

SEQUENCE LISTING

<110> National University of Singapore

<120> ENGINEERED DYSBIOSIS-SENSING PROBIOTIC FOR CLOSTRIDIUM DIFFICILE
INFECTIONS AND RECURRING INFECTIONS MANAGEMENT

<130> SP100743W0

<150> SG10201902947W

<151> 2019-04-02

<160> 14

<170> PatentIn version 3.5

<210> 1

<211> 3641

<212> DNA

<213> Artificial Sequence

<220>

<223> pEaat plasmid

<400> 1

aagcttaagt gatcattggc gcgccgtgta ggctggagct gcttcgaagt tcctatactt	60
tctagagaat aggaacttcg gaataggaac ttcaagatcc ccttattaga agaactcgtc	120
aagaaggcga tagaaggcga tgcgctgcga atcgggagcg gcgataccgt aaagcacgag	180
gaagcggta gccatttcg cgccaagctc ttacgaata tcacgggtag ccaacgctat	240
gtcctgatag cgggtccgca caccagccg gccacagtcg atgaatccag aaaagcggcc	300
atcttccacc atgatattcg gcaagcaggc atcgccatgg gtcacgacga gatcctcgcc	360
gtcgggcatg cgcgccttga gcctggcgaa cagttcggct ggcgcgagcc cctgatgctc	420
ttcgtccaga tcacctgat cgacaagacc ggcttccatc cgagtacgtg ctgctcgat	480
gcgatgtttc gcttggtggt cgaatgggca ggtagccgga tcaagcgat gcagccgccg	540
cattgcatca gccatgatgg atactttctc ggaggagca aggtgagatg acaggagatc	600
ctgccccggc acttcgcca atagcagcca gtcccttccc gcttcagtga caacgtcgag	660
cacagctgcg caaggaacgc ccgtcgtggc cagccacgat agccgcgctg cctcgtcctg	720
cagttcattc agggcaccgg acaggtcggc cttgacaaaa agaaccgggc gcccctgcgc	780

tgacagccgg aacacggcgg catcagagca gccgattgtc tgttgtgccc agtcatagcc	840
gaatagcctc tccaccaag cggccggaga acctgcgtgc aatccatctt gttcaatcat	900
gcgaaacgat cctcatcctg tctcttgttc agatcatgat cccctgcgcc atcagatcct	960
tggcggcaag aaagccatcc agtttacttt gcagggcttc ccaaccttac cagagggcgc	1020
cccagctggc aattccggtt cgcttgctgt ccataaaacc gcccagtcta gctatcgcca	1080
tgtaagccca ctgcaagcta cctgctttct ctttgcgctt gcgttttccc ttgtccagat	1140
agcccagtag ctgacattca tccggggctca gcaccgtttc tgcggactgg ctttctacgt	1200
gttccgcttc ctttagcagc ctttgcgccc tgagtgcctt cggcagcgtg agcttcaaaa	1260
gcgctctgaa gttcctatac tttctagaga ataggaactt cgaactgcag gtcgacggat	1320
ctccggaata cgcgtttcga attcaagaga tctaaaggat cctaactcga gtaaggatct	1380
ccaggcatca aataaaacga aaggctcagt cgaaagactg ggcctttcgt tttatctgtt	1440
gtttgtcggg gaacgctctc tactagagtc acactggctc accttcgggt gggcctttct	1500
gcgtttatac ctagggcggt cggtgcggc gagcggatc agctcactca aaggcggtaa	1560
tacggttatc cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc	1620
aaaaggccag gaaccgtaaa aaggccgctg tgctggcggt tttccatagg ctccgcccc	1680
ctgacgagca tcacaaaaat cgacgctcaa gtcagagggt gcgaaacccg acaggactat	1740
aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc	1800
cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcatagct	1860
cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg	1920
aacccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc	1980
cggtaaagaca cgacttatcg ccaactggcag cagccactgg taacaggatt agcagagcga	2040
ggtatgtagg cgggtgtaca gagttcttga agtggtggcc taactacggc tacactagaa	2100
ggacagtatt tggtatctgc gctctgtga agccagttac cttcggaaaa agagttggta	2160
gctcttgatc cggcaaacia accaccgctg gtagcgggtg tttttttgtt tgcaagcagc	2220
agattacgcg cagaaaaaaaa ggatctcaag aagatccttt gatcttttct acgggggtctg	2280

