

SEQUENCE LISTING

<110> ROCKET PHARMACEUTICALS, LTD.

<120> GENE THERAPY VECTORS FOR TREATMENT OF DANON DISEASE

<130> ROPA-011/01WO 329592-2071

<150> US 62/697,302

<151> 2018-07-12

<160> 38

<170> PatentIn version 3.5

<210> 1

<211> 409

<212> PRT

<213> Homo sapiens

<400> 1

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

Val Cys Leu Val Leu Gly Ala Val Arg Ser Tyr Ala Leu Glu Leu Asn
20 25 30

Leu Thr Asp Ser Glu Asn Ala Thr Cys Leu Tyr Ala Lys Trp Gln Met
35 40 45

Asn Phe Thr Val Arg Tyr Glu Thr Thr Asn Lys Thr Tyr Lys Thr Val
50 55 60

Thr Ile Ser Asp His Gly Thr Val Thr Tyr Asn Gly Ser Ile Cys Gly
65 70 75 80

Asp Asp Gln Asn Gly Pro Lys Ile Ala Val Gln Phe Gly Pro Gly Phe
85 90 95

Ser Trp Ile Ala Asn Phe Thr Lys Ala Ala Ser Thr Tyr Ser Ile Asp
100 105 110

Ser Val Ser Phe Ser Tyr Asn Thr Gly Asp Asn Thr Thr Phe Pro Asp
115 120 125

Ala Glu Asp Lys Gly Ile Leu Thr Val Asp Glu Leu Leu Ala Ile Arg
130 135 140

Ile Pro Leu Asn Asp Leu Phe Arg Cys Asn Ser Leu Ser Thr Leu Glu
145 150 155 160

Lys Asn Asp Val Val Gln His Tyr Trp Asp Val Leu Val Gln Ala Phe
165 170 175

Val Gln Asn Gly Thr Val Ser Thr Asn Glu Phe Leu Cys Asp Lys Asp
180 185 190

Lys Thr Ser Thr Val Ala Pro Thr Ile His Thr Thr Val Pro Ser Pro
195 200 205

Thr Thr Thr Pro Thr Pro Lys Glu Lys Pro Glu Ala Gly Thr Tyr Ser
210 215 220

Val Asn Asn Gly Asn Asp Thr Cys Leu Leu Ala Thr Met Gly Leu Gln
225 230 235 240

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

Asn Thr Thr His Ser Thr Gly Ser Cys Arg Ser His Thr Ala Leu Leu
260 265 270

Arg Leu Asn Ser Ser Thr Ile Lys Tyr Leu Asp Phe Val Phe Ala Val
275 280 285

Lys Asn Glu Asn Arg Phe Tyr Leu Lys Glu Val Asn Ile Ser Met Tyr
290 295 300

Leu Val Asn Gly Ser Val Phe Ser Ile Ala Asn Asn Asn Leu Ser Tyr
305 310 315 320

Trp Asp Ala Pro Leu Gly Ser Ser Tyr Met Cys Asn Lys Glu Gln Thr
325 330 335

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

Gln Pro Phe Asn Val Thr Gln Gly Lys Tyr Ser Thr Ala Gln Glu Cys
355 360 365

Ser Leu Asp Asp Asp Thr Ile Leu Ile Pro Ile Ile Val Gly Ala Gly
370 375 380

Leu Ser Gly Leu Ile Ile Val Ile Val Ile Ala Tyr Val Ile Gly Arg
385 390 395 400

Arg Lys Ser Tyr Ala Gly Tyr Gln Thr
405

<210> 2
<211> 1233
<212> DNA
<213> Homo sapiens

<400> 2
atggtgtgct tccgcctctt cccggttccg ggctcagggc tcgttctggt ctgcctagtc 60
ctgggagctg tgcggtctta tgcattggaa cttaatttga cagattcaga aaatgccact 120
tgcctttatg caaaatggca gatgaatttc acagttcgct atgaaactac aaataaaaact 180
tataaaaactg taaccatttc agaccatggc actgtgacat ataatggaag catttgtggg 240
gatgatcaga atggtcccaa aatagcagtg cagttcggac ctggcttttc ctggattgcg 300
aattttacca aggcagcatc tacttattca attgacagcg tctcattttc ctacaacact 360
ggtgataaca caacatttcc tgatgctgaa gataaaggaa ttcttactgt tgatgaactt 420
ttggccatca gaattccatt gaatgacctt tttagatgca atagtttatc aactttggaa 480
aagaatgatg ttgtccaaca ctactgggat gttcttgtac aagcttttgt ccaaaatggc 540

acagtgagca caaatgagtt cctgtgtgat aaagacaaaa cttcaacagt ggcacccacc	600
atacacacca ctgtgccatc tcctactaca acacctactc caaaggaaaa accagaagct	660
ggaacctatt cagttaataa tggcaatgat acttgtctgc tggctaccat ggggctgcag	720
ctgaacatca ctcaggataa ggttgcttca gttattaaca tcaaccccaa tacaactcac	780
tccacaggca gctgccgttc tcacactgct ctacttagac tcaatagcag caccattaag	840
tatctagact ttgtctttgc tgtgaaaaat gaaaaccgat tttatctgaa ggaagtgaac	900
atcagcatgt atttggttaa tggctccgtt ttcagcattg caaataacaa tctcagctac	960
tgggatgccc ccctgggaag ttcttatatg tgcaacaaag agcagactgt ttcagtgtct	1020
ggagcatttc agataaatac ctttgatcta agggttcagc ctttcaatgt gacacaagga	1080
aagtattcta cagcccaaga gtgttcgctg gatgatgaca ccattctaata cccaattata	1140
gttggtgctg gtctttcagg cttgattatc gttatagtga ttgcttacgt aattggcaga	1200
agaaaaagtt atgctggata tcagactctg taa	1233

<210> 3
 <211> 1233
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> variant of LAMP-2

<400> 3	
atggtctgct tcagactggt ccctgtccct ggatctggtc tggtgcttgt gtgcttggtg	60
ctgggtgctg tgagatccta tgcccttgag ctgaacctga ctgactcaga aaatgccact	120
tgcctgtatg ccaagtggca gatgaacttc actgtgagat atgagactac caacaagacc	180
tacaagactg tgaccatctc agaccatggc actgtcacct acaatggatc aatctgtggt	240
gatgatcaga atggcccaaa gatagcagtg cagtttgggc ccggtttttc ctggattgct	300
aacttcacca aggcagctc cacctacagc attgactcag tcagcttcag ctacaacact	360
ggggataaca ccaccttccc tgacgcagag gacaagggaa tccttactgt ggacgaactc	420
ctggcaatca gaatccccct taacgacctg ttcagatgca actccctttc aacccttgaa	480

aagaatgatg tgggtgcaaca ctattgggac gtcctggtgc aagcctttgt gcagaatggg	540
acagtgagta ccaacgagtt cctctgtgac aaggacaaga ccagcactgt ggccccact	600
atccacacca ctgtgcccag ccctaccact acccccaccc ctaaagagaa gccagaagct	660
ggaacctact cagtcaacaa tggaaatgac acatgcctcc ttgccaccat gggactgcag	720
ctgaacatca ctcaggacaa ggtggcctca gtgattaaca tcaaccctaa caccactcat	780
agcactggga gctgcagatc acatacagct ctgctgaggc tcaactcctc caccatcaag	840
tacctggact ttgtgtttgc tgtgaagaat gagaacaggt tctacctcaa ggaagtgaac	900
atttccatgt acctggtcaa tggttcagtg ttctctattg ccaacaacaa tctgagctac	960
tgggatgcac ccctgggatc ctcctacatg tgcaacaagg agcagactgt gagtgtgtca	1020
ggtgcttttc agatcaacac ttttgacctg aggggtgcagc cttcaatgt gactcaggga	1080
aagtactcca ctgcacaaga gtgttccttg gatgatgaca ctatcctcat cccattatt	1140
gtgggagctg gactgtcagg attgattata gtgattgtga ttgcttatgt gattggaagg	1200
agaaagagct atgctggcta ccagaccctg taa	1233

<210> 4
 <211> 1233
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> variant of LAMP-2

