SEQUENCE LISTING

<110> Beth Israel Deaconess Medical Center, Inc.
Crucell Holland B.V.

<120> Methods and Compositions for Inducing Protective Immunity Against Human Immunodeficiency Virus Infection

<130> 688097-53WO (CRU5171WOPCT/0246 WO 00 ORD)

<160> 7

<170> PatentIn version 3.5

<210> 1
<211> 685
<212> PRT
<213> Artificial Sequence

<220>
<223> Mos1.Env mosaic antigen sequence

<400> 1

Met Arg Val Thr Gly Ile Arg Lys Asn Tyr Gln His Leu Trp Arg Trp
1 5 10 15

Gly Thr Met Leu Leu Gly Ile Leu Met Ile Cys Ser Ala Ala Gly Lys
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr
35 40 45

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

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
65 70 75 80

Gln Glu Val Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys
85 90 95

Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
115 120 125

Asn Cys Thr Asp Asp Val Arg Asn Val Thr Asn Asn Ala Thr Asn Thr
130 135 140

Asn Ser Ser Trp Gly Glu Pro Met Glu Lys Gly Glu Ile Lys Asn Cys
145 150 155 160

Ser Phe Asn Ile Thr Thr Ser Ile Arg Asn Lys Val Gln Lys Gln Tyr
165 170 175

```
Ala Leu Phe Tyr Lys Leu Asp Val Val Pro Ile Asp Asn Asp Ser Asn
            180                 185                 190

Asn Thr Asn Tyr Arg Leu Ile Ser Cys Asn Thr Ser Val Ile Thr Gln
        195                 200                 205

Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala
    210                 215                 220

Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Lys Phe Asn Gly
225                 230                 235                 240

Thr Gly Pro Cys Thr Asn Val Ser Thr Val Gln Cys Thr His Gly Ile
                245                 250                 255

Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu
            260                 265                 270

Glu Glu Val Val Ile Arg Ser Glu Asn Phe Thr Asn Asn Ala Lys Thr
        275                 280                 285

Ile Met Val Gln Leu Asn Val Ser Val Glu Ile Asn Cys Thr Arg Pro
    290                 295                 300

Asn Asn Asn Thr Arg Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe
305                 310                 315                 320

Tyr Thr Ala Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn
                325                 330                 335

Ile Ser Arg Ala Asn Trp Asn Asn Thr Leu Arg Gln Ile Val Glu Lys
            340                 345                 350

Leu Gly Lys Gln Phe Gly Asn Asn Lys Thr Ile Val Phe Asn His Ser
        355                 360                 365

Ser Gly Gly Asp Pro Glu Ile Val Met His Ser Phe Asn Cys Gly Gly
    370                 375                 380

Glu Phe Phe Tyr Cys Asn Ser Thr Lys Leu Phe Asn Ser Thr Trp Thr
385                 390                 395                 400

Trp Asn Asn Ser Thr Trp Asn Asn Thr Lys Arg Ser Asn Asp Thr Glu
                405                 410                 415

Glu His Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp
            420                 425                 430

Gln Glu Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln Ile
        435                 440                 445
```

```
Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly
    450                 455                 460

Asn Asp Thr Ser Gly Thr Glu Ile Phe Arg Pro Gly Gly Gly Asp Met
465                 470                 475                 480

Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Lys Ile
                485                 490                 495

Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln
            500                 505                 510

Ser Glu Lys Ser Ala Val Gly Ile Gly Ala Val Phe Leu Gly Phe Leu
        515                 520                 525

Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Met Thr Leu Thr Val
    530                 535                 540

Gln Ala Arg Leu Leu Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu
545                 550                 555                 560

Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp
                565                 570                 575

Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu
            580                 585                 590

Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile
        595                 600                 605

Cys Thr Thr Thr Val Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu
    610                 615                 620

Asp Lys Ile Trp Asn Asn Met Thr Trp Met Glu Trp Glu Arg Glu Ile
625                 630                 635                 640

Asn Asn Tyr Thr Ser Leu Ile Tyr Thr Leu Ile Glu Glu Ser Gln Asn
                645                 650                 655

Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala
            660                 665                 670

Ser Leu Trp Asn Trp Phe Asp Ile Ser Asn Trp Leu Trp
        675                 680                 685
```

