

eolf-othd-000001.txt
SEQUENCE LISTING

<110> GlaxoSmithKline Biologicals S.A.

<120> Vaccine

<130> VB62727

<150> US 61/015767

<151> 2007-12-21

<150> US 61/019951

<151> 2008-01-09

<160> 12

<170> PatentIn version 3.5

<210> 1

<211> 3411

<212> DNA

<213> Artificial Sequence

<220>

<223> Nucleotide sequence for p24-RT-Nef-p17 fusion protein

<400> 1

atggttatcg tgcagaacat ccaggggcaa atggtacatc aggccatatc acctagaact	60
ttaaattgcat gggtaaaagt agtagaagag aaggctttca gcccagaagt aatacccatg	120
ttttcagcat tatcagaagg agccacccca caagatttaa acaccatgct aaacacagtg	180
gggggacatc aagcagccat gcaaattgta aaagagacca tcaatgagga agctgcagaa	240
tgggatagag tacatccagt gcatgcaggg cctattgcac caggccagat gagagaacca	300
aggggaagtg acatagcagg aactactagt acccttcagg aacaaatagg atggatgaca	360
aataatccac ctatcccagt aggagaaatt tataaaagat ggataatcct gggattaaat	420
aaaatagtaa gaatgtatag ccctaccagc attctggaca taagacaagg accaaaagaa	480
ccttttagag actatgtaga ccggttctat aaaactctaa gagccgagca agcttcacag	540
gaggtaaaaa attggatgac agaaaccttg ttggtccaaa atgcgaacc agattgtaag	600
actattttaa aagcattggg accagcggct acactagaag aaatgatgac agcatgtcag	660
ggagtaggag gacccggcca taaggcaaga gttttgcata tgggccccat tagccctatt	720
gagactgtgt cagtaaaatt aaagccagga atggatggcc caaaagttaa acaatggcca	780
ttgacagaag aaaaaataaa agcattagta gaaatttgta cagagatgga aaaggaaggg	840
aaaatttcaa aaattgggcc tgaaaatcca tacaatactc cagtatttgc cataaagaaa	900
aaagacagta ctaaattggag aaaattagta gatttcagag aacttaataa gagaactcaa	960
gacttctggg aagttcaatt aggaatacca catcccgcag gggttaaaaaa gaaaaaatca	1020
gtaacagtac tggatgtggg tgatgcatat ttttcagttc ccttagatga agacttcagg	1080
aaatatactg catttaccat acctagtata aacaatgaga caccagggat tagatatcag	1140
tacaatgtgc ttccacaggg atggaaagga tcaccagcaa tattccaaag tagcatgaca	1200
aaaatcttag agccttttag aaaacaaaat ccagacatag ttatctatca atacatggat	1260

eo1f-othd-000001.txt

gatttgtatg	taggatctga	cttagaaata	gggcagcata	gaacaaaaat	agaggagctg	1320
agacaacatc	tgttgaggtg	gggacttacc	acaccagaca	aaaaacatca	gaaagaacct	1380
ccattcctta	aaatgggtta	tgaactccat	cctgataaat	ggacagtaca	gcctatagtg	1440
ctgccagaaa	aagacagctg	gactgtcaat	gacatacaga	agttagtggg	gaaattgaat	1500
tgggcaagtc	agattttacc	agggattaaa	gtaaggcaat	tatgtaaact	ccttagagga	1560
accaaagcac	taacagaagt	aataccacta	acagaagaag	cagagctaga	actggcagaa	1620
aacagagaga	ttctaaaaga	accagtacat	ggagtgtatt	atgacctatc	aaaagactta	1680
atagcagaaa	tacagaagca	ggggcaaggc	caatggacat	atcaaattta	tcaagagcca	1740
tttaaaaaatc	tgaaaacagg	aaaatatgca	agaatgaggg	gtgcccacac	taatgatgta	1800
aaacaattaa	cagaggcagt	gcaaaaaata	accacagaaa	gcatagtaat	atggggaaag	1860
actcctaaat	ttaaactgcc	catacaaaag	gaaacatggg	aaacatggtg	gacagagtat	1920
tggcaagcca	cctggattcc	tgagtgggag	tttgttaata	cccctccttt	agtgaattaa	1980
tgggtaccagt	tagagaaaga	acccatagta	ggagcagaaa	ccttctatgt	agatggggca	2040
gctaacaggg	agactaaatt	aggaaaagca	ggatatgtta	ctaatagagg	aagacaaaaa	2100
gttgtcacc	taactgacac	aacaaatcag	aagactgagt	tacaagcaat	ttatctagct	2160
ttgcaggatt	cggttagtag	agtaaacata	gtaacagact	cacaatatgc	attaggaatc	2220
attcaagcac	aaccagatca	aagtgaatca	gagttagtca	atcaaataat	agagcagtta	2280
ataaaaaagg	aaaaggctta	tctggcatgg	gtaccagcac	acaaaggaat	tggaggaaat	2340
gaacaagtag	ataaattagt	cagtgtctga	atcaggaaag	tgctagctat	gggtggcaag	2400
tgggtcaaaaa	gtagtgtggt	tggatggcct	actgtaaggg	aaagaatgag	acgagctgag	2460
ccagcagcag	atgggggtgg	agcagcatct	cgagacctgg	aaaaacatgg	agcaatcaca	2520
agtagcaata	cagcagctac	caatgctgct	tgtgcctggc	tagaagcaca	agaggaggag	2580
gaggtgggtt	ttccagtcac	acctcaggta	cctttaagac	caatgactta	caaggcagct	2640
gtagatctta	gccacttttt	aaaagaaaag	gggggactgg	aagggttaat	tactcccaa	2700
cgaagacaag	atatccttga	tctgtggatc	taccacacac	aagggtactt	ccctgattgg	2760
cagaactaca	caccagggcc	aggggtcaga	tatccactga	cctttggatg	gtgctacaag	2820
ctagtaccag	ttgagccaga	taaggtagaa	gaggccaata	aaggagagaa	caccagcttg	2880
ttacaccctg	tgagcctgca	tggaatggat	gaccctgaga	gagaagtgtt	agagtggagg	2940
tttgacagcc	gcctagcatt	tcatcacgtg	gcccagagagc	tgcatccgga	gtacttcaag	3000
aactgcaggc	ctatgggtgc	gagagcgtca	gtattaagcg	ggggagaatt	agatcgatgg	3060
gaaaaaattc	ggttaaggcc	agggggaaag	aaaaaatata	aattaaaaca	tatagtatgg	3120
gcaagcaggg	agctagaacg	attcgcagtt	aatcctggcc	tgttagaaac	atcagaaggc	3180
tgtagacaaa	tactgggaca	gctacaacca	tcccttcaga	caggatcaga	agaacttaga	3240
tcattatata	atacagtagc	aaccctctat	tgtgtgcatc	aaaggataga	gataaaagac	3300
accaaggaag	ctttagacaa	gatagaggaa	gagcaaaaaca	aaagtaagaa	aaaagcacag	3360

