

4765635_1
SEQUENCE LISTING

<110> BONUMOSE LLC

<120> ENZYMATIC SYNTHESIS OF D-TAGATOSE

<130> 146.0001-US00

<140> 15/743,481

<141> 2018-01-10

<150> PCT/US2016/054838

<151> 2016-09-30

<160> 17

<170> PatentIn version 3.5

<210> 1

<211> 1287

<212> DNA

<213> Anaerolinea thermophila UNI-1

<400> 1

atgttcggct cgcctgctcc cctgctggat atggtcaccg cgcagaaaca gggcatggcg	60
cggggtatcc catccatttg ttcggcacat ccggtggtgc tgagtgccgc ctgccatctt	120
gcccgccgga gcggcgcgcc cctgctcatc gaaaccacct gcaatcaggc caaccaccaa	180
ggtgggtaca gcggcatgac ccccgccgat tttgtccgct ttctgcgcga aattctggaa	240
cggggaaggtt tcccccgca acaggtcatc ctgggcgggg atcacctggg tccttacctc	300
tggcggaag agcctgcccga aaccgccata gcacaagcgc tggaaatggt gcgggcatac	360
gtgcaggcag gctacaccaa aattcatctg gacgcttcca tgccctgcgc cgatgacgac	420
cccgagcgtc cctgcccgt ggagcgcata gcccgcggg cggcgcagtt gtgcgccgcc	480
gccgaagccg ccgcgggagc ggttcagccg gtgtacgtaa ttggcagtga ggtgcccccg	540
cccggcggcg cgcagggtca ggaggcaaga cttcacgtca cactccgca ggaagcccaa	600
gccgcgctgg atgcctttcg ggaagccttt ctgcaggcag gcttgactcc cgtttgaggag	660
cggggtcattg cgctggtagt ccagccgggg gtggagtttg gcgtggacag cattcacgcc	720
tatcagcgcg aagccgcccc cccgctgaag accttcatcg agggcgtgcc cggcatggtg	780
tatgaagccc actcgaccga ttaccagacc cgtgcctccc tgcgtgcgct ggtggaagac	840
cactttttcca ttctcaagggt tgggtccggca ctaacctttg cctaccgcca agccgtgttc	900
gccctggaac acatcgaacg ggaaatattg ggcaggcagg atatgcctct ctcccgccgtg	960
agtgaagtcc tcgacgaggt gatgctgaac gatccacgcc actggcaggg atactttgccc	1020
ggcgctcccc ccgaacaggc gctggcgcgc cgctacagtt tcagcgaccg cattcgctat	1080
tactggcacc atccccgcc gcaggaagcc gtgcggagac tgctcgccaa cctgatcgaa	1140
accccgccgc cgctgagttt gctcagccag tacctgccgc gcgagtatga gatggtgcgc	1200
gcgggggaaa tctccagcca cccgcaggac ctgattcggg cacatatcca gcacacgctg	1260
gaagattacg ctgcggcgctg cgggtaa	1287

4765635_1

<210> 2
 <211> 428
 <212> PRT
 <213> Anaerolinea thermophila UNI-1

<400> 2

```

Met Phe Gly Ser Pro Ala Pro Leu Leu Asp Met Val Thr Ala Gln Lys
1      5      10      15

Gln Gly Met Ala Arg Gly Ile Pro Ser Ile Cys Ser Ala His Pro Val
      20      25      30

Val Leu Ser Ala Ala Cys His Leu Ala Arg Arg Ser Gly Ala Pro Leu
      35      40      45

Leu Ile Glu Thr Thr Cys Asn Gln Val Asn His Gln Gly Gly Tyr Ser
      50      55      60

Gly Met Thr Pro Ala Asp Phe Val Arg Phe Leu Arg Glu Ile Leu Glu
      65      70      75      80

Arg Glu Gly Ile Pro Pro Gln Gln Val Ile Leu Gly Gly Asp His Leu
      85      90      95

Gly Pro Tyr Pro Trp Arg Lys Glu Pro Ala Glu Thr Ala Ile Ala Gln
      100      105      110

Ala Leu Glu Met Val Arg Ala Tyr Val Gln Ala Gly Tyr Thr Lys Ile
      115      120      125

His Leu Asp Ala Ser Met Pro Cys Ala Asp Asp Asp Pro Glu Arg Pro
      130      135      140

Leu Pro Leu Glu Arg Ile Ala Arg Arg Ala Ala Gln Leu Cys Ala Ala
      145      150      155      160

Ala Glu Ala Ala Ala Gly Ala Val Gln Pro Val Tyr Val Ile Gly Ser
      165      170      175

Glu Val Pro Pro Pro Gly Gly Ala Gln Gly Gln Glu Ala Arg Leu His
      180      185      190

Val Thr Thr Pro Gln Glu Ala Gln Ala Ala Leu Asp Ala Phe Arg Glu
      195      200      205

Ala Phe Leu Gln Ala Gly Leu Thr Pro Val Trp Glu Arg Val Ile Ala
      210      215      220

Leu Val Val Gln Pro Gly Val Glu Phe Gly Val Asp Ser Ile His Ala
      225      230      235      240
  
