

SEQUENCE LISTING

<110> BASF SE

<120> New Malonate Decarboxylases for Industrial Biocatalysis

<130> M/48290-PCT

<160> 36

<170> PatentIn version 3.3

<210> 1

<211> 744

<212> DNA

<213> Bordetella bronchiseptica

<220>

<221> CDS

<222> (1)..(744)

<400> 1

atg	cag	cag	gcg	agc	acc	ccg	acc	att	ggc	atg	att	gtg	ccg	ccg	gca	48
Met	Gln	Gln	Ala	Ser	Thr	Pro	Thr	Ile	Gly	Met	Ile	Val	Pro	Pro	Ala	
1				5					10					15		

gcg	ggt	ctg	gtt	ccg	gcg	gat	ggc	gcg	cgt	ctg	tat	ccg	gat	ctg	ccg	96
Ala	Gly	Leu	Val	Pro	Ala	Asp	Gly	Ala	Arg	Leu	Tyr	Pro	Asp	Leu	Pro	
			20					25					30			

ttt	att	gcg	agc	ggc	ctg	ggc	ctg	ggt	agc	gtt	acc	ccg	gaa	ggc	tat	144
Phe	Ile	Ala	Ser	Gly	Leu	Gly	Leu	Gly	Ser	Val	Thr	Pro	Glu	Gly	Tyr	
		35					40					45				

gat	gcg	gtg	att	gaa	agc	gtg	gtg	gat	cat	gcg	cgt	cgt	ctg	cag	aaa	192
Asp	Ala	Val	Ile	Glu	Ser	Val	Val	Asp	His	Ala	Arg	Arg	Leu	Gln	Lys	
		50				55					60					

cag	ggt	gcg	gcg	gtg	gtg	agc	ctg	atg	ggc	acc	agc	ctg	agc	ttt	tat	240
Gln	Gly	Ala	Ala	Val	Val	Ser	Leu	Met	Gly	Thr	Ser	Leu	Ser	Phe	Tyr	
65					70					75					80	

cgt	ggt	gcg	gcg	ttt	aac	gcg	gcg	ctg	acc	gtg	gcg	atg	cgt	gaa	gcg	288
Arg	Gly	Ala	Ala	Phe	Asn	Ala	Ala	Leu	Thr	Val	Ala	Met	Arg	Glu	Ala	
				85					90					95		

acc	ggt	ctg	ccg	tgc	acc	acc	atg	tct	acc	gcg	gtg	ctg	aat	ggt	ctg	336
Thr	Gly	Leu	Pro	Cys	Thr	Thr	Met	Ser	Thr	Ala	Val	Leu	Asn	Gly	Leu	
			100					105					110			

cgt	gcg	ctg	ggt	gtg	cgt	cgt	gtt	gcg	ctg	gcc	acc	gcg	tat	atc	gat	384
Arg	Ala	Leu	Gly	Val	Arg	Arg	Val	Ala	Leu	Ala	Thr	Ala	Tyr	Ile	Asp	
		115					120					125				

gat	gtg	aac	gaa	cgt	ctg	gcc	gcg	ttt	ctg	gcc	gaa	gaa	agc	ctg	gtg	432
Asp	Val	Asn	Glu	Arg	Leu	Ala	Ala	Phe	Leu	Ala	Glu	Glu	Ser	Leu	Val	
		130				135					140					

ccg	acc	ggt	tgt	cgt	agc	ctg	ggc	att	acc	ggc	gtg	gaa	gcg	atg	gcg	480
Pro	Thr	Gly	Cys	Arg	Ser	Leu	Gly	Ile	Thr	Gly	Val	Glu	Ala	Met	Ala	
145					150					155					160	

cgt gtg gat acc gcg acc ctg gtg gat ctg tgc gtg cgt gcg ttt gaa 528
 Arg Val Asp Thr Ala Thr Leu Val Asp Leu Cys Val Arg Ala Phe Glu
 165 170 175

gcg gca ccg gat agc gat ggc att ctg ctg tct tgc ggc ggt ctg ctg 576
 Ala Ala Pro Asp Ser Asp Gly Ile Leu Leu Ser Cys Gly Gly Leu Leu
 180 185 190

acc ctg gat gcg att ccg gaa gtg gaa cgt cgt ctg ggc gtg ccg gtg 624
 Thr Leu Asp Ala Ile Pro Glu Val Glu Arg Arg Leu Gly Val Pro Val
 195 200 205

gtt agc agc agc ccg gca ggc ttt tgg gat gcg gtg cgt ctg gcc ggt 672
 Val Ser Ser Ser Pro Ala Gly Phe Trp Asp Ala Val Arg Leu Ala Gly
 210 215 220

ggt ggt gcg aaa gcg cgt ccg ggc tat ggc cgt ctg ttt gat gaa agc 720
 Gly Gly Ala Lys Ala Arg Pro Gly Tyr Gly Arg Leu Phe Asp Glu Ser
 225 230 235 240

