

SEQUENCE LISTING

<110> DSM IP Assets B.V.

Raamsdonk, L.M.

Mueller, U.

Wu, L.

<120> Butanol production in eukaryotic cell

<130> 26265WO

<140> 26265WO

<141> 2008-07-11

<160> 21

<170> PatentIn version 3.2

<210> 1

<211> 24

<212> DNA

<213> artificial sequence

<220>

<223> Primer p1

<400> 1

gaattgaagg atatctacat caag

24

<210> 2

<211> 25

<212> DNA

<213> Artificial sequence

<220>

<223> primer P2

<400> 2

cccattctacg gaaccctgat caagc

25

<210> 3

<211> 26

<212> DNA

<213> Artificial sequence

<220>

<223> Primer P3

<400> 3

gatggtgtca ccattaccag gtctag

26

<210> 4

<211> 56

<212> DNA

<213> Artificial sequence

<220>

<223> Primer P4

<400> 4

gttctctggt caagttgaag tccatttga ttgatttgac tgtgttattt tgcgtg

56

<210> 5

<211> 25

<212> DNA

<213> Artificial sequence

<220>

<223> Primer P5

<400> 5

gaacaataga gcgaccatga ccttg

25

<210> 6

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer P6

<400> 6

gacatcagcg tcaccagcct tgatg

25

<210> 7

<211> 27

<212> DNA

<213> Artificial sequence

<220>

<223> Primer P7

<400> 7

gattgaaggt ttcaagaaca ggtgatg

27

<210> 8

<211> 25

<212> DNA

<213> Artificial sequence

<220>

<223> Primer P8

<400> 8

ggcgatcaga gttgaaaaaa aaatg

25

<210> 9

<211> 392

<212> PRT

<213> Clostridium acetobutylicum

<400> 9

Met Lys Glu Val Val Ile Ala Ser Ala Val Arg Thr Ala Ile Gly Ser

1 5 10 15

Tyr Gly Lys Ser Leu Lys Asp Val Pro Ala Val Asp Leu Gly Ala Thr

20 25 30

Ala Ile Lys Glu Ala Val Lys Lys Ala Gly Ile Lys Pro Glu Asp Val

35 40 45

Asn Glu Val Ile Leu Gly Asn Val Leu Gln Ala Gly Leu Gly Gln Asn

50 55 60

Pro Ala Arg Gln Ala Ser Phe Lys Ala Gly Leu Pro Val Glu Ile Pro

65 70 75 80

Ala Met Thr Ile Asn Lys Val Cys Gly Ser Gly Leu Arg Thr Val Ser

85 90 95

Leu Ala Ala Gln Ile Ile Lys Ala Gly Asp Ala Asp Val Ile Ile Ala
100 105 110

Gly Gly Met Glu Asn Met Ser Arg Ala Pro Tyr Leu Ala Asn Asn Ala
115 120 125

Arg Trp Gly Tyr Arg Met Gly Asn Ala Lys Phe Val Asp Glu Met Ile
130 135 140

Thr Asp Gly Leu Trp Asp Ala Phe Asn Asp Tyr His Met Gly Ile Thr
145 150 155 160

Ala Glu Asn Ile Ala Glu Arg Trp Asn Ile Ser Arg Glu Glu Gln Asp
165 170 175

Glu Phe Ala Leu Ala Ser Gln Lys Lys Ala Glu Glu Ala Ile Lys Ser
180 185 190

Gly Gln Phe Lys Asp Glu Ile Val Pro Val Val Ile Lys Gly Arg Lys
195 200 205

Gly Glu Thr Val Val Asp Thr Asp Glu His Pro Arg Phe Gly Ser Thr
210 215 220

Ile Glu Gly Leu Ala Lys Leu Lys Pro Ala Phe Lys Lys Asp Gly Thr
225 230 235 240

Val Thr Ala Gly Asn Ala Ser Gly Leu Asn Asp Cys Ala Ala Val Leu
245 250 255

Val Ile Met Ser Ala Glu Lys Ala Lys Glu Leu Gly Val Lys Pro Leu
260 265 270

Ala Lys Ile Val Ser Tyr Gly Ser Ala Gly Val Asp Pro Ala Ile Met
275 280 285

Gly Tyr Gly Pro Phe Tyr Ala Thr Lys Ala Ala Ile Glu Lys Ala Gly
290 295 300

Trp Thr Val Asp Glu Leu Asp Leu Ile Glu Ser Asn Glu Ala Phe Ala
305 310 315 320

Ala Gln Ser Leu Ala Val Ala Lys Asp Leu Lys Phe Asp Met Asn Lys
325 330 335

Val Asn Val Asn Gly Gly Ala Ile Ala Leu Gly His Pro Ile Gly Ala
340 345 350

Ser Gly Ala Arg Ile Leu Val Thr Leu Val His Ala Met Gln Lys Arg
355 360 365

Asp Ala Lys Lys Gly Leu Ala Thr Leu Cys Ile Gly Gly Gly Gln Gly
370 375 380

Thr Ala Ile Leu Leu Glu Lys Cys
385 390

<210> 10
<211> 282
<212> PRT
<213> Clostridium acetobutylicum

<400> 10

Met Lys Lys Val Cys Val Ile Gly Ala Gly Thr Met Gly Ser Gly Ile
1 5 10 15

Ala Gln Ala Phe Ala Ala Lys Gly Phe Glu Val Val Leu Arg Asp Ile
20 25 30

Lys Asp Glu Phe Val Asp Arg Gly Leu Asp Phe Ile Asn Lys Asn Leu
35 40 45

Ser Lys Leu Val Lys Lys Gly Lys Ile Glu Glu Ala Thr Lys Val Glu
50 55 60

Ile Leu Thr Arg Ile Ser Gly Thr Val Asp Leu Asn Met Ala Ala Asp
65 70 75 80

Cys Asp Leu Val Ile Glu Ala Ala Val Glu Arg Met Asp Ile Lys Lys
85 90 95

Gln Ile Phe Ala Asp Leu Asp Asn Ile Cys Lys Pro Glu Thr Ile Leu
100 105 110

Ala Ser Asn Thr Ser Ser Leu Ser Ile Thr Glu Val Ala Ser Ala Thr
115 120 125

Lys Arg Pro Asp Lys Val Ile Gly Met His Phe Phe Asn Pro Ala Pro
130 135 140

Val Met Lys Leu Val Glu Val Ile Arg Gly Ile Ala Thr Ser Gln Glu
145 150 155 160

Thr Phe Asp Ala Val Lys Glu Thr Ser Ile Ala Ile Gly Lys Asp Pro
165 170 175

Val Glu Val Ala Glu Ala Pro Gly Phe Val Val Asn Arg Ile Leu Ile
180 185 190

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

Val Glu Asp Ile Asp Lys Ala Met Lys Leu Gly Ala Asn His Pro Met
210 215 220

Gly Pro Leu Glu Leu Gly Asp Phe Ile Gly Leu Asp Ile Cys Leu Ala
225 230 235 240

Ile Met Asp Val Leu Tyr Ser Glu Thr Gly Asp Ser Lys Tyr Arg Pro
245 250 255

His Thr Leu Leu Lys Lys Tyr Val Arg Ala Gly Trp Leu Gly Arg Lys
260 265 270

Ser Gly Lys Gly Phe Tyr Asp Tyr Ser Lys
275 280

<210> 11

<211> 261

<212> PRT

<213> Clostridium acetobutylicum

<400> 11

Met Glu Leu Asn Asn Val Ile Leu Glu Lys Glu Gly Lys Val Ala Val
1 5 10 15

Val Thr Ile Asn Arg Pro Lys Ala Leu Asn Ala Leu Asn Ser Asp Thr
20 25 30

Leu Lys Glu Met Asp Tyr Val Ile Gly Glu Ile Glu Asn Asp Ser Glu
35 40 45

Val Leu Ala Val Ile Leu Thr Gly Ala Gly Glu Lys Ser Phe Val Ala
50 55 60

Gly Ala Asp Ile Ser Glu Met Lys Glu Met Asn Thr Ile Glu Gly Arg
65 70 75 80

Lys Phe Gly Ile Leu Gly Asn Lys Val Phe Arg Arg Leu Glu Leu Leu
85 90 95

Glu Lys Pro Val Ile Ala Ala Val Asn Gly Phe Ala Leu Gly Gly Gly
100 105 110

Cys Glu Ile Ala Met Ser Cys Asp Ile Arg Ile Ala Ser Ser Asn Ala

115 120 125

Arg Phe Gly Gln Pro Glu Val Gly Leu Gly Ile Thr Pro Gly Phe Gly
130 135 140

Gly Thr Gln Arg Leu Ser Arg Leu Val Gly Met Gly Met Ala Lys Gln
145 150 155 160

Leu Ile Phe Thr Ala Gln Asn Ile Lys Ala Asp Glu Ala Leu Arg Ile
165 170 175

Gly Leu Val Asn Lys Val Val Glu Pro Ser Glu Leu Met Asn Thr Ala
180 185 190

Lys Glu Ile Ala Asn Lys Ile Val Ser Asn Ala Pro Val Ala Val Lys
195 200 205

Leu Ser Lys Gln Ala Ile Asn Arg Gly Met Gln Cys Asp Ile Asp Thr
210 215 220

Ala Leu Ala Phe Glu Ser Glu Ala Phe Gly Glu Cys Phe Ser Thr Glu
225 230 235 240

Asp Gln Lys Asp Ala Met Thr Ala Phe Ile Glu Lys Arg Lys Ile Glu
245 250 255

Gly Phe Lys Asn Arg
260

<210> 12

<211> 379

<212> PRT

<213> Clostridium acetobutylicum

<400> 12

Met Asp Phe Asn Leu Thr Arg Glu Gln Glu Leu Val Arg Gln Met Val
1 5 10 15

Arg Glu Phe Ala Glu Asn Glu Val Lys Pro Ile Ala Ala Glu Ile Asp
20 25 30

Glu Thr Glu Arg Phe Pro Met Glu Asn Val Lys Lys Met Gly Gln Tyr
35 40 45

Gly Met Met Gly Ile Pro Phe Ser Lys Glu Tyr Gly Gly Ala Gly Gly
50 55 60

Asp Val Leu Ser Tyr Ile Ile Ala Val Glu Glu Leu Ser Lys Val Cys
65 70 75 80

Gly Thr Thr Gly Val Ile Leu Ser Ala His Thr Ser Leu Cys Ala Ser
85 90 95

Leu Ile Asn Glu His Gly Thr Glu Glu Gln Lys Gln Lys Tyr Leu Val
100 105 110

Pro Leu Ala Lys Gly Glu Lys Ile Gly Ala Tyr Gly Leu Thr Glu Pro
115 120 125

Asn Ala Gly Thr Asp Ser Gly Ala Gln Gln Thr Val Ala Val Leu Glu
130 135 140

Gly Asp His Tyr Val Ile Asn Gly Ser Lys Ile Phe Ile Thr Asn Gly
145 150 155 160

Gly Val Ala Asp Thr Phe Val Ile Phe Ala Met Thr Asp Arg Thr Lys
165 170 175

Gly Thr Lys Gly Ile Ser Ala Phe Ile Ile Glu Lys Gly Phe Lys Gly
180 185 190

Phe Ser Ile Gly Lys Val Glu Gln Lys Leu Gly Ile Arg Ala Ser Ser
195 200 205

Thr Thr Glu Leu Val Phe Glu Asp Met Ile Val Pro Val Glu Asn Met
210 215 220

Ile Gly Lys Glu Gly Lys Gly Phe Pro Ile Ala Met Lys Thr Leu Asp
225 230 235 240

Gly Gly Arg Ile Gly Ile Ala Ala Gln Ala Leu Gly Ile Ala Glu Gly
245 250 255

Ala Phe Asn Glu Ala Arg Ala Tyr Met Lys Glu Arg Lys Gln Phe Gly
260 265 270

Arg Ser Leu Asp Lys Phe Gln Gly Leu Ala Trp Met Met Ala Asp Met
275 280 285

Asp Val Ala Ile Glu Ser Ala Arg Tyr Leu Val Tyr Lys Ala Ala Tyr
290 295 300

Leu Lys Gln Ala Gly Leu Pro Tyr Thr Val Asp Ala Ala Arg Ala Lys
