtacgctc atccgccatg ccgaagagct tctacggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 169

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 169

```

atgattgaag tcaaaccaat aaacgcggag gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcag gggcgcggtt 120
caccttgggtg gatattaccg gggcaagctg gtcagcatcg cctcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240
taccgcgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag acctcttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgactggggc ccatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 170

&lt;211&gt; 441

&lt;212&gt; DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 170

```
atgattgaag tcaaaccaat aaacgcggag gatacgtatg agatcaggca ccgcattctc 60
cgcccgaaatc agccgctgga agcatgcaag tatgaaaccg atttgctcag gggtagcttt 120
caccttggtg gatattaccg gggcaagctg gtcagcatcg cctcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240
taccgcgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag acctcttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaaa 360
ctcggtttca gcgaacaggg cgaagtctac gacataccgc cgactggggc ccatattttg 420
atgtataaga aattgacgta a 441
```

<210> 171

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 171

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cgcccgaaatc agccgctgga cgcattgcaag tatgaaaccg atttgctcgg gggcacgttt 120
cacctcggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag acctottatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaaa 360
ctcggtttca gcgaacaggg cgaagtctac gacataccgc cggtcggacc tcatattttg 420
atgtataaga aattgacgta a 441
```

<210> 172

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 172

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcatactc 60
cgcccgaaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtagcttt 120
cacctcggtg gatattaccg gggcaagctg gtcagcatcg cctcctttca tcaagctgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagcgct tcttcgggag 300
aaaggcgctg accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaaa 360
ctcggtttca gcgaacaggg cgaagtctac gacacaccgc cggtcggacc ccatattttg 420
atgtataaga agttgacgta a 441
```

<210> 173

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 173

```

atgattgaag tcaagccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cgcccggaatc agccgcttga agcatgtatg tatgaaaccg atatgctcag gggcgcttt 120
cacctcggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240
taccacgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttta gcgaacaggg cgaagtctac aacacaccgc cggttggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 174

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 174

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cgcccggaatc agccgctgga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggc gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180
catccagagc ttgtaggcca aaaacagtat cagctgagag ggatggcgac acttgaagga 240
taccgtgagc aaaaagcggg cagtacactt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cggtcgggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 175

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 175

```

atgattgaag taaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcgttctc 60
cgcccggaatc agccgctgga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcggtg gatattaccg gggcgagctg gtcagcatcg cttcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttggt gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca gcgaacaagg cgggtctac gacataccgc cggtcgggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 176

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 176

415	86918	416	
atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg agatcaggca ccgcattctc 60
cgccgaatc	agccgctgga	agcatgtatg	tatgaaactg atttgctcgg gggcacgttt 120
cacctcggtg	gatattaccg	gggcaagctg	atcagcatcg cctcctttca tcaagccgaa 180
catccagagc	ttgaaggcca	aaaacagtat	cagctgagag ggatggcgac actcgaagga 240
taccgcgagc	aaaaagcggg	cagtacgcta	atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag	acctcttatg	gtgcaacgcc	aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca	gcgatcaggg	cgaagtctac	gacataccgc cggtcggacc tcatattttg 420
atgtataaga	aattgacgta	a	441

<210> 177

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 177

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg agataaggca ccgcattctc 60
cgccgaatc	agccgctgga	agcatgcaag	tatgaaaccg atttgctcgg gggcacgttt 120
cacctcggtg	gatattaccg	gggcaagctg	gtcagcatcg cctcctttca tcaagccgaa 180
catccagagc	ttgaaggcca	aaaacagtat	cagctgagag ggatggcgac acttgaaggg 240
taccgtgagc	aaaaagcggg	cagtacgctt	atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag	accttttatg	gtgcaacgcc	aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca	gcgaacaggg	cgaagtctac	gacttaaccgc cgaccggacc ccatattttg 420
atgtataaga	aattgacgta	a	441

<210> 178

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 178

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg agatcaggca ccgcattctc 60
cgccgaatc	agccgctgga	agcatgcaag	tatgaaaccg atttgctcgg gggcacgctt 120
cacctcggtg	gatattaccg	gggcaagctg	gtcagcattg cttcctttca tcaagccgaa 180
catccagagc	ttgagggcca	aaaacagtat	cagctgagag ggatggcgac acttgaaggg 240
taccgcgggc	aaaaagcggg	cagtacgctt	atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag	accttttatg	gtgcaatgcc	aggacatctg cgagcgggta ctataaaaag 360
ctcggcttca	gcgaacaagg	cggggtctac	gacataccgc cggtcggacc tcatattttg 420
atgtataaga	aattgacgta	a	441

<210> 179

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 179

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg agatcaggca ccgcattctc 60
agccgaatc	agccgctaga	agcatgcaag	tatgaaaccg atttgctcag gggcgcgttt 120
cacctcggtg	gatattaccg	gggcaagctg	atcagcatcg cctcctttca tcaagccgaa 180

417	86918	418	
cattcagagc	ttgaaggcca	aaaacagtat	cagctgagag ggatggcgac acttgaagga 240
taccgtgagc	aaaaagcggg	cagtaacgctt	atccgccatg ccgaagcgct tcttcggaaa 300
aaaggcgagg	acctttttgtg	gtgcaacgcc	aggacgtctg cgagcgggta ctataaaaaag 360
ctcggcttca	gcgaacaggg	cgaagtctac	gacataaccgc cgaccgggacc ccatattttg 420
atgtataaga	aattgacgta	a	441

<210> 180

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 180

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg	agatcaggca	ccgcattctc	60
cggccgaatc	agccgctgga	agcatgcaag	tatgaaactg	atttgctagg	gggtacgctt	120
cacctcgggtg	gatattaccg	gggcaagctg	atcagcatcg	cctcctttca	tcaagccgaa	180
catccagagc	ttgaaggcca	aaaacagtat	cagctgagag	ggatggcgac	acttgaaggg	240
taccgtgagc	aaaaagcggg	cagtaacgctt	atccgccatg	ccgaagagct	tcttcggaaa	300
aaggggcgagg	accttatatg	gtgcaacgcc	aggacatctg	cgagcgggta	ctataaaaaag	360
ctcggcttca	gcgaacaggg	cgaagtctac	gacataaccgc	cggtcgggacc	tcatattttg	420
atgtataaga	aattgacata	a				441

<210> 181

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 181

atgattgaag	tcaaaccaat	aaatgcggaa	gatacgtatg	agatcaggca	ccgcattctc	60
cggccgaatc	agccgcttga	agcatgtatg	tatgaaaccg	atttgctcgg	aggcacgttt	120
cacctcgggtg	gatattaccg	gggcaagctg	atcagcatcg	cctcctttta	tcaagccgaa	180
catccagagc	ttgaaggcca	aaaacagtat	cagctgagag	ggatggcgac	acttgaaggg	240
taccgtgagc	aaaaagcggg	cagtaacgctt	atccgccatg	ccgaagagct	tcttcggaaa	300
aagggggcag	acctttttatg	gtgcaacgcc	aggacatctg	cgagcgggta	ctataaaaaag	360
ctcggcttca	gcgaacaggg	cgaagtctac	gacataaccgc	cggtcgggacc	tcatattttg	420
atgtataaga	aattgacata	a				441

<210> 182

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 182

atgattgaag	tcaaaccaat	aaccgcggaa	gatacgtatg	agatcaggca	ccgcattctc	60
cggccgaatc	agccgctgga	agcatgcaag	tatgaaaccg	atttgctcgg	gggtacgttt	120
cacctcgggtg	gatattaccg	gggcaagctg	atcagcatcg	cctcctttca	tcaagccgaa	180
catccagagc	ttgaaggcca	aaaacagtat	cagctgagag	ggatggcgac	actagaaggg	240
taccgcgagc	aaaaagcggg	cagtaacgctc	atccgccatg	ccgaagagct	tcttcggaaa	300
aagggggcag	acctttttatg	gtgcaacgcc	agaacatctg	cgagcgggta	ctataaaaaag	360

419

86918

420

ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgaccggacc ccatattttg 420  
atgtataaga aattgacgta a 441

&lt;210&gt; 183

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 183

atgattgaag tcaaaccaat aaacgcggaa gatgcgtatg agatcaggca ccgcattctc 60  
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gagcacgttt 120  
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cctcctttca tcaagccgaa 180  
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaagag 240  
taccgcgagc aaaaagcggg aagtacgctt atccgccatg ccgaagagct tcttcggaaa 300  
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg cgagcgggta ctataaaaag 360  
ctcggcttca gcgaacaggg cgaagtctac gacataccgc ctgtcggacc tcatattttg 420  
atgtataaga aattgacgta a 441

&lt;210&gt; 184

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 184

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggtagcttt 120  
cacctcgggtg gatattaccg gggcaagctg atctgcacatg cctcctttca tcaagccgaa 180  
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaaggg 240  
taccgcgagc aaaaagcggg aagtacgctt atccgccatg ccgaagagct tcttcggaaa 300  
aagggggcag accttttatg gtgcaatgcc aggacatctg tgagaggcta ctatgaaaag 360  
ctcggcttca gcgaacaagg cggggtctac gatataccgc cgatcggacc tcatattttg 420  
atgtataaga aattggcgta a 441

&lt;210&gt; 185

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 185

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60  
cggccgaatc agccgttgga agcatgcaag tatgaaaccg atttgctcgg gggtagcttt 120  
cacctcggcg gatattaccg gggcaagctg atcagcatcg cttcctttca taaagccgaa 180  
cattcagagc ttgagggcga agaacagtat cagctgagag ggatggcgac gcttgaagga 240  
taccgtgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300  
aagggggcag accttttatg gtgcaatgcc aggacatctg tgagcgggta ctataaaaag 360  
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420  
atgtataaga aattgacgta a 441

&lt;210&gt; 186

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 186

```

atgatagaag tgaaaccgat taacgcagag gatacctatg aactaaggca taaaatactc 60
agaccaaacc agccgataga agcgtgtatg tatgaaagcg atttacttcg tgggtgcattt 120
cacttaggcg gcttttacag gggcaaactg atttccatag cttcattcca ccaggccgag 180
cactcagacc tcgaaggcca gaaacagtac cagctccgag gtatggctac cttggaagggt 240
tatcgtgatac agaaagcggg atcgactcta attaaacacg ctgaagaaat tcttcgtaag 300
aggggggcgg acatgctttg gtgcaatgcg cggacaaccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agagatattt gatacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggctcacata a                                     441

```

&lt;210&gt; 187

&lt;211&gt; 438

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 187

```

atgatagagg tgaaaccgat taacgcagag gatacctatg aactaaggca taaaatactc 60
agaccaaacc agccgataga agcgtgtatg tatgaaagcg atttacttcg tgggtgcattt 120
cacttaggcg gcttttacag gggcaaactg atttccatag cttcattcca ccaggccgag 180
cactcagacc tcgaaggcca gaaacagtac cagctccgag gtatggctac cttggaagggt 240
tatcgtgatac agaaagcggg atcgactcta attaaacacg ctgaacaact tcttcgtaag 300
aggggggcgg acatgctttg gtgcaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agaggtattt gaaacgccgc cagtaggacc tcacatcctg 420
atgtataaaa agatcaca                                     438

```

&lt;210&gt; 188

&lt;211&gt; 438

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 188

```

atgctagagg tgaaaccgat taacgcagag gatacctatg aactaaggca tagaatactc 60
agaccaaacc agccgataga agcgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120
cacttaggcg gcttttacag gggcaaactg atttccatag cttcattcca ccaggccgag 180
cactcagaac tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaagggt 240
tatcgtgatac agaaagcggg atcgagtcta attaaacacg ctgaacaact tcttcgtaag 300
aggggggcgg acttgctttg gtgcaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agaggtattt gatacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggatcaca                                     438

```

&lt;210&gt; 189

&lt;211&gt; 438

&lt;212&gt; DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 189

```
atgctagagg tgaaactgat taacgcagag gatacctatg aactaaggca tagaatactc 60
agaccaaacc agccgtaga agcgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120
cacttaggcg gcttttacag gggcaaactg atttccatag cttcattcca ccaggccgag 180
cactcagacc tcgaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tttcgtgata agaaagcggg atcgagtcta attaaacacg ctgaagaaat tcttcgtaag 300
agggggggcg aacttgctttg gtgtaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agaggtattt gatacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggatcaca 438
```

<210> 190

<211> 438

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 190

```
atgatagagg tgaaaccgat taacgcagag gatacctatg aactaaggca taaaatactc 60
agaccaaacc agccgataga agcgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120
cacttaggcg gcttttacag gggcaaactg atttccatag cttcattcca ccaggccgag 180
cactcagacc tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tategtgata agaaagcggg atcgagtcta attagacacg ctgaacaaat tcttcgtaag 300
agggggggcg aacttgctttg gtgcaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agagatattt gatacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggatcaca 438
```

<210> 191

<211> 438

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 191

```
atgatagagg tgaaaccgat taacgcagag gatacctatg aactaaggca taaaatactc 60
agaccaaacc agccgtaga agcgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120
cacttaggcg gcttttacag gggcaaactg atttccatag cttcattcca ccaggccgag 180
cactcaggcc tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tategtgata agaaagcggg atcgagtata attaaacacg ctgaagaaat tcttcgtaag 300
aagggggggc aacttgctttg gtgcaatgcg cggacgtccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agagatattt gacacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggatcaca 438
```

<210> 192

<211> 438

<212> DNA

<213> Artificial Sequence

<220>



## &lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 192

```

atgatagagg tgaaaccgat taacgcagag gatacctatg aactaaggca tagaatactc 60
agaccaaacc agccgataga agcgtgtatg tatgaaagcg atttacttcg tgggtgcattt 120
cacttaggcg gcttttacag gggcaaaactg atttccatag cttcattcca ccaggccgag 180
cactcagacc tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaagggt 240
tatcgtgagc agaaagcggg atcgactcta attagacacg ctgaacaact tcttcgtaag 300
aggggggcgg acttgctttg gtgcaatgcg cggacatccg cctcaggcta ctacaaaagg 360
ttaggcttca gcgagcaggg agaggtatgt gatacgcgcg cagtaggacc tcacatcctg 420
atgtataaaa ggctcaca                                     438

```

&lt;210&gt; 193

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 193

```

atgctagagg tgaaaccgat taacgcagag gatacctatg aactaaggca tagaatactc 60
agaccaaacc agccgataga agcgtgtatg tatgaaagcg atttacttcg tgggtgcattt 120
cacttaggcg gctattacag gggcaaaactg atttccatag cttcattcca ccaggccgag 180
cactcagaac tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaagggt 240
tatcgtgagc agaaagcggg atcgagtcta attaaacacg ctgaagaaat tcttcgtaag 300
aggggggcgg acttgctttg gtgtaatgcg cggacatccg cctcaggcta ctacaaaagg 360
ttaggcttca gcgagcaggg agagatattt gaaacgcgcg cagtaggacc tcacatcctg 420
atgtataaaa ggatcacata a                                     441

```

&lt;210&gt; 194

&lt;211&gt; 438

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 194

```

atgatagaag tgaaaccgat taacgcagag gagacctatg aactaaggca tagaatactc 60
agaccaaacc agccgataga agcgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120
cacttaggcg gcttttacag gggcaaaactg atttccatag cttcattcca ccaggccgag 180
cactcagaac tgaaggcca gaaacagtac cagctccgag gtatggctac cttggaagggt 240
tatcgtgagc agaaagcggg atcgagtcta attagacacg ctgaacaact tcttcgtaag 300
aagggggcga atatgctttg gtgtaatgcg cggacaaccg cctcaggcta ctacaaaagg 360
ttaggcttca gcgagcaggg agagatattt gatacgcgcg cagtaggacc tcacatcctg 420
atgtataaaa ggatcaca                                     438

```

&lt;210&gt; 195

&lt;211&gt; 438

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 195

427	86918	428	
atgctagagg	tgaaaccgat	taacgcagag	gatacctatg aactaaggca tagaatactc 60
agaccaaacc	agccgataga	agcgtgtatg	tatgaaagcg atttacttcg tgggtgcactt 120
cacttaggcg	gcttttacag	gggcaaactg	atttccatag cttcattcca ccaggccgag 180
cactcagacc	tccaaggcca	gaaacagtac	cagctccgag gtatggctac cttggaagggt 240
tttcgtgatc	agaaagcggg	atcgagtcta	attagacacg ctgaacaaat tcttcgtaag 300
agggggggcg	acttgctttg	gtgcaatgcg	cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca	gcgagcaggg	aaaggatatt	gatacgcgcg cagtaggacc tcacatcctg 420
atgtataaaa	ggatcaca		438

<210> 196

<211> 438

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 196

atgctagagg	tgaaaccgat	taacgcagag	gatacctatg	aactaaggca	taaaatactc	60
agaccaaacc	agccgttaga	agtgtgtatg	tatgaaaccg	atttacttcg	tgggtgcattt	120
cacttaggcg	gcttttacag	gggcaaactg	atttccatag	cttcattcca	ccaggccgag	180
cactcagacc	tccaaggcca	gaaacagtac	cagctccgag	gtatggctac	cttgggaagggt	240
tatcgtgatc	agaaagcggg	atcgagtcta	attagacacg	ctgaacaaat	tcttcgtaag	300
agggggggcg	acatgctttg	gtgcaatgcg	cggacatccg	cctcaggcta	ctacaaaaag	360
ttaggcttca	gcgagcaggg	agaggatatt	gaaacgcgcg	cagtaggacc	tcacatcctg	420
atgtataaaa	ggatcaca					438

<210> 197

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 197

atgattgaag	tcaaaccaat	aaacgcggaa	gatacgtatg	agatcaggca	ccgcattctc	60
cggccgaatc	agccgctgga	agcatgcaag	tatgaaaccg	atttgcctcg	gggcacgctt	120
cacctcggtg	gatattaccg	gggcaagctg	gtcagcatcg	cttcctttca	tcaagccgaa	180
catccagagc	ttgaaggcca	aaaacagtat	cagctgagag	ggatggcgac	gcttgaagggt	240
taccgtgagc	aaaaagcggg	cagtacgctt	atccgccatg	ccgaagagct	tcttcggaaa	300
aaggggggcag	accttttatg	gtgcaacgcc	aggacatctg	cgagcgggta	ctataaaaag	360
ctcggttca	gcgaacaggg	cgaagtctac	gacataccgc	cggtcggacc	tcataatttg	420
atgtataaga	aattgacgta	a				441

<210> 198

<211> 438

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 198

atgatagagg	tgaaaccgat	taacgcagag	gatacctatg	aactaaggca	tagaatactc	60
agaccaaacc	agccgataga	agcgtgtatg	tatgaaagcg	atttacttcg	tgggtgcattt	120
cacttaggcg	gotattacag	gggcaaactg	atttccatag	cttcattcca	ccaggccgag	180

429

86918

430

```

cactcagaac tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tttcgtgagc agaaagcggg atcgactcta attagacacg ctgaacaaat tcttcgtaag 300
agggggggcgg acatgctttg gtgcaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agagatattt gatacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggatcaca                                     438

```

&lt;210&gt; 199

&lt;211&gt; 438

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 199

```

atgatagagg tgaaaccgat taacgcagag gatacctatg aactaaggca taaaatactc 60
agaccaaacc agccgataga agcgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120
cacttaggcg gcttttacgg gggcaaaactg atttccatag cttcattcca ccaggccgag 180
cactcagacc tcgaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tatcgtgagc agaaagcggg atcgactcta attagacacg ctgaagaaat tcttcgtaag 300
aagggggcga acttgctttg gtgcaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agaggatattt gacacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggctcaca                                     438

```

&lt;210&gt; 200

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;221&gt; misc\_feature

&lt;222&gt; 134, 313

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 200