```

4765635_1

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

Pro Gly Met Val Tyr Glu Ala His Ser Thr Asp Tyr Gln Thr Arg Ala
260 265 270

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

Pro Ala Leu Thr Phe Ala Tyr Arg Glu Ala Val Phe Ala Leu Glu His
290 295 300

Ile Glu Arg Glu Ile Leu Gly Arg Gln Asp Met Pro Leu Ser Arg Leu
305 310 315 320

Ser Glu Val Leu Asp Glu Val Met Leu Asn Asp Pro Arg His Trp Gln
325 330 335

Gly Tyr Phe Ala Gly Ala Pro Ala Glu Gln Ala Leu Ala Arg Arg Tyr
340 345 350

Ser Phe Ser Asp Arg Ile Arg Tyr Tyr Trp His His Pro Ala Ala Gln
355 360 365

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

Leu Ser Leu Leu Ser Gln Tyr Leu Pro Arg Glu Tyr Glu Met Val Arg
385 390 395 400

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

Gln His Thr Leu Glu Asp Tyr Ala Ala Ala Cys Gly
420 425

<210> 3
<211> 1290
<212> DNA
<213> Caldicellulosiruptor kronotskyensis

<400> 3
atgagtcctc aaaatccatt gattggttta tttaagaata gagaaaaaga gtttaagggt 60
attatttcag tttgttcttc aaatgaaata gtcttagaag cagtttttaa aagaatgaaa 120
gatacaaacc taccaattat tattgaagcc acagcgaacc aggtaaatca atttggcggg 180
tattctgggt tgacaccgtc tcagttcaaa gaacgagtta taaaaattgc tcaaaaagtt 240
gattttccac ttgagagaat aattcttggt ggggaccatc ttggaccatt tgtgtggcgt 300
gaccaggaac cagaaattgc tatggagtat gctaagcaaa tgataaaaaga atacataaaa 360
gcagggtttta ccaaaattca catcgacacg agtatgcctt taaaagggga gaacagcata 420

4765635_1

gatgatgaaa taattgctaa aagaactgct gtgctctgca ggattgcgga ggagtgtttt	480
gagaagattht ctataaacia tccctatatt acaaggccag tttatgtgat aggagctgat	540
gtgccacctc ccggcggaga gtcttctatt tgtcaaacia ttactactaa agatgaatta	600
gaaagaagtt tagaatattht caaagaagca tttaaaaagg aaggaattga gcatgtattc	660
gattatgtag ttgctgttgt tgcaaatttht ggagttgaat ttgggagcga tgaaatttgt	720
gattttgata tggaaaaagt aaagccgcta aaagaacttht tggcaaagta caatatagta	780
tttgaaggcc attctacaga ttatcaaaca aaagaaaact taaaaagaat ggtcgaatgt	840
ggtattgcaa ttttaaagggt tggctctgct ctaacattta cattgcgcga agcgttagta	900
gcacttagtc atattgaaga agaaatttat agcaatgaaa aggagaaaact gtcaagattht	960
agagaagtht tattgaatac tatgctaaca tgcaaagatc actggagtaa atattttgat	1020
gagaatgata agttaattaa gtcaaagctc ctatatagct atcttgacag atggagatac	1080
tattttgaaa acgagagtgt gaaaagtgtc gtttattctc ttattggaaa tttagagaat	1140
gttaaaattc caccttggtc tgtaagtcag tattttcctt ctcagtacca aaagatgaga	1200
aaaaaagatt taaaaaacgg tgctgccgac ctaatattgg ataaaatagg ggaagtcatt	1260
gaccattatg tttatgcggt aaaagaataa	1290

<210> 4
 <211> 429
 <212> PRT
 <213> Caldicellulosiruptor kronotskyensis
 <400> 4

