



US 20190194295A1

(19) **United States**

(12) **Patent Application Publication**

Falkner et al.

(10) **Pub. No.: US 2019/0194295 A1**

(43) **Pub. Date:** **Jun. 27, 2019**

(54) **VIRAL VECTORS ENCODING
RECOMBINANT FVIII VARIANTS WITH
INCREASED EXPRESSION FOR GENE
THERAPY OF HEMOPHILIA A**

(71) Applicants: **Baxalta Incorporated**, Bannockburn, IL (US); **Baxalta GmbH**, Zug (CH)

(72) Inventors: **Falko-Günter Falkner**, Orth/Donau (AT); **Franziska Horling**, Gaenserndorf (AT); **Johannes Lengler**, Vienna (AT); **Hanspeter Rottensteiner**, Vienna (AT); **Friedrich Scheiflinger**, Vienna (AT)

(21) Appl. No.: **16/211,202**

(22) Filed: **Dec. 5, 2018**

Related U.S. Application Data

(63) Continuation of application No. 15/349,930, filed on Nov. 11, 2016, now Pat. No. 10,189,888.

(60) Provisional application No. 62/255,317, filed on Nov. 13, 2015.

Publication Classification

(51) **Int. Cl.**

C07K 14/755 (2006.01)
A61K 48/00 (2006.01)
C12N 15/86 (2006.01)

(52) **U.S. Cl.**

CPC **C07K 14/755** (2013.01); **A61K 48/0066** (2013.01); **A61K 48/0058** (2013.01); **C12N 2800/22** (2013.01); **C12N 15/86** (2013.01); **C12N 2840/007** (2013.01); **C12N 2750/14143** (2013.01); **A61K 48/0008** (2013.01)

(57) **ABSTRACT**

The present disclosure provides, among other aspects, codon-altered polynucleotides encoding Factor VIII variants for expression in mammalian cells. In some embodiments, the disclosure also provides mammalian gene therapy vectors and methods for treating hemophilia A.

Specification includes a Sequence Listing.

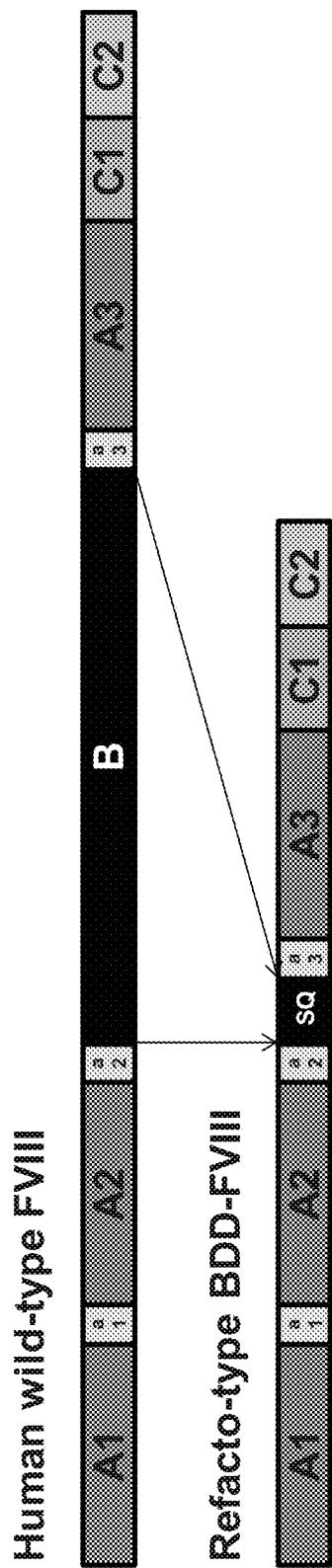


Figure 1

CS04-FL-NA

atgcagattgagctgagcacctgcttcttctgtgcctgtcggatgttttcgtccaccaggaa
gatactacotggggctgtggagcttcttggactacatgcagtctgacccctggggagctgcctgt
ggatgccaggttcccacccagagtgcaccaatccttcccatcaacacacccatgtggatacaagaag
accctttgtggagttcaactgaccacctgttcaacattgccaaccaggccaccctggatggac
tcctggaccaccattcaggctgagggttatgacactgtgtcatcaccctcaagaacatggcctc
ccaccctgtgaggcctgcattgtggggctcagactactggaaaggcctgtggggctgaggatgtatgat
gaccagaccccccagaggagaaggaggatgacaaagtgttccctggggcagccacacccatgtgt
ggcaggtcctaaggagaatggccccatggcctgtggccactctgcctgacccactccttacccat
tcatgtggacctggtaaggacacttgcactgtggactgattggggccctgtggatgtgcaggaggc
tccctggccaaagagaagacccagaccctgcacaagttcattctctgtttgtctttgtatgagg
gcaagagctggcacttgcaccaagaactccctgatgcaggacaggatgtgcctctggcaggc
ctggcccaagatgcacactgtgaatggctatgtgaacaggagcctgtggactcattggctgccc
agggaaatctgtctactggcatgtgatggcatgggacaaccctgagggtgcactccatttctgg
agggccacacccctoactggcaggaaccacacagacaggccagctggagatcagcccatcactt
caactggccagaccctgctgtgatggacccctcgagttcctgtgttctggccacatcagctccacc
catgtggcatggaggcctatgtcaaggtggacagctgcctgaggagccacagctcaggatgaaga
acaataggaggctgaggactatgtatgacactgtgactctgagatggatgtggccctttga
tgatgacaacagcccatccttcattcagatcaggctgtggccaagaaaacaccccaagacccatgg
caactacattgtgtgaggaggactggactatggccacttgcctgtgttctggccatcagatc
gctacaagagccagttacctaacaatggccacagaggatggacgcagaactacaagaaagtca
catggccatcactgtgatgaaacottcaagaccaggaggccattcagcatgagtctggcatc
ccactcctgtatgggaggtggggacaccctgtcatcatcttcaagaaccaggccctcaggcc
acaacatctaccacatggcatcaactgtgatgtcaggccctgtacagccgcaggctgcca
gaaacacotcaaggacttccccatttcgcctggggagatcttcaagtacaatggactgtcact
gaggatggacccaaccaaactgtgacccctggctcaccagataactacttccagcttgc
agaggacccctggctctggctgattggccactgtcatctgttcaaggatgtggacttgc
ggaaaccaggatcatgtgtgacaagaggatgtgattctgttgc
tggtaacctgactgagaacattcagcgttctggccaaaccctgtgggtgcagctggaggacc
agttccaggccagcaacatcatgcactccatcaatggctatgtgttgcacagccctc
ctgcctgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgcatttgc
ctggggagactgtgttcatgagcatggagaaccctggctgtggattctggatgc
cttccgcaacaggggcatgactgccctgtcaaaactgttccctgtgacaaga
tatgaggacagctatgaggacatctgcctactgtc
gottcagccagaatccacccatttcgcctgaaacgccaccaggaggatc
tgaccaggaggatgtactatgtgacaccatttgcatttgc
tatgacgaggacagaaccaggccaaaggagcttcc
ctgtggactatggcatgagctcc
tggctctgtgc
ctgtacaggaggactgtgatgagcacccctgg
acaacatcatgg
ctatgagGaggaccagaggc
acccat
(Continued)

Figure 2A

cctacttctctgatgtggacctggagaaggatgtcactctggcctgattggccactctggctcg
ccacaccaacaccctgaaccctgcccatggaaggcaagtgactgtgcaggagttgcctcttc
accatcttgatgaaaccaagagactggtacttcaactgagaacatggagcgcaactgcaggccccat
gcaacattcagatggaggaccccaccccaaagagaactaccgcctccatgccatcaatggctacat
catggacaccctgcctgggcttgtcatggccaggaccaggatcagggtgtacctgccttctatg
ggctccaatgagaacattcaactccatccacttctggcatgtcttcaactgtgcgcagaaggagg
agtacaagatgcccgttacaaccttacccctgggtctttgagactgtggagatgtgcctccaa
agctggcatctgggggtggagtgccttgcattggggagcacctgcattgtggcatgaggcaccctgttc
ctggctacagaacaactgcccaggccccctggaatggcctctggccacatcaggactccaga
tcactgcctctggccagtatggccagtggggcccaagctggccaggctccactactgttatcat
caatgcctggagcaccaaggagccattcagctggatcaaagtggacotgtggccccatgatcatc
catgcattcaagaccaggccccccaggcagaagtgcctccagctgtacatcaggcacttcatcatca
tgtacagcctggatggcaagaaatggcagacctacagaggcaactccacttggactcatgtt
ctttggcaatgtggacagctggcatcaagcacaacatcttcaaccccccataatcgccagatac
atcaggctgcacccaccactacagcatccgcagcacccctcaggatggagctgtatggctgtgacc
tgaactcctgcagcatgcctggcatggagagcaaggccatttctgtatgcctccagatcactgcctc
cagctacttaccaacatgttgcacccctggagcccaagcaaggccaggctgcacccctccaggaaagg
agcaatgcctggagggcccaaggtaacaaacccaaaggagtggtgcagggtggacttccagaagacca
tgaaggactgggtgaccaccaggccccgggtcaagagcctgcaccaggcatgtatgtgaaggagtt
cctgatcagctccagccaggatggccaccaggctggacccttotcttccagaatggcaaggtcaagggt
ttccaggccaaaccaggacagcttacCcctgtggtaacagcctggaccccccctctgaccaggat
acctgaggattaccccccagagactgggtccaccaggattgcctgaggatggaggcttggatgtga
ggccaggacctgtactga (SEQ ID NO:1)

Figure 2B

CS04-FL-AA

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDA' FPPRVPKSFPFNTSVVYK
KTLFVEFTDHALFNIAKPRPPWMGLLGPTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKASEGAEY
DDQTSQREKEDDKVFPGGSHTYVWQVLKENGPMA SDPLCLTYSYLSHVDLVKDLNSGLIGALLVCRE
GSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGC
HRKSVYWHVIGMGTPEVHSIFLEGHFLVRNHRQASLEISPITFLTAQTLLMDLGQFLFCHISSH
QHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLDSEMDVVRFDDDNSPSFIQIRSVAKKHPKTW
VHYIAAEEEWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQHESGIL
GPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSSRLPKGVKHLKDFPILPGEIFKYKWTVT
VEDGPTKSDPRCLTRYSSFVNMERDLASGLIGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENR
SWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDLQLSVCLHEVAYWYIISIGAQTDPLS
VFFSGYTFKHKMVYEDTLLFPFSGETVFMMSMENPGLWILGCHNSDFRNRRGMTALLKVSSCDKNTGD
YYEDSYEDIISAYLLSKNNIAIEPRSFQSQQNPVLRHQREITRTTLQSDQEEIDYDDTISVEMKKEDFD
IYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSPHVLRNRAQSGSVPQFKVVFQEFTDGSFTQ
PLYRGELNEHLGLLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNET
KTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLEKDVSGLIGPLLVCHTNTLNPAHGRQVTQEFALF
FTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLS
MGSNENIHSIHSGHVFTRKKEEYKMALYNLYPGVFETVEMLPSKAGIWRVECLIGEHLHAGMSTL
FLVYSNKCQTPLGMASGHI RDFQITASGQYQQWAPKLARLHYSGSINAWSTKEPFSWIKVVDLLAPMI
IHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPIIIAR
YIRLHPTHYSIRSTLREMELMGCDLNSCSMPLGMEKAISDAQITASSYFTNMFAWTWSPSKARLHLQG
RSNAWRPQVNNPKEWLQDFQKTMKVTVGTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVK
VFQGNQDSFTPVNNSLDPPLLTRYLRIHPQS梧HQIALRMEVLGCEAQDLY (SEQ ID NO:2)

Figure 3

CS04-HC-NA

ggc
accaggagat actacacctggg ggctgtggag ctttcttggg actacatgc gtctgacact
ggggagactgc ctgtggatgc caggccccca cccagagactgc ccaaattcctt cccattcaac
acacctgtgg tctacaagaa gaccctttt gtggagttca ctgaccacact gttcaacatt
gccaaaccca ggcacccctg gatggactc ctgggaccca ccattcaggg tgaggtgtat
gacactgtgg tcataccctt caagaacatg qcctcccccacc ctgtgagccct gcatgctgtg
ggggtcagct actggaaaggc ctctgagggg gctgagttatg atgaccagac ctcccagagg
gagaaggagg atgacaaagt gttccctggg ggcagccaca cctatgtgtg gcaggtccctc
aaggagaatg gccccatggc ctctgaccca ctctgcctga cctactccta cctttctcat
gtggacctgg tcaaggacct caactctggc ctgattgggg ccctgtgtgt gtgcagggag
ggctccctgg ccaaagagaa gaccaggacc ctgcacaatg tcatttctt gttgtgtc
tttgcgtgagg gcaagagctg gcactctgaa accaagaact cctgtatgc ggacagggat
gctgcctctg ccaggccctg gcccaagatg cacactgtga atggctatgt gaacaggagc
ctgcctggac tcattggctg ccacagaaaa tctgttact ggcatgtat tggcatgggg
acaaccctg aggtgcacte cattttctg qagggccaca ctttctgtt caggaaccac
agacaggcca gcttggagat cagccccatc accttctca ctgcccagac cctgtgtatg
gacctcggac agttctgtt gttctgccc atcagctccc accagcatga tggcatggag
gcttatgtca aggtggacag ctgccttgag gagccacacgc tcaggatgaa gaacaatgag
gaggctgagg actatgtatg tgacctgact gactctgaga tggatgtgtt cctgtttgt
gtatgacaaca gcccattctt cattcagatc aggtctgtgg ccaagaaaaca ccccaagacc
tgggtgcact acattgtctg tgaggaggag gactggact atgccccact ggtctggcc
cctgtatgaca ggagctacaa gagccagttac ctcaacaatg gcccacagag gattggacgc
aagtacaaga aagttaggtt catggctac actgtatgaaa ctttcaagac cagggaggcc
attcagcatg agtctggcat cttggccca ctccctgtatg gggaggtggg ggacaccctg
ctcatcatct tcaagaacca ggcctcoagg ccctacaaca totaccacata tggcatact
gtatgcaggc ccctgtacag ccgcagctg cccaaaggggg taaaacacact caaggacttc
cccattctgc ctggggagat cttcaagtac aagtggactg tcactgtgg ggtggacca
accaaaatctg accccaggtg ctcaccaga tactactcca gctttgtgaa catggagagg
gacctggcct ctggcctgtat gggccactg ctcatctgtt acaaggagtc tggaccag
agggggaaacc agatcatgtc tgacaagagg aatgtgattc tggatctctgt ctgtatgag
aacaggagct ggtacctgac tgagaacatt cagcgcttc tgcccaaccc tgctgggggt
cagctggagg accctgagtt ccaggccaggc aacatcatgc actccatcaa tggctatgt
tttgcacagcc tccagtttc tgtctgctg catgaggtgg cctactggta catttttt
atggggggcc agactgactt ctttctgtc ttcttctctg gotacacccctt caaacacaag
atgggtgtatg aggacaccct gaccctttc ccattctctg gggagactgt gttcatgagc
atggagaacc ctggcctgtg gattctgggta tggccacaact ctgactccg caacaggggc
atgactgccc tgcctcaaatg ctctctctgt gacaagaaca ctggggacta ctatgaggac
agctatgagg acatctctgc ctacctgtc agcaagaaca atgcccattga gcccagg
(SEQ ID NO:3)

Figure 4

CS04-LC-NA

g agatcaccag gaccaccctc
cagtctgacc aggaggagat tgactatgtat gacaccattt ctgtggagat gaagaaagag
gactttgaca tctatgacga ggacgagaac cagagccaa ggagcttcca gaagaagacc
aggcactact tcattgtgc tgtggagcgc ctgtggact atggcatgag ctccagcccc
catgtctca ggaacagggc ccagtctggc tctgtgccac agttcaagaa agtgttctc
caagagttca ctgatggcag cttcacccag cccctgtaca gagggggagct gaatgagcac
ctgggactcc tggccctata catcagggtct gaggtggagg acaacatcat ggtgaccttc
cgcaaccagg cctccaggcc ctacagcttc tacagotccc tcatcagcta tgaggaggac
cagaggcagg gggctgagcc acgcaagaac tttgtgaaac ccaatgaaac caagacctac
ttctggaaag tccagcacca catggccccc accaaggatg agtttactg caaggcctgg
gcctacttct ctgatgtgga cctggagaag gatgtgcact ctggcctgat tggccactc
ctggctctgcc acaccaaacac octgaaccct gccatggaa ggcgaagtgao tggcaggag
tttgcctct tcttcacccat ctttgatgaa accaagagct ggtacttcac tgagaacatg
gagcgcact gcagggccccc atgcaacatt cagatggagg accccacctt caaagagaac
taccqcttcc atgccatcaa tggctacatc atggacacccc tgcctggct tgtcatggcc
caggaccaga ggatcaggtg gtacctgctt tctatggct ccaatgagaa cattcactcc
atccacttct ctgggcatgt cttcaactgt cgcaagaagg aggactacaa gatggccctg
tacaacctct accctgggt ctttgagact gtggagatgc tgcctccaa agctggcattc
tggagggtgg agtgccatc tggggagcac ctgcattgtg gcatgagcac cctgttccctg
gtctacagca acaagtgcca gaccccccgt ggaatggct ctggccacat cagggacttc
cagatcactg ccctggcca gtagggccag tggccccc agctggccag gttccactac
tctggatcca tcaatgcctg gagcaccaag gagccattca gctggatcaa agtggacctg
ctggccccc tgcattca tggcatcaag acccaggggg ccaggcagaa gttctccage
ctgtacatca gccagttcat catcatgtac agcctggatg gcaagaaatg gcagacctac
agaggcaact ccactggaaac actcatggtc ttctttggca atgtggacag ctctggcattc
aagacaaca tcttcaaccc cccaaatcatc gccagatata tcaggctgca ccccaacccac
tacagcatcc gcagcacccct caggatggag ctgatggct gtgacctgaa ctctgcage
atgcccctgg gcatggagag caaggccatt tctgatggcc agatcactgc ctccagctac
ttcaccaaca tggggccac ctggagccca agcaaggcca ggctgcaccc ccagggaaagg
agcaatgcct ggaggccccc ggtcaacaac cccaaaggagt ggctgcaggt ggacttccag
aagaccatga aggtcaactgg ggtgaccacc cagggggtca agagcctgct caccagcatg
tatgtgaagg agttctgtat cagtcctcagc caggatggcc accagtgaa cctcttcttc
cagaatggca aggtcaaggt gttccaggcc aaccaggaca gtttcacccct tgggtgaac
agcctggacc cccccccttcc gaccagatac ctgaggattc acccccaagag ctgggtccac
cagattggcc tgaggatgga ggtcctggga tgtgaggccc aggacctgtac
(SEQ ID NO:4)

Figure 5

EDL001 - agc ttcttcaga atccacatgt cttggaaaga caccagaga (SEQ ID NO:5)
EDL004 - agc ttcaaggaa atccacatgt cttggaaacgc caccagagg (SEQ ID NO:6)
EDL023 - agc ttcaaggaa acccccccggt gttggaaagg caccagagg (SEQ ID NO:7)
EDLING1 - agtttcaaggaaatgttggaaacaataatgttggaaatccacatgttggaaacgcaccagg (SEQ ID NO:36)
EDLING4 - agtttcaaggaaatgttggaaacaataatgttggaaatccacatgttggaaacgcaccagg (SEQ ID NO:37)
EDLING5 - agtttcaaggaaatgttggaaacaataatgttggaaacgcaccagg (SEQ ID NO:38)
EDLING6 - agtttcaaggaaatgttggaaataatccacatgttggaaacgcaccagg (SEQ ID NO:39)
EDLING9 - agtttcaaggaaatgttggaaatccacatgttggaaacgcaccagg (SEQ ID NO:40)
EDLING10 - agtttcaaggaaatgttggaaatccacatgttggaaacgcaccagg (SEQ ID NO:41)
EDLING16 - agtttcaaggaaatgttggaaatccacatgttggaaacgcaccagg (SEQ ID NO:42)
EDLING17 - agtttcaaggaaatgttggaaatccacatgttggaaacgcaccagg (SEQ ID NO:43)
EDLING18 - agtttcaaggaaatgttggaaatccacatgttggaaacgcaccagg (SEQ ID NO:44)
EDLING19 - agtttcaaggaaatccacatgttggaaatccacatgttggaaacgcaccagg (SEQ ID NO:45)
EDLING20 - agtttcaaggaaatccacatgttggaaacgcaccagg (SEQ ID NO:46)
EDLING21 - agtttcaaggaaatccacatgttggaaacgcaccagg (SEQ ID NO:47)
EDLINGV - agtttcaaggaaatgttggaaatccacatgttggaaacgcaccagg (SEQ ID NO:48)

Figure 6

CS04-AV-NA

1 tcgcgcgtt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggta
 61 cagcttgbt gtaagcggat gcccggagca gacaaggcccg tcagggcgcg tcaggggtg
 121 ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgtt ctgagatgc
 181 accatatgcg gtgtgaaata ccgcacagat gcgtaaaggag aaaataccgc atcaggcgcc
 241 attcgccatt caggctgcgc aactgttggg aaggcgatc ggtgcgggccc tcttcgttat
 301 tacgccagct ggcgaaagggg ggatgtgtc caaggcgatt aagtgggtt acggcagggt
 361 ttcccaggc acgacgttgt aaaacgcacgg ccagtgaatt cctcgagatt taaatgacgt
 421 tggccactcc ctctctgcgc gctcgctcgc tcactgaggc cggcgacca aaggtcgccc
 481 gacgcccggg ctggcccccgg gcccccttag tgagcgagcg agcgcgcaga gagggagtg
 541 ccaactccat cactaggggt tcttgaggtt aaacttcgtc gacgattcga qcttgggt
 601 caggtcgagg gcactgggag gatgtttagt aagatggaaa actactgtat accottgcag
 661 agacagagta ttaggacatg tttgaacagg ggccgggcga tcagcaggta getctagagg
 721 atccccgtct gtctgcacat ttctgttagc gatgttccg atactctaatt ctcccttaggc
 781 aagggtcata tttgtgttagg ttacttattt tcctttgtt gactaagtca ataattcagaa
 841 tcagcagggtt tggagtcage ttggcagggg tcagcaggct ggggttggaaag gagggggtat
 901 aaaagccccc tcaccaggag aagccgtcac acagactagg cgcgcacccg ccaccatgca
 961 gattgagctg agcacctgtc tcttcctgtt cctgtcgagg ttctgtttt ctgccaccag
 1021 gagatactac ctgggggctg tggagcttc ttgggactac atgcagtcgt acctggggga
 1081 gctgcctgtg gatgccaggt toccaccagg agtgcacaaa tccttcccat tcaacacctc
 1141 tgtggtctac aagaagaccc tttttgttgg gttcactgtac cacctgttca acattgcca
 1201 acccaggcoca ccctggatgg gactcttggg acccaccatt caggctgagg tggatgacac
 1261 tgtggtcatc accctcaaga acatggcctc ccaccctgtg agcctgcatg ctgtgggggt
 1321 cagctactgg aaggcccttg agggggctga gtatgtatc acagacccc agagggagaa
 1381 ggaggatgac aaagtgttcc ctggggcag ccacacccat gtgtggcagg tcctcaagga
 1441 gaatggccccc atggcctgtg acccactctg cctgacctac tcctaccatt ctcatgttga
 1501 cctggtaag gacctaact ctggactgtat tggggccctg ctgggtgtca gggagggtc
 1561 cctggccaaa gagaagaccc agaccctgtca caagttcatt ctccctgttt ctgtctttga
 1621 tgaggcaag agctggcaact ctgaaaccaa gaactccctg atgcaggaca gggatgctgc
 1681 ctctgccagg gcctggccca agatgcacac tggatgttgc tatgtgaaca ggagcctgcc
 1741 tggacttatt ggctgccaca gggaaatctgt ctactggcat gtgtggca tggggacaac
 1801 ccctgagggtg cactccatt tcctggaggg ccacacccat ctggtcagga accacagaca
 1861 ggccagctg gagatcagcc ccatcacctt cctactgtc cagaccctgc tgatggaccc
 1921 cggacagttc ctgtgttct gccacatcag ctcccaccag catgtggca tggaggccta
 1981 tgtcaagggtg gacagctgcc ctgaggagcc acagctcagg atgaagaaca atgaggaggc
 2041 tgaggactat gatgtatgacc tgactgactc tgagatggat gtggccctgt ttgtatgt
 2101 caacagccca tccttcattt agatcagtc tggcccaag aaacacccca agacccctgg
 2161 gcactacatt gctgtgagg aggaggactg ggactatgcc ccactggtcc tggccctgt
 2221 tgacaggagc tacaagagcc agtacccatca caatggccca cagaggattt gacgcaagta
 2281 caagaaagtc aggttcatgg cctacactga tggaaacctt aagaccagggg agggccattca
 2341 gcatgagttt ggcatactgg gcccacttct gtatggggag gtgggggaca ccctgtcat
 2401 catcttcaag aaccaggccct ccaggcccta caacatctac ccacatggca tcactgtatgt
 2461 caggccccctg tacagccgca ggctgccaaa ggggtgaaa caccctcaagg acttccccat

Figure 7A

2521 tctgcctggg gagatottca agtacaagt gactgtcaact gtggaggatg gaccaaccaa
2581 atctgacccc aggtgcctca ccagatacta ctccagcttt gtgaacatgg agagggacct
2641 ggccctctggc ctgattggcc cactgctcat ctqctacaag qagtctgtgg accagagggg
2701 aaaccagatc atgtctgaca agaggaatgt gattctgttc totgtctttg atgagaacag
2761 gagctggta c tgactgaga acattcagcg cttccctgccc aaccctgtg ggggtcagct
2821 ggaggaccct gagttccagg ccagcaacat catgcaetcc atcaatggct atgtgtttga
2881 cagccctccag ctttctgtct gcctgcata ggtggectac tggtaacattc tttcttattgg
2941 gqcccagact gacttccctt ctgtcttctt ctctggctac accttcaaaac acaagatgg
3001 gtatgaggac accctgaccc tcttcccatt ctctggggag actgtgttca tgaggatgg
3061 gaaccctggc ctgtgatcc tggatgcca caactctgac ttccgcaaca ggggcatgac
3121 tgccctgctc aaagtctctt cctgtgacaa gaacactggg gactactatg aggacagcta
3181 tgaggacatc tctgcctacc tgctcagcaa gaacaatgcc attgagccca ggagcttcag
3241 ccagaatcca cctgtctgta aacgocacca gagggagatc accaggacca ccctccagtc
3301 tgaccaggag gagattgact atgatgacac catttctgtg gagatgaaga aagaggactt
3361 tgacatctat gacgaggacg agaaccagag cccaaaggagc ttccagaaga agaccaggca
3421 ctacttcatt gctgtgtgg aqgcctgtg ggactatggc atgagctcca gccccatgt
3481 cctcaggaac agggcccagt ctggotctgt gccacagttc aagaaaagtgg tcttccaaga
3541 gttcaactgat ggcagcttca cccagccct gtacagaggg gagctgaatg agcacctggg
3601 actoctgggg ccatacatca gggctgaggt ggaggacaac atcatggta ccttccgcaa
3661 ccaggccctcc aggcctaca gcttctacag ctccctcata agctatgagg aggaccagag
3721 gcagggggct gagccacgca agaactttgt gaaacccaat gaaaccaaga cctacttctg
3781 gaaagtccag caccacatgg ccccccaccaa ggatgagttt gactgcaagg cctgggccta
3841 cttctctgat gtggacctgg agaaggatgt gcaactctggc ctgattggcc cactccttgt
3901 ctgccacacc aacacctgta accctgccc tggaaaggca gtgactgtgc aggagttgc
3961 cctttcttc accatcttg atgaaaccaa gagctggta ttcactgaga acatggagcg
4021 caactgcagg gccccatgca acattcagat ggaggacccc accttcaaaag agaactaccg
4081 cttccatgcc atcaatggct acatcatgga caccctgcct gggcttgc tggccccagga
4141 ccagaggatc aggtggtaacc tgctttctat gggctccaat gagaacatcc actccatcca
4201 ctctctggg catgtttca ctgtgcgoaa gaaggaggag tacaagatgg ccctgtacaa
4261 cctctaccct ggggtctttg agactgtgga gatgtgccttcc tccaaagctg gcatctggag
4321 ggtggagtgc ctcatgggg aqcacctgca tgctggcatg agcaccctgt tcttggctta
4381 caqcaacaag tgccagaccc ccctggaaat ggcctctggc cacatcaggg acttccagat
4441 cactgcctct ggccagtatg gccagtggc ccccaagctg gccaggctcc actactctgg
4501 atccatcaat gcttgagca ccaaggagcc attcaagctgg atcaaagttt acctgctggc
4561 ccccatgatc atccatggca tcaagaccca gggggccagg cagaaggatctt ccagcctgtta
4621 catcagccag ttcatcatca ttttacagctt ggatggcaag aatggcaga cctacagagg
4681 caactccact ggaacactca tggtcttctt tggcaatgtg gacagctctg gcatcaagca
4741 caacatcttc aaccccccacca tcatacgccag atacatcagg ctgcacccca cccactacag
4801 catccgcagc accctcagga tggagctgat gggctgtgac ctgaactctt gcagcatgcc
4861 cctggcatg gagagcaagg ccatttctga tgcccaagatc actgccttca gctacttcc
4921 caacatgttt gccacctgga qcccaagcaa ggccaggctg cacctccagg gaaggagcaa
4981 tgcctggagg ccccaaggta acaacccaaa ggatggctg caggtggact tccagaagac

Figure 7B

5041 catgaaggtc actgggtga ccacccaggg ggtcaagagc ctgctcacca gcatgtatgt
5101 gaaggagttc ctgatcagct ccagccagga tggccaccag tggaccctct tttccagaa
5161 tggcaaggtc aagggtttcc agggcaacca ggacagcttc accccctgtgg tgaacagcct
5221 ggacccccc ctcctgacca gataacctgag gattcacccc cagagctggg tcaccagat
5281 tgccctgagg atggaggtcc tggatgtga ggcccaggac ctgtactgat gacgagcgcc
5341 cgctttagt agcagtatcg ataataaaag atctttatct tcattagato tttgtgttgg
5401 tttttgtgt gttaattaag ctcgcgaagg aacccttagt gatggagttg gcactccct
5461 ctctgcgcgc tcgctcgctc actgaggccg ggccgacaaa ggtcgcccgaa cccccgggct
5521 ttgcccggc ggcctcagtg agcgagcgag cgccgagaga gggagttggc aagacgattt
5581 aaatgacaag cttggcgtaa tcatggtcat agctgtttcc tttgtgaaat ttttatccgc
5641 tcacaatcc acacaacata cgagccggaa ycataaaagtg taaagcttg ggtgcctaatt
5701 gagtgagcta actcacatta attgogtgc gctcaactgcc cgctttccag tcggaaacc
5761 tgcgtgcca gctgcattaa tgaatcgcc aacgcgcggg gagaggccgt ttgcgtatttq
5821 ggcgcetttc cgcttcctcg ctcactgact cgctgcgcgc ggtcgltcgg ctgcggcgag
5881 cggtatcago tcactcaaag gcgtaataac gttatccac agaatcaggg gataacgcag
5941 gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa cctaaaaag gccgcgttgc
6001 tggcggtttt ccataggctc cgcccccctg acgagcatca caaaaatcga cgctcaagtc
6061 agaggtggcg aaacccgaca ggactataaa gataccaggc gtttccctt ggaagctccc
6121 tcgtgcgcgc tcctgttccg accctgcgc ttacccgata cctgtccgc ttctccctt
6181 cgggaagcgt ggcgccttct catagotcac gctgttaggt ttcagttcg gtgttaggtcg
6241 ttgcgtccaa gctggctgt gtgcacgaaac ccccgctca gcccggccgc tgcccttat
6301 ccggtaacta tcgtcttgag tccaaaccgg taagacacga ctatcgcca ctggcagcag
6361 ccactggtaa caggattagc agagcgaggt atgtaggccg tgctacagag ttcttgaagt
6421 ggtggcctaa ctacggctac actagaagaa cagtatttgg tatctgcgt ctgctgaagc
6481 cagttacattt cggaaaaaga gttggtagct cttgatccgg caaacaaacc accgctggta
6541 gccgtgggtt tttgtttgc aagcagcaga ttacccgcag aaaaaaaagga tctcaagaag
6601 atcccttgc ttttctacg gggctgcacg ctcagttggaa cgaaaactca cgttaaggga
6661 ttttggcat gagattatca aaaaggatct tcaccttagat cttttaaat taaaatgaa
6721 gttttaaatc aatctaaagt atatatgagt aaacttggc tgacagttac caatgcattaa
6781 tcagtggc acctatctca gcgatctgc tatttcgttc atccatagtt gcctgactcc
6841 ccgtcgta gataactacg atacgggagg gcttaccatc tggcccccgt gctgcaatga
6901 taccgcgaga cccacgctca ccggctccag atttacgcg aataaaccag ccagccggaa
6961 gggccgagcg cagaagtggc cctgcaactt tatccgcctc catccagttt attaattgtt
7021 gccgggaagc tagatgtt agttcgccag ttaatagtt ggcacacgtt gtgcatttg
7081 ctacaggcat cgtgggtcata cgctcggtt ttgttatggc ttcatccgc tcgggttcccc
7141 aacgateaag gcgagttaca tgatccccaa tttttttttt aaaaagggtt agcttccctcg
7201 gtcctccgat cgttgtcaga agtaagtggc ccgcagtgtt atcactcatg gttatggcag
7261 cactgcataa ttcttact gtcatgcac ccgtaaatgt cttttctgtg actggtagt
7321 actcaaccaa gtoattctga gaatagtgtt tgccggcgacc gagttgtct tgcccggt
7381 caatacggga taataccgcg ccacatagca gaactttaaa agtgcgcattc attggaaaaac
7441 gttttccggc gcgaaaaactc tcaaggatct taccgcgttt gagatccagt togatgtaa
7501 ccactcggtc acccaactga tcttcgcac cttttacttt caccagcggt tctgggt
7561 caaaaacagg aaggcaaaat gcccggaaaa agggaaataag ggcgacacgg aatgttgaa
7621 tactcataact cttccctttt caatattttaa gaagcatttta tcagggttat tttttttt
7681 gcccatacat atttgaatgt atttagaaaa ataaacaaat aggggtccg cgacatttc
7741 cccggaaaaat gcccacgtac gtctaaagaaa ccattattat catgacattttaa
7801 ataggcgat cacgaggccc ttgcgc (SEQ ID NO:8)

Figure 7C

CS01m1-FL-NA

ATGCAGATTGAGCTGTCCACCTGCTTCTTGTGCCCTGCTGAGATTCTGCTTCCTGCCACCAGGA
GATACTACCTGGGGCTGTGGAACCTTCTTGGACTACATGCAGTCTGACCTGGGAGAGCTGCCTGT
GGATGCCAGGTTCCCACCCAGAGTGCCAAGTCCTCCCATTCAACACCTCTGTTGAGTACAAGAAG
ACACTCTTGTGGAATTCACTGACCACTGTTCAACATTGAAAACCCAGACACCACCCCTGGATGGGAC
TCCTGGGACCCACCATTCACTGAGGTGTATGACACTGTTGAGTCAACCTCAAGAACATGGCATC
CCACCCCTGTGTCTGCATGCTGTGGAGTCTCATACTGAAAGGCTCTGAAGGGCTGAGTATGAT
GACCAGACATCCCAGAGAGAAAGAGGATGACAAGGTGTTCCCTGGGGATCTCACACCTATGTGT
GCCAAGTCCTCAAGGAGAATGGACCCATGGCATCTGACCCACTCTGCCTGACATACTCCTACCTTC
TCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGACTGCTGGTGTGCAGGGAAAGGA
TCCCTGGCCAAGGAGAAAACCCAGACACTGCACAAGTTCTCTCTGTTGCTGTCTTGATGAGG
GCAAGTCTTGGCACTCTGAAACAAAGAAACTCCCTGATGCAAGACAGGATGCTGCCCTGCCAGGGC
ATGGCCAAGATGCACACTGTGAATGGCTATGTGAACAGATCACTGCCCTGGACTCATGGCTGCCAC
AGGAAATCTGTCTACTGGCATGTGATTGGCATGGGACAACCCCTGAAGTGCACCTCCATTCTGG
AGGGACACACCTTCTGGTCAGGAACCACAGACAGACAGCCTCTGGAGATCTCTCCATCACCTCCT
CACTGCACAGACACTGCTGATGGACCTTGGACAGTTCTGTCATGCCACATCTCTTCCACCAG
CATGATGGCATGGAAGCCTATGTCAAGGTGGACTCATGCCCTGAGGAACCACAGCTCAGGATGAAGA
ACAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATGGATGGTCAAGATTGA
TGATGACAACCTCTCCATCOTTCAATTGAGATCAGGCTGTGGCAAAGAAACACCCCAAGACATGGGTG
CACTACATTGCTGCTGAGGAAGAGGACTGGGACTATGCAACACTGGTCTGGCCCTGATGACAGGA
GCTACAAGTCTCAGTACCTCAACAATGGCCACAAAGAATTGGAAGAAAGTACAAGAAAGTCAGATT
CATGGCCTACACTGATGAAACCTTCAAGACAAGAGAACGCATTCACTGAGTCAGGCTGGCATTCTGGGA
CCACTCCTGTATGGGAAGTGGGAGACACCCCTGTCATCATCTTCAAGAACCCAGGCCCTCCAGGCCCT
ACAACATCTACCCACATGGCATCACTGATGTCAGGCCCTGTACAGCAGGAGACTGCCAAAGGGGT
GAAACACCTCAAGGACTTCCCCATTCTGCCCTGGAGAGATCTCAAGTACAAGTGGACTGTCAGTGT
GAGGATGGACCAACAAAGTCTGACCCAGGTGCCCTCACCAGATACTACTCCTCTTGTGACATGG
AGAGAGACCTGGCATCTGGACTGATTGGACCACTGCTCATCTGCTACAAGGAGTCTGTGGACCAAG
AGGCAACCAGATCATGTCTGACAAGAGAAATGTGATTCTGTTCTGTCTTGATGAGAACAGATCA
TGGTACCTGACTGAGAACATTCACTGAGAGATTGCCAACCTGCTGGGATGCCACTGGAAAGACCCCTG
AGTTCCAGGCAAGCAACATCATGCACTCCATCAATGGCTATGTGTTGACTCTCTCCAGCTTCTGT
CTGCCCTGCATGAGGTGGCCTACTGGTACATTCTCTATTGGGACAAAAGTCACTTCTCTGTGAC
TTCTCTCTGGATACACCTCAAGCACAAGATGGTGTATGAGGACACCCCTGACACTCTTCCATTCT
CTGGGAAACTGTGTTCATGAGCATGGAGAACCCCTGGACTGTGGATTCTGGATGCCACAACACTGT
CTTCAGAAAACAGGGGAATGACTGCACTGCTCAAAGTCTCCCTGTGACAAGAACACTGGGACTAC
TATGAGGACTCTTATGAGGACATCTGCTCCTACCTGTCAGCAAGAACATGCCATTGAGCCAGAA
GCTTCTCTCAGAAATCCACCTGCTCTGAAGAGACACCAGAGAGATCACCAGGACAACCCCTCCAGTC
TGACCAGGAAGAGATTGACTATGATGACACCATTCTGAGGAGATGAGAACAGGAGGACTTTGACATC
TATGATGAGGACGAGAACCAAGTCTCCAAGATCATTCCAGAAGAACAGACACTACTTCATTGCTG
CTGTGGAAAGACTGTGGACTATGGCATGTCTCTCCCCATGTCTCAGGAACAGGGCACAGTC
TGGCTCTGTGCCACAGTTCAAGAAAAGTGGCTTCCAGGAGTTCACTGATGGCTCATCACCAGCCC
CTGTACAGAGGGGAACATGAAATGAGCACCTGGGACTCCTGGGACCATACATCAGGGCTGAGGTGGAAG
ACAACATCATGGTGACATTCAAGAACCCAGGGCTCCAGGCCCTACAGCTTCACTCTCCCTCATCAG
CTATGAGGAAGACCAGAGAACAGGGCTGAGCCAAGAAAAGAACTTGTGAAACCAATGAAACCAAG
ACCTACTTCTGGAAAGTCCAGCACCATGGCACCCACCAAGGATGAGTTGACTGCAAGGCCTGGG

(Continued)

Figure 8A

CATACTTCTCTGATGTGGACCTGGAGAAAGATGTGCACTCTGGCCTGATTGGCCCCACTCCTGGTCTG
CCACACCAACACCCCTGAACCCCTGCACATGGAAGGCAGTGAATGTGCAAGGAGTTGCCCTCTTC
ACCATCTTGATGAAACCAAGTCATGGTACTTCAGTGAGAACATGGAGAGAAAATGCAAGAGCACCACAT
GCAACATTCAAGATGGAAGACCCCACCTCAAGGAGAACATACAGGTTCCATGCCATCAATGGCTACAT
CATGGACACCCCTGCCTGGCTGTCAATGGCACAGGACCAGAGAACATGAGATGGTACCTGCTTCTATG
GGATCCAATGAGAACATTCACTCCACTTCTCTGGGCATGTCTCACTGTGAGAAAGAAGGAGG
AATACAAGATGCCCTGTAACACCTCTACCTGGGTCTTGAGACTGTGGAGATGCTGCCCTCAA
AGCTGGCATCTGGAGGGTGGAAATGCCCTCATTGGGGACCTGCATGCTGGCATGTCAACCCGTTC
CTGGCTACAGCAACAAGGCCAGACACCCCTGGGAATGCCCTCTGCCACATCAGGACTCCAGA
TCACTGCCTCTGCCAGTATGCCAGTGGGCACCCAAACTGCCAGGCTCCACTACTCTGGCTCCAT
CAATGCATGGTCAACCAAGGAGCCATTCTCTGGATCAAGGTGGACCTGCTGGCACCCATGATCATT
CATGGCATCAAGACACAGGGGCAAGACAGAAATTCTCCTCTGTACATCTCACAGTTCATCATCA
TGTACTCTCTGGATGGCAAGAAGTGGCAGACATACAGAGGCAACTCCACTGGCACCCCTCATGGTCTT
CTTGGCAATGTGGACAGCTCTGGCATCAAGCACAACATCTCAACCCCTCCATCATTGCCAGATA
ATCAGGCTGCACCCCACTACTCAATCAGATCAACCCCTCAGGATGGAACGTGGATGTGACC
TGAACCTCTGCTCAATGCCCTGGGAATGGAGAGCAAGGCCATTCTGATGCCAGATCACTGCATC
CTCTTACTTCACCAACATGTTGCCACCTGGTCAACATCAAAGCCAGGCTGCACCTCCAGGGAAGA
AGCAATGCCCTGGAGACCCAGGTACAACCCAAAGGAATGGCTGCAAGTGGACTTCCAGAAGACAA
TGAAAGTCACTGGGTGACAACCCAGGGGTCAAGTCTGCTCACCTCAATGTATGTGAAGGAGTT
CCTGATCTCTCCACAGGATGCCACCGAGTGGACACTCTTCTTCCAGAATGGCAAAGTCAGGTG
TTCCAGGGCAACCAGGACTTTCACACCTGTGGTGAACACTGGAACCCCCCTCCTGACAAGAT
ACCTGAGAATTCAACCCCCAGTCTTGGTCCACAGATTGCCCTGAGAATGGAAGTCCTGGATGTGA
GGCACACAAGACCTGTACTGA (SEQ ID NO: 49)

Figure 8B

CS04Δ (760-1667) - CS04-SC1-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCTGTGCCCTGCTGAGGGTCTGCTTCTCTGCCACCAGGAGATA
TACCTGGGGCTGTGGAGCTTCTTGGACTACATGCAGTCTGACCTGGGGAGCTGCCCTGTGGATGCCAGG
TTCCCACCCAGAGTGCCAAATCCTTCCCATTCAACACCTCTGTGGTCTACAAGAAGACCCCTCTTGAG
TTCACTGACCACCTGTTCAACATTGCCAAACCCAGGCCACCCCTGGATGGGACTCCTGGGACCCACATTCA
GCTGAGGTGTATGACACTGTGGTCATCACCTCAAGAACATGCCCTCCACCCCTGTGAGGCTGATGCTGT
GGGGTCACTGGAAGGCTCTGAGGGGCTGAGTATGATGACCAAGACCTCCAGAGGGAGAAGGAGGAT
GACAAAGTGTCCCTGGGGCAGCCACACCTATGTGTGGCAGGTCTCAAGGAGAAATGGCCCCATGCCCT
GACCCACTCTGCCCTGACCTACTCTACCTTCTCATGTGGACCTGGTCAAAGGACCTCAACTCTGACTGATT
GGGGCCCTGCTGGTGTGCAGGGAGGGCTCCCTGGCCAAGAGAGAACCCAGACCCCTGACAAGTTCATTCTC
CTGTTGCTGTCTTGATGAGGGCAAGAGCTGGACTCTGAAACCAAGAACCTGGTGTGAACAGGAGCCTGGACTC
GCTGCTCTGCCAGGGCTGGCCAAGATGCACACTGTGAATGGCTATGTGAACAGGAGCCTGGACTC
ATTGGCTGCCACAGGAATCTGTCTACTGGCATGTGATTGGCATGGGACAACCCCTGAGGTGCACTCCATT
TTCTGGAGGCCACACCTTCTGGTCAGGAACCACAGACAGGCCAGCCTGGAGATCAGCCCCATCACCTTC
CTCACTGCCAGACCCCTGCTGATGGACCTCGGACAGTCCCTGCTGTGTTCTGCCACATCAGCTCCACAGCAT
GATGGCATGGAGGCATATGTCAGGTGGACAGCTGCCCTGAGGAGCCACAGCTCAGGATGAAGAACATGAG
GAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATGGATGTGGTCCGCTTGATGATGACAACAGC
CCATCCTTCATTCAAGATCAGGTCTGTGCCAAGAACACCCCCAAGACCTGGTGCACTAATTGCTGCTGAG
GAGGAGGACTGGACTATGCCCACTGGTCTGGCCCTGATGACAGGAGCTACAAGAGCCAGTACCTCAAC
AATGGCCCACAGAGGATTGGACGCAAGTACAAGAACATGAGGTATGGCCTACACTGATGAAACATTCAAG
ACCAGGGAGGCCATTCAAGCATGAGTCTGGCATCTGGGCCACTCCTGTATGGGAGGTGGGGACACCCCTG
CTCATCATCTTCAGAACCCAGGCCCTCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCC
CTGTCAGCCCGCAGGCTGCCAAAGGGGTGAAACACCTCAAGGACTCTCCCATTCTGCTGGGAGATCTTC
AACTACAAGTGGACTGTCACTGTGGAGGATGGACCAACAAATCTGACCCAGGTGCCCTACCAAGATACTAC
TCCAGCTTGTGAAACATGGAGGGACCTGGCCTCTGGCCTGATTGGCCACTGCTCATGCTACAAGGAG
TCTGTGGACCAGAGGGAAACCAAGATCATGTCAGCAAGAGGAATGTGATCTGTTCTGTCTTGTGAG
AACAGGAGCTGGTACCTGACTGAGAACATTCAGCGCTTCTGCCAACCTGCTGGGAGTGCAGCTGGAGGAC
CCTGAGTTCCAGGCCAGCAACATCATGCACTCCATCAATGGCATGTGTTGACAGCCCTCCAGCTTCTGTC
TGCCCTGCATGAGGTGGCCTACTGGTACATTCTTCTATTGGGGCCAGACTGACTTCCCTCTGCTCTTC
TCTGGCTACACCTCAAACACAAGATGGTGTATGAGGACACCCCTGACCCCTTCCCATTCTCTGGGAGACT
GTGTTCATGAGCATGGAGAACCTGGCCTGTGGATTCTGGGATGCCACAACCTGACTTCCGCAACAGGGC
ATGACTGCCCTGCTAAAGTCTCTCTGTGACAAGAACACTGGGACTACTATGAGGACAGCTATGAGGAC
ATCTCTGCCCTACCTGCTCAGCAAGAACATGCCATTGAGCCACAGGGAGATCACCAGGACACCCCTCAGTCT
GACCAGGAGGAGATTGACTATGATGACACCAATTCTGTGGAGGATGAAGAACAGGAGACTTGTACACTATGAC
GAGGACGAGAACCAAGAGGCCAACGGAGCTCCAGAAGAACAGGACAGGCACTACTTCATTGCTGTGGAGC
CTGTGGGACTATGGCATGAGCTCCAGGCCCTATGTCTCAGGAACAGGGCCAGTCTGGCTCTGTCACAG
TTCAAGAACAGTGGCTTCCAAAGAGTTCACTGATGGCAGCTTCACCCAGGCCCTGTACAGAGGGAGCTGAAT
GAGCACCTGGGACTCCTGGGCCCATACATCAGGGCTGAGGTGGAGGACAACATCATGGTGCACCTCCGCAAC
CAGGCCCTCAGGCCCTACAGCTTCTACAGCTCCCTCATCAGCTATGAGGAGGACAGAGGCCAGGGCTGAG
CCACGCAAGAACATTGTGAAACCAATGAAACCAAGACCTACTTCTGGAAAGTCCAGCACCACATGCC

(Continued)

Figure 9A

ACCAAGGATGAGTTGACTGCAAGGCCTGGGCCTACTTCTCTGATGTGGACCTGGAGAAGGATGTGCACTCT
GGCCTGATTGGCCCACCTCTGGTCTGCCACACCAACACCCCTGAACCCCTGCCCATGGAAGGCAAGTGACTGTG
CAGGAGTTGCCCTCTTCACCATCTTGATGAAACCAAGAGCTGGTACTTCACTGAGAACATGGAGCGC
AACTGCAGGGCCCCATGCAACATTCAAGATGGAGGACCCCACCTCAAAGAGAACTACCGCTTCCATGCCATC
AATGGCTACATCATGGACACCCCTGCTGGGCTTGTCTGGCCAGGACCAGAGGATCAGGTGGTACCTGCTT
TCTATGGGCTCCAATGAGAACATTCACTCCATCCACTTCTCTGGCATGCTTCACTGTGCGCAAGAAGGAG
GAGTACAAGATGGCCCTGTACAACCTCTACCCCTGGGTCTTGAGACTGTGGAGATGCTGCCCTCAAAGCT
GCCATCTGGAGGGTGGAGTGCCTCATTGGGAGCACCTGCATGCTGGCATGASCACCCCTGTTCTGGTCTAC
AGCAACAAGTGCCAGACCCCCCTGGGAATGGCTCTGGCCACATCAGGGACTCCAGATCACTGCCCTGGC
CAGTATGCCAGTGGGCCCCAAGCTGGCCAGGCTCCACTACTCTGGATCCATCAATGCTGGAGCACCAAG
GAGCCATTCACTGGGATCAAAGTGGACCTGCTGGCCCCATGATCATCCATGGCATCAAGACCCAGGGGCC
AGGCAGAAGTTCTCCAGGCTGTACATCAGCCAGTTCATCATGTACAGGCTGGATGGCAAGAAATGGCAG
ACCTACAGAGGCAACTCCACTGGAACACTCATGGCTTCTTGGCAATGTGGACAGCTCTGGCATCAAGCAC
AACATCTCAACCCCCCAATCATGCCAGATACTCAGGCTGCACCCCACCCACTACAGCATCCGCAGCACC
CTCAGGATGGAGCTGATGGGTGTGACCTGAACCTCTGCAGCATGCCCTGGCATGGAGAGCAAGGCCATT
TCTGATGCCAGATCACTGCTCCAGCTACTTCACCAACATGTTGCCACCTGGAGGCCAAGCAAGGCCAGG
CTGCACCTCCAGGGAAAGGAGCAATGCCCTGGAGGCCAGGTCAACAACCCAAAGGAGTGGCTGCAGGTGGAC
TTCCAGAAGACCATGAAGGTCACTGGGTGACCACCCAGGGGGTCAAGAGGCTGCTCACCAGCATGTATGTC
AAGGAGTTCTGATCAGCTCAGCCAGGATGGCCACCAGTGGACCCCTCTTCTTCAAGAATGGCAAGGTCAAG
GTGTTCCAGGGCAACCAGGGACAGCTTCACCCCTGTGGTGAACAGCCTGGACCCCCCCCCTGACCAGATA
CTGAGGATTCAACCCCCAGAGCTGGTCCACCAGATTGCCCTGAGGATGGAGGTCTGGATGTGAGGCCAG
GACCTGTACTGA (SEQ ID NO:9)

Figure 9B

CS04A(760-1667) - CS04-SC1-AA

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDFPPRVPKSFPFNTSVVYKK
TLFVEFTDHLFNIAKPRPPWMGLLGTIQAEVYDTVVITLKNMASHPVSLLHAVGVSYWKASEGAEYD
DQTSQREKEEDDKVFPGGSHTYVWQVLKENGPMASDPLCLTYSYLSHVDLVKDLNSGLIGALLVCREG
SLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGCH
RKSVYWHVIGMGTPEVHSIFLEGHTFLVRNHRQASLEISPTFLTAQTLMDLGQFLIFCHISSHQ
HDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSEIQIRSVAKKHPKTWV
HYIAEEEDWDYAPLVLAPDDRSYKSQYLNNNGPQRIGRKYKKVRFMAYTDETFKTREAIQHESGILG
PLLYGEVGDTLLIIFKNQASRPYNTIYPHGIDVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVT
EDGPTKSDPRCLTRYSSFVNMERDLASGLIGPLLICYKESVDQRGNQIMSDKRNVILFSVFENRS
WYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSLQLSVCVLCHEVAYWIILSIGAQTDFLSV
FFSGYTFKHKMVYEDTLLTFPFSGETVFMMSMENPGLWILGHNSDFRNRCMTALLKVSSCDKNTGDY
YEDSYEDIISAYLLSKNNATEPRETTRTTLQSDQEETDYDDTTSEMKEFDTYDEDENQSPRSFQK
KTRHYFIAAVERLWDYGMSSSPHVLRNRAQSGSPQFKKVVFOEFTDGSFQPLYRGELNEHGLLG
PYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHHMAPTK
DEFDCCKAWAYFSDVDLEKDVGSLIGPLLVCHTNTLNPAHGRQVTVOEFALFFTIFDETKSWEFTEN
MERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIPWYLLSMGSNENIHSIHFSGH
VFTVRKKEEYKMALYNLYPGVFETVEMPSKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMA
SGHIRDFQITASGQYQWAPKLARLHYSGSINAWSTKEPPSWIKVDLLAPMTIHGIKTQGRQKFSS
LYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIARYIRLHPTHYSIRSTL
RMELMGCDLNSCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEW
LQVDFQKTMKVTVGTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQONGKVKVFQGNQDSFTPVVNS
LDPPPLLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY (SEQ ID NO:10)

Figure 10

CS04A (772-1667) - CS04-SC2-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCGTGCCTGCTGAGGTTCTGCTTCTGCCACCAGG
AGATACTACCTGGGGCTGTGGAGCTTCTTGGACTACATGCAGTCTGACCTGGGGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGACTGCCAAATCCTCCATTCAACACACCTCTGGTCTACAAG
AAGACCTCTTGTGGAGTTCACTGACCACCTGTTCAACATTGCCAAACCCAGGCCACCTGGATG
GGACTCCTGGGACCCACCATTCAAGGTGAGGTGATGACACTGTGGTCATCACCTCAAGAACATG
GCCTCCCACCCCTGTGAGCCTGCATGCTGTGGGGTCAGCTACTGGAAGGCCCTGAGGGGGCTGAG
TATGATGACCAGACCTCCCAGAGGGAGAAGGGAGGATGACAAAGTGTCCCTGGGGCAGCCACACC
TATGTGTGGCAGGTCTCAAGGAGAATGGCCCCATGGCCTCTGACCCACTCTGCTGACCTACTCC
TACCTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATGGGGCCCTGCTGGTGTGC
AGGGAGGGCTCCCTGGCCAAGAGAAGACCCAGACCTGCACAAGTTCAATTCTCCTGTTGCTGTC
TTTGATGAGGGCAAGAGCTGGCACTCTGAAACCAAGAACACTCCCTGATGCAAGGACAGGGATGCTGCC
TCTGCCAGGGCTGGCCAAGATGCACACTGTGAATGGCTATGTGAACAGGAGCCTGCCCTGGACTC
ATTGGCTGCCACAGGAATCTGTCTACTGGCATGTGATTGGCATGGGACAACCCCTGAGGTGCAC
TCCATTTCTGGAGGGCACACCTTCTGGTCAGGAACCACAGACAGGCCAGCCTGGAGATCAGC
CCCATCACCTCCTCACTGCCAGACCCCTGCTGATGGACCTCGGACAGTCTGCTGTTCTGCCAC
ATCAGCTCCCACCAGCATGGCATGGAGGCATGTCAAGGTGGACAGCTGCCCTGAGGAGCCA
CAGCTCAGGATGAAGAACAAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCCGTTGATGATGACAACAGCCCATCCTCAATTCAAGATCAGGTCTGTGGCCAAGAAA
CACCCCAAGACCTGGGTGACTACATTGCTGCTGAGGAGGAGCTGGACTATGCCCACTGGTC
CTGGCCCTGATGACAGGAGCTACAAGAGCCAGTACCTCAACAATGCCCAACAGAGGATTGGACGC
AAGTACAAGAAAGTCAGGTCATGGCCTACACTGATGAAACCTCAAGAACCCAGGGAGGCCATTCA
CATGAGTCTGGCATCCTGGCCCACTCCTGATGGGAGGTGGGGACACCCCTGCTCATCATCTTC
AAGAACCCAGGCCCTCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCCTGTAC
AGCCCAAGGCTGCCAAAGGGGGTGAACACACCTCAAGGACTTCCCCATTCTGCTGGGGAGATCTTC
AAAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACCAAATCTGACCCAGGTGCCCTACCCAGA
TACTACTCCAGCTTGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATGGGCCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGGGAAACCAAGATCATGTCTGACAAGAGGAATGTGATTCTG
TTCTCTGTCTTGATGAGAACAGGAGCTGGTACCTGACTGAGAACATTCAAGCCTTCTGCCCAAC
CCTGCTGGGTGCAAGCTGGAGGACCTGAGTCCAGGCCAGCAACATCATGCACTCCATCAATGGC
TATGTGTTGACAGCCTCCAGCTTCTGCTGCTGATGAGGTGGCTACTGGTACATTCTTCT
ATTGGGGCCAGACTGACTCCTTCTGCTTCTGCTACACCTTCAAAACACAAGATGGTG
TATGAGGACACCCCTGACCCATTCTGAGGGAGACTGTGTTCATGAGCATGGAGAACCC
GGCCCTGGAATTCTGGGATGCCACAACACTCTGACTTCCGCAACAGGGCATGACTGCCCTGCTCAA
GTCTCCTCTGTGACAAGAACACTGGGAGACTATGAGGACAGCTATGAGGACATCTCTGCCCTAC
CTGCTCAGCAAGAACAAATGCCATTGAGGCCAGGGAGCTCAGCCAGAAATTCCAGACACCCCCAGCACC
AGGGAGGATCACCAGGACCAACCCCTCCAGTCTGACCAGGAGGAGATTGACTATGATGACACCATTCT
GTGGAGATGAAGAAAGAGGACTTGAACATCTATGACGAGGACGAGAACAGGCCAAGGAGCTTC

(Continued)

Figure 11A

CAGAAGAAGACCAGGCACTTCTATTGCTGCTGTGGAGCGCCTGTGGGACTATGGCATGAGCTCC
AGCCCCCATGTCCTCAGGAACAGGGCCAGTCTGGCTCTGTGCCACAGTTCAAGAAAAGTGGTCTTC
CAAGAGTTCACTGATGGCAGCTCACCCAGCCCCGTACAGAGGGAGCTGAATGAGCACCTGGGA
CTCCTGGGCCATACATCAGGGCTGAGGTGGAGGACAACATCATGGTACCTTCCGCAACCAGGCC
TCCAGGCCCTACAGCTCTACAGCTCCCTCATCAGCTATGAGGAGGACCAGAGGCAGGGGCTGAG
CCACGCAAGAACCTTGAAACCCAATGAAACCAAGACTACTTCTGGAAAGTCCAGCACACATG
CCCCCCCACCAAGGGATGAGTTGACTGCAAGGCCCTGGGCTACTTCTCTGATGTGGACCTGGAGAAG
GATGTGCACTCTGGCTGATTGGCCACTCCTGGTCTGCCACACCAACACCCTGAACCCCTGGCCAT
GGAAGGCAAGTGAUTGTGCAGGGAGTTGCCCTCTTCTCACCATCTTGATGAAACCAAGAGCTGG
TACTTCACTGAGAACATGGAGCGCAACTGCAGGGCCCATGCAACATTCAGATGGAGGACCCACC
TTCAAAGAGAACATACCGCTTCCATGCCATCAATGGCTACATCATGGACACCCCTGCCCTGGGCTTGTC
ATGGCCCAGGACCAGAGGATCAGGTGGTACCTGCTTCTATGGGCTCCAATGAGAACATTCACTCC
ATCCACTTCTCTGGCATGTCTTCACTGTGCCAAGAAGGGAGGAGTACAAGATGCCCTGTACAAC
CTCTACCCCTGGGTCTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGGAGGGTGGAG
TGCCTCATGGGGAGCACCTGCATGCTGGCATGAGCACCCCTGTTCTGGTCTACAGCAACAAAGTGC
CAGACCCCCCTGGGAATGGCCTCTGCCACATCAGGGACTTCCAGATCACTGCCCTGGCAGTAT
GCCAGTGGCCCCCAAGCTGCCAGGCTCCACTACTCTGGATCCATCAATGCCCTGGAGCACCAAG
GAGCCATTCAAGCTGGATCAAAGTGGACCTGCTGGCCCCCATGATCATCCATGGCATCAAGACCCAG
GGGGCCAGGCAGAAGTCTCCAGGCTGTACATCAGCCAGTTCATCATGTACAGCCTGGATGGC
AAGAAATGGCAGACCTACAGAGGCAACTCCACTGGAACACTCATGGTCTTGGCAATGTGGAC
AGCTCTGGCATCAAGCACACATCTCAACCCCCCAATCATGCCAGATACTCAGGCTGCCACCC
ACCCACTACAGCATCCGCAGCACCCCTCAGGATGGAGCTGATGGGCTGTGACCTGAACCTCTGCAGC
ATGCCCTGGCATGGAGAGCAAGGCCATTCTGATGCCAGATCACTGCCCTCAGCTACTTCACC
AACATGTTGCCACCTGGAGGCCAAGCAAGGCCAGGCTGCACCTCAGGGAGGAGCAATGCCCTGG
AGGCCCCAGGTCAACACCCAAAGGAGTGGCTGCAGGTGGACTTCCAGAAGACCATGAAGGTCACT
GGGGTGACCACCCAGGGGGTCAAGAGCCTGCTCACCAGCATGTATGTGAAGGAGTTCTGATCAGC
TCCAGCCAGGATGGCCACCAAGTGGACCCCTCTTCCAGAATGGCAAGGTCAAGGTGTTCCAGGGC
AACCAGGACAGCTTCACCCCTGTGGTGAACAGCCTGGACCCCCCCTCTGACCAGATACTGAGG
ATTCAACCCCCAGAGCTGGTCCACCAGATTGCCCTGAGGATGGAGGTGTTGGATGTGAGGCCAG
GACCTGTACTGA (SEQ ID NO:11)

Figure 11B

CS04A(772-1667) - CS04-SC2-AA

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMOSDLGELPVNDARFPPRVPKSFPFNTSVVYKKTLFVEFTDHLFNIAKPRPPWMGLGPQTQAEVYDTVVITLKNMASHPVSLHAVGVSYWKASEGAEYDDQTSQREKEEDKVFPGGSHTYVWQVLKENGPMSADPLCLTYSYLSHVDLVKDLSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETRNSLMQDRDAASARAWPKMHTVNNGYVNRSLPGLIGCHRKSVYWHVIGMGTTPEVHSIFLEGHTFLVRNHRQASLEISPITFLTAQTLIMDLGQFLFCHISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKHPKTWVHYIAAEEEDWDYAPLVLAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETETKTREARIQHESGLLGPLLYGEVGDTLLIFKNQASRPYNIYPHGIDVRLYSRRLPKGVKHLKDFPILPGEIFKYKWTVTVEDGPTKSDPRCLTRYSSFVNMERDIASGLIGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENRSWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSDLQLSVCLHEVAYWYILSIGAQTDLFSVFFSGYTFKKHMVYEDTLTLPFSGETVFMSSMENPGLWILGCHNSDFRNNGMTALLKVSSCDKNTGDYYEDSYEDISAYLLSKNNATEPRSFSQSNSRHPSTREITRTTLQSDQEEIDYDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSSFHVRNRAQSGSVPQFKVVFQEFTDGSFTQPLYRGELNHHLGLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQROGAEPRKNFVKPNETKTYFWKVQHHMAPTKDEFDCCKAWAYFSDVDLEKDVHSGLIGPLLVCHTNLPAHGRQVTVOEFALFFTIFDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIHSIHFSGHVFTVRKKEEYKMALYNLYPGVFETVEMPLPSKAGIWRVECLIGEHLHAGMSTLFLVYSNKCQTPLGMASGHIRDQFQITASGQYGCQWAPKLARLHYSGSINAWSSTKEPFWSIKV DLLAPMIHGIKTQGARQKFSSLIYSQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPEPIIAEYIRLHPHTYSIRSTLREMELMGCDLNCSMPLGMESKAISDAQITASSYFTNMFMATWSPSKARLHLQGRSNAWRFQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLFFQNGKVKVFGNQDSFTPVVNSLDPILLTRYLRIHPQSWVHQIALRMEVLGCEAQDLY(SEQ ID NO:12)

Figure 12

NG1: V S N N V S N N A T N N A T N (SEQ ID NO:51)
GTC AGC AAC AAT GTG AGC AAC AAT GTC ACC AAC AAT GCT ACC AAC (SEQ ID NO:50)

NG4: V S N N A T N N V S N (SEQ ID NO:53)
GTG AGC AAC AAT GCC ACC AAC AAT GTG AGC AAC (SEQ ID NO:52)

NG5: V S N N A T N (SEQ ID NO:55)
GTG AGC AAC AAT GCC ACC AAC (SEQ ID NO:54)

NG6: V S N N (SEQ ID NO:57)
GTG AGC AAC AAT (SEQ ID NO:56)

NG9: R S L (SEQ ID NO:59)
AGG AGC CTG (SEQ ID NO:58)

NG10: A T N V S N N S A T S A D S A V S (SEQ ID NO:61)
GCC ACT AAT GTG TCT AAC AAC TCT GCT ACC TCT GCT GAC TCT GCT GTC AGC (SEQ ID NO:60)

NG16: A T N Y V N R S L (SEQ ID NO:63)
GCC ACC AAC TAT GTG AAC AGG AGC CTG (SEQ ID NO:62)

Figure 13A

NG17: A T N Y V N R S I S A T S A D S A V S Q N (SEQ ID NO:65)
GCC ACC AAC TAT GTG AAC AGC CTG TCT GCC ACC TCT GCT GAC TCT GCT GTG AGC CAG AAT (SEQ ID NO:64)

NG18: V S N N V S N A V S A V S A (SEQ ID NO:67)
GTG AGC AAC AAT GTG AGC AAC GCT GTG TCT GCT GCG TCT GCT (SEQ ID NO:66)

NG19: I T V A S A T S N I T V A S A D (SEQ ID NO:69)
ATC ACT GTG GCC TCT GCC ACC TCT AAC ATC ACT GTG GCC TCT GCT GAC (SEQ ID NO:68)

NG20: I T V T N I T V T A (SEQ ID NO:71)
ATC ACT GTG ACC AAC ATC ACT GTG ACT GCC (SEQ ID NO:70)

NG21: Q T V T N I T V T A (SEQ ID NO:73)
CAG ACT GTG ACC AAC ATC ACT GTG ACT GCC (SEQ ID NO:72)

NGv: A T N V S N N S N T S N D S N V S (SEQ ID NO:75)
GCC ACT AAT GTG TCT AAC AAC AGC AAC ACC AGC AAT GAC AGC AAC TCT GTG TCT (SEQ ID NO:74)

Figure 13B

Figure 14

Name: Sequence length: 41
 LSEHMAKKEPESQKATVYDQDQSTTSDSSTVSSPVTLKQKDKR

 (Threshold=0.5)

SeqName	Position	Potential	Jury agreement	N-Glyc result
Sequence	18	0.63224	(8/9)	+
Sequence	19	0.64333	(9/9)	++
Sequence	21	0.41358	(8/9)	-
Sequence	24	0.41558	(8/9)	-
Sequence	27	0.36119	(8/9)	-
Sequence	38	0.31349	(9/9)	----

Figure 15

CS01-FL-NA

ATGCAGATTGAGCTGTCCACCTGCTCTTCTGTGCCCTGCTGAGATTCTGCTCTGCCACCAGG
AGATACTACCTGGGGCTGTGGAACCTTCTTGGACTACATGCAGTCTGACCTGGGAGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGAGTGCCCAAGTCCTCCCATTCAACACACCTCTGGGTCTACAAG
AAGACACTCTTGTGGAATTCACTGACCACCTGTTCAACATTGCAAACACCCAGACCACCTGGATG
GGACTCCTGGGACCCACCATTCAGGCTGAGGTGTATGACACTGTGGTCATCACCCCTAAGAACATG
GCATCCCACCCCTGTGTCTGCATGCTGGGAGTCTCATACTGGAAAGCCTCTGAAGGGGCTGAG
TATGATGACCAGACATCCCAGAGAGAGAAAGAGGATGACAAGGTGTTCCCTGGGGATCTCACACC
TATGTGTGGCAAGTCCTCAAGGAGAATGGACCCATGGCATCTGACCCACTCTGCCTGACATACTCC
TACCTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGCACTGCTGGTGTGC
AGGGAGGATCCCTGGCCAAGGAGAAAACCCAGACACTGCACAAGTTCATTCTCCTGTTGCTGTC
TTTGATGAGGGCAAGTCTGGCACTCTGAAACAAAGAACACTCCCTGATGCAAGACAGGGATGCTGCC
TCTGCCAGGGCATGGCCCAAGATGCAACACTGTGAATGGCTATGTGAACAGATCACTGCCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCACTGGCATGTGATTGGCATGGGACAACCCCTGAAGTGCAC
TCCATTTCTGGAGGGACACACCTCCTGGTCAGGAACCACAGACAAGCCTCTGGAGATCTCT
CCCATCACCTCCTCACTGCACAGACACTGCTGATGGACCTGGACAGTTCCCTGCTGTTGCCAC
ATCTCTCCCACCAAGCATGGCATGGAAGCCTATGTCAGGTGGACTCATGCCCTGAGGAACCA
CAGCTCAGGATGAAGAACAAATGAGGAGGCTGAGGACTATGATGATGACCTGACTCTGAGATG
GATGTTGTCAGATTGATGACAACTCTCCATCCTTCATTGAGATCAGGCTGTTGCAAAGAAA
CACCCCAAGACATGGGTGCACTACATGCTGAGGAAGAGGACTGGACTATGCACCACTGGTC
CTGGCCCTGATGACAGGAGCTACAAGTCTCAGTACCTCAACAATGGCCCAAAAGAATTGGAAGA
AAAGTACAAGAAAGTCAGATTGATGGCCTACACTGATGAAACCTTCAAGACAAGAGAAGCCATTGAG
CATGAGTCTGGCATTCTGGGACCACCTCTGATGGGAAGTGGAGACACCCCTGCTCATCATCTTC
AAGAACCAAGGCCTCCAGGGCCTACAAACATCTACCCACATGGCATCACTGATGTCAGGCCCTGTAC
AGCAGGAGACTGCCAAAGGGTGAACACACCTCAAGGACTTCCCCATTCTGCCCTGGAGAGATCTTC
AAGTACAAGTGGACTGCACTGTGGAGGATGGACCAACAAAGTCTGACCCCAAGGTGCCCTACCAGA
TAATGACTCTCTTTGTGAACATGGAGAGAGACCTGGCATCTGGACTGATTGGACCACTGCTCATC
TGCTACAAGGAGTCTGTGACCAAGAGGCAACCAGATCATGTCAGAACAGAGAAATGTGATTCTG
TTCTCTGTCTTGATGAGAACAGATCATGGTACCTGACTGAGAACATTGAGGATTCCTGCCAAC
CCTGCTGGGTGCAACTGGAAGACCCCTGAGTTCCAGGCAAGCAACATCATGCACACTCCATCAATGGC
TATGTGTTGACTCTCTCAGCTTCTGTCTGCCCTGCATGAGGTGCCACTGGTACATTCTTCT
ATTGGGCACAAACTGACTTCTTCTGTCTTCTCTGGATACACCTCAAGCACAAAGATGGTG
TATGAGGACACCTGACACTCTTCCCATTCTCTGGGAAACTGTGTTCATGAGCATGGAGAACCC
GGACTGTGGATTCTGGATGCCACAACACTCTGACTTCAGAAACAGGGGAATGACTGCACTGCTCAA
GTCTCCTCTGTGACAAGAACACTGGGACTACTATGAGGACTCTTATGAGGACATCTGCCCTAC
CTGCTCAGCAAGAACAAATGCCATTGAGCCAGAAGCTCTCAGAATCCACCTGTCTGAAGAGA
CACCAAGAGAGAGATCACCAGGACAACCCCTCCAGTCTGACCAAGGAAGAGATTGACTATGATGACACC
ATTTCTGTGGAGATGAAGAAGGAGACTTGACATCTATGATGAGGACGAGAACCGACTCTCCAAGA
TCATTCAGAAGAACAAAGACACTACTTCATGCTGCTGTGGAAAGACTGTGGACTATGGCATG
TCTTCCCTCTCCCCATTGTCCTCAGGAACAGGGCACAGTCTGGCTCTGTGCCACAGTTCAAGAAAGTG
GTCTCCAGGAGTTCACTGATGGCTCATTCACCCAGGCCCTGTACAGAGGGAACTGAATGAGCAC

(Continued)

Figure 16A

CTGGGACTCCTGGGACCATAACATCAGGGCTGAGGTGGAAGACAACATCATGGTGACATTCAAGAAAC
CAGGCCTCCAGGCCCTACAGCTTACTCTTCCCTCATCAGCTATGAGGAAGACCAGAGACAAGGG
GCTGAGCCAAGAAGAACCTTGTGAAACCCAATGAAACCAAGACCTACTCTGGAAAGTCCAGCAC
CACATGGCACCCACCAAGGATGAGTTGACTGCAAGGCTGGGCATACTCTCTGATGTGGACCTG
GAGAAAGATGTGCACTCTGGCCTGATTGGCCACTCCTGGTCTGCCACACCAACACCCCTGAACCC
GCACATGGAAGGCAAGTGAUTGCAAGGAGTTGCCCCCTTCACCATCTTGATGAAACCAAG
TCATGGTACTTCACTGAGAACATGGAGAGAACTGCAGAGCACCATGCAACATTCAAGATGGAAGAC
CCCACCTCAAGGAGAACTACAGGTTCCATGCCATCAATGGCTACATCATGGACACCCCTGCGTGG
CTTGTATGGCACAGGACCAGAGAACATCAGATGGTACCTGCTTCTATGGGATCCAATGAGAACATT
CACTCCATCCACTTCTCTGGCATGTCTTCACTGTGAGAACAGGAGGAATACAAGATGGCCCTG
TACAACCTCTACCCCTGGGGTCTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGGAGG
GTGGAATGCCCTCATGGGGACCACCTGCATGCTGGCATGTCAACCCCTGTTCTGGTCTACAGAAC
AAGTGCCAGACACCCCTGGGAATGCCCTCTGCCACATCAGGGACTTCCAGATCACTGCCCTG
CAGTATGCCAGTGGCACCCAAACTGCCAGGCTCCACTACTCTGGCTCCATCAATGCATGGTCA
ACCAAGGAGCCATTCTCTGGATCAAGGTGGACCTGCTGGCACCCATGATCATTCAATGGCATCAAG
ACACAGGGGCAAGACAGAAATTCTCTCTCTGTACATCTCACAGTTCATCATGTACTCTCTG
GATGGCAAGAAGTGGCAGACATACAGAGGCAACTCCACTGGCACCCCTCATGGTCTTCTGGCAAT
GTGGACAGCTGTCATCAAGCACACATCTCAACCCCTCATGTCAGGAGATACATCAGGCTG
CACCCCAACCAACTACTCAATCAGATCAACCCCTCAGGATGGAACGTGATGGGATGTGACCTGAACCTCC
TGCTCAATGCCCTGGGAATGGAGAGCAAGGCCATTCTGATGCCAGATCACTGCATCTCTTAC
TTCACCAACATGTTGCCACCTGGTCAACATCAAAGCAGGCTGCACCTCCAGGGAAAGAACAT
GCCTGGAGACCCCAAGGTCAACAACCCAAAGGAATGGCTGCAAGTGGACTTCCAGAAGACAATGAAA
GTCACTGGGGTACAACCCAGGGGTCAAGTCTCTGCTCACCTCAATGTATGTGAAGGAGTTCC
ATCTCTTCTCACAGGATGGCCACCAAGTGGACACTCTCTTCCAGAATGGCAAAGTCAAGGTGTT
CAGGGCAACCAGGACTCTTCACACCTGTGGTAACACTGGACCCCCCCCCCTGACAAGATA
CTGAGAATTCAACCCCAAGTCTGGGTCCACCAAGATTGCCCTGAGAATGGAAGTCCCTGGGATGTGAG
GCACAAGACCTGTACTGA (SEQ ID NO:13)

Figure 16B

CS08-FL-NA

ATGCAGATCGAACTGAGCACTTGCTTCTCCTGTGTCFCCTGCGCTTTGCTCTCCGCCACAAGG
AGATACTATCTCGGTGCCGTGGAGCTCAGCTGGACTACATGCAGAGCAGCTGGGTGAACTGCCT
GTGGACGCCAGGTTCCACCCCGCCGCCAAGAGTTCCCGTTCAACACCAGTGTCTGTACAAG
AAAACCCCTCTCGTGGATTACCGACCACCTGTTCAACATGCCAACCGGCCCTCCCTGGATG
GGGCTGCTCGGCCGACGATCCAGGCTGAGGTCTATGACACGGTGGTATTACCCCTAAGAACATG
GCTAGCCACCCGGTGGAGCTGCACGCCGTGGCGTGTCTATTGGAAAGCGTCCGAGGGTGCAGG
TACGATGACCAGACTTCACAGCGGGAGAAGGAAGACGACAAAGTGTCCCCGGGGTCCCACACC
TATGTCTGGCAGGTCTGAAGGAGAATGGTCTATGCCCTCCGACCCATTGTGCCTCACCTACTCT
TACCTAACGCATGTGGATCTCGTCAAGGACCTGAACCTGGGCTGATCGGCCCTGCTCGTGC
CGGGAGGGCTCACTGGCCAAGGAGAAGACCCAAACTCTGCACAAGTTCATCTGCTGTTCGGGTA
TTCGACGAGGGGAAGTCCCTGGCACTCCGAGACCAAGAACAGCCTGATGCAGGACCGCAGCAGCC
TCGGCCCGTGGCAAAGATGCACACCGTGAACGGCTACGTTAACAGGAGCCTACCCGGCCTG
ATCGGCTGCCACCGCAAATCGGTCTACTGGCATGTGATCGGAATGGGCACAACGCCGAGGTCCAC
AGTATCTTCTCGAGGGCCACACTTCTGGTCCGGAAATCACGCCAGGCCAGCCTGGAGATCAGC
CCCATAACCTTCTGACGGCGCAGACCTACTCATGGATCTGCCCTGGTCCCTGTTCTGCCAC
ATTCGTCACCAGCACGATGGGATGGAAGCATATGTGAAAGTGGACTCCTGCCCGAGGAACCC
CAGCTTAGGATGAAGAACATGAGGAGGCGAGGACTACGACGATGACCTTACCGATTAGAAATG
GACGTTAGTACGCTTGACGACAACTCTCATACAGATTGCTCCGTGCCAAGAAC
CACCTAACGACTTGGGTGCACTACATCGGCCGAGGAGGAGCTGGATTATGCTCCCTGGT
CTGGCCCCCGACGACCGCAGCTACAAGAGCCAGTACCTGAATAACGGGCCAGCGCATGGCCGG
AACTACAAGAACGACTGGGTTCATGGCTTACACGGACGAGACCTCAAGACCCGGAGGCTATCCAG
CATGAGAGCGGCATCTGGGCCCTCCTGTACGGCAAGTGGAGACACACTGCTGATCATCTTC
AAGAACCGGGGAGCAGGCCCTACAACATCTACCCACGGCATTACCGATGTCGGCCCTGTTGAC
AGCCGACGGCTGCCAAGGGCGTGAAGCACCTGAAGGACTTTCGATCTGCCGGCGAGATCTTC
AACTACAAGTGGACTGTGACCGTGGAGGATGGCCGACCAAGAGCGATCCGGCTGCCGTACCG
TACTACTCCAGCTTGCAATATGGAGCGCAGCTCGTAGCGGCTTGATTGCCCTCTGCTGATC
TGCTACAAGGAGTCCGTGGACAGAGGGGAATCAGATCATGAGTACAAGAGGAACGTGATCCTG
TTCTCCGTGTTGACGAAAACCGCAGCTGGTATCTCACCGAGAAATATCCAGCGCTCTGCCAAC
CCGGCCGGTGTGCAGCTGGAGGACCCGAGTTTCAGGCCAGCAACATCATGCAATTCTATCAACGG
TATGTGTTGATTCCCTGCAGCTCTCAGTGTCTGCAACGGTGCCTACTGGTATATCCTCAGC
ATTGGGGCACAGACCGACTTCCCTGAGCGTGTCTTCTCCGGGTATACCTCAAGCACAAGATGGT
TACGAGGATACCTGACCCCTGTTAGCGCGAAACCGTGTATGTCTATGGAGAACCC
GGGCTCTGGATCCTGGCTGCCATAACTCCGACTTCCGCAACCGCGGAATGACCGCGCTCTGAAA
GTGTCGAGTTGTGACAAGAACACCGCGACTATTACGAGGACAGTTACGAGGACATCTCTGCGTAC
CTCCTTAGCAAGAATAACGCCATCGACGCCAGATCCTCAGGCCAGAACCCCCAGTGTGAAGAGG
CATCAGCGGGAGATCACCCGACGCCCTGCAGTGGATCAGGAGGAGATTGATTACGACGACACG
ATCAGTGTGGAGATGAAGAAGGAGGACTTCGACATCTACGACGAAGATGAAACCCAGTCCCTCGG
TCCTCCAAAAGAAGACCCGGCACTACTCATGCCGTGTGGAACGCCGTGGAGCTATGGAATG

(Continued)

Figure 17A

TCTTCTAGCCCTCACGTTTGAGGAACCGCGCCCAGTCGGGCAGCGTCCCCAGTCAAGAAAGTG
GTGTCAGGAGTTCACCGACGGCTCCTCACCCAGCCACTTACCGGGCGAGCTCAATGAACAT
CTGGGCCTGCTGGGACCCCTACATCAGGGCTGAGGTGGAGGACAACATCATGGTACATTCCGGAAT
CAGGCCAGCAGACCATAACAGTTCTACAGTCAGTCACTCATCTCTAACGAGGAGGACCAGCGCCAGGG
GCTGAACCCCCGTAAGAACCTCGTGAAGCCAACGAAACAAAGACCTACTTCTGAAAGGTCCAGCAC
CACATGGCACCTACCAAGGAGCAGTTCGATTGCAAGGCCTGGGCCTACTTCTCCGACGTGGACCTG
GAGAAAGATGTGCACAGGGCCTGATTGGCCCTCTGCTGGTGTGACACGAACACACTCAACCC
GCACACGGCGGCAGGTCACTGTGCAAGGAAATTGCCCTGTTTACCATCTTGATGAGACGAAG
TCCTGGTATTTCACCGAAAACATGGAGAGGAAC TGCCGCGCACCCCTGCAACATCCAGATGGAAGAT
CCGACATTCAAGGAGAACTACCGGTTCCATGCCATCAATGGCTACATCATGGACACCCCTGCCTGGC
CTCGTGATGGCCAAGACCAGCGTATCGCTGGTATCTGCTGTCATGGCTCCAACGAGAACATC
CATAGTATCCACTTCAGGGCATGTCCTCACGGTAGGAAAAAGGAGGAGTACAAGATGGCACTG
TACAACCTCTATCCC GGCGTTCGAGACCGTGGAGATGCTGCCCTCCAAGGCCGCATCTGGAGA
GTGGAATGCCGTATCGGCAGCACCTCCACGCTGGATGTCACGCTGTTCTCGTTACAGCAAT
AAGTGCCAGACCCCTCTGGGATGGCGAGCGGCCACATCCCGCACTTCCAGATTACGCCAGGGC
CAGTACGGTCAGTGGGCTCCAAAGCTGGCCCTCTGCACTACTCCGGATCCATCAACGCCCTGGTCC
ACCAAGGAACCGTTCTCCTGGATCAAAGTAGACCTGCTAGCCCCATGATCATTCAGGCATCAAG
ACACAAGGCGCCGACAGAAAGTCTCGAGCCTCTATATCTCCAGTTCATCATCATGTATAGCCTG
GACGGAAAGAAGTGGCAGACTTACCGGAAAACCTGACAGGGACCCCTGATGGTATTCTCGGTAAAC
GTGGACAGCTCCGGAAATCAAGCACAACATCTCAACCCACCCATTATGCCCGCTACATCCGCTG
CACCCCACTCACTATAGCATTAGGTCCACCTCGGAATGGAGCTCATGGCTGTGACCTGAACAGC
TGTAGCATGCCCTCGGCATGGAGTCTAAGCGATCTCGACAGCAGATAACGGCATCATCCTAC
TTTACCAACATGTCGCTACCTGGTCCCCCTCCAAGGCCCAGCTCCACCTGCAAGGGAGATCCAAC
GCCTGGCGGCCACAGGTCAACAATCCAAGGAGTGGCTGCAAGTGGACTTCAGAAAATATGAAA
GTCACCGGAGTGGCAGCACACAGGGAGTGAAGTCTCTGCTGACCGATGTACGTGAAGGAGTTCTC
ATCTCCAGTTCGCAGGATGGCCACCAGTGGACGTTCTCCAAAACGGTAAAGTCAAAGTCTTC
CAAGGGAAACCAGGACAGCTTACACCCGTCGTGAACCTCCCTGGACCCCCCGCTTCTCACTAGATAC
CTCCGCATCCACCCCTCAGAGCTGGGTGCACCGAGTGGCTGCGCATGGAGGTTCTGGGGTGTGAA
GCCCAGGACCTGTACTAA (SEQ ID NO:14)

Figure 17B

CS10-FL-NA

ATGCAGATTGAGCTCTCACCTGTTCTCGCTTCTGCCACACGC
AGGTACTATTGGGAGCAGTGAAGTGGATTACATGCAGAGTGACCTGGTAACCTCT
GTGGACGCTCGTTCCACCTAGAGTCCAAAGTCCTCCCTCAACACCTCAGGGTCTACAAG
AAAACGCTTTGTGGAGTTCACTGACCACCTCTCAACATTGCCAACCAGACCCCCCTGGATG
GGATTGCTGGGACCCACAATACAAGCAGAAGTCTACGACACGGTGGTATTACCTGAAGAACATG
GCGTCACACCCTGTTCACTTCACGCTGGGGTCAGTTATGGAAAGCCTCAGAGGGTGGGAA
TACGATGATCAAACCAAGCCAGAGGGAGAAGGAAGATGACAAGGTCTTCTGGGGTAGCCATACC
TATGTTGGCAGGTGCTGAAAGAGAATGGGCCTATGGCCTCTGATCCCTTGCTCACATACTCT
TACCTGAGTCACGTCGACCTGGTGAAGAGACCTGAATAGCGGTCTGATTGGTCAGTGCTGTTGT
AGAGAGGGAGTTGGCAAGGGAGAAAACTCAGACTCTCCACAAGTTATCCTCTGTTGCTGTG
TTCGACGAGGGCAACTCTGGCACTCTGAAACAAGAACACTCCCTGATGCAGGACAGAGATGCTGCA
TCTGCAAGGGCTTGGCCAAAAATGCACACAGTGAACGGCTATGTGAATCGATCACTGCCAGGACTG
ATAGGCTGTCACTGCAAGTCAGTGTATTGGCACGTTATGGGATGGGAACAACCTCAGAAGTGCAC
AGCATTCTCCTGGAGGGCCACACTTCTGGTCCGAATCATAGACAGGCCAGCCTGAGATCAGC
CCAATCACCTTCTGACTGCCAACCTTGCTGATGGATCTGGACAGTTCCCTGTTGTCAC
ATCTCCCTCCACCAACATGACGGATGGAGGCTTATGTGAAGGTCGATAGCTGCCGGAGGAACCA
CAACTGAGGATGAAGAACAGAACAGAAGAGGCAGAGGACTATGACGACGATCTGACTGACAGTGAATG
GACGTGTTGGTTCGCTGACGATGACAATTCTCCTCATTTATCCAGATCCGTTCCGGCCAAGAAG
CACCCCAAGACTTGGGTTCAATTACATCGCTGCTGAGGAGGAGTGGGACTACGCCCTGGTG
TTGGCCCCAGACGATCGCTCATACAAGAGCCAGTACCTTAACAATGGTCCACAAAGGATGCCCGG
AAGTACAAGAAGGTTAGATTATGGCTTATACCGACGAGACTTTAAACTAGGAAGCAATTCA
CATGAAAGTGGCATTCTGGACCCCTGCTGATGGCAGGTTGGCGACACCCTGCTGATTATCTT
AAGAACCCAGGCAAGGCCGCCCCCTACAACATCTACCCGCACGGCATAACCGATGTACGCCCTGTAC
AGTCGAGACTTCTAAAGGGGTGAAACACCTGAAGGACTTCCCAATTCTGCCGGGGAGATCTTC
AAAGTATAAAATGGACCGTGACGGTTGAGGATGGTCCCACAAAGTCGATCCGAGATGCCCTACCCGA
TATTATTCCAGCTCGTGAACATGGAAAGGGACCTGGCAGCGGGCTGATTGGCCCACTGCTGATT
TGTACAGGAGTCTGTCGATCAAAGAGGAAACCAAATAATGAGCGACAAACGTAACGTACCTG
TTCAAGCGTCTTGATGAGAAATAGAAGCTGGTACCTCACAGAAAATATTCAAGCGTTCTGCC
ACCCGCAGGCGTCCAGCTGGAGATCCCGAGTTCAAGCCTCAAACATCATGCAAGCATCAACGG
TACGTATTGATAGCCTGCAGCTGTCCTGTCATGAAGTGGCATATTGGTACATCCTGAGT
ATCGGGCGCAGACCGACTTCTGAGCGTGTCTTCTGGATACAGTTCAAACACAAAATGGTC
TATGAAGATACCCCTGACTCTGTTCCATTCTCAGGAGAGACAGTCTTATGAGTATGGAAAATCT
GGACTGTGGATCTGGCTGTCACAATTCTGATTTCGGAACAGAGGCATGACAGCCCTGCTTAAA
GTGAGCTCATGCGACAAGAACACCGGTGATTACTACGAAGATAGCTATGAGGACATCAGTGC
TTGCTCTCCAAAGAACACCGTATCGAGCCACGGTCTTCACTGAGAACAGAGGAAATCG
ACATCAGCGCAGAACACACCGCACAACCCPTCAGTCAGACCAAGAGGAAATCGACTACG
ATCTCTGAGATGAAGAAGGAGGATTGACATTTACGACGAGGACGAGAATCAGTCCCCAAGG
AGCTTCAAGAACACAGACACTATTCTGAGCGCCGTGGAGCGACTGTGGACTACGGCATG

(Continued)

Figure 18A

TCTAGCTCTCCGCATGTACTTAGAAATAGGGACAAAGCGGATCCGTGCCCTCAGTTAAGAAAAGTT
GTCTTCAGGAGTTACAGATGGCTCCTTCACCCAGCCCTGTATCGCGGGAACTCAATGAACAC
CTGGGCCTCCTGGGTCTTATATTAGGGCCGAAGTCGAGGACAATATCATGGTGACCTTAGGAAC
CAGGCATCTAGACCTTACTCTTCTACTCCTCCTGATATCCTATGAGGAGGACCAGCGGCAAGGC
GCTGAGCCTCGGAAGAACCTTGTGAAGCAAATGAAACCAAAACATACTTTGGAAAGTCAGCAC
CACATGGCTCCCACGAAGGACGAATTGACTGTAAAGCCTGGCCTACTTCTCAGATGTAGATCTC
GAGAAAGACGTGCACTCAGGGCTCATTGGTCCCTCCTGGTCTGTOATACTAATACCCCTCAATCCA
GCACACGGACGTCAAGGTAACCGTCCAGGAATTGCCCTGTTACCATTTGATGAGACTAAA
TCCTGGTACTTTACCGAAAACATGGAGAGGAATTGCAGAGCCCCATGCAACATCCAGATGGAGGAC
CCTACCTCAAAGAGAACTATCGCTTCCATGCCATTAACGGTTACATTATGGATACTCTCCAGGA
CTTGTGATGGCACAGGATCAGCGGATAAGATGGTATCTGTTGAGCATGGCCTCAACGAGAATATT
CACAGCATCCATTCTCCGGTCACGTGTTACAGTGAGAAAGAAAGAGTACAAGATGGCTCTG
TATAATCTCTATCCAGGGGTATTGAAACGGTGGAGATGTTGCCTAGCAAGGCCGCATTGGCGA
GTAGAATGCCATTATCGGGGAACATCTGCATGCCGAATGAGCACGCTCTCCTGGTGTATAGTAAC
AAAGTGCCAGACTCCGCTGGGCATGGCATCTGGCCATATAACGGGACTTCAGATTACGGCTAGCGGG
CACTATGGGCAGTGGGCACCCAAACTTGCAGCAGTGCACATTAGGCTCTATCAATGCATGGTCC
ACCAAGGAACCCCTCTTGGATTAAGGTGGACCTTTGGCGCCCATGATAATCCATGGATCAAA
ACCCAGGGCGCTCGTCAGAAATTCTCATCACTCTACATCTCAGTCATAATAATGTATTCACTG
GATGGGAAGAAATGGCAGACTTACAGAGGAAACAGCACCGGGACGCTGATGGTGTCTTGGCAAC
GTGGACAGCAGCGGCATCAAACACAATCTCAATCCTCCATTATTGCCCGTTATATTAGACTG
CATCCCACACTCACTCTATACGCAGCACACTAGGATGGAGCTATGGGATGCGACCTGAACAGT
TGTAGTATGCCCTGGGGATGGAGTCAAAGCTATAAGCGACGCACAAATTACAGCTAGCTTAC
TTTACGAATATGTTGCCACGTGGAGCCCAAGCAAAGCCGGCTGATTTGCAGGGTCGGAGTAAT
GCTTGGCGCCCAAGGTGAATAACCTAAGGAATGGTGTGCAAGTAGATTCCAGAAAACATATGAAG
GTAACCGGCCTCACTACACAGGGAGTCAAGTCCTCTTGCACCTCTATGTACGTCAAGGAGTTCTG
ATTAGCAGCAGTCAGGATGGGCACCAATGGACACTGTTCTCCAGAATGGAAAGTTAAAGTATT
CAGGGTAACCAGGACTCCTTACACCTGTGGTGAATAGCCTCGACCCACCCCTGCTGACACGATAAC
CTCCGCATCCACCCCTCAGTCTGGGTGATCAAATTGCCCTGCGAATGGAGGTGTTGGGATGCGAA
GCTCAGGACCTCTACTGA (SEQ ID NO:15)

Figure 18B

CS11-FL-NA

ATGCAGATCGAACTCTACTTGCTTCTCCTGTGCCCTCTGAGGTTCTGCTTCTGCCACTCGC
CGATATTACCTCGGGGCCGTGGAGTTGAGTTGGACTACATGCAATCAGATCTGGCGAACCTCCCT
GTGGATGCCGATTCCCACCGCGCGTCCCCAAGTCCTTCCCATTAAACTCTGTGGTGTACAAG
AAGACATTGTTGTGGAGTTACCGATCACCTGTTCAACATGCCAACCGCGGCCCATGGATG
GGTCTGCTTGGGCCACCATTCAAGCGGAGGTCTATGATAACAGTGGTATAACCGCTTAAGAACATG
GCGAGCCACCCAGTGTCTGCATGCCGTTGGTGTATCATATTGAAAGGCCAGCGAAGGAGCGGAG
TACGATGACCAGACCTCTCAGAGAGAAGGAAGACGATAAGGTTTCTGGCGGAAGTCATACA
TATGTATGGCAGGTCTGAAAGAGAATGGGCCATGGCTTCTGACCCCTTGTCTTACCTATAGT
TATCTGAGCCACGTGGACCTGGTCAAGGACCTCAACAGTGGTCTGATTGGGCTCTGCTTGTGTTG
AGAGAGGGTAGCTTGGCTAAGGAGAAAACCCAAACACTCCATAAGTCATTTGCTGTTCCGGTG
TTCGACGAGGGAAAGAGTTGGCACAGCGAAACAAAGAATTCACTGATGCAAGACAGGGACGCCGCT
TCCGCAAGGGCTGGCTAAGATGCAACGGTGAATGGGTATGTGAACCGGAGCCTCCGGGCTG
ATCGGGTGCATCGCAAGTCTGTTACTGGCACGTCAATTGGAATGGGACAACCCAGAGGTACAT
AGTATATTCTGAAGGCCACACGTTCTCGTACCGAACCCGACAGGCTCCCTGGAGATAAGC
CCCATTACCTTCTGACCGCTCAGACTCTGCTGATGGACCTTGGCCAGTTCTGTTCTGCCAT
ATTAGCAGCCACCAGCACGGTATGGAAGCATACTGAAAGTCGATAGCTGTCCTGAGGAGCCT
CAGCTCAGAATGAAAGAACACGAGGGCCGAAGACTATGACGATGACCTTACAGATTCCGAGATG
GACGTGGTGCCTTGACGACGATAACAGTCCTAGTTCAATTCAAATCAGATCCGTAGCCAAAAAG
CATCCAAAGACATGGGTGCATTACATTGCAAGCCGAAGAGGAGGATTGGGATTATGCCCTTGTT
CTGGCTCCAGATGACAGGAGCTATAAGTCCCAGTACTTGACACAACGGCCACAGCGAATCGGTAGA
AAATATAAGAAGTAAGATTCACTGGCTACACTGACGAAACATTAAACCCAGGAAAGCTATCCAA
CACGAATCTGAAATTCTCGCCCTCTGCTCTACGGTGAGGTGGGGACACCTTGCTGATCATTTT
AAAAATCAGGCATCCAGGCCTACACATACCCCCATGGCATCACCAGTGTCCGCCCCGTGTAT
TCCAGAAGACTCCCCAAGGGAGTGAACACATCTGAAAGATTTCCTCATCTGCCGGGAGATCTT
AAATACAAATGGACTGTGACTGTAGAGGACGGGCTACAAATCAGACCCACGGTGCCTGACAAGG
TATTACAGTAGCTCGTCAACATGGAACGGGACCTCGCCAGCGGACTCATTGGCCACTGTGATC
TGTACAAAGAGTCAGTGGATCAGAGGGAAATCAGATCATGAGCGATAAGAGAAACGTTATCCTG
TTTAGTGTCTCGACGAGAACGGTCTTGGTACCTTACTGAGAACATCCAGAGGTTCTGCCGAAT
CCGGCTGGCGTTAGCTCGAGGACCCAGAGTTCCAGGCCAGTAATATAATGCACTCAATCACGGT
TATGTGTTGATAGCCTGCAGCTGAGCGTCTGCCACAGGTTAGCCTATTGGTACATATTGTCC
ATCGGGGCTCAGACCGATTTCCTGTCGTGTTAGCGGGTACCTTAAACATAAAATGGTC
TATGAAGACACCCCTGACCCCTGTTCCATTCTCCGGTGAGACTGTGTTCATGTCCATGGAGAACCCA
GGGCTGTGGATCTGGGTGTCACAATAGTGACTTTAGGAATCGGGAAATGACGGCACTGCTGAAG
GTGAGTTCTTGCATAAAACAGGAGATTACTATGAGGAGATTGAGGATAGTTACGAGGATATCAGTGCCTAT
CTGCTTCAAAAAACACGCAATTGAGCCCCGGCTTTCTCACAAAACCCCCGGTGTGAAGCGC
CACCAAGCGCAAATTACCCGGACAACCTGCACTGCCGACCAGGAGGAATGATTATGACGATACT
ATCAGTGTAGAAATGAAAAGGAGGATTGATATTACGACGAAGACGAGAACAGTCTCCGCGA

(Continued)

Figure 19A

AGTTTCAGAACAAAACCGGACACTACTTATAGCTGCCGTGGAACGACTCTGGATTATGGCATG
TCCTCCAGCCCTCATGTCCTTAGGAATCGAGCGCAGAGTGGCTCTGTCCTCAGTTCAAAAGGTT
GTGTTCCAGGAATTCACCGACGGCTCATTTACCCAGCCGCTGTACAGAGGCAGACTCAACGAACAC
CTTGGGCTGCTTGGGCCATATAATTCGAGCAGAGGTGGAAGATAATATCATGGTAACCTTAGAAC
CAGGCGTCAAGACCCATTCCCTCATACAGTCTCTGATCAGCTACGAGGAGGACCAAAGACAGGGA
GCTGAACCCAGGAAGAACTTGTGAAACCTAATGAGACCAAGACCTACTTCTGGAAGGTCCAGCAC
CATATGGCCCCAACTAAAGATGAATTGATTCGAAAGGCTGGCTTATTCAGCGACGTGGATCTC
GAAAAGGATGTGCACAGCGGGTTGATCGGACCGCTTTGGTGTGCCACACAAATACCCCTCAATCCT
GCCACGGGCGGCAGGTACAGTTCAAGAGTTGCACTCTCTTACAATATTGACGAGACAAAG
TCATGGTATTTACAGAGAATATGGAGAGAATTGTCGCGCACCTTGCAACATTCAGATGGAGGAC
CCCACATTAAAGGAGAATTACAGATTTCATGCTATCAATGGTACATTATGGATACTCTGCCTGGT
CTGGTCATGGCCCAAGGATCAGCGCATAAGGTGGACTTGTGAGCATGGATCTAATGAGAATATA
CACAGCATTCACTCAGTGGCACGTTTACTGTTAGAAAGAAGGAGGAGTACAAAATGGCGCTC
TACAACCTTACCCGGGTGTGTTGAGACAGTGGAGATGCTGCCAAGCAAGGCAGGCATCTGGAGG
GTTGAGTGTCTTATTGGGAGCATCTGCATGCTGGAATGTCCACCCCTTTCTGTGTACAGCAAT
AAAGTGCCAGACACCGCTTGCATGGCCAGCGGGCACATTAGGGACTTCAAGATAACTGCCAGTGG
CAGTACGGCCAGTGGCTCCAAAGCTTGCAGACTCCACTACTCCGGAAGCATAACGCATGGAGC
ACCAAGGAACCCCTCTTGGATTAAGGTGGACCTGCTGGCCAAATGATCATTCAACGGATAAAA
ACCCAAGGGGCACGACAGAAATTTCATCTTGATATTAGTCAGTTATCATCATGTACAGCTG
GATGGAAAGAAGTGGCAGACGTACAGGGCAATTCTACAGGAACACTTATGGTGTGTTGGAAAT
GTCGATTCCAGCGGGATCAAACATAACATCTCAATCCTCTTATTATGCCGATATATCCGCTG
CACCCATCGCATTACTCCATCAGGTCCACATTGAGAATGAACTGATGGGTGCGACCTGAATAGT
TGTAGTATGCCACTGGCATGGAGTCTAAAGCCATCAGCGATGCACAGATCACTGCCAGCTCTTAC
TTCACCAACATGTTGCAACTTGGTCCCCCTCTAAAGCTCGCCTGCATCTGCAGGGACGOTCAAAT
GCATGGCGACCACAGGTGAAACATCCAAAAGAGTGGCTCAGGTGACTTCAAGACATGAAG
GTAACAGGAGTGACAACCCAGGGTGTAAAAAGCCTCCTACGACTATGTACGTTAAGGAGTTCTG
ATTTCTAGCTCCACAGGACGGACACCAGTGGACTCTGTTCTCCAGAACGGCAAAGTGAAGGTATTT
CAGGGAAACCAGGATTCTTTACCCGGTAGTGAATAGCCTGGATCCACCGTTGCTGACCCGCTAT
CTGAGAATTCCATCCACAATCCTGGGTGCATCAGATTGCCCTCCGGATGGAAGTGCTCGGCTGTGAA
GCTCAGGATCTGTATTAG (SEQ ID NO:16)

Figure 19B

CS40-FL-NA

ATGCAAATAGAGCTCTCACCTGCTTCTTGCGATTCTGCCATTGTGAGCTGGCACCAGA
AGATACTACCTGGGTGCAGTGGAACTGTCACTGGACTATATGCAAAGTGAATCGGTGAGCTGCCT
GTGGACGCAAGATTCCTCCTAGAGTGCCAAATCTTTCCATTCAACACCTCAGTCGTGTACAAA
AAGACTCTGTTGTAGAATTACCGGATCACCTTCAACATCGCTAACCCAAGGCCACCCCTGGATG
GGTCGCTAGGTCTACCATCCAGGCTGAGGTTATGATACAGTGGTCATTACACTTAAGAACATG
GCTTCCCACATCGTCAGTCTTCATGCTGTTGGTGTATCCTACTGGAAAGCTCTGAGGGAGCTGAA
TATGATGATCAGACCGAGTCAAAGGGAGAAAAGAAGATGATAAGTCTTCCCTGGTGAAGCCATACA
TATGTCGGCAGGTCTGAAAGAGAATGGTCCAATGGCCTCTGACCCACTGTGCCCTACCTACTCA
TATCTTCTCATGTGGACCTGGTAAAAGACTTGAATTCAAGGCTCATTGGAGGCCACTAGTATGT
AGAGAAGGGAGTCTGGCAAGGAAAAGACACAGACCTGCACAAATTATACTACTTTTGTGTA
TTTGATGAAGGGAAAAGTTGGCACTCAGAAACAAAGAACTCCTGATGCAAGGATAGGGATGCTGCA
TCTGCTCGGGCCTGGCCTAAAATGCACACAGTCAATGGTTATGTAACAGGTCTCTGCCAGGTCTG
ATTGGATGCCACAGGAAATCAGTCTATTGGCATGTGATTGGAATGGGCACCACTCCTGAAGTGCAC
TCAATATTCCCTCGAAGGTACACACATTCTGTGAGGAACCATGCCAGGCTCTGGAAATCTG
CCAATAACTTCCCTACTGCTAAACACTCTTGATGGACCTTGGACAGTTCTACTGTTTGTAT
ATCTCTTCCCACCAACATGATGGCATGGAAGCTTATGTCAAAGTAGACAGCTGTCCAGAGGAACCC
CAACTACGAATGAAAATAATGAAGAACGGAAAGACTATGATGATGATCTTACTGATTCTGAAATG
GATGTGGTCAGGTTGATGATGACAACCTCCCTCCTTATCCAAATTGCTCAGTTGCCAAGAAG
CATCCTAAAACCTGGGTACATTACATTGCTGCTGAAGAGGAGACTGGACTATGCTCCCTAGTC
CTCGCCCCGATGACAGAAGTTATAAAAGTCAATATTGAAACAATGGCCTCAGCGGATTGGTAGG
AACTACAAAAAGTCCGATTTATGGCATAACACAGATGAAACCTTTAAGACTCGTGAAGCTATTCA
CATGAATCAGGAATCTGGACCTTACTTTATGGGAAGTTGGAGACACACTGTTGATTATATT
AAGAATCAAGCAAGCAGACCATATAACATCTACCCCTACGGAACTGATGTCCTGCTTGTAT
TCAAGGAGATTACCAAAAGGTGAAACATTGAAAGGATTTCACATTGCTCAGGAGAAATATT
AAATATAATGGACAGTGAATGACTGAGAAGATGGCAACTAAATCAGATCCTCGGTGCCTGACCCGC
TATTACTCTAGTTGTTAATATGGAGAGAGATCTAGCTTCAGGACTCATTGGCCTCTCCTCATC
TGCTACAAAGAATCTGTAGATCAAAGAGGAAACAGATAATGTCAGACAAGAGGAATGTCACTCTG
TTTCTGTATTGATGAGAACCGAAGCTGGTACCTCAGAGATAACCTCAAAACACAAAATGGTC
CCAGCTGGAGTGCAGCTGAGGATCCAGAGTTCCAAGCCTCAACATCATGCACAGCATCAATGGC
TATGTTTGATAGTTGCAGTTGTCAGTTGTTGCATGAGGTGGCATACTGTCACATTCTAAGC
ATTGGAGCACAGACTGACTTCCTTCTGTCTTCTGATATAACCTCAAAACACAAAATGGTC
TATGAAGACACACTCACCCATTCCATTCTCAGGAGAAACTGTCCTCATGTCATGGAAACCCA
GGTCTATGGATTCTGGGTGCCACAACTCAGACTTCGGAACAGAGGCATGACCGCCTACTGAAG
GTTCTAGTTGACAAGAACACTGGTGAATTACGAGGACAGTTATGAAAGATAATTCAAGC
TTGCTGAGTAAAACAATGCCATTGAACCAAGAAGCTCTCCAGAATCCACCCAGTCTGAAACGC
CATCAACGGAAATAACTCGTACTACTCTTCAGTCAGATCAAGAGGAAATTGACTATGATGATACC
ATATCAGTTGAAATGAAGAAGGAAAGATTTCACATTATGATGAGGATGAAAATCAGAGCCCCGC
AGCTTCAAAAGAAAACACGACACTATTTATTGCTGCAGTGGAGAGGCTCTGGATTATGGATG

(Continued)

Figure 20A

AGTAGCTCCCCACATGTTCAAGAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTT
GTTTCCAGGAATTACTGATGGCTCCTTACTCAGCCCTTATACCGTGGAGAACTAAATGAACAT
TTGGGACTCCTGGGCCATATAAGAGCAGAAGTTGAAGATAATATCATGGTAACTTICAGAAAT
CAGGCCTCTCGCCCTATTCCCTCTATTCTAGCCTTATTCTTATGAGGAAGATCAGAGGCAAGGA
GCAGAACCTAGAAAAAACTTGTCAAGCCTAATGAAACAAAACACTACTTTGAAAGTGCAACAT
CATATGGCACCCACTAAAGATGAGTTGACTGCAAAGCCTGGCTTATTCCTGATGTTGACCTG
GAAAAAGATGTGCACTCAGGCCTGATTGGACCCCTCTGGTCTGCCACACTAACACACACTGAACCC
GCTCATGGGAGACAAGTGACAGTACAGGAATTGCTCTGTTTACCATCTTGATGAGACCAA
AGCTGGTACTTCACTGAAAATATGAAAGAAACTGCAGGGCTCCCTGCAATATCCAGATGGAAGAT
CCCACTTAAAGAGAATTATCGCTTCCATGCAATCAATGGCTACATAATGGATACACTACCTGGC
TTAGTAATGGCTCAGGATCAAAGGATTGATGGTATCTGCTCAGCATGGCAGCAATGAAAACATC
CATTCTATTCAATTCACTGGACATGTGTTCACTGTACGAAAAAAAGAGGAGTATAAAATGGCACTG
TACAATCTCTATCCAGGTGTTTGAGACAGTGGAAATGTTACCATCCAAGCTGGAATTGGCG
GTGGAATGCCTTATTGGCGAGCATCTACATGCTGGGATGAGCACACTTTCTGGTGTACAGCAAT
AAAGTGTCAAGACTCCCTGGGAATGGCTCTGGACACATAGAGATTTCAGATTACAGCTTCAGGA
CAATATGGACAGTGGGCCAAAGCTGGCCAGACTTCATTATTCCGGATCAATCAATGCCTGGAGC
ACCAAGGAGGCCCTTTCTGGATCAAGGTGGATCTGTTGGCACCAATGATTATTCAAGGCATCAAG
ACCCAGGGTGCCTCGTCAGAAGTCTCCAGCCCTACATCTCAGTTATCATCATGTATAGTCTT
GATGGAAGAAGTGGCAGACTTATCGAGGAATTCCACTGGAACCTTAATGGCTTCTTGGCAAT
GTGGATTTCATCTGGATAAAACACAATATTAAACCCCTCAATTATTGCTCGATACATCCGGTTG
CACCCAACTCATTAGCATTGCCACTGGTCTCCPTCAAAAGCTCGACTTCACCTCCAAGGGAGGAGTAAT
GCCTGGAGACCTCAGGTGAATAATCCAAAAGAGTGGCTGCAAGTGGACTTCCAGAACAGACAATGAAA
GTCACAGGAGTAACTACTCAGGGAGTAAAGCAATATCAGATGCACAGATTACTGCTTACATCCTAC
TTTACCAATATGTTGCCACCTGGTCTCCPTCAAAAGCTCGACTTCACCTCCAAGGGAGGAGTAAT
GCCTGGAGACCTCAGGTGAATAATCCAAAAGAGTGGCTGCAAGTGGACTTCCAGAACAGACAATGAAA
GTCACAGGAGTAACTACTCAGGGAGTAAAGCAATATCAGATGTATGTGAAGGGAGTCCCTC
ATCTCCAGCAGTCAGATGCCATCAGTGGACTCTTTTTCAGAATGGCAAAGTAAAGGTTTT
CAGGGAAATCAAGACTCCTTCACACCTGTGGTAACCTCTAGACCCACCCTACTGACTCGCTAC
CTTCAATTCAACCCCCAGAGTTGGGTGCACCAAGATTGCCCTGAGGATGGAGGTTCTGGGCTGCGAG
GCACAGGACCTCTACTGA (SEQ ID NO:17)

Figure 20B

CH25-FL-NA

ATGCAGATCGAGCTGTCCACATGCTTTCTGTGCCCTGCGGTTCTGCTTCAGGCCACCCGG
CGGTACTACCTGGCGCCGTGGAGCTGTCTGGACTACATGCAGAGCAGCTGGCGAGCTGCC
GTGGACGCCGGTCCCCCCCAGAGTGCCAAGAGCTCCCCCTCAACACCAGCGTGGTGTACAAG
AAAACCTGTTCTGGAGTCACCGACCACCTGTTAACATGCCAACCCCCAGGGCCCCCTGGATG
GGCCTGCTGGCCCCACCATTCCAGGCCAGGTGTACGACACCGTGATCACCTGAAGAACATG
GCCAGCCACCCCGTGAGCTGCACGCCGTGGCGTGAGCTACTGAAAGGCTCCGAGGGCGCGAG
TACGACGACCAGACCAGCCAGCGGGAGAAAGAGGACGACAAAGTCTTCCTGGCGGCAGCCACACC
TACGTGTGGCAGGTCTGAAAGAAAAGCCCATGGCCTCCGACCCCTGTGCCCTGACCTACAGC
TACCTGAGCCACGTGGACCTGGTGAAGGACCTGAACAGCAGGGCTGATTGGGCCCTGCTGGTCTGC
CGGGAGGGCAGCCTGGCCAAGAGAAAAACCCAGACCCCTGCACAAGTTCATCCTGCTGTTGCCGTG
TTCGACGAGGGCAAGAGCTGGCACAGCGAGACCAAGAACAGCCTGATGCAGGACCGGGACGCCGCC
TCTGCCAGAGCCTGGCCAAGATGCACACCGTGAACGGCTACGTGAACAGAACGCTGCCGGCTG
ATTGGCTGCCACCGAAGAGCGTGTACTGGCACGTGATGGCATGGCACACACCGCAGGCCATGGAAATCAGC
ACCATCTTCTGGAAGGGCACACCTTCTGGTGCAGAACCCACCGCAGGCCATGGAAATCAGC
CCTATCACCTTCTGACCAGCAGACACTGCTGATGGACCTGGCCAGTTCTGCTGTTGCCAC
ATCAGCTCTCACCAGCACGACGGCATGGAAAGCCTACGTGAAGGTGGACTCCTGCCCGAGGAACCC
CAGCTGGATGAAGAACAAACGAGGAAGCCGAGGACTACGACGACCTGACCCACAGCAGATG
GACGTGGTGGCTGACGACGACAACAGCCCCAGCTTACATCCAGATCAGAACGCTGGCAAGAAC
CACCCAAGACCTGGGTGACTACATGCCCGAGGAAGAGGACTGGACTACGCCCTGGT
CTGGCCCCCGACGACAGAACGCTACAAGAGCCAGTACCTGAACAATGGCCCCCAGCGGATGGCCGG
AACTACAAGAACGCTGGGTTCATGGCCTACACCGACGAGACCTTAAGACCCGGAGGCCATCCAG
CACGAGAGCGGCATCTGGCCCCCTGCTGTACGGCGAAGTGGCGACACACTGCTGATCATCTTC
AAGAACCAAGGCCAGCCGGCCCTACAACATCTACCCCCACGGCATCACCACGCTGCCGGCCCTGTAC
AGCAGGGCTGCCAAGGGCGTGAAGCACCTGAAGGACTTCCCCATCCTGCCGGCGAGATCTTC
AACTACAAGTGGACCGTGGACCGTGGAGGACGCCAACCAAGAGCGACCCAGATGCCCTGCCGG
TACTACAGCAGCTTCTGTAACATGGAACGGGACCTGGCTCCGGCTGATGGACCTCTGCTGATC
TGCTACAAAGAACGCGTGGACCGAGCAGCGGGCAACAGATCATGAGCGACAAGCGAACGTGATCCTG
TTCAGCGTGTGATGAGAACCGGTCTGGTATCTGACCGAGAACATCCAGCGGTTCTGCCAAC
CCTGCCGGGGTGCAGCTGGAAGATCCCAGTTCAGGCCAGCAACATCATGCACTCCATCAATGGC
TACGTGTTGACAGCCTGCAGCTGTCCGTGTCTGCACGAGGTGGCCTACTGGTACATCCTGAGC
ATCGGGGCCAGACCGACTTCTGAGCGTGTCTCAGCCGCTACACCTCAAGCACAAGATGGTG
TACGAGGACACCTGACCCCTGTTCCCTTCAAGCGGCGAGAACCGTGTCTGAGCATGGAAAACCC
GGCCTGTTGACAGCCTGCCACACAGCGACTTCCGGAACCGGGCATGCCCTGCTGAAG
GTGTCCAGCTGCCACAAGAACACCGGGCGACTACTACGAGGACAGCTACGAGGATATCAGCGCCTAC
CTGCTGTCAGAACACGCCATCGAGCCAGAACGCTTCAAGCAGAACCCCCCTGTGCTGAAGCGG
CACCAAGAGAGATCACCCGGACCACCCCTGCAGTCCGACCAGGAAGAGATCGATTACGACCGACACC

(Continued)

