SEQUENCE LISTING

<110> Commonwealth Scientific and Industrial Research Organisation
Nourish Ingredients Pty Ltd

<120> Production of saturated fats in microbes

<130> 534616PRV

<160> 176

<170> PatentIn version 3.5

<210> 1
<211> 419
<212> PRT
<213> Yarrowia lipolytica

<400> 1

```
Met Asp Ser Thr Thr Gln Thr Asn Thr Gly Thr Gly Lys Val Ala Val
1               5                  10                  15


Gln Pro Pro Thr Ala Phe Ile Lys Pro Ile Glu Lys Val Ser Glu Pro
            20                  25                  30


Val Tyr Asp Thr Phe Gly Asn Glu Phe Thr Pro Pro Asp Tyr Ser Ile
        35                  40                  45


Lys Asp Ile Leu Asp Ala Ile Pro Gln Glu Cys Tyr Lys Arg Ser Tyr
    50                  55                  60


Val Lys Ser Tyr Ser Tyr Val Ala Arg Asp Cys Phe Phe Ile Ala Val
65                  70                  75                  80


Phe Ala Tyr Met Ala Tyr Ala Tyr Leu Pro Leu Ile Pro Ser Ala Ser
                85                  90                  95


Gly Arg Ala Val Ala Trp Ala Met Tyr Ser Ile Val Gln Gly Leu Phe
            100                 105                 110


Gly Thr Gly Leu Trp Val Leu Ala His Glu Cys Gly His Ser Ala Phe
```

```
        115                 120                 125

Ser Asp Ser Asn Thr Val Asn Asn Val Thr Gly Trp Val Leu His Ser
    130                 135                 140

Ser Met Leu Val Pro Tyr Tyr Ala Trp Lys Leu Thr His Ser Met His
145                 150                 155                 160

His Lys Ser Thr Gly His Leu Thr Arg Asp Met Val Phe Val Pro Lys
                165                 170                 175

Asp Arg Lys Glu Phe Met Glu Asn Arg Gly Ala His Asp Trp Ser Glu
            180                 185                 190

Leu Ala Glu Asp Ala Pro Leu Met Thr Leu Tyr Gly Leu Ile Thr Gln
        195                 200                 205

Gln Val Phe Gly Trp Pro Leu Tyr Leu Leu Ser Asn Val Thr Gly Gln
    210                 215                 220

Lys Tyr Pro Lys Leu Asn Lys Trp Ala Val Asn His Phe Asn Pro Asn
225                 230                 235                 240

Ala Pro Leu Phe Glu Lys Lys Asp Trp Phe Asn Ile Trp Ile Ser Asn
                245                 250                 255

Val Gly Ile Gly Ile Thr Met Ser Val Ile Ala Tyr Ser Ile Asn Arg
            260                 265                 270

Trp Gly Leu Ala Ser Val Thr Leu Tyr Tyr Leu Ile Pro Tyr Leu Trp
        275                 280                 285

Val Asn His Trp Leu Val Ala Ile Thr Tyr Leu Gln His Thr Asp Pro
    290                 295                 300

Thr Leu Pro His Tyr His Ala Asp Gln Trp Asn Phe Thr Arg Gly Ala
305                 310                 315                 320
```

```
Ala Ala Thr Ile Asp Arg Glu Phe Gly Phe Ile Gly Ser Phe Cys Phe
                325                 330                 335

His Asp Ile Ile Glu Thr His Val Leu His His Tyr Val Ser Arg Ile
            340                 345                 350

Pro Phe Tyr Asn Ala Arg Ile Ala Thr Glu Lys Ile Lys Lys Val Met
        355                 360                 365

Gly Lys His Tyr Arg His Asp Asp Thr Asn Phe Ile Lys Ser Leu Tyr
    370                 375                 380

Thr Val Ala Arg Thr Cys Gln Phe Val Glu Gly Lys Glu Gly Ile Gln
385                 390                 395                 400

Met Phe Arg Asn Val Asn Gly Val Gly Val Ala Pro Asp Gly Leu Pro
                405                 410                 415

Ser Lys Lys
```

<210> 2
<211> 3260
<212> DNA
<213> Yarrowia lipolytica

<400> 2

```
aactatacaa gtatcaagtt gcccattgtg ttgtatgttc cagaaccatt gtacccgtac       60

aagtagttaa gcatagccct tgtgcaagag tgtgtacgat caatgtctca gatacggcgg      120

tgactgccgc ttctggcagt ctttcaacgc tgcgtttctt cgtcgatgga ccggaagagc      180

tctctttacg agccaaaggc tccaagagtt tggttgtctt gtactctcta tctttccaca      240

ccgccgccca caaggttcgg aacgacatct cccaactcta actttaagca tggcctgttg      300

agaagccatg ctattcatcg gacacctgtt caacgagtca acaagcggac atagtcgtca      360

agtagtggca gcacagagag agagtcgtgc caatgggggt catgtagtag cacatgacga      420

aataagagac gttggattca tgtagaaccg acgtagtcac gtgaacacgt ccaaggtctt      480
```

```
cccctcttca cgtgcacgtg taactggttc tttcactctc ccccccgcca taagtcgacc      540

catgtcacgg ttgagacgtg gtccctccac actatccacg ctgagagcag cagtataccc      600

cgggcggttc cgttcggagc tctcttggcc aatccgcaca tgcgcattat gcgtttttca      660

ggtggtaaca gacgaggagt acgagcgtgt taaagtggcg gttaggagag gcggggcccg      720

tagttatata caagaggtag atgcgtgctg gtgttagagg ggctctcagg attaggagga      780

aaatttgaca ttggccctca acatataacc tcgggtgtgc ctctgtttac cctcagcttt      840

tgcttgtccc caagtcagtc acgccaggcc aaaaaggttg gtggattgac agggagaaaa      900

aaaaaagcct agtgggttta aactcgaggt aagacattga aatatatacc ggtcggcatc      960

ctgagtccct ttctcgtatt ccaacagacc gaccatagaa atggattcga ccacgcagac     1020

caacaccggc accggcaagg tggccgtgca gcccccccacg gccttcatta agcccattga     1080

gaaggtgtcc gagcccgtct acgacacctt tggcaacgag ttcactcctc cagactactc     1140

tatcaaggat attctggatg ccattcccca ggagtgctac aagcggtcct acgttaagtc     1200

ctactcgtac gtggcccgag actgcttctt tatcgccgtt tttgcctaca tggcctacgc     1260

gtacctgcct cttattccct cggcttccgg ccgagctgtg gcctgggcca tgtactccat     1320

tgtccagggt ctgtttggca ccggtctgtg ggttcttgcc cacgagtgtg gccactctgc     1380

tttctccgac tctaacaccg tcaacaacgt caccggatgg gttctgcact cctccatgct     1440

ggtcccttac tacgcctgga agctgaccca ctccatgcac cacaagtcca ctggtcacct     1500

cacccgtgat atggtgtttg tgcccaagga ccgaaaggag tttatggaga accgaggcgc     1560

ccatgactgg tctgagcttg ctgaggacgc tcccctcatg accctctacg gcctcatcac     1620

ccagcaggtg tttggatggc ctctgtatct gctgtctaac gttaccggac agaagtaccc     1680

caagctcaac aaatgggctg tcaaccactt caaccccaac gccccgctgt ttgagaagaa     1740

ggactggttc aacatctgga tctctaacgt cggtattggt atcaccatgt ccgtcatcgc     1800

atactccatc aaccgatggg gcctggcttc cgtcaccctc tactacctga tcccctacct     1860

gtgggtcaac cactggctcg tggccatcac ctacctgcag cacaccgacc ccactctgcc     1920

ccactaccac gccgaccagt ggaacttcac ccgaggagcc gccgccacca tcgaccgaga     1980
```

gtttggcttc atcggctcct tctgcttcca tgacatcatc gagacccacg ttctgcacca 2040

ctacgtgtct cgaattccct tctacaacgc ccgaatcgcc actgagaaga tcaagaaggt 2100

catgggcaag cactaccgac acgacgacac caacttcatc aagtctcttt acactgtcgc 2160

ccgaacctgc cagtttgttg aaggtaagga aggcattcag atgtttagaa acgtcaatgg 2220

agtcggagtt gctcctgacg gcctgccttc taaaaagtag agctagaaat gttatttgat 2280

tgtgttttaa ctgaacagca ccgagcccga ggctaagcca agcgaagccg aggggttgtg 2340

tagtccatgg acgtaacgag taggcgatat caccgcactc ggcactgcgt gtctgcgttc 2400

atgggcgaag tcacattacg ctgacaaccg ttgtagtttc cctttagtat caatactgtt 2460

acaagtaccg gtactcgtac tcgtactgat acgaatctgt gggaagaagt cacccttatc 2520

agaccttcat actgatgttt cggatatcaa tagaactggc atagagccgt taaagaagtt 2580

tcacttaatc actccaaccc tcctacttgt agattcaagc agatcgataa gatggatttg 2640

atggtcagtg ctagcctggg cccgagcgtc accatcagtc attctttaca agttgttcca 2700

ctcttctccc tacttgaatc gatctgtctg ggcgcctgat attcgttcga cttgtacgat 2760

agcggggtgt atttcgagct cgccgaaacc ggttttccat gaatttccac cgataattcc 2820

attggcaatt caatcgacca taagcgtatc cggtaacggt gagacattga gacaccgatc 2880

gattagagag actactgtac gagcagtcta ctgtatatta gaaggaacaa taggtagtcc 2940

aactgagctt ttaactagtg ggtgtttttt gaacgtgttt ggagatctta ttaagcaaaa 3000

aaaaagctgt gttcacatta acaccaacaa accctgccgg tagttccttc aaaaacgcca 3060

catgcatcaa tcatcttcag ctttgccgat ctcctcctgc cgaggttaca accaactgga 3120

gccatcttta tcgtgccgta gagaagcgca gccagggaga gtttgagacc ggcaaaacga 3180

gggaatcggt agattctcag aaataatcca tatggagtca cgtgatcaga cttcccatgg 3240

acttttatat tgatatttgt 3260

<210> 3
<211> 1620
<212> DNA
<213> Artificial Sequence

<220>
<223> Construct encoding hygromycin resistance selectable marker gene

<400> 3
tagggataac agggtaatga taacttcgta tagcatacat tatacgaagt tatgggttgg 60

cggcgcattt gtgtcccaaa aaacagcccc aattgcccca ttgacccca aattgaccca 120

gtagcggacc caaccccggc gagagccccc ttcaccccac atatcaaacc tcccccggtt 180

cccacacttg ccgttaaggg cgtagggtac tgcagtctgg aatctacgct tgttcagact 240

ttgtactagt ttctttgtct ggccatccgg gtaacccatg ccggacgcaa aatagactac 300

tgaaaatttt tttgctttgt ggttgggact ttagccaagg gtataaaaga ccaccgtccc 360

cgaattacct ttcctcttct tttctctctc tccttgtcaa ctcacacccg aagcacaatg 420

aagaagcccg agctgaccgc tacctctgtt gagaagttcc tgattgagaa gtttgattcc 480

gtttccgacc tgatgcagct gtccgagggc gaggagtctc gagccttctc ctttgacgtg 540

ggcggacgag gttacgttct gcgagtgaac tcgtgtgccg acggcttcta caaggatcga 600

tacgtctacc gacactttgc ttctgccgct ctgcccatcc ctgaggttct cgacattggc 660

gagttctctg agtccctcac ctactgcatc tctcgacgag ctcagggagt caccctgcag 720

gacctccctg agactgagct gcctgctgtc ctccagcctg ttgctgaggc catggacgct 780

atcgctgctg ctgatctgtc ccagacctcg ggtttcggcc cctttggacc tcagggaatt 840

ggacagtaca ccacttggcg agacttcatc tgtgctattg ccgatcctca cgtctaccat 900

tggcagaccg ttatggacga tactgtgtcg gcttctgtcg ctcaggctct ggacgagctg 960

atgctctggg ccgaggattg ccccgaggtt cgacacctgg tgcatgctga cttcggttcc 1020

aacaacgttc tcaccgacaa cggccgaatc actgccgtga ttgactggtc cgaggctatg 1080

tttggcgact cgcagtacga ggtggccaac atcttctttt ggcgaccctg gctggcttgt 1140

atggagcagc agacccgata cttcgagcga cgacatcctg agctcgctgg ctcccctcga 1200

ctgcgagctt acatgctccg aattggtctg gaccagctct accagtcgct ggtggatggc 1260

aactttgacg atgctgcctg ggctcaggga cgatgtgacg ccatcgtgcg atctggcgct 1320

ggaaccgtcg gacgaactca gattgcccga cgatccgctg ctgtctggac cgacggatgc 1380

gtggaggtcc tggctgattc gggtaaccga cgaccctcta ctcgacctcg agctaaggag 1440

taagtgtctg tggtatctaa gctatttatc actctttaca acttctacct caactatcta 1500

ctttaataaa tgaatatcgt ttattctcta tgattactgt atatgcgttc ctctaagaca 1560

aatcgaatat aacttcgtat agcaggagtt atccgaagcg agtagggata acagggtaat 1620

<210> 4
<211> 341
<212> PRT
<213> Escherichia coli

<400> 4

Met Lys Lys Pro Glu Leu Thr Ala Thr Ser Val Glu Lys Phe Leu Ile
1 5 10 15

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

Glu Ser Arg Ala Phe Ser Phe Asp Val Gly Gly Arg Gly Tyr Val Leu
35 40 45

Arg Val Asn Ser Cys Ala Asp Gly Phe Tyr Lys Asp Arg Tyr Val Tyr
50 55 60

Arg His Phe Ala Ser Ala Ala Leu Pro Ile Pro Glu Val Leu Asp Ile
65 70 75 80

Gly Glu Phe Ser Glu Ser Leu Thr Tyr Cys Ile Ser Arg Arg Ala Gln
85 90 95

Gly Val Thr Leu Gln Asp Leu Pro Glu Thr Glu Leu Pro Ala Val Leu
100 105 110

Gln Pro Val Ala Glu Ala Met Asp Ala Ile Ala Ala Ala Asp Leu Ser
115 120 125

Gln Thr Ser Gly Phe Gly Pro Phe Gly Pro Gln Gly Ile Gly Gln Tyr

```
    130                 135                 140

Thr Thr Trp Arg Asp Phe Ile Cys Ala Ile Ala Asp Pro His Val Tyr
145                 150                 155                 160

His Trp Gln Thr Val Met Asp Asp Thr Val Ser Ala Ser Val Ala Gln
                165                 170                 175

Ala Leu Asp Glu Leu Met Leu Trp Ala Glu Asp Cys Pro Glu Val Arg
            180                 185                 190

His Leu Val His Ala Asp Phe Gly Ser Asn Asn Val Leu Thr Asp Asn
        195                 200                 205

Gly Arg Ile Thr Ala Val Ile Asp Trp Ser Glu Ala Met Phe Gly Asp
    210                 215                 220

Ser Gln Tyr Glu Val Ala Asn Ile Phe Phe Trp Arg Pro Trp Leu Ala
225                 230                 235                 240

Cys Met Glu Gln Gln Thr Arg Tyr Phe Glu Arg Arg His Pro Glu Leu
                245                 250                 255

Ala Gly Ser Pro Arg Leu Arg Ala Tyr Met Leu Arg Ile Gly Leu Asp
            260                 265                 270

Gln Leu Tyr Gln Ser Leu Val Asp Gly Asn Phe Asp Asp Ala Ala Trp
        275                 280                 285

Ala Gln Gly Arg Cys Asp Ala Ile Val Arg Ser Gly Ala Gly Thr Val
    290                 295                 300

Gly Arg Thr Gln Ile Ala Arg Arg Ser Ala Ala Val Trp Thr Asp Gly
305                 310                 315                 320

Cys Val Glu Val Leu Ala Asp Ser Gly Asn Arg Arg Pro Ser Thr Arg
                325                 330                 335
```

Pro Arg Ala Lys Glu
340

<210> 5
<211> 1165
<212> DNA
<213> Artificial Sequence

<220>
<223> Construct encoding nourseothricin resistance selectable marker gene

<400> 5
tagggataac agggtaatga taacttcgta tagcatacat tatacgaagt tatggggttg 60

gcggcgcatt tgtgtcccaa aaaacagccc caattgcccc aattgacccc aaattgaccc 120

agtagcggac ccaaccccgg cgagagcccc cttcacccca catatcaaac ctcccccggt 180

tcccacactt gccgttaagg gcgtagggta ctgcagtctg gaatctacgc ttgttcagac 240

tttgtactag tttctttgtc tggccatccg ggtaacccat gccggacgca aaatagacta 300

ctgaaaattt ttttgctttg tggttgggac tttagccaag ggtataaaag accaccgtcc 360

ccgaattacc tttcctcttc ttttctctct ctccttgtca actcacaccc gaagcacaat 420

gaccactctg gatgacaccg cttaccgata ccgaacttcc gttcctggcg atgccgaggc 480

tattgaggct ctggatggat ctttcaccac tgacaccgtt ttccgagtga ccgctactgg 540

cgacggcttc accctgcgag aggtgcctgt cgaccctcct ctcaccaagg ttttccctga 600

cgatgagtcg gacgatgagt ctgacgctgg agaggacggc gaccctgact ctcgaacttt 660

cgtggcttac ggcgacgatg gcgacctggc cggctttgtg gtcgtttctt actccggatg 720

gaaccgacga ctgaccgtgg aggacatcga ggtcgctcct gagcaccgag gtcatggtgt 780

cggacgagct ctgatgggtc tggctactga gttcgctcga gagcgaggtg ctggccacct 840

gtggctcgag gtcaccaacg ttaacgcccc tgctattcat gcctaccgac gaatgggttt 900

taccctgtgt ggcctcgata ctgccctgta cgacggaacc gcttccgatg gagagcaggc 960

cctctacatg tcgatgccct gcccttaagt gtctgtggta tctaagctat ttatcactct 1020

ttacaacttc tacctcaact atctacttta ataaatgaat atcgtttatt ctctatgatt 1080

actgtatatg cgttcctcta agacaaatcg aatataactt cgtatagcag gagttatccg 1140

aagcgagtag ggataacagg gtaat 1165

<210> 6
<211> 189
<212> PRT
<213> Streptomyces noursei

<400> 6

Met Thr Thr Leu Asp Asp Thr Ala Tyr Arg Tyr Arg Thr Ser Val Pro
1 5 10 15

Gly Asp Ala Glu Ala Ile Glu Ala Leu Asp Gly Ser Phe Thr Thr Asp
20 25 30

Thr Val Phe Arg Val Thr Ala Thr Gly Asp Gly Phe Thr Leu Arg Glu
35 40 45

Val Pro Val Asp Pro Pro Leu Thr Lys Val Phe Pro Asp Asp Glu Ser
50 55 60

Asp Asp Glu Ser Asp Ala Gly Glu Asp Gly Asp Pro Asp Ser Arg Thr
65 70 75 80

Phe Val Ala Tyr Gly Asp Asp Gly Asp Leu Ala Gly Phe Val Val Val
85 90 95

Ser Tyr Ser Gly Trp Asn Arg Arg Leu Thr Val Glu Asp Ile Glu Val
100 105 110

Ala Pro Glu His Arg Gly His Gly Val Gly Arg Ala Leu Met Gly Leu
115 120 125

Ala Thr Glu Phe Ala Arg Glu Arg Gly Ala Gly His Leu Trp Leu Glu
130 135 140

Val Thr Asn Val Asn Ala Pro Ala Ile His Ala Tyr Arg Arg Met Gly

145 150 155 160

Phe Thr Leu Cys Gly Leu Asp Thr Ala Leu Tyr Asp Gly Thr Ala Ser
165 170 175

Asp Gly Glu Gln Ala Leu Tyr Met Ser Met Pro Cys Pro
180 185

<210> 7
<211> 417
<212> DNA
<213> Yarrowia lipolytica

<400> 7
tagggataac agggtaatga taacttcgta tagcatacat tatacgaagt tatgggttgg 60

cggcgcattt gtgtcccaaa aaacagcccc aattgcccca ttgacccca aattgaccca 120

gtagcggacc caaccccggc gagagccccc ttcaccccac atatcaaacc tcccccggtt 180

cccacacttg ccgttaaggg cgtagggtac tgcagtctgg aatctacgct tgttcagact 240

ttgtactagt ttctttgtct ggccatccgg gtaacccatg ccggacgcaa aatagactac 300

tgaaaatttt tttgctttgt ggttgggact ttagccaagg gtataaaaga ccaccgtccc 360

cgaattacct ttcctcttct tttctctctc tccttgtcaa ctcacacccg aagcaca 417

<210> 8
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 063 SacII-Hyg/Nos_F

<400> 8
gggccgcggt agggataaca gggtaatgat aac 33

<210> 9
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 064 SacII-Hyg/Nos_R

<400> 9
cccccgcgga ttaccctgtt atccctactc gc 32

<210> 10
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 213 LRO_DelCas5F

<400> 10
gggggcgcgc cgagtcgact ttccgataca gaag 34

<210> 11
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 214 LRO_DelCas5R

<400> 11
cccgcgatcg ctgatggttg tgatcaaccg gaaa 34

<210> 12
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 215 LRO_DelCas3F

<400> 12
ggggcgatcg cagagtccgt tttgtagagt aatatg 36

<210> 13
<211> 35
<212> DNA

<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 216
LRO_DelCas3R

<400> 13
cccgcggccg cacggacctc gcaagttgca aaaag 35

<210> 14
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 217
DGA1_DelCas5F

<400> 14
gggggcgcgc catgctgcgg gcggatcctg g 31

<210> 15
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 218
DGA1_DelCas5R

<400> 15
cccgcgatcg cagcttttgt tttgtgtgac ttgtc 35

<210> 16
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 219
DGA1_DelCas3F

<400> 16
ggggcgatcg cggaaaactg cctgggttag gc 32

<210> 17

<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 220
DGA1_DelCas3R

<400> 17
cccgcggccg ctctgatggc ctggagcgag tt 32

<210> 18
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 221
DGA2_DelCas5F

<400> 18
gggggcgcgc ctgggagtgt atttggaaaa tgact 35

<210> 19
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 222
DGA2_DelCas5R

<400> 19
cccgcgatcg ctttgcgggc ggtacgggta c 31

<210> 20
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 223
DGA2_DelCas3F

<400> 20
ggggcgatcg ccataacact catcagtagc ctttac 36

<210> 21
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 224
DGA2_DelCas3R

<400> 21
cccgcggccg ctcttgtaat tccatagata atatatacg 39

<210> 22
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 225
ARE1_DelCas5F

<400> 22
gggggcgcgc cgccgacaat ctctcttcct catt 34

<210> 23
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 226
ARE1_DelCas5R

<400> 23
cccgcgatcg ctgtgtgtgc ggagagtgtc ttg 33

<210> 24
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 227
ARE1_DelCas3F

<400> 24

ggggcgatcg cgcacagtcg cttcaccact tg 32

<210> 25
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 228 ARE1_DelCas3R

<400> 25
cccgcggccg ctaaaacaat gacaaataca actctag 37

<210> 26
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 229 AsiSI-Hyg/Nos_F

<400> 26
ggggcgatcg ctagggataa cagggtaatg ataac 35

<210> 27
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 230 AsiSI-Hyg/Nos_R

<400> 27
cccgcgatcg cattaccctg ttatccctac tcgc 34

<210> 28
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 239 Fad2-gene_F

<400> 28
gtgtccgagc ccgtctacg 19

<210> 29
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 240
Fad2-gene_R

<400> 29
acatggtgat accaataccg ac 22

<210> 30
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 241
ARE1KOScreen_F1

<400> 30
gatgcaaatg acgcacggcc 20

<210> 31
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 242
ARE1KOScreen_F2

<400> 31
ccaaacgaca ctcgatagtg 20

<210> 32
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Nucleotide sequence of artificial oligonucleotide primer at 243 ARE1KOScreen_R1

<400> 32
tgctttgcgt gggagagctc 20

<210> 33
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 244 ARE1KOScreen_R2

<400> 33
cagtggttga cattctatag gcac 24

<210> 34
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 245 DGA1KOScreen_F1

<400> 34
caggcaacag acaagtcaca c 21

<210> 35
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 246 DGA1KOScreen_F2

<400> 35
catctggctc caacggccac 20

<210> 36
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 247
DGA1KOScreen_R1

<400> 36
cagctgtggg tcgttttcgc 20

<210> 37
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 248
DGA1KOScreen_R2

<400> 37
gctatttgcc taacccaggc 20

<210> 38
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 249
DGA2KOScreen_F1

<400> 38
catagctaca cccaaagtcg c 21

<210> 39
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 250
DGA2KOScreen_F2

<400> 39
cgctcagcgg actgaccatg 20

<210> 40
<211> 20

<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 251 DGA2KOScreen_R1

<400> 40
tcttggggta cacgggctgg 20

<210> 41
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 252 DGA2KOScreen_R2

<400> 41
ctgtaaaggc tactgatgag tg 22

<210> 42
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 257 LRO1KOScreen_F1

<400> 42
catgactaat cattcacgcg ac 22

<210> 43
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 258 LRO1KOScreen_F2

<400> 43
cgtggtcggt gtgatagccg 20

<210> 44
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 259 LRO1KOScreen_R1

<400> 44
gttcgcagca gatcggctcg 20

<210> 45
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 260 LRO1KOScreen_R2

<400> 45
cgaacgccct tctccatcag tg 22

<210> 46
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 270 URA3KOScrnF1

<400> 46
cgttggaggc tgtgggtctg 20

<210> 47
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 271 URA3KOScrnF2

<400> 47
agtgctcaag ctcgtggcag 20

<210> 48
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 272
URA3KOScrnR1

<400> 48
gatctcggtt ctggccgtac 20

<210> 49
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 273
URA3KOScrnR2

<400> 49
gcctccagga agtccatggg 20

<210> 50
<211> 286
<212> PRT
<213> Yarrowia lipolytica

<400> 50

Met Pro Ser Tyr Glu Ala Arg Ala Asn Val His Lys Ser Ala Phe Ala
1 5 10 15

Ala Arg Val Leu Lys Leu Val Ala Ala Lys Lys Thr Asn Leu Cys Ala
20 25 30

Ser Leu Asp Val Thr Thr Thr Lys Glu Leu Ile Glu Leu Ala Asp Lys
35 40 45

Val Gly Pro Tyr Val Cys Met Ile Lys Thr His Ile Asp Ile Ile Asp
50 55 60

```
Asp Phe Thr Tyr Ala Gly Thr Val Leu Pro Leu Lys Glu Leu Ala Leu
65                  70                  75                  80

Lys His Gly Phe Phe Leu Phe Glu Asp Arg Lys Phe Ala Asp Ile Gly
                85                  90                  95

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

Trp Ser Asp Ile Thr Asn Ala His Gly Val Pro Gly Thr Gly Ile Ile
        115                 120                 125

Ala Gly Leu Arg Ala Gly Ala Glu Glu Thr Val Ser Glu Gln Lys Lys
    130                 135                 140

Glu Asp Val Ser Asp Tyr Glu Asn Ser Gln Tyr Lys Glu Phe Leu Val
145                 150                 155                 160

Pro Ser Pro Asn Glu Lys Leu Ala Arg Gly Leu Leu Met Leu Ala Glu
                165                 170                 175

Leu Ser Cys Lys Gly Ser Leu Ala Thr Gly Glu Tyr Ser Lys Gln Thr
            180                 185                 190

Ile Glu Leu Ala Arg Ser Asp Pro Glu Phe Val Val Gly Phe Ile Ala
        195                 200                 205

Gln Asn Arg Pro Lys Gly Asp Ser Glu Asp Trp Leu Ile Leu Thr Pro
    210                 215                 220

Gly Val Gly Leu Asp Asp Lys Gly Asp Ala Leu Gly Gln Gln Tyr Arg
225                 230                 235                 240

Thr Val Glu Asp Val Met Ser Thr Gly Thr Asp Ile Ile Ile Val Gly
                245                 250                 255

Arg Gly Leu Tyr Gly Gln Asn Arg Asp Pro Ile Glu Glu Ala Lys Arg
```

260 265 270

Tyr Gln Lys Ala Gly Trp Glu Ala Tyr Gln Lys Ile Asn Cys
275 280 285

<210> 51
<211> 2861
<212> DNA
<213> Yarrowia lipolytica

<400> 51
ttaccatttt aaagcaattg ttaacaagat actactgtag aaaactgatt aaacattcgt 60

aacttgcctg tggctcgact tacagccaaa tgttcgcggc aaaatttgaa tattttgcac 120

acttattgtc ccaaaaccaa ataatgccaa caaaatccga tatagtgtag gggctatcac 180

ataccctctc ataagaatac ttttaaccgt ggagacgggg ttcgattcgt atcggagtgt 240

ttatttttg ctcaaccata ccctggggtg tgttctgtgg agcattctca cttttggtaa 300

acgacattgc ttcaagtgca gcggaatcaa aaagtataaa gtgggcagcg agtatacctg 360

tacagactgt aggcgataac tcaatccaat taccccccac aacatgactg gccaaactga 420

tctcaagact ttattgaaat cagcaacacc gattctcaat gaaggcacat acttcttctg 480

caacattcac ttgacgccta aagttggtga gaaatggacc gacaagacat attctgctat 540

ccacggactg ttgcctgtgt cggtggctac aatacgtgag tcagaagggc tgacggtggt 600

ggttcccaag gaaaaggtcg acgagtatct gtctgactcg tcattgccgc ctttggagta 660

cgactccaac tatgagtgtg cttggatcac tttgacgata cattcttcgt tggaggctgt 720

gggtctgaca gctgcgtttt cggcgcggtt ggccgacaac aatatcagct gcaacgtcat 780

tgctggcttt catcatgatc acatttttgt cggcaaaggc gacgcccaga gagccattga 840

cgttctttct aatttggacc gatagccgta tagtccagtc tatctataag ttcaactaac 900

tcgtaactat taccataaca tatacttcac tgccccagat aaggttccga taaaaagttc 960

tgcagactaa atttatttca gtctcctctt caccaccaaa atgccctcct acgaagctcg 1020

agctaacgtc cacaagtccg cctttgccgc tcgagtgctc aagctcgtgg cagccaagaa 1080

aaccaacctg tgtgcttctc tggatgttac caccaccaag gagctcattg agcttgccga 1140

taaggtcgga ccttatgtgt gcatgatcaa gacccatatc gacatcattg acgacttcac 1200

ctacgccggc actgtgctcc ccctcaagga acttgctctt aagcacggtt tcttcctgtt 1260

cgaggacaga aagttcgcag atattggcaa cactgtcaag caccagtaca agaacggtgt 1320

ctaccgaatc gccgagtggt ccgatatcac caacgcccac ggtgtacccg gaaccggaat 1380

cattgctggc ctgcgagctg gtgccgagga aactgtctct gaacagaaga aggaggacgt 1440

ctctgactac gagaactccc agtacaagga gttcctggtc ccctctccca acgagaagct 1500

ggccagaggt ctgctcatgc tggccgagct gtcttgcaag ggctctctgg ccactggcga 1560

gtactccaag cagaccattg agcttgcccg atccgacccc gagtttgtgg ttggcttcat 1620

tgcccagaac cgacctaagg gcgactctga ggactggctt attctgaccc ccggggtggg 1680

tcttgacgac aagggagacg ctctcggaca gcagtaccga actgttgagg atgtcatgtc 1740

taccggaacg gatatcataa ttgtcggccg aggtctgtac ggccagaacc gagatcctat 1800

tgaggaggcc aagcgatacc agaaggctgg ctgggaggct taccagaaga ttaactgtta 1860

gaggttagac tatggatatg taatttaact gtgtatatag agagcgtgca agtatggagc 1920

gcttgttcag cttgtatgat ggtcagacga cctgtctgat cgagtatgta tgatactgca 1980

caacctgtgt atccgcatga tctgtccaat ggggcatgtt gttgtgtttc tcgatacgga 2040

gatgctgggt acaagtagct aatacgattg aactacttat acttatatga ggcttgaaga 2100

aagctgactt gtgtatgact tattctcaac tacatcccca gtcacaatac caccactgca 2160

ctaccactac accaaaacca tgatcaaacc acccatggac ttcctggagg cagaagaact 2220

tgttatggaa aagctcaaga gagagaagcc aagatactat caagacatgt gtcgcaactt 2280

caaggaggac caagctctgt acaccgagaa acaggccttt gtcgacgaag ctctaaagct 2340

gaaacgggac atcgaaaagt tggagaagtc aatctggaaa ctacacgtag aaaacagaca 2400

ccatctccaa caagtacagg acttcaaatt tgcactgaag acattctgtg ggtggagcca 2460

taagtaacaa tcgcctgctc caggaactta aaaaagaggc attataggcc ccttcatctg 2520

aatattcttt gagtaccaac tgccgccaac tccccaatcg ttaattagtg ttatctattc 2580

tatctatgct attagcacat atttatcatg tgctgctgtt atgtacgggt tcaatataca 2640

aatgcttagt cctcaaacac cttcataggt gggtcgtctt cctggccaaa tcgaattctc 2700

atgatttcaa acacgaacag accgatagac gtagacggaa gcgcactgag ggtcgtccag 2760

ccacaaccat ggtataggta tctgggccaa ctcaaccgca gacccttttcg gatgttattg 2820

gctgtatgca gagtctcaac gtaggctcgg tagtagagat g 2861

<210> 52
<211> 3545
<212> DNA
<213> Yarrowia lipolytica

<400> 52
atgctgcggg cggatcctgg tgcatttttg cttgcgattg ttgtgaaaat tagcacggtt 60

atattagccc gtaataaatg cccgtctcca tcttcatatg gccatcaccc cgcaaatagc 120

cggccaatca ggcgtatgtc acctgttgct cacacgcatg tcctcggacc gttgtattgt 180

gcaagtaggg gtacctcccc gatccatcct cgaccagtgg cacgctcaac cccatggttc 240

gcttttctct tttcgtctat cctgaactga gtttttttcc acgccaactg atatcccctt 300

acgttacccc ctcatcacct ggtgaggcga aactgtaagg tgaaagctaa aaatgacatc 360

tcagctgcac gaaggaccgg ggcttaaaag acgggctggt gcttgtgatt taaaactgga 420

caaatctcag cttgcttgaa attttggtct ccaactgttt ccgagcgaat cgcacacaaa 480

ccgggcttct ctctgcagac cacgcccccg aaactctttc tcccaccacc accaacactc 540

cctttccatt cccacaccgt tcctctctcg tgcgtcatcc ttgcgcaatc atcttcgtct 600

gcgacatatt gtacgacata cagtaccacg gaacgtttca gaccgtcgac gtgaacacat 660

cttaggaaca gcaacctgag ctacagaaat ctatctatag gcggataaaa aaacgcaccc 720

actgctcgtc ctccttgctc ctcgaaaccg actcctctac acacgtcaaa tccgaggttg 780

aaatcttccc cacatttggc agccaaacca gcacatccca gcaacctcgc acagcgccga 840

aatcgacctg tcgacttggc cacaaaaaaa agcaccggct ctgcaacagt tctcacgacc 900

aattacgtac aagtacgaaa tcgttcgtgg accgtgactg ataagctccc acttttttctt 960

ctaacaacag gcaacagaca agtcacacaa aacaaaagct atgactatcg actcacaata 1020

ctacaagtcg cgagacaaaa acgacacggc acccaaaatc gcgggaatcc gatatgcccc 1080

gctatcgaca ccattactca accgatgtga gaccttctct ctggtctggc acattttcag 1140

cattcccact ttcctcacaa ttttcatgct atgctgcgca attccactgc tctggccatt 1200

tgtgattgcg tatgtagtgt acgctgttaa agacgactcc ccgtccaacg gaggagtggt 1260

caagcgatac tcgcctattt caagaaactt cttcatctgg aagctctttg gccgctactt 1320

ccccataact ctgcacaaga cggtggatct ggagcccacg cacacatact accctctgga 1380

cgtccaggag tatcacctga ttgctgagag atactggccg cagaacaagt acctccgagc 1440

aatcatctcc accatcgagt actttctgcc cgccttcatg aaacggtctc tttctatcaa 1500

cgagcaggag cagcctgccg agcgagatcc tctcctgtct cccgtttctc ccagctctcc 1560

gggttctcaa cctgacaagt ggattaacca cgacagcaga tatagccgtg gagaatcatc 1620

tggctccaac ggccacgcct cgggctccga acttaacggc aacggcaaca atggcaccac 1680

taaccgacga cctttgtcgt ccgcctctgc tggctccact gcatctgatt ccacgcttct 1740

taacgggtcc ctcaactcct acgccaacca gatcattggc gaaaacgacc cacagctgtc 1800

gcccacaaaa ctcaagccca ctggcagaaa atacatcttc ggctaccacc cccacggcat 1860

tatcggcatg ggagcctttg gtggaattgc caccgaggga gctggatggt ccaagctctt 1920

tccgggcatc cctgtttctc ttatgactct caccaacaac ttccgagtgc ctctctacag 1980

agagtacctc atgagtctgg gagtcgcttc tgtctccaag aagtcctgca aggccctcct 2040

caagcgaaac cagtctatct gcattgtcgt tggtggagca caggaaagtc ttctggccag 2100

acccggtgtc atggacctgg tgctactcaa gcgaaagggt tttgttcgac ttggtatgga 2160

ggtcggaaat gtcgcccttg ttcccatcat ggcctttggt gagaacgacc tctatgacca 2220

ggttagcaac gacaagtcgt ccaagctgta ccgattccag cagtttgtca agaacttcct 2280

tggattcacc cttcctttga tgcatgcccg aggcgtcttc aactacgatg tcggtcttgt 2340

cccctacagg cgacccgtca acattgtggt tggttccccc attgacttgc cttatctccc 2400

acaccccacc gacgaagaag tgtccgaata ccacgaccga tacatcgccg agctgcagcg 2460

aatctacaac gagcacaagg atgaatattt catcgattgg accgaggagg gcaaaggagc 2520

cccagagttc cgaatgattg agtaaggaaa actgcctggg ttaggcaaat agctaatgag 2580

tatttttttg atggcaacca aatgtagaaa gaaaaaaaaa aaaaaagaaa aaaaaaagag 2640

aatattatat ctatgtaatt ctattaaaag ctctgttgag tgagcggaat aaatactgtt 2700

gaagagggga ttgtgtagag atctgtttac tcaatggcaa actcatctgg gggagatcct 2760

tccactgtgg gaagctcctg gatagccttt gcatcgggtt caagaagacc attgtgaaca 2820

gccttgacac tgtcgacaat gacagcatca acgccatgct tgatctgggt cttggcaatc 2880

acagggttgt tattcagaac accgtaggag acacagacca gcccagcccg cttgacgacc 2940

tggataagtc gagggcacat gacaagggga ctacactcgg acacaagacc caaaaggttc 3000

caatccttgc agaaccgaat ggcctcctgg agagaagagg ctcgaatgtc atacatgggg 3060

gcagttccag actcggtgag gaacagggtg gggtatccga cagggttgac ttcagtcttg 3120

ccaatcatag ggacacctcc ttgcttgtgg ttcagcatga ggcacacgtc agggtggaaa 3180

gaggagaaga taatctctcg gtttgctcca tactcaaaga cacactcgag gaccttgtcc 3240

gcccaaacgt tgaggtcaag ggcaaactct tccatctcct cagcttgtcc ttcgtcaagc 3300

atggggtact tgcactcgat gttgcagcca acgttgggag ggagagactg tagaacctct 3360

cgtagagtgg tgaaaggggc ctgaatggaa ataccacgag aattaccctt gaaacccttc 3420

atcttgtagt cacgagtgta cttcattcga tgagtgatgt ggtcatctgc gtcagggtca 3480

tcttgctgaa gtgcagttcc ttcaagtatc ttgtcgccct cagaaactcg ctccaggcca 3540

tcaga 3545

<210> 53
<211> 514
<212> PRT
<213> Yarrowia lipolytica

<400> 53

Met Thr Ile Asp Ser Gln Tyr Tyr Lys Ser Arg Asp Lys Asn Asp Thr
1 5 10 15

Ala Pro Lys Ile Ala Gly Ile Arg Tyr Ala Pro Leu Ser Thr Pro Leu
20 25 30

```
Leu Asn Arg Cys Glu Thr Phe Ser Leu Val Trp His Ile Phe Ser Ile
        35                  40                  45

Pro Thr Phe Leu Thr Ile Phe Met Leu Cys Cys Ala Ile Pro Leu Leu
    50                  55                  60

Trp Pro Phe Val Ile Ala Tyr Val Val Tyr Ala Val Lys Asp Asp Ser
65                  70                  75                  80

Pro Ser Asn Gly Gly Val Val Lys Arg Tyr Ser Pro Ile Ser Arg Asn
                85                  90                  95

Phe Phe Ile Trp Lys Leu Phe Gly Arg Tyr Phe Pro Ile Thr Leu His
            100                 105                 110

Lys Thr Val Asp Leu Glu Pro Thr His Thr Tyr Tyr Pro Leu Asp Val
        115                 120                 125

Gln Glu Tyr His Leu Ile Ala Glu Arg Tyr Trp Pro Gln Asn Lys Tyr
    130                 135                 140

Leu Arg Ala Ile Ile Ser Thr Ile Glu Tyr Phe Leu Pro Ala Phe Met
145                 150                 155                 160

Lys Arg Ser Leu Ser Ile Asn Glu Gln Glu Gln Pro Ala Glu Arg Asp
                165                 170                 175

Pro Leu Leu Ser Pro Val Ser Pro Ser Ser Pro Gly Ser Gln Pro Asp
            180                 185                 190

Lys Trp Ile Asn His Asp Ser Arg Tyr Ser Arg Gly Glu Ser Ser Gly
        195                 200                 205

Ser Asn Gly His Ala Ser Gly Ser Glu Leu Asn Gly Asn Gly Asn Asn
    210                 215                 220
```

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

Ala Ser Asp Ser Thr Leu Leu Asn Gly Ser Leu Asn Ser Tyr Ala Asn
                245                 250                 255

Gln Ile Ile Gly Glu Asn Asp Pro Gln Leu Ser Pro Thr Lys Leu Lys
            260                 265                 270

Pro Thr Gly Arg Lys Tyr Ile Phe Gly Tyr His Pro His Gly Ile Ile
        275                 280                 285

Gly Met Gly Ala Phe Gly Gly Ile Ala Thr Glu Gly Ala Gly Trp Ser
    290                 295                 300

Lys Leu Phe Pro Gly Ile Pro Val Ser Leu Met Thr Leu Thr Asn Asn
305                 310                 315                 320

Phe Arg Val Pro Leu Tyr Arg Glu Tyr Leu Met Ser Leu Gly Val Ala
                325                 330                 335

Ser Val Ser Lys Lys Ser Cys Lys Ala Leu Leu Lys Arg Asn Gln Ser
            340                 345                 350

Ile Cys Ile Val Val Gly Gly Ala Gln Glu Ser Leu Leu Ala Arg Pro
        355                 360                 365

Gly Val Met Asp Leu Val Leu Leu Lys Arg Lys Gly Phe Val Arg Leu
    370                 375                 380

Gly Met Glu Val Gly Asn Val Ala Leu Val Pro Ile Met Ala Phe Gly
385                 390                 395                 400

Glu Asn Asp Leu Tyr Asp Gln Val Ser Asn Asp Lys Ser Ser Lys Leu
                405                 410                 415

Tyr Arg Phe Gln Gln Phe Val Lys Asn Phe Leu Gly Phe Thr Leu Pro
            420                 425                 430
```

Leu Met His Ala Arg Gly Val Phe Asn Tyr Asp Val Gly Leu Val Pro
435 440 445

Tyr Arg Arg Pro Val Asn Ile Val Val Gly Ser Pro Ile Asp Leu Pro
450 455 460

Tyr Leu Pro His Pro Thr Asp Glu Glu Val Ser Glu Tyr His Asp Arg
465 470 475 480

Tyr Ile Ala Glu Leu Gln Arg Ile Tyr Asn Glu His Lys Asp Glu Tyr
485 490 495

Phe Ile Asp Trp Thr Glu Glu Gly Lys Gly Ala Pro Glu Phe Arg Met
500 505 510

Ile Glu

<210> 54
<211> 3581
<212> DNA
<213> Yarrowia lipolytica

<400> 54
tgggagtgta tttggaaaat gacttggcct ggaccccgac taccgctagt gctacccatg 60

gggtgtctcc agaacccttg tacatcagcc tatatgcgtg tgtgtctacg tctactgtac 120

aaatctgctt ggcgtcaaat ggcactacaa gtacaaagac caaatgccac aagttcagta 180

cgagtacaag agtctgctac ttgtagctcc cgcaggtatt tagttcatac aaacgagtat 240

acttgtagct gtccaacagt agttttcaga cacctgatca gaaataggaa cacgctccac 300

ggacctcgta tccactctaa gaccataata tatgggtttc tggaacattg gtgcccctca 360

cctctcatct ctttttttt tacctctctg accaccacca cccttgttgt tccgagtggt 420

tagctgcacc ccacctagca aactttaggg ccccatacgc ataagcatac gtccggtgtg 480

cgacttgtgt gtgaacgcgg tccacacaag gtgtctttca agtccacaca gggctttgtc 540

tagagcctgg gctatcttcc cacgtttgta tatatagact ggatccacac ccagcaacgc 600

aacgtcctcc cacccaaccc tatcctccac gttacacgcc caagatcttc caaacaagaa 660

gacacattcc acagcaagaa taaccttggc aagaacgtgc gacgaggtga ataaaagagc 720

ccagagtact aacattttgt ggacctgttg gaaactaatt gccgccgcaa ctcgactgga 780

accaagttca agaccactga caagcgcccc ggtccaacca gaacactaaa cggagccgca 840

cagcaagcac cacggacgat cgtcatcatc cctatggcca acctgcatgg atactcttga 900

taacagttaa gctatattta ccaacaccac ctccgttcac gacataacat ccccacacca 960

tagctacacc caaagtcgct gtacccgtac cgcccgcaaa atggaagtcc gacgacgaaa 1020

aatcgacgtg ctcaaggccc agaaaaacgg ctacgaatcg ggcccaccat ctcgacaatc 1080

gtcgcagccc tcctcaagag catcgtccag aacccgcaac aaacactcct cgtccaccct 1140

gtcgctcagc ggactgacca tgaaagtcca gaagaaacct gcgggacccc cggcgaactc 1200

caaaacgcca ttcctacaca tcaagcccgt gcacacgtgc tgctccacat caatgctttc 1260

gcgcgattat gacggctcca accccagctt caagggcttc aaaaacatcg gcatgatcat 1320

tctcattgtg ggaaatctac ggctcgcatt cgaaaactac ctcaaatacg gcatttccaa 1380

cccgttcttc gaccccaaaa ttactccttc cgagtggcag ctctcaggct tgctcatagt 1440

cgtggcctac gcacatatcc tcatggccta cgctattgag agcgctgcca agctgctgtt 1500

cctctctagc aaacaccact acatggccgt ggggcttctg cataccatga acactttgtc 1560

gtccatctcg ttgctgtcct acgtcgtcta ctactacctg cccaaccccg tggcaggcac 1620

aatagtcgag tttgtggccg ttattctgtc tctcaaactc gcctcatacg ccctcactaa 1680

ctcggatctc cgaaaagccg caattcatgc ccagaagctc gacaagacgc aagacgataa 1740

cgaaaaggaa tccacctcgt cttcctcttc ttcagatgac gcagagactt tggcagacat 1800

tgacgtcatt cctgcatact acgcacagct gccctacccc cagaatgtga cgctgtcgaa 1860

cctgctgtac ttctggtttg ctcccacact ggtctaccag cccgtgtacc ccaagacgga 1920

gcgtattcga cccaagcacg tgatccgaaa cctgtttgag ctcgtctctc tgtgcatgct 1980

tattcagttt ctcatcttcc agtacgccta ccccatcatg cagtcgtgtc tggctctgtt 2040

cttccagccc aagctcgatt atgccaacat ctccgagcgc ctcatgaagt tggcctccgt 2100

gtctatgatg gtctggctca ttggattcta cgctttcttc cagaacggtc tcaatcttat 2160

tgccgagctc acctgttttg gaaacagaac cttctaccag cagtggtgga attcccgctc 2220

cattggccag tactggactc tatggaacaa gccagtcaac cagtacttta gacaccacgt 2280

ctacgtgcct cttctcgctc ggggcatgtc gcggttcaat gcgtcggtgg tggttttctt 2340

tttctccgcc gtcatccatg aactgcttgt cggcatcccc actcacaaca tcatcggagc 2400

cgccttcttc ggcatgatgt cgcaggtgcc tctgatcatg gctactgaga accttcagca 2460

tattaactcc tctctgggcc ccttccttgg caactgtgca ttctggttca ccttttttcct 2520

gggacaaccc acttgtgcat tcctttatta tctggcttac aactacaagc agaaccagta 2580

gcataacact catcagtagc ctttacagtg ataaataata ttgacattga cgtaatctga 2640

tgattgtaga caaacgtgag acggcattag cgataatatt ttgacgatgc ccacgggcat 2700

ataataacca ttctgtaata gaactttgag catgttcaga agtaatgagt ttcatacata 2760

ggagtacaga tcgtggcgat gtgtaattac cggtaccgca tatgttgaac ctgttcgtaa 2820

aaggtgcctt caacacagtt ttgttgccgt cttacctgtg tataagcgtt cagcttctgg 2880

tgtgtagcag aaggaggtga agttgaggaa caaataatgg agacttgcag cagactccgt 2940

gcctcaagca aaaaagatga tccagcagag gatcgaactc tgagcctcct cggcactcgt 3000

agattgaagc cactttgtga gcgagacgtg ataaccatta caccactaga ccgggggtga 3060

agatcagagc ccgtattgtc gtattcaaag ttaaaaatgg ggatttttta tattcagaag 3120

tgtgtgattt attgaacgcc aataacccgg cttactacag taggtatata gctgtaatga 3180

atcattagaa taaaaaaaaa tccgcctcaa agtagctgaa atgtatccct agaataatga 3240

ggggaaaaaa ctgataaatc aggtgaaata cggcatttag ggacaggtaa gagttcaatt 3300

gaatttatgt cgtattattt aagctgtcgg agaaaacttg tgtgcttttt tgtatggttt 3360

tcgtgactgt gtacatttct gtacccctga ttccgaacat cttatactga cacaagacat 3420

tggagataaa tcatacagta taaagtagct acgtcttact tgtacagtgc ttttaattac 3480

aatatgcccc actatacata catgtgattc attttataaa cccctcctca ggctaattac 3540

ttctaatgcg tttcgtatat attatctatg gaattacaag a 3581

<210> 55
<211> 526
<212> PRT
<213> Yarrowia lipolytica

<400> 55

Met Glu Val Arg Arg Arg Lys Ile Asp Val Leu Lys Ala Gln Lys Asn
1 5 10 15

Gly Tyr Glu Ser Gly Pro Pro Ser Arg Gln Ser Ser Gln Pro Ser Ser
20 25 30

Arg Ala Ser Ser Arg Thr Arg Asn Lys His Ser Ser Ser Thr Leu Ser
35 40 45

Leu Ser Gly Leu Thr Met Lys Val Gln Lys Lys Pro Ala Gly Pro Pro
50 55 60

Ala Asn Ser Lys Thr Pro Phe Leu His Ile Lys Pro Val His Thr Cys
65 70 75 80

Cys Ser Thr Ser Met Leu Ser Arg Asp Tyr Asp Gly Ser Asn Pro Ser
85 90 95

Phe Lys Gly Phe Lys Asn Ile Gly Met Ile Ile Leu Ile Val Gly Asn
100 105 110

Leu Arg Leu Ala Phe Glu Asn Tyr Leu Lys Tyr Gly Ile Ser Asn Pro
115 120 125

Phe Phe Asp Pro Lys Ile Thr Pro Ser Glu Trp Gln Leu Ser Gly Leu
130 135 140

Leu Ile Val Val Ala Tyr Ala His Ile Leu Met Ala Tyr Ala Ile Glu
145 150 155 160

```
Ser Ala Ala Lys Leu Leu Phe Leu Ser Ser Lys His His Tyr Met Ala
                165                 170                 175

Val Gly Leu Leu His Thr Met Asn Thr Leu Ser Ser Ile Ser Leu Leu
            180                 185                 190

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

Val Glu Phe Val Ala Val Ile Leu Ser Leu Lys Leu Ala Ser Tyr Ala
    210                 215                 220

Leu Thr Asn Ser Asp Leu Arg Lys Ala Ala Ile His Ala Gln Lys Leu
225                 230                 235                 240

Asp Lys Thr Gln Asp Asp Asn Glu Lys Glu Ser Thr Ser Ser Ser Ser
                245                 250                 255

Ser Ser Asp Asp Ala Glu Thr Leu Ala Asp Ile Asp Val Ile Pro Ala
            260                 265                 270

Tyr Tyr Ala Gln Leu Pro Tyr Pro Gln Asn Val Thr Leu Ser Asn Leu
        275                 280                 285

Leu Tyr Phe Trp Phe Ala Pro Thr Leu Val Tyr Gln Pro Val Tyr Pro
    290                 295                 300

Lys Thr Glu Arg Ile Arg Pro Lys His Val Ile Arg Asn Leu Phe Glu
305                 310                 315                 320

Leu Val Ser Leu Cys Met Leu Ile Gln Phe Leu Ile Phe Gln Tyr Ala
                325                 330                 335

Tyr Pro Ile Met Gln Ser Cys Leu Ala Leu Phe Phe Gln Pro Lys Leu
            340                 345                 350

Asp Tyr Ala Asn Ile Ser Glu Arg Leu Met Lys Leu Ala Ser Val Ser
        355                 360                 365
```

```
Met Met Val Trp Leu Ile Gly Phe Tyr Ala Phe Phe Gln Asn Gly Leu
    370                 375                 380

Asn Leu Ile Ala Glu Leu Thr Cys Phe Gly Asn Arg Thr Phe Tyr Gln
385                 390                 395                 400

Gln Trp Trp Asn Ser Arg Ser Ile Gly Gln Tyr Trp Thr Leu Trp Asn
                405                 410                 415

Lys Pro Val Asn Gln Tyr Phe Arg His His Val Tyr Val Pro Leu Leu
            420                 425                 430

Ala Arg Gly Met Ser Arg Phe Asn Ala Ser Val Val Val Phe Phe Phe
        435                 440                 445

Ser Ala Val Ile His Glu Leu Leu Val Gly Ile Pro Thr His Asn Ile
    450                 455                 460

Ile Gly Ala Ala Phe Phe Gly Met Met Ser Gln Val Pro Leu Ile Met
465                 470                 475                 480

Ala Thr Glu Asn Leu Gln His Ile Asn Ser Ser Leu Gly Pro Phe Leu
                485                 490                 495

Gly Asn Cys Ala Phe Trp Phe Thr Phe Phe Leu Gly Gln Pro Thr Cys
            500                 505                 510

Ala Phe Leu Tyr Tyr Leu Ala Tyr Asn Tyr Lys Gln Asn Gln
        515                 520                 525
```

<210> 56
<211> 3947
<212> DNA
<213> Yarrowia lipolytica

<400> 56
gagtcgactt tccgatacag aagtggaacc agacttgctc cacattcatc tttcccccat 60

ttttctatag tggttcagtg gagttgacta cttgataaca attggtacat gtaatactga 120

tacatctccg tggatatacc atacacattc agtgaccaag cagcgtcgaa catgaccttc 180

aagaatgggt tccctgtcat ctggagggat gacggctctg gtttgtcaag tctgctgaat 240

gcatctactt gtacgttgta gtaacttcct ctgtcctccc tagaaacgac ccttcaatgc 300

agcacagacg ataaatcagc gtatatcgtc cagtccaact gtgagaccta cgatgaccta 360

gagaatggtt ctcttaatct ggagtctgct tcttctccac acaactgaca gcgtatcccc 420

catggcaccg aagctccagt actgaccgac ctttgctaca agtgctgtgt ccagagttct 480

ctagtaggta cgatacgaca gcagcttgat actgtgcttg gacctaccct tgcatgaggt 540

cttgctgttg tccactgaca caggaccatt ctgcggggtc cactgaaaca agaccattct 600

gacgcgattt ctttgtacag atgaagtcag aaatacaagc aatggcacca ctcttttggt 660

tcgaccatga ctgtttttac atccatttcc agtaccaacc gcacttttt catgttcgcc 720

ccacatttat taatattatg ctcttcatgc accagcaaaa taaccgaaac gcgcatatga 780

tagtgggatt ctcgatttgc ccggcagaca aacgccgcta aaatcgccac agtatcgaat 840

tttaattgaa tacgaacgtc aattccggct tatccttcta gcagttgtct cccgcagctc 900

gctccatgac taatcattca cgcgacatgt ctcagctacc ccggtctggc tcatgtaaaa 960

aaagtgtaat cggctttttt ccggttgatc acaaccatca atgacacaac ctgtgaatcg 1020

gaaggcgact gtcgagcggg tcgagccagc agtggaggtg gctgactccg agtccgaggc 1080

caagaccgac gtccacgttc accaccatca tcaccaccac aagcgaaaat ccgtcaaggg 1140

caagattctc aacttcttca cccgaagtcg acgtatcacc ttcgtcctcg gcgccgtggt 1200

cggtgtgata gccgcgggat actacgctgc gccaccggag ctcagcattg atatcgatgc 1260

tcttctcggc gacttgccct cgttcgactt tgacgctcta tctctcgaca acttgtccat 1320

ggacagtgtg tcggactttg tacaagacat gaaatcgcgg tttccgacca agattctgca 1380

ggaggcggcc aagatcgaga agcaccagaa aagcgaacag aaggctgccc cttttgctgt 1440

gggcaaggct atgaagagcg agggactcaa cgccaagtac ccggtggtgc tggtgcccgg 1500

cgtcatctcc acgggactgg agagctggtc cctggaggga accgaggagt gtcccaccga 1560

gtcgcacttc agaaagcgaa tgtggggctc ctggtacatg atccgagtca tgctgctgga 1620

caagtactgc tggctgcaga acctgatgct ggacacagag accggtctag accctcccca 1680

tttcaagctg cgagccgccc agggatttgc ctccgccgac ttctttatgg caggctactg 1740

gctgtggaac aagctgctcg agaacctggc tgttattgga tacgatacgg atacaatgtc 1800

tgctgcggcg tacgactgga gactgtccta ccctgatttg gagcaccgag acggatactt 1860

ctccaagctc aaagcttcaa tcgaagagac taagcgtatg acaggtgaga agacagttct 1920

gacgggccat tccatgggct cccaggtcat cttctacttc atgaagtggg ctgaggccga 1980

gggatatgga ggaggaggtc ccaactgggt caatgaccat attgaatcct ttgtcgacat 2040

ttccggctcc atgctgggta ctcccaagac cctggttgct cttctgtctg gagaaatgaa 2100

ggataccgtg cagctgaacg cgatggctgt gtatggactg gagcagttct tctctcgacg 2160

agagcgagcc gatctgctgc gaacatgggg aggaattgct tccatgattc ccaagggtgg 2220

taaggctatc tggggtgatc attctggagc ccctgatgac gagcctggcc agaatgtcac 2280

ctttggcaac ttcatcaagt tcaaggagtc cttgaccgag tactctgcta agaacctcac 2340

catggatgaa accgttgact tcctgtattc tcagtctccc gagtggtttg tgaaccgaac 2400

cgagggtgct tactcctttg gaattgccaa gactcgaaag caggttgagc agaatgagaa 2460

gcgaccttct acctggagca accctctgga agctgctctc cccaatgccc ccgatctcaa 2520

gatctactgc ttctatggag tcggtaagga taccgagcga gcctactact accaggatga 2580

gcccaatccc gagcagacca acttgaacgt cagtatcgct ggaaacgacc ctgatggtgt 2640

gcttatgggc cagggcgatg gaaccgtctc ccttgtgacc cataccatgt gtcaccgatg 2700

gaaggacgag aattccaagt tcaaccctgg taacgcccag gtcaaggttg tggagatgtt 2760

gcaccagcct gatcgacttg atattcgagg cggtgctcag actgccgagc atgtggacat 2820

tctggggcgt tctgagttga acgagatggt tctgaaggtg gccagtggaa agggaaatga 2880

gattgaagag agagtcatct ccaacattga tgagtgggtg tggaagattg atctcggcag 2940

caattagaga gtccgttttg tagagtaata tgttttgtat atcacactga tggagaaggg 3000

cgttcgtcct ctcatgattc catgtggttg tttaatgaac acgtagaacg acgaacgacg 3060

aacgacgaac gacgaacgaa cgagcaacaa acgtcatcaa cgcgtgtttt cccgcaaaat 3120

gaagaatgtt tcccccacaa tatgcgcctt tctccagatc tgcgcgttct gataaagaat 3180

atatttcaat tgtttggaat aaatcgggaa aattcaggag aaaagaggag aaacagggaa 3240

atcactgaaa tctactcctg aaaatgttcg acaaaaaacg aattgcttcc aatatacaac 3300

cactcctgtc acctcctcaa caaaatatat tttactccgc caccaatgcc aacccatggt 3360

gacgccaaga ttcacaagac tgcttggagc ggctgaaaag agtgttttgc ggtcttctaa 3420

gcgaccccga gccgctcaaa gacccattgg aggagatctt ctttggccat ttcggttctg 3480

ttggctatca tccgcctcca aaaagcaacc ccatttgctt ctttttccaa cccgttaaaa 3540

tacctcatct cgcttcacat tccagggttt cattttaagc gtagcaccat gcaaagtgca 3600

ggtgtagggc gtacaataaa gcctgtgaga cactctttga caaggacagg acgctcgtat 3660

ttttgagacg ttgatgttgg gggctagtgg ggcgatttgg gggcaataaa cggggtaaat 3720

ctgtgctttt gaggctgtat ttttgattcc agccttttt aatttattgt tatctgcttt 3780

actacaaatt tctgccgtat ttttagcccc tcaatttcgt ctgtcacttc tcatgtcgca 3840

caaggctgtg ccatcttgca ttaggttatc caaccctact gtctcttctc tgtctgtctg 3900

gaatctgaaa gcaagccgtg cgacttttg caacttgcga ggtccgt 3947

<210> 57
<211> 648
<212> PRT
<213> Yarrowia lipolytica

<400> 57

Met Thr Gln Pro Val Asn Arg Lys Ala Thr Val Glu Arg Val Glu Pro
1 5 10 15

Ala Val Glu Val Ala Asp Ser Glu Ser Glu Ala Lys Thr Asp Val His
20 25 30

Val His His His His His His His Lys Arg Lys Ser Val Lys Gly Lys
35 40 45

```
Ile Leu Asn Phe Phe Thr Arg Ser Arg Arg Ile Thr Phe Val Leu Gly
    50                  55                  60

Ala Val Val Gly Val Ile Ala Ala Gly Tyr Tyr Ala Ala Pro Pro Glu
65                  70                  75                  80

Leu Ser Ile Asp Ile Asp Ala Leu Leu Gly Asp Leu Pro Ser Phe Asp
                85                  90                  95

Phe Asp Ala Leu Ser Leu Asp Asn Leu Ser Met Asp Ser Val Ser Asp
            100                 105                 110

Phe Val Gln Asp Met Lys Ser Arg Phe Pro Thr Lys Ile Leu Gln Glu
        115                 120                 125

Ala Ala Lys Ile Glu Lys His Gln Lys Ser Glu Gln Lys Ala Ala Pro
    130                 135                 140

Phe Ala Val Gly Lys Ala Met Lys Ser Glu Gly Leu Asn Ala Lys Tyr
145                 150                 155                 160

Pro Val Val Leu Val Pro Gly Val Ile Ser Thr Gly Leu Glu Ser Trp
                165                 170                 175

Ser Leu Glu Gly Thr Glu Glu Cys Pro Thr Glu Ser His Phe Arg Lys
            180                 185                 190

Arg Met Trp Gly Ser Trp Tyr Met Ile Arg Val Met Leu Leu Asp Lys
        195                 200                 205

Tyr Cys Trp Leu Gln Asn Leu Met Leu Asp Thr Glu Thr Gly Leu Asp
    210                 215                 220

Pro Pro His Phe Lys Leu Arg Ala Ala Gln Gly Phe Ala Ser Ala Asp
225                 230                 235                 240

Phe Phe Met Ala Gly Tyr Trp Leu Trp Asn Lys Leu Leu Glu Asn Leu
                245                 250                 255
```

```
Ala Val Ile Gly Tyr Asp Thr Asp Thr Met Ser Ala Ala Ala Tyr Asp
            260                 265                 270

Trp Arg Leu Ser Tyr Pro Asp Leu Glu His Arg Asp Gly Tyr Phe Ser
        275                 280                 285

Lys Leu Lys Ala Ser Ile Glu Glu Thr Lys Arg Met Thr Gly Glu Lys
    290                 295                 300

Thr Val Leu Thr Gly His Ser Met Gly Ser Gln Val Ile Phe Tyr Phe
305                 310                 315                 320

Met Lys Trp Ala Glu Ala Glu Gly Tyr Gly Gly Gly Gly Pro Asn Trp
                325                 330                 335

Val Asn Asp His Ile Glu Ser Phe Val Asp Ile Ser Gly Ser Met Leu
            340                 345                 350

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

Thr Val Gln Leu Asn Ala Met Ala Val Tyr Gly Leu Glu Gln Phe Phe
    370                 375                 380

Ser Arg Arg Glu Arg Ala Asp Leu Leu Arg Thr Trp Gly Gly Ile Ala
385                 390                 395                 400

Ser Met Ile Pro Lys Gly Gly Lys Ala Ile Trp Gly Asp His Ser Gly
                405                 410                 415

Ala Pro Asp Asp Glu Pro Gly Gln Asn Val Thr Phe Gly Asn Phe Ile
            420                 425                 430

Lys Phe Lys Glu Ser Leu Thr Glu Tyr Ser Ala Lys Asn Leu Thr Met
        435                 440                 445
```

```
Asp Glu Thr Val Asp Phe Leu Tyr Ser Gln Ser Pro Glu Trp Phe Val
    450                 455                 460

Asn Arg Thr Glu Gly Ala Tyr Ser Phe Gly Ile Ala Lys Thr Arg Lys
465                 470                 475                 480

Gln Val Glu Gln Asn Glu Lys Arg Pro Ser Thr Trp Ser Asn Pro Leu
                485                 490                 495

Glu Ala Ala Leu Pro Asn Ala Pro Asp Leu Lys Ile Tyr Cys Phe Tyr
            500                 505                 510

Gly Val Gly Lys Asp Thr Glu Arg Ala Tyr Tyr Tyr Gln Asp Glu Pro
        515                 520                 525

Asn Pro Glu Gln Thr Asn Leu Asn Val Ser Ile Ala Gly Asn Asp Pro
    530                 535                 540

Asp Gly Val Leu Met Gly Gln Gly Asp Gly Thr Val Ser Leu Val Thr
545                 550                 555                 560

His Thr Met Cys His Arg Trp Lys Asp Glu Asn Ser Lys Phe Asn Pro
                565                 570                 575

Gly Asn Ala Gln Val Lys Val Val Glu Met Leu His Gln Pro Asp Arg
            580                 585                 590

Leu Asp Ile Arg Gly Gly Ala Gln Thr Ala Glu His Val Asp Ile Leu
        595                 600                 605

Gly Arg Ser Glu Leu Asn Glu Met Val Leu Lys Val Ala Ser Gly Lys
    610                 615                 620

Gly Asn Glu Ile Glu Glu Arg Val Ile Ser Asn Ile Asp Glu Trp Val
625                 630                 635                 640

Trp Lys Ile Asp Leu Gly Ser Asn
                645
```

<210> 58
<211> 3632
<212> DNA
<213> Yarrowia lipolytica

<400> 58
gccgacaatc tctcttcctc attaccgcgc tttaatagcg aggagctatt agtttgattt 60

catccaaacc ccgcgggacg tattggcggg ccaaccaaaa aaagagtctt catttagcgc 120

gaaaaaaaca agaggttgaa ggcaatattt ctgctcaatt agacgacaca tgatctcttg 180

tccaacagca atacggttcc tggacgccgc gggtcacgtg aagggtatcc ccggcttgtg 240

gatctctgga cacctgtgga tctatcgacc ctatggactt agctggtctg tgatcatgcg 300

atgaggccaa gacttttgtg cagatgaatc atgtgtctgg gtgagacttg ggtcggtctt 360

gtagcaaggc tggcggggcg ccacaccgac ggcgaggaag ggggggttct cgttgactcc 420

ggcctccgac tgcgagctgt gaaatcgcgc gaccgttcgc tggtgcctta cagatgcatc 480

taaaaaaagc ctcgttgttg ctgggtaagt caaacctcgg taagcacgat gggctgccca 540

gagagagcga gtgcactgca gcccggaagg ttgactgaag aaatcttata ccgtttatct 600

tctttaattt ttttgtcagc gcctttaaac tcggcggctt aattagtgac cgttttgggc 660

acctaaaatg gcgcacgagg ttattgttga ggtcagtagg tgtgagagag tctgagggcg 720

ggggccggac aagtgctcac ttatatagtt gtctcccatg acacgggtct ggtaacacgc 780

cggcgcctga cacggctcag atttggttca gcttggcaca atttggtttc aagcttgcac 840

tttagcagtt tgtacactgg cagcaagttt gaatttcagt gcatgagaac tctatgtata 900

agatgcaaat gacgcacggc cctacatata gcggggtcta tcgtggcgac ggggcaaaaa 960

ggatacaagt cgttcttaca agacactctc cgcacacaca atggccacac tccaccccga 1020

agacgccgca ggacggcccg tgcgacgacg acctcgtccc tccagttcgg gcggctccag 1080

atcgccgtcc accaaacgac actcgatagt gcgggagcat ctcggagaag agctcaatgt 1140

gcccgacggc caggaaatgg acctgggcca ggtcaacaag aacctcaatg ccgcatacgc 1200

caaggccgag aaggactcgg acgacgagaa ggaaaagaag gaggagggcg tggtggacga 1260

```
gctgccagag aagtattcct accctcgatt ctcaaagaac aaccgacgct acagattcac     1320

cgacatcaag ttcaagccaa caccgtcgat tctcgacaag ttcgcccaca aggactcgga     1380

gttctttggc ttctacaccc tgctgtggat ggtgtttgcc ttctgcgtct tccgaaccgg     1440

cctgctcaac tacacaaacg aaggcatcct gttccggggc cagattttcg ccattctcag     1500

caaagatctc tggaaagtcg cattggtcga tctgggcatg tacctgacca cctatctgtc     1560

tgtgtttctg caattggccg tcaagcacgg tctggtcgac tggaactcgt ttggctggat     1620

catccagaac gtgcaccaga ccctgttcct cttcttctac ctttgggtcg ccaagtcgag     1680

taacctgcct tggatcggta acatcttcat tgtgcttcat gcctttgtca tgctcatgaa     1740

acaacactcg tacgccttct acaatggcta cctatggact gtcgaggacg agctctccca     1800

cgcaaagcag cgtctcaccg aagacattcc tgtttcagag aaggaggatc tcaagctgga     1860

catcgagttc tgcgagacag agctcaaggt ccaatccaga cacacccctt tccccaccaa     1920

catcaccttt tctaactact tctggtactc catgttccca acgctcgtct acgaaattga     1980

gttccctcga acccccccgaa tcaagtggac atacgtgctg gagaaggtcg ccgcagtctt     2040

tggcgtcttc ttccttatga tctgggtcgc agagtcgtac ctgtatcccc ctgtggtggc     2100

tgttattcaa atgcgagacg aacccttctg gaacaaggtc cgaatctatc ccattttcct     2160

gtcggacatt ctgctgccct ttgtcattga gtacatgctt gttttctaca tcatctggga     2220

cgccattctc aacggcattg ccgagctcac tcgcttcgcc gacagagact tttatggccc     2280

ctggtggaac tgtaccagct gggagcagtt tagccgagaa tggaacattc ctgtctacca     2340

gttcctcaag cgacacgtct accactcgtc catctctgct ttcaagttct ccaagggcgc     2400

agctaccctc accaccttct tgctgtcttc tcttgtccac gagctggtca tgtttgccat     2460

ctttaagaag ttccgaggat acctgctgtt gctgcagatg acccagctgc ccctggccat     2520

gctgcagaaa accaaatgga tccaggacag acccgttttt ggcaacgctt tcttctggtt     2580

ctcgctcatg atcggacctt ctctcatgtg ttccatgtac ctcctcttct aagcacagtc     2640

gcttcaccac ttgtgcctat agaatgtcaa ccactgtcaa cttcttgtat ggccatcaac     2700

tatgaacaag tcagttgaca gctacaatgc ttccactgaa gttatcctcc cccactgtat     2760
```

tttattctta acactgatca tctacctaac gagacactgt agagccactg gtattgtatg 2820

aaacttcgcc aagagatgcc gacgatgaaa gtctgagttt aagaaaatca aaaaggggta 2880

aaaaaaacgg tttacgccct ccgagccgat actacggaga tatcggagtt acagacactt 2940

tgggagcaaa gtctgcaata caagcccacg gaactagagg tgggtcgtca gtaaaggccg 3000

acgatactca gctggctgca taacttccgg attgaaagag agccagtctt tcgaacctcg 3060

gctcagaaaa gaagagaagg aaggtacaaa catcatacga tatagagtct ttttagcacg 3120

gcggtcagtt ggattgcaaa gtggaagggc aagagtactg ctgccttgtt taattggtat 3180

ccactatgca ttttgctttt tgcaagtgga actgggagaa cttgtagcct tgggttgatg 3240

cggtcggaga ttagagttga tctgaaaaat aaaaaataaa aaaacgagat gcggattgtg 3300

gaagctgtag aggatatgtg actacaagag tagtatctac gagtgtgtac ttgctgcgaa 3360

agttgatttc gaacgaacca gcttcattga ctaaacagtg atgattttcg agacagcgga 3420

aaacttcaag ccagtatgca gcatttaggc tcatcggcac ctattttagc cgcaactagc 3480

cgcttctttc tgtaacatac gatgatttgt caggtacagt acagtaccgt tttaagcacc 3540

tgaactagcg gtggcctacg tgcacatgct cgtattgtag taccgagtta aagattgttt 3600

tggagtctag agttgtattt gtcattgttt ta 3632

<210> 59
<211> 543
<212> PRT
<213> Yarrowia lipolytica

<400> 59

Met Ala Thr Leu His Pro Glu Asp Ala Ala Gly Arg Pro Val Arg Arg
1 5 10 15

Arg Pro Arg Pro Ser Ser Ser Gly Gly Ser Arg Ser Pro Ser Thr Lys
20 25 30

Arg His Ser Ile Val Arg Glu His Leu Gly Glu Glu Leu Asn Val Pro
35 40 45

```
Asp Gly Gln Glu Met Asp Leu Gly Gln Val Asn Lys Asn Leu Asn Ala
    50                  55                  60

Ala Tyr Ala Lys Ala Glu Lys Asp Ser Asp Asp Glu Lys Glu Lys Lys
65                  70                  75                  80

Glu Glu Gly Val Val Asp Glu Leu Pro Glu Lys Tyr Ser Tyr Pro Arg
                85                  90                  95

Phe Ser Lys Asn Asn Arg Arg Tyr Arg Phe Thr Asp Ile Lys Phe Lys
            100                 105                 110

Pro Thr Pro Ser Ile Leu Asp Lys Phe Ala His Lys Asp Ser Glu Phe
        115                 120                 125

Phe Gly Phe Tyr Thr Leu Leu Trp Met Val Phe Ala Phe Cys Val Phe
    130                 135                 140

Arg Thr Gly Leu Leu Asn Tyr Thr Asn Glu Gly Ile Leu Phe Arg Gly
145                 150                 155                 160

Gln Ile Phe Ala Ile Leu Ser Lys Asp Leu Trp Lys Val Ala Leu Val
                165                 170                 175

Asp Leu Gly Met Tyr Leu Thr Thr Tyr Leu Ser Val Phe Leu Gln Leu
            180                 185                 190

Ala Val Lys His Gly Leu Val Asp Trp Asn Ser Phe Gly Trp Ile Ile
        195                 200                 205

Gln Asn Val His Gln Thr Leu Phe Leu Phe Phe Tyr Leu Trp Val Ala
    210                 215                 220

Lys Ser Ser Asn Leu Pro Trp Ile Gly Asn Ile Phe Ile Val Leu His
225                 230                 235                 240

Ala Phe Val Met Leu Met Lys Gln His Ser Tyr Ala Phe Tyr Asn Gly
                245                 250                 255
```

```
Tyr Leu Trp Thr Val Glu Asp Glu Leu Ser His Ala Lys Gln Arg Leu
            260                 265                 270

Thr Glu Asp Ile Pro Val Ser Glu Lys Glu Asp Leu Lys Leu Asp Ile
        275                 280                 285

Glu Phe Cys Glu Thr Glu Leu Lys Val Gln Ser Arg His Thr Pro Phe
    290                 295                 300

Pro Thr Asn Ile Thr Phe Ser Asn Tyr Phe Trp Tyr Ser Met Phe Pro
305                 310                 315                 320

Thr Leu Val Tyr Glu Ile Glu Phe Pro Arg Thr Pro Arg Ile Lys Trp
                325                 330                 335

Thr Tyr Val Leu Glu Lys Val Ala Ala Val Phe Gly Val Phe Phe Leu
            340                 345                 350

Met Ile Trp Val Ala Glu Ser Tyr Leu Tyr Pro Pro Val Val Ala Val
        355                 360                 365

Ile Gln Met Arg Asp Glu Pro Phe Trp Asn Lys Val Arg Ile Tyr Pro
    370                 375                 380

Ile Phe Leu Ser Asp Ile Leu Leu Pro Phe Val Ile Glu Tyr Met Leu
385                 390                 395                 400

Val Phe Tyr Ile Ile Trp Asp Ala Ile Leu Asn Gly Ile Ala Glu Leu
                405                 410                 415

Thr Arg Phe Ala Asp Arg Asp Phe Tyr Gly Pro Trp Trp Asn Cys Thr
            420                 425                 430

Ser Trp Glu Gln Phe Ser Arg Glu Trp Asn Ile Pro Val Tyr Gln Phe
        435                 440                 445
```

```
Leu Lys Arg His Val Tyr His Ser Ser Ile Ser Ala Phe Lys Phe Ser
    450                 455                 460

Lys Gly Ala Ala Thr Leu Thr Thr Phe Leu Leu Ser Ser Leu Val His
465                 470                 475                 480

Glu Leu Val Met Phe Ala Ile Phe Lys Lys Phe Arg Gly Tyr Leu Leu
                485                 490                 495

Leu Leu Gln Met Thr Gln Leu Pro Leu Ala Met Leu Gln Lys Thr Lys
            500                 505                 510

Trp Ile Gln Asp Arg Pro Val Phe Gly Asn Ala Phe Phe Trp Phe Ser
        515                 520                 525

Leu Met Ile Gly Pro Ser Leu Met Cys Ser Met Tyr Leu Leu Phe
    530                 535                 540
```

<210> 60
<211> 4034
<212> DNA
<213> Yarrowia lipolytica

<400> 60

```
tgagtataat ccgtctccgt tcccacttga ggggaggact ttggacgggt ttcggtgcaa       60

cctcgtcagg tgaataatca tgattcattt gcgtgatctg gctgggacaa aatacaagtg      120

cttgtatatg gtataccagg aggaccactt atgagaatgc caaagaaaat tgaaggggaa      180

aataaatgct gccacccgat actcgaaata taccattgga cgtgttttcc acataggtaa      240

tcgccactcc atacccgaga tctgaagtga aacattcgtg ttctgtagcc gttagcccac      300

ccccacacac tgagacactc ggctgatgtt tcgaaattca tgctactgta cggtacggac      360

ttaacactta tccccattgt cctgggacta ctacattgta gtatactcta cacatccaac      420

aagcaatact ttttcgtaca agtaggtagt tgtgcgagta ctcgtacaat ggaaaaacat      480

tatacatgaa gcagggcgtg cgtcggtctg tagggtgaaa tagggagccg acagtgcatt      540

tccagcattc tagattgatc tgaatcaata atttgtgggg gaaattggca ctagtgcttt      600
```

gtttattagc tatttgacag tctgatgatg agaaagtcgt tcttcggctt agcccaaccc 660

cacactcact acgcacctgg cacagggtgg tgcatataat atgcaagtac ggagatgtac 720

taacgctagc caggagagag tgacgaaacg tagccatccg ttgagcgggc agacgtggga 780

gtaatatagt cgtttagggc acgtgattgc gaccccagtg gtatcatctt tgtctcccaa 840

actcttcgta atgcccaccc aggacgatcc tctctgaact ttactacccc acattcccga 900

gactataaag tatagacggc tcctacaaaa gagagcaggt ggaataacac actattgacc 960

gacaatgaca accaacacat tcaccgatcc tcctgtggag atggccaagg agcgaggtaa 1020

gactcaattc actgtccgag atgtgaccaa cttcctcaat ggtggagaag aagagaccca 1080

gattgtcgag aagatcatga gcagtattga acgtgatcca gtactgtctg tcactgctga 1140

ctacgactgc aaccttcagc aggcccgaaa acagaccatg gagcgggtgg ctgctctgtc 1200

gccttatctg gtcaccgata ctgagaagct atctctgtgg cgtgcgcaac tgcatggaat 1260

ggttgatatg tctactcgta cgcggttgtc gatccacaac aacctgttca ttggttccat 1320

caggggatct ggtactcctg aacagttcaa gtactgggtc aagaagggag cggtggctgt 1380

taagcagttc tatggatgct ttgccatgac agagttgggc catggaagca acctcaaggg 1440

actagagaca accgccactt atgaccagga cagtgaccag ttcattatca acactcctca 1500

tattggtgct accaagtggt ggattggcgg tgcagcccac acttccaccc attgtgtttg 1560

tttcgcgaaa ctgattgtgc atggcaagga ctatggtact cgaaactttg tggtacctct 1620

ccgaaatgtc cacgatcaca gtctcaaggt cggtgtttca attggagaca ttggaaagaa 1680

gatgggcaga gatggtgttg acaatggctg gatccagttc accaatgttc gaatccccag 1740

acagaacatg ctaatgagat atgccaaggt gtctgatact ggagtggtaa ccaaacccgc 1800

tcttgaccaa ctcacttatg gagccctcat tcgaggtcga gtgtccatga ttgccgactc 1860

gttccacgtc tccaaacgat tcctcacaat tgctcttcgg tacgcttgtg tccgacgaca 1920

gtttggaacc tctggagaca ctaaggagac caagatcatc gactaccctt accaccagcg 1980

acgattgctg cctcttctgg cctactgcta cgctatgaag atgggtgctg atgaggctca 2040

gaagacttgg attgagacca ccgatcgaat tctggctctc aatcccaacg accccgccca 2100

gaagaacgat ctggagaagg ccgtcaccga cacaaaggag ctgtttgctg cgtctgcagg 2160

aatgaaggca tttaccacgt ggggatgtgc caaaatcatt gatgagtgcc gacaggcctg 2220

tggaggtcat ggatactctg gatataacgg atttggccag ggctacgctg actgggttgt 2280

ccagtgtacc tgggaaggag acaacaacgt tctgtgtctg tcaatgggcc gagggctggt 2340

tcagtcagct ctacagattt tggctggaaa gcacgtcggt gcttctattc agtacgtagg 2400

agacaagtct aaaatctccc agaacggcca gggtaccccc agagagcaac ttctgtcccc 2460

cgagtttcta gtagaagctt tcagaacggc ttctcgaaac aacattctca gaaccaccga 2520

taaataccaa gagcttgtca aaactctcaa tcccgaccag gcctttgagg agctgtctca 2580

gcagagattc cagtgtgctc gaatccacac acgacagcat cttatctctt cattctatgc 2640

ccgaattgcc actgccaaag acgatatcaa gccccatctg ctgaaactgg ccaatctgtt 2700

tgccctctgg tcaattgagg aggacactgg aatcttcctg cgggagaaca tcctcacccc 2760

tggagacatt gacctgatca acagtcttgt ggacgagctc tgtgttgcag ttcgagatca 2820

ggtaattgga ctcactgatg cctttggtct ctctgacttc ttcattaacg ctcccatcgg 2880

ctcctacgat ggtaatgttt acgaaaagta ctttgccaag gtcaaccagc aaaaccccgc 2940

tactaaccct cgtcctccct actacgagtc gactctcaag cccttcttgt tccgagaaga 3000

ggaggacgat gaaatttgcg atctcgatga gtgatgttag taattatgag atacattctt 3060

tgcaatgatg agggctcttg tagagtattc tgctgtaaag gagatcttga attttgttgt 3120

tgaatttgtg taattgtctt atgattattg tcttgtgtag cgcttttcaa gcctgcagga 3180

aaagtgaaat gaacatgtta ttgccatatc atccgttatt ggtgtttagt tcctattggg 3240

cttgattacc aacgatccgg gtcgaaagac ccctccgttt aggggtggta agtctcaagt 3300

ggctgaacac gtgcatgcac tgtacatcgc cgagaccact caacaccatc aagggctact 3360

gtattcgtag agcatacggc gttcaaggct tcaccgataa ttagagctca gcgcaatcaa 3420

ttatttgaag gcttaatggc actgtcagct gcttgatctt gcacgcaatt actacaatat 3480

atagtccaac ataataacat aattagaatt atgatcatcc gtatcaatca tacacaaaat 3540

ccttacacct aatctcgtag ctcctccaca aatcccacaa cgcgtcctgc ctccacaatt 3600

ctaagctcct catctgcaga ctcgtgcaga ccaagccagg tagcgtcctg aacaggcagt 3660

tcatttgtcc aggtctggaa atcagccaca cgcgagccag tagtcagctt gcatggaccg 3720

agagcacgcg actcgttgtc aaaataatcc acctgcatca cctgttggac catctcttgt 3780

aatcgagcca ggtcagcagc gatgtcaatc ttggcgctgt aaatgtcaga cagcaccgag 3840

actacagcat gccaaggaat ggaatgggga gcaatattgt ctctctgacc actcacatga 3900

cgaacccatc cttcaatagc tcgcaaactg gcagccaggt ctccctcatg gaagccataa 3960

tccttggcaa agcccagagg cacaaaacga gccaggtcaa caatgcagac atgcagccaa 4020

ctggccagga agtt 4034

<210> 61
<211> 677
<212> PRT
<213> Yarrowia lipolytica

<400> 61

Met Ala Lys Glu Arg Gly Lys Thr Gln Phe Thr Val Arg Asp Val Thr
1 5 10 15

Asn Phe Leu Asn Gly Gly Glu Glu Glu Thr Gln Ile Val Glu Lys Ile
20 25 30

Met Ser Ser Ile Glu Arg Asp Pro Val Leu Ser Val Thr Ala Asp Tyr
35 40 45

Asp Cys Asn Leu Gln Gln Ala Arg Lys Gln Thr Met Glu Arg Val Ala
50 55 60

Ala Leu Ser Pro Tyr Leu Val Thr Asp Thr Glu Lys Leu Ser Leu Trp
65 70 75 80

Arg Ala Gln Leu His Gly Met Val Asp Met Ser Thr Arg Thr Arg Leu
85 90 95

Ser Ile His Asn Asn Leu Phe Ile Gly Ser Ile Arg Gly Ser Gly Thr
100 105 110

