SEQUENCE LISTING

<110> Pairwise Plants Services, Inc.
Watts, Joseph Matthew
Jali, Sathya Sheela

<120> VARIANTS OF CAS12A NUCLEASES AND METHODS OF MAKING AND USE THEREOF

<130> 1499.7.WO

<150> US 62/916,392
<151> 2019-10-17

<160> 62

<170> PatentIn version 3.5

<210> 1
<211> 1228
<212> PRT
<213> Unknown

<220>
<223> Lachnospiraceae bacterium

<400> 1

Met Ser Lys Leu Glu Lys Phe Thr Asn Cys Tyr Ser Leu Ser Lys Thr
1 5 10 15

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

Asn Lys Arg Leu Leu Val Glu Asp Glu Lys Arg Ala Glu Asp Tyr Lys
35 40 45

Gly Val Lys Lys Leu Leu Asp Arg Tyr Tyr Leu Ser Phe Ile Asn Asp
50 55 60

Val Leu His Ser Ile Lys Leu Lys Asn Leu Asn Asn Tyr Ile Ser Leu
65 70 75 80

Phe Arg Lys Lys Thr Arg Thr Glu Lys Glu Asn Lys Glu Leu Glu Asn
85 90 95

Leu Glu Ile Asn Leu Arg Lys Glu Ile Ala Lys Ala Phe Lys Gly Asn
100 105 110

```
Glu Gly Tyr Lys Ser Leu Phe Lys Lys Asp Ile Ile Glu Thr Ile Leu
        115                 120                 125

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

Asn Gly Phe Thr Thr Ala Phe Thr Gly Phe Phe Asp Asn Arg Glu Asn
145                 150                 155                 160

Met Phe Ser Glu Glu Ala Lys Ser Thr Ser Ile Ala Phe Arg Cys Ile
                165                 170                 175

Asn Glu Asn Leu Thr Arg Tyr Ile Ser Asn Met Asp Ile Phe Glu Lys
            180                 185                 190

Val Asp Ala Ile Phe Asp Lys His Glu Val Gln Glu Ile Lys Glu Lys
        195                 200                 205

Ile Leu Asn Ser Asp Tyr Asp Val Glu Asp Phe Phe Glu Gly Glu Phe
    210                 215                 220

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

Ile Gly Gly Phe Val Thr Glu Ser Gly Glu Lys Ile Lys Gly Leu Asn
                245                 250                 255

Glu Tyr Ile Asn Leu Tyr Asn Gln Lys Thr Lys Gln Lys Leu Pro Lys
            260                 265                 270

Phe Lys Pro Leu Tyr Lys Gln Val Leu Ser Asp Arg Glu Ser Leu Ser
        275                 280                 285

Phe Tyr Gly Glu Gly Tyr Thr Ser Asp Glu Glu Val Leu Glu Val Phe
    290                 295                 300

Arg Asn Thr Leu Asn Lys Asn Ser Glu Ile Phe Ser Ser Ile Lys Lys
305                 310                 315                 320
```

```
Leu Glu Lys Leu Phe Lys Asn Phe Asp Glu Tyr Ser Ser Ala Gly Ile
                325                 330                 335

Phe Val Lys Asn Gly Pro Ala Ile Ser Thr Ile Ser Lys Asp Ile Phe
            340                 345                 350

Gly Glu Trp Asn Val Ile Arg Asp Lys Trp Asn Ala Glu Tyr Asp Asp
        355                 360                 365

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

Arg Arg Lys Ser Phe Lys Lys Ile Gly Ser Phe Ser Leu Glu Gln Leu
385                 390                 395                 400

Gln Glu Tyr Ala Asp Ala Asp Leu Ser Val Val Glu Lys Leu Lys Glu
                405                 410                 415

Ile Ile Ile Gln Lys Val Asp Glu Ile Tyr Lys Val Tyr Gly Ser Ser
            420                 425                 430

Glu Lys Leu Phe Asp Ala Asp Phe Val Leu Glu Lys Ser Leu Lys Lys
        435                 440                 445

Asn Asp Ala Val Val Ala Ile Met Lys Asp Leu Leu Asp Ser Val Lys
    450                 455                 460

Ser Phe Glu Asn Tyr Ile Lys Ala Phe Phe Gly Glu Gly Lys Glu Thr
465                 470                 475                 480

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

Leu Leu Lys Val Asp His Ile Tyr Asp Ala Ile Arg Asn Tyr Val Thr
            500                 505                 510

Gln Lys Pro Tyr Ser Lys Asp Lys Phe Lys Leu Tyr Phe Gln Asn Pro
        515                 520                 525

Gln Phe Met Gly Gly Trp Asp Lys Asp Lys Glu Thr Asp Tyr Arg Ala
    530                 535                 540
```

```
Thr Ile Leu Arg Tyr Gly Ser Lys Tyr Tyr Leu Ala Ile Met Asp Lys
545                 550                 555                 560

Lys Tyr Ala Lys Cys Leu Gln Lys Ile Asp Lys Asp Asp Val Asn Gly
                565                 570                 575

Asn Tyr Glu Lys Ile Asn Tyr Lys Leu Leu Pro Gly Pro Asn Lys Met
            580                 585                 590

Leu Pro Lys Val Phe Phe Ser Lys Lys Trp Met Ala Tyr Tyr Asn Pro
        595                 600                 605

Ser Glu Asp Ile Gln Lys Ile Tyr Lys Asn Gly Thr Phe Lys Lys Gly
    610                 615                 620

Asp Met Phe Asn Leu Asn Asp Cys His Lys Leu Ile Asp Phe Phe Lys
625                 630                 635                 640

Asp Ser Ile Ser Arg Tyr Pro Lys Trp Ser Asn Ala Tyr Asp Phe Asn
                645                 650                 655

Phe Ser Glu Thr Glu Lys Tyr Lys Asp Ile Ala Gly Phe Tyr Arg Glu
            660                 665                 670

Val Glu Glu Gln Gly Tyr Lys Val Ser Phe Glu Ser Ala Ser Lys Lys
        675                 680                 685

Glu Val Asp Lys Leu Val Glu Glu Gly Lys Leu Tyr Met Phe Gln Ile
    690                 695                 700

Tyr Asn Lys Asp Phe Ser Asp Lys Ser His Gly Thr Pro Asn Leu His
705                 710                 715                 720

Thr Met Tyr Phe Lys Leu Leu Phe Asp Glu Asn Asn His Gly Gln Ile
                725                 730                 735

Arg Leu Ser Gly Gly Ala Glu Leu Phe Met Arg Arg Ala Ser Leu Lys
            740                 745                 750
```

```
Lys Glu Glu Leu Val Val His Pro Ala Asn Ser Pro Ile Ala Asn Lys
        755                 760                 765

Asn Pro Asp Asn Pro Lys Lys Thr Thr Thr Leu Ser Tyr Asp Val Tyr
    770                 775                 780

Lys Asp Lys Arg Phe Ser Glu Asp Gln Tyr Glu Leu His Ile Pro Ile
785                 790                 795                 800

Ala Ile Asn Lys Cys Pro Lys Asn Ile Phe Lys Ile Asn Thr Glu Val
                805                 810                 815

Arg Val Leu Leu Lys His Asp Asp Asn Pro Tyr Val Ile Gly Ile Asp
            820                 825                 830

Arg Gly Glu Arg Asn Leu Leu Tyr Ile Val Val Val Asp Gly Lys Gly
        835                 840                 845

Asn Ile Val Glu Gln Tyr Ser Leu Asn Glu Ile Ile Asn Asn Phe Asn
    850                 855                 860

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

Lys Glu Arg Phe Glu Ala Arg Gln Asn Trp Thr Ser Ile Glu Asn Ile
                885                 890                 895

Lys Glu Leu Lys Ala Gly Tyr Ile Ser Gln Val Val His Lys Ile Cys
            900                 905                 910

Glu Leu Val Glu Lys Tyr Asp Ala Val Ile Ala Leu Glu Asp Leu Asn
        915                 920                 925

Ser Gly Phe Lys Asn Ser Arg Val Lys Val Glu Lys Gln Val Tyr Gln
    930                 935                 940

Lys Phe Glu Lys Met Leu Ile Asp Lys Leu Asn Tyr Met Val Asp Lys
945                 950                 955                 960

Lys Ser Asn Pro Cys Ala Thr Gly Gly Ala Leu Lys Gly Tyr Gln Ile
                965                 970                 975
```

```
Thr Asn Lys Phe Glu Ser Phe Lys Ser Met Ser Thr Gln Asn Gly Phe
            980                 985                 990

Ile Phe Tyr Ile Pro Ala Trp Leu  Thr Ser Lys Ile Asp  Pro Ser Thr
        995                 1000                1005

Gly Phe  Val Asn Leu Leu Lys  Thr Lys Tyr Thr Ser  Ile Ala Asp
    1010                1015                1020

Ser Lys  Lys Phe Ile Ser Ser  Phe Asp Arg Ile Met  Tyr Val Pro
    1025                1030                1035

Glu Glu  Asp Leu Phe Glu Phe  Ala Leu Asp Tyr Lys  Asn Phe Ser
    1040                1045                1050

Arg Thr  Asp Ala Asp Tyr Ile  Lys Lys Trp Lys Leu  Tyr Ser Tyr
    1055                1060                1065

Gly Asn  Arg Ile Arg Ile Phe  Arg Asn Pro Lys Lys  Asn Asn Val
    1070                1075                1080

Phe Asp  Trp Glu Glu Val Cys  Leu Thr Ser Ala Tyr  Lys Glu Leu
    1085                1090                1095

Phe Asn  Lys Tyr Gly Ile Asn  Tyr Gln Gln Gly Asp  Ile Arg Ala
    1100                1105                1110

Leu Leu  Cys Glu Gln Ser Asp  Lys Ala Phe Tyr Ser  Ser Phe Met
    1115                1120                1125

Ala Leu  Met Ser Leu Met Leu  Gln Met Arg Asn Ser  Ile Thr Gly
    1130                1135                1140

Arg Thr  Asp Val Asp Phe Leu  Ile Ser Pro Val Lys  Asn Ser Asp
    1145                1150                1155

Gly Ile  Phe Tyr Asp Ser Arg  Asn Tyr Glu Ala Gln  Glu Asn Ala
    1160                1165                1170
```

```
Ile Leu  Pro Lys Asn Ala Asp  Ala Asn Gly Ala Tyr  Asn Ile Ala
    1175                1180                1185


Arg Lys  Val Leu Trp Ala Ile  Gly Gln Phe Lys Lys  Ala Glu Asp
    1190                1195                1200


Glu Lys  Leu Asp Lys Val Lys  Ile Ala Ile Ser Asn  Lys Glu Trp
    1205                1210                1215


Leu Glu  Tyr Ala Gln Thr Ser  Val Lys His
    1220                1225


<210>  2
<211>  1307
<212>  PRT
<213>  Acidaminococcus sp.

<400>  2

Met Thr Gln Phe Glu Gly Phe Thr Asn Leu Tyr Gln Val Ser Lys Thr
1               5                   10                  15


Leu Arg Phe Glu Leu Ile Pro Gln Gly Lys Thr Leu Lys His Ile Gln
            20                  25                  30


Glu Gln Gly Phe Ile Glu Glu Asp Lys Ala Arg Asn Asp His Tyr Lys
        35                  40                  45


Glu Leu Lys Pro Ile Ile Asp Arg Ile Tyr Lys Thr Tyr Ala Asp Gln
    50                  55                  60


Cys Leu Gln Leu Val Gln Leu Asp Trp Glu Asn Leu Ser Ala Ala Ile
65                  70                  75                  80


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


Glu Glu Gln Ala Thr Tyr Arg Asn Ala Ile His Asp Tyr Phe Ile Gly
            100                 105                 110


Arg Thr Asp Asn Leu Thr Asp Ala Ile Asn Lys Arg His Ala Glu Ile
        115                 120                 125
```

```
Tyr Lys Gly Leu Phe Lys Ala Glu Leu Phe Asn Gly Lys Val Leu Lys
    130                 135                 140

Gln Leu Gly Thr Val Thr Thr Thr Glu His Glu Asn Ala Leu Leu Arg
145                 150                 155                 160

Ser Phe Asp Lys Phe Thr Thr Tyr Phe Ser Gly Phe Tyr Glu Asn Arg
                165                 170                 175

Lys Asn Val Phe Ser Ala Glu Asp Ile Ser Thr Ala Ile Pro His Arg
            180                 185                 190

Ile Val Gln Asp Asn Phe Pro Lys Phe Lys Glu Asn Cys His Ile Phe
        195                 200                 205

Thr Arg Leu Ile Thr Ala Val Pro Ser Leu Arg Glu His Phe Glu Asn
    210                 215                 220

Val Lys Lys Ala Ile Gly Ile Phe Val Ser Thr Ser Ile Glu Glu Val
225                 230                 235                 240

Phe Ser Phe Pro Phe Tyr Asn Gln Leu Leu Thr Gln Thr Gln Ile Asp
                245                 250                 255

Leu Tyr Asn Gln Leu Leu Gly Gly Ile Ser Arg Glu Ala Gly Thr Glu
            260                 265                 270

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

Asp Glu Thr Ala His Ile Ile Ala Ser Leu Pro His Arg Phe Ile Pro
    290                 295                 300

Leu Phe Lys Gln Ile Leu Ser Asp Arg Asn Thr Leu Ser Phe Ile Leu
305                 310                 315                 320

Glu Glu Phe Lys Ser Asp Glu Glu Val Ile Gln Ser Phe Cys Lys Tyr
                325                 330                 335

Lys Thr Leu Leu Arg Asn Glu Asn Val Leu Glu Thr Ala Glu Ala Leu
```

```
            340                 345                 350

Phe Asn Glu Leu Asn Ser Ile Asp Leu Thr His Ile Phe Ile Ser His
        355                 360                 365

Lys Lys Leu Glu Thr Ile Ser Ser Ala Leu Cys Asp His Trp Asp Thr
    370                 375                 380

Leu Arg Asn Ala Leu Tyr Glu Arg Arg Ile Ser Glu Leu Thr Gly Lys
385                 390                 395                 400

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

Asp Ile Asn Leu Gln Glu Ile Ile Ser Ala Ala Gly Lys Glu Leu Ser
            420                 425                 430

Glu Ala Phe Lys Gln Lys Thr Ser Glu Ile Leu Ser His Ala His Ala
        435                 440                 445

Ala Leu Asp Gln Pro Leu Pro Thr Thr Leu Lys Lys Gln Glu Glu Lys
    450                 455                 460

Glu Ile Leu Lys Ser Gln Leu Asp Ser Leu Leu Gly Leu Tyr His Leu
465                 470                 475                 480

Leu Asp Trp Phe Ala Val Asp Glu Ser Asn Glu Val Asp Pro Glu Phe
                485                 490                 495

Ser Ala Arg Leu Thr Gly Ile Lys Leu Glu Met Glu Pro Ser Leu Ser
            500                 505                 510

Phe Tyr Asn Lys Ala Arg Asn Tyr Ala Thr Lys Lys Pro Tyr Ser Val
        515                 520                 525

Glu Lys Phe Lys Leu Asn Phe Gln Met Pro Thr Leu Ala Ser Gly Trp
    530                 535                 540

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

```
Gly Leu Tyr Tyr Leu Gly Ile Met Pro Lys Gln Lys Gly Arg Tyr Lys
                565                 570                 575

Ala Leu Ser Phe Glu Pro Thr Glu Lys Thr Ser Glu Gly Phe Asp Lys
            580                 585                 590

Met Tyr Tyr Asp Tyr Phe Pro Asp Ala Ala Lys Met Ile Pro Lys Cys
        595                 600                 605

Ser Thr Gln Leu Lys Ala Val Thr Ala His Phe Gln Thr His Thr Thr
    610                 615                 620

Pro Ile Leu Leu Ser Asn Asn Phe Ile Glu Pro Leu Glu Ile Thr Lys
625                 630                 635                 640

Glu Ile Tyr Asp Leu Asn Asn Pro Glu Lys Glu Pro Lys Lys Phe Gln
                645                 650                 655

Thr Ala Tyr Ala Lys Lys Thr Gly Asp Gln Lys Gly Tyr Arg Glu Ala
            660                 665                 670

Leu Cys Lys Trp Ile Asp Phe Thr Arg Asp Phe Leu Ser Lys Tyr Thr
        675                 680                 685

Lys Thr Thr Ser Ile Asp Leu Ser Ser Leu Arg Pro Ser Ser Gln Tyr
    690                 695                 700

Lys Asp Leu Gly Glu Tyr Tyr Ala Glu Leu Asn Pro Leu Leu Tyr His
705                 710                 715                 720

Ile Ser Phe Gln Arg Ile Ala Glu Lys Glu Ile Met Asp Ala Val Glu
                725                 730                 735

Thr Gly Lys Leu Tyr Leu Phe Gln Ile Tyr Asn Lys Asp Phe Ala Lys
            740                 745                 750

Gly His His Gly Lys Pro Asn Leu His Thr Leu Tyr Trp Thr Gly Leu
        755                 760                 765

Phe Ser Pro Glu Asn Leu Ala Lys Thr Ser Ile Lys Leu Asn Gly Gln
```

```
    770                 775                 780

Ala Glu Leu Phe Tyr Arg Pro Lys Ser Arg Met Lys Arg Met Ala His
785                 790                 795                 800

Arg Leu Gly Glu Lys Met Leu Asn Lys Lys Leu Lys Asp Gln Lys Thr
                805                 810                 815

Pro Ile Pro Asp Thr Leu Tyr Gln Glu Leu Tyr Asp Tyr Val Asn His
            820                 825                 830

Arg Leu Ser His Asp Leu Ser Asp Glu Ala Arg Ala Leu Leu Pro Asn
        835                 840                 845

Val Ile Thr Lys Glu Val Ser His Glu Ile Ile Lys Asp Arg Arg Phe
    850                 855                 860

Thr Ser Asp Lys Phe Phe Phe His Val Pro Ile Thr Leu Asn Tyr Gln
865                 870                 875                 880

Ala Ala Asn Ser Pro Ser Lys Phe Asn Gln Arg Val Asn Ala Tyr Leu
                885                 890                 895

Lys Glu His Pro Glu Thr Pro Ile Ile Gly Ile Asp Arg Gly Glu Arg
            900                 905                 910

Asn Leu Ile Tyr Ile Thr Val Ile Asp Ser Thr Gly Lys Ile Leu Glu
        915                 920                 925

Gln Arg Ser Leu Asn Thr Ile Gln Gln Phe Asp Tyr Gln Lys Lys Leu
    930                 935                 940

Asp Asn Arg Glu Lys Glu Arg Val Ala Ala Arg Gln Ala Trp Ser Val
945                 950                 955                 960

Val Gly Thr Ile Lys Asp Leu Lys Gln Gly Tyr Leu Ser Gln Val Ile
                965                 970                 975

His Glu Ile Val Asp Leu Met Ile His Tyr Gln Ala Val Val Val Leu
            980                 985                 990
```

```
Glu Asn Leu Asn Phe Gly Phe Lys  Ser Lys Arg Thr Gly  Ile Ala Glu
        995                 1000                1005


Lys Ala  Val Tyr Gln Gln Phe  Glu Lys Met Leu Ile  Asp Lys Leu
    1010                1015                1020


Asn Cys  Leu Val Leu Lys Asp  Tyr Pro Ala Glu Lys  Val Gly Gly
    1025                1030                1035


Val Leu  Asn Pro Tyr Gln Leu  Thr Asp Gln Phe Thr  Ser Phe Ala
    1040                1045                1050


Lys Met  Gly Thr Gln Ser Gly  Phe Leu Phe Tyr Val  Pro Ala Pro
    1055                1060                1065


Tyr Thr  Ser Lys Ile Asp Pro  Leu Thr Gly Phe Val  Asp Pro Phe
    1070                1075                1080


Val Trp  Lys Thr Ile Lys Asn  His Glu Ser Arg Lys  His Phe Leu
    1085                1090                1095


Glu Gly  Phe Asp Phe Leu His  Tyr Asp Val Lys Thr  Gly Asp Phe
    1100                1105                1110


Ile Leu  His Phe Lys Met Asn  Arg Asn Leu Ser Phe  Gln Arg Gly
    1115                1120                1125


Leu Pro  Gly Phe Met Pro Ala  Trp Asp Ile Val Phe  Glu Lys Asn
    1130                1135                1140


Glu Thr  Gln Phe Asp Ala Lys  Gly Thr Pro Phe Ile  Ala Gly Lys
    1145                1150                1155


Arg Ile  Val Pro Val Ile Glu  Asn His Arg Phe Thr  Gly Arg Tyr
    1160                1165                1170


Arg Asp  Leu Tyr Pro Ala Asn  Glu Leu Ile Ala Leu  Leu Glu Glu
    1175                1180                1185


Lys Gly  Ile Val Phe Arg Asp  Gly Ser Asn Ile Leu  Pro Lys Leu
```

```
    1190                1195                1200


Leu Glu  Asn Asp Asp Ser His  Ala Ile Asp Thr Met  Val Ala Leu
    1205                1210                1215


Ile Arg  Ser Val Leu Gln Met  Arg Asn Ser Asn Ala  Ala Thr Gly
    1220                1225                1230


Glu Asp  Tyr Ile Asn Ser Pro  Val Arg Asp Leu Asn  Gly Val Cys
    1235                1240                1245


Phe Asp  Ser Arg Phe Gln Asn  Pro Glu Trp Pro Met  Asp Ala Asp
    1250                1255                1260


Ala Asn  Gly Ala Tyr His Ile  Ala Leu Lys Gly Gln  Leu Leu Leu
    1265                1270                1275


Asn His  Leu Lys Glu Ser Lys  Asp Leu Lys Leu Gln  Asn Gly Ile
    1280                1285                1290


Ser Asn  Gln Asp Trp Leu Ala  Tyr Ile Gln Glu Leu  Arg Asn
    1295                1300                1305


<210>  3
<211>  1241
<212>  PRT
<213>  utyrivibrio proteoclasticus

<400>  3

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


Leu Arg Leu Glu Leu Arg Pro Gln Gly Lys Thr Leu Arg Asn Ile Lys
            20                  25                  30


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


Arg Leu Lys Pro Val Ile Asp Glu Gly Ile Lys Asp Ile Ala Arg Asp
    50                  55                  60
```

```
Thr Leu Lys Asn Cys Glu Leu Ser Phe Glu Lys Leu Tyr Glu His Phe
65                  70                  75                  80

Leu Ser Gly Asp Lys Lys Ala Tyr Ala Lys Glu Ser Glu Arg Leu Lys
                85                  90                  95

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

Lys Ile Ser Glu Ile Asn Ser Ala Lys Tyr Leu Asn Gly Val Leu Tyr
        115                 120                 125

Asp Phe Ile Asp Lys Thr His Lys Asp Ser Glu Glu Lys Gln Asn Ile
    130                 135                 140

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

Phe Leu Thr Ser Arg Ile Thr Thr Leu Glu Gln Ser Met Pro Lys Arg
                165                 170                 175

Val Ile Glu Asn Phe Glu Ile Tyr Ala Ala Asn Ile Pro Lys Met Gln
            180                 185                 190

Asp Ala Leu Glu Arg Gly Ala Val Ser Phe Ala Ile Glu Tyr Glu Ser
        195                 200                 205

Ile Cys Ser Val Asp Tyr Tyr Asn Gln Ile Leu Ser Gln Glu Asp Ile
    210                 215                 220

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

Lys Glu Lys Gly Ile Asn Gln Thr Ile Ser Glu Lys Asn Ile Lys Ile
                245                 250                 255

Lys Ser Glu His Leu Glu Glu Lys Pro Phe Arg Ile Leu Lys Gln Leu
            260                 265                 270

His Lys Gln Ile Leu Glu Glu Arg Glu Lys Ala Phe Thr Ile Asp His
        275                 280                 285
```

```
Ile Asp Ser Asp Glu Glu Val Val Gln Val Thr Lys Glu Ala Phe Glu
    290                 295                 300

Gln Thr Lys Glu Gln Trp Glu Asn Ile Lys Lys Ile Asn Gly Phe Tyr
305                 310                 315                 320

Ala Lys Asp Pro Gly Asp Ile Thr Leu Phe Ile Val Val Gly Pro Asn
                325                 330                 335

Gln Thr His Val Leu Ser Gln Leu Ile Tyr Gly Glu His Asp Arg Ile
            340                 345                 350

Arg Leu Leu Leu Glu Glu Tyr Glu Lys Asn Thr Leu Glu Val Leu Pro
        355                 360                 365

Arg Arg Thr Lys Ser Glu Asp Ala Arg Tyr Asp Lys Phe Val Asn Ala
    370                 375                 380

Val Pro Lys Lys Val Ala Lys Glu Ser His Thr Phe Asp Gly Leu Gln
385                 390                 395                 400

Lys Met Thr Gly Asp Asp Arg Leu Phe Ile Leu Tyr Arg Asp Glu Leu
                405                 410                 415

Ala Arg Asn Tyr Met Arg Ile Lys Glu Ala Tyr Gly Thr Phe Glu Arg
            420                 425                 430

Asp Ile Leu Lys Ser Arg Arg Gly Ile Lys Gly Asn Arg Asp Val Gln
        435                 440                 445

Glu Ser Leu Val Ser Phe Tyr Asp Glu Leu Thr Lys Phe Arg Ser Ala
    450                 455                 460

Leu Arg Ile Ile Asn Ser Gly Asn Asp Glu Lys Ala Asp Pro Ile Phe
465                 470                 475                 480

Tyr Asn Thr Phe Asp Gly Ile Phe Glu Lys Ala Asn Arg Thr Tyr Lys
                485                 490                 495
```

```
Ala Glu Asn Leu Cys Arg Asn Tyr Val Thr Lys Ser Pro Ala Asp Asp
            500                 505                 510

Ala Arg Ile Met Ala Ser Cys Leu Gly Thr Pro Ala Arg Leu Arg Thr
        515                 520                 525

His Trp Trp Asn Gly Glu Glu Asn Phe Ala Ile Asn Asp Val Ala Met
    530                 535                 540

Ile Arg Arg Gly Asp Glu Tyr Tyr Tyr Phe Val Leu Thr Pro Asp Val
545                 550                 555                 560

Lys Pro Val Asp Leu Lys Thr Lys Asp Glu Thr Asp Ala Gln Ile Phe
                565                 570                 575

Val Gln Arg Lys Gly Ala Lys Ser Phe Leu Gly Leu Pro Lys Ala Leu
            580                 585                 590

Phe Lys Cys Ile Leu Glu Pro Tyr Phe Glu Ser Pro Glu His Lys Asn
        595                 600                 605

Asp Lys Asn Cys Val Ile Glu Glu Tyr Val Ser Lys Pro Leu Thr Ile
    610                 615                 620

Asp Arg Arg Ala Tyr Asp Ile Phe Lys Asn Gly Thr Phe Lys Lys Thr
625                 630                 635                 640

Asn Ile Gly Ile Asp Gly Leu Thr Glu Glu Lys Phe Lys Asp Asp Cys
                645                 650                 655

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

Tyr Ser Cys Phe Asn Met Ser Gly Leu Lys Arg Ala Asp Glu Tyr Asn
        675                 680                 685

Asp Ile Gly Glu Phe Phe Ser Asp Val Asp Thr Arg Leu Cys Thr Met
    690                 695                 700

Glu Trp Ile Pro Val Ser Phe Glu Arg Ile Asn Asp Met Val Asp Lys
705                 710                 715                 720
```

```
Lys Glu Gly Leu Leu Phe Leu Val Arg Ser Met Phe Leu Tyr Asn Arg
                725                 730                 735

Pro Arg Lys Pro Tyr Glu Arg Thr Phe Ile Gln Leu Phe Ser Asp Ser
            740                 745                 750

Asn Met Glu His Thr Ser Met Leu Leu Asn Ser Arg Ala Met Ile Gln
        755                 760                 765

Tyr Arg Ala Ala Ser Leu Pro Arg Arg Val Thr His Lys Lys Gly Ser
    770                 775                 780

Ile Leu Val Ala Leu Arg Asp Ser Asn Gly Glu His Ile Pro Met His
785                 790                 795                 800

Ile Arg Glu Ala Ile Tyr Lys Met Lys Asn Asn Phe Asp Ile Ser Ser
                805                 810                 815

Glu Asp Phe Ile Met Ala Lys Ala Tyr Leu Ala Glu His Asp Val Ala
            820                 825                 830

Ile Lys Lys Ala Asn Glu Asp Ile Ile Arg Asn Arg Arg Tyr Thr Glu
        835                 840                 845

Asp Lys Phe Phe Leu Ser Leu Ser Tyr Thr Lys Asn Ala Asp Ile Ser
    850                 855                 860

Ala Arg Thr Leu Asp Tyr Ile Asn Asp Lys Val Glu Glu Asp Thr Gln
865                 870                 875                 880

Asp Ser Arg Met Ala Val Ile Val Thr Arg Asn Leu Lys Asp Leu Thr
                885                 890                 895

Tyr Val Ala Val Val Asp Glu Lys Asn Asn Val Leu Glu Glu Lys Ser
            900                 905                 910

Leu Asn Glu Ile Asp Gly Val Asn Tyr Arg Glu Leu Leu Lys Glu Arg
        915                 920                 925
```

```
Thr Lys Ile Lys Tyr His Asp Lys Thr Arg Leu Trp Gln Tyr Asp Val
    930                 935                 940

Ser Ser Lys Gly Leu Lys Glu Ala Tyr Val Glu Leu Ala Val Thr Gln
945                 950                 955                 960

Ile Ser Lys Leu Ala Thr Lys Tyr Asn Ala Val Val Val Val Glu Ser
                965                 970                 975

Met Ser Ser Thr Phe Lys Asp Lys Phe Ser Phe Leu Asp Glu Gln Ile
            980                 985                 990

Phe Lys Ala Phe Glu Ala Arg Leu  Cys Ala Arg Met Ser  Asp Leu Ser
        995                 1000                1005

Phe Asn  Thr Ile Lys Glu Gly  Glu Ala Gly Ser Ile  Ser Asn Pro
    1010                1015                1020

Ile Gln  Val Ser Asn Asn Asn  Gly Asn Ser Tyr Gln  Asp Gly Val
    1025                1030                1035

Ile Tyr  Phe Leu Asn Asn Ala  Tyr Thr Arg Thr Leu  Cys Pro Asp
    1040                1045                1050

Thr Gly  Phe Val Asp Val Phe  Asp Lys Thr Arg Leu  Ile Thr Met
    1055                1060                1065

Gln Ser  Lys Arg Gln Phe Phe  Ala Lys Met Lys Asp  Ile Arg Ile
    1070                1075                1080

Asp Asp  Gly Glu Met Leu Phe  Thr Phe Asn Leu Glu  Glu Tyr Pro
    1085                1090                1095

Thr Lys  Arg Leu Leu Asp Arg  Lys Glu Trp Thr Val  Lys Ile Ala
    1100                1105                1110

Gly Asp  Gly Ser Tyr Phe Asp  Lys Asp Lys Gly Glu  Tyr Val Tyr
    1115                1120                1125

Val Asn  Asp Ile Val Arg Glu  Gln Ile Ile Pro Ala  Leu Leu Glu
    1130                1135                1140
```

```
Asp Lys  Ala Val Phe Asp Gly  Asn Met Ala Glu Lys  Phe Leu Asp
    1145                1150                1155


Lys Thr  Ala Ile Ser Gly Lys  Ser Val Glu Leu Ile  Tyr Lys Trp
    1160                1165                1170


Phe Ala  Asn Ala Leu Tyr Gly  Ile Ile Thr Lys Lys  Asp Gly Glu
    1175                1180                1185


Lys Ile  Tyr Arg Ser Pro Ile  Thr Gly Thr Glu Ile  Asp Val Ser
    1190                1195                1200


Lys Asn  Thr Thr Tyr Asn Phe  Gly Lys Lys Phe Met  Phe Lys Gln
    1205                1210                1215


Glu Tyr  Arg Gly Asp Gly Asp  Phe Leu Asp Ala Phe  Leu Asn Tyr
    1220                1225                1230


Met Gln  Ala Gln Asp Ile Ala  Val
    1235                1240


<210>  4
<211>  1238
<212>  PRT
<213>  Candidatus Methanoplasma termitum

<400>  4

Met Asn Asn Tyr Asp Glu Phe Thr Lys Leu Tyr Pro Ile Gln Lys Thr
1               5                  10                  15


Ile Arg Phe Glu Leu Lys Pro Gln Gly Arg Thr Met Glu His Leu Glu
            20                  25                  30


Thr Phe Asn Phe Phe Glu Glu Asp Arg Asp Arg Ala Glu Lys Tyr Lys
        35                  40                  45


Ile Leu Lys Glu Ala Ile Asp Glu Tyr His Lys Lys Phe Ile Asp Glu
    50                  55                  60


His Leu Thr Asn Met Ser Leu Asp Trp Asn Ser Leu Lys Gln Ile Ser
```

```
65                  70                  75                  80

Glu Lys Tyr Tyr Lys Ser Arg Glu Glu Lys Asp Lys Lys Val Phe Leu
                85                  90                  95

Ser Glu Gln Lys Arg Met Arg Gln Glu Ile Val Ser Glu Phe Lys Lys
            100                 105                 110