ctc gag cac cac cac cac cac cac 744
 Leu Glu His His His His His His
 245

<210> 2
 <211> 248
 <212> PRT
 <213> Bordetella bronchiseptica

<400> 2

Met Gln Gln Ala Ser Thr Pro Thr Ile Gly Met Ile Val Pro Pro Ala
 1 5 10 15

Ala Gly Leu Val Pro Ala Asp Gly Ala Arg Leu Tyr Pro Asp Leu Pro
 20 25 30

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

Asp Ala Val Ile Glu Ser Val Val Asp His Ala Arg Arg Leu Gln Lys
 50 55 60

Gln Gly Ala Ala Val Val Ser Leu Met Gly Thr Ser Leu Ser Phe Tyr
 65 70 75 80

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

Thr Gly Leu Pro Cys Thr Thr Met Ser Thr Ala Val Leu Asn Gly Leu
 100 105 110

Arg Ala Leu Gly Val Arg Arg Val Ala Leu Ala Thr Ala Tyr Ile Asp

115	120	125
Asp Val Asn Glu Arg Leu Ala Ala Phe Leu Ala Glu Glu Ser Leu Val 130 135 140		
Pro Thr Gly Cys Arg Ser Leu Gly Ile Thr Gly Val Glu Ala Met Ala 145 150 155 160		
Arg Val Asp Thr Ala Thr Leu Val Asp Leu Cys Val Arg Ala Phe Glu 165 170 175		
Ala Ala Pro Asp Ser Asp Gly Ile Leu Leu Ser Cys Gly Gly Leu Leu 180 185 190		
Thr Leu Asp Ala Ile Pro Glu Val Glu Arg Arg Leu Gly Val Pro Val 195 200 205		
Val Ser Ser Ser Pro Ala Gly Phe Trp Asp Ala Val Arg Leu Ala Gly 210 215 220		
Gly Gly Ala Lys Ala Arg Pro Gly Tyr Gly Arg Leu Phe Asp Glu Ser 225 230 235 240		
Leu Glu His His His His His His 245		

<210> 3
 <211> 747
 <212> DNA
 <213> Mesorhizobium sp.

<220>
 <221> CDS
 <222> (1)..(747)

<400> 3	
atg agc gcg gaa gat ccg gtg att ggc ctg att gtt ccg ccg gca gca Met Ser Ala Glu Asp Pro Val Ile Gly Leu Ile Val Pro Pro Ala Ala 1 5 10 15	48
ggc ctg gtt ccg ccg gaa gcg gtg ggc atg tat ccg gaa gtg acc ttt Gly Leu Val Pro Pro Glu Ala Val Ala Met Tyr Pro Glu Val Thr Phe 20 25 30	96
cat gcg agc ggc ctg ggc ctg aaa gaa atg acc ccg tgc ggc tat gcg His Ala Ser Gly Leu Gly Leu Lys Glu Met Thr Pro Cys Gly Tyr Ala 35 40 45	144
ggc gtg att gat ctg gtg ggc gat cat gcg gaa cgt cat gcg cag gcg Gly Val Ile Asp Leu Val Gly Asp His Ala Glu Arg His Ala Gln Ala 50 55 60	192

```

ggt gcg cag gcg gtg gcg ctg atg ggc acc agc ctg agc ttt ttt cgt      240
Gly Ala Gln Ala Val Ala Leu Met Gly Thr Ser Leu Ser Phe Phe Arg
65                      70                      75                      80

ggt gcg gcg ttt aac gcg gaa ctg att acc cac atg agc aac cgt agc      288
Gly Ala Ala Phe Asn Ala Glu Leu Ile Thr His Met Ser Asn Arg Ser
85                      90                      95

ggt ctg ccg gcg acc acc atg agc cag gcg gtg gtg gat gaa ctg aaa      336
Gly Leu Pro Ala Thr Thr Met Ser Gln Ala Val Val Asp Glu Leu Lys
100                     105                     110

agc cat ggc gcg cgt cgt att gcg gtg gtg acc gcg tat cgc cag gat      384
Ser His Gly Ala Arg Arg Ile Ala Val Val Thr Ala Tyr Arg Gln Asp
115                     120                     125

gtg aac aac ctg ctg gcc gcg ttt ctg aac gaa cat ggc att gaa gcg      432
Val Asn Asn Leu Leu Ala Ala Phe Leu Asn Glu His Gly Ile Glu Ala
130                     135                     140

cgt agc ctg aaa agc ctg ggc att acc agc gtg gcg gat gtt gcg ggc      480
Arg Ser Leu Lys Ser Leu Gly Ile Thr Ser Val Ala Asp Val Ala Gly
145                     150                     155                     160

acc ccg gcg aaa cgt ctg ctg cag ctg tgc aaa gaa gcg gtg cat gat      528
Thr Pro Ala Lys Arg Leu Leu Gln Leu Cys Lys Glu Ala Val His Asp
165                     170                     175

gcg ggt ccg gtg gat gcg gtg ctg att agc tgc ggc ggt ctg cat acc      576
Ala Gly Pro Val Asp Ala Val Leu Ile Ser Cys Gly Gly Leu His Thr
180                      185                      190

ctg gat gtg att gtg gat gtg gaa atc agc acc ggc ctg ccg gtt gtt      624
Leu Asp Val Ile Val Asp Val Glu Ile Ser Thr Gly Leu Pro Val Val
195                      200                      205

acc agc gcg acc gcg ggt gtg cgt ggt gcg gtg ggc ctg ctg aaa cgt      672
Thr Ser Ala Thr Ala Gly Val Arg Gly Ala Val Gly Leu Leu Lys Arg
210                     215                     220

agc gcg gaa acc gcg gaa cat gcg acc agc aaa ttt ctg acc ggc cgt      720
Ser Ala Glu Thr Ala Glu His Ala Thr Ser Lys Phe Leu Thr Gly Arg
225                     230                     235                     240

gcg ctc gag cac cac cac cac cac cac      747
Ala Leu Glu His His His His His His
245

```

```

<210> 4
<211> 249
<212> PRT
<213> Mesorhizobium sp.

