

SEQUENCE LISTING

<110> Oxford Genetics Limited
Oxford University Innovation Limited

<120> PROCESS FOR MAKING ADENOASSOCIATED VIRAL VECTORS

<130> 489.147164/04

<160> 23

<170> PatentIn version 3.5

<210> 1

<211> 1866

<212> DNA

<213> Rep nucleotide sequence (adeno-associated virus 2)

<400> 1

atgccgggggt tttacgagat tgtgattaag gtccccagcg accttgacga gcatctgccc	60
ggcatttctg acagctttgt gaactgggtg gccgagaagg aatgggagtt gccgccagat	120
tctgacatgg atctgaatct gattgagcag gcacccctga ccgtggccga gaagctgcag	180
cgcgactttc tgacggaatg gcgccgtgtg agtaaggccc cggaggccct tttctttgtg	240
caatttgaga agggagagag ctacttccac atgcacgtgc tcgtggaaac caccgggggtg	300
aaatccatgg ttttgggacg tttcctgagt cagattcgcg aaaaactgat tcagagaatt	360
taccgcggga tcgagccgac ttgccaaac tggttcgcg tcacaaagac cagaaatggc	420
gccggaggcg ggaacaaggt ggtggatgag tgctacatcc ccaattactt gctcccaaa	480
accagcctg agctccagt ggcgtggact aatatggaac agtatttaag cgcctgtttg	540
aatctcacgg agcgtaaagc gttggtggcg cagcatctga cgcacgtgtc gcagacgcag	600
gagcagaaca aagagaatca gaatccaat tctgatgcgc cggatgatcag atcaaaaact	660
tcagccaggt acatggagct ggtcgggtgg ctctgggaca aggggattac ctcggagaag	720
cagtggatcc aggaggacca ggcctcatac atctccttca atgcggcctc caactcgcg	780
tcccaaatca aggctgcctt ggacaatgcg ggaaagatta tgagcctgac taaaaccgcc	840
cccgactacc tgggtgggcca gcagcccgtg gaggacattt ccagcaatcg gatttataaa	900
attttggaaac taaacgggta cgatcccaa tatgcggctt ccgtctttct gggatgggcc	960
acgaaaaagt tcggcaagag gaacaccatc tggctgtttg ggcctgcaac taccgggaag	1020
accaacatcg cggaggccat agcccacact gtgcccttct acgggtgcgt aaactggacc	1080

aatgagaact ttcccttcaa cgactgtgtc gacaagatgg tgatctggtg ggaggagggg	1140
aagatgaccg ccaaggtcgt ggagtcggcc aaagccattc tcggaggaag caaggtgcgc	1200
gtggaccaga aatgcaagtc ctcggcccag atagaccga ctcccgtgat cgtcacctcc	1260
aacaccaaca tgtgcgccgt gattgacggg aactcaacga ctttcgaaca ccagcagccg	1320
ttgcaagacc ggatgttcaa atttgaactc acccgccgtc tggatcatga ctttggaag	1380
gtcaccaagc aggaagtcaa agactttttc cgggtgggcaa aggatcacgt ggttgaggtg	1440
gagcatgaat tctacgtcaa aaagggtgga gccaaagaaa gacccgcccc cagtgcgcga	1500
gatataagtg agcccaaacg ggtgcgcgag tcagttgcgc agccatcgac gtcagacgcg	1560
gaagcttcga tcaactacgc agacaggtac caaaacaaat gttctcgtca cgtgggcatg	1620
aatctgatgc tgtttccctg cagacaatgc gagagaatga atcagaattc aaatatctgc	1680
ttcactcacg gacagaaaga ctgttttagag tgctttcccg tgtcagaatc tcaaccggtt	1740
tctgtcgtca aaaaggcgta tcagaaactg tgctacattc atcatatcat gggaaaggtg	1800
ccagacgctt gcactgcctg cgatctggtc aatgtggatt tggatgactg catctttgaa	1860
caatag	1866

<210> 2
 <211> 2559
 <212> DNA
 <213> Cap nucleotide sequence (adeno-associated virus 2)

