

SEQUENCE LISTING

<110> DSM IP ASSETS B.V.
 Raamsdonk, L.M.
 Mueller, U.
 Wu, L.
 Winkler A.A.

<120> Acetyl CoA producing enzymes in yeast

<130> 26480WO

<140> 26480WO
 <141> 2008-07-11

<160> 52

<170> PatentIn version 3.3

<210> 1
 <211> 64
 <212> DNA
 <213> Artificial

<220>
 <223> primer 5'acs2

<400> 1
 atacacaaac agaatacagg aaagtaaadc aatacaataa taaaacagct gaagcttcgt 60
 acgc 64

<210> 2
 <211> 67
 <212> DNA
 <213> Artificial

<220>
 <223> primer 3'acs2

<400> 2
 tctcattacg aaattttttct catttaagtt atttcttttt ttgaggcata ggccactagt 60
 ggatctg 67

<210> 3
 <211> 20
 <212> DNA
 <213> Artificial

<220>
 <223> probe 5'acs2

<400> 3
 gatattcggg agccgattcc 20

```

<210> 4
<211> 20
<212> DNA
<213> Artificial

<220>
<223> probe 3'acs2

<400> 4
ccgtaacctt ctcgtaatgc 20

<210> 5
<211> 20
<212> DNA
<213> Artificial

<220>
<223> probe ACS2internal

<400> 5
cggattcgtc atcagcttca 20

<210> 6
<211> 20
<212> DNA
<213> Artificial

<220>
<223> probe KanA

<400> 6
cgcacgtcaa gactgtcaag 20

<210> 7
<211> 20
<212> DNA
<213> Artificial

<220>
<223> probe KanB

<400> 7
tcgtatgtga atgctggtcg 20

<210> 8
<211> 64
<212> DNA
<213> Artificial

<220>
<223> primer MCS-5'Tadh

```

<400> 8
aaggtacctc tagactagtc ccgggctgca gtcgactcga gcgaatttct tatgatttat 60
gatt 64

<210> 9
<211> 39
<212> DNA
<213> Artificial

<220>
<223> primer Tadh1-Hind

<400> 9
aggaagctta ggcctgtgtg gaagaacgat tacaacagg 39

<210> 10
<211> 24
<212> DNA
<213> Artificial

<220>
<223> primer P1

<400> 10
gaattgaagg atatctacat caag 24

<210> 11
<211> 25
<212> DNA
<213> Artificial

<220>
<223> primer P2

<400> 11
cccatctacg gaaccctgat caagc 25

<210> 12
<211> 26
<212> DNA
<213> Artificial

<220>
<223> primer P3

<400> 12
gatggtgtca ccattaccag gtctag 26

<210> 13
<211> 56

<212> DNA
 <213> Artificial

<220>
 <223> Primer P4

<400> 13
 gttctctggt caagttgaag tccattttga ttgatttgac tgtgttattt tgcgtg 56

<210> 14
 <211> 25
 <212> DNA
 <213> Artificial

<220>
 <223> Primer P5

<400> 14
 gaacaataga gcgaccatga ccttg 25

<210> 15
 <211> 25
 <212> DNA
 <213> Artificial

<220>
 <223> Primer P6

<400> 15
 gacatcagcg tcaccagcct tgatg 25

<210> 16
 <211> 27
 <212> DNA
 <213> Artificial

<220>
 <223> Primer P7

<400> 16
 gattgaaggt ttcaagaaca ggtgatg 27

<210> 17
 <211> 25
 <212> DNA
 <213> Artificial

<220>
 <223> Primer P8

<400> 17
 ggcgatcaga gttgaaaaaa aaatg 25

<210> 18
 <211> 1404
 <212> DNA
 <213> Escherichia coli

<220>
 <221> CDS
 <222> (1)..(1404)

<400> 18
 atg aat caa cag gat att gaa cag gtg gtg aaa gcg gta ctg ctg aaa 48
 Met Asn Gln Gln Asp Ile Glu Gln Val Val Lys Ala Val Leu Leu Lys
 1 5 10 15

atg caa agc agt gac acg ccg tcc gcc gcc gtt cat gag atg ggc gtt 96
 Met Gln Ser Ser Asp Thr Pro Ser Ala Val His Glu Met Gly Val
 20 25 30

ttc gcg tcc ctg gat gac gcc gtt gcg gca gcc aaa gtc gcc cag caa 144
 Phe Ala Ser Leu Asp Asp Ala Val Ala Ala Lys Val Ala Gln Gln
 35 40 45

ggg tta aaa agc gtg gca atg cgc cag tta gcc att gct gcc att cgt 192
 Gly Leu Lys Ser Val Ala Met Arg Gln Leu Ala Ile Ala Ala Ile Arg
 50 55 60

gaa gca ggc gaa aaa cac gcc aga gat tta gcg gaa ctt gcc gtc agt 240
 Glu Ala Gly Glu Lys His Ala Arg Asp Leu Ala Glu Leu Ala Val Ser
 65 70 75 80

gaa acc ggc atg ggg cgc gtt gaa gat aaa ttt gca aaa aac gtc gct 288
 Glu Thr Gly Met Gly Arg Val Glu Asp Lys Phe Ala Lys Asn Val Ala
 85 90 95

cag gcg cgc ggc aca cca ggc gtt gag tgc ctc tct ccg caa gtg ctg 336
 Gln Ala Arg Gly Thr Pro Gly Val Glu Cys Leu Ser Pro Gln Val Leu
 100 105 110

act ggc gac aac ggc ctg acc cta att gaa aac gca ccc tgg ggc gtg 384
 Thr Gly Asp Asn Gly Leu Thr Leu Ile Glu Asn Ala Pro Trp Gly Val
 115 120 125

gtg gct tcg gtg acg cct tcc act aac ccg gcg gca acc gta att aac 432
 Val Ala Ser Val Thr Pro Ser Thr Asn Pro Ala Ala Thr Val Ile Asn
 130 135 140

aac gcc atc agc ctg att gcc gcg ggc aac agc gtc att ttt gcc ccg 480
 Asn Ala Ile Ser Leu Ile Ala Ala Gly Asn Ser Val Ile Phe Ala Pro
 145 150 155 160

cat ccg gcg gcg aaa aaa gtc tcc cag cgg gcg att acg ctg ctc aac 528
 His Pro Ala Ala Lys Lys Val Ser Gln Arg Ala Ile Thr Leu Leu Asn
 165 170 175

cag gcg att gtt gcc gca ggt ggg ccg gaa aac tta ctg gtt act gtg 576

Gln	Ala	Ile	Val	Ala	Ala	Gly	Gly	Pro	Glu	Asn	Leu	Leu	Val	Thr	Val	
			180					185					190			
gca	aat	ccg	gat	atc	gaa	acc	gcg	caa	cgc	ttg	ttc	aag	ttt	ccg	ggt	624
Ala	Asn	Pro	Asp	Ile	Glu	Thr	Ala	Gln	Arg	Leu	Phe	Lys	Phe	Pro	Gly	
		195					200					205				
atc	ggc	ctg	ctg	gtg	gta	acc	ggc	ggc	gaa	gcg	gta	gta	gaa	gcg	gcg	672
Ile	Gly	Leu	Leu	Val	Val	Thr	Gly	Gly	Glu	Ala	Val	Val	Glu	Ala	Ala	
	210					215					220					
cgt	aaa	cac	acc	aat	aaa	cgt	ctg	att	gcc	gca	ggc	gct	ggc	aac	ccg	720
Arg	Lys	His	Thr	Asn	Lys	Arg	Leu	Ile	Ala	Ala	Gly	Ala	Gly	Asn	Pro	
225					230				235						240	
ccg	gta	gtg	gtg	gat	gaa	acc	gcc	gac	ctc	gcc	cgt	gcc	gct	cag	tcc	768
Pro	Val	Val	Val	Asp	Glu	Thr	Ala	Asp	Leu	Ala	Arg	Ala	Ala	Gln	Ser	
				245					250					255		
atc	gtc	aaa	ggc	gct	tct	ttc	gat	aac	aac	atc	att	tgt	gcc	gac	gaa	816
Ile	Val	Lys	Gly	Ala	Ser	Phe	Asp	Asn	Asn	Ile	Ile	Cys	Ala	Asp	Glu	
		260					265						270			
aag	gta	ctg	att	gtt	gtt	gat	agc	gta	gcc	gat	gaa	ctg	atg	cgt	ctg	864
Lys	Val	Leu	Ile	Val	Val	Asp	Ser	Val	Ala	Asp	Glu	Leu	Met	Arg	Leu	
	275					280						285				
atg	gaa	ggc	cag	cac	gcg	gtg	aaa	ctg	acc	gca	gaa	cag	gcg	cag	cag	912
Met	Glu	Gly	Gln	His	Ala	Val	Lys	Leu	Thr	Ala	Glu	Gln	Ala	Gln	Gln	
	290					295					300					
ctg	caa	ccg	gtg	ttg	ctg	aaa	aat	atc	gac	gag	cgc	gga	aaa	ggc	acc	960
Leu	Gln	Pro	Val	Leu	Leu	Lys	Asn	Ile	Asp	Glu	Arg	Gly	Lys	Gly	Thr	
305				310					315						320	
gtc	agc	cgt	gac	tgg	gtt	ggc	cgc	gac	gca	ggc	aaa	atc	gcg	gcg	gca	1008
Val	Ser	Arg	Asp	Trp	Val	Gly	Arg	Asp	Ala	Gly	Lys	Ile	Ala	Ala	Ala	
				325					330					335		
atc	ggc	ctt	aaa	gtt	ccg	caa	gaa	acg	cgc	ctg	ctg	ttt	gtg	gaa	acc	1056
Ile	Gly	Leu	Lys	Val	Pro	Gln	Glu	Thr	Arg	Leu	Leu	Phe	Val	Glu	Thr	
			340					345					350			
acc	gca	gaa	cat	ccg	ttt	gcc	gtg	act	gaa	ctg	atg	atg	ccg	gtg	ttg	1104
Thr	Ala	Glu	His	Pro	Phe	Ala	Val	Thr	Glu	Leu	Met	Met	Pro	Val	Leu	
		355				360						365				
ccc	gtc	gtg	cgc	gtc	gcc	aac	gtg	gcg	gat	gcc	att	gcg	cta	gcg	gtg	1152
Pro	Val	Val	Arg	Val	Ala	Asn	Val	Ala	Asp	Ala	Ile	Ala	Leu	Ala	Val	
	370					375					380					
aaa	ctg	gaa	ggc	ggc	tgc	cac	cac	acg	gcg	gca	atg	cac	tcg	cgc	aac	1200
Lys	Leu	Glu	Gly	Gly	Cys	His	His	Thr	Ala	Ala	Met	His	Ser	Arg	Asn	
385					390				395						400	
atc	gaa	aac	atg	aac	cag	atg	gcg	aat	gct	att	gat	acc	agc	att	ttc	1248
Ile	Glu	Asn	Met	Asn	Gln	Met	Ala	Asn	Ala	Ile	Asp	Thr	Ser	Ile	Phe	

405										410					415					
gtt	aag	aac	gga	ccg	tgc	att	gcc	ggg	ctg	ggg	ctg	ggc	ggg	gaa	ggc	1296				
Val	Lys	Asn	Gly	Pro	Cys	Ile	Ala	Gly	Leu	Gly	Leu	Gly	Gly	Glu	Gly					
			420					425					430							
tgg	acc	acc	atg	acc	atc	acc	acg	cca	acc	ggg	gaa	ggg	gta	acc	agc	1344				
Trp	Thr	Thr	Met	Thr	Ile	Thr	Thr	Pro	Thr	Gly	Glu	Gly	Val	Thr	Ser					
		435					440					445								
gcg	cgt	acg	ttt	gtc	cgt	ctg	cgt	cgc	tgt	gta	tta	gtc	gat	gcg	ttt	1392				
Ala	Arg	Thr	Phe	Val	Arg	Leu	Arg	Arg	Cys	Val	Leu	Val	Asp	Ala	Phe					
	450					455					460									
cgc	att	gtt	taa													1404				
Arg	Ile	Val																		
465																				

<210> 19
 <211> 467
 <212> PRT
 <213> Escherichia coli

<400> 19

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

Met	Gln	Ser	Ser	Asp	Thr	Pro	Ser	Ala	Ala	Val	His	Glu	Met	Gly	Val
			20					25					30		

Phe	Ala	Ser	Leu	Asp	Asp	Ala	Val	Ala	Ala	Ala	Lys	Val	Ala	Gln	Gln
		35					40					45			

Gly	Leu	Lys	Ser	Val	Ala	Met	Arg	Gln	Leu	Ala	Ile	Ala	Ala	Ile	Arg
	50					55					60				

Glu	Ala	Gly	Glu	Lys	His	Ala	Arg	Asp	Leu	Ala	Glu	Leu	Ala	Val	Ser
65					70					75					80

Glu	Thr	Gly	Met	Gly	Arg	Val	Glu	Asp	Lys	Phe	Ala	Lys	Asn	Val	Ala
				85					90					95	

Gln	Ala	Arg	Gly	Thr	Pro	Gly	Val	Glu	Cys	Leu	Ser	Pro	Gln	Val	Leu
			100					105					110		

Thr	Gly	Asp	Asn	Gly	Leu	Thr	Leu	Ile	Glu	Asn	Ala	Pro	Trp	Gly	Val
		115					120					125			

Val Ala Ser Val Thr Pro Ser Thr Asn Pro Ala Ala Thr Val Ile Asn
130 135 140

Asn Ala Ile Ser Leu Ile Ala Ala Gly Asn Ser Val Ile Phe Ala Pro
145 150 155 160

His Pro Ala Ala Lys Lys Val Ser Gln Arg Ala Ile Thr Leu Leu Asn
165 170 175

Gln Ala Ile Val Ala Ala Gly Gly Pro Glu Asn Leu Leu Val Thr Val
180 185 190

Ala Asn Pro Asp Ile Glu Thr Ala Gln Arg Leu Phe Lys Phe Pro Gly
195 200 205

Ile Gly Leu Leu Val Val Thr Gly Gly Glu Ala Val Val Glu Ala Ala
210 215 220

Arg Lys His Thr Asn Lys Arg Leu Ile Ala Ala Gly Ala Gly Asn Pro
225 230 235 240

Pro Val Val Val Asp Glu Thr Ala Asp Leu Ala Arg Ala Ala Gln Ser
245 250 255

Ile Val Lys Gly Ala Ser Phe Asp Asn Asn Ile Ile Cys Ala Asp Glu
260 265 270

Lys Val Leu Ile Val Val Asp Ser Val Ala Asp Glu Leu Met Arg Leu
275 280 285

Met Glu Gly Gln His Ala Val Lys Leu Thr Ala Glu Gln Ala Gln Gln
290 295 300

Leu Gln Pro Val Leu Leu Lys Asn Ile Asp Glu Arg Gly Lys Gly Thr
305 310 315 320

Val Ser Arg Asp Trp Val Gly Arg Asp Ala Gly Lys Ile Ala Ala Ala
325 330 335

Ile Gly Leu Lys Val Pro Gln Glu Thr Arg Leu Leu Phe Val Glu Thr
340 345 350

Thr Ala Glu His Pro Phe Ala Val Thr Glu Leu Met Met Pro Val Leu
 355 360 365

Pro Val Val Arg Val Ala Asn Val Ala Asp Ala Ile Ala Leu Ala Val
 370 375 380

Lys Leu Glu Gly Gly Cys His His Thr Ala Ala Met His Ser Arg Asn
 385 390 395 400

Ile Glu Asn Met Asn Gln Met Ala Asn Ala Ile Asp Thr Ser Ile Phe
 405 410 415

Val Lys Asn Gly Pro Cys Ile Ala Gly Leu Gly Leu Gly Gly Glu Gly
 420 425 430

Trp Thr Thr Met Thr Ile Thr Thr Pro Thr Gly Glu Gly Val Thr Ser
 435 440 445

Ala Arg Thr Phe Val Arg Leu Arg Arg Cys Val Leu Val Asp Ala Phe
 450 455 460

Arg Ile Val
 465

<210> 20
 <211> 1401
 <212> DNA
 <213> Artificial

<220>
 <223> optimised sequence

<400> 20
 atgaaccaac aagatatcga acaagttgtc aaggctgtct tgttgaaaat gcaatcttct 60
 gacactccat ctgctgctgt ccacgaaatg ggtgttttcg cttctttgga cgacgctgtt 120
 gctgctgcc aagttgctca acaaggtttg aaatctgttg ccatgagaca attggccatt 180
 gctgccatca gagaagctgg tgaaaagcat gccagagact tggctgaatt ggctgtctcc 240
 gaaaccggtg tgggtagagt tgaagacaaa ttcgctaaga acgttgctca agctagaggt 300
 actccaggtg tcgaatgttt gtctccacaa gtcttgaccg gtgataatgg tttgactttg 360
 attgaaaatg ctccatgggg tgttgttgct tccgtcacc catctaccaa cccagctgct 420
 actgtcatca acaacgccat ctctttgatt gctgctggta actccgttat cttcgtccca 480

```

caccagctg ccaagaaggt ttctcaaaga gccatcactc tattgaacca agccattgtt 540
gctgctgggt gtccagaaaa cttggtgggc actggttgcca acccagatat cgaaactgct 600
caaagattat tcaagttccc aggtatcggg ctattagtcg tctactgggtg tgaagctgtt 660
gttgaagctg ccagaaagca caccaacaag agattgattg ctgctgggtgc tggtaaccct 720
cctgttggtg tcgatgaaac cgctgatttg gccagagctg ctcaatccat tgtcaagggt 780
gcttctttcg acaacaacat catctgtgct gacgaaaagg ttttgattgt tgttgactcc 840
gttgctgacg aattgatgag attgatggaa ggtcaacatg ccgtcaagtt gactgctgaa 900
caagctcaac aattgcaacc agttttgttg aagaacatcg atgaaagagg taaggggtacc 960
gtctccagag actggggttg tagagatgct ggtaagattg ctgctgccat cggtttgaag 1020
gttcacaag aaaccagatt attattcgtc gaaaccaccg ctgaacaccc atttgctgtc 1080
actgaattga tgatgccagt cttaccagtt gtccgtgttg ctaacgttgc tgacgctatt 1140
gctttggctg tcaaattgga aggtgggtgt caccacactg ctgccatgca ctccagaaac 1200
atcgaaaaca tgaaccaa at ggctaacgcc attgacactt ccattcttgt caagaacggt 1260
ccatgtatcg ctggtttggg tttgggtggg gaagggttga ccaccatgac catcaccacc 1320
ccaactgggt aaggtgtcac ttctgccaga actttcgtca gattacgtcg ttgtgttttg 1380
gtcgatgctt tcagaattgt t 1401

```

```

<210> 21
<211> 1410
<212> DNA
<213> Listeria innocua

```

```

<220>
<221> CDS
<222> (1)..(1410)

```

```

<400> 21
atg gaa tca tta gaa ctc gaa caa ctg gta aaa aaa gtt ctc tta gaa 48
Met Glu Ser Leu Glu Leu Glu Gln Leu Val Lys Lys Val Leu Leu Glu
1 5 10 15

aaa tta gca gaa caa aaa gaa gta cca aca aaa aca act aca caa ggc 96
Lys Leu Ala Glu Gln Lys Glu Val Pro Thr Lys Thr Thr Thr Gln Gly
20 25 30

gcg aaa agt ggc gtt ttt gat aca gtt gac gag gct gtt caa gca gca 144
Ala Lys Ser Gly Val Phe Asp Thr Val Asp Glu Ala Val Gln Ala Ala
35 40 45

gtt ata gcg cag aat tgc tat aaa gaa aaa tca ctt gaa gaa cgc cgc 192

```

Val	Ile	Ala	Gln	Asn	Cys	Tyr	Lys	Glu	Lys	Ser	Leu	Glu	Glu	Arg	Arg	
50						55					60					
aat	ggt	gta	aaa	gca	att	cgt	gaa	gca	ctt	tat	cca	gaa	att	gaa	aca	240
Asn	Val	Val	Lys	Ala	Ile	Arg	Glu	Ala	Leu	Tyr	Pro	Glu	Ile	Glu	Thr	
65					70					75					80	
att	gcg	aca	aga	gca	ggt	gca	gag	act	ggt	atg	gga	aat	gtg	aca	gat	288
Ile	Ala	Thr	Arg	Ala	Val	Ala	Glu	Thr	Gly	Met	Gly	Asn	Val	Thr	Asp	
				85					90					95		
aaa	att	ttg	aaa	aac	acg	tta	gca	atc	gaa	aaa	acg	cca	ggg	gta	gaa	336
Lys	Ile	Leu	Lys	Asn	Thr	Leu	Ala	Ile	Glu	Lys	Thr	Pro	Gly	Val	Glu	
			100					105					110			
gat	tta	tat	aca	gaa	gta	gct	aca	ggt	gat	aac	ggt	atg	aca	cta	tat	384
Asp	Leu	Tyr	Thr	Glu	Val	Ala	Thr	Gly	Asp	Asn	Gly	Met	Thr	Leu	Tyr	
		115					120					125				
gaa	ctc	tct	ccg	tat	ggc	gta	att	ggt	gca	gta	gcg	ccg	agc	aca	aac	432
Glu	Leu	Ser	Pro	Tyr	Gly	Val	Ile	Gly	Ala	Val	Ala	Pro	Ser	Thr	Asn	
	130					135					140					
cca	acg	gaa	aca	ttg	att	tgt	aat	tca	atc	ggt	atg	ctc	gca	gct	gga	480
Pro	Thr	Glu	Thr	Leu	Ile	Cys	Asn	Ser	Ile	Gly	Met	Leu	Ala	Ala	Gly	
145					150					155					160	
aat	gcc	ggt	ttt	tat	agc	cct	cat	cca	ggg	gca	aaa	aac	att	tca	ctg	528
Asn	Ala	Val	Phe	Tyr	Ser	Pro	His	Pro	Gly	Ala	Lys	Asn	Ile	Ser	Leu	
				165					170					175		
tgg	ttg	att	gaa	aaa	cta	aac	aca	att	ggt	cgc	gat	agt	tgt	ggt	ata	576
Trp	Leu	Ile	Glu	Lys	Leu	Asn	Thr	Ile	Val	Arg	Asp	Ser	Cys	Gly	Ile	
			180					185					190			
gat	aat	cta	att	gtc	acc	gtg	gct	aaa	cca	tcc	atc	caa	gca	gct	caa	624
Asp	Asn	Leu	Ile	Val	Thr	Val	Ala	Lys	Pro	Ser	Ile	Gln	Ala	Ala	Gln	
		195					200					205				
gaa	atg	atg	aac	cat	cca	aaa	gta	ccg	cta	ctt	ggt	att	aca	ggt	ggt	672
Glu	Met	Met	Asn	His	Pro	Lys	Val	Pro	Leu	Leu	Val	Ile	Thr	Gly	Gly	
	210					215					220					
ccg	ggc	ggt	ggt	ctc	caa	gcg	atg	caa	tca	ggt	aaa	aaa	gtg	att	gga	720
Pro	Gly	Val	Val	Leu	Gln	Ala	Met	Gln	Ser	Gly	Lys	Lys	Val	Ile	Gly	
225					230					235					240	
gca	gga	gca	ggg	aac	ccg	cct	tct	att	ggt	gac	gaa	aca	gct	aat	atc	768
Ala	Gly	Ala	Gly	Asn	Pro	Pro	Ser	Ile	Val	Asp	Glu	Thr	Ala	Asn	Ile	
				245					250					255		
gaa	aaa	gcg	gct	gct	gac	atc	gta	gac	gga	gca	tct	ttt	gac	cat	aat	816
Glu	Lys	Ala	Ala	Ala	Asp	Ile	Val	Asp	Gly	Ala	Ser	Phe	Asp	His	Asn	
			260				265						270			
att	tta	tgt	att	gct	gaa	aaa	agt	gtg	gta	gct	ggt	gat	agc	att	gct	864
Ile	Leu	Cys	Ile	Ala	Glu	Lys	Ser	Val	Val	Ala	Val	Asp	Ser	Ile	Ala	

275	280	285	
gat ttc ttg tta ttc caa atg gaa aaa aat ggt gcc ctt cat gtt act Asp Phe Leu Leu Phe Gln Met Glu Lys Asn Gly Ala Leu His Val Thr 290 295 300			912
aat cca agt gat att caa aaa tta gaa aaa gta gcc gtt acc gat aaa Asn Pro Ser Asp Ile Gln Lys Leu Glu Lys Val Ala Val Thr Asp Lys 305 310 315 320			960
ggt gta act aat aaa aaa tta gtc gga aaa agt gca act gaa atc tta Gly Val Thr Asn Lys Lys Leu Val Gly Lys Ser Ala Thr Glu Ile Leu 325 330 335			1008
aaa gaa gca gga ata gct tgt gat ttt aca cca cgt tta atc att gtg Lys Glu Ala Gly Ile Ala Cys Asp Phe Thr Pro Arg Leu Ile Ile Val 340 345 350			1056
gaa acg gag aaa tct cat cca ttt gca aca gta gag cta tta atg cca Glu Thr Glu Lys Ser His Pro Phe Ala Thr Val Glu Leu Leu Met Pro 355 360 365			1104
atc gtt cca gtt gta agg gtg cct gat ttt gac gaa gcc ctt gaa gtg Ile Val Pro Val Val Arg Val Pro Asp Phe Asp Glu Ala Leu Glu Val 370 375 380			1152
gct att gaa ctc gaa caa ggc tta cat cat aca gca aca atg cat tca Ala Ile Glu Leu Glu Gln Gly Leu His His Thr Ala Thr Met His Ser 385 390 395 400			1200
caa aat atc tcg aga tta aac aaa gct gca aga gat atg caa act tcc Gln Asn Ile Ser Arg Leu Asn Lys Ala Ala Arg Asp Met Gln Thr Ser 405 410 415			1248
atc ttt gtc aaa aat ggt ccg tcc ttt gcg gga tta ggc ttt aga gga Ile Phe Val Lys Asn Gly Pro Ser Phe Ala Gly Leu Gly Phe Arg Gly 420 425 430			1296
gaa ggt agt act act ttc act att gca acg cct act gga gaa gga aca Glu Gly Ser Thr Thr Phe Thr Ile Ala Thr Pro Thr Gly Glu Gly Thr 435 440 445			1344
act aca gca cgt cat ttt gct aga cgc cgc cgc tgt gtt tta aca gat Thr Thr Ala Arg His Phe Ala Arg Arg Arg Arg Cys Val Leu Thr Asp 450 455 460			1392
ggt ttt tcg att cgt taa Gly Phe Ser Ile Arg 465			1410

<210> 22
 <211> 469
 <212> PRT
 <213> *Listeria innocua*
 <400> 22

Met Glu Ser Leu Glu Leu Glu Gln Leu Val Lys Lys Val Leu Leu Glu
1 5 10 15

Lys Leu Ala Glu Gln Lys Glu Val Pro Thr Lys Thr Thr Thr Gln Gly
20 25 30

Ala Lys Ser Gly Val Phe Asp Thr Val Asp Glu Ala Val Gln Ala Ala
35 40 45

Val Ile Ala Gln Asn Cys Tyr Lys Glu Lys Ser Leu Glu Glu Arg Arg
50 55 60

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

Ile Ala Thr Arg Ala Val Ala Glu Thr Gly Met Gly Asn Val Thr Asp
85 90 95

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

Asp Leu Tyr Thr Glu Val Ala Thr Gly Asp Asn Gly Met Thr Leu Tyr
115 120 125

Glu Leu Ser Pro Tyr Gly Val Ile Gly Ala Val Ala Pro Ser Thr Asn
130 135 140

Pro Thr Glu Thr Leu Ile Cys Asn Ser Ile Gly Met Leu Ala Ala Gly
145 150 155 160

Asn Ala Val Phe Tyr Ser Pro His Pro Gly Ala Lys Asn Ile Ser Leu
165 170 175

Trp Leu Ile Glu Lys Leu Asn Thr Ile Val Arg Asp Ser Cys Gly Ile
180 185 190

Asp Asn Leu Ile Val Thr Val Ala Lys Pro Ser Ile Gln Ala Ala Gln
195 200 205

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

Pro Gly Val Val Leu Gln Ala Met Gln Ser Gly Lys Lys Val Ile Gly
225 230 235 240

Ala Gly Ala Gly Asn Pro Pro Ser Ile Val Asp Glu Thr Ala Asn Ile
245 250 255

Glu Lys Ala Ala Ala Asp Ile Val Asp Gly Ala Ser Phe Asp His Asn
260 265 270

Ile Leu Cys Ile Ala Glu Lys Ser Val Val Ala Val Asp Ser Ile Ala
275 280 285

Asp Phe Leu Leu Phe Gln Met Glu Lys Asn Gly Ala Leu His Val Thr
290 295 300

Asn Pro Ser Asp Ile Gln Lys Leu Glu Lys Val Ala Val Thr Asp Lys
305 310 315 320

Gly Val Thr Asn Lys Lys Leu Val Gly Lys Ser Ala Thr Glu Ile Leu
325 330 335

Lys Glu Ala Gly Ile Ala Cys Asp Phe Thr Pro Arg Leu Ile Ile Val
340 345 350

Glu Thr Glu Lys Ser His Pro Phe Ala Thr Val Glu Leu Leu Met Pro
355 360 365

Ile Val Pro Val Val Arg Val Pro Asp Phe Asp Glu Ala Leu Glu Val
370 375 380

Ala Ile Glu Leu Glu Gln Gly Leu His His Thr Ala Thr Met His Ser
385 390 395 400

Gln Asn Ile Ser Arg Leu Asn Lys Ala Ala Arg Asp Met Gln Thr Ser
405 410 415

Ile Phe Val Lys Asn Gly Pro Ser Phe Ala Gly Leu Gly Phe Arg Gly
420 425 430

Glu Gly Ser Thr Thr Phe Thr Ile Ala Thr Pro Thr Gly Glu Gly Thr
435 440 445

Thr Thr Ala Arg His Phe Ala Arg Arg Arg Arg Cys Val Leu Thr Asp

450

455

460

Gly Phe Ser Ile Arg
465

<210> 23
<211> 1407
<212> DNA
<213> Artificial

<220>
<223> optimised sequence

<400> 23
atggaatctt tggaattgga acaattagtc aagaaggttt tgttggaata attggctgaa 60
caaaaggaag ttccaaccaa gaccaccacc caagggtgcca agtccggtgt tttcgatacc 120
gtcgatgaag ctgtccaagc tgccgtcatt gctcaaaact gttacaagga aaaatctttg 180
gaagaaagaa gaaacgttgt caaggccatc agagaagctt tataaccaga aatcgaaacc 240
attgctacca gagctgttgc tgaaaccggt atgggtaatg tcaccgataa aatcttgaag 300
aacacttttag ctatcgaaaa gactccaggt gttgaagact tgtacactga agttgctacc 360
ggtgacaacg gtatgacttt atacgaatta tctccatacg gtgtcatcgg tgctgttgct 420
ccatctacca acccaactga aactttgatc tgtaactcca tcggatatgtt ggctgctggt 480
aacgccgttt tctactctcc tcaccaggt gccagaaca tctctttatg gttgattgaa 540
aagttgaaca ctatcgtcag agattcttgt ggtattgaca acttgattgt caccgttgcc 600
aagccatcta tccaagctgc tcaagaaatg atgaaccacc caaagggtcc attgttggtc 660
atcactggtg gtccaggtgt tgtcttgcaa gctatgcaat ctggtaagaa ggttatcggc 720
gctggtgctg gtaaccctcc atccatcggt gacgaaaccg ctaacattga aaaggctgct 780
gctgacattg tcgacggtgc ttcctttgac cataatatct tgtgtatcgc tgaaaagtct 840
gttggtgccg ttgactccat tgctgacttc ttgttggtcc aaatggaaaa gaacggtgct 900
ttgcacgtca ctaaccatc tgatatccaa aaattggaaa aggttgccgt cactgacaag 960
gggtgcacca acaagaaatt gggtggtaag tctgccactg aaatcttgaa agaagctggt 1020
attgcttggtg atttcacccc aagattgatc attgtcgaaa ctgaaaagtc ccaccattc 1080
gctactgttg aattgttgat gccaatgtt ccagttgtca gagttccaga cttcgatgaa 1140
gctttggaag ttgccattga attggaacaa ggtctacatc aactgctac catgcactct 1200
caaacatct ccagattgaa caaggctgcc cgtgacatgc aaacctccat ctttgtcaag 1260

```

aacggtccat ctttcgctgg tttaggtttc agaggtgaag gttccaccac tttcaccatt 1320
gctactccaa ctggtgaagg tactaccact gcccgtcact tcgctagaag aagaagatgt 1380
gtcttgactg atggtttctc cattaga 1407

```

```

<210> 24
<211> 1476
<212> DNA
<213> Clostridium kluyveri

```

```

<220>
<221> CDS
<222> (1)..(1476)

```

```

<400> 24
atg gag ata atg gat aag gac tta cag tca ata cag gaa gta aga act 48
Met Glu Ile Met Asp Lys Asp Leu Gln Ser Ile Gln Glu Val Arg Thr
1 5 10 15

ctt ata gca aaa gca aag aaa gct caa gca gaa ttt aaa aat ttt tct 96
Leu Ile Ala Lys Ala Lys Lys Ala Gln Ala Glu Phe Lys Asn Phe Ser
20 25 30

caa gaa gct gta aac aag gta ata gaa aaa ata gct aag gct aca gaa 144
Gln Glu Ala Val Asn Lys Val Ile Glu Lys Ile Ala Lys Ala Thr Glu
35 40 45

gtt gaa gct gta aaa ctt gca aaa ttg gca tat gaa gat aca gga tat 192
Val Glu Ala Val Lys Leu Ala Lys Leu Ala Tyr Glu Asp Thr Gly Tyr
50 55 60

gga aaa tgg gaa gat aaa gta ata aag aat aag ttt tca agt ata gta 240
Gly Lys Trp Glu Asp Lys Val Ile Lys Asn Lys Phe Ser Ser Ile Val
65 70 75 80

gtt tat aac tat att aaa gat ttg aaa acg gtt gga att tta aaa gaa 288
Val Tyr Asn Tyr Ile Lys Asp Leu Lys Thr Val Gly Ile Leu Lys Glu
85 90 95

gac aag gaa aag aaa tta ata gat ata gct gtt cca ctt gga gtt ata 336
Asp Lys Glu Lys Lys Leu Ile Asp Ile Ala Val Pro Leu Gly Val Ile
100 105 110

gca gga ctt ata cct tca act aac cca act tca aca gca ata ttc aag 384
Ala Gly Leu Ile Pro Ser Thr Asn Pro Thr Ser Thr Ala Ile Phe Lys
115 120 125

gta tta ata gca tta aag gca gga aat gca ata gta ttc tca cca cat 432
Val Leu Ile Ala Leu Lys Ala Gly Asn Ala Ile Val Phe Ser Pro His
130 135 140

cca aca gca gta aga agt att aca gaa act gta aag ata atg cag aaa 480
Pro Thr Ala Val Arg Ser Ile Thr Glu Thr Val Lys Ile Met Gln Lys