```

```

<400> 4

```

```

Met Ser Ala Glu Asp Pro Val Ile Gly Leu Ile Val Pro Pro Ala Ala
1                      5                      10                      15

Gly Leu Val Pro Pro Glu Ala Val Ala Met Tyr Pro Glu Val Thr Phe
20                      25                      30

```

His Ala Ser Gly Leu Gly Leu Lys Glu Met Thr Pro Cys Gly Tyr Ala
 35 40 45

Gly Val Ile Asp Leu Val Gly Asp His Ala Glu Arg His Ala Gln Ala
 50 55 60

Gly Ala Gln Ala Val Ala Leu Met Gly Thr Ser Leu Ser Phe Phe Arg
 65 70 75 80

Gly Ala Ala Phe Asn Ala Glu Leu Ile Thr His Met Ser Asn Arg Ser
 85 90 95

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

Ser His Gly Ala Arg Arg Ile Ala Val Val Thr Ala Tyr Arg Gln Asp
 115 120 125

Val Asn Asn Leu Leu Ala Ala Phe Leu Asn Glu His Gly Ile Glu Ala
 130 135 140

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

Thr Pro Ala Lys Arg Leu Leu Gln Leu Cys Lys Glu Ala Val His Asp
 165 170 175

Ala Gly Pro Val Asp Ala Val Leu Ile Ser Cys Gly Gly Leu His Thr
 180 185 190

Leu Asp Val Ile Val Asp Val Glu Ile Ser Thr Gly Leu Pro Val Val
 195 200 205

Thr Ser Ala Thr Ala Gly Val Arg Gly Ala Val Gly Leu Leu Lys Arg
 210 215 220

Ser Ala Glu Thr Ala Glu His Ala Thr Ser Lys Phe Leu Thr Gly Arg
 225 230 235 240

Ala Leu Glu His His His His His His
 245

<210> 5
 <211> 738
 <212> DNA
 <213> Pyrococcus horikoshii

<220>

<221> CDS

<222> (1) .. (738)

<400> 5

atg tat ggc tgg cgt cgt cgt att ggc ctg att gtg ccg agc agc aac	48
Met Tyr Gly Trp Arg Arg Arg Ile Gly Leu Ile Val Pro Ser Ser Asn	
1 5 10 15	
acc acc atg gaa ccg gaa ttt tgg aaa atg gcg ccg gaa ggc gtg agc	96
Thr Thr Met Glu Pro Glu Phe Trp Lys Met Ala Pro Glu Gly Val Ser	
20 25 30	
att cat acc gcg cgt atg cgt ctg cgt gaa gtg acc gaa gaa agc ctg	144
Ile His Thr Ala Arg Met Arg Leu Arg Glu Val Thr Glu Glu Ser Leu	
35 40 45	
ctg gaa atg gaa cgt cat gcg aaa gat gcg gcg ctg gaa ctg gcc gat	192
Leu Glu Met Glu Arg His Ala Lys Asp Ala Ala Leu Glu Leu Ala Asp	
50 55 60	
gcg gaa gtg gat gtg att gtg tat ggc tgc acc agc ggc agc ctg att	240
Ala Glu Val Asp Val Ile Val Tyr Gly Cys Thr Ser Gly Ser Leu Ile	
65 70 75 80	
aaa ggc aaa ggc tat gat aaa gaa atc gcg aaa aac ctg gaa gaa gcg	288
Lys Gly Lys Gly Tyr Asp Lys Glu Ile Ala Lys Asn Leu Glu Glu Ala	
85 90 95	
agc ggc att aaa acc att acc acc tct acc gcg gtt ctg aac gcg ctg	336
Ser Gly Ile Lys Thr Ile Thr Thr Ser Thr Ala Val Leu Asn Ala Leu	
100 105 110	
aac acc ctg gat att cag aaa gtg gtg gtg gcg acc ccg tat att gat	384
Asn Thr Leu Asp Ile Gln Lys Val Val Val Ala Thr Pro Tyr Ile Asp	
115 120 125	
agc gtg aac gaa cgc gaa aaa gaa ttt ctg gaa gcg aac ggc att gaa	432
Ser Val Asn Glu Arg Glu Lys Glu Phe Leu Glu Ala Asn Gly Ile Glu	
130 135 140	
gtg ctg cgt att aaa ggc ctg ggc att gtg aaa aac acc gaa att ggc	480
Val Leu Arg Ile Lys Gly Leu Gly Ile Val Lys Asn Thr Glu Ile Gly	
145 150 155 160	
cgt cag gaa ccg tgg gtt gcg tat aaa ctg gcc ctg gaa gtg tat acc	528
Arg Gln Glu Pro Trp Val Ala Tyr Lys Leu Ala Leu Glu Val Tyr Thr	
165 170 175	
ccg gaa gcg gat ggc ctg ttt att agc tgc acc aac ttt cgc acc att	576
Pro Glu Ala Asp Gly Leu Phe Ile Ser Cys Thr Asn Phe Arg Thr Ile	
180 185 190	
gaa atc atc gat aaa ctg gaa gtg gaa ctg ggc gtg ccg gtg gtg acc	624
Glu Ile Ile Asp Lys Leu Glu Val Glu Leu Gly Val Pro Val Val Thr	
195 200 205	
agc aac cag gcg acc atg tgg tat gcg ctg aaa acc atg aaa gtg aaa	672
Ser Asn Gln Ala Thr Met Trp Tyr Ala Leu Lys Thr Met Lys Val Lys	
210 215 220	

gaa aaa tat gat atg tac ggc acc ctg ctg aaa gaa aaa ctg ctc gag 720
 Glu Lys Tyr Asp Met Tyr Gly Thr Leu Leu Lys Glu Lys Leu Leu Glu
 225 230 235 240

