SEQUENCE LISTING

<110> THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA
Weiner, David B.
Yan, Jian
Obeng-Adjei, Nyamakye

<120> NUCLEIC ACID MOLECULES ENCODING HEPATITIS B VIRUS CORE PROTEIN AND SURFACE ANTIGEN PROTEIN AND VACCINE COMPRISING THE SAME

<130> UPVG0037 WO2

<140> PCT/US13/60618
<141> 2013-09-19

<150> US 13/622,965
<151> 2012-09-19

<160> 21

<170> PatentIn version 3.5

<210> 1
<211> 552
<212> DNA
<213> Artificial Sequence

<220>
<223> nucleic acid sequence M-core

<220>
<221> misc_feature

<222> (1)..(3)

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

<400> 1

nnngacatcg acccctacaa agaattcggc gccaccgtgg aactgctgag cttcctgccc 60

agcgacttct tcccctccgt gcgggacctg ctggataccg ccagcgccct gtacagagag 120

gccctggaaa gccccgagca ctgcagccct caccacacag ccctgcggca ggccatcctg 180

tgctggggcg agctgatgac cctggccacc tgggtcggaa gcaacctgga agatcccgcc 240

agccgggacc tggtggtgtc ctacgtgaac accaacatgg gcctgaagat ccggcagctg 300

ctgtggttcc acatctcctg cctgaccttc ggccgggaaa ccgtgctgga atacctggtg 360

tccttcggcg tgtggatcag aacccccccct gcctacagac cccccaacgc ccctatcctg 420

agcaccctgc ccgagacaac cgtggtccgc agacggggca gaagccccag aagaagaacc 480

cccagcccta gacggcggag atctcagagc cccaggcgga gaagatccca gagccgcgag 540

agccagtgct ga 552

<210> 2

<211> 183

<212> PRT

<213> Artificial Sequence

<220>

<223> Amino Acid Sequence of M-Core

<220>

<221> MISC_FEATURE

<222> (1)..(1)

<223> Xaa can be any naturally occurring amino acid

<400> 2

Xaa Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu Leu
1 5 10 15

Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu Asp
20 25 30

Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His Cys
35 40 45

Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu
50 55 60

Leu Met Thr Leu Ala Thr Trp Val Gly Ser Asn Leu Glu Asp Pro Ala
65 70 75 80

Ser Arg Asp Leu Val Val Ser Tyr Val Asn Thr Asn Met Gly Leu Lys
85 90 95

Ile Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly Arg
100 105 110

Glu Thr Val Leu Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg Thr
115 120 125

Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu Pro
130 135 140

Glu Thr Thr Val Val Arg Arg Arg Gly Arg Ser Pro Arg Arg Arg Thr
145 150 155 160

Pro Ser Pro Arg Arg Arg Arg Ser Gln Ser Pro Arg Arg Arg Arg Ser
165 170 175

Gln Ser Arg Glu Ser Gln Cys
180

<210> 3
<211> 603
<212> DNA
<213> Artificial Sequence

<220>
<223> nucleic acid sequence IgE leader - M-Core

<400> 3
atggactgga cctggattct gttcctggtg gccgctgcca caagggtgca cagcgacatc 60

gacccctaca aagaattcgg cgccaccgtg gaactgctga gcttcctgcc cagcgacttc 120

ttcccctccg tgcgggacct gctggatacc gccagcgccc tgtacagaga ggccctggaa 180

agccccgagc actgcagccc tcaccacaca gccctgcggc aggccatcct gtgctggggc 240

gagctgatga ccctggccac ctgggtcgga agcaacctgg aagatcccgc cagccgggac 300

ctggtggtgt cctacgtgaa caccaacatg ggcctgaaga tccggcagct gctgtggttc 360

cacatctcct gcctgacctt cggccgggaa accgtgctgg aatacctggt gtccttcggc 420

gtgtggatca gaacccccccc tgcctacaga ccccccaacg cccctatcct gagcaccctg 480

cccgagacaa ccgtggtccg cagacggggc agaagcccca gaagaagaac ccccagccct 540

agacggcgga gatctcagag ccccaggcgg agaagatccc agagccgcga gagccagtgc 600

tga 603

<210> 4
<211> 200
<212> PRT
<213> Artificial Sequence

<220>
<223> Amino Acid Sequence of IgE leader - M-Core

<400> 4

Met Asp Trp Thr Trp Ile Leu Phe Leu Val Ala Ala Ala Thr Arg Val
1 5 10 15

His Ser Asp Ile Asp Pro Tyr Lys Glu Phe Gly Ala Thr Val Glu Leu

20 25 30

Leu Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu
35 40 45

Asp Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His
50 55 60

Cys Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly
65 70 75 80

Glu Leu Met Thr Leu Ala Thr Trp Val Gly Ser Asn Leu Glu Asp Pro
85 90 95

Ala Ser Arg Asp Leu Val Val Ser Tyr Val Asn Thr Asn Met Gly Leu
100 105 110

Lys Ile Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly
115 120 125

Arg Glu Thr Val Leu Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg
130 135 140

Thr Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu
145 150 155 160

Pro Glu Thr Thr Val Val Arg Arg Arg Gly Arg Ser Pro Arg Arg Arg
165 170 175

Thr Pro Ser Pro Arg Arg Arg Arg Ser Gln Ser Pro Arg Arg Arg Arg
180 185 190

Ser Gln Ser Arg Glu Ser Gln Cys
195 200

<210> 5
<211> 630
<212> DNA
<213> Artificial Sequence

<220>
<223> nucleic acid sequence IgE leader - M-Core + HA Tag

<400> 5
atggactgga cctggattct gttcctggtg gccgctgcca caagggtgca cagcgacatc 60

gacccctaca aagaattcgg cgccaccgtg gaactgctga gcttcctgcc cagcgacttc 120

ttcccctccg tgcgggacct gctggatacc gccagcgccc tgtacagaga ggccctggaa 180

agccccgagc actgcagccc tcaccacaca gccctgcggc aggccatcct gtgctggggc 240

gagctgatga ccctggccac ctgggtcgga agcaacctgg aagatcccgc cagccgggac 300

ctggtggtgt cctacgtgaa caccaacatg ggcctgaaga tccggcagct gctgtggttc 360

cacatctcct gcctgacctt cggccgggaa accgtgctgg aatacctggt gtccttcggc 420

gtgtggatca gaacccccc tgcctacaga ccccccaacg cccctatcct gagcaccctg 480

cccgagacaa ccgtggtccg cagacggggc agaagcccca gaagaagaac ccccagccct 540

agacggcgga gatctcagag ccccaggcgg agaagatccc agagccgcga gagccagtgc 600

taccctacg acgtgcccga ctacgcctga 630

<210> 6
<211> 209
<212> PRT
<213> Artificial Sequence

<220>
<223> Amino Acid Sequence of IgE leader - M-Core + HA Tag

<400> 6

Met Asp Trp Thr Trp Ile Leu Phe Leu Val Ala Ala Ala Thr Arg Val
1 5 10 15

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

Leu Ser Phe Leu Pro Ser Asp Phe Phe Pro Ser Val Arg Asp Leu Leu
35 40 45

Asp Thr Ala Ser Ala Leu Tyr Arg Glu Ala Leu Glu Ser Pro Glu His

```
    50                  55                  60

Cys Ser Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly
65                  70                  75                  80

Glu Leu Met Thr Leu Ala Thr Trp Val Gly Ser Asn Leu Glu Asp Pro
                85                  90                  95

Ala Ser Arg Asp Leu Val Val Ser Tyr Val Asn Thr Asn Met Gly Leu
            100                 105                 110

Lys Ile Arg Gln Leu Leu Trp Phe His Ile Ser Cys Leu Thr Phe Gly
        115                 120                 125

Arg Glu Thr Val Leu Glu Tyr Leu Val Ser Phe Gly Val Trp Ile Arg
    130                 135                 140

Thr Pro Pro Ala Tyr Arg Pro Pro Asn Ala Pro Ile Leu Ser Thr Leu
145                 150                 155                 160

Pro Glu Thr Thr Val Val Arg Arg Arg Gly Arg Ser Pro Arg Arg Arg
                165                 170                 175

Thr Pro Ser Pro Arg Arg Arg Arg Ser Gln Ser Pro Arg Arg Arg Arg
            180                 185                 190
```