<210> 2
<211> 684
<212> PRT
<213> Artificial Sequence

<220>
<223> Mos2.Env mosaic antigen sequence

<400> 2

```
Met Arg Val Arg Gly Ile Gln Arg Asn Trp Pro Gln Trp Trp Ile Trp
1               5                  10                  15

Gly Ile Leu Gly Phe Trp Met Ile Ile Ile Cys Arg Val Met Gly Asn
            20                  25                  30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Lys
        35                  40                  45

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

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
65                  70                  75                  80

Gln Glu Met Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys
                85                  90                  95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Arg Leu Trp Asp
            100                 105                 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
        115                 120                 125

Glu Cys Arg Asn Val Arg Asn Val Ser Ser Asn Gly Thr Tyr Asn Ile
    130                 135                 140

Ile His Asn Glu Thr Tyr Lys Glu Met Lys Asn Cys Ser Phe Asn Ala
145                 150                 155                 160

Thr Thr Val Val Glu Asp Arg Lys Gln Lys Val His Ala Leu Phe Tyr
                165                 170                 175

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

Ser Glu Asn Ser Ser Glu Tyr Tyr Arg Leu Ile Asn Cys Asn Thr Ser
        195                 200                 205

Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Asp Pro Ile Pro Ile
    210                 215                 220

His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys
225                 230                 235                 240

Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn Val Ser Thr Val Gln Cys
                245                 250                 255

Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly
            260                 265                 270
```

```
Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr Asn
        275                 280                 285

Asn Ala Lys Thr Ile Ile Val His Leu Asn Glu Thr Val Asn Ile Thr
    290                 295                 300

Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro
305                 310                 315                 320

Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln
                325                 330                 335

Ala His Cys Asn Leu Ser Arg Asp Gly Trp Asn Lys Thr Leu Gln Gly
            340                 345                 350

Val Lys Lys Lys Leu Ala Glu His Phe Pro Asn Lys Thr Ile Asn Phe
        355                 360                 365

Thr Ser Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn
    370                 375                 380

Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly Leu Phe Asn Gly
385                 390                 395                 400

Thr Tyr Met Pro Asn Gly Thr Asn Ser Asn Ser Ser Ser Asn Ile Thr
                405                 410                 415

Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly
            420                 425                 430

Arg Ala Met Tyr Ala Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser
        435                 440                 445

Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Ser Asn Asn Gly
    450                 455                 460

Val Pro Asn Asp Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg
465                 470                 475                 480

Asn Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Val Lys
                485                 490                 495

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

Glu Lys Ser Ala Val Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly
        515                 520                 525

Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln
    530                 535                 540
```

```
Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu
545                 550                 555                 560

Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly
                565                 570                 575

Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln
            580                 585                 590

Asp Gln Gln Leu Leu Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys
        595                 600                 605

Thr Thr Ala Val Pro Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr
    610                 615                 620

Asp Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly
625                 630                 635                 640

Asn Tyr Thr Gly Glu Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln
                645                 650                 655

Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp Lys Asn
            660                 665                 670

Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp
        675                 680
```

