

<110> CJ CheilJedang Corporation  
 <120> Novel citrate synthase variant and a method for producing  
 O-Acetyl-L-Homoserine or L-Methionine using same  
 <130> OPA21399-PCT  
 <150> KR 10-2021-0031643  
 <151> 2021-03-10  
 <160> 33  
 <170> KoPatentIn 3.0  
 <210> 1  
 <211> 437  
 <212> PRT  
 <213> Unknown  
 <220>  
 <223> ATCC 13032 GltA AA

<400> 1  
 Met Phe Glu Arg Asp Ile Val Ala Thr Asp Asn Asn Lys Ala Val Leu  
 1 5 10 15  
 His Tyr Pro Gly Gly Glu Phe Glu Met Asp Ile Ile Glu Ala Ser Glu  
 20 25 30  
 Gly Asn Asn Gly Val Val Leu Gly Lys Met Leu Ser Glu Thr Gly Leu  
 35 40 45  
 Ile Thr Phe Asp Pro Gly Tyr Val Ser Thr Gly Ser Thr Glu Ser Lys  
 50 55 60  
 Ile Thr Tyr Ile Asp Gly Asp Ala Gly Ile Leu Arg Tyr Arg Gly Tyr  
 65 70 75 80  
 Asp Ile Ala Asp Leu Ala Glu Asn Ala Thr Phe Asn Glu Val Ser Tyr  
 85 90 95  
 Leu Leu Ile Asn Gly Glu Leu Pro Thr Pro Asp Glu Leu His Lys Phe  
 100 105 110  
 Asn Asp Glu Ile Arg His His Thr Leu Leu Asp Glu Asp Phe Lys Ser  
 115 120 125  
 Gln Phe Asn Val Phe Pro Arg Asp Ala His Pro Met Ala Thr Leu Ala  
 130 135 140

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

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|     | 405 |     | 410 |     | 415 |     |     |     |     |     |     |     |     |     |     |
| Asn | Arg | Pro | Arg | Gln | Val | Tyr | Thr | Gly | Asn | Glu | Ser | Arg | Lys | Leu | Val |
|     | 420 |     |     |     |     |     |     | 425 |     |     |     |     | 430 |     |     |

Pro Arg Glu Glu Arg  
435

<210> 2  
<211> 1314  
<212> DNA  
<213> Unknown

<220>  
<223> ATCC13032 GltA NT

<400> 2

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| atgtttgaaa gggatatcgt ggctactgat aacaacaagg ctgtcctgca ctaccccgg  | 60  |
| ggcgagtctg aaatggacat catcgaggct tctgagggtg acaacggtgt tgtcctgggc | 120 |
| aagatgctgt ctgagactgg actgatcact tttgaccag gttatgtgag cactggctcc  | 180 |
| accgagtcga agatcaccta catcgatggc gatgcgggaa tcctgcgtta ccgcggctat | 240 |
| gacatcgctg atctggctga gaatgccacc ttcaacgagg tttcttacct acttatcaac | 300 |
| ggtgagctac caaccccaga tgagcttcac aagtttaacg acgagattcg ccaccacacc | 360 |
| cttctggacg aggacttcaa gtcccagttc aacgtgttcc cacgcgacgc tcaccaatg  | 420 |
| gcaaccttgg cttcctcggg taacattttg tctacctact accaggacca gctgaaccca | 480 |
| ctcgatgagg cacagcttga taaggcaacc gttcgcctca tggcaaaggt tccaatgctg | 540 |
| gctgcgtacg cacaccgcgc acgcaagggt gtccttaca tgtaccaga caactccctc   | 600 |
| aatgcgcgtg agaacttcct gcgcatgatg ttcggttacc caaccgagcc atacgagatc | 660 |
| gaccaatca tggtaaggc tctggacaag ctgctcatcc tgcacgctga ccacgagcag   | 720 |
| aactgctcca cctccaccgt tcgtatgatc ggttccgcac aggccaacat gtttgtctcc | 780 |
| atcgctggtg gcatcaacgc tctgtccggc cactgcacg gtggcgcaaa ccaggctggt  | 840 |
| ctggagatgc tcgaagacat caagagcaac cacggtggcg acgcaaccga gttcatgaac | 900 |
| aaggtaaga acaaggaaga cggcgtccgc ctcatgggct tcggacaccg cgttttacaag | 960 |

aactacgatac cacgtgcagc aatcgtcaag gagaccgcac acgagatcct cgagcacctc 1020  
ggtggcgacg atcttctgga tctggcaatc aagctggaag aaattgcact ggctgatgat 1080  
tacttcatct cccgcaagct ctaccgaac gtagacttct acaccggcct gatctaccgc 1140  
gcaatgggct tcccaactga cttcttcacc gtattgttcg caatcggtcg tctgccagga 1200  
tggatcgctc actaccgcga gcagctcggg gcagcaggca acaagatcaa ccgcccacgc 1260  
caggtctaca ccggcaacga atcccgcaag ttggttcctc gcgaggagcg ctaa 1314

<210> 3  
<211> 54  
<212> PRT  
<213> Unknown

<220>  
<223> ATCC 13032 GltA K415H 362~415

<400> 3  
Phe Ile Ser Arg Lys Leu Tyr Pro Asn Val Asp Phe Tyr Thr Gly Leu  
1 5 10 15  
Ile Tyr Arg Ala Met Gly Phe Pro Thr Asp Phe Phe Thr Val Leu Phe  
20 25 30  
Ala Ile Gly Arg Leu Pro Gly Trp Ile Ala His Tyr Arg Glu Gln Leu  
35 40 45  
Gly Ala Ala Gly Asn His  
50

<210> 4  
<211> 437  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> ATCC13032 GltA K415H AA