Ser Gln Ser Arg Glu Ser Gln Cys Tyr Pro Tyr Asp Val Pro Asp Tyr
195 200 205

Ala

<210> 7
<211> 18
<212> PRT
<213> Artificial Sequence

<220>
<223> IgE leader amino acid sequence

<400> 7

Met Asp Trp Thr Trp Ile Leu Phe Leu Val Ala Ala Ala Thr Arg Val
1 5 10 15

His Ser

<210> 8
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> HA Tag amino acid sequence

<400> 8

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala
1 5

<210> 9
<211> 1317
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic (LHBs-A)

<400> 9
ggatccgcca ccatggattg gacttggatt ctgttcctgg tcgccgctgc tacacgggtg 60

cattcagggg gctggtcttc aaaacctaga aaaggcatgg gcaccaacct gagcgtgccc 120

aatcctctgg ggttctttcc agaccaccag ctggaccccg ctttcggcgc aaactccaac 180

aatcctgact gggacttcaa cccaatcaaa gaccactggc cagcagctaa ccaagtggga 240

gtcggagctt tcggaccagg actgactccc cctcatggcg ggattctggg ctggtctccc 300

caggctcagg gcatcctgac cacagtgagc actattccac cccctgcaag caccaacagg 360

cagtccggac gccagccaac cccaatctca ccacccctgc gagacagcca ccctcaggcc 420

agaggcagga aacggagatc tatgcagtgg aatagtacag ccttccatca ggctctgcag 480

gacccccggg tgcggggact gtactttcca gccggaggca gctcctctgg cactgtcaac 540

cctgcaccaa atatcgcctc ccacatcagt tcaatttctg ctcgaactgg agaccccgtg 600

accaaccggg gcagaaagag gcgcagtatg gagaatatta cctcagggtt cctgggacct 660

ctgctggtcc tgcaggcagg cttctttctg ctgacccgca tcctgacaat tcctcagtca 720

ctggatagct ggtggaccag cctgaacttc ctgggcggca gccccgtgtg cctgggacag 780

aactctcaga gtcctacctc caatcattct ccaacaagtt gtcctccaat ctgcccaggc 840

tacagatgga tgtgcctgcg acggttcatc attttcctgt ttatcctgct gctgtgcctg 900

atttttctgc tggtgctgct ggactatcag ggcatgctgc cagtctgccc cctgattcct 960

gggtccacta ccacatctac aggaccctgt aagacttgca ctacccctgc ccaggggaac 1020

agtatgtttc catcatgctg ttgcacaaaa cccactgatg gaaattgtac atgcatcccc 1080

attcctagct cctgggcatt cgccaagtat ctgtgggaat gggcaagcgt gaggttttca 1140

tggctgagcc tgctggtgcc cttcgtccag tggtttgtgg gactgagccc taccgtctgg 1200

ctgtccgcca tctggatgat gtggtactgg gggcccagcc tgtattcaat cgtgtctcca 1260

ttcatccccc tgctgccaat cttcttttgt ctgtgggtct acatttgata actcgag 1317

<210> 10

<211> 431

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic (LHBs-A)

<400> 10

```
Met Asp Trp Thr Trp Ile Leu Phe Leu Val Ala Ala Ala Thr Arg Val
1               5                  10                  15

His Ser Gly Gly Trp Ser Ser Lys Pro Arg Lys Gly Met Gly Thr Asn
            20                  25                  30

Leu Ser Val Pro Asn Pro Leu Gly Phe Phe Pro Asp His Gln Leu Asp
        35                  40                  45

Pro Ala Phe Gly Ala Asn Ser Asn Asn Pro Asp Trp Asp Phe Asn Pro
    50                  55                  60

Ile Lys Asp His Trp Pro Ala Ala Asn Gln Val Gly Val Gly Ala Phe
65                  70                  75                  80

Gly Pro Gly Leu Thr Pro Pro His Gly Gly Ile Leu Gly Trp Ser Pro
                85                  90                  95

Gln Ala Gln Gly Ile Leu Thr Thr Val Ser Thr Ile Pro Pro Pro Ala
            100                 105                 110

Ser Thr Asn Arg Gln Ser Gly Arg Gln Pro Thr Pro Ile Ser Pro Pro
        115                 120                 125
```

```
Leu Arg Asp Ser His Pro Gln Ala Arg Gly Arg Lys Arg Arg Ser Met
    130                 135                 140

Gln Trp Asn Ser Thr Ala Phe His Gln Ala Leu Gln Asp Pro Arg Val
145                 150                 155                 160

Arg Gly Leu Tyr Phe Pro Ala Gly Gly Ser Ser Ser Gly Thr Val Asn
                165                 170                 175

Pro Ala Pro Asn Ile Ala Ser His Ile Ser Ser Ile Ser Ala Arg Thr
            180                 185                 190

Gly Asp Pro Val Thr Asn Arg Gly Arg Lys Arg Arg Ser Met Glu Asn
        195                 200                 205

Ile Thr Ser Gly Phe Leu Gly Pro Leu Leu Val Leu Gln Ala Gly Phe
    210                 215                 220

Phe Leu Leu Thr Arg Ile Leu Thr Ile Pro Gln Ser Leu Asp Ser Trp
225                 230                 235                 240

Trp Thr Ser Leu Asn Phe Leu Gly Gly Ser Pro Val Cys Leu Gly Gln
                245                 250                 255

Asn Ser Gln Ser Pro Thr Ser Asn His Ser Pro Thr Ser Cys Pro Pro
```

260 265 270

Ile Cys Pro Gly Tyr Arg Trp Met Cys Leu Arg Arg Phe Ile Ile Phe
275 280 285

Leu Phe Ile Leu Leu Leu Cys Leu Ile Phe Leu Leu Val Leu Leu Asp
290 295 300

Tyr Gln Gly Met Leu Pro Val Cys Pro Leu Ile Pro Gly Ser Thr Thr
305 310 315 320

Thr Ser Thr Gly Pro Cys Lys Thr Cys Thr Thr Pro Ala Gln Gly Asn
325 330 335

Ser Met Phe Pro Ser Cys Cys Cys Thr Lys Pro Thr Asp Gly Asn Cys
340 345 350

Thr Cys Ile Pro Ile Pro Ser Ser Trp Ala Phe Ala Lys Tyr Leu Trp
355 360 365

Glu Trp Ala Ser Val Arg Phe Ser Trp Leu Ser Leu Leu Val Pro Phe
370 375 380

Val Gln Trp Phe Val Gly Leu Ser Pro Thr Val Trp Leu Ser Ala Ile
385 390 395 400

Trp Met Met Trp Tyr Trp Gly Pro Ser Leu Tyr Ser Ile Val Ser Pro
405 410 415

Phe Ile Pro Leu Leu Pro Ile Phe Phe Cys Leu Trp Val Tyr Ile
420 425 430

<210> 11
<211> 1319
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic (LHBs-C)

