ADENO-ASSOCIATED VIRAL VECTORS FOR THE TREATMENT AND PREVENTION OF DIABETES

Inventors: Terence R. Flotte, Gainesville, FL (US); Sihong Song, Gainesville, FL (US); Barry J. Byrne, Gainesville, FL (US); Michael Morgan, Gainesville, FL (US)

Correspondence Address:
Mark D. Moore, Ph.D.
WILLIAMS, MORGAN & AMERSON, P.C.
Suite 1100
10333 Richmond
Houston, TX 77042 (US)

Assignee: University of Florida Research Foundation

Appl. No.: 10/340,112
Filed: Jan. 10, 2003

Related U.S. Application Data
Continuation of application No. 10/267,117, filed on Oct. 8, 2002, which is a continuation of application No. 09/299,141, filed on Apr. 23, 1999, now Pat. No. 6,461,606.

Provisional application No. 60/083,025, filed on Apr. 24, 1998.

Publication Classification
Int. Cl. 7 A61K 48/00; C12N 15/861
U.S. Cl. 424/93.2; 435/456

ABSTRACT

The subject invention concerns materials and methods for gene therapy. One aspect of the invention pertains to vectors which can be used to effect genetic therapy in animals or humans having genetic disorders where expression of high levels of a protein of interest are required to treat or correct the disorder. The subject invention also pertains to methods for treating animals or humans in need of gene therapy to treat or correct a genetic disorder. The materials and methods of the invention can be used to provide therapeutically effective levels of a protein that is non-functional, or that is absent or deficient in the animal or human to be treated. In one embodiment, the materials and methods can be used to treat alpha-1-antitrypsin deficiency.
FIG. 1

FIG. 2
FIG. 3
FIG. 5C
FIG. 6
<table>
<thead>
<tr>
<th>Genomic DNA</th>
<th>Episomal DNA</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>-</td>
</tr>
</tbody>
</table>

FIG. 8
![Graph showing hAAT (µg/mL) over weeks post-injection for C-AT and p43CB-AT.](image)

**FIG. 9**
FIG. 10

hAAT (ng/ml)

p43C-AT  p43rmsENC-AT

0  5000  10000  15000  20000  25000

FIG. 12

hAAT (ng/ml)

SuperFect  FuGENE  Lipofectin  LipofectAMINE  Ca-PO4  Control
FIG. 14
FIG. 15A
C-AT (Ligation of pTR and aat) (cont.)

gccgcaggccc gccgcgtttct tttttgtcaag aagcaacagt gccgtgcccc gaatgacgtg 2700
caggacaggg cgcgcggtgct atcgtcttggt gccacgaagg ggtcttcttg gcgcgtgctg 2760
cgctgttttg gcagccgtttc ttttttgatc gccgctttgc ggacagaagg ggacgcagtg 2820
gatctctgtgg ctcctgctcgc gagaaccttg ccaatgcagc tgcgcgctcg ggcggccag 2880
cggtggtggt atacgcttcaa tccgcgtacg tgcgcattcg gccacgcaagcc ccaatgcagc 2940
cagctctccgc cagcgctgtgg ggcggtgtgct gcacgctgct gagctgtgtgc cgggtggtggt 3000
gaggtcgactt cggctctgctgg ttcgcgtttag ctcggtcgct gcacgctgct gcacgctgct 3060
ggcgcgtctgc gcgtggtgtgct gcacgctgct gcacgctgct gcacgctgct gcacgctgct 3120
ggcgcgcttt ttcctgatcg tcgcgtcttg gcacgctgct gcacgctgct gcacgctgct 3180
ccgtgcgggg ctgcgtcctgg ctgcgtctgg gcacgctgct gcacgctgct gcacgctgct 3240
tcgtctctgg ggcggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 3300
gacggtgtctt gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 3360
tctgtgctgc gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 3420
gccacttcgca ttcctgatcg ttcctgatcg ttcctgatcg ttcctgatcg ttcctgatcg 3480
tgcgttcctga caggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 3540
aatgcgtcctg gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 3600
tcgtctggtc gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 3660
tggctggctgc gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 3720
tgcgttcctg ctgcgtctgg ctgcgtctgg ctgcgtctgg ctgcgtctgg ctgcgtctgg 3780
gttctggtttc gcggctgtttc gcggctgtttc gcggctgtttc gcggctgtttc gcggctgtttc 3840
gggcgaggg ggcgttggttgc gcggctgtttc gcggctgtttc gcggctgtttc gcggctgtttc 3900
agcgctgctgc gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 3960
tgcgttggtc gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 4020
ccgtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct gcacgctgct 4080
ctgccggcgc gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 4140
ctcgtggtgg gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 4200
agcggctgg gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 4260
ctgccggcgc gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 4320
cggctgctgg gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 4380	
tggctggtgtg gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 4440
tgcgtgttgc gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 4500
tggctggtgtg gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 4560
tgcgtgttgc gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 4620
ctgctggtgtg gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 4680
ctgccggcgc gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 4740
tgcgtgttgc gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 4800
tgcgtgttgc gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 4860
ctgccggcgc gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 4920
ctgccggcgc gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 4980
ctgccggcgc gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 5040
ctgccggcgc gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 5100
ctgccggcgc gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 5160
ctgccggcgc gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 5220
ctgccggcgc gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 5280
ctgccggcgc gcgtggtggtggt atacgcttgg gcacgctgct gcacgctgct gcacgctgct 5340

FIG. 15B
C-AT (Ligation of pTR and aat) (cont.)

cggcgtcaat acgggataat acgcgcac atagcagaac ttttaaatgg ctcattcatg 5400
gaanaagcag tcctgggcca aaaaacctaa ggatctttac gctgtgtaga tccagtctga 5460
tgtaaccaccc caattgcacc caggcattccca tactcttaac acgggtctgt 5520
gttgacaaa acaagaggaag caaaaacgct caaaaaggg aaataaggcc aacaaggaat 5580
gaeaatctct ctaactcttt cttttttcat tattatgaag cattctcag ggttattgtc 5640
tcatacagg cattctattt gattgtattt aagaaaaataa ac aaataggg gtctccagca 5700
cattcccccg aaagtgccca ctgactgctt aagaaaccct attattacag acatttaccc 5760
ataaaatatg gctttactacg actgttccttc ttctctgctg gaacggggaa 5820
acgctgaca caatcgagctt ccggagacgt tcacacgttt tgtgtacaacg ggaggtcgg 5880
gcagacaagc ccgctcaggc ggccaagcctt tggtgagcgc gttcgggcc tggctaaact 5940
atgctgcctggacctgct getgctgtgct gaccatatgt ggggtagcga atacaacagca 6000
gatgtcatg gagaaataac cgcattcagg aattttgaaa actttatttt tggctttattt 6060
cyggttaaact tttttattttt ttcagcctatt tttaacccaa tgggagaaaat aggctggcnaat 6120
ccctttataaa tccaaagaaat agacacgagat aggggtgggt gttgtdctcc agttaaaca 6180
gactggacta taagaacagc tggacttaca cgcctcaaggg cggaaacgcgt tctatcagg 6240
cgtggtcctc atcctggtac gtcaccctta atcaagtttt gttggtcgcga gttctccgtaa 6300
agcataataa cgcacactta aaggggcccc cggatttatga gctttacgcgg gagacccgac 6360
gacagtggcc agaagggggc ggaaggaagc ganaaggagc gggcgctggg cggcttgcaag 6420
tgtactgctgc acgctgctcgc taaccacac ccgccgcccc ttaatatgcgc egctacaggg 6480
cgcgtgccgc cattctgact tcaggctgag caactgtttgg gaaggagcac eggtgccgca 6540
ctcctctgcttatagcgggcg cggcactgca 6565

FIG. 15C
FIG. 16A
acatatttacca aaactgttcca ttacctgaaac ctatgatctcg aagagctgct tgggctcaact 2700
gggcatatcc aaggttttcca gcaatgggggc tgcattctctgc cgggtcaccag aggaggccacc 2760
cctgaagcgtc tcgcaagggcg tgcataaggg tgtgtcgtctcg atcggacgaga aagggacgta 2820
agctgctgagg cacctaggct tatagcctcgtctc acatacccccag agttcatcgttct 2880
caacacatccc ttgcttccttctaatgatgta acaaaataac ccgattccccc ctcctcctccta tttcttctgg 2940
aaaaggttggtt aaccttttacactcctgctgc cctcataaccc ccttttctcc 3000
ccctgctgccc ccctctggtgcc acatttaagag aaggtgtgagctgtgtaaccc ccccccwvcccc 3060
tcgactgctgcc tcggctgatagac gtcgcaagcgag aatggccttg cccttctcttttgtcctc 3120
cctggaaggtg cttcaccacac cttcgaagcgg ctgtgtgggtgtgaagtcctgccttc 3180
ctctacacca ctggctctttactcctgctc gtcgtaaccc gttctgctgcct ctatggtcatg 3240
cattggccaag ttgcaaaccc cagatggctg cgggggctgcc cgggtctccag tcacactctcgc 3300
ataattttgcc gacggcgtgtgt gcctggaacca cgtccgacacc cctgcaagcctcattggtgatt 3360
gccctatggat gacagcttggc gacgtgtggtcttcacccagac cggcttctggtgcttgcctct 3420
ggctgtaggct ggcgtgaggtc tttttttctttt acgcgcttggtt cttgtttggtttcttgcttc 3480
ggcgcgaggcc gcgcggttcttg ttttttttttttgcggtgctgcg tttttttttttgtggttt 3540
cagggacgggg cagcgcggtctg atcggtgcttg gcctcagggg gctctgccttt cgcagctcttg 3600
ctcgcaggtgg tacactggaac gggagggccag ctggtgcctgt ctgcgggagtgc ggccggtctg 3660
gatctctcgtg catcactaccc tgcgccctctc gagaattactagcattgcgt cgtatgccata tggcctgaat 3720
cgccgctgcgc atacgcctgga tcgcggtactgc tgcgcttcctc gatacccaacac gacccggagga 3780
atcgacgagc caagtcctgct gcgggtgtgcct cggcggcttcct ctcattctgcag ctcggtgatgc ggagccaggg 3840
gcccacgcgg gcgcgctgccc ggcaagagcct gcgtggggtttt gcgggcaggg gcagacgcggg 3900
ggcgaggggatt cgcggtgctcg gcctggagcc catgttgacgc cgcctgctctc gctgcggggg 3960
ggcggtgtttt cggggtatggc gcgcggtgtcc cggcgggtttg tggggggcgc gccggcggtgctgctc 4020
atagcgttgggc ttcacccctgg tattggtggaa gctgttgctgg gcgatagggcg gcacggcggac gatggcgcttg 4080
cgcgtgttttt atcgctggcgt gcgtttgcctgc tcggctgctgc gctggtggagtc ggcgggtggg cgtgggcttc 4140
gacagcttgc gcgggtctggc ttcagcgcctctg gcgcggtgttt gctgtgctggt gctgcgggagt gcctgcggtgct 4200
tctgtttgcc tcggggttctg tcggtgtggct ctttgtttgc gctggtgtttt gcgcgggagtc ggcggggttgct 4260
gccacatccac ccggtttccttc ctgcattgta ggtggggggg cggggtgtttt gcgcggtgtttt gcgcggtgtttt 4320
tgtgctctcgtg catctggtggt cagagggggg gcggggtttg gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt 4380
aatacaggggc atcggtgaaaag gctgggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt 4440
cctccgtggac cgcggccggt gcgtgggggt cggggttttgc gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt 4500
tgggtggcgcc gcgggttttgcc agatgggttg gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt 4560
tggggtgggaa ctcctctcgcct gttttctcgtgc gcgtgtgtgtgt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt 4620
gtgcgtttttt gatcggtagtc gcgggtttttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt 4680
agttggcaaca gttggctggc cggggtttttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt 4740
aagccgaacaagaacatgta gcggaaggggc gtttttgcgcag gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt 4800
agtctgggctc gcgggggggc ggcgggtttttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt 4860
ccggggtttttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt 4920
acggttggga ggcgggtttttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt 4980
aggggccttg cggggtttttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt 5040
agctggggtc acggggtttttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt 5100
tgtggccgtcg tcggtgtttttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt 5160
tggggtttttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt 5220
aaggggtttttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt 5280
tggggtttttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt 5340
tgtgggtgttg gttgtgggttg ccggtttttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt 5400
aaaaaggttggg gtcggtttttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt gcgcggtgtttt 5460

FIG. 16B
E-AT (Ligation of AAT and elf) (cont.)

aagggatattt ggctcatgaga tttatcaaaaa ggatcttccac ctatgcatttt ttaaatattaa 5520
aatgaagtttt taataactaact taagactatat atgagtaaaaa tgggtgtgac ctgttacatct 5580
gcttaactcc ttaacacccact tcgccatccact cctgtgttac gtttcgttctc 5640
gacctgctgac cttggtatatgacctg gggaggtcag accatcagcc accatcagcc 5700
cattatgaacct cagccaggcc aggccagcact cccctgctag cccctgctag 5760
tctggatgcc ccttattttt cttggtttgag ggttactcgg gtcttggctc 5820
ctcaatctcc ccaattttcct cctctctctt cctctctctt cctctctctt 5880
gtggagcaac tctctctctt cattacattt cattacattt cattacattt 5940
tggtaaatcct ctatactctct ctatactctct ctatactctct ctatactctct 6000
ccctgctgcc ttgctgctgag ttgctgctgag ttgctgctgag ttgctgctgag 6060
tggatattcct cttatatatttt tttatataaatct aatttaggtt ggttacttc 6120
ttttaataattc accaaatccttt cactgataat atggtgactt gcttcaagtgg 6180
tctgtcctctt acctcttttct aggccaccctt ttgctttcttt taaaaatggc 6240
gaaaaaggtc cactttctttt ctatcacaat ttctacgatttt tctactttaaa 6300
tctgctccact gtaatttcctt tattttaaat ttatatcactt attatatatttt 6360
gctgtagcaag gttgatacag ctttttctttt cctctctctt gctgtcctctt 6420
ccgggatctcc cggagcgggac cggagcgggac cggagcgggac cggagcgggac 6480
agcttcaaaa ttttatttttt ttttatttttt ttttatttttt ttttatttttt 6540
cccttaataa ctctactctct tctacaactct cattttttttt cattttttttt 6600
gcttttttttt cgttttttttt cgttttttttt cgttttttttt cgttttttttt 6660
tctgttccaac ccgggatctcc cggagcgggac cggagcgggac cggagcgggac 6720
gcttcaaaa ttttatttttt ttttatttttt ttttatttttt ttttatttttt 6780
gcttttttttt cgttttttttt cgttttttttt cgttttttttt cgttttttttt 6840
tacttttttttt cgttttttttt cgttttttttt cgttttttttt cgttttttttt 6900
cgggagcttt ttttttttttt ttttttttttt ttttttttttt ttttttttttt 6960
gcttttttttt cgttttttttt cgttttttttt cgttttttttt cgttttttttt 7020
gcttttttttt cgttttttttt cgttttttttt cgttttttttt cgttttttttt 7080
gcttttttttt cgttttttttt cgttttttttt cgttttttttt cgttttttttt 7140
gcttttttttt cgttttttttt cgttttttttt cgttttttttt cgttttttttt 7200
gcttttttttt cgttttttttt cgttttttttt cgttttttttt cgttttttttt 7260
gcttttttttt cgttttttttt cgttttttttt cgttttttttt cgttttttttt 7320
gcttttttttt cgttttttttt cgttttttttt cgttttttttt cgttttttttt 7380
gcttttttttt cgttttttttt cgttttttttt cgttttttttt cgttttttttt 7440

FIG. 16C
dE-A (Fragment 2 Circularized)

```
ggggacatgg gcccaacacc cccccgggcc ctggctgcgt gcctggagcc 60
ggccccgagc caccccaagg ccgccccggcc actctgccga cctgccccag 120
ggccgagag caccccaagg ccgccccggcc actctgccga cctgccccag 180
gtggagacct gaaagaagaa cctgctgcgt gcctggagcc 240
ccgcccccgc cccccccgcc cccccccgcc cccccccgcc cccccccgcc 300
ggctggagcc cccccccgcc cccccccgcc cccccccgcc cccccccgcc 360
agagtcgcct ggctgcgggc ccgcccccgc cccccccgcc cccccccgcc 420
gatgcgtct ggctgcgggc ccgcccccgc cccccccgcc cccccccgcc 480
gccgagctgg gggagaggcc agcggaggttc agaagaggttc gcgggaggttc 540
aactgagcga ggtgcgtcgc cttggtccct cttggtccct cttggtccct 600
tatatagttg cagtgatgcg cgtgcaagct ctatccctcc gcgctggaacct 660
caggtgtagtg ccggtgagtc tggtccgggg gcggagggcc gcggagggcc 720
agtggctgct gcggag ggctgcgtgt gcgtggaggttc gcgtggaggttc 780
gctgccgcc gcggggcgc gcgtggaggttc gcgtggaggttc gcgtggaggttc 840
gcgagccgtg ccggtgagtc tggtccgggg gcggagggcc gcggagggcc 900
tgcgtgctgg gcgtggaggttc gcgtggaggttc gcgtggaggttc gcgtggaggttc 960
caacccgagg gactgccacc cttggtgccct cttggtgccct cttggtgccct 1020
cgggagcccg ccggtgagtc tggtccgggg gcggagggcc gcggagggcc 1080
tgcgtgctgg gcgtggaggttc gcgtggaggttc gcgtggaggttc gcgtggaggttc 1140

tgcgagccac cttggtgccct cttggtgccct cttggtgccct cttggtgccct 1200
```
dE-A (Fragment 2 Circularized) (cont.)

aaccagccagc ccggaagggcc gcagccagcag aaagcgtctgct gcccttacgcc aactttatcc gccttcatcc 5460
tagctattaa ttggttgcgg gaagctagag taagtgattc ggcagtaatt agtttgcgcag 5520
accttgctgcc cattgctacga ggatctgagg gcctgacgctc gtcggttctgc gtcggttcat 5580
tcaactccag gtaggtctcc ctcaagggct gtaatcagtt ccatcagtcgcc cccatgttgg tggccaaagg 5640
cggttagtct cttccccccagc ccagattttc cctcagaaacct ccagatctcct ctctgccagcag 5700
tcatttgtta ggcagacgag gctacactac ttcagctctct gctacacagcc atctactcctct 5760
tgtgaacttg gtagtataac accaagtcag tctgtaatac gtcggtattac gacgagctt 5820
gctctagcgg cagccctaat ccggagtaata ccgctggcaca ctagcgaagaa ttaaaatgca 5880
tcataattgg aaaaagttct tcggggcgaa aactctcaag gctcttcacc cttcttgagat 5940
tcagttctag caagcccatct cggcagcaccc actgatctcc acttctctta cagttccacca 6000
gcctttctgg gtaggcacaa ctaaagggcg ctttaatgagc gccaaaggct ctaattgg 6060
cacggaaatag ttaattgcta taatttatttt cctttctaat ggatttgagg aatatttagc 6120
gttaggctct cttcagtgag lataacttta tatattatttt gaaaaatatt ccaaataggg 6180
tccgcggcct aatcctccag ctaagccctct cttctggagct aagagacattt aatatgca 6240
cattacacta taatggcttg gtttttaatt ttttttttaattt ttttttaattt cctttttttt 6300
agctgtaaaa cctcctgacac atgcagctcc cggagacggt cagctgatttt ctgtaagcgc 6360
atgcccggcc cagctgacac ccagctggggt gtcgctgcgg gttctgttgg gttgtggggcg 6420
gctctgttct gtcgagcagc gctgatattc ttaaatcatt ttttaatcatt accatatg 6480
tacccaaagc atctctatag gcacactgaa aatctgttaa atatggaaat 6540
gttaaatatt gctgtaatat ttttttaatt cctttttttt aaaaaatatt gtttaaatatt 6600
cggagaaac gtaaatatat gataaatgaaaat gtaaataggaatt gtttctaggt cccagtcag 6660
tgtagttagc gtagtctacc taaaagatag ggcacgccag ctgaaccagc gccaaaccctg 6720
ccccacctag gtagtctacc ccccacccag ccccagcagc ccccagcagc ccccagcagc 6780
gtctctactaa ccgagatctaat gatgaagcccg cggtttgagtt tgaaggtggg gtagtctacc 6840
aaagaagcgg ctagtctgcag gtagttgctg cctgctgctg cccagcagc ccccagcagc 6900
gctgctgctg cctgctgctg cccagcagc ccccagcagc ccccagcagc ccccagcagc 6960
gtagtctagg cctgctgctg cccagcagc ccccagcagc ccccagcagc ccccagcagc 7020
gggggttggg cctgctgctg cccagcagc ccccagcagc ccccagcagc ccccagcagc 7080
FIG. 17C
FIG. 18
FIG. 18A
p43C-AT (Ligation of TR and aat) (cont.)

gctggtaacc ccccccceccc ctgccaggggc cccctgacccg ggcgcccgggt tgcagcaaac 2700
atgataagat acacttgatga gttggcaca aaccacaacat aatcagctgt aaaaaatgc 2760
tttatattaca acacaaatttg cattcatttt atgtaagcag ttcacagaga gatgtggagg 2880
cttttttcata acgtaatcaga cttaaatctt cccctttctt gacacagcag tgcagcagcag 2940
ccctagcttag gacctgcagccaccttgctaggcctgaagac ggcgcccgggc 3000
gcacaagccc ggcgctgagg gcagctttgg tagccaggggc taaggccgcag aacagctgccc 3060
cagagagggg gttgtctggcc cccctcctgg cctgctgctgta atacgagagag 3120
gcgcgtcacc gatgcgcttt cccacacagct tgcagggcgag atgcggccgata ggcgcccgggc 3180
gccgtcgatcc gcggcgggct cccctgtcctt cctgcttctt cctgctccttc gcggcggcagc 3300
gccggtctttt ccccgctcaag cttcataatcg ggggtctccct ttgcggttcc caattgagc 3360
tttacgcacca ctcgctgacca aacaggtgtga attagggcttc gttcgacata ctggctcattc 3420
gccgctgatac acgcttttcc gccttggagc ctttgaggtcc acgtttctctt ataagttgact 3480
catttcgccc aacctgacca cattcacttg tttccttttattttttttt attatatag 3540
gattttgccc ctttccggctt attggattaa aacatgtactg atttacacaa atatttaacc 3600
gaatatatata cgtaatatg atctctagcctc ttggtgttttt ccatttgcatg 3660
ctggtcgccttt ttcacaccgc ctttggtgtc aacgctctgta actatcagtt taattgctc tgcagcagcag 3720
tagtgacgac gccctggcagca cccggcacaac cccgctgacag cgcccctgacg ggcctggtcg 3780
cctggtcgccttt gaccctgacg acctggctctgac gcggcggcagc gttcgacatg 3840
cttccacgcttt cagctgagaa cccgggactg gctgggtctttt tgcagttattt attgatatt 3900
taggttaatct cttggtcataa aacgctctgt taggctgaatc tggcgaactt cggggagaaat 3960
gttgcaaggg ccccttgtgga ttttcccctt ttaaatcatt caaatagttta cccgggtgctt 4020
agacaaactac cccgcaatata gcttcagataa tttggaatgaa gggaggtatg gatgtattcccc 4080
cattgtcgttc tggcccttttt tcccttcttc ccggcttttt ggcctggttt gttgtgctcc 4140
cgcacagccc ccagggcattt cccctgtgcc atgggtctttc aagagctttt gttgtgctcc 4200
atgcatgtag accatccagc cctgacgata ctttgaggttt cctgcggcag aacagcttttt 4260
ccataatggc gaccttttaa ctgtgtgccgc cttgctgtgc ggtgctgctt aacacgctctt 4320
ggcgcagagc cccgtgttccg cccctgatac ccttcctcctt cttcttccag atgctgtttt 4380
ccacctgacag cccgtgttccg cccctgatac ccttcctcctt cttcttccag atgctgtttt 4440
atacatgatcc tcgggggtgc tggcttttac ccgtggtcataa cccgtctgctt ggcctggttt 4500
agacctgcatc cttttttgtca ccaacagggg cggctttgtaa gacagctgtg tggcttttac 4560
ggcgcctgta aacagctctt cctggtgttt cccgctgctt gcgggttttt ccccgtgctt 4620
cagctacgct cctgcgctat gatagctgcg ggcctggtgcgt gggctgttccg gacagctgttt 4680
ttatagctgag cggctgttccg gggctgttccg ggcctggtgcgt gggctgttccg gacagctgttt 4740
ggcgctgtgcttt ctttgaggttt cggctggttggt cggctggttggt cggctggttggt cggctggttggt 4800
gacagctgagc cccgctggttt gacagctgttt cggctggttt cggctggttt cggctggttt 4860
cagcagatct ccaggtatg tttttgagtt cggctggttt cggctggttt cggctggttt 4920
catttgctctt cttggtgtggg cggctggttt cggctggttt cggctggttt cggctggttt 4980
ctttatatgta ccactgatc cggctgttttt cggctggttt cggctggttt cggctggttt 5040
taagcggatgc cttttttgata cggcttccg cccgctggttt cggctggttt cggctggttt 5100

FIG. 18B
p43C-AT (Ligation of TR and aat) (cont.)

tgagaatatc ttctctcctgct ctaaatctgct tcattgcacaa caaaaaaacc accgctacca 5160
gcggctgttgg tatttgccggag tcaagagcta ccaactcttt ttgccaaggtt aactggtgcc 5220
agcagacgcc agataccaaactactctccc tcttgattgctt acattttg gccaccctctc 5280
aagaacctctg tagcactcgcct caataccctctc gcttgcgcttaacctctgttt acgtggtgtc 5340
gcggagtggc gtaaagcttgg ttttaccggg ttggaacttcaag ctgctgatagttt accggataag 5400
gcggcagcggt cggtctgacat ggccccggcagg tgcacacagc ccagccccgct gccgaaagacc 5460
tacacccgcc tggataacct aacgggtgag caattgagaa gagccccgct tccggaaggg 5520
agaagaacggcg acaggtatctgc gcgttaacgccccggagggctggtaaaggccga caggagcggc 5580
ctccccaggg gcacgcccctctctccctt actctcttgtc ggccaagcgcct gccgacagcttc 5640
gcgccgtcagtt ttggtgagta cctgctagcgg gccccggagcc tatggaaaaa cgcgcagcaac 5700
gcgcctctct ttcgctctttt ggctcccttttg cctcactctt cttctctctccgct 5760
tttacccctctg attccgtgtag gacagcttccccgctgtctgcttaccgcttc aactggtgct 5820
gcggccgccag ccgccagcagc cagcgagctca gtcagcggaggg cagccagaaga gcgccccaaa 5880
cgcaaacgcc ctgccctcctg gctggtgctgc attccatttg gcagggctgctc ag 5932

FIG. 18C
p43C-AT-IN

7492 bps

Exon 2 (649)

Exon 3 (270)

Exon 4 (147)

Exon 5 (267)

polyA

itr

FIG. 19
FIG. 19A
p43C-AT-IN (Ligation of p43-C into IN) (cont.)