<400> 4  
Met Phe Glu Arg Asp Ile Val Ala Thr Asp Asn Asn Lys Ala Val Leu  
1 5 10 15

His Tyr Pro Gly Gly Glu Phe Glu Met Asp Ile Ile Glu Ala Ser Glu  
 20 25 30  
 Gly Asn Asn Gly Val Val Leu Gly Lys Met Leu Ser Glu Thr Gly Leu  
 35 40 45  
 Ile Thr Phe Asp Pro Gly Tyr Val Ser Thr Gly Ser Thr Glu Ser Lys  
 50 55 60  
 Ile Thr Tyr Ile Asp Gly Asp Ala Gly Ile Leu Arg Tyr Arg Gly Tyr  
 65 70 75 80  
 Asp Ile Ala Asp Leu Ala Glu Asn Ala Thr Phe Asn Glu Val Ser Tyr  
 85 90 95  
 Leu Leu Ile Asn Gly Glu Leu Pro Thr Pro Asp Glu Leu His Lys Phe  
 100 105 110  
 Asn Asp Glu Ile Arg His His Thr Leu Leu Asp Glu Asp Phe Lys Ser  
 115 120 125  
 Gln Phe Asn Val Phe Pro Arg Asp Ala His Pro Met Ala Thr Leu Ala  
 130 135 140  
 Ser Ser Val Asn Ile Leu Ser Thr Tyr Tyr Gln Asp Gln Leu Asn Pro  
 145 150 155 160  
 Leu Asp Glu Ala Gln Leu Asp Lys Ala Thr Val Arg Leu Met Ala Lys  
 165 170 175  
 Val Pro Met Leu Ala Ala Tyr Ala His Arg Ala Arg Lys Gly Ala Pro  
 180 185 190  
 Tyr Met Tyr Pro Asp Asn Ser Leu Asn Ala Arg Glu Asn Phe Leu Arg  
 195 200 205  
 Met Met Phe Gly Tyr Pro Thr Glu Pro Tyr Glu Ile Asp Pro Ile Met  
 210 215 220  
 Val Lys Ala Leu Asp Lys Leu Leu Ile Leu His Ala Asp His Glu Gln  
 225 230 235 240  
 Asn Cys Ser Thr Ser Thr Val Arg Met Ile Gly Ser Ala Gln Ala Asn  
 245 250 255  
 Met Phe Val Ser Ile Ala Gly Gly Ile Asn Ala Leu Ser Gly Pro Leu  
 260 265 270  
 His Gly Gly Ala Asn Gln Ala Val Leu Glu Met Leu Glu Asp Ile Lys  
 275 280 285

Ser Asn His Gly Gly Asp Ala Thr Glu Phe Met Asn Lys Val Lys Asn  
 290 295 300  
 Lys Glu Asp Gly Val Arg Leu Met Gly Phe Gly His Arg Val Tyr Lys  
 305 310 315 320  
 Asn Tyr Asp Pro Arg Ala Ala Ile Val Lys Glu Thr Ala His Glu Ile  
 325 330 335  
 Leu Glu His Leu Gly Gly Asp Asp Leu Leu Asp Leu Ala Ile Lys Leu  
 340 345 350  
 Glu Glu Ile Ala Leu Ala Asp Asp Tyr Phe Ile Ser Arg Lys Leu Tyr  
 355 360 365  
 Pro Asn Val Asp Phe Tyr Thr Gly Leu Ile Tyr Arg Ala Met Gly Phe  
 370 375 380  
 Pro Thr Asp Phe Phe Thr Val Leu Phe Ala Ile Gly Arg Leu Pro Gly  
 385 390 395 400  
 Trp Ile Ala His Tyr Arg Glu Gln Leu Gly Ala Ala Gly Asn His Ile  
 405 410 415  
 Asn Arg Pro Arg Gln Val Tyr Thr Gly Asn Glu Ser Arg Lys Leu Val  
 420 425 430  
 Pro Arg Glu Glu Arg  
 435

<210> 5  
 <211> 1314  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> ATCC13032 GltA K415H NT

<400> 5  
 atgtttgaaa gggatatcgt ggctactgat aacaacaagg ctgtcctgca ctaccccggt 60  
 ggcgagttcg aaatggacat catcgaggct tctgagggtta acaacgggtgt tgtcctgggc 120  
 aagatgctgt ctgagactgg actgatcact tttgaccag gttatgtgag cactggctcc 180  
 accgagtcga agatcaccta catcgatggc gatgcgggaa tcctgcgtta ccgcggctat 240

gacatcgctg atctggctga gaatgccacc ttcaacgagg tttcttacct acttatcaac 300  
ggtgagctac caaccccaga tgagcttcac aagtttaacg acgagattcg ccaccacacc 360  
cttctggacg aggacttcaa gtcccagttc aacgtgttcc cacgcgacgc tcaccaatg 420  
gcaaccttgg cttcctcggg taacattttg tctacctact accaggacca gctgaaccca 480  
ctcgatgagg cacagcttga taaggcaacc gttcgcctca tggcaaagggt tccaatgctg 540  
gctgcgtacg cacaccgcgc acgcaagggt gtccttaca tgtaccaga caactccctc 600  
aatgcgcgtg agaacttcct gcgcatgatg ttcggttacc caaccgagcc atacgagatc 660  
gaccaatca tgggtcaaggc tctggacaag ctgctcatcc tgcacgctga ccacgagcag 720  
aactgctcca cctccaccgt tcgtatgatc ggttccgcac aggccaacat gtttgtctcc 780  
atcgctgggtg gcatcaacgc tctgtccggc cactgcacg gtggcgcaaa ccaggctggt 840  
ctggagatgc tcgaagacat caagagcaac cacggtggcg acgcaaccga gttcatgaac 900  
aagggtcaaga acaaggaaga cggcgtccgc ctcattgggtc tcggacaccg cgtttacaag 960  
aactacgatc cacgtgcagc aatcgtcaag gagaccgcac acgagatcct cgagcacctc 1020  
ggtggcgacg atcttctgga tctggcaatc aagctggaag aaattgcact ggctgatgat 1080  
tacttcatct cccgcaagct ctaccgaac gtagacttct acaccggcct gatctaccgc 1140  
gcaatgggct tccaactga cttcttcacc gtattgttcg caatcggtcg tctgccagga 1200  
tggatcgctc actaccgcga gcagctcggg gcagcaggca accacatcaa ccgcccacgc 1260  
cagggtctaca ccggcaacga atcccgcaag ttggttcctc gcgaggagcg ctaa 1314

