

SEQUENCE LISTING

<110> ZHAO, Xinhua
 KEALEY, James
 HUYNH, Tina

<120> METHODS AND MICROORGANISMS FOR MAKING 1,4-BUTANEDIOL AND
 DERIVATIVES THEREOF FROM C1 CARBONS

<130> INX00386

<160> 58

<170> PatentIn version 3.5

<210> 1
 <211> 1231
 <212> PRT
 <213> Mycobacterium bovis

<400> 1

Met Ala Asn Ile Ser Ser Pro Phe Gly Gln Asn Glu Trp Leu Val Glu
 1 5 10 15

Ala Met Tyr Arg Lys Phe Arg Asp Asp Pro Ser Ser Val Asp Pro Ser
 20 25 30

Trp His Glu Phe Leu Val Asp Tyr Ser Pro Glu Pro Thr Ser Gln Pro
 35 40 45

Ala Ala Glu Pro Thr Arg Val Thr Ser Pro Leu Val Ala Glu Arg Ala
 50 55 60

Ala Ala Ala Ala Pro Gln Ala Pro Pro Lys Pro Ala Asp Thr Ala Ala
 65 70 75 80

Ala Gly Asn Gly Val Val Ala Ala Leu Ala Ala Lys Thr Ala Val Pro
 85 90 95

Pro Pro Ala Glu Gly Asp Glu Val Ala Val Leu Arg Gly Ala Ala Ala
 100 105 110

Ala Val Val Lys Asn Met Ser Ala Ser Leu Glu Val Pro Thr Ala Thr
 115 120 125

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

Ile Asn Asn Gln Leu Lys Arg Thr Arg Gly Gly Lys Ile Ser Phe Thr
145 150 155 160

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

Met Asn Arg His Tyr Thr Glu Val Asp Gly Lys Pro Thr Ala Val Thr
180 185 190

Pro Ala His Thr Asn Leu Gly Leu Ala Ile Asp Leu Gln Gly Lys Asp
195 200 205

Gly Lys Arg Ser Leu Val Val Ala Gly Ile Lys Arg Cys Glu Thr Met
210 215 220

Arg Phe Ala Gln Phe Val Thr Ala Tyr Glu Asp Ile Val Arg Arg Ala
225 230 235 240

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

Leu Thr Asn Pro Gly Thr Ile Gly Thr Val His Ser Val Pro Arg Leu
260 265 270

Met Pro Gly Gln Gly Ala Ile Ile Gly Val Gly Ala Met Glu Tyr Pro
275 280 285

Ala Glu Phe Gln Gly Ala Ser Glu Glu Arg Ile Ala Glu Leu Gly Ile
290 295 300

Gly Lys Leu Ile Thr Leu Thr Ser Thr Tyr Asp His Arg Ile Ile Gln
305 310 315 320

Gly Ala Glu Ser Gly Asp Phe Leu Arg Thr Ile His Glu Leu Leu Leu
325 330 335

Ser Asp Gly Phe Trp Asp Glu Val Phe Arg Glu Leu Ser Ile Pro Tyr
340 345 350

Leu Pro Val Arg Trp Ser Thr Asp Asn Pro Asp Ser Ile Val Asp Lys
355 360 365

Asn Ala Arg Val Met Asn Leu Ile Ala Ala Tyr Arg Asn Arg Gly His
370 375 380

Leu Met Ala Asp Thr Asp Pro Leu Arg Leu Asp Lys Ala Arg Phe Arg
385 390 395 400

Ser His Pro Asp Leu Glu Val Leu Thr His Gly Leu Thr Leu Trp Asp
405 410 415

Leu Asp Arg Val Phe Lys Val Asp Gly Phe Ala Gly Ala Gln Tyr Lys
420 425 430

Lys Leu Arg Asp Val Leu Gly Leu Leu Arg Asp Ala Tyr Cys Arg His
435 440 445

Ile Gly Val Glu Tyr Ala His Ile Leu Asp Pro Glu Gln Lys Glu Trp
450 455 460

Leu Glu Gln Arg Val Glu Thr Lys His Val Lys Pro Thr Val Ala Gln
465 470 475 480

Gln Lys Tyr Ile Leu Ser Lys Leu Asn Ala Ala Glu Ala Phe Glu Thr
485 490 495

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

Ala Glu Ser Val Ile Pro Met Met Asp Ala Ala Ile Asp Gln Cys Ala
515 520 525

Glu His Gly Leu Asp Glu Val Val Ile Gly Met Pro His Arg Gly Arg
530 535 540

Leu Asn Val Leu Ala Asn Ile Val Gly Lys Pro Tyr Ser Gln Ile Phe
545 550 555 560

Thr Glu Phe Glu Gly Asn Leu Asn Pro Ser Gln Ala His Gly Ser Gly
565 570 575

Asp Val Lys Tyr His Leu Gly Ala Thr Gly Leu Tyr Leu Gln Met Phe
580 585 590

Gly Asp Asn Asp Ile Gln Val Ser Leu Thr Ala Asn Pro Ser His Leu
595 600 605

Glu Ala Val Asp Pro Val Leu Glu Gly Leu Val Arg Ala Lys Gln Asp
610 615 620

Leu Leu Asp His Gly Ser Ile Asp Ser Asp Gly Gln Arg Ala Phe Ser
625 630 635 640

Val Val Pro Leu Met Leu His Gly Asp Ala Ala Phe Ala Gly Gln Gly
645 650 655

Val Val Ala Glu Thr Leu Asn Leu Ala Asn Leu Pro Gly Tyr Arg Val
660 665 670

Gly Gly Thr Ile His Ile Ile Val Asn Asn Gln Ile Gly Phe Thr Thr
675 680 685

Ala Pro Glu Tyr Ser Arg Ser Ser Glu Tyr Cys Thr Asp Val Ala Lys
690 695 700

Met Ile Gly Ala Pro Ile Phe His Val Asn Gly Asp Asp Pro Glu Ala
705 710 715 720

Cys Val Trp Val Ala Arg Leu Ala Val Asp Phe Arg Gln Arg Phe Lys
725 730 735

Lys Asp Val Val Ile Asp Met Leu Cys Tyr Arg Arg Arg Gly His Asn
740 745 750

Glu Gly Asp Asp Pro Ser Met Thr Asn Pro Tyr Met Tyr Asp Val Val
755 760 765

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

Arg Gly Asp Ile Ser Met Lys Glu Ala Glu Asp Ala Leu Arg Asp Tyr
785 790 795 800

Gln Gly Gln Leu Glu Arg Val Phe Asn Glu Val Arg Glu Leu Glu Lys
805 810 815

His Gly Val Gln Pro Ser Glu Ser Val Glu Ser Asp Gln Met Ile Pro
820 825 830

Ala Gly Leu Ala Thr Ala Val Asp Lys Ser Leu Leu Ala Arg Ile Gly
835 840 845

Asp Ala Phe Leu Ala Leu Pro Asn Gly Phe Thr Ala His Pro Arg Val
850 855 860

Gln Pro Val Leu Glu Lys Arg Arg Glu Met Ala Tyr Glu Gly Lys Ile
865 870 875 880

Asp Trp Ala Phe Gly Glu Leu Leu Ala Leu Gly Ser Leu Val Ala Glu
885 890 895

Gly Lys Leu Val Arg Leu Ser Gly Gln Asp Ser Arg Arg Gly Thr Phe
900 905 910

Ser Gln Arg His Ser Val Leu Ile Asp Arg His Thr Gly Glu Glu Phe
915 920 925

Thr Pro Leu Gln Leu Leu Ala Thr Asn Ser Asp Gly Ser Pro Thr Gly
930 935 940

Gly Lys Phe Leu Val Tyr Asp Ser Pro Leu Ser Glu Tyr Ala Ala Val
945 950 955 960

Gly Phe Glu Tyr Gly Tyr Thr Val Gly Asn Pro Asp Ala Val Val Leu
965 970 975

Trp Glu Ala Gln Phe Gly Asp Phe Val Asn Gly Ala Gln Ser Ile Ile
980 985 990

Asp Glu Phe Ile Ser Ser Gly Glu Ala Lys Trp Gly Gln Leu Ser Asn
995 1000 1005

Val Val Leu Leu Leu Pro His Gly His Glu Gly Gln Gly Pro Asp
1010 1015 1020

His Thr Ser Ala Arg Ile Glu Arg Phe Leu Gln Leu Trp Ala Glu
1025 1030 1035

Gly Ser Met Thr Ile Ala Met Pro Ser Thr Pro Ser Asn Tyr Phe
1040 1045 1050

His Leu Leu Arg Arg His Ala Leu Asp Gly Ile Gln Arg Pro Leu
1055 1060 1065

Ile Val Phe Thr Pro Lys Ser Met Leu Arg His Lys Ala Ala Val
1070 1075 1080

Ser Glu Ile Lys Asp Phe Thr Glu Ile Lys Phe Arg Ser Val Leu
1085 1090 1095

Glu Glu Pro Thr Tyr Glu Asp Gly Ile Gly Asp Arg Asn Lys Val
1100 1105 1110

Ser Arg Ile Leu Leu Thr Ser Gly Lys Leu Tyr Tyr Glu Leu Ala
1115 1120 1125

Ala Arg Lys Ala Lys Asp Asn Arg Asn Asp Leu Ala Ile Val Arg
1130 1135 1140

Leu Glu Gln Leu Ala Pro Leu Pro Arg Arg Arg Leu Arg Glu Thr
1145 1150 1155

Leu Asp Arg Tyr Glu Asn Val Lys Glu Phe Phe Trp Val Gln Glu
1160 1165 1170

Glu Pro Ala Asn Gln Gly Ala Trp Pro Arg Phe Gly Leu Glu Leu
1175 1180 1185

Pro Glu Leu Leu Pro Asp Lys Leu Ala Gly Ile Lys Arg Ile Ser
1190 1195 1200

Arg Arg Ala Met Ser Ala Pro Ser Ser Gly Ser Ser Lys Val His
 1205 1210 1215

Ala Val Glu Gln Gln Glu Ile Leu Asp Glu Ala Phe Gly
 1220 1225 1230

<210> 2
 <211> 3696
 <212> DNA
 <213> Mycobacterium bovis

<400> 2
 atggccaata tctcgtcccc gttcggccag aacgagtggc tcgtggaggc catgtatcgg 60
 aagttccggg acgaccgctc cagcgtcgac ccgtcctggc acgagttcct cgtggactac 120
 tccccgagc cgaccagcca gccggcggcc gagccgaccc gggtcacgtc gccgctcgtc 180
 gccgaacggg ccgccgccgc cgccccccag gcgcccccca agccggccga cacggccgcc 240
 gccggcaacg gcgtggtggc ggccctggcc gccaaagacc ccgtgccgcc cccggccgaa 300
 ggcgacgaag tcgccgtcct ccggggcgcg gcggcgggcg tcgtgaagaa catgagcgcg 360
 tccctggagg tgccgaccgc cacgtcggtc cgggccgtcc ccgcgaaact cctcatcgac 420
 aaccgcatcg tcatcaataa ccagctgaag cggacgcggg gcggcaagat ctcgttcacg 480
 cacctcctcg gctacgcgct cgtgcaggcc gtcaagaagt tcccgaacat gaaccggcac 540
 tacacggaag tggacggcaa acccacggcc gtcacccccg ccatacgaa cctcggcctg 600
 gccatcgacc tgcagggcaa ggacggcaag cggtcgctgg tggtcgccgg catcaagcgc 660
 tgcgagacga tgcgcttcgc gcagttcgtc acggcctatg aagacatcgt ccgccgggcc 720
 cgggacggca agctcaccac ggaagatttc gcgggcgtca ccatctcgct caccaatccc 780
 ggacgatcg gcaccgtcca tagcgtgccg cgcctgatgc ccggccaggg cgcgatcatc 840
 ggcgtcggcg cgatggaata ccccgccgag ttccagggcg cgtccgagga gcgcatcgcg 900
 gaactgggca tcggcaagct catcaccctc accagcacgt atgaccaccg catcatccag 960
 ggcgccgagt cgggcgactt cctccggacg atccatgagc tgctgctgag cgacggcttc 1020
 tgggatgagg tgttccggga actgtccatc ccgtatctgc cggtcgggtg gagcacggac 1080
 aaccccgata gcatcgtgga caaaaacgcc cgcgtgatga acctcatcgc cgcctaccgc 1140

aaccgcgggc	acctgatggc	cgacaccgat	ccgctgcggc	tcgacaaagc	ccgcttccgg	1200
agccatcccc	acctggaagt	cctgaccac	ggcctcacc	tgtgggatct	ggatcgggtc	1260
ttcaaggtcg	acggcttcgc	cggcgcccag	tacaagaagc	tgcgcgatgt	gctcggcctg	1320
ctgcgggacg	cctattgccg	ccacatcggc	gtggagtacg	cgacatcct	cgacccggaa	1380
cagaaagaat	ggctggaaca	gcgcgtcgaa	accaagcatg	tgaagcccac	ggtggcgcag	1440
cagaagtaca	tcctgtcgaa	gctgaatgcg	gccgaagcct	tcgaaacctt	cctgcagacc	1500
aagtacgtcg	gccagaagcg	cttctccctg	gaaggcgccg	agtcctcat	cccgatgatg	1560
gatgccgcga	tcgaccagt	cgccgaacac	ggcctggacg	aagtcgtgat	cgcatgccc	1620
catcgcgggc	ggctcaacgt	cctcgcgaac	atcgtcggca	agccgtactc	gcagatcttc	1680
acggagtctg	aaggcaacct	gaacccctcc	caggcccacg	gcagcggcga	cgtgaaatat	1740
cacctgggcg	ccaccggcct	gtatctgcag	atgttcggcg	acaacgacat	ccaggtcagc	1800
ctgaccgcca	acccctcgca	tctggaagcc	gtcgaccccg	tcctggaagg	cctggtccgc	1860
gcgaagcagg	acctcctgga	ccacggcagc	atcgatagcg	acggccagcg	ggccttctcc	1920
gtggtgcccc	tcattgctgca	cggcgacgcc	gccttcgccg	gccagggcgt	ggtcgccgag	1980
acgtcaacc	tggcgaacct	gcccggctac	cgcgctggcg	gcacgatcca	catcatcgtc	2040
aacaatcaga	tcggcttcac	caccgcccc	gagtactccc	gcagctcgga	atactgcacc	2100
gatgtggcga	aatgatcgg	cgccccgatc	ttccatgtca	atggcgatga	tccggaagcg	2160
tgcgtgtggg	tggcccgct	ggcgggtgat	ttccgccagc	ggttcaagaa	ggatgtggtg	2220
atcgacatgc	tgtgctaccg	ccgccggggc	cataatgaag	gcgacgatcc	ctccatgacc	2280
aaccctaca	tgtacgatgt	cgtggacacg	aagcgcgggc	ccgcaagtc	ctataccgag	2340
gccctcatcg	gccggggcga	tatcagcatg	aaagaagccg	aggatgcgct	ccgcgactat	2400
cagggccagc	tggaacgggt	gttcaacgag	gtgcgggaac	tcgaaaaaca	cggcgtccag	2460
ccgtccgaga	gcgtggagtc	cgatcagatg	atcccggcgg	gcctcgccac	ggcggtcgat	2520
aagtccctgc	tggcccggat	cggcgatgcc	ttcctggccc	tcccaatgg	cttcacggcg	2580
caccgcggg	tgcagccggt	gctcgaaaaa	cgccgggaaa	tggcctacga	aggcaagatc	2640
gactgggcct	tcggcgagct	gctggcgctg	ggctccctgg	tggccgaagg	caaactggtg	2700
cgctgtccg	gccaggactc	gcgcccgggc	accttctccc	agcgccactc	ggtcctgatc	2760

gatcgccata ccggcgagga gttcaccccg ctccagctcc tggccacgaa ctccgacggc	2820
tcgcccaccg gcggcaagtt cctggtgtat gacagccccc tgagcgagta tgcggcggtc	2880
ggcttcgagt acggctacac cgtcggcaat cccgacgccg tcgtcctgtg ggaagcgcag	2940
ttcggcgact tcgtcaatgg cgcccagtcg atcatcgacg agttcatctc gtcgggcgag	3000
gccaaatggg gccagctctc gaacgtcgtg ctgctgctcc cgcacggcca cgaaggccag	3060
ggccccgcatc acacctccgc ccggatcgaa cggttcctgc agctctgggc cgagggtctc	3120
atgacgatcg ccatgcccag caccctgtcc aactatttcc acctcctgcg gcgccatgcg	3180
ctcgacggca tccagcggcc cctgatcgtc ttcacgccga agagcatgct gcgccataag	3240
gccgccgtct cggagatcaa ggacttcacc gagatcaaat tccggtcggt gctggaagaa	3300
ccgacctatg aagatggcat cggcgaccgc aataaggtgt cgcgcacacct gctgacctcg	3360
ggcaagctgt actacgagct ggccgcccgc aaggccaaag acaaccggaa cgacctggcc	3420
atcgtccggc tggagcagct ggcgcccctc ccgcgccggc gcctgcgcga aaccttgat	3480
cgctatgaga acgtcaagga gttcttctgg gtgcaggagg aaccgcgaa tcagggcgcc	3540
tggccccgct tcggcctgga gctgccggag ctgctccccg acaagctggc cggcatcaag	3600
cggatcagcc gccgcgcgat gtcggccccg agctcgggct ccagcaaggt ccatgcggtc	3660
gagcagcagg agatcctgga tgaagcgttc ggctga	3696

<210> 3
 <211> 1231
 <212> PRT
 <213> Mycobacterium tuberculosis

<400> 3

Met	Ala	Asn	Ile	Ser	Ser	Pro	Phe	Gly	Gln	Asn	Glu	Trp	Leu	Val	Glu
1				5				10					15		

Glu	Met	Tyr	Arg	Lys	Phe	Arg	Asp	Asp	Pro	Ser	Ser	Val	Asp	Pro	Ser
		20					25					30			

Trp	His	Glu	Phe	Leu	Val	Asp	Tyr	Ser	Pro	Glu	Pro	Thr	Ser	Gln	Pro
		35				40						45			

Ala	Ala	Glu	Pro	Thr	Arg	Val	Thr	Ser	Pro	Leu	Val	Ala	Glu	Arg	Ala
	50					55					60				