Asp Asp Arg Phe Lys Asp Leu Phe Ser Lys Lys Leu Phe Ser Glu Leu
        115                 120                 125

Leu Lys Glu Glu Ile Tyr Lys Lys Gly Asn His Gln Glu Ile Asp Ala
    130                 135                 140

Leu Lys Ser Phe Asp Lys Phe Ser Gly Tyr Phe Ile Gly Leu His Glu
145                 150                 155                 160

Asn Arg Lys Asn Met Tyr Ser Asp Gly Asp Glu Ile Thr Ala Ile Ser
                165                 170                 175

Asn Arg Ile Val Asn Glu Asn Phe Pro Lys Phe Leu Asp Asn Leu Gln
            180                 185                 190

Lys Tyr Gln Glu Ala Arg Lys Lys Tyr Pro Glu Trp Ile Ile Lys Ala
        195                 200                 205

Glu Ser Ala Leu Val Ala His Asn Ile Lys Met Asp Ile Val Phe Ser
    210                 215                 220

Leu Glu Tyr Phe Asn Lys Val Leu Asn Gln Glu Gly Ile Gln Arg Tyr
225                 230                 235                 240

Asn Leu Ala Leu Gly Gly Tyr Val Thr Lys Ser Gly Glu Lys Met Met
                245                 250                 255

Gly Leu Asn Asp Ala Leu Asn Leu Ala His Gln Ser Glu Lys Ser Ser
            260                 265                 270

Lys Gly Arg Ile His Met Thr Pro Leu Phe Lys Gln Ile Leu Ser Glu
        275                 280                 285
```

```
Lys Glu Ser Phe Ser Tyr Ile Pro Asp Val Phe Thr Glu Asp Ser Gln
    290                 295                 300

Leu Leu Pro Ser Ile Gly Gly Phe Phe Ala Gln Ile Glu Asn Asp Lys
305                 310                 315                 320

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

Glu Tyr Asp Thr Glu Arg Ile Tyr Ile Arg Gln Ala Asp Ile Asn Arg
            340                 345                 350

Val Ser Asn Val Ile Phe Gly Glu Trp Gly Thr Leu Gly Gly Leu Met
        355                 360                 365

Arg Glu Tyr Lys Ala Asp Ser Ile Asn Asp Ile Asn Leu Glu Arg Thr
    370                 375                 380

Cys Lys Lys Val Asp Lys Trp Leu Asp Ser Lys Glu Phe Ala Leu Ser
385                 390                 395                 400

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

Glu Tyr Ile Ser Lys Met Arg Thr Ala Arg Glu Lys Ile Asp Ala Ala
            420                 425                 430

Arg Lys Glu Met Lys Phe Ile Ser Glu Lys Ile Ser Gly Asp Glu Glu
        435                 440                 445

Ser Ile His Ile Ile Lys Thr Leu Leu Asp Ser Val Gln Gln Phe Leu
    450                 455                 460

His Phe Phe Asn Leu Phe Lys Ala Arg Gln Asp Ile Pro Leu Asp Gly
465                 470                 475                 480

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

Val Pro Leu Tyr Asn Lys Val Arg Asn Tyr Leu Thr Lys Asn Asn Leu
```

```
            500                 505                 510

Asn Thr Lys Lys Ile Lys Leu Asn Phe Lys Asn Pro Thr Leu Ala Asn
        515                 520                 525

Gly Trp Asp Gln Asn Lys Val Tyr Asp Tyr Ala Ser Leu Ile Phe Leu
    530                 535                 540

Arg Asp Gly Asn Tyr Tyr Leu Gly Ile Ile Asn Pro Lys Arg Lys Lys
545                 550                 555                 560

Asn Ile Lys Phe Glu Gln Gly Ser Gly Asn Gly Pro Phe Tyr Arg Lys
                565                 570                 575

Met Val Tyr Lys Gln Ile Pro Gly Pro Asn Lys Asn Leu Arg Pro Val
            580                 585                 590

Phe Leu Thr Ser Thr Lys Gly Lys Lys Glu Tyr Lys Pro Ser Lys Glu
        595                 600                 605

Ile Ile Glu Gly Tyr Glu Ala Asp Lys His Ile Arg Gly Asp Lys Phe
    610                 615                 620

Asp Leu Asp Phe Cys His Lys Leu Ile Asp Phe Phe Lys Glu Ser Ile
625                 630                 635                 640

Glu Lys His Lys Asp Trp Ser Lys Phe Asn Phe Tyr Phe Ser Pro Thr
                645                 650                 655

Glu Ser Tyr Gly Asp Ile Ser Glu Phe Tyr Leu Asp Val Glu Lys Gln
            660                 665                 670

Gly Tyr Arg Met His Phe Glu Asn Ile Ser Ala Glu Thr Ile Asp Glu
        675                 680                 685

Tyr Val Glu Lys Gly Asp Leu Phe Leu Phe Gln Ile Tyr Asn Lys Asp
    690                 695                 700

Phe Val Lys Ala Ala Thr Gly Lys Lys Asp Met His Thr Ile Tyr Trp
705                 710                 715                 720
```

```
Asn Ala Ala Phe Ser Pro Glu Asn Leu Gln Asp Val Val Val Lys Leu
                725                 730                 735

Asn Gly Glu Ala Glu Leu Phe Tyr Arg Asp Lys Ser Asp Ile Lys Glu
            740                 745                 750

Ile Val His Arg Glu Gly Glu Ile Leu Val Asn Arg Thr Tyr Asn Gly
        755                 760                 765

Arg Thr Pro Val Pro Asp Lys Ile His Lys Lys Leu Thr Asp Tyr His
    770                 775                 780

Asn Gly Arg Thr Lys Asp Leu Gly Glu Ala Lys Glu Tyr Leu Asp Lys
785                 790                 795                 800

Val Arg Tyr Phe Lys Ala His Tyr Asp Ile Thr Lys Asp Arg Arg Tyr
                805                 810                 815

Leu Asn Asp Lys Ile Tyr Phe His Val Pro Leu Thr Leu Asn Phe Lys
            820                 825                 830

Ala Asn Gly Lys Lys Asn Leu Asn Lys Met Val Ile Glu Lys Phe Leu
        835                 840                 845

Ser Asp Glu Lys Ala His Ile Ile Gly Ile Asp Arg Gly Glu Arg Asn
    850                 855                 860

Leu Leu Tyr Tyr Ser Ile Ile Asp Arg Ser Gly Lys Ile Ile Asp Gln
865                 870                 875                 880

Gln Ser Leu Asn Val Ile Asp Gly Phe Asp Tyr Arg Glu Lys Leu Asn
                885                 890                 895

Gln Arg Glu Ile Glu Met Lys Asp Ala Arg Gln Ser Trp Asn Ala Ile
            900                 905                 910

Gly Lys Ile Lys Asp Leu Lys Glu Gly Tyr Leu Ser Lys Ala Val His
        915                 920                 925

Glu Ile Thr Lys Met Ala Ile Gln Tyr Asn Ala Ile Val Val Met Glu
```

```
    930                 935                 940

Glu Leu Asn Tyr Gly Phe Lys Arg Gly Arg Phe Lys Val Glu Lys Gln
945                 950                 955                 960

Ile Tyr Gln Lys Phe Glu Asn Met Leu Ile Asp Lys Met Asn Tyr Leu
                965                 970                 975

Val Phe Lys Asp Ala Pro Asp Glu Ser Pro Gly Gly Val Leu Asn Ala
            980                 985                 990

Tyr Gln Leu Thr Asn Pro Leu Glu  Ser Phe Ala Lys Leu  Gly Lys Gln
        995                 1000                1005

Thr Gly  Ile Leu Phe Tyr Val  Pro Ala Ala Tyr Thr  Ser Lys Ile
    1010                1015                1020

Asp Pro  Thr Thr Gly Phe Val  Asn Leu Phe Asn Thr  Ser Ser Lys
    1025                1030                1035

Thr Asn  Ala Gln Glu Arg Lys  Glu Phe Leu Gln Lys  Phe Glu Ser
    1040                1045                1050

Ile Ser  Tyr Ser Ala Lys Asp  Gly Gly Ile Phe Ala  Phe Ala Phe
    1055                1060                1065

Asp Tyr  Arg Lys Phe Gly Thr  Ser Lys Thr Asp His  Lys Asn Val
    1070                1075                1080

Trp Thr  Ala Tyr Thr Asn Gly  Glu Arg Met Arg Tyr  Ile Lys Glu
    1085                1090                1095

Lys Lys  Arg Asn Glu Leu Phe  Asp Pro Ser Lys Glu  Ile Lys Glu
    1100                1105                1110

Ala Leu  Thr Ser Ser Gly Ile  Lys Tyr Asp Gly Gly  Gln Asn Ile
    1115                1120                1125

Leu Pro  Asp Ile Leu Arg Ser  Asn Asn Asn Gly Leu  Ile Tyr Thr
    1130                1135                1140
```

```
Met Tyr  Ser Ser Phe Ile Ala  Ala Ile Gln Met Arg  Val Tyr Asp
    1145                1150                1155


Gly Lys  Glu Asp Tyr Ile Ile  Ser Pro Ile Lys Asn  Ser Lys Gly
    1160                1165                1170


Glu Phe  Phe Arg Thr Asp Pro  Lys Arg Arg Glu Leu  Pro Ile Asp
    1175                1180                1185


Ala Asp  Ala Asn Gly Ala Tyr  Asn Ile Ala Leu Arg  Gly Glu Leu
    1190                1195                1200


Thr Met  Arg Ala Ile Ala Glu  Lys Phe Asp Pro Asp  Ser Glu Lys
    1205                1210                1215


Met Ala  Lys Leu Glu Leu Lys  His Lys Asp Trp Phe  Glu Phe Met
    1220                1225                1230


Gln Thr  Arg Gly Asp
    1235
```

<210> 5
<211> 1281
<212> PRT
<213> Eubacterium eligens

<400> 5

```
Met Asn Gly Asn Arg Ser Ile Val Tyr Arg Glu Phe Val Gly Val Ile
1               5                   10                  15


Pro Val Ala Lys Thr Leu Arg Asn Glu Leu Arg Pro Val Gly His Thr
            20                  25                  30


Gln Glu His Ile Ile Gln Asn Gly Leu Ile Gln Glu Asp Glu Leu Arg
        35                  40                  45


Gln Glu Lys Ser Thr Glu Leu Lys Asn Ile Met Asp Asp Tyr Tyr Arg
    50                  55                  60


Glu Tyr Ile Asp Lys Ser Leu Ser Gly Val Thr Asp Leu Asp Phe Thr
65                  70                  75                  80
```

```
Leu Leu Phe Glu Leu Met Asn Leu Val Gln Ser Ser Pro Ser Lys Asp
                85                  90                  95

Asn Lys Lys Ala Leu Glu Lys Glu Gln Ser Lys Met Arg Glu Gln Ile
            100                 105                 110

Cys Thr His Leu Gln Ser Asp Ser Asn Tyr Lys Asn Ile Phe Asn Ala
        115                 120                 125

Lys Leu Leu Lys Glu Ile Leu Pro Asp Phe Ile Lys Asn Tyr Asn Gln
    130                 135                 140

Tyr Asp Val Lys Asp Lys Ala Gly Lys Leu Glu Thr Leu Ala Leu Phe
145                 150                 155                 160

Asn Gly Phe Ser Thr Tyr Phe Thr Asp Phe Phe Glu Lys Arg Lys Asn
                165                 170                 175

Val Phe Thr Lys Glu Ala Val Ser Thr Ser Ile Ala Tyr Arg Ile Val
            180                 185                 190

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

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

Gln Asp Lys Met Gly Asp Trp Glu Leu Asn Gln Ile Phe Asn Pro Asp
225                 230                 235                 240

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

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

Lys Asn Asn Tyr Asn Leu Phe Lys Met Arg Lys Leu His Lys Gln Ile
        275                 280                 285
```

```
Leu Ala Tyr Thr Ser Thr Ser Phe Glu Val Pro Lys Met Phe Glu Asp
    290                 295                 300

Asp Met Ser Val Tyr Asn Ala Val Asn Ala Phe Ile Asp Glu Thr Glu
305                 310                 315                 320

Lys Gly Asn Ile Ile Gly Lys Leu Lys Asp Ile Val Asn Lys Tyr Asp
                325                 330                 335

Glu Leu Asp Glu Lys Arg Ile Tyr Ile Ser Lys Asp Phe Tyr Glu Thr
            340                 345                 350

Leu Ser Cys Phe Met Ser Gly Asn Trp Asn Leu Ile Thr Gly Cys Val
        355                 360                 365

Glu Asn Phe Tyr Asp Glu Asn Ile His Ala Lys Gly Lys Ser Lys Glu
    370                 375                 380

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

Asp Val Asn Asp Leu Val Glu Lys Tyr Ile Asp Glu Lys Glu Arg Asn
                405                 410                 415

Glu Phe Lys Asn Ser Asn Ala Lys Gln Tyr Ile Arg Glu Ile Ser Asn
            420                 425                 430

Ile Ile Thr Asp Thr Glu Thr Ala His Leu Glu Tyr Asp Asp His Ile
        435                 440                 445

Ser Leu Ile Glu Ser Glu Glu Lys Ala Asp Glu Met Lys Lys Arg Leu
    450                 455                 460

Asp Met Tyr Met Asn Met Tyr His Trp Ala Lys Ala Phe Ile Val Asp
465                 470                 475                 480

Glu Val Leu Asp Arg Asp Glu Met Phe Tyr Ser Asp Ile Asp Asp Ile
                485                 490                 495

Tyr Asn Ile Leu Glu Asn Ile Val Pro Leu Tyr Asn Arg Val Arg Asn
            500                 505                 510
```

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

Gln Ser Pro Thr Leu Ala Asn Gly Trp Ser Gln Ser Lys Glu Phe Asp
    530                 535                 540

Asn Asn Ala Ile Ile Leu Ile Arg Asp Asn Lys Tyr Tyr Leu Ala Ile
545                 550                 555                 560

Phe Asn Ala Lys Asn Lys Pro Asp Lys Lys Ile Ile Gln Gly Asn Ser
                565                 570                 575

Asp Lys Lys Asn Asp Asn Asp Tyr Lys Lys Met Val Tyr Asn Leu Leu
            580                 585                 590

Pro Gly Ala Asn Lys Met Leu Pro Lys Val Phe Leu Ser Lys Lys Gly
        595                 600                 605

Ile Glu Thr Phe Lys Pro Ser Asp Tyr Ile Ile Ser Gly Tyr Asn Ala
    610                 615                 620

His Lys His Ile Lys Thr Ser Glu Asn Phe Asp Ile Ser Phe Cys Arg
625                 630                 635                 640

Asp Leu Ile Asp Tyr Phe Lys Asn Ser Ile Glu Lys His Ala Glu Trp
                645                 650                 655

Arg Lys Tyr Glu Phe Lys Phe Ser Ala Thr Asp Ser Tyr Ser Asp Ile
            660                 665                 670

Ser Glu Phe Tyr Arg Glu Val Glu Met Gln Gly Tyr Arg Ile Asp Trp
        675                 680                 685

Thr Tyr Ile Ser Glu Ala Asp Ile Asn Lys Leu Asp Glu Glu Gly Lys
    690                 695                 700

Ile Tyr Leu Phe Gln Ile Tyr Asn Lys Asp Phe Ala Glu Asn Ser Thr
705                 710                 715                 720
```

```
Gly Lys Glu Asn Leu His Thr Met Tyr Phe Lys Asn Ile Phe Ser Glu
                725                 730                 735

Glu Asn Leu Asp Lys Ile Ile Lys Leu Asn Gly Gln Ala Glu Leu Phe
            740                 745                 750

Tyr Arg Arg Ala Ser Val Lys Asn Pro Val Lys His Lys Lys Asp Ser
        755                 760                 765

Val Leu Val Asn Lys Thr Tyr Lys Asn Gln Leu Asp Asn Gly Asp Val
    770                 775                 780

Val Arg Ile Pro Ile Pro Asp Asp Ile Tyr Asn Glu Ile Tyr Lys Met
785                 790                 795                 800

Tyr Asn Gly Tyr Ile Lys Glu Ser Asp Leu Ser Glu Ala Ala Lys Glu
                805                 810                 815

Tyr Leu Asp Lys Val Glu Val Arg Thr Ala Gln Lys Asp Ile Val Lys
            820                 825                 830

Asp Tyr Arg Tyr Thr Val Asp Lys Tyr Phe Ile His Thr Pro Ile Thr
        835                 840                 845

Ile Asn Tyr Lys Val Thr Ala Arg Asn Asn Val Asn Asp Met Val Val
    850                 855                 860

Lys Tyr Ile Ala Gln Asn Asp Asp Ile His Val Ile Gly Ile Asp Arg
865                 870                 875                 880

Gly Glu Arg Asn Leu Ile Tyr Ile Ser Val Ile Asp Ser His Gly Asn
                885                 890                 895

Ile Val Lys Gln Lys Ser Tyr Asn Ile Leu Asn Asn Tyr Asp Tyr Lys
            900                 905                 910

Lys Lys Leu Val Glu Lys Glu Lys Thr Arg Glu Tyr Ala Arg Lys Asn
        915                 920                 925

Trp Lys Ser Ile Gly Asn Ile Lys Glu Leu Lys Glu Gly Tyr Ile Ser
    930                 935                 940
```

```
Gly Val Val His Glu Ile Ala Met Leu Ile Val Glu Tyr Asn Ala Ile
945                 950                 955                 960


Ile Ala Met Glu Asp Leu Asn Tyr Gly Phe Lys Arg Gly Arg Phe Lys
                965                 970                 975


Val Glu Arg Gln Val Tyr Gln Lys Phe Glu Ser Met Leu Ile Asn Lys
            980                 985                 990


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


Leu Leu  Lys Gly Tyr Gln Leu  Thr Tyr Val Pro Asp  Asn Ile Lys
    1010                1015                1020


Asn Leu  Gly Lys Gln Cys Gly  Val Ile Phe Tyr Val  Pro Ala Ala
    1025                1030                1035


Phe Thr  Ser Lys Ile Asp Pro  Ser Thr Gly Phe Ile  Ser Ala Phe
    1040                1045                1050


Asn Phe  Lys Ser Ile Ser Thr  Asn Ala Ser Arg Lys  Gln Phe Phe
    1055                1060                1065


Met Gln  Phe Asp Glu Ile Arg  Tyr Cys Ala Glu Lys  Asp Met Phe
    1070                1075                1080


Ser Phe  Gly Phe Asp Tyr Asn  Asn Phe Asp Thr Tyr  Asn Ile Thr
    1085                1090                1095


Met Gly  Lys Thr Gln Trp Thr  Val Tyr Thr Asn Gly  Glu Arg Leu
    1100                1105                1110


Gln Ser  Glu Phe Asn Asn Ala  Arg Arg Thr Gly Lys  Thr Lys Ser
    1115                1120                1125


Ile Asn  Leu Thr Glu Thr Ile  Lys Leu Leu Leu Glu  Asp Asn Glu
    1130                1135                1140
```

```
Ile Asn  Tyr Ala Asp Gly His  Asp Ile Arg Ile Asp  Met Glu Lys
    1145                1150                1155


Met Asp  Glu Asp Lys Lys Ser  Glu Phe Phe Ala Gln  Leu Leu Ser
    1160                1165                1170


Leu Tyr  Lys Leu Thr Val Gln  Met Arg Asn Ser Tyr  Thr Glu Ala
    1175                1180                1185


Glu Glu  Gln Glu Asn Gly Ile  Ser Tyr Asp Lys Ile  Ile Ser Pro
    1190                1195                1200


Val Ile  Asn Asp Glu Gly Glu  Phe Phe Asp Ser Asp  Asn Tyr Lys
    1205                1210                1215


Glu Ser  Asp Asp Lys Glu Cys  Lys Met Pro Lys Asp  Ala Asp Ala
    1220                1225                1230


Asn Gly  Ala Tyr Cys Ile Ala  Leu Lys Gly Leu Tyr  Glu Val Leu
    1235                1240                1245


Lys Ile  Lys Ser Glu Trp Thr  Glu Asp Gly Phe Asp  Arg Asn Cys
    1250                1255                1260


Leu Lys  Leu Pro His Ala Glu  Trp Leu Asp Phe Ile  Gln Asn Lys
    1265                1270                1275


Arg Tyr  Glu
    1280


<210>  6
<211>  1300
<212>  PRT
<213>  Francisella novicida

<400>  6

Met Ser Ile Tyr Gln Glu Phe Val Asn Lys Tyr Ser Leu Ser Lys Thr
1               5                  10                  15


Leu Arg Phe Glu Leu Ile Pro Gln Gly Lys Thr Leu Glu Asn Ile Lys
            20                  25                  30
```

```
Ala Arg Gly Leu Ile Leu Asp Asp Glu Lys Arg Ala Lys Asp Tyr Lys
        35                  40                  45

Lys Ala Lys Gln Ile Ile Asp Lys Tyr His Gln Phe Phe Ile Glu Glu
    50                  55                  60

Ile Leu Ser Ser Val Cys Ile Ser Glu Asp Leu Leu Gln Asn Tyr Ser
65                  70                  75                  80

Asp Val Tyr Phe Lys Leu Lys Lys Ser Asp Asp Asp Asn Leu Gln Lys
                85                  90                  95

Asp Phe Lys Ser Ala Lys Asp Thr Ile Lys Lys Gln Ile Ser Glu Tyr
            100                 105                 110

Ile Lys Asp Ser Glu Lys Phe Lys Asn Leu Phe Asn Gln Asn Leu Ile
        115                 120                 125

Asp Ala Lys Lys Gly Gln Glu Ser Asp Leu Ile Leu Trp Leu Lys Gln
    130                 135                 140

Ser Lys Asp Asn Gly Ile Glu Leu Phe Lys Ala Asn Ser Asp Ile Thr
145                 150                 155                 160

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

Thr Tyr Phe Lys Gly Phe His Glu Asn Arg Lys Val Asn Tyr Ser Ser
            180                 185                 190

Asn Asp Ile Pro Thr Ser Ile Ile Tyr Arg Ile Val Asp Asp Asn Leu
        195                 200                 205

Pro Lys Phe Leu Glu Asn Lys Ala Lys Tyr Glu Ser Leu Lys Asp Lys
    210                 215                 220

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

Glu Leu Thr Phe Asp Ile Asp Tyr Lys Thr Ser Glu Val Asn Gln Arg
```

```
                245                 250                 255

Val Phe Ser Leu Asp Glu Val Phe Glu Ile Ala Asn Phe Asn Asn Tyr
            260                 265                 270

Leu Asn Gln Ser Gly Ile Thr Lys Phe Asn Thr Ile Ile Gly Gly Lys
        275                 280                 285

Phe Val Asn Gly Glu Asn Thr Lys Arg Lys Gly Ile Asn Glu Tyr Ile
    290                 295                 300

Asn Leu Tyr Ser Gln Gln Ile Asn Asp Lys Thr Leu Lys Lys Tyr Lys
305                 310                 315                 320

Met Ser Val Leu Phe Lys Gln Ile Leu Ser Asp Thr Glu Ser Lys Ser
                325                 330                 335

Phe Val Ile Asp Lys Leu Glu Asp Asp Ser Asp Val Val Thr Thr Met
            340                 345                 350

Gln Ser Phe Tyr Glu Gln Ile Ala Ala Phe Lys Thr Val Glu Glu Lys
        355                 360                 365

Ser Ile Lys Glu Thr Leu Ser Leu Leu Phe Asp Asp Leu Lys Ala Gln
    370                 375                 380

Lys Leu Asp Leu Ser Lys Ile Tyr Phe Lys Asn Asp Lys Ser Leu Thr
385                 390                 395                 400

Asp Leu Ser Gln Gln Val Phe Asp Asp Tyr Ser Val Ile Gly Thr Ala
                405                 410                 415

Val Leu Glu Tyr Ile Thr Gln Gln Ile Ala Pro Lys Asn Leu Asp Asn
            420                 425                 430

Pro Ser Lys Lys Glu Gln Glu Leu Ile Ala Lys Lys Thr Glu Lys Ala
        435                 440                 445

Lys Tyr Leu Ser Leu Glu Thr Ile Lys Leu Ala Leu Glu Glu Phe Asn
    450                 455                 460
```

```
Lys His Arg Asp Ile Asp Lys Gln Cys Arg Phe Glu Glu Ile Leu Ala
465                 470                 475                 480

Asn Phe Ala Ala Ile Pro Met Ile Phe Asp Glu Ile Ala Gln Asn Lys
                485                 490                 495

Asp Asn Leu Ala Gln Ile Ser Ile Lys Tyr Gln Asn Gln Gly Lys Lys
            500                 505                 510

Asp Leu Leu Gln Ala Ser Ala Glu Asp Asp Val Lys Ala Ile Lys Asp
        515                 520                 525

Leu Leu Asp Gln Thr Asn Asn Leu Leu His Lys Leu Lys Ile Phe His
    530                 535                 540

Ile Ser Gln Ser Glu Asp Lys Ala Asn Ile Leu Asp Lys Asp Glu His
545                 550                 555                 560

Phe Tyr Leu Val Phe Glu Glu Cys Tyr Phe Glu Leu Ala Asn Ile Val
                565                 570                 575

Pro Leu Tyr Asn Lys Ile Arg Asn Tyr Ile Thr Gln Lys Pro Tyr Ser
            580                 585                 590

Asp Glu Lys Phe Lys Leu Asn Phe Glu Asn Ser Thr Leu Ala Asn Gly
        595                 600                 605

Trp Asp Lys Asn Lys Glu Pro Asp Asn Thr Ala Ile Leu Phe Ile Lys
    610                 615                 620

Asp Asp Lys Tyr Tyr Leu Gly Val Met Asn Lys Lys Asn Asn Lys Ile
625                 630                 635                 640

Phe Asp Asp Lys Ala Ile Lys Glu Asn Lys Gly Glu Gly Tyr Lys Lys
                645                 650                 655

Ile Val Tyr Lys Leu Leu Pro Gly Ala Asn Lys Met Leu Pro Lys Val
            660                 665                 670

Phe Phe Ser Ala Lys Ser Ile Lys Phe Tyr Asn Pro Ser Glu Asp Ile
```

```
        675                 680                 685

Leu Arg Ile Arg Asn His Ser Thr His Thr Lys Asn Gly Ser Pro Gln
    690                 695                 700

Lys Gly Tyr Glu Lys Phe Glu Phe Asn Ile Glu Asp Cys Arg Lys Phe
705                 710                 715                 720

Ile Asp Phe Tyr Lys Gln Ser Ile Ser Lys His Pro Glu Trp Lys Asp
                725                 730                 735

Phe Gly Phe Arg Phe Ser Asp Thr Gln Arg Tyr Asn Ser Ile Asp Glu
            740                 745                 750

Phe Tyr Arg Glu Val Glu Asn Gln Gly Tyr Lys Leu Thr Phe Glu Asn
        755                 760                 765

Ile Ser Glu Ser Tyr Ile Asp Ser Val Val Asn Gln Gly Lys Leu Tyr
    770                 775                 780

Leu Phe Gln Ile Tyr Asn Lys Asp Phe Ser Ala Tyr Ser Lys Gly Arg
785                 790                 795                 800

Pro Asn Leu His Thr Leu Tyr Trp Lys Ala Leu Phe Asp Glu Arg Asn
                805                 810                 815

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

Arg Lys Gln Ser Ile Pro Lys Lys Ile Thr His Pro Ala Lys Glu Ala
        835                 840                 845

Ile Ala Asn Lys Asn Lys Asp Asn Pro Lys Lys Glu Ser Val Phe Glu
    850                 855                 860

Tyr Asp Leu Ile Lys Asp Lys Arg Phe Thr Glu Asp Lys Phe Phe Phe
865                 870                 875                 880

His Cys Pro Ile Thr Ile Asn Phe Lys Ser Ser Gly Ala Asn Lys Phe
                885                 890                 895
```

```
Asn Asp Glu Ile Asn Leu Leu Leu Lys Glu Lys Ala Asn Asp Val His
            900                 905                 910

Ile Leu Ser Ile Asp Arg Gly Glu Arg His Leu Ala Tyr Tyr Thr Leu
        915                 920                 925

Val Asp Gly Lys Gly Asn Ile Ile Lys Gln Asp Thr Phe Asn Ile Ile
    930                 935                 940

Gly Asn Asp Arg Met Lys Thr Asn Tyr His Asp Lys Leu Ala Ala Ile
945                 950                 955                 960

Glu Lys Asp Arg Asp Ser Ala Arg Lys Asp Trp Lys Lys Ile Asn Asn
                965                 970                 975

Ile Lys Glu Met Lys Glu Gly Tyr Leu Ser Gln Val Val His Glu Ile
            980                 985                 990

Ala Lys Leu Val Ile Glu Tyr Asn  Ala Ile Val Val Phe  Glu Asp Leu
        995                 1000                1005

Asn Phe  Gly Phe Lys Arg Gly  Arg Phe Lys Val Glu  Lys Gln Val
    1010                1015                1020

Tyr Gln  Lys Leu Glu Lys Met  Leu Ile Glu Lys Leu  Asn Tyr Leu
    1025                1030                1035

Val Phe  Lys Asp Asn Glu Phe  Asp Lys Thr Gly Gly  Val Leu Arg
    1040                1045                1050

Ala Tyr  Gln Leu Thr Ala Pro  Phe Glu Thr Phe Lys  Lys Met Gly
    1055                1060                1065

Lys Gln  Thr Gly Ile Ile Tyr  Tyr Val Pro Ala Gly  Phe Thr Ser
    1070                1075                1080

Lys Ile  Cys Pro Val Thr Gly  Phe Val Asn Gln Leu  Tyr Pro Lys
    1085                1090                1095

Tyr Glu  Ser Val Ser Lys Ser  Gln Glu Phe Phe Ser  Lys Phe Asp
```

```
    1100                1105                1110

Lys Ile  Cys Tyr Asn Leu Asp  Lys Gly Tyr Phe Glu  Phe Ser Phe
    1115                1120                1125

Asp Tyr  Lys Asn Phe Gly Asp  Lys Ala Ala Lys Gly  Lys Trp Thr
    1130                1135                1140

Ile Ala  Ser Phe Gly Ser Arg  Leu Ile Asn Phe Arg  Asn Ser Asp
    1145                1150                1155

Lys Asn  His Asn Trp Asp Thr  Arg Glu Val Tyr Pro  Thr Lys Glu
    1160                1165                1170

Leu Glu  Lys Leu Leu Lys Asp  Tyr Ser Ile Glu Tyr  Gly His Gly
    1175                1180                1185

Glu Cys  Ile Lys Ala Ala Ile  Cys Gly Glu Ser Asp  Lys Lys Phe
    1190                1195                1200

Phe Ala  Lys Leu Thr Ser Val  Leu Asn Thr Ile Leu  Gln Met Arg
    1205                1210                1215

Asn Ser  Lys Thr Gly Thr Glu  Leu Asp Tyr Leu Ile  Ser Pro Val
    1220                1225                1230

Ala Asp  Val Asn Gly Asn Phe  Phe Asp Ser Arg Gln  Ala Pro Lys
    1235                1240                1245

Asn Met  Pro Gln Asp Ala Asp  Ala Asn Gly Ala Tyr  His Ile Gly
    1250                1255                1260

Leu Lys  Gly Leu Met Leu Leu  Gly Arg Ile Lys Asn  Asn Gln Glu
    1265                1270                1275

Gly Lys  Lys Leu Asn Leu Val  Ile Lys Asn Glu Glu  Tyr Phe Glu
    1280                1285                1290

Phe Val  Gln Asn Arg Asn Asn
    1295                1300
```