<400> 4	
atggtgtgct ttagactgtt tcctgtgcct gggtcagggc tggtcctggt ctgtctggtg	60
ctgggggctg tcagaagcta tgccttggag ctgaacctca ctgatagtga aaatgccact	120
tgtctgtatg ctaagtggca gatgaacttc actgtgagat atgaaaccac caacaagact	180
tacaaaacag tgaccatctc agatcatgga actgtgacct acaacggcag catttgtgga	240
gacgaccaga acggacaaaa aatcgctgtc caatttgggc ctggattctc ctggattgcc	300
aatttcacta aagctgcctc cacatattca attgactcag tgtccttctc ctacaacact	360
ggggacaaca ctactttccc tgatgctgaa gataagggaa tcttgacagt ggatgagctg	420

ctggctatca ggatcccttt gaatgacctg tttagggtgta attcactgag cactctggag	480
aagaacgacg tgggtgcagca ctactgggac gtgctgggtgc aggcccttgt gcagaacggc	540
actgtgtcca ccaacgaatt cctgtgtgat aaggacaaaa cttccactgt ggcacctaca	600
attcacacta ctgtgccttc acctaccacc actccaactc caaaggaaaa gcctgaagca	660
ggaacctact ctgtgaacaa tggcaatgat acctgtctgt tggccaccat gggcctccaa	720
ctgaacatta ctcaggacaa ggtggcctca gtgattaaca ttaaccccaa cactaccac	780
tccactggca gctgtagatc acacacagcc ttgctcagac tgaatagcag caccatcaag	840
tatttggatt ttgtgtttgc agtgaagaat gaaaacaggt tctacctgaa ggaagtcaac	900
atctcaatgt acctggtgaa cggctcagtg ttcagcattg ccaacaacaa cctctcctat	960
tgggacgctc cactggggag cagctacatg tgtaacaagg aacagactgt gtcagtgtca	1020
ggagccttcc agattaacac ctttgatctg agggccaac cttttaatgt cactcaagga	1080
aagtatagca ctgcccagga gtgctccctg gatgatgaca ccattctgat tccaatcatt	1140
gtgggtgcag gactttctgg gcttattatt gtgattgtga ttgcctatgt gattggcaga	1200
aggaaatcct atgcagggtg ccaaactctg taa	1233

<210> 5
 <211> 1233
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> variant of LAMP-2

<400> 5	
atgggtctgtt ttaggctgtt ccctgtccct gggtcaggac tggctctagt gtgtctgggtg	60
cttggagctg tcagaagcta tgccctggag ctgaacctga ctgactcaga aaatgccact	120
tgctgtatg ccaagtggca gatgaacttc actgtcagat atgaaaccac caacaagacc	180
tataagactg tgaccatctc agaccatggc actgtgactt acaatgggtc aatttgtgga	240
gatgaccaga atggccctaa gatagctgtc cagtttgggtc caggattcag ctggattgcc	300
aacttcacca aggcagccag cacctacagc attgactctg tgtccttctc ctacaacaca	360

ggagacaaca ccactttccc tgatgcagag gacaaaggta tcctgactgt ggatgagttg	420
ctggcaatca ggatcccact gaacgatctg ttcaggtgca actcactgtc cactctggaa	480
aagaatgatg tgggtgcagca ctattgggat gtgctagtcc aggcctttgt ccagaatggg	540
actgtgtcaa ctaatgagtt cctgtgtgac aaggacaaga caagcactgt agccccact	600
atccatacca cagtacctag cccaccact actccaaccc ccaaggagaa gcctgaggct	660
ggcacctact cagtgaacaa tgggaatgac acctgtttgc tggccactat gggactccaa	720
ctgaacatca cccaggacaa agtggcctct gtgatcaata tcaatcccaa caccaccac	780
agcactgggt cctgcagaag ccacactgcc ctctgaggc tcaactcatc aactatcaag	840
tacttggatt ttgtgtttgc agtgaagaat gagaacagat tctacctcaa agaggtcaac	900
atttcaatgt acctggtgaa tgggagtgtg ttctccattg ctaacaacaa cctgagctac	960
tgggatgccc ctctgggctc ctcatatg tgcaacaagg aacagactgt gagtgtgtca	1020
ggggccttcc agatcaacac ttttgacctg agagtgcagc cttttaatgt gacacaggga	1080
aagtacagca ctgctcagga gtgcagcctg gatgatgaca ctatcctgat ccctatcatt	1140
gtgggggag gcctgtctgg actcattatt gtgattgtga ttgcctatgt gatagggaga	1200
aggaagtctt atgctggata ccagaccctg taa	1233

<210> 6
 <211> 13
 <212> DNA
 <213> Unknown

<220>
 <223> Kozak sequence

<400> 6	
gccgccacca tgg	13

<210> 7
 <211> 387
 <212> DNA
 <213> Unknown

<220>
 <223> full-length polyA sequence

<400> 7
 tggctaataa aggaaattta ttttcattgc aatagtgtgt tggaattttt tgtgtctctc 60
 actcgggaagg acatatggga gggcaaatca tttaaaacat cagaatgagt atttggttta 120
 gagtttggca acatatgccc atatgctggc tgccatgaac aaaggttggc tataaagagg 180
 tcatcagtat atgaaacagc cccctgctgt ccattcctta ttccatagaa aagccttgac 240
 ttgaggttag atttttttta tattttgttt tgtgttattt ttttctttaa catccctaaa 300
 attttcctta catgttttac tagccagatt tttcctcctc tcctgactac tcccagtcac 360
 agctgtccct cttctcttat ggagatc 387

<210> 8
 <211> 4549
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> expression cassette

<400> 8
 ctgcgcgctc gctcgtcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt 60
 ggtcgcccgg cctcagtgag cgagcgagcg cgagagagg gagtggccaa ctccatcact 120
 aggggttcct tgtagttaat gattaacccg ccatgctact tatctaccag ggtaatgggg 180
 atcctctaga actatagcta gtcgacattg attattgact agttattaat agtaatcaat 240
 tacgggggtca ttagttcata gcccatatat ggagttccgc gttacataac ttacggtaaa 300
 tggcccgctt ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt 360
 tcccatagta acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta 420
 aactgcccac ttggcagtac atcaagtgtg tcatatgccca agtacgcccc ctattgacgt 480
 caatgacggg aaatggcccg cctggcatta tgcccagtac atgaccttat gggactttcc 540
 tacttggcag tacatctacg tattagtcac cgctattacc atggtcgagg tgagccccac 600
 gttctgcttc actctcccca tctccccccc ctccccaccc ccaattttgt atttatttat 660
 tttttaatta ttttgtgcag cgatggggggc ggggggggggg ggggggcgcg cgccaggcgg 720

ggcggggcgg ggcgaggggc ggggcggggc gaggcggaga ggtgcggcgg cagccaatca	780
gagcggcgcg ctccgaaagt ttctttttat ggcgagggcg cggcggcggc ggccctataa	840
aaagcgaagc gcgcggcggg cgggagtcgc tgcgcgctgc cttcgccccg tgccccgctc	900
cgccgccgcc tcgcgccgcc cggcccggt ctgactgacc gcgttactcc cacaggtagag	960
cgggcgggac ggcccttctc ctccgggctg taattagcgc ttggtttaat gacggcttgt	1020
ttcttttctg tggctgcgtg aaagccttga ggggctccgg gagggccctt tgtgcggggg	1080
gagcggctcg gggggtgcgt gcgtgtgtgt gtgcgtgggg agcgccgcgt gcggctccgc	1140
gctgccccgc ggctgtgagc gctgcgggcg cggcgcgggg ctttgtgcgc tccgcagtgt	1200
gcgcgagggg agcgcgccg ggggcgggtgc cccgcgggtgc ggggggggct gcgaggggaa	1260
caaaggctgc gtgcgggggtg tgtgcgtggg ggggtgagca gggggtgtgg gcgcgtcgg	1320
cgggctgcaa cccccctgc accccctcc ccgagttgct gagcacggcc cggcttcggg	1380
tgccgggctc cgtacggggc gtggcgcggg gctcgccgtg ccgggcgggg ggtggcggca	1440
ggtgggggtg ccgggcgggg cggggccgcc tcgggccggg gagggctcgg gggaggggcg	1500
cggcggcccc cggagcgccg gcggctgtcg aggcgcggcg agccgcagcc attgcctttt	1560
atggtaatcg tgcgagaggg cgaggggact tcctttgtcc caaatctgtg cggagccgaa	1620
atctgggagg cgccgccga cccctctag cgggcgcggg gcgaagcggg gcggcgccgg	1680
caggaaggaa atgggcgggg agggccttcg tgcgtcgccg cgccgccgtc cccttctccc	1740
tctccagcct cggggctgtc gcggggggga cggctgcctt cgggggggac ggggcagggc	1800
ggggttcggc ttctggcgtg tgaccggcgg ctctagagcc tctgctaacc atgttcatgc	1860
cttcttcttt ttcttacagc tcctgggcaa cgtgctgggt attgtgctgt ctcatcattt	1920
tggcaaagaa ttcgagcggc cgccagccgc caccatggtc tgcttcagac tgttccctgt	1980
ccctggatct ggtctgggtg ttgtgtgctt ggtgctgggt gctgtgagat cctatgccct	2040
tgagctgaac ctgactgact cagaaaaatgc cacttgccctg tatgccaaagt ggcatatgaa	2100
cttactgtg agatatgaga ctaccaacaa gacctacaag actgtgacca tctcagacca	2160
tggcactgtc acctacaatg gatcaatctg tgggtgatgat cagaatggcc caaagatagc	2220

