<110> TOOLGEN INCORPORATED

<120> Composition for cleaving a target DNA comprising a guide RNA specific for the target DNA and Cas protein-encoding nucleic acid or Cas protein, and use thereof

<130> OPA13141PCT

<150> US 61/717,324
<151> 2012-10-23

<150> US 61/803,599
<151> 2013-03-20

<150> US 61/837,481
<151> 2013-06-20

<160> 111

<170> KopatentIn 2.0

<210> 1
<211> 4107
<212> DNA
<213> Artificial Sequence

<220>
<223> Cas9-coding sequence

<400> 1
atggacaaga agtacagcat cggcctggac atcggtacca acagcgtggg ctgggccgtg 60

atcaccgacg agtacaaggt gcccagcaag aagttcaagg tgctgggcaa caccgaccgc 120

cacagcatca agaagaacct gatcggcgcc ctgctgttcg acagcggcga gaccgccgag 180

gccacccgcc tgaagcgcac cgcccgccgc cgctacaccc gccgcaagaa ccgcatctgc 240

tacctgcagg agatcttcag caacgagatg gccaaggtgg acgacagctt cttccaccgc 300

ctggaggaga gcttcctggt ggaggaggac aagaagcacg agcgccaccc catcttcggc 360

aacatcgtgg acgaggtggc ctaccacgag aagtacccca ccatctacca cctgcgcaag 420

aagctggtgg acagcaccga caaggccgac ctgcgcctga tctacctggc cctggcccac 480

atgatcaagt tccgcggcca cttcctgatc gagggcgacc tgaaccccga caacagcgac 540

gtggacaagc tgttcatcca gctggtgcag acctacaacc agctgttcga ggagaacccc 600

atcaacgcca gcggcgtgga cgccaaggcc atcctgagcg cccgcctgag caagagccgc 660

cgcctggaga acctgatcgc ccagctgccc ggcgagaaga agaacggcct gttcggcaac 720

ctgatcgccc tgagcctggg cctgaccccc aacttcaaga gcaacttcga cctggccgag 780

gacgccaagc tgcagctgag caaggacacc tacgacgacg acctggacaa cctgctggcc 840

cagatcggcg accagtacgc cgacctgttc ctggccgcca agaacctgag cgacgccatc 900

ctgctgagcg acatcctgcg cgtgaacacc gagatcacca aggcccccct gagcgccagc 960

atgatcaagc gctacgacga gcaccaccag gacctgaccc tgctgaaggc cctggtgcgc 1020

cagcagctgc ccgagaagta caaggagatc ttcttcgacc agagcaagaa cggctacgcc 1080

ggctacatcg acggcggcgc cagccaggag gagttctaca agttcatcaa gcccatcctg 1140

gagaagatgg acggcaccga ggagctgctg gtgaagctga accgcgagga cctgctgcgc 1200

aagcagcgca ccttcgacaa cggcagcatc ccccaccaga tccacctggg cgagctgcac 1260

gccatcctgc gccgccagga ggacttctac cccttcctga aggacaaccg cgagaagatc 1320

gagaagatcc tgaccttccg catcccctac tacgtgggcc ccctggcccg cggcaacagc 1380

cgcttcgcct ggatgacccg caagagcgag gagaccatca ccccctggaa cttcgaggag 1440

gtggtggaca agggcgccag cgcccagagc ttcatcgagc gcatgaccaa cttcgacaag 1500

aacctgccca acgagaaggt gctgcccaag cacagcctgc tgtacgagta cttcaccgtg 1560

tacaacgagc tgaccaaggt gaagtacgtg accgagggca tgcgcaagcc cgccttcctg 1620

agcggcgagc agaagaaggc catcgtggac ctgctgttca agaccaaccg caaggtgacc 1680

gtgaagcagc tgaaggagga ctacttcaag aagatcgagt gcttcgacag cgtggagatc 1740

agcggcgtgg aggaccgctt caacgccagc ctgggcacct accacgacct gctgaagatc 1800

atcaaggaca aggacttcct ggacaacgag gagaacgagg acatcctgga ggacatcgtg 1860

ctgaccctga ccctgttcga ggaccgcgag atgatcgagg agcgcctgaa gacctacgcc 1920

cacctgttcg acgacaaggt gatgaagcag ctgaagcgcc gccgctacac cggctggggc 1980

cgcctgagcc gcaagcttat caacggcatc cgcgacaagc agagcggcaa gaccatcctg 2040

gacttcctga agagcgacgg cttcgccaac cgcaacttca tgcagctgat ccacgacgac 2100

agcctgacct tcaaggagga catccagaag gcccaggtga gcggccaggg cgacagcctg 2160

cacgagcaca tcgccaacct ggccggcagc cccgccatca agaagggcat cctgcagacc 2220

gtgaaggtgg tggacgagct ggtgaaggtg atgggccgcc acaagcccga gaacatcgtg 2280

atcgagatgg cccgcgagaa ccagaccacc cagaagggcc agaagaacag ccgcgagcgc 2340

atgaagcgca tcgaggaggg catcaaggag ctgggcagcc agatcctgaa ggagcacccc 2400

gtggagaaca cccagctgca gaacgagaag ctgtacctgt actacctgca gaacggccgc 2460

gacatgtacg tggaccagga gctggacatc aaccgcctga gcgactacga cgtggaccac 2520

atcgtgcccc agagcttcct gaaggacgac agcatcgaca acaaggtgct gacccgcagc 2580

gacaagaacc gcggcaagag cgacaacgtg cccagcgagg aggtggtgaa gaagatgaag 2640

aactactggc gccagctgct gaacgccaag ctgatcaccc agcgcaagtt cgacaacctg 2700

accaaggccg agcgcggcgg cctgagcgag ctggacaagg ccggcttcat caagcgccag 2760

ctggtggaga cccgccagat caccaagcac gtggcccaga tcctggacag ccgcatgaac 2820

accaagtacg acgagaacga caagctgatc cgcgaggtga aggtgatcac cctgaagagc 2880

aagctggtga gcgacttccg caaggacttc cagttctaca aggtgcgcga gatcaacaac 2940

taccaccacg cccacgacgc ctacctgaac gccgtggtgg gcaccgccct gatcaagaag 3000

taccccaagc tggagagcga gttcgtgtac ggcgactaca aggtgtacga cgtgcgcaag 3060

atgatcgcca agagcgagca ggagatcggc aaggccaccg ccaagtactt cttctacagc 3120

aacatcatga acttcttcaa gaccgagatc accctggcca acggcgagat ccgcaagcgc 3180

cccctgatcg agaccaacgg cgagaccggc gagatcgtgt gggacaaggg ccgcgacttc 3240

gccaccgtgc gcaaggtgct gagcatgccc caggtgaaca tcgtgaagaa gaccgaggtg 3300

cagaccggcg gcttcagcaa ggagagcatc ctgcccaagc gcaacagcga caagctgatc 3360

gcccgcaaga aggactggga ccccaagaag tacggcggct tcgacagccc caccgtggcc 3420

tacagcgtgc tggtggtggc caaggtggag aagggcaaga gcaagaagct gaagagcgtg 3480

aaggagctgc tgggcatcac catcatggag cgcagcagct tcgagaagaa ccccatcgac 3540

ttcctggagg ccaagggcta caaggaggtg aagaaggacc tgatcatcaa gctgcccaag 3600

tacagcctgt tcgagctgga gaacggccgc aagcgcatgc tggccagcgc cggcgagctg 3660

cagaagggca acgagctggc cctgcccagc aagtacgtga acttcctgta cctggccagc 3720

cactacgaga agctgaaggg cagccccgag gacaacgagc agaagcagct gttcgtggag 3780

cagcacaagc actacctgga cgagatcatc gagcagatca gcgagttcag caagcgcgtg 3840

atcctggccg acgccaacct ggacaaggtg ctgagcgcct acaacaagca ccgcgacaag 3900

cccatccgcg agcaggccga gaacatcatc cacctgttca ccctgaccaa cctgggcgcc 3960

cccgccgcct tcaagtactt cgacaccacc atcgaccgca agcgctacac cagcaccaag 4020

gaggtgctgg acgccaccct gatccaccag agcatcaccg gtctgtacga gacccgcatc 4080

gacctgagcc agctgggcgg cgactaa 4107

<210> 2
<211> 21
<212> PRT
<213> Artificial Sequence

<220>
<223> peptide tag

<400> 2
Gly Gly Ser Gly Pro Pro Lys Lys Lys Arg Lys Val Tyr Pro Tyr Asp
1 5 10 15