cac cac cac cac cac cac 738
 His His His His His His
 245

<210> 6
 <211> 246
 <212> PRT
 <213> Pyrococcus horikoshii

<400> 6

Met Tyr Gly Trp Arg Arg Arg Ile Gly Leu Ile Val Pro Ser Ser Asn
 1 5 10 15

Thr Thr Met Glu Pro Glu Phe Trp Lys Met Ala Pro Glu Gly Val Ser
 20 25 30

Ile His Thr Ala Arg Met Arg Leu Arg Glu Val Thr Glu Glu Ser Leu
 35 40 45

Leu Glu Met Glu Arg His Ala Lys Asp Ala Ala Leu Glu Leu Ala Asp
 50 55 60

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

Lys Gly Lys Gly Tyr Asp Lys Glu Ile Ala Lys Asn Leu Glu Glu Ala
 85 90 95

Ser Gly Ile Lys Thr Ile Thr Thr Ser Thr Ala Val Leu Asn Ala Leu
 100 105 110

Asn Thr Leu Asp Ile Gln Lys Val Val Val Ala Thr Pro Tyr Ile Asp
 115 120 125

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

Val Leu Arg Ile Lys Gly Leu Gly Ile Val Lys Asn Thr Glu Ile Gly
 145 150 155 160

Arg Gln Glu Pro Trp Val Ala Tyr Lys Leu Ala Leu Glu Val Tyr Thr
 165 170 175

Pro Glu Ala Asp Gly Leu Phe Ile Ser Cys Thr Asn Phe Arg Thr Ile

180										185										190									
Glu	Ile	Ile	Asp	Lys	Leu	Glu	Val	Glu	Leu	Gly	Val	Pro	Val	Val	Thr														
		195					200					205																	
Ser	Asn	Gln	Ala	Thr	Met	Trp	Tyr	Ala	Leu	Lys	Thr	Met	Lys	Val	Lys														
	210					215					220																		
Glu	Lys	Tyr	Asp	Met	Tyr	Gly	Thr	Leu	Leu	Lys	Glu	Lys	Leu	Leu	Glu														
225					230					235					240														
His	His	His	His	His	His																								
				245																									
<210> 7																													
<211> 819																													
<212> DNA																													
<213> Bordetella bronchiseptica																													
<220>																													
<221> CDS																													
<222> (1)..(819)																													
<400> 7																													
atg	ccg	acc	ccg	acc	cag	atg	aaa	acc	acc	gcg	cgt	tat	cgt	ccg	tat				48										
Met	Pro	Thr	Pro	Thr	Gln	Met	Lys	Thr	Thr	Ala	Arg	Tyr	Arg	Pro	Tyr														
1				5					10					15															
gcg	tgg	cgt	gcg	aaa	gtg	ggc	ctg	att	gtg	ccg	agc	acc	aac	acc	att				96										
Ala	Trp	Arg	Ala	Lys	Val	Gly	Leu	Ile	Val	Pro	Ser	Thr	Asn	Thr	Ile														
			20				25						30																
aac	gaa	ccg	gaa	ttc	tat	cgt	ctg	gcc	ccg	gat	ggc	gtg	acc	att	cat				144										
Asn	Glu	Pro	Glu	Phe	Tyr	Arg	Leu	Ala	Pro	Asp	Gly	Val	Thr	Ile	His														
		35					40					45																	
acc	ggc	cgt	gcg	att	aat	gcg	ggt	ccg	gcg	acc	caa	gaa	aac	tat	gat				192										
Thr	Gly	Arg	Ala	Ile	Asn	Ala	Gly	Pro	Ala	Thr	Gln	Glu	Asn	Tyr	Asp														
	50					55					60																		
cgt	atg	gcg	cgt	ggc	gtt	ctg	gaa	gcg	gcg	gat	ctg	att	aaa	acc	gcg				240										
Arg	Met	Ala	Arg	Gly	Val	Leu	Glu	Ala	Ala	Asp	Leu	Ile	Lys	Thr	Ala														
65					70				75						80														
gaa	gtg	gat	gtg	gtg	gcg	tat	ggc	tgc	acc	agc	ggc	agc	att	gtg	tgc				288										
Glu	Val	Asp	Val	Val	Ala	Tyr	Gly	Cys	Thr	Ser	Gly	Ser	Ile	Val	Cys														
			85					90					95																
agc	ctg	gat	gaa	ctg	tgc	gat	ggc	atg	agc	gaa	cgt	gtg	ggc	gtt	ccg				336										
Ser	Leu	Asp	Glu	Leu	Cys	Asp	Gly	Met	Ser	Glu	Arg	Val	Gly	Val	Pro														
			100					105					110																
gcg	att	gcg	acc	gcg	ggt	gcg	gtt	gtt	gcg	gcg	ctg	cgt	gcg	ctg	ggt				384										
Ala	Ile	Ala	Thr	Ala	Gly	Ala	Val	Val	Ala	Ala	Leu	Arg	Ala	Leu	Gly														
		115					120					125																	