```
Pro Glu Gln Phe Lys Tyr Trp Val Lys Lys Gly Ala Val Ala Val Lys
        115                 120                 125

Gln Phe Tyr Gly Cys Phe Ala Met Thr Glu Leu Gly His Gly Ser Asn
    130                 135                 140

Leu Lys Gly Leu Glu Thr Thr Ala Thr Tyr Asp Gln Asp Ser Asp Gln
145                 150                 155                 160

Phe Ile Ile Asn Thr Pro His Ile Gly Ala Thr Lys Trp Trp Ile Gly
                165                 170                 175

Gly Ala Ala His Thr Ser Thr His Cys Val Cys Phe Ala Lys Leu Ile
            180                 185                 190

Val His Gly Lys Asp Tyr Gly Thr Arg Asn Phe Val Val Pro Leu Arg
        195                 200                 205

Asn Val His Asp His Ser Leu Lys Val Gly Val Ser Ile Gly Asp Ile
    210                 215                 220

Gly Lys Lys Met Gly Arg Asp Gly Val Asp Asn Gly Trp Ile Gln Phe
225                 230                 235                 240

Thr Asn Val Arg Ile Pro Arg Gln Asn Met Leu Met Arg Tyr Ala Lys
                245                 250                 255

Val Ser Asp Thr Gly Val Val Thr Lys Pro Ala Leu Asp Gln Leu Thr
            260                 265                 270

Tyr Gly Ala Leu Ile Arg Gly Arg Val Ser Met Ile Ala Asp Ser Phe
        275                 280                 285

His Val Ser Lys Arg Phe Leu Thr Ile Ala Leu Arg Tyr Ala Cys Val
    290                 295                 300
```

```
Arg Arg Gln Phe Gly Thr Ser Gly Asp Thr Lys Glu Thr Lys Ile Ile
305                 310                 315                 320

Asp Tyr Pro Tyr His Gln Arg Arg Leu Leu Pro Leu Leu Ala Tyr Cys
                325                 330                 335

Tyr Ala Met Lys Met Gly Ala Asp Glu Ala Gln Lys Thr Trp Ile Glu
            340                 345                 350

Thr Thr Asp Arg Ile Leu Ala Leu Asn Pro Asn Asp Pro Ala Gln Lys
        355                 360                 365

Asn Asp Leu Glu Lys Ala Val Thr Asp Thr Lys Glu Leu Phe Ala Ala
    370                 375                 380

Ser Ala Gly Met Lys Ala Phe Thr Thr Trp Gly Cys Ala Lys Ile Ile
385                 390                 395                 400

Asp Glu Cys Arg Gln Ala Cys Gly Gly His Gly Tyr Ser Gly Tyr Asn
                405                 410                 415

Gly Phe Gly Gln Gly Tyr Ala Asp Trp Val Val Gln Cys Thr Trp Glu
            420                 425                 430

Gly Asp Asn Asn Val Leu Cys Leu Ser Met Gly Arg Gly Leu Val Gln
        435                 440                 445

Ser Ala Leu Gln Ile Leu Ala Gly Lys His Val Gly Ala Ser Ile Gln
    450                 455                 460

Tyr Val Gly Asp Lys Ser Lys Ile Ser Gln Asn Gly Gln Gly Thr Pro
465                 470                 475                 480

Arg Glu Gln Leu Leu Ser Pro Glu Phe Leu Val Glu Ala Phe Arg Thr
                485                 490                 495

Ala Ser Arg Asn Asn Ile Leu Arg Thr Thr Asp Lys Tyr Gln Glu Leu
            500                 505                 510
```

```
Val Lys Thr Leu Asn Pro Asp Gln Ala Phe Glu Glu Leu Ser Gln Gln
        515                 520                 525

Arg Phe Gln Cys Ala Arg Ile His Thr Arg Gln His Leu Ile Ser Ser
    530                 535                 540

Phe Tyr Ala Arg Ile Ala Thr Ala Lys Asp Asp Ile Lys Pro His Leu
545                 550                 555                 560

Leu Lys Leu Ala Asn Leu Phe Ala Leu Trp Ser Ile Glu Glu Asp Thr
                565                 570                 575

Gly Ile Phe Leu Arg Glu Asn Ile Leu Thr Pro Gly Asp Ile Asp Leu
            580                 585                 590

Ile Asn Ser Leu Val Asp Glu Leu Cys Val Ala Val Arg Asp Gln Val
        595                 600                 605

Ile Gly Leu Thr Asp Ala Phe Gly Leu Ser Asp Phe Phe Ile Asn Ala
    610                 615                 620

Pro Ile Gly Ser Tyr Asp Gly Asn Val Tyr Glu Lys Tyr Phe Ala Lys
625                 630                 635                 640

Val Asn Gln Gln Asn Pro Ala Thr Asn Pro Arg Pro Pro Tyr Tyr Glu
                645                 650                 655

Ser Thr Leu Lys Pro Phe Leu Phe Arg Glu Glu Glu Asp Asp Glu Ile
            660                 665                 670

Cys Asp Leu Asp Glu
        675
```

<210> 62
<211> 4103
<212> DNA
<213> Yarrowia lipolytica

<400> 62
gattccgcca agtgagactg gcgatcggga gaagggttgg tggtcatggg ggatagaatt 60

tgtacaagtg gaaaaaccac tacgagtagc ggatttgata ccacaagtag cagagatata 120

cagcaatggt gggagtgcaa gtatcggaat gtactgtacc tcctgtactc gtactcgtac 180

ggcactcgta gaaacggggc aatacggggg agaagcgatc gcccgtctgt tcaatcgcca 240

caagtccgag taatgctcga gtatcgaagt cttgtacctc cctgtcaatc atggcaccac 300

tggtcttgac ttgtctattc atactggaca agcgccagag ttaagcttgt agcgaatttc 360

gccctcggac atcaccccat acgacggaca cacatgcccg acaaacagcc tctcttattg 420

tagctgaaag tatattgaat gtgaacgtgt acaatatcag gtaccagcgg gaggttacgg 480

ccaaggtgat accggaataa ccctggcttg gagatggtcg gtccattgta ctgaagtgtc 540

cgtgtcgttt ccgtcactgc cccaattgga catgtttgtt tttccgatct ttcgggcgcc 600

ctctccttgt ctccttgtct gtctcctgga ctgttgctac cccatttctt tggcctccat 660

tggttcctcc ccgtctttca cgtcgtctat ggttgcatgg tttcccttat acttttcccc 720

acagtcacat gttatggagg ggtctagatg gaggcctaat tttgacgtgc aaggggcgaa 780

ttggggcgag aaacacgtcg tggacatggt gcaaggcccg cagggttgat tcgacgcttt 840

tccgcgaaaa aaacaagtcc aaatacccccc gtttattctc cctcggctct cggtatttca 900

catgaaaact ataacctaga ctacacgggc aaccttaacc ccagagtata cttatatacc 960

aaagggatgg gtcctcaaaa atcacacaag caacgacgcc atgaacccca acaacactgg 1020

caccattgaa atcaacggta aggagtacaa caccttcacc gagccccccg tggccatggc 1080

tcaggagcga gccaagacct ccttccccgt gcgagagatg acctacttcc tcgacggtgg 1140

cgagaagaac accctcaaaa acgagcagat catggaggag attgagcgag accctctttt 1200

caacaacgac aactactacg atctcaacaa ggagcagatc cgagagctca ccatggagcg 1260

agtcgccaag ctgtctctgt ttgtgcgtga tcagcccgag gacgacatca agaagcgatt 1320

tgctctcatt ggtatcgccg atatgggaac ctacacccga cttggtgtcc actacggcct 1380

cttctttggc gccgtccgag gtaccggaac tgccgagcag tttggccact ggatctccaa 1440

gggagccgga gacctgcgaa agttctacgg atgtttctcc atgaccgagc tgggccatgg 1500

ctccaacctg gctggtctcg agaccaccgc catctacgat gaggagaccg acgagttcat 1560

catcaacacc cctcacattg ccgccaccaa gtggtggatt ggaggagccg cccacaccgc 1620

cacccacact gtcgtgttcg cccgactcat tgtcaagggc aaggactacg gtgtcaagac 1680

ctttgttgtc cagctgcgaa acatcaacga ccacagcctc aaggtcggta tctctattgg 1740

tgatatcgga aagaagatgg gccgagacgg tatcgataac ggatggatcc agttcaccaa 1800

cgtgcgaatc ccccgacaga acctgctcat gaagtacaca aaggtcgacc gagagggtaa 1860

cgtgacccag cctcctctgg ctcagcttac ctacggttct cttatcactg gtcgagtctc 1920

catggcctct gattctcacc aggtcggaaa gcgattcatc accattgctc tgcgatacgc 1980

ctgcattcga cgacagttct ccaccacccc cggccagccc gagaccaaga tcatcgacta 2040

cccctaccat cagcgacgac ttctgcctct tctggcctat gtctatgctc ttaagatgac 2100

tgccgatgag gttggagctc tcttctcccg aaccatgctt aagatggacg acctcaagcc 2160

cgacgacaag gccggcctca atgaggttgt ttccgacgtc aaggagctct tctccgtctc 2220

cgccggtctc aaggccttct ccacctgggc ttgtgccgac gtcattgaca agacccgaca 2280

ggcttgcggt ggccacggtt actctggata caacggtttc ggccaggcct acgccgactg 2340

ggttgtccag tgcacctggg agggtgacaa caacattctc acccttttctg ccggccgagc 2400

tcttatccag tctgccgttg ctctgcgaaa gggcgagcct gttggtaacg ccgtttctta 2460

cctgaagcga tacaaggatc tggccaacgc taagctcaat ggccgatctc tcaccgaccc 2520

caaggtcctc gtcgaggcct gggaggttgc tgccggtaac atcatcaacc gagccaccga 2580

ccagtacgag aagctcattg gcgagggtct taacgccgac caggcctttg aggttctgtc 2640

tcagcagcga ttccaggccg ccaaggtcca cacacgacga cacctcattg ccgctttctt 2700

ctcccgaatt gacaccgagg ctggcgaggc catcaagcag cccctgctta acctggctct 2760

gctgtttgcc ctgtggtcca tcgaagagga ctctggtctg ttcctgcgag agggcttcct 2820

cgagcccaag gatatcgaca ccgtcaccga gctcgtcaac aagtactgca ccactgtgcg 2880

agaggaggtc attggctaca ccgatgcctt caacctgtcc gactacttca tcaacgctcc 2940

tattggatgc tacgatggtg acgcttaccg acactacttc cagaaggtca acgagcagaa 3000

ccctgcccga gacccccgac ctccttacta cgcctctact ctcaagccct tccttttccg 3060

agaggaggag gatgatgaca tttgcgagct tgatgaggaa tagacaagcg ggtatttatt 3120

gtatgaataa agattatgta ttgattgcaa aaaagtgcat ttgtagatgt ggtttattgt 3180

agagagtacg gtatgtactg tacgaacatt aggagctact tctacaagta gattttctta 3240

acaagggtga aatttactag gaagtacatg catatttcgt tagtagaatc acaaaagaaa 3300

tgtacaagca cgtactactt gtactccaca atgtggagtg ggagcaaaaa aattggacga 3360

caccggaatc gaaccgggga cctcgcgcat gctaagcgca tgtgataacc aactacacca 3420

gacgcccaag aactttcttg gtgattatgg aatacgtggt ctgctatatc tcaattttgc 3480

tgtaatgaat cattagaatt aaaaaaaaaa ccccattttt gtgtgattgt cggccaagag 3540

atggaacagg aagaatacgt gaacaagcga gcacgaatgc catatgctct tctgaacaac 3600

cgagtccgaa tccgatttgt gggtatcaca tgtctcaagt agctgaaatg tatttcgcta 3660

gaataaaata aatgagatta agaattaaaa atattggaat atattttcct agaatagaaa 3720

ctttggattt tttttcggct attacagtct gaactggaca aacggctgac tatatataaa 3780

tattattggg tctgttttct tgtttatgtc gaaattatct gggttttact actgtgtcgt 3840

cgagtataga gtggcctgac tggagaaaat gcagtagtat ggacagtagg tactgccagc 3900

cagagaagtt tttggaattg atacttgagt cattttcca ttccccattc cccattccaa 3960

cacaatcaac tgtttctgaa cattttccaa aacgcggaga tgtatgtcac ttggcactgc 4020

aagtctcgat tcaaaatgca tctctttcag accaaagtgt catcagcttt gtttggcccc 4080

aaattaccgc aaatacttgt cga 4103

<210> 63
<211> 700
<212> PRT
<213> Yarrowia lipolytica

<400> 63

Met Asn Pro Asn Asn Thr Gly Thr Ile Glu Ile Asn Gly Lys Glu Tyr
1 5 10 15

```
Asn Thr Phe Thr Glu Pro Pro Val Ala Met Ala Gln Glu Arg Ala Lys
            20                  25                  30

Thr Ser Phe Pro Val Arg Glu Met Thr Tyr Phe Leu Asp Gly Gly Glu
        35                  40                  45

Lys Asn Thr Leu Lys Asn Glu Gln Ile Met Glu Glu Ile Glu Arg Asp
    50                  55                  60

Pro Leu Phe Asn Asn Asp Asn Tyr Tyr Asp Leu Asn Lys Glu Gln Ile
65                  70                  75                  80

Arg Glu Leu Thr Met Glu Arg Val Ala Lys Leu Ser Leu Phe Val Arg
                85                  90                  95

Asp Gln Pro Glu Asp Asp Ile Lys Lys Arg Phe Ala Leu Ile Gly Ile
            100                 105                 110

Ala Asp Met Gly Thr Tyr Thr Arg Leu Gly Val His Tyr Gly Leu Phe
        115                 120                 125

Phe Gly Ala Val Arg Gly Thr Gly Thr Ala Glu Gln Phe Gly His Trp
    130                 135                 140

Ile Ser Lys Gly Ala Gly Asp Leu Arg Lys Phe Tyr Gly Cys Phe Ser
145                 150                 155                 160

Met Thr Glu Leu Gly His Gly Ser Asn Leu Ala Gly Leu Glu Thr Thr
                165                 170                 175

Ala Ile Tyr Asp Glu Glu Thr Asp Glu Phe Ile Ile Asn Thr Pro His
            180                 185                 190

Ile Ala Ala Thr Lys Trp Trp Ile Gly Gly Ala Ala His Thr Ala Thr
        195                 200                 205
```

```
His Thr Val Val Phe Ala Arg Leu Ile Val Lys Gly Lys Asp Tyr Gly
    210                 215                 220

Val Lys Thr Phe Val Val Gln Leu Arg Asn Ile Asn Asp His Ser Leu
225                 230                 235                 240

Lys Val Gly Ile Ser Ile Gly Asp Ile Gly Lys Lys Met Gly Arg Asp
                245                 250                 255

Gly Ile Asp Asn Gly Trp Ile Gln Phe Thr Asn Val Arg Ile Pro Arg
            260                 265                 270

Gln Asn Leu Leu Met Lys Tyr Thr Lys Val Asp Arg Glu Gly Asn Val
        275                 280                 285

Thr Gln Pro Pro Leu Ala Gln Leu Thr Tyr Gly Ser Leu Ile Thr Gly
    290                 295                 300

Arg Val Ser Met Ala Ser Asp Ser His Gln Val Gly Lys Arg Phe Ile
305                 310                 315                 320

Thr Ile Ala Leu Arg Tyr Ala Cys Ile Arg Arg Gln Phe Ser Thr Thr
                325                 330                 335

Pro Gly Gln Pro Glu Thr Lys Ile Ile Asp Tyr Pro Tyr His Gln Arg
            340                 345                 350

Arg Leu Leu Pro Leu Leu Ala Tyr Val Tyr Ala Leu Lys Met Thr Ala
        355                 360                 365

Asp Glu Val Gly Ala Leu Phe Ser Arg Thr Met Leu Lys Met Asp Asp
    370                 375                 380

Leu Lys Pro Asp Asp Lys Ala Gly Leu Asn Glu Val Val Ser Asp Val
385                 390                 395                 400

Lys Glu Leu Phe Ser Val Ser Ala Gly Leu Lys Ala Phe Ser Thr Trp
                405                 410                 415
```

```
Ala Cys Ala Asp Val Ile Asp Lys Thr Arg Gln Ala Cys Gly Gly His
            420                 425                 430

Gly Tyr Ser Gly Tyr Asn Gly Phe Gly Gln Ala Tyr Ala Asp Trp Val
        435                 440                 445

Val Gln Cys Thr Trp Glu Gly Asp Asn Asn Ile Leu Thr Leu Ser Ala
    450                 455                 460

Gly Arg Ala Leu Ile Gln Ser Ala Val Ala Leu Arg Lys Gly Glu Pro
465                 470                 475                 480

Val Gly Asn Ala Val Ser Tyr Leu Lys Arg Tyr Lys Asp Leu Ala Asn
                485                 490                 495

Ala Lys Leu Asn Gly Arg Ser Leu Thr Asp Pro Lys Val Leu Val Glu
            500                 505                 510

Ala Trp Glu Val Ala Ala Gly Asn Ile Ile Asn Arg Ala Thr Asp Gln
        515                 520                 525

Tyr Glu Lys Leu Ile Gly Glu Gly Leu Asn Ala Asp Gln Ala Phe Glu
    530                 535                 540

Val Leu Ser Gln Gln Arg Phe Gln Ala Ala Lys Val His Thr Arg Arg
545                 550                 555                 560

His Leu Ile Ala Ala Phe Phe Ser Arg Ile Asp Thr Glu Ala Gly Glu
                565                 570                 575

Ala Ile Lys Gln Pro Leu Leu Asn Leu Ala Leu Leu Phe Ala Leu Trp
            580                 585                 590

Ser Ile Glu Glu Asp Ser Gly Leu Phe Leu Arg Glu Gly Phe Leu Glu
        595                 600                 605
```

Pro Lys Asp Ile Asp Thr Val Thr Glu Leu Val Asn Lys Tyr Cys Thr
610 615 620

Thr Val Arg Glu Glu Val Ile Gly Tyr Thr Asp Ala Phe Asn Leu Ser
625 630 635 640

Asp Tyr Phe Ile Asn Ala Pro Ile Gly Cys Tyr Asp Gly Asp Ala Tyr
645 650 655

Arg His Tyr Phe Gln Lys Val Asn Glu Gln Asn Pro Ala Arg Asp Pro
660 665 670

Arg Pro Pro Tyr Tyr Ala Ser Thr Leu Lys Pro Phe Leu Phe Arg Glu
675 680 685

Glu Glu Asp Asp Asp Ile Cys Glu Leu Asp Glu Glu
690 695 700

<210> 64
<211> 4103
<212> DNA
<213> Yarrowia lipolytica

<400> 64
tagtgttttt gttggttttt atttgatttg ttggtttaat atttatttta ttttaatatt 60

tattttattt tatttcaatt aattatttta tttatttgct taaatcaaat attttggacc 120

aataaaaacc gacacagttt tcagacatgt ttgtagctca tatgtatcag gtgtggttgt 180

ggggtatcaa agccgaggtt ggcaaatttt actggggcct tcagaacggt aaaaatagac 240

caatcagaat tctgaaaagc acatcttgat ctcctcattg cggggagtcc aacggtggtc 300

ttattccccc gaatttcccg ctcaatatcg ttccagaccg acccggacac agtgcttaac 360

gccgttccga aactctaccg cagatatgct ccaacggact gggctgcata gatgtgatcc 420

tcggcttgga gaaatggata aaagccggcc aaaaaaaaag cggaaaaaag cggaaaaaaa 480

gagaaaaaaa atcgcaaaat cggaaaaata gggggaaaag acgcaaaaac gcaaggaggg 540

gggagtatat gacactgata agcaagctca caacggttcc tcttattttt ttcctcatct 600

tctgcctagg ttcccaaaat cccagatgct tctctccagt gccaaaagta agtaccccac 660

aggttttcgg ccgaaaattc cacgtgcagc aacgtcgtgt ggggtgttaa aatgtggggg 720

ggggaccagg acaagaggct cttgtgggag ccgaatgaga gcacaaagcg ggcgggtgtg 780

ataagggcat ttttgcccat tttcccttct cctgtctctc cgacggtgat ggcgttgtgc 840

gtcctctatt tctttttatt tctttttgtt ttatttctct gactaccgat ttggtttgat 900

ttcctcaacc ccacacaaat aagctcgggc cgaggaatat atatatacac ggacacagtc 960

gccctgtgga caacacgtca ctacctctac gatacacaca atgatctccc ccaacctcac 1020

agctaacgtc gagattgacg gcaagcagta caacaccttc acagagccac ccaaggcgct 1080

cgccggcgag cgagccaagg tcaagttccc catcaaggac atgacggagt ttctgcacgg 1140

tggcgaggag aacgtgacca tgatcgagcg actgatgacg gagctcgagc gagaccccgt 1200

gctcaacgtg tcgggcgact acgacatgcc caaggagcag ctgcgagaga cggccgtggc 1260

gcgaattgcg gcgctgtccg gccactggaa gaaggacaca gaaaaggagg cgctgctgcg 1320

gtcccagctg cacggcattg tggacatggg cacccgaatc cgactcggtg tgcacacggg 1380

cctgttcatg ggcgccatcc ggggttccgg caccaaggag cagtacgact actgggtgcg 1440

aaagggcgcc gcggacgtca agggcttcta cggctgcttt gctatgaccg agctgggcca 1500

tggctccaac gtggccggtc ttgagaccac cgccacctac atccaggaca cggacgagtt 1560

catcatcaac acccccaaca ctggagccac caagtggtgg attggaggag ccgcccactc 1620

ggccacccac accgcctgct ttgctcgtct gcttgtcgac ggcaaggact acggcgtcaa 1680

gatctttgtt gtccagctgc gagacgtctc ttctcactct ctcatgcccg gcatcgctct 1740

cggcgacatt ggaaagaaga tgggccgaga cgccatcgac aacggctgga tccagttcac 1800

caatgtgcga atcccccgac agaacatgct catgaagtac gccaaggtct cgtctaccgg 1860

caaggtgtcg cagcctcctc tggcccagct cacctacggc gctctcattg gcggccgagt 1920

caccatgatt gccgactcct tctttgtctc ccagcgattc atcaccattg ctctgcgata 1980

cgcctgtgtg cgacgacagt ttggcaccac ccccggccag cccgagacta agatcatcga 2040

ctacccctac catcagcgac gtctgctgcc tcttctggcc ttcacctacg ccatgaagat 2100

ggccgccgac cagtcccaga ttcagtacga tcagaccacc gatctgctgc agaccatcga 2160

ccctaaggac aagggcgctc tgggcaaggc cattgtcgac ctcaaggagc tgtttgcctc 2220

ttctgctggt ctcaaggcct tcaccacctg gacctgtgcc aacatcattg accagtgccg 2280

acaggcctgc ggtggccacg gctactctgg ctacaacggc tttggccagg cctacgccga 2340

ctgggttgtc cagtgcacct gggagggtga caacaacgtc ctgtgtctgt ccatgggccg 2400

aggtctcatc cagtcgtgtc tgggccaccg aaagggtaag cctctgggct cttctgtcgg 2460

ctacctggct aacaagggtc ttgagcaggc tactctgagc ggccgagacc tcaaggaccc 2520

caaggttctc atcgaggcct gggagaaggt cgccaacggc gccatccagc gggccactga 2580

caaatttgtc gagctcacca agggcggcct ctctcctgac caggcctttg aggagctgtc 2640

gcagcagcga ttccagtgtg ccaagatcca cacccgaaag cacctggtga ctgccttcta 2700

cgagcgaatc aacgcctctg cgaaggccga cgtcaagcct tacctcatca acctcgccaa 2760

cctcttcact ctgtggtcca ttgaggagga ctctggtctc ttcctgcgag agggtttcct 2820

gcagcccaag gacattgacc aggtgactga gctggtgaac cactactgca aggaggttcg 2880

agaccaggtt gccggctaca ccgatgcctt tggtctgtct gactggttca tcaacgctcc 2940

cattggaaac tacgatggtg acgtttacaa gcattacttt gccaaggtta accagcagaa 3000

ccctgctcag aacccccgac ctccttacta tgagagcact cttcgacctt tcctgttccg 3060

agaggatgag gatgacgaca tttgcgagct ggacgaggaa tagatggagc gtgtgttctg 3120

agtcgatgtt ttctatggag ttgtgagtgt tagtagacat gatgggttta tatatgatga 3180

atgaatagat gtgattttga tttgcacgat ggaattgaga actttgtaaa cgtacatggg 3240

aatgtatgaa tgtgggggtt ttgtgactgg ataactgacg gtcagtggac gccgttgttc 3300

aaatatccaa gagatgcgag aaactttggg tcaagtgaac atgtcctctc tgttcaagta 3360

aaccatcaac tatgggtagt atatttagta aggacaggag ttgagagagg aaagttgcca 3420

ttctttggag tcccagaaac gtattttcgc gttccaagat caaattagta gagtaatacg 3480

ggcacgggaa tccattcata gtctcaattt tcccataggt gtgctacaag gtgttgagat 3540

gtggtacagt accaccatga ttcgagataa agagcccaga agtcattgat gaggtcaaga 3600

aatacacaga tctacagctc aatacaatga atatcttctt tcatattctt caggtgacac 3660

caagggtgtc tattttcccc agaaatgcgt gaaaaggcgc gtgtgtagcg tggagtatgg 3720

gttcggttgg cgtatccttc atatatcgac gaaatagtag ggcaagagat gacaaaaagt 3780

atctatatgt agacagcgta gaatatggat ttgattggta taaattcatt tattgcgtgt 3840

ctcacaaata ctctcgataa gttggggtta aactggagat ggaacaatgt cgatatctcg 3900

acatattttg atatttgtac tgttgatagt gataaaaagt agaccgttcg aatctcgaca 3960

aggagaagag tccaatgaaa taggttccat catcattcgt catagttaaa cgccgctggt 4020

tgccattact atccgtcttg actacaaccc caactcagcc tagaccacag cgaagagaat 4080

cagtttggag accgaaaatg agc 4103

<210> 65
<211> 700
<212> PRT
<213> Yarrowia lipolytica

<400> 65

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

Tyr Asn Thr Phe Thr Glu Pro Pro Lys Ala Leu Ala Gly Glu Arg Ala
20 25 30

Lys Val Lys Phe Pro Ile Lys Asp Met Thr Glu Phe Leu His Gly Gly
35 40 45

Glu Glu Asn Val Thr Met Ile Glu Arg Leu Met Thr Glu Leu Glu Arg
50 55 60

Asp Pro Val Leu Asn Val Ser Gly Asp Tyr Asp Met Pro Lys Glu Gln
65 70 75 80

Leu Arg Glu Thr Ala Val Ala Arg Ile Ala Ala Leu Ser Gly His Trp
85 90 95

```
Lys Lys Asp Thr Glu Lys Glu Ala Leu Leu Arg Ser Gln Leu His Gly
            100                 105                 110

Ile Val Asp Met Gly Thr Arg Ile Arg Leu Gly Val His Thr Gly Leu
        115                 120                 125

Phe Met Gly Ala Ile Arg Gly Ser Gly Thr Lys Glu Gln Tyr Asp Tyr
    130                 135                 140

Trp Val Arg Lys Gly Ala Ala Asp Val Lys Gly Phe Tyr Gly Cys Phe
145                 150                 155                 160

Ala Met Thr Glu Leu Gly His Gly Ser Asn Val Ala Gly Leu Glu Thr
                165                 170                 175

Thr Ala Thr Tyr Ile Gln Asp Thr Asp Glu Phe Ile Ile Asn Thr Pro
            180                 185                 190

Asn Thr Gly Ala Thr Lys Trp Trp Ile Gly Gly Ala Ala His Ser Ala
        195                 200                 205

Thr His Thr Ala Cys Phe Ala Arg Leu Leu Val Asp Gly Lys Asp Tyr
    210                 215                 220

Gly Val Lys Ile Phe Val Val Gln Leu Arg Asp Val Ser Ser His Ser
225                 230                 235                 240

Leu Met Pro Gly Ile Ala Leu Gly Asp Ile Gly Lys Lys Met Gly Arg
                245                 250                 255

Asp Ala Ile Asp Asn Gly Trp Ile Gln Phe Thr Asn Val Arg Ile Pro
            260                 265                 270

Arg Gln Asn Met Leu Met Lys Tyr Ala Lys Val Ser Ser Thr Gly Lys
        275                 280                 285

Val Ser Gln Pro Pro Leu Ala Gln Leu Thr Tyr Gly Ala Leu Ile Gly
    290                 295                 300
```

```
Gly Arg Val Thr Met Ile Ala Asp Ser Phe Phe Val Ser Gln Arg Phe
305                 310                 315                 320

Ile Thr Ile Ala Leu Arg Tyr Ala Cys Val Arg Arg Gln Phe Gly Thr
                325                 330                 335

Thr Pro Gly Gln Pro Glu Thr Lys Ile Ile Asp Tyr Pro Tyr His Gln
            340                 345                 350

Arg Arg Leu Leu Pro Leu Leu Ala Phe Thr Tyr Ala Met Lys Met Ala
        355                 360                 365

Ala Asp Gln Ser Gln Ile Gln Tyr Asp Gln Thr Thr Asp Leu Leu Gln
    370                 375                 380

Thr Ile Asp Pro Lys Asp Lys Gly Ala Leu Gly Lys Ala Ile Val Asp
385                 390                 395                 400

Leu Lys Glu Leu Phe Ala Ser Ser Ala Gly Leu Lys Ala Phe Thr Thr
                405                 410                 415

Trp Thr Cys Ala Asn Ile Ile Asp Gln Cys Arg Gln Ala Cys Gly Gly
            420                 425                 430

His Gly Tyr Ser Gly Tyr Asn Gly Phe Gly Gln Ala Tyr Ala Asp Trp
        435                 440                 445

Val Val Gln Cys Thr Trp Glu Gly Asp Asn Asn Val Leu Cys Leu Ser
    450                 455                 460

Met Gly Arg Gly Leu Ile Gln Ser Cys Leu Gly His Arg Lys Gly Lys
465                 470                 475                 480

Pro Leu Gly Ser Ser Val Gly Tyr Leu Ala Asn Lys Gly Leu Glu Gln
                485                 490                 495
```

```
Ala Thr Leu Ser Gly Arg Asp Leu Lys Asp Pro Lys Val Leu Ile Glu
            500                 505                 510

Ala Trp Glu Lys Val Ala Asn Gly Ala Ile Gln Arg Ala Thr Asp Lys
        515                 520                 525

Phe Val Glu Leu Thr Lys Gly Gly Leu Ser Pro Asp Gln Ala Phe Glu
    530                 535                 540

Glu Leu Ser Gln Gln Arg Phe Gln Cys Ala Lys Ile His Thr Arg Lys
545                 550                 555                 560

His Leu Val Thr Ala Phe Tyr Glu Arg Ile Asn Ala Ser Ala Lys Ala
                565                 570                 575

Asp Val Lys Pro Tyr Leu Ile Asn Leu Ala Asn Leu Phe Thr Leu Trp
            580                 585                 590

Ser Ile Glu Glu Asp Ser Gly Leu Phe Leu Arg Glu Gly Phe Leu Gln
        595                 600                 605

Pro Lys Asp Ile Asp Gln Val Thr Glu Leu Val Asn His Tyr Cys Lys
    610                 615                 620

Glu Val Arg Asp Gln Val Ala Gly Tyr Thr Asp Ala Phe Gly Leu Ser
625                 630                 635                 640

Asp Trp Phe Ile Asn Ala Pro Ile Gly Asn Tyr Asp Gly Asp Val Tyr
                645                 650                 655

Lys His Tyr Phe Ala Lys Val Asn Gln Gln Asn Pro Ala Gln Asn Pro
            660                 665                 670

Arg Pro Pro Tyr Tyr Glu Ser Thr Leu Arg Pro Phe Leu Phe Arg Glu
        675                 680                 685

Asp Glu Asp Asp Asp Ile Cys Glu Leu Asp Glu Glu
    690                 695                 700
```

<210> 66
<211> 4706
<212> DNA
<213> Yarrowia lipolytica

<400> 66
acgacgtaca agccaaagag atgttgactc cagcagcgaa ttagagtaag actacagtac 60

tggtagaagt aggcctattg gactgtaggg agtactgcaa tgaccagtgt tttcggtata 120

aatatgagaa ttaccgagaa aaggtaaaaa atctgaattt cggagttacg atgggtggcg 180

ttccctagtg gaacatagat agacctacaa ttgatgggct actctgtcgc aattgtcatt 240

tctgtcacta ttaccacatg tgtcaagtgg cagcaattgt ctagtaatca tttccgttgt 300

cgcaggccaa cggctgattt gtgtaatgtg gggttgagtt ccagcacggg gcttatttgg 360

attggatgag ctaatgcaag ccagagtgag catgaaaacg ggcagatccc catggggagg 420

ttggccaacc ccgagccaac ctctttgggg agagatgggg aaaccctcca agaagaggag 480

aaggatttcg taggtcatgc cgtttcccag tcagaattat ttaaacgtaa aatcggttat 540

aaccatttcc tggttcatgt tgaaatttcc acaagcttct gtatggatgt ggggttctac 600

aactttgaaa ttgatgacat tttcgagaaa atgagaacaa ccgaatattt gtactccata 660

ttgatacgga aaaattcaca atggaatcac acaagatagc tctgaacaat tgaggaccct 720

tttctcccta aattcattgt ggttgcggtt gtattgccaa catggattac gcccgcccga 780

tgacgagcga cttgcttgag ccgaggcaga gtctaaatcg ggacaacgaa accataaacc 840

ccataggaag ccaaagtgac cactgcatga ccccagtata taacctgccc cacccctcca 900

tacttgctgc tttctagatc aagttatacc ggacctgcta actcaattga ggtagaagat 960

acgttattaa cacaatctat ttattcaaca cacacacaca atgtctggag aactaagata 1020

cgacggaaag gtcgtcattg ttaccggtgc cggtggcggt ctcggtaagg catacgccct 1080

tttctacggc tctcgaggag cctctgttgt tgtcaacgat cttggtggcg acttcaaggg 1140

cgacggtgcc caggctggca gtggcaagcg agttgccgat gttgtcgtcg acgagattgt 1200

ttccaaggga ggcaaggctg ttgctaacta cgactctgtc gagaacggtg acaagattgt 1260

cgagactgcc gtcaaggctt ttggctccgt ccacattgtc atcaacaacg ccggtattct 1320

ccgagatatt tccttcaaga agatgaccga caaggactgg gatcttgtct acaaggtcca 1380

cgttttcggt gcctacaagg ttacccgagc tgcctggcct tacttccgaa agcagaagta 1440

cggtcgagtt atctctacct cttccgctgc tggtctttac ggaaacttcg gccagaccaa 1500

ctactccgct gccaagctcg ccctggttgg tttcggtgag actctcgcca aggagggtgc 1560

caagtacaac attacttcca acgtcatcgc tcctcttgct gcttcccgaa tgaccgagac 1620

agtcatgccc gaggatatcc tcaagctcct caagcctgag tacgttgttc ctctggtcgg 1680

ctacctcacc cacgactctg tcaccgagtc ttatggtatt tacgaggtcg gtgctggtta 1740

catggctaaa atccgatggg agcgaggcaa cggtgctgtt ttcaagggcg acgacacttt 1800

caccccgtct gctattctga agcgatggga tgaggtcacc tcttttgaga gccccaccta 1860

ccctaacggc cctgctgact tcttcaaata cgctgaggag tctgttaagc gacccgagaa 1920

cccccaggga cccaccgtct ccttcaagga ccaggttgtc attgtcactg gagccggtgc 1980

tggcattggc cgagcttact ctcacctcct tgctaagctt ggtgccaagg tcgttgttaa 2040

cgatttcggt aaccctcaga aggttgtcga tgaaattaag gccctcggtg gtatcgccgt 2100

cgctgacaag aacaacgtca tccacggtga gaaggttgtt cagaccgcta tcgacgcctt 2160

cggtgctgtc cacgccgttg tcaacaacgc tggtattctc cgagacaagt ctttcgccaa 2220

catggatgat gagatgtggc agctgatctt tgatgtccac ctcaacggta cttactccgt 2280

taccaaggcc gcgtggcccc acttccttaa gcagaagtac ggccgtgtca tcaacaccac 2340

ctcaacttct ggtatctacg gtaacttcgg ccaggccaac tactctgccg ccaaggctgg 2400

tatcctcggt ttctcccgag ctcttgctcg agagggtgag aagtacaaca ttcttgtcaa 2460

caccattgcc cctaacgctg gtactgccat gactgcttct gtcttcactg aggagatgct 2520

cgagctcttc aagcccgatt tcatcgcacc catcaccgtc ctgcttgctt ccgatcaggc 2580

tcccgtcacc ggtgatctgt ttgagactgg ttctgcttgg atcggacaga ctcgatggca 2640

gcgagctggt ggtaaggcct tcaacaccaa gaagggtgtc acccccgaaa tggttcgaga 2700

cagctgggct aagatcgtcg acttcgatga tggtaactcc acccatccca ccactccctc 2760

cgagtctact actcagattc ttgagaacat cttcaacgtg cctgatgagg aggttgagga 2820

gactgctctc gttgctggtc ccggtggtcc cggtatcctc aacaaggagg gcgaaccttt 2880

cgactacact tacacttacc gagacctcat tctttacaac cttggtctcg gtgccaaggc 2940

taatgagctc aagtatgtct tcgagggtga tgatgacttc cagaccgtgc ccactttcgg 3000

tgttatccct tacatgggtg gcctcatcac taccaactat ggcgacttcg ttcctaactt 3060

caaccctatg atgcttctcc acggtgagca gtaccttgaa atccgacagt ggcctattcc 3120

taccaatgct acattggaga acaaggctaa ggtcatcgat gtcgttgaca agggcaaggc 3180

tgccctcctt gtcactgcta ccaccaccac gaacaaggag actggtgagg aggttttcta 3240

caacgagtct tctctcttca tccgaggctc tggtggtttc ggtggtaagt ctaccggtac 3300

tgaccgtggc gctgccactg ctgccaacaa gccccctgct cgagctcctg acttcgttaa 3360

ggagatcaag atccaggagg accaggctgc catttaccga ctttctggtg attacaaccc 3420

tcttcacatc gaccctgctt ttgctgctgt tggtaacttt gaccgaccta ttctccacgg 3480

tctctgctct tttggtgtct ccggtaaggc tctttacgat cagtttggtc ctttcaagaa 3540

cgctaaggtc cgatttgctg gtcacgtctt ccctggtgag accctgaagg ttgagggctg 3600

gaaggagggc aacaaggtca ttttccagac caaggttgtt gagcgaggta ctaccgccat 3660

cagcaatgcc gccattgagc tcttccccaa ggatgctaag ctctaagcta ttatctgacc 3720

aagtgatacg agttaagtaa attttaagaa atgtatgagt taagattagt taatactgtt 3780

atagtttcag ttaggaccac caccataaaa aaaaacgtag agtaatgtat cgcgctgtgt 3840

actgcaagta cagtacatac tgtaccggaa gtgtgagatc agtagagggt ctgatgtgtt 3900

aatagttcaa caaagtgttg ttcaagtact gtacgaatac ttgtactgta ttatgtgatt 3960

ataaaatata tatattggat ctcgcccccc cccccaatat aatgttatta caggtataaa 4020

aatttttatg gatgaaaaaa taaaataaaa caagttgcga tcaagagact tcgtggtggg 4080

actttttttat tactaacaac ctaaattgat cgtgcataat cgtgttgaat tctcacacat 4140

cgcgccaaat cgacagcttt catgaacaga gtgaatcaac aagaagcgga ccagaaccct 4200

gaggctactc tttatatagg caaccttgac gagaacgtca cagaagcgat tctgtatgaa 4260

cttatgcttc aagcgggccc tatcgtgcac atcaatcttc ctaaagacag agtcagccag 4320

acacaccaag gctatggctt cgtggagtac aaaacagagg ctgatgcaaa ttacgcagcg 4380

tctattatga accaaatctg gctgttcgga aaatccattc gagtcagcaa atctgcatcc 4440

gacaagcaaa acgggtttga gattggagca acacttttcg tgggaagttt ggatcctctc 4500

gtagatgagt ctacactaca gcaaacgttc tctgtgtttg gaccatttgc caaacccccct 4560

cgaatatctc gtgacacaga tggttcttca aaagggtacg gttttgtgtc ttttactgat 4620

tttgaacatt cagaccgtgc attggaatca atggacaaac agtatctcat gaaccagcaa 4680

ataacagtca catatgcctt caagca 4706

<210> 67
<211> 901
<212> PRT
<213> Yarrowia lipolytica

<400> 67

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

Ala Gly Gly Gly Leu Gly Lys Ala Tyr Ala Leu Phe Tyr Gly Ser Arg
20 25 30

Gly Ala Ser Val Val Val Asn Asp Leu Gly Gly Asp Phe Lys Gly Asp
35 40 45

Gly Ala Gln Ala Gly Ser Gly Lys Arg Val Ala Asp Val Val Val Asp
50 55 60

Glu Ile Val Ser Lys Gly Gly Lys Ala Val Ala Asn Tyr Asp Ser Val
65 70 75 80

Glu Asn Gly Asp Lys Ile Val Glu Thr Ala Val Lys Ala Phe Gly Ser
85 90 95

Val His Ile Val Ile Asn Asn Ala Gly Ile Leu Arg Asp Ile Ser Phe
100 105 110

```
Lys Lys Met Thr Asp Lys Asp Trp Asp Leu Val Tyr Lys Val His Val
        115                 120                 125

Phe Gly Ala Tyr Lys Val Thr Arg Ala Ala Trp Pro Tyr Phe Arg Lys
    130                 135                 140

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

Gly Asn Phe Gly Gln Thr Asn Tyr Ser Ala Ala Lys Leu Ala Leu Val
                165                 170                 175

Gly Phe Gly Glu Thr Leu Ala Lys Glu Gly Ala Lys Tyr Asn Ile Thr
            180                 185                 190

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

Met Pro Glu Asp Ile Leu Lys Leu Leu Lys Pro Glu Tyr Val Val Pro
    210                 215                 220

Leu Val Gly Tyr Leu Thr His Asp Ser Val Thr Glu Ser Tyr Gly Ile
225                 230                 235                 240

Tyr Glu Val Gly Ala Gly Tyr Met Ala Lys Ile Arg Trp Glu Arg Gly
                245                 250                 255

Asn Gly Ala Val Phe Lys Gly Asp Asp Thr Phe Thr Pro Ser Ala Ile
            260                 265                 270

Leu Lys Arg Trp Asp Glu Val Thr Ser Phe Glu Ser Pro Thr Tyr Pro
        275                 280                 285

Asn Gly Pro Ala Asp Phe Phe Lys Tyr Ala Glu Glu Ser Val Lys Arg
    290                 295                 300
```

```
Pro Glu Asn Pro Gln Gly Pro Thr Val Ser Phe Lys Asp Gln Val Val
305                 310                 315                 320

Ile Val Thr Gly Ala Gly Ala Gly Ile Gly Arg Ala Tyr Ser His Leu
                325                 330                 335

Leu Ala Lys Leu Gly Ala Lys Val Val Val Asn Asp Phe Gly Asn Pro
            340                 345                 350

Gln Lys Val Val Asp Glu Ile Lys Ala Leu Gly Gly Ile Ala Val Ala
        355                 360                 365

Asp Lys Asn Asn Val Ile His Gly Glu Lys Val Val Gln Thr Ala Ile
    370                 375                 380

Asp Ala Phe Gly Ala Val His Ala Val Val Asn Asn Ala Gly Ile Leu
385                 390                 395                 400

Arg Asp Lys Ser Phe Ala Asn Met Asp Asp Glu Met Trp Gln Leu Ile
                405                 410                 415

Phe Asp Val His Leu Asn Gly Thr Tyr Ser Val Thr Lys Ala Ala Trp
            420                 425                 430

Pro His Phe Leu Lys Gln Lys Tyr Gly Arg Val Ile Asn Thr Thr Ser
        435                 440                 445

Thr Ser Gly Ile Tyr Gly Asn Phe Gly Gln Ala Asn Tyr Ser Ala Ala
    450                 455                 460

Lys Ala Gly Ile Leu Gly Phe Ser Arg Ala Leu Ala Arg Glu Gly Glu
465                 470                 475                 480

Lys Tyr Asn Ile Leu Val Asn Thr Ile Ala Pro Asn Ala Gly Thr Ala
                485                 490                 495

Met Thr Ala Ser Val Phe Thr Glu Glu Met Leu Glu Leu Phe Lys Pro
            500                 505                 510
```

```
Asp Phe Ile Ala Pro Ile Thr Val Leu Leu Ala Ser Asp Gln Ala Pro
        515                 520                 525

Val Thr Gly Asp Leu Phe Glu Thr Gly Ser Ala Trp Ile Gly Gln Thr
    530                 535                 540

Arg Trp Gln Arg Ala Gly Gly Lys Ala Phe Asn Thr Lys Lys Gly Val
545                 550                 555                 560

Thr Pro Glu Met Val Arg Asp Ser Trp Ala Lys Ile Val Asp Phe Asp
                565                 570                 575

Asp Gly Asn Ser Thr His Pro Thr Thr Pro Ser Glu Ser Thr Thr Gln
            580                 585                 590

Ile Leu Glu Asn Ile Phe Asn Val Pro Asp Glu Glu Val Glu Glu Thr
        595                 600                 605

Ala Leu Val Ala Gly Pro Gly Gly Pro Gly Ile Leu Asn Lys Glu Gly
    610                 615                 620

Glu Pro Phe Asp Tyr Thr Tyr Thr Tyr Arg Asp Leu Ile Leu Tyr Asn
625                 630                 635                 640

Leu Gly Leu Gly Ala Lys Ala Asn Glu Leu Lys Tyr Val Phe Glu Gly
                645                 650                 655

Asp Asp Asp Phe Gln Thr Val Pro Thr Phe Gly Val Ile Pro Tyr Met
            660                 665                 670

Gly Gly Leu Ile Thr Thr Asn Tyr Gly Asp Phe Val Pro Asn Phe Asn
        675                 680                 685

Pro Met Met Leu Leu His Gly Glu Gln Tyr Leu Glu Ile Arg Gln Trp
    690                 695                 700
```

```
Pro Ile Pro Thr Asn Ala Thr Leu Glu Asn Lys Ala Lys Val Ile Asp
705                 710                 715                 720

Val Val Asp Lys Gly Lys Ala Ala Leu Leu Val Thr Ala Thr Thr Thr
                725                 730                 735

Thr Asn Lys Glu Thr Gly Glu Glu Val Phe Tyr Asn Glu Ser Ser Leu
            740                 745                 750

Phe Ile Arg Gly Ser Gly Gly Phe Gly Gly Lys Ser Thr Gly Thr Asp
        755                 760                 765

Arg Gly Ala Ala Thr Ala Ala Asn Lys Pro Pro Ala Arg Ala Pro Asp
    770                 775                 780

Phe Val Lys Glu Ile Lys Ile Gln Glu Asp Gln Ala Ala Ile Tyr Arg
785                 790                 795                 800

Leu Ser Gly Asp Tyr Asn Pro Leu His Ile Asp Pro Ala Phe Ala Ala
                805                 810                 815

Val Gly Asn Phe Asp Arg Pro Ile Leu His Gly Leu Cys Ser Phe Gly
            820                 825                 830

Val Ser Gly Lys Ala Leu Tyr Asp Gln Phe Gly Pro Phe Lys Asn Ala
        835                 840                 845

Lys Val Arg Phe Ala Gly His Val Phe Pro Gly Glu Thr Leu Lys Val
    850                 855                 860

Glu Gly Trp Lys Glu Gly Asn Lys Val Ile Phe Gln Thr Lys Val Val
865                 870                 875                 880

Glu Arg Gly Thr Thr Ala Ile Ser Asn Ala Ala Ile Glu Leu Phe Pro
                885                 890                 895

Lys Asp Ala Lys Leu
            900
```

<210> 68
<211> 3134
<212> DNA
<213> Yarrowia lipolytica

<400> 68
tgtgttttcg tcttagcgtc atgtattctc aagcttagtc agagagaagg actatggagg 60

agaaggggag aattgagaag ggtatttgaa gggactttga aggtcgcgtg gaagaggtac 120

ttgaagaggt atttgaaggt cacgtggaag aggtatttga agatcacgtg gaagaagtac 180

ttgttttaca gagaatatcg gggtgatttt gacagtggga ttgtctccca agtcctaatc 240

gtttgacatg ggagcagtga aaagtcgggc taaaaaaggg aatatcggaa atcggaaaga 300

cggaaagaat tactggactc atgtttagta gatctgagca cttcaaattt gaaaatatct 360

cttcaaacag cagatcggtt ggtcgtggag gtaccatcaa gggtaaaatc aaggctatca 420

tcaagggcca tatatcgcaa gtttgggggа agataatatg ttcatagtga atcagggttg 480

tggatttcct catctaacgg cattgtaact agtcctggag ggtctttttt atggataacc 540

tccatgtacg atgtatccaa gatctccacg tactgtgttc tgtttcctaa gtaataccca 600

acaacctctc caacaaacac ttgggaagat gcacttgtgc tgagatgtca agatgttagt 660

actgtactgg atggagagaa tattaataaa taattgttac ccaactacat cttgtcgatt 720

gaaagagata cccctaagac agataggata tctgcaaccc gaggaatgaa ccccccagca 780

ccggcaccct ttctattaac aaaatgccaa ctgaaatttg aaaagttcaa ctaaacttat 840

ttgacccaca aaaactcgtc aaaagtggcg gcgaaagctg gcaaatgatg acatcccctt 900

ggaactatga tatcccctcg gaatcttcgt ccccatttgc cacatctact tgcaacgcca 960

cgtctgctta ctaagcaacc caaatctgcc tcggctcaaa atgtggggaa gttcacatgc 1020

attcgctggt gaatctgatc tgacactaca actacacacc aggtccaaca tgagcgacaa 1080

tacgacaatc aaaaagccga tccgacccaa accgatccgg acggaacgcc tgccttacgc 1140

tggggccgca gaaatcatcc gagccaacca gaaagaccac tactttgagt ccgtgcttga 1200

acagcatctc gtcacgtttc tgcagaaatg gaagggagta cgatttatcc accagtacaa 1260

ggaggagctg gagacggcgt ccaagtttgc atatctcggt ttgtgtacgc ttgtgggctc 1320

caagactctc ggagaagagt acaccaatct catgtacact atcagagacc gaacagctct 1380

accgggggtg gtgagacggt ttggctacgt gctttccaac actctgtttc catacctgtt 1440

tgtgcgctac atgggcaagt tgcgcgccaa actgatgcgc gagtatcccc atctggtgga 1500

gtacgacgaa gatgagcctg tgcccagccc ggaaacatgg aaggagcggg tcatcaagac 1560

gtttgtgaac aagtttgaca agttcacggc gctggagggg tttaccgcga tccacttggc 1620

gattttctac gtctacggct cgtactacca gctcagtaag cggatctggg gcatgcgtta 1680

tgtatttgga caccgactgg acaagaatga gcctcgaatc ggttacgaga tgctcggtct 1740

gctgattttc gcccggtttg ccacgtcatt tgtgcagacg ggaagagagt acctcggagc 1800

gctgctggaa aagagcgtgg agaaagaggc aggggagaag gaagatgaaa aggaagcggt 1860

tgtgccgaaa aagaagtcgt caattccgtt cattgaggat acagaagggg agacggaaga 1920

caagatcgat ctggaggacc ctcgacagct caagttcatt cctgaggcgt ccagagcgtg 1980

cactctgtgt ctgtcataca ttagtgcgcc ggcatgtacg ccatgtggac acttttctg 2040

ttgggactgt atttccgaat gggtgagaga gaagcccgag tgtcccttgt gtcggcaggg 2100

tgtgagagag cagaacttgt tgcctatcag ataatgacga ggtctggatg gaaggactag 2160

tcagcgagac acagagcatc agggaccaga cacgaccaat tcaatcgaca acactgtgct 2220

gcatagcagt gcacagaggt cctgggcatg aatatatttt agcattggag atatgagtgg 2280

tagagcgtat acagtattaa ttgtggaggt atctcgtcgc attgatagag caatacagtt 2340

actgctgaag ggaatgatac cgagtatttc ggcccgattc agttcttgat atcgtcattt 2400

tgtctctatt gtctactttt cagataacct caacaaatct tcaacaaatc tcccagtaaa 2460

cagtcagaga tcatatccga gatcatatca gatatgtcac gatccgagta caataatgga 2520

tattaatctg cttgattttg aattctgttg cgattatgat ttctttgatt tcgatatgaa 2580

cacatacggc gactcccaga cctttagaag ctccagtttg gattcttagc aatggttaca 2640

ctcaactata tcccaagtaa tacttggtaa caatatgcca agttagtcat tcattcgtta 2700

taggagttag caagtgtttg tcagctaaaa atggttagtc ggtcgattac cacttagatc 2760

ttttcagcgt ggaacttgat ggtacgcttg aaccgacact tggagtagtc ggggctgttg 2820

atgacgtaga tgacgtttcg ctcagggtga ggagtgcaat agtagtactc cttggggccg 2880

tctctcagct caaaggttcc atcggcggca atgtcaaaga ccgagccctg gagcttgtag 2940

ccgtagtcgc cggtccagaa caaagcctgc agctccagat aggcgatggg catgtcgtta 3000

acagagaagg tgttgccctc gccctcggtg atggtgatgg gttcgccgtc ggtggaggcg 3060

gtgatcaggt catcttggta ggtgacgggc agagattcga ccgattgggc gtctgatctg 3120

gtataggtca gctt 3134

<210> 69
<211> 377
<212> PRT
<213> Yarrowia lipolytica

<400> 69

Met Trp Gly Ser Ser His Ala Phe Ala Gly Glu Ser Asp Leu Thr Leu
1 5 10 15

Gln Leu His Thr Arg Ser Asn Met Ser Asp Asn Thr Thr Ile Lys Lys
20 25 30

Pro Ile Arg Pro Lys Pro Ile Arg Thr Glu Arg Leu Pro Tyr Ala Gly
35 40 45

Ala Ala Glu Ile Ile Arg Ala Asn Gln Lys Asp His Tyr Phe Glu Ser
50 55 60

Val Leu Glu Gln His Leu Val Thr Phe Leu Gln Lys Trp Lys Gly Val
65 70 75 80

Arg Phe Ile His Gln Tyr Lys Glu Glu Leu Glu Thr Ala Ser Lys Phe
85 90 95

Ala Tyr Leu Gly Leu Cys Thr Leu Val Gly Ser Lys Thr Leu Gly Glu
100 105 110

```
Glu Tyr Thr Asn Leu Met Tyr Thr Ile Arg Asp Arg Thr Ala Leu Pro
        115                 120                 125

Gly Val Val Arg Arg Phe Gly Tyr Val Leu Ser Asn Thr Leu Phe Pro
    130                 135                 140

Tyr Leu Phe Val Arg Tyr Met Gly Lys Leu Arg Ala Lys Leu Met Arg
145                 150                 155                 160

Glu Tyr Pro His Leu Val Glu Tyr Asp Glu Asp Glu Pro Val Pro Ser
                165                 170                 175

Pro Glu Thr Trp Lys Glu Arg Val Ile Lys Thr Phe Val Asn Lys Phe
            180                 185                 190

Asp Lys Phe Thr Ala Leu Glu Gly Phe Thr Ala Ile His Leu Ala Ile
        195                 200                 205

Phe Tyr Val Tyr Gly Ser Tyr Tyr Gln Leu Ser Lys Arg Ile Trp Gly
    210                 215                 220

Met Arg Tyr Val Phe Gly His Arg Leu Asp Lys Asn Glu Pro Arg Ile
225                 230                 235                 240

Gly Tyr Glu Met Leu Gly Leu Leu Ile Phe Ala Arg Phe Ala Thr Ser
                245                 250                 255

Phe Val Gln Thr Gly Arg Glu Tyr Leu Gly Ala Leu Leu Glu Lys Ser
            260                 265                 270

Val Glu Lys Glu Ala Gly Glu Lys Glu Asp Glu Lys Glu Ala Val Val
        275                 280                 285

Pro Lys Lys Lys Ser Ser Ile Pro Phe Ile Glu Asp Thr Glu Gly Glu
    290                 295                 300

Thr Glu Asp Lys Ile Asp Leu Glu Asp Pro Arg Gln Leu Lys Phe Ile
305                 310                 315                 320
```

Pro Glu Ala Ser Arg Ala Cys Thr Leu Cys Leu Ser Tyr Ile Ser Ala
325 330 335

Pro Ala Cys Thr Pro Cys Gly His Phe Phe Cys Trp Asp Cys Ile Ser
340 345 350

Glu Trp Val Arg Glu Lys Pro Glu Cys Pro Leu Cys Arg Gln Gly Val
355 360 365

Arg Glu Gln Asn Leu Leu Pro Ile Arg
370 375

<210> 70
<211> 3740
<212> DNA
<213> Yarrowia lipolytica

<400> 70
tgaagcggga aatcaagatt gagagttgga gtggtggggg cggtggtgga agagacagag 60

gtgatagaag ggaaagaaac gacagaagag acgcaaaagg aagaaacaga agagatggag 120

ataggaggaa tggcgatcga agagacagag acagagatag agatagagac gacaggtcaa 180

gtcgttcaag tcggtacaac cgatcaagat caagatctcc gactgtcaaa aaggaaacag 240

atgaatatgg ccgggaccgg aaggactaac tgcattatat ataaatatag attcttgatt 300

agtgatcaaa gcatgagata ctgtcaaggc aagcaagcct tacaatatct gttggcttag 360

ttaaatcttt tttgagatgt aataattggg gaacctctgt tagatgaaca gttgcggctt 420

actcatctta gacatggtcc agtttaggcg taattaatta tcagttttaa tgcgcctcta 480

cgtaataagc gttgccacaa gttacagtag agatgatacc attgctacag gaccgtatag 540

cgtatataat gacgaggcta attgatgatg atcgacagcg aagacatatc tacgtcaatc 600

tgacgcttca ttttgtgctt cttgttcgga tttgtgtttc tcccagtgca cccagttttt 660

tgctgaccat tttttgctta ccagtttttg ctcgcgcgtt ttctcgattg aactctgctg 720

cgagggtggt gtaagacaca ctccctgtag aacatgcatt tcgaggtgtg acatttccat 780

ctctgttaaa tcaaacgagt gacacaaacg gaataaatct acatatacgt cgaaccctct 840

tcttctcttc tctctcttcc ttaatagaca ccctcgctat cgactctgct cctccttatc 900

acaccatcaa agtttagggt gtgtcccaac accacctcca caccacactc tattatcacc 960

acctccacga ctacatacca ctccttccac cacctcacaa atggcgaccg aacacgtgga 1020

acacgccgaa aacgccaagc agtcgcccca accgtccttc tcgggctccg agagcggccg 1080

aattggccgg taccaaatca tcaagactct gggagagggc tcctttggca aggtcaagct 1140

cgcctaccat ctggccaccc acgaaaaagt agcgctcaaa atcatcaaca gaaagacgct 1200

ggccaagtcg gatatgcagg gccgagtcga gcgggaaatc tcgtacctgc gactgctcag 1260

acacccccac atcatcaagc tgtacgacgt gatcaagtcc aaggatgaaa tcatcatggt 1320

gattgaattt gccggcaagg agctctttga ctacattgtg caacgtggaa agatgcccga 1380

agacgaggct aggcgattct tccaacaaat catctccgca gtcgagtact gccatagaca 1440

caagattgtg cacagagacc tcaagcccga gaacctgttg ctcgacgaaa acctcaacgt 1500

caagattgcc gactttggac tctccaacat tatgaccgat ggtaacttcc tcaagaccag 1560

ttgtggctct cccaactatg ctgctcccga ggttatcagt ggcaagctgt atgctggccc 1620

cgaggtcgac gtgtggtcct gtggtgtcat tctatacgtc atgctgtgcg gccgtctccc 1680

ctttgacgac gaattcatcc ccaacctctt caagaagatc tccaatggtg tctacaccat 1740

tccccctat ctgtctgccg gcgcaaaaca cctgcttcaa cagatgctgg tggtcaaccc 1800

cctgaaccga atcaccgtcc agggcatcat ggaagacccc tggtttaaac agggcattgc 1860

cgagtacctt gttccacgtg acctgaagcc tgatcaggtg gaaatcgacg acaaggttgt 1920

gggtgctctt tcagacacca tgggctacga ccgagaccag gtctacgagg ctctcaaggc 1980

tcctaaatcc gacggaaccg aaatccgaga tgcctacgac ctcatgaagg agaacagaca 2040

gcagttggag aaggactcga ccgtgttcga caaggtggac gaatccggca agcctaagca 2100

ccagttctcg cgaaagaacc cgccggctct gacgtttgac tcacatcaca tgcccaccat 2160

tgtcgaccat tctgatcttc gaacccctaa cacaaccatt gctgttcttc cctcctcgct 2220

gccagcatac catcgggcca acatgatggc ccatggtccc gctaccctgc gtaagctgaa 2280

tcccatctca tcgcgaaaat caaagacccg atggcatttc ggcattcggt ccaagtcgta 2340

ccctctcgac gtgatgggtg agctctaccg agctctcaag aacctgggag ccgagtgggc 2400

aaagccatca gagaacgacc tgtggaccat cagggtgcga tggagataca cccctgaaaa 2460

ggatgagtcg gacagcacaa ctaagccaga cagcggactg cttaagatgc agattcaact 2520

gtatctgctg gagcagaact cgtacttggt ggacttcaag tttgacggct gggagtgcga 2580

gacggagagc ggtgagatta agacccgcga tgtcgcccag gtgtcgaatt caaagtccgc 2640

caacgagacc ttcacgtttt ccgcataccc cttcttgcat ctggccaccc ggctcattat 2700

ggagctggcc gtgagttctc agaaggagag tgagaagtaa gcacttgtag agcacactag 2760

ggatttagag gggattatgg cacgtacaat atagataatt aagcagtagc tgagtcagtt 2820

gaatgatcag aggtgtaaca tgagtgtgga tggattgtgt agagtcgttg tttaaaataa 2880

tgagttaaga ataattatac gactacagga tacgatgtac ttgtattgta tcgatacagt 2940

acatacagta catacgtgta acatactcct aaactgttgc atcacctaca actccaacta 3000

gtcggtcata attcattaat acgtttccct ggtgtagtct aggccaaaca ctccttggtg 3060

aacttcttga gaaagattcg ctctctagac tccttgaatt ctctaggctg gatttcggcc 3120

caaaaagaac ccacaaacgg cgcaatactc ttcttgtagc cggaatagtt gtctcgaatg 3180

acaatatccg ggttggcaga ccgctgctga tgctccttga cgatgatact ctcaaagtgg 3240

tccacggtag acatgacaat aagagccgac gaggttttgc tcttgtagtt ggcgatgagc 3300

gttttaagtg ccccagagtt ctccaccacc ttgtggatca gctgatgctt ttgttcctcg 3360

ctgagtcctt ccgaggccag ggggatagcc aagcaggtga gaatagcgcc tgtgtctcgt 3420

ttagaagagg tgctgtttgt atcatccgca tgcttgttga tgaactgaac taaattcagt 3480

agcgtgctcc acagctcatt ccagtggtaa tttgagggag tgtccctgta aaccatgagt 3540

gttgcttcag tgacgaccag agcgagatca taaatgtctg ggagaggttt cttcatattg 3600

tatctgagac agcaaagaag tgcgtccaaa aggccggttc caaaggtgta ggcctctgtt 3660

tctggaaggg ggggtttccg ttgcttggcg ttgattttaa tagatgcttt tgtgttgaac 3720

gactggtgca attcgggcag 3740

<210> 71
<211> 579
<212> PRT
<213> Yarrowia lipolytica

<400> 71

```
Met Ala Thr Glu His Val Glu His Ala Glu Asn Ala Lys Gln Ser Pro
1               5                   10                  15


Gln Pro Ser Phe Ser Gly Ser Glu Ser Gly Arg Ile Gly Arg Tyr Gln
            20                  25                  30


Ile Ile Lys Thr Leu Gly Glu Gly Ser Phe Gly Lys Val Lys Leu Ala
        35                  40                  45


Tyr His Leu Ala Thr His Glu Lys Val Ala Leu Lys Ile Ile Asn Arg
    50                  55                  60


Lys Thr Leu Ala Lys Ser Asp Met Gln Gly Arg Val Glu Arg Glu Ile
65                  70                  75                  80


Ser Tyr Leu Arg Leu Leu Arg His Pro His Ile Ile Lys Leu Tyr Asp
                85                  90                  95


Val Ile Lys Ser Lys Asp Glu Ile Ile Met Val Ile Glu Phe Ala Gly
            100                 105                 110


Lys Glu Leu Phe Asp Tyr Ile Val Gln Arg Gly Lys Met Pro Glu Asp
        115                 120                 125


Glu Ala Arg Arg Phe Phe Gln Gln Ile Ile Ser Ala Val Glu Tyr Cys
    130                 135                 140


His Arg His Lys Ile Val His Arg Asp Leu Lys Pro Glu Asn Leu Leu
145                 150                 155                 160


Leu Asp Glu Asn Leu Asn Val Lys Ile Ala Asp Phe Gly Leu Ser Asn
                165                 170                 175
```

```
Ile Met Thr Asp Gly Asn Phe Leu Lys Thr Ser Cys Gly Ser Pro Asn
            180                 185                 190

Tyr Ala Ala Pro Glu Val Ile Ser Gly Lys Leu Tyr Ala Gly Pro Glu
        195                 200                 205

Val Asp Val Trp Ser Cys Gly Val Ile Leu Tyr Val Met Leu Cys Gly
    210                 215                 220

Arg Leu Pro Phe Asp Asp Glu Phe Ile Pro Asn Leu Phe Lys Lys Ile
225                 230                 235                 240

Ser Asn Gly Val Tyr Thr Ile Pro Pro Tyr Leu Ser Ala Gly Ala Lys
                245                 250                 255

His Leu Leu Gln Gln Met Leu Val Val Asn Pro Leu Asn Arg Ile Thr
            260                 265                 270

Val Gln Gly Ile Met Glu Asp Pro Trp Phe Lys Gln Gly Ile Ala Glu
        275                 280                 285

Tyr Leu Val Pro Arg Asp Leu Lys Pro Asp Gln Val Glu Ile Asp Asp
    290                 295                 300

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

Val Tyr Glu Ala Leu Lys Ala Pro Lys Ser Asp Gly Thr Glu Ile Arg
                325                 330                 335

Asp Ala Tyr Asp Leu Met Lys Glu Asn Arg Gln Gln Leu Glu Lys Asp
            340                 345                 350

Ser Thr Val Phe Asp Lys Val Asp Glu Ser Gly Lys Pro Lys His Gln
        355                 360                 365
```

```
Phe Ser Arg Lys Asn Pro Pro Ala Leu Thr Phe Asp Ser His His Met
    370                 375                 380

Pro Thr Ile Val Asp His Ser Asp Leu Arg Thr Pro Asn Thr Thr Ile
385                 390                 395                 400

Ala Val Leu Pro Ser Ser Leu Pro Ala Tyr His Arg Ala Asn Met Met
                405                 410                 415

Ala His Gly Pro Ala Thr Leu Arg Lys Leu Asn Pro Ile Ser Ser Arg
            420                 425                 430

Lys Ser Lys Thr Arg Trp His Phe Gly Ile Arg Ser Lys Ser Tyr Pro
        435                 440                 445

Leu Asp Val Met Gly Glu Leu Tyr Arg Ala Leu Lys Asn Leu Gly Ala
    450                 455                 460

Glu Trp Ala Lys Pro Ser Glu Asn Asp Leu Trp Thr Ile Arg Val Arg
465                 470                 475                 480

Trp Arg Tyr Thr Pro Glu Lys Asp Glu Ser Asp Ser Thr Thr Lys Pro
                485                 490                 495

Asp Ser Gly Leu Leu Lys Met Gln Ile Gln Leu Tyr Leu Leu Glu Gln
            500                 505                 510

Asn Ser Tyr Leu Val Asp Phe Lys Phe Asp Gly Trp Glu Cys Glu Thr
        515                 520                 525

Glu Ser Gly Glu Ile Lys Thr Arg Asp Val Ala Gln Val Ser Asn Ser
    530                 535                 540

Lys Ser Ala Asn Glu Thr Phe Thr Phe Ser Ala Tyr Pro Phe Leu His
545                 550                 555                 560

Leu Ala Thr Arg Leu Ile Met Glu Leu Ala Val Ser Ser Gln Lys Glu
                565                 570                 575
```

Ser Glu Lys

<210> 72
<211> 7490
<212> DNA
<213> Yarrowia lipolytica

<400> 72
gcgtccaaaa aattcaaatt tcgacatttt tcggcttaaa atagaacatc tacaagacga 60

tgttgcagac tcctgagtct gggaaaaggt ggttgaggaa cctacactgg agctgcgatt 120

ggagagcttc gtggcggccg tgggccaaga ttgggcagag tatgggtcga gcatttctat 180

tgcttctact gtacttgtac ttgtacttgt acttacttac ttgtacaagt acagacagga 240

ggcttgctat ttcgatgaaa ctaaagcacg tttgttatgt accgaagttc ggaacgaggc 300

tgtggagatg gcaagtagtt aagtgtcatt aattggagta ggttaaaaca tcaactgtgc 360

tcgttctaac tgcttactgc gcctgttata ccaccgctag cgccctctaa cggatcctag 420

gtgcgggttg aactggttta tgattgctgt ttttacactt tttagacccc cagagcgtgc 480

aggtgatgga gtgaacgcgc ggtgctacag taatactcta agacaagttt gggcagggaa 540

ggaaacacct acaggaggcc aggtggccaa gatttagacg gtttggggag ttgtatttgt 600

tgccacgtac aaagggacaa aaacacgggc ttgtagtggc cgcaattcca ctgctcttgt 660

gcttgtcagc tgctactctt cgtctaagag gacccttct taatgtgaac tcagctgcat 720

gaccgcgcat acttctttcc tactgcagcc cctgacaggc aatactagca agtgaagcat 780

ccattctacg acagaggtat tttcaggacg aaaaaacact cttgagtcgc cacaaaaagt 840

gagtgttccg acaacgcggg cggagcagtt tggagtttgg gagcagtgta ttgatctgac 900

tgagcttgag agagtgtgtg acattctgga gggagaaggg aaaggggaga tgatatgatg 960

gatatgggtt ggggagatga tggtatagga ctggttgtat atggaaacga tgcttgatcg 1020

tgttattcca ctcagctgtg tacctcctta cgccacaccc tacagtcacg ttttcaccac 1080

ccccatctca aacaacacta acacagtgtt ctacgaagac accgccgact ccgattggca 1140

gtcgtccgac tacgacagtg aacatgcgac cgtcaacgtg cccatccgga gcaaacggaa 1200

aaagaatgct gctaaggagg gcaatccagt ggatcccatc tctgaggatc cgcaccagga 1260

atccttttca gatggaagtg aagctgcggg acccagcgct cccccggccg acggatacat 1320

caacaccgcg cccaatggcg tctactcgga ggaagatctt cctggaaccc gccaggtgct 1380

tccaggacct caggaggtac ccgacgtgac ctccactccc cagatgaacc gccatcccac 1440

agcagatgaa tctactcctc agggctcacg gtggagccga ggcagagagg gatttggaca 1500

tcttggaagt ccctcgttct ttagacgaca cagtgccaca gctgaagata tgacccctgg 1560

aggaggacct cgacggctgt tcaagcgacc acagctcaac tcttccaaca ccacatcggg 1620

ccagcgttgg agagatatca aagccgggct tcgtcagcgg ttccgaaaga agaaggacgc 1680

gaaattcagc gctgagaagg catacaatca agccattctg ggccagcttg ctgctggttc 1740

tcctgcggct ctcatggttg cttcctccat ggttcgcgac gaaaagaaca tgaagcgagt 1800

gcccattctg ctggagcagg tccgtttcgc tctcagagat gtgacaaaga accctcacga 1860

gaagaatcga cagtatcgaa tcgatctttc gtacggttcc ggtcacacga tgctcgcatg 1920

gacaatctac aaggactaca aggactttat gtttctgcac agtcggctgc gaaccatggc 1980

cttcaacgcc aagaacccct tcaactcgtc aaaacagcct cttcagatcc ccatctttcc 2040

tacccgaaac aaggtgggca agaaaaaaaa aggcaagaga gacaagtcca accctggcga 2100

tgaagattac acagacacag atggagagtc tgagtgtgat cccgcagagc agcaacaaaa 2160

ccagcacact ccaggagggg gatcttcgca tggacctagg atggtgaacc agttctttga 2220

caagcgacgg aaaaaccgca acatgatcaa cgagcgcatt aagggtgaac tggaacgata 2280

cctgcgtcag ctgttcaaac tgctgctctt ccgaggtgaa gctaacaagc tgttccagtt 2340

tttggagctg tccaacatgt ccattcgact tgctcctgag gaccatttcc acggcaaggc 2400

tggattcctg gtcaaccgta cgtctgctcg aaaggctgga tggagagtgt cgcattggcg 2460

agtcgacgat atcagacaga tggtggctcg tcacacgtcg aagtggttca tggtgcgcca 2520

ctcgtacatt gtgtgtgtcg acaacctgta ttccacaaca cctcttgagg tgttccttct 2580

cgattccaag ttcaaggtgt cacacaatct caaaaagaag aacaaggacg gcgttctttc 2640

gggtggatcc aaccccaata cttctggaga gtccgatttg gaagaacttc agaaaccggg 2700

aagtacccat attactctca aggtggagaa cgcatcccga cagcttaagc tggttgccac 2760

atctgacaag cagctggccc agtggatgga gagtatcaat ctgatcaagg agcggtccat 2820

ttgggctcag cagaagcggt ttgactcatt tgcgcctgtg cgaaccaact gcaaggccca 2880

atggttcgtc gacgctcgag actacttttg gacgctgtcg tgcgctctgg atatggccaa 2940

ggaggtcatt tacattcacg actggtggct gtctcctgag atatacctac gtcgaccccc 3000

tgagggtaac caggagtggc gtctggatcg tgttttgaag cgaaaggccg agcagggagt 3060

caagattttt gttattgtgt accgaaatgt cggtcagact attcccattg actcgcagta 3120

caccaagttc tcgctgctgg acctgtctcc taacatttat gtcatgcggt ctcccaacca 3180

gctcattcag aatacctact tttgggccca tcacgaaaag ctgtgtctta ttgatcacac 3240

gtgtgccttt gtgggtggta ttgatctctg ttttggtcgt tacgatactg cggaacacgt 3300

tctcgtcgac gatgcgccca cgtttgtgaa tgccgacaag aaggacggct acaccggtcg 3360

aactcagctc ttccctggta aggattactc taacccccgt accaaggact tcttctcgtt 3420

ggacaagcct tttgaagata tgtacgaccg tcagaaggtg ccccgcatgc cctggcacga 3480

tatccacatg atggtagtag gccaacctgc acgtgatctg gtgcgtcatt tcgtgcagcg 3540

atggaactat gttcttcgac agaagcgtcc ttcccgattc acgcctcttc tgctgcctcc 3600

tccagacttc aaggaggagg agctggccga gtacaagctg aatggtacgt gtgaggttca 3660

gattctgcgg tctgcttgtt cctggaacac tggagtcaag gagcacgagc agagtatcca 3720

gaatgcgtac atcaagagca ttgagcagag tgagcatttc gtctacattg agaaccagtt 3780

ttttatcact cacacctcgt ggcacaatat tgtgattgaa aacaagatcg gagacgctct 3840

ggtcaaccga atcatcaagg cccatcagaa tgacgaggat tggcgtgcca tcattgtcat 3900

tcctctcatg cctggttttg aagccgaggt cgacgagtcc gagggttcgt cggtgcgtgt 3960

catcatgcag tgccagtaca tgtccatttc tcgaggtgcc aactgcattt ttgcacgtct 4020

tgagaacgct ggcatccatc ccgaagacta catttgcttc ttttcactgc gaaagtgggg 4080

caagatcggt ccccacgaca agctggtcac cgagcagctg tacatccacg ccaaggctat 4140

ggttgtcgat gaccgtattg ccattattgg ttcggcgaac atcaacgaac ggtctcaacg 4200

aggtactcga gactctgagg tcgcagctat tgttcgagac acccatactg tggactctgt 4260

catggctggg cgtccctaca aggtgggcca ctttgcacac accctgcgtc tccatctgat 4320

gcgagagcac atgggcgttg atgtggacca tgtggagttc attgagcgaa aggctgagct 4380

tcttgagaaa cacaagagag ctgaggctac ccaagatgag ccccaggagc ttcctatcat 4440

caagaacaag tccacacgcc gagcacgggg gcgtggacga cgttcccagc gaggaggaga 4500

ggattccgat ggagcttccg gctctgaaga tgagcgagtt ctcgatatct caacctctat 4560

cttcggggag acggttcagg ctgccaaggc cgctaacggt aatggcaatg gcaacggtaa 4620

cggtgctgat gcgtctgctg cggagtctgt cccttccgtc atggctccca cccgaagata 4680

ctcttaccag gatggagtac ctgcctctcc ttctattgac gaggactgtg ttgtcgaagg 4740

cacaactctg gacaatgctg accagcagct agccaacttg gacgatgcgt acgattccga 4800

cgccgaggtg ctactgcgag actatgatga ggatgtggat cagatcaacc ggggcttgga 4860

catctacacc ttcaactctc tagccggtgt ggacaacaag ggcatccgag agaagaagct 4920

tgtctcttct gataaccgta tccagggcaa cgagaagcac cgtgctgacg ttgagggccg 4980

gggtctggac cgaggcgcca actcgaagga gtctgcggag gacttccgaa tccgaatcca 5040

gcgacacgcc cagtttgcca acaagctgac ggagcacgag ctgcagcaga ccagcgagac 5100

ggtggacgat gtgacggaac ttaagaagcg cgtctactac gagcagatgg agctcaagca 5160

gcggtaccac aattcgcaca tggaggccga gaaggcgttc aacaatcccc cacagttgga 5220

ggactgcgag aacgagaaga acggcgttaa tggagtcaac ggcaacagtg gctccactga 5280

ttccactgca cctgtaaatc ttgtctccga tgctaacggc aacattcacc ctgaacatcc 5340

tgagacatct gatcattctt cggactcttt tgaagatgag ggcttcgaga tccccaagat 5400

ccctcctccc gtagacccct actgttttaa ggacccttg gacgacgact tttactacga 5460

catctggctg gccacagccg aaaagaacac caggctcttc cgggaagtgt ttcgatgcca 5520

gcctgacgac gaggtgacca cctggagaga ttacaaggat ttcaccactt acgccgagcg 5580

gttctccatg tctcaggaca ccggcgactt gaagtcttcg tatgcagatg aacaccaagc 5640

tcattccccc ttggctgtca agatcaatgg caatgacaag gaagacgaga agaaacagaa 5700

caagcacaag gccacacatg actacgacaa gtattatgcc cgtcaccagc accaccgatc 5760

gtcttcgcct cactaccgaa ggtcccatgt gcctctcaag accggtgcca acttcactga 5820

tcgacttgca cacggtgttg atgatgccgc tggtcatctc aagtacgccc tcggacgggc 5880

caaccgccga ggcagccacc atgacactgt tgatgtggat gattctcctc ccagggactt 5940

ccttcctgac gatttcaacg acccgatccc tgtcgaggag gtggacaaca ccaacctcaa 6000

caacaaccag cagctcaacc aggagcttta cgagtcgcag gtggagggcg agaaccctcg 6060

atttgtgcct agccacgcca ttgatggcca gcaccctcga aagcgtgaga agcgacacat 6120

tggtgaggtt gatcagcctg ccgagaagcc ccccatcaac gaggcttaca gcactggggt 6180

ctcaggccgg cttgggcagc accagcagcc ccatcttaag caggaagttg atactgattt 6240

taaggatacc tcttctgctg ctcatacgcc attccaccat acccccaacg ccaggcgaag 6300

acgacgacga cgggctcgac agagatctgg catggacaag gtgcacgaca aggcaacggc 6360

cgagaagatt ctcaacggta tccaggggca tctggtcatg ttccccaccc actggctgta 6420

caaggaactc gacagcggca actggttcta caacctggat cgaatccccc cgatcgaaat 6480

ttatgattag aatggtaata tattggtatg ataaataatg caagcacttg catgagtgta 6540

ctgtacgtaa gcaagtagtg acagttcagt taatcaaatg caaacccgct gcgacaaaca 6600

cgtgctaagc tcttcagatt cttcatcact ggtctcgaaa tcaaagagct cacaggtgtc 6660

ttcttcctct ttctggtcac tggcaaagtt ccagtattga cgacggacga ttcccgaaag 6720

cgtccagatt tccacctgtt gtttgatctt gtcgaagaag attagtgcaa tgatgaccga 6780

ggccagctcg ggaatgatgt acgggagtcg agacatgatc tcgtctggtt ttaggggagt 6840

cttgtcgtct tcggggagat ctgtgaaggt ggtgtacaga aacgtgaatg caatatggat 6900

ctgaagaatc ttttgtgttt ccagagcatt gtgaacgatt ttcacgatgg tggcgtcgta 6960

ggcagccttc ttatgagcgt caggaggtcc attcagcccg gtggagaaag agtccacgat 7020

cacgtatctg attttcgagc tgttgtttga cttttcaatg tgtccgatga ttgtgtctga 7080

ttgcagctcg gtggcgagct ggtcgaaggt ggttttcaca atcacgaggt tctggcgaat 7140

ataactggca agcacatttt cacccagacc agtggattga tccagcacaa agtagttgat 7200

gacgttgttt ttcgtccatg agctgtcgag atccagaacc agaaccgtct tgttctcgcc 7260

tcctggagtg tcgaatggca aacacacacg gcacagctcg tacagcgccg tggttttgcc 7320

tgatcttgga gctccggtca gagtgatcag ccctgcaagc tgtgctgatg gcaaatgcag 7380

ccagtgtaac aagttttctg ccgcatcaga cggaccgtaa tgctgaatgg cgtcgaacaa 7440

gttttccttc ttgttttccg tttccatgat ggtagttgtc agtgatgttc 7490

<210> 73
<211> 1829
<212> PRT
<213> Yarrowia lipolytica

<400> 73

Met Glu Thr Met Leu Asp Arg Val Ile Pro Leu Ser Cys Val Pro Pro
1 5 10 15

Tyr Ala Thr Pro Tyr Ser His Val Phe Thr Thr Pro Ile Ser Asn Asn
20 25 30

Thr Asn Thr Val Phe Tyr Glu Asp Thr Ala Asp Ser Asp Trp Gln Ser
35 40 45

Ser Asp Tyr Asp Ser Glu His Ala Thr Val Asn Val Pro Ile Arg Ser
50 55 60

Lys Arg Lys Lys Asn Ala Ala Lys Glu Gly Asn Pro Val Asp Pro Ile
65 70 75 80

Ser Glu Asp Pro His Gln Glu Ser Phe Ser Asp Gly Ser Glu Ala Ala
85 90 95

Gly Pro Ser Ala Pro Pro Ala Asp Gly Tyr Ile Asn Thr Ala Pro Asn
100 105 110

Gly Val Tyr Ser Glu Glu Asp Leu Pro Gly Thr Arg Gln Val Leu Pro
115 120 125