acgctcagtg gaacgaaaac tcacgttaag ggatttttgg	catgactagt gcttggattc	2340
tcaccaataa aaaacgccc	gcggaaccg agcgttctga acaaatccag atggagtct	2400
gaggtcatta ctggatctat caacaggagt ccaagcgagc	tctcgaaccc cagagtgata	2460
tcttaatcca cgtatttcat cgcgactctt gaagtcaggc	gcgtaataag ttcgtaagcg	2520
cttactttcg tcatttcagc gatacgttct acgggcaaac	cttcgcccc aaaaatgacc	2580
gggtccccgg ctttgtcctg cgctgtgga cctaagtcta	cgcagatcat atccatcgcg	2640
actcgcccga caatcggcac ttcgcgaccg ttcaccagca	ctggcggtacc ggacggcgcg	2700
gcgcgcggat aaccatcgcc atagcccac gcgactacgc	caagacgagt atcacgttcg	2760
cttaccaggg ttccaccata accgacaggc tctccggctt	tatgctcacg cacggcaatc	2820
aggctggagg ttaaagacat gactggctga cagccaaaat	cagcccctgt cgtgccatct	2880
tccagcggcg agacaccgta aagaatgatc cccggacgcg	cccagtcaaa atgcgactgc	2940
ggccataaca aaatgccgcc tgatgccgca attgagcggt	gccccggtt accttcacaa	3000
aaggtgttga aaatatcgag ctgcttttca gtcgcgccgc	tttgcggttc atcggcacgg	3060
gcgaagtgac taacaatgtt caccggctgg cggacatttt	tacactggct cagacgtga	3120
taaaacgcct cggcctgttc cggcaatacg cccaaacgg	gcattccggt atcgagcttc	3180
atccagacgg tgacaggctc tttaagttca gcgttttcga	ggcgacaag ctgctcttca	3240
ttgtggactg cggatatgcag atgttcagcg gagatcgtcg	gcaaactgctc tgcttcaaaa	3300
aaaccttcca gtaacaaaat aggtcgcgtg ataccccccg	cccgcagccg tagggcttct	3360
tcgagacggg caacgcaaaa ggcgtcagca tcggggagcg	ttcgcgcggt ctcaatcaga	3420
ccgtgaccgt aggcgttcgc tttcaccacc gcaaccagtt	tactggcggg ggccagttca	3480
cgcagacgtt gcaggttggt tcgcagagcg cggcgggttaa	tcaaaacagt tgccgcttgc	3540
atttgtattc ctttttttca ggttctgccc accagtgcaa	aacctcgcta aacagatatg	3600
accggagtat gctattccac atccagggat gggtttataa	a	3641

<210> 2

<211> 99

<212> DNA
<213> Artificial Sequence

<220>
<223> alr promoter

<400> 2
tttataaacc catccctgga tgtggaatag cataactccgg tcatatctgt ttagcgaggt 60
tttgacttgg tgggcagaac ctgaaaaaaaa ggaatacaa 99