<400> 11
gaattcgcca ccatggattg gacatggatt ctgtttctgg tcgccgccgc aacccgcgtg 60

cactcagggg gatggtcatc aaaacctaga cagggaatgg gcactaacct gagtgtgccc 120

aatcctctgg ggttctttcc cgaccaccag ctggatcctg ccttcggcgc taactctaac 180

aatccagact gggacttcaa ccccaataag gaccactggc ctgaggcaaa tcaagtggga 240

gcaggagcct tcggaccagg ctttacaccc cctcatggcg gactgctggg atggtcccct 300

caggctcagg ggatcctgac cacagtccca gcagctccac cccctgcaag tactaacagg 360

cagtcaggac gccagccaac ccccatttct ccacccctga gggacagtca ccctcaggcc 420

agaggcagga agcggagaag catgcagtgg aacagcacta ccttccatca ggcactgctg 480

gatccacgcg tgcgaggact gtactttcca gccggaggca gctcctctgg aaccgtgaac 540

cctgtcccaa caactgcctc cccaatcagt tcaattttct ctcggacagg agaccccgct 600

cctaatcggg gcagaaaaag gcgctcaatg gaaagcacca catccgggtt tctgggacca 660

ctgctggtgc tgcaggcagg cttctttctg ctgaccagaa tcctgacaat tccccagtct 720

ctggatagtt ggtggaccag cctgaacttc ctgggcggcg cccctacttg tccaggacag 780

aactctcaga gtccaacatc aaatcatagc cccacttcct gtcctccaat ctgccctggc 840

taccgctgga tgtgcctgcg acggttcatc attttcctgt ttatcctgct gctgtgcctg 900

atttttctgc tggtgctgct ggactatcag ggaatgctgc ccgtctgccc tctgctgcca 960

gggacttcta ctaccagtac cggaccttgt aagacatgca ctattccagc tcaggggacc 1020

tccatgttcc cctcttgctg ttgcaccaaa cctagcgatg gaaattgtac atgcatccca 1080

attcccagct cctgggcttt cgcacgattt ctgtgggagt gggccagcgt gcgcttttca 1140

tggctgagcc tgctggtgcc cttcgtccag tggtttgtcg gcctgtcacc taccgtgtgg 1200

ctgagcgtca tctggatgat gtggtactgg gggcccagcc tgtataatat cctgtcacca 1260

ttcctgccac tgctgccaat cttcttttgt ctgtgggtct acatttgata agcggccgc 1319

<210> 12

<211> 431

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic (LHBs-C)

<400> 12

```
Met Asp Trp Thr Trp Ile Leu Phe Leu Val Ala Ala Ala Thr Arg Val
1               5                  10                  15

His Ser Gly Gly Trp Ser Ser Lys Pro Arg Gln Gly Met Gly Thr Asn
            20                  25                  30

Leu Ser Val Pro Asn Pro Leu Gly Phe Phe Pro Asp His Gln Leu Asp
        35                  40                  45

Pro Ala Phe Gly Ala Asn Ser Asn Asn Pro Asp Trp Asp Phe Asn Pro
    50                  55                  60

Asn Lys Asp His Trp Pro Glu Ala Asn Gln Val Gly Ala Gly Ala Phe
65                  70                  75                  80

Gly Pro Gly Phe Thr Pro Pro His Gly Gly Leu Leu Gly Trp Ser Pro
                85                  90                  95

Gln Ala Gln Gly Ile Leu Thr Thr Val Pro Ala Ala Pro Pro Pro Ala
            100                 105                 110

Ser Thr Asn Arg Gln Ser Gly Arg Gln Pro Thr Pro Ile Ser Pro Pro
```

115 120 125

Leu Arg Asp Ser His Pro Gln Ala Arg Gly Arg Lys Arg Arg Ser Met
130 135 140

Gln Trp Asn Ser Thr Thr Phe His Gln Ala Leu Leu Asp Pro Arg Val
145 150 155 160

Arg Gly Leu Tyr Phe Pro Ala Gly Gly Ser Ser Ser Gly Thr Val Asn
165 170 175

Pro Val Pro Thr Thr Ala Ser Pro Ile Ser Ser Ile Phe Ser Arg Thr
180 185 190

Gly Asp Pro Ala Pro Asn Arg Gly Arg Lys Arg Arg Ser Met Glu Ser
195 200 205

Thr Thr Ser Gly Phe Leu Gly Pro Leu Leu Val Leu Gln Ala Gly Phe
210 215 220

Phe Leu Leu Thr Arg Ile Leu Thr Ile Pro Gln Ser Leu Asp Ser Trp
225 230 235 240

Trp Thr Ser Leu Asn Phe Leu Gly Gly Ala Pro Thr Cys Pro Gly Gln
245 250 255

```
Asn Ser Gln Ser Pro Thr Ser Asn His Ser Pro Thr Ser Cys Pro Pro
            260                 265                 270

Ile Cys Pro Gly Tyr Arg Trp Met Cys Leu Arg Arg Phe Ile Ile Phe
        275                 280                 285

Leu Phe Ile Leu Leu Leu Cys Leu Ile Phe Leu Leu Val Leu Leu Asp
    290                 295                 300

Tyr Gln Gly Met Leu Pro Val Cys Pro Leu Leu Pro Gly Thr Ser Thr
305                 310                 315                 320

Thr Ser Thr Gly Pro Cys Lys Thr Cys Thr Ile Pro Ala Gln Gly Thr
                325                 330                 335

Ser Met Phe Pro Ser Cys Cys Cys Thr Lys Pro Ser Asp Gly Asn Cys
            340                 345                 350

Thr Cys Ile Pro Ile Pro Ser Ser Trp Ala Phe Ala Arg Phe Leu Trp
        355                 360                 365

Glu Trp Ala Ser Val Arg Phe Ser Trp Leu Ser Leu Leu Val Pro Phe
    370                 375                 380

Val Gln Trp Phe Val Gly Leu Ser Pro Thr Val Trp Leu Ser Val Ile
```

385 390 395 400

Trp Met Met Trp Tyr Trp Gly Pro Ser Leu Tyr Asn Ile Leu Ser Pro
405 410 415

Phe Leu Pro Leu Leu Pro Ile Phe Phe Cys Leu Trp Val Tyr Ile
420 425 430

<210> 13
<211> 753
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic (SHBs-A)

<400> 13
ggatccgcca ccatggactg gacctggatt ctgttcctgg tggctgccgc tacacgggtg 60

cattctgaaa atatcacatc tggattcctg ggacctctgc tggtgctgca ggctgggttc 120

tttctgctga caagaatcct gactattccc cagtcactgg acagctggtg gacatctctg 180

aacttcctgg gcgggagtcc tgtctgtctg ggacagaact ctcagagtcc tacttccaat 240

cactctccaa ccagttgtcc ccctatctgc ccaggctacc gctggatgtg cctgcggaga 300

ttcatcattt tcctgtttat cctgctgctg tgcctgattt ttctgctggt gctgctggac 360

tatcagggca tgctgcctgt ctgcccactg attcccggca gcaccacaac ttctaccggc 420

ccctgtaaga catgcaccac acctgcccag gggaacagta tgtttccatc atgctgttgc 480

actaaaccca ccgatggaaa ttgtacatgc atcccaattc ccagctcctg ggccttcgct 540

aagtacctgt gggagtgggc aagcgtgcga ttttcatggc tgagcctgct ggtgcctttc 600

gtccagtggt ttgtgggcct gagcccaact gtctggctgt ccgccatctg gatgatgtgg 660

tactgggggc ccagcctgta ttccatcgtg tcaccattca ttcccctgct gccaatcttt 720

ttctgcctgt gggtctacat ctgataactc gag 753

<210> 14

<211> 243

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic (SHBs-A)