tgctgggcccc atgtagggagc ggcggcatac gccgtgctact cccccgcaag tcagaaggtc aa 2760
caaaacccctt gctctctttta tgatgacaaac aaatcacaag ctcctctcttc tcaatgggaa a 2820
agtgggtgaat ccccccccaaa aataactggt cctccgttcc ctacccctcc cctcctccca c 2880
	gcccctccct cctgttcaaa attaaggttc gggtggtgcat gtaagccgta cccccctcg cc 2940
agggccctgca gacgcggcgg atgacagcct ggtagacgcc tgcgaatgttt ctagaaggg c 3000
cgccgagtgtg cttcgaaact gctctctggtg cctccgtttcc gcaccgcggt gcggcttgt gg 3060
gacgcagcatg taagctgatag tgtagttttg ggacaaacc aactgtagaat gcgctgtaaa a 3120
aaatctgttt ctaagttgaaat tttgttgatg cttctggtat ttagttaaact tataaagctg c 3180
atatatacaag ttaacctaa cccagccgtt ctttctttag gggtgggtc aaggggatc g 3240
tgaggccctg ttaaactggcc gtaaaatcctg taaacagacc gtaaagccaa c 3300
gaaaccctta tggagacgctt ggcggcactct ttctctggtg gcgctgctctg agcagtggt g 3360
cgccggtgca aacgcgggcat gcggcccgtg ttctgctttcc gcggcggctg tggccgagc g 3420
aggcgcacag gggagacggg ccaactccag tgcgctattt aatccacgca acgcgcgggg g 3480
agagccgctt ggctctctgt cgccttcgg cgccttgtgt gcgcgctggt gcgcgctgg g 3540
gctggctggc tgcggcggcg ggtcttaggct cctcctgaggg gtttagctac cggcctagg c 3600
gaataggcag ttaacgaggag aagaagtagt cgacaaagag ggcggcgaac ggcggcggac g 3660
cctaaagagg cctgggttccc gggcgtgccc ctaggctccc gggcctccgtg ggcggcgtt c 3720
aaaaatcgac gtcactgaag cggtagggga aacgccagag cactataagg ataccagcgg c 3780
	tcctcttcgt gaaacgccct tgtagcttttg gcgccgcttgc aataccggtc taccgctagt c 3840

tctgcggtct ttcctccctcc gggagagcgt gcgtctctctt aatcttgctct cttgtaggt c 3900
ctcagtcttgc ttcggctctgc cgcctgcttg gcgcgtggag tcacgcggag tcagtcgagt g 3960
cgccgacct ggcggctgtt gcggaggttg cgcctcgcg aacgctggag ggtggtggag c 4020


FIG. 19B
p43C-AT-IN (Ligation of p43C into IN) (cont.)

```
Ggggttcggc gcacatttcgc ccgaaaaagtgc ccacgtgcggc tcctaaaggaac cattattactgc 5460
agcactattaa cctatataaa tagggctctac acggagccttt tcctgctgctgc ggttcccggg 5520
gatggcattg gcacactccttg acacactcgc accctcggag cgttcaccgc ttgctgttga 5580
gggagagcggc gcagagcggag ccggagcctcgg cgggttgctcag cgggttgctcag ggggttgctcag 5640
ggctgttcctgt acttaggtgcg agaatagcggc attgtgacgg gatggtgcgtg tataggtgcgtg 5700
ggaatacgcc acagatgctgt aagggagaaattccgatac gaggaattgta aacgttaata 5760
tttgggttaa atttgcagttata aattcagctc attttttataa caattcggcc 5820
aaatacgccg aaattccctaaa aatctaaagag aataagcggc gatagagttgg aagtggtgattgg 5880
cagtttggaaga cagagctccaa ctattaaagag aagttggactc caacgtcggc gggcggaaaga 5940
eeccccttcga cccgagcctgc caataagctgt acaccaagcct ctaaactcgag ttattagggtt 6000
cagagagctgc ctaaagcacttc aatgcggagc cttaaggggag ccccccagattt agacgtgtgac 6060
ggggagagcc gcagagcggag ccggagccttcg cgggttgctcag cgggttgctcag ggggttgctcag 6120
ggctgttcctgt acttaggtgcg agaatagcggc attgtgagctc gatggtgcgtg tataggtgcgtg 6180
cagctcttcgc cctcctctcg gcctgccgctgc cttcctgctgc cttcctgctgc cttcctgctgc 6240
gcgctgcttgc gctctggtacg gcctgccgctgc cttcctgtgcg cttcctgctgc cttcctgctgc 6300
cgtgctgcttgc gctctggtacg gcctgccgctgc cttcctgctgc cttcctgctgc cttcctgctgc 6360
gcgctgcttgc gctctggtacg gcctgccgctgc cttcctgctgc cttcctgctgc cttcctgctgc 6420
tattggtgtactt gcctccgttgct gcctccgttgct gcctccgttgct gcctccgttgct gcctccgttgct 6480
tattggtgtactt gcctccgttgct gcctccgttgct gcctccgttgct gcctccgttgct gcctccgttgct 6540
tcctgcttgcc atatgacgccc cgagtgtggcct tcgattttgtg actgatgtattt aatagttaactt 6600
aatctgctgg tcattttggtcc atagcggtttg cggctgtatac aacttcggcat 6660
aatctgccgg ccctggtctcc gcgcccccaagc ccccccggccaa tcggtcggctg tgcagctga 6720
tgtccccata gtaaccctccg cagggcacttt cacattgcagtc caattggtcgg cattattacgt 6780
gtaatctgccc eactctgctgc atacatcaggt gactacagctt cccattcattg cccattcattg 6840
tgcctactgc attactgcctgc tggctgcactgc tggctgcctgc tggctgcctgc tggctgcctgc 6900
ttcctactgc atgtatatgt tagcctgtatc ctatgtatgtg aaccaggtgta tgcgggtttg 6960
gcctacggc aaggtagctgc gatagtcggct gcattgcctgc ggttccccaa gatttcctta 7020
catggctgctgc aatgggaggtgc tgcgccttcgca ccctttctcc ggggaattgc cccattgctcg 7080
atatacggcgc ccccccggtgac ccgaaatcagc gggtggtgcgt gcgtgtcgacg aatgtagctt 7140
aaggtagctgc gttttaatttc aacgtctcgat caattgaacgc tttattctcg tagtttaata 7200
cagctttaaat gctaaaaacc ttggagggctg tgcacagcag tgaattagctg tgcacagaatc 7260
aattgggctgg tggctgcagag atcatttgctgc tggctgcctgc cagccgttgcct gcgctgcctgc 7320
ccatagacata ccctgcttgct gcggagtctcg agtagcttcag cattcagcgag caattgctgc cattcagcgag 7380
tgtgctttatc gctatcattg tctcctcttt cttcagactgc ggcacccctgcagt gacccataa 7440
cagctttaaag ggtttagagta cttatatcag ctaattacgata ctggtgctgg atgggtgtgg 7492
```
FIG. 20
p43CB-AT (Ligation of Fragment 2 into Fragment 2)

GGGGGGGGGG GGGGGGGGCTG GGGCGTCTCCT CTGCGCGCGC TGGCGCTGCTG aCGCGGCGCG 60
GCGGAGCGAA GGCGCGCCGG TGCGCGCGCG GGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 120
CAGGGGAGCC AAGCGCAGCG CAGCGCGCGC CGCAGCGCGC CGCAGCGCGC CGCAGCGCGC CGCAGCGCGC 180
TCGCGCTGATG TCGCGCTGATG TCGCGCTGATG TCGCGCTGATG TCGCGCTGATG TCGCGCTGATG TCGCGCTGATG 240
ATCGCGCGATG TCGCGCTGATG TCGCGCTGATG TCGCGCTGATG TCGCGCTGATG TCGCGCTGATG TCGCGCTGATG 300
GGCGACTCGG CAGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 360
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 420
GGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 480
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 540
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 600
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 660
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 720
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 780
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 840
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 900
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 960
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 1020
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 1080
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 1140
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 1200
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 1260
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 1320
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 1380
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 1440
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 1500
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 1560
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 1620
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 1680
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 1740
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 1800
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 1860
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 1920
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 1980
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 2040
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 2100
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 2160
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 2220
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 2280
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 2340
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 2400
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 2460
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 2520
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 2580
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 2640
GGGCGCGCGC CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG 2700

FIG. 20A
p43CB-AT (Ligation of Fragment 2 into Fragment 2) (cont.)

gacacagtttttgaggtcaagggacacggagggagaagctaagggagcactggctcgagtgatgtctctttgactgaaagttgctccgtggtgagttgcttgatttatttttcatgcttctctgtttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
p43CB-AT (Ligation of Fragment 2 into Fragment 2) (cont.)

```
aattaataga ctggatggag gcggataaaag ttgcaaggacc aotctctgcgc tccgccccctc c5520
cggctggtcg gtttattgtc gataatactg gcgcgggtgga gcgggctgtct cggcggtatca c5580
ttgcaacact ggggcagcgat ggttaagccct cccgatctcag aattttctcag acagacgggga c5640
gtcaggcaaac tatggtgaaa gcgaataagac agatcgcgtgaa gataggggcc ttacagcatta c5700
gacaattgta acctgcagac caagtttact cattatatact ttagttctag ttaaaaacctc c5760
atatttaatt ttgaaaagatt tgggtggaaa ctcccccttgta taattttctag ccagaaaaatctc c5820
cttaactgta gttttctggct caactgagct cagaccccccgt agaaaaagcat aagggatctt c5880
cttgagatcc cttttttctgg ctgcttgctg cccgttggcca aacaaaaaa cacccgctac c5940
cagcggttgtt ttggttggcgc gataaacagc taccatactc ttctccgagag gtaactgtgc c6000
tcaagcagag gcagataacc aatattgcct ttctgattgta gctgtaggtta ggcaccaact c6060
teaagaactc tggcgccacc cctcatatacc tcgctctgctc aatctctgtta cccagtgcctg c6120
cgcccaggg gataagtcgc tgtctacgctt ggttggacct aagacgtatg ttaacgggata c6180
agggcagcgc gtcgggctgta agggggggct cttggcaacaaga cgggagagtg cagccagcga c6240
cctcaccgga acctagagat ctcagagggtg agcattgaga aagggcccagc tttccgagac c6300
ggagaaaaac gcggatgttag cgggttaacgc gcagggctgg gacagagggag agcagcgagg c6360
actttcagg ggagagcccc ttgttatctct tattcaggtc cctgttagttc caagctctgac c6420
lttgagctcg aattttctgtg tggctgtgccag ggggggcaag cctatggaaca aagggcagca c6480
acggcgcttt ttacgggcttc tggccctttctac tggctctctg tttccttctct c6540
caggtacccc tgatttcggt gataacccat gcgcgcttcag ttcagagctt ttcagagctt c6600
ggcgcgacgg acggcgcagag gcagcgcggcttag ggtgagagtag gcggccgaaa ggcgcgcggag c6660
tacgcagcag ggcgcgtgcttg ggatacctta agtcagggagt gccagagagag c6714
```

FIG. 20C
C-AT2 (Ligation of Fragment 1 and Fragment 2)

c-tagaactag tggatccccc gcggtcacagg aatctgatat caagcgtggg gattttcagg 60
ccaccacact gcacctggagc agtgaatcga caatgcttct gccgtctcct gcggccacctc 120
tctctgtgagc aggctgctgc cgctgcttct ctgcgccttgc ggctgagagat ccggagagag 180
atgtgcccag caagcagtac caatcaccct atgatcagga tcaaaaaaaaccttgc gcacacacagc 240
tcaccaccacgc ctcggctcgag gcctatccag ccacgcccgc ccagcataagc 300
acacgccacg ctcgctttcc tcccggctga gcctgcgggg tggggctcact cagctgttcc 360
acggtgccag ccggctgcgc cctgcttctg ccgtgatgtg cttgttcagg ctggtccagag 420
agtttgcacg cagctggtgag cttgcagtg ccggagagag ctgggtggctcg ctcggtggtg 480
acagcccccg gctcggagtc acaccccggt cctgggtggg ctctggtgctc ctggtgtgatc 540
tggtgtaaat gtttctgagc gatgttttaa aagttgttcag ctcggtggtc ctcggtgctc 600
actctggagga cccgtgacag ggcacagaaac gatcaaggtc ttcggctcag gacgttcagt 660
acgagagaaat tgtgattttg gttcagaggg ctgactacag ggaggtagtc ctggtggtgc 720
attacatctt cttttaaggg aataggagga gcctctggag aatctggtgg gcacccgtcag 780
agaggctcca cgtgggacag gttaccagcg lgaaagttcg tctgggttcgc ggcttgcttgc 840
tgtttaattcg ccagcacttc tggagagcgt ccagagggctgg agctgtgatt ccaatactgc 900
gcaattcagc ccagctgcttc ttcctgcttg agtggaaggg accttcagag cttggaataag 960
acaccccccg cgttacatctt cacccttcac taccctgttt cagcaggttg gacgttcagt 1020
taactttacc cagacacagt acacagttcg tggagagcgt ccagsctggt cccctggtcc 1080
tggcgcatcac taaggcttcgg acgactatttgc gttcagaggg ctgactacag ggaggtagtc 1140
ccctggaagtc ctcaggaagggt ttgctgactc gcctgcttttc gctggtgcag gaggaggttc 1200
aacgctggct gggagctgttc ttaagagcgg tccagctggc ttccttccgg ctcagaggtct 1260
tccacccggt ctttctcttc ttataggtgc aacataacac gacgttcagc ctccttgctg 1320
ggaaagatgtg gatggtcacc ccaaaataac gttctgcttc tgcctgccag ctcagctgcag 1380
tctctggagga ccacccgttc gaagctgggt gcctgtttcc gctggtgttc cccctggtcc 1440
tcgcagaggcc ctcggagaccg gcgcggaggc agagatctga acagctattt gctggttttg 1500
acacacccca actacggtatg gttgctgatag gctgttttttg cagctgcttc gcctggtttc 1560
ttgcttttttt tgggtgcttc ccagctgatg ggaggttttt tggctgttttt gatgtgcttc 1620
ctttagttttgt cagcgttctg gggagttgtg ggaggttttt tttactctgtg ctcagaggtc 1680
cttaccttgca gggagctgcc acacccgttc ccacccggttg ggacgttcag gctggtgtttc 1740
cagagccagt gtggtggtgc aagagagcc aagagccggtt cccctgccac gggagttgtg 1800
tccacccgact gcacccggtgc ccacccggttg ggacgttcag gctggtgtttc 1860
gatgcagcttg gcgcggagac gttcagagtt cccctgccac gggagttgtg 1920
tcgcacccag cggagacccct gcctggagaa tcctggtggc gcctggtggg gacgttcagc 1980
acgacaggttc tccggcgacag gcgggtctgg gcctggtgcag tcctggtggg gccctgggctg 2040
caatcgtttt ctgctgtgtg gtcctgccgg gcgggttttt gcgcggagac 2100
ctgctgccag gcagccgttc ctgctgtgtg gcgcggagac 2160
cgtgggtctgc gcagccgttc ctgctgtgtg gcgcggagac 2220
gaaagccgtc gttgggttttc cggcagatgg gcgcggagac 2280
tcgcagcggc gaaatctgtc atctggtttc gcgcggagac 2340
cgctggttttc ccagccgcag ccacccggttg gcgcggagac 2400
tgagaagcgc gcgggttttt gcgcggagac 2460
cgcgggttttt tcctggttttc gcgcggagac 2520
atggaggggc gcgggttttt gcgcggagac 2580
acggttcagc gcgcggagac gcgcggagac 2640
cttacctgct gcgcggagac gcgcggagac gcgcggagac 2700
FIG. 21B

C-AT2 (Ligation of Fragment 1 and Fragment 2) (cont.)

cctcgcgttcc gccagcgttac gcccttcatt gcccttcpga cagcgttccct tcagggggtcg 2760
cctcgcgttcc agctgcatcg ctctgccgca cggctccgct tcagggggtcg 2820
ttgcgcccttc ccccccttta atgtccaatg gccaaagagc cgtcttccttc 2880
cacccctgatg ccctcgttgc aggtggcagc gcctttgtctg cgtcttccttc 2940
ggggctgcttt ccgccggccg ccggccgctt ccgcctcgtt gcctttgtctg 3000
ggggctgcttt ccgccggccg ccggccgctt ccgcctcgtt gcctttgtctg 3060
ttgcgcccttc ccccccttta atgtccaatg gccaaagagc cgtcttccttc 3210
ttgcgcccttc ccccccttta atgtccaatg gccaaagagc cgtcttccttc 3270
ttgcgcccttc ccccccttta atgtccaatg gccaaagagc cgtcttccttc 3330

aaaggggtga atacccctta ccacgaaact gcggagatac gcggagatac 3390
aaaggggtga atacccctta ccacgaaact gcggagatac gcggagatac 3450
ctgccggtgta cgcttggtta caggtgggtt caggtgggtt caggtgggtt 3460
cctcgtccct ggtatcgggt ggtatcgggt ggtatcgggt ggtatcgggt 3510
tcctcgtcgtc gcgttggtgc gcgttggtgc gcgttggtgc gcgttggtgc 3570
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 3630
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 3690
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 3750
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 3810
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 3870
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 3930
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 3990
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 4050
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 4110
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 4170
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 4230
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 4290
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 4350
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 4410
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 4470
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 4530
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 4590
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 4650
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 4710
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 4770
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 4830
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 4890
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 4950
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 5010
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 5070
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 5130
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 5190
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 5250
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 5310
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 5370
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 5430
ccatcgtcggt gcatcggatg gctgatcggat gctgatcggat gctgatcggat 5490
C-AT2 (Ligation of Fragment 1 and Fragment 2) (cont.)

actgagagtg caccaatagc ggtgtaaat acgccacaga tgcttaaggag aaaaatacccg 5460
catccggaaca ttgtaaccgt taatataagg tttaaattcg cgttaaatgg ttgtaaattc 5520
acgacatatt ttaaaccattt ggccgaagaat gcaaaatact cttaaattgaa aaagaatag 5580
acgccacagt gtaacctagtg tgtttaaggg tcggaaacag agctcactatt aaaaacagc 5640
gtcctcaaag gtaaaggccc aaaaaagcctc ttgctaaagc ataaggtgac acgtgaacc 5700
ttcacaattc taagtttgagc gggtcgagct tttgggcacag ccaaatagc ggacccattg 5760
ggaagccccc agtagttagc gcacctaggg ggaagctggga aaccgctgagc acgtgggacag aaaaagggag 5820
aaagggaagc aagtagctgg cggctaaagc cttctgcttac gttggcggttta 5880
acgccacac ccggcgcgctc taatcgtgag cacccgacgg gccgccgcctc ttgccttatttctc 5940
acgctcagcc acgttggga gggcgatcgc gttgggcgctc cttgcttgac aagccaggcttt 6000
gcaggggggg gcaggggggg gttgccactc ccctctcttc gcgctgcgctc gcgtcactgag 6060
gccgggcgcac ccaagttgcc cggacgcgcgg ggctttgggg gggcgctcgc acgtgagcag 6120
cagagccgcca gagagggagt gcggacaccc atcaactaggg gttcctgatgt ctagattccg 6180
tacgggttac ataatctactc gttaaatgggcc egctggtgctg acgcaccaac gacccccgcc 6240
ccgggattac ataatctagg tattttccc tagtaacgcc aatagggact cttcatttgc 6300
gtcaatggttg ggattttgac cggtaaatgc ccactttgagc agtacatctaa gttgcctctata 6360
tgcacaggtac gccctcattc gagctcaatag acgcgtaaatg ggcggctgctg catatgcctcc 6420
agtcacagac cttttagggc ttctcaattc gccagtaaca atcaactatt gcctcctaattga 6480
ltcagggttgg ggacgggtttt ggttgagcgg cccagcttata cccagcttata gttgaggctc 6540
ggggttttcc aaggtcaccc cccatcagac tcaatggaggg cttgggggta caccacaaac 6600
aacggaattt ctcgaatagt cggccagtggg ccgctcaagt ccgctcaagt gggctagggc 6660
gttgattcactc tamgtagcactct ctagctttatc gaggctcacag atccgccctg 6720
gacgccctcc aagctgttttt gactcattgac ccgacacctg gcggatctgg ccgatcttttg 6780
cttcagagga cctggtactc gcggagccgc aaaaaacgag aagttaactgg taagttttagt 6840
ctttttcttc tttattctctt cttgcttgacct gcggagctagc cggttgagttg ggacccatc 6900
tcagtgtgatt tgtgtcttac ttatgtgctc gaagaggtgag cttcataaac ggagcctctgg 6960
tgcggagattc taccgcgcgc cggctttgtcctgctctc 6981
p43msENC-AT (Ligation of Inverted msEnhancer into p43-AAT*)

FIG. 22A
p43msEncAT (Ligation of Inverted msEnhancer into p43-AAT*) (cont.)

cgatatcatc accaaggtgcc tggaaaatgag aagcagaaaggt ttgctgcaacc tacatttaccc 2460
taactcggata accatgtgct gagacgctgcatc tgggtcaaac tggcagttca 2520
taacagcactgcc cggaggacctg cggaggctcag gacagcagctt tggcagttca 2580
ttcgagctgctc ttcgtgcacc caagtcctcc ttcgtgctcc gctgttcagttc 2700
tacagtagttta ccctggtctccc tacactcactccc ttcctgcttc accaagagtcg 2760
taccttcgctgctc cttctttcagt ccctgctgcct gatcagtcggga ggaggtccacc 2820
ttcgagctgctc ttcgtgcacc caagtcctcc ttcctgcttc accaagagtcg 2880
ttcgagctgctc ttcgtgcacc caagtcctcc ttcctgcttc accaagagtcg 2940
acccacactctgta gccgaggcaggtt cagagatcc tggggtgctgctg 3000
ttcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 3060
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 3120
ttcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 3180
ttcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 3240
ttcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 3300
ttcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 3360
ttcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 3420
ttcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 3480
ttcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 3540
ttcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 3600
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 3660
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 3720
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 3780
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 3840
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 3900
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 3960
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 4020
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 4080
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 4140
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 4200
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 4260
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 4320
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 4380
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 4440
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 4500
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 4560
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 4620
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 4680
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 4740
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 4800
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 4860
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 4920
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 4980
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 5040
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 5100
atcagctctctgta gccgaggcaggtt cagagatcc tggggtgctgctg 5160

FIG. 22B
p43msENC-AT (Ligation of Inverted msEnhancer into p43-AAT) (cont.)

tatatcttt agatggatt aaaaaccttct ttttaattta aaaggatcta ggtgaagact 5220
catggata atctcatgae caaaaacctt taacgtgagtt tttcgcttca ctggagcctca 5280
gaccccgtag aaaaagatcaaa aggatctttct tgagatcttt tttttctgcg cttaatctgc 5340
tgcttcgaaa caaaaaaaccc accggctacc aaggttgggttt gtttgcgga taagagccta 5400
tccacgtcttt ttcgggaaggct aacgtggtctt agcagagcgc agataaccaca taactgtctctt 5460
tctagtgtagc ctagttgtagc ccacattcctt aagaactcttg tagcaccgcc taataacccc 5520
gctctctgtaa ccctgtctacc agtggcctgct gccagctggcg ataagctgctg tttacccgag 5580
tggacgtcacc gcgagatgttt aacgggtaaag gcgcaccgggt cggtgctgaac ggggttgcttct 5640
tgcacacagc ccagcctggga gcrgaagacc tccacccgga acctgcttaacct accgctgagc 5700
cagcggagaa ggcggcaggct cccggaaaggg agaaaagcgg accggatatac cttgtaaccggc 5760
aagggctggga caagagacgc caaggggag cttccagggg gaaacgcctgt gttacttttat 5820
agtctcttgct ggtttccgcca cctctctgatt gaggctcgat tttggtgtag cttcgctaggg 5880
gggggagggc tattggaaaaa cggcagccct cacggcctttt taacgtgcttc gcccttttttc 5940
tggcttttttt ctacatgttt tttttctgcc ttatatcccttc attctcggtta taaccgtattt 6000
acccgcccttgg agtggatctga taccgctgcg cgcagccgaa agaccggagc cagcgagctca 6060
gtagcgagg aagccggaaga gcgccaactta cgcaaccgcc ttctgcccccgc gcgttgccccg 6120
attcattatg gacgggctgc ag 6142

FIG. 22C
p43rmsENC-AT (Ligation of Inverted msEnhancer into p43-LAT*)

FIG. 23A
FIG. 23B
p43rmsENC-AT (Ligation of Inverted msEnhancer into p43-AAT\textsuperscript{*}) (cont.)

gcagtgaactg aagagaattata tcgagtgtgct cccataaccat gagtgataac actgcggccca 5460
acattactctc gacaacgaagc ggaggagccg aggagctaac cgctttttttg cacaacacttg 5520
ggcacagtgt aactcgcctgg gtagctgtggg aacccggagct gatgaaagcc atacaaaccc 5580
acgagcgcgtga ccacccacgtg ccctgaagcaa tggcaacacac cggcaacgaa atattttgtg 5640
gcgaactcttc tcctctagct tccggcgaac aatattatga ctggatggag gcgggataaag 5700
ttgcagggacc atctcagccct tcgcccccttc cgctgctgtgg ctgtatcttct gataaatctg 5760
ggcgcggctga gcgtggggttct cgcggtatca ttgcaagcaat ggggccagat ggtgaacccct 5820
cctgatcgtc atgttatctac aagcacggtgga gtcaaggaac tatggtgatga cgaaatagac 5880
agatcgcgtga gataaggtgg tccactgatta agcttgggtga actgtcagac caagtttact 5940
catatatact ttggattgat ttaaaactctc atatttaatatt taaaaaggtct tagttaaaga 6000
tccttttttaga taacttccttg accaaatccct cttaacgtga gtttttctgc attcgtgagc 6060
cgaccccccgtt agaaaatgctc cccacgccttac cgcgggttgc tttggtgccgc gtaacaagac 6120
gctggctgcta aacaaaaaaacc acccgcttcac cgcgggttgc tttggtgccgc gtaacaagac 6180
tacaacacct ttttcgctgt tcagcagagc gcagataaca accatttacct tacaacatcc 6240
tctagtgtgta gcggtaggttga ggcacacact tcaagaacct tggtaagccg ccatacattc 6300
tcgctcctgcg aactctcgtta cccgggtгccc ttgcnctgag cggtaactgc ctcattaccg 6360
ggttgacgcc aagacgatag ttatcgggata aggccgacgcg gtcggggctg caagggggtt 6420
cgagcacaac gcccgctttgg cagcggagca cccacacccga actgagatac ctacacgctg 6480
agcattgaga aagccgccac cgctccggaag ggaccagaagc ccggcagttat ctcgtaagcc 6540
gcgagggtcgg aacagcaagag cgacaagagcg agctccacag cggaaacgcgc tggtaatctc 6600
atagctcctgt cgggtlctgc acccttcctac ttagcgcgtct atcctgtctg taatctctcac 6660
gggggcgagg ccataaggaat aacgccccca cccgggcttt gttacgctgct tggctcttctt 6720
gcggctgtt cgcacagctg ccgtctccttg cgtcttttttg tgctcttccct ctataaatctg 6780
ttaccccttt tagtgagcctg gatacgcccc gcggcaacgcgc ggacccgagc cggagcgagt 6840
cagtgacgac gcagcgccagaa ggcggcsgga taccacgcaag gcggcgctcgc gcgcgtggcg 6900
cgatttcatt atgcggcgtc gcag 6924

FIG. 23C
FIG. 24
p43msENCB-AT (Ligation of msEnhancer into p43CB-AT*)

---

FIG. 24A
p43msENCB-AT (Ligation of msEnhancer into p43CB-AT*) (cont.)

