

SEQUENCE LISTING

<110> Vical Incorporated

<120> Severe Acute Respiratory Syndrome DNA Vaccine Compositions and Method of Use

<130> 1530.067PC03

<150> US 60/482,505

<151> 2003-06-26

<150> US 60/470,820

<151> 2003-05-16

<160> 69

<170> PatentIn version 3.2

<210> 1

<211> 3588

<212> DNA

<213> SARS-CoV Urbani strain

<400> 1
atgtttattt tcttattatt tcttactctc actagtggta gtgacctga ccggtgcacc 60
acttttgatg atgttcaagc tcctaattac actcaacata cttcatctat gaggggggtt 120
tactatcctg atgaaatttt tagatcagac actcttttatt taactcagga tttatttctt 180
ccatttttatt ctaatgttac agggtttcat actattaatc atacgtttgg caaccctgtc 240
atacctttta aggatgggtat ttattttgct gccacagaga aatcaaagt tgtccgtggt 300
tgggtttttg gttctaccat gaacaacaag tcacagtcgg tgattattat taacaattct 360
actaatgttg ttatacgagc atgtaacttt gaattgtgtg acaacccttt ctttgctggt 420
tctaaaccca tgggtacaca gacacatact atgatattcg ataatgcatt taattgcact 480
ttcgagtaca tatctgatgc cttttcgctt gatgtttcag aaaagtcagg taatttttaa 540
cacttacgag agtttgtgtt taaaaataaa gatgggtttc tctatgttta taagggctat 600
caacctatag atgtagttcg tgatctacct tctgggttta acactttgaa acctattttt 660
aagttgcctc ttggtattaa cattacaaat tttagagcca ttcttacagc cttttcacct 720
gctcaagaca tttggggcac gtcagctgca gcctattttg ttggctattt aaagccaact 780
acatttatgc tcaagtatga tgaaaatggg acaatcacag atgctgttga ttgttctcaa 840
aatccacttg ctgaactcaa atgctctggt aagagctttg agattgacaa aggaatttac 900
cagacctcta atttcagggg tggtccctca ggagatgttg tgagattccc taatattaca 960
aacttgtgtc cttttggaga ggtttttaat gctactaaat tcccttctgt ctatgcatgg 1020
gagagaaaaa aaattttctaa ttgtgttgct gattactctg tgctctacaa ctcaacattt 1080

ttttcaacct	ttaagtgcta	tggcgtttct	gccactaagt	tgaatgatct	ttgcttctcc	1140
aatgtctatg	cagattcttt	tgtagtcaag	ggagatgatg	taagacaaat	agcgccagga	1200
caaactgggtg	ttattgctga	ttataattat	aaattgccag	atgatttcat	gggttggtgc	1260
cttgcttgga	atactaggaa	cattgatgct	acttcaactg	gtaattataa	ttataaatat	1320
aggatatctta	gacatggcaa	gcttaggccc	tttgagagag	acatatctaa	tgtgcctttc	1380
tcccctgatg	gcaaaccttg	cacccacct	gctcttaatt	gttattggcc	attaaatgat	1440
tatgggtttt	acaccactac	tggcattggc	taccaacctt	acagagttgt	agtactttct	1500
tttgaacttt	taaatgcacc	ggccacgggt	tgtggaccaa	aattatccac	tgaccttatt	1560
aagaaccagt	gtgtcaat	taattttaat	ggactcactg	gtactgggtg	gttaactcct	1620
tcttcaaaga	gatttcaacc	atltcaacaa	tttggccgtg	atgtttctga	tttactgat	1680
tccgttcgag	atcctaaaac	atctgaaata	ttagacattt	caccttgctc	ttttgggggt	1740
gtaagtgtaa	ttacacctgg	aacaaatgct	tcactctgaag	ttgctgttct	atatcaagat	1800
gttaactgca	ctgatgtttc	tacagcaatt	catgcagatc	aactcacacc	agcttggcgc	1860
atatattcta	ctggaaacaa	tgtattccag	actcaagcag	gctgtcttat	aggagctgag	1920
catgtcgaca	cttcttatga	gtgcgacatt	cctattggag	ctggcatttg	tgctagttac	1980
catacagttt	ctttattacg	tagtactagc	caaaaatcta	ttgtggctta	tactatgtct	2040
ttaggtgctg	atagttcaat	tgttactct	aataacacca	ttgctatacc	tactaacttt	2100
tcaattagca	ttactacaga	agtaatgcct	gtttctatgg	ctaaaacctc	cgtagattgt	2160
aatatgtaca	tctgcggaga	ttctactgaa	tgtgctaatt	tgcttctcca	atatggtagc	2220
ttttgcacac	aactaaatcg	tgcaactctca	ggatttgctg	ctgaacagga	tcgcaacaca	2280
cgtgaagtgt	tcgctcaagt	caaacaaatg	tacaaaaccc	caactttgaa	atattttggt	2340
ggttttaatt	tttcacaaat	attacctgac	cctctaaagc	caactaagag	gtcttttatt	2400
gaggacttgc	tctttaataa	ggtgacactc	gctgatgctg	gcttcatgaa	gcaatatggc	2460
gaatgcctag	gtgatattaa	tgctagagat	ctcatttgctg	cgcagaagtt	caatggactt	2520
acagtgttgc	cacctctgct	cactgatgat	atgattgctg	cctacactgc	tgctctagtt	2580
agtgggtactg	ccactgctgg	atggacattt	gggtgctggcg	ctgctcttca	aatacctttt	2640
gctatgcaaa	tggcatatag	gttcaatggc	attggagtta	ccaaaaatgt	tctctatgag	2700
aaccaaaaaac	aaatcgccaa	ccaatttaac	aaggcgatta	gtcaaattca	agaatcactt	2760
acaacaacat	caactgcatt	gggcaagctg	caagacgttg	ttaaccagaa	tgctcaagca	2820
ttaaacacac	ttgttaaaca	acttagctct	aattttgggtg	caatttcaag	tgtgctaaat	2880
gatatccttt	cgcgacttga	taaagtcgag	gcggagggtac	aaattgacag	gttaattaca	2940

```

ggcagacttc aaagccttca aacctatgta acacaacaac taatcagggc tgctgaaatc 3000
agggcttctg ctaatcttgc tgctactaaa atgtctgagt gtgttcttgg acaatcaaaa 3060
agagttgact tttgtggaaa gggctaccac cttatgtcct tcccacaagc agccccgcat 3120
ggtgttgtct tcctacatgt cacgtatgtg ccatcccagg agaggaactt caccacagcg 3180
ccagcaattt gtcatgaagg caaagcatac ttccctcgtg aaggtgtttt tgtgtttaat 3240
ggcacttctt gggtttattac acagaggaac ttcttttctc cacaaataat tactacagac 3300
aatacatttg tctcaggaaa ttgtgatgtc gttattggca tcattaacaa cacagtttat 3360
gatcctctgc aacctgagct cgactcattc aaagaagagc tggacaagta cttcaaaaat 3420
catacatcac cagatgttga tcttggcgac atttcaggca ttaacgcttc tgcgtcaac 3480
attcaaaaag aaattgaccg cctcaatgag gtcgctaaaa atttaaatga atcactcatt 3540
gaccttcaag aattgggaaa atatgagcaa tatattaaat ggccttgg 3588

```

<210> 2

<211> 1196

<212> PRT

<213> SARS-CoV Urbani strain

<400> 2

```

Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
1           5           10          15

```

```

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

```

```

His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
35           40          45

```

```

Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
50           55          60

```

```

Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
65           70          75          80

```

```

Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
85           90          95

```

```

Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
100          105         110

```

```

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
115          120         125

```

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

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
 145 150 155 160

Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
 165 170 175

Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
 180 185 190

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

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
 210 215 220

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
 225 230 235 240

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
 245 250 255

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
 260 265 270

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
 275 280 285

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
 290 295 300

Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
 305 310 315 320

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
 325 330 335

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
 340 345 350

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

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
 370 375 380

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
 385 390 395 400

Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
 405 410 415

Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
 420 425 430

Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
 435 440 445

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

Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
 465 470 475 480

Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
 485 490 495

Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
 500 505 510

Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
 515 520 525

Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
 530 535 540

Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
 545 550 555 560

Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
 565 570 575

Ser Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
 580 585 590

Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
 595 600 605

Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
610 615 620

Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
625 630 635 640

His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
645 650 655

Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
660 665 670

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

Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
690 695 700

Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
705 710 715 720

Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
725 730 735

Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
740 745 750

Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys
755 760 765

Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe
770 775 780

Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile
785 790 795 800

Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met
805 810 815

Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile
820 825 830

Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
835 840 845

Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala

850		855		860											
Thr	Ala	Gly	Trp	Thr	Phe	Gly	Ala	Gly	Ala	Ala	Leu	Gln	Ile	Pro	Phe
865					870				875						880
Ala	Met	Gln	Met	Ala	Tyr	Arg	Phe	Asn	Gly	Ile	Gly	Val	Thr	Gln	Asn
			885						890					895	
Val	Leu	Tyr	Glu	Asn	Gln	Lys	Gln	Ile	Ala	Asn	Gln	Phe	Asn	Lys	Ala
			900					905					910		
Ile	Ser	Gln	Ile	Gln	Glu	Ser	Leu	Thr	Thr	Thr	Ser	Thr	Ala	Leu	Gly
		915					920					925			
Lys	Leu	Gln	Asp	Val	Val	Asn	Gln	Asn	Ala	Gln	Ala	Leu	Asn	Thr	Leu
	930					935					940				
Val	Lys	Gln	Leu	Ser	Ser	Asn	Phe	Gly	Ala	Ile	Ser	Ser	Val	Leu	Asn
945				950						955					960
Asp	Ile	Leu	Ser	Arg	Leu	Asp	Lys	Val	Glu	Ala	Glu	Val	Gln	Ile	Asp
			965						970					975	
Arg	Leu	Ile	Thr	Gly	Arg	Leu	Gln	Ser	Leu	Gln	Thr	Tyr	Val	Thr	Gln
		980					985						990		
Gln	Leu	Ile	Arg	Ala	Ala	Glu	Ile	Arg	Ala	Ser	Ala	Asn	Leu	Ala	Ala
	995						1000					1005			
Thr	Lys	Met	Ser	Glu	Cys	Val	Leu	Gly	Gln	Ser	Lys	Arg	Val	Asp	
	1010					1015					1020				
Phe	Cys	Gly	Lys	Gly	Tyr	His	Leu	Met	Ser	Phe	Pro	Gln	Ala	Ala	
	1025					1030					1035				
Pro	His	Gly	Val	Val	Phe	Leu	His	Val	Thr	Tyr	Val	Pro	Ser	Gln	
	1040					1045					1050				
Glu	Arg	Asn	Phe	Thr	Thr	Ala	Pro	Ala	Ile	Cys	His	Glu	Gly	Lys	
	1055					1060					1065				
Ala	Tyr	Phe	Pro	Arg	Glu	Gly	Val	Phe	Val	Phe	Asn	Gly	Thr	Ser	
	1070					1075					1080				
Trp	Phe	Ile	Thr	Gln	Arg	Asn	Phe	Phe	Ser	Pro	Gln	Ile	Ile	Thr	
	1085					1090					1095				

Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly
1100 1105 1110

Ile Ile Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp
1115 1120 1125

Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser
1130 1135 1140

Pro Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val
1145 1150 1155

Val Asn Ile Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala Lys
1160 1165 1170

Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr
1175 1180 1185

Glu Gln Tyr Ile Lys Trp Pro Trp
1190 1195

<210> 3
<211> 2049
<212> DNA
<213> SARS-CoV Urbani strain

<400> 3
atgtttatatt tcttattatt tcttactctc actagtggta gtgacctga ccggtgcacc 60
acttttgatg atgttcaagc tcctaattac actcaacata cttcatctat gagggggggtt 120
tactatcctg atgaaatatt tagatcagac actctttatt taactcagga tttatttctt 180
ccattttatt ctaatgttac agggtttcat actattaatc atacgtttgg caaccctgtc 240
atacctttta aggatgggat ttattttgct gccacagaga aatcaaagt tgtccgtggt 300
tgggtttttg gttctaccat gaacaacaag tcacagtcgg tgattattat taacaattct 360
actaatgttg ttatacgagc atgtaacttt gaattgtgtg acaaccctt ctttgctggt 420
tctaaaccca tgggtacaca gacacatact atgatattcg ataatgcatt taattgcact 480
ttcgagtaca tatctgatgc cttttcgtt gatgtttcag aaaagtcagg taattttaaa 540
cacttacgag agtttgtgtt taaaaataaa gatgggtttc tctatgttta taagggtat 600
caacctatag atgtagttcg tgatctacct tctggtttta acactttgaa acctattttt 660
aagttgcctc ttggtattaa cattacaaat tttagagcca ttcttacagc cttttcacct 720
gctcaagaca tttggggcac gtcagctgca gcctattttg ttggctattt aaagccaact 780

```

acatttatgc tcaagtatga tgaaaatggg acaatcacag atgctgttga ttgttctcaa      840
aatccacttg ctgaactcaa atgctctgtt aagagctttg agattgacaa aggaattttac      900
cagacctcta atttcagggg tgttccctca ggagatgttg tgagattccc taatattaca      960
aacttgtgtc cttttggaga ggtttttaat gctactaaat tcccttctgt ctatgcatgg     1020
gagagaaaaa aaattttctaa ttgtgttgct gattactctg tgctctacaa ctcaacattt     1080
ttttcaacct ttaagtgcta tggcgtttct gccactaagt tgaatgatct ttgcttctcc     1140
aatgtctatg cagattcttt tgtagtcaag ggagatgatg taagacaaat agcgccagga     1200
caaactgggtg ttattgctga ttataattat aaattgccag atgatttcat gggttgtgtc     1260
cttgcttgga atactaggaa cattgatgct acttcaactg gtaattataa ttataaatat     1320
aggtatctta gacatggcaa gcttaggccc tttagagagag acatatctaa tgtgcctttc     1380
tcccctgatg gcaaaccttg caccacacct gctcttaatt gttattggcc attaaatgat     1440
tatggttttt acaccactac tggcattggc taccaacctt acagagttgt agtactttct     1500
tttgaacttt taaatgcacc ggccacgggt tgtggaccaa aattatccac tgaccttatt     1560
aagaaccagt gtgtcaattt taattttaat ggactcactg gtactgggtg gttaactcct     1620
tcttcaaaga gatttcaacc atttcaacaa tttggccgtg atgtttctga tttcactgat     1680
tccgttcgag atcctaaaac atctgaaata ttagacattt caccttgctc ttttgggggt     1740
gtaagtgtaa ttacacctgg aacaaatgct tcatctgaag ttgctgttct atatcaagat     1800
gttaactgca ctgatgtttc tacagcaatt catgcagatc aactcacacc agcttggcgc     1860
atatattcta ctggaaacaa tgtattccag actcaagcag gctgtcttat aggagctgag     1920
catgtcgaca cttcttatga gtgcgacatt cctattggag ctggcatttg tgctagttac     1980
catacagttt ctttattacg tagtactagc caaaaatcta ttgtggctta tactatgtct     2040
ttaggtgct                                     2049

```

```

<210> 4
<211> 683
<212> PRT
<213> SARS-CoV Urbani strain

```

```

<400> 4

```

```

Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
1           5           10          15

```

```

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

```

His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 35 40 45

Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 50 55 60

Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 65 70 75 80

Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 85 90 95

Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
 100 105 110

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
 115 120 125

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

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
 145 150 155 160

Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
 165 170 175

Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
 180 185 190

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

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
 210 215 220

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
 225 230 235 240

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
 245 250 255

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
 260 265 270

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys

275	280	285
Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn 290 295 300		
Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr 305 310 315 320		
Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser 325 330 335		
Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr 340 345 350		
Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly 355 360 365		
Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala 370 375 380		
Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly 385 390 395 400		
Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe 405 410 415		
Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser 420 425 430		
Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu 435 440 445		
Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly 450 455 460		
Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp 465 470 475 480		
Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val 485 490 495		
Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly 500 505 510		
Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn 515 520 525		

Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
 530 535 540

Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
 545 550 555 560

Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
 565 570 575

Ser Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
 580 585 590

Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
 595 600 605

Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
 610 615 620

Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
 625 630 635 640

His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
 645 650 655

Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
 660 665 670

Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala
 675 680

<210> 5
 <211> 1539
 <212> DNA
 <213> SARS-CoV Urbani strain

<400> 5
 gatagttcaa ttgcttactc taataacacc attgctatac ctactaactt ttcaattagc 60
 attactacag aagtaatgcc tgtttctatg gctaaaacct ccgtagattg taatatgtac 120
 atctgcggag atttactga atgtgctaat ttgcttctcc aatatggtag cttttgcaca 180
 caactaaatc gtgcactctc aggtattgct gctgaacagg atcgcaacac acgtgaagtg 240
 ttcgctcaag tcaaacaaat gtacaaaacc ccaactttga aatatttttg tggttttaaat 300
 ttttcacaaa tattacctga ccctctaaag ccaactaaga ggtcttttat tgaggacttg 360
 ctctttaata aggtgacact cgctgatgct ggcttcatga agcaatatgg cgaatgccta 420

```

ggatgatatta atgctagaga tctcatttgt ggcgagaagt tcaatggact tacagtgttg 480
ccacctctgc tcactgatga tatgattgct gcctacactg ctgctctagt tagtggtact 540
gccactgctg gatggacatt tggtgctggc gctgctcttc aaataccttt tgctatgcaa 600
atggcatata ggttcaatgg cattggagtt acccaaatg ttctctatga gaacaaaaaa 660
caaatcgcca accaattta caaggcgatt agtcaaattc aagaatcact tacaacaaca 720
tcaactgcat tgggcaagct gcaagacgtt gttaaccaga atgctcaagc attaaacaca 780
cttggttaaac aacttagctc taatttttgg gcaatttcaa gtgtgctaaa tgatattcctt 840
tcgctgacttg ataaagtcca ggcggaggta caaattgaca ggtaattac aggcagactt 900
caaagccttc aaacctatgt aacacaacaa ctaatcaggg ctgctgaaat cagggtcttct 960
gctaattcttg ctgctactaa aatgtctgag tgtgttcttg gacaatcaaa aagagttgac 1020
ttttgtggaa agggctacca ccttatgtcc ttcccacaag cagccccgca tgggtgtgtc 1080
ttcctacatg tcacgtatgt gccatcccag gagaggaact tcaccacagc gccagcaatt 1140
tgtcatgaag gcaaagcata cttccctcgt gaagggtgtt ttgtgtttta tggcacttct 1200
tggtttatta cacagaggaa cttcttttct ccacaaataa ttactacaga caatacattt 1260
gtctcaggaa atttgtatgt cgttattggc atcattaaca acacagtta tgatcctctg 1320
caacctgagc tcgactcatt caaagaagag ctggacaagt acttcaaaaa tcatacatca 1380
ccagatgttg atcttggcga catttcaggc attaacgctt ctgtcgtcaa cattcaaaaa 1440
gaaattgacc gcctcaatga ggtcgctaaa aatttaaagt aatcactcat tgaccttcaa 1500
gaattgggaa aatatgagca atatattaaa tggccttgg 1539

```

```

<210> 6
<211> 513
<212> PRT
<213> SARS-CoV Urbani strain

```

```

<400> 6

```

```

Asp Ser Ser Ile Ala Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn
1           5           10          15

```

```

Phe Ser Ile Ser Ile Thr Thr Glu Val Met Pro Val Ser Met Ala Lys
20           25           30

```

```

Thr Ser Val Asp Cys Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys
35           40           45

```

```

Ala Asn Leu Leu Leu Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg
50           55           60

```

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

Phe Ala Gln Val Lys Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe
 85 90 95

Gly Gly Phe Asn Phe Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr
 100 105 110

Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala
 115 120 125

Asp Ala Gly Phe Met Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn
 130 135 140

Ala Arg Asp Leu Ile Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu
 145 150 155 160

Pro Pro Leu Leu Thr Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu
 165 170 175

Val Ser Gly Thr Ala Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala
 180 185 190

Leu Gln Ile Pro Phe Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile
 195 200 205

Gly Val Thr Gln Asn Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn
 210 215 220

Gln Phe Asn Lys Ala Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr
 225 230 235 240

Ser Thr Ala Leu Gly Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln
 245 250 255

Ala Leu Asn Thr Leu Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile
 260 265 270

Ser Ser Val Leu Asn Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala
 275 280 285

Glu Val Gln Ile Asp Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln
 290 295 300

Thr Tyr Val Thr Gln Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser
305 310 315 320

Ala Asn Leu Ala Ala Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser
325 330 335

Lys Arg Val Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro
340 345 350

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

Ser Gln Glu Arg Asn Phe Thr Thr Ala Pro Ala Ile Cys His Glu Gly
370 375 380

Lys Ala Tyr Phe Pro Arg Glu Gly Val Phe Val Phe Asn Gly Thr Ser
385 390 395 400

Trp Phe Ile Thr Gln Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr Thr
405 410 415

Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Ile
420 425 430

Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys
435 440 445

Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp
450 455 460

Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile Gln Lys
465 470 475 480

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

Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr Ile Lys Trp Pro
500 505 510

Trp

<210> 7
<211> 3633
<212> DNA
<213> SARS-CoV Urbani strain

<400> 7
atggatgcaa tgaagagagg gctctgctgt gtgctgctgc tgtgtggagc agtcttcggt 60
tcgcccagcg ctagaggatc gggaagtgc cttgaccggg gcaccacttt tgatgatgtt 120
caagctccta attacactca acatacttca tctatgaggg gggtttacta tcctgatgaa 180
atTTTTtagat cagacactct ttatttaact caggatttat ttcttccatt ttattctaatt 240
gttacagggg ttcatactat taatcatagc tttggcaacc ctgtcatacc ttttaaggat 300
ggatatttatt ttgctgccac agagaaatca aatgttgctc gtggttggtt ttttggttct 360
accatgaaca acaagtcaca gtcggtgatt attattaaca attctactaa tgttggttata 420
cgagcatgta actttgaatt gtgtgacaac cctttctttg ctgtttctaa acccatgggt 480
acacagacac atactatgat attcgataat gcatttaatt gcactttcga gtacatatct 540
gatgcctttt cgcttgatgt ttcagaaaag tcaggtaatt ttaaactt acgagagttt 600
gtgttttaaaa ataaagatgg gtttctctat gtttataagg gctatcaacc tatagatgta 660
gttcgtgatc taccttctgg ttttaacact ttgaaacctt tttttaagtt gcctcttggt 720
attaacatta caaatttttag agccattctt acagcctttt cacctgctca agacatttggt 780
ggcacgtcag ctgcagccta ttttggtggc tatttaaagc caactacatt tatgctcaag 840
tatgatgaaa atgggtacaat cacagatgct gttgattggt ctcaaaatcc acttgctgaa 900
ctcaaagct ctgttaagag ctttgagatt gacaaaggaa tttaccagac ctctaatttc 960
agggttggtt cctcaggaga tgttggtgaga ttccctaata ttacaaactt gtgtcctttt 1020
ggagaggttt ttaatgctac taaattccct tctgtctatg catgggagag aaaaaaatt 1080
tctaattgtg ttgctgatta ctctgtgctc tacaactcaa catttttttc aacctttaag 1140
tgctatggcg tttctgccac taagttgaat gatctttgct tctccaatgt ctatgcagat 1200
tcttttgtag tcaagggaga tgatgtaaga caaatagcgc caggacaaac tgggtgttatt 1260
gctgattata attataaatt gccagatgat ttcatgggtt gtgtccttgc ttggaatact 1320
aggaacattg atgctacttc aactggtaat tataattata aatataggta tcttagacat 1380
ggcaagctta ggccctttga gagagacata tctaattgtc ctttctcccc tgatggcaaa 1440
ccttgacccc cacctgctct taattgttat tggccattaa atgattatgg tttttacacc 1500
actactggca ttggctacca accttacaga gttgtagtac tttcttttga acttttaaat 1560
gcaccggcca cggtttggtg accaaaatta tccactgacc ttattaagaa ccagtgtgtc 1620
aatTTTtaatt ttaatggact cactgggtact ggtgtgttaa ctcttcttc aaagagattt 1680
caaccatttc aacaatttgg ccgtgatgtt tctgatttca ctgattccgt tcgagatcct 1740
aaaacatctg aaatattaga catTTTcacct tgctcttttg ggggtgtaag tgtaattaca 1800

cctggaacaa atgcttcac	tgaagttgct gttctatatc	aagatgttaa ctgcactgat	1860
gtttctacag caattcatgc	agatcaactc acaccagctt	ggcgcatata ttctactgga	1920
aacaatgtat tccagactca	agcaggctgt cttataggag	ctgagcatgt cgacacttct	1980
tatgagtgcg acattcctat	tggagctggc atttgtgcta	gttaccatac agtttcttta	2040
ttacgtagta ctagccaaaa	atctattgtg gcttatacta	tgtctttagg tgctgatagt	2100
tcaattgctt actctaataa	caccattgct atacctacta	acttttcaat tagcattact	2160
acagaagtaa tgccgtgttc	tatggctaaa acctccgtag	attgtaatat gtacatctgc	2220
ggagattcta ctgaatgtgc	taatttgctt ctccaatatg	gtagcttttg cacacaacta	2280
aatcgtgcac tctcaggtat	tgctgctgaa caggatcgca	acacacgtga agtggtcgct	2340
caagtcaaac aaatgtacaa	aacccaact ttgaaatatt	ttggtggttt taatttttca	2400
caaataattac ctgaccctct	aaagccaact aagaggtctt	ttattgagga cttgctcttt	2460
aataaggtga cactcgctga	tgctggcttc atgaagcaat	atggcgaatg cctaggtgat	2520
attaatgcta gagatctcat	ttgtgcgcag aagttcaatg	gacttacagt gttgccacct	2580
ctgctcactg atgatatgat	tgctgcctac actgctgctc	tagttagtgg tactgccact	2640
gctggatgga catttggtgc	tggcgtgct cttcaaatac	cttttgctat gcaaattggca	2700
tataggttca atggcattgg	agttacccaa aatgttctct	atgagaacca aaaacaaatc	2760
gccaaccaat ttaacaaggc	gattagtcaa attcaagaat	cacttacaac aacatcaact	2820
gcattgggca agctgcaaga	cgttgttaac cagaatgctc	aagcattaaa cacacttggt	2880
aaacaactta gctctaattt	tggtgcaatt tcaagtgtgc	taaatgatat cctttcgca	2940
cttgataaag tcgaggcgga	ggtacaaatt gacagggtta	ttacaggcag acttcaaagc	3000
cttcaaacct atgtaacaca	acaactaatc agggctgctg	aatcagggc ttctgcta	3060
cttgctgcta ctaaatgtc	tgagtgtgtt cttggacaat	caaaaagagt tgacttttgt	3120
ggaaagggct accaccttat	gtccttccca caagcagccc	cgcatgggtg tgtcttccta	3180
catgtcacgt atgtgccatc	ccaggagagg aacttcacca	cagcgccagc aatttgtcat	3240
gaaggcaaag catacttccc	tcgtgaaggt gtttttgtgt	ttaatggcac ttcttggttt	3300
attacacaga ggaacttctt	ttctccacaa ataattacta	cagacaatac atttgtctca	3360
ggaaattgtg atgtcgttat	tggcatcatt aacaacacag	tttatgatcc tctgcaacct	3420
gagctcgact cattcaaaga	agagctggac aagtacttca	aaaatcatac atcaccagat	3480
gttgatcttg gcgacatttc	aggcattaac gcttctgtcg	tcaacattca aaaagaaatt	3540
gaccgcctca atgaggtcgc	taaaaattta aatgaatcac	tcattgacct tcaagaattg	3600

ggaaaatatg agcaatatat taaatggcct tgg

3633

<210> 8
 <211> 1211
 <212> PRT
 <213> SARS-CoV Urbani strain

<400> 8

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
 1 5 10 15

Ala Val Phe Val Ser Pro Ser Ala Arg Gly Ser Gly Ser Asp Leu Asp
 20 25 30

Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln His
 35 40 45

Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg Ser
 50 55 60

Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser Asn
 65 70 75 80

Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val Ile
 85 90 95

Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn Val
 100 105 110

Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln Ser
 115 120 125

Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys Asn
 130 135 140

Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met Gly
 145 150 155 160

Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr Phe
 165 170 175

Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser Gly
 180 185 190

Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly Phe
 195 200 205

Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp Leu
210 215 220

Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu Gly
225 230 235 240

Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro Ala
245 250 255

Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr Leu
260 265 270

Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile Thr
275 280 285

Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys Ser
290 295 300

Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn Phe
305 310 315 320

Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr Asn
325 330 335

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

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

Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly Val
370 375 380

Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala Asp
385 390 395 400

Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly Gln
405 410 415

Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Met
420 425 430

Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser Thr
435 440 445

Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu Arg
 450 455 460

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

Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp Tyr
 485 490 495

Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val Val
 500 505 510

Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly Pro
 515 520 525

Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn Phe
 530 535 540

Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg Phe
 545 550 555 560

Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp Ser
 565 570 575

Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys Ser
 580 585 590

Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser Glu
 595 600 605

Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr Ala
 610 615 620

Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr Gly
 625 630 635 640

Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu His
 645 650 655

Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys
 660 665 670

Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys Ser
 675 680 685

Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala Tyr

690	695	700
Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile Thr		
705	710	715 720
Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys Asn		
	725	730 735
Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu Gln		
	740	745 750
Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile Ala		
	755	760 765
Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys Gln		
	770	775 780
Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe Ser		
	785	790 795 800
Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile Glu		
	805	810 815
Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met Lys		
	820	825 830
Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile Cys		
	835	840 845
Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp		
	850	855 860
Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala Thr		
	865	870 875 880
Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala		
	885	890 895
Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn Val		
	900	905 910
Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala Ile		
	915	920 925
Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly Lys		
	930	935 940

Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val
 945 950 955 960

Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp
 965 970 975

Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp Arg
 980 985 990

Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln
 995 1000 1005

Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala
 1010 1015 1020

Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp
 1025 1030 1035

Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ala Ala
 1040 1045 1050

Pro His Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ser Gln
 1055 1060 1065

Glu Arg Asn Phe Thr Thr Ala Pro Ala Ile Cys His Glu Gly Lys
 1070 1075 1080

Ala Tyr Phe Pro Arg Glu Gly Val Phe Val Phe Asn Gly Thr Ser
 1085 1090 1095

Trp Phe Ile Thr Gln Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr
 1100 1105 1110

Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly
 1115 1120 1125

Ile Ile Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp
 1130 1135 1140

Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser
 1145 1150 1155

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

Val Asn Ile Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala Lys
 1175 1180 1185

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

Glu Gln Tyr Ile Lys Trp Pro Trp
 1205 1210

<210> 9
 <211> 2093
 <212> DNA
 <213> SARS-CoV Urbani strain

<400> 9
 atgggatgcaa tgaagagagg gctctgctgt gtgctgctgc tgtgtggagc agtcttcggt 60
 tcgcccagcg ctagaggatc gggaagtgc cttgaccggt gcaccacttt tgatgatggt 120
 caagctccta attacactca acatacttca tctatgaggg gggtttacta tcctgatgaa 180
 attttttagat cagacactct ttattttaact caggatttat ttcttccatt ttattctaata 240
 gttacaggggt ttcatactat taatcatatc tttggcaacc ctgtcatacc ttttaaggat 300
 ggtattttatt ttgctgccac agagaaaatca aatgttgtcc gtgggtgggt ttttggttct 360
 accatgaaca acaagtcaca gtcggtgatt attattaaca attctactaa tgttggtata 420
 cgagcatgta actttgaatt gtgtgacaac cctttctttg ctgtttctaa acccatgggt 480
 acacagacac atactatgat attcgataat gcatttaatt gcactttcga gtacatatct 540
 gatgcctttt cgcttgatgt ttcagaaaag tcaggtaatt ttaaactt acgagagttt 600
 gtgttttaaaa ataaagatgg gtttctctat gtttataagg gctatcaacc tatagatgta 660
 gttcgtgatc taccttctgg ttttaacact ttgaaaccta tttttaagtt gcctcttggt 720
 attaacatta caaatcttag agccattctt acagcctttt cacctgctca agacatttgg 780
 ggcacgtcag ctgcagccta ttttggtggc tatttaaagc caactacatt tatgctcaag 840
 tatgatgaaa atggtacaat cacagatgct gttgattggt ctcaaatcc acttgctgaa 900
 ctcaaatgct ctgttaagag ctttgagatt gacaaaggaa tttaccagac ctctaatttc 960
 agggttgttc cctcaggaga tgttggtgaga ttccctaata ttacaaactt gtgtcctttt 1020
 ggagagggtt ttaatgctac taaattccct tctgtctatg catgggagag aaaaaaatt 1080
 totaattgtg ttgctgatta ctctgtgctc tacaactcaa ctttttttc aacctttaag 1140
 tgctatggcg tttctgccac taagttgaat gatctttgct tctccaatgt ctatgcagat 1200
 tctttttag tcaagggaga tgatgtaaga caaatagcgc caggacaaac tgggtgttatt 1260

gctgattata attataaatt gccagatgat ttcatggggt gtgtccttgc ttggaatact 1320
 aggaacattg atgctacttc aactggtaat tataattata aatataggta tcttagacat 1380
 ggcaagctta ggccctttga gagagacata tctaattgtgc ctttctcccc tgatggcaaa 1440
 ccttgcaccc cacctgctct taattgttat tggccattaa atgattatgg tttttacacc 1500
 actactggca ttggctacca accttacaga gttgtagtac tttcttttga acttttaaact 1560
 gcaccggcca cggtttgtgg accaaaatta tccactgacc ttattaagaa ccagtgtgtc 1620
 aattttaatt ttaatggact cactgggtact ggtgtgttaa ctccctcttc aaagagattt 1680
 caaccatttc aacaatttgg ccgtgatgtt tctgatttca ctgattccgt tcgagatcct 1740
 aaaacatctg aaatattaga catttcacct tgctcttttg ggggtgtaag tgtaattaca 1800
 cctggaacaa atgcttcac tgaagttgct gttctatatc aagatgttaa ctgcactgat 1860
 gtttctacag caattcatgc agatcaactc acaccagctt ggcgcatata ttctactgga 1920
 aacaatgtat tccagactca agcaggctgt cttataggag ctgagcatgt cgacacttct 1980
 tatgagtgcg acattcctat tggagctggc atttgtgcta gttaccatac agtttcttta 2040
 ttacgtagta ctagccaaaa atctattgtg gcttatacta tgtctttagg tgc 2093

<210> 10
 <211> 698
 <212> PRT
 <213> SARS-CoV Urbani strain

<400> 10

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
 1 5 10 15

Ala Val Phe Val Ser Pro Ser Ala Arg Gly Ser Gly Ser Asp Leu Asp
 20 25 30

Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln His
 35 40 45

Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg Ser
 50 55 60

Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser Asn
 65 70 75 80

Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val Ile
 85 90 95

Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn Val

100				105				110							
Val	Arg	Gly	Trp	Val	Phe	Gly	Ser	Thr	Met	Asn	Asn	Lys	Ser	Gln	Ser
		115					120						125		
Val	Ile	Ile	Ile	Asn	Asn	Ser	Thr	Asn	Val	Val	Ile	Arg	Ala	Cys	Asn
	130					135					140				
Phe	Glu	Leu	Cys	Asp	Asn	Pro	Phe	Phe	Ala	Val	Ser	Lys	Pro	Met	Gly
145					150					155					160
Thr	Gln	Thr	His	Thr	Met	Ile	Phe	Asp	Asn	Ala	Phe	Asn	Cys	Thr	Phe
			165						170					175	
Glu	Tyr	Ile	Ser	Asp	Ala	Phe	Ser	Leu	Asp	Val	Ser	Glu	Lys	Ser	Gly
			180						185				190		
Asn	Phe	Lys	His	Leu	Arg	Glu	Phe	Val	Phe	Lys	Asn	Lys	Asp	Gly	Phe
		195					200						205		
Leu	Tyr	Val	Tyr	Lys	Gly	Tyr	Gln	Pro	Ile	Asp	Val	Val	Arg	Asp	Leu
	210					215					220				
Pro	Ser	Gly	Phe	Asn	Thr	Leu	Lys	Pro	Ile	Phe	Lys	Leu	Pro	Leu	Gly
225					230					235					240
Ile	Asn	Ile	Thr	Asn	Phe	Arg	Ala	Ile	Leu	Thr	Ala	Phe	Ser	Pro	Ala
			245						250					255	
Gln	Asp	Ile	Trp	Gly	Thr	Ser	Ala	Ala	Ala	Tyr	Phe	Val	Gly	Tyr	Leu
		260							265				270		
Lys	Pro	Thr	Thr	Phe	Met	Leu	Lys	Tyr	Asp	Glu	Asn	Gly	Thr	Ile	Thr
		275					280						285		
Asp	Ala	Val	Asp	Cys	Ser	Gln	Asn	Pro	Leu	Ala	Glu	Leu	Lys	Cys	Ser
	290					295					300				
Val	Lys	Ser	Phe	Glu	Ile	Asp	Lys	Gly	Ile	Tyr	Gln	Thr	Ser	Asn	Phe
305					310					315				320	
Arg	Val	Val	Pro	Ser	Gly	Asp	Val	Val	Arg	Phe	Pro	Asn	Ile	Thr	Asn
			325						330					335	
Leu	Cys	Pro	Phe	Gly	Glu	Val	Phe	Asn	Ala	Thr	Lys	Phe	Pro	Ser	Val
		340							345				350		

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

Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly Val
370 375 380

Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala Asp
385 390 395 400

Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly Gln
405 410 415

Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Met
420 425 430

Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser Thr
435 440 445

Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu Arg
450 455 460

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

Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp Tyr
485 490 495

Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val Val
500 505 510

Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly Pro
515 520 525

Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn Phe
530 535 540

Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg Phe
545 550 555 560

Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp Ser
565 570 575

Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys Ser
580 585 590

Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser Glu
595 600 605

Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr Ala
610 615 620

Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr Gly
625 630 635 640

Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu His
645 650 655

Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys
660 665 670

Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys Ser
675 680 685

Ile Val Ala Tyr Thr Met Ser Leu Gly Ala
690 695

<210> 11
<211> 1623
<212> DNA
<213> SARS-CoV Urbani strain

<400> 11
atggatgcaa tgaagagagg gctctgctgt gtgctgctgc tgtgtggagc agtcttcggt 60
tcgcccagcg ctagaggatc gggagatagt tcaattgctt actctaataa caccattgct 120
atacctacta acttttcaat tagcattact acagaagtaa tgccctgtttc tatggctaaa 180
acctccgtag attgtaatat gtacatctgc ggagattcta ctgaatgtgc taatttgctt 240
ctccaatatg gtagcttttg cacacaacta aatcgtgcac tctcaggtat tgctgctgaa 300
caggatcgca acacacgtga agtggtcgct caagtcaaac aaatgtacaa aacccaact 360
ttgaaatatt ttggtggttt taatttttca caaatattac ctgaccctct aaagccaact 420
aagagggtctt ttattgagga cttgctcttt aataaggatga cactcgctga tgctggcttc 480
atgaagcaat atggcgaatg cctaggatgat attaattgcta gagatctcat ttgtgcgcag 540
aagttcaatg gacttacagt gttgccacct ctgctcactg atgatatgat tgctgcctac 600
actgctgctc tagttagtgg tactgccact gctggatgga catttggtgc tggcgctgct 660
cttcaaatac cttttgctat gcaaattggca tataggttca atggcattgg agttacccaa 720
aatgttctct atgagaacca aaaacaaatc gccaaccaat ttaacaaggc gattagtcaa 780

```

attcaagaat cacttacaac aacatcaact gcattgggca agctgcaaga cgttgttaac      840
cagaatgctc aagcattaaa cacacttggt aaacaactta gctctaattt tgggtgcaatt      900
tcaagtgtgc taaatgatat cctttcgga cttgataaag tcgaggcgga ggtacaaatt      960
gacagggtta ttacaggcag acttcaaagc cttcaaacct atgtaacaca acaactaatc     1020
agggctgctg aaatcagggc ttctgctaatt cttgctgcta ctaaaatgtc tgagtgtggt     1080
cttggacaat caaaaagagt tgacttttgt ggaaagggt accaccttat gtccttccca     1140
caagcagccc cgcattggtg tgtcttccta catgtcacgt atgtgccatc ccaggagagg     1200
aacttcacca cagcgccagc aatttgtcat gaaggcaaag cataacttccc tcgtgaagggt     1260
gtttttgtgt ttaatggcac ttcttggttt attacacaga ggaacttctt ttctccacaa     1320
ataattacta cagacaatac atttgtctca ggaaattgtg atgtcggttat tggcatcatt     1380
aacaacacag tttatgatcc tctgcaacct gagctcgact cattcaaaga agagctggac     1440
aagtacttca aaaatcatac atcaccagat gttgatcttg gcgacatttc aggcattaac     1500
gcttctgtcg tcaacattca aaaagaaatt gaccgcctca atgaggtcgc taaaaattta     1560
aatgaatcac tcattgacct tcaagaattg ggaaaatatg agcaatatat taaatggcct     1620
tgg                                                                    1623

```

```

<210> 12
<211> 541
<212> PRT
<213> SARS-CoV Urbani strain

```

```

<400> 12

```

```

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
1           5           10           15

```

```

Ala Val Phe Val Ser Pro Ser Ala Arg Gly Ser Gly Asp Ser Ser Ile
          20           25           30

```

```

Ala Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser
          35           40           45

```

```

Ile Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp
          50           55           60

```

```

Cys Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu
65           70           75           80

```

```

Leu Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly
          85           90           95

```

Ile Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val
 100 105 110

Lys Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn
 115 120 125

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

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

Met Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu
 165 170 175

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

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

Ala Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro
 210 215 220

Phe Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln
 225 230 235 240

Asn Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys
 245 250 255

Ala Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu
 260 265 270

Gly Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr
 275 280 285

Leu Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu
 290 295 300

Asn Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile
 305 310 315 320

Asp Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr
 325 330 335

Gln Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala
 340 345 350

Ala Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp
 355 360 365

Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ala Ala Pro
 370 375 380

His Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ser Gln Glu Arg
 385 390 395 400

Asn Phe Thr Thr Ala Pro Ala Ile Cys His Glu Gly Lys Ala Tyr Phe
 405 410 415

Pro Arg Glu Gly Val Phe Val Phe Asn Gly Thr Ser Trp Phe Ile Thr
 420 425 430

Gln Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr Thr Asp Asn Thr Phe
 435 440 445

Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Ile Asn Asn Thr Val
 450 455 460

Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu Glu Leu Asp
 465 470 475 480

Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp Leu Gly Asp Ile
 485 490 495

Ser Gly Ile Asn Ala Ser Val Val Asn Ile Gln Lys Glu Ile Asp Arg
 500 505 510

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

Glu Leu Gly Lys Tyr Glu Gln Tyr Ile Lys Trp Pro Trp
 530 535 540

<210> 13

<211> 1269

<212> DNA

<213> SARS-CoV Urbani strain

<400> 13

atgtctgata atggacccca atcaaacc aa cgtagtgccc cccgcattac atttggtgga 60

cccacagatt caactgacaa taaccagaat ggaggacgca atggggcaag gccaaaacag 120

```

cgccgacccc aaggtttacc caataatact gcgtcttggg tcacagctct cactcagcat 180
ggcaaggagg aacttagatt cctcagaggc cagggcggtc caatcaacac caatagtggg 240
ccagatgacc aaattggcta ctaccgaaga gctacccgac gagttcgtgg tggtgacggc 300
aaaatgaaag agctcagccc cagatggtac ttctattacc taggaactgg ccagaagct 360
tcacttcctt acggcgctaa caaagaaggc atcgtatggg ttgcaactga gggagccttg 420
aatacaccca aagaccacat tggcaccggc aatcctaata acaatgctgc caccgtgcta 480
caacttcctc aaggaacaac attgccaaaa ggcttctacg cagaggggag cagagggcggc 540
agtcaagcct cttctcgctc ctcatcacgt agtcgcggta attcaagaaa ttcaactcct 600
ggcagcagta ggggaaattc tcctgctcga atggctagcg gaggtgggta aactgccctc 660
gcgctattgc tgctagacag attgaaccag cttgagagca aagtttctgg taaaggccaa 720
caacaacaag gccaaactgt cactaagaaa tctgctgctg aggcattctaa aaagcctcgc 780
caaaaacgta ctgccacaaa acagtacaac gtcactcaag catttgggag acgtgggtcca 840
gaacaaaccc aaggaaattht cggggaccaa gacctaata gacaaggaac tgattacaaa 900
cattggccgc aaattgcaca atttgctcca agtgccctctg cattctttgg aatgtcacgc 960
attggcatgg aagtcacacc ttcggaaca tggtgactt atcatggagc cattaaattg 1020
gatgacaaag atccacaatt caaagacaac gtcatactgc tgaacaagca cattgacgca 1080
tacaaaacat tcccaccaac agagcctaaa aaggacaaaa agaaaaagac tgatgaagct 1140
cagcctttgc cgcagagaca aaagaagcag ccactgtga ctcttcttcc tgcggctgac 1200
atggatgatt tctccagaca acttcaaaat tccatgagtg gagcttctgc tgattcaact 1260
caggcataa 1269

```

<210> 14

<211> 422

<212> PRT

<213> SARS-CoV Urbani strain

<400> 14

```

Met Ser Asp Asn Gly Pro Gln Ser Asn Gln Arg Ser Ala Pro Arg Ile
1           5           10           15

```

```

Thr Phe Gly Gly Pro Thr Asp Ser Thr Asp Asn Asn Gln Asn Gly Gly
20           25           30

```

```

Arg Asn Gly Ala Arg Pro Lys Gln Arg Arg Pro Gln Gly Leu Pro Asn
35           40           45

```

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

Leu Arg Phe Pro Arg Gly Gln Gly Val Pro Ile Asn Thr Asn Ser Gly
 65 70 75 80

Pro Asp Asp Gln Ile Gly Tyr Tyr Arg Arg Ala Thr Arg Arg Val Arg
 85 90 95

Gly Gly Asp Gly Lys Met Lys Glu Leu Ser Pro Arg Trp Tyr Phe Tyr
 100 105 110

Tyr Leu Gly Thr Gly Pro Glu Ala Ser Leu Pro Tyr Gly Ala Asn Lys
 115 120 125

Glu Gly Ile Val Trp Val Ala Thr Glu Gly Ala Leu Asn Thr Pro Lys
 130 135 140

Asp His Ile Gly Thr Arg Asn Pro Asn Asn Asn Ala Ala Thr Val Leu
 145 150 155 160

Gln Leu Pro Gln Gly Thr Thr Leu Pro Lys Gly Phe Tyr Ala Glu Gly
 165 170 175

Ser Arg Gly Gly Ser Gln Ala Ser Ser Arg Ser Ser Ser Arg Ser Arg
 180 185 190

Gly Asn Ser Arg Asn Ser Thr Pro Gly Ser Ser Arg Gly Asn Ser Pro
 195 200 205

Ala Arg Met Ala Ser Gly Gly Gly Glu Thr Ala Leu Ala Leu Leu Leu
 210 215 220

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

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

Lys Lys Pro Arg Gln Lys Arg Thr Ala Thr Lys Gln Tyr Asn Val Thr
 260 265 270

Gln Ala Phe Gly Arg Arg Gly Pro Glu Gln Thr Gln Gly Asn Phe Gly
 275 280 285

Asp Gln Asp Leu Ile Arg Gln Gly Thr Asp Tyr Lys His Trp Pro Gln

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

<210> 15
 <211> 1209
 <212> DNA
 <213> SARS-CoV Urbani strain

<400> 15
 atgtctgata atggacccca atcaaacc aa cgtagtgccc cccgcattac atttggtgga 60
 cccacagatt caactgacaa taaccagaat ggaggacgca atggggcaag gccaaaacag 120
 cgccgacccc aaggtttacc caataatact gcgtcttggt tcacagctct cactcagcat 180
 ggcaaggagg aacttagatt ccctcgaggc cagggcggtc caatcaacac caatagtggg 240
 ccagatgacc aaattggcta ctaccgaaga gctacccgac gagttcgtgg tggtgacggc 300
 aaaatgaaag agctcagccc cagatgggtac ttctattacc taggaactgg ccagaagct 360
 tcacttcctt acggcgctaa caaagaaggc atcgtatggg ttgcaactga gggagccttg 420
 aatacacc caaagaccacat tggcaccgc aatcctaata acaatgctgc caccgtgcta 480
 caacttcctc aaggaacaac attgccaaaa ggcttctacg cagaggggaag cagaggcggc 540
 agtcaagcct cttctcgctc ctcatcacgt agtcgcggta attcaagaaa ttcaactcct 600

```

ggcagcagta ggggaaattc tcctgctcga atggctagcg gaggtggtga aactgccctc      660
gcgctattgc tgctagacag attgaaccag cttgagagca aagtttctgg taaaggccaa      720
caacaacaag gccaaactgt cactaagaaa tctgctgctg aggcattctaa aaagcctcgc      780
caaaaacgta ctgccacaaa acagtacaac gtcactcaag catttgggag acgtgggtcca      840
gaacaaaccc aaggaaattht cggggaccaa gacctaatca gacaaggaac tgattacaaa      900
cattggccgc aaattgcaca atttgctcca agtgctcttg cattcttttg aatgtcacgc      960
attggcatgg aagtcacacc ttcgggaaca tggctgactt atcatggagc cattaaattg     1020
gatgacaaag atccacaatt caaagacaac gtcatactgc tgaacaagca cattgacgca     1080
taccctttgc cgcagagaca aaagaagcag cccactgtga ctcttcttcc tgcggctgac     1140
atggatgatt tctccagaca acttcaaaat tccatgagtg gagcttctgc tgattcaact     1200
caggcataa                                     1209

```

<210> 16
 <211> 402
 <212> PRT
 <213> SARS-CoV Urbani strain

<400> 16

Met Ser Asp Asn Gly Pro Gln Ser Asn Gln Arg Ser Ala Pro Arg Ile
 1 5 10 15

Thr Phe Gly Gly Pro Thr Asp Ser Thr Asp Asn Asn Gln Asn Gly Gly
 20 25 30

Arg Asn Gly Ala Arg Pro Lys Gln Arg Arg Pro Gln Gly Leu Pro Asn
 35 40 45

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

Leu Arg Phe Pro Arg Gly Gln Gly Val Pro Ile Asn Thr Asn Ser Gly
 65 70 75 80

Pro Asp Asp Gln Ile Gly Tyr Tyr Arg Arg Ala Thr Arg Arg Val Arg
 85 90 95

Gly Gly Asp Gly Lys Met Lys Glu Leu Ser Pro Arg Trp Tyr Phe Tyr
 100 105 110

Tyr Leu Gly Thr Gly Pro Glu Ala Ser Leu Pro Tyr Gly Ala Asn Lys
 115 120 125

Glu Gly Ile Val Trp Val Ala Thr Glu Gly Ala Leu Asn Thr Pro Lys
 130 135 140

Asp His Ile Gly Thr Arg Asn Pro Asn Asn Asn Ala Ala Thr Val Leu
 145 150 155 160

Gln Leu Pro Gln Gly Thr Thr Leu Pro Lys Gly Phe Tyr Ala Glu Gly
 165 170 175

Ser Arg Gly Gly Ser Gln Ala Ser Ser Arg Ser Ser Ser Arg Ser Arg
 180 185 190

Gly Asn Ser Arg Asn Ser Thr Pro Gly Ser Ser Arg Gly Asn Ser Pro
 195 200 205

Ala Arg Met Ala Ser Gly Gly Gly Glu Thr Ala Leu Ala Leu Leu Leu
 210 215 220

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

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

Lys Lys Pro Arg Gln Lys Arg Thr Ala Thr Lys Gln Tyr Asn Val Thr
 260 265 270

Gln Ala Phe Gly Arg Arg Gly Pro Glu Gln Thr Gln Gly Asn Phe Gly
 275 280 285

Asp Gln Asp Leu Ile Arg Gln Gly Thr Asp Tyr Lys His Trp Pro Gln
 290 295 300

Ile Ala Gln Phe Ala Pro Ser Ala Ser Ala Phe Phe Gly Met Ser Arg
 305 310 315 320

Ile Gly Met Glu Val Thr Pro Ser Gly Thr Trp Leu Thr Tyr His Gly
 325 330 335

Ala Ile Lys Leu Asp Asp Lys Asp Pro Gln Phe Lys Asp Asn Val Ile
 340 345 350

Leu Leu Asn Lys His Ile Asp Ala Tyr Pro Leu Pro Gln Arg Gln Lys
 355 360 365

Lys Gln Pro Thr Val Thr Leu Leu Pro Ala Ala Asp Met Asp Asp Phe
 370 375 380

Ser Arg Gln Leu Gln Asn Ser Met Ser Gly Ala Ser Ala Asp Ser Thr
 385 390 395 400

Gln Ala

<210> 17
 <211> 20
 <212> PRT
 <213> SARS-CoV Urbani strain

<400> 17

Lys Thr Phe Pro Pro Thr Glu Pro Lys Lys Asp Lys Lys Lys Lys Thr
 1 5 10 15

Asp Glu Ala Gln
 20

<210> 18
 <211> 666
 <212> DNA
 <213> SARS-CoV Urbani strain

<400> 18
 atggcagaca acggtactat taccgttgag gagcttaaac aactcctgga acaatggaac 60
 ctagtaatag gtttcctatt cctagcctgg attatgttac tacaatttgc ctattctaata 120
 cggaacaggt ttttgtacat aataaagctt gttttcctct ggctcttggt gccagtaaca 180
 cttgcttggt ttgtgcttgc tgctgtctac agaattaatt gggtgactgg cgggattgctg 240
 attgcaatgg cttgtattgt aggcctgatg tggcttagct acttcgttgc ttccttcagg 300
 ctgtttgctc gtacccgctc aatgtggtca ttcaaccag aaacaaacat tcttctcaat 360
 gtgcctctcc gggggacaat tgtgaccaga ccgctcatgg aaagtgaact tgtcattggt 420
 gctgtgatca ttctgtgtca cttgcgaatg gccggacacc ccctagggcg ctgtgacatt 480
 aaggacctgc caaaagagat cactgtggct acatcacgaa cgctttctta ttacaaatta 540
 ggagcgtcgc agcgtgtagg cactgattca ggttttgctg catacaaccg ctaccgtatt 600
 ggaaactata aattaaatac agaccacgcc ggtagcaacg acaatattgc tttgctagta 660
 cagtaa 666