```
Gly Pro Gln Glu Val Pro Asp Val Thr Ser Thr Pro Gln Met Asn Arg
    130                 135                 140

His Pro Thr Ala Asp Glu Ser Thr Pro Gln Gly Ser Arg Trp Ser Arg
145                 150                 155                 160

Gly Arg Glu Gly Phe Gly His Leu Gly Ser Pro Ser Phe Phe Arg Arg
                165                 170                 175

His Ser Ala Thr Ala Glu Asp Met Thr Pro Gly Gly Gly Pro Arg Arg
            180                 185                 190

Leu Phe Lys Arg Pro Gln Leu Asn Ser Ser Asn Thr Thr Ser Gly Gln
        195                 200                 205

Arg Trp Arg Asp Ile Lys Ala Gly Leu Arg Gln Arg Phe Arg Lys Lys
    210                 215                 220

Lys Asp Ala Lys Phe Ser Ala Glu Lys Ala Tyr Asn Gln Ala Ile Leu
225                 230                 235                 240

Gly Gln Leu Ala Ala Gly Ser Pro Ala Ala Leu Met Val Ala Ser Ser
                245                 250                 255

Met Val Arg Asp Glu Lys Asn Met Lys Arg Val Pro Ile Leu Leu Glu
            260                 265                 270

Gln Val Arg Phe Ala Leu Arg Asp Val Thr Lys Asn Pro His Glu Lys
        275                 280                 285

Asn Arg Gln Tyr Arg Ile Asp Leu Ser Tyr Gly Ser Gly His Thr Met
    290                 295                 300

Leu Ala Trp Thr Ile Tyr Lys Asp Tyr Lys Asp Phe Met Phe Leu His
305                 310                 315                 320
```

```
Ser Arg Leu Arg Thr Met Ala Phe Asn Ala Lys Asn Pro Phe Asn Ser
                325                 330                 335

Ser Lys Gln Pro Leu Gln Ile Pro Ile Phe Pro Thr Arg Asn Lys Val
            340                 345                 350

Gly Lys Lys Lys Lys Gly Lys Arg Asp Lys Ser Asn Pro Gly Asp Glu
        355                 360                 365

Asp Tyr Thr Asp Thr Asp Gly Glu Ser Glu Cys Asp Pro Ala Glu Gln
    370                 375                 380

Gln Gln Asn Gln His Thr Pro Gly Gly Gly Ser Ser His Gly Pro Arg
385                 390                 395                 400

Met Val Asn Gln Phe Phe Asp Lys Arg Arg Lys Asn Arg Asn Met Ile
                405                 410                 415

Asn Glu Arg Ile Lys Gly Glu Leu Glu Arg Tyr Leu Arg Gln Leu Phe
            420                 425                 430

Lys Leu Leu Leu Phe Arg Gly Glu Ala Asn Lys Leu Phe Gln Phe Leu
        435                 440                 445

Glu Leu Ser Asn Met Ser Ile Arg Leu Ala Pro Glu Asp His Phe His
    450                 455                 460

Gly Lys Ala Gly Phe Leu Val Asn Arg Thr Ser Ala Arg Lys Ala Gly
465                 470                 475                 480

Trp Arg Val Ser His Trp Arg Val Asp Asp Ile Arg Gln Met Val Ala
                485                 490                 495

Arg His Thr Ser Lys Trp Phe Met Val Arg His Ser Tyr Ile Val Cys
            500                 505                 510

Val Asp Asn Leu Tyr Ser Thr Thr Pro Leu Glu Val Phe Leu Leu Asp
        515                 520                 525
```

```
Ser Lys Phe Lys Val Ser His Asn Leu Lys Lys Lys Asn Lys Asp Gly
    530                 535                 540

Val Leu Ser Gly Gly Ser Asn Pro Asn Thr Ser Gly Glu Ser Asp Leu
545                 550                 555                 560

Glu Glu Leu Gln Lys Pro Gly Ser Thr His Ile Thr Leu Lys Val Glu
                565                 570                 575

Asn Ala Ser Arg Gln Leu Lys Leu Val Ala Thr Ser Asp Lys Gln Leu
            580                 585                 590

Ala Gln Trp Met Glu Ser Ile Asn Leu Ile Lys Glu Arg Ser Ile Trp
        595                 600                 605

Ala Gln Gln Lys Arg Phe Asp Ser Phe Ala Pro Val Arg Thr Asn Cys
    610                 615                 620

Lys Ala Gln Trp Phe Val Asp Ala Arg Asp Tyr Phe Trp Thr Leu Ser
625                 630                 635                 640

Cys Ala Leu Asp Met Ala Lys Glu Val Ile Tyr Ile His Asp Trp Trp
                645                 650                 655

Leu Ser Pro Glu Ile Tyr Leu Arg Arg Pro Pro Glu Gly Asn Gln Glu
            660                 665                 670

Trp Arg Leu Asp Arg Val Leu Lys Arg Lys Ala Glu Gln Gly Val Lys
        675                 680                 685

Ile Phe Val Ile Val Tyr Arg Asn Val Gly Gln Thr Ile Pro Ile Asp
    690                 695                 700

Ser Gln Tyr Thr Lys Phe Ser Leu Leu Asp Leu Ser Pro Asn Ile Tyr
705                 710                 715                 720
```

```
Val Met Arg Ser Pro Asn Gln Leu Ile Gln Asn Thr Tyr Phe Trp Ala
                725                 730                 735

His His Glu Lys Leu Cys Leu Ile Asp His Thr Cys Ala Phe Val Gly
            740                 745                 750

Gly Ile Asp Leu Cys Phe Gly Arg Tyr Asp Thr Ala Glu His Val Leu
        755                 760                 765

Val Asp Asp Ala Pro Thr Phe Val Asn Ala Asp Lys Lys Asp Gly Tyr
    770                 775                 780

Thr Gly Arg Thr Gln Leu Phe Pro Gly Lys Asp Tyr Ser Asn Pro Arg
785                 790                 795                 800

Thr Lys Asp Phe Phe Ser Leu Asp Lys Pro Phe Glu Asp Met Tyr Asp
                805                 810                 815

Arg Gln Lys Val Pro Arg Met Pro Trp His Asp Ile His Met Met Val
            820                 825                 830

Val Gly Gln Pro Ala Arg Asp Leu Val Arg His Phe Val Gln Arg Trp
        835                 840                 845

Asn Tyr Val Leu Arg Gln Lys Arg Pro Ser Arg Phe Thr Pro Leu Leu
    850                 855                 860

Leu Pro Pro Pro Asp Phe Lys Glu Glu Glu Leu Ala Glu Tyr Lys Leu
865                 870                 875                 880

Asn Gly Thr Cys Glu Val Gln Ile Leu Arg Ser Ala Cys Ser Trp Asn
                885                 890                 895

Thr Gly Val Lys Glu His Glu Gln Ser Ile Gln Asn Ala Tyr Ile Lys
            900                 905                 910

Ser Ile Glu Gln Ser Glu His Phe Val Tyr Ile Glu Asn Gln Phe Phe
        915                 920                 925
```

```
Ile Thr His Thr Ser Trp His Asn Ile Val Ile Glu Asn Lys Ile Gly
    930                 935                 940

Asp Ala Leu Val Asn Arg Ile Ile Lys Ala His Gln Asn Asp Glu Asp
945                 950                 955                 960

Trp Arg Ala Ile Ile Val Ile Pro Leu Met Pro Gly Phe Glu Ala Glu
                965                 970                 975

Val Asp Glu Ser Glu Gly Ser Ser Val Arg Val Ile Met Gln Cys Gln
            980                 985                 990

Tyr Met Ser Ile Ser Arg Gly Ala  Asn Cys Ile Phe Ala  Arg Leu Glu
        995                 1000                1005

Asn Ala  Gly Ile His Pro Glu  Asp Tyr Ile Cys Phe  Phe Ser Leu
    1010                1015                1020

Arg Lys  Trp Gly Lys Ile Gly  Pro His Asp Lys Leu  Val Thr Glu
    1025                1030                1035

Gln Leu  Tyr Ile His Ala Lys  Ala Met Val Val Asp  Asp Arg Ile
    1040                1045                1050

Ala Ile  Ile Gly Ser Ala Asn  Ile Asn Glu Arg Ser  Gln Arg Gly
    1055                1060                1065

Thr Arg  Asp Ser Glu Val Ala  Ala Ile Val Arg Asp  Thr His Thr
    1070                1075                1080

Val Asp  Ser Val Met Ala Gly  Arg Pro Tyr Lys Val  Gly His Phe
    1085                1090                1095

Ala His  Thr Leu Arg Leu His  Leu Met Arg Glu His  Met Gly Val
    1100                1105                1110
```

```
Asp Val  Asp His Val Glu Phe  Ile Glu Arg Lys Ala  Glu Leu Leu
    1115                1120                1125

Glu Lys  His Lys Arg Ala Glu  Ala Thr Gln Asp Glu  Pro Gln Glu
    1130                1135                1140

Leu Pro  Ile Ile Lys Asn Lys  Ser Thr Arg Arg Ala  Arg Gly Arg
    1145                1150                1155

Gly Arg  Arg Ser Gln Arg Gly  Gly Glu Asp Ser Asp  Gly Ala Ser
    1160                1165                1170

Gly Ser  Glu Asp Glu Arg Val  Leu Asp Ile Ser Thr  Ser Ile Phe
    1175                1180                1185

Gly Glu  Thr Val Gln Ala Ala  Lys Ala Ala Asn Gly  Asn Gly Asn
    1190                1195                1200

Gly Asn  Gly Asn Gly Ala Asp  Ala Ser Ala Ala Glu  Ser Val Pro
    1205                1210                1215

Ser Val  Met Ala Pro Thr Arg  Arg Tyr Ser Tyr Gln  Asp Gly Val
    1220                1225                1230

Pro Ala  Ser Pro Ser Ile Asp  Glu Asp Cys Val Val  Glu Gly Thr
    1235                1240                1245

Thr Leu  Asp Asn Ala Asp Gln  Gln Leu Ala Asn Leu  Asp Asp Ala
    1250                1255                1260

Tyr Asp  Ser Asp Ala Glu Val  Leu Leu Arg Asp Tyr  Asp Glu Asp
    1265                1270                1275

Val Asp  Gln Ile Asn Arg Gly  Leu Asp Ile Tyr Thr  Phe Asn Ser
    1280                1285                1290

Leu Ala  Gly Val Asp Asn Lys  Gly Ile Arg Glu Lys  Lys Leu Val
    1295                1300                1305
```

```
Ser Ser  Asp Asn Arg Ile Gln  Gly Asn Glu Lys His  Arg Ala Asp
    1310                1315                1320

Val Glu  Gly Arg Gly Leu Asp  Arg Gly Ala Asn Ser  Lys Glu Ser
    1325                1330                1335

Ala Glu  Asp Phe Arg Ile Arg  Ile Gln Arg His Ala  Gln Phe Ala
    1340                1345                1350

Asn Lys  Leu Thr Glu His Glu  Leu Gln Gln Thr Ser  Glu Thr Val
    1355                1360                1365

Asp Asp  Val Thr Glu Leu Lys  Lys Arg Val Tyr Tyr  Glu Gln Met
    1370                1375                1380

Glu Leu  Lys Gln Arg Tyr His  Asn Ser His Met Glu  Ala Glu Lys
    1385                1390                1395

Ala Phe  Asn Asn Pro Pro Gln  Leu Glu Asp Cys Glu  Asn Glu Lys
    1400                1405                1410

Asn Gly  Val Asn Gly Val Asn  Gly Asn Ser Gly Ser  Thr Asp Ser
    1415                1420                1425

Thr Ala  Pro Val Asn Leu Val  Ser Asp Ala Asn Gly  Asn Ile His
    1430                1435                1440

Pro Glu  His Pro Glu Thr Ser  Asp His Ser Ser Asp  Ser Phe Glu
    1445                1450                1455

Asp Glu  Gly Phe Glu Ile Pro  Lys Ile Pro Pro Pro  Val Asp Pro
    1460                1465                1470

Tyr Cys  Phe Lys Asp Pro Leu  Asp Asp Asp Phe Tyr  Tyr Asp Ile
    1475                1480                1485
```

```
Trp Leu  Ala Thr Ala Glu Lys  Asn Thr Arg Leu Phe  Arg Glu Val
    1490                1495                1500

Phe Arg  Cys Gln Pro Asp Asp  Glu Val Thr Thr Trp  Arg Asp Tyr
    1505                1510                1515

Lys Asp  Phe Thr Thr Tyr Ala  Glu Arg Phe Ser Met  Ser Gln Asp
    1520                1525                1530

Thr Gly  Asp Leu Lys Ser Ser  Tyr Ala Asp Glu His  Gln Ala His
    1535                1540                1545

Ser Pro  Leu Ala Val Lys Ile  Asn Gly Asn Asp Lys  Glu Asp Glu
    1550                1555                1560

Lys Lys  Gln Asn Lys His Lys  Ala Thr His Asp Tyr  Asp Lys Tyr
    1565                1570                1575

Tyr Ala  Arg His Gln His His  Arg Ser Ser Ser Pro  His Tyr Arg
    1580                1585                1590

Arg Ser  His Val Pro Leu Lys  Thr Gly Ala Asn Phe  Thr Asp Arg
    1595                1600                1605

Leu Ala  His Gly Val Asp Asp  Ala Ala Gly His Leu  Lys Tyr Ala
    1610                1615                1620

Leu Gly  Arg Ala Asn Arg Arg  Gly Ser His His Asp  Thr Val Asp
    1625                1630                1635

Val Asp  Asp Ser Pro Pro Arg  Asp Phe Leu Pro Asp  Asp Phe Asn
    1640                1645                1650

Asp Pro  Ile Pro Val Glu Glu  Val Asp Asn Thr Asn  Leu Asn Asn
    1655                1660                1665

Asn Gln  Gln Leu Asn Gln Glu  Leu Tyr Glu Ser Gln  Val Glu Gly
    1670                1675                1680
```

```
Glu Asn  Pro Arg Phe Val Pro  Ser His Ala Ile Asp  Gly Gln His
    1685                 1690                 1695

Pro Arg  Lys Arg Glu Lys Arg  His Ile Gly Glu Val  Asp Gln Pro
    1700                 1705                 1710

Ala Glu  Lys Pro Pro Ile Asn  Glu Ala Tyr Ser Thr  Gly Val Ser
    1715                 1720                 1725

Gly Arg  Leu Gly Gln His Gln  Gln Pro His Leu Lys  Gln Glu Val
    1730                 1735                 1740

Asp Thr  Asp Phe Lys Asp Thr  Ser Ser Ala Ala His  Thr Pro Phe
    1745                 1750                 1755

His His  Thr Pro Asn Ala Arg  Arg Arg Arg Arg Arg  Arg Ala Arg
    1760                 1765                 1770

Gln Arg  Ser Gly Met Asp Lys  Val His Asp Lys Ala  Thr Ala Glu
    1775                 1780                 1785

Lys Ile  Leu Asn Gly Ile Gln  Gly His Leu Val Met  Phe Pro Thr
    1790                 1795                 1800

His Trp  Leu Tyr Lys Glu Leu  Asp Ser Gly Asn Trp  Phe Tyr Asn
    1805                 1810                 1815

Leu Asp  Arg Ile Pro Pro Ile  Glu Ile Tyr Asp
    1820                 1825
```

<210> 74
<211> 3863
<212> DNA
<213> Yarrowia lipolytica

<400> 74
aggtgcccat tcagattcag cccgggtacc agcccttgaa gggtctttac agtgacgtga 60

gtatgataac agaaggaaac gacgaccggt gaggaaagtg atggtatttg tggtgggagc 120

catttgtcat tggcggttgg tcattggtca ttggtcattg gtcattggtg attagcgatt 180

ggtgattagc gattggtgat tggtgattgg tatggtctga ttactgaatg gacttgagac 240

acgaagcaga tggaccatga taaaccatga tacagctctt gacttacttg atttgaacgt 300

gtacttgaga aacagtttta ttgttttcga gctattcaac tgcttggtct gagtctgttt 360

ggtgtgtgag ctcacaccgt ttggtatgtt tggtatcact atgtcaatgc ttctgccaat 420

gtcatgtagt ctccagaatg aggcacttgt tttccttctc tgtcgtttgt ggcacctcct 480

tatgtgtctc ggttcttgtg gcacgttttt ttcttttcgt tctatctcgg ttccctcgtc 540

atctcatgat actcaccgtt gtcaggctgg cagctcgcac ttgcaccatg tttcggtctt 600

tgctttcccg gctacgttcc ctgcccacac atatatctgt ttctacgtgc atgtgattcc 660

cccacatgtg cctgtgatat ctatgtaccg tggctaggat tatgcctgtg attgtgtcat 720

gataaccatt ctgcgcagtg tgattgacga tcagggtctt ggtgggtaca aactagcttt 780

ttgggtttct ttttttcgct ccaaaatgtg agcgtgaaaa tgcacgtgag ctaaaagagt 840

cccaacaccc accaactgcg ctggctcgac tgttgagcat gttgagtgag aggagaggag 900

aaacaaggcg cggaaaaaag ggacgagaaa agaaccatag ggccaaagac aaaaaaccac 960

agacacatgt cacatgcccg gattatgccc cgatcagtgc atgcggaaga ccgacccggc 1020

gatagaagga gaagaaatac gaggtgcaga agtgtcatcc tccgccgtga ttgcattttg 1080

cgtggtaata accacgagcc acgcactttg tccacgcaca tgtctatggt gtaattgtac 1140

ctttcctccc aacgcaccca acccaaacct aaaccctaag ttgcacaaaa cacatgaacg 1200

tgatactgca aacttcccct ctgataaagg cgctgtgtcc atatcatctc agcacaacca 1260

catctccaac accaacacga cttttgccac ctcgggagcc cagagccagt accaccggct 1320

ctccaacaac accaagacct tcactcgaga ctccccctac cattcgcgaa ccccttcagt 1380

ggatgacgac ggaataagac aagcatcgtt tctcaaccga gtctaccagt cctccaagtc 1440

ctacagtccc cgattccgct acggagcaga gatagtcgaa cggtgggcaa gttctgaaaa 1500

caccaacgca gcaacacctt cggcgtcaac cacaaacagc aacaacagca accacagcaa 1560

cacgcctgtt tcaaccacac cctcaagcac aaccatgccc aaggctctgt cttcaaagtc 1620

gctggacgcc gccagcatcc acatggcatc gaacggtgcg cctccactga tacaaaagtc 1680

gtcgttgcac gactttgagc gccttaaatc gggaattgac gacattcaaa actcagacac 1740

caacggtacc cagtacagtg tcgatgtcgg ctcccagggc ctgcgcatgc gtatccagac 1800

ccaagggtac gcccccagcg gcaacagcaa cagaggcagc cccgtgccac cgtcaccggc 1860

gctgtccacg tcgacctcct cgggctccac ttctgctgca gcctccagat ccacctcgtc 1920

tcctgtgcca caacaacctc agacagctgc cacagctgca ggaggacctc gacccggaca 1980

acctcggtcg gcatggcagg aggtgctgat cagcgccacg tcgctggcat cattgtcgca 2040

ggactcgcga aagcgactgc gatactgtct gcatctcctt aagctcgcca acgcccatct 2100

ggcttccacc gtcacaaagc tgcaaggcgc cattgccgag gagactgcat acagtctcgc 2160

gcagtccatt gccgccaacc atgtgcctca caacgagcga caggcgtacc tgcatcagcc 2220

tccctccgct gagcctgctc tctccatcac cgctctcaag gccgatgtcg tgtccacgat 2280

ccgcaagatc atcaaggttg tcagccagta cgccggcaac agtctacccg agccggctcg 2340

gtcgcacatc cggacataca ttcttggcct gccctctcgg tgggcctcta ccaccgcttc 2400

cactaacatc accccaaccc ggtctcctgc cggttcccag tctcctgtgg gctctcctaa 2460

ggagtgtgtc acacctgacc accagggtcc tcctcttccc acaaacacag tgtcgtctcc 2520

agtgcccgag ggcccttctg acgagcagtt gtcgcggcat cgaatcgagg ttgaggctgg 2580

aggcaaggtg ttgattcttg ccaacgaggc tctcgatatg ctgggcaaca tcatcagtat 2640

tgtggatggc accctggagc gtgccgaggg ttggtgcgag ggcatcaacc gggtcaagca 2700

gagagttggg ctcggtgaaa ccgctcctgg cgcagctggg gaacctggag cctccaccga 2760

ggcttctgct gaggcttctg ctgaggcttc tgccagtgcc gccgctgcgt actctgctgt 2820

tgatacctcc acggaccagg acgttgagat gggagatgcc taaggctggc atgacacgac 2880

aagggacaca gatgaaatgc atgatggatg atgacacaac acacacagca tgaagcatgg 2940

catgatggag aacgacatac agcacaacac aacagagaga ggatagcata ggcacaacat 3000

atttcatgta ctcaccacac acagaacaag cgacactgac acatacacac acacacacac 3060

acacgcacac acagcagcag cagcagcagc ggcagcagca gtacagcatt atgcaacccc 3120

tcacagatga tgcaacactt aaatggttac acacatggca tatttaatca cctctcttgt 3180

gggtgctcca gaagatggct tttggcgtcg gatgacggca ctggttttca ttttcctcga 3240

ttacatgtct gtctcgggcg ttgtttgtcc cctcggcact ttttcggcct cggctacccg 3300

gtcctcggga aagtgtgtca aggttcacaa cggtaaaaaa agtcaaaggt ggtggcgttt 3360

tgggaggttc ggtttgaaga gcggagctac tgtgggcagc agcgttaaac ttggtgacaa 3420

gacaacggct ggacgctgac gtttaccatt aattggcgct gctgccgtgg gtggctactg 3480

gatttggtct tttgcacgaa gacttatgcg ttgctgcttg cgagtgtgtc tctgtcttgg 3540

ggtggtttgc cacgtgacgg gccggtgtga cgccgagcgg aattcgtgtt tcttgcccct 3600

tttatttccg acatgttccg gctgctgctt ctccagcatg ttaccccaca accgacgggt 3660

aaaatttctg gacgcgcaca aacctttttt ggcttgcggt ccaggggaaa ggaaccgaca 3720

gatgtctacc accaccatca acaagagcac accaccagta ccaggagtac actgttaccg 3780

ccattagcaa gtagagttta tttatataat gcaaagatta tgatgggttt tgggccgtgg 3840

tgttggtgtt ggtgtctgta gaa 3863

<210> 75
<211> 620
<212> PRT
<213> Yarrowia lipolytica

<400> 75

Met Arg Lys Thr Asp Pro Ala Ile Glu Gly Glu Glu Ile Arg Gly Ala
1 5 10 15

Glu Val Ser Ser Ser Ala Val Ile Ala Phe Cys Val Val Ile Thr Thr
20 25 30

Ser His Ala Leu Cys Pro Arg Thr Cys Leu Trp Cys Asn Cys Thr Phe
35 40 45

Pro Pro Asn Ala Pro Asn Pro Asn Leu Asn Pro Lys Leu His Lys Thr
50 55 60

```
His Glu Arg Asp Thr Ala Asn Phe Pro Ser Asp Lys Gly Ala Val Ser
65                  70                  75                  80

Ile Ser Ser Gln His Asn His Ile Ser Asn Thr Asn Thr Thr Phe Ala
                85                  90                  95

Thr Ser Gly Ala Gln Ser Gln Tyr His Arg Leu Ser Asn Asn Thr Lys
            100                 105                 110

Thr Phe Thr Arg Asp Ser Pro Tyr His Ser Arg Thr Pro Ser Val Asp
        115                 120                 125

Asp Asp Gly Ile Arg Gln Ala Ser Phe Leu Asn Arg Val Tyr Gln Ser
    130                 135                 140

Ser Lys Ser Tyr Ser Pro Arg Phe Arg Tyr Gly Ala Glu Ile Val Glu
145                 150                 155                 160

Arg Trp Ala Ser Ser Glu Asn Thr Asn Ala Ala Thr Pro Ser Ala Ser
                165                 170                 175

Thr Thr Asn Ser Asn Asn Ser Asn His Ser Asn Thr Pro Val Ser Thr
            180                 185                 190

Thr Pro Ser Ser Thr Thr Met Pro Lys Ala Leu Ser Ser Lys Ser Leu
        195                 200                 205

Asp Ala Ala Ser Ile His Met Ala Ser Asn Gly Ala Pro Pro Leu Ile
    210                 215                 220

Gln Lys Ser Ser Leu His Asp Phe Glu Arg Leu Lys Ser Gly Ile Asp
225                 230                 235                 240

Asp Ile Gln Asn Ser Asp Thr Asn Gly Thr Gln Tyr Ser Val Asp Val
                245                 250                 255
```

```
                                                105
Gly Ser Gln Gly Leu Arg Met Arg Ile Gln Thr Gln Gly Tyr Ala Pro
            260                 265                 270

Ser Gly Asn Ser Asn Arg Gly Ser Pro Val Pro Pro Ser Pro Ala Leu
        275                 280                 285

Ser Thr Ser Thr Ser Ser Gly Ser Thr Ser Ala Ala Ala Ser Arg Ser
    290                 295                 300

Thr Ser Ser Pro Val Pro Gln Gln Pro Gln Thr Ala Ala Thr Ala Ala
305                 310                 315                 320

Gly Gly Pro Arg Pro Gly Gln Pro Arg Ser Ala Trp Gln Glu Val Leu
                325                 330                 335

Ile Ser Ala Thr Ser Leu Ala Ser Leu Ser Gln Asp Ser Arg Lys Arg
            340                 345                 350

Leu Arg Tyr Cys Leu His Leu Leu Lys Leu Ala Asn Ala His Leu Ala
        355                 360                 365

Ser Thr Val Thr Lys Leu Gln Gly Ala Ile Ala Glu Glu Thr Ala Tyr
    370                 375                 380

Ser Leu Ala Gln Ser Ile Ala Ala Asn His Val Pro His Asn Glu Arg
385                 390                 395                 400

Gln Ala Tyr Leu His Gln Pro Pro Ser Ala Glu Pro Ala Leu Ser Ile
                405                 410                 415

Thr Ala Leu Lys Ala Asp Val Val Ser Thr Ile Arg Lys Ile Ile Lys
            420                 425                 430

Val Val Ser Gln Tyr Ala Gly Asn Ser Leu Pro Glu Pro Ala Arg Ser
        435                 440                 445

His Ile Arg Thr Tyr Ile Leu Gly Leu Pro Ser Arg Trp Ala Ser Thr
    450                 455                 460
```

```
Thr Ala Ser Thr Asn Ile Thr Pro Thr Arg Ser Pro Ala Gly Ser Gln
465                 470                 475                 480

Ser Pro Val Gly Ser Pro Lys Glu Cys Val Thr Pro Asp His Gln Gly
                485                 490                 495

Pro Pro Leu Pro Thr Asn Thr Val Ser Ser Pro Val Pro Glu Gly Pro
            500                 505                 510

Ser Asp Glu Gln Leu Ser Arg His Arg Ile Glu Val Glu Ala Gly Gly
        515                 520                 525

Lys Val Leu Ile Leu Ala Asn Glu Ala Leu Asp Met Leu Gly Asn Ile
    530                 535                 540

Ile Ser Ile Val Asp Gly Thr Leu Glu Arg Ala Glu Gly Trp Cys Glu
545                 550                 555                 560

Gly Ile Asn Arg Val Lys Gln Arg Val Gly Leu Gly Glu Thr Ala Pro
                565                 570                 575

Gly Ala Ala Gly Glu Pro Gly Ala Ser Thr Glu Ala Ser Ala Glu Ala
            580                 585                 590

Ser Ala Glu Ala Ser Ala Ser Ala Ala Ala Ala Tyr Ser Ala Val Asp
        595                 600                 605

Thr Ser Thr Asp Gln Asp Val Glu Met Gly Asp Ala
    610                 615                 620
```

<210> 76
<211> 4247
<212> DNA
<213> Saccharomyces cerevisiae

<400> 76
tcactcaacc acctccaaaa aataacaggt tcatctaaag taaaagactt taacttgctc 60

ttagtttcca aattaaatat ctgcacgata gtaccatttg ctctaacgga ataaccatc 120

tgagatggat gcatgatagc agaatcaccg cccatattct tccttgtcac ttcattgcct 180

ttggccaaat ccacgattgc aacagagttt gtaccgtcct ttgtttctct aacagtgacg 240

aagtggtcac tctcgaaagt agttgatctg aagtcaagga attgagggga aattcctaag 300

gacatcagat cgaccaattc ggtaaattca atgggtaggt cactcattgg ttagaacttt 360

cgtgataatt tatttttata gttgaatatc ttctttctct ctcaactctg atccggattg 420

tcgaggtttc aataagttac tctgaacaac taatcaaaat atctccttat ttctgtagat 480

tccttcagtt ccacttttta cttttcttaa ttctctttgt atttattcct agcgacgaaa 540

aatgcgagat ctcgaccaaa aaaaggggggt agggtaataa aattaaccct attatttttt 600

aactttaaaa cctataatgt gctaatattt tattataaac ctccttttttt tgcgttcaaa 660

ccctgacaca ttttaagccc tatatttacg gtattagttg attaaactcc gaagcgaaag 720

gaattcggtc attagcggct aatagccgtt ggggtaaatc acctacaagc aagtacacaa 780

gagaacgttg gcgttgttaa gtcaaagcac taatacattg gggctttaag agtgtttata 840

aaggtctaac ctgtaaaaat tatttaaaca acttgaacag gccttaaagt tttcctcatt 900

ccgctcatca tcactaatat tgctctccgt ttttgaatac acacttgaca ctaataagta 960

tcacagaaaa aaagaaaata taataaatta gtattgcgat atgacgagac gtactactat 1020

taatcccgat tcggtggttc tgaatcctca aaaatttatc cagaaagaaa gggcggattc 1080

gaaaatcaaa gttgaccaag ttaacacatt tttagagtca tccccggaga ggagaactct 1140

gacgcacgcc ttaatagacc aaatagtgaa tgatcctata ttgaaaactg atacggacta 1200

ttacgatgct aaaaaaatgc aagagagaga aattactgcc aaaaaaatag ctaggcttgc 1260

tagttatatg gagcacgata tcaaaacagt gcgcaaacac tttcgcgaca ctgacctgat 1320

gaaagagttg caagcaaatg atccagacaa agcttcgcct ttaacaaaca aagacctttt 1380

tatattcgat aagagattgt cacttgtagc aaatattgat cctcaattgg gtacgcgcgt 1440

gggtgtacac ttggggctat ttggtaattg tatcaagggc aatggtactg atgagcaaat 1500

ccggtattgg ttgcaggaga gaggtgccac tttgatgaaa ggtatatatg gctgttttgc 1560

aatgactgag ttaggacatg gttccaatgt tgcccagctg cagactaggg ctgtgtacga 1620

taagcaaaat gatacttttg taattgatac acctgatcta actgccacca aatggtggat 1680

tggtggggct gcccattctg ccacgcacgc tgccgtgtac gccagattga tcgttgaagg 1740

taaagactac ggtgtaaaaa cattcgttgt tcctctgaga gacccttcga ctttccaact 1800

gttagctggt gtttccatag gggatattgg agcgaagatg ggtcgtgacg gtattgataa 1860

tggctggatc cagttcagaa acgtagttat ccctagagaa tttatgctaa gtagatttac 1920

caaagttgtc cgttctccag atggttcagt caccgtcaaa actgagccac aattggatca 1980

aatttctggt tatagtgcat tgttaagtgg tagagttaac atggtcatgg attcatttag 2040

gtttggctcc aaatttgcta ctattgctgt acgttacgcg gttggtcgtc agcaattcgc 2100

acctagaaag ggattgtctg aaacacaatt aatcgactat ccccttcacc aatatcgtgt 2160

tttaccacaa ttgtgtgttc catatttggt gtcacctgta gcttttaagt taatggacaa 2220

ctattattcc actttggacg agttatacaa cgcttcctca tctgcataca aagctgctct 2280

ggttaccgtg agtaaaaagt tgaagaattt atttattgat agcgccagct tgaaagccac 2340

caatacttgg ttaattgcta cactgattga tgagttgaga cagacttgcg gaggacatgg 2400

gtattcacag tataacggat ttggtaaagg ctatgacgac tgggtggttc agtgcacatg 2460

ggagggtgat aataatgttt tatctttaac ttcagcaaaa tcaatattga aaaaatttat 2520

cgattcagcc acaaagggta gatttgacaa cacactggat gtggactcat tctcttactt 2580

aaaacctcag tacataggat ctgtggtttc tggagaaata aagagtggtt taaaggagtt 2640

gggtgattat actgaaattt ggtctatcac cttaatcaaa ttactggcac atattggtac 2700

tttagttgaa aaatcaagaa gtattgatag cgtttctaag cttttagtct tagtatccaa 2760

atttcatgcc ttgcgctgca tgttgaaaac ctattacgac aagttaaact ctcgtgattc 2820

acatatttcc gatgaaatta caaaggaatc tatgtggaat gtttataagt tattttcctt 2880

gtattttatt gacaagcatt ccggagaatt ccaacaattc aagatcttca ctcctgatca 2940

gatctctaaa gttgtgcagc cacaactatt ggctcttttg ccaattgtga ggaaagactg 3000

tataggtctg acagactcct ttgaattacc tgacgcgatg ttaaattctc ctataggtta 3060

ctttgatggc gatatctatc acaattactt caatgaagtt tgccgcaata atccagtgga 3120

ggcagatggg gcagggaagc cttcttatca tgcgctgttg agcagcatgc tcggtagagg 3180

tttcgaattt gaccaaaagt taggtggtgc agctaatgcg gaaattttat cgaaaataaa 3240

caagtgagta gaggtttcct gttttccttc gaaccctctg ttttgcgact tttgtttcaa 3300

ttcaactagt gtcgccaagt tttaacaaaa agttacaaaa tcctagtgag aggccatctt 3360

atgtgcataa cggtactctc tatctattta catatctaat actattcaca taactatgac 3420

gaatcaatga catgactaca tttaccaatg tatagtagta aaaaaattgc catctcctag 3480

aatcgaacca gggtttcatc ggccacaacg atgtgtacta accactatac taagatggca 3540

ggtaattatt gagtgtttac tcaatatctc agaattatat acataacaaa taataaatgt 3600

gctgtgttgg cacttgttgg aatgaagctc taaaacatca tgatgattgt gtggattctg 3660

atgtaattgt tgggattcca tttttaataa ggcaatatgt tggaacgaga gtaattaata 3720

gtgacatgag ttgctatggt aacaatctaa tgcttacatc gtatattaat gtacacatcg 3780

tatacgttta agtgtgattg cgcctattgc agaaggaatg ttaaacgaga agctcagaca 3840

atactgaagc tgtgttaaag acctattagc tgaacatgat atggtaggta catatatgag 3900

gaatatgagt cgtcacatca atgtatagta actaccggaa tcactattat attggtcatg 3960

attaatatga ccaatcggcg tgtgttttat atacctctct tatttagtat aagaagatca 4020

gtactcactt cttcattaat actaattttt aacctctaat tatcaacaca ataatattag 4080

gtatgtagat atactagaag ttctcctcca ggatttagga atccataaaa gggagtctgc 4140

aattctacac aattctataa atattatcat catcgtttta tatgttaata ttcattgatc 4200

ctattacatt atcaatcctt gcgtttcagc ttccactaat ttagatg 4247

<210> 77
<211> 748
<212> PRT
<213> Saccharomyces cerevisiae

<400> 77

Met Thr Arg Arg Thr Thr Ile Asn Pro Asp Ser Val Val Leu Asn Pro
1 5 10 15

```
Gln Lys Phe Ile Gln Lys Glu Arg Ala Asp Ser Lys Ile Lys Val Asp
            20                  25                  30

Gln Val Asn Thr Phe Leu Glu Ser Ser Pro Glu Arg Arg Thr Leu Thr
        35                  40                  45

His Ala Leu Ile Asp Gln Ile Val Asn Asp Pro Ile Leu Lys Thr Asp
    50                  55                  60

Thr Asp Tyr Tyr Asp Ala Lys Lys Met Gln Glu Arg Glu Ile Thr Ala
65                  70                  75                  80

Lys Lys Ile Ala Arg Leu Ala Ser Tyr Met Glu His Asp Ile Lys Thr
                85                  90                  95

Val Arg Lys His Phe Arg Asp Thr Asp Leu Met Lys Glu Leu Gln Ala
            100                 105                 110

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

Phe Asp Lys Arg Leu Ser Leu Val Ala Asn Ile Asp Pro Gln Leu Gly
    130                 135                 140

Thr Arg Val Gly Val His Leu Gly Leu Phe Gly Asn Cys Ile Lys Gly
145                 150                 155                 160

Asn Gly Thr Asp Glu Gln Ile Arg Tyr Trp Leu Gln Glu Arg Gly Ala
                165                 170                 175

Thr Leu Met Lys Gly Ile Tyr Gly Cys Phe Ala Met Thr Glu Leu Gly
            180                 185                 190

His Gly Ser Asn Val Ala Gln Leu Gln Thr Arg Ala Val Tyr Asp Lys
        195                 200                 205
```

```
Gln Asn Asp Thr Phe Val Ile Asp Thr Pro Asp Leu Thr Ala Thr Lys
    210                 215                 220

Trp Trp Ile Gly Gly Ala Ala His Ser Ala Thr His Ala Ala Val Tyr
225                 230                 235                 240

Ala Arg Leu Ile Val Glu Gly Lys Asp Tyr Gly Val Lys Thr Phe Val
                245                 250                 255

Val Pro Leu Arg Asp Pro Ser Thr Phe Gln Leu Leu Ala Gly Val Ser
            260                 265                 270

Ile Gly Asp Ile Gly Ala Lys Met Gly Arg Asp Gly Ile Asp Asn Gly
        275                 280                 285

Trp Ile Gln Phe Arg Asn Val Val Ile Pro Arg Glu Phe Met Leu Ser
    290                 295                 300

Arg Phe Thr Lys Val Val Arg Ser Pro Asp Gly Ser Val Thr Val Lys
305                 310                 315                 320

Thr Glu Pro Gln Leu Asp Gln Ile Ser Gly Tyr Ser Ala Leu Leu Ser
                325                 330                 335

Gly Arg Val Asn Met Val Met Asp Ser Phe Arg Phe Gly Ser Lys Phe
            340                 345                 350

Ala Thr Ile Ala Val Arg Tyr Ala Val Gly Arg Gln Gln Phe Ala Pro
        355                 360                 365

Arg Lys Gly Leu Ser Glu Thr Gln Leu Ile Asp Tyr Pro Leu His Gln
    370                 375                 380

Tyr Arg Val Leu Pro Gln Leu Cys Val Pro Tyr Leu Val Ser Pro Val
385                 390                 395                 400

Ala Phe Lys Leu Met Asp Asn Tyr Tyr Ser Thr Leu Asp Glu Leu Tyr
                405                 410                 415
```

```
Asn Ala Ser Ser Ser Ala Tyr Lys Ala Ala Leu Val Thr Val Ser Lys
            420                 425                 430

Lys Leu Lys Asn Leu Phe Ile Asp Ser Ala Ser Leu Lys Ala Thr Asn
        435                 440                 445

Thr Trp Leu Ile Ala Thr Leu Ile Asp Glu Leu Arg Gln Thr Cys Gly
    450                 455                 460

Gly His Gly Tyr Ser Gln Tyr Asn Gly Phe Gly Lys Gly Tyr Asp Asp
465                 470                 475                 480

Trp Val Val Gln Cys Thr Trp Glu Gly Asp Asn Asn Val Leu Ser Leu
                485                 490                 495

Thr Ser Ala Lys Ser Ile Leu Lys Lys Phe Ile Asp Ser Ala Thr Lys
            500                 505                 510

Gly Arg Phe Asp Asn Thr Leu Asp Val Asp Ser Phe Ser Tyr Leu Lys
        515                 520                 525

Pro Gln Tyr Ile Gly Ser Val Val Ser Gly Glu Ile Lys Ser Gly Leu
    530                 535                 540

Lys Glu Leu Gly Asp Tyr Thr Glu Ile Trp Ser Ile Thr Leu Ile Lys
545                 550                 555                 560

Leu Leu Ala His Ile Gly Thr Leu Val Glu Lys Ser Arg Ser Ile Asp
                565                 570                 575

Ser Val Ser Lys Leu Leu Val Leu Val Ser Lys Phe His Ala Leu Arg
            580                 585                 590

Cys Met Leu Lys Thr Tyr Tyr Asp Lys Leu Asn Ser Arg Asp Ser His
        595                 600                 605
```

Ile Ser Asp Glu Ile Thr Lys Glu Ser Met Trp Asn Val Tyr Lys Leu
610 615 620

Phe Ser Leu Tyr Phe Ile Asp Lys His Ser Gly Glu Phe Gln Gln Phe
625 630 635 640

Lys Ile Phe Thr Pro Asp Gln Ile Ser Lys Val Val Gln Pro Gln Leu
645 650 655

Leu Ala Leu Leu Pro Ile Val Arg Lys Asp Cys Ile Gly Leu Thr Asp
660 665 670

Ser Phe Glu Leu Pro Asp Ala Met Leu Asn Ser Pro Ile Gly Tyr Phe
675 680 685

Asp Gly Asp Ile Tyr His Asn Tyr Phe Asn Glu Val Cys Arg Asn Asn
690 695 700

Pro Val Glu Ala Asp Gly Ala Gly Lys Pro Ser Tyr His Ala Leu Leu
705 710 715 720

Ser Ser Met Leu Gly Arg Gly Phe Glu Phe Asp Gln Lys Leu Gly Gly
725 730 735

Ala Ala Asn Ala Glu Ile Leu Ser Lys Ile Asn Lys
740 745

<210> 78
<211> 973
<212> DNA
<213> Yarrowia lipolytica

<400> 78
aaattgcccc ggagaagacg gccaggccgc ctagatgaca aattcaacaa ctcacagctg 60

actttctgcc attgccacta gggggggccc tttttatatg gccaagccaa gctctccacg 120

tcggttgggc tgcacccaac aataaatggg tagggttgca ccaacaaagg gatgggatgg 180

ggggtagaag atacgaggat aacggggctc aatggcacaa ataagaacga atactgccat 240

taagactcgt gatccagcga ctgacaccat tgcatcatct aagggcctca aaactacctc 300

ggaactgctg cgctgatctg gacaccacag aggttccgag cactttaggt tgcaccaaat 360

gtcccaccag gtgcaggcag aaaacgctgg aacagcgtgt acagtttgtc ttaacaaaaa 420

gtgagggcgc tgaggtcgag cagggtggtg tgacttgtta tagcctttag agctgcgaaa 480

gcgcgtatgg atttggctca tcaggccaga ttgagggtct gtggacacat gtcatgttag 540

tgtacttcaa tcgccccctg gatatagccc cgacaatagg ccgtggcctc atttttttgc 600

cttccgcaca tttccattgc tcggtaccca caccttgctt ctcctgcact tgccaacctt 660

aatactggtt tacattgacc aacatcttac aagcgggggg cttgtctagg gtatatataa 720

acagtggctc tcccaatcgg ttgccagtct ctttttttcct ttctttcccc acagattcga 780

aatctaaact acacatcaca caatgcctgt tactgacgtc cttaagcgaa agtccggtgt 840

catcgtcggc gacgatgtcc gagccgtgag tatccacgac aagatcagtg tcgagacgac 900

gcgttttgtg taatgacaca atccgaaagt cgctagcaac acacactctc tacacaaact 960

aacccagctc tcc 973

<210> 79
<211> 927
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of a modified Y. lipolytica FBAIN gene promoter

<400> 79
aaattgcccc ggagaagacg gccaggccgc ctagatgaca aattcaacaa ctcacagctg 60

actttctgcc attgccacta gggggggggcc tttttatatg gccaagccaa gctctccacg 120

tcggttgggc tgcacccaac aataaatggg tagggttgca ccaacaaagg gatgggatgg 180

ggggtagaag atacgaggat aacggggctc aatggcacaa ataagaacga atactgccat 240

taagactcgt gatccagcga ctgacaccat tgcatcatct aagggcctca aaactacctc 300

ggaactgctg cgctgatctg gacaccacag aggttccgag cactttaggt tgcaccaaat 360

gtcccaccag gtgcaggcag aaaacgctgg aacagcgtgt acagtttgtc ttaacaaaaa 420

gtgagggcgc tgaggtcgag cagggtggtg tgacttgtta tagcctttag agctgcgaaa 480

gcgcgtatgg atttggctca tcaggccaga ttgagggtct gtggacacat gtcatgttag 540

tgtacttcaa tcgccccctg gatatagccc cgacaatagg ccgtggcctc atttttttgc 600

cttccgcaca tttccattgc tcggtaccca caccttgctt ctcctgcact tgccaacctt 660

aatactggtt tacattgacc aacatcttac aagcgggggg cttgtctagg gtatatataa 720

acagtggctc tcccaatcgg ttgccagtct ctttttttcct ttctttcccc acagattcga 780

aatctaaact acacatcaca gaattccgag ccgtgagtat ccacgacaag atcagtgtcg 840

agacgacgcg ttttgtgtaa tgacacaatc cgaaagtcgc tagcaacaca cactctctac 900

acaaactaac ccagctctca cacaatg 927

<210> 80
<211> 971
<212> DNA
<213> Yarrowia lipolytica

<400> 80
gacgcagtag gatgtcctgc acgggtcttt ttgtggggtg tggagaaagg ggtgcttgga 60

gatggaagcc ggtagaaccg ggctgcttgt gcttggagat ggaagccggt agaaccgggc 120

tgcttggggg gatttggggc cgctgggctc caaagagggg taggcatttc gttggggtta 180

cgtaattgcg gcatttgggt cctgcgcgca tgtcccattg gtcagaatta gtccggatag 240

gagacttatc agccaatcac agcgccggat ccacctgtag gttgggttgg gtgggagcac 300

ccctccacag agtagagtca aacagcagca gcaacatgat agttgggggt gtgcgtgtta 360

aaggaaaaaa aagaagcttg ggttatattc ccgctctatt tagaggttgc gggatagacg 420

ccgacggagg gcaatggcgc catggaacct tgcggatatc gatacgccgc ggcggactgc 480

gtccgaacca gctccagcag cgtttttttcc gggccattga gccgactgcg accccgccaa 540

cgtgtcttgg cccacgcact catgtcatgt tggtgttggg aggccacttt ttaagtagca 600

caaggcacct agctcgcagc aaggtgtccg aaccaaagaa gcggctgcag tggtgcaaac 660

ggggcggaaa cggcgggaaa aagccacggg ggcacgaatt gaggcacgcc ctcgaatttg 720

agacgagtca cggccccatt cgcccgcgca atggctcgcc aacgcccggt cttttgcacc 780

acatcaggtt accccaagcc aaacctttgt gttaaaaagc ttaacatatt ataccgaacg 840

taggtttggg cgggcttgct ccgtctgtcc aaggcaacat ttatataagg gtctgcatcg 900

ccggctcaat tgaatctttt ttcttcttct cttctctata ttcattcttg aattaaacac 960

acatcaacat g 971

<210> 81
<211> 535
<212> PRT
<213> Macadamia tetraphylla

<400> 81

Met Ser Ile Ser Asn Ser Pro Glu Asn Leu Asp Gln Lys Pro Asn Ile
1 5 10 15

Arg Asn Glu Ser Glu Asp Arg Ser Phe Ser Val Arg Arg Arg Thr Asn
20 25 30

Ser Asn Val Val Asn Ala Val Pro Asp Ser Cys Ser Lys Val Gly Thr
35 40 45

Ile Glu Thr Val Thr Lys Phe Ser Ser Val Asp Ser Ala Ala Glu Met
50 55 60

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

Glu Glu Ala Asn Ser Leu Ser Asn Gly Gln Val Tyr Asn Glu Arg Val
85 90 95

Gly Ile Gly Glu Gly Arg Ser Pro Glu Ala Glu Val Pro Phe Arg Phe
100 105 110

Ser Tyr Arg Ala Ser Ala Pro Ala His Arg Arg Val Lys Glu Ser Pro
115 120 125

```
Leu Ser Ser Asp Ala Ile Phe Arg Gln Ser His Ala Gly Leu Phe Asn
    130                 135                 140

Leu Cys Ile Leu Val Leu Val Ala Val Asn Gly Arg Leu Ile Ile Glu
145                 150                 155                 160

Asn Leu Met Lys Tyr Gly Leu Leu Ile Arg Ala Gly Phe Trp Phe Ser
                165                 170                 175

Ser Lys Ser Leu Arg Asp Trp Pro Leu Leu Met Cys Cys Leu Ser Leu
            180                 185                 190

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

Lys Phe Ile Pro Glu Pro Val Val Met Ser Leu His Val Leu Ile Cys
    210                 215                 220

Thr Ala Ser Leu Leu Tyr Pro Ile Phe Val Ile Leu Arg Cys Gly Ser
225                 230                 235                 240

Ala Val Leu Phe Gly Ile Ala Leu Met Leu Leu Ala Ser Ile Val Trp
                245                 250                 255

Leu Lys Leu Val Ser Tyr Ala His Thr Asn Cys Asp Met Arg Ala Leu
            260                 265                 270

Ala Lys Leu Thr Asp Lys Val Ala Ser Lys Ser Ser Asp Met Asp Tyr
        275                 280                 285

Ser Tyr Asp Val Ser Phe Glu Ser Leu Ala Tyr Phe Met Val Ala Pro
    290                 295                 300

Thr Leu Cys Tyr Gln Ala Ser Tyr Pro Arg Ser Thr Cys Ile Arg Lys
305                 310                 315                 320

Gly Trp Val Ile Arg Gln Leu Val Lys Leu Val Ile Phe Thr Gly Val
                325                 330                 335
```

```
Met Gly Phe Ile Val Glu Gln Tyr Ile Asn Pro Ile Val Gln His Ser
            340                 345                 350

Gln His Pro Leu Lys Gly Asn Leu Leu Tyr Ala Ile Glu Arg Val Leu
        355                 360                 365

Lys Leu Ser Val Pro Thr Leu Tyr Val Trp Leu Cys Met Phe Tyr Cys
    370                 375                 380

Phe Phe His Leu Trp Leu Asn Ile Leu Ala Glu Leu Leu Arg Phe Gly
385                 390                 395                 400

Asp Arg Glu Phe Tyr Lys Asp Trp Trp Asn Ala Lys Thr Val Glu Glu
                405                 410                 415

Tyr Trp Arg Met Trp Asn Met Pro Val His Lys Trp Met Val Arg His
            420                 425                 430

Ile Tyr Phe Pro Cys Leu Arg Asn Gly Leu Pro Arg Glu Val Thr Leu
        435                 440                 445

Val Ile Val Phe Phe Leu Ser Ala Ile Phe His Glu Leu Cys Ile Gly
    450                 455                 460

Val Pro Cys His Ile Phe Lys Phe Trp Ala Phe Ile Gly Ile Met Phe
465                 470                 475                 480

Gln Val Pro Leu Val Leu Ile Thr Asn Tyr Leu Gln Asn Lys Phe Arg
                485                 490                 495

Ser Ser Met Val Gly Asn Met Val Phe Trp Phe Phe Phe Ser Ile Leu
            500                 505                 510

Gly Gln Pro Met Ser Val Leu Leu Tyr Tyr His Asp Val Met Asn Arg
        515                 520                 525
```

Lys Glu Asn Gly Val Ser Lys
530 535

<210> 82
<211> 1634
<212> DNA
<213> Macadamia tetraphylla

<400> 82
ggtctctcca ccacaatgtc tatctctaac tctcccgaga acctggacca gaagcctaac 60

atccgaaacg agtctgagga ccgatctttc tctgtgcgac gacgaactaa ctctaacgtg 120

gtgaacgctg tgcccgactc ttgctctaag gtgggcacca tcgagactgt gaccaagttc 180

tcttctgtgg actctgccgc cgagatgtcc tctggcgagg accgaaaccg aatcggcaac 240

ggtgacgacc gaaaggaaga ggccaactct ctgtctaacg gccaggtgta caacgagcga 300

gtcggcatcg gcgagggacg atctcccgag gccgaggtgc ccttccgatt ctcttaccga 360

gcctctgctc ccgctcaccg acgagtgaag gaatctcccc tgtcctctga cgccatcttc 420

cgacagtctc acgccggcct gttcaacctg tgcatcctgg tgctggtggc cgtgaacggc 480

cgactgatca ttgagaacct gatgaagtac ggcctgctga tccgagccgg attctggttc 540

tcttcgaagt ctctgcgaga ctggcccctg ctgatgtgct gcctgtctct gcctatcttc 600

cctctgggcg ccttcaccgt cgagaagtct gccaagcaca agttcatccc cgagcctgtg 660

gtgatgtctc tgcacgtgct gatctgtacc gcctctctgc tgtaccccat cttcgtgatc 720

ctccgatgcg gatctgccgt gctgttcgga atcgccctga tgctgctggc ctctatcgtg 780

tggctgaagc tggtgtctta cgcccacacc aactgcgaca tgcgagccct ggccaagctg 840

accgacaagg tggcctctaa gtcctccgac atggactact cttacgacgt gtctttcgag 900

tctctggcct acttcatggt ggctcccact ctgtgctacc aggcttctta ccctcgatct 960

acctgcatcc gaaagggctg ggtgatccga cagctggtca agctggtgat tttcaccggc 1020

gtgatgggct tcatcgtcga gcagtacatc aaccccatcg tgcagcactc tcagcaccct 1080

ctgaagggca acctgctgta cgctatcgag cgagtgctga agctctctgt ccctactctg 1140

tacgtctggc tgtgcatgtt ctactgtttc ttccacctgt ggctcaacat cctggccgag 1200

ctgctgcgat tcggcgaccg agagttctac aaggactggt ggaacgccaa gaccgtggaa 1260

gagtactggc gaatgtggaa catgcccgtg cacaagtgga tggtgcgaca catctacttc 1320

ccttgcctgc gaaacggact gccccgagag gtgaccctgg tgatcgtgtt cttcctgtct 1380

gccattttcc acgagctgtg catcggcgtg ccctgccaca tcttcaagtt ctgggccttc 1440

atcggcatca tgttccaggt gcctctggtg ctgatcacca actacctgca gaacaagttc 1500

cgatcttcta tggtgggcaa catggtgttc tggtttttct tctcgatcct gggacagccc 1560

atgtctgtgc tgctgtacta ccacgacgtg atgaacagaa aggaaaacgg cgtttctaag 1620

taggtataga gacc 1634

<210> 83
<211> 352
<212> PRT
<213> Garcinia mangostana

<400> 83

Met Leu Lys Leu Ser Ser Ser Arg Ser Pro Leu Ala Arg Ile Pro Thr
1 5 10 15

Arg Pro Arg Pro Asn Ser Ile Pro Pro Arg Ile Ile Val Val Ser Ser
20 25 30

Ser Ser Ser Lys Val Asn Pro Leu Lys Thr Glu Ala Val Val Ser Ser
35 40 45

Gly Leu Ala Asp Arg Leu Arg Leu Gly Ser Leu Thr Glu Asp Gly Leu
50 55 60

Ser Tyr Lys Glu Lys Phe Ile Val Arg Cys Tyr Glu Val Gly Ile Asn
65 70 75 80

Lys Thr Ala Thr Val Glu Thr Ile Ala Asn Leu Leu Gln Glu Val Gly
85 90 95

Cys Asn His Ala Gln Ser Val Gly Tyr Ser Thr Gly Gly Phe Ser Thr
100 105 110

```
Thr Pro Thr Met Arg Lys Leu Arg Leu Ile Trp Val Thr Ala Arg Met
        115                 120                 125

His Ile Glu Ile Tyr Lys Tyr Pro Ala Trp Ser Asp Val Val Glu Ile
    130                 135                 140

Glu Ser Trp Gly Gln Gly Glu Gly Lys Ile Gly Thr Arg Arg Asp Trp
145                 150                 155                 160

Ile Leu Arg Asp Tyr Ala Thr Gly Gln Val Ile Gly Arg Ala Thr Ser
                165                 170                 175

Lys Trp Val Met Met Asn Gln Asp Thr Arg Arg Leu Gln Lys Val Asp
            180                 185                 190

Val Asp Val Arg Asp Glu Tyr Leu Val His Cys Pro Arg Glu Leu Arg
        195                 200                 205

Leu Ala Phe Pro Glu Glu Asn Asn Ser Ser Leu Lys Lys Ile Ser Lys
    210                 215                 220

Leu Glu Asp Pro Ser Gln Tyr Ser Lys Leu Gly Leu Val Pro Arg Arg
225                 230                 235                 240

Ala Asp Leu Asp Met Asn Gln His Val Asn Asn Val Thr Tyr Ile Gly
                245                 250                 255

Trp Val Leu Glu Ser Met Pro Gln Glu Ile Ile Asp Thr His Glu Leu
            260                 265                 270

Gln Thr Ile Thr Leu Asp Tyr Arg Arg Glu Cys Gln His Asp Asp Val
        275                 280                 285

Val Asp Ser Leu Thr Ser Pro Glu Pro Ser Glu Asp Ala Glu Ala Val
    290                 295                 300
```

Phe Asn His Asn Gly Thr Asn Gly Ser Ala Asn Val Ser Ala Asn Asp
305 310 315 320

His Gly Cys Arg Asn Phe Leu His Leu Leu Arg Leu Ser Gly Asn Gly
325 330 335

Leu Glu Ile Asn Arg Gly Arg Thr Glu Trp Arg Lys Lys Pro Thr Arg
340 345 350

<210> 84
<211> 1081
<212> DNA
<213> Garcinia mangostana

<400> 84
ggtctctaat gatgctgaaa ttatcttctt cacgttcgcc gttggcaaga attccgaccc 60

gacctcgacc taattccatc ccgcccagaa tcattgttgt tagcagttct tctagcaagg 120

ttaatccttt gaaaactgaa gctgttgtca gctcgggatt ggcggataga ttgcgactgg 180

gttcgttgac cgaagatggt ttatcctaca aagagaaatt tattgtgcga tgctatgagg 240

tgggcatcaa taaaacagcc accgtggaga ctattgcaaa cttgctccaa gaagtcggtt 300

gtaaccacgc ccaatccgtg ggctattcca ctggcggctt ttctactacc ccgaccatgc 360

gaaaattgcg gctgatttgg gtcactgcca gaatgcatat cgagatatat aaatatcccg 420

cctggtctga tgtcgttgag attgaatcgt ggggacaagg tgaaggcaag attggtacac 480

gccgagactg gatcttacgt gattatgcta ccggtcaagt tatcggacga gcgacatcga 540

agtgggttat gatgaatcaa gatacccgac gacttcaaaa ggtcgatgtg gatgtgcggg 600

atgagtacct cgtgcattgt ccgcgtgagt tgcgtttggc gttccccgag gagaataatt 660

ctagcttgaa gaagatctcc aagttggaag atccttccca atattcgaaa ctgggcttgg 720

tgccacgacg tgccgacttg gatatgaacc aacacgtcaa taatgttact tatattggat 780

gggtcctgga atctatgcct caagagatca tagacactca tgaattgcag acaattacct 840

tggattatcg acgcgagtgc caacatgatg atgttgtgga tagcttgacc tctcccgaac 900

cctctgagga tgccgaagca gtgttcaatc ataacggaac aaacggctcc gctaatgtct 960

ctgctaatga ccatggttgc agaaatttct tgcatctgtt gcgattaagc ggcaatggtt 1020

tggaaattaa tcggggacga accgaatgga gaaagaaacc cacacggtag tctaagagac 1080

c 1081

<210> 85
<211> 355
<212> PRT
<213> Garcinia mangostana

<400> 85

Met Phe Lys Ile Ser Ser Ser Leu Ser Pro Val Asp Gln Ile Pro Pro
1 5 10 15

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

Leu Ala Val Ser Ser Ser Ser Gly Lys Ile Val Asn Asn Pro Leu Lys
35 40 45

Ala Glu Thr Thr Glu Ala Val Ser Gly Glu Leu Ala Arg Arg Phe Arg
50 55 60

Leu Gly Arg Leu Ala Glu Asp Gly Phe Ser Tyr Lys Glu Lys Phe Ile
65 70 75 80

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

Leu Ala Asn Leu Leu Gln Glu Val Gly Gly Asn His Ala Gln Ser Val
100 105 110

Gly Phe Ser Thr Asp Gly Phe Ala Thr Thr His Ser Met Arg Lys Met
115 120 125

His Leu Ile Trp Val Thr Ala Arg Met His Ile Glu Ile Tyr Lys Tyr
130 135 140

```
Pro Ala Trp Ser Asp Val Ile Glu Val Glu Thr Trp Ile Gly Ala Glu
145                 150                 155                 160

Gly Arg Ile Gly Thr Arg Arg Asn Trp Ile Ile Lys Asp Cys Ala Thr
                165                 170                 175

Asp Glu Val Ile Gly Arg Ala Thr Ser Lys Trp Val Met Met Asn Gln
            180                 185                 190

Asp Thr Arg Arg Leu Glu Lys Val Ser Asp Asp Val Arg Glu Glu His
        195                 200                 205

Leu Val Phe Ser Pro Arg Glu Pro Arg Leu Pro Phe Pro Asp Glu Asn
    210                 215                 220

Asn Ser Ser Leu Lys Lys Ile Ser Lys Leu Asp Asp Pro Ala Gln Tyr
225                 230                 235                 240

Ser Lys Leu Ser Leu Glu Pro Arg Arg Gly Asp Leu Asp Met Asn Gln
                245                 250                 255

His Val Asn Asn Val Thr Tyr Ile Gly Trp Val Leu Glu Ser Met Pro
            260                 265                 270

Gln Glu Ile Ile Asp Thr His Glu Leu Gln Thr Ile Thr Leu Asp Tyr
        275                 280                 285

Arg Arg Glu Cys Gln His Asp Asp Leu Val Asp Ser Leu Thr Ser Pro
    290                 295                 300

Glu Pro Ser Glu Phe Ser Glu Thr Thr Asn Gly Ser Ala Asn Val Ser
305                 310                 315                 320

Pro Asn Asp Asn Arg Cys Leu Asn Phe Leu His Leu Leu Arg Leu Ser
                325                 330                 335

Ser Asp Gly Ser Glu Ile Asn Arg Gly Arg Thr Val Trp Arg Lys Lys
            340                 345                 350
```

Pro Ala Lys
355

<210> 86
<211> 1090
<212> DNA
<213> Garcinia mangostana

<400> 86
ggtctctaat gatgttcaaa atttcgtcgt cactctcacc tgttgaccaa attcccccaa 60

ttagcccttt gcccagacca cgaccccgtc caattacacc aagagttctg gcagtgtcca 120

gctccagcgg caaaatcgtt aataatccct tgaaagccga aaccaccgaa gcggtcagtg 180

gcgaactggc ccgcagattt agattgggtc gactggctga agatggattt agctataaag 240

agaaattcat cgtcagatgt tatgaagttg gtattaataa aactgcgaca gtcgaaacct 300

tggctaattt gctgcaagaa gtgggtggaa atcatgccca atcggttggc tttagcactg 360

atggcttcgc taccactcat tccatgcgga aaatgcattt gatttgggtg accgctcgga 420

tgcatataga aatctataaa tatcctgcct ggtcggacgt cattgaagtc gaaacttgga 480

tcggagccga aggtcgtatt ggtaccagac gaaattggat tattaaagac tgcgcgacag 540

atgaagttat tggtcgggct acaagcaaat gggtcatgat gaatcaagac acacggcgat 600

tggaaaaggt ttcggatgat gttcgtgaag aacatttggt cttttccccc agagaaccgc 660

gcctgccgtt ccccgacgaa aataactctt cattgaagaa aatctcgaaa ttagatgatc 720

cggctcaata ttccaaattg agtttggaac cacgacgggg agacttagat atgaatcaac 780

atgttaacaa tgttacttat ataggctggg ttctcgagag tatgccacaa gagatcattg 840

acacacatga attgcagaca ataactttgg actaccggcg agaatgtcaa cacgatgatt 900

tagtggactc gctgacctct cctgagcctt ctgaattttc tgaaactact aatggtagtg 960

ccaatgtcag ccccaatgat aaccgctgtt taaacttctt gcatttgctg agacttagct 1020

ccgacggctc tgaaattaac cggggtcgta ctgtgtggag aaagaaaccc gcaaaataat 1080

ctaagagacc 1090

<210> 87
<211> 545
<212> PRT
<213> Pseudomonas chlororaphis

<400> 87

```
Met Arg Asp Tyr Glu His Val Val Glu Ser Phe Asp Tyr Leu Gln Ser
1               5                   10                  15


Ala Thr Gln Asp Leu His Gly Glu Leu Thr Ala Leu Asn Ala Cys Val
            20                  25                  30


Glu Cys Cys Asp Arg His Ala His Gly Glu Ala Val Ala Leu Tyr Cys
        35                  40                  45


Glu Ala Gln Asp Gly His Ala Glu Arg Tyr Arg Phe Arg Asp Leu Gln
    50                  55                  60


Arg Gln Ala Ala Arg Phe Gly Asn Phe Leu Arg Glu Gln Gly Val Lys
65                  70                  75                  80


Pro Gly Asp Arg Val Ala Gly Leu Met Pro Arg Thr Val Glu Leu Leu
                85                  90                  95


Ile Ala Ile Leu Gly Thr Trp Arg Ile Gly Ala Val Tyr Gln Pro Leu
            100                 105                 110


Phe Thr Ala Phe Gly Pro Lys Ala Ile Glu Gln Arg Leu Asn Cys Ser
        115                 120                 125


Asn Ala Arg Trp Ile Val Thr Asp Pro His Asn Arg Pro Lys Leu Asp
    130                 135                 140


Asp Val Thr Asp Cys Pro Ser Ile Val Val Thr Gly Gly Ala Pro Gln
145                 150                 155                 160


Asn Pro Ala Asp His His Phe Trp Ser Ala Leu Asn Arg Gln Ala Asp
                165                 170                 175
```

```
Asp Cys Ala Pro Val Leu Leu Asp Ala Ser Ala Pro Phe Leu Leu Met
            180                 185                 190

Cys Thr Ser Gly Thr Thr Gly Pro Ala Lys Pro Leu Glu Val Pro Leu
        195                 200                 205

Ser Ala Ile Leu Ala Phe Lys Gly Tyr Met Arg Asp Ala Ile Asp Leu
    210                 215                 220

Arg Ala Asp Asp Arg Phe Trp Asn Leu Ala Asp Pro Gly Trp Ala Tyr
225                 230                 235                 240

Gly Leu Tyr Tyr Ala Val Thr Gly Pro Leu Ala Cys Gly Tyr Ala Thr
                245                 250                 255

Leu Phe Tyr Asp Gly Pro Phe Thr Val Glu Ser Thr Arg His Ile Ile
            260                 265                 270

Ala Lys Tyr Ala Ile Asn Asn Leu Ala Gly Ser Pro Thr Ala Tyr Arg
        275                 280                 285

Phe Leu Ile Ala Ala Gly Ala Glu Phe Ala Asp Ala Val Arg Gly Arg
    290                 295                 300

Leu Arg Ala Val Ser Ser Ala Gly Glu Pro Leu Asn Pro Gln Val Val
305                 310                 315                 320

Arg Trp Phe Ala Glu Gln Leu Gly Val Val Ile His Asp His Tyr Gly
                325                 330                 335

Gln Thr Glu Ile Gly Met Val Leu Cys Asn His His Gly Leu Arg His
            340                 345                 350

Pro Val Arg Glu Gly Ser Ala Gly Tyr Ala Val Pro Gly Tyr Arg Ile
        355                 360                 365
```

```
Val Val Leu Asp Lys Ala His Arg Glu Leu Pro Ala Gly Gln Pro Gly
    370                 375                 380

Val Leu Ala Val Asp Arg Glu Arg Ser Pro Leu Cys Trp Phe Asp Gly
385                 390                 395                 400

Tyr Leu Gly Met Pro Thr Gln Ala Phe Ala Gly Arg Tyr Tyr Leu Ser
                405                 410                 415

Gly Asp Ile Val Glu Leu Asn Asp Asp Gly Ser Ile Ser Phe Val Gly
            420                 425                 430

Arg Asn Asp Asp Leu Ile Thr Thr Ser Gly Tyr Arg Val Gly Pro Phe
        435                 440                 445

Asp Val Glu Ser Ala Leu Ile Glu His Pro Ala Val Val Glu Ala Ala
    450                 455                 460

Val Ile Gly Lys Pro Asp Pro Gln Arg Thr Glu Leu Ile Lys Ala Phe
465                 470                 475                 480

Val Val Leu Asn Thr Pro Tyr Leu Pro Ser Pro Glu Leu Ala Glu Glu
                485                 490                 495

Leu Arg Leu His Val Arg Gln Arg Leu Ala Ala His Ala Tyr Pro Arg
            500                 505                 510

Glu Met Glu Phe Val Asp His Leu Pro Lys Thr Pro Ser Gly Lys Leu
        515                 520                 525

Gln Arg Phe Ile Leu Arg Asn Gln Glu Ile Ala Lys Gln Gln Ala Leu
    530                 535                 540

Gly
545
```

<210> 88
<211> 1660

<212> DNA
<213> Pseudomonas chlororaphis

<400> 88
ggtctctaca aatgcgagac tacgagcatg ttgtggaaag ttttgactac ttacaatccg 60

caactcaaga ccttcacgga gagttgaccg ccctgaatgc ttgtgttgag tgttgtgata 120

gacacgcgca tggtgaagct gttgcccttt actgtgaggc ccaagacggt cacgcggaac 180

gatatcggtt ccgagatttg cagcgtcaag ccgcacgatt tggtaacttt cttcgagagc 240

agggagtgaa acctggtgac cgagtggctg gtttgatgcc ccgtaccgtc gagttgctta 300

tcgctatcct gggaacatgg cggattggag ctgtgtatca gcccttgttc accgcgtttg 360

gtccgaaagc tattgagcag cggctgaatt gtagcaatgc tcgatggatt gtcactgatc 420

cccataatcg accaaaattg gatgatgtta ccgattgtcc gagtatagtc gtcactggcg 480

gtgcacccca aaatccggca gaccatcact tttggtccgc cttgaatcga caagccgatg 540

actgtgctcc tgtcttatta gacgcttccg ctcccttctt gctcatgtgc acttcgggta 600

ctaccggccc tgcgaagcct ttggaagtcc ccctctcggc catcttagct tttaaaggtt 660

acatgcgaga tgccatcgat cttagagcag atgatcgatt ttggaatttg gccgatccag 720

gatgggcgta cggactgtac tatgccgtca ctggaccctt ggcctgcggc tatgccacct 780

tgttttatga tggaccgttt accgttgagt ccactcgaca tataatcgct aaatatgcca 840

taaataattt ggctggatca ccgacagctt atcgattctt gattgctgcc ggcgctgaat 900

ttgctgatgc agtgcgcggt cgattgcgtg ctgtgtctag tgcgggcgag cccttgaatc 960

cccaggtggt tcgatggttc gctgagcaac tgggagtcgt catccatgat cattatggtc 1020

agacagagat aggtatggtc ctttgtaacc atcatggact tcgacatccc gtgcgtgagg 1080

gttcggcagg atacgcagtt cctggttatc gtattgtggt cctcgacaaa gctcatcgag 1140

agctgccggc aggtcaacct ggagttttgg cagtcgaccg agaacgatcc cctctgtgtt 1200

ggtttgatgg ttatttggga atgcctacac aagcgttcgc cggacgatac tatttgtccg 1260

gagacatcgt cgaactgaac gatgatggtt ctatttcttt tgttggtcga aacgacgatt 1320

tgatcactac ctctggatat cgagtgggtc catttgacgt cgagtccgcc ttgattgagc 1380

atcccgctgt cgttgaagct gcagttatag gaaagccaga ccctcagcgg acagaactga 1440

ttaaagcctt cgttgtcttg aatactccat atctcccttc gccagaatta gccgaagagt 1500

tgcgtttgca cgtgcgacaa cggttggctg cccacgcata cccccgagag atggaattcg 1560

ttgatcactt accgaaaaca ccttcgggaa aactgcagcg atttatcttg agaaatcaag 1620

aaatcgcgaa acagcaagcc ttgggatagt ctaagagacc 1660

<210> 89
<211> 1660
<212> DNA
<213> Pseudomonas chlororaphis

<400> 89
ggtctctaca aatgcgtgac tacgaacatg ttgtggagtc ttttgattac ttgcaatctg 60

caacacaaga tctccacggt gagttgaccg cactgaatgc ttgtgtggag tgttgtgatc 120

gacacgcgca tggagaagct gttgccctct actgtgaggc ccaagacgga cacgcggaac 180

gttatagatt ccgagatctc cagcgacaag ctgctcgttt tggtaacttt ctcagagagc 240

agggtgttaa acccggcgac cgagttgcag gattaatgcc ccgaaccgtg gagttgctta 300

tcgctatcct gggaacttgg agaattggag ccgtttatca gcctttattc acagcttttg 360

gaccaaaagc tattgagcag cgactgaatt gttctaatgc tcgatggatt gtcacagatc 420

ctcataatcg acctaaattg gatgatgtta cagattgtcc ttcgattgtc gtcacaggcg 480

gagctcctca aaatccagcc gaccatcatt tctggtcagc tctcaatcga caagccgatg 540

actgtgcccc tgttctcttg gacgcttcgg ctccctttct gctgatgtgc acttccggta 600

caaccggtcc cgctaagccc ttagaagtcc ctttgtctgc catattagca tttaaaggat 660

acatgcgcga cgctatagat ctgcgagccg atgatcgatt ttggaattta gccgatcccg 720

gctgggcgta cggcttatac tatgctgtca ctggaccctt ggcatgcggc tatgctacac 780

ttttctatga tggtccgttt acagtcgaga gcacacgcca tattatcgcg aaatatgcta 840

taaataattt ggctggaagc ccaactgctt atcgatttct cattgctgcc ggtgctgaat 900

ttgctgatgc tgtccggggc cgtttgcgag ctgtgtcctc tgctggcgag cctttgaatc 960

ctcaggtggt cagatggttc gctgagcaat tgggagttgt catacatgat cattatggtc 1020

agacagagat tggtatggtt ttgtgtaacc atcatggact gcgacatccc gtgcgtgagg 1080

gttctgccgg atacgctgtc cctggttata gaattgtggt cctcgacaaa gcccatcgtg 1140

agctccccgc tggtcaaccc ggagtgttgg cagtggaccg tgaacggtcc cccttatgtt 1200

ggtttgatgg atatttgggt atgccaactc aagctttcgc tggacgttac tatttgtctg 1260

gagacatcgt cgaattgaac gatgatggtt ctatttcgtt tgttggacgg aacgacgatc 1320

tcatcacaac ctcaggatat agagtgggcc catttgacgt cgagtccgcc ttaatagagc 1380

atccggctgt cgtggaagcc gctgtcattg gaaagccaga cccccagaga actgaactga 1440

tcaaagcctt cgtggtcttg aataccccctt atttaccttc tccggaatta gccgaagagt 1500

tgcgattgca cgtgagacaa cggttagcgg ctcacgcata cccccgcgag atggaattcg 1560

tggatcactt gccaaagaca cccagcggta aactgcagag atttatcttg cggaatcaag 1620

aaatcgcaaa acagcaagct ttaggctagt ctaagagacc 1660

<210> 90
<211> 314
<212> PRT
<213> Mortierella alpina

<400> 90

Met Ser Ile Gly Ser Ser Asn Pro Val Leu Leu Ala Ala Ile Pro Phe
1 5 10 15

Val Tyr Leu Phe Val Leu Pro Arg Val Leu Ala Phe Leu Pro Gln Lys
20 25 30

Ala Gln Phe Leu Ala Lys Cys Ile Val Val Leu Ile Ala Thr Leu Ile
35 40 45

Met Ser Val Ala Gly Cys Phe Ile Ser Ile Val Cys Ala Leu Leu Asp
50 55 60

Lys Arg Tyr Val Ile Asn Tyr Val Val Ser Arg Leu Phe Ser Phe Leu
65 70 75 80

```
Ala Ala Arg Pro Cys Gly Val Thr Tyr Lys Ile Val Gly Glu Glu His
                85                  90                  95

Leu Asp Lys Tyr Pro Ala Ile Val Val Cys Asn His Gln Ser Ser Met
            100                 105                 110

Asp Met Met Val Leu Gly Arg Val Phe Pro Lys His Cys Val Val Met
        115                 120                 125

Ala Lys Lys Glu Leu Leu Tyr Phe Pro Phe Leu Gly Met Phe Met Lys
    130                 135                 140

Leu Ser Asn Ala Ile Phe Ile Asp Arg Lys Asn His Lys Lys Ala Ile
145                 150                 155                 160

Glu Ser Thr Thr Gln Ala Val Ala Asp Met Lys Lys His Asn Ser Gly
                165                 170                 175

Ile Trp Ile Phe Pro Glu Gly Thr Arg Ser Arg Leu Asp Lys Ala Asp
            180                 185                 190

Leu Leu Pro Phe Lys Lys Gly Ala Phe His Leu Ala Ile Gln Ala Gln
        195                 200                 205

Leu Pro Ile Leu Pro Ile Ile Ser Gln Gly Tyr Ser His Ile Tyr Asp
    210                 215                 220

Ser Ser Lys Arg Tyr Phe Pro Gly Gly Glu Leu Glu Ile Arg Val Leu
225                 230                 235                 240

Glu Pro Ile Pro Thr Thr Gly Leu Thr Thr Asp Asp Val Asn Asp Leu
                245                 250                 255

Met Asp Lys Thr Arg Asn Leu Met Leu Lys His Leu Lys Glu Met Asp
            260                 265                 270