<210> 6  
<211> 437  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> ATCC13032 GltA K415H+ N241T AA

<400> 6  
Met Phe Glu Arg Asp Ile Val Ala Thr Asp Asn Asn Lys Ala Val Leu  
1 5 10 15

His Tyr Pro Gly Gly Glu Phe Glu Met Asp Ile Ile Glu Ala Ser Glu  
                   20                                  25                                  30  
 Gly Asn Asn Gly Val Val Leu Gly Lys Met Leu Ser Glu Thr Gly Leu  
                   35                                  40                                  45  
 Ile Thr Phe Asp Pro Gly Tyr Val Ser Thr Gly Ser Thr Glu Ser Lys  
           50                                  55                                  60  
 Ile Thr Tyr Ile Asp Gly Asp Ala Gly Ile Leu Arg Tyr Arg Gly Tyr  
   65                                  70                                  75                                  80  
 Asp Ile Ala Asp Leu Ala Glu Asn Ala Thr Phe Asn Glu Val Ser Tyr  
                   85                                  90                                  95  
 Leu Leu Ile Asn Gly Glu Leu Pro Thr Pro Asp Glu Leu His Lys Phe  
                   100                                  105                                  110  
 Asn Asp Glu Ile Arg His His Thr Leu Leu Asp Glu Asp Phe Lys Ser  
           115                                  120                                  125  
 Gln Phe Asn Val Phe Pro Arg Asp Ala His Pro Met Ala Thr Leu Ala  
           130                                  135                                  140  
 Ser Ser Val Asn Ile Leu Ser Thr Tyr Tyr Gln Asp Gln Leu Asn Pro  
   145                                  150                                  155                                  160  
 Leu Asp Glu Ala Gln Leu Asp Lys Ala Thr Val Arg Leu Met Ala Lys  
                   165                                  170                                  175  
 Val Pro Met Leu Ala Ala Tyr Ala His Arg Ala Arg Lys Gly Ala Pro  
                   180                                  185                                  190  
 Tyr Met Tyr Pro Asp Asn Ser Leu Asn Ala Arg Glu Asn Phe Leu Arg  
           195                                  200                                  205  
 Met Met Phe Gly Tyr Pro Thr Glu Pro Tyr Glu Ile Asp Pro Ile Met  
           210                                  215                                  220  
 Val Lys Ala Leu Asp Lys Leu Leu Ile Leu His Ala Asp His Glu Gln  
   225                                  230                                  235                                  240  
 Thr Cys Ser Thr Ser Thr Val Arg Met Ile Gly Ser Ala Gln Ala Asn  
                   245                                  250                                  255  
 Met Phe Val Ser Ile Ala Gly Gly Ile Asn Ala Leu Ser Gly Pro Leu  
                   260                                  265                                  270  
 His Gly Gly Ala Asn Gln Ala Val Leu Glu Met Leu Glu Asp Ile Lys  
           275                                  280                                  285

Ser Asn His Gly Gly Asp Ala Thr Glu Phe Met Asn Lys Val Lys Asn  
 290 295 300  
 Lys Glu Asp Gly Val Arg Leu Met Gly Phe Gly His Arg Val Tyr Lys  
 305 310 315 320  
 Asn Tyr Asp Pro Arg Ala Ala Ile Val Lys Glu Thr Ala His Glu Ile  
 325 330 335  
 Leu Glu His Leu Gly Gly Asp Asp Leu Leu Asp Leu Ala Ile Lys Leu  
 340 345 350  
 Glu Glu Ile Ala Leu Ala Asp Asp Tyr Phe Ile Ser Arg Lys Leu Tyr  
 355 360 365  
 Pro Asn Val Asp Phe Tyr Thr Gly Leu Ile Tyr Arg Ala Met Gly Phe  
 370 375 380  
 Pro Thr Asp Phe Phe Thr Val Leu Phe Ala Ile Gly Arg Leu Pro Gly  
 385 390 395 400  
 Trp Ile Ala His Tyr Arg Glu Gln Leu Gly Ala Ala Gly Asn His Ile  
 405 410 415  
 Asn Arg Pro Arg Gln Val Tyr Thr Gly Asn Glu Ser Arg Lys Leu Val  
 420 425 430  
 Pro Arg Glu Glu Arg  
 435

<210> 7  
 <211> 1314  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> ATCC13032 GltA K415H+ N241T NT