<210> 19
 <211> 221
 <212> PRT

<213> SARS-CoV Urbani strain

<400> 19

Met Ala Asp Asn Gly Thr Ile Thr Val Glu Glu Leu Lys Gln Leu Leu
 1 5 10 15

Glu Gln Trp Asn Leu Val Ile Gly Phe Leu Phe Leu Ala Trp Ile Met
 20 25 30

Leu Leu Gln Phe Ala Tyr Ser Asn Arg Asn Arg Phe Leu Tyr Ile Ile
 35 40 45

Lys Leu Val Phe Leu Trp Leu Leu Trp Pro Val Thr Leu Ala Cys Phe
 50 55 60

Val Leu Ala Ala Val Tyr Arg Ile Asn Trp Val Thr Gly Gly Ile Ala
 65 70 75 80

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

Ala Ser Phe Arg Leu Phe Ala Arg Thr Arg Ser Met Trp Ser Phe Asn
 100 105 110

Pro Glu Thr Asn Ile Leu Leu Asn Val Pro Leu Arg Gly Thr Ile Val
 115 120 125

Thr Arg Pro Leu Met Glu Ser Glu Leu Val Ile Gly Ala Val Ile Ile
 130 135 140

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

Lys Asp Leu Pro Lys Glu Ile Thr Val Ala Thr Ser Arg Thr Leu Ser
 165 170 175

Tyr Tyr Lys Leu Gly Ala Ser Gln Arg Val Gly Thr Asp Ser Gly Phe
 180 185 190

Ala Ala Tyr Asn Arg Tyr Arg Ile Gly Asn Tyr Lys Leu Asn Thr Asp
 195 200 205

His Ala Gly Ser Asn Asp Asn Ile Ala Leu Leu Val Gln
 210 215 220

<210> 20

<211> 231
 <212> DNA
 <213> SARS-CoV Urbani strain

<400> 20
 atgtactcat tcgttttcgga agaaacaggt acgttaatag ttaatagcgt acttcttttt 60
 cttgcttttcg tgggtattctt gctagtcaca ctagccatcc ttactgcgct tcgattgtgt 120
 gcgtactgct gcaatattgt taacgtgagt ttagtaaaac caacggttta cgtctactcg 180
 cgtgttaaaa atctgaactc ttctgaagga gttcctgac ttctgggtcta a 231

<210> 21
 <211> 76
 <212> PRT
 <213> SARS-CoV Urbani strain

<400> 21
 Met Tyr Ser Phe Val Ser Glu Glu Thr Gly Thr Leu Ile Val Asn Ser
 1 5 10 15
 Val Leu Leu Phe Leu Ala Phe Val Val Phe Leu Leu Val Thr Leu Ala
 20 25 30
 Ile Leu Thr Ala Leu Arg Leu Cys Ala Tyr Cys Cys Asn Ile Val Asn
 35 40 45
 Val Ser Leu Val Lys Pro Thr Val Tyr Val Tyr Ser Arg Val Lys Asn
 50 55 60
 Leu Asn Ser Ser Glu Gly Val Pro Asp Leu Leu Val
 65 70 75

<210> 22
 <211> 3768
 <212> DNA
 <213> SARS-CoV Urbani strain

<400> 22
 atgtttatatt tcttattatt tcttactctc actagtggta gtgaccttga ccggtgcacc 60
 acttttgatg atgttcaagc tcctaattac actcaacata cttcatctat gagggggggtt 120
 tactatcctg atgaaatttt tagatcagac actctttatt taactcagga tttatttctt 180
 ccattttatt ctaatgttac agggtttcat actattaatc atacgtttgg caaccctgtc 240
 atacctttta aggatgggtat ttatttttgct gccacagaga aatcaaagt tgtccgtgggt 300
 tggggtttttg gttctaccat gaacaacaag tcacagtcgg tgattattat taacaattct 360
 actaatgttg ttatacgagc atgtaacttt gaattgtgtg acaacccttt ctttgctgtt 420

tctaaaccca tgggtacaca gacacatact atgatattcg ataatgcatt taattgcact 480
ttcgagtaca tatctgatgc cttttcgctt gatgtttcag aaaagtcagg taattttaaa 540
cacttacgag agtttgtgtt taaaaataaa gatgggtttc tctatgttta taagggctat 600
caacctatag atgtagttcg tgatctacct tctggtttta acactttgaa acctattttt 660
aagttgcctc ttggtattaa cattacaaat tttagagcca ttcttacagc cttttcacct 720
gctcaagaca tttggggcac gtcagctgca gcctattttg ttggctattt aaagccaact 780
acatttatgc tcaagtatga tgaaaatggg acaatcacag atgctgttga ttgttctcaa 840
aatccacttg ctgaactcaa atgctctgtt aagagctttg agattgacaa aggaatttac 900
cagacctcta atttcagggg tgttccctca ggagatgttg tgagattccc taatattaca 960
aacttgtgtc cttttggaga ggtttttaat gctactaaat tcccttctgt ctatgcatgg 1020
gagagaaaaa aaattttctaa ttgtgttgct gattactctg tgctctacaa ctcaacattt 1080
ttttcaacct ttaagtgcta tggcgtttct gccactaagt tgaatgatct ttgcttctcc 1140
aatgtctatg cagattcttt tgtagtcaag ggagatgatg taagacaaat agcgccagga 1200
caaactgggtg ttattgctga ttataattat aaattgccag atgatttcat gggttgtgtc 1260
cttgcttgga atactaggaa cattgatgct acttcaactg gtaattataa ttataaatat 1320
aggatatctta gacatggcaa gcttaggccc tttgagagag acatatctaa tgtgcctttc 1380
tcccctgatg gcaaaccttg caccacacct gctcttaatt gttattggcc attaaatgat 1440
tatgggtttt acaccactac tggcattggc taccaacctt acagagttgt agtactttct 1500
tttgaacttt taaatgcacc ggccacgggt tgtggaccaa aattatccac tgaccttatt 1560
aagaaccagt gtgtcaattt taattttaat ggactcactg gtactgggtg gttaactcct 1620
tcttcaaaga gatitcaacc atttcaacaa tttggccgtg atgtttctga tttcactgat 1680
tccgttcgag atcctaaaac atctgaaata ttagacattt caccttgctc ttttgggggt 1740
gtaagtgtaa ttacacctgg aacaaatgct tcatctgaag ttgctgttct atatcaagat 1800
gttaactgca ctgatgtttc tacagcaatt catgcagatc aactcacacc agcttggcgc 1860
atatattcta ctggaaacaa tgtattccag actcaagcag gctgtcttat aggagctgag 1920
catgtcgaca cttcttatga gtgcgacatt cctattggag ctggcatttg tgctagttag 1980
catacagttt ctttattacg tagtactagc caaaaatcta ttgtggctta tactatgtct 2040
ttaggtgctg atagttcaat tgcttactct aataacacca ttgctatacc tactaacttt 2100
tcaattagca ttactacaga agtaatgcct gtttctatgg ctaaaacctc cgtagattgt 2160
aatatgtaca tctgcggaga ttctactgaa tgtgctaatt tgcttctcca atatggtagc 2220
ttttgcacac aactaaatcg tgcactctca ggtattgctg ctgaacagga tcgcaacaca 2280

```

cgtgaagtgt tgcgtcaagt caaacaaatg tacaaaaccc caactttgaa atattttggg 2340
ggttttaatt tttcacaaat attacctgac cctctaaagc caactaagag gtctttttatt 2400
gaggacttgc tctttaataa ggtgacactc gctgatgctg gcttcatgaa gcaatatggc 2460
gaatgcctag gtgatattaa tgctagagat ctcatittgtg cgcagaagtt caatggactt 2520
acagtgttgc cacctctgct cactgatgat atgattgctg cctacactgc tgctctagtt 2580
agtgggtactg ccactgctgg atggacattt ggtgctggcg ctgctcttca aatacctttt 2640
gctatgcaaa tggcatatag gttcaatggc attggagtta cccaaaatgt tctctatgag 2700
aaccaaaaac aaatcgccaa ccaatttaac aaggcgatta gtcaaattca agaatcactt 2760
acaacaacat caactgcatt gggcaagctg caagacgttg ttaaccagaa tgctcaagca 2820
ttaaacacac ttgttaaaca acttagctct aatttttggtg caatttcaag tgtgctaaat 2880
gatatccttt cgcgacttga taaagtcgag gcggaggtac aaattgacag gttaattaca 2940
ggcagacttc aaagccttca aacctatgta acacaacaac taatcagggc tgctgaaatc 3000
agggtcttctg ctaatcttgc tgctactaaa atgtctgagt gtgttcttgg acaatcaaaa 3060
agagttgact tttgtggaaa gggctaccac cttatgtcct tcccacaagc agccccgcat 3120
ggtgttgtct tcctacatgt cacgtatgtg ccatcccagg agaggaactt caccacagcg 3180
ccagcaattt gtcatgaagg caaagcatac ttccctcgtg aagggtgtttt tgtgtttaat 3240
ggcacttctt ggtttattac acagaggaac ttcttttctc cacaataat tactacagac 3300
aatacatttg tctcaggaaa ttgtgatgtc gttattggca tcattaacaa cacagtttat 3360
gatcctctgc aacctgagct cgactcattc aaagaagagc tggacaagta cttcaaaaat 3420
catacatcac cagatgttga tcttggcgac atttcaggca ttaacgcttc tgtcgtcaac 3480
attcaaaaag aaattgaccg cctcaatgag gtcgctaaaa atttaaata atcactcatt 3540
gaccttcaag aattgggaaa atatgagcaa tatattaaat ggccttggtg tggttggtc 3600
ggcttcattg ctggactaat tgccatcgtc atgggtacaa tcttgctttg ttgcatgact 3660
agttgttgca gttgcctcaa gggtgcatgc tcttggtggt cttgctgcaa gtttgatgag 3720
gatgactctg agccagttct caagggtgtc aaattacatt acacataa 3768

```

```

<210> 23
<211> 1255
<212> PRT
<213> SARS-CoV Urbani strain

```

```

<400> 23

```

```

Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
1           5           10          15

```

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

His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 35 40 45

Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 50 55 60

Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 65 70 75 80

Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 85 90 95

Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
 100 105 110

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
 115 120 125

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

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
 145 150 155 160

Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
 165 170 175

Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
 180 185 190

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

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
 210 215 220

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
 225 230 235 240

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
 245 250 255

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
 260 265 270

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
 275 280 285

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
 290 295 300

Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
 305 310 315 320

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
 325 330 335

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
 340 345 350

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

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
 370 375 380

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
 385 390 395 400

Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
 405 410 415

Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
 420 425 430

Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
 435 440 445

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

Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
 465 470 475 480

Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
 485 490 495

Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
 500 505 510

Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
 515 520 525

Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
 530 535 540

Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
 545 550 555 560

Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
 565 570 575

Ser Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
 580 585 590

Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
 595 600 605

Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
 610 615 620

Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
 625 630 635 640

His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
 645 650 655

Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
 660 665 670

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

Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
 690 695 700

Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
 705 710 715 720

Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
 725 730 735

Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile

740	745	750
Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys		
755	760	765
Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe		
770	775	780
Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile		
785	790	795
Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met		
805	810	815
Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile		
820	825	830
Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr		
835	840	845
Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala		
850	855	860
Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe		
865	870	875
Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn		
885	890	895
Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala		
900	905	910
Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly		
915	920	925
Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu		
930	935	940
Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn		
945	950	955
Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp		
965	970	975
Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln		
980	985	990

Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala
 995 1000 1005

Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp
 1010 1015 1020

Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ala Ala
 1025 1030 1035

Pro His Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ser Gln
 1040 1045 1050

Glu Arg Asn Phe Thr Thr Ala Pro Ala Ile Cys His Glu Gly Lys
 1055 1060 1065

Ala Tyr Phe Pro Arg Glu Gly Val Phe Val Phe Asn Gly Thr Ser
 1070 1075 1080

Trp Phe Ile Thr Gln Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr
 1085 1090 1095

Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly
 1100 1105 1110

Ile Ile Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp
 1115 1120 1125

Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser
 1130 1135 1140

Pro Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val
 1145 1150 1155

Val Asn Ile Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala Lys
 1160 1165 1170

Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr
 1175 1180 1185

Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Val Trp Leu Gly Phe Ile
 1190 1195 1200

Ala Gly Leu Ile Ala Ile Val Met Val Thr Ile Leu Leu Cys Cys
 1205 1210 1215

Met Thr Ser Cys Cys Ser Cys Leu Lys Gly Ala Cys Ser Cys Gly
 1220 1225 1230

Ser Cys Cys Lys Phe Asp Glu Asp Asp Ser Glu Pro Val Leu Lys
 1235 1240 1245

Gly Val Lys Leu His Tyr Thr
 1250 1255

<210> 24
 <211> 3588
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Fully optimized soluble S protein

<400> 24
 atgtttatct tcctcctctt cctgacgctc actagcggat ccgacttaga tcggtgtacc 60
 acttttcgacg acgtccaggc ccctaactat actcaacata cctccagtat gcgcgggggtg 120
 tactatccag atgagatctt tcggagcgac actctgtact taacacagga cctgttttcta 180
 ccgtttttatt caaatgtaac cggcttccac accattaacc atacatttgg caatcccgtg 240
 ataccattca aagacggcat ttacttcgcc gcaacagaaa agagcaatgt tgtgagggggg 300
 tgggtcttcg gctccacaat gaacaataaa tctcagtctg tcatcatcat caataacagc 360
 actaacgtgg taatccgtgc ctgcaatttc gagctttgtg acaaccatt cttcgccgtg 420
 tctaagccta tgggcaccca gactcacaca atgatctttg acaatgcttt caactgcacc 480
 ttcgaataca tatcagatgc attctctttg gatgtcagtg aaaagtctgg aaactttaaa 540
 catctgagag agtttgtctt caaaaacaag gacggctttc tctacgttta caagggttat 600
 cagcccattg atgtggtgog ggacctccct tcagggttta acacattgaa accaatattc 660
 aaactgcccc tgggtatcaa tattactaac tttcgagcca tcttgaccgc cttttcccc 720
 gcgcaagaca tatggggaac cagcgcggca gcctatttcg tcggttatct gaagcccact 780
 acatttatgc tgaagtacga cgagaacgga accattaccg atgctgtcga ttgttcacag 840
 aatccactgg ctgaattgaa atgctccgtg aagagctttg agatcgataa ggggatttac 900
 cagacgtcta attttcgagt gggtccctca ggagatgtgg ttagattccc caatatcaca 960
 aatttgtgcc ccttcggtga agtgttcaat gccacaaagt tcccgtctgt ctacgcttgg 1020
 gagcggaaaa agataagcaa ctgtgtcgcg gattacagtg tcctatataa ctcgaccttt 1080
 tttagcacgt tcaagtgtta cggggtgagt gctactaaac tgaatgattt atgttttagt 1140
 aacgtttatg cagactcctt tggtgtaaag ggtgatgacg tgcgccaaat tgcacctggg 1200

cagaccggag	tgatcgaga	ttataactac	aaacttccag	acgactttat	gggatgctg	1260
ctcgctgga	acactcgaa	catcgacgca	accagcaccg	ggaactataa	ttacaaatac	1320
agatacctca	ggcacggcaa	gctgcggcct	tttgagcggg	atatctcaaa	cgtcccattt	1380
agcccggacg	gcaagccctg	tactcctccc	gcacttaact	gttactggcc	actgaacgat	1440
tatggctttt	ataccacaac	cggcacggc	taccagccct	accgggtggg	gggtctatct	1500
ttcgagctgc	tgaacgcgcc	tgccaccgta	tgtgggcccc	agctttcgac	agatctcatc	1560
aagaaccaat	gcgtaaattt	caatttcaat	ggccttacag	gaaccggtgt	gctgacaccc	1620
tcctccaaga	ggtttcaacc	tttcacgag	tttgagcgtg	acgtctcaga	ctttactgac	1680
agtgtgaggg	atcctaagac	ctctgaaatc	ctggatata	ctccctgttc	cttcgggtggg	1740
gttagtgtga	taacccttgg	gacaaatgct	agttccgaag	tggccgtact	ctatcaagac	1800
gtgaactgca	cagacgtgtc	aaccgccatc	cacgctgac	aactcacacc	ggcttggcgg	1860
atctatagca	ctggcaataa	cgtgttccaa	acgcaggccg	gctgccttat	aggggcagag	1920
catgtcgaca	cttcttacga	gtgtgatata	ccaatcggag	cggcatctg	cgcctcatac	1980
cacacggtga	gcttgctgcg	ctccaccagt	cagaagagta	ttgtcgcata	caccatgtca	2040
ctcggcgcag	attcaagtat	cgcctacagc	aataacacta	tcgctattcc	taccaacttt	2100
tcattttcca	tcacaactga	ggttatgcct	gtctccatgg	ctaagacttc	cgtggactgc	2160
aatatgtaca	tttgtgggga	ctctaccgag	tgcgctaacc	ttttactgca	gtatggctcc	2220
ttctgcacac	agctgaatag	agccctgagc	ggaattgccg	ctgagcagga	tagaaatacg	2280
agagaagtgt	ttgcccaggt	gaaacagatg	tataagactc	caaccttgaa	gtatttcgga	2340
gggttcaatt	ttagccagat	ccttcctgac	cccttgaagc	cgacaaaag	gagcttcac	2400
gaagatcttc	tgttcaacaa	agttacttta	gcggacgccg	ggttcatgaa	acagtatggc	2460
gagtgtctcg	gggatattaa	tgccgcgat	ctcatctgtg	ctcagaaatt	caacggcctc	2520
acagtgtctc	ccccacttct	gacggatgat	atgatcgccg	cttacacagc	cgcactcgtg	2580
agcggcaccg	ccacagccgg	ttggacattc	ggagctggag	cgcattaca	gattccattc	2640
gctatgcaga	tggcgtacag	gttcaacgga	ataggcgtga	cccagaacgt	gttgtatgaa	2700
aatcagaagc	agattgcgaa	ccagttcaac	aaagccattt	ctcaaatacca	ggagtccctg	2760
accaccacaa	gcacggcact	gggaaagctg	caagacgtgg	tcaaccagaa	cgcaccaagc	2820
ctaaataccc	tgggttaagca	gctgtctagc	aattttggag	cgatttcac	tgtccttaac	2880
gatatactat	caagactgga	caaagtggag	gcagaggtcc	aatcgaccg	cctgattacg	2940
ggccgcctcc	agagccttca	gacgtatgtg	acacagcagc	tgataagagc	tgctgaaata	3000

cgagcctcgg ctaatctggc cgcaaccaaa atgtccgaat gcgtcctggg gcagtccaaa 3060
 cgtgtcgatt tctgcggcaa aggttaccat ttgatgtcat ttccacaggc ggctcctcac 3120
 ggcgtagtgt ttctgcacgt gacttatgta ccttcgcagg aaaggaactt cacaactgcc 3180
 ccagccatct gccatgaggg aaaagcatat ttcccccgag aagggtgtttt cgttttcaac 3240
 gggacaagct ggttcattac tcaaaggaat tttttttcgc cacagatcat taccactgat 3300
 aacacatttg tatctggtaa ctgcgacgta gttatcggga ttatcaataa tacgggtctat 3360
 gaccccttgc aacctgagct ggatagcttt aaggaagagc tggacaagta ctttaagaat 3420
 cacacctctc cagacgtgga cctgggagac atctccggca ttaatgcaag tgttgtgaat 3480
 attcagaaag agattgatag actaaacgaa gttgctaaga acttgaatga gagtttaatt 3540
 gacctacagg agctcggtaa gtacgaacag tacatcaaatt ggccgtgg 3588

<210> 25
 <211> 3588
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Uniform optimization of S protein

<400> 25
 atgttcatct tcctgctggt cctgaccctg accagcggca gcgacctgga ccggtgcacc 60
 accttcgacg acgtgcaggc cccaactac acccagcaca ccagcagcat gcggggcgtg 120
 tactaccccc acgagatctt ccggagcgac accctgtacc tgaccaggga cctgttcctg 180
 cccttctaca gcaacgtgac cggcttccac accatcaacc acaccttcgg caaccccgtg 240
 atccccttca aggacggcat ctacttcgcc gccaccgaga agagcaacgt ggtgcggggc 300
 tgggtgttcg gcagcaccat gaacaacaag agccagagcg tgatcatcat caacaacagc 360
 accaacgtgg tgatccgggc ctgcaacttc gagctgtgcg acaaccctt cttcgccgtg 420
 agcaagccca tgggcacca gaccacacc atgatcttcg acaacgcctt caactgcacc 480
 ttcgagtaca tcagcgacgc cttcagcctg gacgtgagcg agaagagcgg caacttcaag 540
 cacctgcggg agttcgtggt caagaacaag gacggcttcc tgtacgtgta caagggctac 600
 cagcccatcg acgtggtgcg ggacctgcc agcggcttca acaccctgaa gcccatcttc 660
 aagctgcccc tgggcatcaa catcaccaac ttccgggcca tcctgaccgc cttcagcccc 720
 gcccaggaca tctggggcac cagcgccgcc gcctacttcg tgggctacct gaagcccacc 780
 accttcatgc tgaagtacga cgagaacggc accatcaccc acgccgtgga ctgcagccag 840
 aaccccctgg ccgagctgaa gtgcagcgtg aagagcttcg agatcgacaa gggcatctac 900
 cagaccagca acttccgggt ggtgcccagc ggcgacgtgg tgcggttccc caacatcacc 960

aacctgtgcc	ccttcggcga	ggtgttcaac	gccaccaagt	tccccagcgt	gtacgcctgg	1020
gagcggaaga	agatcagcaa	ctgcgtggcc	gactacagcg	tgctgtacaa	cagcaccttc	1080
ttcagcacct	tcaagtgcta	cggcgtgagc	gccaccaagc	tgaacgacct	gtgcttcagc	1140
aacgtgtacg	ccgacagctt	cgtggtgaag	ggcgacgacg	tgccggcagat	cgcccccggc	1200
cagaccggcg	tgatcgccga	ctacaactac	aagctgcccc	acgacttcat	gggctgcgtg	1260
ctggcctgga	acaccgggaa	catcgacgcc	accagcacccg	gcaactacaa	ctacaagtac	1320
cggtagctgc	ggcacggcaa	gctgcggccc	ttcgagcggg	acatcagcaa	cgtgcccttc	1380
agccccgacg	gcaagccctg	ccccccccc	gccctgaact	gctactggcc	cctgaacgac	1440
tacggcttct	acaccaccac	cggcatcggc	taccagccct	accgggtggg	ggtgctgagc	1500
ttcgagctgc	tgaacgcccc	cgccaccgtg	tgccggccca	agctgagcac	cgacctgac	1560
aagaaccagt	gcgtgaactt	caacttcaac	ggcctgaccg	gcaccggcgt	gctgaccccc	1620
agcagcaagc	ggttccagcc	cttccagcag	ttcgccggg	acgtgagcga	cttcaccgac	1680
agcgtgcggg	acccaagac	cagcgagatc	ctggacatca	gccctgcag	cttcggcggc	1740
gtgagcgtga	tcaccccccg	caccaacgcc	agcagcgagg	tggccgtgct	gtaccaggac	1800
gtgaactgca	ccgacgtgag	caccgccatc	cacgccgacc	agctgacccc	cgcctggcgg	1860
atctacagca	ccggcaacaa	cgtgttccag	accagggccg	gctgcctgat	cggcgccgag	1920
cacgtggaca	ccagctacga	gtgcgacatc	cccatcggcg	ccggcatctg	cgccagctac	1980
cacaccgtga	gcctgctgcg	gagcaccagc	cagaagagca	tcgtggccta	caccatgagc	2040
ctgggcgccg	acagcagcat	cgcctacagc	aacaacacca	tcgccatccc	caccaacttc	2100
agcatcagca	tcaccaccga	ggtgatgccc	gtgagcatgg	ccaagaccag	cgtggactgc	2160
aacatgtaca	tctgcggcga	cagcaccgag	tgccccaacc	tgctgctgca	gtacggcagc	2220
ttctgcaccc	agctgaaccg	ggccctgagc	ggcatcgccg	ccgagcagga	ccggaacacc	2280
cgggaggtgt	tcgcccaggt	gaagcagatg	tacaagaccc	ccaccctgaa	gtacttcggc	2340
ggcttcaact	tcagccagat	cctgccccgac	cccctgaagc	ccaccaagcg	gagcttcac	2400
gaggacctgc	tgttcaacaa	ggtgaccctg	gccgacgccg	gcttcatgaa	gcagtacggc	2460
gagtgcctgg	gcgacatcaa	cggccgggac	ctgatctgcg	cccagaagtt	caacggcctg	2520
accgtgctgc	ccccctgct	gaccgacgac	atgatcgccg	cctacaccgc	cgccttggtg	2580
agcggcaccg	ccaccgccgg	ctggacccttc	ggcgccggcg	ccgccctgca	gatcccccttc	2640
gccatgcaga	tggcctaccg	gttcaacggc	atcggcgtga	cccagaacgt	gctgtacgag	2700
aaccagaagc	agatcgccaa	ccagttcaac	aaggccatca	gccagatcca	ggagagcctg	2760

accaccacca gcaccgccct gggcaagctg caggacgtgg tgaaccagaa cgcccaggcc 2820
ctgaacaccc tgggtgaagca gctgagcagc aacttcggcg ccatcagcag cgtgctgaac 2880
gacatcctga gccggctgga caaggtggag gccgaggtgc agatcgaccg gctgatcacc 2940
ggccggctgc agagcctgca gacctacgtg acccagcagc tgatccgggc cgccgagatc 3000
cgggccagcg ccaacctggc cgccaccaag atgagcaggt gcgtgctggg ccagagcaag 3060
cggggtggact tctgcggcaa gggctaccac ctgatgagct tccccaggc cgccccccac 3120
ggcgtggtgt tcctgcacgt gacctacgtg ccagccagg agcggaactt caccaccgcc 3180
cccgccatct gccacgaggg caaggcctac ttccccggg agggcgtgtt cgtgttcaac 3240
ggcaccagct ggttcatcac ccagcggaac ttcttcagcc ccagatcat caccaccgac 3300
aacaccttcg tgagcggcaa ctgcgacgtg gtgatcggca tcatcaaaa caccgtgtac 3360
gacccccctgc agcccgagct ggacagcttc aaggaggagc tggacaagta cttcaagaac 3420
cacaccagcc ccgacgtgga cctgggagac atcagcggca tcaacgccag cgtggtgaac 3480
atccagaagg agatcgaccg gctgaacgag gtggccaaga acctgaacga gagcctgatc 3540
gacctgcagg agctgggcaa gtacgagcag tacatcaagt ggccttgg 3588