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

4765635_1

Asp Thr Ser Met Pro Leu Lys Gly Glu Asn Ser Ile Asp Asp Glu Ile
 130 135 140
 Ile Ala Lys Arg Thr Ala Val Leu Cys Arg Ile Ala Glu Glu Cys Phe
 145 150 155 160
 Glu Lys Ile Ser Ile Asn Asn Pro Tyr Ile Thr Arg Pro Val Tyr Val
 165 170 175
 Ile Gly Ala Asp Val Pro Pro Pro Gly Gly Glu Ser Ser Ile Cys Gln
 180 185 190
 Thr Ile Thr Thr Lys Asp Glu Leu Glu Arg Ser Leu Glu Tyr Phe Lys
 195 200 205
 Glu Ala Phe Lys Lys Glu Gly Ile Glu His Val Phe Asp Tyr Val Val
 210 215 220
 Ala Val Val Ala Asn Phe Gly Val Glu Phe Gly Ser Asp Glu Ile Val
 225 230 235 240
 Asp Phe Asp Met Glu Lys Val Lys Pro Leu Lys Glu Leu Leu Ala Lys
 245 250 255
 Tyr Asn Ile Val Phe Glu Gly His Ser Thr Asp Tyr Gln Thr Lys Glu
 260 265 270
 Asn Leu Lys Arg Met Val Glu Cys Gly Ile Ala Ile Leu Lys Val Gly
 275 280 285
 Pro Ala Leu Thr Phe Thr Leu Arg Glu Ala Leu Val Ala Leu Ser His
 290 295 300
 Ile Glu Glu Glu Ile Tyr Ser Asn Glu Lys Glu Lys Leu Ser Arg Phe
 305 310 315 320
 Arg Glu Val Leu Leu Asn Thr Met Leu Thr Cys Lys Asp His Trp Ser
 325 330 335
 Lys Tyr Phe Asp Glu Asn Asp Lys Leu Ile Lys Ser Lys Leu Leu Tyr
 340 345 350
 Ser Tyr Leu Asp Arg Trp Arg Tyr Tyr Phe Glu Asn Glu Ser Val Lys
 355 360 365
 Ser Ala Val Tyr Ser Leu Ile Gly Asn Leu Glu Asn Val Lys Ile Pro
 370 375 380
 Pro Trp Leu Val Ser Gln Tyr Phe Pro Ser Gln Tyr Gln Lys Met Arg
 385 390 395 400

4765635_1

Lys Lys Asp Leu Lys Asn Gly Ala Ala Asp Leu Ile Leu Asp Lys Ile
405 410 415

Gly Glu Val Ile Asp His Tyr Val Tyr Ala Val Lys Glu
420 425

<210> 5
<211> 1227
<212> DNA
<213> *Caldilinea aerophila*

<400> 5
atgtcaacac ttcgccacat cattttgcga ctgatcgagc tgcgtgaacg agaacagatc 60
catctcacgc tgctggccgt ctgtcccaac tcggcggcgg tgctggaggc agcggatgaag 120
gtcgccgcgc gctgccacac gccgatgctc ttcgctgcca cgctcaatca agtcgatcgc 180
gacggcggct acaccggttg gacgcctgcg caattcgtcg ccgagatgcg tcgctatgcc 240
gtccgctatg gctgcaccac cccgctctat ccttgccctgg atcacggcgg gccgtggctc 300
aaagatcgcc atgcacagga aaagctaccg ctcgaccagg cgatgcatga ggtcaagctg 360
agcctcacgc cctgtctgga ggccggctac gcgctgctgc acatcgaccc cacggtcgat 420
cgcacgctcc cgcccggaga agcgcgcgtc gtgccgatcg tcgtcgagcg cacggtcgag 480
ctgatcgaac atgccgaaca ggagcgacag cggtgaacc tgccggcggg cgcctatgaa 540
gtcggcaccg aagaagtaca tggcgggctg gtgaatttcg acaattttgt cgccttcttg 600
gatttgctca aggcaaggct tgaacaacgt gccctgatgc acgcctggcc cgccttcgtg 660
gtggcgcagg tcggcactga cctgcataca acgtattttg accccagtgc ggcgcaacgg 720
ctgactgaga tcgtgcgccc taccgggtgca ctgttgaagg ggcactacac cgactgggtc 780
gaaaatcccc ccgactatcc gagggtaggc atgggaggcg ccaacgttgg tccagagttt 840
acggcggccg agttcgaggc gctggaagcg ctggaacggc gggaacaacg gctgtgcgcc 900
aaccggaaat tgcagcccgc ctgttttttg gctgcactgg aagaggcagt agtcgcttca 960
gatcgttggc ggaagtggct ccagcccgat gagatcggca agccctttgc agaattaacg 1020
ccgcacgcc ggcgctggct cgtgcagacc ggggcacgct acgtctggac tgcgccgaaa 1080
gttatcgccg cacgcgaaca gctctatgcg cacctctccc ttgtgcaggc ggatccacat 1140
gcctacgtgg tagagtcagt cgcccgggtca atcgagcgct atatcgatgc cttcaactta 1200
tacgacgccg ctacattgct tggatga 1227