Ala Ala Ala Ala Pro Gln Ala Pro Pro Lys Pro Ala Asp Thr Ala Ala
65 70 75 80

Ala Gly Asn Gly Val Val Ala Ala Leu Ala Ala Lys Thr Ala Val Pro
85 90 95

Pro Pro Ala Glu Gly Asp Glu Val Ala Val Leu Arg Gly Ala Ala Ala
100 105 110

Ala Val Val Lys Asn Met Ser Ala Ser Leu Glu Val Pro Thr Ala Thr
115 120 125

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

Ile Asn Asn Gln Leu Lys Arg Thr Arg Gly Gly Lys Ile Ser Phe Thr
145 150 155 160

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

Met Asn Arg His Tyr Thr Glu Val Asp Gly Lys Pro Thr Ala Val Thr
180 185 190

Pro Ala His Thr Asn Leu Gly Leu Ala Ile Asp Leu Gln Gly Lys Asp
195 200 205

Gly Lys Arg Ser Leu Val Val Ala Gly Ile Lys Arg Cys Glu Thr Met
210 215 220

Arg Phe Ala Gln Phe Val Thr Ala Tyr Glu Asp Ile Val Arg Arg Ala
225 230 235 240

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

Leu Thr Asn Pro Gly Thr Ile Gly Thr Val His Ser Val Pro Arg Leu
260 265 270

Met Pro Gly Gln Gly Ala Ile Ile Gly Val Gly Ala Met Glu Tyr Pro
275 280 285

Ala Glu Phe Gln Gly Ala Ser Glu Glu Arg Ile Ala Glu Leu Gly Ile
290 295 300

Gly Lys Leu Ile Thr Leu Thr Ser Thr Tyr Asp His Arg Ile Ile Gln
305 310 315 320

Gly Ala Glu Ser Gly Asp Phe Leu Arg Thr Ile His Glu Leu Leu Leu
325 330 335

Ser Asp Gly Phe Trp Asp Glu Val Phe Arg Glu Leu Ser Ile Pro Tyr
340 345 350

Leu Pro Val Arg Trp Ser Thr Asp Asn Pro Asp Ser Ile Val Asp Lys
355 360 365

Asn Ala Arg Val Met Asn Leu Ile Ala Ala Tyr Arg Asn Arg Gly His
370 375 380

Leu Met Ala Asp Thr Asp Pro Leu Arg Leu Asp Lys Ala Arg Phe Arg
385 390 395 400

Ser His Pro Asp Leu Glu Val Leu Thr His Gly Leu Thr Leu Trp Asp
405 410 415

Leu Asp Arg Val Phe Lys Val Asp Gly Phe Ala Gly Ala Gln Tyr Lys
420 425 430

Lys Leu Arg Asp Val Leu Gly Leu Leu Arg Asp Ala Tyr Cys Arg His
435 440 445

Ile Gly Val Glu Tyr Ala His Ile Leu Asp Pro Glu Gln Lys Glu Trp
450 455 460

Leu Glu Gln Arg Val Glu Thr Lys His Val Lys Pro Thr Val Ala Gln
465 470 475 480

Gln Lys Tyr Ile Leu Ser Lys Leu Asn Ala Ala Glu Ala Phe Glu Thr
485 490 495

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

Ala Glu Ser Val Ile Pro Met Met Asp Ala Ala Ile Asp Gln Cys Ala
515 520 525

Glu His Gly Leu Asp Glu Val Val Ile Gly Met Pro His Arg Gly Arg
530 535 540

Leu Asn Val Leu Ala Asn Ile Val Gly Lys Pro Tyr Ser Gln Ile Phe
545 550 555 560

Thr Glu Phe Glu Gly Asn Leu Asn Pro Ser Gln Ala His Gly Ser Gly
565 570 575

Asp Val Lys Tyr His Leu Gly Ala Thr Gly Leu Tyr Leu Gln Met Phe
580 585 590

Gly Asp Asn Asp Ile Gln Val Ser Leu Thr Ala Asn Pro Ser His Leu
595 600 605

Glu Ala Val Asp Pro Val Leu Glu Gly Leu Val Arg Ala Lys Gln Asp
610 615 620

Leu Leu Asp His Gly Ser Ile Asp Ser Asp Gly Gln Arg Ala Phe Ser
625 630 635 640

Val Val Pro Leu Met Leu His Gly Asp Ala Ala Phe Ala Gly Gln Gly
645 650 655

Val Val Ala Glu Thr Leu Asn Leu Ala Asn Leu Pro Gly Tyr Arg Val
660 665 670

Gly Gly Thr Ile His Ile Ile Val Asn Asn Gln Ile Gly Phe Thr Thr
675 680 685

Ala Pro Glu Tyr Ser Arg Ser Ser Glu Tyr Cys Thr Asp Val Ala Lys
690 695 700

Met Ile Gly Ala Pro Ile Phe His Val Asn Gly Asp Asp Pro Glu Ala
705 710 715 720

Cys Val Trp Val Ala Arg Leu Ala Val Asp Phe Arg Gln Arg Phe Lys
725 730 735

Lys Asp Val Val Ile Asp Met Leu Cys Tyr Arg Arg Arg Gly His Asn
740 745 750

Glu Gly Asp Asp Pro Ser Met Thr Asn Pro Tyr Val Tyr Asp Val Val
755 760 765

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

Arg Gly Asp Ile Ser Met Lys Glu Ala Glu Asp Ala Leu Arg Asp Tyr
785 790 795 800

Gln Gly Gln Leu Glu Arg Val Phe Asn Glu Val Arg Glu Leu Glu Lys
805 810 815

His Gly Val Gln Pro Ser Glu Ser Val Glu Ser Asp Gln Met Ile Pro
820 825 830

Ala Gly Leu Ala Thr Ala Val Asp Lys Ser Leu Leu Ala Arg Ile Gly
835 840 845

Asp Ala Phe Leu Ala Leu Pro Asn Gly Phe Thr Ala His Pro Arg Val
850 855 860

Gln Pro Val Leu Glu Lys Arg Arg Glu Met Ala Tyr Glu Gly Lys Ile
865 870 875 880

Asp Trp Ala Phe Gly Glu Leu Leu Ala Leu Gly Ser Leu Val Ala Glu
885 890 895

Gly Lys Leu Val Arg Leu Ser Gly Gln Asp Ser Arg Arg Gly Thr Phe
900 905 910

Ser Gln Arg His Ser Val Leu Ile Asp Arg His Thr Gly Glu Glu Phe
915 920 925

Thr Pro Leu Gln Leu Leu Ala Thr Asn Ser Asp Gly Ser Pro Thr Gly
930 935 940

Gly Lys Phe Leu Val Tyr Asp Ser Pro Leu Ser Glu Tyr Ala Ala Val
945 950 955 960

Gly Phe Glu Tyr Gly Tyr Thr Val Gly Asn Pro Asp Ala Val Val Leu
965 970 975

Trp Glu Ala Gln Phe Gly Asp Phe Val Asn Gly Ala Gln Ser Ile Ile
980 985 990

Asp Glu Phe Ile Ser Ser Gly Glu Ala Lys Trp Gly Gln Leu Ser Asn
995 1000 1005

Val Val Leu Leu Leu Pro His Gly His Glu Gly Gln Gly Pro Asp
1010 1015 1020

His Thr Ser Ala Arg Ile Glu Arg Phe Leu Gln Leu Trp Ala Glu
1025 1030 1035

Gly Ser Met Thr Ile Ala Met Pro Ser Thr Pro Ser Asn Tyr Phe
1040 1045 1050

His Leu Leu Arg Arg His Ala Leu Asp Gly Ile Gln Arg Pro Leu
1055 1060 1065

Ile Val Phe Thr Pro Lys Ser Met Leu Arg His Lys Ala Ala Val
1070 1075 1080

Ser Glu Ile Lys Asp Phe Thr Glu Ile Lys Phe Arg Ser Val Leu
1085 1090 1095

Glu Glu Pro Thr Tyr Glu Asp Gly Ile Gly Asp Arg Asn Lys Val
1100 1105 1110

Ser Arg Ile Leu Leu Thr Ser Gly Lys Leu Tyr Tyr Glu Leu Ala
1115 1120 1125

Ala Arg Lys Ala Lys Asp Asn Arg Asn Asp Leu Ala Ile Val Arg
 1130 1135 1140

Leu Glu Gln Leu Ala Pro Leu Pro Arg Arg Arg Leu Arg Glu Thr
 1145 1150 1155

Leu Asp Arg Tyr Glu Asn Val Lys Glu Phe Phe Trp Val Gln Glu
 1160 1165 1170

Glu Pro Ala Asn Gln Gly Ala Trp Pro Arg Phe Gly Leu Glu Leu
 1175 1180 1185

Pro Glu Leu Leu Pro Asp Lys Leu Ala Gly Ile Lys Arg Ile Ser
 1190 1195 1200

Arg Arg Ala Met Ser Ala Pro Ser Ser Gly Ser Ser Lys Val His
 1205 1210 1215

Ala Val Glu Gln Gln Glu Ile Leu Asp Glu Ala Phe Gly
 1220 1225 1230

<210> 4
 <211> 3696
 <212> DNA
 <213> Mycobacterium tuberculosis

<400> 4
 atggccaaca tcagcagccc cttcggccag aatgaatggc tcgtggagga gatgtaccgg 60
 aagttccggg atgatcccag ctcggtcgat ccgtcgtggc acgagttcct ggtcgactat 120
 agccccgaac ccacctcca gccggcggcc gagcccacgc gggtgacctc cccgctggtc 180
 gcggaacgcg ccgcggcggc ggccccccag gccccgccga aaccgccga tacggccgcc 240
 gcgggcaacg gcgtcgtcgc ggccctggcc gcgaagaccg ccgtgcccc cccgccgag 300
 ggcgacgagg tggcggtgct ccggggcgcc gccgcggccg tggtaagaa catgagcgcc 360
 tccctcgaag tgccgaccgc caccagcgtc cgggcggtgc cggccaagct gctgatcgat 420
 aatcggatcg tcatcaaaa tcagctgaaa cggacccgcg gcggcaagat cagcttcacc 480
 catctgctcg gctatgcgt ggtccaggcc gtgaagaagt tccccaat gaatgccac 540
 tataccgaag tcgacggcaa acccaccgcc gtgaccccg cgcacacgaa cctcggcctg 600

gcgatcgacc tgcagggcaa ggacggcaag cggagcctcg tggtaggccc	catcaaacgg	660
tgcgagacga tgcggttcgc gcagttcgtg acggcctatg aggatatcgt	gcgccgcgcc	720
cgggacggca aactgaccac cgaagacttc gccggcgtca cgatctcgct	gaccaaccgg	780
ggcaccatcg gcacgggtcca ttcgggtccc cgcctcatgc cgggccaggg	cgccatcatc	840
ggcgtgggcg cgatggaata tcccgccgag ttccaggggcg cctccgagga	gcggatcgcc	900
gagctgggca tcggcaaact catcacgtg acctccacgt atgaccaccg	catcatccag	960
ggcgcggaaa gcggcgactt cctccggacg atccatgagc tgctgctctc	cgacggcttc	1020
tgggatgaag tgttccgcga gctgtcgatc ccctatctcc cggtagcgctg	gagcaccgac	1080
aacccggact cgatcgtgga caaaaacgcc cgcgtcatga atctgatcgc	cgcgtaccgc	1140
aatcggggcc acctcatggc ggacacggac cccctgcggc tcgataaggc	ccggttcgcg	1200
tcccaccccg acctcgaagt cctgaccac ggctgacct tgtgggacct	ggatcgcgtc	1260
ttcaaggtcg atggcttcgc cggcgcccag tataagaagc tgcgcgatgt	gctgggcctg	1320
ctccgcgatg cctattgccg gcatacggc gtggagtacg cccacatcct	ggaccccgag	1380
cagaaagaat ggctggaaca gcgcgtcgaa acgaagcacg tcaagccac	cgtagcgag	1440
cagaaataca tcctcagcaa actcaacgcc gccgaagcgt tcgaaacgtt	cctccagacg	1500
aagtacgtgg gccagaaacg cttcagcctg gaaggcgcgg agagcgtcat	ccccatgatg	1560
gacgccgcca tcgaccagtgc gccgaacac ggctgggatg aagtggatcat	cggcatgccg	1620
caccgaggcc ggctcaacgt cctggccaat atcgtgggca agccctacag	ccagatcttc	1680
accgagttcg aaggcaacct gaaccgctg caggcgcacg gctcgggcga	tgtcaagtac	1740
catctcggcg ccacgggcct gtatctgcag atgttcggcg acaacgacat	ccaggtgtcc	1800
ctgacggcga accccagcca cctggaggcg gtcgatcccg tcctggaagg	cctggtgcgg	1860
gccaagcagg atctcctcga ccacggctcg atcgactccg atggccagcg	ggcgttcagc	1920
gtggtgcccc tcatgctgca tggcgacgcg gcgttcgcgg gccagggcgt	ggtggcggaa	1980
accctcaacc tcgcgaacct cccgggctat cgggtgggcg gcaccatcca	tatcatcgtg	2040
aacaaccaga tcggcttcac cacggcgccg gaatattccc ggagctcgga	atattgcacg	2100
gacgtggcca agatgatcgg cgcgccgatc ttccacgtca atggcgatga	cccggaggcc	2160
tgtgtgtggg tcgcccggct ggccgtcgat ttccgccagc gcttcaaaaa	agatgtggtc	2220

atcgacatgc tctgctaccg ccgccggggc cataatgagg gcgacgaccc ctccatgacg	2280
aacccgtacg tctacgacgt ggtcgacacc aagcgcgggc cccgcaagtc ctatacggag	2340
gcgctcatcg gccgcggcga catctcgatg aaggaagcgg aagacgcgct ccgcgactac	2400
cagggccagc tggagcgcgt cttcaacgaa gtgcgggaac tcgaaaagca tggcgtgcag	2460
ccctccgaat cgggtggagag cgatcagatg atcccggccg gcctggccac cgccgtcgat	2520
aaaagcctgc tcgcccgcgt cggcgacgcc ttcttgggcc tgccgaacgg cttcacggcc	2580
catccccgcg tccagccggt gctggagaaa cgccgcgaaa tggcctacga gggcaagatc	2640
gattgggcct tcggcgagct gctggcgctc ggcagcctgg tggccgaagg caagctcgtg	2700
cggctctccg gccaggactc gcggcgcggc acgttctcgc agcgccattc cgtgctgac	2760
gaccggcaca ccggcgaaga gttcaccccc ctccagctgc tggccaccaa ttcggacggc	2820
agcccgacgg gcggcaagtt cctggtctat gactccccgc tgagcgagta tgccgccgtc	2880
ggcttcgagt atggctacac cgtcggcaat ccggacgcgg tcgtgctgtg ggaagcgcag	2940
ttcggcgact tcgtgaacgg cgcccagtcg atcatcgatg agttcatctc gtcgggcgaa	3000
gccaagtggg gccagctcag caacgtggtg ctctgctcc cgacaggcca cgagggccag	3060
ggcccgatc atacgagcgc ccgcatcgag cggttcctgc agctgtgggc cgagggctcc	3120
atgacgatcg cgatgccctc gacgccgtcc aactacttcc acctgctgcg ccggcacgcg	3180
ctggacggca tccagcgccc gctgatcgtc ttacccccca agagcatgct ccggcacaaa	3240
gccgccgtgt ccgagatcaa ggacttcacg gagatcaaat tccggagcgt gctggaggag	3300
cccacgtacg aggacggcat cggcgaccgc aacaaggtgt ccgcatacct cctgacgagc	3360
ggcaagctgt actacgagct cgccgcccgc aaggccaagg ataaccggaa cgacctggcc	3420
atcgtgcgcc tcgaacagct ggccccgctg ccgcgccgcc ggctgcggga aacctggac	3480
cgctacgaga acgtcaaaga gttcttctgg gtccaggaag aaccggccaa ccagggcgcg	3540
tggccgcgct tcggcctgga actgcccag ctcctgccgg acaagctggc cggcatcaag	3600
cgcatctcgc ggccggcgat gagcgccccg tcgagcggct cctccaaggt gcatgccgtg	3660
gagcagcagg aaatcctgga cgaggcgttc ggctga	3696

<210> 5
 <211> 1245
 <212> PRT

<213> Corynebacterium terpenotabidum

<400> 5

Met Asn Ser Thr Asn Phe Gly Ser Asn Gly Trp Leu Val Asp Gln Met
1 5 10 15

His Gln Gln Phe Lys Glu Asp Pro Gln Ser Val Asp Lys Glu Trp Arg
20 25 30

Asp Phe Phe Thr Ala Gly Ser Ala Ser Gly Pro Asp Ala Pro Glu Thr
35 40 45

Thr Ala Pro Thr Thr Thr Ala Ala Thr Thr Thr Thr Ala Thr Thr Ser
50 55 60

Ala Ala Pro Ala Thr Ala Gln Thr Ser Gly Val Ala Val Pro Ser Thr
65 70 75 80

Ala Gly Thr Thr His Arg Ala Ser Ala Ala Ala Thr Val Ala Ser Asn
85 90 95

Val Ala Leu Ser Ala Asn Pro Ile Pro Glu Arg Val Ala Pro Pro Ala
100 105 110

Pro Lys Asp Ala Ala Glu Pro Leu Thr Pro Gly Glu Ala Ala Leu Lys
115 120 125

Gly Ala Gln Arg Ala Ile Ala Lys Asn Met Asp Ala Ser Leu Asp Ile
130 135 140

Pro Thr Ala Thr Thr Val Arg Asp Met Pro Ala Lys Leu Met Phe Glu
145 150 155 160

Asn Arg Ala Met Ile Asn Asn His Leu Arg Ser Gln Gly Arg Gly Lys
165 170 175

Ile Ser Phe Thr His Ile Leu Gly Trp Ala Met Val Asn Ala Val Lys
180 185 190

Ala His Pro Thr Met Asn Asn Asn Tyr Lys Val Ile Asp Gly Lys Pro
195 200 205

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

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

Cys Glu Thr Phe Asp Phe Glu Gly Phe Val Asp Ala Tyr Glu Asp Ile
245 250 255

Val Val Arg Ala Arg Lys Gly Lys Leu Thr Met Asp Asp Phe Ser Gly
260 265 270

Val Thr Ile Gln Leu Thr Asn Pro Gly Gly Ile Gly Thr Arg His Ser
275 280 285

Val Pro Arg Leu Thr His Gly Gln Gly Ala Ile Ile Gly Val Gly Ala
290 295 300

Met Asp Tyr Pro Ala Glu Phe Ala Gly Ala Ser Glu Asp Arg Leu Ala
305 310 315 320

Asp Leu Gly Val Gly Lys Leu Val Thr Ile Thr Ser Thr Tyr Asp His
325 330 335

Arg Ile Ile Gln Gly Ala Glu Ser Gly Glu Phe Leu Arg Asp Met Ser
340 345 350

Arg Gln Leu Ile Asp Asp Gly Phe Trp Asp Gly Ile Tyr Ala Ser Leu
355 360 365

Lys Val Pro Tyr Ala Pro Val Arg Trp Ser Gln Asp Val Pro Asn Thr
370 375 380

Gly Val Asp Lys Ser Thr Arg Val Met Gln Leu Ile Glu Ala Tyr Arg
385 390 395 400

Ser Arg Gly His Leu Ile Ala Asp Ile Asp Pro Leu His Trp Thr Gln
405 410 415

Pro Gly Leu Pro Val Pro Asp His Ser Asp Leu Asp Ile Glu Ser His
420 425 430

Gly Leu Thr Leu Trp Asp Phe Asp Arg Arg Phe His Val Gly Gly Phe
435 440 445

Ala Gly Arg Glu Ser Met Thr Leu Arg Glu Val Leu Ala Thr Leu Arg
450 455 460

Lys Ala Tyr Thr Leu Lys Val Gly Ser Glu Tyr Thr His Ile Leu Asp
465 470 475 480

Lys Asp Glu Arg Leu Trp Leu Gln Glu His Ile Glu Ala Gly Gln Gln
485 490 495

Lys Leu Ser Asn Pro Glu Gln Lys Tyr Leu Leu Gln Thr Leu Asn Ser
500 505 510

Ala Glu Ala Phe Glu Asn Phe Leu Gln Thr Lys Tyr Ile Gly Gln Lys
515 520 525

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

Ala Ala Asp Gln Ala Ala Glu Gln Gly Leu Glu Glu Val Val Ile Gly
545 550 555 560

Met Pro His Arg Gly Arg Leu Asn Val Leu Ala Asn Ile Val Gly Lys
565 570 575

Pro Tyr Ser Thr Ile Phe Gly Glu Phe Glu Gly Asn Ile Glu Pro Ala
580 585 590

Ala Ala Gly Gly Ser Gly Asp Val Lys Tyr His Leu Gly Ala Glu Gly
595 600 605

Val Tyr Thr Gln Met Phe Gly Asp Asn Asp Ile Lys Val Thr Leu Thr
610 615 620

Ala Asn Pro Ser His Leu Glu Ala Val Asn Pro Val Met Glu Gly Leu
625 630 635 640

Ala Arg Ala His Gln Asp Ile Ser Pro Arg Ala Glu Asp Arg Pro Ile
645 650 655

Met Pro Ile Leu Met His Gly Asp Ala Ala Phe Thr Gly Leu Gly Ile
660 665 670

Val Pro Glu Thr Ile Asn Met Ala Gln Leu Arg Gly Tyr Ser Val Gly
675 680 685

Gly Thr Val His Val Val Val Asn Asn Gln Ile Gly Phe Thr Thr Thr
690 695 700

Pro Asp Arg Gly Arg Ser Thr His Tyr Ala Thr Asp Ile Ala Lys Gly
705 710 715 720

Phe Asp Cys Pro Val Phe His Val Asn Gly Asp Asp Pro Glu Ala Val
725 730 735

Val Trp Val Ala Arg Leu Ala Val Glu Tyr Arg Arg Arg Phe Gly Lys
740 745 750

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

Ala Asp Asp Pro Ser Met Thr Gln Pro Glu Leu Tyr Ser Ile Ile Glu
770 775 780

Ser Arg Pro Thr Val Arg Ser Leu Tyr His Asp Thr Leu Val Gly Arg
785 790 795 800

Gly Asp Ile Thr Ala Glu Asp Ala Gln Arg Ala Ala Asp Asp Phe His
805 810 815

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

Val Thr Pro Ala Ala Gln Thr Gly Ile Ala Gly Gly Gln Asp Leu Ser
835 840 845

Thr Gly Leu Asp Thr Ser Ile Thr Ala Asp Val Ile Ala Glu Ile Gly
850 855 860

Asp Ser Tyr Thr Val Asp Ala Pro Glu Asp Phe Asn Val His Gln Arg
865 870 875 880

Val Lys Pro Val Val Lys Arg Arg His Gln Met Ser Arg Gln Gly Lys
885 890 895

Ile Asp Trp Ala Phe Gly Glu Leu Leu Ala Phe Gly Ser Leu Ala Arg
900 905 910

Glu Gly Arg Leu Val Arg Leu Ala Gly Glu Asp Ser Gln Arg Gly Thr
915 920 925

Phe Thr Gln Arg His Ala Ile Leu Phe Asp Ser Thr Thr Asn Lys Pro
930 935 940

Phe Ser Pro Leu Glu Met Val Ala Arg Asn Ser Gly Asn Gly Gly Ser
945 950 955 960

Phe Arg Ala Phe Asn Ser Pro Leu Thr Glu Tyr Ala Gly Met Gly Phe
965 970 975

Glu Tyr Gly Tyr Ser Val Gly Asn Leu Asp Ala Val Val Ala Trp Glu
980 985 990

Ala Gln Phe Gly Asp Phe Ala Asp Gly Ala Gln Thr Ile Ile Asp Glu
995 1000 1005

Tyr Ile Ser Ser Gly Glu Ala Lys Trp Gly Gln Leu Ser Ser Val
1010 1015 1020

Ile Leu Leu Leu Pro His Gly Tyr Glu Gly Gln Gly Pro Asp His
1025 1030 1035

Ser Ser Ala Arg Ile Glu Arg Tyr Leu Gln Met Ala Ala Glu Gly
1040 1045 1050

Ser Met Thr Ile Ala Gln Pro Ser Thr Pro Ala Asn His Phe His
1055 1060 1065

Leu Leu Arg Arg His Ala Leu Gly Thr Met Arg Arg Pro Leu Val
1070 1075 1080

Val Phe Thr Pro Lys Ser Met Leu Arg Asn Lys Ala Ala Val Ser
1085 1090 1095

Ser Val Glu Asp Phe Thr Glu Val Thr Lys Phe Arg Ser Val Leu
1100 1105 1110

Asp Asp Pro Arg Phe Ala Asp Gly Thr Ala Asp Arg Ser Gly Val
1115 1120 1125

Lys Thr Val Leu Met Cys Ser Gly Lys Ile Tyr Tyr Asp Leu Glu
1130 1135 1140

Lys Lys Arg Ala Glu Asp Gly Arg Asp Asp Ile Ala Ile Val Arg
1145 1150 1155

Val Glu Met Leu His Pro Ile Pro His Asn Arg Ile Arg Glu Thr
1160 1165 1170

Val Val Glu Gly Tyr Pro Gly Ala Glu Val Arg Trp Val Gln Asp
1175 1180 1185

Glu Pro Ala Asn Gln Gly Ala Trp Pro Phe Leu Ala Leu Asn Leu
1190 1195 1200

Pro Glu Arg Ile Pro Gly Phe Thr Met Lys Arg Val Ser Arg Arg
1205 1210 1215

Ala Gln Ser Ser Thr Ala Thr Gly Val Ala Lys Val His His Leu
1220 1225 1230

Glu Gln Glu Thr Leu Leu Thr Glu Ala Phe Ala Thr
1235 1240 1245

<210> 6

<211> 3738

<212> DNA

<213> Corynebacterium terpenotabidum

<400> 6

atgaactcga ccaacttcgg cagcaatggc tggctggtcg accagatgca ccagcagttc	60
aaggaagatc cgcagtcggt cgacaaggaa tggcgcgact tcttcaccgc cggctcggcg	120
tccggcccgg atgccccgga aacgaccgcc ccgacgacca ccgccgccac gaccacgacc	180
gcgaccacga gcgcggcccc cgccaccgcg cagaccagcg gcgtggccgt cccgtccacc	240
gcgggcacca cccaccgggc ctcggccgcc gcgaccgtgg cctcgaacgt ggccctgagc	300
gcgaaccga tccccgaacg ggtcgcccc ccggccccga aggacgcggc cgagccgctg	360
accccgggcg aggccgccct gaaaggcgcc cagcgggcga tcgccaaaaa catggacgcc	420
agcctggata tccccaccgc gacgacggtc cgcgacatgc cggccaaact catgttcgaa	480
aaccgggcca tgatcaacaa ccatctgcgc agccagggcc ggggcaagat ctcgttcacc	540
cacatcctgg gctgggcgat ggtcaacgcc gtgaaagccc accccaccat gaataacaat	600
tacaaggtca tcgacggcaa gccgtcggtc gtgacccccg aacatatcaa tctcggcctc	660
gccatcgatc tggtcagcaa aaacggctcg cggaacctgg tcgtggccgc catccgcgcg	720
tgcgaaacct tcgacttcga gggcttcgtg gacgcctacg aggacatcgt ggtccgcgcc	780
cggaagggca aactgacgat ggacgacttc agcggcgcta ccatccagct gaccaaccgc	840
ggcggcatcg gcacccgcca ctccgtcccg cgcctgacgc acggccaggg cgcgatcatc	900
ggcgtgggcg cgatggacta cccggcggag ttccgccggcg cctcggaaga ccgcctcgcg	960
gacctgggcg tgggcaaact ggtgaccatc acgagcacgt atgaccaccg gatcatccag	1020
ggcgccgagt cgggcgagtt cctgcgcgac atgagccgcc agctcatcga cgatggcttc	1080
tgggacggca tctacgccag cctgaaagtc ccgtatgcgc ccgtgcggtg gtcgcaggac	1140
gtgccaaca ccggcgtcga caagagcacc cgcgtcatgc agctcatcga ggcgtaccgc	1200
tcccgcggcc acctgatcgc cgacatcgac ccgctgcatt ggaccagcc cggcctgccg	1260
gtgccggacc atagcgacct ggacatcgaa tcccacggcc tgaccctgtg ggacttcgac	1320
cggcgcttcc acgtcggcgg cttcgccggc cgcgaatcga tgaccctgcg ggaagtgctg	1380
gccacgctgc ggaaggcgta tacgctgaag gtcggctccg aatataccca tatcctcgac	1440
aaggacgaac gcctctggct ccaggagcac atcgaggccg gccagcagaa gctcagcaac	1500
cccgaacaga agtatctgct gcagacgtg aactccgccg aggcgttcga aaacttcctg	1560

cagaccaa	at	atc	cg	gcca	ga	agc	gc	ctt	c	agc	ct	gga	ag	gc	gcc	ga	agc	c	ct	gat	ccc	g	1620								
ctgctc	gat	g	cc	gc	gc	gga	tc	agg	cc	gc	ga	cagg	gg	cc	tc	ga	ag	agg	t	ggt	gat	c	ggc	1680							
atgccg	cat	c	gc	gg	cc	gg	ct	ga	tg	tc	ct	gc	ga	ac	at	cg	tc	gg	caa	acc	gt	ac	ag	cacc	1740						
atcttc	gg	gc	g	ag	gg	g	ca	ac	at	cg	ag	cc	gg	cc	gc	cg	cg	gg	gc	gg	ct	c	gg	cg	at	gtc	1800				
aaatac	cat	c	tg	gg	gc	gc	cg	gg	gc	gt	gt	at	ac	gc	ag	at	gt	tc	gg	gc	ga	caa	cg	ac	at	ca	ag	1860			
gtgacg	ct	ga	cc	ca	ac	ccc	ct	ccc	ac	ct	g	ga	ag	cg	gt	ca	at	ccc	gc	t	cat	gg	ag	gg	gc	ct	c	1920			
gcgcgg	gg	gc	g	at	cagg	ata	ct	c	gc	cccc	gg	gc	gg	agg	at	c	gg	cc	gat	cat	g	cc	gat	c	ct	g		1980			
atgcac	gg	gc	g	ac	gc	gc	ctt	c	acc	gg	c	ct	g	gc	at	c	gt	cc	cg	ga	a	c	gat	gg	cc			2040			
cagctg	cg	gg	g	g	ct	act	cc	gt	gg	gc	gg	c	ac	g	gt	cat	gt	gg	t	ca	a	caa	cc	ag	at	c	gg	c	2100		
ttcacc	ac	ga	gc	cc	gg	at	cg	gg	cc	gg	t	cc	ac	gc	act	ac	g	gc	ag	g	ac	at	cg	cg	ga	ag	gg	gc	2160		
ttcgat	t	gcc	cc	gt	ctt	cca	cg	t	ca	ac	gg	c	gc	at	cc	gg		ag	g	cc	gc	t	cg	t	ct	gg	gt	cg	cc	2220	
cggctc	gc	cg	g	tg	ga	ata	cc	g	gc	gc	gc	ttc	gg	caa	ag	ac	g	tg	tt	cat	c	ga	c	ct	g	gt	gt	gt		2280	
tatcgg	gc	g	gc	gg	cc	ata	tg	a	ag	cc	gat	ga	cccc	c	ag	ca	tg	ac	cc	ag	cc	cg	a	a	act	gt	ac			2340	
agcatc	at	cg	ag	cc	gg	cc	ga	cg	gt	t	cc	gg	ag	c	ct	gt	at	at	ga	ca	cg	ct	gg	t	cg	gc	cg			2400	
ggcgac	at	ca	cg	gc	gga	aga	cg	ccc	c	ag	c	gg	gc	cc	gc	ac	g	act	t	cc	ac	gg	cc	ag	ct	gg	aa			2460	
agcgtc	tt	ca	acc	agg	t	caa	gga	agg	gc	gt	g	aagg	gg	c	gt	ga	cccc	gg	c	gg	gc	cc	ag	acc	gg	c				2520	
atcgcc	gg	gc	g	gcc	agg	ac	ct	gc	ac	gg	c	gc	at	ca	cca	gc	at	c	ac	cc	gc	gg	at	gt	gat	c				2580	
gccgag	at	cg	g	at	t	ct	ta	t	acc	gt	gg	at	g	cccc	c	ga	ag	act	t	ca	at	gt	cc	acc	ag	cg	c			2640	
gtgaag	ccc	g	tg	gt	caa	ac	g	cg	gc	ac	cag	at	gt	c	gc	g	agg	g	ca	ag	at	cg	att	gg	gg	cg				2700	
ttcggc	ga	ac	tc	ct	gc	gtt	c	gg	ct	c	gt	g	cccc	gg	ga	ag	g	cc	gg	c	t	gg	t	cc	gg	ct	c	gc			2760
ggcgag	g	act	gc	gg	gg	g	c	ac	ctt	c	acc	c	ag	cg	gg	c	ac	g	cc	at	c	ct	gtt	cg	ac	ag	c	acc			2820
accaac	a	agc	c	ttt	c	ag	ccc	c	ct	g	ga	aat	g	gt	c	gc	gc	gga	act	cc	gg	caa	cg	gg	cg	ct	c	g			2880
ttccgg	gc	ct	t	ca	att	cccc	g	ct	g	ac	gg																				