```


145					150					155				160	
gct gca gta gaa gca gga gca cca gat gga tta atc caa tgt atg tca	528														
Ala Ala Val Glu Ala Gly Ala Pro Asp Gly Leu Ile Gln Cys Met Ser				165				170					175		
ata ttg aca gta gaa ggt act gct gaa ttg atg aag aat aag gat aca	576														
Ile Leu Thr Val Glu Gly Thr Ala Glu Leu Met Lys Asn Lys Asp Thr			180				185					190			
gca ctt atc ctt gca aca ggt gga gaa gga atg gta aga gca gct tac	624														
Ala Leu Ile Leu Ala Thr Gly Gly Glu Gly Met Val Arg Ala Ala Tyr			195			200				205					
agt tca gga aca cca gct ata gga gtt gga cct gga aac ggc cca tgc	672														
Ser Ser Gly Thr Pro Ala Ile Gly Val Gly Pro Gly Asn Gly Pro Cys			210			215			220						
ttt att gaa aga aca gca gat att cct aca gca gta aga aaa gta ata	720														
Phe Ile Glu Arg Thr Ala Asp Ile Pro Thr Ala Val Arg Lys Val Ile				230				235						240	
ggc agt gat act ttt gat aat gga gta ata tgt gct tca gaa caa tca	768														
Gly Ser Asp Thr Phe Asp Asn Gly Val Ile Cys Ala Ser Glu Gln Ser			245				250						255		
ata ata gca gag aca gta aag aaa gca gag ata att gaa gaa ttc aag	816														
Ile Ile Ala Glu Thr Val Lys Lys Ala Glu Ile Ile Glu Glu Phe Lys			260				265					270			
aga caa aaa gga tat ttc tta aat gca gaa gaa tca gaa aaa gta ggc	864														
Arg Gln Lys Gly Tyr Phe Leu Asn Ala Glu Glu Ser Glu Lys Val Gly			275			280					285				
aag att tta tta aga gct aat gga aca cca aac cca gca ata gta gga	912														
Lys Ile Leu Leu Arg Ala Asn Gly Thr Pro Asn Pro Ala Ile Val Gly			290			295				300					
aaa gat gtt caa gca tta gca aaa tta gca gga ata agc ata cca agc	960														
Lys Asp Val Gln Ala Leu Ala Lys Leu Ala Gly Ile Ser Ile Pro Ser			305			310			315				320		
gat gcg gta ata tta ctt tca gag cag aca gat gtg agt cca aag aac	1008														
Asp Ala Val Ile Leu Leu Ser Glu Gln Thr Asp Val Ser Pro Lys Asn			325				330					335			
cct tat gca aag gaa aaa tta gct cca gta ctt gca ttc tat aca gta	1056														
Pro Tyr Ala Lys Glu Lys Leu Ala Pro Val Leu Ala Phe Tyr Thr Val			340				345					350			
gaa gac tgg cat gaa gca tgt gaa aaa tcc tta gca ctt ctt cat aac	1104														
Glu Asp Trp His Glu Ala Cys Glu Lys Ser Leu Ala Leu Leu His Asn			355				360				365				
caa gga agt gga cat aca tta ata att cac tca cag aat gaa gaa atc	1152														
Gln Gly Ser Gly His Thr Leu Ile Ile His Ser Gln Asn Glu Glu Ile			370			375			380						

ata aga gaa ttc gca ttg aag aaa cca gta tca aga ata ctt gta aat 1200
 Ile Arg Glu Phe Ala Leu Lys Lys Pro Val Ser Arg Ile Leu Val Asn
 385 390 395 400

tca cct gga tca ctt gga gga ata ggt gga gct aca aat ctt gta cca 1248
 Ser Pro Gly Ser Leu Gly Gly Ile Gly Gly Ala Thr Asn Leu Val Pro
 405 410 415

tca ctt aca tta ggc tgt gga gca gta ggt gga agt gca act tca gat 1296
 Ser Leu Thr Leu Gly Cys Gly Ala Val Gly Gly Ser Ala Thr Ser Asp
 420 425 430

aac gta gga cca gaa aac tta ttc aac ata aga aaa gta gct tat gga 1344
 Asn Val Gly Pro Glu Asn Leu Phe Asn Ile Arg Lys Val Ala Tyr Gly
 435 440 445

act acg aca gta gaa gaa ata aga gaa gct ttt ggt gta gga gca gct 1392
 Thr Thr Thr Val Glu Glu Ile Arg Glu Ala Phe Gly Val Gly Ala Ala
 450 455 460

tca tca agt gca cca gca gaa cca gaa gat aat gaa gat gta cag gct 1440
 Ser Ser Ser Ala Pro Ala Glu Pro Glu Asp Asn Glu Asp Val Gln Ala
 465 470 475 480

ata gta aaa gct ata atg gct aaa tta aat ctt taa 1476
 Ile Val Lys Ala Ile Met Ala Lys Leu Asn Leu
 485 490

<210> 25
 <211> 491
 <212> PRT
 <213> Clostridium kluyveri

<400> 25

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

Leu Ile Ala Lys Ala Lys Lys Ala Gln Ala Glu Phe Lys Asn Phe Ser
 20 25 30

Gln Glu Ala Val Asn Lys Val Ile Glu Lys Ile Ala Lys Ala Thr Glu
 35 40 45

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

Gly Lys Trp Glu Asp Lys Val Ile Lys Asn Lys Phe Ser Ser Ile Val
 65 70 75 80

Val Tyr Asn Tyr Ile Lys Asp Leu Lys Thr Val Gly Ile Leu Lys Glu
85 90 95

Asp Lys Glu Lys Lys Leu Ile Asp Ile Ala Val Pro Leu Gly Val Ile
100 105 110

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

Val Leu Ile Ala Leu Lys Ala Gly Asn Ala Ile Val Phe Ser Pro His
130 135 140

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

Ala Ala Val Glu Ala Gly Ala Pro Asp Gly Leu Ile Gln Cys Met Ser
165 170 175

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

Ala Leu Ile Leu Ala Thr Gly Gly Glu Gly Met Val Arg Ala Ala Tyr
195 200 205

Ser Ser Gly Thr Pro Ala Ile Gly Val Gly Pro Gly Asn Gly Pro Cys
210 215 220

Phe Ile Glu Arg Thr Ala Asp Ile Pro Thr Ala Val Arg Lys Val Ile
225 230 235 240

Gly Ser Asp Thr Phe Asp Asn Gly Val Ile Cys Ala Ser Glu Gln Ser
245 250 255

Ile Ile Ala Glu Thr Val Lys Lys Ala Glu Ile Ile Glu Glu Phe Lys
260 265 270

Arg Gln Lys Gly Tyr Phe Leu Asn Ala Glu Glu Ser Glu Lys Val Gly
275 280 285

Lys Ile Leu Leu Arg Ala Asn Gly Thr Pro Asn Pro Ala Ile Val Gly
290 295 300

Lys Asp Val Gln Ala Leu Ala Lys Leu Ala Gly Ile Ser Ile Pro Ser

<210>	26
<211>	1473
<212>	DNA
<213>	Artificial

```
<220>
<223> optimised sequence
```

<400> 26
atggaaatca tggacaagga tttgcaatcc atccaagaag ttagaacttt gattgccaag

```

gccaagaagg ctcaagctga attcaagaac ttttcccaag aagctgttaa caaggtcatc 120
gaaaagatcg ccaaggctac tgaagttgaa gctgtcaaat tggccaaatt ggcttacgaa 180
gacaccggtt acggtaaatg ggaagacaag gtcacaaaga acaaattctc ctccattggt 240
gtctacaact acatcaagga tttgaagacc gttggtatct tgaaggaaga caaggaaaag 300
aaattgattg acattgctgt cccattaggt gtcattgctg gtttgattcc atctaccaac 360
ccaacttcca ctgccatttt caaggtcttg attgctttga aggctggtaa cgccattgtc 420
ttctctccac acccaactgc tgtccgttcc atcactgaaa ccgttaagat catgcaaaaag 480
gctgctgttg aagctgggtg tccagatggg ttgatccaat gtatgtccat tttgaccgtt 540
gaaggtactg ctgaattgat gaagaacaag gacaccgctt tgatcttggc taccggtggg 600
gaaggtatgg ttagagctgc ttactcctct ggtactccag ccatcgggtg cgggtccagg 660
aacggtccat gtttcatcga aagaactgct gacattccaa ctgctgttag aaaggttatc 720
ggttctgaca ctttcgacaa cggtgtcatc tgtgcttctg aacaatccat cattgctgaa 780
accgtcaaga aggctgaaat catcgaagaa ttcaagagac aaaagggtta cttcttgaat 840
gctgaagaat ctgaaaaggg ttgtaagatt ctattacgtg ccaacggtag tccaaaccca 900
gccatcggtg gtaaggatgt ccaagctttg gccaaattgg ctggtatttc cattccatct 960
gatgctgtta tcttactatc cgaacaaacc gatgtttctc ctaaaaatcc atacgctaag 1020
gaaaaattgg ctccagtctt ggctttctac accgtcgaag actggcatga agcttgtgaa 1080
aagtctttgg ctttattgca caaccaaggg tctgggcaca ctttgatcat ccactctcaa 1140
aacgaagaaa tcattagaga atttgctttg aagaagcctg tttccagaat tttggttaac 1200
tctccagggt ctttggttgg tatcgggtgg gctaccaact tagtcccatc tttgacttta 1260
ggttgtgggt ctggttgggt ttctgccacc tctgacaacg ttggtccaga aaacttggtc 1320
aacatcagaa aggttgctta cgggtaccac accgtcgaag aaatcagaga agctttcggg 1380
gtcggtgctg cttcttcttc tgctccagct gaaccagaag acaacgaaga tgttcaagcc 1440
attgttaagg ccatcatggc caaattgaac ttg 1473

```

```

<210> 27
<211> 2610
<212> DNA
<213> Staphylococcus aureus

```

```
<220>
```

<221> CDS

<222> (1)..(2610)

<400> 27

atg tta act ata cct gaa aaa gaa aat cgt gga tcg aaa gaa caa gaa	48
Met Leu Thr Ile Pro Glu Lys Glu Asn Arg Gly Ser Lys Glu Gln Glu	
1 5 10 15	
gtg gca att atg att gat gct cta gct gac aaa ggg aaa gca tta	96
Val Ala Ile Met Ile Asp Ala Leu Ala Asp Lys Gly Lys Lys Ala Leu	
20 25 30	
gaa gca tta tct aaa aag tca caa gaa gaa att gat cat att gtt cat	144
Glu Ala Leu Ser Lys Lys Ser Gln Glu Glu Ile Asp His Ile Val His	
35 40 45	
caa atg agc tta gca gct gtt gat caa cat atg gtg cta gca aaa tta	192
Gln Met Ser Leu Ala Ala Val Asp Gln His Met Val Leu Ala Lys Leu	
50 55 60	
gca cat gaa gaa act gga aga ggt ata tac gaa gat aaa gcg att aaa	240
Ala His Glu Glu Thr Gly Arg Gly Ile Tyr Glu Asp Lys Ala Ile Lys	
65 70 75 80	
aat tta tac gct tct gaa tat ata tgg aat tca ata aaa gac aat aag	288
Asn Leu Tyr Ala Ser Glu Tyr Ile Trp Asn Ser Ile Lys Asp Asn Lys	
85 90 95	
aca gta ggg att att ggt gaa gat aaa gaa aaa gga tta acg tat gta	336
Thr Val Gly Ile Ile Gly Glu Asp Lys Glu Lys Gly Leu Thr Tyr Val	
100 105 110	
gcg gaa cca att ggt gtt att tgt ggt gtt acg cca aca aca aat cct	384
Ala Glu Pro Ile Gly Val Ile Cys Gly Val Thr Pro Thr Thr Asn Pro	
115 120 125	
acg tcg aca act att ttt aaa gcg atg att gca att aag aca gga aat	432
Thr Ser Thr Thr Ile Phe Lys Ala Met Ile Ala Ile Lys Thr Gly Asn	
130 135 140	
cca atc att ttt gca ttc cat cca agt gca caa gaa tcg tcg aag cgt	480
Pro Ile Ile Phe Ala Phe His Pro Ser Ala Gln Glu Ser Ser Lys Arg	
145 150 155 160	
gca gca gaa gtt gta tta gaa gcg gca atg aag gca ggt gca cct aaa	528
Ala Ala Glu Val Val Leu Glu Ala Ala Met Lys Ala Gly Ala Pro Lys	
165 170 175	
gat att att cag tgg att gaa gtg cct tct atc gaa gca aca aaa caa	576
Asp Ile Ile Gln Trp Ile Glu Val Pro Ser Ile Glu Ala Thr Lys Gln	
180 185 190	
tta atg aat cac aaa ggt att gca tta gtt cta gca aca ggt ggt tcg	624
Leu Met Asn His Lys Gly Ile Ala Leu Val Leu Ala Thr Gly Gly Ser	
195 200 205	
ggc atg gtt aag tct gca tat tca act ggc aaa ccg gca tta ggt gtg	672

Gly	Met	Val	Lys	Ser	Ala	Tyr	Ser	Thr	Gly	Lys	Pro	Ala	Leu	Gly	Val	
	210					215					220					
gga	cca	ggt	aac	gtg	ccg	tct	tac	att	gaa	aaa	aca	gca	cac	att	aaa	720
Gly	Pro	Gly	Asn	Val	Pro	Ser	Tyr	Ile	Glu	Lys	Thr	Ala	His	Ile	Lys	
225					230					235					240	
cgt	gca	gta	aat	gat	atc	att	ggt	tca	aaa	aca	ttt	gat	aat	ggt	atg	768
Arg	Ala	Val	Asn	Asp	Ile	Ile	Gly	Ser	Lys	Thr	Phe	Asp	Asn	Gly	Met	
				245					250					255		
att	tgt	gct	tct	gaa	caa	ggt	gta	gtc	att	gat	aaa	gaa	att	tat	aaa	816
Ile	Cys	Ala	Ser	Glu	Gln	Val	Val	Val	Ile	Asp	Lys	Glu	Ile	Tyr	Lys	
			260					265						270		
gat	ggt	act	aat	gaa	ttt	aaa	gca	cat	caa	gca	tac	ttt	ggt	aaa	aaa	864
Asp	Val	Thr	Asn	Glu	Phe	Lys	Ala	His	Gln	Ala	Tyr	Phe	Val	Lys	Lys	
		275					280					285				
gat	gaa	tta	caa	cgc	tta	gaa	aat	gca	att	atg	aat	gaa	caa	aaa	aca	912
Asp	Glu	Leu	Gln	Arg	Leu	Glu	Asn	Ala	Ile	Met	Asn	Glu	Gln	Lys	Thr	
	290					295					300					
ggt	att	aag	cct	gat	att	gtc	ggt	aaa	tct	gca	ggt	gaa	ata	gct	gaa	960
Gly	Ile	Lys	Pro	Asp	Ile	Val	Gly	Lys	Ser	Ala	Val	Glu	Ile	Ala	Glu	
305					310					315					320	
tta	gca	ggt	ata	cct	gtc	ccc	gaa	aat	aca	aaa	ctt	atc	ata	gcc	gaa	1008
Leu	Ala	Gly	Ile	Pro	Val	Pro	Glu	Asn	Thr	Lys	Leu	Ile	Ile	Ala	Glu	
				325					330					335		
att	agc	ggt	gta	ggt	tca	gac	tat	ccg	tta	tct	cgt	gaa	aaa	tta	tct	1056
Ile	Ser	Gly	Val	Gly	Ser	Asp	Tyr	Pro	Leu	Ser	Arg	Glu	Lys	Leu	Ser	
			340					345					350			
cca	gta	tta	gcc	tta	gta	aaa	gcc	caa	tct	aca	aaa	caa	gca	ttt	caa	1104
Pro	Val	Leu	Ala	Leu	Val	Lys	Ala	Gln	Ser	Thr	Lys	Gln	Ala	Phe	Gln	
		355					360					365				
att	tgt	gaa	gac	aca	cta	cat	ttt	ggt	gga	tta	gga	cac	aca	gcc	ggt	1152
Ile	Cys	Glu	Asp	Thr	Leu	His	Phe	Gly	Gly	Leu	Gly	His	Thr	Ala	Val	
	370					375					380					
atc	cat	aca	gaa	gat	gaa	aca	tta	caa	aaa	gat	ttt	gga	cta	aga	atg	1200
Ile	His	Thr	Glu	Asp	Glu	Thr	Leu	Gln	Lys	Asp	Phe	Gly	Leu	Arg	Met	
	385				390					395					400	
aaa	gct	tgt	cgt	gta	ctt	gta	aat	aca	cca	tca	gcg	ggt	gga	ggt	att	1248
Lys	Ala	Cys	Arg	Val	Leu	Val	Asn	Thr	Pro	Ser	Ala	Val	Gly	Gly	Ile	
				405					410					415		
ggt	gat	atg	tat	aac	gaa	ttg	att	ccg	tct	tta	aca	tta	ggt	tgt	ggt	1296
Gly	Asp	Met	Tyr	Asn	Glu	Leu	Ile	Pro	Ser	Leu	Thr	Leu	Gly	Cys	Gly	
			420					425					430			
tcg	tac	ggt	aga	aac	tca	att	tca	cat	aat	ggt	agt	gcg	aca	gat	tta	1344
Ser	Tyr	Gly	Arg	Asn	Ser	Ile	Ser	His	Asn	Val	Ser	Ala	Thr	Asp	Leu	

435	440	445	
tta aac att aaa acg att gct aaa cga cgt aat aat act caa att ttc Leu Asn Ile Lys Thr Ile Ala Lys Arg Arg Asn Asn Thr Gln Ile Phe 450 455 460			1392
aag gtg cct gct caa att tat ttt gaa gaa aat gca atc atg agt cta Lys Val Pro Ala Gln Ile Tyr Phe Glu Glu Asn Ala Ile Met Ser Leu 465 470 475 480			1440
aca aca atg gac aag att gaa aaa gtg atg att gtc tgt gac cct ggt Thr Thr Met Asp Lys Ile Glu Lys Val Met Ile Val Cys Asp Pro Gly 485 490 495			1488
atg gta gaa ttc ggt tat aca aaa aca gtt gag aat gta tta aga caa Met Val Glu Phe Gly Tyr Thr Lys Thr Val Glu Asn Val Leu Arg Gln 500 505 510			1536
aga acg gaa cag cct caa att aaa ata ttt agc gaa gtc gaa ccg aac Arg Thr Glu Gln Pro Gln Ile Lys Ile Phe Ser Glu Val Glu Pro Asn 515 520 525			1584
cca tca act aat aca gta tat aaa ggt ctg gaa atg atg gtt gat ttc Pro Ser Thr Asn Thr Val Tyr Lys Gly Leu Glu Met Met Val Asp Phe 530 535 540			1632
caa cca gat aca atc att gca ctt ggt ggt ggt tca gcg atg gat gct Gln Pro Asp Thr Ile Ile Ala Leu Gly Gly Gly Ser Ala Met Asp Ala 545 550 555 560			1680
gca aaa gca atg tgg atg ttc ttt gaa cac cct gag aca tca ttc ttc Ala Lys Ala Met Trp Met Phe Phe Glu His Pro Glu Thr Ser Phe Phe 565 570 575			1728
ggg gct aaa caa aag ttc cta gac atc ggt aaa cgt act tat aaa ata Gly Ala Lys Gln Lys Phe Leu Asp Ile Gly Lys Arg Thr Tyr Lys Ile 580 585 590			1776
ggc atg cct gaa aat gcg acg ttc att tgt atc cct acg aca tca ggt Gly Met Pro Glu Asn Ala Thr Phe Ile Cys Ile Pro Thr Thr Ser Gly 595 600 605			1824
aca ggt tca gaa gta aca cca ttt gca gtt atc aca gat agt gaa aca Thr Gly Ser Glu Val Thr Pro Phe Ala Val Ile Thr Asp Ser Glu Thr 610 615 620			1872
aat gta aaa tat ccg ttg gct gat ttt gct tta aca cct gac gtt gca Asn Val Lys Tyr Pro Leu Ala Asp Phe Ala Leu Thr Pro Asp Val Ala 625 630 635 640			1920
att att gac cct caa ttt gtg atg agt gtg cca aaa agc gtt aca gca Ile Ile Asp Pro Gln Phe Val Met Ser Val Pro Lys Ser Val Thr Ala 645 650 655			1968
gat aca gga atg gat gta cta acg cat gca atg gaa tca tat gta tct Asp Thr Gly Met Asp Val Leu Thr His Ala Met Glu Ser Tyr Val Ser 660 665 670			2016

gta atg gct tca gac tat aca aga ggt ttg agt cta caa gcg att aaa	2064
Val Met Ala Ser Asp Tyr Thr Arg Gly Leu Ser Leu Gln Ala Ile Lys	
675 680 685	
ttg acg ttc gaa tat tta aaa tca tct gtt gaa aag ggt gat aaa gtt	2112
Leu Thr Phe Glu Tyr Leu Lys Ser Ser Val Glu Lys Gly Asp Lys Val	
690 695 700	
tca aga gag aaa atg cat aac gca tca act ttg gct ggt atg gca ttt	2160
Ser Arg Glu Lys Met His Asn Ala Ser Thr Leu Ala Gly Met Ala Phe	
705 710 715 720	
gca aat gca ttc tta ggc att gca cac tca att gca cat aaa att ggt	2208
Ala Asn Ala Phe Leu Gly Ile Ala His Ser Ile Ala His Lys Ile Gly	
725 730 735	
ggc gaa tat ggt att ccg cat ggt aga gcg aat gcg ata tta cta ccg	2256
Gly Glu Tyr Gly Ile Pro His Gly Arg Ala Asn Ala Ile Leu Leu Pro	
740 745 750	
cat att atc cgt tat aat gcc aaa gac ccg caa aaa cat gca tta ttc	2304
His Ile Ile Arg Tyr Asn Ala Lys Asp Pro Gln Lys His Ala Leu Phe	
755 760 765	
cct aaa tat gag ttc ttc aga gca gat aca gat tat gca gat att gcc	2352
Pro Lys Tyr Glu Phe Phe Arg Ala Asp Thr Asp Tyr Ala Asp Ile Ala	
770 775 780	
aaa ttc tta gga tta aaa ggg aat acg aca gaa gca ctc gta gaa tca	2400
Lys Phe Leu Gly Leu Lys Gly Asn Thr Thr Glu Ala Leu Val Glu Ser	
785 790 795 800	
tta gct aaa gct gtc tac gaa tta ggt caa tca gtc gga att gaa atg	2448
Leu Ala Lys Ala Val Tyr Glu Leu Gly Gln Ser Val Gly Ile Glu Met	
805 810 815	
aat ttg aaa tca caa ggt gtg tct gaa gaa gaa tta aat gaa tca att	2496
Asn Leu Lys Ser Gln Gly Val Ser Glu Glu Glu Leu Asn Glu Ser Ile	
820 825 830	
gat aga atg gca gag ctc gca ttt gaa gat caa tgt aca act gct aat	2544
Asp Arg Met Ala Glu Leu Ala Phe Glu Asp Gln Cys Thr Thr Ala Asn	
835 840 845	
cct aaa gaa gca cta atc agt gaa atc aaa gat atc att caa aca tca	2592
Pro Lys Glu Ala Leu Ile Ser Glu Ile Lys Asp Ile Ile Gln Thr Ser	
850 855 860	
tat gat tat aag caa taa	2610
Tyr Asp Tyr Lys Gln	
865	

<210> 28
 <211> 869
 <212> PRT

<213> Staphylococcus aureus

<400> 28

Met Leu Thr Ile Pro Glu Lys Glu Asn Arg Gly Ser Lys Glu Gln Glu
1 5 10 15

Val Ala Ile Met Ile Asp Ala Leu Ala Asp Lys Gly Lys Lys Ala Leu
20 25 30

Glu Ala Leu Ser Lys Lys Ser Gln Glu Glu Ile Asp His Ile Val His
35 40 45

Gln Met Ser Leu Ala Ala Val Asp Gln His Met Val Leu Ala Lys Leu
50 55 60

Ala His Glu Glu Thr Gly Arg Gly Ile Tyr Glu Asp Lys Ala Ile Lys
65 70 75 80

Asn Leu Tyr Ala Ser Glu Tyr Ile Trp Asn Ser Ile Lys Asp Asn Lys
85 90 95

Thr Val Gly Ile Ile Gly Glu Asp Lys Glu Lys Gly Leu Thr Tyr Val
100 105 110

Ala Glu Pro Ile Gly Val Ile Cys Gly Val Thr Pro Thr Thr Asn Pro
115 120 125

Thr Ser Thr Thr Ile Phe Lys Ala Met Ile Ala Ile Lys Thr Gly Asn
130 135 140

Pro Ile Ile Phe Ala Phe His Pro Ser Ala Gln Glu Ser Ser Lys Arg
145 150 155 160

Ala Ala Glu Val Val Leu Glu Ala Ala Met Lys Ala Gly Ala Pro Lys
165 170 175

Asp Ile Ile Gln Trp Ile Glu Val Pro Ser Ile Glu Ala Thr Lys Gln
180 185 190

Leu Met Asn His Lys Gly Ile Ala Leu Val Leu Ala Thr Gly Gly Ser
195 200 205

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

210		215		220
Gly 225	Pro	Gly	Asn	Val
				Pro 230
	Ser	Tyr	Ile	Glu
				Lys 235
	Thr	Ala	His	Ile
				Lys 240
Arg	Ala	Val	Asn	Asp
				Ile 245
	Ile	Ile	Gly	Ser
				Lys 250
	Thr	Phe	Asp	Asn
				Gly 255
Ile	Cys	Ala	Ser	Glu
				Gln 260
	Val	Val	Val	Ile
				Asp 265
	Lys	Glu	Ile	Tyr
				Lys 270
Asp	Val	Thr	Asn	Glu
				Phe 275
	Lys	Ala	His	Gln
				Ala 280
	Tyr	Phe	Val	Lys
				Lys 285
Asp	Glu	Leu	Gln	Arg
				Leu 290
	Glu	Asn	Ala	Ile
				Met 295
	Asn	Glu	Gln	Lys
				Thr 300
Gly	Ile	Lys	Pro	Asp
				Ile 305
	Val	Gly	Lys	Ser
				Ala 310
	Val	Glu	Ile	Ala
				Glu 315
	Glu	Ile	Ala	Glu
				320
Leu	Ala	Gly	Ile	Pro
				Val 325
	Pro	Glu	Asn	Thr
				Lys 330
	Leu	Ile	Ile	Ala
				Glu 335
Ile	Ser	Gly	Val	Gly
				Ser 340
	Asp	Tyr	Pro	Leu
				Ser 345
	Arg	Glu	Lys	Leu
				Ser 350
Pro	Val	Leu	Ala	Leu
				Val 355
	Lys	Ala	Gln	Ser
				Thr 360
	Lys	Gln	Ala	Phe
				Gln 365
Ile	Cys	Glu	Asp	Thr
				Leu 370
	His	Phe	Gly	Gly
				Leu 375
	Gly	His	Thr	Ala
				Val 380
Ile	His	Thr	Glu	Asp
				Glu 385
	Thr	Leu	Gln	Lys
				Asp 390
	Phe	Gly	Leu	Arg
				Met 395
	Gly	Leu	Arg	Met
				400
Lys	Ala	Cys	Arg	Val
				Leu 405
	Leu	Val	Asn	Thr
				Pro 410
	Ser	Ala	Val	Gly
				Gly 415
Gly	Asp	Met	Tyr	Asn
				Glu 420
	Leu	Ile	Pro	Ser
				Leu 425
	Thr	Leu	Gly	Cys
				Gly 430
Ser	Tyr	Gly	Arg	Asn
				Ser 435
	Ile	Ser	His	Asn
				Val 440
	Ser	Ala	Thr	Asp
				Leu 445

Leu Asn Ile Lys Thr Ile Ala Lys Arg Arg Asn Asn Thr Gln Ile Phe
 450 455 460

Lys Val Pro Ala Gln Ile Tyr Phe Glu Glu Asn Ala Ile Met Ser Leu
 465 470 475 480

Thr Thr Met Asp Lys Ile Glu Lys Val Met Ile Val Cys Asp Pro Gly
 485 490 495

Met Val Glu Phe Gly Tyr Thr Lys Thr Val Glu Asn Val Leu Arg Gln
 500 505 510

Arg Thr Glu Gln Pro Gln Ile Lys Ile Phe Ser Glu Val Glu Pro Asn
 515 520 525

Pro Ser Thr Asn Thr Val Tyr Lys Gly Leu Glu Met Met Val Asp Phe
 530 535 540

Gln Pro Asp Thr Ile Ile Ala Leu Gly Gly Gly Ser Ala Met Asp Ala
 545 550 555 560

Ala Lys Ala Met Trp Met Phe Phe Glu His Pro Glu Thr Ser Phe Phe
 565 570 575

Gly Ala Lys Gln Lys Phe Leu Asp Ile Gly Lys Arg Thr Tyr Lys Ile
 580 585 590

Gly Met Pro Glu Asn Ala Thr Phe Ile Cys Ile Pro Thr Thr Ser Gly
 595 600 605

Thr Gly Ser Glu Val Thr Pro Phe Ala Val Ile Thr Asp Ser Glu Thr
 610 615 620

Asn Val Lys Tyr Pro Leu Ala Asp Phe Ala Leu Thr Pro Asp Val Ala
 625 630 635 640

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

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

Val Met Ala Ser Asp Tyr Thr Arg Gly Leu Ser Leu Gln Ala Ile Lys
675 680 685

Leu Thr Phe Glu Tyr Leu Lys Ser Ser Val Glu Lys Gly Asp Lys Val
690 695 700

Ser Arg Glu Lys Met His Asn Ala Ser Thr Leu Ala Gly Met Ala Phe
705 710 715 720

Ala Asn Ala Phe Leu Gly Ile Ala His Ser Ile Ala His Lys Ile Gly
725 730 735

Gly Glu Tyr Gly Ile Pro His Gly Arg Ala Asn Ala Ile Leu Leu Pro
740 745 750

His Ile Ile Arg Tyr Asn Ala Lys Asp Pro Gln Lys His Ala Leu Phe
755 760 765

Pro Lys Tyr Glu Phe Phe Arg Ala Asp Thr Asp Tyr Ala Asp Ile Ala
770 775 780

Lys Phe Leu Gly Leu Lys Gly Asn Thr Thr Glu Ala Leu Val Glu Ser
785 790 795 800

Leu Ala Lys Ala Val Tyr Glu Leu Gly Gln Ser Val Gly Ile Glu Met
805 810 815

Asn Leu Lys Ser Gln Gly Val Ser Glu Glu Glu Leu Asn Glu Ser Ile
820 825 830

Asp Arg Met Ala Glu Leu Ala Phe Glu Asp Gln Cys Thr Thr Ala Asn
835 840 845

Pro Lys Glu Ala Leu Ile Ser Glu Ile Lys Asp Ile Ile Gln Thr Ser
850 855 860

Tyr Asp Tyr Lys Gln
865

<210> 29
<211> 2607
<212> DNA
<213> Artificial

<220>

<223> optimised sequence

<400> 29

atgttgacca ttccagaaaa ggaaaacaga ggttccaagg aacaagaagt tgccatcatg	60
attgatgctt tagctgacaa aggtaagaag gctttggaag ctttgtccaa gaagtctcaa	120
gaagaaattg accacattgt ccaccaaattg tccttggctg ctggtgacca acacatggtt	180
ttggccaagt tggctcatga agaaaccggt agaggatatct acgaagacaa ggctatcaag	240
aacttatacg cctctgaata catctggaac tccatcaagg acaacaagac tgttgggtatc	300
attggtgaag acaaagaaaa gggtttgacc tacgttgctg aaccaattgg tgtcatctgt	360
ggtgtcactc caaccaccaa cccaacttct accaccatct tcaaggctat gattgccatc	420
aagactggta acccaattat tttcgctttc caccatctg ctcaagaatc ttccaagaga	480
gctgctgaag ttgttttgga agctgccatg aaggctgggtg ctccaaagga tatcatccaa	540
tggattgaag ttccatccat tgaagctacc aagcaattga tgaaccacaa gggatttgct	600
ttagtcttgg ctaccggtgg ttctgggtatg gttaagtctg cttactccac tggtaaacca	660
gctttgggtg ttggtccagg taacgttcca tcttacatcg aaaagactgc tcatatcaag	720
cgtgctgtca acgatatcat cggttccaag actttcgata atgggatgat ctgtgcttct	780
gaacaagttg ttgtcattga caaggaaatc tacaaggatg tcaccaatga attcaaggct	840
caccaagctt acttcgtcaa gaaggacgaa ttacaaagat tagaaaacgc catcatgaac	900
gaacaaaaga ctggtatcaa gccagatatc gttggtaagt ctgctgttga aattgctgaa	960
ttggccggtg tcccagttcc agaaaacacc aaattgatca ttgctgaaat ctccggtgtc	1020
ggttctgact acccattgtc cagagaaaag ttgtctccag ttttggcttt agtcaaggct	1080
caatctacca agcaagcttt ccaaactctgt gaagacactt tgcacttcgg tggtttaggt	1140
cacactgctg ttatccacac tgaagacgaa actttgcaaa aggatttcgg tctaagaatg	1200
aaggcttgct gtgttttggt caacactcca tctgctgttg gtggtatcgg tgacatgtac	1260
aacgaattga ttccatcctt gactttgggt tgtgggttctt acggtagaaa ctccatctcc	1320
cacaacgtct ctgctaccga tttggtgaac atcaagacca ttgccaagag aagaaacaac	1380
actcaaactc tcaaggttcc agctcaaatc tatttcgaag aaaacgctat catgtccttg	1440
accaccatgg acaagattga aaaggatcatg atcggttggtg acccaggtat ggttgaattt	1500
ggttacacca aaaccgtcga aaacgtctta cgtcaaagaa ctgaacaacc tcaaatcaag	1560

```

atcttctctg aagttgaacc aaatccatcc accaactctg tctacaaggg tttggaaatg 1620
atggtcgatt tccaaccaga caccatcatt gctttgggtg gtggttctgc catggatgct 1680
gccaaaggcta tgtggatggt cttcgaacat ccagaaactt ctttcttcgg tgccaagcaa 1740
aaattcttgg acattggtaa gagaacctac aagattggta tgccagaaaa cgccactttc 1800
atctgtattc caaccacttc tgggtactgg tctgaagtca ctccatttgc tgttatcact 1860
gactctgaaa ccaacgtcaa ataccattg gctgatttgc ctttgactcc agatgtcgcc 1920
atcattgacc ctcaatttgt catgtccgtc caaaatctg tctactgctga taccggtatg 1980
gacgttttga ctacgctat ggaatcttac gtttctgtca tggcctccga ttacaccaga 2040
ggtttgtccc tacaagctat caaattgacc tttgaatact tgaaatcttc cgttgaaaaa 2100
ggtgacaagg tttccagaga aaagatgcac aacgcttcta ctttggccgg tatggccttt 2160
gctaacgctt tcttgggtat tgctcactcc attgctcaca aaattgggtg tgaatacggc 2220
attccacatg gttagagctaa cgccatcttg ttgcctcaca tcatcagata caacgccaaag 2280
gacctcaaaa agcacgcttt gtcccaaag tacgaattct tcagagctga caccgattac 2340
gctgatatcg ccaagttctt aggtttgaaa ggtaacacca ctgaagcttt ggttgaatct 2400
ttggccaagg ctgtctacga attaggtcaa tctgttggtg ttgaaatgaa cttgaaatct 2460
caaggtgtct ctgaagaaga attgaacgaa tccattgaca gaatggctga attggctttc 2520
gaagaccaat gtaccactgc caacccaaag gaagctttga tttctgaaat caaggatatc 2580
atccaaactt cttacgacta caagcag 2607