caagcagcag ctgacacagg acacagcaat caggtcagcc aaaattacta a

3411

<210> 2

<211> 1136

<212> PRT

<213> Artificial Sequence

<220>

<223> Amino acid sequence for p24-RT-Nef-p17 fusion protein

<400> 2

Met Val Ile Val Gln Asn Ile Gln Gly Gln Met Val His Gln Ala Ile
 1 5 10 15

Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu Glu Lys Ala
 20 25 30

Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser Glu Gly Ala
 35 40 45

Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln
 50 55 60

Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu Ala Ala Glu
 65 70 75 80

Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala Pro Gly Gln
 85 90 95

Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Thr Leu
 100 105 110

Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile Pro Val Gly
 115 120 125

Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg
 130 135 140

Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly Pro Lys Glu
 145 150 155 160

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

Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr Leu Leu Val
 180 185 190

Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala Leu Gly Pro
 195 200 205

Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly
 210 215 220

eo1f-othd-000001.txt

Pro Gly His Lys Ala Arg Val Leu His Met Gly Pro Ile Ser Pro Ile
225 230 235 240

Glu Thr Val Ser Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val
245 250 255

Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile
260 265 270

Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu
275 280 285

Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr
290 295 300

Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln
305 310 315 320

Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys
325 330 335

Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser
340 345 350

Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro
355 360 365

Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu
370 375 380

Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr
385 390 395 400

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

Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln
420 425 430

His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly
435 440 445

Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Lys
450 455 460

Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val
465 470 475 480

Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val
485 490 495

Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg
 500 505 510
 Gln Leu Cys Lys Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile
 515 520 525
 Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile
 530 535 540
 Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu
 545 550 555 560
 Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile
 565 570 575
 Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met
 580 585 590
 Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln
 595 600 605
 Lys Ile Thr Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe
 610 615 620
 Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr
 625 630 635 640
 Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro
 645 650 655
 Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala
 660 665 670
 Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly
 675 680 685
 Lys Ala Gly Tyr Val Thr Asn Arg Gly Arg Gln Lys Val Val Thr Leu
 690 695 700
 Thr Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile Tyr Leu Ala
 705 710 715 720
 Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr
 725 730 735
 Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Gln Ser Glu Ser Glu Leu
 740 745 750
 Val Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu
 755 760 765
 Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp

770

775

780

Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu Ala Met Gly Gly Lys
 785 790 795 800

Trp Ser Lys Ser Ser Val Val Gly Trp Pro Thr Val Arg Glu Arg Met
 805 810 815

Arg Arg Ala Glu Pro Ala Ala Asp Gly Val Gly Ala Ala Ser Arg Asp
 820 825 830

Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr Ala Ala Thr Asn
 835 840 845

Ala Ala Cys Ala Trp Leu Glu Ala Gln Glu Glu Glu Val Gly Phe
 850 855 860

Pro Val Thr Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys Ala Ala
 865 870 875 880

Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu Glu Gly Leu
 885 890 895

Ile His Ser Gln Arg Arg Gln Asp Ile Leu Asp Leu Trp Ile Tyr His
 900 905 910

Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly Pro Gly
 915 920 925

Val Arg Tyr Pro Leu Thr Phe Gly Trp Cys Tyr Lys Leu Val Pro Val
 930 935 940

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

Leu His Pro Val Ser Leu His Gly Met Asp Asp Pro Glu Arg Glu Val
 965 970 975

Leu Glu Trp Arg Phe Asp Ser Arg Leu Ala Phe His His Val Ala Arg
 980 985 990

Glu Leu His Pro Glu Tyr Phe Lys Asn Cys Arg Pro Met Gly Ala Arg
 995 1000 1005

Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp Glu Lys Ile
 1010 1015 1020

Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys Leu Lys His Ile
 1025 1030 1035

Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro Gly
 1040 1045 1050

eo1f-othd-000001.txt

Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu
 1055 1060 1065
 Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr
 1070 1075 1080
 Asn Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Ile
 1085 1090 1095
 Lys Asp Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn
 1100 1105 1110
 Lys Ser Lys Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly His
 1115 1120 1125
 Ser Asn Gln Val Ser Gln Asn Tyr
 1130 1135

<210> 3
 <211> 3411
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Nucleotide sequence for F4 codon-optimised

<400> 3
 atggtcattg ttcagaacat acagggccaa atggtccacc aggcaattag tccgcgaact 60
 cttaatgcat ggggtgaaggt cgtggaggaa aaggcattct ccccgagggt cattccgatg 120
 ttttctgctg tatctgaggg cgcaacgccg caagacctta ataccatgct taacacggta 180
 ggccgggcacc aagccgctat gcaaattgcta aaagagacta taaacgaaga ggccgccgaa 240
 tgggatcgag tgcacccggt gcacgccggc ccaattgcac caggccagat gcgcgagccg 300
 cgcggttctg atattgcagg aactacgtct acccttcagg agcagattgg gtggatgact 360
 aacaatccac caatcccgtt cggagagatc tataagaggt ggatcatact gggactaaac 420
 aagatagtcc gcatgtattc tccgacttct atactggata tacgccaagg cccaaaggag 480
 ccgttcaggg actatgtcga ccgattctat aagacccttc gcgcagagca ggcattccag 540
 gaggtcaaaa attggatgac agaaactctt ttggtgcaga atgcgaatcc ggattgtaaa 600
 acaattttta aggctctagg accggccgca acgctagaag agatgatgac ggcttgtcag 660
 ggagtcggtg gaccggggca taaagcccg cgtttacaca tgggcccgat atctccgata 720
 gaaacagttt cgggtcaagct taaaccaggg atggatggtc caaagggtcaa gcagtggccg 780
 ctaacggaag agaagattaa ggcgctcgta gagatttgta ctgaaatgga gaaggaaggg 840
 aagataagca agatcggggc agagaacccg tacaatacac cggttatttg aataaagaaa 900
 aaggattcaa caaatggcg aaagcttgta gattttaggg aactaaacaa gcgaacccaa 960
 gacttttggg aagtccaact agggatccca catccagccg gtctaaagaa gaagaaatcg 1020