Thr Glu Ala Pro Lys Lys Ala Ala Pro Ala Lys Glu Thr Ala Ala Thr
115 120 125

Asp Ala Lys Ala Ser Ala Pro Ala Pro Ala Val Glu Glu Ser Lys Val
130 135 140

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

Ala Ile Pro Thr Ala Thr Ser Val Arg Ala Ile Pro Ala Lys Leu Met
165 170 175

Phe Asp Asn Arg Ile Val Ile Asn Asn His Leu Ala Arg Thr Arg Gly
180 185 190

Gly Lys Ile Ser Phe Thr His Leu Leu Gly Tyr Ala Ile Val Gln Ala
195 200 205

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

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

Asp Leu Pro Gly Lys Asp Gly Ser Arg Ser Leu Val Val Ala Ala Ile
245 250 255

Lys Asn Thr Asp Thr His Asn Phe Thr Gln Phe Tyr Ser Ala Tyr Glu
260 265 270

Asp Ile Val Arg Arg Ala Arg Asp Gly Lys Leu Thr Ala Glu Asp Phe
275 280 285

Ser Gly Val Thr Ile Ser Leu Thr Asn Pro Gly Gly Ile Gly Thr Val
290 295 300

His Ser Val Pro Arg Leu Met Asn Gly Gln Gly Ala Ile Ile Gly Ala
305 310 315 320

Gly Ala Met Glu Tyr Pro Ala Glu Phe Gln Gly Ala Ser Asp Glu Arg
325 330 335

Leu Ala Glu Ile Gly Val Gly Lys Leu Met Thr Leu Thr Ser Thr Tyr
340 345 350

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

Ile His Asn Leu Leu Ile Ser Asp Glu Phe Tyr Asp Glu Ile Phe His
370 375 380

Ala Leu His Ile Pro Tyr Glu Pro Val Arg Trp Arg Lys Asp Val Pro
385 390 395 400

Glu Gly Ala Val Asp Lys Asn Thr Arg Val Leu Glu Leu Ile Ala Ala
405 410 415

Tyr Arg Asn Arg Gly His Leu Met Ala Asp Thr Asp Pro Leu Gln Phe
420 425 430

Val Lys Asp Lys Phe Arg Ser His Pro Asp Leu Asp Val Arg Thr His
435 440 445

Asp Leu Thr Leu Trp Asp Leu Asp Arg Glu Phe Lys Val Gly Gly Phe
450 455 460

His Gly Gln Glu Lys Met Lys Leu Arg Asp Val Leu Ser Val Leu Arg
465 470 475 480

Asp Ala Tyr Cys Arg His Val Gly Val Glu Tyr Thr His Ile Leu Glu
485 490 495

Pro Glu Gln Gln Gln Trp Leu Gln Asp Arg Val Glu Ala His His Val
500 505 510

Lys Pro Thr Val Ala Gln Gln Lys Tyr Ile Leu Ser Lys Leu Asn Ala
515 520 525

Ala Glu Ala Phe Glu Thr Phe Leu Gln Thr Lys Tyr Val Gly Gln Lys
530 535 540

Arg Phe Ser Leu Glu Gly Ala Glu Ser Val Ile Pro Met Met Asp Ala
545 550 555 560

Val Ile Asp Gln Ala Ala Glu His Gln Leu Asp Glu Val Val Ile Gly
565 570 575

Met Pro His Arg Gly Arg Leu Asn Val Leu Ala Asn Ile Val Gly Lys
580 585 590

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

Ala Ala His Gly Ser Gly Asp Val Lys Tyr His Leu Gly Ala Glu Gly
610 615 620

Thr Tyr Ile Gln Met Phe Gly Asp Asn Asp Ile Thr Val Ser Leu Thr
625 630 635 640

Ala Asn Pro Ser His Leu Glu Ala Val Asp Pro Val Leu Glu Gly Leu
645 650 655

Val Arg Ala Lys Gln Asp Ile Leu Asp Lys Gly Glu Asp Gly Phe Thr
660 665 670

Val Leu Pro Leu Met Leu His Gly Asp Ala Ala Phe Ala Gly Gln Gly
675 680 685

Val Val Ala Glu Thr Leu Asn Leu Ala Leu Leu Arg Gly Tyr Arg Thr
690 695 700

Gly Gly Thr Val His Ile Val Val Asn Asn Gln Val Gly Phe Thr Thr
705 710 715 720

Ala Pro Glu Tyr Ser Arg Ser Ser Glu Tyr Cys Thr Asp Val Ala Lys
725 730 735

Met Ile Gly Ala Pro Ile Phe His Val Asn Gly Asp Asp Pro Glu Ala
740 745 750

Cys Val Trp Val Ala Gln Leu Ala Val Asp Phe Arg Glu Lys Phe Gln
755 760 765

Lys Asp Val Val Ile Asp Met Ile Cys Tyr Arg Arg Arg Gly His Asn
770 775 780

Glu Gly Asp Asp Pro Ser Met Thr Gln Pro Ala Met Tyr Asp Val Ile
785 790 795 800

Asp Thr Lys Arg Ser Val Arg Lys Ser Tyr Thr Glu Ser Leu Ile Gly
805 810 815

Arg Gly Asp Ile Ser Leu Lys Glu Ala Glu Asp Ala Leu Arg Asp Tyr
820 825 830

Gln Gly Gln Leu Glu Arg Val Phe Asn Glu Val Arg Glu Leu Glu Lys
835 840 845

Tyr Thr Pro Glu Pro Ser Glu Ser Val Glu Leu Asp Gln Val Leu Pro
850 855 860

Thr Lys Leu Lys Thr Ser Val Asp Glu Ser Val Leu Glu Arg Ile Gly
865 870 875 880

Asp Ala Phe Val Asn Val Pro Glu Gly Phe Thr Val His Pro Arg Val
885 890 895

Lys Pro Val Ile Glu Lys Arg Arg Glu Met Ser Arg Glu Gly Lys Ile
900 905 910

Asp Trp Ala Phe Ala Glu Leu Leu Ala Phe Gly Ser Leu Val Asp Gln
915 920 925

Gly Lys Met Val Arg Leu Ser Gly Gln Asp Ser Lys Arg Gly Thr Phe
930 935 940

Thr Gln Arg His Ser Val Leu Ile Asp Arg Lys Thr Gly Ala Glu Tyr
945 950 955 960

Thr Pro Leu Gln Asn Leu Gly Ser Glu Asn Pro Gly Lys Phe Leu Val
965 970 975

Tyr Asp Ser Ala Leu Ser Glu Phe Ala Ala Val Gly Phe Glu Tyr Gly
980 985 990

Tyr Ser Val Gly Asn Pro Asp Ala Leu Val Leu Trp Glu Ala Gln Phe
995 1000 1005

Gly Asp Phe Val Asn Gly Ala Gln Ser Ile Ile Asp Glu Phe Ile
1010 1015 1020

Ser Ser Gly Glu Ala Lys Trp Gly Gln Leu Ser Asp Val Val Leu
1025 1030 1035

Leu Leu Pro His Gly His Glu Gly Gln Gly Pro Asp His Thr Ser
1040 1045 1050

Gly Arg Ile Glu Arg Phe Leu Gln Leu Cys Ala Glu Gly Ser Met
1055 1060 1065

Thr Val Ala Val Pro Ser Thr Pro Ala Ser Tyr Phe His Leu Leu
1070 1075 1080

Arg Arg His Ser Leu Asp Gly Ile Arg Arg Pro Leu Val Val Phe
1085 1090 1095

Thr Pro Lys Ser Met Leu Arg Asn Lys Ala Ala Val Ser Asp Val
1100 1105 1110

Glu Asp Phe Thr Thr Gly Lys Phe Arg Ser Val Phe Glu Glu Pro
1115 1120 1125

Thr Tyr Glu Thr Gly Asp Ala Glu Arg Asp Lys Val Arg Arg Val
1130 1135 1140

Leu Leu Val Ser Gly Lys Leu Tyr Trp Glu Leu Leu Ala Lys Lys
1145 1150 1155

Gln Lys Asp Asn Arg Glu Asp Ile Ala Ile Val Arg Ile Glu Gln
1160 1165 1170

Leu Tyr Pro Val Pro Ser Arg Arg Leu Arg Glu Thr Leu Asp Arg
 1175 1180 1185

Tyr Pro Asn Ala Thr Glu Phe Arg Trp Val Gln Glu Glu Pro Ala
 1190 1195 1200

Asn Gln Gly Ala Trp Pro Phe Phe Gly Leu Ala Leu Pro Glu Leu
 1205 1210 1215

Leu Pro Asp Lys Leu Ser Gly Ile Lys Arg Ile Ser Arg Arg Ser
 1220 1225 1230

Met Ser Ala Pro Ser Ser Gly Ser Ser Lys Val His Ala Val Glu
 1235 1240 1245

Gln Gln Glu Ile Ile Asp Glu Ala Phe Gly
 1250 1255

<210> 8
 <211> 3777
 <212> DNA
 <213> Rhodococcus jostii

<400> 8
 atgagctcgt cctcgacctc ccagttcggc cagaaccagt ggctcgtgga tgaaatgtac 60
 cagcgcttcc aggatgatcc gtcgtccgtg gacgcctcct ggcatgagtt cctgacggac 120
 tatagccccg acgccgccgc caaggcgggc gcggcgaacg gccacggcac caacggcacc 180
 accacggcgg cgccggcggc ggcgccgtcc gccaaagccg cgaccccccc ggtccccgag 240
 agcgagacgg ccccgaaacc ccagaccaag acggcgaacg gcgcggcgcc gaaagccgcg 300
 cccaacggcg cggcgccgaa ggccgcggcc cccaagaccg aggcccccaa gaaggccgcc 360
 cccgcgaagg aaacggccgc gacggacgcg aaagcctcgg ccccgcccc cgccgtggag 420
 gaatccaagg tgctgcgggg cgccgccgcc gccgtcgcca agaatatgtc cgcgtcgctc 480
 gccatcccga ccgccacctc cgtccgggcc atccggcgga agctgatgtt cgacaaccgc 540
 atcgtcatca acaaccacct ggcgcggaac cggggcggca agatctcgtt caccatctg 600
 ctgggctacg cgatcgtgca ggccgtgaaa gcgttcccca acatgaaccg ccacttcgcg 660
 gagatcgatg gcaagcccaa tgcggtgacc cccgcccata ccaatctggg cctcgcgatc 720

gatctgccgg gcaaagatgg cagccgctcg ctggctcgtcg cggccatcaa gaacaccgat	780
acgcacaatt tcaccagtt ctatagcgcc tacgaggata tcgtgcgccg ggcgcgggac	840
ggcaaaactga ccgccgaaga cttcagcggc gtcacgatct cgctgaccaa cccgggcggc	900
atcggcaccg tccattccgt cccgcgcctg atgaacggcc agggcgccat catcggcgcg	960
ggcgcgatgg agtaccggc cgagttccag ggcgccctcg acgagcggct ggccgaaatc	1020
ggcgtcggca agctgatgac cctcacgagc acctacgac atcggatcat ccaggcgcc	1080
gaatccggcg acttcctccg gacgatccat aatctgctca tcagcgacga gttctacgac	1140
gaaatcttcc atgccctgca tatcccctat gaaccgtcc ggtggcgcaa agacgtgccc	1200
gaaggcgccg tggataagaa caccgggtc ctggagctca tcgccgccta ccggaaccgc	1260
ggccacctga tggcggacac cgaccgctg cagttcgtga aggacaagtt ccggtcgcac	1320
ccggacctcg acgtgcgcac ccatgatctg acgctctggg acctggatcg ggagttcaaa	1380
gtgggaggct tccacggcca ggagaagatg aagctgcgcg atgtcctgtc ggtgctgcgc	1440
gatgcgtatt gtcgcatgt cggcgtggag tatacccaca tcctggagcc ggagcagcag	1500
cagtggctgc aggatcgct cgaggccac cacgtcaaac cgacggtcgc ccagcagaag	1560
tatatcctgt ccaaactgaa cgccgccgag gccttcgaga cgttcctgca gacgaagtat	1620
gtcggccaga aacgcttctc gctggagggc gcggagagcg tgatcccat gatggacgcc	1680
gtcatcgacc aggccgccga gcatcagctg gacgaggtgg tgatcggcat gccgcatcgg	1740
ggccgcctga atgtcctggc gaacatcgtg ggcaagccgt attccaagat cttcaccgag	1800
ttcgagggca acatgaacc ggcggggcc cacggcagcg gcgacgtgaa gtatcacctg	1860
ggcgccgaag gcacgtatat ccagatgttc ggcgataacg acatcacctg ctcgctcacc	1920
gcgaaccga gccatctgga agccgtggac ccggtgctgg agggcctggt ccgggccaag	1980
caggacatcc tcgacaaggg cgaagacggc ttcacggtgc tccccctcat gctccacggc	2040
gacgcggcct tcgccggcca gggcgtggtc gccgagacgc tcaacctgc cctgctgcgc	2100
ggctaccgga ccggcggcac cgtgcacatc gtggtgaaca accaggctcg cttcacgacc	2160
gccccgagt atagccgctc cagcgagtac tgcacggatg tggcgaagat gatcggcgcc	2220
cccatcttcc acgtcaacgg cgatgacccc gaggcgtgcg tgtgggtggc ccagctggcg	2280
gtggacttcc gcgagaagtt ccagaaggac gtggtcatcg acatgatctg ttatcggcgg	2340

cggggccata atgagggcga cgatccgagc atgaccagc ccgcgatgta cgacgtgac	2400
gacacgaaac gctccgtccg caaatcctat accgagtcgc tgatcggccg cggcgacatc	2460
tcgctgaagg aagcggagga cgccctgcgc gactatcagg gccagctgga acgggtcttc	2520
aatgaggtgc gcgagctgga gaagtacacg cccgaaccga gcgagagcgt cgagctggac	2580
caggtcctgc ccacgaagct gaaaacctcc gtcgatgaat cgggtgctgga gcgcatcggc	2640
gatgccttcg tgaacgtccc ggagggcttc acgggtccatc cgcgggtgaa gccggtgac	2700
gaaaagcgcc gcgagatgag ccgcgaaggc aagatcgact gggccttcgc cgagctgctg	2760
gccttcggct ccctggtgga tcagggcaag atggtgcgc tctccggcca ggacagcaag	2820
cgcggcacgt tcacccagcg cactcggtc ctcatcgacc gcaaaaccgg cgccgaatat	2880
acccccctcc agaacctggg cagcgaaaac ccgggcaagt tcctggtcta cgattccgcg	2940
ctgagcgagt tcgcggcggt cggcttcgag tacggctata gcgtcggcaa ccccgacgcg	3000
ctggtgctgt gggaagccca gttcggcgat ttcgtgaatg gcgcccagag catcatcgac	3060
gagttcatct cgtcgggcga agccaaatgg ggccagctga gcgacgtggt cctgctgctg	3120
ccccacggcc acgaaggcca gggcccggac catacgtccg gccgcatcga acggttcctc	3180
cagctgtgtg ccgaaggctc catgaccgtg gcggtcccca gcacgccggc ctcgtacttc	3240
cacctgctgc gccggcatag cctcgacggc atccgccggc cgctcgtcgt gttcaccccc	3300
aaaagcatgc tgcgcaaca agccgccgtg agcgacgtgg aagacttcac gaccggcaag	3360
ttccggagcg tcttcgagga gcccacctac gagacgggcg atgccgagcg ggacaaggtg	3420
cgccgggtcc tcctggtgtc gggcaagctg tactgggagc tgctggccaa gaagcagaag	3480
gataaccgcg aggacatcgc catcgtgcgc atcgaacagc tgtaccccgt cccctcgcgg	3540
cggctccgcg agacgtgga ccgctacccc aacgccaccg agttccgctg ggtgcaggag	3600
gagcccgcca accagggcgc ctggcccttc ttggcctcg cgctgccga actgctgccc	3660
gacaaactgt cgggcatcaa gcgcatcagc cgccggtcca tgtcggcccc gtcctccggc	3720
tcgtccaagg tgcacgcggt cgaacagcag gaaatcatcg acgaggcctt cggctga	3777

<210> 9

<211> 298

<212> PRT

<213> Escherichia coli

<400> 9

Met Ala Ala Ile Ala Phe Ile Gly Leu Gly Gln Met Gly Ser Pro Met
1 5 10 15

Ala Ser Asn Leu Leu Gln Gln Gly His Gln Leu Arg Val Phe Asp Val
20 25 30

Asn Ala Glu Ala Val Arg His Leu Val Asp Lys Gly Ala Thr Pro Ala
35 40 45

Ala Asn Pro Ala Gln Ala Ala Lys Asp Ala Glu Phe Ile Ile Thr Met
50 55 60

Leu Pro Asn Gly Asp Leu Val Arg Asn Val Leu Phe Gly Glu Asn Gly
65 70 75 80

Val Cys Glu Gly Leu Ser Thr Asp Ala Leu Val Ile Asp Met Ser Thr
85 90 95

Ile His Pro Leu Gln Thr Asp Lys Leu Ile Ala Asp Met Gln Ala Lys
100 105 110

Gly Phe Ser Met Met Asp Val Pro Val Gly Arg Thr Ser Ala Asn Ala
115 120 125

Ile Thr Gly Thr Leu Leu Leu Leu Ala Gly Gly Thr Ala Glu Gln Val
130 135 140

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

Ala Gly Gly Pro Gly Met Gly Ile Arg Val Lys Leu Ile Asn Asn Tyr
165 170 175

Met Ser Ile Ala Leu Asn Ala Leu Ser Ala Glu Ala Ala Val Leu Cys
180 185 190

Glu Ala Leu Asn Leu Pro Phe Asp Val Ala Val Lys Val Met Ser Gly
195 200 205

Thr Ala Ala Gly Lys Gly His Phe Thr Thr Ser Trp Pro Asn Lys Val
 210 215 220

Leu Ser Gly Asp Leu Ser Pro Ala Phe Met Ile Asp Leu Ala His Lys
 225 230 235 240

Asp Leu Gly Ile Ala Leu Asp Val Ala Asn Gln Leu His Val Pro Met
 245 250 255

Pro Leu Gly Ala Ala Ser Arg Glu Val Tyr Ser Gln Ala Arg Ala Ala
 260 265 270

Gly Arg Gly Arg Gln Asp Trp Ser Ala Ile Leu Glu Gln Val Arg Val
 275 280 285

Ser Ala Gly Met Thr Ala Lys Val Lys Met
 290 295

<210> 10
 <211> 897
 <212> DNA
 <213> Escherichia coli

<400> 10
 atggcggcga tcgcgttcat cggcctcggc cagatgggct ccccgatggc cagcaacctc 60
 ctgcagcagg gccatcagct ccgggtgttc gacgtgaacg cggaggcggg ccggcatctg 120
 gtggacaaag gcgcgacccc ggccgccaac cccgcccagg cggccaagga cgcggagtgc 180
 atcatcacca tgctcccaa cggcgacctc gtccgcaatg tgctgttcgg cgagaacggc 240
 gtgtgtgagg gcctgtcgac cgacgccctc gtcacgcaca tgagcaccat ccacccgctg 300
 cagacggata agctcatcgc cgacatgcag gcgaagggtc tcagcatgat ggatgtcccc 360
 gtcggccgca cctccgcgaa cgcgatcacc ggcaccctgc tgctcctggc cggcggcacc 420
 gccgagcagg tggagcgggc cacgcccata ctgatggcga tgggctccga actcatcaac 480
 gcggggcggc ccggcatggg catccgcgtc aagctcatca acaactacat gtcgatcgcg 540
 ctgaatgccc tctccgcgga agcggcggtg ctgtgcgaag cgctgaatct gccgttcgat 600
 gtggcggtga aagtcatgag cggcaccgcg gcgggcaagg gccacttcac cacctcctgg 660
 ccgaacaagg tgctgtccgg cgacctgagc ccggcggttc tgatcgatct ggcccacaag 720

gatctgggca tcgcgctcga cgtggccaac cagctgcacg tcccgatgcc gctcggcgcc 780
gcctccccggg aagtgtattc ccaggcccg cccgccggcc gcggccggca ggattggtcc 840
gccatcctgg agcaggtgcg cgtcagcgcc ggcattgaccg cgaaggtgaa gatgtga 897