<210> 3
<211> 1350
<212> PRT
<213> Artificial Sequence

<220>
<223> Mos1.Gag-Pol mosaic antigen sequence

<400> 3

```
Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp
1               5                   10                  15

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

His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro
        35                  40                  45

Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu
    50                  55                  60

Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn
65                  70                  75                  80
```

```
Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Ile Lys Asp
                85                  90                  95

Thr Lys Glu Ala Leu Glu Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys
            100                 105                 110

Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly Asn Ser Ser Gln Val
        115                 120                 125

Ser Gln Asn Tyr Pro Ile Val Gln Asn Ile Gln Gly Gln Met Val His
    130                 135                 140

Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu
145                 150                 155                 160

Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser
                165                 170                 175

Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly
            180                 185                 190

Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu
        195                 200                 205

Ala Ala Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala
    210                 215                 220

Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr
225                 230                 235                 240

Ser Thr Leu Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile
                245                 250                 255

Pro Val Gly Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys
            260                 265                 270

Ile Val Arg Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Arg Gln Gly
        275                 280                 285

Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu
    290                 295                 300

Arg Ala Glu Gln Ala Ser Gln Asp Val Lys Asn Trp Met Thr Glu Thr
305                 310                 315                 320

Leu Leu Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala
                325                 330                 335

Leu Gly Pro Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly
            340                 345                 350
```

```
Val Gly Gly Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser
        355                 360                 365

Gln Val Thr Asn Ser Ala Thr Ile Met Met Gln Arg Gly Asn Phe Arg
    370                 375                 380

Asn Gln Arg Lys Thr Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His
385                 390                 395                 400

Ile Ala Lys Asn Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys
                405                 410                 415

Gly Lys Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn
            420                 425                 430

Phe Leu Gly Lys Ile Trp Pro Ser Asn Lys Gly Arg Pro Gly Asn Phe
        435                 440                 445

Leu Gln Asn Arg Pro Glu Pro Thr Ala Pro Pro Glu Glu Ser Phe Arg
    450                 455                 460

Phe Gly Glu Glu Thr Thr Thr Pro Ser Gln Lys Gln Glu Pro Ile Asp
465                 470                 475                 480

Lys Glu Met Tyr Pro Leu Ala Ser Leu Lys Ser Leu Phe Gly Asn Asp
                485                 490                 495

Pro Ser Ser Gln Met Ala Pro Ile Ser Pro Ile Glu Thr Val Pro Val
            500                 505                 510

Lys Leu Lys Pro Gly Met Asp Gly Pro Arg Val Lys Gln Trp Pro Leu
        515                 520                 525

Thr Glu Glu Lys Ile Lys Ala Leu Thr Ala Ile Cys Glu Glu Met Glu
    530                 535                 540

Lys Glu Gly Lys Ile Thr Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr
545                 550                 555                 560

Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu
                565                 570                 575

Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val
            580                 585                 590

Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser Val
        595                 600                 605

Thr Val Leu Ala Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu
    610                 615                 620
```

```
Gly Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Thr Asn Asn Glu
625                 630                 635                 640

Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys
                645                 650                 655

Gly Ser Pro Ala Ile Phe Gln Cys Ser Met Thr Arg Ile Leu Glu Pro