Val Pro Asp Tyr Ala
20

<210> 3
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> F primer for CCR5

<400> 3
aattcatgac atcaattatt atacatcgga ggag 34

<210> 4
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> R primer for CCR5

<400> 4
gatcctcctc cgatgtataa taattgatgt catg 34

<210> 5
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> F1 primer for CCR5

<400> 5
ctccatggtg ctatagagca 20

<210> 6
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> F2 primer for CCR5

<400> 6
gagccaagct ctccatctag t 21

<210> 7
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> R primer for CCR5

<400> 7
gccctgtcaa gagttgacac 20

<210> 8
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> F1 primer for C4BPB

<400> 8
tatttggctg gttgaaaggg 20

<210> 9
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> R1 primer for C4BPB

<400> 9
aaagtcatga aataaacaca ccca 24

<210> 10
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> F2 primer for C4BPB

<400> 10
ctgcattgat atggtagtac catg 24

<210> 11
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> R2 primer for C4BPB

<400> 11
gctgttcatt gcaatggaat g 21

<210> 12
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> F1 primer for ADCY5

<400> 12
gctcccacct tagtgctctg 20

<210> 13
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> R1 primer for ADCY5

<400> 13
ggtggcagga acctgtatgt 20

<210> 14
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> F2 primer for ADCY5

<400> 14
gtcattggcc agagatgtgg a 21

<210> 15
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> R2 primer for ADCY5

<400> 15
gtcccatgac aggcgtgtat 20

<210> 16
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> F primer for KCNJ6

<400> 16
gcctggccaa gtttcagtta 20

<210> 17
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> R1 primer for KCNJ6

<400> 17
tggagccatt ggtttgcatc 20

<210> 18
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> R2 primer for KCNJ6

<400> 18
ccagaactaa gccgtttctg ac 22

<210> 19
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> F1 primer for CNTNAP2

<400> 19
atcaccgaca accagtttcc 20

<210> 20
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> F2 primer for CNTNAP2

<400> 20
tgcagtgcag actctttcca 20

<210> 21
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> R primer for CNTNAP2

<400> 21
aaggacacag ggcaactgaa 20

<210> 22
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> F1 primer for N/A Chr. 5

<400> 22
tgtggaacga gtggtgacag 20

<210> 23
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> R1 primer for N/A Chr. 5

<400> 23
gctggattag gaggcaggat tc 22

<210> 24
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> F2 primer for N/A Chr. 5

<400> 24
gtgctgagaa cgcttcatag ag 22

<210> 25
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> R2 primer for N/A Chr. 5

<400> 25
ggaccaaacc acattcttct cac 23

<210> 26
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> F primer for deletion

<400> 26
ccacatctcg ttctcggttt 20

<210> 27
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> R primer for deletion

<400> 27
tcacaagccc acagatattt 20

<210> 28
<211> 105
<212> RNA
<213> Artificial Sequence

<220>
<223> sgRNA for CCR5

<400> 28
ggugacauca auuauuauac auguuuuaga gcuagaaaua gcaaguuaaa auaaggcuag 60

uccguuauca acuugaaaaa guggcaccga gucggugcuu uuuuu 105

<210> 29
<211> 44
<212> RNA
<213> Artificial Sequence

<220>
<223> crRNA for CCR5

<400> 29
ggugacauca auuauuauac auguuuuaga gcuaugcugu uuug 44

<210> 30
<211> 86
<212> RNA
<213> Artificial Sequence

<220>
<223> tracrRNA for CCR5

<400> 30
ggaaccauuc aaaacagcau agcaaguuaa aauaaggcua guccguuauc aacuugaaaa 60

aguggcaccg agucggugcu uuuuuu 86

<210> 31
<211> 86
<212> DNA
<213> Artificial Sequence

<220>
<223> Foxn1 #1 sgRNA

<400> 31
gaaattaata cgactcacta taggcagtct gacgtcacac ttccgtttta gagctagaaa 60

tagcaagtta aaataaggct agtccg 86

<210> 32
<211> 86

<212> DNA
<213> Artificial Sequence

<220>
<223> Foxn1 #2 sgRNA

<400> 32
gaaattaata cgactcacta taggacttcc aggctccacc cgacgtttta gagctagaaa 60

tagcaagtta aaataaggct agtccg 86

<210> 33
<211> 86
<212> DNA
<213> Artificial Sequence

<220>
<223> Foxn1 #3 sgRNA

<400> 33
gaaattaata cgactcacta taggccaggc tccacccgac tggagtttta gagctagaaa 60

tagcaagtta aaataaggct agtccg 86

<210> 34
<211> 86
<212> DNA
<213> Artificial Sequence

<220>
<223> Foxn1 #4 sgRNA

<400> 34
gaaattaata cgactcacta taggactgga gggcgaaccc caaggtttta gagctagaaa 60

tagcaagtta aaataaggct agtccg 86

<210> 35
<211> 86
<212> DNA
<213> Artificial Sequence

<220>
<223> Foxn1 #5 sgRNA

<400> 35
gaaattaata cgactcacta taggacccca aggggacctc atgcgtttta gagctagaaa 60

tagcaagtta aaataaggct agtccg 86

<210> 36
<211> 86
<212> DNA
<213> Artificial Sequence

<220>
<223> Prkdc #1 sgRNA

<400> 36
gaaattaata cgactcacta taggttagtt ttttccagag acttgtttta gagctagaaa 60

tagcaagtta aaataaggct agtccg 86

<210> 37
<211> 86
<212> DNA
<213> Artificial Sequence

<220>
<223> Prkdc #2 sgRNA

<400> 37
gaaattaata cgactcacta taggttggtt tgcttgtgtt tatcgtttta gagctagaaa 60

tagcaagtta aaataaggct agtccg 86

<210> 38
<211> 86
<212> DNA
<213> Artificial Sequence

<220>
<223> Prkdc #3 sgRNA

<400> 38
gaaattaata cgactcacta taggcacaag caaaccaaag tctcgtttta gagctagaaa 60

tagcaagtta aaataaggct agtccg 86

<210> 39

<211> 86
<212> DNA
<213> Artificial Sequence

<220>
<223> Prkdc #4 sgRNA

<400> 39
gaaattaata cgactcacta taggcctcaa tgctaagcga cttcgtttta gagctagaaa 60

tagcaagtta aaataaggct agtccg 86

<210> 40
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> F1 primer for Foxn1

<400> 40
gtctgtctat catctcttcc cttctctcc 29

<210> 41
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> F2 primer for Foxn1

<400> 41
tccctaatcc gatggctagc tccag 25

<210> 42
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> R1 primer for Foxn1

<400> 42
acgagcagct gaagttagca tgc 23

<210> 43
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> R2 primer for Foxn1

<400> 43
ctactcaatg ctcttagagc taccaggctt gc 32

<210> 44
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> F primer for Prkdc