<210> 7
<211> 1206
<212> PRT
<213> Unknown

<220>
<223> Lachnospiraceae bacterium

<400> 7

```
Met Tyr Tyr Glu Ser Leu Thr Lys Gln Tyr Pro Val Ser Lys Thr Ile
1               5                  10                  15


Arg Asn Glu Leu Ile Pro Ile Gly Lys Thr Leu Asp Asn Ile Arg Gln
            20                  25                  30


Asn Asn Ile Leu Glu Ser Asp Val Lys Arg Lys Gln Asn Tyr Glu His
        35                  40                  45


Val Lys Gly Ile Leu Asp Glu Tyr His Lys Gln Leu Ile Asn Glu Ala
    50                  55                  60


Leu Asp Asn Cys Thr Leu Pro Ser Leu Lys Ile Ala Ala Glu Ile Tyr
65                  70                  75                  80


Leu Lys Asn Gln Lys Glu Val Ser Asp Arg Glu Asp Phe Asn Lys Thr
                85                  90                  95


Gln Asp Leu Leu Arg Lys Glu Val Val Glu Lys Leu Lys Ala His Glu
            100                 105                 110


Asn Phe Thr Lys Ile Gly Lys Lys Asp Ile Leu Asp Leu Leu Glu Lys
        115                 120                 125


Leu Pro Ser Ile Ser Glu Asp Asp Tyr Asn Ala Leu Glu Ser Phe Arg
    130                 135                 140


Asn Phe Tyr Thr Tyr Phe Thr Ser Tyr Asn Lys Val Arg Glu Asn Leu
145                 150                 155                 160


Tyr Ser Asp Lys Glu Lys Ser Ser Thr Val Ala Tyr Arg Leu Ile Asn
                165                 170                 175
```

```
Glu Asn Phe Pro Lys Phe Leu Asp Asn Val Lys Ser Tyr Arg Phe Val
            180                 185                 190

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

Ser Leu Phe Ile Val Glu Thr Phe Asn Lys Thr Leu Thr Gln Asp Gly
    210                 215                 220

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

Leu Tyr Asn Gln Lys Asn Gln Lys Ala Asn Gly Phe Arg Lys Ile Pro
                245                 250                 255

Lys Met Lys Met Leu Tyr Lys Gln Ile Leu Ser Asp Arg Glu Glu Ser
            260                 265                 270

Phe Ile Asp Glu Phe Gln Ser Asp Glu Val Leu Ile Asp Asn Val Glu
        275                 280                 285

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

Ala Phe Phe Asp Ala Leu Arg Glu Ser Lys Gly Lys Asn Val Tyr Val
305                 310                 315                 320

Lys Asn Asp Leu Ala Lys Thr Ala Met Ser Val Ile Val Phe Glu Asn
                325                 330                 335

Trp Arg Thr Phe Asp Asp Leu Leu Asn Gln Glu Tyr Asp Leu Ala Asn
            340                 345                 350

Glu Asn Lys Lys Lys Asp Asp Lys Tyr Phe Glu Lys Arg Gln Lys Glu
        355                 360                 365

Leu Lys Lys Asn Lys Ser Tyr Ser Leu Glu His Leu Cys Asn Leu Ser
    370                 375                 380

Glu Asp Ser Cys Asn Leu Ile Glu Asn Tyr Ile His Gln Ile Ser Asp
```

```
385                 390                 395                 400

Asp Ile Glu Asn Ile Ile Ile Asn Asn Glu Thr Phe Leu Arg Ile Val
                405                 410                 415

Ile Asn Glu His Asp Arg Ser Arg Lys Leu Ala Lys Asn Arg Lys Ala
            420                 425                 430

Val Lys Ala Ile Lys Asp Phe Leu Asp Ser Ile Lys Val Leu Glu Arg
        435                 440                 445

Glu Leu Lys Leu Ile Asn Ser Ser Gly Gln Glu Leu Glu Lys Asp Leu
    450                 455                 460

Ile Val Tyr Ser Ala His Glu Glu Leu Leu Val Glu Leu Lys Gln Val
465                 470                 475                 480

Asp Ser Leu Tyr Asn Met Thr Arg Asn Tyr Leu Thr Lys Lys Pro Phe
                485                 490                 495

Ser Thr Glu Lys Val Lys Leu Asn Phe Asn Arg Ser Thr Leu Leu Asn
            500                 505                 510

Gly Trp Asp Arg Asn Lys Glu Thr Asp Asn Leu Gly Val Leu Leu Leu
        515                 520                 525

Lys Asp Gly Lys Tyr Tyr Leu Gly Ile Met Asn Thr Ser Ala Asn Lys
    530                 535                 540

Ala Phe Val Asn Pro Pro Val Ala Lys Thr Glu Lys Val Phe Lys Lys
545                 550                 555                 560

Val Asp Tyr Lys Leu Leu Pro Val Pro Asn Gln Met Leu Pro Lys Val
                565                 570                 575

Phe Phe Ala Lys Ser Asn Ile Asp Phe Tyr Asn Pro Ser Ser Glu Ile
            580                 585                 590

Tyr Ser Asn Tyr Lys Lys Gly Thr His Lys Lys Gly Asn Met Phe Ser
        595                 600                 605
```

```
Leu Glu Asp Cys His Asn Leu Ile Asp Phe Phe Lys Glu Ser Ile Ser
    610                 615                 620

Lys His Glu Asp Trp Ser Lys Phe Gly Phe Lys Phe Asp Thr Gln Ala
625                 630                 635                 640

Ser Tyr Asn Asp Ile Ser Glu Phe Tyr Arg Glu Val Glu Lys Gln Gly
                645                 650                 655

Tyr Lys Leu Thr Tyr Thr Asp Ile Asp Glu Thr Tyr Ile Asn Asp Leu
            660                 665                 670

Ile Glu Arg Asn Glu Leu Tyr Leu Phe Gln Ile Tyr Asn Lys Asp Phe
        675                 680                 685

Ser Met Tyr Ser Lys Gly Lys Leu Asn Leu His Thr Leu Tyr Phe Met
    690                 695                 700

Met Leu Phe Asp Gln Arg Asn Ile Asp Asp Val Val Tyr Lys Leu Asn
705                 710                 715                 720

Gly Glu Ala Glu Val Phe Tyr Arg Pro Ala Ser Ile Ser Glu Asp Glu
                725                 730                 735

Leu Ile Ile His Lys Ala Gly Glu Glu Ile Lys Asn Lys Asn Pro Asn
            740                 745                 750

Arg Ala Arg Thr Lys Glu Thr Ser Thr Phe Ser Tyr Asp Ile Val Lys
        755                 760                 765

Asp Lys Arg Tyr Ser Lys Asp Lys Phe Thr Leu His Ile Pro Ile Thr
    770                 775                 780

Met Asn Phe Gly Val Asp Glu Val Lys Arg Phe Asn Asp Ala Val Asn
785                 790                 795                 800

Ser Ala Ile Arg Ile Asp Glu Asn Val Asn Val Ile Gly Ile Asp Arg
                805                 810                 815

Gly Glu Arg Asn Leu Leu Tyr Val Val Val Ile Asp Ser Lys Gly Asn
```

```
            820                 825                 830

Ile Leu Glu Gln Ile Ser Leu Asn Ser Ile Ile Asn Lys Glu Tyr Asp
        835                 840                 845

Ile Glu Thr Asp Tyr His Ala Leu Leu Asp Glu Arg Glu Gly Gly Arg
    850                 855                 860

Asp Lys Ala Arg Lys Asp Trp Asn Thr Val Glu Asn Ile Arg Asp Leu
865                 870                 875                 880

Lys Ala Gly Leu Tyr Leu Gln Val Val Asn Val Val Ala Lys Leu Val
                885                 890                 895

Leu Lys Tyr Asn Ala Ile Ile Cys Leu Glu Asp Leu Asn Phe Gly Phe
            900                 905                 910

Lys Arg Gly Arg Gln Lys Val Glu Lys Gln Val Tyr Gln Lys Phe Glu
        915                 920                 925

Lys Met Leu Ile Asp Lys Leu Asn Tyr Leu Val Ile Asp Lys Ser Arg
    930                 935                 940

Glu Gln Thr Ser Pro Lys Glu Leu Gly Gly Ala Leu Asn Ala Leu Gln
945                 950                 955                 960

Leu Thr Ser Lys Phe Lys Ser Phe Lys Glu Leu Gly Lys Gln Ser Gly
                965                 970                 975

Val Ile Tyr Tyr Val Pro Ala Tyr Leu Thr Ser Lys Ile Asp Pro Thr
            980                 985                 990

Thr Gly Phe Ala Asn Leu Phe Tyr  Met Lys Cys Glu Asn  Val Glu Lys
        995                 1000                1005

Ser Lys  Arg Phe Phe Asp Gly  Phe Asp Phe Ile Arg  Phe Asn Ala
    1010                1015                1020

Leu Glu  Asn Val Phe Glu Phe  Gly Phe Asp Tyr Arg  Ser Phe Thr
    1025                1030                1035
```

```
Gln Arg  Ala Cys Gly Ile Asn  Ser Lys Trp Thr Val  Cys Thr Asn
    1040                1045                1050

Gly Glu  Arg Ile Ile Lys Tyr  Arg Asn Pro Asp Lys  Asn Asn Met
    1055                1060                1065

Phe Asp  Glu Lys Val Val Val  Val Thr Asp Glu Met  Lys Asn Leu
    1070                1075                1080

Phe Glu  Gln Tyr Lys Ile Pro  Tyr Glu Asp Gly Arg  Asn Val Lys
    1085                1090                1095

Asp Met  Ile Ile Ser Asn Glu  Glu Ala Glu Phe Tyr  Arg Arg Leu
    1100                1105                1110

Tyr Arg  Leu Leu Gln Gln Thr  Leu Gln Met Arg Asn  Ser Thr Ser
    1115                1120                1125

Asp Gly  Thr Arg Asp Tyr Ile  Ile Ser Pro Val Lys  Asn Lys Arg
    1130                1135                1140

Glu Ala  Tyr Phe Asn Ser Glu  Leu Ser Asp Gly Ser  Val Pro Lys
    1145                1150                1155

Asp Ala  Asp Ala Asn Gly Ala  Tyr Asn Ile Ala Arg  Lys Gly Leu
    1160                1165                1170

Trp Val  Leu Glu Gln Ile Arg  Gln Lys Ser Glu Gly  Glu Lys Ile
    1175                1180                1185

Asn Leu  Ala Met Thr Asn Ala  Glu Trp Leu Glu Tyr  Ala Gln Thr
    1190                1195                1200

His Leu  Leu
    1205
```

<210> 8
<211> 1233
<212> PRT
<213> Unknown

<220>
<223> Lachnospiraceae bacterium

<400> 8

```
Met Asp Tyr Gly Asn Gly Gln Phe Glu Arg Arg Ala Pro Leu Thr Lys
1               5                   10                  15


Thr Ile Thr Leu Arg Leu Lys Pro Ile Gly Glu Thr Arg Glu Thr Ile
            20                  25                  30


Arg Glu Gln Lys Leu Leu Glu Gln Asp Ala Ala Phe Arg Lys Leu Val
        35                  40                  45


Glu Thr Val Thr Pro Ile Val Asp Asp Cys Ile Arg Lys Ile Ala Asp
    50                  55                  60


Asn Ala Leu Cys His Phe Gly Thr Glu Tyr Asp Phe Ser Cys Leu Gly
65                  70                  75                  80


Asn Ala Ile Ser Lys Asn Asp Ser Lys Ala Ile Lys Lys Glu Thr Glu
                85                  90                  95


Lys Val Glu Lys Leu Leu Ala Lys Val Leu Thr Glu Asn Leu Pro Asp
            100                 105                 110


Gly Leu Arg Lys Val Asn Asp Ile Asn Ser Ala Ala Phe Ile Gln Asp
        115                 120                 125


Thr Leu Thr Ser Phe Val Gln Asp Asp Ala Asp Lys Arg Val Leu Ile
    130                 135                 140


Gln Glu Leu Lys Gly Lys Thr Val Leu Met Gln Arg Phe Leu Thr Thr
145                 150                 155                 160


Arg Ile Thr Ala Leu Thr Val Trp Leu Pro Asp Arg Val Phe Glu Asn
                165                 170                 175


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


Ser Pro Leu Asn Glu Lys Ile Met Lys Phe Asp Pro Asp Ala Glu Gln
```

```
        195                 200                 205

Tyr Ala Ser Leu Glu Phe Tyr Gly Gln Cys Leu Ser Gln Lys Asp Ile
    210                 215                 220

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

Lys Asn Pro Gly Ile Asn Glu Ile Val Lys Glu Tyr Asn Gln Gln Ile
                245                 250                 255

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

Lys Gln Ile Leu Met Pro Val Glu Lys Ala Phe Phe Val Arg Val Leu
        275                 280                 285

Ser Asn Asp Ser Asp Ala Arg Ser Ile Leu Glu Lys Ile Leu Lys Asp
    290                 295                 300

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

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

Val Ile Tyr Gly Asp His Gly Lys Leu Ser Gln Ile Ile Tyr Asp Lys
            340                 345                 350

Glu Ser Lys Arg Ile Ser Glu Leu Met Glu Thr Leu Ser Pro Lys Glu
        355                 360                 365

Arg Lys Glu Ser Lys Lys Arg Leu Glu Gly Leu Glu Glu His Ile Arg
    370                 375                 380

Lys Ser Thr Tyr Thr Phe Asp Glu Leu Asn Arg Tyr Ala Glu Lys Asn
385                 390                 395                 400

Val Met Ala Ala Tyr Ile Ala Ala Val Glu Glu Ser Cys Ala Glu Ile
                405                 410                 415
```

```
Met Arg Lys Glu Lys Asp Leu Arg Thr Leu Leu Ser Lys Glu Asp Val
            420                 425                 430

Lys Ile Arg Gly Asn Arg His Asn Thr Leu Ile Val Lys Asn Tyr Phe
        435                 440                 445

Asn Ala Trp Thr Val Phe Arg Asn Leu Ile Arg Ile Leu Arg Arg Lys
    450                 455                 460

Ser Glu Ala Glu Ile Asp Ser Asp Phe Tyr Asp Val Leu Asp Asp Ser
465                 470                 475                 480

Val Glu Val Leu Ser Leu Thr Tyr Lys Gly Glu Asn Leu Cys Arg Ser
                485                 490                 495

Tyr Ile Thr Lys Lys Ile Gly Ser Asp Leu Lys Pro Glu Ile Ala Thr
            500                 505                 510

Tyr Gly Ser Ala Leu Arg Pro Asn Ser Arg Trp Trp Ser Pro Gly Glu
        515                 520                 525

Lys Phe Asn Val Lys Phe His Thr Ile Val Arg Arg Asp Gly Arg Leu
    530                 535                 540

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

Met Asp Gly Asp Ile Glu Cys Leu Gln Met Arg Lys Ile Pro Asn Pro
                565                 570                 575

Thr Ile Phe Leu Pro Lys Leu Val Phe Lys Asp Pro Glu Ala Phe Phe
            580                 585                 590

Arg Asp Asn Pro Glu Ala Asp Glu Phe Val Phe Leu Ser Gly Met Lys
        595                 600                 605

Ala Pro Val Thr Ile Thr Arg Glu Thr Tyr Glu Ala Tyr Arg Tyr Lys
    610                 615                 620

Leu Tyr Thr Val Gly Lys Leu Arg Asp Gly Glu Val Ser Glu Glu Glu
```

625 630 635 640

Tyr Lys Arg Ala Leu Leu Gln Val Leu Thr Ala Tyr Lys Glu Phe Leu
645 650 655

Glu Asn Arg Met Ile Tyr Ala Asp Leu Asn Phe Gly Phe Lys Asp Leu
660 665 670

Glu Glu Tyr Lys Asp Ser Ser Glu Phe Ile Lys Gln Val Glu Thr His
675 680 685

Asn Thr Phe Met Cys Trp Ala Lys Val Ser Ser Ser Gln Leu Asp Asp
690 695 700

Leu Val Lys Ser Gly Asn Gly Leu Leu Phe Glu Ile Trp Ser Glu Arg
705 710 715 720

Leu Glu Ser Tyr Tyr Lys Tyr Gly Asn Glu Lys Val Leu Arg Gly Tyr
725 730 735

Glu Gly Val Leu Leu Ser Ile Leu Lys Asp Glu Asn Leu Val Ser Met
740 745 750

Arg Thr Leu Leu Asn Ser Arg Pro Met Leu Val Tyr Arg Pro Lys Glu
755 760 765

Ser Ser Lys Pro Met Val Val His Arg Asp Gly Ser Arg Val Val Asp
770 775 780

Arg Phe Asp Lys Asp Gly Lys Tyr Ile Pro Pro Glu Val His Asp Glu
785 790 795 800

Leu Tyr Arg Phe Phe Asn Asn Leu Leu Ile Lys Glu Lys Leu Gly Glu
805 810 815

Lys Ala Arg Lys Ile Leu Asp Asn Lys Lys Val Lys Val Lys Val Leu
820 825 830

Glu Ser Glu Arg Val Lys Trp Ser Lys Phe Tyr Asp Glu Gln Phe Ala
835 840 845

```
Val Thr Phe Ser Val Lys Lys Asn Ala Asp Cys Leu Asp Thr Thr Lys
    850                 855                 860

Asp Leu Asn Ala Glu Val Met Glu Gln Tyr Ser Glu Ser Asn Arg Leu
865                 870                 875                 880

Ile Leu Ile Arg Asn Thr Thr Asp Ile Leu Tyr Tyr Leu Val Leu Asp
                885                 890                 895

Lys Asn Gly Lys Val Leu Lys Gln Arg Ser Leu Asn Ile Ile Asn Asp
            900                 905                 910

Gly Ala Arg Asp Val Asp Trp Lys Glu Arg Phe Arg Gln Val Thr Lys
        915                 920                 925

Asp Arg Asn Glu Gly Tyr Asn Glu Trp Asp Tyr Ser Arg Thr Ser Asn
    930                 935                 940

Asp Leu Lys Glu Val Tyr Leu Asn Tyr Ala Leu Lys Glu Ile Ala Glu
945                 950                 955                 960

Ala Val Ile Glu Tyr Asn Ala Ile Leu Ile Ile Glu Lys Met Ser Asn
                965                 970                 975

Ala Phe Lys Asp Lys Tyr Ser Phe Leu Asp Asp Val Thr Phe Lys Gly
            980                 985                 990

Phe Glu Thr Lys Lys Leu Ala Lys  Leu Ser Asp Leu His  Phe Arg Gly
        995                 1000                1005

Ile Lys  Asp Gly Glu Pro Cys  Ser Phe Thr Asn Pro  Leu Gln Leu
    1010                1015                1020

Cys Gln  Asn Asp Ser Asn Lys  Ile Leu Gln Asp Gly  Val Ile Phe
    1025                1030                1035

Met Val  Pro Asn Ser Met Thr  Arg Ser Leu Asp Pro  Asp Thr Gly
    1040                1045                1050

Phe Ile  Phe Ala Ile Asn Asp  His Asn Ile Arg Thr  Lys Lys Ala
```

```
    1055                1060                1065

Lys Leu  Asn Phe Leu Ser Lys  Phe Asp Gln Leu Lys  Val Ser Ser
    1070                1075                1080

Glu Gly  Cys Leu Ile Met Lys  Tyr Ser Gly Asp Ser  Leu Pro Thr
    1085                1090                1095

His Asn  Thr Asp Asn Arg Val  Trp Asn Cys Cys Cys  Asn His Pro
    1100                1105                1110

Ile Thr  Asn Tyr Asp Arg Glu  Thr Lys Lys Val Glu  Phe Ile Glu
    1115                1120                1125

Glu Pro  Val Glu Glu Leu Ser  Arg Val Leu Glu Glu  Asn Gly Ile
    1130                1135                1140

Glu Thr  Asp Thr Glu Leu Asn  Lys Leu Asn Glu Arg  Glu Asn Val
    1145                1150                1155

Pro Gly  Lys Val Val Asp Ala  Ile Tyr Ser Leu Val  Leu Asn Tyr
    1160                1165                1170

Leu Arg  Gly Thr Val Ser Gly  Val Ala Gly Gln Arg  Ala Val Tyr
    1175                1180                1185

Tyr Ser  Pro Val Thr Gly Lys  Lys Tyr Asp Ile Ser  Phe Ile Gln
    1190                1195                1200

Ala Met  Asn Leu Asn Arg Lys  Cys Asp Tyr Tyr Arg  Ile Gly Ser
    1205                1210                1215

Lys Glu  Arg Gly Glu Trp Thr  Asp Phe Val Ala Gln  Leu Ile Asn
    1220                1225                1230
```

<210> 9
<211> 1227
<212> PRT
<213> Unknown

<220>
<223> Lachnospiraceae bacterium

<400> 9

```
Met Ser Lys Leu Glu Lys Phe Thr Asn Cys Tyr Ser Leu Ser Lys Thr
1               5                  10                  15

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

Asn Lys Arg Leu Leu Val Glu Asp Glu Lys Arg Ala Glu Asp Tyr Lys
        35                  40                  45

Gly Val Lys Lys Leu Leu Asp Arg Tyr Tyr Leu Ser Phe Ile Asn Asp
    50                  55                  60

Val Leu His Ser Ile Lys Leu Lys Asn Leu Asn Asn Tyr Ile Ser Leu
65                  70                  75                  80

Phe Arg Lys Lys Thr Arg Thr Glu Lys Glu Asn Lys Glu Leu Glu Asn
                85                  90                  95

Leu Glu Ile Asn Leu Arg Lys Glu Ile Ala Lys Ala Phe Lys Gly Asn
            100                 105                 110

Glu Gly Tyr Lys Ser Leu Phe Lys Lys Asp Ile Ile Glu Thr Ile Leu
        115                 120                 125

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

Asn Gly Phe Thr Thr Ala Phe Thr Gly Phe Phe Asp Asn Arg Glu Asn
145                 150                 155                 160

Met Phe Ser Glu Glu Ala Lys Ser Thr Ser Ile Ala Phe Arg Cys Ile
                165                 170                 175

Asn Glu Asn Leu Thr Arg Tyr Ile Ser Asn Met Asp Ile Phe Glu Lys
            180                 185                 190

Val Asp Ala Ile Phe Asp Lys His Glu Val Gln Glu Ile Lys Glu Lys
        195                 200                 205
```

```
Ile Leu Asn Ser Asp Tyr Asp Val Glu Asp Phe Phe Glu Gly Glu Phe
    210                 215                 220

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

Ile Gly Gly Phe Val Thr Glu Ser Gly Glu Lys Ile Lys Gly Leu Asn
                245                 250                 255

Glu Tyr Ile Asn Leu Tyr Asn Gln Lys Thr Lys Gln Lys Leu Pro Lys
            260                 265                 270

Phe Lys Pro Leu Tyr Lys Gln Val Leu Ser Asp Arg Glu Ser Leu Ser
        275                 280                 285

Phe Tyr Gly Glu Gly Tyr Thr Ser Asp Glu Glu Val Leu Glu Val Phe
    290                 295                 300

Arg Asn Thr Leu Asn Lys Asn Ser Glu Ile Phe Ser Ser Ile Lys Lys
305                 310                 315                 320

Leu Glu Lys Leu Phe Lys Asn Phe Asp Glu Tyr Ser Ser Ala Gly Ile
                325                 330                 335

Phe Val Lys Asn Gly Pro Ala Ile Ser Thr Ile Ser Lys Asp Ile Phe
            340                 345                 350

Gly Glu Trp Asn Val Ile Arg Asp Lys Trp Asn Ala Glu Tyr Asp Asp
        355                 360                 365

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

Arg Arg Lys Ser Phe Lys Lys Ile Gly Ser Phe Ser Leu Glu Gln Leu
385                 390                 395                 400

Gln Glu Tyr Ala Asp Ala Asp Leu Ser Val Val Glu Lys Leu Lys Glu
                405                 410                 415

Ile Ile Ile Gln Lys Val Asp Glu Ile Tyr Lys Val Tyr Gly Ser Ser
```

```
            420                 425                 430

Glu Lys Leu Phe Asp Ala Asp Phe Val Leu Glu Lys Ser Leu Lys Lys
        435                 440                 445

Asn Asp Ala Val Val Ala Ile Met Lys Asp Leu Leu Asp Ser Val Lys
    450                 455                 460

Ser Phe Glu Asn Tyr Ile Lys Ala Phe Phe Gly Glu Gly Lys Glu Thr
465                 470                 475                 480

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

Leu Leu Lys Val Asp His Ile Tyr Asp Ala Ile Arg Asn Tyr Val Thr
            500                 505                 510

Gln Lys Pro Tyr Ser Lys Asp Lys Phe Lys Leu Tyr Phe Gln Asn Pro
        515                 520                 525

Gln Phe Met Gly Gly Trp Asp Lys Asp Lys Glu Thr Asp Tyr Arg Ala
    530                 535                 540

Thr Ile Leu Arg Tyr Gly Ser Lys Tyr Tyr Leu Ala Ile Met Asp Lys
545                 550                 555                 560

Lys Tyr Ala Lys Cys Leu Gln Lys Ile Asp Lys Asp Asp Val Asn Gly
                565                 570                 575

Asn Tyr Glu Lys Ile Asn Tyr Lys Leu Leu Pro Gly Pro Asn Lys Met
            580                 585                 590

Leu Pro Lys Val Phe Phe Ser Lys Lys Trp Met Ala Tyr Tyr Asn Pro
        595                 600                 605

Ser Glu Asp Ile Gln Lys Ile Tyr Lys Asn Gly Thr Phe Lys Lys Gly
    610                 615                 620

Asp Met Phe Asn Leu Asn Asp Cys His Lys Leu Ile Asp Phe Phe Lys
625                 630                 635                 640
```

```
Asp Ser Ile Ser Arg Tyr Pro Lys Trp Ser Asn Ala Tyr Asp Phe Asn
                645                 650                 655

Phe Ser Glu Thr Glu Lys Tyr Lys Asp Ile Ala Gly Phe Tyr Arg Glu
            660                 665                 670

Val Glu Glu Gln Gly Tyr Lys Val Ser Phe Glu Ser Ala Ser Lys Lys
        675                 680                 685

Glu Val Asp Lys Leu Val Glu Glu Gly Lys Leu Tyr Met Phe Gln Ile
    690                 695                 700

Tyr Asn Lys Asp Phe Ser Asp Lys Ser His Gly Thr Pro Asn Leu His
705                 710                 715                 720

Thr Met Tyr Phe Lys Leu Leu Phe Asp Glu Asn Asn His Gly Gln Ile
                725                 730                 735

Arg Leu Ser Gly Gly Ala Glu Leu Phe Met Arg Arg Ala Ser Leu Lys
            740                 745                 750

Lys Glu Glu Leu Val Val His Pro Ala Asn Ser Pro Ile Ala Asn Lys
        755                 760                 765

Asn Pro Asp Asn Pro Lys Lys Thr Thr Thr Leu Ser Tyr Asp Val Tyr
    770                 775                 780

Lys Asp Lys Arg Phe Ser Glu Asp Gln Tyr Glu Leu His Ile Pro Ile
785                 790                 795                 800

Ala Asn Ile Asn Lys Cys Pro Lys Asn Ile Phe Lys Ile Asn Thr Glu
                805                 810                 815

Val Arg Val Leu Leu Lys His Asp Asp Asn Pro Tyr Val Ile Gly Ile
            820                 825                 830

Asp Arg Gly Glu Arg Asn Leu Leu Tyr Ile Val Val Val Asp Gly Lys
        835                 840                 845

Gly Asn Ile Val Glu Gln Tyr Ser Leu Asn Glu Ile Ile Asn Asn Phe
```

```
    850                 855                 860

Asn Gly Ile Arg Ile Lys Thr Asp Tyr His Ser Leu Leu Asp Lys Lys
865                 870                 875                 880

Glu Lys Glu Arg Phe Glu Ala Arg Gln Asn Trp Thr Ser Ile Glu Asn
                885                 890                 895

Ile Lys Glu Leu Lys Ala Gly Tyr Ile Ser Gln Val Val His Lys Ile
            900                 905                 910

Cys Glu Leu Val Glu Lys Tyr Asp Ala Val Ile Ala Leu Glu Asp Leu
        915                 920                 925

Asn Ser Gly Phe Lys Asn Ser Arg Val Lys Val Glu Lys Gln Val Tyr
    930                 935                 940

Gln Lys Phe Glu Lys Met Leu Ile Asp Lys Leu Asn Tyr Met Val Asp
945                 950                 955                 960

Lys Lys Ser Asn Pro Cys Ala Thr Gly Gly Ala Leu Lys Gly Tyr Gln
                965                 970                 975

Ile Thr Asn Lys Phe Glu Ser Phe Lys Ser Met Ser Thr Gln Asn Gly
            980                 985                 990

Phe Ile Phe Tyr Ile Pro Ala Trp  Leu Thr Ser Lys Ile  Asp Pro Ser
        995                 1000                1005

Thr Gly  Phe Val Asn Leu Leu  Lys Thr Lys Tyr Thr  Ser Ile Ala
    1010                1015                1020

Asp Lys  Lys Phe Ile Ser Ser  Phe Asp Arg Ile Met  Tyr Val Pro
    1025                1030                1035

Glu Glu  Asp Leu Phe Glu Phe  Ala Leu Asp Tyr Lys  Asn Phe Ser
    1040                1045                1050

Arg Thr  Asp Ala Asp Tyr Ile  Lys Lys Trp Lys Leu  Tyr Ser Tyr
    1055                1060                1065
```

```
Gly Asn  Arg Ile Arg Ile Phe  Arg Asn Pro Lys Lys  Asn Asn Val
    1070                1075                1080


Phe Asp  Trp Glu Glu Val Cys  Leu Thr Ser Ala Tyr  Lys Glu Leu
    1085                1090                1095


Phe Asn  Lys Tyr Gly Ile Asn  Tyr Gln Gln Gly Asp  Ile Arg Ala
    1100                1105                1110


Leu Leu  Cys Glu Gln Ser Asp  Lys Ala Phe Tyr Ser  Ser Phe Met
    1115                1120                1125


Ala Leu  Met Ser Leu Met Leu  Gln Met Arg Asn Ser  Ile Thr Gly
    1130                1135                1140


Arg Thr  Asp Val Asp Phe Leu  Ile Ser Pro Val Lys  Asn Ser Asp
    1145                1150                1155


Gly Ile  Phe Tyr Asp Ser Arg  Asn Tyr Glu Ala Gln  Glu Asn Ala
    1160                1165                1170


Ile Leu  Pro Lys Asn Ala Asp  Ala Asn Gly Ala Tyr  Asn Ile Ala
    1175                1180                1185


Arg Lys  Val Leu Trp Ala Ile  Gly Gln Phe Lys Lys  Ala Glu Asp
    1190                1195                1200


Glu Lys  Leu Asp Lys Val Lys  Ile Ala Ser Asn Lys  Glu Trp Leu
    1205                1210                1215


Glu Tyr  Ala Gln Thr Ser Val  Lys His
    1220                1225


<210>  10
<211>  1264
<212>  PRT
<213>  Leptospira inadai

<400>  10

Met Glu Asp Tyr Ser Gly Phe Val Asn Ile Tyr Ser Ile Gln Lys Thr
1               5                  10                  15
```

```
Leu Arg Phe Glu Leu Lys Pro Val Gly Lys Thr Leu Glu His Ile Glu
            20                  25                  30

Lys Lys Gly Phe Leu Lys Lys Asp Lys Ile Arg Ala Glu Asp Tyr Lys
        35                  40                  45

Ala Val Lys Lys Ile Ile Asp Lys Tyr His Arg Ala Tyr Ile Glu Glu
    50                  55                  60

Val Phe Asp Ser Val Leu His Gln Lys Lys Lys Lys Asp Lys Thr Arg
65                  70                  75                  80

Phe Ser Thr Gln Phe Ile Lys Glu Ile Lys Glu Phe Ser Glu Leu Tyr
                85                  90                  95

Tyr Lys Thr Glu Lys Asn Ile Pro Asp Lys Glu Arg Leu Glu Ala Leu
            100                 105                 110

Ser Glu Lys Leu Arg Lys Met Leu Val Gly Ala Phe Lys Gly Glu Phe
        115                 120                 125

Ser Glu Glu Val Ala Glu Lys Tyr Asn Lys Asn Leu Phe Ser Lys Glu
    130                 135                 140

Leu Ile Arg Asn Glu Ile Glu Lys Phe Cys Glu Thr Asp Glu Glu Arg
145                 150                 155                 160

Lys Gln Val Ser Asn Phe Lys Ser Phe Thr Thr Tyr Phe Thr Gly Phe
                165                 170                 175

His Ser Asn Arg Gln Asn Ile Tyr Ser Asp Glu Lys Lys Ser Thr Ala
            180                 185                 190

Ile Gly Tyr Arg Ile Ile His Gln Asn Leu Pro Lys Phe Leu Asp Asn
        195                 200                 205

Leu Lys Ile Ile Glu Ser Ile Gln Arg Arg Phe Lys Asp Phe Pro Trp
    210                 215                 220
```

```
Ser Asp Leu Lys Lys Asn Leu Lys Lys Ile Asp Lys Asn Ile Lys Leu
225                 230                 235                 240

Thr Glu Tyr Phe Ser Ile Asp Gly Phe Val Asn Val Leu Asn Gln Lys
                245                 250                 255

Gly Ile Asp Ala Tyr Asn Thr Ile Leu Gly Gly Lys Ser Glu Glu Ser
            260                 265                 270

Gly Glu Lys Ile Gln Gly Leu Asn Glu Tyr Ile Asn Leu Tyr Arg Gln
        275                 280                 285

Lys Asn Asn Ile Asp Arg Lys Asn Pro Leu Asn Val Lys Ile Leu Phe
    290                 295                 300

Lys Gln Ile Leu Gly Asp Arg Glu Thr Lys Ser Phe Ile Pro Glu Ala
305                 310                 315                 320

Phe Pro Asp Asp Gln Ser Val Leu Asn Ser Ile Thr Glu Phe Ala Lys
                325                 330                 335

Tyr Leu Lys Leu Asp Lys Lys Lys Lys Ser Ile Ile Ala Glu Leu Lys
            340                 345                 350

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

Ala Asn Asp Asn Ser Leu Ala Ser Ile Ser Thr Phe Leu Phe Asp Asp
    370                 375                 380

Trp Ser Phe Ile Lys Lys Ser Val Ser Phe Lys Tyr Asp Glu Ser Val
385                 390                 395                 400

Gly Asp Pro Lys Lys Lys Ile Lys Ser Pro Leu Lys Tyr Glu Lys Glu
                405                 410                 415

Lys Glu Lys Trp Leu Lys Gln Lys Tyr Tyr Thr Ile Ser Phe Leu Asn
            420                 425                 430

Asp Ala Ile Glu Ser Tyr Ser Lys Ser Gln Asp Glu Lys Arg Val Lys
        435                 440                 445
```

```
Ile Arg Leu Glu Ala Tyr Phe Ala Glu Phe Lys Ser Lys Asp Asp Ala
    450                 455                 460

Lys Lys Gln Phe Asp Leu Leu Glu Arg Ile Glu Glu Ala Tyr Ala Ile
465                 470                 475                 480

Val Glu Pro Leu Leu Gly Ala Glu Tyr Pro Arg Asp Arg Asn Leu Lys
                485                 490                 495

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

Lys Ser Leu Gln Phe Phe Leu Lys Pro Leu Leu Ser Ala Glu Ile Phe
        515                 520                 525

Asp Glu Lys Asp Leu Gly Phe Tyr Asn Gln Leu Glu Gly Tyr Tyr Glu
    530                 535                 540

Glu Ile Asp Ile Ser Gly His Leu Tyr Asn Lys Val Arg Asn Tyr Leu
545                 550                 555                 560

Thr Gly Lys Ile Tyr Ser Lys Glu Lys Phe Lys Leu Asn Phe Glu Asn
                565                 570                 575

Ser Thr Leu Leu Lys Gly Trp Asp Glu Asn Arg Glu Val Ala Asn Leu
            580                 585                 590

Cys Val Ile Phe Arg Glu Asp Gln Lys Tyr Tyr Leu Gly Val Met Asp
        595                 600                 605

Lys Glu Asn Asn Thr Ile Leu Ser Asp Ile Pro Lys Val Lys Pro Asn
    610                 615                 620

Glu Leu Phe Tyr Glu Lys Met Val Tyr Lys Leu Ile Pro Thr Pro His
625                 630                 635                 640

Met Gln Leu Pro Arg Ile Ile Phe Ser Ser Asp Asn Leu Ser Ile Tyr
                645                 650                 655
```

```
Asn Pro Ser Lys Ser Ile Leu Lys Ile Arg Glu Ala Lys Ser Phe Lys
            660                 665                 670

Glu Gly Lys Asn Phe Lys Leu Lys Asp Cys His Lys Phe Ile Asp Phe
        675                 680                 685

Tyr Lys Glu Ser Ile Ser Lys Asn Glu Asp Trp Ser Arg Phe Asp Phe
    690                 695                 700

Lys Phe Ser Lys Thr Ser Ser Tyr Glu Asn Ile Ser Glu Phe Tyr Arg
705                 710                 715                 720

Glu Val Glu Arg Gln Gly Tyr Asn Leu Asp Phe Lys Lys Val Ser Lys
                725                 730                 735

Phe Tyr Ile Asp Ser Leu Val Glu Asp Gly Lys Leu Tyr Leu Phe Gln
            740                 745                 750

Ile Tyr Asn Lys Asp Phe Ser Ile Phe Ser Lys Gly Lys Pro Asn Leu
        755                 760                 765

His Thr Ile Tyr Phe Arg Ser Leu Phe Ser Lys Glu Asn Leu Lys Asp
    770                 775                 780

Val Cys Leu Lys Leu Asn Gly Glu Ala Glu Met Phe Phe Arg Lys Lys
785                 790                 795                 800

Ser Ile Asn Tyr Asp Glu Lys Lys Lys Arg Glu Gly His His Pro Glu
                805                 810                 815

Leu Phe Glu Lys Leu Lys Tyr Pro Ile Leu Lys Asp Lys Arg Tyr Ser
            820                 825                 830

Glu Asp Lys Phe Gln Phe His Leu Pro Ile Ser Leu Asn Phe Lys Ser
        835                 840                 845

Lys Glu Arg Leu Asn Phe Asn Leu Lys Val Asn Glu Phe Leu Lys Arg
    850                 855                 860

Asn Lys Asp Ile Asn Ile Ile Gly Ile Asp Arg Gly Glu Arg Asn Leu
865                 870                 875                 880
```

```
Leu Tyr Leu Val Met Ile Asn Gln Lys Gly Glu Ile Leu Lys Gln Thr
               885                 890                 895

Leu Leu Asp Ser Met Gln Ser Gly Lys Gly Arg Pro Glu Ile Asn Tyr
           900                 905                 910

Lys Glu Lys Leu Gln Glu Lys Glu Ile Glu Arg Asp Lys Ala Arg Lys
        915                 920                 925

Ser Trp Gly Thr Val Glu Asn Ile Lys Glu Leu Lys Glu Gly Tyr Leu
    930                 935                 940

Ser Ile Val Ile His Gln Ile Ser Lys Leu Met Val Glu Asn Asn Ala
945                 950                 955                 960

Ile Val Val Leu Glu Asp Leu Asn Ile Gly Phe Lys Arg Gly Arg Gln
               965                 970                 975

Lys Val Glu Arg Gln Val Tyr Gln Lys Phe Glu Lys Met Leu Ile Asp
           980                 985                 990

Lys Leu Asn Phe Leu Val Phe Lys  Glu Asn Lys Pro Thr  Glu Pro Gly
        995                 1000                1005

Gly Val  Leu Lys Ala Tyr Gln  Leu Thr Asp Glu Phe  Gln Ser Phe
    1010                1015                1020

Glu Lys  Leu Ser Lys Gln Thr  Gly Phe Leu Phe Tyr  Val Pro Ser
    1025                1030                1035

Trp Asn  Thr Ser Lys Ile Asp  Pro Arg Thr Gly Phe  Ile Asp Phe
    1040                1045                1050

Leu His  Pro Ala Tyr Glu Asn  Ile Glu Lys Ala Lys  Gln Trp Ile
    1055                1060                1065

Asn Lys  Phe Asp Ser Ile Arg  Phe Asn Ser Lys Met  Asp Trp Phe
    1070                1075                1080
```

```
Glu Phe  Thr Ala Asp Thr Arg  Lys Phe Ser Glu Asn  Leu Met Leu
    1085                 1090                 1095

Gly Lys  Asn Arg Val Trp Val  Ile Cys Thr Thr Asn  Val Glu Arg
    1100                 1105                 1110

Tyr Phe  Thr Ser Lys Thr Ala  Asn Ser Ser Ile Gln  Tyr Asn Ser
    1115                 1120                 1125

Ile Gln  Ile Thr Glu Lys Leu  Lys Glu Leu Phe Val  Asp Ile Pro
    1130                 1135                 1140

Phe Ser  Asn Gly Gln Asp Leu  Lys Pro Glu Ile Leu  Arg Lys Asn
    1145                 1150                 1155

Asp Ala  Val Phe Phe Lys Ser  Leu Leu Phe Tyr Ile  Lys Thr Thr
    1160                 1165                 1170

Leu Ser  Leu Arg Gln Asn Asn  Gly Lys Lys Gly Glu  Glu Glu Lys
    1175                 1180                 1185

Asp Phe  Ile Leu Ser Pro Val  Val Asp Ser Lys Gly  Arg Phe Phe
    1190                 1195                 1200

Asn Ser  Leu Glu Ala Ser Asp  Asp Glu Pro Lys Asp  Ala Asp Ala
    1205                 1210                 1215

Asn Gly  Ala Tyr His Ile Ala  Leu Lys Gly Leu Met  Asn Leu Leu
    1220                 1225                 1230

Val Leu  Asn Glu Thr Lys Glu  Glu Asn Leu Ser Arg  Pro Lys Trp
    1235                 1240                 1245

Lys Ile  Lys Asn Lys Asp Trp  Leu Glu Phe Val Trp  Glu Arg Asn
    1250                 1255                 1260

Arg
```