gtt cgt cgt gtg gcg gtg ggc acc ccg tat att gat ttt gtg aac cag Val Arg Arg Val Ala Val Gly Thr Pro Tyr Ile Asp Phe Val Asn Gln 130 135 140	432
cgt gaa cgt gaa ttt ctg gaa cag tat ggc ttt gaa gtg gtg agc att Arg Glu Arg Glu Phe Leu Glu Gln Tyr Gly Phe Glu Val Val Ser Ile 145 150 155 160	480
gaa ggc atg gat ctg ggc cat acg cag gaa gaa cgt cgc gat att ggc Glu Gly Met Asp Leu Gly His Thr Gln Glu Glu Arg Arg Asp Ile Gly 165 170 175	528
cgt gtg ccg ccg cag agc acc ttt cag ctg gcc cgt gat att gat cgt Arg Val Pro Pro Gln Ser Thr Phe Gln Leu Ala Arg Asp Ile Asp Arg 180 185 190	576
ccg cag gcg gaa gcg att ttt ctg agc tgc acc aac ctg gcc acc ctg Pro Gln Ala Glu Ala Ile Phe Leu Ser Cys Thr Asn Leu Ala Thr Leu 195 200 205	624
gat atg att gaa cgc atc gaa aac gaa ctg ggc aaa ccg gtg gtg acc Asp Met Ile Glu Arg Ile Glu Asn Glu Leu Gly Lys Pro Val Val Thr 210 215 220	672
agc aac cag gcg acc ttt tgg gcg tgc ctg cgt ctg ctg ggt ctg agc Ser Asn Gln Ala Thr Phe Trp Ala Cys Leu Arg Leu Leu Gly Leu Ser 225 230 235 240	720
tgc gcg att ccg ggc tat ggc cgt ctg ctg acc gat tgc ctg gcc ccg Cys Ala Ile Pro Gly Tyr Gly Arg Leu Leu Thr Asp Cys Leu Ala Pro 245 250 255	768
att acc cgt gcg agc tat gcg ctg gcc ctc gag cac cac cac cac Ile Thr Arg Ala Ser Tyr Ala Leu Ala Leu Glu His His His His His 260 265 270	816
cac His	819

<210> 8
 <211> 273
 <212> PRT
 <213> Bordetella bronchiseptica

<400> 8

Met Pro Thr Pro Thr Gln Met Lys Thr Thr Ala Arg Tyr Arg Pro Tyr 1 5 10 15
Ala Trp Arg Ala Lys Val Gly Leu Ile Val Pro Ser Thr Asn Thr Ile 20 25 30
Asn Glu Pro Glu Phe Tyr Arg Leu Ala Pro Asp Gly Val Thr Ile His 35 40 45
Thr Gly Arg Ala Ile Asn Ala Gly Pro Ala Thr Gln Glu Asn Tyr Asp 50 55 60

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

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

Ser Leu Asp Glu Leu Cys Asp Gly Met Ser Glu Arg Val Gly Val Pro
100 105 110

Ala Ile Ala Thr Ala Gly Ala Val Val Ala Ala Leu Arg Ala Leu Gly
115 120 125

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

Arg Glu Arg Glu Phe Leu Glu Gln Tyr Gly Phe Glu Val Val Ser Ile
145 150 155 160

Glu Gly Met Asp Leu Gly His Thr Gln Glu Glu Arg Arg Asp Ile Gly
165 170 175

Arg Val Pro Pro Gln Ser Thr Phe Gln Leu Ala Arg Asp Ile Asp Arg
180 185 190

Pro Gln Ala Glu Ala Ile Phe Leu Ser Cys Thr Asn Leu Ala Thr Leu
195 200 205

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

Ser Asn Gln Ala Thr Phe Trp Ala Cys Leu Arg Leu Leu Gly Leu Ser
225 230 235 240

Cys Ala Ile Pro Gly Tyr Gly Arg Leu Leu Thr Asp Cys Leu Ala Pro
245 250 255

Ile Thr Arg Ala Ser Tyr Ala Leu Ala Leu Glu His His His His His
260 265 270

His

<210> 9
<211> 756
<212> DNA
<213> Alkaliphilus metalliredigenes

<220>

<221> CDS

<222> (1) .. (756)

<400> 9

atg	agc	gat	ttt	agc	tat	ggc	ggc	cgt	gcg	aaa	att	ggc	ctg	att	tat	48
Met	Ser	Asp	Phe	Ser	Tyr	Gly	Gly	Arg	Ala	Lys	Ile	Gly	Leu	Ile	Tyr	
1				5					10					15		

ccg	gca	ccg	ggc	tgg	gtt	atg	gaa	ccg	gaa	ttc	tat	ctg	atg	gcg	ccg	96
Pro	Ala	Pro	Gly	Trp	Val	Met	Glu	Pro	Glu	Phe	Tyr	Leu	Met	Ala	Pro	
			20					25					30			

tgg	ggc	gtg	agc	acc	tat	acc	acc	cgt	att	agc	ctg	aaa	cat	gtg	aac	144
Trp	Gly	Val	Ser	Thr	Tyr	Thr	Thr	Arg	Ile	Ser	Leu	Lys	His	Val	Asn	
		35					40					45				

gtg	gcg	gaa	ctg	cgt	aaa	ctg	ggc	gat	cag	agc	gtg	gaa	gcg	gcg	gaa	192
Val	Ala	Glu	Leu	Arg	Lys	Leu	Gly	Asp	Gln	Ser	Val	Glu	Ala	Ala	Glu	
	50					55					60					

ctg	ctg	gcc	cag	gcg	ccg	ctg	gat	gtg	att	gcg	ctg	ggc	tgc	acc	agc	240
Leu	Leu	Ala	Gln	Ala	Pro	Leu	Asp	Val	Ile	Ala	Leu	Gly	Cys	Thr	Ser	
65					70					75					80	

ggc	agc	ttc	gtg	aac	ggc	agc	cgt	tat	gat	caa	gaa	ctg	atc	gaa	aaa	288
Gly	Ser	Phe	Val	Asn	Gly	Ser	Arg	Tyr	Asp	Gln	Glu	Leu	Ile	Glu	Lys	
				85				90						95		

atg	gaa	ggc	gtg	agc	ggc	ggt	att	ccg	tgc	acc	acc	acc	tct	acc	gcg	336
Met	Glu	Gly	Val	Ser	Gly	Gly	Ile	Pro	Cys	Thr	Thr	Thr	Ser	Thr	Ala	
			100					105						110		