<400> 7  
 atgtttgaaa gggatatcgt ggctactgat aacaacaagg ctgtcctgca ctaccccggt 60  
 ggcgagttcg aaatggacat catcgaggct tctgagggtta acaacgggtgt tgtcctgggc 120  
 aagatgctgt ctgagactgg actgatcact tttgaccag gttatgtgag cactggctcc 180  
 accgagtcga agatcaccta catcgatggc gatgcgggaa tcctgcgtta ccgcggctat 240

|                                                                    |      |
|--------------------------------------------------------------------|------|
| gacatcgctg atctggctga gaatgccacc ttcaacgagg tttcttacct acttatcaac  | 300  |
| ggtgagctac caaccccaga tgagcttcac aagtttaacg acgagattcg ccaccacacc  | 360  |
| cttctggacg aggacttcaa gtcccagttc aacgtgttcc cacgcgacgc tcaccaatg   | 420  |
| gcaaccttgg cttcctcggg taacattttg tctacctact accaggacca gctgaaccca  | 480  |
| ctcgatgagg cacagcttga taaggcaacc gttcgcctca tggcaaagggt tccaatgctg | 540  |
| gctgcgtacg cacaccgcgc acgcaagggt gtccttaca tgtaccaga caactccctc    | 600  |
| aatgcgcgtg agaacttcct gcgcgatgat ttcggttacc caaccgagcc atacgagatc  | 660  |
| gacccaatca tgggtcaaggc tctggacaag ctgctcatcc tgcacgctga ccacgagcag | 720  |
| acctgctcca cctccaccgt tcgtatgatc ggttccgcac aggccaacat gtttgtctcc  | 780  |
| atcgctgggtg gcatcaacgc tctgtccggc cactgcacg gtggcgcaaa ccaggctggt  | 840  |
| ctggagatgc tcgaagacat caagagcaac cacggtggcg acgcaaccga gttcatgaac  | 900  |
| aagggtcaaga acaaggaaga cggcgtccgc ctcattgggt tcggacaccg cgtttacaag | 960  |
| aactacgatc cacgtgcagc aatcgtcaag gagaccgcac acgagatcct cgagcacctc  | 1020 |
| ggtggcgacg atcttctgga tctggcaatc aagctggaag aaattgcact ggctgatgat  | 1080 |
| tacttcatct cccgcaagct ctacccgaac gtagacttct acaccggcct gatctaccgc  | 1140 |
| gcaatgggct tcccaactga cttcttcacc gtattgttcg caatcggtcg tctgccagga  | 1200 |
| tggatcgctc actaccgcga gcagctcggg gcagcaggca accacatcaa ccgcccacgc  | 1260 |
| cagggtctaca ccggcaacga atcccgcaag ttggttcctc gcgaggagcg ctaa       | 1314 |

<210> 8  
 <211> 44  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> primer 1

|                                                  |    |
|--------------------------------------------------|----|
| <400> 8                                          |    |
| tcgagctcgg tacccccgtt cgtatgatcg gttccgcaca ggcc | 44 |

<210> 9  
<211> 31  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> primer 2

<400> 9  
gtgcagcagg caaccacatc aaccgcccac g 31

<210> 10  
<211> 31  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> primer 3

<400> 10  
cgtgggagggt tgatgtgggt gcctgctgca c 31

<210> 11  
<211> 46  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> primer 4

<400> 11  
ctctagagga tccccgccgt aagcagcctc tggtggaatg gtcagc 46

<210> 12  
<211> 35  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> primer 5

<400> 12  
gctgaccacg agcagacctg ctccacctcc accgt 35

<210> 13  
 <211> 35  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> primer 6

<400> 13  
 acgggtggagg tggagcaggt ctgctcgtgg tcagc

35

<210> 14  
 <211> 418  
 <212> PRT  
 <213> Unknown

<220>  
 <223> YjeH AA

<400> 14  
 Met Ser Gly Leu Lys Gln Glu Leu Gly Leu Ala Gln Gly Ile Gly Leu  
 1 5 10 15  
 Leu Ser Thr Ser Leu Leu Gly Thr Gly Val Phe Ala Val Pro Ala Leu  
 20 25 30  
 Ala Ala Leu Val Ala Gly Asn Asn Ser Leu Trp Ala Trp Pro Val Leu  
 35 40 45  
 Ile Ile Leu Val Phe Pro Ile Ala Ile Val Phe Ala Ile Leu Gly Arg  
 50 55 60  
 His Tyr Pro Ser Ala Gly Gly Val Ala His Phe Val Gly Met Ala Phe  
 65 70 75 80  
 Gly Ser Arg Leu Glu Arg Val Thr Gly Trp Leu Phe Leu Ser Val Ile  
 85 90 95  
 Pro Val Gly Leu Pro Ala Ala Leu Gln Ile Ala Ala Gly Phe Gly Gln  
 100 105 110  
 Ala Met Phe Gly Trp His Ser Trp Gln Leu Leu Leu Ala Glu Leu Gly  
 115 120 125  
 Thr Leu Ala Leu Val Trp Tyr Ile Gly Thr Arg Gly Ala Ser Ser Ser