Figure 21A

ATCAGCGTGGAGATGAAAAAAGAAGATTCGACATCTACGACGAGGACGAGAACCAAGAGCCCCCG
TCCTTCCAGAAGAAAACCCGGCACTACTTTATCGCCGCCGTGGAGCGGCCTGTGGGACTACGGCATG
AGCAGCAGCCCCCACGTGCTGCCAACCGGGCCCAGAGCGGCAGCGTCCCCAGTTCAAGAAAGTG
GTGTTCCAGGAATTACCGCACGGCAGCTCACCCAGCCCCGTACCGGGCGAGCTGAACGAGCAC
CTGGGGCTGCTGGGGCCCTACATCAGGGCGAACGTGGAGGACAACATCATGGTACCTCCGGAAT
CAGGCCAGCAGACCCACTCCTCTACAGCAGCCTGATCAGCTACCAAGAGGACCAGCGGCAGGGC
GCTGAACCCCGAACGAACTTCGTGAAGCCAAATGAGACCAAGACCTACTCTGGAAAGTGCAGCAC
CACATGGCCCCCACCAAGGACGAGTCGACTGCAAGGCCCTGGCCTACTTCAGCAGCTGGATCTG
GAAAAGGACGTGCACTCTGGACTGATTGGCCCTCTGCTGGTGTGCCACACCAACACCCCTGAACCCC
GCCAACGGCCGGCAGGTGACCGTGCAGGAATTGCCCTGTTCTCACCATCTTCGACGAGGACCAAG
TCCTGGTACTTCACCGAGAACATGGAACGAAACTGCAAGGCCCTGCAACATCCAGATGGAAGAT
CCTACCTTCAAAGAGAACCTACCGGTTCCACGCCATCAACGGCTACATCATGGACACCCCTGCCTGGC
CTGGTGATGGCCCAGGACACAGGGATCCGGTGGTATCTGCTGTCCATGGCAGCAACGAGAACATATC
CACAGCATCCACTTCAGCGGCCACGTGTTCACCGTGAGGAAGAACAGACTACAAGATGCCCTG
TACAACCTGTACCCCGGGTGTTCGAGACCGTGGAGATGCTGCCAGCAAGGCCGCATCTGGCG
GTGGAGTGTCTGATCGGCGAGCACCTGCATGCCGGATGAGCACCCCTGTTCTGGTGTACAGCAAC
AAGTGCAGACCCCCCTGGCATGGCAGCGGCCACATCCGGACTTCCAGATCACCGCTCCGGC
CAGTACGGCCAGTGGCCCCCAAGCTGGCCGGCTGCACTACAGCGGCAGCATCACGCCCTGGTCC
ACCAAAGAGGCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCCCTATGATCATCCACGGCATTAAG
ACCCAGGGCGCAGGCAGAACGTTCAAGCTGACATCAGCCAGTTCATCATGTACAGCCTG
GACGGCAAGAAGTGGCAGACCTACCGGGCAACAGCACCCGGCACCCCTGATGGTGTCTCGGCAAC
GTGGACAGCAGCGGCATCAAGCACAACATCTCAACCCCCCATCATGCCCGGTACATCCGGCTG
CACCCCACCCACTACAGCATCAGATCCACCCCTGCGGATGGAACCTGATGGGCTGCGACCTGAACCTCC
TGCAGCATGCCCTGGCATGGAAAGCAAGGCATCAGGCCAGATCACAGCCAGCAGCTAC
TTCACCAACATGTTGCCACCTGGTCCCCCTCCAAGGCAGGTGCACTGCGAGGGCTGACCTG
GCCTGGCGGCCCTCAGGTGAACAAACCCAAAGAATGGCTGCAAGGTGGACTTCAGAAAACCATGAAG
GTGACCGGGCGTGACCACCCAGGGCGTGAAAAGCCTGCTGACCAGCATGTACGTGAAAGAGTTCTG
ATCAGCAGCAGCCAGGACGGCCACCAAGTGGACCCCTGTTCTTCAGAACGGCAAGGTGAAAGTGTTC
CAGGGCAACCAGGACTCCTTCACCCCGTGGTGAACCTCCCTGGACCCCCCCCCTGCTGACCCGCTAC
CTGCGGATCCACCCCACTGTTGGGTGACCAAGATGCCCTGAGGATGGAAGTGTGGATGTGAG
GCCUAGGATCTGTACTGA (SEQ ID NO:18)

Figure 21B

FVIII-FL-AA

mqielstcfff lollrfcfsa trryylgave lswdymqsdl gelpvdarfp prvpksfpfn
tsvvykktlf veftdhlfni akprppwmgl lgptiqaevy dtvvitlknm ashpvslhav
gvsywkaseg aeyddqtsqr ekeddkyfpg gshtywqv1 kengpmasdp lclystyish
vdlvkdlng ligallvcre gslakektqt lhkfillfav fdegkswhse tknslmqdrd
aasarawpkm htvngyvnrs lpgligchrk svywhvigmg ttpevhisifi eghtflvrnh
rzasleispi tfltaqtllm dlqqfllfch issqhgdgme ayvkvdscpe epqlrmknne
eaedydddlt dsemdvvrfd ddnspsfqi rsvakkhpkt wwhyiaaeee dwdyaplvia
pddrsyksqy lnngpqrigr kykkvrfmay tdetfktre aqhesgilgp llygevgdtl
liifknqasr pyniyphgit dvrplysr1 pkvkhkdf pilpgeifky kwtvtvedgp
tksdprctr yyssfvnmer dlasgligpl licykesvdq rgngimsdkr nvilfsvfde
nrswylteni qrflpnpagv qledpefqas nimhsingyv fdslqlsvcl hevaywyils
igaqtdflsv ffsgytfkhh mvyedtltlf pfsgetvfms menpglwlg chnsdfrnrg
mtallkvssc dkntgdyyed syedisayll sknnaieprs fsqnsrhpst rqkqfnatti
pendiektdp wfahrtppmk iqnvsssd11 mlrlqsptph gls1sdlgea kyetfsddps
pgaidsnnsl semthfrpql hhsgdmvftp esglqlrlne klgtaatei kkldfkvsst
snnlistips dnlaagtdnt ss1gppsmpv hydsqldttl fgkkssplte sggplslsee
nndskllesg lmnsqesswg knvssutesgr lfkgkrahgp alltkdnalf kvsisllktn
ktsnnsatnr kthidgpsil ienspsvwgn ilesdtefkk vtplihdrml mdknatalrl
nhmsnkttss knmemvqqkk egpippdaqn pdmsffkmlf lpesarwiqr thgknslnsg
qgpspkqlvs lgpeksvegg nlfseknnvv vvgkgeftkdv glkemvfpss rnlfltndl
lhennthnqe kkiqeeiekk etliqenvvl pqihtvtgtk nfmknlflls trqnvegsyd
gayapvlqdf rsindstnrt kkthahfskk geeenleglg nqtkgiveky acttrispnt
sqqnfvttqrs kralkqfrlp leetelekri ivddtstqws knmkhltpst ltqidyneke
kgaitqspis dcltrshsip qaarsplbia kvssfpsirp iyltrvlfqd nssh1paasy
rkkdsgvqes shflqgakkn nlslailtle mtgdqrevgs lgtusatnsvt ykkventvlp
kpdlpktsqk vellpkvhiy qkdlfptets ngspghlidlv egsslqgkteg aikwneanrp
gkvpflrvat essaktpskl idplawdnhy gtqipkeewk sqekspketa fkkkdtisl
nacesnhaia aineggqnkpe ievtwakqgr terlcqnpp vlkrhgreit rtllqsdqee
idyddtisve mkkedfdiyd edenqsprsf qkktrhyfia averlwdygm sssphvlnr
aqsgsvpqfk kvvfqeftdg sftqplyrge lnehlgllgp yiraevedni mvtfrnqasr
pysfysslis yeedqrqgaae prknfvkpne tktyfwkvqh hmaptkdefd ckawayfsdv
dlekdvhsgl igpllvchtn tlnpahgrqv tvqefalift ifdetkswyf tenmerncra
pcniqmedpt fkenyrfhai ngyimdtlpg lvmaqdqrir wylsmgsne nihsihfsgh
vftvrkkeey kmalynlypg vfetvemlps kagiwrvecl igeohlhagms tlflvysnkc
qtplgmasgh irdfqitasq qygqwapkia rlhyssina wstkepfswi kvdllapmii
hgiktqgarq kfsslyisqf iimysldgkk wqtyrgnstg tlmvffgnvd ssgikhni
ppiaryirl hptphysirst lrmelmgcdl nscsmpigme skaisdaqit assyftnmfa
twpskarlh lqgrsnawrp qvnnpkewiq vdfqktmkvt gvttqgvks1 ltsmyvkefi
issqqdghqw tlffqngkvk vfqgnqdsft pvvnsldpp1 ltrylrihpq swvhqialrm
evlgceaqdl y (SEQ ID NO:19)

Figure 22

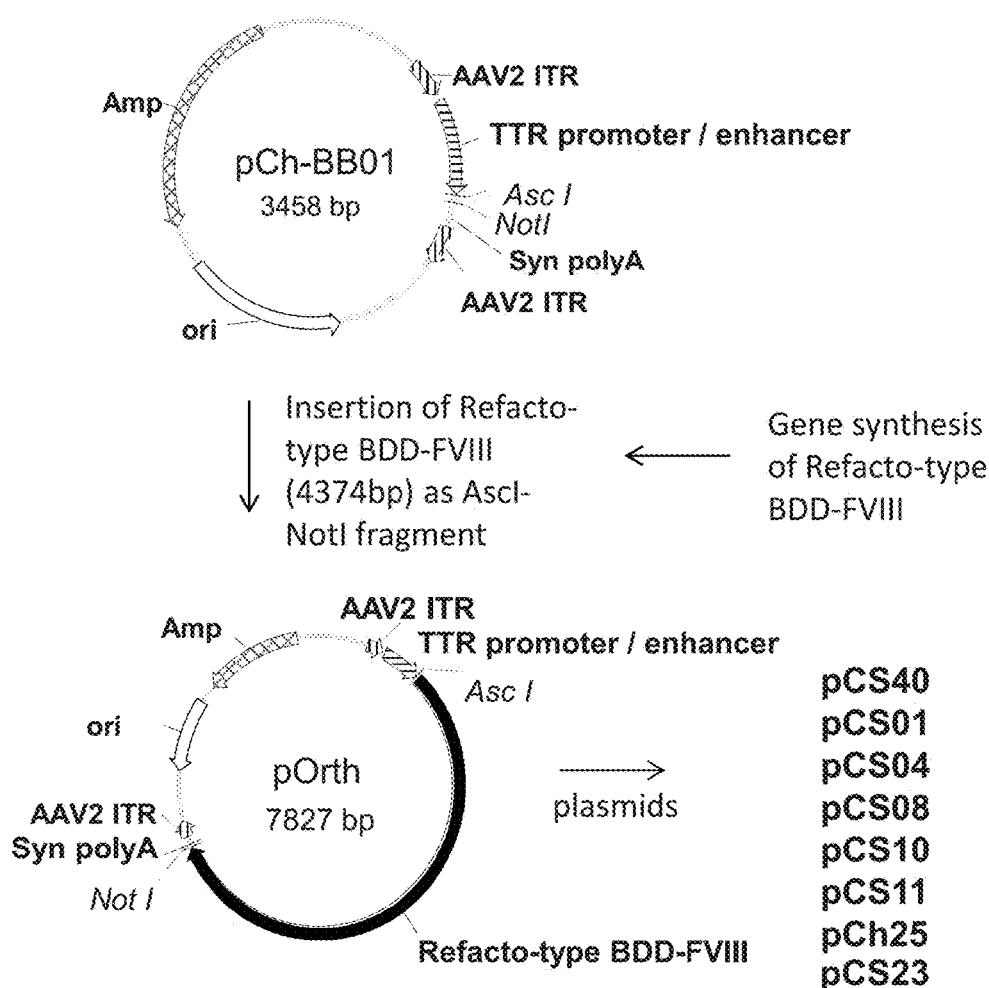


Figure 23

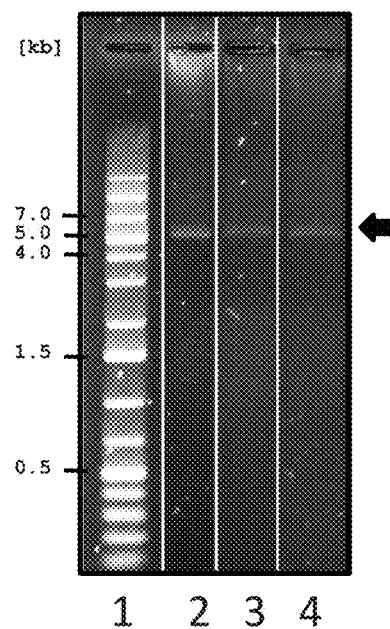


Figure 24

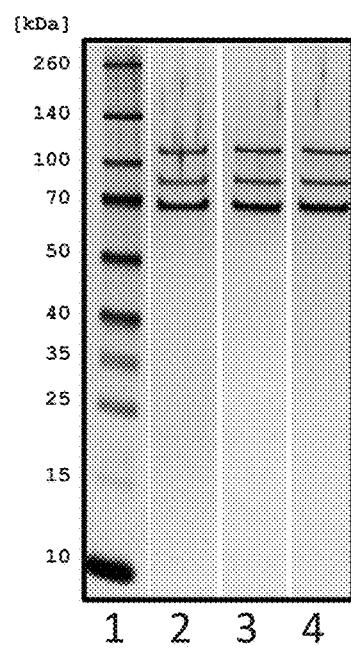


Figure 25

CS23-FL-NA

atgcagattgagctgagcacctgcttccctgtgcctgctgaggttctgcttctgccaccagg
agatactacccggcgccgtggagctgagctggactacatgcagtcgtgcacccggcggactgcct
gtggacgcagggtcccccccaagagtgcacaagagctcccccaccaacacactcaagtggtgtacaag
aagaccctgttcgtggagttccaccgcaccacccgttcaacatgcggaaaggccaggccccctggatg
ggcctgctggcccccaccatccaggccggagggtgtacgcacaccgtggatcacccctgaagaacatg
gcggccaccggcgtgagccctgcaccccggtggcgtggactactggaaaggcctctgagggcgccgag
tatgcacgcaggccaggccaggccaggccaggccaggccaggccaggccaggccaggccaggccaggcc
tacgtgtggcagggtgtgaaggagaacggccatggccagcgcaccctgtggctgatcggccctgtqgtgc
tacctgagccacgtggacccctggtaaggacactctggctgatcggccctgtqgtgc
agggaggggcagccctggccaaggagaagacccagaccctgcacaagtcatccctgtgttcgcccgtg
ttcgatgaggcaagagactggcacagcgagaccaagaacagccctgtatgcaggacaggatgcccgc
tctggccaggccctggcccaagatgcacaccgtgaacggctacgtgaacaggccctgcccggcctg
atcggtgtccacaggaagtctgtgtactggcacgtgtatcggcatggcaccaccccgagggtgcac
agcatcttctggaggccacacccttctggtaaggacccacaggcaggccaggccaggatcggaggatc
cccatcacccctgtaccgcaccgcaccctgtatggacccctggccagttccctgtgttctggccac
atcagcagccaccaggccacgcacggcatggaggctacgtgaagggtggacagactgcccccgaggagccc
cagctgaggatgaagaacaacgcaggccggaggactatgtatgatgacccgtaccgcactctgagatg
gacgtggtgagggttgcgtatgatgacccactggccatccaggatcggatgttgcggccaaagaag
caccggcaagacccctgggtgcactacatgcggccggaggactggactacgcggccctgg
ctggcccccacgcacaggagactacaagagccactacatgtatgacccctggaggatcggcagg
aagtacaagaaggctcagattcatgcctacaccgcacgcacccatccaggccatccag
caccgcgttgcgtatccctggccctgtgtacggcgagggtggccgacaccctgtatcatctt
aagaaccaggccaggccctacaacatctaccggcatcaccgtatgtgaggccctgtac
agcaggaggctggccaaaggccgtgaagcaccctgtacggacttccctgtggccggcggaggatctt
aagtacaaggccgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt
tactacaggccgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt
tgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt
ttctctgtgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt
cccgccggcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt
tacgtgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt
atcgccgcggccaggccacttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt
tacgaggacaccctgtaccctgttcccttgcgttgcgttgcgttgcgttgcgttgcgtt
ggccgtggatcctggctgcacaacagcgacttcaggaaacaggggcatgaccgcctgtgaaa
gtcaggcagctgcgacaagaacacccggcgactactacgcaggacagactacgcaggacatc
ctgtgtgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgtt
caccagaggagatcaccaggaccaccctgcagagcgcaccaggaggatcgactatgtatgacacc

(Continued)

Figure 26A

atcagcgtggagatgaagaaggaggacttcgacatctacgacgaggacgagaaccagagccccagg
agcttccagaagaagaccaggoactacttcatcgccgcgtggagaggctgtggactatggcatg
agcagcagccccacgtgctgaggaacaggcccagagcggcagcgtgcggcacttcagaaggatq
gtgttccaggagttcaccgacggcagcttacccagccccctgtacagaggcgagctgaacgagcac
ctgggcctgtggccctacatcaggcccggaggtggaggacaacatcatggtgcacccatcaggAAC
caggccagcaggcccatacgcttacagcagcgtatcagctacgaggaggaccagaggcaggGC
GCCGAGCCCCAGGAAGAACTTCGTGAAGCCAAACGAGACCAAGACCTACTTCTGGAGGTGAGCAC
CACATGGCCCCACCAAGGACGAGTTCACTGCAGGCCCTGGCCTACTTCTGTATGGACCTG
GAGAAGGACGTGACAGCGGCCTGATCGGCCCTGTGGTGTGCCACACCAACACCCCTGAACCCCC
GCCACCGCAGGGCAGGTGACCGTGCAGGAGTTGCCCTGTTCCACCATCTCGACGAGACCAAG
AGCTGGTACTTCACCGAGAACATGGAGAGGAACCTGCAGGGCCCCCTGCAACACATCCAGATGGAGGAC
CCCACCTTCAGGAGAACACTACAGGTTCCACGCCATCAACGGCTACATCATGGACACCCCTGCCCGC
CTGGTGTGGCCAGGACCAGAGGATCAGGTGGTATCTGCTGAGCATGGCAGCAACGAGAACATC
CACAGCATCCACTTCAGGCCACGTGTTCACCGTGGAGAACAGGAGGAGTACAAGATGGCCTG
TACAACCTGTACCCGGCGTGTTCAGACCGTGGAGATGCTGCCAGCAAGGCCGCATCTGGAGG
GTGGAGTGCCTGATCGCGAGCACCTGCACGCCGATGAGCACCCCTGTTCTGGTGTACAGCAAC
AAGTGCCAGACCCCCCTGGCATGGCAGCGGCCACATCAGGGACTTCAGATCACCGCCTCTGGC
CAGTACGGCAGTGGGCCCAAGCTGGCAGGCTGCACTACAGCGGAGCATCAACGCCCTGGAGC
ACCAAGGAGCCCCCTCAGTGGATCAAGGTGGACCTGCTGGCCCCATGATCATCCACGGCATCAAG
ACCCAGGGCGCCAGGAGGAGTTCACTGCAGGCTGTACATCAGGCCAGTTCACTCATGTACAGCCTG
GACGGCAAGAAGTGGCAGACCTACAGGGCAACACGACCCGGCACCCCTGATGGTGTCTCGGCAAC
GTGGACAGCGGCATCAAGCACAACATCTCAACCCCCCATCATGCCAGGTACATCAGGCTG
CACCCCACTACAGCATCAGGAGCACCTGCGGATGGAACCTGATGGCTGCGACCTGAACAGC
TGCAGCATGCCCTGGCATGGAGAGCAAGGCCATCTGACGCCAGATCACGCCAGCAGCTAC
TTCAACCATGTTGCCACCTGGAGCCCCAGCAAGGCCAGGCTGCACCTGCAGGGCAGGAGCAAC
GCCCTGGAGGCCCCAGGTGAAACAACCCAAGGAGTGGCTGCAGGTGGACTTCAGAAGACCATGAAG
GTGACCGGGGTGACCACCCAGGGCGTGAAGAGCCTGCTGACAGCATGTACGTGAAGGGAGTTCTG
ATCAGCAGCAGGCCAGGACGCCACCACTGGACCCCTGTTCCAGAACCGCAAAGTGAAGGTGTT
CAGGGCAACCAGGACAGCTCACCCCCGTGGTAACAGCCTGGACCCCCCTGCTGACCAAGGT
CTGAGGATCCACCCCCAGAGCTGGGTGCACCAAGATGCCCTGAGAACATGGAAGTGTGGATGCGAG
GCCAGGACCTGTACTGA (SEQ ID NO:20)

Figure 26B

CS23-FL-AA

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDAK' FPPRVPKSFPFNTSVVYK
KTLFVEFTDHLFNIAKPRPPWMGLLGPLTIQAEVYDTVVITLKNMASHPVSLHAVGVSYWKASEGAEY
DDQTSQREKEEDDKVFPGGSHYYWQLKENGPMASDPICLTYSYLSHVDLVKDLNSGLIGALLVCRE
GSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGLIGC
HRKSVYWHVIGHMGTTPEVHSIFLEGHTEFLVRNHRQASLEISPIFLTAQTLLMDLGQFLLFCISSH
QHDGMEAYVKVDSCPEEFQRLMKNNNEAEDYDDDLTDSEMDVVRFDDDSNSPSFIQIRSVAKKHPKTW
VHYIAAEEEDWDYAPLV LAPDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFKTREAIQHESGIL
GPLLYGEVGDTLLIIFKNQASRPYNIYPHGITEVRLPKGVKHLKDFPILPGEIFKYKWTVT
VEDGPTKSDPRCLTRYSSVNMERDLASGLIGPLLICYKESVDQRGNQIMSDKRNVILFSVFDENR
SWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSQLQLSVCLREVAYWIISIGAQTDFLS
VFFSGYTFKHKMVYEDTTLFPFSGETVEMSMENPGLWILGCHNSDFRNRCGMTALLKVSSCDKNTGD
YYEDSYEDISAYLLSKNNALEPRSFQNPPVLUKPHQREITRTTLQSDQEEIDYDDTISVEMKKEFD
IYDEDENQSPRSFQKKTRHYFIAVERLWDYGMSSSPHVLRNRAQSGSGVPQFKVVFQEFTDGSFTQ
PLYRGELNEHLLGPY IRAEVEDNIMVTFRNQASRPYSFYSSLISYEEEDQRQGAEPRKNFVKPNET
KTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLEKDVKHSGLIGPLLVCNTLNPAGRQVTVQEFAFL
FTIFDETKSUYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTPGLVMAQDQRIRWYLLS
MGSNEINIHSIHFGHVFTVRKKEEYKMALYNLYPGVFETVEMPLSKAGIWRVECLIGEHLHAGMSTL
FLVYSNKCQTPLGMASGHIRDFOITASGQYGQWAPKLARLHYSGSINAWS TKEPFWSIKV DLLAPMI
IHGIKTQGARQKFSSLYISQFYIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSGIKHNIFNPPIIAR
YIRLHPHTHYSIRSTLREMELMGCDLNCSMPLGMESKAISDAQITASSYFTNMFATWSPSKARLHLQG
RSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTSMYVKEFLISSSQDGHQWTLLFFQNGKVK
VFQGNQDSFTPVVNSLDPLLTRYLRIHFQSWVHQIALRMEVLGCEAQDLY (SEQ ID NO: 21)

Figure 27

CS23-HC-NA

gcc

accaggagat actaacctggg cgccgtggag ctgagctggg actacatgca gtctgacctg
ggcgagctgc ctgtggacgc caggttcccc cccagagtgc ccaagagctt ccccttcaac
acctcagtgg tgtacaagaa gaccctgttc gtggagttca cccgaccacatc gttcaacatc
gccaagccca ggccccctg gatgggcttg ctggggccca ccatccagge cgaggtgtac
gacaccgtgg tgatcacccct gaagaacatg gccagccacc ccgtgagocgt gcacgcccgtg
ggcgttagct actggaaagcc ctctgagggc gcccagttatg acgaccagac cagccagagg
gagaaggagg acgacaaggt gtccccggc ggcagccaca cctacgttg gcaggtgt
aaggagaacg gccccatggc cagcgaccgc ctgtgcctga cctacagota cctgagccac
gtggacctgg tgaaggacccctt gaactctggc ctgatcgccg ccctgcttgt gtgcagggag
ggcagccctgg ccaaggagaa gaccagacc ctgcacaagt tcatcctgtt gtgcggcgtg
ttcgatgagg gcaagagctg gcacagcgag accaagaaca gcoctgatgca ggacaggat
gcccgcctcg ccagggcctg gcccaagatg cacaccgtga acggctacgt gaacaggagc
ctgccccggcc tgatcggtcg ccacaggaag tctgtgtact ggcacgttat cggcatggc
accacccccc aggtgcacag catcttcctg gagggccaca ctttcttgtt gaggaaaccac
aggcaggcca gcctggagat cagccccatc accttcctga cccgcccagac cctgctgt
gacctggcc agttcctgtt gtctgcccac atcagcagcc accagcacga cggcatgg
gcoctacgtga aggtggacag ctgccccggc gagccccagc tgaggatgaa gaacaacg
gaggccgagg actatgatga tgacctgacc gactctgaga tggacgttgtt gaggtttgt
gatgacaaca gccccagctt catccagatc aggtctgtgg ccaagaagca ccccaagacc
tgggtgcact acatcgccgc cgaggaggag gactggact acggccccctt ggtgtggcc
cccgacgaca ggagctacaa gagccagtac ctgaacaacg gccccccagag gatcgccagg
aagtacaaga aggtcagatt catggccatc accgacgaga ctttcaagac cagggaggcc
atccagcacg agtctggcat cttggggccctt ctgtgttacg gcgaggtggg cgacaccctg
ctgtatcatct tcaagaacca gcccagcagg ccctacaaca ttatccccca cggcatcacc
gatgtgagggc ccctgtacag caggaggctg cccaaaggcg tgaagcacct gaaggactt
cccatccctgc cccggcggat ctcaagttac aagtggaccg tgaccgttga ggtatggcccc
accaagtctg accccaggtg cctgaccagg tactacagca gtttctgtt catggagagg
gacctggccctt ctggccctgtt cgccccctgtt ctgtatctgtt acaaggagag cgtggaccag
aggggcaacc agatcatgtt tgacaagagg aacgtgatec ttgttctctgtt gttcgatg
aacaggagat ggtatctgac cgagaacatc cagaggttcc tgccccaaccc cgcggccgtg
cagctggagg accccaggtt ccaggccagc aacatcatgc acagcatcaa cggctacgt
ttcgacagcc tgcagctgtc ttgtgttgcctg cacgagggtgg ctttctgtt catccctg
atccggccccc agaccgactt cctgtgttgc ttcttctgtt gtttctgtt caagcacaag
atgggtgtacg aggacaccctt gaccctgttc cccttccatcg gcgagaccgt gtttcatg
atggagaacc cccggccctgtt gtttctgttgc tgccacaaca gcgacttca gaaacagg
atgaccggcc ttgtgaaagt cagcagctgc gacaagaaca cccggccacta ctacgagg
agctacgagg acatcagcgc ctacgttg agcaagaaca acggccatcga gcccagg
(SEQ ID NO:22)

Figure 28

CS23-LC-NA

g agatcaccag gaccaccctg
cagagcgacc aggaggagat cgactatgtat gacaccatca gcgtggagat gaagaaggag
gacttcgaca totacgcgca ggacgagaac cagagccccca ggagcttcca gaagaagacc
aggcaactact tcatacgccgc cgtggagagg ctgtggact atggcatgag cagcagcccc
cacgtgctga ggaacagggc ccagagcgcc agcgtgcccc agttcaagaa ggttgttgc
caggagtta cccacggcag cttcacccag cccctgtaca gaggcggact gaacgagcac
ctggccctgc tggggcccta catcaggccg gaggtggagg acaacatcat ggtgacattc
aggaaccagg ccagcaggcc ctacagctt tacagcagcc tgatcagcta cgaggaggac
cagaggcagg ggcgcggagcc caggaagaac ttctgtaaagc ccaacgcgac caagacctac
ttcttggaaagg tgcagcacca catggccccca accaaggacy agtttgactg caaggccctgg
gcctacttct ctgtatgttga cctggagaag gacgtgcaca gggccctgtat cggcccccctg
ctgggtgtgcc acaccaaacac cctgaaccccc gcccacggca ggcagggtgac cgtgcaggag
ttcgcctgt tcttcaccat cttcgacgag accaagagct ggtacttcac cgagaacatg
gagaggaact gcagggcccccc ctgcaacatc cagatggagg accccacccat caaggagaac
tacaggttcc acgcacatcaa cggctacatc atggacaccc tgcccgccct ggtgatggcc
caggaccaga ggatcagggtg gtatctgtg agcatgggc gcaacgcgaaa catccacagc
atccacttca gcccacgt gttcaccgtg aggaagaagg aggagtacaa gatggccctg
tacaacctgt accccggcgt gttcgagacc gtggagatgc tgcccagcaa ggcggccatc
tggaggggtgg agtgcctgtat cggcgagcac ctgcacgccc gcatgagcac cctgttctg
gtgtacagca acaagtgcacca gaccccccotg ggcatggcca gggccacat caggacttc
cagatcaccg cctctggcca gtacggccag tggggccccca agctggccag gctgcactac
agcggcagca tcaacgcctg gagcaccaag gagcccttca gctggatcaa ggtgacactg
ctggccccca tgatcatcca cggcatcaag accccaggccg ccaggcggaaa gttcagcage
ctgtacatca gccagttcat catcatgtac agcctggacy gcaagaagtg gcagacotac
agggcaaca gacccggcac cctgatggtg ttcttcggca acgtggacag cagccgcate
aagcacaaca tcttcacccccc ccccatcatc gccaggtaca tcaggctgca ccccaacccac
tacagcatca ggagcaccct gggatggaa ctgatggct ggcacactgaa cagctgoagc
atgccttgg gcatggagag caaggccatc tctgacgccc agatcaccgc cagcagctac
ttcaccaaca tggcgccac ctggagcccc agcaaggcca ggctgcaccc gcaaggccagg
agcaacgcct ggaggccccca ggtgaacaac cccaaaggagt ggctgcaggt ggacttccag
aagaccatga aggtgaccgg cgtgaccacc cagggcgtga agagcctgt gaccacatg
tacgtgaagg agttcctgtat cagcagcagc caggacggcc accagtggac cctgttcttc
cagaacggca aagtgaaggt gttccaggcc aaccaggaca gcttcacccccc cgtqgtgaac
agcctggacc ccccccctgtat gaccaggtat ctgaggatcc accccccagag ctgggtgcac
cagatcgccc tgagaatgga agtgcgtggaa tgcgaggccc aggacactgta c
(SEQ ID NO:23)

Figure 29

CS01m13-FL-NA

ATGCAGATTGAGCTGTCCACCTGCTTCTTCTGTGCCCTGCTGAGATTCTGCTTCTGCCACCAGG
AGATACTACCTGGGGCTGTGGAACCTTCTTGGACTACATGCAGTCTGACCTGGAGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGAGTGCCCAAGTCCCTCCATTCAACACCTCTGGTCTACAAG
AAGACACTCTTGTGAAATTCACTGACCACCTGTTCAACATTGAAAACCCAGACCACCCCTGGATG
GGACTCCTGGGACCCACCATTCAAGGCTGAGGTGATGACACTGTGGTCATCACCTCAAGAACATG
GCATCCCACCCCTGTGTCTGATGCTGTGGAGTCTCATACTGGAAAGCCTCTGAAGGGCTGAG
TATGATGACCAGACATCCCAGAGAGAGAAAGAGGATGACAAGGTGTCCTGGGGATCTCACACC
TATGTGTGGCAAGTCCTCAAGGAGAATGGACCCATGGCATCTGACCCACTCTGCCCTGACATACTCC
TACCTTCTCATGTGGACCTGGTCAAGGACCTCAACTTGGACTGATTGGGCACTGCTGGTGTGC
AGGGAAAGGATCCTGGCCAAGGAGAAAACCCAGACACTGCACAAGTCATTCTCTGTTGCTGTC
TTTGATGAGGGCAAGTCTGGCACTCTGAAACAAAGAACCTCGATGCAAGACAGGATGCTGCC
TCTGCCAGGGCATGGCCAAGATGCACACTGTGAATGGCTATGTGAACAGATCACTGCCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCACTGGCATGTGATTGGCATGGGACAACCCCTGAAGTGCAC
TCCATTTCCTGGAGGGACACACCTCCCTGGTCAGGAACACAGACAAGCCTCTGGAGATCTCT
CCCATCACCTCCCTCACTGCACAGACACTGCTGATGGACCTGGACAGTCCCTGCTGTCCTGCCAC
ATCTCTCCCACAGCATGTCAGGACCTGCAAGGACTATGTCAGGACTCATGCCCTGAGGAACCA
CAGCTCAGGATGAAGAACATGAGGAGGCTGAGGACTATGATGATGACCTGACTCTGAGATG
GATGTGGTCAGATTGATGATGACAACCTCTCCATTCATTCAAGATCAGGTCATGGCAAAGAAA
CACCCCAAGACATGGGTGCACTACATTCAGGACTGAGGAAGAGGACTGGGACTATGCACCACTGGTC
CTGGCCCTGATGACAGGAGCTACAAGTCTCAGTACCTCAACAATGCCCAACAAAGAATTGGAGA
AAAGTACAAGAAAGTCAGATTGATGGCCTACACTGATGAAACCTTCAAGACAAGAGAAGCCATTCA
CATGAGTCTGGCATTCTGGGACCACCTCTGTATGGGAGACACCCCTGTCATCATCTTC
AAGAACCCAGGCTCCAGGCCCTACAAACATCTACCCACATGGCATCACTGATGTCAGGCCCTGTAC
AGCAGGAGACTGCCAAAGGGTGAACACACCTCAAGGACTTCCCCATTCTGCCCTGGAGAGATCTC
AAAGTACAAGTGGACTGCACTGTGGAGGATGGACCAACAAAGTCTGACCCCAAGGTGCCCTACCAGA
TACTACTCCTCTTGTGAAACATGGAGAGAGACCTGGCATCTGGACTGATTGGACCACTGTCATC
TGCTACAAGGAGCTGTGGACCAAGAGGCAACCAGATCATGTCATGACAAGAGAAATGTGATTCTG
TTCTCTGTCTTGATGAGAACAGATCATGGTACCTGACTGAGAACATTCAAGAGATTCTGCCAAC
CCTGCTGGGTGCAACTGGAAGACCCCTGAGTCCAGGCAAGCAACATCATGCACTCCATCAATGGC
TATGTCTTGACTCTCTCAGCTTCTGTCTGCCATGAGGTGCCCTACTGGTACATTCTTCT
ATTGGGGCACAAACTGACTTCTCTGTCTTCTCTGGATACACCTCAAGCACAAGATGGTG
TATGAGGACACCTGACACTCTTCCCATTCTGGGAAACTGTGTTCATGAGCATGGAGAACCCCT
GGACTGTGGATTCTGGGATGCCACAACACTCTGACTTCAGAAACAGGGAAATGACTGCACTGTC
GTCTCCTCTGTGACAAGAACACTGGGACTACTATGAGGACTCTTATGAGGACATCTCTGCC
CTGCTCAGCAAGAACAAATACCACCTACGTGAACCGCTCCGTCTCAGAAATCCACCTGTCTGAAG
AGACACCAGAGAGAGATCACCAGGACAACCCCTCCAGTCTGACCAGGAAGAGATGACTATGAC
ACCATTCTGTGGAGATGAGAACAGGAGGACTTGACATCTATGATGAGGACGAGAACCAAGTCT
AGATCATTCCAGAAGAACAGACACTACTTCATTGCTGCTGGAAAGAGACTGTGGACTATGGC
ATGTCTCCTCTCCCCATGTCCAGGAACAGGGCACAGTCTGGCTCTGTGCCACAGTTCAAGAAA

(Continued)

Figure 30A

GTGGTCTTCCAGGAGTTCACTGATGGCTCATTACCCAGCCCCGTACAGAGGGAACTGAATGAG
CACCTGGGACTCCTGGGACCATAACATCAGGGCTGAGGTGGAAGACAACATCATGGTGACATTCAA
AACCAGGCCCTCAGGCCCTACAGCTTCACTCTTCCTCATCAGCTATGAGGAAGACCAGAGACAA
GGGGCTGAGCCAAGAAAAGAACCTTGTGAAACCCAATGAAACCAAGACCTACTCTGGAAAGTCCAG
CACACATGGCACCCACCAAGGGATGAGTTGACTGCAAGGCCTGGGCATACTCTCTGATGTGGAC
CTGGAGAAAGATGTGCACTCTGGCCTGATTGGCCACTCTGGTCTGCCACACCAACACCCCTGAAC
CCTGCACATGGAAGGCAAGTGAATGTGCAGGGAGTTGCCCTTCTTCAACATCTTGATGAAACC
AAGTCATGGTACTCACTGAGAACATGGAGAGAAACTGCAGAGCACCATGCAACATTCAAGATGGAA
GACCCCACCTTCAAGGAGAACTACAGGTTCCATGCCATCAATGGCTACATCATGGACACCCCTGCCT
GGGCTTGTATGGCACAGGGACAGAGAATCAGATGGTACCTGCTTCTATGGATCCAATGAGAAC
ATTCACTCCATCCACTTCTCTGGCATGTCTTCACTGTGAGAAAGAAGGAGGAATACAAGATGGCC
CTGTACAACCTCTACCCCTGGGCTTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGG
AGGGTGGAATGCCCTATTGGGAGCACCTGCATGCTGGCATGTCAACCCCTGTCCTGGTCTACAGC
AACAAAGGCCAGACACCCCTGGGAATGCCCTCTGCCACATCAGGGACTTCCAGATCACTGCCTCT
GGCAGTATGCCAGTGGCACCCAAACTGCCAGGCTCCACTACTCTGGCTCCATCAATGCATGG
TCAACCAAGGAGCCATTCTTGGATCAAGGTGGACCTGCTGGCACCCATGATCATTCAATGGCATE
AAGACACAGGGGGCAAGACAGAAAATTCTCCTCTGTACATCTCACAGTTCATCATCTACTCT
CTGGATGGCAAGAAGTGGCAGACATAACAGAGGCAACTCCACTGGCACCCATGGTCTTGGC
AATGTGGACAGCTCTGGCATCAAGCACACATCTCAACCCCTCCATCTGGCAGATACTCAGG
CTGCACCCACCCACTACTCAATCAGATCAACCCCTCAGGATGGAACGTGGGATGTGACCTGAAC
TCCTGCTCAATGCCCTGGGAATGGAGAGCAAGGCCATTCTGATGCCAGATCACTGCATCCTCT
TACTTCACCAACATGTTGCCACCTGGTCAACATGCCAGGCTGCACCTCCAGGGAAAGAAGC
AATGCCCTGGAGACCCAGGTCAACAAACCCAAAGGAATGGCTGCAAGTGGACTTCCAGAAAGACAATG
AAAGTCATGGGTGACAACCCAGGGGTCAAGTCTCTGCTCACCTCAATGTATGTGAAGGAGTTC
CTGATCTCTCCACAGGATGGCACCAGTGGACACTCTTCTCCAGAAATGGCAAAGTCAAGGTG
TTCCAGGGCAACCAGGACTCTTCACACCTGTGGTAACACTGGACCCCCCCCCCTGACAAGA
TACCTGAGAATTCAACCCCCAGTCTGGGTCCACCAGATTGCCCTGAGAATGGAAGTCCTGGGATGT
GAGGCACAAGACCTGTACTGA (SEQ ID NO:90)

Figure 30B

CS01m23-FL-NA

ATGCAGATTGAGCTGCCACCTGCTTCTGTGCCCTGAGATTCTGCTCTGCCACCAGG
AGATACTACCTGGGGCTGTGGAACCTTCTTGGACTACATGCAGTCTGACCTGGAGAGCTGCCT
GTGGATGCCAGGTTCCACCCAGAGTGCCCAAGTCCTCCCATTCAACACCTCTGGCTACAAG
AAGACACTCTTGTGAAATTCACTGACCACTGTTCAAACATTGCAAAAACCCAGACCCACCCCTGGATG
GGACTCCTGGGACCCACCATTCAAGGCTGAGGTGATGACACTGTGGCTCACCTCAAGAACATG
GCATCCCACCCCTGTGTCCTGCACTGCTGAGGTCTCATACTGGAAATCCTCTGAAGGGCTGAG
TATGATGACCAAGACATCCCAGAGAGAGAAAGAGGATGACAAGGTGTCCTGGAAAGTCTCACACC
TATGTGCGCAAGTCTCAAGGAGAATGGACCACTGCATCTGACCCACCCCTGCCTGACATACTCC
TACCTTCTCATGTGGACCTGGTCAGGACCTCAACTCTGGACTGATTGGGCACTGCTGGTGTGC
AGGGAAAGGATCCCTGGCAAGGAGAAAACCCAGACACTGCACAAGTCAATTCTCTGTTGTC
TTTGATGAGGGCAAGTCTGGCACTCTGAAACAAAGAACTCCCTGATGCAAGACAGGGATGTC
TCTGCCAGGGCATGGCCAAGATGCACACTGTGAATGGCTATGTGAAACAGATCAGTGCCTGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGACAACCCCTGAAGTGCAC
TCCATTTCCTGGAGGGACACACCTCCCTGGCAGGAACCACAGACAAGCCTCTGGAGATCTCT
CCCACATCACUTTCCCTACTGCACAGACACTGCTGATGGACCTGGACAGTTCTGCTGTTGCCAC
ATCTCTTCCCACCAAGCATGGCATGGAACGCTATGTCAGGGACTCATGCCCTGAGGAACCA
CAGCTCAGGATGAAGAACAAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTGAGATG
GATGTGGTCAGATTGATGATGACAACACTCTCCATCCTCATTCAGATCAGGCTGTGGCAAAGAAA
CACCCCAAGACATGGGCACTACATTGCTGCTGAGGAAGAGGACTGGGACTATGCCACACTGGTC
CTGGCCCCCTGATGACAGGGACTACAAGTCTCAGTACCTCAACAATGGCCCACAAAGAATTGAAAGA
AACTACAAGAAAGTCAGATTCATGGCCTACACTGATGAAACCTTCAGACAAGAGAACCCATTCA
CATGAGTCTGGCATTCTGGGACCACTCCTGATGGGAAAGTGGGAGACACCCCTGCTCATCATTTC
AGAACACCAGGCTCCAGGCCCTACAAACATCTACCCACATGGCATCACTGATGTCAGGCCCCCTGTAC
AGCAGGAGACTGCCAAAGGGTGAACACCTCAAGGACTTCCCCATTCTGCCCTGGAGAGATCTTC
AACTACAAGTGGACTGCACTGTGGAGGATGGACCAACAAAGTCTGACCCAGGTGCCCTACCAGA
TAATCTCTCTTGTGAACATGGAGAGACCTGGCATCTGGACTGATTGGACCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGAGGCAACCAGATCATGTCATGACAAGAGAAATGTGATTCTG
TTCTCTGCTTTGATGAGAACAGATCATGGTACCTGACTGAGAACATTCAAGAGATTCTGCCAAC
CTGCTGGGGTGCAACTGGAAGACCCCTGAGTCCAGGCAAGCAACATCATGCACTCCATCAATGGC
TATGTGTTGACTCTCTCCAGCTTCTGCTGCTGCACTGAGGTGGCTACTGGTACATTCTTCT
ATTGGGGCACAAACTGACTTCTTCTGTCTCTGGATACACCTTCAGGACAAGATGGTG
TATGAGGACACCCCTGACACTCTTCCCATTCTCTGGGAAACTGTGTCATGAGCATGGAGAACCC
GGACTGTGGATTCTGGATGCCACAACTCTGACTTCAGAACACAGGGAAATGACTGCACTGCTCAA
GTCTCTCTGTGACAAGAACACTGGGACTACTATGAGGACTCTTATGAGGACATCTCTGCC
CTGCTCAGCAAGAACAAATACCACTACGTGAAACCGCTCCCTGTCAGAACATCCACCTGCTGAAAG
AGACACCCAGAGAGACATCACCAGGACAACCCCTCCAGTCTGACCAAGGAAGAGATTGACTATGATGAC
ACCATTCTGTGGAGATGAAGAAGGAGGACTTGCACATCTATGATGAGGACGAGAACCCAGTCTCCA
AGATCAATTCAAGAAGAACAGACACTACTTCATTGCTGCTGAAAGACTGTGGGACTATGGC
ATGTCTTCCCTCCCCATGTCCCTCAGGAACAGGGCACAGTCTGGCTGTGCCACAGTTCAAGAAA

(Continued)

Figure 31A

GTGGTCTTCCAGGAGTTCACTGATGGCTCATTCACCCAGCCCCGTACAGAGGGAACTGAATGAG
CACCTGGACTCCTGGGACCATAACATCAGGGCTGAGGTGGAAGACAACATCATGGTACATTCAA
AACCAAGGCCTCAGGCCCTACAGCTCTACTCTTCCCTCATCAGCTATGAGGAAGACCAGAGACAA
GGGGCTGAGCCAAGAAAGAACTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAG
CACACATGGCACCCACCAAGGATGAGTTGACTGCAAGGCCTGGGCATACTTCTCTGATGTGGAC
CTGGAGAAAGATGTGACTCTGGCCTGATTGGCCACTCTGGTCTGCCACACCAACACCCTGAAC
CCTGCACATGGAAGGCAAGTGAATGTGCAAGGAGTTGCCCTCTTCACCATCTTGATGAAACC
AAGTCATGGTACTTCACTGAGAACATGGAGAGAAACTGCAGAGCACCAGTCAACACATTGAGATGGAA
GACCCCACCTCAAGGAGAACTACAGGTTCCATGCCATCAATGGTACATCATGGACACCCTGCCT
GGGCTTGTATGGCACAGGACCAGAGAATCAGATGGTACCTGTTCTATGGGATCCAATGAGAAC
ATTCACTCCATCCACTTCTCTGGCCTGTCTTCACTGTGAGAAAGAAGGAGGAATACAAGATGGCC
CTGTACAACCTCTACCCCTGGGTCTTGAGACTGTGGAGATGCTGCCCTCAAAGCTGGCATCTGG
AGGGTGGAAATGCCCTATTGGGAGCACCTGCATGCTGGCATTGTCAACCCCTGTTCTGGTCTACAGC
AACAGTGCCAGACACCCCTGGGAATGGCCTCTGCCACATCAGGACTTCCAGATCACTGCCCT
GCCAGTATGCCAGTGGCACCCAACTGCCAGGCTCCACTACTCTGGCTCCATCAATGCATGG
TCAACCAAGGAGCCATTCTCTGGATCAAGGTTGGACCTGCTGGCACCCATGATCATGGCATT
AAGACACAGGGGCAAGACAGAAATTCTCTCTGTACATCTCACAGTTCATCATGTACTCT
CTGGATGGCAAGAAGTGGCAGACATACAGAGGAACTCCACTGGCACCCCTCATGGTCTCTTGGC
AATGTGGACAGCTCTGGCATCAAGCACAACATCTCAACCCCTCCATATTGCCAGATACTCAGG
CTGCACCCCACTACTCAATCAGATCAACCCCTCAGGATGGAACATGATGGATGTGACCTGAAC
TCCCTGCTCAATGCCCTGGGAATGGAGAGCAAGGCCATTCTGATGCCAGATCACTGCATCCCT
TACTCACCAACATGTTGCCACCTGGTACCATCAAAAGCCAGGCTGCACCTCCAGGGAAAGAAGC
AATGCCCTGGAGACCCCAAGGTCAACAAACCCAAAGGAATGGCTGCAAGTGGACTTCCAGAAGACAATG
AAAGTCACTGGGGTGACAACCCAGGGGTCAAGTCTCTGCTCACCTCAATGTATGTGAAGGGAGTTC
CTGATCTCTTCTCACAGGATGCCACCAAGTGGACACTCTTCTTCCAGAATGCCAAAGTCAAGGTG
TTCCAGGGCAACCAGGACTCTTCACACCTGTTGAACCTCACTGGACCCCCCTCCTGACAAGA
TACCTGAGAATTCAACCCCAAGTCTGGGTCCACCAAGATTGCCCTGAGAATGGAAGTCTGGATGT
GAGGCACAAGACCTGTACTGA (SEQ ID NO: 91)

Figure 31B

CS01m3-FL-NA

ATGCAGATTGAGCTGTCCACCTGCTCTTCTGTGCCTGCTGAGATTCTGCTTCTGCCACCAGG
AGATACTACCTGGGGCTGTGAACCTTCTTGGACTACATGCAGTGTACCTGGAGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGAGTGCCAAGTCCTCCATTCAACACCTCTGTGGTCTACAAG
AAGACACTCTTGTGGAATTCACTGACCACCTGTTCAACATTGCAAAACCCAGACCACCCCTGGATG
GGACTCCTGGGACCCACCATTCAAGGCTGAGGTGTATGACACTGTGGTCACTCACCTCAAGAACATG
GCATCCCACCCCTGTGTCTGCATGCTGTGGAGTCTCATACTGGAAAGCCTCTGAAGGGGCTGAG
TATGATGACCAGACATCCCAGAGAGAGAAAGAGGATGACAAGGTGTTCCCTGGGGATCTCACACC
TATGTGTGGCAAGTCCTCAAGGAGAATGGACCCATGGCATCTGACCCACTCTGCCCTGACATACTCC
TACCTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGACTGATTGGGCACTGCTGGTGTGC
AGGGAAAGGATCCCTGGCCAAGGAGAAAACCCAGACACTGCACAAGTCATTCTCCTGTTGCTGTC
TTTGATGAGGGCAAGTCTTGGCACTCTGAAACAAAGAACACTCCCTGATGCAAGACAGGGATGCTGCC
TCTGCCAGGGCATGGCCAAGATGCACACTGTGAATGGCTATGTGAACAGATCACTGCCCTGGACTC
ATTGGCTGCCACAGGAAATCTGCTACTGGCATGTGATTGGCATGGGACAACCCCTGAAGTGCAC
TCCATTTCCTGGAGGGACACACCTTCTGGTCAGGAACCACAGACAAAGCCTCTGGAGATCTCT
CCCACATCACCTTCTCACTGCACAGACACTGCTGATGGACCTTGGACAGTTCCTGCTGTTCTGCCAC
ATCTCTTCCCACCGATGGCATGGAAGCTATGTCAAGGTGGACTCATGCCCTGAGGAACCA
CAGCTCAGGATGAAGAACAAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCAGATTGATGATGACAACACTCTCCATCCTTCATTCAAGATCAGGTCTGTGGCAAAGAAA
CACCCCAAGACATGGGTGCACTACATGCTGCTGAGGAAGAGGACTGGGACTATGCCACACTGGTC
CTGGCCCTGATGACAGGGACTACAAGTCTCAGTACCTCAACAATGCCAACAAAGAATTGGAAGA
AACTACAAGAAAGTCAGATTGATGACAACACTGATGAAACCTTCAAGACAAGAGAACCCATTCA
CATGAGTCTGGATTCTGGGACCACTCTGTATGGGAAGTGGGAGACACCCGCTCATCATCTTC
AAGAACCAAGGCCCTCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCCTGTAC
AGCAGGAGACTGCCAAAAGGGGTGAAACACCTCAAGGACTTCCCCATTCTGCCCTGGAGAGATCTTC
AACTACAAGTGGACTGTCACTGTGGAGGATGGACCAACAAAGTCTGACCCAGGTGCTCACCA
TACTACTCCTCTTGTGACATGGAGAGAGACCTGGCATCTGGACTGATTGGACCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAAGAGAGGCAACCAGATCATGTCATGACAAGAGAAATGTGATTCTG
TTCTCTGTCTTGATGAGAACAGATCATGGTACCTGACTGAGAACATTCAAGAGATTCCCTGCCAAC
CCTGCTGGGGTGCACGGAAAGACCCCTGAGTTCCAGGCAAGCAACATCATGCACTCCATCAATGGC
TATGTGTTGACTCTCTCAGCTTCTGTCTGCCCTGCATGAGGTGGCTACTGGTACATTCTTCT
ATTGGGGCACAAACTGACTTCTCTGTCTTCTCTGGATACACCTTCAAGCACAAGATGGTG
TATGAGGACACCCCTGACACTCTTCCCATTCTCTGGGAAACTGTGTTCATGAGCATGGAGAACCT
GGACTGTGGATTCTGGGATGCCACAACCTCTGACCTCAGAAACAGGGGAATGACTGCACTGCTCAA
GTCTCCTCTGTGACAAGAACACTGGGACTACTATGAGGACTCTTATGAGGACATCTCTGCCCTAC
CTGCTCAGCAAGAACAAATACCACCTACGTGAACCGCTCCCTGTCTCAGAAATCCACCTGTCTGAAG
AGACACCAAGAGAGAGATCACCAGGACAACCCCTCCAGTCTGACCAAGGAGAGATGACTATGATGAC
ACCATTCTGTGAGATGAGAACAGGAGGACTTGTACATCTATGATGAGGACGAGAACCAAGTCTCCA
AGATCATTCCAGAAGAACAGACACTACTTCATTGCTGCTGGAAAGACTGTGGGACTATGGC
ATGTCCTCTCTCCCCATGCTCTAGGAACAGGGCACAGTCTGGCTCTGTGCCACAGTTCAAGAAA

(Continued)

Figure 32A

GTGGTCTTCCAGGAGTTCACTGATGGCTCATTACCCAGCCCCGTACAGAGGGAACTGAATGAG
CACCTGGACTCCTGGGACCATACATCAGGGCTGAGGTGGAAGACAACATCATGGTGACATTGAGA
AACCAAGGCCTCCAGGCCCTACAGCTTCACTCTTCCCTCATCAGCTATGAGGAAGACCAGAGACAA
GGGGCTGAGCCAAGAAAGAACCTTGTGAAACCCAAATGAAACCAAGACCTACTTCTGGAAAGTCCAG
CACCACATGGCACCCACCAAGGATGAGTTGACTGCAAGGCCTGGGCAACTTCTCTGATGTGGAC
CTGGAGAAAGATGTGCACTCTGGCCTGATTGGCCCACCTCTGGTCTGCCACACCAACACCTGAAC
CCTGCACATGGAAGGCAAGTGAUTGTGCAGGAGTTGCCCTTCTTCAACCATTTGATGAAAC
AAGTCATGGTACTTCACTGAGAACATGGAGAGAAACTGCGAGACCCATGCAACATTGAGATGGAA
GACCCCACCTCAAGGAGAACTACAGGTTCCATGCCATCAATGGCTACATCATGGACACCCCTGCCT
GGGCTTGTCAATGGCACAGGACCAGAGAACATGAGATGGTACCTGCTTTCTATGGGATCCAATGAGAAC
ATTCACCTCCATCCACTTCTCTGGGATGTCCTCACTGTGAGAAAGAAGGAGGAATACAAGATGGCC
CTGTACAACCTCTACCCCTGGGTCTTGTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGG
AGGGTGGAAATGCCCTATTGGGAGCACCTGCATGGCATGTCACCCCTGTTCTGGTCTACAGC
AACAACTGCCAGACACCCCTGGGAATGCCCTGGCCACATCAGGGACTTCCAGATCACTGCCCT
GGCCAGTATGCCAGTGGCACCCTAAACTGCCAGGCTCCACTACTCTGGCTCCATCAATGCATGG
TCAACCAAGGAGGCCATTCTCTGGATCAAGGTGGACCTGCTGGCACCCATGATCATGGCATC
AAGACACAGGGGCAAGACAGAAATTCTCTCTGTACATCTCACAGTTCATCATGTACTCT
CTGGATGGCAAGAAGTGGCAGACATAAGAGGAACCTCCACTGGCACCCCTCATGGTCTTCTGGC
AATGTGGACAGCTCTGGCATCAAGCACACATCTCAACCCCTCCATGGCAGATACTCAGG
CTGCACCCACCCACTACTCAATCAGATCAACCCCTCAGGATGGAACATGATGGGATGTGACCTGAAC
TCCTGCTCAATGCCCTGGGAATGGAGAGCAAGGCCATTCTGATGCCAGATCACTGCATCCCT
TACTTCACCAACATGTTGCCACCTGGTACCATCAAAAGCCAGGCTGCACCTCCAGGGAAAGAAC
AATGCCCTGGAGACCCCAAGGTCAACACCCAAAGGAATGGCTGCAAGTGGACTTCCAGAAGACAATG
AAAGTCACTGGGTGACAACCCAGGGGTCAAGTCTCTGCTCACCTCAATGTATGTGAAGGAGTTC
CTGATCTCTCCACAGGATGCCACCAAGTGGACACTCTTCTTCCAGAATGGCAAAGTCAGGTG
TTCCAGGGCAACCAGGACTCTTCACACCTGTGGTGAACACTGGAACCCCCCCTGACAAGA
TACCTGAGAATTCAACCCCAAGTCTGGTCCACCAAGATTGCCCTGAGAATGGAAGTCCTGGGATGT
GAGGCACAAGACCTGTACTGA (SEQ ID NO: 92)

Figure 32B

CS01m2-FL-NA

ATGCAGATTGAGCTGTCCACCTGCTCTTCTGTGCCCTGCTGAGATTCTGCTTCTGCCACCAGG
AGATACTACCTGGGGCTGTGGAACCTTCTTGGACTACATGCAGTCGACCTGGGAGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGAGTGCCCAAGTCCCTCCCATTCACACACCTCTGCGGTCTACAAG
AAGACACTCTTGTGCAATTCACTGACCACCTGTTCAACATTGAAAACCCAGACCAACCCAGGATG
GGACTCCTGGGACCCACCAATTCAAGGCTGAGGTGATGACACTGTGGTCGTCACCCCAAGAACATG
GCATCCCACCCGTGTCATGCTGTGGGAGTCTCATACTGGAAATCCTCTGAAGGGCTGAG
TATGATGACCAGACATCCCAGAGAGAGAAAGAGGATGACAAGGTGTTCCCTGGGAAGTCTCACACC
TATGTGTTGGCAAGTCCTCAAGGAGAATGGACCCACTGCATCTGACCCACCCCTGCCGTGACATACTCC
TACCTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGGACTGCTGGTGTG
AGGGAAGGATCCCTGGCAAGGAGAAAACCCAGACACTGCACAAGTTCATTCTCTGTTGCTGTC
TTTGTGAGGGCAAGTCTTGCACTCTGAAAACAAAGAAACTCCCTGATGCAAGACAGGGATGCTGCC
TCTGCCAGGGCATGGCCAAGATGCACACTGTGAATGGCTATGTGAACAGATCACTGCCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGACAACCCCTGAAGTGCAC
TCCATTTCCTGGAGGGACACACCTCCCTGGTCAGGAACCACAGACAAGCCTCTGGAGATCTCT
CCCATCACCTCCCTCACTGCACAGACACTGCTGATGGACCTTGGACAGTCTCTGCTGTTCTGCCAC
ATCTCTCCCACCAAGCATGATGGCATGGAAGCCTATGTCAGGCTGGACTCATGCCCTGAGGAACCA
CAGCTCAGGATGAAGAACAAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCAGATTGATGACAACTCTCCATCCTTCATTCAAGATCAGGTCTGGCAAAGAAA
CACCCCAAGACATGGGTGCACTACATTGCTGAGGAAGAGGACTGGGACTATGCCACCACTGGTC
CTGGCCCTGATGACAGGAGCTACAAGTCTCAGTACCTCAACAAATGGCCCAAAAGAATTGGAAGA
AACTACAAGAAAGTCAGATTGATGGCTACACTGATGAAACCTTCAGACAAGAGAACCCATTCAG
CATGAGTCTGGCATTCTGGGACCACCTCTGATGGGAAGTGGGAGACACCCCTGCTCATCTTC
AGAACCCAGGCCCTCACACATCTACCCACATGCCATCACTGATGTCAGGCCCTGTAC
AGCAGGAGACTGCCAAAAGGGGTGAAACACCTCAAGGACTTCCATTCTGCCTGGAGAGATCTTC
AAAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACAAAGTCTGACCCCAAGTGCCTCACAGA
TACTACTCCTTTTGTAACATGGAGAGAGACCTGGCATCTGGACTGATTGGACCACTGCTCATC
TGCTACAAGGAGTCTGGGACCAGAGAGGCAACCAAGATCATGTCGACAAGAGAAATGTGATCTG
TTCTCTGTCTTGATGAGAACAGATCATGGTACCTGACTGAGAACATTCAGAGATTCCCTGCCAAC
CCTGCTGGGTGCAACTGGAAGACCCCTGAGTTCCAGGCAAGCAACATCATGCACTCCATCAATGGC
TATGTGTTGACTCTCTCCAGTTCTGTCGCTGCCATGAGGTGGCTACTGGTACATTCTTCT
ATTGGGGCACAAACTGACTTCCTTCTGTCCTCTCTGGATACACCTTCAGACAAGATGGT
TATGAGGACACCCCTGACACTCTCCATTCTCTGGGAAACTGTGTTCATGAGCATGGAGAACCC
GGACTGTGGATTCTGGGATGCCACAACACTGACTTCAGAAACAGGGGAATGACTGCACTGCTCAA
GTCCTCTGTGACAAGAACACTGGGACTACTATGAGGACTCTTATGAGGACATCTCTGCTAC
CTGCTCAGCAAGAACATGCCATTGAGGCCAGAAGCTTCTCAGAACCTGCTGCCATGAGAGA
CACCAGAGAGAGATCACCAGGACAACCCCTCCAGTCTGACCAAGGAAGAGATTGACTATGATGACACC
ATTCTGTGGAGATGAAGAAGGAGACTTTGACATCTATGATGAGGACGAGAACCAAGTCTCCAAGA
TCATTCCAGAAGAACAGACACTACTTCATTGCTGCTGGAAAGACTGTGGACTATGGCATG
TCTTCCTCTCCCCATGTCCTCAGGAACAGGGCACAGTCTGGCTCTGTCACAGTCAAGAACAGT

(Continued)

Figure 33A

GTCTTCCAGGAGTTCACTGATGGCTATTACCCCAGCCCCGTACAGAGGGAACTGAATGAGCAC
CTGGGACTCCTGGGACCATACATCAGGGCTGAGGTGGAAGACAACATCATGGTACATTAGAAAC
CAGGCCTCCAGGCCCTACAGCTTCACTCTCCCTCATCAGCTATGAGGAAGACAGAGACAAGGG
GCTGAGCCAAGAAAGAACTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAGCAC
CACATGGCACCCACCAAGGATGAGTTGACTGCAAGGCCCTGGGCACTACTTCTGTATGTGGACCTG
GAGAAAGATGTGCACTCTGGCCTGATTGGCCACTCCTGGTCTGCCACACCAACACCCCTGAACCC
GCACATGGAAGGCAAGTGACTGTGCAGGGTTGCCCTCTTCACCATCTTGATGAAACCAAG
TCATGGTACTTCACTGAGAACATGGAGAGAAACTGCAGAGCACCAGCAACATTGAGATGAAAGAC
CCCACCTCAAGGAGAACTACAGGTTCCATGCCATCAATGGCTACATCATGGACACCCCTGCCCTGG
CTTGTCTGGCACAGGACCAGAGAACATCAGATGGTACCTGCTTCTATGGGATCCAATGAGAACATT
CACTCCATCCACCTCTGGCATGTCTTCACTGTGAGAAAGAAGGAGGAATAAACAGATGCCCTG
TACAACCTCTACCCCTGGGTCTTGAGACTGTGGAGATGCTGCCCTCAAAGCTGGCATCTGGAGG
GTGGAATGCCCTATTGGGAGCACCTGCATGCTGGCATGTCAACCCCTGTTCTGGTCTACAGCAAC
AAAGTGCAGACACCCCTGGGAATGCCCTCTGGCACATCAGGACTTCACTGCCCTGG
CAGTATGCCAGTGGCACCCAAACTGCCAGGCTCAACTCTGGCTCCATCAATGCAATGGTCA
ACCAAGGAGCCATTCTCTGGATCAAGGTGGACCTGCTGGCACCCATGATCATTGATGGCATCAAG
ACACAGGGGCAAGACAGAAATTCTCCTCTGTACATCTCACAGTCACTCATCATGTACTCTG
GATGGCAAGAAGTGGCAGACATACAGAGGCAACTCCACTGGCACCCATGGTCTTGGCAAT
GTGGACAGCTGGCATCAAGCACAACATCTCAACCCCTCCCATATTGCCAGATACTCAGGCTG
CACCCACCCACTACTCAATCAGATCAACCCCTCAGGATGGAACGTGATGGGATGTGACCTGAAC
TGCTCAATGCCCTGGGAATGGAGAGCAAGGCCATTCTGATGCCAGATCACTGCATCCCTTAC
TTCACCAACATGTTGCCACCTGGTACCCATCAAAGCCAGGCTGCACCTCCAGGAAGAAGCAAT
GCCTGGAGACCCCAGGTCAACAACCCAAAGGAATGGCTGCAAGTGGACTTCCAGAAGACAATGAA
GTCACTGGGTGACAACCCAGGGGGTCAAGTCTCTGCTCACCTCAATGTATGTGAAGGAGTTCTG
ATCTCTTCCCTCACAGGATGGCACCAGTGGACACTCTTCTTCCAGAATGGCAAAGTCAAGGTGTC
CAGGGCAACCAGGACTCTTCACACCTGTGGTGAACTCACGGACCCCCCCCCTCCTGACAAGATA
CTGAGAATTCACCCCCAGTCTGGTCCACCAAGATTGCCCTGAGAATGGAAGTCCCTGGGATGTGAG
GCACAAGACCTGTACTGA (SEQ ID NO:93)

Figure 33B

CS04m2-FL-NA

ATGCAGATTGAGCTGAGCACCTGCCCTGGCTGAGGTTCTGCTTCTGCCACCAAGG
AGATACTACCTGGGGGTGTGGAGCTTCCTGGGACTACATGCAGTCTGAACCTGGGGAGCTGCC
GTGGATGCCAGGTTCCCACCCAGAGTGCCAAATCCTCCCATTCAACACCTCTGGTCTACAAG
AAGACCCCTTTGTGGAGTTCACTGACCACCTGTCACATTGCCAAACCCAGGCCACCCGGATG
GGACTCCGGACCCACCATCGAGCTGAGGTGTATGACACTGTGGCGTCACCCCTCAAGAACATG
GCCCTCCACCCCTGTGAGCCTGCATGCTGTGGGGTCAGCTACTGGAAGTCTCTGAGGGGGCTGAG
TATGATGACCAGACCTCCCAGAGGGAGAAGGGAGATGACAAAGTGTCCCTGGAAAGAGCACACC
TATGTCGGCAGGTCTCAAGGAGATGGCCCCACTGCTCTGACCCACCCCTGCCCTGACCTACTCC
TACCTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGCCCTGCTGGTGTGC
AGGGAGGGCTCCCTGGCAAAGAGAACCCAGACCCCTGCACAAGTCTATTCTCTGTTGCTGTC
TTTGATGAGGGCAAGAGCTGGCACTCTGAAACCAAGAACACTCCCTGATGCAGGACAGGGATGCTGCC
TCTGCCAGGGCTGGCCAAGATGACACACTGTGAATGGCTATGTGAACAGGAGCCTGCCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCACTGGCATGTGATTGGCATGGACAACCCCTGAGGTGAC
TCCATTCTGGAGGGCACACCTTCTGGTCAAGGAAACCACAGACAGGCCAGCCTGGAGATCAGC
CCCATCACCTCCTCACTGCCAGACCCCTGCTGATGGACCTCGGACAGTTCTGCTGTTCTGCCAC
ATCAGCTCCACCCAGCATGATGGCATGGAGGGCTATGTCAGGGTGGACAGCTGCCCTGAGGAGCCA
CAGCTCAGGATGAAGAACAAATGAGGGAGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTTGCGCTTGATGATGACAAACAGCCATCCTTCATTGATCAGATCAGGTCTGTTGCAAGAAA
CACCCCAAGACCTGGTCACTACATTGCTGCTGAGGGAGGACTGGACTATGCCCACTGGTC
CTGGCCCCCTGATGACAGGAGCTACAAGAGCCAGTACCTCAACAATGCCCAACAGAGGATTGGACGC
AAAGTACAAGAAAGTCAGGTTCATGGCTACACTGATGAAACCTCAAGACCAGGGAGGCCATTCA
CATGAGTCTGGCATCCTGGCCCCACTCCTGATGGGAGGTGGGGACACCCCTGCTCATCATCTTC
AAGAACCCAGGCCCTCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCCTGTAC
AGCCGCAGGTGCCAAAGGGGGTGAACACACCTCAAGGACTTCCCTGCTGGGAGATCTTC
AAAGTACAAGTGGACTGCACTGTGGAGGATGGACCAACAAATCTGACCCAGGTGCCCTACCCAGA
TACTACTCCAGCTTGTAACATGGAGAGGGACCTGGCTCTGGCTGATTGGCCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGGGAAACAGATCATGTCAGACAAGAGGAATGTGATTCTG
TTCTCTGCTTGTGATGAGAACAGGAGCTGGTACCTGACTGAGAACATTCAAGCGCTTCTGCCAAC
CCTGCTGGGGTGCAGCTGGAGGACCCCTGAGTCCAGGCCAGCAACATCATGCACTCCATCAATGGC
TATGTTGTTGACAGCCCTCAGCTTCTGTCGCTGCTGATGAGGTGCCCTACTGGTACATTCTTCT
ATTGGGGCCACAGTGAACCTCTTCTGTCTCTTCTCTGGCTACACCTCAAAACACAAGATGGTG
TATGAGGACACCCCTGACCCCTTCTGGGAGACTGTGTTCATGAGCATGGAGAACCC
GCCCTGTGGATTCTGGATGCCACAACACTCTGACTTCCGCAACAGGGCATGACTGCCCTGCTCAAA
GTCTCTCTGTGACAAGAACACTGGGACTACTATGAGGACAGCTATGAGGACATCTGCCCTAC
CTGCTCAGCAAGAACAAATGCCATTGAGCCCAGGGAGCTTCAGGCCAGAACATCCACCTGCTGAAACGC
CACCAGAGGGAGATCACCAGGACACCCCTCCAGTCTGACCCAGGAGGAGATTGACTATGATGACACC
ATTCTGTGGAGATGAAGAAAGAGGACTTTGACATCTATGACGAGGACGAGAACCCAGGCCAAGG
AGCTCCAGAAGAACAGGCCACTACTTCATTGCTGCTGGAGGCCCTGTGGACTATGGCATG
AGCTCCAGCCCCCATGCTCAGGAACAGGGGCCAGTCTGGCTCTGCCCCACAGTTCAAGAAAGTG

(Continued)

Figure 34A

GTCTTCCAAGAGTTCACTGATGGCAGCTCACCCAGCCCCGTACAGAGGGGAGCTGAATGAGCAC
CTGGGACTCCTGGGCCATACATCAGGGCTGAGGTGGAGGACAACATCATGGTGACCTCCGCAAC
CAGGCCTCCAGGCCCTACAGCTTCACTAGCTCCCTCATCAGCTATGAGGAGGACCAGAGGCAGGGG
GCTGAGCCACGCAAGAACCTTGTGAAACCCAATGAAACCAAGACCTACTTCTGAAAGTCCAGCAC
CACATGGCCCCCACCAGGATGAGTTGACTGCAAGGCCTGGGCCACTTCTCTGATGTGGACCTG
GAGAAGGATGTGCACTCTGGCCTGATTGGCCCACTCCTGGCTGCCACACCAACACCCCTGAACCCP
GCCCATGGAAGGCAAGTGAUTGTGCAAGGAGTTGCCCCCTTCTTCAACCATCTTGATGAAACCAAG
AGCTGGTACTTCACTGAGAACATGGAGCGCAACTGCAGGGCCCCATGCAACATTCAGATGGAGGAC
CCCACCTTCAAAGAGAACATACCGCTTCCATGCCATCAATGGCTACATCATGGACACCCCTGCCTGG
CTTGTCACTGGCCCAGGACCAGAGGATCAGGTGGTACCTGCTTCTATGGGCTCCAATGAGAACATT
CACTCCATCCACTCTCTGGGATGCTTCACTGTGCAAGAAGGAGGATACAAGATGGCCCTG
TACAACCTCTACCTCTGGGTCTTGAGACTGTGGAGATGCTGCCCTCAAAGCTGGCATCTGGAGG
GTGGAGTGCCTCATTGGGAGCACCTGCATGCTGGCATGAGCACCTGTTCTGGTCTACAGCAAC
AACTGCCAGACCCCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCTCTGGC
CACTGATGGCCAGTGGGCCCAAGCTGGCCAGGCTCCACTACTCTGGATCCATCAATGCCTGGAGC
ACCAAGGAGCCATTCACTGGGATCAAAGTGGACCTGCTGGCCCCATGATCATCCATGGCATCAAG
ACCCAGGGGCCAGGCAGAACGTTCTCAGCCCTGTACATCAGCCAGTTCATCATGTACAGCTG
GATGGCAAGAAATGGCAGACCTACAGAGGCAACTCCACTGGAACACTCATGGCTTCTGGCAAT
GTGGACAGCTCTGGCATCAAGCACAACATCTCAACCCCCCAATCATGCCAGATACTCAGGCTG
CACCCACCCACTACAGCATCCGCAGCACCTCAGGATGGAGCTGATGGCTGTGACCTGAACCTC
TGCAGCATGCCCTGGGATGGAGAGCAAGGCCATTCTGATGCCAGATCACTGCCTCCAGCTAC
TTCACCAACATGTTGCCACCTGGGAGGCCAAGCAAGGCCAGGCTGCACCTCCAGGGAAAGGAGCAAT
GCCTGGAGGCCAGGTCAACAACCAACAGGAGTGGCTGCAGGTGGACTTCCAGAAAGACCATGAAG
GTCACGGGTGACCACCCAGGGGTCAGAGGCCATTCTGCTACCAGCATGTATGTGAAGGAGTTCTG
ATCAGCTCCAGCCAGGATGGCCACCAGTGGACCTCTTCTCCAGAATGGCAAGGTCAAGGTGTT
CAGGGCAACCAGGACACCTTCACCCCTGTGGTGAACAGGCTGGACCCCCCCCCTGACCAAGATA
CTGAGGATTCAACCCCCAGAGCTGGGTCACCAAGATTGCCCTGAGGATGGAGGTCTGGATGTGAG
GCCCAGGACCTGTACTGA (SEQ ID NO: 94)

Figure 34B

CS04m3-FL-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTCTGTGCCCTGCTGAGGTTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGGCTGTGGAGCTTCTTGGACTACATGCAGTCTGACCTGGGGAGCTGCC
GTGGATGCCAGGTCCCACCCAGACTGCCAAATCCTCCATTCAACACACCTCTGGGTCTACAAG
AAGACCTCTTTGTGGAGTTCACTGACCACCTGTTCAACATTGCCAAACCCAGGCCACCCCTGGATG
GGACTCCTGGGACCCACCATTCAAGGCTGAGGTGATGACACTGTGGTCATCACCCCAAGAACATG
GCCCTCCACCCCTGTGAGCCTGCATGCTGTGGGGTCAGCTACTGGAAGGCCTCTGAGGGGCTGAG
TATGATGACCAGACCTCCCAGAGGGAGAAGGGAGATGACAAAGTGTCCCTGGGGCAGCCACACC
TATGTGTGGCAGGTCTCAAGGAGAATGCCCATGCCCTGACCCACTCTGCCGTACCTACTCC
TACCTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGACTGATTGGGCCTGCTGGTGTGC
AGGGAGGGCTCCCTGGCAAAGAGAAGACCCAGACCCCTGCACAAGTTCAATTCTCTGTTGTC
TTTGATGAGGGCAAGAGCTGGCACTCTGAAACCAAGAACCTCCGTATGCAAGGACAGGGATGTC
TCTGCCAGGGCTGGCCAAGATGCAACTGTGAATGGCTATGTGAACAGGAGCCTGCCGTGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGACAACCCCTGAGGTGCAC
TCCATTTCCTGGAGGGCCACACCTCTGGTCAGGAACCACAGACAGGCCAGCCTGGAGATCAGC
CCCATCACCTCCCTCACTGCCAGACCCCTGCTGATGGACCTCGGAACAGTTCTGCTGTTCTGCCAC
ATCAGCTCCCACCAGCATGATGGCATGGAGGCCTATGTCAAGGTGGACAGCTGCCCTGAGGAGCA
CAGCTCAGGATGAAGAACAAATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCCGCTTGATGATGACAACAGCCCATCCTCATCAGATCAGGTCTGTGGCAAGAAA
CACCCCAAGACCTGGGTGCACTACATTGCTGCTGAGGAGGACTGGACTATGCCCACTGGTC
CTGGCCCTGATGACAGGAGCTACAAGAGCCAGTACCTCAACAATGCCACAGAGGATTGGACGC
AAGTACAAGAAAGTCAGGTTCATGCCCTACACTGATGAAACCTTCAAGACCAGGGAGGCCATTCA
CATGAGTCTGGCATCCTGGGCCACTCCTGTATGGGAGGTGGGACACCCCTGCTCATCATCTTC
AAGAACAGGCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCCTGTAC
AGCCGCAGGCTGCCAAAGGGGTGAAACACCTCAAGGACTTCCCCATTCTGCCCTGGGAGATCTTC
AACTACAAGTGGACTGCACTGTGGAGGATGGACCAACCAAATCTGACCCAGGTGCCCTACCAGA
TACTACTCCAGCTTGATGAACTGGAGAGGGACCTGGCTCTGGCTGATTGGCCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGGGAAACCAAGATCATGTCAGAACAGAGGAATGTGATTCTG
TTCTCTGTCTTGATGAGAACAGGAGCTGGTACCTGACTGAGAACATTCAAGCGCTCCTGCCAAC
CCTGCTGGGTGCACTGGAGGACCCCTGAGTTCCAGGCCAGCAACATCATGCACTCCATCAATGGC
TATGTGTTGACAGCCTCCAGCTTCTGCTGCCCTGCATGAGGTGGCTACTGGTACATTCTTCT
ATTGGGGCCCAGACTGACTTCTTCTGTCTTCTGCTACACCTTCAAACACAAGATGGT
TATGAGGACACCCCTGACCCCTTCTCCATTCTGAGGAGACTGTGTTCATGACCATGGAGAACCT
GGCCTGTGGATTCTGGATGCCACAACACTGACTTCCGCAACAGGGCATGACTGCCCTGCTCAA
GTCTCCCTGTGACAAGAACACTGGGACTACTATGAGGACAGCTATGAGGACATCTGCCCTAC
CTGCTCAGCAAGAACAAATACCACCTACGTGAACCGCTCCCTGAGCCAGAACATCCACCTGTCTGAAA
CGCCACCAGAGGGAGATCACCAGGACCAACCTCCAGTCTGACCAAGGAGGAGATTGACTATGATGAC
ACCATTCTGTGGAGATGAAGAAAGAGGACTTGTACATCTATGACGAGGACAGAACACCAGAGGCCA
AGGAGCTTCCAGAAGAACAGGCCAGGCACTACTTCATTGCTGCTGTGGAGCGCCTGTGGACTATGGC
ATGAGCTCCAGCCCCATGTCCAGGAACAGGGCCAGTCTGGCTCTGTGCCACAGTTCAAGAAA

(Continued)

Figure 35A

GTGGTCTTCCAAGAGTTCACTGATGGCAGCTCACCCAGCCCCGTACAGAGGGGAGCTGAATGAG
CACCTGGGACTCCTGGGCCATACATCAGGGCTGAGGTGGAGGACAACATCATGGTACCTTCGGC
AACCAAGGCCTCCAGGCCCTACAGCTTCTACAGCTCCCTCATCAGCTATGAGGAGGACCAGAGGCAG
GGGGCTGAGCCACGCAAGAACCTTGAAACCCAATGAAACCAAGAACCTACTTCTGGAAAGTCCAG
CACACATGGCCCCCACCAAGGATGAGTTGACTGCAAGGCCTGGGCCTACTCTCTGATGTGGAC
CTGGAGAAGGATGTGACTCTGGCTGATTGGCCACTCCCTGGCTGCCACACCAACACCCCTGAAAC
CCTGCCCATGGAAGGCAAGTGACTGTGCAGGAGTTGCCCTTCTTCACCATTTGATGAAAC
AAGAGCTGGTACTTCAGAACATGGAGCGCAACTGCAGGGCCCATGCAACATTCAAGATGGAG
GACCCCACCTCAAAGAGAACATACCGCTTCCATGCCATCAATGGCTACATCATGGACACCCCTGCC
GGCCTGTCATGGCCAGGACCAGAGGATCAGGTGGTACCTGCTTCTATGGCTCAATGAGAAC
ATTCACTCCATCCACTTCTCTGGCATGTCCTCACTGTGCGCAAGAACGGAGGTACAAGATGGCC
CTGTACAACCTCTACCCCTGGGCTTTGAGACTGTGGAGATGCTGCCCTCAAAGCTGGCATCTGG
AGGGTGGAGTGCCTCATGGGGAGCACUTGCATGCTGGCATGAGCACCCCTGTCCTGGCTACAGC
AACAAAGTGCCAGACCCCCCTGGGAATGCCCTGGCCACATCAGGGACTTCCAGATCACTGCCCT
GCCAGTATGCCAGTGGCCAGGCTCCACTACTCTGGATCCATCAATGCCCTGG
AGCACCAAGGAGGCCATTCAAGCTGGATCAAAGTGGACCTGCTGGCCCCATGATCATCCATGGCATC
AAGACCCAGGGGCCAGGCAGAAGTCTCCAGCCTGTACATCAGCCAGTTCATCATCATGTACAGC
CTGGATGGCAAGAAATGGCAGACCTACAGAGGCAACTCCACTGGAACACTCATGGTCTTGGC
AATGTGGACAGCTGGCATCAAGCACACATCTCAACCCCCCAATCATGCCAGATACTCAGG
CTGCACCCCACCCACTACAGCATCGCAGCACCCCTCAGGATGGAGCTGATGGCTGTGACCTGAAC
TCCTGCAGCATGCCCTGGCATGGAGAGCAAGGCCATTCTGATGCCAGATCACTGCCCTCCAGC
TACCTCACCAACATGTTGCCACCTGGAGGCCAAGCAAGGCCAGGCTGCACCTCCAGGGAGGAGC
AATGCCCTGGAGGCCCCAGGTCAACAACCAAAGGAGTGGCTGCAGGTGGACTTCCAGAAGACCATG
AAGGTCACTGGGTGACCAACCAAGGGGTCAAGAGCCTGCTCACCAGCATGTATGTGAAGGAGTTC
CTGATCAGCTCCAGCCAGGATGGCACCAGTGGACCCCTCTTCTCCAGAATGGCAAGGTCAAGGTG
TTCCAGGGCAACCAGGACAGCTTCACCCCTGTGGTGAACAGCCTGGACCCCCCTCCTGACCAGA
TACCTGAGGATTCAACCCCCAGAGCTGGTCCACCAGATTGCCCTGAGGATGGAGGTCTGGATGT
GAGGCCAGGACCTGTACTGA (SEQ ID NO:95)

Figure 35B

CS04m23-FL-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCTGCTGAGGTTCTGCTTCTCTGCCACCAGG
AGATACTACCTGGGGCTGTGGAGCTTCTGGACTACATGCAGTCAGCTGGGGACCTGCCT
GTGGATGCCAGGGTCCCACCCAGAGTCCCCAAATCCTCCATTCAACACCTCTGTGGCTACAAG
AAGACCCCTTTGTGGAGTTACTGACCACCTGTTAACATTGCCAAACCCAGGCCACCCCTGGATG
GGACTCTGGGACCCACCATTCAAGGTGAGGTGATGACACTGTGGCGTACCCCTCAAGAACATG
GCCTCCCACCCCTGTGAGCTGCATGCTGTGGGGTCAGCTACTGGAAGTCCTCTGAGGGGCTGAG
TATGATGACCAGACCTCCAGAGGGAGAAGGGAGGATGACAAAGTGTCCCTGGGAAGAGCCACACC
TATGTGTGGCAGGTCCCTCAAGGAGAATGGCCCCACTGCCTCTGACCCACCCCTGCCGTACCTACTCC
TACCTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGCCCTGCTGGTGTGC
AGGGAGGGCTCCCTGGCCAAGAGAAGACCCAGACCCCTGCACAAGTCATTCTCCGTGTTGCTGTC
TTTGATGAGGGCAAGAGCTGGCACTCTGAAACCAAGAACCTCCCTGATGCAAGGACAGGGATGCTGCC
TCTGCCAGGGCTGGCCAAGATGCAACTGTGAATGGCTATGTGAACAGGAGCCTGCCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGACAACCCCTGAGGTGCAC
TCCATTTCTGGAGGGCACACCTTCTGGTCAAGGACCTCACAGACAGGCCAGCCTGGAGATCAGC
CCCATCACCTCCCTCACTGCCAGACCCCTGCTGATGGACCTCGGACAGTTCCTGCTGTTGCCAC
ATCAGCTCCCACCCAGCATGATGGCATGGAGGCCATGTCAAGGTGGACAGCTGCCCTGAGGAGCCA
CAGCTCAGGATGAAGAACATGAGGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCCGCTTGTGATGATGACAACAGCCCATCCTCATTCAAGATCAGGTCTGTGGCCAAGAAA
CACCCCAAGACCTGGGTGCACTACATTGCTGCTGAGGGAGGACTGGGACTATGCCCACTGGTC
CTGGCCCTGATGACAGGAGCTACAAGAGCCAGTACCTCAACAATGGCCACAGAGGATTGGACGC
AACTACAAGAAAGTCAGTTCATGGCCTACACTGATGAAACCTTCAGACAGGGAGGCCATTCA
CATGAGTCTGGCATCCTGGCCACTCCTGTATGGGAGGTGGGGACACCCCTGCTCATCATCTC
AAGAACCCAGGCCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCCTGTAC
AGCCGCAGGCTGCCAAAGGGGGTGAACACACCTCAAGGACTTCCCATTCTGCCCTGGGAGATCTC
AACTACAAGTGGACTGCACTGTGGAGGATGGACCAACCAAATCTGACCCAGGTGCTCACCCAGA
TACTACTCCAGCTTGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATTGGCCACTGTCATC
TGCTACAAGGACTCTGTGGACCAGAGGGAAACCAAGATCATGTCTGACAAGAGGAATGTGATTCTG
TTCTCTGTCTTGATGAGAACAGGAGCTGGTACCTGACTGAGAACATTCAAGCCTGCTGCCAAC
CCTGCTGGGGTGCAGCTGGAGGACCCCTGAGTTCCAGGCCAGCAACATCATGCACTCCATCAATGGC
TATGTGTTTGACAGCCTCAGCTTCTGTCTCTGGCTACACCTCAAAACACAAGATGGT
TATGAGGACACCCCTGACCTCTTCTGGGAGACTGTGTTCATGAGCATGGAGAACCC
GGCCTGTGGATTCTGGGATGCCACAACACTCTGACTTCCGCAACAGGGCATGACTGCCCTGCTCAA
GTCTCCTCTGTGACAAGAACACTGGGACTACTATGAGGACAGCTATGAGGACATCTCTGCC
CTGCTCAGCAAGAACAAATACCACTACGTGAACCGCTCCGTGAGCCAGAACCC
CGCCACCAAGAGGGAGATCACCAGGACCAACCTCCAGTCTGACCCAGGAGGAGATTGACTATGAC
ACCATTCTGTGGAGATGAGAACAGGACTTGCACATCTATGACCGAGGACGAGAACCAAGGCC
AGGAGCTCCAGAAGAACACCAGGCACACTTCATTGCTGCTGTGGAGGCCCTGTGGACTATGGC
ATGAGCTCCAGCCCCATGTCCTCAGGAACAGGGCCAGTCTGGCTCTGTGCCACAGTTCAAGAAA

(Continued)

Figure 36A

GTGGTCTTCCAAGAGTTCACTGATGGCAGTTCACCCAGCCCCGTACAGAGGGGAGCTGAATGAG
CACCTGGGACTCCTGGGCCATACATCAGGGCTGAGGTGGAGGACAACATCATGGTACCTTCGC
AACCAAGGCCCTCCAGGCCCTACAGCTTACAGCTCCCTCATCAGCTATGAGGAGGACCAGAGGCAG
GGGGCTGAGCCACGCAAGAACTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAG
CACCAACATGGCCCCCACCAGGATGAGTTGACTGCAAGGCCCTGGCCTACTTCTGTGATGTGGAC
CTGGAGAAGGATGTGACTCTGGCTGATTGGCCACTCTGGCTGCCACACCAACACCCCTGAAC
CCTGCCATGGAAGGCAAGTGAUTGAGGAGCTTGCCCTCTTCTTCAACCATCTTGATGAAACC
AAGAGCTGGTACTTCACTGAGAACATGGAGCGCAACTGCAGGGCCCCATGCAACATTGAGATGGAG
GACCCCACCTTCAAAGAGAACTACCGCTTCCATGCCATCAATGGCTACATCATGGACACCCCTGCCT
GGGCTTGTCAATGCCAGGACCAGAGGATCAGGTGGTACCTGCTTCTATGGGCTCCAATGAGAAC
ATTCACTCCATCCACTCTGGCATGTCTTCACTGTGCGCAAGAAGGAGGAGTACAAGATGGCC
CTGTACAACCTCTACCCCTGGGTCTTGAGACTGTGGAGATGCTGCCCTCAAAGCTGGCATCTGG
AGGGTGGAGTGCCTCATGGGGAGCACCTGCATGCTGGCATGAGCACCCCTGTTCTGGCTACAGC
AACAAAGTGCCAGACCCCCCTGGGAATGCCCTGGCCACATCAGGGACTTCCAGATCACTGCCCT
GGCCAGTATGCCAGTGGGGCCCCAAGCTGCCAGGCTCCACTACTCTGGATCCATCAATGCCCTGG
AGCACCAAGGAGGCATTCACTGGATCAAAGTGGACCTGCTGGCCCCATGATCATCCATGGCATC
AAGACCCAGGGGCCAGGCAGAACAGTTCTCCAGCCTGTACATCAGCCAGTTCATCATGTACAGC
CTGGATGGCAAGAAAATGGCAGACCTACAGAGGCAACTCCACTGGAACACTCATGGCTTCTTGGC
AATGTGGACAGCTCTGGCATCAAGCACACATCTCAACCCCCCAATCATGCCAGATACTCAGG
CTGCACCCACCCACTACAGCATCCGCAGCACCCCTCAGGATGGAGCTGATGGCTGTGACCTGAAC
TCCTGCAGCATGCCCTGGCATGGAGAGCAAGGCCATTCTGATGCCAGATCACTGCCCTCCAGC
TACTTACCAACATGTTGCCACCTGGAGGCCAAGCAAGGCCAGGCTGCACCTCCAGGGAGGAGC
AATGCCCTGGAGGCCAGGTCAACAAACCCAAAGGAGTGGCTGCAGGTGGACTTCCAGAAGACCAG
AAGGTCACTGGGGTGACCACCCAGGGGTCAAGAGCCTGCTCACCAGCATGTATGTGAAGGAGTC
CTGATCAGCTCCAGCCAGGATGCCACCAAGTGGACCTCTTCCAGAATGGCAAGGTCAAGGTG
TTCCAGGGCAACCAGGACAGCTTCAACCCCTGTGGTGAACAGCCTGGACCCCCCCCCCTGACCAGA
TACCTGAGGATTCAACCCCAAGAGCTGGTCCACCAAGATTGCCCTGAGGATGGAGGTCTGGGATGT
GAGGCCAGGACCTGTACTGA (SEQ ID NO: 96)

Figure 36B

CS04ml-FL-NA

ATGCAGATTGAGCTGAGCACCTGCTCTTCCGTGAGGTTCTGCTTCCTGCCACCAAGG
AGATACTACCTGGGGCTGTGGAGCTTCTTGGACTACATGCAGTCTGACCTGGGGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGAGTGCACAAATCCTCCATTCAACACACCTCTGTTCTACAAG
AAGACCCCTTTGTGGAGTTCACTGACCACCTGTTCAACATTGCCAACCCAGGCCACCCCTGGATG
GGACTCTGGGACCCACCAATTCAAGGCTGAGGTGATGACACTGTGGTCATCACCCCAAGAACATG
GCCTCCCACCCCTGTGAGCCTGCATGCTGTGGGGTCAGCTACTGGAAGGGCTCTGAGGGGGCTGAG
TATGATGACCAGACCTCCCAGAGGGAGAAGGAGGATGACAAGTGTCCCTGGGGCAGCCACACC
TATGTTGGCAGGTCTCAAGGAGAATGGCCCCATGGCCTCTGACCCACTCTGCTGACCTACTCC
TACCTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATGGGGCCCTGCTGGTGTG
AGGGAGGGCTCCCTGGCAAAGAGAAGACCCAGCCCTGACAAGTTCAATTCTCTGTTGCTGTC
TTTGATGAGGGCAAGAGCTGGCACTCTGAAACAAAGAACTCCCTGATGCAGGACAGGGATGCTGCC
TCTGCCAGGGCCTGGCCCAAGATGCACACTGTGAATGGCTATGTGAAACAGGAGCCTGCCCTGGACTC
ATTGGCTGCCACAGGAAATCTGPTCACTGGCATGTGATTGGCATGGGACAACCCCTGAGGTGAC
TCCATTTCCTGGAGGGCCACACCTCCCTGGTCAAGGACCCACAGACAGGCCAGCCTGGAGATCAGC
CCCATCACCTCCCTCACTGCCAGACCCCTGCTGATGGACCTCGGACAGTTCCCTGCTGCCCTGCCAC
ATCAGCTCCCACCCAGCATGATGGCATGGAGGCCATGTCAAGGTGGACAGCTGCCCTGAGGAGCCA
CAGCTCAGGATGAAGAACATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCCCGCTTGATGATGACAACAGCCCATCCTTCATTCAAGATCAGGCTGTGGCCAAGAAA
CACCCCAAGACCTGGGTGCACTACATTGCTGCTGAGGAGGAGTGGACTATGCCCACTGGTC
CTGGCCCTGATGACAGGAGCTACAAGAGCCAGTACCTCAACAATGGCCACAGAGGATTGGACGC
AAGTACAAGAAAGTCAGGTTCATGGCCTACACTGATGAAACACCTCAAGACCAGGGAGGCAATTCA
CATGAGTCTGGCATCTGGGCCACTCCTGATGGGAGGTGGGGACACCCCTGCTCATCATTTC
AAGAACCAGGCCCTCCAGGCCCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCCTGTAC
AGCCGCGAGCTGCCAAAGGGGGTAAACACCTCAAGGACTTCCCATCTGCCCTGGGAGATCTTC
AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACAAATCTGACCCAGGTGCCCTACCCAGA
TACTACTCCAGCTTGTGAACATGGAGAGGGACCTGCCCTGTGGCTGATGGCCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGGGAAACCAGATCATGTCAGACAAGAGGAATGTGATTCTG
TTCTCTGTCTTGATGAGAACAGGAGCTGGTAACCTGACTGAGAACATTCAAGCGCTTCCTGCCAAC
CCTGCTGGGGTGCAGCTGGAGGACCTGAGTTCCAGGCCAGCAACATCATGCACTCCATCAATGGC
TATGTTGACAGCCTCCAGTTCTGTCTGCTGATGAGGTGGCTACTGGTACATTCTTCT
ATTGGGGCCAGACTGACTTCCTTCTGCTCTCTCTGGCTACACCTCAAAACACAAGATGGTG
TATGAGGACACCCCTGACCCCTTCCATTCTCTGGGAGACTGTGTTCATGAGCATGGAGAACCT
GGCCTGTGGATTCTGGGATGCCACAACACTGACTTCCGCAACAGGGGAGTGGCTACTGCCCTGCTCAA
GTCTCCTCTGTGACAAGAACACTGGGACTACTATGAGGACAGCTATGAGGACATCTCTGCCCTAC
CTGCTCAGCAAGAACATGCCATTGAGCCCAGGAGCTTCAGCCAGAAATCCACCTGCTCTGAAACGC
CACCAAGAGGGAGATCACCAGGACCAACCTCCAGTCTGACCAAGGAGGAGATTGACTATGATGACACC
ATTCTGTGGAGATGAAGAAAGAGGACTTTGACATCTATGACGAGGACGAGAACCAAGAGGCCAAGG
AGCTCCAGAAAGAACACCAGGCACACTTCATGCTGCTGTGGAGGCCCTGTGGACTATGGCATG
AGCTCCAGCCCCATGTCCCTAGGAACAGGGCCAGTCTGGCTCTGTGCCACAGTTCAAGAAAGTG

(Continued)

Figure 37A

GTCTTCCAAGAGTTCACTGATGGCAGCTCACCCAGCCCCGTACAGAGGGGAGCTGAATGAGCAC
CTGGGACTCCTGGGCCATACATCAGGGCTGAGGTGGAGGACAACATCATGGTACCTCCGCAAC
CAGGCCTCCAGGCCCTACAGCTCTACAGCTCCCTCATCAGCTATGAGGAGGACCAGAGGCAGGGG
GCTGAGCCACGCAAGAACTTGTGAAACCAATGAAACCAAGACCTACTTCTGGAAAGTCCAGCAC
CACATGGCCCCACCAAGGATGAGTTGACTGCAAGGCCTGGCCTACTTCTGTATGTGGACCTG
GAGAAGGATGTGCACTCTGGCCTGATTGGCCACTCCTGGCTGCCACACCAACACCCCTGAACCCCT
GCCCATGGAAGGCAAGTGAATGTGCAGGAGTTGCCCTCTTCACCATCTTGATGAAACCAAG
AGCTGGTACTTCACTGAGAACATGGAGCGCAACTGCAGGGCCCATGCAACATTCAGATGGAGGAC
CCCACCTTCAAAGAGAACATACCGCTTCCATGCCATCAATGGCTACATCATGGACACCCCTGCCTGGG
CTTGTCAATGGCCCAGGACCAGAGGATCAGGTGGTACCTGCTTCTATGGGCTCCAATGAGAACATT
CACTCCATCCACTTCTCTGGCATGTCTTCACTGTGCAGAAGGAGGAGTACAAGATGGCCCTG
TACAACCTCTACCCCTGGGGCTTGTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGGAGG
GTGGAGTGCCTCATGGGGAGCACCTGCATGCTGGCATGAGCACCCCTGTTCTGGTCTACAGCAAC
AAAGTGCAGACCCCCCTGGGAATGGCCTCTGGCACATCAGGGACTTCCAGATCACTGCCTCTGGC
CAGTATGGCCAGTGGGCCCCAAGCTGGCAGGCTCCACTACTCTGGATCCATCAATGCCCTGGAGC
ACCAAGGAGCCATTCACTGGATCAAAGTGGACCTGCTGGGGCCATGATCATCCATGGCATCAAG
ACCCAGGGGCCAGGCAGAAGTTCTCCAGCCTGTACATCAGCCAGTTCACTCATCATGTACAGCCTG
GATGGCAAGAAATGGCAGACCTACAGAGGCAACTCCACTGGAACACTCATGGTCTTGGCAAT
GTGGACAGCTCTGGCATCAAGCACAAACATCTTCAACCCCCCAATCATGCCAGATACATCAGGCTG
CACCCCCACCCACTACAGCATGGCAGCACCCCTCAGGATGGAGCTGATGGGCTGTGACCTGAACCTCC
TGCAGCATGCCCTGGCATGGAGAGCAAGGCCATTCTGATGCCAGATCACTGCCCTCAGCTAC
TTCACCAACATGTTGCCACCTGGAGGCCAAGCAAGGCCAGGCTGCACCTCCAGGGAAAGGAGCAAT
GCCCTGGAGGCCAGGTCAACAACCCAAAGGAGTGGCTGCAGGTGGACTTCCAGAAGACCATGAAG
GTCACTGGGGTGACCACCCAGGGGGTCAAGAGCCTGCTCACCAAGCTGATGTAAGGAGCTCCTG
ATCAGCTCCAGCCAGGATGGCACCAGTGGACCCCTTCTTCCAGAATGGCAAGGTCAAGGTGTT
CAGGGCAACCAGGACAGCTCACCCCTGTGGTGAACAGCCTGGACCCCCCTCCTGACCAGATAAC
CTGAGGATTCACCCCCAGAGCTGGGCTCACCAAGATTGCCCTGAGGATGGAGGTCTGGGATGTGAG
GCCCAAGGACCTGACTGA (SEQ ID NO: 97)

Figure 37B

CS04m13~FL~NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCCTGCTGAGGTTCTGCTTCTCTGCCACCAGG-
AGATACTACCTGGGGCTGTGGAGCTTCTTGGGACTACATGCAGTCTGACCTGGGGAGCTGCCT
GTGGATGCCAGGGTCCCACCCAGAGTGCCTAAATCCTCCATTCAACACACCTCTGTGGTCTACAAG
AAGACCCCTCTTGTGGAGTTCACTGACCACCTGTTCAACATTGCCAACCCAGGCCACCCCTGGATG
GGACTCCTGGGACCCACCATTCAGGCTGAGGTTATGACACTGTGGTCATCACCCCTCAAGAACATG
GCCTCCCACCCCTGTGAGGCTGCATGCTGTGGGGTCAGCTACTGGAAGGCCTCTGAGGGGGCTGAG
TATGATGACCAGACCTCCCAGAGGGAGAAGGAGGATGACAAAGTGTCCCTGGGGCAGCCACACC
TATGTGTGGCAGGTCTCAAGGAGAATGGCCCCATGGCCTCTGACCCACTCTGCCCTGACCTACTCC
TACCTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGCCCTGCTGGTGTGC
AGGGAGGGCTCCCTGGCAAAGAGAAGACCCAGACCCCTGCACAAGTTCATTCTCTGTTGCTGTC
TTTGATGAGGGCAAGAGCTGGCACTCTGAAACCAAGAACCTCCCTGATGCAGGACAGGGATGCTGCC
TCTGCCAGGGCTGGCCAAGATGCCACACTGTGAATGGCTATGTGAAACAGGACCCCTGCCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCTACTGGCATGTGATTGGCATGGGACAACCCCTGAGGTGCAC
TCCATTTCCTGGAGGGCACACCTCCTGGTCAGGAACCACAGACAGGCCAGCCTGGAGATCAGC
CCCATCACCTTCTCACTGCCAGACCCCTGCTGATGGACCTCGGACAGTCCCTGCTGTCCCTGCCAC
ATCAGCTCCCACCAGCATGGCATGGAGGCATGTCAAGGTGGACAGCTGCCCTGAGGAGCCA
CAGCTCAGGATGAAGAACATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCGGCTTGTGATGATGACAACAGCCCACCTTCATTCAAGATCAGGTCTGTGGCCAAGAAA
CACCCCAAGACCTGGGTGCACTACATTGCTGCTGAGGAGGAGACTGGACTATGCCCACTGGTC
CTGGCCCTGATGACAGGAGCTACAAGAGCCAGTACCTCAACAATGGCCACAGAGGATTGGACGC
AAGTACAAGAAAGTCAGGTTCATGGCTACACTGATGAAACCTTCAAGACCAGGGAGGCCATTCA
CATGAGTCTGGCATCTGGGCCACTCCTGTATGGGAGGTGGGGACACCCCTGCTCATCATCTTC
AAGAACCAGGCCTCCAGGCCCTACAAACATCTACCCACATGGCATCACTGATGTCAAGGCCCCTGTAC
AGCCGCAGGCTGCCAAAGGGGGTAAACACCTCAAGGACTTCCCCATTCTGCCCTGGGAGATCTTC
AAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACCAAATCTGACCCAGGTGCCCTCACCAGA
TACTACTCCAGCTTGTGAAACATGGAGAGGGACCTGGCCTCTGGCTGATTGGCCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGGGAAACCAAGATCATGTCTGACAAGAGGAATGTGATTCTG
TTCTCTGTCTTGTGAGAACAGGAGCTGGTACCTGACTGAGAACATTCAAGCCTCTGCCCAAC
CCTGCTGGGTGCACTGGAGGACCTGAGTTCCAGGCCAGCAACATCATGCACTCCATCAATGGC
TATGTGTTGACGCCTCCAGCTTCTGTCTGCCCTGATGAGGTGCCACTGGTACATTCTTCT
ATTGGGGCCAGACTGACTTCTTCTGTCTTCTCTGGCTACACCTTCAAAACACAAGATGGTG
TATGAGGACACCTCTGACCCCTTCCATTCTCTGGGAGACTGTGTTCATGAGCATGGAGAACCCCT
GGCCTGTTGAGTCTGGATGCCACAACACTCTGACTTCCGCAACAGGGCATGACTGCCCTGCTCAAA
GTCTCCTCTGTGACAAGAACACTGGGACTACTATGAGGACAGCTATGAGGACATCTGCCTAC
CTGCTCAGCAAGAACAAATACCACCTACGTGAACCGCTCCCTGAGCCAGAATCCACCTGCTGAAA
CGCCACCAAGAGGGAGATCACCAGGACCAACCCCTCCAGTCTGACCAAGGAGGAGATTGACTATGAC
ACCATTCTGTGGAGATGAAGAAAGAGGACTTGTACATCTATGACGAGGACGAGAACCAAGAGCCCA
AGGAGCTTCCAGAAGAACAGCAGGCACTACTTCATTGCTGCTGTGGAGCGCCTGTGGACTATGGC
ATGAGCTCCAGCCCCATGTCCAGGAACAGGGCCAGTCTGGCTCTGTGCCACAGTTCAAGAAA

(Continued)

Figure 38A

GTGGTCTTCCAAGAGTTCACTGATGGCAGCTCACCCAGCCCCGTACAGAGGGAGCTGAATGAG
CACCTGGGACTCCTGGGCCATACATCAGGGCTGAGGTGGAGGACAACATCATGGTACCTTCGC
AACCAAGGCCTCCAGGCCATACAGCTTCTACAGCTCCCTCATCAGCTATGAGGAGGACCAGAGGCAG
GGGGCTGAGCCACGCAAGAACTTGTGAAACCCATGAAACCAAGACCTACTCTGAAAGTCCAG
CACCAACATGGCCCCCACCAAGGATGAGTTGACTGCAAGGCCCTGGGCTACTCTCTGATGTGGAC
CTGGAGAAGGATGTGACTCTGGCCTGATTGGCCACTCCTGGTCTGCCACACCAACACCCCTGAAC
CCTGCCCATGGAAGGCAAGTGACTGTGCAGGAGTTGCCCTCTTCTTACCCATCTTGATGAAACC
AAGAGCTGGTACTTCACTGAGAACATGGAGCGCAACTGCAGGGCCCCATGCAACATTCAAGATGGAG
GACCCCACCTTCAAAGAGAACATACCGCTTCCATGCCATCAATGGCTACATCATGGACACCCCTGCCT
GGGCTTGTATGGCCAGGACCAGAGGATCAGGTGGTACCTGCTTCTATGGGCTCCAATGAGAAC
ATTCACTCCATCCACTTCTCTGGGCATGTCTTCACTGTGCGCAAGAAGGAGGAGTACAAGATGCC
CTGTACAACCTCTACCCCTGGGTCTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGG
AGGGTGGAGTGCCTCATTGGGAGCACCTGCATGCTGGCATGAGCACCTGTTCTGGTACAGC
AACAAAGTGCCAGACCCCCCTGGGAATGGCCTCTGGCCACATCAGGGACTTCCAGATCACTGCCCT
GGCCAGTATGGCCAGTGGCCCCAAGCTGGCCAGGCTCCACTACTCTGGATCCATCAATGCCATGG
AGCACCAAGGAGGCCATTAGCTGGATCAAAGTGGACCTGCTGGCCCCATGATCATCCATGGCATE
AAGACCCAGGGGCCAGGAGAAGTTCTCCAGCCTGTACATCAGCCAGTTCACTCATCATGTACAGC
CTGGATGGCAAGAAATGGCAGACCTACAGAGGAACCTCCACTGGAACACTCATGGTCTTGGC
AATGTGGACAGCTCTGGCATCAAGCACACATCTCAACCCCCCAATCATGCCAGATACTCAGG
CTGCACCCACCAACTACAGCATCCGACCCCTCAGGATGGAGCTGATGGGCTGTGACCTGAAC
TCCTGCAGCATGCCCTGGCATGGAGAGCAAGGCCATTCTGATGCCAGATCACTGCCCTCCAGC
TACTTCACCAACATGTTGCCACCTGGAGGCCAAGCAAGGCCAGGCTGCACCTCAGGGAGGAGC
AATGCCCTGGAGGCCCAAGCTCAACAAACCAAAGGAGTGGCTGCAGCTGGACTTCAGAACAGCATG
AAGGTCACTGGGGTGAACACCCAGGGGTCAASAGCCTGCTCACCAGCATGTATGTGAAGGAGTTC
CTGATCAGCTCCAGCCAGGATGGCCACCAAGTGGACCCCTCTTCCAGAATGGCAAGGTCAAGGTG
TTCCAGGGCAACCAGGACAGCTTACCCCTGTGGTGAACAGCCTGGACCCCCCCTCTGACCAGA
TACCTGAGGATTCACCCCAAGAGCTGGTCCACCAAGATTGCCCTGAGGATGGAGGTCTGGATGT
GAGGCCAGGACCTGTACTGA (SEQ ID NO: 98)

Figure 38B

CS23m13-FL-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTCCGTGCCCTGCTGAGGTTCTGCTCTGCCACCAAGG
AGATACTACCTGGCGCCGTGGAGCTGAGCTGGACTACATGCAGTCTGACCTGGCGAGCTGCCT
GTGGACGCCAGGTTCCCCCCCAGAGTGCCAAGAGCTTCCCTTCACACACCTCAGTGGTGTACAAG
AAGACCTGTTCTGGAGTTCACCGACCACCTGTTCAACATGCCAACGCCAGGCCCCCTGGATG
GCCCTGCTGGCCCCCACCATCCAGGCCGAGGTGTACGACACCCTGGTGTACCCCTGAAGAACATG
GCCAGCCACCCCGTGGAGCTGCACGCCGTGGCGTGAGCTACTGGAAGGCCCTGAGGGCGCCGAG
TATGACGACCAGACCAGCCAGAGGGAGAAGGAGGACGACAAGGTGTTCCCCGGCGCAGCCACACC
TACGTGTGGCAGGTGCTGAAGGAGAACGGCCCCATGCCAGCGACCCCTGTGCCTGACCTACAGC
TACCTGAGCCACGTGGACCTGGTGAAGGACCTGAACTCTGGCCTGATCGGCCCTGCTGGTGTG
AGGGAGGGCAGCCTGGCCAAGGAGAACCCAGACCCCTGCACAAGTTCATCTGCTGTTGCCGTG
TTCGATGAGGGCAAGAGCTGGCACAGCGAGACCAAGAACAGCCTGATGCAGGACAGGGATGCC
TCTGCCAGGGCCTGGCCAAGATGCACACCGTGAACGGCTACGTGAAACAGGAGCCTGCCGGCCTG
ATCGGCPGCCACAGGAAGTCTGTGACTGGCACGTGATGGCATGGCACCCAGGGCAGCCTGGAGATCAGC
AGCATCTCCTGGAGGGCCACACCTCCTGGTGAAGGACCAAGGCCAGGCTGGAGATCAGC
CCCATCACCTCCTGACCCGCCAGACCCCTGCTGATGGACCTGGCCAGTTCTGCTGTCCTGCCAC
ATCAGCAGCCACCAGCACGGCATGGAGGCTACGTGAAGGTGACAGCTGCCCGAGGAGGCC
CAGCTGAGGATGAAGAACACGAGGAGGCCAGGACTATGATGATGACCTGACCGACTCTGAGATG
GACGTGGTGAGGTTGATGATGACAACAGCCCCAGCTCATCCAGATCAGGTCTGTGGCCAAGAAC
CACCCCAAGACCTGGGTGCACTACATGCCGCCAGGGAGGACTGGGACTACGCCCTGGTG
CTGGCCCCGAGCACAGGAGCTACAAGAGCCAGTACCTGAACAACGGCCCCAGAGGATGCCAGG
AACTACAAGAAGGTGAGATTGATGCCCTACACCGACGAGACCTTCAAGACCAGGGAGGCCATCCAG
CAGCAGTCTGGCATCTGGCCCCCTGCTGTAACGGCAGGTGGCGACACCCCTGCTGATCATCTTC
AAGAACCAAGGCCAGCAGGCCCTACAACATCTACCCCAACGGCATCACCGATGTGAGGCCCTGTAC
AGCAGGAGGCTGCCCAAGGGCGTGAAGCACCTGAAGGACTTCCCCATCTGCCGGCGAGATCTTC
AACTACAAGTGGACCGTGACCGTGAGGATGCCCAAGCTGACCCCAAGGTGCCCTGACCAAGG
TACTACAGCAGCTCGTGAACATGGAGAGGGACCTGGCTCTGGCTGATGCCCTGGCTGATC
TGCTACAAGGAGAGCGTGGACCAGAGGGCAACCAGATCATGTCGACAAGAGGAACGTGATCCTG
TTCTCTGTTGATGAGAACAGGAGCTGGTATCTGACCGAGAACATCCAGAGGTTCTGCCAAC
CCCGCCGGCGTGCAGCTGGAGGACCCGAGTTCCAGGCCAGCAACATCATGACACAGCATCAACGGC
TACGTGTTGACAGCCTGCAGCTGTCGAGCTGTCGAGGACTGGTACATCTGAGC
ATCGGCGCCAGACCGACTTCTGTCGTTCTCTGGCTACACCTTCAAGCACAAGATGGTG
TACGAGGACACCTGACCCCTGTTCCCCCTCAGCGGCCAGGACCGTGGTATGAGCATGGAGAACCCC
GGCCTGTTGATCCTGGCTGCCACACAGCGACTTCAGGAACAGGGCATGACCGCCCTGCTGAAA
GTCAGCAGCTGCAGCAAGAACACCGCGACTACTACGAGGACAGCTACGAGGACATCAGCGCTAC
CTGCTGAGCAAGAACACCCACCTACGTGAACCGCTCCCTGAGCCAGAACCCCCCGTGTGAG
AGGCACCAAGGGAGATCACCAGGACCAACCTGAGGACAGGAGGAGATCGACTATGATGAC
ACCATCAGCGTGGAGATGAAGAACGGACTTCGACATCTACGACGAGGACGAGAACCAAGAGCCCC
AGGAGCTCCAGAAGAACCCAGGACTACTTCATGCCGCCGTGGAGAGGCTGTGGACTATGGC
ATGAGCAGCAGCCCCCACGTGCTGAGGAACAGGGCCAGAGCGGCCAGCGTGCCCTGAGTTCAAGAAC

(Continued)

Figure 39A

GTGGTGTCCAGGAGTTACCGACGGCAGCTTCACCCAGCCCCGTACAGAGGCAGCTGAACGAG
CACCTGGGCCCTGCTGGGCCCTACATCAGGGCCAGGTGGAGGACAACATCATGGTGACCTTCAGG
AACCAGGCCAGCAGGCCCTACAGCTCTACAGCAGCCTGATCAGCTACGGAGGAGGACAGAGGCAG
GGGCCGAGGCCAGGAAGAACCTCGTGAAGCCCACGGAGACCAAGACCTACTTCCTGGAAGGTGCAG
CACACACATGGCCCCCACCAAGGACGAGITTCGACTGCAAGGCCCTGGGCCTACTTCCTGATGTGGAC
CTGGAGAAGGACGTGCACAGCGGCCCTGATCGGCCCCCTGCTGGTGTGCCACACCAACACCCCTGAAC
CCCGCCACGCCAGGCAGGTGACCGTGCAGGAGTTGCCCTGTTCTCACCATCTCGACGAGACC
AAGAGCTGGTACTTCACCGAGAACATGGAGAGGAACACTGCAGGGCCCCCTGCAACATCCAGATGGAG
GACCCCCACCTCAAGGAGAACATCAGGTTCCACCCCATCAACGGCTACATCATGGACACCCCTGCC
GGCCTGGTATGGCCCAGGACCAAGAGGATCAGGTGGTATCTGCTGAGCATGGCAGCAACGAGAAC
ATCCACAGCATCCACTTCAGCGGCCACGTGTTACCGTGAGGAAGAAGGAGGAGTACAAGATGGCC
CTGTACAACCTGTACCCCCGGCTGTTGAGACCGTGGAGATGCTGCCAGCAAGGCCGGCATCTGG
AGGGTGGAGTGCCTGATCGCGAGCACCTGCACGCCGGCATGAGCACCCCTGTTCTGGTGTACAGC
AACAAAGTGCCAGACCCCCCTGGCATGGCCAGCAGGCCACATCAGGGACTTCCAGATCACGCCCT
GGCAGTACGCCAGTGGCCCCAAGCTGCCAGGCTGCACTACAGCGGCAGCATCACGCCCTGG
AGCACCAAGGAGCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCCCCATGATCATCCACGGCATC
AAGACCCAGGCCAGGCAGAAGTCAGCAGCCTGTACATCAGCCAGTTCATCATGTACAGC
CTGGACGGCAAGAAGTGGCAGACCTACAGGGCAACAGCACCGGCACCCCTGATGGTGTCTCGGC
AACGTGGACAGCAGCGGCATCAAGCACACATCTTCAACCCCCCATCATGCCAGGTACATCAGG
CTGCACCCCACCCACTACAGCATCAGGAGCACCCCTGCCATGGAACTGATGGCTGCCACCTGAAC
AGCTGCAGCATGCCCTGGCATGGAGAGCAAGGCCATCTGCTGACGCCAGATCACGCCAGCAGC
TACTTCACCAACATGTTGCCACCTGGAGCCCAAGCAAGGCCAGGCTGCACTGCCAGGGCAGGAGC
AACGCCCTGGAGGCCAGGTGAACAAACCCAAGGAGTGGCTGCAGGTGGACTTCCAGAAGACCATG
AAGGTGACCGGGCTGACCAACCCAGGGCTGAAGAGCCTGCTGACCGAGCATGTACGTGAAGGAGTTC
CTGATCAGCAGCAGCCAGGACGCCACAGTGGACCCCTGTTCTCCAGAACGGCAAAGTGAAGGTG
TTCCAGGGCAACCAGGACAGCTTCACCCCCCTGTTGACAGGCCCTGGACCCCCCTGCTGACCAAGG
TATCTGAGGATCCACCCCCAGAGCTGGGTGCACAGATGCCCTGAGAATGGAAGTGTGGGATGC
GAGGCCAGGACCTGTACTGA (SEQ ID NO: 99)

Figure 39B

CS23m3-FL-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTGCCTGCTGAGGTCTGCTTCTGCCACCAGG
AGATACTACCTGGCGCCGTGGAGCTGAGCTGGACTACATGCAGTCTGACCTGGCGAGCTGCCT
GTGGACGCCAGGTTCCCCCCCCAGAGTGCCCAAGAGCTTCCCCTCAACACCTCAGTGGTGACAAG
AAGACCCCTGTTCTGGAGGTTACCGACCACCTGTTCAACATGCCAAGGCCAGGGCCCCCTGGATG
GCCCTGCTGGGGCCCCACCATCCAGGGCAGGTGTACGACACCCTGATCACCTGAAGAACATG
GCCAGCCACCCCGTGAGCTGCACGCCGTGGCGTAGCTACTGGAAGGCCCTGAGGGCGCCGAG
TATGACGACCAGACCAGCCAGAGGGAGAAGGAGGACGACAAGGTGTCCCCGGGGCAGCCACACC
TACGTGTGGCAGGTGCTGAAGGAGAACGGCCCATGGCCAGCGACCCCTGTGCCTGACCTACAGC
TACCTGAGCCACGTGGACCTGGTGAAGGACCTGAACCTGCTGGCCTGATCGGCGCCCTGCTGGTGTG
AGGGAGGGCAGCCTGGCCAAGGAGAAGACCCAGACCCCTGACAAAGTTCATCTGCTGTTGCCGTG
TTCGATGAGGGCAAGAGCTGGCACAGCGAGACCAAGAACAGCCTGATGCAAGGACAGGGATGCCGCC
TCTGCCAGGGCCTGGCCAAGATGCCACACCGTGAACGGCTACGTGAACAGGAGCCTGCCGGCCTG
ATCGGCTGCCACAGGAAGTCTGTGACTGGCACGTGATGGCATGGCACCAACCCCGAGGTGAC
AGCATCTTCTGGAGGGCACACCTCCTGGTGAAGAACAGGCAAGGCCAGGCTGGAGATCAGC
CCCATCACCTCCTGACCGCCCAGACCCCTGCTGATGGACCTGGCCAGTCCCTGCTGTTGCCAC
ATCAGCAGCCACCAGCACGGATGGAGGCTACGTGAAGGTGGACAGCTGCCCGAGGAGCCC
CAGCTGAGGATGAAGAACACAGGAGGCCAGGGACTATGATGATGACCTGACCGACTCTGAGATG
GACGTGGTGAGGTTTGATGATGACAACAGCCCCAGCTTCATCCAGATCAGGCTGTGGCCAAGAAG
CACCCCAAGACCTGGGTGCACTACATGCCGCCAGGGAGGACTGGGACTACGCCCGCTGGT
CTGGCCCCGACGACAGGAGCTACAAGAGCCAGTACCTGAACACGCCAGGAGGATGGCAGG
AACTACAAGAAGGTCAGATTGATGGCTACACCGACGAGACCTCAAGACCAGGGAGGCCATCCAG
CACGAGTCTGGCATCCTGGCCCCCTGCTGTACGGCGAGGTGGCGACACCCCTGCTGATCATCTTC
AAGAACCCAGGCAGCAGGCCCTACAACATCTACCCCCACGGCATACCGATGTGAGGCCCTGTAC
AGCAGGAGGCTGCCAAGGGCGTGAAGCACCTGAAGGACTTCCCCATCCTGCCGGCGAGATCTTC
AACTACAAGTGGACCGTGAACGTGGAGGATGGCCCCACCAAGTCTGACCCAGGTGCTGACCAGG
TACTACAGCAGCTTGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATCGGCCCTGCTGATC
TGCTACAAGGAGAGCGTGACAGAGGGCAACCAGATCATGCTGACAAGAGAACGTGATCCTG
TTCTCTGTGTTGATGAGAACAGGAGCTGGTATCTGACCGAGAACATCCAGAGGTTCTGCCAAC
CCCCGGCGTGCAGCTGGAGGACCCGAGTTCCAGGCCAGCAACATCATGCACAGCATCACGGC
TACGTGTTGACAGCCTGAGCTGTGTGCTGCACAGGGCTACTGGTACATCTGAGC
ATCGGCCAGACCCCTGACCTGCTGTGTTCTCTGGCTACACCTCAAGCACAAGATGGTG
TACGAGGACACCCCTGACCTGCTTCTCCCTCAGGGCGAGACCGTGTGATGAGCATGGAGAACCCC
GCCCTGTTGATCTGGCTGCCACACAGCGACTTCAGGAACAGGGCATGACGCCCTGCTGAAA
GTCAGCAGCTGCGACAAGAACACCGCGACTACTACGAGGACAGCTACGAGGACATCAGGCCCTAC
CTGCTGAGCAAGAACACACCACTACGTGAACCGCTCCCTGAGGCCAGAACCCCCCTGCTGAAG
AGGCACCAGAGGGAGATCACCAGGACCAACCCCTGCAAGGCCAGGAGGAGATCGACTATGAC
ACCATCAGCGTGGAGATGAAGAACAGGAGCTCGACATCTACGACGAGGACGAGAACAGAGCCC
AGGAGCTTCCAGAAGAACAGGCCACTACTCATGCCGCCGTGGAGAGGCTGTGGACTATGGC
ATGAGCAGCAGCCCCCACGTGCTGAGGAACAGGGCCCAGAGCGCAGCGTCCCCAGTTCAAGAAC