gtt	gtg	gcg	gcg	ctg	aaa	gcg	ctg	cat	gtg	acc	aaa	att	gcg	gtg	gcg	384
Val	Val	Ala	Ala	Leu	Lys	Ala	Leu	His	Val	Thr	Lys	Ile	Ala	Val	Ala	
		115					120					125				

acc	ccg	tat	atc	gat	gaa	gtg	aac	ctg	aaa	gcg	aaa	aac	ttt	ctg	gaa	432
Thr	Pro	Tyr	Ile	Asp	Glu	Val	Asn	Leu	Lys	Ala	Lys	Asn	Phe	Leu	Glu	
	130					135					140					

gcg	gaa	ggc	ctg	gat	gtg	gtg	aac	att	aaa	ggc	ctg	ggc	ctg	ctg	cag	480
Ala	Glu	Gly	Leu	Asp	Val	Val	Asn	Ile	Lys	Gly	Leu	Gly	Leu	Leu	Gln	
145					150					155					160	

gat	att	gaa	att	gat	cgc	cag	gat	atg	gaa	acc	gtg	tat	cgt	ctg	gcc	528
Asp	Ile	Glu	Ile	Asp	Arg	Gln	Asp	Met	Glu	Thr	Val	Tyr	Arg	Leu	Ala	
				165				170						175		

aaa	gaa	gtg	gat	cac	gtt	gaa	gcg	cag	gcg	att	gtg	att	ctg	tgc	acc	576
Lys	Glu	Val	Asp	His	Val	Glu	Ala	Gln	Ala	Ile	Val	Ile	Leu	Cys	Thr	
			180					185					190			

ggc	ctg	cgt	agc	gtg	ccg	att	att	gaa	gcg	ctg	gaa	aac	gat	ctg	ggc	624
Gly	Leu	Arg	Ser	Val	Pro	Ile	Ile	Glu	Ala	Leu	Glu	Asn	Asp	Leu	Gly	
		195				200						205				

aaa	ccg	gtg	att	agc	gcg	att	cag	gcg	acc	ttt	tgg	cat	tgc	ctg	cgt	672
Lys	Pro	Val	Ile	Ser	Ala	Ile	Gln	Ala	Thr	Phe	Trp	His	Cys	Leu	Arg	
	210					215					220					

ctg agc ggc gtg cgt gaa aac gtg gaa ggc tat ggc agc ctg ctg cgt 720
 Leu Ser Gly Val Arg Glu Asn Val Glu Gly Tyr Gly Ser Leu Leu Arg
 225 230 235 240

att gaa ggc ctg ctc gag cac cac cac cac cac cac 756
 Ile Glu Gly Leu Leu Glu His His His His His His
 245 250

<210> 10
 <211> 252
 <212> PRT
 <213> Alkaliphilus metalliredigenes

<400> 10

Met Ser Asp Phe Ser Tyr Gly Gly Arg Ala Lys Ile Gly Leu Ile Tyr
 1 5 10 15

Pro Ala Pro Gly Trp Val Met Glu Pro Glu Phe Tyr Leu Met Ala Pro
 20 25 30

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

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

Leu Leu Ala Gln Ala Pro Leu Asp Val Ile Ala Leu Gly Cys Thr Ser
 65 70 75 80

Gly Ser Phe Val Asn Gly Ser Arg Tyr Asp Gln Glu Leu Ile Glu Lys
 85 90 95

Met Glu Gly Val Ser Gly Gly Ile Pro Cys Thr Thr Thr Ser Thr Ala
 100 105 110

Val Val Ala Ala Leu Lys Ala Leu His Val Thr Lys Ile Ala Val Ala
 115 120 125

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

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

Asp Ile Glu Ile Asp Arg Gln Asp Met Glu Thr Val Tyr Arg Leu Ala
 165 170 175

Lys Glu Val Asp His Val Glu Ala Gln Ala Ile Val Ile Leu Cys Thr

180								185				190			
Gly	Leu	Arg	Ser	Val	Pro	Ile	Ile	Glu	Ala	Leu	Glu	Asn	Asp	Leu	Gly
		195					200					205			
Lys	Pro	Val	Ile	Ser	Ala	Ile	Gln	Ala	Thr	Phe	Trp	His	Cys	Leu	Arg
	210					215					220				
Leu	Ser	Gly	Val	Arg	Glu	Asn	Val	Glu	Gly	Tyr	Gly	Ser	Leu	Leu	Arg
225					230					235					240
Ile	Glu	Gly	Leu	Leu	Glu	His	His	His	His	His	His				
				245						250					

<210> 11
 <211> 786
 <212> DNA
 <213> Aquifex pyrophilus

<220>
 <221> CDS
 <222> (1)..(786)