<400> 44
gactgttgtg gggagggccg 20

<210> 45
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> F2 primer for Prkdc

<400> 45
gggagggccg aaagtcttat tttg 24

<210> 46
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> R1 primer for Prkdc

<400> 46
cctgaagact gaagttggca gaagtgag 28

<210> 47
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> R2 primer for Prkdc

<400> 47
ctttagggct tcttctctac aatcacg 27

<210> 48
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> F primer for Foxn1

<400> 48
ctcggtgtgt agccctgacc tcggtgtgta gccctgac 38

<210> 49
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> R primer for Foxn1

<400> 49
agactggcct ggaactcaca g 21

<210> 50
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> F primer for Foxn1

<400> 50
cactaaagcc tgtcaggaag ccg 23

<210> 51
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> R primer for Foxn1

<400> 51
ctgtggagag cacacagcag c 21

<210> 52
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> F primer for Foxn1

<400> 52
gctgcgacct gagaccatg 19

<210> 53
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> R primer for Foxn1

<400> 53
cttcaatggc ttcctgctta ggctac 26

<210> 54
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> F primer for Foxn1

<400> 54
ggttcagatg aggccatcct ttc 23

<210> 55
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> R primer for Foxn1

<400> 55
cctgatctgc aggcttaacc cttg 24

<210> 56
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> F primer for Prkdc

<400> 56
ctcacctgca catcacatgt gg 22

<210> 57
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> R primer for Prkdc

<400> 57
ggcatccacc ctatggggtc 20

<210> 58
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> F primer for Prkdc

<400> 58
gccttgacct agagcttaaa gagcc 25

<210> 59
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> R primer for Prkdc

<400> 59
ggtcttgtta gcaggaagga cactg 25

<210> 60
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> F primer for Prkdc

<400> 60
aaaactctgc ttgatgggat atgtggg 27

<210> 61
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> R primer for Prkdc

<400> 61
ctctcactgg ttatctgtgc tccttc 26

<210> 62
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> F primer for Prkdc

<400> 62
ggatcaatag gtggtggggg atg 23

<210> 63
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> R primer for Prkdc

<400> 63
gtgaatgaca caatgtgaca gcttcag 27

<210> 64
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> F primer for Prkdc

<400> 64
cacaagacag acctctcaac attcagtc 28

<210> 65
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> R primer for Prkdc

<400> 65
gtgcatgcat ataatccatt ctgattgctc tc 32

<210> 66
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> F1 primer for Prkdc

<400> 66
gggaggcaga ggcaggt 17

<210> 67
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> F2 primer for Prkdc

<400> 67
ggatctctgt gagtttgagg cca 23

<210> 68
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> R1 primer for Prkdc

<400> 68
gctccagaac tcactcttag gctc 24

<210> 69
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for Foxn1

<400> 69
ctactccctc cgcagtctga 20

<210> 70
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for Foxn1

<400> 70
ccaggcctag gttccaggta 20

<210> 71
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for Prkdc

<400> 71
ccccagcatt gcagatttcc 20

<210> 72
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer for Prkdc

<400> 72
agggcttctt ctctacaatc acg 23

<210> 73
<211> 86
<212> DNA
<213> Artificial Sequence

<220>
<223> BRI1 target 1

<400> 73
gaaattaata cgactcacta taggtttgaa agatggaagc gcgggtttta gagctagaaa 60

tagcaagtta aaataaggct agtccg 86

<210> 74
<211> 86
<212> DNA
<213> Artificial Sequence

<220>
<223> BRI1 target 2

<400> 74

gaaattaata cgactcacta taggtgaaac taaactggtc cacagtttta gagctagaaa 60

tagcaagtta aaataaggct agtccg 86

<210> 75
<211> 64
<212> DNA
<213> Artificial Sequence

<220>
<223> Universal

<400> 75
aaaaaagcac cgactcggtg ccacttttc aagttgataa cggactagcc ttattttaac 60

ttgc 64

<210> 76
<211> 65
<212> DNA
<213> Artificial Sequence

<220>
<223> Templates for crRNA

<400> 76
gaaattaata cgactcacta taggnnnnnn nnnnnnnnnn nnnngtttta gagctatgct 60

gtttt 65

<210> 77
<211> 67
<212> DNA
<213> Artificial Sequence

<220>
<223> tracrRNA

<400> 77
gaaattaata cgactcacta taggaaccat tcaaaacagc atagcaagtt aaaataaggc 60

tagtccg 67

<210> 78
<211> 69

<212> DNA
<213> Artificial Sequence

<220>
<223> tracrRNA

<400> 78
aaaaaaagca ccgactcggt gccacttttt caagttgata acggactagc cttattttaa 60

cttgctatg 69

<210> 79
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 79
ctccatggtg ctatagagca 20

<210> 80
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 80
gagccaagct ctccatctag t 21

<210> 81
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 81
gccctgtcaa gagttgacac 20

<210> 82
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 82
gcacagggtg gaacaagatg ga 22

<210> 83
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 83
gccaggtacc tatcgattgt cagg 24

<210> 84
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 84
gagccaagct ctccatctag t 21

<210> 85
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 85
actctgactg ggtcaccagc 20

<210> 86
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 86
tatttggctg gttgaaaggg 20

<210> 87
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 87
aaagtcatga aataaacaca ccca 24

<210> 88
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 88
ctgcattgat atggtagtac catg 24

<210> 89
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 89
gctgttcatt gcaatggaat g 21

<210> 90
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 90
atggagttgg acatggccat gg 22

<210> 91
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 91
actcactatc cacagttcag catttacc 28

<210> 92
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 92
tggagatagc tgtcagcaac ttt 23

<210> 93
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 93
caacaaagca aaggtaaagt tggtaatag 29

<210> 94
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 94
ggtttcagga gatgtgttac aaggc 25

<210> 95
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 95
gattgtgcaa ttcctatgca atcggtc 27

<210> 96
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 96
cactgggtac ttaatctgta gcctc 25

<210> 97
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 97
ggttccaagt cattcccagt agc 23

<210> 98
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 98
catcactgca gttgtaggtt ataactatcc 30

<210> 99
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 99
ttgaaaacca cagatctggt tgaacc 26

<210> 100
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 100
ggagtgccaa gagaatatct gg 22

<210> 101
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 101
ctgaaactgg tttcaaaata ttcgttttaa gg 32

<210> 102
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 102
gctctgtatg ccctgtagta gg 22

<210> 103
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer

<400> 103
tttgcatctg accttacctt tg 22

<210> 104
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Target sequence of RGEN

<400> 104
aatgaccact acatcctcaa ggg 23

<210> 105
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Target sequence of RGEN

<400> 105
agatgatgtc tcatcatcag agg 23

<210> 106
<211> 4170
<212> DNA
<213> Artificial Sequence

<220>
<223> Cas9-coding sequence in p3s-Cas9HC (humanized, C-term tagging, human cell experiments)

