SEQUENCE LISTING

<110> THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA

<120> GENE THERAPY FOR TREATING MUCOPOLYSACCHARIDOSIS TYPE II

<130> SP100062SGM

<150> US 62/323,194
<151> 2016-04-15

<150> US 62/330,938
<151> 2016-05-03

<150> US 62/337,163
<151> 2016-05-16

<150> US 62/367,780
<151> 2016-07-28

<150> US 62/452,494
<151> 2017-01-31

<160> 18

<170> PatentIn version 3.5

<210> 1
<211> 1653
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(1653)
<223> human IDS enzyme

<400> 1

```
atg ccg cca ccc cgg acc ggc cga ggc ctt ctc tgg ctg ggt ctg gtt         48
Met Pro Pro Pro Arg Thr Gly Arg Gly Leu Leu Trp Leu Gly Leu Val
1               5                  10                  15

ctg agc tcc gtc tgc gtc gcc ctc gga tcc gaa acg cag gcc aac tcg         96
Leu Ser Ser Val Cys Val Ala Leu Gly Ser Glu Thr Gln Ala Asn Ser
            20                  25                  30

acc aca gat gct ctg aac gtt ctt ctc atc atc gtg gat gac ctg cgc        144
Thr Thr Asp Ala Leu Asn Val Leu Leu Ile Ile Val Asp Asp Leu Arg
        35                  40                  45

ccc tcc ctg ggc tgt tat ggg gat aag ctg gtg agg tcc cca aat att        192
Pro Ser Leu Gly Cys Tyr Gly Asp Lys Leu Val Arg Ser Pro Asn Ile
    50                  55                  60

gac caa ctg gca tcc cac agc ctc ctc ttc cag aat gcc ttt gcg cag        240
Asp Gln Leu Ala Ser His Ser Leu Leu Phe Gln Asn Ala Phe Ala Gln
65                  70                  75                  80

caa gca gtg tgc gcc ccg agc cgc gtt tct ttc ctc act ggc agg aga        288
Gln Ala Val Cys Ala Pro Ser Arg Val Ser Phe Leu Thr Gly Arg Arg
                85                  90                  95

cct gac acc acc cgc ctg tac gac ttc aac tcc tac tgg agg gtg cac        336
Pro Asp Thr Thr Arg Leu Tyr Asp Phe Asn Ser Tyr Trp Arg Val His
            100                 105                 110

gct gga aac ttc tcc acc atc ccc cag tac ttc aag gag aat ggc tat        384
```

```
Ala Gly Asn Phe Ser Thr Ile Pro Gln Tyr Phe Lys Glu Asn Gly Tyr
        115                 120                 125

gtg acc atg tcg gtg gga aaa gtc ttt cac cct ggg ata tct tct aac        432
Val Thr Met Ser Val Gly Lys Val Phe His Pro Gly Ile Ser Ser Asn
    130                 135                 140

cat acc gat gat tct ccg tat agc tgg tct ttt cca cct tat cat cct        480
His Thr Asp Asp Ser Pro Tyr Ser Trp Ser Phe Pro Pro Tyr His Pro
145                 150                 155                 160

tcc tct gag aag tat gaa aac act aag aca tgt cga ggg cca gat gga        528
Ser Ser Glu Lys Tyr Glu Asn Thr Lys Thr Cys Arg Gly Pro Asp Gly
                165                 170                 175

gaa ctc cat gcc aac ctg ctt tgc cct gtg gat gtg ctg gat gtt ccc        576
Glu Leu His Ala Asn Leu Leu Cys Pro Val Asp Val Leu Asp Val Pro
            180                 185                 190

gag ggc acc ttg cct gac aaa cag agc act gag caa gcc ata cag ttg        624
Glu Gly Thr Leu Pro Asp Lys Gln Ser Thr Glu Gln Ala Ile Gln Leu
        195                 200                 205

ttg gaa aag atg aaa acg tca gcc agt cct ttc ttc ctg gcc gtt ggg        672
Leu Glu Lys Met Lys Thr Ser Ala Ser Pro Phe Phe Leu Ala Val Gly
    210                 215                 220

tat cat aag cca cac atc ccc ttc aga tac ccc aag gaa ttt cag aag        720
Tyr His Lys Pro His Ile Pro Phe Arg Tyr Pro Lys Glu Phe Gln Lys
225                 230                 235                 240

ttg tat ccc ttg gag aac atc acc ctg gcc ccc gat ccc gag gtc cct        768
Leu Tyr Pro Leu Glu Asn Ile Thr Leu Ala Pro Asp Pro Glu Val Pro
                245                 250                 255

gat ggc cta ccc cct gtg gcc tac aac ccc tgg atg gac atc agg caa        816
Asp Gly Leu Pro Pro Val Ala Tyr Asn Pro Trp Met Asp Ile Arg Gln
            260                 265                 270

cgg gaa gac gtc caa gcc tta aac atc agt gtg ccg tat ggt cca att        864
Arg Glu Asp Val Gln Ala Leu Asn Ile Ser Val Pro Tyr Gly Pro Ile
        275                 280                 285

cct gtg gac ttt cag cgg aaa atc cgc cag agc tac ttt gcc tct gtg        912
Pro Val Asp Phe Gln Arg Lys Ile Arg Gln Ser Tyr Phe Ala Ser Val
    290                 295                 300

tca tat ttg gat aca cag gtc ggc cgc ctc ttg agt gct ttg gac gat        960
Ser Tyr Leu Asp Thr Gln Val Gly Arg Leu Leu Ser Ala Leu Asp Asp
305                 310                 315                 320

ctt cag ctg gcc aac agc acc atc att gca ttt acc tcg gat cat ggg       1008
Leu Gln Leu Ala Asn Ser Thr Ile Ile Ala Phe Thr Ser Asp His Gly
                325                 330                 335

tgg gct cta ggt gaa cat gga gaa tgg gcc aaa tac agc aat ttt gat       1056
Trp Ala Leu Gly Glu His Gly Glu Trp Ala Lys Tyr Ser Asn Phe Asp
            340                 345                 350

gtt gct acc cat gtt ccc ctg ata ttc tat gtt cct gga agg acg gct       1104
Val Ala Thr His Val Pro Leu Ile Phe Tyr Val Pro Gly Arg Thr Ala
        355                 360                 365

tca ctt ccg gag gca ggc gag aag ctt ttc cct tac ctc gac cct ttt       1152
Ser Leu Pro Glu Ala Gly Glu Lys Leu Phe Pro Tyr Leu Asp Pro Phe
    370                 375                 380

gat tcc gcc tca cag ttg atg gag cca ggc agg caa tcc atg gac ctt       1200
```

```
Asp Ser Ala Ser Gln Leu Met Glu Pro Gly Arg Gln Ser Met Asp Leu
385                 390                 395                 400

gtg gaa ctt gtg tct ctt ttt ccc acg ctg gct gga ctt gca gga ctg       1248
Val Glu Leu Val Ser Leu Phe Pro Thr Leu Ala Gly Leu Ala Gly Leu
                405                 410                 415

cag gtt cca cct cgc tgc ccc gtt cct tca ttt cac gtt gag ctg tgc       1296
Gln Val Pro Pro Arg Cys Pro Val Pro Ser Phe His Val Glu Leu Cys
            420                 425                 430

aga gaa ggc aag aac ctt ctg aag cat ttt cga ttc cgt gac ttg gaa       1344
Arg Glu Gly Lys Asn Leu Leu Lys His Phe Arg Phe Arg Asp Leu Glu
        435                 440                 445

gag gat ccg tac ctc cct ggt aat ccc cgt gaa ctg att gcc tat agc       1392
Glu Asp Pro Tyr Leu Pro Gly Asn Pro Arg Glu Leu Ile Ala Tyr Ser
    450                 455                 460

cag tat ccc cgg cct tca gac atc cct cag tgg aat tct gac aag ccg       1440
Gln Tyr Pro Arg Pro Ser Asp Ile Pro Gln Trp Asn Ser Asp Lys Pro
465                 470                 475                 480

agt tta aaa gat ata aag atc atg ggc tat tcc ata cgc acc ata gac       1488
Ser Leu Lys Asp Ile Lys Ile Met Gly Tyr Ser Ile Arg Thr Ile Asp
                485                 490                 495

tat agg tat act gtg tgg gtt ggc ttc aat cct gat gaa ttt cta gct       1536
Tyr Arg Tyr Thr Val Trp Val Gly Phe Asn Pro Asp Glu Phe Leu Ala
            500                 505                 510

aac ttt tct gac atc cat gca ggg gaa ctg tat ttt gtg gat tct gac       1584
Asn Phe Ser Asp Ile His Ala Gly Glu Leu Tyr Phe Val Asp Ser Asp
        515                 520                 525

cca ttg cag gat cac aat atg tat aat gat tcc caa ggt gga gat ctt       1632
Pro Leu Gln Asp His Asn Met Tyr Asn Asp Ser Gln Gly Gly Asp Leu
    530                 535                 540

ttc cag ttg ttg atg cct tga                                          1653
Phe Gln Leu Leu Met Pro
545                 550
```

<210> 2
<211> 550
<212> PRT
<213> Homo sapiens

<400> 2

```
Met Pro Pro Pro Arg Thr Gly Arg Gly Leu Leu Trp Leu Gly Leu Val
1               5                   10                  15

Leu Ser Ser Val Cys Val Ala Leu Gly Ser Glu Thr Gln Ala Asn Ser
            20                  25                  30

Thr Thr Asp Ala Leu Asn Val Leu Leu Ile Ile Val Asp Asp Leu Arg
        35                  40                  45

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

Asp Gln Leu Ala Ser His Ser Leu Leu Phe Gln Asn Ala Phe Ala Gln
```

65 70 75 80

Gln Ala Val Cys Ala Pro Ser Arg Val Ser Phe Leu Thr Gly Arg Arg
85 90 95

Pro Asp Thr Thr Arg Leu Tyr Asp Phe Asn Ser Tyr Trp Arg Val His
100 105 110

Ala Gly Asn Phe Ser Thr Ile Pro Gln Tyr Phe Lys Glu Asn Gly Tyr
115 120 125

Val Thr Met Ser Val Gly Lys Val Phe His Pro Gly Ile Ser Ser Asn
130 135 140

His Thr Asp Asp Ser Pro Tyr Ser Trp Ser Phe Pro Pro Tyr His Pro
145 150 155 160

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

Glu Leu His Ala Asn Leu Leu Cys Pro Val Asp Val Leu Asp Val Pro
180 185 190

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

Leu Glu Lys Met Lys Thr Ser Ala Ser Pro Phe Phe Leu Ala Val Gly
210 215 220

Tyr His Lys Pro His Ile Pro Phe Arg Tyr Pro Lys Glu Phe Gln Lys
225 230 235 240

Leu Tyr Pro Leu Glu Asn Ile Thr Leu Ala Pro Asp Pro Glu Val Pro
245 250 255

Asp Gly Leu Pro Pro Val Ala Tyr Asn Pro Trp Met Asp Ile Arg Gln
260 265 270

Arg Glu Asp Val Gln Ala Leu Asn Ile Ser Val Pro Tyr Gly Pro Ile
275 280 285

Pro Val Asp Phe Gln Arg Lys Ile Arg Gln Ser Tyr Phe Ala Ser Val
290 295 300

Ser Tyr Leu Asp Thr Gln Val Gly Arg Leu Leu Ser Ala Leu Asp Asp
305 310 315 320

Leu Gln Leu Ala Asn Ser Thr Ile Ile Ala Phe Thr Ser Asp His Gly
325 330 335

Trp Ala Leu Gly Glu His Gly Glu Trp Ala Lys Tyr Ser Asn Phe Asp

340 345 350

Val Ala Thr His Val Pro Leu Ile Phe Tyr Val Pro Gly Arg Thr Ala
355 360 365

Ser Leu Pro Glu Ala Gly Glu Lys Leu Phe Pro Tyr Leu Asp Pro Phe
370 375 380

Asp Ser Ala Ser Gln Leu Met Glu Pro Gly Arg Gln Ser Met Asp Leu
385 390 395 400

Val Glu Leu Val Ser Leu Phe Pro Thr Leu Ala Gly Leu Ala Gly Leu
405 410 415

Gln Val Pro Pro Arg Cys Pro Val Pro Ser Phe His Val Glu Leu Cys
420 425 430

Arg Glu Gly Lys Asn Leu Leu Lys His Phe Arg Phe Arg Asp Leu Glu
435 440 445

Glu Asp Pro Tyr Leu Pro Gly Asn Pro Arg Glu Leu Ile Ala Tyr Ser
450 455 460

Gln Tyr Pro Arg Pro Ser Asp Ile Pro Gln Trp Asn Ser Asp Lys Pro
465 470 475 480

Ser Leu Lys Asp Ile Lys Ile Met Gly Tyr Ser Ile Arg Thr Ile Asp
485 490 495

Tyr Arg Tyr Thr Val Trp Val Gly Phe Asn Pro Asp Glu Phe Leu Ala
500 505 510

Asn Phe Ser Asp Ile His Ala Gly Glu Leu Tyr Phe Val Asp Ser Asp
515 520 525

Pro Leu Gln Asp His Asn Met Tyr Asn Asp Ser Gln Gly Gly Asp Leu
530 535 540

Phe Gln Leu Leu Met Pro
545 550

<210> 3
<211> 3967
<212> DNA
<213> Artificial Sequence

<220>
<223> CB7.CI.hIDS.RBG

<220>
<221> misc_feature
<222> (2)..(131)
<223> 5' ITR

<220>
<221> promoter
<222> (199)..(580)
<223> CMV IE promoter

<220>
<221> promoter
<222> (583)..(864)
<223> CB promoter

<220>
<221> TATA_signal
<222> (837)..(840)