<400> 2	
cagttgcgca gccatcgacg tcagacgcgg aagcttcgat caactacgca gacaggtacc	60
aaaacaaatg ttctcgtcac gtgggcatga atctgatgct gtttccctgc agacaatgcg	120
agagaatgaa tcagaattca aatatctgct tcactcacgg acagaaagac tgtttagagt	180
gctttcccggt gtcagaatct caaccggtt ctgtcgtcaa aaaggcgat cagaaactgt	240
gctacattca tcatatcatg ggaaaggtgc cagacgcttg cactgcctgc gatctggtca	300
atgtggattt ggatgactgc atctttgaac aataaatgat ttaaatcagg tatggctgcc	360
gatggttatc ttccagattg gctcgaggac actctctctg aaggaataag acagtgggtg	420
aagctcaaac ctggcccacc accaccaaag cccgcagagc ggcataagga cgacagcagg	480
ggtcttgtgc ttctgggta caagtacctc ggacccttca acggactcga caaggagag	540

ccggtcaacg aggcagacgc cgcggccctc gagcacgaca aagcctacga ccggcagctc	600
gacagcggag acaacccgta cctcaagtac aaccacgccg acgcggagtt tcaggagcgc	660
cttaaagaag atacgtcttt tgggggcaac ctcggacgag cagtcttcca ggcgaaaaag	720
agggttcttg aacctctggg cctgggtgag gaacctgtta agacggctcc gggaaaaaag	780
aggccggtag agcactctcc tgtggagcca gactcctcct cggaaccgg aaaggcgggc	840
cagcagcctg caagaaaaag attgaatttt ggtcagactg gagacgcaga ctcagtacct	900
gacccccagc ctctcggaca gccaccagca gccccctctg gtctgggaac taatacgatg	960
gctacaggca gtggcgcacc aatggcagac aataacgagg gcgccgacgg agtgggtaat	1020
tcctcgggaa attggcattg cgattccaca tggatgggcg acagagtcac caccaccagc	1080
acccgaacct gggccctgcc cacctacaac aaccacctct acaaacaat ttccagccaa	1140
tcaggagcct cgaacgaaa tcactacttt ggctacagca ccccttgggg gtattttgac	1200
ttcaacagat tccactgcca cttttcacca cgtgactggc aaagactcat caacaacaac	1260
tggggattcc gaccaagag actcaacttc aagctcttta acattcaagt caaagaggtc	1320
acgcagaatg acggtacgac gacgattgcc aataacctta ccagcacggt tcagggtgtt	1380
actgactcgg agtaccagct cccgtacgtc ctcggctcgg cgcatcaagg atgcctcccg	1440
ccgttcccag cagacgtctt catggtgcca cagtatggat acctcacct gaacaacggg	1500
agtcaggcag taggacgctc ttcatTTTTac tgcctggagt actttccttc tcagatgctg	1560
cgtaccggaa acaactttac cttcagctac acttttgagg acgttccttt ccacagcagc	1620
tacgctcaca gccagagtct ggaccgtctc atgaatcctc tcacgacca gtacctgtat	1680
tacttgagca gaacaaacac tccaagtgga accaccacgc agtcaaggct tcagttttct	1740
caggccggag cgagtacat tcgggaccag tctaggaact ggcttcctgg acctgttac	1800
cgccagcagc gagtatcaaa gacatctgcg gataacaaca acagtgaata ctcgtggact	1860
ggagctacca agtaccacct caatggcaga gactctctgg tgaatccggg cccggccatg	1920
gcaagccaca aggacgatga agaaaagttt tttcctcaga gcggggttct catctttggg	1980
aagcaaggct cagagaaaac aaatgtggac attgaaaagg tcatgattac agacgaagag	2040
gaaatcagga caaccaatcc cgtggctacg gagcagtatg gttctgtatc taccaacctc	2100
cagagaggca acagacaagc agctaccgca gatgtcaaca cacaaggcgt tcttccaggc	2160

atggtctggc aggacagaga tgtgtacctt caggggccc a tctgggcaaa gattccacac 2220
 acggacggac attttcaccc ctctcccctc atgggtggat tcggacttaa acaccctcct 2280
 ccacagattc tcatcaagaa caccgccgta cctgcgaatc cttcgaccac cttcagtgcg 2340
 gcaaagtgtt cttccttcat cacacagtac tccacgggac aggtcagcgt ggagatcgag 2400
 tgggagctgc agaaggaaaa cagcaaacgc tggaatcccg aaattcagta cacttccaac 2460
 tacaacaagt ctgttaatgt ggactttact gtggacacta atggcgtgta ttcagagcct 2520
 cgccccattg gcaccagata cctgactcgt aatctgtaa 2559

<210> 3
 <211> 735
 <212> PRT
 <213> Cap amino acid sequence (adeno-associated virus 2)