            660                 665                 670

Phe Arg Ala Lys Asn Pro Glu Ile Val Ile Tyr Gln Tyr Met Ala Ala
        675                 680                 685

Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Ala Lys Ile
    690                 695                 700

Glu Glu Leu Arg Glu His Leu Leu Lys Trp Gly Phe Thr Thr Pro Asp
705                 710                 715                 720

Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu
                725                 730                 735

His Pro Asp Lys Trp Thr Val Gln Pro Ile Gln Leu Pro Glu Lys Asp
            740                 745                 750

Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp
        755                 760                 765

Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu
    770                 775                 780

Leu Arg Gly Ala Lys Ala Leu Thr Asp Ile Val Pro Leu Thr Glu Glu
785                 790                 795                 800

Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val
                805                 810                 815

His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln
            820                 825                 830

Lys Gln Gly His Asp Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe
        835                 840                 845

Lys Asn Leu Lys Thr Gly Lys Tyr Ala Lys Met Arg Thr Ala His Thr
    850                 855                 860

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

Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe Arg Leu Pro Ile Gln
                885                 890                 895
```

```
Lys Glu Thr Trp Glu Thr Trp Trp Thr Asp Tyr Trp Gln Ala Thr Trp
            900                 905                 910

Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp
        915                 920                 925

Tyr Gln Leu Glu Lys Asp Pro Ile Ala Gly Val Glu Thr Phe Tyr Val
    930                 935                 940

Ala Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val
945                 950                 955                 960

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

Gln Lys Thr Ala Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly
            980                 985                 990

Ser Glu Val Asn Ile Val Thr Ala  Ser Gln Tyr Ala Leu  Gly Ile Ile
        995                 1000                1005

Gln Ala  Gln Pro Asp Lys Ser  Glu Ser Glu Leu Val  Asn Gln Ile
    1010                1015                1020

Ile Glu  Gln Leu Ile Lys Lys  Glu Arg Val Tyr Leu  Ser Trp Val
    1025                1030                1035

Pro Ala  His Lys Gly Ile Gly  Gly Asn Glu Gln Val  Asp Lys Leu
    1040                1045                1050

Val Ser  Ser Gly Ile Arg Lys  Val Leu Phe Leu Asp  Gly Ile Asp
    1055                1060                1065

Lys Ala  Gln Glu Glu His Glu  Lys Tyr His Ser Asn  Trp Arg Ala
    1070                1075                1080

Met Ala  Ser Asp Phe Asn Leu  Pro Pro Val Val Ala  Lys Glu Ile
    1085                1090                1095

Val Ala  Ser Cys Asp Gln Cys  Gln Leu Lys Gly Glu  Ala Met His
    1100                1105                1110

Gly Gln  Val Asp Cys Ser Pro  Gly Ile Trp Gln Leu  Ala Cys Thr
    1115                1120                1125

His Leu  Glu Gly Lys Ile Ile  Leu Val Ala Val His  Val Ala Ser
    1130                1135                1140

Gly Tyr  Ile Glu Ala Glu Val  Ile Pro Ala Glu Thr  Gly Gln Glu
    1145                1150                1155
```

```
Thr Ala  Tyr Phe Ile Leu Lys  Leu Ala Gly Arg Trp  Pro Val Lys
    1160                 1165                 1170

Val Ile  His Thr Ala Asn Gly  Ser Asn Phe Thr Ser  Ala Ala Val
    1175                 1180                 1185

Lys Ala  Ala Cys Trp Trp Ala  Gly Ile Gln Gln Glu  Phe Gly Ile
    1190                 1195                 1200

Pro Tyr  Asn Pro Gln Ser Gln  Gly Val Val Ala Ser  Met Asn Lys
    1205                 1210                 1215

Glu Leu  Lys Lys Ile Ile Gly  Gln Val Arg Asp Gln  Ala Glu His
    1220                 1225                 1230

Leu Lys  Thr Ala Val Gln Met  Ala Val Phe Ile His  Asn Phe Lys
    1235                 1240                 1245

Arg Lys  Gly Gly Ile Gly Gly  Tyr Ser Ala Gly Glu  Arg Ile Ile
    1250                 1255                 1260

Asp Ile  Ile Ala Thr Asp Ile  Gln Thr Lys Glu Leu  Gln Lys Gln
    1265                 1270                 1275

Ile Ile  Lys Ile Gln Asn Phe  Arg Val Tyr Tyr Arg  Asp Ser Arg
    1280                 1285                 1290

Asp Pro  Ile Trp Lys Gly Pro  Ala Lys Leu Leu Trp  Lys Gly Glu
    1295                 1300                 1305

Gly Ala  Val Val Ile Gln Asp  Asn Ser Asp Ile Lys  Val Val Pro
    1310                 1315                 1320

Arg Arg  Lys Val Lys Ile Ile  Lys Asp Tyr Gly Lys  Gln Met Ala
    1325                 1330                 1335

Gly Ala  Asp Cys Val Ala Gly  Arg Gln Asp Glu Asp
    1340                 1345                 1350
```