```

atgatagagg tgaaaccgat taacgcagag gatacctatg aactaaggca taaaatactc 60
agaccaaacc agccgttaga agtgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120
cacttaggcg gctnttacag gggcaaaactg atttccatag cttcattcca ccaggccgag 180
cactcagaac tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tatcgtgatc agaaagcggg atcgagtcta attaaacacg ctgaacaaat tcttcgtaag 300
agggggggcgg acntgctttg gtgcaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agagatattt gatacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggctcacata a                                     441

```

&lt;210&gt; 201

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 201

```

atgatagagg tgaaaccgat taacgcagag gatacctatg aactaaggca tagaatactc 60
agaccaaacc agccgataga agtgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120

```

431	86918	432	
cacttaggcg	gcttttacag	gggcaaactg	atttccatag cttcattcca ccaggccgag 180
cactcagacc	tccaaggcca	gaaacagtac	cagctccgag gtatggctac cttggaaggt 240
tatcgtgagc	agaaagcggg	atcgactcta	attagacacg ctgaacaaat tcttcgtaag 300
aggggggcgg	acttgctttg	gtgcaatgcg	oggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca	gcgagcaggg	agaggtattt	gatacgccgc cagtaggacc tcacatcctg 420
atgtataaaa	ggatcacata	a	441

<210> 202

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 202

atgctagagg	tgaaaccgat	taacgcagag	gatacctatg	aactaaggca	tagaatactc	60
agaccaaaacc	agccgataga	agcgtgtatg	tttgaaagcg	atttacttcg	tggtgcattt	120
cacttaggcg	gcttttacag	gggcaaactg	atttccatag	cttcattcca	ccaggccgag	180
cactcagaac	tccaaggcca	gaaacagtac	cagctccgag	gtatggctac	cttgaaggt	240
tatcgtgagc	agaaagcggg	atcgagtcta	attagacacg	ctgaagaaat	tcttcgtaag	300
aggggggcgg	acatgctttg	gtgtaatgcg	oggacatccg	cctcaggcta	ctacaaaaag	360
ttaggcttca	gcgagcaggg	agagatattt	gaaacgccgc	cagtaggacc	tcacatcctg	420
atgtataaaa	ggctcacata	a				441

<210> 203

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 203

atgatagagg	tgaaaccgat	taacgcagag	gatacctatg	aactaaggca	tagaatactc	60
agaccaaaacc	agccgataga	agcgtgtatg	tatgaaaccg	atttacttcg	tggtgcattt	120
cacttaggcg	gcttttacag	gggcaaactg	atttccatag	cttcattcca	ccaggccgag	180
cactcagaac	tccaaggcca	gaaacagtac	cagctccgag	gtatggctac	cttgaaggt	240
tttcgtgagc	agaaagcggg	atcgagtcta	attagacacg	ctgaacaaat	tcttcgtaag	300
aggggggcgg	acttgctttg	gtgtaatgcg	oggacatccg	cctcaggcta	ctacaaaaag	360
ttaggcttca	gcgagcaggg	agagatattt	gatacgccgc	cagtaggacc	tcacatcctg	420
atgtataaaa	ggctcacata	a				441

<210> 204

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 204

atgatagagg	tgaaaccgat	taacgcagag	gatacctatg	aactaaggca	tagaatactc	60
agaccaaaacc	agccgataga	agcgtgtatg	tttgaaagcg	atttacttcg	tggtgcattt	120
cacttaggcg	gctattacag	gggcaaactg	atttccatag	cttcattcca	ccaggccgag	180
cactcagacc	tccaaggcca	gaaacagtac	cagctccgag	gtatggctac	cttgaaggt	240
taccgcgatac	agaaagcggg	atcgagtcta	attagacacg	ctgaacaaat	tcttcgtaag	300

433	86918	434	
agggggggcgg	acttgctttg	gtgtaatgcg	cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca	gcgagcaggg	agagatat	ttt gatacgccgc cagtaggacc tcacatcctg 420
atgtataaaa	ggctcacata	a	441
<210> 205			
<211> 441			
<212> DNA			
<213> Artificial Sequence			
<220>			
<223> Synthetic DNA Sequence			
<400> 205			
atgctagagg	tgaaaccgat	taacgcagag	gatacctatg aactaaggca taaaatactc 60
agaccaaacc	agccgttaga	agtgtgtatg	tatgaaaccg atttacttcg tgggtgcattt 120
cacttaggcg	gctttttacag	gggcaaactg	atttccatag cttcattcca ccaggccgag 180
cactcagaac	tccaaggcca	gaaacagtac	cagctccgag gtatggctac cttggaagggt 240
tatcgtgagc	agaaagcggg	atcgagtcta	attaaacacg ctgaagaaat tcttcgtaag 300
agggggggcgg	acttgctttg	gtgcaatgcg	cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca	gcgagcaggg	agaggtat	ttt gatacgccgc cagtaggacc tcacatcctg 420
atgtataaaa	ggatcacata	a	441
<210> 206			
<211> 441			
<212> DNA			
<213> Artificial Sequence			
<220>			
<223> Synthetic DNA Sequence			
<400> 206			
atgatagaag	tgaaaccgat	taacgcagag	gatacctatg aactgaggca taaaatactc 60
agaccaaacc	agccgataga	agtgtgtatg	tatgaaagcg atttacttcg tgggtgcattt 120
cacttaggcg	gctttttacag	gggcaaactg	atttccatag cgtcattcca ccaggccgag 180
cacccagacc	tccaaggcca	gaaacagtac	cagctccgag gtatggctac cttggaagggt 240
tatcgtgata	agaaagcggg	atcgagtcta	attaaacacg ctgaacaaat tcttcgtaag 300
agggggggcgg	acttgctttg	gtgcaatgcg	cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca	gcgagcaggg	agaggtat	ttt gaaacgcgc cagtaggacc tcacatcctg 420
atgtataaaa	ggctcacata	a	441
<210> 207			
<211> 441			
<212> DNA			
<213> Artificial Sequence			
<220>			
<223> Synthetic DNA Sequence			
<400> 207			
atgctagagg	tgaaaccgat	taacgcagag	gatacctatg aactgaggca tagaatactc 60
agaccaaacc	agccgataga	agcgtgtatg	ttt gaaaccg atttacttcg tgggtgcattt 120
cacttaggcg	gctttttacag	gggcaaactg	atttccatag cttcattcca ccaggccgag 180
cactcagacc	tccaaggcca	gaaacagtac	caactccgag gtatggctac cttggaagggt 240
tttcgtgagc	agaaagcggg	atcgactcta	attagacacg ctgaagaaat tcttcgtaag 300
agggggggcgg	acttgctttg	gtgcaatgcg	cggacatccg cctcaggcta ctacaaaagg 360
ttaggcttca	gcgagcaggg	agagatat	ttt gacacgcgc cagtagggcc tcacatcctg 420
atgtataaaa	ggctcacata	a	441

&lt;210&gt; 208

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 208

```

atgatagagg tgaaaccgat taacgcagag gatacctatg aactaaggca tagaatactc 60
agaccaaacc agccgataga agcgtgtatg ttgaaagcg atttacttcg tgggtgcattt 120
cacttaggcg gctattacag gggcaaactg atttccatag cgtcattcca ccaggccgag 180
cactcggaac tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tatcgtgac agaaagcggg atcgagtcta attagacacg ctgaacaaat tcttcgtaag 300
aggggggcgg acatgctttg gtgcaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agaggtattt gatacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggatcacata a 441

```

&lt;210&gt; 209

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 209

```

atgatagagg tgaaaccgat taacgcagag gatacctatg aactaaggca tagaatactc 60
agaccaaacc agccgtaga agtgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120
cacttgggcg gcttttacgg gggcaaactg atttccatag cgtcattcca ccaggccgag 180
caccagacc tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tatcgtgac agaaagcggg atcgagtcta attagacacg ctgaacaact tcttcgtaag 300
aggggggcgg acttgctttg gtgcaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agagatattc gaaacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggatcacata a 441

```

&lt;210&gt; 210

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 210

```

atgatagaag tgaaaccgat taacgcagag gatacctatg aactaaggca taaaatactc 60
agaccaaacc agccgataga agcgtgtatg tatgaaagcg atttacttcg tgggtgcattt 120
cacttaggcg gctattacag gggcaaactg atttccatag cgtcattcca ccaggccgag 180
cactcagaac tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tatcgtgagc agaaagcggg atcgagtcta attaaacacg ctgaagaaat tcttcgtaag 300
aggggggcgg acttgctttg gtgcaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agaggtattt gacacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggatcacata a 441

```

&lt;210&gt; 211

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 211

```

atgctagagg tgaaaccgat taacgcagag gatacctatg aactaaggca taaaatactc 60
agaccaaacc agccgataga agcgtgtatg tttgaaagcg atttacttcg tgggtgcattt 120
cacttaggcg gcttttacgg gggcaaactg atttccatag cttcattcca ccaggccgag 180
cactcagacc tcgaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tategtgagc agaaagcggg atcgactcta attagacacg ctgaagaaat tcttcgtaag 300
agggggggcg acttgctttg gtgcaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agagatatatt gatacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggatcacata a 441

```

&lt;210&gt; 212

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 212

```

atgatagagg tgaaaccgat taacgcagag gatacctatg aactaaggca tagaatactc 60
agaccaaacc agccgataga agtgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120
cacttaggcg gcttttacag gggcaaactg atttccatag cttcattcca ccaggccgag 180
cactcagacc tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tategtgagc agaaagcggg atcgagtcta attaaacacg ctgaacaaat tcttcgtaag 300
agggggggcg acttgctttg gtgcaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agaggtatatt gatacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggctcacata a 441

```

&lt;210&gt; 213

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 213

```

atgctagagg tgaaaccgat taacgcagag gatacctatg aactaaggca tagaatactc 60
agaccaaacc agccgataga agtgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120
cacttaggcg gcttttacgg gggcaaactg atttccatag cttcattcca ccaggccgag 180
cactcagaac tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tategtgatc agaaagcggg atcgagtcta attaaacacg ctgaagaaat tcttcgtaag 300
agggggggcg acttgctttg gtgtaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agaggtatatt gaaacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggctcacata a 441

```

&lt;210&gt; 214

&lt;211&gt; 438

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 214

```

atgatagaag tgaaacctat taacgcagag gagacttacg aacttcgaca caagatcctg 60
cgccctaatac agccgataga ggcattgcattg tatgaaagcg atctgctgcg gggctcgttc 120
catttggggcg ggttctatcg tggccaattg atctcgattg cgagtttcca caaagctgaa 180
cactcagaac tgcaagggca aaagcagtat caattacgag ggatggcgac cctcgaagga 240
ttccgtgagc agaaggctgg ctcttcgctt attaggcacg ccgaggagat actacggaat 300
aaaggggagc atctgctttg gtgtaatgca cgcacgacag cctccgggta ctataaaaagg 360
cttggtttta gtgagcagcg cgaagttttc gaaaccccg cggttggggc gcacattctt 420
atgtacaaaa gaactcact                                     438

```

&lt;210&gt; 215

&lt;211&gt; 438

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 215

```

atgatagaag tgaaacctat taacgcagag gatacttacg aacttcgaca caggatcctg 60
cgccctaatac agccgttaga ggcattgcattg tatgaaagcg atctgctgcg gggctcgttc 120
catttggggcg ggttctatcg tggcaaattg atctcgattg cgagtttcca ccaagctgaa 180
cactcagaac tggaagggca aaagcagtat caattacgag ggatggcgac cctcgaagga 240
ttccgtgagc agaaggctgg ctcttcgctt attaggcacg ccgaggagat actacggaaa 300
agaggggagc atctgctttg gtgtaatgca cgcacgacag ccgccgggta ctataaaaagg 360
cttggtttta gtgagcaggg cgaatttttc gacaccccg cggttggggc gcacattctt 420
atgtacaaaa gaactcact                                     438

```

&lt;210&gt; 216

&lt;211&gt; 438

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 216

```

atgatagaag tgaaacctat taacgcagag gatacttacg aaattcgaca caggatcctg 60
cgccctaatac agccgttaga ggcattgcattg tatgaaagcg atctgctgcg gggctcgttc 120
catttggggcg ggttctatcg tggcaaattg atctcgattg cgagtttcca ccaagctgaa 180
cactcagacc tggaagggca aaagcagtat caattacgag ggatggcgac cctcgaagga 240
taccgtgatac agaaggctgg ctcttcgctt attaggcacg ccgagcagat actacggaaa 300
agaggggagc atctgctttg gtgcaatgca cgcacgacag ccgccgggta ctataaaaagg 360
cttggtttta gtgagcaggg cgaagttttc gacaccccg cggttggggc gcacattctt 420
atgtacaaaa aactcact                                     438

```

&lt;210&gt; 217

&lt;211&gt; 438

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence



441

86918

442

&lt;400&gt; 217

```

atgatagaag tgaaacctat taacgcagag gatacttacg aacttcgaca caggatcctg 60
cgccctaatac agccgataga ggcattgcag tatgaaagcg atctgctgcg gggctcgttc 120
catttgggcg ggttctatcg tggcaaattg atctcgattg cgagtttcca ccaagctgaa 180
cagtcagaac tggaaaggca aaagcagtat caattacgag ggatggcgac cctcgaagga 240
taccgtgac agaaaggctgg ctctacgctt attaagcacg ccgaggagat actacggaaa 300
aaaggggcag atctgctttg gtgcaatgca cgcacgtcag ccgccgggta ctataaaaag 360
cttggtttta gtgagcaggg cgaaattttc gacaccccg cggttggggc gcacattctt 420
atgtacaaaa gactcact                                     438

```

&lt;210&gt; 218

&lt;211&gt; 438

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 218

```

atgatagaag tgaaacctat taacgcagag gagacttacg aacttcgaca caagatcctg 60
cgccctaatac agccgtaga ggcattgcag tatgaaacgg atctgctgcg gggctcgttc 120
catttgggcg ggttctatcg tggcaaattg atctcgattg cgagtttcca ccaagctgaa 180
cactcagaac tggaaaggca aaagcagtat caattacgag ggatggcgac cctcgaagga 240
ttccgtgac agaaaggctgg ctcttcgctt attaagcacg ccgaggagat actacggaaa 300
agaggggcag atctgctttg gtgcaatgca cgcacgtcag cctccgggta ctataaaaag 360
cttggtttta gtgagcaggg cgaaattttc gaaaccccg cggttggggc gcacattctt 420
atgtacaaaa gactcact                                     438

```

&lt;210&gt; 219

&lt;211&gt; 438

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 219

```

atgctagaag tgaaacctat taacgcagag gagacttacg aacttcgaca caagatcctg 60
cgccctaatac agccgataga ggcattgcag tatgaaacgg atctgctgcg gggctcgttc 120
catttgggcg ggttctatcg tggccaattg atctcgattg cgagtttcca ccaagctgaa 180
cactcagacc tgcaagggca aaagcagtat caattacgag ggatggcgac cctcgaagga 240
taccgtgagc agaaaggctgg ctctacgctt attaagcacg ccgaggagct actacggaaa 300
aaaggggcag atctgctttg gtgcaatgca cgcacgacag ccgccgggta ctataaaaag 360
cttggtttta gtgagcaggg cgaagttttc gacaccccg cggttggggc gcacattctt 420
atgtacaaaa aaatcact                                     438

```

&lt;210&gt; 220

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 220

```

atgatagaag tgaaacctat taacgcagag gatacttacg aacttcgaca caagatcctg 60
cgccctaatac agccgtaga ggcattgcag tatgaaagcg atctgctgcg gagcgcatc 120

```

443

86918

444

```

catttgggCG ggttctatcg tggcaaattg atctcgattg cgagtttcca caaagctgaa 180
cactcagaac tgcaagggca aaagcagtat caattacgag ggatggcgac cctcgaagga 240
taccgtgac agaaggctgg ctcttcgctt attaggcacg ccgaggagat actacggaaa 300
agaggggCag atatgctttg gtgcaatgca cgcacgtcag ccgccggtta ctataaaagg 360
cttggtttta gtgagcaggg cgaagttttc gacaccccgC cggttgggC gcacattctt 420
atgtacaaaa gaatcactta a 441

```

&lt;210&gt; 221

&lt;211&gt; 438

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 221

```

atgatagaag tgaaacctat taacgcagag gatacttacg aacttcgaca caggatcctg 60
cgccctaata agccgtaga ggcattgcatt tatgaaagcg atctgctgcg gggctcgttc 120
catttgggCG ggttctatcg tggcaaattg atctcgattg cgagtttcca ccaagctgaa 180
cactcagacc tgcaagggca aaagcagtat caattacgag ggatggcgac cctcgaagga 240
taccgtgac agaaggctgg ctcttcgctt attaagcacg ccgagcagct actacggaaa 300
aaaggggCag atatgctttg gtgtaatgca cgcacgtcag ccgccggtta ctataaaagg 360
cttggtttta gtgagcaggg cgaaattttc gaaaccccgC cggttgggC gcacattctt 420
atgtacaaaa gaatcact 438

```

&lt;210&gt; 222

&lt;211&gt; 438

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 222

```

atgctagaag tgaaacctat taacgcagag gatacttacg aacttcgaca caggatcctg 60
cgccctaata agccgtaga ggcattgcatt tatgaaacCG atctgctgcg gggctcgttc 120
catttgggCG ggttctatcg tggcaaattg atctcgattg cgagtttcca ccaagctgaa 180
cattcagaac tggaagggca aaagcagtat caattacgag ggatggcgac tctcgaagga 240
taccgtgac agaaggctgg ctcttcgctt attaggcacg ccgaggagat actacggaaa 300
agaggggCag atatgctttg gtgcaatgca cgcacgacag ccgccggtta ctataaaag 360
cttggtttta gtgagcaggg cgaaattttac gacaccccgC cggttgggC gcacattctt 420
atgtacaaaa aactcact 438

```

&lt;210&gt; 223

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 223

```

atgatagaag tgaaacctat taacgcagag gagacttacg aacttcgaca caagatcctg 60
cgccctaata agccgtaga ggcattgcatt tatgaaacCG atctgctgcg gggcgcgttc 120
catttgggCG ggttctatcg tggcaaattg atctcgattg cgagtttcca ccaagctgac 180
cactcagaac tgcaagggca aaagcagtat caattacgag ggatggcgac cctcgaagga 240
taccgtgagc agaaggctgg ctctacgctt attaggcacg ccgagcagat actacggaaa 300

```

445	86918	446	
agaggggagcag	atctactttg	gtgcaatgca	cgcacgtcag cgcgcgggta ctataaaaag 360
cttggtttta	gtgagcacgg	cgaattttc	gaaaccccg cggttgggccc gcacattctt 420
atgtacaaaa	gactcactta	a	441

<210> 224

<211> 438

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 224

atgatagaag	tgaaacctat	taacgcagag	gatacttacg	aacttcgaca	caggatcctg	60
cgccctaatac	agccgataga	ggcatgcatg	tatgaaagcg	atctgctgcg	gggcgcgttc	120
catttgggcg	ggttctatcg	tggcaaattg	atctcgattg	cgagtttcca	ccaagctgaa	180
cactcagacc	tgcaagggca	aaagcagtat	caattacgag	ggatggcgac	cctcgaagga	240
taccgtgagc	agaaggctgg	ctcttcgctt	attaggcacg	ccgaggagat	actacggaaa	300
aaaggggagc	atatgctttg	gtgcaatgca	cgcacgacag	ccgcgcggta	ctataaaaag	360
cttggtttta	gtgagcagg	cgaagtttc	gacaccccg	cggttgggccc	gcacattctt	420
atgtacaaaa	gaatcact					438

<210> 225

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 225

atgatagagg	tgaaaccgat	taacgcagag	gatacctatg	aactaaggca	tagaatactc	60
agaccaaacc	agccgataga	agtgtgtatg	tatgaaaccg	atttacttcg	tgggtgcattt	120
cacttaggcg	gcttttacag	gggcaaactg	atttccatag	cttcattcca	ccaggccgag	180
cactcagacc	tccaaggcca	gaaacagtac	cagctccgag	gtatggctac	cttggagggt	240
tatcgtgagc	agaaagcggg	atcgagtcta	attaaacacg	ctgaacaaat	tcttcgtaag	300
agggggggcg	acttgctttg	gtgcaatgcg	cggacatccg	cctcaggcta	ctacaaaaag	360
ttaggcttca	gcgagcagg	agaggatatt	gatacgccgc	cagtaggacc	tcacatcctg	420
atgtataaaa	ggctcacata	a				441

<210> 226

<211> 438

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 226

atgatagaag	tgaaacctat	taacgcagag	gagacttacg	aacttcgaca	gaggatcctg	60
cgccctaatac	agccgataga	ggcatgcatg	tatgaaagcg	atctgctgcg	gggctcggttc	120
catttgggcg	ggttctatcg	tggccaattg	atctcgattg	cgagtttcca	ccaagctgaa	180
cactcagaac	tgcaagggca	aaagcagtat	caattacgag	ggatggcgac	cctcgaagga	240
taccgtgagc	agaaggctgg	ctctacgctt	attaagcacg	ccgaggagat	actacggaaa	300
aaaggggagc	atctgctttg	gtgcaatgca	cgcacgtcag	ccgcgcggta	ctataaaaag	360
cttggtttta	gtgagcacgg	cgaatttttc	gacaccccg	cggttgggccc	gcacattctt	420
atgtacaaaa	gaatcact					438

&lt;210&gt; 227

&lt;211&gt; 438

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 227