<220>
<221> Intron
<222> (959)..(1930)
<223> chicken beta-actin intron

<220>
<221> CDS
<222> (1937)..(3589)
<223> hIDS

<220>
<221> polyA_signal
<222> (3623)..(3749)
<223> rabbit globin polyA

<220>
<221> misc_feature
<222> (3838)..(3967)
<223> 3' ITR

<400> 3
gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc gggcgacctt 60

tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca actccatcac 120

taggggttcc ttgtagttaa tgattaaccc gccatgctac ttatctacca gggtaatggg 180

gatcctctag aactatagct agtcgacatt gattattgac tagttattaa tagtaatcaa 240

ttacggggtc attagttcat agcccatata tggagttccg cgttacataa cttacggtaa 300

atggcccgcc tggctgaccg cccaacgacc cccgcccatt gacgtcaata atgacgtatg 360

ttcccatagt aacgccaata gggactttcc attgacgtca atgggtggac tatttacggt 420

aaactgccca cttggcagta catcaagtgt atcatatgcc aagtacgccc cctattgacg 480

tcaatgacgg taaatggccc gcctggcatt atgcccagta catgacctta tgggactttc 540

ctacttggca gtacatctac gtattagtca tcgctattac catggtcgag gtgagcccca 600

cgttctgctt cactctcccc atctcccccc cctccccacc cccaattttg tatttattta 660

ttttttaatt attttgtgca gcgatggggg cgggggggg gggggggcgc gcgccaggcg 720

gggcggggcg gggcgagggg cggggcgggg cgaggcggag aggtgcggcg gcagccaatc 780

agagcggcgc gctccgaaag tttcctttta tggcgaggcg gcggcggcgg cggccctata 840

aaaagcgaag cgcgcggcgg gcggggagtc gctgcgacgc tgccttcgcc ccgtgccccg 900

ctccgccgcc gcctcgcgcc gcccgccccg gctctgactg accgcgttac tcccacaggt 960

```
gagcgggcgg gacggccctt ctcctccggg ctgtaattag cgcttggttt aatgacggct     1020

tgtttctttt ctgtggctgc gtgaaagcct tgaggggctc cgggagggcc ctttgtgcgg     1080

ggggagcggc tcgggggggtg cgtgcgtgtg tgtgtgcgtg gggagcgccg cgtgcggctc     1140

cgcgctgccc ggcggctgtg agcgctgcgg gcgcggcgcg gggctttgtg cgctccgcag     1200

tgtgcgcgag gggagcgcgg ccgggggcgg tgccccgcgg tgcgggggggg gctgcgaggg     1260

gaacaaaggc tgcgtgcggg gtgtgtgcgt ggggggggtga gcagggggtg tgggcgcgtc     1320

ggtcgggctg caacccccccc tgcacccccc tccccgagtt gctgagcacg gcccggcttc     1380

gggtgcgggg ctccgtacgg ggcgtggcgc ggggctcgcc gtgccgggcg gggggtggcg     1440

gcaggtgggg gtgccgggcg gggcggggcc gcctcgggcc ggggagggct cgggggaggg     1500

gcgcggcggc ccccggagcg ccggcggctg tcgaggcgcg gcgagccgca gccattgcct     1560

tttatggtaa tcgtgcgaga gggcgcaggg acttcctttg tcccaaatct gtgcggagcc     1620

gaaatctggg aggcgccgcc gcacccccctc tagcgggcgc ggggcgaagc ggtgcggcgc     1680

cggcaggaag gaaatgggcg gggagggcct tcgtgcgtcg ccgcgccgcc gtccccttct     1740

ccctctccag cctcggggct gtccgcgggg ggacggctgc cttcgggggg gacggggcag     1800

ggcggggttc ggcttctggc gtgtgaccgg cggctctaga gcctctgcta accatgttca     1860

tgccttcttc tttttcctac agctcctggg caacgtgctg gttattgtgc tgtctcatca     1920

ttttggcaaa gaattc atg ccg cca ccc cgg acc ggc cga ggc ctt ctc tgg     1972
                  Met Pro Pro Pro Arg Thr Gly Arg Gly Leu Leu Trp
                  1               5                   10

ctg ggt ctg gtt ctg agc tcc gtc tgc gtc gcc ctc gga tcc gaa acg      2020
Leu Gly Leu Val Leu Ser Ser Val Cys Val Ala Leu Gly Ser Glu Thr
        15                  20                  25

cag gcc aac tcg acc aca gat gct ctg aac gtt ctt ctc atc atc gtg      2068
Gln Ala Asn Ser Thr Thr Asp Ala Leu Asn Val Leu Leu Ile Ile Val
    30                  35                  40

gat gac ctg cgc ccc tcc ctg ggc tgt tat ggg gat aag ctg gtg agg      2116
Asp Asp Leu Arg Pro Ser Leu Gly Cys Tyr Gly Asp Lys Leu Val Arg
45                  50                  55                  60

tcc cca aat att gac caa ctg gca tcc cac agc ctc ctc ttc cag aat      2164
Ser Pro Asn Ile Asp Gln Leu Ala Ser His Ser Leu Leu Phe Gln Asn
                65                  70                  75

gcc ttt gcg cag caa gca gtg tgc gcc ccg agc cgc gtt tct ttc ctc      2212
Ala Phe Ala Gln Gln Ala Val Cys Ala Pro Ser Arg Val Ser Phe Leu
            80                  85                  90

act ggc agg aga cct gac acc acc cgc ctg tac gac ttc aac tcc tac      2260
Thr Gly Arg Arg Pro Asp Thr Thr Arg Leu Tyr Asp Phe Asn Ser Tyr
        95                  100                 105

tgg agg gtg cac gct gga aac ttc tcc acc atc ccc cag tac ttc aag      2308
Trp Arg Val His Ala Gly Asn Phe Ser Thr Ile Pro Gln Tyr Phe Lys
    110                 115                 120

gag aat ggc tat gtg acc atg tcg gtg gga aaa gtc ttt cac cct ggg      2356
Glu Asn Gly Tyr Val Thr Met Ser Val Gly Lys Val Phe His Pro Gly
125                 130                 135                 140
```

```
ata tct tct aac cat acc gat gat tct ccg tat agc tgg tct ttt cca       2404
Ile Ser Ser Asn His Thr Asp Asp Ser Pro Tyr Ser Trp Ser Phe Pro
                145                 150                 155

cct tat cat cct tcc tct gag aag tat gaa aac act aag aca tgt cga       2452
Pro Tyr His Pro Ser Ser Glu Lys Tyr Glu Asn Thr Lys Thr Cys Arg
            160                 165                 170

ggg cca gat gga gaa ctc cat gcc aac ctg ctt tgc cct gtg gat gtg       2500
Gly Pro Asp Gly Glu Leu His Ala Asn Leu Leu Cys Pro Val Asp Val
        175                 180                 185

ctg gat gtt ccc gag ggc acc ttg cct gac aaa cag agc act gag caa       2548
Leu Asp Val Pro Glu Gly Thr Leu Pro Asp Lys Gln Ser Thr Glu Gln
    190                 195                 200

gcc ata cag ttg ttg gaa aag atg aaa acg tca gcc agt cct ttc ttc       2596
Ala Ile Gln Leu Leu Glu Lys Met Lys Thr Ser Ala Ser Pro Phe Phe
205                 210                 215                 220

ctg gcc gtt ggg tat cat aag cca cac atc ccc ttc aga tac ccc aag       2644
Leu Ala Val Gly Tyr His Lys Pro His Ile Pro Phe Arg Tyr Pro Lys
                225                 230                 235

gaa ttt cag aag ttg tat ccc ttg gag aac atc acc ctg gcc ccc gat       2692
Glu Phe Gln Lys Leu Tyr Pro Leu Glu Asn Ile Thr Leu Ala Pro Asp
            240                 245                 250

ccc gag gtc cct gat ggc cta ccc cct gtg gcc tac aac ccc tgg atg       2740
Pro Glu Val Pro Asp Gly Leu Pro Pro Val Ala Tyr Asn Pro Trp Met
        255                 260                 265

gac atc agg caa cgg gaa gac gtc caa gcc tta aac atc agt gtg ccg       2788
Asp Ile Arg Gln Arg Glu Asp Val Gln Ala Leu Asn Ile Ser Val Pro
    270                 275                 280

tat ggt cca att cct gtg gac ttt cag cgg aaa atc cgc cag agc tac       2836
Tyr Gly Pro Ile Pro Val Asp Phe Gln Arg Lys Ile Arg Gln Ser Tyr
285                 290                 295                 300

ttt gcc tct gtg tca tat ttg gat aca cag gtc ggc cgc ctc ttg agt       2884
Phe Ala Ser Val Ser Tyr Leu Asp Thr Gln Val Gly Arg Leu Leu Ser
                305                 310                 315

gct ttg gac gat ctt cag ctg gcc aac agc acc atc att gca ttt acc       2932
Ala Leu Asp Asp Leu Gln Leu Ala Asn Ser Thr Ile Ile Ala Phe Thr
            320                 325                 330

tcg gat cat ggg tgg gct cta ggt gaa cat gga gaa tgg gcc aaa tac       2980
Ser Asp His Gly Trp Ala Leu Gly Glu His Gly Glu Trp Ala Lys Tyr
        335                 340                 345

agc aat ttt gat gtt gct acc cat gtt ccc ctg ata ttc tat gtt cct       3028
Ser Asn Phe Asp Val Ala Thr His Val Pro Leu Ile Phe Tyr Val Pro
    350                 355                 360

gga agg acg gct tca ctt ccg gag gca ggc gag aag ctt ttc cct tac       3076
Gly Arg Thr Ala Ser Leu Pro Glu Ala Gly Glu Lys Leu Phe Pro Tyr
365                 370                 375                 380

ctc gac cct ttt gat tcc gcc tca cag ttg atg gag cca ggc agg caa       3124
Leu Asp Pro Phe Asp Ser Ala Ser Gln Leu Met Glu Pro Gly Arg Gln
                385                 390                 395

tcc atg gac ctt gtg gaa ctt gtg tct ctt ttt ccc acg ctg gct gga       3172
Ser Met Asp Leu Val Glu Leu Val Ser Leu Phe Pro Thr Leu Ala Gly
            400                 405                 410
```

```
ctt gca gga ctg cag gtt cca cct cgc tgc ccc gtt cct tca ttt cac       3220
Leu Ala Gly Leu Gln Val Pro Pro Arg Cys Pro Val Pro Ser Phe His
        415                 420                 425

gtt gag ctg tgc aga gaa ggc aag aac ctt ctg aag cat ttt cga ttc       3268
Val Glu Leu Cys Arg Glu Gly Lys Asn Leu Leu Lys His Phe Arg Phe
    430                 435                 440

cgt gac ttg gaa gag gat ccg tac ctc cct ggt aat ccc cgt gaa ctg       3316
Arg Asp Leu Glu Glu Asp Pro Tyr Leu Pro Gly Asn Pro Arg Glu Leu
445                 450                 455                 460

att gcc tat agc cag tat ccc cgg cct tca gac atc cct cag tgg aat       3364
Ile Ala Tyr Ser Gln Tyr Pro Arg Pro Ser Asp Ile Pro Gln Trp Asn
                465                 470                 475

tct gac aag ccg agt tta aaa gat ata aag atc atg ggc tat tcc ata       3412
Ser Asp Lys Pro Ser Leu Lys Asp Ile Lys Ile Met Gly Tyr Ser Ile
            480                 485                 490

cgc acc ata gac tat agg tat act gtg tgg gtt ggc ttc aat cct gat       3460
Arg Thr Ile Asp Tyr Arg Tyr Thr Val Trp Val Gly Phe Asn Pro Asp
        495                 500                 505

gaa ttt cta gct aac ttt tct gac atc cat gca ggg gaa ctg tat ttt       3508
Glu Phe Leu Ala Asn Phe Ser Asp Ile His Ala Gly Glu Leu Tyr Phe
    510                 515                 520

gtg gat tct gac cca ttg cag gat cac aat atg tat aat gat tcc caa       3556
Val Asp Ser Asp Pro Leu Gln Asp His Asn Met Tyr Asn Asp Ser Gln
525                 530                 535                 540

ggt gga gat ctt ttc cag ttg ttg atg cct tga ctcgaggacg gggtgaacta     3609
Gly Gly Asp Leu Phe Gln Leu Leu Met Pro
                545                 550

cgcctgagga tccgatcttt ttccctctgc caaaaattat ggggacatca tgaagcccct     3669

tgagcatctg acttctggct aataaaggaa atttattttc attgcaatag tgtgttggaa     3729

ttttttgtgt ctctcactcg gaagcaattc gttgatctga atttcgacca cccataatac     3789

ccattaccct ggtagataag tagcatggcg ggttaatcat taactacaag gaacccctag     3849

tgatggagtt ggccactccc tctctgcgcg ctcgctcgct cactgaggcc gggcgaccaa     3909

aggtcgcccg acgcccgggc tttgcccggg cggcctcagt gagcgagcga gcgcgcag       3967
```

<210> 4
<211> 550
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 4

```
Met Pro Pro Pro Arg Thr Gly Arg Gly Leu Leu Trp Leu Gly Leu Val
1               5                   10                  15

Leu Ser Ser Val Cys Val Ala Leu Gly Ser Glu Thr Gln Ala Asn Ser
            20                  25                  30

Thr Thr Asp Ala Leu Asn Val Leu Leu Ile Ile Val Asp Asp Leu Arg
```

35 40 45

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

Asp Gln Leu Ala Ser His Ser Leu Leu Phe Gln Asn Ala Phe Ala Gln
65 70 75 80

Gln Ala Val Cys Ala Pro Ser Arg Val Ser Phe Leu Thr Gly Arg Arg
85 90 95

Pro Asp Thr Thr Arg Leu Tyr Asp Phe Asn Ser Tyr Trp Arg Val His
100 105 110

Ala Gly Asn Phe Ser Thr Ile Pro Gln Tyr Phe Lys Glu Asn Gly Tyr
115 120 125

Val Thr Met Ser Val Gly Lys Val Phe His Pro Gly Ile Ser Ser Asn
130 135 140

His Thr Asp Asp Ser Pro Tyr Ser Trp Ser Phe Pro Pro Tyr His Pro
145 150 155 160

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

Glu Leu His Ala Asn Leu Leu Cys Pro Val Asp Val Leu Asp Val Pro
180 185 190

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

Leu Glu Lys Met Lys Thr Ser Ala Ser Pro Phe Phe Leu Ala Val Gly
210 215 220

Tyr His Lys Pro His Ile Pro Phe Arg Tyr Pro Lys Glu Phe Gln Lys
225 230 235 240

Leu Tyr Pro Leu Glu Asn Ile Thr Leu Ala Pro Asp Pro Glu Val Pro
245 250 255

Asp Gly Leu Pro Pro Val Ala Tyr Asn Pro Trp Met Asp Ile Arg Gln
260 265 270

Arg Glu Asp Val Gln Ala Leu Asn Ile Ser Val Pro Tyr Gly Pro Ile
275 280 285

Pro Val Asp Phe Gln Arg Lys Ile Arg Gln Ser Tyr Phe Ala Ser Val
290 295 300

Ser Tyr Leu Asp Thr Gln Val Gly Arg Leu Leu Ser Ala Leu Asp Asp

305 310 315 320

Leu Gln Leu Ala Asn Ser Thr Ile Ile Ala Phe Thr Ser Asp His Gly
325 330 335

Trp Ala Leu Gly Glu His Gly Glu Trp Ala Lys Tyr Ser Asn Phe Asp
340 345 350

Val Ala Thr His Val Pro Leu Ile Phe Tyr Val Pro Gly Arg Thr Ala
355 360 365

Ser Leu Pro Glu Ala Gly Glu Lys Leu Phe Pro Tyr Leu Asp Pro Phe
370 375 380

Asp Ser Ala Ser Gln Leu Met Glu Pro Gly Arg Gln Ser Met Asp Leu
385 390 395 400

Val Glu Leu Val Ser Leu Phe Pro Thr Leu Ala Gly Leu Ala Gly Leu
405 410 415

Gln Val Pro Pro Arg Cys Pro Val Pro Ser Phe His Val Glu Leu Cys
420 425 430

Arg Glu Gly Lys Asn Leu Leu Lys His Phe Arg Phe Arg Asp Leu Glu
435 440 445

Glu Asp Pro Tyr Leu Pro Gly Asn Pro Arg Glu Leu Ile Ala Tyr Ser
450 455 460

Gln Tyr Pro Arg Pro Ser Asp Ile Pro Gln Trp Asn Ser Asp Lys Pro
465 470 475 480

Ser Leu Lys Asp Ile Lys Ile Met Gly Tyr Ser Ile Arg Thr Ile Asp
485 490 495

Tyr Arg Tyr Thr Val Trp Val Gly Phe Asn Pro Asp Glu Phe Leu Ala
500 505 510

Asn Phe Ser Asp Ile His Ala Gly Glu Leu Tyr Phe Val Asp Ser Asp
515 520 525

Pro Leu Gln Asp His Asn Met Tyr Asn Asp Ser Gln Gly Gly Asp Leu
530 535 540

Phe Gln Leu Leu Met Pro
545 550

<210> 5
<211> 4964
<212> DNA
<213> Artificial Sequence

<220>
<223> CB6.hIDS.IRES.SUMF1

<220>
<221> misc_feature
<222> (11)..(140)
<223> 5'ITR

<220>
<221> enhancer
<222> (208)..(589)
<223> CMV IE enhancer

<220>
<221> promoter
<222> (600)..(859)
<223> CB promoter

<220>
<221> TATA_signal
<222> (835)..(842)