<210> 6
<211> 408
<212> PRT
<213> *Caldilinea aerophila*

<400> 6

Met Ser Thr Leu Arg His Ile Ile Leu Arg Leu Ile Glu Leu Arg Glu
1 5 10 15

4765635_1

Arg Glu Gln Ile His Leu Thr Leu Leu Ala Val Cys Pro Asn Ser Ala
20 25 30

Ala Val Leu Glu Ala Ala Val Lys Val Ala Ala Arg Cys His Thr Pro
35 40 45

Met Leu Phe Ala Ala Thr Leu Asn Gln Val Asp Arg Asp Gly Gly Tyr
50 55 60

Thr Gly Trp Thr Pro Ala Gln Phe Val Ala Glu Met Arg Arg Tyr Ala
65 70 75 80

Val Arg Tyr Gly Cys Thr Thr Pro Leu Tyr Pro Cys Leu Asp His Gly
85 90 95

Gly Pro Trp Leu Lys Asp Arg His Ala Gln Glu Lys Leu Pro Leu Asp
100 105 110

Gln Ala Met His Glu Val Lys Leu Ser Leu Thr Ala Cys Leu Glu Ala
115 120 125

Gly Tyr Ala Leu Leu His Ile Asp Pro Thr Val Asp Arg Thr Leu Pro
130 135 140

Pro Gly Glu Ala Pro Leu Val Pro Ile Val Val Glu Arg Thr Val Glu
145 150 155 160

Leu Ile Glu His Ala Glu Gln Glu Arg Gln Arg Leu Asn Leu Pro Ala
165 170 175

Val Ala Tyr Glu Val Gly Thr Glu Glu Val His Gly Gly Leu Val Asn
180 185 190

Phe Asp Asn Phe Val Ala Phe Leu Asp Leu Leu Lys Ala Arg Leu Glu
195 200 205

Gln Arg Ala Leu Met His Ala Trp Pro Ala Phe Val Val Ala Gln Val
210 215 220

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

Leu Thr Glu Ile Val Arg Pro Thr Gly Ala Leu Leu Lys Gly His Tyr
245 250 255

Thr Asp Trp Val Glu Asn Pro Ala Asp Tyr Pro Arg Val Gly Met Gly
260 265 270

Gly Ala Asn Val Gly Pro Glu Phe Thr Ala Ala Glu Phe Glu Ala Leu
275 280 285

4765635_1

Glu Ala Leu Glu Arg Arg Glu Gln Arg Leu Cys Ala Asn Arg Lys Leu
290 295 300

Gln Pro Ala Cys Phe Leu Ala Ala Leu Glu Glu Ala Val Val Ala Ser
305 310 315 320

Asp Arg Trp Arg Lys Trp Leu Gln Pro Asp Glu Ile Gly Lys Pro Phe
325 330 335

Ala Glu Leu Thr Pro Ala Arg Arg Arg Trp Leu Val Gln Thr Gly Ala
340 345 350

Arg Tyr Val Trp Thr Ala Pro Lys Val Ile Ala Ala Arg Glu Gln Leu
355 360 365

Tyr Ala His Leu Ser Leu Val Gln Ala Asp Pro His Ala Tyr Val Val
370 375 380

Glu Ser Val Ala Arg Ser Ile Glu Arg Tyr Ile Asp Ala Phe Asn Leu
385 390 395 400

Tyr Asp Ala Ala Thr Leu Leu Gly
405

<210> 7
<211> 1362
<212> DNA
<213> *Caldithrix abyssi*

<400> 7
atgagtctgc atcctttaaa taaattaatc gagcgacaca aaaaaggaaac gccggtcgggt 60
atttattccg tctgttcggc caatcccttt gttttgaaag cggccatgct acaggcgcaa 120
aaggatcagt ctttgctact tattgaggcc acttccaacc aggtagatca attcggcgggt 180
tacaccggca tgcggcccga agattttaaa acaatgacgc ttgaactggc agccgaaaac 240
aattacgatc cacagggatt aatcctgggc ggcgaccatc tggggcccaa ccgctggaca 300
aaactgagcg cctcccgggc catggactac gccagagagc agattgccgc ttatgttaaa 360
gccggctttt ccaaaatcca cttagacgcc accatgccct tgcaaaacga tgccacagat 420
tccgccggcc gccttccagt cgaaacaatc gctcaacgta ccgcagaatt atgcgccgtg 480
gccgaacaaa cttaccggca gagcgaccaa ctctttccgc cgcctgttta cattgtcggc 540
agcgacgtgc ccatcccggg cggcgcgcaa gaagcgctga accagatcca tattacggag 600
gtaaaagagg ttcaacagac cattgatcac gtgcggcggg cctttgaaaa aaacggcctg 660
gaagcggctt acgaaagagt ttgcgccgtt gtcgtgcagc caggcgttga attcgccgat 720
caaatcgttt ttgaatacgc tcccgacaga gcggcggcct taaaagattt tattgaaagc 780
cattcgcagc tggtttatga agcgactct actgattacc agaccgcacc tcttttgcgc 840

4765635_1

cagatggtaa aagatcactt tgccatttta aaggtcgggc ctgcgctcac ctttgcctg	900
cgcgaagcca tttttgctct ggcctttatg gaaaaagagc ttttgccatt gcacagagcg	960
ctcaaacctt ctgccattct ggaaacgctg gaccaaacga tggacaaaaa ccctgcttac	1020
tggcaaaagc attacggcgg aacaaaggaa gaagtacgct ttgcgcagcg gtttagcctg	1080
agcgaccgca ttcgttacta ctggccgttt ccaaagggtt aaaaggccct gcgccaattg	1140
ctaaaaaact tgcaacaaat ttccattcct ctaactttgg taagccagtt catgccagag	1200
gaataccaac gtattcgcca aggaacgtta accaacgatc cgcaggcgct gattttgaac	1260
aaaattcaaa gcgtattaaa gcaatacgcg gaggcgacgc aaattcaaaa ctctttgaca	1320
ttcacgcaaa atcaaaattc attagcaatg gagcgactat ga	1362