<400> 14

Met Asp Trp Thr Trp Ile Leu Phe Leu Val Ala Ala Ala Thr Arg Val
1 5 10 15

His Ser Glu Asn Ile Thr Ser Gly Phe Leu Gly Pro Leu Leu Val Leu
20 25 30

Gln Ala Gly Phe Phe Leu Leu Thr Arg Ile Leu Thr Ile Pro Gln Ser
35 40 45

Leu Asp Ser Trp Trp Thr Ser Leu Asn Phe Leu Gly Gly Ser Pro Val
50 55 60

Cys Leu Gly Gln Asn Ser Gln Ser Pro Thr Ser Asn His Ser Pro Thr
65 70 75 80

Ser Cys Pro Pro Ile Cys Pro Gly Tyr Arg Trp Met Cys Leu Arg Arg
85 90 95

Phe Ile Ile Phe Leu Phe Ile Leu Leu Leu Cys Leu Ile Phe Leu Leu
100 105 110

Val Leu Leu Asp Tyr Gln Gly Met Leu Pro Val Cys Pro Leu Ile Pro
115 120 125

Gly Ser Thr Thr Thr Ser Thr Gly Pro Cys Lys Thr Cys Thr Thr Pro
130 135 140

Ala Gln Gly Asn Ser Met Phe Pro Ser Cys Cys Cys Thr Lys Pro Thr
145 150 155 160

Asp Gly Asn Cys Thr Cys Ile Pro Ile Pro Ser Ser Trp Ala Phe Ala
165 170 175

Lys Tyr Leu Trp Glu Trp Ala Ser Val Arg Phe Ser Trp Leu Ser Leu

180 185 190

Leu Val Pro Phe Val Gln Trp Phe Val Gly Leu Ser Pro Thr Val Trp
195 200 205

Leu Ser Ala Ile Trp Met Met Trp Tyr Trp Gly Pro Ser Leu Tyr Ser
210 215 220

Ile Val Ser Pro Phe Ile Pro Leu Leu Pro Ile Phe Phe Cys Leu Trp
225 230 235 240

Val Tyr Ile

<210> 15
<211> 755
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic (SHBs-C)

<400> 15
gaattcgcca ccatggattg gacatggatt ctgttcctgg tcgccgccgc aacacgagtg 60

cattctgaaa gtacaacctc tggcttcctg ggccccctgc tggtgctgca ggcagggttc 120

tttctgctga cacgaatcct gactattcca cagtcactgg acagctggtg gaccagcctg 180

aacttcctgg gcggggcccc tacatgtcca ggacagaact ctcagagtcc cacttccaat 240

cactctccta ccagttgtcc ccctatctgc cccggctaca gatggatgtg cctgcggaga 300

ttcatcattt tcctgtttat cctgctgctg tgcctgattt ttctgctggt gctgctggac 360

tatcagggaa tgctgcccgt ctgccctctg ctgccaggaa cctcaaccac aagcacaggc 420

ccttgtaaga cttgcaccat tcccgctcag gggactagta tgttcccttc atgctgttgc 480

acaaaaccat ctgatggaaa ttgtacttgc atcccaattc ccagctcctg ggccttcgct 540

aggtttctgt gggagtgggc cagtgtgcgc ttttcctggc tgtctctgct ggtgcccttc 600

gtccagtggt ttgtcggcct gagccctaca gtgtggctgt ccgtcatctg gatgatgtgg 660

tactgggggc ctagcctgta taatatcctg tcaccatttc tgccactgct gccaatcttc 720

ttttgcctgt gggtctacat ctgataagcg gccgc 755

<210> 16

<211> 243

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic (SHBs-C)

<400> 16

Met Asp Trp Thr Trp Ile Leu Phe Leu Val Ala Ala Ala Thr Arg Val

1 5 10 15

```
His Ser Glu Ser Thr Thr Ser Gly Phe Leu Gly Pro Leu Leu Val Leu
            20                  25                  30

Gln Ala Gly Phe Phe Leu Leu Thr Arg Ile Leu Thr Ile Pro Gln Ser
        35                  40                  45

Leu Asp Ser Trp Trp Thr Ser Leu Asn Phe Leu Gly Gly Ala Pro Thr
    50                  55                  60

Cys Pro Gly Gln Asn Ser Gln Ser Pro Thr Ser Asn His Ser Pro Thr
65                  70                  75                  80

Ser Cys Pro Pro Ile Cys Pro Gly Tyr Arg Trp Met Cys Leu Arg Arg
                85                  90                  95

Phe Ile Ile Phe Leu Phe Ile Leu Leu Leu Cys Leu Ile Phe Leu Leu
            100                 105                 110

Val Leu Leu Asp Tyr Gln Gly Met Leu Pro Val Cys Pro Leu Leu Pro
        115                 120                 125

Gly Thr Ser Thr Thr Ser Thr Gly Pro Cys Lys Thr Cys Thr Ile Pro
    130                 135                 140

Ala Gln Gly Thr Ser Met Phe Pro Ser Cys Cys Cys Thr Lys Pro Ser
```

145 150 155 160

Asp Gly Asn Cys Thr Cys Ile Pro Ile Pro Ser Ser Trp Ala Phe Ala
165 170 175

Arg Phe Leu Trp Glu Trp Ala Ser Val Arg Phe Ser Trp Leu Ser Leu
180 185 190

Leu Val Pro Phe Val Gln Trp Phe Val Gly Leu Ser Pro Thr Val Trp
195 200 205

Leu Ser Val Ile Trp Met Met Trp Tyr Trp Gly Pro Ser Leu Tyr Asn
210 215 220

Ile Leu Ser Pro Phe Leu Pro Leu Leu Pro Ile Phe Phe Cys Leu Trp
225 230 235 240

Val Tyr Ile

<210> 17
<211> 3585
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic (pGX1801 HepB - MCore)