Ser Gln Tyr Ser Ser Ser Thr Ala Glu Asn Gly Ser Thr His Ile Asp
```

275 280 285

Ala Asp Ile Ala Lys Ser Thr Ala Thr Ser Ile Gly Asn Thr Asp Asp
290 295 300

Ala Ile Thr Lys Arg Arg Thr Pro Lys Glu
305 310

<210> 91
<211> 945
<212> DNA
<213> Mortierella alpina

<400> 91
atgtctatcg gatcttctaa ccccgtgctg ctggccgcta ttcccttcgt gtacctgttc 60

gtgctgcccc gagtgctggc cttcctgcct cagaaggccc agttcctggc caagtgcatc 120

gtggtgctga tcgctaccct gatcatgtct gtggccggct gcttcatctc tatcgtgtgc 180

gccctgctgg acaagcgata cgtgatcaac tacgtggtgt ctcgactgtt ctctttcctg 240

gccgctcgac cctgcggcgt gacctacaag atcgtcggcg aggaacacct ggacaagtac 300

cccgccattg tggtgtgcaa ccaccagtcc tctatggaca tgatggtgct gggccgagtg 360

ttccccaagc actgcgtggt gatggccaag aaggaactgc tgtacttccc cttcctgggc 420

atgttcatga agctgtctaa cgccatcttc atcgaccgaa agaaccacaa gaaggccatc 480

gagtctacca ctcaggccgt ggccgacatg aagaagcaca actctggcat ctggatcttc 540

cccgagggca cccgatctcg actggacaag gccgacctgc tgcctttcaa gaagggcgcc 600

ttccacctgg ccatccaggc tcagctgccc attctgccca tcatctctca gggctactct 660

cacatctacg actcttctaa gcgatacttc cccggtggcg agctggaaat ccgagtgctc 720

gagcccattc ctaccaccgg cctgaccacc gacgacgtga acgacctgat ggacaagacc 780

cgaaacctga tgctgaagca cctgaaggaa atggactctc agtactcttc gtctaccgcc 840

gagaacggct ctacccacat cgacgccgac attgccaagt ccaccgccac ctctatcggc 900

aacactgacg acgccattac caagcgacga acccctaagg aatag 945

<210> 92
<211> 9103
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of the NotI DNA fragment of pAT207 used to transform Y. lipolytica cells.

<400> 92
gcggccgctg tcgggaaccg cgttcaggtg gaacaggacc acctcccttg cacttcttgg 60

tatatcagta taggctgatg tattcatagt ggggtttttc ataataaatt tactaacggc 120

aggcaacatt cactcggctt aaacgcaaaa cggaccgtct tgatatcttc tgacgcattg 180

accaccgaga aatagtgtta gttaccgggt gagttattgt tcttctacac aggcgacgcc 240

catcgtctag agttgatgta ctaactcaga tttcactacc taccctatcc ctggtacgca 300

caaagcactt tgctagaagg tataacttcg tatagcatac attatacgaa gttattctga 360

attccgagaa acacaacaac atgccccatt ggacagatca tgcggataca caggttgtgc 420

agtatcatac atactcgatc agacaggtcg tctgaccatc atacaagctg aacaagcgct 480

ccatacttgc acgctctcta tatacacagt taaattacat atccatagtc taacctctaa 540

cagttaatct tctggtaagc ctcccagcca gccttctggt atcgcttggc ctcctcgata 600

ggatctcggt tctggccgta cagacctcgg ccgacaatta tgatatccgt tccggtagac 660

atgacatcct caacagttcg gtactgctgt ccgagagcgt ctcccttgtc gtcaagaccc 720

accccggggg tcagaataag ccagtcctca gagtcgccct taggtcggtt ctgggcaacg 780

aagccaacca caaactcggg gtcggatcgg gcaagctcaa tggtctgctt ggagtactcg 840

ccagtggcca gagagccctt gcaagacagc tcggccagca tgagcagacc tctggccagc 900

ttctcgttgg gagaggggac caggaactcc ttgtactggg agttctcgta gtcagagacg 960

tcctccttct tctgttcaga gacagtttcc tcggcaccag ctcgcaggcc agcaatgatt 1020

ccggttccgg gtacaccgtg ggcgttggtg atatcggacc actcggcgat tcggtagaca 1080

ccgttcttgt actggtgctt gacagtgttg ccaatatctg cgaactttct gtcctcgaac 1140

aggaagaaac cgtgcttaag agcaagttcc ttgaggggga gcacagtgct ggcgtaggtg 1200

aagtcgtcaa tgatgtcgat atgggtcttg atcatgcaca cataaggtcc gaccttatcg 1260

gcaagctcaa tgagctcctt ggtggtggta acatccagag aagcacacag gttggttttc 1320

ttggctgcca cgagcttgag cactcgagcg gcaaaggcgg acttgtggac gttagctcga 1380

gcttcgtagg agggcatttt ggtggtgaag aggagactga aataaattta gtctgcagaa 1440

ctttttatcg gaaccttatc tggggcagtg aagtatatgt tatggtaata gttacgagtt 1500

agttgaactt atagatagac tggactatac ggctatcggt ccaaattaga aagaacgtca 1560

atggctctct gggcgcgaat tcgtataact tcgtatagca ggagttatcc gaagcgaacg 1620

gcgatacgcg tatcgatacg cgtgcatgct gaggtgtctc acaagtgccg tgcagtcccg 1680

cccccacttg cttctctttg tgtgtagtgt acgtacatta tcgagagggt tgttcccgcc 1740

cacctcgatc cggcatgctg aggtgtctca caagtgccgt gcagtcccgc ccccacttgc 1800

ttctctttgt gtgtagtgta cgtacattat cgagagggtt gttcccgccc acctcgatcc 1860

ggcatgcact gatcacgggc aaaagtgcgt tcgatagaga gggggttggc ggcgcatttg 1920

tgtcccaaaa aacagcccca attgccccaa ttgaccccaa attgacccag tagcggaccc 1980

aaccccggcg agagcccccc tcaccccaca tatcaaacct cccccggttc ccacacttgc 2040

cgttaagggc gtagggtact gcagtctgga atctacgctt gttcagactt tgtactagtt 2100

tctttgtctg gccatccggg taacccatgc cggacgcaaa atagactact gaaaattttt 2160

ttgctttgtg gttgggactt tagccaaggg tataaaagac caccgtcccc gaattacctt 2220

tcctcttctt ttctctctct ccttgtcaac tcacacccga aggatccaat gatgctgaaa 2280

ttatcttctt cacgttcgcc gttggcaaga attccgaccc gacctcgacc taattccatc 2340

ccgcccagaa tcattgttgt tagcagttct tctagcaagg ttaatccttt gaaaactgaa 2400

gctgttgtca gctcgggatt ggcggataga ttgcgactgg gttcgttgac cgaagatggt 2460

ttatcctaca aagagaaatt tattgtgcga tgctatgagg tgggcatcaa taaaacagcc 2520

accgtggaga ctattgcaaa cttgctccaa gaagtcggtt gtaaccacgc ccaatccgtg 2580

ggctattcca ctggcggctt ttctactacc ccgaccatgc gaaaattgcg gctgatttgg 2640

gtcactgcca gaatgcatat cgagatatat aaatatcccg cctggtctga tgtcgttgag 2700

attgaatcgt ggggacaagg tgaaggcaag attggtacac gccgagactg gatcttacgt 2760

gattatgcta ccggtcaagt tatcggacga gcgacatcga agtgggttat gatgaatcaa 2820

gatacccgac gacttcaaaa ggtcgatgtg gatgtgcggg atgagtacct cgtgcattgt 2880

ccgcgtgagt tgcgtttggc gttccccgag gagaataatt ctagcttgaa gaagatctcc 2940

aagttggaag atccttccca atattcgaaa ctgggcttgg tgccacgacg tgccgacttg 3000

gatatgaacc aacacgtcaa taatgttact tatattggat gggtcctgga atctatgcct 3060

caagagatca tagacactca tgaattgcag acaattacct tggattatcg acgcgagtgc 3120

caacatgatg atgttgtgga tagcttgacc tctcccgaac cctctgagga tgccgaagca 3180

gtgttcaatc ataacggaac aaacggctcc gctaatgtct ctgctaatga ccatggttgc 3240

agaaatttct tgcatctgtt gcgattaagc ggcaatggtt tggaaattaa tcggggacga 3300

accgaatgga gaaagaaacc cacacggtag tctagtgtct gtggtatcta agctatttat 3360

cactctttac aacttctacc tcaactatct actttaataa atgaatatcg tttattctct 3420

atgattactg tatatgcgtt cctctaagac aaatcggctt cgatacgcgt atcgatacgc 3480

gtgcatgctg aggtgtctca caagtgccgt gcagtcccgc ccccacttgc ttctctttgt 3540

gtgtagtgta cgtacattat cgagagggtt gttcccgccc acctcgatcc ggcatgctga 3600

ggtgtctcac aagtgccgtg cagtcccgcc cccacttgct tctctttgtg tgtagtgtac 3660

gtacattatc gagagggttg ttcccgccca cctcgatccg gcatgctgag gtgtctcaca 3720

agtgccgtgc agtcccgccc ccacttgctt ctctttgtgt gtagtgtacg tacattatcg 3780

agagggttgt tcccgcccac ctcgatccgg catgctgagg tgtctcacaa gtgccgtgca 3840

gtcccgcccc cacttgcttc tctttgtgtg tagtgtacgt acattatcga gagggttgtt 3900

cccgcccacc tcgatccggc atgcactgat cacgggcaaa agtgcgttcg atagagaggg 3960

ggttggcggc gcatttgtgt cccaaaaaac agccccaatt gccccaattg accccaaatt 4020

gacccagtag cggacccaac cccggcgaga gccccccttca ccccacatat caaacctccc 4080

ccggttccca cacttgccgt taagggcgta gggtactgca gtctggaatc tacgcttgtt 4140

cagactttgt actagtttct ttgtctggcc atccgggtaa cccatgccgg acgcaaaata 4200

gactactgaa aatttttttg ctttgtggtt gggactttag ccaagggtat aaaagaccac 4260

cgtccccgaa ttacctttcc tcttcttttc tctctctcct tgtcaactca cacccgaagg 4320

atccacaaat gcgagactac gagcatgttg tggaaagttt tgactactta caatccgcaa 4380

ctcaagacct tcacggagag ttgaccgccc tgaatgcttg tgttgagtgt tgtgatagac 4440

acgcgcatgg tgaagctgtt gcccttttact gtgaggccca agacggtcac gcggaacgat 4500

atcggttccg agatttgcag cgtcaagccg cacgatttgg taactttctt cgagagcagg 4560

gagtgaaacc tggtgaccga gtggctggtt tgatgccccg taccgtcgag ttgcttatcg 4620

ctatcctggg aacatggcgg attggagctg tgtatcagcc cttgttcacc gcgtttggtc 4680

cgaaagctat tgagcagcgg ctgaattgta gcaatgctcg atggattgtc actgatcccc 4740

ataatcgacc aaaattggat gatgttaccg attgtccgag tatagtcgtc actggcggtg 4800

caccccaaaa tccggcagac catcactttt ggtccgcctt gaatcgacaa gccgatgact 4860

gtgctcctgt cttattagac gcttccgctc ccttcttgct catgtgcact tcgggtacta 4920

ccggccctgc gaagcctttg gaagtccccc tctcggccat cttagctttt aaaggttaca 4980

tgcgagatgc catcgatctt agagcagatg atcgattttg gaatttggcc gatccaggat 5040

gggcgtacgg actgtactat gccgtcactg gacccttggc ctgcggctat gccaccttgt 5100

tttatgatgg accgtttacc gttgagtcca ctcgacatat aatcgctaaa tatgccataa 5160

ataatttggc tggatcaccg acagcttatc gattcttgat tgctgccggc gctgaatttg 5220

ctgatgcagt gcgcggtcga ttgcgtgctg tgtctagtgc gggcgagccc ttgaatcccc 5280

aggtggttcg atggttcgct gagcaactgg gagtcgtcat ccatgatcat tatggtcaga 5340

cagagatagg tatggtcctt tgtaaccatc atggacttcg acatcccgtg cgtgagggtt 5400

cggcaggata cgcagttcct ggttatcgta ttgtggtcct cgacaaagct catcgagagc 5460

tgccggcagg tcaacctgga gttttggcag tcgaccgaga acgatcccct ctgtgttggt 5520

ttgatggtta tttgggaatg cctacacaag cgttcgccgg acgatactat ttgtccggag 5580

acatcgtcga actgaacgat gatggttcta tttcttttgt tggtcgaaac gacgatttga 5640

tcactacctc tggatatcga gtgggtccat ttgacgtcga gtccgccttg attgagcatc 5700

ccgctgtcgt tgaagctgca gttataggaa agccagaccc tcagcggaca gaactgatta 5760

aagccttcgt tgtcttgaat actccatatc tcccttcgcc agaattagcc gaagagttgc 5820

gtttgcacgt gcgacaacgg ttggctgccc acgcataccc ccgagagatg gaattcgttg 5880

atcacttacc gaaaacacct tcgggaaaac tgcagcgatt tatcttgaga aatcaagaaa 5940

tcgcgaaaca gcaagccttg ggatagggat tgtctgtggt atctaagcta tttatcactc 6000

tttacaactt ctacctcaac tatctacttt aataaatgaa tatcgtttat tctctatgat 6060

tactgtatat gcgttcctct aagacaaatc ggtcacgata cgcgtatcga tacgcgtgca 6120

tgctgaggtg tctcacaagt gccgtgcagt cccgccccca cttgcttctc tttgtgtgta 6180

gtgtacgtac attatcgaga gggttgttcc cgcccacctc gatccggcat gctgaggtgt 6240

ctcacaagtg ccgtgcagtc ccgcccccac ttgcttctct ttgtgtgtag tgtacgtaca 6300

ttatcgagag ggttgttccc gcccacctcg atccggcatg ctgaggtgtc tcacaagtgc 6360

cgtgcagtcc cgcccccact tgcttctctt tgtgtgtagt gtacgtacat tatcgagagg 6420

gttgttcccg cccacctcga tccggcatgc tgaggtgtct cacaagtgcc gtgcagtccc 6480

gcccccactt gcttctcttt gtgtgtagtg tacgtacatt atcgagaggg ttgttcccgc 6540

ccacctcgat ccggcatgca ctgatcacgg gcaaaagtgc gttcgataga gaggggttg 6600

gcggcgcatt tgtgtcccaa aaaacagccc caattgcccc aattgacccc aaattgaccc 6660

agtagcggac ccaaccccgg cgagagcccc cttcacccca catatcaaac ctcccccggt 6720

tcccacactt gccgttaagg gcgtagggta ctgcagtctg gaatctacgc ttgttcagac 6780

tttgtactag tttctttgtc tggccatccg ggtaacccat gccggacgca aaatagacta 6840

ctgaaaattt ttttgctttg tggttgggac tttagccaag ggtataaaag accaccgtcc 6900

ccgaattacc tttcctcttc ttttctctct ctccttgtca actcacaccc gaaggatccc 6960

caccacaatg tctatctcta actctcccga gaacctggac cagaagccta acatccgaaa 7020

cgagtctgag gaccgatctt tctctgtgcg acgacgaact aactctaacg tggtgaacgc 7080

tgtgcccgac tcttgctcta aggtgggcac catcgagact gtgaccaagt tctcttctgt 7140

ggactctgcc gccgagatgt cctctggcga ggaccgaaac cgaatcggca acggtgacga 7200

ccgaaaggaa gaggccaact ctctgtctaa cggccaggtg tacaacgagc gagtcggcat 7260

cggcgaggga cgatctcccg aggccgaggt gcccttccga ttctcttacc gagcctctgc 7320

tcccgctcac cgacgagtga aggaatctcc cctgtcctct gacgccatct tccgacagtc 7380

tcacgccggc ctgttcaacc tgtgcatcct ggtgctggtg gccgtgaacg gccgactgat 7440

cattgagaac ctgatgaagt acggcctgct gatccgagcc ggattctggt tctcttcgaa 7500

gtctctgcga gactggcccc tgctgatgtg ctgcctgtct ctgcctatct tccctctggg 7560

cgccttcacc gtcgagaagt ctgccaagca caagttcatc cccgagcctg tggtgatgtc 7620

tctgcacgtg ctgatctgta ccgcctctct gctgtacccc atcttcgtga tcctccgatg 7680

cggatctgcc gtgctgttcg gaatcgccct gatgctgctg gcctctatcg tgtggctgaa 7740

gctggtgtct tacgcccaca ccaactgcga catgcgagcc ctggccaagc tgaccgacaa 7800

ggtggcctct aagtcctccg acatggacta ctcttacgac gtgtctttcg agtctctggc 7860

ctacttcatg gtggctccca ctctgtgcta ccaggcttct taccctcgat ctacctgcat 7920

ccgaaagggc tgggtgatcc gacagctggt caagctggtg attttcaccg gcgtgatggg 7980

cttcatcgtc gagcagtaca tcaaccccat cgtgcagcac tctcagcacc ctctgaaggg 8040

caacctgctg tacgctatcg agcgagtgct gaagctctct gtccctactc tgtacgtctg 8100

gctgtgcatg ttctactgtt tcttccacct gtggctcaac atcctggccg agctgctgcg 8160

attcggcgac cgagagttct acaaggactg gtggaacgcc aagaccgtgg aagagtactg 8220

gcgaatgtgg aacatgcccg tgcacaagtg gatggtgcga cacatctact tcccttgcct 8280

gcgaaacgga ctgccccgag aggtgaccct ggtgatcgtg ttcttcctgt ctgccatttt 8340

ccacgagctg tgcatcggcg tgccctgcca catcttcaag ttctgggcct tcatcggcat 8400

catgttccag gtgcctctgg tgctgatcac caactacctg cagaacaagt tccgatcttc 8460

tatggtgggc aacatggtgt tctggttttt cttctcgatc ctgggacagc ccatgtctgt 8520

gctgctgtac taccacgacg tgatgaacag aaaggaaaac ggcgtttcta agtaggtatg 8580

tgtctgtggt atctaagcta tttatcactc tttacaactt ctacctcaac tatctacttt 8640

aataaatgaa tatcgtttat tctctatgat tactgtatat gcgttcctct aagacagagt 8700

catgtgtaac actcgctctg gagagttagt catccgacag ggtaactcta atctcccaac 8760

accttattaa ctctgcgtaa ctgtaactct tcttgccacg tcgatcttac tcaattttcc 8820

tgctcatcat ctgctggatt gttgtctatc gtctggctct aatacattta ttgtttattg 8880

cccaaacaac tttcattgca cgtaagtgaa ttgttttata acagcgttcg ccaattgctg 8940

cgccatcgtc gtccggctgt cctaccgtta ggggtagtgt gtctcacact accgaggtta 9000

ctagagttgg gaaagcgata ctgcctcgga cacaccacct gggtcttacg actgcagaga 9060

gaatcggcgt tacctctctc acaaagccct tcagtgcggc cgc 9103

<210> 93
<211> 9103
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of the NotI DNA fragment of pAT208 used to transform Y. lipolytica cells.

<400> 93
gcggccgctg tcgggaaccg cgttcaggtg gaacaggacc acctcccttg cacttcttgg 60

tatatcagta taggctgatg tattcatagt ggggtttttc ataataaatt tactaacggc 120

aggcaacatt cactcggctt aaacgcaaaa cggaccgtct tgatatcttc tgacgcattg 180

accaccgaga aatagtgtta gttaccgggt gagttattgt tcttctacac aggcgacgcc 240

catcgtctag agttgatgta ctaactcaga tttcactacc taccctatcc ctggtacgca 300

caaagcactt tgctagaagg tataacttcg tatagcatac attatacgaa gttattctga 360

attccgagaa acacaacaac atgccccatt ggacagatca tgcggataca caggttgtgc 420

agtatcatac atactcgatc agacaggtcg tctgaccatc atacaagctg aacaagcgct 480

ccatacttgc acgctctcta tatacacagt taaattacat atccatagtc taacctctaa 540

cagttaatct tctggtaagc ctcccagcca gccttctggt atcgcttggc ctcctcgata 600

ggatctcggt tctggccgta cagacctcgg ccgacaatta tgatatccgt tccggtagac 660

atgacatcct caacagttcg gtactgctgt ccgagagcgt ctcccttgtc gtcaagaccc 720

accccggggg tcagaataag ccagtcctca gagtcgccct taggtcggtt ctgggcaacg 780

aagccaacca caaactcggg gtcggatcgg gcaagctcaa tggtctgctt ggagtactcg 840

ccagtggcca gagagccctt gcaagacagc tcggccagca tgagcagacc tctggccagc 900

ttctcgttgg gagaggggac caggaactcc ttgtactggg agttctcgta gtcagagacg 960

tcctccttct tctgttcaga gacagtttcc tcggcaccag ctcgcaggcc agcaatgatt 1020

ccggttccgg gtacaccgtg ggcgttggtg atatcggacc actcggcgat tcggtagaca 1080

ccgttcttgt actggtgctt gacagtgttg ccaatatctg cgaactttct gtcctcgaac 1140

aggaagaaac cgtgcttaag agcaagttcc ttgaggggga gcacagtgct ggcgtaggtg 1200

aagtcgtcaa tgatgtcgat atgggtcttg atcatgcaca cataaggtcc gaccttatcg 1260

gcaagctcaa tgagctcctt ggtggtggta acatccagag aagcacacag gttggttttc 1320

ttggctgcca cgagcttgag cactcgagcg gcaaaggcgg acttgtggac gttagctcga 1380

gcttcgtagg agggcatttt ggtggtgaag aggagactga aataaattta gtctgcagaa 1440

ctttttatcg gaaccttatc tggggcagtg aagtatatgt tatggtaata gttacgagtt 1500

agttgaactt atagatagac tggactatac ggctatcggt ccaaattaga aagaacgtca 1560

atggctctct gggcgcgaat tcgtataact tcgtatagca ggagttatcc gaagcgaacg 1620

gcgatacgcg tatcgatacg cgtgcatgct gaggtgtctc acaagtgccg tgcagtcccg 1680

cccccacttg cttctctttg tgtgtagtgt acgtacatta tcgagagggt tgttcccgcc 1740

cacctcgatc cggcatgctg aggtgtctca caagtgccgt gcagtcccgc ccccacttgc 1800

ttctctttgt gtgtagtgta cgtacattat cgagagggtt gttcccgccc acctcgatcc 1860

ggcatgcact gatcacgggc aaaagtgcgt tcgatagaga gggggttggc ggcgcatttg 1920

tgtcccaaaa aacagcccca attgccccaa ttgaccccaa attgacccag tagcggaccc 1980

aaccccggcg agagcccccc tcaccccaca tatcaaacct cccccggttc ccacacttgc 2040

cgttaagggc gtagggtact gcagtctgga atctacgctt gttcagactt tgtactagtt 2100

tctttgtctg gccatccggg taacccatgc cggacgcaaa atagactact gaaaattttt 2160

ttgctttgtg gttgggactt tagccaaggg tataaaagac caccgtcccc gaattacctt 2220

tcctcttctt ttctctctct ccttgtcaac tcacacccga aggatccaat gatgctgaaa 2280

ttatcttctt cacgttcgcc gttggcaaga attccgaccc gacctcgacc taattccatc 2340

ccgcccagaa tcattgttgt tagcagttct tctagcaagg ttaatccttt gaaaactgaa 2400

gctgttgtca gctcgggatt ggcggataga ttgcgactgg gttcgttgac cgaagatggt 2460

ttatcctaca aagagaaatt tattgtgcga tgctatgagg tgggcatcaa taaaacagcc 2520

accgtggaga ctattgcaaa cttgctccaa gaagtcggtt gtaaccacgc ccaatccgtg 2580

ggctattcca ctggcggctt ttctactacc ccgaccatgc gaaaattgcg gctgatttgg 2640

gtcactgcca gaatgcatat cgagatatat aaatatcccg cctggtctga tgtcgttgag 2700

attgaatcgt ggggacaagg tgaaggcaag attggtacac gccgagactg gatcttacgt 2760

gattatgcta ccggtcaagt tatcggacga gcgacatcga agtgggttat gatgaatcaa 2820

gatacccgac gacttcaaaa ggtcgatgtg gatgtgcggg atgagtacct cgtgcattgt 2880

ccgcgtgagt tgcgtttggc gttccccgag gagaataatt ctagcttgaa gaagatctcc 2940

aagttggaag atccttccca atattcgaaa ctgggcttgg tgccacgacg tgccgacttg 3000

gatatgaacc aacacgtcaa taatgttact tatattggat gggtcctgga atctatgcct 3060

caagagatca tagacactca tgaattgcag acaattacct tggattatcg acgcgagtgc 3120

caacatgatg atgttgtgga tagcttgacc tctcccgaac cctctgagga tgccgaagca 3180

gtgttcaatc ataacggaac aaacggctcc gctaatgtct ctgctaatga ccatggttgc 3240

agaaatttct tgcatctgtt gcgattaagc ggcaatggtt tggaaattaa tcggggacga 3300

accgaatgga gaaagaaacc cacacggtag tctagtgtct gtggtatcta agctatttat 3360

cactctttac aacttctacc tcaactatct actttaataa atgaatatcg tttattctct 3420

atgattactg tatatgcgtt cctctaagac aaatcggctt cgatacgcgt atcgatacgc 3480

gtgcatgctg aggtgtctca caagtgccgt gcagtcccgc ccccacttgc ttctctttgt 3540

gtgtagtgta cgtacattat cgagagggtt gttcccgccc acctcgatcc ggcatgctga 3600

ggtgtctcac aagtgccgtg cagtcccgcc cccacttgct tctctttgtg tgtagtgtac 3660

gtacattatc gagagggttg ttcccgccca cctcgatccg gcatgctgag gtgtctcaca 3720

agtgccgtgc agtcccgccc ccacttgctt ctctttgtgt gtagtgtacg tacattatcg 3780

agagggttgt tcccgcccac ctcgatccgg catgctgagg tgtctcacaa gtgccgtgca 3840

gtcccgcccc cacttgcttc tctttgtgtg tagtgtacgt acattatcga gagggttgtt 3900

cccgcccacc tcgatccggc atgcactgat cacgggcaaa agtgcgttcg atagagaggg 3960

ggttggcggc gcatttgtgt cccaaaaaac agccccaatt gccccaattg accccaaatt 4020

gacccagtag cggacccaac cccggcgaga gccccccttca ccccacatat caaacctccc 4080

ccggttccca cacttgccgt taagggcgta gggtactgca gtctggaatc tacgcttgtt 4140

cagactttgt actagtttct ttgtctggcc atccgggtaa cccatgccgg acgcaaaata 4200

gactactgaa aatttttttg ctttgtggtt gggactttag ccaagggtat aaaagaccac 4260

cgtccccgaa ttacctttcc tcttcttttc tctctctcct tgtcaactca cacccgaagg 4320

atccacaaat gcgtgactac gaacatgttg tggagtcttt tgattacttg caatctgcaa 4380

cacaagatct ccacggtgag ttgaccgcac tgaatgcttg tgtggagtgt tgtgatcgac 4440

acgcgcatgg agaagctgtt gccctctact gtgaggccca agacggacac gcggaacgtt 4500

atagattccg agatctccag cgacaagctg ctcgttttgg taactttctc agagagcagg 4560

gtgttaaacc cggcgaccga gttgcaggat taatgccccg aaccgtggag ttgcttatcg 4620

ctatcctggg aacttggaga attggagccg tttatcagcc tttattcaca gcttttggac 4680

caaaagctat tgagcagcga ctgaattgtt ctaatgctcg atggattgtc acagatcctc 4740

ataatcgacc taaattggat gatgttacag attgtccttc gattgtcgtc acaggcggag 4800

ctcctcaaaa tccagccgac catcatttct ggtcagctct caatcgacaa gccgatgact 4860

gtgcccctgt tctcttggac gcttcggctc cctttctgct gatgtgcact tccggtacaa 4920

ccggtcccgc taagccctta gaagtccctt tgtctgccat attagcattt aaaggataca 4980

tgcgcgacgc tatagatctg cgagccgatg atcgattttg gaatttagcc gatcccggct 5040

gggcgtacgg cttatactat gctgtcactg gacccttggc atgcggctat gctacacttt 5100

tctatgatgg tccgtttaca gtcgagagca cacgccatat tatcgcgaaa tatgctataa 5160

ataatttggc tggaagccca actgcttatc gatttctcat tgctgccggt gctgaatttg 5220

ctgatgctgt ccggggccgt ttgcgagctg tgtcctctgc tggcgagcct ttgaatcctc 5280

aggtggtcag atggttcgct gagcaattgg gagttgtcat acatgatcat tatggtcaga 5340

cagagattgg tatggttttg tgtaaccatc atggactgcg acatcccgtg cgtgagggtt 5400

ctgccggata cgctgtccct ggttatagaa ttgtggtcct cgacaaagcc catcgtgagc 5460

tccccgctgg tcaacccgga gtgttggcag tggaccgtga acggtccccc ttatgttggt 5520

ttgatggata tttgggtatg ccaactcaag ctttcgctgg acgttactat ttgtctggag 5580

acatcgtcga attgaacgat gatggttcta tttcgtttgt tggacggaac gacgatctca 5640

tcacaacctc aggatataga gtgggcccat ttgacgtcga gtccgcctta atagagcatc 5700

cggctgtcgt ggaagccgct gtcattggaa agccagaccc ccagagaact gaactgatca 5760

aagccttcgt ggtcttgaat acccccttatt taccttctcc ggaattagcc gaagagttgc 5820

gattgcacgt gagacaacgg ttagcggctc acgcataccc ccgcgagatg gaattcgtgg 5880

atcacttgcc aaagacaccc agcggtaaac tgcagagatt tatcttgcgg aatcaagaaa 5940

tcgcaaaaca gcaagcttta ggctagggat tgtctgtggt atctaagcta tttatcactc 6000

tttacaactt ctacctcaac tatctacttt aataaatgaa tatcgtttat tctctatgat 6060

tactgtatat gcgttcctct aagacaaatc ggtcacgata cgcgtatcga tacgcgtgca 6120

tgctgaggtg tctcacaagt gccgtgcagt cccgccccca cttgcttctc tttgtgtgta 6180

gtgtacgtac attatcgaga gggttgttcc cgcccacctc gatccggcat gctgaggtgt 6240

ctcacaagtg ccgtgcagtc ccgcccccac ttgcttctct ttgtgtgtag tgtacgtaca 6300

ttatcgagag ggttgttccc gcccacctcg atccggcatg ctgaggtgtc tcacaagtgc 6360

cgtgcagtcc cgcccccact tgcttctctt tgtgtgtagt gtacgtacat tatcgagagg 6420

gttgttcccg cccacctcga tccggcatgc tgaggtgtct cacaagtgcc gtgcagtccc 6480

gcccccactt gcttctcttt gtgtgtagtg tacgtacatt atcgagaggg ttgttcccgc 6540

ccacctcgat ccggcatgca ctgatcacgg gcaaaagtgc gttcgataga gaggggggttg 6600

gcggcgcatt tgtgtcccaa aaaacagccc caattgcccc aattgacccc aaattgaccc 6660

agtagcggac ccaaccccgg cgagagcccc cttcacccca catatcaaac ctcccccggt 6720

tcccacactt gccgttaagg gcgtagggta ctgcagtctg gaatctacgc ttgttcagac 6780

tttgtactag tttctttgtc tggccatccg ggtaacccat gccggacgca aaatagacta 6840

ctgaaaattt ttttgctttg tggttgggac tttagccaag ggtataaaag accaccgtcc 6900

ccgaattacc tttcctcttc ttttctctct ctccttgtca actcacaccc gaaggatccc 6960

caccacaatg tctatctcta actctcccga gaacctggac cagaagccta acatccgaaa 7020

cgagtctgag gaccgatctt tctctgtgcg acgacgaact aactctaacg tggtgaacgc 7080

tgtgcccgac tcttgctcta aggtgggcac catcgagact gtgaccaagt tctcttctgt 7140

ggactctgcc gccgagatgt cctctggcga ggaccgaaac cgaatcggca acggtgacga 7200

ccgaaaggaa gaggccaact ctctgtctaa cggccaggtg tacaacgagc gagtcggcat 7260

cggcgaggga cgatctcccg aggccgaggt gcccttccga ttctcttacc gagcctctgc 7320

tcccgctcac cgacgagtga aggaatctcc cctgtcctct gacgccatct tccgacagtc 7380

tcacgccggc ctgttcaacc tgtgcatcct ggtgctggtg gccgtgaacg gccgactgat 7440

cattgagaac ctgatgaagt acggcctgct gatccgagcc ggattctggt tctcttcgaa 7500

gtctctgcga gactggcccc tgctgatgtg ctgcctgtct ctgcctatct tccctctggg 7560

cgccttcacc gtcgagaagt ctgccaagca caagttcatc cccgagcctg tggtgatgtc 7620

tctgcacgtg ctgatctgta ccgcctctct gctgtacccc atcttcgtga tcctccgatg 7680

cggatctgcc gtgctgttcg gaatcgccct gatgctgctg gcctctatcg tgtggctgaa 7740

gctggtgtct tacgcccaca ccaactgcga catgcgagcc ctggccaagc tgaccgacaa 7800

ggtggcctct aagtcctccg acatggacta ctcttacgac gtgtctttcg agtctctggc 7860

ctacttcatg gtggctccca ctctgtgcta ccaggcttct taccctcgat ctacctgcat 7920

ccgaaagggc tgggtgatcc gacagctggt caagctggtg attttcaccg gcgtgatggg 7980

cttcatcgtc gagcagtaca tcaaccccat cgtgcagcac tctcagcacc ctctgaaggg 8040

caacctgctg tacgctatcg agcgagtgct gaagctctct gtccctactc tgtacgtctg 8100

gctgtgcatg ttctactgtt tcttccacct gtggctcaac atcctggccg agctgctgcg 8160

attcggcgac cgagagttct acaaggactg gtggaacgcc aagaccgtgg aagagtactg 8220

gcgaatgtgg aacatgcccg tgcacaagtg gatggtgcga cacatctact tcccttgcct 8280

gcgaaacgga ctgccccgag aggtgaccct ggtgatcgtg ttcttcctgt ctgccatttt 8340

ccacgagctg tgcatcggcg tgccctgcca catcttcaag ttctgggcct tcatcggcat 8400

catgttccag gtgcctctgg tgctgatcac caactacctg cagaacaagt tccgatcttc 8460

tatggtgggc aacatggtgt tctggttttt cttctcgatc ctgggacagc ccatgtctgt 8520

gctgctgtac taccacgacg tgatgaacag aaaggaaaac ggcgtttcta agtaggtatg 8580

tgtctgtggt atctaagcta tttatcactc tttacaactt ctacctcaac tatctacttt 8640

aataaatgaa tatcgtttat tctctatgat tactgtatat gcgttcctct aagacagagt 8700

catgtgtaac actcgctctg gagagttagt catccgacag ggtaactcta atctcccaac 8760

accttattaa ctctgcgtaa ctgtaactct tcttgccacg tcgatcttac tcaattttcc 8820

tgctcatcat ctgctggatt gttgtctatc gtctggctct aatacattta ttgtttattg 8880

cccaaacaac tttcattgca cgtaagtgaa ttgttttata acagcgttcg ccaattgctg 8940

cgccatcgtc gtccggctgt cctaccgtta ggggtagtgt gtctcacact accgaggtta 9000

ctagagttgg gaaagcgata ctgcctcgga cacaccacct gggtcttacg actgcagaga 9060

gaatcggcgt tacctctctc acaaagccct tcagtgcggc cgc 9103

<210> 94
<211> 9112
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of the NotI DNA fragment of pAT209 used to transform Y. lipolytica cells.

<400> 94
gcggccgctg tcgggaaccg cgttcaggtg gaacaggacc acctcccttg cacttcttgg 60

tatatcagta taggctgatg tattcatagt ggggtttttc ataataaatt tactaacggc 120

aggcaacatt cactcggctt aaacgcaaaa cggaccgtct tgatatcttc tgacgcattg 180

accaccgaga aatagtgtta gttaccgggt gagttattgt tcttctacac aggcgacgcc 240

catcgtctag agttgatgta ctaactcaga tttcactacc taccctatcc ctggtacgca 300

caaagcactt tgctagaagg tataacttcg tatagcatac attatacgaa gttattctga 360

attccgagaa acacaacaac atgccccatt ggacagatca tgcggataca caggttgtgc 420

agtatcatac atactcgatc agacaggtcg tctgaccatc atacaagctg aacaagcgct 480

ccatacttgc acgctctcta tatacacagt taaattacat atccatagtc taacctctaa 540

cagttaatct tctggtaagc ctcccagcca gccttctggt atcgcttggc ctcctcgata 600

ggatctcggt tctggccgta cagacctcgg ccgacaatta tgatatccgt tccggtagac 660

atgacatcct caacagttcg gtactgctgt ccgagagcgt ctcccttgtc gtcaagaccc 720

accccggggg tcagaataag ccagtcctca gagtcgccct taggtcggtt ctgggcaacg 780

aagccaacca caaactcggg gtcggatcgg gcaagctcaa tggtctgctt ggagtactcg 840

ccagtggcca gagagccctt gcaagacagc tcggccagca tgagcagacc tctggccagc 900

ttctcgttgg gagaggggac caggaactcc ttgtactggg agttctcgta gtcagagacg 960

tcctccttct tctgttcaga gacagtttcc tcggcaccag ctcgcaggcc agcaatgatt 1020

ccggttccgg gtacaccgtg ggcgttggtg atatcggacc actcggcgat tcggtagaca 1080

ccgttcttgt actggtgctt gacagtgttg ccaatatctg cgaactttct gtcctcgaac 1140

aggaagaaac cgtgcttaag agcaagttcc ttgaggggga gcacagtgct ggcgtaggtg 1200

aagtcgtcaa tgatgtcgat atgggtcttg atcatgcaca cataaggtcc gaccttatcg 1260

gcaagctcaa tgagctcctt ggtggtggta acatccagag aagcacacag gttggttttc 1320

ttggctgcca cgagcttgag cactcgagcg gcaaaggcgg acttgtggac gttagctcga 1380

gcttcgtagg agggcatttt ggtggtgaag aggagactga aataaattta gtctgcagaa 1440

ctttttatcg gaaccttatc tggggcagtg aagtatatgt tatggtaata gttacgagtt 1500

agttgaactt atagatagac tggactatac ggctatcggt ccaaattaga aagaacgtca 1560

atggctctct gggcgcgaat tcgtataact tcgtatagca ggagttatcc gaagcgaacg 1620

gcgatacgcg tatcgatacg cgtgcatgct gaggtgtctc acaagtgccg tgcagtcccg 1680

cccccacttg cttctctttg tgtgtagtgt acgtacatta tcgagagggt tgttcccgcc 1740

cacctcgatc cggcatgctg aggtgtctca caagtgccgt gcagtcccgc ccccacttgc 1800

```
ttctctttgt gtgtagtgta cgtacattat cgagagggtt gttcccgccc acctcgatcc     1860

ggcatgcact gatcacgggc aaaagtgcgt tcgatagaga gggggttggc ggcgcatttg     1920

tgtcccaaaa aacagcccca attgccccaa ttgaccccaa attgacccag tagcggaccc     1980

aaccccggcg agagcccccc tcaccccaca tatcaaacct cccccggttc ccacacttgc     2040

cgttaagggc gtagggtact gcagtctgga atctacgctt gttcagactt tgtactagtt     2100

tctttgtctg gccatccggg taacccatgc cggacgcaaa atagactact gaaaattttt     2160

ttgctttgtg gttgggactt tagccaaggg tataaaagac caccgtcccc gaattacctt     2220

tcctcttctt ttctctctct ccttgtcaac tcacacccga aggatccaat gatgttcaaa     2280

atttcgtcgt cactctcacc tgttgaccaa attcccccaa ttagcccttt gcccagacca     2340

cgaccccgtc caattacacc aagagttctg gcagtgtcca gctccagcgg caaaatcgtt     2400

aataatccct tgaaagccga aaccaccgaa gcggtcagtg gcgaactggc ccgcagattt     2460

agattgggtc gactggctga agatggattt agctataaag agaaattcat cgtcagatgt     2520

tatgaagttg gtattaataa aactgcgaca gtcgaaacct tggctaattt gctgcaagaa     2580

gtgggtggaa atcatgccca atcggttggc tttagcactg atggcttcgc taccactcat     2640

tccatgcgga aaatgcattt gatttgggtg accgctcgga tgcatataga aatctataaa     2700

tatcctgcct ggtcggacgt cattgaagtc gaaacttgga tcggagccga aggtcgtatt     2760

ggtaccagac gaaattggat tattaaagac tgcgcgacag atgaagttat tggtcgggct     2820

acaagcaaat gggtcatgat gaatcaagac acacggcgat tggaaaaggt ttcggatgat     2880

gttcgtgaag aacatttggt cttttccccc agagaaccgc gcctgccgtt ccccgacgaa     2940

aataactctt cattgaagaa aatctcgaaa ttagatgatc cggctcaata ttccaaattg     3000

agtttggaac cacgacgggg agacttagat atgaatcaac atgttaacaa tgttacttat     3060

ataggctggg ttctcgagag tatgccacaa gagatcattg acacacatga attgcagaca     3120

ataactttgg actaccggcg agaatgtcaa cacgatgatt tagtggactc gctgacctct     3180

cctgagcctt ctgaattttc tgaaactact aatggtagtg ccaatgtcag ccccaatgat     3240

aaccgctgtt taaacttctt gcatttgctg agacttagct ccgacggctc tgaaattaac     3300
```

cggggtcgta ctgtgtggag aaagaaaccc gcaaaataat ctagtgtctg tggtatctaa 3360

gctatttatc actctttaca acttctacct caactatcta ctttaataaa tgaatatcgt 3420

ttattctcta tgattactgt atatgcgttc ctctaagaca aatcggcttc gatacgcgta 3480

tcgatacgcg tgcatgctga ggtgtctcac aagtgccgtg cagtcccgcc cccacttgct 3540

tctctttgtg tgtagtgtac gtacattatc gagagggttg ttcccgccca cctcgatccg 3600

gcatgctgag gtgtctcaca agtgccgtgc agtcccgccc ccacttgctt ctctttgtgt 3660

gtagtgtacg tacattatcg agagggttgt tcccgcccac ctcgatccgg catgctgagg 3720

tgtctcacaa gtgccgtgca gtcccgcccc cacttgcttc tctttgtgtg tagtgtacgt 3780

acattatcga gagggttgtt cccgcccacc tcgatccggc atgctgaggt gtctcacaag 3840

tgccgtgcag tcccgccccc acttgcttct ctttgtgtgt agtgtacgta cattatcgag 3900

agggttgttc ccgcccacct cgatccggca tgcactgatc acgggcaaaa gtgcgttcga 3960

tagagagggg gttggcggcg catttgtgtc ccaaaaaaca gccccaattg ccccaattga 4020

ccccaaattg acccagtagc ggacccaacc ccggcgagag ccccccttcac cccacatatc 4080

aaacctcccc cggttcccac acttgccgtt aagggcgtag ggtactgcag tctggaatct 4140

acgcttgttc agactttgta ctagtttctt tgtctggcca tccgggtaac ccatgccgga 4200

cgcaaaatag actactgaaa atttttttgc tttgtggttg ggactttagc caagggtata 4260

aaagaccacc gtccccgaat taccttttcct cttcttttct ctctctcctt gtcaactcac 4320

acccgaagga tccacaaatg cgagactacg agcatgttgt ggaaagtttt gactacttac 4380

aatccgcaac tcaagacctt cacggagagt tgaccgccct gaatgcttgt gttgagtgtt 4440

gtgatagaca cgcgcatggt gaagctgttg ccctttactg tgaggcccaa gacggtcacg 4500

cggaacgata tcggttccga gatttgcagc gtcaagccgc acgatttggt aactttcttc 4560

gagagcaggg agtgaaacct ggtgaccgag tggctggttt gatgccccgt accgtcgagt 4620

tgcttatcgc tatcctggga acatggcgga ttggagctgt gtatcagccc ttgttcaccg 4680

cgtttggtcc gaaagctatt gagcagcggc tgaattgtag caatgctcga tggattgtca 4740

ctgatcccca taatcgacca aaattggatg atgttaccga ttgtccgagt atagtcgtca 4800

ctggcggtgc accccaaaat ccggcagacc atcacttttg gtccgccttg aatcgacaag 4860

ccgatgactg tgctcctgtc ttattagacg cttccgctcc cttcttgctc atgtgcactt 4920

cgggtactac cggccctgcg aagcctttgg aagtcccccc ctcggccatc ttagctttta 4980

aaggttacat gcgagatgcc atcgatctta gagcagatga tcgattttgg aatttggccg 5040

atccaggatg ggcgtacgga ctgtactatg ccgtcactgg acccttggcc tgcggctatg 5100

ccaccttgtt ttatgatgga ccgtttaccg ttgagtccac tcgacatata atcgctaaat 5160

atgccataaa taatttggct ggatcaccga cagcttatcg attcttgatt gctgccggcg 5220

ctgaatttgc tgatgcagtg cgcggtcgat tgcgtgctgt gtctagtgcg ggcgagccct 5280

tgaatcccca ggtggttcga tggttcgctg agcaactggg agtcgtcatc catgatcatt 5340

atggtcagac agagataggt atggtccttt gtaaccatca tggacttcga catcccgtgc 5400

gtgagggttc ggcaggatac gcagttcctg gttatcgtat tgtggtcctc gacaaagctc 5460

atcgagagct gccggcaggt caacctggag ttttggcagt cgaccgagaa cgatcccctc 5520

tgtgttggtt tgatggttat ttgggaatgc ctacacaagc gttcgccgga cgatactatt 5580

tgtccggaga catcgtcgaa ctgaacgatg atggttctat ttcttttgtt ggtcgaaacg 5640

acgatttgat cactacctct ggatatcgag tgggtccatt tgacgtcgag tccgccttga 5700

ttgagcatcc cgctgtcgtt gaagctgcag ttataggaaa gccagaccct cagcggacag 5760

aactgattaa agccttcgtt gtcttgaata ctccatatct cccttcgcca gaattagccg 5820

aagagttgcg tttgcacgtg cgacaacggt tggctgccca cgcatacccc cgagagatgg 5880

aattcgttga tcacttaccg aaaacacctt cgggaaaact gcagcgattt atcttgagaa 5940

atcaagaaat cgcgaaacag caagccttgg gatagggatt gtctgtggta tctaagctat 6000

ttatcactct ttacaacttc tacctcaact atctacttta ataaatgaat atcgtttatt 6060

ctctatgatt actgtatatg cgttcctcta agacaaatcg gtcacgatac gcgtatcgat 6120

acgcgtgcat gctgaggtgt ctcacaagtg ccgtgcagtc ccgcccccac ttgcttctct 6180

ttgtgtgtag tgtacgtaca ttatcgagag ggttgttccc gcccacctcg atccggcatg 6240

ctgaggtgtc tcacaagtgc cgtgcagtcc cgcccccact tgcttctctt tgtgtgtagt 6300

gtacgtacat tatcgagagg gttgttcccg cccacctcga tccggcatgc tgaggtgtct 6360

cacaagtgcc gtgcagtccc gcccccactt gcttctcttt gtgtgtagtg tacgtacatt 6420

atcgagaggg ttgttcccgc ccacctcgat ccggcatgct gaggtgtctc acaagtgccg 6480

tgcagtcccg cccccacttg cttctctttg tgtgtagtgt acgtacatta tcgagagggt 6540

tgttcccgcc cacctcgatc cggcatgcac tgatcacggg caaaagtgcg ttcgatagag 6600

agggggttgg cggcgcattt gtgtcccaaa aaacagcccc aattgcccca attgacccca 6660

aattgaccca gtagcggacc caaccccggc gagagccccc ttcaccccac atatcaaacc 6720

tcccccggtt cccacacttg ccgttaaggg cgtagggtac tgcagtctgg aatctacgct 6780

tgttcagact ttgtactagt ttctttgtct ggccatccgg gtaacccatg ccggacgcaa 6840

aatagactac tgaaaatttt tttgctttgt ggttgggact ttagccaagg gtataaaaga 6900

ccaccgtccc cgaattacct ttcctcttct tttctctctc tccttgtcaa ctcacacccg 6960

aaggatcccc accacaatgt ctatctctaa ctctcccgag aacctggacc agaagcctaa 7020

catccgaaac gagtctgagg accgatcttt ctctgtgcga cgacgaacta actctaacgt 7080

ggtgaacgct gtgcccgact cttgctctaa ggtgggcacc atcgagactg tgaccaagtt 7140

ctcttctgtg gactctgccg ccgagatgtc ctctggcgag gaccgaaacc gaatcggcaa 7200

cggtgacgac cgaaaggaag aggccaactc tctgtctaac ggccaggtgt acaacgagcg 7260

agtcggcatc ggcgagggac gatctcccga ggccgaggtg cccttccgat tctcttaccg 7320

agcctctgct cccgctcacc gacgagtgaa ggaatctccc ctgtcctctg acgccatctt 7380

ccgacagtct cacgccggcc tgttcaacct gtgcatcctg gtgctggtgg ccgtgaacgg 7440

ccgactgatc attgagaacc tgatgaagta cggcctgctg atccgagccg gattctggtt 7500

ctcttcgaag tctctgcgag actggcccct gctgatgtgc tgcctgtctc tgcctatctt 7560

ccctctgggc gccttcaccg tcgagaagtc tgccaagcac aagttcatcc ccgagcctgt 7620

ggtgatgtct ctgcacgtgc tgatctgtac cgcctctctg ctgtacccca tcttcgtgat 7680

cctccgatgc ggatctgccg tgctgttcgg aatcgccctg atgctgctgg cctctatcgt 7740

gtggctgaag ctggtgtctt acgcccacac caactgcgac atgcgagccc tggccaagct 7800

gaccgacaag gtggcctcta agtcctccga catggactac tcttacgacg tgtctttcga 7860

gtctctggcc tacttcatgg tggctcccac tctgtgctac caggcttctt accctcgatc 7920

tacctgcatc cgaaagggct gggtgatccg acagctggtc aagctggtga ttttcaccgg 7980

cgtgatgggc ttcatcgtcg agcagtacat caaccccatc gtgcagcact ctcagcaccc 8040

tctgaagggc aacctgctgt acgctatcga gcgagtgctg aagctctctg tccctactct 8100

gtacgtctgg ctgtgcatgt tctactgttt cttccacctg tggctcaaca tcctggccga 8160

gctgctgcga ttcggcgacc gagagttcta caaggactgg tggaacgcca agaccgtgga 8220

agagtactgg cgaatgtgga acatgcccgt gcacaagtgg atggtgcgac acatctactt 8280

cccttgcctg cgaaacggac tgccccgaga ggtgaccctg gtgatcgtgt tcttcctgtc 8340

tgccattttc cacgagctgt gcatcggcgt gccctgccac atcttcaagt tctgggcctt 8400

catcggcatc atgttccagg tgcctctggt gctgatcacc aactacctgc agaacaagtt 8460

ccgatcttct atggtgggca acatggtgtt ctggtttttc ttctcgatcc tgggacagcc 8520

catgtctgtg ctgctgtact accacgacgt gatgaacaga aaggaaaacg gcgtttctaa 8580

gtaggtatgt gtctgtggta tctaagctat ttatcactct ttacaacttc tacctcaact 8640

atctacttta ataaatgaat atcgtttatt ctctatgatt actgtatatg cgttcctcta 8700

agacagagtc atgtgtaaca ctcgctctgg agagttagtc atccgacagg gtaactctaa 8760

tctcccaaca ccttattaac tctgcgtaac tgtaactctt cttgccacgt cgatcttact 8820

caattttcct gctcatcatc tgctggattg ttgtctatcg tctggctcta atacatttat 8880

tgtttattgc ccaaacaact ttcattgcac gtaagtgaat tgttttataa cagcgttcgc 8940

caattgctgc gccatcgtcg tccggctgtc ctaccgttag gggtagtgtg tctcacacta 9000

ccgaggttac tagagttggg aaagcgatac tgcctcggac acaccacctg ggtcttacga 9060

ctgcagagag aatcggcgtt acctctctca caaagccctt cagtgcggcc gc 9112

<210> 95
<211> 9112
<212> DNA
<213> Artificial Sequence

<220>

<223> Nucleotide sequence of the NotI DNA fragment of pAT210 used to transform Y. lipolytica cells.

<400> 95

gcggccgctg tcgggaaccg cgttcaggtg gaacaggacc acctcccttg cacttcttgg 60

tatatcagta taggctgatg tattcatagt ggggtttttc ataataaatt tactaacggc 120

aggcaacatt cactcggctt aaacgcaaaa cggaccgtct tgatatcttc tgacgcattg 180

accaccgaga aatagtgtta gttaccgggt gagttattgt tcttctacac aggcgacgcc 240

catcgtctag agttgatgta ctaactcaga tttcactacc taccctatcc ctggtacgca 300

caaagcactt tgctagaagg tataacttcg tatagcatac attatacgaa gttattctga 360

attccgagaa acacaacaac atgccccatt ggacagatca tgcggataca caggttgtgc 420

agtatcatac atactcgatc agacaggtcg tctgaccatc atacaagctg aacaagcgct 480

ccatacttgc acgctctcta tatacacagt taaattacat atccatagtc taacctctaa 540

cagttaatct tctggtaagc ctcccagcca gccttctggt atcgcttggc ctcctcgata 600

ggatctcggt tctggccgta cagacctcgg ccgacaatta tgatatccgt tccggtagac 660

atgacatcct caacagttcg gtactgctgt ccgagagcgt ctcccttgtc gtcaagaccc 720

accccggggg tcagaataag ccagtcctca gagtcgccct taggtcggtt ctgggcaacg 780

aagccaacca caaactcggg gtcggatcgg gcaagctcaa tggtctgctt ggagtactcg 840

ccagtggcca gagagccctt gcaagacagc tcggccagca tgagcagacc tctggccagc 900

ttctcgttgg gagaggggac caggaactcc ttgtactggg agttctcgta gtcagagacg 960

tcctccttct tctgttcaga gacagtttcc tcggcaccag ctcgcaggcc agcaatgatt 1020

ccggttccgg gtacaccgtg ggcgttggtg atatcggacc actcggcgat tcggtagaca 1080

ccgttcttgt actggtgctt gacagtgttg ccaatatctg cgaactttct gtcctcgaac 1140

aggaagaaac cgtgcttaag agcaagttcc ttgaggggga gcacagtgct ggcgtaggtg 1200

aagtcgtcaa tgatgtcgat atgggtcttg atcatgcaca cataaggtcc gaccttatcg 1260

gcaagctcaa tgagctcctt ggtggtggta acatccagag aagcacacag gttggttttc 1320

ttggctgcca cgagcttgag cactcgagcg gcaaaggcgg acttgtggac gttagctcga 1380

gcttcgtagg agggcatttt ggtggtgaag aggagactga aataaattta gtctgcagaa 1440

ctttttatcg gaaccttatc tggggcagtg aagtatatgt tatggtaata gttacgagtt 1500

agttgaactt atagatagac tggactatac ggctatcggt ccaaattaga aagaacgtca 1560

atggctctct gggcgcgaat tcgtataact tcgtatagca ggagttatcc gaagcgaacg 1620

gcgatacgcg tatcgatacg cgtgcatgct gaggtgtctc acaagtgccg tgcagtcccg 1680

cccccacttg cttctctttg tgtgtagtgt acgtacatta tcgagagggt tgttcccgcc 1740

cacctcgatc cggcatgctg aggtgtctca caagtgccgt gcagtcccgc ccccacttgc 1800

ttctctttgt gtgtagtgta cgtacattat cgagagggtt gttcccgccc acctcgatcc 1860

ggcatgcact gatcacgggc aaaagtgcgt tcgatagaga gggggttggc ggcgcatttg 1920

tgtcccaaaa aacagcccca attgccccaa ttgaccccaa attgacccag tagcggaccc 1980

aaccccggcg agagcccccc tcaccccaca tatcaaacct cccccggttc ccacacttgc 2040

cgttaagggc gtagggtact gcagtctgga atctacgctt gttcagactt tgtactagtt 2100

tctttgtctg gccatccggg taacccatgc cggacgcaaa atagactact gaaaattttt 2160

ttgctttgtg gttgggactt tagccaaggg tataaaagac caccgtcccc gaattacctt 2220

tcctcttctt ttctctctct ccttgtcaac tcacacccga aggatccaat gatgttcaaa 2280

atttcgtcgt cactctcacc tgttgaccaa attcccccaa ttagcccttt gcccagacca 2340

cgaccccgtc caattacacc aagagttctg gcagtgtcca gctccagcgg caaaatcgtt 2400

aataatccct tgaaagccga aaccaccgaa gcggtcagtg gcgaactggc ccgcagattt 2460

agattgggtc gactggctga agatggattt agctataaag agaaattcat cgtcagatgt 2520

tatgaagttg gtattaataa aactgcgaca gtcgaaacct tggctaattt gctgcaagaa 2580

gtgggtggaa atcatgccca atcggttggc tttagcactg atggcttcgc taccactcat 2640

tccatgcgga aaatgcattt gatttgggtg accgctcgga tgcatataga aatctataaa 2700

tatcctgcct ggtcggacgt cattgaagtc gaaacttgga tcggagccga aggtcgtatt 2760

ggtaccagac gaaattggat tattaaagac tgcgcgacag atgaagttat tggtcgggct 2820

acaagcaaat gggtcatgat gaatcaagac acacggcgat tggaaaaggt ttcggatgat 2880

gttcgtgaag aacatttggt cttttccccc agagaaccgc gcctgccgtt ccccgacgaa 2940

aataactctt cattgaagaa aatctcgaaa ttagatgatc cggctcaata ttccaaattg 3000

agtttggaac cacgacgggg agacttagat atgaatcaac atgttaacaa tgttacttat 3060

ataggctggg ttctcgagag tatgccacaa gagatcattg acacacatga attgcagaca 3120

ataactttgg actaccggcg agaatgtcaa cacgatgatt tagtggactc gctgacctct 3180

cctgagcctt ctgaattttc tgaaactact aatggtagtg ccaatgtcag ccccaatgat 3240

aaccgctgtt taaacttctt gcatttgctg agacttagct ccgacggctc tgaaattaac 3300

cggggtcgta ctgtgtggag aaagaaaccc gcaaaataat ctagtgtctg tggtatctaa 3360

gctatttatc actctttaca acttctacct caactatcta ctttaataaa tgaatatcgt 3420

ttattctcta tgattactgt atatgcgttc ctctaagaca aatcggcttc gatacgcgta 3480

tcgatacgcg tgcatgctga ggtgtctcac aagtgccgtg cagtcccgcc cccacttgct 3540

tctctttgtg tgtagtgtac gtacattatc gagagggttg ttcccgccca cctcgatccg 3600

gcatgctgag gtgtctcaca agtgccgtgc agtcccgccc ccacttgctt ctctttgtgt 3660

gtagtgtacg tacattatcg agagggttgt tcccgcccac ctcgatccgg catgctgagg 3720

tgtctcacaa gtgccgtgca gtcccgcccc cacttgcttc tctttgtgtg tagtgtacgt 3780

acattatcga gagggttgtt cccgcccacc tcgatccggc atgctgaggt gtctcacaag 3840

tgccgtgcag tcccgccccc acttgcttct ctttgtgtgt agtgtacgta cattatcgag 3900

agggttgttc ccgcccacct cgatccggca tgcactgatc acgggcaaaa gtgcgttcga 3960

tagagagggg gttggcggcg catttgtgtc ccaaaaaaca gccccaattg ccccaattga 4020

ccccaaattg acccagtagc ggacccaacc ccggcgagag ccccccttcac cccacatatc 4080

aaacctcccc cggttcccac acttgccgtt aagggcgtag ggtactgcag tctggaatct 4140

acgcttgttc agactttgta ctagtttctt tgtctggcca tccgggtaac ccatgccgga 4200

cgcaaaatag actactgaaa atttttttgc tttgtggttg ggactttagc caagggtata 4260

aaagaccacc gtccccgaat taccttttcct cttcttttct ctctctcctt gtcaactcac 4320

acccgaagga tccacaaatg cgtgactacg aacatgttgt ggagtctttt gattacttgc 4380

aatctgcaac acaagatctc cacggtgagt tgaccgcact gaatgcttgt gtggagtgtt 4440

gtgatcgaca cgcgcatgga gaagctgttg ccctctactg tgaggcccaa gacggacacg 4500

cggaacgtta tagattccga gatctccagc gacaagctgc tcgttttggt aactttctca 4560

gagagcaggg tgttaaaccc ggcgaccgag ttgcaggatt aatgccccga accgtggagt 4620

tgcttatcgc tatcctggga acttggagaa ttggagccgt ttatcagcct ttattcacag 4680

cttttggacc aaaagctatt gagcagcgac tgaattgttc taatgctcga tggattgtca 4740

cagatcctca taatcgacct aaattggatg atgttacaga ttgtccttcg attgtcgtca 4800

caggcggagc tcctcaaaat ccagccgacc atcatttctg gtcagctctc aatcgacaag 4860

ccgatgactg tgcccctgtt ctcttggacg cttcggctcc ctttctgctg atgtgcactt 4920

ccggtacaac cggtcccgct aagcccttag aagtcccttt gtctgccata ttagcattta 4980

aaggatacat gcgcgacgct atagatctgc gagccgatga tcgattttgg aatttagccg 5040

atcccggctg ggcgtacggc ttatactatg ctgtcactgg acccttggca tgcggctatg 5100

ctacactttt ctatgatggt ccgtttacag tcgagagcac acgccatatt atcgcgaaat 5160

atgctataaa taatttggct ggaagcccaa ctgcttatcg atttctcatt gctgccggtg 5220

ctgaatttgc tgatgctgtc cggggccgtt tgcgagctgt gtcctctgct ggcgagcctt 5280

tgaatcctca ggtggtcaga tggttcgctg agcaattggg agttgtcata catgatcatt 5340

atggtcagac agagattggt atggttttgt gtaaccatca tggactgcga catcccgtgc 5400

gtgagggttc tgccggatac gctgtccctg gttatagaat tgtggtcctc gacaaagccc 5460

atcgtgagct ccccgctggt caacccggag tgttggcagt ggaccgtgaa cggtcccct 5520

tatgttggtt tgatggatat ttgggtatgc caactcaagc tttcgctgga cgttactatt 5580

tgtctggaga catcgtcgaa ttgaacgatg atggttctat ttcgtttgtt ggacggaacg 5640

acgatctcat cacaacctca ggatatagag tgggcccatt tgacgtcgag tccgccttaa 5700

tagagcatcc ggctgtcgtg gaagccgctg tcattggaaa gccagacccc cagagaactg 5760

aactgatcaa agccttcgtg gtcttgaata ccccttattt accttctccg gaattagccg 5820

aagagttgcg attgcacgtg agacaacggt tagcggctca cgcatacccc cgcgagatgg 5880

aattcgtgga tcacttgcca aagacaccca gcggtaaact gcagagattt atcttgcgga 5940

atcaagaaat cgcaaaacag caagctttag gctagggatt gtctgtggta tctaagctat 6000

ttatcactct ttacaacttc tacctcaact atctacttta ataaatgaat atcgtttatt 6060

ctctatgatt actgtatatg cgttcctcta agacaaatcg gtcacgatac gcgtatcgat 6120

acgcgtgcat gctgaggtgt ctcacaagtg ccgtgcagtc ccgcccccac ttgcttctct 6180

ttgtgtgtag tgtacgtaca ttatcgagag ggttgttccc gcccacctcg atccggcatg 6240

ctgaggtgtc tcacaagtgc cgtgcagtcc cgcccccact tgcttctctt tgtgtgtagt 6300

gtacgtacat tatcgagagg gttgttcccg cccacctcga tccggcatgc tgaggtgtct 6360

cacaagtgcc gtgcagtccc gcccccactt gcttctcttt gtgtgtagtg tacgtacatt 6420

atcgagaggg ttgttcccgc ccacctcgat ccggcatgct gaggtgtctc acaagtgccg 6480

tgcagtcccg cccccacttg cttctctttg tgtgtagtgt acgtacatta tcgagagggt 6540

tgttcccgcc cacctcgatc cggcatgcac tgatcacggg caaaagtgcg ttcgatagag 6600

agggggttgg cggcgcattt gtgtcccaaa aaacagcccc aattgcccca attgacccca 6660

aattgaccca gtagcggacc caaccccggc gagagccccc ttcaccccac atatcaaacc 6720

tcccccggtt cccacacttg ccgttaaggg cgtagggtac tgcagtctgg aatctacgct 6780

tgttcagact ttgtactagt ttctttgtct ggccatccgg gtaacccatg ccggacgcaa 6840

aatagactac tgaaaatttt tttgctttgt ggttgggact ttagccaagg gtataaaaga 6900

ccaccgtccc cgaattacct ttcctcttct tttctctctc tccttgtcaa ctcacacccg 6960

aaggatcccc accacaatgt ctatctctaa ctctcccgag aacctggacc agaagcctaa 7020

catccgaaac gagtctgagg accgatcttt ctctgtgcga cgacgaacta actctaacgt 7080

ggtgaacgct gtgcccgact cttgctctaa ggtgggcacc atcgagactg tgaccaagtt 7140

ctcttctgtg gactctgccg ccgagatgtc ctctggcgag gaccgaaacc gaatcggcaa 7200

cggtgacgac cgaaaggaag aggccaactc tctgtctaac ggccaggtgt acaacgagcg 7260

agtcggcatc ggcgagggac gatctcccga ggccgaggtg cccttccgat tctcttaccg 7320

agcctctgct cccgctcacc gacgagtgaa ggaatctccc ctgtcctctg acgccatctt 7380

ccgacagtct cacgccggcc tgttcaacct gtgcatcctg gtgctggtgg ccgtgaacgg 7440

ccgactgatc attgagaacc tgatgaagta cggcctgctg atccgagccg gattctggtt 7500

ctcttcgaag tctctgcgag actggcccct gctgatgtgc tgcctgtctc tgcctatctt 7560

ccctctgggc gccttcaccg tcgagaagtc tgccaagcac aagttcatcc ccgagcctgt 7620

ggtgatgtct ctgcacgtgc tgatctgtac cgcctctctg ctgtacccca tcttcgtgat 7680

cctccgatgc ggatctgccg tgctgttcgg aatcgccctg atgctgctgg cctctatcgt 7740

gtggctgaag ctggtgtctt acgcccacac caactgcgac atgcgagccc tggccaagct 7800

gaccgacaag gtggcctcta agtcctccga catggactac tcttacgacg tgtctttcga 7860

gtctctggcc tacttcatgg tggctcccac tctgtgctac caggcttctt accctcgatc 7920

tacctgcatc cgaaagggct gggtgatccg acagctggtc aagctggtga ttttcaccgg 7980

cgtgatgggc ttcatcgtcg agcagtacat caaccccatc gtgcagcact ctcagcaccc 8040

tctgaagggc aacctgctgt acgctatcga gcgagtgctg aagctctctg tccctactct 8100

gtacgtctgg ctgtgcatgt tctactgttt cttccacctg tggctcaaca tcctggccga 8160

gctgctgcga ttcggcgacc gagagttcta caaggactgg tggaacgcca agaccgtgga 8220

agagtactgg cgaatgtgga acatgcccgt gcacaagtgg atggtgcgac acatctactt 8280

cccttgcctg cgaaacggac tgccccgaga ggtgaccctg gtgatcgtgt tcttcctgtc 8340

tgccattttc cacgagctgt gcatcggcgt gccctgccac atcttcaagt tctgggcctt 8400

catcggcatc atgttccagg tgcctctggt gctgatcacc aactacctgc agaacaagtt 8460

ccgatcttct atggtgggca acatggtgtt ctggtttttc ttctcgatcc tgggacagcc 8520

catgtctgtg ctgctgtact accacgacgt gatgaacaga aaggaaaacg gcgtttctaa 8580

gtaggtatgt gtctgtggta tctaagctat ttatcactct ttacaacttc tacctcaact 8640

atctacttta ataaatgaat atcgtttatt ctctatgatt actgtatatg cgttcctcta 8700

agacagagtc atgtgtaaca ctcgctctgg agagttagtc atccgacagg gtaactctaa 8760

tctcccaaca ccttattaac tctgcgtaac tgtaactctt cttgccacgt cgatcttact 8820

caattttcct gctcatcatc tgctggattg ttgtctatcg tctggctcta atacatttat 8880

tgtttattgc ccaaacaact ttcattgcac gtaagtgaat tgttttataa cagcgttcgc 8940

caattgctgc gccatcgtcg tccggctgtc ctaccgttag gggtagtgtg tctcacacta 9000

ccgaggttac tagagttggg aaagcgatac tgcctcggac acaccacctg ggtcttacga 9060

ctgcagagag aatcggcgtt acctctctca caaagccctt cagtgcggcc gc 9112

<210> 96
<211> 8417
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of the NotI DNA fragment of pAT211 used to transform Y. lipolytica cells.

<400> 96
gcggccgctg tcgggaaccg cgttcaggtg gaacaggacc acctcccttg cacttcttgg 60

tatatcagta taggctgatg tattcatagt ggggtttttc ataataaatt tactaacggc 120

aggcaacatt cactcggctt aaacgcaaaa cggaccgtct tgatatcttc tgacgcattg 180

accaccgaga aatagtgtta gttaccgggt gagttattgt tcttctacac aggcgacgcc 240

catcgtctag agttgatgta ctaactcaga tttcactacc taccctatcc ctggtacgca 300

caaagcactt tgctagaagg tataacttcg tatagcatac attatacgaa gttattctga 360

attccgagaa acacaacaac atgccccatt ggacagatca tgcggataca caggttgtgc 420

agtatcatac atactcgatc agacaggtcg tctgaccatc atacaagctg aacaagcgct 480

ccatacttgc acgctctcta tatacacagt taaattacat atccatagtc taacctctaa 540

cagttaatct tctggtaagc ctcccagcca gccttctggt atcgcttggc ctcctcgata 600

ggatctcggt tctggccgta cagacctcgg ccgacaatta tgatatccgt tccggtagac 660

atgacatcct caacagttcg gtactgctgt ccgagagcgt ctcccttgtc gtcaagaccc 720

accccggggg tcagaataag ccagtcctca gagtcgccct taggtcggtt ctgggcaacg 780

aagccaacca caaactcggg gtcggatcgg gcaagctcaa tggtctgctt ggagtactcg 840

ccagtggcca gagagccctt gcaagacagc tcggccagca tgagcagacc tctggccagc 900

ttctcgttgg gagaggggac caggaactcc ttgtactggg agttctcgta gtcagagacg 960

tcctccttct tctgttcaga gacagtttcc tcggcaccag ctcgcaggcc agcaatgatt 1020

ccggttccgg gtacaccgtg ggcgttggtg atatcggacc actcggcgat tcggtagaca 1080

ccgttcttgt actggtgctt gacagtgttg ccaatatctg cgaactttct gtcctcgaac 1140

aggaagaaac cgtgcttaag agcaagttcc ttgaggggga gcacagtgct ggcgtaggtg 1200

aagtcgtcaa tgatgtcgat atgggtcttg atcatgcaca cataaggtcc gaccttatcg 1260

gcaagctcaa tgagctcctt ggtggtggta acatccagag aagcacacag gttggttttc 1320

ttggctgcca cgagcttgag cactcgagcg gcaaaggcgg acttgtggac gttagctcga 1380

gcttcgtagg agggcatttt ggtggtgaag aggagactga aataaattta gtctgcagaa 1440

ctttttatcg gaaccttatc tggggcagtg aagtatatgt tatggtaata gttacgagtt 1500

agttgaactt atagatagac tggactatac ggctatcggt ccaaattaga aagaacgtca 1560

atggctctct gggcgcgaat tcgtataact tcgtatagca ggagttatcc gaagcgaacg 1620

gcgatacgcg tatcgatacg cgtgcatgct gaggtgtctc acaagtgccg tgcagtcccg 1680

cccccacttg cttctctttg tgtgtagtgt acgtacatta tcgagagggt tgttcccgcc 1740

cacctcgatc cggcatgctg aggtgtctca caagtgccgt gcagtcccgc ccccacttgc 1800

ttctctttgt gtgtagtgta cgtacattat cgagagggtt gttcccgccc acctcgatcc 1860

ggcatgcact gatcacgggc aaaagtgcgt tcgatagaga gggggttggc ggcgcatttg 1920

tgtcccaaaa aacagcccca attgccccaa ttgaccccaa attgacccag tagcggaccc 1980

aaccccggcg agagcccccc tcaccccaca tatcaaacct cccccggttc ccacacttgc 2040

cgttaagggc gtagggtact gcagtctgga atctacgctt gttcagactt tgtactagtt 2100

tctttgtctg gccatccggg taacccatgc cggacgcaaa atagactact gaaaattttt 2160

ttgctttgtg gttgggactt tagccaaggg tataaaagac caccgtcccc gaattacctt 2220

tcctcttctt ttctctctct ccttgtcaac tcacacccga aggatccaat gatgctgaaa 2280

ttatcttctt cacgttcgcc gttggcaaga attccgaccc gacctcgacc taattccatc 2340

ccgcccagaa tcattgttgt tagcagttct tctagcaagg ttaatccttt gaaaactgaa 2400

gctgttgtca gctcgggatt ggcggataga ttgcgactgg gttcgttgac cgaagatggt 2460

ttatcctaca aagagaaatt tattgtgcga tgctatgagg tgggcatcaa taaaacagcc 2520

accgtggaga ctattgcaaa cttgctccaa gaagtcggtt gtaaccacgc ccaatccgtg 2580

ggctattcca ctggcggctt ttctactacc ccgaccatgc gaaaattgcg gctgatttgg 2640

gtcactgcca gaatgcatat cgagatatat aaatatcccg cctggtctga tgtcgttgag 2700

attgaatcgt ggggacaagg tgaaggcaag attggtacac gccgagactg gatcttacgt 2760

gattatgcta ccggtcaagt tatcggacga gcgacatcga agtgggttat gatgaatcaa 2820

gatacccgac gacttcaaaa ggtcgatgtg gatgtgcggg atgagtacct cgtgcattgt 2880

ccgcgtgagt tgcgtttggc gttccccgag gagaataatt ctagcttgaa gaagatctcc 2940

aagttggaag atccttccca atattcgaaa ctgggcttgg tgccacgacg tgccgacttg 3000

gatatgaacc aacacgtcaa taatgttact tatattggat gggtcctgga atctatgcct 3060

caagagatca tagacactca tgaattgcag acaattacct tggattatcg acgcgagtgc 3120

caacatgatg atgttgtgga tagcttgacc tctcccgaac cctctgagga tgccgaagca 3180

gtgttcaatc ataacggaac aaacggctcc gctaatgtct ctgctaatga ccatggttgc 3240

agaaatttct tgcatctgtt gcgattaagc ggcaatggtt tggaaattaa tcggggacga 3300

accgaatgga gaaagaaacc cacacggtag tctagtgtct gtggtatcta agctatttat 3360

cactctttac aacttctacc tcaactatct actttaataa atgaatatcg tttattctct 3420

atgattactg tatatgcgtt cctctaagac aaatcggctt cgatacgcgt atcgatacgc 3480

gtgcatgctg aggtgtctca caagtgccgt gcagtcccgc ccccacttgc ttctctttgt 3540

gtgtagtgta cgtacattat cgagagggtt gttcccgccc acctcgatcc ggcatgctga 3600

ggtgtctcac aagtgccgtg cagtcccgcc cccacttgct tctctttgtg tgtagtgtac 3660

gtacattatc gagagggttg ttcccgccca cctcgatccg gcatgctgag gtgtctcaca 3720

agtgccgtgc agtcccgccc ccacttgctt ctctttgtgt gtagtgtacg tacattatcg 3780

agagggttgt tcccgcccac ctcgatccgg catgctgagg tgtctcacaa gtgccgtgca 3840

gtcccgcccc cacttgcttc tctttgtgtg tagtgtacgt acattatcga gagggttgtt 3900

cccgcccacc tcgatccggc atgcactgat cacgggcaaa agtgcgttcg atagagaggg 3960

ggttggcggc gcatttgtgt cccaaaaaac agccccaatt gccccaattg accccaaatt 4020

gacccagtag cggacccaac cccggcgaga gccccccttca ccccacatat caaacctccc 4080

ccggttccca cacttgccgt taagggcgta gggtactgca gtctggaatc tacgcttgtt 4140

cagactttgt actagtttct ttgtctggcc atccgggtaa cccatgccgg acgcaaaata 4200

gactactgaa aatttttttg ctttgtggtt gggactttag ccaagggtat aaaagaccac 4260

cgtccccgaa ttacctttcc tcttcttttc tctctctcct tgtcaactca cacccgaagg 4320

atccacaaca caatgtctat cggatcttct aaccccgtgc tgctggccgc tattcccttc 4380

gtgtacctgt tcgtgctgcc ccgagtgctg gccttcctgc ctcagaaggc ccagttcctg 4440

gccaagtgca tcgtggtgct gatcgctacc ctgatcatgt ctgtggccgg ctgcttcatc 4500

tctatcgtgt gcgccctgct ggacaagcga tacgtgatca actacgtggt gtctcgactg 4560

ttctctttcc tggccgctcg accctgcggc gtgacctaca agatcgtcgg cgaggaacac 4620

ctggacaagt accccgccat tgtggtgtgc aaccaccagt cctctatgga catgatggtg 4680

ctgggccgag tgttccccaa gcactgcgtg gtgatggcca agaaggaact gctgtacttc 4740

cccttcctgg gcatgttcat gaagctgtct aacgccatct tcatcgaccg aaagaaccac 4800

aagaaggcca tcgagtctac cactcaggcc gtggccgaca tgaagaagca caactctggc 4860

atctggatct tccccgaggg cacccgatct cgactggaca aggccgacct gctgcctttc 4920

aagaagggcg ccttccacct ggccatccag gctcagctgc ccattctgcc catcatctct 4980

cagggctact ctcacatcta cgactcttct aagcgatact tccccggtgg cgagctggaa 5040

atccgagtgc tcgagcccat tcctaccacc ggcctgacca ccgacgacgt gaacgacctg 5100

atggacaaga cccgaaacct gatgctgaag cacctgaagg aaatggactc tcagtactct 5160

tcgtctaccg ccgagaacgg ctctacccac atcgacgccg acattgccaa gtccaccgcc 5220

acctctatcg gcaacactga cgacgccatt accaagcgac gaacccctaa ggaatagtaa 5280

ggattgtctg tggtatctaa gctatttatc actctttaca acttctacct caactatcta 5340

ctttaataaa tgaatatcgt ttattctcta tgattactgt atatgcgttc ctctaagaca 5400

aatcggtcac gatacgcgta tcgatacgcg tgcatgctga ggtgtctcac aagtgccgtg 5460

cagtcccgcc cccacttgct tctctttgtg tgtagtgtac gtacattatc gagagggttg 5520

ttcccgccca cctcgatccg gcatgctgag gtgtctcaca agtgccgtgc agtcccgccc 5580

ccacttgctt ctctttgtgt gtagtgtacg tacattatcg agagggttgt tcccgcccac 5640

ctcgatccgg catgctgagg tgtctcacaa gtgccgtgca gtcccgcccc cacttgcttc 5700

tctttgtgtg tagtgtacgt acattatcga gagggttgtt cccgcccacc tcgatccggc 5760

atgctgaggt gtctcacaag tgccgtgcag tcccgccccc acttgcttct ctttgtgtgt 5820

agtgtacgta cattatcgag agggttgttc ccgcccacct cgatccggca tgcactgatc 5880

acgggcaaaa gtgcgttcga tagagagggg gttggcggcg catttgtgtc ccaaaaaaca 5940

gccccaattg ccccaattga ccccaaattg acccagtagc ggacccaacc ccggcgagag 6000

ccccccttcac cccacatatc aaacctcccc cggttcccac acttgccgtt aagggcgtag 6060

ggtactgcag tctggaatct acgcttgttc agactttgta ctagtttctt tgtctggcca 6120

tccgggtaac ccatgccgga cgcaaaatag actactgaaa atttttttgc tttgtggttg 6180

ggactttagc caagggtata aaagaccacc gtccccgaat taccttttcct cttcttttct 6240

ctctctcctt gtcaactcac acccgaagga tccccaccac aatgtctatc tctaactctc 6300

ccgagaacct ggaccagaag cctaacatcc gaaacgagtc tgaggaccga tctttctctg 6360

tgcgacgacg aactaactct aacgtggtga acgctgtgcc cgactcttgc tctaaggtgg 6420

gcaccatcga gactgtgacc aagttctctt ctgtggactc tgccgccgag atgtcctctg 6480

gcgaggaccg aaaccgaatc ggcaacggtg acgaccgaaa ggaagaggcc aactctctgt 6540

ctaacggcca ggtgtacaac gagcgagtcg gcatcggcga gggacgatct cccgaggccg 6600

aggtgccctt ccgattctct taccgagcct ctgctcccgc tcaccgacga gtgaaggaat 6660

ctcccctgtc ctctgacgcc atcttccgac agtctcacgc cggcctgttc aacctgtgca 6720

tcctggtgct ggtggccgtg aacggccgac tgatcattga gaacctgatg aagtacggcc 6780

tgctgatccg agccggattc tggttctctt cgaagtctct gcgagactgg cccctgctga 6840

tgtgctgcct gtctctgcct atcttccctc tgggcgcctt caccgtcgag aagtctgcca 6900

agcacaagtt catccccgag cctgtggtga tgtctctgca cgtgctgatc tgtaccgcct 6960

ctctgctgta ccccatcttc gtgatcctcc gatgcggatc tgccgtgctg ttcggaatcg 7020

ccctgatgct gctggcctct atcgtgtggc tgaagctggt gtcttacgcc cacaccaact 7080

gcgacatgcg agccctggcc aagctgaccg acaaggtggc ctctaagtcc tccgacatgg 7140

actactctta cgacgtgtct ttcgagtctc tggcctactt catggtggct cccactctgt 7200

gctaccaggc ttcttaccct cgatctacct gcatccgaaa gggctgggtg atccgacagc 7260

tggtcaagct ggtgattttc accggcgtga tgggcttcat cgtcgagcag tacatcaacc 7320

ccatcgtgca gcactctcag caccctctga agggcaacct gctgtacgct atcgagcgag 7380

tgctgaagct ctctgtccct actctgtacg tctggctgtg catgttctac tgtttcttcc 7440

acctgtggct caacatcctg gccgagctgc tgcgattcgg cgaccgagag ttctacaagg 7500

actggtggaa cgccaagacc gtggaagagt actggcgaat gtggaacatg cccgtgcaca 7560

agtggatggt gcgacacatc tacttccctt gcctgcgaaa cggactgccc cgagaggtga 7620

ccctggtgat cgtgttcttc ctgtctgcca ttttccacga gctgtgcatc ggcgtgccct 7680

gccacatctt caagttctgg gccttcatcg gcatcatgtt ccaggtgcct ctggtgctga 7740

tcaccaacta cctgcagaac aagttccgat cttctatggt gggcaacatg gtgttctggt 7800

ttttcttctc gatcctggga cagcccatgt ctgtgctgct gtactaccac gacgtgatga 7860

acagaaagga aaacggcgtt tctaagtagg tatgtgtctg tggtatctaa gctatttatc 7920

actctttaca acttctacct caactatcta ctttaataaa tgaatatcgt ttattctcta 7980

tgattactgt atatgcgttc ctctaagaca gagtcatgtg taacactcgc tctggagagt 8040

tagtcatccg acagggtaac tctaatctcc caacacctta ttaactctgc gtaactgtaa 8100

ctcttcttgc cacgtcgatc ttactcaatt ttcctgctca tcatctgctg gattgttgtc 8160

tatcgtctgg ctctaataca tttattgttt attgcccaaa caactttcat tgcacgtaag 8220

tgaattgttt tataacagcg ttcgccaatt gctgcgccat cgtcgtccgg ctgtcctacc 8280

gttaggggta gtgtgtctca cactaccgag gttactagag ttgggaaagc gatactgcct 8340

cggacacacc acctgggtct tacgactgca gagagaatcg gcgttacctc tctcacaaag 8400

ccttcagtg cggccgc 8417

<210> 97
<211> 8426
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of the NotI DNA fragment of pAT212 used to transform Y. lipolytica cells.

<400> 97
gcggccgctg tcgggaaccg cgttcaggtg gaacaggacc acctcccttg cacttcttgg 60

tatatcagta taggctgatg tattcatagt ggggtttttc ataataaatt tactaacggc 120

aggcaacatt cactcggctt aaacgcaaaa cggaccgtct tgatatcttc tgacgcattg 180

accaccgaga aatagtgtta gttaccgggt gagttattgt tcttctacac aggcgacgcc 240

catcgtctag agttgatgta ctaactcaga tttcactacc taccctatcc ctggtacgca 300

caaagcactt tgctagaagg tataacttcg tatagcatac attatacgaa gttattctga 360

attccgagaa acacaacaac atgccccatt ggacagatca tgcggataca caggttgtgc 420

agtatcatac atactcgatc agacaggtcg tctgaccatc atacaagctg aacaagcgct 480

ccatacttgc acgctctcta tatacacagt taaattacat atccatagtc taacctctaa 540

cagttaatct tctggtaagc ctcccagcca gccttctggt atcgcttggc ctcctcgata 600

ggatctcggt tctggccgta cagacctcgg ccgacaatta tgatatccgt tccggtagac 660

atgacatcct caacagttcg gtactgctgt ccgagagcgt ctcccttgtc gtcaagaccc 720

accccggggg tcagaataag ccagtcctca gagtcgccct taggtcggtt ctgggcaacg 780

aagccaacca caaactcggg gtcggatcgg gcaagctcaa tggtctgctt ggagtactcg 840

ccagtggcca gagagccctt gcaagacagc tcggccagca tgagcagacc tctggccagc 900

ttctcgttgg gagaggggac caggaactcc ttgtactggg agttctcgta gtcagagacg 960

tcctccttct tctgttcaga gacagtttcc tcggcaccag ctcgcaggcc agcaatgatt 1020

ccggttccgg gtacaccgtg ggcgttggtg atatcggacc actcggcgat tcggtagaca 1080

ccgttcttgt actggtgctt gacagtgttg ccaatatctg cgaactttct gtcctcgaac 1140

aggaagaaac cgtgcttaag agcaagttcc ttgaggggga gcacagtgct ggcgtaggtg 1200

aagtcgtcaa tgatgtcgat atgggtcttg atcatgcaca cataaggtcc gaccttatcg 1260

gcaagctcaa tgagctcctt ggtggtggta acatccagag aagcacacag gttggttttc 1320

ttggctgcca cgagcttgag cactcgagcg gcaaaggcgg acttgtggac gttagctcga 1380

gcttcgtagg agggcatttt ggtggtgaag aggagactga aataaattta gtctgcagaa 1440

ctttttatcg gaaccttatc tggggcagtg aagtatatgt tatggtaata gttacgagtt 1500

agttgaactt atagatagac tggactatac ggctatcggt ccaaattaga aagaacgtca 1560

atggctctct gggcgcgaat tcgtataact tcgtatagca ggagttatcc gaagcgaacg 1620

gcgatacgcg tatcgatacg cgtgcatgct gaggtgtctc acaagtgccg tgcagtcccg 1680

cccccacttg cttctctttg tgtgtagtgt acgtacatta tcgagagggt tgttcccgcc 1740

cacctcgatc cggcatgctg aggtgtctca caagtgccgt gcagtcccgc ccccacttgc 1800

ttctctttgt gtgtagtgta cgtacattat cgagagggtt gttcccgccc acctcgatcc 1860

ggcatgcact gatcacgggc aaaagtgcgt tcgatagaga gggggttggc ggcgcatttg 1920

tgtcccaaaa aacagcccca attgccccaa ttgaccccaa attgacccag tagcggaccc 1980

aaccccggcg agagcccccc tcaccccaca tatcaaacct cccccggttc ccacacttgc 2040

cgttaagggc gtagggtact gcagtctgga atctacgctt gttcagactt tgtactagtt 2100

tctttgtctg gccatccggg taacccatgc cggacgcaaa atagactact gaaaattttt 2160

ttgctttgtg gttgggactt tagccaaggg tataaaagac caccgtcccc gaattacctt 2220

tcctcttctt ttctctctct ccttgtcaac tcacacccga aggatccaat gatgttcaaa 2280

atttcgtcgt cactctcacc tgttgaccaa attcccccaa ttagcccttt gcccagacca 2340

cgaccccgtc caattacacc aagagttctg gcagtgtcca gctccagcgg caaaatcgtt 2400

aataatccct tgaaagccga aaccaccgaa gcggtcagtg gcgaactggc ccgcagattt 2460

agattgggtc gactggctga agatggattt agctataaag agaaattcat cgtcagatgt 2520

tatgaagttg gtattaataa aactgcgaca gtcgaaacct tggctaattt gctgcaagaa 2580

gtgggtggaa atcatgccca atcggttggc tttagcactg atggcttcgc taccactcat 2640

tccatgcgga aaatgcattt gatttgggtg accgctcgga tgcatataga aatctataaa 2700

tatcctgcct ggtcggacgt cattgaagtc gaaacttgga tcggagccga aggtcgtatt 2760

ggtaccagac gaaattggat tattaaagac tgcgcgacag atgaagttat tggtcgggct 2820

acaagcaaat gggtcatgat gaatcaagac acacggcgat tggaaaaggt ttcggatgat 2880

gttcgtgaag aacatttggt cttttccccc agagaaccgc gcctgccgtt ccccgacgaa 2940

aataactctt cattgaagaa aatctcgaaa ttagatgatc cggctcaata ttccaaattg 3000

agtttggaac cacgacgggg agacttagat atgaatcaac atgttaacaa tgttacttat 3060

ataggctggg ttctcgagag tatgccacaa gagatcattg acacacatga attgcagaca 3120

ataactttgg actaccggcg agaatgtcaa cacgatgatt tagtggactc gctgacctct 3180

cctgagcctt ctgaattttc tgaaactact aatggtagtg ccaatgtcag ccccaatgat 3240

aaccgctgtt taaacttctt gcatttgctg agacttagct ccgacggctc tgaaattaac 3300

cggggtcgta ctgtgtggag aaagaaaccc gcaaaataat ctagtgtctg tggtatctaa 3360

gctatttatc actctttaca acttctacct caactatcta ctttaataaa tgaatatcgt 3420

ttattctcta tgattactgt atatgcgttc ctctaagaca aatcggcttc gatacgcgta 3480

tcgatacgcg tgcatgctga ggtgtctcac aagtgccgtg cagtcccgcc cccacttgct 3540

tctctttgtg tgtagtgtac gtacattatc gagagggttg ttcccgccca cctcgatccg 3600

gcatgctgag gtgtctcaca agtgccgtgc agtcccgccc ccacttgctt ctctttgtgt 3660

gtagtgtacg tacattatcg agagggttgt tcccgcccac ctcgatccgg catgctgagg 3720

tgtctcacaa gtgccgtgca gtcccgcccc cacttgcttc tctttgtgtg tagtgtacgt 3780

acattatcga gagggttgtt cccgcccacc tcgatccggc atgctgaggt gtctcacaag 3840

tgccgtgcag tcccgccccc acttgcttct ctttgtgtgt agtgtacgta cattatcgag 3900

agggttgttc ccgcccacct cgatccggca tgcactgatc acgggcaaaa gtgcgttcga 3960

tagagagggg gttggcggcg catttgtgtc ccaaaaaaca gccccaattg ccccaattga 4020

ccccaaattg acccagtagc ggacccaacc ccggcgagag ccccccttcac cccacatatc 4080

aaacctcccc cggttcccac acttgccgtt aagggcgtag ggtactgcag tctggaatct 4140

acgcttgttc agactttgta ctagtttctt tgtctggcca tccgggtaac ccatgccgga 4200

cgcaaaatag actactgaaa attttttttgc tttgtggttg ggactttagc caagggtata 4260

aaagaccacc gtccccgaat tacctttcct cttcttttct ctctctcctt gtcaactcac 4320

acccgaagga tccacaacac aatgtctatc ggatcttcta accccgtgct gctggccgct 4380

attcccttcg tgtacctgtt cgtgctgccc cgagtgctgg ccttcctgcc tcagaaggcc 4440

cagttcctgg ccaagtgcat cgtggtgctg atcgctaccc tgatcatgtc tgtggccggc 4500

tgcttcatct ctatcgtgtg cgccctgctg gacaagcgat acgtgatcaa ctacgtggtg 4560

tctcgactgt tctctttcct ggccgctcga ccctgcggcg tgacctacaa gatcgtcggc 4620

gaggaacacc tggacaagta ccccgccatt gtggtgtgca accaccagtc ctctatggac 4680

atgatggtgc tgggccgagt gttccccaag cactgcgtgg tgatggccaa gaaggaactg 4740

ctgtacttcc ccttcctggg catgttcatg aagctgtcta acgccatctt catcgaccga 4800

aagaaccaca agaaggccat cgagtctacc actcaggccg tggccgacat gaagaagcac 4860

aactctggca tctggatctt ccccgagggc acccgatctc gactggacaa ggccgacctg 4920

ctgcctttca agaagggcgc cttccacctg gccatccagg ctcagctgcc cattctgccc 4980

atcatctctc agggctactc tcacatctac gactcttcta agcgatactt ccccggtggc 5040

gagctggaaa tccgagtgct cgagcccatt cctaccaccg gcctgaccac cgacgacgtg 5100

aacgacctga tggacaagac ccgaaacctg atgctgaagc acctgaagga aatggactct 5160

cagtactctt cgtctaccgc cgagaacggc tctacccaca tcgacgccga cattgccaag 5220

tccaccgcca cctctatcgg caacactgac gacgccatta ccaagcgacg aacccctaag 5280

gaatagtaag gattgtctgt ggtatctaag ctatttatca ctctttacaa cttctacctc 5340

aactatctac tttaataaat gaatatcgtt tattctctat gattactgta tatgcgttcc 5400

tctaagacaa atcggtcacg atacgcgtat cgatacgcgt gcatgctgag gtgtctcaca 5460

agtgccgtgc agtcccgccc ccacttgctt ctctttgtgt gtagtgtacg tacattatcg 5520

agagggttgt tcccgcccac ctcgatccgg catgctgagg tgtctcacaa gtgccgtgca 5580

gtcccgcccc cacttgcttc tctttgtgtg tagtgtacgt acattatcga gagggttgtt 5640

cccgcccacc tcgatccggc atgctgaggt gtctcacaag tgccgtgcag tcccgccccc 5700

acttgcttct ctttgtgtgt agtgtacgta cattatcgag agggttgttc ccgcccacct 5760

cgatccggca tgctgaggtg tctcacaagt gccgtgcagt cccgccccca cttgcttctc 5820

tttgtgtgta gtgtacgtac attatcgaga gggttgttcc cgcccacctc gatccggcat 5880

gcactgatca cgggcaaaag tgcgttcgat agagaggggg ttggcggcgc atttgtgtcc 5940

caaaaaacag ccccaattgc cccaattgac cccaaattga cccagtagcg gacccaaccc 6000

cggcgagagc ccccttcacc ccacatatca aacctccccc ggttcccaca cttgccgtta 6060

agggcgtagg gtactgcagt ctggaatcta cgcttgttca gactttgtac tagtttcttt 6120

gtctggccat ccgggtaacc catgccggac gcaaaataga ctactgaaaa ttttttgct 6180

ttgtggttgg gactttagcc aagggtataa aagaccaccg tccccgaatt acctttcctc 6240

ttcttttctc tctctccttg tcaactcaca cccgaaggat ccccaccaca atgtctatct 6300

ctaactctcc cgagaacctg gaccagaagc ctaacatccg aaacgagtct gaggaccgat 6360

ctttctctgt gcgacgacga actaactcta acgtggtgaa cgctgtgccc gactcttgct 6420

ctaaggtggg caccatcgag actgtgacca agttctcttc tgtggactct gccgccgaga 6480

tgtcctctgg cgaggaccga aaccgaatcg gcaacggtga cgaccgaaag gaagaggcca 6540

actctctgtc taacggccag gtgtacaacg agcgagtcgg catcggcgag ggacgatctc 6600

ccgaggccga ggtgcccttc cgattctctt accgagcctc tgctcccgct caccgacgag 6660

tgaaggaatc tcccctgtcc tctgacgcca tcttccgaca gtctcacgcc ggcctgttca 6720

acctgtgcat cctggtgctg gtggccgtga acggccgact gatcattgag aacctgatga 6780

agtacggcct gctgatccga gccggattct ggttctcttc gaagtctctg cgagactggc 6840

ccctgctgat gtgctgcctg tctctgccta tcttccctct gggcgccttc accgtcgaga 6900

agtctgccaa gcacaagttc atccccgagc ctgtggtgat gtctctgcac gtgctgatct 6960

gtaccgcctc tctgctgtac cccatcttcg tgatcctccg atgcggatct gccgtgctgt 7020

tcggaatcgc cctgatgctg ctggcctcta tcgtgtggct gaagctggtg tcttacgccc 7080

acaccaactg cgacatgcga gccctggcca agctgaccga caaggtggcc tctaagtcct 7140

ccgacatgga ctactcttac gacgtgtctt tcgagtctct ggcctacttc atggtggctc 7200

ccactctgtg ctaccaggct tcttaccctc gatctacctg catccgaaag ggctgggtga 7260

tccgacagct ggtcaagctg gtgattttca ccggcgtgat gggcttcatc gtcgagcagt 7320

acatcaaccc catcgtgcag cactctcagc accctctgaa gggcaacctg ctgtacgcta 7380

tcgagcgagt gctgaagctc tctgtcccta ctctgtacgt ctggctgtgc atgttctact 7440

gtttcttcca cctgtggctc aacatcctgg ccgagctgct gcgattcggc gaccgagagt 7500

tctacaagga ctggtggaac gccaagaccg tggaagagta ctggcgaatg tggaacatgc 7560

ccgtgcacaa gtggatggtg cgacacatct acttcccttg cctgcgaaac ggactgcccc 7620

gagaggtgac cctggtgatc gtgttcttcc tgtctgccat tttccacgag ctgtgcatcg 7680

gcgtgccctg ccacatcttc aagttctggg ccttcatcgg catcatgttc caggtgcctc 7740

tggtgctgat caccaactac ctgcagaaca agttccgatc ttctatggtg ggcaacatgg 7800

tgttctggtt tttcttctcg atcctgggac agcccatgtc tgtgctgctg tactaccacg 7860

acgtgatgaa cagaaaggaa aacggcgttt ctaagtaggt atgtgtctgt ggtatctaag 7920

ctatttatca ctctttacaa cttctacctc aactatctac tttaataaat gaatatcgtt 7980

tattctctat gattactgta tatgcgttcc tctaagacag agtcatgtgt aacactcgct 8040

ctggagagtt agtcatccga cagggtaact ctaatctccc aacaccttat taactctgcg 8100

taactgtaac tcttcttgcc acgtcgatct tactcaattt tcctgctcat catctgctgg 8160

attgttgtct atcgtctggc tctaatacat ttattgttta ttgcccaaac aactttcatt 8220

gcacgtaagt gaattgtttt ataacagcgt tcgccaattg ctgcgccatc gtcgtccggc 8280

tgtcctaccg ttaggggtag tgtgtctcac actaccgagg ttactagagt tgggaaagcg 8340

atactgcctc ggacacacca cctgggtctt acgactgcag agagaatcgg cgttacctct 8400

ctcacaaagc ccttcagtgc ggccgc 8426

<210> 98
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 001

<400> 98
tgctggcgta ggtgaagtcg 20

<210> 99
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 002

<400> 99
gatcagtgca tgccggatcg 20

<210> 100
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 085

<400> 100
gggggtctct acaaatgcga gactacgagc atgtt 35

<210> 101
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 086

<400> 101
cccggtctct atccctatcc caaggcttgc tgttt 35

<210> 102
<211> 37
<212> DNA
<213> Artificial Sequence

<220>

<223> Nucleotide sequence of artificial oligonucleotide primer at 274

<400> 102
gggggtctct aatgcacaat gtctatctct aactctc 37

<210> 103
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 275

<400> 103
cccggtctct tagactactt agaaacgccg ttttcc 36

<210> 104
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 297

<400> 104
gggggtctct acaacacaat gtctatctct aactctc 37

<210> 105
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 298

<400> 105
cccggtctct atccctactt agaaacgccg ttttcc 36

<210> 106
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 299

<400> 106
gggggtctct ccaccacaat gactatcgac tcacaatac 39

<210> 107
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of artificial oligonucleotide primer at 300

<400> 107
cccggtctct atacttactc aatcattcgg aactctg 37

<210> 108
<211> 4655
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of the NotI DNA fragment of pAT091 used to transform Y. lipolytica cells.

<400> 108
gcggccgctg tcgggaaccg cgttcaggtg gaacaggacc acctcccttg cacttcttgg 60

tatatcagta taggctgatg tattcatagt ggggtttttc ataataaatt tactaacggc 120

aggcaacatt cactcggctt aaacgcaaaa cggaccgtct tgatatcttc tgacgcattg 180

accaccgaga aatagtgtta gttaccgggt gagttattgt tcttctacac aggcgacgcc 240

catcgtctag agttgatgta ctaactcaga tttcactacc taccctatcc ctggtacgca 300

caaagcactt tgctagaagg tataacttcg tatagcatac attatacgaa gttattctga 360

attccgagaa acacaacaac atgccccatt ggacagatca tgcggataca caggttgtgc 420

agtatcatac atactcgatc agacaggtcg tctgaccatc atacaagctg aacaagcgct 480

ccatacttgc acgctctcta tatacacagt taaattacat atccatagtc taacctctaa 540

cagttaatct tctggtaagc ctcccagcca gccttctggt atcgcttggc ctcctcgata 600

ggatctcggt tctggccgta cagacctcgg ccgacaatta tgatatccgt tccggtagac 660

atgacatcct caacagttcg gtactgctgt ccgagagcgt ctcccttgtc gtcaagaccc 720

accccggggg tcagaataag ccagtcctca gagtcgccct taggtcggtt ctgggcaacg 780

aagccaacca caaactcggg gtcggatcgg gcaagctcaa tggtctgctt ggagtactcg 840

ccagtggcca gagagccctt gcaagacagc tcggccagca tgagcagacc tctggccagc 900

ttctcgttgg gagaggggac caggaactcc ttgtactggg agttctcgta gtcagagacg 960

tcctccttct tctgttcaga gacagtttcc tcggcaccag ctcgcaggcc agcaatgatt 1020

ccggttccgg gtacaccgtg ggcgttggtg atatcggacc actcggcgat tcggtagaca 1080

ccgttcttgt actggtgctt gacagtgttg ccaatatctg cgaactttct gtcctcgaac 1140

aggaagaaac cgtgcttaag agcaagttcc ttgaggggga gcacagtgct ggcgtaggtg 1200

aagtcgtcaa tgatgtcgat atgggtcttg atcatgcaca cataaggtcc gaccttatcg 1260

gcaagctcaa tgagctcctt ggtggtggta acatccagag aagcacacag gttggttttc 1320

ttggctgcca cgagcttgag cactcgagcg gcaaaggcgg acttgtggac gttagctcga 1380

gcttcgtagg agggcatttt ggtggtgaag aggagactga aataaattta gtctgcagaa 1440

ctttttatcg gaaccttatc tggggcagtg aagtatatgt tatggtaata gttacgagtt 1500

agttgaactt atagatagac tggactatac ggctatcggt ccaaattaga aagaacgtca 1560

atggctctct gggcgcgaat tcgtataact tcgtatagca ggagttatcc gaagcgaacg 1620

gcgatacgcg tatcgatacg cgtgcatgct gaggtgtctc acaagtgccg tgcagtcccg 1680

cccccacttg cttctctttg tgtgtagtgt acgtacatta tcgagagggt tgttcccgcc 1740

cacctcgatc cggcatgctg aggtgtctca caagtgccgt gcagtcccgc ccccacttgc 1800

ttctctttgt gtgtagtgta cgtacattat cgagagggtt gttcccgccc acctcgatcc 1860

ggcatgctga ggtgtctcac aagtgccgtg cagtcccgcc cccacttgct tctctttgtg 1920

tgtagtgtac gtacattatc gagagggttg ttcccgccca cctcgatccg gcatgctgag 1980

gtgtctcaca agtgccgtgc agtcccgccc ccacttgctt ctctttgtgt gtagtgtacg 2040

tacattatcg agagggttgt tcccgcccac ctcgatccgg catgcactga tcacgggcaa 2100

aagtgcgttc gatagagagg gggttggcgg cgcatttgtg tcccaaaaaa cagccccaat 2160

tgccccaatt gaccccaaat tgacccagta gcggacccaa ccccggcgag agccccccttc 2220

accccacata tcaaacctcc cccggttccc acacttgccg ttaagggcgt agggtactgc 2280

agtctggaat ctacgcttgt tcagactttg tactagtttc tttgtctggc catccgggta 2340

acccatgccg gacgcaaaat agactactga aaatttttt gctttgtggt tgggacttta 2400

gccaagggta taaaagacca ccgtccccga attacctttc ctcttctttt ctctctctcc 2460

ttgtcaactc acacccgaag gatccaatga tgcgagacta cgagcatgtt gtggaaagtt 2520

ttgactactt acaatccgca actcaagacc ttcacggaga gttgaccgcc ctgaatgctt 2580

gtgttgagtg ttgtgataga cacgcgcatg gtgaagctgt tgccctttac tgtgaggccc 2640

aagacggtca cgcggaacga tatcggttcc gagatttgca gcgtcaagcc gcacgatttg 2700

gtaactttct tcgagagcag ggagtgaaac ctggtgaccg agtggctggt ttgatgcccc 2760

gtaccgtcga gttgcttatc gctatcctgg gaacatggcg gattggagct gtgtatcagc 2820

ccttgttcac cgcgtttggt ccgaaagcta ttgagcagcg gctgaattgt agcaatgctc 2880

gatggattgt cactgatccc cataatcgac caaaattgga tgatgttacc gattgtccga 2940

gtatagtcgt cactggcggt gcaccccaaa atccggcaga ccatcacttt tggtccgcct 3000

tgaatcgaca agccgatgac tgtgctcctg tcttattaga cgcttccgct cccttcttgc 3060

tcatgtgcac ttcgggtact accggccctg cgaagccttt ggaagtcccc ctctcggcca 3120

tcttagcttt taaaggttac atgcgagatg ccatcgatct tagagcagat gatcgatttt 3180

ggaatttggc cgatccagga tgggcgtacg gactgtacta tgccgtcact ggacccttgg 3240

cctgcggcta tgccaccttg ttttatgatg gaccgtttac cgttgagtcc actcgacata 3300

taatcgctaa atatgccata aataatttgg ctggatcacc gacagcttat cgattcttga 3360

ttgctgccgg cgctgaattt gctgatgcag tgcgcggtcg attgcgtgct gtgtctagtg 3420

cgggcgagcc cttgaatccc caggtggttc gatggttcgc tgagcaactg ggagtcgtca 3480

tccatgatca ttatggtcag acagagatag gtatggtcct ttgtaaccat catggacttc 3540

gacatcccgt gcgtgagggt tcggcaggat acgcagttcc tggttatcgt attgtggtcc 3600

tcgacaaagc tcatcgagag ctgccggcag gtcaacctgg agttttggca gtcgaccgag 3660

aacgatcccc tctgtgttgg tttgatggtt atttgggaat gcctacacaa gcgttcgccg 3720

gacgatacta tttgtccgga gacatcgtcg aactgaacga tgatggttct atttcttttg 3780

ttggtcgaaa cgacgatttg atcactacct ctggatatcg agtgggtcca tttgacgtcg 3840

agtccgcctt gattgagcat cccgctgtcg ttgaagctgc agttatagga aagccagacc 3900

ctcagcggac agaactgatt aaagccttcg ttgtcttgaa tactccatat ctcccttcgc 3960

cagaattagc cgaagagttg cgtttgcacg tgcgacaacg gttggctgcc cacgcatacc 4020

cccgagagat ggaattcgtt gatcacttac cgaaaacacc ttcgggaaaa ctgcagcgat 4080

ttatcttgag aaatcaagaa atcgcgaaac agcaagcctt gggatagtct agtgtctgtg 4140

gtatctaagc tatttatcac tctttacaac ttctacctca actatctact ttaataaatg 4200

aatatcgttt attctctatg attactgtat atgcgttcct ctaagacaga gtcatgtgta 4260

acactcgctc tggagagtta gtcatccgac agggtaactc taatctccca acaccttatt 4320

aactctgcgt aactgtaact cttcttgcca cgtcgatctt actcaatttt cctgctcatc 4380

atctgctgga ttgttgtcta tcgtctggct ctaatacatt tattgtttat tgcccaaaca 4440

actttcattg cacgtaagtg aattgtttta taacagcgtt cgccaattgc tgcgccatcg 4500

tcgtccggct gtcctaccgt taggggtagt gtgtctcaca ctaccgaggt tactagagtt 4560

gggaaagcga tactgcctcg gacacaccac ctgggtctta cgactgcaga gagaatcggc 4620

gttacctctc tcacaaagcc cttcagtgcg gccgc 4655

<210> 109
<211> 4085
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of the NotI DNA fragment of pAT108 used to transform Y. lipolytica cells.

<400> 109
gcggccgctg tcgggaaccg cgttcaggtg gaacaggacc acctcccttg cacttcttgg 60

tatatcagta taggctgatg tattcatagt ggggtttttc ataataaatt tactaacggc 120

aggcaacatt cactcggctt aaacgcaaaa cggaccgtct tgatatcttc tgacgcattg 180

accaccgaga aatagtgtta gttaccgggt gagttattgt tcttctacac aggcgacgcc 240

catcgtctag agttgatgta ctaactcaga tttcactacc taccctatcc ctggtacgca 300

caaagcactt tgctagaagg tataacttcg tatagcatac attatacgaa gttattctga 360

attccgagaa acacaacaac atgccccatt ggacagatca tgcggataca caggttgtgc 420

agtatcatac atactcgatc agacaggtcg tctgaccatc atacaagctg aacaagcgct 480

ccatacttgc acgctctcta tatacacagt taaattacat atccatagtc taacctctaa 540

cagttaatct tctggtaagc ctcccagcca gccttctggt atcgcttggc ctcctcgata 600

ggatctcggt tctggccgta cagacctcgg ccgacaatta tgatatccgt tccggtagac 660

atgacatcct caacagttcg gtactgctgt ccgagagcgt ctcccttgtc gtcaagaccc 720

accccggggg tcagaataag ccagtcctca gagtcgccct taggtcggtt ctgggcaacg 780

aagccaacca caaactcggg gtcggatcgg gcaagctcaa tggtctgctt ggagtactcg 840

ccagtggcca gagagccctt gcaagacagc tcggccagca tgagcagacc tctggccagc 900

ttctcgttgg gagaggggac caggaactcc ttgtactggg agttctcgta gtcagagacg 960

tcctccttct tctgttcaga gacagtttcc tcggcaccag ctcgcaggcc agcaatgatt 1020

ccggttccgg gtacaccgtg ggcgttggtg atatcggacc actcggcgat tcggtagaca 1080

ccgttcttgt actggtgctt gacagtgttg ccaatatctg cgaactttct gtcctcgaac 1140

aggaagaaac cgtgcttaag agcaagttcc ttgaggggga gcacagtgct ggcgtaggtg 1200

aagtcgtcaa tgatgtcgat atgggtcttg atcatgcaca cataaggtcc gaccttatcg 1260

gcaagctcaa tgagctcctt ggtggtggta acatccagag aagcacacag gttggttttc 1320

ttggctgcca cgagcttgag cactcgagcg gcaaaggcgg acttgtggac gttagctcga 1380

gcttcgtagg agggcatttt ggtggtgaag aggagactga aataaattta gtctgcagaa 1440

ctttttatcg gaaccttatc tggggcagtg aagtatatgt tatggtaata gttacgagtt 1500

agttgaactt atagatagac tggactatac ggctatcggt ccaaattaga aagaacgtca 1560

atggctctct gggcgcgaat tcgtataact tcgtatagca ggagttatcc gaagcgaacg 1620

gcgatacgcg tatcgatacg cgtgcatgct gaggtgtctc acaagtgccg tgcagtcccg 1680

cccccacttg cttctctttg tgtgtagtgt acgtacatta tcgagagggt tgttcccgcc 1740

cacctcgatc cggcatgctg aggtgtctca caagtgccgt gcagtcccgc ccccacttgc 1800

ttctctttgt gtgtagtgta cgtacattat cgagagggtt gttcccgccc acctcgatcc 1860

ggcatgctga ggtgtctcac aagtgccgtg cagtcccgcc cccacttgct tctctttgtg 1920

tgtagtgtac gtacattatc gagagggttg ttcccgccca cctcgatccg gcatgctgag 1980

gtgtctcaca agtgccgtgc agtcccgccc ccacttgctt ctctttgtgt gtagtgtacg 2040

tacattatcg agagggttgt tcccgcccac ctcgatccgg catgcactga tcacgggcaa 2100

aagtgcgttc gatagagagg gggttggcgg cgcatttgtg tcccaaaaaa cagccccaat 2160

tgccccaatt gaccccaaat tgacccagta gcggacccaa ccccggcgag agccccccttc 2220

accccacata tcaaacctcc cccggttccc acacttgccg ttaagggcgt agggtactgc 2280

agtctggaat ctacgcttgt tcagactttg tactagtttc tttgtctggc catccgggta 2340

acccatgccg gacgcaaaat agactactga aaatttttttt gctttgtggt tgggacttta 2400

gccaagggta taaaagacca ccgtccccga attacctttc ctcttctttt ctctctctcc 2460

ttgtcaactc acacccgaag gatccaatga tgttcaaaat ttcgtcgtca ctctcacctg 2520

ttgaccaaat tcccccaatt agccctttgc ccagaccacg accccgtcca attacaccaa 2580

gagttctggc agtgtccagc tccagcggca aaatcgttaa taatcccttg aaagccgaaa 2640

ccaccgaagc ggtcagtggc gaactggccc gcagatttag attgggtcga ctggctgaag 2700

atggatttag ctataaagag aaattcatcg tcagatgtta tgaagttggt attaataaaa 2760

ctgcgacagt cgaaaccttg gctaatttgc tgcaagaagt gggtggaaat catgcccaat 2820

cggttggctt tagcactgat ggcttcgcta ccactcattc catgcggaaa atgcatttga 2880

tttgggtgac cgctcggatg catatagaaa tctataaata tcctgcctgg tcggacgtca 2940

ttgaagtcga aacttggatc ggagccgaag gtcgtattgg taccagacga aattggatta 3000

ttaaagactg cgcgacagat gaagttattg gtcgggctac aagcaaatgg gtcatgatga 3060

atcaagacac acggcgattg gaaaaggttt cggatgatgt tcgtgaagaa catttggtct 3120

tttcccccag agaaccgcgc ctgccgttcc ccgacgaaaa taactcttca ttgaagaaaa 3180

tctcgaaatt agatgatccg gctcaatatt ccaaattgag tttggaacca cgacggggag 3240

acttagatat gaatcaacat gttaacaatg ttacttatat aggctgggtt ctcgagagta 3300

tgccacaaga gatcattgac acacatgaat tgcagacaat aactttggac taccggcgag 3360

aatgtcaaca cgatgattta gtggactcgc tgacctctcc tgagccttct gaattttctg 3420

aaactactaa tggtagtgcc aatgtcagcc ccaatgataa ccgctgttta aacttcttgc 3480

atttgctgag acttagctcc gacggctctg aaattaaccg gggtcgtact gtgtggagaa 3540

agaaacccgc aaaataatct agtgtctgtg gtatctaagc tatttatcac tctttacaac 3600

ttctacctca actatctact ttaataaatg aatatcgttt attctctatg attactgtat 3660

atgcgttcct ctaagacaga gtcatgtgta acactcgctc tggagagtta gtcatccgac 3720

agggtaactc taatctccca acaccttatt aactctgcgt aactgtaact cttcttgcca 3780

cgtcgatctt actcaatttt cctgctcatc atctgctgga ttgttgtcta tcgtctggct 3840

ctaatacatt tattgtttat tgcccaaaca actttcattg cacgtaagtg aattgtttta 3900

taacagcgtt cgccaattgc tgcgccatcg tcgtccggct gtcctaccgt taggggtagt 3960

gtgtctcaca ctaccgaggt tactagagtt gggaaagcga tactgcctcg gacacaccac 4020

ctgggtctta cgactgcaga gagaatcggc gttacctctc tcacaaagcc cttcagtgcg 4080

gccgc 4085

<210> 110
<211> 4629
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of the NotI DNA fragment of pAT135 used to transform Y. lipolytica cells.

<400> 110
gcggccgctg tcgggaaccg cgttcaggtg gaacaggacc acctcccttg cacttcttgg 60

tatatcagta taggctgatg tattcatagt ggggtttttc ataataaatt tactaacggc 120

aggcaacatt cactcggctt aaacgcaaaa cggaccgtct tgatatcttc tgacgcattg 180

accaccgaga aatagtgtta gttaccgggt gagttattgt tcttctacac aggcgacgcc 240

catcgtctag agttgatgta ctaactcaga tttcactacc taccctatcc ctggtacgca 300

caaagcactt tgctagaagg tataacttcg tatagcatac attatacgaa gttattctga 360

attccgagaa acacaacaac atgccccatt ggacagatca tgcggataca caggttgtgc 420

agtatcatac atactcgatc agacaggtcg tctgaccatc atacaagctg aacaagcgct 480

ccatacttgc acgctctcta tatacacagt taaattacat atccatagtc taacctctaa 540

cagttaatct tctggtaagc ctcccagcca gccttctggt atcgcttggc ctcctcgata 600

ggatctcggt tctggccgta cagacctcgg ccgacaatta tgatatccgt tccggtagac 660

atgacatcct caacagttcg gtactgctgt ccgagagcgt ctcccttgtc gtcaagaccc 720

accccggggg tcagaataag ccagtcctca gagtcgccct taggtcggtt ctgggcaacg 780

aagccaacca caaactcggg gtcggatcgg gcaagctcaa tggtctgctt ggagtactcg 840

ccagtggcca gagagccctt gcaagacagc tcggccagca tgagcagacc tctggccagc 900

ttctcgttgg gagaggggac caggaactcc ttgtactggg agttctcgta gtcagagacg 960

tcctccttct tctgttcaga gacagtttcc tcggcaccag ctcgcaggcc agcaatgatt 1020

ccggttccgg gtacaccgtg ggcgttggtg atatcggacc actcggcgat tcggtagaca 1080

ccgttcttgt actggtgctt gacagtgttg ccaatatctg cgaactttct gtcctcgaac 1140

aggaagaaac cgtgcttaag agcaagttcc ttgaggggga gcacagtgct ggcgtaggtg 1200

aagtcgtcaa tgatgtcgat atgggtcttg atcatgcaca cataaggtcc gaccttatcg 1260

gcaagctcaa tgagctcctt ggtggtggta acatccagag aagcacacag gttggttttc 1320

ttggctgcca cgagcttgag cactcgagcg gcaaaggcgg acttgtggac gttagctcga 1380

gcttcgtagg agggcatttt ggtggtgaag aggagactga aataaattta gtctgcagaa 1440

ctttttatcg gaaccttatc tggggcagtg aagtatatgt tatggtaata gttacgagtt 1500

agttgaactt atagatagac tggactatac ggctatcggt ccaaattaga aagaacgtca 1560

atggctctct gggcgcgaat tcgtataact tcgtatagca ggagttatcc gaagcgaacg 1620

gcgatacgcg tatcgatacg cgtgcatgct gaggtgtctc acaagtgccg tgcagtcccg 1680

cccccacttg cttctctttg tgtgtagtgt acgtacatta tcgagagggt tgttcccgcc 1740