(Continued)

Figure 40A

GTGGTGTCCAGGAGTTACCGACGGCAGCTCACCCAGCCCCGTACAGAGGCAGCTGAACGAG
CACCTGGGCCTGCTGGGCCCTACATCAGGGCGAGGTGGAGGACAACATCATGGTACCTTCAGG
AACCAAGGCCAGCAGGCCCTACAGCTTCTACAGCAGCCTGATCAGCTACGAGGAGGACCAGAGGCAG
GGCGCCGAGCCCAGGAAGAACTTCGTGAAGCCCCAACGAGACCAAGACCTACTTCTGGAAAGGTGCAG
CACACATGGCCCCCACCAAGGACGAGTTCGACTGCAAGGCCCTACTTCTCTGATGTGGAC
CTGGAGAAGGACGTGCACAGCGGCCTGATCGGCCCTGCTGGTGTGCCACACCAACACCCCTGAAC
CCCAGCCACGGCAGGCAGGTGACCGTGCAGGAGTTGCCCTGTTCTCACCATCTCGACGAGACC
AAGAGCTGGTACTTCACCGAGAACATGGAGAGGAACCTGCAGGGCCCCCTGCAACATCCAGATGGAG
GACCCCCACCTTCAGGAGAACATACAGGTTCCACGCCATCACGGCTACATCATGGACACCCCTGCC
GGCCTGGTATGGCCAGGACCAAGAGGATCAGGTGGTATCTGCTGAGCATGGCAGCAACGAGAAC
ATCCACAGCATTCACCTCAGCGGCCACGTGTTGACCGTGGAGAACAGGAGGAGTACAAGATGCC
CTGTACAACCTGTACCCCGGCGTGTGAGACCGTGGAGATGCTGCCAGCAAGGCCGATCTGG
AGGGTGGAGTGCCTGATCGCGAGCACCTGCACGCCGATGAGCACCCCTGTTCTGGTACAGC
AACAGTGCCAGACCCCCCTGGCATGGCCAGCGGCCACATCAGGGACTCCAGATCACCGCCTCT
GGCCAGTACGGCAGTGGCCCCAAGCTGGCAGGCTGACTACAGCGGCAGCATCAACGCCCTGG
AGCACCAAGGAGCCCCTCAGCTGGATCAAGGTGGACCTGCTGGCCCCATGATCATCCACGGCATC
AAGACCCAGGGCGCCAGGACAGGTTCAGCAGCCTGTACATCAGCCAGTTCATCATGTACAGC
CTGGACGGCAAGAAGTGGCAGACCTACAGGGCAACAGCACCGGCACCCCTGATGGTGTCTCGGC
AACGTGGACAGCAGCGGCATCAAGCACACATCTCAACCCCCCATCATGCCAGGTACATCAGG
CTGCACCCCAACCACTACAGCATCAGGAGCACCTGCGGATGGAACGTGATGGCTGGCACCTGAAC
AGCTGCAGCATGCCCTGGCATGGAGAGCAAGGCCATCTGACGCCAGATCACGCCAGCAGC
TACTTCACCAACATGTTGCCACCTGGAGGCCAGCAAGGCCAGGCTGCACCTGCAGGGCAGGAGC
AACGCCCTGGAGGCCAGGTGAACAACCCCAAGGAGTGGCTGCAGGTGGACTTCAGAAGACCAG
AAGGTGACCGGGCTGACCACCCAGGGCGTGAAGAGCCTGACCAGCATGTACGTGAAGGAGTTC
CTGATCAGCAGCAGGCCAGGACAGGCCACCGAGTGGACCCCTGTTCTTCCAGAACGGCAAAGTGAAGGTG
TTCCAGGGCAACCAGGACAGCTCACCCCCGTGGTGAACAGCCTGGACCCCCCTGCTGACCAGG
TATCTGAGGATCCACCCCCAGAGCTGGGTGCACAGATGCCCTGAGAATGGAAGTGTGCTGGATGC
GAGGCCAGGACCTGTACTGA (SEQ ID NO:100)

Figure 40B

CS23m2~FL~NA

ATGCAGATTGAGGTGAGCACTGCTTCTTCCTGTGCCCTGCTGAGGTTCTGCTTCTCTGCCACCCAGG
AGATACTACCTGGCGCCGTGGAGCTGAGCTGGACTACATGCAGTCTGACCTGGCGAGCTGCCT
GTGGACGCCAGGTTCCCCCCCAGAGTGCCAAGAGCTTCCCTTCAACACCTCAGTGGTGTACAAG
AAGACCTGTTCTGGAGTTCACCGACCACCTGTTAACATGCCAAGCCCAGGCCCCCTGGATG
GGCTGCTGGCCCCACCATCCAGGCCGAGGTGTACGACACCGTGGTGTACCTGAGGGCGCGAG
GCCAGCCACCCCGTGGAGCTGCACGCCGTGGCGTGAGCTACTGGAAGTCTCTGAGGGCGCGAG
TATGACGACCAGACCAGCCAGAGGGAGAAGGGAGGACGACAAGGTGTTCCCOGGCAAGAGGCCACACC
TACGTGTGGCAGGTGCTGAAGGAGAACGGCCCCACTGCCAGCGACCCCCCTGCCCTGACCTACAGC
TACCTGAGCCACGTGGACCTGGTGAAGGACCTGAACCTGCTGGCCTGATCGGCCCTGCTGGTGTGC
AGGGAGGGAGCCTGGCCAAGGAGAACCCAGACCTGACAAGTTCATCCTGCTGTCGCCGTG
TTCGATGAGGGCAAGAGCTGGCACAGCGAGACCAAGAACAGCCTGATGCAAGGACAGGGATGCCGCC
TCTGCCAGGGCTGGCCAAGATGCACACCGTGAACGGCTAACGTGAACAGGAGCCTGCCGGCTG
ATCGGCTGCCACAGGAAGTCTGTGACTGGCACGTGATGGCATGGCACCCAGGGAGGTGCAAC
AGCATCTTCTGGAGGGCCACACCTTCTGGTGAAGAACCCAGGCCAGCCTGGAGATCAGC
CCCACATCACCTTCTGCCAGACCCCTGCTGATGGACCTGGCCAGTTCTGCTGTTCTGCCAC
ATCAGCAGCCACCAAGCACGACGGCATGGAGGCTAACGTGAAGGTGGACAGCTGCCCGAGGAGGCC
CAGCTGAGGATGAAGAACACGAGGCCAGGACTATGATGATGACCTGACCGACTCTGAGATG
GACGTGGTGGAGGTTGATGATGACAACAGCCCCAGCTTCATCCAGATCAGGCTGTTGGCCAAGAAC
CACCCCAAGACCTGGTGCACATCACGCCGGAGGAGGACTGGACTAACGCCCTGGTG
CTGGCCCCCGACAGCAGGAGCTAACAGAGCCAGTACCTGAACAACGGCCCCAGAGGATGCCAGG
AACTACAAGAAGGTCAAGATTCACTGGCCTAACCGACGAGACCTTCAAGACCAGGGAGGCCATCCAG
CACGAGTCTGGCATCTGGCCCCCTGCTGTACGGCGAGGTGGCGACACCTGCTGATCATCTTC
AAGAACCAAGGCCAGCAGGCCCTAACACATCTACCCCCACGGCATCACCAGATGTGAGGCCCTGTAC
AGCAGGAGGTGCCCCAAGGGCGTGAAGCACCTGAAAGGACTTCCATCTGGCCTGAGGAGATCTTC
AACTACAAGTGGACCGTGAACATGGAGGGACCTGGCTCTGGCTGATCGGCCCTGCTGATC
TGCTACAAGGAGAGCGTGGACCAGAGGGCAACCGACATGTCTGACAAGAGGAACGTGATCTG
TTCTCTGTGTTGATGAGAACAGGAGCTGGTATCTGACCGAGAACATCCAGAGGTTCTGCCAAC
CCCGCCGGCGTGCAGCTGGAGGACCCGAGTTCCAGGCCAGCAACATCATGCACAGCATCAACGGC
TACGTGTTGACAGCCTGCAAGCTGTGCTGTGCTGCCAGGAGTGGCTACTGGTACATCCTGAGC
ATCGGCGCCCAGACCGACTTCCCTGTGTTCTCTGGCTACACCTTCAAGCACAAGATGGTG
TACGAGGACACCCCTGACCCCTGTTCTGAGGAGACCGTGGCTGAGGAGATCAGCAGTGGAGAAC
GGCCTGTGGATCCTGGCTGCCAACACGGCAACTTCAAGAACAGGGCATGACCGCCCTGCTGAAA
GTCAGCAGCTGCGACAAGAACACGGCGACTACTACGAGGAGACAGCTACGAGGACATCAGGCCCTAC
CTGCTGAGCAAGAACACGCCATCGAGCCAGGGAGCTTCAAGCAGAACCCCCCGTGTGAAGAGG
CACCAGAGGGAGATCACCAAGGACCCCTGCAAGAGCGACCCAGGAGGAGATCAGTATGACACC
ATCAGCGTGGAGATGAAGAAGGAGACTTCGACATCTACGACGAGGACGAGAACAGAGGCCAGG
AGCTTCCAGAAGAACAGGCCAGGGCACTACTTCATCGCCGCCGTGGAGAGGGCTGTGGACTATGGCATG
AGCAGCAGCCCCCACGTGCTGAGGAACAGGGCCAGAGCGGCAGCGTCCCCAGTTCAAGAAGGTG

(Continued)

Figure 41A

GTGTTCCAGGAGTTCACCGACGGCAGCTTCACCCAGCCCCGTACAGAGGCAGACTGAACGAGCAC
CTGGGCCTGCTGGGCCCTACATCAGGGCCGAGGTGGAGGACAACATCATGGTACCTTCAGGAAC
CAGGCCAGCAGGCCCTACAGCTCTACAGCAGCCTGATCAGCTACGAGGAGGACCAGAGGCAGGGC
GCCGAGGCCAGGAAGAACTCGTGAAGCCAAACGAGACCAAGACCTACTTCTGGAGGTGCAGCAC
CACATGGCCCCCACCAAGGACGAGTTCGACTGCAAGGCCTGGGCCACTTCTGTGATGTGGACCTG
GAGAAGGACGTGCACAGGCCCTGATGGGCCCTGCTGGTGTGCCACACCAACACCCCTGAACCCC
GCCCACGGCAGGCAGGTGACCGTGAGGAGTTGCCCTGTTCTCACCATCTCGACGAGACCAAG
AGCTGGTACTTCACCGAGAACATGGAGAGGAACCTGCAGGGCCCTGCAACATCCAGATGGAGGAC
CCCACCTCAAGGAGAACATCAGGTTCCACGCCATCAACGGCTACATCATGGACACCCCTGCCCGC
CTGGTGTGGCCAGGACCAAGGATCAGGTGGTATCTGCTGAGCATGGCAGCAACGAGAACATC
CACAGCATCCACTTCAGCGGCCACGTGTTCACCGTGAGGAAGAAGGAGGAGTACAAGATGCCCTG
TACAACCTGTACCCGGCGTGTGAGACCGTGAGGAGATGCTGCCAGCAAGGCCGATCTGGAGG
GTGGAGTGCCTGATCGGCAGCACCTGACGCCGGCATGAGCACCCCTGTTCTGGTGTACAGAAC
AAAGTGCCAGACCCCCCTGGGCATGGCCAGCGGCCACATCAGGACTTCCAGATCACCGCCTCTGGC
CACTACAGGCCAGTGGCCCCCAAGCTGGCAGGCTGCACTACAGCGGCAGCATCACGCCCTGGAGC
ACCAAGGAGCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCCCCATGATCATCCACGGCATCAAG
ACCCAGGGCGCCAGGCAGAACAGTTCAGCAGCCTGTACATCAGCAGTTCATCATGTACAGCCTG
GACGGCAAGAAGTGGCAGACCTACAGGGCAACAGCACCGGCACCCCTGATGGTGTCTTGGCAAC
GTGGACAGCAGCGGCATCAAGCACACATCTCAACCCCCCATCATGCCAGGTACATCAGGCTG
CACCCCCACCCACTACAGCATCAGGAGCACCCCTGGGATGGAACCTGATGGCTGCCACCTGAACAGC
TGCAGCATGCCCTGGGCATGGAGAGCAAGGCCATCTGACGCCAGATCACGCCAGCAGCTAC
TTCACCAACATGTTGCCACCTGGAGCCCCAGCAAGGCCAGGCTGCACCTGCAGGGCAGGAGAAC
GCCTGGAGGCCCAAGGTGAACAACCCAAGGAGTGGCTGCAGGTGGACTTCCAGAACGACATGAAG
GTGACCGCGTGACCACCCAGGGCGTGAAGAGCCTGCTGACCAGCATGTAAGTGAAGGAGTTCTG
ATCAGCAGCAGCCAGGACGCCACCAAGTGGACCCCTGTTCTCCAGAACGGCAAAGTGAAGGTGTC
CAGGGCAACCAGGACAGCTTCACCCCCGTGGTGAACAGCCTGGACCCCCCTGCTGACCAGGTAT
CTGAGGATCCACCCCCAGAGCTGGGTGCACCAGATGCCCTGAGAACGAGTGGCTGGATGCGAG
GCCCAAGGACCTGTACTGA (SEQ ID NO:101)

Figure 41B

CS23m1-FL-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCTGCTGAGGTTCTGCTTCTGCCACCAGG
AGATACTACCTGGGCGCCGTGGAGCTGAGCTGGGACTACATGCAGTCTGACCTGGCGAGCTGCCT
GTGGACGCCAGGTTCCCCCCCCAGAGTGCCAAGAGCTTCCCCTCAACACCTCAGTGGTGACAAG
AAGACCCCTGTTCTGGAGGTTACCGACCACCTGTTCAACATGCCAAGGCCAGGGCCCCCTGGATG
GGCCTGCTGGGCCCCACCATCCAGGGCAGGGTGTACGACACCCTGATCACCCCTGAAGAACATG
GCCAGCCACCCCGTGAGCTGCACGCCGTGGCGTAGCTACTGGAAAGGCCCTGAGGGGCCGAG
TATGACGACCAGCAGGCCAGAGGGAGAAGGAGGACGACAAGGTGTCCCCGGCGCAGCCACACC
TACGTGTGGCAGGTGCTGAAGGAGAACGCCACATGGCCAGCGACCCCTGTGCCTGACCTACAGC
TACCTGAGCCACCGTGGACCTGGTGAAGGACCTGAACCTGCTGATCGGCGCCCTGCTGGTGTGC
AGGGAGGGCAGCCTGGCCAAGGAGAAGACCCAGACCCCTGACAAGTTCATCTGCTGTTGCCGTG
TTCGATGAGGGCAAGAGCTGGCACAGCAGACCAAGAACAGCCTGATGCAAGGACAGGGATGCC
TCTGCCAGGGCCTGGCCAAGATGCACACCGTGAACGGCTACGTGAACAGGAGCCTGCCCGCTG
ATCGGCTGCCACAGGAAGTCTGTTACTGGCACGTGATGGCATGGCACCCAGGGAGGTGCAC
AGCATCTTCTGGAGGGCACACCTCCTGGTGAAGAACACAGGCCAGGCTGGAGATCAGC
CCCACATCACCTCCTGACGCCAGACCCCTGCTGATGGACTGGCCAGTTCTGCTGTCCTGCCAC
ATCAGCAGCCACCGACGGCATGGAGGCTACGTGAAGGTGGACAGCTGCCCGAGGGAGCCC
CAGCTGAGGATGAAGAACAAACGAGGAGGCCAGGGACTATGATGATGACCTGACCGACTCTGAGATG
GACGTGGTGAAGGTTGATGATGACAACAGCCCCAGCTTCATCCAGATCAGGTCTGTGGCCAAGAAC
CACCCCAAGACCTGGGTGCACTACATGCCGCGAGGGAGGACTGGGACTACGCCCGCTGGT
CTGGCCCCCGACGACAGGAGCTACAAGAGCCAGTACCTGAACAAACGCCAGGGAGGATCGGCA
AAAGTACAAGAAGGTCAAGATTGATGGCCATACCCGACGAGACCTTCAAGACCAGGGAGGCCATCCAG
CACGAGTCTGGCATCCTGGCCCCCTGCTGACGGCAGGGTGGCGACACCTGCTGATCATCTTC
AAGAACCAAGGCCAGCAGGCCCTACAACATCTACCCCGACGGCATCACCAGTGTGAGGCCCCCTGTAC
AGCAGGAGGCTGCCAAGGGCGTGAAGCACCTGAAGGACTTCCCTGCTGACCGACTCTGAGGATCTC
AAAGTACAAGTGGACCGTGAACGGTGGAGGATGGCCCCACCAAGTCTGACCCAGGTGCCTGACCA
TACTACAGCAGCTCGTGAACATGGAGAGGGACCTGGCTCTGGCTGATGGCCCCCTGCTGATC
TGCTACAAGGAGAGCGTGGACCGAGAGGGCAACAGATCATGTCAGACAGGAAACGTGATCTG
TTCTCTGTGTTCGATGAGAACAGGGAGCTGGTATCTGACCGAGAACATCCAGAGGTTCTGCCAAC
CCCAGGGCGTGCAGCTGGAGGACCCGAGTTCCAGGCCAGCAACATCATGCACAGCATCAACGGC
TACGTGTCAGCCCTGACGCTGCTGTGCTGCACGGGTGGCTACTGGTACATCTGAGC
ATCGGCGCCAGACCGACTTCCCTGCTGTTCTCTGGCTACACCTTCAAGCACAAGATGGT
TACGAGGACACCCCTGACCCCTGTTCCCTCAGCGGCAGAGACCGTGTCTGAGCATGGAGAACCCC
GGCCCTGTTGATCCTGGCTGCCACAACAGCGACTTCAGGAACAGGGCATGACGCCCTGCTGAAA
GTCAGCAGCTGCCACAAGAACACCGCGACTACTACGAGGACAGCTACGAGGACATCAGGCCCTAC
CTGCTGAGCAAGAACACGCCATCGAGCCAGGGAGCTTCAGCCAGAACCCCCCGTGTGAAGAGG
CACCAGAGGGAGATCACCAGGACCAACCTGCAAGAGCGACAGGAGGAGATCGACTATGATGACACC
ATCAGCGTGGAGATGAAGAACAGGAGGACTTCGACATCTACGACGAGGACGAGAACAGAGCCCCAGG
AGCTTCCAGAAGAACAGGCCACTACTTCATGCCGCCGTGGAGAGGGCTGTGGGACTATGGCATG
AGCAGCAGCCCCACGTGCTGAGGAACAGGGCCAGAGCGGCAGCGTGCCAGTTCAAGAACAGGTG

(Continued)

Figure 42A

GTGTTCCAGGAGTTACCGACGGCAGCTTCACCCAGCCCCGTACAGAGGCAGACTGAACGAGCAC
CTGGGCCTGCTGGCCCCCTACATCAGGGCCGAGGTGGAGGACAACATCATGGTGACCTTCAGGAAC
CAGGCCAGCAGGCCCTACAGCTTCTACAGCAGCCTGATCAGCTACGAGGAGGACCAGAGGCAGGGC
GCCGAGCCCAGGAAGAACTCGTGAAGCCAACGAGACCAAGACCTACTTCTGGAGGTCAGCAC
CACATGGCCCCACCAAGGACGAGTTCGACTGCAAGGCCTGGGCCTACTTCTCTGATGTGGACCTG
GAGAAGGACGTGCACAGGGCCTGATCGGCCCCCTGCTGGTGTGCCACACCAACACCCCTGAACCCC
GCCACGGCAGGCAGGTGACCGTGCAAGGAGTTGCCCTGTTCTTACCATCTCGACGAGACCAAG
AGCTGGTACTTCACCGAGAACATGGAGAGAACCTGCAGGGCCCCCTGCAACATCCAGATGGAGGAC
CCCACCTTCAAGGAGAACATCAGGTTCCACGCCATCAACGGCTACATCATGGACACCCCTGGCCGC
CTGGTGTGGCCAGGACCUAGAGGATCAGGTGGTATCTGCTGAGCATGGCAGCAACGAGAACATC
CACAGCATCCACTTCAGCGGCCACGTGTTCACCGTGAGGAAGAAGGAGGAGTACAAGATGCCCTG
TACAACCTGTACCCGGCGTGGTCAAGGACCGTGGAGATGCTGCCAGCAAGGCCGGCATCTGGAGG
GTGGAGTGCCTGATCGGCCAGCACCTGCACGCCGGCATGAGCACCCCTGTTCTGGTGTACAGCAAC
AAGTGCCAGACCCCCCTGGGCATGGCCAGCGGCCACATCAGGACTTCCAGATCACCGOCTCTGGC
CACTACGGCCAGTGGCCCCCAAGCTGGCCAGGCTGCACTACAGCGGCAGCATCACGCCCTGGAGC
ACCAAGGAGCCCCCTCAGCTGGATCAAGGTGGACCTGCTGGCCCCCATGATCATCCACGGCATCAAG
ACCCAGGGGCCAGGCAGAACAGTTCACCAAGCCTGACATCACGCCAGTTCATCATGTACAGCTG
GACGGCAAGAAGTGGCAGACCTACAGGGCAACAGCACCCGCACCGTGAATGGTGTCTCGGCAAC
GTGGACAGCAGCGGCATCAAGCACAACATCTCAACCCCCCATCATGCCAGGTACATCAGGCTG
CACCCACCCACTACAGCATCAGGAGCACCCCTCGGGATGGAACGTGATGGGCTGCGACCTGAACAGC
TGCAGCATGCCCTGGGCATGGAGAGCAAGGCCATCTCTGACGCCAGATCACGCCAGCAGCTAC
TTCACCAACATGTTGCCACCTGGAGCCCCAGCAAGGCCAGGCTGCACTGCAGGGCAGGAGCAAC
GCCTGGAGGCCCCAGGTGAACAACCCCAAGGAGTGGCTGAGGTGGACTTCCAGAAGACCATGAAG
GTGACCGGGCTGACCACCCAGGGCCTGAAGAGCCTGCTGACCGCATGTACGTGAAGGAGTCCCTG
ATCAGCAGCAGCCAGGACGGCCACCGAGTGGACCCCTGTTCTCCAGAACGGCAAAGTGAAGGTGTT
CAGGGCAACCAGGACAGCTTACCCCCGTGGTGAACAGCCTGGACCCCCCTGCTGACCAGGTAT
CTGAGGATCCACCCCCAGAGCTGGTGCACCAAGATGCCCTGAGAATGGAAGTGTGGATGCGAG
GCCAGGACCTGTACTGA (SEQ ID NO:102)

Figure 42B

CS23m23~FL~NA

ATGCAGATTGAGCTGAGCACCTGCTTCCTGTGCCCTGCTGAGGTTCTGCTCTGCCACCAGG
AGATACTACCTGGCGCCGTGGAGCTGAGCTGGGACTACATGCAGTCTGACCTGGCGAGCTGCCT
GTGGACGCCAGGTTCCCCCCCAGAGTGCCCAAGAGCTTCCCTAACACACTCAGTGGTGTACAAG
AAGACCCCTGTTCTGGAGTTCACCGACCACCTGTTAACATGCCAACGCCCAGGCCCCCTGGATG
GCCCTGCTGGGCCCCACCATCCAGGGCGAGGTGTACGACACCGTGGTGGTACCCCTGAAGAACATG
GCCAGCCACCCGTGAGCTGCACGCCGTGGCGTGGCTACTGGAAGTCTCTGAGGGCGCCGAG
TATGACGACCAGACCAGCAGAGGGAGAAGGAGGACGACAAGGTGTTCCCCGGCAAGAGCCACACC
TACGTGTGGCAGGTGCTGAAGGAGAACGGCCCCACTGCCAGCGACCCCCCTGCCCTGACCTACAGC
TACCTGAGCCACGTGGACCTGGTGAAGGACCTGAACACTTGGCCTGATCGGCCCTGCTGGTGTGC
AGGGAGGGCAGCTGGCCAAGGAGAACGACCCAGACCTGCAAAAGTTCATCCTGCTGTTGCCGTG
TTCGATGAGGGCAAGAGCTGGCACAGCGAGACCAAGAACAGCCTGATGCAGGACAGGGATGCCGCC
TCTGCCAGGGCTGGCCAAGATGCACACCGTGAACGGCTACGTGAACAGGAGCCTGCCCTGAGGTGCAC
ATCGGCTGCCACAGGAAGTCTGTGTACTGGCACGTGATCGGATGGCAGGAGCCTGCCCTGAGGTGCAC
AGCATCTTCTGGAGGGCACACCTTCCCTGGTGAAGGAAACCACAGGAGGCCAGCCTGGAGATCAGC
CCCATCACCTTCTGACCGCCCCAGACCCCTGCTGATGGACCTGGCAGTCCCTGCTGTTCTGCCAC
ATCAGCAGCCACCAGCACGACGGCATGGAGGCTACGTGAAGGTGGACAGCTGCCCGAGGAGCCC
CAGCTGAGGATGAAGAACACGAGGAGGCCGAGGACTATGATGATGACCTGACCGACTCTGAGATG
GACGTGGTGAGGTTGATGATGACAACAGCCCCAGCTCATCCAGATCAGGTCTGTTGCCAAGAAC
CACCCCAAGACTGGGTGCACTACATCGCCGCCAGGGAGGACTGGGACTACGCCCTGGTGC
CTGGCCCCCGACGACAGGAGCTACAAGAGCCAGTACCTGAAACAACGGCCCCAGAGGATCGGCAGG
AAAGTACAAGAAGGTCAAGATTGATGGCCTACACCGACGAGACCTCAAGACCAGGGAGGCCATCCAG
CACGACTCTGGCATCTGGCCCCCTGCTGTACGGCAGGTGGCAGACCCCTGCTGATCATCTTC
AAGAACCAAGGCCAGCAGGGCCCTACAACATCTACCCCCACGGCATCACCAGTGTGAGGCCCCCTGTAC
AGCAGGAGGCTGCCAAGGGCGTGAAGCACCTGAAGGACTTCCCCATCCTGCCGGAGATCTTC
AAAGTACAAGTGGACCGTGAACGGTGGAGGATGGCCCCACCAAGTCTGACCCAGGTGCCCTGACCAAG
TACTACAGCAGCTCGTGAACATGGAGAGGGACCTGGCCTCTGGCCTGATCGGCCCTGCTGATC
TGCTACAAGGAGAGCGTGGACCAGAGGGCAACAGATCATGTCGACAAGAGGAACGTGATCTG
TTCTCTGTGTTGATGAGAACAGGGACTGGTATCTGACCGAGAACATCCAGAGGTTCTGCCAAC
CCCCCGGCCAGCTGGAGGACCCGAGTTCCAGGCCAGCAACATCATGCACAGCATCAACGGC
TACGTGTTGACAGCCTGAGCTGTCGCTGTGTCGCTGCACGGTGGCTACTGGTACATCCTGAGC
ATCGGCGCCCAGACCGACTTCTGTCGTTCTCTGGCTACACCTCAAGCACAAGATGGTGC
TACGAGGACACCCCTGACCTGTTCCCCCTCAGCGCGAGACCGTGGTACATGAGCATGGAGAACCCC
GGCCTGTTGGATCTGGCTGCCACAAACAGCGACTTCAGGAACAGGGCATGACCGCCCTGCTGAAA
GTCAGCAGCTGCGACAAGAACACCGCGACTACTACGAGGACAGCTACGAGGACATCAGCGCCTAC
CTGCTGAGCAAGAACACACCACCTACGTGAACCGCTCCCTGAGCCAGAACCCCCCGTGTGAAAG
AGGCACCAAGAGGGAGATCACCAGGACCAACCTGAGGGAGGAGATCGACTATGATGAC
ACCATCAGCGTGGAGATGAAGAACAGGAGACTTCGACATCTACGACCGAGGACGAGAACACCAGGCC
AGGAGCTTCCAGAAGAACGACCAAGGACTACTTCATGCCCGCTGGAGAGGCTGTGGGACTATGGC
ATGAGCAGCAGCCCCACGTGCTGAGGAACAGGGCCCAGAGCGGCAGCGTGCCCTGAGTTCAAGAAC

(Continued)

Figure 43A

GTGGTGTTCAGGAGTTACCGACGGCAGCTCACCCAGCCCCGTACAGAGGCAGCTGAACGAG
CACCTGGGCCTGCTGGGCCCTACATCAGGGCCGAGGTGGAGGACAACATCATGGTGACCTTCAGG
AACCAAGGCCAGCAGGCCCTACAGCTTCTACAGCAGCCTGATCAGCTACGAGGAGGACCAGGGCAG
GGCGCCGAGCCCAGGAAGAACTTCGTGAAGGCCAACGAGACCAAGACCTACTCTGGAAAGGTGCAG
CACCAACATGGCCCCCACCAAGGACCGAGTTCGACTGCAAGGCCTGGCCTACTCTCATGATGTGGAC
CTGGAGAAGGACGTGCACAGCGGCCTGATCGGCCCTGCTGGTGTGCCACACCAACACCCCTGAAC
CCCGCCCACGGCAGGCAGGTGACCGTGCAGGAGTTGCCCTGTTCTTCACCATCTCGACGAGACC
AAGAGCTGGTACTTCACCCAGAAACATGGAGAGGAACCTGCAGGGCCCTGCAACATCCAGATGGAG
GACCCCCACCTTCAAGGAGAACTACAGGTTCACGCCATCAACGGCTACATCATGGACACCCCTGCC
GGCCTGGTGTGGCCAGGACCAAGAGGATCAGGTGGTATCTGCTGAGCATGGCAGCAACGAGAAC
ATCCACAGCATCCACTTCAGCGGCCACGTGTTCACCGTGAGGAAGAAGGAGGAGTACAAGATGCC
CTGTACAACCTGTACCCCGCGTGGTACGAGACCGTGGAGATGCTGCCAGCAAGGCCGATCTGG
AGGGTGGAGTGCCTGATCGCGAGCACCTGCACGCCGATGAGCACCCCTGTTCTGGTGTACAGC
AACAAAGTGCCAGACCCCCCTGGGCATGGCCAGCGGCCACATCAGGGACTTCCAGATCACCGCCT
GGCCAGTACGCCAGTGGGCCAGCTGGGCCAGGCTGCACTACAGCGGCAGCATTCAACGCCCTGG
AGCACCAAGGAGCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCCCCATGATCATCCACGGCATT
AAGACCCAGGGGCCAGGCAGAAAGTTCAGCAGCCTGTACATCAGCCAGTTCATCATGTACAGC
CTGGACGGCAAGAAGTGGCAGACCTACAGGGCAACAGCACCGGCACCCCTGATGGTGTCTCGGC
AACGTGGACAGCAGCGGCATCAAGCACAACATCTCAACCCCCCATCATGCCAGGTACATCAGG
CTGCACCCCACCCACTACAGCATCAGGAGCACCCCTGCGATGGAACATGATGGGCTGCCACCTGAAC
AGCTGCAGCATGCCCTGGGCATGGAGAGCAAGGCCATCTCTGACGCCAGATCACCGCCAGCAGC
TACTTCACCAACATGTTGCCACCTGGAGCCCCAGCAAGGCCAGGCTGCACCTGCAGGGCAGGAGC
AACCCCTGGAGGCCAGGTGAACAAACCCCAGGACTGGCTGCCAGGTGACTCCAGAAAGACCATG
AAGGTGACCGGGCTGACCAACCCAGGGCGTGAAGAGCCTGCTGACCAGCATGTACGTGAAGGAGTTC
CTGATCAGCAGCAGCCAGGACGCCACCGACTGGACCCCTGTTCTCAGAACGGCAAAGTGAAGGTG
TTCCAGGGCAACCAGGACAGCAGCTTCACCCCGTGGTGAACAGCCTGACCCCCCTGCTGACCAGG
TATCTGAGGATCCACCCCCAGAGCTGGGTGCACCAGATGCCCTGAGAATGGAAGTGCTGGATGCC
GAGGCCAGGACCTGTACTGA (SEQ ID NO:103)

Figure 43B

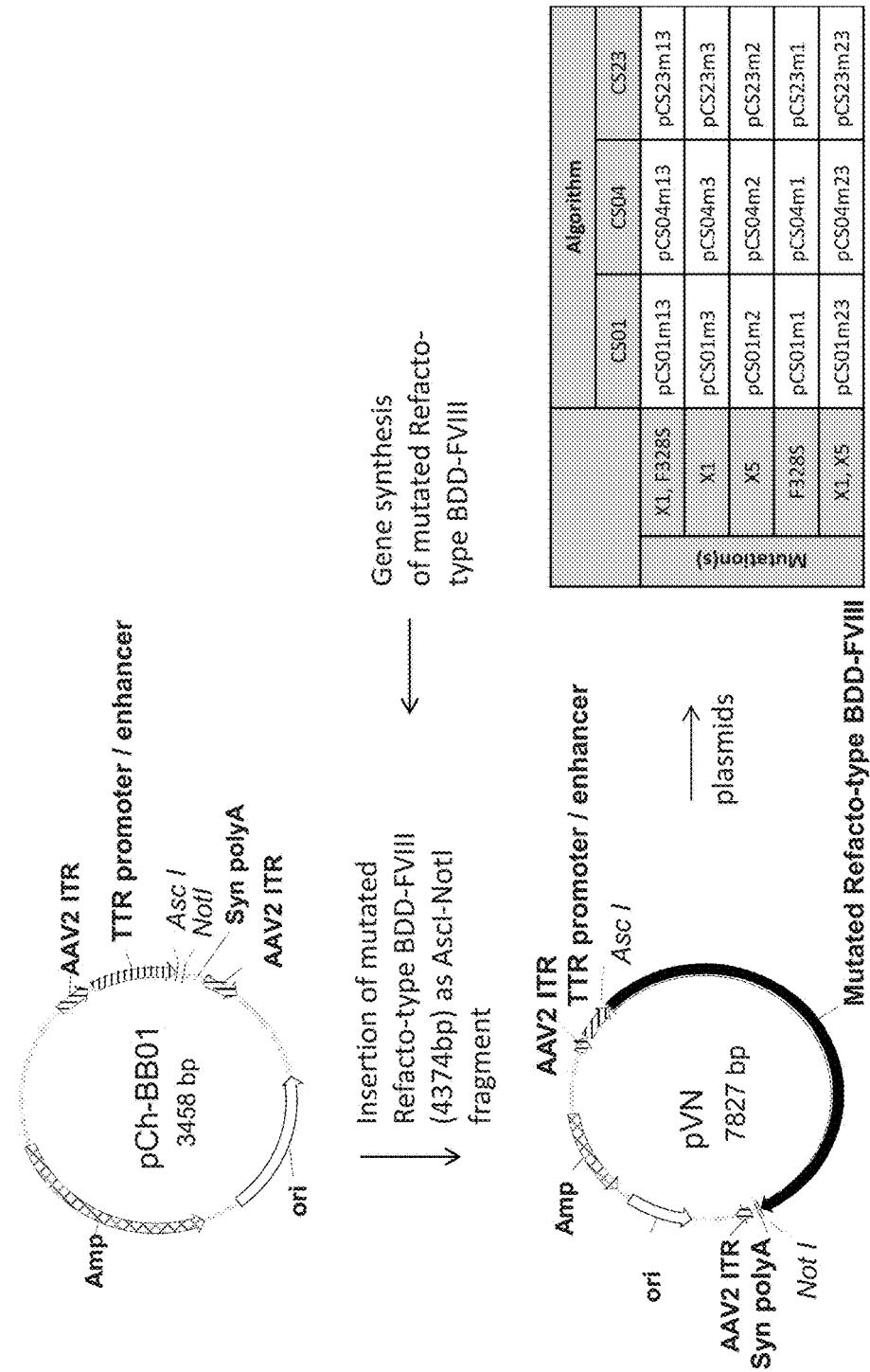


Figure 44

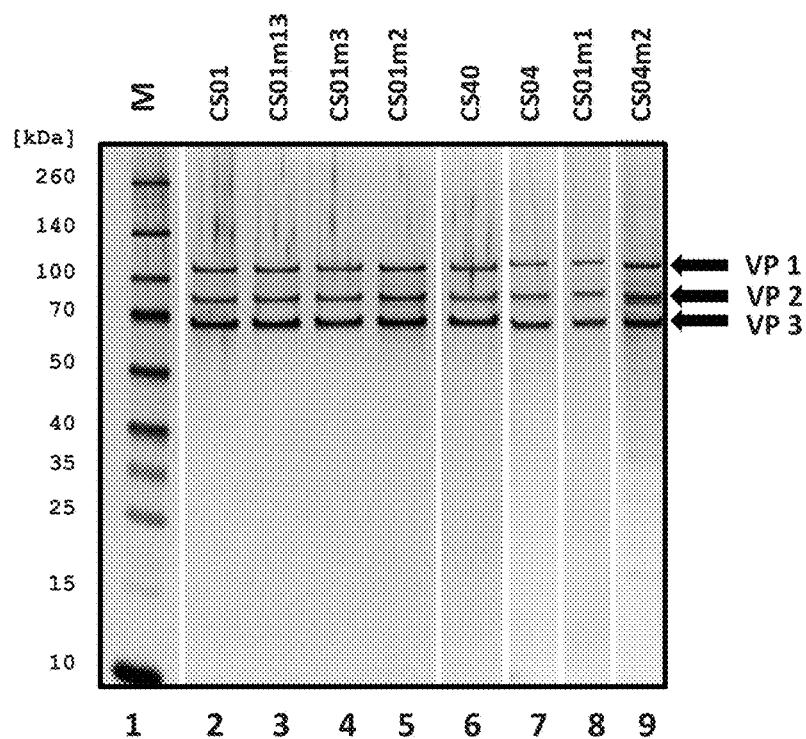


Figure 45

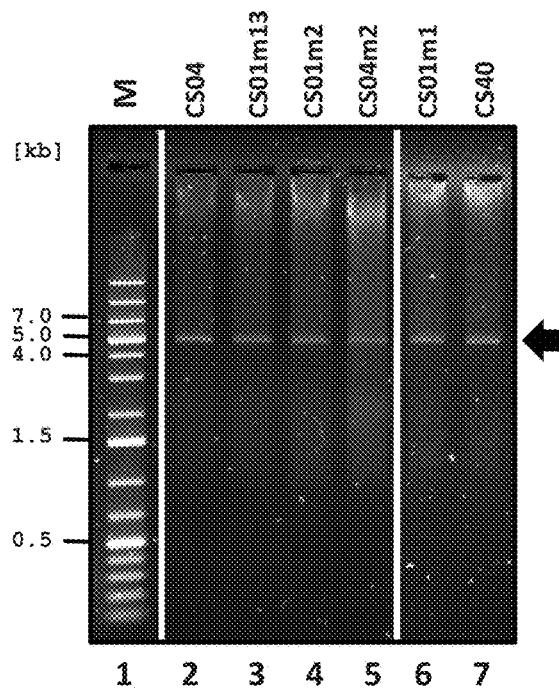


Figure 46

CS01-HC-NA

gcc
accaggagat actaacctggg ggctgtggaa ctttcttggg actacatgca gtcgtacccgc
ggagagctgc ctgtggatgc caggttccca cccagagtgc ccaagtccctt cccattcaac
acctctgtgg tctacaagaa gacactctt gtggattca ctgaccacot gttcaacatt
gcaaaaaccca gaccaccctg gatgggactc ctgggacccca ccattcaggc tgagggttat
gacactgtgg tcatcacccct caagaacatg gcacccacc ctgtgtctct gcatgtgt
ggagtctcat actggaaaggc ctctgaaggg gctgagtgatg atgaccagac atcccagaga
gagaaaagagg atgacaagg gttccctggg gatatctaca cctatgtgt gcaagtccctc
aaggagaatg gaccatggc atctgaccca ctctgcctga catactctta cctttctcat
gtggacctgg tcaaggacccct caactctggc ctgattgggg cactgctggt gtgcaggggaa
ggatccctgg ccaaggagaa aacccagaca ctgcacaagg tcattctct gtttgcgtc
tttgcgttgg gcaagtcctg gcactctgaa acaaaagaact ccctgatgca agacaggat
gctgcctctg ccagggcatg gcccaagatg cacactgtga atggctatgt gaacagatca
ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcatgtgat tggcatgggg
acaaccctg aagtgcactc catttcctg gagggacaca ccttccctggt caggaaccac
agacaaggct ctctggagat ctctcccatc accttccctca ctgcacagac actgctgtat
gacccctggac agttccctgct gttctgcccac atctcttccc accagcatga tggcatggaa
gcctatgtca aggtggactc atgcccctgag gaaccacagc tcaggatgaa gaacaatgag
gaggctgagg actatgtga tgacctgact gactctgaga tggatgtgtt cagattgtat
gatgacaact ctccatctt cattcagatc aggtctgtgg caaaagaaaaca ccccaagaca
tgggtgcact acattgtgc tgaggaagag gactggact atgcaccact ggtcctggcc
cctgatgaca ggagctacaa gtctcagatc ctcaacaatg gcccacaaag aattgaaaga
aagtacaaga aagtcaaggatt catggcctac actgtatgaaa ccttcaagac aagagaagcc
attcagcatg agtctggcat tctggacca ctccctgtatg gggaaagtggg agacaccctg
ctcatcatct tcaagaacca ggcttccagg ccctacaaca totacccaca tggcatcact
gatgtcaggc ccctgtacag caggagactg ccaaaaagggg tggaaacacccat caaggacttc
ccattctgc ctggagagat ctcaagtac aagtggactg tcactgtgg ggtggacca
acaaaagtctg accccaggtg cctcaccaga tactactt cttttgtgaa catggagaga
gacctggcat ctggactgtat tggaccactg ctcatctgat acaaggagtc tgtggaccag
agaggcaacc agatcatgtc tgacaaagaga aatgtgatcc tggatctctgt ctttgtatgag
aacagatcat ggtacctgac tgagaacatt cagagattcc tgcccaaccc tgctgggtg
caactggaaag accctgagtt ccaggcaagc aacatcatgc actccatcaa tggctatgt
tttgcgttgc tccagcttc tgctgtctg catgaggtgg cctactggta cattcttct
atggggcac aaactgactt ctttctgtc ttcttctctg gatacacctt caagcacaag
atggatgtatg aggacaccct gacactcttc ccattctctg gggaaaactgt gttcatgagc
atggagaacc ctggactgtg gattctggaa tgccacaact ctgacttcag aaacagggga
atgactgcac tgctcaaagt ctccctctgt gacaagaaca ctggggacta ctatgaggac
tcttatgagg acatctctgc ctacctgtc agcaagaaca atgcccattga gcccaga
(SEQ ID NO:24X)

Figure 47

CS01-LC-NA

g agatcaccag gacaaccctc
cagtctgacc aggaagagat tgactatgtat gacaccattt ctgtggagat gaagaaggag
gactttgaca tcttatgtat ggacgagaac cagtctccaa gatcattcca gaagaagaca
agacactact tcattgttc tggtggaaaga ctgtgggact atggcatgtc ttccctctcc
catgtcctca ggaacagggc acagtctggc tctgtgccac agttcaagaa agtggcttc
caggagttca ctgatggctc attcacccag cccctgtaca gaggggaact gaatgagcac
ctgggactcc tgggaccata catcagggtc gaggtggaag acaacatcat ggtgacattc
agaaaaccagg cctccaggcc ctacagcttc tactcttccc tcatcagcta tgaggaagac
cagagacaag gggctgagcc aagaaaagaac tttgtgaaac ccaatgaaac caagacctac
ttctgtgaaag tccagcacca catggcaccc accaaggatg agtttgactg caaggcctgg
gcataacttct ctgatgtgga cctggagaaaa gatgtcact ctggctgtat tggccccactc
ctggctgtcc acaccaaacac cctgaaccct gcacatggaa gcacagtgcac tgtgcaggag
tttgcctct tottcaccat otttgcataa accaagtcat gtaacttcac tgagaacatg
gagagaaaact gcaagagcacc atgcaacattt cagatggaag accccaccc caaggagaac
tacaggttcc atgccccatcaa tggctacatc atggacaccc tgcctggct tgcataatggca
caggaccaga gaatcagatg gtacactgtt tctatggat ccaatgagaa cattcactcc
atccacttct ctgggcatgt cttaactgtt agaaaagaagg aggaataaaaaa gatggccctg
tacaacccctt accctggggtt otttgcataa gtttgcataa gtttgcataa agctggcatac
tggaggggtgg aatgcctcat tggggagcac ctgcataatggat gtttgcataa cctgttccctg
gtctacagca acaagtgcac gacaccctt ggaatggctt ctggccacat caggacttc
cagatcacttgc cctctggccat gtatggccat tggccaccc aactggccat gctccactac
tctggctcca tcaatgcata gtcacccatc gagccatttctt cttggatcaa ggtggacctg
ctggccacccat tggatcattca tggatcataa acacaggggg caagacagaaa atttccctt
ctgtacatct cacagttcat catcatgtac tctctggatg gcaagaatgtt gcaagatata
agagggcaactt ccactggcac cctcatggtc tttttggca atgtggacag ctctggcatac
aagcacaaca tcttcacccat tccatcattt gcaatgatata tcaatgtca ccccccaccc
tactcaatca gatcaaccctt caggatggaa ctgtatggat gtcacccat gtcataatgg
atgcccctgg gaatggagat caaggccattt tctgtatggcc agatcactgc atcccttttac
ttcaccatca tggatggccat ctggatcataa tcaatgtca ggtggacccat ccaggaaaga
agcaatgcctt ggagacccat ggtcaacaa ccaaaaggat ggtggacccat ggacttccat
aagacaatgtt aagtcaactgg ggtgacccat caggatggca agtctctgtt cacccatca
tatgtgaagg agttccatgtt ctcttcctca caggatggcc accagtggac actctttttt
cagaatggca aagtcaaggtt gttccaggcc aaccaggactt ctttcacccat tggatggca
tcaactggacc ccccccctt gacaagatata ctgagaattt accccatggcc tgggtccac
cagatggccc tgagaatggaa agtccctggaa tgtgaggcac aagacccatgtt c
(SEQ ID NO:25)

Figure 48

CS01Δ(760-1667) - CS01-SC1-NA

ATGCAGATTGAGCTGTCCACCTGCTTCTTCTGTGCCCTGCTGAGAATTCTGCTTCTGCCACCAGGGAGATACTACCTGGGGGCTGTGGAACTTTCTTGGGACTACATGCAGTCTGACCTGGGAGAGCTGCCCTGTGGATGCCAGGTTCCUACCCAGAGTGCCAACGACTCCTTCCCATTCAACACCTCTGTGGTCTACAAGAAGACACTCTTGTGGAA TTCACTGACCAACCTGTTCAACATTGCAAACACCCAGACACCACCCCTGGATGGGACTCCTGGGACCCACCATTCAGGCTGAGGTGTATGACACTGTGGTCACTCACCCCTCAAGAACATGGCATCCCACCCCTGTGTCTCTGCATGCTGTGGAGACTCTCATACTGGAAAGGCTCTGAAGGGGCTGAGTATGATGACCAAGACATCCCAGAGAGAGAAAAGAGGATGACAAGGGTGTCCCTGGGGATCTCACACCTATGTGTGGCAAGTCCTCAAGGAGAAATGGACCCATGGCATCTGACCCACTCTGCCCTGACATACTCCTACCTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGGCACTGCTGGTGTGCAGGGAAAGGATCCCTGGCAAGGGAGAAAACCCAGACACTGCACAAAGTTCATTCTCCTGTTGCTGTCTGGCAAGGAGACAGGGATGCTGCCCTGCCAGGGCATGGCCAAGATGCAACACTGTGAATGGCTATGTGAACAGATCACTGCCCTGGACTCATTGGGCTGCCACAGGGAAATCTGTCTACTGGCATGTGATTGGCATGGGACAAACCCCTGAAGTGCACCTCATTTCCTGGGAGGGACACACCTTCCTGGTCAGGAACCCACAGACAAAGCCCTCTGGAGATCTCTCCCACATCACCTTCCTCACTGCACAGACACTGCTGATGGACCTTGGACAGTCTGGACAGTCTGCTGTCTGCCACATCTCTCCCACAGGGCATGATGGCATGGAAAGCCTATGTCAAGGGACTCATGCCCTGAGGAACCCAGCTCAGGATGAAAGAACATGAGGAGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATGGATGTGGTCAAGATTTGATGATGACAACCTCTCCATTCATGTTGCTGCTGGCAAGAAAACACCCCCAACAGACATGGGTGCACTACATTGCTGCTGAGGAAGGGACTGGGACTATGCAACCACTGGTCTGGCCCTGTGATGACAGGGAGCTACAAGTCTCAGTACCTCAACAAATGGGCCAACAAAAGAATTGGAAGAAAAGTACAAGAAAAGTCAGATTGATGGCTACACTGATGAAACACCTTCAAGACAAGAGAACCGCAATTGAGCATGACTCTGGCATTCTGGGACCACTCCCTGTATGGGGAAAGTGGGAGACACCCUTGCTCATCTTCAAGAACCCAGGGCTCCAGGCCUCTACAACATCTACCCACATGGCATCACTGATGTCAGGCCUCTGTAAGCAGCAGGGAGACTGCCAACAGGGGTGAAACACCTCAAGGACTTCCCACATCTGCCCTGGAGAGATCTTCAGTACAAGTGGACTGTCACTGTGGAGGGATGGACCAACAAAGCTGACCCCCAGGTGCCCTCACAGATACTACCTCCTTGTGACATGGGACAGGACTGGTCTGTCTGGTCTTGTGAGAACAGATCATGGTACCTGACTGAGAACATTGAGAGATTGAGGACTTCTGCCAACCCCTGCTGGGATCTGACCTCTGCTGACATTGAGAACAGACCTGAGTCCAGGGAAGCAACATCATGCACTCCATCAATGGCTATGTGTTGACTCTCTCCAGCTTCTGTCAGCTGGCATGAGGTGGCTACTGGTACATTCTTCTATTGGGGCACAAACTGACTTCTTCTGTCTTCTTCCTGGGATACACCTTCAAGCACAAGATGGTGTATGAGGACACCCCTGACACTCTCCCATTCTCTGGGGAAACTGTGTCATGAGCATGGGAGAACCCCTGGACTGTGGATTCTGGGATGCCACAACCTCTGACTTCAAGAAAACAGGGGAATGACTGCACTGCTCAAAGTCTCTCTGTGACAAGAACACTGGGACTACTATGAGGACTCTATGAGGACATCTCTGCTAACCTGCTCAGCAAGAACAAATGCCATTGAGGCCAGAGAGATCACCAGGACAACCCCTCCAGTCTGACCCAGGAAAGAGATTGACTATGATGACACCATTCTGTGGAGATGAAGAACGGAGACTTGTGACATCTGATGAGAACAGGAGACTCTCCAGAAGAACAGACACTACTTCAATTGCTGCTGTGGAAAGAAGACATCATGGTGAATTCAAGAACAGGACCTCCAGGCTCCAGGCTCACAGTCTGGCTCTGTGCCACAGTTCAAGAAAAGTGGCTTCCAGGGAGTTCACTGATGGCTCATTCACTGCCAGGCCCCCTGTACAGAGGGGAACCTGAATGAGCACCTGGGACCTCTGGGACCACATCAGGGCTGAGGTGGAAGAACATCATGGTGAATTCAAGAACAGGCTCCAGGCTCACAGTCTGGCTCTGTGCCACAGTTCAAGAAAAGAACTTTGTGAAACCCAATGAAACCAAGAACCTACTTCTGGAAAGTCCAGCACCACTGGCACCC

(Continued)

Figure 49A

ACCAAGGATGAGTTGACTGCAAGGCCTGGGCATACTTCTCTGATGTGGACCTGGAGAAAGATGTGCACTCT
GGCCTGATTGGCCCACTCCTGGTCTGCCAACACCAACACCCCTGAACCCCTGCACATGGAAGGCAAGTGACTGTG
CAGGAGTTGCCCTCTTCAACCACCTTGATGAAACCAAGTCATGGTACITCACTGAGAACATGGAGAGA
AACTGCAGAGCACCATGCAACATTCAGATGGAAGACCCCACCTCAAGGAGAACTACAGGTTCCATGCCATC
AATGGCTACATCATGGACACCCCTGCTGGGCTTGTCTGGCACAGGACAGAGAACATGAGATGGTACCTGCTT
TCTATGGGATCCAATGAGAACATTCACTCCATCCACTTCTCTGGGATGTCCTCACTGTGAGAAAGAAGGAG
GAATAACAAGATGCCCTGTAACACCTTACCCCTGGGCTTTGAGACTGTGAGAACATGCCCTCCAAGCT
GGCATCTGGGAGGTGGAATGCCCTCATGGGAGCACCTGCATGCTGGCATGTCACCCCTGTCCTGGTCTAC
AGCAACAAGTGCAGACACACCCCTGGGAATGCCCTCTGGCACATCAGGGACTTCCAGATCACTGCCCTGGC
CAGTATGCCAGTGGGACCCAAACTGGGCAAGGCTTCACTACTCTGGCCATCAATGCAATGGTCAACCAAG
GAGCCATTCCTCTTGATCAAGGTGGACCTGCTGGCACCCATGATCAATTCAAGACACACGGGRC
AGACAGAAAATTCTCTCTTGATCATCTCACAGTTCACTCATCATGATCTACTCFCTGGATGGCAAGAAGTGGCAG
ACATACAGAGGCAACTTCACCTGGCACCCCTCATGGTCTTGGCAATGTGGACASCTCTGGCATCAAGCAC
AACATCTTCAACCCCTOCATCATTGCCAGATAACATCAGGCTGCACCCCACTACTCAATCAGATAACC
CTCAGGATGGARCTGATGGGATGTGACCTGAACTCCTGCTCAATGCCCTGGAAATGGAGAGCAAGGCCATT
TCTGATGCCAGATCACTGCACTCCTCTTACTTCACCAACATGTTGCCACCTGGTCAACCATCAAAGCCAGG
CTGCACCTCCAGGGAAAGAACATGCCCTGGAGAGACCCCAGGTCAACAAACCCAAAGGAATGGCTGCAAGTGGAC
TTCCAGAAGACAATGAAAGTCACTGGGGTGACAACCCAGGGGGTCAAGTCTGCTCACCTCAATGTATGTG
AAGGAGTTCTGATCTCTTCAACAGGATGGCCACCCAGTGGACACTCTTCTCCAGAAATGGCAAAAGTCAAG
GTGTTCCAGGGCAACCCAGGACTCTTCACACCTGTGGTGAACCTCACTGGACCCCCCCCCCTCTGACAAGATA
CTGAGAATTCAACCCCCAGTCTGGGTCCACAGATTGCCCTGAGAACATGGAAGTCCTGGGATGTGAGGCACAA
GACCTGTACTGA (SEQ ID NO:26)

Figure 49B

CS01A(772-1667) - CS01-SC2-NA

ATGCAGATTGAGCTGTCACCTGCCTTCTGTGCCCTGCTGAGATCTGCTCTGCCACCAGG
AGATACTACCTGGGGCTGTGGAACCTTCTTGGACTACATGCAGTCTGACCTGGAGAGCTGCCT
GTGGATGCCAGGTTCCACCCAGAGTGCCAAGTCCTCCCATTCAACACCTCTGGTACAAG
AAGACACTCTTGGAATTCACTGACCACCTGTTAACATTGCAAAACCCAGACCAACCCCTGGATG
GGACTCCTGGGACCCACCATTCAAGGCTGAGGTGTATGACACTGTGGTCATCACCCCAAGAACATG
GCATCCCACCCGTGTCTGATGCTGTGGAGTCTCATCTGGAAAGCCTCTGAAGGGCTGAG
TATGATGACCAGACATCCCAGAGAGAAAAGAGGTGACAAGGTGTCCCTGGGGATCTCACACC
TATGTGTGGCAAGTCCTCAAGGAGAATGGACCCATGGCATCTGACCCACTCTGCCCTGACATACTCC
TACCTTCTCATGTGGACCTGGTCAAGGAACCTCAACTCTGGACTGATTGGGGCACTGCTGGTGTG
AGGAAGGATCCCTGGCCAAGGAGAAAACCCAGACACTGCACAAGTTCAATTCTCTGGTGTG
TTGATGAGGGCAAGTCTTGGCACTCTGAAACAAAGAACTCCCTGATGCAAGACAGGGATGCTGCC
TCTGCCAGGGCATGGCCAAGATGACACTGTGAATGGCTATGTGAACAGATCACTGCCCTGGACTC
ATTGGCTGCCACAGGAATCTGTCTACTGGCATCTGATTGGCATGGGACAACCCCTGAAGTGCAC
TCCATTTCCTGGAGGGACACACCTCCTGGTCAGGAACCAACAGACAAGCCTCTGGAGATCTCT
CCCATCACCTCCTCACTGCACAGACACTGCTGATGGACCTGGACAGTCCTGCTGTTCTGCCAC
ATCTCTCCCACAGCATGGCATGGAAGCCTATGTCAAGGTGGACTCATGCCCTGAGGAACCA
CAGCTCAGGATGAAGAACATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTGAGATG
GATGGTCAGATTGATGATGACAACACTCTCCATCCTCATTGAGATCAGGTCTGTGGCAAAGAAA
CACCCCAAGACATGGGTGCACTACATTGCTGAGGAAGAGGACTGGGACTATGCAACCAGGGTC
CTGGCCCCCTGATGACAGGGACTACAAGTCTCAGTACCTCAACAATGGCCCAACAAAGAACATGG
AAAGTACAAGAAAGTCAGATTCACTGGCTACACTGATGAAACCTTCAAGACAAGAGAACGCCATT
CATGAGTCTGGCATTCTGGGACACTCCTGTATGGGAAGTGGGAGACACCCCTGCTCATCATCTC
AAGAACCAAGGCTCCAGGCCCTACAACATCTACCCACATGGCATCAGTGATGTCAGGCCCCCTGTAC
AGCAGGAGACTGCCAAAGGGGTGAAACACCTCAAGGACTTCCCCATTCTGCCCTGGAGAGATCTC
AAAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACAAAGTCTGACCCAGGTGCCCTACCC
TACTACTCTCTTGTGAACATGGAGAGAGACCTGGCATCTGGACTGATTGGACCACTGTCATC
TGCTACAAGGAGTCTGTGGACCAAGAGAGGCAACCAGATCATGTCTGACAAGAGAAATGTGATT
TTCTCTGTCTTGTGAGAACAGATCATGGTACCTGACTGAGAACATTCAAGAGATTCTGCCAAC
CCTGCTGGGGTGCAACTGGAAGACCCCTGAGTTCCAGGCAAGCAACATCATGCACTCCATCAAT
TATGTGTTGACTCTCTCCAGCTTCTGTCTGCCCTGCATGAGGTGGCTACTGGTACATTCTTCT
ATTGGGGCACAAACTGACTTCCTTCTGTCTTCTCTGGATACACCTTCAGAACAGCATGGT
TATGAGGACACCCCTGACACTCTTCCCATTCTCTGGGAAACTGTGTTCATGAGCATGGAGAACCC
GGACTGTGGATTCTGGATGCCACAACACTCTGACTTCAGAAACAGGGGAATGACTGCACTGCTCAA
GTCTCCTCTGTGACAAGAACACTGGGACTACTATGAGGACTCTTATGAGGACATCTCTGCC
CTGCTCAGCAAGAACATGCCATTGAGCCAGAAGCTTCAGAATTCCAGACACCCAGCACC
AGGGAGATCACCAGGACAACCCCTCCAGTCTGACCAGGAAGAGATTGACTATGATGACACCATT
GTGGAGATGAAGAACAGGAGGACTTGTGACATCTATGATGAGGACGAGAACCGAGTCTCCAAGATCATT

(Continued)

Figure 50A

CAGAAGAACAGACACTTCAATTGCTGCTGGAAAGACTGTGGACTATGGCATGTCTTCC
TCTCCCCATGTCCTCAGGAACAGGGCACAGTCTGGCTCTGTGCCACAGTTCAAGAAAAGTGGCTTC
CAGGAGTTCACTGATGGCTCATTCAACCCAGCCCCGTACAGAGGGAACTGAATGACCACCTGGGA
CTCCTGGGACCATACATCAGGGCTGAGGTGGAAGACAACATCATGGTACATTGAGAAACCCAGGCC
TCCAGGCCCTACAGCTCTACTCTTCCCTCATCAGCTATGAGGAAGACCAGAGACAAGGGCTGAG
CCAAGAAAGAACCTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAGCACCACATG
GCACCCACCAAGGATGAGTTGACTGCAAGGCCCTGGCATACTTCTGATGTGGACCTGGAGAAA
GATGTGCACTCTGGCCTGATTGGCCACTCTCTGGTCTGCCACACCAACACCCCTGAACCCCTGCACAT
GGAAGGCAAGTGAATGTGCAGGGAGTTGCCCTTCTTCAACCACATTGATGAAACCAAGTCATGG
TACTCACTGAGAACATGGAGAGAACTGCAAGGACCATGCAACATTGAGATGGAAGACCCACC
TTCAAGGAGAACTACAGGTTCCATGCCATCAATGGCTACATCATGGACACCCCTGCCCTGGCTTGTC
ATGGCACAGGACCAGAGAACATCAGATGGTACCTGCTTCTATGGGATCCAATGAGAACATTCACTCC
ATCCACTTCTCTGGGCATGTCTCACTGTGAGAAAGAAGGAGGAATAACAGATGGCCCTGTACAAC
CTCTACCCCTGGGTCTTGAGACTGTGGAGATGCTGCCCTCAAAGCTGGCATCTGGAGGGTGGAA
TGCCTCATTGGGGAGCACCTGCATGCTGGCATGTCAACCCCTGGTCTACAGCAACAAAGTGC
CAGACACCCCTGGGAATGCCCTCTGCCACATCAGGGACTTCCAGATCACTGCCCTGGCCAGTAT
GGCCAGTGGCACCCAAACTGCCAGGGCTCCACTACTCTGGCTCCATCAATGCAATGGTCAACCAAG
GAGCCATTCTCTGGATCAAGGTGGACCTGCTGGCACCCATGATCATGGCATCAAGACACAG
GGGGCAAGACAGAAATTCTCCTCTGTACATCTCACAGTTCACTCATGTAATCTCTGGATGGC
AAGAAGTGGCAGACATAACAGAGGCAACTCCACTGGCACCCATGGCTTCTGGCAATGTGGAC
AGCTCTGGCATCAAGCACAAACATCTCAACCCCTCCATCATTGCCAGATACTCAGGCTGCACCC
ACCCACTACTCAATCAGATCAACCCCTCAGGATGGAACCTGATGGGATGTGACCTGAACCTGCTCA
ATGCCCTGGGAATGGAGAGCAAGGCCATTCTGATGCCAGATCACTGCATCCTCTTACCTCACC
AACATGTTGCCACCTGGTCACCATAAAAGCCAGGCTGCACCTCCAGGGAAAGAACATGCCTGG
AGACCCCAGGTCAACAACCAAAGGAATGGCTGCAAGTGACTTCCAGAACAGAACATGAAAGTCAT
GGGGTGAACAACCCAGGGGTCAAGTCTCTGCTCACCTCAATGATGTAAGGAGTTCCCTGATCTCT
TCCTCACAGGATGGCCACCACTGGACACTCTTCTCCAGAACATGGCAAAGTCAAGGTGTTCCAGGGC
AACCAAGGACTCTTCACACCTGTGGTGAACCTGAGACCCCCCCCCCTGACAAGATACTGAGA
ATTCAACCCCCAGTCTGGGTCCACCAAGATTGCCCTGAGAACATGGAAGTCCTGGGATGTGAGGCACAA
GACCTGTACTGA (SEQ ID NO:27)

Figure 50B

CS23A(760-1667) - CS23-SC1-NA

(Continued)

Figure 51A

CCCAGGAAGAACCTCGTGAAGCCCAACGAGACCAAGACCTACTTGTGGAAAGGTGCAGCACCATGGCCCC
ACCAAGGACGAGTCGACTGCAAGGCCTGGCCTACTTCTCTGATGTGGACCTGGAGAAGGACGTCCACAGC
GGCTGATCGGCCCTGCTGGTGTGCCACACCAACACCCCTGAACCCGCCACGGCAGGCAGGTGACCGTG
CAGGAGTTCGCCCTGTTCTCACCATCTCGACGAGACCAAGAGCTGGTACTTCACCGAGAACATGGAGAGG
AACTGCAGGGCCCCCTGCAACATCCAGATGGAGGACCCACCTCAAGGAGAACATACAGGTCCACGCCATC
AACGGCTACATCATGGACACCTGCCCGCCTGGTGTGGCCACGGACCAGAGGATCAGGTGGTATCTGCTG
AGCATGGCAGCAACGAGAACATCCACAGCATCCACTTCAGCGCCACGTGTTCACCGTGAGGAAGAAGGAG
GAGTACAAGATGGCCCTGTAACACCTGTACCCCCGGCGTGTGAGACCGTGGAGATGCTGCCAGCAAGGCC
GGCATTGAGGGTGGAGTGCCTGATCGCGAGSCACCTGCACGCCGATGAGCACCCGTGTTCTGGTGTAC
AGCAACAAGTGCCAGACCCCCCTGGCATGGCCAGGGCACATCAGGGACTTCCAGATCACCGCCTCTGGC
CACTACGGCCAGTGGGCCAGCTACAGCGGCAGCATCAACGCCCTGGAGCACCAAG
GAGCCCTTCAGCTGGATCAAGGTGGACCTGCTGGCCCCATGATCATCCACGGCATCAAGACCCAGGGCGCC
AGGCAGAAGTTCAGCAGCCTGTACATCAGCCAGTTCATCATGTACAGCCTGGACGCAAGAAGTGGCAG
ACCTACAGGGCAACAGCACCGCACCCCTGATGGTGTCTCGCAACGTTGACAGCAGCGGCATCAAGCAC
AACATCTCAACCCCCCATCATGCCAGGTACATCAGGCTGCACCCCACTACAGCATCAGGAGCAC
CTGCGATGGAATGATGGCTGCGACCTGAAACAGCTGCAGCATGCCCTGGCATGGAGAGCAAGGCCATC
TCTGACGCCAGATCACGCCAGCAGCTACTTCACCAAACATGTTGCCACCTGGAGGCCAGCAAGGCCAG
CTGCACCTGCAAGGGCAGGAGCAACGCCCTGGAGGCCAGGTGAACAACCCCCAAGGAGTGGCTGCAGGTGGAC
TTCCAGAAGACCATGAAGGTGACCGGGGTGACCACCCAGGGCGTGAAGAGCCTGCTGACCAGCATGTACGTG
AAGGAGTTCTGATCAGCAGCAGCCAGGACGGCCACCAAGTGGACCCCTGTTCTCCAGAACGGCAAAGTGAAG
GTGTTCCAGGGCAACCAGGACAGCTTCACCCCCGTGGTGAACAGCCTGGACCCCCCTGCTGACCAGGTAT
CTGAGGATCCACCCCCAGAGCTGGGTGACCAAGATGCCCTGAGAATGAGAAGTGTGGATGCCAGGGCAG
GACCTGTACTGA (SEQ ID NO:28)

Figure 51B

CS23Δ (772-1667) ~ CS23-SC2-NA

ATGCAGATTGAGCTGAGCACCTGCTTCTTCCTGTGCCGTGAGGTTCTGCTTCTGCCACCAGG
AGATACTACCTGGCGCCGTGGAGCTGAGCTGGACTACATGCAGTCTGACCTGGCGAGCTGCCT
GTGGACGCCAGGTTCCCCCCCAGAGTCCCCAAGAGCTCCCCITCAACACCTCAGTGGTACAAG
AAGACCTGTTCTGGAGTTACCGACCACCTGTTAACATGCCAACCCCAGGGCCCCCTGGATG
GGCCTGCTGGGCCCCACCATCCAGGCCAGGTTACGACACCCGTGGTATCACCTGAAGAACATG
GCCAGCCACCCCGTGAGCTGCACGCCGTGGGCGTAGCTACTGGAAGGCCTCTGAGGGGCCAG
TATGACGACCAGACCAGCCAGAGGGAGAAGGAGGACGACAAGGTGTTCCCCGGGGCAGCACACC
TACGTGTGGCAGGTGCTGAAGGAGAACGGCCCCATGCCAGCGACCCCTGTGCCCTGACCTACAGC
TACCTGAGCCACGTGGACCTGGTGAAGGACCTGAACTCTGGCCTGATCGGCCCTGCTGGTGTG
AGGGAGGGCAGCCTGGCCAAGGAGAACGACCCAGACCTGACAAGTTACATCCTGCTGTTGCCGTG
TTCGATGAGGGCAAGAGCTGGCACAGCAGACCAAGAACAGCCTGATGCAGGACAGGGATGCC
TCTGCCAGGGCTGGCCAAGATGCACACCGTGAACGGCTACGTGAACAGGAGCTGCCGGCTG
ATCGGCTGCCACAGGAAGTCTGTGTACTGGCACGTGATCGGATGGCACCCGGAGGTGCA
AGCATCTCCTGGAGGGCACACCTTCTGGTAGGAACACAGGCAGGCCAGCCTGGAGATCAGC
CCCACATCACCTTCTGACGCCAGACCCCTGCTGATGGACCTGGCCAGTTCTGCTGTTGCCAC
ATCAGCAGCCACCAGCACGCCATGGAGGCCTACGTGAAGGTGGACAGCTGCCCGAGGAGCCC
CAGCTGAGGATGAAGAACAAACGAGGAGGCCAGGACTATGATGATGACCTGACCGACTCTGAGATG
GACGTGGTGGAGGTTGATGATGACAACAGCCCCAGCTTCATCCAGATCAGGTCTGAGGAGATG
CACCCCAAGACCTGGGTGACTACATGCCGCCAGGAGGACTGGGACTACGCCCTGGT
CTGGCCCCCGACGACAGGAGCTACAAGAGCCAGTACCTGAACAAACGCCAGAGGATGCCAGG
AAGTACAAGAACGGTCAGATTACATGCCCTACACCGACGAGACCTTCAAGACCAGGGAGGCCATCCAG
CACGAGTCTGGCATCCTGGCCCCCTGCTGTACGGCAGGTGGCGACACCCCTGCTGATCATCTTC
AAGAACCCAGGCCAGCAGGCCCTACACATCTACCCCCACGGCATCACCAGTGTGAGGCCCTGTAC
AGCAGGAGGCTGCCAACGGCGTAAGCACCTGAGGACTTCCCCATCCTGCCGGAGATCTTC
AAAGTACAAGTGGACCGTGACCGTGGAGGATGGCCCCACCAAGTCTGACCCAGGTGCTGACCGAG
TACTACAGCAGCTTGTGAAACATGGAGAGGGACCTGGCCTCTGGCCTGATGGCCCCCTGCTGATC
TGCTACAAGGAGAGCGTGAGGACAGAGGAGCTGGTATCTGACCGAGAACATCCAGAGGTTCTGCCAAC
TTCTCTGTGTTGATGAGAACAGGAGCTGGTATCTGACCGAGAACATCCAGAGGTTCTGCCAAC
CCGCCGGCGTGCAGCTGGAGGACCCGAGTTCCAGGCCAGCAACATCATGCACAGCATCACGGC
TACGTGTTGACAGCCTGCAGCTGTGTCCTGCACGAGGTGGCCTACTGGTACATCCTGAGC
ATCGGCCAGACCGACTTCTGTGTCCTCTGGCTACACCTCAAGCACAAAGATGGT
TACGAGGACACCCCTGACCCCTGTTCCCCCTCAGGCCAGAACAGGGCATGACGCCCTGCTGAAA
GTCAGCAGCTGGACAAGAACACCGGGCAGTACTACGAGGACAGCTACGAGGACATCAGGCCCTAC
CTGCTGAGCAAGAACACGCCATCGAGCCCAGGAGCTTCAGCCAGAACATCCAGACACCCAGCACC

(Continued)

Figure 52A

AGGGAGATCACCAGGACCACCTGCAGAGCGACCAGGAGGAGATCGACTATGATGACACCACATCAGC
GTGGAGATGAAGAAGGGAGCTTCGACATCTACGACGAGGACGAGAACCCAGAGCCCCAGGGAGCTTC
CAGAAGAAGACCAAGGCACACTTACATCGCCGCCGTGGAGAGGCTGTGGACTATGGCATGAGCAGC
AGCCCCCACGTGCTGAGGAACAGGGCCCAGAGCGGCAGCGTGCCTCAGTTCAAGAAGGTGGTGTTC
CAGGAGTTACCGACGGCAGCTTCACCCAGCCCTGTACAGAGGGCAGCTGAACGAGCACCTGGGC
CTGCTGGGCCCTACATCAGGGCCGAGGTGGAGGAACACATCATGGTGACCTTCAGGAACCAGGCC
AGCAGGCCCTACAGCTTACAGCAGCCTGATCAGCTACGGAGGACCAGAGGCAGGGCGCCGAG
CCCAGGAAGAACTTCGTGAAGCCCAACGAGACCAAGACCTACTTCTGGAGGGCAGCACCACATG
GCCCCACCAAGGACGAGTTGACTGCAAGGCCTGGGCCTACTTCTGTATGTGGACCTGGAGAAG
GACGTGCACAGCGGCCTGATGGCCCCCTGCTGGTGTGCACACCAACACCCCTGAACCCCGCCAC
GGCAGGCAGGTGACCGTGAGGAGTTCGCCCCCTGTTCTCACCATCTCGACGAGACCAAGAGCTGG
TACTTCACCGAGAACATGGAGAGGAACCTGCAGGGCCCCCTGCAACATCCAGATGGAGGACCCACC
TTCAAGGAGAACTACAGGTTCCACGCCATCAACGGCTACATCATGGACACCCCTGCCGGCTGGTG
ATGGCCCAGGACCAGAGGATCAGGTGGTATCTGCTGAGCATGGCAGCAACGAGAACATCCACAGC
ATCCACTTCAGCGGCCACGTGTTACCGTGAGGAAGAAGGGAGTACAAGATGGCCCTGTACAAC
CTGTACCCCGCGTGTTCAGACCGTGGAGATGCTGCCAGCAAGGCCGGCATCTGGAGGGTGGAG
TGCCTGATCGCGAGCACCTGCACGCCGGCATGAGCACCCCTGTTCTGGTGTACAGCAACAAGTGC
CAGACCCCCCTGGCATGCCAGCGGCCACATCAGGGACTTCCAGATCACCGCCCTGGCCAGTAC
GGCCAGTGGCCCCAAGCTGCCAGGCTGCACTACAGCGGCAGCATCAACGCCCTGGAGCACCAAG
GAGCCTTCAGCTGGATCAAGGTGGACCTGCTGGCCCCATGATCATCCACGGCATCAAGACCCAG
GGCGCCAGGCAGAAGTTCAGCAGCCTGTACATCAGCCAGTTCATCATGTACAGCCTGGACGGC
AAGAAGTGGCAGACCTACAGGGCAACAGCACCGCACCCGTATGGTGTCTTCGGCAACGTGGAC
AGCAGCGGCATCAAGCACACATCTCAACCCCCCATCATGCCAGGTACATCAGGCTGCACCC
ACCCACTACAGCATCAGGAGCACCTGCGGATGGAACGTATGGCTGCGACCTGAACAGCTGCAGC
ATGCCCTGGCATGGAGAGCAAGGCCATCTGACGCCAGATCACGCCAGCAGCTACTTCACC
AACATGTTGCCACCTGGAGCCCCAGCAAGGCAGGCTGCACCTGCAGGGCAGGAGCAACGCCCTGG
AGGCCCCAGGTGAACAACCCCCAAGGAGTGGCTGCAAGGTGGACTTCCAGAAGACCATGAAGGTGACC
GGCGTGACCACCCAGGGCGTGAAGAGCCTGCTGACCAGCATGTACGTGAAGGGAGTTCTGATCAGC
AGCAGCCAGGACGGCCACCAAGTGGACCCCTGTTCTCCAGAACGGCAAAGTGAAGGTGTTCCAGGGC
AACCAGGACAGCTTCACCCCCGTGGTGAACAGCCTGGACCCCCCTGCTGACCAGGTATCTGAGG
ATCCACCCCCAGAGCTGGGTGACCAAGATGCCCTGAGAATGGAAGTGTGGATGCGAGGCCAG
GACCTGTACTGA (SEQ ID NO:29)

Figure 52B

CS01m23-FL-AA (SEQ ID NO: 104)

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVNDARFPVPRVPKSFPFN
TSVVYKKTLFVEFTDHLFNIAKPRPPWMGLLGPTIQAEVYDTVVVTLKNMASHPVSLHAV
GVSYWKSSEGAEYDDQTSQREKEDDKVFFGKSHTYVWQVLKENGPtasDPPCLTYSYLSH
VDLVKDLNSGLIGALLVCREGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRD
AASARAWPKMHTVNGYVNRSLPGLIGCHRKSvyWHVIGMGTPEVHSIFLEGHTFLVRNH
RQASLEISPITFLTAQTLMDLGQFLLFCISSHQHDGMEAYVKVDSCPEEPQLRMKNNE
EAEDYDDDLDSEMDVVRFDDDNSPSFTQIRSVAKHPKTWVHYIAAEEEWDWYAPLVLA
PDDRSYKSQYLNNGPQRIGRKYKKVRFMAYTDETFTREAIQHESGILGPLLYGEVGDTL
LIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVTVEDGP
TKSDPRCLTRYSSFVNMERDLASGLIGPLLICYKESVDQRGNQIMSDKRNVILFSVFDE
NRSWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSLQLSVCLHEVAYWYILS
IGAQTDLFSVFFSGYTFKHMKVYEDTTLFPFSGETVFMSENPGWLILGCHNSDFRNRG
MTALLKVSSCDKNTGDYYEDSYEDISAYLLSKNNTTVNRSLSQNPPVLRHQREITRTT
LQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSS
PHVLRNRAQSGSVPQFKVVFQEFTDGSFQTQPLYRGELNEHLGLGPYIRAEVEDNIMVT
FRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHHMAPTKDEFDCKA
WAYFSDVDLEKDVSGLIGPLLVCHTNTLNPQAHGRQVTQEFALFFTIFDETWSWYFTEN
MERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIH
SIHFGHVFTVRKKEEYKMALYNLYPGVFETVEMLP SKAGIWRVECLIGEHLHAGMSTLF
LVYSNKCQTPLGMASGHIRDFQITASGQYQWAPKLARLHYSGSINAWS TKEPF SWIKVD
LLAPMI IHGIKTQGARQKFSSLIYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSG
IKHNIFNPPIIARYIIRLHPTHYSIRSTLREMELMGCDLNCSMPLGMESKAISDAQITASS
YFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTQGVKSLLTS
MYVKEFLISSQDGHQWTLFFQNGKVKVFOQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWV
HQIALRMEVLGCEAQDLY

Figure 53

CS04m3-FL-AA (SEQ ID NO: 105)

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDARFPPRVPKSFPFN
TSVVYKKTLFVEFTDHNFNIAKPRPPWMGLGPTIQAEVYDTVVITLKNMASHPVSLHAV
GVSYWKASEGAEYDDQTSQLREKEEDKVFPAGSHTYVWQVLKENGPMASDPLCLTYSYLSH
VDLVKDLNSGLIGALLVCREGSLAKEKTQLHKFILLFAVFDEGKSWHSETKNSLMQDRD
AASARAWPKMHTVNGYVNRSLPGLIGCHRKSVDWHVIGMGTPEVHSIFLEGHTFLVRNH
RQASLEISPITFLTAQTLIMDLQFLLFCCHISSHQHDGMEAYVKVDSCPEEPQILRMKNNE
EAEDYDDDLDSEMDVVRFDNNPSFIQIRSVAKHPKTWVHYIAAEEEEDWDYAPLVLA
PDDRSYKSQYLNNGPQRIGRKVKVRFMAYTDETFKTRAEIQHESGILGPLLYGEVGDTL
LIIFKNQASREPYNIYPHGIDTVRPLYSRRLPKGVKHLKDFPILPGEIFKYKWTVTVEDGP
TKSDPRCLTRYSSFVNMERDLASGLIGPLLICYKESVDQRGNQIMSDKRNVILFSVFDE
NRSWYLTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDSQLSVCLHEVAYWYILS
IGAQTDFLSVFFSGYTFKHKMVYEDTLTLPFFSGETVFMNSMENPGLWILOGCHNSDFRNNG
MTALLKVSSCDKNTGDYYEDSYEDISAYLLSKNNNTYVNRSLSQNPVLRHQREITRTT
LQSDQEEIDYDDTISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAAVERLWDYGMSSS
PHVLRNRAQSGSVPQFKVVFQEFTDGSFTQPLYRGELNEHLGLLGPYIRAEVEDNIMVT
FRNQASRPYSFYSSLISYEEDQRQGAEPRKNFVKPNETKTYFWKVQHHMAPTKDEFDCKA
WAYFSDVDLEKDVKHSGLIGPLLVCHTNLPAHGRQVTVQEFALEFFTIFDETKSWYFTEN
MERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPGLVMAQDQRIRWYLLSMGSNENIH
SIHFSGHVFTVRKKEEYKMALYNLYPGVFTETVEMPLSKAGIWRVECLIGEHLHAGMSTLF
LVYSNKCQTPLGMASGHIRDQITASGQYQWAPKLARLHYSGSINAWSTKEPFWSIKVD
LLAPMIIHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGNVDSSG
IKHNIFNPPIIARYIRLHPHTHYSIRSTLRMELMGCDLNCSMPLGMESKAISDAQITASS
YFTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTQGVKSLLTS
MYVKEFLISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWV
HQIALRMEVLGCEAQDLY

Figure 54

CS01-FL-AAm12 (SEQ ID NO: 106)

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSWDYMQSDLGELPVDFPPRVPKSFPNTSVVYK
KTLFVEFTDHFLNIAKPRPPWMGLLGPТИQAEVYDTVVVTLNMAHPVSLHAVGVSYWK3SEGAE
YDDQTSQREKEDDKVFPGKSHTYVWQVLKENGPASDPPCLTYSYLSHVDLVKDLNSGLIGALLVC
REGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKMHTVNGYVNRSLPGL
IGCHRKSVYWHVIGMGTTPEVHSIFLEIGHTFLVRNHRQASLEISPITFLTAQTLMDLGQFLLSCH
ISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVFDDDNPSFIQIRSVAKK
HPKTWVHYIAAEEEDWDYAPLVLAPDDRSYKSQLNNPQRIGRKYKKVRFMAYTDETFKTREAIQ
HESGILGPLLYGEVGDTLLIIFKNQASRPYNIPHGTIDVRPLYSRRLPKGVKHLKDFPILPGEIF
KYKWTVTVEDGPTKSDPRCCLTRYSSFVNMERDLASGLIGPLLICYKESVDQRCGNQIMSDKRNVIL
FSVFDENRSWYLTTENIQRFLPNPAGVQLEDPEFQASNIMHSINGYVFDQLQLSVCLHEVAYWYILS
IGAQTDLSVFFSGYTFKHKMVYEDTLTLFPFSGETVFMSMENPGLWILGCHNSDFRNRGMTALK
VSSCDKNTGDYYEDSYEDISAYLLSKNNIAEPRFSFSQNPPVLRHQREITRTTLQSDQEEIDYDDT
ISVEMKKEDFDIYDEDENQSPRSFQKKTRHYFIAVERLDYGMSSSPHVLRNRAQSGSVPQFKKV
VFQEFTDGSFTQPLYRGELNEHLGLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQG
AEPRKNFVKPNETKTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLEKDVKHSGLIGPLLVCHTNLP
AHGRQVTQEFALFFTIFDETWSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPG
LVMAQDQRIRWYLLSMGSNENIHSHFGHVFTRKKEEYKMALYNLYPGVFETVEMLPSKAGIWR
VECLIGEHLHAGMSTLFVYNSKCQTPLGMASGHIRDQITASGQYQWAPKLARLHYSGSINAWS
TKEPFISWIKV DLLAPMIIHGIKTQGARQKFSSLYISQFIMYSLDGGKWQTYRGNSTGTLMVFFGN
VDSSGIKHNIFNPIIIARYIRLHPTHYSIRSTLREMELMGCDLNCSMPLGMESKAISDAQITASSY
FTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTGVTTOGVKSLLTSMYVKEFL
ISSSQDGHQWTLEFFQNGKVKVFQGNQDSFTPVNVSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCE
AQDLY

Figure 55

CS04-FL-AAm12 (SEQ ID NO: 107)

MQIELSTCFFLCLLRFCFSATRRYYLGAVELSDYMQSDLGELPNDARFPPRVPKSFPNTSVVYK
KTLFVEFTDHLFNIAKPRPPWMGLLGPLTIQAEVYDTVVVTLKNMASHPVSLHAVGVSYWKSSEGAE
YDDQTSQREKEEDKVFPGKSHTYVWQVLKENGPtasDPPCLTYSYLSHVDLVKDLNSGLIGALLVC
REGSLAKEKTQTLHKFILLFAVFDEGKSWHSETKNSLMQDRDAASARAWPKMHVNGYVNRSLPGL
IGCHRKSVYWHVIGMGTTPEVHSIFLEHTFLVRNHRQASLEISPIFLTAQTLMDLGQFLLSCH
ISSHQHDGMEAYVKVDSCPEEPQLRMKNNEEAEDYDDDLTDSEMDVVRFDDDNSPSFIQIRSVAKK
HPKTWVHYIAAEEEEDWDYAPLVLAPDDRSYKSQYLNNNGPQRIGRKYKKVRFMAYTDETFKTREAIQ
HESGILGPLLYGEVGDTLLIIFKNQASRPYNIYPHGITDVRPLYSRRLPKGVKHLKDFPILPGEIF
KYKWTVTVEDGPTKSDPRCLTRYSSFVNMERDLASGLIGPLLICYKESVDQRGNQIMSDKRNVIL
FSVFDENRSWYLTENIQRFILPNPAGVQLEDPEFQASNIMHSINGYVFDSDLQSVCLHEVAYWYILS
IGAQTDPLSVFSGYTFKHKMVYEDTLTLFPFSGETVFMSENPGWILGCHNSDFRNRGMTALLX
VSSCDKNTGDYYEDSYEDISAYLLSKNNATEPRSFQNPPVLRHQREITRTTLQSDQEEIDYDDT
ISVEMKKEDFDIYDDEDENQSFRSFQKRTRHYFIAVERLWDYGMSSSPHVRNRAQSGSVQFKV
VFQEFTDGTSFTQPLYRGELENEHLGLGPYIRAEVEDNIMVTFRNQASRPYSFYSSLISYEEDQRQG
AEPRKNFVKPNETKTYFWKVQHHMAPTKDEFDCKAWAYFSDVDLEKDVKHSGLIGPLLVCNTLNP
AHGRQTVQEFAFFTIFTDETKSWYFTENMERNCRAPCNIQMEDPTFKENYRFHAINGYIMDTLPG
LVMAQDQRIRWYLLSMGSNENIHSIHFGHVFTRKKEEYKMALYNLYPGVFETVEMLP SKAGIWR
VECLIGERLHAGMSTLFLVYSNKCQTPLGMASGHIRDFOITASGQYQQWAPKLARLHYSGSINAWS
TKEPF SWIKV DLLAPMI IHGIKTQGARQKFSSLYISQFIIMYSLDGKKWQTYRGNSTGTLMVFFGN
VDSSGIKHNIFNPPTIARYIRLHPTHYSIRSTLRMELMGCDLNCSMPLGMESKAISDAQITASSY
FTNMFATWSPSKARLHLQGRSNAWRPQVNNPKEWLQVDFQKTMKVTVGTTQGVKSLLTSMYVKEFL
ISSSQDGHQWTLFFQNGKVKVFQGNQDSFTPVVNSLDPPLLTRYLRIHPQSWVHQIALRMEVLGCE
AQDLY

Figure 56

CS01-FL-NAm12 (SEQ ID NO: 108)

ATGCAGATTGAGCTGTCCACCTGCTTCTTCTGTGCCTGCTGAGATTCTGCTCTGCCACCAGG
AGATACTACCTGGGGCTGTGGAACCTTCTTGGGACTACATPGCAGTCTGACCTGGAGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGAGTCCCCAAGTCCTTCCATTCAACACCTCTGTGGCTACAAG
AAGACACTCTTGTGGAATTCACTGACCACCTGTTCAACATGCAAAACCCAGACCACCCCTGGATG
GGACTCCTGGGACCCACCATTCAAGGCTGAGGTGTATGACACTGTGGCTGTCACCCCTCAAGAACATG
GCATCCCACCCCTGTGTCTGCATGCTGTGGAGTCTCATACTGGAAATCCTCTGAAGGGCTGAG
TATGATGACCAGACATCCCAGAGAGAGAAAGAGGTGACAAGGTGTTCCCTGGGAAGTCTCACACC
TATGTGTGGCAAGTCCCTCAAGGAGAATGGACCCACTGCATCTGACCCACCCCTGCCGTACATACTCC
TACCTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGGACTGCTGGTGTGC
AGGGAAAGGATCCCTGGCCAAGGAGAAAACCCAGACACTGCACAAGTTCAATTCTCCTGTTGCTGTC
TTTGATGAGGGCAAGTCTGGCACTCTGAAACAAAGAACTCCCTGATGCAAGACAGGGATGCTGCC
TCTGCCAGGGCATGGCCAAGATGCACACTGTGAATGGCTATGTGAACAGATCACTGCCTGGACTC
ATTGGCTGCCACAGGAATCTGTCTACTGGCATGTGATTGGCATGGGACAACCCCTGAAGTGCAC
TCCATTTCCTGGAGGGACACACCTCCTGGTCAGGAACCACAGACAAGCCTCTGGAGATCTCT
CCCATCACCTCCTCACTGCACAGACACTGCTGATGGACCTTGGACAGTTCTGCTGTGCCAC
ATCTCTTCCCACCAAGCATGGCATGGAAAGCCTATGTCAAGGTGGACTCATGCCCTGAGGAACCA
CAGCTCAGGATGAAGAACAAATGAGGAGGCTGAGGACTATGATGATGACCTGACTCTGAGATG
GATGTGGTCAGATTGATGATGACAACCTCCTCATTCAGATCAGGTCTGTGGCAAAGAAA
CACCCCAAGACATGGGTGCACTACATTGCTGCTGAGGAAGAGGACTGGGACTATGCACCACTGGTC
CTGGCCCTGATGACAGGAGCTACAAGTCTCAGTACCTCAACAATGCCCAACAAAGAACATTGGAAGA
AACTACAAGAAAGTCAGATTGATGACACTCTGATGAAACCTTCAAGACAAGAGAACCCATTCA
CATGAGTCTGGCATTCTGGGACCACTCCTGTATGGGGAAAGTGGGAGACACCCCTGCTCATCATCTC
AAGAACCAAGGCCCTCAGGCCCTACAACATCTACCCACATGGCATCAGTGTGAGGCCCTGTAC
AGCAGGAGACTGCCAAAAGGGGTGAAACACCTCAAGGACTTCCCCATTCTGCCCTGGAGAGATCTC
AAAGTACAAGTGACTGTCACTGTGGAGGATGGACCAACAAAGTCTGACCCAGGTGCCTCACCAGA
TACTACTCCTCTTTGTGAACATGGAGAGAGACCTGGCATCTGGACTGATTGGACCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGAGGCAACCAGATCATGTCTGACAAGAGAAATGTGATTCTG
TTCTCTGTCTTGATGAGAACAGATCATGGTACCTGACTGAGAACATTCAAGAGATTCTGCCAAC
CCTGCTGGGTGCAACTGGAAGACCCCTGAGTTCCAGGCAAGCAACATCATGCACTCCATCAATGGC
TATGTGTTGACTCTCTCAGCTTCTGTCTGCCATGAGGTGGCTACTGGTACATTCTTCT
ATTGGGGCACAAACTGACTCCTTCTGTCTTCTCTGGATACACCTTCAAGCACAAGATGGTG
TATGAGGACACCCCTGACACTCTTCCATTCTGGGAAACTGTGTTCATGAGCATGGAGAACCCCT
GGACTGTGGATTCTGGATGCCACAACACTGACTTCAGAAACAGGGGAATGACTGCACTGCTCAA
GTCTCCTCTGTGACAAGAACACTGGGACTACTATGAGGACTCTTATGAGGACATCTCTGCCCTAC
CTGCTCAGCAAGAACAAATGCCATTGAGGCCAGAAGGCTCTCAGAACATCCACCTGCTGTAAAGAGA
CACCAAGAGAGGATCACCAGGACAACCCCTCCAGTCTGACCCAGGAAGAGGATTGACTATGATGACACC
ATTTCTGTGGAGATGAAGAAGGAGGACTTTGACATCTATGATGAGGAGCAGAACCAAGTCTCCAAGA
TCATTCCAGAAGAACAGACACTACTTCATTGCTGCTGGAAAGACTGTGGGACTATGGCATG
TCTTCCCTCTCCCCATGTCCCTCAGGAACAGGGCACAGTCTGGCTCTGTGCCACAGTTCAAGAACAGTG

(Continued)

Figure 57A

GTCTTCCAGGAGTTCACTGATGGCTCATTCAACCCAGCCCCGTACAGAGGGAACTGAATGAGCAC
CTGGGACTCCTGGGACCATACTCAGGGCTGAGGTGGAAGACAACATCATGGTACATTAGAAAC
CAGGCCTCCAGGCCCTACAGCTCTACTCTTCCCTCATCAGCTATGAGGAAGACAGAGACAAGGG
GCTGAGCCAAGAAAGAACCTTGTGAAACCCAATGAAACCAAGACCTACTTCTGGAAAGTCCAGCAC
CACATGGCACCCACCAAGGATGAGTTGACTGCAAGGCCGGCATACTTCTGTATGTGGACCTG
GAGAAAGATGTGCACTCTGGCCTGATTGGCCACTCCTGGTCTGCCACACCAACACCCCTGAACCC
GCACATGGAAGGCAAGTGAATGTGAGGAGTTGCCCTTCTTCACCATCTTGATGAAACCAAG
TCATGGTACTTCACTGAGAACATGGAGAGAACACTGCAAGACCATGCAACATTGAGATGGAAGAC
CCCACCTTCAAGGAGAACTACAGGTTCCATGCCATCAATGGCTACATCATGGACACCCCTGCCCTGG
CTTGTCTGGCACAGGACAGAGAACATCAGATGGTACCTGCTTCTATGGGATCCAATGAGAACATT
CACTCCATCCACTTCTCTGGGATGTCTTCACTGTGAGAAAAGAAGGAGGAATAACAGATGCCCTG
TACAACCTCTACCCCTGGGTCTTGAGACTGTGGAGATGCTGCCCTCCAAAGCTGGCATCTGGAGG
GTGGAATGCCCTCATGGGAGCACCTGCATGTCACCTGTCACACCTGTTCTGGTCTACAGCAAC
AAGTGCCAGACACCCCTGGGAATGCCCTCTGCCACATCAGGGACTTCCAGATCACTGCCCTGGC
CAGTATGCCAGTGGCACCCAAACTGCCAGGCTCCACTACTCTGGCTCCATCAATGCAATGGTCA
ACCAAGGAGCCATTCTCTGGATCAAGGTGGACCTGCTGGCACCCATGATCATTGCAAG
ACACAGGGGCAAGACAGAAATTCTCTCTGTACATCTCACAGTCATCATGTACTCTCTG
GATGGCAAGAAGTGGCAGACATACAGAGGCAACTCCACTGGCACCCATGGTCTTCTGGCAAT
GTGGACAGCTGGCATCAAGCACACATCTCAACCCCTCCATGTCAGATACTCAGGCTG
CACCCACCCACTACTCAATCAGATCAACCCCTCAGGATGGAACGTGATGGGATGTGACCTGAACTCC
TGCTCAATGCCCTGGGAATGGAGAGCAAGGCCATTCTGATGCCAGATCACTGCATCCCTTAC
TTCACCAACATGTTGCCACCTGGTCACCATCAAAGCCAGGCTGCACCTCCAGGAAGAACATGAAA
GCCTGGAGACCCAGGTCAACACCCAAAGGAATGGCTGCAAGTGGACTTCCAGAAGACAAATGAAA
GTCACGGGGTACAACCCAGGGGTCAAGTCTGCTCACCTCAATGTATGTAAGGAGTTCTG
ATCTCTCCTCACAGGATGGCACCCAGTGGACACTCTTCTCCAGAATGGCAAAGTCAGGTGTT
CAGGGCAACCAGGACTCTTCACACCTGTGGTGAACTCACTGGACCCCCCCCCTCCTGACAAGATA
CTGAGAATTCAACCCCAAGTCTGGGTCACAGATTGCCCTGAGAATGGAAGTCTGGGATGTGAG
GCACAAGACCTGTACTGA

Figure 57B

CS04-FL-NAm12 (SEQ ID NO: 109)

ATGCAGATTGAGCTGAGCACCTGCTCTTCCCTGTGCCCTGCTGAGGTTCTGCTTCTCTGCCACCAAGG
AGATACTACCTGGGGCTGTGGAGCCTTCTGGGACTACATGCAGTCTGACCTGGGGAGCTGCCT
GTGGATGCCAGGTTCCCACCCAGACTGCCAAATCCTCCATTCAACACACCTCTGTGGTCTACAAG
AAGACCCCTTTGTGGAGTCACTGACCACCTGTTCAACATTGCCAAACCCAGGCCACCCCTGGATG
GGACTCCTGGGACCCACCATTCAAGGCTGAGGTTATGACACTGTGGTCGTACCCCTAAGAACATG
GCCCTCCACCCGTGAGCCTGCATGCTGTGGGGTCAGCTACTGGAAGTCTCTGAGGGGGCTGAG
TATGATGACCAGACCTCCCAGAGGGAGAAGGGAGATGACAAAGTGTCCCTGGGAAGAGGCCACACC
TATGTGTGGCAGGTCTCAAGGAGAATGGCCCCACTGCCTCTGACCCACCCCTGCCGTACCTACTCC
TACCTTCTCATGTGGACCTGGTCAAGGACCTCAACTCTGGACTGATTGGGGCCCTGCTGGTGTGC
AGGGAGGGCTCCCTGGCCAAGAGAGAACCCAGACCCCTGACAAGTCATTCTCCTGTTGCTGTC
TTTGATGAGGGCAAGAGCTGGCACTCTGAAACCAAGAACCTCCGTGACAGGGATGCTGCC
TCTGCCAGGGCCTGGCCAAGATGCACACTGTGAATGGCTATGTGAACAGGAGCCTGCCCTGGACTC
ATTGGCTGCCACAGGAAATCTGTCACTGGCATGTGATTGGCATGGGACAACCCCTGAGGTGCAC
TCCATTTCTGGAGGGCCACACCTCCTGGTCAGGAACCACAGACAGGCCAGCCTGGAGATCAGC
CCCATCACCTTCTCACTGCCAGACCCCTGCTGATGGACCTCGGACAGTTCCCTGCTGCCCTGCCAC
ATCAGCTCCCACCAGCATGATGGCATGGAGGCCTATGTCAAGGTGGACAGCTGCCCTGAGGAGCCA
CAGCTCAGGATGAAAGAACATGAGGAGGCTGAGGACTATGATGATGACCTGACTGACTCTGAGATG
GATGTGGTCCCGCTTGATGATGACAACAGCCCATCCTCATTCAAGATCAGGTCTGTGGCAAGAAA
CACCCCAAGACCTGGGTGCACTACATTGCTGAGGAGGAGACTGGGACTATGCCCACTGGTC
CTGGCCCTGATGACAGGAGCTACAAGAGCCAGTACCTCAACAATGGCCACAGAGGATTGGACGC
AACTACAAGAACAGGTTCATGGCCTACACTGATGAAACCTCAAGACCAGGGAGGCCATTCAAG
CATGAGTCTGGCATCCTGGGCCACTCCTGTATGGGAGGTGGGGACACCCCTGCTCATCATCTTC
AAGAACCAAGGCCCTCCAGGCCCTACACATCTACCCACATGGCATCACTGATGTCAGGCCCTGTAC
AGCCGCAGGCTGCCAAAGGGGGTGAACACACCTCAAGGACTTCCCCATTCTGCCCTGGGAGATCTTC
AAAGTACAAGTGGACTGTCACTGTGGAGGATGGACCAACCAATCTGACCCCAAGGTGCCCTACCCAGA
TACTACTCCAGCTTGATGAAACATGGAGAGGGACCTGGCCTCTGGCCTGATTGGCCACTGCTCATC
TGCTACAAGGAGTCTGTGGACCAGAGGGAAACAGATCATGTCTGACAAGAGGAATGTGATTCTG
TTCTCTGTCTTGATGAGAACAGGAGCTGGTACCTGACTGAGAACATTCAAGGCCCTGCCAAC
CCTGCTGGGGTGCAGCTGGAGGACCCCTGAGTTCCAGGCCAGCAACATCATGCACTCCATCAATGGC
TATGTGTTTGACAGCCTCCAGCTTCTGTCTGCCATGAGGTGCCACTGGTACATTCTTCT
ATTGGGGCCCAAGACTGACATTCTTCTGTCTTCTCTGGCTACACCCITCAAACACAAAGATGGTG
TATGAGGACACCCCTGACCCCTTCCCATTCTCTGGGAGACTGTGTTCATGAGCATGGAGAACCC
GGCCTGTGGATCTGGATGCCACAACACTGTGACTTCCGCAACAGGGCATGACTGCCCTGCTCAA
GTCTCCTCCTGTGACAAGAACACTGGGACTACTATGAGGACAGCTATGAGGACATCTGCCCTAC
CTGCTCAGCAAGAACAAATGCCATTGAGGCCAGGGAGCTCAGCCAGAAATCCACCTGTCCCTGAAACGC
CACCAAGGGAGATCACCAGGACACCCCTCCAGTCTGACCCAGGGAGATTGACTATGATGACACC
ATTCTGTGGAGATGAAGAACAGGACTTGTGACATCTATGACGGAGACGAGAACAGAGGCCAAGG
AGCTTCCAGAACAGGCCAGGACTACTTCAATTGCTGCTGTGGAGGCCCTGTGGGACTATGGCATG
AGCTCCAGCCCCATGTCCCTGAGAACAGGGGCCAGTCTGGCTCTGTGCCACAGTTCAAGAACAGT
GTCTCCAAGAGGTTCACTGATGGCAGCTTCACCCAGCCCTGTACAGAGGGAGCTGAATGAGCAC
CTGGGACTCCTGGGCCATACATCAGGGCTGAGGTGGAGGACAACATCATGGTACCTCCGCAAC

(Continued)

Figure 58A

CAGGCCCTCCAGGCCCTACAGCTTCTACAGCTCCCTCATCAGCTATGAGGAGGACCAGAGGCAGGGG
GCTGAGCCACGCAAGAACCTTGTGAAACCCAATGAAACCAAGAACCTACTTCTGAAAGTCCAGCAC
CACATGGCCCCCACCAAGGATGAGTTGACTGCAAGGCCTGGCCTACTTCTGATGTGGACCTG
GAGAAGGATGTGCACTCTGGCTGATTGGCCCACCTCTGGCTGCCACACCAACACCCCTGAACCC
GCCCATGGAAGGCAAGTGACTGTGCAGGAGTTGCCCTCTTCTTCACCATCTTGATGAAACCAAG
AGCTGGTACTTCACTGAGAACATGGAGCGCACTGCAGGGCCCCATGCAACATTAGATGGAGGAC
CCCACCTCAAAGAGAACCTACCGCTTCCATGCCATCAATGGCTACATCATGGACACCCCTGCCTGG
CTTGTCATGGCCAGGACCAGAGGATCAGGTGGTACCTGCTTCTATGGGCTCCAATGAGAACATT
CACTOCATCCACTTCTGGCATGTTCACTGTGCCAAGAACAGGAGGAGTACAAGATGGCCCTG
TACAACCTCTACCCCTGGGTCTTGAGACTGTGGAGATGCTGCCCTCAAAGCTGGCATCTGGAGG
GTGGACTGCCCTCATGGGAGCACCTGCATGCTGGCATGAGCACCCCTGTTCTGGTCTACAGCAAC
AAGTGCCAGACCCCCCTGGGAATGCCCTGGCCACATCAGGGACTTCCAGATCACTGCCCTGGC
CAGTATGCCAGTGGCCCCAAGCTGCCAGGCTCCACTACTCTGGATCCATCAATGCCCTGGAGC
ACCAAGGAGCCATTAGCTGGATCAAAGTGGACCTGCTGGCCCCCATGATCATCCATGGCATCAAG
ACCCAGGGGGCCAGGCAGAACAGTTCTCCAGGCTGTACATCAGCCAGTTCATCATGTACAGCCTG
GATGGCAAGAAATGGCAGACCTACAGAGGCAACTCCACTGGAACACTCATGGTCTTCTTGGCAAT
GTGGACAGCTCTGGCATCAAGCACAAACATCTTCAACCCCCCAATCATGCCAGATACTCAGGCTG
CACCCCCACCCACTACAGCATCCGCAGCACCCCTCAGGATGGAGCTGATGGCTGTGACCTGAACCT
TGCAGCATGCCCTGGCATGGAGAGCAAGGCCATTCTGATGCCAGATCACTGCCCTCCAGCTAC
TTCACCAACATGTTGCCACCTGGAGGCCAAGCAAGGCCAGGCTGCACCTCCAGGGAGGAGGAAT
GCCCTGGAGGCCAGGTCAACAACCCAAAGGAGTGGCTGCAGGTGGACTTCCAGAAGACCATGAAG
GTCACTGGGGTGACCACCCAGGGGTCAAGAGCCTGCTCACCAGCATGTATGTGAAGGAGTTCCTG
ATCAGCTCCAGCCAGGATGCCACCAGTGGACCCCTTCTTCCAGAATGGCAAGGTCAAGGTGTT
CAGGGCAACCAGGACAGCTTCACCCCTGTGGTGAACAGCCTGGACCCCCCCCCTGACCAGATA
CTGAGGATTCAACCCCAGAGCTGGTCCACCAGATTGCCCTGAGGATGGAGGTCTGGATGTGAG
GCCCAGGACCTGTACTGA

Figure 58B

**VIRAL VECTORS ENCODING
RECOMBINANT FVIII VARIANTS WITH
INCREASED EXPRESSION FOR GENE
THERAPY OF HEMOPHILIA A**

**CROSS REFERENCE TO RELATED
APPLICATIONS**

[0001] This application is a Continuation of U.S. application Ser. No. 15/349,930, filed Nov. 11, 2016, which claims priority to U.S. Provisional Patent Application No. 62/255,317, filed Nov. 13, 2015, the content of which are hereby incorporated by reference in its entirety for all purposes.

SEQUENCE LISTING

[0002] The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on Nov. 9, 2016, is named 008073_5107_US02_Sequence_Listing.txt and is 345 KB bytes in size.

BACKGROUND OF THE DISCLOSURE

[0003] Blood coagulation proceeds through a complex and dynamic biological pathway of interdependent biochemical reactions, referred to as the coagulation cascade. Coagulation Factor VIII (FVIII) is a key component in the cascade. Factor VIII is recruited to bleeding sites, and forms a Xase complex with activated Factor IX (FIXa) and Factor X (FX). The Xase complex activates FX, which in turn activates prothrombin to thrombin, which then activates other components in the coagulation cascade to generate a stable clot (reviewed in Saenko et al., *Trends Cardiovasc. Med.*, 9:185-192 (1999); Lenting et al., *Blood*, 92:3983-3996 (1998)).

[0004] Hemophilia A is a congenital X-linked bleeding disorder characterized by a deficiency in Factor VIII activity. Diminished Factor VIII activity inhibits a positive feedback loop in the coagulation cascade. This causes incomplete coagulation, which manifests as bleeding episodes with increased duration, extensive bruising, spontaneous oral and nasal bleeding, joint stiffness and chronic pain, and possibly internal bleeding and anemia in severe cases (Zhang et al., *Clinic. Rev. Allerg. Immunol.*, 37:114-124 (2009)).

[0005] Conventionally, hemophilia A is treated by Factor VIII replacement therapy, which consists of administering Factor VIII protein (e.g., plasma-derived or recombinantly-produced Factor VIII) to an individual with hemophilia A. Factor VIII is administered prophylactically to prevent or reduce frequency of bleeding episodes, in response to an acute bleeding episode, and/or perioperatively to manage bleeding during surgery. However, there are several undesirable features of Factor VIII replacement therapy.

[0006] First, Factor VIII replacement therapy is used to treat or manage hemophilia A, but does not cure the underlying Factor VIII deficiency. Because of this, individuals with hemophilia A require Factor VIII replacement therapy for the duration of their lives. Continuous treatment is expensive and requires the individual to maintain strict compliance, as missing only a few prophylactic doses can have serious consequences for individuals with severe hemophilia A.

[0007] Second, because Factor VIII has a relatively short half-life in vivo, conventional prophylactic Factor VIII replacement therapy requires administration every second or

third day. This places a burden on the individual to maintain compliance throughout their life. While third generation “long-acting” Factor VIII drugs may reduce the frequency of administration, prophylactic Factor VIII replacement therapy with these drugs still requires monthly, weekly, or more frequent administration in perpetuity. For example, prophylactic treatment with ELOCTATE™ [Antihemophilic Factor (Recombinant), Fc Fusion Protein] requires administration every three to five days (ELOCTATE™ Prescribing Information, Biogen Idec Inc., (2015)). Moreover, the long-term effects of chemically modified biologics (e.g., pegylated polypeptides) are not yet fully understood.

[0008] Third, between 15% and 30% of all individuals receiving Factor VIII replacement therapy form anti-Factor VIII inhibitor antibodies, rendering the therapy inefficient. Factor VIII bypass therapy (e.g., administration of plasma-derived or recombinantly-produced prothrombin complex concentrates) can be used to treat hemophilia in individuals that form inhibitor antibodies. However, Factor VIII bypass therapy is less effective than Factor VIII replacement therapy (Mannucci P. M., *J Thromb Haemost.*, 1(7):1349-55 (2003)) and may be associated with an increased risk of cardiovascular complication (Luu and Ewenstein, *Haemophilia*, 10 Suppl. 2:10-16 (2004)).

[0009] Somatic gene therapy holds great promise for the treatment of hemophilia A because it would remedy the underlying under-expression functional Factor VIII activity (e.g., due to missense or nonsense mutations), rather than provide a one-time dose of Factor VIII activity to the individual. Because of this difference in the mechanism of action, as compared to Factor VIII replacement therapy, one-time administration of a Factor VIII gene therapy vector may provide an individual with Factor VIII for several years, reducing the cost of treatment and eliminating the need for continued patient compliance.

[0010] Coagulation Factor IX (FIX) gene therapy has been used effectively to treat individuals with hemophilia B, a related blood coagulation condition characterized by diminished Factor IX activity (Manno C. S., et al., *Nat Med.*, 12(3):342-47 (2006)). However, Factor VIII gene therapy presents several unique challenges. For example, the full-length, wild-type Factor VIII polypeptide (2351 amino acids; UniProt accession number P00451) is five times larger than the full-length, wild-type Factor IX polypeptide (461 amino acids; UniProt accession number P00740). As such, the coding sequence of wild-type Factor VIII is 7053 base pairs, which is too large to be packaged in conventional AAV gene therapy vectors. Further, reported recombinant expression of B-domain deleted variants of Factor VIII (BDD-FVIII) has been poor. As such, several groups have attempted to alter the codon usage of BDD-FVIII constructs, with limited success.

BRIEF SUMMARY OF DISCLOSURE

[0011] Accordingly, there is a need for Factor VIII variants whose coding sequences are more efficiently packaged into, and delivered via, gene therapy vectors. There is also a need for synthetic, codon-altered nucleic acids which express Factor VIII more efficiently. Such Factor VIII variants and codon-altered nucleic acids allow for improved treatment of Factor VIII deficiencies (e.g., hemophilia A). The above deficiencies and other problems associated with the treat-

ment of Factor VIII deficiencies (e.g., hemophilia A) are reduced or eliminated by the disclosed codon-altered Factor VIII variants.

[0012] In accordance with some embodiments, the present disclosure provides nucleic acids encoding Factor VIII variants that have high sequence identity to the disclosed codon-altered sequences of the Factor VIII heavy chain (e.g., CS01-HC-NA, CS04-HC-NA, or CS23-HC-NA) and light chain (CS01-LC-NA, CS04-LC-NA, or CS23-LC-NA). In some embodiments, these nucleic acids further include a sequence encoding a linker sequence that replaces the native Factor VIII B-domain (e.g., a linker sequences comprising a furin cleavage site), between the sequences coding for the Factor VIII heavy and light chains.

[0013] In one aspect, the disclosure provides a polynucleotide including a nucleotide sequence encoding a Factor VIII polypeptide. The Factor VIII polypeptide includes a light chain, a heavy chain, and a polypeptide linker joining the C-terminus of the heavy chain to the N-terminus of the light chain. The heavy chain of the Factor VIII polypeptide is encoded by a first nucleotide sequence having at least 95% identity to CS04-HC-NA (SEQ ID NO: 3). The light chain of the Factor FVIII polypeptide is encoded by a second nucleotide sequence having at least 95% identity to CS04-LC-NA (SEQ ID NO: 4). The polypeptide linker comprises a furin cleavage site.

[0014] In one embodiment of the polynucleotides described above, the polypeptide linker is encoded by a third nucleotide sequence having at least 95% identity to BDLO04 (SEQ ID NO: 6).

[0015] In one aspect, the disclosure provides a polynucleotide including a nucleotide sequence encoding a Factor VIII polypeptide. The Factor VIII polypeptide includes a light chain, a heavy chain, and a polypeptide linker joining the C-terminus of the heavy chain to the N-terminus of the light chain. The heavy chain of the Factor VIII polypeptide is encoded by a first nucleotide sequence having at least 95% identity to CS01-HC-NA (SEQ ID NO: 24). The light chain of the Factor FVIII polypeptide is encoded by a second nucleotide sequence having at least 95% identity to CS01-LC-NA (SEQ ID NO: 25). The polypeptide linker comprises a furin cleavage site.

[0016] In one embodiment of the polynucleotides described above, the polypeptide linker is encoded by a third nucleotide sequence having at least 95% identity to BDLO01 (SEQ ID NO: 5).

[0017] In one aspect, the disclosure provides a polynucleotide including a nucleotide sequence encoding a Factor VIII polypeptide. The Factor VIII polypeptide includes a light chain, a heavy chain, and a polypeptide linker joining the C-terminus of the heavy chain to the N-terminus of the light chain. The heavy chain of the Factor VIII polypeptide is encoded by a first nucleotide sequence having at least 95% identity to CS23-HC-NA (SEQ ID NO: 22). The light chain of the Factor FVIII polypeptide is encoded by a second nucleotide sequence having at least 95% identity to CS23-LC-NA (SEQ ID NO: 23). The polypeptide linker comprises a furin cleavage site.

[0018] In one embodiment of the polynucleotides described above, the polypeptide linker is encoded by a third nucleotide sequence having at least 95% identity to BDLO23 (SEQ ID NO: 7).

[0019] In one embodiment of the polynucleotides described above, the first nucleotide sequence encoding the

heavy chain of the Factor VIII polypeptide has at least 96% identity to the respective heavy chain sequence (e.g., CS04-HC-NA (SEQ ID NO: 3), CS01-HC-NA (SEQ ID NO: 24), or CS23-HC-NA (SEQ ID NO: 22)), and the second nucleotide sequence encoding the light chain of the Factor FVIII polypeptide has at least 96% identity to the respective light chain sequence (e.g., CS04-LC-NA (SEQ ID NO: 4), CS01-LC-NA (SEQ ID NO: 25), or CS23-LC-NA (SEQ ID NO: 23)).

[0020] In one embodiment of the polynucleotides described above, the first nucleotide sequence encoding the heavy chain of the Factor VIII polypeptide has at least 97% identity to the respective heavy chain sequence (e.g., CS04-HC-NA (SEQ ID NO: 3), CS01-HC-NA (SEQ ID NO: 24), or CS23-HC-NA (SEQ ID NO: 22)), and the second nucleotide sequence encoding the light chain of the Factor FVIII polypeptide has at least 97% identity to the respective light chain sequence (e.g., CS04-LC-NA (SEQ ID NO: 4), CS01-LC-NA (SEQ ID NO: 25), or CS23-LC-NA (SEQ ID NO: 23)).

[0021] In one embodiment of the polynucleotides described above, the first nucleotide sequence encoding the heavy chain of the Factor VIII polypeptide has at least 98% identity to the respective heavy chain sequence (e.g., CS04-HC-NA (SEQ ID NO: 3), CS01-HC-NA (SEQ ID NO: 24), or CS23-HC-NA (SEQ ID NO: 22)), and the second nucleotide sequence encoding the light chain of the Factor FVIII polypeptide has at least 98% identity to the respective light chain sequence (e.g., CS04-LC-NA (SEQ ID NO: 4), CS01-LC-NA (SEQ ID NO: 25), or CS23-LC-NA (SEQ ID NO: 23)).

[0022] In one embodiment of the polynucleotides described above, the first nucleotide sequence encoding the heavy chain of the Factor VIII polypeptide has at least 99% identity to the respective heavy chain sequence (e.g., CS04-HC-NA (SEQ ID NO: 3), CS01-HC-NA (SEQ ID NO: 24), or CS23-HC-NA (SEQ ID NO: 22)), and the second nucleotide sequence encoding the light chain of the Factor FVIII polypeptide has at least 99% identity to the respective light chain sequence (e.g., CS04-LC-NA (SEQ ID NO: 4), CS01-LC-NA (SEQ ID NO: 25), or CS23-LC-NA (SEQ ID NO: 23)).

[0023] In one embodiment of the polynucleotides described above, the first nucleotide sequence encoding the heavy chain of the Factor VIII polypeptide has at least 99.5% identity to the respective heavy chain sequence (e.g., CS04-HC-NA (SEQ ID NO: 3), CS01-HC-NA (SEQ ID NO: 24), or CS23-HC-NA (SEQ ID NO: 22)), and the second nucleotide sequence encoding the light chain of the Factor FVIII polypeptide has at least 99.5% identity to the respective light chain sequence (e.g., CS04-LC-NA (SEQ ID NO: 4), CS01-LC-NA (SEQ ID NO: 25), or CS23-LC-NA (SEQ ID NO: 23)).