<400> 11	
atg aaa atc ggc att ttt gat agc ggc gtg ggc ggt ctg acc gtg ctg	48
Met Lys Ile Gly Ile Phe Asp Ser Gly Val Gly Gly Leu Thr Val Leu	
1 5 10 15	
aaa gcg att cgt aac cgt tat cgc aaa gtg gat att gtg tat ctg ggc	96
Lys Ala Ile Arg Asn Arg Tyr Arg Lys Val Asp Ile Val Tyr Leu Gly	
20 25 30	
gat acc gcg cgt gtg ccg tat ggc att cgt agc aaa gat acc att atc	144
Asp Thr Ala Arg Val Pro Tyr Gly Ile Arg Ser Lys Asp Thr Ile Ile	
35 40 45	
cgc tat agc ctg gaa tgc gcg ggc ttt ctg aaa gat aaa ggc gtg gat	192
Arg Tyr Ser Leu Glu Cys Ala Gly Phe Leu Lys Asp Lys Gly Val Asp	
50 55 60	
att att gtg gtg gcg tgt aac acc gcg agc gcg tat gcg ctg gaa cgt	240
Ile Ile Val Val Ala Cys Asn Thr Ala Ser Ala Tyr Ala Leu Glu Arg	
65 70 75 80	
ctg aaa aaa gag atc aac gtt ccg gtt ttt ggc gtg att gaa ccg ggc	288
Leu Lys Lys Glu Ile Asn Val Pro Val Phe Gly Val Ile Glu Pro Gly	
85 90 95	
gtg aaa gaa gcg ctg aaa aaa agc cgc aac aaa aaa att ggt gtg att	336
Val Lys Glu Ala Leu Lys Lys Ser Arg Asn Lys Lys Ile Gly Val Ile	
100 105 110	
ggc acc ccg gcg acc gtt aaa agc ggc gcg tat cag cgt aaa ctg gaa	384
Gly Thr Pro Ala Thr Val Lys Ser Gly Ala Tyr Gln Arg Lys Leu Glu	
115 120 125	

gaa ggc ggt gcg gat gtt ttt gcg aaa gcg tgc ccg ctg ttt gtg ccg 432
 Glu Gly Gly Ala Asp Val Phe Ala Lys Ala Cys Pro Leu Phe Val Pro
 130 135 140

ctg gcc gaa gaa ggc ctg ctg gaa ggc gaa att acc cgt aaa gtg gtg 480
 Leu Ala Glu Glu Gly Leu Leu Glu Gly Glu Ile Thr Arg Lys Val Val
 145 150 155 160

gaa cac tat ctg aaa gaa ttc aaa ggc aaa atc gat acc ctg att ctg 528
 Glu His Tyr Leu Lys Glu Phe Lys Gly Lys Ile Asp Thr Leu Ile Leu
 165 170 175

ggc tgc acg cac tat ccg ctg ctg aaa aaa gaa atc aaa aaa ttc ctg 576
 Gly Cys Thr His Tyr Pro Leu Leu Lys Lys Glu Ile Lys Lys Phe Leu
 180 185 190

ggc gat gtg gaa gtg gtg gat agc agc gaa gcc ctg agc ctg agc ctg 624
 Gly Asp Val Glu Val Val Asp Ser Ser Glu Ala Leu Ser Leu Ser Leu
 195 200 205

cat aac ttc att aaa gat gat ggc agc agc agc ctg gaa ctg ttt ttt 672
 His Asn Phe Ile Lys Asp Asp Gly Ser Ser Ser Leu Glu Leu Phe Phe
 210 215 220

acc gat ctg agc ccg aac ctg cag ttt ctg att aaa ctg att ctg ggt 720
 Thr Asp Leu Ser Pro Asn Leu Gln Phe Leu Ile Lys Leu Ile Leu Gly
 225 230 235 240

cgt gat tat ccg gtg aaa ctg gcc gaa ggc gtg ttt acc cat ctc gag 768
 Arg Asp Tyr Pro Val Lys Leu Ala Glu Gly Val Phe Thr His Leu Glu
 245 250 255

cac cac cac cac cac cac 786
 His His His His His His
 260

<210> 12
 <211> 262
 <212> PRT
 <213> Aquifex pyrophilus

<400> 12

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

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

Asp Thr Ala Arg Val Pro Tyr Gly Ile Arg Ser Lys Asp Thr Ile Ile
 35 40 45

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

Ile Ile Val Val Ala Cys Asn Thr Ala Ser Ala Tyr Ala Leu Glu Arg
 65 70 75 80

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

Val Lys Glu Ala Leu Lys Lys Ser Arg Asn Lys Lys Ile Gly Val Ile
100 105 110

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

Glu Gly Gly Ala Asp Val Phe Ala Lys Ala Cys Pro Leu Phe Val Pro
130 135 140

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

Glu His Tyr Leu Lys Glu Phe Lys Gly Lys Ile Asp Thr Leu Ile Leu
165 170 175

Gly Cys Thr His Tyr Pro Leu Leu Lys Lys Glu Ile Lys Lys Phe Leu
180 185 190

Gly Asp Val Glu Val Val Asp Ser Ser Glu Ala Leu Ser Leu Ser Leu
195 200 205

His Asn Phe Ile Lys Asp Asp Gly Ser Ser Ser Leu Glu Leu Phe Phe
210 215 220

Thr Asp Leu Ser Pro Asn Leu Gln Phe Leu Ile Lys Leu Ile Leu Gly
225 230 235 240

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

His His His His His His
260

<210> 13
<211> 4
<212> PRT
<213> Artificial

<220>
<223> CS sequence pattern

<220>
<221> MISC_FEATURE
<222> (1)..(1)

<223> Xaa is Ser or Leu

<220>

<221> MISC_FEATURE

<222> (3)..(3)

<223> Xaa is Gly or Thr

<220>

<221> MISC_FEATURE

<222> (4)..(4)

<223> Xaa is Gyl or Asn

<400> 13

Xaa Cys Xaa Xaa

1

<210> 14

<211> 5

<212> PRT

<213> Artificial

<220>

<223> OAH1 sequence pattern

<220>

<221> MISC_FEATURE

<222> (1)..(1)

<223> Xaa is Met or Gly

<220>

<221> MISC_FEATURE

<222> (2)..(2)

<223> Xaa is Cys or Gly

<220>

<221> MISC_FEATURE

<222> (5)..(5)