<220>
<221> Intron
<222> (976)..(1108)

<220>
<221> CDS
<222> (1177)..(2829)
<223> hIDS

<220>
<221> misc_feature
<222> (2830)..(3422)
<223> IRES

<220>
<221> CDS
<222> (3423)..(4547)
<223> hSUMF

<220>
<221> polyA_signal
<222> (4620)..(4746)
<223> rabbit globin polyA

<220>
<221> misc_feature
<222> (4835)..(4964)
<223> 3'ITR

<400> 5
cttaattagg ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg 60

ggcgaccttt ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa 120

ctccatcact aggggttcct tgtagttaat gattaacccg ccatgctact tatctaccag 180

ggtaatgggg atcctctaga actatagcta gtcgacattg attattgact agttattaat 240

agtaatcaat tacggggtca ttagttcata gcccatatat ggagttccgc gttacataac 300

ttacggtaaa tggcccgcct ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa 360

tgacgtatgt tcccatagta acgccaatag ggactttcca ttgacgtcaa tgggtggact 420

atttacggta aactgcccac ttggcagtac atcaagtgta tcatatgcca agtacgcccc 480

```
ctattgacgt caatgacggt aaatggcccg cctggcatta tgcccagtac atgaccttat      540

gggactttcc tacttggcag tacatctacg tattagtcat cgctattacc atgtcgaggc      600

cacgttctgc ttcactctcc ccatctcccc cccctcccca cccccaattt tgtatttatt      660

tatttttaa ttattttgtg cagcgatggg ggcgggggggg gggggcgcgc gccaggcggg      720

gcggggcggg gcgaggggcg gggcggggcg aggcggagag gtgcggcggc agccaatcag      780

agcggcgcgc tccgaaagtt tccttttatg gcgaggcggc ggcggcggcg gccctataaa      840

aagcgaagcg cgcggcgggc gggagcaagc tttattgcgg tagtttatca cagttaaatt      900

gctaacgcag tcagtgcttc tgacacaaca gtctcgaact taagctgcag aagttggtcg      960

tgaggcactg ggcaggtaag tatcaaggtt acaagacagg tttaaggaga ccaatagaaa     1020

ctgggcttgt cgagacagag aagactcttg cgtttctgat aggcacctat tggtcttact     1080

gacatccact ttgcctttct ctccacaggt gtccactccc agttcaatta cagctcttaa     1140

ggctagagta cttaatacga ctcactatag gctagc atg ccg cca ccc cgg acc        1194
                                        Met Pro Pro Pro Arg Thr
                                        1               5

ggc cga ggc ctt ctc tgg ctg ggt ctg gtt ctg agc tcc gtc tgc gtc        1242
Gly Arg Gly Leu Leu Trp Leu Gly Leu Val Leu Ser Ser Val Cys Val
            10                  15                  20

gcc ctc gga tcc gaa acg cag gcc aac tcg acc aca gat gct ctg aac        1290
Ala Leu Gly Ser Glu Thr Gln Ala Asn Ser Thr Thr Asp Ala Leu Asn
        25                  30                  35

gtt ctt ctc atc atc gtg gat gac ctg cgc ccc tcc ctg ggc tgt tat        1338
Val Leu Leu Ile Ile Val Asp Asp Leu Arg Pro Ser Leu Gly Cys Tyr
    40                  45                  50

ggg gat aag ctg gtg agg tcc cca aat att gac caa ctg gca tcc cac        1386
Gly Asp Lys Leu Val Arg Ser Pro Asn Ile Asp Gln Leu Ala Ser His
55                  60                  65                  70

agc ctc ctc ttc cag aat gcc ttt gcg cag caa gca gtg tgc gcc ccg        1434
Ser Leu Leu Phe Gln Asn Ala Phe Ala Gln Gln Ala Val Cys Ala Pro
                75                  80                  85

agc cgc gtt tct ttc ctc act ggc agg aga cct gac acc acc cgc ctg        1482
Ser Arg Val Ser Phe Leu Thr Gly Arg Arg Pro Asp Thr Thr Arg Leu
            90                  95                  100

tac gac ttc aac tcc tac tgg agg gtg cac gct gga aac ttc tcc acc        1530
Tyr Asp Phe Asn Ser Tyr Trp Arg Val His Ala Gly Asn Phe Ser Thr
        105                 110                 115

atc ccc cag tac ttc aag gag aat ggc tat gtg acc atg tcg gtg gga        1578
Ile Pro Gln Tyr Phe Lys Glu Asn Gly Tyr Val Thr Met Ser Val Gly
    120                 125                 130

aaa gtc ttt cac cct ggg ata tct tct aac cat acc gat gat tct ccg        1626
Lys Val Phe His Pro Gly Ile Ser Ser Asn His Thr Asp Asp Ser Pro
135                 140                 145                 150

tat agc tgg tct ttt cca cct tat cat cct tcc tct gag aag tat gaa        1674
Tyr Ser Trp Ser Phe Pro Pro Tyr His Pro Ser Ser Glu Lys Tyr Glu
                155                 160                 165

aac act aag aca tgt cga ggg cca gat gga gaa ctc cat gcc aac ctg        1722
```

```
Asn Thr Lys Thr Cys Arg Gly Pro Asp Gly Glu Leu His Ala Asn Leu
            170                 175                 180

ctt tgc cct gtg gat gtg ctg gat gtt ccc gag ggc acc ttg cct gac       1770
Leu Cys Pro Val Asp Val Leu Asp Val Pro Glu Gly Thr Leu Pro Asp
        185                 190                 195

aaa cag agc act gag caa gcc ata cag ttg ttg gaa aag atg aaa acg       1818
Lys Gln Ser Thr Glu Gln Ala Ile Gln Leu Leu Glu Lys Met Lys Thr
    200                 205                 210

tca gcc agt cct ttc ttc ctg gcc gtt ggg tat cat aag cca cac atc       1866
Ser Ala Ser Pro Phe Phe Leu Ala Val Gly Tyr His Lys Pro His Ile
215                 220                 225                 230

ccc ttc aga tac ccc aag gaa ttt cag aag ttg tat ccc ttg gag aac       1914
Pro Phe Arg Tyr Pro Lys Glu Phe Gln Lys Leu Tyr Pro Leu Glu Asn
                235                 240                 245

atc acc ctg gcc ccc gat ccc gag gtc cct gat ggc cta ccc cct gtg       1962
Ile Thr Leu Ala Pro Asp Pro Glu Val Pro Asp Gly Leu Pro Pro Val
            250                 255                 260

gcc tac aac ccc tgg atg gac atc agg caa cgg gaa gac gtc caa gcc       2010
Ala Tyr Asn Pro Trp Met Asp Ile Arg Gln Arg Glu Asp Val Gln Ala
        265                 270                 275

tta aac atc agt gtg ccg tat ggt cca att cct gtg gac ttt cag cgg       2058
Leu Asn Ile Ser Val Pro Tyr Gly Pro Ile Pro Val Asp Phe Gln Arg
    280                 285                 290

aaa atc cgc cag agc tac ttt gcc tct gtg tca tat ttg gat aca cag       2106
Lys Ile Arg Gln Ser Tyr Phe Ala Ser Val Ser Tyr Leu Asp Thr Gln
295                 300                 305                 310

gtc ggc cgc ctc ttg agt gct ttg gac gat ctt cag ctg gcc aac agc       2154
Val Gly Arg Leu Leu Ser Ala Leu Asp Asp Leu Gln Leu Ala Asn Ser
                315                 320                 325

acc atc att gca ttt acc tcg gat cat ggg tgg gct cta ggt gaa cat       2202
Thr Ile Ile Ala Phe Thr Ser Asp His Gly Trp Ala Leu Gly Glu His
            330                 335                 340

gga gaa tgg gcc aaa tac agc aat ttt gat gtt gct acc cat gtt ccc       2250
Gly Glu Trp Ala Lys Tyr Ser Asn Phe Asp Val Ala Thr His Val Pro
        345                 350                 355

ctg ata ttc tat gtt cct gga agg acg gct tca ctt ccg gag gca ggc       2298
Leu Ile Phe Tyr Val Pro Gly Arg Thr Ala Ser Leu Pro Glu Ala Gly
    360                 365                 370

gag aag ctt ttc cct tac ctc gac cct ttt gat tcc gcc tca cag ttg       2346
Glu Lys Leu Phe Pro Tyr Leu Asp Pro Phe Asp Ser Ala Ser Gln Leu
375                 380                 385                 390

atg gag cca ggc agg caa tcc atg gac ctt gtg gaa ctt gtg tct ctt       2394
Met Glu Pro Gly Arg Gln Ser Met Asp Leu Val Glu Leu Val Ser Leu
                395                 400                 405

ttt ccc acg ctg gct gga ctt gca gga ctg cag gtt cca cct cgc tgc       2442
Phe Pro Thr Leu Ala Gly Leu Ala Gly Leu Gln Val Pro Pro Arg Cys
            410                 415                 420

ccc gtt cct tca ttt cac gtt gag ctg tgc aga gaa ggc aag aac ctt       2490
Pro Val Pro Ser Phe His Val Glu Leu Cys Arg Glu Gly Lys Asn Leu
        425                 430                 435

ctg aag cat ttt cga ttc cgt gac ttg gaa gag gat ccg tac ctc cct       2538
```

```
Leu Lys His Phe Arg Phe Arg Asp Leu Glu Glu Asp Pro Tyr Leu Pro
    440                 445                 450

ggt aat ccc cgt gaa ctg att gcc tat agc cag tat ccc cgg cct tca       2586
Gly Asn Pro Arg Glu Leu Ile Ala Tyr Ser Gln Tyr Pro Arg Pro Ser
455                 460                 465                 470

gac atc cct cag tgg aat tct gac aag ccg agt tta aaa gat ata aag       2634
Asp Ile Pro Gln Trp Asn Ser Asp Lys Pro Ser Leu Lys Asp Ile Lys
                475                 480                 485

atc atg ggc tat tcc ata cgc acc ata gac tat agg tat act gtg tgg       2682
Ile Met Gly Tyr Ser Ile Arg Thr Ile Asp Tyr Arg Tyr Thr Val Trp
            490                 495                 500

gtt ggc ttc aat cct gat gaa ttt cta gct aac ttt tct gac atc cat       2730
Val Gly Phe Asn Pro Asp Glu Phe Leu Ala Asn Phe Ser Asp Ile His
        505                 510                 515

gca ggg gaa ctg tat ttt gtg gat tct gac cca ttg cag gat cac aat       2778
Ala Gly Glu Leu Tyr Phe Val Asp Ser Asp Pro Leu Gln Asp His Asn
    520                 525                 530

atg tat aat gat tcc caa ggt gga gat ctt ttc cag ttg ttg atg cct       2826
Met Tyr Asn Asp Ser Gln Gly Gly Asp Leu Phe Gln Leu Leu Met Pro
535                 540                 545                 550

tga gcccctctcc ctccccccc cctaacgtta ctggccgaag ccgcttggaa             2879

taaggccggt gtgcgtttgt ctatatgtta ttttccacca tattgccgtc ttttggcaat     2939

gtgagggccc ggaaacctgg ccctgtcttc ttgacgagca ttcctagggg tctttcccct     2999

ctcgccaaag gaatgcaagg tctgttgaat gtcgtgaagg aagcagttcc tctggaagct     3059

tcttgaagac aaacaacgtc tgtagcgacc ctttgcaggc agcggaaccc cccacctggc     3119

gacaggtgcc tctgcggcca aaagccacgt gtataagata cacctgcaaa ggcggcacaa     3179

ccccagtgcc acgttgtgag ttggatagtt gtggaaagag tcaaatggct ctcctcaagc     3239

gtattcaaca aggggctgaa ggatgcccag aaggtacccc attgtatggg atctgatctg     3299

gggcctcggt acacatgctt tacatgtgtt tagtcgaggt taaaaaaacg tctaggcccc     3359

ccgaaccacg gggacgtggt tttcctttga aaaacacgat gataatatgg ccacaaccat     3419

ggt atg gct gcg ccc gca cta ggg ctg gtg tgt gga cgt tgc cct gag       3467
    Met Ala Ala Pro Ala Leu Gly Leu Val Cys Gly Arg Cys Pro Glu
                555                 560                 565

ctg ggt ctc gtc ctc ttg ctg ctg ctg ctc tcg ctg ctg tgt gga gcg       3515
Leu Gly Leu Val Leu Leu Leu Leu Leu Leu Ser Leu Leu Cys Gly Ala
            570                 575                 580

gca ggg agc cag gag gcc ggg acc ggt gcg ggc gcg ggg tcc ctt gcg       3563
Ala Gly Ser Gln Glu Ala Gly Thr Gly Ala Gly Ala Gly Ser Leu Ala
        585                 590                 595

ggt tct tgc ggc tgc ggc acg ccc cag cgg cct ggc gcc cat ggc agt       3611
Gly Ser Cys Gly Cys Gly Thr Pro Gln Arg Pro Gly Ala His Gly Ser
    600                 605                 610

tcg gca gcc gct cac cga tac tcg cgg gag gct aac gct ccg ggc ccc       3659
Ser Ala Ala Ala His Arg Tyr Ser Arg Glu Ala Asn Ala Pro Gly Pro
615                 620                 625

gta ccc gga gag cgg caa ctc gcg cac tca aag atg gtc ccc atc cct       3707
```

```
Val Pro Gly Glu Arg Gln Leu Ala His Ser Lys Met Val Pro Ile Pro
630                 635                 640                 645

gct gga gta ttt aca atg ggc aca gat gat cct cag ata aag cag gat       3755
Ala Gly Val Phe Thr Met Gly Thr Asp Asp Pro Gln Ile Lys Gln Asp
                650                 655                 660

ggg gaa gca cct gcg agg aga gtt act att gat gcc ttt tac atg gat       3803
Gly Glu Ala Pro Ala Arg Arg Val Thr Ile Asp Ala Phe Tyr Met Asp
            665                 670                 675

gcc tat gaa gtc agt aat act gaa ttt gag aag ttt gtg aac tca act       3851
Ala Tyr Glu Val Ser Asn Thr Glu Phe Glu Lys Phe Val Asn Ser Thr
        680                 685                 690

ggc tat ttg aca gag gct gag aag ttt ggc gac tcc ttt gtc ttt gaa       3899
Gly Tyr Leu Thr Glu Ala Glu Lys Phe Gly Asp Ser Phe Val Phe Glu
    695                 700                 705

ggc atg ttg agt gag caa gtg aag acc aat att caa cag gca gtt gca       3947
Gly Met Leu Ser Glu Gln Val Lys Thr Asn Ile Gln Gln Ala Val Ala
710                 715                 720                 725

gct gct ccc tgg tgg tta cct gtg aaa ggc gct aac tgg aga cac cca       3995
Ala Ala Pro Trp Trp Leu Pro Val Lys Gly Ala Asn Trp Arg His Pro
                730                 735                 740

gaa ggg cct gac tct act att ctg cac agg ccg gat cat cca gtt ctc       4043
Glu Gly Pro Asp Ser Thr Ile Leu His Arg Pro Asp His Pro Val Leu
            745                 750                 755

cat gtg tcc tgg aat gat gcg gtt gcc tac tgc act tgg gca ggg aag       4091
His Val Ser Trp Asn Asp Ala Val Ala Tyr Cys Thr Trp Ala Gly Lys
        760                 765                 770

cgg ctg ccc acg gaa gct gag tgg gaa tac agc tgt cga gga ggc ctg       4139
Arg Leu Pro Thr Glu Ala Glu Trp Glu Tyr Ser Cys Arg Gly Gly Leu
    775                 780                 785

cat aat aga ctt ttc ccc tgg ggc aac aaa ctg cag ccc aaa ggc cag       4187
His Asn Arg Leu Phe Pro Trp Gly Asn Lys Leu Gln Pro Lys Gly Gln
790                 795                 800                 805

cat tat gcc aac att tgg cag ggc gag ttt ccg gtg acc aac act ggt       4235
His Tyr Ala Asn Ile Trp Gln Gly Glu Phe Pro Val Thr Asn Thr Gly
                810                 815                 820

gag gat ggc ttc caa gga act gcg cct gtt gat gcc ttc cct ccc aat       4283
Glu Asp Gly Phe Gln Gly Thr Ala Pro Val Asp Ala Phe Pro Pro Asn
            825                 830                 835

ggt tat ggc tta tac aac ata gtg ggg aac gca tgg gaa tgg act tca       4331
Gly Tyr Gly Leu Tyr Asn Ile Val Gly Asn Ala Trp Glu Trp Thr Ser
        840                 845                 850

gac tgg tgg act gtt cat cat tct gtt gaa gaa acg ctt aac cca aaa       4379
Asp Trp Trp Thr Val His His Ser Val Glu Glu Thr Leu Asn Pro Lys
    855                 860                 865

ggt ccc cct tct ggg aaa gac cga gtg aag aaa ggt gga tcc tac atg       4427
Gly Pro Pro Ser Gly Lys Asp Arg Val Lys Lys Gly Gly Ser Tyr Met
870                 875                 880                 885

tgc cat agg tct tat tgt tac agg tat cgc tgt gct gct cgg agc cag       4475
Cys His Arg Ser Tyr Cys Tyr Arg Tyr Arg Cys Ala Ala Arg Ser Gln
                890                 895                 900

aac aca cct gat agc tct gct tcg aat ctg gga ttc cgc tgt gca gcc       4523
```