[0024] In one embodiment of the polynucleotides described above, the first nucleotide sequence encoding the heavy chain of the Factor VIII polypeptide has at least 99.9% identity to the respective heavy chain sequence (e.g., CS04-HC-NA (SEQ ID NO: 3), CS01-HC-NA (SEQ ID NO: 24), or CS23-HC-NA (SEQ ID NO: 22)), and the second nucleotide sequence encoding the light chain of the Factor FVIII polypeptide has at least 99.9% identity to the respective light chain sequence (e.g., CS04-LC-NA (SEQ ID NO: 4), CS01-LC-NA (SEQ ID NO: 25), or CS23-LC-NA (SEQ ID NO: 23)).

[0025] In one embodiment of the polynucleotides described above, the first nucleotide sequence encoding the heavy chain of the Factor VIII polypeptide is CS04-HC-NA (SEQ ID NO: 3), and the second nucleotide sequence encoding the light chain of the Factor FVIII polypeptide is CS04-LC-NA (SEQ ID NO: 4).

[0026] In one embodiment of the polynucleotides described above, the first nucleotide sequence encoding the heavy chain of the Factor VIII polypeptide is CS01-HC-NA (SEQ ID NO: 24), and the second nucleotide sequence encoding the light chain of the Factor FVIII polypeptide is CS01-LC-NA (SEQ ID NO: 25).

[0027] In one embodiment of the polynucleotides described above, the first nucleotide sequence encoding the heavy chain of the Factor VIII polypeptide is CS23-HC-NA (SEQ ID NO: 22), and the second nucleotide sequence encoding the light chain of the Factor FVIII polypeptide is CS23-LC-NA (SEQ ID NO: 23).

[0028] In one aspect, the disclosure provides a polynucleotide comprising a nucleotide sequence having at least 95% identity to CS04-FL-NA, wherein the polynucleotide encodes a Factor VIII polypeptide.

[0029] In one aspect, the disclosure provides a polynucleotide comprising a nucleotide sequence having at least 95% identity to CS01-FL-NA, wherein the polynucleotide encodes a Factor VIII polypeptide.

[0030] In one aspect, the disclosure provides a polynucleotide comprising a nucleotide sequence having at least 95% identity to CS23-FL-NA, wherein the polynucleotide encodes a Factor VIII polypeptide.

[0031] In one embodiment of the polynucleotides described above, the nucleotide sequence has at least 96% identity to the respective full-length polynucleotide sequence (e.g., CS04-FL-NA (SEQ ID NO: 1), CS01-FL-NA (SEQ ID NO: 13), or CS23-FL-NA (SEQ ID NO: 20)).

[0032] In one embodiment of the polynucleotides described above, the nucleotide sequence has at least 97% identity to the respective full-length polynucleotide sequence (e.g., CS04-FL-NA (SEQ ID NO: 1), CS01-FL-NA (SEQ ID NO: 13), or CS23-FL-NA (SEQ ID NO: 20)).

[0033] In one embodiment of the polynucleotides described above, the nucleotide sequence has at least 98% identity to the respective full-length polynucleotide sequence (e.g., CS04-FL-NA (SEQ ID NO: 1), CS01-FL-NA (SEQ ID NO: 13), or CS23-FL-NA (SEQ ID NO: 20)).

[0034] In one embodiment of the polynucleotides described above, the nucleotide sequence has at least 99% identity to the respective full-length polynucleotide sequence (e.g., CS04-FL-NA (SEQ ID NO: 1), CS01-FL-NA (SEQ ID NO: 13), or CS23-FL-NA (SEQ ID NO: 20)).

[0035] In one embodiment of the polynucleotides described above, the nucleotide sequence has at least 99.5% identity to the respective full-length polynucleotide sequence (e.g., CS04-FL-NA (SEQ ID NO: 1), CS01-FL-NA (SEQ ID NO: 13), or CS23-FL-NA (SEQ ID NO: 20)).

[0036] In one embodiment of the polynucleotides described above, the nucleotide sequence has at least 99.9% identity to the respective full-length polynucleotide sequence (e.g., CS04-FL-NA (SEQ ID NO: 1), CS01-FL-NA (SEQ ID NO: 13), or CS23-FL-NA (SEQ ID NO: 20)).

[0037] In one embodiment of the polynucleotides described above, the nucleotide sequence is CS04-FL-NA (SEQ ID NO: 1).

[0038] In one embodiment of the polynucleotides described above, the nucleotide sequence is CS01-FL-NA (SEQ ID NO: 13).

[0039] In one embodiment of the polynucleotides described above, the nucleotide sequence is CS23-FL-NA (SEQ ID NO: 20).

[0040] In one embodiment of the polynucleotides described above, the polynucleotide encodes a Factor VIII polypeptide comprising an amino acid sequence having at least 95% identity to CS04-FL-AA (SEQ ID NO: 2).

[0041] In one embodiment of the polynucleotides described above, the polynucleotide encodes a Factor VIII polypeptide comprising an amino acid sequence having at least 96% identity to CS04-FL-AA (SEQ ID NO: 2).

[0042] In one embodiment of the polynucleotides described above, the polynucleotide encodes a Factor VIII polypeptide comprising an amino acid sequence having at least 97% identity to CS04-FL-AA (SEQ ID NO: 2).

[0043] In one embodiment of the polynucleotides described above, the polynucleotide encodes a Factor VIII polypeptide comprising an amino acid sequence having at least 98% identity to CS04-FL-AA (SEQ ID NO: 2).

[0044] In one embodiment of the polynucleotides described above, the polynucleotide encodes a Factor VIII polypeptide comprising an amino acid sequence having at least 99% identity to CS04-FL-AA (SEQ ID NO: 2).

[0045] In one embodiment of the polynucleotides described above, the polynucleotide encodes a Factor VIII polypeptide comprising an amino acid sequence having at least 99.5% identity to CS04-FL-AA (SEQ ID NO: 2).

[0046] In one embodiment of the polynucleotides described above, the polynucleotide encodes a Factor VIII polypeptide comprising an amino acid sequence having at least 99.9% identity to CS04-FL-AA (SEQ ID NO: 2).

[0047] In one embodiment of the polynucleotides described above, the polynucleotide encodes a Factor VIII polypeptide comprising the amino acid sequence of CS04-FL-AA (SEQ ID NO: 2).

[0048] In one aspect, the disclosure provides a polynucleotide comprising a nucleotide sequence having at least 95% identity to CS04-SC1-NA (SEQ ID NO: 9), wherein the polynucleotide encodes a single-chain Factor VIII polypeptide.

[0049] In one aspect, the disclosure provides a polynucleotide comprising a nucleotide sequence having at least 95% identity to CS04-SC2-NA (SEQ ID NO: 11), wherein the polynucleotide encodes a single-chain Factor VIII polypeptide.

[0050] In one aspect, the disclosure provides a polynucleotide comprising a nucleotide sequence having at least 95% identity to CS01-SC1-NA (SEQ ID NO: 26), wherein the polynucleotide encodes a single-chain Factor VIII polypeptide.

[0051] In one aspect, the disclosure provides a polynucleotide comprising a nucleotide sequence having at least 95% identity to CS01-SC2-NA (SEQ ID NO: 27), wherein the polynucleotide encodes a single-chain Factor VIII polypeptide.

[0052] In one aspect, the disclosure provides a polynucleotide comprising a nucleotide sequence having at least 95% identity to CS23-SC1-NA (SEQ ID NO: 28), wherein the polynucleotide encodes a single-chain Factor VIII polypeptide.

[0053] In one aspect, the disclosure provides a polynucleotide comprising a nucleotide sequence having at least 95% identity to CS23-SC2-NA (SEQ ID NO: 29), wherein the polynucleotide encodes a single-chain Factor VIII polypeptide.

[0054] In one embodiment of the polynucleotides described above, the nucleotide sequence has at least 96% identity to the respective full-length polynucleotide sequence (e.g., CS04-SC1-NA (SEQ ID NO: 9), CS04-SC2-NA (SEQ ID NO: 11), CS01-SC1-NA (SEQ ID NO: 26), CS01-SC2-NA (SEQ ID NO: 27), CS23-SC1-NA (SEQ ID NO: 28), or CS23-SC2-NA (SEQ ID NO: 29)).

[0055] In one embodiment of the polynucleotides described above, the nucleotide sequence has at least 97% identity to the respective full-length polynucleotide sequence (e.g., CS04-SC1-NA (SEQ ID NO: 9), CS04-SC2-NA (SEQ ID NO: 11), CS01-SC1-NA (SEQ ID NO: 26), CS01-SC2-NA (SEQ ID NO: 27), CS23-SC1-NA (SEQ ID NO: 28), or CS23-SC2-NA (SEQ ID NO: 29)).

[0056] In one embodiment of the polynucleotides described above, the nucleotide sequence has at least 98% identity to the respective full-length polynucleotide sequence (e.g., CS04-SC1-NA (SEQ ID NO: 9), CS04-SC2-NA (SEQ ID NO: 11), CS01-SC1-NA (SEQ ID NO: 26), CS01-SC2-NA (SEQ ID NO: 27), CS23-SC1-NA (SEQ ID NO: 28), or CS23-SC2-NA (SEQ ID NO: 29)).

[0057] In one embodiment of the polynucleotides described above, the nucleotide sequence has at least 99% identity to the respective full-length polynucleotide sequence (e.g., CS04-SC1-NA (SEQ ID NO: 9), CS04-SC2-NA (SEQ ID NO: 11), CS01-SC1-NA (SEQ ID NO: 26), CS01-SC2-NA (SEQ ID NO: 27), CS23-SC1-NA (SEQ ID NO: 28), or CS23-SC2-NA (SEQ ID NO: 29)).

[0058] In one embodiment of the polynucleotides described above, the nucleotide sequence has at least 99.5% identity to the respective full-length polynucleotide sequence (e.g., CS04-SC1-NA (SEQ ID NO: 9), CS04-SC2-NA (SEQ ID NO: 11), CS01-SC1-NA (SEQ ID NO: 26), CS01-SC2-NA (SEQ ID NO: 27), CS23-SC1-NA (SEQ ID NO: 28), or CS23-SC2-NA (SEQ ID NO: 29)).

[0059] In one embodiment of the polynucleotides described above, the nucleotide sequence has at least 99.9% identity to the respective full-length polynucleotide sequence (e.g., CS04-SC1-NA (SEQ ID NO: 9), CS04-SC2-NA (SEQ ID NO: 11), CS01-SC1-NA (SEQ ID NO: 26), CS01-SC2-NA (SEQ ID NO: 27), CS23-SC1-NA (SEQ ID NO: 28), or CS23-SC2-NA (SEQ ID NO: 29)).

[0060] In one embodiment of the polynucleotides described above, the nucleotide sequence is CS04-SC1-NA (SEQ ID NO: 9).

[0061] In one embodiment of the polynucleotides described above, the nucleotide sequence is CS04-SC2-NA (SEQ ID NO: 11).

[0062] In one embodiment of the polynucleotides described above, the nucleotide sequence is CS01-SC1-NA (SEQ ID NO: 26).

[0063] In one embodiment of the polynucleotides described above, the nucleotide sequence is CS01-SC2-NA (SEQ ID NO: 27).

[0064] In one embodiment of the polynucleotides described above, the nucleotide sequence is CS23-SC1-NA (SEQ ID NO: 28).

[0065] In one embodiment of the polynucleotides described above, the nucleotide sequence is CS23-SC2-NA (SEQ ID NO: 29).

[0066] In one embodiment of the polynucleotides described above, the nucleotide sequence has at least 95% identity to a sequence selected from the group consisting of CS01-FL-NA, CS01-HC-NA, CS01-LC-NA, CS04-FL-NA, CS04-HC-NA, CS04-LC-NA, CS23-FL-NA, CS23-HC-NA, CS23-LC-NA, CS01m1-FL-NA, CS01m2-FL-NA, CS01m3-FL-NA, CS01m4-FL-NA, CS01m12-FL-NA, CS01m13-FL-NA, CS01m23-FL-NA, CS01m24-FL-NA, CS01m234-FL-NA, CS04m1-FL-NA, CS04m2-FL-NA, CS04m3-FL-NA, CS04m4-FL-NA, CS04m12-FL-NA, CS04m13-FL-NA, CS04m23-FL-NA, CS04m24-FL-NA, CS04m34-FL-NA, CS04m123-FL-NA, CS04m234-FL-NA, CS23m1-FL-NA, CS23m2-FL-NA, CS23m3-FL-NA, CS23m4-FL-NA, CS23m12-FL-NA, CS23m13-FL-NA, CS23m13-FL-NA, CS23m23-FL-NA, CS23m24-FL-NA, CS23m34-FL-NA, CS23m123-FL-NA, CS23m234-FL-NA, CS01-SC1-NA, CS04-SC1-NA, CS23-SC1-NA, CS01m1-SC1-NA, CS01m2-SC1-NA, CS01m3-SC1-NA, CS01m4-SC1-NA, CS01m12-SC1-NA, CS01m13-SC1-NA, CS01m23-SC1-NA, CS01m24-SC1-NA, CS01m34-SC1-NA, CS04m1-SC1-NA, CS04m2-SC1-NA, CS04m3-SC1-NA, CS04m4-SC1-NA, CS04m12-SC1-NA, CS04m13-SC1-NA, CS04m23-SC1-NA, CS04m234-SC1-NA, CS04m34-SC1-NA, CS04m123-SC1-NA, CS04m234-SC1-NA, CS23m1-SC1-NA, CS23m2-SC1-NA, CS23m3-SC1-NA, CS23m4-SC1-NA, CS23m12-SC1-NA, CS23m13-SC1-NA, CS23m23-SC1-NA, CS23m24-SC1-NA, CS23m34-SC1-NA, CS23m123-SC1-NA, and CS23m234-SC1-NA.

[0067] In one embodiment of the polynucleotides described above, the nucleotide sequence has at least 96% identity to a sequence selected from the group consisting of CS01-FL-NA, CS01-HC-NA, CS01-LC-NA, CS04-FL-NA, CS04-HC-NA, CS04-LC-NA, CS23-FL-NA, CS23-HC-NA, CS23-LC-NA, CS01m1-FL-NA, CS01m2-FL-NA, CS01m3-FL-NA, CS01m4-FL-NA, CS01m12-FL-NA, CS01m13-FL-NA, CS01m23-FL-NA, CS01m24-FL-NA, CS01m34-FL-NA, CS01m234-FL-NA, CS04m1-FL-NA, CS04m2-FL-NA, CS04m3-FL-NA, CS04m4-FL-NA, CS04m12-FL-NA, CS04m13-FL-NA, CS04m23-FL-NA, CS04m24-FL-NA, CS04m34-FL-NA, CS04m123-FL-NA, CS04m234-FL-NA, CS23m1-FL-NA, CS23m2-FL-NA, CS23m3-FL-NA, CS23m4-FL-NA, CS23m12-FL-NA, CS23m13-FL-NA, CS23m23-FL-NA, CS23m24-FL-NA, CS23m34-FL-NA, CS23m123-FL-NA, CS23m234-FL-NA, CS01-SC1-NA, CS04-SC1-NA, CS23-SC1-NA, CS01m1-SC1-NA, CS01m2-SC1-NA, CS01m3-SC1-NA, CS01m4-SC1-NA, CS01m12-SC1-NA,

CS01m13-SC1-NA, CS01m23-SC1-NA, CS01m24-SC1-NA, CS01m34-SC1-NA, CS01m123-SC1-NA, CS01m234-SC1-NA, CS04m1-SC1-NA, CS04m2-SC1-NA, CS04m3-SC1-NA, CS04m4-SC1-NA, CS04m12-SC1-NA, CS04m13-SC1-NA, CS04m23-SC1-NA, CS04m24-SC1-NA, CS04m34-SC1-NA, CS04m123-SC1-NA, CS04m234-SC1-NA, CS23m1-SC1-NA, CS23m2-SC1-NA, CS23m3-SC1-NA, CS23m4-SC1-NA, CS23m12-SC1-NA, CS23m13-SC1-NA, CS23m23-SC1-NA, CS23m24-SC1-NA, CS23m34-SC1-NA, CS23m123-SC1-NA, CS23m234-SC1-NA, CS01-SC2-NA, CS04-SC2-NA, CS23-SC2-NA, CS01m1-SC2-NA, CS01m2-SC2-NA, CS01m3-SC2-NA, CS01m4-SC2-NA, CS01m12-SC2-NA, CS01m13-SC2-NA, CS01m23-SC2-NA, CS01m24-SC2-NA, CS01m34-SC2-NA, CS01m123-SC2-NA, CS01m234-SC2-NA, CS04m1-SC2-NA, CS04m2-SC2-NA, CS04m3-SC2-NA, CS04m4-SC2-NA, CS04m12-SC2-NA, CS04m13-SC2-NA, CS04m23-SC2-NA, CS04m24-SC2-NA, CS04m34-SC2-NA, CS04m123-SC2-NA, CS04m234-SC2-NA, CS23m1-SC2-NA, CS23m2-SC2-NA, CS23m3-SC2-NA, CS23m4-SC2-NA, CS23m12-SC2-NA, CS23m13-SC2-NA, CS23m23-SC2-NA, CS23m24-SC2-NA, CS23m34-SC2-NA, CS23m123-SC2-NA, and CS23m234-SC2-NA.

[0068] In one embodiment of the polynucleotides described above, the nucleotide sequence has at least 97% identity to a sequence selected from the group consisting of CS01-FL-NA, CS01-HC-NA, CS01-LC-NA, CS04-FL-NA, CS04-HC-NA, CS04-LC-NA, CS23-FL-NA, CS23-HC-NA, CS23-LC-NA, CS01m1-FL-NA, CS01m2-FL-NA, CS01m3-FL-NA, CS01m4-FL-NA, CS01m12-FL-NA, CS01m13-FL-NA, CS01m23-FL-NA, CS01m24-FL-NA, CS01m34-FL-NA, CS04m1-FL-NA, CS04m2-FL-NA, CS04m3-FL-NA, CS04m4-FL-NA, CS04m12-FL-NA, CS04m13-FL-NA, CS04m23-FL-NA, CS04m24-FL-NA, CS04m34-FL-NA, CS04m123-FL-NA, CS04m234-FL-NA, CS23m1-FL-NA, CS23m2-FL-NA, CS23m3-FL-NA, CS23m4-FL-NA, CS23m12-FL-NA, CS23m13-FL-NA, CS23m23-FL-NA, CS23m24-FL-NA, CS23m34-FL-NA, CS23m123-FL-NA, CS23m234-FL-NA, CS01-SC1-NA, CS04-SC1-NA, CS23-SC1-NA, CS01m1-SC1-NA, CS01m2-SC1-NA, CS01m3-SC1-NA, CS01m4-SC1-NA, CS01m12-SC1-NA, CS01m13-SC1-NA, CS01m23-SC1-NA, CS01m24-SC1-NA, CS01m34-SC1-NA, CS01m123-SC1-NA, CS01m234-SC1-NA, CS04m1-SC1-NA, CS04m2-SC1-NA, CS04m3-SC1-NA, CS04m4-SC1-NA, CS04m12-SC1-NA, CS04m13-SC1-NA, CS04m23-SC1-NA, CS04m24-SC1-NA, CS04m34-SC1-NA, CS04m123-SC1-NA, CS04m234-SC1-NA, CS23m1-SC1-NA, CS23m2-SC1-NA, CS23m3-SC1-NA, CS23m4-SC1-NA, CS23m12-SC1-NA, CS23m13-SC1-NA, CS23m23-SC1-NA, CS23m24-SC1-NA, CS23m34-SC1-NA, CS23m123-SC1-NA, and CS23m234-SC1-NA.

NA, CS23m23-SC2-NA, CS23m24-SC2-NA, CS23m34-SC2-NA, CS23m123-SC2-NA, and CS23m234-SC2-NA.

[0069] In one embodiment of the polynucleotides described above, the nucleotide sequence has at least 98% identity to a sequence selected from the group consisting of CS01-FL-NA, CS01-HC-NA, CS01-LC-NA, CS04-FL-NA, CS04-HC-NA, CS04-LC-NA, CS23-FL-NA, CS23-HC-NA, CS23-LC-NA, CS01m1-FL-NA, CS01m2-FL-NA, CS01m3-FL-NA, CS01m4-FL-NA, CS01m12-FL-NA, CS01m13-FL-NA, CS01m23-FL-NA, CS01m24-FL-NA, CS01m34-FL-NA, CS04m1-FL-NA, CS04m2-FL-NA, CS04m3-FL-NA, CS04m4-FL-NA, CS04m12-FL-NA, CS04m13-FL-NA, CS04m23-FL-NA, CS04m24-FL-NA, CS04m34-FL-NA, CS04m123-FL-NA, CS04m234-FL-NA, CS23m1-FL-NA, CS23m2-FL-NA, CS23m3-FL-NA, CS23m4-FL-NA, CS23m12-FL-NA, CS23m13-FL-NA, CS23m23-FL-NA, CS23m24-FL-NA, CS23m34-FL-NA, CS23m123-FL-NA, CS23m234-FL-NA, CS01-SC1-NA, CS04-SC1-NA, CS23-SC1-NA, CS01m1-SC1-NA, CS01m2-SC1-NA, CS01m3-SC1-NA, CS01m4-SC1-NA, CS01m12-SC1-NA, CS01m13-SC1-NA, CS01m23-SC1-NA, CS01m24-SC1-NA, CS01m34-SC1-NA, CS01m123-SC1-NA, CS01m234-SC1-NA, CS04m1-SC1-NA, CS04m2-SC1-NA, CS04m3-SC1-NA, CS04m4-SC1-NA, CS04m12-SC1-NA, CS04m13-SC1-NA, CS04m23-SC1-NA, CS04m24-SC1-NA, CS04m34-SC1-NA, CS04m123-SC1-NA, CS04m234-SC1-NA, CS23m1-SC1-NA, CS23m2-SC1-NA, CS23m3-SC1-NA, CS23m4-SC1-NA, CS23m12-SC1-NA, CS23m13-SC1-NA, CS23m23-SC1-NA, CS23m24-SC1-NA, CS23m34-SC1-NA, CS23m123-SC1-NA, and CS23m234-SC1-NA.

[0070] In one embodiment of the polynucleotides described above, the nucleotide sequence has at least 99% identity to a sequence selected from the group consisting of CS01-FL-NA, CS01-HC-NA, CS01-LC-NA, CS04-FL-NA, CS04-HC-NA, CS04-LC-NA, CS23-FL-NA, CS23-HC-NA, CS23-LC-NA, CS01m1-FL-NA, CS01m2-FL-NA, CS01m3-FL-NA, CS01m4-FL-NA, CS01m12-FL-NA, CS01m13-FL-NA, CS01m23-FL-NA, CS01m24-FL-NA, CS01m34-FL-NA, CS01m123-FL-NA, CS01m234-FL-NA, CS04m1-FL-NA, CS04m2-FL-NA, CS04m3-FL-NA, CS04m4-FL-NA, CS04m12-FL-NA, CS04m13-FL-NA, CS04m23-FL-NA, CS04m24-FL-NA, CS04m34-FL-NA, CS04m123-FL-NA, CS04m234-FL-NA, CS23m1-FL-NA, CS23m2-FL-NA, CS23m3-FL-NA, CS23m4-FL-NA, CS23m12-FL-NA, CS23m13-FL-NA, CS23m23-FL-NA, CS23m24-FL-NA, CS23m34-FL-NA, CS23m123-FL-NA, CS23m234-FL-NA, CS01-SC1-NA, CS04-SC1-NA, CS23-SC1-NA, CS01m1-SC1-NA, CS01m2-SC1-NA, CS01m3-SC1-NA, CS01m4-SC1-NA, CS01m12-SC1-NA, CS01m13-SC1-NA, CS01m23-SC1-NA, CS01m24-SC1-NA, CS01m34-SC1-NA, CS01m123-SC1-NA, and CS01m234-SC1-NA.

CS04m12-SC1-NA, CS04m13-SC1-NA, CS04m23-SC1-NA, CS04m24-SC1-NA, CS04m34-SC1-NA, CS04m123-SC1-NA, CS04m234-SC1-NA, CS23m1-SC1-NA, CS23m2-SC1-NA, CS23m3-SC1-NA, CS23m4-SC1-NA, CS23m12-SC1-NA, CS23m13-SC1-NA, CS23m23-SC1-NA, CS23m24-SC1-NA, CS23m34-SC1-NA, CS23m123-SC1-NA, CS23m234-SC1-NA, CS01-SC2-NA, CS04-SC2-NA, CS23-SC2-NA, CS01m1-SC2-NA, CS01m2-SC2-NA, CS01m3-SC2-NA, CS01m4-SC2-NA, CS01m12-SC2-NA, CS01m13-SC2-NA, CS01m23-SC2-NA, CS01m24-SC2-NA, CS01m34-SC2-NA, CS01m123-SC2-NA, CS01m234-SC2-NA, CS04m1-SC2-NA, CS04m2-SC2-NA, CS04m3-SC2-NA, CS04m4-SC2-NA, CS04m12-SC2-NA, CS04m13-SC2-NA, CS04m23-SC2-NA, CS04m24-SC2-NA, CS04m34-SC2-NA, CS04m123-SC2-NA, CS04m234-SC2-NA, CS23m1-SC2-NA, CS23m2-SC2-NA, CS23m3-SC2-NA, CS23m4-SC2-NA, CS23m12-SC2-NA, CS23m13-SC2-NA, CS23m23-SC2-NA, CS23m24-SC2-NA, CS23m34-SC2-NA, CS23m123-SC2-NA, and CS23m234-SC2-NA.

[0074] In one embodiment of the polynucleotides described above, the encoded Factor VIII polypeptide comprises a glycosylation polypeptide positioned between two consecutive amino acids.

[0075] In one embodiment of the polynucleotides described above, the encoded polypeptide linker includes a glycosylation peptide with an amino acid sequence having at least 92% identity to a glycosylation peptide selected from the group consisting of NG1-AA, NG4-AA, NG5-AA, NG6-AA, NG7-AA, NG9-AA, NG10-AA, NG16-AA, NG17-AA, NG18-AA, NG19-AA, NG20-AA, NG21-AA and NGV-AA.

[0076] In one embodiment of the polynucleotides described above, the encoded polypeptide linker comprises a glycosylation peptide with an amino acid sequence selected from the group consisting of NG1-AA, NG4-AA, NG5-AA, NG6-AA, NG7-AA, NG9-AA, NG10-AA, NG16-AA, NG17-AA, NG18-AA, NG19-AA, NG20-AA, NG21-AA and NGV-AA.

[0077] In one embodiment of the polynucleotides described above, the glycosylation peptide is encoded by a polynucleotide with a nucleotide sequence having at least 95% identity to a sequence selected from the group consisting of NG1-NA, NG4-NA, NG5-NA, NG6-NA, NG7-NA, NG9-NA, NG10-NA, NG16-NA, NG17-NA, NG18-NA, NG19-NA, NG20-NA, NG21-NA and NGV-NA.

[0078] In one embodiment of the polynucleotides described above, the glycosylation peptide is encoded by a polynucleotide with a nucleotide sequence selected from one of NG1-NA, NG4-NA, NG5-NA, NG6-NA, NG7-NA, NG9-NA, NG10-NA, NG16-NA, NG17-NA, NG18-NA, NG19-NA, NG20-NA, NG21-NA and NGV-NA.

[0079] In one embodiment of the polynucleotides described above, the polypeptide linker is encoded by a third nucleotide sequence having at least 95% identity to a sequence selected from the group consisting of BDLNG1-NA, BDLNG3-NA, BDLNG5-NA, BDLNG6-NA, BDLNG9-NA, BDLNG10-NA, BDLNG16-NA, BDLNG17-NA, BDLNG18-NA, BDLNG19-NA, BDLNG20-NA and BDLNG21-NA.

[0080] In one embodiment of the polynucleotides described above, the encoded Factor VIII polypeptide includes an F328S (SPI; F309S SPE) amino acid substitution, relative to FVIII-FL-AA (SEQ ID NO: 19).

[0081] In one embodiment of the polynucleotides described above, the encoded Factor VIII polypeptide includes I105V, A127S, G151K, M166T, and L171P (SPI; I86V, A108S, G132K, M147T, and L152P, SPE, respectively) amino acid substitutions, relative to FVIII-FL-AA (SEQ ID NO: 19).

[0082] In one embodiment of the polynucleotides described above, the encoded Factor VIII polypeptide includes a) a deletion of amino acids AIEPR755-759, relative to FVIII-FL-AA (SEQ ID NO: 19), and b) an insertion of amino acids TTYVNRS (SEQ ID NO: 33) after N754, relative to FVIII-FL-AA (SEQ ID NO: 19). In some embodiments (e.g., where the encoded FVIII molecule includes a portion of the N-terminal region of the wild-type B-domain), the encoded Factor VIII polypeptide also includes a deletion of amino acids SF760-761, relative to FVIII-FL-AA (SEQ ID NO: 19).

[0083] In one embodiment of the polynucleotides described above, the encoded Factor VIII polypeptide includes a) an F328S (SPI; F3095 SPE) amino acid substitution, relative to FVIII-FL-AA (SEQ ID NO: 19), and b) C1918G and C1922G (SPI; C1899G and C1903 SPE, respectively) amino acid substitutions, relative to FVIII-FL-AA (SEQ ID NO: 19).

[0084] In one embodiment of the polynucleotides described above, the encoded Factor VIII polypeptide includes a) an F328S (SPI; F3095 SPE) amino acid substitution, relative to FVIII-FL-AA (SEQ ID NO: 19), and b) I105V, A127S, G151K, M166T, and L171P (SPI; I86V, A108S, G132K, M147T, and L152P, SPE, respectively) amino acid substitutions, relative to FVIII-FL-AA (SEQ ID NO: 19).

[0085] In one embodiment of the polynucleotides described above, the encoded Factor VIII polypeptide includes a) an F328S amino acid substitution, relative to FVIII-FL-AA (SEQ ID NO: 19), b) a deletion of amino acids AIEPR755-759, relative to FVIII-FL-AA (SEQ ID NO: 19), and c) an insertion of amino acids TTYVNRS (SEQ ID NO: 33) after N754, relative to FVIII-FL-AA (SEQ ID NO: 19). In some embodiments (e.g., where the encoded FVIII molecule includes a portion of the N-terminal region of the wild-type B-domain), the encoded Factor VIII polypeptide also includes a deletion of amino acids SF760-761, relative to FVIII-FL-AA (SEQ ID NO: 19).

[0086] In one embodiment of the polynucleotides described above, the encoded Factor VIII polypeptide includes a) I105V, A127S, G151K, M166T, and L171P amino acid substitutions, relative to FVIII-FL-AA (SEQ ID NO: 19), b) a deletion of amino acids AIEPR755-759, relative to FVIII-FL-AA (SEQ ID NO: 19), and c) an insertion of amino acids TTYVNRS (SEQ ID NO: 33) after N754, relative to FVIII-FL-AA (SEQ ID NO: 19). In some embodiments (e.g., where the encoded FVIII molecule includes a portion of the N-terminal region of the wild-type B-domain), the encoded Factor VIII polypeptide also includes a deletion of amino acids SF760-761, relative to FVIII-FL-AA (SEQ ID NO: 19).

[0087] In one embodiment of the polynucleotides described above, the encoded Factor VIII polypeptide includes a) an F328S amino acid substitution, relative to FVIII-FL-AA (SEQ ID NO: 19), b) C1918G and C1922G amino acid substitutions, relative to FVIII-FL-AA (SEQ ID NO: 19).

NO: 19), and c) I105V, A127S, G151K, M166T, and L171P amino acid substitutions, relative to FVIII-FL-AA (SEQ ID NO: 19).

[0088] In one embodiment of the polynucleotides described above, the encoded Factor VIII polypeptide includes a) an F328S amino acid substitution, relative to FVIII-FL-AA (SEQ ID NO: 19), b) C1918G and C1922G amino acid substitutions, relative to FVIII-FL-AA (SEQ ID NO: 19), c) a deletion of amino acids AIEPR755-759, relative to FVIII-FL-AA (SEQ ID NO: 19), and d) an insertion of amino acids TTYVNRSR (SEQ ID NO: 33) after N754, relative to FVIII-FL-AA (SEQ ID NO: 19). In some embodiments (e.g., where the encoded FVIII molecule includes a portion of the N-terminal region of the wild-type B-domain), the encoded Factor VIII polypeptide also includes a deletion of amino acids SF760-761, relative to FVIII-FL-AA (SEQ ID NO: 19).

[0089] In one embodiment of the polynucleotides described above, the encoded Factor VIII polypeptide includes a) I105V, A127S, G151K, M166T, and L171P amino acid substitutions, relative to FVIII-FL-AA (SEQ ID NO: 19), b) an F328S amino acid substitution, relative to FVIII-FL-AA (SEQ ID NO: 19), c) C1918G and C1922G amino acid substitutions, relative to FVIII-FL-AA (SEQ ID NO: 19), d) a deletion of amino acids AIEPR755-759, relative to FVIII-FL-AA (SEQ ID NO: 19), and e) an insertion of amino acids TTYVNRSR (SEQ ID NO: 33) after N754, relative to FVIII-FL-AA (SEQ ID NO: 19). In some embodiments (e.g., where the encoded FVIII molecule includes a portion of the N-terminal region of the wild-type B-domain), the encoded Factor VIII polypeptide also includes a deletion of amino acids SF760-761, relative to FVIII-FL-AA (SEQ ID NO: 19).

[0090] In one embodiment of the polynucleotides described above, the polynucleotide also includes a promoter element operably linked to the polynucleotide encoding the Factor VIII polypeptide.

[0091] In one embodiment of the polynucleotides described above, the polynucleotide also includes an enhancer element operably linked to the polynucleotide encoding the Factor VIII polypeptide.

[0092] In one embodiment of the polynucleotides described above, the polynucleotide also includes a polyadenylation element operably linked to the polynucleotide encoding the Factor VIII polypeptide.

[0093] In one embodiment of the polynucleotides described above, the polynucleotide also includes an intron operatively linked to the nucleotide sequence encoding the Factor VIII polypeptide.

[0094] In one embodiment of the polynucleotides described above, the intron is positioned between a promoter element and the translation initiation site (e.g., the first coding ATG) of the nucleotide sequence encoding a Factor VIII polypeptide.

[0095] In another aspect, the disclosure provides a mammalian gene therapy vector including a polynucleotide as described above.

[0096] In one embodiment of the mammalian gene therapy vector described above, the mammalian gene therapy vector is an adeno-associated virus (AAV) vector.

[0097] In one embodiment of the mammalian gene therapy vector described above, the AAV vector is an AAV-8 vector.

[0098] In another aspect, the disclosure provides a method for treating hemophilia A including administering, to a patient in need thereof, a mammalian gene therapy vector as described above.

[0099] In another aspect, the disclosure provides a mammalian gene therapy vector as described above for treating hemophilia A.

[0100] In another aspect, the disclosure provides the use of a mammalian gene therapy vector as described above for the manufacture of a medicament for treating hemophilia A.

[0101] In another aspect, the disclosure provides a Factor VIII polypeptide including a light chain, a heavy chain, and a polypeptide linker joining the C-terminus of the heavy chain to the N-terminus of the light chain. The heavy chain of the Factor VIII polypeptide has a sequence at least 95% identical to the sequence CS01-HC-AAm23. The light chain of the Factor VIII polypeptide has a sequence at least 95% identical to the sequence CS01-LC-AAm23. The polypeptide linker of the Factor VIII polypeptide includes a furin cleavage site. The Factor VIII polypeptide includes i) I105V, A127S, G151K, M166T, and L171P amino acid substitutions, ii) a deletion of amino acids AIEPR755-759, relative to FVIII-FL-AA (SEQ ID NO: 19), and iii) an insertion of amino acids TTYVNRSR (SEQ ID NO: 33) after N754, relative to FVIII-FL-AA (SEQ ID NO: 19).

[0102] In another aspect, the disclosure provides a Factor VIII polypeptide including a light chain, a heavy chain, and a polypeptide linker joining the C-terminus of the heavy chain to the N-terminus of the light chain. The heavy chain of the Factor VIII polypeptide has a sequence at least 95% identical to the sequence CS01-HC-AAm123. The light chain of the Factor VIII polypeptide has a sequence at least 95% identical to the sequence CS01-LC-AAm123. The polypeptide linker of the Factor VIII polypeptide includes a furin cleavage site. The Factor VIII polypeptide includes i) I105V, A127S, G151K, M166T, and L171P amino acid substitutions, ii) a deletion of amino acids AIEPR755-759, relative to FVIII-FL-AA (SEQ ID NO: 19), iii) an insertion of amino acids TTYVNRSR (SEQ ID NO: 33) after N754, relative to FVIII-FL-AA (SEQ ID NO: 19), and iv) an F328S amino acid substitution.

[0103] In another aspect, the disclosure provides a Factor VIII polypeptide including a light chain, a heavy chain, and a polypeptide linker joining the C-terminus of the heavy chain to the N-terminus of the light chain. The heavy chain of the Factor VIII polypeptide has a sequence at least 95% identical to the sequence CS01-HC-AAm234. The light chain of the Factor VIII polypeptide has a sequence at least 95% identical to the sequence CS01-LC-AAm234. The polypeptide linker of the Factor VIII polypeptide includes a furin cleavage site. The Factor VIII polypeptide includes i) I105V, A127S, G151K, M166T, and L171P amino acid substitutions, ii) a deletion of amino acids AIEPR755-759, relative to FVIII-FL-AA (SEQ ID NO: 19), iii) an insertion of amino acids TTYVNRSR (SEQ ID NO: 33) after N754, relative to FVIII-FL-AA (SEQ ID NO: 19), and iv) F328S/C1918G/C1922G amino acid substitutions.

[0104] In one embodiment of the Factor VIII polypeptides described, the heavy chain of the Factor VIII polypeptide has a sequence at least 96% identical to the respective heavy chain sequence (e.g., CS01-HC-AAm23, CS01-HC-AAm123, or CS01-HC-AAm234), and the light chain of the Factor FVIII polypeptide has a sequence at least 96%

identical to the respective light chain sequence (e.g., CS01-LC-AAm23, CS01-LC-AAm123, or CS01-LC-AAm234).

[0105] In one embodiment of the Factor VIII polypeptides described, the heavy chain of the Factor VIII polypeptide has a sequence at least 97% identical to the respective heavy chain sequence (e.g., CS01-HC-AAm23, CS01-HC-AAm123, or CS01-HC-AAm234), and the light chain of the Factor FVIII polypeptide has a sequence at least 97% identical to the respective light chain sequence (e.g., CS01-LC-AAm23, CS01-LC-AAm123, or CS01-LC-AAm234).

[0106] In one embodiment of the Factor VIII polypeptides described, the heavy chain of the Factor VIII polypeptide has a sequence at least 98% identical to the respective heavy chain sequence (e.g., CS01-HC-AAm23, CS01-HC-AAm123, or CS01-HC-AAm234), and the light chain of the Factor FVIII polypeptide has a sequence at least 98% identical to the respective light chain sequence (e.g., CS01-LC-AAm23, CS01-LC-AAm123, or CS01-LC-AAm234).

[0107] In one embodiment of the Factor VIII polypeptides described, the heavy chain of the Factor VIII polypeptide has a sequence at least 99% identical to the respective heavy chain sequence (e.g., CS01-HC-AAm23, CS01-HC-AAm123, or CS01-HC-AAm234), and the light chain of the Factor FVIII polypeptide has a sequence at least 99% identical to the respective light chain sequence (e.g., CS01-LC-AAm23, CS01-LC-AAm123, or CS01-LC-AAm234).

[0108] In one embodiment of the Factor VIII polypeptides described, the heavy chain of the Factor VIII polypeptide has a sequence at least 99.5% identical to the respective heavy chain sequence (e.g., CS01-HC-AAm23, CS01-HC-AAm123, or CS01-HC-AAm234), and the light chain of the Factor FVIII polypeptide has a sequence at least 99.5% identical to the respective light chain sequence (e.g., CS01-LC-AAm23, CS01-LC-AAm123, or CS01-LC-AAm234).

[0109] In one embodiment of the Factor VIII polypeptides described, the heavy chain of the Factor VIII polypeptide has a sequence identical to the respective heavy chain sequence (e.g., CS01-HC-AAm23, CS01-HC-AAm123, or CS01-HC-AAm234), and the light chain of the Factor FVIII polypeptide has a sequence identical to the respective light chain sequence (e.g., CS01-LC-AAm23, CS01-LC-AAm123, or CS01-LC-AAm234).

[0110] In one embodiment of the Factor VIII polypeptides described above, the polypeptide linker has at least 95% identity to BDL-SQ-AA (SEQ ID NO: 30).

[0111] In one embodiment of the Factor VIII polypeptides described above, the polypeptide linker has the amino acid sequence of BDL-SQ-AA (SEQ ID NO: 30).

[0112] In one embodiment of the Factor VIII polypeptides described above, the polypeptide linker includes a glycosylation peptide with an amino acid sequence having at least 92% identity to a glycosylation peptide selected from the group consisting of NG1-AA, NG4-AA, NG5-AA, NG6-AA, NG7-AA, NG9-AA, NG10-AA, NG16-AA, NG17-AA, NG18-AA, NG19-AA, NG20-AA, NG21-AA and NGV-AA.

[0113] In one embodiment of the Factor VIII polypeptides described above, the polypeptide linker includes a glycosylation peptide selected from the group consisting of NG1-AA, NG4-AA, NG5-AA, NG6-AA, NG7-AA, NG9-AA, NG10-AA, NG16-AA, NG17-AA, NG18-AA, NG19-AA, NG20-AA, NG21-AA and NGV-AA.

[0114] In one embodiment of the Factor VIII polypeptides described above, the polypeptide linker has an amino acid

sequence having at least 95% identity to a sequence selected from the group consisting of BDLNG1-AA, BDLNG3-AA, BDLNGS-AA, BDLNG6-AA, BDLNG9-AA, BDLNG10-AA, BDLNG16-AA, BDLNG17-AA, BDLNG18-AA, BDLNG19-AA, BDLNG20-AA and BDLNG21-AA.

[0115] In one embodiment of the Factor VIII polypeptides described above, the polypeptide linker has an amino acid sequence selected from the group consisting of BDLNG1-AA, BDLNG3-AA, BDLNGS-AA, BDLNG6-AA, BDLNG9-AA, BDLNG10-NA, BDLNG16-AA, BDLNG17-AA, BDLNG18-AA, BDLNG19-AA, BDLNG20-AA and BDLNG21-AA.

[0116] In another aspect, the disclosure provides a Factor VIII polypeptide having an amino acid sequence with at least 95% identity to CS40-FL-AAm23 (SEQ ID NO: 104). The Factor VIII polypeptide includes i) I105V, A127S, G151K, M166T, and L171P amino acid substitutions, ii) a deletion of amino acids AIEPR755-759, relative to FVIII-FL-AA (SEQ ID NO: 19), and iii) an insertion of amino acids TTYVNRS (SEQ ID NO: 33) after N754, relative to FVIII-FL-AA (SEQ ID NO: 19).

[0117] In another aspect, the disclosure provides a Factor VIII polypeptide having an amino acid sequence with at least 95% identity to CS40-FL-AAm123. The Factor VIII polypeptide includes i) I105V, A127S, G151K, M166T, and L171P amino acid substitutions, ii) a deletion of amino acids AIEPR755-759, relative to FVIII-FL-AA (SEQ ID NO: 19), iii) an insertion of amino acids TTYVNRS (SEQ ID NO: 33) after N754, relative to FVIII-FL-AA (SEQ ID NO: 19), and iv) an F328S amino acid substitution.

[0118] In another aspect, the disclosure provides a Factor VIII polypeptide having an amino acid sequence with at least 95% identity to CS40-FL-AAm234. The Factor VIII polypeptide includes i) I105V, A127S, G151K, M166T, and L171P amino acid substitutions, ii) a deletion of amino acids AIEPR755-759, relative to FVIII-FL-AA (SEQ ID NO: 19), iii) an insertion of amino acids TTYVNRS (SEQ ID NO: 33) after N754, relative to FVIII-FL-AA (SEQ ID NO: 19), and iv) F328S/C1918G/C1922G amino acid substitutions.

[0119] In one embodiment of the Factor VIII polypeptides described, the Factor VIII polypeptide has a sequence at least 96% identical to the respective full-length sequence (e.g., CS40-FL-AAm23 (SEQ ID NO: 104), CS40-FL-AAm123, or CS40-FL-AAm234).

[0120] In one embodiment of the Factor VIII polypeptides described, the Factor VIII polypeptide has a sequence at least 97% identical to the respective full-length sequence (e.g., CS40-FL-AAm23 (SEQ ID NO: 104), CS40-FL-AAm123, or CS40-FL-AAm234).

[0121] In one embodiment of the Factor VIII polypeptides described, the Factor VIII polypeptide has a sequence at least 98% identical to the respective full-length sequence (e.g., CS40-FL-AAm23 (SEQ ID NO: 104), CS40-FL-AAm123, or CS40-FL-AAm234).

[0122] In one embodiment of the Factor VIII polypeptides described, the Factor VIII polypeptide has a sequence at least 99% identical to the respective full-length sequence (e.g., CS40-FL-AAm23 (SEQ ID NO: 104), CS40-FL-AAm123, or CS40-FL-AAm234).

[0123] In one embodiment of the Factor VIII polypeptides described, the Factor VIII polypeptide has a sequence at least 99.5% identical to the respective full-length sequence (e.g., CS40-FL-AAm23 (SEQ ID NO: 104), CS40-FL-AAm123, or CS40-FL-AAm234).

[0124] In one embodiment of the Factor VIII polypeptides described, the Factor VIII polypeptide has a sequence identical to the respective full-length sequence (e.g., CS40-FL-AAm23 (SEQ ID NO: 104), CS40-FL-AAm123, or CS40-FL-AAm234).

BRIEF DESCRIPTION OF DRAWINGS

[0125] FIG. 1 shows schematic illustrations of the wild-type and ReFacto-type human Factor VIII protein constructs.

[0126] FIGS. 2A and 2B show the CS04 codon-altered nucleotide sequence (SEQ ID NO: 1) encoding a Factor VIII variant in accordance with some embodiments (“CS04-FL-NA” for full-length coding sequence).

[0127] FIG. 3 shows the Factor VIII variant amino acid sequence (SEQ ID NO: 2) encoded by the CS04 codon-altered nucleotide sequence in accordance with some embodiments (“CS04-FL-AA” for full-length amino acid sequence).

[0128] FIG. 4 shows the portion of the CS04 codon-altered nucleotide sequence (SEQ ID NO: 3) encoding the heavy chain of a Factor VIII variant in accordance with some embodiments (“CS04-HC-NA”).

[0129] FIG. 5 shows the portion of the CS04 codon-altered nucleotide sequence (SEQ ID NO: 4) encoding the light chain of a Factor VIII variant in accordance with some embodiments (“CS04-LC-NA”).

[0130] FIG. 6 shows exemplary coding sequences (SEQ ID NOS 5-7 and 36-48, respectively, in order of appearance) for B-domain substituted linkers in accordance with some embodiments. BDLO01 (SEQ ID NO: 5), BDLO04 (SEQ ID NO: 6), and BDLO23 (SEQ ID NO: 7) are the respective portions of the CS01, CS04, and CS23 codon-altered nucleotide sequences that encode a B-domain substituted linker, respectively.

[0131] FIGS. 7A, 7B, and 7C show an AAV vector sequence (SEQ ID NO: 8) containing an CS04 codon-altered nucleotide sequence in accordance with some embodiments (“CS04-AV-NA”).

[0132] FIGS. 8A and 8B show the CS01m1 codon-altered nucleotide sequence (SEQ ID NO: 49) encoding a Factor VIII variant with an F328S amino acid substitution in accordance with some embodiments (“CS01m1-FL-NA”).

[0133] FIGS. 9A and 9B show the CS04Δ(760-1667) (SPI; CS04Δ(741-1648), SPE) codon-altered nucleotide sequence (SEQ ID NO: 9) encoding a single-chain Factor VIII variant in accordance with some embodiments (“CS04-SC1-NA”).

[0134] FIG. 10 shows the Factor VIII variant amino acid sequence (SEQ ID NO: 10) encoded by the CS01Δ(760-1667) (SPI; CS01Δ(741-1648), SPE), CS04Δ(760-1667) (SPI; CS04Δ(741-1648), SPE), and CS23Δ(760-1667) (SPI; CS23Δ(741-1648), SPE) codon-altered nucleotide sequences in accordance with some embodiments (“CS01-SC1-AA,” “CS04-SC1-AA,” and “CS23-SC1-AA,” respectively).

[0135] FIGS. 11A and 11B show the CS04Δ(772-1667) (SPI; CS04Δ(753-1648), SPE) codon-altered nucleotide sequence (SEQ ID NO: 11) encoding a single-chain Factor VIII variant in accordance with some embodiments (“CS04-SC2-NA”).

[0136] FIG. 12 shows the Factor VIII variant amino acid sequence (SEQ ID NO: 12) encoded by the CS01Δ(772-1667) (SPI; CS01Δ(753-1648), SPE), CS04Δ(772-1667) (SPI; CS04Δ(753-1648), SPE), and CS23Δ(772-1667) (SPI;

CS23Δ(753-1648), SPE) codon-altered nucleotide sequence in accordance with some embodiments (“CS01-SC2-AA,” “CS04-SC2-AA,” and “CS23-SC2-AA,” respectively).

[0137] FIGS. 13A and 13B show amino acid and nucleotide sequences for exemplary glycosylation peptides that are inserted into the B-domain substituted linker in accordance with some embodiments. “NG1” or NG1-AA” is the code for the amino acid sequence, shown in the top line. “NG1-NA” is the code for the nucleic acid sequence, shown in the bottom line for each set. FIGS. 13A and 13B disclose the amino acid sequences as SEQ ID NOS 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, and 75, and the nucleotide sequences as SEQ ID NOS 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, and 74, all respectively, in order of appearance.

[0138] FIG. 14 shows the results of in silico prediction of in vivo N-glycosylation of the wild-type Factor VIII B-domain. Figure discloses SEQ ID NOS 76 and 76-82, respectively, in order of appearance.

[0139] FIG. 15 shows the results of in silico prediction of in vivo N-glycosylation of the V3 peptide linker. Figure discloses SEQ ID NOS 83 and 83-89, respectively, in order of appearance.

[0140] FIGS. 16A and 16B show the CS01 codon-altered nucleotide sequence (SEQ ID NO: 13) encoding a Factor VIII variant in accordance with some embodiments (“CS01-FL-NA”).

[0141] FIGS. 17A and 17B show the CS08 codon-altered nucleotide sequence (SEQ ID NO: 14) encoding a Factor VIII variant in accordance with some embodiments (“CS08-FL-NA”).

[0142] FIGS. 18A and 18B show the CS10 codon-altered nucleotide sequence (SEQ ID NO: 15) encoding a Factor VIII variant in accordance with some embodiments (“CS10-FL-NA”).

[0143] FIGS. 19A and 19B show the CS11 codon-altered nucleotide sequence (SEQ ID NO: 16) encoding a Factor VIII variant in accordance with some embodiments (“CS11-FL-NA”).

[0144] FIGS. 20A and 20B show the CS40 wild-type ReFacto coding sequence (SEQ ID NO: 17), in accordance with some embodiments (“CS40-FL-NA”).

[0145] FIGS. 21A and 21B show the CH25 codon-altered nucleotide sequence (SEQ ID NO: 18) encoding a Factor VIII variant in accordance with some embodiments (“CH25-FL-NA”).

[0146] FIG. 22 shows a wild-type human Factor VIII amino acid sequence (SEQ ID NO: 19), in accordance with some embodiments (“FVIII-FL-AA”).

[0147] FIG. 23 illustrates the scheme for cloning the pCS40, pCS01, pCS04, pCS08, pCS10, pCS11, and pCh25 constructs, by inserting synthetic Refacto-type BDD-FVIII DNA sequences into the vector backbone pCh-BB01 via Ascl and NotI restriction sites.

[0148] FIG. 24 shows the integrity of AAV vector genome preparations, as analyzed by agarose gel electrophoresis. Lane 1, DNA marker; lane 2, vCS40; lane 3, vCS01; lane 4, vCS04. The AAV vectors have all the same-sized genomes, migrating at approximately 5 kb (arrow, right side). The scale on the left side indicates size of the DNA fragments in kilobases (kb).

[0149] FIG. 25 shows the protein analysis of AAV vector preparations by PAGE and silver staining. Lane 1, protein marker (M); lane 2, vCS40, lane 3, vCS01; and lane 4, vCS04. The constructs all have the same AAV8 capsids

consisting of VP1, VP2, and VP3 (arrows right side). The scale on the left side indicates size of the protein marker in kilodaltons (kDa).

[0150] FIGS. 26A and 26B show the CS23 codon-altered nucleotide sequence (SEQ ID NO: 20) encoding a Factor VIII variant in accordance with some embodiments (“CS23-FL-NA”).

[0151] FIG. 27 shows the Factor VIII variant amino acid sequence (SEQ ID NO: 21) encoded by the CS23 codon-altered nucleotide sequence in accordance with some embodiments (“CS23-FL-AA”).

[0152] FIG. 28 shows the portion of the CS23 codon-altered nucleotide sequence (SEQ ID NO: 22) encoding the heavy chain of a Factor VIII variant in accordance with some embodiments (“CS23-HC-NA”).

[0153] FIG. 29 shows the portion of the CS23 codon-altered nucleotide sequence (SEQ ID NO: 23) encoding the light chain of a Factor VIII variant in accordance with some embodiments (“CS23-LC-NA”).

[0154] FIGS. 30A and 30B show the CS01m13 codon-altered nucleotide sequence (SEQ ID NO: 90) encoding a Factor VIII variant with m1 (F328S) and m3 amino acid substitutions in accordance with some embodiments (“CS01-FL-NA-m13”).

[0155] FIGS. 31A and 31B show the CS01m23 codon-altered nucleotide sequence (SEQ ID NO: 91) encoding a Factor VIII variant with the m2 and m3 mutation sets in accordance with some embodiments (“CS01-FL-NA-m23”).

[0156] FIGS. 32A and 32B show the CS01m3 codon-altered nucleotide sequence (SEQ ID NO: 92) encoding a Factor VIII variant with m3 amino acid substitutions in accordance with some embodiments (“CS01-FL-NA-m3”).

[0157] FIGS. 33A and 33B show the CS01m2 codon-altered nucleotide sequence (SEQ ID NO: 93) encoding a Factor VIII variant with the m2 mutation set (I105V/A127S/G151K/M166T/L171P (SPI)) amino acid substitutions in accordance with some embodiments (“CS01-FL-NA-m2”).

[0158] FIGS. 34A and 34B show the CS04m2 codon-altered nucleotide sequence (SEQ ID NO: 94) encoding a Factor VIII variant with the m2 mutants (I105V/A127S/G151K/M166T/L171P (SPI)) amino acid substitutions in accordance with some embodiments (“CS01-FL-NA-m2”).

[0159] FIGS. 35A and 35B show the CS04m3 codon-altered nucleotide sequence (SEQ ID NO: 95) encoding a Factor VIII variant with m3 amino acid substitutions in accordance with some embodiments (“CS04-FL-NA-m3”).

[0160] FIGS. 36A and 36B show the CS04m23 codon-altered nucleotide sequence (SEQ ID NO: 96) encoding a Factor VIII variant with the m2 mutant set (I105V/A127S/G151K/M166T/L171P (SPI)) and m3 amino acid substitutions in accordance with some embodiments (“CS04-FL-NA-m23”).

[0161] FIGS. 37A and 37B show the CS04m1 codon-altered nucleotide sequence (SEQ ID NO: 97) encoding a Factor VIII variant with an m1 (F328S) amino acid substitution in accordance with some embodiments (“CS04-FL-NA-m1”).

[0162] FIGS. 38A and 38B show the CS04m13 codon-altered nucleotide sequence (SEQ ID NO: 98) encoding a Factor VIII variant with m1 and m3 amino acid substitutions in accordance with some embodiments (“CS04-FL-NA-m13”)

[0163] FIGS. 39A and 39B show the CS23m13 codon-altered nucleotide sequence (SEQ ID NO: 99) encoding a

Factor VIII variant with m1 and m3 amino acid substitutions in accordance with some embodiments (“CS23m13-FL-NA”)

[0164] FIGS. 40A and 40B show the CS23m3 codon-altered nucleotide sequence (SEQ ID NO: 100) encoding a Factor VIII variant with m3 amino acid substitutions in accordance with some embodiments (“CS23-FL-NA-m3”)

[0165] FIGS. 41A and 41B show the CS23m2 codon-altered nucleotide sequence (SEQ ID NO: 101) encoding a Factor VIII variant with the m2 mutant set (I105V/A127S/G151K/M166T/L171P amino acid substitutions) in accordance with some embodiments (“CS23-FL-NA-m2”).

[0166] FIGS. 42A and 42B show the CS23m1 codon-altered nucleotide sequence (SEQ ID NO: 102) encoding a Factor VIII variant with an m1 (F328S) amino acid substitution in accordance with some embodiments (“CS23-FL-NA-m1”).

[0167] FIGS. 43A and 43B show the CS23m23 codon-altered nucleotide sequence (SEQ ID NO: 103) encoding a Factor VIII variant with the m2 mutant set (I105V/A127S/G151K/M166T/L171P) and m3 amino acid substitutions in accordance with some embodiments (“CS23-FL-NA-m23”).

[0168] FIG. 44 depicts cloning of the pCS constructs, done by inserting synthetic Refacto-type BDD-FVIII carrying different mutations (see inserted table) into the vector backbone pCh-BB01 via AscI and NotI restriction sites.

[0169] FIG. 45 depicts the protein analysis of AAV vector preparations by PAGE and silver staining. Lane 1, protein marker (M); lane 2, vCS01; lane 3, vCS17; lane 4, vCS19; lane 5, vCS20; lane 6, vCS40; lane 7, vCS04; lane 8, vCS17; lane 9, vCS24 construct. The constructs have all the same AAV8 capsids consisting of VP1, VP2 and VP3 (arrows right side). The scale on the left side indicates size of the protein marker in kilo Daltons (kDa).

[0170] FIG. 46 depicts the integrity of AAV vector genome preparations analyzed by agarose gel electrophoresis. Lane 1, DNA marker (M); lane 2, vCS04, lane 3, vCS17; lane 4, vCS20; lane 5, vCS24; lane 6, vCS16; lane 7, vCS40 construct. Vector load is 1.5E10 vg per lane. The AAV vectors have the same-sized genomes, migrating at approximately 5 kb (arrow, right side). The scale on the left side indicates size of the DNA fragments in kilobases (kb).

[0171] FIG. 47 shows the portion of the CS01 codon-altered nucleotide sequence (SEQ ID NO: 24) encoding the heavy chain of a Factor VIII variant in accordance with some embodiments (“CS01-HC-NA”).

[0172] FIG. 48 shows the portion of the CS01 codon-altered nucleotide sequence (SEQ ID NO: 25) encoding the light chain of a Factor VIII variant in accordance with some embodiments (“CS01-LC-NA”).

[0173] FIGS. 49A and 49B show the CS01Δ(760-1667) (SPI; CS01Δ(741-1648), SPE) codon-altered nucleotide sequence (SEQ ID NO: 26) encoding a single-chain Factor VIII variant in accordance with some embodiments (“CS01-SC1-NA”).

[0174] FIGS. 50A and 50B show the CS01Δ(772-1667) (SPI; CS01Δ(753-1648), SPE) codon-altered nucleotide sequence (SEQ ID NO: 27) encoding a single-chain Factor VIII variant in accordance with some embodiments (“CS01-SC2-NA”).

[0175] FIGS. 51A and 51B show the CS23Δ(760-1667) (SPI; CS23Δ(741-1648), SPE) codon-altered nucleotide

sequence (SEQ ID NO: 28) encoding a single-chain Factor VIII variant in accordance with some embodiments (“CS23-SC1-NA”).

[0176] FIGS. 52A and 52B show the CS23Δ(772-1667) (SPI; CS23Δ(753-1648), SPE) codon-altered nucleotide sequence (SEQ ID NO: 29) encoding a single-chain Factor VIII variant in accordance with some embodiments (“CS23-SC2-NA”).

[0177] FIG. 53 shows the Factor VIII variant amino acid sequence (SEQ ID NO: 104) encoded by the CS01m23 codon-altered nucleotide sequence in accordance with some embodiments (“CS01m23-FL-AA”).

[0178] FIG. 54 shows the Factor VIII variant amino acid sequence (SEQ ID NO: 105) encoded by the CS04m3 codon-altered nucleotide sequence in accordance with some embodiments (“CS01m23-FL-AA”).

[0179] FIG. 55 shows the Factor VIII variant amino acid sequence (SEQ ID NO: 106) encoded by the CS01m12 codon-altered nucleotide sequence in accordance with some embodiments (“CS01m12-FL-AA”).

[0180] FIG. 56 shows the Factor VIII variant amino acid sequence (SEQ ID NO: 107) encoded by the CS04m12 codon-altered nucleotide sequence in accordance with some embodiments (“CS04m12-FL-AA”).

[0181] FIGS. 57A and 57B show the CS01m12 codon-altered nucleotide sequence (SEQ ID NO: 108) encoding a Factor VIII variant with m1 (F328S) and m2 amino acid substitutions in accordance with some embodiments (“CS01-FL-NAm12”).

[0182] FIGS. 58A and 58B show the CS04m12 codon-altered nucleotide sequence (SEQ ID NO: 109) encoding a Factor VIII variant with m1 (F328S) and m2 amino acid substitutions in accordance with some embodiments (“CS04-FL-NAm12”).

DETAILED DESCRIPTION OF DISCLOSURE

I. Introduction

[0183] AAV-based gene therapy holds great promise for the treatment of hemophiliacs. For hemophilia B, first clinical data are encouraging in that FIX levels of about 10% can be maintained in at least some patients for more than 1 year. For hemophilia A however, achieving therapeutic expression levels of 5-10% with AAV vectors remains challenging for various reasons. First, the Factor VIII coding sequence is too large for conventional AAV-based vectors. Second, engineered B-domain deleted or truncated Factor VIII constructs suffer from poor expression in vivo, even when codon-optimized. Third, these B-domain deleted or truncated Factor VIII variant constructs have short half-lives in vivo, exacerbating the effects of poor expression. Fourth, even when expressed, FVIII is not efficiently secreted from cells, as are other coagulation factors, such as Factor IX.

[0184] Moreover, these challenges cannot be addressed by simply administering higher doses of the gene therapy construct. According to current knowledge, the vector dose of an AAV-based gene therapy vector should be increased above 2×10^2 vg/kg bodyweight. This is because at such high doses a T cell immune response is triggered, which destroys transduced cells and, as a consequence, transgene expression is reduced or even eliminated. Therefore, strategies to improve the expression of FVIII are needed to make FVIII gene therapy a viable therapeutic option for hemophilia A patients.

[0185] The present disclosure relates to the discovery of codon-altered Factor VIII variant coding sequences that solve these and other problems associated with Factor VIII gene therapy. For example, the polynucleotides disclosed herein provide markedly improved expression in mammalian cells, and display improved virion packaging due to stabilized packing interactions. In some implementations, these advantages are realized by using coding sequences for the heavy and light chains of Factor VIII with high sequence identity to the codon altered CS01, CS04, and CS23 constructs (e.g., with high sequence identity to one of the CS01-HC, CS04-HC, and CS23-HC heavy chain coding sequences and high sequence identity to one of the CS01-LC, CS04-LC, and CS23-LC light chain coding sequences).

[0186] In some implementations, the Factor VIII molecules encoded by the polynucleotides described herein have been shortened by truncating, deleting, or replacing the wild-type B-domain. As such, the polynucleotides are better suited for expressing Factor VIII via conventional gene therapy vectors, which inefficiently express larger polypeptides, such as the wild-type Factor VIII.

[0187] Advantageously, it is shown herein that the CS01, CS04, and CS23 codon-altered Factor VIII variant coding sequences provide superior expression of a B-domain deleted Factor VIII construct in vivo. For example, it is demonstrated in Example 2 and Example 4 that intravenous administration of AAV-based gene therapy vectors having the CS01 (SEQ ID NO: 13), CS04 (SEQ ID NO: 1), and CS23 (SEQ ID NO: 20) coding sequence provide 18-fold, 74-fold, and 30-fold increases in Factor VIII expression, relative to the corresponding CS40 construct encoded with the wild-type polynucleotide sequence (SEQ ID NO: 17), in Factor VIII knock-out mice (Table 4 and Table 7).

[0188] Further, it is also shown herein that the CS01 and CS04 codon-altered Factor VIII variant coding sequences provide superior virion packaging and virus production. For example, it is demonstrated in Example 1 that AAV vector constructs containing the CS01 and CS04 constructs provided 5 to 7-fold greater viral yield, relative to the corresponding CS40 construct encoded with the wild-type polynucleotide sequence, when isolated from the same amount of cell pellet.

[0189] Advantageously, Applicants also found that the improved Factor VIII activity generated from the CS01, CS04, and CS23 codon altered sequences could be further enhanced by introducing mutations into the underlying Factor VIII polypeptide sequence. For example, as demonstrated in Example 4, the F328S, X5, and X1 mutations, alone and in combination with one another, further increased FVIII activity when expressed in vivo in the CS01 or CS04 codon altered background 2 to 7-fold, relative to the wild type, codon altered constructs (Table 7). More strikingly, these codon altered sequences, encoding the mutant Factor VIII mutants, provided up to 246-fold greater increase as compared to the corresponding CS40 construct encoded with the wild-type polynucleotide sequence (Table 7).

II. Definitions

[0190] As used herein, the following terms have the meanings ascribed to them unless specified otherwise.

[0191] As used herein, the terms “Factor VIII” and “FVIII” are used interchangeably, and refer to any protein with Factor VIII activity (e.g., active FVIII, often referred to as FVIIIa) or protein precursor (e.g., pro-protein or pre-pro-

protein) of a protein with Factor VIII activity, particularly Factor IXa cofactor activity. In an exemplary embodiment, a Factor VIII polypeptide refers to a polypeptide that has sequences with high sequence identity (e.g., at least 70%, 75%, 80%, 85%, 90%, 95%, 99%, or more) to the heavy and light chains of a wild type Factor VIII polypeptide. In some embodiments, the B-domain of a Factor VIII polypeptide is deleted, truncated, or replaced with a linker polypeptide to reduce the size of the polynucleotide encoding the Factor VIII polypeptide. In an exemplary embodiment, amino acids 20-1457 of SEQ ID NO: 2 constitute a Factor VIII polypeptide.

[0192] Non-limiting examples of wild type Factor VIII polypeptides include human pre-pro-Factor VIII (e.g., GenBank accession nos. AAA52485, CAA25619, AAA58466, AAA52484, AAA52420, AAV85964, BAF82636, BAG36452, CAI41660, CAI41666, CAI41672, CAI43241, CA003404, EAW72645, AAH22513, AAH64380, AAH98389, AAI11968, AAI11970, or AAB61261), corresponding pro-Factor VIII, and natural variants thereof; porcine pre-pro-Factor VIII (e.g., UniProt accession nos. F1RZ36 or K7GSZ5), corresponding pro-Factor VIII, and natural variants thereof; mouse pre-pro-Factor VIII (e.g., GenBank accession nos. AAA37385, CAM15581, CAM26492, or EDL29229), corresponding pro-Factor VIII, and natural variants thereof; rat pre-pro-Factor VIII (e.g., GenBank accession no. AAQ21580), corresponding pro-Factor VIII, and natural variants thereof; rat pre-pro-Factor VIII; and other mammalian Factor VIII homologues (e.g., monkey, ape, hamster, guinea pig, etc.).

[0193] As used herein, a Factor VIII polypeptide includes natural variants and artificial constructs with Factor IX cofactor activity. As used in the present disclosure, Factor VIII encompasses any natural variants, alternative sequences, isoforms, or mutant proteins that retain some basal Factor IX cofactor activity (e.g., at least 5%, 10%, 25%, 50%, 75%, or more of the corresponding wild type activity). Examples of Factor VIII amino acid variations (relative to FVIII-FL-AA (SEQ ID NO: 19)) found in the human population include, without limitation, S19R, R22T, Y24C, Y25C, L26P/R, E30V, W33G, Y35C/H, G41C, R48C/K, K67E/N, L69P, E72K, D75E/V/Y, P83R, G89D/V, G92A/V, A97P, E98K, V99D, D101G/H/V, V104D, K108T, M110V, A111T/V, H113R/Y, L117F/R, G121S, E129V, G130R, E132D, Y133C, D135G/Y, T137A/I, S138R, E141K, D145H, V147D, Y155H, V159A, N163K, G164D/V, P165S, C172W, S176P, S179P, V181E/M, K185T, D186G/N/Y, S189L, L191F, G193R, L195P, C198G, S202N/R, F214V, L217H, A219D/T, V220G, D222V, E223K, G224W, T252I, V253F, N254I, G255V, L261P, P262L, G263S, G266F, C267Y, W274C, H275L, G278R, G280D, E284K, V285G, E291G/K, T294I, F295L, V297A, N299I, R301C/H/L, A303E/P, I307S, S308L, F312S, T314A/I, A315V, G323E, L326P, L327P/V, C329F, I331V, M339T, E340K, V345A/L, C348R/S/Y, Y365C, R391C/H/P, S392L/P, A394S, W401G, I405F/S, E409G, W412G/R, K427I, L431F/S, R437P/W, I438F, G439D/S/V, Y442C, K444R, Y450D/N, T454I, F455C, G466E, P470L/R/T, G474E/R/V, E475K, G477V, D478N, T479R, F484C, A488G, R490G, Y492C/H, Y492H, I494T, P496R, G498R, R503H, G513S/V, I522Y, K529E, W532G, P540T, T541S, D544N, R546W, R550C/G/H, S553P, S554C/G, V556D, R560T, D561G/H/Y, I567T, P569R, S577F, V578A, D579A/H, N583S, Q584H/K/R, I585R/T, M586V, D588G/Y,

L594Q, S596P, N601D/K, R602G, S603I/R, W604C, Y605H/S, N609I, R612C, N631K/S, M633I, S635N, N637D/I/S, Y639C, L644V, L650F, V653A/M, L659P, A663V, Q664P, F677L, M681I, V682F, Y683C/N, T686R, F698L, M699T/V, M701I, G705V, G710W, N713I, R717L/W, G720D/S, M721I/L, A723T, L725Q, V727F, E739K, Y742C, R795G, P947R, V1012L, E1057K, H1066Y, D1260E, K1289Q, Q1336K, N1460K, L1481P, A1610S, I1698T, Y1699C/F, E1701K, Q1705H, R1708C/H, T1714S, R1715G, A1720V, E1723K, D1727V, Y1728C, R1740G, K1751Q, F1762L, R1768H, G1769R, L1771P, L1775F/V, L1777P, G1779E/R, P1780L, I1782R, D1788H, M1791T, A1798P, S1799H, R1800C/G/H, P1801A, Y1802C, S1803Y, F1804S, L1808F, M1842I, P1844S, T1845P, E1848G, A1853T/V, S1858C, K1864E, D1865N/Y, H1867P/R, G1869D/V, G1872E, P1873R, L1875P, V1876L, C1877R/Y, L1882P, R1888I, E1894G, I1901F, E1904D/K, S1907C/R, W1908L, Y1909C, A1939T/V, N1941D/S, G1942A, M1945V, L1951F, R1960L/Q, L1963P, S1965I, M1966I/V, G1967D, S1968R, N1971T, H1973L, G1979V, H1980P/Y, F1982I, R1985Q, L1994P, Y1998C, G2000A, T2004R, M2007I, G2013R, W2015C, R2016P/W, E2018G, G2022D, G2028R, S2030N, V2035A, Y2036C, N2038S, 2040Y, G2045E/V, L2051S, I2056N, A2058P, W2065R, P2067L, A2070V, S2082N, S2088F, D2093G/Y, H2101D, T2105N, Q2106E/P/R, G2107S, R2109C, 12117F/S, Q2119R, F2120C/L, Y2124C, R2135P, S2138Y, T2141N, M2143V, F2145C, N2148S, N2157D, P2162L, R2169C/H, P2172L/Q/R, T2173A/I, H2174D, R2178C/H/L, R2182C/H/P, M2183R/V, L2185S/W, S2192I, C2193G, P2196R, G2198V, E2200D, 12204T, 12209N, A2211P, A2220P, P2224L, R2228G/L/P/Q, L2229F, V2242M, W2248C/S, V2251A/E, M2257V, T2264A, Q2265R, F2279C/I, I2281T, D2286G, W2290L, G2304V, D2307A, P2319L/S, R2323C/G/H/L, R2326G/L/P/Q, Q2330P, W2332R, I2336F, R2339T, G2344C/D/S, and C2345S/Y. Factor VIII proteins also include polypeptides containing post-translational modifications.

[0194] Generally, polynucleotides encoding Factor VIII encode for an inactive single-chain polypeptide (e.g., a pre-pro-protein) that undergoes post-translational processing to form an active Factor VIII protein (e.g., FVIIIa). For example, referring to FIG. 1, the wild type human Factor VIII pre-pro-protein is first cleaved to release the encoded signal peptide (not shown), forming a first single-chain pro-protein (shown as “human wild-type FVIII”). The pro-protein is then cleaved between the B and A3 domains to form a first polypeptide that includes the Factor VIII heavy chain (e.g., the A1 and A2 domains) and B-domain, and a second polypeptide that includes the Factor VIII light chain (e.g., including the A3, C1, and C3 domains). The first polypeptide is further cleaved to remove the B-domain, and also to separate the A1 and A2 domains, which remain associated with the Factor VIII light chain in the mature Factor VIIIa protein. For review of the Factor VIII maturation process, see Graw et al., Nat Rev Genet., 6(6):488-501 (2005), the content of which is incorporated herein by reference in its entirety for all purposes.

[0195] However, in some embodiments, the Factor VIII polypeptide is a single-chain Factor VIII polypeptide. Single-chain Factor VIII polypeptides are engineered to remove natural cleavage sites, and optionally remove, truncate, or replace the B-domain of Factor VIII. As such, they are not matured by cleavage (other than cleavage of an

optional signal and/or leader peptide), and are active as a single chain. Non-limiting examples of single-chain Factor VIII polypeptides are described in Zollner et al. (Thromb Res, 134(1):125-31 (2014)) and Donath et al. (Biochem J., 312(1):49-55 (1995)), the disclosures of which are hereby incorporated by reference in their entireties for all purposes.

[0196] As used herein, the terms “Factor VIII heavy chain,” or simply “heavy chain,” refers to the aggregate of the A1 and A2 domains of a Factor VIII polypeptide. In an exemplary embodiment, amino acids 20-759 of CS04-FL-AA (SEQ ID NO: 2) constitute a Factor VIII heavy chain.

[0197] As used herein, the term “Factor VIII light chain,” or simply “light chain,” refers to the aggregate of the A3, C1, and C2 domains of a Factor VIII polypeptide. In an exemplary embodiment, amino acids 774-1457 CS04-FL-AA (SEQ ID NO: 2) constitute a Factor VIII light chain. In some embodiments, a Factor VIII light chain excludes the acidic a3 peptide, which is released during maturation in vivo.

[0198] Generally, Factor VIII heavy and light chains are expressed as a single polypeptide chain, e.g., along with an optional B-domain or B-domain substituted linker. However, in some embodiments, a Factor VIII heavy chain and Factor VIII light chain are expressed as separate polypeptide chains (e.g., co-expressed), and reconstituted to form a Factor VIII protein (e.g., in vivo or in vitro).

[0199] As used herein, the terms “B-domain substituted linker” and “Factor VIII linker” are used interchangeably, and refer to truncated versions of a wild type Factor VIII B-domain (e.g., amino acids 760-1667 of FVIII-FL-AA (SEQ ID NO: 19)) or peptides engineered to replace the B-domain of a Factor VIII polypeptide. As used herein, a Factor VIII linker is positioned between the C-terminus of a Factor VIII heavy chain and the N-terminus of a Factor VIII light chain in a Factor VIII variant polypeptide in accordance with some embodiments. Non-limiting examples of B-domain substituted linkers are disclosed in U.S. Pat. Nos. 4,868,112, 5,112,950, 5,171,844, 5,543,502, 5,595,886, 5,610,278, 5,789,203, 5,972,885, 6,048,720, 6,060,447, 6,114,148, 6,228,620, 6,316,226, 6,346,513, 6,458,563, 6,924,365, 7,041,635, and 7,943,374; U.S. Patent Application Publication Nos. 2013/024960, 2015/0071883, and 2015/0158930; and PCT Publication Nos. WO 2014/064277 and WO 2014/127215, the disclosures of which are hereby incorporated by reference, in their entireties, for all purposes.

[0200] Unless otherwise specified herein, the numbering of Factor VIII amino acids refers to the corresponding amino acid in the full-length, wild-type human Factor VIII sequence (FVIII-FL-AA), presented as SEQ ID NO: 19 in FIG. 22. As such, when referring to an amino acid substitution in a Factor VIII variant protein disclosed herein, the recited amino acid number refers to the analogous (e.g., structurally or functionally equivalent) and/or homologous (e.g., evolutionarily conserved in the primary amino acid sequence) amino acid in the full-length, wild-type Factor VIII sequence. For example, a T2105N amino acid substitution refers to a T to N substitution at position 2105 of the full-length, wild-type human Factor VIII sequence (FVIII-FL-AA; SEQ ID NO: 19), a T to N substitution at position 1211 of the Factor VIII variant protein encoded by CS04 (CS04-FL-AA; SEQ ID NO: 2), and a T to N substitution at position 1212 of the Factor VIII variant encoded by CS04m3 (CS04m3-FL-AA; SEQ ID NO: 105).

[0201] As described herein, the Factor VIII amino acid numbering system is dependent on whether the Factor VIII signal peptide (e.g., amino acids 1-19 of the full-length, wild-type human Factor VIII sequence) is included. Where the signal peptide is included, the numbering is referred to as “signal peptide inclusive” or “SPI”. Where the signal peptide is not included, the numbering is referred to as “signal peptide exclusive” or “SPE.” For example, F328S is SPI numbering for the same amino acid as F3095, in SPE numbering. Unless otherwise indicated, all amino acid numbering refers to the corresponding amino acid in the full-length, wild-type human Factor VIII sequence (FVIII-FL-AA), presented as SEQ ID NO: 19 in FIG. 22.

[0202] As described herein, the codon-altered polynucleotides provide increased expression of transgenic Factor VIII in vivo (e.g., when administered as part of a gene therapy vector), as compared to the level of Factor VIII expression provided by a natively-coded Factor VIII construct (e.g., a polynucleotide encoding the same Factor VIII construct using the wild-type human codons). As used herein, the term “increased expression” refers to an increased level of transgenic Factor VIII activity in the blood of an animal administered the codon-altered polynucleotide encoding Factor VIII, as compared to the level of transgenic Factor VIII activity in the blood of an animal administered a natively-coded Factor VIII construct. The activity levels can be measured using any Factor VIII activity known in the art. An exemplary assay for determining Factor VIII activity is the Technochrome FVIII assay (Technoclone, Vienna, Austria).

[0203] In some embodiments, increased expression refers to at least 25% greater transgenic Factor VIII activity in the blood of an animal administered the codon-altered Factor VIII polynucleotide, as compared to the level of transgenic Factor VIII activity in the blood of an animal administered a natively coded Factor VIII polynucleotide. In some embodiments, increased expression refers to at least 50% greater, at least 75% greater, at least 100% greater, at least 3-fold greater, at least 4-fold greater, at least 5-fold greater, at least 6-fold greater, at least 7-fold greater, at least 8-fold greater, at least 9-fold greater, at least 10-fold greater, at least 15-fold greater, at least 20-fold greater, at least 25-fold greater, at least 30-fold greater, at least 40-fold greater, at least 50-fold greater, at least 60-fold greater, at least 70-fold greater, at least 80-fold greater, at least 90-fold greater, at least 100-fold greater, at least 125-fold greater, at least 150-fold greater, at least 175-fold greater, at least 200-fold greater, at least 225-fold greater, or at least 250-fold greater transgenic Factor VIII activity in the blood of an animal administered the codon-altered Factor VIII polynucleotide, as compared to the level of transgenic Factor VIII activity in the blood of an animal administered a natively coded Factor VIII polynucleotide.

[0204] As described herein, the codon-altered polynucleotides provide increased vector production, as compared to the level of vector production provided by a natively-coded Factor VIII construct (e.g., a polynucleotide encoding the same Factor VIII construct using the wild-type human codons). As used herein, the term “increased virus production” refers to an increased vector yield in cell culture (e.g., titer per liter culture) inoculated with the codon-altered polynucleotide encoding Factor VIII, as compared to the vector yield in cell culture inoculated with a natively-coded Factor VIII construct. The vector yields can be measured

using any vector titer assay known in the art. An exemplary assay for determining vector yield (e.g., of an AAV vector) is qPCR targeting the AAV2 inverted terminal repeats (Aurnhammer, Human Gene Therapy Methods: Part B 23:18-28 (2012)).

[0205] In some embodiments, increased virus production refers to at least 25% greater codon-altered vector yield, as compared to the yield of a natively-coded Factor VIII construct in the same type of culture. In some embodiments, increased vector production refers to at least 50% greater, at least 75% greater, at least 100% greater, at least 3-fold greater, at least 4-fold greater, at least 5-fold greater, at least 6-fold greater, at least 7-fold greater, at least 8-fold greater, at least 9-fold greater, at least 10-fold greater, at least 15-fold greater, or at least 20-fold greater codon-altered vector yield, as compared to the yield of a natively-coded Factor VIII construct in the same type of culture.

[0206] As used herein, the term "hemophilia" refers to a group of disease states broadly characterized by reduced blood clotting or coagulation. Hemophilia may refer to Type A, Type B, or Type C hemophilia, or to the composite of all three diseases types. Type A hemophilia (hemophilia A) is caused by a reduction or loss of factor VIII (FVIII) activity and is the most prominent of the hemophilia subtypes. Type B hemophilia (hemophilia B) results from the loss or reduction of factor IX (FIX) clotting function. Type C hemophilia (hemophilia C) is a consequence of the loss or reduction in factor XI (FXI) clotting activity. Hemophilia A and B are X-linked diseases, while hemophilia C is autosomal. Conventional treatments for hemophilia include both prophylactic and on-demand administration of clotting factors, such as FVIII, FIX, including Bebulin®-VH, and FXI, as well as FEIBA-VH, desmopressin, and plasma infusions.

[0207] As used herein, the term "FVIII gene therapy" includes any therapeutic approach of providing a nucleic acid encoding Factor VIII to a patient to relieve, diminish, or prevent the reoccurrence of one or more symptoms (e.g., clinical factors) associated with hemophilia. The term encompasses administering any compound, drug, procedure, or regimen comprising a nucleic acid encoding a Factor VIII molecule, including any modified form of Factor VIII (e.g., Factor VIII variant), for maintaining or improving the health of an individual with hemophilia. One skilled in the art will appreciate that either the course of FVIII therapy or the dose of a FVIII therapeutic agent can be changed, e.g., based upon the results obtained in accordance with the present disclosure.

[0208] As used herein, the term "bypass therapy" includes any therapeutic approach of providing non-Factor VIII hemostatic agents, compounds or coagulation factors to a patient to relieve, diminish, or prevent the reoccurrence of one or more symptoms (e.g., clinical factors) associated with hemophilia. Non-Factor VIII compounds and coagulation factors include, but are not limited to, Factor VIII Inhibitor Bypass Activity (FEIBA), recombinant activated factor VII (FVIIa), prothrombin complex concentrates, and activated prothrombin complex concentrates. These non-Factor VIII compounds and coagulation factors may be recombinant or plasma-derived. One skilled in the art will appreciate that either the course of bypass therapy or the dose of bypass therapy can be changed, e.g., based upon the results obtained in accordance with the present disclosure.

[0209] As used herein, a "combination therapy" including administration of a nucleic acid encoding a Factor VIII

molecule and a conventional hemophilia A therapeutic agent includes any therapeutic approach of providing both a nucleic acid encoding a Factor VIII molecule and a Factor VIII molecule and/or non-Factor VIII hemostatic agent (e.g., bypass therapeutic agent) to a patient to relieve, diminish, or prevent the reoccurrence of one or more symptoms (e.g., clinical factors) associated with hemophilia. The term encompasses administering any compound, drug, procedure, or regimen including a nucleic acid encoding a Factor VIII molecule, including any modified form of factor VIII, which is useful for maintaining or improving the health of an individual with hemophilia and includes any of the therapeutic agents described herein.

[0210] The terms "therapeutically effective amount or dose" or "therapeutically sufficient amount or dose" or "effective or sufficient amount or dose" refer to a dose that produces therapeutic effects for which it is administered. For example, a therapeutically effective amount of a drug useful for treating hemophilia can be the amount that is capable of preventing or relieving one or more symptoms associated with hemophilia. The exact dose will depend on the purpose of the treatment, and will be ascertainable by one skilled in the art using known techniques (see, e.g., Lieberman, *Pharmaceutical Dosage Forms* (vols. 1-3, 1992); Lloyd, *The Art, Science and Technology of Pharmaceutical Compounding* (1999); Pickar, *Dosage Calculations* (1999); and Remington: *The Science and Practice of Pharmacy*, 20th Edition, 2003, Gennaro, Ed., Lippincott, Williams & Wilkins).

[0211] As used herein, the term "gene" refers to the segment of a DNA molecule that codes for a polypeptide chain (e.g., the coding region). In some embodiments, a gene is positioned by regions immediately preceding, following, and/or intervening the coding region that are involved in producing the polypeptide chain (e.g., regulatory elements such as a promoter, enhancer, polyadenylation sequence, 5'-untranslated region, 3'-untranslated region, or intron).

[0212] As used herein, the term "regulatory elements" refers to nucleotide sequences, such as promoters, enhancers, terminators, polyadenylation sequences, introns, etc, that provide for the expression of a coding sequence in a cell.

[0213] As used herein, the term "promoter element" refers to a nucleotide sequence that assists with controlling expression of a coding sequence. Generally, promoter elements are located 5' of the translation start site of a gene. However, in certain embodiments, a promoter element may be located within an intron sequence, or 3' of the coding sequence. In some embodiments, a promoter useful for a gene therapy vector is derived from the native gene of the target protein (e.g., a Factor VIII promoter). In some embodiments, a promoter useful for a gene therapy vector is specific for expression in a particular cell or tissue of the target organism (e.g., a liver-specific promoter). In yet other embodiments, one of a plurality of well characterized promoter elements is used in a gene therapy vector described herein. Non-limiting examples of well-characterized promoter elements include the CMV early promoter, the β-actin promoter, and the methyl CpG binding protein 2 (MeCP2) promoter. In some embodiments, the promoter is a constitutive promoter, which drives substantially constant expression of the target protein. In other embodiments, the promoter is an inducible promoter, which drives expression of the target protein in response to a particular stimulus (e.g., exposure to a particular treatment or agent). For a review of designing promoters for AAV-mediated gene therapy, see Gray et al.

(Human Gene Therapy 22:1143-53 (2011)), the contents of which are expressly incorporated by reference in their entirety for all purposes.

[0214] As used herein, the term “vector” refers to any vehicle used to transfer a nucleic acid (e.g., encoding a Factor VIII gene therapy construct) into a host cell. In some embodiments, a vector includes a replicon, which functions to replicate the vehicle, along with the target nucleic acid. Non-limiting examples of vectors useful for gene therapy include plasmids, phages, cosmids, artificial chromosomes, and viruses, which function as autonomous units of replication in vivo. In some embodiments, a vector is a viral vehicle for introducing a target nucleic acid (e.g., a codon-altered polynucleotide encoding a Factor VIII variant). Many modified eukaryotic viruses useful for gene therapy are known in the art. For example, adeno-associated viruses (AAVs) are particularly well suited for use in human gene therapy because humans are a natural host for the virus, the native viruses are not known to contribute to any diseases, and the viruses illicit a mild immune response.

[0215] As used herein, the term “CpG island” refers to a region within a polynucleotide having a statistically elevated density of CpG dinucleotides. As used herein, a region of a polynucleotide (e.g., a polynucleotide encoding a codon-altered Factor VIII protein) is a CpG island if, over a 200-base pair window: (i) the region has GC content of greater than 50%, and (ii) the ratio of observed CpG dinucleotides per expected CpG dinucleotides is at least 0.6, as defined by the relationship:

$$\frac{N[CpG] * N[\text{length of window}]}{N[C] * N[G]} \geq 0.6.$$

For additional information on methods for identifying CpG islands, see Gardiner-Garden M. et al., J Mol Biol., 196(2): 261-82 (1987), the content of which is expressly incorporated herein by reference, in its entirety, for all purposes.

[0216] As used herein, the term “nucleic acid” refers to deoxyribonucleotides or ribonucleotides and polymers thereof in either single- or double-stranded form, and complements thereof. The term encompasses nucleic acids containing known nucleotide analogs or modified backbone residues or linkages, which are synthetic, naturally occurring, and non-naturally occurring, which have similar binding properties as the reference nucleic acid, and which are metabolized in a manner similar to the reference nucleotides. Examples of such analogs include, without limitation, phosphorothioates, phosphoramidates, methyl phosphonates, chiral-methyl phosphonates, 2-O-methyl ribonucleotides, and peptide-nucleic acids (PNAs).

[0217] The term “amino acid” refers to naturally occurring and non-natural amino acids, including amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids include those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxypyroline, γ -carboxyglutamate, and O-phosphoserine. Naturally occurring amino acids can include, e.g., D- and L-amino acids. The amino acids used herein can also include non-natural amino acids. Amino acid analogs refer to compounds that have the same basic chemical structure as a naturally occurring amino acid, i.e., any carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R

group, e.g., homoserine, norleucine, methionine sulfoxide, or methionine methyl sulfonium. Such analogs have modified R groups (e.g., norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refer to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that function in a manner similar to a naturally occurring amino acid. Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

[0218] The nucleotide sequences that encode the mutant Factor VIII constructs herein may be identical to the coding sequence provided herein or may be a different coding sequence, which sequence, as a result of the redundancy or degeneracy of the genetic code, encodes the same polypeptides as the coding sequences provided herein. One of ordinary skill in the art will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule. Accordingly, each variation of a nucleic acid which encodes a same polypeptide is implicit in each described sequence with respect to the expression product, but not with respect to actual gene therapy constructs.

[0219] As to amino acid sequences, one of ordinary skill in the art will recognize that individual substitutions, deletions or additions to a nucleic acid or peptide sequence that alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a “conservatively modified variant” where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the disclosure.

[0220] Conservative amino acid substitutions providing functionally similar amino acids are well known in the art. Dependent on the functionality of the particular amino acid, e.g., catalytic, structural, or sterically important amino acids, different groupings of amino acid may be considered conservative substitutions for each other. Table 1 provides groupings of amino acids that are considered conservative substitutions based on the charge and polarity of the amino acid, the hydrophobicity of the amino acid, the surface exposure/structural nature of the amino acid, and the secondary structure propensity of the amino acid.

TABLE 1

Groupings of conservative amino acid substitutions based on the functionality of the residue in the protein.

Important Feature	Conservative Groupings
Charge/Polarity	1. H, R, and K 2. D and E 3. C, T, S, G, N, Q, and Y 4. A, P, M, L, I, V, F, and W

TABLE 1-continued

Groupings of conservative amino acid substitutions based on the functionality of the residue in the protein.	
Important Feature	Conservative Groupings
Hydrophobicity	1. D, E, N, Q, R, and K 2. C, S, T, P, G, H, and Y 3. A, M, I, L, V, F, and W
Structural/Surface Exposure	1. D, E, N, Q, H, R, and K 2. C, S, T, P, A, G, W, and Y 3. M, I, L, V, and F
Secondary Structure Propensity	1. A, E, Q, H, K, M, L, and R 2. C, T, I, V, F, Y, and W 3. S, G, P, D, and N
Evolutionary Conservation	1. D and E 2. H, K, and R 3. N and Q 4. S and T 5. L, I, and V 6. F, Y, and W 7. A and G 8. M and C

[0221] The terms “identical” or percent “identity,” in the context of two or more nucleic acids or peptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same (i.e., about 60% identity, preferably 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or higher identity over a specified region, when compared and aligned for maximum correspondence over a comparison window or designated region) as measured using a BLAST or BLAST 2.0 sequence comparison algorithms with default parameters described below, or by manual alignment and visual inspection.

[0222] As is known in the art, a number of different programs may be used to identify whether a protein (or nucleic acid as discussed below) has sequence identity or similarity to a known sequence. Sequence identity and/or similarity is determined using standard techniques known in the art, including, but not limited to, the local sequence identity algorithm of Smith & Waterman, *Adv. Appl. Math.*, 2:482 (1981), by the sequence identity alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.*, 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Natl. Acad. Sci. U.S.A.*, 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Drive, Madison, Wis.), the Best Fit sequence program described by Devereux et al., *Nucl. Acid Res.*, 12:387-395 (1984), preferably using the default settings, or by inspection. Preferably, percent identity is calculated by FastDB based upon the following parameters: mismatch penalty of 1; gap penalty of 1; gap size penalty of 0.33; and joining penalty of 30, “Current Methods in Sequence Comparison and Analysis,” Macromolecule Sequencing and Synthesis, Selected Methods and Applications, pp 127-149 (1988), Alan R. Liss, Inc, all of which are incorporated by reference.

[0223] An example of a useful algorithm is PILEUP. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pair wise alignments. It may also plot a tree showing the clustering relationships used to create the alignment. PILEUP uses a simplification of the progressive alignment method of Feng

& Doolittle, *J. Mol. Evol.* 35:351-360 (1987); the method is similar to that described by Higgins & Sharp CABIOS 5:151-153 (1989), both incorporated by reference. Useful PILEUP parameters including a default gap weight of 3.00, a default gap length weight of 0.10, and weighted end gaps. Another example of a useful algorithm is the BLAST algorithm, described in: Altschul et al., *J. Mol. Biol.* 215, 403-410, (1990); Altschul et al., *Nucleic Acids Res.* 25:3389-3402 (1997); and Karlin et al., *Proc. Natl. Acad. Sci. U.S.A.* 90:5873-5787 (1993), both incorporated by reference. A particularly useful BLAST program is the WU-BLAST-2 program which was obtained from Altschul et al., *Methods in Enzymology*, 266:460-480 (1996); <http://blast.wustl.edu/blast/README.html>. WU-BLAST-2 uses several search parameters, most of which are set to the default values. The adjustable parameters are set with the following values: overlap span=1, overlap fraction=0.125, word threshold (T)=11. The HSP S and HSP S2 parameters are dynamic values and are established by the program itself depending upon the composition of the particular sequence and composition of the particular database against which the sequence of interest is being searched; however, the values may be adjusted to increase sensitivity.

[0224] An additional useful algorithm is gapped BLAST, as reported by Altschul et al., *Nucl. Acids Res.*, 25:3389-3402, incorporated by reference. Gapped BLAST uses BLOSUM-62 substitution scores; threshold T parameter set to 9; the two-hit method to trigger ungapped extensions; charges gap lengths of k a cost of 10+k; Xu set to 16, and Xg set to 40 for database search stage and to 67 for the output stage of the algorithms. Gapped alignments are triggered by a score corresponding to ~22 bits.

[0225] A % amino acid sequence identity value is determined by the number of matching identical residues divided by the total number of residues of the “longer” sequence in the aligned region. The “longer” sequence is the one having the most actual residues in the aligned region (gaps introduced by WU-Blast-2 to maximize the alignment score are ignored). In a similar manner, “percent (%)” nucleic acid sequence identity” with respect to the coding sequence of the polypeptides identified is defined as the percentage of nucleotide residues in a candidate sequence that are identical with the nucleotide residues in the coding sequence of the cell cycle protein. A preferred method utilizes the BLASTN module of WU-BLAST-2 set to the default parameters, with overlap span and overlap fraction set to 1 and 0.125, respectively.

[0226] The alignment may include the introduction of gaps in the sequences to be aligned. In addition, for sequences which contain either more or fewer amino acids than the protein encoded by the sequence of FIG. 2 (SEQ ID NO:1), it is understood that in one embodiment, the percentage of sequence identity will be determined based on the number of identical amino acids or nucleotides in relation to the total number of amino acids or nucleotides. Thus, for example, sequence identity of sequences shorter than that shown in FIG. 2 (SEQ ID NO:1), as discussed below, will be determined using the number of nucleotides in the shorter sequence, in one embodiment. In percent identity calculations relative weight is not assigned to various manifestations of sequence variation, such as, insertions, deletions, substitutions, etc.

[0227] In one embodiment, only identities are scored positively (+1) and all forms of sequence variation including

gaps are assigned a value of “0”, which obviates the need for a weighted scale or parameters as described below for sequence similarity calculations. Percent sequence identity may be calculated, for example, by dividing the number of matching identical residues by the total number of residues of the “shorter” sequence in the aligned region and multiplying by 100. The “longer” sequence is the one having the most actual residues in the aligned region.

[0228] The term “allelic variants” refers to polymorphic forms of a gene at a particular genetic locus, as well as cDNAs derived from mRNA transcripts of the genes, and the polypeptides encoded by them. The term “preferred mammalian codon” refers a subset of codons from among the set of codons encoding an amino acid that are most frequently used in proteins expressed in mammalian cells as chosen from the following list: Gly (GGC, GGG); Glu (GAG); Asp (GAC); Val (GTG, GTC); Ala (GCC, GCT); Ser (AGC, TCC); Lys (AAG); Asn (AAC); Met (ATG); Ile (ATC); Thr (ACC); Trp (TGG); Cys (TGC); Tyr (TAT, TAC); Leu (CTG); Phe (TTC); Arg (CGC, AGG, AGA); Gln (CAG); His (CAC); and Pro (CCC).

[0229] As used herein, the term codon-altered refers to a polynucleotide sequence encoding a polypeptide (e.g., a Factor VIII variant protein), where at least one codon of the native polynucleotide encoding the polypeptide has been changed to improve a property of the polynucleotide sequence. In some embodiments, the improved property promotes increased transcription of mRNA coding for the polypeptide, increased stability of the mRNA (e.g., improved mRNA half-life), increased translation of the polypeptide, and/or increased packaging of the polynucleotide within the vector. Non-limiting examples of alterations that can be used to achieve the improved properties include changing the usage and/or distribution of codons for particular amino acids, adjusting global and/or local GC content, removing AT-rich sequences, removing repeated sequence elements, adjusting global and/or local CpG dinucleotide content, removing cryptic regulatory elements (e.g., TATA box and CCAAT box elements), removing of intron/exon splice sites, improving regulatory sequences (e.g., introduction of a Kozak consensus sequence), and removing sequence elements capable of forming secondary structure (e.g., stem-loops) in the transcribed mRNA.

[0230] As discussed herein, there are various nomenclatures to refer to components of the disclosure herein. “CS-number” (e.g. “CS04”, “CS01”, “CS23”, etc.) refer to codon altered polynucleotides encoding FVIII polypeptides and/or the encoded polypeptides, including variants. For example, CS01-FL refers to the Full Length codon altered CS01 polynucleotide sequence or amino acid sequence (sometimes referred to herein as “CS01-FL-AA” for the Amino Acid sequence and “CS01-FL-NA” for the Nucleic Acid sequence) encoded by the CS01 polynucleotide sequence. Similarly, “CS01-LC” refers to either the codon altered nucleic acid sequence (“CS01-LC-NA”) encoding the light chain of a FVIII polypeptide or the amino acid sequence (also sometimes referred to herein as “CS01-LC-AA”) of the FVIII light chain encoded by the CS01 polynucleotide sequence. Likewise, CS01-HC, CS01-HC-AA and CS01-HC-NA are the same for the FVIII heavy chain. As will be appreciated by those in the art, for constructs such as CS01, CS04, CS23, etc., that are only codon-altered (e.g. they do not contain additional amino acid substitutions as compared to Refacto), the amino acid sequences will be identical, as

the amino acid sequences are not altered by the codon optimization. Thus, sequence constructs of the disclosure include, but are not limited to, CS01-FL-NA, CS01-FL-AA, CS01-LC-NA, CS01-LC-AA, CS01-HC-AA, CS01-HC-NA, CS04-FL-NA, CS04-FL-AA, CS04-LC-NA, CS04-LC-AA, CS04-HC-AA, CS04-HC-NA, CS23-FL-NA, CS23-FL-AA, CS23-LC-NA, CS23-LC-AA, CS23-HC-AA and CS23-HC-NA.

[0231] This nomenclature also applies to glycosylation peptides as shown in FIG. 13, such that “NGA1-AA” refers to the amino acid sequence and NGA1-NA refers to the nucleic acid sequence.

[0232] The disclosure also includes additional new Factor VIII variants, as described below, with the appropriate nomenclature.

III. Codon-Altered Factor VIII Variants

[0233] In some embodiments, the present disclosure provides codon-altered polynucleotides encoding Factor VIII variants. These codon-altered polynucleotides provide markedly improved expression of Factor VIII when administered in an AAV-based gene therapy construct. The codon-altered polynucleotides also demonstrate improved AAV-virion packaging, as compared to conventionally codon-optimized constructs. As demonstrated in Example 2 and Example 4, Applicants have achieved these advantages through the discovery of three codon-altered polynucleotides (CS01-FL-NA, CS04-FL-NA, and CS23-FL-NA) encoding a Factor VIII polypeptide with human wild-type Factor VIII heavy and light chains, and a short, 14 amino acid, B-domain substituted linker (the “SQ” linker) containing a furin cleavage site to facilitate maturation of an active FVIIIa protein in vivo. As further demonstrated in Example 4, incorporation of various combinations of the F328S, X5, and X1 amino acid mutations into the encoded Factor VIII molecule further increased the in vivo expression of Factor VIII activity.

[0234] In one embodiment, a codon-altered polynucleotide provided herein has nucleotide sequences with high sequence identity to at least the sequences within CS01, CS04, or CS23 (SEQ ID NOS 13, 1, and 20, respectively) encoding the Factor VIII heavy chain and Factor VIII light chains. As known in the art, the B-domain of Factor VIII is dispensable for activity in vivo. Thus, in some embodiments, the codon-altered polynucleotides provided herein completely lack a Factor VIII B-domain. In some embodiments, the native Factor VIII B-domain is replaced with a short amino acid linker containing a furin cleavage site, e.g., the “SQ” linker consisting of amino acids 760-773 of the CS01, CS04, or CS23 (SEQ ID NOS 2, 2, and 21, respectively) constructs. The “SQ” linker is also referred to as BDLO04, (−AA for the amino acid sequence and −NA for the nucleotide sequence shown in FIG. 6).

[0235] In one embodiment, the Factor VIII heavy and light chains encoded by the codon-altered polynucleotide are human Factor VIII heavy and light chains, respectively. In other embodiments, the Factor VIII heavy and light chains encoded by the codon-altered polynucleotide are heavy and light chain sequences from another mammal (e.g., porcine Factor VIII). In yet other embodiments, the Factor VIII heavy and light chains are chimeric heavy and light chains (e.g., a combination of human and a second mammalian sequence). In yet other embodiments, the Factor VIII heavy and light chains are humanized version of the heavy and

light chains from another mammal, e.g., heavy and light chain sequences from another mammal in which human residues are substituted at select positions to reduce the immunogenicity of the resulting peptide when administered to a human.

[0236] The GC content of human genes varies widely, from less than 25% to greater than 90%. However, in general, human genes with higher GC contents are expressed at higher levels. For example, Kudla et al. (PLoS Biol., 4(6):80 (2006)) demonstrate that increasing a gene's GC content increases expression of the encoded polypeptide, primarily by increasing transcription and effecting a higher steady state level of the mRNA transcript. Generally, the desired GC content of a codon-optimized gene construct is equal or greater than 60%. However, native AAV genomes have GC contents of around 56%.

[0237] Accordingly, in some embodiments, the codon-altered polynucleotides provided herein have a CG content that more closely matches the GC content of native AAV virions (e.g., around 56% GC), which is lower than the preferred CG contents of polynucleotides that are conventionally codon-optimized for expression in mammalian cells (e.g., at or above 60% GC). As outlined in Example 1, CS04-FL-NA (SEQ ID NO: 1), which has a GC content of about 56%, has improved virion packaging as compared to similarly codon-altered coding sequences with higher GC content.

[0238] Thus, in some embodiments, the overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is less than 60%. In some embodiments, the overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is less than 59%. In some embodiments, the overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is less than 58%. In some embodiments, the overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is less than 57%. In some embodiments, the overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is no more than 56%.

[0239] In some embodiments, the overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is from 54% to 59%. In some embodiments, the overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is from 55% to 59%. In some embodiments, the overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is from 56% to 59%. In some embodiments, the overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is from 54% to 58%. In some embodiments, the overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is from 55% to 58%. In some embodiments, the overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is from 56% to 58%. In some embodiments, the overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is from 54% to 57%. In some embodiments, the overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is from 55% to 57%. In some embodiments, the overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is from 56% to 57%. In some embodiments, the overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is from 54% to 56%. In some embodiments, the

overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is from 55% to 56%.

[0240] In some embodiments, the overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is $56\pm0.5\%$. In some embodiments, the overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is $56\pm0.4\%$. In some embodiments, the overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is $56\pm0.3\%$. In some embodiments, the overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is $56\pm0.2\%$. In some embodiments, the overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is $56\pm0.1\%$. In some embodiments, the overall GC content of a codon-altered polynucleotide encoding a Factor VIII polypeptide is 56%.

[0241] A. Factor VIII Amino Acid Substitutions

[0242] To further increase the efficiency of AAV-vector based expression of the Factor VIII constructs described herein, amino acid substitutions known to improve secretion, increase specific activity, and/or enhanced the stability of Factor VIII are further incorporated, in some implementations. A number of potential variants were identified that increase the plasma levels of FVIII activity at a given vector dose. These variants include those with a more efficient signal peptide, amino acid substitutions that prevent BiP interactions, amino acid substitutions resembling more efficiently secreted Factor VIII orthologs (e.g., porcine Factor VIII), single-chain Factor VIII variants, and amino acid substitutions that stabilize Factor VIII and/or reduce subunit dissociation.

[0243] Mutation of residues A108, R121, and L2302 (SPE), located at the interface between the A1 and C2 domains, increases the stability of Factor VIII. For example, the A108I amino acid substitution introduces a hydrophobic residue that better fills the inter-domain space, stabilizing the interaction. Likewise, an R121C/L2302C (SPE) double amino acid substitution introduces a disulfide bond spanning the A1-C2 domains, further stabilizing the interaction. Taken together, all three amino acid substitutions increase the thermal stability of Factor VIII by 3 to 4-fold. For review, see Wakabayashi et al., J Biol Chem. 286(29):25748-55 (2011) and Wakabayashi et al., Thromb Haemost. 103: 492-95 (2012). Accordingly, in some embodiments, the encoded Factor VIII polypeptide includes A108I and/or R121C/L2302C amino acid substitutions.

[0244] Mutation of E113 (SPE), located within the calcium binding domain of Factor VIII, increases the specific FVIII clotting activity. For example, E113A appears to increase FXase formation through increased FVIII affinity for Factor IXa. Specifically, the E113A amino acid substitution increases specific FVIII clotting activity two-fold and increases affinity for Factor IXa by four-fold (Biochemistry, 41:8485 (2002); J. Biol. Chem., 279:12677 (2004); and Biochemistry, 44:10298 (2005)). Accordingly, in some embodiments, the encoded Factor VIII polypeptides include an E113A amino acid substitution.

[0245] Substitution of one or more amino acid residues surrounding the Factor VIII APC cleavage site (residues 331-341 (SPE)) reduce Factor VIIIa inactivation by activated protein C, without affecting FVIII activity. For example PQL333-335VDQ (SPE) amino acid substitutions reduce Factor VIII inactivation by 16-fold. Likewise, MKN336-339GNQ amino acid substitutions reduce Factor

VIII inactivation by 9-fold. When combined, the two triple amino acid substitutions (e.g., PQLRMKN333-339VDQRGNQ) (SEQ ID NOS 34 and 35, respectively) reduce Factor VIII inactivation by 100-fold (*J. Biol. Chem.*, 282:20264 (2007)). Accordingly, in some embodiments, the encoded Factor VIII polypeptide include PQL333-335VDQ and/or MKN337-339GNQ (SPE) amino acid substitutions.

[0246] Mutations within the A2 domain interface also increase Factor VIII stability. Specifically, mutating charged residues in the A1-A2 and A2-A3 domain interfaces increases stability and retention of the A2 subunit in Factor VIIIa. For example, mutation of D519, E665, and E1984 to V or A yields up to 2-fold increased stability in Factor VIII and up to 5-fold stability in Factor VIIIa. Specifically, D519A/E665V amino acid substitutions provide a 3-fold increase in stability; D519V/E665V amino acid substitutions provide a 2-fold increase in stability, an 8-fold decrease in A2 dissociation, and a 2-4-fold increase in thrombin generation potential; D519V/E1984A amino acid substitutions provide a 2-fold increase in stability; and D519V/E665V/E1984A amino acid substitution provide a 2-fold increase in stability (*Blood* 112:2761-69 (2008); *J. Thromb. Haemost.*, 7:438-44 (2009)). Accordingly, in some embodiments, the encoded Factor VIII polypeptides include one or more of D519A/V, E665A/V, and E1984A/V amino acid substitutions.

[0247] Of particular relevance to the present disclosure are a number of specific mutations that can be included separately or in combinations with other variants described herein. These variants are coded as sets herein as follows: “m1” refers to a single amino acid change, “m2” is a set of 5 amino acid variants, “m3” is a combination of a deletion of 7 amino acids and an insertion of six amino acids that span the junction between the polypeptide linker and the heavy chain, “m4” is a combination of the m1 single mutation and the m5 double mutation, and “m5” is a set of two cysteine ablations. These mutations are described below. These can be included in any particular construct alone or in combination with other variants, and they are coded accordingly. For example, “m23” is a combination of the m2 and m3 variants onto a particular scaffold, as outlined herein; thus “CS01m23-FL-NA” or “CS01-FL-NAm23” refers to the CS01 codon-altered polynucleotide sequence with the nucleotides encoding the m2 and m3 mutations included, and “CS01m23-FL-AA” or “CS01-FL-AAm23” refers to the amino acid sequence. As CS01 is codon-altered but does not change the amino acid sequence of Refacto, these can be thought of on the amino acid level as mutations as compared to the Refacto amino acid sequence of CS01-FL-AA (SEQ ID NO: 2).

[0248] In many embodiments, the polypeptides of the disclosure are made with the “m1” variant included. Mutations within an 11 amino acid hydrophobic β-sheet in the A1 domain, which interacts with BiP, increase secretion of Factor VIII. For example, an F328S (SPI, F3095 SPE) amino acid substitution within the pocket increased Factor VIII secretion 3-fold. The F328S variant is referred to herein as the “m1” mutation and is within the heavy chain. Again, as described herein, the number of the variants can be done inclusive of the signal peptide, “Signal Peptide Inclusive”, or “SPI”, or starting from the processed final protein sequence, “Signal Peptide Exclusive”, or “SPE”. Thus, using SPI numbering, the mutation F328S is the same as the F309 SPE mutant. Generally the specification uses the SPI

numbering, but as will be appreciated by those in the art, either numbering system results in the same mutation(s).

[0249] Accordingly, included in the present disclosure are polypeptides that include the m1 mutation, including CS01-FL-AAm1, CS01-HC-AAm1, CS04-FL-AAm1, CS04-HC-AAm1 CS23-FL-AAm1, CS23-HC-AAm1, CS40-FL-AAm1 and CS40-HC-AAm1 (all of which encode the same corresponding protein sequences).

[0250] In addition, included in the present disclosure are not only polypeptide sequences that include the m1 mutation, but also those codon-altered polynucleotide sequences that encode proteins with the m1 mutation, such as CS01-FL-NAm1, CS01-HC-NAm1, CS04-FL-NAm1, CS04-HC-NAm1, CS23-FL-NAm1, CS23-HC-NA-m1, CS40-FL-NAm1 and CS40-HC-NAm1.

[0251] In many embodiments, the polypeptides of the disclosure are made with the “m2” variant set included, which is the I105V/A127S/G151K/M166T/L171P mutations (SPI numbering; (SPE numbering is V861/S108A/K132G/T147M/P152L, respectively). The m2 mutation set is based on the fact that substitution of porcine amino acids 82-176 for the corresponding human amino acids in a B-domain deleted gene therapy construct increased Factor VIII activity when expressed in HEK293 cells (W. Xiao, communication). Id. Back-mutation of single porcine amino acids into the human BDD-FVIII construct identified five amino acids within the A1 domain that contribute to this phenomenon: I105V, A127S, G151K, M166T, and L171P (SPI). Introduction of the combination of these mutations into the human construct recapitulated the improved activity of the larger porcine substitution. Id. Accordingly, in some embodiments, the encoded Factor VIII polypeptides include one or more amino acid substitutions selected from I105V, A127S, G151K, M166T, and L171P, with the entire 5 amino acid set, m2, finding particular use in many embodiments. As for the m1 mutation, the m2 variants are in the heavy chain, and thus the present disclosure includes polypeptides that include the m2 mutation, including CS01-FL-AAm2, CS01-HC-AAm2, CS04-FL-AAm2, CS04-HC-AAm2, CS23-FL-AAm2, CS23-HC-AAm2, CS40-FL-AAm2 and CS40-HC-AAm2 (all of which encode the same corresponding protein sequences).

[0252] In addition, included in the present disclosure are not only polypeptide sequences that include the m2 mutation, but also those codon-altered polynucleotide sequences that encode proteins with the m2 mutation, such as CS01-FL-NAm2, CS01-HC-NAm2, CS04-FL-NAm2, CS04-HC-NAm2, CS23-FL-NAm2, CS23-HC-NA-m2, CS40-FL-NAm2 and CS40-HC-NAm2.

[0253] In additional embodiments, the polypeptides and polynucleotides of the disclosure include m3 mutations. m3 is the substitution of seven amino acids for six across the HC-B domain interface that introduces an additional glycosylation site introduced close to the interface. Accordingly, in some embodiments, m3 is the deletion of amino acids AIEPRSF755-761 and the insertion of amino acids TTYVNRSL (SEQ ID NO: 33) after N754, relative to FVIII-FL-AA (SEQ ID NO: 19) (e.g., AIEPRSF755-761TTYVNRSL (“TTYVNRSL” disclosed as SEQ ID NO: 33). Residues AIEPRSF755-759, relative to SEQ ID NO: 19, fall within the end of the heavy chain, while residues 5760 and F761 fall within the B-domain. In some embodiments, where the FVIII B-domain is deleted, truncated, or replaced, residues 5760 and F761 may not be present in the underlying

amino acid sequence being mutated. Accordingly, in some embodiments, m3 is the deletion of amino acids AIEPR755-759 and the insertion of amino acids TTYVNRS (SEQ ID NO: 33) after N754, relative to FVIII-FL-AA (SEQ ID NO: 19) (e.g., AIEPR755-759TTYVNRS ("TTYVNRS" disclosed as SEQ ID NO: 33)

[0254] The m3 variants are in the junction between the heavy chain and the B domain, and thus the present disclosure includes polypeptides that include the m3 mutation, including CS01-FL-AAm3, CS01-HC-AAm3, CS04-FL-AAm3, CS04-HC-AAm3, CS23-FL-AAm3, CS23-HC-AAm3, CS40-FL-AAm3 and CS40-HC-AAm3 (all of which encode the same corresponding protein sequences).

[0255] In addition, included in the present disclosure are not only polypeptide sequences that include the m3 mutation, but also those codon-altered polynucleotide sequences that encode proteins with the m3 mutations, such as CS01-FL-NAm3, CS01-HC-NAm3, CS04-FL-NAm3, CS04-HC-NAm3, CS23-FL-NAm3, CS23-HC-NA-m3, CS40-FL-NAm3 and CS40-HC-NAm3.

[0256] In additional embodiments, the polypeptides and polynucleotides of the disclosure include m4 mutations. Elimination of the C1899-C1903 disulfide bond in Factor VIII also increased secretion. Moreover, the increases in Factor VIII secretion are additive for the combination of F328S (SPI, F3095 SPE) and C1918G/C1922G amino acid substitutions (Miao et al., Blood, 103:3412-19 (2004); Selvaraj et al., J. Thromb. Haemost., 10:107-15 (2012)). Accordingly, in some embodiments, the encoded Factor VIII polypeptides include m4 mutations, which is the F328S (SPI, F3095 SPE) and C1918G/C1922G (SPI) amino acid substitutions. As the F328S variant is in the heavy chain and the two cysteine variants are in the light chain, polypeptide sequences that include m4 mutations are CS01-FL-AAm4, CS01-HC-AAm4, CS01-LC-AAm4, CS04-FL-AAm4, CS04-HC-AAm4, CS04-LC-AAm4, CS23-FL-AAm4, CS23-HC-AAm4 and CS23-LC-AAm4.

[0257] In addition, included in the present disclosure are not only polypeptide sequences that include the m4 mutation, but also those codon-altered polynucleotide sequences that encode proteins with the m4 mutations, such as CS01-FL-NAm4, CS01-HC-NAm4, CS01-LC-NAm4, CS04-FL-NAm4, CS04-HC-NAm4, CS04-LC-NAm4, CS23-FL-NAm4, CS23-HC-NAm4, CS23-LC-NAm4, CS40-FL-NA-m4, CS40-HC-NA-m4 and CS40-LC-NA-m4.

[0258] In additional embodiments, the polypeptides and polynucleotides of the disclosure include m5 mutations. As above, elimination of the C1899-C1903 disulfide bond in Factor VIII also increased secretion. C1918G/C1922G (SPI) amino acid substitutions, contained within the light chain, referred to herein as the m5 mutation set.

[0259] The m5 variants are in the light chain, and thus the present disclosure includes polypeptides that include the m5 mutation, including CS01-FL-AAm5, CS01-LC-AAm5, CS04-FL-AAm5, CS04-LC-AAm5, CS23-FL-AAm5, CS23-LC-AAm5, CS40-FL-AAm5 and CS40-LC-AAm5 (all of which encode the same corresponding protein sequences).

[0260] In addition, included in the present disclosure are not only polypeptide sequences that include the m5 mutation, but also those codon-altered polynucleotide sequences that encode proteins with the m5 mutations, such as CS01-FL-NAm5, CS01-LC-NAm5, CS04-FL-NAm5, CS04-LC-

NAm5, CS23-FL-NA-m5, CS23-LC-NA-m5, CS40-FL-NA-m5 and CS40-LC-NA-m5.

[0261] In addition to specific constructs (both amino acid and nucleic acid) that include m1, m2, m3, m4 and m5 individually, combinations of mutation sets can be made as outlined herein. As noted herein, these are noted as "m12", which is the combination of m1 and m2 sets, or "m123" which is the combination of m1, m2 and m3 sets. Thus, included in the disclosure are dual combinations including m12, m13, m14, m15, m23, m24, m25, m34, m35 and m45. Also included are triple combinations, m123, m124, m125, m234, m235 and m345. Further included are quad combinations, m1234, m1235, m1345 and the m12345 combination.

[0262] Of particular interest in some embodiments are the following mutation sets: m1, m2, m3 and m4, m23, m123, and m234.

[0263] B. Factor VIII B-Domain Substituted Linkers

[0264] In some embodiments, the linkage between the FVIII heavy chain and the light chain (e.g., the B-domain in wild-type Factor VIII) is further altered. Due to size constraints of AAV packaging capacity, B-domain deleted, truncated, and/or linker substituted variants should improve the efficacy of the FVIII gene therapy construct. The most conventionally used B-domain substituted linker is that of SQ FVIII, which retains only 14 amino acids of the B domain as linker sequence. Another variant of porcine VIII ("OBI-1," described in U.S. Pat. No. 6,458,563) is well expressed in CHO cells, and has a slightly longer linker of 24 amino acids. In some embodiments, the Factor VIII constructs encoded by the codon-altered polynucleotides described herein include an SQ-type B-domain linker sequence. In other embodiments, the Factor VIII constructs encoded by the codon-altered polynucleotides described herein include an OBI-1-type B-domain linker sequence.

[0265] In some embodiments, the encoded Factor VIII polypeptides described herein include an SQ-type B-domain linker, including amino acids 760-762/1657-1667 of the wild-type human Factor VIII B-domain (FVIII-FL-AA; SEQ ID NO: 19) (Sandberg et al. Thromb. Haemost. 85:93 (2001)). In some embodiments, the SQ-type B-domain linker has one amino acid substitution relative to the corresponding wild-type sequence. In some embodiments, the SQ-type B-domain linker has two amino acid substitutions relative to the corresponding wild-type sequence. In some embodiments, a glycosylation peptide is inserted into the SQ-type B-domain linker. In some embodiments, the glycosylation peptide is selected from those shown in FIG. 13 (SEQ ID NOS 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, and 75, respectively, in order of appearance).

