

27-12-2011 ALG-B-0002 PCT -SEQ  
SEQUENCE LISTING

<110> ALGENICS  
CARLIER, Aude  
MICHEL, Rémy  
DUFOURMANTEL, Nathalie  
CADORET, Jean-Paul  
LEJEUNE, Alexandre

<120> PRODUCTION OF HIGH MANNOSE GLYCOSYLATED PROTEINS STORED IN THE  
PLASTID OF MICROALGAE

<130> ALG-B-0002 PCT

<150> EP 10016162.9  
<151> 2010-12-29

<160> 45

<170> PatentIn version 3.5

<210> 1  
<211> 46  
<212> PRT  
<213> Guillardia theta

<400> 1

Met Ile Arg Ala Cys Ala Leu Leu Gly Leu Ala Ala Ser Ala Ala Ala  
1 5 10 15

Phe Ala Pro Ser Ser Leu Pro Ile Arg Ala Asn Arg Ala Ser Ala Val  
20 25 30

Ser Lys Met Ser Met Gln Ser Asn Arg Phe Ser Tyr Arg Ser  
35 40 45

<210> 2  
<211> 55  
<212> PRT  
<213> Phaeodactylum tricornutum

<400> 2

Met Arg Ser Phe Cys Ile Ala Ala Leu Leu Ala Val Ala Ser Ala Phe  
1 5 10 15

Thr Thr Gln Pro Thr Ser Phe Thr Val Lys Thr Ala Asn Val Gly Glu  
20 25 30

Arg Ala Ser Gly Val Phe Pro Glu Gln Ser Ser Ala His Arg Thr Arg  
35 40 45

Lys Ala Thr Ile Val Met Asp  
50 55

<210> 3  
<211> 54  
<212> PRT  
<213> Phaeodactylum tricornutum

&lt;400&gt; 3

Met Lys Val Ala Thr Thr Leu Thr Leu Ala Phe Ile Cys Cys Ala Ser  
 1 5 10 15

Ala Phe Gly Leu Asn Gly Gln Thr Thr Ser Val Met Lys Lys Val Gly  
 20 25 30

Phe Asp Ala Gly Ser Lys Pro Met Val Gln Ala Ile Asp Val Gln Gly  
 35 40 45

Asn Arg Leu Gly Ser Asn  
 50

&lt;210&gt; 4

&lt;211&gt; 30

&lt;212&gt; PRT

&lt;213&gt; Phaeodactylum tricornutum

&lt;400&gt; 4

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

Ala Pro Ala Lys Asn Ala Ala Arg Thr Ser Val Ala Thr Asn  
 20 25 30

&lt;210&gt; 5

&lt;211&gt; 104

&lt;212&gt; PRT

&lt;213&gt; Phaeodactylum tricornutum

&lt;400&gt; 5

Met Gly Arg Gly Val Ile Ile Phe Cys Val Lys Asn Phe Ala Val Trp  
 1 5 10 15

Leu Leu Ile Ile Thr Ser Ala Val Ser Ile Gln Ala Trp Ile Pro Leu  
 20 25 30

Pro Leu Ser Ala Thr Val Lys Ala Arg Ile Asp Ser Thr Thr Leu Phe  
 35 40 45

Phe Ser Arg Tyr Lys Thr Pro Leu Tyr His Gly Gly Asn Glu Glu Ser  
 50 55 60

Tyr Gly Pro Pro Ala Pro Ala Val Asp Ser Arg Tyr Tyr Thr Tyr Val  
 65 70 75 80

Glu Ala Pro Val Gln Ser Ser Arg Ser Arg Asp Thr Lys Gln Pro Ile  
 85 90 95

Thr Leu Ser Arg Phe Leu Ser Asp  
 100

27-12-2011 ALG-B-0002 PCT -SEQ

<210> 6  
 <211> 43  
 <212> PRT  
 <213> Phaeodactylum tricornutum

<400> 6

Met Lys Phe Thr Ala Ala Cys Ser Ile Ala Leu Ala Ala Ser Ala Ser  
 1 5 10 15

Ala Phe Ala Pro Ile Pro Ser Val Ser Arg Thr Thr Asp Leu Ser Met  
 20 25 30

Ser Leu Gln Lys Asp Leu Ala Asn Val Gly Lys  
 35 40

<210> 7  
 <211> 497  
 <212> PRT  
 <213> Homo sapiens

<400> 7

Ala Arg Pro Cys Ile Pro Lys Ser Phe Gly Tyr Ser Ser Val Val Cys  
 1 5 10 15

Val Cys Asn Ala Thr Tyr Cys Asp Ser Phe Asp Pro Pro Thr Phe Pro  
 20 25 30

Ala Leu Gly Thr Phe Ser Arg Tyr Glu Ser Thr Arg Ser Gly Arg Arg  
 35 40 45

Met Glu Leu Ser Met Gly Pro Ile Gln Ala Asn His Thr Gly Thr Gly  
 50 55 60

Leu Leu Leu Thr Leu Gln Pro Glu Gln Lys Phe Gln Lys Val Lys Gly  
 65 70 75 80

Phe Gly Gly Ala Met Thr Asp Ala Ala Ala Leu Asn Ile Leu Ala Leu  
 85 90 95

Ser Pro Pro Ala Gln Asn Leu Leu Leu Lys Ser Tyr Phe Ser Glu Glu  
 100 105 110

Gly Ile Gly Tyr Asn Ile Ile Arg Val Pro Met Ala Ser Cys Asp Phe  
 115 120 125

Ser Ile Arg Thr Tyr Thr Tyr Ala Asp Thr Pro Asp Asp Phe Gln Leu  
 130 135 140

His Asn Phe Ser Leu Pro Glu Glu Asp Thr Lys Leu Lys Ile Pro Leu  
 145 150 155 160

Ile His Arg Ala Leu Gln Leu Ala Gln Arg Pro Val Ser Leu Leu Ala  
 165 170 175

27-12-2011 ALG-B-0002 PCT -SEQ

Ser Pro Trp Thr 180 Ser Pro Thr Trp 185 Lys Thr Asn Gly 190 Ala Val Asn  
 Gly Lys Gly 195 Ser Leu Lys Gly Gln 200 Pro Gly Asp Ile Tyr 205 His Gln Thr  
 Trp Ala Arg Tyr Phe Val 210 Lys 215 Phe Leu Asp Ala Tyr 220 Ala Glu His Lys  
 Leu Gln Phe Trp Ala Val 230 Thr Ala Glu Asn Glu 235 Pro Ser Ala Gly Leu 240  
 Leu Ser Gly Tyr 245 Phe Gln Cys Leu Gly 250 Phe Thr Pro Glu His 255 Gln  
 Arg Asp Phe Ile 260 Ala Arg Asp Leu Gly 265 Pro Thr Leu Ala Asn 270 Ser Thr  
 His His Asn Val 275 Arg Leu Leu Met 280 Leu Asp Asp Gln Arg 285 Leu Leu Leu  
 Pro His Trp Ala Lys Val 290 Val 295 Leu Thr Asp Pro Glu 300 Ala Ala Lys Tyr  
 Val 305 His Gly Ile Ala Val 310 His Trp Tyr Leu Asp 315 Phe Leu Ala Pro Ala 320  
 Lys Ala Thr Leu Gly 325 Glu Thr His Arg Leu 330 Phe Pro Asn Thr Met 335 Leu  
 Phe Ala Ser Glu 340 Ala Cys Val Gly Ser 345 Lys Phe Trp Glu Gln 350 Ser Val  
 Arg Leu Gly 355 Ser Trp Asp Arg Gly 360 Met Gln Tyr Ser His 365 Ser Ile Ile  
 Thr Asn 370 Leu Leu Tyr His Val 375 Val Gly Trp Thr Asp 380 Trp Asn Leu Ala  
 Leu 385 Asn Pro Glu Gly Gly 390 Pro Asn Trp Val Arg 395 Asn Phe Val Asp Ser 400  
 Pro Ile Ile Val 405 Asp Ile Thr Lys Asp Thr 410 Phe Tyr Lys Gln Pro Met 415  
 Phe Tyr His Leu 420 Gly His Phe Ser Lys 425 Phe Ile Pro Glu Gly 430 Ser Gln  
 Arg Val Gly 435 Leu Val Ala Ser Gln 440 Lys Asn Asp Leu Asp 445 Ala Val Ala

27-12-2011 ALG-B-0002 PCT -SEQ

Leu Met His Pro Asp Gly Ser Ala Val Val Val Val Leu Asn Arg Ser  
450 455 460

Ser Lys Asp Val Pro Leu Thr Ile Lys Asp Pro Ala Val Gly Phe Leu  
465 470 475 480

Glu Thr Ile Ser Pro Gly Tyr Ser Ile His Thr Tyr Leu Trp His Arg  
485 490 495

Gln

<210> 8  
<211> 398  
<212> PRT  
<213> Homo sapiens

<400> 8

Leu Asp Asn Gly Leu Ala Arg Thr Pro Thr Met Gly Trp Leu His Trp  
1 5 10 15

Glu Arg Phe Met Cys Asn Leu Asp Cys Gln Glu Glu Pro Asp Ser Cys  
20 25 30

Ile Ser Glu Lys Leu Phe Met Glu Met Ala Glu Leu Met Val Ser Glu  
35 40 45

Gly Trp Lys Asp Ala Gly Tyr Glu Tyr Leu Cys Ile Asp Asp Cys Trp  
50 55 60

Met Ala Pro Gln Arg Asp Ser Glu Gly Arg Leu Gln Ala Asp Pro Gln  
65 70 75 80

Arg Phe Pro His Gly Ile Arg Gln Leu Ala Asn Tyr Val His Ser Lys  
85 90 95

Gly Leu Lys Leu Gly Ile Tyr Ala Asp Val Gly Asn Lys Thr Cys Ala  
100 105 110

Gly Phe Pro Gly Ser Phe Gly Tyr Tyr Asp Ile Asp Ala Gln Thr Phe  
115 120 125

Ala Asp Trp Gly Val Asp Leu Leu Lys Phe Asp Gly Cys Tyr Cys Asp  
130 135 140

Ser Leu Glu Asn Leu Ala Asp Gly Tyr Lys His Met Ser Leu Ala Leu  
145 150 155 160

Asn Arg Thr Gly Arg Ser Ile Val Tyr Ser Cys Glu Trp Pro Leu Tyr  
165 170 175

27-12-2011 ALG-B-0002 PCT -SEQ

Met Trp Pro Phe Gln Lys Pro Asn Tyr Thr Glu Ile Arg Gln Tyr Cys  
180 185 190

Asn His Trp Arg Asn Phe Ala Asp Ile Asp Asp Ser Trp Lys Ser Ile  
195 200 205

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

Val Ala Gly Pro Gly Gly Trp Asn Asp Pro Asp Met Leu Val Ile Gly  
225 230 235 240

Asn Phe Gly Leu Ser Trp Asn Gln Gln Val Thr Gln Met Ala Leu Trp  
245 250 255

Ala Ile Met Ala Ala Pro Leu Phe Met Ser Asn Asp Leu Arg His Ile  
260 265 270

Ser Pro Gln Ala Lys Ala Leu Leu Gln Asp Lys Asp Val Ile Ala Ile  
275 280 285

Asn Gln Asp Pro Leu Gly Lys Gln Gly Tyr Gln Leu Arg Gln Gly Asp  
290 295 300

Asn Phe Glu Val Trp Glu Arg Pro Leu Ser Gly Leu Ala Trp Ala Val  
305 310 315 320