<223> Xaa is Leu or Gly

<400> 14

Xaa Xaa Thr Ser Xaa

1

5

<210> 15

<211> 5

<212> PRT

<213> Artificial

<220>

<223> OAH2 sequence pattern

<220>

<221> MISC_FEATURE

<222> (1)..(1)

<223> Thr is optionally missing

<220>
 <221> MISC_FEATURE
 <222> (2)..(2)
 <223> Xaa is Ala or Pro

 <220>
 <221> MISC_FEATURE
 <222> (4)..(4)
 <223> Xaa is Ile or Arg

 <220>
 <221> MISC_FEATURE
 <222> (5)..(5)
 <223> Xaa is Asp or Gln or missing

 <400> 15

Thr Xaa Tyr Xaa Xaa
 1 5

<210> 16
 <211> 6
 <212> PRT
 <213> Artificial

<220>
 <223> LBP sequence pattern

<220>
 <221> MISC_FEATURE
 <222> (4)..(4)
 <223> Xaa is Phe, Ser or Ala

<220>
 <221> MISC_FEATURE
 <222> (5)..(5)
 <223> Xaa is Phe, Ser, Thr or Ala

<220>
 <221> MISC_FEATURE
 <222> (6)..(6)
 <223> Xaa is Ala, Asn or Gly

<400> 16

Ile Val Pro Xaa Xaa Xaa
 1 5

<210> 17
 <211> 27
 <212> DNA
 <213> Artificial

<220>
 <223> forward primer

<400> 17
 gagcggcctg ggcgcgggta gcgttac

<210> 18
<211> 27
<212> DNA
<213> Artificial

<220>
<223> Reverse Primer

<400> 18
gtaacgctac ccgcgcccag gccgctc 27

<210> 19
<211> 25
<212> DNA
<213> Artificial

<220>
<223> Forward Primer

<400> 19
cctgggtagc gcaaccccg aaggc 25

<210> 20
<211> 25
<212> DNA
<213> Artificial

<220>
<223> Reverse Primer

<400> 20
gccttcgagg gttgcgctac ccagg 25

<210> 21
<211> 25
<212> DNA
<213> Artificial

<220>
<223> Forward Primer

<400> 21
gcattaccgg cgcggaagcg atggc 25

<210> 22
<211> 25
<212> DNA
<213> Artificial

<220>
<223> Reverse Primer

<400> 22
gccatcgctt ccgcgccggt aatgc 25

<210> 23

<211> 32
 <212> DNA
 <213> Artificial
 <220>
 <223> PCR Primer
 <400> 23
 cattggcatg attgtggcgc cggcagcggg tc 32

<210> 24
 <211> 32
 <212> DNA
 <213> Artificial
 <220>
 <223> PCR Primer
 <400> 24
 gaccgcgtgc cggcgccaca atcatgcca tg 32

<210> 25
 <211> 34
 <212> DNA
 <213> Artificial
 <220>
 <223> PCR primer
 <400> 25
 cattggcatg attgtgggtc cggcagcggg tctg 34

<210> 26
 <211> 34
 <212> DNA
 <213> Artificial
 <220>
 <223> PCR Primer
 <400> 26
 cagaccgct gccggaccca caatcatgcc aatg 34

<210> 27
 <211> 27
 <212> DNA
 <213> Artificial
 <220>
 <223> PCR primer
 <400> 27
 ctgctgtctt gcgcgggtct gctgacc 27

<210> 28
 <211> 27
 <212> DNA

<213> Artificial

<220>

<223> PCR primer

<400> 28

ggtcagcaga cccgcgcaag acagcag

27

<210> 29

<211> 34

<212> DNA

<213> Artificial

<220>

<223> PCR primer

<400> 29

ctgctgtctt gcggcgcgct gctgaccctg gatg

34

<210> 30

<211> 34

<212> DNA

<213> Artificial

<220>

<223> PCR primer

<400> 30

catccagggt cagcagcgcg ccgcaagaca gcag

34

<210> 31

<211> 38

<212> DNA

<213> Artificial

<220>

<223> PCR primer

<400> 31

cattctgctg tcttgccgcg cgctgctgac cctggatg

38

<210> 32

<211> 38

<212> DNA

<213> Artificial

<220>

<223> PCR primer

<400> 32

catccagggt cagcagcgcc gcgcaagaca gcagaatg

38

<210> 33

<211> 45

<212> DNA

<213> Artificial

<220>

<223> PCR primer

<220>

<221> misc_feature

<222> (21)..(21)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (24)..(24)

<223> n is a, c, g, or t

<400> 33

ccccgaccat tggcatgatt ndtndtccgg cagcgggtct ggttc

45

<210> 34

<211> 45

<212> DNA

<213> Artificial

<220>

<223> PCR primer

<220>

<221> misc_feature

<222> (22)..(22)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (25)..(25)

<223> n is a, c, g, or t

<400> 34

gaaccagacc cgctgccgga hnahnaatca tgccaatggt cgggg

45

<210> 35

<211> 45

<212> DNA

<213> Artificial

<220>

<223> PCR primer

<220>

<221> misc_feature

<222> (21)..(21)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (24)..(24)

<223> n is a, c, g, or t

<400> 35

cgaccattgg catgattgtg ndtndtgcag cgggtctggt tccgg

45

<210> 36
 <211> 45
 <212> DNA
 <213> Artificial

<220>
 <223> PCR primer

<220>
 <221> misc_feature
 <222> (22)..(22)
 <223> n is a, c, g, or t

<220>
 <221> misc_feature
 <222> (25)..(25)
 <223> n is a, c, g, or t

<400> 36
 ccggaaccag acccgctgca hnahncacaa tcatgccaat ggtcg