<400> 3

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

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

Lys Pro Ala Glu Arg His Lys Asp Asp Ser Arg Gly Leu Val Leu Pro
 35 40 45

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

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

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

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

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

Leu Gly Leu Val Glu Glu Pro Val Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu His Ser Pro Val Glu Pro Asp Ser Ser Ser Gly Thr Gly
145 150 155 160

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

Gly Asp Ala Asp Ser Val Pro Asp Pro Gln Pro Leu Gly Gln Pro Pro
180 185 190

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

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

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Met Gly Asp Arg Val Ile
225 230 235 240

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

Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Ser Asn Asp Asn His Tyr
260 265 270

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

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

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

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

Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro Tyr
340 345 350

Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala Asp
355 360 365

Val Phe Met Val Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser
370 375 380

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

Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe Glu
405 410 415

Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp Arg
420 425 430

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

Asn Thr Pro Ser Gly Thr Thr Thr Gln Ser Arg Leu Gln Phe Ser Gln
450 455 460

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

Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ser Ala Asp Asn Asn
485 490 495

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

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

Asp Glu Glu Lys Phe Phe Pro Gln Ser Gly Val Leu Ile Phe Gly Lys
530 535 540

Gln Gly Ser Glu Lys Thr Asn Val Asp Ile Glu Lys Val Met Ile Thr
545 550 555 560

Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln Tyr
565 570 575

Gly Ser Val Ser Thr Asn Leu Gln Arg Gly Asn Arg Gln Ala Ala Thr
580 585 590

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

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

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

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

Pro Ser Thr Thr Phe Ser Ala Ala Lys Phe Ala Ser Phe Ile Thr Gln
660 665 670

Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys
675 680 685

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

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

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

<210> 4

<211> 130

<212> DNA

<213> Wild-type adeno-associated virus 2 p5 promoter

<400> 4

gtcctgtatt agaggtcacg tgagtgtttt gcgacatttt gcgacaccat gtggtcacgc 60

tgggtattta agcccgagtg agcacgcagg gtctccattt tgaagcggga ggtttgaacg 120

cgcagccgcc 130

<210> 5

<211> 75

<212> DNA

<213> Sequence that forms part of the 5'-untranslated region (UTR) of the Homo sapiens beta-globin gene

<400> 5

agatctttgt cgatcctacc atccactcga cacacccgcc agcggccgct gccaaagcttc 60

cgagctctcg aattc 75

<210> 6

<211> 273

<212> DNA

<213> Wild-type adeno-associated virus 2 p19 promoter

<400> 6

gtcacaaaga ccagaaatgg cgccggaggc gggaacaagg tggtggatga gtgctacatc 60

cccaattact tgctcccaa aaccagcct gagctccagt gggcgtggac taatatggaa 120

cagtatttaa ggcctgttt gaatctcac gagcgtaaac ggttggtggc gcagcatctg 180

acgcacgtgt cgagacgca ggagcagaac aaagagaatc agaatcccaa ttctgatgcg 240

ccggtgatca gatcaaaaac ttcagccagg tac 273

<210> 7

<211> 273

<212> DNA

<213> Artificial Sequence

<220>

<223> Non-functional p19 sequence

<400> 7

gtcacaaaga ccagaaatgg cgccggaggc gggaacaagg tggtggatga gtgctacatc 60

cccaattact tgctcccaa aaccagcct gagctccagt gggcgtggac taatatggaa 120

cagtacctca ggcctgttt gaatctcac gagcgtaaac ggttggtggc gcagcatctg 180

acgcacgtgt cgagacgca ggagcagaac aaagagaatc agaatcccaa ttctgatgcg 240

ccggtgatca gatcaaaaac ttcagccagg tac 273

<210> 8

<211> 209

<212> DNA
<213> Wild-type adeno-associated virus 2 p40 promoter

<400> 8
ggtcaccaag caggaagtca aagacttttt cgggtgggca aaggatcacg tggttgaggt 60
ggagcatgaa ttctacgtca aaaagggtgg agccaagaaa agaccgccc ccagtgcgc 120
agatataagt gagcccaaac ggggtgcgca gtcagttgcg cagccatcga cgtcagacgc 180
ggaagcttcg atcaactacg cagacaggt 209