```

atgatagaag tgaaacctat taacgcagag gatacttacg aacttcgaca caggatcctg 60
cgccctaata agccgataga ggcatgcatg tatgaaaccg atctgctgcg gggcgcggttc 120
catttgggcg ggtactatcg tggccaattg atctcgattg cgagtttcca caaagctgaa 180
cactcagaac tgcaagggca aaagcagtat caattacgag ggatggcgac cctcgaagga 240
taccgtgagc agaaggctgg ctctacgctt attaagcacg ccgagcagct actacgggaa 300
aaaggggcag atatgctttg gtgcaatgca cgcacgtcag ccgccggtta ctataaaaagg 360
cttggtttta gtgagcaggg cgaagttttc gacaccccg cggttggggc gcacattctt 420
atgtacaaaa aactcact                                     438

```

&lt;210&gt; 228

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 228

```

atgatagaag tgaaacctat taacgcagag gatacttacg aacttcgaca caagatcctg 60
cgccctaata agccgataga ggcatgcatg tatgaaaccg atctgctgcg gggctcggttc 120
catttgggcg ggttctatcg tggcaaattg atctcgattg cgagtttcca caaagctgaa 180
cactcagacc tggaagggca aaaccagtat caattacgag ggatggcgac cctcgaagga 240
taccgtgagc agaaggctgg ctctacgctt attaggcacg ccgaggagat actacggaaa 300
agaggggcag atatgctttg gtgcaatgca cgcacgtcag cctccggtta ctataaaaagg 360
cttggtttta gtgagcacgg cgaatttttc gacaccccg cggttggggc gcacattctt 420
atgtacaaaa gactcactta a                                     441

```

&lt;210&gt; 229

&lt;211&gt; 438

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 229

```

atgctagaag tgaaacctat taacgcagag gatacttacg aacttcgaca caggatcctg 60
cgccctaata agccgataga ggcatgcatg tatgaaaccg atctgctgcg gggctcggttc 120
catttgggcg ggttctatcg tggccaattg atctcgattg cgagtttcca caaagctgaa 180
cactcagacc tggaagggca aaagcagtat caattacgag ggatggcgac cctcgaagga 240
taccgtgagc agaaggctgg ctctacgctt attaggcacg ccgagcagat actacggaaa 300
agaggggcag atatgctctg gtgcaatgca cgcacgtcag ccgccggtta ctataaaaagg 360
cttggtttta gtgagcaggg cgaagttttc gaaaccccg cggttggggc gcacattctt 420
atgtacaaaa gactcact                                     438

```

&lt;210&gt; 230

&lt;211&gt; 438

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 230

```

atgatagaag tgaaacctat taacgcagag gatacttacg aacttcgaca caggatcctg 60
cgccctaata agccgttaga ggcatgcatg tatgaaaccg atctgctgcg gggctcgttc 120
catttgggcg ggttctatcg tggcaaattg atctcgattg cgagtttcca ccaagctgaa 180
cactcagacc tgcaagggca aaagcagtat caattacgag ggatggcgac cctcgaagga 240
taccgtgagc agaaggctgg ctctacgctt attaggcacg ccgagcagct actacggaaa 300
agagggggcag atctgctttg gtgcaatgca cgcacgtcag cctccgggta ctataaaaagg 360
cttggtttta gtgagcacgg cgaagttttc gacaccccg cggttggggc gcacattctt 420
atgtacaaaa gactcact                                     438

```

&lt;210&gt; 231

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 231

```

atgctagaag tgaaacctat taacgcagag gagacttacg aacttcgaca caagatcctg 60
cgccctaata agccgttaga ggcatgcatg tatgaaagcg atctgctgcg gggctcgttc 120
catttgggcg ggtactatcg tggccaattg atctcgattg cgagtttcca ccaagctgaa 180
cactcagaac tggaagggca aaagcagtat caattacgag ggatggcgac cctcgaagga 240
ttccgtgagc agaaggctgg ctctacgctt attaagcacg ccgagcagat actacggaaa 300
agagggggcag atatgctttg gtgcaatgca cgcacgtcag ccgccgggta ctataaaaag 360
cttggtttta gtgagcacgg cgaaattttc gacaccccg cggttggggc gcacattctt 420
atgtacaaaa aactcactta a                                     441

```

&lt;210&gt; 232

&lt;211&gt; 438

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 232

```

atgatagaag tgaaacctat taacgcagag gagacttacg aacttcgaca caggatcctg 60
cgccctaata agccgataga ggcatgcatg tatgaaagcg atctgctgcg gggctcgttc 120
catttgggcg ggttctatcg tggcaaattg atctcgattg cgagtttcca ccaagctgaa 180
cactcagacc tagaagggca aaagcagtat caattacgag ggatggcgac cctcgaagga 240
taccgtgatc agaaggctgg ctctacgctt attaagcacg ccgaggagct actacggaaa 300
agagggggcag atatgctttg gtgcaatgca cgcacgtcag ccgccgggta ctataaaaag 360
cttggtttta gtgagcacgg cgaaattttac gaaaccccg cggttggggc gcacattctt 420
atgtacaaaa aaatcact                                     438

```

&lt;210&gt; 233

&lt;211&gt; 438

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 233

```

atgatagaag tgaaacctat taacgcagag gatacttacg aacttcgaca caagatcctg 60
cgccctaatac agccgataga ggcatgcatg tatgaaagcg atctgctgcg gggctcgttc 120
catttgggcg ggttctatcg tggccaattg atctcgattg cgagtttcca ccaagctgaa 180
cactcagacc tggaaaggca aaagcagtat caattacgag ggatggcgac cctcgaagga 240
taccgtgatac agaaggctgg ctcttcgctt attaaagcac ccgaggagat actacggaaa 300
agaggggagc atctgctttg gtgcaatgca cgcacgtcag ccgccgggta ctataaaagg 360
cttggtttta gtgagcaggg cgaaattttc gacaccccg cggttggggc gcacattctt 420
atgtacaaaa aaatcact

```

438

&lt;210&gt; 234

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 234

```

atgctagagg tgaaaccgat taacgcagag gatacctatg aactaaggca tagaatactc 60
agaccaaacc agccgataga agcgtgtatg tatgaaagcg atttacttcg tgggtgcattt 120
cacttaggcg gctattacag gggcaaactg atttccatag cttcattcca ccaggccgag 180
cactcagaac tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tatcgtgagc agaaagcggg atcgagtcta gttaaacacg ctgaagaaat tcttcgtaag 300
aggggggcgg acttgctttg gtgtaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agagatatatt gaaacgcgcg cagtaggacc tcacatcctg 420
atgtataaaa ggatcacata a

```

441

&lt;210&gt; 235

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 235

```

atgctagagg tgaaaccgat taacgcagag gatacctatg aactaaggca tagaatactc 60
agaccaaacc agccgataga agcgtgtatg tatgaaagcg atttacttcg tgggtgcattt 120
cacttaggcg gcttttacag gggcaaactg atttccatag cttcattcca ccaggccgag 180
cactcagaac tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tatcgtgagc agaaagcggg atcgagtcta attaaacacg ctgaagaaat tcttcgtaag 300
aggggggcgg acttgctttg gtgtaatgcg cggacgtccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agagatatatt gaaacgcgcg cagtaggacc tcacatcctg 420
atgtataaaa ggctcacata a

```

441

&lt;210&gt; 236

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 236

```

atgctagagg tgaaaccgat taacgcagag gatacctatg aactaaggca taaaatactc 60
agaccaaacc agcgttaga agtgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120
cacttaggcg gctattacag gggcaaactg atttccatag cttcattcca ccaggccgag 180
cactcagaac tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tategtgagc agaaagcggg atcgagtcta attaaacacg ctgaagaaat tcttcgtaag 300
agggggggcg acttgctttg gtgtaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agagatattt gaaacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggatcacata a                                     441

```

&lt;210&gt; 237

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 237

```

atgatagaag tgaaaccgat taacgcagag gatacctatg aactaaggca tagaatactc 60
agaccaaacc agccgataga agcgtgtatg tatgaaagcg atttacttcg tgggtgcattt 120
cacttaggcg gctattacag gggcaaactg atttccatag cttcattcca ccaggccgag 180
cactcagaac tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tategtgagc agaaagcggg atcgagtcta attaaacacg ctgaagaaat tcttcgtaag 300
agggggggcg acttgctttg gtgtaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agagatattt gaaacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggatcacata a                                     441

```

&lt;210&gt; 238

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 238

```

atgctagagg tgaaaccgat taacgcagag gatacctatg aactaaggca tagaatactc 60
agaccaaacc agccgataga agcgtgtatg tatgaaagcg atttacttcg tgggtgcattt 120
cacttaggcg gctattacag gggcaaactg atttccatag cttcattcca ccaggccgag 180
caccagaac tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tategtgagc agaaagcggg atcgagtcta attaaacacg ctgaagaaat tcttcgtaag 300
agggggggcg acttgctttg gtgtaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agaggtattt gaaacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggatcacata a                                     441

```

&lt;210&gt; 239

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 239

```

atgctagagg tgaaaccgat taacgcagag gatacctatg aactaaggca tagaatactc 60
agaccaaacc agccgataga agcgtgtatg tatgaaagcg atttacttcg tgggtgcattt 120

```

455	86918	456	
cacttaggcg	gctattacgg	gggcaaactg	atttccatag cttcattcca ccaggccgag 180
cactcagacc	tccaaggcca	gaaacagtac	cagctccgag gtatggctac cttggaaggt 240
tatcgtgagc	agaaagcggg	atcgagtcta	attaaacacg ctgaagaaat tcttcgtaag 300
aggggggcgg	acttgctttg	gtgtaatgcg	cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca	gcgagcaggg	agagatatatt	gaaacgccgc cagtaggacc tcacatcctg 420
atgtataaaa	ggatcacata	a	441

<210> 240

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 240

atgctagagg	tgaaaccgat	taacgcagag	gatacctatg	aactaaggca	tagaatactc	60
agaccaaac	agccgataga	agtgtgtatg	tatgaaagcg	atttacttcg	tggtgcattt	120
cacttaggcg	gctattacag	gggcaaactg	atttccatag	cttcattcca	ccaggccgag	180
cactcagaac	tccaaggcca	gaaacagtac	cagctccgag	gtatggctac	cttgaaggt	240
tatcgtgagc	agaaagcggg	atcgagtcta	attaaacacg	ctgaagaaat	tcttcgtaag	300
aggggggcgg	acttgctttg	gtgtaatgcg	cggacatccg	cctcaggcta	ctacaaaaag	360
ttaggcttca	gcgagcaggg	agagatatatt	gaaacgccgc	cagtaggacc	tcacatcctg	420
atgtataaaa	ggatcacata	a				441

<210> 241

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 241

atgatagaag	tgaaacctat	taacgcagag	gatacctatg	aactaaggca	tagaatactc	60
agaccaaac	agccgataga	agtgtgtatg	tatgaaaccg	atttacttcg	tggtgcattt	120
cacttaggcg	gcttttacgg	gggcaaactg	atttccatag	cttcattcca	ccaggccgag	180
cactcagacc	tccaaggcca	gaaacagtac	cagctccgag	gtatggctac	cttgaaggt	240
tatcgtgagc	agaaagcggg	atcgagtcta	attagacacg	ctgaacaaat	tcttcgtaag	300
aggggggcgg	acatgctttg	gtgcaatgcg	cggacatccg	cctcaggcta	ctacaaaaag	360
ttaggcttca	gcgagcaggg	agagatatatt	gaaacgccgc	cagtaggacc	tcacatcctg	420
atgtataaaa	ggatcacata	a				441

<210> 242

<211> 441

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA Sequence

<400> 242

atgatagaag	tgaaacctat	taacgcagag	gatacctatg	aactaaggca	tagaatactc	60
agaccaaac	agccgttaga	agtgtgtatg	tatgaaaccg	atttacttcg	tggtgcattt	120
cacttaggcg	gcttttacgg	gggcaaactg	atttccatag	cttcattcca	ccaggccgag	180
cactcagacc	tccaaggcca	gaaacagtac	cagctccgag	gtatggctac	cttgaaggt	240
tatcgtgagc	agaaagcggg	atcgagtcta	attagacacg	ctgaacaaat	tcttcgtaag	300



457

86918

458

```

agggggggcgg acatgctttg gtgtaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agagatattt gaaacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggatcacata a 441

```

&lt;210&gt; 243

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;221&gt; misc\_feature

&lt;222&gt; 9, 76, 98

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 243

```

atgctagang tgaaaccgat taacgcagag gatacctatg aactaaggca taaaatactc 60
agaccaaacc agccgntaga agtgtgtatg tatgaaancg atttacttcg tgggtgcattt 120
cacttaggcg gcttttacag gggcaaaactg atttccatag cttcattcca ccaggccgag 180
cactcagacc tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tatcgtgatc agaaagcggg atcgagtcta attaaacacg ctgaacaaat tcttcgtgag 300
agggggggcgg acatgctttg gtgcaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agaggtattt gacacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggctcacata a 441

```

&lt;210&gt; 244

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 244

```

atgctagaag tgaaacctat taacgcagag gatacctatg aactaaggca taaaatactc 60
agaccaaacc agccgataga agtgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120
cacttaggcg gcttttacgg gggcaaaactg atttccatag cttcattcca ccaggccgag 180
cactcagacc tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tatcgtgatc agaaagcggg atcgagtcta attagacacg ctgaacaaat tcttcgtaag 300
agggggggcgg acatgctttg gtgcaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agagatattt gaaacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggctcacata a 441

```

&lt;210&gt; 245

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 245

```

atgctagagg tgaaaccgat taacgcagag gatacctatg aactaaggca taaaatactc 60
agaccaaacc agccgttaga agtgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120
cacttaggcg gcttttacag gggcaaaactg atttccatag cgtcattcca ccaggccgag 180
cactcagacc tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240

```

459

86918

460

```
tatcgtgatac agaaagcggg atcgagtcta attagacacg ctgaacaaat tcttcgtaag 300
agggggggcgg acatgctttg gtgcaatgcg cggacatccg cctcaggcta ctacagaaag 360
ttaggcttca gcgagcaggg agaggtatgt gaaacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggctcacata a 441
```

&lt;210&gt; 246

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 246

```
atgatagagg tgaaaccgat taacgcagag gatacctatg aactaaggca taaaatactc 60
agaccaaacc agccgttaga agtgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120
cacttagggc gcttttacag gggcaaactg atttccatag cttcattcca ccaggccgag 180
cactcagacc tccaaggcca gaaacagtag cagctccgag gtatggctac cttggaagggt 240
tatcgtgatac agaaagcggg atcgagtcta attagacacg ctgaacaaat tcttcgtaag 300
agggggggcgg acatgctttg gtgcaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg ggaggtatgt gaaacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggatcacata a 441
```

&lt;210&gt; 247

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 247

```
atgatagagg tgaaaccgat taacgcagag gatacctatg aactaaggca taaaatactc 60
agaccaaacc agccgttaga agtgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120
cacttagggc gcttttacag gggcaaactg atttccatag cttcattcca ccaggccgag 180
cactcagacc tccaaggcca gaaacagtag cagctccgag gtatggctac cttggaagggt 240
tatcgtgagc agaaagcggg atcgagtcta attagacacg ctgaacaaat tcttcgtaag 300
agggggggcgg acatgctttg gtgcaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agaggtatgt gaaacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggctcacata a 441
```

&lt;210&gt; 248

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 248

```
atgatagagg tgaaaccgat taacgcagag gatacctatg aactaaggca taaaatactc 60
agaccaaacc agccgttaga agtgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120
cacttagggc gcttttacag gggcaaactg atttccatag cttcattcca ccaggccgag 180
cactcagacc tccaaggcca gaaacagtag cagctccgag gtatggctac cttggaagggt 240
tatcgtgagc agaaagcggg atcgagtcta attagacacg ctgaacaaat tcttcgtaag 300
agggggggcgg acatgctttg gtgcaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agaggtatgt gaaacgccgc cagtaggacc tcacatcctg 420
```

461 86918 462

atgtataaaaa ggctcacata a 441

<210> 249  
 <211> 441  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic DNA Sequence

<400> 249  
 atgctagagg tgaaaccgat taacgcagag gatacctatg aactaaggca taaaatactc 60  
 agaccaaacc agccgttaga agtgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120  
 cacttaggcg gctttttacgg gggcaaaactg atttccatag cttcattcca ccaggccgag 180  
 cactcggacc tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240  
 tatcgtgac agaaagcggg atcgagtcta attagacacg ctgaagaaat tcttcgtaag 300  
 aggggggcgg acatgctttg gtgcaatgcg cggacatccg cctcaggcta ctacaaaaag 360  
 ttaggcttca gcgagcaggg agaggtattt gacacgccgc cagtaggacc tcacatcctg 420  
 atgtataaaaa agatcacata a 441

<210> 250  
 <211> 441  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic DNA Sequence

<400> 250  
 atgctagaag tgaaaccgat taacgcagag gatacctatg aactaaggca taaaatactc 60  
 agaccaaacc agccgttaga agtgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120  
 cacttaggcg gctttttacgg gggcaaaactg atttccatag cttcattcca ccaggccgag 180  
 cactcagacc tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240  
 tatcgtgac agaaagcggg atcgactcta attaaacacg ctgaacaaat tcttcgtaag 300  
 aggggggcgg acatgctttg gtgcaatgcg cggacatccg cctcaggcta ctacaaaaag 360  
 ttaggcttca gcgagcaggg agaggtattt gaaacgccgc cagtaggacc tcacatcctg 420  
 atgtataaaaa ggctcacata a 441

<210> 251  
 <211> 441  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic DNA Sequence

<400> 251  
 atgatagagg tgaaaccgat taacgcagag gatacctatg aactaaggca tagaatactc 60  
 agaccaaacc agccgttaga agtgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120  
 cacttaggcg gctttttacag gggcaaaactg atttccatag cttcattcca ccaggccgag 180  
 cactcagacc tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240  
 tatcgtgac agaaagcggg atcgagtcta attaaacacg ctgaagaaat tcttcgtaag 300  
 aggggggcgg acttgctttg gtgtaatgcg cggacatccg cctcaggcta ctacaaaaag 360  
 ttaggcttca gcgagcaggg agagatattt gatacgcgcg cagtaggacc tcacatcctg 420  
 atgtataaaaa ggatcacata a 441

<210> 252

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 252