Ala Met Ile Asn Arg Gln Glu Ile Gly Gly Pro Arg Ser Tyr Thr Ile  
325 330 335

Ala Val Ala Ser Leu Gly Lys Gly Val Ala Cys Asn Pro Ala Cys Phe  
340 345 350

Ile Thr Gln Leu Leu Pro Val Lys Arg Lys Leu Gly Phe Tyr Glu Trp  
355 360 365

Thr Ser Arg Leu Arg Ser His Ile Asn Pro Thr Gly Thr Val Leu Leu  
370 375 380

Gln Leu Glu Asn Thr Met Gln Met Ser Leu Lys Asp Leu Leu  
385 390 395

<210> 9  
<211> 883  
<212> PRT  
<213> Homo sapiens  
<400> 9

Ala His Pro Gly Arg Pro Arg Ala Val Pro Thr Gln Cys Asp Val Pro  
1 5 10 15

27-12-2011 ALG-B-0002 PCT -SEQ

Pro Asn Ser Arg Phe Asp Cys Ala Pro Asp Lys Ala Ile Thr Gln Glu  
20 25 30

Gln Cys Glu Ala Arg Gly Cys Cys Tyr Ile Pro Ala Lys Gln Gly Leu  
35 40 45

Gln Gly Ala Gln Met Gly Gln Pro Trp Cys Phe Phe Pro Pro Ser Tyr  
50 55 60

Pro Ser Tyr Lys Leu Glu Asn Leu Ser Ser Ser Glu Met Gly Tyr Thr  
65 70 75 80

Ala Thr Leu Thr Arg Thr Thr Pro Thr Phe Phe Pro Lys Asp Ile Leu  
85 90 95

Thr Leu Arg Leu Asp Val Met Met Glu Thr Glu Asn Arg Leu His Phe  
100 105 110

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

Pro His Val His Ser Arg Ala Pro Ser Pro Leu Tyr Ser Val Glu Phe  
130 135 140

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

Val Leu Leu Asn Thr Thr Val Ala Pro Leu Phe Phe Ala Asp Gln Phe  
165 170 175

Leu Gln Leu Ser Thr Ser Leu Pro Ser Gln Tyr Ile Thr Gly Leu Ala  
180 185 190

Glu His Leu Ser Pro Leu Met Leu Ser Thr Ser Trp Thr Arg Ile Thr  
195 200 205

Leu Trp Asn Arg Asp Leu Ala Pro Thr Pro Gly Ala Asn Leu Tyr Gly  
210 215 220

Ser His Pro Phe Tyr Leu Ala Leu Glu Asp Gly Gly Ser Ala His Gly  
225 230 235 240

Val Phe Leu Leu Asn Ser Asn Ala Met Asp Val Val Leu Gln Pro Ser  
245 250 255

Pro Ala Leu Ser Trp Arg Ser Thr Gly Gly Ile Leu Asp Val Tyr Ile  
260 265 270

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

Val Gly Tyr Pro Phe Met Pro Pro Tyr Trp Gly Leu Gly Phe His Leu  
 290 295 300

Cys Arg Trp Gly Tyr Ser Ser Thr Ala Ile Thr Arg Gln Val Val Glu  
 305 310 315 320

Asn Met Thr Arg Ala His Phe Pro Leu Asp Val Gln Trp Asn Asp Leu  
 325 330 335

Asp Tyr Met Asp Ser Arg Arg Asp Phe Thr Phe Asn Lys Asp Gly Phe  
 340 345 350

Arg Asp Phe Pro Ala Met Val Gln Glu Leu His Gln Gly Gly Arg Arg  
 355 360 365

Tyr Met Met Ile Val Asp Pro Ala Ile Ser Ser Ser Gly Pro Ala Gly  
 370 375 380

Ser Tyr Arg Pro Tyr Asp Glu Gly Leu Arg Arg Gly Val Phe Ile Thr  
 385 390 395 400

Asn Glu Thr Gly Gln Pro Leu Ile Gly Lys Val Trp Pro Gly Ser Thr  
 405 410 415

Ala Phe Pro Asp Phe Thr Asn Pro Thr Ala Leu Ala Trp Trp Glu Asp  
 420 425 430

Met Val Ala Glu Phe His Asp Gln Val Pro Phe Asp Gly Met Trp Ile  
 435 440 445

Asp Met Asn Glu Pro Ser Asn Phe Ile Arg Gly Ser Glu Asp Gly Cys  
 450 455 460

Pro Asn Asn Glu Leu Glu Asn Pro Pro Tyr Val Pro Gly Val Val Gly  
 465 470 475 480

Gly Thr Leu Gln Ala Ala Thr Ile Cys Ala Ser Ser His Gln Phe Leu  
 485 490 495

Ser Thr His Tyr Asn Leu His Asn Leu Tyr Gly Leu Thr Glu Ala Ile  
 500 505 510

Ala Ser His Arg Ala Leu Val Lys Ala Arg Gly Thr Arg Pro Phe Val  
 515 520 525

Ile Ser Arg Ser Thr Phe Ala Gly His Gly Arg Tyr Ala Gly His Trp  
 530 535 540

Thr Gly Asp Val Trp Ser Ser Trp Glu Gln Leu Ala Ser Ser Val Pro  
 545 550 555 560



Glu Ile Leu Gln Phe Asn Leu Leu Gly Val Pro Leu Val Gly Ala Asp  
 565 570 575  
 Val Cys Gly Phe Leu Gly Asn Thr Ser Glu Glu Leu Cys Val Arg Trp  
 580 585 590  
 Thr Gln Leu Gly Ala Phe Tyr Pro Phe Met Arg Asn His Asn Ser Leu  
 595 600 605  
 Leu Ser Leu Pro Gln Glu Pro Tyr Ser Phe Ser Glu Pro Ala Gln Gln  
 610 615 620  
 Ala Met Arg Lys Ala Leu Thr Leu Arg Tyr Ala Leu Leu Pro His Leu  
 625 630 635 640  
 Tyr Thr Leu Phe His Gln Ala His Val Ala Gly Glu Thr Val Ala Arg  
 645 650 655  
 Pro Leu Phe Leu Glu Phe Pro Lys Asp Ser Ser Thr Trp Thr Val Asp  
 660 665 670  
 His Gln Leu Leu Trp Gly Glu Ala Leu Leu Ile Thr Pro Val Leu Gln  
 675 680 685  
 Ala Gly Lys Ala Glu Val Thr Gly Tyr Phe Pro Leu Gly Thr Trp Tyr  
 690 695 700  
 Asp Leu Gln Thr Val Pro Val Glu Ala Leu Gly Ser Leu Pro Pro Pro  
 705 710 715 720  
 Pro Ala Ala Pro Arg Glu Pro Ala Ile His Ser Glu Gly Gln Trp Val  
 725 730 735  
 Thr Leu Pro Ala Pro Leu Asp Thr Ile Asn Val His Leu Arg Ala Gly  
 740 745 750  
 Tyr Ile Ile Pro Leu Gln Gly Pro Gly Leu Thr Thr Thr Glu Ser Arg  
 755 760 765  
 Gln Gln Pro Met Ala Leu Ala Val Ala Leu Thr Lys Gly Gly Glu Ala  
 770 775 780  
 Arg Gly Glu Leu Phe Trp Asp Asp Gly Glu Ser Leu Glu Val Leu Glu  
 785 790 795 800  
 Arg Gly Ala Tyr Thr Gln Val Ile Phe Leu Ala Arg Asn Asn Thr Ile  
 805 810 815  
 Val Asn Glu Leu Val Arg Val Thr Ser Glu Gly Ala Gly Leu Gln Leu  
 820 825 830

Gln Lys Val Thr Val Leu Gly Val Ala Thr Ala Pro Gln Gln Val Leu  
835 840 845

Ser Asn Gly Val Pro Val Ser Asn Phe Thr Tyr Ser Pro Asp Thr Lys  
850 855 860

Val Leu Asp Ile Cys Val Ser Leu Leu Met Gly Glu Gln Phe Leu Val  
865 870 875 880

Ser Trp Cys

<210> 10  
<211> 626  
<212> PRT  
<213> Homo sapiens

<400> 10

Ala Pro His Leu Val His Val Asp Ala Ala Arg Ala Leu Trp Pro Leu  
1 5 10 15

Arg Arg Phe Trp Arg Ser Thr Gly Phe Cys Pro Pro Leu Pro His Ser  
20 25 30

Gln Ala Asp Gln Tyr Val Leu Ser Trp Asp Gln Gln Leu Asn Leu Ala  
35 40 45

Tyr Val Gly Ala Val Pro His Arg Gly Ile Lys Gln Val Arg Thr His  
50 55 60

Trp Leu Leu Glu Leu Val Thr Thr Arg Gly Ser Thr Gly Arg Gly Leu  
65 70 75 80

Ser Tyr Asn Phe Thr His Leu Asp Gly Tyr Leu Asp Leu Leu Arg Glu  
85 90 95

Asn Gln Leu Leu Pro Gly Phe Glu Leu Met Gly Ser Ala Ser Gly His  
100 105 110

Phe Thr Asp Phe Glu Asp Lys Gln Gln Val Phe Glu Trp Lys Asp Leu  
115 120 125

Val Ser Ser Leu Ala Arg Arg Tyr Ile Gly Arg Tyr Gly Leu Ala His  
130 135 140

Val Ser Lys Trp Asn Phe Glu Thr Trp Asn Glu Pro Asp His His Asp  
145 150 155 160

Phe Asp Asn Val Ser Met Thr Met Gln Gly Phe Leu Asn Tyr Tyr Asp  
165 170 175

Ala Cys Ser Glu Gly Leu Arg Ala Ala Ser Pro Ala Leu Arg Leu Gly

180

Gly Pro Gly Asp Ser Phe His Thr Pro Pro Arg Ser Pro Leu Ser Trp  
195 200 205

Gly Leu Leu Arg His Cys His Asp Gly Thr Asn Phe Phe Thr Gly Glu  
210 215 220

Ala Gly Val Arg Leu Asp Tyr Ile Ser Leu His Arg Lys Gly Ala Arg  
225 230 235 240

Ser Ser Ile Ser Ile Leu Glu Gln Glu Lys Val Val Ala Gln Gln Ile  
245 250 255

Arg Gln Leu Phe Pro Lys Phe Ala Asp Thr Pro Ile Tyr Asn Asp Glu  
260 265 270

Ala Asp Pro Leu Val Gly Trp Ser Leu Pro Gln Pro Trp Arg Ala Asp  
275 280 285

Val Thr Tyr Ala Ala Met Val Val Lys Val Ile Ala Gln His Gln Asn  
290 295 300

Leu Leu Leu Ala Asn Thr Thr Ser Ala Phe Pro Tyr Ala Leu Leu Ser  
305 310 315 320

Asn Asp Asn Ala Phe Leu Ser Tyr His Pro His Pro Phe Ala Gln Arg  
325 330 335

Thr Leu Thr Ala Arg Phe Gln Val Asn Asn Thr Arg Pro Pro His Val  
340 345 350

Gln Leu Leu Arg Lys Pro Val Leu Thr Ala Met Gly Leu Leu Ala Leu  
355 360 365

Leu Asp Glu Glu Gln Leu Trp Ala Glu Val Ser Gln Ala Gly Thr Val  
370 375 380

Leu Asp Ser Asn His Thr Val Gly Val Leu Ala Ser Ala His Arg Pro  
385 390 395 400

Gln Gly Pro Ala Asp Ala Trp Arg Ala Ala Val Leu Ile Tyr Ala Ser  
405 410 415

Asp Asp Thr Arg Ala His Pro Asn Arg Ser Val Ala Val Thr Leu Arg  
420 425 430

Leu Arg Gly Val Pro Pro Gly Pro Gly Leu Val Tyr Val Thr Arg Tyr  
435 440 445

Leu Asp Asn Gly Leu Cys Ser Pro Asp Gly Glu Trp Arg Arg Leu Gly

450

Arg Pro Val Phe Pro Thr Ala Glu Gln Phe Arg Arg Met Arg Ala Ala  
465 470 475 480

Glu Asp Pro Val Ala Ala Ala Pro Arg Pro Leu Pro Ala Gly Gly Arg  
485 490 495

Leu Thr Leu Arg Pro Ala Leu Arg Leu Pro Ser Leu Leu Leu Val His  
500 505 510

Val Cys Ala Arg Pro Glu Lys Pro Pro Gly Gln Val Thr Arg Leu Arg  
515 520 525

Ala Leu Pro Leu Thr Gln Gly Gln Leu Val Leu Val Trp Ser Asp Glu  
530 535 540

His Val Gly Ser Lys Cys Leu Trp Thr Tyr Glu Ile Gln Phe Ser Gln  
545 550 555 560