<210> 9
<211> 208
<212> DNA
<213> Artificial Sequence

<220>
<223> Non-functional p40 promoter sequence

<400> 9
gtgacaaaac aagaggtgaa ggacttcttt cgttgggcca aagaccatgt ggtcgaggtc 60
gaacacgagt tctatgtgaa gaaaggaggc gcgaagaagc gccagcgcc atcggacgct 120
gacatctccg aaccgaagcg cgtgagagag agcgtggcac aaccatcaac ctcggatgcc 180
gaggcatcca tcaattatgc ggacaggt 208

<210> 10
<211> 208
<212> DNA
<213> Wild-type adeno-associated virus 2 adenovirus inhibitor

<400> 10
gtcaccaagc aggaagtcaa agactttttc cgggtgggcaa aggatcacgt ggttgaggtg 60
gagcatgaat tctacgtcaa aaagggtgga gccaaagaaa gaccgcccc cagtgcgcga 120
gatataagtg agcccaaacg ggtgcgagag tcagttgcgc agccatcgac gtcagacgcg 180
gaagcttcga tcaactacgc agacaggt 208

<210> 11
<211> 208
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleotide sequence without a functional p40 promoter sequence

<400> 11
 gtgacaaaac aagaggtgaa ggacttcttt cgttgggcca aagaccatgt ggtcgaggtc 60
 gaacacgagt tctatgtgaa gaaaggaggc gcgaagaagc gcccagcgcc atcggacgct 120
 gacatctccg aaccgaagcg cgtgagagag agcgtggcac aaccatcaac ctcggatgcc 180
 gaggcattcca tcaattatgc ggacaggt 208

<210> 12
 <211> 1932
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> rep gene sequence with p19 and p40 promoters ablated, retaining the Rep78 and Rep68 coding sequences

<400> 12
 atgccgggggt tttacgagat tgtgattaag gtccccagcg accttgacga gcatctgccc 60
 ggcatcttctg acagctttgt gaactgggtg gccgagaagg aatgggagtt gccgccagat 120
 tctgacatgg atctgaatct gattgagcag gcacccctga ccgtggccga gaagctgcag 180
 cgcgactttc tgacggaatg gcgccgtgtg agtaaggccc cggaggccct tttctttgtg 240
 caatttgaga agggagagag ctacttcac atgcacgtgc tcgtggaaac caccggggtg 300
 aaatccatgg ttttgggacg tttcctgagt cagattcgcg aaaaactgat tcagagaatt 360
 taccgcggga tcgagccgac tttgccaaac tggttcgcgg tcacaaagac cagaaatggc 420
 gccggaggcg ggaacaaggt ggtggatgag tgctacatcc ccaattactt gctcccaaa 480
 acccagcctg agctccagtg ggcgtggact aatatggaac agtacctcag cgcctgtttg 540
 aatctcacgg agcgtaaagc gttggtggcg cagcatctga cgcacgtgtc gcagacgcag 600
 gagcagaaca aagagaatca gaatcccaat tctgatgcgc cggatgatcag atcaaaaact 660
 tcagccaggt acatggagct ggtcgggtgg ctctgggaca aggggattac ctcggagaag 720
 cagtggatcc aggaggacca ggcctcatal atctccttca atgcggcctc caactcgcgg 780
 tcccaaatca aggctgcctt ggacaatgcg ggaaagatta tgagcctgac taaaaccgcc 840
 cccgactacc tgggtgggcca gcagcccgtg gaggacattt ccagcaatcg gatttataaa 900
 attttggaaac taaacgggta cgatcccaa tatgcggctt ccgtctttct gggatgggcc 960
 acgaaaaagt tcggcaagag gaacaccatc tggctgtttg ggcctgcaac taccgggaag 1020

accaacatcg cggaggccat agcccacact gtgcccttct acgggtgcgt aaactggacc	1080
aatgagaact ttcccttcaa cgactgtgtc gacaagatgg tgatctggtg ggaggagggg	1140
aagatgaccg ccaaggtcgt ggagtcggcc aaagccattc tcggaggaag caaggtgcgc	1200
gtggaccaga aatgcaagtc ctcggcccag atagaccga ctcccgtgat cgtcacctcc	1260
aacaccaaca tgtgcgccgt catcgacgga aatagcacca ctttcgagca tcaacagcct	1320
ctgcaggatc ggatgtttta gttcgagctg acgaggcggc tcgaccatga tttcgggaaa	1380
gtgacaaaac aagaggtgaa ggacttcttt cgttgggcca aagaccatgt ggtcgaggtc	1440
gaacacgagt tctatgtgaa gaaaggaggc gcgaagaagc gccagcgcc atcggacgct	1500
gacatctccg aaccgaagcg cgtgagagag agcgtggcac aaccatcaac ctcggatgcc	1560
gaggcatcca tcaattatgc ggacaggtac caaaacaaat gttctcgtca cgtgggcatg	1620
aatctgatgc tgtttccctg cagacaatgc gagagaatga atcagaattc aaatatctgc	1680
ttcactcacg gacagaaaga ctgttttagag tgctttcccg tgtcagaatc tcaaccggtt	1740
tctgtcgtca aaaaggcgta tcagaaactg tgctacattc atcatatcat gggaaagggtg	1800
ccagatgctt gcactgcctg cgatctggtc aatgtggatt tggatgactg catctttgaa	1860
caataaatga tttaattcag gtatggctgc cgatggttat cttccagatt ggctcgagga	1920
cactctctct ga	1932