eo1f-othd-000001.txt

gtcacagtcc	tggatgtagg	agacgcatat	tttagtgtac	cgcttgatga	ggacttccga	1080
aagtatactg	cgtttactat	accgagcata	aacaatgaaa	cgccaggcat	tcgctatcag	1140
tacaacgtgc	ttccgcaggg	ctggaagggg	tctccggcga	tatttcagag	ctgtatgaca	1200
aaaatacttg	aaccattccg	aaagcagaat	ccggatattg	taattttacca	atacatggac	1260
gatctctatg	tggtgctcga	tctagaaatt	gggcagcatc	gcactaagat	tgaggaactg	1320
aggcaacatc	tgcttcgatg	gggcctcact	actcccgaca	agaagcacca	gaaggagccg	1380
ccgttcctaa	agatgggcta	cgagcttcat	ccggacaagt	ggacagtaca	gccgatagtg	1440
ctgcccga	aggattcttg	gaccgtaaat	gatattcaga	aactagtcgg	caagcttaac	1500
tgggcctctc	agattttacc	aggcattaag	gtccgacagc	tttgcaagct	actgagggga	1560
actaaggctc	taacagaggt	catcccatta	acggaggaag	cagagcttga	gctggcagag	1620
aatcgcgaaa	ttcttaagga	gccggtgcac	gggtataact	acgaccctc	caaggacctt	1680
atagccgaga	tccagaagca	ggggcagggc	caatggacgt	accagatata	tcaagaaccg	1740
tttaagaatc	tgaagactgg	gaagtacgcg	cgcatgagcg	gggctcatal	taatgatgta	1800
aagcaactta	cggaagcagt	acaaaagatt	actactgagt	ctattgtgat	atggggcaag	1860
accccaaagt	tcaagctgcc	catacagaag	gaaacatggg	aaacatggtg	gactgaatat	1920
tggcaagcta	cctggattcc	agaatgggaa	tttgtcaaca	cgccgccact	tgттаagctt	1980
tggtaccagc	ttgaaaagga	gccgatagta	ggggcagaga	ccttctatgt	cgatggcgcc	2040
gcgaatcgcg	aaacgaagct	aggcaaggcg	ggatacgtga	ctaatagggg	ccgccaaaag	2100
gtcgtaaccc	ttacggatac	caccaatcag	aagactgaac	tacaagcgat	ttaccttgca	2160
cttcaggata	gtggcctaga	ggtcaacata	gtcacggact	ctcaatatgc	gcttggcatt	2220
attcaagcgc	agccagatca	aagcgaaagc	gagcttgtaa	accaaataat	agaacagctt	2280
ataaagaaa	agaaggtata	tctggcctgg	gtccccgctc	acaaggggaat	tggcggcaat	2340
gagcaagtgg	acaagctagt	cagcgctggg	attcgcaagg	ttcttgcgat	ggggggtaag	2400
tggctctaagt	ctagcgtagt	cggctggccg	acagtccgcg	agcgcatgcg	acgcgccgaa	2460
ccagccgcag	atggcgtggg	ggcagcgtct	agggatctgg	agaagcacgg	ggctataact	2520
tccagtaaca	cggcggcgac	gaacgccgca	tgcgcatggt	tagaagccca	agaagaggaa	2580
gaagtagggg	ttccggtaac	tccccagggt	ccgttaaggc	cgatgacct	taaggcagcg	2640
gtggatcttt	ctcacttcct	taaggagaaa	ggggggctgg	agggcttaat	tcacagccag	2700
aggcgacagg	atattcttga	tctgtggatt	taccataccc	aggggtactt	tccggactgg	2760
cagaattaca	ccccggggcc	aggcgtgctc	tatcccccta	ctttcggttg	gtgctacaaa	2820
ctagtcccag	tggaacccga	caaggctcaa	gaggctaata	agggcgagaa	cacttctctt	2880
cttcacccgg	taagcctgca	cgggatggat	gaccagaaac	gagaggttct	agaatggagg	2940
ttcgactctc	gacttgctgt	ccatcacgta	gcacgcgagc	tgcatccaga	atatttcaag	3000
aactgccgcc	caatgggcgc	cagggccagt	gtacttagtg	gcggagaact	agatcgatgg	3060
gaaaagatac	gcctacgccc	ggggggcaag	aagaagtaca	agcttaagca	cattgtgtgg	3120

eo1f-othd-000001.txt

gcctctcgcg aacttgagcg attcgcagtg aatccaggcc tgcttgagac gagtgaaggc	3180
tgtaggcaaa ttctggggca gctacagccg agcctacaga ctggcagcga ggagcttcgt	3240
agtctttata ataccgtcgc gactctctac tgcgttcac t aacgaattga aataaaggat	3300
actaaagagg cccttgataa aattgaggag gaacagaata agtcgaaaaa gaaggcccag	3360
caggccgccc cgcacaccgg gcacagcaac caggtgtccc aaaactacta a	3411

<210> 4
 <211> 1302
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Nucleotide sequence for P51 RT

<400> 4	
atgagtactg gtccgatctc tccgatagaa acagtttcgg tcaagcttaa accagggatg	60
gatggtccaa aggtcaagca gtggccgcta acggaagaga agattaaggc gctcgtagag	120
atttgtactg aaatggagaa ggaaggcaag ataagcaaga tcggggccaga gaaccggtac	180
aatacaccgg tatttgcaat aaagaagaag gattcaacaa aatggcgaaa gcttgtagat	240
tttagggaaac taaacaagcg aaccaagac ttttggaag tccaactagg tatcccat	300
ccagccggtc taaagaagaa gaaatcggtc acagtcctgg atgtaggaga cgcataTTTT	360
agtgtaccgc ttgatgagga cttccgaaag tatactgcgt ttactatacc gagcataaac	420
aatgaaacgc caggcattcg ctatcagtac aacgtgctcc cgcagggctg gaaggggtct	480
ccggcgatat ttcagagctc tatgacaaaa atacttgaac cattccgaaa gcagaatccg	540
gatattgtaa tttaccaata catggacgat ctctatgtgg gctcggatct agaaattggg	600
cagcatcgca ctaagattga ggaactgagg caacatctgc ttcgatgggg ctcactact	660
cccgacaaga agcaccagaa ggagccgccc ttcctaaaga tgggctacga gcttcatccg	720
gacaagtgga cagtacagcc gatagtgtcg cccgaaaagg attcttggac cgtaaagtat	780
attcagaaac tagtcggcaa gcttaactgg gcctctcaga tttaccagg cattaaggctc	840
cgacagcttt gcaagctact gaggggaaact aaggctctaa cagagggtcat cccattaacg	900
gaggaagcag agcttgagct ggagagaaat cgcgaaattc ttaaggagcc ggtgcacggg	960
gtatactacg acccctccaa ggaccttata gccgagatcc agaagcaggg gcagggccaa	1020
tggacgtacc agatatatca agaaccgttt aagaatctga agactgggaa gtacgcgcgc	1080
atgcgagggg ctcatactaa tgatgtaaag caacttacgg aagcagtaca aaagattact	1140
actgagtcta ttgtgatatg gggcaagacc ccaaagttca agctgcccac acagaaggaa	1200
acatgggaaa catggtggac tgaatattgg caagctacct ggattccaga atgggaattt	1260
gtcaacacgc cgccgctggt aaaactgagg cctgctagct aa	1302