```

```

<210> 30
<211> 392
<212> PRT
<213> Clostridium acetobutylicum

```

```
<400> 30
```

```

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

```

```

Tyr Gly Lys Ser Leu Lys Asp Val Pro Ala Val Asp Leu Gly Ala Thr
20           25           30

```

```

Ala Ile Lys Glu Ala Val Lys Lys Ala Gly Ile Lys Pro Glu Asp Val
35           40           45

```

```

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

```

50		55		60											
Pro 65	Ala	Arg	Gln	Ala	Ser 70	Phe	Lys	Ala	Gly	Leu 75	Pro	Val	Glu	Ile	Pro 80
Ala	Met	Thr	Ile	Asn 85	Lys	Val	Cys	Gly	Ser 90	Gly	Leu	Arg	Thr	Val 95	Ser
Leu	Ala	Ala	Gln	Ile 100	Ile	Lys	Ala	Gly	Asp 105	Ala	Asp	Val	Ile 110	Ile	Ala
Gly	Gly	Met 115	Glu	Asn	Met	Ser	Arg	Ala	Pro	Tyr	Leu	Ala 125	Asn	Asn	Ala
Arg	Trp 130	Gly	Tyr	Arg	Met	Gly	Asn	Ala	Lys	Phe	Val	Asp 140	Glu	Met	Ile
Thr 145	Asp	Gly	Leu	Trp	Asp 150	Ala	Phe	Asn	Asp	Tyr	His	Met	Gly	Ile	Thr 160
Ala	Glu	Asn	Ile 165	Ala	Glu	Arg	Trp	Asn	Ile 170	Ser	Arg	Glu	Glu	Gln 175	Asp
Glu	Phe	Ala	Leu 180	Ala	Ser	Gln	Lys	Lys	Ala	Glu	Glu	Ala	Ile 190	Lys	Ser
Gly	Gln	Phe	Lys	Asp 195	Glu	Ile	Val	Pro	Val	Val	Ile	Lys	Gly	Arg	Lys
Gly	Glu	Thr	Val	Val	Asp 210	Thr	Asp	Glu	His	Pro	Arg	Phe	Gly	Ser	Thr
Ile 225	Glu	Gly	Leu	Ala	Lys 230	Leu	Lys	Pro	Ala	Phe	Lys	Lys	Asp	Gly	Thr 240
Val	Thr	Ala	Gly	Asn 245	Ala	Ser	Gly	Leu	Asn	Asp	Cys	Ala	Ala	Val	Leu
Val	Ile	Met	Ser	Ala	Glu	Lys	Ala	Lys	Glu	Leu	Gly	Val	Lys	Pro	Leu
Ala	Lys	Ile	Val	Ser	Tyr	Gly	Ser	Ala	Gly	Val	Asp	Pro	Ala	Ile	Met
		275					280					285			

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

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

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

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

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

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

Thr Ala Ile Leu Leu Glu Lys Cys
 385 390

<210> 31
 <211> 282
 <212> PRT
 <213> Clostridium acetobutylicum

<400> 31

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

<210> 32
<211> 261
<212> PRT
<213> Clostridium acetobutylicum

<400> 32

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

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

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

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

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

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

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

Cys Glu Ile Ala Met Ser Cys Asp Ile Arg Ile Ala Ser Ser Asn Ala
 115 120 125

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

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

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

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

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

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

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

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

Gly Phe Lys Asn Arg
260

<210> 33
<211> 379
<212> PRT
<213> Clostridium acetobutylicum

<400> 33

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

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

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

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

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

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

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

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

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

Gly Asp His Tyr Val Ile Asn Gly Ser Lys Ile Phe Ile Thr Asn Gly

<210> 34
 <211> 858
 <212> PRT
 <213> Clostridium acetobutylicum

<400> 34

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Leu Gly Cys Gly Thr Trp Gly Gly Asn Ser Val Ser Gln Asn Val Glu

420	425	430																	
Pro	Lys	His	Leu	Leu	Asn	Ile	Lys	Ser	Val	Ala	Glu	Arg	Arg	Glu	Asn				
	435						440					445							
Met	Leu	Trp	Phe	Lys	Val	Pro	Gln	Lys	Ile	Tyr	Phe	Lys	Tyr	Gly	Cys				
	450					455					460								
Leu	Arg	Phe	Ala	Leu	Lys	Glu	Leu	Lys	Asp	Met	Asn	Lys	Lys	Arg	Ala				
465					470					475					480				
Phe	Ile	Val	Thr	Asp	Lys	Asp	Leu	Phe	Lys	Leu	Gly	Tyr	Val	Asn	Lys				
				485					490					495					
Ile	Thr	Lys	Val	Leu	Asp	Glu	Ile	Asp	Ile	Lys	Tyr	Ser	Ile	Phe	Thr				
			500					505					510						
Asp	Ile	Lys	Ser	Asp	Pro	Thr	Ile	Asp	Ser	Val	Lys	Lys	Gly	Ala	Lys				
	515						520					525							
Glu	Met	Leu	Asn	Phe	Glu	Pro	Asp	Thr	Ile	Ile	Ser	Ile	Gly	Gly	Gly				
	530					535					540								
Ser	Pro	Met	Asp	Ala	Ala	Lys	Val	Met	His	Leu	Leu	Tyr	Glu	Tyr	Pro				
545					550					555					560				
Glu	Ala	Glu	Ile	Glu	Asn	Leu	Ala	Ile	Asn	Phe	Met	Asp	Ile	Arg	Lys				
				565					570					575					
Arg	Ile	Cys	Asn	Phe	Pro	Lys	Leu	Gly	Thr	Lys	Ala	Ile	Ser	Val	Ala				
			580					585					590						
Ile	Pro	Thr	Thr	Ala	Gly	Thr	Gly	Ser	Glu	Ala	Thr	Pro	Phe	Ala	Val				
	595						600					605							
Ile	Thr	Asn	Asp	Glu	Thr	Gly	Met	Lys	Tyr	Pro	Leu	Thr	Ser	Tyr	Glu				
	610					615					620								
Leu	Thr	Pro	Asn	Met	Ala	Ile	Ile	Asp	Thr	Glu	Leu	Met	Leu	Asn	Met				
625					630					635					640				
Pro	Arg	Lys	Leu	Thr	Ala	Ala	Thr	Gly	Ile	Asp	Ala	Leu	Val	His	Ala				
				645					650					655					

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

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

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

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

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

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

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

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

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

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

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

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

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

<210> 35
<211> 862
<212> PRT

<213> Clostridium acetobutylicum

<400> 35

Met Lys Val Thr Thr Val Lys Glu Leu Asp Glu Lys Leu Lys Val Ile
1 5 10 15

Lys Glu Ala Gln Lys Lys Phe Ser Cys Tyr Ser Gln Glu Met Val Asp
20 25 30

Glu Ile Phe Arg Asn Ala Ala Met Ala Ala Ile Asp Ala Arg Ile Glu
35 40 45

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

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

Lys Asp Glu Lys Thr Cys Gly Ile Ile Glu Arg Asn Glu Pro Tyr Gly
85 90 95

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

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

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

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

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

Leu Thr Gln Tyr Leu Met Gln Lys Ala Asp Ile Thr Leu Ala Thr Gly
180 185 190

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

Gly Val Gly Pro Gly Asn Thr Pro Val Ile Ile Asp Glu Ser Ala His

210		215		220
Ile Lys Met Ala Val Ser Ser Ile Ile Leu Ser Lys Thr Tyr Asp Asn				
225		230		235 240
Gly Val Ile Cys Ala Ser Glu Gln Ser Val Ile Val Leu Lys Ser Ile				
		245		250 255
Tyr Asn Lys Val Lys Asp Glu Phe Gln Glu Arg Gly Ala Tyr Ile Ile				
		260		265 270
Lys Lys Asn Glu Leu Asp Lys Val Arg Glu Val Ile Phe Lys Asp Gly				
		275		280 285
Ser Val Asn Pro Lys Ile Val Gly Gln Ser Ala Tyr Thr Ile Ala Ala				
		290		295 300
Met Ala Gly Ile Lys Val Pro Lys Thr Thr Arg Ile Leu Ile Gly Glu				
305		310		315 320
Val Thr Ser Leu Gly Glu Glu Glu Pro Phe Ala His Glu Lys Leu Ser				
		325		330 335
Pro Val Leu Ala Met Tyr Glu Ala Asp Asn Phe Asp Asp Ala Leu Lys				
		340		345 350
Lys Ala Val Thr Leu Ile Asn Leu Gly Gly Leu Gly His Thr Ser Gly				
		355		360 365
Ile Tyr Ala Asp Glu Ile Lys Ala Arg Asp Lys Ile Asp Arg Phe Ser				
		370		375 380
Ser Ala Met Lys Thr Val Arg Thr Phe Val Asn Ile Pro Thr Ser Gln				
385		390		395 400
Gly Ala Ser Gly Asp Leu Tyr Asn Phe Arg Ile Pro Pro Ser Phe Thr				
		405		410 415
Leu Gly Cys Gly Phe Trp Gly Gly Asn Ser Val Ser Glu Asn Val Gly				
		420		425 430
Pro Lys His Leu Leu Asn Ile Lys Thr Val Ala Glu Arg Arg Glu Asn				
		435		440 445

Met Leu Trp Phe Arg Val Pro His Lys Val Tyr Phe Lys Phe Gly Cys
 450 455 460

Leu Gln Phe Ala Leu Lys Asp Leu Lys Asp Leu Lys Lys Lys Arg Ala
 465 470 475 480

Phe Ile Val Thr Asp Ser Asp Pro Tyr Asn Leu Asn Tyr Val Asp Ser
 485 490 495

Ile Ile Lys Ile Leu Glu His Leu Asp Ile Asp Phe Lys Val Phe Asn
 500 505 510

Lys Val Gly Arg Glu Ala Asp Leu Lys Thr Ile Lys Lys Ala Thr Glu
 515 520 525

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

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

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

Arg Ile Tyr Thr Phe Pro Lys Leu Gly Lys Lys Ala Met Leu Val Ala
 580 585 590

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

Val Thr Asp Asn Asn Thr Gly Asn Lys Tyr Met Leu Ala Asp Tyr Glu
 610 615 620

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

Pro Lys Gly Leu Thr Ala Tyr Ser Gly Ile Asp Ala Leu Val Asn Ser
 645 650 655

Ile Glu Ala Tyr Thr Ser Val Tyr Ala Ser Glu Tyr Thr Asn Gly Leu
 660 665 670

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

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

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

His Ser Met Ala Ile Lys Leu Ser Ser Glu His Asn Ile Pro Ser Gly
725 730 735

Ile Ala Asn Ala Leu Leu Ile Glu Glu Val Ile Lys Phe Asn Ala Val
740 745 750

Asp Asn Pro Val Lys Gln Ala Pro Cys Pro Gln Tyr Lys Tyr Pro Asn
755 760 765

Thr Ile Phe Arg Tyr Ala Arg Ile Ala Asp Tyr Ile Lys Leu Gly Gly
770 775 780

Asn Thr Asp Glu Glu Lys Val Asp Leu Leu Ile Asn Lys Ile His Glu
785 790 795 800

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

Leu Glu Glu Asn Phe Tyr Ser Ser Leu Asp Arg Ile Ser Glu Leu Ala
820 825 830

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

Glu Ile Lys Glu Met Tyr Ile Asn Cys Phe Lys Lys Gln Pro
850 855 860

<210> 36
<211> 389
<212> PRT
<213> Clostridium acetobutylicum

<400> 36

Met Leu Ser Phe Asp Tyr Ser Ile Pro Thr Lys Val Phe Phe Gly Lys

1		5		10		15													
Gly	Lys	Ile	Asp	Val	Ile	Gly	Glu	Glu	Ile	Lys	Lys	Tyr	Gly	Ser	Arg				
			20					25					30						
Val	Leu	Ile	Val	Tyr	Gly	Gly	Gly	Ser	Ile	Lys	Arg	Asn	Gly	Ile	Tyr				
		35					40					45							
Asp	Arg	Ala	Thr	Ala	Ile	Leu	Lys	Glu	Asn	Asn	Ile	Ala	Phe	Tyr	Glu				
	50					55					60								
Leu	Ser	Gly	Val	Glu	Pro	Asn	Pro	Arg	Ile	Thr	Thr	Val	Lys	Lys	Gly				
65					70					75					80				
Ile	Glu	Ile	Cys	Arg	Glu	Asn	Asn	Val	Asp	Leu	Val	Leu	Ala	Ile	Gly				
				85					90					95					
Gly	Gly	Ser	Ala	Ile	Asp	Cys	Ser	Lys	Val	Ile	Ala	Ala	Gly	Val	Tyr				
			100					105					110						
Tyr	Asp	Gly	Asp	Thr	Trp	Asp	Met	Val	Lys	Asp	Pro	Ser	Lys	Ile	Thr				
		115					120					125							
Lys	Val	Leu	Pro	Ile	Ala	Ser	Ile	Leu	Thr	Leu	Ser	Ala	Thr	Gly	Ser				
	130					135					140								
Glu	Met	Asp	Gln	Ile	Ala	Val	Ile	Ser	Asn	Met	Glu	Thr	Asn	Glu	Lys				
145					150				155						160				
Leu	Gly	Val	Gly	His	Asp	Asp	Met	Arg	Pro	Lys	Phe	Ser	Val	Leu	Asp				
				165				170						175					
Pro	Thr	Tyr	Thr	Phe	Thr	Val	Pro	Lys	Asn	Gln	Thr	Ala	Ala	Gly	Thr				
			180					185					190						
Ala	Asp	Ile	Met	Ser	His	Thr	Phe	Glu	Ser	Tyr	Phe	Ser	Gly	Val	Glu				
		195					200					205							
Gly	Ala	Tyr	Val	Gln	Asp	Gly	Ile	Ala	Glu	Ala	Ile	Leu	Arg	Thr	Cys				
	210					215					220								
Ile	Lys	Tyr	Gly	Lys	Ile	Ala	Met	Glu	Lys	Thr	Asp	Asp	Tyr	Glu	Ala				
225					230					235					240				

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

Ser Leu Gly Lys Asp Arg Lys Trp Ser Cys His Pro Met Glu His Glu
260 265 270

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

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

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

Asn Tyr Glu Ile Ala Arg Glu Ala Ile Lys Asn Thr Arg Glu Tyr Phe
325 330 335

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

Asp Lys Leu Glu Leu Met Ala Lys Gln Ala Val Arg Asn Ser Gly Gly
355 360 365

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

Phe Lys Lys Ser Tyr
385

```
<210> 37
<211> 390
<212> PRT
<213> Clostridium acetobutylicum
```

<400> 37

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Ala Ser Asp Ile Arg Thr Lys Val Ser Lys Val Val Lys Leu Ala Lys
195 200 205

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

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

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

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

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

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

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

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

<210> 39
 <211> 259
 <212> PRT
 <213> Clostridium acetobutylicum

<400> 39

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

His Tyr Ile

<210> 40
 <211> 5976
 <212> DNA
 <213> Artificial

<220>
 <223> plasmid YEplac112PtdhTadh

<400> 40
 gcccggggga tccactagtt ctagaatccg tcgaaactaa gttctggtgt tttaaaacta 60
 aaaaaaagac taactataaa agtagaattt aagaagttta agaaatagat ttacagaatt 120
 acaatcaata cctaccgtct ttatatactt attagtcaag taggggaata atttcagggg 180
 actggtttca accttttttt tcagcttttt ccaaatacaga gagagcagaa ggtaatagaa 240
 ggtgtaagaa aatgagatag atacatgcgt gggatcaattg ccttggtgtca tcatttactc 300
 caggcagggt gcatcactcc attgagggtt tgcccgtttt ttgcctgttt gtgccctgt 360
 tctctgtagt tgcgctaaga gaatggacct atgaactgat ggttggtgaa gaaaacaata 420
 ttttggtgct gggattcttt ttttttctgg atgccagctt aaaaagcggg ctccattata 480
 tttagtggat gccaggaata aactgttcac ccagacacct acgatgttat atattctgtg 540

taacccgccc cctatTTTTgG gcatgtacgg gttacagcag aattaaaagg ctaatTTTTt	600
gactaaataa agttaggaaa atcactacta ttaattatTT acgtattctt tgaaatggcg	660
agtattgata atgataaaact gagctcgaat tcactggccg tcgTTTTaca acgtcgtgac	720
tgggaaaacc ctggcgTTac ccaacttaat cgccttgCag cacatcccc tttcgccagc	780
tggcgtaata gcgaagaggc ccgcaccgat cgcccttccc aacagttGcg cagcctgaat	840
ggcgaatggc gcctgatGcg gtatTTTTctc cttacgcacTc tgtGcggtat ttcacaccgc	900
atatatcgga tcgtacttGt taccatcat tgaatTTtga acatccgaac ctgggagttt	960
tccctgaaac agatagtata tttgaacctg tataataata tatagtctag cgctttacgg	1020
aagacaatgt atgtatTTcg gttcctggag aaactattGc atctattGca taggtaatct	1080
tgcacgtGc atccccggtt catTTTTctgc gtttccatct tgcacttcaa tagcatatct	1140
ttgttaacga agcatctgtg cttcatTTtG tagaacaaaa atgcaacGcg agagcgctaa	1200
TTTTTcaaac aaagaatctg agctgcattt ttacagaaca gaaatGcaac gcgaaagGc	1260
tattttacca acgaagaatc tgtgcttcat ttttgtaaaa caaaaatGca acgcgagagc	1320
gctaatTTTT caaacaaaga atctgagctg catTTTTtaca gaacagaaat gcaacGcgag	1380
agcgctatTT taccaacaaa gaatctatac ttctTTTTtG ttctacaaaa atgcatcccg	1440
agagcgctat ttttctaaca aagcatctta gattactTTt tttctccttt gtgcgctcta	1500
taatgcagtc tcttgataac tttttgcact gtaggtccgt taaggttaga agaaggctac	1560
tttggtgtct attttctctt ccataaaaaa agcctgactc cacttcccgc gtttactgat	1620
tactagcgaa gtcgcgggtg catTTTTtca agataaaggc atccccgatt atattctata	1680
ccgatgtgga ttgcgcatac tttgtgaaca gaaagtgata gcgttgatga ttcttcattg	1740
gtcagaaaat tatgaacggT ttcttctatt ttgtctctat atactacgta taggaaatgt	1800
ttacattttc gtattgtttt cgattcactc tatgaatagt tcttactaca atttttttgt	1860
ctaaagagta atactagaga taaacataaa aaatgtagag gtcgagttta gatgcaagtt	1920
caaggagcga aaggTggatg ggtaggTtat atagggatat agcacagaga tatatagcaa	1980
agagatactt ttgagcaatg tttgtggaag cggTattcgc aatatTTtag tagctcgTTa	2040
cagTccggtg cgTTTTtggT tttttgaaag tgcgTcttca gagcgctttt ggTTTTcaaa	2100
agcgctctga agttcctata ctttctagct agagaatagg aacttcggaa taggaacttc	2160
aaagcgTTtc cgaaaacgag cgcttccgaa aatGcaacgc gagctgcgca catacagctc	2220
actgttcacg tcgcacctat atctgcgtgt tgccTgtata tatatataca tgagaagaac	2280

ggcatagtc	gtgtttatgc	ttaaatacggt	acttatatgc	gtctatttat	gtaggatgaa	2340
aggtagtcta	gtacctcctg	tgatattatc	ccattccatg	cggggtatcg	tatgcttcct	2400
tcagcactac	cctttagctg	ttctatatgc	tgccactcct	caattggatt	agtctcatcc	2460
ttcaatgcta	tcatttcctt	tgatattgga	tcgatccgat	gataagctgt	caaacatgag	2520
aattgatcct	ttatgcttgc	ttttcaaaaag	gcttgcaggc	aagtgcacaa	acaataactta	2580
aataaatact	actcagtaat	aacctatttc	ttagcatttt	tgacgaaatt	tgctattttg	2640
ttagagtctt	ttacaccatt	tgtctccaca	cctccgctta	catcaacacc	aataacgcc	2700
tttaatctaa	gcgcatac	aacattttct	ggcgtcagtc	caccagctaa	cataaaatgt	2760
aagctctcgg	ggctctcttg	ccttccaacc	cagtcagaaa	tcgagttcca	atccaaaagt	2820
tcacctgtcc	cacctgcttc	tgaatcaaac	aaggaataa	acgaatgagg	tttctgtgaa	2880
gctgcactga	gtagtatgtt	gcagtctttt	ggaaatacga	gtcttttaat	aactggcaaa	2940
ccgaggaact	cttggtatcc	ttgccacgac	tcctctccat	gcagttggac	gatatcaatg	3000
ccgtaatcat	tgaccagagc	caaaacatcc	tccttaggtt	gattacgaaa	cacgccaacc	3060
aagtatttcg	gagtgccctga	actattttta	tatgctttta	caagacttga	aattttcctt	3120
gcaataaccg	ggtcaattgt	tctctttcta	ttgggcacac	atataatacc	cagcaagtca	3180
gcatcggaat	ctagtgcaca	ttctgcggcc	tctgtgctct	gcaagccgca	aactttcacc	3240
aatggaccag	aactacctgt	gaaattaata	acagacatac	tccaagctgc	ctttgtgtgc	3300
ttaatcacgt	atactcacgt	gctcaatagt	caccaatgcc	ctccctcttg	gccctctcct	3360
tttctttttt	cgaccgaatt	aattcttgaa	gacgaaagg	cctcgtgata	cgctatttt	3420
tataggttaa	tgtcatgata	ataatggttt	cttagacgtc	aggtggcact	tttcggggaa	3480
atgtgcgcgg	aaccctatt	tgtttatttt	tctaaataca	ttcaaataatg	tatccgctca	3540
tgagacaata	accctgataa	atgcttcaat	aatattgaaa	aaggaagagt	atgagtattc	3600
aacatttcgg	tgctgccctt	attccctttt	ttgcggcatt	ttgccttcct	gtttttgctc	3660
accagaaac	gctggtgaaa	gtaaaagatg	ctgaagatca	gttgggtgca	cgagtgggtt	3720
acatcgaact	ggatctcaac	agcggtaaga	tccttgagag	ttttcgcccc	gaagaacgtt	3780
ttccaatgat	gagcactttt	aaagtctctg	tatgtggcgc	ggtattatcc	cgtattgacg	3840
ccgggcaaga	gcaactcggg	cggcgcatat	actattctca	gaatgacttg	gttgagtact	3900
caccagtcac	agaaaagcat	cttacgggatg	gcatgacagt	aagagaatta	tgagtgctg	3960

ccataacccat	gagtgataac	actgcggcca	acttacttct	gacaacgatc	ggaggaccga	4020
aggagctaac	cgcttttttg	cacaacatgg	gggatcatgt	aactcgcctt	gatcgttggg	4080
aaccggagct	gaatgaagcc	ataccaaacg	acgagcgtga	caccacgatg	cctgtagcaa	4140
tggcaacaac	gttgcgcaaa	ctattaactg	gcgaactact	tactctagct	tcccggcaac	4200
aattaataga	ctggatggag	gcggataaag	ttgcaggacc	acttctgcgc	tcggcccttc	4260
cggttggtg	gtttattgct	gataaatctg	gagccggtga	gcgtgggtct	cgcggtatca	4320
ttgcagcact	ggggccagat	ggtaagccct	cccgtatcgt	agttatctac	acgacgggga	4380
gtcaggcaac	tatggatgaa	cgaaatagac	agatcgcgtga	gatagggtgcc	tcactgatta	4440
agcattggta	actgtcagac	caagtttact	catatatact	ttagattgat	ttaaaacttc	4500
atttttaatt	taaaaggatc	taggtgaaga	tcctttttga	taatctcatg	acaaaaatcc	4560
cttaacgtga	gttttcgttc	cactgagcgt	cagaccccg	agaaaagatc	aaaggatcct	4620
cttgagatcc	tttttttctg	cgcgtaatct	gctgcttgca	aacaaaaaaaa	ccaccgctac	4680
cagcgggtgt	ttgtttgccg	gatcaagagc	taccaactct	ttttccgaag	gtaactggct	4740
tcagcagagc	gcagatacca	aatactgtcc	ttctagtgtg	gccgtagtta	ggccaccact	4800
tcaagaactc	tgtagcaccg	cctacatacc	tcgctctgct	aatcctgtta	ccagtggctg	4860
ctgccagtgg	cgataagtcg	tgtcttaccg	ggttggactc	aagacgatag	ttaccggata	4920
aggcgcagcg	gtcgggctga	acgggggggt	cgtgcacaca	gccagcttg	gagcgaacga	4980
cctacaccga	actgagatac	ctacagcgtg	agctatgaga	aagcgccacg	cttcccgaag	5040
ggagaaaggc	ggacaggtat	ccggtaagcg	gcagggtcgg	aacaggagag	cgcacgaggg	5100
agcttccagg	gggaaacgcc	tggtatcttt	atagtcctgt	cgggtttcgc	cacctctgac	5160
ttgagcgtcg	atttttgtga	tgctcgtcag	gggggcggag	cctatggaaa	aacgccagca	5220
acgcggcctt	tttacggttc	ctggcctttt	gctggccttt	tgctcacatg	ttctttcctg	5280
cgttatcccc	tgattctgtg	gataaccgta	ttaccgcctt	tgagtgagct	gataccgctc	5340
gccgcagccg	aacgaccgag	cgcagcgagt	cagtgagcga	ggaagcggaa	gagcgcccaa	5400
tacgcaaacc	gcctctcccc	gcgcgttggc	cgattcatta	atgcagctgg	cacgacaggt	5460
ttcccgaactg	gaaagcgggc	agtgagcgca	acgcaattaa	tgtgagttag	ctcactcatt	5520
aggcacccca	ggctttacac	tttatgcttc	cggctcgtat	gttgtgtgga	attgtgagcg	5580
gataacaatt	tcacacagga	aacagctatg	accatgatta	cgccaagctt	aggcctgtgt	5640
ggaagaacga	ttacaacagg	tgttgtcctc	tgaggacata	aaatacacac	cgagattcat	5700

caactcattg ctggagttag catatctaca attgggtgaa atggggagcg atttgcaggc 5760
 atttgctcgg catgccggta gaggtgtggt caataagagc gacctcatgc tatacctgag 5820
 aaagcaacct gacctacagg aaagagttac tcaagaataa gaattttcgt tttaaaacct 5880
 aagagtcact ttaaaatttg tatacactta ttttttttat aacttattta ataataaaaa 5940
 tcataaatca taagaaattc gctcgagtcg actgca 5976

<210> 41
 <211> 13286
 <212> DNA
 <213> Artificial

<220>
 <223> pBOL34

<400> 41
 aagcttgcat gcctgcaggc cgacggcgcg cggggcccggt ttaaaccggcc ggccaagggtg 60
 agacgcgcat aaccgctaga gtactttgaa gaggaacag caatagggtt gctaccagta 120
 taaatagaca ggtacatata acactggaaa tgggtgtctg tttgagtacg ctttcaattc 180
 atttgggtgt gcactttatt atgttacaat atggaagggg actttacact tctcctatgc 240
 acatatatta attaaagtcc aatgctagta gagaaggggg gtaacacccc tccgcgctct 300
 tttccgattt ttttctaaac cgtggaatat ttccgatata cttttgttgt ttccgggtgt 360
 acaatatgga cttcctcttt tctggcaacc aaaccatac atcgggattc ctataatacc 420
 ttcgttggtc tccctaakat gtaggtggcg gaggggagat atacaataga acagatacca 480
 gacaagacat aatgggctaa acaagactac accaattaca ctgcctcatt gatggtggta 540
 cataacgaac taatactgta gccctagact tgatagccat catcatatcg aagtttcact 600
 accctttttc catttgccat ctattgaagt aataataggc gcatgcaact tcttttcttt 660
 ttttttcttt tctctctccc cgttggttgt ctcaccatat ccgcaatgac aaaaaaatga 720
 tggaagacac taaaggaaaa aattaacgac aaagacagca ccaacagatg tcgttggtcc 780
 agagctgatg aggggtatct cgaagcacac gaaacttttt ccttccttca ttcacgcaca 840
 ctactctcta atgagcaacg gtatacggcc ttccttccag ttacttgaat ttgaaataaa 900
 aaaaagtttg ctgtcttgct atcaagtata aatagacctg caattattaa tcttttgttt 960
 cctcgtcatt gttctcgttc cctttcttcc ttgtttcttt ttctgcacaa tatttcaagc 1020
 tataccaagc atacaatcaa ctatctcata tacaatgaag gaagttgtta ttgcttctgc 1080

tgtcagaact gccattggtt cttacggtaa gtctttgaag gacgtcccag ctgtcgactt	1140
gggtgctacc gccatcaagg aagctgtcaa gaaggctggt atcaagccag aagatgttaa	1200
cgaagttatc ttaggtaacg ttttgcaagc tggtttaggt caaaaccag ctctgcaagc	1260
ttctttcaag gctggtttgc cagttgaaat tccagccatg accatcaaca aggtttgtgg	1320
ttctggtttg agaactgttt ctttggtgc tcaaatacgc aaggctggtg acgctgatgt	1380
catcattgct ggtggtatgg aaaacatgtc cagagctcca tacttggtta acaatgctag	1440
atgggggttac agaatgggta acgccaagtt cgctgatgaa atgatcactg acggtttatg	1500
ggacgctttc aacgactacc acatgggtat cactgctgaa aacattgctg aaagatggaa	1560
catctccaga gaagaacaag atgaatttgc tttggcttct caaaagaagg ctgaagaagc	1620
catcaaactt ggtcaattca aggacgaaat tgtcccagtt gtcatacaagg gtagaaagg	1680
tgaaccggtt gtcgacaccg atgaacaccc aagattcggt tccaccattg aaggtttggc	1740
caagttgaaa ccagctttca agaaggatgg taccgtcact gctggtaacg cttccggttt	1800
gaacgactgt gctgctgttt tggttatcat gtctgctgaa aaggccaagg aattgggtgt	1860
caagccattg gccaaagattg tctcctacgg ttctgctggt gttgaccag ccatcatggg	1920
ttacggtcct ttctacgcta ccaaggctgc tatcgaaaag gctggttga cgttgacga	1980
attggatttg attgaatcca acgaagcttt cgctgctcaa tctttggctg ttgccaagga	2040
cttgaaattc gacatgaaca aggtcaacgt taacgggtgt gccattgctt tgggtcacc	2100
aattgggtgt tccggtgcca gaatcttgggt tacttttagtc cacgctatgc aaaagcgtga	2160
tgccaagaag ggtttggcta ctctatgtat cgggtggtgt caagggtactg ccatcttatt	2220
ggaaaagtgt taggcccggg cataaagcaa tcttgatgag gataatgatt tttttttgaa	2280
tatacataaa tactaccgtt tttctgctag attttgtgaa gacgtaaata agtacatatt	2340
actttttaag ccaagacaag attaagcatt aactttaccc ttttctcttc taagtttcaa	2400
tactagtatt cactgtttta aagttatggc gagaacgtcg gcggttaaaa tatattaccc	2460
tgaacgtggt gaattgaagt tctaggatgg tttaaagatt tttccttttt gggaaataag	2520
taaacaatat attgctgcct ttgcaaaacg cacataccca caatatgtga ctattggcaa	2580
agaacgcatt atcctttgaa gaggtggata ctgatactaa gagagtctct attccggctc	2640
cacttttagt ccagagatta cttgtcttct tacgtatcag aacaagaaag catttccaaa	2700
gtaattgcat ttgcccttga gcagtatata tatactaaga agtttaaaaca tttaaactgt	2760
tgtgtgcatt atatatatta aaaattaaga attagactaa ataaagtgtt tctaaaaaaa	2820

tattaaagtt gaaatgtgcg tgttgtgaat tgtgctctat tagaataatt atgacttgtg	2880
tgcgtttcat attttaaaat aggaaataac caagaaagaa aaagtaccat ccagagaaac	2940
caattatatac aaatcaaata aaacaaccag cttcggtgtg tgtgtgtgtg tgaagctaag	3000
agttgatgcc atttaatatc aaaattttaa ggtgtgtgtg tggataaaat attagaatga	3060
caattcgaga tgaaatttta agcaaaactct agtaggaaat aagcggctta ttcttgttgg	3120
ctcctaattc tttttagtgt atcagttccc attgataaaa aaattaaaat taaaattaga	3180
aaaattaaac cagaaaaatc aagttgatta aaatgtgaca aaaattatga ttaaattgcta	3240
cttcaacagg agcccgggcc tatttggagt agtcgtagaa acccttacca gactttctac	3300
ctaaccaacc agctctaacg tacttcttca ataaagtgtg aggtctgtac ttagagtcac	3360
cggtttcaga gtataagaca tccatgatgg ccaaacagat atccaaaccg atgaagtcac	3420
ctaattccaa tggacccatt ggggtggttag cacccaattt catggccttg tcgatatctt	3480
caacagaagc aataccttca gccaaaatac cgacagcttc gttgatcatt ggaatcaaga	3540
ttctgttgac aacgaaacct ggagcttcag caacttcaac tgggtcctta ccaatggcaa	3600
tggaagtttc cttgacagca tcgaaagttt cttgagaggt ggcaatacct ctgatgactt	3660
cgaccaactt catgactgga gctgggttga agaagtgcac accgataacc ttgtctggtc	3720
tcttggtagc agaagcaact tcagtgatgg acaaagaaga agtgttggaa gccaaaatgg	3780
tttctggctt acagatgttg tccaaatcag caaagatttg cttcttgatg tccattcttt	3840
caacggcagc ttcaatgacc aaatcacagt cagcagccat gttcaagtca acagtaccgg	3900
agattctggt caagatttcg accttggttag cttcttcaat cttacccttc ttgaccaact	3960
tggacaagtt cttgttgatg aaatccaaac cacggtcaac gaattcgtcc ttgatatctc	4020
tcaaaacaac ttcgaaaccc ttggcagcga aagcttgagc aataccagaa cccatggtac	4080
cggcaccaat gacacaaacc ttcttcattt tgatttagtg tttgtgtgtt gataagcagt	4140
tgcttggttt tttatgaaaa atagctagaa ggaataaggg attacaagag agatgttaca	4200
agaaagaagt aaaataaatt tgattaatat tgccattatc aaaagctatt tatatgttga	4260
aatcgtggag atcatgtgtg ccagaaaagg ccacagtttc cggggagagg cataccttga	4320
ggtggctagg aatcacggag acctcttgac ttgcagggtg ggctagctag aattaagtga	4380
ggtgacaagg tttccataca gttttgacct tgagacgttg ctacttacga tttgcagtat	4440
gcaagtctca tgctgcaaac aaaagaggac cgctcaggta atcgctcaat tagtggacgt	4500