<210> 13
 <211> 19
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> TetR binding site

<400> 13	
tccctatcag tgatagaga	19

<210> 14
 <211> 107
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Modified MLP

<400> 14

cgccctcttc ggcacatcaagg aaggtgattg gttttaggt gtaggccacg tgaccgggtg	60
ttcctgaagg ggggctataa aaggtcccta tcagtgatag agactca	107

<210> 15
 <211> 107
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Modified MLP

<400> 15	
cgccctcttc ggcacatcaagg aaggtgattg gttttaggt gtaggccacg tgactcccta	60
tcagtgatag agaactataa aaggtcccta tcagtgatag agactca	107

<210> 16
 <211> 1195
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Rep52 nucleotide sequence

<400> 16	
catggagctg gtcgggtggc tcgtggacaa ggggattacc tcggagaagc agtggatcca	60
ggaggaccag gcctcataca tctccttcaa tgcggcctcc aactcgcggt cccaaatcaa	120
ggctgccttg gacaatgcgg gaaagattat gaggctgact aaaaccgccc ccgactacct	180
ggtgggccag cagcccgtgg aggacatttc cagcaatcgg atttataaaa ttttggaact	240
aaacgggtac gatccccaat atgcggcttc cgtctttctg ggatgggcca cgaaaaagtt	300
cggcaagagg aacaccatct ggctgtttgg gcctgcaact accgggaaga ccaacatcgc	360
ggaggccata gcccacactg tgcccttcta cgggtgcgta aactggacca atgagaactt	420
tcccttcaac gactgtgtcg acaagatggg gatctggtgg gaggagggga agatgaccgc	480
caaggtcgtg gagtgcggca aagccattct cggaggaagc aaggtgcgcg tggaccagaa	540
atgcaagtcc tcggcccaga tagaccgac tcccgtgac gtcacctcca acaccaacat	600
gtgcgccgtg attgacggga actcaacgac cttcgaacac cagcagccgt tgcaagaccg	660
gatgttcaaa tttgaactca cccgccgtct ggatcatgac tttgggaagg tcaccaagca	720
ggaagtcaaa gactttttcc ggtgggcaaa ggatcacgtg gttgaggtgg agcatgaatt	780

ctacgtcaaa aagggtggag ccaagaaaag acccgcccc agtgacgcag atataagtga	840
gccccaaacgg gtgcgcgagt cagttgcgca gccatcgacg tcagacgcgg aagcttcgat	900
caactacgca gacaggtacc aaaacaaatg ttctcgtcac gtgggcatga atctgatgct	960
gtttccctgc agacaatgcg agagaatgaa tcagaattca aatatctgct tcactcacgg	1020
acagaaagac tgtttagagt gctttcccgt gtcagaatct caaccgttt ctgtcgtcaa	1080
aaaggcgtat cagaaactgt gctacattca tcatatcatg ggaaagggtgc cagacgcttg	1140
cactgcctgc gatctggtca atgtggattt ggatgactgc atctttgaac aatag	1195