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

Gly Leu Trp Leu Leu Leu Pro Lys Arg Lys Thr Pro Glu Asn Gly Ile  
405 410 415

Thr Thr

<210> 15  
<211> 1257  
<212> DNA  
<213> Unknown

<220>  
<223> YjeH NT

<400> 15  
atgagtggac tcaaacaaga actggggctg gcccagggca ttggcctgct atcgacgtca 60  
ttattaggca ctggcgtgtt tgccgttcct gcgttagctg cgctggtagc gggcaataac 120  
agcctgtggg cgtggcccggt tttgattatc ttagtggtcc cgattgcgat tgtgtttgctg 180  
attctgggtc gccactatcc cagcgcaggc ggctgcgcgc acttcgtcgg tatggcggtt 240  
ggttcgcggc ttgagcgagt caccggctgg ctgtttttat cggtcattcc cgtgggtttg 300  
cctgccgcac tacaaattgc cgccgggttc ggccaggcga tgtttggtg gcatagctgg 360  
caactgttgt tggcagaact cggtacgctg gcgctggtgt ggtatatcgg tactcgcggt 420  
gccagttcca gtgctaattc acaaaccgtt attgccggac ttatcgtcgc gctgattgtc 480  
gctatctggt gggcgggcga tatcaaacct gcgaatatcc cttttccggc acctggtaat 540  
atcgaactta ccgggttatt tgctgcgtta tcagtgatgt tctggtgttt tgtcggctctg 600  
gaggcatttg cccatctcgc ctccgaattt aaaaatccag agcgtgattt tcctcgtgct 660  
ttgatgattg gtctgctgct ggcaggatta gtctactggg gctgtacggt agtcgtctta 720  
cacttcgacg cctatggtga aaaaatggcg gcggcagcat cgcttccaaa aattgtagtg 780  
cagttgttcg gtgtaggagc gttatggatt gcctgcgtga ttggctatct ggcctgcttt 840  
gccagtctca acatttatat acagagcttc gcccgcctgg tctggtcgca ggcgaacat 900  
aatcctgacc actacctggc acgcctctct tctcgccata tcccgaataa tgccctcaat 960  
gcggtgctcg gctgctgtgt ggtgagcact ttggtgattc atgctttaga gatcaatctg 1020

gacgctctta ttatttatgc caatggcatc tttattatga tttatctggt atgcatgctg 1080  
gcaggctgta aattattgca aggacgttat cgactactgg cggtggttgg cgggctgtta 1140  
tgcgttctgt tactggcaat ggtcggctgg aaaagtctct atgcgctgat catgctggcg 1200  
gggttatggc tgttgctgcc aaaacgaaaa acgccgaaa atggcataac cacataa 1257

<210> 16  
<211> 88  
<212> PRT  
<213> Unknown

<220>  
<223> NCgl2335 AA

<400> 16  
Met Ala Tyr Thr Phe Asp His Val Val Ala Trp Arg Trp Cys Thr Lys  
1 5 10 15  
Glu Asp Ala Tyr Asn Tyr Thr His Leu Phe Asp Gln Leu Gln Pro Pro  
20 25 30  
Leu Ile Val Thr Thr Asp Gly Gln Lys Arg Arg Thr Gln Ser His His  
35 40 45  
His Asp Leu Ala Asp Asn Glu Asn Pro Thr Leu Pro Arg Pro Arg Gln  
50 55 60  
Thr Gln Arg Pro Lys Thr Arg His Pro Lys Thr Arg Ala Glu Leu Ala  
65 70 75 80  
Glu Lys His Ser Gly Val Ser Pro  
85

<210> 17  
<211> 267  
<212> DNA  
<213> Unknown

<220>  
<223> NCgl2335 NT

<400> 17  
gtggcctaca ccttcgacca cgtcgtcgcc tggcgtggt gcaccaaaga agacgcctac 60

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| aactacaccc acctcttcga tcaactccaa ccacccttaa tcgtgaccac cgacggacaa | 120 |
| aaaaggcgca ctcaaagcca tcaccacgac ctggccgaca acgaaaatcc aacgctgcct | 180 |
| cgtccacgtc aaacgcaacg tccaaaaaca cgtcacccta agaccctgc tgagctcgcc  | 240 |
| gaaaagcact ccggggtctc tccttga                                     | 267 |

<210> 18  
 <211> 39  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Tn\_5 F

|                                            |    |
|--------------------------------------------|----|
| <400> 18                                   |    |
| tgaattcgag ctcggtaccc caccgacgcg catctgcct | 39 |

<210> 19  
 <211> 50  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Tn\_5 R

|                                                        |    |
|--------------------------------------------------------|----|
| <400> 19                                               |    |
| ggtgtggtga ctttcagcag ttcccggggg ggaggaggca tgtggtgttg | 50 |

<210> 20  
 <211> 50  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Tn\_3 F

|                                                        |    |
|--------------------------------------------------------|----|
| <400> 20                                               |    |
| caacaccaca tgcctcctcc cccccgggaa ctgctgaaag tcaccacacc | 50 |

<210> 21

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

<220>  
<223> Tn\_3 R

<400> 21  
gtcgactcta gaggatcccc ctcccaaacc attgaggaat gg 42

<210> 22  
<211> 39  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> CJ7\_yjeH F

<400> 22  
acaccacatg cctcctcccc agaaacatcc cagcgctac 39

<210> 23  
<211> 43  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> CJ7\_yjeH R

<400> 23  
agttcttgtt tgagtccact catagtgttt cctttcgttg ggt 43

<210> 24  
<211> 44  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> yjeH F

<400> 24  
acccaacgaa aggaaacact atgagtggac tcaaacaaga actg 44

<210> 25  
 <211> 44  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> yjeH R

<400> 25  
 gactttcagc agttcccggg ttatgtgggtt atgccatttt ccgg

44

<210> 26  
 <211> 418  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> yjeH F351L AA

<400> 26  
 Met Ser Gly Leu Lys Gln Glu Leu Gly Leu Ala Gln Gly Ile Gly Leu  
 1 5 10 15  
 Leu Ser Thr Ser Leu Leu Gly Thr Gly Val Phe Ala Val Pro Ala Leu  
 20 25 30  
 Ala Ala Leu Val Ala Gly Asn Asn Ser Leu Trp Ala Trp Pro Val Leu  
 35 40 45  
 Ile Ile Leu Val Phe Pro Ile Ala Ile Val Phe Ala Ile Leu Gly Arg  
 50 55 60  
 His Tyr Pro Ser Ala Gly Gly Val Ala His Phe Val Gly Met Ala Phe  
 65 70 75 80  
 Gly Ser Arg Leu Glu Arg Val Thr Gly Trp Leu Phe Leu Ser Val Ile  
 85 90 95  
 Pro Val Gly Leu Pro Ala Ala Leu Gln Ile Ala Ala Gly Phe Gly Gln  
 100 105 110  
 Ala Met Phe Gly Trp His Ser Trp Gln Leu Leu Leu Ala Glu Leu Gly  
 115 120 125  
 Thr Leu Ala Leu Val Trp Tyr Ile Gly Thr Arg Gly Ala Ser Ser Ser  
 130 135 140