tatcaggggc	gggagaggcg	aaagtggttt	ttggtggtgt	aagtaaaggt	cgtccaaata	4560
tgcaggtgtt	tgggtgctat	cctagtggaa	gctcggatca	gtagataacc	cgcctagaag	4620
cggtattttt	cttttttttt	cttcctttct	tttcgtcatt	atttcaaacg	cttttgcgct	4680
aagtaatgaa	tatctggcgg	ttccgcggta	atgcgacaat	ttgtgatatg	cactcttaaa	4740
accccgccac	gatgatcgca	cgtgccggca	tttatagacg	acttttctgg	ttgtcccgtc	4800
tcacggcaca	tgcattgcac	aatgaccgaa	ttcaggttgc	tactaacctt	tgtgttgtgt	4860
tattgctgtg	catgaggtgc	tcaagtcccc	gcggcatctg	actagtggta	actctagacg	4920
gcttcgatgc	agagagttcc	tcaaaatttt	tcttttcaat	tgtttgcctg	gtttccgcgg	4980
cgtatatcag	tttttggcga	tatggtaacg	cgatactcta	cggcaccttc	acggtagatg	5040
tcttttttaa	aagtgactgt	taattccagg	attgaaagga	agtgtcgaat	agtatagtat	5100
gctttctagc	ccggccggtt	aaatgggccc	gcggcccgtt	taaacggccg	gcccttcctt	5160
tttacagtgc	ttcggaaaag	cacagcgttg	tccaagggaa	caatttttct	tcaagttaat	5220
gcataagaaa	tatctttttt	tatgttttagc	taagtaaaag	cagcttggag	taaaaaaaaa	5280
aatgagtaaa	tttctcgatg	gattagtttc	tcacaggtaa	cataacaaaa	accaagaaaa	5340
gcccgcttct	gaaaactaca	gttgacttgt	atgctaaagg	gccagactaa	tgggaggaga	5400
aaaagaaacg	aatgtatatg	ctcattttaca	ctctatatca	ccatatggag	gataagttgg	5460
gctgagcttc	tgatccaatt	tattctatcc	attagttgct	gatatgtccc	accagccaac	5520
acttgatagt	atctactcgc	cattcacttc	cagcagcgcc	agtaggggtg	ttgagcttag	5580
taaaaatgtg	cgcaccacaa	gcctacatga	ctccacgtca	catgaaacca	caccgtgggg	5640
ccttgttgcg	ctaggaatag	gatatgcgac	gaagacgctt	ctgcttagta	accacaccac	5700
attttcaggg	ggtcgatctg	cttgcttcct	ttactgtcac	gagcggccca	taatcgcgct	5760
ttttttttta	aaggcgcgag	acagcaaaca	ggaagctcgg	gtttcaacct	tcggagtggg	5820
cgcagatctg	gagactggat	ctttacaata	cagtaaggca	agccaccatc	tgcttcttag	5880
gtgcatgcga	cggtatccac	gtgcagaaca	acatagtctg	aagaaggggg	ggaggagcat	5940
gttcattctc	tgtagcagta	agagcttggg	gataatgacc	aaaactggag	tctcgaaatc	6000
atataaatag	acaatatatt	ttcacacaat	gagatttgta	gtacagttct	attctctctc	6060
ttgcataaat	aagaaattca	tcaagaactt	ggtttgatat	ttcaccaaca	cacacaaaaa	6120
acagtacttc	actaaattta	cacacaaaac	aaaatggaat	tgaacaacgt	tatcttggaa	6180
aaggaaggta	aggttgccgt	tgtcaccatc	aacagaccaa	aggctttgaa	tgctttgaac	6240

tctgacactt tgaaggaaat ggactacgtc attggtgaaa ttgaaaacga ttctgaagtt	6300
ttggctgtca tcttgaccgg tgccggtgaa aagtctttcg ttgctggtgc tgatatctct	6360
gaaatgaagg aaatgaacac cattgaaggt agaaagttcg gtatcttagg taacaaggtt	6420
ttcagaagat tggaattggt ggaaaagcca gtcattgctg ctgtcaacgg tttcgctttg	6480
ggtggtggtt gtgaaattgc catgtcctgt gacatcagaa ttgcttcttc taacgctcgt	6540
ttcgggtcaac cagaagtcgg tctaggtatc actccaggtt tcggtggtac tcaaagatta	6600
tccagattgg ttggtatggg tatggccaag caattgatct tcaccgctca aaacatcaag	6660
gctgacgaag ctttgagaat tggtttagtc aacaaggttg ttgaaccatc tgaattgatg	6720
aacactgcca aggaaattgc taacaagatc gtctccaacg ctccagttgc tgtcaaattg	6780
tccaagcaag ccatcaacag aggtatgcaa tgtgatatcg acaccgcttt ggcttttgaa	6840
tctgaagctt tcggtgaatg tttctccact gaagacccaa aggatgctat gaccgctttc	6900
atcgaaaaga gaaagattga aggtttcaag aacagggtgat gagcccgggc gcgaatttct	6960
tatgatttat gatttttatt attaaataag ttataaaaaa aataagtgtg taaaaatttt	7020
aaagtgactc ttaggtttta aaacgaaaat tcttattctt gagtaactct ttcctgtagg	7080
tcaggttgct ttctcaggta tagcatgagg tcgctcttat tgaccacacc tctaccggca	7140
tgccgagcaa atgcctgcaa atcgctcccc atttcaccca attgtagata tgctaactcc	7200
agcaatgagt tgatgaatct cgggtgtgtat tttatgtcct cagaggacaa cacctgttgt	7260
aatcgttctt ccacacggat ccacagccta gccttcagtt gggctctatc ttcacgtca	7320
ttcattgcat ctactagccc cttacctgag cttcaagacg ttatatcgct tttatgtatc	7380
atgatcttat cttgagatat gaatacataa atatatttac tcaagtgtat acgtgcatgc	7440
tttttttacg gtttaaacat ttaaattgggc cgctctagag gatccccggg taccgagctc	7500
gggccagcg ctactagttc cggtaatttg aaaacaaacc cggctctgaa gcggagatcc	7560
ggcgataatt accgcagaaa taaaccata cagcagacgt agaaccagcc gcacatggcc	7620
ggagaaactc ctgcgagaat ttcgtaaact cgcgcgcatt gcatctgtat ttcctaatagc	7680
ggcacttcca ggcctcgaga cctctgacat gcttttgaca ggaatagaca ttttcagaat	7740
gttatccata tgcctttcgg gtttttttcc ttccttttcc atcatgaaaa atctctcgag	7800
accgtttatc cattgctttt ttgttgtctt tttccctcgt tcacagaaaag tctgaagaag	7860
ctatagtaga actatgagct ttttttggtt ctgttttctt tttttttttt tttacctctg	7920

tggaaattgt tactctcaca ctcttttagtt cgtttgtttg ttttgtttat tccaattatg	7980
accggtgacg aaacgtgggc gatgggtgggt accgcttatg ctccccctcca ttagtttcga	8040
ttatataaaa aggccaaata ttgtattatt ttcaaagtgc ctatcattat cgtctaakat	8100
ctaattttctc ttaaatttttt tctcttttctt tcctataaca ccaatagtga aaatctttttt	8160
ttctttctata tctacaaaaa ctttttttttt ctatcaacct cggttgataaa ttttttcttt	8220
aacaatcggt aataattaat taattggaaa ataaccattt tttctctctt ttatacacac	8280
attcaaaaaga aagaaaaaaa atatacccca gctagttaaa gaaaatcatt gaaaagaata	8340
agaagataag aaagatttaa ttatcaaaca atatcaatat gcctcaatcc tgggaagaac	8400
tggccgctga taagcgcgcc cgctcgcga aaaccatccc tgatgaatgg aaagtccaga	8460
cgctgcctgc ggaagacagc gttattgatt tcccaaagaa atcggggatc ctttcagagg	8520
ccgaactgaa gatcacagag gcctccgctg cagatcttgt gtccaagctg gcggccggag	8580
agttgacctc ggtggaagtt acgctagcat tctgtaaacy ggcagcaatc gccagcagt	8640
taacaaactg cgccacgag ttcttccctg acgcccgtct cgcgaggca agggaactcg	8700
atgaatacta cgcaaagcac aagagacccg ttggtccact ccatggcctc cccatctctc	8760
tcaaagacca gtttcgagtc aagggtacg aaacatcaat gggctacatc tcatggctaa	8820
acaagtacga cgaaggggac tcggttctga caaccatgct ccgcaaagcc ggtgccgtct	8880
tctacgtcaa gacctctgtc ccgcagacct tgatgggtctg cgagacagtc aacaacatca	8940
tcgggcgcac cgtcaaccca cgcaacaaga actgggtcgtg cggcggcagt tctggtggtg	9000
agggtgcgat cgttgggatt cgtgggtggcg tcatcgggtg aggaacggat atcgggtggt	9060
cgattcgagt gccggccgcg ttcaacttcc tgtacggtct aaggccgagt catgggcggc	9120
tgccgtatgc aaagatggcg aacagcatgg agggtcagga gacggtgcac agcgttgctg	9180
ggccgattac gactctgtt gaggacctcc gcctcttcac caaatccgtc ctcggtcagg	9240
agccatggaa atacgactcc aaggtcatcc ccatgccctg gcgccagtcc gagtccgaca	9300
ttattgcctc caagatcaag aacggcgggc tcaatatcgg ctactacaac ttcgacggca	9360
atgtccttcc acaccctcct atcctgcgcg gcgtggaaac caccgtcgcc gactcgcca	9420
aagccggtca caccgtgacc ccgtggacgc catacaagca cgatttcggc cacgatctca	9480
tctcccatat ctacgcggct gacggcagcg ccgacgtaat gcgcgatatc agtgcacccg	9540
gcgagccggc gattccaaat atcaaagacc tactgaacct gaacatcaaa gctgttaaca	9600
tgaacgagct ctgggacacg catctccaga agtggaaatta ccagatggag taccttgaga	9660

aatggcggga	ggctgaagaa	aaggccggga	aggaactgga	cgccatcatc	gcgccgatta	9720
cgcctaccgc	tgcggtacgg	catgaccagt	tccggtacta	tgggtatgcc	tctgtgatca	9780
acctgctgga	tttcacgagc	gtggttggtc	cggttacctt	tgcggataag	aacatcgata	9840
agaagaatga	gagtttcaag	gcggttagtg	agcttgatgc	cctcgtgcag	gaagagtatg	9900
atccggaggc	gtaccatggg	gcaccggttg	cagtgcaggt	tatcggacgg	agactcagtg	9960
aagagaggac	gttggcgatt	gcagaggaag	tggggaagtt	gctgggaaat	gtggtgactc	10020
cataggtcga	gaatttatac	ttagataagt	atgtacttac	aggtatattt	ctatgagata	10080
ctgatgtata	catgcatgat	aatattttaa	cggttattag	tgccgattgt	cttgtgcgat	10140
aatgacgttc	ctatcaaagc	aatacactta	ccacctatta	catgggccaa	gaaaatattt	10200
tcgaacttgt	ttagaatatt	agcacagagt	atatgatgat	atccgttaga	ttatgcatga	10260
ttcatttccta	caactttttc	gtagcataag	gattaattac	ttggatgcca	ataaaaaaaaa	10320
aaaacatcga	gaaaatttca	gcatgctcag	aaacaattgc	agtgtatcaa	agtaaaaaaaaa	10380
agattttcgc	tacatgttcc	ttttgaagaa	agaaaatcat	ggaacattag	atttacaaaa	10440
atttaaccac	cgctgattaa	cgattagacc	gttaagcgca	caacaggtta	ttagtacaga	10500
gaaagcattc	tgtggtggtg	ccccggactt	tcttttgcca	cataggtaaa	tcgaatacca	10560
tcatactatc	ttttccaatg	actccctaaa	gaaagactct	tcttcgatgt	tgtatacgtt	10620
ggagcatagg	gcaagaattg	tggcttgaga	tgaattcact	ggccgtcggt	ttacaacgtc	10680
gtgactggga	aaaccctggc	gttaccacaac	ttaatcgctt	tgcagcacat	ccccctttcg	10740
ccagctggcg	taatagcgaa	gaggcccgca	ccgatcgccc	ttccaacag	ttgcgcagcc	10800
tgaatggcga	atggcgcttg	atgcggtatt	ttctccttac	gcatctgtgc	ggtatttcac	10860
accgcatatg	gtgcactctc	agtacaatct	gctctgatgc	cgcatagtta	agccagcccc	10920
gacacccgcc	aacacccgct	gacgcgccct	gacgggcttg	tctgctcccg	gcatccgctt	10980
acagacaagc	tgtgaccgtc	tccgggagct	gcatgtgtca	gaggttttca	ccgtcatcac	11040
cgaaacgcgc	gagacgaaag	ggcctcgtga	tacgcctatt	tttatagggt	aatgtcatga	11100
taataatggt	ttcttagacg	tcagggtggca	cttttcgggg	aaatgtgcgc	ggaaccctta	11160
tttgtttatt	tttctaaata	cattcaaata	tgtatccgct	catgagacaa	taaccctgat	11220
aaatgcttca	ataatattga	aaaaggaaga	gtatgagtat	tcaacatttc	cgtgtcgccc	11280
ttattccctt	ttttgcggca	ttttgccttc	ctgtttttgc	tcaccagaaa	acgctggtga	11340

aagtaaaaga tgctgaagat cagttgggtg cacgagtggg ttacatcgaa ctggatctca 11400
acagcggtaa gatccttgag agttttcgcc ccgaagaacg ttttccaatg atgagcactt 11460
ttaaagttct gctatgtggc gcggtattat cccgtattga cgccgggcaa gagcaactcg 11520
gtcgccgcat aactatttct cagaatgact tggttgagta ctcaccagtc acagaaaagc 11580
atcttacgga tggcatgaca gtaagagaat tatgcagtgc tgccataacc atgagtgata 11640
aactgcggc caacttactt ctgacaacga tcggaggacc gaaggagcta accgcttttt 11700
tgcacaacat gggggatcat gtaactcgcc ttgatcgttg ggaaccggag ctgaatgaag 11760
ccataccaaa cgacgagcgt gacaccacga tgccctgtagc aatggcaaca acgttgcgca 11820
aactattaac tggcgaacta cttactctag cttcccgga acaattaata gactggatgg 11880
aggcggataa agttgcagga ccacttctgc gctcgccct tccggctggc tggtttattg 11940
ctgataaatc tggagccggt gagcgtgggt ctgcggtat cattgcagca ctggggccag 12000
atggtaagcc ctcccgatc gtagttatct acacgacggg gagtcaggca actatggatg 12060
aacgaaatag acagatcgct gagataggtg cctcactgat taagcattgg taactgtcag 12120
accaagttta ctcatatata ctttagattg atttaaaact tcatttttaa tttaaaagga 12180
tctagtgaa gatccttttt gataatctca tgacccaaat cccttaacgt gagttttcgt 12240
tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat ctttttttc 12300
tgcgcgtaat ctgctgcttg caaacaaaaa aaccaccgct accagcggtg gtttgtttg 12360
cggatcaaga gctaccaact ctttttccga aggtaactgg cttcagcaga gcgagatac 12420
caaatactgt ccttctagt tagccgtagt taggccacca cttcaagaac tctgtagcac 12480
cgctacata cctcgctctg ctaatcctgt taccagtggc tgctgccagt ggcgataagt 12540
cgtgtcttac cgggttgac tcaagacgat agttaccgga taaggcgag cggtcgggt 12600
gaacgggggg ttcgtgcaca cagcccagct tggagcgaac gacctacacc gaactgagat 12660
acctacagcg tgagctatga gaaagcgcca cgcttccga agggagaaag gcggacaggt 12720
atccggtaa cggcagggtc ggaacaggag agcgcacgag ggagcttcca ggggaaacg 12780
cctggatatc ttatagtcct gtcgggtttc gccacctctg acttgagcgt cgatttttgt 12840
gatgctcgtc aggggggcgg agcctatgga aaaacgccag caacgcggcc tttttacggt 12900
tcctggcctt ttgctggcct tttgctcaca tgttctttcc tgcgttatcc cctgattctg 12960
tggaataaccg tattaccgcc tttgagttag ctgataaccg tcgccgcagc cgaacgaccg 13020
agcgagcga gtcagtgagc gaggaagcgg aagagcgccc aatacgcaaa ccgcctctcc 13080

ccgcgcgttg gccgattcat taatgcagct ggcacgacag gtttcccgac tggaaagcgg 13140
gcagtgagcg caacgcaatt aatgtgagtt agctcactca ttaggcaccc caggctttac 13200
actttatgct tccggctcgt atgttgtgtg gaattgtgag cggataacaa tttcacacag 13260
gaaacagcta tgaccatgat tacgcc 13286

<210> 42
<211> 16359
<212> DNA
<213> Artificial

<220>
<223> pBOL36

<400> 42
aagcttgcat gcctgcaggt cgacggcgcg ccggggcccg ttaaacaatg gcaaaactgag 60
cacaacaata ccagtccgga tcaactggca ccatctctcc cgtagtctca tctaattttt 120
cttccggatg aggttccaga tataccgcaa cacctttatt atgggttccc tgagggaata 180
atagaatgtc ccattcgaaa tcaccaattc taaacctggg cgaattgtat ttcggggttg 240
ttaactcgtt ccagtcagga atgttccacg tgaagctatc ttccagcaaa gtctccactt 300
cttcatcaaa ttgtgggaga atactcccaa tgctcttata tatgggactt ccgggaaaca 360
cagtaccgat acttcccaat tCGTcttcag agctcattgt ttgtttgaag agactaatca 420
aagaatcgtt ttctcaaaaa aattaatatc ttaactgata gtttgatcaa aggggcaaaa 480
cgtaggggca aacaaacgga aaaatcgttt ctcaaatttt ctgatgccaa gaactctaac 540
cagtcttata taaaaattgc cttatgatcc gtctctccgg ttacagcctg tgtaactgat 600
taatcctgcc tttctaatac ccattctaata gttttaatta agggattttg tcttcattaa 660
cggctttcgc tcataaaaat gttatgacgt ttgcccgcga ggcgggaaac catccacttc 720
acgagactga tctcctctgc cggaacaccg ggcattctca acttataagt tggagaaata 780
agagaatttc agattgagag aatgaaaaaa aaaaaaaaaa aaaaggcaga ggagagcata 840
gaaatggggt tcactttttg gtaaagctat agcatgccta tcacatataa atagagtgcc 900
agtagcgact tttttcacac tcgaaatact cttactactg ctctcttggt gtttttatca 960
cttcttggtt cttcttggtg aatagaatat caagctacaa aaagcataca atcaactatc 1020
aactattaac tatatcgtaa tacacaggcc ggccaaaatg aaggccaaat caaggcggga 1080
agggacaacc aggacgtaaa gggtagcctc ccataacat aaactcaata aaatatatag 1140

tcttcaactt	gaaaaaggaa	caagctcatg	caaagaggtg	gtacccgcac	gccgaaatgc	1200
atgcaagtaa	cctattcaaa	gtaatatctc	atacatgttt	catgagggta	acaacatgcg	1260
actgggtgag	catatgttcc	gctgatgtga	tgtgcaagat	aaacaagcaa	gacagaaact	1320
aacttcttct	tcatgtaata	aacacacccc	gcgtttattt	acctatcttt	aaacttcaac	1380
accttatatc	ataactaata	tttcttgaga	taagcacact	gcaccatac	cttccttaaa	1440
aacgtagctt	ccagtttttg	gtggttcttg	cttccttccc	gattccgccc	gctaaacgca	1500
taattttggt	gcctggtggc	atttgcaaaa	tgcataacct	atgcatttaa	aagattatgt	1560
atgctcttct	gacttttcgt	gtgatgaggc	tcgtggaaaa	aatgaataat	ttatgaattt	1620
gagaacaatt	ttgtgttggt	acggtatttt	actatggaat	aatcaatcaa	ttgaggattt	1680
tatgcaaata	tcgtttgaat	atttttccga	ccctttgagt	acttttcttc	ataattgcat	1740
aatattgtcc	gctgcccggt	tttctgttag	acggtgtctt	gatctacttg	ctatcgttca	1800
acaccacctt	attttctaac	tatttttttt	ttagctcatt	tgaatcagct	tatggtgatg	1860
gcacattttt	gcataaacct	agctgtcctc	gttgaacata	ggaaaaaaaa	atatataaac	1920
aaggctcttt	cactctcctt	ggaatcagat	ttgggtttgt	tccctttatt	ttcataattc	1980
ttgtcatatt	cttttctcaa	ttattatctt	ctactcataa	cctcacgcaa	aataacacag	2040
tcaaatcaat	caaaatggac	ttcaacttga	ccagagaaca	agaattgggtc	agacaaatgg	2100
ttagagaatt	tgctgaaaac	gaagttaagc	caattgctgc	tgaaatcgat	gaaactgaaa	2160
gattcccaat	ggaaaacgtc	aagaagatgg	gtcaatacgg	tatgatgggt	attccattct	2220
ctaaggaata	cggtggtgct	ggtggtgacg	tcttgtctta	catcattgct	gtcgaagaat	2280
tgtccaaggt	ttgtggtacc	actggtgtca	tcttatctgc	tcacacttct	ctatgtgcct	2340
ccttgatcaa	cgaacacggt	actgaagaac	aaaagcaaaa	gtacttggtt	ccattggcca	2400
agggtgaaaa	gattggtgcc	tacggtttga	ctgaacaaaa	cgctggtact	gactctggtg	2460
ctcaacaaac	tggtgccggt	ttggaagggtg	accactacgt	catcaacggt	tccaagatct	2520
tcatcaccaa	cggtggtggt	gctgacacct	ttgtcatctt	cgctatgacc	gatcgtacca	2580
agggtaccaa	gggtatctct	gctttcatta	ttgaaaagggt	tttcaagggt	ttctccatcg	2640
gtaaggtcga	acaaaagttg	ggtatcagag	cttcctctac	caactgaattg	gttttcgaag	2700
acatgattgt	tccagttgaa	aacatgatcg	gtaaggaagg	taagggtttc	ccaattgcca	2760
tgaagacttt	agatggtggt	agaattggta	ttgctgctca	agctttgggt	attgctgaag	2820
gtgccttcaa	cgaagctaga	gcttacatga	aggaaagaaa	gcaattcggt	agatctttgg	2880

acaaattcca	aggtttggct	tggatgatgg	ctgacatgga	cgttgccatc	gaatctgctc	2940
gttacttggg	ctacaaggct	gcttacttga	agcaagctgg	tttgccatac	accgtcgatg	3000
ctgccagagc	taagttgcac	gctgccaacg	ttgccatgga	tgtcaccacc	aaggctgtcc	3060
aattattcgg	tggttacggg	tacaccaagg	actaccagtg	tgaagaagt	atgagagatg	3120
ctaagatcac	tgaaatctac	gaaggtactt	ctgaagttca	aaagttgggt	atctccggta	3180
agatcttcag	ataggcccgg	gcataaagca	atcttgatga	ggataatgat	ttttttttga	3240
atatacataa	atactaccgt	ttttctgcta	gattttgtga	agacgtaaat	aagtacatat	3300
tactttttta	gccaagacaa	gattaagcat	taactttacc	cttttctctt	ctaagtttca	3360
atactagtta	tcactgttta	aaagttatgg	cgagaacgtc	ggcgggttaa	atatattacc	3420
ctgaacgtgg	tgaattgaag	ttctaggatg	gtttaaagat	ttttcctttt	tgggaaataa	3480
gtaaacaata	tattgctgcc	tttgcaaaac	gcacataccc	acaatatgtg	actattggca	3540
aagaacgcat	tatcctttga	agaggtggat	actgatacta	agagagtctc	tattccggct	3600
ccacttttag	tccagagatt	acttgtcttc	ttacgtatca	gaacaagaaa	gcatttccaa	3660
agtaattgca	tttgcccttg	agcagtatat	atatactaag	aagtttaaac	atttaaacgg	3720
ccggcctaga	aagcactata	tactattcga	cacttccttt	caatcctgga	attaacagtc	3780
acttttaaaa	aagacatcta	ccgtgaagggt	gccgtagagt	atcgcgttac	catatcgcca	3840
aaaactgata	tacgccggcg	aaaccaggca	aacaattgaa	aagaaaaatt	ttgaggaact	3900
ctctgcatcg	aagccgtcta	gagttaccac	tagtcagatg	ccgcggggac	ttgagcacct	3960
catgcacagc	aataacacaa	cacaatgggt	agtagcaacc	tgaattcggg	cattgatgca	4020
tgcattgtgcc	gtgaagcggg	acaaccagaa	aagtcgtcta	taaatgccgg	cacgtgcgat	4080
catcgtggcg	gggttttaag	agtgcataat	acaaattgtc	gcattaccgc	ggaaccgcca	4140
gatattcatt	acttgacgca	aaagcgtttg	aaataatgac	gaaaaagaag	gaagaaaaaa	4200
aaagaaaaat	accgcttcta	ggcgggttat	ctactgatcc	gagcttccac	taggatagca	4260
cccaaacacc	tgcataattg	gacgaccttt	acttacacca	ccaaaaacca	ctttcgcttc	4320
tcccgcccct	gataacgtcc	actaattgag	cgattacctg	agcggtcctc	ttttgtttgc	4380
agcatgagac	ttgcatactg	caaatcgtaa	gtagcaacgt	ctcaagggtc	aaactgtatg	4440
gaaaccttgt	cacctcactt	aattctagct	agcctaccct	gcaagtcaag	aggtctccgt	4500
gattcctagc	cacctcaagg	tatgcctctc	ccgggaaact	gtggcctttt	ctggcacaca	4560

tgatctccac gatttcaaca tataaatagc ttttgataat ggcaatatta atcaaattta	4620
ttttacttct ttcttgtaac atctctcttg taatccctta ttccttctag ctatttttca	4680
taaaaaacca agcaactgct tatcaacaca caaacactaa atcaaaatgg tcgatttcga	4740
atactctatc ccaaccagaa tcttcttcgg taaggacaag atcaacgttt tgggtagaga	4800
attgaagaaa tacggttcca aggttttgat tgtctacggt ggtggttcca tcaagagaaa	4860
cggtatctac gacaaggctg tctccatttt ggaaaagaac tctatcaa at tctacgaatt	4920
ggctggtggt gaaccaaacc caagagttac caccgtcgaa aagggtgtca agatctgtcg	4980
tgaaaacggt gttgaagttg ttttggccat cgggtggtggt tctgccattg actgtgcca	5040
ggtcattgct gctgcctgtg aatacgatgg taacccatgg gacattgtct tggatggttc	5100
taagatcaag cgtgtcttac caattgcttc catcttgact atcgctgcta ctggttctga	5160
aatggacacc tgggctgtta tcaacaacat ggacactaac gaaaagttga ttgctgtc	5220
cccagatatg gcccacaaagt tctctatttt ggacccaacc tacacttaca ctgttccaac	5280
caaccaaact gctgctggta ctgctgat at catgtctcac atctttgaag tttacttctc	5340
caacaccaag accgcttact tgcaagacag aatggctgaa gctctattaa gaacctgtat	5400
caagtacggt ggtattgctt tggaaaagcc agatgactac gaagccagag ctaacttgat	5460
gtgggcttcc tctttggcta tcaacggttt attgacttac ggtaaggaca ccaactggtc	5520
cgttcatttg atggaacacg aattgtctgc ttactacgat atcactcacg gtgtcggttt	5580
ggccatcttg actccaaact ggatggaata cattttgaac aacgacactg tctacaagtt	5640
cgtcgaatac ggtgttaacg tctggggtat tgacaaggaa aagaaccact acgacattgc	5700
tcaccaagcc atccaaaaga ccagagacta tttcgtcaac gttttgggtt taccatccag	5760
attaagagat gttggtattg aagaagaaaa attggatata atggctaagg aatctgtcaa	5820
attgactggt ggtaccattg gtaacttgag acctgttaac gcttctgaag ttttgcaaat	5880
cttcaagaaa tctgttttag cccgggctcc tgttgaagta gcatttaatc ataatttttg	5940
tcacatttta atcaacttga tttttctggt ttaatttttc taattttaat ttttaatttt	6000
ttatcaatgg gaactgatac actaaaaaga attaggagcc aacaagaata agccgcttat	6060
ttctacttag agtttgctta aaatttcac tcgaattgtc attctaatat tttatccaca	6120
cacacacctt aaaattttta gattaaatgg catcaactct tagcttcaca cacacacaca	6180
caccgaagct ggttgtttta tttgatttga tataattggt ttctctggat ggtacttttt	6240
ctttcttggt tatttccat tttaaaatat gaaacgcaca caagtcataa ttatttcta	6300

agagcacaat	tcacaacacg	cacatttcaa	ctttaatat	tttttagaaa	cactttat	6360
agtctaattc	ttaatTTTTA	atatatataa	tgcacacaca	cgtttaaatg	ggcccgcggc	6420
ccgtttaaac	ggccggccct	tccctttttac	agtgtctcgg	aaaagcacag	cgttgtccaa	6480
gggaacaatt	tttcttcaag	ttaatgcata	agaaatatct	ttttttatgt	ttagctaagt	6540
aaaagcagct	tggagtataa	aaaaaaatga	gtaaatttct	cgatggatta	gtttctcaca	6600
ggtaacatag	caaaaaccaa	gaaaagcccg	cttctgaaaa	ctacagttga	cttgtatgct	6660
aaagggccag	actaatggga	ggagaaaaag	aaacgaatgt	atatgctcat	ttacactcta	6720
tatcaccata	tggaggataa	gttgggctga	gcttctgata	caatttat	tatccattag	6780
ttgctgatat	gtcccaccag	ccaacacttg	atagtatcta	ctcgccattc	acttccagca	6840
gcgccagtag	ggttggtgag	cttagtaaaa	atgtgcgcac	cacaagccta	catgactcca	6900
cgtcacatga	aaccacaccg	tggggccttg	ttgcgctagg	aataggatat	gcgacgaaga	6960
cgttctgct	tagtaaccac	accacat	cagggggctg	atctgcttgc	ttcctttact	7020
gtcacgagcg	gcccataatc	gcgctTTTTT	tttaaaaggc	gcgagacagc	aaacaggaag	7080
ctcgggtttc	aaccttcgga	gtggtcgcag	atctggagac	tggatcttta	caatacagta	7140
aggcaagcca	ccatctgctt	cttaggtgca	tgcgacggta	tccacgtgca	gaacaacata	7200
gtctgaagaa	gggggggagg	agcatgttca	ttctctgtag	cagtaagagc	ttggtgataa	7260
tgacaaaaac	tggagtctcg	aatcatata	aatagacaat	atattttcac	acaatgagat	7320
ttgtagtaca	gttctattct	ctctcttgca	taaataagaa	attcatcaag	aacttggttt	7380
gatatttcac	caacacacac	aaaaaacagt	acttcactaa	atttacacac	aaaacaaaat	7440
gaaggttacc	aacaaaaagg	aattgaagca	aaagttgaac	gaattgagag	aagctcaaaa	7500
gaagttcgct	acctacactc	aagaacaagt	tgacaagatc	ttcaagcaat	gtgccattgc	7560
tgctgccaa	gaacgtatca	acttggccaa	gttggctgtc	gaagaaaccg	gtattgggtt	7620
ggttgaagac	aagatcatca	agaaccactt	cgctgctgaa	tacatctaca	acaagtacaa	7680
gaacgaaaag	acctgtggta	tcatcgacca	cgatgactct	ttgggtatca	ccaaggttgc	7740
tgaaccaatc	ggtattgtcg	cgcctattgt	cccaaccact	aaccaactt	ccactgccat	7800
cttcaaatct	ttgatctcct	tgaagaccag	aaacgctatc	ttcttctccc	cacaccaag	7860
agccaagaag	tccaccattg	ctgctgccaa	attaatcttg	gatgctgctg	ttaaggctgg	7920
tgccccaag	aacattattg	gttggatcga	tgaaccttcc	attgaattgt	ctcaagactt	7980

gatgtctgaa gctgatatca tcttggctac cggtgggtcca tccatgggtca aggccgctta	8040
ctcttctggt aagccagcta ttggtggttg tgctggtaac actccagcta tcatcgatga	8100
atctgctgac attgacatgg ctgtctctc cattatcttg tccaagactt atgacaacgg	8160
tgtcatctgt gcctctgaac aatccatctt gggtatgaac tctatctacg aaaaggtcaa	8220
ggaagaattt gtaagagag gttcctacat cttaaacc aaatgaaattg ccaagatcaa	8280
ggaaccatg ttcaagaacg gtgccatcaa cgctgacatt gtcggtaa atctgcttacat	8340
cattgccaa atggctggta ttgaagttcc acaaccact aagattttga tcggtgaagt	8400
tcaatctgtc gaaaagtctg aattattctc tcacgaaaag ttgtctccag tcttggctat	8460
gtacaaggtc aaggatttcg acgaagcttt gaagaaggct caaagattaa ttgaattagg	8520
tggttctggt cacacctctt ctctatacat tgactctcaa aacaacaagg acaaggtcaa	8580
ggaattcgg ttagctatga agacttccag aactttcatc aacatgccat cttctcaagg	8640
tgcttctggt gatgtgtaca actttgccat tgctccatct ttcactttag gttgtggtac	8700
ctgggggtggt aactctgttt ctcaaaacgt tgaaccaaag catttgctaa acatcaagtc	8760
cgttgctgaa agaagagaaa acatgttgtg gttcaagggt ccacaaaaga tctacttcaa	8820
atacggttgt ttgagatttg ctttgaagga attgaaagat atgaacaaga agcgtgcttt	8880
catcgttact gacaaggatt tgttcaaatt gggttacgtt aacaagatca ctaaggtttt	8940
ggatgaaatt gatatcaagt actccatctt cactgatatc aaatctgacc caaccattga	9000
ctccgtcaag aagggtgcta aggaaatggt gaacttcgaa ccagatacca ttatctccat	9060
tggtgggtggt tctccaatgg atgctgcaa gggtatgcat ttgttgtagc aataaccaga	9120
agctgaaatc gaaaacttgg ccatcaactt catggacatc agaaagagaa tctgtaactt	9180
cccaaagttg ggtaccaagg ccatttctgt tgccattcca accaccgtg gtaccggttc	9240
tgaagctact ccatttgctg tcatcaccaa cgacgaaacc ggtatgaagt acccattgac	9300
ctcttacgaa ttgactccaa acatggccat cattgacact gaattgatgt tgaacatgcc	9360
aagaaagttg actgctgcta ccggtattga cgcttttagtc cacgctatcg aagcttacgt	9420
ctccgttatg gccactgact acactgacga attggctttg agagctatca agatgatctt	9480
caagtacttg ccaagagctt acaagaacgg tactaacgat atcgaagctc gtgaaaagat	9540
ggctcacgct tccaacattg ctggtatggc ttctgctaac gctttcttgg gtgtttgtca	9600
ctccatggcc cacaagttgg gtgctatgca ccacgttcct cacggtattg cttgtgctgt	9660
tttgattgaa gaagtcatca agtacaacgc tactgactgt ccaaccaagc aaactgcttt	9720

```

cccacaatac aagtctccaa acgccaagag aaagtagcgt gaaattgctg aatacttgaa 9780
cttgaaaggt acttctgaca ctgaaaaggt cactgcttta atcgaagcta tctccaagtt 9840
gaagattgac ttatctattc ctcaaaacat ctctgctgct ggtattaaca agaaggactt 9900
ctacaacact ttagacaaga tgtccgaatt ggctttcgat gaccaatgta ccaccgctaa 9960
cccaagatac ccattgatct ctgaattgaa ggatatctac atcaagtcct ttttaagcccg 10020
ggcgcggtac tcttatgtct ttacgattta tagttttcat tatcaagtat gcctatatta 10080
gtatatagca tcttttagatg acagtgttcg aagtttcacg aataaaagat aatattctac 10140
tttttgcctc caccgcgttt gctagcacga gtgaacacca tccctcgcct gtgagttgta 10200
cccattcctc taaactgtag acatggtagc ttcagcagtg ttcgttatgt acggcatcct 10260
ccaacaaaca gtcggttata gtttgcctg ctctctgaa tcgtctccct cgatatttct 10320
cattttcctt cgcattgccag cattgaaatg atcgaagtgc aatgatgaaa cggtaattct 10380
tctgtcattt actcatctca tctcatcaag ttatataatt ctatacggat gtaatttttc 10440
acttttcgtc ttgacgtcca ccctataatt tcaattattg aaccctcaca aatgatgcac 10500
tgcaatgtac acaccctcat atagtttaaa catttaaagtg ggccgctcta gaggatcccc 10560
gggtaccgag ctcgggccca gcgctactag ttccggtaat ttgaaaacaa acccggtctc 10620
gaagcggaga tccggcgata attaccgcag aaataaaccc atacacgaga cgtagaacca 10680
gccgcacatg gccggagaaa ctctgcgag aatttcgtaa actcgcgcgc attgcatctg 10740
tatttcctaa tgcggcactt ccaggcctcg agacctctga catgcttttg acaggaatag 10800
acattttcag aatgttatcc atatgccttt cgggtttttt tccttccttt tccatcatga 10860
aaaatctctc gagaccgttt atccattgct tttttgttgt ctttttcctt cgttcacaga 10920
aagtctgaag aagctatagt agaactatga gctttttttg tttctgtttt cttttttttt 10980
ttttttacct ctgtggaaat tgttactctc acactcttta gttcgtttgt ttgttttgtt 11040
tattccaatt atgaccggtg acgaaacgtg gtcgatggtg ggtaccgctt atgctcccct 11100
ccattagttt cgattatata aaaaggccaa atattgtatt attttcaa atgtcctatcat 11160
tatcgtctaa catctaattt ctcttaaat ttttctcttt ctttcctata acaccaatag 11220
tgaaaatctt tttttcttct atatctacaa aaactttttt tttctatcaa cctcgttgat 11280
aaattttttc ttttaacaatc gttaataatt aattaattgg aaaataacca ttttttctct 11340
cttttataca cacattcaaa agaaagaaaa aaaatatacc ccagctagtt aaagaaaatc 11400