```

atgatagagg tgaaaccgat taacgcagag gatacctatg aactaaggca tagagtactc 60
agaccaaacc agccgtaga agtgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120
cacttaggcg gctattacag gggcaaactg atttccatag cttcattcca ccaggccgag 180
cactcagaac tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tatcgtgagc agaaagcggg atcgagtcta attaaacacg ctgaagaaat tcttcgtaag 300
aggggggcgg acttgctttg gtgcaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agaggtattt gagacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggctcacgta a                                     441

```

&lt;210&gt; 253

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 253

```

atgctagagg tgaaaccgat taacgcagag gatacttacg aactaaggca taaaatactc 60
agaccaaacc agccgataga agtgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120
cacttaggcg gctattacag gggcaaactg atttccatag cttcattcca ccaggccgag 180
cactcagaac tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tatcgtgagc agaaagcggg atcgagtcta attagacacg ctgaagaaat tcttcgtaag 300
aggggggcgg acttgctttg gtgtaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agagatattt gatacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggatcacgta a                                     441

```

&lt;210&gt; 254

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 254

```

atgctagagg tgaaaccgat taacgcagag gatacctatg aactaaggca tagaatactc 60
agaccaaacc agccgataga agtgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120
cacttaggcg gctttttacag gggcaaactg atttccatag cttcattcca ccaggccgag 180
cactcagaac tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tatcgtgagc agaaagcggg atcgagtcta attaaacacg ctgaagaaat tcttcgtaag 300
aggggggcgg acatgctttg gtgcaatgcg cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agagatattt gaaacgccgc cagtaggacc tcacatcctg 420
atgtataaaa ggctcacata a                                     441

```

&lt;210&gt; 255

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 255

```
atgctagagg tgaaaccgat taacgcagag gatacctatg aactaaggca taaaatactc 60
agaccaaacc agccgttaga agtgtgtatg tatgaaaccg atttacttcg tgggtgcattt 120
cacttaggag gctattacag gggcaaactg atttccatag cttcattcca ccaggccgag 180
cactcagacc tccaaggcca gaaacagtac cagctccgag gtatggctac cttggaaggt 240
tatcgtgagc agaaagcggg atcgagtcta attagacaag ctgaagaaat tcttcgtaag 300
aggggggagg acttgctttg gtgtaatgag cggacatccg cctcaggcta ctacaaaaag 360
ttaggcttca gcgagcaggg agagggtattt gatacgcgcg cagtaggacc tcacatcctg 420
atgtataaaa ggctcacata a 441
```

&lt;210&gt; 256

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 256

```
atgattgaag tcaaacctat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgctgga agcatgcaag tatgaaaccg atttgctcgg gggcacgttt 120
cactcgggag gatattaccg gggcaagctg atcagcatcg cttcctttca taatgccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaaggga 240
taccgtgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcgg accttttatg gtgcaacgcc aggacatctg tgagcgggta ctataaaaaag 360
ctcggttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacgta a 441
```

&lt;210&gt; 257

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 257

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggt gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggga 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaatgcc aggacatttg tgagcgggta ctatgaaaag 360
ctcggttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc ttatattttg 420
atgtattaga aattgacata a 441
```

&lt;210&gt; 258

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

467

86918

468

&lt;400&gt; 258

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cgcccggaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggtagcttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgcgagc aaaaagcggg tagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg tgagcgggta ctataaaaag 360
ctcggcttca gcgaacaagg cggggtctgc gatataccgc cgatcggacc tcatattttg 420
atgtataaga aattggcata a 441
```

&lt;210&gt; 259

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 259

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcatactc 60
cgcccggaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgcgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcag accttttatg gtgcaacgcc aggacatctg tgagcgggta ctatgaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacata a 441
```

&lt;210&gt; 260

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 260

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cgcccggaatc agccgcttga agcatgcaag tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca taatgccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac gcttgaaggg 240
taccgcgagc aaaaagcggg aagcacgctc atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcag accttttatg gtgcaacgcc aggacatctg tgagcgggta ctataaaaag 360
ctcggcttca gcgaacaggg cgaagtctac gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacgta a 441
```

&lt;210&gt; 261

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 261

```
atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcatactc 60
```

469

86918

470

```

cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180
catccagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgagagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aaaggcgcag accttttatg gtgcaacgcc aggacatctg tgagcggcta ctatgaaaag 360
ctcggcttca gcgaacaggg cgaagtctgc gacataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacata a                                     441

```

&lt;210&gt; 262

&lt;211&gt; 441

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 262

```

atgattgaag tcaaaccaat aaacgcggaa gatacgtatg agatcaggca ccgcattctc 60
cggccgaatc agccgcttga agcatgtatg tatgaaaccg atttgctcgg gggcacgttt 120
cacctcgggtg gatattaccg gggcaagctg atcagcatcg cttcctttca tcaagccgaa 180
cattcagagc ttgaaggcca aaaacagtat cagctgagag ggatggcgac acttgaaggg 240
taccgtgagc aaaaagcggg cagtacgctt atccgccatg ccgaagagct tcttcggaaa 300
aagggggcag accttttatg gtgcaacgcc aggacatctg tgagcgggta ctataaaaag 360
ctcggcttca gcgaacaagg cgggggtctac gatataccgc cgatcggacc tcatattttg 420
atgtataaga aattgacgta a                                     441

```

&lt;210&gt; 263

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 263

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
          20           25           30
Thr Asp Leu Leu Arg Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
          50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
          85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
          115          120          125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
          130          135          140
Leu Thr
145

```

<210> 264  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 264  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Arg Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Tyr Asp Ile Pro Pro Thr Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

<210> 265  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 265  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Ser Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe Asn Gln Ala Glu His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Tyr Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys



473

86918

474

130

135

140

Leu Thr

145

&lt;210&gt; 266

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 266

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
          20           25           30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
          35           40           45
Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
          50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
          85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
          115          120          125
Val Tyr Asp Thr Pro Pro Thr Gly Pro His Ile Leu Met Tyr Lys Lys
          130          135          140
Leu Thr
145

```

&lt;210&gt; 267

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 267

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
          20           25           30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
          35           40           45
Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
          50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Ala
          85           90           95

```

475

86918

476

```

Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100      105      110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115      120      125
Val Tyr Asp Thr Pro Pro Ala Gly Pro His Ile Leu Met Tyr Lys Lys
      130      135      140
Leu Thr
145

```

&lt;210&gt; 268

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 268

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
  1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
      20      25      30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
      35      40      45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
      50      55      60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
      65      70      75      80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85      90      95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100      105      110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115      120      125
Val Tyr Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
      130      135      140
Leu Thr
145

```

&lt;210&gt; 269

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 269

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
  1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
      20      25      30
Thr Asp Leu Leu Gly Gly Thr Leu His Leu Gly Gly Tyr Tyr Arg Gly
      35      40      45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu

```

477					86918					478				
50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu
65					70					75				80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu
				85					90					95
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg
			100					105					110	
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly
		115					120					125		
Val	Tyr	Asp	Ile	Pro	Pro	Thr	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys
	130					135					140			
Leu	Thr													
145														

<210> 270  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 270															
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Arg	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Val	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Glu
65					70					75					80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Gly
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 271  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 271															
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	

479

86918

480

```

His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
      20      25      30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
      35      40      45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
      50      55      60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65      70      75      80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85      90      95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100      105      110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115      120      125
Val Tyr Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
      130      135      140
Leu Thr
145

```

&lt;210&gt; 272

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 272

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
  1      5      10      15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
      20      25      30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
      35      40      45
Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
      50      55      60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu
65      70      75      80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85      90      95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100      105      110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115      120      125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
      130      135      140
Leu Thr
145

```

&lt;210&gt; 273

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

## &lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 273

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly
           115          120          125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Thr
145

```

&lt;210&gt; 274

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

## &lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 274

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
           50           55           60
Glu Gly Arg Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
           115          120          125
Val Tyr Asp Ile Pro Pro Thr Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Thr
145

```

&lt;210&gt; 275

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 275

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1             5             10             15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Leu His Leu Gly Gly Tyr Tyr Arg Gly
       35       40       45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
   50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
       100       105       110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly
   115           120           125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
   130           135           140
Leu Thr
145

```

&lt;210&gt; 276

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 276

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1             5             10             15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
       35       40       45
Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
   50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
       100       105       110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
   115           120           125
Val Tyr Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
   130           135           140

```

Leu Thr  
145

<210> 277

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 277

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Glu
65					70					75				80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Tyr	Asp	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135						140			

Leu Thr  
145

<210> 278

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 278

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Arg	Gly	Ala	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75				80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr

487

86918

488

		100					105					110	
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln
		115					120					125	
Val	Tyr	Asp	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr
		130					135					140	
Leu	Thr												
145													

&lt;210&gt; 279

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 279

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20				25						30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35				40						45			
Lys	Leu	Val	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
	50					55				60					
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65				70					75					80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
			85					90					95		
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100				105						110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Thr	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
		130					135					140			
Leu	Thr														
145															

&lt;210&gt; 280

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 280

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20				25						30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35				40						45			
Lys	Leu	Val	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
	50					55				60					



489

86918

490

```

Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65      70      75      80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85      90      95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100      105      110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115      120      125
Val Tyr Asp Ile Pro Pro Thr Gly Pro His Ile Leu Met Tyr Lys Lys
      130      135      140
Leu Thr
145

```

&lt;210&gt; 281

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 281

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
1      5      10      15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
      20      25      30
Thr Asp Leu Leu Gly Gly Thr Leu His Leu Gly Gly Tyr Tyr Arg Gly
      35      40      45
Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
      50      55      60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65      70      75      80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85      90      95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100      105      110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115      120      125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
      130      135      140
Leu Thr
145

```

&lt;210&gt; 282

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 282

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
1      5      10      15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu

```

491					86918					492					
20					25					30					
Thr	Asp	Leu	Leu	Gly	Gly	Ala	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65						70					75				80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Gly
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 283  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 283															
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20					25				30			
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Gln	Gly
	35						40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65						70				75					80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Val	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Gly
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 284  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

493

86918

494

&lt;400&gt; 284

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Gln Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Glu Lys Leu Gly Phe Ser Glu Gln Gly Gly
           115          120          125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Ala
145

```

&lt;210&gt; 285

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 285

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Ala Phe His Leu Gly Gly Tyr Tyr Gln Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Lys Ala Glu His Ser Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Val Arg Gly Tyr Tyr Glu Lys Leu Gly Phe Ser Glu Gln Gly Gly
           115          120          125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Thr
145

```

&lt;210&gt; 286

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 286

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
           115          120          125
Val His Asp Ile Pro Pro Thr Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Thr
145

```

&lt;210&gt; 287

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 287

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
           115          120          125
Val Tyr Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Thr

```

145

&lt;210&gt; 288

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 288

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
          20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
          50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
          85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly
          115          120          125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
          130          135          140
Leu Thr
145

```

&lt;210&gt; 289

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 289

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
          20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
          50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
          85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110

```

499

86918

500

Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Tyr Asp Ala Pro Pro Thr Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

&lt;210&gt; 290

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 290

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly  
 115 120 125  
 Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

&lt;210&gt; 291

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 291

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu

501					86918					502					
65				70				75				80			
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85				90						95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Gly
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 292  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 292															
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65				70					75					80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
			85					90						95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Thr	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 293  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 293															
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20					25					30		

503					86918					504					
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Val	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75				80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Ala
			85						90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Thr	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 294

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 294

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10				15		
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20				25					30			
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
	35					40					45				
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75				80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
			85						90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Gly
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Thr	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 295

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence



505

86918

506

&lt;400&gt; 295

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
      20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
      35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
      50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115          120          125
Val Tyr Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
      130          135          140
Leu Thr
145

```

&lt;210&gt; 296

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 296

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
      20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Gln Gly
      35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Lys Ala Glu His Ser Glu Leu
      50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100          105          110
Ser Val Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly
      115          120          125
Ile Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
      130          135          140
Leu Thr
145

```

&lt;210&gt; 297

&lt;211&gt; 146

&lt;212&gt; PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 297

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Gln Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Lys Ala Glu His Ser Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Val Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
           115          120          125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Thr
145

```

<210> 298

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 298

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly
           115          120          125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Thr
145

```

<210> 299  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 299  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe Asn Gln Ala Glu His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met His Lys Lys  
 130 135 140  
 Leu Thr  
 145

<210> 300  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 300  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Ser Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe Asn Gln Ala Glu His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu

511		86918		512
115		120		125
Val Tyr Asp Thr Pro Pro	Val Gly Pro His Ile	Leu Met Tyr Lys Lys		
130	135	140		
Leu Thr				
145				

<210> 301  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 301

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg	
1 5 10 15	
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu	
20 25 30	
Thr Asp Leu Leu Arg Ser Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly	
35 40 45	
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu	
50 55 60	
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu	
65 70 75 80	
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu	
85 90 95	
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr	
100 105 110	
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu	
115 120 125	
Val Tyr Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys	
130 135 140	
Leu Thr	
145	

<210> 302  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 302

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg	
1 5 10 15	
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu	
20 25 30	
Thr Asp Leu Leu Gly Ser Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly	
35 40 45	
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu	
50 55 60	
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly	
65 70 75 80	

513

86918

514

```

Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85      90      95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100      105      110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115      120      125
Val Tyr Asp Ile Pro Pro Thr Gly Pro His Ile Leu Met Tyr Lys Lys
      130      135      140
Leu Thr
145

```

&lt;210&gt; 303

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 303

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
  1      5      10      15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
      20      25      30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
      35      40      45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
      50      55      60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
      65      70      75      80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85      90      95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100      105      110
Ser Ala Arg Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115      120      125
Val Tyr Asp Ile Pro Pro Thr Gly Pro His Ile Leu Met Tyr Lys Lys
      130      135      140
Leu Thr
145

```

&lt;210&gt; 304

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 304

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
  1      5      10      15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
      20      25      30
Thr Asp Ser Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly

```

515					86918					516				
	35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	Asn	Gln	Ala	Glu	His	Pro	Glu
	50					55					60			
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu
65						70					75			80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu
				85					90					95
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg
			100				105						110	
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly
		115				120						125		
Val	Tyr	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys
	130					135					140			
Leu	Thr													
145														

<210> 305

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 305

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20				25						30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35				40						45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Glu
65					70					75					80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100				105						110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115				120						125			
Val	Tyr	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 306

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 306

517

86918

518

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
           20           25           30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
           115          120          125
Val Tyr Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Thr
145

```

&lt;210&gt; 307

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 307

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
           20           25           30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
           115          120          125
Val Tyr Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Thr
145

```

&lt;210&gt; 308

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 308

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
          20           25           30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
          35           40           45
Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
          50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
          85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
          115          120          125
Val Tyr Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
          130          135          140
Leu Thr
145

```

&lt;210&gt; 309

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 309

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
          20           25           30
Thr Asp Leu Leu Gly Ser Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
          35           40           45
Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
          50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Ala
          85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
          115          120          125
Val Tyr Asp Thr Pro Pro Thr Gly Pro His Ile Leu Met Tyr Lys Lys
          130          135          140
Leu Thr
145

```



<210> 310  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 310  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Ala  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Arg Glu Gln Gly Gly  
 115 120 125  
 Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

<210> 311  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 311  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125

Val Tyr Asp Ile Pro Pro Thr Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

<210> 312

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 312

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe Asn Gln Ala Glu His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly  
 115 120 125  
 Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

<210> 313

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 313

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu

525

86918

526

			85					90					95				
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr		
			100					105					110				
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu		
		115					120					125					
Val	Tyr	Asp	Ile	Pro	Pro	Thr	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys		
	130					135					140						
Leu	Thr																
145																	

&lt;210&gt; 314

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 314

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg		
1				5					10					15			
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu		
			20				25						30				
Thr	Asp	Leu	Leu	Arg	Gly	Ala	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly		
	35					40						45					
Lys	Leu	Val	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu		
	50				55					60							
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly		
65				70					75					80			
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Ala		
			85					90					95				
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr		
			100				105						110				
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Gly		
		115				120						125					
Val	Tyr	Asp	Ile	Pro	Pro	Ala	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys		
	130					135					140						
Leu	Thr																
145																	

&lt;210&gt; 315

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 315

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg		
1				5					10					15			
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu		
			20				25						30				
Thr	Asp	Leu	Leu	Gly	Ser	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly		
	35					40						45					

527

86918

528

```

Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
  50                      55                      60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
  65                      70                      75                      80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
                      85                      90                      95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
                      100                      105                      110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
                      115                      120                      125
Val Tyr Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
                      130                      135                      140
Leu Thr
145

```

&lt;210&gt; 316

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 316

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
  1                      5                      10                      15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
                      20                      25                      30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
                      35                      40                      45
Lys Leu Val Ser Ile Ala Ser Phe Asn Gln Ala Glu His Pro Glu Leu
                      50                      55                      60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
  65                      70                      75                      80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
                      85                      90                      95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
                      100                      105                      110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
                      115                      120                      125
Val Tyr Asp Thr Pro Pro Val Gly Pro His Val Leu Met Tyr Lys Lys
                      130                      135                      140
Leu Thr
145

```

&lt;210&gt; 317

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 317

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg

```

529					86918					530					
1			5					10				15			
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
		50				55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75					80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
			85						90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Gly
		115					120					125			
Val	Tyr	Asp	Met	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
		130				135					140				
Leu	Thr														
145															

<210> 318

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 318

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	Asn	Gln	Ala	Glu	His	Pro	Glu	Leu
		50				55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75					80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Ala
			85						90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Tyr	Asp	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
		130				135					140				
Leu	Thr														
145															

<210> 319

<211> 146

<212> PRT

<213> Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 319

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
      20           25           30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
      35           40           45
Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
      50           55           60
Glu Gly Leu Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115          120          125
Val Tyr Asp Thr Pro Pro Thr Gly Pro His Ile Leu Met Tyr Lys Lys
      130          135          140
Leu Thr
145

```

&lt;210&gt; 320

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 320

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
      20           25           30
Thr Asp Leu Leu Arg Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
      35           40           45
Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
      50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115          120          125
Val Tyr Asp Thr Pro Pro Ala Gly Pro His Ile Leu Met Tyr Lys Lys
      130          135          140
Leu Thr
145

```

<210> 321  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 321  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Lys His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Tyr Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

<210> 322  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 322  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe Asn Gln Ala Glu His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Ala  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Tyr Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys

535 86918 536

130 135 140

Leu Thr  
145

<210> 323  
<211> 146  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Synthetic Protein Sequence

<400> 323  
Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
1 5 10 15  
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
20 25 30  
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
35 40 45  
Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu  
50 55 60  
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
65 70 75 80  
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
85 90 95  
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
100 105 110  
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
115 120 125  
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys  
130 135 140  
Leu Thr  
145

<210> 324  
<211> 146  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Synthetic Protein Sequence

<400> 324  
Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
1 5 10 15  
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
20 25 30  
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
35 40 45  
Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu  
50 55 60  
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu  
65 70 75 80  
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Ala  
85 90 95



537

86918

538

Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
			115				120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Thr	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
			130			135					140				
Leu	Thr														
145															

&lt;210&gt; 325

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 325

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20				25					30			
Thr	Asp	Leu	Leu	Arg	Gly	Ala	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35				40					45				
Lys	Leu	Val	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
	50				55				60						
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65				70					75					80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
			85					90				95			
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105				110			
Thr	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115				120					125				
Val	Phe	Asp	Ile	Pro	Pro	Thr	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

&lt;210&gt; 326

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 326

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20				25					30			
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35				40					45				
Lys	Leu	Val	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu

539					86918					540					
50					55				60						
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Glu
65					70				75					80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr	
			100				105					110			
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 327

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 327

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Arg	Gly	Ala	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35				40						45			
Lys	Leu	Val	Ser	Ile	Ala	Ser	Phe	Asn	Gln	Ala	Glu	His	Pro	Glu	Leu
	50					55				60					
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75				80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
			85						90					95	
Leu	Leu	Arg	Arg	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Thr	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 328

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 328

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	

541					86918					542					
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Arg	Gly	Ala	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
		50					55					60			
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65						70					75				80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Ala
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Thr	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
		130					135					140			
Leu	Thr														
145															

<210> 329

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 329

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
		50					55				60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65						70				75					80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Ala
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Thr	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
		130					135					140			
Leu	Thr														
145															

<210> 330

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

## &lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 330

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1         5         10         15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
      20         25         30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
      35         40         45
Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
      50         55         60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65         70         75         80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85         90         95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100        105        110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115        120        125
Val Tyr Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
      130        135        140
Leu Thr
145