[0266] In some embodiments, the encoded Factor VIII polypeptides described herein include a Greengene-type B-domain linker, including amino acids 760/1582-1667 of the wild-type human Factor VIII B-domain (FVIII-FL-AA; SEQ ID NO: 19) (Oh et al., Biotechnol. Prog., 17:1999 (2001)). In some embodiments, the Greengene-type B-domain linker has one amino acid substitution relative to the corresponding wild-type sequence. In some embodiments, the Greengene-type B-domain linker has two amino acid substitutions relative to the corresponding wild-type sequence. In some embodiments, a glycosylation peptide is inserted into the Greengene-type B-domain linker. In some embodiments, the glycosylation peptide is selected from

those shown in FIG. 13 (SEQ ID NOS 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, and 75, respectively, in order of appearance).

[0267] In some embodiments, the encoded Factor VIII polypeptides described herein include an extended SQ-type B-domain linker (SFSQNPPVLKRHQR; BDL-SQ-AA; SEQ ID NO: 30), including amino acids 760-769/1657-1667 of the wild-type human Factor VIII B-domain (FVIII-FL-AA; SEQ ID NO: 19) (Thim et al., *Haemophilia*, 16:349 (2010)). In some embodiments, the extended SQ-type B-domain linker has one amino acid substitution relative to the corresponding wild-type sequence. In some embodiments, the extended SQ-type B-domain linker has two amino acid substitutions relative to the corresponding wild-type sequence. In some embodiments, a glycosylation peptide is inserted into the extended SQ-type B-domain linker. In some embodiments, the glycosylation peptide is selected from those shown in FIG. 13 (SEQ ID NOS 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, and 75, respectively, in order of appearance).

[0268] In some embodiments, the encoded Factor VIII polypeptides described herein include a porcine OBI-1-type B-domain linker, including the amino acids SFAQNSR-PPSASAPKPPVLRHHQR (SEQ ID NO: 31) from the wild-type porcine Factor VIII B-domain (Toschi et al., *Curr. Opin. Mol. Ther.* 12:517 (2010)). In some embodiments, the porcine OBI-1-type B-domain linker has one amino acid substitution relative to the corresponding wild-type sequence. In some embodiments, the porcine OBI-1-type B-domain linker has two amino acid substitutions relative to the corresponding wild-type sequence. In some embodiments, a glycosylation peptide is inserted into the porcine OBI-1-type B-domain linker. In some embodiments, the glycosylation peptide is selected from those shown in FIG. 13 (SEQ ID NOS 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, and 75, respectively, in order of appearance).

[0269] In some embodiments, the encoded Factor VIII polypeptides described herein include a human OBI-1-type B-domain linker, including amino acids 760-772/1655-1667 of the wild-type human Factor VIII B-domain (FVIII-FL-AA; SEQ ID NO: 19). In some embodiments, the human OBI-1-type B-domain linker has one amino acid substitution relative to the corresponding wild-type sequence. In some embodiments, the human OBI-1-type B-domain linker has two amino acid substitutions relative to the corresponding wild-type sequence. In some embodiments, a glycosylation peptide is inserted into the human OBI-1-type B-domain linker. In some embodiments, the glycosylation peptide is selected from those shown in FIG. 13 (SEQ ID NOS 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, and 75, respectively, in order of appearance).

[0270] In some embodiments, the encoded Factor VIII polypeptides described herein include an 08-type B-domain linker, including the amino acids SFSQNSRHQAYRYRRG (SEQ ID NO: 32) from the wild-type porcine Factor VIII B-domain (Toschi et al., *Curr. Opin. Mol. Ther.* 12:517 (2010)). In some embodiments, the porcine OBI-1-type B-domain linker has one amino acid substitution relative to the corresponding wild-type sequence. In some embodiments, the porcine OBI-1-type B-domain linker has two amino acid substitutions relative to the corresponding wild-type sequence. In some embodiments, a glycosylation peptide is inserted into the porcine OBI-1-type B-domain linker. In some embodiments, the glycosylation peptide is selected

from those shown in FIG. 13 (SEQ ID NOS 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, and 75, respectively, in order of appearance).

[0271] Removal of the B-domain from Factor VIII constructs does not appear to affect the activity of the activated enzyme (e.g., FVIIa), presumably because the B-domain is removed during activation. However, the B-domain of Factor VIII contains several residues that are post-translationally modified, e.g., by N- or O-linked glycosylation. In silico analysis (Prediction of N-glycosylation sites in human proteins, R. Gupta, E. Jung and S. Brunak, *in preparation* (2004)) of the wild-type Factor VIII B-domain predicts that at least four of these sites are glycosylated in vivo (FIG. 14). It is thought that these modifications within the B-domain contribute to the post-translational regulation and/or half-life of Factor VIII in vivo.

[0272] While the Factor VIII B-domain is absent in mature Factor VIIIa protein, glycosylation within the B-domain of the precursor Factor VIII molecule may increase the circulating half-life of the protein prior to activation. Thus, in some embodiments, the polypeptide linker of the encoded Factor VIII constructs described herein includes one or more glycosylation sequences, to allow for glycosylation in vivo. In some embodiments, the polypeptide linker includes at least one consensus glycosylation sequence (e.g., an N- or O-linked glycosylation consensus sequence). In some embodiments, the polypeptide linker includes at least two consensus glycosylation sequences. In some embodiments, the polypeptide linker includes at least three consensus glycosylation sequences. In some embodiments, the polypeptide linker includes at least four consensus glycosylation sequences. In some embodiments, the polypeptide linker includes at least five consensus glycosylation sequences. In some embodiments, the polypeptide linker includes at least six, 7, 8, 9, 10, or more consensus glycosylation sequences.

[0273] In some embodiments, the polypeptide linker contains at least one N-linked glycosylation sequence N—X-S/T, where X is any amino acid other than P, S, or T. In some embodiments, the polypeptide linker contains at least two N-linked glycosylation sequences N—X-S/T, where X is any amino acid other than P, S, or T. In some embodiments, the polypeptide linker contains at least three N-linked glycosylation sequences N—X-S/T, where X is any amino acid other than P, S, or T. In some embodiments, the polypeptide linker contains at least four N-linked glycosylation sequences N—X-S/T, where X is any amino acid other than P, S, or T. In some embodiments, the polypeptide linker contains at least five N-linked glycosylation sequences N—X-S/T, where X is any amino acid other than P, S, or T. In some embodiments, the polypeptide linker contains at least six, 7, 8, 9, 10, or more N-linked glycosylation sequences N—X-S/T, where X is any amino acid other than P, S, or T.

[0274] In some embodiments, the polypeptide linker includes a glycosylation peptide with high sequence identity to any one of SEQ ID NOS 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, and 75, respectively, in order of appearance, as shown in FIGS. 13A-13B. In some embodiments, the glycosylation polypeptide has at least 92% identity to any one of SEQ ID NOS 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, and 75, respectively, in order of appearance, as shown in FIGS. 13A-13B. In some embodiments, the glycosylation peptide has no more than two amino acid substitutions relative to any one of SEQ ID NOS 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, and 75, respectively, in order

amino acid substitutions relative to any one of SEQ ID NOS 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, and 75, respectively, in order of appearance, as shown in FIGS. 13A-13B, and a glycosylation peptide having no more than one amino acid substitution relative to any one of SEQ ID NOS 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, and 75, respectively, in order of appearance, as shown in FIGS. 13A-13B. In some embodiments, the glycosylation peptide is inserted in the SQ peptide between residues N768 and P769 (relative to CS04-FL-AA; SEQ ID NO: 2).

[0280] In some embodiments, the polypeptide linker of the Factor VIII construct is encoded by a third nucleotide sequence having high sequence identity to any one of those shown in FIG. 6 (SEQ ID NOS 5-7 and 36-48, respectively, in order of appearance). In some embodiments, the third nucleotide sequence has at least 95% identity to any one of those shown in FIG. 13 (SEQ ID NOS 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, and 74, respectively, in order of appearance). In some embodiments, the third nucleotide sequence has at least 96% identity to any one of those shown in FIG. 13 (SEQ ID NOS 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, and 74, respectively, in order of appearance). In some embodiments, the third nucleotide sequence has at least 97% identity to any one of those shown in FIG. 13 (SEQ ID NOS 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, and 74, respectively, in order of appearance). In some embodiments, the third nucleotide sequence has at least 98% identity to any one of those shown in FIG. 13 (SEQ ID NOS 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, and 74, respectively, in order of appearance). In some embodiments, the third nucleotide sequence has at least 99% identity to any one of those shown in FIG. 13 (SEQ ID NOS 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, and 74, respectively, in order of appearance). In some embodiments, the third nucleotide sequence has at least 99.5% identity to any one of those shown in FIG. 13 (SEQ ID NOS 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, and 74, respectively, in order of appearance). In some embodiments, the third nucleotide sequence has at least 99.9% identity to any one of those shown in FIG. 13 (SEQ ID NOS 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, and 74, respectively, in order of appearance). In some embodiments, the third nucleotide sequence is identical to any one of those shown in FIG. 13 (SEQ ID NOS 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, and 74, respectively, in order of appearance).

[0281] C. Codon-Altered Polynucleotides Encoding a Factor VIII Variant with a Cleavable Linker

[0282] CS04 Codon Altered Polynucleotides

[0283] In one embodiment, the codon-altered polynucleotides provided herein include a nucleotide sequence encoding a Factor VIII variant polypeptide with a linker that is cleavable in vivo. The Factor VIII polypeptide includes a Factor VIII light chain, a Factor VIII heavy chain, and a polypeptide linker joining the C-terminus of the heavy chain to the N-terminus of the light chain. The heavy chain of the Factor VIII polypeptide is encoded by a first nucleotide sequence having high sequence identity to CS04-HC-NA (SEQ ID NO: 3), which is the portion of CS04-FL-NA (SEQ ID NO: 1) encoding for a Factor VIII heavy chain. The light chain of the Factor VIII polypeptide is encoded by a second nucleotide sequence with high sequence identity to CS04-LC-NA (SEQ ID NO: 4), which is the portion of CS04-FL-NA (SEQ ID NO: 1) encoding for a Factor VIII light chain. The polypeptide linker includes a furin cleavage site, which

allows for maturation in vivo (e.g., after expression in vivo or administration of the precursor polypeptide).

[0284] In some embodiments, the first and second nucleotide sequences have at least 95% sequence identity to CS04-HC-NA and CS04-LC-NA (SEQ ID NOS 3 and 4), respectively. In some embodiments, the first and second nucleotide sequences have at least 96% sequence identity to CS04-HC-NA and CS04-LC-NA (SEQ ID NOS 3 and 4), respectively. In some embodiments, the first and second nucleotide sequences have at least 97% sequence identity to CS04-HC-NA and CS04-LC-NA (SEQ ID NOS 3 and 4), respectively. In some embodiments, the first and second nucleotide sequences have at least 98% sequence identity to CS04-HC-NA and CS04-LC-NA (SEQ ID NOS 3 and 4), respectively. In some embodiments, the first and second nucleotide sequences have at least 99% sequence identity to CS04-HC-NA and CS04-LC-NA (SEQ ID NOS 3 and 4), respectively. In some embodiments, the first and second nucleotide sequences have at least 99.5% sequence identity to CS04-HC-NA and CS04-LC-NA (SEQ ID NOS 3 and 4), respectively. In some embodiments, the first and second nucleotide sequences have at least 99.9% sequence identity to CS04-HC-NA and CS04-LC-NA (SEQ ID NOS 3 and 4), respectively. In some embodiments, the first and second nucleotide sequences are identical to CS04-HC-NA and CS04-LC-NA (SEQ ID NOS 3 and 4), respectively.

[0285] In some embodiments, the polypeptide linker of the Factor VIII construct is encoded by a third nucleotide sequence having high sequence identity to BDLO04 (SEQ ID NO: 6), which encodes the 14-amino acid linker corresponding to amino acids 760-773 of CS04-FL-AA (SEQ ID NO: 2). In some embodiments, the third nucleotide sequence has at least 95% identity to BDLO04 (SEQ ID NO: 6). In some embodiments, the third nucleotide sequence has at least 96% identity to BDLO04 (SEQ ID NO: 6). In some embodiments, the third nucleotide sequence has at least 97% identity to BDLO04 (SEQ ID NO: 6). In some embodiments, the third nucleotide sequence has at least 98% identity to BDLO04 (SEQ ID NO: 6). In some embodiments, the third nucleotide sequence is identical to BDLO04 (SEQ ID NO: 6).

[0286] In some embodiments, the codon-altered polynucleotide has a nucleotide sequence with high sequence identity to CS04-FL-NA (SEQ ID NO: 1). In some embodiments, the nucleotide sequence has at least 95% identity to CS04-FL-NA (SEQ ID NO: 1). In some embodiments, the nucleotide sequence has at least 96% identity to CS04-FL-NA (SEQ ID NO: 1). In some embodiments, the nucleotide sequence has at least 97% identity to CS04-FL-NA (SEQ ID NO: 1). In some embodiments, the nucleotide sequence has at least 98% identity to CS04-FL-NA (SEQ ID NO: 1). In some embodiments, the nucleotide sequence has at least 99% identity to CS04-FL-NA (SEQ ID NO: 1). In some embodiments, the nucleotide sequence has at least 99.5% identity to CS04-FL-NA (SEQ ID NO: 1). In some embodiments, the nucleotide sequence has at least 99.9% identity to CS04-FL-NA (SEQ ID NO: 1). In some embodiments, the nucleotide sequence is identical to CS04-FL-NA (SEQ ID NO: 1).

[0287] In some embodiments, the Factor VIII variant encoded by the codon-altered polynucleotide has an amino acid sequence with high sequence identity to CS04-FL-AA (SEQ ID NO: 2). In some embodiments, the amino acid sequence has at least 97% identity to CS04-FL-AA (SEQ ID NO: 2). In some embodiments, the amino acid sequence has at least 98% identity to CS04-FL-AA (SEQ ID NO: 2). In some embodiments, the amino acid sequence has at least 99% identity to CS04-FL-AA (SEQ ID NO: 2). In some embodiments, the amino acid sequence has at least 99.5% identity to CS04-FL-AA (SEQ ID NO: 2). In some embodiments, the amino acid sequence has at least 99.9% identity to CS04-FL-AA (SEQ ID NO: 2). In some embodiments, the amino acid sequence is identical to CS04-FL-AA (SEQ ID NO: 2).

some embodiments, the amino acid sequence has at least 99% identity to CS04-FL-AA (SEQ ID NO: 2). In some embodiments, the amino acid sequence has at least 99.5% identity to CS04-FL-AA (SEQ ID NO: 2). In some embodiments, the amino acid sequence has at least 99.9% identity to CS04-FL-AA (SEQ ID NO: 2). In some embodiments, the amino acid sequence is identical to CS04-FL-AA (SEQ ID NO: 2).

[0288] In some embodiments, the Factor VIII variant encoded by the CS04 polynucleotide, having high sequence homology to CS04-FL-AA (e.g., at least 95%, 96%, 97%, 98%, 99%, 99.5%, or 99.9% identity), comprises one or more amino acid substitutions selected from m1, m2, m3, m4, and m5.

[0289] In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises an m1 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises an m2 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises an m3 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises an m4 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises an m5 amino acid substitution.

[0290] In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises m12 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises m13 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises m23 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises m24 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises m25 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises m34 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises m35 amino acid substitutions.

[0291] In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises m123 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises m234 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises m125 amino acid substitutions.

[0292] CS01 Codon Altered Polynucleotides

[0293] In one embodiment, the codon-altered polynucleotides provided herein include a nucleotide sequence encoding a Factor VIII variant polypeptide with a linker that is cleavable in vivo. The Factor VIII polypeptide includes a Factor VIII light chain, a Factor VIII heavy chain, and a polypeptide linker joining the C-terminus of the heavy chain to the N-terminus of the light chain. The heavy chain of the Factor VIII polypeptide is encoded by a first nucleotide sequence having high sequence identity to CS01-HC-NA (SEQ ID NO: 24), which is the portion of CS01-FL-NA (SEQ ID NO: 13) encoding for a Factor VIII heavy chain. The light chain of the Factor VIII polypeptide is encoded by a second nucleotide sequence with high sequence identity to CS01-LC-NA (SEQ ID NO: 25), which is the portion of CS01-FL-NA (SEQ ID NO: 13) encoding for a Factor VIII

light chain. The polypeptide linker includes a furin cleavage site, which allows for maturation in vivo (e.g., after expression in vivo or administration of the precursor polypeptide).

[0294] In some embodiments, the first and second nucleotide sequences have at least 95% sequence identity to CS01-HC-NA and CS01-LC-NA (SEQ ID NOS 24 and 25), respectively. In some embodiments, the first and second nucleotide sequences have at least 96% sequence identity to CS01-HC-NA and CS01-LC-NA (SEQ ID NOS 24 and 25), respectively. In some embodiments, the first and second nucleotide sequences have at least 97% sequence identity to CS01-HC-NA and CS01-LC-NA (SEQ ID NOS 24 and 25), respectively. In some embodiments, the first and second nucleotide sequences have at least 98% sequence identity to CS01-HC-NA and CS01-LC-NA (SEQ ID NOS 24 and 25), respectively. In some embodiments, the first and second nucleotide sequences have at least 99% sequence identity to CS01-HC-NA and CS01-LC-NA (SEQ ID NOS 24 and 25), respectively. In some embodiments, the first and second nucleotide sequences have at least 99.5% sequence identity to CS01-HC-NA and CS01-LC-NA (SEQ ID NOS 24 and 25), respectively. In some embodiments, the first and second nucleotide sequences have at least 99.9% sequence identity to CS01-HC-NA and CS01-LC-NA (SEQ ID NOS 24 and 25), respectively. In some embodiments, the first and second nucleotide sequences are identical to CS01-HC-NA and CS01-LC-NA (SEQ ID NOS 24 and 25), respectively.

[0295] In some embodiments, the polypeptide linker of the Factor VIII construct is encoded by a third nucleotide sequence having high sequence identity to BDLO04 (SEQ ID NO: 6), which encodes the 14-amino acid linker corresponding to amino acids 760-773 of CS01-FL-AA (SEQ ID NO: 2). In some embodiments, the third nucleotide sequence has at least 95% identity to BDLO04 (SEQ ID NO: 6). In some embodiments, the third nucleotide sequence has at least 96% identity to BDLO04 (SEQ ID NO: 6). In some embodiments, the third nucleotide sequence has at least 97% identity to BDLO04 (SEQ ID NO: 6). In some embodiments, the third nucleotide sequence has at least 98% identity to BDLO04 (SEQ ID NO: 6). In some embodiments, the third nucleotide sequence is identical to BDLO04 (SEQ ID NO: 6).

[0296] In some embodiments, the codon-altered polynucleotide has a nucleotide sequence with high sequence identity to CS01-FL-NA (SEQ ID NO: 13). In some embodiments, the nucleotide sequence has at least 95% identity to CS01-FL-NA (SEQ ID NO: 13). In some embodiments, the nucleotide sequence has at least 96% identity to CS01-FL-NA (SEQ ID NO: 13). In some embodiments, the nucleotide sequence has at least 97% identity to CS01-FL-NA (SEQ ID NO: 13). In some embodiments, the nucleotide sequence has at least 98% identity to CS01-FL-NA (SEQ ID NO: 13). In some embodiments, the nucleotide sequence has at least 99% identity to CS01-FL-NA (SEQ ID NO: 13). In some embodiments, the nucleotide sequence has at least 99.5% identity to CS01-FL-NA (SEQ ID NO: 13). In some embodiments, the nucleotide sequence has at least 99.9% identity to CS01-FL-NA (SEQ ID NO: 13). In some embodiments, the nucleotide sequence is identical to CS01-FL-NA (SEQ ID NO: 13).

[0297] In some embodiments, the Factor VIII variant encoded by the codon-altered polynucleotide has an amino acid sequence with high sequence identity to CS01-FL-AA (SEQ ID NO: 2). In some embodiments, the amino acid

sequence has at least 97% identity to CS01-FL-AA (SEQ ID NO: 2). In some embodiments, the amino acid sequence has at least 98% identity to CS01-FL-AA (SEQ ID NO: 2). In some embodiments, the amino acid sequence has at least 99% identity to CS01-FL-AA (SEQ ID NO: 2). In some embodiments, the amino acid sequence has at least 99.5% identity to CS01-FL-AA (SEQ ID NO: 2). In some embodiments, the amino acid sequence has at least 99.9% identity to CS01-FL-AA (SEQ ID NO: 2). In some embodiments, the amino acid sequence is identical to CS01-FL-AA (SEQ ID NO: 2).

[0298] In some embodiments, the Factor VIII variant encoded by the CS01 polynucleotide, having high sequence homology to CS01-FL-AA (e.g., at least 95%, 96%, 97%, 98%, 99%, 99.5%, or 99.9% identity), comprises one or more amino acid substitutions selected from m1, m2, m3, m4, and m5.

[0299] In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises an m1 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises an m2 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises an m3 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises an m4 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises an m5 amino acid substitution.

[0300] In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises m12 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises m13 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises m23 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises m24 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises m25 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises m34 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises m35 amino acid substitutions.

[0301] In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises m123 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises m234 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises m125 amino acid substitutions.

[0302] CS23 Codon Altered Polynucleotides

[0303] In one embodiment, the codon-altered polynucleotides provided herein include a nucleotide sequence encoding a Factor VIII variant polypeptide with a linker that is cleavable in vivo. The Factor VIII polypeptide includes a Factor VIII light chain, a Factor VIII heavy chain, and a polypeptide linker joining the C-terminus of the heavy chain to the N-terminus of the light chain. The heavy chain of the Factor VIII polypeptide is encoded by a first nucleotide sequence having high sequence identity to CS23-HC-NA (SEQ ID NO: 22), which is the portion of CS23-FL-NA (SEQ ID NO: 20) encoding for a Factor VIII heavy chain. The light chain of the Factor VIII polypeptide is encoded by

a second nucleotide sequence with high sequence identity to CS23-LC-NA (SEQ ID NO: 23), which is the portion of CS23-FL-NA (SEQ ID NO: 20) encoding for a Factor VIII light chain. The polypeptide linker includes a furin cleavage site, which allows for maturation in vivo (e.g., after expression in vivo or administration of the precursor polypeptide).

[0304] In some embodiments, the first and second nucleotide sequences have at least 95% sequence identity to CS23-HC-NA and CS23-LC-NA (SEQ ID NOS 22 and 23), respectively. In some embodiments, the first and second nucleotide sequences have at least 96% sequence identity to CS23-HC-NA and CS23-LC-NA (SEQ ID NOS 22 and 23), respectively. In some embodiments, the first and second nucleotide sequences have at least 97% sequence identity to CS23-HC-NA and CS23-LC-NA (SEQ ID NOS 22 and 23), respectively. In some embodiments, the first and second nucleotide sequences have at least 98% sequence identity to CS23-HC-NA and CS23-LC-NA (SEQ ID NOS 22 and 23), respectively. In some embodiments, the first and second nucleotide sequences have at least 99% sequence identity to CS23-HC-NA and CS23-LC-NA (SEQ ID NOS 22 and 23), respectively. In some embodiments, the first and second nucleotide sequences have at least 99.5% sequence identity to CS23-HC-NA and CS23-LC-NA (SEQ ID NOS 22 and 23), respectively. In some embodiments, the first and second nucleotide sequences have at least 99.9% sequence identity to CS23-HC-NA and CS23-LC-NA (SEQ ID NOS 22 and 23), respectively. In some embodiments, the first and second nucleotide sequences are identical to CS23-HC-NA and CS23-LC-NA (SEQ ID NOS 22 and 23), respectively.

[0305] In some embodiments, the polypeptide linker of the Factor VIII construct is encoded by a third nucleotide sequence having high sequence identity to BDLO04 (SEQ ID NO: 6), which encodes the 14-amino acid linker corresponding to amino acids 760-773 of CS23-FL-AA (SEQ ID NO: 21). In some embodiments, the third nucleotide sequence has at least 95% identity to BDLO04 (SEQ ID NO: 6). In some embodiments, the third nucleotide sequence has at least 96% identity to BDLO04 (SEQ ID NO: 6). In some embodiments, the third nucleotide sequence has at least 97% identity to BDLO04 (SEQ ID NO: 6). In some embodiments, the third nucleotide sequence has at least 98% identity to BDLO04 (SEQ ID NO: 6). In some embodiments, the third nucleotide sequence is identical to BDLO04 (SEQ ID NO: 6).

[0306] In some embodiments, the codon-altered polynucleotide has a nucleotide sequence with high sequence identity to CS23-FL-NA (SEQ ID NO: 20). In some embodiments, the nucleotide sequence has at least 95% identity to CS23-FL-NA (SEQ ID NO: 20). In some embodiments, the nucleotide sequence has at least 96% identity to CS23-FL-NA (SEQ ID NO: 20). In some embodiments, the nucleotide sequence has at least 97% identity to CS23-FL-NA (SEQ ID NO: 20). In some embodiments, the nucleotide sequence has at least 98% identity to CS23-FL-NA (SEQ ID NO: 20). In some embodiments, the nucleotide sequence has at least 99% identity to CS23-FL-NA (SEQ ID NO: 20). In some embodiments, the nucleotide sequence has at least 99.5% identity to CS23-FL-NA (SEQ ID NO: 20). In some embodiments, the nucleotide sequence has at least 99.9% identity to CS23-FL-NA (SEQ ID NO: 20). In some embodiments, the nucleotide sequence is identical to CS23-FL-NA (SEQ ID NO: 20).

[0307] In some embodiments, the Factor VIII variant encoded by the codon-altered polynucleotide has an amino acid sequence with high sequence identity to CS23-FL-AA (SEQ ID NO: 21). In some embodiments, the amino acid sequence has at least 97% identity to CS23-FL-AA (SEQ ID NO: 21). In some embodiments, the amino acid sequence has at least 98% identity to CS23-FL-AA (SEQ ID NO: 21). In some embodiments, the amino acid sequence has at least 99% identity to CS23-FL-AA (SEQ ID NO: 21). In some embodiments, the amino acid sequence has at least 99.5% identity to CS23-FL-AA (SEQ ID NO: 21). In some embodiments, the amino acid sequence has at least 99.9% identity to CS23-FL-AA (SEQ ID NO: 21). In some embodiments, the amino acid sequence is identical to CS23-FL-AA (SEQ ID NO: 21).

[0308] In some embodiments, the Factor VIII variant encoded by the CS23 polynucleotide, having high sequence homology to CS23-FL-AA (e.g., at least 95%, 96%, 97%, 98%, 99%, 99.5%, or 99.9% identity), comprises one or more amino acid substitutions selected from m1, m2, m3, m4, and m5.

[0309] In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises an m1 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises an m2 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises an m3 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises an m4 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises an m5 amino acid substitution.

[0310] In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises m12 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises m13 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises m23 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises m24 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises m25 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises m34 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises m35 amino acid substitutions.

[0311] In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises m123 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises m234 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises m125 amino acid substitutions.

[0312] D. Codon-Altered Polynucleotides Encoding a Single-Chain Factor VIII Protein

[0313] Factor VIII constructs in which the furin cleavage site located at the C-terminal end of the B-domain is removed retain activity as a single chain polypeptide, despite that normal maturation of the Factor VIII molecule cannot occur (Leyte et al. (1991)). Similarly, a B-domain deleted Factor VIII construct with an attenuated furin site (containing an R1664H amino acid substitution) is more

biologically active than the corresponding Factor VIII construct with a wild-type furin cleavage site (Siner et al. (2013)). Accordingly, in some embodiments, the codon-altered polynucleotides provided herein include a nucleotide sequence encoding a single-chain Factor VIII variant polypeptide. The single-chain Factor VIII polypeptide includes a Factor VIII light chain, a Factor VIII heavy chain, and a polypeptide linker joining the C-terminus of the heavy chain to the N-terminus of the light chain. The polypeptide linker does not include a furin cleavage site.

[0314] Single-Chain CS04 Codon Altered Polynucleotides

[0315] In one embodiment, the codon-altered polynucleotides provided herein include a nucleotide sequence encoding a single-chain Factor VIII variant polypeptide. The Factor VIII polypeptide includes a Factor VIII light chain, a Factor VIII heavy chain, and an optional polypeptide linker joining the C-terminus of the heavy chain to the N-terminus of the light chain. The heavy chain of the Factor VIII polypeptide is encoded by a first nucleotide sequence having high sequence identity to CS04-HC-NA (SEQ ID NO: 3), which is the portion of CS04-FL-NA (SEQ ID NO: 1) encoding for a Factor VIII heavy chain. The light chain of the Factor VIII polypeptide is encoded by a second nucleotide sequence with high sequence identity to CS04-LC-NA (SEQ ID NO: 4), which is the portion of CS04-FL-NA (SEQ ID NO: 1) encoding for a Factor VIII light chain. The optional polypeptide linker does not include a furin cleavage site.

[0316] In some embodiments, the first and second nucleotide sequences have at least 95% sequence identity to CS04-HC-NA and CS04-LC-NA (SEQ ID NOS 3 and 4), respectively. In some embodiments, the first and second nucleotide sequences have at least 96% sequence identity to CS04-HC-NA and CS04-LC-NA (SEQ ID NOS 3 and 4), respectively. In some embodiments, the first and second nucleotide sequences have at least 97% sequence identity to CS04-HC-NA and CS04-LC-NA (SEQ ID NOS 3 and 4), respectively. In some embodiments, the first and second nucleotide sequences have at least 98% sequence identity to CS04-HC-NA and CS04-LC-NA (SEQ ID NOS 3 and 4), respectively. In some embodiments, the first and second nucleotide sequences have at least 99% sequence identity to CS04-HC-NA and CS04-LC-NA (SEQ ID NOS 3 and 4), respectively. In some embodiments, the first and second nucleotide sequences have at least 99.5% sequence identity to CS04-HC-NA and CS04-LC-NA (SEQ ID NOS 3 and 4), respectively. In some embodiments, the first and second nucleotide sequences have at least 99.9% sequence identity to CS04-HC-NA and CS04-LC-NA (SEQ ID NOS 3 and 4), respectively.

[0317] In some embodiments, the codon-altered polynucleotide has a nucleotide sequence with high sequence identity to CS04-SC1-NA (SEQ ID NO: 9). In some embodiments, the nucleotide sequence has at least 95% identity to CS04-SC1-NA (SEQ ID NO: 9). In some embodiments, the nucleotide sequence has at least 96% identity to CS04-SC1-NA (SEQ ID NO: 9). In some embodiments, the nucleotide sequence has at least 97% identity to CS04-SC1-NA (SEQ ID NO: 9). In some embodiments, the nucleotide sequence has at least 98% identity to CS04-SC1-NA (SEQ ID NO: 9). In some embodiments, the nucleotide sequence has at least 99% identity to CS04-SC1-NA (SEQ ID NO: 9). In some

embodiments, the nucleotide sequence has at least 99.5% identity to CS04-SC1-NA (SEQ ID NO: 9). In some embodiments, the nucleotide sequence has at least 99.9% identity to CS04-SC1-NA (SEQ ID NO: 9). In some embodiments, the nucleotide sequence is identical to CS04-SC1-NA (SEQ ID NO: 9).

[0318] In some embodiments, the codon-altered polynucleotide has a nucleotide sequence with high sequence identity to CS04-SC2-NA (SEQ ID NO: 11). In some embodiments, the nucleotide sequence has at least 95% identity to CS04-SC2-NA (SEQ ID NO: 11). In some embodiments, the nucleotide sequence has at least 96% identity to CS04-SC2-NA (SEQ ID NO: 11). In some embodiments, the nucleotide sequence has at least 97% identity to CS04-SC2-NA (SEQ ID NO: 11). In some embodiments, the nucleotide sequence has at least 98% identity to CS04-SC2-NA (SEQ ID NO: 11). In some embodiments, the nucleotide sequence has at least 99% identity to CS04-SC2-NA (SEQ ID NO: 11). In some embodiments, the nucleotide sequence has at least 99.5% identity to CS04-SC2-NA (SEQ ID NO: 11). In some embodiments, the nucleotide sequence has at least 99.9% identity to CS04-SC2-NA (SEQ ID NO: 11). In some embodiments, the nucleotide sequence is identical to CS04-SC2-NA (SEQ ID NO: 11).

[0319] In some embodiments, the single-chain Factor VIII variant encoded by the codon-altered polynucleotide has an amino acid sequence with high sequence identity to CS04-SC1-AA (SEQ ID NO: 10; human Factor VIII Δ (760-1667) (SPI; HsFVIII Δ (741-1648), SPE)). In some embodiments, the Factor VIII variant encoded by the codon-altered polynucleotide has an amino acid sequence with high sequence identity to CS04-SC1-AA (SEQ ID NO: 10). In some embodiments, the amino acid sequence has at least 97% identity to CS04-SC1-AA (SEQ ID NO: 10). In some embodiments, the amino acid sequence has at least 98% identity to CS04-SC1-AA (SEQ ID NO: 10). In some embodiments, the amino acid sequence has at least 99% identity to CS04-SC1-AA (SEQ ID NO: 10). In some embodiments, the amino acid sequence has at least 99.5% identity to CS04-SC1-AA (SEQ ID NO: 10). In some embodiments, the amino acid sequence has at least 99.9% identity to CS04-SC1-AA (SEQ ID NO: 10). In some embodiments, the amino acid sequence is identical to CS04-SC1-AA (SEQ ID NO: 10).

[0320] In some embodiments, the Factor VIII variant encoded by the CS04-SC1 polynucleotide, having high sequence homology to CS04-SC1-AA (e.g., at least 95%, 96%, 97%, 98%, 99%, 99.5%, or 99.9% identity), comprises one or more amino acid substitutions selected from m1, m2, m3, m4, and m5.

[0321] In some embodiments, the single-chain Factor VIII variant encoded by the codon-altered polynucleotide has an amino acid sequence with high sequence identity to CS04-SC2-AA (SEQ ID NO: 12; human Factor VIII Δ (772-1667) (SPI; HsFVIII Δ (753-1648), SPE)). In some embodiments, the Factor VIII variant encoded by the codon-altered polynucleotide has an amino acid sequence with high sequence identity to CS04-SC2-AA (SEQ ID NO: 12). In some embodiments, the amino acid sequence has at least 97% identity to CS04-SC2-AA (SEQ ID NO: 12). In some embodiments, the amino acid sequence has at least 98% identity to CS04-SC2-AA (SEQ ID NO: 12). In some embodiments, the amino acid sequence has at least 99%

identity to CS04-SC2-AA (SEQ ID NO: 12). In some embodiments, the amino acid sequence has at least 99.5% identity to CS04-SC2-AA (SEQ ID NO: 12). In some embodiments, the amino acid sequence has at least 99.9% identity to CS04-SC2-AA (SEQ ID NO: 12). In some embodiments, the amino acid sequence is identical to CS04-SC2-AA (SEQ ID NO: 12).

[0322] In some embodiments, the single-chain Factor VIII variant encoded by the CS04-SC2 polynucleotide, having high sequence homology to CS04-SC2-AA (e.g., at least 95%, 96%, 97%, 98%, 99%, 99.5%, or 99.9% identity), comprises one or more amino acid substitutions selected from m1, m2, m3, m4, and m5.

[0323] In one embodiment, the single-chain Factor VIII variant encoded by the CS04 polynucleotide comprises an m1 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises an m2 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises an m3 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises an m4 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises an m5 amino acid substitution.

[0324] In one embodiment, the single-chain Factor VIII variant encoded by the CS04 polynucleotide comprises m12 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises m13 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises m23 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises m24 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises m25 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises m34 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises m35 amino acid substitutions.

[0325] In one embodiment, the single-chain Factor VIII variant encoded by the CS04 polynucleotide comprises m123 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises m234 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS04 polynucleotide comprises m125 amino acid substitutions.

[0326] Single-Chain CS01 Codon Altered Polynucleotides

[0327] In one embodiment, the codon-altered polynucleotides provided herein include a nucleotide sequence encoding a single-chain Factor VIII variant polypeptide. The Factor VIII polypeptide includes a Factor VIII light chain, a Factor VIII heavy chain, and an optional polypeptide linker joining the C-terminus of the heavy chain to the N-terminus of the light chain. The heavy chain of the Factor VIII polypeptide is encoded by a first nucleotide sequence having high sequence identity to CS01-HC-NA (SEQ ID NO: 24), which is the portion of CS01-FL-NA (SEQ ID NO: 13) encoding for a Factor VIII heavy chain. The light chain of the Factor VIII polypeptide is encoded by a second nucleotide sequence with high sequence identity to CS01-LC-NA (SEQ ID NO: 25), which is the portion of CS01-FL-NA

(SEQ ID NO: 13) encoding for a Factor VIII light chain. The optional polypeptide linker does not include a furin cleavage site.

[0328] In some embodiments, the first and second nucleotide sequences have at least 95% sequence identity to CS01-HC-NA and CS01-LC-NA (SEQ ID NOS 24 and 25), respectively. In some embodiments, the first and second nucleotide sequences have at least 96% sequence identity to CS01-HC-NA and CS01-LC-NA (SEQ ID NOS 24 and 25), respectively. In some embodiments, the first and second nucleotide sequences have at least 97% sequence identity to CS01-HC-NA and CS01-LC-NA (SEQ ID NOS 24 and 25), respectively. In some embodiments, the first and second nucleotide sequences have at least 98% sequence identity to CS01-HC-NA and CS01-LC-NA (SEQ ID NOS 24 and 25), respectively. In some embodiments, the first and second nucleotide sequences have at least 99% sequence identity to CS01-HC-NA and CS01-LC-NA (SEQ ID NOS 24 and 25), respectively. In some embodiments, the first and second nucleotide sequences have at least 99.5% sequence identity to CS01-HC-NA and CS01-LC-NA (SEQ ID NOS 24 and 25), respectively. In some embodiments, the first and second nucleotide sequences have at least 99.9% sequence identity to CS01-HC-NA and CS01-LC-NA (SEQ ID NOS 24 and 25), respectively. In some embodiments, the first and second nucleotide sequences are identical to CS01-HC-NA and CS01-LC-NA (SEQ ID NOS 24 and 25), respectively.

[0329] In some embodiments, the codon-altered polynucleotide has a nucleotide sequence with high sequence identity to CS01-SC1-NA (SEQ ID NO: 26). In some embodiments, the nucleotide sequence has at least 95% identity to CS01-SC1-NA (SEQ ID NO: 26). In some embodiments, the nucleotide sequence has at least 96% identity to CS01-SC1-NA (SEQ ID NO: 26). In some embodiments, the nucleotide sequence has at least 97% identity to CS01-SC1-NA (SEQ ID NO: 26). In some embodiments, the nucleotide sequence has at least 98% identity to CS01-SC1-NA (SEQ ID NO: 26). In some embodiments, the nucleotide sequence has at least 99% identity to CS01-SC1-NA (SEQ ID NO: 26). In some embodiments, the nucleotide sequence has at least 99.5% identity to CS01-SC1-NA (SEQ ID NO: 26). In some embodiments, the nucleotide sequence has at least 99.9% identity to CS01-SC1-NA (SEQ ID NO: 26). In some embodiments, the nucleotide sequence is identical to CS01-SC1-NA (SEQ ID NO: 26).

[0330] In some embodiments, the codon-altered polynucleotide has a nucleotide sequence with high sequence identity to CS01-SC2-NA (SEQ ID NO: 27). In some embodiments, the nucleotide sequence has at least 95% identity to CS01-SC2-NA (SEQ ID NO: 27). In some embodiments, the nucleotide sequence has at least 96% identity to CS01-SC2-NA (SEQ ID NO: 27). In some embodiments, the nucleotide sequence has at least 97% identity to CS01-SC2-NA (SEQ ID NO: 27). In some embodiments, the nucleotide sequence has at least 98% identity to CS01-SC2-NA (SEQ ID NO: 27). In some embodiments, the nucleotide sequence has at least 99% identity to CS01-SC2-NA (SEQ ID NO: 27). In some embodiments, the nucleotide sequence has at least 99.5% identity to CS01-SC2-NA (SEQ ID NO: 27). In some embodiments, the nucleotide sequence has at least 99.9%

identity to CS01-SC2-NA (SEQ ID NO: 27). In some embodiments, the nucleotide sequence is identical to CS01-SC2-NA (SEQ ID NO: 27).

[0331] In some embodiments, the single-chain Factor VIII variant encoded by the codon-altered polynucleotide has an amino acid sequence with high sequence identity to CS01-SC1-AA (SEQ ID NO: 10; human Factor VIII Δ (760-1667) (SPI; HsFVIII Δ (741-1648), SPE)). In some embodiments, the Factor VIII variant encoded by the codon-altered polynucleotide has an amino acid sequence with high sequence identity to CS01-SC1-AA (SEQ ID NO: 10). In some embodiments, the amino acid sequence has at least 97% identity to CS01-SC1-AA (SEQ ID NO: 10). In some embodiments, the amino acid sequence has at least 98% identity to CS01-SC1-AA (SEQ ID NO: 10). In some embodiments, the amino acid sequence has at least 99% identity to CS01-SC1-AA (SEQ ID NO: 10). In some embodiments, the amino acid sequence has at least 99.5% identity to CS01-SC1-AA (SEQ ID NO: 10). In some embodiments, the amino acid sequence has at least 99.9% identity to CS01-SC1-AA (SEQ ID NO: 10). In some embodiments, the amino acid sequence is identical to CS01-SC1-AA (SEQ ID NO: 10).

[0332] In some embodiments, the Factor VIII variant encoded by the CS01-SC1 polynucleotide, having high sequence homology to CS01-SC1-AA (e.g., at least 95%, 96%, 97%, 98%, 99%, 99.5%, or 99.9% identity), comprises one or more amino acid substitutions selected from m1, m2, m3, m4, and m5.

[0333] In some embodiments, the single-chain Factor VIII variant encoded by the codon-altered polynucleotide has an amino acid sequence with high sequence identity to CS01-SC2-AA (SEQ ID NO: 12; human Factor VIII Δ (772-1667) (SPI; HsFVIII Δ (753-1648), SPE)). In some embodiments, the Factor VIII variant encoded by the codon-altered polynucleotide has an amino acid sequence with high sequence identity to CS01-SC2-AA (SEQ ID NO: 12). In some embodiments, the amino acid sequence has at least 97% identity to CS01-SC2-AA (SEQ ID NO: 12). In some embodiments, the amino acid sequence has at least 98% identity to CS01-SC2-AA (SEQ ID NO: 12). In some embodiments, the amino acid sequence has at least 99% identity to CS01-SC2-AA (SEQ ID NO: 12). In some embodiments, the amino acid sequence has at least 99.5% identity to CS01-SC2-AA (SEQ ID NO: 12). In some embodiments, the amino acid sequence has at least 99.9% identity to CS01-SC2-AA (SEQ ID NO: 12). In some embodiments, the amino acid sequence is identical to CS01-SC2-AA (SEQ ID NO: 12).

[0334] In some embodiments, the single-chain Factor VIII variant encoded by the CS01-SC2 polynucleotide, having high sequence homology to CS01-SC2-AA (e.g., at least 95%, 96%, 97%, 98%, 99%, 99.5%, or 99.9% identity), comprises one or more amino acid substitutions selected from m1, m2, m3, m4, and m5.

[0335] In one embodiment, the single-chain Factor VIII variant encoded by the CS01 polynucleotide comprises an m1 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises an m2 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises an m3 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises an m4 amino acid substitution. In one

embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises an m5 amino acid substitution.

[0336] In one embodiment, the single-chain Factor VIII variant encoded by the CS01 polynucleotide comprises m12 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises m13 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises m23 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises m24 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises m25 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises m34 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises m35 amino acid substitutions.

[0337] In one embodiment, the single-chain Factor VIII variant encoded by the CS01 polynucleotide comprises m123 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises m234 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS01 polynucleotide comprises m125 amino acid substitutions.

[0338] Single-Chain CS23 Codon Altered Polynucleotides

[0339] In one embodiment, the codon-altered polynucleotides provided herein include a nucleotide sequence encoding a single-chain Factor VIII variant polypeptide. The Factor VIII polypeptide includes a Factor VIII light chain, a Factor VIII heavy chain, and an optional polypeptide linker joining the C-terminus of the heavy chain to the N-terminus of the light chain. The heavy chain of the Factor VIII polypeptide is encoded by a first nucleotide sequence having high sequence identity to CS23-HC-NA (SEQ ID NO: 22), which is the portion of CS23-FL-NA (SEQ ID NO: 20) encoding for a Factor VIII heavy chain. The light chain of the Factor VIII polypeptide is encoded by a second nucleotide sequence with high sequence identity to CS23-LC-NA (SEQ ID NO: 23), which is the portion of CS23-FL-NA (SEQ ID NO: 20) encoding for a Factor VIII light chain. The optional polypeptide linker does not include a furin cleavage site.

[0340] In some embodiments, the first and second nucleotide sequences have at least 95% sequence identity to CS23-HC-NA and CS23-LC-NA (SEQ ID NOS 22 and 23), respectively. In some embodiments, the first and second nucleotide sequences have at least 96% sequence identity to CS23-HC-NA and CS23-LC-NA (SEQ ID NOS 22 and 23), respectively. In some embodiments, the first and second nucleotide sequences have at least 97% sequence identity to CS23-HC-NA and CS23-LC-NA (SEQ ID NOS 22 and 23), respectively. In some embodiments, the first and second nucleotide sequences have at least 98% sequence identity to CS23-HC-NA and CS23-LC-NA (SEQ ID NOS 22 and 23), respectively. In some embodiments, the first and second nucleotide sequences have at least 99% sequence identity to CS23-HC-NA and CS23-LC-NA (SEQ ID NOS 22 and 23), respectively. In some embodiments, the first and second nucleotide sequences have at least 99.5% sequence identity to CS23-HC-NA and CS23-LC-NA (SEQ ID NOS 22 and 23), respectively. In some embodiments, the first and second nucleotide sequences have at least 99.9% sequence identity to CS23-HC-NA and CS23-LC-NA (SEQ ID NOS 22 and 23), respectively.

23), respectively. In some embodiments, the first and second nucleotide sequences are identical to CS23-HC-NA and CS23-LC-NA (SEQ ID NOS 22 and 23), respectively.

[0341] In some embodiments, the codon-altered polynucleotide has a nucleotide sequence with high sequence identity to CS23-SC1-NA (SEQ ID NO: 28). In some embodiments, the nucleotide sequence has at least 95% identity to CS23-SC1-NA (SEQ ID NO: 28). In some embodiments, the nucleotide sequence has at least 96% identity to CS23-SC1-NA (SEQ ID NO: 28). In some embodiments, the nucleotide sequence has at least 97% identity to CS23-SC1-NA (SEQ ID NO: 28). In some embodiments, the nucleotide sequence has at least 98% identity to CS23-SC1-NA (SEQ ID NO: 28). In some embodiments, the nucleotide sequence has at least 99% identity to CS23-SC1-NA (SEQ ID NO: 28). In some embodiments, the nucleotide sequence has at least 99.5% identity to CS23-SC1-NA (SEQ ID NO: 28). In some embodiments, the nucleotide sequence has at least 99.9% identity to CS23-SC1-NA (SEQ ID NO: 28). In some embodiments, the nucleotide sequence is identical to CS23-SC1-NA (SEQ ID NO: 28).

[0342] In some embodiments, the codon-altered polynucleotide has a nucleotide sequence with high sequence identity to CS23-SC2-NA (SEQ ID NO: 29). In some embodiments, the nucleotide sequence has at least 95% identity to CS23-SC2-NA (SEQ ID NO: 29). In some embodiments, the nucleotide sequence has at least 96% identity to CS23-SC2-NA (SEQ ID NO: 29). In some embodiments, the nucleotide sequence has at least 97% identity to CS23-SC2-NA (SEQ ID NO: 29). In some embodiments, the nucleotide sequence has at least 98% identity to CS23-SC2-NA (SEQ ID NO: 29). In some embodiments, the nucleotide sequence has at least 99% identity to CS23-SC2-NA (SEQ ID NO: 29). In some embodiments, the nucleotide sequence has at least 99.5% identity to CS23-SC2-NA (SEQ ID NO: 29). In some embodiments, the nucleotide sequence has at least 99.9% identity to CS23-SC2-NA (SEQ ID NO: 29). In some embodiments, the nucleotide sequence is identical to CS23-SC2-NA (SEQ ID NO: 29).

[0343] In some embodiments, the single-chain Factor VIII variant encoded by the codon-altered polynucleotide has an amino acid sequence with high sequence identity to CS23-SC1-AA (SEQ ID NO: 10; human Factor VIIIΔ(760-1667) (SPI; CS04Δ(741-1648), SPE)). In some embodiments, the Factor VIII variant encoded by the codon-altered polynucleotide has an amino acid sequence with high sequence identity to CS23-SC1-AA (SEQ ID NO: 10). In some embodiments, the amino acid sequence has at least 97% identity to CS23-SC1-AA (SEQ ID NO: 10). In some embodiments, the amino acid sequence has at least 98% identity to CS23-SC1-AA (SEQ ID NO: 10). In some embodiments, the amino acid sequence has at least 99% identity to CS23-SC1-AA (SEQ ID NO: 10). In some embodiments, the amino acid sequence has at least 99.5% identity to CS23-SC1-AA (SEQ ID NO: 10). In some embodiments, the amino acid sequence has at least 99.9% identity to CS23-SC1-AA (SEQ ID NO: 10). In some embodiments, the amino acid sequence is identical to CS23-SC1-AA (SEQ ID NO: 10).

[0344] In some embodiments, the Factor VIII variant encoded by the CS23-SC1 polynucleotide, having high sequence homology to CS23-SC1-AA (e.g., at least 95%,

96%, 97%, 98%, 99%, 99.5%, or 99.9% identity), comprises one or more amino acid substitutions selected from m1, m2, m3, m4, and m5.

[0345] In some embodiments, the single-chain Factor VIII variant encoded by the codon-altered polynucleotide has an amino acid sequence with high sequence identity to CS23-SC2-AA (SEQ ID NO: 12; human Factor VIIIΔ(772-1667) (SPI; HsFVIIIΔ(753-1648), SPE)). In some embodiments, the Factor VIII variant encoded by the codon-altered polynucleotide has an amino acid sequence with high sequence identity to CS23-SC2-AA (SEQ ID NO: 12). In some embodiments, the amino acid sequence has at least 97% identity to CS23-SC2-AA (SEQ ID NO: 12). In some embodiments, the amino acid sequence has at least 98% identity to CS23-SC2-AA (SEQ ID NO: 12). In some embodiments, the amino acid sequence has at least 99% identity to CS23-SC2-AA (SEQ ID NO: 12). In some embodiments, the amino acid sequence has at least 99.5% identity to CS23-SC2-AA (SEQ ID NO: 12). In some embodiments, the amino acid sequence has at least 99.9% identity to CS23-SC2-AA (SEQ ID NO: 12). In some embodiments, the amino acid sequence is identical to CS23-SC2-AA (SEQ ID NO: 12).

[0346] In some embodiments, the single-chain Factor VIII variant encoded by the CS23-SC2 polynucleotide, having high sequence homology to CS23-SC2-AA (e.g., at least 95%, 96%, 97%, 98%, 99%, 99.5%, or 99.9% identity), comprises one or more amino acid substitutions selected from m1, m2, m3, m4, and m5.

[0347] In one embodiment, the single-chain Factor VIII variant encoded by the CS23 polynucleotide comprises an m1 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises an m2 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises an m3 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises an m4 amino acid substitution. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises an m5 amino acid substitution.

[0348] In one embodiment, the single-chain Factor VIII variant encoded by the CS23 polynucleotide comprises m12 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises m13 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises m23 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises m24 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises m25 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises m34 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises m35 amino acid substitutions.

[0349] In one embodiment, the single-chain Factor VIII variant encoded by the CS23 polynucleotide comprises m123 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises m234 amino acid substitutions. In one embodiment, the Factor VIII variant encoded by the CS23 polynucleotide comprises m125 amino acid substitutions.

[0350] E. Factor VIII Expression Vectors

[0351] In some embodiments, the codon-altered polynucleotides described herein are integrated into expression vectors. Non-limiting examples of expression vectors include viral vectors (e.g., vectors suitable for gene therapy), plasmid vectors, bacteriophage vectors, cosmids, phagemids, artificial chromosomes, and the like.

[0352] Non-limiting examples of viral vectors include: retrovirus, e.g., Moloney murine leukemia virus (MMLV), Harvey murine sarcoma virus, murine mammary tumor virus, and Rous sarcoma virus; adenoviruses, adeno-associated viruses; SV40-type viruses; polyomaviruses; Epstein-Barr viruses; papilloma viruses; herpes viruses; vaccinia viruses; and polio viruses.

[0353] In some embodiments, the codon-altered polynucleotides described herein are integrated into a gene therapy vector. In some embodiments, the gene therapy vector is a retrovirus, and particularly a replication-deficient retrovirus. Protocols for the production of replication-deficient retroviruses are known in the art. For review, see Kriegler, M., Gene Transfer and Expression, A Laboratory Manual, W.H. Freeman Co., New York (1990) and Murry, E. J., Methods in Molecular Biology, Vol. 7, Humana Press, Inc., Clifton, N.J. (1991).

[0354] In one embodiment, the gene therapy vector is an adeno-associated virus (AAV) based gene therapy vector. AAV systems have been described previously and are generally well known in the art (Kelleher and Vos, *Biotechniques*, 17(6):1110-17 (1994); Cotten et al., *Proc Natl Acad Sci USA*, 89(13):6094-98 (1992); Curiel, *Nat Immun*, 13(2-3):141-64 (1994); Muzyczka, *Curr Top Microbiol Immunol*, 158:97-129 (1992); and Asokan A, et al., *Mol. Ther.*, 20(4): 699-708 (2012), each incorporated herein by reference in their entireties for all purposes). Details concerning the generation and use of rAAV vectors are described, for example, in U.S. Pat. Nos. 5,139,941 and 4,797,368, each incorporated herein by reference in their entireties for all purposes. In a particular embodiment, the AAV vector is an AAV-8 vector.

[0355] In some embodiments, the codon-altered polynucleotides described herein are integrated into a retroviral expression vector. These systems have been described previously, and are generally well known in the art (Mann et al., *Cell*, 33:153-159, 1983; Nicolas and Rubinstein, In: *Vectors: A survey of molecular cloning vectors and their uses*, Rodriguez and Denhardt, eds., Stoneham: Butterworth, pp. 494-513, 1988; Temin, In: *Gene Transfer*, Kucherlapati (ed.), New York: Plenum Press, pp. 149-188, 1986). In a specific embodiment, the retroviral vector is a lentiviral vector (see, for example, Naldini et al., *Science*, 272(5259): 263-267, 1996; Zufferey et al., *Nat Biotechnol*, 15(9):871-875, 1997; Blomer et al., *J Virol.*, 71(9):6641-6649, 1997; U.S. Pat. Nos. 6,013,516 and 5,994,136).

[0356] A wide variety of vectors can be used for the expression of a Factor VIII polypeptide from a codon-altered polypeptide in cell culture, including eukaryotic and prokaryotic expression vectors. In certain embodiments, a plasmid vector is contemplated for use in expressing a Factor VIII polypeptide in cell culture. In general, plasmid vectors containing replicon and control sequences which are derived from species compatible with the host cell are used in connection with these hosts. The vector can carry a replication site, as well as marking sequences which are capable of providing phenotypic selection in transformed cells. The plasmid will include the codon-altered polynucle-

otide encoding the Factor VIII polypeptide, operably linked to one or more control sequences, for example, a promoter. [0357] Non-limiting examples of vectors for prokaryotic expression include plasmids such as pRSET, pET, pBAD, etc., wherein the promoters used in prokaryotic expression vectors include lac, trc, trp, recA, arabad, etc. Examples of vectors for eukaryotic expression include: (i) for expression in yeast, vectors such as pAO, pPIC, pYES, pMET, using promoters such as AOX1, GAP, GAL1, AUG1, etc; (ii) for expression in insect cells, vectors such as pMT, pAc5, pIB, pMIB, pBAC, etc., using promoters such as PH, p10, MT, Ac5, OpIE2, gp64, polh, etc., and (iii) for expression in mammalian cells, vectors such as pSVL, pCMV, pRc/RSV, pcDNA3, pBPV, etc., and vectors derived from viral systems such as vaccinia virus, adeno-associated viruses, herpes viruses, retroviruses, etc., using promoters such as CMV, SV40, EF-1, UbC, RSV, ADV, BPV, and β-actin.

IV. Examples

Example 1—Construction of a Codon Altered Factor VIII Variant Expression Sequence

[0358] Two hurdles had to be overcome in order to create a Factor VIII coding sequence that is effective for gene therapy of hemophilia A. First, because of the genomic size limitations of conventional gene therapy delivery vectors (e.g., AAV virions), the encoded Factor VIII polypeptide had to be shortened considerably. Second, the coding sequence had to be altered to: (i) stabilize packaging interactions within the delivery vector, (ii) stabilize the mRNA intermediary, and (iii) improve the robustness of transcription/translation of the mRNA.

[0359] To achieve the first objective, Applicants started with a B-domain deleted Factor VIII variant construct, referred to herein as “FVIII-BDD-SQ.” In this construct, the B-domain is replaced with a fourteen amino acid sequence referred to as the “SQ” sequence. Recombinant FVIII-BDD-SQ is sold under the trade name REFACTO®, and has been shown to be effective for the management of hemophilia A. However, the native coding sequence for FVIII-BDD-SQ, which includes human wild-type nucleic acid sequences for the Factor VIII heavy and light chains, is ineffectively expressed in gene therapy vectors.

[0360] To address the poor expression of the native FVIII-BDD-SQ, the codon optimization algorithm described in Fath et al. (PLoS ONE, 6:e17596 (2011)), modified as described in Ward et al. (Blood, 117:798 (2011)) and in McIntosh et al. (Blood, 121, 3335-3344 (2013)), was applied to the FVIII-BDD-SQ sequence to create first intermediate coding sequence CS04a. However, Applicants recognized that the CS04a sequence created using the modified algorithm could be improved by further modifying the sequence. Accordingly, Applicants re-introduced CpG dinucleotides, re-introduced the CGC codon for arginine, changed the leucine and serine codon distributions, re-introduced highly conserved codon pairs, and removed cryptic TATA box, CCAAT box, and splice site elements, while avoiding CpG islands and local overrepresentation of AT-rich and GC-rich stretches.

[0361] First, the modified algorithm systematically replaces codons containing CpG-dinucleotides (e.g., arginine codons) with non-CpG-dinucleotide codons, and eliminates/avoids CpG-dinucleotides created by neighboring codons. This strict avoidance of CpG dinucleotides is usu-

ally done to prevent TLR-induced immunity after intramuscular injection of DNA vaccines. However, doing so limits the codon optimization possibilities. For example, the modified algorithm excludes use of the complete set of CGX arginine codons. This is particularly disruptive in the coding of genes for expression in human cells, because CGC is the most frequently used arginine codon in highly expressed human genes. Additionally, avoiding the creation of CpGs by neighboring codons further limits the optimization possibilities (e.g., limits the number of codon pairs that may be used together).

[0362] Because TLR-induced immunity is not expected to be a problem associated with liver-directed, AAV-based gene therapy, codons including CpGs, and neighboring codons creating CpGs, were re-introduced into intermediate coding sequence CS04a, preferentially in the sequence coding for the Factor VIII light chain (e.g., at the 3' end of the FVIII-BDD-SQ coding sequence). This allowed for more frequent use of preferred human codons, particularly those for arginine. Care was taken, however, to avoid creation of CpG islands, which are regions of coding sequence having a high frequency of CpG sites. This is contrary to the teachings of Krinner et al. (Nucleic Acids Res., 42(6):3551-64 (2014)), which suggests that CpG domains downstream of transcriptional start sites promote high levels of gene expression.

[0363] Second, the modified algorithm applies certain codons exclusively, such as CTG for leucine, GTG for valine, and CAG for glutamine. However, this offends the principles of balanced codon use, for example, as proposed in Haas et al. (Current Biology, 6(3):315-24 (1996)). To account for the overuse of preferred codons by the modified algorithm, alternate leucine codons were re-introduced where allowed by the other rules applied to the codon alteration (e.g., CpG frequency and GC content).

[0364] Third, the modified algorithm replaces codon pairs without regard to how conserved they are in nature, when certain criteria (e.g., the presence of CG-dinucleotides) are met. To account for beneficial properties which may have been conserved by evolution, the most conserved codon pairs that were replaced by the algorithm and the most conserved preferred codon pairs, e.g., as described in Tats et al. (BMC Genomics 9:463 (2008)), were analyzed and adjusted where allowed by the other rules applied to the codon alteration (e.g., CpG frequency and GC content).

[0365] Fourth, serine codons used in the intermediate coding sequence were also re-engineered. Specifically, AGC, TCC, and TCT serine codons were introduced into the modified coding sequence with higher frequency, to better match overall for human codon usage (Haas et al., *supra*).

[0366] Fifth, TATA box, CCAAT box elements, and intron/exon splice sites were screened and removed from the modified coding sequence. When modifying the coding sequence, care was taken to avoid local overrepresentation of AT-rich or GC rich stretches.

[0367] Finally, in addition to optimizing the codon usage within the coding sequence, the structural requirements of the underlying AAV virion were considered when further refining the intermediate coding sequence CS04a. AAV vectors (e.g., the nucleic acid portion of an AAV virion) are packaged as single stranded DNA molecules into their capsids (for review, see, Daya and Berns, Clin. Microbiol Rev., 21(4):583-93 (2008)). The GC content of the vector is therefore likely to influence packaging of the genome and,

thus, vector yields during production. Like many algorithms, the modified algorithm used here creates an optimized gene sequence with a GC content of at least 60% (see, Fath et al., PLoS One, 6(3):e17596 (2011) (erratum in: PLoS One, (6)3 (2011)). However, the AAV8 capsid protein is encoded by a nucleotide sequence having a lower GC content of about 56%. Thus, to better mimic the native AAV8 capsid protein coding sequence, the GC content of the intermediate coding sequence CS04a was reduced to 56%.

[0368] The resulting CS04 coding sequence, shown in FIG. 2, has an overall GC content of 56%. The CpG-dinucleotide content of the sequence is moderate. However, CpG dinucleotides are predominantly present in the downstream portion of the coding sequence, e.g., the portion coding for the Factor VIII light chain. The CS04 sequence has 79.77% nucleotide sequence identity to the corresponding coding sequences in wild-type Factor VIII (Genbank accession M14113).

[0369] For comparison purposes, several other codon-optimized, ReFacto constructs were prepared. CS01 was constructed by applying the codon-optimization algorithm of Fath et al., as modified by Ward et al., as done for CS04. However, unlike CS04, the CS01 construct does not contain any CpG islands. The CS08 ReFacto construct was codon-optimized as described in Radcliff P. M. et al., Gene Therapy, 15:289-97 (2008), the content of which is hereby expressly incorporated by reference herein, in its entirety, for all purposes. The CS10 codon-optimized ReFacto construct was obtained from Eurofins Genomics (Ebersberg, Germany). The CS11 codon-optimized ReFacto construct was obtained from Integrated DNA Technologies, Inc. (Corvalville, USA). The CH25 codon-optimized ReFacto construct was obtained from ThermoFischer Scientific's GeneArt services (Regensburg, Germany). The CS40 ReFacto construct consists of the wild type Factor VIII coding sequence. The algorithm used to construct CS23 is based on the JCAT tool (www.jcat.de), an on-line tool for codon-optimizations (Grote et al., 2005; Nucl. Acids Res. W526-31). The sequence was further modified to more reflect the codon usage of the albumin superfamily (Mirsafian et al. 2014: Sc. Word Journal 2014, ID 639682). The sequence identities shared between each of the ReFacto coding sequences is shown in Table 2, below.

TABLE 2

Percent identity matrix for codon-altered Factor VIII constructs.							
	CS01	CS04	CS08	CS10	CS11	CS40	CH25
CS01	100%						
CS04	93.0%	100%					
CS08	80.7%	82.2%	100%				
CS10	79.1%	79.4%	78.4%	100%			
CS11	78.3%	78.3%	78.1%	77.5%	100%		
CS40	79.6%	79.8%	76.7%	77.6%	75.4%	100%	
CH25	81.3%	85.1%	85.0%	79.9%	79.4%	75.8%	100%
CS23	84.3%	89.2%	85.1%	80.3%	79.9	76.5%	93.2%
							100%

[0370] Plasmids of each construct were constructed by cloning different synthetic DNA fragments into the same vector backbone plasmid (pCh-BB01). DNA synthesis of the ReFacto-type BDD-FVIII fragments with flanking Ascl and NotI enzyme restriction sites were done by ThermoFischer Scientific (Regensburg, Germany). The vector backbone contains two flanking AAV2-derived inverted terminal

repeats (ITRs) that encompass a promoter/enhancer sequence derived from the liver-specific murine transthyretin gene, Ascl and NotI enzyme restriction sites for insertion of the respective ReFacto-type BDD-FVIII and a synthetic polyA site. After ligation of the prepared vector backbone and inserts via the Ascl and NotI sites, the resulting plasmids were amplified in milligram scale. The ReFacto-type BDD-FVIII sequences of the constructs were verified by direct sequencing (Microsynth, Balgach, Switzerland). The cloning resulted in seven different plasmid constructs named pCS40, pCS01, pCS04, pCS08, pCS10, pCS11, and pCh25 (FIG. 23). The constructs have the same vector backbone and encode the same B-domain deleted FVIII protein (ReFacto-type BDD-FVIII), but differ in their FVIII coding sequence.

[0371] AAV8-based vectors were prepared by the three plasmid transfection method, as described in Grieger J C, et al. (Virus Vectors Using Suspension HEK293 Cells and Continuous Harvest of Vector From the Culture Media for GMP FIX and FLT1 Clinical Vector, Mol Ther., Oct. 6. (2015) doi: 10.1038/mt.2015.187. [Epub ahead of print]), the content of which is hereby expressly incorporated by reference herein, in its entirety, for all purposes. HEK293 suspensions cells were used for plasmid transfections using the corresponding FVIII vector plasmid, the helper plasmid pXX6-80 (carrying adenoviral helper genes), and the packaging plasmid pGSK2/8 (contributing the rep2 and cap8 genes). To isolate the AAV8 constructs, the cell pellets of one liter cultures were processed using iodixanol gradients, as described in Grieger et al. (2015, Supra). The procedure resulted in vector preparations called vCS01, vCS04, vCS08, vCS10, vCS11, and vCH25. Vectors were quantified by qPCR using the universal qPCR procedure targeting the AAV2 inverted terminal repeats (Aurnhammer, Human Gene Therapy Methods: Part B 23:18-28 (2012)). A control vector plasmid carrying AAV2 inverted terminal repeats served for preparing the standard curve. The resulting vCS04 construct is presented as SEQ ID NO: 8 in FIGS. 7A-7C.

[0372] The integrity of the vector genomes was analyzed by AAV agarose gel electrophoresis. The electrophoresis was performed as described in Fagone et al., Human Gene Therapy Methods 23:1-7 (2012). Briefly, AAV vector prepa-

rations were incubated at 75° C. for 10 minutes in the presence of 0.5% SDS and then cooled down to room temperature. Approximately 1.5E10 vector genomes (vg) were loaded per lane on a 1% 1×TAE agarose gel and electrophoresed for 60 min at 7 V/cm of gel length. The gel was then stained in 2× GelRed (Biotium Cat#41003) solution and imaged by ChemiDocTMMP (Biorad). The results

shown in FIG. 24 demonstrate that the vCS01, vCS04, and vCS40 viral vectors have the same-sized genome, indicated by a distinct band in the 5 kb range (FIG. 24, lanes 2-4). Despite a vector size of approx. 5.2 kb, the genome is a homogenous band confirming correct packaging of the somewhat oversized genome (relative to an AAV wild-type genome of 4.7 kb). All other vCS vector preparations show the same genomic size (data not shown).

[0373] In order to confirm the expected pattern of capsid proteins, SDS PAGE followed by silver staining was performed with the vectors vCS01, vCS04, and vCS40 (FIG. 25). As shown in the figure, the downstream purification procedure resulted in highly purified material displaying the expected protein pattern of VP1, VP2 and VP3 (FIG. 25, lanes 2-4). The same pattern was seen with all other viral preparations (not shown). The SDS-PAGE procedure of AAV preparations was done according to standard procedures. Each lane contained 1E10 vg of the respective viral construct, and were separated on a 4-12% Bis-Tris (NuPAGE® Novex, Life Technologies) gel as per manufacturer's instructions. Silver staining was performed with a SilverQuest™ kit (Novex, Life Technologies) according to the manufacturer's instructions.

[0374] Surprisingly, AAV vectors vCS01 and vCS04 had higher virion packaging, measured by higher yields in AAV virus production, as compared to the vCS40 wild-type coding construct and the other codon-optimized constructs. As shown in Table 3, the vCS01 and vCS04 vectors replicated substantially better than vCS40, providing a 5-7 fold yield increase in AAV titer.

TABLE 3

Yields per liter cell culture obtained with AAV vector constructs vCS01, vCS04, and vCD40, as purified from cell pellets.			
Construct	Vector concentration [vg/ml] × 10E12	Yields [vg/liter] × 10E12	Fold increase vs wt
vCS40	2.0	11.0	—
vCS01	9.2	51.4	4.7
vCS04 - Sample 1	17.6	79.2	7.2
vCS04 - Sample 2	15.9	58.8	5.4

Example 2—In Vivo Expression of Codon Altered Factor VIII Variant Expression Sequences

[0375] To test the biological potency of the codon-altered Factor VIII variant sequences, the ReFacto-type FVIII constructs described in Example 1 were administered to mice lacking Factor VIII. Briefly, the assays were performed in C57Bl/6 FVIII knock-out (ko) mice (with 6-8 animals per group) by tail vein injection of 4E12 vector genomes (vg) per kilogram body weight of mouse. Blood was drawn 14 days after injection by retroorbital puncture and plasma was prepared and frozen using standard procedures. Expression levels at day 14 were chosen because there is minimal influence of inhibitory antibodies at this time, which are seen in some animals of this mouse model at later times. FVIII activity in the mouse plasma was determined using the Technochrome FVIII assay performed, with only minor modifications, as suggested by the manufacturer (Technoclone, Vienna, Austria). For the assay, the plasma samples were appropriately diluted and mixed with assay reagents, containing thrombin, activated factor IX (FIXa), phospho-

lipids, factor X and calcium. Following FVIII activation by thrombin a complex with FIXa, phospholipids and calcium is formed. This complex activates FX to activated FX (FXa) which in turn cleaves para-nitroanilide (pNA) from the chromogenic substrate. The kinetics of pNA formation is measured at 405 nm. The rate is directly proportional to the FVIII concentration in the sample. FVIII concentrations are read from a reference curve and results are given in IU FVIII/milliliter.

[0376] The results, presented in Table 4 below, demonstrate that the codon-altered sequences designed using commercial algorithms (CS10, CS11, and CH25) provided only a modest increase in BDD-Factor VIII (3-4 fold) as compared to the wild-type BDD-Factor VIII construct (CS40). Similarly, the codon-altered BDD-Factor VIII construct prepared as described in Radcliffe et al. (CS08), only provided a 3-4 fold increase in BDD-FVIII expression. This result is consistent with the results reported in Radcliff et al. Surprisingly, the CS01, CS04, and CS23 constructs provided much higher BDD-FVIII expression in the in-vivo biopotency assays (18-, 74-, and -30-fold increases, respectively).

TABLE 4

Expression of FVIII in the plasma of FVIII-knock-out mice induced by the different AAV vector constructs.					
Construct	Codon Algorithm	Average FVIII Expression at Day 14 [IU/ml]	Standard deviation	Number of mice	Fold increase vs wt
vCS40	Human wild-type	0.03	0.03	12	—
vCS01	Applicants'	0.55	0.28	22	18.3
vCS04	Applicants'	2.21	1.20	55	73.7
vCS08	Radcliffe et al.	0.11	0.01	6	3.6
vCS10	Eurofins	0.09	0.01	7	3.0
vCS11	IDT	0.08	0.02	8	2.7
vCH25	GeneArt	0.13	0.12	18	4.3
vCS23	Applicants'	0.91	0.32	5	30.3

Example 3—Design of Glycosylation Peptides for the B-Domain Substituted Linker

[0377] Others have shown that inclusion of a small peptide (the "V3 peptide") containing six putative N-linked glycosylation sites from the wild-type Factor VIII B-domain, into a B-domain deleted gene therapy construct, increased Factor VIII levels in the plasma of mice (McIntosh et al., Blood 121(17):3335-44 (2013)). However, in order to maintain the small size of the B-domain substituted linker, the glycosylation sites were taken out of the context of the wild-type B-domain. In silico prediction (Gupta et al., Supra) of the linker containing the V3 peptide suggests that only two of these glycosylation sites in the V3 peptide will be modified in vivo (FIG. 15).

[0378] Thus, Applicants attempted to identify alternative glycosylation peptides that would support higher levels of glycosylation in vivo, which matched wild type glycosylation more closely than the V3 peptide. Applicants designed and tested several alternative glycosylation peptides, in silico. Several of these peptides, shown in FIGS. 13A-13B, were predicted to have equal or greater glycosylation in vivo than the V3 peptide, when placed between amino acids N768 and P769 of the B-domain substituted linker in SEQ

ID NO:2. The results of the in silico predictions are shown in Table 5, below. Table 5 also reports the results of expression experiments performed for several constructs encoding a ReFacto-type Factor VIII protein with a glycosylation peptide incorporated into the B-domain substituted linker, in a CS01 codon-optimized background.

TABLE 5

Sequence	Number of Predicted N-glycosylation sites	Day 28 expression [IU/ml]	SD	Number of mice [n]	Fold expression
vCS01	0	0.74	0.52	5	21
vNG1/CS01	4	n.d.	—	—	—
vNG4/CS01	3	1.93	0.57	6	55
vNG5/CS01	2	n.d.	—	—	—
vNG6/CS01	1	0.80	0.67	5	23
vNG9/CS01	1	n.d.	—	—	—
vNG10/CS01	2	2.66	0.52	6	76
vNG16/CS01	2	1.59	0.57	6	45
vNG17/CS01	2	n.d.	—	—	—
vNG18/CS01	2	n.d.	—	—	—
vNG19/CS01	2	0.88	0.25	5	25
vNG20/CS01	2	n.d.	—	—	—
vNG21/CS01	2	n.d.	—	—	—
vCS40	0	0.035	0.030	12	1

[0379] AAV vectors containing the NG variants were constructed as described in Example 1 and tested in FVIII knock-out mice as described in Example 2. All virus vectors (except the control vector vCS40) shown in Table 5 are based on the algorithm as used in vCS01. A parallel set of constructs using the algorithm of vCS04 was also prepared (vNG/CS04 series) and is tested in the mouse model. Results were compared to the expression levels achieved with the wild-type vCS40 construct. The day 28 expression levels were chosen in this example, because expression levels of the majority of construct reached the highest levels at this time point. Three AAV vectors achieved greater than 40-fold FVIII expression levels including vNG4/CS01, vNG10/CS01 and vNG16/CS01 (Table 5). The corresponding constructs vNG4/CS04, vNG10/CS04 and vNG16/CS04 are expected to show even higher expression because they are based on the superior vCS04 algorithm.

[0380] Surprisingly, the AAV vectors of the vNG/CS01 series had higher virion packaging, measured by higher yields in AAV virus production, as compared to the vCS40 wild-type coding construct. As shown in Table 6, the vNG/CS01-based vectors replicated substantially better than vCS40, providing an approximately 3-fold yield increase in AAV titer.

TABLE 6

Sequence	Vector conc. [vg/ml] ×10 ¹²	Yields [vg/liter] ×10 ¹²	Fold increase vs wild-type
vCS01	9.17	51.35	4.7
vNG1/CS01	2.13	17.04	1.5
vNG4/CS01	5.74	33.01	3.0
vNG5/CS01	6.91	27.29	2.5
vNG6/CS01	7.01	40.66	3.7

TABLE 6-continued

Yields per liter cell culture obtained with AAV vector constructs as purified from cell pellets.			
Sequence	Vector conc. [vg/ml] ×10 ¹²	Yields [vg/liter] ×10 ¹²	Fold increase vs wild-type
vNG9/CS01	6.39	29.39	2.7
vNG10/CS01	8.57	37.71	3.4
vNG16/CS01	5.3	28.36	2.6
vNG17/CS01	4.24	32.22	2.9
vNG18/CS01	6.11	37.88	3.4
vNG19/CS01	9.42	39.56	3.6
vNG20/CS01	4.09	30.27	2.8
vNG21/CS01	n.d.	—	—
vCS40	2.03	11	1.0

Example 4—Construction of Mutant BDD-FVIII Constructs

[0381] Numerous different mutated Refacto-type BDD-FVIII constructs, carrying amino acid mutations within the Factor VIII heavy chain and/or B-domain substituted linker, were cloned and screened. The corresponding vectors, as referred to herein as the “vCS” series of vectors, encode BDD-FVIII variants in the CS01, CS04, and CS23 codon-altered backgrounds. The method used to construct the CS01 and CS04 backgrounds is described in Example 1. The method used to construct CS23 was based on the JCAT tool (www.jcat.de), an on-line tool for codon-optimizations (Grote et al., 2005; Nucl. Acids Res. W526-31). The sequence was further modified to better reflect the codon usage of the albumin superfamily (Mirsafian et al., Sc. Word Journal, ID 639682 (2014)), the content of which is hereby expressly incorporated by reference, in its entirety, for all purposes.

[0382] Combinations of three types of mutations were included in the FVIII sequences of the vCS series of constructs. The first amino acid change introduced into the FVIII sequence is the X1 mutation (TTYVNRSL (SEQ ID NO: 33); X. Xiao), which introduces an additional glycosylation site near the B-domain substituted linker. The X1 mutation is also referred to herein as the “m3” mutation. The second amino acid change made in the FVIII sequence includes the F328S (SPI, F3095 SPE) mutation, an amino acid change known to improve secretion of FVIII (Swaaroop, J. Biol. Chem., 272:24121-24 (1997)). This mutation is also referred to herein as the “m1” mutation. The third change is the so-called X5 mutation, which is a combination of five amino acid changes in the A1 domain of the heavy chain that improves specific activity and secretion of BDD-FVIII (Cao et al., 2014; ASGCT abstract #460; details of mutations disclosed in oral presentation). The X5 mutation is also referred to herein as the “m2” mutation. Next, combinations of X1 and F328S (SPI, F3095 SPE) were made, followed by combinations of X1 and X5, also referred to as “X6,” and yet other combinations of X5 and F328S (SPI, F3095 SPE) were made (Table 7).

[0383] Gene Synthesis and Cloning of the Vector Plasmids.

[0384] The plasmids were constructed by cloning different synthetic DNA fragments into the same vector backbone plasmid (pCh-BB01). DNA synthesis of the Refacto-type BDD-FVIII fragments with flanking AscI and NotI enzyme restriction sites were done by ThermoFischer Scientific

(Regensburg, Germany). The vector backbone contains two flanking AAV2-derived inverted terminal repeats (ITRs) that encompass a promoter/enhancer sequence derived from the liver-specific murine transthyretin gene, AscI and NotI enzyme restriction sites for insertion of the respective Refacto-type BDD-FVIII, and a synthetic polyA site. After ligation of the prepared vector backbone and insertions via the AscI and NotI sites, the resulting plasmids were amplified in milligram scale. The Refacto-type BDD-FVIII sequences of the constructs were verified by direct sequencing (Microsynth, Balgach, Switzerland). The cloning resulted in different plasmid constructs, as shown in FIG. 44.

[0385] Small Scale Vector Preparations and Quantification by Quantitative PCR (qPCR).

[0386] AAV8-based vectors were prepared by the three plasmid transfection method essentially as described in Grieger et al. (2015, Supra). HEK293 suspension cells were used for plasmid transfections using the corresponding FVIII vector plasmid, the helper plasmid pXX6X80 (carrying adenoviral helper genes) and the packaging plasmid pGSK2/8 (contributing the rep2 and cap8 genes). In the downstream process the cell pellet of a one liter culture was processed using iodixanol gradients as described above. The procedure resulted in vector preparations as outlined in Table 8. Vectors were quantified by qPCR using the universal qPCR procedure targeting the AAV2 inverted terminal repeats (Aurnhammer, HUMAN GENE THERAPY METHODS: Part B 23:18-28 (2012)). An accurately quantified vector plasmid carrying AAV2 Inverted terminal repeats served for preparing the standard curve.

[0387] AAV Vector Characterizations.

[0388] The integrity of the vector genome was analyzed by AAV agarose gel electrophoresis. The electrophoresis was done similar as described in Fagone et al. (Human Gene Therapy Methods, 23:1-7 (2012)). AAV vector preparations were incubated at 75° C. for 10 minutes in the presence of 0.5% SDS and then cooled down to room temperature. Approximately 1.5E10 vector genomes (vg) were loaded per lane on a 1% 1×TAE agarose gel and electrophoresed for 60 min at 7 V/cm of gel length. The gel was then stained in 2× GelRed (Biotium Cat#41003) solution and imaged by ChemiDoc™ MP (Biorad). The results of a selection of vectors are shown in FIG. 45. The viral vectors vCS04 (control), vCS17, vCS20, vCS24, vCS16 and vCS40 (control) show all the same-sized genome as a distinct band in the 5 kb range (FIG. 45, lanes 2-7; arrow right side). Despite a vector size of approx. 5.2 kb, the genome is a homogenous band confirming correct packaging of the somewhat oversized genome (relative to an AAV wild-type genome of 4.7 kb).

[0389] In order to confirm purity of the vector and the expected pattern of capsid proteins, SDS PAGE followed by silver staining was performed with the vectors, as shown in FIG. 46. As shown in the figure, the downstream purification procedure resulted in highly purified material displaying the expected protein pattern of VP1, VP2 and VP3 (FIG. 46 lanes 2-9; arrows right hand side). The SDS-PAGE procedure of AAV preparations was done according to standard procedures. The amounts of 1E10 vg per lane were separated on a 4-12% Bis-Tris (NuPAGE® Novex, Life Technologies) gel as per manufacturer's instructions. Silver staining was performed with a SilverQuest™ kit (Novex, Life Technologies) according to the instructions of the manufacturer.

[0390] In-vivo biopotency screening of vectors. The different Refacto-type BDD-FVIII constructs were screened in mice. The assay was performed in C57Bl/6 FVIII knock-out (ko) mice (with 6-8 animals per group) by tail vein injection of 4E12 vector genomes (vg) per kilogram body weight of mouse. Blood was drawn 14 days after injection by retroorbital puncture and plasma was prepared and frozen using standard procedures. FVIII activity in mouse plasma was determined with a chromogenic assay from Technoclone with minor modifications (Technochrome FVIII, Technoclone, Vienna, Austria). In brief, the plasma sample was appropriately diluted and mixed with assay reagents, containing thrombin, activated factor IX (FIXa), phospholipids, factor X and calcium. Following FVIII activation by thrombin a complex with FIXa, phospholipids and calcium is formed. This complex activates FX to activated FX (FXa) which in turn cleaves para-nitroanilide (pNA) from the chromogenic substrate. The kinetics of pNA formation is measured at 405 nm. The rate is directly proportional to the FVIII concentration in the sample. FVIII concentrations are read from a reference curve and results are given in IU FVIII/milliliter.

[0391] The results of the mouse biopotency assay (day 14 expression data of FVIII in international units per milliliter [IU/ml] in mouse plasma and fold expression compared to the wild-type vCS40 control) are shown in Table 7. AAV vectors vCS19, vCS26 and vCS32 all contain the X1 glycosylation site in the CS01, CS04, and CS23 codon-altered backgrounds, respectively. As seen in Table 7, surprisingly high expression levels were obtained, as compared to the wild-type construct vCS40 (level defined as 1). vCS26, for instance, expressed 202-fold higher levels compared to the wild-type vCS40 vector. Another control construct for the X1-series of vectors, vCH111, that contains the X1 mutation in the Geneart codon context, showed a more modest increase in expression (12-fold).

[0392] Vectors vCS16, vCS28, and vCS34 all contain the F328S (SPI, F309S SPE) mutation enhancing secretion in the CS01, CS04, and CS23 codon-altered backgrounds, respectively. As seen in Table 7, high expression levels (45-93-fold higher than the wt vCS40 control) were obtained with vCS16 and vCS28.

[0393] Vectors vCS20, vCS24, and vCS33 contain the X5 mutation in the CS01, CS04, and CS23 codon-altered backgrounds, respectively. The best performing variant in the X5 series was vCS20, achieving levels of >3 units/ml after day 14 and a 121-fold increase over the wt vCS40 control.

[0394] Vectors vCS17, vCS29, and vCS31 contain the combination of the X1 and F328S (SPI, F309S SPE) mutations in the CS01, CS04, and CS23 codon-altered backgrounds, respectively (Table 6). The vCS17 and vCS29 constructs achieved very high expression levels in the mouse studies (115 to 246-fold increase over the vCS40 control). Remarkably, in the FVIII KO mouse model used, the majority of mice treated with the vCS17 construct did not develop neutralizing antibodies over time, evidenced by increasing levels of FVIII at later time points (e.g., day 28 and day 42; data not shown). This is an unexpected finding, because in some other constructs the expression levels began to decrease with time due to the formation of neutralizing antibodies. The CS01 background combined with the secretion-enhancing mutations F328S (SPI, F309S SPE) and X1 resulted in low immunogenicity induction.

[0395] Vectors vCS18, vCS27, and vCS35 contain the combination of the X1 and X5 mutations in the CS01, CS04, and CS23 codon-altered backgrounds, respectively. The combination of these two mutations was also very efficient. A 145-fold increase over the vCS40 control could be achieved with vCS18, for example (Table 7).

[0396] Vectors vCS48 and vCS49 contain the combination of the X5 and F328S (SPI, F309S SPE) mutations in the CS01 and CS04 codon-altered backgrounds, respectively. The combination of these two mutations was also very efficient. One of the largest increases of all mutants, a 239-fold increase over the vCS40 control, could be achieved with vCS49 confirming the special value of the combinations including the F328S (SPI, F309S SPE) mutation.

[0397] A further surprising observation was that the mutant AAV vectors grew substantially better than the vCS40 construct harboring the wild-type BDD-FVIII codons. Sequence optimization resulted in a several-fold yield increase in vector production. In some of the best expressing constructs (e.g., vCS29, vCS17, vCS20, and vCS26) the increase in yields due to codon-alteration and/or mutant sequence was approximately 3-5-fold higher, as compared to the wild-type vector (Table 8).

[0398] Expression of BDD-FVIII in the plasma of FVIII-knock-out mice induced by the different AAV vector constructs is shown in Table 7. The constructs have the same vector backbone, however, encode different types of mutated FVIII, including different codon optimization backgrounds. Expression levels at day 14 were chosen because at this time point there is minimal influence of inhibitory antibodies usually seen in some animals in the mouse model at later times. N.d., not determined.

TABLE 7

In vivo biopotency data of vCS constructs.						
#	Vector	Algorithm, mutations	Day 14 ex- pression		Number of mice [n]	Fold ex- pression
			[IU/ml]	SD		
1	vCS19	CS01, X1	2.34	1.10	13	78
2	vCS26	CS04, X1	6.07	2.72	12	202
3	vCS32	CS23, X1	n.d.	—	—	—
4	vCS16	CS01, F328S	1.35	0.88	6	45
5	vCS28	CS04, F328S	2.78	0.92	7	93
6	vCS34	CS23, F328S	n.d.	—	—	—
7	vCS20	CS01, X5	3.62	1.96	21	121
8	vCS24	CS04, X5	0.79	0.89	18	26
9	vCS33	CS23, X5	n.d.	—	—	n.d.
10	vCS17	CS01, X1, F328S	3.44	1.92	20	115
11	vCS29	CS04, X1, F328S	7.39	2.64	9	246
12	vCS31	CS23, X1, F328S	n.d.	—	—	n.d.
13	vCS18	CS01, X1 + X5 (X6)	4.34	2.50	6	145
14	vCS27	CS04, X1 + X5 (X6)	8.03	3.97-	6-	268-
15	vCS35	CS23, X1 + X5 (X6)	n.d.	—	—	—

TABLE 7-continued

In vivo biopotency data of vCS constructs.						
#	Vector	Algorithm, mutations	Day 14 ex- pression		Number of mice [n]	Fold ex- pression
			[IU/ml]	SD		
19	vCS48	CS01, X5, F328S	2.54	0.72	8	85
20	vCS49	CS04, X5, F328S	7.17	1.30	7	239
<u>controls</u>						
16	vCS40	Human wild-type	0.03	0.03	12	1
17	vCh25	Geneart	0.13	0.12	18	4
18	vCh111	Geneart + X1	0.37	0.21	17	12

TABLE 8

Yields per liter cell culture (packaging efficiency) obtained with the different AAV vector constructs. The vectors were purified out of the cell pellets; n.d., not determined.					
construct	Algorithm, mutations	Vector conc. [vg/ml] × 10 ¹²	Yields [vg /liter] × 10 ¹²	Fold increase vs wt	
1	vCS19	CS01, X1	9.71	36	3.22
2	vCS26	CS04, X1	5.93	32	2.87
3	vCS32	CS23, X1	n.d.	n.d.	n.d.
4	vCS16	CS01, F328S	6.51	29	2.56
5	vCS28	CS04, F328S	5.85	32	2.88
6	vCS34	CS23, F328S	n.d.	n.d.	n.d.
7	vCS20	CS01, X5	9.90	50	4.48
8	vCS24	CS04, X5	3.00	16	1.46
9	vCS33	CS23, X5	n.d.	n.d.	n.d.
10	vCS17	CS01, X1, F328S	8.94	37	3.34
11	vCS29	CS04, X1, F328S	7.42	53	4.72
12	vCS31	CS23, X1, F328S	n.d.	n.d.	n.d.
13	vCS18	CS01, X1 + X5 (X6)	21.20	53	4.75
14	vCS27	CS04, X1 + X5 (X6)	4.15	19	1.67
15	vCS35	CS23, X1 + X5 (X6)	n.d.	n.d.	n.d.
16	vCS48	CS01, X5, F328S	7.14	42.1	3.77
17	vCS49	CS04, X5, F328S	8.27	37.2	3.33
18	vCS40	Human wild-type	2.03	11	1.00

[0399] It is 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 scope of the appended claims. All publications, patents, and patent applications cited herein are hereby incorporated by reference in their entirety for all purposes.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 109

```

<210> SEQ ID NO 1
<211> LENGTH: 4374
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence

```

-continued

```

<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polynucleotide

<400> SEQUENCE: 1
atgcagattg agctgagcac ctgttcttc ctgtgcctgc tgaggttctg cttctctgcc      60
accaggagat actacacctggg ggctgtggag ctttcttggg actacatgca gtctgacactg    120
ggggagctgc ctgtggatgc caggttccca cccagagtgc ccaaattcctt cccattcaac     180
acctctgtgg tctacaagaa gaccctctt gtggagttca ctgaccacct gttcaacatt     240
gceaaaccca ggcacccctg gatgggactc ctgggaccca ccattcaggc tgaggtgtat    300
gacaactgtgg tcatcacccct caagaacatcg gcctcccacc ctgtgagcct gcatgctgtg    360
ggggtcagct actggaaggc ctctgagggg gctgagttatg atgaccagac ctcccagagg    420
gagaaggagg atgacaatagt gttccctggg ggcagccaca cctatgtgtg gcaggtcctc    480
aaggagaatg gccccatggc ctctgaccca ctctgcctga cctactccctt ccttctcat    540
gtggacactgg tcaaggaccc tcaactcttgc ctgttggggg cctgtgttgc gtgcaggag     600
ggctccctgg ccaaagagaa gacccagacc ctgcacaatgt tcaattctctt gtttgcgtgc    660
tttgatgagg gcaagagctg gcactctgaa accaagaact cctgtatgc ggacaggat     720
gtgcctctg ccagggctg gcccaagatg cacactgtga atggctatgt gaacaggagc    780
ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcattgtatg tggcatgggg    840
acaacccctg aggtgcactc cattttctg gagggccaca ctttctgttgc caggaaccac    900
agacaggcca gcctggagat cagccccatc accttcttgc ctgcccagac cttgtgtatg    960
gacctegiac agttctgttgc gtttgcac accatgtatgc tggcatgggg 1020
gcctatgtca aggtggacag ctgcctgttgc gagccacagc tcaggatgaa gaacaatgtg    1080
gaggctgagg actatgtatgc tgacccatgttgc gactctgttgc tggatgttgttgc 1140
gatgacaaca gccccatctt catttgcactt aggttctgttgc ccaagaaaca ccccaagacc    1200
tgggtgcactt acattgttgc tgaggaggat gactggactt atgccccactt ggttgcgttgc    1260
ctgtatgttgc gtagtgcactt ggttgcgttgc ccaagaaaca ccccaagacc 1320
aagtacaaga aagtcaaggat catggccatc actgtatgttgc ctttcaagac cagggaggcc    1380
attcagcatg agtctggcat cctggggccatc ctctgtatgc gggagggtgg ggacaccctg    1440
ctcatcatct tcaagaacca ggcctccagg ccctacaaca tctaccatca tggcatcatc    1500
gatgtcaggc ccctgtatgc ccgcaggatc ccaaaagggggg tgaaacaccc caaggacttc    1560
ccatttgc tggggaggat cttcaagatc aagtggactt tcactgttgc ggttgcgttgc 1620
accaaatctg accccaggatc ccttccatc tactacttca gtttgcgttgc catggaggatg    1680
gacctggccatc ctggccatgttgc tggccatgttgc ctttgcgttgc acaaggatc tggatgttgc    1740
agggaaacc agatcatgttgc tgacaaggatc aatgtgttgc tggatgttgc ctttgcgttgc    1800
aacaggagctt ggttgcgttgc tgagaacatc ctttgcgttgc tggatgttgc 1860
cagctggagg accctgtatgc ccaggccaggc aacatcatgc actccatcaaa tggatgttgc    1920
tttgacatgc tccatgttgc tggatgttgc ctttgcgttgc ctttgcgttgc 1980
attggggccatc agactgtatgc ctttgcgttgc ttcttgcgttgc gtttgcgttgc 2040
atggatgtatgc aggacaccctt gacccatgttgc ccatttgcgttgc gggatgttgc 2100

```

-continued

atggagaacc ctggcctgtg gattctggga tgccacaact ctgacttccg caacaggggc	2160
atgactgccg tgctcaaagt ctccctctgt gacaagaaca ctggggacta ctatgaggac	2220
agctatgagg acatctctgc ctacctgtc agcaagaaca atgcccattga gcccaggagc	2280
ttcagccaga atccacctgt cctgaaacgc caccagaggg agatcaccag gaccaccctc	2340
cagtctgacc aggaggagat tgactatgtat gacaccattt ctgtggagat gaagaaagag	2400
gactttgaca tctatgacga ggacgagaac cagagccaa ggagcttcca gaagaagacc	2460
aggcactact tcattgctgc tggggactc ctgtgggact atggcatgag ctccagcccc	2520
catgtccctca ggaacagggc ccagtctggc tctgtgccac agttcaagaa agtgtcttc	2580
caagagttca ctgatggcag cttcacccag cccctgtaca gagggggagct gaatgaggac	2640
ctgggactcc tggggccata catcagggtt gagggtggagg acaacatcat ggtgaccctc	2700
cgcacccagg cctccaggcc ctacagcttc tacagctccc tcatcagcta tgaggaggac	2760
cagaggcagg gggctgagcc acgcaagaac tttgtgaaac ccaatgaaac caagacctac	2820
ttctggaaag tccagcacca catggccccc accaaggatg agtttgactg caaggcctgg	2880
gcctacttct ctgatgtgga cctggagaag gatgtgactt ctggcctgtat tggccactc	2940
ctgggtctgcc acaccaacac cctgaaccct gcccatggaa ggcaagtgac tggcaggag	3000
tttgcctct ttttaccat ctttgatgaa accaagagct ggtacttccat tgagaacatg	3060
gagcgcact gcagggccccc atgcaacatt cagatggagg acccccaccc caaagagaac	3120
taccgttcc atgccatcaa tggcacatc atggacaccc tgcctgggt tgcacatggcc	3180
caggaccaga ggatcagggtg gtacctgtt tctatgggtt ccaatgagaa cattcactcc	3240
atccacttct ctgggcatgt cttaactgtg cgcaagaagg aggagtacaa gatggccctg	3300
tacaacccctt accctgggtt ctttgagact gtggagatgc tgcctccaa agctggcattc	3360
tggagggtgg agtgcctcat tggggagcac ctgcatgtt gcatgagcac cctgttctg	3420
gtctacagca acaagtggca gaccccccgtt ggaatggccct ctggccacat cagggacttc	3480
cagatcactg cctctggcca gtatggccag tggggccca agctggccag gctccactac	3540
tctggatcca tcaatgcctg gagcaccaag gagccattca gctggatcaa agtggactg	3600
ctggcccccata tgatcatcca tggcatcaag acccagggggg ccaggcagaa gttctccagc	3660
ctgtacatca gccagttcat catcatgtac agcctggatg gcaagaaatg gcagacccatc	3720
agaggcaact ccactggAAC actcatggtc ttctttggca atgtggacag ctctggcattc	3780
aagcacaaca tcttcaaccc cccaaatcatc gccagataca tcaggctgca ccccacccac	3840
tacagcatcc gcagcacccct caggatggag ctgatgggtt gtgacccatc tccctgcagc	3900
atgccccctgg gcatggagag caaggccatt tctgtatgccc agatcactgc ctccagctac	3960
ttcaccaaca tggggccac ctggagccca agcaaggcca ggctgcaccc ccaggaaagg	4020
agcaatgcct ggaggccca ggtcaacaac ccaaaggagt ggctgcaggt ggacttccag	4080
aagaccatga aggtcaactgg ggtgaccacc cagggggtca agagcctgtt caccagcatg	4140
tatgtgaagg agttctgtat cagctccagc caggatggcc accagtggac cctcttcttc	4200
cagaatggca aggtcaaggt gttccaggcc aaccaggaca gtttcacccccc tgggtgaac	4260
agcctggacc cccccccttcc gaccagatac ctgaggattc acccccagag ctgggtccac	4320
cagattgccc tgaggatgga ggttctggga tggaggcccc aggacctgtt ctga	4374

-continued

<210> SEQ ID NO 2
<211> LENGTH: 1457
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polypeptide

<400> SEQUENCE: 2

Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe
1 5 10 15

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

Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg
35 40 45

Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val
50 55 60

Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile
65 70 75 80

Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln
85 90 95

Ala Glu Val Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser
100 105 110

His Pro Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser
115 120 125

Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp
130 135 140

Asp Lys Val Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu
145 150 155 160

Lys Glu Asn Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser
165 170 175

Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile
180 185 190

Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr
195 200 205

Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly
210 215 220

Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp
225 230 235 240

Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr
245 250 255

Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val
260 265 270

Tyr Trp His Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile
275 280 285

Phe Leu Glu Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser
290 295 300

Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met
305 310 315 320

Asp Leu Gly Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His
325 330 335

Asp Gly Met Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro
340 345 350

-continued

Gln Leu Arg Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp
355 360 365

Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser
370 375 380

Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr
385 390 395 400

Trp Val His Tyr Ile Ala Ala Glu Glu Asp Trp Asp Tyr Ala Pro
405 410 415

Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn
420 425 430

Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met
435 440 445

Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu
450 455 460

Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu
465 470 475 480

Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro
485 490 495

His Gly Ile Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys
500 505 510

Gly Val Lys His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe
515 520 525

Lys Tyr Lys Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp
530 535 540

Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg
545 550 555 560

Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu
565 570 575

Ser Val Asp Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val
580 585 590

Ile Leu Phe Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu
595 600 605

Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp
610 615 620

Pro Glu Phe Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val
625 630 635 640

Phe Asp Ser Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp
645 650 655

Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe
660 665 670

Ser Gly Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr
675 680 685

Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro
690 695 700

Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly
705 710 715 720

Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp
725 730 735

Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys
740 745 750

-continued

Asn	Asn	Ala	Ile	Glu	Pro	Arg	Ser	Phe	Ser	Gln	Asn	Pro	Pro	Val	Leu
755				760						765					
Lys	Arg	His	Gln	Arg	Glu	Ile	Thr	Arg	Thr	Thr	Leu	Gln	Ser	Asp	Gln
770				775						780					
Glu	Glu	Ile	Asp	Tyr	Asp	Asp	Thr	Ile	Ser	Val	Glu	Met	Lys	Lys	Glu
785				790						795			800		
Asp	Phe	Asp	Ile	Tyr	Asp	Glu	Asp	Glu	Asn	Gln	Ser	Pro	Arg	Ser	Phe
	805					810						815			
Gln	Lys	Lys	Thr	Arg	His	Tyr	Phe	Ile	Ala	Ala	Val	Glu	Arg	Leu	Trp
	820					825						830			
Asp	Tyr	Gly	Met	Ser	Ser	Ser	Pro	His	Val	Leu	Arg	Asn	Arg	Ala	Gln
	835					840						845			
Ser	Gly	Ser	Val	Pro	Gln	Phe	Lys	Lys	Val	Val	Phe	Gln	Glu	Phe	Thr
	850					855						860			
Asp	Gly	Ser	Phe	Thr	Gln	Pro	Leu	Tyr	Arg	Gly	Glu	Leu	Asn	Glu	His
	865					870					875			880	
Leu	Gly	Leu	Leu	Gly	Pro	Tyr	Ile	Arg	Ala	Glu	Val	Glu	Asp	Asn	Ile
	885					890					895				
Met	Val	Thr	Phe	Arg	Asn	Gln	Ala	Ser	Arg	Pro	Tyr	Ser	Phe	Tyr	Ser
	900					905						910			
Ser	Leu	Ile	Ser	Tyr	Glu	Glu	Asp	Gln	Arg	Gln	Gly	Ala	Glu	Pro	Arg
	915					920						925			
Lys	Asn	Phe	Val	Lys	Pro	Asn	Glu	Thr	Lys	Thr	Tyr	Phe	Trp	Lys	Val
	930					935						940			
Gln	His	His	Met	Ala	Pro	Thr	Lys	Asp	Glu	Phe	Asp	Cys	Lys	Ala	Trp
	945					950					955			960	
Ala	Tyr	Phe	Ser	Asp	Val	Asp	Leu	Glu	Lys	Asp	Val	His	Ser	Gly	Leu
	965					970						975			
Ile	Gly	Pro	Leu	Leu	Val	Cys	His	Thr	Asn	Thr	Leu	Asn	Pro	Ala	His
	980					985						990			
Gly	Arg	Gln	Val	Thr	Val	Gln	Glu	Phe	Ala	Leu	Phe	Phe	Thr	Ile	Phe
	995					1000						1005			
Asp	Glu	Thr	Lys	Ser	Trp	Tyr	Phe	Thr	Glu	Asn	Met	Glu	Arg	Asn	
	1010					1015					1020				
Cys	Arg	Ala	Pro	Cys	Asn	Ile	Gln	Met	Glu	Asp	Pro	Thr	Phe	Lys	
	1025					1030					1035				
Glu	Asn	Tyr	Arg	Phe	His	Ala	Ile	Asn	Gly	Tyr	Ile	Met	Asp	Thr	
	1040					1045					1050				
Leu	Pro	Gly	Leu	Val	Met	Ala	Gln	Asp	Gln	Arg	Ile	Arg	Trp	Tyr	
	1055					1060					1065				
Leu	Leu	Ser	Met	Gly	Ser	Asn	Glu	Asn	Ile	His	Ser	Ile	His	Phe	
	1070					1075					1080				
Ser	Gly	His	Val	Phe	Thr	Val	Arg	Lys	Lys	Glu	Glu	Tyr	Lys	Met	
	1085					1090					1095				
Ala	Leu	Tyr	Asn	Leu	Tyr	Pro	Gly	Val	Phe	Glu	Thr	Val	Glu	Met	
	1100					1105					1110				
Leu	Pro	Ser	Lys	Ala	Gly	Ile	Trp	Arg	Val	Glu	Cys	Leu	Ile	Gly	
	1115					1120					1125				
Glu	His	Leu	His	Ala	Gly	Met	Ser	Thr	Leu	Phe	Leu	Val	Tyr	Ser	
	1130					1135					1140				
Asn	Lys	Cys	Gln	Thr	Pro	Leu	Gly	Met	Ala	Ser	Gly	His	Ile	Arg	

-continued

1145	1150	1155
Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr Gly Gln Trp Ala Pro		
1160	1165	1170
Lys Leu Ala Arg Leu His Tyr Ser Gly Ser Ile Asn Ala Trp Ser		
1175	1180	1185
Thr Lys Glu Pro Phe Ser Trp Ile Lys Val Asp Leu Leu Ala Pro		
1190	1195	1200
Met Ile Ile His Gly Ile Lys Thr Gln Gly Ala Arg Gln Lys Phe		
1205	1210	1215
Ser Ser Leu Tyr Ile Ser Gln Phe Ile Ile Met Tyr Ser Leu Asp		
1220	1225	1230
Gly Lys Lys Trp Gln Thr Tyr Arg Gly Asn Ser Thr Gly Thr Leu		
1235	1240	1245
Met Val Phe Phe Gly Asn Val Asp Ser Ser Gly Ile Lys His Asn		
1250	1255	1260
Ile Phe Asn Pro Pro Ile Ile Ala Arg Tyr Ile Arg Leu His Pro		
1265	1270	1275
Thr His Tyr Ser Ile Arg Ser Thr Leu Arg Met Glu Leu Met Gly		
1280	1285	1290
Cys Asp Leu Asn Ser Cys Ser Met Pro Leu Gly Met Glu Ser Lys		
1295	1300	1305
Ala Ile Ser Asp Ala Gln Ile Thr Ala Ser Ser Tyr Phe Thr Asn		
1310	1315	1320
Met Phe Ala Thr Trp Ser Pro Ser Lys Ala Arg Leu His Leu Gln		
1325	1330	1335
Gly Arg Ser Asn Ala Trp Arg Pro Gln Val Asn Asn Pro Lys Glu		
1340	1345	1350
Trp Leu Gln Val Asp Phe Gln Lys Thr Met Lys Val Thr Gly Val		
1355	1360	1365
Thr Thr Gln Gly Val Lys Ser Leu Leu Thr Ser Met Tyr Val Lys		
1370	1375	1380
Glu Phe Leu Ile Ser Ser Gln Asp Gly His Gln Trp Thr Leu		
1385	1390	1395
Phe Phe Gln Asn Gly Lys Val Lys Val Phe Gln Gly Asn Gln Asp		
1400	1405	1410
Ser Phe Thr Pro Val Val Asn Ser Leu Asp Pro Pro Leu Leu Thr		
1415	1420	1425
Arg Tyr Leu Arg Ile His Pro Gln Ser Trp Val His Gln Ile Ala		
1430	1435	1440
Leu Arg Met Glu Val Leu Gly Cys Glu Ala Gln Asp Leu Tyr		
1445	1450	1455

```

<210> SEQ_ID NO 3
<211> LENGTH: 2220
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
polynucleotide

```

<400> SEQUENCE: 3

```

gccaccagga gatactacct gggggctgtg gagcttctt gggactacat gcagtctgac      60
ctgggggagc tgcctgtgga tgccagggttc ccaccccagag tgccccaaatc cttcccatc     120

```

-continued

aacacctctg	tggctacaa	gaagaccctc	tttggaggat	tcactgacca	cctgttcaac	180
attgc当地	ccaggcccacc	ctggatggga	ctcctggac	ccaccattca	ggctgaggtg	240
tatgacactg	tggctatcac	cctcaagaac	atggcctccc	accctgttag	cctgcatact	300
gtgggggtca	gtaactggaa	ggcctcttag	ggggctgagt	atgatgacca	gacctccag	360
agggagaagg	aggatgacaa	agtgttccct	gggggcagcc	acacctatgt	gtggcaggc	420
ctcaaggaga	atggccccat	ggcctctgac	ccactctgac	tgacctactc	ctacccat	480
catgtggacc	tggtaagga	cctcaactct	ggactgattt	ggccctgtct	ggtgtgcagg	540
gaggggctcc	tggccaaaga	gaagacccag	accctgcaca	agttcattct	cctgtttgt	600
gtctttgatg	agggcaagag	ctggcactct	gaaaccaaga	actccctgtat	gcaggacagg	660
gatgctgcct	ctgccaggc	ctggcccaag	atgcacactt	tgaatggcta	tgtgaacagg	720
agccctgcctg	gactcattgg	ctgcccacagg	aaatctgtct	actggcatgt	gattggcatg	780
gggacaaccc	ctgagggtca	ctccattttc	ctggaggggc	acaccttct	ggtcaggaac	840
cacagacagg	ccagccttgg	gatcagcccc	atcaccttcc	tcactgccc	gaccctgt	900
atggacatcg	gacagttcct	gctgttctgc	cacatcagct	cccaccagca	tgtggcatg	960
gaggcctatg	tcaagggtgg	cagctgcct	gaggagccac	agctcaggat	gaagaacaat	1020
gaggaggctg	aggactatga	tgtgacctg	actgactctg	agatggatgt	ggtccgctt	1080
gatgatgaca	acagccatc	cttcatttc	atcaggctgt	tggcaagaa	acaccccaag	1140
acctgggtgc	actacattgc	tgctgaggag	gaggactggg	actatgcccc	actggctct	1200
gcccctgtat	acaggagctt	caagagccag	tacctcaaca	atggccacaca	gaggatttgg	1260
cgcaagtaca	agaaagtca	gttcatggcc	tacactgtat	aaaccttcaa	gaccaggag	1320
gccatttcagc	atgagtctgg	catcctgggc	ccactcctgt	atggggaggt	gggggacacc	1380
ctgctcatca	tcttcaagaa	ccaggcctcc	aggccctaca	acatctaccc	acatggcatc	1440
actgatgtca	ggccctgtta	cagccgcagg	ctgccaagg	gggtgaaaca	cctcaaggac	1500
ttccccatcc	tgcctgggg	gatcttcaag	tacaagtgg	ctgtcactgt	ggaggatgg	1560
ccaaaccaat	ctgacccctag	gtgcctcacc	agataactact	ccagcttct	gaacatggag	1620
agggacactgg	ctctggccct	gattggccca	ctgctcatct	gctacaagga	gtctgtggac	1680
cagaggggaa	accagatcat	gtctgacaag	aggaatgtga	ttctgttctc	tgtctttgtat	1740
gagaacagga	gctggtacct	gactgagaac	attcagcgct	tcctgeccaa	ccctgctgg	1800
gtgcagctgg	aggacccctga	gttccaggcc	agcaacatca	tgcactccat	caatggctat	1860
gtgtttgaca	gcctccagct	ttctgtctgc	ctgcatgagg	tggcctactg	gtacattct	1920
tctattgggg	cccaactgtca	cttcctttct	gttcttct	ctggctacac	cttcaaacac	1980
aagatgggt	atgaggacac	cctgaccctc	ttcccttct	ctggggagac	tgtgttcatg	2040
agcatggaga	accctggccct	gtggattctg	ggatgccaca	actctgactt	ccgcaacagg	2100
ggcatgactg	ccctgctcaa	agtctccctcc	tgtgacaaga	acactgggg	ctactatgag	2160
gacagctatg	aggacatctc	tgcctacctg	ctcagcaaga	acaatgccc	tgagcccagg	2220

<210> SEQ ID NO 4

<211> LENGTH: 2052

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

-continued

```

<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polynucleotide

<400> SEQUENCE: 4
      gagatcacca ggaccaccct ccagtctgac caggaggaga ttgactatga tgacaccatt    60
      tctgtggaga tgaagaaaga ggactttgac atctatgacg aggacgagaa ccagagccca    120
      agagaccttc agaagaagac caggcactac ttcatgtcg ctgtggagcg cctgtggac     180
      tatggcatga gctccagccc ccatgtcctc aggaacaggg cccagtctgg ctctgtgccca   240
      cagttcaaga aagtggtctt ccaagagttc actgatggca gtttcaccca gcccctgtac   300
      agaggggagc tgaatgagca cctgggactc ctggggccat acatcaggc tgagggtggag   360
      gacaacatca tggtgacctt cgcgaaccag gcctccaggc cctacagctt ctacagctcc   420
      ctcatcagct atgaggagga ccagaggcag ggggctgagc cacgcaagaa ctttgtgaaa   480
      cccaatgaaa ccaagaccta cttctggaaa gtccagcacc acatggcccc caccaaggat   540
      gagtttact gcaaggctg ggcctacttc tctgtatgtgg acctggagaa ggatgtgcac   600
      tctggcctga ttggccact cctgtgtcg cacaccaaca cctgtaccc tgcccatgg     660
      aggcaagtga ctgtgcagga gtttgcctc ttcttcacca tctttatgtga aaccaagagc   720
      tggtaacttca ctgagaacat ggagcgcaac tgcagggccc catgcaacat tcagatggag   780
      gacccacact tcaaagagaa ctaccgcctc catgcatca atggctacat catggacacc   840
      ctgcctggc ttgtcatgac ccaggaccag aggtcagggt ggtacactgt ttctatggc   900
      tccaatgaga acattcactc catccacttc tctgggcatg tcttcactgt ggcgaagaag   960
      gaggagtaca agatggccct gtacaaccc taccctgggg tctttgagac tgtggagatg   1020
      ctgcctccca aagctggcat ctgggggtg gagtgcctca ttggggagca cctgcacgt   1080
      ggcatgagca ccctgttccct ggtctacgc aacaagtgcc agacccccc gggaatggcc   1140
      tctggccaca tcagggactt ccagatcact gcctctggc agtatggcca gtggggcccc   1200
      aagctggcca ggctccacta ctctggatcc atcaatgcct ggagcaccaa ggagccattc   1260
      agctggatca aagtggaccc gctggccccc atgatcatcc atggcatcaa gacccagggg   1320
      gccaggcaga agttctccag cctgtacatc agccagttca tcatcatgta cagcctggat   1380
      ggcaagaaat ggcagaccta cagaggcaac tccactggaa cactcatggt cttcttggc   1440
      aatgtggaca gctctggcat caagcacaac atcttcaacc ccccaatcat cgccagatac   1500
      atcaggctgc acccccacca ctacagcatc cgcagcaccc tcaggatgga gctgtatggc   1560
      tttgtacatca actctctgcag catgcccctg ggcacatggaga gcaaggccat ttctgtatggc   1620
      cagatcactg cctccagacta cttcaccaac atgtttgcca cctggagccc aagcaaggcc   1680
      aggctgcacc tccagggaag gagcaatgcc tggaggcccc aggtcaacaa cccaaaggag   1740
      tggctgcagg tggacttcca gaagaccatg aaggtcactg gggtgaccac ccagggggc   1800
      aagagctgc tcaccagcat gtatgtgaag gagttcctga tcaagctccag ccaggatggc   1860
      caccagtggaa ccctttctt ccagaatggc aaggtcaagg tggccaggc caaccaggac   1920
      agcttcaccc ctgtgggtgaa cagcctggac ccccccctcc tgaccagata cctgaggatt   1980
      caccggcaga gctgggtcca ccagattgcc ctgaggatgg aggtcctggg atgtgaggcc   2040
      caggacactgt ac                                         2052

```

-continued

```
<210> SEQ ID NO 5
<211> LENGTH: 42
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 5
agtttctc tca agaatccacc tgcgttgtaa agacaccaga ga 42

<210> SEQ ID NO 6
<211> LENGTH: 42
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 6
agtttcagcc tca agaatccacc tgcgttgtaaa cgccaccaga gg 42

<210> SEQ ID NO 7
<211> LENGTH: 42
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide

<400> SEQUENCE: 7
agtttcagcc tca agaacccccc cgtgttgtaa aggccaccaga gg 42

<210> SEQ ID NO 8
<211> LENGTH: 7827
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
polynucleotide

<400> SEQUENCE: 8
tcgcgcgttt cgggtatgac ggtgaaaacc tctgacacat gcagctcccg gagacggta 60
cagcttgtct gtaagcgat gcccggggca gacaaggcccg tcagggcgcg tcagcggtg 120
ttggcggttg tcggggctgg cttaactatg cggcatcaga gcagattgtt ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgttaaggag aaaataccgc atcaggcgcc 240
attcgcatt caggctgcgc aactgttggg aaggcgatc ggtgcggggcc tcttcgttat 300
tacgcccagct ggcgaaaggg ggtatgtctg caaggcgatt aagttgggtt acgccagggt 360
tttcccagtc acgacgttgtt aaaacgacgg ccagtgaatt cctcgagatt taaatgacgt 420
tggccactcc ctctctgcgc gctcgctcg tcactgaggc cggcgacca aaggctgccc 480
gacgccccggg ctttgcgggg cggcgctcg tgagcgagcg agcgcgacaga gagggaggtgg 540
ccaactccat cactagggtt tcctgagtt aaacttcgtc gacgattcga gcttggctg 600
caggtcgagg gcactgggg gatgttgagt aagatggaaa actactgtt acccttgcag 660
agacagagta ttaggacatg tttgaacagg ggccggggca tcagcgatgt gctctagagg 720
atccccgtct gtctgcacat ttctgttaggc gagtgttccg atactctaatt ctccttaggc 780
```

-continued

aagggttcata	tttgtttagg	ttacttattc	tcctttgtt	gactaatgtca	ataatcatgaa	840			
tca	cgaggtt	tgagtcage	ttggcaggga	tca	gcagcc	gggttggaaag	gagggggtat	900	
aaa	aggccc	tcaccaggag	aagccgtac	acagactagg	cgcgcac	ccaccatgca	960		
gat	tggactg	agcacctgct	tcttcctgtg	cctgctgagg	ttctgttct	ctgccaccag	1020		
gag	atactac	ctgggggctg	tggagcttc	ttgggactac	atgcagtctg	acctggggga	1080		
gt	gcctgtg	gtgc	ccagg	tcc	ccccag	atgtcccaa	tc	aaacac	1140
tgtgg	tctac	aagaagaccc	tctt	gttga	gttca	tc	acattgcca	1200	
acc	cagg	ccctggatgg	gactcctggg	acc	caccatt	caggctgagg	tgtatgac	1260	
tgtgg	tatc	acc	ctcaaga	acatggc	ccac	ctgtg	agc	ctg	1320
cag	ctactgg	aagg	cct	ggctg	atgt	gttca	agac	ctcc	1380
ggagg	atgttcc	ctgggg	ccac	ac	ctt	gttgg	agg	tc	1440
gaat	ggcccc	atggc	ctct	actctg	cctgac	cttac	ctt	ctcatgtg	1500
cctgg	tcaag	gac	ctca	act	ctgg	actgat	ttc	gggg	1560
cctgg	ccaaa	gaga	agac	cc	agtt	ctt	gtt	gtt	1620
tgagg	ggcaag	agc	tgg	act	ctgaa	act	cc	atgc	1680
ctct	gecagg	gc	ctgg	ccca	agat	gcac	ac	tgt	1740
tggact	catt	gg	ctgg	ccaca	gg	atgtgg	ca	ttt	1800
ccct	gaggt	cac	tgg	ttt	cc	atgtt	gg	gtt	1860
ggcc	agcctg	gag	atgc	ccat	cct	cttgc	ttgc	atgtt	1920
cg	acatgtt	ctg	ctgtt	ccat	ccc	acc	atgtt	gg	1980
tgtca	agg	tgac	actg	cc	ctt	ggat	gg	atgtt	2040
tgagg	actat	tgat	tgact	tgat	tt	gttgg	ttt	gtat	2100
caac	agccca	tc	ttt	atgtt	ttt	gttgg	ttt	gtat	2160
gc	actacatt	gt	gtt	ttt	atgtt	ttt	gttgg	ttt	2220
tgac	agg	aca	ag	atgtt	ttt	gttgg	ttt	gtat	2280
caagaa	agtc	ttt	atgtt	ttt	atgtt	ttt	gttgg	ttt	2340
gc	atgtt	ttt	atgtt	ttt	atgtt	ttt	gttgg	ttt	2400
catctt	caag	ttt	atgtt	ttt	atgtt	ttt	gttgg	ttt	2460
cagg	ccctg	ta	ccat	ttt	atgtt	ttt	gttgg	ttt	2520
tctg	ctggg	ga	at	ttt	atgtt	ttt	gttgg	ttt	2580
atct	gttcccc	ag	gtt	cc	atgtt	ttt	gttgg	ttt	2640
ggc	cttggc	ct	gtt	cc	atgtt	ttt	gttgg	ttt	2700
aaacc	aggatc	at	gtt	cc	atgtt	ttt	gttgg	ttt	2760
gag	ctgttac	ct	gtt	cc	atgtt	ttt	gttgg	ttt	2820
ggagg	ccct	ca	gtt	cc	atgtt	ttt	gttgg	ttt	2880
cag	cttctgtt	cc	gtt	cc	atgtt	ttt	gttgg	ttt	2940
ggccc	ccat	ct	gtt	cc	atgtt	ttt	gttgg	ttt	3000
gtat	gagg	ac	ctt	cc	atgtt	ttt	gttgg	ttt	3060