Asp Gly Lys Ala Tyr Thr Pro Val Ser Arg Lys Pro Ser Thr Phe Asn  
565 570 575

Leu Phe Val Phe Ser Pro Asp Thr Gly Ala Val Ser Gly Ser Tyr Arg  
580 585 590

Val Arg Ala Leu Asp Tyr Trp Ala Arg Pro Gly Pro Phe Ser Asp Pro  
595 600 605

Val Pro Tyr Leu Glu Val Pro Val Pro Arg Gly Pro Pro Ser Pro Gly  
610 615 620

Asn Pro  
625

<210> 11  
<211> 517  
<212> PRT  
<213> Homo sapiens  
<400> 11

Thr Asp Ala Leu Asn Val Leu Leu Ile Ile Val Asp Asp Leu Arg Pro  
1 5 10 15

Ser Leu Gly Cys Tyr Gly Asp Lys Leu Val Arg Ser Pro Asn Ile Asp  
20 25 30

Gln Leu Ala Ser His Ser Leu Leu Phe Gln Asn Ala Phe Ala Gln Gln  
35 40 45

Ala Val Cys Ala Pro Ser Arg Val Ser Phe Leu Thr Gly Arg Arg Pro  
50 55 60

27-12-2011 ALG-B-0002 PCT -SEQ

Asp Thr Thr Arg Leu Tyr Asp Phe Asn Ser Tyr Trp Arg Val His Ala  
 65 70 75 80  
 Gly Asn Phe Ser Thr Ile Pro Gln Tyr Phe Lys Glu Asn Gly Tyr Val  
 85 90 95  
 Thr Met Ser Val Gly Lys Val Phe His Pro Gly Ile Ser Ser Asn His  
 100 105 110  
 Thr Asp Asp Ser Pro Tyr Ser Trp Ser Phe Pro Pro Tyr His Pro Ser  
 115 120 125  
 Ser Glu Lys Tyr Glu Asn Thr Lys Thr Cys Arg Gly Pro Asp Gly Glu  
 130 135 140  
 Leu His Ala Asn Leu Leu Cys Pro Val Asp Val Leu Asp Val Pro Glu  
 145 150 155 160  
 Gly Thr Leu Pro Asp Lys Gln Ser Thr Glu Gln Ala Ile Gln Leu Leu  
 165 170 175  
 Glu Lys Met Lys Thr Ser Ala Ser Pro Phe Phe Leu Ala Val Gly Tyr  
 180 185 190  
 His Lys Pro His Ile Pro Phe Arg Tyr Pro Lys Glu Phe Gln Lys Leu  
 195 200 205  
 Tyr Pro Leu Glu Asn Ile Thr Leu Ala Pro Asp Pro Glu Val Pro Asp  
 210 215 220  
 Gly Leu Pro Pro Val Ala Tyr Asn Pro Trp Met Asp Ile Arg Gln Arg  
 225 230 235 240  
 Glu Asp Val Gln Ala Leu Asn Ile Ser Val Pro Tyr Gly Pro Ile Pro  
 245 250 255  
 Val Asp Phe Gln Arg Lys Ile Arg Gln Ser Tyr Phe Ala Ser Val Ser  
 260 265 270  
 Tyr Leu Asp Thr Gln Val Gly Arg Leu Leu Ser Ala Leu Asp Asp Leu  
 275 280 285  
 Gln Leu Ala Asn Ser Thr Ile Ile Ala Phe Thr Ser Asp His Gly Trp  
 290 295 300  
 Ala Leu Gly Glu His Gly Glu Trp Ala Lys Tyr Ser Asn Phe Asp Val  
 305 310 315 320  
 Ala Thr His Val Pro Leu Ile Phe Tyr Val Pro Gly Arg Thr Ala Ser  
 325 330 335

27-12-2011 ALG-B-0002 PCT -SEQ

Leu Pro Glu Ala Gly Glu Lys Leu Phe Pro Tyr Leu Asp Pro Phe Asp  
340 345 350

Ser Ala Ser Gln Leu Met Glu Pro Gly Arg Gln Ser Met Asp Leu Val  
355 360 365

Glu Leu Val Ser Leu Phe Pro Thr Leu Ala Gly Leu Ala Gly Leu Gln  
370 375 380

Val Pro Pro Arg Cys Pro Val Pro Ser Phe His Val Glu Leu Cys Arg  
385 390 395 400

Glu Gly Lys Asn Leu Leu Lys His Phe Arg Phe Arg Asp Leu Glu Glu  
405 410 415

Asp Pro Tyr Leu Pro Gly Asn Pro Arg Glu Leu Ile Ala Tyr Ser Gln  
420 425 430

Tyr Pro Arg Pro Ser Asp Ile Pro Gln Trp Asn Ser Asp Lys Pro Ser  
435 440 445

Leu Lys Asp Ile Lys Ile Met Gly Tyr Ser Ile Arg Thr Ile Asp Tyr  
450 455 460

Arg Tyr Thr Val Trp Val Gly Phe Asn Pro Asp Glu Phe Leu Ala Asn  
465 470 475 480

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

Leu Gln Asp His Asn Met Tyr Asn Asp Ser Gln Gly Gly Asp Leu Phe  
500 505 510

Gln Leu Leu Met Pro  
515

<210> 12  
<211> 497  
<212> PRT  
<213> Homo sapiens

<400> 12

Ser Gly Ala Gly Ala Ser Arg Pro Pro His Leu Val Phe Leu Leu Ala  
1 5 10 15

Asp Asp Leu Gly Trp Asn Asp Val Gly Phe His Gly Ser Arg Ile Arg  
20 25 30

Thr Pro His Leu Asp Ala Leu Ala Ala Gly Gly Val Leu Leu Asp Asn  
35 40 45

27-12-2011 ALG-B-0002 PCT -SEQ

Tyr Tyr Thr Gln Pro Leu Cys Thr Pro Ser Arg Ser Gln Leu Leu Thr  
 50 55 60  
 Gly Arg Tyr Gln Ile Arg Thr Gly Leu Gln His Gln Ile Ile Trp Pro  
 65 70 75 80  
 Cys Gln Pro Ser Cys Val Pro Leu Asp Glu Lys Leu Leu Pro Gln Leu  
 85 90 95  
 Leu Lys Glu Ala Gly Tyr Thr Thr His Met Val Gly Lys Trp His Leu  
 100 105 110  
 Gly Met Tyr Arg Lys Glu Cys Leu Pro Thr Arg Arg Gly Phe Asp Thr  
 115 120 125  
 Tyr Phe Gly Tyr Leu Leu Gly Ser Glu Asp Tyr Tyr Ser His Glu Arg  
 130 135 140  
 Cys Thr Leu Ile Asp Ala Leu Asn Val Thr Arg Cys Ala Leu Asp Phe  
 145 150 155 160  
 Arg Asp Gly Glu Glu Val Ala Thr Gly Tyr Lys Asn Met Tyr Ser Thr  
 165 170 175  
 Asn Ile Phe Thr Lys Arg Ala Ile Ala Leu Ile Thr Asn His Pro Pro  
 180 185 190  
 Glu Lys Pro Leu Phe Leu Tyr Leu Ala Leu Gln Ser Val His Glu Pro  
 195 200 205  
 Leu Gln Val Pro Glu Glu Tyr Leu Lys Pro Tyr Asp Phe Ile Gln Asp  
 210 215 220  
 Lys Asn Arg His His Tyr Ala Gly Met Val Ser Leu Met Asp Glu Ala  
 225 230 235 240  
 Val Gly Asn Val Thr Ala Ala Leu Lys Ser Ser Gly Leu Trp Asn Asn  
 245 250 255  
 Thr Val Phe Ile Phe Ser Thr Asp Asn Gly Gly Gln Thr Leu Ala Gly  
 260 265 270  
 Gly Asn Asn Trp Pro Leu Arg Gly Arg Lys Trp Ser Leu Trp Glu Gly  
 275 280 285  
 Gly Val Arg Gly Val Gly Phe Val Ala Ser Pro Leu Leu Lys Gln Lys  
 290 295 300  
 Gly Val Lys Asn Arg Glu Leu Ile His Ile Ser Asp Trp Leu Pro Thr  
 305 310 315 320

27-12-2011 ALG-B-0002 PCT -SEQ

Leu Val Lys Leu Ala Arg Gly His Thr Asn Gly Thr Lys Pro Leu Asp  
325 330 335

Gly Phe Asp Val Trp Lys Thr Ile Ser Glu Gly Ser Pro Ser Pro Arg  
340 345 350

Ile Glu Leu Leu His Asn Ile Asp Pro Asn Phe Val Asp Ser Ser Pro  
355 360 365

Cys Pro Arg Asn Ser Met Ala Pro Ala Lys Asp Asp Ser Ser Leu Pro  
370 375 380

Glu Tyr Ser Ala Phe Asn Thr Ser Val His Ala Ala Ile Arg His Gly  
385 390 395 400

Asn Trp Lys Leu Leu Thr Gly Tyr Pro Gly Cys Gly Tyr Trp Phe Pro  
405 410 415

Pro Pro Ser Gln Tyr Asn Val Ser Glu Ile Pro Ser Ser Asp Pro Pro  
420 425 430

Thr Lys Thr Leu Trp Leu Phe Asp Ile Asp Arg Asp Pro Glu Glu Arg  
435 440 445

His Asp Leu Ser Arg Glu Tyr Pro His Ile Val Thr Lys Leu Leu Ser  
450 455 460

Arg Leu Gln Phe Tyr His Lys His Ser Val Pro Val Tyr Phe Pro Ala  
465 470 475 480

Gln Asp Pro Arg Cys Asp Pro Lys Ala Thr Gly Val Trp Gly Pro Trp  
485 490 495

Met

<210> 13  
<211> 583  
<212> PRT  
<213> Homo sapiens

<400> 13

Leu Ser Asp Ser Arg Val Leu Trp Ala Pro Ala Glu Ala His Pro Leu  
1 5 10 15

Ser Pro Gln Gly His Pro Ala Arg Leu His Arg Ile Val Pro Arg Leu  
20 25 30

Arg Asp Val Phe Gly Trp Gly Asn Leu Thr Cys Pro Ile Cys Lys Gly  
35 40 45



27-12-2011 ALG-B-0002 PCT -SEQ

Leu Phe Thr Ala Ile Asn 50 Leu Gly Leu Lys Lys Glu Pro Asn Val Ala 55 60

Arg Val Gly Ser Val Ala Ile Lys Leu Cys Asn Leu Leu Lys Ile Ala 65 70 75 80

Pro Pro Ala Val Cys Gln Ser Ile Val His Leu Phe Glu Asp Asp Met 85 90 95

Val Glu Val Trp Arg Arg Ser Val Leu Ser Pro Ser Glu Ala Cys Gly 100 105 110

Leu Leu Leu Gly Ser Thr Cys Gly His Trp Asp Ile Phe Ser Ser Trp 115 120 125

Asn Ile Ser Leu Pro Thr Val Pro Lys Pro Pro Pro Lys Pro Pro Ser 130 135 140

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

Leu His Trp Asp His Asp Tyr Leu Glu Gly Thr Asp Pro Asp Cys Ala 165 170 175

Asp Pro Leu Cys Cys Arg Arg Gly Ser Gly Leu Pro Pro Ala Ser Arg 180 185 190

Pro Gly Ala Gly Tyr Trp Gly Glu Tyr Ser Lys Cys Asp Leu Pro Leu 195 200 205

Arg Thr Leu Glu Ser Leu Leu Ser Gly Leu Gly Pro Ala Gly Pro Phe 210 215 220

Asp Met Val Tyr Trp Thr Gly Asp Ile Pro Ala His Asp Val Trp His 225 230 235 240

Gln Thr Arg Gln Asp Gln Leu Arg Ala Leu Thr Thr Val Thr Ala Leu 245 250 255

Val Arg Lys Phe Leu Gly Pro Val Pro Val Tyr Pro Ala Val Gly Asn 260 265 270

His Glu Ser Thr Pro Val Asn Ser Phe Pro Pro Pro Phe Ile Glu Gly 275 280 285

Asn His Ser Ser Arg Trp Leu Tyr Glu Ala Met Ala Lys Ala Trp Glu 290 295 300

Pro Trp Leu Pro Ala Glu Ala Leu Arg Thr Leu Arg Ile Gly Gly Phe 305 310 315 320

27-12-2011 ALG-B-0002 PCT -SEQ

Tyr Ala Leu Ser Pro Tyr Pro Gly Leu Arg Leu Ile Ser Leu Asn Met  
325 330 335

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

Pro Ala Gly Gln Leu Gln Trp Leu Val Gly Glu Leu Gln Ala Ala Glu  
355 360 365

Asp Arg Gly Asp Lys Val His Ile Ile Gly His Ile Pro Pro Gly His  
370 375 380

Cys Leu Lys Ser Trp Ser Trp Asn Tyr Tyr Arg Ile Val Ala Arg Tyr  
385 390 395 400

Glu Asn Thr Leu Ala Ala Gln Phe Phe Gly His Thr His Val Asp Glu  
405 410 415

Phe Glu Val Phe Tyr Asp Glu Glu Thr Leu Ser Arg Pro Leu Ala Val  
420 425 430

Ala Phe Leu Ala Pro Ser Ala Thr Thr Tyr Ile Gly Leu Asn Pro Gly  
435 440 445

Tyr Arg Val Tyr Gln Ile Asp Gly Asn Tyr Ser Gly Ser Ser His Val  
450 455 460

Val Leu Asp His Glu Thr Tyr Ile Leu Asn Leu Thr Gln Ala Asn Ile  
465 470 475 480