```

attgaaaaga ataagaagat aagaaagatt taattatcaa acaatatcaa tatgcctcaa 11460
 tcctgggaag aactggccgc tgataagcgc gcccgctcg caaaaaccat ccctgatgaa 11520
 tggaaagtcc agacgctgcc tgcggaagac agcgttattg atttcccaa gaaatcgggg 11580
 atcctttcag aggccgaact gaagatcaca gaggcctccg ctgcagatct tgtgtccaag 11640
 ctggcgggccg gagagttgac ctcggtggaa gttacgctag cattctgtaa acgggcagca 11700
 atcgcccagc agttaacaaa ctgcgcccac gagttcttcc ctgacgccgc tctcgcgag 11760
 gcaagggaac tcgatgaata ctacgcaaag cacaagagac ccgttggtcc actccatggc 11820
 ctccccatct ctctcaaaga ccagcttcga gtcaagggtc acgaaacatc aatgggctac 11880
 atctcatggc taaacaagta cgacgaaggg gactcggttc tgacaaccat gctccgcaaa 11940
 gccggtgccg tcttctacgt caagacctct gtccgcaga ccctgatggt ctgcgagaca 12000
 gtcaacaaca tcatcgggcg caccgtcaac ccacgcaaca agaactggtc gtgcggcggc 12060
 agttctggtg gtgaggggtc gatcggtggg attcgtggtg gcgtcatcgg ttaggaacg 12120
 gatatcggtg gtcgattcg agtgccggcc gcgttcaact tcctgtacgg tctaaggccg 12180
 agtcatgggc ggctgccgta tgcaaagatg gcgaacagca tggaggggtca ggagacggtg 12240
 cacagcgttg tcgggccgat tacgcactct gttgaggacc tccgcctctt caccaaacc 12300
 gtctcggtc aggagccatg gaaatacgac tccaagggtca tccccatgcc ctggcgccag 12360
 tccgagtcgg acattattgc ctccaagatc aagaacggcg ggctcaatat cggctactac 12420
 aacttcgacg gcaatgtcct tccacaccct cctatcctgc gcggcgtgga aaccaccgtc 12480
 gccgcactcg ccaaagccgg tcacaccgtg acccgtgga cgccatacaa gcacgatttc 12540
 ggccacgac tcattctcca tatctacgcg gctgacggca gcgccgacgt aatgcgcgat 12600
 atcagtgcac ccggcgagcc ggcgattcca aatatcaaag acctactgaa cccgaacatc 12660
 aaagctgtta acatgaacga gctctgggac acgcatctcc agaagtggaa ttaccagatg 12720
 gagtaccttg agaaatggcg ggaggtgaa gaaaaggccg ggaaggaact ggacgccatc 12780
 atcgcgccga ttacgcctac cgctgcggta cggcatgacc agttccggta ctatgggtat 12840
 gcctctgtga tcaacctgct ggatttcacg agcgtggtt ttccggttac ctttgcggat 12900
 aagaacatcg ataagaagaa tgagagtttc aaggcgggtta gtgagcttga tgccctcgtg 12960
 caggaagagt atgatccgga ggcgtaccat ggggcaccgg ttgcagtgca ggttatcgga 13020
 cggagactca gtgaagagag gacgttggcg attgcagagg aagtggggaa gttgctggga 13080
 aatgtggtga ctccataggt cgagaattta tacttagata agtatgtact tacaggtata 13140

```

tttctatgag atactgatgt atacatgcat gataatatTT aaacggttat tagtgccgat 13200
tgtcttTgtgc gataatgacg ttcttatcaa agcaatacac ttaccaccta ttacatgggc 13260
caagaaaata ttttcgaact tgtttagaat attagcacag agtatatgat gatatccgtt 13320
agattatgca tgattcattc ctacaacttt ttcgtagcat aaggattaat tacttggatg 13380
ccaataaaaa aaaaaaacat cgagaaaatt tcagcatgct cagaaacaat tgcagtgtat 13440
caaagtaaaa aaaagatttt cgctacatgt tccttttgaa gaaagaaaat catggaacat 13500
tagatttaca aaaattttaac caccgctgat taacgattag accgttaagc gcacaacagg 13560
ttattagtac agagaaagca ttctgtggtg ttgccccgga ctttcttttg cgacataggt 13620
aaatcgaata ccatcatact atcttttcca atgactccct aaagaaagac tcttcttTga 13680
tgttTtatac gttggagcat agggcaagaa ttgtggcttg agatgaattc actggccgtc 13740
gttttacaac gtcgtgactg ggaaaaccct ggcgttacc cacttaatcg ccttgacgca 13800
catccccctt tcgccagctg gcgtaatagc gaagaggccc gcaccgatcg cccttcccaa 13860
cagttgTgca gcctgaatgg cgaatggcgc ctgatgcggt attttctcct tacgcatctg 13920
tgcggtatTT cacaccgcat atggtgcact ctcagtacaa tctgctctga tgccgcatag 13980
ttaagccagc cccgacaccc gccaacaccc gctgacgcgc cctgacgggc ttgtctgctc 14040
ccggcatccg cttacagaca agctgtgacc gtctccggga gctgcatgtg tcagaggttt 14100
tcaccgtcat caccgaaacg cgcgagacga aagggcctcg tgatacgctt atttttatag 14160
gttaatgtca tgataataat ggtttcttag acgtcaggtg gcacttttcg gggaaatgtg 14220
cgcggaaccc ctatttgttt atttttctaa atacattcaa atatgtatcc gtcTatgaga 14280
caataaccct gataaatgct tcaataatat tgaaaaagga agagtatgag tattcaacat 14340
ttccgtgtcg cccttattcc cttttttgcg gcattttgcc ttctgtttt tgctcaccca 14400
gaaacgctgg tgaaagtaaa agatgctgaa gatcagttgg gtgcacgagt gggttacatc 14460
gaactggatc tcaacagcgg taagatcctt gagagttttc gcccgaaga acgttttcca 14520
atgatgagca cttttaaaagt tctgctatgt ggcgcggtat tatcccgtat tgacgccggg 14580
caagagcaac tcggtcgccg catacactat tctcagaatg acttggttga gtactacca 14640
gtcacagaaa agcatcttac ggatggcatg acagtaagag aattatgcag tgctgccata 14700
accatgagtg ataacactgc ggccaactta cttctgacaa cgatcggagg accgaaggag 14760
ctaaccgctt ttttgacaaa catgggggat catgtaactc gccttgatcg ttgggaaccg 14820

```



```

gagctgaatg aagccatacc aaacgacgag cgtgacacca cgatgcctgt agcaatggca 14880
acaacgttgc gcaaactatt aactggcgaa ctacttactc tagcttcccg gcaacaatta 14940
atagactgga tggaggcgga taaagttgca ggaccacttc tgcgctcggc ccttccggct 15000
ggctggttta ttgctgataa atctggagcc ggtgagcgtg ggtctcgcgg tatcattgca 15060
gcactggggc cagatggtaa gccctcccgt atcgtagtta tctacacgac ggggagtcag 15120
gcaactatgg atgaacgaaa tagacagatc gctgagatag gtgcctcact gattaagcat 15180
tggtaaactgt cagaccaagt ttactcatat atactttaga ttgatttaaa acttcatttt 15240
taatttaaaa ggatctaggt gaagatcctt ttgataatc tcatgaccaa aatcccttaa 15300
cgtgagtttt cgttccactg agcgtcagac cccgtagaaa agatcaaagg atcttcttga 15360
gatccttttt ttctgcgcgt aatctgctgc ttgcaaacia aaaaaccacc gctaccagcg 15420
gtggtttgtt tgccggatca agagctacca actctttttc cgaaggtaac tggcttcagc 15480
agagcgcaga taccaaatac tgtccttcta gtgtagccgt agttaggcca ccacttcaag 15540
aactctgtag caccgcctac atacctcgtc ctgctaatac tgttaccagt ggctgctgcc 15600
agtggcgata agtcgtgtct taccgggttg gactcaagac gatagttacc ggataaggcg 15660
cagcggtcgg gctgaacggg gggttcgtgc acacagccca gcttgagcg aacgacctac 15720
accgaactga gatacctaca gcgtgagcta tgagaaagcg ccacgcttcc cgaagggaga 15780
aaggcggaca ggtatccggt aagcggcagg gtcggaacag gagagcgcac gagggagctt 15840
ccagggggaa acgcctggta tctttatagt cctgtcgggt ttcgccacct ctgacttgag 15900
cgtcgatttt tgtgatgctc gtcagggggg cggagcctat ggaaaaacgc cagcaacgcg 15960
gcctttttac ggttcctggc cttttgctgg ccttttgctc acatgttctt tctgctgta 16020
tcccctgatt ctgtggataa ccgtattacc gcctttgagt gagctgatac cgctcgccgc 16080
agccgaacga ccgagcgcag cgagtcagtg agcgaggaag cggaagagcg cccaatacgc 16140
aaaccgcctc tccccgcgcg ttggccgatt cattaatgca gctggcacga caggtttccc 16200
gactggaaaag cgggcagtga gcgcaacgca attaatgtga gttagctcac tcattaggca 16260
ccccaggctt tacactttat gcttccggct cgtatgttgt gtggaattgt gagcggataa 16320
caatttcaca caggaaacag ctatgaccat gattacgcc 16359

```

```

<210> 43
<211> 8684
<212> DNA
<213> Artificial

```

<220>

<223> pBOL113

<400> 43

tcgcgcggtt	cggtgatgac	ggtgaaaacc	tctgacacat	gcagctcccg	gagacggtca	60
cagcttgtct	gtaagcggat	gccgggagca	gacaagcccg	tcagggcgcg	tcagcgggtg	120
ttggcgggtg	tcggggctgg	cttaactatg	cggcatcaga	gcagattgta	ctgagagtgc	180
accataccac	agcttttcaa	ttcaattcat	catttttttt	ttattctttt	ttttgatttc	240
ggtttctttg	aaattttttt	gattcggtaa	tctccgaaca	gaaggaagaa	cgaaggaagg	300
agcacagact	tagattggta	tatatacgca	tatgtagtgt	tgaagaaaca	tgaaattgcc	360
cagtattctt	aaccaactg	cacagaacaa	aaacctgcag	gaaacgaaga	taaatacatgt	420
cgaagctac	atataaggaa	cgtgctgcta	ctcatcctag	tcctgttgct	gccaagctat	480
ttaatatcat	gcacgaaaag	caaacaaact	tgtgtgcttc	attggatggt	cgtaccacca	540
aggaattact	ggagttagtt	gaagcattag	gtcccaaaat	ttgtttacta	aaaacacatg	600
tggatatctt	gactgatttt	tccatggagg	gcacagttaa	gccgctaaag	gcattatccg	660
ccaagtacaa	ttttttactc	ttcgaagaca	gaaaatttgc	tgacattggt	aatacagtca	720
aattgcagta	ctctgcgggt	gtatacagaa	tagcagaatg	ggcagacatt	acgaatgcac	780
acgggtgtgt	gggcccaggt	attgttagcg	gtttgaagca	ggcggcagaa	gaagtaacaa	840
aggaacctag	aggccttttg	atgttagcag	aattgtcatg	caagggctcc	ctatctactg	900
gagaatatac	taagggtact	gttgacattg	cgaagagcga	caaagatttt	gttatcggct	960
ttattgctca	aagagacatg	ggtggaagag	atgaaggtta	cgattggttg	attatgacac	1020
ccggtgtggg	tttagatgac	aaggagagcg	cattgggtca	acagtataga	accgtggatg	1080
atgtggtctc	tacaggatct	gacattatta	ttgttggaag	aggactatth	gcaaagggaa	1140
gggatgctaa	ggtagagggg	gaacgttaca	gaaaagcagg	ctgggaagca	tatttgagaa	1200
gatgcggcca	gcaaaactaa	aaaactgtat	tataagtaaa	tgcatgtata	ctaaactcac	1260
aaattagagc	ttcaatttaa	ttatatcagt	tattacccta	tgcggtgtga	aataccgcac	1320
agatgcgtaa	ggagaaaata	ccgcatcagg	aaattgtaaa	cgttaaatatt	ttgttaaaat	1380
tcgcgttaaa	tttttgttaa	atcagctcat	tttttaacca	ataggccgaa	atcggaacaa	1440
tcccttataa	atcaaaagaa	tagaccgaga	taggggtgag	tggtgttcca	gtttggaaca	1500
agagtccact	attaaagaac	gtggactcca	acgtcaaagg	gcgaaaaacc	gtctatcagg	1560

gcgatggccc	actacgtgaa	ccatcacccct	aatcaagttt	tttggggtcg	aggtgccgta	1620
aagcactaaa	tcggaaccct	aaagggagcc	cccgathtag	agcttgacgg	ggaaagccgg	1680
cgaacgtggc	gagaaaggaa	gggaagaaag	cgaaaggagc	gggcgctagg	gcgctggcaa	1740
gtgtagcgg	cacgctgcgc	gtaaccacca	caccgcgcgc	gcttaatgcg	ccgctacagg	1800
gcgctgcgc	ccattcgcca	ttcaggctgc	gcaactgttg	ggaagggcga	tcgggtgcggg	1860
cctcttcgct	attacgccag	ctggcgaaag	ggggatgtgc	tgcaaggcga	ttaagttggg	1920
taacgccagg	gttttcccag	tcacgacggt	gtaaaacgac	ggccagtgag	cgcgcgtaat	1980
acgactcact	atagggcgaa	ttgggtaccg	ggccccccct	cgaggtcgac	ggtatcgata	2040
agcttgatat	cgaattcctg	cagcccgggg	gatccactag	ttctagagcg	gccattttaa	2100
acggccggcc	ctagatcaga	gggtggtaaa	tgaagtgtaa	tagtattcat	ttttcttata	2160
aatcatccct	tccgtgattt	atacaaaaga	agaggagaat	atgctgaata	cttggtatat	2220
tactctacat	tatactctta	tcttgacggg	tattctgagc	atcttactca	gtttcaagat	2280
cttttaaatgt	ccaaaaacat	ttgagccgat	ctaaatactt	ctgtgttttc	attaatttat	2340
aaattgtact	cttttaagac	atggaaagta	ccaacatcgg	ttgaaacagt	ttttcattta	2400
cttatggttt	attggttttt	ccagtgaatg	attatttgtc	gttacccttt	cgtaaaagtt	2460
caaacacggt	tttaagtatt	gttttagttgc	tctttcgaca	tatatgatta	tccctgcgcg	2520
gctaaagtta	aggatgcaaa	aaacataaga	caactgaagt	taatttacgt	caattaagtt	2580
ttccagggta	atgatgtttt	gggcttccac	taattcaata	agtatgtcat	gaaatacgtt	2640
gtgaagagca	tccagaaata	atgaaaagaa	acaacgaaac	tgggtcggcc	tgttgtttct	2700
tttctttacc	acgtgatctg	cggcatttac	aggaagtcgc	gcgttttgcg	cagttgttgc	2760
aacgcagcta	cggctaacaa	agcctagtgg	aactcgactg	atgtgttagg	gcctaaaact	2820
ggtggtgaca	gctgaagtga	actattcaat	ccaatcatgt	catggctgtc	acaaagacct	2880
tgcgggaccg	acgtacgaac	acatacgtat	gctaatatgt	gttttgatag	taccagtgta	2940
tcgcagacct	gcaatttttt	tgtaggtttg	gaagaatata	taaaggttgc	actcattcaa	3000
gatagttttt	ttcttggtgtg	tctattcatt	ttattattgt	ttgtttaaat	gttaaaaaaa	3060
ccaagaactt	agtttcaaat	taaattcatc	acacaaacaa	acaaaacaaa	atgaacattg	3120
ttgtttgttt	gaagcaagtt	ccagacactg	ctgaagtcag	aattgacca	gtcaagggta	3180
ctttaatcag	agaaggtggt	ccatctatca	tcaaccaga	cgacaagaac	gctttggaag	3240
aagctttggg	tttgaaggac	aactacggtg	ctcacgttac	cgtcatttcc	atgggtccac	3300

ctcaagccaa gaacgctttg gttgaagctt tggccatggg tgctgatgaa gctgtcttat	3360
tgactgacag agctttcggg ggtgctgata ctttagctac ctctcacacc attgctgctg	3420
gatatcaagaa attgaaatac gatatcgtct ttgccggctg tcaagccatc gatggtgata	3480
ccgctcaagt cgggccagaa attgctgaac atttgggtat tccacaagtc acctacgttg	3540
aaaaggttga agttgacggg gacactttga agatcagaaa ggcttgggaa gacggttacg	3600
aagttgttga agtcaagact ccagttctat tgactgccat caaggaattg aacgttccaa	3660
gatacatgtc cgttgaaaaag atcttcgggtg ctttcgacaa ggaagtcaag atgtggactg	3720
ctgatgatat cgatgtcgac aaggccaact tgggtttgaa aggttctcca accaaggtca	3780
agaaatcttc taccaaggaa gtcaaggggc aaggtgaagt cattgacaaa ccagtcaagg	3840
aagctgccgc ttacgttggt tccaagttga aggaagaaca ctacatctaa agcccgggcg	3900
gagattgata agacttttct agttgcatat cttttatatt taaatcttat ctattagtta	3960
atTTTTTgta atttatcctt atatatagtc tggttattct aaaatatcat ttcagtatct	4020
aaaaattccc ctctTTTTtc agttatatct taacaggcga cagtccaaat gttgatttat	4080
cccagtccga ttcatcaggg ttgtgaagca ttttgtcaat ggtcgaaatc acatcagtaa	4140
tagtgcctct tacttgccctc atagaatttc tttctcttaa cgtcaccgtt tggcttttta	4200
tagtttcgaa atctatgggtg ataccaaatg gtgttcccaa ttcatcgtaa cgggcgtatt	4260
ttttaccaat tgaagtattg gaatcgtaa ttttaaagta tatctctctt ttacgtaaag	4320
cctgcgagat cctcttaagt atagcgggga agccatcggt attcgatatt gtcgtaacaa	4380
atactttgat cggcgctatg tttaaatggt taaacatgga cagatatgcg atgaaaacgc	4440
taagtgatac tccaaatggt gaaaggtacg atgcttgga acaatacttg gaaatcaccg	4500
gaaacaccat atgcggcgaa aagccaatta gtgtgatact aagtgcttta tcgaaaatcc	4560
gtgatgccgg tccttcaggc atcaaatttc agtggcctaa ttattcacag agttctcatg	4620
tgacaagtat tgatgatagt agtgtcagtt atgcttcagg ttatgttact ataggataat	4680
gatcacggct aaaacggctg aatgtaagca tatactcttc gattgtataa ttgttcccaa	4740
atactacagc atctcaagga aaaaaaaca aaaacttcca aaaaaatcga atccctgagg	4800
aatctttaat acatTTTcaa tctatttaag ttttataaac gtgtatatga gatgtcatga	4860
gcatgaatta ttaataataa aaactaaatc attaaagtaa ctttaaggagt taaagcccgg	4920
gctttaattg ttagcagcct tgacttgagc aatcaattct ggaacaacct tgttgacatc	4980

accgacaatg	gccaaatcag	caaccttcat	gattggagct	tcgacatctt	tgttgatggc	5040
aatgatgtag	tcagagtctt	gcataccage	caagtgttgg	atggcaccag	agataccaca	5100
agcaatgtac	aaagttggtc	tgacggtctt	accggtttga	ccgacttgca	agtccttgtc	5160
aaccttcc	ttttcaatgg	cagctctgga	agcagcaatg	gtaccacca	aaaagaagc	5220
taattcttcc	aatttttcga	agttttcctt	ggaaccaaca	ccacgaccac	cagcaacca	5280
aaccttggct	tcaccgatat	cagcaatgtc	cttggccaat	ttgacaacct	tggaaacctt	5340
ggttctgata	tcagaagcag	tcaatttgat	ggcaaccttt	tcgatcttgt	catcagaaac	5400
gttagcatcg	ttaactggca	atttttcaaa	gacacctggg	ctgacggtgg	ccatttgagg	5460
tctgtggtca	gaacagacaa	tggtagcaat	caagttacca	ccgaaagctg	gtctggtagc	5520
caacaagtca	cggttttcga	catcgatata	caaagaggta	cagtcagcag	tcaaaccagt	5580
agacaatctg	gcagcaattc	ttggacccaa	gtctctaccg	atgaaagtag	caccgatgaa	5640
taagatttct	ggctttcttt	cgttgaccaa	gtcacagata	accttggcgt	aaccgtcagt	5700
ggagaaatga	gctaataatt	cgttgtcagc	agccaaaacc	ttgtcagcac	cgtgggacaa	5760
caagtccttg	gacatctttt	cagtgttgtg	accaataag	acagcagtca	attcaacacc	5820
caatttttca	gccatttcct	tacccttacc	tagcaattcc	aaagaaacct	tttgaattc	5880
accatctctt	tgttcagcga	aaaccagac	acccttgtag	tcagccttgt	tcattgttag	5940
ttaattatag	ttcgttgacc	gtatatctta	aaaacaagta	ctccttaaaa	aaaaaccttg	6000
aagggaataa	acaagtagaa	tagatagaga	gaaaaataga	aatgcaaga	gaatttatat	6060
attagaaaga	gagaaagaaa	aatggaaaaa	aaaaaatagg	aaaagccaga	aatagcacta	6120
gaaggagcga	caccagaaaa	gaagggtgatg	gaaccaattt	agctatatat	agttaactac	6180
cggctcgatc	atctctgcct	ccagcatagt	cgaagaagaa	tttttttttt	cttgaggctt	6240
ctgtcagcaa	ctcgtatttt	ttctttcttt	tttggtgagc	ctaaaaagtt	cccacgttct	6300
cttgtacgac	gccgtcacaa	acaaccttat	gggtaatttg	tcgcggtctg	ggtgtataaa	6360
tgtgtgggtg	caggccggcc	gtttaaacgg	gccgccaccg	cggtggagct	ccagcttttg	6420
ttccctttag	tgagggttaa	ttgcgcgctt	ggcgtaatca	tggtcatagc	tgtttctgt	6480
gtgaaattgt	tatccgctca	caattccaca	caacatagga	gccggaagca	taaagtgtaa	6540
agcctggggg	gcctaagtga	tgaggtaact	cacattaatt	gcgttgcgct	cactgcccgc	6600
ttccagtcg	ggaaacctgt	cgtgccagct	gcattaatga	atcgccaac	gcgcggggag	6660
aggcggtttg	cgtattgggc	gtctttccgc	ttctctcgctc	actgactcgc	tgcgctcggt	6720

cgttcggctg	cggcgagcgg	tatcagctca	ctcaaaggcg	gtaatacggg	tatccacaga	6780
atcaggggat	aacgcaggaa	agaacatgtg	agcaaaaggc	cagcaaaagg	ccaggaaccg	6840
taaaaaggcc	gcgttgctgg	cgtttttcca	taggctccgc	ccccctgacg	agcatcacia	6900
aaatcgacgc	tcaagtcaga	ggtggcgaaa	cccgcacagga	ctataaagat	accaggcggt	6960
tccccctgga	agctccctcg	tgcgctctcc	tgttccgacc	ctgccgctta	ccggatacct	7020
gtccgccttt	ctcccttcgg	gaagcgtggc	gctttctcat	agctcacgct	gtaggtatct	7080
cagttcgggtg	taggtcgttc	gctccaagct	gggctgtgtg	cacgaacccc	ccgttcagcc	7140
cgaccgctgc	gccttatccg	gtaactatcg	tcttgagtcc	aacccggtaa	gacacgactt	7200
atcgccactg	gcagcagcca	ctggtaacag	gattagcaga	gcgaggtatg	taggcggtgc	7260
tacagagttc	ttgaagtggg	ggcctaacta	cggctacact	agaaggacag	tatttggtat	7320
ctgcgctctg	ctgaagccag	ttaccttcgg	aaaaagagtt	ggtagctctt	gatccggcaa	7380
acaaaccacc	gctggtagcg	gtgggtttttt	tgtttgcaag	cagcagatta	cgcgagaaaa	7440
aaaaggatct	caagaagatc	ctttgatctt	ttctacgggg	tctgacgctc	agtggaacga	7500
aaactcacgt	taagggatth	tggatcatgag	attatcaaaa	aggatcttca	cctagatcct	7560
tttaaattaa	aatgaagtt	ttaaataaat	ctaaagtata	tatgagtaaa	cttggtctga	7620
cagttaccaa	tgtttaatac	gtgaggcacc	tatctcagcg	atctgtctat	ttcgttcatc	7680
catagttgcc	tgactccccg	tcgtgtagat	aactacgata	cgggaggggt	taccatctgg	7740
ccccagtgtc	gcaatgatac	cgcgagaccc	acgctcaccg	gctccagatt	tatcagcaat	7800
aaaccagcca	gccggaagg	ccgagcgcag	aagtggctct	gcaactttat	ccgcctccat	7860
ccagtctatt	aattgttgcc	gggaagctag	agtaagtagt	tcgccagtta	atagtttgcg	7920
caacgttggt	gccattgcta	caggcatcgt	ggtgtcacgc	tcgtcgtttg	gtatggcttc	7980
attcagctcc	ggttcccaac	gatcaaggcg	agttacatga	tccccatgt	tgtgcaaaaa	8040
agcggtagc	tccttcgggc	ctccgatcgt	tgtcagaagt	aagttggccg	cagtgttatc	8100
actcatgggt	atggcagcac	tgcataatcc	tcttactgtc	atgccatccg	taagatgctt	8160
ttctgtgact	ggtgagtact	caaccaagtc	attctgagaa	tagtgtatgc	ggcgaccgag	8220
ttgtctcttg	ccggcgtcaa	tacgggataa	taccgcgcca	catagcagaa	ctttaaaggt	8280
gctcatcatt	ggaaaacgtt	cttcggggcg	aaaactctca	aggatcttac	cgctgttgag	8340
atccagttcg	atgtaaccca	ctcgtgcacc	caactgatct	tcagcatctt	ttactttcac	8400

cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc 8460
 gacacggaaa tgttgaatac tcatactctt cctttttcaa tattattgaa gcatttatca 8520
 gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg 8580
 ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc taagaaacca ttattatcat 8640
 gacattaacc tataaaaaata ggcgtatcac gaggcccttt cgtc 8684

<210> 44
 <211> 12314
 <212> DNA
 <213> Artificial

<220>
 <223> pBOL115

<400> 44
 tcgcgcggtt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca 60
 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
 ttggcgggtg tcggggtg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
 accataccac agcttttcaa ttcaattcat cttttttttt ttattctttt ttttgatttc 240
 ggtttctttg aaattttttt gattcggtaa tctccgaaca gaaggaagaa cgaaggaagg 300
 agcacagact tagattggta tatatacgca tatgtagtgt tgaagaaaca tgaaattgcc 360
 cagtattctt aacccaactg cacagaacaa aaacctgcag gaaacgaaga taaatcatgt 420
 cgaaagctac atataaggaa cgtgctgcta ctcatcctag tcctgttgct gccaaagctat 480
 ttaatatcat gcacgaaaag caaacaact tgtgtgcttc attggatgtt cgtaccacca 540
 aggaattact ggagttagtt gaagcattag gtcccaaat ttgtttacta aaaacacatg 600
 tggatatctt gactgatttt tccatggagg gcacagttaa gccgctaaag gcattatccg 660
 ccaagtacaa ttttttactc ttcgaagaca gaaaatttgc tgacattgggt aatacagtca 720
 aattgcagta ctctgcgggt gtatacagaa tagcagaatg ggcagacatt acgaatgcac 780
 acggtgtggg gggcccagggt attgttagcg gtttgaagca ggcggcagaa gaagtaacaa 840
 aggaacctag aggccttttg atgttagcag aattgtcatg caagggctcc ctatctactg 900
 gagaatatac taagggtact gttgacattg cgaagagcga caaagatttt gttatcggtc 960
 ttattgctca aagagacatg ggtggaagag atgaaggtta cgattggttg attatgacac 1020
 ccggtgtggg tttagatgac aaggagagcg cattgggtca acagtataga accgtggatg 1080
 atgtggtctc tacaggatct gacattatta ttgttgaag aggactatth gcaaagggaa 1140

gggatgctaa	ggtagaggggt	gaacggttaca	gaaaagcagg	ctgggaagca	tatttgagaa	1200
gatgcggccca	gcaaaaactaa	aaaactgtat	tataagtaaa	tgcatgtata	ctaaactcac	1260
aaattagagc	ttcaatttta	ttatatcagt	tattacccta	tgcggtgtga	aataccgcac	1320
agatgcgtaa	ggagaaaata	ccgcatcagg	aaattgtaaa	cgtaaataatt	ttgttaaaat	1380
tcgcgttaaa	tttttgttaa	atcagctcat	tttttaacca	ataggccgaa	atcggcaaaa	1440
tcccttataa	atcaaaaagaa	tagaccgaga	taggggtgag	tggtgttcca	gtttggaaca	1500
agagtccact	attaaagaac	gtggactcca	acgtcaaagg	gcgaaaaacc	gtctatcagg	1560
gcgatggccc	actacgtgaa	ccatcacccct	aatcaagttt	tttggggtcg	aggtgccgta	1620
aagcactaaa	tcggaaccct	aaagggagcc	cccgathtag	agcttgacgg	ggaaagccgg	1680
cgaacgtggc	gagaaaaggaa	gggaagaaaag	cgaaaggagc	gggcgctagg	gcgctggcaa	1740
gtgtagcgggt	cacgctgcgc	gtaaccacca	caccgcgcgc	gcttaatgcg	ccgctacagg	1800
gcgcgtcgcg	ccattcgcca	ttcaggctgc	gcaactgttg	ggaagggcga	tcggtgcggg	1860
cctcttcgct	attacgccag	ctggcgaaaag	ggggatgtgc	tgcaaggcga	ttaagttggg	1920
taacgccagg	gttttcccag	tcacgacggt	gtaaaacgac	ggccagttag	cgcgcgtaat	1980
acgactcact	atagggcgaa	ttgggtaccg	ggccccccct	cgaggtcgac	ggtatcgata	2040
agcttgatat	cgaattcctg	cagcccgggg	gatccactag	ttctagagcg	gccatttta	2100
acggccggcc	ctagatcaga	gggtggtaaa	tgaagtgtaa	tagtattcat	ttttcttata	2160
aatcatccct	tccgtgattt	atacaaaaga	agaggagaat	atgctgaata	cttggtatat	2220
tactctacat	tatactctta	tcttgacggg	tattctgagc	atcttactca	gtttcaagat	2280
cttttaaatgt	ccaaaaacat	ttgagccgat	ctaaatactt	ctgtgttttc	attaatttat	2340
aaattgtact	cttttaagac	atggaaagta	ccaacatcgg	ttgaaacagt	ttttcattta	2400
cttatggttt	attgggtttt	ccagtgaatg	attatttgtc	gttacccttt	cgtaaaagtt	2460
caaacacggt	tttaagtatt	gtttagttgc	tctttcgaca	tatatgatta	tccctgcgcg	2520
gctaaagtta	aggatgcaaa	aaacataaga	caactgaagt	taatttacgt	caattaagtt	2580
ttccagggtta	atgatgtttt	gggcttccac	taattcaata	agtatgtcat	gaaatacgtt	2640
gtgaagagca	tccagaaata	atgaaaagaa	acaacgaaac	tgggtcggcc	tggtgtttct	2700
tttctttacc	acgtgatctg	cggcatttac	aggaagtcgc	gcgttttgcg	cagttgttgc	2760
aacgcagcta	cggctaacaa	agcctagtg	aactcgactg	atgtgttagg	gcctaaaact	2820

ggtggtgaca gctgaagtga actattcaat ccaatcatgt catggctgtc acaaagacct	2880
tgcggaccgc acgtacgaac acatacgtat gctaatatgt gttttgatag taccagtgga	2940
tcgcagacct gcaatTTTTT tgtaggTTTg gaagaatata taaaggTTgc actcattcaa	3000
gatagTTTTT ttcttTgtgtg tctattcatt ttattattgt ttgtttaaat gttaaaaaaa	3060
ccaagaactt agtttcaaTt taaattcatc acacaaacaa acaaaacaaa atgaacattg	3120
ttgtttgttt gaagcaagtt ccagacactg ctgaagtcag aattgacca gtcaagggtg	3180
ctttaatcag agaaggTgtt ccatctatca tcaaccaga cgacaagaac gctttggaag	3240
aagctttggt tttgaaggac aactacggtg ctcacgttac cgtcatttcc atgggtccac	3300
ctcaagccaa gaacgctttg gttgaagctt tggccatggg tgctgatgaa gctgtcttat	3360
tgactgacag agctttcggT ggtgctgata ctttagctac ctctcacacc attgctgctg	3420
gtatcaagaa attgaaatac gatatcgtct ttgccggtcg tcaagccatc gatggtgata	3480
ccgctcaagt cggTccagaa attgctgaac atttggttat tccacaagtc acctacgttg	3540
aaaaggTtga agttgacggt gacactttga agatcagaaa ggcttgggaa gacggttacg	3600
aagttgttga agtcaagact ccagttctat tgactgccat caaggaattg aacgttccaa	3660
gatacatgtc cgTtgaaaag atcttcggtg ctttcgacaa ggaagtcaag atgtggactg	3720
ctgatgatat cgatgtcgac aaggccaact tgggtttgaa aggttctcca accaaggTca	3780
agaaatcttc taccaaggaa gtcaagggtc aaggTgaagt cattgacaaa ccagtcaagg	3840
aagctgccgc ttacgttgTt tccaagTtga aggaagaaca ctacatctaa agcccgggcg	3900
gagattgata agacttttct agttgcatat cttttatatt taaatcttat ctattagtta	3960
atTTTTgtta atttatcctt atatatagtc tggTtattct aaaatatcat ttcagtatct	4020
aaaaattccc ctctTTTTtc agttatatct taacaggcga cagtccaaat gttgatttat	4080
cccagtccga ttcatcaggg ttgtgaagca ttttTgtcaat ggtcgaaatc acatcagtaa	4140
tagtgctctc tacttgctc atagaatttc tttctcttaa cgtcacctgt tggTctttta	4200
tagtttcgaa atctatggTg ataccaaTg gtgttcccaa ttcatcgTta cgggcgtatt	4260
ttttaccaat tgaagtattg gaatcgTcaa ttttaaagta tatctctctt ttacgtaaag	4320
cctgcgagat cctcttaagT atagcgggga agccatcgTt attcgatatt gtcgtaacaa	4380
atactttgat cggcgctatg tttaaatgtt taaacatgga cagatatgcg atgaaaacgc	4440
taagtgtac tccaaatggT gaaaggTacg atgcttgga acaatacttg gaaatcaccg	4500
gaaacaccat atgcggcgaa aagccaatta gtgtgatact aagtgcTtta tcgaaaatcc	4560