Asn Thr Pro Asp Ser Ser Ala Ser Asn Leu Gly Phe Arg Cys Ala Ala
905 910 915

gac cgc ctg ccc acc atg gac tga gaattcacgc gtggtacctc tagagtcgac 4577
Asp Arg Leu Pro Thr Met Asp
920

ccgggcggcc tcgaggacgg ggtgaactac gcctgaggat ccgatctttt tccctctgcc 4637

aaaaattatg gggacatcat gaagcccctt gagcatctga cttctggcta ataaaggaaa 4697

tttattttca ttgcaatagt gtgttggaat tttttgtgtc tctcactcgg aagcaattcg 4757

ttgatctgaa tttcgaccac ccataatacc cattaccctg gtagataagt agcatggcgg 4817

gttaatcatt aactacaagg aacccctagt gatggagttg gccactccct ctctgcgcgc 4877

tcgctcgctc actgaggccg ggcgaccaaa ggtcgcccga cgcccgggct ttgcccgggc 4937

ggcctcagtg agcgagcgag cgcgcag 4964

<210> 6
<211> 550
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 6

Met Pro Pro Pro Arg Thr Gly Arg Gly Leu Leu Trp Leu Gly Leu Val
1 5 10 15

Leu Ser Ser Val Cys Val Ala Leu Gly Ser Glu Thr Gln Ala Asn Ser
20 25 30

Thr Thr Asp Ala Leu Asn Val Leu Leu Ile Ile Val Asp Asp Leu Arg
35 40 45

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

Asp Gln Leu Ala Ser His Ser Leu Leu Phe Gln Asn Ala Phe Ala Gln
65 70 75 80

Gln Ala Val Cys Ala Pro Ser Arg Val Ser Phe Leu Thr Gly Arg Arg
85 90 95

Pro Asp Thr Thr Arg Leu Tyr Asp Phe Asn Ser Tyr Trp Arg Val His
100 105 110

Ala Gly Asn Phe Ser Thr Ile Pro Gln Tyr Phe Lys Glu Asn Gly Tyr
115 120 125

Val Thr Met Ser Val Gly Lys Val Phe His Pro Gly Ile Ser Ser Asn
130 135 140

```
His Thr Asp Asp Ser Pro Tyr Ser Trp Ser Phe Pro Pro Tyr His Pro
145                 150                 155                 160

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

Glu Leu His Ala Asn Leu Leu Cys Pro Val Asp Val Leu Asp Val Pro
            180                 185                 190

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

Leu Glu Lys Met Lys Thr Ser Ala Ser Pro Phe Phe Leu Ala Val Gly
    210                 215                 220

Tyr His Lys Pro His Ile Pro Phe Arg Tyr Pro Lys Glu Phe Gln Lys
225                 230                 235                 240

Leu Tyr Pro Leu Glu Asn Ile Thr Leu Ala Pro Asp Pro Glu Val Pro
                245                 250                 255

Asp Gly Leu Pro Pro Val Ala Tyr Asn Pro Trp Met Asp Ile Arg Gln
            260                 265                 270

Arg Glu Asp Val Gln Ala Leu Asn Ile Ser Val Pro Tyr Gly Pro Ile
        275                 280                 285

Pro Val Asp Phe Gln Arg Lys Ile Arg Gln Ser Tyr Phe Ala Ser Val
    290                 295                 300

Ser Tyr Leu Asp Thr Gln Val Gly Arg Leu Leu Ser Ala Leu Asp Asp
305                 310                 315                 320

Leu Gln Leu Ala Asn Ser Thr Ile Ile Ala Phe Thr Ser Asp His Gly
                325                 330                 335

Trp Ala Leu Gly Glu His Gly Glu Trp Ala Lys Tyr Ser Asn Phe Asp
            340                 345                 350

Val Ala Thr His Val Pro Leu Ile Phe Tyr Val Pro Gly Arg Thr Ala
        355                 360                 365

Ser Leu Pro Glu Ala Gly Glu Lys Leu Phe Pro Tyr Leu Asp Pro Phe
    370                 375                 380

Asp Ser Ala Ser Gln Leu Met Glu Pro Gly Arg Gln Ser Met Asp Leu
385                 390                 395                 400

Val Glu Leu Val Ser Leu Phe Pro Thr Leu Ala Gly Leu Ala Gly Leu
                405                 410                 415
```

Gln Val Pro Pro Arg Cys Pro Val Pro Ser Phe His Val Glu Leu Cys
420 425 430

Arg Glu Gly Lys Asn Leu Leu Lys His Phe Arg Phe Arg Asp Leu Glu
435 440 445

Glu Asp Pro Tyr Leu Pro Gly Asn Pro Arg Glu Leu Ile Ala Tyr Ser
450 455 460

Gln Tyr Pro Arg Pro Ser Asp Ile Pro Gln Trp Asn Ser Asp Lys Pro
465 470 475 480

Ser Leu Lys Asp Ile Lys Ile Met Gly Tyr Ser Ile Arg Thr Ile Asp
485 490 495

Tyr Arg Tyr Thr Val Trp Val Gly Phe Asn Pro Asp Glu Phe Leu Ala
500 505 510

Asn Phe Ser Asp Ile His Ala Gly Glu Leu Tyr Phe Val Asp Ser Asp
515 520 525

Pro Leu Gln Asp His Asn Met Tyr Asn Asp Ser Gln Gly Gly Asp Leu
530 535 540

Phe Gln Leu Leu Met Pro
545 550

<210> 7
<211> 374
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 7

Met Ala Ala Pro Ala Leu Gly Leu Val Cys Gly Arg Cys Pro Glu Leu
1 5 10 15

Gly Leu Val Leu Leu Leu Leu Leu Leu Ser Leu Leu Cys Gly Ala Ala
20 25 30

Gly Ser Gln Glu Ala Gly Thr Gly Ala Gly Ala Gly Ser Leu Ala Gly
35 40 45

Ser Cys Gly Cys Gly Thr Pro Gln Arg Pro Gly Ala His Gly Ser Ser
50 55 60

Ala Ala Ala His Arg Tyr Ser Arg Glu Ala Asn Ala Pro Gly Pro Val
65 70 75 80

Pro Gly Glu Arg Gln Leu Ala His Ser Lys Met Val Pro Ile Pro Ala
85 90 95

```
Gly Val Phe Thr Met Gly Thr Asp Asp Pro Gln Ile Lys Gln Asp Gly
            100                 105                 110

Glu Ala Pro Ala Arg Arg Val Thr Ile Asp Ala Phe Tyr Met Asp Ala
        115                 120                 125

Tyr Glu Val Ser Asn Thr Glu Phe Glu Lys Phe Val Asn Ser Thr Gly
    130                 135                 140

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

Met Leu Ser Glu Gln Val Lys Thr Asn Ile Gln Gln Ala Val Ala Ala
                165                 170                 175

Ala Pro Trp Trp Leu Pro Val Lys Gly Ala Asn Trp Arg His Pro Glu
            180                 185                 190

Gly Pro Asp Ser Thr Ile Leu His Arg Pro Asp His Pro Val Leu His
        195                 200                 205

Val Ser Trp Asn Asp Ala Val Ala Tyr Cys Thr Trp Ala Gly Lys Arg
    210                 215                 220

Leu Pro Thr Glu Ala Glu Trp Glu Tyr Ser Cys Arg Gly Gly Leu His
225                 230                 235                 240

Asn Arg Leu Phe Pro Trp Gly Asn Lys Leu Gln Pro Lys Gly Gln His
                245                 250                 255

Tyr Ala Asn Ile Trp Gln Gly Glu Phe Pro Val Thr Asn Thr Gly Glu
            260                 265                 270

Asp Gly Phe Gln Gly Thr Ala Pro Val Asp Ala Phe Pro Pro Asn Gly
        275                 280                 285

Tyr Gly Leu Tyr Asn Ile Val Gly Asn Ala Trp Glu Trp Thr Ser Asp
    290                 295                 300

Trp Trp Thr Val His His Ser Val Glu Glu Thr Leu Asn Pro Lys Gly
305                 310                 315                 320

Pro Pro Ser Gly Lys Asp Arg Val Lys Lys Gly Gly Ser Tyr Met Cys
                325                 330                 335

His Arg Ser Tyr Cys Tyr Arg Tyr Arg Cys Ala Ala Arg Ser Gln Asn
            340                 345                 350

Thr Pro Asp Ser Ser Ala Ser Asn Leu Gly Phe Arg Cys Ala Ala Asp
        355                 360                 365
```

Arg Leu Pro Thr Met Asp
370

<210> 8
<211> 4964
<212> DNA
<213> Artificial Sequence

<220>
<223> CB6.hIDSco.IRES.hSUMF1co

<220>
<221> misc_feature
<222> (11)..(140)
<223> 5'ITR

<220>
<221> enhancer
<222> (208)..(589)
<223> CMV IE enhancer

<220>
<221> promoter
<222> (600)..(859)
<223> CB promoter

<220>
<221> TATA_signal
<222> (835)..(842)

<220>
<221> Intron
<222> (976)..(1108)