<400> 106
atggacaaga agtacagcat cggcctggac atcggtacca acagcgtggg ctgggccgtg 60

atcaccgacg agtacaaggt gcccagcaag aagttcaagg tgctgggcaa caccgaccgc 120

cacagcatca agaagaacct gatcggcgcc ctgctgttcg acagcggcga gaccgccgag 180

gccacccgcc tgaagcgcac cgcccgccgc cgctacaccc gccgcaagaa ccgcatctgc 240

tacctgcagg agatcttcag caacgagatg gccaaggtgg acgacagctt cttccaccgc 300

ctggaggaga gcttcctggt ggaggaggac aagaagcacg agcgccaccc catcttcggc 360

aacatcgtgg acgaggtggc ctaccacgag aagtacccca ccatctacca cctgcgcaag 420

aagctggtgg acagcaccga caaggccgac ctgcgcctga tctacctggc cctggcccac 480

atgatcaagt tccgcggcca cttcctgatc gagggcgacc tgaaccccga caacagcgac 540

gtggacaagc tgttcatcca gctggtgcag acctacaacc agctgttcga ggagaacccc 600

atcaacgcca gcggcgtgga cgccaaggcc atcctgagcg cccgcctgag caagagccgc 660

cgcctggaga acctgatcgc ccagctgccc ggcgagaaga agaacggcct gttcggcaac 720

ctgatcgccc tgagcctggg cctgaccccc aacttcaaga gcaacttcga cctggccgag 780

gacgccaagc tgcagctgag caaggacacc tacgacgacg acctggacaa cctgctggcc 840

cagatcggcg accagtacgc cgacctgttc ctggccgcca agaacctgag cgacgccatc 900

ctgctgagcg acatcctgcg cgtgaacacc gagatcacca aggcccccct gagcgccagc 960

atgatcaagc gctacgacga gcaccaccag gacctgaccc tgctgaaggc cctggtgcgc 1020

cagcagctgc ccgagaagta caaggagatc ttcttcgacc agagcaagaa cggctacgcc 1080

ggctacatcg acggcggcgc cagccaggag gagttctaca agttcatcaa gcccatcctg 1140

gagaagatgg acggcaccga ggagctgctg gtgaagctga accgcgagga cctgctgcgc 1200

aagcagcgca ccttcgacaa cggcagcatc ccccaccaga tccacctggg cgagctgcac 1260

gccatcctgc gccgccagga ggacttctac cccttcctga aggacaaccg cgagaagatc 1320

gagaagatcc tgaccttccg catcccctac tacgtgggcc ccctggcccg cggcaacagc 1380

cgcttcgcct ggatgacccg caagagcgag gagaccatca ccccctggaa cttcgaggag 1440

gtggtggaca agggcgccag cgcccagagc ttcatcgagc gcatgaccaa cttcgacaag 1500

aacctgccca acgagaaggt gctgcccaag cacagcctgc tgtacgagta cttcaccgtg 1560

tacaacgagc tgaccaaggt gaagtacgtg accgagggca tgcgcaagcc cgccttcctg 1620

agcggcgagc agaagaaggc catcgtggac ctgctgttca agaccaaccg caaggtgacc 1680

gtgaagcagc tgaaggagga ctacttcaag aagatcgagt gcttcgacag cgtggagatc 1740

agcggcgtgg aggaccgctt caacgccagc ctgggcacct accacgacct gctgaagatc 1800

atcaaggaca aggacttcct ggacaacgag gagaacgagg acatcctgga ggacatcgtg 1860

ctgaccctga ccctgttcga ggaccgcgag atgatcgagg agcgcctgaa gacctacgcc 1920

cacctgttcg acgacaaggt gatgaagcag ctgaagcgcc gccgctacac cggctggggc 1980

cgcctgagcc gcaagcttat caacggcatc cgcgacaagc agagcggcaa gaccatcctg 2040

gacttcctga agagcgacgg cttcgccaac cgcaacttca tgcagctgat ccacgacgac 2100

agcctgacct tcaaggagga catccagaag gcccaggtga gcggccaggg cgacagcctg 2160

cacgagcaca tcgccaacct ggccggcagc cccgccatca agaagggcat cctgcagacc 2220

gtgaaggtgg tggacgagct ggtgaaggtg atgggccgcc acaagcccga gaacatcgtg 2280

atcgagatgg cccgcgagaa ccagaccacc cagaagggcc agaagaacag ccgcgagcgc 2340

atgaagcgca tcgaggaggg catcaaggag ctgggcagcc agatcctgaa ggagcacccc 2400

gtggagaaca cccagctgca gaacgagaag ctgtacctgt actacctgca gaacggccgc 2460

gacatgtacg tggaccagga gctggacatc aaccgcctga gcgactacga cgtggaccac 2520

atcgtgcccc agagcttcct gaaggacgac agcatcgaca acaaggtgct gacccgcagc 2580

gacaagaacc gcggcaagag cgacaacgtg cccagcgagg aggtggtgaa gaagatgaag 2640

aactactggc gccagctgct gaacgccaag ctgatcaccc agcgcaagtt cgacaacctg 2700

accaaggccg agcgcggcgg cctgagcgag ctggacaagg ccggcttcat caagcgccag 2760

ctggtggaga cccgccagat caccaagcac gtggcccaga tcctggacag ccgcatgaac 2820

accaagtacg acgagaacga caagctgatc cgcgaggtga aggtgatcac cctgaagagc 2880

aagctggtga gcgacttccg caaggacttc cagttctaca aggtgcgcga gatcaacaac 2940

taccaccacg cccacgacgc ctacctgaac gccgtggtgg gcaccgccct gatcaagaag 3000

taccccaagc tggagagcga gttcgtgtac ggcgactaca aggtgtacga cgtgcgcaag 3060

atgatcgcca agagcgagca ggagatcggc aaggccaccg ccaagtactt cttctacagc 3120

aacatcatga acttcttcaa gaccgagatc accctggcca acggcgagat ccgcaagcgc 3180

cccctgatcg agaccaacgg cgagaccggc gagatcgtgt gggacaaggg ccgcgacttc 3240

gccaccgtgc gcaaggtgct gagcatgccc caggtgaaca tcgtgaagaa gaccgaggtg 3300

cagaccggcg gcttcagcaa ggagagcatc ctgcccaagc gcaacagcga caagctgatc 3360

gcccgcaaga aggactggga ccccaagaag tacggcggct tcgacagccc caccgtggcc 3420

tacagcgtgc tggtggtggc caaggtggag aagggcaaga gcaagaagct gaagagcgtg 3480

aaggagctgc tgggcatcac catcatggag cgcagcagct tcgagaagaa ccccatcgac 3540

ttcctggagg ccaagggcta caaggaggtg aagaaggacc tgatcatcaa gctgcccaag 3600

tacagcctgt tcgagctgga gaacggccgc aagcgcatgc tggccagcgc cggcgagctg 3660

cagaagggca acgagctggc cctgcccagc aagtacgtga acttcctgta cctggccagc 3720

cactacgaga agctgaaggg cagccccgag gacaacgagc agaagcagct gttcgtggag 3780

cagcacaagc actacctgga cgagatcatc gagcagatca gcgagttcag caagcgcgtg 3840

atcctggccg acgccaacct ggacaaggtg ctgagcgcct acaacaagca ccgcgacaag 3900

cccatccgcg agcaggccga gaacatcatc cacctgttca ccctgaccaa cctgggcgcc 3960

cccgccgcct tcaagtactt cgacaccacc atcgaccgca agcgctacac cagcaccaag 4020

gaggtgctgg acgccaccct gatccaccag agcatcaccg gtctgtacga gacccgcatc 4080

gacctgagcc agctgggcgg cgacggcggc tccggacctc caaagaaaaa gagaaaagta 4140

taccctacg acgtgcccga ctacgcctaa 4170

<210> 107
<211> 4194
<212> DNA
<213> Artificial Sequence

<220>
<223> Cas9 coding sequence in p3s-Cas9HN (humanized codon, N-term tagging (underlined), human cell experiments)