<210> 3
<211> 1080
<212> DNA
<213> Artificial Sequence

<220>
<223> alr ORF

<400> 3
atgcaagcgg caactgtttt gattaaccgc cgcgctctgc gacacaacct gcaacgtctg 60
cgtgaactgg ccccgccag taaactgggt gcggtgggtga aagcgaacgc ctacggtcac 120
ggctctgattg agaccgcgcg aacgctcccc gatgctgacg cctttggcgt tgcccgtctc 180
gaagaagccc tacggctgcg ggcgggggggt atcacgcgac ctattttgtt actggaaggt 240
ttttttgaag cagacgattt gccgacgatc tccgctgaac atctgcatac cgaggtccac 300
aatgaagagc agcttgctgc cctcgaaaac gctgaactta aagagcctgt caccgtctgg 360
atgaagctcg ataccggaat gcaccgtttg ggcgtattgc cggaacaggc cgaggcgttt 420
tatcagcgtc tgagccagtg taaaaatgtc cgccagccgg tgaacattgt tagtcacttc 480
gcccgtgccg atgaaccgca aagcggcgcg actgaaaagc agctcgatat tttcaacacc 540
ttttgtgaag gtaaaccggg gcaacgctca attgcggcat caggcggcat tttgttatgg 600
ccgcagtcgc attttgactg ggcgcgtccg gggatcattc tttacggtgt ctgcccgtg 660
gaagatggca cgacaggggc tgattttggc tgtcagccag tcatgtcttt aacctccagc 720
ctgattgccg tgcgtgagca taaagccgga gagcctgtcg gttatgggtg aacctgggta 780
agcgaacgtg atactcgtct tggcgtagtc gcgatgggct atggcgatgg ttatccgcgc 840
gccgcgccgt ccggtacgcc agtgctgggtg aacggtcgcg aagtgccgat tgtcgggcga 900

gtcgcgatgg atatgatctg cgtagactta ggtccacagg cgcaggacaa agccggggac	960
ccggtcattt tatggggcga aggtttgccc gtagaacgta tcgctgaaat gacgaaagta	1020
agcgcttacg aacttattac gcgcttgact tcaagagtcg cgatgaaata cgtggattaa	1080

<210> 4
 <211> 85
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> pNanA promoter

<400> 4	
agatcgcatt ataagctttc tgtatggggg gttgcttaat tgatctggta taacaggtat	60
aaaggtatat cgtttatcag acaag	85

<210> 5
 <211> 792
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> NanR ORF

<400> 5	
atgggcctta tgaacgcatt tgattcgcaa accgaagatt cttcacctgc aattggctgc	60
aacttgcgta gccgcccgtt ggcgcgtaaa aaactctccg aaatgggtgga agaagagctg	120
gaacagatga tccgccgtcg tgaatttggc gaaggatgaac aattaccgtc tgaacgcgaa	180
ctgatggcgt tctttaacgt cgggcgtcct tcggtgcgtg aagcgctggc agcgtaaaaa	240
cgcaaaggctc tgggtgcaaata aaacaacggc gaacgcgctc gcgtctcgcg tccttctgcg	300
gacactatca tcggtgagct ttccggcatg gcgaaagatt tcctttctca tcccgggtggg	360
attgcccatt tcgaacaatt acgtctgttc tttgaatcca gtctgggtgcg ctatgcggct	420
gaacatgccca ccgatgagca aatcgatttg ctggcaaaaag cactgggaaat caacagtcag	480
tcgctggata acaacgcggc attcattcgt tcagacgttg atttccaccg cgtgctggcg	540
gagatccccg gtaacccaat cttcatggcg atccacgttg ccctgctcga ctggcttatt	600

gccgcacgcc caacggttac cgatcaggca ctgcacgaac ataacaacgt tagttatcaa	660
cagcatattg cgatcgttga tgcgatccgc cgtcatgatc ctgacgaagc cgatcgtgcg	720
ttgcaatcgc atctcaacag cgtctctgct acctggcacg ctttcgggtca gaccaccaac	780
aaaaagaaat aa	792

<210> 6
 <211> 1094
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> J23113-rbs4-NanR