<210> 11
<211> 1373

<212> PRT
<213> Moraxella bovoculi

<400> 11

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


Arg Phe Glu Leu Phe Ile Asp Arg Thr Leu Glu His Ile His Ala Lys
            20                  25                  30


Asn Phe Leu Ser Gln Asp Glu Thr Met Ala Asp Met His Gln Lys Val
        35                  40                  45


Lys Val Ile Leu Asp Asp Tyr His Arg Asp Phe Ile Ala Asp Met Met
    50                  55                  60


Gly Glu Val Lys Leu Thr Lys Leu Ala Glu Phe Tyr Asp Val Tyr Leu
65                  70                  75                  80


Lys Phe Arg Lys Asn Pro Lys Asp Asp Glu Leu Gln Lys Ala Gln Leu
                85                  90                  95


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


Asn Gly Gly Lys Tyr Lys Ala Gly Tyr Asp Arg Leu Phe Gly Ala Lys
        115                 120                 125


Leu Phe Lys Asp Gly Lys Glu Leu Gly Asp Leu Ala Lys Phe Val Ile
    130                 135                 140


Ala Gln Glu Gly Glu Ser Ser Pro Lys Leu Ala His Leu Ala His Phe
145                 150                 155                 160


Glu Lys Phe Ser Thr Tyr Phe Thr Gly Phe His Asp Asn Arg Lys Asn
                165                 170                 175


Met Tyr Ser Asp Glu Asp Lys His Thr Ala Ile Ala Tyr Arg Leu Ile
            180                 185                 190


His Glu Asn Leu Pro Arg Phe Ile Asp Asn Leu Gln Ile Leu Thr Thr
```

```
        195                 200                 205

Ile Lys Gln Lys His Ser Ala Leu Tyr Asp Gln Ile Ile Asn Glu Leu
    210                 215                 220

Thr Ala Ser Gly Leu Asp Val Ser Leu Ala Ser His Leu Asp Gly Tyr
225                 230                 235                 240

His Lys Leu Leu Thr Gln Glu Gly Ile Thr Ala Tyr Asn Thr Leu Leu
                245                 250                 255

Gly Gly Ile Ser Gly Glu Ala Gly Ser Pro Lys Ile Gln Gly Ile Asn
            260                 265                 270

Glu Leu Ile Asn Ser His His Asn Gln His Cys His Lys Ser Glu Arg
        275                 280                 285

Ile Ala Lys Leu Arg Pro Leu His Lys Gln Ile Leu Ser Asp Gly Met
    290                 295                 300

Ser Val Ser Phe Leu Pro Ser Lys Phe Ala Asp Asp Ser Glu Met Cys
305                 310                 315                 320

Gln Ala Val Asn Glu Phe Tyr Arg His Tyr Ala Asp Val Phe Ala Lys
                325                 330                 335

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

Tyr Val Glu His Lys Asn Leu Asn Glu Leu Ser Lys Gln Ala Phe Gly
        355                 360                 365

Asp Phe Ala Leu Leu Gly Arg Val Leu Asp Gly Tyr Tyr Val Asp Val
    370                 375                 380

Val Asn Pro Glu Phe Asn Glu Arg Phe Ala Lys Ala Lys Thr Asp Asn
385                 390                 395                 400

Ala Lys Ala Lys Leu Thr Lys Glu Lys Asp Lys Phe Ile Lys Gly Val
                405                 410                 415
```

```
His Ser Leu Ala Ser Leu Glu Gln Ala Ile Glu His Tyr Thr Ala Arg
            420                 425                 430

His Asp Asp Glu Ser Val Gln Ala Gly Lys Leu Gly Gln Tyr Phe Lys
        435                 440                 445

His Gly Leu Ala Gly Val Asp Asn Pro Ile Gln Lys Ile His Asn Asn
    450                 455                 460

His Ser Thr Ile Lys Gly Phe Leu Glu Arg Glu Arg Pro Ala Gly Glu
465                 470                 475                 480

Arg Ala Leu Pro Lys Ile Lys Ser Gly Lys Asn Pro Glu Met Thr Gln
                485                 490                 495

Leu Arg Gln Leu Lys Glu Leu Leu Asp Asn Ala Leu Asn Val Ala His
            500                 505                 510

Phe Ala Lys Leu Leu Thr Thr Lys Thr Thr Leu Asp Asn Gln Asp Gly
        515                 520                 525

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

Pro Thr Leu Tyr Asn Lys Val Arg Asp Tyr Leu Ser Gln Lys Pro Phe
545                 550                 555                 560

Ser Thr Glu Lys Tyr Lys Leu Asn Phe Gly Asn Pro Thr Leu Leu Asn
                565                 570                 575

Gly Trp Asp Leu Asn Lys Glu Lys Asp Asn Phe Gly Val Ile Leu Gln
            580                 585                 590

Lys Asp Gly Cys Tyr Tyr Leu Ala Leu Leu Asp Lys Ala His Lys Lys
        595                 600                 605

Val Phe Asp Asn Ala Pro Asn Thr Gly Lys Ser Ile Tyr Gln Lys Met
    610                 615                 620

Ile Tyr Lys Tyr Leu Glu Val Arg Lys Gln Phe Pro Lys Val Phe Phe
```

625 630 635 640

Ser Lys Glu Ala Ile Ala Ile Asn Tyr His Pro Ser Lys Glu Leu Val
645 650 655

Glu Ile Lys Asp Lys Gly Arg Gln Arg Ser Asp Asp Glu Arg Leu Lys
660 665 670

Leu Tyr Arg Phe Ile Leu Glu Cys Leu Lys Ile His Pro Lys Tyr Asp
675 680 685

Lys Lys Phe Glu Gly Ala Ile Gly Asp Ile Gln Leu Phe Lys Lys Asp
690 695 700

Lys Lys Gly Arg Glu Val Pro Ile Ser Glu Lys Asp Leu Phe Lys Asp
705 710 715 720

Ile Asn Gly Ile Phe Ser Ser Lys Pro Lys Leu Glu Met Glu Asp Phe
725 730 735

Phe Ile Gly Glu Phe Lys Arg Tyr Asn Pro Ser Gln Asp Leu Val Asp
740 745 750

Gln Tyr Asn Ile Tyr Lys Lys Ile Asp Ser Asn Asp Asn Arg Lys Lys
755 760 765

Glu Asn Phe Tyr Asn Asn His Pro Lys Phe Lys Lys Asp Leu Val Arg
770 775 780

Tyr Tyr Tyr Glu Ser Met Cys Lys His Glu Glu Trp Glu Glu Ser Phe
785 790 795 800

Glu Phe Ser Lys Lys Leu Gln Asp Ile Gly Cys Tyr Val Asp Val Asn
805 810 815

Glu Leu Phe Thr Glu Ile Glu Thr Arg Arg Leu Asn Tyr Lys Ile Ser
820 825 830

Phe Cys Asn Ile Asn Ala Asp Tyr Ile Asp Glu Leu Val Glu Gln Gly
835 840 845

```
Gln Leu Tyr Leu Phe Gln Ile Tyr Asn Lys Asp Phe Ser Pro Lys Ala
    850                 855                 860

His Gly Lys Pro Asn Leu His Thr Leu Tyr Phe Lys Ala Leu Phe Ser
865                 870                 875                 880

Glu Asp Asn Leu Ala Asp Pro Ile Tyr Lys Leu Asn Gly Glu Ala Gln
                885                 890                 895

Ile Phe Tyr Arg Lys Ala Ser Leu Asp Met Asn Glu Thr Thr Ile His
            900                 905                 910

Arg Ala Gly Glu Val Leu Glu Asn Lys Asn Pro Asp Asn Pro Lys Lys
        915                 920                 925

Arg Gln Phe Val Tyr Asp Ile Ile Lys Asp Lys Arg Tyr Thr Gln Lys
    930                 935                 940

Asp Phe Met Leu His Val Pro Ile Thr Met Asn Phe Gly Val Gln Gly
945                 950                 955                 960

Met Thr Ile Lys Glu Phe Asn Lys Lys Val Asn Gln Ser Ile Gln Gln
                965                 970                 975

Tyr Asp Glu Val Asn Val Ile Gly Ile Asp Arg Gly Glu Arg His Leu
            980                 985                 990

Leu Tyr Leu Thr Val Ile Asn Ser  Lys Gly Glu Ile Leu  Glu Gln Cys
        995                 1000                1005

Ser Leu  Asn Asp Ile Thr Thr  Ala Ser Ala Asn Gly  Thr Gln Met
    1010                1015                1020

Thr Thr  Pro Tyr His Lys Ile  Leu Asp Lys Arg Glu  Ile Glu Arg
    1025                1030                1035

Leu Asn  Ala Arg Val Gly Trp  Gly Glu Ile Glu Thr  Ile Lys Glu
    1040                1045                1050

Leu Lys  Ser Gly Tyr Leu Ser  His Val Val His Gln  Ile Ser Gln
```

```
    1055                1060                1065

Leu Met  Leu Lys Tyr Asn Ala  Ile Val Val Leu Glu  Asp Leu Asn
    1070                1075                1080

Phe Gly  Phe Lys Arg Gly Arg  Phe Lys Val Glu Lys  Gln Ile Tyr
    1085                1090                1095

Gln Asn  Phe Glu Asn Ala Leu  Ile Lys Lys Leu Asn  His Leu Val
    1100                1105                1110

Leu Lys  Asp Lys Ala Asp Asp  Glu Ile Gly Ser Tyr  Lys Asn Ala
    1115                1120                1125

Leu Gln  Leu Thr Asn Asn Phe  Thr Asp Leu Lys Ser  Ile Gly Lys
    1130                1135                1140

Gln Thr  Gly Phe Leu Phe Tyr  Val Pro Ala Trp Asn  Thr Ser Lys
    1145                1150                1155

Ile Asp  Pro Glu Thr Gly Phe  Val Asp Leu Leu Lys  Pro Arg Tyr
    1160                1165                1170

Glu Asn  Ile Gln Ala Ser Gln  Ala Phe Phe Gly Lys  Phe Asp Lys
    1175                1180                1185

Ile Cys  Tyr Asn Ala Asp Lys  Asp Tyr Phe Glu Phe  His Ile Asp
    1190                1195                1200

Tyr Ala  Lys Phe Thr Asp Lys  Ala Lys Asn Ser Arg  Gln Ile Trp
    1205                1210                1215

Thr Ile  Cys Ser His Gly Asp  Lys Arg Tyr Val Tyr  Asp Lys Thr
    1220                1225                1230

Ala Asn  Gln Asn Lys Gly Ala  Ala Lys Gly Ile Asn  Val Asn Asp
    1235                1240                1245

Ile Leu  Lys Ser Leu Phe Ala  Arg His His Ile Asn  Glu Lys Gln
    1250                1255                1260
```

```
Pro Asn  Leu Val Met Asp Ile  Cys Gln Asn Asn Asp  Lys Glu Phe
    1265                1270                1275


His Lys  Ser Leu Met Tyr Leu  Leu Lys Thr Leu Leu  Ala Leu Arg
    1280                1285                1290


Tyr Ser  Asn Ala Ser Ser Asp  Glu Asp Phe Ile Leu  Ser Pro Val
    1295                1300                1305


Ala Asn  Asp Glu Gly Val Phe  Phe Asn Ser Ala Leu  Ala Asp Asp
    1310                1315                1320


Thr Gln  Pro Gln Asn Ala Asp  Ala Asn Gly Ala Tyr  His Ile Ala
    1325                1330                1335


Leu Lys  Gly Leu Trp Leu Leu  Asn Glu Leu Lys Asn  Ser Asp Asp
    1340                1345                1350


Leu Asn  Lys Val Lys Leu Ala  Ile Asp Asn Gln Thr  Trp Leu Asn
    1355                1360                1365


Phe Ala  Gln Asn Arg
    1370


<210>  12
<211>  1352
<212>  PRT
<213>  Unknown

<220>
<223>  Parcubacteria bacterium

<400>  12

Met Glu Asn Ile Phe Asp Gln Phe Ile Gly Lys Tyr Ser Leu Ser Lys
1               5                  10                  15


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


Lys Ile Asn Lys Val Phe Glu Lys Asp Gln Thr Ile Asp Asp Ser Tyr
        35                  40                  45
```

```
Asn Gln Ala Lys Phe Tyr Phe Asp Ser Leu His Gln Lys Phe Ile Asp
    50                  55                  60

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

Ala Asp Val Leu Glu Lys Gln Asn Lys Ile Ile Leu Asp Lys Lys Arg
                85                  90                  95

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

Arg Leu Gln Lys Glu Ile Asn Asp Ala Glu Asp Ile Ile Gln Lys Glu
        115                 120                 125

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

Glu Ser Trp Lys Thr Tyr Tyr Gln Glu Arg Glu Val Asp Gly Lys Lys
145                 150                 155                 160

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

Thr Ala Ala Gly Ile Leu Lys Val Leu Lys Tyr Glu Phe Pro Glu Glu
            180                 185                 190

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

Glu Lys Glu Asn Pro Gly Gln Lys Arg Tyr Ile Phe Asp Ser Phe Asp
    210                 215                 220

Lys Phe Ala Gly Tyr Leu Thr Lys Phe Gln Gln Thr Lys Lys Asn Leu
225                 230                 235                 240

Tyr Ala Ala Asp Gly Thr Ser Thr Ala Val Ala Thr Arg Ile Ala Asp
                245                 250                 255

Asn Phe Ile Ile Phe His Gln Asn Thr Lys Val Phe Arg Asp Lys Tyr
```

```
            260                 265                 270

Lys Asn Asn His Thr Asp Leu Gly Phe Asp Glu Glu Asn Ile Phe Glu
        275                 280                 285

Ile Glu Arg Tyr Lys Asn Cys Leu Leu Gln Arg Glu Ile Glu His Ile
    290                 295                 300

Lys Asn Glu Asn Ser Tyr Asn Lys Ile Ile Gly Arg Ile Asn Lys Lys
305                 310                 315                 320

Ile Lys Glu Tyr Arg Asp Gln Lys Ala Lys Asp Thr Lys Leu Thr Lys
                325                 330                 335

Ser Asp Phe Pro Phe Phe Lys Asn Leu Asp Lys Gln Ile Leu Gly Glu
            340                 345                 350

Val Glu Lys Glu Lys Gln Leu Ile Glu Lys Thr Arg Glu Lys Thr Glu
        355                 360                 365

Glu Asp Val Leu Ile Glu Arg Phe Lys Glu Phe Ile Glu Asn Asn Glu
    370                 375                 380

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

Glu Phe Glu Ser Glu Tyr Glu Gly Ile Tyr Leu Lys Asn Lys Ala Ile
                405                 410                 415

Asn Thr Ile Ser Arg Arg Trp Phe Val Ser Asp Arg Asp Phe Glu Leu
            420                 425                 430

Lys Leu Pro Gln Gln Lys Ser Lys Asn Lys Ser Glu Lys Asn Glu Pro
        435                 440                 445

Lys Val Lys Lys Phe Ile Ser Ile Ala Glu Ile Lys Asn Ala Val Glu
    450                 455                 460

Glu Leu Asp Gly Asp Ile Phe Lys Ala Val Phe Tyr Asp Lys Lys Ile
465                 470                 475                 480
```

```
Ile Ala Gln Gly Gly Ser Lys Leu Glu Gln Phe Leu Val Ile Trp Lys
                485                 490                 495

Tyr Glu Phe Glu Tyr Leu Phe Arg Asp Ile Glu Arg Glu Asn Gly Glu
            500                 505                 510

Lys Leu Leu Gly Tyr Asp Ser Cys Leu Lys Ile Ala Lys Gln Leu Gly
        515                 520                 525

Ile Phe Pro Gln Glu Lys Glu Ala Arg Glu Lys Ala Thr Ala Val Ile
    530                 535                 540

Lys Asn Tyr Ala Asp Ala Gly Leu Gly Ile Phe Gln Met Met Lys Tyr
545                 550                 555                 560

Phe Ser Leu Asp Asp Lys Asp Arg Lys Asn Thr Pro Gly Gln Leu Ser
                565                 570                 575

Thr Asn Phe Tyr Ala Glu Tyr Asp Gly Tyr Tyr Lys Asp Phe Glu Phe
            580                 585                 590

Ile Lys Tyr Tyr Asn Glu Phe Arg Asn Phe Ile Thr Lys Lys Pro Phe
        595                 600                 605

Asp Glu Asp Lys Ile Lys Leu Asn Phe Glu Asn Gly Ala Leu Leu Lys
    610                 615                 620

Gly Trp Asp Glu Asn Lys Glu Tyr Asp Phe Met Gly Val Ile Leu Lys
625                 630                 635                 640

Lys Glu Gly Arg Leu Tyr Leu Gly Ile Met His Lys Asn His Arg Lys
                645                 650                 655

Leu Phe Gln Ser Met Gly Asn Ala Lys Gly Asp Asn Ala Asn Arg Tyr
            660                 665                 670

Gln Lys Met Ile Tyr Lys Gln Ile Ala Asp Ala Ser Lys Asp Val Pro
        675                 680                 685

Arg Leu Leu Leu Thr Ser Lys Lys Ala Met Glu Lys Phe Lys Pro Ser
```

```
    690                 695                 700

Gln Glu Ile Leu Arg Ile Lys Lys Glu Lys Thr Phe Lys Arg Glu Ser
705                 710                 715                 720

Lys Asn Phe Ser Leu Arg Asp Leu His Ala Leu Ile Glu Tyr Tyr Arg
                725                 730                 735

Asn Cys Ile Pro Gln Tyr Ser Asn Trp Ser Phe Tyr Asp Phe Gln Phe
            740                 745                 750

Gln Asp Thr Gly Lys Tyr Gln Asn Ile Lys Glu Phe Thr Asp Asp Val
        755                 760                 765

Gln Lys Tyr Gly Tyr Lys Ile Ser Phe Arg Asp Ile Asp Asp Glu Tyr
    770                 775                 780

Ile Asn Gln Ala Leu Asn Glu Gly Lys Met Tyr Leu Phe Glu Val Val
785                 790                 795                 800

Asn Lys Asp Ile Tyr Asn Thr Lys Asn Gly Ser Lys Asn Leu His Thr
                805                 810                 815

Leu Tyr Phe Glu His Ile Leu Ser Ala Glu Asn Leu Asn Asp Pro Val
            820                 825                 830

Phe Lys Leu Ser Gly Met Ala Glu Ile Phe Gln Arg Gln Pro Ser Val
        835                 840                 845

Asn Glu Arg Glu Lys Ile Thr Thr Gln Lys Asn Gln Cys Ile Leu Asp
    850                 855                 860

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

Met Phe His Met Ser Leu Val Leu Asn Thr Gly Lys Gly Glu Ile Lys
                885                 890                 895

Gln Val Gln Phe Asn Lys Ile Ile Asn Gln Arg Ile Ser Ser Ser Asp
            900                 905                 910
```

```
Asn Glu Met Arg Val Asn Val Ile Gly Ile Asp Arg Gly Glu Lys Asn
        915                 920                 925

Leu Leu Tyr Tyr Ser Val Val Lys Gln Asn Gly Glu Ile Ile Glu Gln
    930                 935                 940

Ala Ser Leu Asn Glu Ile Asn Gly Val Asn Tyr Arg Asp Lys Leu Ile
945                 950                 955                 960

Glu Arg Glu Lys Glu Arg Leu Lys Asn Arg Gln Ser Trp Lys Pro Val
                965                 970                 975

Val Lys Ile Lys Asp Leu Lys Lys Gly Tyr Ile Ser His Val Ile His
            980                 985                 990

Lys Ile Cys Gln Leu Ile Glu Lys  Tyr Ser Ala Ile Val  Val Leu Glu
        995                 1000                1005

Asp Leu  Asn Met Arg Phe Lys  Gln Ile Arg Gly Gly  Ile Glu Arg
    1010                1015                1020

Ser Val  Tyr Gln Gln Phe Glu  Lys Ala Leu Ile Asp  Lys Leu Gly
    1025                1030                1035

Tyr Leu  Val Phe Lys Asp Asn  Arg Asp Leu Arg Ala  Pro Gly Gly
    1040                1045                1050

Val Leu  Asn Gly Tyr Gln Leu  Ser Ala Pro Phe Val  Ser Phe Glu
    1055                1060                1065

Lys Met  Arg Lys Gln Thr Gly  Ile Leu Phe Tyr Thr  Gln Ala Glu
    1070                1075                1080

Tyr Thr  Ser Lys Thr Asp Pro  Ile Thr Gly Phe Arg  Lys Asn Val
    1085                1090                1095

Tyr Ile  Ser Asn Ser Ala Ser  Leu Asp Lys Ile Lys  Glu Ala Val
    1100                1105                1110

Lys Lys  Phe Asp Ala Ile Gly  Trp Asp Gly Lys Glu  Gln Ser Tyr
```

```
    1115                1120                1125

Phe Phe  Lys Tyr Asn Pro Tyr  Asn Leu Ala Asp Glu  Lys Tyr Lys
    1130                1135                1140

Asn Ser  Thr Val Ser Lys Glu  Trp Ala Ile Phe Ala  Ser Ala Pro
    1145                1150                1155

Arg Ile  Arg Arg Gln Lys Gly  Glu Asp Gly Tyr Trp  Lys Tyr Asp
    1160                1165                1170

Arg Val  Lys Val Asn Glu Glu  Phe Glu Lys Leu Leu  Lys Val Trp
    1175                1180                1185

Asn Phe  Val Asn Pro Lys Ala  Thr Asp Ile Lys Gln  Glu Ile Ile
    1190                1195                1200

Lys Lys  Ile Lys Ala Gly Asp  Leu Gln Gly Glu Lys  Glu Leu Asp
    1205                1210                1215

Gly Arg  Leu Arg Asn Phe Trp  His Ser Phe Ile Tyr  Leu Phe Asn
    1220                1225                1230

Leu Val  Leu Glu Leu Arg Asn  Ser Phe Ser Leu Gln  Ile Lys Ile
    1235                1240                1245

Lys Ala  Gly Glu Val Ile Ala  Val Asp Glu Gly Val  Asp Phe Ile
    1250                1255                1260

Ala Ser  Pro Val Lys Pro Phe  Phe Thr Thr Pro Asn  Pro Tyr Ile
    1265                1270                1275

Pro Ser  Asn Leu Cys Trp Leu  Ala Val Glu Asn Ala  Asp Ala Asn
    1280                1285                1290

Gly Ala  Tyr Asn Ile Ala Arg  Lys Gly Val Met Ile  Leu Lys Lys
    1295                1300                1305

Ile Arg  Glu His Ala Lys Lys  Asp Pro Glu Phe Lys  Lys Leu Pro
    1310                1315                1320
```

```
Asn Leu  Phe Ile Ser Asn Ala  Glu Trp Asp Glu Ala  Ala Arg Asp
    1325                1330                1335


Trp Gly  Lys Tyr Ala Gly Thr  Thr Ala Leu Asn Leu  Asp His
    1340                1345                1350


<210>  13
<211>  1260
<212>  PRT
<213>  Porphyromonas crevioricanis

<400>  13

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


Leu Arg Phe Glu Leu Lys Pro Val Gly Lys Thr Leu Glu Asn Ile Glu
            20                  25                  30


Lys Ala Gly Ile Leu Lys Glu Asp Glu His Arg Ala Glu Ser Tyr Arg
        35                  40                  45


Arg Val Lys Lys Ile Ile Asp Thr Tyr His Lys Val Phe Ile Asp Ser
    50                  55                  60


Ser Leu Glu Asn Met Ala Lys Met Gly Ile Glu Asn Glu Ile Lys Ala
65                  70                  75                  80


Met Leu Gln Ser Phe Cys Glu Leu Tyr Lys Lys Asp His Arg Thr Glu
                85                  90                  95


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


Ile Val Gly Ala Phe Thr Gly Val Cys Gly Arg Arg Glu Asn Thr Val
        115                 120                 125


Gln Asn Glu Lys Tyr Glu Ser Leu Phe Lys Glu Lys Leu Ile Lys Glu
    130                 135                 140


Ile Leu Pro Asp Phe Val Leu Ser Thr Glu Ala Glu Ser Leu Pro Phe
145                 150                 155                 160
```

```
Ser Val Glu Glu Ala Thr Arg Ser Leu Lys Glu Phe Asp Ser Phe Thr
                165                 170                 175

Ser Tyr Phe Ala Gly Phe Tyr Glu Asn Arg Lys Asn Ile Tyr Ser Thr
            180                 185                 190

Lys Pro Gln Ser Thr Ala Ile Ala Tyr Arg Leu Ile His Glu Asn Leu
        195                 200                 205

Pro Lys Phe Ile Asp Asn Ile Leu Val Phe Gln Lys Ile Lys Glu Pro
    210                 215                 220

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

Tyr Ile Lys Lys Asp Glu Arg Leu Glu Asp Ile Phe Ser Leu Asn Tyr
                245                 250                 255

Tyr Ile His Val Leu Ser Gln Ala Gly Ile Glu Lys Tyr Asn Ala Leu
            260                 265                 270

Ile Gly Lys Ile Val Thr Glu Gly Asp Gly Glu Met Lys Gly Leu Asn
        275                 280                 285

Glu His Ile Asn Leu Tyr Asn Gln Gln Arg Gly Arg Glu Asp Arg Leu
    290                 295                 300

Pro Leu Phe Arg Pro Leu Tyr Lys Gln Ile Leu Ser Asp Arg Glu Gln
305                 310                 315                 320

Leu Ser Tyr Leu Pro Glu Ser Phe Glu Lys Asp Glu Glu Leu Leu Arg
                325                 330                 335

Ala Leu Lys Glu Phe Tyr Asp His Ile Ala Glu Asp Ile Leu Gly Arg
            340                 345                 350

Thr Gln Gln Leu Met Thr Ser Ile Ser Glu Tyr Asp Leu Ser Arg Ile
        355                 360                 365
```