<210> 17

<211> 939

<212> DNA

<213> Artificial Sequence

<220>

<223> Rep40 nucleotide sequence

<400> 17

atggagctgg tcgggtggct cgtggacaag gggattacct cggagaagca gtggatccag	60
gaggaccagg cctcatacat ctccttcaat gcggcctcca actcgcggtc ccaaatacaag	120
gctgccttgg acaatgcggg aaagattatg agcctgacta aaaccgcccc cgactacctg	180
gtgggccagc agcccgtgga ggacatttcc agcaatcgga ttataaaaat ttggaacta	240
aacgggtacg atccccaata tgcggcttcc gtctttctgg gatgggccac gaaaaagtgc	300
ggcaagagga acaccatctg gctgtttggg cctgcaacta ccgggaagac caacatcgcg	360
gaggccatag cccacactgt gcccttctac gggcgcgtta actggaccaa tgagaacttt	420
cccttcaacg actgtgtcga caagatggtg atctggtggg aggaggggaa gatgaccgcc	480
aaggtcgtgg agtcggccaa agccattctc ggaggaagca aggtgcgcgt ggaccagaaa	540
tgcaagtcct cggcccagat agacccgact cccgtgatcg tcacctcaa caccaacatg	600
tgcgccgtga ttgacgggaa ctcaacgacc ttcgaacacc agcagccgtt gcaagaccgg	660
atgttcaaata ttgaactcac ccgccgtctg gatcatgact ttgggaaggt caccaagcag	720
gaagtcaaag actttttccg gtgggcaaag gatcacgtgg ttgaggtgga gcatgaattc	780
tacgtcaaaa aggggtggagc caagaaaaga cccgccccca gtgacgcaga tataagttag	840
cccaaacggg tgcgcgagtc agttgcgcag ccatcgacgt cagacgcgga agcttcgatc	900

aactacgcag acagattggc tcgaggacac tctctctag

939

<210> 18

<211> 397

<212> PRT

<213> Rep52 amino acid sequence (adeno-associated virus 2)

<400> 18

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

Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
20 25 30

Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
35 40 45

Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
50 55 60

Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
65 70 75 80

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

Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
100 105 110

Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
115 120 125

Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
130 135 140

Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
145 150 155 160

Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
165 170 175

Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
180 185 190

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

Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
210 215 220

Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln
225 230 235 240

Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val
245 250 255

Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala
260 265 270

Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val
275 280 285

Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp
290 295 300

Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu
305 310 315 320

Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys
325 330 335

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

Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr
355 360 365

Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp
370 375 380

Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln
385 390 395

<210> 19
<211> 312
<212> PRT
<213> Rep40 amino acid sequence (adeno-associated virus 2)

<400> 19

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

Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
20 25 30

Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
35 40 45

Ile Met Ser Leu Thr Lys Thr Ala Pro Asp Tyr Leu Val Gly Gln Gln
50 55 60

Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
65 70 75 80

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

Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
100 105 110

Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
115 120 125

Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
130 135 140

Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
145 150 155 160

Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
165 170 175

Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val

180

185

190

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

Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
 210 215 220

Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln
 225 230 235 240

Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val
 245 250 255

Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala
 260 265 270

Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val
 275 280 285

Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp
 290 295 300

Arg Leu Ala Arg Gly His Ser Leu
 305 310

<210> 20

<211> 1866

<212> DNA

<213> Rep78 nucleotide sequence (adeno-associated virus 2)