<210> 5
 <211> 433
 <212> PRT

<213> Artificial Sequence

<220>

<223> Amino acid sequence for P51 RT

<400> 5

Met Ser Thr Gly Pro Ile Ser Pro Ile Glu Thr Val Ser Val Lys Leu
 1 5 10 15

Lys Pro Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Thr Glu
 20 25 30

Glu Lys Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met Glu Lys Glu
 35 40 45

Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val
 50 55 60

Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp
 65 70 75 80

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

Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser Val Thr Val
 100 105 110

Leu Asp Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Asp Phe
 115 120 125

Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn Glu Thr Pro
 130 135 140

Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Ser
 145 150 155 160

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

Lys Gln Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Asp Asp Leu Tyr
 180 185 190

Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys Ile Glu Glu
 195 200 205

Leu Arg Gln His Leu Leu Arg Trp Gly Leu Thr Thr Pro Asp Lys Lys
 210 215 220

His Gln Lys Glu Pro Pro Phe Leu Lys Met Gly Tyr Glu Leu His Pro
 225 230 235 240

Asp Lys Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys Asp Ser Trp
 245 250 255

eo1f-othd-000001.txt

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

Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg
275 280 285

Gly Thr Lys Ala Leu Thr Glu Val Ile Pro Leu Thr Glu Glu Ala Glu
290 295 300

Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val His Gly
305 310 315 320

Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln Lys Gln
325 330 335

Gly Gln Gly Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn
340 345 350

Leu Lys Thr Gly Lys Tyr Ala Arg Met Arg Gly Ala His Thr Asn Asp
355 360 365

Val Lys Gln Leu Thr Glu Ala Val Gln Lys Ile Thr Thr Glu Ser Ile
370 375 380

Val Ile Trp Gly Lys Thr Pro Lys Phe Lys Leu Pro Ile Gln Lys Glu
385 390 395 400

Thr Trp Glu Thr Trp Trp Thr Glu Tyr Trp Gln Ala Thr Trp Ile Pro
405 410 415

Glu Trp Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Arg Pro Ala
420 425 430

Ser

<210> 6
<211> 1023
<212> DNA
<213> Artificial Sequence

<220>
<223> Nef-p17 nucleotide sequence

<400> 6
atgggtggca agtgggtcaaa aagtagtggtg gttggatggc ctactgtaag ggaaagaatg 60
agacgagctg agccagcagc agatgggggtg ggagcagcat ctcgagacct ggaaaaacat 120
ggagcaatca caagtagcaa tacagcagct accaatgctg cttgtgcctg gctagaagca 180
caagaggagg aggaggtggg tttccagtc acacctcagg tacctttaag accaatgact 240
tacaaggcag ctgtagatct tagccacttt ttaaaagaaa aggggggact ggaagggcta 300

eo1f-othd-000001.txt

attcactccc aacgaagaca agatatacctt gatctgtgga tctaccacac acaaggctac 360
 ttccctgatt ggcagaacta cacaccaggg ccaggggtca gatatccact gacctttgga 420
 tgggtgctaca agctagtacc agttgagcca gataaggtag aagaggccaa taaaggagag 480
 aacaccagct tgttacaccc tgtgagcctg catggaatgg atgaccctga gagagaagtg 540
 ttagagtggga ggtttgacag ccgcctagca tttcatcacg tggcccgaga gctgcatccg 600
 gagtacttca agaactgcag gcctatgggt gcgagagcgt cagtattaag cgggggagaa 660
 ttagatcgat gggaaaaaat tcggttaagg ccagggggaa agaaaaata taaattaaaa 720
 catatagtat gggcaagcag ggagctagaa cgattcgag ttaatcctgg cctgttagaa 780
 acatcagaag gctgtagaca aatactggga cagctacaac catcccttca gacaggatca 840
 gaagaactta gatcattata taatacagta gcaaccctct attgtgtgca tcaaaggata 900
 gagataaaaag acaccaagga agcttttagac aagatagagg aagagcaaaa caaaagtaag 960
 aaaaaagcac agcaagcagc agctgacaca ggacacagca atcagggtcag ccaaattac 1020
 taa 1023

<210> 7
 <211> 340
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> Nef-p17 amino acid sequence
 <400> 7

Met Gly Gly Lys Trp Ser Lys Ser Ser Val Val Gly Trp Pro Thr Val
 1 5 10 15
 Arg Glu Arg Met Arg Arg Ala Glu Pro Ala Ala Asp Gly Val Gly Ala
 20 25 30
 Ala Ser Arg Asp Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr
 35 40 45
 Ala Ala Thr Asn Ala Ala Cys Ala Trp Leu Glu Ala Gln Glu Glu Glu
 50 55 60
 Glu Val Gly Phe Pro Val Thr Pro Gln Val Pro Leu Arg Pro Met Thr
 65 70 75 80
 Tyr Lys Ala Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly
 85 90 95
 Leu Glu Gly Leu Ile His Ser Gln Arg Arg Gln Asp Ile Leu Asp Leu
 100 105 110
 Trp Ile Tyr His Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr
 115 120 125

Pro Gly Pro Gly Val Arg Tyr Pro Leu Thr Phe Gly Trp Cys Tyr Lys
 130 135 140

Leu Val Pro Val Glu Pro Asp Lys Val Glu Glu Ala Asn Lys Gly Glu
 145 150 155 160

Asn Thr Ser Leu Leu His Pro Val Ser Leu His Gly Met Asp Asp Pro
 165 170 175

Glu Arg Glu Val Leu Glu Trp Arg Phe Asp Ser Arg Leu Ala Phe His
 180 185 190

His Val Ala Arg Glu Leu His Pro Glu Tyr Phe Lys Asn Cys Arg Pro
 195 200 205

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp
 210 215 220

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys Leu Lys
 225 230 235 240

His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro
 245 250 255

Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu
 260 265 270

Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn
 275 280 285

Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Ile Lys Asp
 290 295 300

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

Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly His Ser Asn Gln Val
 325 330 335

Ser Gln Asn Tyr
 340

<210> 8
 <211> 1029
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> P17-Nef nucleotide sequence