```

&lt;210&gt; 331

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

## &lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 331

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1         5         10         15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
      20         25         30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
      35         40         45
Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
      50         55         60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65         70         75         80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85         90         95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100        105        110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Lys
      115        120        125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
      130        135        140
Leu Thr
145

```

&lt;210&gt; 332

<211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 332  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Arg Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

<210> 333  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 333  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Val Ser Phe His Gln Ala Glu His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Tyr Asp Thr Pro Pro Ala Gly Pro His Ile Leu Met Tyr Thr Lys  
 130 135 140

Leu Thr  
145

<210> 334  
<211> 146  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Synthetic Protein Sequence

<400> 334  
Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
1 5 10 15  
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
20 25 30  
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
35 40 45  
Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu  
50 55 60  
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu  
65 70 75 80  
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
85 90 95  
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
100 105 110  
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
115 120 125  
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys  
130 135 140  
Leu Thr  
145

<210> 335  
<211> 146  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Synthetic Protein Sequence

<400> 335  
Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Lys Ile Arg  
1 5 10 15  
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu  
20 25 30  
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
35 40 45  
Lys Leu Val Ser Ile Ala Ser Phe Asn Gln Ala Glu His Pro Glu Leu  
50 55 60  
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
65 70 75 80  
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Ala  
85 90 95  
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr

549		86918		550
	100		105	110
Ser	Ala	Ser	Gly	Tyr
	115		120	
Val	Tyr	Asp	Ile	Pro
	130		135	
Leu	Thr			
145				

<210> 336  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 336

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35				40						45			
Lys	Leu	Val	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
	50					55				60					
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Glu
65					70					75					80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Thr	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 337  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 337

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35				40						45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
	50					55					60				

551					86918					552					
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Glu
65					70					75					80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
			85						90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115				120						125			
Val	Tyr	Asp	Ile	Pro	Pro	Thr	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 338

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 338

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20				25						30		
Thr	Asp	Leu	Leu	Arg	Gly	Ala	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35				40						45			
Lys	Leu	Val	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Glu
65				70						75					80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
			85						90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Lys
		115				120						125			
Val	Tyr	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 339

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 339

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu



553					86918					554					
20					25					30					
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Glu
65				70					75					80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
			85						90					95	
Leu	Leu	Arg	Arg	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115				120						125			
Val	Tyr	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 340

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 340

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Lys	Ala	Glu	His	Ser	Glu	Leu
	50					55					60				
Glu	Gly	Glu	Glu	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65				70					75					80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	Tyr	Ala	Glu	Glu
			85						90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Val	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115				120						125			
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 341

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

555

86918

556

&lt;400&gt; 341

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
      20           25           30
Thr Asp Leu Leu Gly Gly Thr Leu His Leu Gly Gly Tyr Tyr Gln Gly
      35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Lys Ala Glu His Ser Gly Leu
      50           55           60
Glu Gly Glu Glu Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100          105          110
Ser Val Ser Gly Tyr Tyr Glu Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115          120          125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
      130          135          140
Leu Thr
145

```

&lt;210&gt; 342

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 342

```

Met Ile Asp Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
      20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Gln Gly
      35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
      50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100          105          110
Ser Val Ser Gly Tyr Tyr Glu Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115          120          125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
      130          135          140
Leu Thr
145

```

&lt;210&gt; 343

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 343

```

Met Ile Glu Val Lys Pro Ile Ser Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Ala Phe His Leu Gly Gly Tyr Tyr Gln Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Val Ser Gly Tyr Tyr Glu Lys Leu Gly Phe Ser Glu Gln Gly Glu
           115          120          125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Thr
145

```

&lt;210&gt; 344

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 344

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Lys Ala Glu His Ser Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Glu Lys Leu Gly Phe Ser Glu Gln Gly Glu
           115          120          125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Ala

```

145

&lt;210&gt; 345

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 345

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
          20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
          35           40           45
Lys Leu Ile Cys Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
          50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
          85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Val Arg Gly Tyr Tyr Glu Lys Leu Gly Phe Ser Glu Gln Gly Gly
          115          120          125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
          130          135          140
Leu Ala
145

```

&lt;210&gt; 346

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 346

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Cys Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
          20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Gln Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Lys Ala Glu His Ser Glu Leu
          50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
          85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110

```

561					86918					562					
Ser	Val	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 347  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 347

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Gln	Gly
		35				40						45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Asn	Ala	Glu	His	Ser	Glu	Leu
	50					55				60					
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65				70					75					80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Val	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Val	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Gly
		115				120						125			
Ile	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Ala														
145															

<210> 348  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 348

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35				40						45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Glu	Leu
	50					55				60					
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly

563

86918

564

```

65          70          75          80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
          85          90
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Val Ser Gly Tyr Tyr Glu Lys Leu Gly Phe Ser Glu Gln Gly Glu
          115          120          125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
          130          135          140
Leu Thr
145

```

&lt;210&gt; 349

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 349

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1          5          10          15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
          20          25          30
Thr Asp Leu Leu Gly Gly Ala Phe His Leu Gly Gly Tyr Tyr Gln Gly
          35          40          45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
          50          55          60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65          70          75          80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
          85          90          95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Val Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
          115          120          125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
          130          135          140
Leu Thr
145

```

&lt;210&gt; 350

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 350

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1          5          10          15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
          20          25          30

```

565					86918					566					
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Gln	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Lys	Ala	Glu	His	Ser	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75					80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Glu	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Ala														
145															

<210> 351  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 351															
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20					25				30			
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Asp
	35						40					45			
Arg	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75					80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105				110			
Ser	Val	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 352  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

567

86918

568

&lt;400&gt; 352

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Ile
           100          105          110
Ser Val Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly
           115          120          125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Ala
145

```

&lt;210&gt; 353

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 353

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Asn Ala Glu His Ser Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Glu Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
           115          120          125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Ala
145

```

&lt;210&gt; 354

&lt;211&gt; 146

&lt;212&gt; PRT



<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 354

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Asp
           35           40           45
Arg Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Glu Lys Leu Gly Phe Ser Glu Gln Gly Glu
           115          120          125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Thr
145

```

<210> 355

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 355

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Gln Gly
           35           40           45
Lys Leu Ile Ser Thr Ala Ser Phe His Gln Ala Gly His Ser Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Arg Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Ile
           100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly
           115          120          125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Thr
145

```

<210> 356  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 356  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Phe Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Ala Phe His Leu Gly Gly Tyr Tyr Gln Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Val Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

<210> 357  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 357  
 Met Ile Glu Val Arg Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu

	573		86918		574
	115		120		125
Ala	Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys				
	130		135		140
Leu Thr					
145					

<210> 358  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 358

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75				80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Gly
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu Thr															
145															

<210> 359  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 359

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75				80	

575					86918					576					
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115				120						125			
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 360  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 360															
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Ala	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35				40						45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Pro	Ala	Glu	His	Ser	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65				70					75					80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
			85						90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Ile
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Glu	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115				120						125			
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 361  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 361															
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly

577					86918					578				
	35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Glu
	50					55					60			
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu
65						70					75			80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu
				85					90					95
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg
			100				105						110	Thr
Ser	Ala	Ser	Gly	Tyr	Tyr	Glu	Lys	Leu	Gly	Phe	Ser	Lys	Gln	Gly
		115					120					125		Glu
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys
	130					135					140			Lys
Leu	Thr													
145														

<210> 362

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 362

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25				30			
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75					80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Ile
			100				105						110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Gly
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 363

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 363

579					86918					580					
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65						70				75					80
Tyr	Arg	Gly	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85						90				95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Glu	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 364

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 364

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Gln	Asp
		35					40					45			
Arg	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65						70				75					80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	Tyr	Ala	Glu	Glu
				85						90				95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Ile
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Glu	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 365

<211> 146

<212> PRT

<213> Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 365

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Glu Lys Leu Gly Phe Ser Glu Gln Gly Glu
           115          120          125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Thr
145

```

&lt;210&gt; 366

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 366

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
           115          120          125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Thr
145

```

<210> 367  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 367  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Ile  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Glu Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

<210> 368  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 368  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly  
 115 120 125



Val Tyr Asp Ile Pro Pro Ile Gly Pro Tyr Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

<210> 369

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 369

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly  
 115 120 125  
 Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

<210> 370

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 370

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu

587		85		90		588		95							
Leu	Leu	Arg	Lys	Gly	Ala	Asp	Leu	Trp	Cys	Asn	Ala	Arg	Thr		
		100				105				110					
Ser	Ala	Ser	Gly	Tyr	Tyr	Glu	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115				120					125				
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 371  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 371															
Met	Ile	Glu	Val	Asn	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Ser	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Asn	Ala	Glu	His	Ser	Glu	Leu
	50					55				60					
Asp	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75				80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Val	Ser	Gly	Tyr	Tyr	Glu	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Ser	Met	Tyr	Lys	Lys
	130					135						140			
Leu	Thr														
145															

<210> 372  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 372															
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			

589					86918					590					
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
50					55					60					
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75					80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Gly
			115				120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Ala														
145															

<210> 373  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 373															
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10				15		
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Ala	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35				40						45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ser	Glu	His	Pro	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75					80
Tyr	Arg	Glu	Leu	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Ile
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
			115				120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 374  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 374															
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg

591					86918					592					
1			5					10				15			
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25				30			
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Glu	Leu
		50				55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Glu
65				70						75				80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
			85						90				95		
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Ile
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Gly
		115				120						125			
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
		130				135					140				
Leu	Thr														
145															

<210> 375

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 375

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10				15		
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25				30			
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Glu	Leu
		50				55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65				70						75				80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
			85						90				95		
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Gly
		115				120						125			
Val	Tyr	Gly	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
		130				135					140				
Leu	Thr														
145															

<210> 376

<211> 146

<212> PRT

<213> Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 376

```

Met Ile Glu Ala Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
           115          120          125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
130           135           140
Leu Thr
145

```

&lt;210&gt; 377

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 377

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu His Gly Glu
           115          120          125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
130           135           140
Leu Thr
145

```

<210> 378  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 378  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Val Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

<210> 379  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 379  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Glu Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys

597

86918

598

130  
Leu Thr  
145

135

140

<210> 380  
<211> 146  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Synthetic Protein Sequence

<400> 380  
Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
1 5 10 15  
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu  
20 25 30  
Thr Asp Leu Leu Gly Gly Ala Phe His Leu Gly Gly Tyr Tyr Gln Gly  
35 40 45  
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Val His Ser Glu Leu  
50 55 60  
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
65 70 75 80  
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
85 90 95  
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
100 105 110  
Ser Val Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly  
115 120 125  
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys  
130 135 140  
Leu Thr  
145

<210> 381  
<211> 146  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Synthetic Protein Sequence

<400> 381  
Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
1 5 10 15  
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu  
20 25 30  
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
35 40 45  
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu  
50 55 60  
Glu Gly His Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu  
65 70 75 80  
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
85 90 95

599					86918					600					
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Gly
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 382  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 382

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35				40						45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Glu	Leu
	50					55				60					
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65				70					75					80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
			85					90					95		
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115				120						125			
Val	Tyr	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 383  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 383

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35				40						45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Glu	Leu



601					86918					602				
50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu
65					70					75				80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu
				85					90					95
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg
			100					105					110	Thr
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly
		115					120					125		Gly
Val	Tyr	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys
	130					135					140			Lys
Leu	Thr													
145														

<210> 384  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 384														
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile
1				5					10				15	Arg
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr
			20					25				30		Glu
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg
		35					40					45		Gly
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Val	His	Ser	Glu
	50					55					60			Leu
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu
65					70					75				80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu
				85					90					95
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg
			100					105					110	Thr
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly
		115					120					125		Gly
Val	Tyr	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys
	130					135					140			Lys
Leu	Thr													
145														

<210> 385  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 385														
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile
1				5					10				15	Arg

603					86918					604					
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20					25					30		
Ala	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
		50				55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65						70					75				80
Asn	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115				120						125			
Val	Tyr	Asp	Val	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
		130				135					140				
Leu	Thr														
145															

<210> 386

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 386

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Arg	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
		50				55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65						70				75					80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Ile
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115				120						125			
Val	Tyr	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
		130				135					140				
Leu	Thr														
145															

<210> 387

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

## &lt;223&gt; Synthetic Protein Sequence

&lt;221&gt; unsure

&lt;222&gt; 17

&lt;223&gt; Xaa = His or Pro

&lt;400&gt; 387

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
Xaa Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
      20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
      35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
      50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly
      115          120          125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
      130          135          140
Leu Thr
145

```

&lt;210&gt; 388

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 388

```

Met Ile Glu Ala Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
      20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
      35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
      50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115          120          125
Val Tyr Asp Ile Pro Pro Thr Gly Pro His Ile Leu Met Tyr Lys Lys
      130          135          140
Leu Thr

```

145

&lt;210&gt; 389

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 389

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1             5             10             15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
             20             25             30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
             35             40             45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
             50             55             60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65             70             75             80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
             85             90             95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
             100            105            110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
             115            120            125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
130            135            140
Leu Thr
145

```

&lt;210&gt; 390

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 390

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1             5             10             15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
             20             25             30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
             35             40             45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
             50             55             60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65             70             75             80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
             85             90             95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
100            105            110

```

609

86918

610

Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

&lt;210&gt; 391

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 391

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Leu His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg Gln Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly  
 115 120 125  
 Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Glu  
 130 135 140  
 Leu Thr  
 145

&lt;210&gt; 392

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 392

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Arg Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu

611					86918					612					
65				70				75				80			
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 393

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 393

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Gly	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20				25						30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35				40						45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65				70					75					80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
			85						90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Gly
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 394

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 394

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20				25						30		

613

86918

614

```

Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
      35      40      45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
      50      55      60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65      70      75      80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85      90      95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100      105      110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115      120      125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
      130      135      140
Leu Thr
145

```

&lt;210&gt; 395

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 395

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
  1      5      10      15
His Arg Val Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
      20      25      30
Thr Asp Leu Leu Gly Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
      35      40      45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
      50      55      60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65      70      75      80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85      90      95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100      105      110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115      120      125
Val Tyr Asp Val Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
      130      135      140
Leu Thr
145

```

&lt;210&gt; 396

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

615

86918

616

&lt;400&gt; 396

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Arg Pro Leu Glu Ala Cys Met Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Gly Leu
           50           55           60
Glu Gly Lys Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly
           115          120          125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Thr
145

```

&lt;210&gt; 397

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 397

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
           20           25           30
Thr Asp Leu Leu Glu Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
           115          120          125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Thr
145

```

&lt;210&gt; 398

&lt;211&gt; 146

&lt;212&gt; PRT



<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 398

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
          20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
          50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Asp Glu
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
          85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
          115          120          125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
          130          135          140
Leu Thr
145

```

<210> 399

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 399

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Gly
          20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu Gln Pro Glu Leu
          50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
          85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly
          115          120          125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
          130          135          140
Leu Thr
145

```

&lt;210&gt; 400

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 400

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Ile Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly
           115          120          125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Thr
145

```

&lt;210&gt; 401

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 401

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Val Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
           50           55           60
Glu Gly Gln Lys Pro Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu

```

621		86918		622
	115		120	
Ile Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys				
130		135		140
Leu Thr				
145				

<210> 402  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 402

Met Ile Glu Ile Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg															
1				5				10						15	
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu															
			20				25						30		
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly															
			35				40						45		
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu															
			50			55				60					
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly															
65					70			75						80	
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu															
				85			90							95	
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr															
			100				105						110		
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu															
			115			120							125		
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys															
			130			135							140		
Leu Thr															
145															

<210> 403  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 403

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg															
1				5				10						15	
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu															
			20				25						30		
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly															
			35				40						45		
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu															
			50			55				60					
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly															
65					70			75						80	

```

Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85                      90                      95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100                      105                      110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115                      120                      125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
      130                      135                      140
Leu Thr
145

```

<210> 404

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 404

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
      20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
      35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
      50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115          120          125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
      130          135          140
Leu Thr
145

```

<210> 405

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 405

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
      20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly

```

625					86918					626				
	35				40					45				
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	
	50					55				60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	
65					70				75				80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	
			85					90				95		
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	
			100				105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	
	115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	
	130					135					140			
Leu	Thr													
145														

<210> 406

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 406

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10				15		
His	Arg	Val	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20				25					30			
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
	35					40					45				
Lys	Leu	Ile	Gly	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
	50					55				60					
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65				70					75					80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
			85					90				95			
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100				105					110			
Ser	Ala	Ser	Gly	Tyr	Tyr	Glu	Lys	Leu	Gly	Phe	Ser	Gly	Gln	Gly	Glu
	115					120					125				
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 407

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 407

627					86918					628					
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Glu
65					70					75					80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 408

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 408

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Glu
65					70					75					80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	His	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 409

<211> 146

<212> PRT

<213> Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 409

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
          20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
          50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu
          65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
          85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Ile
          100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly
          115          120          125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
          130          135          140
Leu Thr
145

```

&lt;210&gt; 410

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 410

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Ala Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
          20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
          50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
          65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
          85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly
          115          120          125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
          130          135          140
Leu Thr
145

```

<210> 411  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 411  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly  
 115 120 125  
 Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

<210> 412  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 412  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly  
 115 120 125



Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

<210> 413

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 413

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Leu His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

<210> 414

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 414

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Val Cys Ile Ala Ser Phe His Lys Ala Glu His Ser Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Asp Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu

635					86918					636				
			85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110	
Ser	Val	Ser	Gly	Tyr	Tyr	Glu	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly
			115					120					125	
Val	Tyr	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys
			130					135					140	
Leu	Thr													
145														

<210> 415  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 415															
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Glu
65				70					75					80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85				90					95		
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Gly
			115				120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
			130				135					140			
Leu	Thr														
145															

<210> 416  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 416															
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			

637

86918

638

Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
50						55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75					80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Gly
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

&lt;210&gt; 417

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 417

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75					80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Gly
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

&lt;210&gt; 418

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 418

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

639					86918					640				
1			5					10				15		
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr
			20					25				30		
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg
		35					40					45		
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu
		50				55					60			
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu
65					70					75				80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu
			85						90				95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg
			100					105					110	
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly
		115					120					125		
Val	Tyr	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys
	130					135					140			
Leu	Thr													
145														

<210> 419

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 419

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10				15		
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25				30			
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
		50				55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75					80
Tyr	Arg	Glu	Gln	Lys	Val	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
			85						90				95		
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Gly
		115					120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Thr														
145															

<210> 420

<211> 146

<212> PRT

<213> Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 420

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1         5         10         15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
      20         25         30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
      35         40         45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
      50         55         60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65         70         75         80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85         90         95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100        105        110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115        120        125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
      130        135        140
Leu Thr
145

```

&lt;210&gt; 421

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 421

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1         5         10         15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
      20         25         30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
      35         40         45
Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
      50         55         60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu
65         70         75         80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85         90         95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100        105        110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115        120        125
Val Tyr Asp Ile Pro Pro Thr Gly Pro His Ile Leu Met Tyr Lys Lys
      130        135        140
Leu Thr
145