```
Tyr Val Arg Asn Asp Ser Gln Leu Thr Asp Ile Ser Lys Lys Met Leu
    370                 375                 380

Gly Asp Trp Asn Ala Ile Tyr Met Ala Arg Glu Arg Ala Tyr Asp His
385                 390                 395                 400

Glu Gln Ala Pro Lys Arg Ile Thr Ala Lys Tyr Glu Arg Asp Arg Ile
                405                 410                 415

Lys Ala Leu Lys Gly Glu Glu Ser Ile Ser Leu Ala Asn Leu Asn Ser
            420                 425                 430

Cys Ile Ala Phe Leu Asp Asn Val Arg Asp Cys Arg Val Asp Thr Tyr
        435                 440                 445

Leu Ser Thr Leu Gly Gln Lys Glu Gly Pro His Gly Leu Ser Asn Leu
    450                 455                 460

Val Glu Asn Val Phe Ala Ser Tyr His Glu Ala Glu Gln Leu Leu Ser
465                 470                 475                 480

Phe Pro Tyr Pro Glu Glu Asn Asn Leu Ile Gln Asp Lys Asp Asn Val
                485                 490                 495

Val Leu Ile Lys Asn Leu Leu Asp Asn Ile Ser Asp Leu Gln Arg Phe
            500                 505                 510

Leu Lys Pro Leu Trp Gly Met Gly Asp Glu Pro Asp Lys Asp Glu Arg
        515                 520                 525

Phe Tyr Gly Glu Tyr Asn Tyr Ile Arg Gly Ala Leu Asp Gln Val Ile
    530                 535                 540

Pro Leu Tyr Asn Lys Val Arg Asn Tyr Leu Thr Arg Lys Pro Tyr Ser
545                 550                 555                 560

Thr Arg Lys Val Lys Leu Asn Phe Gly Asn Ser Gln Leu Leu Ser Gly
                565                 570                 575

Trp Asp Arg Asn Lys Glu Lys Asp Asn Ser Cys Val Ile Leu Arg Lys
            580                 585                 590
```

```
Gly Gln Asn Phe Tyr Leu Ala Ile Met Asn Asn Arg His Lys Arg Ser
        595                 600                 605

Phe Glu Asn Lys Met Leu Pro Glu Tyr Lys Glu Gly Glu Pro Tyr Phe
    610                 615                 620

Glu Lys Met Asp Tyr Lys Phe Leu Pro Asp Pro Asn Lys Met Leu Pro
625                 630                 635                 640

Lys Val Phe Leu Ser Lys Lys Gly Ile Glu Ile Tyr Lys Pro Ser Pro
                645                 650                 655

Lys Leu Leu Glu Gln Tyr Gly His Gly Thr His Lys Lys Gly Asp Thr
            660                 665                 670

Phe Ser Met Asp Asp Leu His Glu Leu Ile Asp Phe Phe Lys His Ser
        675                 680                 685

Ile Glu Ala His Glu Asp Trp Lys Gln Phe Gly Phe Lys Phe Ser Asp
    690                 695                 700

Thr Ala Thr Tyr Glu Asn Val Ser Ser Phe Tyr Arg Glu Val Glu Asp
705                 710                 715                 720

Gln Gly Tyr Lys Leu Ser Phe Arg Lys Val Ser Glu Ser Tyr Val Tyr
                725                 730                 735

Ser Leu Ile Asp Gln Gly Lys Leu Tyr Leu Phe Gln Ile Tyr Asn Lys
            740                 745                 750

Asp Phe Ser Pro Cys Ser Lys Gly Thr Pro Asn Leu His Thr Leu Tyr
        755                 760                 765

Trp Arg Met Leu Phe Asp Glu Arg Asn Leu Ala Asp Val Ile Tyr Lys
    770                 775                 780

Leu Asp Gly Lys Ala Glu Ile Phe Phe Arg Glu Lys Ser Leu Lys Asn
785                 790                 795                 800
```

```
Asp His Pro Thr His Pro Ala Gly Lys Pro Ile Lys Lys Lys Ser Arg
               805                 810                 815

Gln Lys Lys Gly Glu Glu Ser Leu Phe Glu Tyr Asp Leu Val Lys Asp
            820                 825                 830

Arg Arg Tyr Thr Met Asp Lys Phe Gln Phe His Val Pro Ile Thr Met
        835                 840                 845

Asn Phe Lys Cys Ser Ala Gly Ser Lys Val Asn Asp Met Val Asn Ala
    850                 855                 860

His Ile Arg Glu Ala Lys Asp Met His Val Ile Gly Ile Asp Arg Gly
865                 870                 875                 880

Glu Arg Asn Leu Leu Tyr Ile Cys Val Ile Asp Ser Arg Gly Thr Ile
               885                 890                 895

Leu Asp Gln Ile Ser Leu Asn Thr Ile Asn Asp Ile Asp Tyr His Asp
            900                 905                 910

Leu Leu Glu Ser Arg Asp Lys Asp Arg Gln Gln Glu His Arg Asn Trp
        915                 920                 925

Gln Thr Ile Glu Gly Ile Lys Glu Leu Lys Gln Gly Tyr Leu Ser Gln
    930                 935                 940

Ala Val His Arg Ile Ala Glu Leu Met Val Ala Tyr Lys Ala Val Val
945                 950                 955                 960

Ala Leu Glu Asp Leu Asn Met Gly Phe Lys Arg Gly Arg Gln Lys Val
               965                 970                 975

Glu Ser Ser Val Tyr Gln Gln Phe Glu Lys Gln Leu Ile Asp Lys Leu
            980                 985                 990

Asn Tyr Leu Val Asp Lys Lys Lys  Arg Pro Glu Asp Ile  Gly Gly Leu
        995                 1000                1005

Leu Arg  Ala Tyr Gln Phe Thr  Ala Pro Phe Lys Ser  Phe Lys Glu
    1010                1015                1020
```

```
Met Gly  Lys Gln Asn Gly Phe  Leu Phe Tyr Ile Pro  Ala Trp Asn
    1025                1030                1035

Thr Ser  Asn Ile Asp Pro Thr  Thr Gly Phe Val Asn  Leu Phe His
    1040                1045                1050

Val Gln  Tyr Glu Asn Val Asp  Lys Ala Lys Ser Phe  Phe Gln Lys
    1055                1060                1065

Phe Asp  Ser Ile Ser Tyr Asn  Pro Lys Lys Asp Trp  Phe Glu Phe
    1070                1075                1080

Ala Phe  Asp Tyr Lys Asn Phe  Thr Lys Lys Ala Glu  Gly Ser Arg
    1085                1090                1095

Ser Met  Trp Ile Leu Cys Thr  His Gly Ser Arg Ile  Lys Asn Phe
    1100                1105                1110

Arg Asn  Ser Gln Lys Asn Gly  Gln Trp Asp Ser Glu  Glu Phe Ala
    1115                1120                1125

Leu Thr  Glu Ala Phe Lys Ser  Leu Phe Val Arg Tyr  Glu Ile Asp
    1130                1135                1140

Tyr Thr  Ala Asp Leu Lys Thr  Ala Ile Val Asp Glu  Lys Gln Lys
    1145                1150                1155

Asp Phe  Phe Val Asp Leu Leu  Lys Leu Phe Lys Leu  Thr Val Gln
    1160                1165                1170

Met Arg  Asn Ser Trp Lys Glu  Lys Asp Leu Asp Tyr  Leu Ile Ser
    1175                1180                1185

Pro Val  Ala Gly Ala Asp Gly  Arg Phe Phe Asp Thr  Arg Glu Gly
    1190                1195                1200

Asn Lys  Ser Leu Pro Lys Asp  Ala Asp Ala Asn Gly  Ala Tyr Asn
    1205                1210                1215
```

```
Ile Ala  Leu Lys Gly Leu Trp  Ala Leu Arg Gln Ile  Arg Gln Thr
    1220                1225                1230


Ser Glu  Gly Gly Lys Leu Lys  Leu Ala Ile Ser Asn  Lys Glu Trp
    1235                1240                1245


Leu Gln  Phe Val Gln Glu Arg  Ser Tyr Glu Lys Asp
    1250                1255                1260


<210>  14
<211>  1324
<212>  PRT
<213>  Prevotella disiens

<400>  14

Met Glu Asn Tyr Gln Glu Phe Thr Asn Leu Phe Gln Leu Asn Lys Thr
1               5                   10                  15


Leu Arg Phe Glu Leu Lys Pro Ile Gly Lys Thr Cys Glu Leu Leu Glu
            20                  25                  30


Glu Gly Lys Ile Phe Ala Ser Gly Ser Phe Leu Glu Lys Asp Lys Val
        35                  40                  45


Arg Ala Asp Asn Val Ser Tyr Val Lys Lys Glu Ile Asp Lys Lys His
    50                  55                  60


Lys Ile Phe Ile Glu Glu Thr Leu Ser Ser Phe Ser Ile Ser Asn Asp
65                  70                  75                  80


Leu Leu Lys Gln Tyr Phe Asp Cys Tyr Asn Glu Leu Lys Ala Phe Lys
                85                  90                  95


Lys Asp Cys Lys Ser Asp Glu Glu Glu Val Lys Lys Thr Ala Leu Arg
            100                 105                 110


Asn Lys Cys Thr Ser Ile Gln Arg Ala Met Arg Glu Ala Ile Ser Gln
        115                 120                 125


Ala Phe Leu Lys Ser Pro Gln Lys Lys Leu Leu Ala Ile Lys Asn Leu
    130                 135                 140
```

```
Ile Glu Asn Val Phe Lys Ala Asp Glu Asn Val Gln His Phe Ser Glu
145                 150                 155                 160

Phe Thr Ser Tyr Phe Ser Gly Phe Glu Thr Asn Arg Glu Asn Phe Tyr
                165                 170                 175

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

Asn Leu Pro Ile Phe Ile Lys Asn Ile Tyr Ile Phe Glu Lys Leu Lys
        195                 200                 205

Glu Gln Phe Asp Ala Lys Thr Leu Ser Glu Ile Phe Glu Asn Tyr Lys
    210                 215                 220

Leu Tyr Val Ala Gly Ser Ser Leu Asp Glu Val Phe Ser Leu Glu Tyr
225                 230                 235                 240

Phe Asn Asn Thr Leu Thr Gln Lys Gly Ile Asp Asn Tyr Asn Ala Val
                245                 250                 255

Ile Gly Lys Ile Val Lys Glu Asp Lys Gln Glu Ile Gln Gly Leu Asn
            260                 265                 270

Glu His Ile Asn Leu Tyr Asn Gln Lys His Lys Asp Arg Arg Leu Pro
        275                 280                 285

Phe Phe Ile Ser Leu Lys Lys Gln Ile Leu Ser Asp Arg Glu Ala Leu
    290                 295                 300

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

Leu Lys Gly Phe Tyr Ile Glu Asp Gly Phe Glu Asn Asn Val Leu Thr
                325                 330                 335

Pro Leu Ala Thr Leu Leu Ser Ser Leu Asp Lys Tyr Asn Leu Asn Gly
            340                 345                 350

Ile Phe Ile Arg Asn Asn Glu Ala Leu Ser Ser Leu Ser Gln Asn Val
```

```
        355                 360                 365

Tyr Arg Asn Phe Ser Ile Asp Glu Ala Ile Asp Ala Gln Asn Ala Glu
    370                 375                 380

Leu Gln Thr Phe Asn Asn Tyr Glu Leu Ile Ala Asn Ala Leu Arg Ala
385                 390                 395                 400

Lys Ile Lys Lys Glu Thr Lys Gln Gly Arg Lys Ser Phe Glu Lys Tyr
                405                 410                 415

Glu Glu Tyr Ile Asp Lys Lys Val Lys Ala Ile Asp Ser Leu Ser Ile
            420                 425                 430

Gln Glu Ile Asn Glu Leu Val Glu Asn Tyr Val Ser Glu Phe Asn Ser
        435                 440                 445

Asn Ser Gly Asn Met Pro Arg Lys Val Glu Asp Tyr Phe Ser Leu Met
    450                 455                 460

Arg Lys Gly Asp Phe Gly Ser Asn Asp Leu Ile Glu Asn Ile Lys Thr
465                 470                 475                 480

Lys Leu Ser Ala Ala Glu Lys Leu Leu Gly Thr Lys Tyr Gln Glu Thr
                485                 490                 495

Ala Lys Asp Ile Phe Lys Lys Asp Glu Asn Ser Lys Leu Ile Lys Glu
            500                 505                 510

Leu Leu Asp Ala Thr Lys Gln Phe Gln His Phe Ile Lys Pro Leu Leu
        515                 520                 525

Gly Thr Gly Glu Glu Ala Asp Arg Asp Leu Val Phe Tyr Gly Asp Phe
    530                 535                 540

Leu Pro Leu Tyr Glu Lys Phe Glu Glu Leu Thr Leu Leu Tyr Asn Lys
545                 550                 555                 560

Val Arg Asn Arg Leu Thr Gln Lys Pro Tyr Ser Lys Asp Lys Ile Arg
                565                 570                 575
```

```
Leu Cys Phe Asn Lys Pro Lys Leu Met Thr Gly Trp Val Asp Ser Lys
            580                 585                 590

Thr Glu Lys Ser Asp Asn Gly Thr Gln Tyr Gly Gly Tyr Leu Phe Arg
        595                 600                 605

Lys Lys Asn Glu Ile Gly Glu Tyr Asp Tyr Phe Leu Gly Ile Ser Ser
    610                 615                 620

Lys Ala Gln Leu Phe Arg Lys Asn Glu Ala Val Ile Gly Asp Tyr Glu
625                 630                 635                 640

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

Tyr Glu Gly Glu Asn Ser Tyr Lys Glu Asp Lys Lys Arg Leu Asn Lys
            660                 665                 670

Val Ile Ile Ala Tyr Ile Glu Gln Ile Lys Gln Thr Asn Ile Lys Lys
        675                 680                 685

Ser Ile Ile Glu Ser Ile Ser Lys Tyr Pro Asn Ile Ser Asp Asp Asp
    690                 695                 700

Lys Val Thr Pro Ser Ser Leu Leu Glu Lys Ile Lys Lys Val Ser Ile
705                 710                 715                 720

Asp Ser Tyr Asn Gly Ile Leu Ser Phe Lys Ser Phe Gln Ser Val Asn
                725                 730                 735

Lys Glu Val Ile Asp Asn Leu Leu Lys Thr Ile Ser Pro Leu Lys Asn
            740                 745                 750

Lys Ala Glu Phe Leu Asp Leu Ile Asn Lys Asp Tyr Gln Ile Phe Thr
        755                 760                 765

Glu Val Gln Ala Val Ile Asp Glu Ile Cys Lys Gln Lys Thr Phe Ile
    770                 775                 780

Tyr Phe Pro Ile Ser Asn Val Glu Leu Glu Lys Glu Met Gly Asp Lys
```

785 790 795 800

Asp Lys Pro Leu Cys Leu Phe Gln Ile Ser Asn Lys Asp Leu Ser Phe
805 810 815

Ala Lys Thr Phe Ser Ala Asn Leu Arg Lys Lys Arg Gly Ala Glu Asn
820 825 830

Leu His Thr Met Leu Phe Lys Ala Leu Met Glu Gly Asn Gln Asp Asn
835 840 845

Leu Asp Leu Gly Ser Gly Ala Ile Phe Tyr Arg Ala Lys Ser Leu Asp
850 855 860

Gly Asn Lys Pro Thr His Pro Ala Asn Glu Ala Ile Lys Cys Arg Asn
865 870 875 880

Val Ala Asn Lys Asp Lys Val Ser Leu Phe Thr Tyr Asp Ile Tyr Lys
885 890 895

Asn Arg Arg Tyr Met Glu Asn Lys Phe Leu Phe His Leu Ser Ile Val
900 905 910

Gln Asn Tyr Lys Ala Ala Asn Asp Ser Ala Gln Leu Asn Ser Ser Ala
915 920 925

Thr Glu Tyr Ile Arg Lys Ala Asp Asp Leu His Ile Ile Gly Ile Asp
930 935 940

Arg Gly Glu Arg Asn Leu Leu Tyr Tyr Ser Val Ile Asp Met Lys Gly
945 950 955 960

Asn Ile Val Glu Gln Asp Ser Leu Asn Ile Ile Arg Asn Asn Asp Leu
965 970 975

Glu Thr Asp Tyr His Asp Leu Leu Asp Lys Arg Glu Lys Glu Arg Lys
980 985 990

Ala Asn Arg Gln Asn Trp Glu Ala Val Glu Gly Ile Lys Asp Leu Lys
995 1000 1005

```
Lys Gly  Tyr Leu Ser Gln Ala  Val His Gln Ile Ala  Gln Leu Met
    1010                1015                1020

Leu Lys  Tyr Asn Ala Ile Ile  Ala Leu Glu Asp Leu  Gly Gln Met
    1025                1030                1035

Phe Val  Thr Arg Gly Gln Lys  Ile Glu Lys Ala Val  Tyr Gln Gln
    1040                1045                1050

Phe Glu  Lys Ser Leu Val Asp  Lys Leu Ser Tyr Leu  Val Asp Lys
    1055                1060                1065

Lys Arg  Pro Tyr Asn Glu Leu  Gly Gly Ile Leu Lys  Ala Tyr Gln
    1070                1075                1080

Leu Ala  Ser Ser Ile Thr Lys  Asn Asn Ser Asp Lys  Gln Asn Gly
    1085                1090                1095

Phe Leu  Phe Tyr Val Pro Ala  Trp Asn Thr Ser Lys  Ile Asp Pro
    1100                1105                1110

Val Thr  Gly Phe Thr Asp Leu  Leu Arg Pro Lys Ala  Met Thr Ile
    1115                1120                1125

Lys Glu  Ala Gln Asp Phe Phe  Gly Ala Phe Asp Asn  Ile Ser Tyr
    1130                1135                1140

Asn Asp  Lys Gly Tyr Phe Glu  Phe Glu Thr Asn Tyr  Asp Lys Phe
    1145                1150                1155

Lys Ile  Arg Met Lys Ser Ala  Gln Thr Arg Trp Thr  Ile Cys Thr
    1160                1165                1170

Phe Gly  Asn Arg Ile Lys Arg  Lys Lys Asp Lys Asn  Tyr Trp Asn
    1175                1180                1185

Tyr Glu  Glu Val Glu Leu Thr  Glu Glu Phe Lys Lys  Leu Phe Lys
    1190                1195                1200

Asp Ser  Asn Ile Asp Tyr Glu  Asn Cys Asn Leu Lys  Glu Glu Ile
```

```
    1205                1210                1215

Gln Asn  Lys Asp Asn Arg Lys  Phe Phe Asp Asp Leu  Ile Lys Leu
    1220                1225                1230

Leu Gln  Leu Thr Leu Gln Met  Arg Asn Ser Asp Asp  Lys Gly Asn
    1235                1240                1245

Asp Tyr  Ile Ile Ser Pro Val  Ala Asn Ala Glu Gly  Gln Phe Phe
    1250                1255                1260

Asp Ser  Arg Asn Gly Asp Lys  Lys Leu Pro Leu Asp  Ala Asp Ala
    1265                1270                1275

Asn Gly  Ala Tyr Asn Ile Ala  Arg Lys Gly Leu Trp  Asn Ile Arg
    1280                1285                1290

Gln Ile  Lys Gln Thr Lys Asn  Lys Asp Asp Leu Asn  Leu Ser Ile
    1295                1300                1305

Ser Ser  Thr Glu Trp Leu Asp  Phe Val Arg Glu Lys  Pro Tyr Leu
    1310                1315                1320

Lys
```

<210> 15
<211> 1484
<212> PRT
<213> Peregrinibacteria bacterium

<220>
<221> misc_feature
<222> (1073)..(1073)
<223> Xaa can be any naturally occurring amino acid

<400> 15

```
Met Ser Asn Phe Phe Lys Asn Phe Thr Asn Leu Tyr Glu Leu Ser Lys
1               5                   10                  15

Thr Leu Arg Phe Glu Leu Lys Pro Val Gly Asp Thr Leu Thr Asn Met
            20                  25                  30
```

```
Lys Asp His Leu Glu Tyr Asp Glu Lys Leu Gln Thr Phe Leu Lys Asp
        35                  40                  45

Gln Asn Ile Asp Asp Ala Tyr Gln Ala Leu Lys Pro Gln Phe Asp Glu
    50                  55                  60

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

Glu Ile Asp Phe Ser Glu Tyr Leu Asp Leu Phe Gln Glu Lys Lys Glu
                85                  90                  95

Leu Asn Asp Ser Glu Lys Lys Leu Arg Asn Lys Ile Gly Glu Thr Phe
            100                 105                 110

Asn Lys Ala Gly Glu Lys Trp Lys Lys Glu Lys Tyr Pro Gln Tyr Glu
        115                 120                 125

Trp Lys Lys Gly Ser Lys Ile Ala Asn Gly Ala Asp Ile Leu Ser Cys
    130                 135                 140

Gln Asp Met Leu Gln Phe Ile Lys Tyr Lys Asn Pro Glu Asp Glu Lys
145                 150                 155                 160

Ile Lys Asn Tyr Ile Asp Asp Thr Leu Lys Gly Phe Phe Thr Tyr Phe
                165                 170                 175

Gly Gly Phe Asn Gln Asn Arg Ala Asn Tyr Tyr Glu Thr Lys Lys Glu
            180                 185                 190

Ala Ser Thr Ala Val Ala Thr Arg Ile Val His Glu Asn Leu Pro Lys
        195                 200                 205

Phe Cys Asp Asn Val Ile Gln Phe Lys His Ile Ile Lys Arg Lys Lys
    210                 215                 220

Asp Gly Thr Val Glu Lys Thr Glu Arg Lys Thr Glu Tyr Leu Asn Ala
225                 230                 235                 240
```

```
Tyr Gln Tyr Leu Lys Asn Asn Asn Lys Ile Thr Gln Ile Lys Asp Ala
                245                 250                 255

Glu Thr Glu Lys Met Ile Glu Ser Thr Pro Ile Ala Glu Lys Ile Phe
            260                 265                 270

Asp Val Tyr Tyr Phe Ser Ser Cys Leu Ser Gln Lys Gln Ile Glu Glu
        275                 280                 285

Tyr Asn Arg Ile Ile Gly His Tyr Asn Leu Leu Ile Asn Leu Tyr Asn
    290                 295                 300

Gln Ala Lys Arg Ser Glu Gly Lys His Leu Ser Ala Asn Glu Lys Lys
305                 310                 315                 320

Tyr Lys Asp Leu Pro Lys Phe Lys Thr Leu Tyr Lys Gln Ile Gly Cys
                325                 330                 335

Gly Lys Lys Lys Asp Leu Phe Tyr Thr Ile Lys Cys Asp Thr Glu Glu
            340                 345                 350

Glu Ala Asn Lys Ser Arg Asn Glu Gly Lys Glu Ser His Ser Val Glu
        355                 360                 365

Glu Ile Ile Asn Lys Ala Gln Glu Ala Ile Asn Lys Tyr Phe Lys Ser
    370                 375                 380

Asn Asn Asp Cys Glu Asn Ile Asn Thr Val Pro Asp Phe Ile Asn Tyr
385                 390                 395                 400

Ile Leu Thr Lys Glu Asn Tyr Glu Gly Val Tyr Trp Ser Lys Ala Ala
                405                 410                 415

Met Asn Thr Ile Ser Asp Lys Tyr Phe Ala Asn Tyr His Asp Leu Gln
            420                 425                 430

Asp Arg Leu Lys Glu Ala Lys Val Phe Gln Lys Ala Asp Lys Lys Ser
        435                 440                 445

Glu Asp Asp Ile Lys Ile Pro Glu Ala Ile Glu Leu Ser Gly Leu Phe
    450                 455                 460
```

```
Gly Val Leu Asp Ser Leu Ala Asp Trp Gln Thr Thr Leu Phe Lys Ser
465                 470                 475                 480

Ser Ile Leu Ser Asn Glu Lys Leu Lys Ile Ile Thr Asp Ser Gln Thr
                485                 490                 495

Pro Ser Glu Ala Leu Leu Lys Met Ile Phe Asn Asp Ile Glu Lys Asn
            500                 505                 510

Met Glu Ser Phe Leu Lys Glu Thr Asn Asp Ile Ile Thr Leu Lys Lys
        515                 520                 525

Tyr Lys Gly Asn Lys Glu Gly Thr Glu Lys Ile Lys Gln Trp Phe Asp
    530                 535                 540

Tyr Thr Leu Ala Ile Asn Arg Met Leu Lys Tyr Phe Leu Val Lys Glu
545                 550                 555                 560

Asn Lys Ile Lys Gly Asn Ser Leu Asp Thr Asn Ile Ser Glu Ala Leu
                565                 570                 575

Lys Thr Leu Ile Tyr Ser Asp Asp Ala Glu Trp Phe Lys Trp Tyr Asp
            580                 585                 590

Ala Leu Arg Asn Tyr Leu Thr Gln Lys Pro Gln Asp Glu Ala Lys Glu
        595                 600                 605

Asn Lys Leu Lys Leu Asn Phe Asp Asn Pro Ser Leu Ala Gly Gly Trp
    610                 615                 620

Asp Val Asn Lys Glu Cys Ser Asn Phe Cys Val Ile Leu Lys Asp Lys
625                 630                 635                 640

Asn Glu Lys Lys Tyr Leu Ala Met Ile Lys Lys Gly Glu Asn Thr Leu
                645                 650                 655

Phe Gln Lys Glu Trp Thr Glu Gly Arg Gly Lys Asn Leu Thr Lys Lys
            660                 665                 670
```

```
Ser Asn Pro Leu Phe Glu Ile Asn Asn Cys Glu Ile Leu Ser Lys Met
        675                 680                 685

Glu Tyr Asp Phe Trp Ala Asp Val Ser Lys Met Ile Pro Lys Cys Ser
    690                 695                 700

Thr Gln Leu Lys Ala Val Val Asn His Phe Lys Gln Ser Asp Asn Glu
705                 710                 715                 720

Phe Ile Phe Pro Ile Gly Tyr Lys Val Thr Ser Gly Glu Lys Phe Arg
                725                 730                 735

Glu Glu Cys Lys Ile Ser Lys Gln Asp Phe Glu Leu Asn Asn Lys Val
            740                 745                 750

Phe Asn Lys Asn Glu Leu Ser Val Thr Ala Met Arg Tyr Asp Leu Ser
        755                 760                 765

Ser Thr Gln Glu Lys Gln Tyr Ile Lys Ala Phe Gln Lys Glu Tyr Trp
    770                 775                 780

Glu Leu Leu Phe Lys Gln Glu Lys Arg Asp Thr Lys Leu Thr Asn Asn
785                 790                 795                 800

Glu Ile Phe Asn Glu Trp Ile Asn Phe Cys Asn Lys Lys Tyr Ser Glu
                805                 810                 815

Leu Leu Ser Trp Glu Arg Lys Tyr Lys Asp Ala Leu Thr Asn Trp Ile
            820                 825                 830

Asn Phe Cys Lys Tyr Phe Leu Ser Lys Tyr Pro Lys Thr Thr Leu Phe
        835                 840                 845

Asn Tyr Ser Phe Lys Glu Ser Glu Asn Tyr Asn Ser Leu Asp Glu Phe
    850                 855                 860

Tyr Arg Asp Val Asp Ile Cys Ser Tyr Lys Leu Asn Ile Asn Thr Thr
865                 870                 875                 880

Ile Asn Lys Ser Ile Leu Asp Arg Leu Val Glu Glu Gly Lys Leu Tyr
                885                 890                 895
```

```
Leu Phe Glu Ile Lys Asn Gln Asp Ser Asn Asp Gly Lys Ser Ile Gly
            900                 905                 910

His Lys Asn Asn Leu His Thr Ile Tyr Trp Asn Ala Ile Phe Glu Asn
        915                 920                 925

Phe Asp Asn Arg Pro Lys Leu Asn Gly Glu Ala Glu Ile Phe Tyr Arg
    930                 935                 940

Lys Ala Ile Ser Lys Asp Lys Leu Gly Ile Val Lys Gly Lys Lys Thr
945                 950                 955                 960

Lys Asn Gly Thr Trp Ile Ile Lys Asn Tyr Arg Phe Ser Lys Glu Lys
                965                 970                 975

Phe Ile Leu His Val Pro Ile Thr Leu Asn Phe Cys Ser Asn Asn Glu
            980                 985                 990

Tyr Val Asn Asp Ile Val Asn Thr  Lys Phe Tyr Asn Phe  Ser Asn Leu
        995                 1000                1005

His Phe  Leu Gly Ile Asp Arg  Gly Glu Lys His Leu  Ala Tyr Tyr
    1010                1015                1020

Ser Leu  Val Asn Lys Asn Gly  Glu Ile Val Asp Gln  Gly Thr Leu
    1025                1030                1035

Asn Leu  Pro Phe Thr Asp Lys  Asp Gly Asn Gln Arg  Ser Ile Lys
    1040                1045                1050

Lys Glu  Lys Tyr Phe Tyr Asn  Lys Gln Glu Asp Lys  Trp Glu Ala
    1055                1060                1065

Lys Glu  Val Asp Xaa Trp Asn  Tyr Asn Asp Leu Leu  Asp Ala Met
    1070                1075                1080

Ala Ser  Asn Arg Asp Met Ala  Arg Lys Asn Trp Gln  Arg Ile Gly
    1085                1090                1095
```

```
Thr Ile  Lys Glu Ala Lys Asn  Gly Tyr Val Ser Leu  Val Ile Arg
    1100                1105                1110

Lys Ile  Ala Asp Leu Ala Val  Asn Asn Glu Arg Pro  Ala Phe Ile
    1115                1120                1125

Val Leu  Glu Asp Leu Asn Thr  Gly Phe Lys Arg Ser  Arg Gln Lys
    1130                1135                1140

Ile Asp  Lys Ser Val Tyr Gln  Lys Phe Glu Leu Ala  Leu Ala Lys
    1145                1150                1155

Lys Leu  Asn Phe Leu Val Asp  Lys Asn Ala Lys Arg  Asp Glu Ile
    1160                1165                1170

Gly Ser  Pro Thr Lys Ala Leu  Gln Leu Thr Pro Pro  Val Asn Asn
    1175                1180                1185

Tyr Gly  Asp Ile Glu Asn Lys  Lys Gln Ala Gly Ile  Met Leu Tyr
    1190                1195                1200

Thr Arg  Ala Asn Tyr Thr Ser  Gln Thr Asp Pro Ala  Thr Gly Trp
    1205                1210                1215

Arg Lys  Thr Ile Tyr Leu Lys  Ala Gly Pro Glu Glu  Thr Thr Tyr
    1220                1225                1230

Lys Lys  Asp Gly Lys Ile Lys  Asn Lys Ser Val Lys  Asp Gln Ile
    1235                1240                1245

Ile Glu  Thr Phe Thr Asp Ile  Gly Phe Asp Gly Lys  Asp Tyr Tyr
    1250                1255                1260

Phe Glu  Tyr Asp Lys Gly Glu  Phe Val Asp Glu Lys  Thr Gly Glu
    1265                1270                1275

Ile Lys  Pro Lys Lys Trp Arg  Leu Tyr Ser Gly Glu  Asn Gly Lys
    1280                1285                1290

Ser Leu  Asp Arg Phe Arg Gly  Glu Arg Glu Lys Asp  Lys Tyr Glu
    1295                1300                1305
```

```
Trp Lys  Ile Asp Lys Ile Asp  Ile Val Lys Ile Leu  Asp Asp Leu
    1310                1315                1320

Phe Val  Asn Phe Asp Lys Asn  Ile Ser Leu Leu Lys  Gln Leu Lys
    1325                1330                1335

Glu Gly  Val Glu Leu Thr Arg  Asn Asn Glu His Gly  Thr Gly Glu
    1340                1345                1350

Ser Leu  Arg Phe Ala Ile Asn  Leu Ile Gln Gln Ile  Arg Asn Thr
    1355                1360                1365

Gly Asn  Asn Glu Arg Asp Asn  Asp Phe Ile Leu Ser  Pro Val Arg
    1370                1375                1380

Asp Glu  Asn Gly Lys His Phe  Asp Ser Arg Glu Tyr  Trp Asp Lys
    1385                1390                1395

Glu Thr  Lys Gly Glu Lys Ile  Ser Met Pro Ser Ser  Gly Asp Ala
    1400                1405                1410

Asn Gly  Ala Phe Asn Ile Ala  Arg Lys Gly Ile Ile  Met Asn Ala
    1415                1420                1425

His Ile  Leu Ala Asn Ser Asp  Ser Lys Asp Leu Ser  Leu Phe Val
    1430                1435                1440

Ser Asp  Glu Glu Trp Asp Leu  His Leu Asn Asn Lys  Thr Glu Trp
    1445                1450                1455

Lys Lys  Gln Leu Asn Ile Phe  Ser Ser Arg Lys Ala  Met Ala Lys
    1460                1465                1470

Arg Lys  Lys Lys Arg Pro Ala  Ala Thr Lys Lys
    1475                1480
```

<210> 16
<211> 1245
<212> PRT
<213> Porphyromonas macacae

<400> 16

```
Met Lys Thr Gln His Phe Phe Glu Asp Phe Thr Ser Leu Tyr Ser Leu
1               5                  10                  15