Pro Gly Ala Ile Pro His Trp Gln Leu Leu Tyr Arg Ala Arg Glu Thr  
485 490 495

Tyr Gly Leu Pro Asn Thr Leu Pro Thr Ala Trp His Asn Leu Val Tyr  
500 505 510

Arg Met Arg Gly Asp Met Gln Leu Phe Gln Thr Phe Trp Phe Leu Tyr  
515 520 525

His Lys Gly His Pro Pro Ser Glu Pro Cys Gly Thr Pro Cys Arg Leu  
530 535 540

Ala Thr Leu Cys Ala Gln Leu Ser Ala Arg Ala Asp Ser Pro Ala Leu  
545 550 555 560

Cys Arg His Leu Met Pro Asp Gly Ser Leu Pro Glu Ala Gln Ser Leu  
565 570 575

Trp Pro Arg Pro Leu Phe Cys  
580

27-12-2011 ALG-B-0002 PCT -SEQ

<210> 14  
 <211> 378  
 <212> PRT  
 <213> Homo sapiens

<400> 14

Ser Gly Gly Lys Leu Thr Ala Val Asp Pro Glu Thr Asn Met Asn Val  
 1 5 10 15

Ser Glu Ile Ile Ser Tyr Trp Gly Phe Pro Ser Glu Glu Tyr Leu Val  
 20 25 30

Glu Thr Glu Asp Gly Tyr Ile Leu Cys Leu Asn Arg Ile Pro His Gly  
 35 40 45

Arg Lys Asn His Ser Asp Lys Gly Pro Lys Pro Val Val Phe Leu Gln  
 50 55 60

His Gly Leu Leu Ala Asp Ser Ser Asn Trp Val Thr Asn Leu Ala Asn  
 65 70 75 80

Ser Ser Leu Gly Phe Ile Leu Ala Asp Ala Gly Phe Asp Val Trp Met  
 85 90 95

Gly Asn Ser Arg Gly Asn Thr Trp Ser Arg Lys His Lys Thr Leu Ser  
 100 105 110

Val Ser Gln Asp Glu Phe Trp Ala Phe Ser Tyr Asp Glu Met Ala Lys  
 115 120 125

Tyr Asp Leu Pro Ala Ser Ile Asn Phe Ile Leu Asn Lys Thr Gly Gln  
 130 135 140

Glu Gln Val Tyr Tyr Val Gly His Ser Gln Gly Thr Thr Ile Gly Phe  
 145 150 155 160

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

Phe Ala Leu Gly Pro Val Ala Ser Val Ala Phe Cys Thr Ser Pro Met  
 180 185 190

Ala Lys Leu Gly Arg Leu Pro Asp His Leu Ile Lys Asp Leu Phe Gly  
 195 200 205

Asp Lys Glu Phe Leu Pro Gln Ser Ala Phe Leu Lys Trp Leu Gly Thr  
 210 215 220

His Val Cys Thr His Val Ile Leu Lys Glu Leu Cys Gly Asn Leu Cys  
 225 230 235 240

Phe Leu Leu Cys Gly Phe Asn Glu Arg Asn Leu Asn Met Ser Arg Val

245 27-12-2011 ALG-B-0002 PCT -SEQ 255  
250

Asp Val Tyr Thr Thr His Ser Pro Ala Gly Thr Ser Val Gln Asn Met  
260 265 270

Leu His Trp Ser Gln Ala Val Lys Phe Gln Lys Phe Gln Ala Phe Asp  
275 280 285

Trp Gly Ser Ser Ala Lys Asn Tyr Phe His Tyr Asn Gln Ser Tyr Pro  
290 295 300

Pro Thr Tyr Asn Val Lys Asp Met Leu Val Pro Thr Ala Val Trp Ser  
305 310 315 320

Gly Gly His Asp Trp Leu Ala Asp Val Tyr Asp Val Asn Ile Leu Leu  
325 330 335

Thr Gln Ile Thr Asn Leu Val Phe His Glu Ser Ile Pro Glu Trp Glu  
340 345 350

His Leu Asp Phe Ile Trp Gly Leu Asp Ala Pro Trp Arg Leu Tyr Asn  
355 360 365

Lys Ile Ile Asn Leu Met Arg Lys Tyr Gln  
370 375

<210> 15  
<211> 483  
<212> PRT  
<213> Human immunodeficiency virus 1

<400> 15

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

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

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

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

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

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

Leu Cys Val Ser Leu Lys Cys Thr Asp Leu Lys Asn Asp Thr Asn Thr  
100 105 110

27-12-2011 ALG-B-0002 PCT -SEQ

Asn Ser Ser Ser Gly Arg Met Ile Met Glu Lys Gly Glu Ile Lys Asn  
115 120 125

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

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

Thr Ser Tyr Lys Leu Thr Ser Cys Asn Thr Ser Val Ile Thr Gln Ala  
165 170 175

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

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

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

Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu  
225 230 235 240

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

Ile Val Gln Leu Asn Thr Ser Val Glu Ile Asn Cys Thr Arg Pro Asn  
260 265 270

Asn Asn Thr Arg Lys Arg Ile Arg Ile Gln Arg Gly Pro Gly Arg Ala  
275 280 285

Phe Val Thr Ile Gly Lys Ile Gly Asn Met Arg Gln Ala His Cys Asn  
290 295 300

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

Leu Arg Glu Gln Phe Gly Asn Asn Lys Thr Ile Ile Phe Lys Gln Ser  
325 330 335

Ser Gly Gly Asp Pro Glu Ile Val Thr His Ser Phe Asn Cys Gly Gly  
340 345 350

Glu Phe Phe Tyr Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Phe  
355 360 365

Asn Ser Thr Trp Ser Thr Glu Gly Ser Asn Asn Thr Glu Gly Ser Asp  
370 375 380

27-12-2011 ALG-B-0002 PCT -SEQ

Thr Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln  
385 390 395 400

Lys Val Gly Lys Ala Met Tyr Ala Pro Pro Ile Ser Gly Gln Ile Arg  
405 410 415

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

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

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

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

Glu Lys Arg

<210> 16  
<211> 345  
<212> PRT  
<213> Human immunodeficiency virus 1  
<400> 16

Ala Val Gly Ile Gly Ala Leu Phe Leu Gly Phe Leu Gly Ala Ala Gly  
1 5 10 15

Ser Thr Met Gly Ala Ala Ser Met Thr Leu Thr Val Gln Ala Arg Gln  
20 25 30

Leu Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu Leu Arg Ala Ile  
35 40 45

Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln  
50 55 60

Leu Gln Ala Arg Ile Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln  
65 70 75 80

Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala  
85 90 95

Val Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu Glu Gln Ile Trp  
100 105 110

Asn His Thr Thr Trp Met Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr  
115 120 125

27-12-2011 ALG-B-0002 PCT -SEQ

Ser Leu Ile His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys  
130 135 140

Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn  
145 150 155 160

Trp Phe Asn Ile Thr Asn Trp Leu Trp Tyr Ile Lys Leu Phe Ile Met  
165 170 175

Ile Val Gly Gly Leu Val Gly Leu Arg Ile Val Phe Ala Val Leu Ser  
180 185 190

Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr  
195 200 205

His Leu Pro Thr Pro Arg Gly Pro Asp Arg Pro Glu Gly Ile Glu Glu  
210 215 220

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

Ser Leu Ala Leu Ile Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser  
245 250 255

Tyr His Arg Leu Arg Asp Leu Leu Leu Ile Val Thr Arg Ile Val Glu  
260 265 270

Leu Leu Gly Arg Arg Gly Trp Glu Ala Leu Lys Tyr Trp Trp Asn Leu  
275 280 285

Leu Gln Tyr Trp Ser Gln Glu Leu Lys Asn Ser Ala Val Ser Leu Leu  
290 295 300

Asn Ala Thr Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Val Ile Glu  
305 310 315 320

Val Val Gln Gly Ala Cys Arg Ala Ile Arg His Ile Pro Arg Arg Ile  
325 330 335

Arg Gln Gly Leu Glu Arg Ile Leu Leu  
340 345

<210> 17  
<211> 192  
<212> PRT  
<213> Hepatitis C virus  
<400> 17

Tyr Gln Val Arg Asn Ser Ser Gly Leu Tyr His Val Thr Asn Asp Cys  
1 5 10 15

Pro Asn Ser Ser Val Val Tyr Glu Ala Ala Asp Ala Ile Leu His Thr  
20 25 30

Pro Gly Cys Val Pro Cys Val Arg Glu Gly Asn Ala Ser Arg Cys Trp  
35 40 45

Val Ala Val Thr Pro Thr Val Ala Thr Arg Asp Gly Lys Leu Pro Thr  
50 55 60

Thr Gln Leu Arg Arg His Ile Asp Leu Leu Val Gly Ser Ala Thr Leu  
65 70 75 80

Cys Ser Ala Leu Tyr Val Gly Asp Leu Cys Gly Ser Val Phe Leu Val  
85 90 95

Gly Gln Leu Phe Thr Phe Ser Pro Arg His His Trp Thr Thr Gln Asp  
100 105 110

Cys Asn Cys Ser Ile Tyr Pro Gly His Ile Thr Gly His Arg Met Ala  
115 120 125

Trp Asn Met Met Met Asn Trp Ser Pro Thr Ala Ala Leu Val Val Ala  
130 135 140

Gln Leu Leu Arg Ile Pro Gln Ala Ile Met Asp Met Ile Ala Gly Ala  
145 150 155 160

His Trp Gly Val Leu Ala Gly Ile Lys Tyr Phe Ser Met Val Gly Asn  
165 170 175

Trp Ala Lys Val Leu Val Val Leu Leu Leu Phe Ala Gly Val Asp Ala  
180 185 190

<210> 18  
<211> 363  
<212> PRT  
<213> Hepatitis C virus  
<400> 18

Glu Thr His Val Thr Gly Gly Asn Ala Gly Arg Thr Thr Ala Gly Leu  
1 5 10 15

Val Gly Leu Leu Thr Pro Gly Ala Lys Gln Asn Ile Gln Leu Ile Asn  
20 25 30

Thr Asn Gly Ser Trp His Ile Asn Ser Thr Ala Leu Asn Cys Asn Glu  
35 40 45

Ser Leu Asn Thr Gly Trp Leu Ala Gly Leu Phe Tyr Gln His Lys Phe  
50 55 60

Asn Ser Ser Gly Cys Pro Glu Arg Leu Ala Ser Cys Arg Arg Leu Thr



65

70

75

80

Asp Phe Ala Gln Gly Trp Gly Pro Ile Ser Tyr Ala Asn Gly Ser Gly  
85 90 95

Leu Asp Glu Arg Pro Tyr Cys Trp His Tyr Pro Pro Arg Pro Cys Gly  
100 105 110

Ile Val Pro Ala Lys Ser Val Cys Gly Pro Val Tyr Cys Phe Thr Pro  
115 120 125

Ser Pro Val Val Val Gly Thr Thr Asp Arg Ser Gly Ala Pro Thr Tyr  
130 135 140

Ser Trp Gly Ala Asn Asp Thr Asp Val Phe Val Leu Asn Asn Thr Arg  
145 150 155 160

Pro Pro Leu Gly Asn Trp Phe Gly Cys Thr Trp Met Asn Ser Thr Gly  
165 170 175

Phe Thr Lys Val Cys Gly Ala Pro Pro Cys Val Ile Gly Gly Val Gly  
180 185 190

Asn Asn Thr Leu Leu Cys Pro Thr Asp Cys Phe Arg Lys Tyr Pro Glu  
195 200 205

Ala Thr Tyr Ser Arg Cys Gly Ser Gly Pro Arg Ile Thr Pro Arg Cys  
210 215 220

Met Val Asp Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys Thr Ile Asn  
225 230 235 240

Tyr Thr Ile Phe Lys Val Arg Met Tyr Val Gly Gly Val Glu His Arg  
245 250 255

Leu Glu Ala Ala Cys Asn Trp Thr Arg Gly Glu Arg Cys Asp Leu Glu  
260 265 270

Asp Arg Asp Arg Ser Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Gln  
275 280 285

Trp Gln Val Leu Pro Cys Ser Phe Thr Thr Leu Pro Ala Leu Ser Thr  
290 295 300

Gly Leu Ile His Leu His Gln Asn Ile Val Asp Val Gln Tyr Leu Tyr  
305 310 315 320