<400> 107
atggtgtacc cctacgacgt gcccgactac gccgaattgc ctccaaaaaa gaagagaaag 60

gtagggatcc gaattcccgg ggaaaaaccg gacaagaagt acagcatcgg cctggacatc 120

ggtaccaaca gcgtgggctg ggccgtgatc accgacgagt acaaggtgcc cagcaagaag 180

ttcaaggtgc tgggcaacac cgaccgccac agcatcaaga agaacctgat cggcgccctg 240

ctgttcgaca gcggcgagac cgccgaggcc acccgcctga agcgcaccgc ccgccgccgc 300

tacacccgcc gcaagaaccg catctgctac ctgcaggaga tcttcagcaa cgagatggcc 360

aaggtggacg acagcttctt ccaccgcctg gaggagagct tcctggtgga ggaggacaag 420

aagcacgagc gccaccccat cttcggcaac atcgtggacg aggtggccta ccacgagaag 480

taccccacca tctaccacct gcgcaagaag ctggtggaca gcaccgacaa ggccgacctg 540

cgcctgatct acctggccct ggcccacatg atcaagttcc gcggccactt cctgatcgag 600

ggcgacctga accccgacaa cagcgacgtg gacaagctgt tcatccagct ggtgcagacc 660

tacaaccagc tgttcgagga gaaccccatc aacgccagcg gcgtggacgc caaggccatc 720

ctgagcgccc gcctgagcaa gagccgccgc ctggagaacc tgatcgccca gctgcccggc 780

gagaagaaga acggcctgtt cggcaacctg atcgccctga gcctgggcct gaccccaac 840

ttcaagagca acttcgacct ggccgaggac gccaagctgc agctgagcaa ggacacctac 900

gacgacgacc tggacaacct gctggcccag atcggcgacc agtacgccga cctgttcctg 960

gccgccaaga acctgagcga cgccatcctg ctgagcgaca tcctgcgcgt gaacaccgag 1020

atcaccaagg ccccctgag cgccagcatg atcaagcgct acgacgagca ccaccaggac 1080

ctgaccctgc tgaaggccct ggtgcgccag cagctgcccg agaagtacaa ggagatcttc 1140

ttcgaccaga gcaagaacgg ctacgccggc tacatcgacg gcggcgccag ccaggaggag 1200

ttctacaagt tcatcaagcc catcctggag aagatggacg gcaccgagga gctgctggtg 1260

aagctgaacc gcgaggacct gctgcgcaag cagcgcacct tcgacaacgg cagcatcccc 1320

caccagatcc acctgggcga gctgcacgcc atcctgcgcc gccaggagga cttctacccc 1380

ttcctgaagg acaaccgcga gaagatcgag aagatcctga ccttccgcat cccctactac 1440

gtgggccccc tggcccgcgg caacagccgc ttcgcctgga tgacccgcaa gagcgaggag 1500

accatcaccc cctggaactt cgaggaggtg gtggacaagg gcgccagcgc ccagagcttc 1560

atcgagcgca tgaccaactt cgacaagaac ctgcccaacg agaaggtgct gcccaagcac 1620

agcctgctgt acgagtactt caccgtgtac aacgagctga ccaaggtgaa gtacgtgacc 1680

gagggcatgc gcaagcccgc cttcctgagc ggcgagcaga agaaggccat cgtggacctg 1740

ctgttcaaga ccaaccgcaa ggtgaccgtg aagcagctga aggaggacta cttcaagaag 1800

atcgagtgct tcgacagcgt ggagatcagc ggcgtggagg accgcttcaa cgccagcctg 1860

ggcacctacc acgacctgct gaagatcatc aaggacaagg acttcctgga caacgaggag 1920

aacgaggaca tcctggagga catcgtgctg accctgaccc tgttcgagga ccgcgagatg 1980

atcgaggagc gcctgaagac ctacgcccac ctgttcgacg acaaggtgat gaagcagctg 2040

aagcgccgcc gctacaccgg ctggggccgc ctgagccgca agcttatcaa cggcatccgc 2100

gacaagcaga gcggcaagac catcctggac ttcctgaaga gcgacggctt cgccaaccgc 2160

aacttcatgc agctgatcca cgacgacagc ctgaccttca aggaggacat ccagaaggcc 2220

caggtgagcg gccagggcga cagcctgcac gagcacatcg ccaacctggc cggcagcccc 2280

gccatcaaga agggcatcct gcagaccgtg aaggtggtgg acgagctggt gaaggtgatg 2340

ggccgccaca agcccgagaa catcgtgatc gagatggccc gcgagaacca gaccacccag 2400

aagggccaga agaacagccg cgagcgcatg aagcgcatcg aggagggcat caaggagctg 2460

ggcagccaga tcctgaagga gcaccccgtg gagaacaccc agctgcagaa cgagaagctg 2520

tacctgtact acctgcagaa cggccgcgac atgtacgtgg accaggagct ggacatcaac 2580

cgcctgagcg actacgacgt ggaccacatc gtgccccaga gcttcctgaa ggacgacagc 2640

atcgacaaca aggtgctgac ccgcagcgac aagaaccgcg gcaagagcga caacgtgccc 2700

agcgaggagg tggtgaagaa gatgaagaac tactggcgcc agctgctgaa cgccaagctg 2760

atcacccagc gcaagttcga caacctgacc aaggccgagc gcggcggcct gagcgagctg 2820

gacaaggccg gcttcatcaa gcgccagctg gtggagaccc gccagatcac caagcacgtg 2880

gcccagatcc tggacagccg catgaacacc aagtacgacg agaacgacaa gctgatccgc 2940

gaggtgaagg tgatcaccct gaagagcaag ctggtgagcg acttccgcaa ggacttccag 3000

ttctacaagg tgcgcgagat caacaactac caccacgccc acgacgccta cctgaacgcc 3060

gtggtgggca ccgccctgat caagaagtac cccaagctgg agagcgagtt cgtgtacggc 3120

gactacaagg tgtacgacgt gcgcaagatg atcgccaaga gcgagcagga gatcggcaag 3180

gccaccgcca agtacttctt ctacagcaac atcatgaact tcttcaagac cgagatcacc 3240

ctggccaacg gcgagatccg caagcgcccc ctgatcgaga ccaacggcga gaccggcgag 3300

atcgtgtggg acaagggccg cgacttcgcc accgtgcgca aggtgctgag catgccccag 3360

gtgaacatcg tgaagaagac cgaggtgcag accggcggct tcagcaagga gagcatcctg 3420

cccaagcgca acagcgacaa gctgatcgcc cgcaagaagg actgggaccc caagaagtac 3480

ggcggcttcg acagccccac cgtggcctac agcgtgctgg tggtggccaa ggtggagaag 3540

ggcaagagca agaagctgaa gagcgtgaag gagctgctgg gcatcaccat catggagcgc 3600

agcagcttcg agaagaaccc catcgacttc ctggaggcca agggctacaa ggaggtgaag 3660

aaggacctga tcatcaagct gcccaagtac agcctgttcg agctggagaa cggccgcaag 3720

cgcatgctgg ccagcgccgg cgagctgcag aagggcaacg agctggccct gcccagcaag 3780

tacgtgaact tcctgtacct ggccagccac tacgagaagc tgaagggcag ccccgaggac 3840

aacgagcaga agcagctgtt cgtggagcag cacaagcact acctggacga gatcatcgag 3900

cagatcagcg agttcagcaa gcgcgtgatc ctggccgacg ccaacctgga caaggtgctg 3960

agcgcctaca acaagcaccg cgacaagccc atccgcgagc aggccgagaa catcatccac 4020

ctgttcaccc tgaccaacct gggcgccccc gccgccttca agtacttcga caccaccatc 4080

gaccgcaagc gctacaccag caccaaggag gtgctggacg ccaccctgat ccaccagagc 4140

atcaccggtc tgtacgagac ccgcatcgac ctgagccagc tgggcggcga ctaa 4194

<210> 108
<211> 4107
<212> DNA
<213> Artificial Sequence

<220>
<223> Cas9-coding sequence in Streptococcus pyogenes

<400> 108

atggataaga aatactcaat aggcttagat atcggcacaa atagcgtcgg atgggcggtg 60

atcactgatg aatataaggt tccgtctaaa aagttcaagg ttctgggaaa tacagaccgc 120

cacagtatca aaaaaaatct tataggggct cttttatttg acagtggaga gacagcggaa 180

gcgactcgtc tcaaacggac agctcgtaga aggtatacac gtcggaagaa tcgtatttgt 240

tatctacagg agatttttc aaatgagatg gcgaaagtag atgatagttt ctttcatcga 300

cttgaagagt ctttttggt ggaagaagac aagaagcatg aacgtcatcc tatttttgga 360

aatatagtag atgaagttgc ttatcatgag aaatatccaa ctatctatca tctgcgaaaa 420

aaattggtag attctactga taaagcggat ttgcgcttaa tctatttggc cttagcgcat 480

atgattaagt ttcgtggtca ttttttgatt gagggagatt taaatcctga taatagtgat 540

gtggacaaac tatttatcca gttggtacaa acctacaatc aattatttga agaaaaccct 600

attaacgcaa gtggagtaga tgctaaagcg attctttctg cacgattgag taaatcaaga 660

cgattagaaa atctcattgc tcagctcccc ggtgagaaga aaaatggctt atttgggaat 720

ctcattgctt tgtcattggg tttgacccct aattttaaat caaattttga tttggcagaa 780

gatgctaaat tacagctttc aaaagatact tacgatgatg atttagataa tttattggcg 840

caaattggag atcaatatgc tgatttgttt ttggcagcta agaatttatc agatgctatt 900

ttactttcag atatcctaag agtaaatact gaaataacta aggctcccct atcagcttca 960

atgattaaac gctacgatga acatcatcaa gacttgactc ttttaaaagc tttagttcga 1020

caacaacttc cagaaaagta taaagaaatc ttttttgatc aatcaaaaaa cggatatgca 1080

ggttatattg atgggggagc tagccaagaa gaattttata aatttatcaa accaatttta 1140

gaaaaaatgg atggtactga ggaattattg gtgaaactaa atcgtgaaga tttgctgcgc 1200

aagcaacgga cctttgacaa cggctctatt ccccatcaaa ttcacttggg tgagctgcat 1260

gctattttga gaagacaaga agacttttat ccatttttaa aagacaatcg tgagaagatt 1320

gaaaaaatct tgacttttcg aattccttat tatgttggtc cattggcgcg tggcaatagt 1380

cgttttgcat ggatgactcg gaagtctgaa gaaacaatta ccccatggaa ttttgaagaa 1440

gttgtcgata aaggtgcttc agctcaatca tttattgaac gcatgacaaa ctttgataaa 1500

aatcttccaa atgaaaaagt actaccaaaa catagtttgc tttatgagta ttttacggtt 1560

tataacgaat tgacaaaggt caaatatgtt actgaaggaa tgcgaaaacc agcatttctt 1620

tcaggtgaac agaagaaagc cattgttgat ttactcttca aaacaaatcg aaaagtaacc 1680

gttaagcaat taaaagaaga ttatttcaaa aaaatagaat gttttgatag tgttgaaatt 1740

tcaggagttg aagatagatt taatgcttca ttaggtacct accatgattt gctaaaaatt 1800

attaaagata aagatttttt ggataatgaa gaaaatgaag atatcttaga ggatattgtt 1860

ttaacattga ccttatttga agatagggag atgattgagg aaagacttaa aacatatgct 1920

cacctctttg atgataaggt gatgaaacag cttaaacgtc gccgttatac tggttgggga 1980

cgtttgtctc gaaaattgat taatggtatt agggataagc aatctggcaa aacaatatta 2040

gatttttga aatcagatgg ttttgccaat cgcaatttta tgcagctgat ccatgatgat 2100

agtttgacat ttaaagaaga cattcaaaaa gcacaagtgt ctggacaagg cgatagttta 2160

catgaacata ttgcaaattt agctggtagc cctgctatta aaaaaggtat tttacagact 2220

gtaaaagttg ttgatgaatt ggtcaaagta atggggcggc ataagccaga aaatatcgtt 2280

attgaaatgg cacgtgaaaa tcagacaact caaaagggcc agaaaaattc gcgagagcgt 2340

atgaaacgaa tcgaagaagg tatcaaagaa ttaggaagtc agattcttaa agagcatcct 2400

gttgaaaata ctcaattgca aaatgaaaag ctctatctct attatctcca aaatggaaga 2460

gacatgtatg tggaccaaga attagatatt aatcgtttaa gtgattatga tgtcgatcac 2520

attgttccac aaagtttcct taaagacgat tcaatagaca ataaggtctt aacgcgttct 2580

gataaaaatc gtggtaaatc ggataacgtt ccaagtgaag aagtagtcaa aaagatgaaa 2640

aactattgga gacaacttct aaacgccaag ttaatcactc aacgtaagtt tgataattta 2700

acgaaagctg aacgtggagg tttgagtgaa cttgataaag ctggttttat caaacgccaa 2760

ttggttgaaa ctcgccaaat cactaagcat gtggcacaaa ttttggatag tcgcatgaat 2820

actaaatacg atgaaaatga taaacttatt cgagaggtta aagtgattac cttaaaatct 2880

aaattagttt ctgacttccg aaaagatttc caattctata aagtacgtga gattaacaat 2940

taccatcatg cccatgatgc gtatctaaat gccgtcgttg gaactgcttt gattaagaaa 3000

tatccaaaac ttgaatcgga gtttgtctat ggtgattata aagtttatga tgttcgtaaa 3060

atgattgcta agtctgagca agaaataggc aaagcaaccg caaaatattt cttttactct 3120

aatatcatga acttcttcaa aacagaaatt acacttgcaa atggagagat tcgcaaacgc 3180

cctctaatcg aaactaatgg ggaaactgga gaaattgtct gggataaagg gcgagatttt 3240

gccacagtgc gcaaagtatt gtccatgccc caagtcaata ttgtcaagaa aacagaagta 3300

cagacaggcg gattctccaa ggagtcaatt ttaccaaaaa gaaattcgga caagcttatt 3360

gctcgtaaaa aagactggga tccaaaaaaa tatggtggtt ttgatagtcc aacggtagct 3420

tattcagtcc tagtggttgc taaggtggaa aaagggaaat cgaagaagtt aaaatccgtt 3480

aaagagttac tagggatcac aattatggaa agaagttcct ttgaaaaaaa tccgattgac 3540

tttttagaag ctaaaggata taaggaagtt aaaaaagact taatcattaa actacctaaa 3600

tatagtcttt ttgagttaga aaacggtcgt aaacggatgc tggctagtgc cggagaatta 3660

caaaaaggaa atgagctggc tctgccaagc aaatatgtga atttttata tttagctagt 3720

cattatgaaa agttgaaggg tagtccagaa gataacgaac aaaaacaatt gtttgtggag 3780

cagcataagc attatttaga tgagattatt gagcaaatca gtgaattttc taagcgtgtt 3840

attttagcag atgccaattt agataaagtt cttagtgcat ataacaaaca tagagacaaa 3900

ccaatacgtg aacaagcaga aaatattatt catttattta cgttgacgaa tcttggagct 3960