```

<210> 422  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 422  
 Met Ile Glu Ala Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Thr Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

<210> 423  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 423  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Asp Ala Cys Lys Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys

645	86918	646
130	135	140
Leu Thr		
145		

<210> 424  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 424  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Ala  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Tyr Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

<210> 425  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 425  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Met Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu  
 65 70 75 80  
 Tyr His Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95

647

86918

648

```

Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100      105      110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115      120      125
Val Tyr Asn Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
      130      135      140
Leu Thr
145

```

&lt;210&gt; 426

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 426

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
  1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
      20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
      35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
      50           55           60
Val Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
      65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115          120          125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
      130          135          140
Leu Thr
145

```

&lt;210&gt; 427

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 427

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
  1           5           10           15
His Arg Val Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
      20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
      35           40           45
Glu Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu

```



649

86918

650

```

      50              55              60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65              70              75              80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85              90              95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100              105              110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly
      115              120              125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
      130              135              140
Leu Thr
145

```

&lt;210&gt; 428

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 428

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
1              5              10              15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
      20              25              30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
      35              40              45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
      50              55              60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65              70              75              80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85              90              95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100              105              110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Asp Gln Gly Glu
      115              120              125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
      130              135              140
Leu Thr
145

```

&lt;210&gt; 429

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 429

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
1              5              10              15

```

651

86918

652

```

His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
      20      25      30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
      35      40      45
Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
      50      55      60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65      70      75      80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85      90      95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100      105      110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115      120      125
Val Tyr Asp Leu Pro Pro Thr Gly Pro His Ile Leu Met Tyr Lys Lys
      130      135      140
Leu Thr
145

```

&lt;210&gt; 430

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 430

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
  1      5      10      15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
      20      25      30
Thr Asp Leu Leu Gly Gly Thr Leu His Leu Gly Gly Tyr Tyr Arg Gly
      35      40      45
Lys Leu Val Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
      50      55      60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65      70      75      80
Tyr Arg Gly Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85      90      95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100      105      110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly
      115      120      125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
      130      135      140
Leu Thr
145

```

&lt;210&gt; 431

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

## &lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 431

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
           20           25           30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Ala
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
           115          120          125
Val Tyr Asp Ile Pro Pro Thr Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Thr
145

```

&lt;210&gt; 432

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

## &lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 432

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
           20           25           30
Thr Asp Leu Leu Gly Gly Thr Leu His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
           50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
           85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Ile Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
           115          120          125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
           130          135          140
Leu Thr
145

```

&lt;210&gt; 433

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 433

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
          20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe Asn Gln Ala Glu His Pro Glu Leu
          50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
          85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
          115          120          125
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
          130          135          140
Leu Thr
145

```

&lt;210&gt; 434

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 434

```

Met Ile Glu Val Lys Pro Ile Thr Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu
          20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
          50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
          85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
          115          120          125
Val Tyr Asp Ile Pro Pro Thr Gly Pro His Ile Leu Met Tyr Lys Lys
          130          135          140

```

Leu Thr  
145

<210> 435  
<211> 146  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Synthetic Protein Sequence

<400> 435  
Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Ala Tyr Glu Ile Arg  
1 5 10 15  
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
20 25 30  
Thr Asp Leu Leu Gly Ser Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
35 40 45  
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu  
50 55 60  
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Glu  
65 70 75 80  
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
85 90 95  
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
100 105 110  
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
115 120 125  
Val Tyr Asp Ile Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys  
130 135 140  
Leu Thr  
145

<210> 436  
<211> 146  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Synthetic Protein Sequence

<400> 436  
Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
1 5 10 15  
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu  
20 25 30  
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Gln Gly  
35 40 45  
Lys Leu Ile Ser Ile Ala Ser Phe His Asn Ala Glu His Ser Glu Leu  
50 55 60  
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
65 70 75 80  
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
85 90 95  
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr

	659		86918		660
	100		105		110
Ser Val	Ser Gly Tyr Tyr Glu	Lys Leu Gly Phe	Ser Glu Gln Gly Glu		
	115	120	125		
Val Tyr	Asp Ile Pro Pro Val	Gly Pro His Ile	Leu Met Tyr Lys Lys		
	130	135	140		
Leu Ala					
145					

<210> 437  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 437

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg	
1 5 10 15	
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu	
20 25 30	
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly	
35 40 45	
Lys Leu Ile Ser Ile Ala Ser Phe His Lys Ala Glu His Ser Glu Leu	
50 55 60	
Glu Gly Glu Glu Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly	
65 70 75 80	
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu	
85 90 95	
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr	
100 105 110	
Ser Val Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu	
115 120 125	
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys	
130 135 140	
Leu Thr	
145	

<210> 438  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 438

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg	
1 5 10 15	
His Lys Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu	
20 25 30	
Ser Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Arg Gly	
35 40 45	
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu	
50 55 60	

661

86918

662

```

Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65          70          75          80
Tyr Arg Asp Gln Lys Ala Gly Ser Thr Leu Ile Lys His Ala Glu Glu
          85          90          95
Ile Leu Arg Lys Arg Gly Ala Asp Met Leu Trp Cys Asn Ala Arg Thr
          100         105         110
Thr Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
          115         120         125
Ile Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
          130         135         140
Leu Thr
145

```

&lt;210&gt; 439

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 439

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
1          5          10          15
His Lys Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu
          20          25          30
Ser Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Gly Gly
          35          40          45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu
          50          55          60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65          70          75          80
Tyr Arg Asp Gln Lys Ala Gly Ser Thr Leu Ile Lys His Ala Glu Gln
          85          90          95
Leu Leu Arg Lys Arg Gly Ala Asp Met Leu Trp Cys Asn Ala Arg Thr
          100         105         110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
          115         120         125
Val Phe Glu Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
          130         135         140
Ile Thr
145

```

&lt;210&gt; 440

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 440

```

Met Leu Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
1          5          10          15
His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu

```

663					86918					664					
20					25					30					
Thr	Asp	Leu	Leu	Arg	Gly	Ala	Phe	His	Leu	Gly	Gly	Phe	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Glu	Leu
	50					55					60				
Gln	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75					80
Tyr	Arg	Asp	Gln	Lys	Ala	Gly	Ser	Ser	Leu	Ile	Lys	His	Ala	Glu	Gln
			85						90					95	
Leu	Leu	Arg	Lys	Arg	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115				120					125				
Val	Phe	Asp	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Arg
	130					135					140				
Ile	Thr														
145															

<210> 441

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 441

Met	Leu	Glu	Val	Lys	Leu	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Leu	Arg
1				5					10				15		
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Arg	Gly	Ala	Phe	His	Leu	Gly	Gly	Phe	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Asp	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75					80
Phe	Arg	Asp	Gln	Lys	Ala	Gly	Ser	Ser	Leu	Ile	Lys	His	Ala	Glu	Glu
			85						90					95	
Ile	Leu	Arg	Lys	Arg	Gly	Ala	Asn	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115				120					125				
Val	Phe	Asp	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Arg
	130					135					140				
Ile	Thr														
145															

<210> 442

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence



665

86918

666

&lt;400&gt; 442

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1           5           10           15
His Lys Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu
           20           25           30
Ser Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu
           50           55           60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Asp Gln Lys Ala Gly Ser Ser Leu Ile Arg His Ala Glu Gln
           85           90           95
Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
           115          120          125
Ile Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
           130          135          140
Leu Thr
145

```

&lt;210&gt; 443

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 443

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1           5           10           15
His Lys Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
           20           25           30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Gly Leu
           50           55           60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Ser Ile Ile Lys His Ala Glu Glu
           85           90           95
Ile Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
           115          120          125
Ile Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
           130          135          140
Ile Thr
145

```

&lt;210&gt; 444

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 444

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu
           20           25           30
Ser Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu
           50           55           60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Gln
           85           90           95
Leu Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Arg Leu Gly Phe Ser Glu Gln Gly Glu
           115          120          125
Val Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
           130          135          140
Leu Thr
145

```

&lt;210&gt; 445

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 445

```

Met Leu Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu
           20           25           30
Ser Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
           35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
           50           55           60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Ser Leu Ile Lys His Ala Glu Glu
           85           90           95
Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
           115          120          125
Ile Phe Glu Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
           130          135          140
Ile Thr

```

145

&lt;210&gt; 446

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 446

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Glu Thr Tyr Glu Leu Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu
          20           25           30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Arg Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
          50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Asp Gln Lys Ala Gly Ser Ser Leu Ile Arg His Ala Glu Gln
          85           90           95
Leu Leu Arg Lys Lys Gly Ala Asn Met Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Thr Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
          115          120          125
Ile Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
          130          135          140
Ile Thr
145

```

&lt;210&gt; 447

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 447

```

Met Leu Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu
          20           25           30
Ser Asp Leu Leu Arg Gly Ala Leu His Leu Gly Gly Phe Tyr Arg Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu
          50           55           60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Phe Arg Asp Gln Lys Ala Gly Ser Ser Leu Ile Arg His Ala Glu Gln
          85           90           95
Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110

```

671

86918

672

```

Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Lys
    115                      120                      125
Val Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
    130                      135                      140
Ile Thr
145

```

&lt;210&gt; 448

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 448

```

Met Leu Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1          5          10          15
His Lys Ile Leu Arg Pro Asn Gln Pro Leu Glu Val Cys Met Tyr Glu
    20          25          30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Arg Gly
    35          40          45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu
    50          55          60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65          70          75          80
Tyr Arg Asp Gln Lys Ala Gly Ser Ser Leu Ile Arg His Ala Glu Gln
    85          90          95
Ile Leu Arg Lys Arg Gly Ala Asp Met Leu Trp Cys Asn Ala Arg Thr
    100         105         110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
    115         120         125
Val Phe Glu Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
    130         135         140
Leu Thr
145

```

&lt;210&gt; 449

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 449

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1          5          10          15
His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu
    20          25          30
Ser Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
    35          40          45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu
    50          55          60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly

```

673					86918					674					
65				70				75				80			
Phe	Arg	Asp	Gln	Lys	Ala	Gly	Ser	Ser	Leu	Ile	Arg	His	Ala	Glu	Gln
				85				90						95	
Ile	Leu	Arg	Lys	Arg	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Phe	Glu	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Arg
	130					135					140				
Ile	Thr														
145															

<210> 450  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 450															
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Leu	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Ile	Glu	Ala	Cys	Met	Tyr	Glu
			20					25					30		
Ser	Asp	Leu	Leu	Arg	Gly	Ala	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35				40						45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Glu	Leu
	50					55					60				
Gln	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65				70					75					80	
Phe	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Gln
				85					90					95	
Ile	Leu	Arg	Lys	Arg	Gly	Ala	Asp	Met	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115				120						125			
Ile	Phe	Asp	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Arg
	130					135					140				
Ile	Thr														
145															

<210> 451  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 451															
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Leu	Arg
1				5					10					15	
His	Lys	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Ile	Glu	Ala	Cys	Met	Tyr	Glu
			20					25					30		



<210> 453  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 453  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Val Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu  
 50 55 60  
 Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Gln  
 85 90 95  
 Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg  
 130 135 140  
 Ile Thr  
 145

<210> 454  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 454  
 Met Leu Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Phe Glu  
 20 25 30  
 Ser Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu  
 50 55 60  
 Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Ser Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Ile Leu Arg Lys Arg Gly Ala Asp Met Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Ile Phe Glu Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg

679 86918 680

130 135 140

Leu Thr  
145

<210> 455  
<211> 146  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Synthetic Protein Sequence

<400> 455  
Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg  
1 5 10 15  
His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Val Cys Met Tyr Glu  
20 25 30  
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Arg Gly  
35 40 45  
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu  
50 55 60  
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
65 70 75 80  
Phe Arg Glu Gln Lys Ala Gly Ser Ser Leu Ile Arg His Ala Glu Gln  
85 90 95  
Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
100 105 110  
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
115 120 125  
Ile Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg  
130 135 140  
Leu Thr  
145

<210> 456  
<211> 146  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Synthetic Protein Sequence

<400> 456  
Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg  
1 5 10 15  
His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Phe Glu  
20 25 30  
Ser Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly  
35 40 45  
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu  
50 55 60  
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
65 70 75 80  
Tyr Arg Asp Gln Lys Ala Gly Ser Ser Leu Ile Arg His Ala Glu Gln  
85 90 95



681

86918

682

```

Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100      105      110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115      120      125
Ile Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
      130      135      140
Leu Thr
145

```

&lt;210&gt; 457

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 457

```

Met Leu Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1      5      10      15
His Lys Ile Leu Arg Pro Asn Gln Pro Leu Glu Val Cys Met Tyr Glu
      20      25      30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Arg Gly
      35      40      45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
      50      55      60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65      70      75      80
Tyr Arg Glu Gln Lys Ala Gly Ser Ser Leu Ile Lys His Ala Glu Glu
      85      90      95
Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100      105      110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115      120      125
Val Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
      130      135      140
Ile Thr
145

```

&lt;210&gt; 458

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 458

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1      5      10      15
His Lys Ile Leu Arg Pro Asn Gln Pro Ile Glu Val Cys Met Tyr Glu
      20      25      30
Ser Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Arg Gly
      35      40      45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Asp Leu

```

683

86918

684

```

      50              55              60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65              70              75              80
Tyr Arg Asp Gln Lys Ala Gly Ser Ser Leu Ile Lys His Ala Glu Gln
      85              90              95
Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100              105              110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115              120              125
Val Phe Glu Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
      130              135              140
Leu Thr
145

```

&lt;210&gt; 459

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 459

```

Met Leu Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
1              5              10              15
His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Phe Glu
      20              25              30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Arg Gly
      35              40              45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu
      50              55              60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65              70              75              80
Phe Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
      85              90              95
Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100              105              110
Ser Ala Ser Gly Tyr Tyr Lys Arg Leu Gly Phe Ser Glu Gln Gly Glu
      115              120              125
Ile Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
      130              135              140
Leu Thr
145

```

&lt;210&gt; 460

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 460

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
1              5              10              15

```

685					86918					686					
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Ile	Glu	Ala	Cys	Met	Phe	Glu
			20					25					30		
Ser	Asp	Leu	Leu	Arg	Gly	Ala	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Glu	Leu
	50					55					60				
Gln	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65				70					75					80	
Tyr	Arg	Asp	Gln	Lys	Ala	Gly	Ser	Ser	Leu	Ile	Arg	His	Ala	Glu	Gln
			85						90					95	
Ile	Leu	Arg	Lys	Arg	Gly	Ala	Asp	Met	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115				120					125				
Val	Phe	Asp	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Arg
	130					135					140				
Ile	Thr														
145															

<210> 461  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 461															
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Leu	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Val	Cys	Met	Tyr	Glu
			20				25					30			
Thr	Asp	Leu	Leu	Arg	Gly	Ala	Phe	His	Leu	Gly	Gly	Phe	Tyr	Gly	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Asp	Leu
	50					55					60				
Gln	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65				70					75					80	
Tyr	Arg	Asp	Gln	Lys	Ala	Gly	Ser	Ser	Leu	Ile	Arg	His	Ala	Glu	Gln
			85						90					95	
Leu	Leu	Arg	Lys	Arg	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115				120					125				
Ile	Phe	Glu	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Arg
	130					135					140				
Ile	Thr														
145															

<210> 462  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>

## &lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 462

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1           5           10           15
His Lys Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu
          20           25           30
Ser Asp Leu Leu Arg Gly Ala Leu His Leu Gly Gly Tyr Tyr Arg Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
          50           55           60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Ser Leu Ile Lys His Ala Glu Glu
          85           90           95
Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
          115          120          125
Val Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
          130          135          140
Ile Thr
145

```

&lt;210&gt; 463

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 463

```

Met Leu Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1           5           10           15
His Lys Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Phe Glu
          20           25           30
Ser Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Gly Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu
          50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
          85           90           95
Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
          115          120          125
Ile Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
          130          135          140
Ile Thr
145

```

&lt;210&gt; 464

<211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 464  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Val Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu  
 50 55 60  
 Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Ser Leu Ile Lys His Ala Glu Gln  
 85 90 95  
 Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg  
 130 135 140  
 Leu Thr  
 145

<210> 465  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 465  
 Met Leu Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Val Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Gly Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu  
 50 55 60  
 Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Asp Gln Lys Ala Gly Ser Ser Leu Ile Lys His Ala Glu Glu  
 85 90 95  
 Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Phe Glu Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg  
 130 135 140

Leu Thr  
145

<210> 466

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 466

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Glu	Thr	Tyr	Glu	Leu	Arg
1				5					10					15	
His	Lys	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Ile	Glu	Ala	Cys	Met	Tyr	Glu
			20					25					30		
Ser	Asp	Leu	Leu	Arg	Gly	Ser	Phe	His	Leu	Gly	Gly	Phe	Tyr	Arg	Gly
		35					40					45			
Gln	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Lys	Ala	Glu	His	Ser	Glu	Leu
	50					55					60				
Gln	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75				80	
Phe	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Ser	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Ile	Leu	Arg	Asn	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Thr	Ala	Ser	Gly	Tyr	Tyr	Lys	Arg	Leu	Gly	Phe	Ser	Glu	His	Gly	Glu
		115					120					125			
Val	Phe	Glu	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Arg
	130						135					140			
Ile	Thr														
145															

<210> 467

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 467

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Leu	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25					30		
Ser	Asp	Leu	Leu	Arg	Gly	Ser	Phe	His	Leu	Gly	Gly	Phe	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75				80	
Phe	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Ser	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Ile	Leu	Arg	Lys	Arg	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr

693					86918					694				
		100					105					110		
Thr	Ala	Ala	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly
		115					120					125		
Ile	Phe	Asp	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys
		130					135					140		
Ile	Thr													
145														

<210> 468

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 468

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25					30		
Ser	Asp	Leu	Leu	Arg	Gly	Ser	Phe	His	Leu	Gly	Gly	Phe	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Asp	Leu
		50				55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75				80	
Tyr	Arg	Asp	Gln	Lys	Ala	Gly	Ser	Ser	Leu	Ile	Arg	His	Ala	Glu	Gln
				85					90					95	
Ile	Leu	Arg	Lys	Arg	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Thr	Ala	Ala	Gly	Tyr	Tyr	Lys	Arg	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Phe	Asp	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
		130					135					140			
Leu	Thr														
145															

<210> 469

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 469

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Leu	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Ile	Glu	Ala	Cys	Met	Tyr	Glu
			20					25					30		
Ser	Asp	Leu	Leu	Arg	Gly	Ser	Phe	His	Leu	Gly	Gly	Phe	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	Gln	Ser	Glu	Leu
		50				55					60				

695

86918

696

```

Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65          70          75          80
Tyr Arg Asp Gln Lys Ala Gly Ser Thr Leu Ile Lys His Ala Glu Glu
          85          90          95
Ile Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ala Gly Tyr Tyr Lys Arg Leu Gly Phe Ser Glu Gln Gly Glu
          115          120          125
Ile Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
          130          135          140
Leu Thr
145

```

&lt;210&gt; 470

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 470

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Glu Thr Tyr Glu Leu Arg
1          5          10          15
His Lys Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
          20          25          30
Thr Asp Leu Leu Arg Gly Ser Phe His Leu Gly Gly Phe Tyr Arg Gly
          35          40          45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
          50          55          60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65          70          75          80
Phe Arg Asp Gln Lys Ala Gly Ser Ser Leu Ile Lys His Ala Glu Glu
          85          90          95
Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
          115          120          125
Ile Phe Glu Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
          130          135          140
Leu Thr
145

```

&lt;210&gt; 471

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 471

```

Met Leu Glu Val Lys Pro Ile Asn Ala Glu Glu Thr Tyr Glu Leu Arg
1          5          10          15
His Lys Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu

```



697					86918					698					
20					25					30					
Thr	Asp	Leu	Leu	Arg	Gly	Ser	Phe	His	Leu	Gly	Gly	Phe	Tyr	Arg	Gly
		35					40					45			
Gln	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Asp	Leu
	50					55					60				
Gln	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75					80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Lys	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Thr	Ala	Ala	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115				120					125				
Val	Phe	Asp	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Ile	Thr														
145															