<400> 6	
tataaacgca gaaaggccca cccgaagggtg agccagtgtg actctagtag agagcgttca	60
ccgacaaaca acagataaaa cgaaaggccc agtctttcga ctgagccttt cgttttat	120
gatgcctgga gatccttatt tctttttggt ggtgggtctga ccgaaagcgt gccaggtagc	180
agagacgctg ttgagatgcg attgcaacgc acgatcggct tcgtcaggat catgacggcg	240
gatcgcatca acgatcgcaa tatgctgttg ataactaacg ttgttatggt cgtgcagtgc	300
ctgatcggta accgttgggc gtgcggcaat aagccagtcg agcagggcaa cgtggatcgc	360
catgaagatt gggttaccgg ggatctccgc cagcacgcgg tggaaatcaa cgtctgaacg	420
aatgaatgcc gcgttggttat ccagcgactg actgttgatt tccagtgtt ttgccagcaa	480
atcgatttgc tcatcgggtg catgttcagc cgcatagcgc accagactgg attcaaagaa	540
cagacgtaat tgttcgaaat gggcaatccc accgggatga gaaaggaaat ctttcgcat	600
gccggaaagc tcaccgatga tagtgtccgc agaaggacgc gagacgcgag cgcgttcgcc	660
gttgtttatt tgcaccagac ctttgcgttt taacgctgcc agcgcttcac gcaccgaagg	720
acgcccgcg ttaaagaacg ccatcagttc gcgttcagac ggtaattgtt caccttcgcc	780
aaattcacga cggcggatca tctgttccag ctcttcttcc accatttcgg agagtttttt	840
acgcgccagc gggcggctac gcaagttgcg accaattgca ggtgaagaat cttcggtttg	900
cgaatcaaat gcgttcataa ggcccataga tccgtcctgt gtgaagatcc gctagcataa	960

tccctaggac tgagctagcc atcagggatc tagatcgcat tataagcttt ctgtatgggg	1020
tgttgcttaa ttgatctggg ataacaggta taaaggata tcgtttatca gacaagggat	1080
ctaaagagga gaaa	1094

<210> 7
 <211> 1923
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> cadC-pCadBA amplifier

<400> 7	
atgcaacaac ctgtagttcg cgttggcgaa tggcttggtta ctccgtccat aaaccaaatt	60
agccgcaatg ggcgtcaact tacccttgag ccgagattaa tcgatcttct ggttttcttt	120
gctcaacaca gtggcgaagt acttagcagg gatgaactta tcgataatgt ctggaagaga	180
agtattgtca ccaatcacgt tgtgacgcag agtatctcag aactacgtaa gtcattaaaa	240
gataatgatg aagatagtcc tgtctatatc gctactgtac caaagcgcgg ctataaatta	300
atggtgccgg ttatctggta cagcgaagaa gagggagagg aaataatgct atcttcgcct	360
ccccctatac cagaggcggg tcctgccaca gattctccct cccacagtct taacattcaa	420
aacaccgcaa cgccacctga acaatcccca gttaaaagca aacgattcac taccttttgg	480
gtatggtttt ttttcctggt gtcgttaggt atctgtgtag cactggtagc gttttcaagt	540
cttgatacac gtcttcctat gagcaaatcg cgtattttgc tcaatccacg cgatattgac	600
attaatatgg taaataaaag ttgtaacagc tggagttccc cgtatcagct ctcttacgcg	660
ataggcgtgg gtgatttggt ggcgacatca cttaacacct tctccacctt tatggtgcat	720
gacaaaatca actacaacat tgatgaaccg agcagttccg gtaaaacatt atctattgcg	780
tttgttaatc agtgccaata ccgtgctcaa caatgcttta tgtcgataaa attggtagac	840
aatgcagatg gttcaaccat gctggataaa cgttatgtca tcactaacgg taatcagctg	900
gcgattcaaa atgatttact ggagagttaa tcaaaagcgt taaaccaacc gtggccacaa	960
cgaatgcagg agacgctcca gaaaattttg ccgcatcggt gtgcgttatt aactaatttt	1020