cccgctgctt ttaaatattt tgatacaaca attgatcgta aacgatatac gtctacaaaa 4020

gaagttttag atgccactct tatccatcaa tccatcactg gtctttatga aacacgcatt 4080

gatttgagtc agctaggagg tgactaa 4107

<210> 109
<211> 1368
<212> PRT
<213> Artificial Sequence

<220>
<223> Amino acid sequence of Cas9 from S.pyogenes

<400> 109
Met Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val
1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
20 25 30

```
Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
         35                  40                  45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
     50                  55                  60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
 65                  70                  75                  80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
                 85                  90                  95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
            100                 105                 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
        115                 120                 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
    130                 135                 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145                 150                 155                 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
                165                 170                 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
            180                 185                 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
        195                 200                 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
    210                 215                 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225                 230                 235                 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
                245                 250                 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
            260                 265                 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
        275                 280                 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
    290                 295                 300
```

```
Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305                 310                 315                 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
                325                 330                 335

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
            340                 345                 350

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
        355                 360                 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
    370                 375                 380

Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385                 390                 395                 400

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
                405                 410                 415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
            420                 425                 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
        435                 440                 445

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
    450                 455                 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465                 470                 475                 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
                485                 490                 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
            500                 505                 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
        515                 520                 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
    530                 535                 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545                 550                 555                 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
                565                 570                 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
```

```
            580                 585                 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
        595                 600                 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
    610                 615                 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625                 630                 635                 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
                645                 650                 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
            660                 665                 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
        675                 680                 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
    690                 695                 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705                 710                 715                 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
                725                 730                 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
            740                 745                 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
        755                 760                 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
    770                 775                 780

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785                 790                 795                 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
                805                 810                 815

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
            820                 825                 830

Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys
        835                 840                 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
    850                 855                 860
```

```
Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865                 870                 875                 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
                885                 890                 895

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
            900                 905                 910

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
        915                 920                 925

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
    930                 935                 940

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
945                 950                 955                 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
                965                 970                 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
            980                 985                 990

Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
        995                1000                1005

Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala Lys
   1010                1015                1020

Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser
1025                1030                1035                1040

Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu
                1045                1050                1055

Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu Ile
            1060                1065                1070

Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser
       1075                1080                1085

Met Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly Gly
   1090                1095                1100

Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile
1105                1110                1115                1120

Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser
                1125                1130                1135
```

```
Pro Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly
           1140                1145                1150

Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile
       1155                1160                1165

Met Glu Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala
   1170                1175                1180

Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys
1185                1190                1195                1200

Tyr Ser Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser
               1205                1210                1215

Ala Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr
           1220                1225                1230

Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser
       1235                1240                1245

Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys His
   1250                1255                1260

Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val
1265                1270                1275                1280

Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys
               1285                1290                1295

His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu
           1300                1305                1310

Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp
       1315                1320                1325

Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu Val Leu Asp
   1330                1335                1340

Ala Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile
1345                1350                1355                1360

Asp Leu Ser Gln Leu Gly Gly Asp
               1365


<210>    110
<211>    4221
<212>    DNA
<213>    Artificial Sequence

<220>
```

<223> Cas9-coding sequence in pET-Cas9N3T for the production of recombinant Cas9 protein in E. coli (humanized codon; hexa-His-tag and a nuclear localization signal at the N terminus)