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

405

410

415

Thr Thr

<210> 27  
 <211> 1257  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> yjeH F351L NT

<400> 27  
 atgagtggac tcaaacaaga actggggctg gcccagggca ttggcctgct atcgacgtca 60  
 ttattaggca ctggcgtgtt tgccgttcct gcgttagctg cgctggtagc gggcaataac 120  
 agcctgtggg cgtggcccggt tttgattatc ttagtggtcc cgattgcgat tgtgtttgcg 180  
 attctgggtc gccactatcc cagcgcaggc ggcgtcgcgc acttcgtcgg tatggcgttt 240  
 ggttcgcggc ttgagcgagt caccggctgg ctgaatttat cggtcattcc cgtgggtttg 300  
 cctgccgcac taaaattgc cgccgggttc ggccaggcga tgtttggtg gcatagctgg 360  
 caactgttgt tggcagaact cggtacgtg gcgctggtgt ggtatatcgg tactcgcggt 420  
 gccagttcca gtgctaattc acaaaccgtt attgccggac ttatcgtcgc gctgattgtc 480  
 gctatctggt gggcgggcga tatcaaacct gcgaatatcc cctttccggc acctggtaat 540  
 atcgaactta ccgggttatt tgctgcgtta tcagtgatgt tctggtgttt tgcggtctg 600  
 gaggcatttg cccatctcgc ctcggaattt aaaaatccag agcgtgattt tcctcgtgct 660  
 ttgatgattg gtctgctgct ggcaggatta gtctactggg gctgtacggt agtcgtctta 720  
 cacttcgacg cctatggtga aaaaatggcg gcggcagcat cgcttccaaa aattgtagtg 780  
 cagttgttcg gtgtaggagc gttatggatt gcctgcgtga ttggctatct ggcctgcttt 840  
 gccagtctca acatttatat acagagcttc gccgcctgg tctggtcgca ggcgcaacat 900  
 aatcctgacc actacctggc acgcctctct tctcgccata tcccgaataa tgccctcaat 960  
 gcggtgctcg gctgctgtgt ggtgagcact ttggtgattc atgctttaga gatcaatctg 1020

|                                                                   |      |
|-------------------------------------------------------------------|------|
| gacgctctta ttatttatgc caatggcatc cttattatga tttatctgtt atgcatgctg | 1080 |
| gcaggctgta aattattgca aggacgttat cgactactgg cggtggttgg cgggctgtta | 1140 |
| tgcgttctgt tactggcaat ggtcggctgg aaaagtctct atgcgctgat catgctggcg | 1200 |
| gggttatggc tgttgctgcc aaaacgaaaa acgccgaaa atggcataac cacataa     | 1257 |

<210> 28  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> F351L F

|                           |    |
|---------------------------|----|
| <400> 28                  |    |
| caatggcatc cttattatga ttt | 23 |

<210> 29  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> F351L R

|                            |    |
|----------------------------|----|
| <400> 29                   |    |
| aaatcataat aaggatgccca ttg | 23 |

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

<220>  
 <223> gltA-F

|                       |    |
|-----------------------|----|
| <400> 30              |    |
| atgtttgaaa gggatatcgt | 20 |

<210> 31  
 <211> 21

<212> DNA  
 <213> Artificial Sequence

<220>  
 <223> gltA-R

<400> 31  
 ttagcgctcc tcgcgaggaa c 21

<210> 32  
 <211> 5803  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> pDCM2

<400> 32  
 gttcgcttgc tgtccataaa accgcccagt ctagctatcg ccatgtaagc cactgcaag 60  
 ctacctgctt tctcttttgcg cttgcgtttt cccttgtcca gatagcccag tagctgacat 120  
 tcatccgggg tcagcaccgt ttctgcggac tggctttcta cgtgttccgc ttcctttagc 180  
 agcccttgcg ccctgagtgc ttgcggcagc gtgaagctag cttttatcgc cattcgccat 240  
 tcaggctgcg caactgttgg gaagggcgat cgggtgcgggc ctcttcgcta ttacgccagc 300  
 tggcgaaagg gggatgtgct gcaaggcgat taagttgggt aacgccaggg ttttcccagt 360  
 cacgacgttg taaaacgacg gccagtgaat tcgagctcgg taccggggga tcctctagag 420  
 tcgacctgca ggcatgcaag cttggcgtaa tcatggtcac agctgtttcc tgtgtgaaat 480  
 tgttatccgc tcacaattcc acacaacata cgagccggaa gcataaagtg taaagcctgg 540  
 ggtgcctaata gaggtagcta actcacatta attgcgttgc gctcactgcc cgctttccag 600  
 tcgggaaacc tgtcgtgcc gctgcattaa tgaatcggcc aacgcgcggg gagaggcggg 660  
 ttgcgtattg ggcgctcttc cgcttcctcg ctactgact cgctgcgctc ggctcgttcgg 720  
 ctgcggcgag cggatatcagc tactcaaag gcggtataac gggtatccac agaatacagg 780  
 gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag 840  
 gccgcgttgc tggcggtttt ccataggctc cgccccctg acgagcatca caaaaatcga 900