305 310 315 320

Leu His Ala Ala Asn Val Ala Met Asp Val Thr Thr Lys Ala Val Gln
325 330 335

Leu Phe Gly Gly Tyr Gly Tyr Thr Lys Asp Tyr Pro Val Glu Arg Met
340 345 350

Met Arg Asp Ala Lys Ile Thr Glu Ile Tyr Glu Gly Thr Ser Glu Val
355 360 365

Gln Lys Leu Val Ile Ser Gly Lys Ile Phe Arg
370 375

<210> 13
<211> 858

<212> PRT

<213> Clostridium acetobutylicum

<400> 13

Met Lys Val Thr Asn Gln Lys Glu Leu Lys Gln Lys Leu Asn Glu Leu
1 5 10 15

Arg Glu Ala Gln Lys Lys Phe Ala Thr Tyr Thr Gln Glu Gln Val Asp
20 25 30

Lys Ile Phe Lys Gln Cys Ala Ile Ala Ala Ala Lys Glu Arg Ile Asn
35 40 45

Leu Ala Lys Leu Ala Val Glu Glu Thr Gly Ile Gly Leu Val Glu Asp
50 55 60

Lys Ile Ile Lys Asn His Phe Ala Ala Glu Tyr Ile Tyr Asn Lys Tyr
65 70 75 80

Lys Asn Glu Lys Thr Cys Gly Ile Ile Asp His Asp Asp Ser Leu Gly
85 90 95

Ile Thr Lys Val Ala Glu Pro Ile Gly Ile Val Ala Ala Ile Val Pro
100 105 110

Thr Thr Asn Pro Thr Ser Thr Ala Ile Phe Lys Ser Leu Ile Ser Leu
115 120 125

Lys Thr Arg Asn Ala Ile Phe Phe Ser Pro His Pro Arg Ala Lys Lys
130 135 140

Ser Thr Ile Ala Ala Ala Lys Leu Ile Leu Asp Ala Ala Val Lys Ala
145 150 155 160

Gly Ala Pro Lys Asn Ile Ile Gly Trp Ile Asp Glu Pro Ser Ile Glu
165 170 175

Leu Ser Gln Asp Leu Met Ser Glu Ala Asp Ile Ile Leu Ala Thr Gly
180 185 190

Gly Pro Ser Met Val Lys Ala Ala Tyr Ser Ser Gly Lys Pro Ala Ile
195 200 205

Gly Val Gly Ala Gly Asn Thr Pro Ala Ile Ile Asp Glu Ser Ala Asp
210 215 220

Ile Asp Met Ala Val Ser Ser Ile Ile Leu Ser Lys Thr Tyr Asp Asn
225 230 235 240

Gly Val Ile Cys Ala Ser Glu Gln Ser Ile Leu Val Met Asn Ser Ile
245 250 255

Tyr Glu Lys Val Lys Glu Glu Phe Val Lys Arg Gly Ser Tyr Ile Leu
260 265 270

Asn Gln Asn Glu Ile Ala Lys Ile Lys Glu Thr Met Phe Lys Asn Gly
275 280 285

Ala Ile Asn Ala Asp Ile Val Gly Lys Ser Ala Tyr Ile Ile Ala Lys
290 295 300

Met Ala Gly Ile Glu Val Pro Gln Thr Thr Lys Ile Leu Ile Gly Glu
305 310 315 320

Val Gln Ser Val Glu Lys Ser Glu Leu Phe Ser His Glu Lys Leu Ser
325 330 335

Pro Val Leu Ala Met Tyr Lys Val Lys Asp Phe Asp Glu Ala Leu Lys
340 345 350

Lys Ala Gln Arg Leu Ile Glu Leu Gly Gly Ser Gly His Thr Ser Ser
355 360 365

Leu Tyr Ile Asp Ser Gln Asn Asn Lys Asp Lys Val Lys Glu Phe Gly
370 375 380

Leu Ala Met Lys Thr Ser Arg Thr Phe Ile Asn Met Pro Ser Ser Gln
385 390 395 400

Gly Ala Ser Gly Asp Leu Tyr Asn Phe Ala Ile Ala Pro Ser Phe Thr
405 410 415

Leu Gly Cys Gly Thr Trp Gly Gly Asn Ser Val Ser Gln Asn Val Glu
420 425 430

Pro Lys His Leu Leu Asn Ile Lys Ser Val Ala Glu Arg Arg Glu Asn
435 440 445

Met Leu Trp Phe Lys Val Pro Gln Lys Ile Tyr Phe Lys Tyr Gly Cys
450 455 460

Leu Arg Phe Ala Leu Lys Glu Leu Lys Asp Met Asn Lys Lys Arg Ala
465 470 475 480

Phe Ile Val Thr Asp Lys Asp Leu Phe Lys Leu Gly Tyr Val Asn Lys
485 490 495

Ile Thr Lys Val Leu Asp Glu Ile Asp Ile Lys Tyr Ser Ile Phe Thr
500 505 510

Asp Ile Lys Ser Asp Pro Thr Ile Asp Ser Val Lys Lys Gly Ala Lys
515 520 525

Glu Met Leu Asn Phe Glu Pro Asp Thr Ile Ile Ser Ile Gly Gly Gly
530 535 540

Ser Pro Met Asp Ala Ala Lys Val Met His Leu Leu Tyr Glu Tyr Pro
545 550 555 560

Glu Ala Glu Ile Glu Asn Leu Ala Ile Asn Phe Met Asp Ile Arg Lys
565 570 575

Arg Ile Cys Asn Phe Pro Lys Leu Gly Thr Lys Ala Ile Ser Val Ala
580 585 590

Ile Pro Thr Thr Ala Gly Thr Gly Ser Glu Ala Thr Pro Phe Ala Val
595 600 605

Ile Thr Asn Asp Glu Thr Gly Met Lys Tyr Pro Leu Thr Ser Tyr Glu
610 615 620

Leu Thr Pro Asn Met Ala Ile Ile Asp Thr Glu Leu Met Leu Asn Met
625 630 635 640

Pro Arg Lys Leu Thr Ala Ala Thr Gly Ile Asp Ala Leu Val His Ala
645 650 655

Ile Glu Ala Tyr Val Ser Val Met Ala Thr Asp Tyr Thr Asp Glu Leu
660 665 670

Ala Leu Arg Ala Ile Lys Met Ile Phe Lys Tyr Leu Pro Arg Ala Tyr
675 680 685

Lys Asn Gly Thr Asn Asp Ile Glu Ala Arg Glu Lys Met Ala His Ala
690 695 700

Ser Asn Ile Ala Gly Met Ala Phe Ala Asn Ala Phe Leu Gly Val Cys
705 710 715 720

His Ser Met Ala His Lys Leu Gly Ala Met His His Val Pro His Gly
725 730 735

Ile Ala Cys Ala Val Leu Ile Glu Glu Val Ile Lys Tyr Asn Ala Thr
740 745 750

Asp Cys Pro Thr Lys Gln Thr Ala Phe Pro Gln Tyr Lys Ser Pro Asn
755 760 765

Ala Lys Arg Lys Tyr Ala Glu Ile Ala Glu Tyr Leu Asn Leu Lys Gly
770 775 780

Thr Ser Asp Thr Glu Lys Val Thr Ala Leu Ile Glu Ala Ile Ser Lys
785 790 795 800

Leu Lys Ile Asp Leu Ser Ile Pro Gln Asn Ile Ser Ala Ala Gly Ile
805 810 815

Asn Lys Lys Asp Phe Tyr Asn Thr Leu Asp Lys Met Ser Glu Leu Ala
820 825 830

Phe Asp Asp Gln Cys Thr Thr Ala Asn Pro Arg Tyr Pro Leu Ile Ser
835 840 845

Glu Leu Lys Asp Ile Tyr Ile Lys Ser Phe
850 855

<210> 14

<211> 390

<212> PRT

<213> Clostridium acetobutylicum

<400> 14

Met Val Asp Phe Glu Tyr Ser Ile Pro Thr Arg Ile Phe Phe Gly Lys
1 5 10 15

Asp Lys Ile Asn Val Leu Gly Arg Glu Leu Lys Lys Tyr Gly Ser Lys
20 25 30

Val Leu Ile Val Tyr Gly Gly Gly Ser Ile Lys Arg Asn Gly Ile Tyr
35 40 45

Asp Lys Ala Val Ser Ile Leu Glu Lys Asn Ser Ile Lys Phe Tyr Glu
50 55 60

Leu Ala Gly Val Glu Pro Asn Pro Arg Val Thr Thr Val Glu Lys Gly
65 70 75 80

Val Lys Ile Cys Arg Glu Asn Gly Val Glu Val Val Leu Ala Ile Gly
85 90 95

Gly Gly Ser Ala Ile Asp Cys Ala Lys Val Ile Ala Ala Ala Cys Glu
100 105 110

Tyr Asp Gly Asn Pro Trp Asp Ile Val Leu Asp Gly Ser Lys Ile Lys
115 120 125

Arg Val Leu Pro Ile Ala Ser Ile Leu Thr Ile Ala Ala Thr Gly Ser
130 135 140

Glu Met Asp Thr Trp Ala Val Ile Asn Asn Met Asp Thr Asn Glu Lys
145 150 155 160

Leu Ile Ala Ala His Pro Asp Met Ala Pro Lys Phe Ser Ile Leu Asp
165 170 175

Pro Thr Tyr Thr Tyr Thr Val Pro Thr Asn Gln Thr Ala Ala Gly Thr
180 185 190

Ala Asp Ile Met Ser His Ile Phe Glu Val Tyr Phe Ser Asn Thr Lys
195 200 205

Thr Ala Tyr Leu Gln Asp Arg Met Ala Glu Ala Leu Leu Arg Thr Cys
210 215 220

Ile Lys Tyr Gly Gly Ile Ala Leu Glu Lys Pro Asp Asp Tyr Glu Ala
225 230 235 240

Arg Ala Asn Leu Met Trp Ala Ser Ser Leu Ala Ile Asn Gly Leu Leu
245 250 255

Thr Tyr Gly Lys Asp Thr Asn Trp Ser Val His Leu Met Glu His Glu
260 265 270

Leu Ser Ala Tyr Tyr Asp Ile Thr His Gly Val Gly Leu Ala Ile Leu
275 280 285

Thr Pro Asn Trp Met Glu Tyr Ile Leu Asn Asn Asp Thr Val Tyr Lys
290 295 300

Phe Val Glu Tyr Gly Val Asn Val Trp Gly Ile Asp Lys Glu Lys Asn
305 310 315 320

His Tyr Asp Ile Ala His Gln Ala Ile Gln Lys Thr Arg Asp Tyr Phe
325 330 335

Val Asn Val Leu Gly Leu Pro Ser Arg Leu Arg Asp Val Gly Ile Glu
340 345 350

Glu Glu Lys Leu Asp Ile Met Ala Lys Glu Ser Val Lys Leu Thr Gly
355 360 365

Gly Thr Ile Gly Asn Leu Arg Pro Val Asn Ala Ser Glu Val Leu Gln
370 375 380

Ile Phe Lys Lys Ser Val
385 390

<210> 15

<211> 336

<212> PRT

<213> Clostridium acetobutylicm

<400> 15

Met Asn Lys Ala Asp Tyr Lys Gly Val Trp Val Phe Ala Glu Gln Arg

1 5 10 15

Asp Gly Glu Leu Gln Lys Val Ser Leu Glu Leu Leu Gly Lys Gly Lys
20 25 30

Glu Met Ala Glu Lys Leu Gly Val Glu Leu Thr Ala Val Leu Leu Gly
35 40 45

His Asn Thr Glu Lys Met Ser Lys Asp Leu Leu Ser His Gly Ala Asp
50 55 60

Lys Val Leu Ala Ala Asp Asn Glu Leu Leu Ala His Phe Ser Thr Asp
65 70 75 80

Gly Tyr Ala Lys Val Ile Cys Asp Leu Val Asn Glu Arg Lys Pro Glu
85 90 95

Ile Leu Phe Ile Gly Ala Thr Phe Ile Gly Arg Asp Leu Gly Pro Arg
100 105 110

Ile Ala Ala Arg Leu Ser Thr Gly Leu Thr Ala Asp Cys Thr Ser Leu
115 120 125

Asp Ile Asp Val Glu Asn Arg Asp Leu Leu Ala Thr Arg Pro Ala Phe
130 135 140

Gly Gly Asn Leu Ile Ala Thr Ile Val Cys Ser Asp His Arg Pro Gln
145 150 155 160

Met Ala Thr Val Arg Pro Gly Val Phe Glu Lys Leu Pro Val Asn Asp
165 170 175

Ala Asn Val Ser Asp Asp Lys Ile Glu Lys Val Ala Ile Lys Leu Thr
180 185 190