```
cacctcgatc cggcatgctg aggtgtctca caagtgccgt gcagtcccgc ccccacttgc     1800

ttctctttgt gtgtagtgta cgtacattat cgagagggtt gttcccgccc acctcgatcc     1860

ggcatgctga ggtgtctcac aagtgccgtg cagtcccgcc cccacttgct tctctttgtg     1920

tgtagtgtac gtacattatc gagagggttg ttcccgccca cctcgatccg gcatgctgag     1980

gtgtctcaca agtgccgtgc agtcccgccc ccacttgctt ctctttgtgt gtagtgtacg     2040

tacattatcg agagggttgt tcccgcccac ctcgatccgg catgcactga tcacgggcaa     2100

aagtgcgttc gatagagagg gggttggcgg cgcatttgtg tcccaaaaaa cagccccaat     2160

tgccccaatt gaccccaaat tgacccagta gcggacccaa ccccggcgag agccccccttc     2220

accccacata tcaaacctcc cccggttccc acacttgccg ttaagggcgt agggtactgc     2280

agtctggaat ctacgcttgt tcagactttg tactagtttc tttgtctggc catccgggta     2340

acccatgccg gacgcaaaat agactactga aaatttttttt gctttgtggt tgggacttta     2400

gccaagggta taaaagacca ccgtccccga attacctttc ctcttctttt ctctctctcc     2460

ttgtcaactc acacccgaag gatccaatgc acaatgtcta tctctaactc tcccgagaac     2520

ctggaccaga agcctaacat ccgaaacgag tctgaggacc gatctttctc tgtgcgacga     2580

cgaactaact ctaacgtggt gaacgctgtg cccgactctt gctctaaggt gggcaccatc     2640

gagactgtga ccaagttctc ttctgtggac tctgccgccg agatgtcctc tggcgaggac     2700

cgaaaccgaa tcggcaacgg tgacgaccga aaggaagagg ccaactctct gtctaacggc     2760

caggtgtaca acgagcgagt cggcatcggc gagggacgat ctcccgaggc cgaggtgccc     2820

ttccgattct cttaccgagc ctctgctccc gctcaccgac gagtgaagga atctcccctg     2880

tcctctgacg ccatcttccg acagtctcac gccggcctgt tcaacctgtg catcctggtg     2940

ctggtggccg tgaacggccg actgatcatt gagaacctga tgaagtacgg cctgctgatc     3000

cgagccggat tctggttctc ttcgaagtct ctgcgagact ggcccctgct gatgtgctgc     3060

ctgtctctgc ctatcttccc tctgggcgcc ttcaccgtcg agaagtctgc caagcacaag     3120

ttcatccccg agcctgtggt gatgtctctg cacgtgctga tctgtaccgc ctctctgctg     3180

taccccatct tcgtgatcct ccgatgcgga tctgccgtgc tgttcggaat cgccctgatg     3240
```

ctgctggcct ctatcgtgtg gctgaagctg gtgtcttacg cccacaccaa ctgcgacatg 3300

cgagccctgg ccaagctgac cgacaaggtg gcctctaagt cctccgacat ggactactct 3360

tacgacgtgt ctttcgagtc tctggcctac ttcatggtgg ctcccactct gtgctaccag 3420

gcttcttacc ctcgatctac ctgcatccga aagggctggg tgatccgaca gctggtcaag 3480

ctggtgattt tcaccggcgt gatgggcttc atcgtcgagc agtacatcaa ccccatcgtg 3540

cagcactctc agcaccctct gaagggcaac ctgctgtacg ctatcgagcg agtgctgaag 3600

ctctctgtcc ctactctgta cgtctggctg tgcatgttct actgtttctt ccacctgtgg 3660

ctcaacatcc tggccgagct gctgcgattc ggcgaccgag agttctacaa ggactggtgg 3720

aacgccaaga ccgtggaaga gtactggcga atgtggaaca tgcccgtgca caagtggatg 3780

gtgcgacaca tctacttccc ttgcctgcga aacggactgc cccgagaggt gaccctggtg 3840

atcgtgttct tcctgtctgc cattttccac gagctgtgca tcggcgtgcc ctgccacatc 3900

ttcaagttct gggccttcat cggcatcatg ttccaggtgc ctctggtgct gatcaccaac 3960

tacctgcaga acaagttccg atcttctatg gtgggcaaca tggtgttctg gtttttcttc 4020

tcgatcctgg gacagcccat gtctgtgctg ctgtactacc acgacgtgat gaacagaaag 4080

gaaaacggcg tttctaagta gtctagtgtc tgtggtatct aagctattta tcactcttta 4140

caacttctac ctcaactatc tactttaata aatgaatatc gtttattctc tatgattact 4200

gtatatgcgt tcctctaaga cagagtcatg tgtaacactc gctctggaga gttagtcatc 4260

cgacagggta actctaatct cccaacacct tattaactct gcgtaactgt aactcttctt 4320

gccacgtcga tcttactcaa ttttcctgct catcatctgc tggattgttg tctatcgtct 4380

ggctctaata catttattgt ttattgccca aacaactttc attgcacgta agtgaattgt 4440

tttataacag cgttcgccaa ttgctgcgcc atcgtcgtcc ggctgtccta ccgttagggg 4500

tagtgtgtct cacactaccg aggttactag agttgggaaa gcgatactgc ctcggacaca 4560

ccacctgggt cttacgactg cagagagaat cggcgttacc tctctcacaa agcccttcag 4620

tgcggccgc 4629

<210> 111

<211> 6721
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of the NotI DNA fragment of pAT213 used to transform Y. lipolytica cells.

<400> 111
gcggccgctg tcgggaaccg cgttcaggtg gaacaggacc acctcccttg cacttcttgg 60

tatatcagta taggctgatg tattcatagt ggggtttttc ataataaatt tactaacggc 120

aggcaacatt cactcggctt aaacgcaaaa cggaccgtct tgatatcttc tgacgcattg 180

accaccgaga aatagtgtta gttaccgggt gagttattgt tcttctacac aggcgacgcc 240

catcgtctag agttgatgta ctaactcaga tttcactacc taccctatcc ctggtacgca 300

caaagcactt tgctagaagg tataacttcg tatagcatac attatacgaa gttattctga 360

attccgagaa acacaacaac atgccccatt ggacagatca tgcggataca caggttgtgc 420

agtatcatac atactcgatc agacaggtcg tctgaccatc atacaagctg aacaagcgct 480

ccatacttgc acgctctcta tatacacagt taaattacat atccatagtc taacctctaa 540

cagttaatct tctggtaagc ctcccagcca gccttctggt atcgcttggc ctcctcgata 600

ggatctcggt tctggccgta cagacctcgg ccgacaatta tgatatccgt tccggtagac 660

atgacatcct caacagttcg gtactgctgt ccgagagcgt ctcccttgtc gtcaagaccc 720

accccggggg tcagaataag ccagtcctca gagtcgccct taggtcggtt ctgggcaacg 780

aagccaacca caaactcggg gtcggatcgg gcaagctcaa tggtctgctt ggagtactcg 840

ccagtggcca gagagccctt gcaagacagc tcggccagca tgagcagacc tctggccagc 900

ttctcgttgg gagaggggac caggaactcc ttgtactggg agttctcgta gtcagagacg 960

tcctccttct tctgttcaga gacagtttcc tcggcaccag ctcgcaggcc agcaatgatt 1020

ccggttccgg gtacaccgtg ggcgttggtg atatcggacc actcggcgat tcggtagaca 1080

ccgttcttgt actggtgctt gacagtgttg ccaatatctg cgaactttct gtcctcgaac 1140

aggaagaaac cgtgcttaag agcaagttcc ttgaggggga gcacagtgct ggcgtaggtg 1200

aagtcgtcaa tgatgtcgat atgggtcttg atcatgcaca cataaggtcc gaccttatcg 1260

gcaagctcaa tgagctcctt ggtggtggta acatccagag aagcacacag gttggttttc 1320

ttggctgcca cgagcttgag cactcgagcg gcaaaggcgg acttgtggac gttagctcga 1380

gcttcgtagg agggcatttt ggtggtgaag aggagactga aataaattta gtctgcagaa 1440

ctttttatcg gaaccttatc tggggcagtg aagtatatgt tatggtaata gttacgagtt 1500

agttgaactt atagatagac tggactatac ggctatcggt ccaaattaga aagaacgtca 1560

atggctctct gggcgcgaat tcgtataact tcgtatagca ggagttatcc gaagcgaacg 1620

gcgatacgcg tatcgatacg cgtgcatgct gaggtgtctc acaagtgccg tgcagtcccg 1680

cccccacttg cttctctttg tgtgtagtgt acgtacatta tcgagagggt tgttcccgcc 1740

cacctcgatc cggcatgctg aggtgtctca caagtgccgt gcagtcccgc ccccacttgc 1800

ttctctttgt gtgtagtgta cgtacattat cgagagggtt gttcccgccc acctcgatcc 1860

ggcatgctga ggtgtctcac aagtgccgtg cagtcccgcc cccacttgct tctctttgtg 1920

tgtagtgtac gtacattatc gagagggttg ttcccgccca cctcgatccg gcatgctgag 1980

gtgtctcaca agtgccgtgc agtcccgccc ccacttgctt ctctttgtgt gtagtgtacg 2040

tacattatcg agagggttgt tcccgcccac ctcgatccgg catgcactga tcacgggcaa 2100

aagtgcgttc gatagagagg gggttggcgg cgcatttgtg tcccaaaaaa cagccccaat 2160

tgccccaatt gaccccaaat tgacccagta gcggacccaa ccccggcgag agcccccttc 2220

accccacata tcaaacctcc cccggttccc acacttgccg ttaagggcgt agggtactgc 2280

agtctggaat ctacgcttgt tcagactttg tactagtttc tttgtctggc catccgggta 2340

acccatgccg gacgcaaaat agactactga aaatttttt gctttgtggt tgggacttta 2400

gccaagggta taaaagacca ccgtccccga attacctttc ctcttctttt ctctctctcc 2460

ttgtcaactc acacccgaag gatccaatga tgttcaaaat ttcgtcgtca ctctcacctg 2520

ttgaccaaat tcccccaatt agcccttttgc ccagaccacg accccgtcca attacaccaa 2580

gagttctggc agtgtccagc tccagcggca aaatcgttaa taatcccttg aaagccgaaa 2640

ccaccgaagc ggtcagtggc gaactggccc gcagatttag attgggtcga ctggctgaag 2700

atggatttag ctataaagag aaattcatcg tcagatgtta tgaagttggt attaataaaa 2760

ctgcgacagt cgaaaccttg gctaatttgc tgcaagaagt gggtggaaat catgcccaat 2820

cggttggctt tagcactgat ggcttcgcta ccactcattc catgcggaaa atgcatttga 2880

tttgggtgac cgctcggatg catatagaaa tctataaata tcctgcctgg tcggacgtca 2940

ttgaagtcga aacttggatc ggagccgaag gtcgtattgg taccagacga aattggatta 3000

ttaaagactg cgcgacagat gaagttattg gtcgggctac aagcaaatgg gtcatgatga 3060

atcaagacac acggcgattg gaaaaggttt cggatgatgt tcgtgaagaa catttggtct 3120

tttcccccag agaaccgcgc ctgccgttcc ccgacgaaaa taactcttca ttgaagaaaa 3180

tctcgaaatt agatgatccg gctcaatatt ccaaattgag tttggaacca cgacggggag 3240

acttagatat gaatcaacat gttaacaatg ttacttatat aggctgggtt ctcgagagta 3300

tgccacaaga gatcattgac acacatgaat tgcagacaat aactttggac taccggcgag 3360

aatgtcaaca cgatgattta gtggactcgc tgacctctcc tgagccttct gaattttctg 3420

aaactactaa tggtagtgcc aatgtcagcc ccaatgataa ccgctgttta aacttcttgc 3480

atttgctgag acttagctcc gacggctctg aaattaaccg gggtcgtact gtgtggagaa 3540

agaaacccgc aaaataatct agtgtctgtg gtatctaagc tatttatcac tctttacaac 3600

ttctacctca actatctact ttaataaatg aatatcgttt attctctatg attactgtat 3660

atgcgttcct ctaagacaaa tcggcttcga tacgcgtatc gatacgcgtg catgctgagg 3720

tgtctcacaa gtgccgtgca gtcccgcccc cacttgcttc tctttgtgtg tagtgtacgt 3780

acattatcga gagggttgtt cccgcccacc tcgatccggc atgctgaggt gtctcacaag 3840

tgccgtgcag tcccgccccc acttgcttct ctttgtgtgt agtgtacgta cattatcgag 3900

agggttgttc ccgcccacct cgatccggca tgctgaggtg tctcacaagt gccgtgcagt 3960

cccgccccca cttgcttctc tttgtgtgta gtgtacgtac attatcgaga gggttgttcc 4020

cgcccacctc gatccggcat gctgaggtgt ctcacaagtg ccgtgcagtc ccgcccccac 4080

ttgcttctct ttgtgtgtag tgtacgtaca ttatcgagag ggttgttccc gcccacctcg 4140

atccggcatg cactgatcac gggcaaaagt gcgttcgata gagaggggggt tggcggcgca 4200

tttgtgtccc aaaaaacagc cccaattgcc ccaattgacc ccaaattgac ccagtagcgg 4260

acccaacccc ggcgagagcc cccttcaccc cacatatcaa acctcccccg gttcccacac 4320

ttgccgttaa gggcgtaggg tactgcagtc tggaatctac gcttgttcag actttgtact 4380

agtttctttg tctggccatc cgggtaaccc atgccggacg caaaatagac tactgaaaat 4440

ttttttgctt tgtggttggg actttagcca agggtataaa agaccaccgt ccccgaatta 4500

cctttcctct tcttttctct ctctccttgt caactcacac ccgaaggatc cacaaatgcg 4560

agactacgag catgttgtgg aaagttttga ctacttacaa tccgcaactc aagaccttca 4620

cggagagttg accgccctga atgcttgtgt tgagtgttgt gatagacacg cgcatggtga 4680

agctgttgcc ctttactgtg aggcccaaga cggtcacgcg gaacgatatc ggttccgaga 4740

tttgcagcgt caagccgcac gatttggtaa ctttcttcga gagcagggag tgaaacctgg 4800

tgaccgagtg gctggtttga tgccccgtac cgtcgagttg cttatcgcta tcctgggaac 4860

atggcggatt ggagctgtgt atcagccctt gttcaccgcg tttggtccga aagctattga 4920

gcagcggctg aattgtagca atgctcgatg gattgtcact gatccccata atcgaccaaa 4980

attggatgat gttaccgatt gtccgagtat agtcgtcact ggcggtgcac cccaaaatcc 5040

ggcagaccat cacttttggt ccgccttgaa tcgacaagcc gatgactgtg ctcctgtctt 5100

attagacgct tccgctccct tcttgctcat gtgcacttcg ggtactaccg gccctgcgaa 5160

gcctttggaa gtccccctct cggccatctt agcttttaaa ggttacatgc gagatgccat 5220

cgatcttaga gcagatgatc gattttggaa tttggccgat ccaggatggg cgtacggact 5280

gtactatgcc gtcactggac ccttggcctg cggctatgcc accttgtttt atgatggacc 5340

gtttaccgtt gagtccactc gacatataat cgctaaatat gccataaata atttggctgg 5400

atcaccgaca gcttatcgat tcttgattgc tgccggcgct gaatttgctg atgcagtgcg 5460

cggtcgattg cgtgctgtgt ctagtgcggg cgagcccttg aatccccagg tggttcgatg 5520

gttcgctgag caactgggag tcgtcatcca tgatcattat ggtcagacag agataggtat 5580

ggtcctttgt aaccatcatg gacttcgaca tcccgtgcgt gagggttcgg caggatacgc 5640

agttcctggt tatcgtattg tggtcctcga caaagctcat cgagagctgc cggcaggtca 5700

acctggagtt ttggcagtcg accgagaacg atcccctctg tgttggtttg atggttattt 5760

gggaatgcct acacaagcgt tcgccggacg atactatttg tccggagaca tcgtcgaact 5820

gaacgatgat ggttctattt cttttgttgg tcgaaacgac gatttgatca ctacctctgg 5880

atatcgagtg ggtccatttg acgtcgagtc cgccttgatt gagcatcccg ctgtcgttga 5940

agctgcagtt ataggaaagc cagaccctca gcggacagaa ctgattaaag ccttcgttgt 6000

cttgaatact ccatatctcc cttcgccaga attagccgaa gagttgcgtt tgcacgtgcg 6060

acaacggttg gctgcccacg catacccccg agagatggaa ttcgttgatc acttaccgaa 6120

aacaccttcg ggaaaactgc agcgatttat cttgagaaat caagaaatcg cgaaacagca 6180

agccttggga tagggatgtg tctgtggtat ctaagctatt tatcactctt tacaacttct 6240

acctcaacta tctactttaa taaatgaata tcgtttattc tctatgatta ctgtatatgc 6300

gttcctctaa gacagagtca tgtgtaacac tcgctctgga gagttagtca tccgacaggg 6360

taactctaat ctcccaacac cttattaact ctgcgtaact gtaactcttc ttgccacgtc 6420

gatcttactc aattttcctg ctcatcatct gctggattgt tgtctatcgt ctggctctaa 6480

tacatttatt gtttattgcc caaacaactt tcattgcacg taagtgaatt gttttataac 6540

agcgttcgcc aattgctgcg ccatcgtcgt ccggctgtcc taccgttagg ggtagtgtgt 6600

ctcacactac cgaggttact agagttggga aagcgatact gcctcggaca caccacctgg 6660

gtcttacgac tgcagagaga atcggcgtta cctctctcac aaagcccttc agtgcggccg 6720

c 6721

<210> 112
<211> 6695
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of the NotI DNA fragment of pAT214 used to transform Y. lipolytica cells.

<400> 112
gcggccgctg tcgggaaccg cgttcaggtg gaacaggacc acctcccttg cacttcttgg 60

tatatcagta taggctgatg tattcatagt ggggtttttc ataataaatt tactaacggc 120

aggcaacatt cactcggctt aaacgcaaaa cggaccgtct tgatatcttc tgacgcattg 180

accaccgaga aatagtgtta gttaccgggt gagttattgt tcttctacac aggcgacgcc 240

catcgtctag agttgatgta ctaactcaga tttcactacc taccctatcc ctggtacgca 300

caaagcactt tgctagaagg tataacttcg tatagcatac attatacgaa gttattctga 360

attccgagaa acacaacaac atgccccatt ggacagatca tgcggataca caggttgtgc 420

agtatcatac atactcgatc agacaggtcg tctgaccatc atacaagctg aacaagcgct 480

ccatacttgc acgctctcta tatacacagt taaattacat atccatagtc taacctctaa 540

cagttaatct tctggtaagc ctcccagcca gccttctggt atcgcttggc ctcctcgata 600

ggatctcggt tctggccgta cagacctcgg ccgacaatta tgatatccgt tccggtagac 660

atgacatcct caacagttcg gtactgctgt ccgagagcgt ctcccttgtc gtcaagaccc 720

accccggggg tcagaataag ccagtcctca gagtcgccct taggtcggtt ctgggcaacg 780

aagccaacca caaactcggg gtcggatcgg gcaagctcaa tggtctgctt ggagtactcg 840

ccagtggcca gagagccctt gcaagacagc tcggccagca tgagcagacc tctggccagc 900

ttctcgttgg gagaggggac caggaactcc ttgtactggg agttctcgta gtcagagacg 960

tcctccttct tctgttcaga gacagtttcc tcggcaccag ctcgcaggcc agcaatgatt 1020

ccggttccgg gtacaccgtg ggcgttggtg atatcggacc actcggcgat tcggtagaca 1080

ccgttcttgt actggtgctt gacagtgttg ccaatatctg cgaactttct gtcctcgaac 1140

aggaagaaac cgtgcttaag agcaagttcc ttgaggggga gcacagtgct ggcgtaggtg 1200

aagtcgtcaa tgatgtcgat atgggtcttg atcatgcaca cataaggtcc gaccttatcg 1260

gcaagctcaa tgagctcctt ggtggtggta acatccagag aagcacacag gttggttttc 1320

ttggctgcca cgagcttgag cactcgagcg gcaaaggcgg acttgtggac gttagctcga 1380

gcttcgtagg agggcatttt ggtggtgaag aggagactga aataaattta gtctgcagaa 1440

ctttttatcg gaaccttatc tggggcagtg aagtatatgt tatggtaata gttacgagtt 1500

agttgaactt atagatagac tggactatac ggctatcggt ccaaattaga aagaacgtca 1560

atggctctct gggcgcgaat tcgtataact tcgtatagca ggagttatcc gaagcgaacg 1620

gcgatacgcg tatcgatacg cgtgcatgct gaggtgtctc acaagtgccg tgcagtcccg 1680

cccccacttg cttctctttg tgtgtagtgt acgtacatta tcgagagggt tgttcccgcc 1740

cacctcgatc cggcatgctg aggtgtctca caagtgccgt gcagtcccgc ccccacttgc 1800

ttctctttgt gtgtagtgta cgtacattat cgagagggtt gttcccgccc acctcgatcc 1860

ggcatgctga ggtgtctcac aagtgccgtg cagtcccgcc cccacttgct tctctttgtg 1920

tgtagtgtac gtacattatc gagagggttg ttcccgccca cctcgatccg gcatgctgag 1980

gtgtctcaca agtgccgtgc agtcccgccc ccacttgctt ctctttgtgt gtagtgtacg 2040

tacattatcg agagggttgt tcccgcccac ctcgatccgg catgcactga tcacgggcaa 2100

aagtgcgttc gatagagagg gggttggcgg cgcatttgtg tcccaaaaaa cagccccaat 2160

tgccccaatt gaccccaaat tgacccagta gcggacccaa ccccggcgag agccccccttc 2220

accccacata tcaaacctcc cccggttccc acacttgccg ttaagggcgt agggtactgc 2280

agtctggaat ctacgcttgt tcagactttg tactagtttc tttgtctggc catccgggta 2340

acccatgccg gacgcaaaat agactactga aaatttttt gctttgtggt tgggacttta 2400

gccaagggta taaaagacca ccgtccccga attacctttc ctcttctttt ctctctctcc 2460

ttgtcaactc acacccgaag gatccaatga tgttcaaaat ttcgtcgtca ctctcacctg 2520

ttgaccaaat tcccccaatt agcccttttgc ccagaccacg accccgtcca attacaccaa 2580

gagttctggc agtgtccagc tccagcggca aaatcgttaa taatcccttg aaagccgaaa 2640

ccaccgaagc ggtcagtggc gaactggccc gcagatttag attgggtcga ctggctgaag 2700

atggatttag ctataaagag aaattcatcg tcagatgtta tgaagttggt attaataaaa 2760

ctgcgacagt cgaaaccttg gctaatttgc tgcaagaagt gggtggaaat catgcccaat 2820

cggttggctt tagcactgat ggcttcgcta ccactcattc catgcggaaa atgcatttga 2880

tttgggtgac cgctcggatg catatagaaa tctataaata tcctgcctgg tcggacgtca 2940

ttgaagtcga aacttggatc ggagccgaag gtcgtattgg taccagacga aattggatta 3000

ttaaagactg cgcgacagat gaagttattg gtcgggctac aagcaaatgg gtcatgatga 3060

atcaagacac acggcgattg gaaaaggttt cggatgatgt tcgtgaagaa catttggtct 3120

tttcccccag agaaccgcgc ctgccgttcc ccgacgaaaa taactcttca ttgaagaaaa 3180

tctcgaaatt agatgatccg gctcaatatt ccaaattgag tttggaacca cgacggggag 3240

acttagatat gaatcaacat gttaacaatg ttacttatat aggctgggtt ctcgagagta 3300

tgccacaaga gatcattgac acacatgaat tgcagacaat aactttggac taccggcgag 3360

aatgtcaaca cgatgattta gtggactcgc tgacctctcc tgagccttct gaattttctg 3420

aaactactaa tggtagtgcc aatgtcagcc ccaatgataa ccgctgttta aacttcttgc 3480

atttgctgag acttagctcc gacggctctg aaattaaccg gggtcgtact gtgtggagaa 3540

agaaacccgc aaaataatct agtgtctgtg gtatctaagc tatttatcac tctttacaac 3600

ttctacctca actatctact ttaataaatg aatatcgttt attctctatg attactgtat 3660

atgcgttcct ctaagacaaa tcggcttcga tacgcgtatc gatacgcgtg catgctgagg 3720

tgtctcacaa gtgccgtgca gtcccgcccc cacttgcttc tctttgtgtg tagtgtacgt 3780

acattatcga gagggttgtt cccgcccacc tcgatccggc atgctgaggt gtctcacaag 3840

tgccgtgcag tcccgccccc acttgcttct ctttgtgtgt agtgtacgta cattatcgag 3900

agggttgttc ccgcccacct cgatccggca tgctgaggtg tctcacaagt gccgtgcagt 3960

cccgccccca cttgcttctc tttgtgtgta gtgtacgtac attatcgaga gggttgttcc 4020

cgcccacctc gatccggcat gctgaggtgt ctcacaagtg ccgtgcagtc ccgccccac 4080

ttgcttctct ttgtgtgtag tgtacgtaca ttatcgagag ggttgttccc gcccacctcg 4140

atccggcatg cactgatcac gggcaaaagt gcgttcgata gagaggggt tggcggcgca 4200

tttgtgtccc aaaaaacagc cccaattgcc ccaattgacc ccaaattgac ccagtagcgg 4260

acccaacccc ggcgagagcc cccttcaccc cacatatcaa acctcccccg gttcccacac 4320

ttgccgttaa gggcgtaggg tactgcagtc tggaatctac gcttgttcag actttgtact 4380

agtttctttg tctggccatc cgggtaaccc atgccggacg caaaatagac tactgaaaat 4440

ttttttgctt tgtggttggg actttagcca agggtataaa agaccaccgt ccccgaatta 4500

cctttcctct tcttttctct ctctccttgt caactcacac ccgaaggatc cacaacacaa 4560

tgtctatctc taactctccc gagaacctgg accagaagcc taacatccga aacgagtctg 4620

aggaccgatc tttctctgtg cgacgacgaa ctaactctaa cgtggtgaac gctgtgcccg 4680

actcttgctc taaggtgggc accatcgaga ctgtgaccaa gttctcttct gtggactctg 4740

ccgccgagat gtcctctggc gaggaccgaa accgaatcgg caacggtgac gaccgaaagg 4800

aagaggccaa ctctctgtct aacggccagg tgtacaacga gcgagtcggc atcggcgagg 4860

gacgatctcc cgaggccgag gtgcccttcc gattctctta ccgagcctct gctcccgctc 4920

accgacgagt gaaggaatct cccctgtcct ctgacgccat cttccgacag tctcacgccg 4980

gcctgttcaa cctgtgcatc ctggtgctgg tggccgtgaa cggccgactg atcattgaga 5040

acctgatgaa gtacggcctg ctgatccgag ccggattctg gttctcttcg aagtctctgc 5100

gagactggcc cctgctgatg tgctgcctgt ctctgcctat cttccctctg ggcgccttca 5160

ccgtcgagaa gtctgccaag cacaagttca tccccgagcc tgtggtgatg tctctgcacg 5220

tgctgatctg taccgcctct ctgctgtacc ccatcttcgt gatcctccga tgcggatctg 5280

ccgtgctgtt cggaatcgcc ctgatgctgc tggcctctat cgtgtggctg aagctggtgt 5340

cttacgccca caccaactgc gacatgcgag ccctggccaa gctgaccgac aaggtggcct 5400

ctaagtcctc cgacatggac tactcttacg acgtgtcttt cgagtctctg gcctacttca 5460

tggtggctcc cactctgtgc taccaggctt cttaccctcg atctacctgc atccgaaagg 5520

gctgggtgat ccgacagctg gtcaagctgg tgattttcac cggcgtgatg ggcttcatcg 5580

tcgagcagta catcaacccc atcgtgcagc actctcagca ccctctgaag ggcaacctgc 5640

tgtacgctat cgagcgagtg ctgaagctct ctgtccctac tctgtacgtc tggctgtgca 5700

tgttctactg tttcttccac ctgtggctca acatcctggc cgagctgctg cgattcggcg 5760

accgagagtt ctacaaggac tggtggaacg ccaagaccgt ggaagagtac tggcgaatgt 5820

ggaacatgcc cgtgcacaag tggatggtgc gacacatcta cttcccttgc ctgcgaaacg 5880

gactgccccg agaggtgacc ctggtgatcg tgttcttcct gtctgccatt ttccacgagc 5940

tgtgcatcgg cgtgccctgc cacatcttca agttctgggc cttcatcggc atcatgttcc 6000

aggtgcctct ggtgctgatc accaactacc tgcagaacaa gttccgatct tctatggtgg 6060

gcaacatggt gttctggttt ttcttctcga tcctgggaca gcccatgtct gtgctgctgt 6120

actaccacga cgtgatgaac agaaaggaaa acggcgtttc taagtaggga tgtgtctgtg 6180

gtatctaagc tatttatcac tctttacaac ttctacctca actatctact ttaataaatg 6240

aatatcgttt attctctatg attactgtat atgcgttcct ctaagacaga gtcatgtgta 6300

acactcgctc tggagagtta gtcatccgac agggtaactc taatctccca acaccttatt 6360

aactctgcgt aactgtaact cttcttgcca cgtcgatctt actcaatttt cctgctcatc 6420

atctgctgga ttgttgtcta tcgtctggct ctaatacatt tattgtttat tgcccaaaca 6480

actttcattg cacgtaagtg aattgtttta taacagcgtt cgccaattgc tgcgccatcg 6540

tcgtccggct gtcctaccgt taggggtagt gtgtctcaca ctaccgaggt tactagagtt 6600

gggaaagcga tactgcctcg gacacaccac ctgggtctta cgactgcaga gagaatcggc 6660

gttacctctc tcacaaagcc cttcagtgcg gccgc 6695

<210> 113
<211> 7265
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of the NotI DNA fragment of pAT215 used to transform Y. lipolytica cells.

<400> 113
gcggccgctg tcgggaaccg cgttcaggtg gaacaggacc acctcccttg cacttcttgg 60

tatatcagta taggctgatg tattcatagt ggggtttttc ataataaatt tactaacggc 120

aggcaacatt cactcggctt aaacgcaaaa cggaccgtct tgatatcttc tgacgcattg 180

accaccgaga aatagtgtta gttaccgggt gagttattgt tcttctacac aggcgacgcc 240

catcgtctag agttgatgta ctaactcaga tttcactacc taccctatcc ctggtacgca 300

caaagcactt tgctagaagg tataacttcg tatagcatac attatacgaa gttattctga 360

attccgagaa acacaacaac atgccccatt ggacagatca tgcggataca caggttgtgc 420

agtatcatac atactcgatc agacaggtcg tctgaccatc atacaagctg aacaagcgct 480

ccatacttgc acgctctcta tatacacagt taaattacat atccatagtc taacctctaa 540

cagttaatct tctggtaagc ctcccagcca gccttctggt atcgcttggc ctcctcgata 600

ggatctcggt tctggccgta cagacctcgg ccgacaatta tgatatccgt tccggtagac 660

atgacatcct caacagttcg gtactgctgt ccgagagcgt ctcccttgtc gtcaagaccc 720

accccggggg tcagaataag ccagtcctca gagtcgccct taggtcggtt ctgggcaacg 780

aagccaacca caaactcggg gtcggatcgg gcaagctcaa tggtctgctt ggagtactcg 840

ccagtggcca gagagccctt gcaagacagc tcggccagca tgagcagacc tctggccagc 900

ttctcgttgg gagaggggac caggaactcc ttgtactggg agttctcgta gtcagagacg 960

tcctccttct tctgttcaga gacagtttcc tcggcaccag ctcgcaggcc agcaatgatt 1020

ccggttccgg gtacaccgtg ggcgttggtg atatcggacc actcggcgat tcggtagaca 1080

ccgttcttgt actggtgctt gacagtgttg ccaatatctg cgaactttct gtcctcgaac 1140

aggaagaaac cgtgcttaag agcaagttcc ttgaggggga gcacagtgct ggcgtaggtg 1200

aagtcgtcaa tgatgtcgat atgggtcttg atcatgcaca cataaggtcc gaccttatcg 1260

gcaagctcaa tgagctcctt ggtggtggta acatccagag aagcacacag gttggttttc 1320

ttggctgcca cgagcttgag cactcgagcg gcaaaggcgg acttgtggac gttagctcga 1380

gcttcgtagg agggcatttt ggtggtgaag aggagactga aataaattta gtctgcagaa 1440

ctttttatcg gaaccttatc tggggcagtg aagtatatgt tatggtaata gttacgagtt 1500

agttgaactt atagatagac tggactatac ggctatcggt ccaaattaga aagaacgtca 1560

atggctctct gggcgcgaat tcgtataact tcgtatagca ggagttatcc gaagcgaacg 1620

gcgatacgcg tatcgatacg cgtgcatgct gaggtgtctc acaagtgccg tgcagtcccg 1680

cccccacttg cttctctttg tgtgtagtgt acgtacatta tcgagagggt tgttcccgcc 1740

cacctcgatc cggcatgctg aggtgtctca caagtgccgt gcagtcccgc ccccacttgc 1800

ttctctttgt gtgtagtgta cgtacattat cgagagggtt gttcccgccc acctcgatcc 1860

ggcatgctga ggtgtctcac aagtgccgtg cagtcccgcc cccacttgct tctctttgtg 1920

tgtagtgtac gtacattatc gagagggttg ttcccgccca cctcgatccg gcatgctgag 1980

gtgtctcaca agtgccgtgc agtcccgccc ccacttgctt ctctttgtgt gtagtgtacg 2040

tacattatcg agagggttgt tcccgcccac ctcgatccgg catgcactga tcacgggcaa 2100

aagtgcgttc gatagagagg gggttggcgg cgcatttgtg tcccaaaaaa cagccccaat 2160

tgccccaatt gaccccaaat tgacccagta gcggacccaa ccccggcgag agccccccttc 2220

accccacata tcaaacctcc cccggttccc acacttgccg ttaagggcgt agggtactgc 2280

agtctggaat ctacgcttgt tcagactttg tactagtttc tttgtctggc catccgggta 2340

acccatgccg gacgcaaaat agactactga aaatttttt gctttgtggt tgggacttta 2400

gccaagggta taaaagacca ccgtccccga attacctttc ctcttctttt ctctctctcc 2460

ttgtcaactc acacccgaag gatccaatga tgcgagacta cgagcatgtt gtggaaagtt 2520

ttgactactt acaatccgca actcaagacc ttcacggaga gttgaccgcc ctgaatgctt 2580

gtgttgagtg ttgtgataga cacgcgcatg gtgaagctgt tgccctttac tgtgaggccc 2640

aagacggtca cgcggaacga tatcggttcc gagatttgca gcgtcaagcc gcacgatttg 2700

gtaactttct tcgagagcag ggagtgaaac ctggtgaccg agtggctggt ttgatgcccc 2760

gtaccgtcga gttgcttatc gctatcctgg gaacatggcg gattggagct gtgtatcagc 2820

ccttgttcac cgcgtttggt ccgaaagcta ttgagcagcg gctgaattgt agcaatgctc 2880

gatggattgt cactgatccc cataatcgac caaaattgga tgatgttacc gattgtccga 2940

gtatagtcgt cactggcggt gcaccccaaa atccggcaga ccatcacttt tggtccgcct 3000

tgaatcgaca agccgatgac tgtgctcctg tcttattaga cgcttccgct cccttcttgc 3060

tcatgtgcac ttcgggtact accggccctg cgaagccttt ggaagtcccc ctctcggcca 3120

tcttagcttt taaaggttac atgcgagatg ccatcgatct tagagcagat gatcgatttt 3180

ggaatttggc cgatccagga tgggcgtacg gactgtacta tgccgtcact ggacccttgg 3240

cctgcggcta tgccaccttg ttttatgatg gaccgtttac cgttgagtcc actcgacata 3300

taatcgctaa atatgccata aataatttgg ctggatcacc gacagcttat cgattcttga 3360

ttgctgccgg cgctgaattt gctgatgcag tgcgcggtcg attgcgtgct gtgtctagtg 3420

cgggcgagcc cttgaatccc caggtggttc gatggttcgc tgagcaactg ggagtcgtca 3480

tccatgatca ttatggtcag acagagatag gtatggtcct ttgtaaccat catggacttc 3540

gacatcccgt gcgtgagggt tcggcaggat acgcagttcc tggttatcgt attgtggtcc 3600

tcgacaaagc tcatcgagag ctgccggcag gtcaacctgg agttttggca gtcgaccgag 3660

aacgatcccc tctgtgttgg tttgatggtt atttgggaat gcctacacaa gcgttcgccg 3720

gacgatacta tttgtccgga gacatcgtcg aactgaacga tgatggttct atttcttttg 3780

ttggtcgaaa cgacgatttg atcactacct ctggatatcg agtgggtcca tttgacgtcg 3840

agtccgcctt gattgagcat cccgctgtcg ttgaagctgc agttatagga aagccagacc 3900

ctcagcggac agaactgatt aaagccttcg ttgtcttgaa tactccatat ctcccttcgc 3960

cagaattagc cgaagagttg cgtttgcacg tgcgacaacg gttggctgcc cacgcatacc 4020

cccgagagat ggaattcgtt gatcacttac cgaaaacacc ttcgggaaaa ctgcagcgat 4080

ttatcttgag aaatcaagaa atcgcgaaac agcaagcctt gggatagtct agtgtctgtg 4140

gtatctaagc tatttatcac tctttacaac ttctacctca actatctact ttaataaatg 4200

aatatcgttt attctctatg attactgtat atgcgttcct ctaagacaaa tcggcttcga 4260

tacgcgtatc gatacgcgtg catgctgagg tgtctcacaa gtgccgtgca gtcccgcccc 4320

cacttgcttc tctttgtgtg tagtgtacgt acattatcga gagggttgtt cccgcccacc 4380

tcgatccggc atgctgaggt gtctcacaag tgccgtgcag tcccgccccc acttgcttct 4440

ctttgtgtgt agtgtacgta cattatcgag agggttgttc ccgcccacct cgatccggca 4500

tgctgaggtg tctcacaagt gccgtgcagt cccgccccca cttgcttctc tttgtgtgta 4560

gtgtacgtac attatcgaga gggttgttcc cgcccacctc gatccggcat gctgaggtgt 4620

ctcacaagtg ccgtgcagtc ccgcccccac ttgcttctct ttgtgtgtag tgtacgtaca 4680

ttatcgagag ggttgttccc gcccacctcg atccggcatg cactgatcac gggcaaaagt 4740

gcgttcgata gagagggggt tggcggcgca tttgtgtccc aaaaaacagc cccaattgcc 4800

ccaattgacc ccaaattgac ccagtagcgg acccaacccc ggcgagagcc cccttcaccc 4860

cacatatcaa acctcccccg gttcccacac ttgccgttaa gggcgtaggg tactgcagtc 4920

tggaatctac gcttgttcag actttgtact agtttctttg tctggccatc cgggtaaccc 4980

atgccggacg caaaatagac tactgaaaat ttttttgctt tgtggttggg actttagcca 5040

agggtataaa agaccaccgt ccccgaatta cctttcctct tcttttctct ctctccttgt 5100

caactcacac ccgaaggatc cacaacacaa tgtctatctc taactctccc gagaacctgg 5160

accagaagcc taacatccga aacgagtctg aggaccgatc tttctctgtg cgacgacgaa 5220

ctaactctaa cgtggtgaac gctgtgcccg actcttgctc taaggtgggc accatcgaga 5280

ctgtgaccaa gttctcttct gtggactctg ccgccgagat gtcctctggc gaggaccgaa 5340

accgaatcgg caacggtgac gaccgaaagg aagaggccaa ctctctgtct aacggccagg 5400

tgtacaacga gcgagtcggc atcggcgagg gacgatctcc cgaggccgag gtgcccttcc 5460

gattctctta ccgagcctct gctcccgctc accgacgagt gaaggaatct cccctgtcct 5520

ctgacgccat cttccgacag tctcacgccg gcctgttcaa cctgtgcatc ctggtgctgg 5580

tggccgtgaa cggccgactg atcattgaga acctgatgaa gtacggcctg ctgatccgag 5640

ccggattctg gttctcttcg aagtctctgc gagactggcc cctgctgatg tgctgcctgt 5700

ctctgcctat cttccctctg ggcgccttca ccgtcgagaa gtctgccaag cacaagttca 5760

tccccgagcc tgtggtgatg tctctgcacg tgctgatctg taccgcctct ctgctgtacc 5820

ccatcttcgt gatcctccga tgcggatctg ccgtgctgtt cggaatcgcc ctgatgctgc 5880

tggcctctat cgtgtggctg aagctggtgt cttacgccca caccaactgc gacatgcgag 5940

ccctggccaa gctgaccgac aaggtggcct ctaagtcctc cgacatggac tactcttacg 6000

acgtgtcttt cgagtctctg gcctacttca tggtggctcc cactctgtgc taccaggctt 6060

cttaccctcg atctacctgc atccgaaagg gctgggtgat ccgacagctg gtcaagctgg 6120

tgattttcac cggcgtgatg ggcttcatcg tcgagcagta catcaacccc atcgtgcagc 6180

actctcagca ccctctgaag ggcaacctgc tgtacgctat cgagcgagtg ctgaagctct 6240

ctgtccctac tctgtacgtc tggctgtgca tgttctactg tttcttccac ctgtggctca 6300

acatcctggc cgagctgctg cgattcggcg accgagagtt ctacaaggac tggtggaacg 6360

ccaagaccgt ggaagagtac tggcgaatgt ggaacatgcc cgtgcacaag tggatggtgc 6420

gacacatcta cttcccttgc ctgcgaaacg gactgccccg agaggtgacc ctggtgatcg 6480

tgttcttcct gtctgccatt ttccacgagc tgtgcatcgg cgtgccctgc cacatcttca 6540

agttctgggc cttcatcggc atcatgttcc aggtgcctct ggtgctgatc accaactacc 6600

tgcagaacaa gttccgatct tctatggtgg gcaacatggt gttctggttt ttcttctcga 6660

tcctgggaca gcccatgtct gtgctgctgt actaccacga cgtgatgaac agaaaggaaa 6720

acggcgtttc taagtaggga tgtgtctgtg gtatctaagc tatttatcac tctttacaac 6780

ttctacctca actatctact ttaataaatg aatatcgttt attctctatg attactgtat 6840

atgcgttcct ctaagacaga gtcatgtgta acactcgctc tggagagtta gtcatccgac 6900

agggtaactc taatctccca acaccttatt aactctgcgt aactgtaact cttcttgcca 6960

cgtcgatctt actcaatttt cctgctcatc atctgctgga ttgttgtcta tcgtctggct 7020

ctaatacatt tattgtttat tgcccaaaca actttcattg cacgtaagtg aattgtttta 7080

taacagcgtt cgccaattgc tgcgccatcg tcgtccggct gtcctaccgt taggggtagt 7140

gtgtctcaca ctaccgaggt tactagagtt gggaaagcga tactgcctcg gacacaccac 7200

ctgggtctta cgactgcaga gagaatcggc gttacctctc tcacaaagcc cttcagtgcg 7260

gccgc 7265

<210> 114
<211> 9267
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence of the NotI DNA fragment of pAT216 used to transform Y. lipolytica cells.

<400> 114
gcggccgctg tcgggaaccg cgttcaggtg gaacaggacc acctcccttg cacttcttgg 60

tatatcagta taggctgatg tattcatagt ggggtttttc ataataaatt tactaacggc 120

aggcaacatt cactcggctt aaacgcaaaa cggaccgtct tgatatcttc tgacgcattg 180

accaccgaga aatagtgtta gttaccgggt gagttattgt tcttctacac aggcgacgcc 240

catcgtctag agttgatgta ctaactcaga tttcactacc taccctatcc ctggtacgca 300

caaagcactt tgctagaagg tataacttcg tatagcatac attatacgaa gttattctga 360

attccgagaa acacaacaac atgccccatt ggacagatca tgcggataca caggttgtgc 420

agtatcatac atactcgatc agacaggtcg tctgaccatc atacaagctg aacaagcgct 480

ccatacttgc acgctctcta tatacacagt taaattacat atccatagtc taacctctaa 540

cagttaatct tctggtaagc ctcccagcca gccttctggt atcgcttggc ctcctcgata 600

ggatctcggt tctggccgta cagacctcgg ccgacaatta tgatatccgt tccggtagac 660

atgacatcct caacagttcg gtactgctgt ccgagagcgt ctcccttgtc gtcaagaccc 720

accccggggg tcagaataag ccagtcctca gagtcgccct taggtcggtt ctgggcaacg 780

aagccaacca caaactcggg gtcggatcgg gcaagctcaa tggtctgctt ggagtactcg 840

ccagtggcca gagagccctt gcaagacagc tcggccagca tgagcagacc tctggccagc 900

ttctcgttgg gagaggggac caggaactcc ttgtactggg agttctcgta gtcagagacg 960

tcctccttct tctgttcaga gacagtttcc tcggcaccag ctcgcaggcc agcaatgatt 1020

ccggttccgg gtacaccgtg ggcgttggtg atatcggacc actcggcgat tcggtagaca 1080

ccgttcttgt actggtgctt gacagtgttg ccaatatctg cgaactttct gtcctcgaac 1140

aggaagaaac cgtgcttaag agcaagttcc ttgaggggga gcacagtgct ggcgtaggtg 1200

aagtcgtcaa tgatgtcgat atgggtcttg atcatgcaca cataaggtcc gaccttatcg 1260

gcaagctcaa tgagctcctt ggtggtggta acatccagag aagcacacag gttggttttc 1320

ttggctgcca cgagcttgag cactcgagcg gcaaaggcgg acttgtggac gttagctcga 1380

gcttcgtagg agggcatttt ggtggtgaag aggagactga aataaattta gtctgcagaa 1440

ctttttatcg gaaccttatc tggggcagtg aagtatatgt tatggtaata gttacgagtt 1500

agttgaactt atagatagac tggactatac ggctatcggt ccaaattaga aagaacgtca 1560

atggctctct gggcgcgaat tcgtataact tcgtatagca ggagttatcc gaagcgaacg 1620

gcgatacgcg tatcgatacg cgtgcatgct gaggtgtctc acaagtgccg tgcagtcccg 1680

cccccacttg cttctctttg tgtgtagtgt acgtacatta tcgagagggt tgttcccgcc 1740

cacctcgatc cggcatgctg aggtgtctca caagtgccgt gcagtcccgc ccccacttgc 1800

ttctctttgt gtgtagtgta cgtacattat cgagagggtt gttcccgccc acctcgatcc 1860

ggcatgctga ggtgtctcac aagtgccgtg cagtcccgcc cccacttgct tctctttgtg 1920

tgtagtgtac gtacattatc gagagggttg ttcccgccca cctcgatccg gcatgctgag 1980

gtgtctcaca agtgccgtgc agtcccgccc ccacttgctt ctctttgtgt gtagtgtacg 2040

tacattatcg agagggttgt tcccgcccac ctcgatccgg catgcactga tcacgggcaa 2100

aagtgcgttc gatagagagg gggttggcgg cgcatttgtg tcccaaaaaa cagccccaat 2160

tgccccaatt gaccccaaat tgacccagta gcggacccaa ccccggcgag agcccccttc 2220

accccacata tcaaacctcc cccggttccc acacttgccg ttaagggcgt agggtactgc 2280

agtctggaat ctacgcttgt tcagactttg tactagtttc tttgtctggc catccgggta 2340

acccatgccg gacgcaaaat agactactga aaatttttt gctttgtggt tgggacttta 2400

gccaagggta taaaagacca ccgtccccga attacctttc ctcttctttt ctctctctcc 2460

ttgtcaactc acacccgaag gatccaatga tgttcaaaat ttcgtcgtca ctctcacctg 2520

ttgaccaaat tcccccaatt agcccttttgc ccagaccacg accccgtcca attacaccaa 2580

gagttctggc agtgtccagc tccagcggca aaatcgttaa taatcccttg aaagccgaaa 2640

ccaccgaagc ggtcagtggc gaactggccc gcagatttag attgggtcga ctggctgaag 2700

atggatttag ctataaagag aaattcatcg tcagatgtta tgaagttggt attaataaaa 2760

ctgcgacagt cgaaaccttg gctaatttgc tgcaagaagt gggtggaaat catgcccaat 2820

cggttggctt tagcactgat ggcttcgcta ccactcattc catgcggaaa atgcatttga 2880

tttgggtgac cgctcggatg catatagaaa tctataaata tcctgcctgg tcggacgtca 2940

ttgaagtcga aacttggatc ggagccgaag gtcgtattgg taccagacga aattggatta 3000

ttaaagactg cgcgacagat gaagttattg gtcgggctac aagcaaatgg gtcatgatga 3060

atcaagacac acggcgattg gaaaaggttt cggatgatgt tcgtgaagaa catttggtct 3120

tttcccccag agaaccgcgc ctgccgttcc ccgacgaaaa taactcttca ttgaagaaaa 3180

tctcgaaatt agatgatccg gctcaatatt ccaaattgag tttggaacca cgacggggag 3240

acttagatat gaatcaacat gttaacaatg ttacttatat aggctgggtt ctcgagagta 3300

tgccacaaga gatcattgac acacatgaat tgcagacaat aactttggac taccggcgag 3360

aatgtcaaca cgatgattta gtggactcgc tgacctctcc tgagccttct gaattttctg 3420

aaactactaa tggtagtgcc aatgtcagcc ccaatgataa ccgctgttta aacttcttgc 3480

atttgctgag acttagctcc gacggctctg aaattaaccg gggtcgtact gtgtggagaa 3540

agaaacccgc aaaataatct agtgtctgtg gtatctaagc tatttatcac tctttacaac 3600

ttctacctca actatctact ttaataaatg aatatcgttt attctctatg attactgtat 3660

atgcgttcct ctaagacaaa tcggcttcga tacgcgtatc gatacgcgtg catgctgagg 3720

tgtctcacaa gtgccgtgca gtcccgcccc cacttgcttc tctttgtgtg tagtgtacgt 3780

acattatcga gagggttgtt cccgcccacc tcgatccggc atgctgaggt gtctcacaag 3840

tgccgtgcag tcccgccccc acttgcttct ctttgtgtgt agtgtacgta cattatcgag 3900

agggttgttc ccgcccacct cgatccggca tgctgaggtg tctcacaagt gccgtgcagt 3960

cccgccccca cttgcttctc tttgtgtgta gtgtacgtac attatcgaga gggttgttcc 4020

cgcccacctc gatccggcat gctgaggtgt ctcacaagtg ccgtgcagtc ccgccccac 4080

ttgcttctct ttgtgtgtag tgtacgtaca ttatcgagag ggttgttccc gcccacctcg 4140

atccggcatg cactgatcac gggcaaaagt gcgttcgata gagaggggt tggcggcgca 4200

tttgtgtccc aaaaaacagc cccaattgcc ccaattgacc ccaaattgac ccagtagcgg 4260

acccaacccc ggcgagagcc cccttcaccc cacatatcaa acctcccccg gttcccacac 4320

ttgccgttaa gggcgtaggg tactgcagtc tggaatctac gcttgttcag actttgtact 4380

agtttctttg tctggccatc cgggtaaccc atgccggacg caaaatagac tactgaaaat 4440

ttttttgctt tgtggttggg actttagcca agggtataaa agaccaccgt ccccgaatta 4500

cctttcctct tcttttctct ctctccttgt caactcacac ccgaaggatc cacaaatgcg 4560

agactacgag catgttgtgg aaagttttga ctacttacaa tccgcaactc aagaccttca 4620

cggagagttg accgccctga atgcttgtgt tgagtgttgt gatagacacg cgcatggtga 4680

agctgttgcc ctttactgtg aggcccaaga cggtcacgcg gaacgatatc ggttccgaga 4740

tttgcagcgt caagccgcac gatttggtaa ctttcttcga gagcagggag tgaaacctgg 4800

tgaccgagtg gctggtttga tgccccgtac cgtcgagttg cttatcgcta tcctgggaac 4860

atggcggatt ggagctgtgt atcagccctt gttcaccgcg tttggtccga aagctattga 4920

gcagcggctg aattgtagca atgctcgatg gattgtcact gatccccata atcgaccaaa 4980

attggatgat gttaccgatt gtccgagtat agtcgtcact ggcggtgcac cccaaaatcc 5040

ggcagaccat cacttttggt ccgccttgaa tcgacaagcc gatgactgtg ctcctgtctt 5100

attagacgct tccgctccct tcttgctcat gtgcacttcg ggtactaccg gccctgcgaa 5160

gcctttggaa gtccccctct cggccatctt agcttttaaa ggttacatgc gagatgccat 5220

cgatcttaga gcagatgatc gattttggaa tttggccgat ccaggatggg cgtacggact 5280

gtactatgcc gtcactggac ccttggcctg cggctatgcc accttgtttt atgatggacc 5340

gtttaccgtt gagtccactc gacatataat cgctaaatat gccataaata atttggctgg 5400

atcaccgaca gcttatcgat tcttgattgc tgccggcgct gaatttgctg atgcagtgcg 5460

cggtcgattg cgtgctgtgt ctagtgcggg cgagcccttg aatccccagg tggttcgatg 5520

gttcgctgag caactgggag tcgtcatcca tgatcattat ggtcagacag agataggtat 5580

ggtcctttgt aaccatcatg gacttcgaca tcccgtgcgt gagggttcgg caggatacgc 5640

agttcctggt tatcgtattg tggtcctcga caaagctcat cgagagctgc cggcaggtca 5700

acctggagtt ttggcagtcg accgagaacg atcccctctg tgttggtttg atggttattt 5760

gggaatgcct acacaagcgt tcgccggacg atactatttg tccggagaca tcgtcgaact 5820

gaacgatgat ggttctattt cttttgttgg tcgaaacgac gatttgatca ctacctctgg 5880

atatcgagtg ggtccatttg acgtcgagtc cgccttgatt gagcatcccg ctgtcgttga 5940

agctgcagtt ataggaaagc cagaccctca gcggacagaa ctgattaaag ccttcgttgt 6000

cttgaatact ccatatctcc cttcgccaga attagccgaa gagttgcgtt tgcacgtgcg 6060

acaacggttg gctgcccacg catacccccg agagatggaa ttcgttgatc acttaccgaa 6120

aacaccttcg ggaaaactgc agcgatttat cttgagaaat caagaaatcg cgaaacagca 6180

agccttggga tagggattgt ctgtggtatc taagctattt atcactcttt acaacttcta 6240

cctcaactat ctactttaat aaatgaatat cgtttattct ctatgattac tgtatatgcg 6300

ttcctctaag acaaatcggt cacgatacgc gtatcgatac gcgtgcatgc tgaggtgtct 6360

cacaagtgcc gtgcagtccc gcccccactt gcttctcttt gtgtgtagtg tacgtacatt 6420

atcgagaggg ttgttcccgc ccacctcgat ccggcatgct gaggtgtctc acaagtgccg 6480

tgcagtcccg cccccacttg cttctctttg tgtgtagtgt acgtacatta tcgagagggt 6540

tgttcccgcc cacctcgatc cggcatgctg aggtgtctca caagtgccgt gcagtcccgc 6600

ccccacttgc ttctctttgt gtgtagtgta cgtacattat cgagagggtt gttcccgccc 6660

acctcgatcc ggcatgctga ggtgtctcac aagtgccgtg cagtcccgcc cccacttgct 6720

tctctttgtg tgtagtgtac gtacattatc gagagggttg ttcccgccca cctcgatccg 6780

gcatgcactg atcacgggca aaagtgcgtt cgatagagag ggggttggcg gcgcatttgt 6840

gtcccaaaaa acagccccaa ttgccccaat tgaccccaaa ttgacccagt agcggaccca 6900

accccggcga gagccccctt caccccacat atcaaacctc ccccggttcc cacacttgcc 6960

gttaagggcg tagggtactg cagtctggaa tctacgcttg ttcagacttt gtactagttt 7020

ctttgtctgg ccatccgggt aacccatgcc ggacgcaaaa tagactactg aaaatttttt 7080

tgctttgtgg ttgggacttt agccaagggt ataaaagacc accgtccccg aattaccttt 7140

cctcttcttt tctctctctc cttgtcaact cacacccgaa ggatccccac cacaatgact 7200

atcgactcac aatactacaa gtcgcgagac aaaaacgaca cggcacccaa aatcgcggga 7260

atccgatatg ccccgctatc gacaccatta ctcaaccgat gtgagacttt ctctctggtc 7320

tggcacattt tcagcattcc cactttcctc acaattttca tgctatgctg cgcaattcca 7380

ctgctctggc catttgtgat tgcgtatgta gtgtacgctg ttaaagacga ctccccgtcc 7440

aacggaggag tggtcaagcg atactcgcct atttcaagaa acttcttcat ctggaagctc 7500

tttggccgct acttccccat aactctgcac aagacggtgg atctggagcc cacgcacaca 7560

tactaccctc tggacgtcca ggagtatcac ctgattgctg agagatactg gccgcagaac 7620

aagtacctcc gagcaatcat ctccaccatc gagtactttc tgcccgcctt catgaaacga 7680

tctctttcta tcaacgagca ggagcagcct gccgagcgag atcctctcct gtctcccgtt 7740

tctcccagct ctccgggttc tcaacctgac aagtggatta accacgacag cagatatagc 7800

cgtggagaat catctggctc caacggccac gcctcgggct ccgaacttaa cggcaacggc 7860

aacaatggca ccactaaccg acgacctttg tcgtccgcct ctgctggctc cactgcatct 7920

gattccacgc ttcttaacgg gtccctcaac tcctacgcca accagatcat tggcgaaaac 7980

gacccacagc tgtcgcccac aaaactcaag cccactggca gaaaatacat cttcggctac 8040

caccccacg gcattatcgg catgggagcc tttggtggaa ttgccaccga gggagctgga 8100

tggtccaagc tctttccggg catccctgtt tctcttatga ctctcaccaa caacttccga 8160

gtgcctctct acagagagta cctcatgagt ctgggagtcg cttctgtctc caagaagtcc 8220

tgcaaggccc tcctcaagcg aaaccagtct atctgcattg tcgttggtgg agcacaggaa 8280

agtctcctgg ccagacccgg tgtcatggac ctggtgctac tcaagcgaaa gggttttgtt 8340

cgacttggta tggaggtcgg aaatgtcgcc cttgttccca tcatggcctt tggtgagaac 8400

gacctctatg accaggttag caacgacaag tcgtccaagc tgtaccgatt ccagcagttt 8460

gtcaagaact tccttggatt cacccttcct ttgatgcatg cccgaggcgt gttcaactac 8520

gatgtcggtc ttgtccccta caggcgaccc gtcaacattg tggttggttc ccccattgac 8580

ttgccttatc tcccacaccc caccgacgaa gaagtgtccg aataccacga ccgatacatc 8640

gccgagctgc agcgaatcta caacgagcac aaggatgaat atttcatcga ttggaccgag 8700

gagggcaaag gagccccaga gttccgaatg attgagtaag tatgtgtctg tggtatctaa 8760

gctatttatc actctttaca acttctacct caactatcta ctttaataaa tgaatatcgt 8820

ttattctcta tgattactgt atatgcgttc ctctaagaca gagtcatgtg taacactcgc 8880

tctggagagt tagtcatccg acagggtaac tctaatctcc caacacctta ttaactctgc 8940

gtaactgtaa ctcttcttgc cacgtcgatc ttactcaatt ttcctgctca tcatctgctg 9000

gattgttgtc tatcgtctgg ctctaataca tttattgttt attgcccaaa caactttcat 9060

tgcacgtaag tgaattgttt tataacagcg ttcgccaatt gctgcgccat cgtcgtccgg 9120

ctgtcctacc gttaggggta gtgtgtctca cactaccgag gttactagag ttgggaaagc 9180

gatactgcct cggacacacc acctgggtct tacgactgca gagagaatcg gcgttacctc 9240

tctcacaaag cccttcagtg cggccgc 9267

<210> 115
<211> 501
<212> PRT

<213> Theobroma cacao

<400> 115