-continued

gaaccctggc ctgtggattc tggatgcca caactctgac ttccgcaca gggcatgac	3120
tgcctgctc aaagtctcct cctgtgacaa gaacactggg gactactatg aggacagcta	3180
tgaggacatc tctgcctacc tgctcagcaa gaacaatgcc attgagccca ggagcttcag	3240
ccagaatcca cctgtcctga aacgccacca gagggagatc accaggacca ccctccagtc	3300
tgaccaggag gagattgact atgatgacac catttctgtg gagatgaaga aagaggactt	3360
tgacatctat gacgaggacg agaaccagag cccaaggagc ttccagaaga agaccaggca	3420
ctacttcatt gctgctgtgg agcgcctgtg ggactatggc atgagctca gccccatgt	3480
cetcaggaac agggeccagt ctggctctgt gccacagttc aagaaagtgg tcttccaaga	3540
gttcactgat ggcagctca cccagcccc gtacagggg gagctgaatg agcacctggg	3600
actcctgggc ccatatcatca gggctgaggt ggaggacaac atcatggta cttcccgaa	3660
ccaggccctcc agggccctaca gcttctacag ctccctcatac agctatgagg aggaccagag	3720
gcagggggct gagccacgca agaactttgt gaaacccaat gaaaccaaga cctacttctg	3780
gaaaggccatc caccacatgg ccccccaccaa ggatgagttt gactgcaagg cctgggccta	3840
cttctctgat gtggacctgg agaaggatgt gcactctggc ctgattggcc cactctgg	3900
ctgccacacc aacaccctga accctgccc tggaggcata gtgactgtgc aggagttgc	3960
cctcttcttc accatctttg atgaaaccaa gagctggta ttcactgaga acatggagcg	4020
caactgcagg ccccatgca acattcagat ggaggacccc accttcaaag agaactaccg	4080
cttccatgcc atcaatggct acatcatgga caccctgcct gggcttgc tggcccgagga	4140
ccagaggatc aggtggtaacc tgcttctat gggctccaat gagaacatc actccatcca	4200
cttctctggg catgtttca ctgtgcgcaaa gaaggaggag tacaagatgg ccctgtacaa	4260
cctctaccct ggggtctttg agactgtgga gatgctgccc tccaaagctg gcatctggag	4320
ggtggagtgc ctcatgggg agcacctgca tgctggcatg agcacccctgt tcctggctca	4380
cagcaacaag tgccagaccc ccctggaaat ggctctggc cacatcaggg actccatcatc	4440
cactgcctct ggcaggatgt gccaggggc ccccaagctg gccaggctcc actactctgg	4500
atccatcaat gcctggagca ccaaggagcc attcagetgg atcaaagtgg acctgctggc	4560
ccccatgatc atccatggca tcaagacca gggggccagg cagaagttct ccagectgt	4620
catcagccat ttcatcatca tgcacggctt ggatggcaag aaatggcaga cctacagagg	4680
caactccact ggaacactca tggcttttt tggcaatgtg gacagctctg gcatcaagca	4740
caacatcttc aaccccccacca tcatgcggc atacatcagg ctgcacccca cccactacag	4800
catccgcagc accctcagga tggagctgat gggctgtac ctgactctc gcagcatgcc	4860
cctgggcatg gagagcaagg ccatttctga tgcccgatc actgcctcca gctacttcac	4920
caacatgttt gccacctggc gccaagccaa ggccaggctg cacctccagg gaaggagcaa	4980
tgcctggagg ccccaaggta acaacccaaa ggagtggctg caggtggact tccagaagac	5040
catgaagggtc actgggggtga ccacccagg ggtcaagagc ctgctcacca gcatgtatgt	5100
gaaggaggatc ctgatcagct ccagccagga tggccaccag tggaccctct tcttccagaa	5160
tggcaagggtc aagggttcc agggcaacca ggacagcttc accccctgtgg tgaacagcct	5220
ggaccccccctt ctcctgacca gatacctgag gattcacccca cagagctggg tccaccagat	5280
tgcctgagg atggaggtcc tggatgtga ggcccaggac ctgtactgat gacgagcggc	5340

-continued

cgctttagt agcagtatcg ataataaaaag atctttattt tcattagatc tgtgtgttgg	5400
tttttgtgt gttaattaag ctgcgaaagg aacccctagt gatggagttg gccactccct	5460
ctctgcccgc tcgctcgctc actgaggccg ggcgaccaa ggtcgcccga cgccgggct	5520
ttgccccggc ggcctcagtg agcgagcgag cgcgacaga gggagtggcc aagacgattt	5580
aatgacaag cttggcgtaa tcatggtcat agctgttcc tgtgtgaaat tgttatccgc	5640
tcacaattcc acacaacata cgagccgaa gcataaagtg taaagctgg ggtgcctaatt	5700
gagtgagcta actcacatta attgcgttgc gtcactgccc cgcttccag tcgggaaacc	5760
tgtcgtgcca gctgcattaa tgaatcgcc aacgcgcccc gagaggcgtt ttgcgtattt	5820
ggcgcttcg ctgccttcg ctcactgact cgctgcgtc ggtcgttccg ctgcggcgag	5880
cggtatcagc tcactcaaag gcggtaaatac ggttatccac agaatcaggg gataacgcag	5940
gaaagaacat gtgagaaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc	6000
tggcgaaaaatccatggctc cgccccctcg acgagcatca caaaaatcga cgctcaagtc	6060
agaggtggcg aaacccgaca ggactataaa gataccaggc gtttccctt ggaagctccc	6120
tatgtgcgtc tccgttccg accctgcgc ttaccggata cctgtccgcc tttccctt	6180
cgggaaacgt ggccgtttct catagtcac gctgttaggtatccatggcgttgc	6240
ttcgctccaa gctggctgt gtgcacgaac ccccggtca gcccggccgc tgcccttat	6300
ccggtaacta tcgttcttgc tccaaacccgg taagacacga cttatgcgc ctggcagcag	6360
ccactggtaa caggatttgc agagcgaggt atgttaggcgg tgctacagag ttcttgaagt	6420
gggtggctaa ctacggctac actagaagaa cagtattttg tatctgcgtct tgctgaagc	6480
cagttaccc ctggaaaaaga gttggtagctt cttgatccgg caaacaaacc accgtggta	6540
gggggtggtttttttttgc aacgcgcaga ttacgcgcag aaaaaaaaaaaggtctcaagaag	6600
atcccttgat cttttctacg gggctgcacg ctcagttggaa cgaaaaactca cgtaaggaa	6660
ttttggtcat gagattatca aaaaggatct tcacctagat cttttaaat taaaatgaa	6720
gttttaatc aatctaaagt atatatgagt aaacttggtc tgacagttac caatgtttaa	6780
tcagtgggc acctatctca gcgatctgc tatttcgttc atccatagtt gcctgactcc	6840
ccgtcggtta gataactacg atacggggagg gcttaccatc tggcccccagt gctgcata	6900
taccggcaga cccacgctca ccggctccag atttacgc aataaaccag ccagccggaa	6960
ggggccgagcg cagaagtggt cctgcaactt tatccgcctc catccagtct attaattttt	7020
gccggaaagc tagagtaagt agttcgccag ttaatagttt ggcacacgtt gttggcattt	7080
ctacaggcat cgtgggtgtca cgctcgctgt ttggatggc ttcatcgac tccgggtccc	7140
aacgcatacg gcgaggatca tgcattccca tggatgtcaaaa aaaaagggtt agtcctcg	7200
gtccctccat cttttgtcaga agtaagttgg ccgcagtttgc atcactcatg gttatggcag	7260
cactgcataa ttcttctact gtcattccat ccgtaaatgtt cttttctgtg actggtgagt	7320
actcaaccaa gtcattctga gaatagtgtt tgcggcgacc gagttgtct tgccggcg	7380
caataccggtaa taataccgcg ccacatagca gaactttaaa agtgcgtatc attggaaaac	7440
gttcttcggg gcgaaaaactc tcaaggatct taccgtgtt gagatccagt tcgatgtaac	7500
ccactcggtgc acccaactga tcttcagcat cttttactttt caccagcggtt tctgggtgag	7560
caaaaacagg aaggcaaaaat gcccggaaaaa agggataaag ggcgacacgg aaatgttga	7620

-continued

tactcatact	tttcctttt	caatattattt	gaagcattta	tcagggttat	tgtctcatga	7680
gccccatacat	atttgaatgt	attttagaaaa	ataaacaaat	aggggttccg	cgcacattc	7740
cccgaaaagt	gccacctgac	gtctaagaaa	ccattattat	catgacattha	acctataaaaa	7800
ataggcgtat	cacgaggccc	tttcgtc				7827

```

<210> SEQ ID NO 9
<211> LENGTH: 4332
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
polynucleotide

```

```

<400> SEQUENCE: 9

atgcagattg agctgagcac ctgtttcttc ctgtgcctgc tgagggtctg cttctctgcc 60
accaggagat actacatggg ggctgtggag cttcttggg actacatgca gtctgactg 120
ggggagctgc ctgtggatgc caggttccca cccagagtgc ccaaattcctt cccattcaac 180
acctctgtgg tctacaagaa gaccctctt gtggagttca ctgaccacat gttcaacatt 240
gccaacccca ggccaccctg gatgggactc ctgggaccca ccattcaggc tgaggttat 300
gacactgtgg tcataccctt caagaacatg gcctcccacc ctgtgagcct gcatgttg 360
ggggtcagct actggaaaggc ctctgagggg gctgagttat atgaccagac ctcccaagg 420
gagaaggagg atgacaaagt gttccctggg ggcagccaca cctatgttg gcaggctc 480
aaggagaatg gccccatggc ctctgaccca ctctgcctga cctactccta cctttctcat 540
gtggacctgg tcaaggaccc caactcttga ctgtttgggg ccctgttgt gtgcaggag 600
ggttccctgg ccaaagagaa gacccagacc ctgcacaagt tcattctctt gtttgcgtc 660
tttgcgtgg gcaagagctg gcactctgaa accaagaact ccctgtatgca ggacaggat 720
gtgcctctg ccaggccctg gcccaagatg cacactgtga atggctatgt gaacaggagc 780
ctgcctggac tcattggctg ccacaggaaa tctgttactt ggcatgtat tggcatgggg 840
acaacccctg aggtgcactc cattttcttg gaggccaca ctttcttgtt caggaaccac 900
agacaggcca gcctggagat cagccccatc accttcctca ctgcccagac cctgtgtat 960
gacctcggac agtttctgtt gtttgcac accagcatga tggcatgggg 1020
gcctatgtca aggtggacag ctgccttgag gagccacagc tcaggatgaa gaacaatgag 1080
gaggctgagg actatgtatg tgacactgact gactctgaga tggatgttgtt ccgtttgtat 1140
gatgacaaca gccccatctt cattcagatc aggttctgtgg ccaagaaaca ccccaagacc 1200
tgggtgcact acattgtgc tgaggaggag gactggact atgccccact ggtcctggcc 1260
cctgtatca gtagtggatc cctggccca ctctgtatg gggaggtggg ggacaccctg 1320
aagtacaaga aagtcaaggat catggccatc actgtatgaa ctttcaagac caggaggcc 1380
attcagcatg agtctggatc cctggccca ctctgtatg gggaggtggg ggacaccctg 1440
ctcatcatct tcaagaacca ggcctccagg ccctacaaca tctaccacca tggcatcact 1500
gatgtcaggc ccctgtacag ccgcaggctg ccaaaggggg tgaaacacct caaggactc 1560
ccatttgc ctggggagat cttaagatc aagtggactc tcactgttggaa ggttggatc 1620
accaaattctg accccaggtg cctcaccaga tactactcca gctttgtgaa catggaggg 1680
gacctggccct ctggcctgtat tggccactg ctcatctgtat acaaggagtc tggaccac 1740

```

-continued

aggggaaacc agatcatgtc tgacaagagg aatgtgattc ttttctctgt ctttgatgag	1800
aacaggagct ggtacctgac tgagaacatt cagcgctcc tgcccaaccc tgctgggtg	1860
cagctggagg accctgagtt ccaggccago aacatcatgc actccatcaa tggctatgtg	1920
tttgacagcc tccagcttc tgtctgcctg catgagggtgg cctactggta cattttct	1980
attggggccc agactgactt ctttctgtc ttcttctctg gctacacccaa caaacacaag	2040
atgggtgtatg aggacacccct gacectcttc ccattctctg gggagactgt gttcatgagc	2100
atggagaacc ctggcctgtg gattctggga tgccacaact ctgacttccg caacaggggc	2160
atgactgccc tgctcaaagt ctccctctgt gacaagaaca ctggggacta ctatgaggac	2220
agctatgagg acatctctgc ctacctgctc agcaagaaca atgccattga gcccaggag	2280
atcaccagga ccaccccca gtctgaccag gaggagattg actatgtga caccattct	2340
gtggagatga agaaagagga ctttgacatc tatgacgagg acgagaacca gagcccaagg	2400
agcttccaga agaagaccag gcactacttc attgctgtg tggagcgctt gtggactat	2460
ggcatgagct ccagccccca tgtcctcagg aacagggccc agtctggctc tgtccacag	2520
ttcaagaaag tggcttcca agagttcaact gatggcagct tcacccagcc cctgtacaga	2580
ggggagctga atgagcacct gggactcctg ggccatataca tcagggctga ggtggaggac	2640
aacatcatgg tgaccttccg caaccaggcc tccaggccct acagcttcta cagctccctc	2700
atcagctatg aggaggacca gaggcagggg gctgagccac gcaagaactt tgtgaaaccc	2760
aatgaaacca agacctactt ctggaaagtgc cagcaccaca tggcccccac caaggatgag	2820
tttgactgca aggcctggc ctacttctct gatgtggacc tggagaaggaa tgtgcactct	2880
ggcctgattg gcccactcct ggtctgccac accaacaccc tgaaccctgc ccatggaaagg	2940
caagtgactg tgcaggagtt tgccttcttc ttccatctt ttgatgaaac caagagctgg	3000
tacttcaactg agaacatgga gcgcaactgc agggcccccatttca gcaacattca gatggaggac	3060
cccacccatca aagagaacta ccgcttccat gccatcaatg gtcacatcat ggacaccctg	3120
cctgggttg tcatggccca ggaccagagg atcaggttgtt acctgtttc tatggctcc	3180
aatgagaaca ttcactccat ccacttctct gggcatgtct tcactgtgcg caagaaggag	3240
gagtaacaaga tggccctgtt caacccatc cctgggtctt ttgagactgt ggagatgctg	3300
ccctccaaag ctggcatctg gagggtggag tgcctcatgg gggagcacct gcatgctggc	3360
atgagcaccc tggctctgggtt ctacagcaac aagtgcaga ccccccctggg aatggctct	3420
ggccacatca gggacttcca gatcaactgatc tctggccagt atggccagtg ggcccaag	3480
ctggccaggc tccactactc tggatccatc aatgcctggatc gaccaagggatc gccattcagc	3540
tggatcaaaag tggacactgtt gggcccccatttgc atcatccatg gcatcaagac ccagggggcc	3600
aggcagaagt tctccagccct gtacatcagc cagttcatca tcatgtacag cctggatggc	3660
aagaaaatggc agacctacag aggcaactcc actggaaacac tcatggctt ctttggcaat	3720
gtggacagct ctggcatcaa gcacaacatc ttcaacccccc caatcatcgc cagatacatc	3780
aggctgcacc ccacccacta cagcatccgc agcaccctca ggtggagct gatggctgt	3840
gacactgaact octgeagcat gcccctgggc atggagagca aggccatttc tcatggccag	3900
atcactgcct ccagctactt caccaacatg tttgcccaccc ttgagccaaag caaggccagg	3960
ctgcacccatcc agggaaaggag caatgcctgg agggcccccagg tcaacaaccc aaaggatgg	4020

-continued

ctgcagggtgg	acttccagaa	gaccatgaag	gtcactgggg	tgaccaccca	gggggtcaag	4080
agcctgctca	ccagcatgta	tgtgaaggag	ttcctgatca	gtcccagcca	ggatggcac	4140
cagtggaccc	tcttcctcca	gaatggcaag	gtcaagggtg	tccaggggcaa	ccaggacagc	4200
ttcaccctcg	tggtgaacag	cctggacccc	ccccctctga	ccagataacct	gaggattcac	4260
ccccagagct	gggtccacca	gattgccctg	aggatggagg	tcctggatg	tgaggccag	4320
gacctgtact	ga					4332

<210> SEQ ID NO 10
<211> LENGTH: 1443
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polypeptide

<400> SEQUENCE: 10

Met	Gln	Ile	Glu	Leu	Ser	Thr	Cys	Phe	Phe	Leu	Cys	Leu	Leu	Arg	Phe
1															
															15
Cys	Phe	Ser	Ala	Thr	Arg	Arg	Tyr	Tyr	Leu	Gly	Ala	Val	Glu	Leu	Ser
															30
Trp	Asp	Tyr	Met	Gln	Ser	Asp	Leu	Gly	Glu	Leu	Pro	Val	Asp	Ala	Arg
															45
Phe	Pro	Pro	Arg	Val	Pro	Lys	Ser	Phe	Pro	Phe	Asn	Thr	Ser	Val	Val
															60
Tyr	Lys	Lys	Thr	Leu	Phe	Val	Glu	Phe	Thr	Asp	His	Leu	Phe	Asn	Ile
															80
Ala	Lys	Pro	Arg	Pro	Pro	Trp	Met	Gly	Leu	Leu	Gly	Pro	Thr	Ile	Gln
															95
Ala	Glu	Val	Tyr	Asp	Thr	Val	Val	Ile	Thr	Leu	Lys	Asn	Met	Ala	Ser
															110
His	Pro	Val	Ser	Leu	His	Ala	Val	Gly	Val	Ser	Tyr	Trp	Lys	Ala	Ser
															125
Glu	Gly	Ala	Glu	Tyr	Asp	Asp	Gln	Thr	Ser	Gln	Arg	Glu	Lys	Glu	Asp
															140
Asp	Lys	Val	Phe	Pro	Gly	Gly	Ser	His	Thr	Tyr	Val	Trp	Gln	Val	Leu
															145
															150
															155
															160
Lys	Glu	Asn	Gly	Pro	Met	Ala	Ser	Asp	Pro	Leu	Cys	Leu	Thr	Tyr	Ser
															165
															170
															175
Tyr	Leu	Ser	His	Val	Asp	Leu	Val	Lys	Asp	Leu	Asn	Ser	Gly	Leu	Ile
															180
															185
															190
Gly	Ala	Leu	Leu	Val	Cys	Arg	Glu	Gly	Ser	Leu	Ala	Lys	Glu	Lys	Thr
															195
															200
															205
Gln	Thr	Leu	His	Lys	Phe	Ile	Leu	Leu	Phe	Ala	Val	Phe	Asp	Glu	Gly
															210
															215
															220
Lys	Ser	Trp	His	Ser	Glu	Thr	Lys	Asn	Ser	Leu	Met	Gln	Asp	Arg	Asp
															225
															230
															235
															240
Ala	Ala	Ser	Ala	Arg	Ala	Trp	Pro	Lys	Met	His	Thr	Val	Asn	Gly	Tyr
															245
															250
															255
Val	Asn	Arg	Ser	Leu	Pro	Gly	Leu	Ile	Gly	Cys	His	Arg	Lys	Ser	Val
															260
															265
															270
Tyr	Trp	His	Val	Ile	Gly	Met	Gly	Thr	Thr	Pro	Glu	Val	His	Ser	Ile

-continued

275	280	285
Phe Leu Glu Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser		
290	295	300
Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met		
305	310	315
Asp Leu Gly Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His		
325	330	335
Asp Gly Met Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro		
340	345	350
Gln Leu Arg Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp		
355	360	365
Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser		
370	375	380
Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr		
385	390	395
Trp Val His Tyr Ile Ala Ala Glu Glu Asp Trp Asp Tyr Ala Pro		
405	410	415
Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn		
420	425	430
Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met		
435	440	445
Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu		
450	455	460
Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu		
465	470	475
480		
Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro		
485	490	495
His Gly Ile Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys		
500	505	510
Gly Val Lys His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe		
515	520	525
Lys Tyr Lys Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp		
530	535	540
Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg		
545	550	555
560		
Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu		
565	570	575
Ser Val Asp Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val		
580	585	590
Ile Leu Phe Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu		
595	600	605
Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp		
610	615	620
Pro Glu Phe Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val		
625	630	635
640		
Phe Asp Ser Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp		
645	650	655
Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe		
660	665	670
Ser Gly Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr		
675	680	685

-continued

Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro
 690 695 700
 Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly
 705 710 715 720
 Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp
 725 730 735
 Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys
 740 745 750
 Asn Asn Ala Ile Glu Pro Arg Glu Ile Thr Arg Thr Thr Leu Gln Ser
 755 760 765
 Asp Gln Glu Glu Ile Asp Tyr Asp Asp Thr Ile Ser Val Glu Met Lys
 770 775 780
 Lys Glu Asp Phe Asp Ile Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg
 785 790 795 800
 Ser Phe Gln Lys Lys Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg
 805 810 815
 Leu Trp Asp Tyr Gly Met Ser Ser Pro His Val Leu Arg Asn Arg
 820 825 830
 Ala Gln Ser Gly Ser Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu
 835 840 845
 Phe Thr Asp Gly Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn
 850 855 860
 Glu His Leu Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp
 865 870 875 880
 Asn Ile Met Val Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe
 885 890 895
 Tyr Ser Ser Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu
 900 905 910
 Pro Arg Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe Trp
 915 920 925
 Lys Val Gln His His Met Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys
 930 935 940
 Ala Trp Ala Tyr Phe Ser Asp Val Asp Leu Glu Lys Asp Val His Ser
 945 950 955 960
 Gly Leu Ile Gly Pro Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro
 965 970 975
 Ala His Gly Arg Gln Val Thr Val Gln Glu Phe Ala Leu Phe Phe Thr
 980 985 990
 Ile Phe Asp Glu Thr Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg
 995 1000 1005
 Asn Cys Arg Ala Pro Cys Asn Ile Gln Met Glu Asp Pro Thr Phe
 1010 1015 1020
 Lys Glu Asn Tyr Arg Phe His Ala Ile Asn Gly Tyr Ile Met Asp
 1025 1030 1035
 Thr Leu Pro Gly Leu Val Met Ala Gln Asp Gln Arg Ile Arg Trp
 1040 1045 1050
 Tyr Leu Leu Ser Met Gly Ser Asn Glu Asn Ile His Ser Ile His
 1055 1060 1065
 Phe Ser Gly His Val Phe Thr Val Arg Lys Lys Glu Glu Tyr Lys
 1070 1075 1080

-continued

Met	Ala	Leu	Tyr	Asn	Leu	Tyr	Pro	Gly	Val	Phe	Glu	Thr	Val	Glu
1085					1090						1095			
Met	Leu	Pro	Ser	Lys	Ala	Gly	Ile	Trp	Arg	Val	Glu	Cys	Leu	Ile
1100					1105						1110			
Gly	Glu	His	Leu	His	Ala	Gly	Met	Ser	Thr	Leu	Phe	Leu	Val	Tyr
1115					1120						1125			
Ser	Asn	Lys	Cys	Gln	Thr	Pro	Leu	Gly	Met	Ala	Ser	Gly	His	Ile
1130					1135						1140			
Arg	Asp	Phe	Gln	Ile	Thr	Ala	Ser	Gly	Gln	Tyr	Gly	Gln	Trp	Ala
1145					1150						1155			
Pro	Lys	Leu	Ala	Arg	Leu	His	Tyr	Ser	Gly	Ser	Ile	Asn	Ala	Trp
1160					1165						1170			
Ser	Thr	Lys	Glu	Pro	Phe	Ser	Trp	Ile	Lys	Val	Asp	Leu	Leu	Ala
1175					1180						1185			
Pro	Met	Ile	Ile	His	Gly	Ile	Lys	Thr	Gln	Gly	Ala	Arg	Gln	Lys
1190					1195						1200			
Phe	Ser	Ser	Leu	Tyr	Ile	Ser	Gln	Phe	Ile	Ile	Met	Tyr	Ser	Leu
1205					1210						1215			
Asp	Gly	Lys	Lys	Trp	Gln	Thr	Tyr	Arg	Gly	Asn	Ser	Thr	Gly	Thr
1220					1225						1230			
Leu	Met	Val	Phe	Phe	Gly	Asn	Val	Asp	Ser	Ser	Gly	Ile	Lys	His
1235					1240						1245			
Asn	Ile	Phe	Asn	Pro	Pro	Ile	Ile	Ala	Arg	Tyr	Ile	Arg	Leu	His
1250					1255						1260			
Pro	Thr	His	Tyr	Ser	Ile	Arg	Ser	Thr	Leu	Arg	Met	Glu	Leu	Met
1265					1270						1275			
Gly	Cys	Asp	Leu	Asn	Ser	Cys	Ser	Met	Pro	Leu	Gly	Met	Glu	Ser
1280					1285						1290			
Lys	Ala	Ile	Ser	Asp	Ala	Gln	Ile	Thr	Ala	Ser	Ser	Tyr	Phe	Thr
1295					1300						1305			
Asn	Met	Phe	Ala	Thr	Trp	Ser	Pro	Ser	Lys	Ala	Arg	Leu	His	Leu
1310					1315						1320			
Gln	Gly	Arg	Ser	Asn	Ala	Trp	Arg	Pro	Gln	Val	Asn	Asn	Pro	Lys
1325					1330						1335			
Glu	Trp	Leu	Gln	Val	Asp	Phe	Gln	Lys	Thr	Met	Lys	Val	Thr	Gly
1340					1345						1350			
Val	Thr	Thr	Gln	Gly	Val	Lys	Ser	Leu	Leu	Thr	Ser	Met	Tyr	Val
1355					1360						1365			
Lys	Glu	Phe	Leu	Ile	Ser	Ser	Ser	Gln	Asp	Gly	His	Gln	Trp	Thr
1370					1375						1380			
Leu	Phe	Phe	Gln	Asn	Gly	Lys	Val	Lys	Val	Phe	Gln	Gly	Asn	Gln
1385					1390						1395			
Asp	Ser	Phe	Thr	Pro	Val	Val	Asn	Ser	Leu	Asp	Pro	Pro	Leu	Leu
1400					1405						1410			
Thr	Arg	Tyr	Leu	Arg	Ile	His	Pro	Gln	Ser	Trp	Val	His	Gln	Ile
1415					1420						1425			
Ala	Leu	Arg	Met	Glu	Val	Leu	Gly	Cys	Glu	Ala	Gln	Asp	Leu	Tyr
1430					1435						1440			

<210> SEQ ID NO 11
<211> LENGTH: 4368
<212> TYPE: DNA

-continued

```

<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polynucleotide

<400> SEQUENCE: 11
atgcagatt agctgagcac ctgtttcttc ctgtgcctgc tgaggttctg cttctctgcc      60
accaggagat actacactggg ggctgtggag ctttcttggg actacatgca gtctgacactg    120
ggggagatgc ctgtggatgc caggttccca cccagagtgc ccaaattcctt cccattcaac    180
acctctgtgg tctacaagaa gaccctctt gtggagttca ctgaccacct gttcaacatt    240
gccaatccca gccaccctg gatggactc ctgggaccca ccattcaggc tgagggttat    300
gacactgtgg tcatacaccct caagaacatg gcctcccacc ctgtgagcct gcatgctgtg    360
ggggtcagct actggaaaggc ctctgagggg gctgagttatg atgaccagac ctcccaagg    420
gagaaggagg atgacaaagt gttccctggg ggcagccaca cctatgtgtg gcaggctc      480
aaggagaatg gccccatggc ctctgaccca ctctgcctga cctactccctt cctttctcat    540
gtggacacctgg tcaaggaccc taactctggta ctgattgggg ccctgttgtt gtgcaggagg    600
ggctccctgg ccaaagagaa gaccaggacc ctgcacaagt tcatttcctt gtttgcgtc      660
tttgatgagg gcaagagctg gcactctgaa accaagaact ccctgtatgca ggacaggatg    720
gtgcctctg ccaggccctg gcccaagatg cacactgtga atggctatgt gaacaggagc    780
ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcattgtatg tggcatgggg    840
acaaccctg aggtgcactc catttcctg gagggccaca ccttccttgtt caggaaccac    900
agacaggcca gcctggagat cagccccatc accttcctca ctgcccagac cctgtatg      960
gacctcggac agttccctgct gttctgccac atcagctccc accagcatga tggcatgggg   1020
gcctatgtca aggtggacag ctgcccctg gagccacagc tcaggatgaa gaacaatgag    1080
gaggctgagg actatgtga tgacactgact gactctgaga tggatgttgtt ccgccttgat   1140
gatgacaaca gccccatctt cattcagatc aggtctgtgg ccaagaaaca ccccaagacc   1200
tgggtgact acattgtgc tgaggaggag gactggact atgccccact ggtctggcc      1260
cctgtatgaca ggagctacaa gagccagttac ctcaacaatg gcccacagag gattggacgc   1320
aagtacaaga aagtcaagggtt catggccatc actgtatgaaa cttcaagac cagggaggcc  1380
attcagcatg agtctggcat cctggggcca ctctgtatg gggagggtgg ggacaccctg   1440
ctcatcatct tcaagaacca ggcctccagg ccctacaaca tctaccacaca tggcatcaact 1500
gatgtcaggc ccctgtacag ccgcaggctg ccaaagggggg tggatgttgtt caaggactc  1560
ccatttcgc ctggggagat cttcaagtac aagtggactg tcactgtggta ggatggacca   1620
accaaatctg accccagggtg cctcaccaga tactactcca gttttgtgaa catggagagg  1680
gacctggcct ctggcctgat tggcccactg ctcatctgct acaaggagtc tggacccag     1740
agggaaacc agatcatgtc tgacaagagg aatgtgattc tggatgttgtt ctttgcgtatg  1800
aacaggagct ggtacactgac tgagaacatt cagcgcttcc tgcccaaccc tgctgggtg  1860
cagctggagg accctgagtt ccaggccagc aacatcatgc actccatcaa tggctatgtg   1920
tttgacagcc tccagtttc tggatgttgtt ctttgcgtatg ctttgcgtatg ctttgcgtatg 1980
atggggccccc agactgactt cttttctgtc ttcttctctg gctacaccctt caaacacaag 2040
atgggtatg aggacaccct gaccctttc ccattctctg gggagactgt gttcatgagc   2100

```

-continued

atggagaacc ctggectgtg gattctggga tgccacaact ctgacttccg caacagggc	2160
atgactgccc tgctcaaagt ctcctcctgt gacaagaaca ctggggacta ctatgaggac	2220
agctatgagg acatctctgc ctacctgctc agcaagaaca atgccattga gcccaggagc	2280
ttcagecaga attccagaca ccccagcacc agggagatca ccaggaccac cctccagtct	2340
gaccaggagg agattgacta tcatgacacc atttctgtgg agatgaagaa agaggactt	2400
gacatctatg acgaggacga gaaccagagc ccaaggagct tccagaagaa gaccaggac	2460
tacttcattg ctgctgtgg a gcgcctgtgg gactatggca tgagctccag cccccatgtc	2520
ctcaggaaca gggccctgtc tggctctgtg ccacagttca agaaagtgg cttccaagag	2580
ttcaactgatg gcagettcac ccagccctg tacagagggg agctgaatga gcacctggga	2640
ctccctggcc catacatcag ggctgaggtg gaggacaaca tcatggtgac cttccgcaac	2700
caggcctcca ggcctacag cttctacagc tccctcatca gctatgagga ggaccagagg	2760
cagggggctg agccacgcaa gaactttgtg aaacccaatg aaaccaagac ctacttctgg	2820
aaagtcacgc accacatggc ccccaccaag gatgagttt actgcaaggc ctgggcctac	2880
ttctctgtatg tggacctgg a gaaggatgtg cactctggcc tgattggccc actcctggc	2940
tgccacacca acaccctgaa ccctgccc ggaaggcaag tgactgtgca ggagtttgc	3000
ctcttcttca ccatcttga tgaacccaag agctggtaact tcaactgagaa catggagcgc	3060
aactgcaggc ccccatgcaa cattcagatg gaggacccca ctttcaaaaga gaactaccgc	3120
ttccatgcca tcaatggcta catcatggac accctgcctg ggcttgcata ggcaggac	3180
caaggatca ggtggtaacct gcttctatg ggcttcaatg agaacattca ctccatccac	3240
ttctctggc atgtcttac tgcgcgaag aaggaggatg acaagatggc cctgtacaac	3300
ctctaccctg gggctttga gactgtggag atgctgcctt ccaaagctgg catctggagg	3360
gtggagtgcc tcattggga gcaacctgcat gctggcatga gcaacctgtt cctggctac	3420
agcaacaatg gccagacccc cctggaaatg gcttgcgc acatcaggaa cttccagatc	3480
actgcctctg gccagtatgg ccagtggcc cccaaatgtt ccaggctcca ctactctgg	3540
tccatcaatg cctgggcac caaggagcca ttcaatgtgg tcaaaatgtt cctgtggcc	3600
cccatgatca tccatggcat caagacccag gggccaggc agaagtttcc cagcctgtac	3660
atcagccatg tcatcatcat gtacagcctg gatggcaaga aatggcagac ctacagaggc	3720
aactccactg gaacactcat ggtttttt ggcaatgtgg acagctctgg catcaagcac	3780
aacatcttca acccccaat catggccaga tacatcaggc tgcacccac ccactacagc	3840
atccgcagca ccctcaggat ggagctgtatg ggctgtgacc tgaactccctg cagcatgcc	3900
ctgggcattgg agagcaaggc catttctgtat gcccagatca ctgcctccag ctacttacc	3960
aacatgtttt ccacctggag cccaaatgtt gccaggctgc acctccagg aaggagcaat	4020
gcctggaggc cccaggatcaa caacccaaatg gagtggctgc aggtggactt ccagaagacc	4080
atgaaggatca ctgggggtgac cacccagggg gtcaagagcc tgctcaccag catgtatgt	4140
aaggagttcc tgatcagctc cagccaggat ggccaccatg ggaccctt cttccagaat	4200
ggcaaggatca aggtgttcca gggcaaccag gacagttca cccctgtggt gaacagctcg	4260
gacccccc tccgtaccag atacctgagg attcacccca agagctggg ccaccagatt	4320
gcctgagga tggaggctt gggatgtgag gcccaggacc tgcactga	4368

-continued

<210> SEQ ID NO 12
<211> LENGTH: 1455
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polypeptide

<400> SEQUENCE: 12

Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe
1 5 10 15

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

Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg
35 40 45

Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val
50 55 60

Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile
65 70 75 80

Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln
85 90 95

Ala Glu Val Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser
100 105 110

His Pro Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser
115 120 125

Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp
130 135 140

Asp Lys Val Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu
145 150 155 160

Lys Glu Asn Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser
165 170 175

Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile
180 185 190

Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr
195 200 205

Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly
210 215 220

Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp
225 230 235 240

Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr
245 250 255

Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val
260 265 270

Tyr Trp His Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile
275 280 285

Phe Leu Glu Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser
290 295 300

Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met
305 310 315 320

Asp Leu Gly Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His
325 330 335

Asp Gly Met Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro

-continued

340	345	350
Gln Leu Arg Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp		
355	360	365
Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser		
370	375	380
Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr		
385	390	395
Trp Val His Tyr Ile Ala Ala Glu Glu Asp Trp Asp Tyr Ala Pro		
405	410	415
Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn		
420	425	430
Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met		
435	440	445
Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu		
450	455	460
Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu		
465	470	475
Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro		
485	490	495
His Gly Ile Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys		
500	505	510
Gly Val Lys His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe		
515	520	525
Lys Tyr Lys Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp		
530	535	540
Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg		
545	550	555
Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu		
565	570	575
Ser Val Asp Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val		
580	585	590
Ile Leu Phe Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu		
595	600	605
Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp		
610	615	620
Pro Glu Phe Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val		
625	630	635
Phe Asp Ser Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp		
645	650	655
Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe		
660	665	670
Ser Gly Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr		
675	680	685
Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro		
690	695	700
Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly		
705	710	715
Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp		
725	730	735
Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys		
740	745	750

-continued

Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser Gln Asn Ser Arg His Pro
 755 760 765
 Ser Thr Arg Glu Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln Glu Glu
 770 775 780
 Ile Asp Tyr Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu Asp Phe
 785 790 795 800
 Asp Ile Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe Gln Lys
 805 810 815
 Lys Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp Asp Tyr
 820 825 830
 Gly Met Ser Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln Ser Gly
 835 840 845
 Ser Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr Asp Gly
 850 855 860
 Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu His Leu Gly
 865 870 875 880
 Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp Asn Ile Met Val
 885 890 895
 Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe Tyr Ser Ser Leu
 900 905 910
 Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg Lys Asn
 915 920 925
 Phe Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys Val Gln His
 930 935 940
 His Met Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala Trp Ala Tyr
 945 950 955 960
 Phe Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly Leu Ile Gly
 965 970 975
 Pro Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro Ala His Gly Arg
 980 985 990
 Gln Val Thr Val Gln Glu Phe Ala Leu Phe Phe Thr Ile Phe Asp Glu
 995 1000 1005
 Thr Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg Asn Cys Arg
 1010 1015 1020
 Ala Pro Cys Asn Ile Gln Met Glu Asp Pro Thr Phe Lys Glu Asn
 1025 1030 1035
 Tyr Arg Phe His Ala Ile Asn Gly Tyr Ile Met Asp Thr Leu Pro
 1040 1045 1050
 Gly Leu Val Met Ala Gln Asp Gln Arg Ile Arg Trp Tyr Leu Leu
 1055 1060 1065
 Ser Met Gly Ser Asn Glu Asn Ile His Ser Ile His Phe Ser Gly
 1070 1075 1080
 His Val Phe Thr Val Arg Lys Lys Glu Glu Tyr Lys Met Ala Leu
 1085 1090 1095
 Tyr Asn Leu Tyr Pro Gly Val Phe Glu Thr Val Glu Met Leu Pro
 1100 1105 1110
 Ser Lys Ala Gly Ile Trp Arg Val Glu Cys Leu Ile Gly Glu His
 1115 1120 1125
 Leu His Ala Gly Met Ser Thr Leu Phe Leu Val Tyr Ser Asn Lys
 1130 1135 1140

-continued

Cys	Gln	Thr	Pro	Leu	Gly	Met	Ala	Ser	Gly	His	Ile	Arg	Asp	Phe
1145				1150						1155				
Gln	Ile	Thr	Ala	Ser	Gly	Gln	Tyr	Gly	Gln	Trp	Ala	Pro	Lys	Leu
1160				1165						1170				
Ala	Arg	Leu	His	Tyr	Ser	Gly	Ser	Ile	Asn	Ala	Trp	Ser	Thr	Lys
1175				1180						1185				
Glu	Pro	Phe	Ser	Trp	Ile	Lys	Val	Asp	Leu	Leu	Ala	Pro	Met	Ile
1190				1195						1200				
Ile	His	Gly	Ile	Lys	Thr	Gln	Gly	Ala	Arg	Gln	Lys	Phe	Ser	Ser
1205				1210						1215				
Leu	Tyr	Ile	Ser	Gln	Phe	Ile	Ile	Met	Tyr	Ser	Leu	Asp	Gly	Lys
1220				1225						1230				
Lys	Trp	Gln	Thr	Tyr	Arg	Gly	Asn	Ser	Thr	Gly	Thr	Leu	Met	Val
1235				1240						1245				
Phe	Phe	Gly	Asn	Val	Asp	Ser	Ser	Gly	Ile	Lys	His	Asn	Ile	Phe
1250				1255						1260				
Asn	Pro	Pro	Ile	Ile	Ala	Arg	Tyr	Ile	Arg	Leu	His	Pro	Thr	His
1265				1270						1275				
Tyr	Ser	Ile	Arg	Ser	Thr	Leu	Arg	Met	Glu	Leu	Met	Gly	Cys	Asp
1280				1285						1290				
Leu	Asn	Ser	Cys	Ser	Met	Pro	Leu	Gly	Met	Glu	Ser	Lys	Ala	Ile
1295				1300						1305				
Ser	Asp	Ala	Gln	Ile	Thr	Ala	Ser	Ser	Tyr	Phe	Thr	Asn	Met	Phe
1310				1315						1320				
Ala	Thr	Trp	Ser	Pro	Ser	Lys	Ala	Arg	Leu	His	Leu	Gln	Gly	Arg
1325				1330						1335				
Ser	Asn	Ala	Trp	Arg	Pro	Gln	Val	Asn	Asn	Pro	Lys	Glu	Trp	Leu
1340				1345						1350				
Gln	Val	Asp	Phe	Gln	Lys	Thr	Met	Lys	Val	Thr	Gly	Val	Thr	Thr
1355				1360						1365				
Gln	Gly	Val	Lys	Ser	Leu	Leu	Thr	Ser	Met	Tyr	Val	Lys	Glu	Phe
1370				1375						1380				
Leu	Ile	Ser	Ser	Ser	Gln	Asp	Gly	His	Gln	Trp	Thr	Leu	Phe	Phe
1385				1390						1395				
Gln	Asn	Gly	Lys	Val	Lys	Val	Phe	Gln	Gly	Asn	Gln	Asp	Ser	Phe
1400				1405						1410				
Thr	Pro	Val	Val	Asn	Ser	Leu	Asp	Pro	Pro	Leu	Leu	Thr	Arg	Tyr
1415				1420						1425				
Leu	Arg	Ile	His	Pro	Gln	Ser	Trp	Val	His	Gln	Ile	Ala	Leu	Arg
1430				1435						1440				
Met	Glu	Val	Leu	Gly	Cys	Glu	Ala	Gln	Asp	Leu	Tyr			
1445				1450						1455				

```

<210> SEQ ID NO 13
<211> LENGTH: 4374
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polynucleotide

```

<400> SEQUENCE: 13

atgcagattg agctgtccac ctgcttcttt ctgtgcctgc tgagattctg cttctctgcc 60

-continued

accaggagat actacctggg ggctgtggaa cttcttggg actacatgca gtctgacctg	120
ggagagctgc ctgtggatgc caggttccca cccagagtgc ccaagtcctt cccattcaac	180
acctctgtgg tctacaagaa gacactctt gtggattca ctgaccacct gttcaacatt	240
gcaaaaaccca gaccaccctg gatgggactc ctgggaccca ccattcaggc tgaggtgtat	300
gacactgtgg tcataccct caagaacatg gcatcccacc ctgtgtctct gcatgtgtg	360
ggagtctcat actggaaagc ctctgaaggg gctgagttatg atgaccagac atccagaga	420
gagaaagagg atgacaaggt gtccctggg ggatctcaca cctatgtgtg gcaagtctc	480
aaggagaatg gacccatggc atctgaccca ctctgcctga catactccctt ccttctcat	540
gtggacctgg tcaaggaccc caactcttggaa ctgattgggg cactgtgtt gtgcaggaa	600
ggatccctgg ccaaggagaa aacccagaca ctgcacaagt tcattctctt gtttgcgtc	660
tttgtatgggg gcaagtcttgc gcactcttgc acaaagaact cctgtatgca agacaggat	720
gtgcctctg ccaggccatg gcccaagatg cacactgttga atggctatgtt gaacagatca	780
ctgcctggac tcattggctg ccacaggaaa tctgtctactt ggcattgtatg tggcatgggg	840
acaaccctg aagtgcactc cattttctg gagggacaca ctttctgtt caggaaccac	900
agacaaggct ctctggagat ctctccatc accttccatc ctgcacagac actgtgtat	960
gaccttggac agtttctgtt gtttgcac accatgttga tggcatggaa	1020
gcctatgtca aggtggactc atgccttagt gaaaccacatc tcaggatgtt gaaatgtt	1080
gaggctgagg actatgttga tgacctgtt gactctgtt gggatgtgtt cagatttgtt	1140
gatgacaact ctccatctt cattcagatc aggtctgtgg caaagaaaca ccccaagaca	1200
tgggtgcact acattgttgc tgaggaagat gactggactc atgcaccactt ggtctggcc	1260
cctgtatgaca ggagctacaa gtctcagatc ctcaacaatg gcccacaaag aatttggaa	1320
aagtacaaga aagttagt catggccatc actgtatgtt ctttcaagatc aagatgttgc	1380
attcagatc agtctggatc tctggatca ctctgtatg gggaaatgtt agacaccctt	1440
ctcatcatct tcaagaacca ggcctccagg ccctacaaca tctaccacata tggcatcact	1500
gtatgttgc cctgtatgc caggagactt cttttttttt cttttttttt cttttttttt	1560
cccttctgc ctggagat cttcaagatc aagtggactt tcactgttgc ggttggatca	1620
acaaagtctg accccagggtt ctttccatc tactacttctt cttttttttt cttttttttt	1680
gaccttggatc ctggactgtt tggaccactt ctttccatc tactacttctt cttttttttt	1740
agaggcaacc agatcatgtt tgacaagatc aatgttgcattt tttttttttt cttttttttt	1800
aacatgttgc ggttggatc tgagaacattt ctttccatc tactacttctt cttttttttt	1860
caacttggaa acccttgcattt ccaggcaatc aacatcatgtt tactacttctt cttttttttt	1920
tttgcattttc tccatgttgc ttgttgcattt ctttccatc tactacttctt cttttttttt	1980
attttttttt ctttccatc tactacttctt ctttccatc tactacttctt cttttttttt	2040
atgggtgtatg aggacaccctt gacactcttc ccatttctgtt gggaaatgtt gtttcatgtt	2100
atgggatcacc ctggactgtt gatttggatc tgccacaaactt ctgacttgc aaacaggaa	2160
atgactgcac tgctcaatgtt ctcttccatc gacaagaaca ctggggactt ctatgttgc	2220
tcttatgagg acatctctgc ctacatgttgc agcaagaaca atgcccatttgc gcccacaa	2280
ttctctcaga atccacatgtt ctttccatc tactacttctt ctttccatc tactacttctt	2340

-continued

cagtctgacc	aggaagagat	tgactatgt	gacaccattt	ctgtggagat	gaagaaggag	2400
gactttgaca	tctatgtga	ggacgagaac	cagtctccaa	gatcattcca	gaagaagaca	2460
agacactact	tcattgctgc	tgtggaaaga	ctgtgggact	atggcatgtc	ttcctctccc	2520
catgtcctca	ggaacagggc	acagtctggc	tctgtgccac	agttcaagaa	agtggcttc	2580
caggagttca	ctgatggctc	attcacccag	cccctgtaca	gaggggaact	gaatgagcac	2640
ctgggactcc	tgggaccata	catcagggt	gaggtggaaag	acaacatcat	ggtgacattc	2700
agaaaccagg	cctccaggcc	ctacagcttc	tactcttccc	tcatcagcta	tgaggaagac	2760
cagagacaag	gggctgagcc	aagaaagaac	tttgtgaaac	ccaatgaaac	caagacctac	2820
ttctggaaag	tccagcacca	catggcaccc	accaaggatg	agtttgactg	caaggcctgg	2880
geatacttct	ctgatgtgga	cctggagaaa	gatgtgcact	ctggcctgat	tggcccactc	2940
ctggtctgcc	acaccaacac	cctgaaccct	gcacatggaa	ggcaagtgac	tgtgcaggag	3000
tttgcctct	tcttcaccat	ctttgatgaa	accaagtcat	ggtacttcac	tgagaacatg	3060
gagagaaact	gcagagcacc	atgcaacatt	cagatggaa	accccaccc	caaggagaac	3120
tacaggttcc	atgccccatcaa	tggctacatc	atggacaccc	tgcctgggt	tgtcatggca	3180
caggaccaga	gaatcagatg	gtacctgctt	tctatgggt	ccaatgagaa	cattcactcc	3240
atccacttct	ctgggcatgt	cttcaactgt	agaaagaagg	aggaatacaa	gatggcctg	3300
tacaacccct	accctgggtt	ctttgagact	gtggagatgc	tgcctccaa	agctggcattc	3360
tggaggggtgg	aatgectcat	tggggagcac	ctgcatgtg	gcatgtcaac	cctgttctcg	3420
gtctcacagca	acaagtgcac	gacaccctgt	ggaatggcc	ctggccacat	cagggacttc	3480
cagatcaactg	cctctggcca	gtatggccag	tgggcaccca	aactggccag	gctccactac	3540
tctggctcca	tcaatgcatg	gtcaaccaag	gagccattct	cttggatcaa	ggtggactg	3600
ctggcaccca	tgtatcattca	tggcatcaag	acacaggggg	caagacagaa	attctccct	3660
ctgtacatct	cacagttcat	catcatgtac	tctctggatg	gcaagaagtg	gcagacatac	3720
agaggcaact	ccactggcac	cctcatggtc	ttctttggca	atgtggacag	ctctggcattc	3780
aagcacaaca	tcttcaaccc	tcccatcatt	gcccata	tcaaggtgca	ccccaccac	3840
tactcaatca	gatcaaccc	caggatggaa	ctgatggat	gtgacctgaa	ctcctgctca	3900
atgccccctgg	gaatggagag	caaggccatt	tctgtatgccc	agatcaactgc	atcccttac	3960
ttcaccaaca	tgtttgccac	ctggtcacca	tcaaaagcca	ggctgcaccc	ccaggaaaga	4020
agcaatgcct	ggagacccca	ggtcaacaac	ccaaaggaa	ggctgcaagt	ggacttccag	4080
aagacaatga	agtcactgg	ggtgacaacc	cagggggtca	agtctctgt	caccaatg	4140
tatgtgaagg	agttcctgtat	ctcttcctca	caggatggcc	accagtggac	actcttcttc	4200
cagaatggca	aagtcaaggt	gttccaggcc	aaccaggact	ctttcacacc	tgtggtaac	4260
tcactggacc	ccccccctct	gacaagatac	ctgagaattc	accccccagtc	ttgggtccac	4320
cagattgccc	tgagaatgga	agtcctggga	tgtgaggcac	aagacctgta	ctga	4374

```

<210> SEQ ID NO 14
<211> LENGTH: 4374
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polynucleotide
  
```

-continued

<400> SEQUENCE: 14

atgcagatcg aactgagcac ttgtttcttc ctgtgtctcc tgcgccttgc cttctccgcc	60
acaaggagat actatctcg tgccgtggag ctcagctggg actacatgca gagcgacttg	120
ggtaactgc ctgtggacgc caggttcca ccccgcgtgc ccaagagt cccgttcaac	180
accagtgtcg tgtacaagaa aaccctcttc gtgaaattca ccgaccacct gttcaacatc	240
gccaaacgcg cccctccctg gatggggctg ctggcccgca cgatccaggc tgaggcttat	300
gacacggtgg tgattaccct caagaacatg gctagccacc cggtgagcct gcacgcgtg	360
ggcgtgtccct attggaaagc gtcggagggt gcccggatcg atgaccagac ttacagcgg	420
gagaaggaag acgacaaagt gttccccggg gttcccaca cctatgtctg gcaggctctg	480
aaggagaatg gtcctatggc ctccgaccca ttgtgcctca cctactctta cctaagccat	540
gtggatctcg tcaaggaccc gaaactgggg ctgatcgccg ccctgtctgt gtgcgggag	600
ggctcaetgg ccaaggagaa gacccaaact ctgcacaatg tcattctgt gttcgcgta	660
ttcgacgagg ggaagtccctg gcactccgag accaagaaca gcctgatgca ggaccgcac	720
gcagcctcgcc cccgtgcgtg gccaaagatg cacaccgtga acggctacgt taacaggagc	780
ctacccggcc tgatcggtcg ccacccgaaa tcggtctact ggcattgtat cgaaatggc	840
acaacgcggc aggtccacag tatcttcctc gaggccaca ctttccttgtt ccggaatcac	900
cggcaggcca gcctggagat cagccccata acctttctga cggcgcagac cttactcatg	960
gatctcgcc agttccctctt gttctgccac atttcgtccc accagcacga tggatggaa	1020
gcataatgtga aagtggactc ctggcccgag gaaccccagc ttagatgaa gaacaatgag	1080
gaggccgagg actacgacga tgaccttacc gattcagaaa tggacgtagt acgctttgac	1140
gacgacaact ctccatccctt catacagatt cgctccgtcg ccaagaagca ccctaagact	1200
tgggtgact acatcgccgc cgaggaggag gactggatt atgctccctt ggtgtggcc	1260
cccgacgacc gcagctacaa gagccagttac ctgaaataacg ggccccagcg catcgccgg	1320
aagtacaaga aagtgcggtt catggcttac acggacgaga cttcaagac ccggagggct	1380
atccagcatg agagcgccat cttggggccc ctctgtacg gcaaggtgg agacacactg	1440
ctgatcatct tcaagaacca ggccgaggcc ccctacaaca tctacccca cggcattacc	1500
gatgtccggc cgttgtacag ccgacggctg cccaaaggccg tgaagcacct gaaggacttt	1560
ccgatctgc cggcggagat cttcaagtac aagtggactg tgaccgtgaa ggatggccg	1620
accaagagcg atcccgccgt cctgacccgt tactactcca gtttgcata tatggagcgc	1680
gacctcgcta gcgggttgat tggccctctg ctgatctgct acaaggagtc cgtggaccag	1740
aggggaaatc agatcatgag tgacaagagg aacgtgtatcc ttgttccctgtt gttcgaccaa	1800
aaccgcagct ggtatctcac cgagaatatac cagcgcttc tgcccaaccc ggccgggtgt	1860
cagctggagg accccgagtt tcaggccagc aacatcatgc attctatcaa cggatatgtg	1920
tttgattccc tgcagctctc agtgtgtctg caccgggtcg cttactggta tatccctcagc	1980
attggggcac agaccgactt cctgagcggtg ttcttctccg ggtataccctt caagcacaag	2040
atgggtgtacg aggataccct gaccctgttc ccctttagcg gcaaaaccgt ttatgtct	2100
atggagaacc cccggctctg gatccttggc tgccataact ccgacttccg caaccgcgg	2160
atgaccgcgc tcctgaaagt gtcgagttgt gacaagaaca ccggcgacta ttacgaggac	2220

-continued

agttacgagg acatctctgc gtacctcctt agcaagaata acgccatcgaa gccaagatcc	2280
ttcagccaga accccccagt gctgaagagg catcagcggg agatcaccccg cacgaccctg	2340
cagtcggatc aggaggagat tgattacgac gacacgtaca gtgtggagat gaagaaggag	2400
gacttcgaca tctacgacga agataaaaac cagtccctc ggtccttcca aaagaagacc	2460
cggcactact tcategcccgc tgtggAACgc ctgtgggact atggaatgtc ttctagccct	2520
cacgtttga ggaaccgcgc ccagtccggc agcgtcccccc agttcaagaa agtgtgttc	2580
caggagttca ccgacggctc cttcacccag ccactttacc ggggcgagct caatgaacat	2640
ctgggcctgc tgggacccta catcagggtc gaggtggagg acaacatcat ggtgacattc	2700
cggaatcagg ccagcagacc atacagttt tacagttcac tcataccctta cgaggaggac	2760
cagcgcagg gggctgaacc ccgtaaagaaat ttcgtgaagc caaacgaaac aaagacctac	2820
ttctggaaagg tccagcacca catggcacctt accaaggacg agttcgatgtc caaggccctgg	2880
gcctacttctt ccgacgttggc cctggagaaa gatgtgcaca gcccgcctgtat tggccctctg	2940
ctgggtgttc acacgaacac actcaaccct gcacacgggc ggcagggtcac tggcaggaa	3000
ttcgcctgt tctttaccat ctttgatgag acgaagtcctt ggtatccatcg cggaaacatg	3060
gagaggaact gcccgcgcacc ctgcaacatc cagatggaaat atccgacattt caaggagaac	3120
taccggttcc atgcatcaa tggctacatc atggacaccc tggctggctt cgtgatggcc	3180
caagaccaggc gtatccgctg gtatctgctg tcgatgggtt ccaacggagaa catccatgt	3240
atccacttca cggggcatgt ctacgggtc agggaaaagg aggagtacaa gatggactg	3300
tacaacccctt atccggcgtt gttcgagacc gtggagatgc tggctccaa ggccggcatc	3360
tggagagtgg aatgcctgtat cggcgagcac ctccacgctg ggatgtccac gctgttctc	3420
gtttacgca ataagtgcac gaccctctg ggcacatggc gcccgcacat ccgcgacttc	3480
cagattacag ccacggccca gtacggtcag tggctccaa agctggcccg tctgcactac	3540
tccggatcca tcaacgcctg gtccaccaag gaaccgttctt cctggatcaa agtagacatg	3600
ctagccccca tgatcattca cggcatcaag acacaaggcg cccgacagaa gttctcgagc	3660
ctctatatctt cccagttcat catcatgtat agcctggacg gaaagaatgt gcagacttac	3720
cgccggaaact cgacaggacg cctgtatggta ttcttcggta acgtggacag ctccggatc	3780
aagcacaaca tcttcaaccc acccattatc gcccgttaca tccgcctgca ccccaactcac	3840
tatagcatta ggtccacccctt gcaatggag ctcatgggtt gtgacctgaa cagctgttagc	3900
atgcctctcg gcatggagtc taaggcgatc tccgacgcac agataacggc atcatccatc	3960
tttaccaaca tggctgttac ctggcccccc tccaaaggccc gactccaccc gcaaggagaa	4020
tccaaacgcctt ggccggccaca ggtcaacaat cccaaaggagg ggctgcagaat ggactttcag	4080
aaaactatga aagtccacgg agtggaccaca cagggaggatgaa agtctctgtt gaccagcatg	4140
tacgtgaagg agttccatctt ctcccgatgtc caggatggcc accagtggac gttgttctc	4200
caaaaacggta aagtccaaatg cttccaaagggg aaccaggaca gctttacacc cgtcgtgaac	4260
tccctggacc ccccgcttctt cactagatac ctccgcattcc accctcagag ctgggtgcac	4320
cagattggcc tgcgtatggaa ggttctgggg tggtaagcccc aggacctgtatctaa	4374

-continued

```

<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
polynucleotide

<400> SEQUENCE: 15

atgcagattt agctctccac ctgtttcttt ctctgccttc ttgcgttctg cttttctgcc      60
acacgcaggactatattggg aacgtggaa ctgagctggg attacatgca gagtgacatt      120
ggtaacttc ctgtggacgc tcgtttcca cctagatc ccaaggccctt ccccttcaac      180
acctcagtgg tctacaagaa aacgctgttt gtggagttca ctgaccacct cttcaacatt      240
gccaacacaa gaccccttg gatgggattt ctgggaccca caatacaga agaagtctac      300
gacacggtgg tgattaccctt gaagaacatg gctcacacc ctgtttcaact tcacgctgtt      360
ggggtcagtt attggaaagc ctcagagggt gctggaaatc agatcaaacc cagccagagg      420
gagaaggaag atgacaaggat ctttcctggg ggttagccata cctatgttttgcagggtgt      480
aaagagaatg ggcctatggc ctctgatccc ttgtgcctca catacttcta cctgagtcac      540
gtcgacotgg taaaagaccc gaatagcggt ctgattgggt cactgttgtt ttgttagagag      600
gggagtttgg ccaaggagaa aactcagact ctccacaagt ttatcctcctt gtttgctgt      660
ttcgaecagg gcaaggcttg gcactctgaa acaaagaact cctgtatgca ggacagagat      720
gtgcacatcg caagggcttg gccaaaaatg cacacagtga acggctatgt gaatcgatca      780
ctgccaggac tgataggctg tcatcgcaag tcagtgtatt ggcacgttat cgggatggga      840
acaactccatg aagtgcacag catttccctt gagggccaca ctttctgtt tcggaatcat      900
agacaggcca gccttgatgat cagccaaatc accttctgaa ctgccccaaac cttgtgtatg      960
gatctggac agtctccctt gtttgcac atctccccc accaacatgca cgggatggag      1020
gtttatgtga aggtcgatag ctgtccggag gaaccacaac tgaggatgaa gaacaacgaa      1080
gaggcagagg actatgacga cgatctgact gacagtggaa tggacgttgt tcgggtcgac      1140
gatgacaatt ctcccttatt tatccagatc cgttccgtgg ccaagaagca ccccaagact      1200
tgggttattt acatcgatcg tgaggaggag gattggact acgcgeccctt ggtgtggcc      1260
ccagacgatc gtcatacaa gagccagatc cttacaatg gtccacaaag gatcgccgg      1320
aagtacaaga aggttagatt tatggctt accgacgaga cttttaaaac tagggaaagca      1380
attcagcatg aaagtggcat tcttggaccc ctgctgtatg gcgagggttgg cgacaccctg      1440
ctgattatct ttaagaacca ggcaagccgg ccctacaaca tctacccgca cggcataacc      1500
gatgtacgac ccctgtacag tcgcagactt cctaaagggg tgaaacaccc gaaggactt      1560
ccaaattctgc cggggagat cttcaagtat aaatggaccc tgacgggttga ggtatggccc      1620
acaaagtccg atccgagatg ctttacccga tattattcca gtttcgtgaa catggaaagg      1680
gacctggcca gccccgtat tggccactg ctgatttggt acaaggagtc tgcgtatcaa      1740
agaggaaacc aaataatgag cgacaaacgt aacgtcatcc tgcgtatgtt cttgtatgag      1800
aatagaagct ggtacccatc agaaaatatt cagcggttgc tgcctaaacc cgcaggcgatc      1860
cagctggaaatcccgatgtt ccaaggctca aacatcatgc atagcatcaa cggataacgtt      1920
ttcgatagcc tgcgtatgtc cgtctgttc catgaagtgg catattggta catcctgagt      1980
atcgccccgc agaccgactt cctgagcgtt ttctttctg gatacacgtt caaacacaaa      2040

```

-continued

atgggtctatg aagataccct gactctgttt ccattctcag gagagacagt ctttatgagt	2100
atggaaaatc ctggactgtg gatcccggc tgtcacaatt ctgatttcg gaacagaggc	2160
atgacagccc tgcttaaagt gagctcatgc gacaagaaca cccggtgatca ctacgaagat	2220
agctatgagg acatcaatgc gtatggctc tccaagaaca acgctatcga gccacggct	2280
ttcagtcaga atcctccgt tctgaagcg catcagcgaa aaataacacg cacaaccctt	2340
cagtcagacc aagaggaaat cgactacat gatactatct ctgtggagat gaagaaggag	2400
gatttcgaca ttacgacga ggacgagaat cagtcggccaa ggagcttca gaagaaaaca	2460
agacactatt tcattgccgc cgtggagcga ctgtggact acggcatgtc tagctctccg	2520
catgtactta gaaataggc acaaaggcga tccgtgcctc agtttaagaa agttgtcttt	2580
caggagttt aagatggctc cttcacccag cccttgtatc gggggaaact caatgaacac	2640
ctgggcctcc tgggtcctt tattaggccg gaagtcgagg acaatatcat ggtgaccctt	2700
aggaaccagg catctagacc ttactcttcc tactcctccc tgatatocta tgaggaggac	2760
cagcggcaag ggcgtgagcc tcggaagaac ttgtgaaacg caaatgaaac caaaacatac	2820
ttttggaaag tttagcacca catggctccc acgaaggacg aatttgactg taaaggctgg	2880
gcctacttct cagatgtaga tctcgagaaa gacgtgcact cagggtcatc tggtccccctc	2940
ctgggtctgtc atactaatac cctcaatcca gcacacggac gtcaggtaac cgtccaggaa	3000
tttgccctgt tctttaccat tttagatgactaaatccctt ggtactttac cgaaaacatg	3060
gagaggaatt gcagagcccc atgcaacatc cagatggagg accctacccaaagagaac	3120
tatcgcttcc atgcccattaa cggttacattt atggataactc tcccgaggact tggatggca	3180
caggatcagc ggataagatg gtatctgtt agcatggcactt ccaacgagaa tattcacagc	3240
atccatttctt ccggtcacgt gtttacatg agaaagaaag aagagtacaa gatggctctg	3300
tataatctct atccaggcgtt attcgaaacg gtggagatgt tgcctagccaa ggccggcatt	3360
tggcgagtag aatgccttat cggggacat ctgcgtccg gaatgagcac gctttcttg	3420
gtgtatagta acaagtgcga gactccgtc ggcgttatc acgggacttt	3480
cagattacgg ctageggcga gtatggcag tgggcacccaa aacttgcgcg actgeactat	3540
tcaggctcta tcaatgcgt gtcaccaag gaacccttctt ctggattaa ggtggactt	3600
ttggcgccca tgataatcca tggatcaaa acccaggccg ctgcgtacaa attctcatca	3660
ctctacatct ctcagttcat aataatgtat tcactggatg ggaagaaatg gcagacttac	3720
agaggaaaca gcacccggac gctgtatggt ttctttggca acgtggacacg cagcggcatc	3780
aaacacaaca tcttcataatcc tcccattttt gcccgttata ttagactgca tcccactcac	3840
tactctatac gcagcacact taggatggag ctcatggatg ggcacactgaa cagttgtatg	3900
atggcccttgg ggatggagtc caaagctata agcgacgcac aaattacagc tagctttac	3960
tttacgaata tggcgccac gtggagccca agcaagccc ggctgcattt gcagggtcg	4020
agtaatgtt ggcgcacaca ggtgaataac cctaaaggat ggttgcagt agattccag	4080
aaaactatga aggttaaccgg cgtcaactaca cagggagtca agtccctt gacctctatg	4140
tacgtcaagg agttcctgtat tagcagcgtt caggtggcc accaatggac actgttttc	4200
cagaatggga aagttaaagt atttcagggt aaccaggact ctttacacc tgggtgaat	4260
aggctcgacc caccctgtt gacacgatac ctccgcattcc accctcagtc ttgggtgcatt	4320

-continued

caaattggcc tgcgaatgg a ggtttggg tgcgaagctc aggacctcta ctga	4374
<210> SEQ_ID NO 16	
<211> LENGTH: 4374	
<212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide	
<400> SEQUENCE: 16	
atgcagatcg aactctctac ttgttcttc ctgtgccttc tgaggctctg cttctctgcc	60
actcgccat attacctcg ggccgtggag ttgagttggg actacatgca atcagatctg	120
ggcgaactcc ctgtggatgc ccgattccca ccgcgcgtgc ccaagtcttt cccatttaat	180
acttctgtgg tgtacaagaa gacattgttt gtggagttt ccgatcacct gttcaacatc	240
gccaaaccgc ggcggccatg gatgggtctg cttggggcca ccattcaagc ggaggcttat	300
gatacagtgg tgataacgct taagaacatg gcgagccacc cagtgtctct gcatgccgtt	360
ggtgtatcat attggaaggc cagcgaagga gcggagtagc atgaccagac ctctcagaga	420
gagaaggaag acgataaggt tttcctggc ggaagtcata catatgtatg gcaggtcttg	480
aaagagaatg ggccgatggc ttctgacccc cttgtttaa cctatagtt tctgagccac	540
gtggacctgg tcaaggaccc caacagtggt ctgattgggg ctctgcttgt ttgttagagag	600
ggtagcttgg ctaaggagaa aacccaaaca ctccataagt tcattttgtt gttcgccgtg	660
ttcgacgagg gaaagagttt gcacagcgaa acaaagaatt cactgatgca agacaggac	720
gcccgttccg caagggtttt gcctaagatg catacggtga atgggtatgt gaaccggac	780
ctcccgcccc tgatcggtt ccatacgcaag tctgtttact ggcacgtcat tggaatgggg	840
acaacgcac aggtacatag tatatttctt gaaggccaca cgttccctgt acggaaccac	900
cgacaggctt ccctggagat aagccccatt acctttctga ccgctcagac tctgctgatg	960
gaccttggcc agtttccctt gttctgccat attagcagcc accagcacga cggtatggaa	1020
gcatacgtga aagtgcatacg ctgtccttag gaggctcagc tcagaatgaa gaacaacag	1080
gaggccgaag actatgacga tgacccatac gattccgaga tggacgtggt gcgcttgc	1140
gacgataaca gtcctagttt cattcaaatac agatccctgtt caaaaagca tccaaagaca	1200
tgggtgcatt acattgcagc cgaagaggag gattggattt atgcgcctt tggatggct	1260
ccagatgaca ggagctataa gtcctgttcc ttgaacaacg ggccacagcg aatcggtaga	1320
aaatataaga aggtaaagatt catggccatc actgacgaaa cattttaaac cagggaaagct	1380
atccaaacacg aatctggaaat tctcgccct ctgtctacg gtgaggtggg ggacacccctt	1440
ctgatcattt taaaaatca ggcattccagg ctttacaaca tataccccc tggatcacc	1500
gatgtccgcc cgctgttcc cagaagactc cccaaaggag tgaaacatct gaaagat	1560
ccatccctgc cggccgagat ctttataac aaatggactg tgactgtaga ggacggccct	1620
acaaaatcg acccacgggt cctgacaaggatt tattacgtt gttcgatcaa catggaaac	1680
gacctcgcca gcggactcat tggccactg ttgatctgtt acaaagagtc agtggatcg	1740
agggaaatc agatcatgag cgataagaga aacgttatcc tggatgtt gttcgacgag	1800
aaccggctt ggtacccatc tgagaacatc cagaggcttcc tgccgaatcc ggctggcg	1860
cagctcgagg acccagagtt ccaggccagt aatataatgc actcaatcaa cggttatgt	1920

-continued

ttcgatagcc	tgcagctgag	cgtctgcctc	cacgaggtag	cctattggta	catattgtcc	1980
atcggggctc	agaccgattt	tctgtccgtg	ttcttttagcg	ggtataccctt	taaacataaa	2040
atgggtctatg	aagacacccct	gaccctgttc	ccattctccg	gtgagactgt	gttcatgtcc	2100
atggagaacc	cagggctgtg	gatcctgggg	tgtcacaata	gtgactttag	gaatcgggga	2160
atgacggcac	tgctgaaggt	gagttcttgc	gataaaaata	caggagattha	ctatgaggat	2220
agttacgagg	atatacgatgc	ctatctgttt	tcaaaaaaaca	acgcaattga	gccccggct	2280
ttctcacaaa	accccccggt	gctgaagcgc	caccagcgc	aaattacccg	gacaacccct	2340
cagtccgacc	aggaggaaat	cgattatgac	gatactatca	gtgttagaaat	aaaaaaggag	2400
gattttgata	tttacgacga	agacgagaac	cagtctccgc	gaagtttca	gaagaaaacg	2460
cgacactact	ttatagctgc	cgtggaacga	ctctgggatt	atggcatgtc	ctccagccct	2520
catgtcttta	gaaatcgagc	gcagagtgcc	tctgtgcctc	agttcaaaaa	ggttgtgttc	2580
caggaattca	ccgacggctc	atttacccag	ccgctgtaca	gaggcgaact	caacgaacac	2640
cttgggctgc	ttgggccata	tattcgagca	gaggtggaaag	ataatatcat	ggtaacccctt	2700
agaaaccagg	cgtcaagacc	ctattccctc	tacagttctc	tgatcagcta	cgaggaggac	2760
caaagacagg	gagctgaacc	caggaagaac	tttgtgaaac	ctaattgagac	caagacctac	2820
ttctggaaagg	tccagcacca	tatggcccca	actaaagatg	aattcgattg	caaggcctgg	2880
gcttatttca	gcgacgtgga	tctcgaaaag	gatgtgcaca	gcgggttgat	cgaccgctt	2940
ttgggtgccc	acacaaatac	cctcaatcct	gcccacgggc	ggcaggtcac	agttcaagag	3000
tttgcactct	tctttacaat	atttgacgag	acaaagtcat	ggtattttac	agagaatatg	3060
gagagaaatt	gtcgccgacc	ttgcaacatt	cagatggagg	acccacatt	taaggagaat	3120
tacagatttc	atgctatcaa	tgggtacatt	atggatactc	tgccctggct	ggtcatggcc	3180
caggatcagc	gcataaggtg	gtacttgctg	agcatgggat	ctaattgagaa	tatacacagc	3240
attcacttca	gtggccacgt	ttttactgtt	agaaagaagg	aggagtacaa	aatggcgctc	3300
tacaacctt	acccgggtgt	gtttgagaca	gtggagatgc	tgccaagcaa	ggcaggcatc	3360
tggagggttg	agtgtcttat	tggggagcat	ctgcatgtg	gaatgtccac	cctctttctt	3420
gtgtacagca	ataagtgcua	gacaccgctt	ggcatggcca	gcggccacat	tagggactt	3480
cagataactg	ccagtgccaca	gtacggccag	tgggctccca	agcttgcaag	actccactac	3540
tccggaagca	taaacgcatt	gagcaccaag	gaacccttct	cttggattaa	ggtggacctg	3600
ctggcgccaa	tgatcattca	cgccataaaa	acccaagggg	cacgacagaa	attttcatct	3660
ttgtatatta	gtcagtttat	catcatgtac	agcttggatg	gaaagaagtg	gcagacgtac	3720
aggggcaatt	otacaggaac	acttatggtg	ttttttggaa	atgtcgattc	cagcgggatc	3780
aaacataaca	tcttaatcc	tccttattatc	gcccgatata	tccgcctgca	ccctacgcatt	3840
tactccatca	ggtccacatt	gagaatggaa	ctgatgggg	gcgacctgaa	tagttgtagt	3900
atgccactgg	gcatggagtc	taaagccatc	agcgatgcac	agatcactgc	cagctttac	3960
ttcaccaaca	tgtttgcaac	ttggtcccc	tctaaagctc	gcctgcattt	gcagggacgc	4020
tcaaataatca	ggcgaccaca	ggtgaacaat	ccaaaagagt	ggctccaggt	cgactttcag	4080
aagacaatga	aggtaacagg	agtgacaacc	cagggtgtaa	aaagcctcct	tacgagttatg	4140
tacgttaagg	agtttctgtat	ttcttagctcc	caggacggac	accagtggac	tctgttcttc	4200

-continued

cagaacggca aagtgaaggt atttcaggga aaccaggatt ctttacccc ggttgtaat 4260
agcctggatc caccgtgct gacccgctat ctgagaattc atccacaatc ctgggtgcat 4320
cagattgcc tccggatgga agtgctcgcc tgtgaagtc aggatctgta ttag 4374

<210> SEQ ID NO 17
<211> LENGTH: 4374
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 17

atgcaaata agctctccac ctgtttcttt ctgtgccttt tgcgattctg ctttagtgcc 60
accagaagat actacctggg tgcaagtggaa ctgtcatggg actatatgca aagtgtatctc 120
ggtgagctgc ctgtggacgc aagatttcct cctagagtgc caaaatctt tccattcaac 180
acctcagtcg tgtacaaaaa gactctgttt gttagattca cggatcacct tttcaacatc 240
gctaagocaa ggccaccctg gatgggtctg ctaggtctta ccattccaggc tgaggtttat 300
gatacagtgg tcattacact taagaacatg gcttccatc ctgtcagttc tcatgctgtt 360
ggtgtatcct actggaaage ttctgaggga gctgaatatg atgatcagac cagtcaaaagg 420
gagaaaagaag atgataaagt ctcccttgtt ggaagccata catatgtctg gcaggctctg 480
aaagagaatg gtccaatggc ctctgaccca ctgtgcctta cctactcata tctttctcat 540
gtggacactgg taaaagactt gaattcaggc ctcattggag ccctactagt atgttagagaa 600
gggagtctgg ccaaggaaaa gacacagacc ttgcacaaat ttatactact ttttgctgta 660
tttgatgaag ggaaaagttt gcactcagaa acaaagaact ccttgtatgc ggtataggat 720
gtgcacatcg ctccggcctg gcctaaaatg cacacagtca atggttatgt aaacaggct 780
ctgccaggcgc tgattggatg ccacaggaaa tcagtttattt ggcatgtat tggaatggc 840
accactcctg aagtgcactc aatattccctc gaaggtcaca catttttgtt gaggaaccat 900
cgccaggcgt ctttggaaat ctgcacataa actttccctt ctgcacaaac actcttgcatt 960
gaccttggac agtttctact gttttgtcat atctcttccc accaacatga tggcatggaa 1020
gcttatgtca aagttagacag ctgtccagag gaaccccaac tacgaatgaa aaataatgaa 1080
gaagcggaaag actatgtatg tgatcttact gattctgaaa tggatgttgtt caggtttgc 1140
gatgacaact ctccctccct tatccaaatt cgctcagttt ccaagaagca tcctaaaact 1200
tgggtacatt acattgtgc tgaagaggag gactggact atgcctccctt agtctcgcc 1260
cccgatgaca gaagttataa aagtcaatat ttgaacaatg gcccctcagcg gattggtagg 1320
aagtacaaaa aagtccgatt tatggcatac acagatgaaa ctttaagac tcgtgaagct 1380
attcagcatg aatcaggaat ctggggacct ttactttatg gggaaagttgg agacacactg 1440
ttgattat ttaagaatca agcaagcaga ccatataaca totaccotca cggaaatcact 1500
gatgtccgtc ctttgtatc aaggagatta cccaaagggtg taaaacatgtt gaaggattt 1560
ccaaattcgc caggagaaat attcaatat aaatggacag tgactgtaga agatggccca 1620
actaaatcag atcctcggtt cctgacccgc tattactcta gtttcgtttaa tatggagaga 1680
gatctagctt caggactcat tggccctetc ctcatctgct acaaagaatc tgttagatcaa 1740

-continued

agaggaaacc agataatgtc agacaagagg aatgtcatcc tggggatgtt attttatgtg	1800
aaccgaagct ggtacccac agagaatata caacgcatttc tccccaatcc agctggagt	1860
cagcttgagg atccagatgtt ccaaggctcc aacatcatgc acagcatcaa tggctatgtt	1920
tttgatagtt tgcagttgtc agtttgggg catgaggtgg catactggta cattctaagc	1980
atggagcac agactgactt ccttctgtc ttcttctctg gatatacctt caaacacaaa	2040
atggctatg aagacacact caccctattt ccattcttag gagaaactgt cttcatgtcg	2100
atggaaaacc caggctatg gattctgggg tgccacaact cagactttcg gaacagaggc	2160
atgaccgcct tactgaaggt ttcttagttgt gacaagaaca ctggtgatta ttacgaggac	2220
agtttatgaag atatccatgc atatctgtc agaaaaaaca atgccattga accaagaagc	2280
ttctcccaga atccaccagt cttgaaacgc catcaacggg aaataactcg tactactctt	2340
cagtcagatc aagaggaaat tgactatgtat gataccatat cagttgaaat gaagaaggaa	2400
gattttgaca ttatgtatga ggatgaaaat cagagcccccc gcaagcttca aaagaaaaca	2460
cgacactatt ttattgtgc agtggagagg ctctggattt atggatgag tagctccca	2520
catgttctaa gaaacagggc tcagagtgcc agtgtccctc agttcaagaa agttgtttc	2580
caggaattta ctgatggctc cttaactcg cccttataacc gtggagaact aaatgaacat	2640
ttgggactcc tggggccata tataagagca gaagttgaag ataataatcat ggtacttcc	2700
agaaatcagg cctctcgcc ctattccctc tattctagcc ttatttctta tgaggaagat	2760
cagaggcaag gagcagaacc tagaaaaaac tttgtcaagc ctaatgaaac caaaacttac	2820
ttttggaaag tgcaacatca tatggcaccc actaaagatg agtttgactg caaaggctgg	2880
gttttatttct ctgatgttga cctggaaaaa gatgtgcact caggctgtat tggaccctt	2940
ctggctcgcc acactaacac actgaaccct gctcatggg gacaagtgc agtacaggaa	3000
tttgctctgt ttttaccat ctttgtatgatg accaaaagct ggtacttcac tgaaaatatg	3060
gaaagaaaact gcagggctcc ctgcaatatac cagatggaa atcccacttt taaagagaat	3120
tatcgcttcc atgcaatcaa tggctacata atggatacac tacctggctt agtaatggct	3180
caggatcaa ggattcgatg gatatgtcc agcatggca gcaatggaaa catccatct	3240
attcatttca gtggacatgt gttcaactgtc cgaaaaaaag aggagtataa aatggactg	3300
tacaatctct atccaggtgt ttttgagaca gtggaaatgt taccatccaa agctggaaatt	3360
tggcggttgg aatgccttat tggcgagcat ctacatgtc ggtggcac acttttctg	3420
gtgtacagca ataagtgtca gactccctg ggaatggctt ctggacacat tagagattt	3480
cagattacag cttcaggaca atatggacag tggggcccaa agctggccag acttcattat	3540
tccggatcaa tcaatgcctg gagcaccaag gagccctttt ctggatccaa ggtggatctg	3600
ttggcaccaa tgattattca cggcatcaag acccagggtg cccgtcagaa gttctccagc	3660
ctctacatct ctcagtttat catcatgtat agtcttgatg ggaagaagtgc gcaacttat	3720
cgaggaaattt ccactggaaac cttaatggtc ttctttggca atgtggatcc atctggata	3780
aaacacaata ttttaaccc tccaattatt gctcgataca tccgtttgca cccaaactcat	3840
tatagcattc gcagcactct tcgcatggag ttgtatggct gtgatataa tagttgcagc	3900
atgccattgg gaatggagag taaagcaata tcagatgcac agattactgc ttcatccat	3960
tttaccaata tggggccac ctggctctt taaaagctc gacttcacctt ccaagggagg	4020

-continued

agtaatgcct ggagacctca ggtgaataat cccaaagagt ggctgcaagt ggacttcag	4080
aagacaatga aagtccacagg agtaactact cagggagtaa aatctctgt taccagcatg	4140
tatgtgaagg agttcctcat ctccagcagt caagatggcc atcagtggac tctttttt	4200
cagaatggca aagtaaagggt ttttcaggga aatcaagact cttcacacc tgtggtaac	4260
tctcttagacc caccgttaact gactcgctac ctgcgaattc acccccagag ttgggtgcac	4320
cagattgccc tgaggatgga gtttctggc tgcgaggcac aggacctcta ctga	4374

```

<210> SEQ ID NO 18
<211> LENGTH: 4374
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
polynucleotide

```

<400> SEQUENCE: 18	
atgcagatcg agctgtccac atgctttttt ctgtgcctgc tgccgttctg cttcagcgcc	60
acccggcggt actacctggg cgccgtggag ctgtcctggg actacatgca gagcgacctg	120
ggcgagctgc ccgtggacgc ccggttcccc cccagagtgc ccaagagctt ccccttcaac	180
accagcgtgg tgtacaagaa aaccctgttc gtggagttca cccgaccacct gttcaacatc	240
gccaagccca ggcccccctg gatgggcctg ctggggccca ccatccaggc cgaggtgtac	300
gacaccgtgg tgcgttccct gaagaacatg gccagccacc ccgtgagcct gcacgcccgt	360
ggcgtgagct actggaaaggc ctccgagggc gccgagttacg acgaccagac cagccagcgg	420
gagaaaaggagg acgacaaagt ctttcctggc ggcagccaca cctacgtgtg gcaggtctg	480
aaagaaaaacg gccccatggc ctccgaccct ctgtgcctga cctacagctt cctgagccac	540
gtggacctgg tgaaggacct gaacagcggg ctgattgggg ccctgttgtt ctggggag	600
ggcagectgg ccaaagagaa aaccctggacc ctgcacaagt tcacgttgtt gttcccggt	660
ttcgacgagg gcaagagctg gcacagcgtg accaagaaca gctgtatgca ggaccggac	720
gcccgccttg ccacggcttg gcccacatg cacaccgtga acggctacgt gaacagaagc	780
ctggccggcc tgattggctg ccaccggaaag agcgtgtact ggcacgtgat cggcatggc	840
accacacccg aggtgcacag catcttctg gaagggcaca ctttcttgtt gcgaaaccac	900
cggcaggcca gcctggaaat cggcccttac accttctgtg ccccccagac actgtgtatg	960
gacctgggcc agttcctgtt gtttgcac atcagcttc accagcacga cggcatggaa	1020
gcttacgtga aggtggactc ctggccggag gaaccccgac tgccgtatgaa gaacaacgag	1080
gaagccgagg actacgacga cgacactgacc gacagcgaga tggacgttgtt gcggttcac	1140
gacgacaaca gccccagctt catccagatc agaagcgtgg ccaagaagca ccccaagacc	1200
tgggtgact acatgcgcgc cgaggaagag gactggact acggccctt ggtgtggcc	1260
cccgacgaca gaagctacaa gagccagttac ctgaacaatg gccccagcg gatcgccgg	1320
aagtacaaga aagtgcgggtt catggcctac accgacgaga cttcaagac cggggaggcc	1380
atccagcacg agagcggcat cttggggccc ctgtgtacg gcaagtgccc cgacacactg	1440
ctgatcatct tcaagaacca ggccagccgg ccctacaaca totacccca cggcatcacc	1500
gacgtgcggc ccctgtacag caggccggctg cccaaaggcg tgaagcacct gaaggactc	1560
cccatctgc cggcggagat cttcaagttac aagtggaccg tgaccgtgga ggacggcccc	1620

-continued

accaagagcg accccagatg cctgacccgg tactacagca gtttcgtgaa catggAACGG	1680
gacctggcct cggggctgtat cgacccctcg ctgtatctgc acaaAGAAAG cgtggaccAG	1740
cggggcaACC agatcatgag cgacaAGCGG aacgtgtcc tggTCAGCGT gttcgatGAG	1800
aaccggcctt ggtatctgac cgagaACATC cagcggttcc tgcccAACCC tgccgggGTG	1860
cagctggaAG atcccAGTT ccaggCCAGC aacatcatGC actccatCAA tggctacGTG	1920
ttcgacAGCC tgcAGCTGTC cgtgtgtctg cacgaggTGG cctactGGTA catcctGAGC	1980
atcgccGCC AGACCGACTT cctgagcGTG ttcttcAGCG gctacACCTT caAGCACAAG	2040
atgggtGTACG aggacACCTT gaccCTGTc ccttCAGCG gCGAGACCGT gttcatGAGC	2100
atggAAAACC CGGCGCTGTG gatcctGGGC tgccacaACA gCGACTTCG gaaccGGGGC	2160
atgaccGCC CGTGAAGGT gtccAGCTGC gacaAGAACA CGGCGACTA CTACGAGGAC	2220
agctacGAGG atatcAGCGC ctacCTGTG tccaAGAACA acggccATCGA gcccAGAAGC	2280
ttcagecAGA ACCCCCCCTGT gctGAAGCGG caccAGAGAG agatCACCCG gaccACCTG	2340
cagtccGACC AGGAAGAGAT cgattacGAC gacACCATCA gCGTGGAGAT gaaaaAAAGAA	2400
gatttcGACA tctacGACGA ggacGAGAAC cAGAGCCCCC ggtcCTTCCA gaAGAAAACC	2460
cggcactACT ttatcGCCGC cgtggAGCGG ctgtgggACT acggcatGAG cAGCAGCCCC	2520
cacgtGCTGC ggaACCGGGC CCAGAGCGGC agcgtGCCCC agttcaAGAA agtgggttC	2580
caggaATTCA CGACAGGCAG ctTCACCCAG CCCCTGTACC ggggcGAGCT gaACGAGCAC	2640
ctggggCTGC tggggcccta catcAGGGCC gaagtggagg acaACATCAT ggtgacCTTC	2700
cggaatCAGG CCAGCAGACC ctactCCTTC tacAGCAGCC tgatcAGCTA cgaAGAGGAC	2760
cagcggCAGG gCGCTGAACC CGGGAAGAAC ttcgtGAAGC ccaatGAGAC caAGACCTAC	2820
ttctggAAAG tgcAGCACCA catggCCCCC accaAGGACG agttcAGCTG caaggCCTGG	2880
gcctacttCA GCGACGTGGA tctggAAAAG gacgtGACT ctggACTGAT tggccCTCTG	2940
ctgggtGTC ACACCAACAC CCTGAACCCC GCCCACGGCC ggcAGGTGAC cgtcAGGAA	3000
ttcgcctGT tcttaccat ctTCGACGAG accaAGTCTC ggtacttCAC cgagaATATG	3060
gaacggAACT GCAAGAGCCC CTGCAACATC cAGATGGAA atcctACCTT caaAGAGAAC	3120
taccggTTCC acgCCATCAA CGGCTACATC atggACACCC tgcctGGCCT ggtgatGGCC	3180
caggaccAGA ggatccGGTG gtatctGTC tccatGGGCA gcaACGAGAA tatccACAGC	3240
atccacttCA GCGGCCACGT gttcACCGTG aggaAGAAAG aAGAGTACAA gatggccCTG	3300
tacaACCTGT ACCCCGGCGT gtTCGAGACG gttggAGATGC tgcccAGCAA gGCCGGCATC	3360
tggcGGGTGG agtgtctGAT CGGCGAGCAC ctgcATGCCG ggtAGAGCAC cctgtttCTG	3420
gtgtacAGCA ACAAGTGCA GACCCCCCTG GGCAATGGCA gggGCCACAT CGGGACCTC	3480
cagatCACCG CTCGGGCCA GTACGGCCAG TGGGCCCCA AGCTGGCCG gctGACTAC	3540
agcggcGCA TCAACGCTG GTCCACAAA GAGCCCTCA GCTGGATCAA GGTGGACCTG	3600
ctggccccCTA TGATCATCCA CGGCATTAAG ACCCAGGGCG CCAGGCAGAA GTTCAcGAGC	3660
ctgtacATCA GCCAGTTCAt CATCATGTAC AGCCTGGACG GCAAGAAGTG GCAGACCTAC	3720
cggggcaACA GCAACGGCAC CCTGATGGTG ttcttcGGCA acgtggACAG cAGCAGCATC	3780
aagcacaACA tcttcaACCC CCCCATCATC GCCCAGGTACA tccggCTGCA ccccACCCAC	3840
tacagcatCA GATCCACCCt CGGGATGGAA CTGATGGGCT GCGACCTGAA CTCCtGAGC	3900

-continued

```

atgcctctgg gcatggaaag caaggccatc agcgacgccc agatcacagc cagcagctac 3960
ttcacaaca tttcgccac ctggcccccc tccaaggcca ggctgcacct gcagggccgg 4020
tccaacgcct gcggcctca ggtgaacaac cccaaagaat ggctgcaggt ggacttcag 4080
aaaaccatga aggtgaccgg cgtgaccacc cagggcgtga aaagcctgct gaccagcatg 4140
tacgtgaaag agtttctgat cagcagcagc caggacggcc accagtggac cctgttctt 4200
cagaacggca aggtgaaagt gttccagggc aaccaggact ctttcacccc cgtggtaac 4260
tccctggacc ccccccgtct gacccgatcc ctgcggatcc accccagtc ttgggtgcac 4320
cagatcgccc tgaggatgga agtgctgggta tgtgaggccc aggatctgta ctga 4374

```

<210> SEQ_ID NO 19

<211> LENGTH: 2351

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 19

```

Met Gln Ile Glu Leu Ser Thr Cys Phe Leu Cys Leu Leu Arg Phe
1 5 10 15

```

```

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

```

```

Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg
35 40 45

```

```

Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val
50 55 60

```

```

Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile
65 70 75 80

```

```

Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln
85 90 95

```

```

Ala Glu Val Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser
100 105 110

```

```

His Pro Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser
115 120 125

```

```

Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp
130 135 140

```

```

Asp Lys Val Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu
145 150 155 160

```

```

Lys Glu Asn Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser
165 170 175

```

```

Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile
180 185 190

```

```

Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr
195 200 205

```

```

Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly
210 215 220

```

```

Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp
225 230 235 240

```

```

Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr
245 250 255

```

```

Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val
260 265 270

```

-continued

Tyr	Trp	His	Val	Ile	Gly	Met	Gly	Thr	Thr	Pro	Glu	Val	His	Ser	Ile
275						280					285				
Phe	Leu	Glu	Gly	His	Thr	Phe	Leu	Val	Arg	Asn	His	Arg	Gln	Ala	Ser
290						295					300				
Leu	Glu	Ile	Ser	Pro	Ile	Thr	Phe	Leu	Thr	Ala	Gln	Thr	Leu	Leu	Met
305						310					315				320
Asp	Leu	Gly	Gln	Phe	Leu	Leu	Phe	Cys	His	Ile	Ser	Ser	His	Gln	His
						325					330				335
Asp	Gly	Met	Glu	Ala	Tyr	Val	Lys	Val	Asp	Ser	Cys	Pro	Glu	Glu	Pro
						340					345				350
Gln	Leu	Arg	Met	Lys	Asn	Asn	Glu	Glu	Ala	Glu	Asp	Tyr	Asp	Asp	Asp
						355					360				365
Leu	Thr	Asp	Ser	Glu	Met	Asp	Val	Val	Arg	Phe	Asp	Asp	Asp	Asn	Ser
						370					375				380
Pro	Ser	Phe	Ile	Gln	Ile	Arg	Ser	Val	Ala	Lys	Lys	His	Pro	Lys	Thr
385						390					395				400
Trp	Val	His	Tyr	Ile	Ala	Ala	Glu	Glu	Asp	Trp	Asp	Tyr	Ala	Pro	
						405					410				415
Leu	Val	Leu	Ala	Pro	Asp	Asp	Arg	Ser	Tyr	Lys	Ser	Gln	Tyr	Leu	Asn
						420					425				430
Asn	Gly	Pro	Gln	Arg	Ile	Gly	Arg	Lys	Tyr	Lys	Lys	Val	Arg	Phe	Met
						435					440				445
Ala	Tyr	Thr	Asp	Glu	Thr	Phe	Lys	Thr	Arg	Glu	Ala	Ile	Gln	His	Glu
						450					455				460
Ser	Gly	Ile	Leu	Gly	Pro	Leu	Leu	Tyr	Gly	Glu	Val	Gly	Asp	Thr	Leu
465						470					475				480
Leu	Ile	Ile	Phe	Lys	Asn	Gln	Ala	Ser	Arg	Pro	Tyr	Asn	Ile	Tyr	Pro
						485					490				495
His	Gly	Ile	Thr	Asp	Val	Arg	Pro	Leu	Tyr	Ser	Arg	Arg	Leu	Pro	Lys
						500					505				510
Gly	Val	Lys	His	Leu	Lys	Asp	Phe	Pro	Ile	Leu	Pro	Gly	Glu	Ile	Phe
						515					520				525
Lys	Tyr	Lys	Trp	Thr	Val	Thr	Val	Glu	Asp	Gly	Pro	Thr	Lys	Ser	Asp
						530					535				540
Pro	Arg	Cys	Leu	Thr	Arg	Tyr	Tyr	Ser	Ser	Phe	Val	Asn	Met	Glu	Arg
545						550					555				560
Asp	Leu	Ala	Ser	Gly	Leu	Ile	Gly	Pro	Leu	Leu	Ile	Cys	Tyr	Lys	Glu
						565					570				575
Ser	Val	Asp	Gln	Arg	Gly	Asn	Gln	Ile	Met	Ser	Asp	Lys	Arg	Asn	Val
						580					585				590
Ile	Leu	Phe	Ser	Val	Phe	Asp	Glu	Asn	Arg	Ser	Trp	Tyr	Leu	Thr	Glu
						595					600				605
Asn	Ile	Gln	Arg	Phe	Leu	Pro	Asn	Pro	Ala	Gly	Val	Gln	Leu	Glu	Asp
						610					615				620
Pro	Glu	Phe	Gln	Ala	Ser	Asn	Ile	Met	His	Ser	Ile	Asn	Gly	Tyr	Val
625						630					635				640
Phe	Asp	Ser	Leu	Gln	Leu	Ser	Val	Cys	Leu	His	Glu	Val	Ala	Tyr	Trp
						645					650				655
Tyr	Ile	Leu	Ser	Ile	Gly	Ala	Gln	Thr	Asp	Phe	Leu	Ser	Val	Phe	Phe
						660					665				670
Ser	Gly	Tyr	Thr	Phe	Lys	His	Lys	Met	Val	Tyr	Glu	Asp	Thr	Leu	Thr

-continued

675	680	685
Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro		
690	695	700
Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly		
705	710	715
Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp		
725	730	735
Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys		
740	745	750
Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser Gln Asn Ser Arg His Pro		
755	760	765
Ser Thr Arg Gln Lys Gln Phe Asn Ala Thr Thr Ile Pro Glu Asn Asp		
770	775	780
Ile Glu Lys Thr Asp Pro Trp Phe Ala His Arg Thr Pro Met Pro Lys		
785	790	795
800		
Ile Gln Asn Val Ser Ser Ser Asp Leu Leu Met Leu Leu Arg Gln Ser		
805	810	815
Pro Thr Pro His Gly Leu Ser Leu Ser Asp Leu Gln Glu Ala Lys Tyr		
820	825	830
Glu Thr Phe Ser Asp Asp Pro Ser Pro Gly Ala Ile Asp Ser Asn Asn		
835	840	845
Ser Leu Ser Glu Met Thr His Phe Arg Pro Gln Leu His His Ser Gly		
850	855	860
Asp Met Val Phe Thr Pro Glu Ser Gly Leu Gln Leu Arg Leu Asn Glu		
865	870	875
880		
Lys Leu Gly Thr Thr Ala Ala Thr Glu Leu Lys Lys Leu Asp Phe Lys		
885	890	895
Val Ser Ser Thr Ser Asn Asn Leu Ile Ser Thr Ile Pro Ser Asp Asn		
900	905	910
Leu Ala Ala Gly Thr Asp Asn Thr Ser Ser Leu Gly Pro Pro Ser Met		
915	920	925
Pro Val His Tyr Asp Ser Gln Leu Asp Thr Thr Leu Phe Gly Lys Lys		
930	935	940
Ser Ser Pro Leu Thr Glu Ser Gly Gly Pro Leu Ser Leu Ser Glu Glu		
945	950	955
960		
Asn Asn Asp Ser Lys Leu Leu Glu Ser Gly Leu Met Asn Ser Gln Glu		
965	970	975
Ser Ser Trp Gly Lys Asn Val Ser Ser Thr Glu Ser Gly Arg Leu Phe		
980	985	990
Lys Gly Lys Arg Ala His Gly Pro Ala Leu Leu Thr Lys Asp Asn Ala		
995	1000	1005
Leu Phe Lys Val Ser Ile Ser Leu Leu Lys Thr Asn Lys Thr Ser		
1010	1015	1020
Asn Asn Ser Ala Thr Asn Arg Lys Thr His Ile Asp Gly Pro Ser		
1025	1030	1035
Leu Leu Ile Glu Asn Ser Pro Ser Val Trp Gln Asn Ile Leu Glu		
1040	1045	1050
Ser Asp Thr Glu Phe Lys Lys Val Thr Pro Leu Ile His Asp Arg		
1055	1060	1065
Met Leu Met Asp Lys Asn Ala Thr Ala Leu Arg Leu Asn His Met		
1070	1075	1080

-continued

Ser Asn Lys Thr Thr Ser Ser Lys Asn Met Glu Met Val Gln Gln
 1085 1090 1095
 Lys Lys Glu Gly Pro Ile Pro Pro Asp Ala Gln Asn Pro Asp Met
 1100 1105 1110
 Ser Phe Phe Lys Met Leu Phe Leu Pro Glu Ser Ala Arg Trp Ile
 1115 1120 1125
 Gln Arg Thr His Gly Lys Asn Ser Leu Asn Ser Gly Gln Gly Pro
 1130 1135 1140
 Ser Pro Lys Gln Leu Val Ser Leu Gly Pro Glu Lys Ser Val Glu
 1145 1150 1155
 Gly Gln Asn Phe Leu Ser Glu Lys Asn Lys Val Val Val Gly Lys
 1160 1165 1170
 Gly Glu Phe Thr Lys Asp Val Gly Leu Lys Glu Met Val Phe Pro
 1175 1180 1185
 Ser Ser Arg Asn Leu Phe Leu Thr Asn Leu Asp Asn Leu His Glu
 1190 1195 1200
 Asn Asn Thr His Asn Gln Glu Lys Lys Ile Gln Glu Glu Ile Glu
 1205 1210 1215
 Lys Lys Glu Thr Leu Ile Gln Glu Asn Val Val Leu Pro Gln Ile
 1220 1225 1230
 His Thr Val Thr Gly Thr Lys Asn Phe Met Lys Asn Leu Phe Leu
 1235 1240 1245
 Leu Ser Thr Arg Gln Asn Val Glu Gly Ser Tyr Asp Gly Ala Tyr
 1250 1255 1260
 Ala Pro Val Leu Gln Asp Phe Arg Ser Leu Asn Asp Ser Thr Asn
 1265 1270 1275
 Arg Thr Lys Lys His Thr Ala His Phe Ser Lys Lys Gly Glu Glu
 1280 1285 1290
 Glu Asn Leu Glu Gly Leu Gly Asn Gln Thr Lys Gln Ile Val Glu
 1295 1300 1305
 Lys Tyr Ala Cys Thr Thr Arg Ile Ser Pro Asn Thr Ser Gln Gln
 1310 1315 1320
 Asn Phe Val Thr Gln Arg Ser Lys Arg Ala Leu Lys Gln Phe Arg
 1325 1330 1335
 Leu Pro Leu Glu Glu Thr Glu Leu Glu Lys Arg Ile Ile Val Asp
 1340 1345 1350
 Asp Thr Ser Thr Gln Trp Ser Lys Asn Met Lys His Leu Thr Pro
 1355 1360 1365
 Ser Thr Leu Thr Gln Ile Asp Tyr Asn Glu Lys Glu Lys Gly Ala
 1370 1375 1380
 Ile Thr Gln Ser Pro Leu Ser Asp Cys Leu Thr Arg Ser His Ser
 1385 1390 1395
 Ile Pro Gln Ala Asn Arg Ser Pro Leu Pro Ile Ala Lys Val Ser
 1400 1405 1410
 Ser Phe Pro Ser Ile Arg Pro Ile Tyr Leu Thr Arg Val Leu Phe
 1415 1420 1425
 Gln Asp Asn Ser Ser His Leu Pro Ala Ala Ser Tyr Arg Lys Lys
 1430 1435 1440
 Asp Ser Gly Val Gln Glu Ser Ser His Phe Leu Gln Gly Ala Lys
 1445 1450 1455

-continued

Lys	Asn	Asn	Leu	Ser	Leu	Ala	Ile	Leu	Thr	Leu	Glu	Met	Thr	Gly
1460														
							1465					1470		
Asp	Gln	Arg	Glu	Val	Gly	Ser	Leu	Gly	Thr	Ser	Ala	Thr	Asn	Ser
1475							1480					1485		
Val	Thr	Tyr	Lys	Lys	Val	Glu	Asn	Thr	Val	Leu	Pro	Lys	Pro	Asp
1490							1495					1500		
Leu	Pro	Lys	Thr	Ser	Gly	Lys	Val	Glu	Leu	Leu	Pro	Lys	Val	His
1505							1510					1515		
Ile	Tyr	Gln	Lys	Asp	Leu	Phe	Pro	Thr	Glu	Thr	Ser	Asn	Gly	Ser
1520							1525					1530		
Pro	Gly	His	Leu	Asp	Leu	Val	Glu	Gly	Ser	Leu	Leu	Gln	Gly	Thr
1535							1540					1545		
Glu	Gly	Ala	Ile	Lys	Trp	Asn	Glu	Ala	Asn	Arg	Pro	Gly	Lys	Val
1550							1555					1560		
Pro	Phe	Leu	Arg	Val	Ala	Thr	Glu	Ser	Ser	Ala	Lys	Thr	Pro	Ser
1565							1570					1575		
Lys	Leu	Leu	Asp	Pro	Leu	Ala	Trp	Asp	Asn	His	Tyr	Gly	Thr	Gln
1580							1585					1590		
Ile	Pro	Lys	Glu	Glu	Trp	Lys	Ser	Gln	Glu	Lys	Ser	Pro	Glu	Lys
1595							1600					1605		
Thr	Ala	Phe	Lys	Lys	Asp	Thr	Ile	Leu	Ser	Leu	Asn	Ala	Cys	
1610							1615					1620		
Glu	Ser	Asn	His	Ala	Ile	Ala	Ala	Ile	Asn	Glu	Gly	Gln	Asn	Lys
1625							1630					1635		
Pro	Glu	Ile	Glu	Val	Thr	Trp	Ala	Lys	Gln	Gly	Arg	Thr	Glu	Arg
1640							1645					1650		
Leu	Cys	Ser	Gln	Asn	Pro	Pro	Val	Leu	Lys	Arg	His	Gln	Arg	Glu
1655							1660					1665		
Ile	Thr	Arg	Thr	Thr	Leu	Gln	Ser	Asp	Gln	Glu	Ile	Asp	Tyr	
1670							1675					1680		
Asp	Asp	Thr	Ile	Ser	Val	Glu	Met	Lys	Glu	Asp	Phe	Asp	Ile	
1685							1690					1695		
Tyr	Asp	Glu	Asp	Glu	Asn	Gln	Ser	Pro	Arg	Ser	Phe	Gln	Lys	Lys
1700							1705					1710		
Thr	Arg	His	Tyr	Phe	Ile	Ala	Ala	Val	Glu	Arg	Leu	Trp	Asp	Tyr
1715							1720					1725		
Gly	Met	Ser	Ser	Ser	Pro	His	Val	Leu	Arg	Asn	Arg	Ala	Gln	Ser
1730							1735					1740		
Gly	Ser	Val	Pro	Gln	Phe	Lys	Lys	Val	Val	Phe	Gln	Glu	Phe	Thr
1745							1750					1755		
Asp	Gly	Ser	Phe	Thr	Gln	Pro	Leu	Tyr	Arg	Gly	Glu	Leu	Asn	Glu
1760							1765					1770		
His	Leu	Gly	Leu	Leu	Gly	Pro	Tyr	Ile	Arg	Ala	Glu	Val	Glu	Asp
1775							1780					1785		
Asn	Ile	Met	Val	Thr	Phe	Arg	Asn	Gln	Ala	Ser	Arg	Pro	Tyr	Ser
1790							1795					1800		
Phe	Tyr	Ser	Ser	Leu	Ile	Ser	Tyr	Glu	Glu	Asp	Gln	Arg	Gln	Gly
1805							1810					1815		
Ala	Glu	Pro	Arg	Lys	Asn	Phe	Val	Lys	Pro	Asn	Glu	Thr	Lys	Thr
1820							1825					1830		
Tyr	Phe	Trp	Lys	Val	Gln	His	His	Met	Ala	Pro	Thr	Lys	Asp	Glu

-continued

1835	1840	1845
Phe Asp Cys Lys Ala Trp Ala	Tyr Phe Ser Asp Val	Asp Leu Glu
1850	1855	1860
Lys Asp Val His Ser Gly Leu	Ile Gly Pro Leu Leu	Val Cys His
1865	1870	1875
Thr Asn Thr Leu Asn Pro Ala	His Gly Arg Gln Val	Thr Val Gln
1880	1885	1890
Glu Phe Ala Leu Phe Phe	Thr Ile Phe Asp Glu	Thr Lys Ser Trp
1895	1900	1905
Tyr Phe Thr Glu Asn Met Glu	Arg Asn Cys Arg Ala	Pro Cys Asn
1910	1915	1920
Ile Gln Met Glu Asp Pro Thr	Phe Lys Glu Asn Tyr	Arg Phe His
1925	1930	1935
Ala Ile Asn Gly Tyr Ile Met	Asp Thr Leu Pro Gly	Leu Val Met
1940	1945	1950
Ala Gln Asp Gln Arg Ile Arg	Trp Tyr Leu Leu Ser	Met Gly Ser
1955	1960	1965
Asn Glu Asn Ile His Ser Ile	His Phe Ser Gly His	Val Phe Thr
1970	1975	1980
Val Arg Lys Lys Glu Glu Tyr	Lys Met Ala Leu Tyr	Asn Leu Tyr
1985	1990	1995
Pro Gly Val Phe Glu Thr Val	Glu Met Leu Pro Ser	Lys Ala Gly
2000	2005	2010
Ile Trp Arg Val Glu Cys Leu	Ile Gly Glu His Leu	His Ala Gly
2015	2020	2025
Met Ser Thr Leu Phe Leu Val	Tyr Ser Asn Lys Cys	Gln Thr Pro
2030	2035	2040
Leu Gly Met Ala Ser Gly His	Ile Arg Asp Phe Gln	Ile Thr Ala
2045	2050	2055
Ser Gly Gln Tyr Gly Gln Trp	Ala Pro Lys Leu Ala	Arg Leu His
2060	2065	2070
Tyr Ser Gly Ser Ile Asn Ala	Trp Ser Thr Lys Glu	Pro Phe Ser
2075	2080	2085
Trp Ile Lys Val Asp Leu Leu	Ala Pro Met Ile Ile	His Gly Ile
2090	2095	2100
Lys Thr Gln Gly Ala Arg Gln	Lys Phe Ser Ser Leu	Tyr Ile Ser
2105	2110	2115
Gln Phe Ile Ile Met Tyr Ser	Leu Asp Gly Lys Lys	Trp Gln Thr
2120	2125	2130
Tyr Arg Gly Asn Ser Thr Gly	Thr Leu Met Val Phe	Phe Gly Asn
2135	2140	2145
Val Asp Ser Ser Gly Ile Lys	His Asn Ile Phe Asn	Pro Pro Ile
2150	2155	2160
Ile Ala Arg Tyr Ile Arg Leu	His Pro Thr His Tyr	Ser Ile Arg
2165	2170	2175
Ser Thr Leu Arg Met Glu Leu	Met Gly Cys Asp Leu	Asn Ser Cys
2180	2185	2190
Ser Met Pro Leu Gly Met Glu	Ser Lys Ala Ile Ser	Asp Ala Gln
2195	2200	2205
Ile Thr Ala Ser Ser Tyr Phe	Thr Asn Met Phe Ala	Thr Trp Ser
2210	2215	2220

-continued

Pro	Ser	Lys	Ala	Arg	Leu	His	Leu	Gln	Gly	Arg	Ser	Asn	Ala	Trp
2225					2230					2235				
Arg	Pro	Gln	Val	Asn	Asn	Pro	Lys	Glu	Trp	Leu	Gln	Val	Asp	Phe
2240						2245				2250				
Gln	Lys	Thr	Met	Lys	Val	Thr	Gly	Val	Thr	Gln	Gly	Val	Lys	
2255					2260				2265					
Ser	Leu	Leu	Thr	Ser	Met	Tyr	Val	Lys	Glu	Phe	Ile	Ser	Ser	
2270						2275				2280				
Ser	Gln	Asp	Gly	His	Gln	Trp	Thr	Leu	Phe	Phe	Gln	Asn	Gly	Lys
2285						2290				2295				
Val	Lys	Val	Phe	Gln	Gly	Asn	Gln	Asp	Ser	Phe	Thr	Pro	Val	Val
2300						2305				2310				
Asn	Ser	Leu	Asp	Pro	Pro	Leu	Leu	Thr	Arg	Tyr	Leu	Arg	Ile	His
2315						2320				2325				
Pro	Gln	Ser	Trp	Val	His	Gln	Ile	Ala	Leu	Arg	Met	Glu	Val	Leu
2330						2335				2340				
Gly	Cys	Glu	Ala	Gln	Asp	Leu	Tyr							
	2345					2350								

```

<210> SEQ ID NO 20
<211> LENGTH: 4374
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
polynucleotide

```

<400> SEQUENCE: 20

atgcagattg	agctgagcac	ctgtttcttc	ctgtgcctgc	tgaggttctg	cttctctgcc	60
accaggagat	actacctggg	cgcgcgtggag	ctgagctgggg	actacatgca	gtctgacctg	120
ggcgagactgc	ctgtggacgc	caggttcccc	cccaagagtgc	ccaagagctt	ccccttcaac	180
acctcagtgg	tgtacaagaa	gaccctgttc	gtggagttca	ccgaccacct	gttcaacatc	240
gecaagccca	ggcccccctg	gatgggcctg	ctggggccca	ccatccaggc	cgaggtgtac	300
gacaccgtgg	tgatcaccct	gaagaacatg	gccagccacc	ccgtgagcct	gcacgcccgt	360
ggcgtgagct	actggaaaggc	ctctgagggc	gccgagtatg	acgaccagac	cagccagagg	420
gagaaggagg	acgacaaggt	gttcccccgc	ggcagccaca	cctacgtgtg	gcaggtgtc	480
aaggagaacg	gccccatggc	cagcgacccc	ctgtgcctga	cctacagcta	cctgagccac	540
gtggacctgg	tgaaggacct	gaactctggc	ctgatcggcg	ccctgtgtt	gtgcagggag	600
ggcagectgg	ccaaggagaa	gaccaggacc	ctgcacaagt	tcatcctgt	gttcgcctgt	660
ttcgatgagg	gcaagagctg	gcacagcgcag	accaagaaca	gcctgtatgca	ggacaggaggat	720
cccgccctcg	ccaggccctg	gcccaagatg	cacaccgtga	acggctacgt	gaacaggagc	780
ctgccccggcc	tgatcggctg	ccacaggaag	tctgtgtact	ggcacgtgt	cgccatgggc	840
accacccccc	aggtgcacag	catttcctg	gagggccaca	ccttccttgt	gaggaaccac	900
aggcaggcca	gcctggagat	cagccccatc	actttcctga	ccgcccagac	cctgtgtatg	960
gacctggggcc	agtctctgtct	gttctgccac	atcagcgcac	accagcacga	cgccatggag	1020
gcctacgtga	aggtggacag	ctgccccgag	gagccccagc	tgaggatgaa	gaacaacgag	1080
gaggccgagg	actatgatga	tgacctgacc	gactctgaga	tggacgttgt	gaggtttgat	1140

-continued

gatgacaaca	gccccagctt	catccagatc	aggctctgtgg	ccaagaagca	ccccaagacc	1200
tgggtgact	acategccgc	cgaggaggag	gactgggact	acggccccc	gtgtgtggcc	1260
cccgacaca	ggagctacaa	gagccagtagc	ctgaacaacg	gccccagag	gatccggcagg	1320
aagtacaaga	aggtcagatt	catggcctac	accgacgaga	ccttcaagac	cagggaggcc	1380
atccagcacg	agtctggcat	cctggggccc	ctgctgtacg	gcgagggtgg	cgacacccctg	1440
ctgtatct	tcaagaacca	ggccagcagg	ccctacaaca	tctacccca	cgccatcacc	1500
gatgtgaggc	ccctgtacag	caggaggctg	cccaaggcg	tgaagcacct	gaaggactc	1560
cccatctgc	ccggcagat	cttcaagtac	aagtggaccg	tgaccgtgga	ggatggcccc	1620
accaagtctg	accccaggtg	cctgaccagg	tactacagca	gcttcgtgaa	catggagagg	1680
gacctggcct	ctggcctgat	cggccccc	ctgatctgct	acaaggagag	cgtggaccag	1740
aggggcaacc	agatcatgtc	tgacaagagg	aacgtgatcc	ttttctctgt	tttcgtatgag	1800
aacaggagct	ggtatctgac	cgagaacatc	cagaggttc	tgcccaaccc	cgccggcgtg	1860
cagctggagg	accccgagtt	ccaggccago	aacatcatgc	acagcatcaa	cggctacgtg	1920
ttcgacagcc	tgcagctgtc	tgtgtgcctg	cacgagggtgg	cctactggta	catcctgagc	1980
atcggcgccc	agaccgactt	cctgtctgtg	ttttctctgt	gctacacctt	caagcacaag	2040
atgggtgtacg	aggacacccct	gaccctgttc	cccttcagcg	gcgagaccgt	tttcatgagc	2100
atggagaacc	ccggcctgtg	gatcctgggc	tgccacaaca	gcaacttcag	gaacaggggc	2160
atgaccgccc	tgctgaaagt	cagcagctgc	gacaagaaca	ccggcgcacta	ctacgaggac	2220
agctacgagg	acatcagcgc	ctacctgtcg	agcaagaaca	acgccccatcg	gcccaggagc	2280
ttcagccaga	accccccgt	gctgaagagg	caccagagg	agatcaccag	gaccaccctg	2340
cagagcgcacc	aggaggagat	cgactatgt	gacaccatca	gcgtggagat	gaagaaggag	2400
gacttcgaca	tctacgacga	ggacgagaac	cagagcccc	ggagcttcca	gaagaagacc	2460
aggcactact	tcatcgccgc	cgtggagagg	ctgtgggact	atggcatgag	cagcagcccc	2520
cacgtgtga	ggaacagggc	ccagagcggc	agegtgcccc	agttcaagaa	ggtgtgttc	2580
caggagttca	ccgacggcag	cttcacccag	cccctgtaca	gaggcgcagct	gaacgagcac	2640
ctgggcctgc	tgggcctcta	catcagggcc	gaggtggagg	acaacatcat	ggtgaccctc	2700
aggaaccagg	ccagcaggcc	ctacagcttc	tacagcagcc	tgatcagcta	cgaggaggac	2760
cagaggcagg	gcgcgcgcgc	caggaagaac	ttcgtgaagc	ccaacgcagac	caagacccatc	2820
ttcttggaaagg	tgcagcacca	catggccccc	accaaggacg	atttcgactg	caaggccctgg	2880
gctacttct	ctgatgtgga	cctggagaag	gacgtgcaca	ggggcctgat	cgggccccc	2940
ctgggtgtgcc	acaccaacac	cctgaacccc	gcccacggca	ggcagggtgac	cgtgcaggag	3000
ttcgccctgt	tcttaccat	cttcgacag	accaagagct	ggtacttcac	cgagaacatg	3060
gagaggaact	gcagggccccc	ctgcaacatc	cagatggagg	accccacctt	caaggagaac	3120
tacaggttcc	acgccccatcaa	cggtacatc	atggacaccc	tgcccgccct	ggtgtatggcc	3180
caggaccaga	ggatcaggtg	gtatctgtcg	agcatggca	gcaacgagaa	catccacagc	3240
atccactca	cgccgcacgt	tttcacccgt	aggaagaagg	aggagtacaa	gatggccctg	3300
tacaacctgt	accccgccgt	tttcgagacc	gtggagatgc	tgcccagcaa	ggccggcata	3360
tggagggtgg	agtgcctgat	cgccgcagcac	ctgcacgccc	gcatgagcac	cctgttctg	3420

-continued

gtgtacagca acaagtgccca gaccccccgt ggcatggcca gggccacat cagggactc	3480
cagatcacccg cctctggcca gtacggccag tggggcccca agctggccag gctgcactac	3540
agcggcagca tcaacgcctg gagcaccaag gagccctca gctggatcaa ggtggactg	3600
ctggccccca tgatcatcca cggcatcaag acccaggcg ccaggcagaa gttcagcagc	3660
ctgtacatca gccagttcat catcatgtac agectggacg gcaagaagtg gcagacctac	3720
agggcaaca gcacccggcac cctgtatggtg ttcttcggca acgtggacag cagcggcattc	3780
aagcacaaca tcttcaaccc ccccatcatc gccaggtaca tcaggtgca ccccacccac	3840
tacagcatca ggagcacccct gcggatggaa ctgtatggct ggcacctgaa cagctgcagc	3900
atgccccctgg gcatggagag caaggccatc tctgacgccc agatcaccgc cagcagctac	3960
ttcacaaca tggccac cttggagcccc agcaaggcca ggctgcacct gcagggcagg	4020
agcaacgcct ggaggccccca ggtgaacaac cccaggagt ggctgcaggt ggacttcag	4080
aagaccatga aggtgaccgg cgtgaccacc cagggcgtga agagcctgct gaccagcatg	4140
tacgtgaagg agttcctgtat cagcagcago caggacggcc accagtggac cctgttctc	4200
cagaacggca aagtgaaggt gttccagggc aaccaggaca gttcaccccg cgtggtaaac	4260
agcctggacc ccccccgtct gaccaggtat ctgaggatcc acccccagag ctgggtgcac	4320
cagatcgccc tgagaatgga agtgctggga tgcgaggccc aggacctgta ctga	4374

<210> SEQ_ID NO 21
<211> LENGTH: 1457
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER_INFORMATION: Description of Artificial Sequence: Synthetic polypeptide