<210> 11
<211> 371
<212> PRT
<213> Porphyromonas gingivalis
<400> 11

Met Gln Leu Phe Lys Leu Lys Ser Val Thr His His Phe Asp Thr Phe
1 5 10 15

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

Thr Asn Glu Phe Ile Tyr Glu Pro Tyr Met Lys Ala Cys Gln Leu Pro
35 40 45

Cys His Phe Val Met Gln Glu Lys Tyr Gly Gln Gly Glu Pro Ser Asp
50 55 60

Glu Met Met Asn Asn Ile Leu Ala Asp Ile Arg Asn Ile Gln Phe Asp
65 70 75 80

Arg Val Ile Gly Ile Gly Gly Gly Thr Val Ile Asp Ile Ser Lys Leu
85 90 95

Phe Val Leu Lys Gly Leu Asn Asp Val Leu Asp Ala Phe Asp Arg Lys
100 105 110

Ile Pro Leu Ile Lys Glu Lys Glu Leu Ile Ile Val Pro Thr Thr Cys
115 120 125

Gly Thr Gly Ser Glu Val Thr Asn Ile Ser Ile Ala Glu Ile Lys Ser
130 135 140

Arg His Thr Lys Met Gly Leu Ala Asp Asp Ala Ile Val Ala Asp His
145 150 155 160

Ala Ile Ile Ile Pro Glu Leu Leu Lys Ser Leu Pro Phe His Phe Tyr
165 170 175

Ala Cys Ser Ala Ile Asp Ala Leu Ile His Ala Ile Glu Ser Tyr Val
180 185 190

Ser Pro Lys Ala Ser Pro Tyr Ser Arg Leu Phe Ser Glu Ala Ala Trp
195 200 205

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

Tyr Arg Phe Glu Lys Leu Gly Glu Met Ile Met Ala Ser Asn Tyr Ala
225 230 235 240

Gly Ile Ala Phe Gly Asn Ala Gly Val Gly Ala Val His Ala Leu Ser
245 250 255

Tyr Pro Leu Gly Gly Asn Tyr His Val Pro His Gly Glu Ala Asn Tyr
260 265 270

Gln Phe Phe Thr Glu Val Phe Lys Val Tyr Gln Lys Lys Asn Pro Phe
275 280 285

Gly Tyr Ile Val Glu Leu Asn Trp Lys Leu Ser Lys Ile Leu Asn Cys
290 295 300

Gln Pro Glu Tyr Val Tyr Pro Lys Leu Asp Glu Leu Leu Gly Cys Leu
305 310 315 320

Leu Thr Lys Lys Pro Leu His Glu Tyr Gly Met Lys Asp Glu Glu Val
325 330 335

Arg Gly Phe Ala Glu Ser Val Leu Lys Thr Gln Gln Arg Leu Leu Ala
340 345 350

Asn Asn Tyr Val Glu Leu Thr Val Asp Glu Ile Glu Gly Ile Tyr Arg
355 360 365

Arg Leu Tyr
370

<210> 12
 <211> 1116
 <212> DNA
 <213> Porphyromonas gingivalis

<400> 12
 atgcagctgt tcaaactcaa aagcgtgacg caccacttcg atacgttcgc ggagttcgcc 60
 aaggagttct gcctcggcga gcgggacctg gtcatcacca atgagttcat ctacgagccc 120
 tatatgaagg cgtgccagct cccgtgccac ttcgtcatgc aggaaaagta cggccagggc 180
 gaaccgtcgg atgagatgat gaacaacatc ctggcggata tccggaacat ccagttcgac 240
 cgggtcatcg gcatcggcgg cggcaccgtc atcgacatct ccaaactgtt cgtcctgaaa 300
 ggcttcaacg acgtgctgga cgccttcgac cgcaaaatcc cgctcatcaa ggagaaggag 360
 ctcatcatcg tgcccacgac ctgcggcacc ggctcggag tgaccaacat ctcgatcgcc 420
 gagatcaagt cgcgccacac caagatgggc ctggcggatg acgcgatcgt ggccgaccat 480
 gccatcatca tccccgagct gctgaagagc ctgccgttcc acttctatgc gtgctccgcc 540
 atcgatgccc tgatccacgc catcgagagc tacgtctcgc cgaaggcctc cccgtacagc 600
 cggctgttca gcgaggcggc ctgggacatc atcctggaag tcttcaagaa gatcgccgag 660
 cacggccccg aatatcggtt cgaaaagctc ggcgagatga tcatggccag caattacgcc 720
 ggcatcgcct tcggcaacgc cggcgtcggc gccgtccacg ccctgtcgta cccgctgggc 780
 ggcaattatc atgtgccgca tggcgaagcg aattaccagt tcttcaccga ggtgttcaag 840
 gtctatcaga agaaaaacc gttcggctac atcgtcgagc tgaactggaa gctgtcgaag 900
 atcctgaact gccagccgga gtatgtgtat cccaagctgg atgaactcct gggctgcctg 960
 ctgacaaaaa agccgctgca tgaatatggc atgaaggacg aagaggtgcg gggcttcgcc 1020
 gaatcggtgc tgaagacca gcagcgcctg ctgccaaca actacgtgga gctgaccgtc 1080
 gacgaaatcg aaggcatcta ccgccggctg tactga 1116

<210> 13
 <211> 371
 <212> PRT
 <213> Clostridium kluyveri

<400> 13

Met Lys Leu Leu Lys Leu Ala Pro Asp Val Tyr Lys Phe Asp Thr Ala
1 5 10 15

Glu Glu Phe Met Lys Tyr Phe Lys Val Gly Lys Gly Asp Phe Ile Leu
20 25 30

Thr Asn Glu Phe Leu Tyr Lys Pro Phe Leu Glu Lys Phe Asn Asp Gly
35 40 45

Ala Asp Ala Val Phe Gln Glu Lys Tyr Gly Leu Gly Glu Pro Ser Asp
50 55 60

Glu Met Ile Asn Asn Ile Ile Lys Asp Ile Gly Asp Lys Gln Tyr Asn
65 70 75 80

Arg Ile Ile Ala Val Gly Gly Gly Ser Val Ile Asp Ile Ala Lys Ile
85 90 95

Leu Ser Leu Lys Tyr Thr Asp Asp Ser Leu Asp Leu Phe Glu Gly Lys
100 105 110

Val Pro Leu Val Lys Asn Lys Glu Leu Ile Ile Val Pro Thr Thr Cys
115 120 125

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

Arg His Thr Lys Lys Gly Ile Ala Ser Asp Glu Leu Tyr Ala Thr Tyr
145 150 155 160

Ala Val Leu Val Pro Glu Phe Ile Lys Gly Leu Pro Tyr Lys Phe Phe
165 170 175

Val Thr Ser Ser Val Asp Ala Leu Ile His Ala Thr Glu Ala Tyr Val
180 185 190

Ser Pro Asn Ala Asn Pro Tyr Thr Asp Met Phe Ser Val Lys Ala Met
195 200 205

Glu Leu Ile Leu Asn Gly Tyr Met Gln Met Val Glu Lys Gly Asn Asp
210 215 220

Tyr Arg Val Glu Ile Ile Glu Asp Phe Val Ile Gly Ser Asn Tyr Ala
225 230 235 240

Gly Ile Ala Phe Gly Asn Ala Gly Val Gly Ala Val His Ala Leu Ser
245 250 255

Tyr Pro Ile Gly Gly Asn Tyr His Val Pro His Gly Glu Ala Asn Tyr
260 265 270

Leu Phe Phe Thr Glu Ile Phe Lys Thr Tyr Tyr Glu Lys Asn Pro Asn
275 280 285

Gly Lys Ile Lys Asp Val Asn Lys Leu Leu Ala Gly Ile Leu Lys Cys
290 295 300

Asp Glu Ser Glu Ala Tyr Asp Ser Leu Ser Gln Leu Leu Asp Lys Leu
305 310 315 320

Leu Ser Arg Lys Pro Leu Arg Glu Tyr Gly Met Lys Glu Glu Glu Ile
325 330 335

Glu Thr Phe Ala Asp Ser Val Ile Glu Gly Gln Gln Arg Leu Leu Val
340 345 350

Asn Asn Tyr Glu Pro Phe Ser Arg Glu Asp Ile Val Asn Thr Tyr Lys
355 360 365

Lys Leu Tyr
370

<210> 14
<211> 1116
<212> DNA
<213> Clostridium kluyveri

<400> 14
atgaagctcc tgaagctggc cccggatgtg tacaaattcg ataccgccga ggagttcatg 60
aagtatttca aggtcggcaa gggcgacttc atcctgacga acgagttcct ctataaaccc 120
ttcctggaaa agttcaatga tggcgccgac gccgtgttcc aggaaaagta cggcctgggc 180

gagccgtccg acgaaatgat caacaacatc atcaaggata tcggcgacaa acagtacaat 240
cggatcatcg cgggtgggcgg cggcagcgtc atcgatatcg ccaaaatcct gtcgctgaaa 300
tacacggacg attccctgga cctgttcgaa ggcaaggtcc cgctggtgaa gaataaagag 360
ctcatcatcg tcccgaccac ctgcggcacc ggctcggag tgacgaatgt ctcggtcgcc 420
gaactgaagc ggcgccatac gaagaaaggc atgcctccg acgaactcta tgcgacgtac 480
gcggtcctgg tgcccagatt catcaaaggc ctgccctata agttcttcgt gacgagctcc 540
gtcgacgccc tcatccatgc gaccgaggcc tacgtgtcgc cgaacgcaa cccctacacc 600
gacatgttct cggatcaaggc gatggaactg atcctgaacg gctacatgca gatggtggag 660
aagggcaacg attatcgctt ggaaatcatc gaggatttcg tgatcggctc caactacgag 720
ggcatcgcgt tcggcaacgc gggcgtcggc gcggtccatg ccctcagcta ccccatcggc 780
ggcaactacc atgtcccga tggcgaagcc aactacctt tcttcacgga gatcttcaag 840
acctactacg agaagaaccc gaatggcaag atcaaggacg tgaataagct cctggcgggc 900
atcctcaagt gcgatgagtc cgaggcgtag gattcgtctt cgcagctgct ggacaaactg 960
ctctcgcgca agccgctgag cgagtacggc atgaaggagg aagagatcga aaccttcgcc 1020
gacagcgtga tcgaaggcca gcagcggctg ctggtcaaca actacgagcc cttcagccgc 1080
gaggatatcg tcaacaccta caaaaagctc tactga 1116

<210> 15
<211> 431
<212> PRT
<213> Porphyromonas gingivalis

<400> 15

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

Glu Ala Val Lys His Ile Lys Asn Gly Glu Arg Val Ala Leu Ser His
20 25 30

Ala Ala Gly Val Pro Gln Ser Cys Val Asp Ala Leu Val Gln Gln Ala
35 40 45

Asp Leu Phe Gln Asn Val Glu Ile Tyr His Met Leu Cys Leu Gly Glu
50 55 60

Gly Lys Tyr Met Ala Pro Glu Met Ala Pro His Phe Arg His Ile Thr
65 70 75 80

Asn Phe Val Gly Gly Asn Ser Arg Lys Ala Val Glu Glu Asn Arg Ala
85 90 95

Asp Phe Ile Pro Val Phe Phe Tyr Glu Val Pro Ser Met Ile Arg Lys
100 105 110

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

Glu Asn Gly Tyr Cys Ser Phe Gly Val Ser Cys Asp Tyr Ser Lys Pro
130 135 140

Ala Ala Glu Ser Ala His Leu Val Ile Gly Glu Ile Asn Arg Gln Met
145 150 155 160

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

Ile Val Met Ala Asp Tyr Pro Ile Tyr Ser Leu Ala Lys Pro Lys Ile
180 185 190

Gly Glu Val Glu Glu Ala Ile Gly Arg Asn Cys Ala Glu Leu Ile Glu
195 200 205

Asp Gly Ala Thr Leu Gln Leu Gly Ile Gly Ala Ile Pro Asp Ala Ala
210 215 220

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

Phe Ser Asp Gly Val Val Glu Leu Val Arg Ser Gly Val Ile Thr Gly
245 250 255

Lys Lys Lys Thr Leu His Pro Gly Lys Met Val Ala Thr Phe Leu Met
260 265 270

Gly Ser Glu Asp Val Tyr His Phe Ile Asp Lys Asn Pro Asp Val Glu
275 280 285

Leu Tyr Pro Val Asp Tyr Val Asn Asp Pro Arg Val Ile Ala Gln Asn
290 295 300

Asp Asn Met Val Ser Ile Asn Ser Cys Ile Glu Ile Asp Leu Met Gly
305 310 315 320

Gln Val Val Ser Glu Cys Ile Gly Ser Lys Gln Phe Ser Gly Thr Gly
325 330 335

Gly Gln Val Asp Tyr Val Arg Gly Ala Ala Trp Ser Lys Asn Gly Lys
340 345 350

Ser Ile Met Ala Ile Pro Ser Thr Ala Lys Asn Gly Thr Ala Ser Arg
355 360 365

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

Glu Val Asp Tyr Val Val Thr Glu Tyr Gly Ile Ala Gln Leu Lys Gly
385 390 395 400

Lys Ser Leu Arg Gln Arg Ala Glu Ala Leu Ile Ala Ile Ala His Pro
405 410 415

Asp Phe Arg Glu Glu Leu Thr Lys His Leu Arg Lys Arg Phe Gly
420 425 430

<210> 16
<211> 1296
<212> DNA
<213> Porphyromonas gingivalis

<400> 16
atgaaggacg tgctggccga gtacgcctcg cgcatacgtct cggcggaaga agcgggtcaag 60
cacatcaaga acggcgagcg cgtcgcgctc tcgcatgccg ccggcggtccc gcagtcctgc 120
gtcgaatgcc tggtccagca ggccgatctc ttccagaatg tcgaaatcta ccatatgctc 180
tgcttgggagc agggcaagta catggccccc gagatggcgc cgcatttccg ccatatcacg 240

```

aatttcgtcg gcggcaactc gcggaaagcc gtggaggaga accgggccga cttcatcccc      300
gtgttcttct acgaggtccc cagcatgata cgcaaagata tcctccacat cgacgtggcg      360
atcgtgcagc tgagcatgcc ggatgagaac ggctactgca gcttcggcgt gagctgtgac      420
tacagcaagc cggccgccga atccgccac ctcgtcatcg gcgagatcaa ccggcagatg      480
ccgtacgtcc acggcgacaa cctgatccac atcagcaagc tggactacat cgtcatggcc      540
gactatccga tctatagcct ggcgaaaccc aagatcggcg aggtcgagga ggccatcggc      600
cggaactgcg ccgagctgat cgaagacggc gccaccctcc agctgggcat cggcgccatc      660
ccggacgccg ccctgctctt cctgaaagac aagaaagacc tgggcatcca caccgaaatg      720
ttcagcgacg gcgtcgtcga actgggtccg agcggcgtca tcaccggcaa gaaaaaaacc      780
ctgcacccgg gcaaaatggt cgcgacgttc ctgatgggct ccgaagatgt ctatcatttc      840
atcgacaaga accccgacgt ggagctctac ccggtggact atgtgaacga cccgcgcgtc      900
atgcccaga acgacaacat ggtcagcatc aactcctgta tcgagatcga cctcatgggc      960
caggtcgtga gcgaatgtat cggctccaag cagttctcgg gcacgggchg ccaggtggac     1020
tacgtccgcg gcgcggcggtg gtcgaagaac ggcaagtcga tcatggccat cccctccacc     1080
gccaagaacg gcaccgcctc gcgcatcgtg ccgatcatcg cggagggcgc cgcggtgacg     1140
accctgcgga acgaagtgga ttacgtggtc acggagtatg gcatcgccca gctgaagggc     1200
aagtcgctcc gccagcgggc ggaagcgctg atcgcgatcg cgcacccgga cttccgggag     1260
gagctgacca aacacctgcg gaaacggttc ggctga                                1296

```

```

<210> 17
<211> 431
<212> PRT
<213> Porphyromonas gingivalis

```

```
<400> 17
```

```

Met Gln Trp Gln Glu Leu Tyr Arg Gln Arg Val Cys Ser Ala Asp Glu
1           5           10          15

```

```

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

```

```

Ala Ala Ala Ala Pro Val Arg Phe Ser Gln Ala Met Tyr Arg Gln Arg
          35          40          45

```

Glu Lys Leu Glu Asn Ile Thr Val Phe His Met Leu Tyr Phe Gly Asp
50 55 60

Ala Pro His Leu Ala Pro Glu Met Arg Ser His Val His Pro Thr Leu
65 70 75 80

Asn Phe Leu Glu Gly Asn Ser Arg Pro Ala Ser Arg Asp Arg Arg Val
85 90 95

Asp Phe Ile Pro Cys His Phe His Glu Val Pro Glu Leu Phe Arg Gln
100 105 110

Gly Phe Phe Pro Leu Asp Val Ala Val Val Gln Val Ser Thr Pro Asn
115 120 125

Glu Glu Gly Tyr Cys Ser Phe Gly Val Ser Cys Asp Tyr Thr Lys Ala
130 135 140

Ala Ala Glu Cys Ala Pro Val Val Val Ala Glu Val Asn Lys Gln Met
145 150 155 160

Pro Phe Ile Gly Gly Glu Asn Leu Ile His Ile Ser Lys Leu Thr His
165 170 175

Ile Ile Glu Val Asp Glu Pro Ile Ala Glu Val Leu Pro Pro Ala Ile
180 185 190

Ser Asp Leu Glu Leu Arg Ile Gly Gln Asn Cys Ala Ser Leu Ile Lys
195 200 205

Asp Gly Asp Thr Leu Gln Leu Gly Ile Gly Gly Ile Pro Asp Ala Val
210 215 220

Leu Arg Ala Leu Glu Gly His Lys Asp Leu Gly Ile His Thr Glu Met
225 230 235 240

Phe Thr Asp Gly Val Met Arg Met Ile Arg Lys Gly Ile Ile Asn Gly
245 250 255

Lys Lys Lys Thr Leu His Pro Glu Lys Val Val Thr Ser Leu Ile Phe
260 265 270

Gly Ser Lys Glu Leu Tyr Asp Phe Val Asn Asn Asn Pro Val Ile Glu
275 280 285

Cys Tyr Pro Val Asp Tyr Ile Asn Asn Pro Asp Val Ile Gly Lys Asn
290 295 300

Asp Arg Met Val Ser Ile Asn Ser Cys Leu Glu Met Asp Leu Met Gly
305 310 315 320

Gln Ala Ala Ser Glu Ser Ile Gly Tyr Glu Gln Phe Ser Gly Ser Gly
325 330 335

Gly Gln Val Asp Phe Leu Arg Gly Ala Lys Arg Ser Lys Gly Gly Ile
340 345 350

Ser Ile Met Ala Phe Pro Ser Thr Ala Lys Lys Gly Thr Glu Ser Arg
355 360 365

Ile Val Pro Ile Leu Lys Glu Gly Ala Cys Val Thr Thr Gly Arg Asn
370 375 380

Glu Val Asp Tyr Val Val Thr Glu Tyr Gly Val Ala Arg Leu Arg Gly
385 390 395 400

Ala Thr Leu Arg Gln Arg Ala Glu Ala Leu Thr Ala Ile Ala His Pro
405 410 415

Asp Phe Arg Pro Ala Leu Glu Glu Glu Ile Arg Arg Arg Phe Glu
420 425 430

<210> 18

<211> 1296

<212> DNA

<213> Porphyromonas gingivalis

<400> 18

atgcagtggc aggaactcta ccggcagcgc gtgtgctccg cggatgaagc cgtcgtggat 60

agcctgaagc ccggcaccaa ggtggtcttc ggccatgccg ccgccgcgcc cgtgcggttc 120

tcgcaggcca tgtatcgcca gcgcgaaaag ctcgaaaaca tcaccgtggt ccatatgctc	180
tatttcggcg atgcccccca cctggccccc gaaatgcgct cccatgtcca cccgacgctc	240
aacttcctgg aaggcaatag ccgccccgcg tcgcggggtc ggcgcggtga tttcatcccc	300
tgccacttcc acgaagtgcc cgagctgttc cgccagggtc tcttccccct cgacgtggcg	360
gtggtgcagg tgtcgacccc gaacgaggag ggctattgta gcttcggcgt ctcgtgtgac	420
tatacgaagg ccgcggccga atgtgccccg gtcgtgggtg ccgaggtgaa caagcagatg	480
ccgttcatcg gcggcgagaa cctgatccat atctccaagc tgacccatat catcgaagtg	540
gatgagccga tcgccgaggt gctgcccccc gcgatcagcg acctggagct gcgcatcggc	600
cagaactgcg cgtccctcat caaggacggc gatacgtcc agctgggcat cggcggcatc	660
ccggacgcgg tgctgcgggc cctggaaggc cataaagacc tcggcatcca cacggagatg	720
ttcaccgacg gcgtgatgcg gatgatccgg aagggcata tcaacggcaa gaaaaagacc	780
ctgcacccgg aaaaagtggg gaccagcctg atcttcggca gcaaagaact gtacgacttc	840
gtcaacaaca acccggatgat cgagtgttat ccggtcgatt atatcaacaa tccgatgtc	900
atcggcaaga acgaccgcat ggtgagcatc aactcctgcc tggaaatgga cctgatgggc	960
caggccgcga gcgaaagcat cggctacgaa cagttctccg gctcgggcgg ccaggtggat	1020
ttcctgcggg gcgcgaagcg gtccaaaggc ggcatcagca tcatggcctt cccctccacg	1080
gccaagaagg gcacggagtc ccgcatcgtg cccatcctga aggaggggcg gtgcgtcacg	1140
accggccgca atgaggtgga ttacgtcgtc accgagtac gcgtcgcccg cctccggggc	1200
gccaccctcc ggcagcgggc cgaggccctg acggccatcg cccatcccga cttccgcccc	1260
gccctggagg aagagatccg ccggcgcttc gagtga	1296

<210> 19
 <211> 438
 <212> PRT
 <213> Clostridium aminobutyricum

<400> 19

Met	Asp	Trp	Lys	Lys	Ile	Tyr	Glu	Asp	Arg	Thr	Cys	Thr	Ala	Asp	Glu
1				5					10					15	

Ala	Val	Lys	Ser	Ile	Lys	Ser	Gly	Asp	Arg	Val	Leu	Phe	Ala	His	Cys
			20					25					30		

Val Ala Glu Pro Pro Val Leu Val Glu Ala Met Val Ala Asn Ala Ala
35 40 45

Ala Tyr Lys Asn Val Thr Val Ser His Met Val Thr Leu Gly Lys Gly
50 55 60

Glu Tyr Ser Lys Pro Glu Tyr Lys Glu Asn Phe Thr Phe Glu Gly Trp
65 70 75 80

Phe Thr Ser Pro Ser Thr Arg Gly Ser Ile Ala Glu Gly His Gly Gln
85 90 95

Phe Val Pro Val Phe Phe His Glu Val Pro Ser Leu Ile Arg Lys Asp
100 105 110

Ile Phe His Val Asp Val Phe Met Val Met Val Ser Pro Pro Asp His
115 120 125

Asn Gly Phe Cys Cys Val Gly Val Ser Ser Asp Tyr Thr Met Gln Ala
130 135 140

Ile Lys Ser Ala Lys Ile Val Leu Ala Glu Val Asn Asp Gln Val Pro
145 150 155 160

Val Val Tyr Gly Asp Thr Phe Val His Val Ser Asn Ile Asp Lys Phe
165 170 175

Val Glu Thr Ser His Pro Leu Pro Glu Ile Gly Leu Pro Lys Ile Gly
180 185 190

Glu Val Glu Ala Ala Ile Gly Lys His Cys Ala Ser Leu Ile Glu Asp
195 200 205

Gly Ser Thr Leu Gln Leu Gly Ile Gly Ala Ile Pro Asp Ala Val Leu
210 215 220

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

Ser Asp Gly Val Val Asp Leu Tyr Glu Ala Gly Val Ile Asp Cys Ser
245 250 255

Gln Lys Ser Ile Asp Lys Gly Lys Met Ala Ile Thr Phe Leu Met Gly
260 265 270

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

Lys Thr Val Asp Tyr Ile Asn His Pro Ser Val Val Ala Gln Cys Ser
290 295 300

Lys Met Val Cys Ile Asn Ala Cys Leu Gln Val Asp Phe Met Gly Gln
305 310 315 320

Ile Val Ser Asp Ser Ile Gly Thr Lys Arg Phe Ser Gly Val Gly Gly
325 330 335

Gln Val Asp Phe Val Arg Gly Ala Ser Met Ser Ile Asp Gly Lys Gly
340 345 350

Lys Ala Ile Ile Ala Met Pro Ser Val Ala Lys Lys Lys Asp Gly Ser
355 360 365

Met Ile Ser Lys Ile Val Pro Phe Ile Asp His Gly Ala Ala Val Thr
370 375 380

Thr Ser Arg Asn Asp Ala Asp Tyr Val Val Thr Glu Tyr Gly Ile Ala
385 390 395 400

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

Ile Ala His Pro Asp Phe Lys Asp Glu Leu Lys Ala Glu Phe Glu Lys
420 425 430

Arg Phe Asn Ala Ala Phe
435

<210> 20

<211> 1317

<212> DNA

<213> *Clostridium aminobutyricum*

<400> 20

```
atggattgga aaaagatcta cgaagatcgc acctgcaccg ccgatgaagc cgtcaagtcc      60
atcaagtccg gcgaccgcgt cctgttcgcc cactgtgtcg ccgaaccccc ggtgctggtc     120
gaagcgaatg tggccaacgc cgcggcgtag aagaacgtca ccgtgagcca tatggtgacc     180
ctgggcaagg gcgaatatag caagcccgag tacaaggaga acttcacgtt cgagggtggt      240
ttcacctccc cgtcgacgcg gggctcgatc gcggaggggc accgccagtt cgtgcccgtg      300
ttcttccatg aggtgccgtc cctgatccgc aaggatatct tccacgtcga cgtgttcatg     360
gtcatggtgt ccccgcccga tcataacggc ttctgctgtg tcggcgtagc cagcgactac     420
acgatgcagg ccatcaagtc ggccaagatc gtgctggcgg aagtcaacga ccaggtgccc     480
gtggtgtatg gcgatacgtt cgtgcatgtc tcgaacatcg ataagttcgt ggaaacctcg     540
catccgctcc cggaaatcgg cctgccgaaa atcggcgagg tggaagcggc catcggcaaa     600
cactgtgcct ccctgatcga agacggctcg acgctgcagc tgggcatcgg cgccatcccc     660
gacgccgtgc tcagccagct gaaagacaag aagcacctgg gcatccactc ggagatgata     720
tccgacggcg tcgtggacct gtatgaggcg ggcgtgatcg actgctcca gaagtcgata     780
gacaagggca aaatggcgat caccttcttg atgggcacca agcgcctgta cgacttcgcg     840
gcgaacaatc cgaagggtga gctgaaaacg gtggactaca tcaaccatcc gtcggtggtg     900
gcccagtgct ccaagatggt gtgcatcaac gcctgcctgc aggtggattt catgggccag     960
atcgtgagcg acagcatcgg caccaagcgc ttctcgggcg tgggcggcca ggtcgacttc    1020
gtcgggggcg cctccatgag catcgacggc aagggaagg cgatcatcgc catgccgtcg    1080
gtcgccaaga agaaggacgg cagcatgata tcgaaaatcg tccccttcat cgatcacggc    1140
gccgccgtga ccacctcgcg gaacgacgcc gattatgtgg tcacggagta cggcatcgcc    1200
gagatgaaag gcaagtcctt ccaggaccgg gcccgggcgc tcatcaacat cgcgcacccg    1260
gatttcaagg atgagctgaa ggcggagttc gagaagcggc tcaacgcggc gttctga      1317
```