<210> 4
<211> 1341
<212> PRT
<213> Artificial Sequence

<220>
<223> Mos2.Gag-Pol mosaic antigen sequence

<400> 4

```
Met Gly Ala Arg Ala Ser Ile Leu Arg Gly Gly Lys Leu Asp Lys Trp
1               5                  10                  15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys His Tyr Met Leu Lys
            20                  25                  30
```

```
His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro
        35                  40                  45

Gly Leu Leu Glu Thr Ser Glu Gly Cys Lys Gln Ile Ile Lys Gln Leu
    50                  55                  60

Gln Pro Ala Leu Gln Thr Gly Thr Glu Glu Leu Arg Ser Leu Phe Asn
65                  70                  75                  80

Thr Val Ala Thr Leu Tyr Cys Val His Ala Glu Ile Glu Val Arg Asp
                85                  90                  95

Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Gln
            100                 105                 110

Gln Lys Thr Gln Gln Ala Lys Glu Ala Asp Gly Lys Val Ser Gln Asn
        115                 120                 125

Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His Gln Pro Ile
    130                 135                 140

Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Ile Glu Glu Lys Ala
145                 150                 155                 160

Phe Ser Pro Glu Val Ile Pro Met Phe Thr Ala Leu Ser Glu Gly Ala
                165                 170                 175

Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln
            180                 185                 190

Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu Ala Ala Glu
        195                 200                 205

Trp Asp Arg Leu His Pro Val His Ala Gly Pro Val Ala Pro Gly Gln
    210                 215                 220

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

Gln Glu Gln Ile Ala Trp Met Thr Ser Asn Pro Pro Ile Pro Val Gly
                245                 250                 255

Asp Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg
            260                 265                 270

Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Lys Gln Gly Pro Lys Glu
        275                 280                 285

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

```
Gln Ala Thr Gln Asp Val Lys Asn Trp Met Thr Asp Thr Leu Leu Val
305                 310                 315                 320

Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Arg Ala Leu Gly Pro
                325                 330                 335

Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly
            340                 345                 350

Pro Ser His Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln Thr Asn
        355                 360                 365

Ser Thr Ile Leu Met Gln Arg Ser Asn Phe Lys Gly Ser Lys Arg Ile
    370                 375                 380

Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His Ile Ala Arg Asn Cys
385                 390                 395                 400

Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys Glu Gly His
                405                 410                 415

Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu Gly Lys Ile
            420                 425                 430

Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe Leu Gln Ser Arg Pro
        435                 440                 445

Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Arg Phe Glu Glu Thr Thr
    450                 455                 460

Pro Ala Pro Lys Gln Glu Pro Lys Asp Arg Glu Pro Leu Thr Ser Leu
465                 470                 475                 480

Arg Ser Leu Phe Gly Ser Asp Pro Leu Ser Gln Met Ala Pro Ile Ser
                485                 490                 495

Pro Ile Glu Thr Val Pro Val Lys Leu Lys Pro Gly Met Asp Gly Pro
            500                 505                 510

Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val
        515                 520                 525

Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly
    530                 535                 540

Pro Glu Asn Pro Tyr Asn Thr Pro Ile Phe Ala Ile Lys Lys Lys Asp
545                 550                 555                 560

Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg
                565                 570                 575
```

```
Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly
            580                 585                 590

Leu Lys Lys Lys Lys Ser Val Thr Val Leu Ala Val Gly Asp Ala Tyr
        595                 600                 605

Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr
    610                 615                 620

Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn
625                 630                 635                 640

Val Leu Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser
                645                 650                 655

Met Thr Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val
            660                 665                 670

Ile Tyr Gln Tyr Met Ala Ala Leu Tyr Val Gly Ser Asp Leu Glu Ile
        675                 680                 685

Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg
    690                 695                 700

Trp Gly Phe Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe
705                 710                 715                 720

Leu Trp Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro
                725                 730                 735

Ile Val Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys
            740                 745                 750

Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Ala Gly Ile Lys
        755                 760                 765

Val Lys Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu
    770                 775                 780

Val Val Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg
785                 790                 795                 800

Glu Ile Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys
                805                 810                 815

Asp Leu Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr
            820                 825                 830

Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala
        835                 840                 845
```

```
Arg Met Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala
    850                 855                 860

Val Gln Lys Ile Ala Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro
865                 870                 875                 880

Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Ala Trp Trp Thr
                885                 890                 895

Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr
            900                 905                 910

Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val
        915                 920                 925

Gly Ala Glu Thr Phe Tyr Val Ala Gly Ala Ala Asn Arg Glu Thr Lys
    930                 935                 940

Leu Gly Lys Ala Gly Tyr Val Thr Asp Arg Gly Arg Gln Lys Val Val
945                 950                 955                 960

Ser Leu Thr Asp Thr Thr Asn Gln Lys Thr Ala Leu Gln Ala Ile His
                965                 970                 975

Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Ala Ser
            980                 985                 990

Gln Tyr Ala Leu Gly Ile Ile Gln  Ala Gln Pro Asp Lys  Ser Glu Ser
        995                 1000                1005