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

Asn Ile Lys Lys Asn Gly Leu Ile Arg Arg Asp Glu Gln Arg Leu Asp
        35                  40                  45

Asp Tyr Glu Lys Leu Lys Lys Val Ile Asp Glu Tyr His Glu Asp Phe
    50                  55                  60

Ile Ala Asn Ile Leu Ser Ser Phe Ser Phe Ser Glu Glu Ile Leu Gln
65                  70                  75                  80

Ser Tyr Ile Gln Asn Leu Ser Ile Ser Glu Ala Arg Ala Lys Ile Glu
                85                  90                  95

Lys Thr Met Arg Asp Thr Leu Ala Lys Ala Phe Ser Glu Asp Glu Arg
            100                 105                 110

Tyr Lys Ser Ile Phe Lys Lys Glu Leu Val Lys Lys Asp Ile Pro Val
        115                 120                 125

Trp Cys Pro Ala Tyr Lys Ser Leu Cys Lys Lys Phe Asp Asn Phe Thr
    130                 135                 140

Thr Ser Leu Val Pro Phe His Glu Asn Arg Lys Asn Leu Tyr Thr Ser
145                 150                 155                 160

Asn Glu Ile Thr Ala Ser Ile Pro Tyr Arg Ile Val His Val Asn Leu
                165                 170                 175

Pro Lys Phe Ile Gln Asn Ile Glu Ala Leu Cys Glu Leu Gln Lys Lys
            180                 185                 190

Met Gly Ala Asp Leu Tyr Leu Glu Met Met Glu Asn Leu Arg Asn Val
        195                 200                 205
```

```
Trp Pro Ser Phe Val Lys Thr Pro Asp Asp Leu Cys Asn Leu Lys Thr
    210                 215                 220

Tyr Asn His Leu Met Val Gln Ser Ser Ile Ser Glu Tyr Asn Arg Phe
225                 230                 235                 240

Val Gly Gly Tyr Ser Thr Glu Asp Gly Thr Lys His Gln Gly Ile Asn
                245                 250                 255

Glu Trp Ile Asn Ile Tyr Arg Gln Arg Asn Lys Glu Met Arg Leu Pro
            260                 265                 270

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

Ser Phe Ile Ser Asp Thr Leu Glu Asn Asp Asp Gln Val Phe Cys Val
    290                 295                 300

Leu Arg Gln Phe Arg Lys Leu Phe Trp Asn Thr Val Ser Ser Lys Glu
305                 310                 315                 320

Asp Asp Ala Ala Ser Leu Lys Asp Leu Phe Cys Gly Leu Ser Gly Tyr
                325                 330                 335

Asp Pro Glu Ala Ile Tyr Val Ser Asp Ala His Leu Ala Thr Ile Ser
            340                 345                 350

Lys Asn Ile Phe Asp Arg Trp Asn Tyr Ile Ser Asp Ala Ile Arg Arg
        355                 360                 365

Lys Thr Glu Val Leu Met Pro Arg Lys Lys Glu Ser Val Glu Arg Tyr
    370                 375                 380

Ala Glu Lys Ile Ser Lys Gln Ile Lys Lys Arg Gln Ser Tyr Ser Leu
385                 390                 395                 400

Ala Glu Leu Asp Asp Leu Leu Ala His Tyr Ser Glu Glu Ser Leu Pro
                405                 410                 415

Ala Gly Phe Ser Leu Leu Ser Tyr Phe Thr Ser Leu Gly Gly Gln Lys
```

```
            420                 425                 430

Tyr Leu Val Ser Asp Gly Glu Val Ile Leu Tyr Glu Glu Gly Ser Asn
        435                 440                 445

Ile Trp Asp Glu Val Leu Ile Ala Phe Arg Asp Leu Gln Val Ile Leu
    450                 455                 460

Asp Lys Asp Phe Thr Glu Lys Lys Leu Gly Lys Asp Glu Glu Ala Val
465                 470                 475                 480

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

Phe Asp Leu Leu Ser Gly Thr Gly Ala Glu Ile Arg Arg Asp Ser Ser
            500                 505                 510

Phe Tyr Ala Leu Tyr Thr Asp Arg Met Asp Lys Leu Lys Gly Leu Leu
        515                 520                 525

Lys Met Tyr Asp Lys Val Arg Asn Tyr Leu Thr Lys Lys Pro Tyr Ser
    530                 535                 540

Ile Glu Lys Phe Lys Leu His Phe Asp Asn Pro Ser Leu Leu Ser Gly
545                 550                 555                 560

Trp Asp Lys Asn Lys Glu Leu Asn Asn Leu Ser Val Ile Phe Arg Gln
                565                 570                 575

Asn Gly Tyr Tyr Tyr Leu Gly Ile Met Thr Pro Lys Gly Lys Asn Leu
            580                 585                 590

Phe Lys Thr Leu Pro Lys Leu Gly Ala Glu Glu Met Phe Tyr Glu Lys
        595                 600                 605

Met Glu Tyr Lys Gln Ile Ala Glu Pro Met Leu Met Leu Pro Lys Val
    610                 615                 620

Phe Phe Pro Lys Lys Thr Lys Pro Ala Phe Ala Pro Asp Gln Ser Val
625                 630                 635                 640
```

```
Val Asp Ile Tyr Asn Lys Lys Thr Phe Lys Thr Gly Gln Lys Gly Phe
                645                 650                 655

Asn Lys Lys Asp Leu Tyr Arg Leu Ile Asp Phe Tyr Lys Glu Ala Leu
            660                 665                 670

Thr Val His Glu Trp Lys Leu Phe Asn Phe Ser Phe Ser Pro Thr Glu
        675                 680                 685

Gln Tyr Arg Asn Ile Gly Glu Phe Phe Asp Glu Val Arg Glu Gln Ala
    690                 695                 700

Tyr Lys Val Ser Met Val Asn Val Pro Ala Ser Tyr Ile Asp Glu Ala
705                 710                 715                 720

Val Glu Asn Gly Lys Leu Tyr Leu Phe Gln Ile Tyr Asn Lys Asp Phe
                725                 730                 735

Ser Pro Tyr Ser Lys Gly Ile Pro Asn Leu His Thr Leu Tyr Trp Lys
            740                 745                 750

Ala Leu Phe Ser Glu Gln Asn Gln Ser Arg Val Tyr Lys Leu Cys Gly
        755                 760                 765

Gly Gly Glu Leu Phe Tyr Arg Lys Ala Ser Leu His Met Gln Asp Thr
    770                 775                 780

Thr Val His Pro Lys Gly Ile Ser Ile His Lys Lys Asn Leu Asn Lys
785                 790                 795                 800

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

Phe Thr Glu Asp Lys Phe Phe Phe His Val Pro Ile Ser Ile Asn Tyr
            820                 825                 830

Lys Asn Lys Lys Ile Thr Asn Val Asn Gln Met Val Arg Asp Tyr Ile
        835                 840                 845

Ala Gln Asn Asp Asp Leu Gln His Gly Ile Asp Arg Gly Glu Arg Asn
```

```
    850                 855                 860

Leu Leu Tyr Ile Ser Arg Ile Asp Thr Arg Gly Asn Leu Leu Glu Gln
865                 870                 875                 880

Phe Ser Leu Asn Val Ile Glu Ser Asp Lys Gly Asp Leu Arg Thr Asp
                885                 890                 895

Tyr Gln Lys Ile Leu Gly Asp Arg Glu Gln Glu Arg Leu Arg Arg Arg
            900                 905                 910

Gln Glu Trp Lys Ser Ile Glu Ser Ile Lys Asp Leu Lys Asp Gly Tyr
        915                 920                 925

Met Ser Gln Val Val His Lys Ile Cys Asn Met Val Val Glu His Lys
    930                 935                 940

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

Lys Lys Val Glu Lys Ser Val Tyr Glu Lys Phe Glu Arg Met Leu Val
                965                 970                 975

Asp Lys Leu Asn Tyr Leu Val Val Asp Lys Lys Asn Leu Ser Asn Glu
            980                 985                 990

Pro Gly Gly Leu Tyr Ala Ala Tyr  Gln Leu Thr Asn Pro  Leu Phe Ser
        995                 1000                1005

Phe Glu  Glu Leu His Arg Tyr  Pro Gln Ser Gly Ile  Leu Phe Phe
    1010                1015                1020

Val Asp  Pro Trp Asn Thr Ser  Leu Thr Asp Pro Ser  Thr Gly Phe
    1025                1030                1035

Val Asn  Leu Leu Gly Arg Ile  Asn Tyr Thr Asn Val  Gly Asp Ala
    1040                1045                1050

Arg Lys  Phe Phe Asp Arg Phe  Asn Ala Ile Arg Tyr  Asp Gly Lys
    1055                1060                1065
```

```
Gly Asn  Ile Leu Phe Asp Leu  Asp Leu Ser Arg Phe  Asp Val Arg
    1070                1075                1080

Val Glu  Thr Gln Arg Lys Leu  Trp Thr Leu Thr Thr  Phe Gly Ser
    1085                1090                1095

Arg Ile  Ala Lys Ser Lys Lys  Ser Gly Lys Trp Met  Val Glu Arg
    1100                1105                1110

Ile Glu  Asn Leu Ser Leu Cys  Phe Leu Glu Leu Phe  Glu Gln Phe
    1115                1120                1125

Asn Ile  Gly Tyr Arg Val Glu  Lys Asp Leu Lys Lys  Ala Ile Leu
    1130                1135                1140

Ser Gln  Asp Arg Lys Glu Phe  Tyr Val Arg Leu Ile  Tyr Leu Phe
    1145                1150                1155

Asn Leu  Met Met Gln Ile Arg  Asn Ser Asp Gly Glu  Glu Asp Tyr
    1160                1165                1170

Ile Leu  Ser Pro Ala Leu Asn  Glu Lys Asn Leu Gln  Phe Asp Ser
    1175                1180                1185

Arg Leu  Ile Glu Ala Lys Asp  Leu Pro Val Asp Ala  Asp Ala Asn
    1190                1195                1200

Gly Ala  Tyr Asn Val Ala Arg  Lys Gly Leu Met Val  Val Gln Arg
    1205                1210                1215

Ile Lys  Arg Gly Asp His Glu  Ser Ile His Arg Ile  Gly Arg Ala
    1220                1225                1230

Gln Trp  Leu Arg Tyr Val Gln  Glu Gly Ile Val Glu
    1235                1240                1245
```

<210> 17
<211> 1250
<212> PRT
<213> Smithella sp.

<400> 17

```
Met Gln Thr Leu Phe Glu Asn Phe Thr Asn Gln Tyr Pro Val Ser Lys
1               5                  10                  15

Thr Leu Arg Phe Glu Leu Ile Pro Gln Gly Lys Thr Lys Asp Phe Ile
            20                  25                  30

Glu Gln Lys Gly Leu Leu Lys Lys Asp Glu Asp Arg Ala Glu Lys Tyr
        35                  40                  45

Lys Lys Val Lys Asn Ile Ile Asp Glu Tyr His Lys Asp Phe Ile Glu
    50                  55                  60

Lys Ser Leu Asn Gly Leu Lys Leu Asp Gly Leu Glu Lys Tyr Lys Thr
65                  70                  75                  80

Leu Tyr Leu Lys Gln Glu Lys Asp Asp Lys Asp Lys Lys Ala Phe Asp
                85                  90                  95

Lys Glu Lys Glu Asn Leu Arg Lys Gln Ile Ala Asn Ala Phe Arg Asn
            100                 105                 110

Asn Glu Lys Phe Lys Thr Leu Phe Ala Lys Glu Leu Ile Lys Asn Asp
        115                 120                 125

Leu Met Ser Phe Ala Cys Glu Glu Asp Lys Lys Asn Val Lys Glu Phe
    130                 135                 140

Glu Ala Phe Thr Thr Tyr Phe Thr Gly Phe His Gln Asn Arg Ala Asn
145                 150                 155                 160

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

His Glu Asn Leu Pro Lys Phe Ile Asp Asn Ile Lys Ile Phe Glu Lys
            180                 185                 190

Met Lys Lys Glu Ala Pro Glu Leu Leu Ser Pro Phe Asn Gln Thr Leu
        195                 200                 205
```

```
Lys Asp Met Lys Asp Val Ile Lys Gly Thr Thr Leu Glu Glu Ile Phe
    210                 215                 220

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

Tyr Asn Ser Val Ile Gly Gly Arg Thr Pro Glu Glu Gly Lys Thr Lys
                245                 250                 255

Ile Lys Gly Leu Asn Glu Tyr Ile Asn Thr Asp Phe Asn Gln Lys Gln
            260                 265                 270

Thr Asp Lys Lys Lys Arg Gln Pro Lys Phe Lys Gln Leu Tyr Lys Gln
        275                 280                 285

Ile Leu Ser Asp Arg Gln Ser Leu Ser Phe Ile Ala Glu Ala Phe Lys
    290                 295                 300

Asn Asp Thr Glu Ile Leu Glu Ala Ile Glu Lys Phe Tyr Val Asn Glu
305                 310                 315                 320

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

Ile Lys Asn Ala Val Ser Asn Leu Glu Ser Phe Asn Leu Thr Lys Met
            340                 345                 350

Tyr Phe Arg Ser Gly Ala Ser Leu Thr Asp Val Ser Arg Lys Val Phe
        355                 360                 365

Gly Glu Trp Ser Ile Ile Asn Arg Ala Leu Asp Asn Tyr Tyr Ala Thr
    370                 375                 380

Thr Tyr Pro Ile Lys Pro Arg Glu Lys Ser Glu Lys Tyr Glu Glu Arg
385                 390                 395                 400

Lys Glu Lys Trp Leu Lys Gln Asp Phe Asn Val Ser Leu Ile Gln Thr
                405                 410                 415

Ala Ile Asp Glu Tyr Asp Asn Glu Thr Val Lys Gly Lys Asn Ser Gly
            420                 425                 430
```

```
Lys Val Ile Ala Asp Tyr Phe Ala Lys Phe Cys Asp Asp Lys Glu Thr
        435                 440                 445

Asp Leu Ile Gln Lys Val Asn Glu Gly Tyr Ile Ala Val Lys Asp Leu
    450                 455                 460

Leu Asn Thr Pro Cys Pro Glu Asn Glu Lys Leu Gly Ser Asn Lys Asp
465                 470                 475                 480

Gln Val Lys Gln Ile Lys Ala Phe Met Asp Ser Ile Met Asp Ile Met
                485                 490                 495

His Phe Val Arg Pro Leu Ser Leu Lys Asp Thr Asp Lys Glu Lys Asp
            500                 505                 510

Glu Thr Phe Tyr Ser Leu Phe Thr Pro Leu Tyr Asp His Leu Thr Gln
        515                 520                 525

Thr Ile Ala Leu Tyr Asn Lys Val Arg Asn Tyr Leu Thr Gln Lys Pro
    530                 535                 540

Tyr Ser Thr Glu Lys Ile Lys Leu Asn Phe Glu Asn Ser Thr Leu Leu
545                 550                 555                 560

Gly Gly Trp Asp Leu Asn Lys Glu Thr Asp Asn Thr Ala Ile Ile Leu
                565                 570                 575

Arg Lys Asp Asn Leu Tyr Tyr Leu Gly Ile Met Asp Lys Arg His Asn
            580                 585                 590

Arg Ile Phe Arg Asn Val Pro Lys Ala Asp Lys Lys Asp Phe Cys Tyr
        595                 600                 605

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

Lys Val Phe Phe Ser Gln Ser Arg Ile Gln Glu Phe Thr Pro Ser Ala
625                 630                 635                 640
```

```
Lys Leu Leu Glu Asn Tyr Ala Asn Glu Thr His Lys Lys Gly Asp Asn
                645                 650                 655

Phe Asn Leu Asn His Cys His Lys Leu Ile Asp Phe Phe Lys Asp Ser
            660                 665                 670

Ile Asn Lys His Glu Asp Trp Lys Asn Phe Asp Phe Arg Phe Ser Ala
        675                 680                 685

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

Gln Gly Tyr Lys Ile Ser Phe Gln Ser Val Ala Asp Ser Phe Ile Asp
705                 710                 715                 720

Asp Leu Val Asn Glu Gly Lys Leu Tyr Leu Phe Gln Ile Tyr Asn Lys
                725                 730                 735

Asp Phe Ser Pro Phe Ser Lys Gly Lys Pro Asn Leu His Thr Leu Tyr
            740                 745                 750

Trp Lys Met Leu Phe Asp Glu Asn Asn Leu Lys Asp Val Val Tyr Lys
        755                 760                 765

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

Lys Asn Thr Thr Ile His Lys Ala Asn Glu Ser Ile Ile Asn Lys Asn
785                 790                 795                 800

Pro Asp Asn Pro Lys Ala Thr Ser Thr Phe Asn Tyr Asp Ile Val Lys
                805                 810                 815

Asp Lys Arg Tyr Thr Ile Asp Lys Phe Gln Phe His Ile Pro Ile Thr
            820                 825                 830

Met Asn Phe Lys Ala Glu Gly Ile Phe Asn Met Asn Gln Arg Val Asn
        835                 840                 845

Gln Phe Leu Lys Ala Asn Pro Asp Ile Asn Ile Ile Gly Ile Asp Arg
    850                 855                 860
```

```
Gly Glu Arg His Leu Leu Tyr Tyr Ala Leu Ile Asn Gln Lys Gly Lys
865                 870                 875                 880

Ile Leu Lys Gln Asp Thr Leu Asn Val Ile Ala Asn Glu Lys Gln Lys
                885                 890                 895

Val Asp Tyr His Asn Leu Leu Asp Lys Lys Glu Gly Asp Arg Ala Thr
            900                 905                 910

Ala Arg Gln Glu Trp Gly Val Ile Glu Thr Ile Lys Glu Leu Lys Glu
        915                 920                 925

Gly Tyr Leu Ser Gln Val Ile His Lys Leu Thr Asp Leu Met Ile Glu
    930                 935                 940

Asn Asn Ala Ile Ile Val Met Glu Asp Leu Asn Phe Gly Phe Lys Arg
945                 950                 955                 960

Gly Arg Gln Lys Val Glu Lys Gln Val Tyr Gln Lys Phe Glu Lys Met
                965                 970                 975

Leu Ile Asp Lys Leu Asn Tyr Leu Val Asp Lys Asn Lys Lys Ala Asn
            980                 985                 990

Glu Leu Gly Gly Leu Leu Asn Ala  Phe Gln Leu Ala Asn  Lys Phe Glu
        995                 1000                1005

Ser Phe  Gln Lys Met Gly Lys  Gln Asn Gly Phe Ile  Phe Tyr Val
    1010                1015                1020

Pro Ala  Trp Asn Thr Ser Lys  Thr Asp Pro Ala Thr  Gly Phe Ile
    1025                1030                1035

Asp Phe  Leu Lys Pro Arg Tyr  Glu Asn Leu Asn Gln  Ala Lys Asp
    1040                1045                1050

Phe Phe  Glu Lys Phe Asp Ser  Ile Arg Leu Asn Ser  Lys Ala Asp
    1055                1060                1065
```

```
Tyr Phe  Glu Phe Ala Phe Asp  Phe Lys Asn Phe Thr  Glu Lys Ala
    1070                1075                1080

Asp Gly  Gly Arg Thr Lys Trp  Thr Val Cys Thr Thr  Asn Glu Asp
    1085                1090                1095

Arg Tyr  Gln Trp Asn Arg Ala  Leu Asn Asn Asn Arg  Gly Ser Gln
    1100                1105                1110

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

Lys Val  Asp Tyr Lys Ser Gly  Lys Asp Leu Lys Gln  Gln Ile Ala
    1130                1135                1140

Ser Gln  Glu Ser Ala Asp Phe  Phe Lys Ala Leu Met  Lys Asn Leu
    1145                1150                1155

Ser Ile  Thr Leu Ser Leu Arg  His Asn Asn Gly Glu  Lys Gly Asp
    1160                1165                1170

Asn Glu  Gln Asp Tyr Ile Leu  Ser Pro Val Ala Asp  Ser Lys Gly
    1175                1180                1185

Arg Phe  Phe Asp Ser Arg Lys  Ala Asp Asp Asp Met  Pro Lys Asn
    1190                1195                1200

Ala Asp  Ala Asn Gly Ala Tyr  His Ile Ala Leu Lys  Gly Leu Trp
    1205                1210                1215

Cys Leu  Glu Gln Ile Ser Lys  Thr Asp Asp Leu Lys  Lys Val Lys
    1220                1225                1230

Leu Ala  Ile Ser Asn Lys Glu  Trp Leu Glu Phe Val  Gln Thr Leu
    1235                1240                1245

Lys Gly
    1250
```

<210> 18
<211> 166

<212> PRT
<213> Escherichia coli

<400> 18

Ser Glu Val Glu Phe Ser His Glu Tyr Trp Met Arg His Ala Leu Thr
1 5 10 15

Leu Ala Lys Arg Ala Trp Asp Glu Arg Glu Val Pro Val Gly Ala Val
20 25 30

Leu Val His Asn Asn Arg Val Ile Gly Glu Gly Trp Asn Arg Pro Ile
35 40 45

Gly Arg His Asp Pro Thr Ala His Ala Glu Ile Met Ala Leu Arg Gln
50 55 60

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

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

Arg Ile Gly Arg Val Val Phe Gly Ala Arg Asp Ala Lys Thr Gly Ala
100 105 110

Ala Gly Ser Leu Met Asp Val Leu His His Pro Gly Met Asn His Arg
115 120 125

Val Glu Ile Thr Glu Gly Ile Leu Ala Asp Glu Cys Ala Ala Leu Leu
130 135 140

Ser Asp Phe Phe Arg Met Arg Arg Gln Glu Ile Lys Ala Gln Lys Lys
145 150 155 160

Ala Gln Ser Ser Thr Asp
165

<210> 19
<211> 166
<212> PRT
<213> Artificial

<220>
<223> adenosine deaminase

<400> 19

Ser Glu Val Glu Phe Ser His Glu Tyr Trp Met Arg His Ala Leu Thr
1 5 10 15

Leu Ala Lys Arg Ala Arg Asp Glu Arg Glu Val Pro Val Gly Ala Val
20 25 30

Leu Val Leu Asn Asn Arg Val Ile Gly Glu Gly Trp Asn Arg Ala Ile
35 40 45

Gly Leu His Asp Pro Thr Ala His Ala Glu Ile Met Ala Leu Arg Gln
50 55 60

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

Val Thr Phe Glu Pro Cys Val Met Cys Ala Gly Ala Met Ile His Ser
85 90 95

Arg Ile Gly Arg Val Val Phe Gly Val Arg Asn Ala Lys Thr Gly Ala
100 105 110

Ala Gly Ser Leu Met Asp Val Leu His Tyr Pro Gly Met Asn His Arg
115 120 125

Val Glu Ile Thr Glu Gly Ile Leu Ala Asp Glu Cys Ala Ala Leu Leu
130 135 140

Cys Tyr Phe Phe Arg Met Pro Arg Gln Val Phe Asn Ala Gln Lys Lys
145 150 155 160

Ala Gln Ser Ser Thr Asp
165

<210> 20
<211> 166
<212> PRT
<213> Artificial

<220>
<223> adenosine deaminase

<400> 20

Ser Glu Val Glu Phe Ser His Glu Tyr Trp Met Arg His Ala Leu Thr
1 5 10 15

Leu Ala Lys Arg Ala Trp Asp Glu Arg Glu Val Pro Val Gly Ala Val
20 25 30

Leu Val Leu Asn Asn Arg Val Ile Gly Glu Gly Trp Asn Arg Ser Ile
35 40 45

Gly Leu His Asp Pro Thr Ala His Ala Glu Ile Met Ala Leu Arg Gln
50 55 60

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

Val Thr Phe Glu Pro Cys Val Met Cys Ala Gly Ala Met Ile His Ser
85 90 95

Arg Ile Gly Arg Val Val Phe Gly Val Arg Asn Ala Lys Thr Gly Ala
100 105 110

Ala Gly Ser Leu Met Asp Val Leu His Tyr Pro Gly Met Asn His Arg
115 120 125

Val Glu Ile Thr Glu Gly Ile Leu Ala Asp Glu Cys Ala Ala Leu Leu
130 135 140

Cys Tyr Phe Phe Arg Met Arg Arg Gln Val Phe Asn Ala Gln Lys Lys
145 150 155 160

Ala Gln Ser Ser Thr Asp
165

<210> 21
<211> 166
<212> PRT
<213> Artificial

<220>
<223> adenosine deaminase

<400> 21

Ser Glu Val Glu Phe Ser His Glu Tyr Trp Met Arg His Ala Leu Thr
1 5 10 15

Leu Ala Lys Arg Ala Leu Asp Glu Arg Glu Val Pro Val Gly Ala Val
20 25 30

Leu Val Leu Asn Asn Arg Val Ile Gly Glu Gly Trp Asn Arg Ala Ile
35 40 45

Gly Leu His Asp Pro Thr Ala His Ala Glu Ile Met Ala Leu Arg Gln
50 55 60

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

Val Thr Phe Glu Pro Cys Val Met Cys Ala Gly Ala Met Ile His Ser
85 90 95

Arg Ile Gly Arg Val Val Phe Gly Val Arg Asn Ala Lys Thr Gly Ala
100 105 110

Ala Gly Ser Leu Met Asp Val Leu His Tyr Pro Gly Met Asn His Arg
115 120 125

Val Glu Ile Thr Glu Gly Ile Leu Ala Asp Glu Cys Asn Ala Leu Leu
130 135 140

Cys Tyr Phe Phe Arg Met Arg Arg Gln Val Phe Asn Ala Gln Lys Lys
145 150 155 160

Ala Gln Ser Ser Thr Asp
165

<210> 22
<211> 166
<212> PRT
<213> Artificial

<220>
<223> adenosine deaminase

<400> 22

Ser Glu Val Glu Phe Ser His Glu Tyr Trp Met Arg His Ala Leu Thr
1 5 10 15

Leu Ala Lys Arg Ala Leu Asp Glu Arg Glu Val Pro Val Gly Ala Val
20 25 30

Leu Val Leu Asn Asn Arg Val Ile Gly Glu Gly Trp Asn Arg Ala Ile
35 40 45

Gly Leu His Asp Pro Thr Ala His Ala Glu Ile Met Ala Leu Arg Gln
50 55 60

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

Val Thr Phe Glu Pro Cys Val Met Cys Ala Gly Ala Met Ile His Ser
85 90 95

Arg Ile Gly Arg Val Val Phe Gly Val Arg Asn Ala Lys Thr Gly Ala
100 105 110

Ala Gly Ser Leu Met Asp Val Leu His Tyr Pro Gly Met Asn His Arg
115 120 125

Val Glu Ile Thr Glu Gly Ile Leu Ala Asp Glu Cys Asn Ala Leu Leu
130 135 140

Cys Tyr Phe Phe Arg Met Pro Arg Gln Val Phe Asn Ala Gln Lys Lys
145 150 155 160

Ala Gln Ser Ser Thr Asp
165

<210> 23
<211> 228
<212> PRT
<213> Rattus norvegicus

<400> 23

```
Ser Ser Glu Thr Gly Pro Val Ala Val Asp Pro Thr Leu Arg Arg Arg
1               5                  10                  15

Ile Glu Pro His Glu Phe Glu Val Phe Phe Asp Pro Arg Glu Leu Arg
            20                  25                  30

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

Ile Trp Arg His Thr Ser Gln Asn Thr Asn Lys His Val Glu Val Asn
    50                  55                  60

Phe Ile Glu Lys Phe Thr Thr Glu Arg Tyr Phe Cys Pro Asn Thr Arg
65                  70                  75                  80

Cys Ser Ile Thr Trp Phe Leu Ser Trp Ser Pro Cys Gly Glu Cys Ser
                85                  90                  95

Arg Ala Ile Thr Glu Phe Leu Ser Arg Tyr Pro His Val Thr Leu Phe
            100                 105                 110

Ile Tyr Ile Ala Arg Leu Tyr His His Ala Asp Pro Arg Asn Arg Gln
        115                 120                 125

Gly Leu Arg Asp Leu Ile Ser Ser Gly Val Thr Ile Gln Ile Met Thr
    130                 135                 140

Glu Gln Glu Ser Gly Tyr Cys Trp Arg Asn Phe Val Asn Tyr Ser Pro
145                 150                 155                 160

Ser Asn Glu Ala His Trp Pro Arg Tyr Pro His Leu Trp Val Arg Leu
                165                 170                 175

Tyr Val Leu Glu Leu Tyr Cys Ile Ile Leu Gly Leu Pro Pro Cys Leu
            180                 185                 190

Asn Ile Leu Arg Arg Lys Gln Pro Gln Leu Thr Phe Phe Thr Ile Ala
        195                 200                 205
```

```
Leu Gln Ser Cys His Tyr Gln Arg Leu Pro Pro His Ile Leu Trp Ala
    210                 215                 220

Thr Gly Leu Lys
225
```

<210> 24
<211> 199
<212> PRT
<213> Homo sapiens

<400> 24

```
Met Glu Ala Ser Pro Ala Ser Gly Pro Arg His Leu Met Asp Pro His
1               5                   10                  15

Ile Phe Thr Ser Asn Phe Asn Asn Gly Ile Gly Arg His Lys Thr Tyr
            20                  25                  30

Leu Cys Tyr Glu Val Glu Arg Leu Asp Asn Gly Thr Ser Val Lys Met
        35                  40                  45

Asp Gln His Arg Gly Phe Leu His Asn Gln Ala Lys Asn Leu Leu Cys
    50                  55                  60

Gly Phe Tyr Gly Arg His Ala Glu Leu Arg Phe Leu Asp Leu Val Pro
65                  70                  75                  80

Ser Leu Gln Leu Asp Pro Ala Gln Ile Tyr Arg Val Thr Trp Phe Ile
                85                  90                  95

Ser Trp Ser Pro Cys Phe Ser Trp Gly Cys Ala Gly Glu Val Arg Ala
            100                 105                 110

Phe Leu Gln Glu Asn Thr His Val Arg Leu Arg Ile Phe Ala Ala Arg
        115                 120                 125

Ile Tyr Asp Tyr Asp Pro Leu Tyr Lys Glu Ala Leu Gln Met Leu Arg
    130                 135                 140