<400> 8
 atgggtgcga gagcgtcagt attaagcggg ggagaattag atcgatggga aaaaattcgg 60
 ttaaggccag ggggaaagaa aaaatataaa ttaaaacata tagtatgggc aagcaggag 120

eolf-othd-000001.txt

ctagaacgat	tcgcagttaa	tcctggcctg	ttagaaacat	cagaaggctg	tagacaaata	180
ctgggacagc	tacaaccatc	ccttcagaca	ggatcagaag	aacttagatc	attatataat	240
acagtagcaa	ccctctattg	tgtgcatcaa	aggatagaga	taaaagacac	caaggaagct	300
ttagacaaga	tagaggaaga	gcaaaacaaa	agtaagaaaa	aagcacagca	agcagcagct	360
gacacaggac	acagcaatca	ggtcagccaa	aattacctcg	acaggcctat	gggtggcaag	420
tgggtcaaaa	gtagtgtggt	tggatggcct	actgtaaggg	aaagaatgag	acgagctgag	480
ccagcagcag	atgggggtggg	agcagcatct	cgagacctgg	aaaaacatgg	agcaatcaca	540
agtagcaata	cagcagctac	caatgctgct	tgtgcctggc	tagaagcaca	agaggaggag	600
gaggtgggtt	ttccagtcac	acctcaggta	cctttaagac	caatgactta	caaggcagct	660
gtagatctta	gccacttttt	aaaagaaaag	gggggactgg	aagggtctaat	tcactcccaa	720
cgaagacaag	atattcctga	tctgtggatc	taccacacac	aaggctactt	ccctgattgg	780
cagaactaca	caccagggcc	aggggtcaga	tatccactga	cctttggatg	gtgctacaag	840
ctagtaccag	ttgagccaga	taaggtagaa	gaggccaata	aaggagagaa	caccagcttg	900
ttacaccctg	tgagcctgca	tggaatggat	gaccctgaga	gagaagtgtt	agagtggagg	960
tttgacagcc	gcctagcatt	tcatcacgtg	gcccagagagc	tgcatccgga	gtacttcaag	1020
aactgctaa						1029

<210> 9
 <211> 3411
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Nucleotide sequence for mutated version of F4 where the Methionine at position 592 is replaced by Lysine

<400> 9	
atggttatcg	tgacaaacat ccaggggcaa atggtacatc aggccatatc acctagaact 60
ttaaatgcat	gggtaaaagt agtagaagag aaggctttca gcccagaagt aatacccatg 120
ttttcagcat	tatcagaagg agccaccca caagatttaa acaccatgct aaacacagtg 180
gggggacatc	aagcagccat gcaaattgta aaagagacca tcaatgagga agctgcagaa 240
tgggtagag	tacatccagt gcatgcagg cctattgcac caggccagat gagagaacca 300
aggggaagtg	acatagcagg aactactagt acccttcagg aacaaatagg atggatgaca 360
aataatccac	ctatcccagt aggagaaatt tataaaagat ggataatcct gggattaaat 420
aaaatagtaa	gaatgtatag ccctaccagc attctggaca taagacaagg accaaaagaa 480
ccttttagag	actatgtaga ccggttctat aaaactctaa gagccgagca agcttcacag 540
gaggtaaaaa	attggatgac agaaaccttg ttggtccaaa atgcgaacc agattgtaag 600
actattttaa	aagcattggg accagcggct acactagaag aaatgatgac agcatgtcag 660
ggagtaggag	gacccggcca taaggcaaga gttttgcata tgggccccat tagccctatt 720
gagactgtgt	cagtaaaatt aaagccagga atggatggcc caaaagttaa acaatggcca 780

eo1f-othd-000001.txt

ttgacagaag	aaaaaataaa	agcattagta	gaaatttgta	cagagatgga	aaaggaaggg	840
aaaatttcaa	aaattgggcc	tgaaaatcca	tacaatactc	cagtatttgc	cataaagaaa	900
aaagacagta	ctaaatggag	aaaattagta	gatttcagag	aacttaataa	gagaactcaa	960
gacttctggg	aagttcaatt	aggaatacca	catcccgag	ggttaaaaaa	gaaaaaatca	1020
gtaacagtac	tggatgtggg	tgatgcatat	ttttcagttc	ccttagatga	agacttcagg	1080
aaatatactg	catttaccat	acctagtata	aacaatgaga	caccagggat	tagatatcag	1140
tacaatgtgc	ttccacaggg	atggaaagga	tcaccagcaa	tattccaaag	tagcatgaca	1200
aaaatcttag	agccttttag	aaaacaaaat	ccagacatag	ttatctatca	atacatggat	1260
gatttgtatg	taggatctga	cttagaaata	gggcagcata	gaacaaaaat	agaggagctg	1320
agacaacatc	tgttgaggtg	gggacttacc	acaccagaca	aaaaacatca	gaaagaacct	1380
ccattcctta	aaatgggtta	tgaactccat	cctgataaat	ggacagtaca	gcctatagtg	1440
ctgccagaaa	aagacagctg	gactgtcaat	gacatacaga	agttagtggg	gaaattgaat	1500
tgggcaagtc	agatttacc	agggattaaa	gtaaggcaat	tatgtaaact	ccttagagga	1560
accaaagcac	taacagaagt	aataccacta	acagaagaag	cagagctaga	actggcagaa	1620
aacagagaga	ttctaaaaga	accagtacat	ggagtgtatt	atgaccatc	aaaagactta	1680
atagcagaaa	tacagaagca	ggggcaaggc	caatggacat	atcaaattta	tcaagagcca	1740
tttaaaaatc	tgaaaacagg	aaaatatgca	cgtaaacgcg	gtgcccacac	taatgatgta	1800
aaacaattaa	cagaggcagt	gcaaaaaata	accacagaaa	gcatagtaat	atggggaaag	1860
actcctaaat	ttaaactgcc	catacaaaag	gaaacatggg	aaacatggtg	gacagagtat	1920
tggcaagcca	cctggattcc	tgagtgggag	tttgtttaata	cccctccttt	agtgaaatta	1980
tgggtaccagt	tagagaaaga	acccatagta	ggagcagaaa	ccttctatgt	agatggggca	2040
gctaacaggg	agactaaatt	aggaaaagca	ggatatgtta	ctaatagagg	aagacaaaaa	2100
gttgtcacc	taactgacac	aacaaatcag	aagactgagt	tacaagcaat	ttatctagct	2160
ttgcaggatt	cgggattaga	agtaaacata	gtaacagact	cacaatatgc	attaggaatc	2220
attcaagcac	aaccagatca	aagtgaatca	gagttagtca	atcaaataat	agagcagtta	2280
ataaaaaagg	aaaaggctta	tctggcatgg	gtaccagcac	acaaaggaat	tggaggaaat	2340
gaacaagtag	ataaattagt	cagtgtctgga	atcaggaaag	tgctagctat	gggtggcaag	2400
tgggtcaaaa	gtagtgtggt	tggatggcct	actgtaaggg	aaagaatgag	acgagctgag	2460
ccagcagcag	atgggggtggg	agcagcatct	cgagacctgg	aaaaacatgg	agcaatcaca	2520
agtagcaata	cagcagctac	caatgctgct	tgtgcctggc	tagaagcaca	agaggaggag	2580
gaggtgggtt	ttccagtcac	acctcaggta	cctttaagac	caatgactta	caaggcagct	2640
gtagatctta	gccacttttt	aaaagaaaag	gggggactgg	aagggtctaat	tcactcccaa	2700
cgaagacaag	atatccttga	tctgtggatc	taccacacac	aagggtactt	ccctgattgg	2760
cagaactaca	caccagggcc	aggggtcaga	tatccactga	cctttggatg	gtgctacaag	2820
ctagtaccag	ttgagccaga	taaggtagaa	gaggccaata	aaggagagaa	caccagcttg	2880