Glu Leu  Val Ser Gln Ile Ile  Glu Gln Leu Ile Lys  Lys Glu Lys
    1010                1015                1020

Val Tyr  Leu Ala Trp Val Pro  Ala His Lys Gly Ile  Gly Gly Asn
    1025                1030                1035

Glu Gln  Val Asp Lys Leu Val  Ser Arg Gly Ile Arg  Lys Val Leu
    1040                1045                1050

Phe Leu  Asp Gly Ile Asp Lys  Ala Gln Glu Glu His  Glu Lys Tyr
    1055                1060                1065

His Ser  Asn Trp Arg Ala Met  Ala Ser Glu Phe Asn  Leu Pro Pro
    1070                1075                1080

Ile Val  Ala Lys Glu Ile Val  Ala Ser Cys Asp Lys  Cys Gln Leu
    1085                1090                1095

Lys Gly  Glu Ala Ile His Gly  Gln Val Asp Cys Ser  Pro Gly Ile
    1100                1105                1110
```

```
Trp Gln  Leu Ala Cys Thr His  Leu Glu Gly Lys Val  Ile Leu Val
    1115                 1120                 1125

Ala Val  His Val Ala Ser Gly  Tyr Ile Glu Ala Glu  Val Ile Pro
    1130                 1135                 1140

Ala Glu  Thr Gly Gln Glu Thr  Ala Tyr Phe Leu Leu  Lys Leu Ala
    1145                 1150                 1155

Gly Arg  Trp Pro Val Lys Thr  Ile His Thr Ala Asn  Gly Ser Asn
    1160                 1165                 1170

Phe Thr  Ser Ala Thr Val Lys  Ala Ala Cys Trp Trp  Ala Gly Ile
    1175                 1180                 1185

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

Val Ala  Ser Ile Asn Lys Glu  Leu Lys Lys Ile Ile  Gly Gln Val
    1205                 1210                 1215

Arg Asp  Gln Ala Glu His Leu  Lys Thr Ala Val Gln  Met Ala Val
    1220                 1225                 1230

Phe Ile  His Asn Phe Lys Arg  Lys Gly Gly Ile Gly  Glu Tyr Ser
    1235                 1240                 1245

Ala Gly  Glu Arg Ile Val Asp  Ile Ile Ala Ser Asp  Ile Gln Thr
    1250                 1255                 1260

Lys Glu  Leu Gln Lys Gln Ile  Thr Lys Ile Gln Asn  Phe Arg Val
    1265                 1270                 1275

Tyr Tyr  Arg Asp Ser Arg Asp  Pro Leu Trp Lys Gly  Pro Ala Lys
    1280                 1285                 1290

Leu Leu  Trp Lys Gly Glu Gly  Ala Val Val Ile Gln  Asp Asn Ser
    1295                 1300                 1305

Asp Ile  Lys Val Val Pro Arg  Arg Lys Ala Lys Ile  Ile Arg Asp
    1310                 1315                 1320

Tyr Gly  Lys Gln Met Ala Gly  Asp Asp Cys Val Ala  Ser Arg Gln
    1325                 1330                 1335

Asp Glu  Asp
    1340
```

<210> 5
<211> 679

<212> PRT
<213> Artificial Sequence

<220>
<223> clade C gp140 trimeric protein

<400> 5

```
Ala Glu Asn Leu Trp Val Gly Asn Met Trp Val Thr Val Tyr Tyr Gly
1               5                  10                  15

Val Pro Val Trp Thr Asp Ala Lys Thr Thr Leu Phe Cys Ala Ser Asp
            20                  25                  30

Thr Lys Ala Tyr Asp Arg Glu Val His Asn Val Trp Ala Thr His Ala
        35                  40                  45

Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Ile Val Leu Glu Asn Val
    50                  55                  60

Thr Glu Asn Phe Asn Met Trp Lys Asn Asp Met Val Asp Gln Met His
65                  70                  75                  80

Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys
                85                  90                  95

Leu Thr Pro Leu Cys Val Thr Leu His Cys Thr Asn Ala Thr Phe Lys
            100                 105                 110