cagctccacgc tggaccaaggg caaatgcctgc ttcctcagcc agggctctgaa gctagtggtat 2760
aagtattttgg aggatgtaaa aagtgtgtaa ccactcagaa cctctctgtg caaattcgcgg 2820
gacacccagaag acagacacac gagttctgtg acagctgtgac tcaagggaaca 2880
atttggtgatt ttgctcaagg gccctgacag cagacacact ttgcctgtga gaattacatt 2940
ttctcttaag gcacataatgg gacacctccgt gaaatctaaag acacccaggaag aagagactttc 3000
caagtggacc goctgtatcag ctgtaaatgga gtagctttgaa catgtgctaac 3060
atccgaaggact gtaaacaagt gcctggctcag tttggctgtca cctagttgacct ggcacaatcgc 3120
accgcatact cccctctcctc tcagaggggg aacactcagcc accctggaaga tgaactaccc 3180
caagataata ccaaccaagt tcccggaaat gaaagcagaaag ggtctggcagc cttcattatta 3240
cccaaaaccttg cccatcaagt gaaatcgaccg cctgctgacaag cctgctgacaag 3300
actaattggtat ctcagcaattgg ggtctgacctct tccggggtcaca cagaggggagc aaccctcgaag 3360
ctctctcaag cccgtctcaata ggtctgtcagct acactcgacg aagaaagcttc gcagcttggat 3420
gggcccagatg ttttagaggct catcacccagct tctatcccccc cgaggcttcaaa gctctgaacaa 3480
ccctttgcttc tccataatgat tccaaacact gcacactcctccc tccctctctcatt ggcgaagatac 3540
gtgtaattcaca ccaccaataact gctgcctcttc gcctgctcaac cccctctctcacc tcccctggtgc 3600
ccccctgcttcgt gatgcacattt ccaagaggggctcagagggatc cccctgcggcag 3660
cccctgacgcc cggcgcgcacag ctcgctgacg acatggtaag atacatgattg gatgctatattg 3720
aaaaactttagc tgcacatta gtaaaaaatg gcttttttgtg tgaataattgtg gatgcacatta 3780
cttttatattt aacacttaata agttgcaata aacaagtttaa ccaacaacatt tgcacatttt 3840
atatggattcct cgttttcaggg gagaaggggg aggttttttt aagctcagtaa aaccctctcaca 3900
aatgggttaa aatcctgtaag gtagcagcag acctagtgaa cccctctctctg gtagctgtggc 3960
ctgcccctcgc ttcctgctac gcacagctgg ctgctgacatgg caagggctgcgc ggacagtctt 4020
gtgctggcag ctctgatgag cagcagcagg cgcagagagg gattgctgcaa cccctctctctc 4080
ccccctctctcc cgccgctcctg tataagcacag gaggcccgac ccatctgcgctc tccaaacagc 4140
ttcgtctgcca tgaatggcca acgcgcgccg ccgctgctgta ggcgccctatt aagcgcgcagc 4200
gttgctgggct ttcgcagcag cttgctgctgc cactccgctgcc gcgcgctcttt gccgcctctct 4260
ttgctgtcttct ggctgctgctg ggctgctgctg ccgctgcctgtg aaagtccacctctgcct 4320
cggggtctcct ctgagttggttc cggaggttagc gttttctgag cagcagagagc ccaaaacttct 4380
egattaggtct agatgctcagcg acaggttcaggg gctgctgctgctg ctgctgctgctg 4440
acgctgagact cccgctcttc tatagtgcga aacctgctgac ccagctgctgac cagctgctgac 4500
cttctctctctc ttatatatatg tggggtgttg gatggctgctg ucaggtttctc ctatggta 4560
aaaaatatgc gcagctcttcag cccagctcttcag cccagctcttcag cccagctcttcag cccagctcttcag 4620
atctctctctct ccgctgctgctg ccgctgctgctg ccgctgctgctg ccgctgctgctg ccgctgctgctg 4680
gccatctctct gcctgtgcag gcctgtgcag gcctgtgcag gcctgtgcag gcctgtgcag 4740
caacccgctcga cgccgcctctcg ggctgcctgtg ctgcctgcctgc gcctgtgcac gcctgtgcctgc 4800
tgacctgctgt gccggtgcag atgtggtcga gttttgttcct gcataccgcgc cccagctcttcag 4860
gacaaaggggg ccgtgtgtcag ccctctctctc tataggttaa tgcagtgatg ataataagtttt 4920
cctgagctgcc gaggtagcag ctcggggtgaa ctagtttgctgg ctgctgctgctg ctgctgctgctg 4980
ccaatatata ctcgatctat ttgcctcgac cggagttaa aatgctctcag tggctggataa 5040
aatattgaaa aagggagagt ataggtatctt acaatgccgct gtgcctgtgg ctggcgctgtgg 5100
ttggcagcctg ccggatgttg cttggtgctc cccagctcttcag gcctgtgcctgc gcctgtgcctgc 5160
tgtaagctca cgctgctgctg ccgctgctgctg ccgctgctgctg ccgctgctgctg ccgctgctgctg 5220
tcctgagaag ccgggtctctgc gcaaagttcc ttcctatgtgc gacactctctt aagctctctct 5280
atcagtgctg ccgttatctcgg ctgtgtctgag gggcgctctg ccgccgcctccc gcagagagttccc 5340
actactctctc gaatagcggt tcctctctctct gtcgctgatat gcacctcgctc gcagagagttccc 5400

FIG. 24B
p43msENCB-AT (Ligation of msEnhancer into p43CB-AT*) (cont.)

gcatgacagt aagagaatta tgcatggtgc ccataaccat gatgataac acgtgagcga 5460
acctactctg gcaacagctc ggagacgca gaagcgttaac ctgcttttctt ccataacatg 5520
gggatcatgta aacgtccctt gcattctgctgg aacggagctaat gtaaagccc attcacaacacg 5580
acgagctgta caccacgtgag cctgtagccttg gaccacaactac atgtggtccttg gacgctaggg 5700
ttcgagcacc accttctgcctg tcggcctccac gttctgttcgct gttctgtccttg gataaactctg 5760
gagcgggtgta gcgctggtgttt cgcgcgctatc ttgacgacacct gggcgggatg gtaagcctt 5820
cctgccgtgtg aagtcggcaac ctcacccacac taagcttgac caagtttact 5880
agtccgctga gattaggtgcc tacactgattag cagcatttgta actgctgcagcac caagtttact 5940
catatatact ttgaggtgatatt taaaaacctc atttttaattt taaaagggattc taggtgaaac 6000
tcttttttga ttaatctctc aggcaaatccc ttaaactgtgact gcttttcttc cactgagcgt 6060
ccagccccctg aagaaagctc aagagatcttt tttgagataccc ttctttttcctg tgcgttaacctct 6120
gctcttggca aacaaaaaa gccacccgctc caagcgtgttgc tcgtttgctgct gcacgagagc 6180	

cacaactctt ttgccggaag actacgctctag ttcagcagacgc gcagatacaact aactcagcgc 6240
tctctggtatt ggcgtgttatttg cccacactcc actcgccttgc cctggtcgacag 6300
tctctctgct aactcgcgtta ctgctggttcctgtc tggcatgcgttg cgataaggtg 6360
gggtgacgcc aagcgcgtatt ttacccggata ggcgcaagctgc acgctgtgctgg 6420
cgtgcacaca gcccacctgta ggcgaacggct ctcacccactc actgacgatag ctcagcgttg 6480
acgtgagaga aacgcaagcc ctcacccgtct ggcagagttgg gc gagaggtcag 6540
gacggctggtgc aagcggcggag cgcggagagc gcgtctcagct gcgctgtgtgc 6600
atagccgctg cgggttcgcct gtcctccgcag cgtcctcagctc tcgtctcctgc 6660
ggcccggagt cgggttgccttg gtcctccgcag cgtcctcctctc tcggttgcct 6720
gcttggcccttt ggtcctccaga ctgctctgttg gacgatctgctg tgcgctgtgctg 6780
ttacggctct tgtgctgtgc cagactgcgct gcgacccggag acgcacggctgc gcggttgcgag 6840
cagtgcagca ggcagcggcag aggaccccaact gcctgccctgcc gcgctttgcgc gcaagttgtg 6900
cgatttcatt attctcgggtgc gcaagttttgc 6924

FIG. 24C
FIG. 25
p43rmsENC-BAT (Ligation of Inverted msEnhancer into p43CB-AT*)

ggggggggggg ggggggggttg gcggacctccct ctcggcggcg ccctctggctgc tcgctctgcc actggagggcg 60
ggggggcccac ggggcccggcg ggcccggggcg ggcccggggcg ggcccggggcg ggcccggggcg ggcccggggcg 120
cgcaaggcag gggagggcgt gcgggcatcg cgccgagccg cgccgagccg cgccgagccg cgccgagccg cgccgagccg 180
ttactccggg cggaatgtgg gccggcagcgc cgccggctggcc cttgggacgct gcgtgatggg gcggcgcgct 240
cgggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 300
cgggggggggg gcgtgggggg gcgtgaggggg gcgtgaggggg gcgtgaggggg gcgtgaggggg gcgtgaggggg 360
ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 420
ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 480
ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 540
ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 600
ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 660
ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 720
ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 780
ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 840

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 900

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 960

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 1020

cgtgcggggg ggggggggcgg gcgggggaggg cgccgggggg gcgggggggg gcgggggggg gcgggggggg gcgggggggg 1080
cggcgggggg gcgggggggg gcgggggggg gcgggggggg gcgggggggg gcgggggggg gcgggggggg gcgggggggg 1140

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 1200

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 1260

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 1320

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 1380

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 1440

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 1500

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 1560

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 1620

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 1680

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 1740

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 1800

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 1860

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 1920

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 1980

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 2040

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 2100

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 2160

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 2220

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 2280

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 2340

ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg ggggggggggg 2400

**FIG. 25A**
cagaagacag atacattccca ccacagctcag gatcaccaaa ccctcaaaca gaatcccucccc 2460
aacatgctctt actgctctatac cagctctcagt ccgactcgtt cacaagtgct ccagccgacc 2520
aatatatcctt tcctctctct tatacgccgct acaatctctctt cagctctctct ccctcctctt 2580
aacagctctgta cacagctctctt aacatcctctt ctacatctctt ccgactcgtt gcacacgacg 2640
gacagctcgt ttcctctctt cctcctcttt cagctctcgg acatcctcttt ccctcctctt 2700
aatatatcctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 2760
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 2820
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 2880
aatatatcctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 2940
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 3000
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 3060
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 3120
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 3180
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 3240
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 3300
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 3360
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 3420
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 3480
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 3540
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 3600
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 3660
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 3720
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 3780
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 3840
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 3900
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 3960
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 4020
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 4080
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 4140
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 4200
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 4260
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 4320
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 4380
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 4440
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 4500
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 4560
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 4620
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 4680
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 4740
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 4800
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 4860
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 4920
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 4980
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 5040
aacatcctctt tctctctctt cctcctcttt cctcctcttt cctcctcttt ccctcctctt 5100

FIG. 2B
p43rmsENCb-AT (Ligation of Inverted msEnhancer into p43CB-AT*) (cont.)

```
ttgcggcatt ttgccttctct gttttttgtcct accacgaacac gcgtggtgaaac gttaaagagtct 5160
tctgaagatca ttggtggggc caagagttggctt acatacgcaacct ggatctcaac aacggtatag 5220
ttcctgagag ctttccgctcc tcacactgatgatt gacacttgctt ctgctgacctctag 5280
tagtgcggag ctcctttatct cctggtgacag cctgggaagcact gccacgcagct gcctgcaagct 5340
actgttctcat cgtaccttcgt gacgcttccgc cagtagctctg gctcttttggct 5400
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 5460
actgttctcat cgtaccttcgt gacgcttccgc cagtagctctg gctcttttggct 5520
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 5580
actgttctcat cgtaccttcgt gacgcttccgc cagtagctctg gctcttttggct 5640
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 5700
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 5760
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 5820
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 5880
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 5940
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 6000
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 6060
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 6120
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 6180
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 6240
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 6300
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 6360
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 6420
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 6480
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 6540
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 6600
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 6660
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 6720
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 6780
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 6840
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 6900
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc 6960
actgtgagc tcgcagatc catttttactgc gacagagcact gcctgtgattac ctccttttctttc
```

**FIG. 25C**
ADENO-ASSOCIATED VIRAL VECTORS FOR THE TREATMENT AND PREVENTION OF DIABETES

CROSS-REFERENCE TO A RELATED APPLICATION

[0001] This application claims priority from provisional application U.S. Serial No. 60/083,025, filed Apr. 24, 1998.

[0002] The subject invention was made with government support under a research project supported by National Institute of Health NHLBI Grant No. HL 59412. The government has certain rights in this invention.

BACKGROUND OF THE INVENTION

[0003] Alpha-1-antitrypsin (AAT) deficiency is the second most common monogenic lung disease in man, accounting for approximately 3% of all early deaths due to obstructive pulmonary disease. AAT protein is normally produced in the liver, secreted into the serum and circulated to the lung where it protects the fine supporting network of elastin fibers from degradation by neutrophil elastase. Current therapy for AAT deficiency includes avoidance of cigarette smoke exposure and weekly intravenous infusions of recombinant human AAT (hAAT) protein. Attempts to devise gene therapy strategies to replace AAT either in the lung itself or within any of a number of other tissues which are capable of AAT secretion have been limited by the short duration of expression from some vectors and by the relatively high circulating levels of AAT which is required for therapeutic effect. Methods of gene therapy have been described in U.S. Pat. No. 5,399,346.

[0004] It has recently been demonstrated that adeno-associated virus (AAV) vectors are capable of stable in vivo expression and may be less immunogenic than other viral vectors (Flotte et al., 1996; Xiao et al., 1996; Kessler et al., 1996; Joss et al., 1998). AAV is a non-pathogenic human parvovirus whose life cycle naturally includes a mechanism for long-term latency. In the case of wild-type AAV (wtAAV), this persistence is due to site-specific integration into a site on human chromosome 19 (the AAVS1 site) in the majority of cells (Kotin et al., 1990), whereas with recombinant AAV (rAAV) vectors, persistence appears to be due to a combination of episomal persistence and integration into non-chromosomal 19 locations (Affione et al., 1996; Kearns et al., 1996). Recombinant AAV latency also differs from that of wtAAV in that wtAAV is rapidly converted to double-stranded DNA in the absence of helper virus (e.g., adenovirus) infection, while with rAAV leading strand synthesis is delayed in the absence of helper virus (Fisher et al., 1996; Ferrari et al., 1996). U.S. Pat. No. 5,658,785 describes adeno-associated virus vectors and methods for gene transfer to cells.

[0005] Kessler et al. (1996) demonstrated that murine skeletal myofibers transduced by an rAAV vector were capable of sustained secretion of biologically active human erythropoietin (hEpo), apparently without eliciting a significant immune response against the secreted hEpo. See also U.S. Pat. No. 5,858,351 issued to Podolskoff et al. Likewise, Murphy et al. (1997) have observed the expression and secretion of sustained levels of leptin in ob/ob mice after AAV muscle transduction. Brantly et al. (U.S. Pat. No. 5,439,824) disclose methods for increasing expression of AAT using vectors comprising intron II of the human AAT gene. However, the level of leptin expression observed was only in the range of 2 to 5 ng/ml. Therapy for AAT deficiency requires serum levels of at least about 800 μg/ml. Thus, there remains a need in the art for a means of providing therapeutically beneficial levels of a protein to a person in need of such treatment.

BRIEF SUMMARY OF THE INVENTION

[0006] The subject invention concerns materials and methods for gene therapy. One aspect of the invention pertains to vectors which can be used to provide genetic therapy in animals or humans having a genetic disorder where relatively high levels of expression of a protein is required to treat the disorder. The vectors of the invention are based on adeno-associated virus (AAV). The vectors are designed to provide high levels of expression of heterologous DNA contained in the vector. In one embodiment, the vectors comprise AAV inverted terminal repeat sequences and constitutive or regulatable promoters for driving high levels of gene expression. The subject invention also pertains to methods for treating animals or humans in need of gene therapy, e.g., to correct a genetic deficiency disorder.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] FIG. 1 shows rAAV-AAT vector cassettes used according to the subject invention. The A-AT and B-AT constructs contain the promoters from the small nuclear RNA genes, U1a and U1b, respectively. The C-AT construct contains the CMV promoter, whereas the E-AT vector uses the human elongation factor 1-α (τEF in the figure) promoter. TTR refers to AAV inverted terminal repeat; An refers to polyA signal; Tk refers to the HSV thymidine kinase promoter; neo refers to the Tn5 neomycin phosphotransferase gene.

[0008] FIG. 2 shows hAAT secretion rates in vitro from transiently transfected murine C2C12 myoblast cell line using expression vectors according to the subject invention. C-AT does not differ significantly from E-AT, but both differ from A-AT and B-AT (p < 0.05) AAT expression was detected using an ELISA assay specific for human AAT.

[0009] FIG. 3 shows hAAT secretion rates in vitro from stably transfected murine C2C12 myoblast cell line using viral particles comprising expression vectors according to the subject invention. The mean rates of secretion from G418-resistant cultures 1 mo after transfection with either packaged B-AT vector or packaged C-AT vector are shown. In each instance, a “low” multiplicity transduction (4×10⁵ particles/cell) and a high multiplicity transduction (4×10⁶ particles/cell) were performed. E-AT “low” and “high” are greater than “high” multiplicity C-AT (p < 0.02) but are not significantly different from each other (n=3). AAT expression was detected using an ELISA assay specific for human AAT.

[0010] FIG. 4 shows additional constructs tested for hAAT expression. The murine myoblast C2C12 cells were grown in 35-mm wells with approximately 4×10⁵ cell per well and were transfected with 5 μg of the appropriate plasmid DNA using SUPERFECT transfection (Qiagen Inc., CA). Secretion of hAAT into the medium was assessed at 2 days after transfection using an antigen-capture ELISA. Each bar represents the mean of results from three experiments (triplicate in each experiment).
[0011] Data from transfection experiments indicate that the expression from p43CB-AT was at least three times higher than that from C-AT in vitro.

[0012] FIGS. 5A and 5B show sustained secretion of therapeutic levels of hAAT using either the C-AT vector or the E-AT vector in either SCID or C57BL mice. FIG. 5A shows the mean total serum levels of hAAT observed in groups of either SCID (squares) or C57BL (circles) mice receiving either low dose (5x10^18 particles) (open symbols) or high dose (1.4x10^19 particles) (filled symbols) single injections into muscle of the C-AT vector measured at time points ranging from 1 to 16 wk after injection. For each strain, the high-dose curve is significantly different from the low-dose curve (P=0.009 for SCID, P=0.02 for C57BL), but the strains do not differ from each other. FIG. 5B shows analogous data with the E-AT vector. None of these differences were significant.

[0013] FIG. 5C shows long term secretion of hAAT from murine muscle transduced with C-AT. C57BL/6 or C57BL/6-SCID mice received 3.5x10^10 IU, 1.4x10^19 particles/mouse. One year after injection, serum hAAT levels were still 400 μg/ml in C57BL/6-SCID and 200 μg/ml in C57BL/6. This level are comparable with the peak levels observed (800 or 400 μg/ml, respectively).

[0014] FIG. 6 shows an immunoblot of sera taken from several of the C-AT vector-treated mice at 11 weeks after vector administration. Ten microliters of a 1:100 dilution of serum was electrophoresed by 10% SDS/PAGE, blotted, and incubated with 1:1,500 dilution of goat anti-hAAT-horse radish peroxidase conjugate (Cappel/ICN). Samples from three high-dose SCID (h1-h3), one high-dose C57BL (h3), and three low-dose C57BL (h1-3) were included, along with one negative control (saline-injected-sal) serum to indicate the level of reactivity with endogenous mAAT. As a standard, hAAT was added either to negative-control C57BL1 serum (first hAAT lane) or to PBS (second hAAT) lane to final equivalent serum concentration of 100μg/ml.

[0015] FIGS. 7A and 7B show that some BALB/c mice mount humoral immune responses to hAAT, which correlate with lower serum levels but no observable toxicity. FIG. 7A shows serum hAAT levels and FIG. 7B shows serum anti-hAAT antibody levels determined by ELISA performed on serum taken from mice injected with 1x10^11 particles of the C-AT vector. Each set of symbols represents an individual animal (2, no. 1; A, no. 2; , no. 3). Note the inverse correlation between the presence of antibody and the presence of circulating hAAT.

[0016] FIG. 8 shows the persistence of rAAV-AAT vector DNA in high molecular weight form. PCR products were amplified from DNA prepared by Hirt extraction from three SCID mice injected 16 wk earlier with 5x10^11 resistant-particles of C-AT and analyzed by Southern blot. The high molecular weight Hirt pellet (genomic DNA lanes) and the low molecular weight supernatant (episomal DNA lanes) were analyzed separately. Control lanes include a sample in which an hAAT cDNA plasmid was the template DNA (+) and a control in which water was the template (−). In this internal PCR reaction, a 500-bp product is expected regardless of whether or not the vector genome is integrated.

[0017] FIG. 9 shows serum hAAT in C57B 1/6 mice transduced with C-AT and p43CB-AT. C57B1/6 mice were injected in muscle with C-AT (3.5x10^10 IU/mouse, 1x10^12 particles/mouse) or p43CB-AT (6x10^7 IU, 1x10^8 particles/mouse). The level of hAAT from p43CB-AT were projected based on an estimation of the equivalent dosage of infectious unit of C-AT.

[0018] FIG. 10 shows enhancement of CMV promoter activity by a synthetic enhancer in C2C12 cells. The murine myoblast C2C12 cells were grown in 35-mm wells with approximately 4x10^5 cell per well and were transduced with 5 μg of p43msENC-AT vector DNA using SUPERFECT transfection (Qiagen Inc., CA). Secretion of hAAT into the medium was assessed at 2 days after transfection using an antigen-capture ELISA. Each bar represents the mean of results from one experiment (triplicate).

[0019] FIG. 11 shows secretion of hAAT from mouse liver cells (HO 15) transduced with different constructs. The murine liver cells (HO15) were grown in 35-mm wells with approximately 4x10^5 cell per well and were transduced with 5 μg of the plasmid DNA using LIPOFECTAMINE reagents (Life Technologies Inc, MD). Secretion of hAAT into the medium was assessed at 2 days after transfection using an antigen-capture ELISA. Each bar represents the mean of results from two experiments (triplicate).

[0020] FIG. 12 shows secretion of hAAT from mouse liver cells (HO15) transduced using different methods. The murine liver cells (HO15) were grown in 35-mm wells with approximately 4x10^5 cell per well and were transduced with 5 μg of the p43CB-AT vector using SUPERFECT (Qiagen Inc., CA), FuGENE (Boehringer Mannheim Co, IN), LIPOFECTAMINE (Life Technologies Inc, MD) reagents and Calcium phosphate (CA-PO4) transfection. Secretion of hAAT into the medium was assessed at 2 days after transfection using an antigen-capture ELISA. Each bar represents the mean of results from one experiment (triplicate).

[0021] FIG. 13 shows hAAT secretion from mouse liver transduced with rAAV. C57B1/6 mice were injected with either p43CB-AT, C-AT or E-AT vector either by portal vein or tail vein injection. PV=portal vein injection. TV=tail vein injection.

[0022] FIG. 14 shows serum hAAT levels in C57B1/6 mice after intratracheal (IT) injection of C-AT or p43CB-AT vector. Mice received either 10^8 IU of C-AT (open circles), 10^6 IU of p43CB-AT (open triangles) or 10^10 IU of p43CB-AT (open squares).

[0023] FIG. 15 shows a map and nucleotide sequence for the vector of the present invention designated as C-AT.

[0024] FIG. 16 shows a map and nucleotide sequence for the vector of the present invention designated as E-AT.

[0025] FIG. 17 shows a map and nucleotide sequence for the vector of the present invention designated as dE-AT.

[0026] FIG. 18 shows a map and nucleotide sequence for the vector of the present invention designated as p43C-AT.

[0027] FIG. 19 shows a map and nucleotide sequence for the vector of the present invention designated as p43C-AT-IN. This vector includes intron II from human AAT gene to enhance transcription.

[0028] FIG. 20 shows a map and nucleotide sequence for the vector of the present invention designated as p43CB-AT.
Figure 21 shows a map and nucleotide sequence for the vector of the present invention designated as C-AT2.

Figure 22 shows a map and nucleotide sequence for the vector of the present invention designated as p43msENC-AT. This vector is the same as the p43 msENC-AT vector except that the enhancer sequence is in an opposite orientation.

Figure 23 shows a map and nucleotide sequence for the vector of the present invention designated as p43msENC-AT. This vector is similar to p43CB-AT but also comprises an enhancer sequence upstream of the CMV promoter.

Figure 24 shows a map and nucleotide sequence for the vector of the present invention designated as p43msENC-AT. This vector is similar to p43CB-AT but also comprises an enhancer sequence upstream of the CMV promoter.

Figure 25 shows a map and nucleotide sequence for the vector of the present invention designated as p43msENC-AT. This vector is the same as p43 msENC-AT except that the enhancer sequence is in an opposite orientation.

Detailed Disclosure of the Invention

The subject invention pertains to novel materials and methods for providing gene therapy to a mammal or human having a condition or disorder, such as genetic deficiency disorders, where high levels of expression of a protein are required to treat the disorder or condition. In one method of the subject invention, a viral vector is introduced into cells of an animal wherein a therapeutic protein is produced, thereby providing genetic therapy for the animal.

In one embodiment, a method of the invention comprises introducing into an animal cell or tissue an effective amount of viral particles or vector comprising a recombinant genome which includes heterologous polynucleotide encoding a protein useful in genetic therapy and that can be expressed by the cell or tissue. Expression of the heterologous polynucleotide results in production of the protein. Preferably, the therapeutic protein encoded by the heterologous polynucleotide is a serum protein. In a preferred embodiment, vector material comprising the heterologous polynucleotide is integrated into a chromosome of the cell of the host animal.

In one embodiment, a recombinant polynucleotide vector of the present invention is derived from adeno-associated virus (AAV) and comprises a constitutive or regulatable promoter capable of driving sufficient levels of expression of the heterologous DNA in the viral vector. Preferably, a recombinant vector of the invention comprises inverted terminal repeat sequences of AAV, such as those described in WO 93/24641. In a preferred embodiment, a vector of the present invention comprises polynucleotide sequences of the pTR-UF5 plasmid. The pTR-UF5 plasmid is a modified version of the pTRq2-UF/UF1/UF2/UF3 series of plasmids (Zolotukhin et al., 1996; Klein et al., 1998). The pTR-UF5 plasmid contains modifications to the sequence encoding the green fluorescent protein (GFP).

Promoters useful with the subject invention include, for example, the cytomegalovirus immediate early promoter (CMV), the human elongation factor 1-alpha promoter (EF1), the small nuclear RNA promoters (U1a and U1b), α-myosin heavy chain promoter, Simian virus 40 promoter (SV40), Rous sarcoma virus promoter (RSV), adenovirus major late promoter, β-actin promoter and hybrid regulatory element comprising a CMV enhancer/β-actin promoter. These promoters have been shown to be active in a wide range of mammalian cells.

The promoters are operably linked with heterologous DNA encoding the protein of interest. By “operably linked,” it is intended that the promoter element is positioned relative to the coding sequence to be capable of effecting expression of the coding sequence.