Ala Ser Asp Ile Arg Thr Lys Val Ser Lys Val Val Lys Leu Ala Lys

195 200 205

Asp Ile Ala Asp Ile Gly Glu Ala Lys Val Leu Val Ala Gly Gly Arg
210 215 220

Gly Val Gly Ser Lys Glu Asn Phe Glu Lys Leu Glu Glu Leu Ala Ser
225 230 235 240

Leu Leu Gly Gly Thr Ile Ala Ala Ser Arg Ala Ala Ile Glu Lys Glu
245 250 255

Trp Val Asp Lys Asp Leu Gln Val Gly Gln Thr Gly Lys Thr Val Arg
260 265 270

Pro Thr Leu Tyr Ile Ala Cys Gly Ile Ser Gly Ala Ile Gln His Leu
275 280 285

Ala Gly Met Gln Asp Ser Asp Tyr Ile Ile Ala Ile Asn Lys Asp Val
290 295 300

Glu Ala Pro Ile Met Lys Val Ala Asp Leu Ala Ile Val Gly Asp Val
305 310 315 320

Asn Lys Val Val Pro Glu Leu Ile Ala Gln Val Lys Ala Ala Asn Asn
325 330 335

<210> 16

<211> 1011

<212> DNA

<213> Artificial sequence

<220>

<223> ETFa CPO for *S. cerevisiae*

<400> 16

atgaacaagg ctgactacaa ggggtgtctgg gtttcgctg aacaaagaga tggatgaatta 60

caaaagggtt ctttgaatt gctaggttaag ggtaaggaaa tggctgaaaa attgggtgtt 120

gaattgactg ctgtcttatt gggtcacaac actgaaaaga tgtccaagga cttgtgtcc 180
 cacgggtgctg acaaggtttt ggctgctgac aacgaattat tagctcattt ctccactgac 240
 gggttacgcca aggttatctg tgacttgggc aacgaaagaa agccagaaat cttattcatc 300
 ggtgctactt tcacggttag agacttgggt ccaagaattg ctgccagatt gtctactggt 360
 ttgactgctg actgtacctc ttggatatac gatgtcgaaa accgtgactt gttggctacc 420
 agaccagctt tcggtggtaa cttgattgct accattgtct gttctgacca cagacctcaa 480
 atggccaccg tcagaccagg tgtctttgaa aaattgccag ttaacgatgc taacgtttct 540
 gatgacaaga tcgaaaaggt tgccatcaaa ttgactgctt ctgatatcag aaccaagggt 600
 tccaagggtg tcaaattggc caaggacatt gctgatatcg gtgaagccaa ggttttggtt 660
 gctggtgggc gtggtgttgg ttccaaggaa aacttcgaaa aattggaaga attagcttct 720
 ttgtgggtg gtaccattgc tgcttcaga gctgccattg aaaaggaatg gggtgacaag 780
 gacttgcaag tcggtcaaac cggtgaagacc gtcagaccaa cttgtacat tgcttgtggt 840
 atctctggtg ccatccaaca cttggctggt atgcaagact ctgactacat cattgccatc 900
 aacaaagatg tcgaagctcc aatcatgaag gttgctgatt tggccattgt cggtgatgtc 960
 aacaagggtg ttccagaatt gattgtctca gtaaggctg ctaacaatta a 1011

<210> 17

<211> 259

<212> PRT

<213> Clostridium acetobutylicum

<400> 17

Met Asn Ile Val Val Cys Leu Lys Gln Val Pro Asp Thr Ala Glu Val
 1 5 10 15

Arg Ile Asp Pro Val Lys Gly Thr Leu Ile Arg Glu Gly Val Pro Ser
 20 25 30

Ile Ile Asn Pro Asp Asp Lys Asn Ala Leu Glu Glu Ala Leu Val Leu
 35 40 45

Lys Asp Asn Tyr Gly Ala His Val Thr Val Ile Ser Met Gly Pro Pro
50 55 60

Gln Ala Lys Asn Ala Leu Val Glu Ala Leu Ala Met Gly Ala Asp Glu
65 70 75 80

Ala Val Leu Leu Thr Asp Arg Ala Phe Gly Gly Ala Asp Thr Leu Ala
85 90 95

Thr Ser His Thr Ile Ala Ala Gly Ile Lys Lys Leu Lys Tyr Asp Ile
100 105 110

Val Phe Ala Gly Arg Gln Ala Ile Asp Gly Asp Thr Ala Gln Val Gly
115 120 125

Pro Glu Ile Ala Glu His Leu Gly Ile Pro Gln Val Thr Tyr Val Glu
130 135 140

Lys Val Glu Val Asp Gly Asp Thr Leu Lys Ile Arg Lys Ala Trp Glu
145 150 155 160

Asp Gly Tyr Glu Val Val Glu Val Lys Thr Pro Val Leu Leu Thr Ala
165 170 175

Ile Lys Glu Leu Asn Val Pro Arg Tyr Met Ser Val Glu Lys Ile Phe
180 185 190

Gly Ala Phe Asp Lys Glu Val Lys Met Trp Thr Ala Asp Asp Ile Asp
195 200 205

Val Asp Lys Ala Asn Leu Gly Leu Lys Gly Ser Pro Thr Lys Val Lys
210 215 220

Lys Ser Ser Thr Lys Glu Val Lys Gly Gln Gly Glu Val Ile Asp Lys
225 230 235 240

Pro Val Lys Glu Ala Ala Ala Tyr Val Val Ser Lys Leu Lys Glu Glu
 245 250 255

His Tyr Ile

<210> 18

<211> 781

<212> DNA

<213> Artificial sequence

<220>

<223> ETFb CPO for *S. cerevisiae*

<400> 18

atgaacattg ttgtttgtt gaagcaagtt ccagacactg ctgaagtcag aattgaccca 60

gtcaagggtg cttaatcag agaaggtgtt ccatttatca tcaaccaga cgacaagaac 120

gctttggaag aagctttggt ttgaaggac aactacggtg ctcacgttac cgtcatttcc 180

atgggtccac ctcaagccaa gaacgctttg gttgaagctt tggccatggg tgctgatgaa 240

gctgtcttat tgactgacag agctttcggg ggtgctgata cttagctac ctctcacacc 300

attgctgctg gttatcaagaa attgaaatac gatatcgtct ttgccgggtc tcaagccatc 360

gatggtgata ccgctcaagt cgggtccagaa attgctgaac atttgggtat tccacaagtc 420

acctacgttg aaaaggttga agttgacggt gacacttga agatcagaaa ggcttgggaa 480

gacggttacg aagttgttga agtcaagact ccagttctat tgactgcat caaggaattg 540

aacgttccaa gatacatgtc cgttgaaaag atcttcggtg ctttcgacaa ggaagtcaag 600

atgtggactg ctgatgatat cgatgtcgac aaggccaact tgggtttgaa aggttctcca 660

accaaggtca agaaatcttc taccaaggaa gtcaagggtc aaggtgaagt cattgacaaa 720

ccagtcaagg aagctgccgc ttacgttgtt tccaagttga aggaagaaca ctacatctaa 780

a

781

<210> 19
<211> 13286
<212> DNA
<213> Artificial sequence

<220>
<223> pBOL34

<400> 19
aagcttgc at gcctgcaggt cgacggcgcg ccggggcccgt ttaaaccggc ggccaagggtg 60
agacgcgc at aaccgctaga gtactttgaa gaggaacag caatagggtt gctaccagta 120
taaatagaca ggtacataca acactggaaa tgggtgtctg ttgagtacg ctttcaattc 180
atttgggtgt gcactttatt atgttacaat atggaaggga actttacact tctcctatgc 240
acatatatta attaaagtcc aatgctagta gagaaggggg gtaacacccc tccgcgtct 300
tttccgattt ttttctaaac cgtggaatat ttcggatatc cttttgttgt ttccgggtgt 360
acaatatgga ctctctcttt tctggcaacc aaaccatac atcgggatc ctataatacc 420
ttcgttggtc tccctaacat gtagggtggcg gaggggagat atacaataga acagatacca 480
gacaagacat aatgggctaa acaagactac accaattaca ctgcctcatt gatggtgta 540
cataacgaac taatactgta gccctagact ttagtagccat catcatatcg aagtttact 600
acccttttc catttgccat ctattgaagt aataatagge gcatgcaact tctttcttt 660
tttttcttt tctctctccc ccgttgttgt ctcaccatat ccgcaatgac aaaaaaatga 720
tggaagacac taaaggaaaa aattaacgac aaagacagca ccaacagatg tcgttgttcc 780
agagctgatg aggggtatct cgaagcacac gaaactttt ccttcttca ttcacgaca 840
ctactctcta atgagcaacg gtatacggcc ttcctccag ttacttgaat ttgaaataaa 900
aaaaagtttg ctgtcttgct atcaagtata aatagacctg caattattaa tctttgttt 960
cctcgtcatt gttctegttc cttttctcc ttgtttttt ttctgcaca tatttcaagc 1020
tataccaagc atacaatcaa ctatctcata tacaatgaag gaagttgtta ttgcttctgc 1080
tgtcagaact gccattgggt ctacggtaa gtctttgaag gacgtcccag ctgtcgactt 1140
gggtgctacc gccatcaagg aagctgtcaa gaaggctggt atcaagccag aagatgttaa 1200

cgaagttatc ttaggtaacg ttttgcaagc tggtttaggt caaaaccag ctcgtcaagc 1260
ttctttcaag gctgggttgc cagttgaaat tccagccatg accatcaaca aggtttgtgg 1320
ttctggtttg agaactgttt ctttggtgc tcaaatac caggtggtg acgtgatgt 1380
catcattgct ggtggtatgg aaaacatgc cagagctcca tacttggtta acaatgctag 1440
atgggggttac agaatgggta acgccaagt cgtcgatgaa atgactactg acggtttatg 1500
ggacgcttcc aacgactacc acatgggtat cactgctgaa aacattgctg aaagatggaa 1560
catctccaga gaagaacaag atgaattgc ttgggttct caaaagaagg ctgaagaagc 1620
catcaaatct ggtcaattca aggacgaaat tgtccagtt gtcataagg gtagaaaggg 1680
tgaaaccgtt gtcgacaccg atgaacaccc aagattcggg tccaccattg aaggtttggc 1740
caagttgaaa ccagctttca agaaggatgg taccgtcact gctggtaacg cttccggtt 1800
gaacgactgt gctgctgttt tgggtatcat gtctgctgaa aaggccaagg aattgggtgt 1860
caagccattg gccaagattg tctctacgg ttctgctggt gttgaccag ccatcatggg 1920
ttacggtcct ttctacgta ccaaggctgc tatgaaaag gctggttga cgttgacga 1980
attggatttg attgaatcca acgaagcttt cgctgctcaa tcttggctg ttgccaagga 2040
cttgaaatc gacatgaaca aggtcaacgt taacgggtgt gccattgctt tgggtcacc 2100
aattggtgct tccggtgcca gaattctgt tactttagtc cagctatgc aaaagcgtga 2160
tgccaagaag ggtttggtta ctctatgtat cgggtgtgt caaggtactg ccatctatt 2220
ggaaaagtgt taggcccggg cataaagcaa tctgatgag gataatgatt ttttttgaa 2280
tatacataaa tactaccgtt ttctgctag attttgtgaa gacgtaaata agtacatatt 2340
actttttaag ccaagacaag attaagcatt aactttaccc tttctcttc taagttcaa 2400
tactagttat cactgtttaa aagttatggc gagaacgtc gcggttaaaa tatattacc 2460
tgaacgtgtt gaattgaagt tctaggatgg tttaaagatt tttcctttt gggaaataag 2520
taaacaatat attgctgcct ttgcaaacg cacatacca caatatgtga ctattggcaa 2580
agaacgcatt atcctttgaa gaggtggata ctgatactaa gagagtctct attccggctc 2640