Asn Asn Val Thr Asn Asp Met Asn Lys Glu Ile Arg Asn Cys Ser Phe
        115                 120                 125

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

Phe Tyr Arg Pro Asp Ile Val Leu Leu Lys Glu Asn Arg Asn Asn Ser
145                 150                 155                 160

Asn Asn Ser Glu Tyr Ile Leu Ile Asn Cys Asn Ala Ser Thr Ile Thr
                165                 170                 175

Gln Ala Cys Pro Lys Val Asn Phe Asp Pro Ile Pro Ile His Tyr Cys
            180                 185                 190

Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Ser
        195                 200                 205

Gly Lys Gly Pro Cys Asn Asn Val Ser Thr Val Gln Cys Thr His Gly
    210                 215                 220

Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala
225                 230                 235                 240
```

```
Glu Lys Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr Asp Asn Val Lys
                245                 250                 255

Thr Ile Ile Val His Leu Asn Lys Ser Val Glu Ile Val Cys Thr Arg
            260                 265                 270

Pro Asn Asn Asn Thr Arg Lys Ser Met Arg Ile Gly Pro Gly Gln Thr
        275                 280                 285

Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala Tyr Cys
    290                 295                 300

Asn Ile Ser Gly Ser Lys Trp Asn Glu Thr Leu Lys Arg Val Lys Glu
305                 310                 315                 320

Lys Leu Gln Glu Asn Tyr Asn Asn Asn Lys Thr Ile Lys Phe Ala Pro
                325                 330                 335

Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn Cys Arg
            340                 345                 350

Gly Glu Phe Phe Tyr Cys Asn Thr Thr Arg Leu Phe Asn Asn Asn Ala
        355                 360                 365

Thr Glu Asp Glu Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile
    370                 375                 380

Asn Met Trp Gln Gly Val Gly Arg Ala Met Tyr Ala Pro Pro Ile Ala
385                 390                 395                 400

Gly Asn Ile Thr Cys Lys Ser Asn Ile Thr Gly Leu Leu Leu Val Arg
                405                 410                 415

Asp Gly Gly Glu Asp Asn Lys Thr Glu Glu Ile Phe Arg Pro Gly Gly
            420                 425                 430

Gly Asn Met Lys Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val
        435                 440                 445

Ile Glu Leu Lys Pro Leu Gly Ile Ala Pro Thr Gly Ala Lys Glu Arg
    450                 455                 460

Val Val Glu Arg Glu Glu Arg Ala Val Gly Ile Gly Ala Val Phe Leu
465                 470                 475                 480

Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Leu Thr
                485                 490                 495

Leu Thr Val Gln Ala Arg Gln Leu Leu Ser Ser Ile Val Gln Gln Gln
            500                 505                 510
```

Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu
515 520 525

Thr Val Trp Gly Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu
530 535 540

Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly
545 550 555 560

Lys Leu Ile Cys Thr Thr Asn Val Pro Trp Asn Ser Ser Trp Ser Asn
565 570 575

Lys Ser Gln Thr Asp Ile Trp Asn Asn Met Thr Trp Met Glu Trp Asp
580 585 590

Arg Glu Ile Ser Asn Tyr Thr Asp Thr Ile Tyr Arg Leu Leu Glu Asp
595 600 605

Ser Gln Thr Gln Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp
610 615 620

Ser Trp Lys Asn Leu Trp Ser Trp Phe Asp Ile Ser Asn Trp Leu Trp
625 630 635 640

Tyr Ile Lys Ser Arg Ile Glu Gly Arg Gly Ser Gly Gly Tyr Ile Pro
645 650 655

Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp
660 665 670

Val Leu Leu Ser Thr Phe Leu
675

<210> 6
<211> 695
<212> PRT
<213> Artificial Sequence

<220>
<223> mosaic gp140 chimeric protein

<400> 6

Ala Gly Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys
1 5 10 15

Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp
20 25 30

Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp
35 40 45

Pro Asn Pro Gln Glu Val Val Leu Glu Asn Val Thr Glu Asn Phe Asn
50 55 60

```
Met Trp Lys Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser
65                  70                  75                  80

Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys
                85                  90                  95