```
Met Ala Ile Ser Asp Ser Pro Glu Ile Leu Gly Ser Thr Ala Thr Val
1               5                  10                  15


Thr Ser Ser Ser His Ser Asp Ser Asp Leu Asn Leu Leu Ser Ile Arg
            20                  25                  30


Arg Arg Thr Ser Thr Thr Ala Ala Gly Arg Ala Pro Asp Arg Asp Asp
        35                  40                  45


Ser Gly Asn Gly Glu Ala Val Asp Asp Arg Asp Gln Val Glu Ser Ala
    50                  55                  60


Asn Leu Met Ser Asn Val Ala Glu Asn Ala Asn Glu Met Pro Asn Ser
65                  70                  75                  80


Ser Asp Thr Arg Phe Thr Tyr Arg Pro Arg Val Pro Ala His Arg Arg
                85                  90                  95


Ile Lys Glu Ser Pro Leu Ser Ser Gly Ala Ile Phe Lys Gln Ser His
            100                 105                 110


Ala Gly Leu Phe Asn Leu Cys Ile Val Val Leu Val Ala Val Asn Ser
        115                 120                 125


Arg Leu Ile Ile Glu Asn Leu Met Lys Tyr Gly Trp Leu Ile Arg Ser
    130                 135                 140


Gly Phe Trp Phe Ser Ser Arg Ser Leu Ser Asp Trp Pro Leu Phe Met
145                 150                 155                 160


Cys Cys Leu Thr Leu Pro Ile Phe Pro Leu Ala Ala Phe Val Val Glu
                165                 170                 175


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

```
His Ala Ile Ile Ser Thr Thr Ala Val Leu Tyr Pro Val Ile Val Asn
        195                 200                 205

Leu Arg Cys Asp Ser Ala Phe Leu Ser Gly Val Ala Leu Met Leu Phe
    210                 215                 220

Ala Cys Ile Val Trp Leu Lys Leu Val Ser Tyr Ala His Thr Asn Asn
225                 230                 235                 240

Asp Met Arg Ala Leu Ala Lys Ser Ala Glu Lys Gly Asp Val Asp Pro
                245                 250                 255

Ser Tyr Asp Val Ser Phe Lys Ser Leu Ala Tyr Phe Met Val Ala Pro
            260                 265                 270

Thr Leu Cys Tyr Gln Gln Ser Tyr Pro Arg Thr Pro Ala Val Arg Lys
        275                 280                 285

Ser Trp Val Val Arg Gln Phe Ile Lys Leu Ile Val Phe Thr Gly Leu
    290                 295                 300

Met Gly Phe Ile Ile Glu Gln Tyr Ile Asn Pro Ile Val Gln Asn Ser
305                 310                 315                 320

Gln His Pro Leu Lys Gly Asn Leu Leu Tyr Ala Ile Glu Arg Val Leu
                325                 330                 335

Lys Leu Ser Val Pro Asn Leu Tyr Val Trp Leu Cys Met Phe Tyr Cys
            340                 345                 350

Phe Phe His Leu Trp Leu Asn Ile Leu Ala Glu Leu Leu Arg Phe Gly
        355                 360                 365

Asp Arg Glu Phe Tyr Lys Asp Trp Trp Asn Ala Lys Thr Val Glu Glu
    370                 375                 380
```

206
Tyr Trp Arg Met Trp Asn Met Pro Val His Lys Trp Met Val Arg His
385 390 395 400

Ile Tyr Phe Pro Cys Leu Arg Asn Gly Ile Pro Lys Gly Val Ala Ile
405 410 415

Val Ile Ala Phe Leu Val Ser Ala Val Phe His Glu Leu Cys Ile Ala
420 425 430

Val Pro Cys His Ile Phe Lys Leu Trp Ala Phe Ile Gly Ile Met Phe
435 440 445

Gln Val Pro Leu Val Leu Ile Thr Asn Tyr Leu Gln Asp Lys Phe Arg
450 455 460

Ser Ser Met Val Gly Asn Met Ile Phe Trp Phe Ile Phe Ser Ile Leu
465 470 475 480

Gly Gln Pro Met Cys Val Leu Leu Tyr Tyr His Asp Leu Met Asn Arg
485 490 495

Lys Gly Lys Ala Asp
500

<210> 116
<211> 327
<212> PRT
<213> Theobroma cacao

<400> 116

Met Met Gly Glu Glu Met Glu Glu Arg Lys Ala Thr Gly Tyr Arg Glu
1 5 10 15

Phe Ser Gly Arg His Glu Phe Pro Ser Asn Thr Met His Ala Leu Leu
20 25 30

Ala Met Gly Ile Trp Leu Gly Ala Ile His Phe Asn Ala Leu Leu Leu
35 40 45

```
Leu Phe Ser Phe Leu Phe Leu Pro Phe Ser Lys Phe Leu Val Val Phe
    50                  55                  60

Gly Leu Leu Leu Leu Phe Met Ile Leu Pro Ile Asp Pro Tyr Ser Lys
65                  70                  75                  80

Phe Gly Arg Arg Leu Ser Arg Tyr Ile Cys Lys His Ala Cys Ser Tyr
                85                  90                  95

Phe Pro Ile Thr Leu His Val Glu Asp Ile His Ala Phe His Pro Asp
            100                 105                 110

Arg Ala Tyr Val Phe Gly Phe Glu Pro His Ser Val Leu Pro Ile Gly
        115                 120                 125

Val Val Ala Leu Ala Asp Leu Thr Gly Phe Met Pro Leu Pro Lys Ile
    130                 135                 140

Lys Val Leu Ala Ser Ser Ala Val Phe Tyr Thr Pro Phe Leu Arg His
145                 150                 155                 160

Ile Trp Thr Trp Leu Gly Leu Thr Pro Ala Thr Lys Lys Asn Phe Ser
                165                 170                 175

Ser Leu Leu Asp Ala Gly Tyr Ser Cys Ile Leu Val Pro Gly Gly Val
            180                 185                 190

Gln Glu Thr Phe His Met Glu Pro Gly Ser Glu Ile Ala Phe Leu Arg
        195                 200                 205

Ala Arg Arg Gly Phe Val Arg Ile Ala Met Glu Met Gly Ser Pro Leu
    210                 215                 220

Val Pro Val Phe Cys Phe Gly Gln Ser His Val Tyr Lys Trp Trp Lys
225                 230                 235                 240

Pro Gly Gly Lys Phe Tyr Leu Gln Phe Ser Arg Ala Ile Lys Phe Thr
```

208

245 250 255

Pro Ile Phe Phe Trp Gly Ile Phe Gly Ser Pro Leu Pro Tyr Gln His
260 265 270

Pro Met His Val Val Val Gly Lys Pro Ile Asp Val Lys Lys Asn Pro
275 280 285

Gln Pro Ile Val Glu Glu Val Ile Glu Val His Asp Arg Phe Val Glu
290 295 300

Ala Leu Gln Asp Leu Phe Glu Arg His Lys Ala Gln Val Gly Phe Ala
305 310 315 320

Asp Leu Pro Leu Lys Ile Leu
325

<210> 117
<211> 476
<212> PRT
<213> Theobroma cacao

<400> 117

Met Gly Ser Glu Ala Gln Tyr Lys Asn Ala Met Ala Leu Asn Gly Glu
1 5 10 15

Glu Gly Glu Tyr Ser Glu Pro Val Ser Pro Thr Gly Gln Tyr Phe Asn
20 25 30

Ser Ser Ala Leu Ser Ile Gly Val Leu Gly Val Leu Asp Ser Glu Ile
35 40 45

Pro Ile Asp Asp Ser Pro Thr Met Lys Leu Leu Glu Asp Val Phe Leu
50 55 60

Pro Ile Asn Pro Arg Phe Ser Ser Val Met Val Gln Asp Glu Asn Gly
65 70 75 80

```
Ala Lys Gln Trp Lys Lys Val Glu Val Lys Leu Val Asp His Val Asn
                85                  90                  95

Val Pro Val Phe Pro Pro Gly Leu Ser Pro Glu Ser Tyr Asp Asn Tyr
            100                 105                 110

Leu Ser Asp Tyr Leu Ser Lys Ile Ala Thr Glu Gln Leu Pro Gln Asn
        115                 120                 125

Arg Pro Leu Trp Asn Ile His Ile Val Lys Tyr Thr Thr Ser Asn Ala
    130                 135                 140

Ala Gly Ser Leu Ile Phe Lys Leu His His Ala Leu Gly Asp Gly Tyr
145                 150                 155                 160

Ser Leu Met Gly Ala Leu Leu Ser Cys Leu Gln Arg Ala Glu Asn Pro
                165                 170                 175

Ser Val Pro Leu Thr Phe Pro Ala Leu Ser Ser Ala Pro Asn Leu Ser
            180                 185                 190

Glu Asn Ser Asn Ser Val Phe Arg Ser Leu Tyr Ser Ala Phe Asn Thr
        195                 200                 205

Ile Ser Asp Phe Gly Trp Ser Leu Leu Lys Ser Ser Phe Leu Glu Asp
    210                 215                 220

Asp Arg Thr Met Ile Arg Ser Gly Asp Pro Gly Val Gln Phe Lys Pro
225                 230                 235                 240

Val Val Ile Ser Thr Met Thr Phe Ser Leu Asp His Ile Lys Gln Ile
                245                 250                 255

Lys Thr Lys Leu Gly Val Thr Ile Asn Asp Ala Ile Thr Gly Ile Ile
            260                 265                 270

Phe Phe Gly Thr Arg Leu Tyr Met Gln Gly Thr Ser Ser Asp Lys Leu
        275                 280                 285
```

```
Ser Asn Gly Asn Ser Thr Ala Leu Val Leu Leu Asn Thr Arg Thr Ile
    290                 295                 300

Gly Gly Tyr Lys Ser Val Lys Asp Met Val Lys Pro Asp Ala Asp Ser
305                 310                 315                 320

Pro Trp Gly Asn Gln Phe Gly Phe Leu His Val Ser Leu Pro Glu Leu
                325                 330                 335

Ala Ser Thr Ala Glu Ser Cys Asn Pro Leu Glu Phe Ile Trp Lys Ala
            340                 345                 350

Gln Lys Leu Ile Gln Arg Lys Arg Asn Ser Lys Ala Val Phe Phe Thr
        355                 360                 365

Gly Gln Leu Leu Glu Gly Leu Arg Lys Tyr Arg Gly Pro Glu Ala Thr
    370                 375                 380

Ala Lys Tyr Ile His Ser Thr Leu Lys Asn Ser Ser Met Thr Ile Ser
385                 390                 395                 400

Asn Leu Ile Gly Pro Val Glu Arg Met Ala Leu Ala Asp His Pro Val
                405                 410                 415

Lys Ser Leu Tyr Phe Met Val Val Gly Val Pro Gln Ser Leu Thr Ile
            420                 425                 430

Thr Met Val Ser Tyr Met Gly Lys Leu Thr Val Ala Ile Gly Thr Glu
        435                 440                 445

Lys Asp Phe Val Asp Pro Gln Lys Phe Lys Ser Ser Val Glu Asn Ala
    450                 455                 460

Phe Gln Met Met Leu Lys Ala Ala Gln Ala Ile Ala
465                 470                 475
```

<210> 118
<211> 485
<212> PRT
<213> Theobroma cacao

<400> 118

Met Ala Cys Pro Gly Thr Phe Ser Asp Glu Pro Leu Thr Pro Ala Gly
1 5 10 15

Arg Leu Phe Leu Gln Ser Gln Thr Asn Ile Ile Ile His Cys Ile Leu
20 25 30

Arg Gly Lys Asn Pro Ile Asp Ile Asp Ala Val Lys Ser Thr Leu Arg
35 40 45

Ser Ser Leu Met Val Arg His Pro Arg Phe Cys Ser Leu Leu Val Arg
50 55 60

Asp Lys Asn Gly Phe Glu His Trp Arg Lys Ala Gln Val Asp Ile Asp
65 70 75 80

Gln His Val Ile Val Ile Asp Lys Arg Leu Asp Lys Ser Asp Asn Phe
85 90 95

Leu Asp Glu Ser Arg Arg Phe Ser Ser Glu Glu Gly Glu Gly Asp Asp
100 105 110

Asp Asp Glu Ala Ala Val Asn Gln Tyr Val Ala Asp Leu Ser Val Ser
115 120 125

Ser Pro Leu Ser Thr Asp Lys Pro Leu Trp Glu Ile His Ile Leu Val
130 135 140

Pro His Lys Cys Ala Val Phe Arg Ile His His Ala Leu Gly Asp Gly
145 150 155 160

Ile Ser Leu Met Ser Met Leu Met Ala Ser Cys Arg Arg Ala Asp Asp
165 170 175

```
Ala Asp Ala Leu Pro Arg Met Val Pro Glu Lys Ser Ala Glu Phe Lys
            180                 185                 190

Asp Gly Lys Gly Arg Asp Trp Phe Trp Leu Phe Gly Ile Leu Trp Gly
        195                 200                 205

Phe Leu Lys Met Val Cys Phe Thr Val Met Phe Val Met Glu Phe Val
    210                 215                 220

Val Ser Ser Leu Leu Val Cys Asp Arg Lys Thr Val Ile Ser Gly Gly
225                 230                 235                 240

Glu Gly Val Glu Leu Trp Pro Arg Lys Leu Ala Thr Ala Arg Phe Leu
                245                 250                 255

Leu Glu Asp Met Arg Val Val Lys Lys Ala Ile Pro Asn Thr Thr Ile
            260                 265                 270

Asn Asp Val Leu Phe Gly Val Val Ser Ser Gly Leu Ser Arg Tyr Leu
        275                 280                 285

Asp His Arg Thr Pro Asn Ala Leu His Glu Gly Leu Arg Met Thr Gly
    290                 295                 300

Val Ala Met Val Asn Leu Arg Pro Gln Thr Gly Phe Gln Asp Leu Ser
305                 310                 315                 320

Gln Leu Met Lys Gly Ala Ala Glu Ala Arg Trp Gly Asn Lys Phe Gly
                325                 330                 335

Leu Ile Leu Leu Pro Val Tyr Phe Lys Lys Ala Gly Asn Asn Pro Leu
            340                 345                 350

Glu Tyr Val Lys Arg Ala Lys Arg Met Val Asp Arg Lys Lys His Ser
        355                 360                 365

Leu Glu Ala Tyr Phe Ser Tyr Arg Ile Gly Asp Leu Val Met Ser Leu
```

213

370 375 380

Leu Gly Ser Lys Tyr Ala Cys Leu Leu Asn Tyr Arg Leu Leu Cys Asn
385 390 395 400

Thr Thr Phe Thr Ile Ser Asn Ile Ile Gly Pro Leu Glu Glu Ile Thr
405 410 415

Leu Ala Gly Asn Pro Ile Ser Ser Ile Lys Val Asn Thr Ser Ser Leu
420 425 430

Pro Gln Ala Ile Thr Met His Met Leu Ser Tyr Ala Gly Arg Ala Glu
435 440 445

Met Gln Ile Leu Val Ala Lys Asp Ile Val Pro Asp Pro Glu Phe Leu
450 455 460

Ala Lys Cys Phe Glu Asp Ala Leu Leu Glu Met Lys Glu Ala Val Ile
465 470 475 480

Gly Thr Glu Lys Glu
485

<210> 119
<211> 540
<212> PRT
<213> Theobroma cacao

<400> 119

Met Asn Arg Lys Ser Glu Ala Lys Trp Val Pro Lys Leu Phe His Gln
1 5 10 15

Ser Leu Ile Ser Met Asn Ser Thr Glu Ile Gln Glu Ile Asp Gln Val
20 25 30

Glu Ala Lys Ile Gly Thr Glu Glu Ala Ser Pro Val Ser Lys Asp Glu
35 40 45

```
Val Pro Leu Ser Thr Gly Ser Leu Ile Phe His Glu Pro Gly Ile Asn
    50                  55                  60

Cys Cys Ile Ile Ala Ile Met Gly Tyr Lys Ser Lys Leu Asp Pro Ala
65                  70                  75                  80

Val Ile Lys Glu Gly Leu Lys Gln Thr Leu Ile Lys His Pro Cys Phe
                85                  90                  95

Ser Ser Lys Leu Leu Ser Ala Ile Asp Gln Met Met Asn Ser Gly Lys
            100                 105                 110

Lys Lys Trp Ile Arg Thr Gln Val Asn Val Asp Asn His Val Ile Leu
        115                 120                 125

Pro Glu Val Asp Thr Val Thr Glu Ser Pro Ser Gln Phe Ile Glu Asp
    130                 135                 140

Tyr Val Thr Tyr Leu Thr Gly Ile Pro Ile Asp Leu Ser Lys Pro Leu
145                 150                 155                 160

Trp Glu Leu His Leu Val Asn Leu Lys Thr Pro Glu Ala Glu Ala Val
                165                 170                 175

Gly Ile Phe Arg Ile His His Ser Ile Gly Asp Gly Met Ser Leu Ile
            180                 185                 190

Ser Leu Leu Leu Ala Cys Cys Arg Lys Ser Ser Asp Pro Lys Ala Leu
        195                 200                 205

Pro Thr Leu Pro Lys Gln Lys Gln Ala Asp Ser Arg Tyr Pro His Pro
    210                 215                 220

His Gly Phe Phe Arg Leu Phe Leu Ala Ile Trp Ser Val Leu Arg Leu
225                 230                 235                 240

Ile Leu Asn Thr Leu Val Asp Leu Leu Leu Phe Val Ala Thr Val Val
                245                 250                 255
```

```
Phe Leu Lys Asp Thr Lys Thr Pro Leu Lys Gly Ser Ser Gly Val Glu
            260                 265                 270

Gln Asn Ala Lys Lys Ile Val His Arg Thr Val Ser Leu Asp Asp Ile
        275                 280                 285

Lys Leu Val Lys Asp Ala Met Gly Met Thr Val Asn Asp Val Ile Leu
    290                 295                 300

Gly Val Thr Glu Ala Gly Leu Ser Arg Tyr Leu Asn Arg Lys Tyr Gly
305                 310                 315                 320

Glu Val Asp Arg Gly Lys Glu Ala Glu Gln Lys Ser Asn His Leu Pro
                325                 330                 335

Arg Asn Ile Arg Leu Arg Ala Thr Ala Leu Val Asn Ile Arg Gln Thr
            340                 345                 350

Ala Gly Ile Gln Val Ser Phe Leu Ser Pro Phe Ser Arg Ala Ser Glu
        355                 360                 365

Lys Val Glu Tyr Ala Ser Gln Ser Phe Leu Val Phe Lys Ser Lys Val
    370                 375                 380

Lys Trp Gly Asn Gln Ile Gly Tyr Ile Cys Ile Pro Phe Thr Ile Ala
385                 390                 395                 400

Leu Arg Asn Asp Pro Leu Asp Tyr Leu Arg Gly Ala Lys Ala Ala Gly
                405                 410                 415

Asp Arg Lys Lys Leu Ser Leu Glu Ala Ile Cys Thr His Leu Thr Asn
            420                 425                 430

Lys Cys Val Val Lys Leu Phe Gly Ser Lys Leu Ser Ala Ala Leu Val
        435                 440                 445
```

```
Tyr Arg Val Ile Phe Asn Thr Thr Met Thr Leu Ser Asn Val Val Gly
    450                 455                 460

Pro Val Glu Glu Ile Ser Leu Tyr Gly His Pro Ile Ala Phe Ile Ala
465                 470                 475                 480

Pro Ser Val Tyr Gly His Pro Gln Ala Leu Thr Cys His Phe Gln Ser
                485                 490                 495

Tyr Met Asn Lys Met Ser Ile Val Leu Ala Val Asp Pro Asn Val Ile
            500                 505                 510

Pro Asp Pro Arg Leu Leu Cys Asp Asp Leu Glu Glu Ser Leu Lys Ile
        515                 520                 525

Phe Lys Asp Ala Val Val Leu Ala Lys Asp Ala Asp
    530                 535                 540
```

<210> 120
<211> 510
<212> PRT
<213> Theobroma cacao

<400> 120

```
Met Glu Ser Arg Asn Gly Phe Arg Trp Arg Lys Gln Gly Leu Lys Pro
1               5                   10                  15

Ile Glu Thr Lys Thr Gly Ser Lys Gly Glu Glu Glu Asn Lys Ser Ser
            20                  25                  30

Ser Gly Gln Arg Thr Ala Glu Glu Glu Pro Leu Ser Pro Ser Ala Arg
        35                  40                  45

Leu Phe His Glu Pro Asn Phe Asn Val Tyr Ile Ile Ala Thr Leu Gly
    50                  55                  60

Cys Lys Thr Arg Ile Tyr Pro Asp Val Val Lys Ala Asn Leu Gly His
65                  70                  75                  80
```

```
Thr Leu Leu Lys His Pro Arg Phe Ser Ser Leu Gln Val Glu Asp Glu
                85                  90                  95

Asp Asn Lys Gly Glu Met Lys Trp Val Arg Thr Glu Val Asp Leu Glu
            100                 105                 110

Arg His Val Val Val Pro Asp Leu Asp Pro Asn Met Asp Leu Pro Asp
        115                 120                 125

Lys Phe Leu Glu Asp Tyr Ile Tyr Asn Leu Ser Lys Thr Thr Ile Asp
    130                 135                 140

Lys Ser Gln Pro Leu Trp Asp Leu His Leu Leu Asn Leu Arg Thr Ser
145                 150                 155                 160

Gln Ser Glu Ala Val Gly Val Phe Arg Ile His His Ser Leu Gly Asp
                165                 170                 175

Gly Thr Ser Leu Met Ser Leu Leu Leu Ala Cys Thr Arg Gln Met Asn
            180                 185                 190

Asp Pro Leu Ala Leu Pro Thr Ile Pro Ile Arg Lys Lys Gln Glu Arg
        195                 200                 205

Lys Asn Asp Arg Arg Gly Phe Trp Arg Ile Met Phe Met Leu Cys Ser
    210                 215                 220

Val Phe Gln Val Phe Trp Asn Ser Val Val Asp Val Phe Met Phe Met
225                 230                 235                 240

Ala Thr Ala Leu Leu Leu Lys Asp Thr Glu Asn Pro Phe Lys Gly Leu
                245                 250                 255

Pro Gly Val Glu Phe Thr Pro Arg Arg Ile Val Tyr Gly Thr Val Ser
            260                 265                 270

Leu Asp Asp Ile Lys Leu Val Lys Asn Ala Met Asn Thr Thr Val Asn
```

```
        275                 280                 285

Asp Val Ala Leu Gly Ile Thr Gln Ala Gly Leu Ser Arg Tyr Ile Asn
    290                 295                 300

Arg Ile Tyr Gly Gly Asn Lys Asn Asp Gly Ala Ala Ile Glu Met Asp
305                 310                 315                 320

Asn Leu Pro Lys Ser Ile Arg Leu Arg Ser Thr Leu Leu Val Asn Ile
                325                 330                 335

Arg Pro Ser Ala Gly Ile Gln Ala Leu Ala Asp Met Met Glu Lys Asp
            340                 345                 350

Ala Glu Ala Lys Trp Gly Asn Trp Ile Gly Tyr Val Leu Leu Pro Phe
        355                 360                 365

Thr Ile Ala Ile Arg Asp Asp Pro Leu Asp Tyr Val Arg Asp Ala Lys
    370                 375                 380

Ala Thr Ile Asp Arg Lys Lys Arg Ser Leu Glu Ala Ile Tyr Thr Phe
385                 390                 395                 400

Ser Ile Ala Glu Leu Val Leu Lys Leu Phe Gly Val Lys Ala Ala Ser
                405                 410                 415

Ala Leu Ser His Arg Ile Met Ser His Thr Thr Met Cys Phe Ser Asn
            420                 425                 430

Leu Val Gly Pro Leu Glu Glu Ile Gly Phe Tyr Gly His Pro Ile Ala
        435                 440                 445

Phe Leu Ala Pro Ser Ser Tyr Gly Gln Pro His Ala Leu Met Ile Asn
    450                 455                 460

Phe Gln Ser Tyr Ile Asp Lys Met Thr Ile Val Leu Ser Val Asp Glu
465                 470                 475                 480
```

```
Gly Thr Ile Pro Asn Pro His Gln Leu Cys Asp Asp Ile Val Glu Ser
                485                 490                 495


Leu Gly Leu Ile Lys Asp Ala Val Val Thr Arg Gly Leu Val
            500                 505                 510


<210>  121
<211>  597
<212>  PRT
<213>  Theobroma cacao

<400>  121

Met Asp Ile Ile Gly Leu Arg Ser Arg Ser His Leu Gly Leu Lys Gln
1               5                   10                  15


Ile Lys Val Thr Arg Asp Met Lys Glu Gly Gln Ser Val Ser Gly Glu
            20                  25                  30


Glu Glu Glu Pro Leu Ser Pro Met Ala Arg Met Phe His Glu Pro Asp
        35                  40                  45


Ser Asn Val Lys Glu Ile Asn Asp Ala Leu His Ala Ile Pro Met Ile
    50                  55                  60


Leu Gln Thr Leu Thr Cys Ser Asn Met Asn Ser Ala Tyr Arg Gln Gly
65                  70                  75                  80


Arg Arg Val Ala Asp Glu Asn Asn Gly Gly Glu Leu Lys Trp Val Lys
                85                  90                  95


Thr Glu Val Glu Leu Glu Lys His Val Ile Ile Pro Arg Val Asp Glu
            100                 105                 110


Glu Met Ala Ser Gln Gly Ala Ala Asp Lys Phe Val Glu Asp Tyr Ile
        115                 120                 125


Ala Asn Ile Ser Lys Thr Lys Ile Ser Leu Ser Ile Pro Met Trp Asp
    130                 135                 140
```

```
Cys His Ile Leu Asn Ile Lys Thr Ser Asp Ala Glu Ser Val Leu Val
145                 150                 155                 160

Leu Leu Val His His Ser Leu Gly Asp Gly Thr Ser Leu Met Ser Leu
                165                 170                 175

Leu Ile Ser Cys Ser Arg Lys Leu Ser Asp Leu Leu Ala Leu Leu Thr
            180                 185                 190

Phe Pro Ala Met Lys Lys Lys Gln Met Pro Thr Thr Thr Thr Thr Thr
        195                 200                 205

Trp Leu Cys Phe Trp Ile Arg Leu Trp Ser Phe Phe Leu Leu Ile Trp
    210                 215                 220

Asn Thr Ser Val Asp Leu Leu Val Cys Val Ala Thr Leu Phe Phe Tyr
225                 230                 235                 240

Lys Asp Thr Pro Thr Pro Leu Lys Pro Ser Ser Arg Arg Val Ala Cys
                245                 250                 255

Thr Pro Gly Arg Ile Met Arg Arg Thr Phe Thr Leu Asp Asp Ile Lys
            260                 265                 270

Leu Val Lys Asn Ala Thr Asn Met Thr Val Asn Asp Val Val Leu Ala
        275                 280                 285

Ile Thr Gln Ala Gly Leu Ser Arg Tyr Leu Asn Arg Lys Tyr Gly Lys
    290                 295                 300

Thr Lys Arg Asn Glu Ala Gly Arg Lys Trp Glu Asp Asn Leu Pro Asn
305                 310                 315                 320

Asn Ile Arg Leu Arg Ala Thr Leu Phe Ile Asn Leu Arg Ser Ser Pro
                325                 330                 335
```

```
Gly Ile Tyr Ala Leu Gly Glu Met Leu Lys Lys Asn Ser Lys Ala Glu
            340                 345                 350

Trp Gly Asn Lys Ile Gly Tyr Val Leu Phe Pro Phe Thr Ile Ser Leu
        355                 360                 365

Lys Asp Asn Leu Leu Asp Tyr Ile Arg Asp Val Lys Ala Ala Thr Asp
    370                 375                 380

Arg Lys Lys Ala Tyr Leu Glu Ala Lys Phe Arg Leu Leu Met Ala Met
385                 390                 395                 400

Val Phe Val Arg Phe Tyr Arg Thr Lys Phe Ala Lys Phe Pro Ser Thr
                405                 410                 415

Thr Met Trp Phe Ser Asn Val Ala Arg Pro Gln Asp Asp Ile Thr Ile
            420                 425                 430

Phe Gly Asn Gln Val Ala Tyr Ile Ala Pro Ser Leu Tyr Gly Gln Leu
        435                 440                 445

Val Ala Leu Thr Val His Val Val Ser Tyr Ala Lys Met Ile Ser Met
    450                 455                 460

Val Leu Ser Val Asp Asp Asn Ile Ile Ser Asp Pro Tyr Gln Leu Cys
465                 470                 475                 480

Asp Asp Leu Glu Glu Ser Leu Lys Leu Ile Lys Thr Ser Val Val Ser
                485                 490                 495

Pro Leu Val Arg Val Arg Leu Ser Tyr Ser Arg Thr Lys Ser Arg Ser
            500                 505                 510

Thr Leu Val Ile Leu Ala Lys Phe Pro Ser Thr Thr Ile Trp Phe Ser
        515                 520                 525

Asn Val Ala Arg Pro Gln Asp Glu Ile Thr Ile Phe Gly Asn Gln Val
    530                 535                 540
```

```
Ala Tyr Ile Ala Ser Ser Val Tyr Gly Gln Pro Lys Ala Leu Thr Ile
545                 550                 555                 560


His Val Val Ser Phe Ala Lys Lys Ile Ser Gln Pro Val Val Leu Leu
                565                 570                 575


Ala Gly His Pro Val Pro Leu Thr Pro Met Leu Pro Arg Arg Ala Gly
            580                 585                 590


Ala Asp Ala Ser Leu
        595
```

<210> 122
<211> 514
<212> PRT
<213> Theobroma cacao

<400> 122

```
Met Asp Ile Ile Asp Leu Arg Ser Gly Ser His Leu Gly Leu Lys Gln
1               5                   10                  15


Ile Lys Val Thr Arg Glu Met Glu Glu Gly Arg Ser Val Ser Gly Glu
            20                  25                  30


Asp Asn Glu Pro Leu Ser Pro Met Ala Arg Met Phe His Glu Pro Asp
        35                  40                  45


Ser Asn Val Tyr Ile Ile Thr Ile Val Gly Phe Lys Asn Pro Ile Glu
    50                  55                  60


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


Arg Phe Ser Ser Val Gln Val Ala Asp Glu Asn Asn Gly Gly Glu Leu
                85                  90                  95


Lys Trp Val Gln Thr Glu Val Glu Leu Glu Lys His Val Ile Val Pro
```

100 105 110

Lys Val Asp Glu Glu Met Ala Ser Gln Gly Ala Ala Asp Lys Phe Ile
115 120 125

Glu Asp Tyr Ile Ser Asn Met Ser Lys Thr Lys Ile Ser Met Ser Ile
130 135 140

Pro Met Trp Asp Cys His Ile Leu Asn Leu Lys Thr Ser Asp Ala Glu
145 150 155 160

Ser Val Leu Val Leu Arg Val His His Ser Leu Gly Asp Gly Thr Ser
165 170 175

Leu Met Ser Leu Leu Ile Ser Cys Ser Arg Lys Leu Phe Asp Pro Leu
180 185 190

Ala Leu Pro Thr Phe Pro Ala Met Lys Lys Lys Pro Ile Ala Thr Thr
195 200 205

Thr Trp Leu Cys Phe Trp Ile Lys Leu Trp Ser Phe Phe Leu Leu Ile
210 215 220

Trp Asn Thr Leu Val Asp Met Leu Met Cys Val Ala Thr Leu Tyr Phe
225 230 235 240

Tyr Lys Asp Thr Pro Thr Pro Leu Lys Pro Pro Ser Arg Ser Val Ala
245 250 255

Cys Thr Pro Lys Arg Ile Val Arg Arg Thr Phe Ser Leu Asp Asp Val
260 265 270

Lys Leu Val Lys Asn Ala Thr Asn Thr Leu Ile Trp Leu Ile Arg Val
275 280 285

Tyr Leu Lys Cys Phe Ala Ile Gly Pro Phe Val Asp Arg Gln Arg Cys
290 295 300

```
Arg Pro Ser Asn His Ser Ser Arg Leu Ile Ser Leu Asp Ile Ser Thr
305                 310                 315                 320

Val Asn Met Arg His Glu Ala Gly Lys Glu Trp Glu Asp Asn Leu Pro
                325                 330                 335

Asn Asn Ile Arg Leu Arg Ala Thr Leu Phe Ile Asn Leu Arg Ser Ser
            340                 345                 350

Pro Gly Ile Tyr Ala Leu Gly Glu Met Leu Lys Lys Asn Ser Lys Ala
        355                 360                 365

Glu Trp Gly Asn Lys Ile Gly Tyr Val Leu Phe Pro Phe Thr Ile Ala
    370                 375                 380

Leu Lys Asp Asn Pro Leu Glu Tyr Ile Arg Asp Val Lys Glu Ala Met
385                 390                 395                 400

Asp Arg Lys Lys Ala Ser Leu Glu Ala Lys Phe Arg His Leu Met Ala
                405                 410                 415

Thr Val Phe Val Arg Phe Tyr Arg Thr Arg Leu Ala Lys Phe Pro Ser
            420                 425                 430

Thr Thr Met Trp Phe Ser Asn Val Ala Gly Pro Gln Asp Glu Ile Thr
        435                 440                 445

Ile Phe Gly Asn Gln Val Thr Tyr Ile Ala Pro Ser Leu Tyr Gly Gln
    450                 455                 460

Pro Val Ala Leu Thr Ile His Val Val Ser Tyr Ala Lys Lys Met Ser
465                 470                 475                 480

Met Val Leu Ser Val Asp Asp Asn Ile Ile Pro Asp Pro Tyr Gln Leu
                485                 490                 495

Cys Asp Asp Leu Glu Glu Ser Leu Lys Leu Ile Lys Lys Ser Val Ile
```

500 505 510

Ser Gln

<210> 123
<211> 670
<212> PRT
<213> Theobroma cacao

<400> 123

Met Ser Phe Leu Arg Arg Arg Lys Val Thr Asp Ser Ser Lys Ser Gln
1 5 10 15

Asn Ser Asp Ser Asn Ala Gly Lys Glu Asp Asp Lys Lys Lys Gln Tyr
20 25 30

Lys Ser Pro Lys Arg Glu Ser Tyr Ser Ser Lys Lys Trp Ser Cys Trp
35 40 45

Asp Ser Cys Cys Trp Phe Ile Gly Leu Ile Cys Ser Met Trp Trp Phe
50 55 60

Leu Leu Phe Leu Tyr Asn Ala Met Pro Ala Ser Ile Pro Gln Tyr Val
65 70 75 80

Thr Glu Ala Ile Thr Gly Pro Leu Pro Asp Pro Pro Gly Val Lys Leu
85 90 95

Arg Lys Glu Gly Leu Thr Val Asn His Pro Val Val Phe Val Pro Gly
100 105 110

Ile Val Thr Gly Gly Leu Glu Leu Trp Glu Gly His His Cys Ala Asp
115 120 125

Gly Leu Phe Arg Lys Arg Leu Trp Gly Gly Ser Phe Gly Glu Leu Tyr
130 135 140

```
                                                226
Lys Arg Pro Leu Cys Trp Ala Glu His Met Ser Leu Asp Asn Glu Thr
145                 150                 155                 160


Gly Leu Asp Pro Pro Gly Ile Arg Val Arg Pro Val Ser Gly Leu Val
                165                 170                 175


Ala Ala Asp Tyr Phe Ala Ala Gly Tyr Phe Val Trp Ala Val Leu Ile
            180                 185                 190


Ala Asn Leu Ala Gln Ile Gly Tyr Glu Glu Lys Thr Met Tyr Met Ala
        195                 200                 205


Ala Tyr Asp Trp Arg Leu Ser Phe Gln Asn Thr Glu Val Arg Asp Gln
    210                 215                 220


Thr Leu Ser Arg Ile Lys Ser Asn Ile Glu Leu Met Val Ala Thr Asn
225                 230                 235                 240


Gly Gly Lys Lys Val Val Val Ile Pro His Ser Met Gly Val Leu Tyr
                245                 250                 255


Phe Leu His Phe Met Lys Trp Val Glu Ala Pro Val Pro Met Gly Gly
            260                 265                 270


Gly Gly Gly Ser Asp Trp Cys Ala Lys His Ile Lys Ala Val Met Asn
        275                 280                 285


Ile Gly Ala Pro Phe Leu Gly Val Pro Lys Ser Val Ser Gly Leu Phe
    290                 295                 300


Ser Ile Glu Ala Arg Asp Ile Ala Ile Ala Arg Ala Phe Ala Pro Gly
305                 310                 315                 320


Phe Leu Asp Lys Asp Val Phe Gly Leu Gln Thr Phe Gln His Leu Met
                325                 330                 335


Arg Met Thr Arg Thr Trp Asp Ser Thr Met Ser Met Ile Pro Lys Gly
            340                 345                 350
```

```
Gly Asp Thr Ile Trp Gly Gly Leu Asp Trp Ser Pro Glu Gly Gly Ser
        355                 360                 365

Phe Asn Cys Ser Ala Lys Lys Leu Lys Asn Asn Gly Thr Arg Thr Gly
    370                 375                 380

His Asn Ala Asn Ser Asn Leu Gly Asn Met Lys Ser Val Asn Tyr Gly
385                 390                 395                 400

Arg Ile Ile Ser Phe Gly Lys Asp Val Ala Glu Ala His Ser Ser Lys
                405                 410                 415

Ile Glu Arg Val Asp Phe Arg Asp Val Met Lys Gly Gly Lys Leu Ala
            420                 425                 430

Asn Ser Ser Asn Cys Asp Ile Trp Thr Glu Tyr His Glu Met Gly Asn
        435                 440                 445

Gly Ala Ile Lys Ala Val Ala Asp Tyr Lys Val Tyr Thr Ala Gly Ser
    450                 455                 460

Ile Leu Asp Leu Leu His Phe Val Ala Pro Lys Leu Met Ala Arg Gly
465                 470                 475                 480

Gly Ala His Phe Ser Tyr Gly Ile Ala Asp Asn Leu Asp Asp Pro Lys
                485                 490                 495

Tyr Glu His Tyr Lys Tyr Trp Ser Asn Pro Leu Glu Thr Lys Leu Pro
            500                 505                 510

Ser Ala Pro Asp Met Glu Ile Tyr Ser Met Tyr Gly Val Gly Ile Pro
        515                 520                 525

Thr Glu Arg Ala Tyr Val Tyr Lys Leu Thr Thr Ala Thr Asp Cys Tyr
    530                 535                 540
```

```
Ile Pro Phe Gln Ile Asp Thr Ser Ala Glu Gly Gly Ser Glu Gly Ser
545                 550                 555                 560


Cys Leu Lys Gly Gly Val Phe Ser Val Asp Gly Asp Glu Thr Val Pro
                565                 570                 575


Val Leu Ser Ala Gly Phe Met Cys Ala Lys Gly Trp Arg Gly Lys Thr
            580                 585                 590


Arg Phe Asn Pro Ser Gly Ile Arg Thr Tyr Ile Arg Glu Tyr Asn His
        595                 600                 605


Ala Pro Pro Ser Asn Leu Leu Glu Gly Arg Gly Thr Gln Ser Gly Ala
    610                 615                 620


His Val Asp Ile Met Gly Asn Phe Ala Leu Ile Glu Asp Ile Ile Arg
625                 630                 635                 640


Val Ala Ala Gly Ala Thr Gly Gln Asp Leu Gly Gly Asp Arg Val Tyr
                645                 650                 655


Ser Asp Ile Phe Lys Trp Ser Glu Arg Ile Asn Leu Gln Leu
            660                 665                 670


<210>  124
<211>  690
<212>  PRT
<213>  Theobroma cacao

<400>  124

Met Ala Ser Ile Leu Arg Phe Arg Lys Leu Cys Tyr Val Glu Gln Ala
1               5                   10                  15


Val Lys Cys Ala Ser Val Gly Phe Glu Ser Phe Asp Gln Ser Pro Lys
            20                  25                  30


Ile Asp Gln Lys Leu Asp Lys Lys Glu Glu Glu Val Ile Ser Ala Asn
        35                  40                  45
```

```
Asn Phe Ala Leu Glu Ile Ile Lys Lys Arg Lys Gln Pro Arg Lys Gln
    50                  55                  60

Pro Lys Glu Trp Arg Arg Leu Asp Ser Cys Cys Trp Met Ile Gly Tyr
65                  70                  75                  80

Leu Cys Ser Thr Trp Trp Leu Leu Leu Phe Leu Tyr His Ser Leu Pro
                85                  90                  95

Val Thr Leu Leu Gln Val Pro Glu Ser Pro Gly Val Arg Leu Lys Arg
            100                 105                 110

Glu Gly Leu Ala Ala Leu His Pro Val Val Leu Val Pro Gly Ile Val
        115                 120                 125

Thr Gly Gly Leu Glu Leu Trp Glu Gly Gln Pro Cys Ala Asp Gly Leu
    130                 135                 140

Phe Arg Lys Arg Leu Trp Gly Gly Ser Phe Thr Glu Ile Phe Lys Arg
145                 150                 155                 160

Pro Leu Cys Trp Leu Glu His Leu Ser Leu His Ser Glu Thr Gly Leu
                165                 170                 175

Asp Pro Pro Gly Ile Arg Val Arg Ala Val Pro Gly Leu Val Ala Ala
            180                 185                 190

Asp Tyr Phe Ala Pro Gly Tyr Phe Val Trp Ala Val Leu Ile Glu Asn
        195                 200                 205

Leu Ala Lys Ile Gly Tyr Glu Gly Lys Asn Leu His Met Ala Ala Tyr
    210                 215                 220

Asp Trp Arg Leu Ser Phe Gln Asn Thr Glu Ile Arg Asp His Ala Leu
225                 230                 235                 240

Ser Arg Leu Lys Ser Lys Ile Glu Leu Met Tyr Leu Thr Asn Gly Tyr
```

```
                                        230
                245                 250                 255

Lys Lys Val Ala Val Val Pro His Ser Met Gly Val Ile Tyr Phe Leu
            260                 265                 270

His Phe Leu Lys Trp Val Glu Thr Pro Pro Pro Ile Gly Gly Gly Gly
        275                 280                 285

Gly Pro Gly Trp Cys Ala Lys His Ile Lys Ala Ile Met Asn Ile Gly
    290                 295                 300

Pro Ala Phe Leu Gly Val Pro Lys Ala Val Ser Asn Ile Phe Ser Ala
305                 310                 315                 320

Glu Gly Lys Asp Ile Ala Tyr Ile Arg Ala Met Ala Pro Gly Leu Leu
                325                 330                 335

Asp Ser Lys Ile Leu Gly Leu Gln Thr Leu Glu His Val Met Arg Val
            340                 345                 350

Ser Arg Thr Trp Asp Ser Ile Val Ser Leu Val Pro Lys Gly Gly Glu
        355                 360                 365

Thr Ile Trp Gly Asn Met Asp Trp Ser Pro Glu Glu Arg His Ala Cys
    370                 375                 380

Asp Phe Ser Arg Lys Arg His Ser Gln Pro Ser Pro Ile Asp Asn Asn
385                 390                 395                 400

Val Asn Asn Ser Asp Val Lys Arg Gly Phe Arg Val Lys Asp Pro Val
                405                 410                 415

Gln Tyr Gly Arg Ile Ile Ser Phe Gly Lys Thr Ala Ser Gln Leu His
            420                 425                 430

Ser Ser Gln Leu Pro Thr Ile Asp Ser Lys Glu Phe Leu Arg Thr Ser
        435                 440                 445
```

```
Ala Ser Gln Asn Leu Asn Phe Ser Cys Gly Glu Ala Trp Thr Glu Tyr
    450                 455                 460

Asp Glu Met Ser Arg Glu Gly Ile Gln Lys Val Ala Ala Asn Lys Ala
465                 470                 475                 480

Tyr Thr Thr Gln Thr Leu Leu Asp Leu Leu Arg Phe Val Ala Pro Lys
                485                 490                 495

Met Met Leu Arg Ala Glu Ala His Phe Ser His Gly Ile Ala Asp Asp
            500                 505                 510

Leu Asp Asn Pro Lys Tyr Asn His Tyr Lys Tyr Trp Ser Asn Pro Leu
        515                 520                 525

Glu Met Lys Leu Pro Asp Ala Pro Asp Met Glu Ile Tyr Cys Leu Tyr
    530                 535                 540

Gly Val Gly Ile Pro Thr Glu Arg Ser Tyr Val Tyr Lys Leu Ser Pro
545                 550                 555                 560

Thr Ser Gly Cys Lys Ser Ile Pro Tyr Gln Ile Asp Ser Ser Val Asn
                565                 570                 575

Gly Glu Asp Gly Ser Cys Leu Lys Gly Gly Val Tyr Phe Ala Asp Gly
            580                 585                 590

Asp Glu Asn Val Pro Val Leu Ser Ala Gly Phe Met Cys Ala Lys Gly
        595                 600                 605

Trp Arg Gly Arg Thr Arg Phe Asn Pro Ser Gly Ile Ala Thr Tyr Ile
    610                 615                 620

Arg Glu Tyr Arg His Lys Pro Pro Ser Ser Leu Leu Glu Gly Arg Gly
625                 630                 635                 640

Ile Glu Ser Gly Ala His Val Asp Ile Met Gly Asn Val Ala Leu Ile
```

645 650 655

Glu Asp Ile Leu Arg Ile Ala Ser Gly Ala Thr Gly Lys Glu Ile Arg
660 665 670

Gly Asp Arg Ile Tyr Ser Asp Ile Leu Arg Met Ser Glu Arg Ile Asn
675 680 685

Leu Arg
690

<210> 125
<211> 685
<212> PRT
<213> Theobroma cacao

<400> 125

Met Ser Ala Ile Arg Arg Arg Lys Pro Ile Asn Glu Ser Asp Lys Thr
1 5 10 15

Ser Glu Ala Ser Asp Ser Lys Pro His Lys Glu Lys Glu Glu Glu Gln
20 25 30

His Gly Asp Asp Asp Gly Ser Asp Lys Asp Lys Lys Lys Ile Pro Ser
35 40 45

Lys Ile Lys Lys Lys His Glu Glu Lys Gln Pro Lys Trp Ser Cys Leu
50 55 60

Asp Ser Cys Cys Trp Phe Ile Gly Cys Ile Cys Val Thr Trp Trp Leu
65 70 75 80

Leu Leu Phe Leu Tyr Asn Ala Met Pro Ala Ser Phe Pro Gln Tyr Val
85 90 95

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

```
                                                233
Lys Lys Glu Gly Leu Lys Ala Lys His Pro Val Val Phe Val Pro Gly
        115                 120                 125


Ile Val Thr Gly Gly Leu Glu Leu Trp Glu Gly Arg Lys Cys Ala Glu
    130                 135                 140


Gly Leu Phe Arg Lys Arg Phe Trp Gly Gly Thr Phe Gly Glu Val Tyr
145                 150                 155                 160


Lys Arg Pro Leu Cys Trp Val Glu His Met Ser Leu Asp Asn Glu Thr
                165                 170                 175


Gly Leu Asp Pro Cys Gly Ile Arg Val Arg Pro Val Ser Gly Leu Val
            180                 185                 190


Ala Ala Asp Tyr Phe Ala Pro Gly Tyr Phe Val Trp Ala Val Leu Ile
        195                 200                 205


Ala Asn Leu Ala Arg Ile Gly Tyr Glu Glu Lys Thr Met Tyr Met Ala
    210                 215                 220


Ala Tyr Asp Trp Arg Leu Ser Phe Gln Asn Thr Glu Ala Arg Asp Gln
225                 230                 235                 240


Thr Leu Ser Arg Ile Lys Ser Asn Ile Glu Leu Ile Val Ala Thr Asn
                245                 250                 255


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


Phe Leu His Phe Met Lys Trp Val Glu Ala Pro Ala Pro Met Gly Gly
        275                 280                 285


Gly Gly Gly Pro Asp Trp Cys Ser Lys His Ile Lys Ala Val Val Asn
    290                 295                 300


Ile Gly Gly Pro Phe Leu Gly Val Pro Lys Ala Ile Ala Gly Leu Phe
305                 310                 315                 320
```

```
Ser Ala Glu Ala Lys Asp Ile Ala Val Val Arg Ala Ile Ala Pro Gly
                325                 330                 335

Phe Leu Asp Asn Asp Ile Phe Gln Leu Gln Thr Leu Gln His Val Met
            340                 345                 350

Arg Met Ser Arg Ser Trp Asp Ser Thr Met Ser Met Ile Pro Arg Gly
        355                 360                 365

Gly Asp Thr Ile Trp Gly Gly Leu Asp Trp Ser Pro Glu Glu Gly Tyr
    370                 375                 380

Ser Cys Ala Lys Lys Arg Glu Arg Lys Asn Asp Thr Gln Ile Ala Glu
385                 390                 395                 400

Glu Ala Gly Val Glu Ser Ala Val Ser Gln Thr Arg Ser Ala Lys Tyr
                405                 410                 415

Gly Arg Ile Ile Ser Phe Gly Lys Asp Val Ala Glu Ala Pro Ser Ser
            420                 425                 430

Asp Val Glu Arg Ile Asp Phe Arg Asp Ala Val Lys Gly His Ser Ala
        435                 440                 445

Ala Asn Thr Thr Cys Arg Asp Val Trp Thr Glu Tyr His Asp Met Gly
    450                 455                 460

Phe Gly Gly Ile Asn Ala Val Ala Glu Tyr Lys Thr Tyr Thr Ala Glu
465                 470                 475                 480

Ser Ile Val Asp Leu Leu His Phe Val Ala Pro Lys Met Met Ala Arg
                485                 490                 495

Gly Thr Ala His Phe Ser Tyr Gly Ile Ala Asp Asn Leu Asp Asp Pro
            500                 505                 510
```

```
                                                235
Lys Tyr Lys His Tyr Lys Tyr Trp Ser Asn Pro Leu Glu Thr Lys Leu
        515                 520                 525


Pro Asn Ala Pro Glu Met Glu Ile Phe Ser Leu Tyr Gly Val Gly Leu
    530                 535                 540


Pro Thr Glu Arg Ala Tyr Ile Tyr Lys Leu Ser Pro Ala Ala Glu Cys
545                 550                 555                 560


Tyr Ile Pro Phe Gln Ile Asp Thr Ser Ala Asp Asp Glu Glu Thr Cys
                565                 570                 575


Leu Lys Asp Gly Val Tyr Ser Val Asp Gly Asp Glu Thr Val Pro Val
            580                 585                 590


Leu Ser Ala Gly Phe Met Cys Ala Lys Gly Trp Arg Gly Lys Thr Arg
        595                 600                 605


Phe Asn Pro Ser Gly Ile Arg Thr Tyr Ile Arg Glu Tyr Ser His Ser
    610                 615                 620


Pro Pro Ala Asn Leu Leu Glu Gly Arg Gly Thr Leu Ser Gly Ala His
625                 630                 635                 640


Val Asp Ile Met Gly Asn Phe Ala Leu Ile Glu Asp Val Ile Arg Val
                645                 650                 655


Ala Ala Gly Ala Ser Gly Glu Glu Leu Gly Gly Asp Gln Val Tyr Ser
            660                 665                 670


Asn Ile Phe Asn Trp Ser Glu Lys Ile Asp Leu Gln Leu
        675                 680                 685


<210>  126
<211>  500
<212>  PRT
<213>  Theobroma cacao

<400>  126
```

```
Met Ala Pro Ala Lys Ser Gly Arg Ser Phe Pro Ser Ile Thr Lys Cys
1               5                  10                  15

Glu Gly Ser Thr Tyr Glu Ser Ile Ala Ala Asp Leu Asp Gly Thr Leu
            20                  25                  30

Leu Ile Ser Arg Ser Ser Phe Pro Tyr Phe Met Leu Val Ala Val Glu
        35                  40                  45

Ala Gly Ser Leu Leu Arg Gly Leu Ile Leu Leu Leu Ser Leu Pro Leu
    50                  55                  60

Val Ile Val Ser Tyr Leu Phe Ile Ser Glu Ala Ile Gly Ile Gln Ile
65                  70                  75                  80

Leu Ile Phe Ile Ser Phe Ala Gly Leu Lys Ile Arg Asp Ile Glu Leu
                85                  90                  95

Val Ser Arg Ala Val Leu Pro Arg Phe Tyr Ala Ala Asn Val Arg Lys
            100                 105                 110

Glu Ser Phe Glu Val Phe Asp Arg Cys Lys Arg Lys Val Val Val Thr
        115                 120                 125

Ala Asn Pro Thr Phe Met Val Glu Pro Phe Val Lys Asp Phe Leu Gly
    130                 135                 140

Gly Asp Lys Val Leu Gly Thr Glu Ile Glu Val Asn Pro Lys Thr Lys
145                 150                 155                 160

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

Lys Arg Leu Ala Ile Leu Lys Glu Phe Gly Glu Glu Ser Pro Asp Leu
            180                 185                 190

Gly Ile Gly Asp Arg Glu Ser Asp His Asp Phe Met Ser Ile Cys Lys
```

```
                                        237
        195                 200                 205

Glu Gly Tyr Met Val His Pro Ser Lys Ser Ala Thr Pro Val Pro Leu
    210                 215                 220

Asp Arg Leu Lys Ser Arg Ile Ile Phe His Asp Gly Arg Phe Val Gln
225                 230                 235                 240

Arg Pro Asp Pro Leu Asn Ala Leu Ile Thr Tyr Leu Trp Leu Pro Phe
                245                 250                 255

Gly Phe Ile Leu Ser Ile Val Arg Val Tyr Phe Asn Leu Pro Leu Pro
            260                 265                 270

Glu Arg Ile Val Arg Tyr Thr Tyr Glu Met Leu Gly Ile His Leu Val
        275                 280                 285

Ile Arg Gly Lys Arg Pro Pro Pro Pro Ser Pro Gly Thr Pro Gly Asn
    290                 295                 300

Leu Tyr Val Cys Asn His Arg Ser Ala Leu Asp Pro Ile Val Ile Ala
305                 310                 315                 320

Ile Ala Leu Gly Arg Lys Val Ser Cys Val Thr Tyr Ser Val Ser Arg
                325                 330                 335

Leu Ser Arg Phe Leu Ser Pro Ile Pro Ala Val Ala Leu Thr Arg Asp
            340                 345                 350

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

Leu Val Val Cys Pro Glu Gly Thr Thr Cys Arg Glu Gln Phe Leu Leu
    370                 375                 380

Arg Phe Ser Ala Leu Phe Ala Glu Leu Ser Asp Arg Ile Val Pro Val
385                 390                 395                 400
```

```
Ala Val Asn Cys Arg Gln Asn Met Phe Tyr Gly Thr Thr Val Arg Gly
                405                 410                 415

Val Lys Phe Trp Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Thr
            420                 425                 430

Tyr Glu Val Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Val Lys
        435                 440                 445

Ala Gly Gly Lys Ser Ser Ile Glu Val Ala Asn His Val Gln Lys Val
    450                 455                 460

Leu Gly Asp Val Leu Gly Phe Glu Cys Thr Gly Leu Thr Arg Lys Asp
465                 470                 475                 480

Lys Tyr Leu Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Met Tyr
                485                 490                 495

Asn Ala Lys Lys
            500
```

<210> 127
<211> 376
<212> PRT
<213> Theobroma cacao

<400> 127

```
Met Ser Ser Arg Gly Gly Lys Leu Ser Ser Ser Ser Ser Glu Leu Asp
1               5                   10                  15

Leu Asp Gly Pro Asn Ile Glu Asp Tyr Leu Pro Ser Gly Ser Ser Ile
            20                  25                  30

Asn Glu Pro Arg Gly Lys Leu Arg Leu Arg Asp Leu Leu Asp Ile Ser
        35                  40                  45

Pro Thr Leu Thr Glu Ala Ala Gly Ala Ile Val Asp Asp Ser Phe Thr
    50                  55                  60
```

```
Arg Cys Phe Lys Ser Asn Pro Pro Glu Pro Trp Asn Trp Asn Val Tyr
65                  70                  75                  80

Leu Phe Pro Leu Trp Cys Phe Gly Val Ala Val Arg Tyr Leu Ile Leu
                85                  90                  95

Phe Pro Ala Arg Val Val Val Leu Thr Ile Gly Trp Ile Ile Phe Leu
            100                 105                 110

Ser Ser Phe Ile Pro Val His Phe Leu Leu Lys Gly His Asp Lys Leu
        115                 120                 125

Arg Lys Lys Met Glu Arg Val Leu Val Glu Leu Met Cys Ser Phe Phe
    130                 135                 140

Val Ala Ser Trp Thr Gly Val Val Lys Tyr His Gly Pro Arg Pro Ser
145                 150                 155                 160

Ile Arg Pro Lys Gln Val Phe Val Ala Asn His Thr Ser Met Ile Asp
                165                 170                 175

Phe Ile Ile Leu Glu Gln Met Thr Ala Phe Ala Val Ile Met Gln Lys
            180                 185                 190

His Pro Gly Trp Val Gly Leu Leu Gln Ser Thr Ile Leu Glu Ser Val
        195                 200                 205

Gly Cys Ile Trp Phe Asn Arg Ser Glu Ala Lys Asp Arg Glu Ile Val
    210                 215                 220

Ala Lys Lys Leu Arg Asp His Val Gln Gly Val Asp Asn Asn Pro Leu
225                 230                 235                 240

Leu Ile Phe Pro Glu Gly Thr Cys Ile Asn Asn Gln Tyr Ser Val Met
                245                 250                 255
```

```
                                            240
Phe Lys Lys Gly Ala Phe Glu Leu Gly Cys Thr Val Cys Pro Ile Ala
            260                 265                 270


Ile Lys Tyr Asn Lys Ile Phe Val Asp Ala Phe Trp Asn Ser Arg Lys
        275                 280                 285


Gln Ser Phe Thr Met His Leu Leu Gln Leu Met Thr Ser Trp Ala Val
    290                 295                 300


Val Cys Asp Val Trp Tyr Leu Glu Pro Gln Asn Leu Arg Pro Gly Glu
305                 310                 315                 320


Thr Pro Ile Glu Phe Ala Glu Arg Val Arg Asp Ile Ile Ser Val Arg
                325                 330                 335


Ala Gly Leu Lys Lys Val Pro Trp Asp Gly Tyr Leu Lys Tyr Ser Arg
            340                 345                 350


Pro Ser Pro Lys His Arg Glu Arg Lys Gln Gln Ser Phe Ala Glu Ser
        355                 360                 365


Val Leu Leu Arg Leu Glu Glu Lys
    370                 375


<210>  128
<211>  540
<212>  PRT
<213>  Theobroma cacao

<400>  128

Met Val Phe Pro Val Val Phe Leu Lys Leu Ala Asp Trp Val Leu Tyr
1               5                   10                  15


Gln Leu Leu Ala Asn Ser Cys Tyr Arg Ala Ala Arg Lys Met Arg Asn
            20                  25                  30


Tyr Gly Phe Phe Leu Arg Asn Gln Thr Leu Arg Ser Pro Pro Gln Gln
        35                  40                  45
```

```
Gln Ala Ala Ser Leu Phe Pro Ser Val Thr Lys Cys Asp Val Gly Asn
    50                  55                  60

Ser Arg Arg Phe Asp Thr Leu Val Cys Asp Ile His Gly Val Leu Leu
65                  70                  75                  80

Gly Ser Asp Thr Phe Phe Pro Tyr Phe Met Leu Val Ala Phe Glu Gly
                85                  90                  95

Gly Ser Ile Val Arg Ala Phe Leu Leu Leu Leu Ser Cys Ser Phe Leu
            100                 105                 110

Trp Val Leu Asp Ser Glu Leu Lys Leu Arg Ile Met Ile Phe Ile Ser
        115                 120                 125

Phe Cys Gly Leu Arg Lys Lys Asp Ile Glu Ser Val Gly Arg Ala Val
    130                 135                 140

Leu Pro Lys Phe Tyr Leu Glu Asn Leu Asn Leu Gln Val Tyr Glu Val
145                 150                 155                 160

Trp Ser Lys Thr Ser Ser Arg Val Val Phe Thr Ser Ile Pro Arg Val
                165                 170                 175

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

Gly Thr Glu Leu His Thr Val Gly Asn Arg Phe Thr Gly Leu Leu Ser
        195                 200                 205

Ser Ser Gly Leu Leu Val Lys His Asn Ala Leu Lys Glu His Phe Gly
    210                 215                 220

Asp Lys Lys Pro Asp Val Gly Leu Gly Ser Ser Ser Leu His Asp Gln
225                 230                 235                 240

Tyr Phe Ile Ser Leu Cys Lys Glu Ala Tyr Val Val Asn Met Glu Asp
```

245 250 255

Gly Lys Ser Asn Leu Ser Ser Phe Met Pro Arg Asp Lys Tyr Pro Lys
260 265 270

Pro Leu Ile Phe His Asp Gly Arg Leu Ala Phe Leu Pro Thr Pro Phe
275 280 285

Ala Thr Leu Ser Met Phe Leu Trp Leu Pro Phe Gly Ile Val Leu Ser
290 295 300

Ile Leu Arg Ile Phe Val Gly Ile Cys Leu Pro Tyr Lys Leu Ala Val
305 310 315 320

Ile Cys Ala Thr Leu Ser Gly Val Gln Leu Lys Phe Gln Gly Cys Phe
325 330 335

Pro Ser Ser Asn Ser Gln His Lys Lys Gly Val Leu Tyr Val Cys Thr
340 345 350

His Arg Thr Leu Leu Asp Pro Val Phe Leu Ser Thr Ala Leu Cys Lys
355 360 365

Pro Leu Thr Ala Val Thr Tyr Ser Leu Ser Lys Met Ser Glu Leu Ile
370 375 380

Ala Pro Ile Lys Thr Val Arg Leu Thr Arg Asp Arg Lys Gln Asp Gly
385 390 395 400

Glu Thr Met Gln Lys Leu Leu Ser Glu Gly Asp Leu Val Val Cys Pro
405 410 415

Glu Gly Thr Thr Cys Arg Glu Pro Tyr Leu Leu Arg Phe Ser Ser Leu
420 425 430

Phe Ala Glu Leu Ala Asp Glu Ile Val Pro Val Ala Ile Asn Ala His
435 440 445

Val Ser Met Phe Tyr Gly Thr Thr Ala Ser Gly Leu Lys Cys Leu Asp
450 455 460

Pro Ile Phe Phe Leu Met Asn Pro Arg Pro Ser Tyr His Val Gln Ile
465 470 475 480

Leu Gly Lys Val Pro Gln Glu Phe Thr Cys Ala Gly Gly Arg Ser Ser
485 490 495

Leu Glu Val Ala Asn Tyr Ile Gln Arg Lys Leu Ala Asp Ala Leu Gly
500 505 510

Phe Glu Cys Thr Thr Leu Thr Arg Arg Asp Lys Tyr Leu Met Leu Ala
515 520 525

Gly Asn Glu Gly Val Val His Glu Asn Lys Arg Asn
530 535 540

<210> 129
<211> 537
<212> PRT
<213> Theobroma cacao

<400> 129

Met Ala Lys Leu Ser Met Glu Phe Ser Phe Phe Gln Thr Leu Phe Phe
1 5 10 15

Leu Phe Cys Arg Val Val Phe Arg Gln Ser Lys Asn His Lys Ser Leu
20 25 30

His Arg Asn Val Ser Asn Ile His Ala Asn Glu Gly Lys Tyr His Lys
35 40 45

Tyr Pro Ser Phe Val His Arg Ser Asn Leu Ser Asn Gln Thr Leu Val
50 55 60

Phe Ser Val Glu Glu Ala Leu Leu Lys Ser Ser Ser Leu Phe Pro Tyr
65 70 75 80

```
Phe Met Leu Val Ala Phe Glu Ala Gly Gly Leu Leu Arg Ala Phe Ile
                85                  90                  95

Leu Phe Val Leu Tyr Pro Ile Leu Cys Leu Val Ser Glu Glu Met Gly
            100                 105                 110

Leu Lys Ile Met Val Leu Val Cys Phe Phe Gly Ile Lys Lys Lys Ser
        115                 120                 125

Phe Arg Val Gly Ser Ala Val Leu Pro Lys Phe Phe Leu Glu Asp Val
    130                 135                 140

Gly Leu Glu Pro Phe Glu Met Leu Lys Lys Gly Gly Lys Lys Val Ala
145                 150                 155                 160

Val Ser Lys Ile Pro Gln Val Met Ile Glu Ser Phe Leu Lys Asp Tyr
                165                 170                 175

Leu Glu Ile Asp Phe Val Val Gly Arg Glu Leu Lys Glu Phe Cys Gly
            180                 185                 190

Tyr Phe Leu Gly Val Met Glu Glu Lys Lys Arg Ser Lys Ala Ala Leu
        195                 200                 205

Asp Glu Ile Ile Gly Ser Glu Ser Met Gly Phe Asp Val Ile Gly Ile
    210                 215                 220

Ser Gly Leu Lys Lys Ser Leu Asp Tyr His Phe Phe Ser His Cys Lys
225                 230                 235                 240

Glu Ile Tyr Gln Val Arg Lys Ala Asp Lys Arg Asn Trp Arg His Val
                245                 250                 255

Pro Arg Gln Glu Tyr Ser Lys Pro Leu Ile Phe His Asp Gly Arg Leu
            260                 265                 270
```

```
                                            245
Ala Leu Arg Pro Thr Leu Val Ala Ser Leu Thr Met Phe Val Trp Phe
        275                 280                 285

Pro Phe Gly Leu Ala Leu Ser Ile Leu Arg Ala Val Val Gly Leu Met
    290                 295                 300

Leu Pro Tyr Lys Ile Ser Ile Pro Leu Leu Ala Tyr Ser Gly Leu His
305                 310                 315                 320

Leu Phe Leu Ser Thr Pro Glu Ser Ser Leu His Pro Leu Ser Leu Pro
                325                 330                 335

Asn Ser Lys Lys Gln Asn Pro Lys Gly Arg Leu Tyr Val Cys Asn His
            340                 345                 350

Arg Thr Leu Leu Asp Pro Val Tyr Leu Ser Phe Ala Leu Gln Lys Asp
        355                 360                 365

Leu Thr Ala Val Thr Tyr Ser Leu Ser Arg Ile Ser Glu Leu Leu Ala
    370                 375                 380

Pro Ile Lys Thr Val Arg Leu Ala Arg Asp Arg Asp Gln Asp Gly Lys
385                 390                 395                 400

Met Met Glu Lys Met Leu Asn Leu Gly Asp Leu Val Val Cys Pro Glu
                405                 410                 415

Gly Thr Thr Cys Arg Glu Pro Tyr Leu Leu Arg Phe Ser Pro Leu Phe
            420                 425                 430

Ser Glu Met Ser Asp Asp Ile Val Pro Val Ala Met Asp Ser Asn Val
        435                 440                 445

Ser Leu Phe Tyr Gly Thr Thr Ala Ser Gly Leu Lys Cys Leu Asp Pro
    450                 455                 460

Leu Phe Phe Leu Met Asn Pro Arg Pro Ile Tyr Thr Val Gln Ile Leu
465                 470                 475                 480
```

```
Asp Gly Val Ser Gly Leu Tyr Thr Cys His Asp Gly Gln Arg Ser Arg
                485                 490                 495


Phe Lys Val Ala Asn Gln Val Gln Asn Glu Ile Gly Lys Ala Leu Gly
            500                 505                 510


Phe Glu Cys Thr Lys Leu Thr Arg Arg Asp Lys Tyr Leu Ile Met Ala
        515                 520                 525


Gly Asn Glu Gly Ile Ile Ser Gln Thr
    530                 535


<210>  130
<211>  488
<212>  PRT
<213>  Theobroma cacao

<400>  130

Met Glu Ser Val Val Ser Glu Leu Glu Gly Val Leu Leu Lys Asp Pro
1               5                   10                  15


Asp Pro Phe Cys Tyr Phe Met Leu Val Ala Phe Glu Ala Ser Gly Leu
            20                  25                  30


Ile Arg Phe Ala Leu Leu Leu Leu Phe Trp Pro Leu Ile Arg Leu Leu
        35                  40                  45


Asp Met Leu Gly Leu Asp Asp Ala Gly Leu Lys Leu Met Thr Phe Val
    50                  55                  60


Ala Thr Ala Gly Leu Arg Glu Ala Glu Ile Glu Ser Val Ser Arg Ala
65                  70                  75                  80


Val Leu Pro Lys Phe Tyr Met Asp Asp Val Asp Met Glu Ala Trp Lys
                85                  90                  95


Val Phe Ser Ser Tyr Asp Lys Arg Val Val Val Thr Lys Thr Pro Arg
```

100 105 110

Ile Met Val Ala Arg Phe Val Lys Glu His Leu Arg Ala Asp Glu Val
115 120 125

Leu Gly Ser Glu Leu Val Val Asn Arg Phe Gly Phe Ala Thr Gly Phe
130 135 140

Val Lys Gly Asp Ile Ala Ser Ile Ser Ser Arg Val Ala Lys Leu Phe
145 150 155 160

Val Asp Glu Glu Pro Thr Leu Gly Leu Gly Arg Ala Pro Ser Ser Phe
165 170 175

Gln Phe Leu Ser Leu Cys Lys Lys Gln Met His Pro Pro Leu Met Thr
180 185 190

Asn Gln Asn Leu His Asp His Gln Leu Leu Arg Pro Leu Pro Val Ile
195 200 205

Phe His Asp Gly Arg Leu Val Lys Arg Pro Thr Pro Ser Thr Ala Leu
210 215 220

Leu Ile Leu Leu Trp Met Pro Leu Gly Ile Leu Leu Ala Thr Ile Arg
225 230 235 240

Ile Val Val Gly Leu Ile Leu Pro Met Arg Val Ile Pro Tyr Met Ser
245 250 255

Arg Leu Phe Gly Gly Lys Ile Ile Val Lys Gly Lys Gln Pro Pro Pro
260 265 270

Val Ser Gly Ser Ser Ser Gly Val Leu Phe Val Cys Thr His Arg Thr
275 280 285

Leu Met Asp Pro Val Val Leu Ser Ala Val Leu Met Arg Lys Ile Pro
290 295 300

```
Ala Val Thr Tyr Ser Ile Ser Arg Leu Ser Glu Ile Leu Ser Pro Ile
305                 310                 315                 320

Pro Thr Val Arg Leu Thr Arg Ile Arg Glu Val Asp Ala Glu Lys Ile
                325                 330                 335

Lys Arg Glu Leu Ala Lys Gly Asp Leu Val Val Cys Pro Glu Gly Thr
            340                 345                 350

Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu
        355                 360                 365

Leu Thr Asp Arg Ile Val Pro Val Ala Met Asn Tyr Arg Val Gly Phe
    370                 375                 380

Phe His Ala Thr Thr Ala Arg Gly Trp Lys Gly Leu Asp Pro Ile Phe
385                 390                 395                 400

Phe Phe Met Asn Pro Arg Pro Val Tyr Glu Val Thr Phe Leu Asn Gln
                405                 410                 415

Leu Pro Ala Glu Ala Thr Cys Ser Ser Gly Lys Ser Pro His Asp Val
            420                 425                 430

Ala Asn Tyr Val Gln Arg Ile Leu Ala Ala Thr Leu Gly Phe Glu Cys
        435                 440                 445

Thr Asn Phe Thr Arg Lys Asp Lys Tyr Arg Val Leu Ala Gly Asn Asp
    450                 455                 460

Gly Thr Val Ser Cys Thr Ser Phe Val Asp Gln Val Lys Lys Val Val
465                 470                 475                 480

Ser Thr Phe Lys Pro Phe Phe Gln
                485
```

<210> 131

<211> 593
<212> PRT
<213> Theobroma cacao

<400> 131

```
Met Lys Leu Ser Arg Arg Gln Thr Val Leu Thr Asn Asp Asn Asn Asn
1               5                   10                  15


Cys Asn Leu Lys Leu Ala Ala Trp Pro Cys Ser Phe Gly Tyr Lys Arg
            20                  25                  30


Arg Lys Ala Arg Gln Thr Ser Ala Ile Gln Pro Lys Tyr Gln Val Pro
        35                  40                  45


Thr Ala Ser Tyr Ser Ser Ala Thr Ala Gln Arg Ser Lys Ala Ser Pro
    50                  55                  60


Arg Leu Ile His Trp Leu Ala Val Lys Asp Ser Trp Lys Ser Arg Ala
65                  70                  75                  80


Leu Asn Val Leu Ile Leu Leu Phe Leu Asn Met Ala Ile Asn Ala Phe
                85                  90                  95


Pro Thr Val Gln Asn Cys Ala Ser Ile Gly Arg Glu Lys His Thr Met
            100                 105                 110


Val Ala Asp Met Asp Gly Thr Leu Leu Ile Ser Arg Ser Ser Phe Pro
        115                 120                 125


Tyr Phe Ala Leu Val Ala Phe Glu Val Gly Gly Val Leu Arg Leu Phe
    130                 135                 140


Phe Leu Leu Leu Ala Ser Pro Leu Ala Gly Leu Leu Tyr Tyr Phe Val
145                 150                 155                 160


Ser Glu Ser Ala Gly Ile Gln Val Leu Ile Phe Ala Thr Tyr Val Gly
                165                 170                 175
```

```
                                            250
Met Lys Val Ser Asp Ile Glu Ser Val Ala Arg Ala Val Leu Pro Lys
            180                 185                 190


Phe Tyr Ser Ser Asp Leu His Pro Glu Ser Trp Arg Val Phe Ser Ser
        195                 200                 205


Cys Gly Lys Arg Cys Val Leu Thr Ala Asn Pro Arg Ile Met Val Glu
    210                 215                 220


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


Ile Gly Ile Tyr Arg Gly Arg Ala Thr Gly Phe Val Cys Gln Pro Gly
                245                 250                 255


Val Leu Val Gly Lys Asn Lys Ala Asp Ala Leu Gln Lys Ala Phe Gly
            260                 265                 270


Glu Ala Leu Pro Asp Val Gly Leu Gly Asp Arg His Thr Asp Ile Pro
        275                 280                 285


Phe Ile Ala Leu Cys Lys Glu Gly Tyr Ile Val Pro Pro Lys Pro Glu
    290                 295                 300


Ala Lys Ala Val Arg Ser Glu Lys Leu Pro Lys Pro Val Ile Phe His
305                 310                 315                 320


Asp Gly Arg Leu Val Gln Lys Pro Thr Pro Leu Met Ala Leu Leu Ile
                325                 330                 335


Ile Ile Trp Ile Pro Ile Gly Phe Leu Leu Ala Cys Leu Arg Ile Thr
            340                 345                 350


Ala Gly Ser Leu Leu Pro Met Pro Leu Val Tyr Tyr Ala Phe Trp Ala
        355                 360                 365


Leu Gly Val Arg Val Thr Val Lys Gly Thr Pro Pro Pro Pro Val Lys
    370                 375                 380
```

```
Lys Ser Thr Gly Gln Ser Gly Ile Leu Phe Ile Cys Ser His Arg Thr
385                 390                 395                 400

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

Val Val Thr Tyr Ser Val Ser Arg Leu Ser Glu Ile Ile Ser Pro Ile
            420                 425                 430

Lys Thr Val Ser Leu Ser Arg Asp Arg Ala Thr Asp Ala Ser Met Ile
        435                 440                 445

Lys Lys Leu Leu Gln Glu Gly Asp Leu Ala Ile Cys Pro Glu Gly Thr
    450                 455                 460

Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu
465                 470                 475                 480

Leu Thr Asp Gln His Val Pro Val Ala Met Val Asn Arg Met Ser Met
                485                 490                 495

Phe His Gly Thr Thr Ala Arg Gly Trp Lys Gly Met Asp Pro Phe Tyr
            500                 505                 510

Phe Phe Met Asn Pro Ser Pro Ala Tyr Glu Val Thr Phe Leu Asn Lys
        515                 520                 525

Leu Pro Met Glu Leu Thr Cys Ser Ser Gly Lys Ser Ser His Glu Ile
    530                 535                 540

Ala Asn Tyr Ile Gln Arg Val Ile Ala Ala Thr Leu Ser Tyr Glu Cys
545                 550                 555                 560

Thr Ser Phe Thr Arg Lys Asp Lys Tyr Arg Ala Leu Ala Gly Asn Asp
                565                 570                 575
```

Gly Thr Val Val Glu Lys Pro Lys Phe Ser Ala Asn Lys Val Met Gly
580 585 590

Cys

<210> 132
<211> 438
<212> PRT
<213> Theobroma cacao

<400> 132

Met Gly Ser Arg Ile Leu Pro Ser Ala Ser Leu Phe Leu Cys Ser Val
1 5 10 15

Phe Leu Ser Ser Phe Cys Ala Ser Cys Arg Ser Ser Leu Ser Ile Pro
20 25 30

Pro Leu Arg Ala Ser Phe Pro Cys Ala Ser Ser Lys Ser Leu Ala Pro
35 40 45

Ile Cys Arg Leu Pro Leu Pro His Ala Cys Gly Leu Gly Ile Leu Val
50 55 60

Leu Phe Thr Ala Ser Ser Leu Ala Met Ala Asp Val Val Ala Asp Val
65 70 75 80

Val Ala Pro Val Lys Pro Lys Asn Arg Ala Ser Lys Lys Phe Ser Ser
85 90 95

Arg Gly Val Asp Leu Asp Val Leu Leu Glu Met Ser Ser Ser Asn Ser
100 105 110

Ser Pro Leu Leu Ser Pro Lys Val Phe Leu Arg Asp Leu Lys Arg Arg
115 120 125

Pro Leu Ala Leu Thr Lys Lys Ile Leu Ser Trp Ala Leu Glu Gly Ala
130 135 140

```
Arg Ser Thr Ala Phe Ser Leu Gly Thr Gly Gly Lys Phe His Tyr Pro
145                 150                 155                 160

Ala Arg Arg Glu Phe Phe Ala Ser Gly Arg Leu Leu Phe Leu Leu Leu
                165                 170                 175

Ala Ser Pro Leu Ala Gly Thr Leu Tyr Tyr Phe Thr Ser Glu Ser His
            180                 185                 190

Asp Ile Arg Val Leu Ile Phe Ala Thr Phe Val Gly Met Lys Val Ser
        195                 200                 205

Asp Ile Glu Ser Val Ala Arg Ala Val Leu Pro Lys Phe Tyr Ser Ser
    210                 215                 220

Asp Leu His Pro Gly Phe Arg Phe Met Trp Glu Glu Met Cys Ala Tyr
225                 230                 235                 240

Gly Asp Ser Glu Gly His Ser Gly Val Lys Ser Pro Gly Val Leu Val
                245                 250                 255

Gly Lys Asn Lys Ala Asp Ala Leu Glu Asn Glu Ser His Val Val Pro
            260                 265                 270

Ala Lys Ala Glu Ala Gly Ala Val Ser His Asp Lys Leu Pro Lys Pro
        275                 280                 285

Ile Val Phe His Asp Gly Arg Leu Val Gln Asn Pro Thr Pro Phe Met
    290                 295                 300

Ala Leu Leu Thr Ile Leu Trp Ile Pro Val Gly Phe Leu Leu Ala Cys
305                 310                 315                 320

Leu Arg Ile Ala Val Gly Ile Leu Leu Pro Thr Pro Leu Val Tyr Tyr
                325                 330                 335

Ala Met Trp Pro Leu Gly Phe Arg Ile His Ile Lys Gly Thr Pro Pro
```