<210> 26

<211> 2049

<212> DNA

<213> Artificial Sequence

<220>

<223> Fully Optimized soluble S1 protein

<400> 26

atgtttatct ttttgctgtt tctcacatta acttcgggggt ctgacctgga ccggtgcacc 60
acattcgatg acgtccaagc ccccaactac actcagcata catctagcat gcgcggcgtg 120
tactaccag atgagatctt taggtccgac accctttatc tgaccagga ctttttctt 180
cctttctact ctaatgtaac tgggttccat accatcaacc atacctttgg caaccagtg 240
attccattta aggatggtat ttacttcgcc gcgaccgaga aatcaaagt tgtgcgcggc 300
tggtttttcg gctccaccat gaacaataag agtcagtcgg taattatcat taacaatagt 360
acaaacgtgg tgatcagggc atgtaatttt gaattgtgcg acaacccttt cttcgctgta 420
agcaaaccac tggggacgca gactcacacg atgatcttcg ataacgcttt caattgcacg 480
tttgagtaca tatccgatgc cttttctcta gatgtgtccg aaaaatcagg gaattttaag 540
cacctgagag agttcgtctt taagaacaag gacggtttct tgtacgtgta caagggatac 600
cagccgatcg acgtggtgcg ggacctacc agcggattca acaccctcaa gccattttt 660
aagctcccac tgggtatcaa tataactaac ttcagagcca ttctcacagc tttctctcca 720

```

gctcaggata tttggggggac tagtgcgga gcttatttcg tgggatacct taagcccaca      780
accttcatgt tgaaatacga tgagaacgga accataactg acgcagttga ctgctcacag      840
aacccccctcg cagagttgaa atgctcagtt aaatcctttg agatcgacaa gggatatttac      900
cagaccagta acttttagagt cgtgccgtca ggcgacgtcg tgaggtttcc taacatcaca      960
aatctatgtc ctttcggaga agtggtcaat gccacaaagt tcccagcgt gtacgcctgg     1020
gagcgaaaaa agatatctaa ctgcgtcgca gactacagcg tactgtataa cagcactttt     1080
ttcagcacct ttaagtgtta tgggggtgtca gcaacaaaac tgaacgatct ctgcttttca     1140
aacgtttatg ccgattcctt cgttgtcaag ggagacgatg tccgtcaaatt tgctcccggg     1200
caaactggcg ttatcgctga ctataactat aaactgccag acgattttat ggggtgtgtc     1260
ctcgcatgga atacgcgcaa catcgatgcg acctctaccg gaaactacaa ctataaatat     1320
aggatatctt cgcacgggaa attacggccg ttcgagcgag atatttcgaa cgtgcctttc     1380
agtcccgatg gaaaaccatg tactcctcca gccctcaatt gttactggcc attgaatgac     1440
tacgggttct acacgacaac tggaataggc tatcagcctt atcgtgtcgt cgttctttct     1500
ttcgaactgc tgaatgctcc cgccacggtg tgcggtccaa aactcagcac cgacctgac     1560
aagaatcagt gcgtgaatth caatttcaac ggctgacag gcacaggcgt tctgacccca     1620
agctccaagc gcttccagcc cttccagcaa tttggcaggg atgtgtccga ctttaccgat     1680
tcagtgcgag atcccaagac cagtgaata ctagacattt ctccgtgtag ctttggcggc     1740
gtgtctgtca ttactcctgg gacgaatgcc tcgagcgagg tggcgggtgtt atatcaggac     1800
gttaattgta cagacgtcag taccgccata catgctgac agctgactcc tgcattggaga     1860
atctactcca caggaaataa tgtgtttcag acacaagcag gttgcctgat cggagccgaa     1920
cacgtcgaca ccagctacga atgtgatatc cctatcgggt cggcatctg cgctagttat     1980
cacacagtaa gcctgctgcy gagcaccagt cagaagtcca ttgtggccta tactatgtcc     2040
ctgggcgcc                                     2049

```

<210> 27
 <211> 2049
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Uniform optimization of soluble S1 protein

```

<400> 27
atgttcatct tcctgctgtt cctgaccctg accagcggca gcgacctgga cagatgcacc      60
accttcgacg acgtgcaggc ccccaactac acccagcaca ccagcagcat gagaggcgtg     120

```

tactacccccg acgagatctt cagaagcgac accctgtacc tgacccagga cctgttcctg 180
cccttctaca gcaacgtgac cggcttccac accatcaacc acaccttcgg caacccccgtg 240
atcccccttca aggacggcat ctacttcgcc gccaccgaga agagcaacgt ggtgagaggc 300
tggggtgttcg gcagcaccat gaacaacaag agccagagcg tgatcatcat caacaacagc 360
accaacgtgg tgatcagagc ctgcaacttc gagctgtgcg acaacccctt cttcgccgtg 420
agcaagccca tgggcacca gaccacacc atgatcttcg acaacgcctt caactgcacc 480
ttcgagtaca tcagcgagc cttcagcctg gacgtgagcg agaagagcgg caacttcaag 540
cacctgagag agttcgtgtt caagaacaag gacggcttcc tgtacgtgta caagggctac 600
cagcccatcg acgtggtgag agacctgccc agcggcttca acaccctgaa gcccatcttc 660
aagctgcccc tgggcatcaa catcaccaac ttcagagcca tcctgaccgc cttcagcccc 720
gcccaggaca tctggggcac cagcgccgcc gcctacttcg tgggctacct gaagcccacc 780
accttcatgc tgaagtacga cgagaacggc accatcaccg acgccgtgga ctgcagccag 840
aacccccctgg ccgagctgaa gtgcagcgtg aagagcttcg agatcgacaa gggcatctac 900
cagaccagca acttcagagt ggtgcccagc ggcgacgtgg tgagattccc caacatcacc 960
aacctgtgcc ccttcggcga ggtgttcaac gccaccaagt tccccagcgt gtacgcctgg 1020
gagagaaaga agatcagcaa ctgcgtggcc gactacagcg tgctgtacaa cagcaccttc 1080
ttcagcacct tcaagtgcta cggcgtgagc gccaccaagc tgaacgacct gtgcttcagc 1140
aacgtgtacg ccgacagctt cgtggtgaag ggcgacgacg tgagacagat cggccccggc 1200
cagaccggcg tgatcgccga ctacaactac aagctgcccc acgacttcat gggctgcgtg 1260
ctggcctgga acaccagaaa catcgacgcc accagcaccg gcaactacaa ctacaagtac 1320
agatacctga gacacggcaa gctgagaccc ttcgagagag acatcagcaa cgtgcccttc 1380
agccccgacg gcaagccctg ccccccccc gccctgaact gctactggcc cctgaacgac 1440
tacggcttct acaccaccac cggcatcggc taccagccct acagagtggg ggtgctgagc 1500
ttcgagctgc tgaacgcccc cgccaccgtg tgcgggccca agctgagcac cgacctgatc 1560
aagaaccagt gogtgaactt caacttcaac ggctgaccg gcaccggcgt gctgaccccc 1620
agcagcaaga gattccagcc cttccagcag ttcggcagag acgtgagcga cttcaccgac 1680
agcgtgagag accccaagac cagcgagatc ctggacatca gccctgcag cttcggcggc 1740
gtgagcgtga tcacccccgg caccaacgcc agcagcgagg tggccgtgct gtaccaggac 1800
gtgaactgca ccgacgtgag caccgccatc cacgccgacc agctgacccc cgcctggaga 1860
atctacagca ccggcaacaa cgtgttccag acccaggccg gctgcctgat cggcgccgag 1920
cacgtggaca ccagctacga gtgcgacatc cccatcggcg ccggcatctg cgccagctac 1980

cacaccgtga gcctgctgag aagcaccagc cagaagagca tcgtggccta caccatgagc 2040
ctgggcgcc 2049

<210> 28
<211> 1539
<212> DNA
<213> Artificial Sequence

<220>
<223> Fully optimized S2 protein

<400> 28
gacagttcaa tcgcctattc gaacaacact atagcaatcc caacaaattt ttcaatttct 60
ataacaacag aggtgatgcc agtgtccatg gcaaagacta gcgtagactg caatatgtac 120
atctgcggag attctacaga atgtgcaaac ttgctgctac agtatggatc gttctgtacc 180
cagctcaacc gggcgctgag cggcattgct gccgaacagg atcgcaatac gagagagggtg 240
tttgctcaag tgaaacaaat gtataagacc ccaacattga aatacttcgg tggattcaat 300
ttcagtcaga ttctgccaga cccactcaaa cccaccaaga ggagctttat tgaagatctt 360
ctgttcaaca aagttacctt ggccgacgct gggtttatga agcaatacgg tgagtgcctg 420
ggcgacatta acgcacgaga cctgatctgc gcccagaagt ttaacgggct cacggtttta 480
ccgccactgc tgactgatga tatgattgcc gcttacactg cggcccttgt gagtgggtacc 540
gcaactgctg gctggacggt tggcgctggg gcggccttac agatcccttt tgccatgcag 600
atggcctaca ggttcaatgg aattggtgtc actcagaatg tcctgtacga gaaccagaaa 660
cagatcgcca accagttcaa taaagctatt tcacagattc aggaatcact taccacaact 720
tccacggcac tcggtaaact gcaggacgtg gtgaatcaga acgctcaggc actaaataca 780
ctcgtcaagc aactgagttc caatttcggg gccatatcta gcgtattgaa cgacatcctc 840
agtcggctcg acaaagtgga ggccgaagtc caaatagacc gtcttatcac aggcagacta 900
cagtcattgc agacctacgt taccagcag ttgatccgcg ccgctgagat acgagcctcc 960
gccaatctgg ccgctaccaa aatgtctgag tgtgtgctcg gacaaagtaa gcgggtggat 1020
ttttgcggca agggctatca cctcatgtcc ttccctcaag cagcacccca cggagtcgtt 1080
tttctgcatg tgacatacgt gcctagccag gagagaaact ttaccactgc gcctgccatt 1140
tgtcatgaag gcaaagctta ttttccccgc gagggggtgt tcgttttcaa cggaactagc 1200
tggtttatca caciaaggaa tttcttctcc cccagatca tcaccaccga caacaccttt 1260
gtctctggaa actgtgacgt cgttataggc atcatcaata atacagtata cgatcccctg 1320
cagcccgaac ttgactcttt caaggaggaa ctagataagt acttcaagaa tcacaccagc 1380

ccggatgtag atttagggga tattagcggg attaacgcat ccgtgggtcaa catccaaaaa 1440
gagattgaca gactgaacga agtggcgaag aacctgaatg agtccctgat cgatcttcag 1500
gagctgggca agtatgaaca gtatatcaag tggccttgg 1539

<210> 29
<211> 1539
<212> DNA
<213> Artificial Sequence

<220>
<223> Uniform Optimization of S2 protein

<400> 29
gacagcagca tcgcctacag caacaacacc atcgccatcc ccaccaactt cagcatcagc 60
atcaccaccg aggtgatgcc cgtgagcatg gccaaagacca gcgtggactg caacatgtac 120
atctgcggcg acagcaccga gtgcgccaac ctgctgctgc agtacggcag cttctgcacc 180
cagctgaacc gggccctgag cggcatcgcc gccgagcagg accggaacac ccgggaggtg 240
ttcgcccagg tgaagcagat gtacaagacc cccaccctga agtacttcgg cggcttcaac 300
ttcagccaga tcctgcccga cccctgaag cccaccaagc ggagcttcat cgaggacctg 360
ctgttcaaca aggtgaccct ggccgacgcc ggcttcatga agcagtacgg cgagtgcctg 420
ggcgacatca acgcccggga cctgatctgc gcccagaagt tcaacggcct gaccgtgctg 480
ccccccctgc tgaccgacga catgatcgcc gcctacaccg ccgccctggg gagcggcacc 540
gccaccgccg gctggacctt cggcgccggc gccgccctgc agatcccctt cgccatgcag 600
atggcctacc ggttcaacgg catcggcgtg acccagaacg tgctgtacga gaaccagaag 660
cagatcgcca accagttcaa caaggccatc agccagatcc aggagagcct gaccaccacc 720
agcaccgccc tgggcaagct gcaggacgtg gtgaaccaga acgcccaggc cctgaacacc 780
ctggtgaagc agctgagcag caacttcggc gccatcagca gcgtgctgaa cgacatcctg 840
agccggctgg acaagggtga ggccgaggtg cagatcgacc ggctgatcac cggccggctg 900
cagagcctgc agacctacgt gaccagcag ctgatccggg ccgccgagat ccgggcccagc 960
gccaacctgg ccgccaccaa gatgagcgag tgcgtgctgg gccagagcaa gcgggtggac 1020
ttctgcggca agggctacca cctgatgagc ttcccccagg ccgcccccca cggcgtgggtg 1080
ttcctgcacg tgacctacgt gccagccag gagcggaact tcaccaccgc ccccgccatc 1140
tgccacgagg gcaaggccta cttcccccg gagggcgtgt tcgtgttcaa cggcaccagc 1200
tggttcatca cccagcggaa cttcttcagc cccagatca tcaccaccga caacaccttc 1260
gtgagcggca actgcgacgt ggtgatcggc atcatcaaca acaccgtgta cgacccctg 1320
cagcccagac tggacagctt caaggaggag ctggacaagt acttcaagaa ccacaccagc 1380

cccgcagtgg acctgggcga catcagcggc atcaacgccg gcgtgggtgaa catccagaag 1440
gagatcgacc ggctgaacga ggtggccaag aacctgaacg agagcctgat cgacctgcag 1500
gagctgggca agtacgagca gtacatcaag tggccctgg 1539

<210> 30
<211> 3633
<212> DNA
<213> Artificial Sequence

<220>
<223> Fully optimized TPA-S protein

<400> 30
atggatgcaa tgaagcgggg cctgtgctgc gtgctcctgc tctgcggggc ggtgtttgtg 60
agccccagtg ccagaggtag cggcagcgat ttggataggt gcaccacatt tgatgacgtg 120
caggctccca attacaccca gcacaccagt tctatgagag gagtatacta ccctgacgag 180
atcttccgca gtgataccct atatttaaca caagatttat tcttaccctt ctactccaac 240
gtcacagggg ttcacaccat caaccacacc ttcggcaacc ccgtgatccc gtttaaagat 300
ggcatttatt tcgcagccac agagaagtcg aatgtagtgc ggggttgggt gtttggatca 360
acaatgaata ataaatctca gtccgtgatc attattaaca actctacgaa tgtgggttata 420
cgagcctgta atttcgagtt atgcgataat ccattttttcg cggtcagtaa accaatgggc 480
actcagaccc atacgatgat tttcgataac gcattcaatt gtacgtttga atacatttct 540
gatgcttttt cactcgacgt ttcagaaaag tctgggaact tcaagcattt aagagagtcc 600
gtcttttaaa ataaagacgg gttcctgtac gtgtataaag gataccagcc tatcgacgtg 660
gtgcggggacc tgccaagcgg ttttaatacc ctgaagccca tctttaagct gccctggga 720
atcaatatta caaacttcag ggctatcctc accgctttta gccagctca ggacatatgg 780
ggaacctccg ccgccgccta cttcgtcgga tatttgaaac caaccacatt catgctgaag 840
tatgacgaaa atgggacgat taccgacgcc gtagactgta gtcagaaccc tttggcggag 900
ttgaagtgct cagtcaagag ctttgagatc gacaagggaa tttatcaaac tagcaacttc 960
aggggtggtgc cctccggaga tgtagtctgc ttccccaaca tcaccaacct gtgcccgttc 1020
ggtgaggtgt ttaatgcaac taaattcccc tcagtgtatg cctgggaaag aaagaaaatt 1080
agcaactgtg ttgccgatta cagcgtcctt tataactcaa cattcttctc tacctttaag 1140
tgctatggtg tgtccgccac taagttgaac gacctctgct ttagtaacgt gtacgctgat 1200
tccttcgtgg tgaaagggga tgacgtgctg cagattgcac cgggccagac cggagtaatc 1260
gccgattaca attacaagtt gcctgacgac ttcattgggt gcgttctagc atggaatacc 1320

cgcaacatag atgccacctc aacggggaac tacaactaca agtacagata tctgagacac 1380
ggtaagctgc ggcccttttga gcgggatatc tccaatgtgc ctttttagccc cgatggcaaa 1440
ccatgcaccc cacctgccct gaattgttat tggcctttga acgattatgg attctacact 1500
accactggga tcggttatca accctaccgg gtcgtcgtcc tgagttttga actcttgaac 1560
gcgcctgcaa cagtctgcgg acccaagctg tcgacagacc ttatcaagaa tcagtgtgtg 1620
aactttaact tcaatgggct caccgggtacc ggtgttctga ctccatctag taagcgattt 1680
caaccattcc aacagttcgg ccgtgacgtt tccgatttta cggattcggg gcgtgatcca 1740
aaaacatcag agatccttga catatcgccg tgttcttttg gaggcgtgtc tgtgattaca 1800
ccaggcacta atgctagtag cgaagtcgct gtactatacc aggacgtgaa ctgcaccgac 1860
gtgagcacgg caatccacgc tgaccagctg acccccgctt ggcgcatcta cagtacaggc 1920
aataacgtct ttcagaccca ggccggctgt ctgattgggg ctgagcacgt cgacacttcc 1980
tatgaatgtg atattcccat cggcgctgga atttgtgcta gctatcacac agtctccctt 2040
ttaagatcaa ccagccagaa atctattgtg gcttacacaa tgtctctcgg cgcagactca 2100
tcaattgcct atagcaacaa taccattgca atccctacca attttagtat atccataacc 2160
accgaggtga tgcccggtgc tatggcgaaa acttccgtcg attgcaacat gtatatctgc 2220
ggggactcca cagaatgcgc caacctgctt ctgcagtatg gaagcttctg tactcaactc 2280
aaccgcgcat tgtctgggat tgccgccgag caggatagga atactagaga ggtgttcgct 2340
cagggttaaac aaatgtacaa gacaccgaca cttaagtact tcggagggtt taacttttcc 2400
cagatactcc ctgacctctt aaagcctact aaacgcagtt tcatcgagga tctcctgttt 2460
aataaggtga cactcgccga tgctggcttc atgaaacaat acggagaatg cctgggagac 2520
attaacgcca gagacctgat ctgtgcccag aagttcaacg gtctgacagt acttcctccc 2580
cttctgacgg acgacatgat tgctgcatac acagccgccc tagttagcgg cacagccaca 2640
gctgggtgga cctttggcgc tggcgcagcg ttgcagattc cattcgcgat gcagatggct 2700
taccgattta acgggatcgg cgtgactcag aatgttttgt atgagaacca gaaacagatc 2760
gctaatacgt ttaacaaggc aatcagccag atacaagaat ctctgactac cacaagcacc 2820
gctctgggaa aactgcagga cgtggtgaat cagaatgcac aggccctcaa cacgctcgtg 2880
aagcagctta gttccaattt cggggccatc tcctccgttt taaatgatat cctgagtcgc 2940
ctggacaagg tcgaggccga agttcagatc gaccgcctga tcacaggag gctacaatca 3000
ttgcagactt acgtgactca gcagctcata agggctgcag agattagggc ctctgcaaac 3060
cttgccgcga ccaagatgtc cgagtgtgtt ctcggtcagt ccaaacgggt tgacttttgt 3120
ggcaaaggct accatctgat gagcttcccc caggccgcac cccatggcgt agtctttctg 3180

```

cacgtaactt atgtgccatc ccaagaaagg aacttcacta cggcgccagc catatgccat 3240
gaaggtaaag catatttccc tcgagaaggg gtatgtgttt tcaacgggac tagctgggtt 3300
attacgcagc ggaattttctt ctcaccacaa atcatcacta ctgataacac attcgtcagc 3360
ggcaattgtg acgtcgtcat tgggaattata aacaacactg tgtacgatcc tctgcagccg 3420
gaactggatt cttttaagga ggagctcgac aagtacttca aaaaccatac ctcgcccgcac 3480
gtggacctag gcgatatctc tgggattaat gcctcagtag tcaacatcca gaaggagata 3540
gaccgactta atgaggttgc caagaatctg aatgagagtc tcatcgatct gcaagaactt 3600
ggcaagtatg aacaatatat caaatggcca tgg 3633

```

<210> 31

<211> 3633

<212> DNA

<213> Artificial Sequence

<220>

<223> Uniform optimization of TPA-S protein

<400> 31

```

atggacgcca tgaagcgggg cctgtgctgc gtgctgctgc tgtgcggcgc cgtgttcgtg 60
agccccagcg cccggggcag cggcagcgac ctggaccggt gcaccacctt cgacgacgtg 120
caggccccca actacacca gcacaccagc agcatgcggg gcgtgtacta ccccgacgag 180
atcttccgga gcgacaccct gtacctgacc caggacctgt tcctgccctt ctacagcaac 240
gtgaccggct tccacaccat caaccacacc ttcggaacc ccgtgatccc cttcaaggac 300
ggcatctact tcgccgccac cgagaagagc aacgtggtgc ggggctgggt gttcggcagc 360
accatgaaca acaagagcca gagcgtgatc atcatcaaca acagcaccaa cgtggtgatc 420
cgggcctgca acttcgagct gtgcgacaac cccttcttcg ccgtgagcaa gcccatgggc 480
accagacccc acaccatgat cttcgacaac gccttcaact gcaccttcga gtacatcagc 540
gacgccttca gcctggacgt gagcgagaag agcggcaact tcaagcacct gcgggagttc 600
gtgttcaaga acaaggacgg cttcctgtac gtgtacaagg gctaccagcc catcgacgtg 660
gtgcggggacc tgcccagcgg cttcaacacc ctgaagccca tcttcaagct gccctgggc 720
atcaacatca ccaacttccg ggccatcctg accgccttca gccccgcca ggacatctgg 780
ggcaccagcg ccgcgccta cttcgtgggc tacctgaagc ccaccacctt catgctgaag 840
tacgacgaga acggcaccat caccgacgcc gtggactgca gccagaaccc cctggccgag 900
ctgaagtgca gcgtgaagag cttcgagatc gacaagggca tctaccagac cagcaacttc 960
cgggtggtgc ccagcggcga cgtggtgcgg ttccccaaca tcaccaacct gtgccccttc 1020