<210> 8
 <211> 453
 <212> PRT
 <213> *Caldithrix abyssi*

<400> 8

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

4765635_1

Tyr Ile Val Gly Ser Asp Val Pro Ile Pro Gly Gly Ala Gln Glu Ala
 180 185 190
 Leu Asn Gln Ile His Ile Thr Glu Val Lys Glu Val Gln Gln Thr Ile
 195 200 205
 Asp His Val Arg Arg Ala Phe Glu Lys Asn Gly Leu Glu Ala Ala Tyr
 210 215 220
 Glu Arg Val Cys Ala Val Val Val Gln Pro Gly Val Glu Phe Ala Asp
 225 230 235 240
 Gln Ile Val Phe Glu Tyr Ala Pro Asp Arg Ala Ala Ala Leu Lys Asp
 245 250 255
 Phe Ile Glu Ser His Ser Gln Leu Val Tyr Glu Ala His Ser Thr Asp
 260 265 270
 Tyr Gln Thr Ala Pro Leu Leu Arg Gln Met Val Lys Asp His Phe Ala
 275 280 285
 Ile Leu Lys Val Gly Pro Ala Leu Thr Phe Ala Leu Arg Glu Ala Ile
 290 295 300
 Phe Ala Leu Ala Phe Met Glu Lys Glu Leu Leu Pro Leu His Arg Ala
 305 310 315 320
 Leu Lys Pro Ser Ala Ile Leu Glu Thr Leu Asp Gln Thr Met Asp Lys
 325 330 335
 Asn Pro Ala Tyr Trp Gln Lys His Tyr Gly Gly Thr Lys Glu Glu Val
 340 345 350
 Arg Phe Ala Gln Arg Phe Ser Leu Ser Asp Arg Ile Arg Tyr Tyr Trp
 355 360 365
 Pro Phe Pro Lys Val Gln Lys Ala Leu Arg Gln Leu Leu Lys Asn Leu
 370 375 380
 Gln Gln Ile Ser Ile Pro Leu Thr Leu Val Ser Gln Phe Met Pro Glu
 385 390 395 400
 Glu Tyr Gln Arg Ile Arg Gln Gly Thr Leu Thr Asn Asp Pro Gln Ala
 405 410 415
 Leu Ile Leu Asn Lys Ile Gln Ser Val Leu Lys Gln Tyr Ala Glu Ala
 420 425 430
 Thr Gln Ile Gln Asn Ser Leu Thr Phe Thr Gln Asn Gln Asn Ser Leu
 435 440 445

4765635_1

Ala Met Glu Arg Leu
450

<210> 9
<211> 1227
<212> DNA
<213> Dictyoglomus thermophilum

<400> 9
atgtggctta gtaaagatta tttagaaaa aagggagttt attctatatg tagctctaatt 60
ccatatgtga ttgaggcaag tgttgaattt gctaaggaga agaatgatta tattttaatt 120
gaggcgacac ctcatcagat aaaccagttt ggtggatatt caggatgac tcccgaagat 180
tttaaaaact ttgtaatggg aataataaaa gaaaagggaa tagaagagga tagggtgatt 240
cttggagggg accatttagg cctctccct tggcaagatg aaccttcttc ttctgcaatg 300
aaaaaggcaa aagaccttat aagggccttt gtggagagtg gttataagaa gatacacctt 360
gattgtagta tgtctctttc tgatgaccc gtagtgctct ctcccagaa gatagcagaa 420
agggagaggg aacttcttga gggtgcagaa gagactgcta gaaagtacaa ttttcagcct 480
gtgtatgtgg tgggaactga tgtaccggtg gctggaggag gcgaagagga aggtattacc 540
tcagtggagg attttagagt agcaatctcc tctttaaaaa aatattttga ggatgttcca 600
aggatatggg ataggataat tggttttgta ataatgcttg gtatagggtt taattatgaa 660
aaagtgtttg agtatgacag gattaagggt agaaaaattt tagaggaggt aaagaaagag 720
aatctttttg ttgaagggtca ctctactgac tatcagacaa aacgtgcatt gagagatatg 780
gtagaggatg gagtaagaat tcttaagggt ggtcctgctt taacagcaag ttttagaagg 840
ggagtatttt tattaagtag cattgaggat gagcttatat cggaagataa aaggtctaatt 900
attaagaaag ttgtgcttga gactatgtta aaagatgata aatattggag aaagtattat 960
aaggattcag aaagattaga attagatatt tgggtacaact tacttgatag gattagatat 1020
tattgggaat ataaagagat aaaaatagct ttaaataaggc tttttgaaaa tttttcggaa 1080
gggggttgata ttagatacat ctatcaatat ttttatgatt cgtattttta agtaagagaa 1140
ggaaaaataa gaaatgatcc aagggagcta ataaagaatg aaataaagaa ggtcttggag 1200
gactatcact atgctgtaaa cttataa 1227