<400> 17

gactcttcgc gatgtacggg ccagatatac gcgttgacat tgattattga ctagttatta 60

atagtaatca attacggggt cattagttca tagcccatat atggagttcc gcgttacata 120

acttacggta aatggcccgc ctggctgacc gcccaacgac ccccgcccat tgacgtcaat 180

aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc aatgggtgga 240

ctatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc caagtacgcc 300

ccctattgac gtcaatgacg gtaaatggcc cgcctggcat tatgcccagt acatgacctt 360

atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta ccatggtgat 420

gcggttttgg cagtacatca atgggcgtgg atagcggttt gactcacggg gatttccaag 480

tctccacccc attgacgtca atgggagttt gttttggcac caaaatcaac gggactttcc 540

aaaatgtcgt aacaactccg ccccattgac gcaaatgggc ggtaggcgtg tacggtggga 600

ggtctatata agcagagctc tctggctaac tagagaaccc actgcttact ggcttatcga 660

aattaatacg actcactata gggagaccca agctggctag cgtttaaact taagcttggt 720

accgagctcg gatccgccac catggactgg acctggattc tgttcctggt ggccgctgcc 780

acaagggtgc acagcgacat cgacccctac aaagaattcg gcgccaccgt ggaactgctg 840

agcttcctgc ccagcgactt cttcccctcc gtgcgggacc tgctggatac cgccagcgcc 900

ctgtacagag aggccctgga aagccccgag cactgcagcc ctcaccacac agccctgcgg 960

caggccatcc tgtgctgggg cgagctgatg accctggcca cctgggtcgg aagcaacctg 1020

gaagatcccg ccagccggga cctggtggtg tcctacgtga acaccaacat gggcctgaag 1080

atccggcagc tgctgtggtt ccacatctcc tgcctgacct tcggccggga aaccgtgctg 1140

gaatacctgg tgtccttcgg cgtgtggatc agaacccccc ctgcctacag acccccccaac 1200

gcccctatcc tgagcaccct gcccgagaca accgtggtcc gcagacgggg cagaagcccc 1260

agaagaagaa cccccagccc tagacggcgg agatctcaga gccccaggcg gagaagatcc 1320

cagagccgcg agagccagtg ctacccctac gacgtgcccg actacgcctg actcgagtct 1380

agagggcccg tttaaacccg ctgatcagcc tcgactgtgc cttctagttg ccagccatct 1440

gttgtttgcc cctcccccgt gccttccttg accctggaag gtgccactcc cactgtcctt 1500

tcctaataaa atgaggaaat tgcatcgcat tgtctgagta ggtgtcattc tattctgggg 1560

ggtggggtgg ggcaggacag caaggggggag gattgggaag acaatagcag gcatgctggg 1620

gatgcggtgg gctctatggc ttctactggg cggttttatg gacagcaagc gaaccggaat 1680

tgccagctgg ggcgccctct ggtaaggttg ggaagccctg caaagtaaac tggatggctt 1740

tctcgccgcc aaggatctga tggcgcaggg gatcaagctc tgatcaagag acaggatgag 1800

gatcgtttcg catgattgaa caagatggat tgcacgcagg ttctccggcc gcttgggtgg 1860

agaggctatt cggctatgac tgggcacaac agacaatcgg ctgctctgat gccgccgtgt 1920

tccggctgtc agcgcagggg cgcccggttc tttttgtcaa gaccgacctg tccggtgccc 1980

tgaatgaact gcaagacgag gcagcgcggc tatcgtggct ggccacgacg ggcgttcctt 2040

gcgcagctgt gctcgacgtt gtcactgaag cgggaaggga ctggctgcta ttgggcgaag 2100

tgccggggca ggatctcctg tcatctcacc ttgctcctgc cgagaaagta tccatcatgg 2160

ctgatgcaat gcggcggctg catacgcttg atccggctac ctgcccattc gaccaccaag 2220

cgaaacatcg catcgagcga gcacgtactc ggatggaagc cggtcttgtc gatcaggatg 2280

atctggacga agagcatcag gggctcgcgc cagccgaact gttcgccagg ctcaaggcga 2340

gcatgcccga cggcgaggat ctcgtcgtga cccatggcga tgcctgcttg ccgaatatca 2400

tggtggaaaa tggccgcttt tctggattca tcgactgtgg ccggctgggt gtggcggacc 2460

gctatcagga catagcgttg gctacccgtg atattgctga agagcttggc ggcgaatggg 2520

ctgaccgctt cctcgtgctt tacggtatcg ccgctcccga ttcgcagcgc atcgccttct 2580

atcgccttct tgacgagttc ttctgaatta ttaacgctta caatttcctg atgcggtatt 2640

ttctccttac gcatctgtgc ggtatttcac accgcataca ggtggcactt ttcggggaaa 2700

tgtgcgcgga acccctattt gtttatttttt ctaaatacat tcaaatatgt atccgctcat 2760

gagacaataa ccctgataaa tgcttcaata atagcacgtg ctaaaacttc atttttaatt 2820

taaaaggatc taggtgaaga tccttttttga taatctcatg accaaaatcc cttaacgtga 2880

gttttcgttc cactgagcgt cagaccccgt agaaaagatc aaaggatctt cttgagatcc 2940

ttttttttctg cgcgtaatct gctgcttgca aacaaaaaaa ccaccgctac cagcggtggt 3000

ttgtttgccg gatcaagagc taccaactct ttttccgaag gtaactggct tcagcagagc 3060

gcagatacca aatactgtcc ttctagtgta gccgtagtta ggccaccact tcaagaactc 3120

tgtagcaccg cctacatacc tcgctctgct aatcctgtta ccagtggctg ctgccagtgg 3180

cgataagtcg tgtcttaccg ggttggactc aagacgatag ttaccggata aggcgcagcg 3240

gtcgggctga acgggggggtt cgtgcacaca gcccagcttg gagcgaacga cctacaccga 3300

actgagatac ctacagcgtg agctatgaga aagcgccacg cttcccgaag ggagaaaggc 3360

ggacaggtat ccggtaagcg gcagggtcgg aacaggagag cgcacgaggg agcttccagg 3420

gggaaacgcc tggtatcttt atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg 3480

atttttgtga tgctcgtcag gggggcggag cctatggaaa aacgccagca acgcggcctt 3540

tttacggttc ctgggctttt gctggccttt tgctcacatg ttctt 3585

<210> 18
<211> 4254
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic (pGX1802 HepB pLHBs-A)