Asp Ala Gly Ala Gln Val Ser Ile Met Thr Tyr Asp Glu Phe Lys His
145                 150                 155                 160
```

Cys Trp Asp Thr Phe Val Asp His Gln Gly Cys Pro Phe Gln Pro Trp
165 170 175

Asp Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu Arg Ala
180 185 190

Ile Leu Gln Asn Gln Gly Asn
195

<210> 25
<211> 621
<212> DNA
<213> Petromyzon marinus

<400> 25
acagatgcag agtatgtgag aattcacgaa aagctggaca tctatacctt caagaagcag 60

ttctttaaca ataagaagtc tgtgagccat aggtgctacg tgctgttcga gctgaagaga 120

aggggtgaaa gaagggcatg tttttggggg tatgctgtga acaagcccca gtctggaact 180

gagagaggca ttcacgccga aattttcagc atcagaaagg tggaggaata cctgagggat 240

aaccctggac agtttacaat taattggtat tctagctggt ctccatgcgc tgactgtgcc 300

gagaagatcc tggaatggta caaccaggag ctgagaggaa atggccatac cctgaagatt 360

tgggcctgca agctgtacta tgaaaagaac gcaagaaatc agatcggact gtggaacctg 420

agggataatg gtgtggggct gaacgtgatg gtgtccgagc actatcagtg ctgtagaaag 480

attttcattc agtcctcaca taatcagctg aacgagaata gatggctgga aaagactctg 540

aagagggctg agaagagaag gtccgaactg tcaattatga tccaggtgaa gatcctgcac 600

accactaagt cacctgccgt g 621

<210> 26
<211> 83
<212> PRT
<213> Bacillus phage

<400> 26

Thr Asn Leu Ser Asp Ile Ile Glu Lys Glu Thr Gly Lys Gln Leu Val
1 5 10 15

Ile Gln Glu Ser Ile Leu Met Leu Pro Glu Glu Val Glu Glu Val Ile

20 25 30

Gly Asn Lys Pro Glu Ser Asp Ile Leu Val His Thr Ala Tyr Asp Glu
35 40 45

Ser Thr Asp Glu Asn Val Met Leu Leu Thr Ser Asp Ala Pro Glu Tyr
50 55 60

Lys Pro Trp Ala Leu Val Ile Gln Asp Ser Asn Gly Glu Asn Lys Ile
65 70 75 80

Lys Met Leu

<210> 27
<211> 19
<212> RNA
<213> Artificial

<220>
<223> polynucleotide

<220>
<221> misc_feature
<222> (1)..(19)
<223> wherein n is A, C, T or G

<400> 27
nnnnnnnnnn nnnnnnnnn 19

<210> 28
<211> 22
<212> DNA
<213> Artificial

<220>
<223> polynucleotide

<220>
<221> misc_feature
<222> (4)..(22)
<223> wherein n is A, C, T or G

<400> 28
aaannnnnnn nnnnnnnnnn nn 22

<210> 29
<211> 22
<212> DNA
<213> Artificial

<220>
<223> polynucleotide

<220>
<221> misc_feature
<222> (4)..(22)
<223> wherein n is A, C, T or G

<400> 29
tttnnnnnnn nnnnnnnnnn nn 22

<210> 30
<211> 23
<212> DNA
<213> Artificial

<220>
<223> polynucleotide

<400> 30
ggaatccctt ctgcagcacc tgg 23

<210> 31
<211> 23
<212> DNA
<213> Artificial

<220>
<223> polynucleotide

<400> 31
ctgatggtcc atgtctgtta ctc 23

<210> 32
<211> 49
<212> DNA
<213> Artificial

<220>
<223> polynucleotide

<220>
<221> misc_feature

<222> (7)..(11)
<223> wherein n is A, T, C or G

<400> 32
cgatgtnnnn nggaatccct tctgcagcac ctgggcgcag gtcacgagg 49

<210> 33
<211> 57
<212> DNA
<213> Artificial

<220>
<223> polynucleotide

<220>
<221> misc_feature
<222> (43)..(46)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (47)..(51)
<223> wherein n is A, T, C or G

<400> 33
aattcctcgt gacctgcgcc caggtgctgc agaagggatt ccnnnnnaca tcgcatg 57

<210> 34
<211> 49
<212> DNA
<213> Artificial

<220>
<223> polynucleotide

<220>
<221> misc_feature
<222> (7)..(11)
<223> wherein n is A, T, C or G

<400> 34
cgatgtnnnn nctgatggtc catgtctgtt actcgcgcag gtcacgagg 49

<210> 35
<211> 57
<212> DNA
<213> Artificial

<220>

<223> polynucleotide

<220>
<221> misc_feature
<222> (43)..(46)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (47)..(51)
<223> wherein n is A, T, C or G

<400> 35
aattcctcgt gacctgcgcg agtaacagac atggaccatc agnnnnnaca tcgcatg 57

<210> 36
<211> 28
<212> DNA
<213> Artificial

<220>
<223> polynucleotide

<220>
<221> misc_feature
<222> (1)..(5)
<223> wherein n is A, C, T or G

<400> 36
nnnnnggaat cccttctgca gcacctgg 28

<210> 37
<211> 28
<212> DNA
<213> Artificial

<220>
<223> polynucleotide

<220>
<221> misc_feature
<222> (1)..(5)
<223> wherein n is A, C, T or G

<400> 37
nnnnncctta gggaagacgt cgtggacc 28

<210> 38

<211> 23
<212> RNA
<213> Artificial

<220>
<223> polynucleotide

<400> 38
ggaaucccuu cugcagcacc ugg 23

<210> 39
<211> 24
<212> DNA
<213> Artificial

<220>
<223> polynucleotide

<220>
<221> misc_feature
<222> (1)..(5)
<223> wherein n is A, C, T or G

<400> 39
nnnnnggaat cccttctgca gcac 24

<210> 40
<211> 1592
<212> DNA
<213> Medicago truncatula

<400> 40
actgttaata atttttaaac gtcagcgcac taaaaaaacg aaaagacgga cacgtgaaaa 60

taaaaaacac acactagttt atgacgcaat actattttac ttatgatttg ggtacattag 120

acaaaaccgt gaaagagatg tatcagctat gaaacctgta tacttcaata cagagactta 180

ctcatatcgg atacgtacgc acgaagtatc atattaatta ttttaatttt taataaatat 240

tttatcggat acttatgtga tactctacat atacacaagg atatttctaa gatactttat 300

agatacgtat cctagaaaaa catgaagagt aaaaaagtga gacaatgttg taaaaattca 360

ttataaatgt atatgattca attttagata tgcatcagta taattgattc tcgatgaaac 420

acttaaaatt atatttcttg tggaagaacg tagcgagaga ggtgattcag ttagacaaca 480

ttaaataaaa ttaatgttaa gttcttttaa tgatgtttct ctcaatatca catcatatga 540

aaatgtaata tgatttataa gaaaattttt aaaaaattta ttttaataat cacatgtact 600

atttttaaa aattgtatct tttataataa tacaataata aagagtaatc agtgttaatt 660

tttcttcaaa tataagtttt attataaatc attgttaacg tatcataagt cattaccgta 720

tcgtatctta attttttttt aaaaaccgct aattcacgta cccgtattgt attgtacccg 780

cacctgtatc acaatcgatc ttagttagaa gaattgtctc gaggcggtgc aagacagcat 840

ataatagacg tggactctct tataccaaac gttgtcgtat cacaaagggt taggtaacaa 900

gtcacagttt gtccacgtgt cacgttttaa ttggaagagg tgccgttggc gtaatataac 960

agccaatcga tttttgctat aaaagcaaat caggtaaact aaacttcttc attcttttct 1020

tccccatcgc tacaaaaccg gttcctttgg aaaagagatt cattcaaacc tagcacccaa 1080

ttccgtttca aggtataatc tactttctat tcttcgatta ttttattatt attagctact 1140

atcgtttaat cgatcttttc ttttgatccg tcaaatttaa attcaattag ggttttgttc 1200

ttttctttca tctgattgaa atccttctga attgaaccgt ttacttgatt ttactgttta 1260

ttgtatgatt taatcctttg tttttcaaag acagtcttta gattgtgatt aggggttcat 1320

ataaattttt agatttggat ttttgtattg tatgattcaa aaaatacgtc ctttaattag 1380

attagtacat ggatattttt tacccgattt attgattgtc agggagaatt tgatgagcaa 1440

gtttttttga tgtctgttgt aaattgaatt gattataatt gctgatctgc tgcttccagt 1500

tttcataacc catattcttt taaccttgtt gtacacacaa tgaaaaattg gtgattgatt 1560

catttgtttt tctttgtttt ggattataca gg 1592

<210> 41
<211> 2000
<212> DNA
<213> Zea mays

<400> 41
gtcgtgcccc tctctagaga taaagagcat tgcatgtcta aagtataaaa aattaccaca 60

tatttttttg tcacacttat ttgaagtgta gtttatctat ctctatacat atatttaaac 120

ttcactctac aaataatata gtctataata ctaaaataat attagtgttt tagaggatca 180

tataaataaa ctgctagaca tggtctaaag gataattgaa tattttgaca atctacagtt 240

ttatcttttt agtgtgcatg tgatctctct gtttttttg caaatagctt gacctatata 300

atacttcatc cattttatta gtacatccat ttaggattta gggttgatgg tttctataga 360

ctaattttta gtacatccat tttattcttt ttagtctcta aatttttaa aactaaaact 420

ctattttagt tttttattta ataatttaga tataaaatga aataaaataa attgactaca 480

aataaaacaa atacccttta agaaataaaa aaactaagca aacatttttc ttgtttcgag 540

tagataatga caggctgttc aacgccgtcg acgagtctaa cggacaccaa ccagcgaacc 600

agcagcgtcg cgtcgggcca agcgaagcag acggcacggc atctctgtag ctgcctctgg 660

acccctctcg agagttccgc tccaccgttg gacttgctcc gctgtcggca tccagaaatt 720

gcgtggcgga gcggcagacg tgaggcggca cggcaggcgg cctcttcctc ctctcacggc 780

accggcagct acgggggatt cctttcccac cgctccttcg ctttcccttc ctcgcccgcc 840

gtaataaata gacacccccc ccacaccctc tttccccaac ctcgtgttcg ttcggagcgc 900

acacacacgc aaccagatct cccccaaatc cagccgtcgg cacctccgct tcaaggtacg 960

ccgctcatcc tccccccccc cctctctcta ccttctctag atcggcgatc cggtccatgg 1020

ttagggcccg gtagttctac ttctgttcat gtttgtgtta gagcaaacat gttcatgttc 1080

atgtttgtga tgatgtggtc tggttgggcg gtcgttctag atcggagtag gatactgttt 1140

caagctacct ggtggattta ttaattttgt atctgtatgt gtgtgccata catcttcata 1200

gttacgagtt taagatgatg gatggaaata tcgatctagg ataggtatac atgttgatgc 1260

gggttttact gatgcatata cagagatgct ttttttctcg cttggttgtg atgatatggt 1320

ctggttgggc ggtcgttcta gatcggagta gaatactgtt tcaaactacc tggtggattt 1380

attaaaggat aaagggtcgt tctagatcgg agtagaatac tgtttcaaac tacctggtgg 1440

atttattaaa ggatctgtat gtatgtgcct acatcttcat agttacgagt ttaagatgat 1500

ggatggaaat atcgatctag gataggtata catgttgatg cgggttttac tgatgcatat 1560

acagagatgc ttttttcgc ttggttgtga tgatgtggtc tggttgggcg gtcgttctag 1620

atcggagtag aatactgttt caaactacct ggtggattta ttaattttgt atctttatgt 1680

gtgtgccata catcttcata gttacgagtt taagatgatg gatggaaata ttgatctagg 1740

ataggtatac atgttgatgt gggttttact gatgcatata catgatggca tatgcggcat 1800

ctattcatat gctctaacct tgagtaccta tctattataa taaacaagta tgttttataa 1860

ttattttgat cttgatatac ttggatgatg gcatatgcag cagctatatg tggatttttt 1920

agccctgcct tcatacgcta tttatttgct tggtactgtt tcttttgtcc gatgctcacc 1980

ctgttgtttg gtgatacttc 2000

<210> 42
<211> 160
<212> PRT
<213> Unknown

<220>
<223> FERNY deaminase

<400> 42

Phe Glu Arg Asn Tyr Asp Pro Arg Glu Leu Arg Lys Glu Thr Tyr Leu
1 5 10 15

Leu Tyr Glu Ile Lys Trp Gly Lys Ser Gly Lys Leu Trp Arg His Trp
20 25 30

Cys Gln Asn Asn Arg Thr Gln His Ala Glu Val Tyr Phe Leu Glu Asn
35 40 45

Ile Phe Asn Ala Arg Arg Phe Asn Pro Ser Thr His Cys Ser Ile Thr
50 55 60

Trp Tyr Leu Ser Trp Ser Pro Cys Ala Glu Cys Ser Gln Lys Ile Val
65 70 75 80

Asp Phe Leu Lys Glu His Pro Asn Val Leu Glu Ile Tyr Val Ala Arg
85 90 95

Leu Tyr Tyr His Glu Asp Glu Arg Asn Arg Gln Gly Leu Arg Asp Leu
100 105 110

Val Asn Ser Gly Val Thr Ile Arg Ile Met Asp Leu Pro Asp Tyr Asn
115 120 125

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

Trp Pro Gly His Phe Ala Pro Trp Ile Lys Gln Tyr Ser Leu Lys Leu
145 150 155 160

<210> 43

<211> 207
<212> PRT
<213> Artificial

<220>
<223> evolved cytidine deaminase

<400> 43

Thr Asp Ala Glu Tyr Val Arg Ile His Glu Lys Leu Asp Ile Tyr Thr
1 5 10 15

Phe Lys Lys Gln Phe Ser Asn Asn Lys Lys Ser Val Ser His Arg Cys
20 25 30

Tyr Val Leu Phe Glu Leu Lys Arg Arg Gly Glu Arg Arg Ala Cys Phe
35 40 45

Trp Gly Tyr Ala Val Asn Lys Pro Gln Ser Gly Thr Glu Arg Gly Ile
50 55 60

His Ala Glu Ile Phe Ser Ile Arg Lys Val Glu Glu Tyr Leu Arg Asp
65 70 75 80

Asn Pro Gly Gln Phe Thr Ile Asn Trp Tyr Ser Ser Trp Ser Pro Cys
85 90 95

Ala Asp Cys Ala Glu Lys Ile Leu Glu Trp Tyr Asn Gln Glu Leu Arg
100 105 110

Gly Asn Gly His Thr Leu Lys Ile Trp Val Cys Lys Leu Tyr Tyr Glu
115 120 125

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

Val Gly Leu Asn Val Met Val Ser Glu His Tyr Gln Cys Cys Arg Lys
145 150 155 160

Ile Phe Ile Gln Ser Ser His Asn Gln Leu Asn Glu Asn Arg Trp Leu
165 170 175

Glu Lys Thr Leu Lys Arg Ala Glu Lys Arg Arg Ser Glu Leu Ser Ile

```
            180                 185                 190

Met Phe Gln Val Lys Ile Leu His Thr Thr Lys Ser Pro Ala Val
        195                 200                 205


<210>  44
<211>  228
<212>  PRT
<213>  Artificial

<220>
<223>  evolved cytidine deaminase

<400>  44

Ser Ser Lys Thr Gly Pro Val Ala Val Asp Pro Thr Leu Arg Arg Arg
1               5                   10                  15


Ile Glu Pro His Glu Phe Glu Val Phe Phe Asp Pro Arg Glu Leu Arg
            20                  25                  30


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


Ile Trp Arg His Thr Ser Gln Asn Thr Asn Lys His Val Glu Val Asn
    50                  55                  60


Phe Ile Glu Lys Phe Thr Thr Glu Arg Tyr Phe Cys Pro Asn Thr Arg
65                  70                  75                  80


Cys Ser Ile Thr Trp Phe Leu Ser Trp Ser Pro Cys Gly Glu Cys Ser
                85                  90                  95


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


Ile Tyr Ile Ala Arg Leu Tyr His Leu Ala Asn Pro Arg Asn Arg Gln
        115                 120                 125


Gly Leu Arg Asp Leu Ile Ser Ser Gly Val Thr Ile Gln Ile Met Thr
    130                 135                 140


Glu Gln Glu Ser Gly Tyr Cys Trp His Asn Phe Val Asn Tyr Ser Pro
```

```
145                 150                 155                 160

Ser Asn Glu Ser His Trp Pro Arg Tyr Pro His Leu Trp Val Arg Leu
                165                 170                 175

Tyr Val Leu Glu Leu Tyr Cys Ile Ile Leu Gly Leu Pro Pro Cys Leu
            180                 185                 190

Asn Ile Leu Arg Arg Lys Gln Ser Gln Leu Thr Ser Phe Thr Ile Ala
        195                 200                 205

Leu Gln Ser Cys His Tyr Gln Arg Leu Pro Pro His Ile Leu Trp Ala
    210                 215                 220

Thr Gly Leu Lys
225
```

<210> 45
<211> 162
<212> PRT
<213> Artificial

<220>
<223> evolved cytidine deaminase

<400> 45

```
Ser Phe Glu Arg Asn Tyr Asp Pro Arg Glu Leu Arg Lys Glu Thr Tyr
1               5                   10                  15

Leu Leu Tyr Glu Ile Lys Trp Gly Lys Ser Gly Lys Leu Trp Arg His
            20                  25                  30

Trp Cys Gln Asn Asn Arg Thr Gln His Ala Glu Val Tyr Phe Leu Glu
        35                  40                  45

Asn Ile Phe Asn Ala Arg Arg Phe Asn Pro Ser Thr His Cys Ser Ile
    50                  55                  60

Thr Trp Tyr Leu Ser Trp Ser Pro Cys Ala Glu Cys Ser Gln Lys Ile
65                  70                  75                  80

Val Asp Phe Leu Lys Glu His Pro Asn Val Asn Leu Glu Ile Tyr Val
```

```
                85                  90                  95

Ala Arg Leu Tyr Tyr Pro Glu Asn Glu Arg Asn Arg Gln Gly Leu Arg
            100                 105                 110

Asp Leu Val Asn Ser Gly Val Thr Ile Arg Ile Met Asp Leu Pro Asp
        115                 120                 125

Tyr Asn Tyr Cys Trp Lys Thr Phe Val Ser Asp Gln Gly Gly Asp Glu
    130                 135                 140

Asp Tyr Trp Pro Gly His Phe Ala Pro Trp Ile Lys Gln Tyr Ser Leu
145                 150                 155                 160

Lys Leu
```

<210> 46
<211> 229
<212> PRT
<213> Rattus norvegicus

<400> 46

```
Met Ser Ser Glu Thr Gly Pro Val Ala Val Asp Pro Thr Leu Arg Arg
1               5                   10                  15

Arg Ile Glu Pro His Glu Phe Glu Val Phe Phe Asp Pro Arg Glu Leu
            20                  25                  30

Arg Lys Glu Thr Cys Leu Leu Tyr Glu Ile Asn Trp Gly Gly Arg His
        35                  40                  45

Ser Ile Trp Arg His Thr Ser Gln Asn Thr Asn Lys His Val Glu Val
    50                  55                  60

Asn Phe Ile Glu Lys Phe Thr Thr Glu Arg Tyr Phe Cys Pro Asn Thr
65                  70                  75                  80

Arg Cys Ser Ile Thr Trp Phe Leu Ser Trp Ser Pro Cys Gly Glu Cys
                85                  90                  95
```

```
Ser Arg Ala Ile Thr Glu Phe Leu Ser Arg Tyr Pro His Val Thr Leu
            100                 105                 110

Phe Ile Tyr Ile Ala Arg Leu Tyr His His Ala Asp Pro Arg Asn Arg
        115                 120                 125

Gln Gly Leu Arg Asp Leu Ile Ser Ser Gly Val Thr Ile Gln Ile Met
    130                 135                 140

Thr Glu Gln Glu Ser Gly Tyr Cys Trp Arg Asn Phe Val Asn Tyr Ser
145                 150                 155                 160

Pro Ser Asn Glu Ala His Trp Pro Arg Tyr Pro His Leu Trp Val Arg
                165                 170                 175

Leu Tyr Val Leu Glu Leu Tyr Cys Ile Ile Leu Gly Leu Pro Pro Cys
            180                 185                 190

Leu Asn Ile Leu Arg Arg Lys Gln Pro Gln Leu Thr Phe Phe Thr Ile
        195                 200                 205

Ala Leu Gln Ser Cys His Tyr Gln Arg Leu Pro Pro His Ile Leu Trp
    210                 215                 220

Ala Thr Gly Leu Lys
225


<210>  47
<211>  198
<212>  PRT
<213>  Homo sapiens

<400>  47

Met Asp Ser Leu Leu Met Asn Arg Arg Lys Phe Leu Tyr Gln Phe Lys
1               5                   10                  15

Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val
            20                  25                  30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr
        35                  40                  45
```

```
Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr
    50                  55                  60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp
65                  70                  75                  80

Phe Thr Ser Trp Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp
                85                  90                  95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg
            100                 105                 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg
        115                 120                 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr
    130                 135                 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Glu Arg Thr Phe Lys
145                 150                 155                 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu
                165                 170                 175

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

Phe Arg Thr Leu Gly Leu
        195
```

<210> 48
<211> 197
<212> PRT
<213> Artificial

<220>
<223> cytosine deaminase

<400> 48

```
Met Asp Ser Leu Leu Met Asn Arg Arg Glu Phe Leu Tyr Gln Phe Lys
1               5                   10                  15
```

```
Asn Val Arg Trp Ala Lys Gly Arg Arg Glu Thr Tyr Leu Cys Tyr Val
            20                  25                  30

Val Lys Arg Arg Asp Ser Ala Thr Ser Phe Ser Leu Asp Phe Gly Tyr
        35                  40                  45

Leu Arg Asn Lys Asn Gly Cys His Val Glu Leu Leu Phe Leu Arg Tyr
    50                  55                  60

Ile Ser Asp Trp Asp Leu Asp Pro Gly Arg Cys Tyr Arg Val Thr Trp
65                  70                  75                  80

Phe Ile Ser Trp Ser Pro Cys Tyr Asp Cys Ala Arg His Val Ala Asp
                85                  90                  95

Phe Leu Arg Gly Asn Pro Asn Leu Ser Leu Arg Ile Phe Thr Ala Arg
            100                 105                 110

Leu Tyr Phe Cys Glu Asp Arg Lys Ala Glu Pro Glu Gly Leu Arg Arg
        115                 120                 125

Leu His Arg Ala Gly Val Gln Ile Ala Ile Met Thr Phe Lys Asp Tyr
    130                 135                 140

Phe Tyr Cys Trp Asn Thr Phe Val Glu Asn His Gly Arg Thr Phe Lys
145                 150                 155                 160

Ala Trp Glu Gly Leu His Glu Asn Ser Val Arg Leu Ser Arg Gln Leu
                165                 170                 175