eo1f-othd-000001.txt

ttacaccctg tgagcctgca tggaatggat gaccctgaga gagaagtgtt agagtggagg	2940
tttgacagcc gcctagcatt tcatcacgtg gcccgagagc tgcattccgga gtacttcaag	3000
aactgcaggc ctatgggtgc gagagcgtca gtattaagcg ggggagaatt agatcgatgg	3060
gaaaaaattc ggtaaggcc agggggaaaag aaaaaatata aattaaaaca tatagtatgg	3120
gcaagcaggg agctagaacg attcgcagtt aatcctggcc tgtagaaaac atcagaaggc	3180
tgtagacaaa tactgggaca gctacaacca tcccttcaga caggatcaga agaacttaga	3240
tcattatata atacagtagc aaccctctat tgtgtgcatc aaaggataga gataaaagac	3300
accaaggaag ctttagacaa gatagaggaa gagcaaaaca aaagtaagaa aaaagcacag	3360
caagcagcag ctgacacagg acacagcaat caggtcagcc aaaattacta a	3411

<210> 10
 <211> 1136
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Amino acid sequence for mutated version of F4 where the Methionine at position 592 is replaced by Lysine

<400> 10

Met	Val	Ile	Val	Gln	Asn	Ile	Gln	Gly	Gln	Met	Val	His	Gln	Ala	Ile
1				5				10						15	

Ser	Pro	Arg	Thr	Leu	Asn	Ala	Trp	Val	Lys	Val	Val	Glu	Glu	Lys	Ala
			20					25					30		

Phe	Ser	Pro	Glu	Val	Ile	Pro	Met	Phe	Ser	Ala	Leu	Ser	Glu	Gly	Ala
		35					40					45			

Thr	Pro	Gln	Asp	Leu	Asn	Thr	Met	Leu	Asn	Thr	Val	Gly	Gly	His	Gln
	50					55					60				

Ala	Ala	Met	Gln	Met	Leu	Lys	Glu	Thr	Ile	Asn	Glu	Glu	Ala	Ala	Glu
65					70					75					80

Trp	Asp	Arg	Val	His	Pro	Val	His	Ala	Gly	Pro	Ile	Ala	Pro	Gly	Gln
				85					90					95	

Met	Arg	Glu	Pro	Arg	Gly	Ser	Asp	Ile	Ala	Gly	Thr	Thr	Ser	Thr	Leu
			100					105					110		

Gln	Glu	Gln	Ile	Gly	Trp	Met	Thr	Asn	Asn	Pro	Pro	Ile	Pro	Val	Gly
		115					120					125			

Glu	Ile	Tyr	Lys	Arg	Trp	Ile	Ile	Leu	Gly	Leu	Asn	Lys	Ile	Val	Arg
	130					135					140				

Met	Tyr	Ser	Pro	Thr	Ser	Ile	Leu	Asp	Ile	Arg	Gln	Gly	Pro	Lys	Glu
145					150					155					160

Pro Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu Arg Ala Glu
 165 170 175
 Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr Leu Leu Val
 180 185 190
 Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala Leu Gly Pro
 195 200 205
 Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly
 210 215 220
 Pro Gly His Lys Ala Arg Val Leu His Met Gly Pro Ile Ser Pro Ile
 225 230 235 240
 Glu Thr Val Ser Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val
 245 250 255
 Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile
 260 265 270
 Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu
 275 280 285
 Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr
 290 295 300
 Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln
 305 310 315 320
 Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys
 325 330 335
 Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser
 340 345 350
 Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro
 355 360 365
 Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu
 370 375 380
 Pro Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr
 385 390 395 400
 Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val Ile Tyr
 405 410 415
 Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln
 420 425 430

eo1f-othd-000001.txt

His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly
435 440 445

Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Lys
450 455 460

Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val
465 470 475 480

Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val
485 490 495

Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg
500 505 510

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

Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile
530 535 540

Leu Lys Glu Pro Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu
545 550 555 560

Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile
565 570 575

Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Lys
580 585 590

Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln
595 600 605

Lys Ile Thr Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe
610 615 620

Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr
625 630 635 640

Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro
645 650 655

Leu Val Lys Leu Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala
660 665 670

Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly
675 680 685

Lys Ala Gly Tyr Val Thr Asn Arg Gly Arg Gln Lys Val Val Thr Leu
690 695 700

Thr Asp Thr Thr Asn Gln Lys Thr Glu Leu Gln Ala Ile Tyr Leu Ala
 705 710 715 720
 Leu Gln Asp Ser Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr
 725 730 735
 Ala Leu Gly Ile Ile Gln Ala Gln Pro Asp Gln Ser Glu Ser Glu Leu
 740 745 750
 Val Asn Gln Ile Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu
 755 760 765
 Ala Trp Val Pro Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp
 770 775 780
 Lys Leu Val Ser Ala Gly Ile Arg Lys Val Leu Ala Met Gly Gly Lys
 785 790 795 800
 Trp Ser Lys Ser Ser Val Val Gly Trp Pro Thr Val Arg Glu Arg Met
 805 810 815
 Arg Arg Ala Glu Pro Ala Ala Asp Gly Val Gly Ala Ala Ser Arg Asp
 820 825 830
 Leu Glu Lys His Gly Ala Ile Thr Ser Ser Asn Thr Ala Ala Thr Asn
 835 840 845
 Ala Ala Cys Ala Trp Leu Glu Ala Gln Glu Glu Glu Glu Val Gly Phe
 850 855 860
 Pro Val Thr Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys Ala Ala
 865 870 875 880
 Val Asp Leu Ser His Phe Leu Lys Glu Lys Gly Gly Leu Glu Gly Leu
 885 890 895
 Ile His Ser Gln Arg Arg Gln Asp Ile Leu Asp Leu Trp Ile Tyr His
 900 905 910
 Thr Gln Gly Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly Pro Gly
 915 920 925
 Val Arg Tyr Pro Leu Thr Phe Gly Trp Cys Tyr Lys Leu Val Pro Val
 930 935 940
 Glu Pro Asp Lys Val Glu Glu Ala Asn Lys Gly Glu Asn Thr Ser Leu
 945 950 955 960
 Leu His Pro Val Ser Leu His Gly Met Asp Asp Pro Glu Arg Glu Val
 965 970 975
 Leu Glu Trp Arg Phe Asp Ser Arg Leu Ala Phe His His Val Ala Arg