<400> 20

atgccgggggt tttacgagat tgtgattaag gtccccagcg accttgacgg gcatctgccc 60

ggcatttctg acagctttgt gaactgggtg gccgagaagg aatgggagtt gccgccagat 120

tctgacatgg atctgaatct gattgagcag gcacccctga ccgtggccga gaagctgcag 180

cgcgactttc tgacggaatg gcgccgtgtg agtaaggccc cggaggccct tttctttgtg 240

caatttgaga agggagagag ctacttcac atgcacgtgc tcgtggaaac caccgggggtg 300

aaatccatgg ttttgggacg tttcctgagt cagattcgcg aaaaactgat tcagagaatt 360

taccgcggga tcgagccgac tttgccaaac tggttcgcgg tcacaaagac cagaaatggc 420

gccggaggcg ggaacaaggt ggtggatgag tgctacatcc ccaattactt gctcccaaaa	480
acccagcctg agctccagtg ggcgtggact aatatggaac agtacctcag cgcctgtttg	540
aatctcacgg agcgtaaacg gttggtggcg cagcatctga cgcacgtgtc gcagacgcag	600
gagcagaaca aagagaatca gaatcccaat tctgatgcgc cggatgatcag atcaaaaact	660
tcagccaggt acatggagct ggtcgggtgg ctctgtggaca aggggattac ctcggaag	720
cagtggatcc aggaggacca ggcctcatac atctccttca atgcggcctc caactcgcgg	780
tcccaaatac aggctgcctt ggacaatgcg ggaaagatta tgagcctgac taaaaccgcc	840
cccgactacc tgggtgggcca gcagcccgtg gaggacattt ccagcaatcg gatttataaa	900
attttggaaac taaacgggta cgatcccaa tatgcggctt ccgtctttct gggatgggcc	960
acgaaaaagt tcggcaagag gaacaccatc tggctgtttg ggcctgcaac taccgggaag	1020
accaacatcg cggaggccat agcccacact gtgcccttct acgggtgcgt aaactggacc	1080
aatgagaact ttcccttcaa cgactgtgtc gacaagatgg tgatctggtg ggaggagggg	1140
aagatgaccg ccaaggtcgt ggagtcggcc aaagccattc tcggaggaag caaggtgcgc	1200
gtggaccaga aatgcaagtc ctcggccag atagaccga ctcccgtgat cgtcacctcc	1260
aacaccaaca tgtgcgccgt gattgacggg aactcaacga cttcgaaca ccagcagccg	1320
ttgcaagacc ggatgttcaa atttgaactc acccgccgtc tggatcatga ctttgggaag	1380
gtcaccaagc aggaagtcaa agactttttc cgggtgggcaa aggatcacgt ggttgagggtg	1440
gagcatgaat tctacgtcaa aaagggtgga gccaaagaaa gaccgcccc cagtgcgca	1500
gatataagt agcccaaacg ggtgcgcgag tcagttgcgc agccatcgac gtcagacgcg	1560
gaagcttcga tcaactacgc agacaggtac caaaacaaat gttctcgtca cgtgggcatg	1620
aatctgatgc tgtttccctg cagacaatgc gagagaatga atcagaattc aaatatctgc	1680
ttcactcacg gacagaaaga ctgttttagag tgctttcccg tgtcagaatc tcaaccggtt	1740
tctgtcgtca aaaaggcgta tcagaaactg tgctacattc atcatatcat gggaaagggtg	1800
ccagacgctt gactgcctg cgatctggtc aatgtggatt tggatgactg catctttgaa	1860
caatag	1866

<210> 21

<211> 1589

<212> DNA

<213> Rep68 nucleotide sequence (adeno-associated virus 2)

<400> 21

atgccgggggt tttacgagat tgtgattaag gtccccagcg accttgacga gcatctgccc	60
ggcatttctg acagctttgt gaactgggtg gccgagaagg aatgggagtt gccgccagat	120
tctgacatgg atctgaatct gattgagcag gcacccctga ccgtggccga gaagctgcag	180
cgcgactttc tgacggaatg gcgccgtgtg agtaaggccc cggaggccct tttctttgtg	240
caatttgaga agggagagag ctacttcac atgcacgtgc tcgtggaaac caccggggtg	300
aaatccatgg ttttgggacg tttcctgagt cagattcgcg aaaaactgat tcagagaatt	360
taccgcggga tcgagccgac ttgccaac tggttcgcgg tcacaaagac cagaaatggc	420
gccggaggcg ggaacaaggt ggtggatgag tgctacatcc ccaattactt gctcccaaa	480
accagcctg agctccagtg ggcgtggact aatatggaac agtacctcag cgcctgtttg	540
aatctcacgg agcgtaaagc gttggtggcg cagcatctga cgcacgtgtc gcagacgcag	600
gagcagaaca aagagaatca gaatcccaat tctgatgcgc cggatgatcag atcaaaaact	660
tcagccaggt acatggagct ggtcgggtgg ctctgggaca aggggattac ctcgagaaag	720
cagtggatcc aggaggacca ggcctcatac atctccttca atgcggcctc caactcgcg	780
tcccaaatca aggctgcctt ggacaatgcg ggaaagatta tgagcctgac taaaaccgcc	840
cccgactacc tgggtgggcca gcagcccgtg gaggacattt ccagcaatcg gatttataaa	900
atthttggaac taaacgggta cgatcccaa tatgcggctt ccgtctttct gggatgggcc	960
acgaaaaagt tcggcaagag gaacaccatc tggctgtttg ggcctgcaac taccgggaag	1020
accaacatcg cggaggccat agcccacact gtgcccttct acgggtgcgt aaactggacc	1080
aatgagaact ttcccttcaa cgactgtgtc gacaagatgg tgatctggtg ggaggagggg	1140
aagatgaccg ccaaggctgt ggagtcggcc aaagccattc tcggaggaag caaggtgcgc	1200
gtggaccaga aatgcaagtc ctcggcccag atagaccga ctcccgtgat cgtcacctcc	1260
aacaccaaca tgtgcgccgt gattgacggg aactcaacga ccttcgaaca ccagcagccg	1320
ttgcaagacc ggatgttcaa atttgaactc acccgccgtc tggatcatga ctttgggaag	1380
gtcaccaagc aggaagtcaa agactttttc cggtgggcaa aggatcacgt ggttgaggtg	1440
gagcatgaat tctacgtcaa aaagggtgga gccaaagaaa gaccgcccc cagtgcgcga	1500

gatataagtg agcccaaacg ggtgcgcgag tcagttgcgc agccatcgac gtcagacgcg 1560
gaagcttcga tcaactacgc agacagtag 1589