Promoters particularly useful for expression of a protein in muscle cells include, for example, hybrid CMV enhancer/β-actin promoters, CMV promoters, synthetic promoters and EF1 promoter. Promoters particularly useful for expression of a protein in liver cells include, for example, hybrid CMV enhancer/β-actin promoters and EF1 promoters.

Also contemplated for use with the vectors of the present invention are inducible and cell type specific promoters. For example, Tet-inducible promoters (Clontech, Palo Alto, Calif.) and VP16-LexA promoters (Nettelbeck et al., 1998) can be used in the present invention.

The vectors can also include introns inserted into the polynucleotide sequence of the vector as a means for increasing expression of heterologous DNA encoding a protein of interest. For example, an intron can be inserted between a promoter sequence and the region coding for the protein of interest on the vector. Intron can also be inserted in the coding regions. Exemplified in the present invention is the use of intron II from the hAAT gene in a subject vector. Transcriptional enhancer elements which can function to increase levels of transcription from a given promoter can also be included in the vectors of the invention. Enhancers can generally be placed in either orientation, 3' or 5', with respect to promoter sequences. In addition to the natural enhancers, synthetic enhancers can be used in the present invention. For example, a synthetic enhancer randomly assembled from Sp5-12-derived elements including muscle-specific elements, serum response factor binding element (SRE), myocyte-specific enhancer factor-1 (MEF-1), myocyte-specific enhancer factor-2 (MEF-2), transcription enhancer factor-1 (TEF-1) and SP-1 (Li et al., 1999; Deshpande et al., 1997; Stewart et al., 1996; Mitchell et al., 1989; Briggs et al., 1986; Pithak et al., 1991) can be used in vectors of the invention.

Heterologous polynucleotide in the recombinant vector can include, for example, polynucleotides encoding normal, functional proteins which provide therapeutic replacement for normal biological function in animals afflicted with genetic disorders which cause the animal to produce a defective protein, or abnormal or deficient levels of that protein. Proteins, and the polynucleotide sequences that encode them, which can be provided by gene therapy using the subject invention include, but are not limited to, anti-protases, enzymes, structural proteins, coagulase factors, interleukins, cytokines, growth factors, interferons, and lymphokines. In an exemplified embodiment, heterologous DNA in a recombinant AAV vector encodes human alpha-1-antitrypsin protein.

As those of ordinary skill in the art will appreciate, any of a number of different nucleotide sequences can be
used, based on the degeneracy of the genetic code, to produce a protein of interest for use in the present invention. Accordingly, any nucleotide sequence which encodes a protein of interest comes within the scope of this invention. Biologically active fragments and variants of a protein of interest can easily and routinely be produced by techniques well known in the art. For example, time-controlled Bal31 exonuclease digestion of the full-length DNA followed by expression of the resulting fragments and routine screening can be used to readily identify expression products having the desired activity (Wei et al., 1993).

[0043] As used herein, the terms “polynucleotide” and “polynucleotide sequence” refer to a deoxyribonucleotide or ribonucleotide polymer in either single- or double-stranded form, and unless otherwise limited, would encompass known analogs of natural nucleotides that can function in a similar manner as naturally-occurring nucleotides. Polynucleotide sequences can include both DNA strand sequences, such as that which is transcribed into RNA, and RNA sequences. The polynucleotide sequences include both full-length sequences as well as shorter sequences derived from the full-length sequences. It is understood that a particular polynucleotide sequence includes sequences, such as degenerate codons of the native sequence or sequences, which may be introduced to provide codon preference in a specific host cell. Polynucleotides of the invention encompass both the sense and antisense strands as either individual strands or in the duplex.

[0044] The polynucleotides of the subject invention also encompass equivalent and variant sequences containing mutations in the exemplified sequences. These mutations can include, for example, nucleotide substitutions, insertions, and deletions as long as the variant sequence functions in a manner similar to the exemplified sequences.

[0045] The gene therapy methods of the invention can be performed by ex vivo or in vivo treatment of the patient’s cells or tissues. Cells and tissues contemplated within the scope of the invention include, for example, muscle, liver, lung, skin and other cells and tissues that are capable of producing and secreting serum proteins. The vectors of the invention can be introduced into suitable cells, cell lines or tissue using methods known in the art. The viral particles and vectors can be introduced into cells or tissue in vitro or in vivo. Methods contemplated include transfection, transduction, injection and inhalation. For example, vectors can be introduced into cells using liposomes containing the subject vectors, by direct transfection with vectors alone, electroporation or by particle bombardment. In an exemplified embodiment, muscle cells are infected in vivo by injection of viral particles comprising recombinant vector into muscle tissue of an animal. In another embodiment, liver cells are infected in vivo by injection of recombinant virus into either the portal vein or peripheral veins.

[0046] The methods and materials of the subject invention can be used to provide genetic therapy for any conditions or diseases treatable by protein or cytokine infusion such as, for example, alpha-1-antitrypsin deficiency, hemophilia, adenosine deaminase deficiency, and diabetes. The methods and materials of the subject invention can also be used to provide genetic therapy for treating conditions such as, for example, cancer, autoimmune diseases, neurological disorders, immunodeficiency diseases, and bacterial and viral infections. For example, the present invention can be used to provide genetic therapy to a patient wherein cells from the patient are transformed to express and produce interleukins such as interleukin-2.

[0047] Using the materials and methods of the subject invention, the skilled artisan can for the first time provide therapeutically effective levels of a serum protein through genetic therapy. In a preferred embodiment, the therapeutically effective level of serum protein that can be obtained using the subject materials and methods is at least about 1 µg/ml of protein in serum. Preferably, the level of serum protein that can be obtained using the present invention is at least about 100 µg/ml in the serum. Most preferably, the level of serum protein that can be obtained by the present invention is at least about 500 µg/ml of protein in the serum.

[0048] Animals that can be treated with the materials and methods of the invention include mammals such as bovine, porcine, equine, ovine, feline and canine mammals. Preferably, the mammals are primates such as chimpanzees and humans.

[0049] The subject invention also concerns cells containing recombinant vectors of the present invention. The cells can be, for example, animal cells such as mammalian cells. Preferably, the cells are human cells. More preferably, the cells are human myofibers or myoblasts, hepatocytes or lung cells. In a preferred embodiment, a recombinant vector of the present invention is stably integrated into the host cell genome. Cell lines containing the recombinant vectors are also within the scope of the invention.

[0050] In an exemplified embodiment, recombinant AAV vectors comprising the human AAT gene (hAAT) using either the CMV promoter (AAV-C-AT) or the human elongation factor 1-alpha (EF1) promoter (AAV-E-AT) to drive expression were constructed and packaged using standard techniques. A murine myeloblast cell line, C2C12, was transduced with each vector and expression of hAAT into the medium was measured by ELISA. In vitro, the EF1 promoter construct resulted in 10-fold higher hAAT expression than the CMV promoter construct. In vivo transduction was performed by injecting doses of up to 1.4x10^12 transduced vector into each skeletal muscle of a number of different strains of mice (including C57B1/6, Balb/c, and SCID). In vivo, the CMV promoter construct resulted in higher levels of expression, with sustained serum levels up to 800 ng/ml in SCID mice, approximately 10,000-fold higher than those previously observed with proteins secreted from AAV vectors in muscle. At lower doses in both C57B1/6 and SCID mice, expression was delayed for several weeks, but was sustained for over 10 weeks without declining. Thus, increasing dosage AAV vector via transduction of skeletal muscle provides a means for replacing AAT or other serum proteins.

[0051] Transduction of muscle using the vectors of the subject invention presents several advantages in that it is stable, non-toxic, and relatively nonimmunogenic. Furthermore, certain transcription promoters, such as the CMV promoter, which appear to be markedly down-regulated in other contexts have been found to remain active over time as used in the subject invention. Using the materials and methods of the subject invention, microgram/ml serum levels of a therapeutic protein can be achieved. In an
The levels of in vivo protein expression achieved represent a 10,000-fold or more increase over previously published results. In addition, a dose-effect relationship was demonstrable within the range of doses used, providing for further increases in expression levels as vector dose is increased.

In another embodiment of the invention, recombinant AAV vectors i.e., C-AT, p43C-AT, P43CB-AT, E-AT and DE-AT comprising the human AAT gene (hAAT) using were constructed and packaged using standard techniques. A murine liver cell line, HO15, was transfected with each vector and expression of hAAT into the medium was measured by ELISA. In vitro, transduction with the p43CB-AT vector exhibited the highest level of hAAT expression. In vivo, the p43CB-AT vector also gave higher levels of expression. Portal vein administration appeared to be the more efficient route of administration as mice injected in this manner exhibited higher levels of expression than those receiving peripheral vein injections. Transduction of liver offers the same advantages as for muscle, but hepatocytes may be more efficient at secretion of protein.

The dosage of recombinant vector or the virus to be administered to an animal in need of such treatment can be determined by the ordinarily skilled clinician based on various parameters such as mode of administration, duration of treatment, the disease state or condition involved, and the like. Typically, recombinant virus of the invention is administered in doses between 10^5 and 10^14 infectious units. The recombinant vectors and virus of the present invention can be prepared in formulations using methods and materials known in the art. Numerous formulations can be found in Remington’s Pharmaceutical Sciences, 15th Edition (1975).

All publications and patents cited herein are expressly incorporated by reference.

Materials and Methods

Construction of rAAV plasmids. The rAAV-AAT vector plasmids used for these experiments are depicted diagrammatically (FIG. 1). Briefly, the plasmid pN2FAT (Garver et al., 1987) plasmid was digested with Xhol to release 1.8-kb fragment containing the human AAT cDNA along with the SV40 promoter and a polyadenylation signal. This fragment was subcloned into a plasmid, pBlueScript (Stratagene) and, after the removal of the SV40 promoter by Hind III digestion and religation, the hAAT cDNA with its polyA signal was released by XbaI and Xhol digestion. This 1.4-kb XbaI-Xhol fragment was then cloned into the pTR-UPS (an AAV-inverted terminal repeat-containing vector) plasmid (Zolotukhin et al., 1996) between the XbaI site 3’ to the CMV promoter and the Xhol site 5’ to the polyoma virus enhancer/HSV thymidine kinase promoter cassette, which drives neo in that construct. This yielded the pAAV-CMV-AAT construct (C-AT). Analogous constructs using the promoter from the small nuclear RNA proteins, U1a and U1b, (to give the A-AT and B-AT constructs, respectively) and human elongation factor 1-alpha (EF1) promoter (to give the E-AT construct) were constructed by substituting each of these promoter cassettes in place of the CMV promoter, between the KpnI and XbaI sites.

The construct, DE-AT derived from E-AT by deletion of the silencer (352 bp) by SAC II-cut (Wakabayashi-Ito et al., 1994). C-AT2 is similar with C-AT except there are SV40 intron and poly (A) sequences flanking the cDNA of hAAT. The p43C-AT was constructed by insertion of hAAT cDNA to an AAV-vector plasmid (pLS), which has CMV promoter, intron and poly (A) sequences. The p43CB-AT is derived by replacement of CMV promoter with CMV enhancer and chicken β-actin promoter sequences. The p43C-AT-IN is derived from p43C-AT by insertion of intron II sequences of hAAT gene to hAAT cDNA (Branly et al., 1995).

Packaging of rAAV vectors. Vectors were packaged using a modification of the method described by Ferrari et al. (1997). Briefly, plasmids containing the AAV rep and cap genes (Li et al., 1997) and the Ad genes (E2a, E4 and VA-RNA) were co-transfected along with the appropriate AAV-AAT vector plasmid into 293 cells grown in Cell Factories (Nunc). Cells were harvested by trypsination and disrupted by freeze-thaw lysis to release vector virions which were then purified by iodixanol gradient ultracentrifugation followed by heparin sepharose affinity column purification. Alternatively, recombinant virus can be prepared according to methods described in Zolotukhin et al. (1999).

Preparations had their physical titers assessed by quantitative competitive PCR and their biological titer assessed by infectious center assay. The presence of wild-type AAV was also assessed using these same assays with appropriate internal AAV probes. The high-dose C-AT stock had a particle-titer of 2.0×10^14 particles/ml and an infectious titer of 5.0×10^12 infectious units (i.u./ml) (particle to i.u. ratio=400:1). The low-dose C-AT measured 8×10^12 particles/ml and 1.2×10^10 i.u./ml (particle to i.u. =667:1). For the E-AT experiments, the titres were 1×10^12 particles/ml and 2.5×10^10 i.u./ml (particle to i.u. =400:1). The low-dose C-AT stock had a wt-like AAV particle titer (i.e., positive AAV genome PCR) equal to 0.1 times the recombinant titer but no detectable infectious wtAAV. The other two preparations had wt-like AAV particle titers <10^-2 times the recombinant titer and no detectable infectious wtAAV.

In vitro transfection and transduction experiments. The C2C12 murine myoblast line was used for in vitro transfection and transduction experiments. Cells were grown in 35-mm wells with approximately 4×10^5 cells per well and transfected with 5 μg of each plasmid DNA using SUPERFECT (Qiagen Corp.). Secretion of hAAT into the medium was assessed at 2 days after transfection using an antigen-capture ELISA assay with standards (Branly et al., 1991). An SV40 promoter luciferase-expression plasmid, pGL2 (Promega), was used as an internal control. For transduction experiments, cells were grown under similar conditions and were transduced with vector at multiplicities of infection ranging from 4×10^6 to 4×10^8 particles per cell. Cells were then passaged in the presence of genetin sulfate (350 μg/ml) and genetin-resistant clones were isolated for hAAT secretion studies.

In vivo injection of AAV-C-AT and AAV-E-AT vectors into murine muscle. Mice strains (C57BL/6, SCID, and Balb/c) were obtained from Jackson Laboratories (Bar Harbor, Me.) and were handled under specific pathogen-free conditions under a protocol approved by the University of Florida Institutional Animal Care and Use Committee. Animals were anesthetized by metaphane inhalation and aliquots of vector were injected percutaneously into the quad-
riceps femoris muscles of both hind limbs. The volume of vector ranged from 50 to 100 μl per injection site and the total amount of virus injected per animal ranged from 5x10⁵ to 1x10⁶ Dnase-resistant particles.

[0061] Antigen capture ELISA assay for hAAT expression. Microtiter plates (Immulon 4, Dynex Technologies, Chantilly, Va) were coated with 100 μl of a 1:200 dilution of goat anti-human AAT (CAPPEL/ICN) in Vollsers buffer (Na₂CO₃=2.76g, NaHCO₃=1.916g, Na₃=0.2g, d₂H₂O=1 liter, Adjust PH=9.6) overnight at 4 °C. After washing, standards and unknown samples containing hAAT were incubated in the plates at 37°C for 1 hour. After blocking in 3% BSA in PBS-Tween 20 at 37°C for 1 hour, a second antibody (1:1000 dilution of rabbit anti-human AAT, Boehringer Mannheim) was reacted with the captured antigen at 37°C for 1 hour. Detection was performed using a third antibody incubation (1:800 dilution of goat anti-rabbit IgG-peroxidase conjugate, 37°C) followed by o-phenylenediamine (OPD, Sigma). Absorbance was determined at the absorbance of 490 nm.

[0062] ELISA assay for anti-hAAT and anti-AAV VP3 antibodies. Wells were coated with antigen (1 μg of hAAT or 100 ng of VP3) at 4°C overnight, blocked with 3% BSA and then reacted with dilutions of either test serum or with positive control antibodies at 37°C for 1 hour. After washing, a goat-anti-mouse IgG-peroxidase conjugate was used as a secondary antibody (1:1500 dilution) to detect bound anti-AAT antibody, using a standard OPD reaction, as described above. Antibody levels were quantitated by comparison with a standard curve generated by reacting dilutions of known positive monoclonal antibodies against VP3 and hAAT.

[0063] Lymphocyte proliferation assays to detect cell-mediated immune responses. Lymphocyte proliferation assays were performed in order to detect T cell responses to the hAAT and VP3 antigens. Freshly isolated splenocytes were grown in primary culture in 96 well plates coated with 0, 0.1, 1, and 10 μg of either hAAT or VP3 in RPMI-1640 medium. On day three, a pulse of ³¹P-thymidine was added, and the cells were harvested on day 4 for lysis and scintillation counting. Phytobhemagglutinin (PHA) was used as a mitogen for positive control wells. A stimulation index was calculated for each antigen dosage level by dividing the counts per minute (cpm) of ³¹P-thymidine incorporated in the antigen-stimulated cells by the cpm in a control (unstimulated) well.

[0064] Following are examples which illustrate procedures for practicing the invention. These examples should not be construed as limiting. All percentages are by weight and all solvent mixture proportions are by volume unless otherwise noted.

EXAMPLE 1

In vitro Studies in Murine C2C12 Myoblasts

[0065] In order to determine the relative strength of a number of constitutively active promoters in the context of AAV-AAT vectors, packageable AAV-AAT expression vectors containing one of the CMV, EF1, U1a or U1b promoters (FIG. 1) were constructed. Each of these constructs were transfected in to the murine C2C12 myoblast cell line. Both the EF1 and the CMV promoter were active for AAT expression, with EF1 construct (AAV-E-AT) expressing 850 ng/10⁵ cells/day and the CMV construct (AAV-C-AT) expressing approximately 670 ng/10⁵ cells/day, as measured by a human-specific ELISA assay for AAT (FIG. 2). This difference was not statistically significant. The levels of expression from the U1a and U1b constructs were undetectable.

[0066] In order to better characterize the level and duration of expression in the setting of vector transduction, cultures of C2C12 cells were transduced with either AAV-E-AT or AAV-C-AT at multiplicities of infection ranging from 4x10⁴ to 4x10⁵ Dnase-resistant particles per cell. Cells were then selected for expression of the neo gene (present in each of the AAV constructs) by growth in G418-containing medium. Several cell clones and pooled cell populations were independently analyzed for AAT expression at four weeks post-transduction (FIG. 3). There was a clear trend toward higher levels of expression at higher multiplicities of infection, and the E-AT construct expressed at least 10-fold greater quantities under all conditions in these long-term cultures. The most active E-AT clone expressed hAAT at a rate of over 1400 ng/10⁵ cells/day.

EXAMPLE 2

In vivo Expression of hAAT from Murine Skeletal Muscle

[0067] In order to determine whether the AAV-AAT constructs would be active in vivo in skeletal muscle, doses of vector were injected into the quadriceps femoris muscle of mice. Circulating serum levels of hAAT were then measured for 11 to 15 weeks after the initial injection. Four saline-injected animals from each mouse strain served as controls. In the case of the C-AT vector (FIG. 5A), levels of expression were sufficient to achieve serum levels in excess of 800 μg/ml in SCID mice after a single injection of 1.4x10¹⁵ particles. A dose-effect relationship was observed, with expression levels in SCID being at least 20-fold lower at the 5x10¹³ particle dose. The levels of expression increased over the first several weeks after injection and were stable thereafter until the time of sacrifice. Since hAAT has a half-life of less than 1 week, this indicated continuous expression. Levels from C57BL/6 mice were comparable, and also achieved values close to the therapeutic range. In similar studies, two of three Balb/c mice injected with 1x10¹³ particles of the C-AT vector did not express hAAT at detectable levels. Both of these were found to have developed high levels of anti-hAAT antibodies.

[0068] Surprisingly, expression levels from the AAV-E-AT vector after in vivo injection were modestly lower than those seen with the C-AT vector (FIG. 5B), with maximal levels of approximately 250 ng/ml at the 5x10¹³ dose and beyond 7 weeks in SCID mice. When the dose was further increased to 1x10¹³ particles, levels of approximately 1200 ng/ml were observed. These levels were stable for one year post-injection (FIG. 5C). Levels observed in SCID and immune competent C57BL/6 mice were similar.

EXAMPLE 3

Immunologic Studies

[0069] In studies in Balb/c mice, antibody levels against hAAT were high in 2 of 3 animals injected. The one which
did not have circulating anti-hAAT was the only animal with levels of hAAT expression similar to those in the C57BL/6 and SCID groups. The high-dose C57-C-AT injection group had detectable levels of antibody directed against VP3, but not hAAT.

[0070] In order to determine whether any cell-mediated immune responses were mounted, lymphocyte proliferation assays were performed using either hAAT or AAV-VIP5 for antigenic stimulation of primary splenic lymphocytes harvested at the time of animal sacrifice, 16 weeks post-vector injection. Using this method, no immune responses were detectable in any of the mice.

EXAMPLE 4

Lack of Toxicity from Direct Vector Injection

[0071] In order to determine whether there was any direct toxicity, inflammation, or neoplastic change associated with vector injection, animals underwent complete necropsies. Histopathologic examination was performed on 5 μm sections taken from the site of vector injection and from a panel of other organs, including the brain, heart, lungs, trachea, pancreas, spleen, liver, kidney, and jejunum. No histologic abnormalities were observed in any of these sites, even among those mice which developed humoral immune responses against hAAT.

EXAMPLE 5

Molecular Evidence of AAV-AAT Vector Persistence

[0072] To confirm the presence of vector DNA, a vector-specific PCR (neu primers 5'-TAATGGGATCGGCCATGAAAC-3' and 5'-CTGTGCCTCCGCTTCGATC-3') was performed on DNA extracted from 3 SCID mice 16 weeks after injection with the C-AT vector, and PCR products were analyzed by Southern blot analysis with a 32P-labeled vector-specific probe (FIG. 8). The state of vector DNA was analyzed using the Hirt procedure (Carter et al., 1983) to separate the low molecular weight episomal DNA from the high molecular weight fraction, which would contain integrated forms and large concatemers. In each case, vector DNA was present in the high molecular weight DNA fraction, whereas in only one of the animals was there a signal in the episomal fraction. This result indicates that by 16 weeks most of the vector DNA in our animals was either integrated or in large concatemers.

EXAMPLE 6

In vivo Expression of hAAT from Murine Liver

[0073] Portal vein or tail vein injections were performed on 18 female C57BL/6 mice 8-10 weeks of age. The injection volume was 100 μl per mouse.

[0074] Each group had the following parameters:

[0076] 2. Group 2: 100 μl of p43CB-AT (3x10^10 IU/animal) n=3.


[0082] A total of 22 animals were used in this study.

[0083] All animals were anesthetized with 2-2.5% tribromoethanol (Avertin) using a working solution of 20 mg/ml at a dosage of 0.5 mg/g IP. A 2 cm ventral midline abdominal incision was made from the pubic symphysis extending cranially to the xyphoid process through the abdominal and muscle layers. The portal vein was exposed by retracting the intestines and associated mesentery to the left side of the animal. Additionally, the quadrate and right medial lobes of the liver were retracted cranially. Intestines and peritoneal cavity were continuously lavaged with 0.9% NaCl. 1

[0084] Virus or PBS was delivered into the portal vein using a 30 g needle attached to a 100 μl capillary pipette using mouth delivery via rubber tubing and a Drummond self-locking double layer 0.8 μm filter. A small piece of Gel-Foam (0.5x0.5 cm) was applied to the injection site before the needle was removed from the portal vein. The needle was retracted from beneath the Gel-Foam and the piece was held in place with forceps while the intestines were replaced into the peritoneal cavity.

[0085] The muscle and skin were closed in one layer using 2 simple interrupted 3-0 nylon sutures on an FS-1 cutting needle. Surgeries were performed on a thermoregulated operating board designed to maintain a temperature of 37 degrees. For recovery from anesthesia, the animals were placed under a heat lamp adjusted to maintain an ambient temperature of approximately 37 degrees and given subcutaneous fluid if there was a significant amount of blood loss during surgery.

[0086] Serum levels of hAAT in the mice were measured two weeks after injection. Serum levels of about 200-150 μg/ml hAAT were detected in mice receiving the p43CB-AT vector (FIG. 13). Studies using the E-AT vector show that injection of vector by portal vein led to greater levels of hAAT secretion as compared to E-AT administered by tail vein injection.

EXAMPLE 7

In vivo Expression of hAAT from Murine Lung

[0087] Mice were injected intratracheally with either C-AT or p43CB-AT vector. Serum levels of hAAT in the mice were measured at day 3, 14 and 31 after injection (FIG. 14). The p43CB-AT vector mediated high levels of expression of hAAT in lung.