cacttttagt ccagagatta ctgtctctct tacgtatcag aacaagaaag catttccaaa 2700
 gtaattgcat ttgcccttga gcagtatata tatactaaga agtttaaaca tttaaactgt 2760
 tgtgtgcatt atatatatta aaaattaaga attagactaa ataaagtgtt tctaaaaaaa 2820
 tattaagtt gaaatgtgcg tgttgtgaat tgtgctctat tagaataatt atgacttgtg 2880
 tgcgttcat attttaaata aggaaataac caagaaagaa aaagtacat ccagagaaac 2940
 caattatata aatcaaata aaacaaccag ctccggtgtg tgtgtgtgtg tgaagctaag 3000
 agttgatgcc atttaacta aaaattttaa ggtgtgtgtg tggataaaat attagaatga 3060
 caattcgaga tgaaatttta agcaactct agtaggaaat aagcggctta ttctgttg 3120
 ctctaattc ttttagtgt atcagttccc attgataaaa aaattaaat taaaattaga 3180
 aaaattaaac cagaaaaatc aagttgatta aatgtgaca aaaattatga ttaaatgcta 3240
 ctcaacagg agcccgggcc tatttgaggt agtcgtagaa accctacca gactttctac 3300
 ctaaccaacc agctctaacg tactttctca ataaagtgtg aggtctgtac ttagagtcac 3360
 cggtttcaga gtataagaca tccatgatgg ccaaacagat atccaaaccg atgaagtcac 3420
 ctaattcaa tggaccatt ggggtggttag cacccaattt catggccttg tcgatatctt 3480
 caacagaagc aatacctca gccaaaatac cgacagcttc gttgatcatt ggaatcaaga 3540
 ttctgttgac aacgaaacct ggagcttcag caactcaac tgggtcctta ccaatggcaa 3600
 tggaagtgc ctgacagca tcgaaagttt ctgagaggt ggcaatact ctgatgactt 3660
 cgaccaact catgactgga gctgggtga agaagtgcac accgataacc ttgtctggtc 3720
 tcttggtagc agaagcaact tcagtgatgg acaaagaaga agtgttgga gccaaaatgg 3780
 ttctggctt acagatgttg tccaaatcag caaagattg ctcttgatg tccattctt 3840
 caacggcagc ttcaatgacc aatcacagt cagcagccat gttcaagtca acagtaccgg 3900
 agattctggt caagatttcg accttggtag ctcttcaat ctacccttc ttgaccaact 3960
 tggacaagtt ctgttgatg aatccaaac caggtcaac gaattcgctc ttgatctc 4020
 tcaaaacaac ttcgaaacct ttggcagcga aagcttgagc aataccagaa cccatggtac 4080

cggcaccaat gacacaaacc ttcttcattt tgatttagtg tttgtgtgtt gataagcagt 4140
 tgcttggttt ttatgaaaa atagctagaa ggaataaggg attacaagag agatgttaca 4200
 agaaagaagt aaaataaatt tgattaatat tgccattatc aaaagctatt tatatgttga 4260
 aatcgtggag atcatgtgtg ccagaaaagg ccacagtttc cggggagagg catacctga 4320
 ggtggctagg aatcacggag acctcttgac ttgcagggtg ggctagctag aattaagtga 4380
 ggtgacaagg ttccataca gttttgacct tgagacgttg ctacttacga ttgcagtat 4440
 gcaagtctca tgctgcaaac aaaagaggac cgctcaggta atcgctcaat tagtggacgt 4500
 tatcaggggc gggagaggcg aaagtggttt ttggtggtgt aagtaaaggt cgtccaaata 4560
 tgcagggtgt tgggtgctat cctagtggaa gctcggatca gtagataacc cgcctagaag 4620
 cggtattttt cttttttt cttccttctt ttctgcatt atttcaaag cttttgcgtc 4680
 aagtaatgaa tatctggcgg ttccgcggta atgcgacaat ttgtgatatg cactcttaaa 4740
 accccgccac gatgatcgca cgtgccggca ttatagacg acttttctgg ttgtcccgct 4800
 tcacggcaca tgcatgcac aatgaccgaa ttcaggtgc tactaacat tgtgttgtgt 4860
 tattgctgtg catgaggtgc tcaagtgcc gcggcatctg actagtggta actctagacg 4920
 gcttcgatgc agagagtcc tcaaaattt tctttcaat tgttgcctg gtttccgcgg 4980
 cgtatatcag tttttggcga tatggtaacg cgatactcta cggcaccttc acggtagatg 5040
 tctttttaa aagtactgt taattccagg attgaaagga agtgcgaat agtatagtat 5100
 gctttctagg ccggccggtt aaatgggccc gcggcccggt taaacggccg gcccttcct 5160
 ttacagtgc ttcgaaaag cacagcgttg tccaaggga caattttct tcaagttaat 5220
 gcataagaaa tatcttttt tatgtttage taagtaaaag cagcttgag taaaaaaaa 5280
 aatgagtaaa ttctcgatg gattagtttc tcacaggtaa cataacaaaa accaagaaaa 5340
 gcccgcttct gaaaactaca gttgactgt atgctaaagg gccagactaa tgggaggaga 5400
 aaaagaaacg aatgtatatg ctcatctaca ctctatatca ccatatggag gataagttgg 5460
 gctgagcttc tgatccaatt tattctatcc attagttgct gatatgtccc accagccaac 5520

acttgatagt atctactcgc cattcacttc cagcagcgcc agtaggggtg ttgagcttag 5580
 taaaaatgtg cgcaccacaa gcctacatga ctccacgtca catgaaacca caccgtgggg 5640
 cctgttgcg ctaggaatag gatatgcgac gaagacgctt ctgcttagta accacaccac 5700
 atttcaggg ggtcgatctg ctgtcttctt ttactgtcac gagcggccca taatcgcgct 5760
 tttttttaa aaggcgcgag acagcaaaca ggaagctcgg gtttcaacct tcggagtggg 5820
 cgcagatctg gagactggat ctttacaata cagtaaggca agccaccatc tgcttcttag 5880
 gtgcgatcga cggatatccac gtgcagaaca acatagtctg aagaaggggg ggaggagcat 5940
 gttcattctc ttagcagta agagcttggg gataatgacc aaaactggag tctcgaaatc 6000
 atataaatag acaatatatt ttcacacaat gagatttgta gtacagttct attctctc 6060
 ttgcataaat aagaaattca tcaagaactt gggttgatat ttaccaaca cacacaaaaa 6120
 acagtacttc actaaattta cacacaaaac aaaatggaat tgaacaacgt tatcttgga 6180
 aaggaaggta aggttgccgt tgcaccatc aacagacca aggccttgaa tgcttgaa 6240
 tctgacactt tgaaggaaat ggactacgtc attggtgaaa ttgaaaacga ttctgaagt 6300
 ttggctgtca tcttgaccgg tgccggtgaa aagtcttcg ttgctggtgc tgatatct 6360
 gaaatgaagg aaatgaacac cattgaaggt agaaagttcg gtatcttagg taacaaggt 6420
 ttcagaagat tggaattgtt ggaaaagcca gtcattgctg ctgtcaacgg ttctgcttg 6480
 ggtggtggtt gtgaaattgc catgtcctgt gacatcagaa ttgcttctc taacgctcgt 6540
 ttggtcaac cagaagtcgg tctaggtatc actccaggtt tcggtgtgtac tcaaagatta 6600
 tccagattgg ttggtatggg tatggccaag caattgatct tcaccgtca aaacatcaag 6660
 gctgacgaag ctttgagaat tggtttagtc aacaaggttg ttgaaccatc tgaattgatg 6720
 aacactgcc aaggaaattgc taacaagatc gtctccaacg ctccagttgc tgtcaaattg 6780
 tccaagcaag ccatcaacag aggtatgcaa tgtgatatcg acaccgctt ggcccttgaa 6840
 tctgaagctt tcggtgaatg ttctccact gaagacaaa aggatgctat gaccgcttc 6900
 atcgaaaaga gaaagattga aggtttcaag aacaggtgat gagcccgggc gcgaattct 6960

tatgatttat gatttttatt attaaataag ttataaaaaa aataagtga tacaaatttt 7020
 aaagtgactc ttaggtttta aaacgaaaat tcttattctt gagtaactct ttcctgtagg 7080
 tcaggttgct ttctcagga tagcatgagg tcgctcttat tgaccacacc tctaccggca 7140
 tgccgagcaa atgctgcaa atcgctcccc atttcacca attgtagata tgctaactcc 7200
 agcaatgagt tgatgaatct cgggtgtgtat ttatgtcct cagaggacaa cacctgttgt 7260
 aatcgttctt ccacacggat ccacagccta gccttcagtt gggctctatc ttcacgtca 7320
 ttcattgeat ctactagccc ctacctgag cttcaagacg ttatctgct ttatgtatc 7380
 atgatcttat cttgagatat gaatacataa atatatttac tcaagtgtat acgtgcatgc 7440
 ttttttacg gtttaaacat ttaaatgggc cgctctagag gatccccggg taccgagctc 7500
 gggcccagcg ctactagttc cggtaatttg aaaacaaacc cggctctgaa gcggagatcc 7560
 ggcgataatt accgcagaaa taaaccata cagagacgt agaaccagcc gcacatggcc 7620
 ggagaaactc ctgcgagaat ttcgtaaact cgcgcgcatt gcactgtat ttctaattgc 7680
 ggcaattcca ggctcgaga cctctgacat gcttttgaca ggaatagaca tttcagaat 7740
 gttatccata tgcctttcgg gttttttcc ttcttttcc atcatgaaaa atctctcgag 7800
 accgtttatc cattgctttt ttgtgtctt ttccctcgt tcacagaaag tctgaagaag 7860
 ctatagtaga actatgagct tttttgttt ctgttttct tttttttt tttacctctg 7920
 tggaaattgt tactctcaca ctcttagtt cgtttgttg tttgtttat tccaattatg 7980
 accggtgacg aaacgtggc gatgggggt accgcttatg ctcccctcca ttagtttca 8040
 ttatataaaa aggccaaata ttgtattatt ttcaaagtc ctatcattat cgtctaacat 8100
 ctaatttctc ttaaattttt tctctttctt tctataaca ccaatagtga aaatctttt 8160
 ttctctata tctacaaaaa cttttttt ctatcaacct cgttgataaa tttttctt 8220
 aacaatcgtt aataattaat taattggaaa ataaccattt ttctctctt ttatacacac 8280
 attcaaaaga aagaaaaaaa atatacccca gctagttaa gaaatcatt gaaaagaata 8340
 agaagataag aaagatttaa ttatcaaca atatcaatat gcctcaatcc tggaagaac 8400

tggccgctga taagcgcgcc cgcctcgcaa aaacatccc tgatgaatgg aaagtccaga 8460
 cgctgcctgc ggaagacagc gttattgatt tcccaaagaa atcggggatc ctttcagagg 8520
 ccgaactgaa gatcacagag gcctccgctg cagatcttgt gtccaagctg gcggccggag 8580
 agttgacctc ggtggaagtt acgctagcat tctgtaaagc ggcagcaatc gccagcagt 8640
 taacaaactg cgcccacgag ttcttcctg acgccgtct cgcgcaggca agggaactcg 8700
 atgaatacta cgcaaagcac aagagacccg ttgtccact ccatggcctc cccatctctc 8760
 tcaaagacca gcttcgagtc aagggtacg aaacatcaat gggctacatc tcatggctaa 8820
 acaagtacga cgaaggggac tcggttctga caacctgct ccgcaaagcc ggtgccgtct 8880
 tctacgtcaa gacctctgc ccgcagacc tgatggctg cgagacagtc aacaacatca 8940
 tcgggcgcac cgtcaacca cgcaacaaga actggctg cggcggcagt tctggtggtg 9000
 aggggtcgat cgttgggatt cgtggtggcg tcatcggtgt aggaacggat atcggtggct 9060
 cgattcgagt gccggccgcg ttaacttcc tctacggtct aaggccgagt catgggcggc 9120
 tgccgtatgc aaagtggcg aacagcatgg agggtcagga gacggtgcac agcgttctc 9180
 ggccgattac gactctgtt gaggacctcc gcctcttcac caaatccgtc ctcggtcagg 9240
 agccatggaa atacgactcc aaggtcatcc ccatgccctg gcgccagtcc gagtcggaca 9300
 ttattgcctc caagatcaag aacggcgggc tcaatatcgg ctactacaac ttcgacggca 9360
 atgtccttc acacctctc atctgcgcg gcgtggaaac caccgtgcc gactcgcca 9420
 aagccgtca caccgtgacc ccgtggacgc catacaagca cgatttcggc cagatctca 9480
 tctccatat ctacgcggct gacggcagcg ccgacgtaat gcgcgatatc agtgcacccg 9540
 gcgagccggc gattccaaat atcaaagacc tactgaaccc gaacatcaaa gctgttaaca 9600
 tgaacgagct ctgggacag catctccaga agtggaaata ccagatggag taccttgaga 9660
 aatggcggga ggctgaagaa aaggccggga aggaactgga cgccatcatc gcgccgatta 9720
 cgcctaccgc tgcggtacgg catgaccagt tccggtacta tgggtatgcc tctgtgatca 9780
 acctgctgga ttacagagc gtggtgttc cgtttacct tgcggataag aacatcgata 9840