<400> 18
gactcttcgc gatgtacggg ccagatatac gcgttgacat tgattattga ctagttatta 60

atagtaatca attacggggt cattagttca tagcccatat atggagttcc gcgttacata 120

acttacggta aatggcccgc ctggctgacc gcccaacgac ccccgcccat tgacgtcaat 180

aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc aatgggtgga 240

ctatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc caagtacgcc 300

ccctattgac gtcaatgacg gtaaatggcc cgcctggcat tatgcccagt acatgacctt 360

atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta ccatggtgat 420

gcggttttgg cagtacatca atgggcgtgg atagcggttt gactcacggg gatttccaag 480

tctccacccc attgacgtca atgggagttt gttttggcac caaaatcaac gggactttcc 540

aaaatgtcgt aacaactccg ccccattgac gcaaatgggc ggtaggcgtg tacggtggga 600

ggtctatata agcagagctc tctggctaac tagagaaccc actgcttact ggcttatcga 660

aattaatacg actcactata gggagaccca agctggctag cgtttaaact taagcttggt 720

accgagctcg gatccgccac catggattgg acttggattc tgttcctggt cgccgctgct 780

acacgggtgc attcaggggg ctggtcttca aaacctagaa aaggcatggg caccaacctg 840

agcgtgccca atcctctggg gttctttcca gaccaccagc tggaccccgc tttcggcgca 900

aactccaaca atcctgactg ggacttcaac ccaatcaaag accactggcc agcagctaac 960

caagtgggag tcggagcttt cggaccagga ctgactcccc ctcatggcgg gattctgggc 1020

tggtctcccc aggctcaggg catcctgacc acagtgagca ctattccacc ccctgcaagc 1080

accaacaggc agtccggacg ccagccaacc ccaatctcac cacccctgcg agacagccac 1140

cctcaggcca gaggcaggaa acggagatct atgcagtgga atagtacagc cttccatcag 1200

gctctgcagg acccccgggt gcggggactg tactttccag ccggaggcag ctcctctggc 1260

actgtcaacc ctgcaccaaa tatcgcctcc cacatcagtt caatttctgc tcgaactgga 1320

gaccccgtga ccaaccgggg cagaaagagg cgcagtatgg agaatattac ctcagggttc 1380

ctgggacctc tgctggtcct gcaggcaggc ttctttctgc tgacccgcat cctgacaatt 1440

cctcagtcac tggatagctg gtggaccagc ctgaacttcc tgggcggcag ccccgtgtgc 1500

ctgggacaga actctcagag tcctacctcc aatcattctc caacaagttg tcctccaatc 1560

tgcccaggct acagatggat gtgcctgcga cggttcatca ttttcctgtt tatcctgctg 1620

ctgtgcctga tttttctgct ggtgctgctg gactatcagg gcatgctgcc agtctgcccc 1680

ctgattcctg ggtccactac cacatctaca ggaccctgta agacttgcac tacccctgcc 1740

caggggaaca gtatgtttcc atcatgctgt tgcacaaaac ccactgatgg aaattgtaca 1800

tgcatcccca ttcctagctc ctgggcattc gccaagtatc tgtgggaatg ggcaagcgtg 1860

aggttttcat ggctgagcct gctggtgccc ttcgtccagt ggtttgtggg actgagccct 1920

accgtctggc tgtccgccat ctggatgatg tggtactggg ggcccagcct gtattcaatc 1980

gtgtctccat tcatcccccc gctgccaatc ttcttttgtc tgtgggtcta catttgataa 2040

ctcgagtcta gagggcccgt ttaaacccgc tgatcagcct cgactgtgcc ttctagttgc 2100

cagccatctg ttgtttgccc ctcccccgtg ccttccttga ccctggaagg tgccactccc 2160

actgtccttt cctaataaaa tgaggaaatt gcatcgcatt gtctgagtag gtgtcattct 2220

attctggggg gtggggtggg gcaggacagc aaggggggagg attgggaaga caatagcagg 2280

catgctgggg atgcggtggg ctctatggct tctactgggc ggttttatgg acagcaagcg 2340

aaccggaatt gccagctggg gcgccctctg gtaaggttgg gaagccctgc aaagtaaact 2400

ggatggcttt ctcgccgcca aggatctgat ggcgcagggg atcaagctct gatcaagaga 2460

caggatgagg atcgtttcgc atgattgaac aagatggatt gcacgcaggt tctccggccg 2520

cttgggtgga gaggctattc ggctatgact gggcacaaca gacaatcggc tgctctgatg 2580

ccgccgtgtt ccggctgtca gcgcaggggc gcccggttct ttttgtcaag accgacctgt 2640

ccggtgccct gaatgaactg caagacgagg cagcgcggct atcgtggctg gccacgacgg 2700

gcgttccttg cgcagctgtg ctcgacgttg tcactgaagc gggaagggac tggctgctat 2760

tgggcgaagt gccggggcag gatctcctgt catctcacct tgctcctgcc gagaaagtat 2820

ccatcatggc tgatgcaatg cggcggctgc atacgcttga tccggctacc tgcccattcg 2880

accaccaagc gaaacatcgc atcgagcgag cacgtactcg gatggaagcc ggtcttgtcg 2940

atcaggatga tctggacgaa gagcatcagg ggctcgcgcc agccgaactg ttcgccaggc 3000

tcaaggcgag catgcccgac ggcgaggatc tcgtcgtgac ccatggcgat gcctgcttgc 3060

cgaatatcat ggtggaaaat ggccgctttt ctggattcat cgactgtggc cggctgggtg 3120

tggcggaccg ctatcaggac atagcgttgg ctacccgtga tattgctgaa gagcttggcg 3180

gcgaatgggc tgaccgcttc ctcgtgcttt acggtatcgc cgctcccgat tcgcagcgca 3240

tcgccttcta tcgccttctt gacgagttct tctgaattat taacgcttac aatttcctga 3300

tgcggtattt tctccttacg catctgtgcg gtatttcaca ccgcatacag gtggcacttt 3360

tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt caaatatgta 3420

tccgctcatg agacaataac cctgataaat gcttcaataa tagcacgtgc taaaacttca 3480

ttttttaattt aaaaggatct aggtgaagat ccttttttgat aatctcatga ccaaaatccc 3540

ttaacgtgag ttttcgttcc actgagcgtc agaccccgta gaaaagatca aaggatcttc 3600

ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac caccgctacc 3660

agcggtggtt tgtttgccgg atcaagagct accaactctt tttccgaagg taactggctt 3720

cagcagagcg cagataccaa atactgtcct tctagtgtag ccgtagttag gccaccactt 3780

caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac cagtggctgc 3840

tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt taccggataa 3900

ggcgcagcgg tcgggctgaa cgggggttc gtgcacacag cccagcttgg agcgaacgac 3960

ctacaccgaa ctgagatacc tacagcgtga gctatgagaa agcgccacgc ttcccgaagg 4020

gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga acaggagagc gcacgaggga 4080

gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc acctctgact 4140

tgagcgtcga tttttgtgat gctcgtcagg ggggcggagc ctatggaaaa acgccagcaa 4200

cgcggccttt ttacggttcc tgggcttttg ctggcctttt gctcacatgt tctt 4254

<210> 19
<211> 4284
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic (pGX1803 HepB pLHBs-C)

<400> 19
gactcttcgc gatgtacggg ccagatatac gcgttgacat tgattattga ctagttatta 60

atagtaatca attacggggt cattagttca tagcccatat atggagttcc gcgttacata 120

acttacggta aatggcccgc ctggctgacc gcccaacgac ccccgcccat tgacgtcaat 180

aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc aatgggtgga 240

ctatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc caagtacgcc 300

ccctattgac gtcaatgacg gtaaatggcc cgcctggcat tatgcccagt acatgacctt 360

atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta ccatggtgat 420

gcggttttgg cagtacatca atgggcgtgg atagcggttt gactcacggg gatttccaag 480

tctccacccc attgacgtca atgggagttt gttttggcac caaaatcaac gggactttcc 540

aaaatgtcgt aacaactccg ccccattgac gcaaatgggc ggtaggcgtg tacggtggga 600

ggtctatata agcagagctc tctggctaac tagagaaccc actgcttact ggcttatcga 660

aattaatacg actcactata gggagaccca agctggctag cgtttaaact taagcttggt 720

accgagctcg gatccactag tccagtgtgg tggaattcgc caccatggat tggacatgga 780

ttctgtttct ggtcgccgcc gcaacccgcg tgcactcagg gggatggtca tcaaaaccta 840

gacagggaat gggcactaac ctgagtgtgc ccaatcctct ggggttcttt cccgaccacc 900

agctggatcc tgccttcggc gctaactcta acaatccaga ctgggacttc aaccccaata 960

aggaccactg gcctgaggca aatcaagtgg gagcaggagc cttcggacca ggctttacac 1020

cccctcatgg cggactgctg ggatggtccc ctcaggctca ggggatcctg accacagtcc 1080

cagcagctcc acccccctgca agtactaaca ggcagtcagg acgccagcca acccccattt 1140

ctccacccct gagggacagt caccctcagg ccagaggcag gaagcggaga agcatgcagt 1200

ggaacagcac taccttccat caggcactgc tggatccacg cgtgcgagga ctgtactttc 1260

cagccggagg cagctcctct ggaaccgtga accctgtccc aacaactgcc tccccaatca 1320

gttcaatttt ctctcggaca ggagaccccg ctcctaatcg gggcagaaaa aggcgctcaa 1380

tggaaagcac cacatccggg tttctgggac cactgctggt gctgcaggca ggcttctttc 1440

tgctgaccag aatcctgaca attccccagt ctctggatag ttggtggacc agcctgaact 1500

tcctgggcgg cgcccctact tgtccaggac agaactctca gagtccaaca tcaaatcata 1560

gccccacttc ctgtcctcca atctgccctg gctaccgctg gatgtgcctg cgacggttca 1620

tcattttcct gtttatcctg ctgctgtgcc tgatttttct gctggtgctg ctggactatc 1680

agggaatgct gcccgtctgc cctctgctgc cagggacttc tactaccagt accggacctt 1740

gtaagacatg cactattcca gctcagggga cctccatgtt cccctcttgc tgttgcacca 1800

aacctagcga tggaaattgt acatgcatcc caattcccag ctcctgggct ttcgcacgat 1860

ttctgtggga gtgggccagc gtgcgctttt catggctgag cctgctggtg cccttcgtcc 1920

agtggtttgt cggcctgtca cctaccgtgt ggctgagcgt catctggatg atgtggtact 1980

gggggcccag cctgtataat atcctgtcac cattcctgcc actgctgcca atcttctttt 2040

gtctgtgggt ctacatttga taagcggccg ctcgagtcta gagggcccgt ttaaacccgc 2100

tgatcagcct cgactgtgcc ttctagttgc cagccatctg ttgtttgccc ctcccccgtg 2160

ccttccttga ccctggaagg tgccactccc actgtccttt cctaataaaa tgaggaaatt 2220

gcatcgcatt gtctgagtag gtgtcattct attctggggg gtggggtggg gcaggacagc 2280

aagggggagg attgggaaga caatagcagg catgctgggg atgcggtggg ctctatggct 2340

tctactgggc ggttttatgg acagcaagcg aaccggaatt gccagctggg gcgccctctg 2400

gtaaggttgg gaagccctgc aaagtaaact ggatggcttt ctcgccgcca aggatctgat 2460

ggcgcagggg atcaagctct gatcaagaga caggatgagg atcgtttcgc atgattgaac 2520

aagatggatt gcacgcaggt tctccggccg cttgggtgga gaggctattc ggctatgact 2580

gggcacaaca gacaatcggc tgctctgatg ccgccgtgtt ccggctgtca gcgcaggggc 2640

gcccggttct ttttgtcaag accgacctgt ccggtgccct gaatgaactg caagacgagg 2700

cagcgcggct atcgtggctg gccacgacgg gcgttccttg cgcagctgtg ctcgacgttg 2760

tcactgaagc gggaagggac tggctgctat tgggcgaagt gccggggcag gatctcctgt 2820

catctcacct tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg cggcggctgc 2880