[0088] It should be understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and the scope of the appended claims.
**REFERENCES**

[0089] U.S. Pat. No. 5,399,346

[0090] U.S. Pat. No. 5,439,824

[0091] U.S. Pat. No. 5,658,785

[0092] U.S. Pat. No. 5,858,351


---

**SEQUENCE LISTING**

<num>160</num> NUMBER OF SEQ ID NOS: 13

<num>210</num> SEQ ID NO 1

<num>211</num> LENGTH: 6565

<num>212</num> TYPE: DNA

<num>213</num> ORGANISM: Artificial Sequence

<num>220</num> FEATURE:

<num>223</num> OTHER INFORMATION: Description of Artificial Sequence: PLASMID C-AT

<num>400</num> SEQUENCE: 1

```plaintext
gggggggggg ggggggggtt ggcacatccc ttcotgctgct otggtotgcct cacgaggggc 60
gacgcagcaca aagggccccc aacgacgggg tttgcgcgag ggcctctcgt gagacagggc 120
ggcgcaggg aggaggctgg ccactcctca actagggggtt cctagarcg ttagtgcttc 180
gctacaata aotacgctta acatggcgcc otggtotgcac gcacaaagc ccccggccat 240
tggacgycaat aatgyagctat gggcggcctag tggacgccag aagggctttg cggcaggytc 300
```
aattgggctag ctaactacgtg taccactgac actgacggt caacctactg taccactgac 360
cacactgac gctacgttct acctgtacgc gccctgtgtct tcctgtacgc 420
catccactg tgggttgggc actggtctgc tccgctgtct tccgctgtct 480
catccactg tgggttgggc actggtctgc tccgctgtct tccgctgtct 540
gatttcacag cttccacacag agggaggtgt caactgcaat gttgccaccc 600
gggaggtgt caactgcaat gttgccaccc 660
tccgggttg gactgtttta aagcagaaggt gtttagaca gcggcagga 720
gccagacag cggcagcag cggcagcag cggcagcag cggcagcag 780
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 840
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 900
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 960
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 1020
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 1080
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 1140
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 1200
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 1260
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 1320
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 1380
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 1440
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 1500
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 1560
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 1620
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 1680
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 1740
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 1800
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 1860
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 1920
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 1980
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 2040
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 2100
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 2160
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 2220
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 2280
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 2340
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 2400
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 2460
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 2520
tggccagcag cggcagcag cggcagcag cggcagcag cggcagcag 2580
ggctatgact ggccacaaca gaacaatcggc tgctctgatg cccgctgttt ccggctgtca 2640
ggcaaggggc ggcggctttct ttttgtcaag aacgacccgtg cccgctgcct gatgaacttg 2700
cagggcagg cagcgcgttc atcgtggtcg gcgcagcaac ccgcgctcctg cgcaagctgtg 2760
cctcacgcttg tcacagggac ggggagggac tggcgcatac tggcgcaggt gcgaggggacg 2820
gatctccctct cacctcaccc tgcctctgcc gagaagatac ccgcctaggg tctggtacaag 2880
cgccgggctgc atacggattga tgcgctaccc tgcgcaatagg gacaccaagc gacacatgca 2940
atcagagcg cacgtactcg gattggagcc ggtctttgtcg atcagagatga tctggtacaag 3000
gacagcaggg gtcggcggcgc gacgcaacttg ttgcgcgggc tcagagcgcc gctgccggcg 3060
ggcgggagtgc tctcggtgcc ccaagtgccg gctgcttgcc gcaaatctcc gggggaatag 3120
ggcggctttttc ctgacgcttcc ggctgcggcg cgcggctgtcg tcggcgcggc ctcagcggcc 3180
atagcggagc ggcggtggtg ctaacgcctga ttatgtgctga gacgttgccg gcgaagaggc tgcacgttctc 3240
cctgctccct ggctgcgcggg ctcgagctcc tggcgaacgg gatcagcgaac gctgctttctc 3300
gagcaggttc tctcgcgggg tccgctgctc agagctcgtc gtcagcgcctc gtcagcgcctc 3360
totgtgcg cgcactcgcc tgtttgcccc tcccgccgct cccttgctgc cctcggaaagt 3420
gccctctcc gcgcgctttcc ctaataaatac gaggaatcttg ctcgcagcttg ctcgagttcg 3480
tgtccattata tctctggggcg tcgctggcgg gacgcaacgc cgggggaga cgggggagaa 3540
aatcagcaggg gtcgcagcttg gacgcaacgc gacggcgttc ggcggcggcc gaaacgcgcttg 3600
cctcgccgct gcgtcgcgctcc agtcggcgcg cccgggagaa gcgcggcggc gacgcaacgc 3660
ttggtgcgc gcgcgctgtcc gcgcggcggc gcgcggcggc gcgcggcggc gcgcggcggc 3720
cccccccc ccggggcgcg gaggctgggg gacgcaacgc gacggcagc gggggaagta 3780
gtctggcg gcgcgctttcc ctcgagcttc gtctgcttca gcgcggcggc gcgcggcggc 3840
ggcgggagtgc atagcgcctcc tcaagcgccg taatacgggta atcagcgcagc tcgctgggata 3900
gacgagggac gacgtcgcgg gacgcaacgc gcggcgggac gcgcggcggc gacgcaacgc 3960
gtggtgcgc gcgcgctttcc ggcgcacgcc gcgcgcaagc gcgcgcaagc gcgcgcaagc 4020
caatcagcag gcggggcgcg gcgcgcaagc ctaacgcagc gcgcgcaagc gcgcgcaagc 4080
agcctcgcgc gcgcgctttcc gcgcgcaagc gcgcgcaagc gcgcgcaagc gcgcgcaagc 4140
tctctgggg cagctggcgct ctctcctctc gcgcgctttcc gcgcgctttcc gcgcgctttcc 4200
agcttcgct gcgcgctttcc gcgcgctttcc gcgcgctttcc gcgcgctttcc gcgcgctttcc 4260
cattggtgg gcgcgctttcc gcgcgctttcc gcgcgctttcc gcgcgctttcc gcgcgctttcc 4320
cacgagggac gacgtcgcgg gacgcaacgc gcggcgggac gcgcgcaagc gcgcgcaagc 4380
tgagcgtggc gcgcgctttcc gcgcgctttcc gcgcgctttcc gcgcgctttcc gcgcgctttcc 4440
tcggcgcagc gcgcgctttcc gcgcgctttcc gcgcgctttcc gcgcgctttcc gcgcgctttcc 4500
tctgctggtcg ggcgcgctttcc gcgcgctttcc gcgcgctttcc gcgcgctttcc gcgcgctttcc 4560
aagaagcgcgg cggcgcgctttcc gcgcgctttcc gcgcgctttcc gcgcgctttcc gcgcgctttcc 4620
aaggggaggct ggcgcgctttcc gcgcgctttcc gcgcgctttcc gcgcgctttcc gcgcgctttcc 4680
aatagggagc gcgcgctttcc gcgcgctttcc gcgcgctttcc gcgcgctttcc gcgcgctttcc 4740
gcagcgcac ccgcgctttcc gcgcgctttcc gcgcgctttcc gcgcgctttcc gcgcgctttcc 4800
gagggagggct ggcgcgctttcc gcgcgctttcc gcgcgctttcc gcgcgctttcc gcgcgctttcc 4860
Continued

catgtacct gctgcaaccc gctgcaccgg ctccgagatt atacgcaata aaccgcgcgc 4920
cggagaggg cgagagaaga agtggctctg caaacatat gcggctcact cagctctatta 4990
agttgctgcc ggaagctaga gtaagagatt gcgagcctaa tagtggctgc aagcttgtag 5040
ccatgtctac acggctagtg gcgctcagct cgctttgtag ttagcctca ttagcctcgc 5100
gtcccccaag atcaacgggca gttacatgat ccccccagtgt ggcaaaaaa gcgcttgagt 5160
cctctggcct tccgagatt tgcgaaagta agtgagccgg aggctttatca tccatggtta 5220
tgccgagat ctatatactt ctatactgca tgcncatcgg aatagctttt tctgytaactg 5280
ctggaatctt aaccgaactgg ttctgagaa attgatacgc ggcagcggct gtttctgtgc 5340
cgagcttc agcggtatgc acggccgac atagcgaac tttaaatgtg tgtctatccttg 5400
gaaaaaggt cctgcgacgc aactgtattt cagcaatttt tacttccca aaggtcttccg 5460
gttgcaagaa aacaggaaag caaattgcgc ccaaaagaag ttaaaggggc acacaccaat 5520
gttgatatt cacttcttc tttttttcct atttatgaa gatttatcgc ggctatatgtc 5590
tcgagaggg atacaaccttt gaaatatttt aaaaatcaaa caaataagg gttggcagca 5640
cattccctgg aaaaagccag ctgctagctt aagaaaatct tattataagt accaaaactt 5700
ataraaattag gctatctgac agggctccttc gttcagcgag ttgtggtagtt gaaagtaaaa 5760
acccctgaga cttggaactc cggcagacgg ttcacccttg tccgttaagcg cgtgagggga 5820
gcaagaacgc cctgcggggc gcgcgagcgc gattgagggc gttgaggccc cttggctaac 5880
tgctggcact gacagctgacgc attcacgagct gaaaaagct cattttcaaa aaccagaaaa 5940
gatctggag gaaacaataa cgcgctcgaa gattgtataaa gttatatattt ttttacaaat 6000
cgctttaata attttttatat ttactttttc tctcggcgc ttcgaacatt 6060
cocttatata ccacagtata agcgagat agggctgagt gttgctccgg tttgaacaca 6120
gactgcaca ttaaagccgt gcgcgacaag ccaaaaaac cctCACagg cttatagggc 6180
gagcgccca cttatactc ggcgagaagc cagaaaaac ggcagcgcgc tccacctaga 6240
agactcataa cagacgacgag cccagcagcc cccatgatgc gcgctgaaac cggagcctc 6300
gagacgagc gaaacoaacg gagaacagtc ggcagcggc gcggctgatgc ggcggcgaga 6360
tgcagcccgt cagctggcgg taaccacac acgcgcccac attcgccgggc cttagcaggg 6420
gcggcagcgt ctcagccgtg ctcgaactcg caactctgag gcgggctgc gctgtggggc 6480
cattgggct cttacagctc agttgcctgg cccagcaccg cttgttaagct cgtgaagtc 6540
ctcttgctta ttcggctcgtt gcgc 6605

<210> SEQ ID NO 2
<211> LENGTH: 7495
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: PLASMID K-AT
<400> SEQUENCE: 2
9999999999 999999ggtt ggcacatcc ctcctggcgc ctcgctagct ctcagtggcgc 60
gggagcaaaaa aggcggcggc aagccggggt tttcgccggc gcggcctctg gcggcagc 120
gcgacagag aggaggcgcc caacccctag actggggttt ccagagctcg aacccgctac 180
ttcgagccg ccaggccccag agaattcgcag gcgcctctcc cggggcctcc 240
-continued

cocctcaca cccoccccaca ccccccgccg cggaagctga ggaatattca tcaaaacagga 300
ctgccccagt ccttggggag taacccggac cctgttttaa ccttcccaca cctgtcacc 360
agagattcgt gccttcaccc ccccccccag gcccccttcg gctatcactg agtgagagaa 420
gagccctcg ggccttcccg tgtgcccttct gttgccagac gcacactgc ccacagtcccc 490
gagaagtggg gcggcaggttg ccgcaattga accggtcgct acagagagtt gcgcggggtg 540
aacgagggac ggtgcacgtc gtaatggcctc tgcctttttt cccagccgtg acagccatcc 600
tataataagc cagagactgc cgtgacgctt ctttttcgca acggggttgg cgcccggaga 660
cagagatcgg cccctgctgt ggccccggag ctcggtcctt ttacccggctta ggctccttgg 720
gtctcccgga ttatcctcag gcccttgccgt gcagttacgtg attctgtgcac cccagctcctg 780
gtgccgaga gggtgagagc gttcgcagcc ttgatcttaa ggaaccctgtt ccccgttctg 840
ttgatactgg gaccccgcttg ggccttgggg cccggcggtgt ccaacagtgg acggatctcg 900
gccgtcgag gtttcgagtt ataatgctctg acgacatcatt aaaaaagagc gctacgtctg 960
gagcccttttt ttgctcaag atatgccctgt aatgcggggc caagactgtgc acacagctgt 1020
ttcgctttttttt gggcggccag ggcggagtgg gcggccttcg tccccacgca cattgtcgcc 1080
gagcggcggc ctcgagcgcc gcccccagag cattgcggag ggtccctcctc aacgctcccg 1140
gccgtcgag gtttcgagtt ataatgctctg acgacatcatt aaaaaagagc gctacgtctg 1200
ggccccctgc ccctgcgtcc gtcgacgcgg caagtgcccg ccacccggag cccggtcgc 1260
gagctccaccc tggggcggag gcggccttcg aggagactgc ccaccacacaag 1320
gaaaaagaattt tgccttccct cccagctgcc tcctgtcgcac ttcgagtgctg accgagctgc 1380
gtcaacgcac ctcgtttact ctgctcctgt ttggaattgt tctgttcag tagttgggata 1440
gggtgtttctt cggagggagg tccccacac cggagagtctct gcagttgctc ggtgcacgag 1500
ttgccatctg atgtaatct tccgattgaa tggccttttg gagttgggtt cttgctttct 1560
tctacacgt ctgcaaggttc ttctcccgtct ttcagctgtt ctctgtcctctt ccacatcctg 1620
tagaactagt ggtcaccgcc ggtcagcaag attgtatatg aagctgtgggc attttcaaggg 1680
aacacacagt acctgctgcc gttgatcagc aatcgcgact tcctgtctcg ggggcatcct 1740
cctgatggga ggcgtggtcct gcctgccttc gttggaggtg gttcgagata ccaccgaga 1800
tgctgcgcag agacagata catccccaca tgcggcttgc caaacacct tcacacagat 1860
cacccacaccc ctcgctgact cgcctctcat cgcctactg accgctgcga aacaagctgg 1920
cagccacact atttctctct cccagcttctg ctcggctaca gtcgctactg ggctttcctc 1980
gggactgcag ggaacgcagc atcgatcagct cggwgacggtg ctaaatcctc acaccacgcg 2040
gattcccag gctccagtac atgaaggctt cccaggacgac ctcggcacc aacccggac 2100
agacgagctag cttccagctcg ccacgctccag tggctgcttc ttggacggag ggtgcagagt 2160
agttccatag ttggtcggag aatgtaaaaa ggtgctctcc tcaagacgtc tcctgacaa 2220
cctggcgagc agcagacgag ccaagacaca gataacagct taogttgagaga agggcactca 2280
aggggaatct gttttcttgg tcctggagct tcacagagac acatgttctg ttcgggtagaa 2340
ttacatacct ttttaaaagcc aatggagagac agcctcttgaa gtaagcagca gcagggaga 2400
ggctccaccc gtggcaccagc ttacacagct gcagtgatcg gttggaggtg gttggaggtg 2460
gttttcata ctcgacgtga aagagctgtg acagttggttg cttgctagat aataccctggg 2520
caacacgcc gcacattcct tcctgcctga tgggggaaac ctacgacacc tggacacagtga
2580
actaaccgcc gataactcata caagattctc ggaaatggaan gacagaaggt ctgcaagct
2640
acatatataac anactgtcag ttactggaacc ctatctgctg aagacgcttc tggatcaacct
2700
gggcatcact aggacctcaca gcaattgggc tgaacctccgg ggggctaccc aagaggccacc
2760
cotgaagcct tcaaaagcggc tggcatagcc tgggtgaccc atcggcaggaa caggggctga
2820
agctgctgag gcacattttt tcagggccat atcatgtgtct atcccccccaag actgcaagtt
2880
cacaacacct tttgcttcct taatgttgcga acaaaatacc aagtctccccc ctctcattggg
2940
aataagctgt aatoccaacc ccaaaaatacct gcttctcggt ctocaccccc tttcccttcat
3000
cttggcgcaca cttgctgtgat gcacaaaag aaggttgtag cgtgttaacc ccccccocccc
3060
tgaaaggggcc ccgcatagct tgggtttggc aaggggaagg aaaaaaccttc tccaccccag
3120
cotggagatgt tccaaacccaa cttcggagcc gttgtgatitt gcaagagggc aaaaaaaggc
3180
ttcccccagtt tttcctaaacc aatgcggag aacccggcgc aaggtctgctg 3240
cttggcggas ttgggaccgg cagatcgctg cggggcggg cgggctcccag tegcttcctc
3300
atatgtagg gcacaggttg gacctggacca ccgacaggcc cttcggcaca tttggtgctc
3360
gotcctggag aagatggatt gtcctgcgag cctgggggtg gcgtttgacjaa gggcttatc
3420
ggcagatcgacctggcagctg_tgtcgtggc cgggctgtgg ccggcgtttg 3480
gcgccgagc gcgccttcct tttgcttcag acagcagctgt ccggtggctct gttgatcaac
3540
caggagggc ctaggggtcg atcgctgctg gctggggag cgtttccttg cggagcggtg
3600
tcggagcttg cagcttgagt cagctggag caggtctcctg ccgagctgtg 3660
gacatcctg cactggtaca tgggatgaca gaggagattt cactcggtgg ctaggaatg
3720
cgggctgctg atcactggtc tccggtaccct tggcctacct gcacccaaag gcaactgctc
3780
atcaggagc caagactcgc gatggaagcc ggtcttgtag atcagagattg tgggaagcgc
3840
gaccagcgg gctgcgctga cagccaaagc ctcggctggg ctcaggggag catcggcagc
3900
ggagggagc tcgcgtggga cgcagtggct gtcgttgccgc gtaataataat gcgggaaat
3960
ggaggttttt cttgctttct gcaagctttgc ccggctggtg tgggctaaccc ctacaggaac
4020
ataaggtgt gtaaccgctga tatttgccaa ggggtttcgg gcgagatggc gcaacagctt
4080
tcgcttttt clcaggtgctc cgcggctgctc tgcgagcgc gcctgcccttc cggctaaattt
4140
gcagagtttc tcggggtgga tcctgtgact gcagctgcag gtagcctgcc gatcctttggc
4200
tatatctggc acgcatctgt tttggccccccc ttcggctgcgc cttcctgtgac ccggaggtg
4260
gccatcctgc cgtgccttct ctaaatatat gcggaaatct catcggcttg tttggctagg
4320
tgctcttttt tttgggggg gggggagtgg cggccgagca gggggaggag tgggagaaagc
4380
atatagcggc atcggttgagga gagctctagg aaccctagtq gatgcttgag gcacactctt
4440
tctgcggcgg tcgcgctgctg atgcggagcc ccccgggaaa cccggggggt cggggaactt
4500
tgggctggcc gcgcagctgtg aagagggagt gcggagaggt ggggaaggtg aaccggggggc
4560
cccggggccc gcagcgcggga ctcataaataa tgggacnncc cccgggaga gggtttttgg
4620
gttggtggct ccattgtgctg ttcggtcctg accaggagtt gttcgggttg gggcttgtgg
4680
ggggacggtt atcgcgcctac tcaaaagcgcc tattacggtt atocacagtaga cccgggyna
4740
gaacoaggaca gaacacttgtga gcbaaancc aagccaaaagc tagggagcct ccccccagggc
4800
-continued

cggtcgagc gtttttcct cagctcggc cccctgaagc gcattcaca aacctagcgtc 4860
caaaagacg gtaggcgaacbcc cagcagagac tataaagata aacgggttct ccccttagga 4920
gctcctctgt gcgcccttct gttcgacccc tgcgccttac gcgatacctg ttgcctcttc 4980
tccttgcag aacggttcgg cttttctcaat gtcacgcttg tagttatctc aacgagctgt 5040
gcgcgttcag ctccataagc gcgcggagct gcagaccccc cgttccagccc gacgcctgac 5100
cattactcgg taccatatgt cttggagcca aacgggaag aacgaatctt cttgcctcag 5160
cagcagccac tggtaacagag attacacagag gcgggtatgt aggcgggtgcct acagagttct 5220
tgaagttggct gcgaattactc gcgaacacta gaggacagt atttggtatc tcggctctgtc 5280
tgcacgaacg ctaaccagga aacaggttgt gcgcgtccttg acctggcaaca caaaaccocyng 5340
tcgcgttcagc gttttgcagc agcaggtacg gcgcgaaana aacagacagct 5400
aacagagtagc tttgcctttt tgtactgggt ctggaccctg atggacgctaca aacagagctt 5460
aggaatttct gctctacgac tttcattaaaa gcctgtccac ctagctcttc taaaattttata 5520
agtaaggttt ttaaacatcct ctaactatgt atggtaacac tggcttgcgc agttacact 5580
gcctactcag tggccagctg atctgagcga tttgcctttct gtggatcacta atagtgcctt 5640
gactctcctg cgtctagata acatcctagac ggcagggctc acatcctagc cccagttgtg 5700
caatcctac gcggacaacc gctcagagcc ttcagatatt atgcacacta aacagggagc 5760
cgcagacgcc aacggtctcttg cncacattag cgcctccatat cgcgctcttt ca 5820
attgggtcgg ggaacgtctg ctaagaggttt ctgggcactgg aacggttggc 5890
coattcctag acacagccag tgcgacgcgt gcggcccttg tgcggctctca ttccagctcg 5940
gtcttcgagc acaataagggc gttccatgat cccccagtgt gcgcagaaaa gcggttagct 6000
cctcctggtgt gcgactagta aacgccggcc aagttctatca ctaggtagta 6060
tgccctctgt ccgactaccc gcgcgactgt gcgcgactgt cctggtgaatgtcctg 6120
gctcactcag acacagggat cttcctgact gcggcagct gcgcggagct gcgcgcgtgc 6180
ccoctrcaacct cgaagcgacac ctaccagatgc ctggccacttt cttggctttccttctgcctg 6240
gaacctctcggc ggcttttttct gcgtctcaac gccgtcttct gcgtcttgcg ctaatgctca 6300
tgtaacccagcgctgctcacc aacatctatgt ctagatcct tttttcaatct gggagcttg 6360
ggtatgttcgg cagcactggg ccaaggtggc ccaaaaagg cctataaggg gcacgagaaat 6420
gttgtcttctt ccattttcaat atatttaag catttattcgg gttatttggct 6480
tactaggtgg atacataattt gacttgattg agaanannt acaaataggg gcgcgcgcgc 6540
ccatcctccagc ccataccgct cggattctt aacacaaaatt aacattacat ccatcctctt 6600
ataanataag gcgctacgcc agggccctttc gttcgcgtgg ctgggcgggt gccgctttcct 6660
cacgtctgagccc taacgcggcc cggcgtctgtgc tgcgtctgttg gctccggaa 6720
gcagcggag cgctctcgggcc gcgcgtcggg gctgcgctggc tgtggtaact 6780
atacgccagg aacagcagacc ttcgctcctgc gtagcgtatg ctggtagcgg ctggccact 6840
gactgagtgc gganntactc ccctcggctgg ctggctcttt gtaggtttcctt tttttcaca 6900
cgcgtaattt tttgttaaa tcaagctctt tttaacccaa tagcgccgaa gttacaaaat 6960
cocttataa tcaanaagact acagcagatg aggcttgagc gttgctccag tttgcaaaa 7020
gctgccacta ttaaagacgcc cggactccgag ctggcagggtt cggaaaaaccc ctgatatcgg 7080
cgtggcogca ctacgtaaac cttacoccta acnaggtttt ttgggtcoga ggtgcogtna 7140
agcactaat cggactctca aagggagccc cagtttaga gttgacoggg gagaogogc 7200
gacagtgcg agaagcagag ggaagaaaac gagaagagc ggcacttagg cgtggcogga 7260
tgtagcggcg acgcgtgccc taaaccacac acgcctcgcc ctaaagtcgc cgtataaggg 7320
cgcgctgcc ctagcgcctc tccgctccag cactcgttgg gagaagcagt cgtggcoggc 7380
cctctctgca ttacggcagg ctcgca 7405
<210> SEQ ID NO 3
<211> LENGTH: 7054
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE: HLA-A
<223> OTHER INFORMATION: Description of Artificial Sequence:PLASMID
<400> SEQUENCE: 3

gggggggggg gggggggggt ggcgaatcct ctgccgcgct ctgcgtcggt cactgggycc 60
ggcgcacca agcgccgcc acgcgcggg ccgtcccgct ggcgctccag gacgccgnga 120
ggcgagcgag aggaggtgagg ccactctcct aactgggytt gctatgcgtg aactctggtac 180
cctcgagctc agccactcct ggtgaggaga agatttgcga ccgtccctcc aggcgcgctc 240
cocgtaacc cccccccccc ccgtccgccg cgcgactgca gataattccc taaaanagga 300
ctcgccctgc ccgtggggcg cgtccgtaaa cctccactca cctcacaacc 360
agagacgct gcgttccggc ccctccaccc gcgcgtcctc gtaatcaactg aggtggagaa 420
gagcagctg cgggtctgctg tggcctcgcag tgggctacgc gcacatcgcc caacaattcc 480
gaagaagggc gggagggggg ggcgaatttg aacgcggctg agaagaggtgg ggcggaaggt 540
aaaccggaa gtagttgctg gacttgctgt cgcctttttc cggaggggtgg gggagacacc 600
tataaaagc ctagtggtgg cgtgaacgct cttttttttcg gcgggtttgg gcggacaaccac 660
caggaagttc cctgtggtgg ttcgccgggg ggccggaggg ggcggcggtg ccggacccac 720
agtctcgccg aggaggcccc cgcaagccag atcgacgctg gcagagaagg cggtagttcga 780
agctgacggg cctgtctggt tgcgcttgggtgc gcgacgcgctg gtagtcgctgcg 840
gcgaaaggt gcgcgctcct gcacagtgcgc gtaggcggag cgggcacgcct 900
tgctgcaggg agtctcataa gggaggcgg gcggcgtggag cggcggggcg ggtagtcacc 960
ccaaaaccgg aaaaaggccc cttccgcctc cgcctcgctg ccagttggct ccggaggtaa 1020
cgggccgct tccagcctgg ccagatgtgc atccgcttcc ggcggtttgt 1080
tggtgggggg ggtttttgtg cggaggatt gcctcaccgt agcggtggag gtagcgtgcc 1140
tagcgcagtc tggcccttgg cgcgctttttcg gttccgggac 1200
tggttcatt cgcctcgcgt aacgatgtgc atccagttcc ttcctctccc ttcagttgct 1260
gtaaaccctc agatcgtttg gcgtcccgcc gcggagggg ggccagcgcgc tgtatcaactgcgttggg 1320
ccggcccag ccaaacgatc cggggagac gtagcgttct cttgcctgctg 1380
ggcgatcctc cttgcctgtg gcgtctgtcct ctgccttcgc gtaggcgttcc 1440
ccggagagt gcgcctccag agacgattc cttcgccctcg gcgcgtttcc gcgcgggacc 1500
ccaaacatc cccgccacac ccgtgcacgt gcggcgccttc ggcgcttccgc gcgcgggacc 1560
ccgggggtc gcgcctcctc ccgcgggctgc gcgcgtttcc gcgcgggacc 1620
-continued

gcctgccgt ggagccaaag ctcgacacta cgtgaaacct cttgagggcc tgaattc tga 1680
gcctgagc attcgagc ctcgacacta ctgaagcctt ccagacatcc tgcgcctttc 1740
cacgcaag cagcggcgaact gcaagcaaat ggcggtctgc ctcgacaggg 1800
cgtgacag ttgaggc aactgtggga ttttgagga tcgtaacaag tcgggttctttc 1960
cactgtcac atcgagggac ccagacagc caaagaaacag acatagcatt agctttgga 1920
ggtaactac gggaaaaattc tgcgattttgtg caagagcgtta gcaagagaca caagtttttc 1980
tcgctatcat tacatcttcct ttaagccacaa atggagagacttccctgaag tcgaagacac 2040
cgcgagcgag gtcctgccag tgtgcagccttg gatgccaggtt agatgctca ggctgac 2100
cttgaggaatt tctgaacctaa agagctgacg cgcctgggtcg ctggctaacaa 2160
atacttgccc atctggcaac agcagcgccg cattttgctt cttgaagcgatt gcagagcc 2220
ggaaagcgacc ctcccaccag taatctctgg ccaaatgagc tcaagagacagc 2280
tgccccgcttt caattccag cacagcttag ctcctctggc aagagctgctc 2340
ggcttacttc gggctgctact ggcctttctc ggagcctgcc gcttttttgagc ctcgacttac 2400