|                                                                     |      |
|---------------------------------------------------------------------|------|
| cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc gtttccccct   | 960  |
| ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata cctgtccgcc   | 1020 |
| tttctccctt cgggaagcgt ggcgctttct caatgctcac gctgtaggta tctcagttcg   | 1080 |
| gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac cccccgttca gcccgaccgc   | 1140 |
| tgcgcccttat ccggtaaacta tcgtcttgag tccaacccgg taagacacga cttatcgcca | 1200 |
| ctggcagcag cactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag    | 1260 |
| ttcttgaagt ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct   | 1320 |
| ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg caaacaacc    | 1380 |
| accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag aaaaaagga    | 1440 |
| tctcaagaag atcctttgat cttttctacg gggctctgacg ctcaagtggaa cgaaaactca | 1500 |
| cgtaaggga ttttggcat gagattatca aaaaggatct tcacctagat cttttgggg      | 1560 |
| tgggcgaaga actccagcat gagatccccg cgctggagga tcatccagcc ctgatagaaa   | 1620 |
| cagaagccac tggagcacct caaaaacacc atcatacact aaatcagtaa gttggcagca   | 1680 |
| tcaccgacg cactttgcgc cgaataaata cctgtgacgg aagatcactt cgcagaataa    | 1740 |
| ataaatcctg gtgtccctgt tgataccggg aagccctggg ccaacttttg gcgaaaatga   | 1800 |
| gacgttgatc ggcacgtaag aggttccaac tttcaccata atgaaataag atcactaccg   | 1860 |
| ggcgtatttt ttgagttatc gagattttca ggagctgata gaaacagaag cactggagc    | 1920 |
| acctcaaaaa caccatcata cactaaatca gtaagttggc agcatcacc gacgcacttt    | 1980 |
| gcgccgaata aatacctgtg acggaagatc acttcgcaga ataaataaat cctggtgtcc   | 2040 |
| ctgttgatac cgggaagccc tgggccaaact tttggcgaaa atgagacgtt gatcggcacg  | 2100 |
| taagaggttc caactttcac cataatgaaa taagatcact accgggcgta ttttttgagt   | 2160 |
| tatcgagatt ttcaggagct ctttggcatc gtctctcgcc tgtcccctca gttcagtaat   | 2220 |
| ttcctgcatt tgcctgtttc cagtcggtag atattccaca aaacagcagg gaagcagcgc   | 2280 |
| ttttccgctg cataaccctg cttcggggtc attatagcga ttttttcggt atatccatcc   | 2340 |
| tttttcgcac gatatacagg attttgccaa agggttcgtg tagactttcc ttggtgtatc   | 2400 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| caacggcgtc agccgggag gataggtgaa gtagggccac ccgcgagcgg gtgttccttc   | 2460 |
| ttcactgtcc cttattcgca cctggcggtg ctcaacggga atcctgctct gcgaggctgg  | 2520 |
| ccggctaccg ccggcgtaac agatgagggc aagcggatgg ctgatgaaac caagccaacc  | 2580 |
| aggaagggca gcccacctat caaggtgtac tgccttcag acgaacgaag agcgattgag   | 2640 |
| gaaaaggcgg cggcgccgg catgagcctg tcggcctacc tgctggccgt cggccagggc   | 2700 |
| tacaaaatca cgggcgtcgt ggactatgag cacgtccgcg agggcgtccc ggaaaacgat  | 2760 |
| tccgaagccc aacctttcat agaaggcggc ggtggaatcg aaatctcgtg atggcaggtt  | 2820 |
| gggcgtcgtc tggtcggtca tttcgaaaaa ggtaggaat acggttagcc atttgcctgc   | 2880 |
| ttttatatag ttcantatgg gattcacctt tatgttgata agaaataaaa gaaaatgcca  | 2940 |
| ataggatatc ggcatTTTTt tttgcgtttt tatttgtaa ctgttaattg tccttgttca   | 3000 |
| aggatgctgt ctttgacaac agatgttttc ttgcctttga tgttcagcag gaagctcggc  | 3060 |
| gcaaacgttg attgtttgtc tgcgtagaat cctctgtttg tcatatagct tgtaatcacg  | 3120 |
| acattgtttc ctttcgcttg aggtacagcg aagtgtgagt aagtaaaggt tacatcgtta  | 3180 |
| ggcggatcaa gatccatttt taacacaagg ccagttttgt tcagcggctt gtatgggcca  | 3240 |
| gttaaagaat tagaaacata accaagcatg taaatatcgt tagacgtaat gccgtcaatc  | 3300 |
| gtcatttttg atccgcggga gtcagtgaac aggtaccatt tgccgttcat tttaaagacg  | 3360 |
| ttcgcgcgtt caatttcac tgttactgtg ttagatgcaa tcagcggttt catcactttt   | 3420 |
| ttcagtgtgt aatcatcgtt tagctcaatc ataccgagag cgccgtttgc taactcagcc  | 3480 |
| gtgcgttttt tatcgctttg cagaagtttt tgactttctt gacggaagaa tgatgtgctt  | 3540 |
| ttgccatagt atgctttgtt aaataaagat tcttcgcctt ggtagccatc ttcagttcca  | 3600 |
| gtgtttgctt caaatactaa gtatttgtgg cttttatctt ctacgtagtg aggatctctc  | 3660 |
| agcgtatggg tgtcgctga gctgtagttg ctttcacga tgaactgctg tacattttga    | 3720 |
| tacgtttttc cgtcaccgtc aaagattgat ttataatcct ctacaccgtt gatgttcaaa  | 3780 |
| gagctgtctg atgctgatac gttaacttgt gcagttgtca gtgtttgttt gccgtaatgt  | 3840 |
| ttaccggaga aatcagtgtga gaataaacgg atttttccgt cagatgtaaa tgtggctgaa | 3900 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| cctgaccatt cttgtgtttg gtcttttagg atagaatcat ttgcatcgaa tttgtcgctg  | 3960 |
| tcttttaaaga cgcggccagc gtttttccag ctgtcaatag aagtttcgcc gactttttga | 4020 |
| tagaacatgt aaatcgatgt gtcacccgca ttttttaggat ctccggctaa tgcaaagacg | 4080 |
| atgtggtagc cgtgatagtt tgcgacagtg ccgtcagcgt tttgtaatgg ccagctgtcc  | 4140 |
| caaacgtcca ggccttttgc agaagagata tttttaattg tggacgaatc aaattcagaa  | 4200 |
| acttgatatt tttcattttt ttgctgttca gggatttgca gcatatcatg gcgtgtaata  | 4260 |
| tgggaaatgc cgtatgtttc cttatatggc ttttggttcg tttctttcgc aaacgcttga  | 4320 |
| gttgcgcctc ctgccagcag tgcggtagta aaggttaata ctgttgcttg ttttgcaaac  | 4380 |
| tttttgatgt tcatcgttca tgtctccttt tttatgtact gtgtagcgg tctgcttctt   | 4440 |
| ccagccctcc tgtttgaaga tggcaagtta gttacgcaca ataaaaaag acctaaaata   | 4500 |
| tgtaaggggt gacgccaaag tatacacttt gccctttaca catttttaggt cttgcctgct | 4560 |
| ttatcagtaa caaaccgcg cgatttactt ttcgacctca ttctattaga ctctcgtttg   | 4620 |
| gattgcaact ggtctatttt cctcttttgt ttgatagaaa atcataaaag gatttgcaga  | 4680 |
| ctacgggcct aaagaactaa aaaatctatc tgtttctttt cattctctgt attttttata  | 4740 |
| gtttctgttg catgggcata aagttgcctt tttaatcaca attcagaaaa tatcataata  | 4800 |
| tctcatttca ctaaataata gtgaacggca ggtatatgtg atgggttaaa aaggatcacc  | 4860 |
| ccagagtccc gtcagaaga actcgtcaag aaggcgatag aaggcgatgc gctgcgaatc   | 4920 |
| gggagcggcg ataccgtaaa gcacaggaa gcggtcagcc cattcgccgc caagctcttc   | 4980 |
| agcaatatca cgggtagcca acgctatgtc ctgatagcgg tccgccacac ccagccggcc  | 5040 |
| acagtcgatg aatccagaaa agcggccatt ttccaccatg atattcggca agcaggcatc  | 5100 |
| gccatgggtc acgacgagat cctcgccgtc gggcatccgc gccttgagcc tggcgaacag  | 5160 |
| ttcggctggc gcgagcccct gatgctcttc gtccagatca tcctgacga caagaccggc   | 5220 |
| ttccatccga gtacgtgctc gtcgatgcg atgtttcgct tgggtggtcga atgggcaggt  | 5280 |
| agccggatca agcgtatgca gccgccgat tgcacagcc atgatggata ctttctcggc    | 5340 |
| aggagcaagg tgagatgaca ggagatcctg ccccggcact tcgccaata gcagccagtc   | 5400 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| ccttccccgct tcagtgacaa cgtcgagaca gctgcgcaag gaacgcccgt cgtggccagc | 5460 |
| cacgatagcc gcgctgcctc gtcttggagt tcattcaggg caccggacag gtcggtcttg  | 5520 |
| acaaaaagaa ccgggcgccc ctgcgctgac agccggaaca cggcggcatc agagcagccg  | 5580 |
| attgtctgtt gtgcccagtc atagccgaat agcctctcca cccaagcggc cggagaacct  | 5640 |
| gcgtgcaatc catcttgttc aatcatgcga aacgatactc atcctgtctc ttgatcagat  | 5700 |
| cttgatcccc tgcgcatca gatccttggc ggcaagaaag ccatccagtt tactttgcag   | 5760 |
| ggcttcccaa ccttaccaga gggcgcccca gctggcaatt ccg                    | 5803 |