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

Phe Arg Thr Cys Thr
        195
```

<210> 49
<211> 3987
<212> DNA
<213> Artificial

<220>
<223> polynucleotide

<400> 49
atggccggga gcaagaagcg ccggataaag caggacacgc agttcgaggg cttcaccaac 60

ctgtaccaag tctccaagac gctccggttc gagcttatcc cgcaagggaa gaccctgaaa 120

cacatccagg aacaaggttt catcgaggag gacaaggccc gcaacgacca ctacaaggag 180

ctcaagccca taatcgatcg gatctacaag acgtacgccg accagtgcct ccaactggtg 240

cagctcgact gggagaacct gagcgccgcc attgacagct accgcaagga aaagacggag 300

gagacgcgca acgcccttat tgaggagcaa gccacctacc gcaacgccat ccacgactac 360

ttcatcgggc gcaccgacaa cctgacggac gcgatcaaca agcgccacgc ggaaatctac 420

aagggccttt tcaaggccga gctcttcaac gggaaggtcc taaaacagct cgggactgtc 480

acgacaaccg agcatgagaa cgccctcctt cgcagcttcg acaagttcac cacatacttc 540

tcgggcttct accggaaccg caagaacgtt ttcagcgccg aggacatctc caccgccatc 600

ccgcacagga tcgtccagga caacttcccc aagttcaagg agaactgcca catcttcacg 660

cgcctgatta cagccgtacc ttcacttcgt gagcacttcg agaacgtcaa aaaggccatc 720

gggatcttcg tctccacgtc catcgaggag gtattctctt tcccgttcta taaccagctc 780

ctgacccaga cgcagatcga cctctacaac cagctactgg gcggcatcag ccgggaggcc 840

gggaccgaga aaataaaggg cctcaacgaa gttctcaacc tggccatcca gaagaacgac 900

gagaccgcgc atatcatcgc atccctgccg catcgcttca ttcctttgtt caagcagata 960

ttgagcgacc ggaacaccct ctcgttcatc ctcgaagaat tcaagagcga cgaggaggtc 1020

attcagtctt tctgcaagta caagacgctc ctacggaatg agaatgtgct ggagaccgcg 1080

gaggcactct tcaatgagct gaactccatt gacctgaccc acatcttcat tagccacaag 1140

aaactggaga cgatctccag cgccctgtgc gaccactggg acactctccg caacgccctc 1200

tacgaacgcc ggatctccga acttaccggc aagataacta agtcggctaa ggagaaggtg 1260

caacggagcc tcaagcacga ggacatcaac cttcaggaaa tcatctcagc cgcgggcaag 1320

gagctgagcg aggcgtttaa gcagaaaaca tcggagatac tgagccacgc gcacgcggcc 1380

ctggatcaac cgctgccgac gactctcaag aagcaagagg agaaggaaat ccttaagtcc 1440

cagctcgact cgctgctcgg cctctatcac ttgctcgact ggttcgcggt tgatgagtcc 1500

aacgaggtgg acccggagtt ctccgcgcgc ctcacgggta ttaagctgga gatggagcca 1560

agcttaagct tctacaacaa ggcccgcaac tacgcgacca aaaaaccgta ctcagtcgag 1620

aaattcaagc tgaatttcca gatgcctaca ttggcgaggg ggtgggacgt gaaccgcgag 1680

aagaacaatg gagccatcct gttcgtcaaa aatgggttgt actacctggg catcatgccc 1740

aagcagaagg gccgttacaa ggccctgtca ttcgagccta ccgagaagac ctcggagggc 1800

ttcgacaaga tgtactacga ctatttcccg gacgccgcca agatgatccc gaagtgctcc 1860

acgcagctca aagccgtcac ggcccacttc cagacgcata ccacgccgat acttctgagc 1920

aacaacttca ttgagccgct agagatcacg aaggagatat acgacctaaa caaccccgaa 1980

aaggagccca agaagttcca gacagcctac gctaagaaga caggtgatca gaagggatat 2040

agggaggcac tctgcaagtg gatcgacttc acgcgcgact tcctgtcgaa atatacaaag 2100

acgaccagca ttgacctaag ttctctccgc ccatcctccc agtacaagga tctgggcgag 2160

tattatgcgg agctgaaccc attgctgtac cacatcagct tccagaggat cgccgagaag 2220

gagattatgg acgcggtgga gacggggaaa ctatacctgt tccaaatata taacaaggac 2280

ttcgctaaag ggcaccacgg gaagcccaac ctgcacacac tctactggac gggcttgttt 2340

tcgccagaaa atttggccaa gacttcgatc aagctcaacg gccaggcgga gttgttttac 2400

cgtcccaagt ctcgcatgaa gcgcatggcg catcgcctcg gagagaaaat gcttaacaag 2460

aagctcaagg atcagaagac gcccatacct gatacgttgt accaggaatt gtacgactac 2520

gtgaaccacc gcctatcgca cgacctctca gacgaggccc gcgccctcct cccaaacgtg 2580

attactaagg aggtttccca tgaaataatc aaggaccgac ggttcaccag cgacaaattt 2640

tttttccacg tgcctatcac gctcaattac caggcggcca actccccatc gaagttcaac 2700

cagcgcgtga acgcctacct taaggagcac ccggagaccc caatcatcgg gatcgaccgt 2760

ggcgagcgga acctgatcta tattacggtg atcgatagca ccgggaagat cctggagcag 2820

cgctccctga acacaatcca gcagtttgac taccagaaga aactcgacaa ccgggagaag 2880

gagcgcgtcg cagcccggca agcatggagt gtggtcggca ccataaagga cctgaaacag 2940

ggttacctaa gtcaagttat ccacgagatc gttgacctga tgatacacta tcaagccgta 3000

gtcgtgctgg agaacctcaa cttcgggttt aagtccaagc gcaccggcat cgcggagaag 3060

gcggtgtacc agcagttcga gaagatgctg atcgacaagc tgaactgcct ggtgctcaag 3120

gactaccctg cggagaaggt cggcggggtc ttgaacccgt accagctaac cgaccagttc 3180

acgagcttcg ccaaaatggg cacgcagtcc ggattcttgt tttatgtccc ggctccatat 3240

acaagtaaga tcgacccgct gacagggttt gttgacccat tcgtgtggaa gaccatcaag 3300

aaccacgaga gcaggaaaca cttcttagag ggcttcgact tcctgcatta cgacgttaag 3360

acaggcgact tcatcctgca cttcaagatg aaccgcaacc tgtcgttcca gaggggcctg 3420

cccggcttca tgcccgcctg ggatatcgtc tttgagaaga atgagacgca gttcgacgcg 3480

aaggggacgc cgttcatcgc tggaaagcgg atcgtgccgg tcatcgagaa ccaccgcttc 3540

acgggtcgct accgagattt ataccccgcc aacgaactaa ttgcgctgct ggaggagaag 3600

gggatcgtgt tccgagatgg cagcaacatt ctcccgaagc tgctggagaa cgacgactcg 3660

cacgctattg acacgatggt cgccctcata cggagcgtgc ttcagatgcg gaacagtaac 3720

gctgccacgg gcgaggacta cattaactcc cccgtccgcg acctcaacgg ggtctgcttc 3780

gatagccgct tccagaaccc ggagtggcct atggatgcgg acgcgaacgg ggcctaccac 3840

atcgccctca agggccaact cctgctcaac cacttgaagg aaagcaaaga cctcaaattg 3900

cagaatggca tcagtaacca ggactggctc gcgtacatcc aggaactgag aaacgggtcc 3960

aagaagcggc gtatcaagca agattga 3987

<210> 50
<211> 3987
<212> DNA
<213> Artificial

<220>
<223> polynucleotide

<400> 50
atggcgggaa gcaaaaagcg ccggattaag caagacacgc agttcgaggg cttcacgaac 60

ctctaccaag tcagcaagac cctccggttc gagctgatac cacagggaaa gacgctcaag 120

cacatccagg aacagggctt catcgaggag gacaaggcgc gcaacgacca ctacaaggag 180

ttgaaaccga tcatcgaccg catctacaag acgtacgccg accagtgcct ccagctcgtg 240

cagctcgact gggagaacct ctccgccgcc attgactcgt accggaagga gaagactgag 300

gagacccgca acgccctgat cgaggagcaa gcaacctacc ggaacgccat ccacgactac 360

ttcatcggcc gcaccgacaa cctcaccgac gcgatcaaca agcggcacgc ggagatatac 420

aaagggctgt tcaaggcgga gctgttcaac ggcaaggtgc tcaagcagct agggacggtg 480

accacgaccg agcacgagaa cgcgctcctc cgcagcttcg acaagttcac cacctacttc 540

agcggcttct accggaaccg caagaatgtg ttcagcgcgg aggacatcag cacggccatc 600

ccgcaccgca tcgtccagga caacttcccg aagttcaagg agaactgcca catcttcacc 660

cgcctgataa ccgccgtccc ctccctgcgg gagcacttcg agaacgtcaa aaaggcaatt 720

gggatcttcg tctcgaccag cattgaggag gtgttcagct tccccttcta caaccagctc 780

ctcacccaga cgcagatcga cctgtacaat cagttgctcg gcgggataag ccgcgaggcg 840

ggaaccgaaa aaatcaaggg gctgaacgaa gtgttgaacc tcgccatcca gaagaacgac 900

gagaccgcgc acatcatcgc ctccctgccc caccggttca tcccgctgtt caagcagatc 960

ctctctgacc ggaacaccct gtccttcatt cttgaggagt tcaagtcgga cgaggaggtc 1020

atccagagct tctgcaagta caagacgctg ctacggaacg agaacgtgct ggagacggcg 1080

gaggcactgt tcaacgagct aaacagcatc gacctcacgc acatcttcat cagtcacaag 1140

aaactggaga ccatctcctc cgcgctgtgc gaccactggg acacgctcag gaacgcgctc 1200

tacgagcgcc gaatcagtga gctgacgggc aagatcacga agtccgcgaa ggagaaggtg 1260

cagcggtccc tcaagcacga ggacatcaac ctccaggaga tcatctcagc ggctgggaaa 1320

gagctgtccg aggcgttcaa gcagaaaacg agcgaaatcc tgtcccacgc gcacgcggcc 1380

ctggatcagc ctctgccgac gaccctcaag aaacaagaag aaaaggaaat cctcaagtcg 1440

cagctcgact cgctgctggg cctgtaccat ctcctcgact ggttcgccgt ggacgagagc 1500

aacgaggtgg accccgagtt ctccgcgcgg cttacgggga tcaagctgga gatggagccc 1560

agcctgtcct tctacaacaa ggcgcgcaac tacgccacca agaagcccta cagcgtggag 1620

aagttcaagc tcaacttcca gatgcccact ctcgcacgtg ggtgggacgt caaccgcgaa 1680

aaaaataatg gggcgatcct gttcgtcaag aacggcctgt actacttggg catcatgccg 1740

aaacagaagg gccgctacaa ggccctgagc ttcgaaccga ccgagaaaac gagcgagggg 1800

ttcgacaaga tgtactacga ctacttcccc gacgccgcga agatgattcc aaagtgctcc 1860

acgcagctta aggccgtgac ggcccacttc cagacgcaca cgaccccgat cctcctcagc 1920

aacaacttca tcgagcccct ggagatcacg aaggagatat acgacctgaa caacccggag 1980

aaggagccca agaaattcca gaccgcctac gccaagaaga caggcgacca aaagggttac 2040

agggaggccc tctgcaagtg gatcgacttc actagggact tcctgtccaa gtacaccaag 2100

actacctcta tcgacctgtc cagcctccgc ccgtcgtccc agtacaaaga tttgggcgag 2160

tattacgcgg agctgaaccc actgctctac cacatcagct tccagcgcat cgcggagaag 2220

gagatcatgg acgcagtgga gacgggcaag ctatacctat ttcagatata caacaaagac 2280

ttcgctaagg gacaccacgg caagcctaac ctgcacaccc tctactggac ggggctcttc 2340

agcccggaga acctcgccaa gacctcgatc aagctcaacg gccaggccga gctgttctac 2400

cggcccaagt cccgcatgaa gcggatggcc caccggctcg gggagaaaat gctcaacaag 2460

aaattgaagg accaaaaaac gccgataccc gacaccctat accaggagct gtacgactat 2520

gtgaaccacc gcctgagcca cgacctcagc gacgaggcgc gggccctcct gccgaacgtc 2580

atcacaaagg aggtcagcca cgagatcatc aaggaccggc gcttcacctc cgacaagttt 2640

ttctttcacg tgcccatcac gctcaactac caggccgcca actcgccgtc caagttcaac 2700

cagcgcgtga acgcctacct caaggagcac cccgagaccc cgatcatcgg gattgaccga 2760

ggggagcgga acctcatcta catcaccgtc atcgacagca ccgggaagat ccttgaacag 2820

cggtcgctca acaccatcca gcagttcgac taccagaaga aactcgacaa ccgggagaag 2880

gagagagtgg cggcccgcca ggcttggtcc gtcgtcggga cgattaagga cttgaaacaa 2940

ggttacctgt cgcaagtgat ccacgagatc gttgacctga tgatccacta ccaagccgtc 3000

gtggtcctgg agaacctcaa cttcggcttc aagagcaaac gaaccggcat cgcggagaag 3060

gccgtgtacc agcagttcga aaaaatgctg atcgacaagc tgaactgcct cgtgctcaag 3120

gactaccccg ctgagaaggt cggcggggtg ctgaacccgt accagctcac tgaccagttc 3180

accagcttcg caaagatggg cacccagtcc ggcttcctgt tctacgtgcc tgcgccatac 3240

acctcgaaga tcgacccgct caccgggttc gtggacccct tcgtctggaa gaccatcaag 3300

aaccacgaga gccgcaagca cttcctggag ggcttcgact tcctccacta cgacgtcaag 3360

accggggact tcatcctgca cttcaagatg aaccgcaacc tcagtttcca gcgcggcctg 3420

ccggggttca tgcccgcttg ggatatagtc ttcgagaaga atgagacgca gttcgacgcg 3480

aagggcaccc cgttcatcgc cgggaagcgc atcgtgccgg tcatcgagaa ccaccggttc 3540

accgggcgct accgcgacct atacccggcg aacgagttga tcgccctcct ggaggagaag 3600

ggcatcgtgt tccgcgacgg ctccaacatc ctcccgaagc tgctcgaaaa cgacgactcc 3660

cacgccatcg acacgatggt cgcgctgatc cggtcggtgc tccagatgcg gaactccaac 3720

gccgcgacgg gcgaggacta catcaacagt ccggtccgcg atctgaacgg cgtctgcttc 3780

gactcccggt tccagaaccc cgagtggccg atggacgcgg acgcgaacgg cgcataccac 3840

atcgccctaa aagggcaatt gctgctcaac cacctcaagg aatccaaaga cctaaagctc 3900

cagaacggca tctccaacca ggactggctg gcgtacatcc aggaactgcg gaacgggagc 3960

aaaaaacgtc ggatcaagca agattga 3987

<210> 51
<211> 3987
<212> DNA
<213> Artificial

<220>
<223> polynucleotide

<400> 51
atggcgggct ccaagaaacg ccggattaag caagataccc agttcgaggg gttcacgaac 60

ctctaccaag tgagcaagac cctccgattc gaactgattc ctcaggggaa gaccctcaag 120

cacatccagg agcaagggtt catcgaggag gacaaggcgc ggaacgacca ctacaaggaa 180

ctcaaaccca tcatcgaccg catctacaag acctacgccg atcagtgcct ccagctcgtg 240

cagttggact gggagaacct cagcgcggcc attgactcct accggaagga gaaaacggag 300

gagacgcgca acgcgctcat cgaggaacag gcaacctatc gcaacgccat ccacgactac 360

ttcatcggga ggactgacaa cctcactgac gcgattaaca agcgccacgc ggagatatac 420

aagggactct tcaaagcgga gctgtttaac ggcaaggttc tcaagcaact cggcactgtg 480

accacgaccg agcatgagaa cgccctgctc cgctccttcg acaagttcac cacctacttc 540

tccgggttct accgcaaccg caagaatgtc ttcagcgcgg aggacatcag cacggccatt 600

ccacatcgaa tcgtccaaga taacttcccg aagttcaagg agaactgcca catcttcacc 660

cgactcatta ctgctgtacc gtcgttacgc gaacacttcg agaacgtcaa gaaggcaatt 720

ggaatcttcg tctctacgtc aatagaggag gtgttcagct tccctttcta caaccagctc 780

cttacgcaga cccagataga cctgtacaat cagctcctcg gtgggatcag ccgggaggcg 840

gggactgaga agattaaagg gctcaacgag gtcttgaacc tggccatcca aaaaaacgat 900

gagacggcgc acatcatcgc ctcgctgccc caccggttca tcccgctgtt caagcagatc 960

ctcagtgaca ggaacacctt gagctttatc ctagaggagt tcaagagcga cgaggaggtg 1020

atccagagct tctgcaagta caaaaccctg ctgaggaacg agaacgtcct ggagacggcg 1080

gaggcgctgt tcaacgagct gaactctatc gacttaactc acatattcat ctcgcacaag 1140

aagctggaga ctattagctc tgcactctgc gaccactggg acaccctccg caacgcgctc 1200

tacgagcgcc gcatctcgga gctgaccggg aagatcacca aatccgcgaa ggaaaaggtc 1260

cagcgttccc tcaaacacga ggatattaac ttacaggaga ttatctcagc ggctgggaag 1320

gagttgtcag aggcgttcaa gcagaaaact tccgagatcc tgagccacgc gcacgcagcg 1380

ctcgaccagc ctctgcccac caccctcaaa aagcaggaag aaaaagagat cctcaagagc 1440

cagttggact ccctgctggg gctctatcac cttctcgact ggttcgccgt cgatgagtcg 1500

aacgaggtgg accccgagtt ctccgcccgg ctgaccggca tcaagctaga gatggagccg 1560

tccctcagct tctacaataa ggcccgcaac tacgcgacca aaaaacccta cagcgtggag 1620

aagttcaagc tgaacttcca gatgccgacc ttagcacgcg gttgggacgt aaacagggag 1680

aagaacaatg gagccatcct gttcgtcaag aacgggcttt actacctcgg gataatgccc 1740

aagcagaagg gccgctacaa ggccctttcc ttcgagccga cggagaaaac ctccgagggg 1800

ttcgacaaga tgtactacga ctacttcccc gacgccgcca agatgatccc gaagtgctca 1860

acgcagctaa aagccgtgac cgcccacttc cagacccaca cgacgccgat cctgctgagc 1920

aacaacttca tcgagcccct tgagatcact aaggagatat acgacctgaa caaccccgag 1980

aaggagccca agaagtttca aaccgcctac gccaaaaaaa ctggcgacca aaagggctac 2040

agggaggcgc tgtgtaagtg gatcgacttc acacgcgact tcctttcgaa gtatacgaag 2100

acaacctcta ttgacctgag cagcctgcgt cctagctccc agtacaaaga tttgggcgag 2160

tactacgcgg agcttaatcc actactctac cacatctcat tccagcgcat cgctgagaag 2220

gaaatcatgg acgcggtgga gacaggcaaa ctgtacctct tccagatata caacaaagac 2280

ttcgctaagg ggcaccacgg gaagcccaac cttcatacgc tctactggac gggcctattc 2340

agccccgaaa atctggccaa gacctccatc aagctgaacg gccaagcgga gctgttctac 2400

agacccaaga gccggatgaa gcggatggcc cacaggctcg gcgagaaaat gcttaacaaa 2460

aagttgaagg accagaaaac ccctatcccc gacaccctct accaggaact gtacgactac 2520

gtgaaccaca ggctctcgca cgacctttcc gacgaggccc gtgccctact cccgaacgtc 2580

attaccaaag aggtttcgca cgagatcatc aaggaccggc ggttcacgag cgacaagttt 2640

ttctttcacg tccccatcac ccttaactac caggcggcca actccccatc caagttcaac 2700

cagcgtgtga atgcctacct caaggagcac ccagagaccc cgatcattgg gatcgaccgg 2760

ggcgagcgga acctgatcta catcaccgtc atcgactcga cgggcaagat tcttgagcag 2820

agatcgttga ataccataca gcagttcgac taccagaaga aactcgacaa ccgcgagaag 2880

gagcgcgtgg cggcccgcca ggcgtggtcc gtcgttggga cgattaagga cttgaaacaa 2940

ggttatctgt cccaagtcat ccacgagatc gttgatctga tgatccacta tcaggcagtg 3000

gtggtgctgg agaatctcaa cttcggcttc aagagtaagc ggacgggaat cgccgagaag 3060

gccgtgtacc agcagttcga gaagatgctg atcgacaagc tcaactgcct tgtgctgaaa 3120

gactacccgg ccgagaaggt cggcggcgtc ctcaacccgt accaacttac cgaccagttc 3180

acctccttcg ccaagatggg cactcagtcc gggttcttgt tctacgtccc cgcaccttac 3240

acctctaaga tcgaccctct gactggcttc gtagatccat tcgtgtggaa gaccattaag 3300

aaccacgaga gccgcaagca cttcctggag ggcttcgact tcctgcacta cgacgtgaag 3360

accggggact tcatccttca cttcaagatg aaccggaacc tcagcttcca gcggggcctg 3420

ccggggttca tgcccgcctg ggacatcgtg ttcgagaaga acgagaccca gttcgacgcg 3480

aagggcacgc ccttcatcgc cgggaagcgt atcgtgccgg tgatcgagaa ccatcgtttc 3540

acgggtcgct accgtgacct ctacccggcg aacgagctta tcgcactcct ggaggagaag 3600

ggcatcgtct tccgggacgg ctccaacatc ctcccgaaac tgctggaaaa cgacgactct 3660

cacgccatcg acacgatggt ggccctcatc cggtccgtgc tccaaatgcg gaacagcaac 3720

gccgccaccg gtgaggacta catcaacagc ccggtccggg atctgaacgg ggtgtgcttc 3780

gattcgcggt tccagaatcc tgagtggccg atggacgcgg atgcaaacgg ggcgtaccac 3840

atcgcgctca agggccagtt acttctgaac caccttaagg agtctaaaga tttgaaactc 3900

cagaacggga tctcgaacca ggactggctg gcctacatcc aagagttgcg gaacggcagc 3960

aagaagcggc ggattaagca agattag 3987

<210> 52
<211> 3790
<212> DNA
<213> Artificial

<220>
<223> expression cassette

<400> 52
ccatgggcag caaactggaa aaatttacga attgttatag cctgtccaag accctgcgtt 60

tcaaagccat ccccgttggc aaaacccagg agaatattga taataaacgt ctgctggttg 120

aggatgaaaa aagagcagaa gactataagg gagtcaaaaa actgctggat cggtactacc 180

tgagctttat aaatgacgtg ctgcatagca ttaaactgaa aaatctgaat aactatatta 240

gtctgttccg caagaaaacc cgaacagaga aagaaaataa agagctggaa aacctggaga 300

tcaatctgcg taaagagatc gcaaaagctt ttaaaggaaa tgaaggttat aaaagcctgt 360

tcaaaaaaga cattattgaa accatcctgc cggaatttct ggatgataaa gacgagatag 420

cgctcgtgaa cagcttcaac gggttcacga ccgccttcac gggcttttтc gataacaggg 480

aaaatatgtt ttcagaggaa gccaaaagca cctcgatagc gttccgttgc attaatgaaa 540

atttgacaag atatatcagc aacatggata ttttcgagaa agttgatgcg atctttgaca 600

aacatgaagt gcaggagatt aaggaaaaaa ttctgaacag cgattatgat gttgaggatt 660

ttttcgaggg ggaatttttt aactttgtac tgacacagga aggtatagat gtgtataatg 720

ctattatcgg cgggttcgtt accgaatccg gcgagaaaat taagggtctg aatgagtaca 780

tcaatctgta taaccaaaag accaaacaga aactgccaaa attcaaaccg ctgtacaagc 840

aagtcctgag cgatcgggaa agcttgagct tttacggtga aggttatacc agcgacgagg 900

aggtactgga ggtctttcgc aataccctga acaagaacag cgaaattttc agctccatta 960

aaaagctgga gaaactgttt aagaattttg acgagtacag cagcgcaggt atttttgtga 1020

agaacggacc tgccataagc accattagca aggatatttt tggagagtgg aatgttatcc 1080

gtgataaatg gaacgcggaa tatgatgaca tacacctgaa aaagaaggct gtggtaactg 1140

agaaatatga agacgatcgc cgcaaaagct ttaaaaaaat cggcagcttt agcctggagc 1200

agctgcagga atatgcggac gccgacctga gcgtggtcga gaaactgaag gaaattatta 1260

tccaaaaagt ggatgagatt tacaaggtat atggtagcag cgaaaaactg tttgatgcgg 1320

acttcgttct ggaaaaaagc ctgaaaaaaa atgatgctgt tgttgcgatc atgaaagacc 1380

tgctcgatag cgttaagagc tttgaaaatt acattaaagc attctttggc gagggcaaag 1440

aaacaaacag agacgaaagc ttttatggcg acttcgtcct ggcttatgac atcctgttga 1500

aggtagatca tatatatgat gcaattcgta attacgtaac ccaaaagccg tacagcaaag 1560

ataagttcaa actgtatttc cagaacccgc agtttatggg tggctgggac aaagacaagg 1620

agacagacta tcgcgccact attctgcgtt acggcagcaa gtactatctc gccatcatgg 1680

acaaaaaata tgcaaagtgt ctgcagaaaa tcgataaaga cgacgtgaac ggaaattacg 1740

aaaagattaa ttataagctg ctgccagggc ccaacaagat gttaccgaaa gtatttttt 1800

ccaaaaaatg gatggcatac tataacccga gcgaggatat acagaagatt tacaaaaatg 1860

ggaccttcaa aaaggggat atgttcaatc tgaatgactg ccacaaactg atcgattttt 1920

ttaaagatag catcagccgt tatcctaaat ggtcaaacgc gtatgatttt aatttctccg 1980

aaacggagaa atataaagac attgctggtt tctatcgcga agtcgaagaa cagggttata 2040

aagttagctt tgaatcggcc agcaagaaag aggttgataa actggtggag gagggtaagc 2100

tgtatatgtt tcagatttat aacaaagact ttagcgacaa aagccacggt actcctaatc 2160

tgcatacgat gtactttaaa ctgctgtttg atgagaataa ccacggccaa atccgtctct 2220

ccggtggagc agaacttttt atgcggcgtg cgagcctaaa aaaggaagaa ctggtggtgc 2280

atcccgccaa cagcccgatt gctaacaaaa atccagataa tcctaagaag accaccacac 2340

tgtcgtacga tgtctataag gataaacgtt tctcggaaga ccagtatgaa ttgcatatac 2400

cgatagcaat taataaatgc ccaaaaaaca ttttcaaaat caacactgaa gttcgtgtgc 2460

tgctgaaaca tgatgataat ccgtatgtga tcggaattga ccgtggggag agaaatctgc 2520

tgtatattgt agtcgttgat ggcaagggca acatcgttga gcagtatagc ctgaatgaaa 2580

taattaataa ttttaacggt atacgtatta aaaccgacta tcatagcctg ctggataaaa 2640

aggagaaaga gcgttttgag gcacgccaaa attggacgag catcgaaaac atcaaggaac 2700

tgaaggcagg atatatcagc caagtagtcc ataaaatctg tgaactggtg gagaagtacg 2760

acgctgtcat tgccctggaa gacctcaata gcggctttaa aaacagccgg gtgaaggtgg 2820

agaaacaggt ataccaaaag tttgaaaaga tgctcattga taagctgaac tatatggttg 2880

ataaaaagag caacccgtgc gccactggcg gtgcactgaa agggtaccaa attaccaata 2940

aatttgaaag ctttaaaagc atgagcacgc agaatgggtt tatttttat ataccagcat 3000

ggctgacgag caagattgac cccagcactg gttttgtcaa tctgctgaaa accaaataca 3060

caagcattgc ggatagcaaa aaatttattt cgagcttcga ccgtattatg tatgttccgg 3120

aggaagatct gtttgaattt gccctggatt ataaaaactt cagccgcacc gatgcagatt 3180

atatcaaaaa atggaagctg tacagttatg gtaatcgtat acgtatcttc cgtaatccga 3240

agaaaaacaa tgtgttcgat tgggaagagg tctgtctgac cagcgcgtat aaagaactgt 3300

tcaacaagta cggaataaat tatcagcaag gtgacattcg cgcactgctg tgtgaacagt 3360

cagataaagc attttatagc agctttatgg cgctgatgag cctgatgctc cagatgcgca 3420

acagcataac cggtcgcaca gatgttgact ttctgatcag ccctgtgaag aatagcgacg 3480

gcatcttcta cgattccagg aactatgaag cacaggaaaa cgctattctg cctaaaaatg 3540

ccgatgccaa cggcgcctat aatattgcac ggaaggttct gtgggcgatt ggacagttca 3600

agaaagcgga agatgagaag ctggataagg taaaaattgc tattagcaat aaggaatggc 3660

tggagtacgc acagacatcg gttaaacacg gtagtaaaag gccggcggcc acgaaaaagg 3720

ccggccaggc aaaaaagaaa aagggagcgg ccgcactcga gcaccaccac caccaccact 3780

gagcggccgc 3790

<210> 53
<211> 8986
<212> DNA
<213> Artificial

<220>
<223> vector

<400> 53
tggcgaatgg gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg 60

cagcgtgacc gctacacttg ccagcgccct agcgcccgct cctttcgctt tcttcccttc 120

ctttctcgcc acgttcgccg gctttccccg tcaagctcta aatcgggggc tccctttagg 180

gttccgattt agtgctttac ggcacctcga ccccaaaaaa cttgattagg gtgatggttc 240

acgtagtggg ccatcgccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt 300

ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggtctattc 360

ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg agctgattta 420

acaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaatttcag gtggcacttt 480

tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta 540

tccgctcatg aattaattct tagaaaaact catcgagcat caaatgaaac tgcaatttat 600

tcatatcagg attatcaata ccatattttt gaaaaagccg tttctgtaat gaaggagaaa 660

actcaccgag gcagttccat aggatggcaa gatcctggta tcggtctgcg attccgactc 720

gtccaacatc aatacaacct attaatttcc cctcgtcaaa aataaggtta tcaagtgaga 780

aatcaccatg agtgacgact gaatccggtg agaatggcaa aagtttatgc atttctttcc 840

agacttgttc aacaggccag ccattacgct cgtcatcaaa atcactcgca tcaaccaaac 900

cgttattcat tcgtgattgc gcctgagcga gacgaaatac gcgatcgctg ttaaaaggac 960

aattacaaac aggaatcgaa tgcaaccggc gcaggaacac tgccagcgca tcaacaatat 1020

tttcacctga atcaggatat tcttctaata cctggaatgc tgttttcccg gggatcgcag 1080

tggtgagtaa ccatgcatca tcaggagtac ggataaaatg cttgatggtc ggaagaggca 1140

taaattccgt cagccagttt agtctgacca tctcatctgt aacatcattg gcaacgctac 1200

ctttgccatg tttcagaaac aactctggcg catcgggctt cccatacaat cgatagattg 1260

tcgcacctga ttgcccgaca ttatcgcgag cccatttata cccatataaa tcagcatcca 1320

tgttggaatt taatcgcggc ctagagcaag acgtttcccg ttgaatatgg ctcataacac 1380

cccttgtatt actgtttatg taagcagaca gttttattgt tcatgaccaa aatcccttaa 1440

cgtgagtttt cgttccactg agcgtcagac cccgtagaaa agatcaaagg atcttcttga 1500

gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa aaaaaccacc gctaccagcg 1560

gtggtttgtt tgccggatca agagctacca actcttttc cgaaggtaac tggcttcagc 1620

agagcgcaga taccaaatac tgtccttcta gtgtagccgt agttaggcca ccacttcaag 1680

aactctgtag caccgcctac atacctcgct ctgctaatcc tgttaccagt ggctgctgcc 1740

agtggcgata agtcgtgtct taccgggttg gactcaagac gatagttacc ggataaggcg 1800

cagcggtcgg gctgaacggg gggttcgtgc acacagccca gcttggagcg aacgacctac 1860

accgaactga gatacctaca gcgtgagcta tgagaaagcg ccacgcttcc cgaagggaga 1920

aaggcggaca ggtatccggt aagcggcagg gtcggaacag gagagcgcac gagggagctt 1980

ccaggggaa acgcctggta tctttatagt cctgtcgggt ttcgccacct ctgacttgag 2040

cgtcgatttt tgtgatgctc gtcaggggg cggagcctat ggaaaaacgc cagcaacgcg 2100

gccttttac ggttcctggc cttttgctgg ccttttgctc acatgttctt tcctgcgtta 2160

tcccctgatt ctgtggataa ccgtattacc gcctttgagt gagctgatac cgctcgccgc 2220

agccgaacga ccgagcgcag cgagtcagtg agcgaggaag cggaagagcg cctgatgcgg 2280

tattttctcc ttacgcatct gtgcggtatt tcacaccgca tatatggtgc actctcagta 2340

caatctgctc tgatgccgca tagttaagcc agtatacact ccgctatcgc tacgtgactg 2400

ggtcatggct gcgccccgac acccgccaac acccgctgac gcgccctgac gggcttgtct 2460

gctcccggca tccgcttaca gacaagctgt gaccgtctcc gggagctgca tgtgtcagag 2520

gttttcaccg tcatcaccga aacgcgcgag gcagctgcgg taaagctcat cagcgtggtc 2580

gtgaagcgat tcacagatgt ctgcctgttc atccgcgtcc agctcgttga gtttctccag 2640

aagcgttaat gtctggcttc tgataaagcg ggccatgtta agggcggttt tttcctgttt 2700

ggtcactgat gcctccgtgt aaggggggatt tctgttcatg ggggtaatga taccgatgaa 2760

acgagagagg atgctcacga tacgggttac tgatgatgaa catgcccggt tactggaacg 2820

ttgtgagggt aaacaactgg cggtatggat gcggcgggac cagagaaaaa tcactcaggg 2880

tcaatgccag cgcttcgtta atacagatgt aggtgttcca cagggtagcc agcagcatcc 2940

tgcgatgcag atccggaaca taatggtgca gggcgctgac ttccgcgttt ccagacttta 3000

cgaaacacgg aaaccgaaga ccattcatgt tgttgctcag gtcgcagacg ttttgcagca 3060

gcagtcgctt cacgttcgct cgcgtatcgg tgattcattc tgctaaccag taaggcaacc 3120

ccgccagcct agccgggtcc tcaacgacag gagcacgatc atgcgcaccc gtggggccgc 3180

catgccggcg ataatggcct gcttctcgcc gaaacgtttg gtggcgggac cagtgacgaa 3240

ggcttgagcg agggcgtgca agattccgaa taccgcaagc gacaggccga tcatcgtcgc 3300

gctccagcga aagcggtcct cgccgaaaat gacccagagc gctgccggca cctgtcctac 3360

gagttgcatg ataaagaaga cagtcataag tgcggcgacg atagtcatgc cccgcgccca 3420

ccggaaggag ctgactgggt tgaaggctct caagggcatc ggtcgagatc ccggtgccta 3480

atgagtgagc taacttacat taattgcgtt gcgctcactg cccgctttcc agtcgggaaa 3540

cctgtcgtgc cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat 3600

tgggcgccag ggtggttttt cttttcacca gtgagacggg caacagctga ttgcccttca 3660

ccgcctggcc ctgagagagt tgcagcaagc ggtccacgct ggtttgcccc agcaggcgaa 3720

aatcctgttt gatggtggtt aacggcggga tataacatga gctgtcttcg gtatcgtcgt 3780

atcccactac cgagatatcc gcaccaacgc gcagcccgga ctcggtaatg gcgcgcattg 3840

cgcccagcgc catctgatcg ttggcaacca gcatcgcagt gggaacgatg ccctcattca 3900

gcatttgcat ggtttgttga aaaccggaca tggcactcca gtcgccttcc cgttccgcta 3960

tcggctgaat ttgattgcga gtgagatatt tatgccagcc agccagacgc agacgcgccg 4020

agacagaact taatgggccc gctaacagcg cgatttgctg gtgacccaat gcgaccagat 4080

gctccacgcc cagtcgcgta ccgtcttcat gggagaaaat aatactgttg atgggtgtct 4140

ggtcagagac atcaagaaat aacgccggaa cattagtgca ggcagcttcc acagcaatgg 4200

catcctggtc atccagcgga tagttaatga tcagcccact gacgcgttgc gcgagaagat 4260

tgtgcaccgc cgctttacag gcttcgacgc cgcttcgttc taccatcgac accaccacgc 4320

tggcacccag ttgatcggcg cgagatttaa tcgccgcgac aatttgcgac ggcgcgtgca 4380

gggccagact ggaggtggca acgccaatca gcaacgactg tttgcccgcc agttgttgtg 4440

ccacgcggtt gggaatgtaa ttcagctccg ccatcgccgc ttccactttt tcccgcgttt 4500

tcgcagaaac gtggctggcc tggttcacca cgcgggaaac ggtctgataa gagacaccgg 4560

catactctgc gacatcgtat aacgttactg gtttcacatt caccaccctg aattgactct 4620

cttccgggcg ctatcatgcc ataccgcgaa aggttttgcg ccattcgatg gtgtccggga 4680

tctcgacgct ctcccttatg cgactcctgc attaggaagc agcccagtag taggttgagg 4740

ccgttgagca ccgccgccgc aaggaatggt gcatgcaagg agatggcgcc caacagtccc 4800

ccggccacgg ggcctgccac catacccacg ccgaaacaag cgctcatgag cccgaagtgg 4860

cgagcccgat cttccccatc ggtgatgtcg gcgatatagg cgccagcaac cgcacctgtg 4920

gcgccggtga tgccggccac gatgcgtccg gcgtagagga tcgagatctc gatcccgcga 4980

aattaatacg actcactata ggggaattgt gagcggataa caattcccct ctagaaataa 5040

ttttgtttaa ctttaagaag gagatatacc atgggcagca aactggaaaa atttacgaat 5100

tgttatagcc tgtccaagac cctgcgtttc aaagccatcc ccgttggcaa aacccaggag 5160

aatattgata ataaacgtct gctggttgag gatgaaaaaa gagcagaaga ctataaggga 5220

gtcaaaaaac tgctggatcg gtactacctg agctttataa atgacgtgct gcatagcatt 5280

aaactgaaaa atctgaataa ctatattagt ctgttccgca agaaaacccg aacagagaaa 5340

gaaaataaag agctggaaaa cctggagatc aatctgcgta aagagatcgc aaaagctttt 5400

aaaggaaatg aaggttataa aagcctgttc aaaaaagaca ttattgaaac catcctgccg 5460

gaatttctgg atgataaaga cgagatagcg ctcgtgaaca gcttcaacgg gttcacgacc 5520

gccttcacgg gctttttcga taacagggaa aatatgtttt cagaggaagc caaaagcacc 5580

tcgatagcgt tccgttgcat taatgaaaat ttgacaagat atatcagcaa catggatatt 5640

ttcgagaaag ttgatgcgat ctttgacaaa catgaagtgc aggagattaa ggaaaaaatt 5700

ctgaacagcg attatgatgt tgaggatttt ttcgaggggg aatttttaa ctttgtactg 5760

acacaggaag gtatagatgt gtataatgct attatcggcg ggttcgttac cgaatccggc 5820

gagaaaatta agggtctgaa tgagtacatc aatctgtata accaaaagac caaacagaaa 5880

ctgccaaaat tcaaaccgct gtacaagcaa gtcctgagcg atcgggaaag cttgagcttt 5940

tacggtgaag gttataccag cgacgaggag gtactggagg tctttcgcaa taccctgaac 6000

aagaacagcg aaattttcag ctccattaaa aagctggaga aactgtttaa gaattttgac 6060

gagtacagca gcgcaggtat ttttgtgaag aacggacctg ccataagcac cattagcaag 6120

gatatttttg gagagtggaa tgttatccgt gataaatgga acgcggaata tgatgacata 6180

cacctgaaaa agaaggctgt ggtaactgag aaatatgaag acgatcgccg caaaagcttt 6240

aaaaaaatcg gcagctttag cctggagcag ctgcaggaat atgcggacgc cgacctgagc 6300

gtggtcgaga aactgaagga aattattatc caaaaagtgg atgagattta caaggtatat 6360

ggtagcagcg aaaaactgtt tgatgcggac ttcgttctgg aaaaaagcct gaaaaaaaat 6420

gatgctgttg ttgcgatcat gaaagacctg ctcgatagcg ttaagagctt tgaaaattac 6480

attaaagcat tctttggcga gggcaaagaa acaaacagag acgaaagctt ttatggcgac 6540

ttcgtcctgg cttatgacat cctgttgaag gtagatcata tatatgatgc aattcgtaat 6600

tacgtaaccc aaaagccgta cagcaaagat aagttcaaac tgtatttcca gaacccgcag 6660

tttatgggtg gctgggacaa agacaaggag acagactatc gcgccactat tctgcgttac 6720

ggcagcaagt actatctcgc catcatggac aaaaaatatg caaagtgtct gcagaaaatc 6780

gataaagacg acgtgaacgg aaattacgaa aagattaatt ataagctgct gccagggccc 6840

aacaagatgt taccgaaagt attttttttcc aaaaaatgga tggcatacta taacccgagc 6900

gaggatatac agaagattta caaaaatggg accttcaaaa agggggatat gttcaatctg 6960

aatgactgcc acaaactgat cgatttttt aaagatagca tcagccgtta tcctaaatgg 7020

tcaaacgcgt atgattttaa tttctccgaa acggagaaat ataaagacat tgctggtttc 7080

tatcgcgaag tcgaagaaca gggttataaa gttagctttg aatcggccag caagaaagag 7140

gttgataaac tggtggagga gggtaagctg tatatgtttc agatttataa caaagacttt 7200

agcgacaaaa gccacggtac tcctaatctg catacgatgt actttaaact gctgtttgat 7260

gagaataacc acggccaaat ccgtctctcc ggtggagcag aacttttat gcggcgtgcg 7320

agcctaaaaa aggaagaact ggtggtgcat cccgccaaca gcccgattgc taacaaaaat 7380

ccagataatc ctaagaagac caccacactg tcgtacgatg tctataagga taaacgtttc 7440

tcggaagacc agtatgaatt gcatataccg atagcaatta ataaatgccc aaaaaacatt 7500

ttcaaaatca acactgaagt tcgtgtgctg ctgaaacatg atgataatcc gtatgtgatc 7560

ggaattgacc gtggggagag aaatctgctg tatattgtag tcgttgatgg caagggcaac 7620

atcgttgagc agtatagcct gaatgaaata attaataatt ttaacggtat acgtattaaa 7680

accgactatc atagcctgct ggataaaaag gagaaagagc gttttgaggc acgccaaaat 7740

tggacgagca tcgaaaacat caaggaactg aaggcaggat atatcagcca agtagtccat 7800

aaaatctgtg aactggtgga gaagtacgac gctgtcattg ccctggaaga cctcaatagc 7860

ggctttaaaa acagccgggt gaaggtggag aaacaggtat accaaaagtt tgaaaagatg 7920

ctcattgata agctgaacta tatggttgat aaaaagagca acccgtgcgc cactggcggt 7980

gcactgaaag ggtaccaaat taccaataaa tttgaaagct ttaaaagcat gagcacgcag 8040

aatgggttta ttttttatat accagcatgg ctgacgagca agattgaccc cagcactggt 8100

tttgtcaatc tgctgaaaac caaatacaca agcattgcgg atagcaaaaa atttatttcg 8160

agcttcgacc gtattatgta tgttccggag gaagatctgt ttgaatttgc cctggattat 8220

aaaaacttca gccgcaccga tgcagattat atcaaaaaat ggaagctgta cagttatggt 8280

aatcgtatac gtatcttccg taatccgaag aaaaacaatg tgttcgattg ggaagaggtc 8340

tgtctgacca gcgcgtataa agaactgttc aacaagtacg gaataaatta tcagcaaggt 8400

gacattcgcg cactgctgtg tgaacagtca gataaagcat tttatagcag ctttatggcg 8460

ctgatgagcc tgatgctcca gatgcgcaac agcataaccg gtcgcacaga tgttgacttt 8520

ctgatcagcc ctgtgaagaa tagcgacggc atcttctacg attccaggaa ctatgaagca 8580

caggaaaacg ctattctgcc taaaaatgcc gatgccaacg gcgcctataa tattgcacgg 8640

aaggttctgt gggcgattgg acagttcaag aaagcggaag atgagaagct ggataaggta 8700

aaaattgcta ttagcaataa ggaatggctg gagtacgcac agacatcggt taaacacggt 8760

agtaaaaggc cggcggccac gaaaaaggcc ggccaggcaa aaaagaaaaa gggagcggcc 8820

gcactcgagc accaccacca ccaccactga gatccggctg ctaacaaagc ccgaaaggaa 8880

gctgagttgg ctgctgccac cgctgagcaa taactagcat aacccccttgg ggcctctaaa 8940

cgggtcttga ggggtttttt gctgaaagga ggaactatat ccggat 8986

<210> 54
<211> 23
<212> DNA
<213> Artificial

<220>
<223> spacer sequence

<400> 54
tcatctgtgc ccctccctcc ctg 23

<210> 55
<211> 23
<212> DNA
<213> Artificial

<220>
<223> spacer sequence

<400> 55
gcggatgttc caatcagtac gca 23

<210> 56
<211> 23
<212> DNA
<213> Artificial

<220>
<223> spacer sequence

<400> 56
cagcccgctg gccctgtaaa gga 23

<210> 57
<211> 23
<212> DNA
<213> Artificial

<220>
<223> spacer sequence

<400> 57
tatgagttac aacgaacacc tca 23

<210> 58
<211> 23
<212> DNA
<213> Artificial

<220>
<223> spacer sequence

<400> 58
cacgtctcat atgccccttg gca 23

<210> 59
<211> 23
<212> DNA
<213> Artificial

<220>
<223> spacer sequence

<400> 59
gaacatgaaa acttaaatag aac 23

<210> 60
<211> 23
<212> DNA
<213> Artificial

<220>
<223> spacer sequence

<400> 60
gtcaggttgg ctgctgggct ggc 23

<210> 61
<211> 23
<212> DNA
<213> Artificial

<220>
<223> spacer sequence

<400> 61
acagatgggg ctggacaatt ttt 23

<210> 62

<211> 43
<212> RNA
<213> Artificial

<220>
<223> CRISPR RNA

<400> 62
aauuucuacu aaguguagau ggaaucccuu cugcagcacc ugg 43