agtgcagttt gggcccggtt tttcctggat tgctaacttc accaaggcag cctccaccta	2280
cagcattgac tcagtcagct tcagctacaa cactggggat aacaccacct tccctgacgc	2340
agaggacaag ggaatcctta ctgtggacga actcctggca atcagaatcc cccttaacga	2400
cctgttcaga tgcaactccc tttcaaccct tgaaaagaat gatgtggtgc aacactattg	2460
ggacgtcctg gtgcaagcct ttgtgcagaa tgggacagtg agtaccaacg agttcctctg	2520
tgacaaggac aagaccagca ctgtggcccc cactatccac accactgtgc ccagccctac	2580
cactaccccc acccctaaag agaagccaga agctggaacc tactcagtca acaatggaaa	2640
tgacacatgc ctccttgcca ccatgggact gcagctgaac atcactcagg acaagggtggc	2700
ctcagtgatt aacatcaacc ctaacaccac tcatagcact gggagctgca gatcacatac	2760
agctctgctg aggctcaact cctccaccat caagtacctg gactttgtgt ttgctgtgaa	2820
gaatgagaac aggttctacc tcaaggaagt gaacatttcc atgtacctgg tcaatggttc	2880
agtgttctct attgccaaca acaatctgag ctactgggat gcacccctgg gatcctccta	2940
catgtgcaac aaggagcaga ctgtgagtgt gtcaggtgct tttcagatca acacttttga	3000
cctgaggggtg cagcccttca atgtgactca gggaaagtac tccactgcac aagagtgttc	3060
cttggatgat gacactatcc tcatccccat tattgtggga gctggactgt caggattgat	3120
tatagtgatt gtgattgctt atgtgattgg aaggagaaaag agctatgctg gctaccagac	3180
cctgtaaaag ggcgaaattcc agcacacgcg tcctaggagc tcgagtacta ctggcggccg	3240
ttactagtgg atccgcggta caagtaagca tgcaagcttc gaggacgggg tgaactacgc	3300
ctgaatcaag cttatcgata aattcgagca tcttaccgcc atttattccc atatttgttc	3360
tgtttttctt gatttgggta tacatttaaa tgtaataaaa acaaaatggg ggggcaatca	3420
tttacatttt tagggatatg taattactag ttcaggtgta ttgccacaag acaaacatgt	3480
taagaaactt tcccgttatt tacgctctgt tcctgttaat caacctctgg attacaaaat	3540
ttgtgaaaga ttgactgata ttcttaacta tgttgctcct ttacgctgt gtggatatgc	3600
tgctttaatg cctctgtatc atgctattgc ttcccgtagc gctttcgttt tctcctcctt	3660
gtataaatcc tgggtgctgt ctctttatga ggagttgtgg cccgttgtcc gtcaacgtgg	3720

cgtggtgtgc tctgtgtttg ctgacgcaac cccactggc tggggcattg ccaccacctg	3780
tcaactcctt tctgggactt tcgctttccc cctcccgatc gccacggcag aactcatcgc	3840
cgcctgcctt gcccgtgct ggacaggggc taggttgctg ggcaactgata attccgtggt	3900
gttgtcgggg aagggcctcg ataccgtcga tatcgatcct ggctaataaa ggaaatttat	3960
tttcattgca atagtgtgtt ggaatttttt gtgtctctca ctcggaagga catatgggag	4020
ggcaaatcat ttaaaacatc agaatgagta tttggttttag agtttggcaa catatgcccc	4080
tatgctggct gccatgaaca aaggttggct ataaagaggt catcagtata tgaaacagcc	4140
ccctgctgtc cattccttat tccatagaaa agccttgact tgaggtaga ttttttttat	4200
attttgtttt gtgttatttt tttctttaac atccctaaaa ttttccttac atgttttact	4260
agccagattt ttcctcctct cctgactact ccagtcata gctgtccctc ttctcttatg	4320
gagatcgaag caattcgttg atctgaattt cgaccacca taatagatct ccattaccc	4380
tggtagataa gtagcatggc gggttaatca ttaactacaa ggaacccta gtgatggagt	4440
tggccactcc ctctctgcgc gctcgctcgc tctactgaggc cgggcgacca aaggtcgccc	4500
gacgcccggg ctttgcccgg gcggcctcag tgagcgagcg agcgcgag	4549

<210> 9
 <211> 4549
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> expression cassette

<400> 9	
ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt	60
ggtcgccccg cctcagttag cgagcgagcg cgcagagagg gaggggccaa ctccatcact	120
aggggttcct ttagttaat gattaaccg ccatgctact tatctaccag ggtaatgggg	180
atcctctaga actatagcta gtcgacattg attattgact agttattaat agtaatcaat	240
tacgggggtca ttagttcata gcccatatat ggagttccgc gttacataac ttacggtaaa	300
tggcccgctt ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt	360

tcccatagta acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta	420
aactgcccac ttggcagtac atcaagtgtat tcatatgccca agtacgcccc ctattgacgt	480
caatgacgggt aaatggcccc cctggcatta tgcccagtac atgaccttat gggactttcc	540
tacttggcag tacatctacg tattagtcatt cgctattacc atggctcagg tgagccccac	600
gttctgcttc actctcccca tctccccccc ctccccaccc ccaattttgt atttatttat	660
tttttaatta ttttgtgcag cgatgggggc gggggggggg ggggggcgcg cgccaggcgg	720
ggcggggcgcg ggcgaggggc ggggcggggc gaggcggaga ggtgcggcgg cagccaatca	780
gagcggcgcg ctccgaaagt ttccttttat ggcgaggcgg cggcggcggc ggccctataa	840
aaagcgaagc gcgcggcggg cgggagtcgc tgcgcgctgc cttcgccccg tgccccgctc	900
cgccgcccgc tcgcgcgcgc cgccccggct ctgactgacc gcgttactcc cacaggtag	960
cgggcgggac ggcccttctc ctccgggctg taattagcgc ttggtttaat gacggcttgt	1020
ttcttttctg tggctgcgtg aaagccttga ggggctccgg gagggccctt tgtgcggggg	1080
gagcggctcg gggggtgcgt gcgtgtgtgt gtgcgtgggg agcgccgcgt gcggctccgc	1140
gctgcccggc ggctgtgagc gctgcgggcg cggcgcgggg ctttgtgcgc tccgcagtgt	1200
gcgcgagggg agcgcgccg ggggcgggtgc cccgcgggtgc ggggggggct gcgaggggaa	1260
caaaggctgc gtgcgggggtg tgtgcgtggg ggggtgagca ggggtgtgg gcgcgtcgggt	1320
cgggctgcaa cccccctgc accccccctcc ccgagttgct gagcacggcc cggcttcggg	1380
tgccggggctc cgtacggggc gtggcgcggg gctcgccgtg ccgggcgggg ggtggcggca	1440
ggtgggggtg ccgggcgggg cggggccgcc tcgggcccgg gagggctcgg gggaggggcg	1500
cggcggcccc cggagcgccg gcggctgtcg aggcgcggcg agccgcagcc attgcctttt	1560
atggtaatcg tgcgagaggg gcgagggact tcctttgtcc caaatctgtg cggagccgaa	1620
atctgggagg cgccgccga cccctctag cgggcgcggg gcgaagcgggt gcggcgccgg	1680
caggaaggaa atgggcgggg agggccttcg tgcgtcgccg cgccgccgtc cccttctccc	1740
tctccagcct cggggctgtc cgcgggggga cggctgcctt cgggggggac ggggcagggc	1800
gggggttcggc ttctggcgtg tgaccggcgg ctctagagcc tctgctaacc atgttcatgc	1860

cttcttctttt ttcctacagc tcctgggcaa cgtgctgggtt attgtgctgt ctcatcattt	1920
tggcaaagaa ttcgagcggc cgccagccgc caccatgggtg tgcttttagac tgtttcctgt	1980
gcctgggttca gggctgggtcc tggctctgtct ggtgctgggg gctgtcagaa gctatgcctt	2040
ggagctgaac ctactgata gtgaaaatgc cacttgtctg tatgctaagt ggcagatgaa	2100
cttcactgtg agatatgaaa ccaccaacaa gacttacaaa acagtgacca tctcagatca	2160
tggaaactgtg acctacaacg gcagcatttg tggagacgac cagaacggac caaaaatcgc	2220
tgtccaattt gggcctggat tctcctggat tgccaatttc actaaagctg cctccacata	2280
ttcaattgac tcagtgtcct tctcctacaa cactggggac aacactactt tccctgatgc	2340
tgaagataag ggaatcttga cagtggatga gctgctggct atcaggatcc ctttgaatga	2400
cctgttttagg tgtaattcac tgagcactct ggagaagaac gacgtgggtgc agcactactg	2460
ggacgtgctg gtgcaggcct ttgtgcagaa cggcactgtg tccaccaacg aattcctgtg	2520
tgataaggac aaaacttcca ctgtggcacc tacaattcac actactgtgc cttcacctac	2580
caccactcca actccaaagg aaaagcctga agcaggaacc tactctgtga acaatggcaa	2640
tgatacctgt ctgttggcca ccatgggcct ccaactgaac attactcagg acaagggtggc	2700
ctcagtgatt aacattaacc ccaacactac ccactccact ggcagctgta gatcacacac	2760
agccttgctc agactgaata gcagcaccat caagtatttg gatthttgtgt ttgcagtgaa	2820
gaatgaaaac aggttctacc tgaaggaagt caacatctca atgtacctgg tgaacggctc	2880
agtgttcagc attgccaaca acaacctctc ctattgggac gctccactgg ggagcagcta	2940
catgtgtaac aaggaacaga ctgtgtcagt gtcaggagcc ttccagatta acacctttga	3000
tctgagggtc caacccttta atgtcactca aggaaaagtat agcactgccc aggagtgtc	3060
cctggatgat gacaccattc tgattccaat cattgtgggt gcaggacttt ctgggcttat	3120
tattgtgatt gtgattgcct atgtgattgg cagaaggaaa tcctatgcag ggtaccaaac	3180
tctgtaaaag ggcgaaattcc agcacacgcg tcctaggagc tcgagtacta ctggcgccg	3240
ttactagtgg atccgcggta caagtaagca tgcaagcttc gaggacgggg tgaactacgc	3300
ctgaatcaag cttatcgata aattcgagca tcttaccgcc atttattccc atatttgttc	3360