<210> 472

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 472

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Leu	Arg
1				5					10					15	
His	Lys	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20					25				30			
Ser	Asp	Leu	Leu	Arg	Ser	Ala	Phe	His	Leu	Gly	Gly	Phe	Tyr	Arg	Gly
	35					40						45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Lys	Ala	Glu	His	Ser	Glu	Leu
	50					55					60				
Gln	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75					80
Tyr	Arg	Asp	Gln	Lys	Ala	Gly	Ser	Ser	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Ile	Leu	Arg	Lys	Arg	Gly	Ala	Asp	Met	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ala	Gly	Tyr	Tyr	Lys	Arg	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115				120					125				
Val	Phe	Asp	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Arg
	130					135					140				
Ile	Thr														
145															

<210> 473

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

699

86918

700

&lt;400&gt; 473

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu
          20           25           30
Ser Asp Leu Leu Arg Gly Ser Phe His Leu Gly Gly Phe Tyr Arg Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu
          50           55           60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Ser Leu Ile Lys His Ala Glu Gln
          85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Met Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ala Gly Tyr Tyr Lys Arg Leu Gly Phe Ser Glu His Gly Glu
          115          120          125
Ile Phe Glu Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
          130          135          140
Ile Thr
145

```

&lt;210&gt; 474

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 474

```

Met Leu Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
          20           25           30
Thr Asp Leu Leu Arg Gly Ser Phe His Leu Gly Gly Phe Tyr Arg Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
          50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Asp Gln Lys Ala Gly Ser Ser Leu Ile Arg His Ala Glu Glu
          85           90           95
Ile Leu Arg Lys Arg Gly Ala Asp Met Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Thr Ala Ala Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
          115          120          125
Ile Tyr Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
          130          135          140
Leu Thr
145

```

&lt;210&gt; 475

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 475

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Glu Thr Tyr Glu Leu Arg
 1           5           10           15
His Lys Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
      20           25           30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Arg Gly
      35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Asp His Ser Glu Leu
      50           55           60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Gln
      85           90           95
Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100          105          110
Ser Ala Ala Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu His Gly Glu
      115          120          125
Ile Phe Glu Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
      130          135          140
Leu Thr
145

```

&lt;210&gt; 476

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 476

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu
      20           25           30
Ser Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Arg Gly
      35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu
      50           55           60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Ser Leu Ile Arg His Ala Glu Glu
      85           90           95
Ile Leu Arg Lys Lys Gly Ala Asp Met Leu Trp Cys Asn Ala Arg Thr
      100          105          110
Thr Ala Ala Gly Tyr Tyr Lys Arg Leu Gly Phe Ser Glu Gln Gly Glu
      115          120          125
Val Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
      130          135          140
Ile Thr

```

145

&lt;210&gt; 477

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 477

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1           5           10           15
His Lys Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu
          20           25           30
Ser Asp Leu Leu Gln Gly Ser Phe His Leu Gly Gly Phe Tyr Arg Gly
          35           40           45
Gln Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu
          50           55           60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Phe Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Lys His Ala Glu Glu
          85           90           95
Ile Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ala Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu His Gly Glu
          115          120          125
Ile Phe Asp Thr Pro Pro Ala Gly Pro His Ile Leu Met Tyr Lys Lys
          130          135          140
Leu Thr
145

```

&lt;210&gt; 478

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 478

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Glu Thr Tyr Glu Leu Arg
 1           5           10           15
Gln Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu
          20           25           30
Ser Asp Leu Leu Arg Gly Ser Phe His Leu Gly Gly Phe Tyr Arg Gly
          35           40           45
Gln Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
          50           55           60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Lys His Ala Glu Glu
          85           90           95
Ile Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110

```

705

86918

706

```

Ser Ala Ala Gly Tyr Tyr Lys Arg Leu Gly Phe Ser Glu His Gly Glu
    115          120          125
Ile Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
    130          135          140
Ile Thr
145

```

&lt;210&gt; 479

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 479

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
  1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu
           20           25           30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
    35           40           45
Gln Leu Ile Ser Ile Ala Ser Phe His Lys Ala Glu His Ser Glu Leu
    50           55           60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
    65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Lys His Ala Glu Gln
           85           90           95
Leu Leu Arg Glu Lys Gly Ala Asp Met Leu Trp Cys Asn Ala Arg Thr
           100          105          110
Ser Ala Ala Gly Tyr Tyr Lys Arg Leu Gly Phe Ser Glu Gln Gly Glu
    115          120          125
Val Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
    130          135          140
Leu Thr
145

```

&lt;210&gt; 480

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 480

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
  1           5           10           15
His Lys Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu
           20           25           30
Ser Asp Leu Leu Arg Gly Ser Phe His Leu Gly Gly Phe Tyr Arg Gly
    35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Lys Ala Glu His Ser Asp Leu
    50           55           60
Glu Gly Gln Asn Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly

```

707					86918					708					
65				70				75				80			
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Ile	Leu	Arg	Lys	Arg	Gly	Ala	Asp	Met	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Arg	Leu	Gly	Phe	Ser	Glu	His	Gly	Glu
		115					120					125			
Ile	Phe	Asp	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Arg
	130					135					140				
Leu	Thr														
145															

<210> 481  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 481															
Met	Leu	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Leu	Arg
1				5					10				15		
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Ile	Glu	Ala	Cys	Met	Tyr	Glu
			20				25					30			
Thr	Asp	Leu	Leu	Arg	Gly	Ser	Phe	His	Leu	Gly	Gly	Phe	Tyr	Arg	Gly
		35				40					45				
Gln	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Lys	Ala	Glu	His	Ser	Asp	Leu
	50					55				60					
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65				70					75				80		
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Gln
			85					90					95		
Ile	Leu	Arg	Lys	Arg	Gly	Ala	Asp	Met	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100				105					110			
Ser	Ala	Ala	Gly	Tyr	Tyr	Lys	Arg	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115				120					125				
Val	Phe	Glu	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Arg
	130					135					140				
Leu	Thr														
145															

<210> 482  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 482															
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Leu	Arg
1				5					10				15		
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20				25					30			

709

86918

710

```

Thr Asp Leu Leu Arg Gly Ser Phe His Leu Gly Gly Phe Tyr Arg Gly
      35      40      45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu
      50      55      60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65      70      75      80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Gln
      85      90      95
Leu Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100      105      110
Ser Ala Ser Gly Tyr Tyr Lys Arg Leu Gly Phe Ser Glu His Gly Glu
      115      120      125
Val Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
      130      135      140
Leu Thr
145

```

&lt;210&gt; 483

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 483

```

Met Leu Glu Val Lys Pro Ile Asn Ala Glu Glu Thr Tyr Glu Leu Arg
 1      5      10      15
His Lys Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
      20      25      30
Ser Asp Leu Leu Arg Gly Ser Phe His Leu Gly Gly Tyr Tyr Arg Gly
      35      40      45
Gln Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
      50      55      60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65      70      75      80
Phe Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Lys His Ala Glu Gln
      85      90      95
Ile Leu Arg Lys Arg Gly Ala Asp Met Leu Trp Cys Asn Ala Arg Thr
      100      105      110
Ser Ala Ala Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu His Gly Glu
      115      120      125
Ile Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
      130      135      140
Leu Thr
145

```

&lt;210&gt; 484

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

711

86918

712

&lt;400&gt; 484

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Glu Thr Tyr Glu Leu Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu
          20           25           30
Ser Asp Leu Leu Arg Gly Ser Phe His Leu Gly Gly Phe Tyr Arg Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu
          50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Asp Gln Lys Ala Gly Ser Thr Leu Ile Lys His Ala Glu Glu
          85           90           95
Leu Leu Arg Lys Arg Gly Ala Asp Met Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ala Gly Tyr Tyr Lys Arg Leu Gly Phe Ser Glu His Gly Glu
          115          120          125
Ile Tyr Glu Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
          130          135          140
Ile Thr
145

```

&lt;210&gt; 485

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 485

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1           5           10           15
His Lys Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu
          20           25           30
Ser Asp Leu Leu Arg Gly Ser Phe His Leu Gly Gly Phe Tyr Arg Gly
          35           40           45
Gln Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu
          50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Asp Gln Lys Ala Gly Ser Ser Leu Ile Lys His Ala Glu Glu
          85           90           95
Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ala Gly Tyr Tyr Lys Arg Leu Gly Phe Ser Glu Gln Gly Glu
          115          120          125
Ile Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Lys
          130          135          140
Ile Thr
145

```

&lt;210&gt; 486

&lt;211&gt; 146

&lt;212&gt; PRT



<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 486

```

Met Leu Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu
          20           25           30
Ser Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
          50           55           60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Ser Leu Val Lys His Ala Glu Glu
          85           90           95
Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
          115          120          125
Ile Phe Glu Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
          130          135          140
Ile Thr
145

```

<210> 487

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 487

```

Met Leu Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu
          20           25           30
Ser Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Arg Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
          50           55           60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Ser Leu Ile Lys His Ala Glu Glu
          85           90           95
Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
          115          120          125
Ile Phe Glu Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
          130          135          140
Leu Thr
145

```

<210> 488  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 488  
 Met Leu Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg  
 1 5 10 15  
 His Lys Ile Leu Arg Pro Asn Gln Pro Leu Glu Val Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Gly Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu  
 50 55 60  
 Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Ser Leu Ile Lys His Ala Glu Glu  
 85 90 95  
 Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Ile Phe Glu Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg  
 130 135 140  
 Ile Thr  
 145

<210> 489  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 489  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Ala Cys Met Tyr Glu  
 20 25 30  
 Ser Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu  
 50 55 60  
 Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Ser Leu Ile Lys His Ala Glu Glu  
 85 90 95  
 Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu

	717		86918		718										
	115		120		125										
Ile	Phe	Glu	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Arg
	130					135					140				
Ile	Thr														
145															

<210> 490  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 490

Met	Leu	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Leu	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Ile	Glu	Ala	Cys	Met	Tyr	Glu
			20					25				30			
Ser	Asp	Leu	Leu	Arg	Gly	Ala	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
	50					55					60				
Gln	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75				80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Ser	Leu	Ile	Lys	His	Ala	Glu	Glu
				85					90					95	
Ile	Leu	Arg	Lys	Arg	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Phe	Glu	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Arg
	130					135					140				
Ile	Thr														
145															

<210> 491  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 491

Met	Leu	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Leu	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Ile	Glu	Ala	Cys	Met	Tyr	Glu
			20					25				30			
Ser	Asp	Leu	Leu	Arg	Gly	Ala	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Gly	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Asp	Leu
	50					55					60				
Gln	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75				80	

```

Tyr Arg Glu Gln Lys Ala Gly Ser Ser Leu Ile Lys His Ala Glu Glu
                85                90                95
Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
                100                105                110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
                115                120                125
Ile Phe Glu Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
                130                135                140
Ile Thr
145

```

&lt;210&gt; 492

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 492

```

Met Leu Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
  1                5                10                15
His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Val Cys Met Tyr Glu
                20                25                30
Ser Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
                35                40                45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
                50                55                60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
        65                70                75                80
Tyr Arg Glu Gln Lys Ala Gly Ser Ser Leu Ile Lys His Ala Glu Glu
                85                90                95
Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
                100                105                110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
                115                120                125
Ile Phe Glu Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
                130                135                140
Ile Thr
145

```

&lt;210&gt; 493

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 493

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
  1                5                10                15
His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Val Cys Met Tyr Glu
                20                25                30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Gly Gly

```

721					86918					722				
	35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Asp
	50					55					60			
Gln	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu
65						70					75			80
Tyr	Arg	Asp	Gln	Lys	Ala	Gly	Ser	Ser	Leu	Ile	Arg	His	Ala	Glu
				85					90					95
Ile	Leu	Arg	Lys	Arg	Gly	Ala	Asp	Met	Leu	Trp	Cys	Asn	Ala	Arg
			100					105					110	
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly
		115					120					125		
Ile	Phe	Glu	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys
	130					135					140			
Ile	Thr													
145														

<210> 494

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 494

Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Leu	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Val	Cys	Met	Tyr	Glu
			20				25					30			
Thr	Asp	Leu	Leu	Arg	Gly	Ala	Phe	His	Leu	Gly	Gly	Phe	Tyr	Gly	Gly
		35				40					45				
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Asp	Leu
	50					55				60					
Gln	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75				80	
Tyr	Arg	Asp	Gln	Lys	Ala	Gly	Ser	Ser	Leu	Ile	Arg	His	Ala	Glu	Gln
				85					90					95	
Ile	Leu	Arg	Lys	Arg	Gly	Ala	Asp	Met	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Ile	Phe	Glu	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Arg
	130					135					140				
Ile	Thr														
145															

<210> 495

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<221> unsure

&lt;222&gt; 3

&lt;223&gt; Xaa = Asp or Glu

&lt;221&gt; unsure

&lt;222&gt; 26

&lt;223&gt; Xaa = Leu, Ile, or Val

&lt;221&gt; unsure

&lt;222&gt; 33

&lt;223&gt; Xaa = Ile, Thr, Asn, or Ser

&lt;400&gt; 495

Met	Leu	Xaa	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Leu	Arg
1				5					10					15	
His	Lys	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Xaa	Glu	Val	Cys	Met	Tyr	Glu
			20					25					30		
Xaa	Asp	Leu	Leu	Arg	Gly	Ala	Phe	His	Leu	Gly	Gly	Phe	Tyr	Arg	Gly
		35				40						45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Asp	Leu
	50					55					60				
Gln	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65				70						75					80
Tyr	Arg	Asp	Gln	Lys	Ala	Gly	Ser	Ser	Leu	Ile	Lys	His	Ala	Glu	Gln
			85						90					95	
Ile	Leu	Arg	Glu	Arg	Gly	Ala	Asp	Met	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Phe	Asp	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Arg
	130					135					140				
Leu	Thr														
145															

&lt;210&gt; 496

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 496

Met	Leu	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Leu	Arg
1				5					10					15	
His	Lys	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Ile	Glu	Val	Cys	Met	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Arg	Gly	Ala	Phe	His	Leu	Gly	Gly	Phe	Tyr	Gly	Gly
		35				40						45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Asp	Leu
	50					55					60				
Gln	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65				70						75					80
Tyr	Arg	Asp	Gln	Lys	Ala	Gly	Ser	Ser	Leu	Ile	Arg	His	Ala	Glu	Gln
			85						90					95	
Ile	Leu	Arg	Lys	Arg	Gly	Ala	Asp	Met	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		

725

86918

726

Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Ile Phe Glu Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg  
 130 135 140  
 Leu Thr  
 145

&lt;210&gt; 497

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 497

Met Leu Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg  
 1 5 10 15  
 His Lys Ile Leu Arg Pro Asn Gln Pro Leu Glu Val Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu  
 50 55 60  
 Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Asp Gln Lys Ala Gly Ser Ser Leu Ile Arg His Ala Glu Gln  
 85 90 95  
 Ile Leu Arg Lys Arg Gly Ala Asp Met Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Arg Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Phe Glu Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg  
 130 135 140  
 Leu Thr  
 145

&lt;210&gt; 498

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 498

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg  
 1 5 10 15  
 His Lys Ile Leu Arg Pro Asn Gln Pro Leu Glu Val Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu  
 50 55 60  
 Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly

727					86918					728					
65				70				75				80			
Tyr	Arg	Asp	Gln	Lys	Ala	Gly	Ser	Ser	Leu	Ile	Arg	His	Ala	Glu	Gln
				85				90						95	
Ile	Leu	Arg	Lys	Arg	Gly	Ala	Asp	Met	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Phe	Glu	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Arg
	130					135					140				
Ile	Thr														
145															

<210> 499  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 499															
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Leu	Arg
1				5					10					15	
His	Lys	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Val	Cys	Met	Tyr	Glu
			20					25					30		
Thr	Asp	Leu	Leu	Arg	Gly	Ala	Phe	His	Leu	Gly	Gly	Phe	Tyr	Arg	Gly
		35				40						45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Asp	Leu
	50					55					60				
Gln	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75				80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Ser	Leu	Ile	Arg	His	Ala	Glu	Gln
				85					90					95	
Ile	Leu	Arg	Lys	Arg	Gly	Ala	Asp	Met	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Phe	Glu	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Arg
	130					135					140				
Leu	Thr														
145															

<210> 500  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 500															
Met	Leu	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Leu	Arg
1				5					10					15	
His	Lys	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Val	Cys	Met	Tyr	Glu
			20					25					30		



729					86918					730					
Thr	Asp	Leu	Leu	Arg	Gly	Ala	Phe	His	Leu	Gly	Gly	Phe	Tyr	Arg	Gly
		35					40					45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Glu	Leu
	50				55						60				
Gln	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65				70						75				80	
Tyr	Arg	Asp	Gln	Lys	Ala	Gly	Ser	Ser	Leu	Ile	Arg	His	Ala	Glu	Gln
			85						90					95	
Ile	Leu	Arg	Lys	Arg	Gly	Ala	Asp	Met	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Phe	Glu	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Arg
	130					135					140				
Leu	Thr														
145															

<210> 501  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 501															
Met	Leu	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Leu	Arg
1				5					10				15		
His	Lys	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Val	Cys	Met	Tyr	Glu
			20				25					30			
Thr	Asp	Leu	Leu	Arg	Gly	Ala	Phe	His	Leu	Gly	Gly	Phe	Tyr	Gly	Gly
	35					40						45			
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Asp	Leu
	50				55						60				
Gln	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65				70						75				80	
Tyr	Arg	Asp	Gln	Lys	Ala	Gly	Ser	Ser	Leu	Ile	Arg	His	Ala	Glu	Glu
			85						90					95	
Ile	Leu	Arg	Lys	Arg	Gly	Ala	Asp	Met	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Ala	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115					120					125			
Val	Phe	Asp	Thr	Pro	Pro	Val	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Ile	Thr														
145															

<210> 502  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

731

86918

732

&lt;400&gt; 502

```

Met Leu Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1           5           10           15
His Lys Ile Leu Arg Pro Asn Gln Pro Leu Glu Val Cys Met Tyr Glu
      20           25           30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Gly Gly
      35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu
      50           55           60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Asp Gln Lys Ala Gly Ser Thr Leu Ile Lys His Ala Glu Gln
      85           90           95
Ile Leu Arg Lys Arg Gly Ala Asp Met Leu Trp Cys Asn Ala Arg Thr
      100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115          120          125
Val Phe Glu Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
      130          135          140
Leu Thr
145

```

&lt;210&gt; 503

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 503

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Val Cys Met Tyr Glu
      20           25           30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Arg Gly
      35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu
      50           55           60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Ser Leu Ile Lys His Ala Glu Glu
      85           90           95
Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
      100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
      115          120          125
Ile Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
      130          135          140
Ile Thr
145

```

&lt;210&gt; 504

&lt;211&gt; 146

&lt;212&gt; PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 504

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1           5           10           15
His Arg Val Leu Arg Pro Asn Gln Pro Leu Glu Val Cys Met Tyr Glu
          20           25           30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
          50           55           60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Ser Leu Ile Lys His Ala Glu Glu
          85           90           95
Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
          115          120          125
Val Phe Glu Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
          130          135          140
Leu Thr
145

```

<210> 505

<211> 146

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Protein Sequence

<400> 505

```

Met Leu Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg
 1           5           10           15
His Lys Ile Leu Arg Pro Asn Gln Pro Ile Glu Val Cys Met Tyr Glu
          20           25           30
Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
          50           55           60
Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Ser Leu Ile Arg His Ala Glu Glu
          85           90           95
Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu
          115          120          125
Ile Phe Asp Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg
          130          135          140
Ile Thr
145

```

<210> 506  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 506  
 Met Leu Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Ile Glu Val Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Phe Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu  
 50 55 60  
 Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Ser Leu Ile Lys His Ala Glu Glu  
 85 90 95  
 Ile Leu Arg Lys Arg Gly Ala Asp Met Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Ile Phe Glu Thr Pro Pro Val Gly Pro His Ile Leu Met Tyr Lys Arg  
 130 135 140  
 Leu Thr  
 145