<210> 10
<211> 1227
<212> DNA
<213> Artificial Sequence

<220>
<223> Codon optimized nucleotide sequence

<400> 10
atgtggctga gcaaggacta cctgcgtaag aagggcgttt acagcatttg cagcagcaac 60
ccgtatgtta ttgaagcgag cgtggagttc gcgaaggaga aaaacgatta catcctgatt 120

4765635_1

gaagcgaccc cgcaccagat caaccaattt ggtggctata gcggcatgac cccggaggac	180
ttcaagaact ttgttatggg catcattaag gaaaaaggta tcgaggaaga ccgtgtgatt	240
ctgggtggcg atcacctggg tccgctgccg tggcaggatg agccgagcag cagcgcgatg	300
aagaaagcga aagacctgat ccgtgcgttc gttgaaagcg gttacaagaa aattcacctg	360
gattgcagca tgagcctgag cgacgatccg gtggttctga gcccggagaa gatcgcgaa	420
cgtgagcgtg aactgctgga agttgcggag gaaaccgcgc gtaaatacaa ctttcaaccg	480
gtgtatgtgg tgggtaccga tggtccggtt gcgggtggcg gtgaggaaga gggatcacc	540
agcgtggagg acttccgtgt tgcgattagc agcctgaaga aatactttga agacgttccg	600
cgtatttggg atcgtatcat tggtttcgtg atcatgctgg gcattggttt caactacgag	660
aaggtgtttg aatatgatcg tatcaaagtg cgtaaaattc tggaagaggt taagaaagag	720
aacctgtttg tggaaggcca cagcaccgac tatcagacca agcgtgcgct gcgtgacatg	780
gtggaggatg gcgttcgtat cctgaaagtg ggtccggcgc tgaccgcgag cttccgtcgt	840
ggtgtgtttc tgctgagcag catcaggagc gaactgatta gcgaggataa acgtagcaac	900
attaagaaag tggttctgga aaccatgctg aaggacgata aatactggcg taagtactat	960
aaagacagcg agcgtctgga actggatata tggtaaac tgctggaccg tattcgttac	1020
tactgggagt acaaggaaat caagattgcg ctgaaccgtc tgttcgagaa ctttagcgaa	1080
ggcgttgata tccgttacat ctaccaatac ttctacgaca gctacttcaa agtgctgag	1140
ggtaaaatcc gtaacgaccc gcgtgaactg attaagaacg agattaagaa agtgctggaa	1200
gactaccatt atgcggtgaa cctgtaa	1227

<210> 11
 <211> 408
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Codon optimized nucleotide sequence

<400> 11

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

Cys Ser Ser Asn Pro Tyr Val Ile Glu Ala Ser Val Glu Phe Ala Lys
 20 25 30

Glu Lys Asn Asp Tyr Ile Leu Ile Glu Ala Thr Pro His Gln Ile Asn
 35 40 45

Gln Phe Gly Gly Tyr Ser Gly Met Thr Pro Glu Asp Phe Lys Asn Phe
 50 55 60

Val Met Gly Ile Ile Lys Glu Lys Gly Ile Glu Glu Asp Arg Val Ile
 65 70 75 80

4765635_1

Leu Gly Gly Asp His Leu Gly Pro Leu Pro Trp Gln Asp Glu Pro Ser
 85 90
 Ser Ser Ala Met Lys Lys Ala Lys Asp Leu Ile Arg Ala Phe Val Glu
 100 105 110
 Ser Gly Tyr Lys Lys Ile His Leu Asp Cys Ser Met Ser Leu Ser Asp
 115 120 125
 Asp Pro Val Val Leu Ser Pro Glu Lys Ile Ala Glu Arg Glu Arg Glu
 130 135 140
 Leu Leu Glu Val Ala Glu Glu Thr Ala Arg Lys Tyr Asn Phe Gln Pro
 145 150 155 160
 Val Tyr Val Val Gly Thr Asp Val Pro Val Ala Gly Gly Gly Glu Glu
 165 170 175
 Glu Gly Ile Thr Ser Val Glu Asp Phe Arg Val Ala Ile Ser Ser Leu
 180 185 190
 Lys Lys Tyr Phe Glu Asp Val Pro Arg Ile Trp Asp Arg Ile Ile Gly
 195 200 205
 Phe Val Ile Met Leu Gly Ile Gly Phe Asn Tyr Glu Lys Val Phe Glu
 210 215 220
 Tyr Asp Arg Ile Lys Val Arg Lys Ile Leu Glu Glu Val Lys Lys Glu
 225 230 235 240
 Asn Leu Phe Val Glu Gly His Ser Thr Asp Tyr Gln Thr Lys Arg Ala
 245 250 255
 Leu Arg Asp Met Val Glu Asp Gly Val Arg Ile Leu Lys Val Gly Pro
 260 265 270
 Ala Leu Thr Ala Ser Phe Arg Arg Gly Val Phe Leu Leu Ser Ser Ile
 275 280 285
 Glu Asp Glu Leu Ile Ser Glu Asp Lys Arg Ser Asn Ile Lys Lys Val
 290 295 300
 Val Leu Glu Thr Met Leu Lys Asp Asp Lys Tyr Trp Arg Lys Tyr Tyr
 305 310 315 320
 Lys Asp Ser Glu Arg Leu Glu Leu Asp Ile Trp Tyr Asn Leu Leu Asp
 325 330 335
 Arg Ile Arg Tyr Tyr Trp Glu Tyr Lys Glu Ile Lys Ile Ala Leu Asn
 340 345 350