tgtttttctt gatttgggta tacatttaaa tgtaataaaa aaaaaatggt ggggcaatca	3420
tttacatttt tagggatatg taattactag ttcagggtga ttgccacaag acaaacatgt	3480
taagaaactt tcccgttatt tacgctctgt tcctgttaat caacctctgg attacaaaat	3540
tttgaaaga ttgactgata ttcttaacta tgttgctcct ttacgctgt gtggatatgc	3600
tgctttaatg cctctgtatc atgctattgc ttcccgtagc gctttcgttt tctcctcctt	3660
gtataaatcc tggttgctgt ctctttatga ggagttgtgg cccgttgctc gtcaacgtgg	3720
cgtggtgtgc tctgtgtttg ctgacgcaac cccactggc tggggcattg ccaccacctg	3780
tcaactcctt tctgggactt tcgctttccc cctcccgatc gccacggcag aactcatcgc	3840
cgctgcctt gcccgtgct ggacaggggc taggttgctg ggactgata attccgtggt	3900
gtgtcgggg aagggcctcg ataccgtcga tatcgatcct ggctaataaa ggaaatttat	3960
tttcattgca atagtgtgtt ggaatttttt gtgtctctca ctcggaagga catatgggag	4020
ggcaaatcat ttaaaacatc agaatgagta tttggtttag agtttgcaa catatgcca	4080
tatgctggct gccatgaaca aaggttggct ataaagaggt catcagtata tgaaacagcc	4140
ccctgctgtc cattccttat tccatagaaa agccttgact tgaggttaga ttttttttat	4200
attttgtttt gtgttatttt tttctttaac atccctaaaa ttttccttac atgttttact	4260
agccagattt ttcctcctct cctgactact ccagtcata gctgtccctc ttctcttatg	4320
gagatcgaag caattcgttg atctgaattt cgaccacca taatagatct ccattatccc	4380
tggtagataa gtagcatggc gggttaatca ttaactacaa ggaacccta gtgatggagt	4440
tggccactcc ctctctgcgc gctcgtcgc tactgaggc cgggcgacca aaggctgccc	4500
gacggcggg ctttgcccgg gcggcctcag tgagcgagcg agcgcgag	4549

<210> 10
 <211> 4549
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> expression cassette

 <400> 10

ctgcgcgctc gctcgtcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt	60
ggtcgcccgg cctcagttag cgagcgagcg cgagagagg gaggggccaa ctccatcact	120
aggggttcct tgtagttaat gattaacccg ccatgctact tatctaccag ggtaatgggg	180
atcctctaga actatagcta gtcgacattg attattgact agttattaat agtaatcaat	240
tacgggggtca ttagttcata gcccataatat ggagttccgc gttacataac ttacggtaaa	300
tggcccgctt ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt	360
tcccatagta acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta	420
aactgcccac ttggcagtag atcaagtgtg tcatatgccg agtacgcccc ctattgacgt	480
caatgacggg aaatggcccg cctggcatta tgcccagtag atgaccttat gggactttcc	540
tacttggcag tacatctacg tattagtcac cgctattacc atggctgagg tgagccccac	600
gttctgcttc actctcccca tctccccccc ctccccaccc ccaattttgt atttatttat	660
tttttaatta ttttgtgcag cgatgggggc gggggggggg gggggggcgc cgccaggcgg	720
ggcggggcgc ggcgaggggc ggggcggggc gaggcggaga ggtgcggcgc cagccaatca	780
gagcggcgcg ctccgaaagt ttccttttat ggcgagggcgc cggcggcggc ggccctataa	840
aaagcgaagc gcgcggcggg cgggagtcgc tgcgcgctgc cttcgccccg tgccccgctc	900
cgccgccgcc tcgcgccgcc cgccccggct ctgactgacc gcgttactcc cacaggtag	960
cgggcgggac ggcccttctc ctccgggctg taattagcgc ttggtttaat gacggcttgt	1020
ttcttttctg tggctgcgtg aaagccttga ggggctccgg gagggccctt tgtgcggggg	1080
gagcggctcg gggggtgcgt gcgtgtgtgt gtgcgtgggg agcgccgcgt gcggctccgc	1140
gctgcccggc ggctgtgagc gctgcgggcg cggcgcgggg ctttgtgcgc tccgagtg	1200
gcgcgagggg agcgcgcccg ggggcgggtgc cccgcgggtgc ggggggggct gcgaggggaa	1260
caaaggctgc gtgcggggtg tgtgcgtggg ggggtgagca gggggtgtgg gcgcgtcgg	1320
cgggctgcaa cccccctgc accccccctc ccgagttgct gagcacggcc cggcttcggg	1380
tgccgggctc cgtacggggc gtggcgcggg gctcgccgtg ccgggcgggg ggtggcggca	1440
ggtgggggtg ccgggcgggg cggggccgcc tcgggcccgg gagggctcgg gggaggggcg	1500

cggcggcccc cggagcgccg gcggctgtcg aggcgcggcg agccgcagcc attgcctttt	1560
atggtaatcg tgcgagaggg cgaggggact tcctttgtcc caaatctgtg cggagccgaa	1620
atctgggagg cgccgccgca cccctctag cgggcgcggg gcgaagcggg gcggcgccgg	1680
caggaaggaa atgggcgggg agggccttcg tgcgtcgccg cgccgccgtc cccttctccc	1740
tctccagcct cggggctgtc cgcgggggga cggctgcctt cgggggggac ggggcagggc	1800
ggggttcggc ttctggcgtg tgaccggcgg ctctagagcc tctgctaacc atgttcatgc	1860
cttcttcttt ttctacagc tcctgggcaa cgtgctgggtt attgtgctgt ctcatcattt	1920
tggcaaagaa ttcgagcggc cgccagccgc caccatggtc tgttttaggc tgttcctgt	1980
ccctgggtca ggactggctt tagtgtgtct ggtgcttgga gctgtcagaa gctatgccct	2040
ggagctgaac ctgactgact cagaaaatgc cacttgccctg tatgccaagt ggcagatgaa	2100
cttcactgtc agatatgaaa ccaccaacaa gacctataag actgtgacca tctcagacca	2160
tggcactgtg acttacaatg ggtcaatttg tggagatgac cagaatggcc ctaagatagc	2220
tgtccagttt ggtccaggat tcagctggat tgccaacttc accaaggcag ccagcaccta	2280
cagcattgac tctgtgtcct tctcctacaa cacaggagac aacaccactt tccctgatgc	2340
agaggacaaa ggtatcctga ctgtggatga gttgctggca atcaggatcc cactgaacga	2400
tctgttcagg tgcaactcac tgtccactct ggaaaagaat gatgtgggtgc agcactattg	2460
ggatgtgcta gtccaggcct ttgtccagaa tgggactgtg tcaactaatg agttcctgtg	2520
tgacaaggac aagacaagca ctgtagcccc cactatccat accacagtac ctagccccac	2580
cactactcca accccaagg agaagcctga ggctggcacc tactcagtga acaatgggaa	2640
tgacacctgt ttgctggcca ctatgggact ccaactgaac atcaccagg acaaagtggc	2700
ctctgtgatc aatatcaatc ccaacaccac ccacagcact gggtcctgca gaagccacac	2760
tgccctcctg aggctcaact catcaactat caagtacttg gattttgtgt ttgcagtga	2820
gaatgagaac agattctacc tcaaagaggt caacatttca atgtacctgg tgaatgggag	2880
tgtgttctcc attgctaaca acaacctgag ctactgggat gcccctctgg gtcctcata	2940
catgtgcaac aaggaacaga ctgtgagtgt gtcaggggcc ttccagatca acacttttga	3000