tatcaggcac atgattat	1080
ttattaggtg agattgttca atcatcccca gaatttacct acgcgagagc agaaaaagca	1140
ttagttgata tcgtgcgcca ttctcaacat ccttttagatg aaaaacaatt agcagcactg	1200
aacacagaaa tagataacat tgttacactg ccggaattga acaacctgtc cattatatat	1260
caaataaaag cggtcagtgc tctggtaaaa ggtaaaacag atgagtctta ccaggcgata	1320
aatactggca ttgatcttga aatgtcctgg ctaaattatg tgttgcttgg caaggtttat	1380
gaaatgaagg ggatgaaccg ggaagcagct gatgcatatc tcaccgcctt taatttacgc	1440
ccaggggcaa acacccttta ctggattgaa aatggatatat tccagacttc tgttccttat	1500
gttgtacctt atctcgacaa atttcttgct tcagaataag gatctagact tctgttcctt	1560
atgttgtacc ttatctcgac aaatttcttg cttcagaata agtaactccg ggttgattta	1620
tgctcggaaa tatttgttgt tgagtttttg tatgttcctg ttggtataat atgttgcggc	1680
aatttatttg ccgcataatt tttattacat aaatttaacc agagaatgtc acgcaatcca	1740
ttgtaaacat taaatgttta tcttttcatg atatcaactt gcgatcctga tgtgttaata	1800
aaaaacctca agttctcact tacagaaact tttgtgttat ttcacctaat ctttaggatt	1860
aatccttttt tcgtgagtaa tcttatcgcc agtttgggtc ggtcaggaaa taaagaggag	1920
aaa	1923

<210> 8
 <211> 990
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> cbh ORF native

<400> 8	
atgtgtacag gattagcctt agaaacaaaa gatggattac atttgtttgg aagaaatatg	60
gatattgaat attcatttaa tcaatctatt atatttattc ctaggaattt taaatgtgta	120
aacaaatcaa acaaaaaaga attaacaaca aaatatgctg ttcttggaat gggaactatt	180
tttgatgatt atcctacctt tgcagatggg atgaatgaaa agggattagg gtgtgctggc	240

ttaaatttcc ctgtttatgt tagctattct aaagaagata tagaaggtaa aactaatatt	300
ccagtatata atttcttatt atgggtttta gctaatttta gctcagtaga agaggtaaag	360
gaagcattaa aaaatgctaa tatagtggat atacctatta gcgaaaatat tcctaataca	420
actcttcatt ggatgataag cgatataaca ggaaagtcta ttgtggttga acaaacaag	480
gaaaaattaa atgtatttga taataatatt ggagtattaa ctaattcacc tacttttgat	540
tggcatgtag caaatthaaa tcaatatgta ggtttgagat ataatcaagt tccagaattt	600
aagttaggag atcaatcttt aactgcttta ggtcaaggaa ctggtttagt aggattacca	660
ggggacttta cacctgcac tagatttata agagtagcat ttttaagaga tgcaatgata	720
aaaaatgata aagattcaat agacttaatt gaatttttcc atatattaaa taatgttgct	780
atggtaagag gatcaactag aactgtagaa gaaaaaagtg atcttactca atatacaagt	840
tgcatgtgtt tagaaaaagg aatttattat tataataacct atgaaaataa tcaaattaat	900
gcaatagaca tgaataaaga aaacttagat ggaaatgaaa ttaaaacata taaatacaac	960
aaaactttta gtattaatca tgtaaattag	990

<210> 9
 <211> 990
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> cbh ORF codon optimised

<400> 9	
atgtgtaccg gtttggcatt ggagaccaag gatggctctcc acttatttgg tcgcaatatg	60
gatattgagt atagctttta ccaatcgatt atttttatcc cgcgcaactt taaatgcgta	120
aataaatcta ataagaaaga actgactacg aaatatgcgg tcctcggatg ggggacgatt	180
ttcgaatgatt atccacggt tgcagacggc atgaacgaaa agggctctggg gtgtgcgggt	240
cttaattttc ctgtgtacgt cagttatagt aaggaagaca tcgagggaaa aaccaatatt	300
ccggtatata acttcttgct gtgggttctg gcaaatttta gctcagtcga agaagtgaag	360
gaagcgttta aaaatgccaa tatcgtggat attccgatta gcgaaaacat tccgaatact	420

```

acgttgcact ggatgatctc ggacattact ggcaaaagca ttgtggtaga acagactaaa      480
gaaaaactga atgtcttcga caacaatatc ggggttttaa ccaattctcc gacttttgac      540
tggcatgtag ctaacttgaa tcagtatgtg ggactgcgtt ataaccaagt cccggagttc      600
aaactgggcg accagtcttt aaccgcgctg ggccagggca ccggcctggg ggggctgccg      660
ggcgacttca cccctgcgtc acgcttcatt cgcgtagcat tccttcgcga tgcgatgatt      720
aaaaatgaca aagacagcat tgacctgac gagttctttc atattttaaa taatgtggct      780
atggtacggg gctctacgcg cactgtggaa gaaaagagcg acttgacca gtatacctca      840
tgcattgtgcc tggaaaaagg catttactac tacaatactt atgaaaataa tcagatcaat      900
gccatcgata tgaacaaaga gaacctggac ggtaatgaaa ttaaaaccta taaatacaat      960
aaaacgctgt cgatcaatca tgtcaactaa      990