```

ggcgaggtgt	tcaacgccac	caagttcccc	agcgtgtacg	cctgggagcg	gaagaagatc	1080
agcaactgcg	tggccgacta	cagcgtgctg	tacaacagca	ccttcttcag	caccttcaag	1140
tgctacggcg	tgagcgccac	caagctgaac	gacctgtgct	tcagcaacgt	gtacgccgac	1200
agcttcgtgg	tgaagggcga	cgacgtgcgg	cagatcgccc	cgggccagac	cggcgtgatc	1260
gccgactaca	actacaagct	gcccgcgcac	ttcatgggct	gcgtgctggc	ctggaacacc	1320
cggaacatcg	acgccaccag	caccggcaac	tacaactaca	agtaccggta	cctgcggcac	1380
ggcaagctgc	ggcccttcga	gcgggacatc	agcaacgtgc	ccttcagccc	cgacggcaag	1440
ccctgcaccc	cccccgccct	gaactgctac	tggcccctga	acgactacgg	cttctacacc	1500
accaccggca	tgggctacca	gccctaccgg	gtggtggtgc	tgagcttcga	gctgctgaac	1560
gccccgcga	ccgtgtgcgg	ccccaagctg	agcaccgacc	tgatcaagaa	ccagtgcgtg	1620
aacttcaact	tcaacggcct	gaccggcacc	ggcgtgctga	ccccagcag	caagcggttc	1680
cagcccttcc	agcagttcgg	ccgggacgtg	agcgacttca	ccgacagcgt	gcgggacccc	1740
aagaccagcg	agatcctgga	catcagcccc	tgcagcttcg	gcggcgtgag	cgtgatcacc	1800
ccgggcacca	acgccagcag	cgaggtggcc	gtgctgtacc	aggacgtgaa	ctgcaccgac	1860
gtgagcaccg	ccatccacgc	cgaccagctg	acccccgcct	ggcggatcta	cagcaccggc	1920
aacaacgtgt	tccagaccca	ggccggctgc	ctgatcggcg	ccgagcacgt	ggacaccagc	1980
tacgagtgcg	acatccccat	cggcgccggc	atctgcgcga	gctaccacac	cgtgagcctg	2040
ctgcggagca	ccagccagaa	gagcatcgtg	gcctacacca	tgagcctggg	cgccgacagc	2100
agcatcgcct	acagcaacaa	caccatcgcc	atccccacca	acttcagcat	cagcatcacc	2160
accgaggtga	tgcccggtgag	catggccaag	accagcgtgg	actgcaacat	gtacatctgc	2220
ggcgacagca	ccgagtgcgc	caacctgctg	ctgcagtacg	gcagcttctg	caccagctg	2280
aaccggggccc	tgagcggcat	cgccgcccag	caggaccgga	acaccgggga	ggtgttcgcc	2340
caggtgaagc	agatgtacaa	gacccccacc	ctgaagtact	tcggcggctt	caacttcagc	2400
cagatcctgc	ccgacccccct	gaagcccacc	aagcggagct	tcacgcagga	cctgctgttc	2460
aacaagggtga	ccctggccga	cgccggcttc	atgaagcagt	acggcgagtg	cctgggagac	2520
atcaacgccc	gggacctgat	ctgcgcccag	aagttcaacg	gcctgaccgt	gctgcccccc	2580
ctgctgaccg	acgacatgat	cgccgcctac	accgccgccc	tggtgagcgg	caccgccacc	2640
gcgggctgga	ccttcggcgc	cggcgccgcc	ctgcagatcc	ccttcgccat	gcagatggcc	2700
taccggttca	acggcatcgg	cgtgaccag	aacgtgctgt	acgagaacca	gaagcagatc	2760
gccaaaccagt	tcaacaaggc	catcagccag	atccaggaga	gcctgaccac	caccagcacc	2820
gccctgggca	agctgcagga	cgtggtgaac	cagaacgccc	aggccctgaa	caccctggtg	2880

```

aagcagctga gcagcaactt cggcgccatc agcagcgtgc tgaacgacat cctgagccgg 2940
ctggacaagg tggaggccga ggtgcagatc gaccggctga tcaccggccg gctgcagagc 3000
ctgcagacct acgtgaccca gcagctgatc cgggcccgcg agatccgggc cagcgccaac 3060
ctggccgcca ccaagatgag cgagtgcgtg ctgggcccaga gcaagcgggt ggacttctgc 3120
ggcaagggtt accacctgat gagcttcccc caggccgccc cccacggcgt ggtgttcctg 3180
cacgtgacct acgtgcccag ccaggagcgg aacttcacca ccgccccgc catctgccac 3240
gagggcaagg cctacttccc ccgggagggc gtgttcgtgt tcaacggcac cagctggttc 3300
atcaccacgc ggaacttctt cagccccag atcatcacca ccgacaacac cttcgtgagc 3360
ggcaactgcg acgtggtgat cggcatcatc aacaacaccg tgtacgacct cctgcagccc 3420
gagctggaca gcttcaagga ggagctggac aagtacttca agaaccacac cagccccgac 3480
gtggacctgg gcgacatcag cggcatcaac gccagcgtgg tgaacatcca gaaggagatc 3540
gaccggctga acgaggtggc caagaacctg aacgagagcc tgatcgacct gcaggagctg 3600
ggcaagtacg agcagtacat caagtggccc tgg 3633

```

<210> 32

<211> 2094

<212> DNA

<213> Artificial Sequence

<220>

<223> Fully optimized soluble TPA-S1 protein

<400> 32

```

atggacgcca tgaagcgagg actgtgctgc gttttgttgc tgtgcggcgc agtttttgtc 60
agtccatccg cccgggggtc gggatctgac ctagatagat gcacgacctt cgatgacgtg 120
caggcaccaa attacacca acatacttca tccatgcgcg gcgtttacta tcccgcgaa 180
atcttccgga gtgacaccct gtatctgact caggacctgt ttctgccctt ctacagcaat 240
gtgacaggct ttcacaccat taaccatacc ttcgggaatc cagtaatccc ttttaaggat 300
gggatttact ttgctgctac tgagaaaagt aatgttgtca ggggggtgggt ttttggctca 360
acaatgaaca ataagtctca gagtgtcatc atcattaaca attctaccaa tgtagtcac 420
agagcatgca acttcgagct ctgtgataac cttttctttg ctgtgtctaa gcccatgggc 480
actcaaacac ataccatgat cttcgacaat gcgttcaatt gtacctttga gtatatatca 540
gacgccttca gcctagacgt ctcggaaaag tccggaaact ttaaacacct gcgggaattc 600
gtgtttaaga acaaagatgg atttttgtac gtatacaagg gttatcagcc tatcgatgtc 660
gtgcgtgata tgccctccgg cttcaacacc ctgaagccta tattcaaact acccctaggg 720

```

```

atcaacatca ccaatthttag ggcaatactt acggcatttt cccagacca ggacatctgg      780
ggaacttccg ccgctgccta ctttgtgggc tatctcaagc ctactacttt catgcttaag      840
tatgatgaga atggcacaat cacggatgca gtggattgct cgcagaatcc acttgctgag      900
ctgaaatgct ccgtaaagag cttcgaaatt gataaaggaa tctatcagac cagcaacttc      960
cgggtcgtgc cctctggcga cgttgtccgg ttccccaaca tcaccaacct ctgcccattc     1020
ggcgagggtgt tcaacgctac aaaattccca agtgtctacg cctgggagag gaaaaagatc     1080
tctaattgtg tggcagatta ttccgtgtta tacaacagca cattcttctc aacgttcaag     1140
tgttatggcg tgagcgccac caagcttaac gacctctgct tctccaatgt atacgctgac     1200
tcttttgtgg ttaagggaga cgatgtgcga cagatcgccc cggggcaaac cggagtgatt     1260
gcggactaca actataaact gcccgcgat ttcatgggtt gtgtgcttgc ttggaatacg     1320
aggaacattg acgcaacgag caccgggaac tataattaca aatatcgta cctgcgccat     1380
gggaaactca gaccttttga acgagatatt agcaacgtcc ctttctcacc ggatgggaag     1440
ccctgtaccc cacctgccct gaactgctat tggcctctca acgactacgg cttctacact     1500
accacagggg tcgggtacca gccctatcgc gtggtgggtc tctccttga actccttaat     1560
gctcccgcga ctgtgtgtgg gccgaagttg agtactgact taataaaaaa tcaatgcgta     1620
aactttaact ttaatggctt gacaggtaca ggtgtgctca caccgagtag caaaagggtc     1680
cagccatttc agcaatthtg cagagatgtg tctgacttta cagacagcgt gcgcgaccc     1740
aagacttctg agatthttag catctcacct tgthccttg gaggagttag cgtgataact     1800
cccgttacca acgcctcatc cgaagtggct gtcctgtatc aggacgttaa ttgcaccgat     1860
gtctctacag ccattcacgc cgatcagctg acaccagctt ggcgcatcta cagtaccggt     1920
aacaatgttt tccagactca ggccggttgt ctgattggcg ccgagcacgt cgacacatct     1980
tacgagtgcg atattcccat aggtgccggc atthgtgcga gctaccacac tgtatcactg     2040
ctgagaagca caagccagaa atcaattgtg gcatacacia tgtccttggg agca          2094

```

<210> 33

<211> 2091

<212> DNA

<213> Artificial Sequence

<220>

<223> Uniform optimization of soluble TPA-S1 protein

<400> 33

```

atggacgcca tgaagcgggg cctgtgctgc gtgctgctgc tgtgcggcgc cgtgttcgtg      60
agccccagcg cccggggcag cggcagcgac ctggaccggt gcaccacctt cgacgacgtg     120
caggccccca actacacca gcacaccagc agcatgcggg gcgtgtacta ccccgacgag     180

```

atcttccgga gcgacaccct gtacctgacc caggacctgt tcctgccctt ctacagcaac	240
gtgaccggct tccacaccat caaccacacc ttccggcaacc ccgtgatccc cttcaaggac	300
ggcatctact tcgccgccac cgagaagagc aacgtggtgc ggggctgggt gttcggcagc	360
accatgaaca acaagagcca gagcgtgac atcatcaaca acagcaccaa cgtggtgac	420
cgggcctgca acttcgagct gtgcgacaac cccttcttcg ccgtgagcaa gcccatgggc	480
accagaccc acaccatgat cttcgacaac gccttcaact gcaccttcga gtacatcagc	540
gacgccttca gcctggacgt gagcgagaag agcggcaact tcaagcacct gcgggagttc	600
gtgttcaaga acaaggacgg cttcctgtac gtgtacaagg gctaccagcc catcgacgtg	660
gtgcgggacc tgcccagcgg cttcaacacc ctgaagccca tcttcaagct gcccctgggc	720
atcaacatca ccaacttccg ggccatcctg accgccttca gcccgcacca ggacatctgg	780
ggcaccagcg ccgccgcta cttcgtgggc tacctgaagc ccaccacctt catgctgaag	840
tacgacgaga acggcaccat caccgacgcc gtggactgca gccagaacct cctggccgag	900
ctgaagtgca gcgtgaagag cttcgagatc gacaagggca tctaccagac cagcaacttc	960
cgggtggtgc ccagcggcga cgtggtgcgg ttccccaaca tcaccaacct gtgccccttc	1020
ggcgaggtgt tcaacgccac caagttcccc agcgtgtacg cctgggagcg gaagaagatc	1080
agcaactgcg tggccgacta cagcgtgctg tacaacagca ccttcttcag caccttcaag	1140
tgctacggcg tgagcgccac caagctgaac gacctgtgct tcagcaacgt gtacgccgac	1200
agcttcgtgg tgaaggcgga cgacgtgcgg cagatcgccc ccggccagac cggcgtgac	1260
gccgactaca actacaagct gcccgacgac ttcattgggt gcgtgctggc ctggaacacc	1320
cggaacatcg acgccaccag caccggcaac tacaactaca agtaccggta cctgcggcac	1380
ggcaagctgc ggcccttcga gcgggacatc agcaacgtgc ccttcagccc cgacggcaag	1440
ccctgcaccc ccccgccct gaactgtac tggcccctga acgactacgg cttctacacc	1500
accaccggca tcggctacca gccctaccgg gtggtggtgc tgagcttcga gctgctgaac	1560
gccccgcca ccgtgtgcgg cccaagctg agcaccgacc tgatcaagaa ccagtgcgtg	1620
aacttcaact tcaacggcct gaccggcacc ggctgtctga ccccagcag caagcggttc	1680
cagcccttcc agcagttcgg ccgggacgtg agcgacttca ccgacagcgt gcgggacccc	1740
aagaccagcg agatcctgga catcagcccc tgcagcttcg gcggcgtgag cgtgatcacc	1800
cccggcacca acgccagcag cgaggtggcc gtgctgtacc aggacgtgaa ctgcaccgac	1860
gtgagcaccg ccatccacgc cgaccagctg accccgcct ggcggtacta cagcaccggc	1920
aacaacgtgt tccagaccca ggccggctgc ctgatcggcg ccgagcacgt ggacaccagc	1980

tacgagtgcg acatccccat cggcgccggc atctgcgcca gctaccacac cgtgagcctg 2040
 ctgcggagca ccagccagaa gagcatcgtg gcctacacca tgagcctggg c 2091

<210> 34
 <211> 1623
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Fully optimized soluble TPA-S2

<400> 34
 atggatgcaa tgaaaagagg cctgtgttgt gttctgctgc tgtgtggggc ggtatttgtg 60
 agtccctctg ccaggggaag cggcgacagc agtatagcct actcaaacia taccatcgcc 120
 attcctacaa atttttccat ctcaatcacg acggaagtca tgccagttag catggccaaa 180
 acctctgtcg actgcaacat gtacatctgc ggagactcta ctgagtgcgc aaacctgctc 240
 ttgcagtatg gctcgttttg caccagttg aatcggggcc tcagtggcat tgccgcagaa 300
 caagatcgga ataccaggga ggtcttcgcg caagtcaagc agatgtacaa aaccctaca 360
 ctcaaatact tcgggggggt caactttagc caaatcctgc cagacccct caagcctact 420
 aagcgcagtt ttatcgaaga cttactcttt aataagggtga cattagctga tgccggattc 480
 atgaagcagt acggagagt cctgggggat atcaacgcgc gggaccta atctgtgccag 540
 aagttcaacg gtctgacagt gcttcgcct ctctgaccg atgatatgat cgcagcttac 600
 accgccgcac tggttagtgg tacggccaca gcaggctgga ccttcggtgc cgggtgctgcc 660
 ctgcaaatac cattcgcgat gcagatggca tacagattta acggcattgg agtcaccag 720
 aatgtcctat acgagaacca gaagcaaata gctaaccagt tcaacaaagc catatcccag 780
 attcaggagt cccttactac aaccagtact gctttaggta aactgcaaga ttagtgtaac 840
 cagaacgctc aggccttaaa tacccttggt aaacagctat cctcaaactt tggggctatc 900
 tcctccgtgc tcaacgatat cctgagccgc ctcgataagg tggaagcgga ggtccagatc 960
 gatagactta ttacaggcag gcttcagtct ctccagacct atgtcacaca acagctcatt 1020
 cgtgctgcag agatccgcgc ttccgccaac ttggctgcaa caaagatgtc tgaatgtgtg 1080
 ctgggacaga gcaagagagt ggacttttgt gggaaaggct atcacttgat gagcttcccc 1140
 caggccgccc cccatggagt ggtattccta cacgtgacgt acgttccatc tcaagaacga 1200
 aatttcacca ccgcacctgc catttgccac gaagggaagg cttatttccc tcgagagggc 1260
 gtgttcgttt ttaacgggac ttcatgggtt ataactcaaa ggaatttctt ctgcgccag 1320
 ataattacaa cagacaacac ttttgtgagc ggcaattgcg acgtgggtcat aggtattatt 1380
 aataatactg tgtatgacct gctgcagccc gaactggaca gctttaaaga ggagctggac 1440

aaatacttca	agaatcatac	ttcacccgac	gtggatctgg	gcgacatatc	cggaatcaat	1500
gcctctgtgg	taaacattca	gaaggagatc	gatcggctga	acgaagtggc	taagaatctg	1560
aatgaatcat	tgattgacct	tcaggagttg	ggcaagtatg	agcagtatat	taaatggcca	1620
tgg						1623

<210> 35
 <211> 1623
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Uniform optimization of TPA-S2 protein

<400> 35	
atggacgcca	tgaagcgggg cctgtgctgc gtgctgctgc tgtgcggcgc cgtgttcgtg 60
agccccagcg	cccggggcag cggcgacagc agcatcgctt acagcaacaa caccatcgcc 120
atccccacca	acttcagcat cagcatcacc accgaggtga tgcccgtgag catggccaag 180
accagcgtgg	actgcaacat gtacatctgc ggcgacagca ccgagtgcgc caacctgctg 240
ctgcagtacg	gcagcttctg caccagctg aaccggggccc tgagcggcat cgccgccgag 300
caggaccgga	acaccgggga ggtgttcgcc caggtgaagc agatgtacaa gacccccacc 360
ctgaagtact	tcggcggtt caacttcagc cagatcctgc ccgacccccct gaagcccacc 420
aagcggagct	tcatcgagga cctgctgttc aacaaggtga ccctggccga cgccggcttc 480
atgaagcagt	acggcgagtg cctgggcgac atcaacgccc gggacctgat ctgcgcccag 540
aagttcaacg	gcctgaccgt gctgcccccc ctgctgaccg acgacatgat cgccgcctac 600
accgccgccc	tggtgagcgg caccgccacc gccggctgga ccttcggcgc cggcgccgcc 660
ctgcagatcc	ccttcgccat gcagatggcc taccggttca acggcatcgg cgtgaccag 720
aacgtgctgt	acgagaacca gaagcagatc gccaaaccagt tcaacaaggc catcagccag 780
atccaggaga	gcctgaccac caccagcacc gccctgggca agctgcagga cgtggtgaac 840
cagaacgccc	aggccctgaa caccctggtg aagcagctga gcagcaactt cggcgccatc 900
agcagcgtgc	tgaacgacat cctgagccgg ctggacaagg tggaggccga ggtgcagatc 960
gaccggctga	tcaccggccg gctgcagagc ctgcagacct acgtgacca gcagctgac 1020
cgggcccgcg	agatccgggc cagcgccaac ctggccgcca ccaagatgag cgagtgcgtg 1080
ctggggccaga	gcaagcgggt ggacttctgc ggcaagggt accacctgat gagcttcccc 1140
caggccgccc	cccacggcgt ggtgttcctg cacgtgacct acgtgccag ccaggagcgg 1200
aacttcacca	ccgccccgc catctgccac gagggcaagg cctacttccc ccgggagggc 1260

```

gtgttcgtgt tcaacggcac cagctgggtc atcaccacagc ggaacttctt cagccccag 1320
atcatcacca ccgacaacac cttcgtgagc ggcaactgcg acgtggtgat cggcatcatc 1380
aacaacaccg tgtacgaccc cctgcagccc gagctggaca gcttcaagga ggagctggac 1440
aagtacttca agaaccacac cagccccgac gtggacctgg gcgacatcag cggcatcaac 1500
gccagcgtgg tgaacatcca gaaggagatc gaccggctga acgaggtggc caagaacctg 1560
aacgagagcc tgatcgacct gcaggagctg ggcaagtacg agcagtacat caagtggccc 1620
tgg 1623

```

<210> 36

<211> 1269

<212> DNA

<213> Artificial Sequence

<220>

<223> Fully optimized N protein

<400> 36

```

atgtccgata atggtcccca gtctaaccag aggtcggcgc caagaatcac attcgggggc 60
ccaacagaca gtaccgataa caaccagaac ggcggaagaa acggggccag gcccaagcag 120
cggagacctc agggattacc aaataatacc gcaagctggt tcacagccct gaccagcat 180
ggaaaagagg aactgagatt ccctagagga caaggggtgc ctattaatac taatagcggg 240
cctgacgatc aaattggcta ttatcgacgt gcgactcgcc gtgttagagg gggggacggg 300
aagatgaagg agcttagccc acgctggtac ttttactatc tgggaaccgg acctgaagct 360
agtctgccct acggcgctaa caaggaggga atagtatggg tcgccacgga aggtgcgttg 420
aatactccga aagatcacat cggcaccaga aatcctaaca ataacgccgc aaccgtgcta 480
caattacccc agggaaactac tctgccgaag gggttctatg cggaggggaag ccgcggcggc 540
tcacaagcca gttcacgctc cagctcccgg tcgaggggta attcccgaag cagcaccg 600
ggatcatcta ggggaaactc tcccgcccgg atggcctcag gcggcggcga aacagctctg 660
gctctgctat tgctggaccg gctcaaccag ctcgagtcca aagtctctgg taaaggctcag 720
cagcagcagg gtcaaacagt gaccaaaaaa agtgcagccg aggccagcaa gaaaccacgc 780
cagaaacgta cggccacaaa gcaatacaat gtgaccaag cctttggaag gcggggggccc 840
gaacagacac agggcaatth cggcgatcaa gatthgatac gacagggcac tgactacaaa 900
cactggccgc agatcgctca gtttgcaact agcgctcccg ctttctttgg catgagtcgg 960
attggcatgg aggtgacacc atcaggtact tgggtaacgt accacggggc aatcaaactt 1020
gatgataaag atccccagtt taaggacaac gttatcctcc tgaataagca tattgacgcc 1080
tataagacct tcccccaac cgaaccaaag aaggacaaga agaagaagac agacgaggca 1140

```

cagcctctcc	cccagaggca	gaaaaagcag	cctactgtca	cccttctgcc	cgctgcagac	1200
atggatgact	tttcccgcc	actccagaac	tctatgagtg	gggcttccgc	tgactctacg	1260
caggcctga						1269

<210> 37
 <211> 1266
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Uniform optimization of N protein

<400> 37	
atgagcgaca	acggccccca gagcaaccag agaagcgccc ccagaatcac cttcggcggc 60
cccaccgaca	gcaccgacaa caaccagaac ggcggcagaa acggcgccag acccaagcag 120
agaagacccc	agggcctgcc caacaacacc gccagctggt tcaccgccct gaccagcac 180
ggcaaggagg	agctgagatt ccccagaggc cagggcgtgc ccatcaacac caacagcggc 240
cccgacgacc	agatcggcta ctacagaaga gccaccagaa gagtgagagg cggcgacggc 300
aagatgaagg	agctgagccc cagatggtac ttctactacc tgggcaccgg ccccgaggcc 360
agcctgccct	acggcgccaa caaggagggc atcgtgtggg tggccaccga gggcgccctg 420
aacaccccc	aggaccacat cggcaccaga aacccaaca acaacgccgc caccgtgctg 480
cagctgcccc	agggcaccac cctgcccag ggcttctacg ccgagggcag cagaggcggc 540
agccaggcca	gcagcagaag cagcagcaga agcagaggca acagcagaaa cagcaccccc 600
ggcagcagca	gaggcaacag ccccgccaga atggccagcg gcggcggcga gaccgccctg 660
gccctgctgc	tgctggacag actgaaccag ctggagagca aggtgagcgg caagggccag 720
cagcagcagg	gccagaccgt gaccaagaag agcgccgccc aggccagcaa gaagcccaga 780
cagaagagaa	ccgccaccaa gcagtacaac gtgaccagg ccttcggcag aagaggcccc 840
gagcagaccc	agggcaactt cggcgaccag gacctgatca gacagggcac cgactacaag 900
cactggcccc	agatcgccca gttcgcccc agcgccagcg ctttcttcgg catgagcaga 960
atcggcatgg	aggtgacccc cagcggcacc tggctgacct accacggcgc catcaagctg 1020
gacgacaagg	acccccagtt caaggacaac gtgatcctgc tgaacaagca catcgacgcc 1080
tacaagacct	tcccccccac cgagcccaag aaggacaaga agaagaagac cgacgaggcc 1140
cagcccctgc	cccagagaca gaagaagcag cccaccgtga ccctgctgcc cgccgccgac 1200
atggacgact	tcagcagaca gctgcagaac agcatgagcg gcgccagcgc cgacagcacc 1260
caggcc	
	1266

<210> 38
<211> 1209
<212> DNA
<213> Artificial Sequence

<220>
<223> Fully optimized N protein lacking NLS

<400> 38
atgagtgata atggcccccga gtctaaccag aggagcgcac cgcggatcac gttcgggtggc 60
ccaaccgact caacagacaa taatcagaac ggaggacgca atggtgcacg tcctaagcag 120
agacgccccc aagggtgcc taataataca gcaagttggt ttaccgcact cacacaacat 180
ggaaaggaag agttgcggtt cccccgcggc cagggcgtgc ccatcaacac aaatagcgga 240
cccgacgata agatcggata ttaccgaaga gctacaagga gagttcgcgg cggggatggc 300
aagatgaagg agctatcacc acgatggtac ttctattacc tcgggacagg ccagagggcc 360
tcgctaccat acggggccaa caaggagggt attgtctggg tcgctaccga aggggccctg 420
aatacaccta aagaccacat aggtaccaga aatcccaaca ataacgccgc gaccgtgtta 480
cagcttcctc agggaacgac cttccaaaaa gggttttacg ccgaaggatc tcggggaggg 540
tcacaggcta gctcccgtag ctctcaagg tccaggggga attctagaaa cagtacaccc 600
ggctctagcc gtggtaactc ccagctcgc atggcatccg gcggagggga aaccgctctg 660
gctctgctcc tgtagatcg gttgaaccaa ctggaatcga aggtatccgg aaagggacag 720
cagcagcaag gccagactgt gactaagaag tccgcggccg aggccagtaa gaaacccgc 780
cagaaacgaa ctgccaccaa acagtataat gtgacacagg cttcggcag acggggtcca 840
gagcagaccc aaggcaactt cggggatcag gacctgattc ggaggggtac cgactataag 900
cactggccgc aaattgctca gtttgctccc agtgcgagtg ctttcttcgg catgtctagg 960
atcgggatgg aggttactcc tagcggcact tggcttactt atcacggagc catcaaactc 1020
gatgataagg acccacagtt taaggataac gtgattctgc tgaacaaaca tatagacgcg 1080
taccctctcc cgcaaaggca gaaaaaacag cctaccgtca cgttactgcc tgccgcagac 1140
atggacgact tttctagaca gttgcaaaac agcatgtcag gcgcatccgc cgatagcact 1200
caagcttga 1209