Glu Leu His Pro Glu Tyr Phe Lys Asn Cys Arg Pro Met Gly Ala Arg
 995 1000 1005

Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp Glu Lys Ile
 1010 1015 1020

Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Lys Leu Lys His Ile
 1025 1030 1035

Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro Gly
 1040 1045 1050

Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu
 1055 1060 1065

Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr
 1070 1075 1080

Asn Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Ile
 1085 1090 1095

Lys Asp Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn
 1100 1105 1110

Lys Ser Lys Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly His
 1115 1120 1125

Ser Asn Gln Val Ser Gln Asn Tyr
 1130 1135

<210> 11
 <211> 3018
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Nucleotide sequence for mutated F4(p51) where putative internal
 Methionine initiation site (present in RT portion) replaced by
 Lysine

<400> 11
 atggttatcg tgcagaacat ccaggggcaa atggtacatc aggccatatc acctagaact 60
 ttaaattgcat gggtaaaagt agtagaagag aaggctttca gcccagaagt aatacccatg 120
 ttttcagcat tatcagaagg agccaccca caagatttaa acaccatgct aaacacagtg 180
 gggggacatc aagcagccat gcaaattgta aaagagacca tcaatgagga agctgcagaa 240
 tgggatagag tacatccagt gcatgcagg cctattgcac caggccagat gagagaacca 300
 aggggaagtg acatagcagg aactactagt acccttcagg aacaaatagg atggatgaca 360
 aataatccac ctatcccagt aggagaaatt tataaaagat ggataatcct gggattaaat 420
 aaaatagtaa gaatgtatag ccctaccagc attctggaca taagacaagg accaaaagaa 480

eo1f-othd-000001.txt

ccttttagag	actatgtaga	ccggttctat	aaaactctaa	gagccgagca	agcttcacag	540
gaggtaaaaa	attggatgac	agaaaccttg	ttgggtccaaa	atgcgaaccc	agattgtaag	600
actattttaa	aagcattggg	accagcggct	acactagaag	aaatgatgac	agcatgtcag	660
ggagtaggag	gacccggcca	taaggcaaga	gttttgcata	tgaggcctgg	tccgatctct	720
ccgatagaaa	cagtttcggt	caagcttaaa	ccagggatgg	atgggtccaaa	ggtcaagcag	780
tggccgctaa	cggaagagaa	gattaaggcg	ctcgtagaga	tttgtactga	aatggagaag	840
gaaggcaaga	taagcaagat	cgggccagag	aacccgtaca	atacaccggt	atttgcaata	900
aagaagaagg	attcaacaaa	atggcgaaag	cttgtagatt	ttagggaaact	aaacaagcga	960
acccaagact	tttgggaagt	ccaactaggt	atcccacatc	cagccggtct	aaagaagaag	1020
aaatcgggtca	cagtcctgga	tgtaggagac	gcatatttta	gtgtaccgct	tgatgaggac	1080
ttccgaaagt	atactgcgtt	tactataccg	agcataaaca	atgaaacgcc	aggcattcgc	1140
tatcagtaca	acgtgctccc	gcagggctgg	aaggggtctc	cggcgatatt	tcagagctct	1200
atgacaaaaa	tacttgaacc	attccgaaag	cagaatccgg	atattgtaat	ttaccaatac	1260
atggacgatc	tctatgtggg	ctcggatcta	gaaattgggc	agcatcgcac	taagattgag	1320
gaactgaggc	aacatctgct	tcgatggggc	ctcactactc	ccgacaagaa	gcaccagaag	1380
gagccgccgt	tcctaaagat	gggctacgag	cttcatccgg	acaagtggac	agtacagccg	1440
atagtgctgc	ccgaaaagga	ttcttggacc	gtaaatgata	ttcagaaact	agtcggcaag	1500
cttaactggg	cctctcagat	ttacccaggc	attaagggtcc	gacagctttg	caagctactg	1560
aggggaacta	aggctctaac	agaggtcatc	ccattaacgg	aggaagcaga	gcttgagctg	1620
gcagagaatc	gcgaaattct	taaggagccg	gtgcacaggg	tatactacga	cccctccaag	1680
gaccttatag	ccgagatcca	gaagcagggg	cagggccaat	ggacgtacca	gatatatcaa	1740
gaaccgttta	agaatctgaa	gactgggaag	tacgcgcgca	aacgaggggc	tcataactaat	1800
gatgtaaagc	aacttacgga	agcagtacaa	aagattacta	ctgagtctat	tgtgatatgg	1860
ggcaagaccc	caaagttcaa	gctgcccata	cagaaggaaa	catgggaaac	atgggtggact	1920
gaatattggc	aagctacctg	gattccagaa	tgggaatttg	tcaacacgcc	gccgctggta	1980
aaactggccc	tagctatggg	tggcaagtgg	tcaaaaagta	gtgtggttgg	atggcctact	2040
gtaagggaaa	gaatgagacg	agctgagcca	gcagcagatg	gggtgggagc	agcatctcga	2100
gacctggaaa	aacatggagc	aatcacaagt	agcaatacag	cagctaccaa	tgctgcttgt	2160
gcctggctag	aagcacaaga	ggaggaggag	gtgggttttc	cagtcacacc	tcaggtaacct	2220
ttaagaccaa	tgacttacaa	ggcagctgta	gatcttagcc	acttttttaa	agaaaagggg	2280
ggactggaag	ggctaattca	ctcccaacga	agacaagata	tccttgatct	gtggatctac	2340
cacacacaag	gctacttccc	tgattggcag	aactacacac	cagggccagg	ggtcagatat	2400
ccactgacct	ttggatgggtg	ctacaagcta	gtaccagttg	agccagataa	ggtagaagag	2460
gccaataaag	gagagaacac	cagcttggtta	caccctgtga	gcctgcatgg	aatggatgac	2520

eo1f-othd-000001.txt

cctgagagag aagtgttaga gtggagggtt gacagccgcc tagcatttca tcacgtggcc 2580
cgagagctgc atccggagta cttcaagaac tgcaggccta tgggtgagag agcgctcagta 2640
ttaagcgggg gagaattaga tcgatgggaa aaaattcggg taaggccagg gggaaagaaa 2700
aaatataaat taaaacatat agtatgggca agcaggggagc tagaacgatt cgcagttaat 2760
cctggcctgt tagaaacatc agaaggctgt agacaaatac tgggacagct acaaccatcc 2820
cttcagacag gatcagaaga acttagatca ttatataata cagtagcaac cctctattgt 2880
gtgcatcaaa ggatagagat aaaagacacc aaggaagctt tagacaagat agaggaagag 2940
caaaacaaaa gtaagaaaaa agcacagcaa gcagcagctg acacaggaca cagcaatcag 3000
gtcagccaaa attactaa 3018