gtgatgccgg	tccttcaggc	atcaaatttc	agtggcctaa	ttattcacag	agttctcatg	4620
tgacaagtat	tgatgatagt	agtgtcagtt	atgcttcagg	ttatgttact	ataggataat	4680
gatcacggct	aaaacggtcg	aatgtaagca	tatatctttc	gattgtataa	ttgttcccaa	4740
atactacagc	atctcaagga	aaaaaaaaa	aaaacttcca	aaaaaatcga	atccctgagg	4800
aatctttaat	acattttcaa	tctattttaag	ttttataaac	gtgtatatga	gatgtcatga	4860
gcatgaatta	ttaataataa	aaactaaatc	attaaagtaa	cttaaggagt	taaagcccg	4920
gctttaattg	ttagcagcct	tgacttgagc	aatcaattct	ggaacaacct	tgttgacatc	4980
accgacaatg	gccaaatcag	caaccttcat	gattggagct	tcgacatctt	tgttgatggc	5040
aatgatgtag	tcagagtctt	gcataccagc	caagtgttgg	atggcaccag	agataccaca	5100
agcaatgtac	aaagttagtc	tgacgggtct	accggtttga	ccgacttgca	agtccttgtc	5160
aaccatttcc	ttttcaatgg	cagctctgga	agcagcaatg	gtaccacca	acaagaagc	5220
taattcttcc	aatttttcga	agttttcctt	ggaaccaaca	ccacgaccac	cagcaaccaa	5280
aaccttggtc	tcaccgatat	cagcaatgtc	cttggccaat	ttgacaacct	tggaaacctt	5340
ggttctgata	tcagaagcag	tcaatttgat	ggcaaccttt	tcgatcttgt	catcagaaac	5400
gttagcatcg	ttaactggca	atttttcaaa	gacacctggt	ctgacggtgg	ccatttgagg	5460
tctgtggtca	gaacagacaa	tggtagcaat	caagttacca	ccgaaagctg	gtctggtagc	5520
caacaagtca	cggttttcga	catcgatatc	caaagaggta	cagtcagcag	tcaaaccagt	5580
agacaatctg	gcagcaattc	ttggacccaa	gtctctaccg	atgaaagtag	caccgatgaa	5640
taagatttct	ggctttcttt	cgttgaccaa	gtcacagata	accttggcgt	aaccgtcagt	5700
ggagaaatga	gctaataatt	cgttgtcagc	agccaaaacc	ttgtcagcac	cgtgggacaa	5760
caagtccttg	gacatctttt	cagtgttgtg	acccaataag	acagcagtca	attcaacacc	5820
caatttttca	gccatttcct	tacccttacc	tagcaattcc	aaagaaacct	tttgtaattc	5880
accatctctt	tgttcagcga	aaaccagac	accctttag	tcagccttgt	tcatgtttag	5940
ttaattatag	ttcgttgacc	gtatattcta	aaaacaagta	ctccttaaaa	aaaaaccttg	6000
aagggaataa	acaagtagaa	tagatagaga	gaaaaataga	aatgcaaga	gaatttatat	6060
attagaaaga	gagaaagaaa	aatggaaaaa	aaaaaatagg	aaaagccaga	aatagcacta	6120
gaaggagcga	caccagaaaa	gaaggtagatg	gaaccaatth	agctatatat	agttaactac	6180
cggctcgatc	atctctgcct	ccagcatagt	cgaagaagaa	tttttttttt	cttgaggctt	6240

ctgtcagcaa	ctcgtatttt	ttcttttcttt	tttggtgagc	ctaaaaagtt	cccacgttct	6300
cttgtacgac	gccgtcacaa	acaaccttat	gggtaatttg	tcgcggtctg	ggtgtataaa	6360
tgtgtgggtg	caggccggcc	gtttaaacgg	gccgccaccg	cgggtggagcc	tgtgtggaag	6420
aacgattaca	acaggtgttg	tcctctgagg	acataaaaata	cacaccgaga	ttcatcaact	6480
cattgctgga	gttagcatat	ctacaattgg	gtgaaatggg	gagcgatttg	caggcatttg	6540
ctcggcatgc	cggtagaggt	gtggtcaata	agagcgacct	catgctatac	ctgagaaaagc	6600
aacctgacct	acaggaaaga	gttactcaag	aataagaatt	ttcgttttaa	aacctaaagag	6660
tcactttaaa	atttgataac	acttattttt	tttataactt	atttaataat	aaaaatcata	6720
aatcataaga	aattcgctcg	agtcgactgc	agtttactgc	ttgtagtcgt	aagaagtttg	6780
gatgatatcc	ttgatttcag	aaatcaaagc	ttcctttggg	ttggcagtg	tacattggtc	6840
ttcgaaagcc	aattcagcca	ttctgtcaat	ggattcgttc	aattcttctt	cagagacacc	6900
ttgagatttc	aagttcattt	caataccaac	agattgacct	aattcgtaga	cagccttggc	6960
caaagattca	accaaagctt	cagtgggtgt	acctttcaaa	cctaagaact	tggcgatatac	7020
agcgtaatcg	gtgtcagctc	tgaagaattc	gtactttggg	aacaaagcgt	gcttttgagg	7080
gtccttggcg	ttgtatctga	tgatgtgagg	caacaagatg	gcgttagctc	taccatgtgg	7140
aataccgtat	tcaccaccaa	ttttgtgagc	aatggagtga	gcaataccca	agaaagcgtt	7200
agcaaaggcc	ataccggcca	aagtagaagc	gttgtgcatac	ttttctctgg	aaaccttgctc	7260
acctttttca	acggaagatt	tcaagtattc	aaaggtcaat	ttgatagctt	gtagggacaa	7320
acctctgggtg	taatcggagg	ccatgacaga	aacgtaagat	tccatagcgt	gagtcaaaac	7380
gtccataaccg	gtatcagcag	tgacagattt	tgggacggac	atgacaaatt	gagggtcaat	7440
gatggcgaca	tctggagtca	aagcgaaatc	agccaatggg	tatttgacgt	tggtttcaga	7500
gtcagtgata	acagcaaata	gagtgacttc	agaaccagta	ccagaagtgg	ttggaatata	7560
gatgaaagtg	gcgtttttctg	gcataccaat	cttgtaggtt	ctcttaccac	tgtccaagaa	7620
tttttgcttg	gcaccgaaga	aagaagtttc	tggatgttcg	aagaacatcc	acatagcctt	7680
ggcagcatcc	atggcagaac	caccacccaa	agcaatgatg	gtgtctgggt	ggaaatcgac	7740
catcatttcc	aaaccttgt	agacagtgtt	ggtggatgga	tttggttcaa	cttcagagaa	7800
gatcttgatt	tgaggttgtt	cagttctttg	acgtaagacg	ttttcgacgg	ttttggtgta	7860
accaaattca	accataacctg	ggtcacaaac	gatcatgacc	ttttcaatct	tgtccatggt	7920
ggtcaaggac	atgatagcgt	tttcttcgaa	atagatttga	gctggaacct	tgaagatttg	7980

agtgttgttt cttctcttgg caatggtctt gatgttcaac aaatcggtag cagagacggt	8040
gtgggagatg gagtttctac cgtaagaacc acaacccaaa gtcaaggatg gaatcaattc	8100
gttgtacatg tcaccgatac caccaacagc agatggagtg ttgaccaaaa cagacaagc	8160
cttcattctt agaccgaaat ccttttgcaa agtttcgtct tcagtgtgga taacagcagt	8220
gtgacctaaa ccaccgaagt gcaaagtgtc ttcacagatt tggaaagctt gcttggtaga	8280
ttgagccttg actaaagcca aaactggaga caacttttct ctggacaatg ggtagtcaga	8340
accgacaccg gagatttcag caatgatcaa tttggtgttt tctggaactg ggataccggc	8400
caattcagca atttcaacag cagacttacc aacgatatct ggcttgatac cagtcttttg	8460
ttcgttcattg atggcgtttt ctaatctttg taattcgtcc ttcttgacga agtaagcttg	8520
gtgagccttg aattcattgg tgacatcctt gtagatttcc ttgtcaatga caacaacttg	8580
ttcagaagca cagatcatac cattatcgaa agtcttggaa ccgatgatat cgttgacagc	8640
acgcttgata tgagcagtct tttcgatgta agatggaacg ttacctggac caacacccaa	8700
agctggttta ccagtggagt aagcagactt aaccatacca gaaccaccgg tagccaagac	8760
taaagcaata cccttgtggg tcatcaattg cttggtagct tcaatggatg gaacttcaat	8820
ccattggatg atatcctttg gagcaccagc cttcatggca gcttccaaaa caacttcagc	8880
agctctcttg gaagattctt gagcagatgg gtggaaagcg aaaataattg ggttaccagt	8940
cttgatggca atcatagcct tgaagatggg ggtagaagtt ggggttggtg ttggagtgc	9000
accacagatg acaccaattg gttcagcaac gtaggtcaaa cccttttctt tgtcttcacc	9060
aatgatacca acagtcttgg tgccttgat ggagttccag atgtattcag aggcgtataa	9120
gttcttgata gccttgtctt cgtagatacc tctaccggtt tcttcatgag ccaacttggc	9180
caaaaccatg tgttggtcaa cagcagccaa ggacatttgg tggacaatgt ggtcaatttc	9240
ttcttgagac ttcttgga aagcttccaa agccttctta cctttgtcag cttaaagcatc	9300
aatcatgatg gcaacttctt gttccttggg acctctgttt tccttttctg gaatggtcaa	9360
cattttttac tagttctaga atccgtcgaa actaagttct ggtgttttaa aactaaaaaa	9420
aagactaact ataaaagtag aatttaagaa gtttaagaaa tagatttaca gaattacaat	9480
caatacctac cgtctttata tacttattag tcaagtaggg gaataatttc aggggaactgg	9540
tttcaacctt ttttttcagc tttttccaaa tcagagagag cagaaggtaa tagaaggtgt	9600
aagaaaatga gatagataca tgcgtaggtc aattgccttg tgtcatcatt tactccaggc	9660

aggttgcatc actccattga gggtgtgccc gttttttgcc tgtttggtgcc cctgttctct 9720
 gtagttgcgc taagagaatg gacctatgaa ctgatggttg gtgaagaaaa caatatatttg 9780
 gtgctgggat tctttttttt tctggatgcc agcttaaaaa gcgggctcca ttatatattag 9840
 tggatgccag gaataaactg ttcacccaga cacctacgat gttatatatt ctgtgtaacc 9900
 cgccccctat tttgggcatg tacgggttac agcagaatta aaaggctaatt tttttgacta 9960
 aataaagtta ggaaaatcac tactattaat tatttacgta ttctttgaaa tggcgagtat 10020
 tgataatgat aaactgagct ccagcttttg ttcccttttag tgagggttaa ttgcgcgctt 10080
 ggcgtaatca tggatcatagc tgtttcctgt gtgaaattgt tatccgctca caattccaca 10140
 caacatagga gccggaagca taaagtgtaa agcctggggg gcctaataag tgaggtaact 10200
 cacattaatt gcgttgcgct cactgcccgc tttccagtcg ggaaacctgt cgtgccagct 10260
 gcattaatga atcggccaac gcgcggggag aggcgggttg cgtattgggc gctcttccgc 10320
 ttctctgctc actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca 10380
 ctcaaaggcg gtaatacggg tatccacaga atcaggggat aacgcaggaa agaacatgtg 10440
 agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca 10500
 taggctccgc cccctgacg agcatcacia aaatcgacgc tcaagtcaga ggtggcgaaa 10560
 cccgacagga ctataaagat accaggcggt tccccctgga agctccctcg tgcgctctcc 10620
 tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc 10680
 gctttctcat agctcacgct gtaggtatct cagttcgggtg taggtcggtc gctccaagct 10740
 gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg 10800
 tcttgagtcc aaccgggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag 10860
 gattagcaga gcgaggtatg taggcgggtg tacagagttc ttgaagtggg gccctaacta 10920
 cggctacact agaaggacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg 10980
 aaaaagagtt ggtagctctt gatccggcaa acaaacacc gctggtagcg gtggtttttt 11040
 tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt 11100
 ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt tggatcatgag 11160
 attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaataaat 11220
 ctaaagtata tatgagtaaa cttgggtctga cagttaccaa tgcttaataca gtgaggcacc 11280
 tatctcagcg atctgtctat ttctgtctac catagttgcc tgactccccg tcgtgtagat 11340
 aactacgata cgggaggggc taccatctgg cccagtgct gcaatgatac cgcgagaccc 11400

```

acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag 11460
aagtggtcct gcaactttat ccgcctccat ccagtctatt aattggtgcc gggaagctag 11520
agtaagtagt tcgccagtta atagtttgcg caacgttggt gccattgcta caggcatcgt 11580
ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg 11640
agttacatga tcccccatgt tgtgcaaaaa agcgggttagc tccttcggtc ctccgatcgt 11700
tgtcagaagt aagttggccg cagtgttatt actcatgggt atggcagcac tgcataattc 11760
tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc 11820
attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa 11880
taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg 11940
aaaactctca aggatcttac cgctgttgag atccagttcg atgtaacca ctcgatgcacc 12000
caactgatct tcagcatctt ttactttcac cagcgtttct gggtagagcaa aaacaggaag 12060
gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactctt 12120
cctttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt 12180
tgaatgtatt tagaaaaata aacaaatagg ggttcgcgcg acatttcccc gaaaagtgcc 12240
acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac 12300
gaggcccttt cgtc 12314

```

```

<210> 45
<211> 11180
<212> DNA
<213> Artificial

```

```

<220>
<223> pBOL116

```

```

<400> 45
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accataccac agcttttcaa ttcaattcat cttttttttt ttattctttt ttttgatttc 240
ggtttctttg aaattttttt gattcggtaa tctccgaaca gaaggaagaa cgaaggaagg 300
agcacagact tagattggta tatatacgca tatgtagtgt tgaagaaaca tgaaattgcc 360
cagtattctt aacccaactg cacagaacaa aaacctgcag gaaacgaaga taaatcatgt 420

```

cgaagctac	atataaggaa	cgtgctgcta	ctcatcctag	tcctgttgct	gccaaagctat	480
ttaatatcat	gcacgaaaag	caaacaaact	tgtgtgcttc	attggatggt	cgtaccacca	540
aggaattact	ggagttagtt	gaagcattag	gtcccaaaat	ttgtttacta	aaaacacatg	600
tggaatctt	gactgatttt	tccatggagg	gcacagttaa	gccgctaaag	gcattatccg	660
ccaagtacaa	ttttttactc	ttcgaagaca	gaaaatttgc	tgacattggt	aatacagtca	720
aattgcagta	ctctgcgggt	gtatacagaa	tagcagaatg	ggcagacatt	acgaatgcac	780
acggtgtggt	gggcccagg	attgttagcg	gtttgaagca	ggcggcagaa	gaagtaacaa	840
aggaacctag	aggccttttg	atgttagcag	aattgtcatg	caagggctcc	ctatctactg	900
gagaatatac	taagggtact	gttgacattg	cgaagagcga	caaagatttt	gttatcggct	960
ttattgctca	aagagacatg	ggtggaagag	atgaaggtta	cgattggttg	attatgacac	1020
ccggtgtggg	tttagatgac	aagggagacg	cattgggtca	acagtataga	accgtggatg	1080
atgtggtctc	tacaggatct	gacattatta	ttgttggaag	aggactattt	gcaaagggaa	1140
gggatgctaa	ggtagaggg	gaacgttaca	gaaaagcagg	ctgggaagca	tatttgagaa	1200
gatgcggcca	gcaaaactaa	aaaactgtat	tataagtaaa	tgcatgtata	ctaaactcac	1260
aaattagagc	ttcaatttaa	ttatatcagt	tattacccta	tgcggtgtga	aataccgcac	1320
agatgcgtaa	ggagaaaata	ccgcatcagg	aaattgtaaa	cgtaaataat	ttgttaaaat	1380
tcgcgttaaa	tttttgtaa	atcagctcat	tttttaacca	ataggccgaa	atcggcaaaa	1440
tccttataa	atcaaaaagaa	tagaccgaga	taggggtgag	tggtgttcca	gtttggaaca	1500
agagtccact	attaaagaac	gtggactcca	acgtcaaagg	gcgaaaaacc	gtctatcagg	1560
gcgatggccc	actacgtgaa	ccatcacct	aatcaagttt	tttggggtcg	aggtgccgta	1620
aagcactaaa	tcggaaccct	aaaggagacc	cccgatttag	agcttgacgg	ggaaagccgg	1680
cgaacgtggc	gagaaaggaa	gggaagaaag	cgaaaggagc	gggcgctagg	gcgctggcaa	1740
gtgtagcgg	cacgctgcgc	gtaaccacca	caccgcgcgc	gcttaatgcg	ccgctacagg	1800
gcgcgtcgcg	ccattcgcca	ttcaggctgc	gcaactgttg	ggaagggcga	tcggtgcggg	1860
cctcttcgct	attacgccag	ctggcgaaaag	ggggatgtgc	tgcaaggcga	ttaagttggg	1920
taacgccagg	gttttcccag	tcacgacgtt	gtaaaacgac	ggccagttag	cgcgcgtaat	1980
acgactcact	atagggcgaa	ttgggtaccg	ggccccccct	cgaggtcgac	ggtatcgata	2040
agcttgatat	cgaattcctg	cagcccgggg	gatccactag	ttctagagcg	gccattttaa	2100
acggccggcc	ctagatcaga	gggtggtaaa	tgaagtgtaa	tagtattcat	ttttcttata	2160

aatcatccct	tccgtgattt	atacaaaaga	agaggagaat	atgctgaata	cttggtatat	2220
tactctacat	tatactctta	tcttgacggg	tattctgagc	atcttactca	gtttcaagat	2280
cttttaaatgt	ccaaaaacat	ttgagccgat	ctaaatactt	ctgtgttttc	attaatttat	2340
aaattgtact	cttttaagac	atggaaagta	ccaacatcgg	ttgaaacagt	ttttcattta	2400
cttatggttt	attggttttt	ccagtgaatg	attatttgtc	gttacccttt	cgtaaaagtt	2460
caaacacggt	tttaagtatt	gtttagttgc	tctttcgaca	tatatgatta	tccctgcgcg	2520
gctaaagtta	aggatgcaaa	aaacataaga	caactgaagt	taatttacgt	caattaagtt	2580
ttccagggtta	atgatgtttt	gggcttccac	taattcaata	agtatgtcat	gaaatacgtt	2640
gtgaagagca	tccagaaata	atgaaaagaa	acaacgaaac	tgggtcggcc	tgttgtttct	2700
tttctttacc	acgtgatctg	cggcattttac	aggaagtcgc	gcgttttgcg	cagttgttgc	2760
aacgcagcta	cggctaacaa	agcctagtgg	aactcgactg	atgtgttagg	gcctaaaact	2820
ggtggtgaca	gctgaagtga	actattcaat	ccaatcatgt	catggctgtc	acaaagacct	2880
tgcggaccgc	acgtacgaac	acatacgtat	gctaatatgt	gttttgatag	taccagtgga	2940
tgcgagacct	gcaatttttt	tgtaggtttg	gaagaatata	taaaggttgc	actcattcaa	3000
gatagttttt	ttcttggtgtg	tctattcatt	ttattattgt	ttgtttaaat	gttaaaaaaa	3060
ccaagaactt	agtttcaaat	taaattcatc	acacaaacaa	acaaaacaaa	atgaacattg	3120
ttgtttgttt	gaagcaagtt	ccagacactg	ctgaagtcag	aattgacca	gtcaagggtta	3180
ctttaatcag	agaagggtgt	ccatctatca	tcaaccaga	cgacaagaac	gctttggaag	3240
aagctttggt	tttgaaggac	aactacggtg	ctcacgttac	cgtcatttcc	atgggtccac	3300
ctcaagccaa	gaacgctttg	gttgaagctt	tggccatggg	tgctgatgaa	gctgtcttat	3360
tgactgacag	agctttcggg	ggtgctgata	ctttagctac	ctctcacacc	attgctgctg	3420
gtatcaagaa	attgaaatac	gatatcgtct	ttgccggtcg	tcaagccatc	gatggtgata	3480
ccgctcaagt	cgggccagaa	attgctgaac	atttgggtat	tccacaagtc	acctacgttg	3540
aaaagggtga	agttgacggt	gacactttga	agatcagaaa	ggcttgggaa	gacggttacg	3600
aagttgttga	agtcaagact	ccagttctat	tgactgccat	caaggaattg	aacgttccaa	3660
gatacatgtc	cgttgaaaag	atcttcggtg	ctttcgacaa	ggaagtcaag	atgtggactg	3720
ctgatgatat	cgatgtcgac	aaggccaact	tgggtttgaa	aggttctcca	accaaggtca	3780
agaaatcttc	taccaaggaa	gtcaagggtc	aagggtgaagt	cattgacaaa	ccagtcaagg	3840

aagctgccgc ttacgttggt tccaagttga aggaagaaca ctacatctaa agcccgggcg	3900
gagattgata agacttttct agttgcatat cttttatatt taaatcttat ctattagtta	3960
atTTTTtgta atttatcctt atatatagtc tggttattct aaaatatcat ttcagtatct	4020
aaaaattccc ctcttttttc agttatatct taacaggcga cagtccaaat gttgatttat	4080
cccagtcgga ttcatcaggg ttgtgaagca ttttgtcaat ggtcgaaatc acatcagtaa	4140
tagtgccctc tacttgccctc atagaatttc tttctcttaa cgtcaccgtt tggcttttta	4200
tagtttcgaa atctatggtg ataccaaagtg gtgttcccaa ttcatcgta cgggcgtatt	4260
ttttaccaat tgaagtattg gaatcgtaaa ttttaaagta tatctctctt ttacgtaaag	4320
cctgcgagat cctcttaagt atagcgggga agccatcggt attcgatatt gtcgtaacaa	4380
atactttgat cggcgctatg tttaaatggt taaacatgga cagatatgcg atgaaaacgc	4440
taagtatac tccaaatggt gaaaggtaag atgcttgga acaatacttg gaaatcaccg	4500
gaaacaccat atgcggcgaa aagccaatta gtgtgatact aagtgcctta tcgaaaatcc	4560
gtgatgccgg tccttcaggc atcaaatttc agtggcctaa ttattcacag agttctcatg	4620
tgacaagtat tgatgatagt agtgtcagtt atgcttcagg ttatgttact ataggataat	4680
gatcacggct aaaacggtcg aatgtaagca tatacttttc gattgtataa ttgttcccaa	4740
atactacagc atctcaagga aaaaaaaca aaaacttcca aaaaaatcga atccctgagg	4800
aatctttaat acattttcaa tctatttaag ttttataaac gtgtatatga gatgtcatga	4860
gcatgaatta ttaataataa aaactaaatc attaaagtaa cttaaggagt taaagcccgg	4920
gctttaattg ttagcagcct tgacttgagc aatcaattct ggaacaacct tgttgacatc	4980
accgacaatg gccaaatcag caaccttcat gattggagct tcgacatctt tgttgatggc	5040
aatgatgtag tcagagtctt gcataccagc caagtgttg atggcaccag agataccaca	5100
agcaatgtac aaagttggtc tgacggtctt accggtttga ccgacttgca agtccttgtc	5160
aaccatttcc ttttcaatgg cagctctgga agcagcaatg gtaccacca acaaagaagc	5220
taattcttcc aatttttcga agttttcctt ggaaccaaca ccacgaccac cagcaaccaa	5280
aaccttggct tcaccgatat cagcaatgtc cttggccaat ttgacaacct tggaaacctt	5340
ggttctgata tcagaagcag tcaatttgat ggcaaccttt tcgatcttgt catcagaaac	5400
gtagcatcg ttaactggca atttttcaaa gacacctggt ctgacggtgg ccatttgagg	5460
tctgtggtca gaacagacaa tggtagcaat caagttacca ccgaaagctg gtctggtagc	5520
caacaagtca cggttttcga catcgatatc caaagaggta cagtcagcag tcaaaccagt	5580

agacaatctg	gcagcaattc	ttggacccaa	gtctctaccg	atgaaagtag	caccgatgaa	5640
taagatttct	ggctttcttt	cgttgaccaa	gtcacagata	accttggcgt	aaccgtcagt	5700
ggagaaatga	gctaataatt	cgttgtcagc	agccaaaacc	ttgtcagcac	cgtgggacaa	5760
caagtccttg	gacatctttt	cagtgttgtg	acccaataag	acagcagtca	attcaacacc	5820
caatthttca	gccatttcct	tacccttacc	tagcaattcc	aaagaaacct	tttgtaattc	5880
accatctctt	tgttcagcga	aaaccagac	acccttgtag	tcagccttgt	tcatgtttag	5940
ttaattatag	ttcgttgacc	gtatatctta	aaaacaagta	ctccttaaaa	aaaaaccttg	6000
aagggaataa	acaagtagaa	tagatagaga	gaaaaataga	aatgcaaga	gaatttatat	6060
attagaaaga	gagaaagaaa	aatggaaaaa	aaaaaatagg	aaaagccaga	aatagcacta	6120
gaaggagcga	caccagaaaa	gaagggtgatg	gaaccaatth	agctatatat	agttaactac	6180
cggctcgatc	atctctgcct	ccagcatagt	cgaagaagaa	tttttttttt	cttgaggcctt	6240
ctgtcagcaa	ctcgtatttt	ttctttcttt	tttggtgagc	ctaaaaagtt	cccacgttct	6300
cttgtagcac	gccgtcacia	acaaccttat	gggtaatttg	tcgcggctctg	gggtgtataaa	6360
tgtgtgggtg	caggccggcc	gtttaaacgg	gccgccaccg	cgggtggagcc	tgtgtggaag	6420
aacgattaca	acaggtgttg	tcctctgagg	acataaaaata	cacaccgaga	ttcatcaact	6480
cattgctgga	gttagcatat	ctacaattgg	gtgaaatggg	gagcgatttg	caggcatttg	6540
ctcggcatgc	cggtagaggt	gtggtcaata	agagcgacct	catgctatac	ctgagaaagc	6600
aacctgacct	acaggaaaga	gttactcaag	aataagaatt	ttcgttttaa	aacctaaagag	6660
tcactttaaa	atttgatatac	acttattttt	tttataactt	atttaataat	aaaaatcata	6720
aatcataaga	aattcgctcg	agtcgactgc	agttttacaag	ttcaatttg	ccatgatggc	6780
cttaacaatg	gcttgaacat	cttcgttgtc	ttctggttca	gctggagcag	aagaagaagc	6840
agcaccgaca	ccgaaagctt	ctctgatttc	ttcgacgggtg	gtggtaccgt	aagcaacctt	6900
tctgatgttg	aacaagtttt	ctggaccaac	gttgtcagag	gtggcagaac	caccaacagc	6960
accacaacct	aaagtcaaag	atgggactaa	gttggttagca	ccaccgatac	cacccaaaga	7020
acctggagag	ttaacaaaaa	ttctggaaac	aggcttcttc	aaagcaaatt	ctctaattgat	7080
ttcttcgttt	tgagagtgga	tgatcaaagt	gtgaccagaa	ccttggttgt	gcaataaagc	7140
caaagacttt	tcacaagctt	catgccagtc	ttcgacgggtg	tagaaagcca	agactggagc	7200
caatthttcc	ttagcgtatg	gattthttagg	agaaacatcg	gtttgttcgg	atagtaagat	7260

aacagcatca	gatggaatgg	aaataccagc	caatttggcc	aaagcttggg	catccttacc	7320
aacgatggct	gggtttggag	taccgttggc	acgtaataga	atcttaccaa	ccttttcaga	7380
ttcttcagca	ttcaagaagt	aacccttttg	tctcttgaat	tcttcgatga	tttcagcctt	7440
cttgacggtt	tcagcaatga	tggattgttc	agaagcacag	atgacaccgt	tgtcgaaagt	7500
gtcagaaccg	ataacctttc	taacagcagt	tggaatgtca	gcagttcttt	cgatgaaaca	7560
tggaccgtta	cctggaccga	caccgatggc	tggagtacca	gaggagtaag	cagctctaac	7620
cataccttca	ccaccggtag	ccaagatcaa	agcgggtgtc	ttgttcttca	tcaattcagc	7680
agtaccttca	acggtcaaaa	tggacataca	ttggatcaaa	ccatctggag	caccagcttc	7740
aacagcagcc	ttttgcatga	tcttaacggt	ttcagtgatg	gaacggacag	cagttgggtg	7800
tggagagaag	acaatggcgt	taccagcctt	caaagcaatc	aagaccttga	aaatggcagt	7860
ggaagtgggg	ttggtagatg	gaatcaaacc	agcaatgaca	cctaattggga	cagcaatgtc	7920
aatcaatttc	ttttccttgt	cttccttcaa	gataccaacg	gtcttcaaata	ccttgatgta	7980
gttgtagaca	acaatggagg	agaatttggt	cttgatgacc	ttgtcttccc	atttaccgta	8040
accggtgtct	tcgtaagcca	atttggccaa	tttgacagct	tcaacttcag	tagccttggc	8100
gatcttttcg	atgaccttgt	taacagcttc	ttgggaaaag	ttcttgaatt	cagcttgagc	8160
cttcttggcc	ttggcaatca	aagtcttaac	ttcttggatg	gattgcaaata	ccttgtccat	8220
gatttccatt	ttttactagt	tctagaatcc	gtcgaaacta	agttctgggtg	ttttaaaact	8280
aaaaaaaaaga	ctaactataa	aagtagaatt	taagaagttt	aagaaataga	tttacagaat	8340
tacaatcaat	acctaccgtc	tttatatact	tattagtcaa	gtaggggaat	aatttcaggg	8400
aactggtttc	aacctttttt	ttcagctttt	tccaaatcag	agagagcaga	aggtaataga	8460
aggtgtaaga	aaatgagata	gatacatgcg	tgggtcaatt	gccttgtgtc	atcatttact	8520
ccaggcaggt	tgcactactc	cattgagggt	gtgcccgttt	tttgccctgtt	tgtgcccttg	8580
ttctctgtag	ttgcgctaag	agaatggacc	tatgaactga	tggttgggtga	agaaaacaat	8640
atthtgggtgc	tgggattctt	tttttttctg	gatgccagct	taaaaagcgg	gctccattat	8700
atthagtggg	tgccagggaat	aaactgttca	cccagacacc	tacgatgtta	tatatctgtg	8760
gtaacccgcc	ccctatthttg	ggcatgtacg	ggttacagca	gaattaaaag	gctaathttt	8820
tgactaaata	aagttaggaa	aatcactact	attaattatt	tacgtattct	ttgaaatggc	8880
gagtattgat	aatgataaac	tgagctccag	ctthtgttcc	ctthagttag	ggttaattgc	8940
gcgcttggcg	taatcatggg	catagctgtt	tcctgtgtga	aattgttatc	cgctcacaat	9000

tccacacaac	ataggagccg	gaagcataaa	gtgtaaagcc	tggggtgcct	aatgagtgag	9060
gtaactcaca	ttaattgcgt	tgcgtcact	gcccgccttc	cagtcgggaa	acctgtcgtg	9120
ccagctgcat	taatgaatcg	gccaacgcgc	ggggagaggc	ggtttgcgta	ttgggcgctc	9180
ttccgccttc	tcgtcactg	actcgtgcg	ctcggtcggt	cggctgcggc	gagcggatc	9240
agctcactca	aaggcggtaa	tacggttatc	cacagaatca	ggggataacg	caggaaagaa	9300
catgtgagca	aaaggccagc	aaaaggccag	gaaccgtaaa	aaggccgcgt	tgctggcggt	9360
tttccatagg	ctccgcccc	ctgacgagca	tcacaaaaat	cgacgctcaa	gtcagagggtg	9420
gcgaaacccg	acaggactat	aaagatacca	ggcgcttccc	cctggaagct	ccctcgtgcg	9480
ctctcctgtt	ccgaccctgc	cgtttaccgg	atacctgtcc	gcctttctcc	cttcgggaag	9540
cgtggcgctt	tctcatagct	cacgctgtag	gtatctcagt	tcgggtgtagg	tcgttcgctc	9600
caagctgggc	tgtgtgcacg	aaccccccg	tcagcccgac	cgtgcgcct	tatccggtaa	9660
ctatcgtctt	gagtccaacc	cggtaagaca	cgacttatcg	ccactggcag	cagccactgg	9720
taacaggatt	agcagagcga	ggatatgtagg	cggtgctaca	gagttcttga	agtgggtggc	9780
taactacggc	tacactagaa	ggacagtatt	tggatatctgc	gctctgctga	agccagttac	9840
cttcggaaaa	agagttggta	gctcttgatc	cggcaaaaa	accaccgctg	gtagcggtagg	9900
tttttttgtt	tgcaagcagc	agattacgcg	cagaaaaaaa	ggatctcaag	aagatccttt	9960
gatcttttct	acggggtctg	acgtcagtg	gaacgaaaac	tcacgttaag	ggatttttgt	10020
catgagatta	tcaaaaagga	tcttcacct	gatcctttta	aattaaaaat	gaagtttta	10080
atcaatctaa	agtatatatg	agtaaacttg	gtctgacagt	taccaatgct	taatcagtga	10140
ggcacctatc	tcagcgatct	gtctatctcg	ttcatccata	gttgctgac	tccccgctgt	10200
gtagataact	acgatacggg	agggcttacc	atctggcccc	agtgcgcaa	tgataccgcg	10260
agaccacgc	tcaccggctc	cagatttatc	agcaataaac	cagccagccg	gaagggccga	10320
gcgcagaagt	ggcctgcaa	ctttatccgc	ctccatccag	tctattaatt	gttgccggga	10380
agctagagta	agtagttcgc	cagttaatag	tttgcgcaac	gttggtgcca	ttgctacagg	10440
catcgtgggt	tcacgctcgt	cgtttgggt	ggcttcattc	agctccggtt	ccaacgatc	10500
aaggcgagtt	acatgatccc	ccatgttggt	caaaaaagcg	gttagctcct	tcggtcctcc	10560
gatcgttggt	agaagtaagt	tggccgcagt	gttatcactc	atggttatgg	cagcactgca	10620
taattctctt	actgtcatgc	catccgtaag	atgcttttct	gtgactgggt	agtactcaac	10680

```

caagtcattc tgagaatagt gtatgcggcg accgagttgc tcttgcccgg cgtcaatacg 10740
ggataatacc gcgccacata gcagaacttt aaaagtgttc atcattggaa aacgttcttc 10800
ggggcgaaaa ctctcaagga tcttaccgct gttgagatcc agttcgatgt aaccactcgc 10860
tgcaccaaac tgatcttcag catctttttac tttcaccagc gtttctgggt gagcaaaaac 10920
aggaaggcaa aatgccgcaa aaaagggaat aagggcgaca cggaatgtt gaatactcat 10980
actcttcctt tttcaatatt attgaagcat ttatcagggt tattgtctca tgagcggata 11040
catatttgaa tgtattttaga aaaataaaca aataggggtt ccgcgcacat ttccccgaaa 11100
agtgccacct gacgtctaag aaaccattat tatcatgaca ttaacctata aaaataggcg 11160
tatcacgagg cccttttcgtc 11180