<210> 39
<211> 1206
<212> DNA
<213> Artificial Sequence

<220>
<223> Uniform optimization of N protein lacking NLS

```

<400> 39
atgagcgaca acggcccccga gagcaaccag agaagcgccc ccagaatcac cttcggcggc 60
cccaccgaca gcaccgacaa caaccagaac ggcggcagaa acggcgccag acccaagcag 120
agaagacccc agggcctgcc caacaacacc gccagctggt tcaccgccct gaccagcac 180
ggcaaggagg agctgagatt ccccagaggc cagggcgtgc ccatcaacac caacagcggc 240
cccgacgacc agatcggcta ctacagaaga gccaccagaa gagtgagagg cggcgacggc 300
aagatgaagg agctgagccc cagatggtac ttctactacc tgggcaccgg ccccgaggcc 360
agcctgccct acggcgccaa caaggagggc atcgtgtggg tggccaccga gggcgccctg 420
aacacccccca aggaccacat cggcaccaga aacccaaca acaacgccgc caccgtgctg 480
cagctgcccc agggcaccac cctgccaaag ggcttctacg ccgagggcag cagaggcggc 540
agccaggcca gcagcagaag cagcagcaga agcagaggca acagcagaaa cagcaccccc 600
ggcagcagca gaggcaacag ccccgccaga atggccagcg gcggcggcga gaccgccctg 660
gccctgctgc tgctggacag actgaaccag ctggagagca aggtgagcgg caagggccag 720
cagcagcagg gccagaccgt gaccaagaag agcgccgcg aggccagcaa gaagcccaga 780
cagaagagaa ccgccaccaa gcagtacaac gtgaccagg ccttcggcag aagaggcccc 840
gagcagaccc agggcaactt cggcgaccag gacctgatca gacagggcac cgactacaag 900
cactggcccc agatcgccca gttcgcccc agcgccagcg cttctctcgg catgagcaga 960
atcggcattg aggtgacccc cagcggcacc tggctgacct accacggcgc catcaagctg 1020
gacgacaagg acccccagtt caaggacaac gtgatcctgc tgaacaagca catcgacgcc 1080
tacccccctgc ccagagaca gaagaagcag cccaccgtga ccctgctgcc cgccgccgac 1140
atggacgact tcagcagaca gctgcagaac agcatgagcg gcgccagcgc cgacagcacc 1200
caggcc 1206

```

<210> 40

<211> 666

<212> DNA

<213> Artificial Sequence

<220>

<223> Fully optimized M protein

<400> 40

```

atggctgaca acggcaccat aaccgtcgag gagcttaaac agttattaga acaatggaac 60
ttggtgatag gattcctctt tctggcatgg atcatgttgc ttcagttcgc ctattctaac 120
cgcaataggt ttttgtacat tatcaagctg gtcttccttt ggctgctctg gcccgtaaca 180
ctagcctgtt ttgttttggc ggccgtgtat cggatcaatt gggtgacagg tggcattgct 240

```

```

attgcatggtg cttgcatcgt ggggctgatg tggctgtcgt atttcgttgc ctcattccgg      300
ctgttttggcc gaacaaggag tatgtggtct ttttaaccccg agaccaatat tctgctcaat      360
gtgccttttac gcggcactat cgtgaccccg cctctaattg aatccgagct ggtaattggc      420
gcagtcacatca taaggggggca cctcagaatg gccggggcacc cacttgggag atgcgacatc      480
aaggatctgc cgaaggaaat tactgttgca acttcacgaa cgctgagcta ttacaaactg      540
ggagctagcc agagagtggg taccgactcc ggcttcgctg cctacaaccg ctaccgtatc      600
ggaaattaca aactcaacac agatcatgca ggaagcaatg ataacatcgc ctccttggtc      660
cagtga

```

```

<210> 41
<211> 663
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Uniform optimization of M protein

```

```

<400> 41
atggccgaca acggcaccat caccgtggag gagctgaagc agctgctgga gcagtggaaac      60
ctgggtgatcg gcttcctgtt cctggcctgg atcatgctgc tgcagttcgc ctacagcaac      120
agaaacagat tcctgtacat catcaagctg gtgttcctgt ggctgctgtg gcccgtagacc      180
ctggcctgct tcgtgctggc cgccgtgtac agaatcaact gggtagaccg cgccatcgcc      240
atcgccatgg cctgcatcgt gggcctgatg tggctgagct acttcgtggc cagcttcaga      300
ctgttcgcca gaaccagaag catgtggagc ttcaaccccg agaccaacat cctgctgaac      360
gtgcccctga gaggcaccat cgtgaccaga cccctgatgg agagcgagct ggtgatcggc      420
gccgtgatca tcagaggcca cctgagaatg gccggccacc ccctgggcag atgcgacatc      480
aaggacctgc ccaaggagat caccgtggcc accagcagaa ccctgagcta ctacaagctg      540
ggcgccagcc agagagtggg caccgacagc ggcttcgccg cctacaacag atacagaatc      600
ggcaactaca agctgaacac cgaccacgcc ggcagcaacg acaacatcgc cctgctgggtg      660
cag

```

```

<210> 42
<211> 231
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Fully optimized E protein

```

```

<400> 42
atgtacagct ttgtgtctga agaaacagga acgttgatag ttaatagtgt tttgcttttc      60

```

ttagcggttcg tagtcttcct tcttgtcaca cttgccattt taactgcgct tcgtctatgc 120
gcttactggt gcaatatcgt aaacgtgtcg cttgttaaac caacggttta cgtatactcg 180
cgagttaaaa acctgaattc ttcagaaggt gttcctgatc tgctagtcta a 231

<210> 43
<211> 231
<212> DNA
<213> Artificial Sequence

<220>
<223> Uniform optimization of E protein

<400> 43
atgtacagct tcgtgagcga ggagaccggc accctgatcg tgaacagcgt gctgctgttc 60
ctggccttcg tgggtgttcct gctggtgacc ctggccatcc tgaccgccct gcggctgtgc 120
gcctactgct gcaacatcgt gaacgtgagc ctggtgaagc ccaccgtgta cgtgtacagc 180
cgggtgaaga acctgaacag cagcaggggc gtgcccagacc tgctggtgtg a 231

<210> 44
<211> 3588
<212> DNA
<213> Artificial Sequence

<220>
<223> Minimal optimization of soluble S protein

<400> 44
atgtttatct tcttctgtgt tctgacactg acaagcggca gtgacctgga tagatgcaca 60
acgtttgacg acgtgcaggc ccccaactac acccagcata catccagcat gaggggcgtt 120
tactaccccg atgagatctt tagaagtgat actctgtatc tgactcagga cctgtttctg 180
cccttctatt ctaacgttac tggcttccat acaatcaacc acaccttcgg caaccccgta 240
atacccttta aggatggcat ctactttgcc gccaccgaga agtctaacgt agtgagaggc 300
tgggtgttcg gcagtactat gaacaacaag tctcagtctg tgataataat caacaactcc 360
actaacgtcg tcatcagagc ctgtaacttc gagctgtgcg ataaccctt cttcgccgtt 420
tcgaagccca tgggcactca gaccatata atgatctttg ataacgcctt caactgcacc 480
tttgagtata tctctgatgc cttcagtctg gatgtgtccg agaagtcagg caacttcaag 540
catctgagag agtttgtgtt caagaacaag gatggctttc tgtacgtcta caagggtac 600
cagcccatag atgtggtacg tgacctgcc agcggcttca acactctgaa gcccatattc 660
aagctgcccc tgggcataaa cattaccaac tttagagcca ttctgacggc cttctcccc 720
gccagcata tctggggcac aagtgccgcc gcctacttcg tgggctacct gaagcccaca 780

acttttatgc tgaagtacga cgagaacggc accataacag atgccgtgga ctgttctcag	840
aacccccctgg ccgagctgaa gtgctcagtt aagagttttg agatagataa gggcatctat	900
cagacaagca acttccgcgt ggtccccagc ggcgatgtgg tgaggtttcc caacattacc	960
aacctgtgcc ccttcggcga ggtattcaac gccacaaagt tcccctccgt ttacgcctgg	1020
gagaggaaga agatttcaaa ctgcgtggcc gactactcgg tgctgtataa ctctactttc	1080
ttcagtacct ttaagtgcta cggcgtgtct gccacaaagc tgaacgatct gtgctttagc	1140
aacgtgtatg ccgatagctt cgtcgtcaag ggcgacgacg tcagacagat cgtccccggc	1200
cagacaggcg tcatcgccga ctacaactac aagctgcccg acgatttcat gggctgcgtg	1260
ctggcctgga acacgaggaa catagatgcc accagcactg gcaactacaa ctacaagtac	1320
agatatctgc ggcacggcaa gctgaggccc ttogagagag acatctctaa cgttcccttt	1380
tccccgatg gcaagccctg cactcccccc gccctgaact gctactggcc cctgaacgac	1440
tatggcttct acaccacaac tggcatcggc tatcagccct accgcgtagt cgtgctgtcg	1500
ttcgagctgc tgaacgcccc cgccacagtc tgcggcccca agctgtccac tgacctgatt	1560
aagaaccagt gtgtgaactt caactttaac ggccctgactg gcaccggcgt gctgacaccc	1620
agcagcaagc ggttccagcc cttccagcag tttggcagag acgtgtctga tttcacagat	1680
tccgtgagag atcccaagac ttccgagata ctggatatca gtccctgctc cttcggcggc	1740
gtgtcagtta ttacaccggc cactaacgcc tegtccgagg tagccgttct gtatcaggac	1800
gtgaactgca ctgatgtgag tacagccatc cacgccgacc agctgacccc cgcctggcgg	1860
atztatagta cgggcaacaa cgtctttcag actcaggccg gctgcctgat cggcgccgag	1920
catgtagata cgtcttatga gtgcgacatc cccatcggcg ccggcatctg cgccagctat	1980
cacaccgttt ctctgctgcg aagtacttct cagaagtota tagtggccta caccatgtct	2040
ctgggcgccg atagctctat cgcctatagc aacaacacta tagccatccc cacaaacttc	2100
tctatttcta tcactacaga ggtgatgccc gtctccatgg ccaagaccag cgttgattgc	2160
aacatgtaca tctgcggcga tagtacagag tgcgccaacc tgctgctgca gtatggcagc	2220
ttctgcaccc agctgaacag agccctgtct ggcacgcgcg ccgagcagga taggaacaca	2280
agagaggttt tcgcccaggt taagcagatg tacaagactc ccactctgaa gtactttggc	2340
ggctttaact tttctcagat tctgcccgat cccctgaagc ccactaagag gagtttcata	2400
gaggacctgc tgttcaacaa ggtgactctg gccgacgcgc gctttatgaa gcagtacggc	2460
gagtgcctgg gcgatatcaa cgccagagac ctgatctgtg ccagagaagtt taacggcctg	2520
acagtactgc cccccctgct gactgatgac atgattgccg cctatacggc cgccctgggtg	2580
tctggcactg ccaccgccgg ctggaccttt ggcgcggcg ccgccctgca gatacccttt	2640

gccatgcaga tggcctaccg attcaacggc ataggcgtaa ccagaaacgt tctgtatgag 2700
aaccagaagc agatagccaa ccagttcaac aaggccatct ctcagattca ggagtctctg 2760
accactacat ctactgccct gggcaagctg caggacgtag tgaaccagaa cgcccaggcc 2820
ctgaacaccc tgggttaagca gctgtcaagt aacttcggcg ccatctctag cgttctgaac 2880
gatatactga gtcggctgga taagggtggag gccgaggtgc agattgacag actgatcaca 2940
ggcagactgc agtctctgca gacatatgtt actcagcagc tgataagggc cgccgagatt 3000
agagccagtg ccaacctggc cgccactaag atgtccgagt gcgtcctggg ccagagtaag 3060
agggtagact tttgtggcaa gggctatcac ctgatgtcct tccccaggc cgccccccac 3120
ggcgctcgtgt ttctgcatgt cacttatgtt ccctcacagg agaggaactt cacgaccgcc 3180
cccgccatct gccacgaggg caaggcctat ttccccaggg agggcgctctt cgtattcaac 3240
ggcacgagtt ggttcatcac ccagcgaaac ttcttttcgc ccagataat tacaacggac 3300
aacacttttg taagtggcaa ctgcatgtc gtcacggca taatcaaca cacggtttac 3360
gacccccctgc agcccgagct ggattcattc aaggaggagc tggacaagta cttcaagaac 3420
catactagcc ccgacgttga tctgggcgac ataagcggca tcaacgccag tgtagtcaac 3480
atacagaagg agatcgatag actgaacgag gtggccaaga acctgaacga gtctctgata 3540
gacctgcagg agctgggcaa gtacgagcag tacatcaagt ggccctgg 3588

<210> 45

<211> 2049

<212> DNA

<213> Artificial Sequence

<220>

<223> Minimal optimization of soluble S1 protein

<400> 45

atgttcatct tcctgctgtt tctgacactg acttctggct cagatctgga tagatgcact 60
acctttgacg atgtacaggc cccaactac actcagcaca catcgtccat gcgaggcgtg 120
tattaccccg acgagatctt cagaagtgac actctgtacc tgacacagga cctgttcctg 180
cccttttact ctaacgtgac tggctttcac actatcaacc ataccttcgg caaccccgta 240
atccccttca aggatggcat ctattttgcc gccaccgaga agtccaacgt ggtgaggggc 300
tggggtcttcg gcagtacgat gaacaacaag tctcagtcctg tgataatcat aaacaacagt 360
actaacgtgg ttataagagc ctgcaacttc gagctgtgcg acaaccctt cttcgccgtg 420
tccaagccca tgggacacaca gaccacacc atgatattcg acaacgcctt taactgtact 480
ttcgagtata taagcgatgc cttcagtcctg gatgtttctg agaagtcagg caactttaag 540

catctgagag agttcgtatt caagaacaag gacggctttc tgtatgttta taagggctac 600
cagcccatag atgtcgtgcg ggatctgccc agcggcttca acacactgaa gcccatTTTT 660
aagctgcccc tgggcatcaa cataaccaac tttagagcca tcctgactgc ctttagcccc 720
gcccaggata tatggggcac tagcgccgcc gcctatttcg tcggctacct gaagcccacc 780
acattcatgc tgaagtacga tgagaacggc acaattacgg atgccgtaga ttgcagtcag 840
aaccctctgg ccgagctgaa gtgcagtgtg aagtctttcg agatcgacaa gggcatatac 900
cagacttcta actttcgggt ggttcccagc ggcgacgttg ttaggtttcc caacatcacc 960
aacctgtgcc ccttcggcga ggtgtttaac gccacaaagt tcccctccgt atatgcctgg 1020
gagaggaaga agatttcgaa ctgcgtggcc gactatagcg tcctgtacaa ctctacattc 1080
ttttctacat tcaagtgcta cggcgtcagt gccactaagc tgaacgacct gtgcttcagc 1140
aacgtgtatg ccgactcatt tgtagttaag ggcgatgatg tgagacagat tgcccccggc 1200
cagacaggcg tgatcgccga ttataactat aagctgcccg acgatttcatt gggctgcgtt 1260
ctggcctgga acacaaggaa catcgatgcc actagcactg gcaactacaa ctacaagtac 1320
aggtatctga gacacggcaa gctgaggccc ttcgagcgag atatcagtaa cgtacccttc 1380
agtcccgacg gcaagccctg cactcccccc gccctgaact gctattggcc cctgaacgac 1440
tacggctttt ataccactac aggcatcggc taccagccct acagggttgt ggtgctgagc 1500
ttcgagctgc tgaacgcccc cgccactggt tgcgggccca agctgtcaac ggatctgac 1560
aagaaccagt gcgtaaactt taactttaac ggctgacag gcacaggcgt cctgactccc 1620
tctagtaaga gattccagcc ctttcagcag ttcggccgcg acgtcagcga ttttacggat 1680
agtgtgagag atcccaagac cagcgagatc ctggacatta gtccctgttc tttcggcggc 1740
gtgtctgtca taacgcccgg cacgaacgcc tcttctgagg tcgccgttct gtaccaggac 1800
gtcaactgta cagacgtctc cacagccata cacgccgatc agctgactcc cgctggaga 1860
atttactcta ccggcaacaa cgtcttcag acccaggccg gctgcctgat cggcgccgag 1920
catgtggata cttcctacga gtgcgacata cccatcggcg ccggcatttg cgctcgtac 1980
cataccgtgt ctctgctgag atctacctct cagaagagta tcgttgcta cactatgtcc 2040
ctgggcgcc 2049

<210> 46

<211> 1539

<212> DNA

<213> Artificial Sequence

<220>

<223> Minimal optimization of soluble S2 protein

<400> 46
gatagcagca tagcctactc aaacaacacg atcgccatcc ccacaaactt ttccatttcc 60
ataactaccg aggtgatgcc cgtgagcatg gccaaagacat cggtagattg caacatgtac 120
atctgtggcg attctacaga gtgtgccaac ctgctgctgc agtacggctc tttctgcacg 180
cagctgaaca gggccctgtc tggcatcgcc gccgagcagg atcggaacac acgggaggtt 240
ttcgcccagg taaagcagat gtataagacg cccactctga agtacttcgg cggcttcaac 300
ttctctcaga tactgcccga cccctgaag cccactaaga ggtcttttat cgaggatctg 360
ctgttcaaca aggttaccct ggccgatgcc ggctttatga agcagtatgg cgagtgcctg 420
ggcgacatca acgccagaga tctgatatgc gcccagaagt tcaacggcct gactgtgctg 480
ccccccctgc tgactgacga catgatcgcc gcctataaccg ccgccctggg gagtggcaca 540
gccactgccg gctggacatt cggcgccggc gccgccctgc agatcccctt cgccatgcag 600
atggcctaca gatttaacgg cattggcgtc actcagaacg tcctgtatga gaaccagaag 660
cagatcgcca accagtttaa caaggccata agccagatcc aggagtcact gacaacgaca 720
agtaccgccc tgggcaagct gcaggatgta gtgaaccaga acgcccaggc cctgaacact 780
ctggttaagc agctgtctag caacttcggc gccatcagta gtgttctgaa cgatattctg 840
tctaggctgg acaaggctga ggccgaggtg cagattgatc gcctgattac cggcagactg 900
cagagtctgc agacttatgt aactcagcag ctgatcagag ccgccgagat tcgagcctcc 960
gccaacctgg ccgccacaaa gatgtctgag tgcgtcctgg gccagagtaa gaggggttgac 1020
ttctgcggca agggctatca tctgatgtct tttccccagg ccgcccccca cggcgtcgtg 1080
ttcctgcacg taacttacgt gccagtcag gagagaaact ttaccactgc ccccgccatc 1140
tgccacgagg gcaaggccta cttccccaga gagggcgtgt ttgtgttcaa cggcacatct 1200
tggttcatca ccagaggaa ctttttcagc cccagatca taacaactga caacactttc 1260
gtttcgggca actgcgacgt agtgatcggc ataataaaca acaccgtgta cgatcccctg 1320
cagcccagac tggacagctt taaggaggag ctggacaagt actttaagaa ccatacctca 1380
cccgatgtgg acctgggcca cttttctggc ataaacgcct ccgtcgtcaa catccagaag 1440
gagatagata gactgaacga ggttgccaag aacctgaacg agtccctgat cgatctgcag 1500
gagctgggca agtacgagca gtatataaag tggccctgg 1539

<210> 47

<211> 1620

<212> DNA

<213> Artificial Sequence

<220>

<223> Minimal optimization of TPA-S protein

<400> 47
atggatgcc a tgaagcgagg cctgtgttgc gtactgctgc tgtgcggcgc cgtgtttgtg 60
agccccagcg cccggggcag tggcgacagc agcatcgctt attcgaacaa cactattgcc 120
atacccacaa acttctctat atctataact acggaggtga tgcccgtgtc tatggccaag 180
actagtgtag actgcaacat gtacatctgc ggcgactcta ctgagtgcgc caacctgctg 240
ctgcagtatg gctctttctg caccagctg aacagagccc tgagtggcat cgccgccgag 300
caggaccgga acacaagaga ggttttcgcc caggtaaagc agatgtacaa gacccccact 360
ctgaagtatt ttggcggctt caacttctct cagatcctgc ccgatcccct gaagcccacc 420
aagaggtctt tcatcgagga cctgctgttc aacaaggcca ctctggccga tgccggcttc 480
atgaagcagt acggcgagtg cctggggcgac attaacgccc gcgacctgat ctgtgccag 540
aagtttaacg gcctgacggc cctgcccccc ctgctgacag atgatatgat cgccgcctac 600
actgccgccc tggctctctg caccgccacc gccggctgga ctttcggcgc cggcgccgcc 660
ctgcagatcc ccttcgccat gcagatggcc tatagattta acggcatagg cgtaactcag 720
aacgtcctgt acgagaacca gaagcagatc gccaacagc ttaacaaggc catctcccag 780
attcaggaga gcctgacaac cactagcact gccctgggca agctgcagga cgtggtgaac 840
cagaacgccc aggcctgaa cacttggtt aagcagctga gttctaactt tggcgccata 900
tcctcgggtgc tgaacgacat actgtcaagg ctggacaagg tcgaggccga ggttcagata 960
gatagactga tcacaggcag actgcagagc ctgcagacct acgttacaca gcagctgac 1020
agagccgccg agatcagagc ctacagcaac ctggccgcc cgaagatgtc tgagtgcgtc 1080
ctgggcccagt ctaagagagt cgatttctgc ggcaagggt accacctgat gagtttccc 1140
caggccgccc cccatggcgt tgtattcctg catgtgacat atgttccctc ccaggagagg 1200
aactttacca cggccccgc catctgccac gagggcaagg cctacttccc cagagagggc 1260
gtgttcgttt ttaacggcac tagctggttt attaccaga ggaacttctt ctccccccag 1320
attataacaa cagataaac tttcgtgtcc ggcaactgcg atgttgtgat aggcattcatt 1380
aacaacacag tgtacgatcc cctgcagccc gagctggata gttttaagga ggagctggac 1440
aagtatttta agaaccacac tttcccgat gtagacctgg gcgatatcag tggcataaac 1500
gccagtgtcg tgaacataca gaaggagatc gataggctga acgaggtggc caagaacctg 1560
aacgagtcac tgatcgatct gcaggagctg ggcaagtacg agcagtatat taagtggccc 1620

<210> 48
<211> 231
<212> DNA
<213> Artificial Sequence

<220>

<223> Minimal optimization of E protein

<400> 48

atgtatagtt ttgtgagtga ggagacgggc accctgattg tcaactcagt gctgctgttc 60
ctggcctttg ttgtcttcct gctggtaact ctggccatcc tgactgccct gagactgtgc 120
gcctactgct gcaacatcgt gaacgtctct ctggtaaagc ccacagtta cgtgtattct 180
aggggtgaaga acctgaactc cagcgagggc gttcccgatc tgctggtatg a 231

<210> 49

<211> 1620

<212> DNA

<213> Artificial Sequence

<220>

<223> Minimal optimization of TPA-S2 protein

<400> 49

atggatgcca tgaagcgagg cctgtgttgc gtactgctgc tgtgcggcgc cgtgtttgtg 60
agccccagcg cccggggcag tggcgacagc agcatcgcct attcgaacaa cactattgcc 120
atacccacaa acttctctat atctataact acggaggtga tgcccgtgtc tatggccaag 180
actagtgtag actgcaacat gtacatctgc ggcgactcta ctgagtgcgc caacctgctg 240
ctgcagtatg gctctttctg caccagctg aacagagccc tgagtggcat cgccgccgag 300
caggaccgga acacaagaga ggttttcgcc caggtaaagc agatgtacaa gacccccact 360
ctgaagtatt ttggcggtt caacttctct cagatcctgc ccgatcccct gaagcccacc 420
aagaggtctt tcatcgagga cctgctgttc aacaaggcca ctctggccga tgccggcttc 480
atgaagcagt acggcgagtg cctgggacgc attaacgccc gcgacctgat ctgtgccag 540
aagtttaacg gcctgacggt cctgcccccc ctgctgacag atgatatgat cgccgcctac 600
actgccgccc tgggtctctg caccgccacc gccggctgga ctttcggcgc cggcgccgcc 660
ctgcagatcc ccttcgccat gcagatggcc tatagattta acggcatagg cgtaactcag 720
aacgtcctgt acgagaacca gaagcagatc gccaaccaagg ttaacaaggc catctcccag 780
attcaggaga gcctgacaac cactagcact gccctgggca agctgcagga cgtgggtgaac 840
cagaacgccc aggcctgaa cactctggtt aagcagctga gttctaactt tggcgccata 900
tcctcggtgc tgaacgacat actgtcaagg ctggacaagg tcgaggccga ggttcagata 960
gatagactga tcacaggcag actgcagagc ctgcagacct acgttacaca gcagctgac 1020
agagccgccg agatcagagc ctcagccaac ctggccgcca cgaagatgtc tgagtgcgtc 1080
ctggggccagt ctaagagagt cgatttctgc ggcaagggt accacctgat gagtttcccc 1140

cagggccgccc cccatggcgt tgtattcctg catgtgacat atgttccctc ccaggagagg 1200
 aactttacca cgccccccgc catctgccac gagggcaagg cctacttccc cagagagggc 1260
 gtgttcgttt ttaacggcac tagctggttt attaccaga ggaacttctt ctccccccag 1320
 attataaaa cagataaac tttcgtgtcc ggcaactgcg atgttgtgat aggcattcatt 1380
 aacaacacag tgtacgatcc cctgcagccc gagctggata gttttaagga ggagctggac 1440
 aagtatttta agaaccacac ttccccgat gtagacctgg gcgatatcag tggcataaac 1500
 gccagtgtcg tgaacataca gaaggagatc gataggctga acgaggtggc caagaacctg 1560
 aacgagtcac tgatcgatct gcaggagctg ggcaagtacg agcagtatat taagtggccc 1620

<210> 50
 <211> 2052
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Sequence contain in VR9208