ggagccac ccagcttcct ccctcaagctgt gctgaaacct cttgaggggct gccttctgamgc 2460
ggctacgaact gccggagtgc cgggttccccg agagcactca ccagttctca tcccctccag 2520
ggctgacagc aacaacccct ttcctctctt cattttgga caaaatacaag cctctcctcttg ctgctctcc 2580
ctctctctct cactggagggg cttttttctcc aaactagtc tctctcctcttg ctgctctcc 2640
ccctctctct cctgctgccc tccctctctct cattttgga caaaatacaag cctctcctcttg ctgctctcc 2700
cccctctct cctgctgccc tccctctctct cattttgga caaaatacaag cctctcctcttg ctgctctcc 2760
ccccctct cctgctgccc tccctctctct cattttgga caaaatacaag cctctcctcttg ctgctctcc 2820
ccacgacact ctcgacacta ggcctggggc tgtttttttc cagcgaagc ctcgacacta 2880
ggctcttcgg atgctgctggt ctgaanagcg tagtgaaggt gcggctgcgctg gtcggagctgt 2940
cacgtgcaac tccaataact ccagctgggtg ctcgaacac gcagacgaccc tgcagcacta 3000
atgctgccttc ctttctctct gctctggagag cagagcaggt cggctgctctct ctttctctct 3060
agctcggac gcctgctcttc ctttctctct gctctggagag cagagcaggt cggctgctctct 3120
cgctggaggct ccaggcttgct ccgagccagct ctgctgctcttc cggctgctctct 3180
atactgcag ccgctgctcttc ctttctctct gctctggagag cagagcaggt cggctgctctct 3240
gcagctggctc tctctgctgtt ccagacagc ggcggtcttc gctgctgctcttc cggctgctctct 3300
cgctggaggct ccaggcttgct ccgagccagct ctgctgctcttc cggctgctctct 3360
gctctggagag cagagcaggt cggctgctcttc ctttctctct gctctggagag cagagcaggt cggctgctctct 3420
acacgacact ctcgacacta ggcctggggc tgtttttttc cagcgaagc ctcgacacta 3480
ctgctgctcttc ctttctctct gctctggagag cagagcaggt cggctgctcttc cggctgctctct 3540
atgctgagc gcggctgctcttc ctttctctct gctctggagag cagagcaggt cggctgctcttc 3600
atagctgagc gcggctgctcttc ctttctctct gctctggagag cagagcaggt cggctgctcttc 3660
taagcagcact ctcgacacta ggcctggggc tgtttttttc cagcgaagc ctcgacacta 3720
ccctctctct cctgctgccc tccctctctct cattttgga caaaatacaag cctctcctcttg ctgctctcc 3780
ccctctctct cctgctgccc tccctctctct cattttgga caaaatacaag cctctcctcttg ctgctctcc 3840
ctctctctct cctgctgccc tccctctctct cattttgga caaaatacaag cctctcctcttg ctgctctcc 3900
ctggaaaggt ccaactccac tgccttttcc taataaaaatg agaasaattgc atcgcatattg 3960
cagatactgt gctcatctat cttggggggt gggtgtggggc aggaacagcna gggggagagt 4020
tggaaacata ataagcaggca tgcctggggag agatcctaga aacctccttg atggagttgg 4080
caactccctc ttcgctgcct gcgctgctca ctgagggcgc gcggcgaang cccggggcgc 4140
ggggacccct tgcctgggag gcctcagttga gcggagcgac gcgcaagagag gagggtggc 4200
acaccccccacc ccccccccct gcagcggcgcc attatgaatt gcgccaagcc gcggggagag 4260
gggctcttgc tgggggagcc ttcctcgttt attcgccact ggacggcctg cgcgcgct 4320
ttcggctgcg gcgaggctga tcagctcact caagngcctg acatacgctta tcacagctgg 4380
caggggagc gcagcagacgc aaacagcgtg ccaagcagcgc gcgaangcgc gcggagcag 4440
acacggcggc gttccgctcgg ttttcccata gcgcgccccg cccctgcgcg cctcacaaga 4500
atcgactgtc aagctttaggc tggcaaaacc gcggcaagact ataagctata cccgggctt 4560
coccttggag ctcctctcttg cgcgcaacct gcggctttg acggctttcgc gtggcacttc 4620
coccttctc ctccctggag gcggctttgc ttcctcattc tccacccttg atgcactctc 4680
acggctttg aacgctggcc tccaagcctg gcctgctgta cgaacccccg gcggccgccc 4740
acgcttgccc ctttacccgt aacatccttc ttgacgcaaa cccggttaaga cagcgactac 4800
cgacaacttg ccgcagacct gcgtaacgag ttcagacag ggctatgagt gcgcgcgcctg 4860
caggtcgtc gcggctggag cccactcag gcgtaacgag ggctatgagt gcgcgcgcctg 4920
ggcgctgcgt gacagcaggtt aacgagctga ttgctttggt gctctcggac 4980
acacggcggc gccgctggag cccactcag gcgtaacgag ggctatgagt gcgcgcgcctg 5040
ccgcagacct gcggctggag cccactcag gcgtaacgag ggctatgagt gcgcgcgcctg 5100
acgcagcgc aagcccttgc gctctcggac ttcagacag gcggccgccc gcggccgccc 5160
acgcagcgc aagcccttgc gctctcggac ttcagacag gcggccgccc gcggccgccc 5220
taaatctaa cgggtattgg gctgcggtatr tacataaaag gcttcccacc tagtaccttt 5280
taaatctaa cgggtattgg gctgcggtatr tacataaaag gcttcccacc tagtaccttt 5340
acgcagcgc aagcccttgc gctctcggac ttcagacag gcggccgccc gcggccgccc 5400
acgcagcgc aagcccttgc gctctcggac ttcagacag gcggccgccc gcggccgccc 5460
acgcagcgc aagcccttgc gctctcggac ttcagacag gcggccgccc gcggccgccc 5520
acgcagcgc aagcccttgc gctctcggac ttcagacag gcggccgccc gcggccgccc 5580
acgcagcgc aagcccttgc gctctcggac ttcagacag gcggccgccc gcggccgccc 5640
acgcagcgc aagcccttgc gctctcggac ttcagacag gcggccgccc gcggccgccc 5700
acgcagcgc aagcccttgc gctctcggac ttcagacag gcggccgccc gcggccgccc 5760
acgcagcgc aagcccttgc gctctcggac ttcagacag gcggccgccc gcggccgccc 5820
acgcagcgc aagcccttgc gctctcggac ttcagacag gcggccgccc gcggccgccc 5880
acgcagcgc aagcccttgc gctctcggac ttcagacag gcggccgccc gcggccgccc 5940
acgcagcgc aagcccttgc gctctcggac ttcagacag gcggccgccc gcggccgccc 6000
acgcagcgc aagcccttgc gctctcggac ttcagacag gcggccgccc gcggccgccc 6060
acgcagcgc aagcccttgc gctctcggac ttcagacag gcggccgccc gcggccgccc 6120
acgcagcgc aagcccttgc gctctcggac ttcagacag gcggccgccc gcggccgccc 6180
ttcgcgcac attccccccg aagaattcgc acctgctctta agaaccacc ttaataatga 6240
cattaccct aaaaaatagg cgtatcgcgc ggccctcttcc tttcgcgcgt ttgggtgtgc 6300
aggtgacaa acctgctcct acgtcgcgcc ccggagccct cccagctctg cttaacgccg 6360
attgcggag gcacagcgcc gcgcgcggc gcgcgcgac tgtgcgggcggt tgtgcggggtc 6420
ggttacatc tcggcgacsta gaggagcttg tactgtggct gcagcactat agggtgtgaa 6480
taccgcagag atcgatgag agaanaactcg gcacagacacatt gttgaaacag ttaaatatat 6540
gttatatcc gcgttaacat ttttgtatgt cagctatcgg tttataactag agggcagact 6600
cggcacaacc cccattataa ccacagacta gacccgatat ggttggagct tgtgatcactg 6660
tttgaacag atcgggatat ctaaaactgc ggcgtgatcg caggtggaat ctttctctgt 6720
catatgacg ggtgcccacc taacgtgacc attacocctaa toaggttttttt tgggggttagc 6780
gttcgcctac gcgaatgact gacacccata aagggaccac cgcaggtcgc cttgacgggg 6840
aagcgcgac acgctggcga gaaagagag ggaacacac aggagcagcgc cgcgtacggc 6900
gtgccagat gtacggctcg cgcgtggcgc acacccagc cccgcgcccc ttaatgcgcc 6960
gotacgccgc gcgtcgggcg atctgcaccct caggtcatgc acaatgtgctgg aaggagacgg 7020
gtgccgcgcc tcctcgactat accgcagac gcacacac tcga 7054

<210> SEQ ID NO 4
<211> LENGTH: 5932
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence Tp3C-AT
<400> SEQUENCE: 4

gggggggggg gggggggttg gcacactctt ctttgagcgc ttcgtctcgt actgccgctg 60
ggogacacag gcacggcagct ggcggcggggt tgtgacccgc ggcggagcagag gcgcgggg 120
ggcacaggag gggacctgac gcactctacga ctggggtcttt ctgctctcctt ccctgtg 180
attggacaatcttatattctgacatctactgtattacgttttt ggcattgtgca 240
tacctggttat caatactctc atatgtacttttatgtgtgct gctagtcacatt ggtctacgc 300
atttcggtcttctgagatcagctaatctactgactgtctttcttcta tattcggggtc atttgacgta 360
attggcccac gcgcgtgatcct cttgagactgtag tttgactgttt ggtctagtcgagatccctatcattacaca 420
acgaccggag gcacatcctg gctactcctg cctgtgtcgg cctgcagtgaag gcgggtgggg 480
acgccgctg tagcgtcgaac gtcggcggcctgcag ccggtgatctacattggtcgg 540
acgtacagag ttcctcgtag actccgtagg cctgcctgtc ggtggtggtgg 600
ggtctcgactgtcacagtc tgggacttggtt cccctggtccg cggctgac 660
ctggattgtgc atcgatggtat cctgtgtgac ggtctaggttct cagctacagct atgggtaggg 720
attgcggtgtac gcgtcaggg ccacactctc ttcgtcactgc atagacttta ggtctggtttt 780
ctggggtcc caataataactc gccgctctgtc ggtctcagttg gcctctgtc ggtctgggttt 840
ggcacatccag gcacatcctg ggtgtggtggtt gcgtctgcttc ggtctggtttt 900
cgcgtgac gggtgtggtggtgtcagagtc cggtctcagtttc ggtctggtttt 960
ctgggtctgc gcacatccag gcacatcctg ggtgtggtgtcagagtc gcgtgtcgtgtggtctgggttt 1020
ccgtgtgcgc gcacatcctg cttgtggtggt cctgtgctggt gcgtgtcgtgtggtctgg 1080
gagacagaga agacctttgc gttttgtgata gcacccatt ggcttactctg acatccactt 1140
tgcttttct tcacaggttg tcaccctcaca gttcaaatcg agctctttag gctagagtac 1200
ttaaacagcct taacaactag ctgaacctag tggaatcocc gcgggtgcagg aatggctaat 1260
cagctttgagg gatttccccg ccacacccct gaccctggcg gtagacattg ccaatggcctc 1320
tttgctctcg ggggtgaatcg toctgctggc aggctgtgac tggctgctgc tcgtctctct 1380
ggctgaggt cccagggggag atgctgcaac accagacag atacccacccc aggctacga 1440
tccacccac ccctacacgc acaccaccac ccctgccgct gctaatcacc gcctatacg 1500
ccagctgcga cccagctcga acacccacca tatacctctc tcaccaatcg gcacgcagac 1560
agcctttgca atgcctctcc tcggggaaccc gcgtcgaact ccagctgaaa ttctgggagg 1620
cctgatttc ctaccacaaggg agatctcggg ggtctcgact ccctgaagct ctcaggaact 1680
cctcaagacg cccacgcgtca gaggacgaga atgcctcaat gcacaggggt cttggtcttt 1740
cctcaagagc ggctctgacg ctacgtggtact aaatgggtgag aagcagcata aagtgtgacg 1800
cctcagagcc aacacgtcga aaccctggga cccagcgagc aacaccggca aacactcaca 1860
tsagctgagg aaggtgacttc aagggaaaat tgggtgattt gctgacccag gtcagagaga 1920
cacagttctt gcctgcggtag cttacatotton ccttaaagcc caattggaga gacccctttga 1980
agctgaagacc aagcacggagc agagttgcttc cgcgagcagc tgcaggaccgc tgaaggtgcc 2040
tatgtgtagg cgctttgtagt gcgggtctaan ccaccgcttg gataggcttg ccacacgggtt 2100
gctgtgagc agacactcag gcctaaccgc accctctcctt tcctggtctg cttgaggggga 2160
actacagacg ctggaasaat caccacaaca cacctcaccg aagcagtttt gaccaaatga 2220
agoacaacgt ttcgtgcagct ttcacttacc aacactcacc ttaacactcc gttatatctc 2280
gaaagcaagt cttgctggac cggctgaacc taaggtgcttt aagcattgggg cctacccctct 2340
cgggtcaca cggggttcaac tgggctctgc cgctgactcc cttgctggaa cttgctggac 2400
catcagcag aaaaaagctc gctctgggtgc cggctgttgg ttgaggccc ttcacagtct 2460
taccccccacc gaagctcaag tcataacccct ctccgagctcc tggacatagc cttgctgtcc 2520
caagcttccc ctctctctgg gaaaaaggttt gcacccaccc caaaaaatan tcctcctgctg 2580
tctcaccac ccctcctctct tcctctgcttc ccctcctgag gcacccatgaa gacgggtgtga 2640
gtctgctcc ccoccccccc ccctgtgggg ctctccagcc gcggggccct tgcagacagcc 2700
atgataagag accattgaga gttgtgaaac aacaaacctt gacccctcgt aaaaaattcgg 2760
ttatatttgg anattttgta tcattttgct cttctttcct cctttttcct ccctttttcct 2820
cactacacg aacaaacctt actcctattt atgattccttg ttcagggggc gtttggggag 2880
ttttcttttc gcaagtaaaa cttctccttc tgggttaaca cctaataggg ctttaagacc 2940
catactagt gattggccca cttcctctct gcgtggcagt gcgtgtcagtgg gcgtggagct 3000
ggaaagcccg cgggtgatgg cgaatcctgg ggctgcggcc tcagtgagc gcgggctcgg 3060
cagagcggcc gttgggcacc ccccccacc ccocctgca gctgtggctga atagcagaaa 3120
ggctggccag cttcgcctct ccacaaagtt gctgtgcttg aatggccaaa ggcctgacgg 3180
gcgtggtgg ggc gagactaaag gcggtggccc gttgggtggt aagcagctgc gcacggccat 3240
actggcagc gcctctcgct ccgctctcct ccttttctctt cctttttttc cttaagctct 3300
cgctgggttt cccgcttacg ctctacggcg ttcaggggtcc gttttgtgtag gcatttgtcgc 3360
tttaaggccc cctgacccca aaaaacctga ttgggttgat ggctacgtag ttgcggcac 3420
gccctgatac aagcttttccc gccttttggac gttggagttc agctttttaa aatagtgact 3480
cctgttccaa aacctgacaa cacttacaaacct tattttgttc tatttttttg attttataag 3540
ctatgcgcgg attctgacgct attctttaaa aatcagcttg aatctacaca aatctnacgc 3600
gatggtagaa aaataatataa cgtttacact tctctgtagt ggttttttct cttacgacat 3660
cggccgcggta ttcccaccacg cattctgggc actctcagta caacctcgctc tgcgtagcga 3720
tagttttagg aggccagca cccggcgac cccgctgagc gcggctgacgc ggcgccttc 3780
tccgctcgat cgcctttaacg acaagtctggt aacgtcctgc gcggtctgtct gcgctgacg 3840
ccccccccac ccctacggaa acgctggaga gcgcttttcgg ctgctgacgc ggttttttct 3900
ttagttcgaag tcagttgtagt tagtttttagg ggcgtttctt ttggctttaga 3960
gtgccggaaga ccroctctttt gttttttaaatttaaaa taatttaacttt ctaatattga 4020
ttgctgctgc ccatctctttt ggtttttttt ggtttttttt gttttttttt 4140
cggccagcgag ttggataagc aaaaagatgc gaagctgtctg tgggctccag aclgcgttag 4200
ctagctgact ctGTGCTTTGGGT ccgtctcggct tgcgctggatgc agaaccgttc 4260
ccagtttgac gcacccccgtttaag tgcgtcctgc ggctttttccat tttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
gagcgtcagc ttgtctgtgg ctggcaggg gcgggagccc ttggaaagaa cggccagccac 5700

<210> SEQ ID NO 5
<211> LENGTH: 7492
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence:p43C-AT-IN

<400> SEQUENCE: 5

aatccgcaag ctcagagata atggtaggtta aagcagtgct ggcgtcagc ttgatacact cagcagacct 60
tgcaccatca ccaccattggt ttgcaactcc acagcagcgc ctagctgcct ggcggcgccc 120
tctgtctcg gcggagcaact gctctcctct gcgtctggcc gttcgcagtcg 180
ctgcgtctgc ctggtcctgg cggagccca acaggagcttg cggagcagag 240
acacccata cccacactgta cccagcagct cccagctatt ccacccacct 300
ggtgcgtcct ccagcagctg cccgcagctg cccatctctct cgcgcagctg 360
ttcctccct tcacgagaag ctcctgctct gcggagccag ccggtctcct 420
gccgttcag atgggtagct gcggggtcct atgcctgagc ctggcagccac cagcagcagc cagcagcagc 480
ccgtcgtcag aagcagctg ccgcgtcagc ctcctcctg ctcctgctct cccgcagct 540
cagctgctcg cgcgcagcagcc ctcctcctg ctcctgctct cccgcagctg 600

ttgagcagct tttagacagt gcagcagcagc ctcctgctct cccgcagctg 660
ggcgagcagc cagcagcagc ctcctgctct cccgcagctg 720
gttcctcg ccagcagcagc ctcctgctct cccgcagctg 780
aacgagcttg gcggagcagc ggtgcgtcct ccagcagcagc ctcctgctct cccgcagctg 840
ttcctgctg cccgcagctg ctcctgctct cccgcagctg 900
cagcagccag ctcctgctct cccgcagctg ctcctgctct cccgcagctg 960
ttgagcagct tttagacagt gcagcagcagc ctcctgctct cccgcagctg 1020
cctttctct ccaagacctg ctcctgctct cccgcagctg 1080
acagctgctct ctcctgctct ccaagacctg ctcctgctct cccgcagctg 1140
gggtgcgtcct ccagcagcagc ctcctgctct cccgcagctg 1200
acacccata cccacactgta cccagcagct cccatctctct cgcgcagctg 1260
ttagcgtgga ccaaaaaacat ttacgctgtt gatgttggct ttcctgctct cccgcagctg 1320
aacagggcag ctcctgctct cccgcagctg 1380
cctttctct ccaagacctg ctcctgctct cccgcagctg 1440
acacccata cccacactgta cccagcagct cccatctctct cgcgcagctg 1500
gggggggtg gcggagcagc ggtgcgtcct ccagcagcagc ctcctgctct cccgcagctg 1560
gggtgcgtcct ctcctgctct cccgcagctg 1620
gggtgcgtcct ctcctgctct cccgcagctg 1680
cctggagcaco ctgtcccactc tgcattgccc tgaactgcag atctcttggt atggcagagg 1740
ggcgggtgc cgggacattcg ctttccocct gccacatcag aggacagcct agagctgagc 1900
gaatagtaag ccagatgttt gccctgctag aataactgag tgcgaaatca acaacgggca 1860
ggacgcctta gcgcgaaggt accaaatagct cctgatgcct caactattta tctgttaaat 1920
coaagacaco cttcaastaa acaatgacct caataaaaat gcaacagccoc gcacggatct 1980
cctagcctga cctcgccatag atgtctcaat tctgactctt cctctagtttc tgaatattcc 2040
cggccacoct ttttctcgca ggttggatag ttgggcttcaag accttttattc atagctgaa 2100
ggcgtgggcct gcggagcttg tgcgctcaag gggtctcagg ggtgatggct ggagggctc 2160
atgctttctt atctgcgtag acctctgcat caacgctctg accccctggt acgcacccct 2220
ttcctctct cagggcnaat ggacagacoo ctttgaagtc aangcagcag aggaaagaga 2280
cctccaatgt gcgcagctta ccaacgctaa agggctctag atgaagcttc tagccctggt 2340
taacatcaag aagcttgcaag cctggtgctg ctgattgaat acctggcacaw 2400
tgcacacgat atcttctctc tggccatgaga gcgcacactgtaa aaaaacatct 2460
coaagacag atacccctga agtcgctctg gaaataagag gaaacgctgg 2520
tttcaccac aatccacacta ctgggaccctc tgcagactga aggctctagg gtcacaaagg 2580
cacoactag aagctgctcag atggcgcacat cctctcgcttg agcgacccct 2640
ggcggaact ggcggatcctg cattctgctag gcacagacac gcgcacacag 2700
tgcggggccg atgtttttag agcattacac actgctatag cccccctggg tcaaattccaa 2760
cacaccccttt cagtttcatc gataattacca aataacgccat ttcctctcat cttgagaaaa 2820
agttggttag ccaccacaaaa aataactgggc ttcgctctcg tctacccccc 2880
tgcctccgctt cgggtagctg attaagacag ggttgagctg gtaaaccacc cccccccctgc 2940
agggcctgca gcgcggctggct cgccgacatt gcggacagcag cgtgctccctc gctgaacggc 3000
cggccagtgct ccgagxtctc ccgctgtaac ctctacaaga gacccgctgg gcgggtgctg 3060
gcgacgtagta ttcagactg gttggtagtt gcgcacacac ccacatctgt gaaagtgtaaa 3120
aataaggctt aatttatagt ttgagctgt atttattattc ttgaaactct ttaaagctag 3180
ataaaccag tcctacccaa caattctttcc catatttcctt ttcgctcttc gcgggctagc 3240
tggcagctt tcacacacaaa gtaaaaaatt gtaaaaatgtg gtaaatctga tagagcttagt 3300
ggcacccctta gttcgttagg tcggacactc ttttctgtgc gcggctcagc ttcgagaggc 3360
cgcggcgccg acgcgctgctgc gtcggtggctc ctgtctccctg tcgagctgctg 3420
agggcccagc gcggaggtgt cccacctgca cggccattct gacctgacct gacgccgagg 3480
agggggctgt tggcatttgg gcgctttctct cggctcttctg gcaacggctg 3540
gtgctgctgc tgggcgctgg gcgcataagc cactcagacc cgggataact gcctatctca 3600
gacgctggg ctaacccagt cccaaadagtg tagcacaaga gcaacgccac cggccaggaac 3660
cgtgatagct gcggagtcccc cagcaggctg gcgcctccgt gcgcagatgc gcgcacacag 3720
aaaaactatca gcggcgaggg gcggagctgc gacgccagag gatataag ttcaccggcg 3780
ttcctctctg gcacgcttcc cggtgcgtctc ccgctcctgt ccgcagcggctg tccagctaat 3840
tgcgcgctt ctctctgtct gcggagccct cggcctgctg tggcggcttg 3900
tccacgctg tgggcggctgc cgggctcttt gcggagctgg gcgggttagtt ccatcgttaat 3960
cogacagct gcccttatac cgtaaatct atgctttgag tcgaaccgcgt aagcaacagc 4020
ttatcccaag tggccgcaac ccccgttaac agggatagca gacgaggtta tggagcggt 4080
gtcgagcgt tccgttacat gttcggttaac tacggttaac cctgaaggg aaccttttgtt 4140
atggctcgg tgcgatggag ggtccttccc ggataaagg tgggttacct cttacccggc 4200
aaacaaccc cgcctccttg cgggttattt tttgcattca acctacaagt taacgcgaga 4260
aaaaaggt cacactccaa ttctttttgt ctttaagga ggtcgtacgc tcagagtgaac 4320
gaaaaactaag gttaaggagat ttcgcgtcag agatatacaca aaagttcttt caacttagc 4380
crrrrrtaaa aaaaataagat tttaantca aactaaagta taatatagta aacctttgct 4440
gacagttacc atgggttatac cgctgagcgc cttcattccag cggatgtctt atgggcttaa 4500
tcctagtgtg gcccattccc gttcgtgttag atanatacgc aagggaggcc cttaoctctt 4560
ggcggcgatt cgcgagggca actagacgac cccgagccg ccgctccaga tttaatacva 4620
atatgaccgc gcacggagac ggcagagcag gaaaaagttc cgganacttt atccggtcctcc 4680
atccgctgatt ctaaattgtg ccggaacgtg agaatacta gttccgagat caagtctgctc 4740
cgacgctgg ttggtcctgg tacacccagt gttggtcgtg gctgatggct 4800
tcctcctttg cctggttccac acgcgtacac aggatagaat gttctccact tggagtccaaa 4860
aagcgcattt gccctctcgg ccgctcctgg gttctcggg atgaagtgcgg gaaagtgtta 4920
tcctctctgat ccctgtggttt ctctctctgt tcctctcgc gcgtctagtg 4980
cttctctgac ccgctatgat cctcaccacag ctatctgtag ccacagttctt gcccagaccg 5040
agttatcctt gcggggtgag ctaactaggt aatccgccg cccatagcag aaatctaaaa 5100
gtctcctccc ttcgagaagg ttcgagtggc cgaacagctt cccagctctt acgcgctgtc 5160
atgacattg acgtacgga ccctgtgagc cccagctctt ccctagctct cttttctttc 5220
accgagttt gtcggttgcag aaacagccga gggaaatctt cgygacccaa ggggttaag 5280
ggacaggg agttcgcata acttactact ttccttttctt atatatttga aacccattat 5340
cagggttatt gccctactcg ccgataataa tttgatgtag ttgaaaaanaa taacaaaaa 5400
ggggtcctgg gcacatctcc cggaaaaagtc ccacgttgag tcctaaagaca acactattac 5460
agccataac cttccttatt agctggtatc atggagcctt ttcgctcctcc gcttctctttgt 5520
gatcagcgtgc aacacctttg acaccctgac ctcctgagga cggctcagag tgcgtctga 5580
gggcagcgg gcagccgaca acgctcgtgag ggcccgtgctg ccgcttttggcc ggctgttgcgg 5640
gcggtcttat acttacgag ctcacacag accacatga ggtgacccga tctgggtcgtt 5700
gaacacgacc acagacgtct aaggtggaana ctggcgacat caagttgtaa aacgttataa 5760
ttttttttaa atcctggtga aatattttgtt aatacgtctt attttttacnga ctaaccggcgc 5820
aarogcggac aactctttat aatacagaca agtaccggcag ggtatctctgt agtgggttctt 5880
cattttgga aagagacactt cttataagga acgtggaactc cccacgtcagaa gggccaaaaa 5940
cgtttctctcg gcgtctgggc accctttttct capcataaag ttttttgggt 6000
cggtttggcg taaccactctt aacggaacgc aataccgggag gcccaggatt taaccctttag 6060
gggccacgcc gcagggacgtg gggccaggag aaggaagaga aagccagaag gggtgctcctt 6120
gccgctgggc aatgtgagac gccacgcgtg gctcatgacgc gcctaacac ccaacccggcc gcccctatag 6180
gcgggtctg ctgcacgcgct gcctacacat acggcctcg ccgcttctgg 6240
---continued---

gatogctgc ggccttcccg ctattacgcc agtgcagtt tgcacactcc ccctctgcgcg 6300
cctgctgct catctagggcc ggggacacaa aagtcgcgcc agccgcgggg gttgcgcggg 6360
ggcgaagcct gagcgcacgg cggcgcagc aggcagttgc gcactcacta actaggcggt 6420
cattctctc cattagggg cttagttcctag gttattatag gttatatcgcgttatc 6480
tattttgttat tcgcatttag aacgtttgta tccatctacat aatgttgata tttatatg 6540
cactgctca atatsacggc cagtttcgcct tgtatttagc aataagacat aatagatatc 6600
aatcaagggg ccattagttc ctatacatta taataggttc cgggttaacct aatatacctg 6660
aatgcgcgcc cctggtggtc gcgcacagc gcccgcgcct ctcgctctac ccattagcgt 6720
tgtcctcagg gtaaagccaa tagggatctt cctctgaagtt ccaatggtgtg aagtatttcag 6780
taatcctgctatgtt cacgctactg gataactctg cccacagcgc cccctatgga 6840
gtcaacttaac ggtaacgtta cccgctgcca taatgcccaag taccaggtgt caagggactt 6900
tocactctg cgtacactat acgtattgtg catcgtattg aacgttacgg tgcgttcttg 6960
caaggtaccc ggttgcagtt gataggggtgt ggcctcgagg ggtttccgaat gttccacc 7020
cattttgctct gaccttcgca cccataatac gcgaacttctt cacaaatgtg 7080
tatcttccac cccgcctgtga ccaacttggg ceatggtgtg gtaagcttcat 7140
aatgcagct gcttttagta acggcgact cactaagaac tttcgggtg tagttttcct 7200
catttactt gtaacgccgg cagctgcgtt gtaaccacca gttcgactc taaaccgg 7260
gcttgcctag tgacgcaagc ggcgcgcagc aagagctttg cgtttgcttg gcgctcatt 7320
ccactagaaa ctgctggtctg ggcacagag aagatccttg gtttcttctg cgggaactc 7380
tgcgtctact gccacccact tggctcttct cccacagttc gtcacactcc agtctacatt 7440
cgtacctagag cggatacagc cttatagcga cttcctattag gctagccctcg acg 7492