4765635_1

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

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

Asn Asp Pro Arg Glu Leu Ile Lys Asn Glu Ile Lys Lys Val Leu Glu
385 390 395 400

Asp Tyr His Tyr Ala Val Asn Leu
405

<210> 12
<211> 672
<212> DNA
<213> Archaeoglobus fulgidus

<400> 12
atgttcaaac caaaggccat cgcagttgac atagatggca ccctcaccga cagaaagagg 60
gctctgaact gcagggctgt tgaagctctc cgcaaggtaa aaattcccgt gattttggcc 120
actggttaaca tatcttgttt tgcgagggct gcagcaaagc tgattggagt ctcagacgtg 180
gtaatctgcg agaatggggg cgtggtgagg ttcgagtagc atggggagga tattgtttta 240
ggagataaag agaaatgcgt tgaggctgtg aggggtgctt agaaacacta tgaggttgag 300
ctgctggact tcgaatacag gaagtcggaa gtgtgcatga ggaggagctt tgacatcaac 360
gaggcgagaa agctcattga ggggatgggg gttaagcttg tggattcagg ctttgcctac 420
cacattatgg atgctgatgt tagcaaggga aaagctttga agttcgttgc cgagaggctt 480
ggtatcagtt cagcggagtt tgcagttatc ggcgactcag agaacgacat agacatgttc 540
agagttgctg gattcggaaat tgctgttgcc aatgccgatg agaggctgaa ggagtatgct 600
gatttagtta cgccatcacc agacggcgag ggggttggtg aggctttgca gtttctggga 660
ttgttgcggt ga 672

<210> 13
<211> 223
<212> PRT
<213> Archaeoglobus fulgidus

<400> 13

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

Asp Arg Lys Arg Ala Leu Asn Cys Arg Ala Val Glu Ala Leu Arg Lys
20 25 30

Val Lys Ile Pro Val Ile Leu Ala Thr Gly Asn Ile Ser Cys Phe Ala
35 40 45

Arg Ala Ala Ala Lys Leu Ile Gly Val Ser Asp Val Val Ile Cys Glu
Page 14

50

55

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

Gly Asp Lys Glu Lys Cys Val Glu Ala Val Arg Val Leu Glu Lys His
85 90 95

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

Met Arg Arg Ser Phe Asp Ile Asn Glu Ala Arg Lys Leu Ile Glu Gly
115 120 125

Met Gly Val Lys Leu Val Asp Ser Gly Phe Ala Tyr His Ile Met Asp
130 135 140

Ala Asp Val Ser Lys Gly Lys Ala Leu Lys Phe Val Ala Glu Arg Leu
145 150 155 160

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

Ile Asp Met Phe Arg Val Ala Gly Phe Gly Ile Ala Val Ala Asn Ala
180 185 190

Asp Glu Arg Leu Lys Glu Tyr Ala Asp Leu Val Thr Pro Ser Pro Asp
195 200 205

Gly Glu Gly Val Val Glu Ala Leu Gln Phe Leu Gly Leu Leu Arg
210 215 220

<210> 14
<211> 660
<212> DNA
<213> Archaeoglobus profundus

<400> 14
gtgttcaagg ctttggtagt tgatatagac ggaactttga cggataagaa gagggcaata 60
aactgcagag cggtcgaagc acttagaaaa ctaaagattc ctgttgtctt ggcaaccgga 120
aacatttcac gctttgcaag ggctgtagct aagattatag gtgtttccga tattgtaata 180
gctgagaacg gaggtgttgt cagattcagc tacgacggag aggacatagt tctgggggat 240
agaagtaaag gcttaagagc tttggagaca ctagaaaac gcttcaaagt agagcttctc 300
gacaacgaat ataggaagtc tgaggtctgc atgaggagga acttccttat agaggaagct 360
agaaagatac tgccaaaaga tgttagaata gtcgatacag gcttcgcata ccacataatc 420
gatgcaaatg tcagcaaggg gaaggctttg atgttcatac cggataagct tggcttggac 480
gttaaggatt tcattgcgat aggtgattcc gaaaacgaca ttgaaatgtt ggaagttgca 540
ggttttggcg ttgcagttgc gaatgcggat gaaaagctta aggaggtagc ggatttggtc 600