<210> 21

<211> 858

<212> PRT

<213> *Clostridium acetobutylicum*

<400> 21

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> 22
 <211> 2577
 <212> DNA
 <213> Clostridium acetobutylicum

<400> 22
 atgaaggtga ccaaccagaa agaactgaag cagaaactga acgaactgcg ggaggcgcag 60
 aaaaagtctg ccacctatac ccaggagcag gtggacaaaa tcttcaagca gtgtgcgatac 120
 gccgccgcca aagaacggat caacctggcg aaactcgccg tcgaagaaac cggcatcggc 180
 ctcgtggagg acaagatcat caaaaatcac ttcgccgccg aatatactta taacaaatac 240
 aagaacgaga agacctgcgg catcatcgac catgacgact cgctgggcat caccaaagtc 300
 gccgagccca tcggcatcgt ggccgcgatac gtgcccacca ccaacccgac gagcaccgcc 360
 atcttcaagt cgctgatctc cctgaagacg cgcaatgcga tcttcttcag ccccatccg 420
 cgggccaaaa agtcgacgat cgcggcggcc aaactgatcc tggacgccgc ggtcaaggcc 480
 ggcccccca aaaatatcat cggctggatac gacgaaccgt ccatcgagct gtcccaggac 540
 ctcatgtccg aagccgacat catcctggcg accggcggcc cgagcatggt gaaggccgcc 600
 tattccagcg gcaaaccggc catcggcgta ggccgccgca acacccccgc gatcatcgac 660
 gagtccgccg acatcgacat ggccgtctcc tcgatcatcc tgagcaagac ctacgataac 720
 ggcgatcatc gcgcgtcgga acagagcatt ctggatcatga actccatcta tgaaaaggta 780
 aaggaagagt tcgtgaagcg cggctcgtat atcctgaacc agaataaat cgcgaagatc 840
 aaggagacga tgttcaagaa cggcgccatc aacgccgaca tcgtcggcaa gagcgcctat 900
 atcatcgca agatggccgg catcgaagtg ccgcagacca ccaagatcct gatcggcgaa 960
 gtccagtccg tcgaaaagtc cgaactgttc tcccacgaaa agctgagccc ggtgctcgcc 1020
 atgtacaaag tcaaagactt cgacgaggcc ctcaaaaagg cgacgaggct catcgaactg 1080
 ggccggcagc gccatacctc gtcgctgtac atcgatagcc agaacaacaa ggacaaagtc 1140
 aaagagtctg gcctggcgat gaagacgagc cgcaccttca tcaacatgcc gagcagccag 1200
 ggccgctccg gcgacctgta caacttcgcc atcgccccga gcttcacgct cggctgcggc 1260
 acctggggcg gcaattcggg gtcgcagaac gtggaaccga aacatctgct gaacatcaaa 1320
 tccgtcgcgg agcgcgcga aaatatgctc tggttcaagg tccgcagaa gatctatttc 1380

```

aaatatggct gcctgcgctt cgccctcaaa gaactgaaag acatgaataa gaaacgcgcc      1440
ttcatcgtca ccgataaaga tctgttcaag ctgggctacg tgaacaaaat cacgaaggct      1500
ctggatgaga tcgatatcaa gtacagcatc ttaccgaca tcaagtccga tcccaccatc      1560
gactccgtca agaagggcgc caaagagatg ctcaacttcg aaccggacac catcatcagc      1620
atcggcggcg gctcgccgat ggacgcggcc aaggtcatgc atctgctcta tgaatatccg      1680
gaggccgaga tcgaaaacct ggccatcaac ttcatggaca tccggaaacg catctgcaac      1740
ttcccgaagc tcggcacgaa agccatcagc gtggccatcc ccacgaccgc gggcaccggc      1800
tcggaagcca cgccgttcgc ggtcatcacg aatgacgaaa ccggcatgaa ataccccctg      1860
acctcctacg agctaccccc caatatggcc atcatcgaca cggagctcat gctcaacatg      1920
ccgcggaagc tcaccgccgc caccggcatc gatgccctgg tgcacgccat cgaagcctac      1980
gtgagcgtga tggcgaccga ctataccgat gagctggccc tccgggcat caagatgatc      2040
ttcaagtatc tgccgcgggc ctacaaaaac ggcacgaatg acatcgaggc ccgcgaaaag      2100
atggcgcgat cgagcaacat cgcgggcatg gccttcgcga acgccttcct gggcgtgtgc      2160
catagcatgg cgcacaaaact gggcgcgatg catcacgtcc cgcattggcat cgcctgtgcc      2220
gtcctgatcg aggaggtcat caagtacaac gccacggatt gtccgaccaa acagaccgcc      2280
ttcccgagt acaagtcccc gaacgcgaag cggaataacg cggagatcgc ggagtacctg      2340
aatctgaaag gcaccagcga caccgaaaaa gtcacggccc tcatcgaagc catctcgaaa      2400
ctgaaaatcg atctgagcat cccccagaac atctccgcgg ccggcatcaa taagaaggat      2460
ttctataaca cgctggacaa gatgagcgag ctggcgttcg atgaccagtg caccaccgcc      2520
aacccccggt atccgctgat ctccgaactc aaggacatct acatcaagtc cttctga      2577

```

```

<210> 23
<211> 891
<212> PRT
<213> Escherichia coli

```

```
<400> 23
```

```

Met Ala Val Thr Asn Val Ala Glu Leu Asn Ala Leu Val Glu Arg Val
1           5           10          15

```

```

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

```

Lys Ile Phe Arg Ala Ala Ala Leu Ala Ala Ala Asp Ala Arg Ile Pro
35 40 45

Leu Ala Lys Met Ala Val Ala Glu Ser Gly Met Gly Ile Val Glu Asp
50 55 60

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

Lys Asp Glu Lys Thr Cys Gly Val Leu Ser Glu Asp Asp Thr Phe Gly
85 90 95

Thr Ile Thr Ile Ala Glu Pro Ile Gly Ile Ile Cys Gly 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 Ile Phe Ser Pro His Pro Arg Ala Lys Asp
130 135 140

Ala Thr Asn Lys Ala Ala Asp Ile Val Leu Gln Ala Ala Ile Ala Ala
145 150 155 160

Gly Ala Pro Lys Asp Leu Ile Gly Trp Ile Asp Gln Pro Ser Val Glu
165 170 175

Leu Ser Asn Ala Leu Met His His Pro Asp Ile Asn Leu Ile Leu Ala
180 185 190

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

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

Ala Asp Ile Lys Arg Ala Val Ala Ser Val Leu Met Ser Lys Thr Phe
225 230 235 240

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

Ser Val Tyr Asp Ala Val Arg Glu Arg Phe Ala Thr His Gly Gly Tyr
260 265 270

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

Asn Gly Ala Leu Asn Ala Ala Ile Val Gly Gln Pro Ala Tyr Lys Ile
290 295 300

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

Gly Glu Val Thr Val Val Asp Glu Ser Glu Pro Phe Ala His Glu Lys
325 330 335

Leu Ser Pro Thr Leu Ala Met Tyr Arg Ala Lys Asp Phe Glu Asp Ala
340 345 350

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

Ser Cys Leu Tyr Thr Asp Gln Asp Asn Gln Pro Ala Arg Val Ser Tyr
370 375 380

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

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

Leu Thr Leu Gly Cys Gly Ser Trp Gly Gly Asn Ser Ile Ser Glu Asn
420 425 430

Val Gly Pro Lys His Leu Ile Asn Lys Lys Thr Val Ala Lys Arg Ala
435 440 445

Glu Asn Met Leu Trp His Lys Leu Pro Lys Ser Ile Tyr Phe Arg Arg
450 455 460

Gly Ser Leu Pro Ile Ala Leu Asp Glu Val Ile Thr Asp Gly His Lys
465 470 475 480

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

Asp Gln Ile Thr Ser Val Leu Lys Ala Ala Gly Val Glu Thr Glu Val
500 505 510

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

Ala Glu Leu Ala Asn Ser Phe Lys Pro Asp Val Ile Ile Ala Leu Gly
530 535 540

Gly Gly Ser Pro Met Asp Ala Ala Lys Ile Met Trp Val Met Tyr Glu
545 550 555 560

His Pro Glu Thr His Phe Glu Glu Leu Ala Leu Arg Phe Met Asp Ile
565 570 575

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

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

Ala Val Val Thr Asp Asp Ala Thr Gly Gln Lys Tyr Pro Leu Ala Asp
610 615 620

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

Asp Met Pro Lys Ser Leu Cys Ala Phe Gly Gly Leu Asp Ala Val Thr
645 650 655

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

Gly Gln Ala Leu Gln Ala Leu Lys Leu Leu Lys Glu Tyr Leu Pro Ala
675 680 685

Ser Tyr His Glu Gly Ser Lys Asn Pro Val Ala Arg Glu Arg Val His
690 695 700

Ser Ala Ala Thr Ile Ala Gly Ile Ala Phe Ala Asn Ala Phe Leu Gly
705 710 715 720

Val Cys His Ser Met Ala His Lys Leu Gly Ser Gln Phe His Ile Pro
725 730 735

His Gly Leu Ala Asn Ala Leu Leu Ile Cys Asn Val Ile Arg Tyr Asn
740 745 750

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

Pro Gln Ala Arg Arg Arg Tyr Ala Glu Ile Ala Asp His Leu Gly Leu
770 775 780

Ser Ala Pro Gly Asp Arg Thr Ala Ala Lys Ile Glu Lys Leu Leu Ala
785 790 795 800

Trp Leu Glu Thr Leu Lys Ala Glu Leu Gly Ile Pro Lys Ser Ile Arg
805 810 815

Glu Ala Gly Val Gln Glu Ala Asp Phe Leu Ala Asn Val Asp Lys Leu
820 825 830

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

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

Arg Asp Tyr Val Glu Gly Glu Thr Ala Ala Lys Lys Glu Ala Ala Pro
865 870 875 880

Ala Lys Ala Glu Lys Lys Ala Lys Lys Ser Ala
885 890

<210> 24
<211> 2676
<212> DNA
<213> Escherichia coli

<400> 24
atggcggtga cgaacgtcgc cgaactgaac gccctcgtcg aacgggtgaa aaaggcgcag 60
cgggagtacg cctcgttcac gcaggagcag gtggataaaa tcttccgcgc cgccgccctg 120
gccgcggccg acgcccgcac cccgctcgcc aaaatggcgg tggccgagtc cggcatgggc 180
atcgtggagg acaaggtcat caagaacat ttcgcctccg agtacatcta caacgcctac 240
aaggacgaga agacctgcgg cgtcctgtcg gaggacgata ccttcggcac catcacgac 300
gccgagccca tcggcatcat ctgcggcatc gtcccgacca ccaatccac ctccaccgcc 360
atcttcaaaa gcctcatctc cctgaaaacc cgcaatgcga tcattctcag ccccatccc 420
cgcgccaaag acgccacgaa caaggcggcc gacatcgtcc tccaggccgc catcgcggcc 480
ggcgccccca aagacctgat cggctggatc gaccagccgt cgggtggagct gtccaacgcc 540
ctcatgcac acccggacat caatctcatc ctggccacgg gcggcccggg catggtgaag 600
gccgcctact ccagcggcaa gccggcgatc ggcgtgggcg cgggcaacac gcccgtagtc 660
atcgacgaaa cggccgatat caaacgcgcc gtcgcctccg tgctcatgag caagaccttc 720
gacaacggcg tgatctgcgc gtcggaacag tcggtcgtgg tcgtcgactc cgtgtacgac 780
gcggtgcgcg agcgcttcgc cacgcatggc ggctacctgc tgcagggcaa ggagctgaaa 840
gccgtgcagg acgtgatcct caaaaacggc gccctgaacg ccgccatcgt cggccagccc 900
gcgtataaga tcgcggagct ggccggcttc tcgggtcccgg agaacacgaa gatcctgac 960
ggcgagggtca cggtcgtcga cgagagcga cccttcgccc acgagaagct gtccccgacc 1020
ctggcgatgt accgggcaaa agacttcgaa gacgcggtcg agaaggccga gaaactggtg 1080
gcgatgggcg gcatcggcca tacgagctgt ctgtacaccg accaggacaa tcagcccgcc 1140
cgcggtgtcct acttcggcca gaaaatgaag accgcccga tcctcatcaa tacgcccgcg 1200
agccagggcg gcatcggcga tctctacaat ttcaagctcg cgccgtcgct gaccctcggc 1260
tgcggctcct ggggaggcaa ctccatctcc gagaacgtgg gcccgaaaca cctcatcaac 1320
aagaagaccg tggccaagcg ggcggagaac atgctgtggc ataagctgcc caagtcgac 1380

tacttccgcc ggggctcgct gcccacgcg ctggacgaag tcatcacgga cgccacaag	1440
cgggcgctga tcgtgacgga ccgcttcctg ttcaataacg gctacgccga ccagatcacc	1500
tccgtcctga aggccgcggg cgtcgaaacg gaagtcttct tcgaagtcga agccgacccg	1560
accctctcca tcgtccggaa gggcgccgag ctggccaatt cttcaagcc ggatgtcatc	1620
atcgcgctgg gcggcggctc gccgatggac gccgcaaaa tcatgtgggt gatgtacgag	1680
cacccgaaa cgcacttcga agaactggcg ctgcgcttca tggatatccg caagcggatc	1740
tacaaattcc cgaaaatggg cgtgaaacg aagatgatcg cggtagcgac caccagcggc	1800
acgggctccg aggtgacgcc gttcgcggtc gtcaccgatg atgccaccgg ccagaaatac	1860
cccctggccg actacgccct gacccccgac atggccatcg tggacgcaa tctcgtgatg	1920
gacatgccga agagcctctg cgccttcggc ggcctggatg ccgtgacgca tgcgatggag	1980
gcgtatgtca gcgtgctggc gtccgagttc tcggacggcc aggcgctgca ggccctgaag	2040
ctgctgaagg agtacctgcc ggcctcctac cacgagggt ccaaaaacc ggtggcccgg	2100
gagcgggtcc actcggcggc caccatcgcg ggcacgcct tcgcgaacgc cttcctcggc	2160
gtgtgccatt cgatggccca caagctgggc tcccagttcc atatcccca cggcctggcg	2220
aacgcgctgc tgatctgtaa cgtcatccgg tacaacgcga acgataacc gaccaaacag	2280
accgccttct cccagtatga ccgccccag gcccggcgc gctatgcgga aatcgccgac	2340
catctcggcc tgtcggcccc gggcgaccgc accgccgcca agatcgagaa gctgctggcc	2400
tggctcgaag cgctgaaggc cgaactgggc atcccgaagt cgatccgca agccggcgtc	2460
caggaagcgg acttcctggc gaatgtggat aaactgagcg aggacgcctt cgatgaccag	2520
tgcacgggcg ccaaccgcg ctaccgctc atcagcgagc tgaaacagat cctgctcgac	2580
acctattatg gccgcgatta cgtcgagggc gaaaccgcc cgaagaagga ggccgccccg	2640
gccaaagccg aaaagaaggc caagaaaagc gcctga	2676

<210> 25
 <211> 387
 <212> PRT
 <213> Escherichia coli

<400> 25

Met	Asn	Asn	Phe	Asn	Leu	His	Thr	Pro	Thr	Arg	Ile	Leu	Phe	Gly	Lys
1				5					10					15	

Gly Ala Ile Ala Gly Leu Arg Glu Gln Ile Pro His Asp Ala Arg Val
20 25 30

Leu Ile Thr Tyr Gly Gly Gly Ser Val Lys Lys Thr Gly Val Leu Asp
35 40 45

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

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

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

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

Asn Ile Asp Pro Trp His Ile Leu Gln Thr Gly Gly Lys Glu Ile Lys
115 120 125

Ser Ala Ile Pro Met Gly Cys Val Leu Thr Leu Pro Ala Thr Gly Ser
130 135 140

Glu Ser Asn Ala Gly Ala Val Ile Ser Arg Lys Thr Thr Gly Asp Lys
145 150 155 160

Gln Ala Phe His Ser Ala His Val Gln Pro Val Phe Ala Val Leu Asp
165 170 175

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

Val Asp Ala Phe Val His Thr Val Glu Gln Tyr Val Thr Lys Pro Val
195 200 205

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

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

Arg Ala Asn Val Met Trp Gly Ala Thr Gln Ala Leu Asn Gly Leu Ile
245 250 255

Gly Ala Gly Val Pro Gln Asp Trp Ala Thr His Met Leu Gly His Glu
260 265 270

Leu Thr Ala Met His Gly Leu Asp His Ala Gln Thr Leu Ala Ile Val
275 280 285

Leu Pro Ala Leu Trp Asn Glu Lys Arg Glu Thr Lys Arg Ala Lys Leu
290 295 300

Leu Gln Tyr Ala Glu Arg Val Trp Asn Ile Thr Glu Gly Ser Asp Asp
305 310 315 320

Glu Arg Ile Asp Ala Ala Ile Ala Ala Thr Arg Asn Phe Phe Glu Gln
325 330 335

Leu Gly Val Pro Thr His Leu Ser Asp Tyr Gly Leu Asp Gly Ser Ser
340 345 350

Ile Pro Ala Leu Leu Lys Lys Leu Glu Glu His Gly Met Thr Gln Leu
355 360 365

Gly Glu Asn His Asp Ile Thr Leu Asp Val Ser Arg Arg Ile Tyr Glu
370 375 380

Ala Ala Arg
385

<210> 26
<211> 1164
<212> DNA
<213> Escherichia coli

<400> 26
atgaacaact ttaatctgca caccccaacc cgcattctgt ttggtaaagg cgcaatcgct 60
ggtttacgcg aacaaattcc tcacgatgct cgcgtattga ttacctacgg cggcggcagc 120

gtgaaaaaaaa ccggcgttct cgatcaagtt ctggatgccc tgaaaggcat ggacgtgctg	180
gaatttggcg gtattgagcc aaacccggct tatgaaacgc tgatgaacgc cgtgaaactg	240
gttcgcgaac agaaagtac tttcctgctg gcggttggcg gcggttctgt actggacggc	300
accaaattta tcgccgcagc ggctaactat ccggaaaata tcgatccgtg gcacattctg	360
caaacgggcg gtaaagagat taaaagcgcc atcccgatgg gctgtgtgct gacgctgcca	420
gcaaccgggtt cagaatccaa cgcaggcgcg gtgatctccc gtaaaaccac aggcgacaag	480
caggcgttcc attctgcccc tgttcagccg gtatttgccg tgctcgatcc ggtttatacc	540
tacaccctgc cgccgcgtca ggtggctaac ggcgtagtgg acgcctttgt acacaccgtg	600
gaacagtatg ttaccaaac ggttgatgcc aaaattcagg accgtttcgc agaaggcatt	660
ttgctgacgc taatcgaaga tgggccgaaa gccctgaaag agccagaaaa ctacgatgtg	720
cgcgccaacg tcatgtgggc ggcgactcag gcgctgaacg gtttgattgg cgctggcgta	780
ccgcaggact gggcaacgca tatgctgggc cacgaactga ctgcgatgca cggctctggat	840
cacgcgcaaa cactggctat cgtcctgcct gcactgtgga atgaaaaacg cgataccaag	900
cgcgctaagc tgctgcaata tgctgaacgc gtctggaaca tcactgaagg ttccgatgat	960
gagcgtattg acgccgcgat tgccgcaacc cgcaatttct ttgagcaatt aggcgtgccg	1020
acccacctct ccgactacgg tctggacggc agctccatcc cggctttgct gaaaaaactg	1080
gaagagcacg gcatgacca actgggcgaa aatcatgaca ttacgttgga tgtcagccgc	1140
cgtatatacg aagccgcccc ctaa	1164

<210> 27
 <211> 383
 <212> PRT
 <213> Escherichia coli

<400> 27

Met	Met	Ala	Asn	Arg	Met	Ile	Leu	Asn	Glu	Thr	Ala	Trp	Phe	Gly	Arg
1			5					10						15	

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

Lys	Ala	Leu	Ile	Val	Thr	Asp	Lys	Thr	Leu	Val	Gln	Cys	Gly	Val	Val
		35					40				45				

Ala Lys Val Thr Asp Lys Met Asp Ala Ala Gly Leu Ala Trp Ala Ile
50 55 60

Tyr Asp Gly Val Val Pro Asn Pro Thr Ile Thr Val Val Lys Glu Gly
65 70 75 80

Leu Asp Val Phe Gln Asn Ser Gly Ala Asp Tyr Leu Ile Ala Ile Gly
85 90 95

Gly Gly Ser Pro Gln Asp Thr Cys Lys Ala Ile Gly Ile Ile Ser Asn
100 105 110

Asn Pro Glu Phe Ala Asp Val Arg Ser Leu Glu Gly Leu Ser Pro Thr
115 120 125

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

Ala Ala Glu Val Thr Ile Asn Tyr Val Ile Thr Asp Glu Glu Lys Arg
145 150 155 160

Arg Lys Phe Val Cys Val Asp Pro His Asp Ile Pro Gln Val Ala Phe
165 170 175

Ile Asp Ala Asp Met Met Asp Gly Met Pro Pro Ala Leu Lys Ala Ala
180 185 190

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

Gly Ala Trp Ala Leu Thr Asp Ala Leu His Ile Lys Ala Ile Glu Ile
210 215 220

Ile Ala Gly Ala Leu Arg Gly Ser Val Ala Gly Asp Lys Asp Ala Gly
225 230 235 240

Glu Glu Met Ala Leu Gly Gln Tyr Val Ala Gly Met Gly Phe Ser Asn
245 250 255

Val Gly Leu Gly Leu Val His Gly Met Ala His Pro Leu Gly Ala Phe
260 265 270

Tyr Asn Thr Pro His Gly Val Ala Asn Ala Ile Leu Leu Pro His Val
275 280 285

Met Arg Tyr Asn Ala Asp Phe Thr Asp Glu Lys Tyr Arg Asp Ile Ala
290 295 300

Arg Val Met Gly Val Lys Val Glu Gly Met Ser Leu Glu Glu Ala Arg
305 310 315 320

Asn Ala Ala Val Glu Ala Val Phe Ala Leu Asn Arg Asp Val Gly Ile
325 330 335

Pro Pro His Leu Arg Asp Val Gly Val Arg Lys Glu Asp Ile Pro Ala
340 345 350

Leu Ala Gln Ala Ala Leu Asp Asp Val Cys Thr Gly Gly Asn Pro Arg
355 360 365

Glu Ala Thr Leu Glu Asp Ile Val Glu Leu Tyr His Thr Ala Trp
370 375 380

<210> 28
<211> 1149
<212> DNA
<213> Escherichia coli

<400> 28
atggcgaatc ggatgatacct caatgaaacg gcctgggttcg gccgcggcgc ggtcggcgcc 60
ctcaccgatg aggtcaagcg gcggggctac cagaaggccc tgatcgtcac ggataaaacc 120
ctggtgcagt gcggcgctcgt cgccaagggtg accgacaaga tggatgcggc cggcctggcc 180
tgggcgatct acgacggcgt ggtgcccac cccaccatca ccgtggtgaa ggaaggcctg 240
ggcgtgttcc agaactcggg cgcggattat ctcatcgca tcggcggcgg cagccccag 300
gacacctgca aggccatcgg catcatctcg aacaaccccg agttcgcgga cgtgcgtcc 360
ctggagggcc tgtcgccgac gaacaagccc tccgtcccga tcctcgccat cccgacgacg 420
gccggcaccg cggccgaggt gaccatcaat tacgtcatca ccgacgagga aaagcggcgc 480

aagttcgtgt gtgtggaccc ccatgacatc ccccaggtcg ccttcacga cgccgacatg	540
atggatggca tgccccccgc cctcaaggcc gcgacgggcg tggacgcgct gacgcatgcc	600
atcgaaggct acatcacccg gggcgcctgg gccctgacgg atgccctgca tatcaaggcc	660
atcgaaatca tcgccggcgc cctgcgcggc tccgtggccg gcgacaagga tgcgggcgag	720
gagatggcgc tgggccagta cgtggccggc atgggcttct ccaatgtggg cctgggcctg	780
gtgcatggca tggcccatcc gctcggcgcc ttctacaaca cgccgcatgg cgtcgcgaac	840
gcgatcctcc tgccgcatgt catgcgctac aatgcggact tcacgggcga gaaataccgc	900
gatatcgccc gggatcatgg cgtgaaggct gagggcatgt cgctggaaga ggcgcggaac	960
gccgcggctc aagccgtctt cgccctgaac cgggatgtgg gcatcccgcc gcacctgcgc	1020
gatgtcggcg tccgcaagga agacatcccc gcgctggcgc aggccgccct ggacgatgtg	1080
tgcaccggcg gcaacccccg cgaggcgacg ctggaagaca tcgtcgaact ctaccatacc	1140
gcgtggtga	1149