cctgagagtg cagcccttta atgtgacaca gggaaaagtac agcactgctc aggagtgcag	3060
cctggatgat gacactatcc tgatccctat cattgtgggg gcaggcctgt ctggactcat	3120
tattgtgatt gtgattgcct atgtgatagg gagaaggaag tcttatgctg gataccagac	3180
cctgtaaaag ggcgaattcc agcacacgcg tcctaggagc tcgagtacta ctggcggccg	3240
ttactagtgg atccgcggta caagtaagca tgcaagcttc gaggacgggg tgaactacgc	3300
ctgaatcaag cttatcgata aattcgagca tcttaccgcc atttattccc atatttgttc	3360
tgtttttctt gatttgggta tacatttaaa tgtaataaaa acaaaatggg ggggcaatca	3420
tttacatttt tagggatatg taattactag ttcagggtga ttgccacaag acaaacatgt	3480
taagaaactt tcccgttatt tacgctctgt tcctgttaat caacctctgg attacaaaaat	3540
ttgtgaaaga ttgactgata ttcttaacta tgttgctcct ttacgctgt gtggatatgc	3600
tgctttaatg cctctgtatc atgctattgc ttcccgtacg gctttcgttt tctcctcctt	3660
gtataaatcc tggttgctgt ctctttatga ggagttgtgg cccgttgtcc gtcaacgtgg	3720
cgtggtgtgc tctgtgtttg ctgacgcaac cccactggc tggggcattg ccaccacctg	3780
tcaactcctt tctgggactt tcgctttccc cctcccgatc gccacggcag aactcatcgc	3840
cgctgcctt gcccgctgct ggacaggggc taggttgctg ggactgata attccgtggg	3900
gttgtcgggg aagggcctcg ataccgtcga tatcgatcct ggctaataaaa ggaaatttat	3960
tttcattgca atagtgtgtt ggaatttttt gtgtctctca ctcggaagga catatgggag	4020
ggcaaactcat ttaaaacatc agaatgagta tttggtttag agtttggcaa catatgcca	4080
tatgctggct gccatgaaca aaggttggct ataaagaggt catcagtata tgaaacagcc	4140
ccctgctgtc cattccttat tccatagaaa agccttgact tgaggtaga ttttttttat	4200
attttgtttt gtgttatttt tttctttaac atccctaaaa ttttccttac atgttttact	4260
agccagatth ttcctcctct cctgactact cccagtcata gctgtccctc ttctcttatg	4320
gagatcgaag caattcgttg atctgaattt cgaccacca taatagatct ccattatccc	4380
tggtagataa gtagcatggc gggttaatca ttaactacaa ggaacccta gtgatggagt	4440
tggccactcc ctctctgcgc gctcgtcgc tcactgaggc cgggcgacca aaggtcgccc	4500

gacgcccggg ctttgcccgg gcggcctcag tgagcgagcg agcgcgag 4549

<210> 11
<211> 130
<212> DNA
<213> Unknown

<220>
<223> inverted terminal repeat (ITR) sequence

<400> 11
ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt 60
ggtcgcccgg cctcagtgag cgagcgagcg cgagagagg gaggggccaa ctccatcact 120
aggggttcct 130

<210> 12
<211> 130
<212> DNA
<213> Unknown

<220>
<223> inverted terminal repeat (ITR) sequence

<400> 12
aggaaccctt agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcactgagg 60
ccgggcgacc aaaggctgcc cgacgcccgg gctttgcccg ggccggcctca gtgagcgagc 120
gagcgcgagc 130

<210> 13

<400> 13
000

<210> 14
<211> 13
<212> RNA
<213> Unknown

<220>
<223> alternative Kozak sequence

<220>

<221> misc_feature
<222> (1)..(3)
<223> May be absent

<220>
<221> misc_feature
<222> (1)..(1)
<223> g is the most common base, but may vary

<220>
<221> misc_feature
<222> (2)..(3)
<223> c is the most common base, but may vary

<220>
<221> misc_feature
<222> (4)..(4)
<223> g is the most common base, but may vary

<220>
<221> misc_feature
<222> (5)..(6)
<223> c is the most common base, but may vary

<220>
<221> misc_feature
<222> (8)..(9)
<223> c is the most common base, but may vary

<400> 14
nnnnnnrnna ugg

13

<210> 15
<211> 8
<212> RNA
<213> Unknown

<220>
<223> alternative Kozak sequence

<220>
<221> misc_feature
<222> (3)..(4)
<223> n is a, c, g, or u

<220>
<221> misc_feature
<222> (8)..(8)

<223> n is a, c, g, or u

<400> 15

agnnaugn

8

<210> 16

<211> 7

<212> RNA

<213> Unknown

<220>

<223> alternative Kozak sequence

<220>

<221> misc_feature

<222> (2)..(3)

<223> n is a, c, g, or u

<400> 16

annaugg

7

<210> 17

<211> 7

<212> RNA

<213> Unknown

<220>

<223> alternative Kozak sequence

<400> 17

accaugg

7

<210> 18

<211> 10

<212> RNA

<213> Unknown

<220>

<223> alternative Kozak sequence

<400> 18

gacaccaugg

10

<210> 19

<211> 235

<212> DNA
 <213> Bos taurus

 <400> 19
 tcgactgtgc cttctagttg ccagccatct gttgtttgcc cctcccccggt gccttccttg 60
 accctggaag gtgccactcc cactgtcctt tcctaataaa atgaggaaat tgcacgcgat 120
 tgtctgagta ggtgtcattc tattctgggg ggtgggggtgg ggaggacag caagggggag 180
 gattgggagg acaatagcag gcatgctggg gatgcgggtgg gctctatggc ttctg 235

<210> 20
 <211> 222
 <212> DNA
 <213> Simian virus 40

 <400> 20
 cagacatgat aagatacatt gatgagtttg gacaaaccac aactagaatg cagtgaaaaa 60
 aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca 120
 ataaacaagt taacaacaac aattgcattc attttatgtt tcaggttcag ggggagatgt 180
 gggaggtttt ttaaagcaag taaaacctct acaaatgtgg ta 222

<210> 21
 <211> 202
 <212> DNA
 <213> Homo sapiens

 <400> 21
 ctgcccgggt ggcatccctg tgacccctcc ccagtgcctc tcctggccct ggaagttgcc 60
 actccagtgc ccaccagcct tgcctaata aaattaagtt gcatcatttt gtctgactag 120
 gtgtccttct ataataattat ggggtggagg ggggtggat ggagcaaggg gcccaagttg 180
 ggaagaaacc tgtaggcct gc 202

<210> 22
 <211> 1730
 <212> DNA
 <213> Unknown

<220>
 <223> CAG promoter

<400> 22

ctagtcgaca ttgattattg actagttatt aatagtaatc aattacgggg tcattagttc	60
atagcccata tatggagttc cgcgttacat aacttacggt aaatggcccg cctggctgac	120
cgcccaacga ccccgccca ttgacgtcaa taatgacgta tgttcccata gtaacgccaa	180
tagggacttt ccattgacgt caatgggtgg agtatttacg gtaaactgcc cacttggcag	240
tacatcaagt gtatcatatg ccaagtacgc cccctattga cgtcaatgac ggtaaatggc	300
ccgcctggca ttatgcccag tacatgacct tatgggactt tcctacttgg cagtacatct	360
acgtattagt catcgctatt accatggtcg aggtgagccc cacgttctgc ttcactctcc	420
ccatctcccc cccctcccca cccccaattt tgtatttatt tattttttaa ttattttgtg	480
cagcgatggg ggcggggggg gggggggggc gcgcgccagg cggggcgggg cggggcgagg	540
ggcggggagg ggcgaggcgg agaggtgcgg cggcagccaa tcagagcggc gcgctccgaa	600
agtttccttt tatggcgagg cggcgggcgg ggcgggcccta taaaaagcga agcgcgcggc	660
gggcgggagt cgctgcgcgc tgccttcgcc ccgtgccccg ctccgccgcc gcctcgcgcc	720
gccccccccg gctctgactg accgcgttac tcccacaggt gagcgggcgg gacggccctt	780
ctcctccggg ctgtaattag cgcttggttt aatgacggct tgtttctttt ctgtggctgc	840
gtgaaagcct tgaggggctc cgggagggcc ctttgtgcgg ggggagcggc tcgggggggtg	900
cgtgcgtgtg tgtgtgcgtg gggagcgccg cgtgcggctc cgcgctgccc ggcggtgtg	960
agcgctgcgg gcgcggcgcg gggctttgtg cgctccgcag tgtgcgcgag gggagcgcg	1020
ccgggggagg tgccccgcgg tgcggggggg gctgcgaggg gaacaaaggc tgcgtgcggg	1080
gtgtgtgcgt ggggggggtga gcaggggggtg tgggcgcgtc ggtcgggctg caaccccccc	1140
tgcaccccc tccccgagtt gctgagcacg gcccggcttc gggcggggg ctccgtacgg	1200
ggcgtggcgc ggggctcgcc gtgccgggcg gggggtggcg gcaggcgggg gtgccgggcg	1260
gggcggggcc gcctcgggcc ggggagggtc cgggggaggg gcgcggcggc ccccgagcgc	1320
ccggcggctg tcgaggcgcg gcgagccgca gccattgcct tttatggtaa tcgtgcgaga	1380
gggcgcaggg acttcctttg tcccaaactc gtgcggagcc gaaatctggg aggcgccgcc	1440

gcacccccctc tagcgggcgc ggggcgaagc ggtgcggcgc cggcaggaag gaaatgggcg	1500
gggagggcct tcgtgcgtcg ccgcgccgcc gtcccccttct ccctctccag cctcggggct	1560
gtccgcggggg ggacggctgc cttcggggggg gacggggcag ggcgggggttc ggcttctggc	1620
gtgtgaccgg cggctctaga gcctctgcta accatgttca tgccttcttc tttttcctac	1680
agctcctggg caacgtgctg gttattgtgc tgtctcatca ttttggcaaa	1730