<210> SEQ ID NO 6
<211> LENGTH: 6714
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE: OTHER INFORMATION: Description of Artificial Sequence: PLASMID
<223> P43CB-AT
<400> SEQUENCE: 6

gggggggtgg gggggtggag gcgcacctct gtaagctttc gatccgaccat ggtgcgcg 60
ggcacacaa gctgctggctt ccgccccggt tgccgcgggc gcgtctactg ggcgcgcgag 120
cgcggacgctg cagctgcgttc cctccctct cctacgtgtaa cactgctgcct gattttgc 180
attagcccatt ctattgctgg gtttatagc atatacatct atttgctatt gcgcacacc 240
taacgtctaat cttatcgcttg ctagttattg ttcataactg ggcgcacacc 300
attggctcatt cattacatag atatgaacta tattaaggggt catttaggc 360
tagccagccat ggtgctggcc gctgtaagcttat aactagctggt atgggtggcctcctgt 420
gcgcgacac cccccgctgt gcagctgctatt gtttatagc taaccgcctg aggactttcc 480
agggactttg cctgctggctg aatctggtta gtatattacgg taataactgccttctggag 540
acactctggtc acgctgcgtct cccatgttac gcgctcggt gtaatagctgcc 600
cgcggcagagt cgccttcagc ccgcacccgt ctgcacagct gcgtgctgcct gcgtgcg 660
-continued

cgtattagtc atcgctatta ccatacggtga ggtgaacccc acocttcgct tcaactcctcc 720
cctctccccc cctctcccac ccoccaaatatac gtattttatatttttaaat tattttttttctg 780
agcgagacggc gcgggcggggc ggggggggggg ccaggccccct cgggggggagg gcgggggggg 840
ggggggggg ggaggggcggc cgccgccccct gctggccgag ggggcccggc ggctgggggg 900
ccgcttgtttat atgggagggc gcgggcggggc gcggccctct aaaaacgctg gcggcgggg 960
ggggagggc gtcgagcgaga tgccttcggcc cctggcccggc cctggcggcg ggcctgcggc 1020
gcccgggggg gtcgagctct gcgcctacgtg tccaaacagt gagcggtcgg gacggcccctt 1080
tctccggggt cttcctatgcgc cctgtggtttt aataacgcaag tgccttttttt tctggtgactg 1140
gtatacgctc ctgagggacgc ccgggggagcc cctggtgacg gggggaggttc cttcgaggggtg 1200
ggtgaagcggt gtgtgctgctg ggggggggag ggtggtcggt gctggccgag gcggccctctg 1260
agcgctggg ggcgcggaggg ggggtttttgt gcgtgcagag agctgctgctg gcggccctctg 1320
cctgggggg ggcctgctggctgg ggggggggag ggcctgcggc gcggcggcgc gggggggggg 1380
gtctggtctg ggggggcttgag gcgcgggttgag tggccggtgg gcgcgggtctg tacaaccccc 1440
tctccggggg cctccgcggtg gtcgtgcttc ggccgtgctg gtcgagctct 1500
ggcggggcc gcggcggcgc gcggcggcgc gcggcggcgc gcggcggcgc gcggcggcgc 1560
ggggggggg gcgcggcgcgc gcggcggcgc gcggcggcgc gcggcggcgc gcggcggcgc 1620
ggcgcggcg gcgcggcgag gcgcggcgag gcgcggcgag gcgcggcgag gcgcggcgag 1680
ggcgccggc gcgcggcgag gcgcggcgag gcgcggcgag gcgcggcgag gcgcggcgag 1740
cgcaccccct ccagggggcg ggcgggcggcg gcgggagggg cgggcgggcgg gcgggagggg 1800
ggggggggg gttcggtcgt ggcgcggcg gcgcggcgcgc gcgcggcgcgc gcgcggcgcgc 1860
tcccgggggg ggcgcggcg gcgcggcgcgc gcgcgggaggg cggggggggg cggggggggg 1920
cgtcttcggag gcgcgggcgag aggcgtcagc agcgcctccct aagctgcggc agcgcctccct 1980
cgcggccgg cgcggcggcgc gcgttttttt gcttttctttct gcttttctttct gcttttctttct 2040
atcaagctgg gggattttttct aggagacgcc tcggccgcgg gcagctccgtt gcagctccgtt 2100
ttcgcttgc cggggggttg cctgcttcgg gcgcgggcgc gcgcgggtgg gcgcgggcgc 2160
tgctgctgggc acctccaggc ggaggggtgg cgcctgcttt ctggctgctgg gcgcgggcgc 2220
gatcggggtg gtcggggtgg cttgctgggc acctccaggc ggaggggtgg cgcctgcttt ctggctgctgg 2280
ggcgacccgg gcgggctgctg cttgctgggc acctccaggc ggaggggtgg cgcctgcttt ctggctgctgg 2340
cgcggaggg gcgcgggggc gcgggggggg gcgggggggg gcgggggggg gcgggggggg 2400
cgcgggggg gcgggggggg gcgggggggg gcgggggggg gcgggggggg gcgggggggg 2460
cgcgggggg gcgggggggg gcgggggggg gcgggggggg gcgggggggg gcgggggggg 2520
ctccagggg gccggtggccg cggggggtgg cggggggtgg cggggggtgg cggggggtgg 2580
cgcgggggg gcgggggggg gcgggggggg gcgggggggg gcgggggggg gcgggggggg 2640
gatcggggtg gtcggggtgg cttgctgggc acctccaggc ggaggggtgg cgcctgcttt ctggctgctgg 2700
gacgcgggg gcgggggggg gcgggggggg gcgggggggg gcgggggggg gcgggggggg 2760
gacgcgggg gcgggggggg gcgggggggg gcgggggggg gcgggggggg gcgggggggg 2820
cgctgctgtg ctgcttgggg ccgggggggg gcgggggggg gcgggggggg gcgggggggg 2880
ggctgcttg cgggggggg gcgggggggg gcgggggggg gcgggggggg gcgggggggg 2940
aacactacgc acctggaaaaa tgaactccccc cacgcatacc tcaccaattt ccttgaaant 3000
gagacagaa ggctcgcagc ttctacattta cccaaacttgt cacattaactg gacocatag 3060
tgaagaagctc tctgcggccaa acctgcgtct ctaaacagtt ctcgaatgaa ggctgtcctc 3120
tccggggctc cacaggaggg aaccccctgaag cttccacaggt cggccaaattg ggtgtctgtc 3180
aacccgctacag aaaaagggcct tgaagctgctg gggccactgt ttttagaggcc ccgccagctcag 3240
tatataccoc ccggagttccaa gtttccaataa cccttttgtct ttttaagttat tgaaccaaat 3300
aaccaagtttct ccctcttctct gggaaaaagt tggactcaca cccaaanataa actgcttctc 3360
gctctccac ccctgctgctc cccctctcccgt ctrgtcattat ctgctactct aacaggggctt 3420
gagctctggat ccccccctccc cccctggtcgg gcccctgcccc cccgcgcgtcc cctgcagcag 3480
aacccgctcc aatacttctga gaccccggtc acacccaaac cggcccacgtc aggagctcgaa 3540
gttttattg tgaataattcg gacgccattcc ccaaccattacta aggtgctataa 3600
aacccgctcc aacccgtccctc cctagtattatactgttccgctccttctct aggccccggg 3660
gagctctggat aacccgctccctc cctagtattatactgttccgctccttctct aggccccggg 3720
cocccctgctg tggagttggc cccctctcccgt ccagccttgctc gacccgtccctc cctgcagcag 3780
cggccagcgc ccgcgtccgcc ggccccggctt ccctgcagcag cggccagcgc ccgcgtccgcc 3840
cgcgtccgcc ggccccggctt ccctgcagcag cggccagcgc ccgcgtccgcc ggccccggctt 3900
gacccgtccctc cctagtattatactgttccgctccttctct aggccccggg 3960
gagctctggat aacccgctccctc cctagtattatactgttccgctccttctct aggccccggg 4020
aaccttcgaa ccctgtctcccgg ccctgtctcccgg ccctgtctcccgg ccctgtctcccgg ccctgtctcccgg 4080
ttcggggctt ccctggggctt ccctggggctt ccctggggctt ccctggggctt ccctggggctt 4140
gacccgtccctc cctagtattatactgttccgctccttctct aggccccggg 4200
ttcggggctt ccctggggctt ccctggggctt ccctggggctt ccctggggctt ccctggggctt 4260
cctcttcgaa ccctggggctt ccctggggctt ccctggggctt ccctggggctt ccctggggctt 4320
ggagtattgag cattggggat cccacaaaaa cggccagcgc ccctgtctcccgg ccctgtctcccgg 4380
gacccgtccctc cctagtattatactgttccgctccttctct aggccccggg 4440
ttcggggctt ccctggggctt ccctggggctt ccctggggctt ccctggggctt ccctggggctt 4500
cctcttcgaa ccctggggctt ccctggggctt ccctggggctt ccctggggctt ccctggggctt 4560
ttggtgtccct cctctgtctcccgg ccctgtctcccgg ccctgtctcccgg ccctgtctcccgg ccctgtctcccgg 4620
ggagtattgag cattggggat cccacaaaaa cggccagcgc ccctgtctcccgg ccctgtctcccgg 4680
ttcggggctt ccctggggctt ccctggggctt ccctggggctt ccctggggctt ccctggggctt 4740
cctcttcgaa ccctggggctt ccctggggctt ccctggggctt ccctggggctt ccctggggctt 4800
ttcggggctt ccctggggctt ccctggggctt ccctggggctt ccctggggctt ccctggggctt 4860
cctcttcgaa ccctggggctt ccctggggctt ccctggggctt ccctggggctt ccctggggctt 4920
aaccttcgaa ccctgtctcccgg ccctgtctcccgg ccctgtctcccgg ccctgtctcccgg ccctgtctcccgg 4980
acccgctccctc cctagtattatactgttccgctccttctct aggccccggg 5040
ttcggggctt ccctggggctt ccctggggctt ccctggggctt ccctggggctt ccctggggctt 5100
cocccctgctg tggagttggc cccctctcccgt ccagccttgctc gacccgtccctc cctgcagcag 5160
cgccccgctt ccctggggctt ccctggggctt ccctggggctt ccctggggctt ccctggggctt 5220
-continued

cctntaacat ggtgtgaac actgcggcga acctactctt gcaacaacag gcgaagacga 5280
gagagctac cgctttttta cccaaactgg gggccagctt aaactctctt gatctttggg 5340
aaccgggagt gtnaagacc ataaacacgc acgcgtgtga caccagacct ctgtagcaa 5400
tggcacaac gttgccacaa cttaactctg gcaacactct taactctact tocggccaaa 5460
aatattata ctgtaggag gcggataaag ttgtagaacc acctctgccg tcggcccttc 5520
cgggtcgctt gttatactgt gataaactct gcacggtgga ggtgtgcctt cgcgtatc 5580
ttgcaagatc ggcccacgat gtaagcctt cccctatgtgc agttatctac acgcagcggga 5640
gtcgaagac actctattgc ccgatagggc agtagtact gcggctgc gatagtgtac gatct 5700
agcctggta actgcaagac caatgttact ccctatactct ttgctgat gtaaactctc 5760
atttttat cccacactgact aggctgactt gctcttctg ataactctgac aaccacaccc 5820
cataactct gtttggtcct cactgagcct tgaacccctt cagaaacgtac agaagactct 5880
ctgtagacg ttttctgttg ccgtctaatc gtcgctgca gacacaaata caaacctgtc 5940
cagcgggtgtc ttcttcccag gtaacgcgcc taaccactt ctttccagcag gtaactgtc 6000
toaacgaagac gcagactact cttctagtgta gcgtctaggc ggcgcacact 6060
toaacgact cttgatccag cctcataaact ctgctctgtgct aacatcgtta ccagtgggt 6120
tctgctctgg cgtataagcg ctgctaactcc cccctatgtgc ctgctgatag ctaaccgata 6180
cgtgcagcgc cggccagcct ggtgggttga ccgtgcacacc ggcagccttg gcgcagacga 6240
cctacaacga actgcagact ctgctctgg gcctctgagc aacgcgcacgc ctgccgag 6300
gggaacaggg ccggtgatagc ccctggtgag cggccagtgc aacgcgcag 6360
agcctggta cggccataac ccctgatatt atagctcagt ctgcagttct gcacgctgg 6420
ttgagctag ctgtagattag ggggagggact gtttaggaaa ccctgacagc gacacgga 6480
aaccgggagt tcggctgtgt cttgctggct ttgctgctgt ctgctgctgtg aaccacaccc 6540
tgctacttt ccagttacatgc gtagaactct tacaacgcttt ctgagctacta ctgctgctg 6600
gagcagcagac gcgcagcagct gcagagcaga gggactggaa ggcgcacact 6660
tacgacacg gcgcacgcat ccctctaggt gcgcagcagct gcgcagcagct gcgcacacgc 6714

<210> SEQ ID NO 7
<211> LENGTH: 6981
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE: Description of Artificial Sequence:PLASMID C-A2
<400> SEQUENCE: 7

tctgacactag tggtagcctccc gggtgcagg aattgcctat ccagctttggg gattttcctag 60
cacaccaacct gatgtgacaa gcctttcttg ggccttgagc ttgctgaga 120
ttcctgctgct ccgatcctgg ccccttcttg ccagccctgtt tctgggtac 180
atgctgcacc gcagatagtc atataaaccctg gctgctgactt ctgctgctgat cccctgtt 240
tcactctcact gcggctgata gcctgatagc ccacgactca gacaagcatc gcgcacacgc 300
acgacaccct catcctccttc ccccttccgttg gctgctgactt cctgctgctgat cccctgtt 360
ttgctgctgct cgctttcttt cccttgctgat ccccttcttt ccctgtgctgctgat cccctgtt 420
agctgacact gcggctgata gcctgatagc ccacgactca gacaagcatc gcgcacacgc 480
cagacagcca gttccagtgg accacagggca atgcccgttt tctcagcggg ggcctgaacc 540
tagtgctas gtttgggag gatgcttataa agttgtcata ctcacgaagcc ttcactgtaa 600
atctccggga caacgaggca ggccagaacc agaacaagct ttcagcgtgg aggctcctctc 660
aagggaaat tctgcttcttg gctcaaggag tgcgagcaga aggagcttcctt gtcgtctgtgc 720
attactccttt cttaaaaggg caaatgggaga gacccccgatt actaaagggc acggaggaac 780
agggctccga ggtggaacag gtcgcaacccg tggagttcag cttgattcag gcggttaaggca 840
tgttttaacac caagcaagctt aaaggagtgc tcaagctgtgg gttcgqcgtg aataaatctgg 900
gcgatgccac gcgaactttct ttcctgtcag atggggagaa actacagccg cttcccaattct 960
aactccacc gctgtaatcc accaaatgcc ctgagcaagg ctgggagcgcg ttcgacacagc 1020
taccatattc cnaactctgccc attactggaag cctnigtcatc gaagagcgtc ctgcttcacc 1080
tggggatacc taaagcttggc ttcgataaggg cttgcttcac gcgggtacaag ggggaggagc 1140
cctctagctct tccaacggcc gtcgataggg cttgctgctac ccgagtgctca gggagccagt 1200
aagctctgag ggcagcttctt ttaagggcga ccaaccttct ctctctcttgg 1260
tonacaccctt cattttttctct taattaaggct aacaaatcct cagttctctcc cttctctctg 1320
gaaacttgtg gatcctccac cccaaataac ctgctcctcc tccctcctcc cccctctcct 1380
tccctggccc cttcctgctgg tgcattttctaa gaggcttgaa gttgcttaccc ccccccccc 1440
tctggaggg ggcgtcggag cgggggtcag cagagctgatct ataatgttctag ataatgttctag 1500
acacactaca actacactgtc agtgaaccac ctgcttattctct tcgtgaatgtg ctgcgtatctac 1560
tgtttttttc gtaaccattt ttaagcttgg acctaaagct cagctctctc atgaagctat 1620
atatcacccct ctagtcttacag ggaggttggag gaggtttttc tggctctccc cggagccagt 1680
ggtttggata gggagcaacc aatcctcctt caccgaccgg tggagttctct ccaaacagat 1740
cgaagcgagt gggtgggttct aagggagcgg ccacaaacct tccaccccag ccctggaattg 1800
tccccacacat tcgcacgagg cccgcaaggg ctgggatttgact gcagggctca 1860
gagagcctcg ggaggctgac ggtcggcgct cctggaggtt gcctggtctg acgagcgttc 1920
ctgcaacacc gcggacccct ctgaggaaaa ctttctttcttt tggctgcagtga gacccaacaga 1980
dgccctgctc ttccgctgcttc ccgggtggag cgggctgccct gcccggctgc cccggggctgc 2040
caatcgctgc ctggtgtgca gcgcctgctga ggggggctc cctttttttc 2100
tgtagccact cccagctctct cggccgcttc atgagctgct ggcagcgcctt ggaagccctg 2160
cgctgtgctg cggtcttctt cgctgctgtct cggagttgtct caagagctgcc cttgagctctc 2220
gagagcctcg ggagttagtg cgccagctca gccggtcagc ttcctctctctt tggctggtctg 2280
tccccacacat tcgcacgagg cccgcaaggg ctgggatttgact gcagggctca 2340
cggcgtctcc gcggaggacac cccagagaga aagactccta cggagcgcgg ccgtactgcag 2400
tggagcggtgg tttggggctct ccaggtggct ctcgaggaag ctggctcagcgg cggcgctgtc 2460
cgcagactgt cgcgcgtgtg ggggcttcag gggaggcagc ggggttccttt tcggagcagtcc 2520
taggtgctgctc cggtctgccct cggagctctc gtcgtctctc ctgctggatgc 2580
acttggtgctc gccggggttg cggagcgtct ccggcctctct cggaggtggtt acgagctctc 2640
tggctgcta gcgtgctgctc gagaagctgt cggagcgtct ctcgagtcag ttcctctctctt 2700
cgcggtcag cgcggcctg cgcgtctctc cgcgtctctc gcctctgccct 2760
ttttcaaat tatgtaagcct tttatcagggt tatattgtcct atgtagctgt atatattttga
atgtatttag aaaaaaaaat aataaggggt tcggccgact tttcccgagga aagtgcgcc
atatattatt gaaactataa aactcagtc gattactaat tcagttactg acatatttga
5100

tgacgcttat aaactagtc attaactat aaaaaaagaggt gattcggaag
5160
gctcttct ctggcgttga tcgggtgcttc cgtgtagac ccgtgagcact ctgctgtctcc
gttgcaag ccacgtggct agccaggtgg ccagccaggtg agccaggtgg
5220
ggacagctgc acacgtttct gcgtgacggt gacgagcagc gacagcagc
gttcaggggt gttggtgggt gttgggggtg gttttactat gggccatggt atgagagttt
5280

gccagagct tcagcagctag gctgtagagat gctgtagagat gctgtagagat
5340
gctgagct gtttctctag cttttactgt gttcgcgctag agtccgaccaga ttaccctgt
5400
cacaagagca gggctgtgtgt tttccgtgtag ggccagcagc gacacagcagc
gttctgttaat ggaagagggt ggggagcaatt ggggagcaatt ggggagcaatt
5460
cacaagagca gggctgtgtgt tttccgtgtag ggccagcagc gacacagcagc
cacaagagca gggctgtgtgt tttccgtgtag ggccagcagc gacacagcagc
5520
ggggagcaatt ggggagcaatt ggggagcaatt ggggagcaatt ggggagcaatt
5580

tggccctgtg aacggtttgcc acagcgctgtg acacgcagctg acacgcagctg
tggccctgtg aacggtttgcc acagcgctgtg acacgcagctg
5640
tgcacagctgtg acacgcagctg acacgcagctg acacgcagctg
5700

ttttactgtg gggctgtgtgt tttccgtgtag ggccagcagc gacacagcagc
ttttactgtg gggctgtgtgt tttccgtgtag ggccagcagc gacacagcagc
5760

gtggggctgtg aacggtttgcc acagcgctgtg acacgcagctg acacgcagctg
tggccctgtg aacggtttgcc acagcgctgtg acacgcagctg
5820

<210> SEQ ID NO 8
<211> LENGTH: 6142
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence:PLASMID
p43maENC-AT
<400> SEQUENCE: 8

gggggggggg  gggggggggg  ggggacccct  tgcggcgccc  tggctgcccct  acggagccgg  60

ggcggacgaa  ggtggcgccc  ggcggcgggt  tgcggcggcc  ggcggcggag  aggcggagag  120
ggcggccgaa  gggggccggt  acgcctccct  tggctggcctt  ctagatcctca  ccctctctca  180
cgacccacaa  atagggcggcc  ggtggagggg  tggctggaggg  tttttttttag  acggctgggg  240

tatggcgcc  agccgagccgg  tggctgcccct  ctattatctat  gtcgctgggag  ttttttttag  300

gagcggctag  cgctgctggg  accaaattca  ggagcggca  caactctcct  cccacgctc  360
tatattggg  attgtagattc  tgggcagctgta  tgaattatatc  atagatccat  420

ttatatatc  atagcggact  tgggcagctgta  tgaattatatc  atagatccat  480
tatatattc  atagcggact  tgggcagctgta  tgaattatatc  atagatccat  540

tatattattc  atagcggact  tgggcagctgta  tgaattatatc  atagatccat  600

tatatattc  atagcggact  tgggcagctgta  tgaattatatc  atagatccat  660

tatattattc  atagcggact  tgggcagctgta  tgaattatatc  atagatccat  720

tatattattc  atagcggact  tgggcagctgta  tgaattatatc  atagatccat  780

tatattattc  atagcggact  tgggcagctgta  tgaattatatc  atagatccat  840

tatattattc  atagcggact  tgggcagctgta  tgaattatatc  atagatccat  900

tatattattc  atagcggact  tgggcagctgta  tgaattatatc  atagatccat  960

tatattattc  atagcggact  tgggcagctgta  tgaattatatc  atagatccat  1020

tatattattc  atagcggact  tgggcagctgta  tgaattatatc  atagatccat  1080

tatattattc  atagcggact  tgggcagctgta  tgaattatatc  atagatccat  1140

tatattattc  atagcggact  tgggcagctgta  tgaattatatc  atagatccat  1200

tatattattc  atagcggact  tgggcagctgta  tgaattatatc  atagatccat  1260

tatattattc  atagcggact  tgggcagctgta  tgaattatatc  atagatccat  1320

tatattattc  atagcggact  tgggcagctgta  tgaattatatc  atagatccat  1380

tatattattc  atagcggact  tgggcagctgta  tgaattatatc  atagatccat  1440

tatattattc  atagcggact  tgggcagctgta  tgaattatatc  atagatccat  1500

gcggctggg  gcggctgccc  atctggccc  ctactgcgct  ctgtccctct  cgcggcgggt  1560

gcggctggg  gcggctgccc  atctggccc  ctactgcgct  ctgtccctct  cgcggcgggt  1620

gcggctggg  gcggctgccc  atctggccc  ctactgcgct  ctgtccctct  cgcggcgggt  1680

gcggctggg  gcggctgccc  atctggccc  ctactgcgct  ctgtccctct  cgcggcgggt  1740

gcggctggg  gcggctgccc  atctggccc  ctactgcgct  ctgtccctct  cgcggcgggt  1800

gcggctggg  gcggctgccc  atctggccc  ctactgcgct  ctgtccctct  cgcggcgggt  1860

gcggctggg  gcggctgccc  atctggccc  ctactgcgct  ctgtccctct  cgcggcgggt  1920

gcggctggg  gcggctgccc  atctggccc  ctactgcgct  ctgtccctct  cgcggcgggt  1980

gcggctggg  gcggctgccc  atctggccc  ctactgcgct  ctgtccctct  cgcggcgggt  2040

gcggctggg  gcggctgccc  atctggccc  ctactgcgct  ctgtccctct  cgcggcgggt  2100

gcggctggg  gcggctgccc  atctggccc  ctactgcgct  ctgtccctct  cgcggcgggt  2160

gcggctggg  gcggctgccc  atctggccc  ctactgcgct  ctgtccctct  cgcggcgggt  2220
cgtggaccag gttaccacccg tgaagggtcc tatgatgaag cggttgggaa tgttatcctcet 2280
cagcgcctctg aaaaaagacct tggctcgggt gcctgcagtg aaataccttg gcgaatgccac 2340
cgcactcttc ttcctgctctg gtagggggaa actacccac ctggaaaaatg actccacccca 2400
cgtatctca acaagcccttc tggaaataag aagccaaaag tgtggcaagt tacatttacc 2460
caaagctcgc ataatgggaa cctatgatct gaaagagtcc tgtggtcanc tgtggcataac 2520
taagcctcgc acaatgaggag ctgagctcctc ggaggtcacc gcggcagcgg gcacagaaaat 2580
ctccacgagg gttgatagcc cttgctgtag ctcagacgag aagggcactg aagcgtgtgg aagctggcact 2640
ggcgtttt tttgagccac tacccatttc taccccccoc gggtaagttg ttagacaaccoc 2700
cctgcttctc tttagtgcag aaaaataaac caagcttccc ttcttcattg gaaagtggtg 2760
gactccacco aaaaatatca ctgctctgac tttctaaacc cctccccctcaca ctccctccca 2820
cocctcggga tgacatttac gaaaggttga gctgtcatac coccccccoc cccgagggg 2880
cotatcagccg ggcgtcgccc ttcgggcccc actgatgatg acatagtga cttggacacaa 2940
acccondac gattgtcgag aaaaataaag gtttttttgt gaaatgtgga tggatatttgtgt 3000
ctatccgaat ccattaagc ctggtcaattt aaaaataatc aacgataaacc tctcccatttt 3060
atgttctcgg ttctaggggag gatgtgagg gtttttttac gcgaatgaac cttcccatttt 3120
tgctgtaacaa ctgaggacct cttgagttag cggattggca ctctctctct 3180
gacgagcac gcagctcctcg gccagcggccg ggcnaagccc gggcgtcgcc gcagctcttg 3240
tgagccagcc tccatgagcc aacctgaggg ggaagaagcc ggccagccgg ccoccccc 3300
cocctcggga gctggggtat ataggagaga ggccagccct ccagccctct ccacacgtt 3360
grctgcagct ctgacctgct ggcgagacg gcgtgagtaa ggcgtggtgg gggggttg 3420
tggttgggt tgggtgcttc gcgttccatg ggcgtgctgtc cccctctcgt cacaccacct 3480
cttttcttctc cttttccttc tgtgctctgc ggcggctttt cccctctctc atttttagct 3540
ctggggctc ttttggtggt gtcgcgcccc ctgacccccca aaaaacttga 3600
ctgttggat gattgggtca tgtggttcag agccttgtct gccccttgttcc 3660
tgggtgctgc gcgtgacttta ataggagagt cttttcattct ctttgccaa cntgaccccc 3720
tatgttctgct tatctttttc atatatggg gtttggggct atgggtgcc atgggttg 3780
ahatagctgt attataacaa aaatatccaa aaaaatctttt ccgctttcctt 3840
ctctctgcag ggtgcttcctc cttggtggat gtaattcacc aaaaactttttt ccgtattccat 3900
accctctg aaacttgctc tggagtgctt cggagttcag cggagttgaatt ggggccaaaaa 3960
ctgggtcagt gcagcctggc gctttggcag ccagctcttggt ctcgccccct ccgctttccc 4020
acgccgttc ggcgtgctgt tgttcacgatt cccgcccgg ccggcgacta 4080
cggagggccg tgtttagttagcta atttttttaa tagttattaag tgtttgaatt 4140
tacgctgtc cttggcattttt cttggacagca gctgttttgcattttttt ccgtgggggg 4200
tacatattt aaaaatgctt tccgttccag acagtttacc ctttgtaagtcc atgggttgat 4260
taatggaacct ggaagagtct gccgtttaaag ccgtatattt cttttttttt 4320
gggttccttg acctggtctc caaccttttc ccgtggggagt 4380
gaaactagt tgtggggacag atgggtgtttt atcgaatctgc gcgcgctatttgc 4440
ctggagacttt ttggggttcgg aaaaaatcct ctaattgtc gagcaattttaatatggttaa 4500
<table>
<thead>
<tr>
<th>Sequence</th>
<th>Start</th>
<th>End</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGGCGCGG TATATACCG TATTAGCGC GGGCGAGGC AACTTGTCG CGCGATAAC</td>
<td>4560</td>
<td></td>
</tr>
<tr>
<td>AATTCGATA ATGACTTTGG TGAATGAC AAGATACAG AAGATACAG TACGGATGG</td>
<td>4620</td>
<td></td>
</tr>
<tr>
<td>CGGCGTAT CAACTTGTTC TGCGTATCAG CTGGTCTGG CTCGAAGAG CACCGTGGC</td>
<td>4680</td>
<td></td>
</tr>
<tr>
<td>CAACTTCAG AAGATGGC TATATCAG CTGGGATCT TATATCAG CTGGGATCT TATATCAG</td>
<td>4740</td>
<td></td>
</tr>
<tr>
<td>TCTGTCGAC GTCGTCGAC GTCGTCGAC GTCGTCGAC GTCGTCGAC GTCGTCGAC GTCGTCGAC</td>
<td>4800</td>
<td></td>
</tr>
<tr>
<td>GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>4860</td>
<td></td>
</tr>
<tr>
<td>CCTGTCGAC GTCGTCGAC GTCGTCGAC GTCGTCGAC GTCGTCGAC GTCGTCGAC GTCGTCGAC</td>
<td>4920</td>
<td></td>
</tr>
<tr>
<td>GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>4980</td>
<td></td>
</tr>
<tr>
<td>GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>5040</td>
<td></td>
</tr>
<tr>
<td>CCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>5100</td>
<td></td>
</tr>
<tr>
<td>TCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>5160</td>
<td></td>
</tr>
<tr>
<td>CCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>5220</td>
<td></td>
</tr>
<tr>
<td>TCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>5280</td>
<td></td>
</tr>
<tr>
<td>GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>5340</td>
<td></td>
</tr>
<tr>
<td>CCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>5400</td>
<td></td>
</tr>
<tr>
<td>CCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>5460</td>
<td></td>
</tr>
<tr>
<td>CCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>5520</td>
<td></td>
</tr>
<tr>
<td>CCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>5580</td>
<td></td>
</tr>
<tr>
<td>CCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>5640</td>
<td></td>
</tr>
<tr>
<td>CCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>5700</td>
<td></td>
</tr>
<tr>
<td>CCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>5760</td>
<td></td>
</tr>
<tr>
<td>CCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>5820</td>
<td></td>
</tr>
<tr>
<td>CCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>5880</td>
<td></td>
</tr>
<tr>
<td>CCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>5940</td>
<td></td>
</tr>
<tr>
<td>CCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>6000</td>
<td></td>
</tr>
<tr>
<td>CCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>6060</td>
<td></td>
</tr>
<tr>
<td>CCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>6120</td>
<td></td>
</tr>
<tr>
<td>CCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG GCGGCGGCG</td>
<td>6182</td>
<td></td>
</tr>
</tbody>
</table>