<400> 110

atgggcagca gccatcatca tcatcatcat gtgtacccct acgacgtgcc cgactacgcc 60

gaattgcctc caaaaaagaa gagaaaggta gggatcgaga acctgtactt ccagggcgac 120

aagaagtaca gcatcggcct ggacatcggt accaacagcg tgggctgggc cgtgatcacc 180

gacgagtaca aggtgcccag caagaagttc aaggtgctgg gcaacaccga ccgccacagc 240

atcaagaaga acctgatcgg cgccctgctg ttcgacagcg gcgagaccgc cgaggccacc 300

cgcctgaagc gcaccgcccg ccgccgctac acccgccgca agaaccgcat ctgctacctg 360

caggagatct tcagcaacga gatggccaag gtggacgaca gcttcttcca ccgcctggag 420

gagagcttcc tggtggagga ggacaagaag cacgagcgcc accccatctt cggcaacatc 480

gtggacgagg tggcctacca cgagaagtac cccaccatct accacctgcg caagaagctg 540

gtggacagca ccgacaaggc cgacctgcgc ctgatctacc tggccctggc ccacatgatc 600

aagttccgcg gccacttcct gatcgagggc gacctgaacc ccgacaacag cgacgtggac 660

aagctgttca tccagctggt gcagacctac aaccagctgt tcgaggagaa ccccatcaac 720

gccagcggcg tggacgccaa ggccatcctg agcgcccgcc tgagcaagag ccgccgcctg 780

gagaacctga tcgcccagct gcccggcgag aagaagaacg gcctgttcgg caacctgatc 840

gccctgagcc tgggcctgac ccccaacttc aagagcaact tcgacctggc cgaggacgcc 900

aagctgcagc tgagcaagga cacctacgac gacgacctgg acaacctgct ggcccagatc 960

ggcgaccagt acgccgacct gttcctggcc gccaagaacc tgagcgacgc catcctgctg 1020

agcgacatcc tgcgcgtgaa caccgagatc accaaggccc ccctgagcgc cagcatgatc 1080

aagcgctacg acgagcacca ccaggacctg accctgctga aggccctggt gcgccagcag 1140

ctgcccgaga agtacaagga gatcttcttc gaccagagca agaacggcta cgccggctac 1200

atcgacggcg gcgccagcca ggaggagttc tacaagttca tcaagcccat cctggagaag 1260

atggacggca ccgaggagct gctggtgaag ctgaaccgcg aggacctgct gcgcaagcag 1320

cgcaccttcg acaacggcag catccccac cagatccacc tgggcgagct gcacgccatc 1380

ctgcgccgcc aggaggactt ctacccctte ctgaaggaca accgcgagaa gatcgagaag 1440

atcctgacct tccgcatccc ctactacgtg ggcccctgg cccgcggcaa cagccgcttc 1500

gcctggatga cccgcaagag cgaggagacc atcaccccct ggaacttcga ggaggtggtg 1560

gacaagggcg ccagcgccca gagcttcatc gagcgcatga ccaacttcga caagaacctg 1620

cccaacgaga aggtgctgcc caagcacagc ctgctgtacg agtacttcac cgtgtacaac 1680

gagctgacca aggtgaagta cgtgaccgag ggcatgcgca agcccgcctt cctgagcggc 1740

gagcagaaga aggccatcgt ggacctgctg ttcaagacca accgcaaggt gaccgtgaag 1800

cagctgaagg aggactactt caagaagatc gagtgcttcg acagcgtgga gatcagcggc 1860

gtggaggacc gcttcaacgc cagcctgggc acctaccacg acctgctgaa gatcatcaag 1920

gacaaggact tcctggacaa cgaggagaac gaggacatcc tggaggacat cgtgctgacc 1980

ctgaccctgt tcgaggaccg cgagatgatc gaggagcgcc tgaagaccta cgcccacctg 2040

ttcgacgaca aggtgatgaa gcagctgaag cgccgccgct acaccggctg gggccgcctg 2100

agccgcaagc ttatcaacgg catccgcgac aagcagagcg gcaagaccat cctggacttc 2160

ctgaagagcg acggcttcgc caaccgcaac ttcatgcagc tgatccacga cgacagcctg 2220

accttcaagg aggacatcca gaaggcccag gtgagcggcc agggcgacag cctgcacgag 2280

cacatcgcca acctggccgg cagccccgcc atcaagaagg gcatcctgca gaccgtgaag 2340

gtggtggacg agctggtgaa ggtgatgggc cgccacaagc ccgagaacat cgtgatcgag 2400

atggcccgcg agaaccagac cacccagaag ggccagaaga acagccgcga gcgcatgaag 2460

cgcatcgagg agggcatcaa ggagctgggc agccagatcc tgaaggagca ccccgtggag 2520

aacacccagc tgcagaacga gaagctgtac ctgtactacc tgcagaacgg ccgcgacatg 2580

tacgtggacc aggagctgga catcaaccgc ctgagcgact acgacgtgga ccacatcgtg 2640

ccccagagct tcctgaagga cgacagcatc gacaacaagg tgctgacccg cagcgacaag 2700

aaccgcggca agagcgacaa cgtgcccagc gaggaggtgg tgaagaagat gaagaactac 2760

tggcgccagc tgctgaacgc caagctgatc acccagcgca agttcgacaa cctgaccaag 2820

gccgagcgcg gcggcctgag cgagctggac aaggccggct tcatcaagcg ccagctggtg 2880

gagacccgcc agatcaccaa gcacgtggcc cagatcctgg acagccgcat gaacaccaag 2940

tacgacgaga acgacaagct gatccgcgag gtgaaggtga tcaccctgaa gagcaagctg 3000

gtgagcgact tccgcaagga cttccagttc tacaaggtgc gcgagatcaa caactaccac 3060

cacgcccacg acgcctacct gaacgccgtg gtgggcaccg ccctgatcaa gaagtacccc 3120

aagctggaga gcgagttcgt gtacggcgac tacaaggtgt acgacgtgcg caagatgatc 3180

gccaagagcg agcaggagat cggcaaggcc accgccaagt acttcttcta cagcaacatc 3240

atgaacttct tcaagaccga gatcaccctg gccaacggcg agatccgcaa gcgcccctg 3300

atcgagacca acggcgagac cggcgagatc gtgtgggaca agggccgcga cttcgccacc 3360

gtgcgcaagg tgctgagcat gccccaggtg aacatcgtga agaagaccga ggtgcagacc 3420

ggcggcttca gcaaggagag catcctgccc aagcgcaaca gcgacaagct gatcgcccgc 3480

aagaaggact gggaccccaa gaagtacggc ggcttcgaca gccccaccgt ggcctacagc 3540

gtgctggtgg tggccaaggt ggagaagggc aagagcaaga agctgaagag cgtgaaggag 3600

ctgctgggca tcaccatcat ggagcgcagc agcttcgaga agaaccccat cgacttcctg 3660

gaggccaagg gctacaagga ggtgaagaag gacctgatca tcaagctgcc caagtacagc 3720

ctgttcgagc tggagaacgg ccgcaagcgc atgctggcca gcgccggcga gctgcagaag 3780

ggcaacgagc tggccctgcc cagcaagtac gtgaacttcc tgtacctggc cagccactac 3840

gagaagctga agggcagccc cgaggacaac gagcagaagc agctgttcgt ggagcagcac 3900

aagcactacc tggacgagat catcgagcag atcagcgagt tcagcaagcg cgtgatcctg 3960

gccgacgcca acctggacaa ggtgctgagc gcctacaaca agcaccgcga caagcccatc 4020

cgcgagcagg ccgagaacat catccacctg ttcaccctga ccaacctggg cgcccccgcc 4080

gccttcaagt acttcgacac caccatcgac cgcaagcgct acaccagcac caaggaggtg 4140

ctggacgcca ccctgatcca ccagagcatc accggtctgt acgagacccg catcgacctg 4200

agccagctgg gcggcgacta a 4221

<210> 111
<211> 1406
<212> PRT
<213> Artificial Sequence

<220>
<223> Amino acid sequence of Cas9 (pET-Cas9N3T)

<400> 111

```
Met Gly Ser Ser His His His His His His Val Tyr Pro Tyr Asp Val
  1               5                  10                  15

Pro Asp Tyr Ala Glu Leu Pro Pro Lys Lys Lys Arg Lys Val Gly Ile
             20                  25                  30

Glu Asn Leu Tyr Phe Gln Gly Asp Lys Lys Tyr Ser Ile Gly Leu Asp
         35                  40                  45

Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys
     50                  55                  60

Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser
 65                  70                  75                  80

Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr
                 85                  90                  95

Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg
            100                 105                 110

Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met
        115                 120                 125

Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu
    130                 135                 140

Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile
145                 150                 155                 160

Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu
                165                 170                 175

Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile
            180                 185                 190

Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile
        195                 200                 205

Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile
    210                 215                 220

Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn
225                 230                 235                 240

Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys
                245                 250                 255

Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys
```

```
            260                 265                 270

Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro
        275                 280                 285

Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu
    290                 295                 300

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

Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp
                325                 330                 335

Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys
            340                 345                 350

Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln
        355                 360                 365

Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys
    370                 375                 380

Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr
385                 390                 395                 400

Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro
                405                 410                 415

Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn
            420                 425                 430

Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile
        435                 440                 445

Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln
    450                 455                 460

Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys
465                 470                 475                 480

Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly
                485                 490                 495

Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr
            500                 505                 510

Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser
        515                 520                 525

Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys
    530                 535                 540
```

```
Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn
545                 550                 555                 560

Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala
                565                 570                 575

Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys
            580                 585                 590

Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys
        595                 600                 605

Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg
    610                 615                 620

Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys
625                 630                 635                 640

Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp
                645                 650                 655

Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu
            660                 665                 670

Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln
        675                 680                 685

Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu
    690                 695                 700

Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe
705                 710                 715                 720

Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His
                725                 730                 735

Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser
            740                 745                 750

Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser
        755                 760                 765

Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu
    770                 775                 780

Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu
785                 790                 795                 800

Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg
                805                 810                 815
```

```
Glu Arg Met Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln
            820                 825                 830

Ile Leu Lys Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn Glu Lys
        835                 840                 845

Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val Asp Gln
    850                 855                 860

Glu Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp His Ile Val
865                 870                 875                 880

Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn Lys Val Leu Thr
                885                 890                 895

Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser Glu Glu
            900                 905                 910

Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys
        915                 920                 925

Leu Ile Thr Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly
    930                 935                 940

Gly Leu Ser Glu Leu Asp Lys Ala Gly Phe Ile Lys Arg Gln Leu Val
945                 950                 955                 960

Glu Thr Arg Gln Ile Thr Lys His Val Ala Gln Ile Leu Asp Ser Arg
                965                 970                 975

Met Asn Thr Lys Tyr Asp Glu Asn Asp Lys Leu Ile Arg Glu Val Lys
            980                 985                 990

Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys Asp Phe
        995                1000                1005

Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr His His Ala His Asp
   1010                1015                1020

Ala Tyr Leu Asn Ala Val Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro
1025                1030                1035                1040

Lys Leu Glu Ser Glu Phe Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val
                1045                1050                1055

Arg Lys Met Ile Ala Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala
           1060                1065                1070

Lys Tyr Phe Phe Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile
       1075                1080                1085

Thr Leu Ala Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn
```

```
1090                1095                1100

Gly Glu Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr
1105                1110                1115                1120

Val Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr
                1125                1130                1135

Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg
            1140                1145                1150

Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys
        1155                1160                1165

Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val Leu Val Val
    1170                1175                1180

Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys Ser Val Lys Glu
1185                1190                1195                1200

Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser Phe Glu Lys Asn Pro
                1205                1210                1215

Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys Glu Val Lys Lys Asp Leu
            1220                1225                1230

Ile Ile Lys Leu Pro Lys Tyr Ser Leu Phe Glu Leu Glu Asn Gly Arg
        1235                1240                1245

Lys Arg Met Leu Ala Ser Ala Gly Glu Leu Gln Lys Gly Asn Glu Leu
    1250                1255                1260

Ala Leu Pro Ser Lys Tyr Val Asn Phe Leu Tyr Leu Ala Ser His Tyr
1265                1270                1275                1280

Glu Lys Leu Lys Gly Ser Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe
                1285                1290                1295

Val Glu Gln His Lys His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser
            1300                1305                1310

Glu Phe Ser Lys Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val
        1315                1320                1325

Leu Ser Ala Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala
    1330                1335                1340

Glu Asn Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala
1345                1350                1355                1360

Ala Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser
                1365                1370                1375
```

```
Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr Gly
          1380                1385                1390

Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp
       1395                1400                1405
```