<210> 23
 <211> 204
 <212> DNA
 <213> Human cytomegalovirus

<400> 23	
gtgatgcggg tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt	60
ccaagtctcc accccattga cgtcaatggg agtttgtttt ggccacaaaa tcaacgggac	120
tttccaaaat gtcgtaacaa ctccgccccca ttgacgcaaa tgggcggtag gcgtgtacgg	180
tgggaggtct atataagcag agct	204

<210> 24
 <211> 317
 <212> DNA
 <213> Simian virus 40

<400> 24	
ggtgtggaaa gtccccaggc tccccagcag gcagaagtat gcaaagcatg catctcaatt	60
agtcagcaac caggtgtgga aagtccccag gctccccagc aggcagaagt atgcaaagca	120
tgcattctcaa ttagtcagca accatagtcc cgcccctaac tccgcccata ccgcccctaa	180
ctccgcccag ttccgcccata tctccgcccc atggctgact aatttttttt atttatgcag	240
aggccgaggc cgcctcggcc tctgagctat tccagaagta gtgaggaggc ttttttggag	300
gcctaggctt ttgcaaa	317

<210> 25
 <211> 500
 <212> DNA
 <213> Unknown

<220>

<223> PGK promoter

<400> 25

gggtagggga	ggcgcttttc	ccaaggcagt	ctggagcatg	cgcttttagca	gccccgctgg	60
gcacttggcg	ctacacaagt	ggcctctggc	ctcgcacaca	ttccacatcc	accggtaggc	120
gccaaccggc	tccgttcttt	ggtggccccct	tcgcgccacc	ttctactcct	cccctagtca	180
ggaagtcccc	ccccgccccg	cagctcgcgt	cgtgcaggac	gtgacaaatg	gaagtagcac	240
gtctcactag	tctcgtgcag	atggacagca	ccgctgagca	atggaagcgg	gtaggccttt	300
ggggcagcgg	ccaatagcag	ctttgctcct	tcgctttctg	ggctcagagg	ctgggaaggg	360
gtgggtccgg	gggcgggctc	aggggcgggc	tcaggggcgg	ggcgggcgcc	cgaaggctct	420
ccggaggccc	ggcattctgc	acgcttcaaa	agcgcacgtc	tgccgcgctg	ttctcctctt	480
cctcatctcc	gggcctttcg					500

<210> 26

<211> 2403

<212> DNA

<213> Homo sapiens

<400> 26

cctgcagggc	ccactagttc	catgtcctta	tatggactca	tctttgccta	ttgcgacaca	60
cactcaatga	acacctacta	cgcgctgcaa	agagccccgc	aggcctgagg	tgccccacc	120
tcaccactct	tcctatTTTT	gtgtaaaaat	ccagcttctt	gtcaccacct	ccaaggaggg	180
ggaggaggag	gaaggcaggt	tcctctaggc	tgagccgaat	gcccctctgt	ggtcccacgc	240
cactgatcgc	tgcatgcccc	ccacctgggt	acacacagtc	tgtgattccc	ggagcagaac	300
ggaccctgcc	cacccggctc	tgtgtgctac	tcagtggaca	gacccaaggc	aagaaaagggt	360
gacaaggaca	gggtcttccc	aggctggctt	tgagttccta	gcaccgcccc	gcccccaatc	420
ctctgtggca	catggagtct	tgggtccccag	agtccccag	cggcctccag	atggtctggg	480
agggcagttc	agctgtggct	gcgcatagca	gacatacaac	ggacgggtggg	cccagaccca	540
ggctgtgtag	acccagcccc	cccgccccgc	agtgccctagg	tcacccacta	acgccccagg	600
cctggtcttg	gctgggcgtg	actgttaccc	tcaaaagcag	gcagctccag	ggtaaaagggt	660

gccctgccct gtagagccca ccttccttcc cagggtcg gctgggtagg tttgtagcct	720
tcatcacggg ccacctccag cactggacc gctggcccct gccctgtcct ggggagtgtg	780
gtcctgcgac ttctaagtgg ccgcaagcca cctgactccc ccaacaccac actctacctc	840
tcaagcccag gtctctccct agtgaccac ccagcacatt tagctagctg agccccacag	900
ccagagggtcc tcaggccctg ctttcagggc agttgctctg aagtcggcaa gggggagtga	960
ctgcctggcc actccatgcc ctccaagagc tccttctgca ggagcgtaca gaacccaggg	1020
ccctggcacc cgtgcagacc ctggcccacc ccacctgggc gctcagtgcc caagagatgt	1080
ccacacctag gatgtcccg cgtgggtggg gggcccgaga gacgggcagg ccgggggcag	1140
gcctggccat gcggggccga accgggcact gcccagcgtg gggcgcgggg gccacggcgc	1200
gcgccccag ccccgggcc cagcaccca aggcggccaa cgccaaaact ctccctcctc	1260
ctcttcctca atctcgtct cgtcttttt ttttttcgca aaaggagggg agagggggta	1320
aaaaaatgct gcactgtgcg gcgaagccgg tgagtgcg gcgcggggcc aatcagcgtg	1380
cgccgttccg aaagttgcct tttatggctc gagcggccgc ggcggcgccc tataaaaccc	1440
agcggcgca cgcgccacca ccgccgagac cgcgtccgcc ccgcgagcac agagcctcgc	1500
ctttgccgat ccgccccc tccacacccg ccgccaggta agcccggcca gccgaccggg	1560
gcaggcggct cacggcccgg ccgcaggcgg ccgcggcccc ttgcccgtg cagagccgcc	1620
gtctgggccc cagcgggggg cgcatggggg gggaaccgga ccgccgtggg gggcgcgggga	1680
gaagcccctg ggcctccgga gatgggggac accccacgcc agttcggagg cgcgaggccg	1740
cgctcgggag gcgcgctccg ggggtgccgc tctcggggcg ggggcaaccg gcggggtctt	1800
tgtctgagcc gggctcttgc caatggggat cgcagggtgg gcgcggcgga gccccgcca	1860
ggcccgggtgg gggctggggc gccattgcgc gtgcgcgctg gtcctttggg cgctaactgc	1920
gtgcgcgctg ggaattggcg ctaattgcgc gtgcgcgctg ggactcaagg cgctaactgc	1980
gcgtgcgttc tggggcccgg ggtgccgcgg cctgggctgg ggcgaaggcg ggctcggccg	2040
gaaggggtgg ggtcgccgc gctccgggc gcttgccgc acttcctgcc cgagccgctg	2100
gccgcccag ggtgtggccg ctgcgtgcgc gcgcgccgac ccggcgctgt ttgaaccggg	2160

cggaggcggg gctggcgccc gggtgggagg ggggtggggc ctggcttcct gccgcgcgcc	2220
gcgggggacgc ctccgaccag tgtttgcctt ttatggtaat aacgcggccg gcccggcttc	2280
ctttgtcccc aatctgggcg cgcgccggcg ccccttggcg gcctaaggac tcggcgcgcc	2340
ggaagtggcc agggcggggg cgacctcggc tcacagcgcg cccggctatt ctcgcagctc	2400
acc	2403

<210> 27
 <211> 596
 <212> DNA
 <213> Woodchuck hepatitis virus

<400> 27	
attcgagcat cttaccgcca ttattccca tatttgttct gtttttcttg atttgggtat	60
acatttaaataa gttaataaaa caaaatgggtg gggcaatcat ttacattttt agggatatgt	120
aattactagt tcagggtgat tgccacaaga caaacatgtt aagaaacttt cccgttat	180
acgctctgtt cctgttaatc aacctctgga ttacaaaatt tgtgaaagat tgactgat	240
tcttaactat gttgctcctt ttacgctgtg tggatatgct gctttaatgc ctctgtatca	300
tgctattgct tcccgtacgg ctttcgtttt ctctccttg tataaatcct ggttgctgtc	360
tctttatgag gagttgtggc ccgttgtccg tcaacgtggc gtggtgtgct ctgtgtttgc	420
tgacgcaacc cccactggct ggggcattgc caccacctgt caactccttt ctgggacttt	480
cgctttcccc ctccgatcg ccacggcaga actcatcgcc gcctgccttg cccgctgctg	540
gacaggggct aggttgctgg gcactgataa ttccgtgggtg ttgtcgggga agggcc	596

<210> 28
 <211> 736
 <212> PRT
 <213> adeno-associated virus 9

<400> 28

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

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

Lys Ala Asn Gln Gln His Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
35 40 45

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

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

Gln Gln Leu Lys Ala 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 Leu Leu Glu Pro
115 120 125

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

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ala Gly Ile Gly
145 150 155 160

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

Gly Asp Thr Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro Pro
180 185 190

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

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

Ser Gly Asn Trp His Cys Asp Ser Gln Trp Leu 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 Asn Ser Thr Ser Gly Gly Ser Ser Asn Asp Asn
260 265 270

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

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

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

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

Asn Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Asp Tyr Gln Leu
340 345 350