<210> 507  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 507  
 Met Leu Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Leu Arg  
 1 5 10 15  
 His Lys Ile Leu Arg Pro Asn Gln Pro Leu Glu Val Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Arg Gly Ala Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Asp Leu  
 50 55 60  
 Gln Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Ser Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Ile Leu Arg Lys Arg Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Ala Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu

	737		86918		738
	115		120		125
Val	Phe Asp Thr Pro Pro	Val	Gly Pro His Ile	Leu	Met Tyr Lys Arg
	130		135		140
Leu	Thr				
	145				

<210> 508  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 508  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Lys Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Asn Ala Glu His Ser Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80  
 Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu  
 85 90 95  
 Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr  
 100 105 110  
 Ser Val Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Glu  
 115 120 125  
 Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys  
 130 135 140  
 Leu Thr  
 145

<210> 509  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 509  
 Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg  
 1 5 10 15  
 His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu  
 20 25 30  
 Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly  
 35 40 45  
 Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu  
 50 55 60  
 Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly  
 65 70 75 80

739					86918					740					
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr	
			100				105					110			
Phe	Val	Ser	Gly	Tyr	Tyr	Glu	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
		115				120						125			
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	Tyr	Ile	Leu	Met	Tyr	Glu	Lys
	130					135					140				
Leu	Thr														
145															

<210> 510  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 510															
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20				25					30			
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35				40					45				
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Ser	Glu	Leu
	50					55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65				70					75					80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
			85						90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100				105					110			
Ser	Val	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Gly
		115				120						125			
Val	Cys	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
	130					135					140				
Leu	Ala														
145															

<210> 511  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 511															
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Met	Tyr	Glu
			20				25					30			
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly

741					86918					742					
35					40					45					
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Gln	Ala	Glu	His	Pro	Glu	Leu
50					55					60					
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70					75					80
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
				85					90					95	
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100					105					110		
Ser	Val	Ser	Gly	Tyr	Tyr	Glu	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
			115				120					125			
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
130					135					140					
Leu	Thr														
145															

<210> 512  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 512															
Met	Ile	Glu	Val	Lys	Pro	Ile	Asn	Ala	Glu	Asp	Thr	Tyr	Glu	Ile	Arg
1				5					10					15	
His	Arg	Ile	Leu	Arg	Pro	Asn	Gln	Pro	Leu	Glu	Ala	Cys	Lys	Tyr	Glu
			20				25					30			
Thr	Asp	Leu	Leu	Gly	Gly	Thr	Phe	His	Leu	Gly	Gly	Tyr	Tyr	Arg	Gly
		35				40					45				
Lys	Leu	Ile	Ser	Ile	Ala	Ser	Phe	His	Asn	Ala	Glu	His	Ser	Glu	Leu
50						55					60				
Glu	Gly	Gln	Lys	Gln	Tyr	Gln	Leu	Arg	Gly	Met	Ala	Thr	Leu	Glu	Gly
65					70				75					80	
Tyr	Arg	Glu	Gln	Lys	Ala	Gly	Ser	Thr	Leu	Ile	Arg	His	Ala	Glu	Glu
			85					90					95		
Leu	Leu	Arg	Lys	Lys	Gly	Ala	Asp	Leu	Leu	Trp	Cys	Asn	Ala	Arg	Thr
			100				105						110		
Ser	Val	Ser	Gly	Tyr	Tyr	Lys	Lys	Leu	Gly	Phe	Ser	Glu	Gln	Gly	Glu
			115			120						125			
Val	Tyr	Asp	Ile	Pro	Pro	Ile	Gly	Pro	His	Ile	Leu	Met	Tyr	Lys	Lys
130						135					140				
Leu	Thr														
145															

<210> 513  
 <211> 146  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Protein Sequence

<400> 513

743

86918

744

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
          20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Pro Glu Leu
 50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
          85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Val Ser Gly Tyr Tyr Glu Lys Leu Gly Phe Ser Glu Gln Gly Glu
          115          120          125
Val Cys Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
 130           135           140
Leu Thr
145

```

&lt;210&gt; 514

&lt;211&gt; 146

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Synthetic Protein Sequence

&lt;400&gt; 514

```

Met Ile Glu Val Lys Pro Ile Asn Ala Glu Asp Thr Tyr Glu Ile Arg
 1           5           10           15
His Arg Ile Leu Arg Pro Asn Gln Pro Leu Glu Ala Cys Met Tyr Glu
          20           25           30
Thr Asp Leu Leu Gly Gly Thr Phe His Leu Gly Gly Tyr Tyr Arg Gly
          35           40           45
Lys Leu Ile Ser Ile Ala Ser Phe His Gln Ala Glu His Ser Glu Leu
 50           55           60
Glu Gly Gln Lys Gln Tyr Gln Leu Arg Gly Met Ala Thr Leu Glu Gly
65           70           75           80
Tyr Arg Glu Gln Lys Ala Gly Ser Thr Leu Ile Arg His Ala Glu Glu
          85           90           95
Leu Leu Arg Lys Lys Gly Ala Asp Leu Leu Trp Cys Asn Ala Arg Thr
          100          105          110
Ser Val Ser Gly Tyr Tyr Lys Lys Leu Gly Phe Ser Glu Gln Gly Gly
          115          120          125
Val Tyr Asp Ile Pro Pro Ile Gly Pro His Ile Leu Met Tyr Lys Lys
 130           135           140
Leu Thr
145

```

&lt;210&gt; 515

&lt;211&gt; 18

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence



&lt;220&gt;

&lt;223&gt; Synthetic DNA Sequence

&lt;400&gt; 515

aactgaagga ggaatctc

18

N-ацетилювання гліфосату

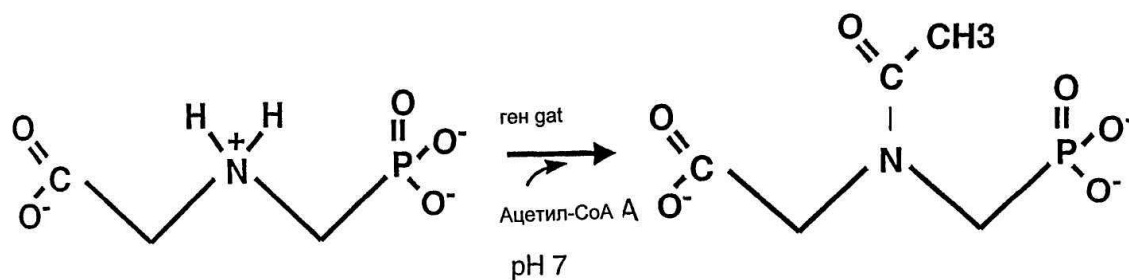
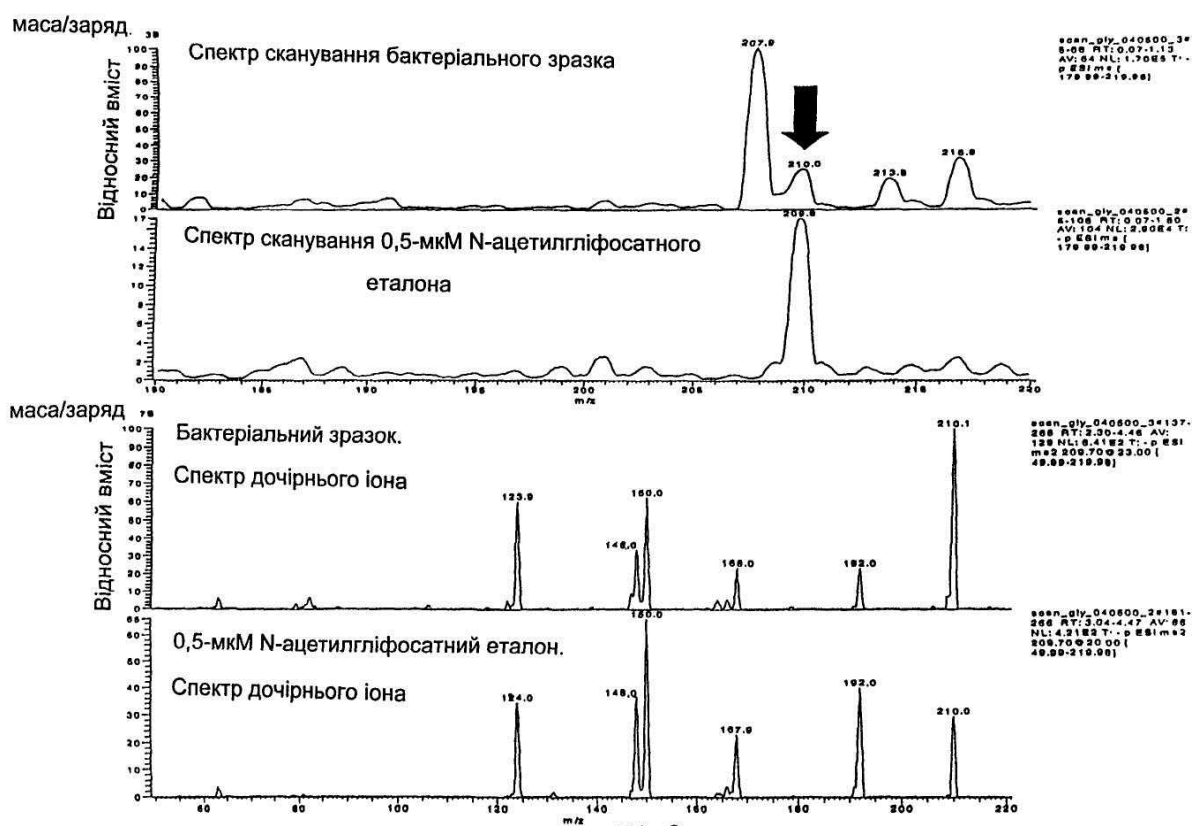


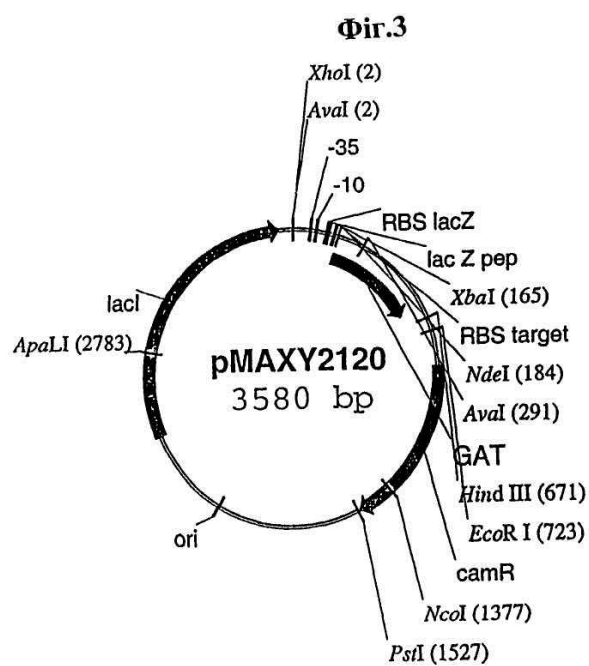
Fig.1

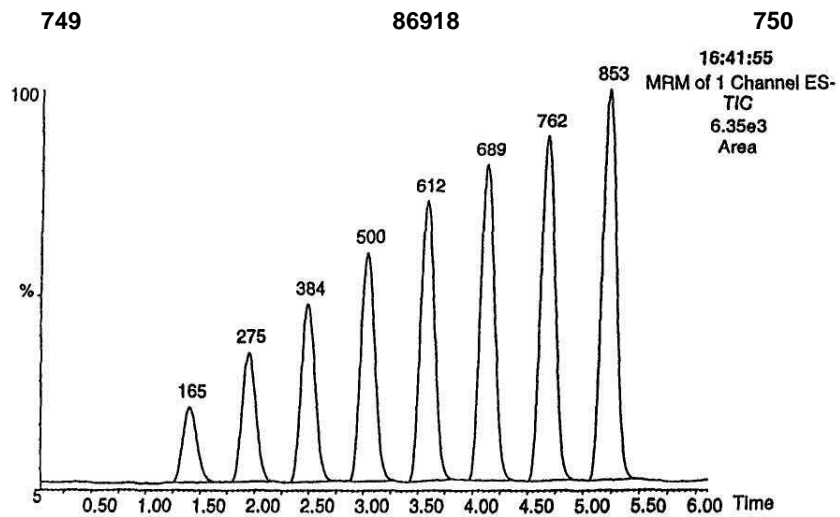
Продукт реакції, ідентифікований як N-ацетилгліфосат



Нуклеотидна ідентичність між парами послідовностей

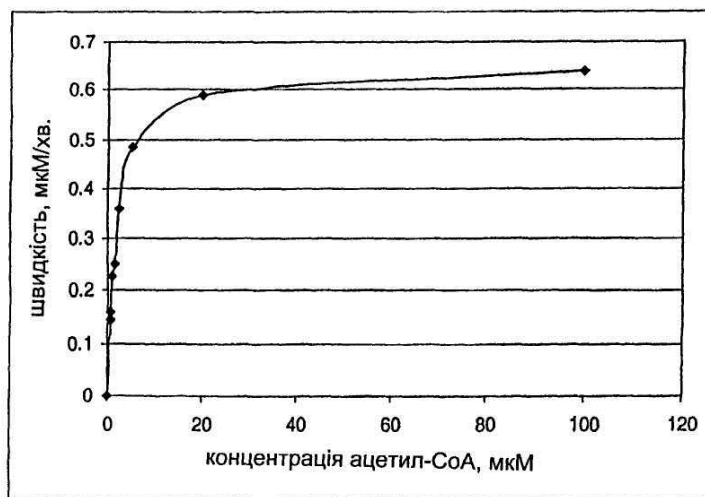
	B6	DS3	ST401	NHA-2	NH5-2	yitI
B6	***	95.0	93.2	94.8	92.8	62.0
DS3		***	95.5	99.8	95.0	61.8
ST401			***	95.2	99.5	62.4
NHA-2				***	94.8	61.5
NH5-2					***	62.9
yitI						***





Очищений B6 GAT каталізує вироблення N-ацетилгліфосату

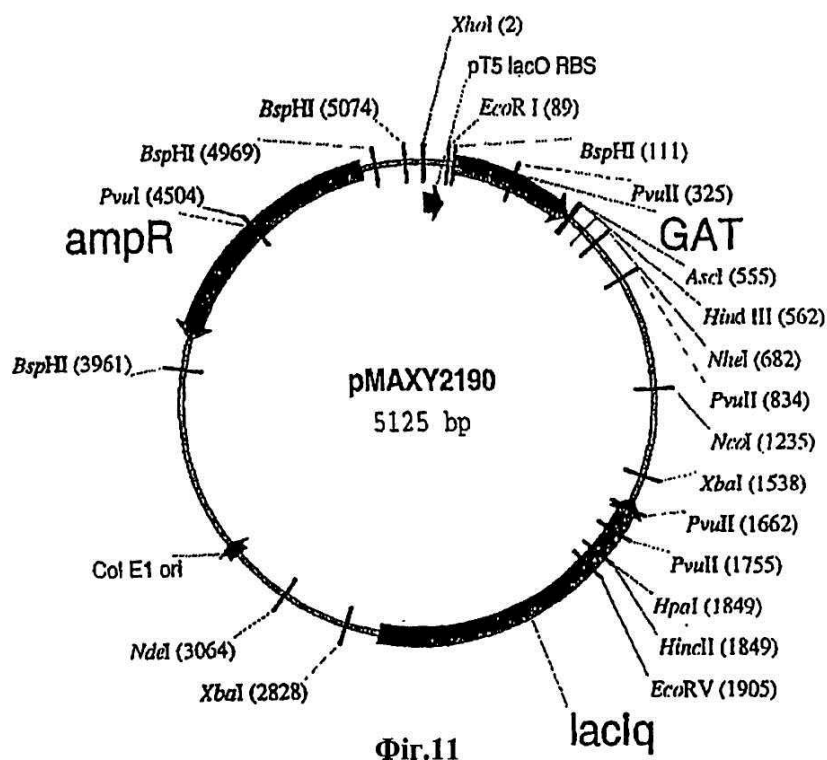
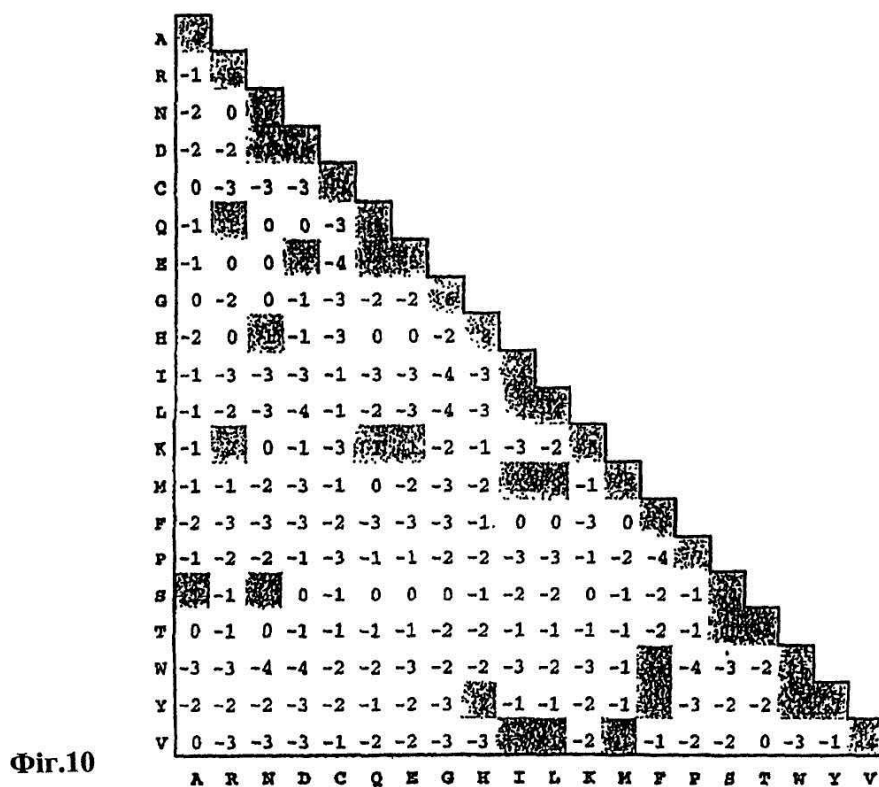
Fig.5

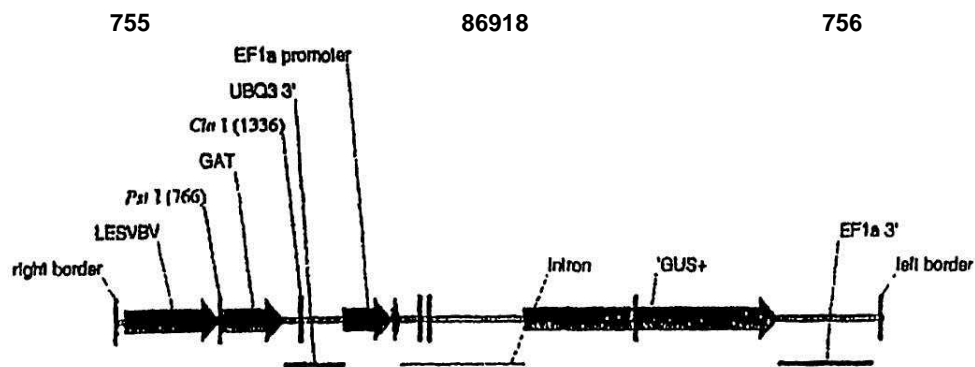


$K_M$  GAT B6 до ацетил-коферменту A = 2 мкМ

Fig.6







**GAT T-DNA**  
5546 bp