<210> 22
<211> 621
<212> PRT
<213> Rep78 amino acid sequence (adeno-associated virus 2)
<400> 22

Met Pro Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp
1 5 10 15

Gly His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu
20 25 30

Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile
35 40 45

Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu
50 55 60

Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
65 70 75 80

Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu
85 90 95

Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile
100 105 110

Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu
115 120 125

Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
130 135 140

Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
145 150 155 160

Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu
165 170 175

Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His
180 185 190

Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn
195 200 205

Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
210 215 220

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
225 230 235 240

Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
245 250 255

Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
260 265 270

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

Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
290 295 300

Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
305 310 315 320

Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
325 330 335

Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
340 345 350

Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
355 360 365

Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
370 375 380

Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
385 390 395 400

Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
405 410 415

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

Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
435 440 445

Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln
450 455 460

Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val
465 470 475 480

Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala
485 490 495

Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val
500 505 510

Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp
515 520 525

Arg Tyr Gln Asn Lys Cys Ser Arg His Val Gly Met Asn Leu Met Leu
530 535 540

Phe Pro Cys Arg Gln Cys Glu Arg Met Asn Gln Asn Ser Asn Ile Cys
545 550 555 560

Phe Thr His Gly Gln Lys Asp Cys Leu Glu Cys Phe Pro Val Ser Glu
565 570 575

Ser Gln Pro Val Ser Val Val Lys Lys Ala Tyr Gln Lys Leu Cys Tyr
580 585 590

Ile His His Ile Met Gly Lys Val Pro Asp Ala Cys Thr Ala Cys Asp
595 600 605

Leu Val Asn Val Asp Leu Asp Asp Cys Ile Phe Glu Gln
610 615 620

<210> 23

<211> 528

<212> PRT

<213> Rep68 amino acid sequence (adeno-associated virus 2)

<400> 23

Met Pro Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp
1 5 10 15

Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Asn Trp Val Ala Glu
20 25 30

Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile
35 40 45

Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu
50 55 60

Thr Glu Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val
65 70 75 80

Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Met His Val Leu Val Glu
85 90 95

Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile
100 105 110

Arg Glu Lys Leu Ile Gln Arg Ile Tyr Arg Gly Ile Glu Pro Thr Leu
115 120 125

Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly
130 135 140

Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys
145 150 155 160

Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Gln Tyr Leu

165

170

175

Ser Ala Cys Leu Asn Leu Thr Glu Arg Lys Arg Leu Val Ala Gln His
 180 185 190

Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Gln Asn
 195 200 205

Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr
 210 215 220

Met Glu Leu Val Gly Trp Leu Val Asp Lys Gly Ile Thr Ser Glu Lys
 225 230 235 240

Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala
 245 250 255

Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys
 260 265 270

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

Pro Val Glu Asp Ile Ser Ser Asn Arg Ile Tyr Lys Ile Leu Glu Leu
 290 295 300

Asn Gly Tyr Asp Pro Gln Tyr Ala Ala Ser Val Phe Leu Gly Trp Ala
 305 310 315 320

Thr Lys Lys Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
 325 330 335

Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Thr Val Pro
 340 345 350

Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp
 355 360 365

Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala
 370 375 380

Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg
385 390 395 400

Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val
405 410 415

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

Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
435 440 445

Glu Leu Thr Arg Arg Leu Asp His Asp Phe Gly Lys Val Thr Lys Gln
450 455 460

Glu Val Lys Asp Phe Phe Arg Trp Ala Lys Asp His Val Val Glu Val
465 470 475 480

Glu His Glu Phe Tyr Val Lys Lys Gly Gly Ala Lys Lys Arg Pro Ala
485 490 495

Pro Ser Asp Ala Asp Ile Ser Glu Pro Lys Arg Val Arg Glu Ser Val
500 505 510

Ala Gln Pro Ser Thr Ser Asp Ala Glu Ala Ser Ile Asn Tyr Ala Asp
515 520 525