<400> SEQUENCE: 21

Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe			
1	5	10	15

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

Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg			
35	40	45	

Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val			
50	55	60	

Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile			
65	70	75	80

Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln			
85	90	95	

Ala Glu Val Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser			
100	105	110	

His Pro Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser			
115	120	125	

Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp			
130	135	140	

Asp Lys Val Phe Pro Gly Gly Ser His Thr Tyr Val Trp Lys Ala Ser			
145	150	155	160

Lys Glu Asn Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser			
165	170	175	

-continued

Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile
180 185 190

Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr
195 200 205

Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly
210 215 220

Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp
225 230 235 240

Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr
245 250 255

Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val
260 265 270

Tyr Trp His Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile
275 280 285

Phe Leu Glu Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser
290 295 300

Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met
305 310 315 320

Asp Leu Gly Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His
325 330 335

Asp Gly Met Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro
340 345 350

Gln Leu Arg Met Lys Asn Asn Glu Ala Glu Asp Tyr Asp Asp Asp
355 360 365

Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser
370 375 380

Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr
385 390 395 400

Trp Val His Tyr Ile Ala Ala Glu Glu Asp Trp Asp Tyr Ala Pro
405 410 415

Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn
420 425 430

Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met
435 440 445

Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu
450 455 460

Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu
465 470 475 480

Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro
485 490 495

His Gly Ile Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys
500 505 510

Gly Val Lys His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe
515 520 525

Lys Tyr Lys Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp
530 535 540

Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg
545 550 555 560

Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu
565 570 575

-continued

Ser Val Asp Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val
 580 585 590
 Ile Leu Phe Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu
 595 600 605
 Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp
 610 615 620
 Pro Glu Phe Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val
 625 630 635 640
 Phe Asp Ser Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp
 645 650 655
 Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe
 660 665 670
 Ser Gly Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr
 675 680 685
 Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro
 690 695 700
 Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly
 705 710 715 720
 Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp
 725 730 735
 Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Ser Lys
 740 745 750
 Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser Gln Asn Pro Pro Val Leu
 755 760 765
 Lys Arg His Gln Arg Glu Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln
 770 775 780
 Glu Glu Ile Asp Tyr Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu
 785 790 795 800
 Asp Phe Asp Ile Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe
 805 810 815
 Gln Lys Lys Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp
 820 825 830
 Asp Tyr Gly Met Ser Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln
 835 840 845
 Ser Gly Ser Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr
 850 855 860
 Asp Gly Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu His
 865 870 875 880
 Leu Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp Asn Ile
 885 890 895
 Met Val Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe Tyr Ser
 900 905 910
 Ser Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg
 915 920 925
 Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys Val
 930 935 940
 Gln His His Met Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala Trp
 945 950 955 960
 Ala Tyr Phe Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly Leu
 965 970 975
 Ile Gly Pro Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro Ala His

-continued

980	985	990
Gly Arg Gln Val Thr Val Gln Glu Phe Ala Leu Phe Phe		Thr Ile Phe
995	1000	1005
Asp Glu Thr Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg Asn		
1010	1015	1020
Cys Arg Ala Pro Cys Asn Ile Gln Met Glu Asp Pro Thr Phe Lys		
1025	1030	1035
Glu Asn Tyr Arg Phe His Ala Ile Asn Gly Tyr Ile Met Asp Thr		
1040	1045	1050
Leu Pro Gly Leu Val Met Ala Gln Asp Gln Arg Ile Arg Trp Tyr		
1055	1060	1065
Leu Leu Ser Met Gly Ser Asn Glu Asn Ile His Ser Ile His Phe		
1070	1075	1080
Ser Gly His Val Phe Thr Val Arg Lys Lys Glu Glu Tyr Lys Met		
1085	1090	1095
Ala Leu Tyr Asn Leu Tyr Pro Gly Val Phe Glu Thr Val Glu Met		
1100	1105	1110
Leu Pro Ser Lys Ala Gly Ile Trp Arg Val Glu Cys Leu Ile Gly		
1115	1120	1125
Glu His Leu His Ala Gly Met Ser Thr Leu Phe Leu Val Tyr Ser		
1130	1135	1140
Asn Lys Cys Gln Thr Pro Leu Gly Met Ala Ser Gly His Ile Arg		
1145	1150	1155
Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr Gly Gln Trp Ala Pro		
1160	1165	1170
Lys Leu Ala Arg Leu His Tyr Ser Gly Ser Ile Asn Ala Trp Ser		
1175	1180	1185
Thr Lys Glu Pro Phe Ser Trp Ile Lys Val Asp Leu Leu Ala Pro		
1190	1195	1200
Met Ile Ile His Gly Ile Lys Thr Gln Gly Ala Arg Gln Lys Phe		
1205	1210	1215
Ser Ser Leu Tyr Ile Ser Gln Phe Ile Ile Met Tyr Ser Leu Asp		
1220	1225	1230
Gly Lys Lys Trp Gln Thr Tyr Arg Gly Asn Ser Thr Gly Thr Leu		
1235	1240	1245
Met Val Phe Phe Gly Asn Val Asp Ser Ser Gly Ile Lys His Asn		
1250	1255	1260
Ile Phe Asn Pro Pro Ile Ile Ala Arg Tyr Ile Arg Leu His Pro		
1265	1270	1275
Thr His Tyr Ser Ile Arg Ser Thr Leu Arg Met Glu Leu Met Gly		
1280	1285	1290
Cys Asp Leu Asn Ser Cys Ser Met Pro Leu Gly Met Glu Ser Lys		
1295	1300	1305
Ala Ile Ser Asp Ala Gln Ile Thr Ala Ser Ser Tyr Phe Thr Asn		
1310	1315	1320
Met Phe Ala Thr Trp Ser Pro Ser Lys Ala Arg Leu His Leu Gln		
1325	1330	1335
Gly Arg Ser Asn Ala Trp Arg Pro Gln Val Asn Asn Pro Lys Glu		
1340	1345	1350
Trp Leu Gln Val Asp Phe Gln Lys Thr Met Lys Val Thr Gly Val		
1355	1360	1365

-continued

Thr	Thr	Gln	Gly	Val	Lys	Ser	Leu	Leu	Thr	Ser	Met	Tyr	Val	Lys
1370					1375						1380			
Glu	Phe	Leu	Ile	Ser	Ser	Ser	Gln	Asp	Gly	His	Gln	Trp	Thr	Leu
1385					1390						1395			
Phe	Phe	Gln	Asn	Gly	Lys	Val	Lys	Val	Phe	Gln	Gly	Asn	Gln	Asp
1400					1405						1410			
Ser	Phe	Thr	Pro	Val	Val	Asn	Ser	Leu	Asp	Pro	Pro	Leu	Leu	Thr
1415					1420						1425			
Arg	Tyr	Leu	Arg	Ile	His	Pro	Gln	Ser	Trp	Val	His	Gln	Ile	Ala
1430					1435						1440			
Leu	Arg	Met	Glu	Val	Leu	Gly	Cys	Glu	Ala	Gln	Asp	Leu	Tyr	
1445					1450						1455			

<210> SEQ_ID NO 22
<211> LENGTH: 2220
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 22

gcacccagga	gatactacct	ggggccgtg	gagctgagct	gggactacat	gcagtctgac	60
ctggggagc	tgcctgtgga	cgcagggttc	ccccccagag	tgcccaagag	cttccccc	120
aacacccatcg	tggtgtacaa	gaagaccctg	ttcgtggagt	tcaccgacca	cctgttcaac	180
atcgccaagc	ccaggcccccc	ctggatgggc	ctgctggcc	ccaccatcca	ggccgagggt	240
tacgacaccg	tggtgatcac	cctgaagaac	atggccagcc	acccctgtgag	cctgcacgcc	300
gtgggggtga	gtaactggaa	ggcctctgag	ggcgccgagt	atgacgacca	gaccagccag	360
agggagaagg	aggacgacaa	ggtgttcccc	ggggcgaccc	acacctacgt	gtggcagggt	420
ctgaaggaga	acggccccat	ggccagcgac	ccctctgccc	tgacctacag	ctacccgtac	480
cacgtggacc	tggtaagga	cctgaactct	ggcctgtatcg	gcccctgtct	ggtgtgcagg	540
gagggcagcc	tggcaagga	gaagacccag	accctgcaca	atgtcatcct	gctgtcgcc	600
gtgttcgatg	agggcaagag	ctggcacacgc	gagaccaaga	acacccgtat	gcaggacagg	660
gtggccgcct	ctggccaggcc	ctggcccaag	atgcacaccc	tgaacggcta	cgtgaacacgg	720
gcctgtcccc	gcctgtatcg	ctgcccacagg	aagtctgtgt	actggcacgt	gatccggatg	780
ggcaccaccc	ccgagggtgca	cagcatcttc	ctggaggggcc	acacccctct	ggtgaggaac	840
cacaggcagg	ccagccctgga	gatcggccccc	atcaccttcc	tgaccggcc	gaccctgt	900
atggacccctgg	gccagttccct	gctgttctgc	cacatcggca	gccaccagca	cgacggcatg	960
gaggcctacg	tgaagggtgga	cagctgtcccc	gaggagcccc	agctgaggat	gaagaacaac	1020
gaggaggccg	aggactatga	tgtatgaccc	accgactctg	agatggacgt	ggtgagggtt	1080
gtatgtacca	acagcccccag	cttcatccag	atcagggtctg	tggccaaagaa	gcaccccaag	1140
acctgggtgc	actacatcgc	cgccgaggag	gaggactggg	actacccccc	cctgggtctg	1200
ccccccgacg	acaggagcta	caagagccag	tacctgaaca	acggccccca	gaggatccgc	1260
aggaagtaca	agaagggtcag	attcatggcc	tacaccgacg	agaccccaa	gaccaggagg	1320
gccccatccagc	acgagtctgg	catactggcc	ccctgtctgt	acggcgagggt	gggcgacacc	1380

-continued

ctgctgtatca tcttcaagaa ccaggccagc agggccatac acatctaccc ccacggcatc	1440
accgatgtga ggccccgtta cagcaggagg ctgcccaggc gctgtaaagca cctgaaggac	1500
tcccccattcc tgcccgccga gatcttcaag tacaagtggc ccgtgaccgt ggaggatggc	1560
cccaccaagt ctgaccccaag gtgcctgacc aggtactaca gcagcttcgt gaacatggag	1620
agggacctgg cctctggcct gatcgcccc ctgctgtatct gctacaagga gagcgtggac	1680
cagaggggca accagatcat gtctgacaag aggaacgtga tcctgttctc tgtgttcgtat	1740
gagaacagga gctggtatct gaccgagaac atccagaggt tcctgeccaa ccccgccgc	1800
gtgcagctgg aggacccga gttccaggcc agcaacatca tgcacagcat caacggctac	1860
gtgttcgaca gcctgcagct gtctgtgtgc ctgcacgagg tggcctactg gtacatctg	1920
agcatcggcg cccagaccga cttectgtct gtgttcttct ctggctacac cttcaagcac	1980
aagatggtgt acgaggacac cctgaccctg ttcccttca gggcgagac cgtgttcatg	2040
agcatggaga accccggctt gtggatcttg ggctgccaca acagcgactt caggaacagg	2100
ggcatgaccg ccctgctgaa agtcagcgc tgcgacaaga acaccggcga ctactacgag	2160
gacagctacg aggacatca ggcctacctg ctgagcaaga acaacgocat cgagccagg	2220

<210> SEQ ID NO 23
<211> LENGTH: 2052
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 23

gagatcacca ggaccacccgc gtagagcgac caggaggaga tcgactatga tgacaccatc	60
agcgtggaga tgaagaagga ggacttcgac atctacgacg aggacgagaa ccagacccc	120
aggagcttcc agaagaagac cagggactac ttcatcgccg ccgtggagag gctgtggac	180
tatggcatga gcagcagccc ccacgtgtcg aggaacaggg cccagacggc cagcgtgccc	240
catgttcaaga aggtggtgtt ccaggagttc accgacggca gttcaccca gcccctgtac	300
agaggcgacg tgaacgagca cctggccctg ctggggccct acatcaggc cgagggtggag	360
gacaacatca tggtgacctt caggaaccag gccagcaggc cctacagctt ctacagcagc	420
ctgatcagct acgaggagga ccagaggcag ggccgcgcgac ccaggaagaa cttctgtgaag	480
cccaacgaga ccaagaccta cttctggaag gtgcagcacc acatggcccc caccaaggac	540
gagttcact gcaaggctg ggcctacttc tctgtatgtgg acctggagaa ggacgtgcac	600
agcggcctga tcggccccc gctgggtgtc cacaccaaca ccctgaaccc cgccacggc	660
aggcagggtga ccgtgcagga gttcggccctg ttcttcacca ttttcacga gaccaagac	720
tggtaactca ccgagaacat ggagaggaac tgcaggccc cctgcaacat ccagatggag	780
gaccccccact tcaaggagaa ctacagggttc cacgcccatac acggctacat catggacacc	840
ctgccccggcc tgggtatggc ccaggaccag aggtatcggt ggtatctgtct gagcatggc	900
agcaacgaga acatccacag catccacttc agcggccacg tggttccatgtt gaggaagaag	960
gaggagttaca agatggccct gtacaacctg taccggccggc tggttcgagac cgtggatgt	1020
ctgccccagca aggccggcat ctggagggtg gagtgcctga tcggcgagca cctgcacgccc	1080
ggcatgagca ccctgttctt ggtgtacagc aacaagtgcc agacccccc gggcatggcc	1140

-continued

agcgccaca tcagggactt ccagatcacc gcctctggcc agtacggcca gtgggcccc	1200
aagctggcca ggctgcacta cagcggcagc atcaacgcct ggagcaccaa ggagcccttc	1260
agctggatca aggtggacct gctggccccc atgatcatcc acggcatcaa gaccaggc	1320
gccaggcaga agttcagcag cctgtacatc agccagttc tcatcatgta cagcctggac	1380
ggcaagaagt ggcagaccta caggggcaac agcacccgca ccctgatggt gttttcgcc	1440
aacgtggaca gcagcggcat caagcacaac atcttcaacc ccccccattcat cgccaggat	1500
atcaggctgc acccccaccca ctacagcattc aggagcaccc tgccgatgga actgatggc	1560
tgcgacactga acagctgcag catgcccctg ggcattggaga gcaaggccat ctctgacgcc	1620
cagatcacccg ccagcagacta cttcaccaac atgttcgcca cctggagccc cagcaaggcc	1680
aggctgcacc tgcaggcag gagcaacgccc tggaggcccc aggtgaacaa ccccaaggag	1740
tggctgcagg tggacttcca gaagaccatg aaggtgaccg gcgtgaccac ccagggcg	1800
aagagectgc tgaccagcat gtacgtgaag gagttcctga ttagcagcag ccaggacggc	1860
caccagtggc ccctgttctt ccagaacggc aaagtgaagg tggccagggg caaccaggac	1920
agtttcaccc ccgtggtgaa cagcctggac ccccccctgc tgaccaggta tctgaggatc	1980
caccccccaga gctgggtgca ccagatcgcc ctgagaatgg aagtgtggg atgcgaggcc	2040
caggacctgt ac	2052

<210> SEQ ID NO 24
<211> LENGTH: 2220
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 24

gcaccaggaga gatactaccc gggggctgtg gaactttctt gggactacat gcagtctgac	60
ctggggagac tgcctgtgga tgccaggatcc ccacccagag tgcccaagtgc ttcccatcc	120
aacacccctctg tggctcacaa gaagacactc tttgtggaaat tcactgacca cctgttcaac	180
attgcacaaac ccagaccacc ctggatggga ctccctggac ccaccattca ggctgaggatg	240
tatgacactg tggctcatcac cctcaagaac atggcatccc accctgtgtc tctgcatgct	300
gtggggagtct catactggaa agcctctgaa ggggctgagt atgatgacca gacatcccg	360
agagagaaaag aggtgaccaa ggtgtccct gggggatctc acacctatgt gtggcaagtc	420
ctcaaggaga atggaccat ggcacatctgac ccactctgccc tgacatactc ctacccatct	480
catgtggacc tggctcaagga cctcaactct ggactgatgg gggactgtct ggtgtgcagg	540
gaaggatccc tggccaaaggaa gaaaaccccg acactgcaca agttcattct cctgttgct	600
gtctttgatg agggcaagtc ttggcactct gaaacaaaga actccctgtat gcaagacagg	660
gatgtgcctt ctgcggggc atggcccaag atgcacactg tgaatggcta tgtgaacaga	720
tcactgcctg gactcattgg ctgcacagg aaatctgtct actggcatgt gatggcatg	780
gggacaaccc ctgaagtgca ctccatccc ctggaggcc acacccctt ggtcaggAAC	840
cacagacaag cctctctggaa gatctctccc atcaccttcc tcactgcaca gacactgt	900
atggacacctg gacagttccct gctgttctgc cacatcttcc cccaccagca tgatggcatg	960

-continued

gaaggctatg tcaagggttga ctcatgccct gaggaaaccac agctcaggat gaagaacaat	1020
gaggaggctg aggactatga tcatgacctg actgactctg agatggatgt ggtcagattt	1080
gatgtatgaca actctccatc ctccatttagt atcaggctctg tggcaaaagaa acaccccaag	1140
acatgggtgc actacattgc tgctgaggaa gaggactggg actatgcacc actggcctg	1200
gcccctgatg acaggagcta caagtctcg tacctaaca atggccaca aagaatttggaa	1260
agaaagtaca agaaagtctg attcatggcc tacactgatg aaaccttcaa gacaagagaa	1320
gcatttcagc atgagtctgg cattctggaa ccactcctgt atggggaaat gggagacacc	1380
ctgctcatca ttctcaagaa ccaggccctcc agggccctaca acatctaccc acatggcattc	1440
actgtatgtca ggccctgtta cagcaggaga ctgccaatggaa ggggtgaaaca cctcaaggac	1500
ttccccatcc tgcctggaga gatcttcaag tacaagtggaa ctgtcactgt ggaggatggaa	1560
ccaaacaaatg ctgacccttggatgtcacc agataactact cctctttgtt gaacatggag	1620
agagacotgg catctggact gattggacca ctgctcatct gotacaagga gtctgtggac	1680
cagagaggca accagatcat gtctgacaag agaaatgtga ttctgttctc tgcactccat	1740
gagaacacatg catggtacct gactgagaac attcagagat toctgeccaa ccctgctggg	1800
gtgcaactgg aagaccctga gtccaggca agcaacatca tgcactccat caatggctat	1860
gtgtttgact ctctccagct ttctgtctgc ctgcatttggggacttggctactg gtacatttt	1920
tctattgggg cacaaaactga ctccctttct gtcttctct ctggatacac cttcaaggcac	1980
aagatgggtt atgaggacac cctgacactc ttccattct ctggggaaac tgtgttcatg	2040
agcatggaga accctggact gtggattctg ggatgccaca actctgactt cagaaacagg	2100
ggaatgactg cactgctcaa agtctccctcc tgcataaca acactggggacttggatggaa	2160
gactcttatg aggacatctc tgcctacctg ctcagcaaga acaatggccat tgagccaga	2220

```

<210> SEQ ID NO 25
<211> LENGTH: 2052
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polynucleotide
  
```

```

<400> SEQUENCE: 25
  
```

gagatcacca ggacaaccct ccagtctgac caggaagaga ttgactatga tgacaccatt	60
tctgtggaga tgaagaagga ggactttgac atctatgtatg aggacgagaa ccagtctcca	120
agatcattcc agaagaagac aagacactac ttcatgtctg ctgtggaaag actgtgggac	180
tatggcatgt ctccctctcc ccatgtctcc aggaacaggc cacagtctgg ctctgtgcc	240
cagttcaaga aagtggcttt ccaggagttc actgtatggctt cattcaccca gcccctgtac	300
agaggggaac tgaatgagca cctgggactc ctgggaccat acatcaggc tgagggtggaa	360
gacaacatca tggtgacatt cagaaaccag gcctccaggc cctacagctt ctactcttcc	420
ctcatcagct atgaggaaga ccagagacaa ggggctgagc caagaaagaa ctttgtgaaa	480
cccaatgaaa ccaagaccta cttctggaaa gtccagcacc acatggcacc caccaaggat	540
gagtttggact gcaaggctg ggcatacttc tctgtatgtgg acctggagaa agatgtgcac	600
tctggcctga ttggccact cctggctctgc cacaccaaca ccctgaaccc tgcacatggaa	660
aggcaagtga ctgtgcagga gtttgcctcc ttcttcacca tctttgtatga aaccaagtca	720

-continued

tggtaacttca	ctgagaacat	ggagagaaaac	tgcagagcac	catgcaacat	tcagatggaa	780
gacccccacct	tcaaggagaa	ctacagggtc	catgccatca	atggctacat	catggacacc	840
ctgcctggc	ttgtcatgge	acaggaccag	agaatcgat	ggtacctgt	ttctatggaa	900
tccaaatgaga	acattcactc	catccacttc	tctggcatg	tcttcactgt	gagaaagaag	960
gaggaataaca	agatggccct	gtacaacetc	taccctgggg	tctttgagac	tgtggagatg	1020
ctgcccctcca	aagctggcat	ctgggggtg	aatgcctca	ttggggagca	cctgcatgt	1080
ggcatgtcaa	ccctgttctt	ggtctacagc	aacaagtgc	agacacccct	ggaaatggcc	1140
tctggccaca	tcagggactt	ccagatcact	gcctctggc	agtatggca	gtgggcaccc	1200
aaactggcca	ggctccacta	ctctggctcc	atcaatgcat	ggtcaaccaa	ggagccattc	1260
tcttggatca	aggtggaccc	gctggcaccc	atgatcattc	atggcatcaa	gacacagggg	1320
gcaagacaga	atttctccctc	tctgtacatc	tcacagtca	tcatcatgt	ctctctggat	1380
ggcaagaagt	ggcagacata	cagaggcaac	tccactggca	ccctcatgtt	cttcttggc	1440
aatgtggaca	gctctggcat	caagcacaac	atcttcaacc	ctcccatcat	tgccagatac	1500
atcaggctgc	accccccacca	ctactcaatc	agatcaaccc	tcaggatgga	actgtatggaa	1560
tgtgacactga	actcctgctc	aatgcccctg	ggaatggaga	gcaaggccat	ttctgtatgcc	1620
cagatcaactg	catcctctta	cttcaccaac	atgttgcca	cctggtcacc	atcaaaagcc	1680
aggctgcacc	tccagggaaag	aagcaatgcc	tggagacccc	aggtcaacaa	cccaaaggaa	1740
tggctgcaag	tggacttcca	gaagacaatg	aaagtcaactg	gggtgacaac	ccagggggtc	1800
aagtctctgc	tcacactcaat	gtatgtgaag	gagttccatg	tctcttcctc	acaggatggc	1860
caccagtgga	cactcttctt	ccagaatggc	aaagtcaagg	tgttccaggg	caaccaggac	1920
tctttcacac	ctgtggatgaa	ctcaactggac	ccccccctcc	tgacaagata	cctgagaatt	1980
caccccccagt	cttgggtcca	ccagattgcc	ctgagaatgg	aagtccctggg	atgtgaggca	2040
caagacactgt	ac					2052

<210> SEQ ID NO 26
<211> LENGTH: 4332
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 26

atgcagattg	agctgtccac	ctgtttcttt	ctgtgcctgc	tgagattctg	cttctctgcc	60
accaggagat	actacactggg	ggctgtggaa	ctttcttggg	actacatgca	gtctgacactg	120
ggagagctgc	ctgtggatgc	caggttccca	cccagagtgc	ccaagtccctt	cccattcaac	180
acctctgtgg	tctacaagaa	gacactcttt	gtggaattca	ctgaccaccc	gttcaacatt	240
gaaaaaccca	gaccacccctg	gatgggactc	ctggggaccca	ccattcaggc	tgagggttat	300
gacactgtgg	tcatcaccc	caagaacatg	gcatccacc	ctgtgtctct	gcatgtgtg	360
ggagtctcat	actggaaagc	ctctgaaggg	gctgagttatg	atgaccagac	atcccagaga	420
gagaaagagg	atgacaaggt	gttccctggg	ggatctcaca	cctatgtgtg	gcaagtccctc	480
aaggagaatg	gaccatggc	atctgaccca	ctctgcctga	cataactccct	cctttctcat	540

-continued

gtggacctgg tcaaggacct caactctgga ctgattgggg cactgtggc gtgcagggaa	600
ggatccctgg ccaaggagaa aaccaggaca ctgcacaagt tcattctcct gtttgcgtc	660
tttgatgagg gcaagtcttgc acactctgaa acaaagaact ccctgtatgca agacagggat	720
gtgcctctg ccagggcatg gcccaagatg cacactgtga atggctatgt gaacagatca	780
ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcatgtatg tggcatgggg	840
acaacccctg aagtgcactc catttcctg gagggacaca ctttcctgtt caggaaccac	900
agacaaggct ctctggagat ctctcccata accttcctca ctgcacagac actgtgtatg	960
gacccctggac agtttcctgtt gtttcgccac atcttcctcc accagcatga tggcatggaa	1020
gcctatgtca aggtggactc atgccttgtag gaaccacagc tcaggatgaa gaacaatgag	1080
gaggctgagg actatgtatg tgacactgact gactctgaga tggatgtggc cagatttgat	1140
gatgacaact ctccatctt cattcagatc aggtctgtgg caaagaaaca ccccaagaca	1200
tgggtgact acattgtgtc tgaggaagag gactggactc atgcaccact ggtcctggcc	1260
cctgatgaca ggagctacaa gtctcagtagc ctcaacaatg gcccacaagaa aatttggaa	1320
aagtacaaga aagttagatt catggctac actgtatgaa cttcaagac aagagaagcc	1380
attcagatc agtctggcat tctggatcca ctctctgtatg gggaaatggg agacaccctg	1440
ctcatcatct tcaagaacca ggcctccagg ccctacaaca ttatccccaca tggcatca	1500
gatgtcaggc ccctgtacag caggagactg cccaaagggg tgaaacaccc caaggactc	1560
cccatctgc ctggagatgt ctcaagtagc aagtggactg tcactgtgg gatggacca	1620
acaaagtctg accccagggtg cctcaccaga tactactctt cttttgtgaa catggagaga	1680
gacctggcat ctggactgtatg tggaccactg ctcatctgtt acaaggagtc tggaccag	1740
agaggcaacc agatcatgtc tgacaagaga aatgtgattc tggatctgtt ctttgcgtat	1800
aacagatcat ggtacactgac tgagaacatt cagagattcc tgcccaaccc tgctgggtg	1860
caactggaa accctgagtt ccaggcaagc aacatcatgc actccatcaa tggctatgt	1920
tttgcacttc tccagtttc tgtctgcctg catgaggatgg cttactggta cattttct	1980
attggggcac aaactgactt ctttctgtc ttcttctctg gatacacctt caagcacaag	2040
atgggtgtatg aggacaccct gacactcttc ccattctgt gggaaactgt gttcatgagc	2100
atggagaacc ctggactgtg gattctggaa tgccacaact ctgacttcg aaacagggg	2160
atgactgcac tgctcaaagt ctccctctgt gacaagaaca ctggggacta ctatgaggac	2220
tcttatgagg acatctctgc ctacactgtc agcaagaaca atgcccattga gcccagagag	2280
atcaccaggaa caaccctcca gtctgaccag gaagagatgg actatgtatg caccattct	2340
gtggagatga agaaggagga ctttgacatc tatgtatgagg acgagaacca gtctccaaga	2400
tcattccaga agaagacaag acactacttc attgctgtg tggaaagact gtggactat	2460
ggcatgtctt cctctccca tggatctcagg aacaggcac agtctggctc tggccacag	2520
ttcaagaaag tggatctcca ggaggactatc gatggctatc tcaacccagcc cttgtacaga	2580
ggggaaactga atgagcacct gggactctg ggaccataca tcagggctga ggtggaaagac	2640
aacatcatgg tgacattcg aaaccaggcc tccaggccct acagcttcta ctctccctc	2700
atcagctatg aggaagacca gagacaaggg gctgagccaa gaaagaactt tggatggaa	2760
aatgaaacca agacacttctt ctggaaagtc cagcaccaca tggcacccac caaggatgag	2820

-continued

tttgactgca	aggcctggc	atacttctct	gatgtggacc	tggagaaaaga	tgtgcactct	2880
ggcctgattg	gcccaactcct	ggtctgccac	accaacaccc	tgaaccctgc	acatggaaagg	2940
caagtgactg	tgcaggagtt	tgcctcttc	ttcaccatct	ttgatgaaac	caagtcatgg	3000
tacttcactg	agaacatgga	gagaaactgc	agagcaccat	gcaacattca	gatggaaagac	3060
cccacccatca	aggagaacta	caggttccat	gccccatcg	gctacatcat	ggacaccctg	3120
cctgggcttg	tcatggcaca	ggaccagaga	atcagatggt	acctgcttgc	tatggatcc	3180
aatgagaaca	ttcactccat	ccacttctct	gggcattgtct	tcactgtgag	aaagaaggag	3240
gaataacaaga	tggccctgtta	caacctctac	cctggggctct	ttgagactgt	ggagatgctg	3300
ccctccaaag	ctggcatctg	gagggtggaa	tgcctcattg	gggagcacct	gcatgctggc	3360
atgtcaaccc	tgttccctgg	ctacagcaac	aagtgcaga	cacccctggg	aatggctct	3420
ggccacatca	gggacttcca	gatcaactgac	tctggccagt	atggccagtg	ggcacccaaa	3480
ctggccaggc	tccactactc	tggctccatc	aatgcattgt	caaccaagga	gccattctct	3540
tggatcaagg	tggacctgct	ggcacccatg	atcattcatg	gcatcaagac	acagggggca	3600
agacagaaat	tctccctctct	gtacatctca	cagttcatca	tcatgtactc	tctggatggc	3660
aagaagtggc	agacatacag	aggcaactcc	actggcaccc	tcatggcttt	ctttggcaat	3720
gtggacagct	ctggcatcaa	gcacaacatc	ttcaaccctc	ccatcattgc	cagatacatc	3780
aggctgcacc	ccacccacta	ctcaatcaga	tcaaccctca	ggatggact	gatggatgt	3840
gacctgaact	cctgetcaat	gccccctgg	atggagagca	aggccatttc	tcatggccag	3900
atcaactgcat	cctcttactt	caccaacatg	tttggccac	ggtcaccatc	aaaagccagg	3960
ctgcacccctc	agggagaag	caatgcctgg	agaccccagg	tcaacaaccc	aaaggaatgg	4020
ctgcaagtgg	acttccagaa	gacaatgaa	gtcaactgggg	tgacaaccca	gggggtcaag	4080
tctctgtca	cctcaatgtt	tgtgaaggag	tccctgtatc	ttccctcaca	ggatggcac	4140
cagtggacac	tcttcttcca	gaatggcaaa	gtcaagggtt	tccaggccaa	ccaggactct	4200
ttcacacactg	tggtaactc	actggacccc	ccctcttgc	caagataacct	gagaattcac	4260
ccccagctt	gggtccacca	gattgcctg	agaatggaa	tcctggatg	tgaggcacaa	4320
gacctgtact	ga					4332

```

<210> SEQ_ID NO 27
<211> LENGTH: 4368
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
polynucleotide

```

<400> SEQUENCE: 27

atgcagattg	agctgtccac	ctgtttcttt	ctgtgcctgc	tgagattctg	tttctctgcc	60
accaggagat	actacctggg	ggctgtggaa	cttcttggg	actacatgca	gtctgacactg	120
ggagagctgc	ctgtggatgc	caggttccca	cccagagtgc	ccaagtcctt	cccattcaac	180
acctctgtgg	tctacaagaa	gacactcttt	gtggaattca	ctgaccaccc	gttcaacatt	240
gaaaaaccca	gaccaccctg	gatgggactc	ctgggaccca	ccattcaggc	tgagggttat	300
gacactgtgg	tcatcaccct	caagaacatg	gcatcccacc	ctgtgtctct	gcatgctgt	360
ggagtgctcat	actggaaagc	ctctgaaggg	gctgagttatg	atgaccagac	atcccagaga	420

-continued

gagaaaagggt atgacaagggt gttccctggg ggatctcaca cctatgtgtg gcaagtctc	480
aaggagaatg gacccatggc atctgaccca ctctgcctga catactccctt ccttcctcat	540
gtggacctgg tcaaggacctt caactctggaa ctgattgggg cactgctgtt gtgcaggaa	600
ggatccctgg ccaaggagaa aacccagaca ctgcacaagt tcatttcctt gtttgctgtc	660
tttgatgagg gcaagtcttgc gcactctgaa acaaagaact ccctgatgca agacaggat	720
gtgcctctg ccagggcatg gcccaagatg cacactgtga atggctatgt gaacagatca	780
ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcatgttat tggcatgggg	840
acaacccctg aagtgcactc catttcctggaa gagggacaca ctttcctgggtt caggaaccac	900
agacaagcctt ctctggagat ctctccatc accttcctca ctgcacagac actgctgtat	960
gaccttggac agtttcctgtt gttctgccac atcttcctccc accagcatgaa tggcatggaa	1020
gccttatgtca aggtggactc atgccttgag gaaccacagc tcaggatgaa gaacaatgag	1080
gaggctgagg actatgtatgaa tgacactgact gactctgaga tggatgttgtt cagatttgat	1140
gtatgacaactt ctccatcctt cattcagatc aggtctgtgg caaagaaaca ccccaagaca	1200
tgggtgactt acatttgcgtc tgaggaagag gactggactt atgcaccactt ggtcctggcc	1260
cctgtatgaca ggagctacaa gtctcagttt ctcaacaatg gcccacaaag aatttggaaaga	1320
aagtacaaga aagttagtcatc catggcctac actgtatgaaa cttcaagac aagagaagcc	1380
atttcagatc agtctggcat tctggggatca ctccctgtatg gggaaatgggg agacaccctg	1440
cttcatactt tcaagaacca ggcctccagg ccctacaaca tctaccacata tggcatact	1500
gtatgtcaggc ccctgtacatc caggagactg cccaaagggg tgaaacacactt caaggacttc	1560
ccatttcgtc ctggagagat cttcaagttt aagtggactt gactgttgaa ggttggacca	1620
acaaagtctg accccggtt cctcaccaga tactactctt cttttgtt gaaatggagata	1680
gaccttggcat ctggactgtatg tggaccactg ctcatctgtt acaaggatgtc tggaccacat	1740
agaggcaacc agatcatgtc tgacaagatgaa aatgtgtt tggatgttgtt cttttgtt gaaatggagata	1800
aacagatcat ggtacactgac tgagaacattt cagagattcc tgcccaaccc tggatgttgtt	1860
caactggaaatg accctgtatc ccaggcaaccc aacatcatgtc actccatcaa tggatgttgtt	1920
tttgcacttc tccagtttc tggatgttgtt catgaggtgg cttactggta cattttttttt	1980
attggggcac aaactgtactt ctttcgttc ttcttcgtt gatacacctt caagcacaag	2040
atgggtgtatg aggacaccctt gacactcttc ccatttcgtt gggaaatgtt gttcatgtt	2100
atggagaacc ctggactgtt gatttggatca tgccacaactt ctgacttcag aaacaggggaa	2160
atgactgcac tggatgttgtt ctttcgttc gataacaactt ctggggactt ctatgaggac	2220
tctttatgtt gacatcttcgtc ctacatgttca agcaagaaca atgcccattttt gcccacaag	2280
tttctctcaga attccagacca ccccaaggacc accggatgtt ccaggacaac cttccatgtt	2340
gaccaggaaatg agatttgcgtt tggatgttgtt atttctgtt gatgttgtt ggaggacttt	2400
gacatctatg atgaggacgtt gacccatgtt ccaagatcat tccagaagaa gacaagacac	2460
tacttcattt gtttgcgtt gtttggatgtt gactatggca tggatgttgtt tcccatgtt	2520
ctcaggaaaca gggcacatgtt tggatgttgtt ccacatgttca agaaatgtt cttccaggat	2580
tttgcactgtt gtttgcattttt ccaggccctt tacaggggg aactgtatgaa gtttgcgtt	2640
ctccttggac catacatgtt ggttgcgtt gtttgcgtt gtttgcgtt gtttgcgtt	2700

-continued

caggcctcca	ggccctacag	cttctactct	tccctcatca	gctatgagga	agaccagaga	2760
caaggggctg	agccaagaaa	gaaccttgtg	aaacccaatg	aaaccaagac	ctacttctgg	2820
aaagtccagc	accacatggc	acccaccaag	gatgagtttgc	actgcaaggc	ctgggcatac	2880
ttctctgatg	tggaccttgg	gaaagatgtg	cactctggc	tgattggccc	actcttggtc	2940
tgccacacca	acaccctgaa	ccctgcacat	ggaaggcaag	tgactgtgca	ggagttgcc	3000
ctcttcttca	ccatcttga	tgaaaccaag	tcatggact	tcactgagaa	catggagaga	3060
aactgcagag	caccatgcaa	cattcagatg	gaagacccca	ccttcaagga	gaactacagg	3120
ttccatgcca	tcaatggcta	catcatggac	accctgccttgc	ggcttgcata	ggcacaggac	3180
cagagaatca	gatggtacct	gctttctatg	ggatccaatg	agaacattca	ctccatccac	3240
ttctctgggc	atgtcttcac	tgtgagaaag	aaggaggaat	acaagatggc	cctgtacaac	3300
ctcttacccctg	gggtctttga	gactgtggag	atgctgccc	ccaaagctgg	catctggagg	3360
gtggaatgcc	tcattgggaa	gcacactgcat	gctggcatgt	caaccctgtt	cctggcttac	3420
agcaacaagt	gccagacacc	cctgggaatg	gcctctggcc	acatcaggga	cttccagatc	3480
actgcctctg	gccagtatgg	ccagtgccca	ccaaactgg	ccaggctca	ctactctggc	3540
tccatcaatg	catggtcaac	caaggagcc	ttctcttgg	tcaaggtgg	cctgtggca	3600
cccatgatca	ttcatggcat	caagacacag	ggggcaagac	agaaattctc	ctctctgtac	3660
atctcacatg	tcatcatcat	gtactctctg	gatggcaaga	agtggcagac	atacagaggc	3720
aactccactg	gcacccctat	ggtcttcttgc	ggcaatgtgg	acagctctgg	catcaagcac	3780
aacatcttca	accctccccat	cattggccaga	tacatcaggc	tgcacccac	ccactactca	3840
atcagatcaa	ccctcaggat	ggaactgtatg	ggatgtgacc	tgaactccctg	ctcaatgccc	3900
ctgggaatgg	agagcaaggc	catttctgtat	gcccagatca	ctgcacccctc	ttacttcacc	3960
aacatgtttg	ccacctggtc	accatcaaaa	gccaggctgc	acctccagg	aagaagcaat	4020
gcctggagac	cccaggtcaa	caacccaaag	aatggctgc	aagtggactt	ccagaagaca	4080
atgaaagtca	ctggggtgac	aacccagggg	gtcaagtctc	tgctcacctc	aatgtatgt	4140
aggaggatcc	tgatctcttc	ctcacaggat	ggccaccagg	ggacactctt	cttccagaaat	4200
ggcaaagtca	agggtttcca	gggcaaccag	gactcttca	cacctgtgg	gaactcactg	4260
gacccccc	tcctgacaag	atacctgaga	attcaccc	agtcttgg	ccaccagatt	4320
gccttgagaa	tggaagtctt	gggatgtgag	gcacaagacc	tgtactga		4368

```

<210> SEQ ID NO 28
<211> LENGTH: 4332
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polynucleotide

```

```

<400> SEQUENCE: 28

atgcagattg agctgagcac ctgtttcttc ctgtgcctgc tgagggtctg cttctctgccc 60
accaggagat actaccttggg cgccgtggag ctgagctggg actacatgca gtctgacactg 120
ggcgagctgc ctgtggacgc caggttcccc cccagagtgc ccaagagctt ccccttcaac 180
acctcagtgg tgcataagaa gacccgttca gtggagttca ccgaccaccc gttcaacatc 240

```

-continued

gccaaggcca ggccccctg gatgggcctg ctggggccca ccatccagc cgagggtac	300
gacaccgtgg tgatcacccct gaagaacatg gccagccacc ccgtgagcct gcacgcccgt	360
ggcgtgagct acttggaaaggc ctctgagggc gccgagttatg acgaccagac cagccagagg	420
gagaaggagg acgacaagggt gtccccggc ggcagccaca cctacgtgtg gcaggtgctg	480
aaggagaacg gccccatggc cagcgacccc ctgtgcctga cctacagcta cctgagccac	540
gtggacactgg tgaaggacct gaactctggc ctgatcggcg ccctgtgtt gtgcaggag	600
ggcagecctgg ccaaggagaa gaccagacc ctgcacaagt tcatacctgtt gttcggcgt	660
ttcgatgagg gcaagagctg gcacagcgag accaagaaca gcctgatgca ggacaggat	720
gcgcctctg ccaggcctgg gcccaagatg cacaccgtga acggctacgt gaacaggagc	780
ctgccccggcc tgatcggctg ccacaggaag tctgtgtact ggcacgtgtat cggcatggc	840
accacccccc aggtgcacag catcttcctg gaggcccaca ctttccttgtt gaggaaccac	900
aggcaggcca gcctggagat cagccccatc accttcctga ccccccacac cctgtgtatg	960
gacactgggcc agttctgtct gttctgccac atcagcagcc accagcacga cggcatggag	1020
gcctacgtga aggtggacag ctgccccgag gagcccccgc tgaggatgaa gaacaacgag	1080
gaggccgagg actatgtatga tgacctgacc gactctgttga tggacgttgtt gaggtttgt	1140
gtatgacaaca gccccagctt catccagatc aggtctgtgg ccaagaagca ccccaagacc	1200
tgggtgcaact acatgcgcg cggaggaggag gactggact acggcccccctt ggtgtggcc	1260
cccgaeacaca ggagetacaa gagecagtagt ctgaacaacg gccccagag gatcgccagg	1320
aagtacaaga aggtcagatt catggcctac accgacgaga cttcaagac cagggaggcc	1380
atccagcact agtctggcat cctggggccc ctgctgtacg gcgagggtgg cgacaccctg	1440
ctgatcatct tcaagaacca ggccagcagg ccctacaaca tctacccca cggcatcacc	1500
gtatgtgggc ccctgtacat caggaggctg cccaaaggccg tgaagcacct gaaggactt	1560
cccatectgc ccggcggagat cttcaagttt aagtggaccg tgaccgttga ggttggcc	1620
accaagtctg accccaggtg cctgaccagg tactacagca gtttgcgttga catggaggg	1680
gacctggccct ctggectgtat cgccccctgg ctgatctgtt acaaggagag cgtggacac	1740
aggggcaacc agatcatgtc tgacaagagg aacgtgtatcc tttttctctgtt gttcgtatg	1800
aacaggagct ggtatctgtac cgagaacatc cagagggttcc tgcccaaccc cggccggctg	1860
cagctggagg accccgagtt ccaggccagc aacatcatgc acagcatcaa cggctacgt	1920
ttcgacagcc tgcagctgtc tttgtgcctg cacgagggttcc cttactgttta catctgtac	1980
atcggccccc agaccgactt cctgtctgtt ttcttctctgtt gctacacccctt caagcacaag	2040
atgggttacg aggacaccctt gaccctgttc cccttcagcg gcgagaccgtt gttcatgtac	2100
atggagaacc cccgcctgtt gatcctgggc tgccacaaca gcgacttcag gaacaggggc	2160
atgaccgccc tgctgaaaatg cagcagctgc gacaagaaca cccggcacta ctacgaggac	2220
agctacgagg acatcagcgc ctactgtgtt agcaagaaca acgccatcga gcccaggag	2280
atcaccagga ccaccctgca gagcgaccag gagggatcg actatgtatga caccatcagc	2340
gtggagatga agaaggagga ctgcacatc tacgacgagg acgagaacca gagccccagg	2400
agcttccaga agaagaccag gcactacttc atcgccggcc tggagaggct gtggacttat	2460
ggcatgagca gcageccccca cgtgctgagg aacaggccc agagcggcag cgtgccccag	2520

-continued

ttaagaagg tggtgttcca ggagttcacc gacggcaget tcacccagcc cctgtacaga	2580
ggcgagctga acgagcacct gggcctgctg ggcccctaca tcagggccga ggtggaggac	2640
aacatcatgg tgaccttcag gaaccaggcc agcaggccct acagcttcta cagcagcctg	2700
atcagctacg aggaggacca gaggcagggc gccgagccca ggaagaactt cgtgaagccc	2760
aacgagacca agacctactt ctggaaggtg cagcaccaca tggcccccac caaggacgag	2820
ttcgactgca aggctgggc ctacttctt gatgtggacc tggagaagga cgtcacacgc	2880
ggcctgatcg gccccctgct ggtgtgccac accaacaccc tgaaccccgcc cacggcagg	2940
caggtgaccg tgcaggagtt cgccctgttc ttcaccatct tcgacgagac caagagctgg	3000
tacttcacccg agaacatgga gaggaactgc agggccccc gcaacatcca gatggaggac	3060
cccaccttca aggagaacta caggttccac gccatcaacg gctacatcat ggacacctg	3120
cccggectgg ttagggccca ggaccagagg atcaggtgg atctgtgag catggcagc	3180
aacgagaaca tccacagcat ccacttcagc ggccacgtgt tcaccgtgag gaagaaggag	3240
gagtacaaga tggccctgta caacctgtac cccggcgtgt tcgagaccgt ggagatgctg	3300
cccgcaagg cggcatctg gagggtggag tgccctgatcg gogagcacct gcaegccggc	3360
atgagcaccc tggccctgggt gtacagcaac aagtggcaga ccccccctggg catggccagc	3420
ggccacatca gggacttcca gatcaccgc tctggccagt acggccagt ggccccaaag	3480
ctggccaggc tgcactacag cggcagcatc aacgcctgga gcaccaagga gcccctcage	3540
tggataagg tggacctgct ggcccccattg atcatccacg gcatcaagac ccagggcgcc	3600
aggcagaagt tcagcagcct gtacatcagc cagttcatca tcatgtacag cctggacggc	3660
aagaagtggc agacctacag gggcaacagc accggcaccc ttaggtgtt ctccggcaac	3720
gtggacagca gcggcatcaa gcacaacatc ttcaaccccc ccatcatcgc caggtacatc	3780
aggctgcacc ccacccacta cagcatcagg agcaccctgc ggatggaact gatggcgtgc	3840
gacctgaaca gctgcagcat gcccctgggc atggagagca aggccatctc tgacgcccag	3900
atcacccgcca gcagctactt caccaacatg ttccgcaccc ggagccccag caaggccagg	3960
ctgcacccctgc agggcaggag caaegcctgg aggccccagg tgaacaaccc caaggagtgg	4020
ctgcaggctgg acttccagaa gaccatgaag tgacccggcg tgaccaccca gggcgtgaag	4080
agcctgtga ccagcatgtt cgtgaaggag ttccctgatca gcagcagcca ggacggcac	4140
cagtggaccc tggcccttcca gaacggcaaa gtgaagggtt tccagggcaa ccaggacagc	4200
ttcacccccc tggtaaacag cctggacccc cccctgtca ccaggtatct gaggatccac	4260
ccccagagct gggtgacca gatcgcctg agaatggaag tgctggatg cgaggccag	4320
gacctgtact ga	4332

<210> SEQ ID NO 29
<211> LENGTH: 4368
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 29
atgcagattg agctgagcac ctgtttttc ctgtgcctgc tgagggtctg cttctctgcc 60
accaggagat actacctggg cgccgtggag ctgagctggg actacatgca gtctgacctg 120

-continued

ggcgagctgc ctgtggacgc caggttcccc cccagagtgc ccaagagctt ccccttcaac	180
acctcagtgg tgtacaagaa gaccctgttc gtggagttca ccgaccacct gttcaacatc	240
gccaagccca ggccccctg gatgggcctg ctggggccca ccatccaggc cgagggtgtac	300
gacaccegtgg tgatcacccct gaagaacatg gecagccacc ccgtgagcct gcacgcccgt	360
ggcgtgagct actggaaggc ctctgagggc gccgagtatg acgaccagac cagccagagg	420
gagaaggagg acgacaaggc gttcccccgc ggcagccaca cctacgttg gcaggtgctg	480
aaggagaacg gccccatggc cagcgacccc ctgtgcctga cctacagcta cctgagccac	540
gtggacactgg tgaaggacct gaactctggc ctgateggc ccctgtgtt gtgcaggag	600
ggcagcctgg ccaaggagaa gacccagacc ctgcacaagt tcatactgtt gtgcgcctg	660
ttcgatgagg gcaagagctg gcacagcgag accaagaaca gcctgtatgc ggacaggat	720
gcccgcctcg ccaggccctg gcccaagatg cacaccgtga acggctacgt gaacaggagc	780
ctgccccggcc tgatcggctg ccacaggaag tctgtgtact ggcacgtgtat cggcatggc	840
accacccccc aggtgcacag catttcctg gagggccaca cttctgtt gaggaaccac	900
aggcaggcca gcctggagat cagccccatc accttcctga cggcccgac cctgtgtat	960
gacctggggcc agttccctgtt gttctgccac atcagcagcc accagcacga cggcatggag	1020
gcctacgtga aggtggacag ctgccccgag gagccccagc tgaggatgaa gaacaacag	1080
gaggccgagg actatgtatg tgacctgacc gactctgaga tggacgttgtt gaggtttag	1140
gatgacaaca gccccagctt catccagatc aggtctgtgg ccaagaagca ccccaagacc	1200
tgggtgcaact acatcgccgc cgaggaggag gactgggact acggccccc ggtgtggcc	1260
cccgacgaca ggagctacaa gagccagttac ctgaacaacg gccccagag gatggcagg	1320
aagtacaaga aggtcgatatt catggcctac accgacgaga cttcaagac caggaggcc	1380
atccagcact agtctggcat cttggggcccc ctgtgtacg gcgagggtgg cgacaccctg	1440
ctgatcatct tcaagaacca ggccagcagg ccctacaaca tctacccca cggcatcacc	1500
gatgtgaggc ccctgtacag caggaggctg cccaaaggcc tgaagcacct gaaggacttc	1560
cccatcctgc cccggcagat cttcaagttac aagtggaccg tgaccgtgga ggtggccccc	1620
accaagtctg accccaggtg cctgaccagg tactacagca gttcgtgaa catggagagg	1680
gacctggctt ctggcctgtat cggcccccctg ctgtatctgtt acaaggagag cgtggaccag	1740
aggggcaacc agatcatgtc tgacaagagg aacgtgtatcc tttttctgtt gttcgatgag	1800
aacaggagct ggtatctgac cgagaacatc cagaggttcc tgcccaaccc cggccggctg	1860
cagctggagg accccaggtt ccaggccagc aacatcatgc acagcatcaa cggctacgt	1920
ttcgacagcc tgcagctgtc tgggtgcctg caccgggtgg ctactctgtt catctgtac	1980
atcggccccc agaccgactt cctgtctgtt ttcttctgtt gctacacctt caagcacaag	2040
atgggtgtacg aggacaccct gaccctgttc cccttcagcc gcgagaccgt gttcatgagc	2100
atggagaacc cccggcctgtg gatcctgggc tgccacaaca gcgacttcag gaacaggggc	2160
atgaccgccc tgctgaaaatg cagcagctgc gacaagaaca cccggcacta ctacgaggac	2220
agctacgagg acatcagcgc ctactctgtt agcaagaaca acgccatcga gcccaggagc	2280
ttcagccaga actccagaca ccccaagcacc agggagatca ccaggaccac cctgcagagc	2340
gaccaggagg agatcgacta tgatgacacc atcagcgtgg agatgaagaa ggaggacttc	2400

-continued

gacatctacg acgaggacga gaaccagagc cccaggagct tccagaagaa gaccaggcac	2460
tacctcatcg ccgcgttgg aaggctgtgg gactatggca tgagcagcag cccccacgtg	2520
ctgaggaaca gggcccagag cgccgcgtg ccccagtca agaaggtgtgt gttccaggag	2580
ttcacccgacg gcacgttac ccacccctgt tacagaggcg agctgaacga gcacctggc	2640
ctgctggcc cctacatcg ggccgagggtg gaggacaaca tcatggtgac cttcaggaac	2700
caggccgca ggcctacag cttctacagc agcctgtatca gctacgagga ggaccagagg	2760
caggggcgccg agcccaggaa gaacttcgtg aagcccaacg agaccaagac ctacttctgg	2820
aagggtgcagc accacatggc cccccaacg gacgagttc actgcaaggc ctggccctac	2880
ttctctgtatg tggaccttgg aaggacgtg cacagccggcc tgatggccc cctgtgttg	2940
tgccacacca acaccctgaa ccccccac ggcaggcagg tgaccgtgca ggagttcgcc	3000
ctgttcttca ccatcttcga cgagaccaag agctggactt tcacccgagaa catggagagg	3060
aactgcaggg cccctgcaa catccagatg gaggacccca cttcaagga gaactacagg	3120
ttccacggca tcaacggcta catcatggac accctgccc gctgtgtatg ggcccaggac	3180
cagaggatca ggtgttatct gctgagcatg ggcagcaacg agaacatcca cagcatccac	3240
ttcagccggcc acgtgttac cgtgaggaag aaggaggagt acaagatggc cctgtacaac	3300
ctgtaccccg gcgtgttgcg aaccgtggag atgctgccc gcaaggccgg catctggagg	3360
gtggagtggcc tgatcgccgca gcacccgtac gccggcatga gcacccgtt cctgtgtac	3420
agcaacaagt gccagacccca cctggccatg gccagccggc acatcaggaa cttccagatc	3480
accgcctctg gccagtaacgg ccagtggcc cccaaatggc ccaggctgca ctacagccgc	3540
agcatcaacg cctggagcac caaggagccc ttcaaggtgaa tcaaggtgg cctgtggcc	3600
cccatatgtca tccacggcat caagacccag ggccggccaggc agaaggatcg cagcctgtac	3660
atcagccatg tcatcatcat gtacagccctg gacggcaaga agtggcagac ctacaggggc	3720
aacagccacg gcacccctgat ggtgttcttc ggcaacgtgg acagcagccg catcaagcac	3780
aacatcttca acccccccattt catcgccagg tacatcaggc tgacccaccc ccaactacac	3840
atcaggagca ccctggggat ggaactgtatg ggctgcgacc tgaacagctg cagcatggcc	3900
ctggggatgg agagcaaggc catctctgac gcccagatca cccggccaggc ctacttcacc	3960
aacatgttgc ccacccgtgg ccccaagcaag gccaggctgc acctgcaggc caggagcaac	4020
gcctggggcc cccagggtgaa caaccccaag gagtggctgc aggtggactt ccagaagacc	4080
atgaaggatgaa cccggcgatgac caccggccggc gtgaagggc tgctgaccag catgtacgt	4140
aaggagttcc tgatcagcag cagccaggac ggccaccagg ggaccctgtt cttccagaac	4200
ggcaaaatgtaa aggtgttcca gggcaaccag gacagttca ccccccgttggt gaacagctcg	4260
gaccccccggc tgctgaccag gtatctgagg atccacccca agagctgggt gcaccagatc	4320
gcctgagaa tggaaatgtctt gggatgcgag gcccaggacc tgtactgaa	4368

```

<210> SEQ_ID NO 30
<211> LENGTH: 14
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      peptide

```

-continued

<400> SEQUENCE: 30

Ser Phe Ser Gln Asn Pro Pro Val Leu Lys Arg His Gln Arg
1 5 10

<210> SEQ ID NO 31

<211> LENGTH: 24

<212> TYPE: PRT

<213> ORGANISM: Sus sp.

<400> SEQUENCE: 31

Ser Phe Ala Gln Asn Ser Arg Pro Pro Ser Ala Ser Ala Pro Lys Pro
1 5 10 15

Pro Val Leu Arg Arg His Gln Arg
20

<210> SEQ ID NO 32

<211> LENGTH: 16

<212> TYPE: PRT

<213> ORGANISM: Sus sp.

<400> SEQUENCE: 32

Ser Phe Ser Gln Asn Ser Arg His Gln Ala Tyr Arg Tyr Arg Arg Gly
1 5 10 15

<210> SEQ ID NO 33

<211> LENGTH: 8

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide

<400> SEQUENCE: 33

Thr Thr Tyr Val Asn Arg Ser Leu
1 5

<210> SEQ ID NO 34

<211> LENGTH: 7

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 34

Pro Gln Leu Arg Met Lys Asn
1 5

<210> SEQ ID NO 35

<211> LENGTH: 7

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide

<400> SEQUENCE: 35

Val Asp Gln Arg Gly Asn Gln
1 5

<210> SEQ ID NO 36

<211> LENGTH: 87

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic

-continued

oligonucleotide

<400> SEQUENCE: 36

agcttcagcc agaatgtgag caacaatgtg agcaacaatg ccaccaataa tgctaccaac 60
ccacacctgtcc tgaaacgccca ccagagg 87

<210> SEQ ID NO 37

<211> LENGTH: 75

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 37

agcttcagcc agaatgtgag caacaatgcc accaacaatg tgagcaaccc acctgtcctg 60
aacacgccacc agagg 75

<210> SEQ ID NO 38

<211> LENGTH: 63

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 38

agcttcagcc agaatgtgag caataatgcc accaaccac ctgtcctgaa acgccaccag 60
agg 63

<210> SEQ ID NO 39

<211> LENGTH: 54

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 39

agcttcagcc agaatgtgag caataatcca cctgtcctga aacgccacca gagg 54

<210> SEQ ID NO 40

<211> LENGTH: 51

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 40

agcttcagcc agaataggag cctgccacct gtcctgaaac gccaccagag g 51

<210> SEQ ID NO 41

<211> LENGTH: 93

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 41

agcttcagcc agaatgccac taatgtgtct aacaactctg ctacacctgc tgactctgct 60

-continued

gtgagccac ctgtcctgaa acgccaccag agg 93

<210> SEQ ID NO 42
<211> LENGTH: 69
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 42

agcttcagcc agaatgccac caactatgtg aacaggagcc tgccacctgt cctgaaaacgc 60
caccagagg 69

<210> SEQ ID NO 43
<211> LENGTH: 105
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 43

agcttcagcc agaatgccac caactatgtg aacaggagcc tgtctgccac ctctgctgac 60
tctgctgtga gccagaatcc acctgtcctg aaacgccacc agagg 105

<210> SEQ ID NO 44
<211> LENGTH: 84
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 44

agcttcagcc agaatgtgag caacaatgtg agcaatgctg tgtctgctgt gtctgctcca 60
cctgtcctga aacgccacca gagg 84

<210> SEQ ID NO 45
<211> LENGTH: 90
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 45

agcttcagcc agaatatcac tgtggcctct gccacctcta acatcactgt ggcctctgct 60
gaccacacctg tcctgaaacg ccaccagagg 90

<210> SEQ ID NO 46
<211> LENGTH: 72
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400> SEQUENCE: 46

agcttcagcc agaatatcac tgtgaccaac atcactgtga ctgccccacc tgtcctgaaa 60

-continued

cggccaccaga gg	72
<210> SEQ ID NO 47	
<211> LENGTH: 72	
<212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide	
<400> SEQUENCE: 47	
agcttcagcc agaatcagac tgtgaccaac atcactgtga ctgccccacc tgtcctgaaa	60
cggccaccaga gg	72
<210> SEQ ID NO 48	
<211> LENGTH: 93	
<212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide	
<400> SEQUENCE: 48	
agcttcagcc agaatgccac taatgtgtct aacaacagca acaccagcaa tgacagcaat	60
gtgtctccac ctgtcctgaa acggcaccag agg	93
<210> SEQ ID NO 49	
<211> LENGTH: 4374	
<212> TYPE: DNA	
<213> ORGANISM: Artificial Sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide	
<400> SEQUENCE: 49	
atgcagattg agctgtccac ctgcttctt ctgtgcctgc tgagattctg cttctctgcc	60
accaggagat actacatggg ggctgtggaa ctttcttggg actacatgca gtctgacctg	120
ggagagatgc ctgtggatgc caggttccca cccagagtgc ccaagtcctt cccattcaac	180
acctctgtgg tctacaagaa gacactcttt gtgaaattca ctgaccacct gttcaacatt	240
gcaaaaaccca gaccaccctg gatgggactc ctgggaccca ccattcaggc tgaggtgtat	300
gacactgtgg tcataccct caagaacatg gcatcccacc ctgtgtctct gcatgctgtg	360
ggagtctcat actggaaagc ctctgaaggg gctgagtgatg atgaccagac atcccgagaga	420
gagaaagagg atgacaaggt gttccctggg ggatctcaca cctatgtgtg gcaagtctc	480
aaggagaatg gacccatggc atctgaccca ctctgcctga catactccctt ctttctcat	540
gtggacactgg tcaaggaccc caactctgga ctgattgggg cactgtgtt gtgcaggaa	600
ggatccctgg ccaaggagaa aacccagaca ctgcacaagt tcattctct gtttgctgtc	660
tttgatgagg gcaagtcttg gcactctgaa acaaagaact ccctgtatgca agacaggat	720
gctgcctctg ccagggcatg gcccaagatg cacactgtga atggctatgt gaacagatca	780
ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcattgtat tggcatgggg	840
acaacccctg aagtgcactc catttctgg gagggacaca ctttctgggt caggaaccac	900
agacaaggct ctctggagat ctctccatc accttctca ctgcacagac actgtgtatg	960
gaccctggac agttctgtct gtcctgccac atctcttccc accagcatga tggcatggaa	1020

-continued

gcctatgtca	aggtgttactc	atgccttagag	gaaccacagc	tcaggatgaa	gaacaatgag	1080	
gaggctgagg	actatgtatga	tgacctgact	gactctgaga	tggatgttgt	cagatttgat	1140	
gtatgacaact	ctccatcctt	cattcagatc	aggctgttgc	caaagaaaca	ccccaaagaca	1200	
tgggtgcact	acattgtgc	tgagaaagag	gactgggact	atgcaccact	ggctctggcc	1260	
cctgtatgaca	ggagctacaa	gtctcagttac	ctcaacaatg	gcccacaaag	aatttggaaaga	1320	
aagtacaaga	aagttagatt	catggcctac	actgtatgaaa	ccttcaagac	aagagaagcc	1380	
attcagcatg	agtctggcat	tctggacca	ctcctgtatg	ggaaagtgg	agacaccctg	1440	
ctcatcatct	tcaagaacca	ggcctccagg	ccctacaaca	tctaccacaca	tggcatcact	1500	
gatgtcaggc	ccctgtacag	caggagactg	ccaaaagggg	tgaaacacct	caaggacttc	1560	
ccattctgc	ctggagagat	cttcaagttac	aagtggactg	tcactgttgc	ggatggacca	1620	
acaaaagtctg	accccagggt	cctcaccaga	tactactctt	cttttgc	aaatggagaga	1680	
gacctggcat	ctggactgtat	tggaccactg	ctcatctgt	acaaggagtc	tgtggaccag	1740	
agaggcaacc	agatcatgtc	tgacaagaga	aatgtgattc	tgttctctgt	ctttgtatgag	1800	
aacagatcat	ggtacactgac	tgagaacatt	cagagattcc	tgcaccaaccc	tgctggggtg	1860	
caactggaag	accctgagtt	ccaggcaagc	aacatcatgc	actccatcaa	tggctatgt	1920	
tttgcacttc	tccagcttgc	tgtctgcctg	catgagggtg	cctactggta	cattttct	1980	
attggggcac	aaactgactt	ccttctgtc	ttcttctgt	gatacacctt	caagcacaag	2040	
atgggttatg	aggacaccct	gacactcttc	ccattctctg	gggaaactgt	gttcatgagc	2100	
atggagaacc	ctggactgtg	gattctggga	tgccacaact	ctgacttcag	aaacagggga	2160	
atgactgcac	tgctcaaagt	ctcctctgt	gacaagaaca	ctggggacta	ctatgaggac	2220	
tcttatgagg	acatctctgc	ctacatgttc	agcaagaaca	atgcccattga	gcccaagaac	2280	
ttctctcaga	atccacctgt	cctgaagaga	caccagagag	agatcaccag	gacaaccctc	2340	
cagtctgacc	aggaagagat	tgactatgtat	gacaccattt	ctgtggagat	gaagaaggag	2400	
gactttgaca	tctatgtatg	ggacgagaac	cagtctccaa	gatcattcca	gaagaagaca	2460	
agacactact	tcattgtgc	tgtggaaaga	ctgtggact	atggcatgtc	ttcctctccc	2520	
catgtcctca	ggaacagggc	acagtctggc	tctgtgccac	agttcaagaa	agtggcttc	2580	
caggagttca	ctgatggctc	attcacccag	cccctgtaca	gaggggaaact	gaatgagcac	2640	
ctgggactcc	tgggaccata	catcagggt	gagggtggaa	acaacatcat	ggtgacattc	2700	
agaaaccagg	cctccaggcc	ctacagcttc	tactcttccc	tcatcagct	tgaggaaagac	2760	
cagagacaag	gggctgagcc	aagaaagaac	tttgc	aaac	ccaatgaaac	caagacccat	2820
ttctggaaag	tccagcacca	catggcaccc	accaaggatg	agtttgactg	caaggccctgg	2880	
gcataacttct	ctgatgttgc	cctggagaaa	gatgtcact	ctggcctgtat	tggccactc	2940	
ctgggtctgc	acaccaacac	cctgaaccct	gcacatggaa	ggcaagtgc	tgtgcaggag	3000	
tttgcctct	tcttcaccat	ctttgtatgaa	accaagtcat	ggtacttcac	tgagaacatg	3060	
gagagaaaact	gcagagcacc	atgcaacatt	cagatggaa	accccacctt	caaggagaac	3120	
taacggatcc	atgcccataa	tggctacatc	atggacacc	tgcctgggt	tgtcatggca	3180	
caggaccaga	aatcagatg	gtacactgtt	tctatggat	ccaatgagaa	cattcactcc	3240	
atccacttct	ctgggcatgt	cttcaactgt	agaaagaagg	aggaatacaa	gatggccctg	3300	

-continued

tacaacctct accctgggtt ctttgagact gtggagatgc tgccctccaa agctggcattc	3360
tgggggtgg aatgcctcat tggggagcac ctgcattgtg gcatgtcaac cctgttctg	3420
gtctacagca acaagtgcac gacaccctgt ggaatggcct ctggccacat cagggacttc	3480
cagatcactg cctctggcca gtatggccag tgggcacccaa aactggccag gctccactac	3540
tctggctcca tcaatgcattg gtcaaccaag gagccattctt cttggatcaa ggtggactg	3600
ctggcacccaa tgatcattca tggcatcaag acacaggggg caagacagaa attctccctt	3660
ctgtacatct cacagttcat catcatgtac tctctggatg gcaagaagtgc gtagacatac	3720
agaggcaact ccactggcac cctcatggtc ttctttggca atgtggacag ctctggcattc	3780
aagcacaaca tcttcaaccc tcccatcatt gccagataca tcaggtgcac ccccacccac	3840
tactcaatca gatcaaccctt caggatggaa ctgatggatgt gtagacatgc ctcctgctca	3900
atggccctgg gaatggagag caaggccatt tctgatggcc agatcactgc atcctcttac	3960
ttcaccacaa tgggtccac ctggtcacca tcaaaagccca ggctgcaccc ccagggaaaga	4020
agcaatgcctt ggagacccca ggtcaacaac ccaaaggaaat ggctgcagg ggacttccag	4080
aagacaatgaa aagtcaactgg ggtgacaacc cagggggcata agtctctgtt cacctcaatg	4140
tatgtgaagg agttcctgtat ctcttcctca caggatggcc accagtggac actcttcttc	4200
cagaatggca aagtcaagggtt gttccaggcc accaggactt cttcacacc tgggtgaac	4260
tcactggacc ccccccttccat gacaagatata ctgagaatttcc acccccagtc ttgggtccac	4320
cagattgccctt tgagaatggaa agtcctggaa tgtgaggcac aagacatgtatcttgc	4374

```

<210> SEQ ID NO 50
<211> LENGTH: 45
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
oligonucleotide
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1) .. (45)

<400> SEQUENCE: 50

```

gtg agc aac aat gtg agc aac aat gcc acc aat aat gct acc aac	45
Val Ser Asn Asn Val Ser Asn Asn Ala Thr Asn Asn Ala Thr Asn	
1 5 10 15	

```

<210> SEQ ID NO 51
<211> LENGTH: 15
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
peptide

<400> SEQUENCE: 51

```

Val Ser Asn Asn Val Ser Asn Asn Ala Thr Asn Asn Ala Thr Asn	
1 5 10 15	

```

<210> SEQ ID NO 52
<211> LENGTH: 33
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic

```

-continued

oligonucleotide

<220> FEATURE:

<221> NAME/KEY: CDS

<222> LOCATION: (1)..(33)

<400> SEQUENCE: 52

gtg agc aac aat gcc acc aac aat gtg agc aac
Val Ser Asn Asn Ala Thr Asn Asn Val Ser Asn
1 5 10

33

<210> SEQ ID NO 53

<211> LENGTH: 11

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide

<400> SEQUENCE: 53

Val Ser Asn Asn Ala Thr Asn Asn Val Ser Asn
1 5 10

<210> SEQ ID NO 54

<211> LENGTH: 21

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<220> FEATURE:

<221> NAME/KEY: CDS

<222> LOCATION: (1)..(21)

<400> SEQUENCE: 54

gtg agc aat aat gcc acc aac
Val Ser Asn Asn Ala Thr Asn
1 5

21

<210> SEQ ID NO 55

<211> LENGTH: 7

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide

<400> SEQUENCE: 55

Val Ser Asn Asn Ala Thr Asn
1 5

<210> SEQ ID NO 56

<211> LENGTH: 12

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<220> FEATURE:

<221> NAME/KEY: CDS

<222> LOCATION: (1)..(12)

<400> SEQUENCE: 56

gtg agc aat aat
Val Ser Asn Asn
1

12

-continued

```
<210> SEQ ID NO 57
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      peptide
```

Vol. 2000, No. 2000

1

```
<210> SEQ ID NO: 58
<211> LENGTH: 9
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
          oligonucleotide
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1) .. (9)

<400> SEQUENCE: 58
```

agg agc ctg

10

agg agc ctg
Arg Ser Leu
1

```
<210> SEQ_ID: NC_00  
<211> LENGTH: 3  
<212> TYPE: PRT  
<213> ORGANISM: Artificial Sequence  
<220> FEATURE:  
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
peptide
```

<400> SEQUENCE: 59

Arg Ser Leu
1

```
<210> SEQ ID NO 60
<211> LENGTH: 51
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      oligonucleotide
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)..(51)

 100  SEQUENCE  60
```

```
gcc actaatgttctaacaacgtctaccgtctgacgtctgtg 48
Ala Thr Asn Val Ser Asn Asn Ser Ala Thr Ser Ala Asp Ser Ala Val
1 5 10 15
```

agc 51

<210> SEQ ID NO 61
<211> LENGTH: 17
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide

<400> SEQUENCE: 61

-continued

```

Ala Thr Asn Val Ser Asn Asn Ser Ala Thr Ser Ala Asp Ser Ala Val
1          5           10          15

```

Ser

```
<210> SEQ ID NO 62
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      oligonucleotide
<220> FEATURE:
<221> NAME/KEY: CDS
```

```
gcc acc aac tat gtg aac agg agc ctg  
Ala Thr Asn Tyr Val Asn Arg Ser Leu  
1           5
```

27

<210> SEQ ID NO 63

<211> LENGTH: 9

<400> SEQUENCE: 63

Ala Thr Asn Tyr Val Asn Arg Ser Leu
1 5

<210> SEQ ID NO 64

```
<211> LENGTH: 63
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description
          oligonucleotide
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1) .. (63)
```

<400> SEQUENCE: 64

```

gcc acc aac tat gtg aac agg agc ctg tct gcc acc tct gct gac tct
Ala Thr Asn Tyr Val Asn Arg Ser Leu Ser Ala Thr Ser Ala Asp Ser
1           5                   10                  15

```

48

```

gct gtg agc cag aat
Ala Val Ser Gln Asn
                         20

```

63

<210> SEQ ID NO 65

```
<211> LENGTH: 21
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description
      peptide
```

<400> SEQUENCE: 65

```

Ala Thr Asn Tyr Val Asn Arg Ser Leu Ser Ala Thr Ser Ala Asp Ser
1           5           10          15

```

Ala Val Ser Gln Asn
20

-continued

```
<210> SEQ ID NO 66
<211> LENGTH: 42
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      oligonucleotide
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1) .. (42)

<400> SEQUENCE: 66
```

gtg agc aac aat gtg agc aat gct gtg tct gct gtg tct gct	42
Val Ser Asn Asn Val Ser Asn Ala Val Ser Ala Val Ser Ala	
1	5
	10

```
<210> SEQ ID NO 67
<211> LENGTH: 14
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      peptide
<400> SEQUENCE: 67
```

Val Ser Asn Asn Val Ser Asn Ala Val Ser Ala Val Ser Ala	
1	5
	10

```
<210> SEQ ID NO 68
<211> LENGTH: 48
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      oligonucleotide
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1) .. (48)

<400> SEQUENCE: 68
```

atc act gtg gcc tct gcc acc tct aac atc act gtg gcc tct gct gac	48
Ile Thr Val Ala Ser Ala Thr Ser Asn Ile Thr Val Ala Ser Ala Asp	
1	5
	10
	15

```
<210> SEQ ID NO 69
<211> LENGTH: 16
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      peptide
<400> SEQUENCE: 69
```

Ile Thr Val Ala Ser Ala Thr Ser Asn Ile Thr Val Ala Ser Ala Asp	
1	5
	10
	15

```
<210> SEQ ID NO 70
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      oligonucleotide
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1) .. (30)
```

-continued

<400> SEQUENCE: 70

atc act gtg acc aac atc act gtg act gcc 30
Ile Thr Val Thr Asn Ile Thr Val Thr Ala
1 5 10

<210> SEQ ID NO 71

<211> LENGTH: 10
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide

<400> SEQUENCE: 71

Ile Thr Val Thr Asn Ile Thr Val Thr Ala
1 5 10

<210> SEQ ID NO 72

<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)..(30)

<400> SEQUENCE: 72

cag act gtg acc aac atc act gtg act gcc 30
Gln Thr Val Thr Asn Ile Thr Val Thr Ala
1 5 10

<210> SEQ ID NO 73

<211> LENGTH: 10
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide

<400> SEQUENCE: 73

Gln Thr Val Thr Asn Ile Thr Val Thr Ala
1 5 10

<210> SEQ ID NO 74

<211> LENGTH: 51
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)..(51)

<400> SEQUENCE: 74

gcc act aat gtg tct aac aac agc aac acc agc aat gac agc aat gtg 48
Ala Thr Asn Val Ser Asn Asn Ser Asn Thr Ser Asn Asp Ser Asn Val
1 5 10 15

tct 51
Ser

<210> SEQ ID NO 75

<211> LENGTH: 17
<212> TYPE: PRT

-continued

<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide
<400> SEQUENCE: 75

Ala Thr Asn Val Ser Asn Asn Ser Asn Thr Ser Asn Asp Ser Asn Val
1 5 10 15

Ser

<210> SEQ ID NO 76
<211> LENGTH: 405
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 76

Met Pro Leu Leu Leu Tyr Thr Cys Leu Leu Trp Leu Pro Thr Ser Gly
1 5 10 15

Leu Trp Thr Val Gln Ala Met Asp Pro Asn Ala Ala Tyr Val Asn Met
20 25 30

Ser Asn His His Arg Gly Leu Ala Ser Ala Asn Val Asp Phe Ala Phe
35 40 45

Ser Leu Tyr Lys His Leu Val Ala Leu Ser Pro Lys Lys Asn Ile Phe
50 55 60

Ile Ser Pro Val Ser Ile Ser Met Ala Leu Ala Met Leu Ser Leu Gly
65 70 75 80

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

Leu Thr Glu Arg Ser Glu Thr Glu Ile His Gln Gly Phe Gln His Leu
100 105 110

His Gln Leu Phe Ala Lys Ser Asp Thr Ser Leu Glu Met Thr Met Gly
115 120 125

Asn Ala Leu Phe Leu Asp Gly Ser Leu Glu Leu Leu Glu Ser Phe Ser
130 135 140

Ala Asp Ile Lys His Tyr Tyr Glu Ser Glu Val Leu Ala Met Asn Phe
145 150 155 160

Gln Asp Trp Ala Thr Ala Ser Arg Gln Ile Asn Ser Tyr Val Lys Asn
165 170 175

Lys Thr Gln Gly Lys Ile Val Asp Leu Phe Ser Gly Leu Asp Ser Pro
180 185 190

Ala Ile Leu Val Leu Val Asn Tyr Ile Phe Phe Lys Gly Thr Trp Thr
195 200 205

Gln Pro Phe Asp Leu Ala Ser Thr Arg Glu Glu Asn Phe Tyr Val Asp
210 215 220

Glu Thr Thr Val Val Lys Val Pro Met Met Leu Gln Ser Ser Thr Ile
225 230 235 240

Ser Tyr Leu His Asp Ser Glu Leu Pro Cys Gln Leu Val Gln Met Asn
245 250 255

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

Met Asn Thr Val Ile Ala Ala Leu Ser Arg Asp Thr Ile Asn Arg Trp
275 280 285

Ser Ala Gly Leu Thr Ser Ser Gln Val Asp Leu Tyr Ile Pro Lys Val
290 295 300

-continued

Thr Ile Ser Gly Val Tyr Asp Leu Gly Asp Val Leu Glu Glu Met Gly
305 310 315 320

Ile Ala Asp Leu Phe Thr Asn Gln Ala Asn Phe Ser Arg Ile Thr Gln
325 330 335

Asp Ala Gln Leu Lys Ser Ser Lys Val Val His Lys Ala Val Leu Gln
340 345 350

Leu Asn Glu Glu Gly Val Asp Thr Ala Gly Ser Thr Gly Val Thr Leu
355 360 365

Asn Leu Thr Ser Lys Pro Ile Ile Leu Arg Phe Asn Gln Pro Phe Ile
370 375 380

Ile Met Ile Phe Asp His Phe Thr Trp Ser Ser Leu Phe Leu Ala Arg
385 390 395 400

Val Met Asn Pro Val
405

<210> SEQ ID NO 77

<211> LENGTH: 4

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 77

Asn Met Ser Asn

1

<210> SEQ ID NO 78

<211> LENGTH: 4

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 78

Asn Leu Thr Glu

1

<210> SEQ ID NO 79

<211> LENGTH: 4

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 79

Asn Lys Thr Gln

1

<210> SEQ ID NO 80

<211> LENGTH: 4

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 80

Asn Gly Thr Val

1

<210> SEQ ID NO 81

<211> LENGTH: 4

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 81

Asn Phe Ser Arg

1

-continued

<210> SEQ ID NO 82
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 82

Asn Leu Thr Ser
1

<210> SEQ ID NO 83
<211> LENGTH: 41
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polypeptide

<400> SEQUENCE: 83

Leu Ser Lys Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser Gln Asn Ala
1 5 10 15

Thr Asn Val Ser Asn Asn Ser Asn Thr Ser Asn Asp Ser Asn Val Ser
20 25 30

Pro Pro Val Leu Lys Arg His Gln Arg
35 40

<210> SEQ ID NO 84
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide

<400> SEQUENCE: 84

Asn Ala Thr Asn
1

<210> SEQ ID NO 85
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide

<400> SEQUENCE: 85

Asn Val Ser Asn
1

<210> SEQ ID NO 86
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide

<400> SEQUENCE: 86

Asn Asn Ser Asn
1

<210> SEQ ID NO 87

-continued

<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide

<400> SEQUENCE: 87

Asn Thr Ser Asn
1

<210> SEQ ID NO 88
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide

<400> SEQUENCE: 88

Asn Asp Ser Asn
1

<210> SEQ ID NO 89
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide

<400> SEQUENCE: 89

Asn Val Ser Pro
1

<210> SEQ ID NO 90
<211> LENGTH: 4377
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 90

atgcagattg	agctgtccac	ctgcttcttt	ctgtgcctgc	tgagattctg	cttctctgcc	60
accaggagat	actacctggg	ggctgtggaa	ctttcttggg	actacatgca	gtctgacctg	120
ggagagctgc	ctgtggatgc	caggttccca	cccagagtgc	ccaagtccctt	cccattcaac	180
acctctgtgg	tctacaagaa	gacactcttt	gtgaaattca	ctgaccacct	gttcaacatt	240
gaaaaaccca	gaccaccctg	gatgggactc	ctgggaccca	ccattcaggc	tgaggtgtat	300
gacactgtgg	tcatccccct	caagaacatg	gcatcccacc	ctgtgtctct	gcatgctgtg	360
ggagtctcat	actggaaagc	ctctgaaggg	gctgagttatg	atgaccagac	atcccaagaga	420
gagaaagagg	atgacaaggt	gttccctggg	ggatctcaca	cctatgtgtg	gcaagtctc	480
aaggagaatg	gaccatggc	atctgaccca	ctctgcctga	catactccct	cctttctcat	540
gtggacctgg	tcaaggacct	caactctgga	ctgattgggg	cactgctgg	gtgcaggaa	600
ggatccctgg	ccaaggagaa	aaccaggaca	ctgcacaagt	tcattctcct	gtttgctgtc	660
tttgcgttgg	gcaagtcttg	gcactctgaa	acaaagaact	ccctgatgca	agacaggat	720
gtgcctctg	ccagggcatg	gcccaagatg	cacactgtga	atggctatgt	gaacagatca	780

-continued

ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcatgtgat tggcatgggg	840
acaacccctg aagtgcactc catttcctg gagggcacaca ctttccttgtt caggaaccac	900
agacaaggct ctctggagat ctctccatc accttcctca ctgcacagac actgtgtatg	960
gaccttggac agttctgtct gtctgtccac atcttctccc accagcatga tggcatggaa	1020
gcctatgtca aggtggactc atgcctgttag gaaccacagc tcaggatgaa gaacaatgag	1080
gaggctgagg actatgtatga tgacctgact gactctgaga tggatgttgtt cagatttgat	1140
gatgacaact ctccatcctt cattcagatc aggtctgtgg caaagaaaca ccccaagaca	1200
tgggtgact acattgtctgc tgaggaagag gactgggact atgcaccact ggtctggcc	1260
cctgatgaca ggagctacaa gtctcagtagt ctcaacaatg gcccacaaag aattggaaga	1320
aagtacaaga aagtcaagatt catggcctac actgtatgaaa cttcaagac aagagaagcc	1380
attcagcatg agtctggcat tctgggacca ctccctgtatg gggaaagtggg agacaccctg	1440
ctcatcatct tcaagaacca ggcctccagg ccctacaaca tctaccacaca tggcatcact	1500
gatgtcaggc ccctgtacag caggagactg cccaaagggg taaaaacaccc caaggactc	1560
cccattctgc ctggagagat cttcaagtagc aagtggactg tcaactgtgga ggtatggacca	1620
acaaaagtctg accccaggtg cctcaccaga tactactcct cttttgtgaa catggagaga	1680
gacctggcat ctggactgtatg tggaccactg ctcatctgtc acaaggagtc tggaccacag	1740
agaggcaacc agatcatgtc tgacaagaga aatgtgatcc tggatgttgtt ctttgcgtat	1800
aacagatcat ggtacctgac tgagaacatt cagagattcc tgcccaaccc tgctgggtg	1860
caactggaaag accctgagtt ccaggcaago aacatcatgc actccatcaa tggctatgt	1920
tttgcgtatc tccagtttc tggatgttgtt ctttgcgtatg ctttgcgtatg ctttgcgtat	1980
atggggcact aaactgactt ctttgcgtatc ttttgcgtatg gatacacctt caagcacaag	2040
atgggttatg aggacaccct gacactcttc ccattctctg gggaaactgtt gttcatgagc	2100
atggagaacc ctggactgtg gattctggga tgccacaact ctgacttcag aaacaggggaa	2160
atgactgcac tgctcaaagt ctctcctgt gacaagaaca ctggggacta ctatgaggac	2220
tcttatgagg acatctctgc ctacatgtc agcaagaaca ataccaccta cgtgaaccgc	2280
tccctgtctc agaatccacc tggatgttgtt agacaccaga gagagatcac caggacaacc	2340
ctccagtctg accaggaaga gattgactat gatgacacca tttctgtgga gatgaagaag	2400
gaggactttg acatctatga tgaggacgag aaccagtctc caagatcatt ccagaagaag	2460
acaagacact acttcatgtc tgctgtggaa agactgtggg actatggcat gtcttcctct	2520
ccccatgtcc ttaggaacag ggcacagtct ggtctgtgc cacagttcaa gaaagtggc	2580
ttccaggagt tcaatgtatgg ctcattcacc cagccctgtt acagaggggaa actgaatgag	2640
cacctgggac tcctgggacc atacatcagg gctgaggatgg aagacaacat catggatgac	2700
ttcagaaacc aggccctccag gccctacagc ttctactctt ccctcatcag ctatgaggaa	2760
gaccagagac aaggggctga gccaagaaag aactttgtga aacccaaatga aaccaagacc	2820
tacttctggaa aagtccagca ccacatggca cccaccaagg atgagttgtt gatgtgtatg	2880
tgggcataact tctctgtatgtt ggacctggag aaagatgtgc actctggcct gatggccca	2940
ctccctggatc gccacaccaa caccctgaac cctgcacatg gaaggcaagt gactgtgcag	3000
gagtttgcctt tcttcttcac catctttgtat gaaaccaagt catggatctt cactgagaac	3060

-continued

atggagagaa	actgcagagc	accatgcaac	attcagatgg	aagacccac	cttcaaggag	3120
aactacagg	tccatgccat	caatggctac	atcatggaca	ccctgectgg	gcttgtcatg	3180
gcacaggacc	agagaatcag	atggcacctg	ctttctatgg	gatccaatga	gaacattcac	3240
tccatccact	tctctggca	tgtcttcact	tgagaaaaga	aggaggaaata	caagatggcc	3300
ctgtacaacc	tctaccctgg	ggtcttttag	actgtggaga	tgctgcctc	caaagctggc	3360
atctggagg	tggaatgcct	cattggggag	cacctgcatg	ctggcatgtc	aaccctgttc	3420
ctggcttaca	gcaacaagt	ccagacaccc	ctgggaatgg	cctctggcca	catcaggac	3480
ttccagatca	ctgcctctgg	ccagtatggc	cagtggcac	ccaaactggc	caggctccac	3540
tactctggct	ccatcaatgc	atggtcaacc	aaggagccat	tctcttggat	caagggtggac	3600
ctgctggcac	ccatgatcat	tcatggcatc	aagacacagg	gggcaagaca	gaaattctcc	3660
tctctgtaca	tctcacagtt	catcatcatg	tactctctgg	atggcaagaa	gtggcagaca	3720
taaaaaaaaaaaaa	actccactgg	caccctcatg	gtcttctttg	gcaatgtgga	cagctctggc	3780
atcaaggaca	acatttcaa	ccctccccatc	attgccagat	acatcaggct	gcacccacc	3840
cactactcaa	tcaaatcaac	cctcaggatg	gaactgtatgg	gatgtgacct	gaactcctgc	3900
tcaatgcccc	tggaaatgga	gagcaaggcc	atttctgtatg	cccagatcac	tgcacccct	3960
tacttcacca	acatgtttgc	cacctggtca	ccataaaaa	ccaggctgca	cctccaggga	4020
agaagcaatg	cctggagacc	ccaggtcaac	aacccaaagg	aatggctgca	agtggacttc	4080
cagaagacaa	tgaaagtca	tgggtgaca	accagggggg	tcaagtctct	gctcacctca	4140
atgtatgtga	aggagttcct	gatctttcc	tcacaggatg	gccaccagtg	gacacttttc	4200
ttccagaatg	gcaaaagtcaa	ggtgttccag	ggcaaccagg	actctttcac	acctgtggtg	4260
aactcaactgg	accccccct	cctgacaaga	tacctgagaa	ttcacccca	gtcttgggtc	4320
caccagattg	ccctgagaat	ggaagtccctg	ggatgtgagg	cacaagacct	gtactga	4377

<210> SEQ ID NO 91
<211> LENGTH: 4377
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 91

atgcagattg	agctgtccac	ctgtttcttt	ctgtgcctgc	tgagattctg	cttctctgccc	60
accaggagat	actacctggg	ggctgtggaa	ctttcttggg	actacatgca	gtctgacactg	120
ggagagatgc	ctgtggatgc	caggttccca	cccagagtgc	ccaagtcctt	cccattcaac	180
acctctgtgg	tctacaagaa	gacactcttt	gtggaaattca	ctgaccacat	gttcaacatt	240
gcaaaaacca	gaccaccctg	gatgggactc	ctgggaccca	ccattcaggc	tgaggtgtat	300
gacactgtgg	tcgtcaccc	caagaacatg	gcatcccacc	ctgtgtctct	gcatgtgtg	360
ggaggtctcat	actggaaatc	ctctgaaggg	gctgagttatg	atgaccagac	atcccgagaa	420
gagaaagagg	atgacaaggt	gttccctggg	aagtctcaca	cctatgtgtg	gcaagtctcc	480
aaggagaatg	gaccactgc	atctgaccca	ccctgcctga	catactccct	cctttctcat	540
gtggacactgg	tcaaggacct	caactctgga	ctgattgggg	cactgtgtgt	gtgcaggaa	600

-continued

ggatccctgg ccaaggagaa aacccagaca ctgcacaagt tcatttcctt gtttgcgtc 660
tttgcgagg gcaagtcttgcactctgaa acaaagaact ccctgtatgc agacaggat 720
gtgcctctg ccagggcatg gcccagatg cacactgtga atggctatgt gaacagatca 780
ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcatgtat tggcatgggg 840
acaaccctg aagtgcactc catttcctg gagggacaca ctttcctggc caggaaccac 900
agacaaggct ctctggagat ctctccatc accttcctca ctgcacagac actgtgtatg 960
gaccttggac agttctgtct gttctgcac atctcttccc accagcatga tggcatggaa 1020
gcctatgtca aggtggactc atgcctgttag gaaccacagc tcaggatgaa gaacaatgag 1080
gaggctgagg actatgtatg tgacactgact gactctgaga tggatgtggc cagatttgat 1140
gtatgacaact ctccatccctt cattcagatc aggtctgtgg caaagaaaaca ccccaagaca 1200
tgggtgcact acattgtgc tgaggaagag gactggact atgcaccaact ggtctggcc 1260
cctgtatgaca ggagctacaa gtctcagtagt ctcacaatg gcccacaaag aatttggaa 1320
aagtacaaga aagttagatt catggctac actgtatgaaa cttcaagac aagagaagcc 1380
attcagatc agtctggcat tctggacca ctctgtatg gggaaatggg agacaccctg 1440
ctcatcatct tcaagaacca ggcctccagg ccctacaaca tctaccacca tggcatcaact 1500
gatgtcaggc ccctgtacag caggagactg cccaaagggg tgaaacaccc caaggacttc 1560
ccattctgc ctggagagat cttcaagtagt aagtggactg tcactgtgg ggttggacca 1620
acaaagtctg accccagggtg cctcaccaga tactactctt cttttgtgaa catggagaga 1680
gacctggcat ctggactgat tggaccactg ctcatctgtc acaaggagtc tgtggaccag 1740
agaggcaacc agatcatgtc tgacaagaga aatgtgattc tggtctctgt ctgtatgag 1800
aacagatcat ggtacactgac tgagaacatt cagagattcc tgcccaaccc tgctgggtg 1860
caacttggaa accctgagtt ccaggcaacg aacatcatgc actccatcaa tggctatgt 1920
tttgcgtctc tccagtttc tgcgtgcctg catgaggatgg cctactggta cattttctt 1980
atggggcac aaactgtactt ctttctgtc ttcttctctg gatacacctt caagcacaag 2040
atgggtgtatg aggacaccct gacactttc ccattctgtc gggaaactgt gttcatgac 2100
atggagaacc ctggactgtg gattctgggta tgccacaact ctgacttcgaa aacagggg 2160
atgactgcac tgctcaaagt ctccctctgt gacaagaaca ctggggacta ctatgaggac 2220
tcttatgagg acatctctgc ctacactgtc agcaagaaca ataccaccta cgtgaaccgc 2280
tccctgtctc agaatccacc tgcgtgttgg agacaccaga gagagatcac caggacaacc 2340
ctccagttcg accaggaaga gattgactat gatgacacca ttctgtggat gatgaagaag 2400
gaggactttg acatctatgtc tgaggacgag aaccagtctc caagatcatt ccagaagaag 2460
acaagacact acttcatttc tgcgtgtggaa agactgtggg actatggcat gtcttcctt 2520
ccccatgtcc tcagggacac ggcacagtct ggctctgtc cacatgttcaaa gaaatgggtc 2580
ttccaggatgttgcactgtatgg ctcatttacc cagccctgtt acagaggggaa actgaatgag 2640
caccttggac tccttgggacc atacatcagg gctgaggatgg aagacaacat catggtgaca 2700
ttcagaaacc aggcctccag gcccatacgc ttctacttcc cctcatcag ctatgaggaa 2760
gaccagagac aaggggctgaa gccaagaag aactttgtgaa aacccatgaa aaccaagacc 2820
tacttctggaa aagtccagca ccacatggca cccaccaagg atqagttgtgaa ctgcaaggcc 2880

-continued

tgggcatact tctctgtatgt ggacactggag aaagatgtgc actctggcct gattggccca	2940
ctcctggctt gccacaccaa caccctgaac cctgcacatg gaaggcaagt gactgtgcag	3000
gagtttgcctt ctcttcac catcttcatgaa gaaaaccaagt catggactt cactgagaac	3060
atggagagaa actgcagagc accatgcaac attcagatgg aagaccccac cttcaaggag	3120
aactacaggt tccatgccat caatggctac atcatggaca ccctgectgg gcttgcatg	3180
geacaggacc agagaatcg atggtacctg ctttctatgg gatccaatga gaacattcac	3240
tcacatccact tctctggca tgcgttcaact gtgagaaaga aggaggaaata caagatggcc	3300
ctgtacaacc tctaccctgg ggtcttttagt actgtggaga tgctgcctc caaagctggc	3360
atctgggggg tggaaatgcct cattggggag cacctgcatg ctggcatgtc aaccctgttc	3420
ctggcttaca gcaacaatgtt ccagacaccc ctggaaatgg cctctggccca catcaggac	3480
ttccagatca ctgcctctgg ccagtatggc cagtgccac cc当地actggc caggctccac	3540
tactctggctt ccatcaatgc atggtaacc aaggagccat tctcttgat caaggtggac	3600
ctgctggcac ccatgtatcat tcatggcattt aagacacagg gggcaagaca gaaattctcc	3660
tctctgtaca tctcacatgtt catcatcatg tactctctgg atggcaagaa gtggcagaca	3720
tacagaggca actccactgg caccctcatg gtcttctttt gcaatgtggc cagctctggc	3780
atcaagcaca acatcttcaa ccctccatc attgccatg acatcaggct gcacccacc	3840
cactactcaa tcaatcaac cctcaggatg gaactgtatgg gatgtgaccc gaaactctgc	3900
tcaatgcctt tggaaatggc gageaaggcc atttctgtatg cccagatcac tgcatccct	3960
tacttcacca acatgtttgc cacctggta ccatcaaaa ccaggctgca cctccaggga	4020
agaagcaatg cctggagacc ccaggtcaac aacccaaagg aatggctgca agtggacttc	4080
cagaagacaa tgaaagtac tgggtgaca acccaggggg tcaagtctct gtcacccat	4140
atgtatgtga aggagttctt gatctcttc tcacaggatg gccaccatgt gacactcttc	4200
ttccagaatg gcaaaatgaa ggtgttccag ggcaaccagg actctttcac acctgtggc	4260
aactcactgg accccccccctt cctgacaaga tacctgagaa ttccacccca gtcttggc	4320
caccagattt ccctgagaat ggaagtccctg ggatgtgagg cacaagacccctt gtactga	4377

<210> SEQ ID NO 92
 <211> LENGTH: 4377
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 polynucleotide

<400> SEQUENCE: 92

atgcagattt agctgtccac ctgtttctttt ctgtgcctgc tgagattctg cttctctg	60
accaggagat actacactggg ggctgtggaa ctttcttggg actacatgca gtctgaccc	120
ggagagatgc ctgtggatgc caggatccca cccagatgc ccaagtcctt cccattcaac	180
acctctgtgg tctacaagaa gacactctttt gtggattca ctgaccaccc gttcaacatt	240
gcaaaaacca gaccaccctg gatggactc ctgggaccca ccattcaggc tgaggtgtat	300
gacactgtgg tcatcacccctt caagaacatg gcatccacc ctgtgtctct gcatgtgtg	360
ggagatctcat actggaaagc ctctgaaggg gctgagatgt atgaccagac atcccaagaga	420
gagaaagagg atgacaaggat gttccctggg ggatctcaca cctatgtgtg gcaagtctc	480

-continued

aaggagaatg gacccatggc atctgaccca ctctgcctga cataactccta cctttctcat	540
gtggacctgg tcaaggaccc caactctgga ctgattgggg cactgttgtt gtgcagggaa	600
ggatccctgg ccaaggagaa aaccaggaca ctgcacaagt tcattctct gtttgctgtc	660
tttgatgagg gcaagtcttg gcactctgaa acaaagaact ccctgtatgc agacaggat	720
gtgcctctg ccagggcatg gcccaagatg cacactgtga atggctatgt gaacagatca	780
ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcatgtat tggcatgggg	840
acaacccctg aagtgcactc catttcctg gagggacaca ctttccttgtt caggaaccac	900
agacaaggct ctctggagat ctctccatc accttcctca ctgcacagac actgtgtat	960
gaccttggac agttctgtct gttctgccac atctcttccc accagcatga tggcatggaa	1020
gcctatgtca aggtggactc atgcctgttag gaaccacagc tcaggatgaa gaacaatgag	1080
gaggctgagg actatgtatgc tgacccatgact gactctgaga tggatgttgtt cagatttgat	1140
gtgacaact ctccatcctt cattcagatc aggtctgttg caaagaaaca ccccaagaca	1200
tgggtgcact acattgtctc tgaggaagag gactggactc atgcaccact ggtctggcc	1260
cctgatgaca ggagctacaa gtctcagatc ctcaacaatg gcccacaaag aattggaaga	1320
aagtacaaga aagttagatt catggcctac actgtatgaaa cttcaagac aagagaagcc	1380
attcagcatg agtctggcat tctggacca ctccctgtatg gggaaatgttgg agacaccctg	1440
ctccatcatct tcaagaacca ggcctccagg ccctacaaca tctaccacca tggcatca	1500
gatgtcaggc ccctgtacag caggagactg cccaaagggg tgaaacacact caaggactc	1560
ccatttcgtc ctggagagat cttcaagatc aagtggactg tcactgtgaa ggtggacca	1620
acaaagtctg accccaggtg cctcaccaga tactactctt cttttgtgaa catggagaga	1680
gacctggcat ctggactgtat tggaccactg ctcatctgtc acaaggatgc tggaccacag	1740
agaggcaacc agatcatgtc tgacaagaga aatgtgattc tggatgttgtt ctttgtatg	1800
aacagatcat ggtacccatgac tgagaacatt cagagattcc tgcccaaccc tgctgggtg	1860
caactggaaag accctgagtt ccaggcaagc aacatcatgc actccatcaa tggctatgt	1920
tttgactctc tccagtttc tggatgttgtt catggatgg cctactggta cattttct	1980
atggggcac aaactgactt ctttcctgtc ttcttcctgt gatacacctt caagcacaag	2040
atggatgtatg aggacaccct gacactcttc ccattctctg gggaaactgtt gttcatgagc	2100
atggagaacc ctggactgtg gattctggaa tgccacaact ctgacttcag aaacaggggaa	2160
atgactgcac tgctcaaatg ctccctctgt gacaagaaca ctggggacta ctatgaggac	2220
tcttatgagg acatctctgc ctacactgtc agcaagaaca ataccacca cgtgaaccgc	2280
tcctctgtc tggatccacc tggatgttgtt agacaccaga gagatgttcc caggacaacc	2340
ctccagttcg accaggaaga gattgactat gatgacacca tttctgtgaa gatgaagaag	2400
gaggactttg acatctatgtc tgaggacccg aaccactctc caagatcatt ccagaagaag	2460
acaagacact acttcattgc tgctgtggaa agactgttgtt actatggcat gtcttcctct	2520
ccccatgtcc tggatccacc ggcacactt ggctctgtgc cacagttcaa gaaatggc	2580
ttccaggatg tggatgttgtt ctcatttcacc cagccctgtt acagagggaa actgtatgag	2640
cacctggac tggatgttgtt atacatcagg gctgaggatgg aagacaacat catggatgaca	2700
ttcagaaacc aggccctccag ggcctacacc ttctactctt ccctcatcag ctatgaggaa	2760

-continued

gaccagagac aaggggctga gccaagaag aacttgta aaccaatga aaccaagacc	2820
tacctctgga aagtccagca ccacatggca cccaccaagg atgagttgtc ctgcaggcc	2880
tgggcatact tctctgtatgt ggacctggag aaagatgtgc actctggcct gattggccca	2940
c当地tggctt gccacaccaa caccctgtac cctgcacatg gaaggcaagt gactgtgcag	3000
gagtttgcct tcttcttcac catcttgcat gaaaccaagt catggtaactt cactgagaac	3060
atggagagaa actgcagagc accatgcac accatgcac attcagatgg aagaccccac cttcaaggag	3120
aactacaggt tccatgccat caatggctac atcatggaca ccctgcctgg gcttgcata	3180
gcacaggacc agagaatcgat atggtaatcgat ctttctatgg gatccaatga gaacattcac	3240
tccatccact tctctggca tgtcttcaact gtgagaaaga aggaggaata caagatggcc	3300
ctgtacaacc tctaccctgg ggtctttgag actgtggaga tgctgcctc caaagctggc	3360
atctgggggg tggaaatgcct cattggggag cacctgcatg ctggcatgtc aaccctgttc	3420
ctggcttaca gcaacaagtgc ccagacaccc ctggaaatgg cctctggcca catcaggac	3480
ttccagatca ctgcctctgg ccagtatggc cagtgccac ccaaactggc caggctccac	3540
tactctggct ccatcaatgc atggtaacc aaggagccat tctcttggat caaggtggac	3600
ctgctggcac ccatgatcat tcatggcatc aagacacagg gggcaagaca gaaattctcc	3660
tctctgtaca tctcacatgtt catcatcatg tactctctgg atggcaagaa gtggcagaca	3720
tacagaggca actccactgg cacccctcatg gtcttcttg gcaatgtgga cagctctggc	3780
atcaaggaca acatcttcaa ccctccatc attgccatg acatcaggct gcacccacc	3840
cactactcaa tcagatcaac cctcaggatg gaactgtatgg gatgtgaccc gaaactctgc	3900
tcaatgcccc tggaaatggc gagcaaggcc atttctgtatgg cccagatcac tgcatctct	3960
tacttcacca acatgtttgc cacctggta ccatcaaaag ccaggctgca cctccaggga	4020
agaagaatgc cctggagacc ccaggtcaac aacccaaagg aatggctgca agtggacttc	4080
cagaagacaa tgaaagtcaac tgggtgtaca acccagggggg tcaagtctct gctcacctca	4140
atgtatgtga aggagttctt gatctcttcc tcaacaggatg gccaccatgt gacactttc	4200
ttccagaatgc gaaatgtcaac ggtgttccat ggcaaccagg actctttcac acctgtgggt	4260
aactcaatgg accccccctt cctgacaaga tacatgagaa ttcacccca gtctgggtc	4320
caccagattg ccctgagaat ggaagtccatg ggatgtgaggc cacaagacccgtactga	4377

<210> SEQ ID NO 93
<211> LENGTH: 4374
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 93

atgcagattg agctgtccac ctgtttcttt ctgtgcctgc tgagattctg cttctctgccc	60
accaggagat actacactggg ggctgtggaa ctttcttgg actacatgc gatctgaccc	120
ggagagatgc ctgtggatgc caggttccca cccagatgc ccaagtcctt cccattcaac	180
acctctgtgg tctacaagaa gacactctt gtggaaattca ctgaccaccc gttcaacatt	240
gcaaaaacca gaccaccctg gatggactc ctgggaccca ccattcaggc tgaggtgtat	300

-continued

gacactgtgg tcgtcaccct caagaacatg gcatcccacc ctgtgtctct gcatgtgtg	360
ggagtctcat actggaaatc ctctgaaggg gctgagtgat atgaccagac atcccagaga	420
gagaaagagg atgacaaggat gttccctggg aagtctcaca cctatgtgtg gcaagtctc	480
aaggagaatg gacccactgc atctgaccca ccctgcctga catactccta cctttctcat	540
gtggacctgg tcaaggacct caactctggat ctgattgggg cactgtgtt gtgcaggaa	600
ggatccctgg ccaaggagaa aacccagaca ctgcacaagt tcatttcctt gtttgcgtc	660
tttgatgagg gcaagtcttgc gcactctgaa acaaagaact ccctgtatgc agacaggat	720
getgcctctg ccagggcatg gcccaagatg cacactgtga atggctatgt gaacagatca	780
ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcatgtat tggcatggg	840
acaacccctg aagtgcactc catttcctg gagggacaca ctttcctggt caggaaccac	900
agacaagect ctctggagat ctctccatc accttcctca ctgcacagac actgtgtatg	960
gaccttggac agtttcctgtt gttctgccac atcttcctcc accagcatga tggcatggaa	1020
gcctatgtca aggtggactc atgcctctgaa gaaccacagc tcaggatgaa gaacaatgag	1080
gaggctgagg actatgtatgatgatgatgact gactctgaga tggatgtgtt cagatttgat	1140
gatgacaact ctccatcctt cattcagatc aggtctgtgg caaagaaaca ccccaagaca	1200
tgggtgcact acattgtgc tgaggaagag gactggactc atgcaccact ggtcctggcc	1260
cctgatgaca ggagctacaa gtctcagtagc ctcaacaatg gcccacaaag aatttggaaaga	1320
aagtacaaga aagttagattt catggcctac actgtatgaaa cttcaagac aagagaagcc	1380
attcagcat agtctggcat tctggatcca ctctgtatgatggatgtgg agacaccctg	1440
ctcatacatct tcaagaacca ggcctccagg ccctacaaca tctaccacata tggcatca	1500
gatgtcaggc ccctgtacatc caggagactg cccaaagggg tggaaacacctt caaggacttc	1560
ccatttctgc ctggagatgt cttaagatc aagtggactg tcactgtgaa ggatggacca	1620
acaaagtctg ccccccaggatg cctcaccaga tactactctt cttttgtgaa catggagaga	1680
gacctggcat ctggactgtatgatgatgact gactgtgtt acaaggagtc tggaccacag	1740
agaggcaacc agatcatgtc tgacaagaga aatgtgttgc tggatgtgtt cttttgtgaa	1800
aacagatcat ggtacactgatc tgagaacattt cagagattcc tggccaaaccc tggatggat	1860
caactggaaag accctgtatgtt ccaggcaagc aacatcatgc actccatcaa tggctatgt	1920
tttgcacttc tccagtttc tggatgtgtt catggatgtgg cttactgttgc tttttttct	1980
attggggcac aaactgtactt cttttctgtc ttcttctctg gatacacctt caagcacaag	2040
atgggtgtatg aggacaccctt gacactcttc ccatttcctgtt gggaaaactgtt gttcatgac	2100
atggagaacc ctggactgtt gattctggaa tggccacaact ctgacttcag aaacaggaa	2160
atgactgcac tgctcaaagt ctccctctgt gacaagaaca ctggggacta ctatgaggac	2220
tcttatgagg acatctctgc ctacactgttc agcaagaaca atgcccatttgc gcccagaac	2280
ttctctcaga atccacactgtt cctgaagaga caccagagat agatcaccatg gacaaccctc	2340
cagtctgacc aggaagagat tgactatgtatgatgactt ctgtggatgtt gaagaaggag	2400
gactttgaca tctatgtatgatgactt ggacgagaac cagtctccaa gatcatttca gaagaagaca	2460
agacactact tcattgtgc tggatgtggactt atggcatgtc ttcctctccc	2520
catgttcctca ggaacaggac acagtctggc tctgtgccac agttcaagaa agtggcttc	2580

-continued

caggagttca ctgatggctc attcacccag cccctgtaca gaggggaact gaatgagcac	2640
ctgggactcc tgggaccata catcagggct gaggtggaaag acaacatcat ggtgacattc	2700
agaaaaccagg cctccaggcc ctacagcttc tactcttccc tcatcagctc tgaggaagac	2760
cagagacaag gggctgagcc aagaaagaac tttgtgaaac ccaatgaaac caagacctac	2820
ttcttggaaag tccagcacca catggcaccc accaaggatg agtttgactg caaggcctgg	2880
gcatacttct ctgatgtgga cctggagaaa gatgtgact ctggcctgat tggccactc	2940
ctgggtctgcc acaccaaacac cctgaaccct gcacatggaa ggcaagtgac tgtgcaggag	3000
tttgcctct ttttaccat ctttgatgaa accaagtcat ggtacttcac tgagaacatg	3060
gagagaaact gcagagcacc atgcaacatt cagatggaa acccccaccc caaggagaac	3120
tacaggttcc atgccccatcaa tggctacatc atggacaccc tgcctggct tgcacatggca	3180
caggaccaga gaatcagatg gtacctgctt tctatggat ccaatgagaa cattcactcc	3240
atccacttct ctgggcatgt cttaactgtg agaaagaagg aggaatacaa gatggccctg	3300
tacaacccctt accctggggctt ctttgagact gtggagatgc tgcctccaa agctggcatc	3360
tggagggtgg aatgcctcat tggggagcac ctgcacatgtg gcatgtcaac octgttctg	3420
gtctacagca acaagtgcac gacaccctgt ggaatggccct ctggccacat cagggacttc	3480
cagatcactg cctctggcca gtatggccag tgggcaccca aactggccag gctccactac	3540
tctggctcca tcaatgcac gtcaccaag gagccattct ctggatcaaa ggtggactg	3600
ctggcacccca tgatcattca tggcatcaag acacaggggg caagacagaa attctccct	3660
ctgtacatct cacatgtcat catcatgtac tctctggatg gcaagaatgt gcagacatac	3720
agaggcaact ccactggcac cctcatggtc ttctttggca atgtggacag ctctggcatc	3780
aagcacaaca tcttcaaccc tcccatcatt gccagataca tcaggctgca ccccacccac	3840
tactcaatca gatcaacccct caggatggaa ctgatggat gtgacccatc ctccctgctca	3900
atgcacctgg gaatggagag caaggccatt tctgtatgccc agatcactgc atcctcttac	3960
ttcaccaaca tggggccac ctggtcacca tcaaaagcca ggctgcaccc ccagggaaaga	4020
agcaatgcct ggagacccca ggtcaacaac ccaaaggaaat ggctgcacgt ggactccag	4080
aagacaatga aagtcaactgg ggtgacaacc cagggggtca agtctctgtcacatc	4140
tatgtgaagg agttctgtat ctcttcctca caggatggcc accaggatggac actcttttc	4200
cagaatggca aagtcaaggt gttccaggcc aaccaggact ctttccatcc tgggtgaac	4260
tcactggacc ccccccctt gacaagatac ctgagaattc acccccagtc ttgggtccac	4320
cagattggcc tgagaatgga agtcctggga tggaggcac aagacctgtatc	4374

<210> SEQ ID NO 94
<211> LENGTH: 4374
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 94

atgcagattg agctgagcac ctgtttcttc ctgtgcctgc tgagggtctg cttctgtcc	60
accaggagat actacccctggg ggctgtggag ctttcttggg actacatgca gtctgactg	120
ggggagatgc ctgtggatgc caggttccca cccagatgc ccaaattcccccattcaac	180

-continued

acctctgtgg tctacaagaa gaccctttt gtggagttca ctgaccacct gttcaacatt	240
gccaaaccca ggccaccctg gatgggactc ctgggaccca ccattcaggc tgaggtgtat	300
gacactgtgg tcgtcaccc caagaacatg gcctcccacc ctgtgagcct gcatgctgtg	360
ggggtcagct actggaaagtc ctctgagggg gctgagttat atgaccagac ctcccagagg	420
gagaaggagg atgacaaagt gttccctggg aagagccaca cctatgttg gcaggtctc	480
aaggagaatg gccccactgc ctctgaccca ccctgcctga cctactccata cctttctcat	540
gtggacctgg tcaaggaccc caactcttga ctgattgggg ccctgctggt gtgcagggag	600
ggctccctgg ccaaagagaa gacccagacc ctgcacaagt tcatttcctt gtttgctgtc	660
ttttagtggg gcaagagactg gcactcttga accaagaact ccctgatgca ggacaggat	720
gtgcctctg ccagggcctg gcccaagatg cacactgtga atggctatgt gaacaggagc	780
ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcatgttat tggcatgggg	840
acaacccctg aggtgcactc catttcctg gagggccaca ctttctggc caggaaccac	900
agacaggccca gcctggagat cagccccatc accttcctca ctgcccagac cctgctgtat	960
gacctcggac agttccctgct gttctgccac atcagctccc accagcatga tggcatggag	1020
gcctatgtca aggtggacag ctgcccctg gagccacagc tcaggatgaa gaacaatgag	1080
gaggctgagg actatgtatgatgatgact gactctgaga tggatgttgtt ccgccttgat	1140
gtatgacaaca gccccatcattt cattcagatc aggtctgtgg ccaagaaaca ccccaagacc	1200
tgggtgcact acattgctgc tgaggaggag gactggact atgccccact ggtcctggcc	1260
cctgatgaca ggagctacaa gagccagttac ctcaacaatg gcccacagag gattggacgc	1320
aagtacaaga aagtcaaggat catggccctac actgtatgaaa cttcaagac cagggaggcc	1380
attcagcatg agtctggcat cctggggccca ctccctgtatg gggaggtggg ggacaccctg	1440
ctcatcatct tcaagaacca ggccctccagg ccctacaaca tctaccacata tggcatca	1500
gatgtcaggc ccctgtacag ccgcaggctg ccaaaggggg taaaacaccc caaggacttc	1560
ccattctgc ctggggagat cttcaagtac aagtggactg tcactgtgaa ggatggacca	1620
accaaatactg accccaggtg cctcaccaga tactactcca gttttgtgaa catggagagg	1680
gacctggccct ctggcctgtat tggccactg ctcatctgatc acaaggagtc tgtggaccag	1740
agggggaaacc agatcatgtc tgacaagagg aatgtgattc tggatgttgtt ctttgtatgag	1800
aacaggagct ggtacctgac tgagaacatt cagcgcttcc tgcccaaccc tgctgggtg	1860
cagctggagg accctgagtt ccaggccagc aacatcatgc actccatcaa tggctatgtg	1920
tttgacagcc tccagcttc tggatgtgttgc catgaggtgg ccactggta catttttct	1980
attggggccca agactgactt cttttctgtc tttttctgtc gttttttttt cttttttttt	2040
atgggtgtatg aggacaccct gaccctcttc ccattctctg gggagactgt gttcatgagc	2100
atggagaacc ctggccctgtg gattctggga tgccacaact ctgacttccg caacaggggc	2160
atgactgccc tgctcaaagt ctccctctgt gacaagaaca ctggggacta ctatgaggac	2220
agctatgagg acatctctgc ctacactgtc agcaagaaca atgccattga gcccaggagc	2280
ttcageccaga atccacactgt cctgaaacgc caccagggg agatcaccag gaccaccctc	2340
cagtctgacc aggaggagat tgactatgat gacaccattt ctgtggagat gaagaaagag	2400
gactttgaca tctatgacga ggacgagaac cagagccaa ggagcttcca gaagaagacc	2460

-continued

aggcactact tcattgctgc tgtggagcgc ctgtggact atggcatgag ctccagcccc	2520
catgtctca ggaacagggc ccagtctggc tctgtgccac agttcaagaa agtggcttc	2580
caagagttca ctgtatggcag cttcacccag cccctgtaca gaggggagct gaatgagcac	2640
ctgggactcc tgggeccata catcagggtc gaggtggagg acaacatcat ggtgaccctc	2700
cgcacaccagg cttccaggcc ctacagcttc tacagctccc tcatcagcta tgaggaggac	2760
cagaggcagg gggctgagcc acgcaagaac tttgtgaaac ccaatgaaac caagacctac	2820
ttctggaaag tccagcacca catggccccc accaaggatg agtttgactg caaggcctgg	2880
gcctacttct ctgtatgttgc cctggagaag gatgtgcact ctggcctgat tggcccactc	2940
ctggctgccc acaccaacac cctgaaccct gcccattggaa ggcaagtgc tgcaggcagg	3000
tttgcctct tcttcaccat ctttgatgaa accaagagct ggtacttcac tgagaacatg	3060
gagcgcaact gcagggccccc atgcaacatt cagatggagg acccccacccaa caaagagaac	3120
taccgcttcc atgcccattcaa tggctacatc atggacaccc tgcctgggtc tgcattggcc	3180
caggaccaga ggatcagggtg gtacactgct tctatgggtc ccaatgagaa cattcactcc	3240
atccacttct ctgggcatgt cttcactgtg cgcaagaagg aggagtacaa gatggccctg	3300
tacaacctct acccctgggtt ctttgagact gtggagatgc tgcctccaa agctggcattc	3360
tgggggttgg agtgcctcat tggggagcac ctgcattgtc gcatgagcac cctgttctg	3420
gtctacagca acaagtgcac gaccccccgt ggaatggcct ctggccacat cagggacttc	3480
cagatcactg cctctggcca gtatggccag tggggccccc agctggccag gctccactac	3540
tctggatcca tcaatgcctg gagcaccaag gagccattca gtcggatcaa agtggactg	3600
ctggcccccata tgatcatcca tggcatcaag acccagggggg ccaggcagaa gttctccagc	3660
ctgtatcatca gccaggatcatc catcatgtac agcctggatg gcaagaaatg gcagacccat	3720
agaggcaact ccactggAAC actcatggtc ttctttggca atgtggacag ctctggcatc	3780
aagcacaaca tcttcaaccc cccatcatc gccagataca tcaggctgca ccccccaccc	3840
tacagcatcc gcagcacccct caggatggag ctgtatgggtc gtgacactgaa ctccctgcagc	3900
atgcccctgg gcatggagag caaggccatt tctgtatggcc agatcactgc ctccagctac	3960
ttcaccaaca tggggccac ctggagccca agcaaggccca ggctgcaccc ccaggaaagg	4020
agcaatgcct ggaggccccc ggtcaacaac ccaaaggagt ggctgcaggt ggacttccag	4080
aagaccatga aggtcactgg ggtgaccacc cagggggtca agagcctgct caccagcatg	4140
tatgtgaagg agttctgtat cagctccagc caggatggcc accagtggac ccttttttc	4200
cagaatggca aggtcaaggt gttccagggc aaccaggaca gtttcaccc tgggtgaac	4260
agcctggacc cccccccttcc gaccagatac ctggaggattc acccccagag ctgggtccac	4320
cagattgccc tgaggatgga ggtcctggga tggaggccc aggacctgta ctga	4374

```

<210> SEQ_ID NO 95
<211> LENGTH: 4377
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polynucleotide

<400> SEQUENCE: 95