<210> 33  
 <211> 127  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> General formula 1

<220>  
 <221> MISC\_FEATURE  
 <222> (1)  
 <223> Xaa is N or S.

<220>  
 <221> MISC\_FEATURE  
 <222> (9)  
 <223> Xaa is A or E.

<220>  
 <221> MISC\_FEATURE  
 <222> (73)  
 <223> Xaa is Y or C.

<400> 33  
 Xaa Asn His Gly Gly Asp Ala Thr Xaa Phe Met Asn Lys Val Lys Asn  
 1 5 10 15  
 Lys Glu Asp Gly Val Arg Leu Met Gly Phe Gly His Arg Val Tyr Lys  
 20 25 30

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | Tyr | Asp | Pro | Arg | Ala | Ala | Ile | Val | Lys | Glu | Thr | Ala | His | Glu | Ile |
|     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |
| Leu | Glu | His | Leu | Gly | Gly | Asp | Asp | Leu | Leu | Asp | Leu | Ala | Ile | Lys | Leu |
|     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |
| Glu | Glu | Ile | Ala | Leu | Ala | Asp | Asp | Xaa | Phe | Ile | Ser | Arg | Lys | Leu | Tyr |
| 65  |     |     |     |     | 70  |     |     |     |     | 75  |     |     |     |     | 80  |
| Pro | Asn | Val | Asp | Phe | Tyr | Thr | Gly | Leu | Ile | Tyr | Arg | Ala | Met | Gly | Phe |
|     |     |     |     | 85  |     |     |     |     | 90  |     |     |     |     | 95  |     |
| Pro | Thr | Asp | Phe | Phe | Thr | Val | Leu | Phe | Ala | Ile | Gly | Arg | Leu | Pro | Gly |
|     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |
| Trp | Ile | Ala | His | Tyr | Arg | Glu | Gln | Leu | Gly | Ala | Ala | Gly | Asn | His |     |
|     |     | 115 |     |     |     |     | 120 |     |     |     |     | 125 |     |     |     |