```

```

<210> 46
<211> 11108
<212> DNA
<213> Artificial

```

```

<220>
<223> pBOL118

```

```

<400> 46
tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accataccac agcttttcaa ttcaattcat catTTTTTTT ttattctttt ttttgatttc 240
ggtttctttg aaattttttt gattcggtaa tctccgaaca gaaggaagaa cgaaggaagg 300
agcacagact tagattggta tatatacgca tatgtagtgt tgaagaaaca tgaaattgcc 360
cagtattctt aaccaactg cacagaacaa aaacctgcag gaaacgaaga taaatcatgt 420
cgaaagctac atataaggaa cgtgctgcta ctcatcctag tcctgttgct gccaaagtat 480
ttaatatcat gcacgaaaag caaacaaact tgtgtgcttc attggatgtt cgtaccacca 540
aggaattact ggagttagtt gaagcattag gtcccaaaat ttgtttacta aaaacacatg 600
tggtatctt gactgatttt tccatggagg gcacagttaa gccgctaaag gcattatccg 660
ccaagtacaa ttttttactc ttcgagaca gaaaatttgc tgacattggt aatacagtca 720
aattgcagta ctctgcgggt gtatacagaa tagcagaatg ggcagacatt acgaatgcac 780
acgggtgtgt gggcccagggt attgttagcg gtttgaagca ggcggcagaa gaagtaacaa 840
aggaacctag aggccttttg atgttagcag aattgtcatg caagggctcc ctatctactg 900

```

gagaatatac taagggtact gttgacattg cgaagagcga caaagatfff gttatcggt	960
ttattgctca aagagacatg ggtggaagag atgaaggtta cgattgggtg attatgacac	1020
ccggtgtggg tttagatgac aaggagagacg cattgggtca acagtataga accgtggatg	1080
atgtggtctc tacaggatct gacattatta ttgttggaag aggactatff gcaaagggaa	1140
gggatgctaa ggtagagggt gaacgttaca gaaaagcagg ctgggaagca tatttgagaa	1200
gatgcggcca gcaaaactaa aaaactgtat tataagtaaa tgcattgata ctaaactcac	1260
aaattagagc ttcaatttaa ttatatcagt tattacccta tgcgggtgtga aataccgcac	1320
agatgcgtaa ggagaaaata ccgcatcagg aaattgtaaa cgttaatat ttgttaaaat	1380
tcgcgttaaa tttttgttaa atcagctcat tttttaacca ataggccgaa atcggaaaaa	1440
tcccttataa atcaaaaagaa tagaccgaga taggggtgag tgttggtcca gtttgaaca	1500
agagtccact attaaagaac gtggactcca acgtcaaagg gcgaaaaacc gtctatcagg	1560
gcgatggccc actacgtgaa ccatcacct aatcaagttt tttggggctg aggtgccgta	1620
aagcactaaa tcggaacct aaaggagacc cccgatttag agcttgacgg ggaaagccgg	1680
cgaacgtggc gagaaaggaa gggaagaaag cgaaaggagc gggcgctagg gcgctggcaa	1740
gtgtagcgt cacgctgct gtaaccacca caccgcctg gcttaatgcg ccgctacagg	1800
gcgcgtcgcg ccattcgcca ttcaggctgc gcaactgttg ggaagggcga tcggtgcggg	1860
cctcttcgct attacgccag ctggcgaaag ggggatgtgc tgcaaggcga ttaagttggg	1920
taacgccagg gttttcccag tcacgacgtt gtaaaacgac ggccagttag cgcgcgtaat	1980
acgactcact ataggcgcaa ttgggtaccg ggccccct cgaggctgac ggtatcgata	2040
agcttgatat cgaattcctg cagcccggg gatccactag ttctagagcg gccatttaa	2100
acggccggcc ctagatcaga gggtggtaaa tgaagtgtaa tagtattcat tttcttata	2160
aatcatccct tccgtgattt atacaaaaga agaggagaat atgctgaata cttggtatat	2220
tactctacat tatactctta tcttgacggg tattctgagc atcttactca gtttcaagat	2280
cttttaatgt ccaaaaacat ttgagccgat ctaaatactt ctgtgttttc attaatatat	2340
aaattgtact cttttaagac atggaaagta ccaacatcg ttgaaacagt tttcattta	2400
cttatggttt attggttttt ccagtgaatg attatttgtc gttacccttt cgtaaaagtt	2460
caaacacgtt ttaagtatt gtttagttgc tctttcgaca tatatgatta tccctgcgcg	2520
gctaaagtta aggatgcaaa aaacataaga caactgaagt taatttacgt caattaagtt	2580

ttccagggta atgatgtttt gggcttccac taattcaata agtatgtcat gaaatacgtt	2640
gtgaagagca tccagaaata atgaaaagaa acaacgaaac tgggtcggcc tggtgtttct	2700
tttctttacc acgtgatctg cggcatttac aggaagtcgc gcgttttgcg cagttgttgc	2760
aacgcagcta cggctaacaa agcctagtgg aactcgactg atgtgttagg gcctaaaact	2820
ggtggtgaca gctgaagtga actattcaat ccaatcatgt catggctgtc acaaagacct	2880
tgcggaccgc acgtacgaac acatacgtat gctaatatgt gttttgatag taccagtgga	2940
tcgcagacct gcaatttttt tgtaggtttg gaagaatata taaagggtgc actcattcaa	3000
gatagttttt ttcttggtgtg tctattcatt ttattattgt ttgtttaaat gttaaaaaaa	3060
ccaagaactt agtttcaaata taaattcatc acacaaacaa acaaaacaaa atgaacattg	3120
ttgtttgttt gaagcaagtt ccagacactg ctgaagtcag aattgacca gtcaagggtga	3180
ctttaatcag agaaggtgtt ccatctatca tcaaccaga cgacaagaac gctttggaag	3240
aagctttggt tttgaaggac aactacggtg ctacggttac cgtcatttcc atgggtccac	3300
ctcaagccaa gaacgctttg gttgaagctt tggccatggg tgctgatgaa gctgtcttat	3360
tgactgacag agctttcggg ggtgctgata ctttagctac ctctcacacc attgctgctg	3420
gtatcaagaa attgaaatac gatatcgtct ttgccggtcg tcaagccatc gatggtgata	3480
ccgctcaagt cgggtccagaa attgctgaac atttgggtat tccacaagtc acctacgttg	3540
aaaaggttga agttgacggg gacactttga agatcagaaa ggcttgggaa gacggttacg	3600
aagttgttga agtcaagact ccagttctat tgactgccat caaggaattg aacgttccaa	3660
gatacatgtc cgttgaaaag atcttcggtg ctttcgacaa ggaagtcaag atgtggactg	3720
ctgatgatat cgatgtcgac aaggccaact tgggtttgaa aggttctcca accaaggtca	3780
agaaatcttc taccaaggaa gtcaagggtc aaggtgaagt cattgacaaa ccagtcaagg	3840
aagctgccgc ttacgttggt tccaagttga aggaagaaca ctacatctaa agcccgggcg	3900
gagattgata agacttttct agttgcata cttttatatt taaatcttat ctattagtta	3960
atTTTTtGta atttatcctt atatatagtc tggttattct aaaatatcat ttcagtatct	4020
aaaaattccc ctcttttttc agttatatct taacaggcga cagtccaaat gttgatttat	4080
cccagtccga ttcatcaggg ttgtgaagca ttttgtcaat ggtcgaaatc acatcagtaa	4140
tagtgctctt tacttgctc atagaatttc tttctcttaa cgtcaccgtt tggcttttta	4200
tagtttcgaa atctatggtg ataccaaatg gtgttcccaa ttcacgttta cgggcgtatt	4260
ttttaccaat tgaagtattg gaatcgtcaa ttttaaagta tatctctctt ttacgtaaag	4320

cctgcgagat	cctcttaagt	atagcgggga	agccatcggt	attcgatatt	gtcgtaacaa	4380
atactttgat	cggcgctatg	tttaaagtgt	taaacatgga	cagatatgcg	atgaaaacgc	4440
taagtgatac	tccaaatggt	gaaagggtacg	atgcttgga	acaatacttg	gaaatcaccg	4500
gaaacaccat	atgcggcgaa	aagccaatta	gtgtgatact	aagtgcctta	tcgaaaatcc	4560
gtgatgccgg	tccttcaggc	atcaaatttc	agtggcctaa	ttattcacag	agttctcatg	4620
tgacaagtat	tgatgatagt	agtgtcagtt	atgcttcagg	ttatgttact	ataggataat	4680
gatcacggct	aaaacggtcg	aatgtaagca	tatatctttc	gattgtataa	ttgttcccaa	4740
atactacagc	atctcaagga	aaaaaaaaaca	aaaacttcca	aaaaaatcga	atccctgagg	4800
aatctttaat	acattttcaa	tctattttaag	ttttataaac	gtgtatatga	gatgtcatga	4860
gcatgaatta	ttaataataa	aaactaaatc	attaaagtaa	cttaaggagt	taaagcccgg	4920
gctttaattg	ttagcagcct	tgacttgagc	aatcaattct	ggaacaacct	tggtgacatc	4980
accgacaatg	gccaaatcag	caaccttcat	gattggagct	tcgacatctt	tggtgatggc	5040
aatgatgtag	tcagagtctt	gcataccagc	caagtgttg	atggcaccag	agataccaca	5100
agcaatgtac	aaagttggtc	tgacggtctt	accggtttga	ccgacttgca	agtccttgtc	5160
aaccatttcc	ttttcaatgg	cagctctgga	agcagcaatg	gtaccacca	aaaagaagc	5220
taattcttcc	aatttttcga	agttttcctt	ggaaccaaca	ccacgaccac	cagcaaccaa	5280
aaccttggt	tcaccgatat	cagcaatgtc	cttggccaat	ttgacaacct	tggaacctt	5340
ggttctgata	tcagaagcag	tcaatttgat	ggcaaccttt	tcgatcttgt	catcagaaac	5400
gtagcatcg	ttaactggca	atttttcaaa	gacacctggt	ctgacggtgg	ccatttgagg	5460
tctgtggtca	gaacagacaa	tggtagcaat	caagttacca	ccgaaagctg	gtctggtagc	5520
caacaagtca	cggttttcga	catcgatata	caaagaggta	cagtcagcag	tcaaaccagt	5580
agacaatctg	gcagcaattc	ttggacccaa	gtctctaccg	atgaaagtag	caccgatgaa	5640
taagatttct	ggcttttctt	cgttgaccaa	gtcacagata	accttggcgt	aaccgtcagt	5700
ggagaaatga	gctaataatt	cgttgtcagc	agccaaaacc	ttgtcagcac	cgtgggacaa	5760
caagtccttg	gacatctttt	cagtgttgtg	accaataag	acagcagtca	attcaacacc	5820
caatttttca	gccatttcct	tacccttacc	tagcaattcc	aaagaaacct	tttgtaattc	5880
accatctctt	tggtcagcga	aaaccagac	acccttgtag	tcagccttgt	tcatgtttag	5940
ttaattatag	ttcggtgacc	gtatattcta	aaaacaagta	ctccttaaaa	aaaaaccttg	6000

aagggaataa	acaagtagaa	tagatagaga	gaaaaataga	aatgcaaga	gaatttatat	6060
attagaaaga	gagaaagaaa	aatggaaaaa	aaaaaatagg	aaaagccaga	aatagcacta	6120
gaaggagcga	caccagaaaa	gaaggtgatg	gaaccaatth	agctatatat	agttaactac	6180
cggctcgatc	atctctgcct	ccagcatagt	cgaagaagaa	tttttttttt	cttgaggctt	6240
ctgtcagcaa	ctcgtattht	ttctttcttt	tttggtgagc	ctaaaaagtt	cccacgttct	6300
cttgtagcac	gccgtcacia	acaaccttat	gggtaatthg	tcgcggtctg	ggtgtataaa	6360
tgtgtgggtg	caggccggcc	gtttaaaccg	gccgccaccg	cgggtggagcc	tgtgtggaag	6420
aacgattaca	acaggtgthg	tcctctgagg	acataaaata	cacaccgaga	ttcatcaact	6480
cattgctgga	gtagcatat	ctacaattgg	gtgaaatggg	gagcgatthg	caggcattthg	6540
ctcggcatgc	cggtagaggt	gtggtcaata	agagcgacct	catgctatac	ctgagaaagc	6600
aacctgacct	acaggaaaga	gttactcaag	aataagaatt	ttcgtthttaa	aacctaaagag	6660
tcactthttaa	atthgtatac	acttatthtth	tttataactt	atthtaataat	aaaaatcata	6720
aatcataaga	aattcgctcg	agtcgactgc	agthttaaaca	attctgaaag	catcgaccaa	6780
aacacaacga	cgtaatctga	cgaaagthct	ggcagaagtg	acaccttcac	cagttgggggt	6840
ggtgatggtc	atggtggtcc	aaccttcacc	acccaaaccc	aaaccagcga	tacatggacc	6900
gttcttgaca	aagatggaag	tgtcaatggc	gtagccatt	tggttcatgt	tttcgatgth	6960
tctggagtgc	atggcagcag	tgtggtgaca	accaccttcc	aatttgacag	ccaaagcaat	7020
agcgtcagca	acgtagcaa	cacggacaac	tggtaaagact	ggcatcatca	attcagtgac	7080
agcaaatggg	tgthcagcgg	tggtthtcgac	gaataataat	ctggthtctt	gtggaacctt	7140
caaaccgatg	gcagcagcaa	tcttaccagc	atctctacca	accagthctc	tggagacggt	7200
accttacct	ctthcatcga	tgttcttcaa	caaaactgggt	tgcaattgth	gagcttgthc	7260
agcagtcaac	ttgacggcat	gttgaccttc	catcaatctc	atcaattcgt	cagcaacgga	7320
gtcaacaaca	atcaaaacct	tttcgtcagc	acagatgatg	ttgttgtcga	aagaagcacc	7380
cttgacaatg	gattgagcag	ctctggccaa	atcagcggtt	tcacgcagaa	caacaggagg	7440
gttaccagca	ccagcagcaa	tcaatctctt	gttggtgtgc	tttctggcag	cttcaacaac	7500
agcttcacca	ccagtgcaga	ctaatagacc	gataacctggg	aacttgaata	atctthtgagc	7560
agthtctgata	tctgggtthg	caacagtgac	caacaagtht	tctggaccac	cagcagcaac	7620
aatggctthg	thcaatagag	tgatggctct	thgagaaacc	thctthggcag	ctgggtgtgg	7680
agcgaagata	acggagthac	cagcagcaat	caaagagatg	gcgtthgttg	tgacagtagc	7740

agctggggttg gtagatgggg tgacggaagc aacaacaccc catggagcat tttcaatcaa	7800
agtcaaacca ttatcacccg tcaagacttg tggagacaaa cattcgacac ctggagtacc	7860
tctagcttga gcaacgttct tagcgaatth gtcttcaact ctaccatac cggtttcggg	7920
gacagccaat tcagccaagt ctctggcatg cttttcacca gcttctctga tggcagcaat	7980
ggccaattgt ctcatggcaa cagatttcaa accttggtga gcaaccttgg cagcagcaac	8040
agcgtcgtcc aaagaagcga aaacacccat ttcgtggaca gcagcagatg gagtgtcaga	8100
agattgcatt ttcaacaaga cagccttgac aacttggtcg atatcttggt ggttcatttt	8160
ttactagtgc tagaatccgt cgaaactaag ttctgggtgt ttaaaactaa aaaaaagact	8220
aactataaaa gtagaattta agaagtttaa gaaatagatt tacagaatta caatcaatac	8280
ctaccgtctt tatatactta ttagtcaagt aggggaataa tttcagggaa ctggtttcaa	8340
cctttttttt cagctttttc caaatcagag agagcagaag gtaatagaag gtgtaagaaa	8400
atgagataga tacatgcgtg ggtcaattgc cttgtgtcat catttactcc aggcaggttg	8460
catcactcca ttgaggttgt gcccggtttt tgccctgttg tgcccctgtt ctctgtagtt	8520
gcgctaagag aatggaccta tgaactgatg gttggtgaag aaaacaatat tttggtgctg	8580
ggattctttt tttttctgga tgccagctta aaaagcgggc tccattatat ttagtggatg	8640
ccaggaataa actgttcacc cagacaccta cgatgttata tattctgtgt aaccgcggcc	8700
ctattttggg catgtacggg ttacagcaga attaaaaggc taattttttg actaaataaa	8760
gttaggaaaa tcactactat taattattta cgtattcttt gaaatggcga gtattgataa	8820
tgataaactg agctccagct tttgttccct ttagtgaggg ttaattgcgc gcttggcgta	8880
atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg ctcaaatc cacacaacat	8940
aggagccgga agcataaagt gtaaagcctg ggggtgcctaa tgagtgaggt aactcacatt	9000
aattgcgttg cgctcactgc ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta	9060
atgaatcggc caacgcgcgg ggagaggcgg ttgctgtatt gggcgctctt ccgcttcctc	9120
gctcactgac tcgctgcgct cggctgttcg gctgcggcga gcggtatcag ctactcaaa	9180
ggcggtataa cgtttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa	9240
aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt tccataggct	9300
ccgccccctt gacgagcatc aaaaaaatcg acgctcaagt cagaggtggc gaaacccgac	9360
aggactataa agataaccagg cgtttccccc tggaagctcc ctctgtgcgt ctctgttcc	9420

gaccctgccg	cttaccggat	acctgtccgc	ctttctccct	tcgggaagcg	tggecgcttc	9480
tcatagtca	cgctgtaggt	atctcagttc	gggtgtaggtc	gttcgctcca	agctgggctg	9540
tgtgcacgaa	cccccggttc	agcccgaccg	ctgcgcctta	tccggttaact	atcgtcttga	9600
gtccaacccg	gtaagacacg	acttatcgcc	actggcagca	gccactggta	acaggattag	9660
cagagcgagg	tatgtaggcg	gtgctacaga	gttcttgaag	tggtggccta	actacggcta	9720
cactagaagg	acagtatttg	gtatctgcgc	tctgctgaag	ccagttacct	tcggaaaaag	9780
agttggtagc	tcttgatccg	gcaaacaac	caccgctggg	agcgggtggt	tttttgttg	9840
caagcagcag	attacgcgca	gaaaaaaagg	atctcaagaa	gacaccttga	tcttttctac	9900
ggggtctgac	gctcagtgga	acgaaaactc	acgttaaggg	attttggtca	tgagattatc	9960
aaaaaggatc	ttcacctaga	tcctttttaa	ttaaaaatga	agttttaaat	caatctaaag	10020
tatatatgag	taaacttggt	ctgacagtta	ccaatgctta	atcagtgagg	cacctatctc	10080
agcgatctgt	ctatttcggt	catccatagt	tgcttgactc	cccgctcgtg	agataactac	10140
gatacgggag	ggcttaccat	ctggccccag	tgctgcaatg	ataccgcgag	accacgctc	10200
accggctcca	gatttatcag	caataaacca	gccagccgga	agggccgagc	gcagaagtgg	10260
tcctgcaact	ttatccgcct	ccatccagtc	tattaattgt	tgccgggaag	ctagagtaag	10320
tagttcgcca	gttaatagtt	tgcgcaacgt	tggtgccatt	gctacaggca	tcgtggtgtc	10380
acgctcgtcg	tttggtatgg	cttcattcag	ctccggttcc	caacgatcaa	ggcgagttac	10440
atgatcccc	atgttggtgca	aaaaagcggg	tagctccttc	ggctctccga	tcgttgctcag	10500
aagtaagttg	gccgcagtgt	tatcactcat	ggttatggca	gcactgcata	attctcttac	10560
tgtcatgcc	tccgtaagat	gcttttctgt	gactggtgag	tactcaacca	agtcattctg	10620
agaatagtgt	atgcggcgac	cgagttgctc	ttgcccggcg	tcaatacggg	ataataccgc	10680
gccacatagc	agaacttta	aagtgtcat	cattggaaaa	cgttcttcgg	ggcgaaaact	10740
ctcaaggatc	ttaccgctgt	tgagatccag	ttcgatgtaa	cccactcgtg	cacccaactg	10800
atcttcagca	tcttttactt	tcaccagcgt	ttctgggtga	gcaaaaacag	gaaggcaaaa	10860
tgccgcaaaa	aagggaataa	gggcgacacg	gaaatgttga	atactcatac	tcttcctttt	10920
tcaatattat	tgaagcattt	atcagggtta	ttgtctcatg	agcggataca	tatttgaatg	10980
tatttagaaa	aataaacaaa	taggggttcc	gcgcacattt	ccccgaaaag	tgccacctga	11040
cgtctaagaa	accattatta	tcatgacatt	aacctataaa	aataggcgta	tcacgaggcc	11100
ctttcgtc						11108

<210> 47
 <211> 11114
 <212> DNA
 <213> Artificial

<220>
 <223> pBOL120

<400> 47
 tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
 accataccac agcttttcaa ttcaattcat cttttttttt ttattctttt ttttgatttc 240
 ggtttctttg aaattttttt gattcggtaa tctccgaaca gaaggaagaa cgaaggaagg 300
 agcacagact tagattggta tatatacgca tatgtagtgt tgaagaaaca tgaaattgcc 360
 cagtattctt aaccaactg cacagaacaa aaacctgcag gaaacgaaga taaatcatgt 420
 cgaaagctac atataaggaa cgtgctgcta ctcatcctag tcctgttgct gccaaagctat 480
 ttaatatcat gcacgaaaag caaacaact tgtgtgcttc attggatgtt cgtaccacca 540
 aggaattact ggagttagtt gaagcattag gtcccaaat ttgtttacta aaaacacatg 600
 tggatatctt gactgatttt tccatggagg gcacagttaa gccgctaaag gcattatccg 660
 ccaagtacaa ttttttactc ttcgaagaca gaaaatttgc tgacattggg aatacagtca 720
 aattgcagta ctctgcgggt gtatacagaa tagcagaatg ggcagacatt acgaatgcac 780
 acgggtgtgtt gggcccagggt attgttagcg gtttgaagca ggcggcagaa gaagtaacaa 840
 aggaacctag aggccttttg atgttagcag aattgtcatg caagggctcc ctatctactg 900
 gagaatatac taagggtact gttgacattg cgaagagcga caaagatttt gttatcggct 960
 ttattgctca aagagacatg ggtggaagag atgaaggtta cgattgggtg attatgacac 1020
 ccggtgtggg tttagatgac aaggagagcg cattgggtca acagtataga accgtggatg 1080
 atgtggtctc tacaggatct gacattatta ttgttgaag aggactattt gcaaagggaa 1140
 gggatgctaa ggtagaggggt gaacgttaca gaaaagcagg ctgggaagca tatttgagaa 1200
 gatgcggcca gcaaaactaa aaaactgtat tataagtaaa tgcattgata ctaaaactcac 1260
 aaattagagc ttcaatttaa ttatatcagt tattacccta tgcgggtgtga aataccgcac 1320
 agatgcgtaa ggagaaaata ccgcatcagg aaattgtaaa cgttaaatatt ttgttaaaat 1380

tcgcgtaaaa	tttttgttaa	atcagctcat	tttttaacca	ataggccgaa	atcggcaaaa	1440
tcccttataa	atcaaaaagaa	tagaccgaga	tagggttgag	tgttgttcca	gtttggaaca	1500
agagtccact	attaaagaac	gtggactcca	acgtcaaagg	gcgaaaaacc	gtctatcagg	1560
gcgatggccc	actacgtgaa	ccatcacccct	aatcaagttt	tttggggtcg	aggtgccgta	1620
aagcactaaa	tcggaaccct	aaagggagcc	cccgatttag	agcttgacgg	ggaaagccgg	1680
cgaacgtggc	gagaaaggaa	gggaagaaag	cgaaggagc	gggcgctagg	gcgctggcaa	1740
gtgtagcgg	cacgctgcgc	gtaaccacca	caccgcgcgc	gcttaatgcg	ccgctacagg	1800
gcgcgtcgcg	ccattcgcca	ttcaggctgc	gcaactgttg	ggaagggcga	tcgggtgcggg	1860
cctcttcgct	attacgccag	ctggcgaaaag	ggggatgtgc	tgcaaggcga	ttaagttggg	1920
taacgccagg	gttttcccag	tcacgacgtt	gtaaaacgac	ggccagttag	cgcgcgtaat	1980
acgactcact	atagggcgaa	ttgggtaccg	ggccccccct	cgaggtcgac	ggtatcgata	2040
agcttgatat	cgaattcctg	cagcccgggg	gatccactag	ttctagagcg	gccattttaa	2100
acggccggcc	ctagatcaga	gggtggtaaa	tgaagtgtaa	tagtattcat	ttttcttata	2160
aatcatccct	tccgtgattt	atacaaaaga	agaggagaat	atgctgaata	cttggtatat	2220
tactctacat	tatactctta	tcttgacggg	tattctgagc	atcttactca	gtttcaagat	2280
cttttaaatgt	ccaaaaacat	ttgagccgat	ctaaatactt	ctgtgttttc	attaatttat	2340
aaattgtact	cttttaagac	atggaaagta	ccaacatcgg	ttgaaacagt	ttttcattta	2400
cttatggttt	attggttttt	ccagtgaatg	attatttgtc	gttacccttt	cgtaaaagtt	2460
caaacacgtt	tttaagtatt	gtttagttgc	tctttcgaca	tatatgatta	tccctgcgcg	2520
gctaaagtta	aggatgcaaa	aaacataaga	caactgaagt	taatttacgt	caattaagtt	2580
ttccagggtta	atgatgtttt	gggcttccac	taattcaata	agtatgtcat	gaaatacgtt	2640
gtgaagagca	tccagaaata	atgaaaagaa	acaacgaaac	tgggtcggcc	tggtgtttct	2700
tttctttacc	acgtgatctg	cggcattttac	aggaagtcgc	gcgtttttgcg	cagttgttgc	2760
aacgcagcta	cggctaacaa	agcctagtgg	aactcgactg	atgtgttagg	gcctaaaact	2820
ggtggtgaca	gctgaagtga	actattcaat	ccaatcatgt	catggctgtc	acaaagacct	2880
tgcggaccgc	acgtacgaac	acatacgtat	gctaatatgt	gttttgatag	taccagtgga	2940
tcgcagacct	gcaatttttt	tgtaggtttg	gaagaatata	taaaggttgc	actcattcaa	3000
gatagttttt	ttcttgtgtg	tctattcatt	ttattattgt	ttgtttaaat	gttaaaaaaa	3060
ccaagaactt	agtttcaaat	taaattcatc	acacaaacaa	acaaaacaaa	atgaacattg	3120

ttgtttgttt	gaagcaagtt	ccagacactg	ctgaagtcag	aattgaccca	gtcaagggtta	3180
ctttaatcag	agaaggtggt	ccatctatca	tcaaccacaga	cgacaagaac	gctttggaag	3240
aagcttttgt	tttgaaggac	aactacggtg	ctcacgttac	cgtcatttcc	atgggtccac	3300
ctcaagccaa	gaacgctttg	gttgaagctt	tggccatggg	tgctgatgaa	gctgtcttat	3360
tgactgacag	agctttcggg	gggtgctgata	cttttagctac	ctctcacacc	attgctgctg	3420
gtatcaagaa	attgaaatac	gatatcgtct	ttgccgggtcg	tcaagccatc	gatggtgata	3480
ccgctcaagt	cgggtccagaa	attgctgaac	atttgggtat	tccacaagtc	acctacgttg	3540
aaaaggttga	agttgacggt	gacactttga	agatcagaaa	ggcttgggaa	gacggttacg	3600
aagttgttga	agtcaagact	ccagttctat	tgactgccat	caaggaattg	aacgttccaa	3660
gatacatgtc	cgttgaaaaag	atcttcgggtg	ctttcgacaa	ggaagtcaag	atgtggactg	3720
ctgatgatat	cgatgtcgac	aaggccaact	tgggtttgaa	aggttctcca	accaaggtca	3780
agaaatcttc	taccaaggaa	gtcaaggggc	aaggtgaagt	cattgacaaa	ccagtcaagg	3840
aagctgccgc	ttacgttggt	tccaagttga	aggaagaaca	ctacatctaa	agcccgggcg	3900
gagattgata	agacttttct	agttgcatat	cttttatatt	taaatcttat	ctattagtta	3960
atTTTTTgta	atTTatcctt	atatatagtc	tggTtattct	aaaatatcat	ttcagtatct	4020
aaaaattccc	ctctTTTTtc	agttatatct	taacaggcga	cagtccaaat	gttgatttat	4080
cccagtccga	ttcatcaggg	ttgtgaagca	ttttgtcaat	ggtcgaaatc	acatcagtaa	4140
tagtgccctc	tacttgccctc	atagaatttc	tttctcttaa	cgtcaccggt	tggctcttta	4200
tagtttcgaa	atctatgggtg	ataccaaatg	gtgttcccaa	ttcatcggtta	cgggcgtatt	4260
ttttaccaat	tgaagtattg	gaatcgtcaa	ttttaaaagta	tatctctctt	ttacgtaaag	4320
cctgcgagat	cctcttaagt	atagcgggga	agccatcggt	attcgatatt	gtcgtaacaa	4380
atactttgat	cggcgctatg	tttaaagtgt	taaacatgga	cagatatgcg	atgaaaacgc	4440
taagtgatac	tccaaatggg	gaaagggtacg	atgcttggaa	acaatacttg	gaaatcaccg	4500
gaaacaccat	atgcggcgaa	aagccaatta	gtgtgatact	aagtgcttta	tcgaaaatcc	4560
gtgatgccgg	tccttcaggc	atcaaatttc	agtggcctaa	ttattcacag	agttctcatg	4620
tgacaagtat	tgatgatagt	agtgtcagtt	atgcttcagg	ttatgttact	ataggataat	4680
gatcacggct	aaaacgggtcg	aatgtaagca	tatatctttc	gattgtataa	ttgttcccaa	4740
atactacagc	atctcaagga	aaaaaaaaaca	aaaacttcca	aaaaaatcga	atccctgagg	4800

aatctttaat	acatthttcaa	tctatthttaag	thttataaac	gtgtatatga	gatgtcatga	4860
gcatgaatta	ttataataaa	aaactaaatc	attaaagtaa	cttaaggagt	taaagcccg	4920
gctthtaattg	ttagcagcct	tgacttgagc	aatcaattct	ggaacaacct	tgthgacatc	4980
accgacaatg	gccaaatcag	caaccttcat	gattggagct	tcgacatctt	tgthgatggc	5040
aatgatgtag	tcagagtctt	gcataaccagc	caagtgttg	atggcaccag	agataccaca	5100
agcaatgtac	aaagthggtc	tgacggctctt	accggtthtga	ccgacttgca	agtccttgtc	5160
aaacctattcc	thttcaatgg	cagctctgga	agcagcaatg	gtaccacca	aaaagaagc	5220
taattcttcc	aaththtcga	agththctt	ggaaccaaca	ccagaccac	cagcaacca	5280
aaccttggt	tcaccgatat	cagcaatgtc	cttggccaat	ttgacaacct	tggaacctt	5340
ggttctgata	tcagaagcag	tcaathtgat	ggcaacctth	tcgatcttgt	catcagaaac	5400
gttagcatcg	ttactggca	aththtcaaa	gacacctggt	ctgacggtg	ccathtgagg	5460
tctgtggtca	gaacagacaa	tggtagcaat	caagttacca	ccgaaagctg	gtctggtagc	5520
caacaagtca	cggtthtcga	catcgatatc	caaagaggta	cagtcagcag	tcaaaccagt	5580
agacaatctg	gcagcaattc	ttggacccaa	gtctctaccg	atgaaagtag	caccgatgaa	5640
taagatttct	ggctthctth	cgthgaccaa	gtcacagata	accttggcgt	aaccgtcagt	5700
ggagaaatga	gctaataatt	cgthgtcagc	agccaaaacc	ttgtcagcac	cgtgggacaa	5760
caagtccttg	gacatctth	cagtgthtg	accaataag	acagcagtca	attcaacacc	5820
caaththtca	gccathtct	taccttacc	tagcaattcc	aaagaaacct	thtgtaattc	5880
accatctctt	tgthcagcga	aaacctcagac	accttgtag	tcagccttgt	tcatgtthtag	5940
ttatattatag	ttcgthgacc	gtatattcta	aaaacaagta	ctcttataaa	aaaaccttg	6000
aagggaataa	acaagtagaa	tagatagaga	gaaaaataga	aatgcaaga	gaatttatat	6060
attagaaaga	gagaaagaaa	aatggaaaaa	aaaaaatagg	aaaagccaga	aatagcacta	6120
gaaggagcga	caccagaaaa	gaagggtgatg	gaaccaattt	agctatatat	agthtaactac	6180
cggtctgatc	atctctgcct	ccagcatagt	cgaagaagaa	ththththth	cttgaggctt	6240
ctgtcagcaa	ctcgththth	thctthctth	thtggtgagc	ctaaaaagth	cccacgttct	6300
cttgtagcag	gccgtcacia	acaaccttat	gggtaatttg	tcgcggtctg	ggtgtataaa	6360
tgtgtgggtg	caggccggcc	gtthaaacgg	gccgccaccg	cggtggagcc	tgtgtggaag	6420
aacgattaca	acaggtgttg	tcctctgagg	acataaaaata	cacaccgaga	thcatcaact	6480
cattgctgga	gttagcatat	ctacaattgg	gtgaaatggg	gagcgatttg	caggcatttg	6540