Val Thr Leu Asn Cys Thr Asp Asp Val Arg Asn Val Thr Asn Asn Ala
            100                 105                 110

Thr Asn Thr Asn Ser Ser Trp Gly Glu Pro Met Glu Lys Gly Glu Ile
        115                 120                 125

Lys Asn Cys Ser Phe Asn Ile Thr Thr Ser Ile Arg Asn Lys Val Gln
    130                 135                 140

Lys Gln Tyr Ala Leu Phe Tyr Lys Leu Asp Val Val Pro Ile Asp Asn
145                 150                 155                 160

Asp Ser Asn Asn Thr Asn Tyr Arg Leu Ile Ser Cys Asn Thr Ser Val
                165                 170                 175

Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro Ile His
            180                 185                 190

Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Lys
        195                 200                 205

Phe Asn Gly Thr Gly Pro Cys Thr Asn Val Ser Thr Val Gln Cys Thr
    210                 215                 220

His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser
225                 230                 235                 240

Leu Ala Glu Glu Glu Val Val Ile Arg Ser Glu Asn Phe Thr Asn Asn
                245                 250                 255

Ala Lys Thr Ile Met Val Gln Leu Asn Val Ser Val Glu Ile Asn Cys
            260                 265                 270

Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile His Ile Gly Pro Gly
        275                 280                 285

Arg Ala Phe Tyr Thr Ala Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala
    290                 295                 300

His Cys Asn Ile Ser Arg Ala Asn Trp Asn Asn Thr Leu Arg Gln Ile
305                 310                 315                 320

Val Glu Lys Leu Gly Lys Gln Phe Gly Asn Asn Lys Thr Ile Val Phe
                325                 330                 335
```

```
Asn His Ser Ser Gly Gly Asp Pro Glu Ile Val Met His Ser Phe Asn
            340                 345                 350

Cys Gly Gly Glu Phe Phe Tyr Cys Asn Ser Thr Lys Leu Phe Asn Ser
        355                 360                 365

Thr Trp Thr Trp Asn Asn Ser Thr Trp Asn Asn Thr Lys Arg Ser Asn
    370                 375                 380

Asp Thr Glu Glu His Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile
385                 390                 395                 400

Asn Met Trp Gln Glu Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg
                405                 410                 415

Gly Gln Ile Arg Cys Ser Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg
            420                 425                 430

Asp Gly Gly Asn Asp Thr Ser Gly Thr Glu Ile Phe Arg Pro Gly Gly
        435                 440                 445

Gly Asp Met Arg Asp Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val
    450                 455                 460

Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr Lys Ala Lys Glu Arg
465                 470                 475                 480

Val Val Gln Arg Glu Glu Arg Ala Val Gly Ile Gly Ala Val Phe Leu
                485                 490                 495

Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Met Thr
            500                 505                 510

Leu Thr Val Gln Ala Arg Leu Leu Leu Ser Gly Ile Val Gln Gln Gln
        515                 520                 525

Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu
    530                 535                 540

Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu
545                 550                 555                 560

Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly
                565                 570                 575

Lys Leu Ile Cys Thr Thr Thr Val Pro Trp Asn Ala Ser Trp Ser Asn
            580                 585                 590

Lys Ser Leu Asp Lys Ile Trp Asn Asn Met Thr Trp Met Glu Trp Glu
        595                 600                 605
```

```
Arg Glu Ile Asn Asn Tyr Thr Ser Leu Ile Tyr Thr Leu Ile Glu Glu
    610                 615                 620

Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp
625                 630                 635                 640

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

Tyr Ile Lys Ser Arg Ile Glu Gly Arg Gly Ser Gly Gly Tyr Ile Pro
            660                 665                 670

Glu Ala Pro Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp
        675                 680                 685

Val Leu Leu Ser Thr Phe Leu
    690                 695
```

<210> 7
<211> 29
<212> PRT
<213> Artificial Sequence

<220>
<223> exemplary leader sequence for gp140 stabilized trimeric protein production

<400> 7

```
Met Arg Val Arg Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp
1               5                   10                  15

Gly Thr Leu Ile Leu Gly Met Leu Met Ile Cys Ser Ala
            20                  25
```