<220>
<221> CDS
<222> (1177)..(2829)
<223> hIDSco

<220>
<221> misc_feature
<222> (2830)..(3422)
<223> IRES

<220>
<221> CDS
<222> (3423)..(4553)
<223> hSUMF1co

<220>
<221> polyA_signal
<222> (4620)..(4746)
<223> rabbit globin polyA

<220>
<221> misc_feature
<222> (4835)..(4964)
<223> 3'ITR

<400> 8
cttaattagg ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg 60

ggcgaccttt ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa 120

```
ctccatcact aggggttcct tgtagttaat gattaacccg ccatgctact tatctaccag      180

ggtaatgggg atcctctaga actatagcta gtcgacattg attattgact agttattaat      240

agtaatcaat tacggggtca ttagttcata gcccatatat ggagttccgc gttacataac      300

ttacggtaaa tggcccgcct ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa      360

tgacgtatgt tcccatagta acgccaatag ggactttcca ttgacgtcaa tgggtggact      420

atttacggta aactgcccac ttggcagtac atcaagtgta tcatatgcca agtacgcccc      480

ctattgacgt caatgacggt aaatggcccg cctggcatta tgcccagtac atgaccttat      540

gggactttcc tacttggcag tacatctacg tattagtcat cgctattacc atgtcgaggc      600

cacgttctgc ttcactctcc ccatctcccc cccctcccca cccccaattt tgtatttatt      660

tatttttaa ttattttgtg cagcgatggg ggcgggggg gggggcgcgc gccaggcggg      720

gcggggcggg gcgaggggcg gggcggggcg aggcggagag gtgcggcggc agccaatcag      780

agcggcgcgc tccgaaagtt tccttttatg gcgaggcggc ggcggcggcg gccctataaa      840

aagcgaagcg cgcggcgggc gggagcaagc tttattgcgg tagtttatca cagttaaatt      900

gctaacgcag tcagtgcttc tgacacaaca gtctcgaact taagctgcag aagttggtcg      960

tgaggcactg ggcaggtaag tatcaaggtt acaagacagg tttaaggaga ccaatagaaa     1020

ctgggcttgt cgagacagag aagactcttg cgtttctgat aggcacctat tggtcttact     1080

gacatccact ttgcctttct ctccacaggt gtccactccc agttcaatta cagctcttaa     1140

ggctagagta cttaatacga ctcactatag gctagc atg ccc ccc ccc agg aca       1194
                                        Met Pro Pro Pro Arg Thr
                                        1               5

gga aga gga ctg ctg tgg ctc gga ctc gtc ctc tcc agc gtg tgc gtg       1242
Gly Arg Gly Leu Leu Trp Leu Gly Leu Val Leu Ser Ser Val Cys Val
            10                  15                  20

gct ctc gga tcc gaa acc cag gct aat tcc aca acc gat gcc ctg aac       1290
Ala Leu Gly Ser Glu Thr Gln Ala Asn Ser Thr Thr Asp Ala Leu Asn
        25                  30                  35

gtc ctg ctc atc atc gtg gat gac ctg agg ccc tcc ctg ggc tgt tac       1338
Val Leu Leu Ile Ile Val Asp Asp Leu Arg Pro Ser Leu Gly Cys Tyr
    40                  45                  50

ggc gac aag ctc gtc agg tcc ccc aac att gat cag ctc gcc tcc cat       1386
Gly Asp Lys Leu Val Arg Ser Pro Asn Ile Asp Gln Leu Ala Ser His
55                  60                  65                  70

tcc ctg ctg ttt caa aac gcc ttc gct cag caa gcc gtg tgc gct cct       1434
Ser Leu Leu Phe Gln Asn Ala Phe Ala Gln Gln Ala Val Cys Ala Pro
                75                  80                  85

agc aga gtc agc ttc ctc acc gga agg aga cct gac acc acc agg ctg       1482
Ser Arg Val Ser Phe Leu Thr Gly Arg Arg Pro Asp Thr Thr Arg Leu
            90                  95                  100

tac gac ttc aac tcc tat tgg agg gtc cac gct ggc aac ttc agc acc       1530
Tyr Asp Phe Asn Ser Tyr Trp Arg Val His Ala Gly Asn Phe Ser Thr
        105                 110                 115

att ccc cag tac ttc aag gag aac ggc tac gtc acc atg tcc gtc ggc       1578
Ile Pro Gln Tyr Phe Lys Glu Asn Gly Tyr Val Thr Met Ser Val Gly
```

```
    120                 125                 130

aag gtg ttc cac cct ggc att tcc tcc aac cac acc gat gac agc ccc       1626
Lys Val Phe His Pro Gly Ile Ser Ser Asn His Thr Asp Asp Ser Pro
135                 140                 145                 150

tac tcc tgg tcc ttc cct ccc tac cac ccc agc agc gaa aag tac gag       1674
Tyr Ser Trp Ser Phe Pro Pro Tyr His Pro Ser Ser Glu Lys Tyr Glu
                155                 160                 165

aat acc aag acc tgt aga ggc ccc gac ggc gaa ctc cac gct aac ctc       1722
Asn Thr Lys Thr Cys Arg Gly Pro Asp Gly Glu Leu His Ala Asn Leu
            170                 175                 180

ctg tgt cct gtc gat gtc ctc gac gtg ccc gaa ggc aca ctc ccc gac       1770
Leu Cys Pro Val Asp Val Leu Asp Val Pro Glu Gly Thr Leu Pro Asp
        185                 190                 195

aag cag tcc aca gaa caa gcc atc cag ctc ctg gag aag atg aag acc       1818
Lys Gln Ser Thr Glu Gln Ala Ile Gln Leu Leu Glu Lys Met Lys Thr
    200                 205                 210

tcc gct tcc ccc ttt ttc ctg gcc gtg ggc tac cac aaa ccc cac att       1866
Ser Ala Ser Pro Phe Phe Leu Ala Val Gly Tyr His Lys Pro His Ile
215                 220                 225                 230

ccc ttc agg tac ccc aag gag ttc cag aag ctg tat ccc ctg gag aac       1914
Pro Phe Arg Tyr Pro Lys Glu Phe Gln Lys Leu Tyr Pro Leu Glu Asn
                235                 240                 245

att acc ctc gcc cct gat ccc gaa gtg cct gac ggc ctg cct cct gtc       1962
Ile Thr Leu Ala Pro Asp Pro Glu Val Pro Asp Gly Leu Pro Pro Val
            250                 255                 260

gct tac aac ccc tgg atg gac atc agg cag aga gaa gac gtg caa gcc       2010
Ala Tyr Asn Pro Trp Met Asp Ile Arg Gln Arg Glu Asp Val Gln Ala
        265                 270                 275

ctc aac atc agc gtg ccc tac ggc cct atc ccc gtg gac ttc cag aga       2058
Leu Asn Ile Ser Val Pro Tyr Gly Pro Ile Pro Val Asp Phe Gln Arg
    280                 285                 290

aag atc agg cag agc tac ttc gcc tcc gtc agc tac ctg gac acc cag       2106
Lys Ile Arg Gln Ser Tyr Phe Ala Ser Val Ser Tyr Leu Asp Thr Gln
295                 300                 305                 310

gtg gga aga ctc ctg tcc gcc ctc gac gat ctc cag ctg gcc aac agc       2154
Val Gly Arg Leu Leu Ser Ala Leu Asp Asp Leu Gln Leu Ala Asn Ser
                315                 320                 325

acc atc att gcc ttc acc agc gac cat ggc tgg gct ctc gga gag cat       2202
Thr Ile Ile Ala Phe Thr Ser Asp His Gly Trp Ala Leu Gly Glu His
            330                 335                 340

ggc gaa tgg gcc aag tac tcc aac ttc gat gtg gcc acc cac gtc ccc       2250
Gly Glu Trp Ala Lys Tyr Ser Asn Phe Asp Val Ala Thr His Val Pro
        345                 350                 355

ctc atc ttc tac gtc ccc gga agg acc gcc agc ctg cct gaa gct gga       2298
Leu Ile Phe Tyr Val Pro Gly Arg Thr Ala Ser Leu Pro Glu Ala Gly
    360                 365                 370

gag aag ctc ttc ccc tac ctg gac cct ttt gat tcc gcc agc caa ctc       2346
Glu Lys Leu Phe Pro Tyr Leu Asp Pro Phe Asp Ser Ala Ser Gln Leu
375                 380                 385                 390

atg gaa ccc ggc agg caa agc atg gac ctc gtg gag ctg gtc tcc ctg       2394
Met Glu Pro Gly Arg Gln Ser Met Asp Leu Val Glu Leu Val Ser Leu
```

395 400 405

ttc cct acc ctg gct gga ctc gcc gga ctc caa gtc cct ccc aga tgc 2442
Phe Pro Thr Leu Ala Gly Leu Ala Gly Leu Gln Val Pro Pro Arg Cys
410 415 420

cct gtg cct agc ttc cat gtg gag ctc tgc aga gag ggc aag aac ctg 2490
Pro Val Pro Ser Phe His Val Glu Leu Cys Arg Glu Gly Lys Asn Leu
425 430 435

ctg aag cac ttc agg ttt agg gac ctg gag gag gat cct tac ctc ccc 2538
Leu Lys His Phe Arg Phe Arg Asp Leu Glu Glu Asp Pro Tyr Leu Pro
440 445 450

gga aac cct aga gag ctc atc gct tac tcc cag tat cct agg ccc tcc 2586
Gly Asn Pro Arg Glu Leu Ile Ala Tyr Ser Gln Tyr Pro Arg Pro Ser
455 460 465 470

gac atc ccc cag tgg aac agc gat aag ccc agc ctg aag gat atc aaa 2634
Asp Ile Pro Gln Trp Asn Ser Asp Lys Pro Ser Leu Lys Asp Ile Lys
475 480 485

atc atg gga tac tcc att agg aca atc gac tac agg tac acc gtc tgg 2682
Ile Met Gly Tyr Ser Ile Arg Thr Ile Asp Tyr Arg Tyr Thr Val Trp
490 495 500

gtc gga ttc aac ccc gat gag ttc ctg gcc aac ttc tcc gac atc cac 2730
Val Gly Phe Asn Pro Asp Glu Phe Leu Ala Asn Phe Ser Asp Ile His
505 510 515

gct ggc gag ctg tat ttc gtg gac agc gac cct ctg cag gac cac aac 2778
Ala Gly Glu Leu Tyr Phe Val Asp Ser Asp Pro Leu Gln Asp His Asn
520 525 530

atg tac aac gac tcc cag gga ggc gat ctg ttt caa ctc ctg atg ccc 2826
Met Tyr Asn Asp Ser Gln Gly Gly Asp Leu Phe Gln Leu Leu Met Pro
535 540 545 550

tga gcccctctcc ctccccccc cctaacgtta ctggccgaag ccgcttggaa 2879

taaggccggt gtgcgtttgt ctatatgtta ttttccacca tattgccgtc ttttggcaat 2939

gtgagggccc ggaaacctgg ccctgtcttc ttgacgagca ttcctagggg tctttcccct 2999

ctcgccaaag gaatgcaagg tctgttgaat gtcgtgaagg aagcagttcc tctggaagct 3059

tcttgaagac aaacaacgtc tgtagcgacc ctttgcaggc agcggaaccc cccacctggc 3119

gacaggtgcc tctgcggcca aaagccacgt gtataagata cacctgcaaa ggcggcacaa 3179

ccccagtgcc acgttgtgag ttggatagtt gtggaaagag tcaaatggct ctcctcaagc 3239

gtattcaaca aggggctgaa ggatgcccag aaggtacccc attgtatggg atctgatctg 3299

gggcctcggt acacatgctt tacatgtgtt tagtcgaggt taaaaaaacg tctaggcccc 3359

ccgaaccacg gggacgtggt tttcctttga aaaacacgat gataatatgg ccacaaccat 3419

ggt atg gcc gcc cct gcc ctc gga ctg gtg tgt ggc agg tgc cct gag 3467
Met Ala Ala Pro Ala Leu Gly Leu Val Cys Gly Arg Cys Pro Glu
555 560 565

ctg gga ctg gtg ctg ctg ctg ctc ctg ctg agc ctc ctg tgt ggc gct 3515
Leu Gly Leu Val Leu Leu Leu Leu Leu Leu Ser Leu Leu Cys Gly Ala
570 575 580

gcc gga agc caa gag gct gga aca gga gct ggc gcc gga agc ctg gct 3563
Ala Gly Ser Gln Glu Ala Gly Thr Gly Ala Gly Ala Gly Ser Leu Ala

585 590 595

ggc agc tgc gga tgt ggc aca ccc cag agg cct gga gct cat gga agc 3611
Gly Ser Cys Gly Cys Gly Thr Pro Gln Arg Pro Gly Ala His Gly Ser
600 605 610

agc gcc gct gcc cac agg tac tcc aga gaa gcc aat gcc cct gga cct 3659
Ser Ala Ala Ala His Arg Tyr Ser Arg Glu Ala Asn Ala Pro Gly Pro
615 620 625

gtg ccc ggc gaa aga cag ctg gcc cac tcc aag atg gtg ccc atc cct 3707
Val Pro Gly Glu Arg Gln Leu Ala His Ser Lys Met Val Pro Ile Pro
630 635 640 645

gcc gga gtg ttt aca atg ggc acc gat gac cct cag atc aag cag gat 3755
Ala Gly Val Phe Thr Met Gly Thr Asp Asp Pro Gln Ile Lys Gln Asp
650 655 660

gga gag gct ccc gcc agg agg gtg aca atc gac gcc ttc tac atg gac 3803
Gly Glu Ala Pro Ala Arg Arg Val Thr Ile Asp Ala Phe Tyr Met Asp
665 670 675

gcc tac gag gtg agc aac acc gag ttc gaa aaa ttc gtg aac agc acc 3851
Ala Tyr Glu Val Ser Asn Thr Glu Phe Glu Lys Phe Val Asn Ser Thr
680 685 690

ggc tac ctg acc gag gcc gaa aag ttc ggc gat agc ttc gtc ttc gag 3899
Gly Tyr Leu Thr Glu Ala Glu Lys Phe Gly Asp Ser Phe Val Phe Glu
695 700 705

ggc atg ctg agc gag cag gtg aag acc aac atc cag cag gcc gtg gcc 3947
Gly Met Leu Ser Glu Gln Val Lys Thr Asn Ile Gln Gln Ala Val Ala
710 715 720 725

gct gct cct tgg tgg ctg ccc gtg aag ggc gcc aat tgg agg cat ccc 3995
Ala Ala Pro Trp Trp Leu Pro Val Lys Gly Ala Asn Trp Arg His Pro
730 735 740

gag ggc ccc gat agc acc atc ctg cac agg ccc gac cat cct gtg ctg 4043
Glu Gly Pro Asp Ser Thr Ile Leu His Arg Pro Asp His Pro Val Leu
745 750 755

cac gtc agc tgg aac gac gct gtg gcc tac tgc aca tgg gcc ggc aaa 4091
His Val Ser Trp Asn Asp Ala Val Ala Tyr Cys Thr Trp Ala Gly Lys
760 765 770

agg ctg cct acc gaa gcc gag tgg gag tac agc tgc agg ggc ggc ctg 4139
Arg Leu Pro Thr Glu Ala Glu Trp Glu Tyr Ser Cys Arg Gly Gly Leu
775 780 785

cac aat agg ctc ttc ccc tgg ggc aac aaa ctc cag ccc aag ggc cag 4187
His Asn Arg Leu Phe Pro Trp Gly Asn Lys Leu Gln Pro Lys Gly Gln
790 795 800 805

cat tat gcc aat atc tgg cag ggc gag ttc cct gtg aca aac acc ggc 4235
His Tyr Ala Asn Ile Trp Gln Gly Glu Phe Pro Val Thr Asn Thr Gly
810 815 820

gag gat ggc ttt caa ggc acc gct ccc gtg gac gcc ttt ccc ccc aat 4283
Glu Asp Gly Phe Gln Gly Thr Ala Pro Val Asp Ala Phe Pro Pro Asn
825 830 835

ggc tac ggc ctg tac aat atc gtg ggc aac gcc tgg gag tgg aca tcc 4331
Gly Tyr Gly Leu Tyr Asn Ile Val Gly Asn Ala Trp Glu Trp Thr Ser
840 845 850

gac tgg tgg acc gtg cat cac agc gtg gag gaa aca ctc aac ccc aag 4379
Asp Trp Trp Thr Val His His Ser Val Glu Glu Thr Leu Asn Pro Lys

```
    855                 860                 865

ggc cct ccc agc gga aag gat agg gtc aag aag ggc ggc agc tac atg       4427
Gly Pro Pro Ser Gly Lys Asp Arg Val Lys Lys Gly Gly Ser Tyr Met
870                 875                 880                 885

tgc cac aga tcc tac tgc tac aga tac agg tgc gcc gcc agg agc cag       4475
Cys His Arg Ser Tyr Cys Tyr Arg Tyr Arg Cys Ala Ala Arg Ser Gln
                890                 895                 900

aac acc ccc gat agc agc gcc tcc aat ctg ggc ttt agg tgt gcc gcc       4523
Asn Thr Pro Asp Ser Ser Ala Ser Asn Leu Gly Phe Arg Cys Ala Ala
            905                 910                 915

gat agg ctg ccc acc atg gac tga gaa ttc acgcgtggta cctctagagt         4573
Asp Arg Leu Pro Thr Met Asp     Glu Phe
        920                 925

cgacccgggc ggcctcgagg acggggtgaa ctacgcctga ggatccgatc tttttccctc     4633

tgccaaaaat tatggggaca tcatgaagcc ccttgagcat ctgacttctg gctaataaag     4693

gaaatttatt ttcattgcaa tagtgtgttg gaatttttg tgtctctcac tcggaagcaa     4753

ttcgttgatc tgaatttcga ccacccataa tacccattac cctggtagat aagtagcatg     4813

gcgggttaat cattaactac aaggaacccc tagtgatgga gttggccact ccctctctgc     4873

gcgctcgctc gctcactgag gccgggcgac caaaggtcgc ccgacgcccg ggctttgccc     4933

gggcggcctc agtgagcgag cgagcgcgca g                                    4964


<210>  9
<211>  550
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic Construct

<400>  9

Met Pro Pro Pro Arg Thr Gly Arg Gly Leu Leu Trp Leu Gly Leu Val
1               5                   10                  15


Leu Ser Ser Val Cys Val Ala Leu Gly Ser Glu Thr Gln Ala Asn Ser
            20                  25                  30


Thr Thr Asp Ala Leu Asn Val Leu Leu Ile Ile Val Asp Asp Leu Arg
        35                  40                  45


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


Asp Gln Leu Ala Ser His Ser Leu Leu Phe Gln Asn Ala Phe Ala Gln
65                  70                  75                  80


Gln Ala Val Cys Ala Pro Ser Arg Val Ser Phe Leu Thr Gly Arg Arg
                85                  90                  95


Pro Asp Thr Thr Arg Leu Tyr Asp Phe Asn Ser Tyr Trp Arg Val His
```