Gly Val Gly Ser Ser Ile Ala Ser Trp Ala Ile Lys Trp Glu Tyr Val  
325 330 335

Val Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg Val Cys Ser Cys Leu

Trp Met Met Leu Leu Ile Ser Gln Ala Glu Ala  
355 360

<210> 19  
<211> 495  
<212> PRT  
<213> Dengue virus 1  
<400> 19

Met Arg Cys Val Gly Ile Gly Asn Arg Asp Phe Val Glu Gly Leu Ser  
1 5 10 15

Gly Ala Thr Trp Val Asp Val Val Leu Glu His Gly Ser Cys Val Thr  
20 25 30

Thr Met Ala Lys Asn Lys Pro Thr Leu Asp Ile Glu Leu Lys Thr  
35 40 45

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

Ile Ser Asn Thr Thr Thr Asp Ser Arg Cys Pro Thr Gln Gly Glu Ala  
65 70 75 80

Thr Leu Val Glu Glu Gln Asp Ala Asn Phe Val Cys Arg Arg Thr Phe  
85 90 95

Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser  
100 105 110

Leu Leu Thr Cys Ala Lys Phe Lys Cys Val Thr Lys Leu Glu Gly Lys  
115 120 125

Ile Val Gln Tyr Glu Asn Leu Lys Tyr Ser Val Ile Val Thr Val His  
130 135 140

Thr Gly Asp Gln His Gln Val Gly Asn Glu Thr Thr Glu His Gly Thr  
145 150 155 160

Ile Ala Thr Ile Thr Pro Gln Ala Pro Thr Ser Glu Ile Gln Leu Thr  
165 170 175

Asp Tyr Gly Ala Leu Thr Leu Asp Cys Ser Pro Arg Thr Gly Leu Asp  
180 185 190

Phe Asn Glu Met Val Leu Leu Thr Met Lys Glu Lys Ser Trp Leu Val  
195 200 205

His Lys Gln Trp Phe Leu Asp Leu Pro Leu Pro Trp Thr Ser Gly Ala  
210 215 220

27-12-2011 ALG-B-0002 PCT -SEQ

Ser Thr Ser Gln Glu Thr Trp Asn Arg Gln Asp Leu Leu Val Thr Phe  
225 230 235 240

Lys Thr Ala His Ala Lys Lys Gln Glu Val Val Val Leu Gly Ser Gln  
245 250 255

Glu Gly Ala Met His Thr Ala Leu Thr Gly Ala Thr Glu Ile Gln Thr  
260 265 270

Ser Gly Thr Thr Thr Ile Phe Ala Gly His Leu Lys Cys Arg Leu Lys  
275 280 285

Met Asp Lys Leu Thr Leu Lys Gly Thr Ser Tyr Val Met Cys Thr Gly  
290 295 300

Ser Phe Lys Leu Glu Lys Glu Val Ala Glu Thr Gln His Gly Thr Val  
305 310 315 320

Leu Val Gln Val Lys Tyr Glu Gly Thr Asp Ala Pro Cys Lys Ile Pro  
325 330 335

Phe Ser Thr Gln Asp Glu Lys Gly Val Thr Gln Asn Gly Arg Leu Ile  
340 345 350

Thr Ala Asn Pro Ile Val Thr Asp Lys Glu Lys Pro Val Asn Ile Glu  
355 360 365

Thr Glu Pro Pro Phe Gly Glu Ser Tyr Ile Val Val Gly Ala Gly Glu  
370 375 380

Lys Ala Leu Lys Leu Ser Trp Phe Lys Lys Gly Ser Ser Ile Gly Lys  
385 390 395 400

Met Phe Glu Ala Thr Ala Arg Gly Ala Arg Arg Met Ala Ile Leu Gly  
405 410 415

Asp Thr Ala Trp Asp Phe Gly Ser Ile Gly Gly Val Phe Thr Ser Val  
420 425 430

Gly Lys Leu Val His Gln Val Phe Gly Thr Ala Tyr Gly Val Leu Phe  
435 440 445

Ser Gly Val Ser Trp Thr Met Lys Ile Gly Ile Gly Ile Leu Leu Thr  
450 455 460

Trp Leu Gly Leu Asn Ser Arg Ser Thr Ser Leu Ser Met Thr Cys Ile  
465 470 475 480

Ala Val Gly Met Val Thr Leu Tyr Leu Gly Val Met Val Gln Ala  
485 490 495

27-12-2011 ALG-B-0002 PCT -SEQ

<210> 20  
 <211> 501  
 <212> PRT  
 <213> West Nile Virus

<400> 20

Phe Asn Cys Leu Gly Met Ser Asn Arg Asp Phe Leu Glu Gly Val Ser  
 1 5 10 15

Gly Ala Thr Trp Val Asp Leu Val Leu Glu Gly Asp Ser Cys Val Thr  
 20 25 30

Ile Met Ser Lys Asp Lys Pro Thr Ile Asp Val Lys Met Met Asn Met  
 35 40 45

Glu Ala Ala Asn Leu Ala Glu Val Arg Ser Tyr Cys Tyr Leu Ala Thr  
 50 55 60

Val Ser Asp Leu Ser Thr Lys Ala Ala Cys Pro Thr Met Gly Glu Ala  
 65 70 75 80

His Asn Asp Lys Arg Ala Asp Pro Ala Phe Val Cys Arg Gln Gly Val  
 85 90 95

Val Asp Arg Gly Trp Gly Asn Gly Cys Gly Leu Phe Gly Lys Gly Ser  
 100 105 110

Ile Asp Thr Cys Ala Lys Phe Ala Cys Ser Thr Lys Ala Ile Gly Arg  
 115 120 125

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

Gly Pro Thr Thr Val Glu Ser His Gly Asn Tyr Ser Thr Gln Ala Gly  
 145 150 155 160

Ala Thr Gln Ala Gly Arg Phe Ser Ile Thr Pro Ala Ala Pro Ser Tyr  
 165 170 175

Thr Leu Lys Leu Gly Glu Tyr Gly Glu Val Thr Val Asp Cys Glu Pro  
 180 185 190

Arg Ser Gly Ile Asp Thr Asn Ala Tyr Tyr Val Met Thr Val Gly Thr  
 195 200 205

Lys Thr Phe Leu Val His Arg Glu Trp Phe Met Asp Leu Asn Leu Pro  
 210 215 220

Trp Ser Ser Ala Gly Ser Thr Val Trp Arg Asn Arg Glu Thr Leu Met  
 225 230 235 240

27-12-2011 ALG-B-0002 PCT -SEQ

Glu Phe Glu Glu Pro His Ala Thr Lys Gln Ser Val Ile Ala Leu Gly  
245 250 255

Ser Gln Glu Gly Ala Leu His Gln Ala Leu Ala Gly Ala Ile Pro Val  
260 265 270

Glu Phe Ser Ser Asn Thr Val Lys Leu Thr Ser Gly His Leu Lys Cys  
275 280 285

Arg Val Lys Met Glu Lys Leu Gln Leu Lys Gly Thr Thr Tyr Gly Val  
290 295 300

Cys Ser Lys Ala Phe Lys Phe Leu Gly Thr Pro Ala Asp Thr Gly His  
305 310 315 320

Gly Thr Val Val Leu Glu Leu Gln Tyr Thr Gly Thr Asp Gly Pro Cys  
325 330 335

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

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

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

Val Val Gly Arg Gly Glu Gln Gln Ile Asn His His Trp His Lys Ser  
385 390 395 400

Gly Ser Ser Ile Gly Lys Ala Phe Thr Thr Thr Leu Lys Gly Ala Gln  
405 410 415

Arg Leu Ala Ala Leu Gly Asp Thr Ala Trp Asp Phe Gly Ser Val Gly  
420 425 430

Gly Val Phe Thr Ser Val Gly Lys Ala Val His Gln Val Phe Gly Gly  
435 440 445

Ala Phe Arg Ser Leu Phe Gly Gly Met Ser Trp Ile Thr Gln Gly Leu  
450 455 460

Leu Gly Ala Leu Leu Leu Trp Met Gly Ile Asn Ala Arg Asp Arg Ser  
465 470 475 480

Ile Ala Leu Thr Phe Leu Ala Val Gly Gly Val Leu Leu Phe Leu Ser  
485 490 495

Val Asn Val His Ala  
500

27-12-2011 ALG-B-0002 PCT -SEQ

<210> 21  
 <211> 676  
 <212> PRT  
 <213> Cote d'Ivoire ebolavirus

<400> 21

```

Met Gly Ala Ser Gly Ile Leu Gln Leu Pro Arg Glu Arg Phe Arg Lys
1          5          10          15

Thr Ser Phe Phe Val Trp Val Ile Ile Leu Phe His Lys Val Phe Ser
          20          25          30

Ile Pro Leu Gly Val Val His Asn Asn Thr Leu Gln Val Ser Asp Ile
          35          40          45

Asp Lys Phe Val Cys Arg Asp Lys Leu Ser Ser Thr Ser Gln Leu Lys
50          55          60

Ser Val Gly Leu Asn Leu Glu Gly Asn Gly Val Ala Thr Asp Val Pro
65          70          75          80

Thr Ala Thr Lys Arg Trp Gly Phe Arg Ala Gly Val Pro Pro Lys Val
          85          90          95

Val Asn Cys Glu Ala Gly Glu Trp Ala Glu Asn Cys Tyr Asn Leu Ala
100          105          110

Ile Lys Lys Val Asp Gly Ser Glu Cys Leu Pro Glu Ala Pro Glu Gly
115          120          125

Val Arg Asp Phe Pro Arg Cys Arg Tyr Val His Lys Val Ser Gly Thr
130          135          140

Gly Pro Cys Pro Gly Gly Leu Ala Phe His Lys Glu Gly Ala Phe Phe
145          150          155          160

Leu Tyr Asp Arg Leu Ala Ser Thr Ile Ile Tyr Arg Gly Thr Thr Phe
165          170          175

Ala Glu Gly Val Ile Ala Phe Leu Ile Leu Pro Lys Ala Arg Lys Asp
180          185          190

Phe Phe Gln Ser Pro Pro Leu His Glu Pro Ala Asn Met Thr Thr Asp
195          200          205

Pro Ser Ser Tyr Tyr His Thr Thr Thr Ile Asn Tyr Val Val Asp Asn
210          215          220

Phe Gly Thr Asn Thr Thr Glu Phe Leu Phe Gln Val Asp His Leu Thr
225          230          235          240
    
```

27-12-2011 ALG-B-0002 PCT -SEQ

Tyr Val Gln Leu Glu Ala Arg Phe Thr Pro Gln Phe Leu Val Leu Leu  
245 250 255

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

Leu Ile Trp Lys Ile Asn Pro Thr Val Asp Thr Ser Met Gly Glu Trp  
275 280 285

Ala Phe Trp Glu Asn Lys Lys Asn Phe Thr Lys Thr Leu Ser Ser Glu  
290 295 300

Glu Leu Ser Phe Val Pro Val Pro Glu Thr Gln Asn Gln Val Leu Asp  
305 310 315 320

Thr Thr Ala Thr Val Ser Pro Pro Ile Ser Ala His Asn His Ala Ala  
325 330 335

Glu Asp His Lys Glu Leu Val Ser Glu Asp Ser Thr Pro Val Val Gln  
340 345 350

Met Gln Asn Ile Lys Gly Lys Asp Thr Met Pro Thr Thr Val Thr Gly  
355 360 365

Val Pro Thr Thr Thr Pro Ser Pro Phe Pro Ile Asn Ala Arg Asn Thr  
370 375 380

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

Ser Thr Thr Gln Pro Ala Lys Thr Thr Ser Gln Pro Thr Asn Ser Thr  
405 410 415

Glu Ser Thr Thr Leu Asn Pro Thr Ser Glu Pro Ser Ser Arg Gly Thr  
420 425 430

Gly Pro Ser Ser Pro Thr Val Pro Asn Thr Thr Glu Ser His Ala Glu  
435 440 445

Leu Gly Lys Thr Thr Pro Thr Thr Leu Pro Glu Gln His Thr Ala Ala  
450 455 460

Ser Ala Ile Pro Arg Ala Val His Pro Asp Glu Leu Ser Gly Pro Gly  
465 470 475 480

Phe Leu Thr Asn Thr Ile Arg Gly Val Thr Asn Leu Leu Thr Gly Ser  
485 490 495

Arg Arg Lys Arg Arg Asp Val Thr Pro Asn Thr Gln Pro Lys Cys Asn  
500 505 510

27-12-2011 ALG-B-0002 PCT -SEQ

Pro Asn Leu His Tyr Trp Thr Ala Leu Asp Glu Gly Ala Ala Ile Gly  
515 520 525

Leu Ala Trp Ile Pro Tyr Phe Gly Pro Ala Ala Glu Gly Ile Tyr Thr  
530 535 540

Glu Gly Ile Met Glu Asn Gln Asn Gly Leu Ile Cys Gly Leu Arg Gln  
545 550 555 560