<210> 12
<211> 1005
<212> PRT
<213> Artificial Sequence

<220>
<223> Amino acid sequence for mutated F4(p51) where putative internal Methionine initiation site (present in RT portion) replaced by Lysine

<400> 12

Met Val Ile Val Gln Asn Ile Gln Gly Gln Met Val His Gln Ala Ile
1 5 10 15

Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu Glu Lys Ala
20 25 30

Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser Glu Gly Ala
35 40 45

Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln
50 55 60

Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu Ala Ala Glu
65 70 75 80

Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala Pro Gly Gln
85 90 95

Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Thr Leu
100 105 110

Gln Glu Gln Ile Gly Trp Met Thr Asn Asn Pro Pro Ile Pro Val Gly
115 120 125

Glu Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg
130 135 140

Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Arg Gln Gly Pro Lys Glu
145 150 155 160

eo1f-othd-000001.txt

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

Gln Ala Ser Gln Glu Val Lys Asn Trp Met Thr Glu Thr Leu Leu Val
180 185 190

Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala Leu Gly Pro
195 200 205

Ala Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly
210 215 220

Pro Gly His Lys Ala Arg Val Leu His Met Arg Pro Gly Pro Ile Ser
225 230 235 240

Pro Ile Glu Thr Val Ser Val Lys Leu Lys Pro Gly Met Asp Gly Pro
245 250 255

Lys Val Lys Gln Trp Pro Leu Thr Glu Glu Lys Ile Lys Ala Leu Val
260 265 270

Glu Ile Cys Thr Glu Met Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly
275 280 285

Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Asp
290 295 300

Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg Glu Leu Asn Lys Arg
305 310 315 320

Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly
325 330 335

Leu Lys Lys Lys Lys Ser Val Thr Val Leu Asp Val Gly Asp Ala Tyr
340 345 350

Phe Ser Val Pro Leu Asp Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr
355 360 365

Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn
370 375 380

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

Met Thr Lys Ile Leu Glu Pro Phe Arg Lys Gln Asn Pro Asp Ile Val
405 410 415

Ile Tyr Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile
420 425 430

eo1f-othd-000001.txt

Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg Gln His Leu Leu Arg
435 440 445

Trp Gly Leu Thr Thr Pro Asp Lys Lys His Gln Lys Glu Pro Pro Phe
450 455 460

Leu Lys Met Gly Tyr Glu Leu His Pro Asp Lys Trp Thr Val Gln Pro
465 470 475 480

Ile Val Leu Pro Glu Lys Asp Ser Trp Thr Val Asn Asp Ile Gln Lys
485 490 495

Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys
500 505 510

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

Val Ile Pro Leu Thr Glu Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg
530 535 540

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

Asp Leu Ile Ala Glu Ile Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr
565 570 575

Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala
580 585 590

Arg Lys Arg Gly Ala His Thr Asn Asp Val Lys Gln Leu Thr Glu Ala
595 600 605

Val Gln Lys Ile Thr Thr Glu Ser Ile Val Ile Trp Gly Lys Thr Pro
610 615 620

Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr
625 630 635 640

Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp Glu Phe Val Asn Thr
645 650 655

Pro Pro Leu Val Lys Leu Ala Leu Ala Met Gly Gly Lys Trp Ser Lys
660 665 670

Ser Ser Val Val Gly Trp Pro Thr Val Arg Glu Arg Met Arg Arg Ala
675 680 685

Glu Pro Ala Ala Asp Gly Val Gly Ala Ala Ser Arg Asp Leu Glu Lys
690 695 700

His Gly Ala Ile Thr Ser Ser Asn Thr Ala Ala Thr Asn Ala Ala Cys
 705 710 715 720
 Ala Trp Leu Glu Ala Gln Glu Glu Glu Glu Val Gly Phe Pro Val Thr
 725 730 735
 Pro Gln Val Pro Leu Arg Pro Met Thr Tyr Lys Ala Ala Val Asp Leu
 740 745 750
 Ser His Phe Leu Lys Glu Lys Gly Gly Leu Glu Gly Leu Ile His Ser
 755 760 765
 Gln Arg Arg Gln Asp Ile Leu Asp Leu Trp Ile Tyr His Thr Gln Gly
 770 775 780
 Tyr Phe Pro Asp Trp Gln Asn Tyr Thr Pro Gly Pro Gly Val Arg Tyr
 785 790 795 800
 Pro Leu Thr Phe Gly Trp Cys Tyr Lys Leu Val Pro Val Glu Pro Asp
 805 810 815
 Lys Val Glu Glu Ala Asn Lys Gly Glu Asn Thr Ser Leu Leu His Pro
 820 825 830
 Val Ser Leu His Gly Met Asp Asp Pro Glu Arg Glu Val Leu Glu Trp
 835 840 845
 Arg Phe Asp Ser Arg Leu Ala Phe His His Val Ala Arg Glu Leu His
 850 855 860
 Pro Glu Tyr Phe Lys Asn Cys Arg Pro Met Gly Ala Arg Ala Ser Val
 865 870 875 880
 Leu Ser Gly Gly Glu Leu Asp Arg Trp Glu Lys Ile Arg Leu Arg Pro
 885 890 895
 Gly Gly Lys Lys Lys Tyr Lys Leu Lys His Ile Val Trp Ala Ser Arg
 900 905 910
 Glu Leu Glu Arg Phe Ala Val Asn Pro Gly Leu Leu Glu Thr Ser Glu
 915 920 925
 Gly Cys Arg Gln Ile Leu Gly Gln Leu Gln Pro Ser Leu Gln Thr Gly
 930 935 940
 Ser Glu Glu Leu Arg Ser Leu Tyr Asn Thr Val Ala Thr Leu Tyr Cys
 945 950 955 960
 Val His Gln Arg Ile Glu Ile Lys Asp Thr Lys Glu Ala Leu Asp Lys
 965 970 975
 Ile Glu Glu Glu Gln Asn Lys Ser Lys Lys Ala Gln Gln Ala Ala

980 eolf-othd-000001.txt 990
985

Ala Asp Thr Gly His Ser Asn Gln Val Ser Gln Asn Tyr
995 1000 1005