agaagaatga gagtttcaag gcggttagtg agcttgatgc cctcgtgcag gaagagtatg 9900
 atccggaggc gtacatggg gcaccggtg cagtgcaggt tatcggacgg agactcagtg 9960
 aagagaggac gttggcgatt gcagaggaag tggggaagtt gctgggaaat gtggtgactc 10020
 cataggtega gaatttatac ttagataagt atgtacttac aggtatattt ctatgagata 10080
 ctgatgtata catgcatgat aatattttaa cggttattag tgccgattgt cttgtgcgat 10140
 aatgacgttc ctatcaaagc aatacactta ccacctatta catgggcca gaaaatattt 10200
 tcgaacttgt ttagaatatt agcacagagt atatgatgat atccgtaga ttatgcatga 10260
 ttcatccta caacttttc gtagcataag gattaattac ttggatgcca ataaaaaaaa 10320
 aaaacatga gaaaattca gcatgctcag aaacaattgc agtgtatcaa agtaaaaaaa 10380
 agattttgc tacatgttcc tttgaagaa agaaaatcat ggaacattag attacaaaa 10440
 atttaaccac cgctgattaa cgattagacc gtttaagcga caacaggta ttagtacaga 10500
 gaaagcattc tgtggtgtg ccccgactt tctttgcga cataggtaaa tcgaatacca 10560
 tcatactatc tttccaatg actccctaaa gaaagactct tcttcgatgt tgtatacgtt 10620
 ggagcatagg gcaagaattg tggcttgaga tgaattcact ggccgctgtt ttacaacgtc 10680
 gtgactggga aaacctggc gttaccaac ttaatgcct tgcagcacat cccctttcg 10740
 ccagctggcg taatagcgaa gaggcccgca ccgacgccc ttccaacag ttgcgcagcc 10800
 tgaatggcg atggcgctg atgcggtatt ttctccttac gcactgtgc ggtatttcac 10860
 accgcatatg gtgcactctc agtacaatct gctctgatgc cgcatagtta agccagcccc 10920
 gacacccgcc aacacccgct gacgcgcct gacgggcttg tctgtcccg gcatccgctt 10980
 acagacaagc tgtgaccgtc tccgggagct gcatgtgtca gaggttttca ccgtcatcac 11040
 cgaaacgcgc gagacgaaag ggcctcgtga tacgcctatt tttatagggt aatgtcatga 11100
 taataatggt ttcttagacg tcagggtggca ctttcgggg aaatgtgcgc ggaacccta 11160
 tttgttatt ttctaaata cattcaaata tgtatccgct catgagacaa taaccctgat 11220
 aaatgcttca ataatttga aaaaggaaga gtatgagtat tcaacattc cgtgtcgc 11280

ttattccctt ttttgcggca ttttgccttc ctgtttttgc tcaccagaa acgtggtga 11340
 aagtaaaaga tgctgaagat cagttgggtg cacgagtggg ttacatcgaa ctggatctca 11400
 acagcggtaa gatccttgag agttttcgcc ccgaagaacg tttccaatg atgagcactt 11460
 ttaaagttct gctatgtggc gcggtattat cccgtattga cgccgggcaa gagcaactcg 11520
 gtcgccgcat acactattct cagaatgact tggttgagta ctcaccagtc acagaaaagc 11580
 atcttacgga tggcatgaca gtaagagaat tatgcagtgc tgccataacc atgagtata 11640
 aactgcggc caacttactt ctgacaacga tcggaggacc gaaggagcta accgcttttt 11700
 tgcacaacat gggggatcat gtaactcgcc ttgatcgttg ggaaccggag ctgaatgaag 11760
 ccataccaaa cgacgagcgt gacaccacga tgcctgtagc aatggcaaca acgttgcgca 11820
 aactattaac tggcgaacta ctactctag ctccccgca acaattaata gactggatgg 11880
 aggcggataa agttgcagga ccactctgc gctcggccct tccggctggc tggtttattg 11940
 ctgataaatc tggagccggt gagcgtgggt ctcgcggtat cattgcagca ctggggccag 12000
 atggttaagc ctccegtatc gtagttatct acacgacggg gagtcaggca actatggatg 12060
 aacgaaatag acagatcgct gagataggtg cctcactgat taagcattgg taactgtcag 12120
 accaagtta ctcatatata ctttagattg atttaaaact tcatttttaa tttaaaagga 12180
 tctagtgaa gatccttttt gataatctca tgacaaaat cccttaacgt gagttttcgt 12240
 tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat ccttttttc 12300
 tgcgcgtaat ctgctgcttg caaacaaaaa aaccaccgct accagcgggt gtttgtttgc 12360
 cgcatcaaga gctaccaact cttttccga aggtaactgg cttcagcaga gcgcagatac 12420
 caaatactgt ctttctagt tagccgtagt taggccacca cttcaagaac tctgtagcac 12480
 cgctacata cctcgtctg ctaatcctgt taccagtggc tgctgccagt ggcgataagt 12540
 cgtgtcttac cgggttgac tcaagacgat agttaccgga taaggcgag cggtcgggct 12600
 gaacggggggg ttctgcaca cagcccagct tggagcgaac gacctacacc gaactgagat 12660
 acctacagcg tgagctatga gaaagcgcca cgcttcccga agggagaaaag gcggacaggt 12720

atccggttaag cggcagggtc ggaacaggag agcgcacgag ggagcttcca gggggaaacg 12780
 cctggatatct ttatagtcct gtcgggttcc gccacctctg acttgagcgt cgatttttgt 12840
 gatgctcgtc agggggggcgg agcctatgga aaaacgccag caacgcggcc tttttacggt 12900
 tcctggcctt ttgctggcct tttgctcaca tgttcttcc tgcgttatcc cctgattctg 12960
 tggataaccg tattaccgcc tttgagttag ctgataccgc tcgccgcagc cgaacgaccg 13020
 agcgcagcga gtcagttagc gaggaagcgg aagagcgccc aatacgcaaa ccgcctctcc 13080
 ccgcgcgttg gccgattcat taatgcagct ggcacgacag gtttcccgac tggaaagcgg 13140
 gcagttagcg caacgcaatt aatgtgagtt agctcactca ttaggcaccc caggctttac 13200
 actttatgct tccggctcgt atgttggtg gaattgtgag cggataacaa ttcaacacag 13260
 gaaacagcta tgaccatgat tacgcc 13286

<210> 20
 <211> 16359
 <212> DNA
 <213> Artificial sequence

<220>
 <223> pBOL36

<400> 20
 aagcttgcac gcttcaggt cgacggcgcg ccgggcccgt taaacaatg gcaaactgag 60
 cacaacaata ccagtcgga tcaactggca ccctctctcc cgtagtctca tctaatttt 120
 ctccggatg aggttcaga tataccgcaa cacctttatt atggttccc tgagggaata 180
 atagaatgac ccattcgaaa tcaccaattc taaacctggg cgaattgtat ttcgggtttg 240
 ttaactcgtt ccagtcagga atgttcacg tgaagctatc ttccagcaaa gtctccactt 300
 cttcatcaaa ttgtgggaga atactccaa tgctcttacc tatgggactt ccgggaaaca 360
 cagtaccgat acttccaat tcgtcttcag agctcattgt ttgttgaag agactaatca 420
 aagaatcgtt ttctcaaaaa aattaatc ttaactgata gtttgatcaa aggggcaaaa 480
 cgtaggggca aacaaacgga aaatcgtt ctcaaattt ctgatgcaa gaacttaac 540

cagtcttate taaaaattgc cttatgatcc gtctctccgg ttacagcctg tgtaactgat 600
taatctgcc ttctaatca ccattctaatt gttttaatta agggattttg tcttcattaa 660
cggtttcgc tcataaaaat gttatgacgt ttgccccgca ggcgggaaac catccacttc 720
acgagactga tctcctctgc cggaacaccg ggcatctcca acttataagt tggagaaata 780
agagaatttc agattgagag aatgaaaaaa aaaaaaaaaa aaaaggcaga ggagagcata 840
gaaatggggg tcaacttttg gtaaagctat agcatgccta tcacatataa atagagtgcc 900
agtagcgact ttttcacac tcgaaafact ctactactg ctctcttgtt gttttatca 960
cttctgttt cttcttggt aatagaatat caagctacaa aaagcataca atcaactatc 1020
aactattaac tatatcgtaa tacacaggcc ggccaaaatg aaggccaaat caaggcggga 1080
agggacaacc aggacgtaaa gggtagcctc ccataacat aaactcaata aaatatatag 1140
tcttcaactt gaaaaaggaa caagctcatg caaagagggtg gtacccgcac gccgaaatgc 1200
atgcaagtaa cctattcaaa gtaatatctc atacatgttt catgagggtg acaacatgcg 1260
actgggtgag catatgttcc gctgatgtga tgtgcaagat aaacaagcaa gacagaaact 1320
aacttctct tcattgaata aacacacccc gcgtttattt acctatcttt aaactcaac 1380
acctatate ataactaata tttcttgaga taagcacact gcaccatac cttccttaa 1440
aacgtagctt ccagttttg gtggttctgg cttcttccc gattccgccc gctaaacgca 1500
taattttgtt gcctggtggc atttgcaaaa tgcataacct atgcatttaa aagattatgt 1560
atgctctct gactttcgt gtgatgaggc tcgtggaaaa aatgaataat ttatgaattt 1620
gagaacaatt ttgtgtgtt acggtatttt actatggaat aatcaatcaa ttgaggattt 1680
tatgcaaata tcgtttgaat atttttccga ccctttgagt acttttcttc ataattgcat 1740
aatattgtcc gctgcccgtt ttctgttag acggtgtctt gatctacttg ctatcggtca 1800
acaccacctt atttctaac tattttttt ttagctcatt tgaatcagct tatggtgatg 1860
gcacattttt gcataaacct agctgtcctc gtgaacata ggaaaaaaaa atatataaac 1920
aaggctcttt cactctcctt ggaatcagat ttgggtttgt tccctttatt tcatatttc 1980

ttgtcatatt ctttttctcaa ttattatctt ctactcataa cctcacgcaa aataacacag 2040
 tcaaatcaat caaaatggac ttcaacttga ccagagaaca agaattgggc agacaaatgg 2100
 ttagagaatt tgctgaaaac gaagttaagc caattgctgc tgaaatcgat gaaactgaaa 2160
 gattcccaat ggaaaacgtc aagaagatgg gtcaatacgg tatgatgggt attccattct 2220
 ctaaggaata cgggtggtgct ggtggtgacg tcttgtctta catcattgct gtcgaagaat 2280
 tgtccaaggt ttgtggtacc actggtgtca tcttatctgc tcacacttct ctatgtgcct 2340
 ccttgatcaa cgaacacggc actgaagaac aaaagcaaaa gtacttggtt ccattggcca 2400
 aggggtgaaaa gattggtgcc tacggtttga ctgaacaaaa cgctggtact gactctggtg 2460
 ctcaacaaac tgttgccgtt ttggaagggtg accactacgt catcaacggc tccaagatct 2520
 tcacaccaa cgggtggtgt gctgacacct ttgtcatctt cgctatgacc gatcgtacca 2580
 aggggtacca gggatatctt gctttcatta ttgaaaaggg ttcaagggt ttcccatcg 2640
 gtaaggctga acaaaagttg ggtatcagag ctctctctac cactgaattg gtttctgaag 2700
 acatgattgt tccagttgaa aacatgatcg gtaaggaagg taagggttc ccaattgcca 2760
 tgaagacttt agatgggtgt agaattggta ttgctgctca agctttgggt attgctgaag 2820
 gtgccttcaa cgaagctaga gcttacctga aggaaagaaa gcaattcggt agatctttgg 2880
 acaaattcca aggtttggct tggatgatgg ctgacatgga cgttgccatc gaatctgctc 2940
 gttacttggc ctacaaggct gcttacttga agcaagctgg ttgccatac accgtcgatg 3000
 ctgccagagc taagttgcac gctgccaacg ttgcatgga tgtcaccacc aaggctgtcc 3060
 aattattcgg tggttacggt tacaccaagg actaccagt tgaaagaatg atgagagatg 3120
 ctaagatcac tgaaatctac gaaggtactt ctgaagtica aaagtgggtt atctccggt 3180
 agatcttcag ataggccccg gcataaagca atcttgatga ggataatgat tttttttga 3240
 atatacataa atactaccgt ttttctgcta gattttgtga agacgtaaat aagtacatat 3300
 tactttttaa gccaaagcaa gattaagcat taactttacc cttttctctt ctaagtttca 3360
 atactagtta tcaactgttta aaagttatgg cgagaacgtc ggcgggttaa atatattacc 3420