100 105 110

Ala Gly Asn Phe Ser Thr Ile Pro Gln Tyr Phe Lys Glu Asn Gly Tyr
115 120 125

Val Thr Met Ser Val Gly Lys Val Phe His Pro Gly Ile Ser Ser Asn
130 135 140

His Thr Asp Asp Ser Pro Tyr Ser Trp Ser Phe Pro Pro Tyr His Pro
145 150 155 160

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

Glu Leu His Ala Asn Leu Leu Cys Pro Val Asp Val Leu Asp Val Pro
180 185 190

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

Leu Glu Lys Met Lys Thr Ser Ala Ser Pro Phe Phe Leu Ala Val Gly
210 215 220

Tyr His Lys Pro His Ile Pro Phe Arg Tyr Pro Lys Glu Phe Gln Lys
225 230 235 240

Leu Tyr Pro Leu Glu Asn Ile Thr Leu Ala Pro Asp Pro Glu Val Pro
245 250 255

Asp Gly Leu Pro Pro Val Ala Tyr Asn Pro Trp Met Asp Ile Arg Gln
260 265 270

Arg Glu Asp Val Gln Ala Leu Asn Ile Ser Val Pro Tyr Gly Pro Ile
275 280 285

Pro Val Asp Phe Gln Arg Lys Ile Arg Gln Ser Tyr Phe Ala Ser Val
290 295 300

Ser Tyr Leu Asp Thr Gln Val Gly Arg Leu Leu Ser Ala Leu Asp Asp
305 310 315 320

Leu Gln Leu Ala Asn Ser Thr Ile Ile Ala Phe Thr Ser Asp His Gly
325 330 335

Trp Ala Leu Gly Glu His Gly Glu Trp Ala Lys Tyr Ser Asn Phe Asp
340 345 350

Val Ala Thr His Val Pro Leu Ile Phe Tyr Val Pro Gly Arg Thr Ala
355 360 365

Ser Leu Pro Glu Ala Gly Glu Lys Leu Phe Pro Tyr Leu Asp Pro Phe

370 375 380

Asp Ser Ala Ser Gln Leu Met Glu Pro Gly Arg Gln Ser Met Asp Leu
385 390 395 400

Val Glu Leu Val Ser Leu Phe Pro Thr Leu Ala Gly Leu Ala Gly Leu
405 410 415

Gln Val Pro Pro Arg Cys Pro Val Pro Ser Phe His Val Glu Leu Cys
420 425 430

Arg Glu Gly Lys Asn Leu Leu Lys His Phe Arg Phe Arg Asp Leu Glu
435 440 445

Glu Asp Pro Tyr Leu Pro Gly Asn Pro Arg Glu Leu Ile Ala Tyr Ser
450 455 460

Gln Tyr Pro Arg Pro Ser Asp Ile Pro Gln Trp Asn Ser Asp Lys Pro
465 470 475 480

Ser Leu Lys Asp Ile Lys Ile Met Gly Tyr Ser Ile Arg Thr Ile Asp
485 490 495

Tyr Arg Tyr Thr Val Trp Val Gly Phe Asn Pro Asp Glu Phe Leu Ala
500 505 510

Asn Phe Ser Asp Ile His Ala Gly Glu Leu Tyr Phe Val Asp Ser Asp
515 520 525

Pro Leu Gln Asp His Asn Met Tyr Asn Asp Ser Gln Gly Gly Asp Leu
530 535 540

Phe Gln Leu Leu Met Pro
545 550

<210> 10
<211> 374
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 10

Met Ala Ala Pro Ala Leu Gly Leu Val Cys Gly Arg Cys Pro Glu Leu
1 5 10 15

Gly Leu Val Leu Leu Leu Leu Leu Leu Ser Leu Leu Cys Gly Ala Ala
20 25 30

Gly Ser Gln Glu Ala Gly Thr Gly Ala Gly Ala Gly Ser Leu Ala Gly
35 40 45

```
Ser Cys Gly Cys Gly Thr Pro Gln Arg Pro Gly Ala His Gly Ser Ser
    50                  55                  60

Ala Ala Ala His Arg Tyr Ser Arg Glu Ala Asn Ala Pro Gly Pro Val
65                  70                  75                  80

Pro Gly Glu Arg Gln Leu Ala His Ser Lys Met Val Pro Ile Pro Ala
                85                  90                  95

Gly Val Phe Thr Met Gly Thr Asp Asp Pro Gln Ile Lys Gln Asp Gly
            100                 105                 110

Glu Ala Pro Ala Arg Arg Val Thr Ile Asp Ala Phe Tyr Met Asp Ala
        115                 120                 125

Tyr Glu Val Ser Asn Thr Glu Phe Glu Lys Phe Val Asn Ser Thr Gly
    130                 135                 140

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

Met Leu Ser Glu Gln Val Lys Thr Asn Ile Gln Gln Ala Val Ala Ala
                165                 170                 175

Ala Pro Trp Trp Leu Pro Val Lys Gly Ala Asn Trp Arg His Pro Glu
            180                 185                 190

Gly Pro Asp Ser Thr Ile Leu His Arg Pro Asp His Pro Val Leu His
        195                 200                 205

Val Ser Trp Asn Asp Ala Val Ala Tyr Cys Thr Trp Ala Gly Lys Arg
    210                 215                 220

Leu Pro Thr Glu Ala Glu Trp Glu Tyr Ser Cys Arg Gly Gly Leu His
225                 230                 235                 240

Asn Arg Leu Phe Pro Trp Gly Asn Lys Leu Gln Pro Lys Gly Gln His
                245                 250                 255

Tyr Ala Asn Ile Trp Gln Gly Glu Phe Pro Val Thr Asn Thr Gly Glu
            260                 265                 270

Asp Gly Phe Gln Gly Thr Ala Pro Val Asp Ala Phe Pro Pro Asn Gly
        275                 280                 285

Tyr Gly Leu Tyr Asn Ile Val Gly Asn Ala Trp Glu Trp Thr Ser Asp
    290                 295                 300

Trp Trp Thr Val His His Ser Val Glu Glu Thr Leu Asn Pro Lys Gly
305                 310                 315                 320
```

Pro Pro Ser Gly Lys Asp Arg Val Lys Lys Gly Gly Ser Tyr Met Cys
325 330 335

His Arg Ser Tyr Cys Tyr Arg Tyr Arg Cys Ala Ala Arg Ser Gln Asn
340 345 350

Thr Pro Asp Ser Ser Ala Ser Asn Leu Gly Phe Arg Cys Ala Ala Asp
355 360 365

Arg Leu Pro Thr Met Asp
370

<210> 11
<211> 3967
<212> DNA
<213> Artificial Sequence

<220>
<223> CB7.hIDSco.RBG

<220>
<221> misc_feature
<222> (2)..(131)
<223> 5'ITR

<220>
<221> promoter
<222> (199)..(580)
<223> CMV IE promoter

<220>
<221> promoter
<222> (583)..(864)
<223> CB promoter

<220>
<221> TATA_signal
<222> (837)..(840)

<220>
<221> Intron
<222> (959)..(1930)
<223> chicken beta-actin intron

<220>
<221> CDS
<222> (1937)..(3589)
<223> hIDSco

<220>
<221> polyA_signal
<222> (3623)..(3749)
<223> rabbit globin polyA

<220>
<221> misc_feature
<222> (3838)..(3967)
<223> 3' ITR

<400> 11
gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc gggcgacctt 60

tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca actccatcac 120

taggggttcc ttgtagttaa tgattaaccc gccatgctac ttatctacca gggtaatggg 180

gatcctctag aactatagct agtcgacatt gattattgac tagttattaa tagtaatcaa 240

ttacggggtc attagttcat agcccatata tggagttccg cgttacataa cttacggtaa 300

atggcccgcc tggctgaccg cccaacgacc cccgcccatt gacgtcaata atgacgtatg 360

ttcccatagt aacgccaata gggactttcc attgacgtca atgggtggac tatttacggt 420

aaactgccca cttggcagta catcaagtgt atcatatgcc aagtacgccc cctattgacg 480

tcaatgacgg taaatggccc gcctggcatt atgcccagta catgacctta tgggactttc 540

ctacttggca gtacatctac gtattagtca tcgctattac catggtcgag gtgagcccca 600

cgttctgctt cactctcccc atctcccccc cctccccacc cccaattttg tatttattta 660

ttttttaatt attttgtgca gcgatggggg cggggggggg gggggggcgc gcgccaggcg 720

gggcggggcg gggcgagggg cggggcgggg cgaggcggag aggtgcggcg gcagccaatc 780

agagcggcgc gctccgaaag tttcctttta tggcgaggcg gcggcggcgg cggccctata 840

aaaagcgaag cgcgcggcgg gcggggagtc gctgcgacgc tgccttcgcc ccgtgccccg 900

ctccgccgcc gcctcgcgcc gcccgccccg gctctgactg accgcgttac tcccacaggt 960

gagcgggcgg gacggccctt ctcctccggg ctgtaattag cgcttggttt aatgacggct 1020

tgtttctttt ctgtggctgc gtgaaagcct tgaggggctc cgggagggcc ctttgtgcgg 1080

ggggagcggc tcgggggggtg cgtgcgtgtg tgtgtgcgtg gggagcgccg cgtgcggctc 1140

cgcgctgccc ggcggctgtg agcgctgcgg gcgcggcgcg gggctttgtg cgctccgcag 1200

tgtgcgcgag gggagcgcgg ccgggggcgg tgccccgcgg tgcgggggggg gctgcgaggg 1260

gaacaaaggc tgcgtgcggg gtgtgtgcgt gggggggtga gcaggggtg tgggcgcgtc 1320

ggtcgggctg caacccccccc tgcacccccc tccccgagtt gctgagcacg gcccggcttc 1380

gggtgcgggg ctccgtacgg ggcgtggcgc ggggctcgcc gtgccgggcg gggggtggcg 1440

gcaggtgggg gtgccgggcg gggcggggcc gcctcgggcc ggggagggct cgggggaggg 1500

gcgcggcggc ccccggagcg ccggcggctg tcgaggcgcg gcgagccgca gccattgcct 1560

tttatggtaa tcgtgcgaga gggcgcaggg acttcctttg tcccaaatct gtgcggagcc 1620

gaaatctggg aggcgccgcc gcacccccctc tagcgggcgc ggggcgaagc ggtgcggcgc 1680

cggcaggaag gaaatgggcg gggagggcct tcgtgcgtcg ccgcgccgcc gtccccttct 1740

ccctctccag cctcggggct gtccgcgggg ggacggctgc cttcgggggg gacggggcag 1800

ggcggggttc ggcttctggc gtgtgaccgg cggctctaga gcctctgcta accatgttca 1860

tgccttcttc tttttcctac agctcctggg caacgtgctg gttattgtgc tgtctcatca 1920

ttttggcaaa gaattc atg ccc ccc ccc agg aca gga aga gga ctg ctg tgg 1972
Met Pro Pro Pro Arg Thr Gly Arg Gly Leu Leu Trp
1 5 10