<400> 50
 atggttatct ttctgctgtt cctcaccctc accagcggca gcgatctgga taggtgcacc 60
 accttcgacg acgtgcaggc ccccaactac acccagcaca ccagcagcat gaggggcgtg 120
 tactaccccg acgagatttt cagaagcgac accctgtacc tcaccagga cctgttcctg 180
 cccttctaca gcaacgtgac cggcttccac accatcaacc acaccttcgg caaccccggtg 240
 atccctttca aggacggcat ctacttcgcc gccaccgaga agagcaatgt ggtgcggggc 300
 tgggtgttcg gcagcaccat gaacaacaag agccagagcg tgatcatcat caacaacagc 360
 accaacgtgg tgatccgggc ctgcaatttc gagctgtgcg acaacccttt cttcgcggtg 420
 tccaaaccta tgggcaccca gaccacacc atgatcttcg acaacgcctt caactgcacc 480
 ttcgagtaca tcagcgacgc cttcagcctg gatgtgagcg agaagagcgg caacttcaag 540
 cacctgcggg agttcgtgtt caagaacaag gacggcttcc tgtacgtgta caagggtac 600
 cagcccatcg acgtggtgag agacctgcc agcggcttca acacctgaa gccatcttc 660
 aagctgcccc tgggcatcaa catcaccaac ttccgggcca tcctaccgc ctttagcoct 720
 gccaggata tctggggcac cagcgccgct gcctacttcg tgggctacct gaagcctacc 780
 accttcatgc tgaagtacga cgagaacggc accatcaccg atgccgtgga ctgcagccag 840
 aacccccctg ccgagctgaa gtgcagcgtg aagagcttcg agatcgaaa gggcatctac 900
 cagaccagca acttcagagt ggtgcctagc ggcgatgtgg tgaggttccc caatatcacc 960
 aacctgtgcc ccttcggcga ggtgttcaac gccaccaagt tccctagcgt gtacgcctgg 1020
 gagcggaaga agatcagcaa ctgcgtggcc gattacagcg tgctgtacaa ctccaccttc 1080

ttcagcacct tcaagtgcta cggcgtgagc gccaccaagc tgaacgacct gtgcttcagc 1140
aacgtgtacg ccgactcatt cgtggtgaag ggcgacgacg tgagacagat cgcccctggc 1200
cagaccggcg tgatcgccga ctacaactac aagcttcccg acgacttcat gggctgctg 1260
ctggcctgga acaccagaaa catcgacgcc acctccaccg gcaactacaa ttacaagtac 1320
cgctacctga ggcacggcaa gctgagaccc ttcgagcggg acatctccaa cgtgcccttc 1380
agccccgacg gcaagccctg cccccccct gccctgaact gctactggcc cctgaacgac 1440
tacggcttct acaccaccac cggcatcggc tatcagccct acagagtggg ggtgctgagc 1500
ttcagactgc tgaacgcccc tgccaccgtg tgcgcccca agctgagcac cgacctcatc 1560
aagaaccagt gcgtgaactt caacttcaac ggctcaccg gcaccggcgt gtcaccccc 1620
agcagcaaga gattccagcc cttccagcag ttcggcaggg acgtgagcga tttcaccgac 1680
agcgtgaggg atcctaagac cagcgagatc ctggacatca gcccttgag cttcggcggc 1740
gtgtccgtga tcacccccgg caccaacgcc agcagcgagg tggccgtgct gtaccaggac 1800
gtgaactgca ccgacgtgag caccgccatc cacgccgacc agctcacccc cgctggaga 1860
atctacagca ccggcaacaa cgtgttccag acccaggccg gctgcctcat cggcgccgag 1920
cacgtggaca ccagctacga gtgcgacatc cccatcggag ccggcatctg cgccagctac 1980
cacaccgtga gcctgctgag aagcaccagc cagaagagca tcgtggccta caccatgagc 2040
ctgggcgcct ga 2052

<210> 51

<400> 51
000

<210> 52

<400> 52
000

<210> 53

<211> 231

<212> DNA

<213> Artificial Sequence

<220>

<223> Minimal optimization of E protein

<400> 53

atgtatagtt ttgtgagtga ggagacgggc accctgattg tcaactcagt gctgctgttc 60
ctggcctttg ttgtcttct gctggtaaact ctggccatcc tgactgccct gagactgtgc 120

gcctactgct gcaacatcgt gaacgtctct ctggtaaagc ccacagttta cgtgtattct 180
agggtgaaga acctgaactc cagcgagggc gttcccgatc tgctgggatg a 231

<210> 54
<211> 1542
<212> DNA
<213> Artificial Sequence

<220>
<223> Optimized soluble S2 protein with MET

<400> 54
atggatagtt caattgctta ctctaataac accattgcta tacctactaa cttttcaatt 60
agcattacta cagaagtaat gcctgtttct atggctaaaa cctccgtaga ttgtaatatg 120
tacatctgcg gagattctac tgaatgtgct aatttgcttc tccaatatgg tagcttttgc 180
acacaactaa atcgtgcact ctcaggtatt gctgctgaac aggatcgcaa cacacgtgaa 240
gtgttcgctc aagtcaaaca aatgtacaaa accccaactt tgaaatattt tgggtggtttt 300
aattttttcac aaatattacc tgaccctcta aagccaacta agaggtcttt tattgaggac 360
ttgctcttta ataaggtgac actcgctgat gctggcttca tgaagcaata tggcgaatgc 420
ctagggtgata ttaatgctag agatctcatt tgtgcgcaga agttcaatgg acttacagtg 480
ttgccacctc tgctcactga tgatatgatt gctgcctaca ctgctgctct agttagtggg 540
actgccactg ctggatggac atttgggtgct ggcgctgctc ttcaaatacc ttttgctatg 600
caaatggcat ataggttcaa tggcattgga gttacccaaa atgttctcta tgagaaccaa 660
aaacaaatcg ccaaccaatt taacaaggcg attagtcaaa ttcaagaatc acttacaaca 720
acatcaactg cattgggcaa gctgcaagac gttgttaacc agaatgctca agcattaaac 780
acacttggtta aacaacttag ctctaatttt ggtgcaattt caagtgtgct aaatgatatc 840
ctttcgcgac ttgataaagt cgaggcggag gtacaaattg acaggttaat tacaggcaga 900
cttcaaagcc ttcaaacctc tgtaacacaa caactaatca gggctgctga aatcagggct 960
tctgctaatc ttgctgctac taaaatgtct gagtgtgttc ttggacaatc aaaaagagtt 1020
gacttttgtg gaaagggcta ccaccttatg tccttcccac aagcagcccc gcatgggtgtt 1080
gtcttcctac atgtcacgta tgtgccatcc caggagagga acttcaccac agcgccagca 1140
at ttgtcatg aaggcaaagc atacttcctt cgtgaagggtg tttttgtgtt taatggcact 1200
tcttggttta ttacacagag gaacttcttt tctccacaaa taattactac agacaatata 1260
tttgtctcag gaaattgtga tgtcgttatt ggcattatta acaacacagt ttatgaccc 1320
ctgcaacctg agctcgactc attcaaagaa gagctggaca agtacttcaa aaatcatata 1380
tcaccagatg ttgatcttgg cgacatttca ggcattaacg cttctgtcgt caacattcaa 1440

aaagaaattg accgcctcaa tgaggtcgct aaaaatttaa atgaatcact cattgacctt 1500

caagaattgg gaaaatatga gcaatatatt aaatggcctt gg 1542

<210> 55

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> H-2Kd binding peptide

<400> 55

Thr Tyr Gln Arg Thr Arg Ala Leu Val

1 5

<210> 56

<211> 514

<212> PRT

<213> Artificial Sequence

<220>

<223> Optimized S2 protein with MET

<400> 56

Met Asp Ser Ser Ile Ala Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr

1 5 10 15

Asn Phe Ser Ile Ser Ile Thr Thr Glu Val Met Pro Val Ser Met Ala

20 25 30

Lys Thr Ser Val Asp Cys Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu

35 40 45

Cys Ala Asn Leu Leu Leu Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn

50 55 60

Arg Ala Leu Ser Gly Ile Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu

65 70 75 80

Val Phe Ala Gln Val Lys Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr

85 90 95

Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro

100 105 110

Thr Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu

115 120 125

Ala Asp Ala Gly Phe Met Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile
130 135 140

Asn Ala Arg Asp Leu Ile Cys Ala Gln Lys Phe Asn Gly Leu Thr Val
145 150 155 160

Leu Pro Pro Leu Leu Thr Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala
165 170 175

Leu Val Ser Gly Thr Ala Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala
180 185 190

Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala Tyr Arg Phe Asn Gly
195 200 205

Ile Gly Val Thr Gln Asn Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala
210 215 220

Asn Gln Phe Asn Lys Ala Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr
225 230 235 240

Thr Ser Thr Ala Leu Gly Lys Leu Gln Asp Val Val Asn Gln Asn Ala
245 250 255

Gln Ala Leu Asn Thr Leu Val Lys Gln Leu Ser Ser Asn Phe Gly Ala
260 265 270

Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg Leu Asp Lys Val Glu
275 280 285

Ala Glu Val Gln Ile Asp Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu
290 295 300

Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala
305 310 315 320

Ser Ala Asn Leu Ala Ala Thr Lys Met Ser Glu Cys Val Leu Gly Gln
325 330 335

Ser Lys Arg Val Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe
340 345 350

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

Pro Ser Gln Glu Arg Asn Phe Thr Thr Ala Pro Ala Ile Cys His Glu
370 375 380

Gly Lys Ala Tyr Phe Pro Arg Glu Gly Val Phe Val Phe Asn Gly Thr
385 390 395 400

Ser Trp Phe Ile Thr Gln Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr
405 410 415

Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile
420 425 430

Ile Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe
435 440 445

Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val
450 455 460

Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile Gln
465 470 475 480

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

Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr Ile Lys Trp
500 505 510

Pro Trp

<210> 57
<211> 1242
<212> DNA
<213> Artificial Sequence

<220>
<223> Fragment of S protein

<400> 57
gtcgacatgg ttatcttttct gctgttcctc accctcacca gcggcagcga tctggatagg 60
tgcaccacct tcgacgacgt gcaggccccc aactacaccc agcacaccag cagcatgagg 120
ggcgtgtact accccgacga gattttcaga agcgacaccc tgtacctcac ccaggacctg 180
ttcctgcctt tctacagcaa cgtgaccggc ttccacacca tcaaccacac cttcggcaac 240
cccgtgatcc ctttcaagga cggcatctac ttcgccgcca ccgagaagag caatgtggtg 300
cggggctggg tggttcggcag caccatgaac aacaagagcc agagcgtgat catcatcaac 360

```

aacagcacca acgtggtgat ccgggcctgc aatttcgagc tgtgcgacaa ccctttcttc      420
gccgtgtcca aacctatggg caccagacc cacaccatga tcttcgacaa cgccttcaac      480
tgcaccttcg agtacatcag cgacgccttc agcctggatg tgagcgagaa gagcggcaac      540
ttcaagcacc tgcgggagtt cgtgttcaag aacaaggacg gcttcctgta cgtgtacaag      600
ggctaccagc ccatcgacgt ggtgagagac ctgccagcg gcttcaacac cctgaagccc      660
atcttcaagc tgccctggg catcaacatc accaacttcc gggccatcct caccgccttt      720
agccctgccc aggatatctg gggcaccagc gccgctgcct acttcgtggg ctacctgaag      780
cctaccacct tcatgctgaa gtacgacgag aacggcacca tcaccgatgc cgtggactgc      840
agccagaacc ccctggccga gctgaagtgc agcgtgaaga gcttcgagat cgacaagggc      900
atctaccaga ccagcaactt cagagtgggtg cctagcggcg atgtggtgag gttccccaat      960
atcaccaacc tgtgcccctt cggcgagggtg ttcaacgcc ccaagttccc tagcgtgtac     1020
gcctgggagc ggaagaagat cagcaactgc gtggccgatt acagcgtgct gtacaactcc     1080
accttcttca gcaccttcaa gtgctacggc gtgagcgcca ccaagctgaa cgacctgtgc     1140
ttcagcaacg tgtacgccga ctcatctgtg gtgaagggcg acgacgtgag acagatcgcc     1200
cctggccaga ccggcgtgat cgccgactac aactacaagc tt                               1242

```

<210> 58
 <211> 412
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Fragment of S protein

<400> 58

```

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

```

```

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

```

```

His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
35              40              45

```

```

Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
50              55              60

```

```

Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
65              70              75              80

```

Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
85 90 95

Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
100 105 110

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
115 120 125

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

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
145 150 155 160

Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
165 170 175

Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
180 185 190

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

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
210 215 220

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
225 230 235 240

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
245 250 255

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
260 265 270

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
275 280 285

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
290 295 300

Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
305 310 315 320

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser

<400>	59						
aagcttccccg	acgacttcat	gggctgcgctg	ctggcctgga	acaccagaaa	catcgacgcc		60
acctccaccg	gcaactacaa	ttacaagtac	cgctacctga	ggcacggcaa	gctgagaccc		120
ttcgagcggg	acatctccaa	cgtgcccttc	agccccgacg	gcaagccctg	cacccccctt		180
gccctgaact	gctactggcc	cctgaacgac	tacggcttct	acaccaccac	cggcatcggc		240
tatcagccct	acagagtgg	ggtgctgagc	ttcgagctgc	tgaacgcccc	tgccaccgtg		300
tgcggcccca	agctgagcac	cgacctcatc	aagaaccagt	gcgtaactt	caacttcaac		360
ggcctcaccg	gcaccggcgt	gctcaccccc	agcagcaaga	gattccagcc	cttcagcag		420
ttcggcaggg	acgtgagcga	tttcaccgac	agcgtgaggg	atcctaagac	cagcgagatc		480
ctggacatca	gcccttgcag	cttcggcggc	gtgtccgtga	tcacccccgg	caccaacgcc		540
agcagcgagg	tggccgtgct	gtaccaggac	gtgaactgca	ccgacgtgag	caccgccatc		600
cacgccgacc	agctcacccc	cgccctggaga	atctacagca	ccggcaacaa	cgtgttccag		660
accagggccg	gctgcctcat	cggcgccgag	cacgtggaca	ccagctacga	gtgcgacatc		720
cccatcggag	ccggcatctg	cgccagctac	cacaccgtga	gcctgctgag	aagcaccagc		780
cagaagagca	tcgtggccta	caccatgagc	ctggggcgccg	acagcagcat	cgcctacagc		840
aacaacacca	tcgccatccc	caccaacttc	agcatctcca	tcaccaccga	ggtgatgccc		900

```

gtgagcatgg ccaagaccag cgtggattgc aacatgtaca tctgcggcga cagcaccgag      960
tgcgccaacc tgctgctgca gtacggcagc ttctgcaccc agctgaacag agccctgagc      1020
ggcattgccg ccgagcagga cagaaacacc agggaggtgt tcgcccaggt gaagcagatg      1080
tataagaccc ccaccctgaa gtacttcggc gggttcaact tcagccagat cctgcccgat      1140
cctctgaagc ccaccaagcg gagcttcacg gaggacctgc tgttcaacaa ggtgaccctg      1200
gccgacgccg gctttatgaa gcagtacggc gagtgacctg gcgatatcaa cgccagggac      1260
ctcatctgcg ccagaagtt caacggcttg accgtgctgc cccctctgct caccgatgat      1320
atgatcgccg cctatacagc cgccctggtg tcaggcaccg ccaccgccgg ctggaccttt      1380
ggcgccggag ccgccctgca gatccccctc gccatgcaga tggcctaccg gt              1432

```

```

<210> 60
<211> 1118
<212> DNA
<213> Artificial Sequence

```

```

<220>
<223> Fragment of S protein

```

```

<400> 60
accggttcaa tggcatcggc gtgaccaga acgtgctgta cgagaaccag aagcagatcg      60
ccaaccagtt caataaggcc atctcccaga tccaggagag cctcaccacc acaagcaccg      120
ccctgggcaa gctgcaggac gtggtgaacc agaacgcca ggccctgaat accctggtga      180
agcagctgag cagcaacttc ggcgcatca gcagcgtgct gaacgacatc ctgagcaggc      240
tggataaggt ggaggccgag gtgcagatcg acagactcat caccggcaga ctgcagagcc      300
tgcagaccta cgtgaccag cagctcatca gagccgccga gatcagagcc agcgccaatc      360
tggccgccac caagatgagc gagtgcgtgc tgggcccagag caagagagtg gacttctgog      420
gcaagggcta tcacctcatg agcttccttc aggccgctcc ccacggcgtg gtgttcctgc      480
acgtgaccta cgtgcctagc caggagagga atttcaccac cgccccagcc atctgccacg      540
agggcaaggc ctacttcccc agagagggcg tgttcgtgtt taacggcacc agctggttca      600
tcaccagcg gaacttcttc agccccaga tcatcaccac agacaacacc ttcgtgtccg      660
gcaattgcga cgtggtcatc ggcacatca ataacaccgt gtacgacccc ctgcagcccg      720
agctggatag cttcaaggag gagctggaca agtacttcaa gaaccacacc tccccgacg      780
tggacctggg cgacatcagc ggcacatg ccagcgtggt gaacatccag aaggagatcg      840
accggctgaa cgaggtggcc aagaacctga acgagagcct catcgacctg caggagctgg      900
gaaagtacga gcagtacatc aagtggccct ggtacgtgtg gctgggcttc atcgccggcc      960

```

tcacgcgcat cgtgatggtg accatcctgc tgtgctgcat gaccagctgc tgctcctgcc 1020
tgaagggcgc ctgcagctgt ggcagctgct gcaagttcga cgaggacgac tcagagcccg 1080
tgctgaaggg cgtgaagctg cactacacct gaagatct 1118

<210> 61

<211> 3780

<212> DNA

<213> Artificial Sequence

<220>

<223> Mutated S protein

<400> 61

gtcgacatgg ttatctttct gctgttcctc accctcacca gcggcagcga tctggatagg 60
tgcaccacct tcgacgacgt gcaggccccc aactacaccc agcacaccag cagcatgagg 120
ggcgtgtact accccgacga gatcttcaga agcgacaccc tgtacctcac ccaggacctg 180
ttcctgccct tctacagcaa cgtgaccggc ttccacacca tcaaccacac cttcggcaac 240
cccgtgatcc ctttcaagga cggcatctac ttccgcccca ccgagaagag caatgtggtg 300
cggggctggg tgttcggcag caccatgaac aacaagagcc agagcgtgat catcatcaac 360
aacagcacca acgtggtgat ccgggcctgc aatttcgagc tgtgcgacaa ccctttcttc 420
gccgtgtcca aacctatggg caccagacc cacaccatga tcttcgacaa cgcttcaac 480
tgcaccttcg agtacatcag cgacgccttc agcctggatg tgagcgagaa gagcggcaac 540
ttcaagcacc tgcgggagtt cgtgttcaag aacaaggacg gcttcctgta cgtgtacaag 600
ggctaccagc ccatcgacgt ggtgagagac ctgccagcg gcttcaacac cctgaagccc 660
atcttcaagc tgccctggg catcaacatc accaacttcc gggccatcct caccgccttt 720
agccctgccc aggatatctg gggcaccagc gccgtgcct acttcgtggg ctacctgaag 780
cctaccacct tcacgtgaa gtacgacgag aacggcacca tcaccgatgc cgtggactgc 840
agccagaacc ccctggccga gctgaagtgc agcgtgaaga gcttcgagat cgacaagggc 900
atctaccaga ccagcaactt cagagtgggtg cctagcggcg atgtggtgag gttccccaat 960
atcaccaacc tgtgcccctt cggcgagggtg ttcaacgcca ccaagttccc tagcgtgtac 1020
gcctgggagc ggaagaagat cagcaactgc gtggccgatt acagcgtgct gtacaactcc 1080
accttcttca gcaccttcaa gtgctacggc gtgagcgcca ccaagctgaa cgacctgtgc 1140
ttcagcaacg tgtacgccga ctcatctctg gtgaagggcg acgacgtgag acagatcgcc 1200
cctggccaga ccggcgtgat cgccgactac aactacaagc ttcccagcga cttcatgggc 1260
tgctgtctgg cctggaacac cagaaacatc gacgccacct ccaccggcaa ctacaattac 1320
aagtaccgct acctgaggca cggcaagctg agacccttcg agcgggacat ctccaacgtg 1380

cccttcagcc	ccgacggcaa	gccctgcacc	ccccctgccc	tgaactgcta	ctggcccctg	1440
aacgactacg	gcttctacac	caccaccggc	atcggtatc	agccctacag	agtgggtggtg	1500
ctgagcttcg	agctgctgaa	cgccccctgcc	accgtgtgcg	gccccaaagct	gagcaccgac	1560
ctcatcaaga	accagtgcgt	gaacttcaac	ttcaacggcc	tcaccggcac	cggcgtgctc	1620
acccccagca	gcaagagatt	ccagcccttc	cagcagttcg	gcagggacgt	gagcgatttc	1680
accgacagcg	tgagggatcc	taagaccagc	gagatcctgg	acatcagccc	ttgcagcttc	1740
ggcggcgtgt	ccgtgatcac	ccccggcacc	aacgccagca	gcgaggtggc	cgtgctgtac	1800
caggacgtga	actgcaccga	cgtgagcacc	gccatccacg	ccgaccagct	cacccccgcc	1860
tggagaatct	acagcaccgg	caacaacgtg	ttccagaccc	aggccggctg	cctcatcggc	1920
gccgagcacg	tggacaccag	ctacgagtgc	gacatcccca	tcggagccgg	catctgcgcc	1980
agctaccaca	ccgtgagcct	gctgagaagc	accagccaga	agagcatcgt	ggcctacacc	2040
atgagcctgg	gcgccgacag	cagcatcgcc	tacagcaaca	acaccatcgc	catccccacc	2100
aacttcagca	tctccatcac	caccgaggtg	atgcccgctga	gcatggccaa	gaccagcgtg	2160
gattgcaaca	tgtacatctg	cggcgacagc	accgagtgcg	ccaacctgct	gctgcagtac	2220
ggcagcttct	gcaccagct	gaacagagcc	ctgagcggca	ttgccgccga	gcaggacaga	2280
aacaccaggg	aggtgttcgc	ccaggtgaag	cagatgtata	agacccccac	cctgaagtac	2340
ttcggcgggt	tcaacttcag	ccagatcctg	cccgatcctc	tgaagcccac	caagcggagc	2400
ttcatcgagg	acctgctggt	caacaaggtg	accctggccg	acgccggctt	tatgaagcag	2460
tacggcgagt	gcctgggcca	tatcaacgcc	agggaacctca	tctgcgcca	gaagttcaac	2520
ggcttgaccg	tgctgcccc	tctgtcacc	gatgatatga	tcgccgccta	tacagccgcc	2580
ctggtgtcag	gcaccgccac	cgcggctgg	acctttggcg	ccggagccgc	cctgcagatc	2640
cccttcgcca	tgcatatggc	ctaccggttc	aatggcatcg	gcgtgacca	gaacgtgctg	2700
tacgagaacc	agaagcagat	cgccaaccag	ttcaataagg	ccatctccca	gatccaggag	2760
agcctcacca	ccacaagcac	cgccctgggc	aagctgcagg	acgtgggtgaa	ccagaacgcc	2820
caggccctga	ataccctggt	gaagcagctg	agcagcaact	tcggcgccat	cagcagcgtg	2880
ctgaacgaca	tcctgagcag	gctggataag	gtggaggccg	aggtgcagat	cgacagactc	2940
atcaccggca	gactgcagag	cctgcagacc	tacgtgacct	agcagctcat	cagagccgcc	3000
gagatcagag	ccagcgccaa	tctggccgcc	accaagatga	gcgagtgcgt	gctgggccag	3060
agcaagagag	tggacttctg	cggcaagggc	tatcacctca	tgagcttccc	tcaggccgct	3120
ccccacggcg	tggtgttcct	gcacgtgacc	tacgtgccta	gccaggagag	gaatttcacc	3180

accgccccag ccattctgcc cgagggcaag gcctacttcc ccagagaggg cgtgttcgtg 3240
 ttttaacggca ccagctgggt catcacccag cggaacttct tcagccccca gatcatcacc 3300
 acagacaaca ccttcgtgtc cggcaattgc gacgtggtca tcggcatcat caataacacc 3360
 gtgtacgacc ccctgcagcc cgagctggat agcttcaagg aggagctgga caagtacttc 3420
 aagaaccaca cctccccga cgtggacctg ggcgacatca gcggcatcaa tgccagcgtg 3480
 gtgaacatcc agaaggagat cgaccggctg aacgaggtgg ccaagaacct gaacgagagc 3540
 ctcatcgacc tgcaggagct gggaaagtac gagcagtaca tcaagtggcc ctggtacgtg 3600
 tggctgggct tcatcgccgg cctcatcgcc atcgtgatgg tgaccatcct gctgtgctgc 3660
 atgaccagct gctgctcctg cctgaagggc gcctgcagct gtggcagctg ctgcaagttc 3720
 gacgaggacg actcagagcc cgtgctgaag ggcgtgaagc tgcactacac ctgaagatct 3780

<210> 62
 <211> 1255
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Mutated S protein

<400> 62

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

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

His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
35 40 45

Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
50 55 60

Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
65 70 75 80

Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
85 90 95

Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
100 105 110

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
115 120 125

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

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
 145 150 155 160

Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
 165 170 175

Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
 180 185 190

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

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
 210 215 220

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
 225 230 235 240

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
 245 250 255

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
 260 265 270

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
 275 280 285

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
 290 295 300

Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
 305 310 315 320

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
 325 330 335

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
 340 345 350

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

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
 370 375 380

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
 385 390 395 400

Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
 405 410 415

Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
 420 425 430

Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
 435 440 445

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

Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
 465 470 475 480

Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
 485 490 495

Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
 500 505 510

Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
 515 520 525

Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
 530 535 540

Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
 545 550 555 560

Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
 565 570 575

Ser Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
 580 585 590

Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
 595 600 605

Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
 610 615 620

Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
 625 630 635 640

His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
 645 650 655

Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
 660 665 670

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

Tyr Ser Asn Asn Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile
 690 695 700

Thr Thr Glu Val Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys
 705 710 715 720

Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
 725 730 735

Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
 740 745 750

Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys
 755 760 765

Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe
 770 775 780

Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile
 785 790 795 800

Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met
 805 810 815

Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile
 820 825 830

Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
 835 840 845

Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala

850		855		860
Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe				
865		870		875 880
Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn				
	885		890	895
Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala				
	900		905	910
Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly				
	915		920	925
Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu				
	930		935	940
Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn				
	945		950	955 960
Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp				
	965		970	975
Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln				
	980		985	990
Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala				
	995		1000	1005
Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp				
	1010		1015	1020
Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ala Ala				
	1025		1030	1035
Pro His Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ser Gln				
	1040		1045	1050
Glu Arg Asn Phe Thr Thr Ala Pro Ala Ile Cys His Glu Gly Lys				
	1055		1060	1065
Ala Tyr Phe Pro Arg Glu Gly Val Phe Val Phe Asn Gly Thr Ser				
	1070		1075	1080
Trp Phe Ile Thr Gln Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr				
	1085		1090	1095

Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly
 1100 1105 1110
 Ile Ile Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp
 1115 1120 1125
 Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser
 1130 1135 1140
 Pro Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val
 1145 1150 1155
 Val Asn Ile Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala Lys
 1160 1165 1170
 Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr
 1175 1180 1185
 Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Val Trp Leu Gly Phe Ile
 1190 1195 1200
 Ala Gly Leu Ile Ala Ile Val Met Val Thr Ile Leu Leu Cys Cys
 1205 1210 1215
 Met Thr Ser Cys Cys Ser Cys Leu Lys Gly Ala Cys Ser Cys Gly
 1220 1225 1230
 Ser Cys Cys Lys Phe Asp Glu Asp Asp Ser Glu Pro Val Leu Lys
 1235 1240 1245
 Gly Val Lys Leu His Tyr Thr
 1250 1255

<210> 63

<211> 1281

<212> DNA

<213> Artificial Sequence

<220>

<223> Mutated N protein

<400> 63

gtcgacatga gcgacaacgg cccccagagc aaccagagaa gcgccccag aatcaccttt 60

ggcgggcccta ccgacagcac cgacaacaac cagaacggcg gcagaaacgg cgccagaccc 120

aagcagagga gaccccaggg cctgccaac aacaccgcca gctgggtcac cgccctcacc 180

cagcacggca aggaggagct gagattcccc agaggccagg gcgtgcccac caataccaac 240
 agcggcccag acgatcagat cggctactac cggagggcca ccagaagagt gagaggcggc 300
 gacggcaaga tgaaggagct gagcccccg tggtacttct actacctggg caccggccct 360
 gagggcagcc tgccctacgg cgccaacaag gagggcatcg tgtgggtggc caccgagggc 420
 gccctgaata cccccaagga ccacatcggc accaggaacc ccaacaacaa tgccgccacc 480
 gtgctgcagc tgccccaggg caccaccctg cccaagggct tctacgccga gggcagcaga 540
 ggcggcagcc agggcagcag cagaagcagc agcaggagca ggggcaacag cagaaatagc 600
 acccccggca gcagcagagg aaattcacc gccagaatgg ccagcggcgg aggcgagacc 660
 gccctggccc tgctgctcct ggacaggctg aatcagctgg agagcaaggc gagcggcaag 720
 ggccagcaac agcagggaca gaccgtgacc aagaagtctg ccgccgaggc cagcaagaag 780
 cccaggcaga agagaaccgc caccaagcag tacaatgtga cccaggcctt cggcagaaga 840
 ggccccgagc agaccaggg caatttcggc gaccaggacc tcatcagaca gggcaccgac 900
 tacaagcact ggcctcagat cgcccagttc gcccacagcg ccagcgcctt cttcggcatg 960
 agccggatcg gcatggaggc gacccccagc ggcacctggc tcacctacca cggcgccatc 1020
 aagctggagc acaaggaccc ccagttcaag gacaacgtga tcctgctgaa caagcacatc 1080
 gacgcctaca agaccttccc accaccgag cccaagaagg acaagaagaa gaaaaccgac 1140
 gagggccagc ccctgcccc aagacagaag aagcagccca ccgtgaccct gctgctgcc 1200
 gccgacatgg acgacttcag ccgccagctg cagaatagca tgagcggcgc ctctgccgat 1260
 tcaaccaggc cctgaagatc t 1281

<210> 64

<211> 1542

<212> DNA

<213> Artificial Sequence

<220>

<223> Uniform optimization of S2 protein with MET

<400> 64

atggacagca gcatcgccta cagcaacaac accatcgcca tccccaccaa cttcagcatc 60
 agcatcacca ccgaggtgat gcccgtgagc atggccaaga ccagcgtgga ctgcaacatg 120
 tacatctgcg gcgacagcac cgagtgcgcc aacctgctgc tgcagtacgg cagcttctgc 180
 acccagctga accgggcccct gagcggcatc gccgccgagc aggaccggaa caccggggag 240
 gtgttcgccc aggtgaagca gatgtacaag accccacccc tgaagtactt cggcggcttc 300
 aacttcagcc agatcctgcc cgacccccctg aagcccacca agcggagctt catcaggagc 360
 ctgctgttca acaagggtgac cctggccgac gccggcttca tgaagcagta cggcgagtgc 420

ctgggcgaca tcaacgcccg ggacctgacg tgcgcccaga agttcaacgg cctgaccgtg	480
ctgccccccc tgctgaccga cgacatgacg gccgcctaca ccgccgccct ggtgagcggc	540
accgccaccg ccggctggac cttcggcgcc ggccgccccc tgcagatccc cttcgccatg	600
cagatggcct accggttcaa cggcatcggc gtgaccgaga acgtgctgta cgagaaccag	660
aagcagatcg ccaaccagtt caacaaggcc atcagccaga tccaggagag cctgaccacc	720
accagcaccg ccctgggcaa gctgcaggac gtggtgaacc agaacgcccc ggccctgaac	780
accctggtga agcagctgag cagcaacttc ggccgccatca gcagcgtgct gaacgacatc	840
ctgagccggc tggacaaggt ggaggccgag gtgcagatcg accggctgat caccggccgg	900
ctgcagagcc tgcagacctg cgtgaccgag cagctgatcc gggccgccga gatccgggcc	960
agcgccaacc tggccgccac caagatgagc gagtgcgtgc tgggccagag caagcgggtg	1020
gacttctgcg gcaagggcta ccacctgatg agcttcccc aggccgcccc ccacggcgtg	1080
gtgttcctgc acgtgacctg cgtgcccagc caggagcggg acttcaccac cgccccgcc	1140
atctgccacg agggcaaggc ctacttcccc cgggagggcg tgttcgtggt caacggcacc	1200
agctggttca tcaccagcg gaacttcttc agccccaga tcatcaccac cgacaacacc	1260
ttcgtgagcg gcaactgcga cgtggtgacg ggcatcatca acaacaccgt gtacgacccc	1320
ctgcagcccc agctggacag cttcaaggag gagctggaga agtacttcaa gaaccacacc	1380
agccccgacg tggacctggg cgacatcagc ggcatcaacg ccagcgtggt gaacatccag	1440
aaggagatcg accggctgaa cgaggtggcc aagaacctga acgagagcct gatcgacctg	1500
caggagctgg gcaagtacga gcagtacatc aagtggccct gg	1542

<210> 65

<211> 1542

<212> DNA

<213> Artificial Sequence

<220>

<223> Fully optimized S2 protein with MET

<400> 65

atggacagtt caatcgcta ttcgaacaac actatagcaa tcccaacaaa tttttcaatt	60
tctataacaa cagaggtgat gccagtgtcc atggcaaaga ctagcgtaga ctgcaatatg	120
tacatctgcg gagattctac agaatgtgca aacttgctgc tacagtatgg atcgttctgt	180
accagctca accggcgct gagcggcatt gctgccgaac aggatcgcaa tacgagagag	240
gtgtttgctc aagtgaacaa aatgtataag accccaacat tgaaatactt cgggtggattc	300
aatttcagtc agattctgcc agaccactc aaaccaccca agaggagctt tattgaagat	360

cttctgttca acaaagttac cttggccgac gctgggttta tgaagcaata cggtagtggtc 420
ctgggcgaca ttaacgcacg agacctgacg tgcgcccaga agtttaacgg gctcacgggtt 480
ttaccgccac tgctgactga tgatatgatt gccgcttaca ctgcggccct tgtgagtgggt 540
accgcaactg ctggctggac gtttggcgct ggggcggcct tacagatccc ttttgccatg 600
cagatggcct acaggttcaa tggaattgggt gtcactcaga atgtcctgta cgagaaccag 660
aaacagatcg ccaaccagtt caataaagct atttcacaga ttcaggaatc acttaccaca 720
acttccacgg cactcggtaa actgcaggac gtggtgaatc agaacgctca ggcactaaat 780
acactcgtca agcaactgag ttccaatttc ggggccatat ctagcgtatt gaacgacatc 840
ctcagtcggc tcgacaaagt ggaggccgaa gtccaaatag accgtcttat cacaggcaga 900
ctacagtcac tgacagaccta cgttaccag cagttgatcc gcgccgctga gatacgagcc 960
tccgccaatc tggccgctac caaaatgtct gagtgtgtgc tcggacaaaag taagcgggtg 1020
gattttttgcg gcaagggcta tcacctcatg tccttccctc aagcagcacc ccacggagtc 1080
gtttttctgc atgtgacata cgtgcctagc caggagagaa actttaccac tgcgcctgcc 1140
atgtgtcatg aaggcaaagc ttattttccc cgcgaggggg tgttcgtttt caacggaact 1200
agctgggttta tcacacaaaag gaattttctt tccccccaga tcatcaccac cgacaacacc 1260
tttgtctctg gaaactgtga cgtcggtata ggcacatca ataatacagt atacgatccc 1320
ctgcagcccg aacttgactc tttcaaggag gaactagata agtacttcaa gaatcacacc 1380
agcccggatg tagatttagg ggatattagc gggattaacg catccgtgggt caacatccaa 1440
aaagagattg acagactgaa cgaagtggcg aagaacctga atgagtcctt gatcgatctt 1500
caggagctgg gcaagtatga acagtatatc aagtggcctt gg 1542

<210> 66

<211> 1542

<212> DNA

<213> Artificial Sequence

<220>

<223> Minimal optimization of S2 protein with MET

<400> 66

atggatagca gcatagccta ctcaaacaac acgatcgcca tccccacaaa cttttccatt 60
tccataacta ccgaggtgat gcccgtagac atggccaaga catcggtaga ttgcaacatg 120
tacatctgtg gcgattctac agagtgtgcc aacctgctgc tgcagtacgg ctctttctgc 180
acgcagctga acagggccct gtctggcatc gccgcccagc aggatcggaa cacacgggag 240
gttttcgccc aggtaaagca gatgtataag acgcccactc tgaagtactt cggcggcttc 300
aacttctctc agatactgcc cgaccccctg aagcccacta agaggtcttt tatcgaggat 360

```

ctgctgttca acaaggttac cctggccgat gccggcttta tgaagcagta tggcgagtgc      420
ctgggcgaca tcaacgccag agatctgata tgcgccaga agttcaacgg cctgactgtg      480
ctgccccccc tgctgactga cgacatgata gccgcctata ccgccgccct ggtgagtggc      540
acagccactg ccggctggac attcggcgcc ggcgcgcgcc tgcagatccc cttcgccatg      600
cagatggcct acagatttaa cggcattggc gtcactcaga acgtcctgta tgagaaccag      660
aagcagatcg ccaaccagtt taacaaggcc ataagccaga tccaggagtc actgacaacg      720
acaagtaccg ccctgggcaa gctgcaggat gtagtgaacc agaacgcca ggccctgaac      780
actctggtta agcagctgtc tagcaacttc ggcgccatca gtagtgttct gaacgatatt      840
ctgtctaggc tggacaaggt cgaggccgag gtgcagattg atcgccgat taccggcaga      900
ctgcagagtc tgcagactta tgtaactcag cagctgatca gagccgccga gattcgagcc      960
tccgccaacc tggccgccac aaagatgtct gagtgcgtcc tgggccagag taagaggggt      1020
gacttctgcy gcaagggcta tcatctgatg tcttttcccc aggccgcccc ccacggcgtc      1080
gtgttcctgc acgtaactta cgtgccagtc caggagagaa actttaccac tgcccccgcc      1140
atctgccacg agggcaaggc ctacttcccc agagagggcg tgtttggtgt caacggcaca      1200
tcttggttca tccccagag gaactttttc agccccaga tcataacaac tgacaacact      1260
ttcgtttcgg gcaactgcga cgtagtgatc ggcataataa acaacaccgt gtacgatccc      1320
ctgcagcccc agctggacag ctttaaggag gagctggaca agtactttaa gaaccatacc      1380
tcacccgatg tggacctggg cgacatttct ggcataaacg cctccgtcgt caacatccag      1440
aaggagatag atagactgaa cgaggttgcc aagaacctga acgagtcctt gatcgatctg      1500
caggagctgg gcaagtacga gcagtatata aagtggccct gg                          1542

```

<210> 67

<211> 3588

<212> DNA

<213> Artificial Sequence

<220>

<223> Standardized optimization of soluble S protein

<400> 67

```

atgttcatct tcctgctgtt cctgaccctg accagcggca gcgacctgga tcgctgcacc      60
accttcgatg acgtgcaggc cccaactac acccagcata ccagcagcat gcgcggcgctg      120
tactaccccg atgagatctt ccgcagcgac accctgtacc tgaccagga cctgttcctg      180
cccttctaca gcaacgtgac cggcttcac accatcaacc ataccttcgg caaccccgctg      240
atccccttca aggacggcat ctacttcgcc gccaccgaga agagcaacgt ggtgcgcggc      300

```

tgggtgttcg	gcagcaccat	gaacaacaag	agccagagcg	tgatcatcat	caacaacagc	360
accaacgtgg	tgatccgcgc	ctgcaacttc	gagctgtgcg	acaacccctt	cttcgccgtg	420
agcaagccca	tgggcaccca	gaccataacc	atgatcttcg	ataacgcctt	caactgcacc	480
ttcgagtaca	tcagcgacgc	cttcagcctg	gacgtgagcg	agaagagcgg	caacttcaag	540
catctgcgcg	agttcgtggt	caagaacaag	gatggcttcc	tgtacgtgta	caagggctac	600
cagcccatcg	acgtggtgcg	cgatctgccc	agcggcttca	acacccctgaa	gcccattctt	660
aagctgcccc	tgggcatcaa	catcaccaac	ttccgcgcca	tcctgaccgc	cttcagcccc	720
gcccaggaca	tctggggcac	cagcgccgcg	gcctacttcg	tgggctacct	gaagcccacc	780
accttcatgc	tgaagtacga	tgagaacggc	accatcacccg	acgccgtgga	ctgcagccag	840
aacccccctg	ccgagctgaa	gtgcagcgtg	aagagcttcg	agatcgataa	gggcatctac	900
cagaccagca	acttccgcgt	ggtgcccagc	ggcgacgtgg	tgcgcttccc	caacatcacc	960
aacctgtgtc	ccttcggcga	ggtgttcaac	gccaccaagt	tccccagcgt	gtacgcctgg	1020
gagcgcaaga	agatcagcaa	ctgcgtggcc	gactacagcg	tgctgtacaa	cagcaccttc	1080
ttcagcacct	tcaagtgcta	cggcgtgagc	gccaccaagc	tgaacgatct	gtgcttcagc	1140
aacgtgtacg	ccgacagctt	cgtggtgaag	ggcgatgatg	tgcgccagat	cgcccccggc	1200
cagaccggcg	tgatcgccga	ttacaactac	aagctgcccc	acgaacttc	gggctgcgtg	1260
ctggcctgga	acacccgcaa	catcgacgcc	accagcacccg	gcaactacaa	ctacaagtac	1320
cgctacctgc	gcatggcaa	gctgcgcccc	ttcgagcgcg	atatcagcaa	cgtgcccttc	1380
agccccgatg	gcaagccctg	cacccccccc	gccctgaact	gttactggcc	cctgaacgac	1440
tacggcttct	acaccaccac	cggcatcggc	taccagccct	accgcgtgg	ggtgctgagc	1500
ttcgagctgc	tgaacgcccc	cgccaccgtg	tgcggcccca	agctgagcac	cgacctgatc	1560
aagaaccagt	gcgtgaactt	caacttcaac	ggcctgaccg	gcaccggcgt	gctgaccccc	1620
agcagcaagc	gcttccagcc	cttccagcag	ttcggccgcg	atgtgagcga	cttcaccgat	1680
agcgtgcgcg	acccaagac	cagcgagatc	ctggatatca	gcccctgcag	cttcggcggc	1740
gtgagcgtga	tcacccccgg	caccaacgcc	agcagcgagg	tggccgtgct	gtaccaggat	1800
gtgaactgta	ccgatgtgag	caccgccatc	cacgccgatc	agctgacccc	cgcctggcgc	1860
atctacagca	ccggcaacaa	cgtgttccag	accaggccg	gctgcctgat	cggcgccgag	1920
catgtggaca	ccagctacga	gtgtgacatc	cccatcggcg	ccggcatctg	tgccagctac	1980
cacaccgtga	gcctgctgcg	cagcaccagc	cagaagagca	tcgtggccta	caccatgagc	2040
ctgggcgccg	atagcagcat	cgcctacagc	aacaacacca	tcgccatccc	caccaacttc	2100
agcatcagca	tcaccaccga	ggtgatgccc	gtgagcatgg	ccaagaccag	cgtggactgc	2160

```

aacatgtaca tctgcggcga tagcaccgag tgcgccaaacc tgctgctgca gtacggcagc 2220
ttctgcaccc agctgaaccg cgccctgagc ggcacgcgcg ccgagcagga tcgcaacacc 2280
cgcgaggtgt tcgcccaggt gaagcagatg tacaagaccc ccaccctgaa gtacttcggc 2340
ggcttcaact tcagccagat cctgcccgat cccctgaagc ccaccaagcg cagcttcac 2400
gaggatctgc tgttcaacaa ggtgaccctg gccgatgcg gcttcatgaa gcagtacggc 2460
gagtgcctgg gcgatatcaa cgcccgcat ctgatctgcg ccagaagtt caacggcctg 2520
accgtgctgc ccccttgct gaccgacgac atgatcgccg cctacaccgc cgccctgggtg 2580
agcggcaccg ccaccgcgg ctggaccttc ggcgcggcg ccgccctgca gatccccttc 2640
gccatgcaga tggcctaccg cttcaacggc atcggcgtga ccagaacgt gctgtacgag 2700
aaccagaagc agatcgccaa ccagttcaac aaggccatca gccagatcca ggagagcctg 2760
accaccacca gcaccgcctt gggcaagctg caggacgtgg tgaaccagaa cgcccaggcc 2820
ctgaacaccc tgggtaagca gctgagcagc aacttcggcg ccatcagcag cgtgctgaac 2880
gacatcctga gccgcctgga taagggtggag gccgaggtgc agatcgatcg cctgatcacc 2940
ggccgcctgc agagcctgca gacctacgtg accagcagc tgatccgcgc cgccgagatc 3000
cgcgccagcg ccaacctggc cgccaccaag atgagcgagt gcgtgctggg ccagagcaag 3060
cgcgtggatt tctgcggcaa gggctaccac ctgatgagct tccccaggc cgccccccat 3120
ggcgtggtgt tcctgcacgt gacctacgtg ccagccagg agcgcaactt caccaccgcc 3180
cccgccatct gccacgaggg caaggcctac ttccccgcg agggcgtggt cgtgttcaac 3240
ggcaccagct ggttcatcac ccagcgcaac ttcttcagcc ccagatcat caccaccgat 3300
aacaccttcg tgagcggcaa ctgcgatgtg gtgatcggca tcatcaacaa caccgtgtac 3360
gatcccctgc agcccgagct ggacagcttc aaggaggagc tggataagta cttcaagaac 3420
cacaccagcc ccgacgtgga tctgggcgat atcagcggca tcaacgccag cgtggtgaac 3480
atccagaagg agatcgatcg cctgaacgag gtggccaaga acctgaacga gagcctgatc 3540
gacctgcagg agctgggcaa gtacgagcag tacatcaagt ggccttg 3588

```

<210> 68

<211> 2049

<212> DNA

<213> Artificial Sequence

<220>

<223> Standardized optimization of soluble S1 protein

<400> 68

```

atgttcatct tcctgctgtt cctgaccctg accagcggca gcgatctgga ccgctgcacc 60

```

accttcgacg atgtgcaggc ccccaactac acccagcaca ccagcagcat gcgcggcgtg 120
tactaccccg atgagatctt ccgcagcgat accctgtacc tgacccagga tctgttcctg 180
cccttctaca gcaacgtgac cggcttccat accatcaacc acaccttcgg caaccccggtg 240
atcccccttca aggatggcat ctacttcgcc gccaccgaga agagcaacgt ggtgcgcggc 300
tggtgttctg gcagcaccat gaacaacaag agccagagcg tgatcatcat caacaacagc 360
accaacgtgg tgatccgcgc ctgcaacttc gagctgtgcg acaacccctt cttcgccgtg 420
agcaagccca tgggcaccca gaccacacc atgatcttcg acaacgcctt caactgcacc 480
ttcgagtaca tcagcgatgc cttcagcctg gacgtgagcg agaagagcgg caacttcaag 540
catctgcgcg agttcgtgtt caagaacaag gatggcttcc tgtacgtgta caagggctac 600
cagcccatcg acgtggtgcg cgacctgcc agcggcttca acacctgaa gcccatcttc 660
aagctgcccc tgggcatcaa catcaccaac ttccgcgcca tcctgaccgc cttcagcccc 720
gcccaggata tctggggcac cagcgccgcc gcctacttcg tgggctacct gaagcccacc 780
accttcatgc tgaagtacga cgagaacggc accatcacccg atgccgtgga ttgcagccag 840
aaccccttgg ccgagctgaa gtgcagcgtg aagagcttcg agatcgataa gggcatctac 900
cagaccagca acttccgcgt ggtgcccagc ggcgacgtgg tgcgcttccc caacatcacc 960
aacctgtgcc ccttcggcga ggtgttcaac gccaccaagt tccccagcgt gtacgcctgg 1020
gagcgcaaga agatcagcaa ctgcgtggcc gattacagcg tgctgtacaa cagcaccttc 1080
ttcagcacct tcaagtgcta cggcgtgagc gccaccaagc tgaacgacct gtgcttcagc 1140
aacgtgtacg ccgacagctt cgtggtgaag ggcgacgacg tgcgccagat cgcggccggc 1200
cagaccggcg tgatcgccga ttacaactac aagctgcccg atgacttcat gggctgcgtg 1260
ctggcctgga acacccgcaa catcgatgcc accagcacccg gcaactacaa ctacaagtac 1320
cgctacctgc gccacggcaa gctgcgcgcc ttcgagcgcg atatcagcaa cgtgcccttc 1380
agccccgatg gcaagccctg ccccccccc gccctgaact gttactggcc cctgaacgat 1440
tacggcttct acaccaccac cggcatcgcc taccagccct accgcgtggg ggtgctgagc 1500
ttcgagctgc tgaacgcccc cgccaccgtg tgcggcccca agctgagcac cgacctgatc 1560
aagaaccagt gcgtgaactt caacttcaac ggcctgaccg gcaccggcgt gctgaccccc 1620
agcagcaagc gcttccagcc cttccagcag ttgggcgcgc acgtgagcga cttcaccgac 1680
agcgtgcgcg atcccaagac cagcgagatc ctggatatca gcccctgcag cttcggcggc 1740
gtgagcgtga tcacccccgg caccaacgcc agcagcgagg tggccgtgct gtaccaggac 1800
gtgaactgca ccgatgtgag caccgccatc cagccgatc agctgacccc cgctggcgc 1860
atctacagca ccggcaacaa cgtgttccag acccaggccg gctgtctgat cggcgccgag 1920

catgtggaca ccagctacga gtgtgatata cccatcggcg ccggcatctg cgccagctac	1980
cataccgtga gcctgctgcg cagcaccagc cagaagagca tcgtggccta caccatgagc	2040
ctgggcgcc	2049

<210> 69

<211> 1623

<212> DNA

<213> Artificial Sequence

<220>

<223> Standardized optimization of TPA-S2 protein

<400> 69

atggatgcca tgaagcgcg cctgtgctgt gtgctgctgc tgtgtggcgc cgtgttcgtg	60
agccccagcg cccgcggcag cggcgatagc agcatcgcct acagcaacaa caccatcgcc	120
atccccacca acttcagcat cagcatcacc accgaggtga tgcccgtgag catggccaag	180
accagcgtgg attgcaacat gtacatctgc ggcgacagca ccgagtgcgc caacctgctg	240
ctgcagtacg gcagcttctg caccagctg aaccgcgccc tgagcggcat cgccgccgag	300
caggaccgca acaccgcga ggtgttcgccc caggtgaagc agatgtacaa gacccccacc	360
ctgaagtact tcggcggtt caacttcagc cagatcctgc ccgacccct gaagcccacc	420
aagcgcagct tcatcgagga tctgctgttc aacaaggtga ccctggccga cgccggcttc	480
atgaagcagt acggcgagt cctgggcgac atcaacgccc gcgacctgat ctgcgcccag	540
aagttcaacg gcctgaccgt gctgcccccc ctgctgaccg atgacatgat cgccgcctac	600
accgccgccc tggtgagcgg caccgccacc gccggctgga ccttcggcgc cggcgccgcc	660
ctgcagatcc ccttcgccat gcagatggcc taccgcttca acggcatcgg cgtgaccag	720
aacgtgctgt acgagaacca gaagcagatc gccaacagct tcaacaaggc catcagccag	780
atccaggaga gcctgaccac caccagcacc gccctgggca agctgcagga tgtggtgaac	840
cagaacgccc aggcctgaa caccctggtg aagcagctga gcagcaactt cggcgccatc	900
agcagcgtgc tgaacgatat cctgagccgc ctggataagg tggaggccga ggtgcagatc	960
gaccgcctga tcaccggccg cctgcagagc ctgcagacct acgtgaccca gcagctgatc	1020
cgcgccgccc agatccgcgc cagcgccaac ctggccgcca ccaagatgag cgagtgcgtg	1080
ctggggccaga gcaagcgcgt ggatttctgc ggcaagggct accacctgat gagcttcccc	1140
caggccgccc cccacggcgt ggtgttcctg catgtgacct acgtgccag ccaggagcgc	1200
aacttcacca ccgccccgc catctgccac gagggcaagg cctacttccc ccgcgagggc	1260
gtgttcgtgt tcaacggcac cagctggttc atcaccagc gcaacttctt cagccccag	1320

atcatcacca	ccgacaacac	cttcgtgagc	ggcaactgcg	acgtggtgat	cggcatcatc	1380
aacaacaccg	tgtacgatcc	cctgcagccc	gagctggata	gcttcaagga	ggagctggac	1440
aagtacttca	agaaccatac	cagccccgat	gtggatctgg	gcgacatcag	cggcatcaac	1500
gccagcgtgg	tgaacatcca	gaaggagatc	gatcgcttga	acgaggtggc	caagaacctg	1560
aacgagagcc	tgatcgatct	gcaggagctg	ggcaagtacg	agcagtacat	caagtggccc	1620
tgg						1623