ctcggcatgc	cggtagaggt	gtggtcaata	agagcgacct	catgctatac	ctgagaaaagc	6600
aacctgacct	acaggaaaaga	gttactcaag	aataagaatt	ttcgttttaa	aacctaaagag	6660
tcactttaaa	atttgtatac	acttatTTTT	tttataactt	atttaataat	aaaaatcata	6720
aatcataaga	aattcgctcg	agtcgactgc	agtttatcta	atggagaaac	catcagtcaa	6780
gacacatctt	cttcttctag	cgaagtgcg	ggcagtggta	gtaccttcac	cagttggagt	6840
agcaatgggtg	aaagtgggtg	aaccttcacc	tctgaaacct	aaaccagcga	aagatggacc	6900
gttcttgaca	aagatggagg	tttgcattgc	acgggcagcc	ttgttcaatc	tggagatgtt	6960
ttgagagtgc	atggtagcag	tgtgatgtag	accttgttcc	aattcaatgg	caacttccaa	7020
agcttcatcg	aagtctggaa	ctctgacaac	tggacaatt	ggcatcaaca	attcaacagt	7080
agcgaatggg	tgggactttt	cagtttgcg	aatgatcaat	cttgggggtga	aatcacaagc	7140
aataccagct	tctttcaaga	tttcagtggc	agacttacca	accaatttct	tgttggtgac	7200
acccttgatc	gtgacggcaa	ccttttccaa	tttttgata	tcagatgggt	tagtgacgtg	7260
caaagcaccg	ttcttttcca	tttggaacaa	caagaagtca	gcaatggagt	caacggcaac	7320
aacagacttt	tcagcgatac	acaagatatt	atggtcaaag	gaagcaccgt	cgacaatgtc	7380
agcagcagcc	ttttcaatgt	tagcggtttc	gtcaacgatg	gatggagggt	taccagcacc	7440
agcaccgata	accttcttac	cagattgcat	agcttgcaag	acaacacctg	gaccaccagt	7500
gatgaccaac	aatggaacct	ttgggtggtt	catcatttct	tgagcagctt	ggatagatgg	7560
cttggaacg	gtgacaatca	agttgtcaat	accacaagaa	tctctgacga	tagtgttcaa	7620
cttttcaatc	aaccataaag	agatgttctt	ggcacctggg	tgaggagagt	agaaaacggc	7680
gttaccagca	gccaacatac	cgatggagtt	acagatcaaa	gtttcagttg	ggttggtaga	7740
tggagcaaca	gcaccgatga	caccgatatg	agataattcg	tataaagtca	taccgttgtc	7800
accggtagca	acttcagtgt	acaagtcttc	aacacctgga	gtcttttcga	tagctaaagt	7860
gttcttcaag	attttatcgg	tgacattacc	cataccggtt	tcagcaacag	ctctggtagc	7920
aatggtttcg	atttctgggt	ataaagcttc	tctgatggcc	ttgacaacgt	ttcttctttc	7980
ttcaaagat	ttttccttgt	aacagttttg	agcaatgacg	gcagcttggg	cagcttcatt	8040
gacggtatcg	aaaacaccgg	acttggcacc	ttgggtggtg	gtcttggttg	gaacttcctt	8100
ttgttcagcc	aatttttcca	acaaaacctt	cttgactaat	tgttccaatt	ccaaagattc	8160
cattttttac	tagttctaga	atccgtcgaa	actaagttct	ggtgttttaa	aactaaaaaa	8220

aagactaact	ataaaagtag	aatttaagaa	gtttaagaaa	tagatttaca	gaattacaat	8280
caatacctac	cgtctttata	tacttattag	tcaagtaggg	gaataatttc	agggaactgg	8340
tttcaacctt	ttttttcagc	tttttccaaa	tcagagagag	cagaaggtaa	tagaagggtg	8400
aagaaaatga	gatagataca	tgcgtgggtc	aattgccttg	tgtcatcatt	tactccaggc	8460
aggttgcatc	actccattga	ggttgtgccc	gttttttgcc	tgtttgtgcc	cctgttctct	8520
gtagttgctc	taagagaatg	gacctatgaa	ctgatgggtg	gtgaagaaaa	caatattttg	8580
gtgctgggat	tctttttttt	tctggatgcc	agcttaaaaa	gcgggctcca	ttatatattag	8640
tggatgccag	gaataaactg	ttcaccacaga	cacctacgat	gttatatatt	ctgtgtaacc	8700
cgccccctat	tttgggcatg	tacgggttac	agcagaatta	aaaggcta	tttttgacta	8760
aataaagtta	ggaaaatcac	tactattaat	tatttacgta	ttctttgaaa	tggcgagtat	8820
tgataatgat	aaactgagct	ccagcttttg	ttccctttag	tgagggttaa	ttgcgcgctt	8880
ggcgtaatca	tggtcatagc	tgtttcctgt	gtgaaattgt	tatccgctca	caattccaca	8940
caacatagga	gccggaagca	taaagtgtaa	agcctggggg	gcctaagtag	tgaggtaact	9000
cacattaatt	gcgttgctct	cactgcccgc	tttccagtcg	ggaaacctgt	cgtgccagct	9060
gcattaatga	atcggccaac	gcgcggggag	aggcggtttg	cgtattgggc	gctcttccgc	9120
ttcctcgctc	actgactcgc	tgcgctcggt	cgttcggctg	cggcgagcgg	tatcagctca	9180
ctcaaaggcg	gtaatacggg	tatccacaga	atcaggggat	aacgcaggaa	agaacatgtg	9240
agcaaaaggc	cagcaaaagg	ccaggaaccg	taaaaaggcc	gcgttgctgg	cgtttttcca	9300
taggctccgc	ccccctgacg	agcatcacia	aaatcgacgc	tcaagtcaga	ggtggcgaaa	9360
cccgacagga	ctataaagat	accaggcggt	tccccctgga	agctccctcg	tgcgctctcc	9420
tgttccgacc	ctgccgctta	ccggatacct	gtccgccttt	ctcccttcgg	gaagcgtggc	9480
gctttctcat	agctcacgct	gtaggtatct	cagttcgggtg	taggtcggtc	gctccaagct	9540
gggctgtgtg	cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	9600
tcttgagtcc	aaccgggtaa	gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	9660
gattagcaga	gcgaggtatg	taggcgggtg	tacagagtcc	ttgaagtggg	ggcctaacta	9720
cggctacact	agaaggacag	tatttggtat	ctgcgctctg	ctgaagccag	ttaccttcgg	9780
aaaaagagtt	gtagctctt	gatccggcaa	acaaaccacc	gctggtagcg	gtggtttttt	9840
tgtttgcaag	cagcagatta	cgcgcagaaa	aaaaggatct	caagaagatc	ctttgatctt	9900
ttctacgggg	tctgacgctc	agtggaaacga	aaactcacgt	taagggattt	tggtcatgag	9960

```

attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat 10020
ctaaagtata tatgagtaaa cttggtctga cagttaccaa tgcttaatca gtgaggcacc 10080
tatctcagcg atctgtctat ttcgttcac ccatagttgcc tgactccccg tcgtgtagat 10140
aactacgata cgggaggggct taccatctgg cccagtgct gcaatgatac cgcgagaccc 10200
acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgag 10260
aagtggctct gcaactttat ccgcctccat ccagtctatt aattggtgcc gggaagctag 10320
agtaagtagt tcgccagtta atagtttgcg caacgttggt gccattgcta caggcatcgt 10380
ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg 10440
agttacatga tccccatgt tgtgcaaaaa agcgggttagc tccttcggtc ctccgatcgt 10500
tgtcagaagt aagttggccg cagtgttata actcatgggt atggcagcac tgcataattc 10560
tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc 10620
attctgagaa tagtgatgc ggcgaccgag ttgctcttgc ccggcggtcaa tacgggataa 10680
taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggag 10740
aaaactctca aggatcttac cgctgttgag atccagttcg atgtaacca ctcgtgcacc 10800
caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag 10860
gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactctt 10920
cctttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt 10980
tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc 11040
acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaaata ggcgtatcac 11100
gaggcccttt cgtc 11114

```

```

<210> 48
<211> 2613
<212> DNA
<213> Entamoeba histolytica

```

```

<400> 48
atgtcaacac aacaaactat gactgtagat gaacatatta atcaacttgt tgctaaagca 60
caagttgcac ttaaagaata tcttaaacca gaatatacac aagaaaaaat agattatatt 120
gtaaagaaag catcagttgc agcacttgat caacattgtg cacttgcagc agctgcagtt 180
gaagaaacag gaagagggtat ttttgaagat aaagctacta aaaatatatt tgcattgtgaa 240
catgttacac atgaaatgag acatgctaaa acagttggta ttattaatgt agatccactt 300

```

tatggaatta	cagaaattgc	agaaccagtt	ggagttgttt	gtggagttac	accagttact	360
aatccaacat	caacagctat	tttcaagtca	cttattttcaa	ttaaaacaag	aatccaatt	420
gtattttcat	tccatccatc	agcacttaaa	tgttctatta	tggcagctaa	aattgtaga	480
gatgcagcta	ttgcagcagg	agcaccagaa	aattgtattc	aatggattga	atttggagga	540
attgaagcat	caaataaatt	aatgaatcat	ccaggagttg	ctactattct	tgctacagga	600
ggaaatgcta	tggttaaagc	agcatattca	tcaggaaaac	cagcacttgg	agtaggagca	660
ggaaatgtac	caacatatat	tgaaaaaaca	tgtaatatta	aacaagcagc	aatgatgta	720
gttatgtcaa	aatcatttga	taatggtatg	atttgtgcat	cagaacaagc	agcaattatt	780
gataaagaaa	tttatgatca	agtagttgaa	gaaatgaaaa	cacttggagc	atatttcatt	840
aatgaagaag	aaaaagctaa	attagaaaag	tttatgtttg	gagttaatgc	atattcagca	900
gatgttaata	atgcaagact	taatccaaaa	tgtccaggta	tgtcaccaca	atggtttgct	960
gaacaagttg	gaattaaagt	tccagaagat	tgtaatatta	tttgtgcagt	ttgtaaagaa	1020
gttggaccaa	atgaaccatt	aacaagagaa	aaattatcac	cagttcttgc	tattcttaaa	1080
gcagaaaata	cacaagatgg	tattgataaa	gctgaagcta	tggttgaatt	taatggtaga	1140
ggacattcag	cagctattca	ttcgaatgat	aaagcagtag	ttgaaaagta	tgcaactaca	1200
atgaaagcat	gcagaatttt	acataataca	ccatcatcac	aaggaggaat	tggtatcaatt	1260
tataactata	tttggccatc	atttacactt	ggatgtggat	catatggagg	aaattcggta	1320
tcagctaattg	ttacatatca	taattttatta	aatatttaaaa	gacttgcaga	tagaagaaac	1380
aaccttcaat	ggttcagagt	tccaccaaag	atthttctttg	aaccacattc	tattagatat	1440
cttgctgaac	ttaaggaact	tagtaaaata	ttcattgttt	cagatagaat	gatgtataaa	1500
ttaggatatg	tagatagagt	tatggatgta	ttgaaaagaa	gaagtaatga	agtagaaatt	1560
gaaatthtca	ttgatgtaga	accagatcca	tctattcaaa	ccgttcaaaa	aggacttgct	1620
gttatgaata	catttggacc	agataatatt	attgctattg	gaggaggatc	agctatggat	1680
gcagctaaga	ttatgtggtt	actttatgaa	catccagaag	ccgatttctt	tgcaatgaaa	1740
caaaaattca	ttgatcttag	aaagagagca	tttaaattcc	caacaatggg	taagaaagct	1800
agattaatth	gtattccaac	aacatcagga	actggatcag	aagttacacc	atttgcagtt	1860
atttcagatc	atgaaacagg	taagaaatat	ccacttgctg	attattcact	tacaccatca	1920
gttgctattg	ttgatccaat	gtttactatg	tcacttccaa	agagagctat	tgctgatact	1980

```

ggacttgatg tattggttca tgcaacagaa gcatatgttt cagttatggc taatgaatat 2040
actgatggac ttgctagaga agcagttaaa ttagtctttg aaaatcttct taaatcatat 2100
aatggagatt tagaagcaag agaaaagatg cacaatgctg caacaattgc aggtatggca 2160
tttgcacag cattccttgg tatggaccat tccatggcac ataaagttgg agcagcattc 2220
catcttccac atggtagatg tgtagcagta ttattaccac atgtcattag atataatgga 2280
caaaaaccaa gaaagcttgc aatgtggcca aaatataatt tctataaggc agaccaaaga 2340
tatatggaac ttgcacaaat ggttggactt aaatgtaata caccagctga aggagttaa 2400
gcatttgcta aagcatgtga agaattaatg aaagccacag agactattac tggattcaag 2460
caagcaaata ttgatgaagc agcatggatg agtaaagtac cagaaatggc acttcttgca 2520
tttgaagatc aatgttcacc agctaatacca agagtcccaa tggttaagga tatggaaaag 2580
attctcaaag ctgcatatta tccaattgct tga 2613

```

```

<210> 49
<211> 2610
<212> DNA
<213> Artificial sequence

<220>
<223> adh2 E. histolytica codon pair optimised

```

```

<400> 49
atgtccactc aacaaaccat gaccgttgat gaacacatta accaattggc cagaaagget 60
caagttgctt tgaaggaata cttgaaacca gaatacactc aagaaaagat cgattacatt 120
gtcaagaagg cttctgttgc tgctctagac caacactgtg ctttggctgc tgctgctgct 180
gaagaaactg gtcgtggtat ctttgaagac aaagctacca agaacatttt cgcttgtaaa 240
cacgtcactc acgaaatgag acacgccaag accgttggta tcatcaacgt tgatccatta 300
tacggtatca ctgaaattgc tgaaccagtc ggtgttgtct gtggtgtcac ccagttacc 360
aaccacaatt ctactgccat tttcaaactt ttgatttcca tcaagaccag aaaccaatt 420
gttttctcct tccaccatc tgctttgaaa tgttccatca tggctgcca gatcgtcaga 480
gatgctgcca ttgctgctgg tgctccagaa aactgtatcc aatggatcga atttgggtgt 540
attgaagctt ccaacaaatt gatgaaccat cctggtgttg ctaccatctt agctactggt 600
ggtaacgcta tggtaaggc tgcttactct tctggtaagc cagctttggg tgcggtgct 660
ggtaacgtcc caacttacat cgaaaagacc tgtaatatca agcaagctgc taacgatgtt 720
gtcatgtcca agtctttcga caacggtatg atctgtgctt ccgaacaagc tgccatcatc 780

```

gacaaagaaa	tctacgacca	agttgttgaa	gaaatgaaga	ctttgggtgc	ttacttcac	840
aacgaagaag	aaaaggccaa	attggaaaaa	ttcatgttcg	gtgttaatgc	ttactctgct	900
gatgtcaaca	acgccagatt	gaacccaaag	tgtccaggta	tgtctccaca	atgggttcgct	960
gaacaagtgc	gtatcaaggt	tccagaagac	tgtaacatca	tctgtgccgt	ttgtaaggaa	1020
gttgggtccaa	acgaaccatt	gaccagagaa	aagttgtctc	cagttttggc	cattttgaag	1080
gctgaaaaca	ctcaagatgg	tattgacaag	gctgaagcta	tggtcgaatt	caacggtcgt	1140
ggtcactctg	ctgccattca	ctccaatgac	aaggctgttg	ttgaaaaata	cgctttgacc	1200
atgaaggctt	gtcgtatctt	gcacaacact	ccatcttctc	aaggtgggat	cggttccatt	1260
tacaactaca	tctggccatc	tttcaacttta	ggttgtgggt	cttacgggtg	taactccgtt	1320
tctgccaatg	ttacctacca	caacttggtg	aacatcaaga	gattggctga	cagaagaaac	1380
aacttacaat	ggttcagagt	cccaccaaag	atcttcttcg	aacctcactc	cattagatac	1440
ttggctgaat	tgaaggaatt	gtccaagatt	ttcattgtct	ctgacagaat	gatgtacaaa	1500
ttgggttacg	ttgacagagt	tatggatgtc	ttgaagagaa	gatccaacga	agttgaaatt	1560
gaaatcttca	tcgatgttga	accagaccca	tccattcaaa	ccgtccaaaa	gggtttggct	1620
gtcatgaaca	ctttcgggtc	agacaacatc	attgccattg	gtgggtgggtc	tgccatggat	1680
gctgccaaga	tcatgtgggt	attatacgaa	catccagaag	ctgatttctt	cgctatgaag	1740
caaaaattca	tcgatttaag	aaagagagct	ttcaagttcc	caaccatggg	taagaaggcc	1800
agattaatct	gtatcccaac	cacttctggg	accggttctg	aagtcacccc	attcgctgtc	1860
atctctgacc	acgaaactgg	taagaagtat	ccattggctg	actactcttt	gaccccatcc	1920
gttgccattg	ttgacccaat	gtttaccatg	tccttgccct	agagagccat	tgctgacact	1980
ggtttgatg	tcttagtcca	cgctactgaa	gcttacgttt	ctgttatggc	taacgaatac	2040
actgacgggt	tggccagaga	agctgtcaaa	ttggtttttcg	aaaacttggt	gaaatcttac	2100
aacggtgact	tggaagctcg	tgaaaagatg	cacaacgctg	ctaccattgc	tggtatggcc	2160
tttgcttctg	ctttcttggg	tatggaccat	tccatggctc	acaaggctcg	tgctgctttc	2220
catttgccac	acggtagatg	tggtgcccgt	ttgttgccctc	acgttatcag	atacaacggt	2280
caaaagccaa	gaaagttggc	catgtggcca	aagtacaact	tctacaaggc	tgatcaaaga	2340
tacatggaat	tggtctaaat	ggtcggtttg	aagtgttaaca	ccccagctga	aggtgtcgaa	2400
gcctttgcc	aggcttgatg	agaattgatg	aaggctactg	aaaccatcac	tggtttcaag	2460

```

aaggccaaca ttgatgaagc tgcttggatg tccaagggtc cagaaatggc tctattggct 2520
ttcgaagacc aatgttctcc agctaaccce agagtcccaa tggttaagga catggaaaag 2580
atattgaagg ctgcttacta cccaatcgct 2610

```

```

<210> 50
<211> 870
<212> PRT
<213> Entamoeba histolytica

```

```

<400> 50

```

```

Met Ser Thr Gln Gln Thr Met Thr Val Asp Glu His Ile Asn Gln Leu
1          5          10          15

```

```

Val Arg Lys Ala Gln Val Ala Leu Lys Glu Tyr Leu Lys Pro Glu Tyr
          20          25          30

```

```

Thr Gln Glu Lys Ile Asp Tyr Ile Val Lys Lys Ala Ser Val Ala Ala
          35          40          45

```

```

Leu Asp Gln His Cys Ala Leu Ala Ala Ala Ala Val Glu Glu Thr Gly
          50          55          60

```

```

Arg Gly Ile Phe Glu Asp Lys Ala Thr Lys Asn Ile Phe Ala Cys Glu
65          70          75          80

```

```

His Val Thr His Glu Met Arg His Ala Lys Thr Val Gly Ile Ile Asn
          85          90          95

```

```

Val Asp Pro Leu Tyr Gly Ile Thr Glu Ile Ala Glu Pro Val Gly Val
          100          105          110

```

```

Val Cys Gly Val Thr Pro Val Thr Asn Pro Thr Ser Thr Ala Ile Phe
          115          120          125

```

```

Lys Ser Leu Ile Ser Ile Lys Thr Arg Asn Pro Ile Val Phe Ser Phe
          130          135          140

```

```

His Pro Ser Ala Leu Lys Cys Ser Ile Met Ala Ala Lys Ile Val Arg
          145          150          155          160

```

```

Asp Ala Ala Ile Ala Ala Gly Ala Pro Glu Asn Cys Ile Gln Trp Ile
          165          170          175

```

Glu Phe Gly Gly Ile Glu Ala Ser Asn Lys Leu Met Asn His Pro Gly
 180 185 190

Val Ala Thr Ile Leu Ala Thr Gly Gly Asn Ala Met Val Lys Ala Ala
 195 200 205

Tyr Ser Ser Gly Lys Pro Ala Leu Gly Val Gly Ala Gly Asn Val Pro
 210 215 220

Thr Tyr Ile Glu Lys Thr Cys Asn Ile Lys Gln Ala Ala Asn Asp Val
 225 230 235 240

Val Met Ser Lys Ser Phe Asp Asn Gly Met Ile Cys Ala Ser Glu Gln
 245 250 255

Ala Ala Ile Ile Asp Lys Glu Ile Tyr Asp Gln Val Val Glu Glu Met
 260 265 270

Lys Thr Leu Gly Ala Tyr Phe Ile Asn Glu Glu Glu Lys Ala Lys Leu
 275 280 285

Glu Lys Phe Met Phe Gly Val Asn Ala Tyr Ser Ala Asp Val Asn Asn
 290 295 300

Ala Arg Leu Asn Pro Lys Cys Pro Gly Met Ser Pro Gln Trp Phe Ala
 305 310 315 320

Glu Gln Val Gly Ile Lys Val Pro Glu Asp Cys Asn Ile Ile Cys Ala
 325 330 335

Val Cys Lys Glu Val Gly Pro Asn Glu Pro Leu Thr Arg Glu Lys Leu
 340 345 350

Ser Pro Val Leu Ala Ile Leu Lys Ala Glu Asn Thr Gln Asp Gly Ile
 355 360 365

Asp Lys Ala Glu Ala Met Val Glu Phe Asn Gly Arg Gly His Ser Ala
 370 375 380

Ala Ile His Ser Asn Asp Lys Ala Val Val Glu Lys Tyr Ala Leu Thr
 385 390 395 400

Met Lys Ala Cys Arg Ile Leu His Asn Thr Pro Ser Ser Gln Gly Gly
 405 410 415

Ile Gly Ser Ile Tyr Asn Tyr Ile Trp Pro Ser Phe Thr Leu Gly Cys
 420 425 430

Gly Ser Tyr Gly Gly Asn Ser Val Ser Ala Asn Val Thr Tyr His Asn
 435 440 445

Leu Leu Asn Ile Lys Arg Leu Ala Asp Arg Arg Asn Asn Leu Gln Trp
 450 455 460

Phe Arg Val Pro Pro Lys Ile Phe Phe Glu Pro His Ser Ile Arg Tyr
 465 470 475 480

Leu Ala Glu Leu Lys Glu Leu Ser Lys Ile Phe Ile Val Ser Asp Arg
 485 490 495

Met Met Tyr Lys Leu Gly Tyr Val Asp Arg Val Met Asp Val Leu Lys
 500 505 510

Arg Arg Ser Asn Glu Val Glu Ile Glu Ile Phe Ile Asp Val Glu Pro
 515 520 525

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

Phe Gly Pro Asp Asn Ile Ile Ala Ile Gly Gly Gly Ser Ala Met Asp
 545 550 555 560

Ala Ala Lys Ile Met Trp Leu Leu Tyr Glu His Pro Glu Ala Asp Phe
 565 570 575

Phe Ala Met Lys Gln Lys Phe Ile Asp Leu Arg Lys Arg Ala Phe Lys
 580 585 590

Phe Pro Thr Met Gly Lys Lys Ala Arg Leu Ile Cys Ile Pro Thr Thr
 595 600 605

Ser Gly Thr Gly Ser Glu Val Thr Pro Phe Ala Val Ile Ser Asp His
 610 615 620

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

625					630					635				640	
Val	Ala	Ile	Val	Asp	Pro	Met	Phe	Thr	Met	Ser	Leu	Pro	Lys	Arg	Ala
				645					650					655	
Ile	Ala	Asp	Thr	Gly	Leu	Asp	Val	Leu	Val	His	Ala	Thr	Glu	Ala	Tyr
			660					665					670		
Val	Ser	Val	Met	Ala	Asn	Glu	Tyr	Thr	Asp	Gly	Leu	Ala	Arg	Glu	Ala
		675					680					685			
Val	Lys	Leu	Val	Phe	Glu	Asn	Leu	Leu	Lys	Ser	Tyr	Asn	Gly	Asp	Leu
	690					695					700				
Glu	Ala	Arg	Glu	Lys	Met	His	Asn	Ala	Ala	Thr	Ile	Ala	Gly	Met	Ala
705					710					715					720
Phe	Ala	Ser	Ala	Phe	Leu	Gly	Met	Asp	His	Ser	Met	Ala	His	Lys	Val
				725					730					735	
Gly	Ala	Ala	Phe	His	Leu	Pro	His	Gly	Arg	Cys	Val	Ala	Val	Leu	Leu
			740					745					750		
Pro	His	Val	Ile	Arg	Tyr	Asn	Gly	Gln	Lys	Pro	Arg	Lys	Leu	Ala	Met
		755					760					765			
Trp	Pro	Lys	Tyr	Asn	Phe	Tyr	Lys	Ala	Asp	Gln	Arg	Tyr	Met	Glu	Leu
	770					775					780				
Ala	Gln	Met	Val	Gly	Leu	Lys	Cys	Asn	Thr	Pro	Ala	Glu	Gly	Val	Glu
785					790					795					800
Ala	Phe	Ala	Lys	Ala	Cys	Glu	Glu	Leu	Met	Lys	Ala	Thr	Glu	Thr	Ile
				805					810					815	
Thr	Gly	Phe	Lys	Lys	Ala	Asn	Ile	Asp	Glu	Ala	Ala	Trp	Met	Ser	Lys
			820					825					830		
Val	Pro	Glu	Met	Ala	Leu	Leu	Ala	Phe	Glu	Asp	Gln	Cys	Ser	Pro	Ala
		835					840					845			
Asn	Pro	Arg	Val	Pro	Met	Val	Lys	Asp	Met	Glu	Lys	Ile	Leu	Lys	Ala
	850					855					860				

Ala Tyr Tyr Pro Ile Ala
865 870

<210> 51
<211> 2658
<212> DNA
<213> *Piromyces* sp. E2

<400> 51
atgtccggat tacaaatggt ccaaaacctt tctctttacg gtagtctcgc cgaaatcgat 60
actagcgaaa agcttaacga agctatggac aaattaactg ctgccaaga acaattcaga 120
gaatacaacc aagaacaagt tgacaaaatc ttcaaggctg ttgcttttagc tgcttctcaa 180
aaccgtgttg ctttcgctaa gtacgcacac gaagaaaccc aaaaggggtgt tttcgaagat 240
aaggttatca agaacgaatt cgctgctgat tacatttacc acaagtactg caatgacaag 300
accgccggta tcattgaata tgatgaagcc aatgggtctta tggaaattgc tgaaccagtt 360
gggccagttg ttggtattgc tccagttact aacccaactt ctactatcat ctacaagtct 420
ttaattgcct taaagacccg taactgtatt atcttctcac cacatccagg agctcacaag 480
gcctctgttt tcgttggttaa ggtcttacac caagctgctg ttaaggctgg tgccccagaa 540
aactgtattc aaatcatctt cccaaagatg gatttaacta ctgaattatt acaccaccaa 600
aagactcggt tcatttgggc tactgggtgg ccagggttag ttcacgcctc ttacacttct 660
ggtaagccag ctcttggtgg tgggtccagg aatgctccag ctcttattga tgaaacttgt 720
gatatgaacg aagctgttgg ttctatcggt gtttctaaga ctttcgattg tggtatgatc 780
tgtgccactg aaaacgctgt tgcgttggtc gaatctgtct acgaaaactt cgttgctacc 840
atgaagaagc gtggtgccta cttcatgact ccagaagaaa ccaagaaggc ttctaacctt 900
cttttcggag aaggatatgag attaaatgct aaggctgttg gtcaaactgc caagacttta 960
gctgaaatgg ccggtttcga agtcccagaa aacaccgttg ttctctgtgg tgaagcttct 1020
gaagttaa at tcgaagaacc aatggctcac gaaaagttaa ctactatcct cggatatctac 1080
aaggctaagg actttgacga tgggtgtcaga ttatgtaagg aattagttac tttcgggtgg 1140
aagggtcaca ctgctgttct ctacaccaac caaaacaaca aggaccgtat tgaaaagtac 1200
caaaacgaag ttccagcctt ccacatctta gttgacatgc catcttcctt cggttgattt 1260
ggatgatgtt acaacttccg tcttgctcca gctcttacca ttacttgtgg tactatgggt 1320
gggtggttct cctctgataa cattgggtcca aagcacttac ttaacatcaa gcgtgttggt 1380

```

atgagacgcg aaaacatgct ttggttcaag attccaaagt ctgtctactt caagcgtgct 1440
atccttttctg aagctttatc tgacttacgt gacaccaca agcgtgctat cattattacc 1500
gatagaacta tgactatggt aggtcaaact gacaagatca ttaaggcttg tgaaggcat 1560
ggtatggtct gcaactgtcta cgataagggt gtcccagatc caactatcaa gtgtattatg 1620
gaagggtgta atgaaatgaa cgtcttcaag ccagatttag ctattgctct tgggtggtggt 1680
tctgctatgg atgccgctaa gatgatgcgt ttattctacg aatacccaga ccaagactta 1740
caagatattg ctactcgttt cgtcgatata cgtaagcgtg ttggttggtg tccaaagctt 1800
ggtagactta ttaagactct tgtctgtatc ccaactacct ctggtactgg tgccgaagtt 1860
actccattcg ctgtcgttac ctctgaagaa ggtcgtaagt acccattagt cgactacgaa 1920
cttactccag atatggctat tgttgatcca gaattcgctg ttggtatgcc aaagcgttta 1980
acttcttgga ctggtattga tgctcttacc cagccattg aatcttacgt ttctattatg 2040
gctactgact tcactagacc atactctctc cgtgctggtg gtcttatctt cgaatccctt 2100
tcccttgctt acaacaacgg taaggatatt gaagctcgtg aaaagatgca caatgcttct 2160
gctattgctg gtatggcctt tgccaacgct ttccttggtt gttgtcactc tgttgctcac 2220
caacttggtt ccgtctacca cattccacac ggtcttgcca acgctttaat gctttctcac 2280
atcattaagt acaacgctac tgactctcca gttaagatgg gtaccttccc acaatacaag 2340
taccacaag ctatgcgtca ctacgtgaa attgctgaac tcttattacc accaactcaa 2400
gttgттаага tgactgatgt tgataagggt caatacttaa ttgaccgtgt tgaacaatta 2460
aaggctgacg ttggtattcc aaagtctatt aaggaaactg gaatggttac tgaagaagac 2520
ttcttcaaca aggttgacca agttgctatc atggccttcg atgaccaatg tactggtgct 2580
aaccacggt acccattagt ttctgaatta aaacaattaa tgattgatgc ctggaacggt 2640
gttggtccaa agctctaa 2658

```

```

<210> 52
<211> 885
<212> PRT
<213> Piromyces sp. E2

```

```
<400> 52
```

```

Met Ser Gly Leu Gln Met Phe Gln Asn Leu Ser Leu Tyr Gly Ser Leu
1           5           10          15

```

Ala Glu Ile Asp Thr Ser Glu Lys Leu Asn Glu Ala Met Asp Lys Leu
 20 25 30
 Thr Ala Ala Gln Glu Gln Phe Arg Glu Tyr Asn Gln Glu Gln Val Asp
 35 40 45
 Lys Ile Phe Lys Ala Val Ala Leu Ala Ala Ser Gln Asn Arg Val Ala
 50 55 60
 Phe Ala Lys Tyr Ala His Glu Glu Thr Gln Lys Gly Val Phe Glu Asp
 65 70 75 80
 Lys Val Ile Lys Asn Glu Phe Ala Ala Asp Tyr Ile Tyr His Lys Tyr
 85 90 95
 Cys Asn Asp Lys Thr Ala Gly Ile Ile Glu Tyr Asp Glu Ala Asn Gly
 100 105 110
 Leu Met Glu Ile Ala Glu Pro Val Gly Pro Val Val Gly Ile Ala Pro
 115 120 125
 Val Thr Asn Pro Thr Ser Thr Ile Ile Tyr Lys Ser Leu Ile Ala Leu
 130 135 140
 Lys Thr Arg Asn Cys Ile Ile Phe Ser Pro His Pro Gly Ala His Lys
 145 150 155 160
 Ala Ser Val Phe Val Val Lys Val Leu His Gln Ala Ala Val Lys Ala
 165 170 175
 Gly Ala Pro Glu Asn Cys Ile Gln Ile Ile Phe Pro Lys Met Asp Leu
 180 185 190
 Thr Thr Glu Leu Leu His His Gln Lys Thr Arg Phe Ile Trp Ala Thr
 195 200 205
 Gly Gly Pro Gly Leu Val His Ala Ser Tyr Thr Ser Gly Lys Pro Ala
 210 215 220
 Leu Gly Gly Gly Pro Gly Asn Ala Pro Ala Leu Ile Asp Glu Thr Cys
 225 230 235 240
 Asp Met Asn Glu Ala Val Gly Ser Ile Val Val Ser Lys Thr Phe Asp

Ile Leu Ser Glu Ala Leu Ser Asp Leu Arg Asp Thr His Lys Arg Ala
485 490 495

Ile Ile Ile Thr Asp Arg Thr Met Thr Met Leu Gly Gln Thr Asp Lys
500 505 510

Ile Ile Lys Ala Cys Glu Gly His Gly Met Val Cys Thr Val Tyr Asp
515 520 525

Lys Val Val Pro Asp Pro Thr Ile Lys Cys Ile Met Glu Gly Val Asn
530 535 540

Glu Met Asn Val Phe Lys Pro Asp Leu Ala Ile Ala Leu Gly Gly Gly
545 550 555 560

Ser Ala Met Asp Ala Ala Lys Met Met Arg Leu Phe Tyr Glu Tyr Pro
565 570 575

Asp Gln Asp Leu Gln Asp Ile Ala Thr Arg Phe Val Asp Ile Arg Lys
580 585 590

Arg Val Val Gly Cys Pro Lys Leu Gly Arg Leu Ile Lys Thr Leu Val
595 600 605

Cys Ile Pro Thr Thr Ser Gly Thr Gly Ala Glu Val Thr Pro Phe Ala
610 615 620

Val Val Thr Ser Glu Glu Gly Arg Lys Tyr Pro Leu Val Asp Tyr Glu
625 630 635 640

Leu Thr Pro Asp Met Ala Ile Val Asp Pro Glu Phe Ala Val Gly Met
645 650 655

Pro Lys Arg Leu Thr Ser Trp Thr Gly Ile Asp Ala Leu Thr His Ala
660 665 670

Ile Glu Ser Tyr Val Ser Ile Met Ala Thr Asp Phe Thr Arg Pro Tyr
675 680 685

Ser Leu Arg Ala Val Gly Leu Ile Phe Glu Ser Leu Ser Leu Ala Tyr
690 695 700

Asn Asn Gly Lys Asp Ile Glu Ala Arg Glu Lys Met His Asn Ala Ser
 705 710 715 720

Ala Ile Ala Gly Met Ala Phe Ala Asn Ala Phe Leu Gly Cys Cys His
 725 730 735

Ser Val Ala His Gln Leu Gly Ser Val Tyr His Ile Pro His Gly Leu
 740 745 750

Ala Asn Ala Leu Met Leu Ser His Ile Ile Lys Tyr Asn Ala Thr Asp
 755 760 765

Ser Pro Val Lys Met Gly Thr Phe Pro Gln Tyr Lys Tyr Pro Gln Ala
 770 775 780

Met Arg His Tyr Ala Glu Ile Ala Glu Leu Leu Leu Pro Pro Thr Gln
 785 790 795 800

Val Val Lys Met Thr Asp Val Asp Lys Val Gln Tyr Leu Ile Asp Arg
 805 810 815

Val Glu Gln Leu Lys Ala Asp Val Gly Ile Pro Lys Ser Ile Lys Glu
 820 825 830

Thr Gly Met Val Thr Glu Glu Asp Phe Phe Asn Lys Val Asp Gln Val
 835 840 845

Ala Ile Met Ala Phe Asp Asp Gln Cys Thr Gly Ala Asn Pro Arg Tyr
 850 855 860

Pro Leu Val Ser Glu Leu Lys Gln Leu Met Ile Asp Ala Trp Asn Gly
 865 870 875 880

Val Val Pro Lys Leu
 885