ctc gga ctc gtc ctc tcc agc gtg tgc gtg gct ctc gga tcc gaa acc 2020
Leu Gly Leu Val Leu Ser Ser Val Cys Val Ala Leu Gly Ser Glu Thr

```
        15                  20                  25

cag gct aat tcc aca acc gat gcc ctg aac gtc ctg ctc atc atc gtg       2068
Gln Ala Asn Ser Thr Thr Asp Ala Leu Asn Val Leu Leu Ile Ile Val
    30                  35                  40

gat gac ctg agg ccc tcc ctg ggc tgt tac ggc gac aag ctc gtc agg       2116
Asp Asp Leu Arg Pro Ser Leu Gly Cys Tyr Gly Asp Lys Leu Val Arg
45                  50                  55                  60

tcc ccc aac att gat cag ctc gcc tcc cat tcc ctg ctg ttt caa aac       2164
Ser Pro Asn Ile Asp Gln Leu Ala Ser His Ser Leu Leu Phe Gln Asn
                65                  70                  75

gcc ttc gct cag caa gcc gtg tgc gct cct agc aga gtc agc ttc ctc       2212
Ala Phe Ala Gln Gln Ala Val Cys Ala Pro Ser Arg Val Ser Phe Leu
            80                  85                  90

acc gga agg aga cct gac acc acc agg ctg tac gac ttc aac tcc tat       2260
Thr Gly Arg Arg Pro Asp Thr Thr Arg Leu Tyr Asp Phe Asn Ser Tyr
        95                  100                 105

tgg agg gtc cac gct ggc aac ttc agc acc att ccc cag tac ttc aag       2308
Trp Arg Val His Ala Gly Asn Phe Ser Thr Ile Pro Gln Tyr Phe Lys
    110                 115                 120

gag aac ggc tac gtc acc atg tcc gtc ggc aag gtg ttc cac cct ggc       2356
Glu Asn Gly Tyr Val Thr Met Ser Val Gly Lys Val Phe His Pro Gly
125                 130                 135                 140

att tcc tcc aac cac acc gat gac agc ccc tac tcc tgg tcc ttc cct       2404
Ile Ser Ser Asn His Thr Asp Asp Ser Pro Tyr Ser Trp Ser Phe Pro
                145                 150                 155

ccc tac cac ccc agc agc gaa aag tac gag aat acc aag acc tgt aga       2452
Pro Tyr His Pro Ser Ser Glu Lys Tyr Glu Asn Thr Lys Thr Cys Arg
            160                 165                 170

ggc ccc gac ggc gaa ctc cac gct aac ctc ctg tgt cct gtc gat gtc       2500
Gly Pro Asp Gly Glu Leu His Ala Asn Leu Leu Cys Pro Val Asp Val
        175                 180                 185

ctc gac gtg ccc gaa ggc aca ctc ccc gac aag cag tcc aca gaa caa       2548
Leu Asp Val Pro Glu Gly Thr Leu Pro Asp Lys Gln Ser Thr Glu Gln
    190                 195                 200

gcc atc cag ctc ctg gag aag atg aag acc tcc gct tcc ccc ttt ttc       2596
Ala Ile Gln Leu Leu Glu Lys Met Lys Thr Ser Ala Ser Pro Phe Phe
205                 210                 215                 220

ctg gcc gtg ggc tac cac aaa ccc cac att ccc ttc agg tac ccc aag       2644
Leu Ala Val Gly Tyr His Lys Pro His Ile Pro Phe Arg Tyr Pro Lys
                225                 230                 235

gag ttc cag aag ctg tat ccc ctg gag aac att acc ctc gcc cct gat       2692
Glu Phe Gln Lys Leu Tyr Pro Leu Glu Asn Ile Thr Leu Ala Pro Asp
            240                 245                 250

ccc gaa gtg cct gac ggc ctg cct cct gtc gct tac aac ccc tgg atg       2740
Pro Glu Val Pro Asp Gly Leu Pro Pro Val Ala Tyr Asn Pro Trp Met
        255                 260                 265

gac atc agg cag aga gaa gac gtg caa gcc ctc aac atc agc gtg ccc       2788
Asp Ile Arg Gln Arg Glu Asp Val Gln Ala Leu Asn Ile Ser Val Pro
    270                 275                 280

tac ggc cct atc ccc gtg gac ttc cag aga aag atc agg cag agc tac       2836
Tyr Gly Pro Ile Pro Val Asp Phe Gln Arg Lys Ile Arg Gln Ser Tyr
```

```
285                 290                 295                 300

ttc gcc tcc gtc agc tac ctg gac acc cag gtg gga aga ctc ctg tcc       2884
Phe Ala Ser Val Ser Tyr Leu Asp Thr Gln Val Gly Arg Leu Leu Ser
                305                 310                 315

gcc ctc gac gat ctc cag ctg gcc aac agc acc atc att gcc ttc acc       2932
Ala Leu Asp Asp Leu Gln Leu Ala Asn Ser Thr Ile Ile Ala Phe Thr
            320                 325                 330

agc gac cat ggc tgg gct ctc gga gag cat ggc gaa tgg gcc aag tac       2980
Ser Asp His Gly Trp Ala Leu Gly Glu His Gly Glu Trp Ala Lys Tyr
        335                 340                 345

tcc aac ttc gat gtg gcc acc cac gtc ccc ctc atc ttc tac gtc ccc       3028
Ser Asn Phe Asp Val Ala Thr His Val Pro Leu Ile Phe Tyr Val Pro
    350                 355                 360

gga agg acc gcc agc ctg cct gaa gct gga gag aag ctc ttc ccc tac       3076
Gly Arg Thr Ala Ser Leu Pro Glu Ala Gly Glu Lys Leu Phe Pro Tyr
365                 370                 375                 380

ctg gac cct ttt gat tcc gcc agc caa ctc atg gaa ccc ggc agg caa       3124
Leu Asp Pro Phe Asp Ser Ala Ser Gln Leu Met Glu Pro Gly Arg Gln
                385                 390                 395

agc atg gac ctc gtg gag ctg gtc tcc ctg ttc cct acc ctg gct gga       3172
Ser Met Asp Leu Val Glu Leu Val Ser Leu Phe Pro Thr Leu Ala Gly
            400                 405                 410

ctc gcc gga ctc caa gtc cct ccc aga tgc cct gtg cct agc ttc cat       3220
Leu Ala Gly Leu Gln Val Pro Pro Arg Cys Pro Val Pro Ser Phe His
        415                 420                 425

gtg gag ctc tgc aga gag ggc aag aac ctg ctg aag cac ttc agg ttt       3268
Val Glu Leu Cys Arg Glu Gly Lys Asn Leu Leu Lys His Phe Arg Phe
    430                 435                 440

agg gac ctg gag gag gat cct tac ctc ccc gga aac cct aga gag ctc       3316
Arg Asp Leu Glu Glu Asp Pro Tyr Leu Pro Gly Asn Pro Arg Glu Leu
445                 450                 455                 460

atc gct tac tcc cag tat cct agg ccc tcc gac atc ccc cag tgg aac       3364
Ile Ala Tyr Ser Gln Tyr Pro Arg Pro Ser Asp Ile Pro Gln Trp Asn
                465                 470                 475

agc gat aag ccc agc ctg aag gat atc aaa atc atg gga tac tcc att       3412
Ser Asp Lys Pro Ser Leu Lys Asp Ile Lys Ile Met Gly Tyr Ser Ile
            480                 485                 490

agg aca atc gac tac agg tac acc gtc tgg gtc gga ttc aac ccc gat       3460
Arg Thr Ile Asp Tyr Arg Tyr Thr Val Trp Val Gly Phe Asn Pro Asp
        495                 500                 505

gag ttc ctg gcc aac ttc tcc gac atc cac gct ggc gag ctg tat ttc       3508
Glu Phe Leu Ala Asn Phe Ser Asp Ile His Ala Gly Glu Leu Tyr Phe
    510                 515                 520

gtg gac agc gac cct ctg cag gac cac aac atg tac aac gac tcc cag       3556
Val Asp Ser Asp Pro Leu Gln Asp His Asn Met Tyr Asn Asp Ser Gln
525                 530                 535                 540

gga ggc gat ctg ttt caa ctc ctg atg ccc tga ctcgaggacg gggtgaacta     3609
Gly Gly Asp Leu Phe Gln Leu Leu Met Pro
                545                 550

cgcctgagga tccgatcttt ttccctctgc caaaaattat ggggacatca tgaagcccct     3669
```

tgagcatctg acttctggct aataaaggaa atttattttc attgcaatag tgtgttggaa 3729

ttttttgtgt ctctcactcg gaagcaattc gttgatctga atttcgacca cccataatac 3789

ccattaccct ggtagataag tagcatggcg ggttaatcat taactacaag gaacccctag 3849

tgatggagtt ggccactccc tctctgcgcg ctcgctcgct cactgaggcc gggcgaccaa 3909

aggtcgcccg acgcccgggc tttgcccggg cggcctcagt gagcgagcga gcgcgcag 3967

<210> 12
<211> 550
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 12

```
Met Pro Pro Pro Arg Thr Gly Arg Gly Leu Leu Trp Leu Gly Leu Val
1               5                  10                  15

Leu Ser Ser Val Cys Val Ala Leu Gly Ser Glu Thr Gln Ala Asn Ser
            20                  25                  30

Thr Thr Asp Ala Leu Asn Val Leu Leu Ile Ile Val Asp Asp Leu Arg
        35                  40                  45

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

Asp Gln Leu Ala Ser His Ser Leu Leu Phe Gln Asn Ala Phe Ala Gln
65                  70                  75                  80

Gln Ala Val Cys Ala Pro Ser Arg Val Ser Phe Leu Thr Gly Arg Arg
                85                  90                  95

Pro Asp Thr Thr Arg Leu Tyr Asp Phe Asn Ser Tyr Trp Arg Val His
            100                 105                 110

Ala Gly Asn Phe Ser Thr Ile Pro Gln Tyr Phe Lys Glu Asn Gly Tyr
        115                 120                 125

Val Thr Met Ser Val Gly Lys Val Phe His Pro Gly Ile Ser Ser Asn
    130                 135                 140

His Thr Asp Asp Ser Pro Tyr Ser Trp Ser Phe Pro Pro Tyr His Pro
145                 150                 155                 160

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

Glu Leu His Ala Asn Leu Leu Cys Pro Val Asp Val Leu Asp Val Pro
            180                 185                 190
```

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

Leu Glu Lys Met Lys Thr Ser Ala Ser Pro Phe Phe Leu Ala Val Gly
    210                 215                 220

Tyr His Lys Pro His Ile Pro Phe Arg Tyr Pro Lys Glu Phe Gln Lys
225                 230                 235                 240

Leu Tyr Pro Leu Glu Asn Ile Thr Leu Ala Pro Asp Pro Glu Val Pro
                245                 250                 255

Asp Gly Leu Pro Pro Val Ala Tyr Asn Pro Trp Met Asp Ile Arg Gln
            260                 265                 270

Arg Glu Asp Val Gln Ala Leu Asn Ile Ser Val Pro Tyr Gly Pro Ile
        275                 280                 285

Pro Val Asp Phe Gln Arg Lys Ile Arg Gln Ser Tyr Phe Ala Ser Val
    290                 295                 300

Ser Tyr Leu Asp Thr Gln Val Gly Arg Leu Leu Ser Ala Leu Asp Asp
305                 310                 315                 320

Leu Gln Leu Ala Asn Ser Thr Ile Ile Ala Phe Thr Ser Asp His Gly
                325                 330                 335

Trp Ala Leu Gly Glu His Gly Glu Trp Ala Lys Tyr Ser Asn Phe Asp
            340                 345                 350

Val Ala Thr His Val Pro Leu Ile Phe Tyr Val Pro Gly Arg Thr Ala
        355                 360                 365

Ser Leu Pro Glu Ala Gly Glu Lys Leu Phe Pro Tyr Leu Asp Pro Phe
    370                 375                 380

Asp Ser Ala Ser Gln Leu Met Glu Pro Gly Arg Gln Ser Met Asp Leu
385                 390                 395                 400

Val Glu Leu Val Ser Leu Phe Pro Thr Leu Ala Gly Leu Ala Gly Leu
                405                 410                 415

Gln Val Pro Pro Arg Cys Pro Val Pro Ser Phe His Val Glu Leu Cys
            420                 425                 430

Arg Glu Gly Lys Asn Leu Leu Lys His Phe Arg Phe Arg Asp Leu Glu
        435                 440                 445

Glu Asp Pro Tyr Leu Pro Gly Asn Pro Arg Glu Leu Ile Ala Tyr Ser
    450                 455                 460
```

Gln Tyr Pro Arg Pro Ser Asp Ile Pro Gln Trp Asn Ser Asp Lys Pro
465 470 475 480

Ser Leu Lys Asp Ile Lys Ile Met Gly Tyr Ser Ile Arg Thr Ile Asp
485 490 495

Tyr Arg Tyr Thr Val Trp Val Gly Phe Asn Pro Asp Glu Phe Leu Ala
500 505 510

Asn Phe Ser Asp Ile His Ala Gly Glu Leu Tyr Phe Val Asp Ser Asp
515 520 525

Pro Leu Gln Asp His Asn Met Tyr Asn Asp Ser Gln Gly Gly Asp Leu
530 535 540

Phe Gln Leu Leu Met Pro
545 550

<210> 13
<211> 736
<212> PRT
<213> Unknown

<220>
<223> AAV9 capsid amino acid sequence

<400> 13

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Ala Leu Lys Pro Gly Ala Pro Gln Pro
20 25 30

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

Gly Tyr Lys Tyr Leu Gly Pro Gly Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Lys Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Leu Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg

130 135 140

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
145 150 155 160

Lys Ser Gly Ala Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr
165 170 175

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190

Ala Ala Pro Ser Gly Val Gly Ser Leu Thr Met Ala Ser Gly Gly Gly
195 200 205

Ala Pro Val Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser Ser
210 215 220

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu Gly Asp Arg Val Ile
225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
245 250 255

Tyr Lys Gln Ile Ser Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270

Ala Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
275 280 285

Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
290 295 300

Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile
305 310 315 320

Gln Val Lys Glu Val Thr Asp Asn Asn Gly Val Lys Thr Ile Ala Asn
325 330 335

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380

Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
385 390 395 400

Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Glu

405 410 415

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430

Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser
435 440 445

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
500 505 510

Gly Arg Asn Ser Leu Met Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

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

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560

Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
565 570 575

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590

Thr Gly Trp Val Gln Asn Gln Gly Ile Leu Pro Gly Met Val Trp Gln
595 600 605

Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
610 615 620

Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Met
625 630 635 640

Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
645 650 655

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670

Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln

675 680 685

Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
690 695 700

Tyr Tyr Lys Ser Asn Asn Val Glu Phe Ala Val Asn Thr Glu Gly Val
705 710 715 720

Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
725 730 735

<210> 14
<211> 3965
<212> DNA
<213> Artificial Sequence

<220>
<223> CB7.CI.hIDS.RBG

<220>
<221> misc_feature
<222> (2)..(131)
<223> 5'ITR

<220>
<221> promoter
<222> (199)..(580)
<223> CMV IE promoter

<220>
<221> promoter
<222> (583)..(864)
<223> CB promoter

<220>
<221> TATA_signal
<222> (837)..(840)

<220>
<221> Intron
<222> (957)..(1928)
<223> chicken beta-actin intron

<220>
<221> CDS
<222> (1935)..(3587)
<223> hIDS

<220>
<221> polyA_signal
<222> (3621)..(3747)
<223> rabbit globin polyA

<220>
<221> misc_feature
<222> (3836)..(3965)
<223> 3'ITR

<400> 14
gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc gggcgacctt 60

tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca actccatcac 120

taggggttcc ttgtagttaa tgattaaccc gccatgctac ttatctacca gggtaatggg 180

gatcctctag aactatagct agtcgacatt gattattgac tagttattaa tagtaatcaa 240

ttacggggtc attagttcat agcccatata tggagttccg cgttacataa cttacggtaa 300

atggcccgcc tggctgaccg cccaacgacc cccgcccatt gacgtcaata atgacgtatg 360

ttcccatagt aacgccaata gggactttcc attgacgtca atgggtggag tatttacggt 420

aaactgccca cttggcagta catcaagtgt atcatatgcc aagtacgccc cctattgacg 480

tcaatgacgg taaatggccc gcctggcatt atgcccagta catgacctta tgggactttc 540

ctacttggca gtacatctac gtattagtca tcgctattac catggtcgag gtgagcccca 600

cgttctgctt cactctcccc atctcccccc cctccccacc cccaattttg tatttattta 660

ttttttaatt attttgtgca gcgatggggg cggggggggg gggggggcgc gcgccaggcg 720

gggcggggcg gggcgagggg cggggcgggg cgaggcggag aggtgcggcg gcagccaatc 780

agagcggcgc gctccgaaag tttcctttta tggcgaggcg gcggcggcgg cggccctata 840

aaaagcgaag cgcgcggcgg gcgggagtcg ctgcgcgctg ccttcgcccc gtgccccgct 900

ccgccgccgc ctcgcgccgc ccgccccggc tctgactgac cgcgttactc ccacaggtga 960

gcgggcggga cggcccttct cctccgggct gtaattagcg cttggtttaa tgacggcttg 1020

tttcttttct gtggctgcgt gaaagccttg aggggctccg ggagggccct ttgtgcgggg 1080

ggagcggctc ggggggtgcg tgcgtgtgtg tgtgcgtggg gagcgccgcg tgcggctccg 1140

cgctgcccgg cggctgtgag cgctgcgggc gcggcgcggg gctttgtgcg ctccgcagtg 1200

tgcgcgaggg gagcgcggcc gggggcggtg ccccgcggtg cggggggggc tgcgagggga 1260

acaaaggctg cgtgcggggt gtgtgcgtgg gggggtgagc agggggtgtg ggcgcgtcgg 1320

tcgggctgca acccccccctg cacccccctc cccgagttgc tgagcacggc ccggcttcgg 1380

gtgcggggct ccgtacgggg cgtggcgcgg ggctcgccgt gccgggcggg gggtggcggc 1440

aggtgggggt gccgggcggg gcggggccgc ctcgggccgg ggagggctcg ggggaggggc 1500

gcggcggccc ccggagcgcc ggcggctgtc gaggcgcggc gagccgcagc cattgccttt 1560

tatggtaatc gtgcgagagg gcgcagggac ttcctttgtc ccaaatctgt gcggagccga 1620

aatctgggag gcgccgccgc acccccctcta gcgggcgcgg ggcgaagcgg tgcggcgccg 1680

gcaggaagga aatgggcggg gagggccttc gtgcgtcgcc gcgccgccgt ccccttctcc 1740

ctctccagcc tcggggctgt ccgcgggggg acggctgcct tcgggggggа cggggcaggg 1800

cggggttcgg cttctggcgt gtgaccggcg gctctagagc ctctgctaac catgttcatg 1860

ccttcttctt tttcctacag ctcctgggca acgtgctggt tattgtgctg tctcatcatt 1920

ttggcaaaga attc atg ccg cca ccc cgg acc ggc cga ggc ctt ctc tgg 1970
Met Pro Pro Pro Arg Thr Gly Arg Gly Leu Leu Trp
1 5 10

ctg ggt ctg gtt ctg agc tcc gtc tgc gtc gcc ctc gga tcc gaa acg 2018
Leu Gly Leu Val Leu Ser Ser Val Cys Val Ala Leu Gly Ser Glu Thr
15 20 25