Pro Tyr Val Leu Gly Ser Ala His Glu Gly Cys Leu Pro Pro Phe Pro
355 360 365

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asp
370 375 380

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

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

Phe Glu Asn Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
420 425 430

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

Lys Thr Ile Asn Gly Ser Gly Gln Asn Gln Gln Thr Leu Lys Phe Ser
450 455 460

Val Ala Gly Pro Ser Asn Met Ala Val Gln Gly Arg Asn Tyr Ile Pro
465 470 475 480

Gly Pro Ser Tyr Arg Gln Gln Arg Val Ser Thr Thr Val Thr Gln Asn
485 490 495

Asn Asn Ser Glu Phe Ala Trp Pro Gly Ala Ser Ser Trp Ala Leu Asn
500 505 510

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

Glu Gly Glu Asp Arg Phe Phe Pro Leu Ser Gly Ser Leu Ile Phe Gly
530 535 540

Lys Gln Gly Thr Gly Arg Asp Asn Val Asp Ala Asp Lys Val Met Ile
545 550 555 560

Thr Asn Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr Glu Ser
565 570 575

Tyr Gly Gln Val Ala Thr Asn His Gln Ser Ala Gln Ala Gln Ala Gln
580 585 590

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

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

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

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

Asp Pro Pro Thr Ala Phe Asn Lys Asp Lys Leu Asn Ser Phe Ile Thr
660 665 670

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

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

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

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

<210> 29
<211> 1233
<212> DNA
<213> Homo sapiens

<400> 29
atggtgtgct tccgcctctt cccggttccg ggctcagggc tcgttctggt ctgcctagtc 60
ctgggagctg tgcggtctta tgcattggaa cttaatttga cagattcaga aaatgccact 120
tgccttttatg caaaatggca gatgaatttc acagtacgct atgaaactac aaataaaact 180
tataaaactg taaccatttc agaccatggc actgtgacat ataatggaag catttgtggg 240
gatgatcaga atggtcccaa aatagcagtg cagttcggac ctggcttttc ctggattgcg 300
aattttacca aggcagcatc tacttattca attgacagcg tctcattttc ctacaacact 360
ggtgataaca caacatttcc tgatgctgaa gataaaggaa ttcttactgt tgatgaactt 420

ttggccatca gaattccatt gaatgacctt tttagatgca atagtttatc aacttttgaa	480
aagaatgatg ttgtccaaca ctactgggat gttcttgtac aagcttttgt ccaaaatggc	540
acagtgagca caaatgagtt cctgtgtgat aaagacaaaa cttcaacagt ggcacccacc	600
atacacacca ctgtgccatc tcctactaca acacctactc caaaggaaaa accagaagct	660
ggaacctatt cagttaataa tggcaatgat acttgtctgc tggctaccat ggggctgcag	720
ctgaacatca ctcaggataa ggttgcttca gttattaaca tcaaccccaa tacaactcac	780
tccacaggca gctgccgttc tcacactgct ctacttagac tcaatagcag caccattaag	840
tatctagact ttgtctttgc tgtgaaaaat gaaaaccgat tttatctgaa ggaagtgaac	900
atcagcatgt atttggttaa tggctccgtt ttcagcattg caaataacaa tctcagctac	960
tgggatgccc ccctgggaag ttcttatatg tgcaacaaag agcagactgt ttcagtgtct	1020
ggagcatttc agataaatac ctttgatcta agggttcagc ctttcaatgt gacacaagga	1080
aagtattcta cagctcaaga ctgcagtgca gatgacgaca acttccttgt gcccatagcg	1140
gtgggagctg ccttggcagg agtacttatt ctagtgttgc tggcttattt tatttgtctc	1200
aagcaccatc atgctggata tgagcaattt tag	1233

<210> 30
 <211> 1236
 <212> DNA
 <213> Homo sapiens

<400> 30	
atggtgtgct tccgcctctt cccggttccg ggctcagggc tcgttctggt ctgcctagtc	60
ctgggagctg tgcggtctta tgcattggaa ctttaattga cagattcaga aaatgccact	120
tgcctttatg caaaatggca gatgaatttc acagtacgct atgaaactac aaataaaaact	180
tataaaaactg taaccatttc agaccatggc actgtgacat ataatggaag catttgtggg	240
gatgatcaga atggtcccaa aatagcagtg cagttcggac ctggcttttc ctggattgcg	300
aattttacca aggcagcatc tacttattca attgacagcg tctcattttc ctacaacact	360
ggtgataaca caacatttcc tgatgctgaa gataaaggaa ttcttactgt tgatgaactt	420
ttggccatca gaattccatt gaatgacctt tttagatgca atagtttatc aacttttgaa	480

aagaatgatg ttgtccaaca ctactgggat gttcttgtag aagcttttgt ccaaaatggc	540
acagtgagca caaatgagtt cctgtgtgat aaagacaaaa cttcaacagt ggcacccacc	600
atacacacca ctgtgccatc tcctactaca acacctactc caaaggaaaa accagaagct	660
ggaacctatt cagttaataa tggcaatgat acttgtctgc tggctaccat ggggctgcag	720
ctgaacatca ctcaggataa ggttgcttca gttattaaca tcaaccccaa tacaactcac	780
tccacaggca gctgccgttc tcacactgct ctacttagac tcaatagcag caccattaag	840
tatctagact ttgtctttgc tgtgaaaaat gaaaaccgat tttatctgaa ggaagtgaac	900
atcagcatgt atttggttaa tggctccgtt ttcagcattg caaataacaa tctcagctac	960
tgggatgccc ccctgggaag ttcttatatg tgcaacaaaag agcagactgt ttcagtgtct	1020
ggagcatttc agataaatac ctttgatcta agggttcagc ctttcaatgt gacacaagga	1080
aagtattcta cagctgaaga atgttctgct gactctgacc tcaactttct tattcctggt	1140
gcagtgggtg tggccttggg cttccttata attgttgtct ttatctctta tatgattgga	1200
agaaggaaaa gtcgtactgg ttatcagtct gtgtaa	1236

<210> 31
 <211> 1000
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> variant of LAMP-2

<400> 31	
atgggtctgct tcagactggt ccctgtccct ggatctggtc tggtgcttgt gtgcttggtg	60
ctgggtgctg tgagatccta tgcccttgag ctgaacctga ctgactcaga aaatgccact	120
tgcctgtatg ccaagtggca gatgaacttc actgtgagat atgagactac caacaagacc	180
tacaagactg tgaccatctc agaccatggc actgtcacct acaatggatc aatctgtggt	240
gatgatcaga atggcccaaa gatagcagtg cagtttgggc ccggtttttc ctggattgct	300
aacttcacca aggcagcctc cacctacagc attgactcag tcagcttcag ctacaacact	360
ggggataaca ccaccttccc tgacgcagag gacaaggga tccttactgt ggacgaactc	420

ctggcaatca gaatccccct taacgacctg ttcagatgca actccctttc aacccttgaa	480
aagaatgatg tgggtgcaaca ctattgggac gtcctgggtgc aagcctttgt gcagaatggg	540
acagtgagta ccaacgagtt cctctgtgac aaggacaaga ccagcactgt ggccccact	600
atccacacca ctgtgcccag ccctaccact acccccaccc ctaaagagaa gccagaagct	660
ggaacctact cagtcaacaa tggaaatgac acatgcctcc ttgccaccat gggactgcag	720
ctgaacatca ctcaggacaa ggtggcctca gtgattaaca tcaaccctaa caccactcat	780
agcactggga gctgcagatc acatacagct ctgctgaggc tcaactcctc caccatcaag	840
tacctggact ttgtgtttgc tgtgaagaat gagaacaggt tctacctcaa ggaagtgaac	900
atttccatgt acctgggtcaa tggttcagtg ttctctattg ccaacaacaa tctgagctac	960
tgggatgcac ccctgggatc ctctacatg tgcaacaagg	1000

<210> 32
 <211> 1000
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> variant of LAMP-2

<400> 32	
atgggtgtgct ttagactgtt tcctgtgcct gggtcagggc tggtcctgggt ctgtctgggtg	60
ctgggggctg tcagaagcta tgccttggag ctgaacctca ctgatagtga aaatgccact	120
tgtctgtatg ctaagtggca gatgaacttc actgtgagat atgaaaccac caacaagact	180
tacaaaacag tgaccatctc agatcatgga actgtgacct acaacggcag catttgtgga	240
gacgaccaga acggacaaaa aatcgctgtc caatttgggc ctggattctc ctggattgcc	300
aatttcacta aagctgcctc cacatattca attgactcag tgtccttctc ctacaacact	360
ggggacaaca ctactttccc tgatgctgaa gataaggga tcttgacagt ggatgagctg	420
ctggctatca ggatcccttt gaatgacctg tttaggtgta attcactgag cactctggag	480
aagaacgacg tgggtgcagca ctactgggac gtgctgggtgc aggcctttgt gcagaacggc	540
actgtgtcca ccaacgaatt cctgtgtgat aaggacaaaa cttccactgt ggcacctaca	600

attcacacta ctgtgccttc acctaccacc actccaactc caaaggaaaa gcctgaagca	660
ggaacctact ctgtgaacaa tggcaatgat acctgtctgt tggccaccat gggcctccaa	720
ctgaacatta ctcaggacaa ggtggcctca gtgattaaca ttaaccccaa cactaccac	780
tccactggca gctgtagatc acacacagcc ttgctcagac tgaatagcag caccatcaag	840
tatttggatt ttgtgtttgc agtgaagaat gaaaacaggt tctacctgaa ggaagtcaac	900
atctcaatgt acctggtgaa cggctcagtg ttcagcattg ccaacaacaa cctctcctat	960
tgggacgctc cactggggag cagctacatg tgtaacaagg	1000