ctgaacgtgg tgaattgaag ttctaggatg gtttaaagat ttttctttt tgggaaataa 3480
 gtaacaata tattgctgcc ttgcaaaac gcacataccc acaatatgtg actattggca 3540
 aagaacgcat tatectttga agaggtggat actgatacta agagagtctc tattccggct 3600
 ccacttttag tccagagatt acttgcttc ttacgtatca gaacaagaaa gcatttcaa 3660
 agtaattgca ttgcccttg agcagtatat atatactaag aagttaaac atttaaacgg 3720
 ccggcctaga aagcatacta tactattcga cacttcctt caatcctgga attaacagtc 3780
 acttttaaaa aagacatcta ccgtgaaggt gccgtagagt atcgcgttac catatcgcca 3840
 aaaactgata tacgccgcgg aaaccaggca acaattgaa aagaaaaatt ttgaggaact 3900
 ctctgcatcg aagccgtcta gagttaccac tagtcagatg ccgcggggcac ttgagcacct 3960
 catgcacagc aataacacaa cacaatggtt agtagcaacc tgaattcggc cattgatgca 4020
 tgcatgtgcc gtgaagcggg acaaccagaa aagtcgtcta taaatgccgg cacgtgcgat 4080
 catcgtggcg gggttttaag agtgcatac acaaattgac gcattaccgc ggaaccgcca 4140
 gatattcatt acttgacgca aaagcgttg aaataatgac gaaaaagaag gaagaaaaaa 4200
 aaagaaaaat accgcttcta ggccgggttat ctactgatcc gagcttcac taggatagca 4260
 cccaaacacc tgcataattg gacgacctt acttacacca ccaaaaacca ctttcgctc 4320
 tccgcccct gataacgtcc actaattgag cgattacctg agcggctctc tttgtttgc 4380
 agcatgagac ttgcatactg caaatcgtaa gtagcaacgt ctcaaggta aaactgtatg 4440
 gaaacctgt cacctcactt aattctagct agcctaccct gcaagtcaag aggtctccgt 4500
 gattcctagc cacctcaagg tatgccttc cccggaaact gtggccttt ctggcacaca 4560
 tgatctccac gatttcaaca tataaatagc ttttgataat ggcaatatta atcaaattta 4620
 ttttactct tcttgaac atctctctg taatccctta ttcttctag ctattttca 4680
 taaaaacca agcaactgct tatcaacaca caaacactaa atcaaatgg tcgatttga 4740
 atactctatc ccaaccagaa tcttcttcgg taaggacaag atcaacgtt tgggtagaga 4800
 attgaagaaa tacgggtcca aggttttgat tgtctacggg ggtggtcca tcaagagaaa 4860

cggtatctac gacaaggctg tctccattt ggaaaagaac tctatcaat tctacgaatt 4920
 ggctgggtgtt gaaccaaacc caagagttac caccgtcgaa aagggtgtca agatctgtcg 4980
 tgaaaacggt gttgaagttg tttggccat cggtgggtgtt tctgccattg actgtgcaa 5040
 ggtcattgct gctgcctgtg aatacgatgg taacccatgg gacattgtct tggatggtc 5100
 taagatcaag cgtgtcttac caattgcttc catcttgact atcgtgcta ctggttctga 5160
 aatggacacc tgggctgtta tcaacaacat ggacactaac gaaaagttga ttgctgtca 5220
 cccagatatg gccccaaagt tctctattt ggaccaacc tacactaca ctgtccaac 5280
 caaccaaact gctgctggtta ctgctgatat catgtctcac atctttgaag tttacttctc 5340
 caacaccaag accgcttact tgcaagacag aatggctgaa gctctattaa gaacctgtat 5400
 caagtacggt ggtattgctt tgaaaagcc agatgactac gaagccagag ctaactgat 5460
 gtgggcttcc tctttggcta tcaacggtt attgacttac ggtaaggaca ccaactggtc 5520
 cggtcatttg atggaacacg aattgtctgc ttactacgat atcactcacg gtgtcggtt 5580
 ggccatcttg actccaaact ggatggaata cttttgaac aacgacactg tctacaagtt 5640
 cgtcgaatac ggtgttaacg tctggggtat tgacaaggaa aagaaccact acgacattgc 5700
 tcaccaagcc atccaaaaga ccagagacta ttctgtcaac gttttgggtt taccatccag 5760
 attaagagat gttggtattg aagaagaaaa attggatata atggctaagg aatctgtcaa 5820
 attgactggt ggtaccattg gtaacttgag acctgttaac gcttctgaag tttgcaa 5880
 cttcaagaaa tctgtttagg cccgggctcc tgtgaagta gcatttaac ataattttg 5940
 tcacattta atcaactga ttttctggt ttaattttc taattttaat ttaattttt 6000
 ttatcaatgg gaactgatac actaaaaaga attaggagcc aacaagaata agccgcttat 6060
 ttctactag agtttgctta aaatttcac tgaattgtc attctaata tttatccaca 6120
 cacacacctt aaaattttta gattaaatgg catcaactct tagcttcaca cacacacaca 6180
 caccgaagct ggtgtttta tttgattga tataattggt ttctctggat ggtactttt 6240
 ctttcttggt tatttctat tttaaaatat gaaacgcaca caagcataa ttattcta 6300

agagcacaat tcacaacacg cacatttcaa cttaatat ttttagaaa cactttatt 6360
 agtctaattc ttaatttta atatatataa tgcacacaca cgtttaaag ggcccgcggc 6420
 ccgtttaaac ggccggccct tccctttac agtgcttcgg aaaagcacag cgttgtccaa 6480
 gggaacaatt ttcttcaag ttaatgcata agaaatatct tttttatgt ttagctaagt 6540
 aaaagcagct tggagtaaaa aaaaaaatga gtaaatttct cgatggatta gtttctaca 6600
 ggtaacatag caaaaaccaa gaaaagcccg ctctgaaaa ctacagttga ctgtatgct 6660
 aaagggccag actaatggga ggagaaaaag aaacgaatgt atatgctcat ttacttcta 6720
 tatcaccata tggaggataa gttgggctga gcttctgac caatttattc tatcattag 6780
 ttgctgatat gtcccaccag ccaacacttg atagtatcta ctgccattc acttcagca 6840
 gcgccagtag ggtgttgag cttagtaaaa atgtgcgcac cacaagccta catgactcca 6900
 cgtcacatga aaccacaccg tggggccttg ttgcgctagg aataggatat gcgacgaaga 6960
 cgcttctgct tagtaaccac accacatttt cagggggctg atctgcttc ttcctttact 7020
 gtcacgagcg gcccataatc gcgcttttt tttaaaaggc gcgagacagc aaacaggaag 7080
 ctggggttc aaccttcgga gtggtcgag atctggagac tggatcttta caatacagta 7140
 aggcaagcca ccactgctt cttaggtgca tgcgacggta tccacgtgca gaacaacata 7200
 gtctgaagaa gggggggagg agcatgttca ttctctgtag cagtaagagc ttggtgataa 7260
 tgacaaaac tggagtctcg aaatcatata aatagacaat atatttcac acaatgagat 7320
 ttgtagtaca gttctattct ctctctgca taaataagaa attcatcaag aacttggtt 7380
 gatatttcac caacacacac aaaaaacagt acttcactaa atttacacac aaaacaaaat 7440
 gaaggttacc aacaaaagg aattgaagca aaagtgaac gaattgagag aagctcaaaa 7500
 gaagttcgt acctacactc aagaacaagt tgacaagatc ttcaagcaat gtgccattgc 7560
 tgctccaag gaacgtatca acttgccaa gttggctgac gaagaaaccg gtattggtt 7620
 ggttgaagac aagatcatca agaaccactt cgctgctgaa tacatctaca acaagtaca 7680
 gaacgaaaag acctgtgta tcatcgacca cgatgactct ttgggtatca ccaaggtgc 7740

tgaaccaatc ggtattgtcg ccgccattgt cccaaccact aaccaactt ccaactgccaat 7800
 cttcaaattct ttgatctcct tgaagaccag aaacgctatc ttcttctccc cacaccaag 7860
 agccaagaag tccaccattg ctgctgccaa attaattctg gatgctgctg ttaaggctgg 7920
 tgcccaaaag aacattattg gttggatcga tgaaccttcc attgaattgt ctcaagactt 7980
 gatgtctgaa gctgatatca tcttggtctac cgggtggcca tccatggta aggccgctta 8040
 ctcttctggt aagccagcta ttggtgttgg tgctggtaac actccagcta tcatcgatga 8100
 atctgctgac attgacatgg ctgtctctc cattatcttg tccaagactt atgacaacgg 8160
 tgtcatctgt gcctctgaac aatccatctt ggttatgaac tctatctacg aaaaggtaaa 8220
 ggaagaattt gtaagagag gttctacat cttaaaccac aatgaaattg ccaagatcaa 8280
 ggaaacatg ttcaagaacg gtgccaatca cgctgacatt gtcggtaaat ctgcttcat 8340
 cattgccaag atggctggtta ttgaagtcc acaaacact aagattttga tcggtgaagt 8400
 tcaatctgtc gaaaagtctg aattattctc tcacgaaaag ttgtctccag tcttggtat 8460
 gtacaaggtc aaggatttcg acgaagcttt gaagaaggct caaagattaa ttgaattagg 8520
 tggttctggt cacacctctt ctctatacat tgactctcaa aacaacaagg acaaggtaaa 8580
 ggaattcggc ctagctatga agacttccag aactttcatc aacatgccat cttctcaagg 8640
 tgcttctggt gatttgtaca actttgccat tgcctcatct ttcactttag gttgtgttac 8700
 ctgggggtgt aactctgttt ctcaaacgt tgaaccaaag catttgctaa acatcaagtc 8760
 cggtgctgaa agaagagaaa acatgttgtg gttcaagggt ccacaaaaga tctactcaa 8820
 atacgggtgt ttgagatttg ctttgaagga attgaaagat atgaacaaga agcgtgcttt 8880
 catcgttact gacaaggatt tgttcaaatt gggttacgtt aacaagatca ctaaggtttt 8940
 ggatgaaatt gatataagt actccatctt cactgatatc aaatctgacc caaccattga 9000
 ctccgtcaag aagggtgcta aggaaatgtt gaacttcgaa ccagatacca ttatctccat 9060
 tgggtgtgtt tctccaatgg atgtgccaa gggtatgcat ttgtgtacg aataccaga 9120
 agctgaaatc gaaaacttgg ccatcaactt catggacatc agaaagagaa tctgtaactt 9180

cccaaagtgt ggtaccaagg ccatttctgt tgccattcca accaccgctg gtaccgggtc 9240
 tgaagctact ccatttctgt tcatcacaa cgacgaaacc ggtatgaagt acccattgac 9300
 ctcttacgaa ttgactccaa acatggccat cattgacact gaattgatgt tgaacatgcc 9360
 aagaaagtgt actgtgcta ccggtattga cgctttagtc cagctatcg aagcttacgt 9420
 ctccgttatg gccactgact aactgacga attggctttg agagctatca agatgatctt 9480
 caagtacttg ccaagagctt acaagaacgg tactaacgat atcgaagctc gtgaaaagat 9540
 ggctcacgct tccaacattg ctggtatggc ttctgctaac gctttcttgg gtgtttgtca 9600
 ctccatggcc cacaagtgg gtgctatgca ccacgttct caggtattg cttgtgctgt 9660
 ttgtattgaa gaagtcata agtacaacgc tactgactgt ccaaccaagc aaactgctt 9720
 cccacaatac aagtctcaa acgccaagag aaagtacgt gaaattgctg aatactgaa 9780
 ctgaaaggt acttctgaca ctgaaaaggt cactgcttta atcgaagcta tctccaagtt 9840
 gaagattgac ttatctatc ctcaaaacat ctctgctgt ggtattaaca agaaggactt 9900
 ctacaacact ttagacaaga tgccgaatt ggcttcgat gaccaatga ccaccgctaa 9960
 cccaagatac ccattgatct ctgaattgaa ggatatctac atcaagtcct ttaagcccc 10020
 ggcgcggatc tcttatgtct ttacgattta tagtttcat tatcaagtat gcctatatta 10080
 gtatatagca tctttagatg acagtgttcg aagtttcacg aataaaagat aatattctac 10140
 ttttgctcc caccgcgtt gctagcacga gtgaacacca tccctgcct gtgagttgta 10200
 cccattctc taaactgtag acatggtagc ttcagcagtg ttcgttatgt acggcatcct 10260
 ccaacaaaca gtcggttata gttgtcctg ctctctgaa tcgtctccct cgatattct 10320
 catttccct cgcatgccag cattgaaatg atcgaagtc aatgatgaaa cggtattct 10380
 tctgtcattt actcatctca tctcatcaag ttatataatt ctatacgat gtaattttc 10440
 actttctgc ttgactcca cctataatt tcaattattg aacctcaca aatgatgcac 10500
 tgcaatgtac acacctcat atagtttaa catttaaag ggccgctcta gaggatcccc 10560
 gggtagcgag ctcgggccca gcgtactag ttccggtaat ttgaaaacaa acccggtctc 10620