**SEQ_ID_NO**: 9

**LENGTH**: 6924

**ORGANISM**: Human

**FEATURES**: Description of Artificial Sequence

**PLASMID**: p43rmaENC-AT

**SEQUENCE**: 9
aagctgaca ctcacgtgat aatgctgag ggctctgatt tcacacccac ggagatcgcg 2640
gagctcag ctcacgtgaag cttccaggaa cttcccgta ccccaaccac gcagacacgc 2700
cagcctccgc tgcaccacgg caatggtctgc tcaccccgag aggcctgattga ccagcgtgat 2760
aagtcttttg agatgtttaa aagattttgc cttcagcaagt ctcctacttt ctcctctcggg 2820
gacaccggaag aggccagaa aagagctcaac gattacgtag gaaaggttac ccacaagggaa 2880
aattgttttc tcggactacag ggtgctcgag gaccaagttt cgtctctgt gatctccattc 2940
tcttttaaac gccacctgga gcacagccctt gaggtcacaag acacccagga agagaccttc 3000
cagctggacgc agtgcacccg cgtgaaggtg cctcagtaga aagcttccag ccagcccaaac 3060
atccacgcgt ctaagagcgt gcctcagctgc gcagcttgag gcggagctgtcc ttcaataactg 3120
acgccatct tcttcctgcgc tgctaggaggg aaccctacgc aocctgaagaa tcgaactccg 3180
cagatatac tccacacggg cttgggaagcc gacacacgga gttcgctgcc gcacatattaa 3240
cocaccaagt cctttactgct aaccctatgc ctacagagcg cctcggtagc aacctggcaac 3300
aactggttg ctcacatacg cttgctgtgc tcggaggtca caggggttcg aacccctgaa 3360
cctcctccagg cctgtctgag accctgagcg acaggctgagc aacacccagag ggcttgctgc 3420
cccgtccttg aatgccgtttt gcaggtgcttc tatctctcct cccagagggg gacgaggtcgt 3480
cccttcgttc tctttaggtg ttgcaaaattt aaccctttctc aacccctttt cgggagggta 3540
gtattccgaa ccccttttttt cttccctccg cgtccttcag cccctttttt aacccttcggc 3600
ccocctccgg ggctcttcat aagaggggtt gcagtttctaa ccocccccttc ccccctggg 3660
gccttcagc cgggctggcc cttcagcgag acagggaa gattattttgct gttggggccc 3720
aaaccccaac tgcattagcc tgaataaaaa gttttttttg tgtttttttg gttgcttttgg 3780
ctttattttt aacccttattt aocgctcaatt accaagtttt caccacaaaatt tgcacctttt 3840
tatagtctac ggttcctggg gagatggtggt agtgttttttta acaagcaaaaa aacccttttt 3900
aatgtgctaa aacccgattac gcctagggaa cccctagtgca tggagtggcc cactcccacc 3960
ctggacccgg gcctggaccc ccacaagggc gcggggctcc gcggcgccttt gcgcaaccttt 4020
ggctcggcc ccttcctggag ccagcagagc gcggagcaga agcttgccaa cccctttttt 4080
ccccccgggg cggcctggcg ctaaacgga gagcccccca cgggtgcccc ttcocacccg 4140
ctgctgactg gttcctctgg gtagggggtg cggcctctgg taattataac gacagctcggcc 4200
gggtgtagtt tcacacggcg ctattgctct cagccagacac gcctcctgact ggcggttctc 4260
ctcttcttct tctggcctac gctgctctgg ccgccggttgg cgttgatttt cattgttttt 4320
cggggctttcct ccttttggcttg aacctctggc ccccaacattt aaccccttttt 4380
gaagggctg gatgggttag gttcctacag gcctgcctgg ccacccattt 4440
aagctgggt tgccgctcttt tgcacgtgct cacccctg 4500
catcctgcc ctattttctgt gtttcttttc aacccctgaa aaccccttcg 4560
aaaatggggc gatttttcctt gattcctttt gttttttttt gtttcttctttt 4620
aattctgt ccaccatttc cttcccgag cgcctggttc gtttcttttc gttttttttt 4680
aagcctcag tccacacggc cctgctgcttg ctaggtcag cagagccgaga ccoccccccc 4740
ccagctggtc cgggctggcc cttcagcgag gcctcctgcc ctgccttttc gacagctcggc 4800
tgcacccrgc cgggctggcc atgtgtctgc gtttccacgc gcctccttcg ccagcccgcc 4860
gacgaaggg cctcggtga  cgcctatttt ttatagtttaa  tgtgtcgata  ataatggtttt  4920
cattagcctc aggctgccct acctggtggc  agttgagcgg acccctatt  tgttttatatt  4980
ttatataataa tccaaatagt tattcgtcct  tgaagactaa acccgtgata  atgcctcatt  5040
aatattgaag  aggagaaatt  attggtatat  accattccctg  tggcgccctt  attctctttt  5100
attggtcatt ttgctctctc gttttgggtct  acccagaacac  gctggtgaaa  gtaaagagtg  5160
tgaaagctct  gttgccgctc  cagatggtcc  aatcagacat  ggtaccccac  agcgggtgaa  5220
tctcttgagc tttaacgccc  gaaccagcct  ttcaaatgat  gacacttttaa  aaagttctgc  5280
tatgctgggtct ggtatattctc  cgtatgacgc  cggccgaca  gcaactctgtg  cgccgctaac  5340
acatcctca  gtaatgcttg  gttgagtgct  cccagcgtaa  acgaacagctt  cttaacgtgc  5400
gatggtatg  angaataata  tgcagtgctg  ccaataacgt  gatgtgacac  acgtgcgcca  5460
actactactc  gagaatacgc  ggaagacgca  agagatcacc  gcgcttgcgct  ccaacattgg  5520
aggcctacgt  acotcgctct  gatgttgctg  aaccgacgtt  gatagaacgc  ataacacaccg  5580
acagcggta  cccagcgcag  ccttgagcaca  tgcagcaca  cttgcacgaa  ccaatctgct  5640
gacaaactct  taacctagct  tcgagcgcac  aataataaga  ctggatgagg  gggatgaaga  5700
tcggagacgc  acctctgtgc  tctggccttc  ccggtggtctg  gttatactgtg  gataacacttg  5760
gcggcggtgca  gctgcggct  cccggtatca  tgtacgcatt  ggggagcagt  gtaaagcctt  5820
cgcctgatct  acgagcggta  gtcgcgcaac  ttgagagctg  cgaataagcgc  5880
agacgcgccgt  gatagagcgc  taactgagta  accttggtga  acgtgcagcc  caagttcact  5940
cataaccttt  ttaagctgct  tttaaaccttc  aattttatct  taaaaggtgctc  tagtgaaga  6000
tctctttttc  taaatcagtc  acccaactcc  cttacagcag  gtttctggcc  taattgagct  6060
cagacccgct  aagaaagacg  aagagctttt  cttagatgtcc  tttttttttgctg  cgggtatcc  6120
gctgctgcct  aacaaaaaaac  ccaacgcctac  cagcggtgtct  tggtggcggc  gtaaagacgc  6180	taaccactct  ttccccgaga  gtaaagctgt  tcgaagagcc  gaagatatcaca  aataactgctc  6240
ttactgattg  gctgattgtc  ggcacacctc  tcaaaacatc  tgtacgacctg  ctaacataacc  6300
tgctctgtct  aactccggtta  ccagtggtctg  ctggaagactg  gataaagttg  tgtttttgcg  6360
gcttgacgct  angaagacag  ttaaagcgtgta  aggggagcgg  gggcggtgtgca  acgggggggtt  6420
cgcacacacc  gcggccgtg  gacgcgaaac  ctcgacacccg  acgtgacagtc  taacagctgcg  6480
agctgttgaga  aagcgccacgc  cttccccgag  ggtgaaagacg  ggcagctgtc  cggtaaagcc  6540
gcggagctggc  aacagcgaggg  cgccgctcccag  cccagcggctg  ggggaagccgc  tggtatatctt  6600
atcgtccgct  cggttttccgc  caaccttgac  ttgagcggcgt  attttgtgaga  tgtgcatgcg  6660
gggggcgcgg  cctatggtgac  aaccgcacccg  agcggtctttt  tttacggttct  ttgctctttct  6720
gctggttttt  tgtcattcaat  ttctcttccttg  gcttatcccgc  tcgctctctgcgtg  gtaaagacgta  6780	taacgctctc  taagtctggtc  gataaaggcc  gcggacagcgc  aacgaagcgg  caagaagctg  6840
cgcgaagcga  gcggcggtaa  gcgcgacaccc  taacacccgc  gcgcctccccc  gcgcgtggc  6900
cgtacacccgatcgggtgct  gcag  6924

<210> SEQ ID NO 10
<211> LENGTH: 6924
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence; PLASMID
<400> SEQUENCE: 10

```
5' gggagggggg gggagggggg gggagggggg gggagggggg gggagggggg gggagggggg gggagggggg gggagggggg gggagggggg gggagggggg
3' tcgggtggtag tcgggtggtag tcgggtggtag tcgggtggtag tcgggtggtag tcgggtggtag tcgggtggtag tcgggtggtag tcgggtggtag tcgggtggtag
```

120
180
240
300
360
420
480
540
600
660
720
780
840
900
960
1020
1080
1140
1200
1260
1320
1380
1440
1500
1560
1620
1680
1740
1800
1860
1920
1980
2040
2100

37 May 22, 2003
-continued

gattaggtg atggttcacc tagtggcaca tgcocctgat agacgttttt ttgocctttag 4440
agttgaggt cccagtttct taatagttga ctctggacct aaactggaac aaactcacoa 4500
catatccgc tgttattttc ggctttttcg gatccatcga ctatactgta 4560
anaggtacg tgaattcaca aaaaatttac agcagctttc aaataaatg aacgggtaaa 4620
attctctgat gcggtatatt ctcctctacgc acctgctgag tattcttccag cgcatactggt 4680
gcacttcac taagatctgtc ttcgacattg catgatttaag ccaggcggcga cccggccgaa 4740
cacggcttcg cacggcctgc ggggtgttgc gctgccagtg aagccaagctg 4800
tgctcgcttc cggtgcgcttg aagcccattt ggtttttnaac gcacatccag aagcgcgcgga 4860
gacgacatg ccccctgtaga gctgattttc tagcgtttttc ttagcagttat ataatgctttt 4920
tgctgggacct ttaggtctttc gtttgggttc gcacagacag gctggtgaa gcacagagtt 4980
ttgaagacaa ggggagatag gtagctatac ttcctctaacgctgcctctt actctctttc 5040
atatctgtctc gtaagacag aagctgaagtt aacaatcagtc ctctctcttct tctctctttt 5100
ctgcaagtgt gttcagctgc ggctgggttc aacgaagcgc cagcgtggct gccagcatgc 5160
tctgcatatg gtagaaggct ccagagcaag cagggcagat gtaagagcat gcagcataagc 5220
tctcttttag ggttttgtctgc aagagcgtgt gtagactgta gataagtcag 5280
atctctgtct ggtgctgtgt tggcagcttg cagcagcttg gcagctttat tctctctcttt 5340
agcagagactg gtagatgtat ttcgctaaag cggcgaagcc cagagcagat gtaagagcat 5400
tagcagatc agaagctgcg gtagctgtgt ctctctctct ctctctcttt ctctctctct 5460
agatcagctg ggttctgtct ctcagctctc ggaagccttt tcccatactg ccatcgtgatg 5520
aggtccagtg cggctgtgtt tggctgtttgc ggttgttttt ccccaacctg 5580
agagctgtag cccacagctgc cgggtgacag tattctctct ggtgctgttt gcacagctgtc 5640
agttgatcag tgaacctgcttt aaaaactata ttagtttttt tcacatccag 5700
tgctgggacgc acctgctgacg tggcctttctc gcgtgtgttg gtttatcg catataccttg 5760
agacagcgtgc ggtgctgcttt cgcagctctg tcagagcaag ctggcagctgc gcagagcagat 5820
tgctgctgct gtctgttacgc ccagagctgtc ccagagctgtc ccagagctgtc 5880
agatcagctg gtagctgtgt tggctgtttgc ggttgttttt ccccaacctg 5940
catatactgc ttagatgctg tcctcttgttt aaaaaaacttc attattttat tcatactgctc 6000
tctctctctgc cttcagtgac tctctctctgc gtttcttttt ccccaacctg 6060
cagacccgctg agaagagcag ccaccagatttt ctctctctct cttctctctct cctctctctct 6120
gtggctgggt gcacacatc cccgaagtgc ttgcttttttc gctcttttttc gctcttttttc 6180
tacactgaat ttttttttt gcagagcagat ttcacagagc gctcttttttc gctcttttttc 6240
ttttagtaag tcagcttctgc cacagctgtc ccagagctgtc ccagagctgtc 6300
tgctgttacctgctgatg cacagctgtc ccagagctgtc ccagagctgtc 6360
aggtcagctgc ggggtgctgggc ggttgctggtgc gcagcagctgtc ccagagctgtc 6420
cgcacggagag ctgctgtgtc aggctgttacgc cctctgttacgc ccagagctgtc 6480
aggtcagctgc ggggtgctgggc ggttgctggtgc gcagcagctgtc ccagagctgtc 6540
gggtgctggtgc gcagcagctgtc ccagagctgtc ccagagctgtc ccagagctgtc 6600
atagctgtgt gcgtgtgtgct cacgctgtc cttcagtgac tctctctctgc 6660
gaggagggag cctaggaaga aacgccagca acgccgcttt tttaggtgcc ctcggtcttt 6720
gtggccttttt tggcataacttg ttcatctctctg gttatatctgt gattacgtaga 6790
ttacccgcttg aagctaggtg gctacgtgcc ggcagccgag cagcagcagt 6840
cagtgacgga gagaagccgaa gagaccccaaa tgcctccccccc ggcagtgggc 6900
cgatatctta atccagggct gcag 6924

<210> SEQ ID NO: 11
<211> LENGTH: 6924
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: PLASMID pRhIME1106-AT

<400> SEQUENCE: 11

gagggggggg gaggagggagc ggcctgtggt gctgctgcttc tccctgccggc actgtgctgcg 60
gggccggaga ggcctgtggt gcgtctgtgcg ctatccagctg gcgtgtgagt gcggagagag 120
cggagagagc gcgtgtgagt gccggggcctc agagaagcttc ctatccagctg gcgtgtgagt 180
tgggtgctgg ggctgagtcg gctgctgcttc tccctgccggc actgtgctgcg 240
tccatccactt ccgtggtggg tgaagtcgcc cccctgcctc ggccctgccggc 300
coccagctgcc tccctgccggc tccctgccggc gctatcggcc atggccgctgcg ctatccagctg 360
tgggctgagtcg gctgctgcttc tccctgccggc actgtgctgcg 420
tccatccactt ccgtggtggg tgaagtcgcc cccctgcctc ggccctgccggc 480
tccctgccggc tccctgccggc gctatcggcc atggccgctgcg ctatccagctg 540
tccatccactt ccgtggtggg tgaagtcgcc cccctgcctc ggccctgccggc 600
tccctgccggc tccctgccggc gctatcggcc atggccgctgcg ctatccagctg 660
tccctgccggc tccctgccggc gctatcggcc atggccgctgcg ctatccagctg 720
tccctgccggc tccctgccggc gctatcggcc atggccgctgcg ctatccagctg 780
tccctgccggc tccctgccggc gctatcggcc atggccgctgcg ctatccagctg 840
tccctgccggc tccctgccggc gctatcggcc atggccgctgcg ctatccagctg 900
tccctgccggc tccctgccggc gctatcggcc atggccgctgcg ctatccagctg 960
tccctgccggc tccctgccggc gctatcggcc atggccgctgcg ctatccagctg 1020
tccctgccggc tccctgccggc gctatcggcc atggccgctgcg ctatccagctg 1080
tccctgccggc tccctgccggc gctatcggcc atggccgctgcg ctatccagctg 1140
tccctgccggc tccctgccggc gctatcggcc atggccgctgcg ctatccagctg 1200
tccctgccggc tccctgccggc gctatcggcc atggccgctgcg ctatccagctg 1260
tccctgccggc tccctgccggc gctatcggcc atggccgctgcg ctatccagctg 1320
tccctgccggc tccctgccggc gctatcggcc atggccgctgcg ctatccagctg 1380
tccctgccggc tccctgccggc gctatcggcc atggccgctgcg ctatccagctg 1440
tccctgccggc tccctgccggc gctatcggcc atggccgctgcg ctatccagctg 1500
tccctgccggc tccctgccggc gctatcggcc atggccgctgcg ctatccagctg 1560
tccctgccggc tccctgccggc gctatcggcc atggccgctgcg ctatccagctg 1620
tccctgccggc tccctgccggc gctatcggcc atggccgctgcg ctatccagctg 1680
ggccogggc ccggggtggg gcggcttgac gcggcgttgg cggggcgtgg cggtggcgggc 1740
gggggggtgc gccggcgggc gcgggtgccc gcggcgtggg cgggggggtgc 1900
tgggggggg ccggccggcc ccccccggcc gcggcgcggc ggccccgggct ggccgcgggc 1860
gcggcgtgcc tttttttggt atctgctgggcc aggggcaagg cccgtcctctt ctgcccacatt 1920
tgtggggacccc ccaggcgtgg ccggcgcggc cggccccccc ctcagggggc ggggggaggggc 1980
cggcgttgg cccgggccc gccccggccc gggggggccc ttcgggtggt cccggcgggc 2040
cgtoctcctc tcctctcctg gcggcggggtt gcgtcgccgg gggacggctgc ccccttggggc 2100
gccggggtcc cggcgggtgg cggttttttc gcgtggtcgg gcggcgctgc gcgggtctgcgct 2160
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 2220
cgtggctcct cttttttgca aacggtgctg aaactcgctgt gggtattttccc gcggcggccc 2280
ccggcgtgg gcgggtcggc gggacggcgg ccccttgc gcgtcgcgtcgc gttctttggtg 2340
gccgggcttc gcgggcttc ctggggtcgc ctgcgtggcg ccccttggccc cccgtcctctt 2400
cggccccgg ccagcggacc ccccttgcgc gcgtgcggcct gccgttggg gttcggccccc 2460
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 2520
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 2580
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 2640
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 2700
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 2760
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 2820
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 2880
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 2940
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 3000
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 3060
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 3120
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 3180
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 3240
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 3300
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 3360
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 3420
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 3480
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 3540
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 3600
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 3660
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 3720
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 3780
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 3840
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 3900
acagagctt atctctctgc cttttttctg acagctctcg gcggcccggg gcggggtgtgtg 3960
We claim:

1. A method for providing an animal with a therapeutically effective amount of a serum protein, said method comprising introducing into cells of said animal an effective amount of viral particles or vector, wherein said viral particles or viral vector comprises a polynucleotide encoding said protein.

2. The method according to claim 1, wherein said animal is a mammal.

3. The method according to claim 2, wherein said mammal is a human.

4. The method according to claim 1, wherein said vector is an adenovirus-associated virus vector.

5. The method according to claim 1, wherein said vector comprises a promoter sequence capable of driving expression of said polynucleotide encoding said protein.

6. The method according to claim 5, wherein said promoter sequence is selected from the group consisting of CMV promoter sequences, hybrid CMV enhancer/β-actin promoter sequences, EF1 promoter sequences, U1a promoter sequences and U1b promoter sequences.

7. The method according to claim 5, wherein said promoter sequence is an inducible promoter selected from the group consisting of Tet-inducible promoters and VP16-LexA promoter.

8. The method according to claim 5, wherein said vector further comprises an enhancer sequence.

9. The method according to claim 8, wherein said enhancer is a synthetic enhancer.

10. The method according to claim 1, wherein said animal has a condition that results in a defective protein or a deficiency of said protein encoded by said polynucleotide.

11. The method according to claim 1, wherein said animal has a condition that can be ameliorated or treated by said protein encoded by said polynucleotide.

12. The method according to claim 1, wherein said protein encoded by said polynucleotide is selected from the group consisting of anti-proteases, enzymes, structural proteins,
coagulase factors, interleukins, cytokines, growth factors, interferons, and lymphokines.

13. The method according to claim 1, wherein said cells are myofibers, myoblasts, hepatocytes, or lung cells.

14. The method according to claim 1, wherein said polynucleotide encodes human alpha-1-antitrypsin protein, or a biologically active fragment or variant thereof.

15. The method according to claim 4, wherein said polynucleotide encodes human alpha-1-antitrypsin protein, or a biologically active fragment or variant thereof.

16. The method according to claim 1, wherein said viral particles are introduced into said cells or tissue by infection or injection.

17. The method according to claim 1, wherein said vector is introduced into said cells by transfection or injection.

18. The method according to claim 1, wherein said viral particles or vector is introduced into said cells in vitro and said treated cells are introduced into said animal.

19. The method according to claim 1, wherein said viral particles or vector is introduced into said cells in vivo.

20. The method according to claim 19, wherein said viral particles or vector is injected into muscle.

21. The method according to claim 19, wherein said viral particles or vector is injected into portal or peripheral vein.

22. The method according to claim 19, wherein said viral particles or vector is injected intrahepatically or inhaled into the lungs.

23. The method according to claim 15, wherein said vector is selected from the group consisting of dE-AT, E-AT, C-AT, C-AT2, p43C-AT, p43CB-AT, p43C-AT-IN, p43msENC-AT, p43msENC-AT, p43msENC-AT, and p43msENC-AT.

24. A recombinant viral vector comprising a polynucleotide encoding a protein capable of providing a therapeutic effect to an animal when expressed in said animal.

25. The vector according to claim 24, wherein said animal is a mammal.

26. The vector according to claim 25, wherein said mammal is a human.

27. The vector according to claim 26, wherein said vector is an adeno-associated virus vector.

28. The vector according to claim 24, wherein said vector comprises a promoter sequence capable of driving expression of said polynucleotide encoding said protein.

29. The vector according to claim 28, wherein said promoter sequence is selected from the group consisting of CMV promoter sequences, hybrid CMV enhancer/β-actin promoter sequences, EF1 promoter sequences, U1a promoter sequences and U1b promoter sequences.

30. The vector according to claim 24, wherein said polynucleotide encodes human alpha-1-antitrypsin protein, or a biologically active fragment or variant thereof.

31. The vector according to claim 27, wherein said polynucleotide encodes human alpha-1-antitrypsin protein, or a biologically active fragment or variant thereof.

32. The vector according to claim 31, wherein said vector is selected from the group consisting of dE-AT, E-AT, C-AT, C-AT2, p43C-AT, p43CB-AT, p43C-AT-IN, p43msENC-AT, p43msENC-AT, p43msENC-AT, and p43msENC-AT.

33. The vector according to claim 28, wherein said promoter sequence is an inducible promoter selected from the group consisting of Tet-inducible promoters and VP16-LexA promoter.

34. The vector according to claim 28, wherein said vector further comprises an enhancer sequence.

35. The vector according to claim 34, wherein said enhancer is a synthetic enhancer.

36. The vector according to claim 24, wherein said protein encoded by said polynucleotide is selected from the group consisting of anti-proteases, enzymes, structural proteins, coagulase factors, interleukins, cytokines, growth factors, interferons, and lymphokines.

37. A viral particle comprising the vector of claim 24.

38. A cell comprising the vector of claim 24.

39. The cell according to claim 38, wherein said cell is a myofiber, myoblast, hepatocyte, or lung cell.

40. A method for treating alpha-1-antitrypsin deficiency in an animal, said method comprising introducing into cells of said animal a vector according to claim 24, wherein said polynucleotide of said vector encodes alpha-1-antitrypsin protein, or a biologically active fragment or variant thereof.