<210> 29
 <211> 860
 <212> PRT
 <213> Clostridium beijerinckii

<400> 29

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

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

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

Leu	Ala	Lys	Met	Ala	Val	Glu	Glu	Ser	Gly	Met	Gly	Ile	Val	Glu	Asp
	50					55					60				

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

Lys	Asp	Thr	Lys	Thr	Cys	Gly	Val	Ile	Glu	Arg	Asp	Glu	Met	Phe	Gly
				85						90					95

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

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

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

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

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

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

Gly Pro Gly Met Val Arg 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 Asp Thr Ala His
210 215 220

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

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

Tyr Asp Glu Val Arg Lys Glu Leu Asp Glu Arg Gly Ala Tyr Ile Leu
260 265 270

Lys Gly Asp Glu Val Asp Lys Val Arg Ser Ile Ile Leu Asp Pro Lys
275 280 285

Gly Ser Leu Asn Ser Glu Ile Val Gly Gln Ser Ala Tyr Lys Ile Ala
290 295 300

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

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

Ser Pro Ile Leu Gly Met Tyr Lys Ala Lys Thr Phe Asp Asp Ala Leu
340 345 350

Arg Leu Ala Ser Arg Met Ile Glu Leu Gly Gly Phe Gly His Thr Ser
355 360 365

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

Gly Val Ala Met Lys Thr Ala Arg Thr Leu Ile Asn Met Pro Ala Ser
385 390 395 400

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

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

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

Asn Met Leu Trp Phe Arg Val Pro Asp Lys Ile Tyr Phe Lys Phe Gly
450 455 460

Cys Leu Pro Ile Ala Leu Glu Glu Leu Asn Ala Met Lys Lys Lys Arg
465 470 475 480

Ala Phe Ile Val Thr Asp Arg Val Leu Phe Asp Leu Gly Tyr Thr His
485 490 495

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

Ser Asp Val Glu Pro Asp Pro Thr Leu Lys Ala Ala Lys Leu Gly Ala
515 520 525

Asp Ala Met Arg Asp Phe Asn Pro Asp Val Ile Ile Ala Ile Gly Gly
530 535 540

Gly Ser Pro Met Asp Ala Ala Lys Ile Met Trp Val Met Tyr Glu His
545 550 555 560

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

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

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

Val Ile Thr Asp Gln Gln Thr Gly Val Lys Tyr Pro Leu Ala Asp Tyr
610 615 620

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

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

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

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

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

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

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

Gly Ile Ala Asn Ala Leu Leu Ile Asp Glu Val Ile Arg Phe Asn Ala
740 745 750

Thr Asp Ala Pro Arg Lys Gln Ala Ala Phe Pro Gln Tyr Lys Tyr Pro
755 760 765

Asn Ala Gly Trp Arg Tyr Ala Arg Ile Ala Asp Tyr Leu Asn Leu Gly
770 775 780

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

Asp Leu Lys Val Lys Val Arg Ile Pro Lys Ser Ile Lys Glu Phe Gly
805 810 815

Val Ser Glu Glu Lys Phe Tyr Asp Ser Met Asp Glu Met Val Glu Gln
820 825 830

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

Ser Glu Ile Lys Glu Met Tyr Ile Lys Ser Tyr Asn
850 855 860

<210> 30
<211> 1407
<212> DNA
<213> Clostridium beijerinckii

<400> 30
atgaacaagg acacgtgat ccccaccacg aaagacctca aggtgaagac caacggcgag 60
aacatcaacc tgaaaaacta caaggacaac agcagctgct tcggcgtggt cgaaaacgtc 120
gaaaacgcca tcagctcggc ggtgcatgcc cagaaaatcc tcagcctgca ctatacgaag 180
gaacagcgcg aaaagatcat caccgaaatc cgcaaggccg cgctccagaa caaagaggtg 240
ctggccacga tgatcctgga ggaaacccat atgggcccgt acgaagataa aatcctgaag 300
cacgagctgg tggcgaaata cacgcccggc accgaagacc tgacgaccac cgcctggagc 360
ggcgataacg gcctgacggt ggtggagatg tcgccgtatg gcgtgatcgg cgcgatcacc 420
ccgtccacca acccgaccga aacggtcatc tgtaactcga tcggcatgat cgcggccggc 480

aacgcggtcg tgttcaatgg ccacccgtgc gccaaagaagt gcgtcgcctt cgccgtcgag 540
 atgatcaaca aggccatcat ctcttcgagg ggcccggaga acctcgtgac gacgatcaaa 600
 aacccgacga tggagtcctt ggatgccatc atcaagcacc cgtccatcaa gctgctctgc 660
 ggacccggcg gcccgggcat ggtcaaaacg ctgctgaact ccggcaaaaa agccatcggc 720
 gcggggcgcg gcaaccccc cgtcatcgtg gatgacaccg ccgacatcga gaaagccggc 780
 cggtcgatca tcgaaggctg ctcttcgac aataacctcc cgtgcatcgc cgagaaagag 840
 gtcttcgtct tcgagaacgt cgcggacgat ctgatctcca acatgctgaa gaacaacgcc 900
 gtcatcatca acgaagacca ggtgtccaag ctgatcgacc tcgtgctgca gaagaacaat 960
 gaaacccagg agtacttcat caacaagaag tgggtcggca aggacgcgaa gctgttcctg 1020
 gacgaaatcg atgtggagtc cccgagcaat gtgaagtga tcactctgca ggtcaacgcg 1080
 aaccaccctt tcgtcatgac cgagctcatg atgccgatcc tgccgatcgt gcgggtgaag 1140
 gatatcgacg aagccatcaa gtacgcgaaa atcgcggaac agaatcgga gcattcggcc 1200
 tatatctaca gcaagaacat cgataacctc aaccgcttcg aacgggagat cgacaccacc 1260
 atcttcgtca agaacgcaa atccttcgcc ggcgtgggct atgaggcgga aggcttcacc 1320
 accttcacga tcgccggctc caccggcgag ggcacacct ccgcccgga cttcaccgcg 1380
 cagcgccgct gcgtgctggc gggctga 1407

<210> 31
 <211> 360
 <212> PRT
 <213> *Saccharomyces cerevisiae*

<400> 31

Met Ser Tyr Pro Glu Lys Phe Glu Gly Ile Ala Ile Gln Ser His Glu
 1 5 10 15

Asp Trp Lys Asn Pro Lys Lys Thr Lys Tyr Asp Pro Lys Pro Phe Tyr
 20 25 30

Asp His Asp Ile Asp Ile Lys Ile Glu Ala Cys Gly Val Cys Gly Ser
 35 40 45

Asp Ile His Cys Ala Ala Gly His Trp Gly Asn Met Lys Met Pro Leu
 50 55 60

Val Val Gly His Glu Ile Val Gly Lys Val Val Lys Leu Gly Pro Lys
65 70 75 80

Ser Asn Ser Gly Leu Lys Val Gly Gln Arg Val Gly Val Gly Ala Gln
85 90 95

Val Phe Ser Cys Leu Glu Cys Asp Arg Cys Lys Asn Asp Asn Glu Pro
100 105 110

Tyr Cys Thr Lys Phe Val Thr Thr Tyr Ser Gln Pro Tyr Glu Asp Gly
115 120 125

Tyr Val Ser Gln Gly Gly Tyr Ala Asn Tyr Val Arg Val His Glu His
130 135 140

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

Leu Leu Cys Gly Gly Leu Thr Val Tyr Ser Pro Leu Val Arg Asn Gly
165 170 175

Cys Gly Pro Gly Lys Lys Val Gly Ile Val Gly Leu Gly Gly Ile Gly
180 185 190

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

Ile Ser Arg Ser Ser Arg Lys Arg Glu Asp Ala Met Lys Met Gly Ala
210 215 220

Asp His Tyr Ile Ala Thr Leu Glu Glu Gly Asp Trp Gly Glu Lys Tyr
225 230 235 240

Phe Asp Thr Phe Asp Leu Ile Val Val Cys Ala Ser Ser Leu Thr Asp
245 250 255

Ile Asp Phe Asn Ile Met Pro Lys Ala Met Lys Val Gly Gly Arg Ile
260 265 270

Val Ser Ile Ser Ile Pro Glu Gln His Glu Met Leu Ser Leu Lys Pro
 275 280 285

Tyr Gly Leu Lys Ala Val Ser Ile Ser Tyr Ser Ala Leu Gly Ser Ile
 290 295 300

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

Ile Trp Val Glu Thr Leu Pro Val Gly Glu Ala Gly Val His Glu Ala
 325 330 335

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

Gly Tyr Asp Lys Glu Phe Ser Asp
 355 360

<210> 32
 <211> 1083
 <212> DNA
 <213> *Saccharomyces cerevisiae*

<400> 32
 atgagctatc ccgagaagtt cgaggggatc gccatccaga gccacgagga ctggaagaac 60
 ccgaaaaaga ccaagtatga tccgaagccc ttctacgatc acgacatcga catcaagatc 120
 gaggcctgcg gcgtctgcgg cagcgatatc cattgtgcgg ctggccactg gggcaacatg 180
 aagatgccgt tggctgctcg ccacgagatc gtgggcaagg tcgtgaagtt aggcccgaaa 240
 agcaacagcg gcttgaaggt gggccagcgc gtgggtgtgg gtgcgcaggt cttcagctgt 300
 ctggagtgcg accgttgcaa gaacgacaac gaaccgtact gcaccaagtt cgtcaccacc 360
 tactcgcagc cctacgagga cggctacgtc tcgcagggcg gttacgcaa ctatgtccga 420
 gtccacgaac acttcgtggt gcccattccg gaaaatatcc ccagccatct ggcggctccc 480
 ctgctgtgcg gtggcttgac cgtctacagc cccctcgtcc gcaatggctg cgggtcccggc 540
 aagaaggtgg gtatcgtggg cctcggcggg ataggctcta tgggcacgct gatctcgaaa 600
 gcgatgggcg cagaaacgta cgtgatctcg cgttcctcgc gcaagcgcga ggatgcgatg 660
 aagatgggtg cggaccacta catcgccacg ctggaggagg gtgactgggg tgagaagtac 720

Glu Asn Thr Ala Thr Ile Ile Gly His Ala Gly Gly Phe Ala Asp Lys
115 120 125

Val Arg Ala Gly Trp Gln Trp Val Ile Pro Leu Pro Asp Glu Leu Asp
130 135 140

Pro Thr Ser Ala Gly Pro Leu Leu Cys Gly Gly Ile Thr Val Phe Asp
145 150 155 160

Pro Ile Leu Lys His Gln Ile Gln Ala Ile His His Val Ala Val Ile
165 170 175

Gly Ile Gly Gly Leu Gly His Met Ala Ile Lys Leu Leu Lys Ala Trp
180 185 190

Gly Cys Glu Ile Thr Ala Phe Ser Ser Asn Pro Asn Lys Thr Asp Glu
195 200 205

Leu Lys Ala Met Gly Ala Asp His Val Val Asn Ser Arg Asp Asp Ala
210 215 220

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

Asn Val Pro Leu Asn Trp Asn Ala Tyr Leu Asn Thr Leu Ala Pro Asn
245 250 255

Gly Thr Phe His Phe Leu Gly Val Val Met Glu Pro Ile Pro Val Pro
260 265 270

Val Gly Ala Leu Leu Gly Gly Ala Lys Ser Leu Thr Ala Ser Pro Thr
275 280 285

Gly Ser Pro Ala Ala Leu Arg Lys Leu Leu Glu Phe Ala Ala Arg Lys
290 295 300

Asn Ile Ala Pro Gln Ile Glu Met Tyr Pro Met Ser Glu Leu Asn Glu
305 310 315 320

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

Lys Ala Asp Phe
340

<210> 34
 <211> 1023
 <212> DNA
 <213> Acinetobacter baylyi

<400> 34
 atgacgacca acgtcatcca tgcctatgcg gccatgcagg ccggcgaggc gctggtgccg 60
 tatagcttcg acgcggggcga gctccagccg catcaggtcg aggtgaaggt ggaataactgc 120
 ggctgtgtc attccgatgt gtcggtcctg aataacgaat ggcactcctc cgtctatccg 180
 gtggtcgccg gccacgaggt catcggcacg atcaccagc tgggcagcga agccaagggc 240
 ctgaagatcg gccagcgcgt cggcatcggc tggacggccg agagctgcca ggcctgcgat 300
 cagtgcattc ccggccagca ggtgctgtgc accggcgaga acaccgccac catcatcggc 360
 catgccggcg gcttcgcgga taaagtgcgg gccggctggc agtgggtgat cccgctcccg 420
 gatgagctcg atcccacgag cgccggcccg ctgctgtgtg gcggcatcac cgtgttcgac 480
 ccgattctga agcaccagat ccaggccatc caccatgtcg ccgtgatcgg catcggcggc 540
 ctggggccaca tggcgatcaa gctgctcaag gcctggggct gtgaaatcac cgccttcagc 600
 agcaatccca acaagaccga cgaactgaag gcgatgggag ccgacctgtt ggtcaactcc 660
 cgcgatgacg cggaaatcaa gagccagcag ggcaagttcg acctgctgct gtcgacggtg 720
 aatgtccccc tcaactggaa cgcctacctg aataccctgg ccccgaaagg cacgttccac 780
 ttcctgggag tcgtcatgga gccgatcccg gtgccggtgg gcgccctcct gggcggcgcg 840
 aaaagcctga ccgcctcgcc cacgggcagc cccgccgccc tccgcaagct gctggagtgc 900
 gccgcccga agaacatcg gccccagatc gaaatgtatc ccatgagcga gctgaacgaa 960
 gccatcgagc gcctgcatag cggccaggcc cgctatcgga tcgtgctcaa agcggacttc 1020
 tga 1023

<210> 35
 <211> 339
 <212> PRT
 <213> Escherichia coli

<400> 35
 Met Ser Met Ile Lys Ser Tyr Ala Ala Lys Glu Ala Gly Gly Glu Leu
 1 5 10 15

Glu Val Tyr Glu Tyr Asp Pro Gly Glu Leu Arg Pro Gln Asp Val Glu
20 25 30

Val Gln Val Asp Tyr Cys Gly Ile Cys His Ser Asp Leu Ser Met Ile
35 40 45

Asp Asn Glu Trp Gly Phe Ser Gln Tyr Pro Leu Val Ala Gly His Glu
50 55 60

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

Leu Gln Val Gly Gln Arg Val Gly Ile Gly Trp Thr Ala Arg Ser Cys
85 90 95

Gly His Cys Asp Ala Cys Ile Ser Gly Asn Gln Ile Asn Cys Glu Gln
100 105 110

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

Arg Ala Asp Trp Gln Trp Val Ile Pro Leu Pro Glu Asn Ile Asp Ile
130 135 140

Glu Ser Ala Gly Pro Leu Leu Cys Gly Gly Ile Thr Val Phe Lys Pro
145 150 155 160

Leu Leu Met His His Ile Thr Ala Thr Ser Arg Val Gly Val Ile Gly
165 170 175

Ile Gly Gly Leu Gly His Ile Ala Ile Lys Leu Leu His Ala Met Gly
180 185 190

Cys Glu Val Thr Ala Phe Ser Ser Asn Pro Ala Lys Glu Gln Glu Val
195 200 205

Leu Ala Met Gly Ala Asp Lys Val Val Asn Ser Arg Asp Pro Gln Ala
210 215 220

Leu Lys Ala Leu Ala Gly Gln Phe Asp Leu Ile Ile Asn Thr Val Asn
 225 230 235 240

Val Ser Leu Asp Trp Gln Pro Tyr Phe Glu Ala Leu Thr Tyr Gly Gly
 245 250 255

Asn Phe His Thr Val Gly Ala Val Leu Thr Pro Leu Ser Val Pro Ala
 260 265 270

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

Thr Pro Tyr Glu Leu Arg Lys Leu Met Arg Phe Ala Ala Arg Ser Lys
 290 295 300

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

Ile Gln His Val Arg Asp Gly Lys Ala Arg Tyr Arg Val Val Leu Lys
 325 330 335

Ala Asp Phe

<210> 36
 <211> 1020
 <212> DNA
 <213> Escherichia coli

<400> 36
 atgtccatga tcaaaagcta cgccgcgaaa gaggcgggcg gcgagctgga ggtgtatgag 60
 tatgaccg ggcgagctgcg gccccaggac gtggaagtgc aggtcgacta ctgcggcatc 120
 tgccattcgg acctctcgat gatcgataac gaggggggct tcagccagta ccccttggtg 180
 gccggccacg aggtgatcgg ccgcgtggtc gccctgggct cggccgcgca ggataaaggc 240
 ctgcaggctg gccagcgcgt cggcatcggc tggacggccc gcagctgcgg ccattgcatc 300
 gcctgcatca gcggcaatca gatcaattgc gaacaggggc cgggtcccgac catcatgaac 360
 cggggcggct tcgccgaaaa gctgcggggc gattggcagt gggtgatccc gctgccggag 420
 aacatcgata tcgagtcggc cggccccctg ctgtgcggcg gcatcaccgt cttcaagccg 480

```

ctcctgatgc atcatatcac ggcgaccagc cgggtcggcg tgatcggcat cggcggcctc      540
ggccacatcg cgatcaaact gctgcacgcg atgggctgcg aggtcaccgc gttctcctcg      600
aaccccgcca aggagcagga agtgctggcg atgggcgccg ataaagtcgt gaactcgcgc      660
gacccccagg ccctcaaagc cctggccggc cagttcgatc tcatcatcaa cacggtgaac      720
gtgtcgctgg actggcagcc ctacttcgaa gccctgacct atggcggcaa cttcataacc      780
gtcggcgccg tgctgacccc gctgtccgtc ccggccttca ccctgatcgc cggcgaccgc      840
agcgtgtccg gcagcgccac cggcacgccg tatgagctgc gcaagctgat gcgcttcgcc      900
gcccgcagca aggtcgcccc gaccaccgag ctgttcccca tgtccaagat caatgacgcg      960
atccagcatg tccgggacgg caaggcccg c tgcgctcg tcctcaaggc ggacttctga     1020

```

```

<210> 37
<211> 388
<212> PRT
<213> Escherichia coli

```

```

<400> 37

```

```

Met Asn Leu His Glu Tyr Gln Ala Lys Gln Leu Phe Ala Arg Tyr Gly
1           5           10           15

```

```

Leu Pro Ala Pro Val Gly Tyr Ala Cys Thr Thr Pro Arg Glu Ala Glu
          20           25           30

```

```

Glu Ala Ala Ser Lys Ile Gly Ala Gly Pro Trp Val Val Lys Cys Gln
          35           40           45

```

```

Val His Ala Gly Gly Arg Gly Lys Ala Gly Gly Val Lys Val Val Asn
          50           55           60

```

```

Ser Lys Glu Asp Ile Arg Ala Phe Ala Glu Asn Trp Leu Gly Lys Arg
65           70           75           80

```

```

Leu Val Thr Tyr Gln Thr Asp Ala Asn Gly Gln Pro Val Asn Gln Ile
          85           90           95

```

```

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

```

Val Val Asp Arg Ser Ser Arg Arg Val Val Phe Met Ala Ser Thr Glu
115 120 125

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

His Lys Val Ala Leu Asp Pro Leu Thr Gly Pro Met Pro Tyr Gln Gly
145 150 155 160

Arg Glu Leu Ala Phe Lys Leu Gly Leu Glu Gly Lys Leu Val Gln Gln
165 170 175

Phe Thr Lys Ile Phe Met Gly Leu Ala Thr Ile Phe Leu Glu Arg Asp
180 185 190

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

Leu Ile Cys Leu Asp Gly Lys Leu Gly Ala Asp Gly Asn Ala Leu Phe
210 215 220

Arg Gln Pro Asp Leu Arg Glu Met Arg Asp Gln Ser Gln Glu Asp Pro
225 230 235 240

Arg Glu Ala Gln Ala Ala Gln Trp Glu Leu Asn Tyr Val Ala Leu Asp
245 250 255

Gly Asn Ile Gly Cys Met Val Asn Gly Ala Gly Leu Ala Met Gly Thr
260 265 270

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

Val Gly Gly Gly Ala Thr Lys Glu Arg Val Thr Glu Ala Phe Lys Ile
290 295 300

Ile Leu Ser Asp Asp Lys Val Lys Ala Val Leu Val Asn Ile Phe Gly
305 310 315 320

Gly Ile Val Arg Cys Asp Leu Ile Ala Asp Gly Ile Ile Gly Ala Val
325 330 335

Ala Glu Val Gly Val Asn Val Pro Val Val Val Arg Leu Glu Gly Asn
340 345 350

Asn Ala Glu Leu Gly Ala Lys Lys Leu Ala Asp Ser Gly Leu Asn Ile
355 360 365

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

Val Glu Gly Lys
385

<210> 38
<211> 1167
<212> DNA
<213> Escherichia coli

<400> 38
atgaacctgc atgaatatca ggccaagcag ctgttcgcgc ggtatggcct cccggcgccg 60
gtcgggtacg cctgtacgac cccgcgggaa gcggaggagg ccgcctccaa gatcggcgcc 120
ggcccgtggg tgggtcaaag ccaggtgcat gcgggcggcc ggggcaaggc gggcggcgtg 180
aaggtcgtca actccaagga ggacatccgc gccttcgccg agaactggct gggcaagcgg 240
ctggtgacct atcagacgga cgccaatggc cagcccgtca atcagatcct ggtcgaggcg 300
gccacggaca tcgcgaaaga actgtacctc ggcgccgtcg tggaccggag cagccggcgg 360
gtggtgttca tggcggtccac cgagggcggc gtggaaatcg aaaaagtggc cgaggaaacc 420
ccgcacctga tccataaagt cgcgctggac ccgctgaccg gccccatgcc gtatcagggc 480
cgggaactcg cgttcaagct cggcctggag ggcaagctgg tgcagcagtt cacgaaaatc 540
ttcatgggcc tggcgacct cttcctggag cgcgacctgg ccctgatcga aatcaaccgc 600
ctggtcatca cgaagcaggg cgacctgac tgcctggacg gcaagctcgg cgccgacggc 660
aacgccctgt tccgccagcc ggacctgcgg gaaatgcgcg atcagtcgca ggaggacccc 720
cgggaggccc aggcggccca gtgggagctg aattatgtgg cgctcgatgg caatatcggc 780
tgcattgtca atggcgcggg cctggcgatg ggcacgatgg acatcgtgaa gctgcatggc 840
ggcgagcccc ccaacttcct ggacgtgggc ggcggcgcga ccaaagagcg ggtgacggaa 900

gcgttcaaga tcatcctgag cgacgataaa gtcaaggccg tgctggtcaa catcttcggc 960
ggcatcgtcc gctgcgacct gatcgccgac ggcatcatcg gcgcggtggc ggaggtcggc 1020
gtcaatgtgc cggtcgtggt ccgcctggag ggcaacaacg ccgaactggg cgccaagaag 1080
ctggccgatt ccggcctgaa catcatcgcg gcgaagggcc tgaccgatgc cgcgcagcag 1140
gtcgtggccg ccgtcgaggg caagtga 1167