gaagcggaga tccggcgata attacgcag aaataaaccc atacagaga cgtagaacca 10680
 gccgcacatg gccggagaaa ctctgcgag aatttcgtaa actcgcgcgc attgcatctg 10740
 tatttcctaa tgcggcactt ccaggcctcg agacctctga catgcttttg acaggaatag 10800
 acatttcag aatgttatcc atatgccttt cgggttttt tcttccctt tccatcatga 10860
 aaaatctctc gagaccgttt atccattgct ttttgtgtgt cttttccct cggtcacaga 10920
 aagtctgaag aagctatagt agaactatga gcttttttg ttctgtttt ccttttttt 10980
 tttttacct ctgtggaaat tgttactctc acactcttta gtctgtttgt ttgtttgtt 11040
 tattccaatt atgaccgggtg acgaaacgtg gtcgatgggtg ggtaccgctt atgctcccct 11100
 ccattagtgt cgattatata aaaaggccaa atattgtatt atttcaaat gtcctatcat 11160
 tategtctaa catctaattt ctctaaatt tttctcttt ctttctata acaccaatag 11220
 tgaaaatctt tttttctct atactacaa aaacttttt ttctatcaa cctcgttgat 11280
 aaatttttc ttaacaatc gtaataatt aattaattgg aaaataacca tttttctct 11340
 cttttataca cacattcaaa agaaagaaaa aaaatatacc ccagctagt aaagaaaatc 11400
 attgaaaaga ataagaagat aagaagatt taattatcaa acaatatcaa tatgcctcaa 11460
 tcttggaag aactggccgc tgataagcgc gccgcctcg caaaaacct cctgatgaa 11520
 tggaagtc agacgtgcc tgcggaagac agcgttattg atttccaaa gaaatcgggg 11580
 atccttcag aggccgaact gaagatcaca gaggcctccg ctgcagatct tgtgtccaag 11640
 ctggcggccg gagagtgac ctcggtggaa gttacgctag cattctgtaa acgggcagca 11700
 atgccccagc agttaacaaa ctgcgccac gagttcttcc ctgacgccgc tctcgcgag 11760
 gcaagggaac tcgatgaata ctacgcaaag cacaagagac ccgttggtcc actccatggc 11820
 ctccccatct ctctaaaaga ccagcttga gtcaagggt acgaaacatc aatgggtac 11880
 atctcatggc taaacaagta cgacgaagg gactcgggtc tgacaacct gctccgcaa 11940
 gccggtgccg tcttctact caagacctct gtccgcaga ccctgatggt ctgcgagaca 12000
 gtcaacaaca tcatcgggcg caccgtcaac ccacgcaaca agaactggtc gtgcggcggc 12060

agttctggtg gtgaggggtgc gatcgttggg attcgtggtg gcgtcatcgg ttaggaacg 12120
 gatacgggtg gctcgattcg agtgccggcc gcgttcaact tcctgtacgg tctaaggccg 12180
 agtcatgggc ggctgccgta tgcaaagatg gcgaacagca tggaggggtca ggagaeggtg 12240
 cacagcgttg tcgggccgat tacgcactct gttgaggacc tccgcctctt caccaaatcc 12300
 gtctcgggtc aggagccatg gaaatacgac tccaagggtca tccccatgcc ctggcgccag 12360
 tccgagtcgg acattattgc ctccaagatc aagaacggcg ggctcaatat cggtactac 12420
 aacttegacg gcaatgtctt tccacacct cctatctgc gcggcgtgga aaccaccgtc 12480
 gccgcactcg ccaaagccgg tcacaccgtg acccctgga cgccatacaa gcacgatttc 12540
 ggccacgac tcactccca tatctacgcg gctgacggca gcgccgacgt aatgcgcgat 12600
 atcagtgcac ccggcgagcc ggcgattcca aatatcaaag acctactgaa cccgaacatc 12660
 aaagctgtta acatgaacga gctctgggac acgcatctcc agaagtggaa ttaccagatg 12720
 gagtaccttg agaaatggcg ggaggtgaa gaaaaggccg ggaaggaact ggacgccatc 12780
 atcgcgccga ttacgcctac cgctgcggta cggcatgacc agttccggtg ctatgggtat 12840
 gcctctgtga tcaacctgct ggatttcacg agcgtggttg ttccggttac cttgcggat 12900
 aagaacatcg ataagaagaa tgagagtgc aaggcgggtg gtgagcttga tgcctcgtg 12960
 caggaagagt atgatccgga ggcgtaacct ggggcaccgg ttgcagtga ggttatcgga 13020
 cggagactca gtgaagagag gacgttggcg attgcagagg aagtggggaa gttgctggga 13080
 aatgtggtga ctccataggt cgagaattta tacttagata agtatgtact tacaggtata 13140
 ttctatgag atactgatgt atacatgat gataatattt aaacggttat tagtgccgat 13200
 tgtcttgc gataatgacg ttctatcaa agcaatacac ttaccaccta ttacatgggc 13260
 caagaaaata tttcgaact tgtttagaat attagcacag agtatatgat gatatccgtt 13320
 agattatgca tgattcattc ctacaacttt ttcgtagcat aaggattaat tacttggatg 13380
 ccaataaaaa aaaaaaacat cgagaaaatt tcagcatgct cagaaacaat tgcagtgtat 13440
 caaagtaaaa aaaagatttt cgctacatgt tccttttgaa gaaagaaaat catggaacat 13500

tagatttaca aaaatttaac caccgctgat taacgattag accgttaagc gcacaacagg 13560
 ttattagtag agagaaagca ttctgtggtg ttgccccgga ctttcttttg cgacataggt 13620
 aaatcgaata ccatcatact atcttttcca atgactccct aaagaaagac tcttcttcga 13680
 tgttgatatac gttggagcat agggcaagaa ttgtggcttg agatgaattc actggccgtc 13740
 gttttacaac gtcgtgactg ggaaaaccct ggcgttacct aacttaatcg ccttcagca 13800
 catccccctt tcgccagctg gcgtaatagc gaagaggccc gcaccgatcg cccttccaa 13860
 cagttgcgca gcctgaatgg cgaatggcgc ctgatgcggt atttctcct tacgcatctg 13920
 tgcggtatct cacaccgcat atggtgcact ctacgtacaa tctgctctga tgccgcatag 13980
 ttaagccagc cccgacaccc gccaacaccc gctgacgcgc cctgacgggc ttgtctgctc 14040
 ccggcatccg ctacagaca agctgtgacc gtctccggga gctgcatgtg tcagagggtt 14100
 tcaccgtcat caccgaaacg cgcgagacga aagggcctcg tgatacgcct attttatag 14160
 gttaatgtca tgataataat gggttcttag acgtcaggtg gcacttttcg gggaaatgtg 14220
 cgcggaaccc ctattgtttt attttctaa atacattcaa atatgtatcc gctcatgaga 14280
 caataaccct gataaatgct tcaataatat tgaaaaagga agagtatgag tattcaacat 14340
 ttccgtgtcg ccctattcc ctttttgcg gcattttgcc ttctgttt tgctcacca 14400
 gaaacgctgg tgaaagtaaa agatgctgaa gatcagttgg gtgcacgagt gggttacatc 14460
 gaactggatc tcaacagcgg taagatcctt gagagttttc gcccgaaga acgtttcca 14520
 atgatgagca cttttaagt tctgctatgt ggcgcggtat tatcccgat tgacgccggg 14580
 caagagcaac tcggtcgccg catacactat tctcagaatg acttggtga gtactacca 14640
 gtcacagaaa agcatcttac ggatggcatg acagtaagag aattatgcag tgctgccata 14700
 accatgagtg ataactgc ggccaactta cttctgacaa cgatcgagg accgaaggag 14760
 ctaaccgctt tttgcacaa catgggggat catgtaact gccttgatcg ttgggaaccg 14820
 gagctgaatg aagccatacc aaacgacgag cgtgacacca cgatgcctgt agcaatggca 14880
 acaacgttgc gcaaactatt aactggcgaa ctacttact tagcttccc gcaacaatta 14940

atagactgga tggaggcgga taaagttgca ggaccacttc tgcgctcggc ccttccggct 15000
 ggctgggtta ttgctgataa atctggagcc ggtgagcgtg ggtctcgcg taccattgca 15060
 gcactggggc cagatggtaa gccctcccg atcgtagtta tctacacgac ggggagtcag 15120
 gcaactatgg atgaacgaaa tagacagatc gctgagatag gtgcctcact gattaagcat 15180
 tggttaactgt cagaccaagt ttactcatat atactttaga ttgatttaa acttcatttt 15240
 taatttaaaa ggatctaggt gaagatcctt ttgataatc tcatgaccaa aatccctaa 15300
 cgtgagtttt cgttccactg agcgtcagac cccgtagaaa agatcaaagg atcttctga 15360
 gatccttttt ttctgcgctg aatctgctgc ttgcaaaaa aaaaaccacc gctaccagcg 15420
 gtggtttgtt tgccggatca agagctacca actcttttc cgaaggtaac tggttcagc 15480
 agagcgcaga taccaaatac tgccttcta gttagccgt agttaggcca ccactcaag 15540
 aactctgtag caccgcctac atacctcgct ctgctaatec tgttaccagt ggctgctgcc 15600
 agtggcgata agtcgtgtct taccgggttg gactcaagac gatagttacc ggataaggcg 15660
 cagcggtcgg gctgaacggg gggttcgtgc acacagccca gcttgagcg aacgacctac 15720
 accgaactga gatactaca gcgtgagcta tgagaaagcg ccacgcttc cgaagggaga 15780
 aaggcggaca ggtatccggt aagcggcagg gtcggaacag gagagcgac gagggagctt 15840
 ccagggggaa acgcctggta tctttatagt cctgtcgggt ttgccacct ctgacttgag 15900
 cgtcgatttt tgtgatgctc gtcagggggg cggagcctat ggaaaaacgc cagcaacgcg 15960
 gcctttttac ggttctggc cttttgctg cctttgctc acatgttct tctgcgta 16020
 tcccctgatt ctgtggataa ccgtattacc gcctttgagt gagctgatac cgctcgccgc 16080
 agccgaacga ccgagcgcag cgagtcagtg agcgaggaag cggaagagcg cccaatacgc 16140
 aaaccgctc tccccgcgcg ttggccgatt cattaatgca gctggcacga caggttccc 16200
 gactggaaag cgggcagtga gcgcaacgca attaattga gttagctcac tcattaggca 16260
 cccaggtt tacactttat gcttccggct cgtatgtgt gtggaattgt gagcggataa 16320
 caattcaca caggaaacag ctatgacct gattacgcc 16359