<210> 33
 <211> 1000
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> variant of LAMP-2

<400> 33	
atggctctgtt ttaggctgtt ccctgtccct gggttcaggac tggctcttagt gtgtctgggtg	60
cttggagctg tcagaagcta tgccctggag ctgaacctga ctgactcaga aaatgccact	120
tgctgtatg ccaagtggca gatgaacttc actgtcagat atgaaaccac caacaagacc	180
tataagactg tgaccatctc agaccatggc actgtgactt acaatgggtc aatttgtgga	240
gatgaccaga atggccctaa gatagctgtc cagtttgggtc caggattcag ctggattgcc	300
aacttcacca aggagccag cacctacagc attgactctg tgtccttctc ctacaacaca	360
ggagacaaca ccactttccc tgatgcagag gacaaaggta tcctgactgt ggatgagttg	420
ctggcaatca ggatcccact gaacgatctg ttcagggtgca actcactgtc cactctggaa	480
aagaatgatg tgggtgcagca ctattgggat gtgctagtcc aggcctttgt ccagaatggg	540
actgtgtcaa ctaatgagtt cctgtgtgac aaggacaaga caagcactgt agccccact	600
atccatacca cagtacctag cccaccact actccaaccc ccaaggagaa gcctgaggct	660
ggcacctact cagtgaacaa tgggaatgac acctgtttgc tggccactat gggactccaa	720
ctgaacatca cccaggacaa agtggcctct gtgatcaata tcaatcccaa caccaccac	780

```

agcactgggt cctgcagaag ccacactgcc ctctgaggc tcaactcatc aactatcaag      840
tacttggatt ttgtgtttgc agtgaagaat gagaacagat tctacctcaa agagggtcaac      900
atttcaatgt acctggtgaa tgggagtgtg ttctccattg ctaacaacaa cctgagctac      960
tgggatgccc ctctgggctc ctcatatcatg tgcaacaagg      1000

```

```

<210> 34
<211> 410
<212> PRT
<213> Homo sapiens

```

```

<400> 34

```

```

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

```

```

Val Cys Leu Val Leu Gly Ala Val Arg Ser Tyr Ala Leu Glu Leu Asn
20           25           30

```

```

Leu Thr Asp Ser Glu Asn Ala Thr Cys Leu Tyr Ala Lys Trp Gln Met
35           40           45

```

```

Asn Phe Thr Val Arg Tyr Glu Thr Thr Asn Lys Thr Tyr Lys Thr Val
50           55           60

```

```

Thr Ile Ser Asp His Gly Thr Val Thr Tyr Asn Gly Ser Ile Cys Gly
65           70           75           80

```

```

Asp Asp Gln Asn Gly Pro Lys Ile Ala Val Gln Phe Gly Pro Gly Phe
85           90           95

```

```

Ser Trp Ile Ala Asn Phe Thr Lys Ala Ala Ser Thr Tyr Ser Ile Asp
100          105          110

```

```

Ser Val Ser Phe Ser Tyr Asn Thr Gly Asp Asn Thr Thr Phe Pro Asp
115          120          125

```

```

Ala Glu Asp Lys Gly Ile Leu Thr Val Asp Glu Leu Leu Ala Ile Arg

```

130

135

140

Ile Pro Leu Asn Asp Leu Phe Arg Cys Asn Ser Leu Ser Thr Leu Glu
 145 150 155 160

Lys Asn Asp Val Val Gln His Tyr Trp Asp Val Leu Val Gln Ala Phe
 165 170 175

Val Gln Asn Gly Thr Val Ser Thr Asn Glu Phe Leu Cys Asp Lys Asp
 180 185 190

Lys Thr Ser Thr Val Ala Pro Thr Ile His Thr Thr Val Pro Ser Pro
 195 200 205

Thr Thr Thr Pro Thr Pro Lys Glu Lys Pro Glu Ala Gly Thr Tyr Ser
 210 215 220

Val Asn Asn Gly Asn Asp Thr Cys Leu Leu Ala Thr Met Gly Leu Gln
 225 230 235 240

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

Asn Thr Thr His Ser Thr Gly Ser Cys Arg Ser His Thr Ala Leu Leu
 260 265 270

Arg Leu Asn Ser Ser Thr Ile Lys Tyr Leu Asp Phe Val Phe Ala Val
 275 280 285

Lys Asn Glu Asn Arg Phe Tyr Leu Lys Glu Val Asn Ile Ser Met Tyr
 290 295 300

Leu Val Asn Gly Ser Val Phe Ser Ile Ala Asn Asn Asn Leu Ser Tyr
 305 310 315 320

Trp Asp Ala Pro Leu Gly Ser Ser Tyr Met Cys Asn Lys Glu Gln Thr
 325 330 335

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

Gln Pro Phe Asn Val Thr Gln Gly Lys Tyr Ser Thr Ala Gln Asp Cys
355 360 365

Ser Ala Asp Asp Asp Asn Phe Leu Val Pro Ile Ala Val Gly Ala Ala
370 375 380

Leu Ala Gly Val Leu Ile Leu Val Leu Leu Ala Tyr Phe Ile Gly Leu
385 390 395 400

Lys His His His Ala Gly Tyr Glu Gln Phe
405 410

<210> 35
<211> 411
<212> PRT
<213> Homo sapiens

<400> 35

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

Val Cys Leu Val Leu Gly Ala Val Arg Ser Tyr Ala Leu Glu Leu Asn
20 25 30

Leu Thr Asp Ser Glu Asn Ala Thr Cys Leu Tyr Ala Lys Trp Gln Met
35 40 45

Asn Phe Thr Val Arg Tyr Glu Thr Thr Asn Lys Thr Tyr Lys Thr Val
50 55 60

Thr Ile Ser Asp His Gly Thr Val Thr Tyr Asn Gly Ser Ile Cys Gly
65 70 75 80

Asp Asp Gln Asn Gly Pro Lys Ile Ala Val Gln Phe Gly Pro Gly Phe
85 90 95

Ser Trp Ile Ala Asn Phe Thr Lys Ala Ala Ser Thr Tyr Ser Ile Asp
100 105 110

Ser Val Ser Phe Ser Tyr Asn Thr Gly Asp Asn Thr Thr Phe Pro Asp
115 120 125

Ala Glu Asp Lys Gly Ile Leu Thr Val Asp Glu Leu Leu Ala Ile Arg
130 135 140

Ile Pro Leu Asn Asp Leu Phe Arg Cys Asn Ser Leu Ser Thr Leu Glu
145 150 155 160

Lys Asn Asp Val Val Gln His Tyr Trp Asp Val Leu Val Gln Ala Phe
165 170 175

Val Gln Asn Gly Thr Val Ser Thr Asn Glu Phe Leu Cys Asp Lys Asp
180 185 190

Lys Thr Ser Thr Val Ala Pro Thr Ile His Thr Thr Val Pro Ser Pro
195 200 205

Thr Thr Thr Pro Thr Pro Lys Glu Lys Pro Glu Ala Gly Thr Tyr Ser
210 215 220

Val Asn Asn Gly Asn Asp Thr Cys Leu Leu Ala Thr Met Gly Leu Gln
225 230 235 240

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

Asn Thr Thr His Ser Thr Gly Ser Cys Arg Ser His Thr Ala Leu Leu
260 265 270

Arg Leu Asn Ser Ser Thr Ile Lys Tyr Leu Asp Phe Val Phe Ala Val
275 280 285

Lys Asn Glu Asn Arg Phe Tyr Leu Lys Glu Val Asn Ile Ser Met Tyr
290 295 300

Leu Val Asn Gly Ser Val Phe Ser Ile Ala Asn Asn Asn Leu Ser Tyr
305 310 315 320

Trp Asp Ala Pro Leu Gly Ser Ser Tyr Met Cys Asn Lys Glu Gln Thr
325 330 335

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

Gln Pro Phe Asn Val Thr Gln Gly Lys Tyr Ser Thr Ala Glu Glu Cys
355 360 365

Ser Ala Asp Ser Asp Leu Asn Phe Leu Ile Pro Val Ala Val Gly Val
370 375 380

Ala Leu Gly Phe Leu Ile Ile Val Val Phe Ile Ser Tyr Met Ile Gly
385 390 395 400

Arg Arg Lys Ser Arg Thr Gly Tyr Gln Ser Val
405 410

<210> 36
<211> 21
<212> DNA
<213> Woodchuck hepatitis virus

<400> 36
atcatgctat tgcttccccgt a 21

<210> 37
<211> 20
<212> DNA
<213> Woodchuck hepatitis virus

<400> 37
gggccacaac tcctcataaa 20

<210> 38
<211> 29
<212> DNA
<213> Woodchuck hepatitis virus

<400> 38
cctccttgta taaatcctgg ttgctgtct