<210> 39
<211> 289
<212> PRT
<213> Escherichia coli

<400> 39

Met Ser Ile Leu Ile Asp Lys Asn Thr Lys Val Ile Cys Gln Gly Phe
1 5 10 15

Thr Gly Ser Gln Gly Thr Phe His Ser Glu Gln Ala Ile Ala Tyr Gly
20 25 30

Thr Lys Met Val Gly Gly Val Thr Pro Gly Lys Gly Gly Thr Thr His
35 40 45

Leu Gly Leu Pro Val Phe Asn Thr Val Arg Glu Ala Val Ala Ala Thr
50 55 60

Gly Ala Thr Ala Ser Val Ile Tyr Val Pro Ala Pro Phe Cys Lys Asp
65 70 75 80

Ser Ile Leu Glu Ala Ile Asp Ala Gly Ile Lys Leu Ile Ile Thr Ile
85 90 95

Thr Glu Gly Ile Pro Thr Leu Asp Met Leu Thr Val Lys Val Lys Leu
100 105 110

Asp Glu Ala Gly Val Arg Met Ile Gly Pro Asn Cys Pro Gly Val Ile
115 120 125

Thr Pro Gly Glu Cys Lys Ile Gly Ile Gln Pro Gly His Ile His Lys
130 135 140

Pro Gly Lys Val Gly Ile Val Ser Arg Ser Gly Thr Leu Thr Tyr Glu
 145 150 155 160

Ala Val Lys Gln Thr Thr Asp Tyr Gly Phe Gly Gln Ser Thr Cys Val
 165 170 175

Gly Ile Gly Gly Asp Pro Ile Pro Gly Ser Asn Phe Ile Asp Ile Leu
 180 185 190

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

Glu Ile Gly Gly Ser Ala Glu Glu Glu Ala Ala Ala Tyr Ile Lys Glu
 210 215 220

His Val Thr Lys Pro Val Val Gly Tyr Ile Ala Gly Val Thr Ala Pro
 225 230 235 240

Lys Gly Lys Arg Met Gly His Ala Gly Ala Ile Ile Ala Gly Gly Lys
 245 250 255

Gly Thr Ala Asp Glu Lys Phe Ala Ala Leu Glu Ala Ala Gly Val Lys
 260 265 270

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

Lys

<210> 40
 <211> 870
 <212> DNA
 <213> Escherichia coli

<400> 40
 atgtccatcc tcatcgacaa gaacaccaag gtgatctgcc agggcttcac cggctcccag 60
 ggcaccttcc acagcgagca ggcgatcgcc tacggcacca agatggtggg cggcgtcacc 120
 cccggcaagg gcggcaccac ccatctcggc ctcccggtgt tcaatacggg gcgggaagcc 180
 gtggcggcca ccggcgcgac cgccagcgtc atctatgtgc ccgcgccgtt ctgcaaggac 240

tccatcctgg aagccatcga cgccggcatc aagctcatca tcaccatcac cgagggcac	300
cccaccctcg acatgctgac ggtgaaagtc aagctggacg aggcgggctg cgggatgac	360
ggccccaact gcccgggcgt catcaccccc ggcgagtga aaatcggcac ccagccgggc	420
cacatccaca aaccgggcaa ggtgggcac gtctcgcgt ccggcacct cacctatgaa	480
gccgtcaagc agaccaccga ctatggcttc ggccagtcga cctgcgtcgg catcggcggc	540
gaccccatcc cgggctcga cttcatcgac atcctggaga tgttcgagaa agacccccag	600
accgaggcca tcgtgatgat cggcgagatc ggcggtcgg cggaggagga ggcggcggcg	660
tacatcaaag agcatgtgac caagccggtc gtgggctata tcgcgggctg gacggcgccc	720
aagggaagc gcatgggcca tgcgggcgcc atcatcgcg gcggcaaagg caccgcggat	780
gagaagtctg ccgccctgga ggccgccggc gtcaagaccg tccgcagcct cgccgatatc	840
ggcgaagccc tgaagaccgt gctcaagtga	870

<210> 41
 <211> 453
 <212> PRT
 <213> Clostridium kluyveri

<400> 41

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

Ala	Gln	Lys	Lys	Leu	Glu	Ala	Tyr	Ser	Gln	Glu	Gln	Val	Asp	Val	Leu
			20					25					30		

Val	Lys	Ala	Leu	Gly	Lys	Val	Val	Tyr	Asp	Asn	Ala	Glu	Met	Phe	Ala
		35					40					45			

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

Ala	Lys	Cys	His	Leu	Lys	Ser	Gly	Ala	Ile	Trp	Asn	His	Ile	Lys	Asp
65					70				75					80	

Lys	Lys	Thr	Val	Gly	Ile	Ile	Lys	Glu	Glu	Pro	Glu	Arg	Ala	Leu	Val
				85					90					95	

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

Asn Pro Val Val Thr Pro Met Cys Asn Ala Met Ala Ala Ile Lys Gly
115 120 125

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

Ala His Thr Val Glu Leu Met Asn Ala Glu Leu Lys Lys Leu Gly Ala
145 150 155 160

Pro Glu Asn Ile Ile Gln Ile Val Glu Ala Pro Ser Arg Glu Ala Ala
165 170 175

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

Gly Arg Val Lys Ala Ala Tyr Ser Ser Gly Arg Pro Ala Tyr Gly Val
195 200 205

Gly Pro Gly Asn Ser Gln Val Ile Val Asp Lys Gly Tyr Asp Tyr Asn
210 215 220

Lys Ala Ala Gln Asp Ile Ile Thr Gly Arg Lys Tyr Asp Asn Gly Ile
225 230 235 240

Ile Cys Ser Ser Glu Gln Ser Val Ile Ala Pro Ala Glu Asp Tyr Asp
245 250 255

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

Glu Glu Thr Val Glu Lys Phe Arg Ser Thr Leu Phe Lys Asp Gly Lys
275 280 285

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

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

Lys Gly Ala Gly Glu Lys Asp Val Leu Cys Lys Glu Lys Met Cys Pro
325 330 335

Val Leu Val Ala Leu Lys Tyr Asp Thr Phe Glu Glu Ala Val Glu Ile
340 345 350

Ala Met Ala Asn Tyr Met Tyr Glu Gly Ala Gly His Thr Ala Gly Ile
355 360 365

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

Ile Ser Arg Leu Val Val Asn Gln Pro Ala Thr Thr Ala Gly Gly Ser
385 390 395 400

Phe Asn Asn Gly Phe Asn Pro Thr Thr Thr Leu Gly Cys Gly Ser Trp
405 410 415

Gly Arg Asn Ser Ile Ser Glu Asn Leu Thr Tyr Glu His Leu Ile Asn
420 425 430

Val Ser Arg Ile Gly Tyr Phe Asn Lys Glu Ala Lys Val Pro Ser Tyr
435 440 445

Glu Glu Ile Trp Gly
450

<210> 42
<211> 1362
<212> DNA
<213> Clostridium kluyveri

<400> 42
atgtcgaatg aagtctcgat caaggaactg atcgaaaagg cgaaggtggc ccagaagaaa 60
ctggaagcct actcgcagga gcaggtggat gtcctgggtca aggccctggg caaggtcgtg 120
tatgataacg cggagatgtt cgccaaagaa gcggtggaag aaacggaaat gggcgtgtac 180
gaggacaagg tggcgaagtg ccacctgaaa tccggcgcca tctggaacca tatcaaggac 240
aagaaaaccg tgggcatcat caaggaagag ccggagcggg ccctggtcta cgtcgccaaag 300

```

cccaagggcg tggtagggcg caccacgccc atcaccaacc ccgtcgtcac gccgatgtgt      360
aatgcatggg ccgcatcaa gggccgcaat acgatcatcg tggccccgca ccccaaagcg      420
aaaaaggtgt ccgcccacac cgtcgagctg atgaacgccg agctgaagaa gctggggcgcc      480
ccggagaaca tcatccagat cgtggaagcg ccctcgcggg aggcggcgaa ggaactcatg      540
gagagcgccg acgtgggtcat cgccaccggc ggcgcgggcc gggtgaaggc cgcctactcg      600
tccggccgcc ccggtacagg cgtcggcccc ggcaactcgc aggtcatcgt cgacaagggc      660
tatgactaca acaaggccgc gcaggacatc atcaccggcc gcaagtatga caacggcatc      720
atctgctcct ccgagcagag cgtgatcgcg cccgccgaag actatgacaa ggtgatcgcc      780
gccttcgtcg agaattggcg gttctacgtg gaggatgaag aaaccgtgga aaagttccgc      840
tccaccctgt tcaaagacgg caagatcaat tccaagatca tcggcaagag cgtccagatc      900
atgcccgacc tggcgggctg caaggtcccc gagggcacga aggtcatcgt gctgaagggc      960
aagggcgcgg gcgagaagga tgtcctgtgc aaggaaaaga tgtgcccggg gctggtggcc     1020
ctgaagtacg acaccttcga agaggccgtc gaaatcgca tggccaacta tatgtatgaa     1080
ggcgccggcc acaccgccgg catccattcg gacaacgacg agaacatccg ctacgcgggc     1140
accgtgctcc ccatcagccg cctcgtggtg aaccagccgg ccaccaccgc cggcggcagc     1200
ttcaataacg gcttcaacc caccaccacc ctgggctgcg gcagctgggg ccgcaacagc     1260
atctcggaac acctcacgta cgaacatctc atcaacgtct cccgcatcgg ctattttaac     1320
aaggaagcga aggtgccctc ctatgaagag atctggggct ga                        1362

```

```

<210> 43
<211> 451
<212> PRT
<213> Porphyromonas gingivalis

```

```
<400> 43
```

```

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

```

```

Tyr Gln Ala Thr His Asn Gln Glu Ala Val Asp Asn Ile Cys Arg Ala
          20          25          30

```

```

Ala Ala Lys Val Ile Tyr Glu Asn Ala Ala Ile Leu Ala Arg Glu Ala
          35          40          45

```

Val Asp Glu Thr Gly Met Gly Val Tyr Glu His Lys Val Ala Lys Asn
50 55 60

Gln Gly Lys Ser Lys Gly Val Trp Tyr Asn Leu His Asn Lys Lys Ser
65 70 75 80

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

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

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

Ile Ile Ile Ala Pro His Pro Arg Ser Lys Lys Cys Ser Ala His Ala
130 135 140

Val Arg Leu Ile Lys Glu Ala Ile Ala Pro Phe Asn Val Pro Glu Gly
145 150 155 160

Met Val Gln Ile Ile Glu Glu Pro Ser Ile Glu Lys Thr Gln Glu Leu
165 170 175

Met Gly Ala Val Asp Val Val Val Ala Thr Gly Gly Met Gly Met Val
180 185 190

Lys Ser Ala Tyr Ser Ser Gly Lys Pro Ser Phe Gly Val Gly Ala Gly
195 200 205

Asn Val Gln Val Ile Val Asp Ser Asn Ile Asp Phe Glu Ala Ala Ala
210 215 220

Glu Lys Ile Ile Thr Gly Arg Ala Phe Asp Asn Gly Ile Ile Cys Ser
225 230 235 240

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

Thr Ala Phe Arg Asn His Gly Ala Tyr Phe Cys Asp Glu Ala Glu Gly
260 265 270

Asp Arg Ala Arg Ala Ala Ile Phe Glu Asn Gly Ala Ile Ala Lys Asp
275 280 285

Val Val Gly Gln Ser Val Ala Phe Ile Ala Lys Lys Ala Asn Ile Asn
290 295 300

Ile Pro Glu Gly Thr Arg Ile Leu Val Val Glu Ala Arg Gly Val Gly
305 310 315 320

Ala Glu Asp Val Ile Cys Lys Glu Lys Met Cys Pro Val Met Cys Ala
325 330 335

Leu Ser Tyr Lys His Phe Glu Glu Gly Val Glu Ile Ala Arg Thr Asn
340 345 350

Leu Ala Asn Glu Gly Asn Gly His Thr Cys Ala Ile His Ser Asn Asn
355 360 365

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

Val Val Asn Ala Pro Ser Ala Thr Thr Ala Gly Gly His Ile Gln Asn
385 390 395 400

Gly Leu Ala Val Thr Asn Thr Leu Gly Cys Gly Ser Trp Gly Asn Asn
405 410 415

Ser Ile Ser Glu Asn Phe Thr Tyr Lys His Leu Leu Asn Ile Ser Arg
420 425 430

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

Trp Glu Leu
450

<210> 44

<211> 1356

<212> DNA

<213> Porphyromonas gingivalis

<400> 44

atggagatca aggaaatggt gtccctggcc cgcaaggccc agaaggagta ccaggccacc	60
cataatcagg aagccgtgga caacatctgc cgcgccgcgg cgaaagtcac ctacgaaaat	120
gccgccatcc tcgcccgcga ggcggtcgat gaaaccggca tgggcgtgta cgaacacaag	180
gtggccaaga atcagggcaa aagcaagggc gtgtggtaca acctgcataa caagaaaagc	240
atcggcatcc tgaacatcga cgagcgcacc ggcatgatcg agatcgccaa gcccatcggc	300
gtcgtcggcg cggtcacgcc gaccaccaat cccatcgtca ccccgatgtc caacatcatc	360
ttcgcgctga aaacctgcaa cgcgatcatc atcgcccccc acccgcgag caaaaaatgc	420
tcggcccacg ccgtgcggct catcaaggaa gccatcgcgc cgttcaacgt gccggagggc	480
atggtgcaga tcatcgaaga gccgtcgatc gagaagacc aggaactgat gggcgccgtc	540
gacgtggtgg tggccaccgg cggcatgggc atggtgaagt cggcgatatag ctcgggcaag	600
ccgagcttcg gcgtgggcgc cggcaacgtg caggtcatcg tcgattcgaa catcgatttc	660
gaagccgccg cggaaaagat catcacgggc cgcgcgttcg ataacggcat catctgctcc	720
ggcgaacaga gcatcatcta caacgaggcg gataaggagg ccgtcttcac ggccttcagg	780
aaccacggcg cgtacttctg cgatgaagcc gaaggcgatc gcgcccgggc ggcgatcttc	840
gaaaacggcg ccatcgccaa agacgtggtg ggccagagcg tggcgttcat cgccaagaaa	900
gcgaacatca atatccccga aggcacgcgc atcctggtgg tggaagcccg cggcgtcggc	960
gcggaagatg tgatctgcaa ggaaaagatg tgccccgtga tgtgcgccct gtcctacaaa	1020
cacttcgaag agggcgtcga gatcgccgc accaacctcg cgaacgaggg caacggccac	1080
acctgcgcca tccattcgaa taatcaggcc cacatcatcc tggccggcag cgagctcacc	1140
gtcagccgca tcgtggtcaa cgcgccctcc gccaccaccg ccggcggccca tatccagaac	1200
ggcctggccg tgacgaacac cctgggctgc ggctcgtggg gcaacaacag catctcggaa	1260
aacttcacct ataagcatct gctcaatatc agccgcatcg ccccgctgaa cagcagcatc	1320
catatcccg acgacaaaga gatctgggaa ctctga	1356

<210> 45

<211> 469

<212> PRT

<213> Methylococcus capsulatus

<400> 45

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

Arg Val Pro Val Ser Ala Leu Tyr Ala Ala Gln Thr Gln Arg Ala Ile
20 25 30

Asp Asn Phe Pro Val Ser Gly Leu Ala Leu Pro Pro Ala Phe Ile Arg
35 40 45

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

Gly His Leu Asp Ala Gly Lys Gly Glu Ala Ile Met Ala Ala Ala Glu
65 70 75 80

Glu Ile Ile Gly Gly Ala Tyr Ala Asp Gln Phe Pro Val Asp Val Phe
85 90 95

Gln Thr Gly Ser Gly Thr Ser Thr Asn Met Asn Val Asn Glu Val Leu
100 105 110

Ala Thr Leu Ala Ser Arg Arg Ala Gly Thr Pro Val Ser Ala Asn Asp
115 120 125

Asp Val Asn Met Gly Gln Ser Ser Asn Asp Val Ile Pro Thr Ala Ile
130 135 140

His Val Ser Ala Ala Leu Ala Val Asn Glu His Leu Ile Pro Ala Leu
145 150 155 160

Glu His Leu Ala Ala Thr Ile Arg Tyr Lys Gly Leu Ala His Pro Lys
165 170 175

Ala Val Lys Thr Gly Arg Thr His Leu Met Asp Ala Met Pro Val Arg
180 185 190

Leu Glu Gln Glu Leu Ser Gly Trp Ala Leu Gln Val Ala Asn Gly Val
195 200 205

Asp Arg Leu Arg Ala Ser Leu Pro Arg Leu Tyr Lys Leu Gly Gln Gly
210 215 220

Gly Thr Ala Val Gly Thr Gly Ile Asn Ala Asp Pro Ala Phe Ala Ser
225 230 235 240

Ser Phe Ala Glu Ala Leu Ala Asp Ala Thr Gly Leu Pro Phe Arg Pro
245 250 255

Asn Asp Ser Phe Phe Glu Ser Leu Ser Cys Gln Asp Ala Ala Val Glu
260 265 270

Leu Ser Gly Gln Leu Lys Thr Ile Ala Val Gly Ile Met Lys Ile Ala
275 280 285

Asn Asp Leu Arg Trp Met Asn Ser Gly Pro Leu Ala Gly Leu Gly Glu
290 295 300

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

Val Asn Pro Val Ile Pro Glu Ala Ala Cys Met Val Ala Ala Gln Val
325 330 335

Ile Gly Asn Asp Ala Ala Ile Thr Val Ala Gly Gln Ser Gly Ser Phe
340 345 350

Gln Leu Asn Val Met Leu Pro Val Ile Ala Tyr Asn Leu Leu Gln Ser
355 360 365

Ile Glu Leu Leu Ala Asn Ile Ala Arg Leu Leu Ala Asp Lys Ala Ile
370 375 380

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

Asn Pro Ile Leu Val Thr Ala Leu Asn Pro Val Ile Gly Tyr Leu Lys
405 410 415

Ala Ala Glu Ile Ala Lys Thr Ala Tyr Arg Thr Gly Arg Pro Ile Leu
420 425 430

Glu Val Ala Ile Glu Met Ser Gly Leu Asp Arg Ala Glu Leu Glu Arg
435 440 445

Leu Leu Asp Pro Ala Ala Leu Thr Thr Gly Gly Ile His Gly Val Pro
450 455 460

Ala Gly Ala Ala Gly
465

<210> 46
<211> 1410
<212> DNA
<213> *Methylococcus capsulatus*

<400> 46
atgagccaag acgactatcg catcgaaaaa gacagcatgg gtgagctgcg ggttccggtt 60
tccgcgctct acgcggctca gaccagcgt gccatcgaca attttccggt gagcgggctg 120
gctttgccgc cggcgttcat ccgcgcgatc gcccggatca agcagtgcgc ggcccgggtc 180
aatgtggcgc tggggcatct ggatgcaggc aagggcgagg cgatcatggc ggcggccgag 240
gagatcatcg gcggggcata cgccgaccag tttcctgtcg acgtgttcca gaccggttcc 300
ggcaccagca ccaacatgaa cgtcaacgaa gtgctggcga ctttggcttc ccggcgcgcg 360
ggaacgcccg tgagcgccaa cgacgacgtc aacatggggc agagcagcaa tgacgtcatc 420
cccacggcga ttcacgtcag tgccgcgctg gcagtcaacg aacatctgat cccggcgctg 480
gaacacctgg cggcgacgat acgtacaag ggcttggcgc atccgaaggc ggtcaagacc 540
gggcgggacc atctgatgga cgccatgcca gtgcggctgg agcaagaact ctcgggctgg 600
gcgctgcagg tcgccaacgg cgtcgaccgc ctgcgggcga gtctgccgcg gctttacaaa 660
ttggggcagg gcggcacggc cgtgggcact ggcatcaacg ccgatccggc cttcgcttcc 720
tccttcgccg aagctctggc cgatgccacg ggcctgcctt ttcgccccaa cgattcgttc 780
ttcgagagcc tgagttgcca ggacgcggca gtcgaactgt cgggccagct caagaccatc 840
gccgtcggca tcatgaagat cgccaacgac ctgcgctgga tgaattcggg tcccctggct 900
ggcctggggg agatcgagct gcccgcgctc cagcccggca gtcctcatcat gcccggaag 960

gtgaatccgg tgattcccga agccgcctgc atggtcgcgg cgcaggtcat cggcaacgac	1020
gccgccatca cggtaggctgg tcagtcgggt tctttccaac tcaacgtgat gctgccggtg	1080
atcgcctaca acttgctgca gagcatcgag ttgctggcca acattgctcg tctgctggcc	1140
gacaaggcca tcgccgactt ccgggtgcgt gaggacaatc tgcgccgggc tctggcaacg	1200
aatccgatcc tggtcaccgc actcaacccg gtgatcggct acctgaaggc ggcggaaatc	1260
gccaagaccg cctaccggac ggggcggccg attctggagg tggccatcga aatgagcgg	1320
ctggaccgcg ccgaactgga gcgactgctc gatccggcgg ccctgaccac cggcggcatc	1380
cacggtgtac ccgccggcgc ggcgggctga	1410

<210> 47
 <211> 467
 <212> PRT
 <213> Escherichia coli

<400> 47

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

Pro	Ala	Asp	Lys	Leu	Trp	Gly	Ala	Gln	Thr	Gln	Arg	Ser	Leu	Glu	His
			20					25					30		

Phe	Arg	Ile	Ser	Thr	Glu	Lys	Met	Pro	Thr	Ser	Leu	Ile	His	Ala	Leu
		35					40					45			

Ala	Leu	Thr	Lys	Arg	Ala	Ala	Ala	Lys	Val	Asn	Glu	Asp	Leu	Gly	Leu
	50					55					60				

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

Leu	Ala	Gly	Gln	His	Asp	Asp	Glu	Phe	Pro	Leu	Ala	Ile	Trp	Gln	Thr
				85					90					95	

Gly	Ser	Gly	Thr	Gln	Ser	Asn	Met	Asn	Met	Asn	Glu	Val	Leu	Ala	Asn
			100					105					110		

Arg	Ala	Ser	Glu	Leu	Leu	Gly	Gly	Val	Arg	Gly	Met	Glu	Arg	Lys	Val
		115					120					125			