<210> 21
<211> 8684
<212> DNA
<213> Artificial sequence

<220>
<223> pBOL113

<400> 21
tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accataccac agcttttcaa ttcaattcat cttttttt ttattcttt ttttgatttc 240
ggtttctttg aaatttttt gattcggtaa tctccgaaca gaaggaagaa cgaaggaagg 300
agcacagact tagattggta tatatacgca tatgtagtgt tgaagaaaca tgaaattgcc 360
cagtattctt aacccaactg cacagaacaa aaacctgcag gaaacgaaga taaatcatgt 420
cgaaagctac atataaggaa cgtgctgcta ctcatcctag tcctgttgct gccaaagctat 480
ttaatatcat gcacgaaaag caaacaact tgtgtgcttc attggatgtt cgtaccacca 540
aggaattact ggagttagtt gaagcattag gtcccaaat ttgtttacta aaaacacatg 600
tggatatctt gactgatttt tccatggagg gcacagttaa gccgctaaag gcattatccg 660
ccaagtacaa tttttactc ttgaagaca gaaaattgc tgacattggt aatacagtca 720
aattgcagta ctctgcgggt gtatacagaa tagcagaatg ggcagacatt acgaatgcac 780
acggtgtggt gggcccaggt attgttagcg gtttgaagca ggcggcagaa gaagtaacaa 840
aggaacctag aggccttttg atgttagcag aattgtcatg caagggtcc ctatctactg 900
gagaatatac taagggtact gttgacattg cgaagagcga caaagatttt gttatcggt 960
ttattgctca aagagacatg ggtggaagag atgaaggta cgattgggtg attatgacac 1020
ccggtgtggg ttagatgac aaggagagac cattgggtca acagtataga accgtggatg 1080
atgtggtctc tacaggatct gacattatta ttgttgaag aggactattt gcaaagggaa 1140

gggatgctaa ggtagagggt gaacgttaca gaaaagcagg ctgggaagca tatttgagaa 1200
 gatgcggccca gcaaaactaa aaaactgtat tataagtaaa tgcattgata ctaaactcac 1260
 aaattagagc ttcaatttaa ttatatcagt tattacceta tgcgggtgta aataccgcac 1320
 agatgcgtaa ggagaaaata ccgcatcagg aaattgtaaa cgtaaatatt ttgttaaaat 1380
 tcgcgttaaa ttttgttaa atcagctcat ttttaacca ataggccgaa atcgcaaaa 1440
 tcccttataa atcaaaagaa tagaccgaga taggggtgag tgtgttcca gtttgaaca 1500
 agagtccact attaaagaac gtggactcca acgtcaaagg gcgaaaaacc gtctatcagg 1560
 gcgatggccc actacgtgaa ccatcaccct aatcaagttt ttgggggtcg aggtgccgta 1620
 aagcactaaa tcggaaccct aaaggaggcc cccgatttag agcttgacgg ggaaagccgg 1680
 cgaacgtggc gagaaaggaa gggaagaaag cgaaaggagc gggcgctagg gcgctggcaa 1740
 gtgtagcggc cagctgcgc gtaaccacca caccgcccgc gcttaatgcg ccgctacagg 1800
 gcgcgtcgcg ccattcgcca ttcaggctgc gcaactgttg ggaagggcga tcggtgcggg 1860
 cctcttcgct attacgccag ctggcgaaag ggggatgtgc tgcaaggcga ttaagttggg 1920
 taacgccagg gtttccag tcacgacgtt gtaaacgac ggccagttag cgcgcgtaat 1980
 acgactcact ataggcgaa ttgggtaccg ggccccccct cgaggtcgac ggtatcgata 2040
 agcttgatat cgaattcctg cagcccgggg gatccactag ttctagagcg gccatttaa 2100
 acggccggcc ctagtcaga ggggtgtaaa tgaagttaa tagtattcat tttcttata 2160
 aatcatccct tccgtgatt atacaaaaga agaggagaat atgctgaata ctgtgtat 2220
 tactctacat tatactctta tcttgacggg tattctgagc atcttactca gttcaagat 2280
 ctttaatgt ccaaaaacat ttgagccgat ctaaatactt ctgtgtttc attaatatt 2340
 aaattgtact ctttaagac atggaaagta ccaacatcgg ttgaaacagt tttcattta 2400
 cttatggtt attggtttt ccagtgaatg attattgtc gttaccctt cgtaaaagt 2460
 caaacacgtt ttaagtatt gtttagttgc tcttcgaca tatatgatta tcctgcgcg 2520
 gctaaagta aggatgcaaa aaacataaga caactgaagt taattacgt caattaagt 2580

ttccagggta atgatgtttt gggcttcac taattcaata agtatgtcat gaaatacgtt 2640
 gtgaagagca tccagaaata atgaaaagaa acaacgaaac tgggtcggcc tgtgtttct 2700
 tttctttacc acgtgatctg cggcatttac aggaagtcgc gcgttttgcg cagttgttgc 2760
 aacgcagcta cggctaacaa agcctagtgg aactcgactg atgtgttagg gcctaaaact 2820
 ggtggtgaca gctgaagtga actattcaat ccaatcatgt catggctgtc acaaagacct 2880
 tgcggaccgc acgtacgaac acatacgtat gctaatatgt gttttgatag taccagtgga 2940
 tcgcagacct gcaattttt ttaggtttg gaagaatata taaaggttgc actcattcaa 3000
 gatagttttt ttctgtgtg tctattcatt ttattattgt ttgtttaa atgttaaaaaa 3060
 ccaagaactt agtttcaaat taaattcatt acacaaacaa acaaaacaaa atgaacattg 3120
 ttgtttgtt gaagcaagtt ccagacactg ctgaagtcag aattgacca gtcaagggtg 3180
 cttaatcag agaaggtgtt ccatctatca tcaaccaga cgacaagaac gctttggaag 3240
 aagctttggt tttgaaggac aactacggtg ctcacgttac cgtcatttc atgggtccac 3300
 ctcaagccaa gaacgcttg gttgaagctt tggccatggg tgctgatgaa gctgtcttat 3360
 tgactgacag agcttcggt ggtgctgata cttagctac ctctcacacc attgctgctg 3420
 gtatcaagaa attgaaatac gatatcgtt ttgccggtcg tcaagccatc gatggtgata 3480
 ccgctcaagt cgggtccagaa attgctgaac atttgggtat tccacaagtc acctacgtt 3540
 aaaaggttga agttgacgtt gacacttga agatcagaaa ggcttgggaa gacggttacg 3600
 aagttgttga agtcaagact ccagttctat tgactgcat caaggaattg aacgttcaa 3660
 gatacatgtc cgttgaaaag atcttcggtg ctctcgacaa ggaagtcaag atgtggactg 3720
 ctgatgatat cgatgtcgac aaggccaact tgggtttgaa aggttctcca accaaggtca 3780
 agaaatctt taccaaggaa gtcaagggtc aaggtgaagt cattgacaaa ccagtcaagg 3840
 aagctgccgc ttacgttgtt tccaagtga aggaagaaca ctacatctaa agcccgggcg 3900
 gagattgata agactttct agttgcatat cttttatatt taaatctat ctattagtta 3960
 atttttga atttattctt atatatagtc tggttattct aaaatatcat ttcagtatct 4020

aaaaattccc ctctttttc agttatatct taacaggcga cagtccaaat gttgatttat 4080
cccagtcgga ttcatcaggg ttgtgaagca tttgtcaat ggtcgaaatc acatcagtaa 4140
tagtgcctct tacttgctc atagaatttc ttctcttaa cgtcaccgtt tggtcttta 4200
tagtttcgaa atctatggg ataccaaatg gtgtcccaa ttcatcgta cgggcgtatt 4260
ttttaccaat tgaagtattg gaatcgtaa ttttaaagta tatctctct ttacgtaaag 4320
cctgcgagat cctcttaagt atagcgggga agccatcgtt attcgatatt gtcgtaaca 4380
atactttgat cggcgctatg tttaaattg taaacatgga cagatatgcg atgaaaacgc 4440
taagtatac tccaaatggt gaaaggtag atgcttgga acaatactg gaaatcaccg 4500
gaaacaccat atgcggcgaa aagccaatta gtgtgatact aagtgcctta tcgaaaatcc 4560
gtgatgccgg tccttcaggc atcaaattc agtggcctaa ttattcacag agttctcatg 4620
tgacaagtat tgatgatagt agtgtcagtt atgcttcagg ttatgttact ataggataat 4680
gatcacggct aaaacggctg aatgtaagca tatatcttc gattgtataa ttgtcccaa 4740
atactacagc atctcaagga aaaaaaaca aaaactcca aaaaaatcga atccctgagg 4800
aatctttaat acattttcaa tctatttaag tttataaac gtgtatatga gatgtcatga 4860
gcatgaatta ttaataataa aaactaaatc attaaagtaa cttaggagt taaagcccgg 4920
gctttaattg ttgacgcct tgacttgagc aatcaattct ggaacaacct tgttgacatc 4980
accgacaatg gccaaatcag caaccttcatt gattggagct tcgacatctt tgttgatggc 5040
aatgatgtag tcagagtctt gcataccagc caagtgttg atggcaccag agataccaca 5100
agcaatgtac aaagtgggc tgacggctt accggttga ccgacttgca agtccttgtc 5160
aaccattcc tttcaatgg cagctctgga agcagcaatg gtaccacca acaaagaagc 5220
taattctcc aattttcga agtttctt ggaaccaaca ccacgaccac cagcaacca 5280
aaccttggt tcaccgatat cagcaatgtc ctggccaat ttgacaacct tggaacctt 5340
ggttctgata tcagaagcag tcaattgat ggcaacctt tcgatctgt catcagaaac 5400
gtagcatcg ttaactggca attttcaaa gacacctggt ctgacggtgg ccatttgagg 5460

tctgtgtgca gaacagacaa tggtagcaat caagttacca ccgaaagctg gtctggtagc 5520
caacaagtca cggttttcga catcgataac caaagaggta cagtcagcag tcaaaccagt 5580
agacaatctg gcagcaattc ttggacccaa gtctctaccg atgaaagtag caccgatgaa 5640
taagatttct ggcttttctt cgttgaccaa gtcacagata accttggcgt aaccgtcagt 5700
ggagaaatga gctaataatt cgtgtgcagc agccaaaacc ttgtcagcac cgtgggacaa 5760
caagtccttg gacatctttt cagtgttgag acccaataag acagcagtca attcaacacc 5820
caattttca gccatttctt tacccttacc tagcaattcc aaagaaacct ttgtaatc 5880
accatctctt tgttcagcga aaaccagac accctttag tcagccttgt tcatgtttag 5940
ttaattatag ttcgttgacc gtatatctta aaaacaagta ctcttaaaa aaaaacctg 6000
aagggataa acaagtagaa tagatagaga gaaaaataga aaatgcaaga gaatttatat 6060
attagaaaga gagaaagaaa aatggaaaaa aaaaaatagg aaaagccaga aatagcacta 6120
gaaggagcga caccagaaaa gaaggatgag gaaccaattt agctatatat agttaactac 6180
cggctcgatc atctctgcct ccagcatagt cgaagaagaa tttttttt cttgaggctt 6240
ctgtcagcaa ctcttatctt ttctttctt ttgttgagc ctaaaaagtt cccacgttct 6300
ctgtacgac gccgtcaca acaaccttat gggttaattg tcgcggtctg ggtgtataaa 6360
tgtgtgggtg caggccggcc gtttaaacgg gccgccaccg cgttgagctt ccagctttg 6420
ttcccttag tgagggttaa ttgcgcgtt ggcgtaata tggtcatagc tgttctctg 6480
gtgaaattgt tatccgtca caattccaca caacatagga gccggaagca taaagttaa 6540
agcctggggt gcctaagag tgaggtaact cacattaatt gcgttcgct cactgcccgc 6600
ttccagtcg ggaaacctgt cgtgccagct gcattaatga atcgccaac gcgcggggag 6660
aggcggttg cgtattgggc gctctccgc ttctcgtc actgactgc tgcgtcgtg 6720
cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacgtt tatccacaga 6780
atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg 6840
taaaaaggcc gcgttgctgg cgtttttcca taggtccgc cccctgacg agcatcaca 6900

aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt 6960
 tccccctgga agtccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct 7020
 gtccgccttt ctcccctcgg gaagcgtggc gctttctcat agtcacgct gtaggtatct 7080
 cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc 7140
 cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt 7200
 atgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc 7260
 tacagagttc ttgaagtggg ggcctaacta cggctacact agaaggacag tatttggat 7320
 ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa 7380
 acaaaccacc gctggtagcg gtggtttttt tgttgcaag cagcagatta cgcgcagaaa 7440
 aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga 7500
 aaactcacgt taagggatit tggatcatgag attatcaaaa aggatcttca cctagatcct 7560
 tttaaattaa aaatgaagtt taaatcaat ctaaagtata tatgagtaaa cttggtctga 7620
 cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcac 7680
 catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggct taccatctgg 7740
 cccagtgct gcaatgatac cgcgagaccc acgctcaccg gctccagatt taccagcaat 7800
 aaaccagcca gccggaaggg ccgagcgcag aagtggctct gcaactttat ccgcctccat 7860
 ccagttctatt aattgttgc gggaagctag agtaagtagt tcgccagta atagtttgcg 7920
 caacgttgtt gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc 7980
 attcagctcc ggttcccaac gatcaaggcg agttacatga tccccatgt tgtgcaaaaa 8040
 agcgggttagc tccttcggtc ctccgatcgt tgtcagaagt aagttggccg cagtgttacc 8100
 actcatggtt atggcagcac tgcataatc tcttactgtc atgcatccg taagatgctt 8160
 ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag 8220
 ttgtcttgc ccggcgtcaa tacgggataa taccgcgcca catagcagaa ctttaaaagt 8280
 gctcatcatt ggaaaacgtt cttcggggcg aaaactctca aggatcttac cgctgttgag 8340

atccagttcg atgtaacca ctcgtgcacc caactgatct tcagcatctt ttactttcac 8400

cagcgtttct gggtagagcaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc 8460

gacacggaaa tgtgaatac tcatactctt ccttttcaa tattattgaa gcatttatca 8520

gggttattgt ctcagagcg gatacatatt tgaatgtatt tagaaaaata acaaatagg 8580

ggttccgcgc acatttccc gaaaagtgcc acctgacgtc taagaaacca ttattatcat 8640

gacattaacc tataaaaata ggcgtatcac gaggcccttt cgtc 8684