Leu Ala Asn Glu Thr Thr Gln Ala Leu Gln Leu Phe Leu Arg Ala Thr  
565 570 575

Thr Glu Leu Arg Thr Phe Ser Ile Leu Asn Arg Lys Ala Ile Asp Phe  
580 585 590

Leu Leu Gln Arg Trp Gly Gly Thr Cys His Ile Leu Gly Pro Asp Cys  
595 600 605

Cys Ile Glu Pro Gln Asp Trp Thr Lys Asn Ile Thr Asp Lys Ile Asp  
610 615 620

Gln Ile Ile His Asp Phe Val Asp Asn Asn Leu Pro Asn Gln Asn Asp  
625 630 635 640

Gly Ser Asn Trp Trp Thr Gly Trp Lys Gln Trp Val Pro Ala Gly Ile  
645 650 655

Gly Ile Thr Gly Val Ile Ile Ala Ile Ile Ala Leu Leu Cys Ile Cys  
660 665 670

Lys Phe Met Leu  
675

<210> 22  
<211> 330  
<212> PRT  
<213> Homo sapiens  
<400> 22

Ala Ser Phe Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys  
1 5 10 15

Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr  
20 25 30

Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser  
35 40 45

Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser  
50 55 60

Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr



65

70

75

80

Tyr Ile Cys Asn Val<sub>85</sub> Asn His Lys Pro Ser<sub>90</sub> Asn Thr Lys Val<sub>95</sub> Asp Lys

Lys Val Glu Pro<sub>100</sub> Lys Ser Cys Asp<sub>105</sub> Thr His Thr Cys Pro<sub>110</sub> Pro Cys

Pro Ala Pro<sub>115</sub> Glu Leu Leu Gly<sub>120</sub> Pro Ser Val Phe<sub>125</sub> Leu Phe Pro Pro

Lys Pro<sub>130</sub> Lys Asp Thr Leu Met<sub>135</sub> Ile Ser Arg Thr Pro<sub>140</sub> Glu Val Thr Cys

Val<sub>145</sub> Val Val Asp Val<sub>150</sub> Ser His Glu Asp Pro<sub>155</sub> Val Lys Phe Asn Trp<sub>160</sub>

Tyr Val Asp Gly<sub>165</sub> Glu Val His Asn Ala<sub>170</sub> Lys Thr Lys Pro Arg<sub>175</sub> Glu

Glu Gln Tyr Asn<sub>180</sub> Ser Thr Tyr Arg Val<sub>185</sub> Val Ser Val Leu Thr<sub>190</sub> Val Leu

His Gln Asp<sub>195</sub> Trp Leu Asn Gly<sub>200</sub> Lys Glu Tyr Lys Cys Lys<sub>205</sub> Val Ser Asn

Lys Ala<sub>210</sub> Leu Pro Ala Pro<sub>215</sub> Ile Glu Lys Thr Ile Ser<sub>220</sub> Lys Ala Lys Gly

Gln Pro Arg Glu Pro<sub>230</sub> Gln Val Tyr Thr Leu Pro<sub>235</sub> Pro Ser Arg Asp Glu<sub>240</sub>

Leu Thr Lys Asn<sub>245</sub> Gln Val Ser Leu Thr Cys<sub>250</sub> Leu Val Lys Gly Phe<sub>255</sub> Tyr

Pro Ser Asp Ile<sub>260</sub> Ala Val Glu Trp Glu<sub>265</sub> Ser Asn Gly Gln Pro<sub>270</sub> Glu Asn

Asn Tyr Lys<sub>275</sub> Thr Thr Pro Pro Val<sub>280</sub> Leu Asp Ser Asp Gly<sub>285</sub> Ser Phe Phe

Leu Tyr Ser Lys Leu Thr Val<sub>295</sub> Asp Lys Ser Arg Trp<sub>300</sub> Gln Gln Gly Asn

Val<sub>305</sub> Phe Ser Cys Ser Val<sub>310</sub> Met His Glu Ala Leu<sub>315</sub> His Asn His Tyr Thr<sub>320</sub>

Gln Lys Ser Leu Ser<sub>325</sub> Leu Ser Pro Gly Lys<sub>330</sub>

&lt;210&gt; 23

27-12-2011 ALG-B-0002 PCT -SEQ

<211> 327  
<212> PRT  
<213> Homo sapiens

<400> 23

Ala Ser Phe Lys Gly Pro Ser Val Phe Pro Leu Val Pro Cys Ser Arg  
1 5 10 15

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

Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Cys Ala Leu Thr Ser  
35 40 45

Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser  
50 55 60

Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Lys Thr  
65 70 75 80

Tyr Thr Cys Asn Val Asp His Lys Pro Ser Asn Thr Lys Val Asp Lys  
85 90 95

Arg Val Glu Ser Lys Tyr Gly Pro Pro Cys Pro Ser Cys Pro Ala Pro  
100 105 110

Glu Phe Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys  
115 120 125

Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val  
130 135 140

Asp Val Ser Gln Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val Asp  
145 150 155 160

Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Phe  
165 170 175

Asn Ser Thr Tyr Arg Val Val Arg Val Leu Thr Val Leu His Gln Asp  
180 185 190

Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Gly Leu  
195 200 205

Pro Ser Ser Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg  
210 215 220

Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Gln Glu Glu Met Thr Lys  
225 230 235 240

Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp  
245 250 255

27-12-2011 ALG-B-0002 PCT -SEQ

Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asp Asn Tyr Lys  
260 265 270

Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser  
275 280 285

Arg Leu Thr Val Asp Lys Ser Arg Trp Gln Glu Gly Asn Val Phe Ser  
290 295 300

Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser  
305 310 315 320

Leu Ser Leu Ser Pro Gly Lys  
325

<210> 24  
<211> 97  
<212> PRT  
<213> Homo sapiens

<400> 24

Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly  
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Ser Tyr  
20 25 30

Ala Met Ser Trp Val Arg Gln Ser Pro Gly Lys Gly Leu Gln Trp Val  
35 40 45

Ser Ala Ile Ser Gly Ser Gly Ile Ser Thr Tyr Tyr Ala Asp Ser Val  
50 55 60

Arg Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr  
65 70 75 80

Leu Gln Met Ser Ser Leu Ser Arg Gly His Gly Arg Ile Leu Leu Cys  
85 90 95

Glu

<210> 25  
<211> 214  
<212> PRT  
<213> Homo sapiens

<400> 25

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

Asp Arg Val Thr Ile Ser Cys Arg Met Ser Gln Gly Ile Ser Asn Tyr  
20 25 30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Asp Leu Leu Ile  
35 40 45

Tyr Ala Ala Ser Thr Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly  
50 55 60

Ser Gly Ser Gly Thr Asp Phe Ile Leu Thr Ile Ser Arg Leu Gln Ser  
65 70 75 80

Glu Asp Phe Ala Ile Tyr Tyr Cys Gln Gln Tyr Tyr Ser Phe Pro Phe  
85 90 95

Thr Phe Gly Pro Gly Thr Lys Val Asp Ile Lys Arg Thr Val Ala Ala  
100 105 110

Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly  
115 120 125

Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala  
130 135 140

Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln  
145 150 155 160

Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser  
165 170 175

Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr  
180 185 190

Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser  
195 200 205

Phe Asn Arg Gly Glu Cys  
210

<210> 26  
<211> 1404  
<212> DNA  
<213> artificial sequence

<220>  
<223> CDS coding for a fusion protein between EPO and GFP

<400> 26  
atgaaggctg ctaccacgct aactctcgct tttatctgct gcgcactctgc gtttggtgta 60  
aatggccaaa ctactagcgt catgaagaag gtcggattcg gcgccggaag caagccgatg 120  
gtgcaggcaa tcgatgttca aggcaaccgt cttggctcca acgctcccc acgcctcatc 180  
tgcgacagtc gagttctgga gaggtacatc ttagaggcca aggaggcaga aaatgtcacg 240

27-12-2011 ALG-B-0002 PCT -SEQ

atgggttgtg cagaaggtcc cagactgagt gaaaatatta cagtcccaga taccaaagtc	300
aacttctatg cttggaaaag aatggaggtg gaagaacagg ccatagaagt ttggcaaggc	360
ctgtccctgc tctcagaagc catcctgcag gcccaggccc tgctagccaa ttctctccag	420
ccaccagaga cccttcagct tcatatagac aaagccatca gtggtctacg tagcctcact	480
tactgcttc ggggtactggg agctcagaag gaattgatgt cgcctccaga taccacccca	540
cctgtccac tccgaacact cacagtggat actttctgca agctcttccg ggtctacgcc	600
aacttctcc gggggaaact gaagctgtac acgggagagg tctgcaggag aggggacagg	660
ctggaagtcc tgttccaggg gcccatgggtg agcaagggcg aggagctgtt caccgggggtg	720
gtgcccattcc tgggtcgagct ggacggcgac gtaaaccggcc acaagttcag cgtgtccggc	780
gagggcgagg gcgatgccac ctacggcaag ctgaccctga agttcatctg caccaccggc	840
aagctgcccg tgccctggcc caccctcgtg accaccttga cctacggcgt gcagtgttc	900
gcccgtacc ccgaccacat gaagcagcac gacttcttca agtccgccat gccggaaggc	960
tacgtccagg agcgcaccat cttcttcaag gacgacggca actacaagac ccgcgccgag	1020
gtgaagtccg agggcgacac cctggtgaac cgcatcgagc tgaagggcat cgacttcaag	1080
gaggacggca acatcctggg gcacaagctg gagtacaact acaacagcca caaggtctat	1140
atcaccgccg acaagcagaa gaacggcatc aagggtgaact tcaagacctg ccacaacatc	1200
gaggacggca gcgtgcagct cgccgaccac taccagcaga acaccccat cggcgacggc	1260
cccgtgctgc tgcccagaaa ccactacctg agcaccctg cgccttgag caaagacccc	1320
aacgagaagc gcgatcacat ggtcctgctg gagtctgtga ccgccgccgg gatcactctc	1380
ggcatggacg agctgtacaa gtaa	1404

<210> 27  
 <211> 1422  
 <212> DNA  
 <213> artificial sequence

<220>  
 <223> CDS coding for a fusion protein EPO-GFP-HisTag

<400> 27	
atgaaggctc ctaccacgct aactctcgct tttatctgct gcgcatctgc gtttgggtta	60
aatggccaaa ctactagcgt catgaagaag gtcggattcg gcgccggaag caagccgatg	120
gtgcaggcaa tcgatgttca aggcaaccgt cttggctcca acgctcccc acgcctcatc	180
tgcgacagtc gagttctgga gaggtacatc ttagaggcca aggaggcaga aaatgtcacg	240
atgggttgtg cagaaggtcc cagactgagt gaaaatatta cagtcccaga taccaaagtc	300
aacttctatg cttggaaaag aatggaggtg gaagaacagg ccatagaagt ttggcaaggc	360
ctgtccctgc tctcagaagc catcctgcag gcccaggccc tgctagccaa ttctctccag	420
ccaccagaga cccttcagct tcatatagac aaagccatca gtggtctacg tagcctcact	480
tactgcttc ggggtactggg agctcagaag gaattgatgt cgcctccaga taccacccca	540

27-12-2011 ALG-B-0002 PCT -SEQ

cctgctccac tccgaacact cacagtggat actttctgca agctcttccg ggtctacgcc	600
aacttcctcc gggggaaact gaagctgtac acgggagagg tctgcaggag aggggacagg	660
ctggaagtcc tggtccaggg gcccatggtg agcaagggcg aggagctgtt caccgggggtg	720
gtgcccattc tggtcgagct ggacggcgac gtaaaccggcc acaagttcag cgtgtccggc	780
gagggcgagg gcgatgccac ctacggcaag ctgaccctga agttcatctg caccaccggc	840
aagctgcccg tgccctggcc caccctcgtg accaccttga cctacggcgt gcagtgttcc	900
gcccgtacc cggaccacat gaagcagcac gacttcttca agtccgccat gcccgaaggc	960
tacgtccagg agcgcaccat cttcttcaag gacgacggca actacaagac ccgcgccgag	1020
gtgaagtccg agggcgacac cctggtgaac cgcacgcagc tgaagggcat cgacttcaag	1080
gaggacggca acatcctggg gcacaagctg gagtacaact acaacagcca caaggtctat	1140
atcaccgccg acaagcagaa gaacggcatc aagggtgaact tcaagacccg ccacaacatc	1200
gaggacggca gcgatgcagc cgccgaccac taccagcaga acacccccat cggcgacggc	1260
cccgtgctgc tgcccagaca cactacctg agcaccacgt ccgccctgag caaagacccc	1320
aacgagaagc gcgatcacat ggtcctgctg gaggctcgtga ccgccgccgg gatcactctc	1380
ggcatggacg agctgtacaa gcaccacat caccaccatt aa	1422