His Pro Asn Asp Asp Val Asn Lys Ser Gln Ser Ser Asn Asp Val Phe
130 135 140

Pro Thr Ala Met His Val Ala Ala Leu Leu Ala Leu Arg Lys Gln Leu
145 150 155 160

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

Ala Phe Ala Asp Ile Val Lys Ile Gly Arg Thr His Leu Gln Asp Ala
180 185 190

Thr Pro Leu Thr Leu Gly Gln Glu Ile Ser Gly Trp Val Ala Met Leu
195 200 205

Glu His Asn Leu Lys His Ile Glu Tyr Ser Leu Pro His Val Ala Glu
210 215 220

Leu Ala Leu Gly Gly Thr Ala Val Gly Thr Gly Leu Asn Thr His Pro
225 230 235 240

Glu Tyr Ala Arg Arg Val Ala Asp Glu Leu Ala Val Ile Thr Cys Ala
245 250 255

Pro Phe Val Thr Ala Pro Asn Lys Phe Glu Ala Leu Ala Thr Cys Asp
260 265 270

Ala Leu Val Gln Ala His Gly Ala Leu Lys Gly Leu Ala Ala Ser Leu
275 280 285

Met Lys Ile Ala Asn Asp Val Arg Trp Leu Ala Ser Gly Pro Arg Cys
290 295 300

Gly Ile Gly Glu Ile Ser Ile Pro Glu Asn Glu Pro Gly Ser Ser Ile
305 310 315 320

Met Pro Gly Lys Val Asn Pro Thr Gln Cys Glu Ala Leu Thr Met Leu
325 330 335

Cys Cys Gln Val Met Gly Asn Asp Val Ala Ile Asn Met Gly Gly Ala
340 345 350

Ser Gly Asn Phe Glu Leu Asn Val Phe Arg Pro Met Val Ile His Asn
355 360 365

Phe Leu Gln Ser Val Arg Leu Leu Ala Asp Gly Met Glu Ser Phe Asn
370 375 380

Lys His Cys Ala Val Gly Ile Glu Pro Asn Arg Glu Arg Ile Asn Gln
385 390 395 400

Leu Leu Asn Glu Ser Leu Met Leu Val Thr Ala Leu Asn Thr His Ile
405 410 415

Gly Tyr Asp Lys Ala Ala Glu Ile Ala Lys Lys Ala His Lys Glu Gly
420 425 430

Leu Thr Leu Lys Ala Ala Ala Leu Ala Leu Gly Tyr Leu Ser Glu Ala
435 440 445

Glu Phe Asp Ser Trp Val Arg Pro Glu Gln Met Val Gly Ser Met Lys
450 455 460

Ala Gly Arg
465

<210> 48
<211> 1404
<212> DNA
<213> Escherichia coli

<400> 48
atgaacaccg tgcggagcga aaaggacagc atgggcgcca tcgacgtgcc cgccgataag 60
ctgtggggcg cccagaccca gcggtccctc gaacacttcc ggatctcgac cgaaaagatg 120
ccgaccagcc tgatccatgc gctcgccctg acgaaacggg ccgccgcaa ggtcaacgag 180
gacctcggcc tgctgagcga agagaaggcc tccgcgatcc gccaggccgc cgatgaagtc 240
ctcgccggcc agcacgacga cgagttcccg ctggccatct ggcagaccgg ctcgggcacc 300
cagtcgaaca tgaacatgaa cgaagtgtc gccaacgcg cgtcggaact gctgggcggc 360

```

gtccggggca tggaacgcaa ggtccatccc aacgacgacg tgaacaagtc gcagagctcc      420
aatgatgtct tccccaccgc catgcatgtc gcggccctcc tggccctgcg caaacagctg      480
atcccccagc tgaaaaccct gaccagacc ctgaatgaaa aatcccgcgc gttcgccgac      540
atcgtgaaga tcggccggac gcacctccag gacgcgaccc ccctgaccct cggccaggag      600
atctccggct gggctggcgat gctggaacac aacctgaaac atatcgagta cagcctgccg      660
cacgtcgcgg aactggcgct gggcggcacg gcggctggca ccggcctgaa cacgcatccg      720
gagtatgcgc ggcgcgtggc cgacgaactg gccgtcatca cctgtgcccc gttcgtgacc      780
gccccgaata aattcgaagc cctggcgacc tgtgacgccc tgggtccaggc gcatggcgcc      840
ctcaagggcc tggccgccag cctcatgaaa atcgcgaaatg acgtgcgctg gctggcctcc      900
ggcccccgct gcggcatcgg cgaaatctcg atccccgaaa acgagccggg ctcgtcgatc      960
atgccgggca aggtgaacc gaccagtg caggcgctca cgatgctctg ctgccaggtc     1020
atgggcaatg acgtggccat caatatgggc ggcgcgtccg gcaacttcga actgaatgtc     1080
ttccgcccga tggatcatcca caatttcctg cagtccgtgc gcctgctggc ggacggcatg     1140
gagagcttca ataagcactg cgcggtgggc atcgagccca accgcgagcg catcaaccag     1200
ctgctgaatg agagcctgat gctggtcacc gcgctgaata cccacatcgg ctacgacaag     1260
gcggccgaaa tcgccaagaa agcgcacaaa gaaggcctga cgctgaaagc cgccgccctg     1320
gccctgggct acctgtccga ggccgagttc gacagctggg tccgcccgga gcagatgggtg     1380
ggctcgatga aagccggccg gtga                                           1404

```

```

<210> 49
<211> 474
<212> PRT
<213> Mycobacterium tuberculosis

```

```
<400> 49
```

```

Met Ala Val Asp Ala Asp Ser Ala Asn Tyr Arg Ile Glu His Asp Thr
1           5           10          15

```

```

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

```

```

Gln Arg Ala Val Glu Asn Phe Pro Ile Ser Gly Arg Gly Leu Glu Arg
          35          40          45

```

Thr Gln Ile Arg Ala Leu Gly Leu Leu Lys Gly Ala Cys Ala Gln Val
50 55 60

Asn Ser Asp Leu Gly Leu Leu Ala Pro Glu Lys Ala Asp Ala Ile Ile
65 70 75 80

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

Ile Asp Val Phe Gln Thr Gly Ser Gly Thr Ser Ser Asn Met Asn Thr
100 105 110

Asn Glu Val Ile Ala Ser Ile Ala Ala Lys Gly Gly Val Thr Leu His
115 120 125

Pro Asn Asp Asp Val Asn Met Ser Gln Ser Ser Asn Asp Thr Phe Pro
130 135 140

Thr Ala Thr His Ile Ala Ala Thr Glu Ala Ala Val Ala His Leu Ile
145 150 155 160

Pro Ala Leu Gln Gln Leu His Asp Ala Leu Ala Ala Lys Ala Leu Asp
165 170 175

Trp His Thr Val Val Lys Ser Gly Arg Thr His Leu Met Asp Ala Val
180 185 190

Pro Val Thr Leu Gly Gln Glu Phe Ser Gly Tyr Ala Arg Gln Ile Glu
195 200 205

Ala Gly Ile Glu Arg Val Arg Ala Cys Leu Pro Arg Leu Gly Glu Leu
210 215 220

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

Phe Gly Val Arg Val Val Ala Val Leu Val Ala Gln Thr Gly Leu Ser
245 250 255

Glu Leu Arg Thr Ala Ala Asn Ser Phe Glu Ala Gln Ala Ala Arg Asp
260 265 270

Gly Leu Val Glu Ala Ser Gly Ala Leu Arg Thr Ile Ala Val Ser Leu
275 280 285

Thr Lys Ile Ala Asn Asp Ile Arg Trp Met Gly Ser Gly Pro Leu Thr
290 295 300

Gly Leu Ala Glu Ile Gln Leu Pro Asp Leu Gln Pro Gly Ser Ser Ile
305 310 315 320

Met Pro Gly Lys Val Asn Pro Val Leu Pro Glu Ala Val Thr Gln Val
325 330 335

Ala Ala Gln Val Ile Gly Asn Asp Ala Ala Ile Ala Trp Gly Gly Ala
340 345 350

Asn Gly Ala Phe Glu Leu Asn Val Tyr Ile Pro Met Met Ala Arg Asn
355 360 365

Ile Leu Glu Ser Phe Lys Leu Leu Thr Asn Val Ser Arg Leu Phe Ala
370 375 380

Gln Arg Cys Ile Ala Gly Leu Thr Ala Asn Val Glu His Leu Arg Arg
385 390 395 400

Leu Ala Glu Ser Ser Pro Ser Ile Val Thr Pro Leu Asn Ser Ala Ile
405 410 415

Gly Tyr Glu Glu Ala Ala Ala Val Ala Lys Gln Ala Leu Lys Glu Arg
420 425 430

Lys Thr Ile Arg Gln Thr Val Ile Asp Arg Gly Leu Ile Gly Asp Arg
435 440 445

Leu Ser Ile Glu Asp Leu Asp Arg Arg Leu Asp Val Leu Ala Met Ala
450 455 460

Lys Ala Glu Gln Leu Asp Ser Asp Arg Leu
465 470

<210> 50
<211> 1425
<212> DNA
<213> Mycobacterium tuberculosis

<400> 50
atggcggtgg atgcggacag cgcgaactac cgcacatgaac acgatacgaat gggcgaagtg 60
cgcgtccccg ccaaggccct ctggcggggc cagacgcagc gggccgtcga gaacttcccc 120
atcagcggcc ggggcctgga acgcacgcag atccgcgccc tgggcctcct gaaggcgcc 180
tgcgcccagg tgaactccga tctgggcctg ctgcgcgccc aaaaggcgga tgccatcatc 240
gccgccgcgg cggagatcgc cgacggccag cagcagatc agttcccgat cgatgtcttc 300
cagaccggca gcggcacctc cagcaatatg aacaccaacg aggtcatcgc gagcatcgcg 360
gccaagggcg gcgtgaccct gcacccgaac gacgacgtga atatgagcca gtcgtccaac 420
gacaccttcc ccaccgccac ccatatcgcg gcgaccgagg cggccgtcgc gcatctgac 480
cccgccctgc agcagctcca cgatgccctg gccgcgaaag cgctggactg gcacaccgtc 540
gtgaaatccg gccggacgca tctcatggac gccgtgccgg tgaccctggg ccaggagttc 600
tccggctacg cccggcagat cgaagccggc atcgaacggg tccgcgcgtg cctcccgagg 660
ctgggcgagc tggccatcgg cggcaccgcc gtgggcaccg gcctcaacgc gcccagcagat 720
ttcggcgtgc ggggtgggtggc ggtgctcgtg gccagaccg gcctgtcggg actccggacg 780
gcggccaatt ccttcgaagc ccaggccgcc cgcgatggcc tgggtggaggc gtcgggcgcc 840
ctgcgcacga tcgccgtgag cctgacgaag atcgcgaacg acatccggtg gatgggctcc 900
ggcccgtca ccggcctggc ggaaatccag ctcccggacc tgacagcccgg cagctccatc 960
atgcccggca aggtcaaccc cgtgctgccg gaagcgggtc cccaggtcgc ggcccaggtg 1020
atcggcaacg acgccgcat cgcctggggc ggcgcgaacg gcgcgttcga actgaatgtc 1080
tatatcccga tgatggcccc gaacatcctg gagtcgttca agctgctcac caacgtgagc 1140
cggctgttcg cgcagcgtg catcgccggc ctcaccgcca acgtcgaaca tctgcggcgg 1200
ctcgcggaat cctcgcccag catcgtcacg ccgctcaaca gcgcgatcgg ctacgaagaa 1260
gcggcggccg tggccaagca ggcgctcaaa gagcgcaaga ccatccgcca gacggtgac 1320
gaccggggcc tgatcggcga tcggctgtcg atcaggacc tggatcgccg cctggatgtg 1380

ctcgcgatgg ccaaagccga acagctggac tccgaccgcc tgtga

1425

<210> 51

<211> 602

<212> PRT

<213> Escherichia coli

<400> 51

Met Gln Thr Phe Gln Ala Asp Leu Ala Ile Val Gly Ala Gly Gly Ala
1 5 10 15

Gly Leu Arg Ala Ala Ile Ala Ala Ala Gln Ala Asn Pro Asn Ala Lys
20 25 30

Ile Ala Leu Ile Ser Lys Val Tyr Pro Met Arg Ser His Thr Val Ala
35 40 45

Ala Glu Gly Gly Ser Ala Ala Val Ala Gln Asp His Asp Ser Phe Glu
50 55 60

Tyr His Phe His Asp Thr Val Ala Gly Gly Asp Trp Leu Cys Glu Gln
65 70 75 80

Asp Val Val Asp Tyr Phe Val His His Cys Pro Thr Glu Met Thr Gln
85 90 95

Leu Glu Leu Trp Gly Cys Pro Trp Ser Arg Arg Pro Asp Gly Ser Val
100 105 110

Asn Val Arg Arg Phe Gly Gly Met Lys Ile Glu Arg Thr Trp Phe Ala
115 120 125

Ala Asp Lys Thr Gly Phe His Met Leu His Thr Leu Phe Gln Thr Ser
130 135 140

Leu Gln Phe Pro Gln Ile Gln Arg Phe Asp Glu His Phe Val Leu Asp
145 150 155 160

Ile Leu Val Asp Asp Gly His Val Arg Gly Leu Val Ala Met Asn Met
165 170 175

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

Thr Gly Gly Ala Gly Arg Val Tyr Arg Tyr Asn Thr Asn Gly Gly Ile
195 200 205

Val Thr Gly Asp Gly Met Gly Met Ala Leu Ser His Gly Val Pro Leu
210 215 220

Arg Asp Met Glu Phe Val Gln Tyr His Pro Thr Gly Leu Pro Gly Ser
225 230 235 240

Gly Ile Leu Met Thr Glu Gly Cys Arg Gly Glu Gly Gly Ile Leu Val
245 250 255

Asn Lys Asn Gly Tyr Arg Tyr Leu Gln Asp Tyr Gly Met Gly Pro Glu
260 265 270

Thr Pro Leu Gly Glu Pro Lys Asn Lys Tyr Met Glu Leu Gly Pro Arg
275 280 285

Asp Lys Val Ser Gln Ala Phe Trp His Glu Trp Arg Lys Gly Asn Thr
290 295 300

Ile Ser Thr Pro Arg Gly Asp Val Val Tyr Leu Asp Leu Arg His Leu
305 310 315 320

Gly Glu Lys Lys Leu His Glu Arg Leu Pro Phe Ile Cys Glu Leu Ala
325 330 335

Lys Ala Tyr Val Gly Val Asp Pro Val Lys Glu Pro Ile Pro Val Arg
340 345 350

Pro Thr Ala His Tyr Thr Met Gly Gly Ile Glu Thr Asp Gln Asn Cys
355 360 365

Glu Thr Arg Ile Lys Gly Leu Phe Ala Val Gly Glu Cys Ser Ser Val
370 375 380

Gly Leu His Gly Ala Asn Arg Leu Gly Ser Asn Ser Leu Ala Glu Leu
385 390 395 400

Val Val Phe Gly Arg Leu Ala Gly Glu Gln Ala Thr Glu Arg Ala Ala
405 410 415

Thr Ala Gly Asn Gly Asn Glu Ala Ala Ile Glu Ala Gln Ala Ala Gly
420 425 430

Val Glu Gln Arg Leu Lys Asp Leu Val Asn Gln Asp Gly Gly Glu Asn
435 440 445

Trp Ala Lys Ile Arg Asp Glu Met Gly Leu Ala Met Glu Glu Gly Cys
450 455 460

Gly Ile Tyr Arg Thr Pro Glu Leu Met Gln Lys Thr Ile Asp Lys Leu
465 470 475 480

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

Ser Val Phe Asn Thr Asp Leu Leu Tyr Thr Ile Glu Leu Gly His Gly
500 505 510

Leu Asn Val Ala Glu Cys Met Ala His Ser Ala Met Ala Arg Lys Glu
515 520 525

Ser Arg Gly Ala His Gln Arg Leu Asp Glu Gly Cys Thr Glu Arg Asp
530 535 540

Asp Val Asn Phe Leu Lys His Thr Leu Ala Phe Arg Asp Ala Asp Gly
545 550 555 560

Thr Thr Arg Leu Glu Tyr Ser Asp Val Lys Ile Thr Thr Leu Pro Pro
565 570 575

Ala Lys Arg Val Tyr Gly Gly Glu Ala Asp Ala Ala Asp Lys Ala Glu
580 585 590

Ala Ala Asn Lys Lys Glu Lys Ala Asn Gly
595 600

<210> 52
<211> 1809
<212> DNA
<213> Escherichia coli

<400> 52
atgcagacct tccaggccga tctggccatc gtgggtgcag gcggtgcagg tctccgggca 60
gcaatcgctg cagcccaggc gaacccaac gccaaagatcg cgctcatctc caaggtctac 120
ccgatgcggt cccataccgt ggcggcagag ggcggttcgg cggcagtggc gcaggaccac 180
gattcgttcg aataccactt ccacgacacg gtggccggcg gcgactggct gtgcgaacag 240
gatgtggtgg actacttcgt gcatcattgc cccaccgaaa tgaccagct ggaactgtgg 300
ggctgcccgt ggtcgcgccg gcccgatggc tccgtgaatg tgcgccggtt cggcggcatg 360
aagatcgaac gcacctggtt cgcggccgac aagacgggct tccacatgct ccataccctg 420
ttccagacga gcctgcagtt cccccagatc cagcggttcg acgagcactt cgtcctggac 480
atcctggtgg acgatggcca tgtgcgcggc ctcgtcgcca tgaacatgat ggagggcacc 540
ctcgtgcaga tccgcgcaa cgccgtggtc atggcgaccg gcggcgccgg ccgctgttac 600
cgctacaata ccaacggcgg catcgtgacc ggcgacggca tgggcatggc cctcagccac 660
ggcgtgcccc tccgcgacat ggagtctctg cagtaccacc ccaccggcct cccgggcagc 720
ggcatcctga tgaccgaggg ctgccggggc gaaggcggca tcctcgtcaa caagaacggc 780
tatcgctacc tgcaggacta cggcatgggc ccggaaaccc cgctcggcga gcccaagaat 840
aaatatatgg agctggggcc ccgggataag gtcagccagg cgttctggca cgaatggcgg 900
aagggaaca ccatctccac cccccgcggc gatgtcgtgt acctcgacct gcggcatctg 960
ggcgaaaaaa agctccacga acgcctgccg ttcattctgc agctggccaa ggcctacgtg 1020
ggcgtcgacc cggatgaagga gccgatcccg gtgcggccga ccgcgcatta cacgatgggc 1080
ggcatcgaac ccgaccagaa ctgcgaaacg cgcataagg gcctgttcgc ggtgggcgag 1140
tgagctccg tcggcctcca cggcgcgaa cgcctgggca gcaatagcct cgccgaactc 1200
gtggtcttcg gccggctggc cggcgagcag gccacggaac gggcgggccac cgcgggcaac 1260
ggcaacgagg ccgccatcga ggcccaggcc gccggcgtgg aacagcggct gaaggatctg 1320
gtcaaccagg atggcggcga aaactgggcg aagatccgcg acgaaatggg cctggcgatg 1380
gaagagggct gcggcatcta ccgcacgcc gagctgatgc agaagaccat cgataagctg 1440

gcggaactgc aggaacggtt caaacgggtc cggatcacgg atacctcctc ggtgttcaac 1500
accgacctcc tgtacacat cgagctgggc cacggcctca acgtggcgga gtgtatggcg 1560
cactccgcga tggcgcgga ggaagccgc ggtgcccacc agcgcctcga cgaaggctgc 1620
accgagcgcg acgatgtgaa cttcctcaaa cacacgtgg ctttcggga cgcggatggc 1680
acgacccgcc tcgaatatag cgatgtgaag atcacgacc tgcctcctgc gaagcgcgtc 1740
tacggcggtg aagccgacgc ggcagacaag gccgaggcag ccaataagaa ggagaaagcc 1800
aatggctga 1809

<210> 53
<211> 244
<212> PRT
<213> Escherichia coli

<400> 53

Met Ala Glu Met Lys Asn Leu Lys Ile Glu Val Val Arg Tyr Asn Pro
1 5 10 15

Glu Val Asp Thr Ala Pro His Ser Ala Phe Tyr Glu Val Pro Tyr Asp
20 25 30

Ala Thr Thr Ser Leu Leu Asp Ala Leu Gly Tyr Ile Lys Asp Asn Leu
35 40 45

Ala Pro Asp Leu Ser Tyr Arg Trp Ser Cys Arg Met Ala Ile Cys Gly
50 55 60

Ser Cys Gly Met Met Val Asn Asn Val Pro Lys Leu Ala Cys Lys Thr
65 70 75 80

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

Phe Pro Ile Glu Arg Asp Leu Val Val Asp Met Thr His Phe Ile Glu
100 105 110

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

Asp Gln Gly Thr Asn Ile Gln Thr Pro Ala Gln Met Ala Lys Tyr His
130 135 140

Gln Phe Ser Gly Cys Ile Asn Cys Gly Leu Cys Tyr Ala Ala Cys Pro
145 150 155 160

Gln Phe Gly Leu Asn Pro Glu Phe Ile Gly Pro Ala Ala Ile Thr Leu
165 170 175

Ala His Arg Tyr Asn Glu Asp Ser Arg Asp His Gly Lys Lys Glu Arg
180 185 190

Met Ala Gln Leu Asn Ser Gln Asn Gly Val Trp Ser Cys Thr Phe Val
195 200 205

Gly Tyr Cys Ser Glu Val Cys Pro Lys His Val Asp Pro Ala Ala Ala
210 215 220

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

Leu Lys Pro Arg

<210> 54
<211> 735
<212> DNA
<213> Escherichia coli

<400> 54
atggcgga tgaagaacct gaagatcgag gtcgtccggt ataatccgga ggtggatacc 60
gccccccaca gcgcgttcta cgaagtcccg tacgacgcca ccaccagcct cctggatgcg 120
ctgggctaca tcaaggataa cctggcccc gatctgtcgt accgctggtc ctgtcgcgatg 180
gccatctgtg gtcctgcgg catgatggtg aataacgtgc cgaagctggc ctgcaagacc 240
ttcctgcggg actataccga cggcatgaaa gtggaggccc tcgcgaactt cccgatcgag 300
cgcgacctgg tggtcgatat gacgcacttc atcgagagcc tggaagcgat caaacctat 360
atcatcggca acagccggac cgccgaccag ggcacgaaca tccagacgcc cgcgcatgatg 420
gcgaagtatc accagttctc cggctgtatc aattgcggcc tgtgctatgc ggcgtgcccg 480

cagttcggcc tgaaccccga gttcatcggc ccggcggcca tcacgctcgc ccaccgctac 540
aacgaggaca gccgcgatca cggcaagaag gagcgcgatgg cccagctcaa ttgcgagaat 600
ggcgtgtgga gctgcacctt cgtggggctat tgctccgagg tctgccccaa gcatgtcgac 660
ccggcggcgg cgatccagca gggcaaggtc gagtcgtcga aagacttcct gatcgccacg 720
ctcaagccgc ggtga 735

<210> 55
<211> 131
<212> PRT
<213> Escherichia coli

<400> 55

Met Thr Thr Lys Arg Lys Pro Tyr Val Arg Pro Met Thr Ser Thr Trp
1 5 10 15

Trp Lys Lys Leu Pro Phe Tyr Arg Phe Tyr Met Leu Arg Glu Gly Thr
20 25 30

Ala Val Pro Ala Val Trp Phe Ser Ile Glu Leu Ile Phe Gly Leu Phe
35 40 45

Ala Leu Lys Asn Gly Pro Glu Ala Trp Ala Gly Phe Val Asp Phe Leu
50 55 60

Gln Asn Pro Val Ile Val Ile Ile Asn Leu Ile Thr Leu Ala Ala Ala
65 70 75 80

Leu Leu His Thr Lys Thr Trp Phe Glu Leu Ala Pro Lys Ala Ala Asn
85 90 95

Ile Ile Val Lys Asp Glu Lys Met Gly Pro Glu Pro Ile Ile Lys Ser
100 105 110

Leu Trp Ala Val Thr Val Val Ala Thr Ile Val Ile Leu Phe Val Ala
115 120 125

Leu Tyr Trp
130

<210> 56
 <211> 396
 <212> DNA
 <213> Escherichia coli

<400> 56
 atgacgacca agcgcaagcc gtacgtccgg ccgatgacct cgacctggtg gaaaaagctc 60
 cccttctacc gcttctatat gctgcgcgag ggcaccgcgg tgccggccgt ctggttctcg 120
 atcgaactga tcttcggcct gttcgcgctg aagaatggcc cggaggcgtg ggccggcttc 180
 gtggacttcc tgcagaaccc cgtcatcgtc atcatcaatc tcatcacgct cgccgcggcc 240
 ctctgcaca ccaagacctg gttcagctg gccccgaagg cggcgaacat catcgtgaag 300
 gacgagaaga tgggcccgga gccgatcatc aaatccctgt gggcggtcac ggtggtcgcc 360
 acgatcgtca tcctgttcgt ggcgctgtac tggtga 396

<210> 57
 <211> 119
 <212> PRT
 <213> Escherichia coli

<400> 57

Met Ile Asn Pro Asn Pro Lys Arg Ser Asp Glu Pro Val Phe Trp Gly
 1 5 10 15

Leu Phe Gly Ala Gly Gly Met Trp Ser Ala Ile Ile Ala Pro Val Met
 20 25 30

Ile Leu Leu Val Gly Ile Leu Leu Pro Leu Gly Leu Phe Pro Gly Asp
 35 40 45

Ala Leu Ser Tyr Glu Arg Val Leu Ala Phe Ala Gln Ser Phe Ile Gly
 50 55 60

Arg Val Phe Leu Phe Leu Met Ile Val Leu Pro Leu Trp Cys Gly Leu
 65 70 75 80

His Arg Met His His Ala Met His Asp Leu Lys Ile His Val Pro Ala
 85 90 95

Gly Lys Trp Val Phe Tyr Gly Leu Ala Ala Ile Leu Thr Val Val Thr
 100 105 110

Leu Ile Gly Val Val Thr Ile
115

<210> 58
<211> 360
<212> DNA
<213> Escherichia coli

<400> 58
atgatcaatc cgaaccccaa acggagcgcg gagcccgtct tctggggcct gttcggcgcg 60
ggcggcatgt ggagcgccat catcgcgccg gtcgatgatcc tcctcgtggg catcctcctc 120
cccctggggc tgttccccgg cgacgccctg tcgtacgagc gcgtgctggc cttcgcccag 180
tccttcatcg gccgcgtggt cctcttcctc atgatcgtgc tccccctctg gtgcggcctg 240
caccgcatgc accacgccat gcacgacctg aagatccatg tccccgccgg caagtgggtc 300
ttctacggcc tcgcggcgat cctgacggtc gtgacgctga tcggcgtggt cacgatctga 360