340 345 350

Gln Pro Thr Lys Lys Ser Ile Gly Gln Thr Val Val Leu Phe Ala Cys
355 360 365

Ser His Arg Thr Leu Leu Asp Pro Ile Ala Leu Ser Val Ala Leu Cys
370 375 380

Arg Pro Ile Pro Thr Ala Thr Tyr Ser Ile Ser Gly Leu Ser Glu Ile
385 390 395 400

Ile Ser Pro Ile Lys Thr Ala Arg Leu Cys Leu Asp Arg Ala Thr Asp
405 410 415

Ala Ser Ile Ile Lys Lys Leu Leu Glu Lys Gly Asp Leu Val Ile Cys
420 425 430

Pro Glu Gly Thr Lys Thr
435

<210> 133
<211> 507
<212> PRT
<213> Theobroma cacao

<400> 133

Met Ala Lys Thr Lys Glu Asn Pro Phe Pro Ser Ile Ala Gln Cys Ala
1 5 10 15

Ser Ile Gly Arg Asp Lys His Thr Val Val Ala Asp Met Asp Gly Thr
20 25 30

Leu Leu Arg Gly Arg Ser Ser Phe Pro Tyr Phe Ala Leu Val Ala Phe
35 40 45

Glu Val Gly Gly Ile Leu Arg Leu Leu Phe Leu Leu Leu Ala Ser Pro
50 55 60

```
                                                255
Val Ala Gly Ile Leu Tyr Tyr Phe Ile Ser Glu Ala Ala Gly Ile Arg
65                  70                  75                  80


Val Leu Ile Phe Ala Thr Phe Val Gly Met Lys Val Ser Asp Ile Glu
                85                  90                  95


Ser Val Ala Arg Ala Val Leu Pro Lys Phe Tyr Ser Ser Asp Leu His
            100                 105                 110


Pro Glu Thr Trp Arg Val Phe Ser Ser Cys Gly Lys Ser Cys Val Leu
        115                 120                 125


Thr Ala Asn Pro Arg Val Met Val Glu Ala Phe Leu Lys Asp Tyr Leu
    130                 135                 140


Gly Ala Asp Met Val Thr Gly Thr Glu Ile His Val Phe Arg Gly Arg
145                 150                 155                 160


Ala Thr Gly Leu Val Lys Ser Pro Gly Ile Leu Val Gly Lys Asn Lys
                165                 170                 175


Val Asp Ala Leu Glu Asn Ala Phe Lys Asp Lys Pro Val Ala Asp Ile
            180                 185                 190


Ala Leu Gly Asp Arg Lys Thr Asp Tyr Pro Phe Met Lys Leu Cys Lys
        195                 200                 205


Glu Ser Tyr Val Val Pro Ala Lys Pro Lys Val Glu Ala Val Ser His
    210                 215                 220


Asp Lys Leu Pro Lys Pro Ile Val Phe His Asp Gly Arg Leu Val Gln
225                 230                 235                 240


Lys Pro Thr Pro Phe Met Ala Leu Leu Thr Ile Leu Trp Ile Pro Val
                245                 250                 255


Gly Phe Leu Leu Ala Cys Leu Arg Ile Ala Ala Gly Ala Leu Leu Pro
            260                 265                 270
```

```
Met Pro Leu Val Tyr Tyr Ala Phe Trp Ala Leu Gly Val Arg Val His
        275                 280                 285

Ile Lys Gly Thr Pro Pro Pro Pro Ala Lys Lys Ser Ile Gly Gln Thr
    290                 295                 300

Gly Val Leu Phe Ile Cys Ser His Arg Thr Leu Leu Asp Pro Ile Phe
305                 310                 315                 320

Leu Ser Thr Ala Leu Gly Arg Lys Ile Pro Ala Val Thr Tyr Ser Leu
                325                 330                 335

Ser Arg Leu Ser Glu Val Ile Ser Pro Ile Lys Thr Ile Arg Leu Ser
            340                 345                 350

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

Gly Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe
    370                 375                 380

Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Asn Glu Leu Val
385                 390                 395                 400

Pro Val Ala Met Ala Asn Arg Met Ser Met Phe His Gly Thr Thr Ala
                405                 410                 415

Arg Gly Trp Lys Gly Met Asp Pro Phe Tyr Phe Phe Met Asn Pro Ser
            420                 425                 430

Pro Ala Tyr Glu Val Thr Phe Leu Asn Lys Leu Pro His Glu Leu Thr
        435                 440                 445

Cys Ser Ala Gly Lys Ser Ser His Glu Val Ala Asn Tyr Ile Gln Arg
    450                 455                 460
```

```
                                                257
Met Ile Ala Ala Ser Leu Ser Tyr Glu Cys Thr Ser Phe Thr Arg Lys
465                 470                 475                 480


Asp Lys Tyr Arg Ala Leu Ala Gly Asn Asp Gly Thr Val Val Glu Lys
                485                 490                 495


Pro Lys Leu Ser Pro Asn Lys Val Met Gly Cys
            500                 505
```

<210> 134
<211> 503
<212> PRT
<213> Theobroma cacao

<400> 134

```
Met Val Met Gly Ala His Arg Arg Phe Glu Pro Ile Ser Lys Cys Ser
1               5                   10                  15


Asn Glu Gly Arg Ser Asn Gln Thr Val Ala Ala Asp Leu Asp Gly Thr
            20                  25                  30


Leu Leu Leu Ser Arg Ser Ala Phe Pro Tyr Phe Met Leu Val Ala Leu
        35                  40                  45


Glu Ala Gly Ser Leu Val Arg Ala Leu Val Leu Leu Ala Ser Val Pro
    50                  55                  60


Phe Val Tyr Phe Thr Tyr Leu Phe Ile Ser Glu Ser Val Ala Ile Asn
65                  70                  75                  80


Thr Phe Ile Phe Ile Thr Phe Ala Gly Leu Lys Val Lys Asp Ile Glu
                85                  90                  95


Leu Val Ser Arg Ser Val Leu Pro Lys Phe Tyr Ala Glu Asp Val His
            100                 105                 110


Pro Glu Ser Trp Arg Val Phe Ser Ser Phe Gly Lys Arg Tyr Ile Ile
        115                 120                 125
```

```
Thr Ala Ser Pro Arg Ile Met Val Glu Pro Phe Val Lys Thr Phe Leu
    130                 135                 140

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

Arg Ala Thr Gly Phe Thr Ile Lys Pro Gly Val Leu Val Gly Glu His
                165                 170                 175

Lys Arg Ala Ala Ile Ser Lys Glu Phe Gly Met Lys Leu Pro Asp Leu
            180                 185                 190

Gly Leu Gly Asp Arg Glu Thr Asp His Asp Phe Met Ser Leu Cys Lys
        195                 200                 205

Glu Gly Tyr Met Val Pro Arg Ser Lys Cys Asp Pro Leu Pro Arg Asn
    210                 215                 220

Lys Leu Leu Ser Pro Ile Ile Phe His Glu Gly Arg Leu Val Gln Arg
225                 230                 235                 240

Pro Thr Pro Val Ala Ala Leu Leu Thr Phe Leu Trp Leu Pro Ile Gly
                245                 250                 255

Phe Ile Leu Ser Leu Leu Arg Val Tyr Leu Asn Ile Pro Leu Pro Glu
            260                 265                 270

Arg Ile Ala Arg Tyr Asn Tyr Lys Ile Leu Gly Ile Lys Leu Ile Val
        275                 280                 285

Lys Gly Asn Pro Pro Pro Ala Pro Lys Lys Gly Gln Ser Gly Val Leu
    290                 295                 300

Phe Val Cys Asn His Arg Thr Val Leu Asp Pro Val Val Thr Ala Val
305                 310                 315                 320

Ala Leu Gly Arg Lys Ile Ser Cys Val Thr Tyr Ser Ile Ser Lys Phe
```

325 330 335

Thr Glu Ile Ile Ser Pro Ile Arg Ala Val Ala Leu Ser Arg Glu Arg
340 345 350

Glu Lys Asp Ala Ala Asn Ile Lys Arg Leu Leu Glu Glu Gly Asp Leu
355 360 365

Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg
370 375 380

Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Arg Ile Val Pro Val Ala
385 390 395 400

Ile Asn Thr Lys Gln Thr Val Phe His Gly Thr Thr Val Arg Gly His
405 410 415

Lys Leu Leu Asp Pro Tyr Phe Val Phe Met Asn Pro Met Pro Thr Tyr
420 425 430

Glu Ile Thr Phe Leu Asn Gln Leu Pro Thr Glu Leu Thr Cys Lys Gly
435 440 445

Gly Lys Ser Ala Ile Glu Val Ala Asn Tyr Ile Gln Arg Val Leu Ala
450 455 460

Gly Thr Leu Gly Phe Glu Cys Thr Asn Leu Thr Arg Lys Asp Lys Tyr
465 470 475 480

Ala Met Met Ala Gly Thr Asp Gly Arg Val Pro Ser Lys Lys Glu Lys
485 490 495

Glu Gln Glu Lys Glu Lys Ala
500

<210> 135
<211> 531
<212> PRT

<213> Theobroma cacao

<400> 135

```
Met Ala Gly Lys Pro Gln Asn Leu Asn Ala Phe Leu Phe Leu Gln Lys
1               5                   10                  15


Ile Leu Leu Thr Lys Thr Trp Ser Lys Ile Tyr Leu Gln Arg Lys Ala
            20                  25                  30


Ser Asn Ser His Gly Phe Gln Pro Gln Phe Asn Lys Tyr Pro Phe Leu
        35                  40                  45


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


Glu Gly Ala Leu Leu Lys Ser Ser Ser Val Phe Ser Tyr Phe Phe Leu
65                  70                  75                  80


Val Ala Phe Glu Ala Gly Gly Leu Leu Arg Ala Phe Ile Leu Phe Leu
                85                  90                  95


Leu His Pro Ile Ile Cys Leu Val Gly Lys Glu Trp Gly Leu Lys Ile
            100                 105                 110


Met Val Phe Val Cys Phe Val Gly Leu Lys Glu Glu Ser Phe Arg Ile
        115                 120                 125


Gly Arg Thr Val Leu Pro Lys Phe Phe Leu Glu Asp Val Gly Asn Glu
    130                 135                 140


Gly Phe Asp Met Val Met Lys Cys Gly Gly Lys Lys Ile Gly Val Thr
145                 150                 155                 160


Asp Met Pro Arg Val Met Val Asp Cys Phe Leu Lys Asp Tyr Leu Arg
                165                 170                 175


Val Glu Ala Val Val Gly Arg Glu Leu Lys Val Val Cys Gly His Phe
            180                 185                 190
```

```
Val Gly Leu Met Glu Glu Lys Lys Pro Ser Gly Thr Ala Phe Leu Glu
        195                 200                 205

Leu Asp Glu Leu Leu Gly Lys Glu Asn Met Asp Ser Ser Val Ile Gly
    210                 215                 220

Ile Ser Cys Phe Asp Asn Ser Leu Asp Gln Gln Leu Phe Ser Cys Cys
225                 230                 235                 240

Glu Val Leu Tyr Met Val Ser Glu Ala Asp Lys Lys Asn Trp Gln Ala
                245                 250                 255

Leu Pro Arg Glu Lys Tyr Pro Arg Pro Leu Ile Phe His Asp Gly Arg
            260                 265                 270

Leu Ala Phe Arg Pro Thr Pro Leu Ala Ser Leu Arg Met Phe Met Trp
        275                 280                 285

Ile Pro Phe Gly Ile Phe Leu Leu Ile Val Arg Ala Thr Val Gly Tyr
    290                 295                 300

Leu Ile Pro Tyr Lys Phe Ser Thr Pro Phe Leu Cys Phe Thr Gly Leu
305                 310                 315                 320

Thr Gly Thr Leu Ile Arg Pro Ser Ser Phe Pro Asp Arg Ala Asp Asp
                325                 330                 335

Glu Gln Lys Gln Gly Gly Val Leu Tyr Val Cys Asn His Arg Thr Leu
            340                 345                 350

Leu Asp Pro Ile Tyr Val Ser Ile Val Leu Asp Lys Pro Phe Asn Ala
        355                 360                 365

Val Thr Tyr Ser Leu Ser Arg Met Thr Glu Leu Leu Ala Pro Ile Lys
    370                 375                 380
```

```
                                                262
Thr Val Arg Leu Thr Arg Asn Arg Glu Glu Asp Ser Lys Met Ile Gly
385                 390                 395                 400


Lys Leu Leu Lys Gln Gly Asp Leu Val Ile Cys Pro Glu Gly Thr Thr
                405                 410                 415


Cys Arg Glu Pro Tyr Leu Leu Arg Phe Ser Pro Leu Phe Ala Glu Ile
            420                 425                 430


Thr Asp Glu Ile Val Pro Val Ala Ile Asp Phe Gln Val Ser Met Phe
        435                 440                 445


Tyr Gly Thr Thr Ala Ser Gly Phe Lys Phe Leu Asp Pro Phe Phe Leu
    450                 455                 460


Leu Met Asn Pro Thr Ser His Cys Ser Ala Ser Ile Leu Glu Lys Leu
465                 470                 475                 480


Pro Lys Ser Leu Thr Cys Gln Gly Gly Gly Lys Ser Lys Phe Glu Val
                485                 490                 495


Ala Asn Arg Val Gln Ala Ile Ile Ala Lys Ala Leu Asn Phe Glu Cys
            500                 505                 510


Thr Ser Leu Thr Arg Arg Asp Lys Tyr Met Ile Leu Ala Gly Asn Gln
        515                 520                 525


Gly Ile Val
    530
```

<210> 136
<211> 553
<212> PRT
<213> Theobroma cacao

<400> 136

```
Met Ala Asn Arg Gln Ser Val Leu Trp Gly Leu Lys Thr Val Cys Leu
1               5                   10                  15
```

```
Val His Gln His Leu Asn Ser Thr Ile Tyr Gly Asn Leu Cys Gln Gln
            20                  25                  30

Ala Ile Lys Val Ile Lys Asn Asp Asp Val Ala Ser Gly Ile Asn Glu
        35                  40                  45

Pro Leu Ser Ile Leu Asn Gln Arg Lys Leu Pro Tyr Leu Arg Phe Lys
    50                  55                  60

Val Ser Asn Gly Pro Ala Thr Gln Phe Lys Phe Leu Lys His Ser Thr
65                  70                  75                  80

Leu Val His Arg Thr Asp Glu Leu Ala Asn Gln Thr Leu Ala Phe His
                85                  90                  95

Leu Glu Gly Ala Leu Leu Lys Ser Cys Ser Leu Phe Pro Tyr Phe Met
            100                 105                 110

Leu Met Ala Phe Glu Ala Gly Gly Leu Phe Arg Ala Leu Ile Leu Leu
        115                 120                 125

Leu Leu Tyr Pro Leu Val Leu Leu Ile Gly Lys Glu Leu Gly Leu Arg
    130                 135                 140

Ile Leu Val Phe Val Ser Phe Val Gly Ile Arg Lys Glu Lys Phe Arg
145                 150                 155                 160

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

Glu Gly Phe Asp Ile Val Met Ser Tyr Lys Lys Arg Val Ala Val Thr
            180                 185                 190

Gly Met Pro Lys Phe Met Val Glu Gly Phe Leu Arg Asp Tyr Leu Gly
        195                 200                 205

Ile Asp Ala Val Val Ala Arg Glu Leu Lys Glu Phe Arg Gly Tyr Phe
```

210 215 220

Leu Gly Leu Met Glu Glu Lys Met Asp Ala Gly Leu Val Ile Ser Glu
225 230 235 240

Cys Thr His His Ile Gly Leu Gly Cys Phe Arg Lys Ser His Asp Gln
245 250 255

Lys Ile Phe Ser His Cys Lys Glu Ile Tyr Leu Val Thr Glu Ala Glu
260 265 270

Lys Lys Asn Trp Gln Val Leu Pro Arg Lys Arg Tyr Pro Lys Pro Leu
275 280 285

Ile Phe His Asp Gly Arg Leu Ala Phe Arg Pro Thr Pro Leu Ala Ala
290 295 300

Leu Ile Met Phe Ile Trp Leu Pro Phe Gly Phe Leu Leu Asn Ile Thr
305 310 315 320

Arg Thr Ile Val Phe Ile Ser Leu Pro Phe Lys Leu Ser Ile Pro Leu
325 330 335

Leu Ala Leu Ser Gly Cys Ile Ile Thr Val Ser Lys Pro Glu Pro Thr
340 345 350

Thr Ala Ser Ile Asn Gly Glu Asn Lys Pro Gly Gly Met Leu Tyr Val
355 360 365

Cys Asn His Arg Thr Leu Leu Asp Pro Leu Phe Leu Ser Ala Val Leu
370 375 380

Met Lys Ser Val Cys Ala Val Thr Tyr Ser Ile Ser Arg Phe Ser Glu
385 390 395 400

Val Ile Ser Pro Ile Lys Thr Val Arg Leu Ser Arg Asp Arg Glu Lys
405 410 415

```
Asp Gly Lys Ile Met Lys Lys Leu Leu Ser Gln Gly Asp Leu Val Val
            420                 425                 430

Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Tyr Leu Leu Arg Phe Ser
        435                 440                 445

Pro Leu Phe Ala Glu Met Thr Asp Glu Ile Val Pro Val Ala Ile Lys
    450                 455                 460

Leu Gln Val Ser Leu Phe Tyr Gly Ser Thr Ala Ser Gly Leu Lys Cys
465                 470                 475                 480

Leu Asp Ser Thr Phe His Leu Met Asn Pro Asn Pro Met Cys Ser Val
                485                 490                 495

Lys Ile Leu Asn Lys Leu Pro Ser Cys Gln Thr His Asn Thr Gly Gly
            500                 505                 510

Lys Ser Lys Phe Glu Val Thr Asn His Val Gln Asn Gln Ile Ala Ala
        515                 520                 525

Ala Leu Gly Phe Glu Cys Thr Asn Leu Thr Arg Lys Asp Lys Tyr Ala
    530                 535                 540

Ile Leu Ala Gly Asn Glu Gly Ile Val
545                 550
```

<210> 137
<211> 461
<212> PRT
<213> Theobroma cacao

<400> 137

```
Met Ser Thr Leu Pro Leu Pro Phe Phe Ser Gly Thr Ser Thr Arg Ala
1               5                   10                  15

Pro Phe Ser Phe Ser Leu Lys Ser Ser Pro Pro Ser Ala Ala Ser Ser
            20                  25                  30
```

```
Ser Ser Ser Ser Ser Leu Val Pro Phe Gln Arg Val Glu Leu Phe Arg
        35                  40                  45

Ser Thr Ser Thr Arg Thr Thr Arg Ser Phe Ser Cys Val Phe Tyr Ser
    50                  55                  60

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

Val Ala Ala Thr Gly Ser Ser Gly Gly Gly Arg Pro Glu Val Glu His
                85                  90                  95

Ser Arg Thr Phe Leu Glu Ala Arg Ser Glu Gln Glu Leu Leu Ser Gly
            100                 105                 110

Ile Arg Lys Glu Val Glu Ala Gly Arg Leu Pro Pro Asn Ile Ala Ala
        115                 120                 125

Gly Met Glu Glu Leu Tyr Gln Asn Tyr Arg Asn Ala Val Phe Gln Ser
    130                 135                 140

Gly Asp Pro Ser Ala Val Glu Ile Val Leu Ser Asn Met Ala Val Ala
145                 150                 155                 160

Phe Asp Arg Met Leu Leu Asp Val Glu Asp Pro Phe Val Phe Glu Pro
                165                 170                 175

Tyr His Lys Ala Leu Arg Glu Pro Phe Asp Tyr Tyr Met Phe Gly Gln
            180                 185                 190

Asn Tyr Ile Arg Pro Leu Thr Asp Phe Arg Asn Ser Tyr Val Gly Asn
        195                 200                 205

Leu Ser Leu Phe Tyr Glu Ile Glu Glu Lys Leu Lys Gln Gly His Asn
    210                 215                 220
```

```
                                                267
Val Val Leu Ile Ser Asn His Gln Thr Glu Ala Asp Pro Ala Ile Ile
225                 230                 235                 240


Thr Leu Leu Leu Glu Lys Thr Asn Pro His Val Ala Glu Asn Met Ile
                245                 250                 255


Tyr Val Ala Gly Asp Arg Val Ile Thr Asp Pro Leu Cys Lys Pro Phe
            260                 265                 270


Ser Met Gly Arg Asn Leu Ile Cys Val Tyr Ser Lys Lys His Met Tyr
        275                 280                 285


Asp Val Pro Glu Leu Val Glu Met Lys Arg Lys Ala Asn Thr Gln Ser
    290                 295                 300


Leu Lys Gln Met Ala Leu Leu Leu Arg Gly Gly Ser Lys Ile Val Trp
305                 310                 315                 320


Ile Ala Pro Ser Gly Gly Arg Asp Arg Pro Asp Pro Leu Thr Glu Glu
                325                 330                 335


Trp Tyr Pro Ala Pro Phe Asp Ser Ser Ser Val Asp Asn Met Arg Arg
            340                 345                 350


Leu Ile Glu His Ser Gly Ala Pro Gly His Ile Tyr Pro Leu Ala Leu
        355                 360                 365


Leu Cys Tyr Asn Ile Met Pro Pro Pro Pro Gln Val Glu Lys Gln Ile
    370                 375                 380


Gly Glu Arg Arg Asn Ile Ser Phe His Gly Ala Gly Leu Ser Val Ala
385                 390                 395                 400


Pro Lys Ile Ser Phe Gln Glu Ile Ala Ala Ala Cys Glu Lys Ser Glu
                405                 410                 415


Glu Ala Lys Asp Val Tyr Ser Gln Ala Leu Tyr Lys Ser Val Thr Lys
            420                 425                 430
```

```
Gln Tyr Asn Val Leu Lys Ser Ala Val His Gly Lys Gln Gly Leu Glu
        435                 440                 445


Ala Ser Thr Ala Gly Val Ser Leu Ser Gln Pro Trp Asn
    450                 455                 460


<210>  138
<211>  522
<212>  PRT
<213>  Theobroma cacao

<400>  138

Met Ala Lys Ile Lys Lys Leu Phe Pro Leu Lys Ala Leu Ser Leu Leu
1               5                   10                  15


Phe Glu Thr Leu Leu Lys Ser Ser Gly Lys Leu Pro Tyr Leu Arg Phe
            20                  25                  30


Lys Val Ser Asn Gly Pro Ala Thr Gln Phe Lys Phe Leu Lys His Ser
        35                  40                  45


Thr Leu Val His Arg Thr Asp Glu Leu Ala Asn Gln Thr Leu Val Phe
    50                  55                  60


His Leu Glu Gly Ala Leu Leu Lys Ser Cys Ser Leu Phe Pro Tyr Phe
65                  70                  75                  80


Met Leu Val Ala Phe Glu Ala Gly Gly Leu Phe Arg Ala Leu Ile Leu
                85                  90                  95


Leu Leu Leu Tyr Pro Leu Val Leu Leu Ile Gly Lys Glu Leu Gly Leu
            100                 105                 110


Arg Ile Leu Val Phe Val Ser Phe Val Gly Ile Thr Lys Glu Lys Phe
        115                 120                 125


Arg Ala Gly Thr Thr Ile Leu Pro Asn Phe Phe Leu Glu Asp Val Gly
```

```
                                                269
    130                 135                 140

Cys Glu Gly Phe Asp Ile Val Met Ser Tyr Lys Lys Arg Val Ala Val
145                 150                 155                 160

Thr Gly Met Pro Lys Phe Met Ile Glu Gly Phe Leu Arg Asp Tyr Leu
                165                 170                 175

Gly Ile Asp Ala Val Val Ala Arg Glu Leu Lys Glu Phe Arg Gly Tyr
            180                 185                 190

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

Glu Cys Thr His Asn Ile Gly Leu Gly Cys Phe Arg Lys Ser His Asp
    210                 215                 220

Gln Lys Ile Phe Ser His Cys Lys Glu Ile Tyr Leu Val Thr Glu Ala
225                 230                 235                 240

Glu Lys Lys Asn Trp Gln Val Leu Pro Arg Lys Arg Tyr Pro Lys Pro
                245                 250                 255

Leu Ile Phe His Asp Gly Arg Leu Ala Phe Arg Pro Thr Pro Phe Ala
            260                 265                 270

Ala Leu Ile Met Phe Ile Trp Leu Pro Phe Gly Phe Leu Leu Asn Ile
        275                 280                 285

Thr Arg Ala Ile Ile Phe Ile Ser Leu Pro Phe Asn Leu Ser Thr Pro
    290                 295                 300

Leu Leu Ala Leu Ser Gly Phe Ile Val Thr Val Ser Lys Pro Glu Pro
305                 310                 315                 320

Thr Thr Ala Ser Ile Asn Gly Glu Asn Lys Pro Gly Gly Met Leu Tyr
                325                 330                 335
```

```
Val Cys Asn His Arg Thr Leu Leu Asp Pro Leu Phe Leu Ser Gly Val
            340                 345                 350

Leu Met Asn Ser Val Ser Ala Val Thr Tyr Ser Ile Ser Lys Phe Ser
        355                 360                 365

Glu Val Ile Ser Pro Ile Lys Thr Val Gly Leu Ser Arg Asp Arg Glu
    370                 375                 380

Lys Asp Gly Lys Ile Met Lys Lys Leu Leu Ser Gln Gly Asp Leu Val
385                 390                 395                 400

Val Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Tyr Leu Leu Arg Phe
                405                 410                 415

Ser Pro Leu Phe Ala Glu Met Thr Asn Glu Ile Val Pro Val Ala Ile
            420                 425                 430

Lys Leu Gln Val Ser Leu Phe Tyr Gly Ser Thr Ala Ser Gly Leu Lys
        435                 440                 445

Cys Leu Asp Ser Thr Phe His Leu Met Asn Pro Asn Pro Met Cys Ser
    450                 455                 460

Val Lys Ile Leu Asn Lys Leu Pro Ser Cys Gln Thr His Asn Thr Gly
465                 470                 475                 480

Gly Lys Ser Lys Phe Glu Val Asn Asn His Val Gln Asn Gln Ile Ala
                485                 490                 495

Ala Ala Leu Gly Phe Glu Cys Thr Asn Leu Thr Arg Lys Asp Lys Tyr
            500                 505                 510

Ala Ile Leu Ala Gly Asn Glu Gly Ile Val
        515                 520
```

<210> 139

<211> 370
<212> PRT
<213> Theobroma cacao

<400> 139

```
Met Ala Asn Leu Arg Ala His Val Cys Phe Leu Leu Ser Phe Val Thr
1               5                   10                  15


Phe Leu Asn Leu Ile Ser Cys His Thr Asn Leu Pro Ala Asn His Val
            20                  25                  30


Ala Leu Phe Val Phe Gly Asp Ser Leu Phe Asp Pro Gly Asn Asn Asn
        35                  40                  45


Tyr Ile Asn Thr Thr Phe Asp Phe Gln Ala Asn Phe Arg Pro Tyr Gly
    50                  55                  60


Glu Thr Phe Phe Lys Tyr Pro Thr Gly Arg Phe Ser Asp Gly Arg Leu
65                  70                  75                  80


Ile Pro Asp Phe Leu Ala Gln Phe Ala Gly Leu Pro Ile Ile Pro Thr
                85                  90                  95


Tyr Leu Gln Pro Gly Asn His Lys Phe Thr Asp Gly Val Asn Phe Ala
            100                 105                 110


Ser Gly Gly Ala Gly Ala Leu Val Glu Ser His Gln Gly Phe Val Val
        115                 120                 125


Asp Leu Glu Thr Gln Ile Arg Tyr Phe Lys Lys Val Glu Lys Ser Leu
    130                 135                 140


Arg Gln Glu Leu Gly Asp Glu Glu Ala Lys Lys Leu Leu Ser Arg Ala
145                 150                 155                 160


Val Tyr Leu Ile Ser Val Gly Gly Asn Asp Tyr Leu Thr Arg Asn Ser
                165                 170                 175
```

```
Ser Ala Ser Asp Glu Glu Tyr Ala Ser Met Val Leu Gly Asn Leu Thr
            180                 185                 190

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

Pro Asn Met Met Pro Leu Gly Cys Leu Pro Tyr Met Lys Ala Lys Ala
    210                 215                 220

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

Lys Glu Leu Pro Lys Thr Leu Gln Lys Leu Glu Thr Gln Leu Glu Gly
                245                 250                 255

Phe Lys Tyr Ala Tyr Tyr Asn Phe Tyr Lys Ser Val Ser Glu Arg Leu
            260                 265                 270

Asn Asn Pro Ser Lys Tyr Gly Phe Lys Asp Ala Thr Thr Ala Cys Cys
        275                 280                 285

Gly Ser Gly Leu Tyr Gly Gly Val Tyr Ser Cys Gly Gly Lys Arg Gly
    290                 295                 300

Ile Thr Glu Phe His Leu Cys Glu Asn Pro Ser Glu Tyr Leu Phe Phe
305                 310                 315                 320

Asp Ser Tyr His Pro Ser Glu Lys Ala Tyr Gln Gln Phe Ala Glu Leu
                325                 330                 335

Met Trp Ser Gly Thr Thr Asp Phe Val Trp Pro Tyr Asn Leu Lys Thr
            340                 345                 350

Leu Phe Glu Ala Lys Thr Ser Asn Leu Tyr Asp His Phe Ile Ile Pro
        355                 360                 365

Pro Pro
    370
```

<210> 140
<211> 398
<212> PRT
<213> Theobroma cacao

<400> 140

```
Met Leu Thr Tyr Leu Pro Leu Arg Cys Leu Gly Leu Lys Ser Leu Val
1               5                  10                  15


Cys Ile Ser Phe Asp Cys Gly Val Ile Gly Ala Met Ala Ala Arg Val
            20                  25                  30


Lys Tyr Ser Trp Ser Leu Ala Leu Val Leu Leu Cys Ile Val Ser Leu
        35                  40                  45


Ser Tyr Phe Gly Asn Ala Ala Asp Leu Gly Ala Ser Phe Val Phe Gly
    50                  55                  60


Asp Ser Leu Val Asp Ala Gly Asn Asn Asn Tyr Leu Pro Thr Leu Ser
65                  70                  75                  80


Lys Ala Asn Ile Pro Pro Asn Gly Ile Asp Phe Lys Ala Ser Gly Gly
                85                  90                  95


Gln Pro Thr Gly Arg Tyr Thr Asn Gly Arg Thr Ile Gly Asp Ile Val
            100                 105                 110


Gly Glu Glu Leu Gly Val Pro Asn Tyr Ala Val Pro Phe Leu Ser Pro
        115                 120                 125


Asn Ser Thr Gly Lys Ala Ile Leu Tyr Gly Val Asn Tyr Ala Ser Gly
    130                 135                 140


Gly Gly Gly Ile Met Asn Ala Thr Gly Arg Ile Phe Val Asn Arg Leu
145                 150                 155                 160


Gly Leu Asp Ile Gln Ile Asp Phe Tyr Asn Thr Thr Arg Lys Gln Phe
```

```
                165                 170                 175

Asp Lys Leu Leu Gly Pro Ser Lys Ala Lys Asp Tyr Ile Arg Lys Ser
            180                 185                 190

Ser Ile Phe Ser Ile Thr Ile Gly Ala Asn Asp Phe Leu Asn Asn Tyr
        195                 200                 205

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

Gly Phe Ile Asp Asp Met Ile Asn His Leu Arg Asp Gln Leu Thr Arg
225                 230                 235                 240

Leu Tyr Gln Leu Asp Ala Arg Lys Phe Val Ile Gly Asn Val Gly Pro
                245                 250                 255

Ile Gly Cys Ile Pro Tyr Gln Lys Thr Ile Asn Gln Leu Asn Glu Asn
            260                 265                 270

Glu Cys Val Asp Leu Ala Asn Lys Leu Ala Lys Gln Tyr Asn Ser Arg
        275                 280                 285

Leu Gln Asn Leu Leu Thr Glu Leu Asn Gly Lys Leu Pro Gly Ala Thr
    290                 295                 300

Phe Val Val Ala Asn Val Tyr Asp Leu Val Met Glu Leu Ile Thr Asn
305                 310                 315                 320

Tyr Glu Lys Tyr Gly Phe Ile Thr Ala Ser Lys Ala Cys Cys Gly Asn
                325                 330                 335

Gly Gly Gln Phe Ala Gly Ile Ile Pro Cys Gly Pro Thr Ser Thr Met
            340                 345                 350

Cys Lys Asp Arg Asp Lys His Val Phe Trp Asp Pro Tyr His Pro Ser
        355                 360                 365
```

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

Lys Tyr Ile Ser Pro Val Asn Leu Lys Glu Leu Arg Asn Leu
385 390 395

<210> 141
<211> 350
<212> PRT
<213> Theobroma cacao

<400> 141

Met Arg Val Ser Ser Ser Phe Leu Ala Ser Leu Val Leu Val Leu Val
1 5 10 15

Phe Ser Ala Ala Asn Gly Asp Pro Leu Val Pro Ala Leu Ile Ser Phe
20 25 30

Gly Asp Ser Val Val Asp Val Gly Asn Asn Asn Asn Leu Asn Thr Leu
35 40 45

Ile Lys Ala Asn Phe Pro Pro Tyr Gly Arg Asp Phe Val Thr His Arg
50 55 60

Pro Thr Gly Arg Phe Cys Asn Gly Lys Leu Ala Thr Asp Phe Thr Ala
65 70 75 80

Glu Tyr Leu Gly Phe Thr Ser Tyr Pro Pro Ala Tyr Leu Ser Gln Asp
85 90 95

Ala Arg Gly Asn Asp Leu Leu Thr Gly Ala Ser Phe Ala Ser Ala Ala
100 105 110

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

Thr Arg Gln Val Asn Tyr Tyr Lys Glu Tyr Gln Thr Lys Val Val Asn
130 135 140

```
Thr Val Gly Lys Ala Lys Ala Asn Asp Ile Phe Ala Gly Ala Ile His
145                 150                 155                 160

Leu Leu Ser Ala Gly Ser Ser Asp Tyr Ile Gln Asn Tyr Tyr Ile Asn
                165                 170                 175

Pro Met Leu Asn Arg Val Tyr Thr Pro Asp Gln Phe Ser Asp Ile Leu
            180                 185                 190

Ile Arg Ser Tyr Ser Thr Phe Ile Gln Asn Leu Tyr Gly Leu Gly Ala
        195                 200                 205

Arg Arg Ile Gly Val Thr Thr Leu Pro Pro Thr Gly Cys Leu Pro Ala
    210                 215                 220

Ala Ile Thr Leu Phe Gly Gly Gly Ser Asn Gln Cys Val Thr Arg Leu
225                 230                 235                 240

Asn Gln Asp Ala Val Ala Phe Asn Lys Lys Leu Asn Asn Thr Ser Glu
                245                 250                 255

Ser Leu Leu Asp Lys Met Pro Gly Leu Lys Leu Leu Val Phe Asp Ile
            260                 265                 270

Tyr Gln Pro Leu Leu Asp Met Val Met Lys Pro Ala Asp Asn Gly Phe
        275                 280                 285

Phe Glu Ser Arg Arg Ala Cys Cys Gly Thr Gly Thr Leu Glu Thr Ser
    290                 295                 300

Leu Leu Cys Asn Asp Arg Ser Leu Gly Thr Cys Thr Asn Ala Thr Ser
305                 310                 315                 320

Tyr Val Phe Trp Asp Gly Phe His Pro Ser Glu Ala Ala Asn Gln Val
                325                 330                 335
```

Leu Ala Gly Asn Leu Leu Ala Gln Gly Val Ser Leu Ile Ser
340 345 350

<210> 142
<211> 388
<212> PRT
<213> Theobroma cacao

<400> 142

Met Ala Lys Leu Ile Ser Tyr Ile Asn Phe Phe Lys Val Ser Val Ile
1 5 10 15

Leu Leu Leu Ser Thr Thr Ser Arg Cys Phe Gly Thr Pro Leu Glu Lys
20 25 30

Gly Ser Ala Leu Phe Ile Phe Gly Asp Ser Thr Val Asp Pro Gly Asn
35 40 45

Asn Asn Tyr Ile Asn Thr Ile Pro Glu Asn Arg Ala His Tyr Lys Pro
50 55 60

Tyr Gly Gln Asn Gly Phe Phe Ala Glu Pro Thr Gly Arg Phe Ser Asp
65 70 75 80

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

Ile Pro Pro Phe Leu Gln Pro Ser Ala Asp Phe Thr Ser Gly Ala Asn
100 105 110

Phe Ala Ser Gly Gly Gly Gly Val Leu Pro Glu Thr Asn Gln Gly Leu
115 120 125

Val Ile Asp Leu Pro Thr Gln Leu Lys Asn Phe Glu Glu Val Gln Lys
130 135 140

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

```
Glu Ala Val Tyr Phe Ile Ser Ile Gly Ser Asn Asp Tyr Met Gly Gly
                165                 170                 175

Tyr Leu Gly Asn Pro Lys Met Gln Glu Phe Tyr Pro Pro Glu Ala Tyr
            180                 185                 190

Val Gly Met Val Ile Gly Asn Leu Thr Gln Thr Leu Gln Val Leu Tyr
        195                 200                 205

Glu Lys Gly Ala Arg Lys Phe Gly Phe Leu Ser Leu Ser Pro Leu Gly
    210                 215                 220

Cys Leu Pro Ala Leu Arg Ala Leu Asn Phe Lys Asp Asn Glu Gly Lys
225                 230                 235                 240

Cys Leu Glu Ala Ala Ser Ala Leu Ala Arg Ala His Asn Asp Ala Leu
                245                 250                 255

Asn Ala Val Leu Thr Asn Leu Glu His Ile Leu Pro Gly Leu Lys Tyr
            260                 265                 270

Cys Asn Ser Asn Phe Tyr Glu Trp Leu Asp Asp Arg Met Asn Asn Pro
        275                 280                 285

Ser Lys Tyr Gly Phe Lys Asp Gly Val Asn Ala Cys Cys Gly Thr Gly
    290                 295                 300

Pro Tyr Gly Gly Val Phe Thr Cys Gly Gly Thr Lys Asn Val Lys Asp
305                 310                 315                 320

Tyr Gln Leu Cys Asp Asn Ala Asp Asp His Val Trp Trp Asp Ser Phe
                325                 330                 335

His Pro Thr Glu Arg Ile His Glu Gln Phe Ala Lys Ala Leu Trp Ser
            340                 345                 350

Gly His Pro Ser Ser Val Gly Pro Tyr Asn Leu Glu Asp Leu Phe Phe
```

355 360 365

Asn Lys Glu Lys Leu Thr Ile Ala Asp Ile Val Asp Ser Pro Glu Thr
370 375 380

Asp Pro Phe Gln
385

<210> 143
<211> 360
<212> PRT
<213> Theobroma cacao

<400> 143

Met Gly Lys Leu Tyr Thr Ile Ile Val Val Ile Leu Ile Phe Ser Ser
1 5 10 15

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

Ser Ile Phe Asp Ala Gly Asn Asn His Phe Asn Lys Asn Cys Ser Val
35 40 45

Gln Ala Asp Phe Ser Pro Tyr Gly Ser Thr Phe Phe His Gln Pro Thr
50 55 60

Gly Arg Phe Thr Asn Gly Arg Thr Val Ala Asp Phe Ile Ser Gln Phe
65 70 75 80

Ile Gly Ile Asp Met Gln Lys Pro Tyr Leu Glu Ala Gln Ile Ala Val
85 90 95

Ile Asn Gly Ser Arg Lys Asp Tyr Pro Ser Asn Gly Ile Asn Phe Ala
100 105 110

Ser Ala Gly Ser Gly Val Leu Gln Gly Thr Asn Met Glu Gln Gly Val
115 120 125

```
                                                    280
Ile Ser Ile Gln Glu Gln Leu Arg Gln Phe Gln Thr Leu Val Glu Gln
    130                 135                 140

Asn Gln Ile Asp Lys Asn Leu Val Gln Lys Ser Leu Phe Phe Leu Glu
145                 150                 155                 160

Ser Gly Ser Asn Asp Ile Phe Asn Tyr Phe Ile Pro Phe Asp Pro Pro
                165                 170                 175

Thr Leu Asp Pro Asp Ala Tyr Val Gln Ala Met Met Lys Glu Val Ala
            180                 185                 190

Thr Leu Ile Asp Gln Ile Tyr Lys Leu Gly Ala Arg Arg Ile Ala Val
        195                 200                 205

Phe Ser Leu Gly Pro Val Gly Cys Val Pro Ala Arg Ala Leu Leu Leu
    210                 215                 220

Gly Ala Pro Ile Asn Arg Cys Tyr Gly Lys Met Asn Val Met Val Lys
225                 230                 235                 240

Lys Tyr Asn Lys Gly Leu Glu Ser Leu Ile Asn Asn Ile Pro Ile Arg
                245                 250                 255

Tyr Pro Gly Ala Ile Gly Val Tyr Gly Ala Ile Tyr Asp Ile Val Gln
            260                 265                 270

Arg Phe Arg Ala Ile Pro Ala Arg Tyr Gly Phe Ser Asp Val Thr Ser
        275                 280                 285

Ala Cys Cys Gly Asn Gly Pro Leu Arg Gly Lys Leu Gln Cys Gly Ile
    290                 295                 300

Glu Gly Tyr Gln Ile Cys Leu Asn Pro Asn Glu Val Leu Phe Trp Asp
305                 310                 315                 320

Tyr Phe His Pro Thr Glu His Ala Tyr Lys Leu Ile Ser Glu Ala Leu
                325                 330                 335
```

Trp Gly Gly Asn Lys Ser Gln Ile Arg Pro Val Asn Leu Lys Asn Leu
340 345 350

Ala Asp Ile Thr Phe Thr Leu Val
355 360

<210> 144
<211> 1506
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 144
atggctatct cggacagccc cgagattctc ggcagcaccg ccactgtcac ttcttcttcg 60

cactctgatt ctgacctcaa tctccttagc attcgacgtc ggacgagcac tacggcagcc 120

ggtcgggctc ctgacagaga tgatagcgga aatggagaag cagtggatga tcgtgaccaa 180

gtcgaatctg caaatctgat gtcgaacgtg gcagaaaatg ccaacgaaat gcccaactcc 240

agcgatacgc gattcaccta ccgtccccga gtccctgcac accgacggat caaagagtcc 300

cctctgtctt ctggcgcaat ctttaaacaa tcgcatgcag gacttttcaa tctgtgcatt 360

gttgttctcg tcgccgtcaa ttcgcgactc attatcgaga atctcatgaa atacggatgg 420

ctgatccgat cgggtttttg gttttctagc cggagcctct ctgactggcc tctgttcatg 480

tgttgcctta cactgcccat cttccccctt gcagctttcg ttgtggaaaa acttgtgcaa 540

agaaactaca tctcggaacc tgttgtggtg tttctgcacg caattatttc tactacagcc 600

gtgctttacc ccgtcatcgt gaatcttaga tgcgattcgg cttttctttc tggtgttgct 660

ctcatgctct ttgcctgcat cgtgtggctg aagctggtgt cttacgcaca tacaaacaat 720

gatatgcgtg ccctcgccaa aagcgctgag aagggagatg tcgatccttc ttacgacgtt 780

agcttcaaaa gcctggcata tttcatggtt gcacctacac tctgctacca acagtcttat 840

cctagaacgc ccgcagtgcg gaagagctgg gtggtgcgac agtttattaa gcttatcgtg 900

ttcacaggac tcatgggttt catcattgag caatacatta accccatcgt ccaaaactcg 960

cagcaccccc ttaagggtaa cctgctttat gccattgagc gtgtgctgaa actgagcgtc 1020

cccaatctct atgtgtggct ctgtatgttc tattgttttt ttcacctttg gctcaacatt 1080

cttgccgagc ttctgcgatt cggcgatcgt gagttttata aagactggtg gaatgctaag 1140

acggttgaag aatactggcg tatgtggaat atgcccgtcc acaagtggat ggtgagacac 1200

atctactttc cctgtctccg taacggtatt cctaagggcg ttgcaatcgt cattgcattt 1260

ctggtgtctg ccgtcttcca tgaactctgt attgctgtgc cttgccatat tttcaaactc 1320

tgggcattca tcggtattat gttccaggtc cctctggttc ttatcacaaa ttatctgcaa 1380

gataagttta gaagctctat ggtcggaaac atgattttct ggttcatctt ctctattctc 1440

ggacaaccta tgtgtgtcct tctctattat catgacctga tgaatcggaa aggaaaggcc 1500

gattaa 1506

<210> 145
<211> 984
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 145
atgatgggag aggagatgga agagcgaaaa gctacaggct atcgtgaatt ttctggacgg 60

catgagtttc cctccaatac tatgcatgct cttcttgcta tgggcatttg gctcggtgca 120

attcacttta atgccctgct gctcctcttc tcttttcttt tcctcccttt ctcgaaattc 180

ctggtcgtct ttggactcct tctccttttc atgattcttc ccatcgaccc ttattccaaa 240

tttggaagaa gactgtcccg atacatctgc aaacatgcat gcagctactt tcctattaca 300

ctgcatgtgg aggacatcca tgcattccac cccgacagag catatgtctt tggatttgaa 360

ccccattctg tgctccccat cggagtggtt gctctcgccg atcttaccgg cttcatgcct 420

ctgcccaaga tcaaagtcct tgcatcttcg gcagtcttct acaccccttt cctgcgtcat 480

atctggactt ggctcggact gacacccgct acgaagaaga atttctcttc tctgctcgac 540

gcaggttata gctgtatcct ggttcctggc ggcgtccagg aaacttttca tatggaaccc 600

ggctccgaga ttgctttcct tcgggcacgt agaggcttcg tgcgtattgc aatggaaatg 660

ggttctcccc tggtgcctgt cttttgtttt ggtcagtcgc acgtctataa atggtggaag 720

cctggtggca aattttacct tcaattttct cgagccatca agtttacacc tatcttcttt 780

tggggcattt tcggtagccc cctgccctac caacatccca tgcatgttgt cgtgggtaaa 840

cccattgatg tcaagaagaa cccccaaccc attgttgaag aggtcatcga agttcacgat 900

cggtttgtcg aagctctgca agacctcttt gaaagacata aagcacaagt gggatttgcc 960

gatcttcccc tgaagatcct gtaa 984

<210> 146
<211> 1431
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 146
atgggatcgg aagctcaata taagaatgcc atggcactga acggtgagga aggagagtat 60

tccgagcctg tcagccccac tggtcagtat ttcaactctt ccgccctgtc cattggtgtt 120

ctcggcgtcc tcgattcgga aattcccatt gatgactctc ccacaatgaa actccttgaa 180

gatgtctttc tgcctattaa tcctcggttt tctagcgtca tggtgcaaga tgagaacggc 240

gctaaacagt ggaaaaaggt tgaggtgaaa ctggttgacc atgtcaacgt ccctgtgttt 300

cctcctggac tttctcccga atcctatgat aattatcttt cggactatct ttctaaaatt 360

gcaacagaac aactgcctca aaaccgaccc ctttggaata tccacattgt taaatacaca 420

acttcgaacg ctgccggatc gctcattttc aaacttcacc acgccctcgg cgacggttat 480

tccctcatgg gagcacttct ttcgtgcctg cagcgagctg agaaccctc cgttcctctc 540

acgttccctg ctctctccag cgcacccaat ctttcggaga actctaactc ggtttttcgg 600

agcctgtact ctgctttcaa caccattagc gacttcggtt ggagccttct taaatcctcg 660

tttctggagg acgacagaac gatgattcga tcgggtgacc ccggagttca gttcaagccc 720

gttgttattt cgacaatgac tttctcgctt gaccatatta agcaaattaa gacgaaactt 780

ggtgtcacta ttaacgatgc aatcaccggc attatcttct tcggaaccag actctatatg 840

cagggtacat ccagcgataa actgtcgaac ggtaactcta ccgcacttgt gctgctcaat 900

acacgaacta tcggcggata taaatctgtg aaggatatgg tcaaacctga cgccgattct 960

ccttggggta atcagttcgg cttcctgcat gtttccctgc ctgagcttgc ctccacggca 1020

gagtcttgta accctctgga gttcatctgg aaggcccaga agctgattca gagaaagcga 1080

aattctaagg ctgtgttctt caccggacaa ctcctggagg gcctccgaaa gtacagaggc 1140

cctgaagcaa ctgctaaata catccattct acacttaaga actcctcgat gactatttcg 1200

aatctgatcg gtcccgtcga acgaatggca ctcgcagacc atcccgttaa aagcctgtac 1260

ttcatggttg tgggtgtgcc ccagagcctt acgattacga tggtttcgta tatgggcaag 1320

cttacggtgg ctattggtac ggagaaagac ttcgtggacc cccagaagtt caaatccagc 1380

gttgagaacg cattccaaat gatgctgaaa gcagcccaag caatcgccta a 1431

<210> 147
<211> 1458
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 147
atggcttgtc ccggaacttt ctctgatgaa ccccttaccc ctgctggtcg actgtttctg 60

cagagccaga cgaatatcat tattcactgc atccttcggg gaaagaatcc catcgacatc 120

gacgcagtta agtcgactct gcgatcgtcc cttatggtcc gtcatcccag attttgctct 180

cttctcgtcc gggacaaaaa cggtttcgaa cactggcgaa aggcacaggt tgacatcgac 240

cagcatgtca tcgtcattga taaacgactc gataagtccg ataatttcct cgacgaaagc 300

cgacgtttta gctcggagga gggagaaggt gacgacgatg acgaggccgc cgtcaaccag 360

tatgtggccg acctgtctgt ttcgtcgcct ctttcgactg acaagcccct ttgggagatt 420

catattcttg ttccccacaa gtgtgctgtg tttagaatcc accacgcact tggtgatggt 480

atctcgctga tgtcgatgct tatggcatcc tgccgacggg ctgatgacgc agatgcactg 540

cctcgtatgg tccctgagaa atctgccgag tttaaggatg gcaaaggacg ggattggttt 600

tggctgttcg gcatcctttg gggtttcctt aaaatggttt gcttcacagt gatgttcgtg 660

atggaatttg tggtgagctc tctccttgtt tgcgaccgga aaacggttat tagcggcggc 720

gaaggcgtcg agctctggcc tcgaaaactg gctacggccc ggttccttct ggaagatatg 780

cgagtcgtga agaaggcaat tcctaacaca acgattaatg acgtgctctt cggtgtcgtc 840

tcctccggac tttcgcggta tctggatcac cgtaccccca atgcccttca tgaaggactt 900

cggatgacgg gcgtcgcaat ggtcaatctt cgaccccaaa caggctttca ggatcttagc 960

caactcatga agggagctgc tgaagcacgt tggggaaata agttcggact cattctcctt 1020

cccgtttact tcaagaaggc tggtaataac cctctggaat atgtcaagcg ggcaaagcgg 1080

atggtggacc gtaagaagca ttccctggaa gcttacttct cctatcgtat cggcgatctc 1140

gttatgtccc ttctcggtag caaatacgcc tgtctcctca attatcgtct cctgtgtaat 1200

acaactttca ctatttctaa tattatcggc cctctcgaag aaattaccct cgctggtaat 1260

cctatcagca gcattaaggt taacacttcg tcgctgcccc aggcaattac gatgcacatg 1320

ctttcctacg ccggtagagc agaaatgcaa attctggtgg ctaaggacat cgtccctgac 1380

cccgagttcc tcgctaagtg ttttgaagac gccctcctgg agatgaagga agcagtcatc 1440

ggcaccgaga aagagtaa 1458

<210> 148
<211> 1623
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 148
atgaatcgaa agtcggaagc aaaatgggtg cctaagcttt tccaccagag cctgatctcc 60

atgaactcca ccgagattca ggaaatcgac caggttgagg ctaagatcgg aacggaggaa 120

gcctctcctg tgtcgaagga tgaggttcct ctctcgaccg gctctctgat tttccacgag 180

cccggtatca attgttgcat tatcgcaatc atgggataca aaagcaagct tgatcccgca 240

gtcatcaaag aaggtcttaa acagacgctg attaaacatc cctgtttttc cagcaaactc 300

ctgtctgcta ttgaccaaat gatgaactct ggtaagaaaa agtggatcag aactcaggtc 360

aatgtggaca atcatgtgat ccttcccgag gtcgataccg ttacagaatc gccctctcaa 420

ttcatcgagg attatgtgac ctatctgaca ggcattccta tcgacctttc caaacctctt 480

tgggaactgc acctggtgaa ccttaaaacc cccgaggcag aagcagtcgg cattttccga 540

atccatcatt cgattggcga cggtatgagc cttatttctc ttcttctcgc ctgctgtcga 600

aagtccagcg accccaaggc actgcctacg ctccccaaac aaaagcaagc agattcccgt 660

tacccccacc ctcacggatt tttccggctc tttctggcca tttggtctgt tcttcggctt 720

atcctgaaca ccctcgtgga cctccttctg tttgtggcca cagttgtttt cctgaaagac 780

accaaaactc cccttaaggg cagcagcgga gtcgagcaga acgctaagaa gatcgtccat 840

cggactgtgt cgcttgacga catcaaactt gtcaaggacg ccatgggaat gacggtcaac 900

gacgtcattc ttggagtgac agaagccgga ctctcccgat atctcaatag aaaatacggt 960

gaggttgatc gaggtaaaga ggcagaacaa aagtctaatc accttccccg taacattcga 1020

cttagagcca cggccctggt gaacattcgt cagaccgctg gaatccaggt gtctttcctg 1080

tcgccttttt ctcgggcctc tgaaaaagtc gagtatgcat ctcagagctt tcttgttttc 1140

aagtcgaagg ttaagtgggg aaaccagatt ggctacatct gtatcccctt cacgatcgcc 1200

cttcggaatg accctctcga ctacctgcga ggtgctaaag cagcaggaga tcgaaagaaa 1260

ctgtccctcg aagccatttg tacccatctg actaataaat gtgtggttaa actttcgga 1320

tcgaaacttt ccgccgcact tgtctaccgg gttatcttta atacgacaat gacactgtcc 1380

aatgtcgtcg gacctgttga agaaatcagc ctgtatggcc accctatcgc ctttattgct 1440

ccttcggtct acggtcaccc tcaggcactt acgtgccatt ttcagtcgta catgaacaag 1500

atgagcattg tcctcgccgt cgatcctaat gtgatccctg atccccgtct gctttgcgat 1560

gatctcgagg aatcccttaa aatctttaag gatgccgtcg ttctcgcaaa ggacgcagat 1620

taa 1623

<210> 149
<211> 1533
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 149
atggagagcc ggaacggatt ccgttggaga aaacagggcc tcaaacccat tgagacaaaa 60

acaggcagca aaggtgaaga agagaacaaa tcctcctctg gccaaagaac cgccgaggag 120

gagcctctct ctccttcggc tcgtcttttc catgaaccta atttcaacgt gtacatcatt 180

gcaacccttg gttgcaaaac gcgtatttac cctgatgtgg tcaaggccaa tctgggccat 240

acccttctga aacaccccag attcagcagc cttcaggtcg aagacgagga taacaagggc 300

gaaatgaagt gggtgcgaac tgaagtggac ctcgagcgac acgtggtcgt tcccgacctg 360

gaccccaata tggaccttcc tgataaattt ctcgaggact acatctacaa cctttccaag 420

acgacgattg ataagagcca acctctgtgg gatctccatc tgcttaacct tcggacctcg 480

cagtctgaag ctgttggcgt ttttagaatc catcacagcc tgggcgacgg tacctcgctc 540

atgagccttc tcctcgcatg tacgcgacaa atgaatgatc cccttgctct tcccaccatc 600

cctattagaa aaaagcagga aagaaagaat gatcgtcgag gtttttggcg gatcatgttc 660

atgctttgtt cggttttcca ggtgttttgg aactctgttg tggacgtttt tatgttcatg 720

gcaactgcac tccttcttaa agacacagag aaccctttca agggactgcc cggagtcgag 780

tttactccta gaagaattgt ttatggcacg gtttcgctgg acgacatcaa gctcgttaag 840

aacgctatga acacaacagt caatgatgtg gcacttggta ttacccaagc tggtctgtct 900

cgttacatta atagaatcta cggcggcaat aagaatgatg gcgcagctat tgaaatggat 960

aaccttccta aatccattcg actgcgatct acgctcctgg tgaacatccg gcccagcgcc 1020

ggcattcaag ccctcgccga catgatggag aaagatgctg aagcaaaatg gggcaactgg 1080

atcggctacg tgcttctccc cttcaccatt gcaatccggg atgaccctct tgattacgtc 1140

cgtgatgcca aagctacgat tgatcggaag aaaagatcgc tcgaagcaat ttacactttc 1200

tcgattgccg aactcgtgct gaaactgttc ggcgttaagg ccgcttcggc tctctcccat 1260

agaattatga gccatacaac gatgtgtttc agcaatctgg tcggtcccct cgaggagatc 1320

ggattctatg gccaccccat cgctttcctc gctccctctt cgtatggcca accccatgca 1380

cttatgatca atttccaatc gtatattgat aagatgacta tcgttctcag cgttgatgag 1440

ggtaccattc ccaatcctca ccaactttgc gatgatatcg tcgagtcgct tggccttatt 1500

aaagacgcag tcgtcacgcg tggcctcgtc taa 1533

<210> 150
<211> 1794
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 150
atggacatca tcggacttcg gagccggagc cacctgggcc tgaaacagat caaggtcact 60

agagacatga aggagggtca gtcggtgtcc ggagaagaag aagagcctct gagccccatg 120

gcccgaatgt tccacgagcc cgactctaat gtcaaggaaa tcaacgacgc actgcacgcc 180

attcctatga tcctgcaaac actcacgtgt agcaacatga actctgccta tagacagggt 240

agacgtgtcg cagacgaaaa caacggtgga gagcttaaat gggtcaagac ggaggtcgaa 300

cttgaaaagc atgttattat cccccgtgtt gatgaggaga tggcaagcca aggcgccgct 360

gacaagttcg tcgaagatta tatcgctaac atctccaaga ctaaaatttc gctctccatc 420

cccatgtggg attgtcatat tcttaacatc aaaacatccg atgccgaatc tgtcctggtt 480

ctgctggttc atcactcgct gggagatgga acttccctga tgtccctcct tatctcctgt 540

tcccggaagc tgtcggatct tctcgccctc ctcactttcc ccgctatgaa gaagaagcag 600

atgcctacca ccacgacaac tacttggctt tgtttctgga tccggctctg gagcttcttt 660

cttcttattt ggaatacgtc ggttgacctc ctggtgtgtg ttgctactct ctttttctac 720

aaggacacgc ccacgcccct caagcctagc tctagacgag tggcttgtac ccctggtcga 780

attatgcgac gaacgttcac tcttgacgac atcaagcttg tgaagaacgc cacgaatatg 840

actgtcaacg acgtcgtgct ggctatcacg caggctggtc tttcccgtta ccttaaccgg 900

aaatatggaa agacgaagcg taatgaggcc ggtcgaaagt gggaagataa cctccctaac 960

aatattcggc tccgtgctac acttttcatt aacctcagat cctcgcctgg catttatgca 1020

cttggtgaga tgctcaagaa gaactctaag gctgaatggg gcaataaaat tggttacgtg 1080

cttttcccct ttaccatcag cctgaaggat aatctccttg actacatccg ggacgtcaaa 1140

gcagccactg atcgtaagaa agcctacctt gaagcaaagt tccgtcttct gatggccatg 1200

gtcttcgtca gattttaccg gacaaaattc gccaaattcc cttctacaac aatgtggttt 1260

tcgaacgtgg cacgtcctca ggatgacatc accattttcg gaaatcaagt cgcctacatt 1320

gctccctctc tgtacggtca gcttgttgct ctgacagtcc acgtggtttc gtacgctaaa 1380

atgatctcta tggtcctctc tgtggatgat aacatcatct cggaccccta ccaactctgt 1440

gacgaccttg aagagtcgct taaactgatt aaaacgtcgg tggtgtcgcc tcttgtccga 1500

gtccgacttt cctacagcag aactaagtcc cggtccacgc ttgtgatcct ggctaagttc 1560

ccctcgacca cgatttggtt tagcaatgtg gcaagacccc aggatgagat tacaattttc 1620

ggcaatcagg tcgcctatat tgcaagcagc gtctacggtc aacctaaggc acttactatt 1680

catgtggtta gcttcgctaa aaaaatctcc caacctgttg ttctgcttgc cggacatccc 1740

gtgcctctga cacccatgct tcctcggcga gctggtgccg acgctagcct ttaa 1794

<210> 151
<211> 1545
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 151
atggatatta ttgacctgcg gagcggatct caccttggac tcaagcagat taaagtcaca 60

agagagatgg aagagggtcg aagcgtttct ggcgaagata atgaacccct gtcgcctatg 120

gctcgaatgt tccatgagcc cgattccaac gtttacatta ttactattgt cggatttaaa 180

aaccctattg agcccaactc ctttaaagca aaccttgttc acactctgct gaagcatccc 240

agattctcgt cggttcaagt cgccgacgag aacaacggag gagagctcaa gtgggttcaa 300

accgaggttg agctggaaaa gcatgtgatc gtccctaagg ttgatgaaga gatggcatcg 360

caaggcgccg cagataaatt catcgaggat tacatttcca atatgtccaa gactaagatt 420

tccatgtcta tccctatgtg ggactgccac atccttaatc tgaagacgtc cgacgctgag 480

tcggtcctcg ttctgcgtgt tcatcattcg ctcggagacg gaacatctct catgtccctg 540

cttatctcct gttcgagaaa acttttgac cctctggcac tgcccacgtt tcctgccatg 600

aaaaagaagc ctattgctac aacgacatgg ctgtgcttct ggatcaaact gtggtccttt 660

ttccttctta tctggaacac gctggttgat atgcttatgt gcgttgctac actctatttc 720

tataaagaca cgcctacacc cctcaaacct cctagccggt cggttgcatg tacgcctaaa 780

agaattgtcc gaagaacttt ttccctcgat gatgtgaaac ttgttaagaa tgccacaaat 840

actctgatct ggcttattag agtttacctt aaatgttttg ctattggccc tttcgtggac 900

cgacagcgat gtcggccctc caatcattct tccagactga tctccctgga catttctacg 960

gtcaatatgc ggcacgaggc aggtaaagaa tgggaagata accttcccaa caatattaga 1020

ctcagagcca cgcttttat caatctgcgt agcagccctg gaatctacgc cctcggtgag 1080

atgctcaaaa aaaattcgaa agccgagtgg ggaaacaaga tcggctatgt gctttttcct 1140

tttactatcg cccttaagga taatcctctc gagtatatcc gagacgttaa agaggctatg 1200

gatagaaaga aggcatctct cgaggctaaa tttagacatc tgatggcaac cgtgttcgtg 1260

cggttttacc gtacgagact cgcaaagttc ccctccacta cgatgtggtt ctccaacgtg 1320

gctggtcctc aggacgaaat caccattttc ggaaatcagg ttacttacat cgccccctcg 1380

ctgtatggtc agcctgttgc ccttactatc catgtcgtct cgtatgccaa gaaaatgtct 1440

atggtcctca gcgtggacga taacatcatt cctgatcctt atcagctttg cgatgatctc 1500

gaggagtcgc tgaagctcat taaaaaatct gtcatttcgc aataa 1545

<210> 152
<211> 2013
<212> DNA
<213> Artificial Sequence

<220>

<223> codon optimised ORF

<400> 152

atgtcttttc tgcgacgtag aaaagtcacc gattcgagca aatctcaaaa ctccgactct 60

aacgccggta aggaagacga caagaagaag cagtataaat cccccaagcg agagtcctat 120

agctccaaaa aatggagctg ctgggactcc tgttgctggt ttattggtct gatttgctcc 180

atgtggtggt ttcttctttt tctttataac gctatgcccg cttctattcc ccaatacgtt 240

acggaagcaa tcacaggccc cctgcctgat ccccccggag ttaagcttcg taaagaagga 300

ctgacagtca accatcccgt tgtcttcgtg cccggcatcg tcactggcgg cctcgagctc 360

tgggaaggcc atcactgtgc tgacggcctt ttccggaagc gactttgggg tggatctttc 420

ggtgaacttt ataaacgtcc cctctgctgg gctgagcaca tgtccctcga caacgaaacg 480

ggactcgacc ctcccggtat tcgagtgcga cctgtgtccg gcctggtcgc agccgactac 540

tttgccgcag gctactttgt ctgggccgtc cttatcgcca accttgcaca aatcggttac 600

gaggaaaaaa ctatgtacat ggccgcatac gattggcgac tcagcttcca aaatacagaa 660

gtgcgggacc agactctctc tcggatcaag tctaatattg agctgatggt cgctacgaat 720

ggtggaaaga aggtggtggt tattcctcac tccatgggtg ttctctactt tctccacttt 780

atgaaatggg ttgaagctcc tgtccccatg ggtggcggag gtggctcgga ttggtgtgca 840

aaacatatca aggccgtgat gaatattgga gccccctttcc tgggcgttcc taagtctgtt 900

tccggactct ttagcattga ggcacgggat attgcaattg ctcgagcatt tgcacctggc 960

ttccttgaca aggacgtttt tggtctgcag acgttccaac atctcatgcg gatgacccga 1020

acttgggaca gcaccatgtc gatgatccct aagggaggag acacgatctg gggcggcctg 1080

gattggtccc ccgagggagg ctcctttaac tgctctgcaa agaaactcaa aaataacgga 1140

acccgaactg gacataatgc caacagcaac ctgggtaata tgaagtcggt caattatggt 1200

cggattatct cgttcggcaa ggacgtggct gaagcacatt cttccaaaat cgaaagagtc 1260

gattttcgag acgtcatgaa gggaggcaag ctggcaaata gctctaactg tgatatctgg 1320

acagaatatc acgaaatggg taacggagcc attaaggctg tggcagatta taaagtgtac 1380

acggctggct ctatccttga tctgctgcat tttgtggccc ctaaacttat ggctcgtgga 1440

ggcgcacatt tctcgtatgg tatcgcagac aatctggacg accccaagta cgaacactat 1500

aagtattggt ctaaccccct ggaaaccaag ctgcctagcg ctcccgatat ggagatttat 1560

tcgatgtacg gagttggaat ccccacagag agagcctacg tctataaact cacaacggct 1620

actgattgct atatcccttt tcagattgac acttctgcag aaggaggatc ggaaggttcg 1680

tgtcttaagg gaggagtctt ttcggtcgat ggtgacgaga ctgtccccgt cctgtctgcc 1740

ggctttatgt gtgcaaaggg ctggcgagga aagacgcgat tcaatccctc gggtatccga 1800

acttatatcc gagaatacaa tcatgctcct ccttctaacc tcctcgaagg tcggggaaca 1860

cagtcgggcg cccatgtgga catcatgggt aactttgccc tgattgagga catcattcgg 1920

gtcgccgctg gcgccacggg tcaagatctg ggcggcgata gagtctattc ggacatcttc 1980

aagtggtcgg aaagaattaa ccttcaactg taa 2013

<210> 153
<211> 2076
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 153
atggcttcga ttcttagatt tcggaagctg tgctacgtcg aacaggctgt gaaatgtgct 60

tcggtcggct ttgaatcctt tgaccagtcg cccaaaattg accaaaagct cgacaagaag 120

gaggaggaag tcattagcgc aaataatttc gcacttgaga ttatcaaaaa gcggaaacaa 180

cctcggaagc aacctaagga gtggcggcga ctggattcgt gttgctggat gatcggctat 240

ctgtgttcta cgtggtggct ccttctgttc ctttatcact ctcttcctgt tacgctcctt 300

caagtccccg aatctcccgg tgttcggctt aaacgagagg gactcgcagc ccttcaccct 360

gtggtccttg tccccggcat cgtcacgggt ggccttgaac tctgggaggg tcaaccctgc 420

gcagatggcc tgtttcgtaa gcggctttgg ggtggcagct ttactgagat ctttaagcga 480

cccctgtgct ggcttgaaca cctgtctctc cactcggaga caggtcttga ccccccctggt 540

attcgggtgc gtgccgttcc cggactggtc gctgcagact acttcgcacc tggctacttt 600

gtctgggccg tcctcattga gaacctggct aagatcggct acgagggtaa aaacctgcac 660

atggcagcat acgactggag actttcgttt caaaacaccg aaatccggga tcacgctctg 720

agccgactca aaagcaaaat tgagcttatg tatcttacga acggctataa gaaagtcgca 780

gtggtccccc actcgatggg cgtcatctat tttctccatt ttctgaagtg ggtggagact 840

cctcccccca tcggcggcgg cggtggaccc ggctggtgcg ctaaacatat taaggcaatt 900

atgaatattg gccctgcttt cctgggcgtt cctaaagctg tgtccaacat tttctccgct 960

gaaggtaagg atattgccta catccgtgct atggctcccg gtcttctgga cagcaaaatt 1020

ctgggcctcc aaacgctgga gcatgtcatg cgtgtttctc ggacctggga ctccattgtc 1080

agcctggtcc ctaagggagg tgagacgatt tggggcaaca tggactggag ccctgaagag 1140

cgtcatgcat gcgatttttc tcgtaagcgg cactcccagc ccagccccat tgataataat 1200

gttaacaatt ctgacgtgaa gcgaggcttc agagtgaagg accccgtcca gtatggacgg 1260

attattagct tcggaaagac ggcctctcag ctgcatagct cccaacttcc cactatcgac 1320

tctaaggaat ttctccggac cagcgcctcg caaaatctga acttctcctg cggagaagca 1380

tggaccgaat acgacgagat gagccgggag ggcattcaaa aagtcgctgc aaataaagcc 1440

tacaccaccc aaactctgct tgatctgctc cgtttcgttg cccctaagat gatgcttcga 1500

gccgaggccc attttagcca tggcatcgct gacgacctgg acaaccctaa gtataatcat 1560

tataaatatt ggtcgaatcc cctcgagatg aaactgcccg atgccccga tatggagatc 1620

tactgtcttt acggcgtggg catccctacc gagcggagct atgtgtataa gctttccccc 1680

acgtctggct gcaaatccat cccttaccag attgatagct ctgttaatgg agaggacgga 1740

tcttgtctca agggtggcgt gtacttcgct gacggagacg aaaacgttcc cgtgcttagc 1800

gctggtttca tgtgcgctaa aggttggcgg ggacggacac gatttaaccc tagcggaatc 1860

gctacctata ttcgagaata ccgtcataaa ccccccctcgt ctctccttga gggacgtgga 1920

atcgagtccg gtgctcacgt cgacatcatg ggcaatgtgg ccctcatcga ggacattctc 1980

cgaattgctt ctggcgctac cggtaaagaa attcgtggcg accgaattta ctccgacatt 2040

ctcagaatga gcgaacgtat taaccttaga ctttaa 2076

<210> 154
<211> 2058
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 154
atgagcgcca tccgtagaag aaaacctatc aacgagtcgg acaaaactag cgaagcctcg 60

gattcgaaac cccacaaaga gaaggaggag gaacagcacg gtgacgatga tggctcggac 120

aaagataaaa aaaaaatccc cagcaagatc aaaaagaaac acgaagaaaa gcaacccaaa 180

tggtcctgcc tcgattcctg ctgctggttc attggatgta tttgcgtcac gtggtggctc 240

ctcctctttc tctacaacgc tatgcctgct tcctttcccc aatacgtgac tgaggcaatt 300

actggccccc ttcccgaccc ccctggcgtt aagctcaaga aggaaggtct gaaagcaaag 360

catcccgttg tgttcgttcc tggcattgtg acgggtggcc tggagctctg ggagggaaga 420

aaatgtgcag aaggactctt tcggaagcgt ttctggggtg gtaccttcgg tgaagtgtat 480

aagcgacccc tctgttgggt ggaacacatg agcctggata atgagacggg tcttgaccct 540

tgtggcatcc gggtcagacc cgtgtccgga ctggttgctg ccgattattt tgcccctgga 600

tactttgtct gggccgtgct cattgcaaac ctggctagaa ttggatacga agagaagaca 660

atgtatatgg ctgcctatga ttggcggctc tccttccaga acacagaagc acgagatcaa 720

acactcagcc gtattaaatc gaacatcgaa cttatcgtgg ccaccaacgg cggaaaaaag 780

gcagttgtca tcccccattc catgggcgtc ctctattttc tccatttcat gaaatgggtg 840

gaggcccctg cccctatggg aggaggcggt ggacctgact ggtgctccaa gcacattaag 900

gctgtggtca acattggagg tcctttcctg ggcgtcccta aggcaattgc cggacttttc 960

agcgccgaag ctaaggatat tgccgtggtt cgtgcaatcg cacccggatt tctcgacaat 1020

gacatttttc agctccaaac ccttcaacac gttatgcgga tgtcccgatc gtgggattct 1080

accatgtcta tgatcccccg aggaggtgac acgatctggg gcggcctgga ttggtcccct 1140

gaggagggtt acagctgcgc caaaaaacgt gagcgaaaaa atgacacaca gatcgcagag 1200

gaggccggcg tggaatcggc tgtgtcccag acacggtctg caaaatatgg aagaatcatt 1260

tctttcggca aagatgtggc tgaggcacct tcgtctgacg tcgagagaat tgacttcaga 1320

gacgctgtta aaggccactc ggcagcaaac acgacctgtc gtgatgtctg gacggaatac 1380

catgacatgg gtttcggcgg tatcaatgct gtcgctgaat acaaaacgta cacagccgag 1440

agcattgtgg accttctcca cttcgtggcc cccaagatga tggctcgtgg aactgcccat 1500

ttttcttacg gcatcgccga caatctggat gatcccaagt acaagcacta taaatattgg 1560

agcaaccctc tcgagactaa gcttcctaac gcccccgaaa tggaaatctt ttccctttat 1620

ggtgtgggcc tccccaccga aagagcatac atttacaaac tgtcccccgc tgcagagtgc 1680

tacatcccct ttcaaattga tacctcggca gacgatgaag agacatgcct gaaggacgga 1740

gtgtatagcg tcgacggcga tgagactgtc cctgttctgt cggctggctt catgtgcgct 1800

aaaggttggc gtggaaagac tcggtttaat ccctctggta tccggactta catccgtgaa 1860

tactcccact cgccccctgc taacctcctc gaaggccgag gaactctgtc cggcgcccat 1920

gtggacatta tgggaaattt tgcactgatc gaagacgtga tcagagtcgc cgccggagct 1980

tcgggtgaag agcttggcgg cgatcaagtg tactctaaca tctttaactg gtcggaaaaa 2040

atcgatcttc agctttaa 2058

<210> 155
<211> 1503
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 155
atggcccctg ctaaatccgg cagatcgttt ccctctatca cgaaatgcga gggtagcacc 60

tatgagagca ttgccgctga tctggacggc acactgctca tcagccgatc ttcgtttcct 120

tactttatgc tcgtggcagt ggaggcagga tctctgctcc gaggactcat tcttctgctg 180

tctcttcctc tcgttattgt ttcctatctt ttcatttccg aggcaattgg tatccagatc 240

ctgatcttta tcagcttcgc tggcctgaag atcagagaca tcgagctcgt tagcagagct 300

gttctgcccc gtttttatgc cgctaacgtc cgaaaggaaa gctttgaagt gttcgatcgg 360

tgcaaacgta aagtggttgt gacagcaaat cccacattta tggtggaacc ctttgttaag 420

gacttccttg gcggcgacaa agtgctcggt acggagatcg aagtgaaccc caaaacaaaa 480

aaggctactg gatttgtgaa gaaacccggc gttcttgtgt cggagctgaa acggctggcc 540

atcctgaagg aatttggtga ggaatctccc gacctcggca ttggcgatcg tgagtctgat 600

catgatttca tgtcgatctg taaagaaggc tacatggtcc acccttccaa gagcgccact 660

cccgtgcccc tggatagact taaatctcgg atcattttcc atgacggccg ttttgttcag 720

cgtcctgacc ccctgaacgc ccttatcacc tacctttggc ttcctttcgg ttttattctc 780

tctatcgtcc gtgtctactt taatcttccc ctgcctgagc gtattgtgag atacacgtac 840

gagatgctcg gcatccacct tgtcattaga ggaaagcgtc cccctcctcc ctctcccggt 900

acacccggaa atctttatgt ttgtaaccac cgttccgccc tcgaccccat tgttatcgca 960

attgctcttg gtcggaaggt ttcgtgcgtt acctattctg tctcgcgact gtcgagattt 1020

ctctccccta ttcccgcagt cgccctgact cgagatcgtg ccgccgacgc cgccagaatt 1080

agcgagcttc tccaaaaagg tgatctggtc gtctgccctg aaggtactac atgccgagaa 1140

cagttcctgc ttcgttttag cgccctgttt gcagagctgt cggatcgaat cgtgcctgtg 1200

gccgtcaact gccgacagaa tatgttttat ggtacgacgg tgcgtggtgt caaattttgg 1260

gacccttatt tcttcttcat gaaccctaga cctacttatg aagtgacgtt cctggatcgt 1320

ctgcctgagg aaatgacggt caaagctggc ggaaaatcgt cgattgaagt tgcaaaccat 1380

gtccaaaagg tgcttggcga cgtgctgggc ttcgagtgta cgggcctcac tcgaaaggac 1440

aagtacctcc tcctcggtgg taacgacgga aaggtcgagt ccatgtacaa cgcaaaaaag 1500

taa 1503

<210> 156
<211> 1131
<212> DNA
<213> Artificial Sequence

<220>

<223> codon optimised ORF

<400> 156
atgagctccc gtggcggtaa gctgtcttcg tcctcgtccg agctcgatct ggacggaccc 60

aacatcgagg attatctccc ttcgggttct tcgatcaacg aacctcgtgg caagctgcgg 120

cttcgtgacc tgcttgacat ctctcccacg ctgacggagg ctgccggtgc aattgtcgac 180

gattctttta cccgttgttt caagtcgaat cctcccgagc cctggaactg gaacgtttac 240

ctcttccccc tctggtgttt cggagttgcc gtccggtatc tcattctttt ccctgcccgt 300

gtcgttgttc ttactatcgg ttggatcatc tttctcagca gcttcatccc tgtccacttc 360

cttctcaaag gtcacgataa gctcagaaaa aaaatggagc gtgtcctcgt tgaactgatg 420

tgctcctttt ttgtggctag ctggacagga gttgtcaaat atcacggacc ccgaccctct 480

attcgtccta aacaggtgtt tgttgcaaat cataccagca tgatcgattt tatcattctt 540

gagcagatga cagcatttgc cgtcattatg cagaagcatc ctggatgggt cggactgctg 600

caaagcacca ttctggaatc tgttggatgt atttggttca acagatcgga agccaaggat 660

agagagatcg tggctaagaa acttagagat cacgtccaag gagtcgacaa caacccctg 720

ctcatcttcc ctgagggaac ctgcattaac aaccagtatt cggtgatgtt taaaaagggc 780

gcctttgagc ttggctgtac agtgtgcccc atcgcaatta agtacaacaa gatttttgtt 840

gatgccttct ggaactctag aaagcagtcg ttcacgatgc accttcttca actcatgaca 900

tcctgggccg tcgtctgtga cgtttggtat ctggagcctc agaaccttcg acccggcgaa 960

actcccatcg aatttgccga acgggtcaga gacattattt ctgttcgagc cggcctcaag 1020

aaggttcctt gggacggtta ccttaaatat tccagaccct cgcccaaaca tcgggagcga 1080

aaacaacaat ccttcgctga gtctgtgctg cttcgacttg aggaaaaata a 1131

<210> 157
<211> 1623
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 157

atggtgttcc ccgtggtgtt cctcaaactt gctgactggg ttctttacca actccttgca 60

aactcgtgtt atcgagcagc ccgtaaaatg agaaattatg gctttttcct gagaaatcaa 120

acgctgagat ctccccccca acagcaagcc gcatcccttt ttccttccgt tacaaagtgt 180

gacgtgggaa actctcgtag attcgataca cttgtctgcg acattcacgg cgtgctcctg 240

ggctccgaca cattttccc ttatttcatg ctcgtggctt tcgagggtgg ttccattgtg 300

cgggctttcc ttctcctcct ttcgtgtagc tttctctggg ttcttgactc cgagctgaag 360

ctgagaatca tgatcttcat ttccttttgc ggtcttcgga agaaggacat tgagtctgtg 420

ggccgagccg tgcttcctaa gttctatctt gagaacctta atctccaggt gtatgaagtt 480

tggtcgaaaa cttctagccg tgtcgtgttc acaagcattc ctcgagttat ggtggagggt 540

tttctccatg agtatatgtc cgcatcgggc gtcgttggta ctgagctcca cacagtcgga 600

aaccgattta ctggcctcct ctcctcttcg ggccttctcg tcaaacataa cgcactcaag 660

gaacattttg gcgacaaaaa gcctgatgtg ggactcggtt cttctagcct ccacgatcag 720

tactttatct ccctgtgcaa ggaggcttac gtggttaaca tggaggacgg taaatccaac 780

ctctcctcct ttatgcccag agacaaatac cccaaacctc tgatttttca cgacggccga 840

ctcgcattcc ttcctacacc ctttgctacc ctgagcatgt tcctttggct tccttttgga 900

attgtcctct ccattcttcg tattttcgtc ggaatctgtc tgccttataa actcgccgtg 960

atctgtgcaa cactttctgg tgtgcagctc aagtttcaag gttgctttcc ctccagcaac 1020

agccagcata agaaaggcgt cctgtacgtc tgtacccaca gaacacttct ggatcccgtg 1080

ttcctgtcga cagcccttg taagcctctg accgcagtga cgtattcgct ctcgaagatg 1140

tccgaactga tcgcacctat caagacagtt agacttacac gggatcggaa acaggacgga 1200

gagactatgc agaagcttct ttcggaaggc gaccttgtgg tgtgccctga gggtacaacc 1260

tgccgggaac cttatctcct gcgtttctcc agcctcttcg ctgagctggc cgatgagatc 1320

gtccctgtcg caatcaatgc ccacgtgtcc atgttctacg gaactacggc ttctggcctc 1380

aaatgtctcg atcccatttt cttcctgatg aatccccggc cctcctatca cgttcaaatc 1440

ctgggaaaag ttccccaaga atttacgtgc gctggaggcc ggtcgtctct ggaagtggcc 1500

aactacatcc agcggaagct tgccgatgca cttggatttg agtgtacgac ccttactcgg 1560

cgagataagt atcttatgct tgcaggaaat gagggagtcg tccacgagaa caagagaaac 1620

taa 1623

<210> 158
<211> 1614
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 158
atggccaaac ttagcatgga gttctcgttc ttccaaacac tgttcttcct cttctgtaga 60

gttgtctttc gacagtcgaa aaaccacaaa tctcttcata gaaacgtcag caacattcac 120

gcaaatgaag gcaaatacca taaatatcct tcgttcgttc accggagcaa cctctcgaac 180

caaactctcg ttttctccgt tgaggaagcc ctgctcaagt cgtcttccct ctttccttac 240

ttcatgctgg tcgcctttga agctggtggc cttctccggg cattcattct ctttgtcctg 300

taccccattc tttgtctggt ttctgaagaa atgggactta agattatggt cctggtgtgc 360

ttcttcggta ttaagaaaaa atcgttccgt gtgggttctg ccgttcttcc caagttcttt 420

cttgaagatg tcggccttga gccctttgag atgctgaaga aaggcggaaa aaaggttgca 480

gtgtctaaaa tcccccaagt gatgatcgag agctttctta aagactatct tgaaattgac 540

ttcgtggttg gtcgagagct gaaggagttc tgcggatact tcctgggtgt tatggaagag 600

aaaaaacggt cgaaggctgc tcttgatgag attatcggtt cggaaagcat gggttttgat 660

gtgattggta tctccggact caagaagtcg cttgactatc acttttttag ccactgcaaa 720

gaaatctatc aagtccgtaa ggcagataag cgaaactggc gacatgtgcc cagacaagag 780

tactccaagc ccctgatttt tcacgatgga cgactcgccc tgcgacccac tctcgttgcc 840

tcgctcacta tgtttgtttg gtttcctttc ggcctcgctc tgtctattct gagagctgtg 900

gtcggcctga tgctgcctta caagatttcg atccctctcc tcgcttattc gggcctccat 960

ctgttcctct cgacacctga aagcagcctc catcctcttt ctctgcccaa ttctaaaaaa 1020

cagaacccta aaggaagact gtacgtgtgt aaccaccgta ccctgctgga tcccgtgtat 1080

ctttcttttg cccttcagaa ggatctcact gcagtcacct attccctgtc tcggatttcc 1140

gaactcctcg cacccattaa aacggtccgt ctcgcccgtg accgagatca agatggcaag 1200

atgatggaaa agatgcttaa cctgggcgac cttgtcgtgt gccccgaggg aacaacgtgc 1260

cgggagcctt accttcttag attttcgcct ctcttttctg agatgtctga tgacatcgtg 1320

cctgtggcaa tggactccaa tgtgtccctt ttctacggaa ccacggcatc tggtctgaag 1380

tgtctcgatc ctctcttttt cctgatgaat cctcggccca tctacacggt gcaaatcctt 1440

gatggagtgt cgggactcta tacctgccat gacggccaac gaagccgatt taaagtggcc 1500

aaccaagtcc agaacgaaat tggtaaagca ctcggtttcg aatgcacgaa actgactcga 1560

cgagataaat acctcatcat ggctggtaac gagggtatca tttcgcagac gtaa 1614

<210> 159
<211> 1467
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 159
atggagtccg ttgtcagcga actggaaggt gtgcttctta aagatcctga ccctttttgt 60