atacgcttga tccggctacc tgcccattcg accaccaagc gaaacatcgc atcgagcgag 2940

cacgtactcg gatggaagcc ggtcttgtcg atcaggatga tctggacgaa gagcatcagg 3000

ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgag catgcccgac ggcgaggatc 3060

tcgtcgtgac ccatggcgat gcctgcttgc cgaatatcat ggtggaaaat ggccgctttt 3120

ctggattcat cgactgtggc cggctgggtg tggcggaccg ctatcaggac atagcgttgg 3180

ctacccgtga tattgctgaa gagcttggcg gcgaatgggc tgaccgcttc ctcgtgcttt 3240

acggtatcgc cgctcccgat tcgcagcgca tcgccttcta tcgccttctt gacgagttct 3300

tctgaattat taacgcttac aatttcctga tgcggtattt tctccttacg catctgtgcg 3360

gtatttcaca ccgcatacag gtggcacttt tcggggaaat gtgcgcggaa cccctatttg 3420

tttatttttc taaatacatt caaatatgta tccgctcatg agacaataac cctgataaat 3480

gcttcaataa tagcacgtgc taaaacttca tttttaattt aaaaggatct aggtgaagat 3540

ccttttttgat aatctcatga ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc 3600

agaccccgta gaaaagatca aaggatcttc ttgagatcct ttttttctgc gcgtaatctg 3660

ctgcttgcaa acaaaaaaac caccgctacc agcggtggtt tgtttgccgg atcaagagct 3720

accaactctt tttccgaagg taactggctt cagcagagcg cagataccaa atactgtcct 3780

tctagtgtag ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct 3840

cgctctgcta atcctgttac cagtggctgc tgccagtggc gataagtcgt gtcttaccgg 3900

gttggactca agacgatagt taccggataa ggcgcagcgg tcgggctgaa cgggggttc 3960

gtgcacacag cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga 4020

gctatgagaa agcgccacgc ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg 4080

cagggtcgga acaggagagc gcacgaggga gcttccaggg ggaaacgcct ggtatcttta 4140

tagtcctgtc gggtttcgcc acctctgact tgagcgtcga tttttgtgat gctcgtcagg 4200

ggggcggagc ctatggaaaa acgccagcaa cgcggccttt ttacggttcc tgggcttttg 4260

ctggcctttt gctcacatgt tctt 4284

<210> 20
<211> 3690
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic (pGX1804 HepB pSHBs-A)