acatcgaagc ctaatggaga cggagttgtc gaagctcttg agttcttggg actcatttag

660

<210> 15
 <211> 219
 <212> PRT
 <213> Archaeoglobus profundus

<400> 15

Met Phe Lys Ala Leu Val Val Asp Ile Asp Gly Thr Leu Thr Asp Lys
 1 5 10 15

Lys Arg Ala Ile Asn Cys Arg Ala Val Glu Ala Leu Arg Lys Leu Lys
 20 25 30

Ile Pro Val Val Leu Ala Thr Gly Asn Ile Ser Cys Phe Ala Arg Ala
 35 40 45

Val Ala Lys Ile Ile Gly Val Ser Asp Ile Val Ile Ala Glu Asn Gly
 50 55 60

Gly Val Val Arg Phe Ser Tyr Asp Gly Glu Asp Ile Val Leu Gly Asp
 65 70 75 80

Arg Ser Lys Cys Leu Arg Ala Leu Glu Thr Leu Arg Lys Arg Phe Lys
 85 90 95

Val Glu Leu Leu Asp Asn Glu Tyr Arg Lys Ser Glu Val Cys Met Arg
 100 105 110

Arg Asn Phe Pro Ile Glu Glu Ala Arg Lys Ile Leu Pro Lys Asp Val
 115 120 125

Arg Ile Val Asp Thr Gly Phe Ala Tyr His Ile Ile Asp Ala Asn Val
 130 135 140

Ser Lys Gly Lys Ala Leu Met Phe Ile Ala Asp Lys Leu Gly Leu Asp
 145 150 155 160

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

Leu Glu Val Ala Gly Phe Gly Val Ala Val Ala Asn Ala Asp Glu Lys
 180 185 190

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

Val Val Glu Ala Leu Glu Phe Leu Gly Leu Ile
 210 215

<210> 16

4765635_1

<211> 675
 <212> DNA
 <213> Archaeoglobus veneficus

<400> 16
 atgctccgtc caaaggtct cgccattgac atcgacggaa ccataacata caggaatcga 60
 agcctgaact gtaaggccgt tgaagctctc aggaaggtaa aaatccctgt agttcttgca 120
 actggcaaca taccctgttt cgcaagaact gctgcaaagc ttataggcgt ctcagacatt 180
 gttatatgcg aaaatggagg tattgttcga ttcagctacg atggcgacga catagtgcct 240
 ggggacataa gcaaatgcct taaagcggct gaaattctca aagagtactt tgaaatcgaa 300
 ttccttgacg ctgagtacag gaagtcggag gtctgtcttc gcagaaactt tcctattgaa 360
 gaggcgagga aaattcttca cgatgcaaag cttgatgtta aaatcgtcga ttcagggttt 420
 gcgtaccaca taatggatgc gaaggtcagc aaaggaaggg ctcttgagta catagctgat 480
 gaacttgta taagtccgaa ggagttcgct gcaattggtg attctgagaa cgacatagac 540
 ctgattaagg ctgccggcct cggatttgcc gttggagatg ctgacttaaa gctcaaaatg 600
 gaggccgacg tggtagtctc gaagaagaat ggcatggag ttgttgaagc acttgagctt 660
 ctgggccttaa tttaa 675

<210> 17
 <211> 224
 <212> PRT
 <213> Archaeoglobus veneficus

<400> 17
 Met Leu Arg Pro Lys Gly Leu Ala Ile Asp Ile Asp Gly Thr Ile Thr
 1 5 10 15
 Tyr Arg Asn Arg Ser Leu Asn Cys Lys Ala Val Glu Ala Leu Arg Lys
 20 25 30
 Val Lys Ile Pro Val Val Leu Ala Thr Gly Asn Ile Ser Cys Phe Ala
 35 40 45
 Arg Thr Ala Ala Lys Leu Ile Gly Val Ser Asp Ile Val Ile Cys Glu
 50 55 60
 Asn Gly Gly Ile Val Arg Phe Ser Tyr Asp Gly Asp Asp Ile Val Leu
 65 70 75 80
 Gly Asp Ile Ser Lys Cys Leu Lys Ala Ala Glu Ile Leu Lys Glu Tyr
 85 90 95
 Phe Glu Ile Glu Phe Leu Asp Ala Glu Tyr Arg Lys Ser Glu Val Cys
 100 105 110
 Leu Arg Arg Asn Phe Pro Ile Glu Glu Ala Arg Lys Ile Leu His Asp
 115 120 125

4765635_1

Ala Lys Leu Asp Val Lys Ile Val Asp Ser Gly Phe Ala Tyr His Ile
130 135 140

Met Asp Ala Lys Val Ser Lys Gly Arg Ala Leu Glu Tyr Ile Ala Asp
145 150 155 160

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

Asn Asp Ile Asp Leu Ile Lys Ala Ala Gly Leu Gly Ile Ala Val Gly
180 185 190

Asp Ala Asp Leu Lys Leu Lys Met Glu Ala Asp Val Val Val Ser Lys
195 200 205

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