```
cag gcc aac tcg acc aca gat gct ctg aac gtt ctt ctc atc atc gtg       2066
Gln Ala Asn Ser Thr Thr Asp Ala Leu Asn Val Leu Leu Ile Ile Val
    30                  35                  40

gat gac ctg cgc ccc tcc ctg ggc tgt tat ggg gat aag ctg gtg agg       2114
Asp Asp Leu Arg Pro Ser Leu Gly Cys Tyr Gly Asp Lys Leu Val Arg
45                  50                  55                  60

tcc cca aat att gac caa ctg gca tcc cac agc ctc ctc ttc cag aat       2162
Ser Pro Asn Ile Asp Gln Leu Ala Ser His Ser Leu Leu Phe Gln Asn
                65                  70                  75

gcc ttt gcg cag caa gca gtg tgc gcc ccg agc cgc gtt tct ttc ctc       2210
Ala Phe Ala Gln Gln Ala Val Cys Ala Pro Ser Arg Val Ser Phe Leu
            80                  85                  90

act ggc agg aga cct gac acc acc cgc ctg tac gac ttc aac tcc tac       2258
Thr Gly Arg Arg Pro Asp Thr Thr Arg Leu Tyr Asp Phe Asn Ser Tyr
        95                  100                 105

tgg agg gtg cac gct gga aac ttc tcc acc atc ccc cag tac ttc aag       2306
Trp Arg Val His Ala Gly Asn Phe Ser Thr Ile Pro Gln Tyr Phe Lys
    110                 115                 120

gag aat ggc tat gtg acc atg tcg gtg gga aaa gtc ttt cac cct ggg       2354
Glu Asn Gly Tyr Val Thr Met Ser Val Gly Lys Val Phe His Pro Gly
125                 130                 135                 140

ata tct tct aac cat acc gat gat tct ccg tat agc tgg tct ttt cca       2402
Ile Ser Ser Asn His Thr Asp Asp Ser Pro Tyr Ser Trp Ser Phe Pro
                145                 150                 155

cct tat cat cct tcc tct gag aag tat gaa aac act aag aca tgt cga       2450
Pro Tyr His Pro Ser Ser Glu Lys Tyr Glu Asn Thr Lys Thr Cys Arg
            160                 165                 170

ggg cca gat gga gaa ctc cat gcc aac ctg ctt tgc cct gtg gat gtg       2498
Gly Pro Asp Gly Glu Leu His Ala Asn Leu Leu Cys Pro Val Asp Val
        175                 180                 185

ctg gat gtt ccc gag ggc acc ttg cct gac aaa cag agc act gag caa       2546
Leu Asp Val Pro Glu Gly Thr Leu Pro Asp Lys Gln Ser Thr Glu Gln
    190                 195                 200

gcc ata cag ttg ttg gaa aag atg aaa acg tca gcc agt cct ttc ttc       2594
Ala Ile Gln Leu Leu Glu Lys Met Lys Thr Ser Ala Ser Pro Phe Phe
205                 210                 215                 220

ctg gcc gtt ggg tat cat aag cca cac atc ccc ttc aga tac ccc aag       2642
Leu Ala Val Gly Tyr His Lys Pro His Ile Pro Phe Arg Tyr Pro Lys
                225                 230                 235

gaa ttt cag aag ttg tat ccc ttg gag aac atc acc ctg gcc ccc gat       2690
Glu Phe Gln Lys Leu Tyr Pro Leu Glu Asn Ile Thr Leu Ala Pro Asp
            240                 245                 250

ccc gag gtc cct gat ggc cta ccc cct gtg gcc tac aac ccc tgg atg       2738
Pro Glu Val Pro Asp Gly Leu Pro Pro Val Ala Tyr Asn Pro Trp Met
        255                 260                 265

gac atc agg caa cgg gaa gac gtc caa gcc tta aac atc agt gtg ccg       2786
Asp Ile Arg Gln Arg Glu Asp Val Gln Ala Leu Asn Ile Ser Val Pro
    270                 275                 280

tat ggt cca att cct gtg gac ttt cag cgg aaa atc cgc cag agc tac       2834
Tyr Gly Pro Ile Pro Val Asp Phe Gln Arg Lys Ile Arg Gln Ser Tyr
285                 290                 295                 300
```

```
ttt gcc tct gtg tca tat ttg gat aca cag gtc ggc cgc ctc ttg agt       2882
Phe Ala Ser Val Ser Tyr Leu Asp Thr Gln Val Gly Arg Leu Leu Ser
                305                 310                 315

gct ttg gac gat ctt cag ctg gcc aac agc acc atc att gca ttt acc       2930
Ala Leu Asp Asp Leu Gln Leu Ala Asn Ser Thr Ile Ile Ala Phe Thr
            320                 325                 330

tcg gat cat ggg tgg gct cta ggt gaa cat gga gaa tgg gcc aaa tac       2978
Ser Asp His Gly Trp Ala Leu Gly Glu His Gly Glu Trp Ala Lys Tyr
        335                 340                 345

agc aat ttt gat gtt gct acc cat gtt ccc ctg ata ttc tat gtt cct       3026
Ser Asn Phe Asp Val Ala Thr His Val Pro Leu Ile Phe Tyr Val Pro
    350                 355                 360

gga agg acg gct tca ctt ccg gag gca ggc gag aag ctt ttc cct tac       3074
Gly Arg Thr Ala Ser Leu Pro Glu Ala Gly Glu Lys Leu Phe Pro Tyr
365                 370                 375                 380

ctc gac cct ttt gat tcc gcc tca cag ttg atg gag cca ggc agg caa       3122
Leu Asp Pro Phe Asp Ser Ala Ser Gln Leu Met Glu Pro Gly Arg Gln
                385                 390                 395

tcc atg gac ctt gtg gaa ctt gtg tct ctt ttt ccc acg ctg gct gga       3170
Ser Met Asp Leu Val Glu Leu Val Ser Leu Phe Pro Thr Leu Ala Gly
            400                 405                 410

ctt gca gga ctg cag gtt cca cct cgc tgc ccc gtt cct tca ttt cac       3218
Leu Ala Gly Leu Gln Val Pro Pro Arg Cys Pro Val Pro Ser Phe His
        415                 420                 425

gtt gag ctg tgc aga gaa ggc aag aac ctt ctg aag cat ttt cga ttc       3266
Val Glu Leu Cys Arg Glu Gly Lys Asn Leu Leu Lys His Phe Arg Phe
    430                 435                 440

cgt gac ttg gaa gag gat ccg tac ctc cct ggt aat ccc cgt gaa ctg       3314
Arg Asp Leu Glu Glu Asp Pro Tyr Leu Pro Gly Asn Pro Arg Glu Leu
445                 450                 455                 460

att gcc tat agc cag tat ccc cgg cct tca gac atc cct cag tgg aat       3362
Ile Ala Tyr Ser Gln Tyr Pro Arg Pro Ser Asp Ile Pro Gln Trp Asn
                465                 470                 475

tct gac aag ccg agt tta aaa gat ata aag atc atg ggc tat tcc ata       3410
Ser Asp Lys Pro Ser Leu Lys Asp Ile Lys Ile Met Gly Tyr Ser Ile
            480                 485                 490

cgc acc ata gac tat agg tat act gtg tgg gtt ggc ttc aat cct gat       3458
Arg Thr Ile Asp Tyr Arg Tyr Thr Val Trp Val Gly Phe Asn Pro Asp
        495                 500                 505

gaa ttt cta gct aac ttt tct gac atc cat gca ggg gaa ctg tat ttt       3506
Glu Phe Leu Ala Asn Phe Ser Asp Ile His Ala Gly Glu Leu Tyr Phe
    510                 515                 520

gtg gat tct gac cca ttg cag gat cac aat atg tat aat gat tcc caa       3554
Val Asp Ser Asp Pro Leu Gln Asp His Asn Met Tyr Asn Asp Ser Gln
525                 530                 535                 540

ggt gga gat ctt ttc cag ttg ttg atg cct tga ctcgaggacg gggtgaacta     3607
Gly Gly Asp Leu Phe Gln Leu Leu Met Pro
                545                 550

cgcctgagga tccgatcttt ttccctctgc caaaaattat ggggacatca tgaagcccct     3667

tgagcatctg acttctggct aataaaggaa atttattttc attgcaatag tgtgttggaa     3727
```

tttttgtgt ctctcactcg gaagcaattc gttgatctga atttcgacca cccataatac 3787

ccattaccct ggtagataag tagcatggcg ggttaatcat taactacaag gaacccctag 3847

tgatggagtt ggccactccc tctctgcgcg ctcgctcgct cactgaggcc gggcgaccaa 3907

aggtcgcccg acgcccgggc tttgcccggg cggcctcagt gagcgagcga gcgcgcag 3965

<210> 15
<211> 550
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 15

Met Pro Pro Pro Arg Thr Gly Arg Gly Leu Leu Trp Leu Gly Leu Val
1 5 10 15

Leu Ser Ser Val Cys Val Ala Leu Gly Ser Glu Thr Gln Ala Asn Ser
20 25 30

Thr Thr Asp Ala Leu Asn Val Leu Leu Ile Ile Val Asp Asp Leu Arg
35 40 45

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

Asp Gln Leu Ala Ser His Ser Leu Leu Phe Gln Asn Ala Phe Ala Gln
65 70 75 80

Gln Ala Val Cys Ala Pro Ser Arg Val Ser Phe Leu Thr Gly Arg Arg
85 90 95

Pro Asp Thr Thr Arg Leu Tyr Asp Phe Asn Ser Tyr Trp Arg Val His
100 105 110

Ala Gly Asn Phe Ser Thr Ile Pro Gln Tyr Phe Lys Glu Asn Gly Tyr
115 120 125

Val Thr Met Ser Val Gly Lys Val Phe His Pro Gly Ile Ser Ser Asn
130 135 140

His Thr Asp Asp Ser Pro Tyr Ser Trp Ser Phe Pro Pro Tyr His Pro
145 150 155 160

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

Glu Leu His Ala Asn Leu Leu Cys Pro Val Asp Val Leu Asp Val Pro
180 185 190

Glu Gly Thr Leu Pro Asp Lys Gln Ser Thr Glu Gln Ala Ile Gln Leu

195 200 205

Leu Glu Lys Met Lys Thr Ser Ala Ser Pro Phe Phe Leu Ala Val Gly
210 215 220

Tyr His Lys Pro His Ile Pro Phe Arg Tyr Pro Lys Glu Phe Gln Lys
225 230 235 240

Leu Tyr Pro Leu Glu Asn Ile Thr Leu Ala Pro Asp Pro Glu Val Pro
245 250 255

Asp Gly Leu Pro Pro Val Ala Tyr Asn Pro Trp Met Asp Ile Arg Gln
260 265 270

Arg Glu Asp Val Gln Ala Leu Asn Ile Ser Val Pro Tyr Gly Pro Ile
275 280 285

Pro Val Asp Phe Gln Arg Lys Ile Arg Gln Ser Tyr Phe Ala Ser Val
290 295 300

Ser Tyr Leu Asp Thr Gln Val Gly Arg Leu Leu Ser Ala Leu Asp Asp
305 310 315 320

Leu Gln Leu Ala Asn Ser Thr Ile Ile Ala Phe Thr Ser Asp His Gly
325 330 335

Trp Ala Leu Gly Glu His Gly Glu Trp Ala Lys Tyr Ser Asn Phe Asp
340 345 350

Val Ala Thr His Val Pro Leu Ile Phe Tyr Val Pro Gly Arg Thr Ala
355 360 365

Ser Leu Pro Glu Ala Gly Glu Lys Leu Phe Pro Tyr Leu Asp Pro Phe
370 375 380

Asp Ser Ala Ser Gln Leu Met Glu Pro Gly Arg Gln Ser Met Asp Leu
385 390 395 400

Val Glu Leu Val Ser Leu Phe Pro Thr Leu Ala Gly Leu Ala Gly Leu
405 410 415

Gln Val Pro Pro Arg Cys Pro Val Pro Ser Phe His Val Glu Leu Cys
420 425 430

Arg Glu Gly Lys Asn Leu Leu Lys His Phe Arg Phe Arg Asp Leu Glu
435 440 445

Glu Asp Pro Tyr Leu Pro Gly Asn Pro Arg Glu Leu Ile Ala Tyr Ser
450 455 460

Gln Tyr Pro Arg Pro Ser Asp Ile Pro Gln Trp Asn Ser Asp Lys Pro

465 470 475 480

Ser Leu Lys Asp Ile Lys Ile Met Gly Tyr Ser Ile Arg Thr Ile Asp
485 490 495

Tyr Arg Tyr Thr Val Trp Val Gly Phe Asn Pro Asp Glu Phe Leu Ala
500 505 510

Asn Phe Ser Asp Ile His Ala Gly Glu Leu Tyr Phe Val Asp Ser Asp
515 520 525

Pro Leu Gln Asp His Asn Met Tyr Asn Asp Ser Gln Gly Gly Asp Leu
530 535 540

Phe Gln Leu Leu Met Pro
545 550

<210> 16
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> synthesized sequence

<400> 16
ttccctctgc caaaaattat gg 22

<210> 17
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> synthesized sequence

<400> 17
cctttattag ccagaagtca gatgct 26

<210> 18
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthesized sequence

<400> 18
acatcatgaa gcccc 15