<400> 20
gactcttcgc gatgtacggg ccagatatac gcgttgacat tgattattga ctagttatta 60

atagtaatca attacggggt cattagttca tagcccatat atggagttcc gcgttacata 120

acttacggta aatggcccgc ctggctgacc gcccaacgac ccccgcccat tgacgtcaat 180

aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc aatgggtgga 240

ctatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc caagtacgcc 300

ccctattgac gtcaatgacg gtaaatggcc cgcctggcat tatgcccagt acatgacctt 360

atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta ccatggtgat 420

gcggttttgg cagtacatca atgggcgtgg atagcggttt gactcacggg gatttccaag 480

tctccacccc attgacgtca atgggagttt gttttggcac caaaatcaac gggactttcc 540

aaaatgtcgt aacaactccg ccccattgac gcaaatgggc ggtaggcgtg tacggtggga 600

ggtctatata agcagagctc tctggctaac tagagaaccc actgcttact ggcttatcga 660

aattaatacg actcactata gggagaccca agctggctag cgtttaaact taagcttggt 720

accgagctcg gatccgccac catggactgg acctggattc tgttcctggt ggctgccgct 780

acacgggtgc attctgaaaa tatcacatct ggattcctgg gacctctgct ggtgctgcag 840

gctgggttct ttctgctgac aagaatcctg actattcccc agtcactgga cagctggtgg 900

acatctctga acttcctggg cgggagtcct gtctgtctgg gacagaactc tcagagtcct 960

acttccaatc actctccaac cagttgtccc cctatctgcc caggctaccg ctggatgtgc 1020

ctgcggagat tcatcatttt cctgtttatc ctgctgctgt gcctgatttt tctgctggtg 1080

ctgctggact atcagggcat gctgcctgtc tgcccactga ttcccggcag caccacaact 1140

tctaccggcc cctgtaagac atgcaccaca cctgcccagg ggaacagtat gtttccatca 1200

tgctgttgca ctaaacccac cgatggaaat tgtacatgca tcccaattcc cagctcctgg 1260

gccttcgcta agtacctgtg ggagtgggca agcgtgcgat tttcatggct gagcctgctg 1320

gtgcctttcg tccagtggtt tgtgggcctg agcccaactg tctggctgtc cgccatctgg 1380

atgatgtggt actgggggcc cagcctgtat tccatcgtgt caccattcat tcccctgctg 1440

ccaatctttt tctgcctgtg ggtctacatc tgataactcg agtctagagg gcccgtttaa 1500

acccgctgat cagcctcgac tgtgccttct agttgccagc catctgttgt ttgcccctcc 1560

cccgtgcctt ccttgaccct ggaaggtgcc actcccactg tcctttccta ataaaatgag 1620

gaaattgcat cgcattgtct gagtaggtgt cattctattc tgggggggtgg ggtggggcag 1680

gacagcaagg gggaggattg ggaagacaat agcaggcatg ctggggatgc ggtgggctct 1740

atggcttcta ctgggcggtt ttatggacag caagcgaacc ggaattgcca gctggggcgc 1800

cctctggtaa ggttgggaag ccctgcaaag taaactggat ggctttctcg ccgccaagga 1860

tctgatggcg caggggatca agctctgatc aagagacagg atgaggatcg tttcgcatga 1920

ttgaacaaga tggattgcac gcaggttctc cggccgcttg ggtggagagg ctattcggct 1980

atgactgggc acaacagaca atcggctgct ctgatgccgc cgtgttccgg ctgtcagcgc 2040

aggggcgccc ggttcttttt gtcaagaccg acctgtccgg tgccctgaat gaactgcaag 2100

acgaggcagc gcggctatcg tggctggcca cgacgggcgt tccttgcgca gctgtgctcg 2160

acgttgtcac tgaagcggga agggactggc tgctattggg cgaagtgccg gggcaggatc 2220

tcctgtcatc tcaccttgct cctgccgaga aagtatccat catggctgat gcaatgcggc 2280

ggctgcatac gcttgatccg gctacctgcc cattcgacca ccaagcgaaa catcgcatcg 2340

agcgagcacg tactcggatg gaagccggtc ttgtcgatca ggatgatctg gacgaagagc 2400

atcaggggct cgcgccagcc gaactgttcg ccaggctcaa ggcgagcatg cccgacggcg 2460

aggatctcgt cgtgacccat ggcgatgcct gcttgccgaa tatcatggtg gaaaatggcc 2520

gcttttctgg attcatcgac tgtggccggc tgggtgtggc ggaccgctat caggacatag 2580

cgttggctac ccgtgatatt gctgaagagc ttggcggcga atgggctgac cgcttcctcg 2640

tgctttacgg tatcgccgct cccgattcgc agcgcatcgc cttctatcgc cttcttgacg 2700

agttcttctg aattattaac gcttacaatt tcctgatgcg gtattttctc cttacgcatc 2760

tgtgcggtat ttcacaccgc atacaggtgg cacttttcgg ggaaatgtgc gcggaacccc 2820

tatttgttta tttttctaaa tacattcaaa tatgtatccg ctcatgagac aataaccctg 2880

ataaatgctt caataatagc acgtgctaaa acttcatttt taatttaaaa ggatctaggt 2940

gaagatcctt tttgataatc tcatgaccaa aatcccttaa cgtgagtttt cgttccactg 3000

agcgtcagac cccgtagaaa agatcaaagg atcttcttga gatccttttt ttctgcgcgt 3060

aatctgctgc ttgcaaacaa aaaaaccacc gctaccagcg gtggtttgtt tgccggatca 3120

agagctacca actcttttc cgaaggtaac tggcttcagc agagcgcaga taccaaatac 3180

tgtccttcta gtgtagccgt agttaggcca ccacttcaag aactctgtag caccgcctac 3240

atacctcgct ctgctaatcc tgttaccagt ggctgctgcc agtggcgata agtcgtgtct 3300

taccgggttg gactcaagac gatagttacc ggataaggcg cagcggtcgg gctgaacggg 3360

gggttcgtgc acacagccca gcttggagcg aacgacctac accgaactga gatacctaca 3420

gcgtgagcta tgagaaagcg ccacgcttcc cgaagggaga aaggcggaca ggtatccggt 3480

aagcggcagg gtcggaacag gagagcgcac gagggagctt ccagggggaa acgcctggta 3540

tctttatagt cctgtcgggt ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc 3600

gtcaggggg cggagcctat ggaaaaacgc cagcaacgcg gccttttac ggttcctggg 3660

cttttgctgg ccttttgctc acatgttctt 3690

<210> 21
<211> 3720
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic (pGX1805 HepB SHBs-C)

<400> 21
gactcttcgc gatgtacggg ccagatatac gcgttgacat tgattattga ctagttatta 60

atagtaatca attacggggt cattagttca tagcccatat atggagttcc gcgttacata 120

acttacggta aatggcccgc ctggctgacc gcccaacgac ccccgcccat tgacgtcaat 180

aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc aatgggtgga 240

ctatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc caagtacgcc 300

ccctattgac gtcaatgacg gtaaatggcc cgcctggcat tatgcccagt acatgacctt 360

atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta ccatggtgat 420

gcggttttgg cagtacatca atgggcgtgg atagcggttt gactcacggg gatttccaag 480

tctccacccc attgacgtca atgggagttt gttttggcac caaaatcaac gggactttcc 540

aaaatgtcgt aacaactccg ccccattgac gcaaatgggc ggtaggcgtg tacggtggga 600

ggtctatata agcagagctc tctggctaac tagagaaccc actgcttact ggcttatcga 660

aattaatacg actcactata gggagaccca agctggctag cgtttaaact taagcttggt 720

accgagctcg gatccactag tccagtgtgg tggaattcgc caccatggat tggacatgga 780

ttctgttcct ggtcgccgcc gcaacacgag tgcattctga aagtacaacc tctggcttcc 840

tgggcccccc gctggtgctg caggcagggt tctttctgct gacacgaatc ctgactattc 900

cacagtcact ggacagctgg tggaccagcc tgaacttcct gggcggggcc cctacatgtc 960

caggacagaa ctctcagagt cccacttcca atcactctcc taccagttgt ccccctatct 1020

gccccggcta cagatggatg tgcctgcgga gattcatcat tttcctgttt atcctgctgc 1080

tgtgcctgat ttttctgctg gtgctgctgg actatcaggg aatgctgccc gtctgccctc 1140

tgctgccagg aacctcaacc acaagcacag gcccttgtaa gacttgcacc attcccgctc 1200

aggggactag tatgttccct tcatgctgtt gcacaaaacc atctgatgga aattgtactt 1260

gcatcccaat tcccagctcc tgggccttcg ctaggtttct gtgggagtgg gccagtgtgc 1320

gcttttcctg gctgtctctg ctggtgccct tcgtccagtg gtttgtcggc ctgagcccta 1380

cagtgtggct gtccgtcatc tggatgatgt ggtactgggg gcctagcctg tataatatcc 1440

tgtcaccatt tctgccactg ctgccaatct tcttttgcct gtgggtctac atctgataag 1500

cggccgctcg agtctagagg gcccgtttaa acccgctgat cagcctcgac tgtgccttct 1560

agttgccagc catctgttgt ttgcccctcc cccgtgcctt ccttgaccct ggaaggtgcc 1620

actcccactg tcctttccta ataaaatgag gaaattgcat cgcattgtct gagtaggtgt 1680

cattctattc tgggggggtgg ggtggggcag gacagcaagg gggaggattg ggaagacaat 1740

agcaggcatg ctggggatgc ggtgggctct atggcttcta ctgggcggtt ttatggacag 1800

caagcgaacc ggaattgcca gctggggcgc cctctggtaa ggttgggaag ccctgcaaag 1860

taaactggat ggctttctcg ccgccaagga tctgatggcg caggggatca agctctgatc 1920

aagagacagg atgaggatcg tttcgcatga ttgaacaaga tggattgcac gcaggttctc 1980

cggccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca atcggctgct 2040

ctgatgccgc cgtgttccgg ctgtcagcgc aggggcgccc ggttcttttt gtcaagaccg 2100

acctgtccgg tgccctgaat gaactgcaag acgaggcagc gcggctatcg tggctggcca 2160

cgacgggcgt tccttgcgca gctgtgctcg acgttgtcac tgaagcggga agggactggc 2220

tgctattggg cgaagtgccg gggcaggatc tcctgtcatc tcaccttgct cctgccgaga 2280

aagtatccat catggctgat gcaatgcggc ggctgcatac gcttgatccg gctacctgcc 2340

cattcgacca ccaagcgaaa catcgcatcg agcgagcacg tactcggatg gaagccggtc 2400

ttgtcgatca ggatgatctg gacgaagagc atcaggggct cgcgccagcc gaactgttcg 2460

ccaggctcaa ggcgagcatg cccgacggcg aggatctcgt cgtgacccat ggcgatgcct 2520

gcttgccgaa tatcatggtg gaaaatggcc gcttttctgg attcatcgac tgtggccggc 2580

tgggtgtggc ggaccgctat caggacatag cgttggctac ccgtgatatt gctgaagagc 2640

ttggcggcga atgggctgac cgcttcctcg tgctttacgg tatcgccgct cccgattcgc 2700

agcgcatcgc cttctatcgc cttcttgacg agttcttctg aattattaac gcttacaatt 2760

tcctgatgcg gtattttctc cttacgcatc tgtgcggtat ttcacaccgc atacaggtgg 2820

cacttttcgg ggaaatgtgc gcggaacccc tatttgttta tttttctaaa tacattcaaa 2880

tatgtatccg ctcatgagac aataaccctg ataaatgctt caataatagc acgtgctaaa 2940

acttcatttt taatttaaaa ggatctaggt gaagatcctt tttgataatc tcatgaccaa 3000

aatcccttaa cgtgagtttt cgttccactg agcgtcagac cccgtagaaa agatcaaagg 3060

atcttcttga gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa aaaaaccacc 3120

gctaccagcg gtggtttgtt tgccggatca agagctacca actcttttc cgaaggtaac 3180

tggcttcagc agagcgcaga taccaaatac tgtccttcta gtgtagccgt agttaggcca 3240

ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaatcc tgttaccagt 3300

ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac gatagttacc 3360

ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca gcttggagcg 3420

aacgacctac accgaactga gatacctaca gcgtgagcta tgagaaagcg ccacgcttcc 3480

cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag gagagcgcac 3540

gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt ttcgccacct 3600

ctgacttgag cgtcgatttt tgtgatgctc gtcaggggggg cggagcctat ggaaaaacgc 3660

cagcaacgcg gccttttac ggttcctggg cttttgctgg ccttttgctc acatgttctt 3720