<210> 28  
 <211> 20  
 <212> DNA  
 <213> artificial sequence

<220>  
 <223> PCR primer

<400> 28	
gtctatatga agctgaaggg	20

<210> 29  
 <211> 19  
 <212> DNA  
 <213> artificial sequence

<220>  
 <223> PCR primer

<400> 29	
gtgagcaagg gcgaggagc	19

<210> 30  
 <211> 1659  
 <212> DNA  
 <213> artificial sequence

<220>  
 <223> CDS coding for a chimeric protein containing the bipartite topogenic signal sequence fused in-frame with the human GBA

<400> 30	
atgagatcct tttgcatcgc agcccttttt gctgtggcat ctgccttcac cacacagcca	60

27-12-2011 ALG-B-0002 PCT -SEQ

acttccttca ctgtgaagac tgcgaatgtg ggcgaacggg cgagtggggg tttccctgag	120
cagagctctg ctcatcgac gcgtaaagca acgattgtca tggatgcccg cccctgcatc	180
cctaaaagct tcggctacag ctcgggtggtg tgtgtctgca atgccacata ctgtgactcc	240
tttgaccccc cgacctttcc tgcccttggt acccttcagcc gctatgagag tacacgcagt	300
gggcgacgga tggagctgag tatggggccc atccaggcta atcacacggg cacaggcctg	360
ctactgacc tgcagccaga acagaagttc cagaaagtga agggatttgg aggggccatg	420
acagatgctg ctgctctcaa catccttgcc ctgtcacccc ctgccccaaa tttgctactt	480
aaatcgtact tctctgaaga aggaatcgga tataacatca tccgggtacc catggccagc	540
tgtgacttct ccatccgcac ctacacctat gcagacaccc ctgatgattt ccagttgcac	600
aacttcagcc tcccagagga agataccaag ctcaagatac ccctgattca ccgagccctg	660
cagttggccc agcgtcccgt ttcactcctt gccagcccct ggacatcacc cacttggctc	720
aagaccaatg gagcggtgaa tgggaagggg tcaactcaagg gacagcccgg agacatctac	780
caccagacct gggccagata ctttgtgaag ttcttgatg cctatgctga gcacaagtta	840
cagttctggg cagtgcagc tgaaaatgag ccttctgctg ggctgttgag tggatacccc	900
ttccagtgcc tgggcttcac ccctgaacat cagcgagact tcattgcccg tgacctaggt	960
cctaccctcg ccaacagtac tcaccacaat gtccgcctac tcatgctgga tgaccaacgc	1020
ttgctgctgc cccactgggc aaagggtgta ctgacagacc cagaagcagc taaatatgtt	1080
catggcattg ctgtacattg gtacctggac tttctggctc cagccaaagc caccctaggg	1140
gagacacacc gcctgttccc caacaccatg ctctttgcct cagaggcctg tgtgggctcc	1200
aagttctggg agcagagtgt gcggctaggc tcctgggatc gagggatgca gtacagccac	1260
agcatcatca cgaacctcct gtacctgtg gtcggctgga ccgactggaa ccttgccctg	1320
aacccgaag gaggaccaa ttgggtgctg aactttgtcg acagtcccat cattgtagac	1380
atcaccaagg acacgtttta caaacagccc atgttctacc accttggcca cttcagcaag	1440
ttcattcctg agggctccca gagagtggg ctggttgcca gtcagaagaa cgacctggac	1500
gcagtggcac tgatgcatcc cgatggctct gctgttgctg tcgtgctaaa ccgctcctct	1560
aaggatgtgc ctcttaccat caaggatcct gctgtgggct tcctggagac aatctcacct	1620
ggctactcca ttcacaccta cctgtggcat cgccagtga	1659

<210> 31  
 <211> 1677  
 <212> DNA  
 <213> artificial sequence

<220>  
 <223> CDS coding for a chimeric GBA-HisTag protein

<400> 31	
atgagatcct tttgcatcgc agcccttttt gctgtggcat ctgccttcac cacacagcca	60
acttccttca ctgtgaagac tgcgaatgtg ggcgaacggg cgagtggggg tttccctgag	120

27-12-2011 ALG-B-0002 PCT -SEQ

cagagctctg ctcacgcac gcgtaaagca acgattgtca tggatgcccg cccctgcatc	180
cctaaaagct tcggctacag ctcggtggtg tgtgtctgca atgccacata ctgtgactcc	240
tttgaccccc cgacctttcc tgcccttggg accttcagcc gctatgagag tacacgcagt	300
gggcgacgga tggagctgag tatggggccc atccaggcta atcacacggg cacaggcctg	360
ctactgaccc tgcagccaga acagaagttc cagaaagtga agggatttgg aggggcatg	420
acagatgctg ctgctctcaa catccttgcc ctgtcacccc ctgccccaaa tttgctactt	480
aaatcgtact tctctgaaga aggaatcgga tataacatca tccgggtacc catggccagc	540
tgtgacttct ccatccgcac ctacacctat gcagacaccc ctgatgattt ccagttgcac	600
aacttcagcc tcccagagga agataccaag ctcaagatac ccctgattca ccgagccctg	660
cagttggccc agcgtcccgt ttcactcctt gccagcccct ggacatcacc cacttggctc	720
aagaccaatg gagcggtgaa tgggaagggg tcaactcaagg gacagcccgg agacatctac	780
caccagacct gggccagata ctttgtgaag ttcctggatg cctatgctga gcacaagtta	840
cagttctggg cagtgcagc tgaaaatgag ccttctgctg ggctgttgag tggatacccc	900
ttccagtgcc tgggcttcac ccctgaacat cagcgagact tcattgcccg tgacctaggt	960
cctaccctcg ccaacagtac tcaccacaat gtccgcctac tcatgctgga tgaccaacgc	1020
ttgctgctgc cccactgggc aaagggtgga ctgacagacc cagaagcagc taaatatgtt	1080
catggcattg ctgtacattg gtacctggac tttctggctc cagccaaagc caccctaggg	1140
gagacacacc gcctgttccc caacaccatg ctctttgcct cagaggcctg tgtgggctcc	1200
aagttctggg agcagagtgt gcggctaggc tcctgggatc gagggatgca gtacagccac	1260
agcatcatca cgaacctcct gtaccatgtg gtcggctgga ccgactggaa ccttgccctg	1320
aaccccgaa gaggacccaa ttgggtgctg aactttgtcg acagtcccat cattgtagac	1380
atcaccaagg acacgtttta caaacagccc atgttctacc accttgcca cttcagcaag	1440
ttcattcctg agggctccca gagagtgggg ctggttgcca gtcagaagaa cgacctggac	1500
gcagtggcac tgatgcatcc cgatggctct gctgttggtg tcgtgctaaa ccgctcctct	1560
aaggatgtgc ctcttaccat caaggatcct gctgtgggct tcctggagac aatctcacct	1620
ggctactcca ttcacaccta cctgtggcat cgccagcacc accatcacca ccattga	1677

<210> 32  
 <211> 19  
 <212> DNA  
 <213> artificial sequence

<220>  
 <223> PCR primer

<400> 32  
 ataccaagct caagatacc

19

<210> 33  
 <211> 20  
 <212> DNA



&lt;213&gt; artificial sequence

&lt;220&gt;

&lt;223&gt; PCR primer

&lt;400&gt; 33

aactgtaact tgtgctcagc

20

&lt;210&gt; 34

&lt;211&gt; 1605

&lt;212&gt; DNA

&lt;213&gt; artificial sequence

&lt;220&gt;

&lt;223&gt; CDS coding for a chimeric protein containing the bipartite topogenic signal sequence fused in-frame with gp120 coding sequence containing a stop codon

&lt;400&gt; 34

atgagatcct tttgcatcgc agcccttttt gctgtggcat ctgccttcac cacacagcca	60
acttccttca ctgtgaagac tgcgaatgtg ggcgaacggg cgagtgggggt tttccctgag	120
cagagctctg ctcatcgcac gcgtaaagca acgattgtca tggataaatt gtgggtcaca	180
gtctattatg ggggtacctgt gtggaaggaa gcaaccacca ctctattttg tgcacagat	240
gctaaagcat atgatacaga ggtacataat gtttgggcca cacatgcctg tgtaccaca	300
gacccaacc cacaagaagt agtattggta aatgtgacag aaaattttta catgtggaaa	360
aatgacatgg tagaacagat gcatgaggat ataatcagtt tatgggatca aagcctaaag	420
ccatgtgtaa aattaacccc actctgtgtt agtttaaagt gcactgattt gaagaatgat	480
actaatacca atagtagtag cgggagaatg ataatggaga aaggagagat aaaaaactgc	540
tctttcaata tcagcacaag cataagaggt aagggtgcaga aagaatatgc atttttttat	600
aaacttgata taataccaat agataatgat actaccagct ataagttgac aagttgtaac	660
acctcagtca ttacacaggc ctgtccaaag gtatcctttg agccaattcc catacattat	720
tgtgccccgg ctggttttgc gattctaaaa tgtaataata agacgttcaa tggaacagga	780
ccatgtacaa atgtcagcac agtacaatgt acacatggaa ttaggccagt agtatcaact	840
caactgctgt taaatggcag tctagcagaa gaagaggtag taattagatc tgtcaatttc	900
acggacaatg ctaaaacat aatagtagag ctgaacacat ctgtagaaat taattgtaca	960
agaccaaca acaatacaag aaaaagaatc cgtatccaga gaggaccagg gagagcattt	1020
gttacaatag gaaaaatagg aaatatgaga caagcacatt gtaacattag tagagcaaaa	1080
tggaataaca ctttaaaaca gatagctagc aaattaagag aacaatttgg aaataataaa	1140
acaataatct ttaagcaatc ctgaggagg gaccagaaa ttgtaacgca cagttttaat	1200
tgtggagggg aatttttcta ctgtaattca acacaactgt ttaatagtac ttggtttaat	1260
agtacttgga gtactgaagg gtcaataaac actgaaggaa gtgacacaat caccctccca	1320
tgcagaataa aacaaattat aaacatgtgg cagaaagtag gaaaagcaat gtatgccctt	1380
cccatcagtg gacaaattag atgttcatca aatattacag ggctgctatt aacaagagat	1440

27-12-2011 ALG-B-0002 PCT -SEQ

ggtggtaata gcaacaatga gtccgagatc ttcagacctg gaggaggaga tatgaggggac	1500
aattggagaa gtgaattata taaatataaa gtagtaaaaa ttgaaccatt aggagtagca	1560
cccaccaagg caaagagaag agtgggtgcag agagaaaaaa gatga	1605

<210> 35  
 <211> 1623  
 <212> DNA  
 <213> artificial sequence

<220>  
 <223> CDS coding for a chimeric gp120-HisTag protein

<400> 35 atgagatcct tttgcatcgc agcccttttt gctgtggcat ctgccttcac cacacagcca	60
acttccttca ctgtgaagac tgcgaatgtg ggcgaacggg cgagtggggt tttccctgag	120
cagagctctg ctcatcgcac gcgtaaagca acgattgtca tggataaatt gtgggtcaca	180
gtctattatg ggggtacctgt gtggaaggaa gcaaccacca ctctattttg tgcacagat	240
gctaaagcat atgatacaga ggtacataat gtttgggccca cacatgcctg tgtaccaca	300
gacccaacc cacaagaagt agtattggta aatgtgacag aaaattttta catgtggaaa	360
aatgacatgg tagaacagat gcatgaggat ataatcagtt tatgggatca aagcctaaag	420
ccatgtgtaa aattaacccc actctgtgtt agtttaaagt gcactgattt gaagaatgat	480
actaatacca atagtagtag cgggagaatg ataatggaga aaggagagat aaaaaactgc	540
tctttcaata tcagcacaag cataagaggt aagggtgcaga aagaatatgc atttttttat	600
aaacttgata taataccaat agataatgat actaccagct ataagttgac aagttgtaac	660
acctcagtca ttacacaggc ctgtccaaag gtatcctttg agccaattcc catacattat	720
tgtgccccgg ctggttttgc gattctaaaa tgtaataata agacgttcaa tggaacagga	780
ccatgtacaa atgtcagcac agtacaatgt acacatggaa ttaggccagt agtatcaact	840
caactgctgt taaatggcag tctagcagaa gaagaggtag taattagatc tgtcaatttc	900
acggacaatg ctaaaacat aatagtacag ctgaacacat ctgtagaaat taattgtaca	960
agaccaaca acaatacaag aaaaagaatc cgtatccaga gaggaccagg gagagcattt	1020
gttacaatag gaaaaatagg aaatatgaga caagcacatt gtaacattag tagagcaaaa	1080
tggaataaca ctttaaaaca gatagctagc aaattaagag aacaatttgg aaataataaa	1140
acaataatct ttaagcaatc ctcaggaggg gaccagaaa ttgtaacgca cagttttaat	1200
tgtggagggg aatttttcta ctgtaattca acacaactgt ttaatagtac ttggtttaat	1260
agtacttgga gtactgaagg gtcaaataac actgaaggaa gtgacacaat caccctccca	1320
tgcagaataa aacaaattat aaacatgtgg cagaaagtag gaaaagcaat gtatgccct	1380
cccatcagtg gacaaattag atgttcatca aatattacag ggctgctatt aacaagagat	1440
ggtggtaata gcaacaatga gtccgagatc ttcagacctg gaggaggaga tatgaggggac	1500
aattggagaa gtgaattata taaatataaa gtagtaaaaa ttgaaccatt aggagtagca	1560

cccaccaagg caaagagaag agtggtgcag agagaaaaaa gacaccacca tcaccaccat 1620  
tga 1623