tattttatgc tggtggcctt tgaagcatcg ggacttattc gatttgctct tctccttctt 120

ttctggcccc ttattcgtct ccttgacatg ctgggacttg atgatgctgg acttaagctg 180

atgacgttcg tcgctacagc cggactgcga gaggcagaga ttgaatctgt ctctagagca 240

gtgctgccta agttctacat ggatgacgtg gacatggagg cttggaaggt tttctcctct 300

tatgataaac gagtggtggt gacgaagact cccagaatca tggttgcaag attcgttaaa 360

gagcacctcc gggcagacga agtgctgggt tcggaactgg tcgtcaatcg gttcggattt 420

gctactggat ttgtgaaggg tgatattgcc tccatttcgt cgcgtgtggc caagctcttc 480

gttgatgaag aacctaccct gggcctggga cgagcaccct cgtccttcca attcctgtcg 540

ctctgtaaaa aacagatgca ccctcccctt atgactaatc aaaacctcca tgatcaccag 600

cttctgcggc ccctgcccgt cattttccat gatggacgtc ttgttaaaag acctactcct 660

tcgactgccc tcctcattct tctttggatg cccctcggca ttcttctggc tacaattcga 720

attgtggtcg gtcttattct ccccatgcga gtcattcctt acatgtctcg actgtttggt 780

ggtaaaatca ttgtgaaggg taagcagcct cctcctgttt ctggctcctc ctcgggtgtt 840

ctgtttgtct gtacacatcg aacgctgatg gatcccgttg tgctttccgc tgttcttatg 900

cgaaagattc ccgctgtgac gtactccatt tctcggctca gcgagattct ctcgcctatt 960

cccacggtcc gacttactcg gatcagagag gttgacgcag agaaaattaa gcgtgagctg 1020

gcaaaaggtg atcttgtggt ttgccccgag ggtacaacgt gccgagagcc ctttctcctt 1080

cggttcagcg ccctttcgc agagctgacc gaccggatcg tgcctgtggc tatgaactac 1140

cgggtgggct tctttcacgc tacaactgct cgtggctgga agggactcga tcccattttt 1200

ttttttatga accctcgtcc cgtctacgag gttacatttc tgaaccagct gcctgctgaa 1260

gctacttgta gctctggcaa atctcctcac gatgttgcaa actatgttca acgtatcctt 1320

gcagccaccc ttggtttcga atgtacgaat ttcactcgaa aagacaagta tcgtgttctc 1380

gcaggcaacg atggaacggt ttcctgcact tctttcgtcg atcaggtgaa gaaggtcgtc 1440

tcgactttca agcccttctt tcaataa 1467

<210> 160
<211> 1782
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 160
atgaaactgt ctagacgaca aactgttctt accaatgaca acaacaactg caacctcaaa 60

ctcgcagctt ggccttgttc gtttggatat aagagacgga aagcccggca gacatccgca 120

atccagccta agtatcaagt tcccaccgcc tcgtactctt ctgccaccgc ccaacggtcg 180

aaagcatctc ccagactcat ccactggctc gccgtgaagg attcgtggaa gtctcgagcc 240

cttaacgttc ttattctgct gttcctcaac atggccatca atgctttccc tactgtgcag 300

aactgtgcct cgatcggccg agaaaaacac accatggttg ctgatatgga cggaacactc 360

ctcatcagcc gtagcagctt cccttacttc gctcttgttg catttgaggt cggcggagtt 420

ctgagactgt tcttccttct ccttgcatcc cctctggcag gccttctgta ttactttgtc 480

tcggagtcgg ctggaatcca agtccttatc ttcgccacct atgtcggcat gaaggttagc 540

gacattgaat ccgtggctcg tgctgttctc cctaaattct acagctcgga tcttcatcct 600

gagtcttggc gggtttttc cagctgtgga aagagatgtg ttcttaccgc taaccccaga 660

atcatggtcg agccttttct taaagatttt ctgggcgcag acatggttct tggtactgaa 720

atcggcattt acagaggccg tgctactgga ttcgtctgtc agcccggagt ccttgttggc 780

aaaaacaaag ctgacgcact ccaaaaagcc ttcggcgagg ctcttcccga tgtcggactc 840

ggcgaccgac atactgatat tccctttatc gctctgtgta aggaaggata catcgttccc 900

cctaagcctg aggcaaaagc cgtcagatcg gagaagcttc ccaagcctgt catttttcac 960

gacggccggc ttgtgcagaa acctacacct ctgatggccc tccttattat tatttggatt 1020

cctatcggat tcctgctggc ctgccttaga attaccgccg gcagcctcct ccctatgcct 1080

ctcgtctatt acgctttctg ggctctcgga gttagagtta ctgtgaaggg cactcctccc 1140

cctcccgtca agaagtccac cggccagtcc ggcattctgt ttatctgttc tcaccgtaca 1200

ctgcttgatc ctatcttcct cagcacagca ctgggccgac ccattcctgt tgtgacctac 1260

tccgtttcca gactttctga gatcatcagc cctattaaaa ccgtctccct ttctcgtgat 1320

cgggccactg atgcatcgat gattaaaaag cttcttcagg agggtgacct tgccatttgt 1380

cccgagggaa caacctgccg ggagcccttt ctgctccggt tctcggcact cttcgctgag 1440

cttacagacc aacacgttcc tgtcgcaatg gtgaatagaa tgtcgatgtt ccatggaact 1500

accgccagag gctggaaggg catggatccc ttttactttt ttatgaatcc cagccccgcc 1560

tatgaggtga cattcctcaa caaacttccc atggaactca cctgctcgtc cggaaaatcc 1620

tcgcatgaga tcgcaaacta cattcaacgt gtcattgctg ctactcttag ctatgagtgt 1680

acgtctttca ctcgtaagga taagtataga gctctggcag gcaacgacgg aacggttgtc 1740

gaaaaaccta aattcagcgc caacaaagtg atgggttgtt aa 1782

<210> 161
<211> 1317
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 161
atgggtagcc gtattcttcc ctctgcctcg ctttttctct gtagcgtctt tctttcgagc 60

ttttgtgcta gctgccgatc ttccctcagc attcctcccc ttcgggccag ctttccttgc 120

gcctctagca agtctcttgc ccccatttgc cggctccctc tgcctcatgc ctgtggactc 180

ggtatcctcg tccttttcac tgctagctcg ctcgccatgg ctgatgtcgt ggcagatgtg 240

gtcgcccccg tcaagcccaa aaatcgtgca tccaaaaaat ttagctcccg tggagttgac 300

ctcgacgtgc tgcttgagat gtcttctagc aactcctctc ccctcctgtc gcccaaagtg 360

ttcctccgtg acctcaagcg aagacctctt gctctcacca agaagattct ttcctgggca 420

ctggaaggtg ccagaagcac agcattctct ctgggtaccg gtggtaaatt tcactatccc 480

gcccgtcggg aatttttcgc aagcggccgg ctcctctttc ttctgcttgc cagccctctc 540

gcaggtacgc tgtactattt cacgagcgag tcccatgaca tccgtgtgct tatttttgca 600

acctttgtgg gaatgaaagt gtcggacatt gaatcggtcg cccgggcagt cctccccaag 660

ttctacagct ccgacctgca ccctggattt cgattcatgt gggaggagat gtgcgcctat 720

ggtgattcgg agggccattc tggcgtgaag tctcccggtg tcctcgttgg aaagaacaaa 780

gcagatgcac tcgaaaatga gtcccatgtg gtgcctgcta aagcagaggc aggtgcagtt 840

tcgcatgaca aactccctaa acccatcgtt ttccatgatg gccgtcttgt ccaaaatcct 900

acccccttca tggctctcct tacgatcctg tggatccctg tgggctttct gcttgcctgt 960

cttagaattg ccgtgggcat cctgctcccc acccctctgg tctattacgc tatgtggcct 1020

cttggattta gaatccatat taagggcacc cctccccaac ctacaaaaaa gtctatcggt 1080

cagacggtcg tcctgttcgc ctgctcgcat agaacactcc tggaccctat tgctctcagc 1140

gtcgcacttt gcagacctat tcccacagcc acgtattcca ttagcggact ctcggagatc 1200

atctcgccta tcaaaaccgc acggctctgt ctggatcgtg ctactgacgc aagcattatc 1260

aagaaacttc tcgaaaaggg tgatctcgtc atctgccccg agggtactaa aacataa 1317

<210> 162
<211> 1524
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 162
atggctaaga ccaaggaaaa ccccttttcct agcatcgccc aatgtgcttc gatcggtcga 60

gacaagcata cagtcgttgc tgatatggac ggcaccctgc tgagaggaag aagcagcttc 120

ccctactttg ctcttgtcgc atttgaggtc ggtggcatcc tgcggctgct ctttcttctg 180

ctggcttccc ccgtggctgg catcctgtac tactttatct ccgaggccgc aggcattcga 240

gtcctcattt ttgctacctt tgtcggtatg aaggttagcg atatcgaatc tgtcgccaga 300

gcagttcttc ctaagtttta ttctagcgat ctgcaccctg aaacctggag agtcttttct 360

tcgtgcggta aaagctgcgt tctcacggct aatcctcgag tcatggtgga agcatttctt 420

aaggattacc ttggcgcaga catggtgacc ggcaccgaga ttcatgtttt ccgaggccga 480

gcaaccggcc tggttaaatc ccctggtatt ctggtgggta aaaacaaggt tgatgctctg 540

gaaaatgcat ttaaggataa gcctgtggct gacatcgcac tcggagacag aaagaccgat 600

tatcctttca tgaagctctg taaggagtcg tatgtggttc ctgccaaacc caaagttgag 660

gccgtgagcc acgacaaact ccccaaaccc attgtttttc atgacggtcg gcttgttcaa 720

aagcccaccc cttttatggc acttctgacc attctttgga ttcccgtcgg tttccttctc 780

gcctgcctgc gaatcgccgc aggagctctt ctccccatgc ccctggtcta ctacgctttc 840

tgggctctgg gagtccgtgt gcatattaag ggaactcctc cccctcctgc aaaaaaatct 900

attggccaga ccggtgtgct ttttatttgc tcgcatcgta ctctcctcga ccccattttc 960

ctttctactg ccctgggcag aaagatcccc gctgtgacct actcgctttc ccgtctcagc 1020

gaagtgattt ctcctatcaa aactattaga ctttctcgag atcgggctac agacgcctcg 1080

atgattaaga aacttctgga ggcaggagat ctggtcattt gccccgaggg caccacgtgt 1140

cgggaacctt tcctccttag attctctgcc ctgttcgcag aactcactaa cgagctcgtc 1200

cctgttgcaa tggccaacag aatgtctatg ttccatggta ctacggcccg tggttggaaa 1260

ggaatggacc ctttctactt cttcatgaat cctagccctg cttacgaagt caccttcctg 1320

aacaagctgc ctcatgagct tacttgttct gctggaaagt cttctcatga ggtcgccaac 1380

tatatccaac ggatgatcgc cgcatctctg agctatgagt gtacttcgtt tacacggaaa 1440

gacaagtacc gtgctctggc tggcaacgac ggtaccgtcg ttgagaagcc taagctgtcg 1500

cccaataagg ttatgggttg ctaa 1524

<210> 163
<211> 1512
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 163
atggttatgg gcgcacaccg acggtttgag cccatttcga agtgctcgaa cgagggccga 60

tcgaatcaaa cggttgccgc tgacctcgat ggcacgctcc tcctctctcg gtccgctttc 120

ccttatttca tgctcgtggc tctcgaagct ggttctcttg tgagagccct tgtcctcctt 180

gctagcgttc ccttcgtgta cttcacctat ctgttcattt ctgagtccgt ggctattaac 240

acgtttattt tcatcacatt tgcaggcctg aaggtcaaag atatcgaact cgtctcgcgt 300

tctgtcctcc ctaagtttta tgctgaggac gtgcatcctg agagctggcg agttttttct 360

tcgttcggca aaagatatat catcacggca tctccccgga ttatggttga gccttttgtt 420

aagacctttc ttggtgcaga taaagttatt ggtacggaac tggaggtcac caaatcgggt 480

cgagctacag gttttactat taagcccggc gtgctcgttg gagagcacaa gcgagccgct 540

atcagcaagg agttcggtat gaagctgccc gatctcggcc tcggtgatag agagactgac 600

catgacttca tgagcctttg caaagaggga tatatggtgc ctcgatctaa gtgcgacccc 660

cttccccgta acaagctgct ctcccccatc atttttcacg agggacgtct cgtccaaaga 720

cctacccccg tggctgctct gctgacattc ctctggctgc ctatcggttt tatcctcagc 780

ctccttcgag tttatcttaa tatccccctg cctgaacgta ttgcacgtta caactataaa 840

atcctgggta tcaaacttat tgtcaaaggc aaccctcctc ctgcacctaa aaagggtcaa 900

tcgggagttc tcttcgtgtg caaccacaga acagtcctcg atcctgtcgt gacagcagtt 960

gccctcggtc gaaaaatctc ctgcgttacc tattcgatta gcaagttcac tgagatcatc 1020

tcgcccatca gagctgtcgc cctcagccgg gagcgagaaa aagacgccgc caatattaag 1080

agacttctcg aagaaggcga cctcgtcatc tgtcccgaag gtactacctg ccgagaaccc 1140

ttccttctcc gattttcggc cctgtttgcc gagctcactg atcgtattgt gcccgttgca 1200

attaatacca agcaaacggt gttccacggc acaaccgtca gaggccataa gctgcttgat 1260

ccttactttg tcttcatgaa tcctatgccc acttacgaga ttacattcct caaccagctt 1320

cccacggaac tgacatgtaa gggaggaaag tcggccatcg aagttgctaa ttatatccaa 1380

cgtgttctcg caggaactct tggctttgaa tgcactaacc ttacacgtaa agataagtac 1440

gcaatgatgg ctggaaccga tggccgagtc ccttcgaaaa aagaaaagga gcaagagaag 1500

gagaaggcat aa 1512

<210> 164
<211> 1596
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 164
atggccggca aaccccagaa tctcaatgct ttcctcttcc tccagaagat tctcctgacc 60

aaaacttggt cgaaaattta ccttcagcga aaggcaagca atagccacgg attccaacct 120

cagttcaaca aatatccctt tcttgctaat ggtctggaga ctcactcgga caaagcactc 180

atgtttcaca cggagggtgc tctccttaaa tcttcgagcg tgttttctta cttctttctt 240

gtcgccttcg aagctggcgg cctcctgcga gccttcatcc tttttctgct gcaccccatt 300

atctgcctgg tgggcaagga gtggggtctt aaaattatgg tcttcgtgtg tttcgtcggc 360

ctgaaggagg aaagcttcag aattggtcgg actgttcttc ccaagttttt ccttgaagat 420

gttggcaatg agggtttcga tatggtgatg aaatgtggcg gtaagaagat tggtgtgacc 480

gatatgcctc gtgtgatggt cgattgtttc ctcaaagact atctgcgtgt ggaggccgtg 540

gtgggcagag agctgaaagt ggtgtgtggc cactttgtcg gtcttatgga agaaaaaaaa 600

cccagcggta ccgcctttct cgaactggac gaactcctgg gcaaagagaa catggatagc 660

tctgttattg gtatctcgtg ttttgataat tccctcgatc agcagctctt ttcctgctgc 720

gaggttctgt acatggtttc tgaggcagac aagaagaatt ggcaggccct gcctcgtgag 780

aagtacccta gacccctgat ttttcatgac ggacgactcg cttttagacc tacacccctc 840

gcctctctta gaatgtttat gtggatcccc ttcggcatct ttcttcttat cgttcgtgca 900

accgttggtt atcttatccc ttacaagttc tcgactccct ttctgtgttt tactggtctg 960

acaggcacac tcattcggcc tagctcgttt cccgatcggg ctgacgacga acagaaacaa 1020

ggcggcgtgc tgtacgtctg taatcaccga acccttctgg atcccattta cgtctccatt 1080

gtcctggata agcccttaa tgcagtcacg tactctctca gccgaatgac agaactgctc 1140

gcccccatca aaacagtccg gcttactaga aaccgtgagg aagactccaa aatgattggt 1200

aagcttctga aacaaggcga cctggtgatt tgccccgaag gtacgacctg tcgagaacct 1260

taccttctcc gtttctcccc ccttttgcc gaaattacag acgagatcgt tcctgttgca 1320

attgacttcc aagtcagcat gttttatggc actacagcct ccggtttcaa gttcctggat 1380

cctttctttc tccttatgaa tcctacctcg cactgttctg cctctattct cgagaaactc 1440

cccaaatccc tgacatgcca gggtggaggc aagtcgaaat ttgaggtcgc aaaccgtgtg 1500

caggccatta tcgccaaagc cctgaatttt gaatgcactt cgctcacgag aagagataag 1560

tacatgattc ttgccggtaa ccagggaatc gtttaa 1596

<210> 165
<211> 1662
<212> DNA
<213> Artificial Sequence

<220>

<223> codon optimised ORF

<400> 165

atggcaaacc gacagtccgt gctttgggga ctcaagacgg tttgtctcgt ccatcagcac 60

ctcaattcta ccatctatgg caatctgtgc caacaggcca tcaaagtcat taaaaacgat 120

gacgttgcct ctggtattaa tgaacctctt agcattctta atcagcgaaa gctcccttat 180

ctgcgattta aagtctctaa cggtcctgct acgcaattca aattcctgaa acactctact 240

ctcgtgcata gaactgacga gcttgctaat caaaccctcg cctttcacct ggagggcgcc 300

cttctcaaaa gctgttctct gttcccctat ttcatgctta tggcctttga ggctggtgga 360

cttttccgag cccttattct tctccttctg tacccccttg tcctcctgat tggaaaggaa 420

ctgggcctgc ggattctcgt cttcgtctcc ttcgtgggca tcagaaagga gaaattcaga 480

gctggtacag ctattctgcc taaattcttt ctcgaggacg tgggatgtga gggttttgac 540

atcgttatgt cgtacaaaaa gagagtggcc gttactggaa tgcccaaatt tatggttgag 600

ggattccttc gagattatct cggtattgat gccgttgttg cacgtgaact gaaagagttc 660

cgtggctact tcctgggcct gatggaggag aagatggatg ccggtcttgt tatcagcgag 720

tgtacacatc atattggcct cggatgcttt agaaaatcgc atgaccaaaa gattttctct 780

cattgtaaag aaatttatct ggttactgaa gcagagaaga agaactggca agtgctgccc 840

cggaagagat accctaagcc tcttattttt catgacggca gactcgcctt ccgacccacg 900

cctctcgcag ctcttatcat gttcatttgg cttccttttg gatttctgct gaacattacc 960

cgaaccatcg tcttcatttc cctccctttt aagctctcta tccccctgct cgccctcagc 1020

ggatgcatca ttaccgtctc gaagcctgag cctacgactg caagcattaa cggtgagaac 1080

aagcctggag gaatgctgta cgtctgtaat catagaacac ttctggatcc tctcttcctt 1140

tccgcagttc tgatgaagtc tgtgtgtgcc gtgacctatt ccatttcccg gttctcggaa 1200

gttatctctc ccattaaaac tgtgcggctg tcccgtgaca gagagaagga cggtaagatc 1260

atgaaaaaac tcctgtctca aggtgacctg gttgtgtgtc ccgagggaac aacatgtcga 1320

gagccttatc tcctgcggtt ttcccccctt tttgccgaga tgacggatga aattgtcccc 1380

gtcgcaatca agctgcaagt tagcctgttc tacggctcca cggctagcgg actgaagtgc 1440

ctcgactcta ctttccatct gatgaacccc aatcccatgt gctccgtcaa aattctgaat 1500

aaactgcctt cgtgccaaac ccacaatacg ggaggtaaga gcaaatttga agttaccaac 1560

cacgtccaga atcagatcgc agcagcactg ggcttcgaat gtactaatct cacgcgaaaa 1620

gataaatatg ctattctcgc tggaaatgaa ggtattgtgt aa 1662

<210> 166
<211> 1386
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 166
atgagcacgc tgcctcttcc cttcttctct ggaacctcga cacgtgctcc cttttcgttc 60

tctcttaagt cttctcctcc tagcgctgcc agctctagct cttcttcttc tctcgtgccc 120

tttcaacgag tggaactgtt tagatcgaca tcgacgagaa ccacacggtc gttttcgtgt 180

gtcttctatt ccctcaaagc taaggctatg gcagaactcg tgcaagataa agaaagcggt 240

gtcgccgcca ccggttcctc cggtggaggt cgtcctgagg tggagcacag ccgaaccttt 300

ctcgaggcac gaagcgaaca ggaacttctt tccggcattc gaaaagaagt cgaagcaggt 360

agacttcccc ctaatattgc tgcaggtatg gaagaactct accaaaatta ccgaaacgct 420

gttttccaaa gcggtgatcc cagcgcagtc gagattgtcc tttctaatat ggccgtggca 480

ttcgacagaa tgctgctcga tgtggaggac cctttcgttt tcgaacccta ccacaaggcc 540

cttcgggagc ccttcgatta ctatatgttc ggtcagaatt acatcagacc tctgacggac 600

tttcggaact cgtatgtcgg caatctgtct ctcttttatg aaatcgaaga gaagctgaag 660

cagggccata acgtggtcct catctctaac caccaaacag aggccgaccc cgctattatc 720

acactcctcc tggagaaaac aaacccccac gtcgcagaaa acatgattta tgtcgccggt 780

gatcgtgtga ttaccgatcc cctttgcaaa ccctttttcta tgggaagaaa cctcatctgc 840

gtttactcca agaagcacat gtacgatgtc cccgagctgg tcgagatgaa gcgtaaggct 900

aacacgcaat ctcttaagca gatggctctt ctcctgcgag gcggttctaa aatcgtttgg 960

atcgcccctt ctggaggtag agatcgtccc gaccctctca ctgaagaatg gtatcccgct 1020

cctttcgatt cctcgtccgt cgacaatatg cggcggctta tcgagcacag cggcgctcct 1080

ggccacattt acccccttgc actcctgtgt tataacatca tgccccctcc tcctcaagtt 1140

gagaagcaga ttggtgaacg gagaaacatt agcttccacg gtgctggact ctccgttgct 1200

cccaagatta gcttccagga gattgcagct gcttgcgaga aatccgaaga agcaaaagat 1260

gtctactcgc aagcactgta caaaagcgtt acaaaacaat acaacgttct gaaatcggca 1320

gtccatggta aacagggtct ggaggcttcg acagctggcg tttctcttag ccagccctgg 1380

aactaa 1386

<210> 167
<211> 1569
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 167
atggctaaaa ttaaaaaact cttccccctc aaggcactct ctctcctctt cgagactctc 60

cttaaatcct ctggaaaact gccttacctc cgtttcaagg tcagcaatgg ccccgccacg 120

caattcaaat ttctcaagca cagcacgctt gttcaccgta ccgacgaact cgctaaccaa 180

acacttgttt ttcatcttga aggcgctctt ctgaaatcct gttcgctgtt tccctatttt 240

atgctggttg cttttgaagc tggcggcctt ttccgggccc tcatcctgct tcttctgtat 300

cctctggtcc tgctcattgg taaggagctt ggcctgcgga tcctggtgtt cgtctccttc 360

gtcggaatca caaaggaaaa attccgagca ggtacgacta tccttcccaa ttttttttctg 420

gaggacgtgg gctgtgaggg atttgacatt gttatgagct ataaaaaacg agtggcagtc 480

acgggaatgc ctaaattcat gattgaagga tttctcagag attatcttgg cattgatgcc 540

gtggtcgcca gagaacttaa ggagtttcgt ggctactttc ttggcctggt cgaagaaaat 600

atggacgccg gactggtgat ctccgagtgt acgcacaata tcggactcgg atgttttcgg 660

aagtctcacg atcagaaaat tttctcgcat tgtaaggaga tctatctggt cactgaggct 720

gaaaaaaaga attggcaggt cctccctcgg aagcggtatc ctaaacccct gatcttccac 780

gatggaagac tcgcctttcg tcctacaccc tttgctgccc tcattatgtt tatctggctg 840

cccttcggat ttctgcttaa catcactcgg gccattatct ttattagcct ccctttaac 900

ctctctacgc ctctgcttgc tctctccggt ttcattgtga ccgtctctaa gcctgaaccc 960

actactgctt ccattaatgg tgagaataag cctggcggta tgctttacgt gtgcaatcac 1020

cggacgctcc ttgaccccct gttcctttcc ggtgttctga tgaactccgt ttcggctgtg 1080

acatactcga ttagcaagtt ctctgaggtt atttccccta tcaaaactgt tggtctttcg 1140

cgggacagag aaaaagatgg aaaaatcatg aagaaactcc tctcgcaggg tgacctggtc 1200

gtttgtcctg aaggcactac ctgtagagaa ccctaccttc tgcggttctc gcctctcttc 1260

gccgaaatga ctaatgagat tgtgcctgtc gcaatcaagc tccaagtttc tctcttctac 1320

ggttccactg cttcgggact taagtgtctg gattctacct tccacctgat gaatcccaac 1380

cctatgtgtt ccgttaaaat tctgaacaaa cttccctcgt gccagacgca caacacggga 1440

ggcaagagca aattcgaggt gaacaatcat gtccaaaacc agatcgccgc tgcccttgga 1500

ttcgaatgca cgaaccttac gcgtaaagac aaatacgcaa ttcttgctgg aaacgagggt 1560

atcgtgtaa 1569

<210> 168
<211> 1113
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 168
atggctaatc tgcgtgctca cgtttgtttt ctcctgtcgt ttgttacttt cctgaatctg 60

atctcctgcc atactaatct tcctgcaaac cacgttgctc tcttcgtgtt cggagattcc 120

ctgttcgacc ctggtaataa taactacatc aacacaacct ttgattttca ggcaaatttc 180

cgtccttacg gtgagacgtt cttcaagtac cccacaggtc ggttcagcga tggtagactc 240

attcctgact ttcttgcaca gttcgcagga ctccctatta tccccactta ccttcaacct 300

ggcaatcata aattcacgga tggtgtcaac tttgcatctg gcggcgccgg cgccctcgtg 360

gagtctcacc aaggatttgt tgtcgacctt gagactcaaa ttagatactt caaaaaagtt 420

gaaaaaagcc ttcgtcaaga gctcggcgac gaggaggcca aaaagcttct cagccgtgca 480

gtctatctca tctcggttgg aggtaatgat tacctcacac ggaactcgtc tgcatccgat 540

gaggagtatg cctcgatggt tcttggtaat cttactgttg cacttaagga gatctacaaa 600

aagggaggaa gaaaatttgg tttccccaat atgatgcctc ttggttgtct tccctatatg 660

aaagcaaaag ccggcggacc ttgtattgac gagtttacag ccattgctaa gctccacaat 720

aaagaactcc ccaaaacgct ccagaaactt gagactcaac ttgagggctt taaatacgca 780

tactataatt tttacaaatc cgttagcgaa cggctcaata atccctctaa gtatggcttc 840

aaggatgcca ccaccgcatg ctgtggaagc ggactttacg gcggtgtcta ttcgtgtggt 900

ggaaaacgag gaatcacgga gtttcacctt tgtgagaacc cttcggaata cctcttttc 960

gactcttatc accctagcga aaaagcttac cagcaattcg cagagcttat gtggagcggc 1020

acaacggact tcgtctggcc ctataatctt aaaactcttt tcgaggctaa aaccagcaac 1080

ctgtatgacc acttcatcat tccccccccct taa 1113

<210> 169
<211> 1197
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 169
atgcttacgt accttcctct gagatgtctg ggtcttaaat ctctcgtttg catctctttt 60

gattgtggag ttattggtgc tatggctgct cgagtcaaat attcgtggtc cctggccctc 120

gtgctcctgt gtattgtgag ccttagctat tttggaaacg ccgctgacct gggagcttct 180

ttcgtttttg gcgactctct ggttgatgcc ggaaacaata actacctccc tacgctgtcc 240

aaagctaata ttccccccaa cggtatcgac tttaaggcaa gcggcggaca acctacaggt 300

cgatatacca atggcagaac catcggagac atcgtcggag aggaactcgg cgtccccaac 360

tacgctgttc ccttcctttc tcctaattcc acgggaaagg ctatcctcta cggagtgaat 420

tatgcatcgg gaggaggtgg catcatgaat gcaactggaa gaatctttgt caatagactc 480

ggcctggaca ttcagattga tttctacaac actactagaa aacagttcga taagctgctg 540

ggtcccagca aggccaaaga ctacattaga aagtcttcca tctttagcat tacgattggt 600

gcaaatgatt ttctgaacaa ttacctgctc cctgtccttt ctattggcgc ccggattagc 660

gagacacctg atggatttat cgacgatatg atcaaccatc ttagagatca gcttacccgt 720

ctctatcaac tggatgctcg gaaattcgtg attggcaacg ttggacctat cggatgcatc 780

ccctatcaga aaaccatcaa tcaactgaat gaaaacgagt gtgtcgacct tgcaaacaag 840

ctggccaaac agtataacag ccgtcttcag aatctcctta cggaacttaa tggcaaactg 900

cctggagcaa catttgtcgt tgctaacgtg tacgaccttg ttatggaact gattaccaat 960

tatgaaaaat atggcttcat cacagcctct aaagcttgtt gtggtaacgg tggccaattc 1020

gccggtatca tcccctgcgg acccacatcg acgatgtgca aagatcgaga caagcacgtg 1080

ttttgggacc cctatcatcc tagcgaggca gctaatctga ttattgccag acaactcctt 1140

tatggtagca ctaaatatat cagccccgtc aacctgaaag agctccgaaa cctgtaa 1197

<210> 170
<211> 1053
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 170
atgagagtta gctcttcgtt cctggcttcg ctcgtgcttg ttctcgtttt ctctgcagcc 60

aacggcgacc ctcttgtccc tgcccttatt tcctttggtg attccgtcgt cgatgtgggc 120

aacaacaata acctcaacac gctcatcaaa gccaatttcc ctccctatgg ccgtgacttc 180

gttacccacc ggcctacagg acgtttttgt aatggtaaac ttgccacgga ctttactgca 240

gagtacctgg gctttacctc ttaccctccc gcataccttt cccaggatgc acgaggtaat 300

gaccttctta caggagcaag cttcgcttcc gctgcctctg gcctttatga tagaaccgct 360

aacctttgcc gagctattcc ccttacacga caagtcaact attacaaaga gtatcagaca 420

aaggtcgtga atacggttgg aaaagccaag gcaaatgata tctttgcagg agcaattcac 480

ctcctgagcg ctggctctag cgattatatc caaaactact acatcaaccc catgctgaat 540

cgagtttata ctcccgacca gttctccgat attctgattc gatcctactc taccttcatt 600

caaaaccttt acggactcgg tgctcgtcgg atcggcgtta ccacccttcc ccccactgga 660

tgccttcccg ccgctatcac gcttttcggc ggaggctcta accagtgtgt cactcggctc 720

aatcaagacg cagtcgcttt caataagaag cttaataata cttctgagtc gctcctcgac 780

aaaatgcccg gacttaagct tctcgttttc gatatctatc aacccctcct cgatatggtt 840

atgaagcctg cagacaatgg tttcttcgag tctcggcgtg cctgttgcgg tactggaacg 900

cttgaaacat cccttctttg taatgaccgt tctcttggca catgcactaa tgctacaagc 960

tatgttttct gggatggatt ccatcctagc gaggccgcaa accaggtcct cgcaggaaat 1020

ctgctcgcac agggagtgtc gctcatttcg taa 1053

<210> 171
<211> 1167
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 171
atggccaaac tgatctctta cattaacttc tttaaagtgt cggttatcct tctgctttct 60

acaacatcgc gatgctttgg tactcccctg gagaaaggtt ccgctctttt catcttcggc 120

gatagcacgg tcgaccctgg caataataat tacatcaaca ctatccctga gaaccgtgct 180

cattacaaac cttatggaca gaacggcttt tttgccgagc ctactggacg tttttccgac 240

ggacgtgtga ttgttgacta tgtcgcagaa tatgcaaagc tccccccttat ccctcccttc 300

ctccaaccct ccgctgactt cacaagcggt gcaaatttcg ctagcggagg tggtggcgtt 360

ctccccgaaa caaaccaggg actcgtcatt gatcttccca cgcagctgaa gaattttgaa 420

gaggttcaga agagccttgt ggaaaaactc ggcgaagcac agacgaagga acttatttcg 480

gaggctgtgt actttatttc catcggttcg aacgactaca tgggtggtta tctcggaaac 540

cccaaaatgc aagagttcta ccctcctgaa gcctatgttg gtatggttat cggaaacctt 600

acacagactc tgcaagtcct gtacgagaag ggcgctcgta aattcggctt tcttagcctt 660

tcgcccctcg gatgcctccc cgccctgaga gccctgaact tcaaggacaa tgaaggaaag 720

tgtcttgaag ctgcaagcgc cctggctcga gcccataatg atgccctcaa tgctgttctg 780

acgaacctgg aacacatcct ccctggcctt aaatattgta attcgaattt ctatgaatgg 840

ctcgacgatc gtatgaacaa ccccagcaaa tatggattta aggacggagt gaacgcttgt 900

tgcggcacgg gtccttacgg aggcgtcttc acctgtggag gcacaaagaa tgtgaaggat 960

tatcagcttt gcgacaatgc cgacgatcat gtgtggtggg actctttcca ccctactgaa 1020

cggatccacg aacaatttgc caaagcactt tggtccggac atccttcttc tgttggtcct 1080

tacaacctcg aggatctttt cttcaataag gagaaactga caatcgcaga tatcgtcgac 1140

tctcctgaaa cagatccctt ccaataa 1167

<210> 172
<211> 1083
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimised ORF

<400> 172
atgggaaaac tctatacaat tattgttgtc attctgatct tttcgtccac gctcgcttct 60

agccgtaacg ttcccgctat ctttattttt ggcgacagca tttttgacgc cggtaacaac 120

cacttcaata aaaattgctc cgtgcaggcc gacttctcgc cctacggatc tacgttcttc 180

catcagccta cgggccggtt cacaaatgga cgaactgtgg ccgattttat ctctcaattt 240

atcggtatcg atatgcaaaa accttacctg gaagcccaga ttgcagtcat caatggatct 300

cgtaaggact atcccagcaa tggtatcaat tttgcctccg ccggaagcgg agttctgcag 360

ggaacaaaca tggagcaggg agttatcagc atccaagaac agctgcgtca gtttcaaacg 420

ctcgtcgaac aaaaccaaat tgacaagaat ctggttcaga aatcgctgtt ctttctcgaa 480

agcggatcca acgacatctt caattatttt atccccttg accctcctac cctggaccct 540

gacgcatacg tgcaagctat gatgaaagag gtggcaactc ttatcgatca gatctataag 600

ctgggagcca gaagaattgc agtcttcagc cttggacccg ttggatgcgt gcctgctcgt 660

gcccttctgc ttggcgctcc tatcaaccga tgttatggaa agatgaacgt gatggttaaa 720

aaatacaata aaggactcga atcgcttatt aacaacattc ctatcagata tcccggcgca 780

attggtgttt atggcgcaat ctacgacatt gttcagcgtt ttcgtgcaat tcccgctcgg 840

tatggcttct ctgacgtcac cagcgcctgc tgcggaaatg gccctctgag aggaaaactc 900

cagtgcggta tcgagggata ccagatttgc ctgaatccca atgaagttct cttctgggac 960

tatttccacc ctacagaaca tgcatataaa ctcatctctg aagccctttg gggaggcaac 1020

aaatcccaga ttcgtcccgt taatctgaag aatctggccg atatcacctt cacgcttgtg 1080

taa 1083

<210> 173
<211> 8397
<212> DNA
<213> Artificial Sequence

<220>
<223> DNA construct

<400> 173
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagactgtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240

attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat 300

tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt 360

tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt gacgcgtatt gggatgcggc 420

cgctgtcggg aaccgcgttc aggtggaaca ggaccacctc ccttgcactt cttggtatat 480

cagtataggc tgatgtattc atagtggggt ttttcataat aaatttacta acggcaggca 540

acattcactc ggcttaaacg caaaacggac cgtcttgata tcttctgacg cattgaccac 600

cgagaaatag tgttagttac cgggtgagtt attgttcttc tacacaggcg acgcccatcg 660

tctagagttg atgtactaac tcagatttca ctacctaccc tatccctggt acgcacaaag 720

cactttgcta gaaggtggtc tccgctagtg ctacccatgg ggtgtctcca gaacccttgt 780

acatcagcct atatgcgtgt gtgtctacgt ctactgtaca aatctgcttg gcgtcaaatg 840

gcactacaag tacaaagacc aaatgccaca agttcagtac gagtacaaga gtctgctact 900

tgtagctccc gcaggtattt agttcataca aacgagtata cttgtagctg tccaacagta 960

gttttcagac acctgatcag aaataggaac acgctccacg gacctcgtat ccactctaag 1020

accataatat atgggtttct ggaacattgg tgcccctcac ctctcatctc ttttttttt 1080

acctctctga ccaccaccac ccttgttgtt ccgagtggtt agctgcaccc cacctagcaa 1140

actttagggc cccatacgca taagcatacg tccggtgtgc gacttgtgtg tgaacgcggt 1200

ccacacaagg tgtctttcaa gtccacacag ggctttgtct agagcctggg ctatcttccc 1260

acgtttgtat atatagactg gatccacacc cagcaacgca acgtcctccc acccaaccct 1320

atcctccacg ttacacgccc aagatcttcc aaacaagaag acacattcca cagcaagaat 1380

aaccttggca agaacgtgcg acgaggtgaa taaaagagcc cagagtacta acattttgtg 1440

gacctgttgg aaactaattg ccgccgcaac tcgactggaa ccaagttcaa gaccactgac 1500

aagcgccccg gtccaaccag aacactaaac ggagccgcac agcaagcacc acggacgatc 1560

gtcatcatcc ctatggccaa cctgcatgga tactcttgat aacagttaag ctatatttac 1620

caacaccacc tccgttcacg acataacatc cccacaccat agctacaccc aaagtcgctg 1680

tacccgtacc gcccgcaaaa tggaagtccg acgacgaaaa atcgacgtgc tcaaggccca 1740

gaggcccccc cggcccccaa atttaaaggg ggccttttg gccataactt cgtataatgt 1800

acattatacg aagttatgaa ttctaggtta gggataacag ggtaatgata acttcgtata 1860

gcatacatta tacgaagtta tgggttggcg gcgcatttgt gtcccaaaaa acagccccaa 1920

ttgccccaat tgaccccaaa ttgacccagt agcggaccca accccggcga gagccccctt 1980

caccccacat atcaaacctc ccccggttcc cacacttgcc gttaagggcg tagggtactg 2040

cagtctggaa tctacgcttg ttcagacttt gtactagttt ctttgtctgg ccatccgggt 2100

aacccatgcc ggacgcaaaa tagactactg aaaatttttt tgctttgtgg ttgggacttt 2160

agccaagggt ataaaagacc accgtccccg aattaccttt cctcttcttt tctctctctc 2220

cttgtcaact cacacccgaa gcacaatgaa gaagcccgag ctgaccgcta cctctgttga 2280

gaagttcctg attgagaagt ttgattccgt ttccgacctg atgcagctgt ccgagggcga 2340

ggagtctcga gccttctcct ttgacgtggg cggacgaggt tacgttctgc gagtgaactc 2400

gtgtgccgac ggcttctaca aggatcgata cgtctaccga cactttgctt ctgccgctct 2460

gcccatccct gaggttctcg acattggcga gttctctgag tccctcacct actgcatctc 2520

tcgacgagct cagggagtca ccctgcagga cctccctgag actgagctgc ctgctgtcct 2580

ccagcctgtt gctgaggcca tggacgctat cgctgctgct gatctgtccc agacctcggg 2640

tttcggcccc tttggacctc agggaattgg acagtacacc acttggcgag acttcatctg 2700

tgctattgcc gatcctcacg tctaccattg gcagaccgtt atggacgata ctgtgtcggc 2760

ttctgtcgct caggctctgg acgagctgat gctctgggcc gaggattgcc ccgaggttcg 2820

acacctggtg catgctgact tcggttccaa caacgttctc accgacaacg gccgaatcac 2880

tgccgtgatt gactggtccg aggctatgtt tggcgactcg cagtacgagg tggccaacat 2940

cttcttttgg cgaccctggc tggcttgtat ggagcagcag acccgatact tcgagcgacg 3000

acatcctgag ctcgctggct cccctcgact gcgagcttac atgctccgaa ttggtctgga 3060

ccagctctac cagtcgctgg tggatggcaa ctttgacgat gctgcctggg ctcagggacg 3120

atgtgacgcc atcgtgcgat ctggcgctgg aaccgtcgga cgaactcaga ttgcccgacg 3180

atccgctgct gtctggaccg acggatgcgt ggaggtcctg gctgattcgg gtaaccgacg 3240

accctctact cgacctcgag ctaaggagta agtgtctgtg gtatctaagc tatttatcac 3300

tctttacaac ttctacctca actatctact ttaataaatg aatatcgttt attctctatg 3360

attactgtat atgcgttcct ctaagacaaa tcgaatataa cttcgtatag caggagttat 3420

ccgaagcgag tagggataac agggtaatac ggagaattcg ggttggcggc gtatttgtgt 3480

cccaaaaaac agccccaatt gccccaattg accccaaatt gacccagtag cgggcccaac 3540

cccggcgaga gccccttca ccccacatat caaacctccc ccggttccca cacttgccgt 3600

taagggcgta gggtactgca gtctggaatc tacgcttgtt cagactttgt actagtttct 3660

ttgtctggcc atccgggtaa cccatgccgg acgcaaaata gactactgaa aatttttttg 3720

ctttgtggtt gggactttag ccaagggtat aaaagaccac cgtccccgaa ttacctttcc 3780

tcttcttttc tctctctcct tgtcaactca cacccgaaat cgttaagcat ttccttctga 3840

gtataagaat cattcaaaat ggtcagctcg ggtgaggaca ttttcagcgg actcgtcccc 3900

attctgattg aactcgaggg tgacgttaat ggacaccgtt tctctgtccg aggtgaaggt 3960

tacggcgacg cttctaacgg caaactggaa atcaagttca tctgcacgac aggacggctt 4020

cccgtgccct ggcctactct ggtcacaacc ctctcctatg gagtgcaatg ttttgctaag 4080

taccctgagc acatgcggca aaatgatttt tttaagagcg ccatgcctga tggatacgtc 4140

caagaacgaa ccatctcctt taaagaagac ggaacgtata aaacacgagc cgaagtcaaa 4200

tttgaaggcg aagctctcgt caatcgaatt gatctgaaag gcctggagtt taaggaggac 4260

ggcaacatcc tcggacataa actcgagtac agctttaatt cccattacgt ttacattaca 4320

gcagataaga atagaaatgg actggaggct cagtttcgga ttcgtcataa cgtggacgat 4380

ggaagcgtcc aacttgcaga ccattatcaa cagaatacac ctatcggcga aggccccgtc 4440

cttctccccg aacagcatta cctgactacc aattccgtcc ttagcaaaga tcctcaggaa 4500

agacgagatc acatggtgct ggttgagttt gttaccgctg ccggactcag cctgggaatg 4560

gatgagctct ataaatccta agtgtctgtg gtatctaagc tatttatcac tctttacaac 4620

ttctacctca actatctact ttaataaatg aatatcgttt attctctatg attactgtat 4680

atgcgttcct ctaagacaat aacttcgtat aatgtacatt atacgaagtt ataacccact 4740

tgtgcattcc tttattatct ggcttacaac tacaagcaga accagtagca taacactcat 4800

cagtagcctt tacagtgata aataatattg acattgacgt aatctgatga ttgtagacaa 4860

acgtgagacg gcattagcga taatattttg acgatgccca cgggcatata ataaccattc 4920

tgtaatagaa ctttgagcat gttcagaagt aatgagtttc atacatagga gtacagatcg 4980

tggcgatgtg taattaccgg taccgcatat gttgaacctg ttcgtaaaag gtgccttcaa 5040

cacagttttg ttgccgtctt acctgtgtat aagcgttcag cttctggtgt gtagcagaag 5100

gaggtgaagt tgaggaacaa ataatggaga cttgcagcag actccgtgcc tcaagcaaaa 5160

aagatgatcc agcagaggat cgaactctga gcctcctcgg cactcgtaga ttgaagccac 5220

tttgtgagcg agacgtgata accattacac cactagaccg ggggtgaaga tcagagcccg 5280

tattgtcgta ttcaaagtta aaaatgggga ttttttatat tcagaagtgt gtgatttatt 5340

gaacgccaat aacccggctt actacagtag gtatatagct gtaatgaatc attagaataa 5400

aaaaaaatcc gcctcaaagt agctgaaatg tatccctaga ataatgaggg gaaaaaactg 5460

ataaatcagg tgaaatacgg catttaggga caggtaagag ttcaattgaa tttatgtcgt 5520

attatttaag ctgtcggaga aaacttgtgt gctttttgt atggttttcg tgactgtgta 5580

catttctgta cccctgattc cgaacatctt atactgacac aagacattgg agataaatca 5640

tacagtataa agtagctacg tcttacttgt acagtgcttt taattacaat atgccccact 5700

atacatacat gtgattcatt ttataaaccc ctgagaccca tgtgtaacac tcgctctgga 5760

gagttagtca tccgacaggg taactctaat ctcccaacac cttattaact ctgcgtaact 5820

gtaactcttc ttgccacgtc gatcttactc aattttcctg ctcatcatct gctggattgt 5880

tgtctatcgt ctggctctaa tacatttatt gtttattgcc caaacaactt tcattgcacg 5940

taagtgaatt gttttataac agcgttcgcc aattgctgcg ccatcgtcgt ccggctgtcc 6000

taccgttagg ggtagtgtgt ctcacactac cgaggttact agagttggga aagcgatact 6060

gcctcggaca caccacctgg gtcttacgac tgcagagaga atcggcgtta cctctctcac 6120

aaagcccttc agtgcggccg catcccaatg gcgcgccgag cttggcgtaa tcatggtcat 6180

agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacaacata cgagccggaa 6240

gcataaagtg taaagcctgg ggtgcctaat gagtgagcta actcacatta attgcgttgc 6300

gctcactgcc cgctttccag tcgggaaacc tgtcgtgcca gctgcattaa tgaatcggcc 6360

aacgcgcggg gagaggcggt ttgcgtattg ggcgctgttc cgcttcctcg ctcactgact 6420

cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc tcactcaaag gcggtaatac 6480

ggttatccac agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa 6540

aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgcccccctg 6600

acgagcatca caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa 6660

gataccaggc gtttcccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc 6720

ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac 6780

gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac 6840

ccccgttca gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag tccaacccgg 6900

taagacacga cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt 6960

atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa 7020

cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct 7080

cttgatccgg caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga 7140

ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg gggtctgacg 7200

ctcagtggaa cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct 7260

tcacctagat ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt 7320

aaacttggtc tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc 7380

tatttcgttc atccatagtt gcctgactcc ccgtcgtgta gataactacg atacgggagg 7440

gcttaccatc tggccccagt gctgcaatga taccgcgaga accacgctca ccggctccag 7500

atttatcagc aataaaccag ccagccggaa gggccgagcg cagaagtggt cctgcaactt 7560

tatccgcctc catccagtct attaattgtt gccgggaagc tagagtaagt agttcgccag 7620

ttaatagttt gcgcaacgtt gttgccattg ctacaggcat cgtggtgtca cgctcgtcgt 7680

ttggtatggc ttcattcagc tccggttccc aacgatcaag gcgagttaca tgatccccca 7740

tgttgtgcaa aaaagcggtt agctccttcg gtcctccgat cgttgtcaga agtaagttgg 7800

ccgcagtgtt atcactcatg gttatggcag cactgcataa ttctcttact gtcatgccat 7860

ccgtaagatg cttttctgtg actggtgagt actcaaccaa gtcattctga gaatagtgta 7920

tgcggcgacc gagttgctct tgcccggcgt caatacggga taataccgcg ccacatagca 7980

gaactttaaa agtgctcatc attggaaaac gttcttcggg gcgaaaactc tcaaggatct 8040

taccgctgtt gagatccagt tcgatgtaac ccactcgtgc acccaactga tcttcagcat 8100

cttttacttt caccagcgtt tctgggtgag caaaaacagg aaggcaaaat gccgcaaaaa 8160

agggaataag ggcgacacgg aaatgttgaa tactcatact cttccttttt caatattatt 8220

gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa 8280

ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac gtctaagaaa 8340

ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc ttttgtc 8397

<210> 174
<211> 7984
<212> DNA
<213> Artificial Sequence

<220>
<223> DNA construct

<400> 174
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagactgtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240

attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat 300

tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt 360

tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt gacgcgtatt gggatgcggc 420

cgctgtcggg aaccgcgttc aggtggaaca ggaccacctc ccttgcactt cttggtatat 480

cagtataggc tgatgtattc atagtggggt ttttcataat aaatttacta acggcaggca 540

acattcactc ggcttaaacg caaaacggac cgtcttgata tcttctgacg cattgaccac 600

cgagaaatag tgttagttac cgggtgagtt attgttcttc tacacaggcg acgcccatcg 660

tctagagttg atgtactaac tcagatttca ctacctaccc tatccctggt acgcacaaag 720

cactttgcta gaaggtggtc tctgtaacct cggcaggagg agatcggcaa agctgaagat 780

gattgatgca tgtggcgttt ttgaaggaac taccggcagg gtttgttggt gttaatgtga 840

acacagcttt ttttttgctt aataagatct ccaaacacgt tcaaaaaaca cccactagtt 900

aaaagctcag ttggactacc tattgttcct tctaatatac agtagactgc tcgtacagta 960

gtctctctaa tcgatcggtg tctcaatgtc tcaccgttac cggatacgct tatggtcgat 1020

tgaattgcca atggaattat cggtggaaat tcatggaaaa ccggtttcgg cgagctcgaa 1080

atacaccccg ctatcgtaca agtcgaacga atatcaggcg cccagacaga tcgattcaag 1140

tagggagaag agtggaacaa cttgtaaaga atgactgatg gtgacgctcg ggcccaggct 1200

agcactgacc atcaaatcca tcttatcgat ctgcttgaat ctacaagtag gagggttgga 1260

gtgattaagt gaaacttctt taacggctct atgccagttc tattgatatc cgaaacatca 1320

gtatgaaggt ctgataaggg tgacttcttc ccacagattc gtatcagtac gagtacgagt 1380

accggtactt gtaacagtat tgatactaaa gggaaactac aacggttgtc agcgtaatgt 1440

gacttcgccc atgaacgcag acacgcagtg ccgagtgcgg tgatatcgcc tactcgttac 1500

gtccatggac tacacaaccc ctcggcttcg cttggcttag cctcgggctc ggtgctgttc 1560

agttaaaaca caatcaaata acatttctag ctctactttt tagaaggcag gccgtcagga 1620

gcaactccga ctccattgac gtttctaaac atctgaatgc cttccttacc ttcaacaaac 1680

tggcaggttc gggcgacagt gtaaagagac ttgatgaagt tggtgtcgtc gtgtcggtag 1740

tgggcccccc cggcccccaa atttaaaggg ggccttttg gccataactt cgtataatgt 1800

acattatacg aagttattct gaattccgag aaacacaaca acatgcccca ttggacagat 1860

catgcggata cacaggttgt gcagtatcat acatactcga tcagacaggt cgtctgacca 1920

tcatacaagc tgaacaagcg ctccatactt gcacgctctc tatatacaca gttaaattac 1980

atatccatag tctaacctct aacagttaat cttctggtaa gcctcccagc cagccttctg 2040

gtatcgcttg gcctcctcga taggatctcg gttctggccg tacagacctc ggccgacaat 2100

tatgatatcc gttccggtag acatgacatc ctcaacagtt cggtactgct gtccgagagc 2160

gtctcccttg tcgtcaagac ccaccccggg ggtcagaata agccagtcct cagagtcgcc 2220

cttaggtcgg ttctgggcaa cgaagccaac cacaaactcg gggtcggatc gggcaagctc 2280

aatggtctgc ttggagtact cgccagtggc cagagagccc ttgcaagaca gctcggccag 2340

catgagcaga cctctggcca gcttctcgtt gggagagggg accaggaact ccttgtactg 2400

ggagttctcg tagtcagaga cgtcctcctt cttctgttca gagacagttt cctcggcacc 2460

agctcgcagg ccagcaatga ttccggttcc gggtacaccg tgggcgttgg tgatatcgga 2520

ccactcggcg attcggtaga caccgttctt gtactggtgc ttgacagtgt tgccaatatc 2580

tgcgaacttt ctgtcctcga acaggaagaa accgtgctta agagcaagtt ccttgagggg 2640

gagcacagtg ctggcgtagg tgaagtcgtc aatgatgtcg atatgggtct tgatcatgca 2700

cacataaggt ccgaccttat cggcaagctc aatgagctcc ttggtggtgg taacatccag 2760

agaagcacac aggttggttt tcttggctgc cacgagcttg agcactcgag cggcaaaggc 2820

ggacttgtgg acgttagctc gagcttcgta ggagggcatt ttggtggtga agaggagact 2880

gaaataaatt tagtctgcag aacttttat cggaacctta tctggggcag tgaagtatat 2940

gttatggtaa tagttacgag ttagttgaac ttatagatag actggactat acggctatcg 3000

gtccaaatta gaaagaacgt caatggctct ctgggcgcga attcgtgggt tggcggcgta 3060

tttgtgtccc aaaaaacagc cccaattgcc ccaattgacc ccaaattgac ccagtagcgg 3120

gcccaacccc ggcgagagcc cccttcaccc cacatatcaa acctcccccg gttcccacac 3180

ttgccgttaa gggcgtaggg tactgcagtc tggaatctac gcttgttcag actttgtact 3240

agtttctttg tctggccatc cgggtaaccc atgccggacg caaaatagac tactgaaaat 3300

ttttttgctt tgtggttggg actttagcca agggtataaa agaccaccgt ccccgaatta 3360

cctttcctct tcttttctct ctctccttgt caactcacac ccgaaatcgt taagcatttc 3420

cttctgagta taagaatcat tcaaaatggt cagctcgggt gaggacattt tcagcggact 3480

cgtccccatt ctgattgaac tcgagggtga cgttaatgga caccgtttct ctgtccgagg 3540

tgaaggttac ggcgacgctt ctaacggcaa actggaaatc aagttcatct gcacgacagg 3600

acggcttccc gtgccctggc ctactctggt cacaaccctc tcctatggag tgcaatgttt 3660

tgctaagtac cctgagcaca tgcggcaaaa tgatttttt aagagcgcca tgcctgatgg 3720

atacgtccaa gaacgaacca tctcctttaa agaagacgga acgtataaaa cacgagccga 3780

agtcaaattt gaaggcgaag ctctcgtcaa tcgaattgat ctgaaaggcc tggagtttaa 3840

ggaggacggc aacatcctcg gacataaact cgagtacagc tttaattccc attacgttta 3900

cattacagca gataagaata gaaatggact ggaggctcag tttcggattc gtcataacgt 3960

ggacgatgga agcgtccaac ttgcagacca ttatcaacag aatacaccta tcggcgaagg 4020

ccccgtcctt ctccccgaac agcattacct gactaccaat tccgtcctta gcaaagatcc 4080

tcaggaaaga cgagatcaca tggtgctggt tgagtttgtt accgctgccg gactcagcct 4140

gggaatggat gagctctata aatcctaagt gtctgtggta tctaagctat ttatcactct 4200

ttacaacttc tacctcaact atctacttta ataaatgaat atcgtttatt ctctatgatt 4260

actgtatatg cgttcctcta agacaataac ttcgtataat gtacattata cgaagttatc 4320

tgcacggcca ccttgccggt gccggtgttg gtctgcgtgg tcgaatccat ttctatggtc 4380

ggtctgttgg aatacgagaa agggactcag gatgccgacc ggtatatatt tcaatgtctt 4440

acctcgagtt taaacccact aggctttttt ttttctccct gtcaatccac caacctttt 4500

ggcctggcgt gactgacttg gggacaagca aaagctgagg gtaaacagag gcacacccga 4560

ggttatatgt tgagggccaa tgtcaaattt tcctcctaat cctgagagcc cctctaacac 4620

cagcacgcat ctacctcttg tatataacta cgggccccgc ctctcctaac cgccacttta 4680

acacgctcgt actcctcgtc tgttaccacc tgaaaaacgc ataatgcgca tgtgcggatt 4740

ggccaagaga gctccgaacg gaaccgcccg gggtatactg ctgctctcag cgtggatagt 4800

gtggagggac cacgtctcaa ccgtgacatg ggtcgactta tggcgggggg gagagtgaaa 4860

gaaccagtta cacgtgcagg tgaagagggg aagaccttgg acgtgttcac gtgactacgt 4920

cggttctaca tgaatccaac gtctcttatt tcgtcatgtg ctactacatg acccccattg 4980

gcacgactct ctctctgtgc tgccactact tgacgactat gtccgcttgt tgactcgttg 5040

aacaggtgtc cgatgaatag catggcttct caacaggcca tgcttaaagt tagagttggg 5100

agatgtcgtt ccgaaccttg tgggcggcgg tgtggaaaga tagagagtac aagacaacca 5160

aactcttgga gcctttggct cgtaaagaga gctcttccgg tccatcgacg aagaaacgca 5220

gcgttgaaag actgccagaa gcggcagtca ccgccgtatc tgagacattg atcgtacaca 5280

ctcttgcaca agggctatgc ttaactactt gtacgggtag agacccatgt gtaacactcg 5340

ctctggagag ttagtcatcc gacagggtaa ctctaatctc ccaacacctt attaactctg 5400

cgtaactgta actcttcttg ccacgtcgat cttactcaat tttcctgctc atcatctgct 5460

ggattgttgt ctatcgtctg gctctaatac atttattgtt tattgcccaa acaactttca 5520

ttgcacgtaa gtgaattgtt ttataacagc gttcgccaat tgctgcgcca tcgtcgtccg 5580

gctgtcctac cgttaggggt agtgtgtctc acactaccga ggttactaga gttgggaaag 5640

cgatactgcc tcggacacac cacctgggtc ttacgactgc agagagaatc ggcgttacct 5700

ctctcacaaa gcccttcagt gcggccgcat cccaatggcg cgccgagctt ggcgtaatca 5760

tggtcatagc tgtttcctgt gtgaaattgt tatccgctca caattccaca caacatacga 5820

gccggaagca taaagtgtaa agcctggggt gcctaatgag tgagctaact cacattaatt 5880

gcgttgcgct cactgcccgc tttccagtcg ggaaacctgt cgtgccagct gcattaatga 5940

atcggccaac gcgcggggag aggcggtttg cgtattgggc gctgttccgc ttcctcgctc 6000

actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg 6060

gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc 6120

cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc 6180

ccccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga 6240

ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc tgttccgacc 6300

ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat 6360

agctcacgct gtaggtatct cagttcggtg taggtcgttc gctccaagct gggctgtgtg 6420

cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc 6480

aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga 6540

gcgaggtatg taggcggtgc tacagagttc ttgaagtggt ggcctaacta cggctacact 6600

agaagaacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt 6660

ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag 6720

cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg 6780

tctgacgctc agtggaacga aaactcacgt taagggattt tggtcatgag attatcaaaa 6840

aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata 6900

tatgagtaaa cttggtctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg 6960

atctgtctat ttcgttcatc catagttgcc tgactccccg tcgtgtagat aactacgata 7020

cgggagggct taccatctgg ccccagtgct gcaatgatac cgcgagaacc acgctcaccg 7080

gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag aagtggtcct 7140

gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaagctag agtaagtagt 7200

tcgccagtta atagtttgcg caacgttgtt gccattgcta caggcatcgt ggtgtcacgc 7260

tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga 7320

tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt tgtcagaagt 7380

aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc tcttactgtc 7440

atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa 7500

tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa taccgcgcca 7560

catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg aaaactctca 7620

aggatcttac cgctgttgag atccagttcg atgtaaccca ctcgtgcacc caactgatct 7680

tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc 7740

gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactctt cctttttcaa 7800

tattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt 7860

tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc 7920

taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac gaggcccttt 7980

tgtc 7984

<210> 175
<211> 7942
<212> DNA
<213> Artificial Sequence

<220>
<223> DNA construct

<400> 175

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagactgtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240

attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat 300

tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt 360

tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt gacgcgtatt gggatgcggc 420

cgctgtcggg aaccgcgttc aggtggaaca ggaccacctc ccttgcactt cttggtatat 480

cagtataggc tgatgtattc atagtggggt ttttcataat aaatttacta acggcaggca 540

acattcactc ggcttaaacg caaaacggac cgtcttgata tcttctgacg cattgaccac 600

cgagaaatag tgttagttac cgggtgagtt attgttcttc tacacaggcg acgcccatcg 660

tctagagttg atgtactaac tcagatttca ctacctaccc tatccctggt acgcacaaag 720

cactttgcta gaaggtggtc tccagcctca aaagcacaga tttaccccgt ttattgcccc 780

caaatcgccc cactagcccc caacatcaac gtctcaaaaa tacgagcgtc ctgtccttgt 840

caaagagtgt ctcacaggct ttattgtacg ccctacacct gcactttgca tggtgctacg 900

cttaaaatga aaccctggaa tgtgaagcga gatgaggtat tttaacgggt tggaaaaaga 960

agcaaatggg gttgcttttt ggaggcggat gatagccaac agaaccgaaa tggccaaaga 1020

agatctcctc caatgggtct ttgagcggct cggggtcgct tagaagaccg caaaacactc 1080

ttttcagccg ctccaagcag tcttgtgaat cttggcgtca ccatgggttg gcattggtgg 1140

cggagtaaaa tatattttgt tgaggaggtg acaggagtgg ttgtatattg gaagcaattc 1200

gtttttgtc gaacattttc aggagtagat ttcagtgatt tccctgtttc ccctcttttc 1260

tcctgaattt cccgatttat tccaaacaat tgaaatatat tctttatcag aacgcgcaga 1320

tctggagaaa ggcgcatatt gtgggggaaa cattcttcat tttgcgggaa aacacgcgtt 1380

gatgacgttt gttgctcgtt cgttcgtcgt tcgtcgttcg tcgttcgtcg ttctacgtgc 1440

tcattaaaca accacatgga atcatgagaa atcgaacgcc cttctccatc agtgtgatat 1500

acaaaacata ttactctaca aaacggactc tctaattgct gccgagatca atcttccaca 1560

cccactcatc aatgttggag atgactctct cttcaatctc atttcccttt ccactagcca 1620

ccttcagaac catctcgttc aactcagaac gccccagaat gtccacatgc tcggcagtct 1680

gagcaccgcc tcgaatatca agtcgatcag gctggtgcaa catctccaca accttgacct 1740

ggggccccc cggcccccaa atttaaaggg ggccttttg gccataactt cgtataatgt 1800

acattatacg aagttatgaa ttctaggtta gggataacag ggtaatgata acttcgtata 1860

gcatacatta tacgaagtta tggggttggc ggcgcatttg tgtcccaaaa aacagcccca 1920

attgccccaa ttgaccccaa attgacccag tagcggaccc aaccccggcg agagcccct 1980

tcaccccaca tatcaaacct cccccggttc ccacacttgc cgttaagggc gtagggtact 2040

gcagtctgga atctacgctt gttcagactt tgtactagtt tctttgtctg gccatccggg 2100

taacccatgc cggacgcaaa atagactact gaaaattttt ttgctttgtg gttgggactt 2160

tagccaaggg tataaaagac caccgtcccc gaattacctt tcctcttctt ttctctctct 2220

ccttgtcaac tcacacccga agcacaatga ccactctgga tgacaccgct taccgatacc 2280

gaacttccgt tcctggcgat gccgaggcta ttgaggctct ggatggatct ttcaccactg 2340

acaccgtttt ccgagtgacc gctactggcg acggcttcac cctgcgagag gtgcctgtcg 2400

accctcctct caccaaggtt ttccctgacg atgagtcgga cgatgagtct gacgctggag 2460

aggacggcga ccctgactct cgaactttcg tggcttacgg cgacgatggc gacctggccg 2520

gctttgtggt cgtttcttac tccggatgga accgacgact gaccgtggag gacatcgagg 2580

tcgctcctga gcaccgaggt catggtgtcg gacgagctct gatgggtctg gctactgagt 2640

tcgctcgaga gcgaggtgct ggccacctgt ggctcgaggt caccaacgtt aacgcccctg 2700

ctattcatgc ctaccgacga atgggtttta ccctgtgtgg cctcgatact gccctgtacg 2760

acggaaccgc ttccgatgga gagcaggccc tctacatgtc gatgccctgc ccttaagtgt 2820

ctgtggtatc taagctattt atcactcttt acaacttcta cctcaactat ctactttaat 2880

aaatgaatat cgtttattct ctatgattac tgtatatgcg ttcctctaag acaaatcgaa 2940

tataacttcg tatagcagga gttatccgaa gcgagtaggg ataacagggt aatacggaga 3000

attcgggttg gcggcgtatt tgtgtcccaa aaaacagccc caattgcccc aattgacccc 3060

aaattgaccc agtagcgggc ccaaccccgg cgagagcccc cttcacccca catatcaaac 3120

ctcccccggt tcccacactt gccgttaagg gcgtagggta ctgcagtctg gaatctacgc 3180

ttgttcagac tttgtactag tttctttgtc tggccatccg ggtaacccat gccggacgca 3240

aaatagacta ctgaaaattt ttttgctttg tggttgggac tttagccaag ggtataaaag 3300

accaccgtcc ccgaattacc tttcctcttc ttttctctct ctccttgtca actcacaccc 3360

gaaatcgtta agcatttcct tctgagtata agaatcattc aaaatggtca gctcgggtga 3420

ggacattttc agcggactcg tccccattct gattgaactc gagggtgacg ttaatggaca 3480

ccgtttctct gtccgaggtg aaggttacgg cgacgcttct aacggcaaac tggaaatcaa 3540

gttcatctgc acgacaggac ggcttcccgt gccctggcct actctggtca caaccctctc 3600

ctatggagtg caatgttttg ctaagtaccc tgagcacatg cggcaaaatg atttttttaa 3660

gagcgccatg cctgatggat acgtccaaga acgaaccatc tcctttaaag aagacggaac 3720

gtataaaaca cgagccgaag tcaaatttga aggcgaagct ctcgtcaatc gaattgatct 3780

gaaaggcctg gagtttaagg aggacggcaa catcctcgga cataaactcg agtacagctt 3840

taattcccat tacgtttaca ttacagcaga taagaataga aatggactgg aggctcagtt 3900

tcggattcgt cataacgtgg acgatggaag cgtccaactt gcagaccatt atcaacagaa 3960

tacacctatc ggcgaaggcc ccgtccttct ccccgaacag cattacctga ctaccaattc 4020

cgtccttagc aaagatcctc aggaaagacg agatcacatg gtgctggttg agtttgttac 4080

cgctgccgga ctcagcctgg gaatggatga gctctataaa tcctaagtgt ctgtggtatc 4140

taagctattt atcactcttt acaacttcta cctcaactat ctactttaat aaatgaatat 4200

cgtttattct ctatgattac tgtatatgcg ttcctctaag acaataactt cgtataatgt 4260

acattatacg aagttattgt cttgtacaaa gtccgacaca ctgtccatgg acaagttgtc 4320

gagagataga gcgtcaaagt cgaacgaggg caagtcgccg agaagagcat cgatatcaat 4380

gctgagctcc ggtggcgcag cgtagtatcc cgcggctatc acaccgacca cggcgccgag 4440

gacgaaggtg atacgtcgac ttcgggtgaa gaagttgaga atcttgccct tgacggattt 4500

tcgcttgtgg tggtgatgat ggtggtgaac gtggacgtcg gtcttggcct cggactcgga 4560

gtcagccacc tccactgctg gctcgacccg ctcgacagtc gccttccgat tcacaggttg 4620

tgtcattgat ggttgtgatc aaccggaaaa aagccgatta cacttttttt acatgagcca 4680

gaccggggta gctgagacat gtcgcgtgaa tgattagtca tggagcgagc tgcgggagac 4740

aactgctaga aggataagcc ggaattgacg ttcgtattca attaaaattc gatactgtgg 4800

cgattttagc ggcgtttgtc tgccgggcaa atcgagaatc ccactatcat atgcgcgttt 4860

cggttatttt gctggtgcat gaagagcata atattaataa atgtggggcg aacatgaaaa 4920

aagtgcggtt ggtactggaa atggatgtaa aaacagtcat ggtcgaacca aaagagtggt 4980

gccattgctt gtatttctga cttcatctgt acaaagaaat cgcgtcagaa tggtcttgtt 5040

tcagtggacc ccgcagaatg gtcctgtgtc agtggacaac agcaagacct catgcaaggg 5100

taggtccaag cacagtatca agctgctgtc gtatcgtacc tactagagaa ctctggacac 5160

agcacttgta gcaaaggtcg gtcagtactg gagcttcggt gccatggggg atacgctgtc 5220

agttgtgtgg agaagaagca gactccagat taagagaacc attctctagg tcatcgtgag 5280

acccatgtgt aacactcgct ctggagagtt agtcatccga cagggtaact ctaatctccc 5340

aacaccttat taactctgcg taactgtaac tcttcttgcc acgtcgatct tactcaattt 5400

tcctgctcat catctgctgg attgttgtct atcgtctggc tctaatacat ttattgttta 5460

ttgcccaaac aactttcatt gcacgtaagt gaattgtttt ataacagcgt tcgccaattg 5520

ctgcgccatc gtcgtccggc tgtcctaccg ttaggggtag tgtgtctcac actaccgagg 5580

ttactagagt tgggaaagcg atactgcctc ggacacacca cctgggtctt acgactgcag 5640

agagaatcgg cgttacctct ctcacaaagc ccttcagtgc ggccgcatcc caatggcgcg 5700

ccgagcttgg cgtaatcatg gtcatagctg tttcctgtgt gaaattgtta tccgctcaca 5760

attccacaca acatacgagc cggaagcata aagtgtaaag cctggggtgc ctaatgagtg 5820

agctaactca cattaattgc gttgcgctca ctgcccgctt tccagtcggg aaacctgtcg 5880

tgccagctgc attaatgaat cggccaacgc gcggggagag gcggtttgcg tattgggcgc 5940

tgttccgctt cctcgctcac tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta 6000

tcagctcact caaaggcggt aatacggtta tccacagaat caggggataa cgcaggaaag 6060

aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg 6120

tttttccata ggctccgccc ccctgacgag catcacaaaa atcgacgctc aagtcagagg 6180

tggcgaaacc cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgtg 6240

cgctctcctg ttccgaccct gccgcttacc ggatacctgt ccgcctttct cccttcggga 6300

agcgtggcgc tttctcatag ctcacgctgt aggtatctca gttcggtgta ggtcgttcgc 6360

tccaagctgg gctgtgtgca cgaacccccc gttcagcccg accgctgcgc cttatccggt 6420

aactatcgtc ttgagtccaa cccggtaaga cacgacttat cgccactggc agcagccact 6480

ggtaacagga ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtggtgg 6540

cctaactacg gctacactag aagaacagta tttggtatct gcgctctgct gaagccagtt 6600

accttcggaa aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggt 6660

ggtttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct 6720

ttgatctttt ctacggggtc tgacgctcag tggaacgaaa actcacgtta agggattttg 6780

gtcatgagat tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt 6840

aaatcaatct aaagtatata tgagtaaact tggtctgaca gttaccaatg cttaatcagt 6900

gaggcaccta tctcagcgat ctgtctattt cgttcatcca tagttgcctg actccccgtc 6960

gtgtagataa ctacgatacg ggagggctta ccatctggcc ccagtgctgc aatgataccg 7020

cgagaaccac gctcaccggc tccagattta tcagcaataa accagccagc cggaagggcc 7080

gagcgcagaa gtggtcctgc aactttatcc gcctccatcc agtctattaa ttgttgccgg 7140

gaagctagag taagtagttc gccagttaat agtttgcgca acgttgttgc cattgctaca 7200

ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat tcagctccgg ttcccaacga 7260

tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag cggttagctc cttcggtcct 7320

ccgatcgttg tcagaagtaa gttggccgca gtgttatcac tcatggttat ggcagcactg 7380

cataattctc ttactgtcat gccatccgta agatgctttt ctgtgactgg tgagtactca 7440

accaagtcat tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaata 7500

cgggataata ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg aaaacgttct 7560

tcggggcgaa aactctcaag gatcttaccg ctgttgagat ccagttcgat gtaacccact 7620

cgtgcaccca actgatcttc agcatctttt actttcacca gcgtttctgg gtgagcaaaa 7680

acaggaaggc aaaatgccgc aaaaaaggga ataagggcga cacggaaatg ttgaatactc 7740

atactcttcc tttttcaata ttattgaagc atttatcagg gttattgtct catgagcgga 7800

tacatatttg aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga 7860

aaagtgccac ctgacgtcta agaaaccatt attatcatga cattaaccta taaaaatagg 7920

cgtatcacga ggccctttg tc 7942

<210> 176
<211> 8397
<212> DNA
<213> Artificial Sequence

<220>
<223> DNA construct

<400> 176
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagactgtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240

attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat 300

tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt 360

tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt gacgcgtatt gggatgcggc 420

cgctgtcggg aaccgcgttc aggtggaaca ggaccacctc ccttgcactt cttggtatat 480

cagtataggc tgatgtattc atagtggggt ttttcataat aaatttacta acggcaggca 540

acattcactc ggcttaaacg caaaacggac cgtcttgata tcttctgacg cattgaccac 600

cgagaaatag tgttagttac cgggtgagtt attgttcttc tacacaggcg acgcccatcg 660

tctagagttg atgtactaac tcagatttca ctacctaccc tatccctggt acgcacaaag 720

cactttgcta gaaggtggtc tcaaaaacca ctacgagtag cggatttgat accacaagta 780

gcagagatat acagcaatgg tgggagtgca agtatcggaa tgtactgtac ctcctgtact 840

cgtactcgta cggcactcgt agaaacgggg caatacgggg gagaagcgat cgcccgtctg 900

ttcaatcgcc acaagtccga gtaatgctcg agtatcgaag tcttgtacct ccctgtcaat 960

catggcacca ctggtcttga cttgtctatt catactggac aagcgccaga gttaagcttg 1020

tagcgaattt cgccctcgga catcacccca tacgacggac acacatgccc gacaaacagc 1080

ctctcttatt gtagctgaaa gtatattgaa tgtgaacgtg tacaatatca ggtaccagcg 1140

ggaggttacg gccaaggtga taccggaata accctggctt ggagatggtc ggtccattgt 1200

actgaagtgt ccgtgtcgtt tccgtcactg ccccaattgg acatgtttgt ttttccgatc 1260

tttcgggcgc cctctccttg tctccttgtc tgtctcctgg actgttgcta ccccatttct 1320

ttggcctcca ttggttcctc cccgtctttc acgtcgtcta tggttgcatg gtttccctta 1380

tacttttccc cacagtcaca tgttatggag gggtctagat ggaggcctaa ttttgacgtg 1440

caaggggcga attggggcga gaaacacgtc gtggacatgg tgcaaggccc gcagggttga 1500

ttcgacgctt ttccgcgaaa aaaacaagtc caaatacccc cgtttattct ccctcggctc 1560

tcggtatttc acatgaaaac tataacctag actacacggg caaccttaac cccagagtat 1620

acttatatac caaagggatg ggtcctcaaa aatcacacaa gcaacgacgc catgaacccc 1680

aacaacactg gcaccattga aatcaacggt aaggagtaca acaccttcac cgagcccccc 1740

gtggcccccc cggcccccaa atttaaaggg ggccttttg gccataactt cgtataatgt 1800

acattatacg aagttatgaa ttctaggtta gggataacag ggtaatgata acttcgtata 1860

gcatacatta tacgaagtta tgggttggcg gcgcatttgt gtcccaaaaa acagccccaa 1920

ttgccccaat tgaccccaaa ttgacccagt agcggaccca accccggcga gagccccctt 1980

caccccacat atcaaacctc ccccggttcc cacacttgcc gttaagggcg tagggtactg 2040

cagtctggaa tctacgcttg ttcagacttt gtactagttt ctttgtctgg ccatccgggt 2100

aacccatgcc ggacgcaaaa tagactactg aaaatttttt tgctttgtgg ttgggacttt 2160

agccaagggt ataaaagacc accgtccccg aattaccttt cctcttcttt tctctctctc 2220

cttgtcaact cacacccgaa gcacaatgaa gaagcccgag ctgaccgcta cctctgttga 2280

gaagttcctg attgagaagt ttgattccgt ttccgacctg atgcagctgt ccgagggcga 2340

ggagtctcga gccttctcct ttgacgtggg cggacgaggt tacgttctgc gagtgaactc 2400

gtgtgccgac ggcttctaca aggatcgata cgtctaccga cactttgctt ctgccgctct 2460

gcccatccct gaggttctcg acattggcga gttctctgag tccctcacct actgcatctc 2520

tcgacgagct cagggagtca ccctgcagga cctccctgag actgagctgc ctgctgtcct 2580

ccagcctgtt gctgaggcca tggacgctat cgctgctgct gatctgtccc agacctcggg 2640

tttcggcccc tttggacctc agggaattgg acagtacacc acttggcgag acttcatctg 2700

tgctattgcc gatcctcacg tctaccattg gcagaccgtt atggacgata ctgtgtcggc 2760

ttctgtcgct caggctctgg acgagctgat gctctgggcc gaggattgcc ccgaggttcg 2820

acacctggtg catgctgact tcggttccaa caacgttctc accgacaacg gccgaatcac 2880

tgccgtgatt gactggtccg aggctatgtt tggcgactcg cagtacgagg tggccaacat 2940

cttcttttgg cgaccctggc tggcttgtat ggagcagcag acccgatact tcgagcgacg 3000

acatcctgag ctcgctggct cccctcgact gcgagcttac atgctccgaa ttggtctgga 3060

ccagctctac cagtcgctgg tggatggcaa ctttgacgat gctgcctggg ctcagggacg 3120

atgtgacgcc atcgtgcgat ctggcgctgg aaccgtcgga cgaactcaga ttgcccgacg 3180

atccgctgct gtctggaccg acggatgcgt ggaggtcctg gctgattcgg gtaaccgacg 3240

accctctact cgacctcgag ctaaggagta agtgtctgtg gtatctaagc tatttatcac 3300

tctttacaac ttctacctca actatctact ttaataaatg aatatcgttt attctctatg 3360

attactgtat atgcgttcct ctaagacaaa tcgaatataa cttcgtatag caggagttat 3420

ccgaagcgag tagggataac agggtaatac ggagaattcg ggttggcggc gtatttgtgt 3480

cccaaaaaac agccccaatt gccccaattg accccaaatt gacccagtag cgggcccaac 3540

cccggcgaga gccccccttca ccccacatat caaacctccc ccggttccca cacttgccgt 3600

taagggcgta gggtactgca gtctggaatc tacgcttgtt cagactttgt actagtttct 3660

ttgtctggcc atccgggtaa cccatgccgg acgcaaaata gactactgaa aatttttttg 3720

ctttgtggtt gggactttag ccaagggtat aaaagaccac cgtccccgaa ttacctttcc 3780

tcttcttttc tctctctcct tgtcaactca cacccgaaat cgttaagcat ttccttctga 3840

gtataagaat cattcaaaat ggtcagctcg ggtgaggaca ttttcagcgg actcgtcccc 3900

attctgattg aactcgaggg tgacgttaat ggacaccgtt tctctgtccg aggtgaaggt 3960

tacggcgacg cttctaacgg caaactggaa atcaagttca tctgcacgac aggacggctt 4020

cccgtgccct ggcctactct ggtcacaacc ctctcctatg gagtgcaatg ttttgctaag 4080

taccctgagc acatgcggca aaatgatttt tttaagagcg ccatgcctga tggatacgtc 4140

caagaacgaa ccatctcctt taaagaagac ggaacgtata aaacacgagc cgaagtcaaa 4200

tttgaaggcg aagctctcgt caatcgaatt gatctgaaag gcctggagtt taaggaggac 4260

ggcaacatcc tcggacataa actcgagtac agctttaatt cccattacgt ttacattaca 4320

gcagataaga atagaaatgg actggaggct cagtttcgga ttcgtcataa cgtggacgat 4380

ggaagcgtcc aacttgcaga ccattatcaa cagaatacac ctatcggcga aggccccgtc 4440

cttctccccg aacagcatta cctgactacc aattccgtcc ttagcaaaga tcctcaggaa 4500

agacgagatc acatggtgct ggttgagttt gttaccgctg ccggactcag cctgggaatg 4560

gatgagctct ataaatccta agtgtctgtg gtatctaagc tatttatcac tctttacaac 4620

ttctacctca actatctact ttaataaatg aatatcgttt attctctatg attactgtat 4680

atgcgttcct ctaagacaat aacttcgtat aatgtacatt atacgaagtt attcaagccc 4740

ttccttttcc gagaggagga ggatgatgac atttgcgagc ttgatgagga atagacaagc 4800

gggtatttat tgtatgaata aagattatgt attgattgca aaaaagtgca tttgtagatg 4860

tggtttattg tagagagtac ggtatgtact gtacgaacat taggagctac ttctacaagt 4920

agattttctt aacaagggtg aaatttacta ggaagtacat gcatatttcg ttagtagaat 4980

cacaaaagaa atgtacaagc acgtactact tgtactccac aatgtggagt gggagcaaaa 5040

aaattggacg acaccggaat cgaaccgggg acctcgcgca tgctaagcgc atgtgataac 5100

caactacacc agacgcccaa gaactttctt ggtgattatg gaatacgtgg tctgctatat 5160

ctcaattttg ctgtaatgaa tcattagaat taaaaaaaaa accccatttt tgtgtgattg 5220

tcggccaaga gatggaacag gaagaatacg tgaacaagcg agcacgaatg ccatatgctc 5280

ttctgaacaa ccgagtccga atccgatttg tgggtatcac atgtctcaag tagctgaaat 5340

gtatttcgct agaataaaat aaatgagatt aagaattaaa aatattggaa tatattttcc 5400

tagaatagaa actttggatt ttttttcggc tattacagtc tgaactggac aaacggctga 5460

ctatatataa atattattgg gtctgttttc ttgtttatgt cgaaattatc tgggttttac 5520

tactgtgtcg tcgagtatag agtggcctga ctggagaaaa tgcagtagta tggacagtag 5580

gtactgccag ccagagaagt ttttggaatt gatacttgag tcatttttcc attccccatt 5640

ccccattcca acacaatcaa ctgtttctga acattttcca aaacgcggag atgtatgtca 5700

cttggcactg caagtctcga ttcaaaatgc atgagaccca tgtgtaacac tcgctctgga 5760

gagttagtca tccgacaggg taactctaat ctcccaacac cttattaact ctgcgtaact 5820

gtaactcttc ttgccacgtc gatcttactc aattttcctg ctcatcatct gctggattgt 5880

tgtctatcgt ctggctctaa tacatttatt gtttattgcc caaacaactt tcattgcacg 5940

taagtgaatt gttttataac agcgttcgcc aattgctgcg ccatcgtcgt ccggctgtcc 6000

taccgttagg ggtagtgtgt ctcacactac cgaggttact agagttggga aagcgatact 6060

gcctcggaca caccacctgg gtcttacgac tgcagagaga atcggcgtta cctctctcac 6120

aaagcccttc agtgcggccg catcccaatg gcgcgccgag cttggcgtaa tcatggtcat 6180

agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacaacata cgagccggaa 6240

gcataaagtg taaagcctgg ggtgcctaat gagtgagcta actcacatta attgcgttgc 6300

gctcactgcc cgctttccag tcgggaaacc tgtcgtgcca gctgcattaa tgaatcggcc 6360

aacgcgcggg gagaggcggt ttgcgtattg ggcgctgttc cgcttcctcg ctcactgact 6420

cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc tcactcaaag gcggtaatac 6480

ggttatccac agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa 6540

aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg 6600

acgagcatca caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa 6660

gataccaggc gtttcccccc ggaagctccc tcgtgcgctc tcctgttccg accctgccgc 6720

ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac 6780

gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac 6840

cccccgttca gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag tccaacccgg 6900

taagacacga cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt 6960

atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa 7020

cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct 7080

cttgatccgg caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga 7140

ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg gggtctgacg 7200

ctcagtggaa cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct 7260

tcacctagat ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt 7320

aaacttggtc tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc 7380

tatttcgttc atccatagtt gcctgactcc ccgtcgtgta gataactacg atacgggagg 7440

gcttaccatc tggccccagt gctgcaatga taccgcgaga accacgctca ccggctccag 7500

atttatcagc aataaaccag ccagccggaa gggccgagcg cagaagtggt cctgcaactt 7560

tatccgcctc catccagtct attaattgtt gccgggaagc tagagtaagt agttcgccag 7620

ttaatagttt gcgcaacgtt gttgccattg ctacaggcat cgtggtgtca cgctcgtcgt 7680

ttggtatggc ttcattcagc tccggttccc aacgatcaag gcgagttaca tgatccccca 7740

tgttgtgcaa aaaagcggtt agctccttcg gtcctccgat cgttgtcaga agtaagttgg 7800

ccgcagtgtt atcactcatg gttatggcag cactgcataa ttctcttact gtcatgccat 7860

ccgtaagatg cttttctgtg actggtgagt actcaaccaa gtcattctga gaatagtgta 7920

tgcggcgacc gagttgctct tgcccggcgt caatacggga taataccgcg ccacatagca 7980

gaactttaaa agtgctcatc attggaaaac gttcttcggg gcgaaaactc tcaaggatct 8040

taccgctgtt gagatccagt tcgatgtaac ccactcgtgc acccaactga tcttcagcat 8100

cttttacttt caccagcgtt tctgggtgag caaaaacagg aaggcaaaat gccgcaaaaa 8160

agggaataag ggcgacacgg aaatgttgaa tactcatact cttccttttt caatattatt 8220

gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa 8280

ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac gtctaagaaa 8340

ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc ttttgtc 8397