```

-continued

atgcagattg agctgagcac ctgcttcttc ctgtgcctgc tgaggttctg cttctctgcc	60
accaggagat actacactggg ggctgtggag ctttcttggg actacatgca gtctgacactg	120
ggggagctgc ctgtggatgc caggttccca cccagagtgc ccaaattctt cccattcaac	180
acctctgtgg tctacaagaa gaccctctt gtggagttca ctgaccacct gttcaacatt	240
gceaaaccca ggccaccctg gatgggactc ctgggaccca ccattcaggc tgagggttat	300
gacactgtgg tcatcacccct caagaacatg gcctcccacc ctgtgagct gcatgtgtg	360
ggggtcagct acttggaaaggc ctctgagggg gctgagttatg atgaccagac ctcccagagg	420
gagaaggagg atgacaaaagt gttccctggg ggcagccaca cctatgttg gcaggtcctc	480
aaggagaatg gccccatggc ctctgaccca ctctgcctga cctactccca cctttctcat	540
gtggacctgg tcaaggacct caactcttggc ctgattgggg ccctgtgtt gtgcaggag	600
ggctccctgg ccaaagagaa gacccagacc ctgcacaagt tcattctctt gtttgctgtc	660
tttgatgagg gcaagagctg gcactctgaa accaagaact ccctgtatgca ggacaggat	720
gtgcctctg ccaggccctg gcccaagatg cacactgtga atggctatgt gaacaggagc	780
ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcatgtat tggcatgggg	840
acaacccctg aggtgcactc catttcctgg gagggccaca ctttctgtt caggaaccac	900
agacaggcca gcctggagat cageccccatc accttcctca ctgcccagac cctgtgtatg	960
gacctteggac agttctgtct gttctgccac atcagctccc accagcatga tggcatggag	1020
gcctatgtca aggtggacag ctgccttag gagccacagc tcaggatgaa gaacaatgag	1080
gaggctgagg actatgtatc tgactctgact gactctgtt gttgtgtt ccgtttgtat	1140
gatgacaaca gccccatctt cattcagatc aggtctgtgg ccaagaaaca ccccaagacc	1200
tgggtgcact acattgtgc tgaggaggag gactgggact atgccccact ggtcctggcc	1260
cctgtatgaca ggagctacaa gagecagtagc ctcaacaatg gcccacagag gattggacgc	1320
aagtacaaga aagttaggtt catggccctac actgtatgaa ctttcaagac cagggaggcc	1380
attcagcatg agtctggcat cctggggcca ctctgtatg gggaggtggg ggacaccctg	1440
ctcatcatct tcaagaacca ggcctccagg ccctacaaca tctaccacaca tggcatcact	1500
gatgtcaggc ccctgtacag ccgcaggctg ccaaaggggg tggaaacaccc caaggactc	1560
ccatttctgc ctggggagat cttcaagtac aagtggactc tcactgttga ggttggacca	1620
acccaaatctg accccagggtg cctcaccaga tactactcca gttttgttcatgaa catggagagg	1680
gacctggcct ctggcctgtat tggcccactg ctcatctgtc acaaggagtc tggaccacag	1740
agggggaaacc agatcatgtc tgacaagagg aatgtgattc tggatgtt ctttgtatgag	1800
aacaggagct ggtacactgac tgagaacatt cagcgcttcc tgcccaaccc tggatgggtt	1860
cagctggagg acccctgagtt ccaggccagc aacatcatgc actccatcaa tggctatgt	1920
tttgacagcc tccagcttcc tggatgttgcctg catgagggtgg cctactggta cattttct	1980
attggggccct agactgactt cttttctgtc tttttctgtt gotacacccctt caaacacaag	2040
atgggtatg aggacaccct gaccctcttc ccattctctg gggagactgt gttcatgagc	2100
atggagaacc ctggccctgtc gattctggga tgccacaact ctgacttccg caacaggggc	2160
atgactgccc tgctcaaagt ctcctcctgt gacaagaaca ctggggacta ctatgaggac	2220
agctatgagg acatctctgc ctacactgtc agcaagaaca ataccaccta cgtgaaccgc	2280

-continued

tcctgagcc agaatccacc tgtcctgaaa cgccaccaga gggagatcac caggaccacc	2340
ctccagtctg accaggagga gattgactat gatgacacca tttctgtgga gatgaagaaa	2400
gaggacttg acatctatga cgaggacgag aaccagaccc caaggagctt ccagaagaag	2460
accaggcact acttcattgc tgctgtggag cgctgtggg actatggcat gagctccagc	2520
ccccatgtcc tcaggaacag ggcccaagtct ggctctgtgc cacagtcaa gaaagtggtc	2580
ttccaagagt tcactgatgg cagttcacc cagccccgt acagagggg gctgaatgag	2640
cacctggac tcctgggccc atacatcagg gctgagggtgg aggacaacat catggtgacc	2700
ttccgcaacc aggccctccag gccctacagc ttctacagct ccctcatcag ctatgaggag	2760
gaccagggc agggggctga gccacgcag aactttgtga aacccaatga aaccaagacc	2820
tacttctgga aagtccagca ccacatggcc cccaccaagg atgagtttgta ctgcaaggcc	2880
tgggcctact tctctgtatgt ggacctggag aaggatgtgc actctggcct gattggccca	2940
ctcctggctc gccacaccaa caccctgaac cctgccccatg gaaggcaagt gactgtgcag	3000
gagtttgcctc ttttttcac catctttgtat gaaaccaaga gttggactt cactgagaac	3060
atggagcgca actgcagggc cccatgcac attcagatgg aggacccac cttcaaagag	3120
aactaccgc tccatgccat caatggctac atcatggaca ccctgectgg gcttgcata	3180
gcccaggacc agaggatcag gtggtacctg ctttctatgg gtcacatga gaacattcac	3240
tccatccact tctctggca tgtcttcact gtgcgcagaagg aggaggagta caagatggcc	3300
ctgtacaacc tctaccctgg ggtcttttag actgtggaga tgctgeccctc caaagctggc	3360
atctgggggg tggagtgcct cattggggag cacctgcata ctggcatgag caccctgttc	3420
ctggctaca gcaacaagtg ccagaccccc ctggaaatgg cctctggcca catcaggac	3480
ttccagatca ctgcctctgg ccagatggc cagttggccca ccaagctggc caggetccac	3540
tactctggat ccatcaatgc ctggagcacc aaggagccat tcagctggat caaagtggac	3600
ctgctggccc ccatgatcat ccatggcata aagacccagg gggccaggca gaagttctcc	3660
agcctgtaca tcagecagtt catcatcatg tacagctgg atggcaagaa atggcagacc	3720
taacagggca actccactgg aacactcatg gtcttctttg gcaatgtgga cagctctggc	3780
atcaagcaca acatcttcaa cccccaatc atgcgcagat acatcaggt gcacccacc	3840
cactacagca tccgcagcac cctcaggatg gagctgtatgg gctgtgaccc gaaactctgc	3900
agcatgcccc tgggcatgga gagcaaggcc atttctgtatg cccagatcac tgcctccagc	3960
tacttcacca acatgtttgc cacctggagc ccaagcaagg ccaggctgca cctccaggaa	4020
aggagcaatg cctggaggcc ccaggtaac aacccaaagg agtggctgca ggtggacttc	4080
cagaagacca tgaaggtaac tgggtgtacc acccaggggg tcaagagcct gtcaccac	4140
atgtatgtga aggagttccct gatcagctcc agccaggatg gccaccaggat gaccctttc	4200
ttccagaatg gcaaggtaaa ggtgttccag ggcaaccagg acagcttcac ccctgtggtg	4260
aacagcgtgg accccccccct cctgaccaga tacctgagga ttcacccca gagctgggtc	4320
caccagattg ccctgaggat ggaggtcctg ggatgtgagg cccaggaccc gtactga	4377

<210> SEQ ID NO 96

<211> LENGTH: 4377

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

-continued

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 96

atgcagattg agctgagcac ctgtttcttc ctgtgcctgc tgaggttctg cttctctgcc 60
accaggagat actaacctggg ggctgtggag ctttcttggg actacatgca gtctgacactg 120
ggggagctgc ctgtggatgc caggttccca cccagagtgc ccaaattcctt cccattcaac 180
acctctgtgg tctacaagaa gaccctcttt gtggagttca ctgaccacct gttcaacatt 240
gccaaaccca ggccaccctg gatgggactc ctgggaccca ccattcaggc tgagggttat 300
gacactgtgg tcgtcaccct caagaacatg gcctcccacc ctgtgagcct gcatgctgtg 360
ggggtcagct actggaaagtc ctctgagggg gctgagttatg atgaccagac ctcccagagg 420
gagaaggagg atgacaaaatg gttccctggg aagagccaca cctatgttg gcaggtcctc 480
aaggagaatg gccccactgc ctctgaccca ccctgcctga cctactccctt cctttctcat 540
gtggacctgg tcaaggaccc caactctgga ctgattgggg ccctgttgtt gtgcaggag 600
ggctccctgg ccaaagagaa gacccagacc ctgcacaatg tcatttcctt gtttgcgtgc 660
tttgcattggagg gcaagagactg gcactctgaa accaagaact ccctgtatgca ggacaggat 720
gctgcctctg ccagggcctg gcccaagatg cacactgtga atggctatgt gaacaggagc 780
ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcatgtat tggcatgggg 840
acaacccctg aggtgcactc catttcctg gagggccaca ctttccttgtt caggaaccac 900
agacaggcca gcctggagat cagccccatc accttcctca ctgcccagac cctgtatg 960
gaccttcggac agttccctgct gttctgccac atcagctccc accagcatga tggcatggag 1020
gcctatgtca aggtggacag ctgccttag gagccacagc tcaggatgaa gaacaatgag 1080
gaggctgagg actatgtatg tgacctgact gactctgaga tggatgttgtt ccgccttgat 1140
gatgacaaca gcccatactt cattcagatc aggtctgtgg ccaagaaaca ccccaagacc 1200
tgggtgcact acattgtgc tgaggaggag gactggact atgccccact ggtccctggcc 1260
ctgtatgaca ggagctacaa gagccagatc ctcacaatg gcccacagag gattggacgc 1320
aagtacaaga aagtcaaggat catggcctac actgtatgaa cttcaagac cagggaggcc 1380
attcagcatg agtctggcat cttggccca ctctgtatg gggaggtggg ggacaccctg 1440
ctcatcatct tcaagaacca ggcctccagg ccctacaaca tctaccacca tggcatcaact 1500
gatgtcaggc ccctgtacatc ccgcaggctg ccaaagggggg tggaaacacctt caaggacttc 1560
cccatatttcctgc ctggggagat cttcaagatc aagtggactg tcactgtgga ggttggacca 1620
acccaaatctg accccaggatc cttcaccaga tactactcca gttttgtgaa catggagagg 1680
gaccttggctt ctggccctgat tggccactg ctcatctgtt acaaggagtc tggatggaccc 1740
aggggaaacc agatcatgtc tgacaagagg aatgtgattc tggatgttgtt ctgtatgatg 1800
aacaggagct ggtacactgac tgagaacatt cagcgcttc tgcccaaccc tggatgggtt 1860
cagctggagg acccttgatc ccaggccagc aacatcatgc actccatcaa tggatgttg 1920
tttgcacggcc tccagtttc tggatgttgtt catggatgtt gttttgttcc tttttttttt 1980
atggggccca agactgtactt cttttctgtt ttttttttttt gttttttttttt cttttttttt 2040
atggatgtatg aggacaccctt gacccctttt ccattctgtt gggatgttgtt gttttttttt 2100
atggagaacc ctggccctgtt gattctggta tgccacaact ctgacttccg caacaggggc 2160

-continued

atgactgccc tgctcaaagt ctcctcctgt gacaagaaca ctggggacta ctatgaggac	2220
agctatgagg acatctctgc ctacctgctc agcaagaaca ataccaccta cgtgaaccgc	2280
tccctgagcc agaatccacc tgcctgaaa cgccaccaga gggagatcac caggaccacc	2340
ctccagtctc accaggagga gattgactat gatgacca tttctgtgga gatgaagaaa	2400
gaggactttg acatctatga cgaggacgag aaccagagcc caaggagctt ccagaagaag	2460
accaggact acttcattgc tgctgtggag cgccctgtgg actatggcat gagctccagc	2520
ccccatgtcc tcaggaacag ggcccagtct ggctctgtgc cacagttcaa gaaagtggtc	2580
ttccaagagt tcactgtatgg cagttcacc cagccccgt acagagggga gctgaatgag	2640
cacctggac tcctgggccc atacatcagg gctgagggtgg aggacaacat catggtagcc	2700
ttccgcAACC aggccctccag gccctacagc ttctacagct ccctcatcag ctatgaggag	2760
gaccagggc agggggctga gccacgcAAC aactttgtaa aacccaaatgaa aaccaagacc	2820
taacttctgaa aagtccagca ccacatggcc cccaccaagg atgagttgtaa ctgcaaggcc	2880
tgggcctact tctctgtatgt ggacctggag aaggatgtgc actctggcctt gattggccca	2940
ctcctggctt gccacaccaa caccctgaac cctgcccattt gaaggcaagt gactgtgcag	3000
gagtttggcc tcttcttcac catctttgtat gaaaccaaga gctggactt cactgagaac	3060
atggagcgca actgcagggc cccatgcAAC attcagatgg aggacccac cttcaaagag	3120
aactaccgct tccatgccat caatggctac atcatggaca ccctgcctgg gcttgcattt	3180
gcccaggacc agaggatcag gtggtagctt ctttctatgg gctccaaatgaa gaacatttcac	3240
tccatccact tctctggca tgccttactt gtgcgcAAAG aggaggagta caagatggcc	3300
ctgtacaacc tctaccctgg ggtcttttagt actgtggata tgctgccttc caaagctggc	3360
atctggaggg tggagtgcctt cattggggag cacctgcattt ctggcatgag caccctgttc	3420
ctggcttaca gcaacaagtcc agggggccctt ctgggaatgg cctctggccca catcaggacc	3480
ttccagatca tgcctctgg ccagtatggc cagtgccccc ccaagctggc caggtccac	3540
tactctggat ccatcaatgc ctggaggacc aaggaggccat tcaagctggat caaaatggac	3600
ctgctggccc ccatgatcatt ccatggcattt aagacccagg gggccaggca gaagttctcc	3660
agcctgtaca tcagecagttt catcatcatg tacagcctgg atggcaagaa atggcagacc	3720
tacagaggca actccactgg aacactcatg gtcttctttt gcaatgtggaa cagctctggc	3780
atcaaggcaca acatcttcaa ccccccattt atcgccagat acatcaggctt gcacccacc	3840
cactacagca tccgcagcac cctcaggatg gagctgtatgg gctgtgaccc gaaactcctgc	3900
agcatgcccc tgggcatttggaa gagcaaggcc atttctgtatgg cccagatcac tgcctccagc	3960
tacttccatca acatgtttgc cacctggacc ccaagcaagg ccaggctgca cctccaggaa	4020
aggagcaatgcctggaggcc ccaggtaac aacccaaagg agtggctgca ggtggacttc	4080
cagaagccatca tgaaggtaac tgggtgtacc acccagggggg tcaagagcctt gctcaccagc	4140
atgtatgtaa aggagttccctt gatcagctcc agccaggatg gccaccagtg gaccctttc	4200
ttccagaatgcgaaggtaaa ggtgttccag ggcaaccagg acagcttcac ccctgtgggt	4260
aacagectgg acccccccctt cctgaccaga tacctgaggaa ttccacccca gagctgggtc	4320
caccagatttgcctgaggat ggatgtgaggccc cccaggactt gtactgaa	4377

-continued

```
<210> SEQ ID NO 97
<211> LENGTH: 4374
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
polynucleotide

<400> SEQUENCE: 97

atgcagattg agcttagcac ctgttttgc ctgtgcattc tgaggttctg cttttctgcc      60
accaggagat actacactggg ggctgtggag ctttcttggg actacatgca gtctgacccctg    120
ggggagctgc ctgtggatgc caggttccca cccagagtgc ccaaattccctt cccattcaac    180
acctctgtgg tctacaagaa gaccctcttt gtggagttca ctgaccaccc gttcaacatt     240
gccaaaccca ggccaccctg gatgggactc ctgggacccca ccattcaggc tgaggtgtat     300
gacactgtgg tcatcacccct caagaacatcg gcctcccacc ctgtgagccct gcatgctgtg   360
ggggtcagct actggaaggc ctctgagggg gctgagttatg atgaccagac ctcccaaggagg  420
gagaaggagg atgacaaatgt ttccctggg ggccagccaca cctatgtgtg gcaggccctc     480
aaggagaatg gccccatggc ctctgacccca ctctgcctga cctactocta cctttctcat    540
gtggacccctgg tcaaggacccct caactctggc ctgattttgggg ccctgtgtt gtgcaggggag 600
ggctccctgg ccaaagagaa gacccagacc ctgcacaatgt tcatttcctt gtttgcgtgc    660
tttgatgagg gcaagagctg gcactctgaa accaagaact ccctgtatgc ggacagggtat   720
gtgcctctg ccagggccctg gcccaagatg cacactgtga atggctatgt gaacaggaggc  780
ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcatgtgtat tggcatgggg    840
acaacccctg aggtgcactc cattttctggc gagggccaca ccttcctggt caggaaccac    900
agacaggcca gcctggagat cagccccatc accttcctca ctgcccagac cctgtgtatg    960
gacctcggac agtttctgtt gtccctggccat atcagctccc accagcatga tggcatggag 1020
gcctatgtca aggtggacag ctgcctgtt gagccacaggc tcaggatgtt gaaaaatgtg    1080
gaggctgagg actatgtatg tgaccccttact gactctgtt gtttgcgtgtt ccgtttgtat 1140
gtgacaaca gccccatctt cattcagatc aggttctgtt ccaagaaaca ccccaagacc    1200
tgggtgtact acattgttgc tgaggaggag gactgggact atgccccact ggtccctggcc 1260
cctgtatgaca ggagctacaa gagccagtttac ctcaacaatgt gcccacaggag gattggacgc 1320
aagtacaaga aagtccatgtt catggccatc actgtatgtt cttcaagac cagggaggcc    1380
attcagcatg agtctggcat cctggggccca ctccctgtatg gggagggtggg ggacaccctg 1440
ctccatcatct tcaagaacca ggcctccagg ccctacaaca tctaccacca tggcatcact 1500
gatgtcaggc ccctgtatcag ccgcaggctg ccaaaaggggg tgaaacaccc caaggacttc 1560
cccattctgc ctggggagat cttcaagtttac aagtggactt gtttgcgtgtt ggttggacca 1620
acccaaatctg accccagggtt cctcaccaga tactactcca gttttgttgc catggagagg 1680
gacctggccctt otggccctgtat tggccctgtt ctcatctgtt acaaggaggc ttttgcgtgtt 1740
aggggaaacc agatcatgtt tgacaaggagg aatgtgattt gtttctgtt ctttgcgtgtt 1800
aacaggaggctt ggttacccatc tgagaacattt cagccgttcc tggcccaaccc tggctgggtt 1860
cagctggagg accctgtatgtt ccaggccaggc aacatcatgc actccatcaa tggctatgtt 1920
tttgacagcc tccagtttc ttttgcgtgtt catggagggtt gtttgcgtgtt ctttgcgtgtt 1980
```

-continued

atggggccc agactgactt ccttctgtc ttcttcctcg gtcacccctt caaacacaag	2040
atgggttatg aggacaccct gaccctcttc ccattctctg gggagactgt gttcatgagc	2100
atggagaacc ctggcctgtg gattctggga tgccacaact ctgacttccg caacaggggc	2160
atgactgccc tgctcaaagt ctccctctgt gacaagaaca ctggggacta ctatgaggac	2220
agctatgagg acatctctgc ctacctgtc agcaagaaca atgccattga gcccaggagc	2280
ttcagccaga atccacctgt cctgaaacgc caccagaggg agatcaccag gaccaccctc	2340
cagtctgacc aggaggagat tgactatgt gacaccattt ctgtggagat gaagaagag	2400
gactttgaca tctatgacga ggacgagaac cagagccaa ggagcttcca gaagaagacc	2460
aggcactact tcattgtgc tggggcgc ctgtggact atggcatgag ctccagcccc	2520
catgtcctca ggaacagggc ccagtctggc tctgtgccac agttcaagaa agtggcttc	2580
caagagttca ctgatggcag cttcacccag cccctgtaca gaggggagct gaatgagcac	2640
ctgggactcc tggggccata catcagggtt gaggtggagg acaacatcat ggtgaccctc	2700
cgcaaccagg cctccaggcc ctacagcttc tacagctccc tcatcagcta tgaggaggac	2760
cagagggcagg gggctgagcc acgcaagaac tttgtgaaac ccaatgaaac caagacctac	2820
ttctggaaag tccagcacca catggccccc accaaggatg agtttgactg caaggctgg	2880
gcctacttct ctgatgtgga cctggagaag gatgtgcact ctggcctgat tggccactc	2940
ctgggtctgc acaccaacac cctgaaccct gccatggaa ggcaagtgac tggcaggag	3000
tttgcctct tttcaccat ctttgcata accaagagct ggtacttcac tgagaacatg	3060
gagcgcacta gcagggccccc atgcaacatt cagatggagg accccaccc caaagagaac	3120
taccgcttcc atgcccataa tggcacatc atggacaccc tgcctggct tgtcatggcc	3180
caggaccaga ggatcagggtg gtactgttt tctatggctt ccaatgagaa cattcactcc	3240
atccacttct ctgggcatgt cttaactgtg cgcaagaagg aggagttacaa gatggccctg	3300
tacaacccctt accctgggtt ctttgcact gtggagatgc tggccctccaa agctggcatc	3360
tggaggggtgg agtgcctcat tggggagcac ctgcattgtc gcatgagcac cctgttctg	3420
gtctacagca acaagtgcca gacccctgtt ggaatggct ctggccacat cagggacttc	3480
cagatcactg cctctggcca gtatggccag tggggccca agctggccag gctccactac	3540
tctggatcca tcaatgcctg gaggaccaag gagccattca gctggatcaa agtggactg	3600
ctggcccccata tgatcatcca tggcatcaag acccaggggg ccaggcagaa gttctccagc	3660
ctgtacatca gccagttcat catcatgtac agcctggatg gcaagaaatg gcagacctac	3720
agaggcaact ccactggAAC actcatggTC ttctttggCA atgtggacAG ctctggcatC	3780
aagcacaaca tcttcaaccc cccaaatcatc gccagataca tcaggctgca ccccaaccac	3840
tacagcatcc gcagcaccct caggatggag ctgatggctt gtgacccat ctcctgcAGC	3900
atgcccctgg gcatggagag caaggccatt tctgatgccc agatcactgc ctccagctac	3960
ttcaccaaca tggggccac ctggggccAA agcaaggccA ggctgcaccc ccaggaaagg	4020
agcaatgcct ggaggccccc ggtcaacaAC cccaaaggAG ggctgcaggT ggacttcAG	4080
aagccatga aggtcaactgg ggtgaccacc cagggggTC agagcctgtt caccagcatG	4140
tatgtgaagg agttcctgtat cagctccAGC caggatggcc accagtgac cctttcttc	4200
cagaatggca aggtcaaggT gttccaggGC aaccaggaca gtttcacccc tgggtgaac	4260

-continued

agcctggacc	ccccctcct	gaccagatac	ctgaggattc	accccccagag	ctgggtccac	4320
cagattgcc	tgaggatgga	ggtccctggga	tgtgaggccc	aggacctgta	ctga	4374
<210> SEQ_ID NO 98						
<211> LENGTH: 4377						
<212> TYPE: DNA						
<213> ORGANISM: Artificial Sequence						
<220> FEATURE:						
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide						
<400> SEQUENCE: 98						
atgcagattg	agctgagcac	ctgtttcttc	ctgtgcctgc	tgaggttctg	cttctctgcc	60
accaggagat	actacactggg	ggctgtggag	ctttcttggg	actacatgca	gtctgacactg	120
ggggagctgc	ctgtggatgc	caggttccca	cccagagtgc	ccaaatcctt	cccattcaac	180
acctctgtgg	tctacaagaa	gaccctcttt	gtggagttca	ctgaccacct	gttcaacatt	240
gccaacccca	ggccaccctg	gatgggactc	ctgggaccca	ccattcaggc	tgagggttat	300
gacactgtgg	tcatcaccct	caagaacatg	gcctcccacc	ctgtgagcct	gcatgctgtg	360
ggggtcagct	actggaaaggc	ctctgagggg	gctgagttatg	atgaccagac	ctcccagagg	420
gagaaggagg	atgacaaggat	gttccctggg	ggcagccaca	cctatgtgtg	gcaggtcctc	480
aaggagaatg	gccccatggc	ctctgaccca	ctctgcctga	cctactccct	ccttctcat	540
gtggacctgg	tcaaggacct	caactcttga	ctgattgggg	ccctgctgg	gtgcagggag	600
ggctccctgg	ccaaagagaa	gaccaggacc	ctgcacaagg	tcatttcct	gtttgtgtc	660
ttttaggagg	gcaagagctg	gcactctgaa	accaagaact	ccctgatgca	ggacagggat	720
gctgcctctg	ccagggctg	gcccaagatg	cacactgtga	atggctatgt	gaacaggagc	780
ctgcctggac	tcattggctg	ccacaggaaa	tctgtctact	ggcatgtgt	tggcatgggg	840
acaacccctg	aggtgcactc	catttcctg	gaggccaca	ccttcctgg	caggaaccac	900
agacaggcca	gcctggagat	cagccccatc	accttcctca	ctgcccagac	cctgtgtatg	960
gacctcggac	agttcctgt	gtctgtccac	atcagctccc	accagcatga	tggcatggag	1020
gcctatgtca	aggtggacag	ctgcccctgag	gagccacagc	tcaggatgaa	gaacaatggag	1080
gaggctgagg	actatgtatg	tgacactgact	gactctgaga	tggatgtgg	ccgctttgtat	1140
gatgacaaca	gccccatctt	cattcagatc	aggtctgtgg	ccaaagaaaca	ccccaaagacc	1200
tgggtgcact	acattgtgc	tgaggaggag	gactggact	atgccccact	ggtcctggcc	1260
cctgtatgaca	ggagctacaa	gagccagttac	ctcaacaatg	gcccacagag	gattggacgc	1320
aagtacaaga	aagtcaagg	catggcttac	actgtatgaaa	ccttcaagac	cagggaggcc	1380
attcagcatg	agtctggcat	cctggccca	ctcctgtatg	gggaggtgg	ggacacccctg	1440
ctcatcatct	tcaagaacca	ggcctccagg	ccctacaaca	tctaccacaca	tggcatact	1500
gatgtcaggc	ccctgtacag	ccgcaggctg	ccaaaggggg	tgaaacacct	caaggacttc	1560
cccattctgc	ctggggagat	cttcaagtac	aagtggactg	tcactgtgga	ggatggacca	1620
acccaaatctg	accccagg	cctcaccaga	tactactcca	gttttgtaa	catggagagg	1680
gacctggcct	ctggcctgtat	tggeccactg	ctcatctgt	acaaggagtc	tgtggaccag	1740
aggggaaacc	agatcatgtc	tgacaagagg	aatgtgattc	tgttctgtgt	ctttgtatgag	1800
aacaggagct	ggtacactgac	tgagaacatt	cagcgcttcc	tgcccaaccc	tgtgggggtg	1860

-continued

cagctggagg accctgagtt ccaggccago aacatcatgc actccatcaa tggctatgtg	1920
tttgacagcc tccagtttc tgtctgcctg catgaggtgg cctactggta cattttct	1980
atggggcccc agactgactt ccttctgtc ttcttctctg gctacacctt caaacacaag	2040
atgggtgtatg aggacacccct gacectcttc ccattctctg gggagactgt gttcatgagc	2100
atggagaacc ctggcctgtg gattctggga tgccacaact ctgacttccg caacaggggc	2160
atgactgccc tgctcaagt ctccctctgt gacaagaaca ctggggacta ctatgaggac	2220
agctatgagg acatctctgc ctacctgctc agcaagaaca ataccaccta cgtgaaccgc	2280
tccctgagcc agaatccacc tgcgttggaaa cgccaccaga gggagatcac caggaccacc	2340
ctccagtcg accaggagga gattgactat gatgacacca tttctgtgga gatgaagaaa	2400
gaggactttg acatctatga cgaggacgag aaccagagcc caaggagctt ccagaagaag	2460
accaggactt acttcattgc tgctgtggag cgccctgtggg actatggcat gagctccagc	2520
ccccatgtcc tcaggaacag ggcccagtct ggctctgtgc cacagttcaa gaaagtggc	2580
ttccaagagt tcactgtatgg cagttcacc cagccctgt acagagggga gctgaatgag	2640
cacctgggac tcctgggccc atacatcagg gctgaggtgg aggacaacat catggtagcc	2700
ttccgcAACC accgcctccag gcctacagc ttctacagct ccctcatcag ctatgaggag	2760
gaccagaggc agggggctgaa gccacgcaag aactttgtaa aacccaaatgaa aaccaagacc	2820
tacttctgaa aagtccagca ccacatggcc cccaccaagg atgagttgaa ctgcaaggcc	2880
tgggcctact tctctgtatgt ggacctggag aaggatgtgc actctggctt gattggccca	2940
cctctgtctt gccacaccaa caccctgaac cctgcccattt gaaggcaagt gactgtgcag	3000
gagtttgcctt tcttcttcac catctttgtt gaaaccaaga gctggactt cactgagaac	3060
atggagcgca actgcagggc cccatgcaac attcagatgg aggaccccac cttcaaaagag	3120
aactaccgtt tccatgccat caatggctac atcatggaca ccctgcctgg gcttgcattt	3180
gcgcaggacc agaggatcag gtggtagctt ctttctatgg gctccaaatgaa gaacattcac	3240
tccatccact tctctggca tgccttactt gtgcgcagaagg aggaggagta caagatggcc	3300
ctgtacaacc tctaccctgg ggtcttttagt actgtggaga tgctgccttc caaagctggc	3360
atctggaggg tggagtgcctt cattggggag cacctgcattt ctggcatgag caccctgttc	3420
ctggcttaca gcaacaagtgc ccagacccccc ctgggaatgg cctctggcca catcaggag	3480
ttccagatca ctgcctctgg ccagttatggc cagttggccc ccaagctggc caggtccac	3540
tacttctgtatgc ctggagcacc aaggagccat tcaagctggat caaagtggac	3600
ctgctggccc ccatgatcatt ccatggcatc aagacccagg gggccaggca gaagttctcc	3660
agcctgtaca tcagccagttt catcatcatg tacagctggat tggcaagaa atggcagacc	3720
tacagaggca actccactgg aacactcatg gtcttctttt gcaatgtggaa cagctctggc	3780
atcaaggcaca acatcttcaa ccccccatac atgccttgc acatcaggctt gcacccacc	3840
cactacagca tccgcagcac cctcaggatg gagctgtatgg gctgtgaccc gaaactcctgc	3900
agcatgcccc tggcatgga gagcaaggcc atttctgtatgg cccagatcac tgcctccagc	3960
tacttccatca acatgtttgc cacctggagc ccaagcaagg ccaggctgca cctccaggaa	4020
aggagcaatg cctggaggcc ccaggtcaac aacccaaagg agtggctgca ggtggacttc	4080
cagaagacca tgaaggtcac tgggtgacc acccaggggg tcaagagcct gctcaccagc	4140

-continued

atgttatgtga	aggagttcct	gatcagctcc	agccaggatg	gccaccagg	gacccttcc	4200
ttccagaatg	gcaaggtaa	ggtgtccag	ggcaaccagg	acagcttac	ccctgtgg	4260
aacagctgg	accccccct	cctgaccaga	tacctgagga	ttcacccca	gagctgg	4320
caccagattg	ccctgaggat	ggaggtcctg	ggatgtgagg	cccaggacct	gtactga	4377

<210> SEQ ID NO 99
<211> LENGTH: 4377
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 99						
atgcagattg	agctgagcac	ctgtttcttc	ctgtgcctgc	tgaggttctg	cttctctgcc	60
accaggagat	actacactggg	cgccgtggag	ctgagctggg	actacatgca	gtctgacactg	120
ggcgagctgc	ctgtggacgc	caggttcccc	cccagagtgc	ccaagagctt	ccccccta	180
acctca	gttacaagaa	gaccctgttc	gtggagttca	cogaccac	gttcaacatc	240
gccaagccca	ggcccccctg	gatgggcctg	ctggggccca	ccatccaggc	cgagggtgtac	300
gacaccegtgg	tgatcaccc	gaagaacatg	gccagccacc	ccgtgagcct	gcacgcccgt	360
ggcgtgagct	actggaaggc	ctctgaggc	gccgagttat	acgaccagac	cagccagagg	420
gagaaggagg	acgacaagg	gttccccggc	ggcagccaca	cctacgtgt	gcaggtgt	480
aggagaac	gccccatggc	cagegacccc	ctgtgcctga	cctacagcta	cctgagccac	540
gtggac	ctgg	tgaaggac	gaactctggc	ctgatcg	ccctgtgt	600
ggcagec	ctgg	ccaaggagaa	gaccagacc	ctgcaca	gttgc	660
ttcgatgagg	gcaagagctg	gcacagcgag	accaagaaca	gcctgatgca	ggacaggat	720
gccc	ctgg	ccagggctg	gcccaagat	cacaccgt	acggctac	780
ctgc	ccgg	tgateggctg	ccacagga	tctgt	actgtgt	840
accaccc	ccgg	agg	catcttctg	gagggccaca	ccttctgt	900
aggcaggc	ccgg	ggat	catcccata	ac	ccgtgt	960
gac	ccgg	ttt	cttctgt	cc	ccgtgt	1020
gcctac	ccgg	ttt	gtc	cc	ccgtgt	1080
gaggcc	ccgg	ttt	atc	cc	ccgtgt	1140
gat	ccgg	ttt	atc	cc	ccgtgt	1200
tggt	ccgg	ttt	atc	cc	ccgtgt	1260
cccg	ccgg	ttt	atc	cc	ccgtgt	1320
aagt	ccgg	ttt	atc	cc	ccgtgt	1380
atcc	ccgg	ttt	atc	cc	ccgtgt	1440
ctgat	ccgg	ttt	atc	cc	ccgtgt	1500
gat	ccgg	ttt	atc	cc	ccgtgt	1560
cccat	ccgg	ttt	atc	cc	ccgtgt	1620
accaag	ccgg	ttt	atc	cc	ccgtgt	1680

-continued

gacctggcct ctggccctgat cggccccctg ctgatctgtc acaaggagag cgtggaccag	1740
agggccaacc agatcatgtc tgacaagagg aacgtgtacc ttttcttgtt gttcgatgag	1800
aacaggagct ggttatctgac cgagaacatc cagaggttcc tgcccaaccc cgccggcgt	1860
cagctggagg accccgagtt ccaggccagc aacatcatgc acagcatcaa cggctacgt	1920
ttcgacagcc tgcagctgtc tgtgtgcctg cacgagggtgg cctactggta catctgagc	1980
atcgccgccc agaccgactt cctgtctgtt ttcttctctg gtcacacccaa caagcacaag	2040
atgggtgtacg aggacacccct gaccctgttc cccttcagc gcgagacccgt gttcatgagc	2100
atggagaacc ccggectgtg gatcctgggc tgccacaaca gcgacttcag gaacaggggc	2160
atgaccgccc tgctgaaagt cagcagctgc gacaagaaca cccggcacta ctacgaggac	2220
agctacgagg acatcagcgc ctacactgtc agcaagaaca acaccaccta cgtgaaccgc	2280
tccttgagcc agaacccttcc cgtgtgaag aggcaccaga gggagatcac caggaccacc	2340
ctgcagagcg accaggagga gatcgactat gatgacacca tcagcgtgga gatgaagaag	2400
gaggacttcg acatctacga cgaggacgag aaccagaccc ccaggagctt ccagaagaag	2460
accaggact acttcatcgc cccctgtggag aggctgtggg actatggcat gagcagcagc	2520
ccccaegtgc tgaggaacag ggcccaagac ggcagcgtgc cccagttcaa gaagggtgt	2580
ttccaggagt tcaccegacgg cagttcacc cagccctgt acagaggcga gctgaacgag	2640
cacctggccc tgctggggcc ctacatcagg gcccggatgg aggacaacat catggtacc	2700
ttcaggaacc aggccagcag gcccatacgc ttctacagca gcctgatcag ctacgaggag	2760
gaccagaggc agggcccccga gcccaggaa aacttcgtga agcccaacga gaccaagacc	2820
tacttctgga aggtgcagca ccacatggcc cccaccaagg acgagttcga ctgcaaggcc	2880
tgggcctact tctctgtatgt ggacctggag aaggacgtgc acagcggct gatggcccc	2940
ctgctgggtgt gccacaccaa caccctgaac cccgcccacg gcaggcaggt gaccgtcag	3000
gaggctgccccc tgttttcac catcttcgac gagaccaaga gctggactt caccgagaac	3060
atggagagga actgcagggc cccctgcaac atccagatgg aggacccac cttcaaggag	3120
aactacaggc tccacgcccataa caccggctac atcatggaca ccctgecccg cctgggtatg	3180
gcccaggacc agaggatcag gtggtatctg ctgagcatgg gcagcaacga gaacatccac	3240
agcatccact tcagcggcca cgtgttcacc gtgaggaaga aggaggagta caagatggcc	3300
ctgtacaacc tggatcccggtt cgtgttcag accgtggaga tgctgcccac caaggccggc	3360
atctggaggc tggatgtccct gatcggcgag cacctgcacg cccggcatgag caccctgttc	3420
ctgggttaca gcaacaagtcc accggccatgg ccagccggcca catcaggagc	3480
ttccagatca ccgcctctgg ccagttacggc cagtggccc ccaagctggc caggctgcac	3540
tacagcggca gcatcaacgc ctggagcacc aaggagccct tcagctggat caagggtggac	3600
ctgctggccc ccatgatcat ccacggcatc aagaccaggc ggcggaggca gaagttcagc	3660
gcctgtaca tcagccagtt catcatcatg tacagctggc acggcaagaa gtggcagacc	3720
tacaggggca acagcaccgg caccctgtat gtgttcttcg gcaacgtgga cagcggc	3780
atcaaggaca acatcttcaa ccccccattt atcggccaggat acatcaggct gcaacccacc	3840
cactacagca tcaggagcac cctgcggatg gaaactgtatgg gctgcgaccc gaaacagctgc	3900
agcatgcccc tggcatgga gagcaaggcc atctctgacg cccagatcac cgccaggcagc	3960

-continued

tacttcacca acatgttcgc cacctggagc cccagcaagg ccaggctgca cctgcagggc	4020
aggagcaacg cctggaggcc ccaggtgaac aaccccaagg agtggctgca ggtggacttc	4080
cagaagacca tgaaggtgac cggcggtgacc acccaggcg tgaagagcct gctgaccagc	4140
atgtacgtga aggagttcct gatcagcagc agccaggacg gccaccagtg gaccctgttc	4200
ttccagaacg gcaaagtgaa ggtgttccag ggcaaccagg acagttcac ccccggttg	4260
aacagcctgg accccccct gctgaccagg tatctgagga tccacccca gagctgggtg	4320
caccagatcg ccctgagaat ggaagtgtcg ggatgcgagg cccaggacct gtactga	4377

<210> SEQ ID NO 100
<211> LENGTH: 4377
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 100

atgcagattg agctgagcac ctgtttttc ctgtgcctgc tgagggtctg cttctctgcc	60
accaggagat actacctggg cgccgtggag ctgagctggg actacatgca gtctgacactg	120
ggcgagctgc ctgtggacgc caggttcccc cccagagtgc ccaagagctt ccccttcaac	180
acctcagtgg tgtacaagaa gaccctgttc gtggagttca ccgaccacct gttcaacatc	240
gccaagccca ggccccctg gatgggcctg ctggggccca ccatccaggc cgaggtgtac	300
gacaccgtgg tgatcaccct gaagaacatg gccagccacc ccgtgagcct gcacgcccgt	360
ggcgtgagct acttggaaaggc ctctgagggc gccgagttatg acgaccagac cagccagagg	420
gagaaggagg acgacaaggt gttccccggc ggcagccaca cctacgtgtg gcaggtgctg	480
aaggagaacg gccccatggc cagcgacccc ctgtgcctga cctacagcta cctgagccac	540
gtggacctgg tgaaggacct gaactctggc ctgatcgccg ccctgtgtt gtgcagggag	600
ggcagctgg ccaaggagaa gaccagacc ctgcacaagt tcatcctgtt gttcgccgt	660
ttcgatgagg gcaagagctg gcacagcgag accaagaaca gcctgtatgca ggacagggat	720
gcccgccttg ccagggcttg gcccaagatg cacaccgtga acggctacgt gaacaggagc	780
ctgccccggcc tgatcggttg ccacaggaag tctgtgtact ggcacgtgtat cggcatgggc	840
accacccccc aggtgcacag catcttcctg gagggccaca cttctgtt gaggaaaccac	900
aggcaggcca gcctggagat cagccccatc accttcctga ccccccagac cctgtgtatg	960
gacctggggcc agttctgtt gttctgccac atcagcagcc accagcacga cggcatggag	1020
gcctacgtga aggtggacag ctgccccgag gagccccagc tgaggatgaa gaacaacgag	1080
gaggccgagg actatgtatgta tgacctgacc gactctgtt gggacgtgtt gagggtttag	1140
gtatgacaaca gccccagctt catccagatc aggtctgtgg ccaagaagca ccccaagacc	1200
tggtgtcact acatcgccgc cgaggaggag gactggact acgccccctt ggtgtggcc	1260
cccgacgaca ggagctacaa gagccagttac ctgaacaacg gccccagag gatcgccagg	1320
aagtacaaga aggtcagatt catggctac accgacgaga cttcaagac cagggaggcc	1380
atccagcacc agtctggcat cctggggccc ctgtgtac gggagggtgg cgacaccctg	1440
ctgatcatct tcaagaacca ggccagcagg ccctacaaca tctacccca cggcatcacc	1500
gtatgtgggc ccctgtacag caggaggctg cccaaaggccg tgaagcacct gaaggacttc	1560

-continued

cccatectgc	ccggcagat	cttcaagtac	aagtggaccg	tgaccgtgga	ggatggcccc	1620
accaagtctg	accccagggtg	cctgaccagg	tactacagca	gcttcgtgaa	catggagagg	1680
gacctggccct	ctggcctgtat	cgccccctcg	ctgatctgct	acaaggagag	cgtggacacag	1740
aggggcaacc	agatcatgtc	tgacaagagg	aacgtgtacc	tgttctctgt	gttcgatgag	1800
aacaggagct	ggtatctgac	cgagaacatc	cagaggttcc	tgcccaaccc	cgcggcgctg	1860
cagctggagg	accccgagtt	ccaggccagc	aacatcatgc	acagcatcaa	cggctacgtg	1920
ttcgacagcc	tgcagctgtc	tgtgtgcctg	cacgagggtgg	cctactggta	catcctgagc	1980
atcggcgccc	agaccgactt	cctgtctgt	ttcttctctg	gctacacccct	caagcacaag	2040
atgggttacg	aggacacccct	gaccctgttc	cccttcagcg	gcgagaccgt	gttcatgagc	2100
atggagaacc	cggcctgtg	gatcctgggc	tgccacaaca	gcgacttcag	gaacaggggc	2160
atgaccggccc	tgctgaaagt	cagcagctgc	gacaagaaca	ccggcgacta	ctacgaggac	2220
agctacgagg	acatcagcgc	ctacctgctg	agcaagaaca	acaccaccta	cgtgaaccgc	2280
tccctgagcc	agaacccccc	cgtgctgaag	aggcaccaga	gggagatcac	caggaccacc	2340
ctgcagagcg	accaggagga	gatcactat	gatgacacca	tca gctgg	gatgaagaag	2400
gaggacttcg	acatctacga	cgaggacgag	aaccagagcc	ccaggagctt	ccagaagaag	2460
accaggcact	acttcatcgc	cgccgtggag	aggctgtgg	actatggcat	gagcagcagc	2520
ccccacgtgc	tgaggaacag	ggcccagacg	ggcagcgtgc	cccagtcaa	gaagggtgg	2580
ttccaggagt	tcaccgacgg	cagttcacc	cagcccctgt	acagaggcga	gctgaacag	2640
cacctgggcc	tgctggccccc	ctacatcagg	gccgaggatgg	aggacaacat	catggtgacc	2700
ttcaggaacc	aggccagcag	gccctacago	ttctacagca	gcctgatcag	ctacgaggag	2760
gaccagaggc	agggcgccga	gccccaggaag	aacttcgtga	agcccaacga	gaccaagacc	2820
tacttctgga	aggtgcagca	ccacatggcc	cccaccaagg	acgagttcga	ctgcaaggcc	2880
tgggcctact	tctctgtatgt	ggacctggag	aaggacgtgc	acagcggcct	gatcgcccc	2940
ctgctggtgt	gccacaccaa	caccctgaac	cccgccccacg	gcaggcaggt	gaccgtgcag	3000
gagttcgccc	tgttcttcac	catcttcgac	gagaccaaga	gctggactt	caccgagaac	3060
atggagagga	actgcagggc	cccctgcaac	atccagatgg	aggacccac	cttcaaggag	3120
aactacagg	tccacgccc	caacggctac	atcatggaca	ccctgcccc	cctggtgatg	3180
geccaggacc	agaggatcag	gtggtatctg	ctgagcatgg	gcagcaacga	gaacatccac	3240
agcatccact	tcageggcca	cgtgttcacc	gtgaggaaga	aggaggagta	caagatggcc	3300
ctgtacaacc	tgtaccccg	cgtgttcag	accgtggaga	tgctgcccc	caaggccggc	3360
atctggggg	tggagtgcct	gatcgccag	cacctgcacg	coggcatgag	caccctgttc	3420
ctggtgtaca	gcaacaagt	ccagacccccc	ctgggcattg	ccagcggcca	catcaggac	3480
ttccagatca	ccgcctctgg	ccagtagcgc	cagtggcccc	ccaagctggc	caggctgcac	3540
tacagcggca	gcatcaacgc	ctggagcacc	aaggagccct	tca gctggat	caagggtggac	3600
ctgctggccc	ccatgatcat	ccacggcatc	aagacccagg	gcccaggca	gaagttcagc	3660
agcctgtaca	tcagecagtt	catcatcatg	tacagctgg	acggcaagaa	gtggcagacc	3720
tacaggggca	acagcaccgg	caccctgtat	gtgttctcg	gcaacgtgga	cagcagccgc	3780
atcaaggaca	acatcttcaa	ccccccatc	atcgccaggt	acatcaggct	gcacccacc	3840

-continued

cactacagca tcaggagcac cctgcggatg	gaactgatgg gctgcgacct gaacagctc	3900
agcatcccc tggcatgga gagcaaggcc atctctgacg	cccagatcac cgccagcagc	3960
tacttcacca acatgttcgc cacctggacg	cccagcaagg ccaggctgca cctgcaggc	4020
aggagcaacg cctggaggcc ccaggtgaaac aaccccaagg	agtggctgca ggtggacttc	4080
cagaagacca tgaaggtgac cggcgtgacc acccaggcg	tgaagagcct gctgaccagc	4140
atgtacgtga aggagttcct gatcagcgcg	agccaggacg gccaccatg gaccctgttc	4200
ttccagaacg gcaaagtgaa ggtgttccag ggcaaccagg	acagcttcac cccctgttg	4260
aacagcctgg accccccct gctgaccagg tatctgagga	tccacccca gagctgggtg	4320
caccagatcg ccctgagaat ggaagtgctg	ggatgcgagg cccaggacct gtactga	4377

<210> SEQ_ID NO 101
<211> LENGTH: 4374
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 101

atgcagattg agctgagcac ctgtttttc	ctgtgcctgc tgaggttctg cttctctgcc	60
accaggagat actacctggg cgccgtggag	ctgagctggg actacatgca gtctgacactg	120
ggcgagctgc ctgtggacgc caggttcccc	cccagagtgc ccaagagctt ccccttcaac	180
acctcagtgg tgtacaagaa gaccctgttc	gtggagttca cccgaccacct gttcaacatc	240
gccaagccca ggccccctg gatgggcctg	ctggggccca ccatccaggc cgaggtgtac	300
gacacccgtgg tggtaaccct gaagaacatg	gccagccacc ccgtgagcct gcacgcgtg	360
ggcgtgagct acttggaaatc ctctgagggc	gccgagttatg acgaccagac cagccagagg	420
gagaaggagg acgacaaggt gttcccccgc	aagagccaca cctacgttg gcaggtgt	480
aaggagaacg gccccactgc cagegacccc	ccctgcctga cctacagcta cctgagccac	540
gtggacctgg tgaaggacct gaactctggc	ctgatcgccg ccctgttgtt gtgcaggag	600
ggcagectgg ccaaggagaa gacccagacc	ctgcacaagt tcatactgtt gtgcgcgtg	660
ttcgatgagg gcaagagctg gcacagcgt	accaagaaca gctgtatgca ggacaggat	720
gcccgccttg ccagggcctg gccaagatg	cacaccgtga acggctacgt gaacaggagc	780
ctgccccggcc tgatcggtt ccacaggaag	tctgtgtact ggcacgtgat cggcatggc	840
accaccccccgg aggtgcacag catttcttg	gagggccaca cttcttgtt gaggaaccac	900
aggcaggccca gcctggagat cagcccccattc	accttcctga cggcccagac cctgtgtatg	960
gacctggggcc agttcctgtt gttctgccac	atcagcagcc accagcacga cggcatggag	1020
gcctacgtga aggtggacag ctgccccgag	gagccccagc tgaggatgaa gaacaacag	1080
gaggccgagg actatgtatgca tgacctgacc	gactctgaga tggacgttgtt gaggtttgt	1140
gatgacaaca gccccagctt catccagatc	aggctgttg ccaagaagca ccccaagacc	1200
tgggtgcact acatcgccgc cgaggaggag	gactgggact acgccccctt ggtgtggcc	1260
cccgacgaca ggagctacaa gagccagttac	ctgaacaacg gccccagag gatcgccagg	1320
aagtacaaga aggtcagatt catggcctac	accgacgaga cttcaagac cagggaggcc	1380

-continued

atccagcacg agtctggcat cctggggccc ctgctgtacg gcgagggtgg cgacaccctg	1440
ctgatcatct tcaagaacca ggccagcagg ccctacaaca tctacccca cggcatcacc	1500
gatgtgaggc ccctgtacag caggaggctg cccaaggggcg tgaaggcacct gaaggacttc	1560
cccatctgc cggcggagat cttcaagtac aagtggaccg tgaccgtgga ggatggcccc	1620
accaagtctg accccagggtg cctgaccagg tactacagca gttctgtgaa catggagagg	1680
gacctggcct ctggcctgat cggcccccctg ctgatctgtc acaaggagag cgtggaccag	1740
aggggcaacc agatcatgtc tgacaagagg aacgtgtatcc ttgttctctgt gttcgtatgag	1800
aacaggagct ggttatctgac cgagaacatc cagaggttcc tgcccaaccc cgccggcgtg	1860
cagctggagg acccccgagtt ccaggccagc aacatcatgc acagcatcaa cggctacgtg	1920
ttcgacagcc tgcagctgtc tgtgtgcctg cacgaggtgg cctactggta catcctgagc	1980
atcggegccc agaccgactt cctgtctgtc ttcttctctg gtcacacccctt caagcacaag	2040
atgggttacg aggacaccctt gacccctgttc cccttcagcg gogagaccgt gttcatgagc	2100
atggagaacc cggcgcgttg gatectgggc tgccacaaca gogacttcag gaacaggggc	2160
atgaccgccc tgctgaaaagt cagcagctgc gacaagaaca coggcgacta ctacgaggac	2220
agctaegagg acatcagcgc ctacctgtcg agcaagaaca acgccatcga gcccaggagc	2280
ttcageccaga acccccccgt gctgaagagg caccagaggg agatcaccag gaccaccctg	2340
cagagegacc aggaggagat cgactatgt gacaccatca gcgtggagat gaagaaggag	2400
gacttegaca tctacgacga ggacgagaac cagagcccc ggagcttcca gaagaagacc	2460
aggcaactact tcataccgcg cgtggaggg ctgtggact atggcatgag cagcagcccc	2520
cacgtgtcga ggaacagggc ccagagcggc agcgtgcccc agttcaagaa ggtgggtttc	2580
caggagttca cggcggcag cttcacccag cccctgtaca gagggcgagct gaacgagcac	2640
ctgggcctgc tggcccccta catcaggggcc gaggtggagg acaacatcat ggtgacccct	2700
aggaaccagg ccageaggcc ctacagcttc tacagcagcc tgatcagcta cgaggaggac	2760
cagaggcagg ggcggcggcc caggaagaac ttctgtgaagc ccaacgagac caagacccat	2820
ttcttggaaagg tgcagcacca catggcccc accaaggacg agttcgactg caaggccctgg	2880
gcctacttct ctgtatgtgga cctggagaag gacgtgcaca gcccggctgtat cggccccctg	2940
ctgggtgtgcc acaccaacac cctgaacccc gcccacggc ggcagggtgac cgtgcaggag	3000
ttcgccctgt ttttaccat cttcgacgag accaagagct ggtacttcac cgagaacatg	3060
gagaggaact gcagggcccc ctgcaacatc cagatggagg accccacccctt caaggagaac	3120
tacaggttcc acggccatcaa cggctacatc atggacaccg tgccggccct ggtgtggcc	3180
caggaccaga ggatcagggtg gtatctgtcg agcatgggc gcaacgagaa catccacagc	3240
atccacttca gcccggcacgt gttcaccgtg aggaagaagg aggagtacaa gatggccctg	3300
tacaacctgt accccggcgt gttcgagacc gtggagatgc tgcccaagcaa ggccggcatc	3360
tggagggtgg agtgcctgat cggcggcgcac ctgcacccggc gcatggacac cctgttctg	3420
gtgtacagca acaagtgcga gaccccccgt ggcacatggca gcccggccat cagggacttc	3480
cagatcaccg cctctggcca gtacggccag tggggccccca agctggccag gctgcactac	3540
agcggcagca tcaacgcctg gagcaccaag gagccctca gctggatcaa ggtggacctg	3600
ctggccccca tgatcatcca cggcatcaag acccaggggcg ccaggcagaa gttcagcagc	3660

-continued

ctgtacatca	gccagttcat	catcatgtac	agcctggacg	gcaagaagtgc	3720
agggccaaca	gcacccggcac	cctgatggtg	ttttcgca	acgtggacag	3780
aagcacaaca	tcttcaaccc	ccccatcatc	gccaggtaca	tcaaggctgca	3840
tacagcatca	ggagcaccc	tcggatggaa	ctgatgggct	gctgacactgaa	3900
atgcacctgg	gcatggagag	caaggccatc	tctgacgccc	agatcaccgc	3960
ttcacaaca	tgttgcac	ctggagcccc	agcaaggcca	ggctgacact	4020
agcaacgcct	ggaggcccc	ggtgaacaac	cccaaggagt	ggctgcaggt	4080
aagaccatga	aggtgaccgg	cgtgaccacc	cagggcgtga	agagcctgct	4140
tacgtgaagg	agttctgtat	cagcagcgc	caggacggcc	accagtggac	4200
cagaacggca	aagtgaaggt	gttccagggc	aaccaggaca	gttccaccc	4260
agcctggacc	ccccctgtct	gaccaggat	ctgaggatcc	accccccagag	4320
cagatcgccc	tgagaatgga	agtgctggga	tgcgaggccc	aggacctgta	4374

<210> SEQ ID NO 102
<211> LENGTH: 4374
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 102

atgcagattg	agctgagcac	ctgttcttc	ctgtgcctgc	tgaggttctg	cttctctgcc	60
accaggagat	actacatgggg	cgccgtggag	ctgagctggg	actacatgca	gtctgacactg	120
ggcgagctgc	ctgtggacgc	caggttcccc	cccaagagtgc	ccaagagctt	cccttcaac	180
acctcagtgg	tgtacaagaa	gaccctgttc	gtggagttca	ccgaccacat	gttcaacatc	240
gccaagccca	ggcccccctg	gatgggcctg	ctggggccca	ccatccaggc	cgaggtgtac	300
gacaccgtgg	tgatcaccct	gaagaacatg	gccagccacc	ccgtgagcct	gcacgcccgt	360
ggcgtgagct	actggaaaggc	ctctgagggc	gccgagttatg	acgaccagac	cagccagagg	420
gagaaggagg	acgacaaggt	gttcccccgc	ggcagccaca	cctacgtgtg	gcaggtgt	480
aaggagaacg	gccccatggc	cagcgacccc	ctgtgcctga	cctacagcta	cctgagccac	540
gtggacctgg	tgaaggacct	gaactctggc	ctgatcggcg	ccctgtgtt	gtgcaggagg	600
ggcagectgg	ccaaggagaa	gaccaggacc	ctgcacaagt	tcatcctgt	gttcgcccgt	660
ttcgatgagg	gcaagagctg	gcacagcgc	accaagaaca	gcctgtatgca	ggacaggaggat	720
gcccctctg	ccagggcctg	gcccaagatg	cacaccgtga	acggctacgt	gaacaggaggc	780
ctgccccggcc	tgatcggtg	ccacaggaag	tctgtgtact	ggcacgtgt	cgccatgggc	840
accacccccc	aggtgcacag	catcttctg	gagggccaca	cttctgtgt	gaggaacac	900
aggcaggccc	gcctggagat	cagccccc	accttctga	ccgcccagac	cctgtgtat	960
gacctggggcc	agttcctgtct	gtcctgcccac	atcagcagcc	accagcacga	cgccatggag	1020
gcctacgtga	aggtggacag	ctgccccgag	gagccccagc	tgaggatgaa	gaacaacgag	1080
gaggccgagg	actatgtatg	tgacctgacc	gactctgaga	tggacgtgtt	gagggtttag	1140
gatgacaaca	gccccagctt	catccagatc	aggtctgtgg	ccaagaagca	ccccaaagacc	1200
tgggtgcact	acatcgccgc	cgaggaggag	gactgggact	acgccccct	ggtgtggcc	1260

-continued

cccgacgaca ggagctacaa gagecagtagt ctgaacaacg gccccagag gatggcagg	1320
aagtacaaga aggtcagatt catggctac accgacgaga cttcaagac cagggaggcc	1380
atccagcacg agtctggcat cctggggccc ctgctgtacg gcgagggtgg cgacacctg	1440
ctgtatct tcaagaacca ggccagcagg cctacaaca tctacccca cggcatcacc	1500
gatgtgaggc ccctgtacag caggaggctg cccaaggcg tgaagcacct gaaggactc	1560
ccatcttcgc cggcgagat cttcaagtac aagtggacc tgaccgtgga ggtggccccc	1620
accaagtctg accccaggtg cctgaccagg tactacagca gttcgtgaa catggaggg	1680
gacctggct ctggctgtat cggccctgtg ctgtatctgt acaaggagag cgtggacag	1740
aggggcaacc agatcatgtc tgacaagagg aacgtgatcc ttttctctgt ttcatgatgag	1800
aacaggagct ggtatctgac cgagaacatc cagaggttcc tgcccaaccc cgccggcgt	1860
cagctggagg accccaggtt ccaggccago aacatcatgc acagcatcaa cggctacgt	1920
ttcgacagcc tgcagctgtc tttgtgcctg cacgagggttg cttactggta catcctgagc	1980
atcgccgccc agaccgactt cctgtctgtt ttcttctctg gttacaccc caagcacaag	2040
atgggttacg aggacaccct gaccctgttc cccttacgtcg gcgagaccgt gttcatgagc	2100
atggagaacc cccgcctgtg gttcctggcc tgccacaaca gcaacttcag gaacaggggc	2160
atgaccgccc tgctgaaagt cagcagctgc gacaagaaca cccggacta ctacgaggac	2220
agctacgagg acatcagcgc ctacccgtcg agcaagaaca acgccatcga gccaggagc	2280
ttcagccaga accccccgt gctgaagagg caccagagg agatcaccag gaccaccctg	2340
cagagcggacc aggaggagat cgactatgtt gacaccatca gcgtggagat gaagaaggag	2400
gacttcgaca tctacgacga ggacgagaac cagagccca ggagcttcca gaagaagacc	2460
aggcactact tcatacgccgc cgtggaggagg ctgtggact atggcatgag cagcggccc	2520
cacgtgtga ggaacagggc ccagagcggc agcgtgcccc agttcaagaa ggtgggttcc	2580
caggagttca ccgcggcag ctccatcccg cccctgtaca gaggcgagct gaacgagcac	2640
ctggcctgc tggccctta catcaggggcc gaggtggagg acaacatcat ggtgaccctc	2700
aggaaccagg ccagcaggcc ctacagcttc tacagcagcc tgatcagcta cgaggaggac	2760
cagaggcagg gcgcggagcc caggaagaac ttctgtgaagc ccaacgagac caagacccatc	2820
ttcttggaaagg tgcagcacca catggccccc accaaggacg agttcgactg caaggccctgg	2880
gctacttct ctgtatgtgaa cctggagaag gacgtgcaca gggccctgtat cggccctgt	2940
ctgggtgtcc acaccaacac cctgaacccccc gcccacggca ggcagggtgac cgtcaggag	3000
ttcgccctgt ttttccat cttcgacgag accaaggagct ggtacttccat cgagaacatg	3060
gagaggaact gcaggggcccc ctgcacatc cagatggagg accccaccc caaggagaac	3120
tacagggttcc acgccccatcaa cggctacatc atggacaccct gttccggctt ggtgtatggcc	3180
caggaccaga ggatcagggtg gtatctgtcg agcatggca gcaacgagaa catccacagc	3240
atccacttca gcggccacgt gttcaccgtg aggaagaagg aggaggtaaa gatggccctg	3300
tacaacccgtt accccggcgt gttcgagacc gtggagatgc tgcccaagcaaa ggccggcattc	3360
tggagggtgg agtgcgtat cggcgagcac ctgcacggcc gcatgagcac cctgttccctg	3420
gtgtacagca acaagtgcac gaccccccgt ggcacatggccca gggccacat cagggactc	3480
cagatcaccg cctctggcca gtacggccag tggggccca agctggccag gctgcactac	3540

-continued

agcggcagca tcaaegcctg gagcaccaag gagccctca gctggatcaa ggtggacctg	3600
ctggccccca tgatcatcca cggcatcaag acccaggcgcc ccaaggcagaa gttcagcagc	3660
ctgtacatca gccagttcat catcatgtac agectggacg gcaagaagtgc cagacactac	3720
aggggcaaca gcacccggcac cctgtatggtg ttcttcggca acgtggacag cagcggcatc	3780
aagcacaaca tttcaaccc cccatcatc gccaggtaca tcaggctgca cccaccac	3840
tacagcatca ggagcacccct gcggatggaa ctgtatggct gcgacactgaa cagctgcagc	3900
atgccccctgg gcatggagag caaggccatc tctgacgccc agatcaccgc cagcagctac	3960
ttcacaaca tttcgccac ctggggcccc agcaaggccca ggctgcaccc gcaggccagg	4020
agcaacgcct ggaggccccca ggtgaacaac cccaggagt ggctgcaggt ggacttccag	4080
aagaccatga aggtgaccgg cgtgaccacc cagggcgtga agagcctgct gaccagcatg	4140
tacgtgaagg agttctgtat cagcagcagc caggacggcc accagtgac cctgttctc	4200
cagaacggca aagtgaaggt gttccaggcc aaccaggaca gttcacccccc cgtggtaaac	4260
agcctggacc ccccccgtct gaccaggatct ctgaggatcc acccccagag ctgggtgcac	4320
cagatcgcccc tgagaatgga agtgctggga tgcgaggccc aggacctgta ctga	4374

<210> SEQ ID NO 103
<211> LENGTH: 4377
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 103	
atgcagattt agctgagcac ctgtttcttc ctgtgcctgc tgaggctctg cttctctgcc	60
accaggagat actacctggg cgccgtggag ctgagctggg actacatgca gtctgactg	120
ggcgagatgc ctgtggacgc caggttcccc cccagatgc ccaagagctt ccccttcaac	180
acctcagtgg tgtacaagaa gaccctgttc gtggagttca ccgaccaccc gttcaacatc	240
gccaaggccca ggcccccctg gatggccctg ctggggccca ccatccaggc cgagggtgtac	300
gacaccgtgg tggcacccct gaagaacatg gccagccacc ccgtgagctt gcacgcccgt	360
ggcgtgagct actggaaagtc ctctgaggcc gcccggatgt acgaccagac cagccagagg	420
gagaaggagg acgacaagggt gttcccccggc aagagccaca cctacgtgt gcaggtgtc	480
aaggagaacg gccccactgc cagcgaccacc ccctgcctga cctacagcta cctgagccac	540
gtggacctgg tgaaggaccc tgaactctggc ctgatcgccg ccctgtgtt gtgcaggag	600
ggcagccctgg ccaaggagaa gacccagacc ctgcacaaggat tcatactgtt gtccgcgt	660
ttcgatgagg gcaagagctg gcacagcgag accaagaaca gcctgtatgca ggacaggat	720
gcccgcctcg ccagggcctg gcccaagatg cacaccgtga acggctacgt gaacaggagc	780
ctgccccggcc tgatcggtcg ccacagggaaat tctgtgtact ggcacgtat cggcatggc	840
accacccccc aggtgcacag catcttctcg gagggccaca ctttctgtt gaggaaccac	900
aggcaggccca gcctggagat cagccccatc accttctgtt ccgcccagac cctgtgtat	960
gacctggggcc agttcctgtt gttctgccac atcagcagcc accagcacga cggcatggag	1020
gcctacgtga aggtggacag ctgcggccag gagccccagc tgaggatgaa gaacaacag	1080

-continued

gaggccgagg actatgatga tgacctgacc gactctgaga tggacgtgg tgggttggat	1140
gatgacaaca gccccagctt catccagatc aggtctgtgg ccaagaagca ccccaagacc	1200
tgggtcaact acatcgccgc cgaggaggag gactgggact acggcccccct ggtgtggcc	1260
cccgacgaca ggagctacaa gagccagtagc ctgaacaacg gccccagag gatcgccagg	1320
aagtacaaga aggtcagatt catggctac accgacgaga cttcaagac cagggaggcc	1380
atccagcact agtctggcat cctggggccc ctgctgtacg gcgaggtggg cgacacctg	1440
ctgatcatct tcaagaacca ggccagcagg ccctacaaca tctacccca cggcatcacc	1500
gatgtgaggc ccctgtacag caggaggctg cccaaggccg tgaagcacct gaaggacttc	1560
ccatccctgc cggcgagat cttcaagtac aagtggaccg tgaccgtgga ggatggcccc	1620
accaagtctg accccaggtg cctgaccagg tactacagca gttcgtgaa catggagagg	1680
gacctggcct otggcctgat cggcccccctg ctgatctgat acaaggagag cgtggaccag	1740
agggccaacc agatcatgtc tgacaagagg aacgtgatcc ttttctctgt gttcgatgag	1800
aacaggagct ggtatctgac cgagaacatc cagaggttcc tgcccaaccc cgccggcgtg	1860
cagctggagg accccaggtt ccaggccagc aacatcatgc acagcatcaa cggctacgtg	1920
ttcgacagcc tgcagctgtc tttgtgcctg cacgaggtgg cctactggta catctgagc	1980
atcggegccc agacggactt cctgtctgtg ttttctctgt gtttacaccc caagcacaag	2040
atggtgtacg aggacacccct gacccctgttc cccttcagtg cggagaccgt gttcatgagc	2100
atggagaacc ccggcctgtg gatccctggc tgccacaaca gggacttcag gaacaggggc	2160
atgacggccc tgctgaaaatg cagcagctgc gacaagaaca cccggcacta ctacgaggac	2220
agctaegagg acatcagcgc ctacctgtc agcaagaaca acaccaccta cgtgaaccgc	2280
tccctgagcc agaacccttcc cgtgtcaag aggcaccaga gggagatcac caggaccacc	2340
ctgcagagcg accaggagga gatcgactat gatgacacca tcagcgtgga gatgaagaag	2400
gaggacttcg acatctacga cgaggacgag aaccagacg ccaggagctt ccagaagaag	2460
accaggcact acttcatcgc cgcgtggag aggctgtggg actatggcat gagcagcagc	2520
cccccaegtgc tgaggaacag ggcccaagac ggcagcgtgc cccagttcaa gaagggtgg	2580
ttccaggagt tcacccgacgg cagttcacc cagccctgt acagaggcga gctgaacgag	2640
cacctggccc tgctggccc ctacatcagg gccgaggtgg aggacaacat catggtacc	2700
ttcaggaacc aggccagcag gcccacacgc ttctacagca gctgtatcag ctacgaggag	2760
gaccagaggc agggcgcccga gcccaggaag aacttcgtga agcccaacga gaccaagacc	2820
tacttctgga aggtgcagca ccacatggcc cccaccaagg acgagttcga ctgcaaggcc	2880
tgggcctact tctctgtatgt ggacctggag aaggacgtgc acagcggcct gatcgccccc	2940
ctgctgggtgt gcccacccaa caccctgaac cccgccccacg gcaggcaggt gaccgtgcag	3000
gagttcgccc tttttttcac catcttcgac gagaccaaga gctggacttt caccgagaac	3060
atggagagga actgcgggac cccctgcaac atccagatgg aggacccac cttcaaggag	3120
aactacaggt tccacggccat caacggctac atcatggaca ccctgcccgg cctggatg	3180
gcccaaggacc agaggatcag gtggtatctg ctgagcatgg gcagcaacga gaacatccac	3240
agcatccact tcagcggccca cgtgttcacc gtgaggaaga aggaggagta caagatggcc	3300
ctgtacaacc tttttttccgg cgtgttcgag accgtggaga tgctgeccag caaggccggc	3360

-continued

atctggggg tggagtgcct gatcgccgag cacctgcacg ccggcatgag caccctttc	3420
ctgggttaca gcaacaagtg ccagaccccc ctgggcattgg ccagcggcca catcaggac	3480
ttccagatca ccgcctctgg ccagtcggc cagtggccc ccaagctggc caggctgcac	3540
tacagccgca gcatcaacgc ctggagcacc aaggagccct tcagctggat caaggtggac	3600
ctgctggccc ccatgatcat ccacggcata aagacccagg ggcgcaggca gaagttcagc	3660
agcctgtaca tcagccagtt catcatcatg tacagctgg acggcaagaa gtggcagacc	3720
tcacggggca acagcacccgg caccctgtat gtgttcttcg gcaacgtgga cagcagcggc	3780
atcaagcaca acatcttcaa cccccccatc atcgccaggt acatcaggct gcacccacc	3840
cactacagca tcaggagcac cctgcggatg gaactgtatgg gtcgcacct gaacagctgc	3900
agcatgcccc tggcatgga gagcaaggcc atctctgacg cccagatcac cgccagcagc	3960
tacttcacca acatgttcgc cacctggcgc cccagcaagg ccaggctgca cctgcaggc	4020
aggagcaacg octggaggcc ccaggtgaac aaccccaagg agtggctgca ggtggacttc	4080
cagaagacca tgaaggtgac cggcgtgacc acccaggccg tgaagagcct gtcgaccagc	4140
atgtacgtga aggagttctt gatcagcagc agccaggacg gocaccaggc gaccctgttc	4200
ttccagaacg gcaaagtgaa ggtgttccag ggcaaccagg acagcttac cccctgttg	4260
aacagectgg accccccctt gctgaccagg tatctgagga tecacccca gagctgggtg	4320
caccagatcg ccctgagaat ggaagtgcgtg ggatgcgagg cccaggact gtactga	4377

<210> SEQ ID NO 104
<211> LENGTH: 1458
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polypeptide

<400> SEQUENCE: 104

Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe
1 5 10 15

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

Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg
35 40 45

Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val
50 55 60

Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile
65 70 75 80

Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln
85 90 95

Ala Glu Val Tyr Asp Thr Val Val Thr Leu Lys Asn Met Ala Ser
100 105 110

His Pro Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ser Ser
115 120 125

Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp
130 135 140

Asp Lys Val Phe Pro Gly Lys Ser His Thr Tyr Val Trp Gln Val Leu
145 150 155 160

Lys Glu Asn Gly Pro Thr Ala Ser Asp Pro Pro Cys Leu Thr Tyr Ser

-continued

165	170	175
Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile		
180	185	190
Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr		
195	200	205
Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly		
210	215	220
Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp		
225	230	235
Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr		
245	250	255
Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val		
260	265	270
Tyr Trp His Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile		
275	280	285
Phe Leu Glu Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser		
290	295	300
Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met		
305	310	315
Asp Leu Gly Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His		
325	330	335
Asp Gly Met Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro		
340	345	350
Gln Leu Arg Met Lys Asn Asn Glu Glu Ala Asp Tyr Asp Asp Asp		
355	360	365
Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asn Ser		
370	375	380
Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr		
385	390	395
400		
Trp Val His Tyr Ile Ala Ala Glu Glu Asp Trp Asp Tyr Ala Pro		
405	410	415
Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn		
420	425	430
Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met		
435	440	445
Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu		
450	455	460
Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu		
465	470	475
480		
Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro		
485	490	495
His Gly Ile Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys		
500	505	510
Gly Val Lys His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe		
515	520	525
Lys Tyr Lys Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp		
530	535	540
Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg		
545	550	555
560		
Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu		
565	570	575

-continued

Ser Val Asp Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val
 580 585 590
 Ile Leu Phe Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu
 595 600 605
 Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp
 610 615 620
 Pro Glu Phe Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val
 625 630 635 640
 Phe Asp Ser Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp
 645 650 655
 Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe
 660 665 670
 Ser Gly Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr
 675 680 685
 Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro
 690 695 700
 Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly
 705 710 715 720
 Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp
 725 730 735
 Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys
 740 745 750
 Asn Asn Thr Thr Tyr Val Asn Arg Ser Leu Ser Gln Asn Pro Pro Val
 755 760 765
 Leu Lys Arg His Gln Arg Glu Ile Thr Arg Thr Thr Leu Gln Ser Asp
 770 775 780
 Gln Glu Glu Ile Asp Tyr Asp Asp Thr Ile Ser Val Glu Met Lys Lys
 785 790 795 800
 Glu Asp Phe Asp Ile Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser
 805 810 815
 Phe Gln Lys Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu
 820 825 830
 Trp Asp Tyr Gly Met Ser Ser Pro His Val Leu Arg Asn Arg Ala
 835 840 845
 Gln Ser Gly Ser Val Pro Gln Phe Lys Val Val Phe Gln Glu Phe
 850 855 860
 Thr Asp Gly Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu
 865 870 875 880
 His Leu Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp Asn
 885 890 895
 Ile Met Val Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe Tyr
 900 905 910
 Ser Ser Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu Pro
 915 920 925
 Arg Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys
 930 935 940
 Val Gln His His Met Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala
 945 950 955 960
 Trp Ala Tyr Phe Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly
 965 970 975

-continued

Leu	Ile	Gly	Pro	Leu	Leu	Val	Cys	His	Thr	Asn	Thr	Leu	Asn	Pro	Ala
980							985								990
<hr/>															
His	Gly	Arg	Gln	Val	Thr	Val	Gln	Glu	Phe	Ala	Leu	Phe	Phe	Thr	Ile
995							1000								1005
<hr/>															
Phe	Asp	Glu	Thr	Lys	Ser	Trp	Tyr	Phe	Thr	Glu	Asn	Met	Glu	Arg	
1010							1015								1020
<hr/>															
Asn	Cys	Arg	Ala	Pro	Cys	Asn	Ile	Gln	Met	Glu	Asp	Pro	Thr	Phe	
1025							1030								1035
<hr/>															
Lys	Glu	Asn	Tyr	Arg	Phe	His	Ala	Ile	Asn	Gly	Tyr	Ile	Met	Asp	
1040							1045								1050
<hr/>															
Thr	Leu	Pro	Gly	Leu	Val	Met	Ala	Gln	Asp	Gln	Arg	Ile	Arg	Trp	
1055							1060								1065
<hr/>															
Tyr	Leu	Leu	Ser	Met	Gly	Ser	Asn	Glu	Asn	Ile	His	Ser	Ile	His	
1070							1075								1080
<hr/>															
Phe	Ser	Gly	His	Val	Phe	Thr	Val	Arg	Lys	Lys	Glu	Glu	Tyr	Lys	
1085							1090								1095
<hr/>															
Met	Ala	Leu	Tyr	Asn	Leu	Tyr	Pro	Gly	Val	Phe	Glu	Thr	Val	Glu	
1100							1105								1110
<hr/>															
Met	Leu	Pro	Ser	Lys	Ala	Gly	Ile	Trp	Arg	Val	Glu	Cys	Leu	Ile	
1115							1120								1125
<hr/>															
Gly	Glu	His	Leu	His	Ala	Gly	Met	Ser	Thr	Leu	Phe	Leu	Val	Tyr	
1130							1135								1140
<hr/>															
Ser	Asn	Lys	Cys	Gln	Thr	Pro	Leu	Gly	Met	Ala	Ser	Gly	His	Ile	
1145							1150								1155
<hr/>															
Arg	Asp	Phe	Gln	Ile	Thr	Ala	Ser	Gly	Gln	Tyr	Gly	Gln	Trp	Ala	
1160							1165								1170
<hr/>															
Pro	Lys	Leu	Ala	Arg	Leu	His	Tyr	Ser	Gly	Ser	Ile	Asn	Ala	Trp	
1175							1180								1185
<hr/>															
Ser	Thr	Lys	Glu	Pro	Phe	Ser	Trp	Ile	Lys	Val	Asp	Leu	Leu	Ala	
1190							1195								1200
<hr/>															
Pro	Met	Ile	Ile	His	Gly	Ile	Lys	Thr	Gln	Gly	Ala	Arg	Gln	Lys	
1205							1210								1215
<hr/>															
Phe	Ser	Ser	Leu	Tyr	Ile	Ser	Gln	Phe	Ile	Ile	Met	Tyr	Ser	Leu	
1220							1225								1230
<hr/>															
Asp	Gly	Lys	Lys	Trp	Gln	Thr	Tyr	Arg	Gly	Asn	Ser	Thr	Gly	Thr	
1235							1240								1245
<hr/>															
Leu	Met	Val	Phe	Phe	Gly	Asn	Val	Asp	Ser	Ser	Gly	Ile	Lys	His	
1250							1255								1260
<hr/>															
Asn	Ile	Phe	Asn	Pro	Pro	Ile	Ile	Ala	Arg	Tyr	Ile	Arg	Leu	His	
1265							1270								1275
<hr/>															
Pro	Thr	His	Tyr	Ser	Ile	Arg	Ser	Thr	Leu	Arg	Met	Glu	Leu	Met	
1280							1285								1290
<hr/>															
Gly	Cys	Asp	Leu	Asn	Ser	Cys	Ser	Met	Pro	Leu	Gly	Met	Glu	Ser	
1295							1300								1305
<hr/>															
Lys	Ala	Ile	Ser	Asp	Ala	Gln	Ile	Thr	Ala	Ser	Ser	Tyr	Phe	Thr	
1310							1315								1320
<hr/>															
Asn	Met	Phe	Ala	Thr	Trp	Ser	Pro	Ser	Lys	Ala	Arg	Leu	His	Leu	
1325							1330								1335
<hr/>															
Gln	Gly	Arg	Ser	Asn	Ala	Trp	Arg	Pro	Gln	Val	Asn	Asn	Pro	Lys	
1340							1345								1350
<hr/>															
Glu	Trp	Leu	Gln	Val	Asp	Phe	Gln	Lys	Thr	Met	Lys	Val	Thr	Gly	

-continued

1355	1360	1365
Val Thr Thr Gln Gly Val Lys Ser Leu Leu Thr Ser Met Tyr Val		
1370	1375	1380
Lys Glu Phe Leu Ile Ser Ser Gln Asp Gly His Gln Trp Thr		
1385	1390	1395
Leu Phe Phe Gln Asn Gly Lys Val Lys Val Phe Gln Gly Asn Gln		
1400	1405	1410
Asp Ser Phe Thr Pro Val Val Asn Ser Leu Asp Pro Pro Leu Leu		
1415	1420	1425
Thr Arg Tyr Leu Arg Ile His Pro Gln Ser Trp Val His Gln Ile		
1430	1435	1440
Ala Leu Arg Met Glu Val Leu Gly Cys Glu Ala Gln Asp Leu Tyr		
1445	1450	1455

```

<210> SEQ_ID NO 105
<211> LENGTH: 1458
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
polypeptide

<400> SEQUENCE: 105

```

Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe			
1	5	10	15

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

Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg		
35	40	45

Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val		
50	55	60

Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile			
65	70	75	80

Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln		
85	90	95

Ala Glu Val Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser		
100	105	110

His Pro Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser		
115	120	125

Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp		
130	135	140

Asp Lys Val Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu			
145	150	155	160

Lys Glu Asn Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser		
165	170	175

Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile		
180	185	190

Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr		
195	200	205

Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly		
210	215	220

Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp			
225	230	235	240

-continued

Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr
 245 250 255

Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val
 260 265 270

Tyr Trp His Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile
 275 280 285

Phe Leu Glu Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser
 290 295 300

Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met
 305 310 315 320

Asp Leu Gly Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His
 325 330 335

Asp Gly Met Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro
 340 345 350

Gln Leu Arg Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp
 355 360 365

Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser
 370 375 380

Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr
 385 390 395 400

Trp Val His Tyr Ile Ala Ala Glu Glu Asp Trp Asp Tyr Ala Pro
 405 410 415

Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn
 420 425 430

Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met
 435 440 445

Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu
 450 455 460

Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu
 465 470 475 480

Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro
 485 490 495

His Gly Ile Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys
 500 505 510

Gly Val Lys His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe
 515 520 525

Lys Tyr Lys Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp
 530 535 540

Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg
 545 550 555 560

Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu
 565 570 575

Ser Val Asp Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val
 580 585 590

Ile Leu Phe Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu
 595 600 605

Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp
 610 615 620

Pro Glu Phe Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val
 625 630 635 640

Phe Asp Ser Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp

-continued

645	650	655
Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe		
660	665	670
Ser Gly Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr		
675	680	685
Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro		
690	695	700
Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly		
705	710	715
Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp		
725	730	735
Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys		
740	745	750
Asn Asn Thr Thr Tyr Val Asn Arg Ser Leu Ser Gln Asn Pro Pro Val		
755	760	765
Leu Lys Arg His Gln Arg Glu Ile Thr Arg Thr Thr Leu Gln Ser Asp		
770	775	780
Gln Glu Glu Ile Asp Tyr Asp Asp Thr Ile Ser Val Glu Met Lys Lys		
785	790	795
Glu Asp Phe Asp Ile Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser		
805	810	815
Phe Gln Lys Lys Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu		
820	825	830
Trp Asp Tyr Gly Met Ser Ser Pro His Val Leu Arg Asn Arg Ala		
835	840	845
Gln Ser Gly Ser Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe		
850	855	860
Thr Asp Gly Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu		
865	870	875
His Leu Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp Asn		
885	890	895
Ile Met Val Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe Tyr		
900	905	910
Ser Ser Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu Pro		
915	920	925
Arg Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys		
930	935	940
Val Gln His His Met Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala		
945	950	955
Trp Ala Tyr Phe Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly		
965	970	975
Leu Ile Gly Pro Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro Ala		
980	985	990
His Gly Arg Gln Val Thr Val Gln Glu Phe Ala Leu Phe Phe Thr Ile		
995	1000	1005
Phe Asp Glu Thr Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg		
1010	1015	1020
Asn Cys Arg Ala Pro Cys Asn Ile Gln Met Glu Asp Pro Thr Phe		
1025	1030	1035
Lys Glu Asn Tyr Arg Phe His Ala Ile Asn Gly Tyr Ile Met Asp		
1040	1045	1050

-continued

Thr Leu Pro Gly Leu Val Met Ala Gln Asp Gln Arg Ile Arg Trp
 1055 1060 1065
 Tyr Leu Leu Ser Met Gly Ser Asn Glu Asn Ile His Ser Ile His
 1070 1075 1080
 Phe Ser Gly His Val Phe Thr Val Arg Lys Lys Glu Glu Tyr Lys
 1085 1090 1095
 Met Ala Leu Tyr Asn Leu Tyr Pro Gly Val Phe Glu Thr Val Glu
 1100 1105 1110
 Met Leu Pro Ser Lys Ala Gly Ile Trp Arg Val Glu Cys Leu Ile
 1115 1120 1125
 Gly Glu His Leu His Ala Gly Met Ser Thr Leu Phe Leu Val Tyr
 1130 1135 1140
 Ser Asn Lys Cys Gln Thr Pro Leu Gly Met Ala Ser Gly His Ile
 1145 1150 1155
 Arg Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr Gly Gln Trp Ala
 1160 1165 1170
 Pro Lys Leu Ala Arg Leu His Tyr Ser Gly Ser Ile Asn Ala Trp
 1175 1180 1185
 Ser Thr Lys Glu Pro Phe Ser Trp Ile Lys Val Asp Leu Leu Ala
 1190 1195 1200
 Pro Met Ile Ile His Gly Ile Lys Thr Gln Gly Ala Arg Gln Lys
 1205 1210 1215
 Phe Ser Ser Leu Tyr Ile Ser Gln Phe Ile Ile Met Tyr Ser Leu
 1220 1225 1230
 Asp Gly Lys Lys Trp Gln Thr Tyr Arg Gly Asn Ser Thr Gly Thr
 1235 1240 1245
 Leu Met Val Phe Phe Gly Asn Val Asp Ser Ser Gly Ile Lys His
 1250 1255 1260
 Asn Ile Phe Asn Pro Pro Ile Ile Ala Arg Tyr Ile Arg Leu His
 1265 1270 1275
 Pro Thr His Tyr Ser Ile Arg Ser Thr Leu Arg Met Glu Leu Met
 1280 1285 1290
 Gly Cys Asp Leu Asn Ser Cys Ser Met Pro Leu Gly Met Glu Ser
 1295 1300 1305
 Lys Ala Ile Ser Asp Ala Gln Ile Thr Ala Ser Ser Tyr Phe Thr
 1310 1315 1320
 Asn Met Phe Ala Thr Trp Ser Pro Ser Lys Ala Arg Leu His Leu
 1325 1330 1335
 Gln Gly Arg Ser Asn Ala Trp Arg Pro Gln Val Asn Asn Pro Lys
 1340 1345 1350
 Glu Trp Leu Gln Val Asp Phe Gln Lys Thr Met Lys Val Thr Gly
 1355 1360 1365
 Val Thr Thr Gln Gly Val Lys Ser Leu Leu Thr Ser Met Tyr Val
 1370 1375 1380
 Lys Glu Phe Leu Ile Ser Ser Ser Gln Asp Gly His Gln Trp Thr
 1385 1390 1395
 Leu Phe Phe Gln Asn Gly Lys Val Lys Val Phe Gln Gly Asn Gln
 1400 1405 1410
 Asp Ser Phe Thr Pro Val Val Asn Ser Leu Asp Pro Pro Leu Leu
 1415 1420 1425

-continued

Thr Arg Tyr Leu Arg Ile His Pro Gln Ser Trp Val His Gln Ile
1430 1435 1440

Ala Leu Arg Met Glu Val Leu Gly Cys Glu Ala Gln Asp Leu Tyr
1445 1450 1455

<210> SEQ ID NO 106

<211> LENGTH: 1457

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polypeptide

<400> SEQUENCE: 106

Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe
1 5 10 15

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

Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg
35 40 45

Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val
50 55 60

Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile
65 70 75 80

Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln
85 90 95

Ala Glu Val Tyr Asp Thr Val Val Thr Leu Lys Asn Met Ala Ser
100 105 110

His Pro Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ser Ser
115 120 125

Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp
130 135 140

Asp Lys Val Phe Pro Gly Lys Ser His Thr Tyr Val Trp Gln Val Leu
145 150 155 160

Lys Glu Asn Gly Pro Thr Ala Ser Asp Pro Pro Cys Leu Thr Tyr Ser
165 170 175

Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile
180 185 190

Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr
195 200 205

Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly
210 215 220

Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp
225 230 235 240

Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr
245 250 255

Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val
260 265 270

Tyr Trp His Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile
275 280 285

Phe Leu Glu Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser
290 295 300

Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met
305 310 315 320

-continued

Asp Leu Gly Gln Phe Leu Leu Ser Cys His Ile Ser Ser His Gln His
 325 330 335

 Asp Gly Met Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro
 340 345 350

 Gln Leu Arg Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp
 355 360 365

 Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser
 370 375 380

 Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr
 385 390 395 400

 Trp Val His Tyr Ile Ala Ala Glu Glu Asp Trp Asp Tyr Ala Pro
 405 410 415

 Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn
 420 425 430

 Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met
 435 440 445

 Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu
 450 455 460

 Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu
 465 470 475 480

 Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro
 485 490 495

 His Gly Ile Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys
 500 505 510

 Gly Val Lys His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe
 515 520 525

 Lys Tyr Lys Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp
 530 535 540

 Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg
 545 550 555 560

 Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu
 565 570 575

 Ser Val Asp Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val
 580 585 590

 Ile Leu Phe Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu
 595 600 605

 Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp
 610 615 620

 Pro Glu Phe Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val
 625 630 635 640

 Phe Asp Ser Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp
 645 650 655

 Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe
 660 665 670

 Ser Gly Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr
 675 680 685

 Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro
 690 695 700

 Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly
 705 710 715 720

-continued

Met	Thr	Ala	Leu	Leu	Lys	Val	Ser	Ser	Cys	Asp	Lys	Asn	Thr	Gly	Asp
725						730									735
Tyr	Tyr	Glu	Asp	Ser	Tyr	Glu	Asp	Ile	Ser	Ala	Tyr	Leu	Ser	Lys	
		740				745									750
Asn	Asn	Ala	Ile	Glu	Pro	Arg	Ser	Phe	Ser	Gln	Asn	Pro	Pro	Val	Leu
755				760											765
Lys	Arg	His	Gln	Arg	Glu	Ile	Thr	Arg	Thr	Thr	Leu	Gln	Ser	Asp	Gln
770			775												780
Glu	Glu	Ile	Asp	Tyr	Asp	Asp	Thr	Ile	Ser	Val	Glu	Met	Lys	Lys	Glu
785			790					795							800
Asp	Phe	Asp	Ile	Tyr	Asp	Glu	Asp	Glu	Asn	Gln	Ser	Pro	Arg	Ser	Phe
	805					810									815
Gln	Lys	Lys	Thr	Arg	His	Tyr	Phe	Ile	Ala	Ala	Val	Glu	Arg	Leu	Trp
	820				825										830
Asp	Tyr	Gly	Met	Ser	Ser	Ser	Pro	His	Val	Leu	Arg	Asn	Arg	Ala	Gln
	835				840										845
Ser	Gly	Ser	Val	Pro	Gln	Phe	Lys	Lys	Val	Val	Phe	Gln	Glu	Phe	Thr
	850				855										860
Asp	Gly	Ser	Phe	Thr	Gln	Pro	Leu	Tyr	Arg	Gly	Glu	Leu	Asn	Glu	His
	865				870				875						880
Leu	Gly	Leu	Leu	Gly	Pro	Tyr	Ile	Arg	Ala	Glu	Val	Glu	Asp	Asn	Ile
		885				890									895
Met	Val	Thr	Phe	Arg	Asn	Gln	Ala	Ser	Arg	Pro	Tyr	Ser	Phe	Tyr	Ser
		900				905									910
Ser	Leu	Ile	Ser	Tyr	Glu	Glu	Asp	Gln	Arg	Gln	Gly	Ala	Glu	Pro	Arg
		915				920									925
Lys	Asn	Phe	Val	Lys	Pro	Asn	Glu	Thr	Lys	Thr	Tyr	Phe	Trp	Lys	Val
		930				935			940						
Gln	His	His	Met	Ala	Pro	Thr	Lys	Asp	Glu	Phe	Asp	Cys	Lys	Ala	Trp
	945				950				955						960
Ala	Tyr	Phe	Ser	Asp	Val	Asp	Leu	Glu	Lys	Asp	Val	His	Ser	Gly	Leu
		965				970									975
Ile	Gly	Pro	Leu	Leu	Val	Cys	His	Thr	Asn	Thr	Leu	Asn	Pro	Ala	His
		980				985									990
Gly	Arg	Gln	Val	Thr	Val	Gln	Glu	Phe	Ala	Leu	Phe	Phe	Thr	Ile	Phe
		995				1000									1005
Asp	Glu	Thr	Lys	Ser	Trp	Tyr	Phe	Thr	Glu	Asn	Met	Glu	Arg	Asn	
		1010			1015										1020
Cys	Arg	Ala	Pro	Cys	Asn	Ile	Gln	Met	Glu	Asp	Pro	Thr	Phe	Lys	
		1025			1030										1035
Glu	Asn	Tyr	Arg	Phe	His	Ala	Ile	Asn	Gly	Tyr	Ile	Met	Asp	Thr	
		1040			1045				1050						
Leu	Pro	Gly	Leu	Val	Met	Ala	Gln	Asp	Gln	Arg	Ile	Arg	Trp	Tyr	
		1055			1060										1065
Leu	Leu	Ser	Met	Gly	Ser	Asn	Glu	Asn	Ile	His	Ser	Ile	His	Phe	
		1070			1075				1080						
Ser	Gly	His	Val	Phe	Thr	Val	Arg	Lys	Lys	Glu	Glu	Tyr	Lys	Met	
		1085			1090										1095
Ala	Leu	Tyr	Asn	Leu	Tyr	Pro	Gly	Val	Phe	Glu	Thr	Val	Glu	Met	
		1100			1105										1110
Leu	Pro	Ser	Lys	Ala	Gly	Ile	Trp	Arg	Val	Glu	Cys	Leu	Ile	Gly	

-continued

1115	1120	1125
Glu His Leu His Ala Gly Met Ser Thr Leu Phe Leu Val Tyr Ser		
1130	1135	1140
Asn Lys Cys Gln Thr Pro Leu Gly Met Ala Ser Gly His Ile Arg		
1145	1150	1155
Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr Gly Gln Trp Ala Pro		
1160	1165	1170
Lys Leu Ala Arg Leu His Tyr Ser Gly Ser Ile Asn Ala Trp Ser		
1175	1180	1185
Thr Lys Glu Pro Phe Ser Trp Ile Lys Val Asp Leu Leu Ala Pro		
1190	1195	1200
Met Ile Ile His Gly Ile Lys Thr Gln Gly Ala Arg Gln Lys Phe		
1205	1210	1215
Ser Ser Leu Tyr Ile Ser Gln Phe Ile Ile Met Tyr Ser Leu Asp		
1220	1225	1230
Gly Lys Lys Trp Gln Thr Tyr Arg Gly Asn Ser Thr Gly Thr Leu		
1235	1240	1245
Met Val Phe Phe Gly Asn Val Asp Ser Ser Gly Ile Lys His Asn		
1250	1255	1260
Ile Phe Asn Pro Pro Ile Ile Ala Arg Tyr Ile Arg Leu His Pro		
1265	1270	1275
Thr His Tyr Ser Ile Arg Ser Thr Leu Arg Met Glu Leu Met Gly		
1280	1285	1290
Cys Asp Leu Asn Ser Cys Ser Met Pro Leu Gly Met Glu Ser Lys		
1295	1300	1305
Ala Ile Ser Asp Ala Gln Ile Thr Ala Ser Ser Tyr Phe Thr Asn		
1310	1315	1320
Met Phe Ala Thr Trp Ser Pro Ser Lys Ala Arg Leu His Leu Gln		
1325	1330	1335
Gly Arg Ser Asn Ala Trp Arg Pro Gln Val Asn Asn Pro Lys Glu		
1340	1345	1350
Trp Leu Gln Val Asp Phe Gln Lys Thr Met Lys Val Thr Gly Val		
1355	1360	1365
Thr Thr Gln Gly Val Lys Ser Leu Leu Thr Ser Met Tyr Val Lys		
1370	1375	1380
Glu Phe Leu Ile Ser Ser Gln Asp Gly His Gln Trp Thr Leu		
1385	1390	1395
Phe Phe Gln Asn Gly Lys Val Lys Val Phe Gln Gly Asn Gln Asp		
1400	1405	1410
Ser Phe Thr Pro Val Val Asn Ser Leu Asp Pro Pro Leu Leu Thr		
1415	1420	1425
Arg Tyr Leu Arg Ile His Pro Gln Ser Trp Val His Gln Ile Ala		
1430	1435	1440
Leu Arg Met Glu Val Leu Gly Cys Glu Ala Gln Asp Leu Tyr		
1445	1450	1455

```

<210> SEQ ID NO 107
<211> LENGTH: 1457
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polypeptide

```

-continued

<400> SEQUENCE: 107

Met Gln Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe
1 5 10 15

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

Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg
35 40 45

Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val
50 55 60

Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Asp His Leu Phe Asn Ile
65 70 75 80

Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln
85 90 95

Ala Glu Val Tyr Asp Thr Val Val Val Thr Leu Lys Asn Met Ala Ser
100 105 110

His Pro Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ser Ser
115 120 125

Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp
130 135 140

Asp Lys Val Phe Pro Gly Lys Ser His Thr Tyr Val Trp Gln Val Leu
145 150 155 160

Lys Glu Asn Gly Pro Thr Ala Ser Asp Pro Pro Cys Leu Thr Tyr Ser
165 170 175

Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile
180 185 190

Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr
195 200 205

Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly
210 215 220

Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp
225 230 235 240

Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr
245 250 255

Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val
260 265 270

Tyr Trp His Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile
275 280 285

Phe Leu Glu Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser
290 295 300

Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met
305 310 315 320

Asp Leu Gly Gln Phe Leu Leu Ser Cys His Ile Ser Ser His Gln His
325 330 335

Asp Gly Met Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro
340 345 350

Gln Leu Arg Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp
355 360 365

Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser
370 375 380

Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr

-continued

385	390	395	400
Trp Val His Tyr Ile Ala Ala Glu Glu Asp Trp Asp Tyr Ala Pro			
405	410	415	
Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn			
420	425	430	
Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met			
435	440	445	
Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu			
450	455	460	
Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu			
465	470	475	480
Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro			
485	490	495	
His Gly Ile Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys			
500	505	510	
Gly Val Lys His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe			
515	520	525	
Lys Tyr Lys Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp			
530	535	540	
Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg			
545	550	555	560
Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu			
565	570	575	
Ser Val Asp Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val			
580	585	590	
Ile Leu Phe Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu			
595	600	605	
Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp			
610	615	620	
Pro Glu Phe Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val			
625	630	635	640
Phe Asp Ser Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp			
645	650	655	
Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe			
660	665	670	
Ser Gly Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr			
675	680	685	
Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro			
690	695	700	
Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly			
705	710	715	720
Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp			
725	730	735	
Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys			
740	745	750	
Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser Gln Asn Pro Pro Val Leu			
755	760	765	
Lys Arg His Gln Arg Glu Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln			
770	775	780	
Glu Glu Ile Asp Tyr Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu			
785	790	795	800

-continued

Asp Phe Asp Ile Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe
 805 810 815

 Gln Lys Lys Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp
 820 825 830

 Asp Tyr Gly Met Ser Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln
 835 840 845

 Ser Gly Ser Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr
 850 855 860

 Asp Gly Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu His
 865 870 875 880

 Leu Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp Asn Ile
 885 890 895

 Met Val Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe Tyr Ser
 900 905 910

 Ser Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg
 915 920 925

 Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys Val
 930 935 940

 Gln His His Met Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala Trp
 945 950 955 960

 Ala Tyr Phe Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly Leu
 965 970 975

 Ile Gly Pro Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro Ala His
 980 985 990

 Gly Arg Gln Val Thr Val Gln Glu Phe Ala Leu Phe Phe Thr Ile Phe
 995 1000 1005

 Asp Glu Thr Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg Asn
 1010 1015 1020

 Cys Arg Ala Pro Cys Asn Ile Gln Met Glu Asp Pro Thr Phe Lys
 1025 1030 1035

 Glu Asn Tyr Arg Phe His Ala Ile Asn Gly Tyr Ile Met Asp Thr
 1040 1045 1050

 Leu Pro Gly Leu Val Met Ala Gln Asp Gln Arg Ile Arg Trp Tyr
 1055 1060 1065

 Leu Leu Ser Met Gly Ser Asn Glu Asn Ile His Ser Ile His Phe
 1070 1075 1080

 Ser Gly His Val Phe Thr Val Arg Lys Lys Glu Glu Tyr Lys Met
 1085 1090 1095

 Ala Leu Tyr Asn Leu Tyr Pro Gly Val Phe Glu Thr Val Glu Met
 1100 1105 1110

 Leu Pro Ser Lys Ala Gly Ile Trp Arg Val Glu Cys Leu Ile Gly
 1115 1120 1125

 Glu His Leu His Ala Gly Met Ser Thr Leu Phe Leu Val Tyr Ser
 1130 1135 1140

 Asn Lys Cys Gln Thr Pro Leu Gly Met Ala Ser Gly His Ile Arg
 1145 1150 1155

 Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr Gly Gln Trp Ala Pro
 1160 1165 1170

 Lys Leu Ala Arg Leu His Tyr Ser Gly Ser Ile Asn Ala Trp Ser
 1175 1180 1185

-continued

Thr	Lys	Glu	Pro	Phe	Ser	Trp	Ile	Lys	Val	Asp	Leu	Leu	Ala	Pro
1190						1195					1200			
Met	Ile	Ile	His	Gly	Ile	Lys	Thr	Gln	Gly	Ala	Arg	Gln	Lys	Phe
1205						1210					1215			
Ser	Ser	Leu	Tyr	Ile	Ser	Gln	Phe	Ile	Ile	Met	Tyr	Ser	Leu	Asp
1220						1225					1230			
Gly	Lys	Lys	Trp	Gln	Thr	Tyr	Arg	Gly	Asn	Ser	Thr	Gly	Thr	Leu
1235						1240					1245			
Met	Val	Phe	Phe	Gly	Asn	Val	Asp	Ser	Ser	Gly	Ile	Lys	His	Asn
1250						1255					1260			
Ile	Phe	Asn	Pro	Pro	Ile	Ile	Ala	Arg	Tyr	Ile	Arg	Leu	His	Pro
1265						1270					1275			
Thr	His	Tyr	Ser	Ile	Arg	Ser	Thr	Leu	Arg	Met	Glu	Leu	Met	Gly
1280						1285					1290			
Cys	Asp	Leu	Asn	Ser	Cys	Ser	Met	Pro	Leu	Gly	Met	Glu	Ser	Lys
1295						1300					1305			
Ala	Ile	Ser	Asp	Ala	Gln	Ile	Thr	Ala	Ser	Ser	Tyr	Phe	Thr	Asn
1310						1315					1320			
Met	Phe	Ala	Thr	Trp	Ser	Pro	Ser	Lys	Ala	Arg	Leu	His	Leu	Gln
1325						1330					1335			
Gly	Arg	Ser	Asn	Ala	Trp	Arg	Pro	Gln	Val	Asn	Asn	Pro	Lys	Glu
1340						1345					1350			
Trp	Leu	Gln	Val	Asp	Phe	Gln	Lys	Thr	Met	Lys	Val	Thr	Gly	Val
1355						1360					1365			
Thr	Thr	Gln	Gly	Val	Lys	Ser	Leu	Leu	Thr	Ser	Met	Tyr	Val	Lys
1370						1375					1380			
Glu	Phe	Leu	Ile	Ser	Ser	Gln	Asp	Gly	His	Gln	Trp	Thr	Leu	
1385						1390					1395			
Phe	Phe	Gln	Asn	Gly	Lys	Val	Lys	Val	Phe	Gln	Gly	Asn	Gln	Asp
1400						1405					1410			
Ser	Phe	Thr	Pro	Val	Val	Asn	Ser	Leu	Asp	Pro	Pro	Leu	Leu	Thr
1415						1420					1425			
Arg	Tyr	Leu	Arg	Ile	His	Pro	Gln	Ser	Trp	Val	His	Gln	Ile	Ala
1430						1435					1440			
Leu	Arg	Met	Glu	Val	Leu	Gly	Cys	Glu	Ala	Gln	Asp	Leu	Tyr	
1445						1450					1455			

```

<210> SEQ ID NO 108
<211> LENGTH: 4374
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
      polynucleotide

```

<400> SEQUENCE: 108

atgcagattg	agctgtccac	ctgtttcttt	ctgtgcctgc	ttagattctg	cttctctgcc	60
accaggagat	actacctggg	ggctgtggaa	ctttcttgggg	actacatgca	gtctgacactg	120
ggagagctgc	ctgtggatgc	caggttccca	cccagagtgc	ccaagtcctt	cccatcaac	180
acctctgtgg	tctacaagaa	gacactcttt	gtggattca	ctgaccacct	gttcaacatt	240
gaaaaaccca	gaccaccctg	gatgggactc	ctgggaccctt	ccattcaggc	tgagggttat	300
gacactgtgg	tcgtcacccct	caagaacatg	gcatcccacc	ctgtgtctct	gcatgctgt	360

-continued

ggagtctcat	actggaaatc	ctctgaaggg	gctgagttatg	atgaccagac	atcccagaga	420
gagaaaaggagg	atgacaagggt	gttccctggg	aagtctcacca	cctatgtgtg	gcaagtctc	480
aaggagaatg	gaccactgc	atctgaccca	ccctgcctga	catactccctaa	cctttctcat	540
gtggacctgg	tcaaggacct	caactcttga	ctgattgggg	cactgttgtt	gtgcagggaa	600
ggatccctgg	ccaaggagaa	aaccagaca	ctgcacaatgt	tcattctctt	gtttgtgtc	660
tttgatgagg	gcaagtcttgc	gcactcttga	acaaagaact	ccctgtatgc	agacaggatgt	720
gtcgctctg	ccagggcatg	gcccaagatg	cacactgtga	atggctatgt	gaacagatca	780
ctgcctggac	tcatggctg	ccacaggaaa	tctgtctact	ggcatgtgt	tggcatgggg	840
acaaccctg	aagtgcactc	catttcctg	gagggacaca	ccttcctgg	caggaaccac	900
agacaaggct	ctctggagat	ctctccatc	accttcctca	ctgcacagac	actgtgtatg	960
gaccctggac	agttctgtct	gtcctgccac	atctcttccc	accagcatga	tggcatggaa	1020
gcctatgtca	aggtggactc	atgcctgtgag	gaaccacagc	tcaaggatgaa	gaacaatgag	1080
gaggctgagg	actatgtatg	tgacactgact	gactctgaga	tggatgttgtt	cagatttgat	1140
gtgacaact	ctccatcctt	cattcagatc	aggctgtgtgg	caaagaaaca	ccccaaagaca	1200
tgggtgcact	acattgctgc	tgaggaagag	gactgggact	atgcaccact	ggtcctggcc	1260
cctgatgaca	ggagctacaa	gtctcagtagc	ctcaacaatg	gcccacaaag	aattggaaaga	1320
aagtacaaga	aagtcaaggatt	catggcctac	actgtatgaaa	ccttcaagac	aagagaagcc	1380
attcagcatg	agtctggcat	tctgggacca	ctcctgtatg	gggaagtggg	agacaccctg	1440
ctccatcatct	tcaagaacca	ggcctccagg	ccctacaaca	tctaccacca	tggcatca	1500
gatgtcaggc	ccctgtacag	caggagactg	ccaaaagggg	tgaaacacct	caaggacttc	1560
cccattctgc	ctggagagat	cttcaagtagc	aagtggactg	tcaactgtgga	ggatggacca	1620
acaaagtctg	accccagggt	cctcaccaga	tactactct	cttttgtgaa	catggagaga	1680
gacctggcat	ctggactgtat	tggaccactg	ctcatctgt	acaaggagtc	tgtggaccag	1740
agaggcaacc	agatcatgtc	tgacaagaga	aatgtgattc	tgttctctgt	ctttgtatgag	1800
aacagatcat	ggtacctgac	tgagaacatt	cagagattcc	tgcccaaccc	tgctggggtg	1860
caactggaaag	accctgagtt	ccaggcaagc	aacatcatgc	actccatcaa	tggctatgt	1920
tttgactctc	tccagtttc	tgtctgcctg	catgagggtgg	cctactggta	cattcttct	1980
attggggcac	aaactgactt	cctttctgtc	ttttctctgt	gatacacctt	caagcacaag	2040
atgggtgtatg	aggacaccct	gacactcttc	ccattctctg	gggaaactgt	gttcatgagc	2100
atggagaacc	ctggactgtg	gattctggga	tgccacaact	ctgacttcag	aaacaggggaa	2160
atgactgcac	tgctcaaaatg	ctcctctgt	gacaagaaca	ctggggacta	ctatgaggac	2220
tcttatgagg	acatctctgc	ctacactgctc	agcaagaaca	atgcccattga	gcccaagaagc	2280
ttctctcaga	atccacactgt	cctgaagaga	caccagagag	agatcaccag	gacaaccctc	2340
cagtctgacc	aggaagagat	tgactatgtat	gacaccattt	ctgtggagat	gaagaaggag	2400
gactttgaca	tctatgtatg	ggacgagaac	cagtctccaa	gatcattcca	gaagaagaca	2460
agacactact	tcattgtctgc	tgtggaaaga	ctgtgggact	atggcatgtc	ttcccttccc	2520
catgtcctca	ggaacagggc	acagtctggc	tctgtccac	agttcaagaa	agtggcttc	2580
caggagttca	ctgatggctc	attcacccag	cccctgtaca	gagggaaact	gaatgagcac	2640

-continued

ctgggactcc	tgggaccata	catcagggt	gaggtggaa	acaacatcat	ggtgacattc	2700
agaaaaccagg	cctccaggcc	ctacagcttc	tactcttccc	tcatcagcta	tgaggaagac	2760
cagagacaag	gggctgagcc	aagaagaac	tttgtgaaac	ccaatgaaac	caagacctac	2820
ttctggaaag	tccagcacca	catggcaccc	accaaggatg	agtttgactg	caaggcctgg	2880
gcataacttct	ctgatgtgga	cctggagaaa	gatgtgcact	ctggcctgat	tggcccactc	2940
ctggctcgtcc	acaccaaacac	cctgaaccct	gcacatggaa	ggcaagtgac	tgtgcaggag	3000
tttgcctct	tcttcaccat	ctttgatgaa	accaagtcat	ggtacttcac	tgagaacatg	3060
gagagaaaact	gcagagcacc	atgcaacatt	cagatggaa	accccaccc	caaggagaac	3120
tacagggttcc	atgccatcaa	tggctacatc	atggacaccc	tgcctggct	tgtcatggca	3180
caggaccaga	gaatcagatg	gtacactgctt	tctatggat	ccaatgagaa	cattcactcc	3240
atccacttct	ctggcatgt	cttcactgtg	agaaagaagg	aggaatacaa	gatggccctg	3300
tacaacctct	accctgggg	ctttgagact	gtggagatgc	tgcctccaa	agctggcatac	3360
tggaggggtgg	aatgcctcat	tggggagcac	ctgcatgctg	gcatgtcaac	cctgttcctg	3420
gtctacagca	acaagtgcac	gacacccctg	ggaatggct	ctggccacat	cagggacttc	3480
cagatcactg	cctctggcca	gtatggccag	tgggcaccca	aactggccag	gctccactac	3540
tctggctcca	tcaatgcatg	gtcaaccaag	gagccattct	cttggatcaa	ggtggactg	3600
ctggcaccca	tgatcattca	tggcatcaag	acacaggggg	caagacagaa	attctccct	3660
ctgtacatct	cacagttcat	catcatgtac	tctctggatg	gcaagaagtg	gcagacatac	3720
agaggcaact	ccactggcac	cctcatggtc	ttctttggca	atgtggacag	ctctggcatc	3780
aagcacaaca	tcttcaaccc	tcccatcatt	gccagataca	tcaggctgca	ccccacccac	3840
tactcaatca	gatcaaccc	caggatggaa	ctgatggat	gtgacctgaa	ctcctgctca	3900
atgcccctgg	gaatggagag	caaggccatt	tctgatgcc	agatcactgc	atcctttac	3960
ttcaccaaca	tgtttgccac	ctggtcacca	tcaaaagcca	ggctgcaccc	ccagggaaaga	4020
agcaatgcct	ggagacccca	ggtaacaacac	ccaaaggaat	ggctgcaagt	ggacttcag	4080
aagacaatga	aagtcaactgg	ggtgacaacc	cagggggtca	agtctctgct	cacctcaatg	4140
tatgtgaagg	agttcctgtat	ctcttcctca	caggatggcc	accagtggac	actttcttc	4200
cagaatggca	aagtcaaggt	gttccaggc	aaccaggact	ctttcacacc	tgtggtaac	4260
tcactggacc	ccccctctt	gacaagatac	ctgagaattc	accccccagtc	ttgggtccac	4320
cagattgccc	tgagaatgga	agtcctggga	tgtgaggcac	aagacctgta	ctga	4374

<210> SEQ ID NO 109
<211> LENGTH: 4374
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 109

atgcagattg	agctgagcac	ctgtttcttc	ctgtgcctgc	tgaggttctg	cttctctgcc	60
accaggagat	actacctggg	ggctgtggag	ctttcttggg	actacatgca	gtctgacctg	120
ggggagctgc	ctgtggatgc	caggttccca	cccagagtgc	ccaaatccct	cccatcaac	180

-continued

acctctgtgg tctacaagaa gaccctctt gtggagttca ctgaccacat gttcaacatt	240
gccaaaccca ggccaccctg gatgggactc ctgggaccca ccattcaggc tgaggttat	300
gacactgtgg tcgtcacccct caagaacatg gcctcccacc ctgtgagcct gcatgttg	360
ggggtcagct actggaagtc ctctgagggg gctgagttatg atgaccagac ctcccagagg	420
gagaaggagg atgacaaaatg ttccctggg aagagccaca cctatgttg gcaggtctc	480
aaggagaatg gccccactgc ctctgaccca ccctgcctga cctactccta cctttctcat	540
gtggacctgg tcaaggaccc caactctggc ctgattgggg ccctgttgtt gtgcagggag	600
ggctccctgg ccaaagagaa gaccagacc ctgcacaagt tcatttcctt gtttgcgttc	660
tttgatgagg gcaagagctg gcactctgaa accaagaact ccctgtatgc ggacagggat	720
gtgcctctg ccagggctg gcccaagatg cacactgtga atggctatgt gaacaggagc	780
ctgcctggac tcattggctg ccacaggaaa tctgtctact ggcatgtatg tggcatgggg	840
acaacccctg aggtgcactc catttcctg gagggccaca ctttccttgtt caggaaccac	900
agacaggcca gcctggagat cagccccatc accttcctca ctgcccagac cctgtgtatg	960
gacctcgac agttctgtt gtccctgccat atcagctccc accagcatga tggcatggag	1020
gcctatgtca aggtggacag ctgcctctg gagccacagc tcaggatgaa gaacaatgag	1080
gaggctgagg actatgtatg tgacctgact gactctgaga tggatgttgtt ccgccttgat	1140
gtgacaaca gccccatctt cattcagatc aggtctgtgg ccaagaaaca ccccaagacc	1200
tgggtgact acattgtgc tgaggaggag gactggact atgcctact ggtcctggcc	1260
cctgatgaca ggagctacaa gagccagttc ctcaacaatg gcacccacatg gattggacgc	1320
aagtacaaga aagttaggtt catggctac actgtatgaa cttcaagac cagggaggcc	1380
attcagcatg agtctggcat cctggccca ctctgtatg gggaggtggg ggacaccctg	1440
ctcatcatct tcaagaacca ggcctccagg ccctacaaca tctaccacata tggcatact	1500
gtatgtaccc ccctgtacat ccgcaggctg ccaaaaggggg tgaaacaccc caaggacttc	1560
cccatctgc ctggggagat cttcaagtac aagtggactg tcactgtgaa ggatggacca	1620
accaaatactg accccagggtg cctcaccaga tactactcca gtttgcgtt catggagagg	1680
gacctggct ctggcctgtat tgcccactg ctcatctgtt acaaggagtc tggatggacag	1740
agggggaaacc agatcatgtc tgacaagagg aatgtgattt tggatgttgtt ctggatgag	1800
aacaggagct ggtacctgac tgagaacatt cagcgcttcc tgcccaaccc tgctgggtt	1860
cagctggagg acccttagtt ccaggccagc aacatcatgc actccatcaa tggatgttg	1920
tttgacagcc tccagcttcc tggatgttgtt catggatgtt cttactgtt cattttct	1980
attggggccct agactgactt cttttctgtc ttcttctctg gctacacccctt caaacacaag	2040
atgggtgtatg aggacaccctt gaccctcttc ccattctgtt gggagactgtt gttcatgagc	2100
atggagaacc ctggcctgtt gattctggga tgccacaact ctgacttccg caacaggggc	2160
atgactgccc tgctcaaatg ctccctctgt gacaagaaca ctggggacta ctatgaggac	2220
agctatgagg acatctctgc ctacactgtc agcaagaaca atgcccatttga gcccaggagc	2280
ttcagccaga atccacctgt cctgaaacccg caccagaggag agatcaccatg gaccaccctc	2340
cagtctgacc aggaggagat tgactatgtt gacaccattt ctgtggagat gaagaagag	2400
gactttgaca tctatgacga ggacgagaac cagagccaa ggagcttcca gaagaagacc	2460

-continued

aggcaactact tcattgtgc tggggact atggcatgag ctccagcccc	2520
catgtcctca ggaacagggc ccagtctggc tctgtgccac agttcaagaa agtggcttc	2580
caagagttca ctgatggcag cttcacccag cccctgtaca gaggggagct gaatgagcac	2640
ctgggactcc tgggcccata catcagggtc gaggtggagg acaacatcat ggtgacctc	2700
cgcaaccagg cctccaggcc ctacagcttc tacagctccc tcatcagcta tgaggaggac	2760
cagaggcagg gggctgagcc acgcaagaac tttgtgaaac ccaatgaaac caagacctac	2820
ttctggaaag tccagcacca catggccccc accaaggatg agtttgactg caaggcctgg	2880
gcctacttct ctgatgtgga cctggagaag gatgtgcact ctggcctgat tggccactc	2940
ctggctgccc acaccaacac cctgaaccct gcccatggaa ggcaagtgac tgtgcaggag	3000
tttgcctct tttcaccat ctttgatgaa accaagagct ggtacttcac tgagaacatg	3060
gagcgcacta gcagggccccc atgcaacatt cagatggagg accccaccccaaagagaac	3120
taccgcttcc atgcccataa tggctacatc atggacaccc tgcctggct tgcacatggcc	3180
caggaccaga ggtcagggtg gtacctgctt tctatggctt ccaatgagaa cattcactcc	3240
atccacttct ctggcatgtt cttcactgtg cgcaagaagg aggagtacaa gatggcctg	3300
tacaacccctt accctgggggtt ctttgagact gtggagatgc tgcctccaa agctggcatc	3360
tggagggtgg agtgcctcat tggggagcac ctgcacatgc gcatgacac cctgttcctg	3420
gtctacagca acaagtgcac gaccccccgtt ggaatggcct ctggccacat cagggactc	3480
cagatcactg cctctggcca gtatggccag tggggccccc agctggccag gctccactac	3540
tctggatcca tcaatgcctg gagcaccaag gagccattca gctggatcaa agtggacactg	3600
ctggcccccata tgatcatcca tggcatcaag acccagggggg ccaggcagaa gttctccagc	3660
ctgtacatca gccaggatcat catcatgtac agcctggatg gcaagaaatg gcagacctac	3720
agaggcaact ccactggAAC actcatggTC ttctttggCA atgtggacAG ctctggcatc	3780
aagcacaaca tttcaaccc cccaaatcatc gccagataca tcaggctgca ccccccac	3840
tacagcatcc gcagcacccct caggatggag ctgatgggtt gtgacctgaa ctcctgcagc	3900
atggccctgg gcatggagag caaggccatt tctgtatggcc agatcactgc ctccagctac	3960
ttcaccaaca tgtttgcac ctggggccca agcaaggcca ggctgcaccc ccaggaaagg	4020
agcaatgcct ggaggccccc ggtcaacaac ccaaaggagt ggctgcagggt ggacttccag	4080
aagaccatga aggtcaactgg ggtgaccacc caggggtca agagcctgct caccagcatg	4140
tatgtgaagg agttcctgat cagctccagc caggatggcc accagtggac cctcttc	4200
cagaatggca aggtcaaggt gttccaggc aaccaggaca gttcaccccc tgggtgaac	4260
acccctggacc ccccccctt gaccagatac ctgaggattc accccagag ctgggtccac	4320
cagattggcc tgaggatgga ggtcctggga tggaggccc aggacctgta ctga	4374

1. A polynucleotide comprising the nucleotide sequence of SEQ ID NO: 13, wherein the polynucleotide encodes a Factor VIII polypeptide.

2-65. (canceled)

66. The polynucleotide of claim **1**, further comprising a promoter element operably linked to the polynucleotide encoding the Factor VIII polypeptide.

67. The polynucleotide of claim **66**, wherein the promoter element is a liver-specific promoter sequence upstream of the nucleotide sequence encoding the Factor VIII polypeptide.

68. The polynucleotide of claim **67**, further comprising an intron sequence positioned between the liver-specific promoter sequence and the nucleotide sequence encoding the Factor VIII polypeptide.

69. An adeno-associated virus (AAV) vector comprising a polynucleotide of claim **1**.

70. An adeno-associated virus (AAV) particle comprising a polynucleotide of claim **1**.

71. A host cell infected with an adeno-associated virus (AAV) particle comprising a polynucleotide of claim **1**.

72. A method for producing an adeno-associated virus (AAV) particle comprising introducing a polynucleotide of claim **1** into a mammalian host cell, wherein the polynucleotide is competent for replication in the mammalian host cell.

73. A method for treating hemophilia A comprising administering, to a patient in need thereof, an adeno-associated virus (AAV) particle according to claim **70**.

74. A method for transducing a host cell comprising contacting the host cell with an adeno-associated virus (AAV) particle according to claim **70**.

* * * * *