<210> 36  
<211> 21  
<212> DNA  
<213> artificial sequence

<220>  
<223> PCR primer

<400> 36  
cacctcagtc attacacagg c 21

<210> 37  
<211> 19  
<212> DNA  
<213> artificial sequence

<220>  
<223> PCR primer

<400> 37  
cctcctgagg attgcttaa 19

<210> 38  
<211> 142  
<212> DNA  
<213> Phaeodactylum tricornutum

<400> 38  
gggctgcagg acgcaatgga ggattatcac cgcaaaaatg aacttcgaaa aaaacttttcg 60  
agcgaccatg gaaaaggagg atcagattca gattacaaca gtggattgct ctggtagcaa 120  
atatcttctg ctagattggc tc 142

<210> 39  
<211> 245  
<212> DNA  
<213> Phaeodactylum tricornutum

<400> 39  
acataccttc agcgtcgtct tcactgtcac agtcaactga cagtaatcgt tgatccggag 60  
agattcaaaa ttcaatctgt ttggacctgg ataagacaca agagcgacat cctgacatga 120  
acgccgtaaa cagcaaattcc tggttgaaca cgtatccttt tgggggcctc cgctacgacg 180  
ctcgtccag ctggggcctt cttactatac acagcgcgca tatttcacgg ttgccagatg 240  
tcaag 245

<210> 40  
<211> 835  
<212> DNA  
<213> cauliflower mosaic virus

<400> 40  
agattagcct tttcaatttc agaaagaatg ctaaccaca gatggtaga gaggcttacg 60  
cagcaggtct catcaagacg atctaccgca gcaataatct ccaggaaatc aaataccttc 120

27-12-2011 ALG-B-0002 PCT -SEQ

ccaagaaggt taaagatgca gtcaaaagat tcaggactaa ctgcatcaag aacacagaga	180
aagatatatt tctcaagatc agaagtacta ttccagtatg gacgattcaa ggcttgcttc	240
acaaaccaag gcaagtaata gagattggag tctctaaaaa ggtagttccc actgaatcaa	300
aggccatgga gtcaaagatt caaatagagg acctaacaga actcgccgta aagactggcg	360
aacagttcat acagagtctc ttacgactca atgacaagaa gaaaatcttc gtcaacatgg	420
tggagcacga cacacttgtc tactccaaaa atatcaaaga tacagtctca gaagaccaa	480
gggcaattga gacttttcaa caaagggtaa tatccggaaa cctcctcgga ttccattgcc	540
cagctatctg tcacttttatt gtgaagatag tggaaaagga aggtggctcc tacaaatgcc	600
atcattgcga taaaggaaag gccatcggtg aagatgcctc tgccgacagt ggccccaaag	660
atggaccccc acccacgagg agcatcggtg aaaaagaaga cgtccaacc acgtcttcaa	720
agcaagtgga ttgatgtgat atctccactg acgtaaggga tgacgcacaa tcccactatc	780
cttcgcaaga cccttcctct atataaggaa gttcatttca tttggagaga acacg	835

<210> 41  
 <211> 240  
 <212> DNA  
 <213> Agrobacterium tumefaciens

<400> 41	
acgattgaag gagccactca gccgcgggtt tctggagttt aatgagctaa gcacatacgt	60
cagaaaccat tattgcgcgt tcaaaagtcg cctaagggtca ctatcagcta gcaaattttt	120
cttgtcaaaa atgctccact gacgttccat aaattcccct cggatatcaa ttagagtctc	180
atattcactc tcaatccaaa taatctgcaa tggcaattac cttattcgca acttctttac	240

<210> 42  
 <211> 242  
 <212> DNA  
 <213> Phaeodactylum tricornutum

<400> 42	
accttcctta aaaatttaat ttctattagt tgcagtcact ccgctttggt ttcacagtca	60
ggaataacac tagctcgtct tcaccatgga tgccaatctc gcctattcat ggtgtataaa	120
agttcaacat ccaaagctag aacttttgga aagagaaaga atatccgaat agggcacggc	180
gtgccgtatt gttggagtgg actagcagaa agtgaggaag gcacaggatg agttttctcg	240
ag	242

<210> 43  
 <211> 253  
 <212> DNA  
 <213> Agrobacterium tumefaciens

<400> 43	
gatcggtcaa acatttgga ataaagtctt ttaagattga atcctgttgc cggcttgcg	60
atgattatca tataatttct gttgaattac gtttaagcatg taataattaa catgtaatgc	120

27-12-2011 ALG-B-0002 PCT -SEQ

atgacgttat ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac 180  
gcgatagaaa acaaaatata gcgcgcaaac taggataaat tatcgcgcgcg ggtgtcatct 240  
atgttactag atc 253

<210> 44  
<211> 1353  
<212> DNA  
<213> artificial sequence

<220>  
<223> BiP-eGFP-EPO-DDEL construction

<400> 44  
atgatattca tgagaattgc cgtagcagca ctggccttgc tggctgctcc ctccattcgt 60  
gccgaagagg ccggtgaaga ggccaagatg ggtaccgtga tggtagagcaa gggcgaggag 120  
ctgttcaccg ggggtggtgcc catcctgggtc gagctggacg gcgacgtaaa cggccacaag 180  
ttcagcgtgt ccggcgaggg cgagggcgat gccacctacg gcaagctgac cctgaagtcc 240  
atctgcacca ccggcaagct gcccgtgccc tggcccaccc tcgtgaccac cttgacctac 300  
ggcgtgcagt gcttcgcccc ctacccccgac cacatgaagc agcacgactt cttcaagtcc 360  
gccatgcccc aaggctacgt ccaggagcgc accatcttct tcaaggacga cggcaactac 420  
aagacccgcg ccgaggtgaa gttcgagggc gacaccctgg tgaaccgcat cgagctgaag 480  
ggcatcgact tcaaggagga cggcaacatc ctggggcaca agctggagta caactacaac 540  
agccacaagg tctatatcac cgccgacaag cagaagaacg gcatcaaggt gaacttcaag 600  
acccgccaca acatcgagga cggcagcgtg cagctcgccg accactacca gcagaacacc 660  
cccctcgccg acggccccgt gctgctgccc gacaaccact acctgagcac ccagtccgcc 720  
ctgagcaaag accccaacga gaagcgcgat cacatgggtc tgctggagtt cgtgaccgcc 780  
gccgggatca ctctcgcat ggacgagctg tacaagctgg aagttctgtt ccagggggccc 840  
gctccccac gcctcatctg cgacagtcga gttctggaga ggtacatctt agaggccaag 900  
gaggcagaaa atgtcacgat gggttgtgca gaaggtccca gactgagtga aaatattaca 960  
gtcccagata ccaaagtcaa cttctatgct tggaaaagaa tggaggtgga agaacaggcc 1020  
atagaagttt ggcaaggcct gtccctgctc tcagaagcca tcctgcaggc ccaggccctg 1080  
ctagccaatt cctcccagcc accagagacc cttcagcttc atatagacaa agccatcagt 1140  
gggtctacgta gcctcacttc actgcttcgg gtactgggag ctgagaagga attgatgtcg 1200  
cctccagata ccacccacc tgctccactc cgaacactca cagtggatac tttctgcaag 1260  
ctcttccggg tctacgcaa cttcctccgg gggaaactga agctgtacac gggagaggtc 1320  
tgcaggagag gggacaggga cgatgagttg tga 1353

<210> 45  
<211> 2346  
<212> DNA  
<213> artificial sequence

&lt;220&gt;

&lt;223&gt; BTS-GP120-GFP construction

&lt;400&gt; 45

atgagatcct tttgcatcgc agcccttttt gctgtggcat ctgccttcac cacacagcca	60
acttccttca ctgtgaagac tgcgaatgtg ggcgaacggg cgagtgggggt tttccctgag	120
cagagctctg ctcatcgac gcgtaaagca acgattgtca tggataaatt gtgggtcaca	180
gtctattatg ggggtacctgt gtggaaggaa gcaaccacca ctctattttg tgcatacagat	240
gctaaagcat atgatacaga ggtacataat gtttgggcca cacatgcctg tgtaccacaca	300
gacccaacc cacaagaagt agtattggta aatgtgacag aaaattttta catgtggaaa	360
aatgacatgg tagaacagat gcatgaggat ataatcagtt tatgggatca aagcctaaag	420
ccatgtgtaa aattaacccc actctgtgtt agtttaaagt gactgattt gaagaatgat	480
actaatacca atagtagtag cgggagaatg ataatggaga aaggagagat aaaaaactgc	540
tctttcaata tcagcacaag cataagaggt aagggtgcaga aagaatatgc atttttttat	600
aaacttgata taataccaat agataatgat actaccagct ataagttgac aagttgtaac	660
acctcagtca ttacacaggc ctgtccaaag gtatcctttg agccaattcc catacattat	720
tgtgccccgg ctggttttgc gattctaaaa tgtaataata agacgttcaa tggaacagga	780
ccatgtacaa atgtcagcac agtacaatgt acacatggaa ttaggccagt agtatcaact	840
caactgctgt taaatggcag tctagcagaa gaagaggtag taattagatc tgtcaatttc	900
acggacaatg ctaaaacccat aatagtacag ctgaacacat ctgtagaaat taattgtaca	960
agaccaaca acaatacaag aaaaagaatc cgtatccaga gaggaccagg gagagcattt	1020
gttacaatag gaaaaatagg aaatatgaga caagcacatt gtaacattag tagagcaaaa	1080
tggaataaca ctttaaaaca gatagctagc aaattaagag aacaatttgg aaataataaa	1140
acaataatct ttaagcaatc ctgaggaggg gaccagaaa ttgtaacgca cagttttaat	1200
tgtggagggg aatttttcta ctgtaattca acacaactgt ttaatagtac ttggtttaat	1260
agtacttggg gtactgaagg gtcaaataac actgaaggaa gtgacacaat caccctccca	1320
tgcagaataa aacaaattat aaacatgtgg cagaaagtag gaaaagcaat gtatgcccct	1380
cccatcagtg gacaaattag atgttcatca aatattacag ggctgctatt aacaagagat	1440
gggtggtata gcaacaatga gtccgagatc ttcagacctg gaggaggaga tatgaggggac	1500
aattggagaa gtgaattata taaatataaa gtagtaaaaa ttgaaccatt aggagtagca	1560
cccaccaagg caaagagaag agtgggtgcag agagaaaaaa gactggaagt tctgttccag	1620
gggcccattg tgagcaaggg cgaggagctg ttcaccgggg tgggtgcccatt cctgggtcgag	1680
ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg gcgagggcga gggcgatgcc	1740
acctacggca agctgaccct gaagttcatc tgcaccaccg gcaagctgcc cgtgccctgg	1800
cccaccctcg tgaccacctt gacctacggc gtgcagtgtc tcgcccgccta ccccgaccac	1860
atgaagcagc acgacttctt caagtccgcc atgcccgaag gctacgtcca ggagcgcacc	1920

27-12-2011 ALG-B-0002 PCT -SEQ

atctttcttca aggacgacgg caactacaag acccgcgccg aggtgaagtt cgagggcgac	1980
accctggtga accgcatcga gctgaagggc atcgacttca aggaggacgg caacatcctg	2040
gggcacaagc tggagtacaa ctacaacagc cacaaggtct atatcaccgc cgacaagcag	2100
aagaacggca tcaaggtgaa cttcaagacc cgccacaaca tcgaggacgg cagcgtgcag	2160
ctcgccgacc actaccagca gaacaccccc atcggcgacg gccccgtgct gctgcccgcac	2220
aaccactacc tgagcaccca gtccgcccctg agcaaagacc ccaacgagaa gcgcgatcac	2280
atggtcctgc tggagtctgt gaccgccgcc gggatcactc tcggcatgga cgagctgtac	2340
aagtaa	2346