```

```

<210> 10
<211> 359
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Alr

```

```

<400> 10

```

```

Met Gln Ala Ala Thr Val Leu Ile Asn Arg Arg Ala Leu Arg His Asn
1           5           10          15

```

```

Leu Gln Arg Leu Arg Glu Leu Ala Pro Ala Ser Lys Leu Val Ala Val
20           25           30

```

```

Val Lys Ala Asn Ala Tyr Gly His Gly Leu Ile Glu Thr Ala Arg Thr
35           40           45

```

```

Leu Pro Asp Ala Asp Ala Phe Gly Val Ala Arg Leu Glu Glu Ala Leu
50           55           60

```

```

Arg Leu Arg Ala Gly Gly Ile Thr Arg Pro Ile Leu Leu Leu Glu Gly
65           70           75           80

```

Phe Phe Glu Ala Asp Asp Leu Pro Thr Ile Ser Ala Glu His Leu His
85 90 95

Thr Ala Val His Asn Glu Glu Gln Leu Val Ala Leu Glu Asn Ala Glu
100 105 110

Leu Lys Glu Pro Val Thr Val Trp Met Lys Leu Asp Thr Gly Met His
115 120 125

Arg Leu Gly Val Leu Pro Glu Gln Ala Glu Ala Phe Tyr Gln Arg Leu
130 135 140

Ser Gln Cys Lys Asn Val Arg Gln Pro Val Asn Ile Val Ser His Phe
145 150 155 160

Ala Arg Ala Asp Glu Pro Gln Ser Gly Ala Thr Glu Lys Gln Leu Asp
165 170 175

Ile Phe Asn Thr Phe Cys Glu Gly Lys Pro Gly Gln Arg Ser Ile Ala
180 185 190

Ala Ser Gly Gly Ile Leu Leu Trp Pro Gln Ser His Phe Asp Trp Ala
195 200 205

Arg Pro Gly Ile Ile Leu Tyr Gly Val Ser Pro Leu Glu Asp Gly Thr
210 215 220

Thr Gly Ala Asp Phe Gly Cys Gln Pro Val Met Ser Leu Thr Ser Ser
225 230 235 240

Leu Ile Ala Val Arg Glu His Lys Ala Gly Glu Pro Val Gly Tyr Gly
245 250 255

Gly Thr Trp Val Ser Glu Arg Asp Thr Arg Leu Gly Val Val Ala Met
260 265 270

Gly Tyr Gly Asp Gly Tyr Pro Arg Ala Ala Pro Ser Gly Thr Pro Val
275 280 285

Leu Val Asn Gly Arg Glu Val Pro Ile Val Gly Arg Val Ala Met Asp
290 295 300

Met Ile Cys Val Asp Leu Gly Pro Gln Ala Gln Asp Lys Ala Gly Asp
305 310 315 320

Pro Val Ile Leu Trp Gly Glu Gly Leu Pro Val Glu Arg Ile Ala Glu
325 330 335

Met Thr Lys Val Ser Ala Tyr Glu Leu Ile Thr Arg Leu Thr Ser Arg
340 345 350

Val Ala Met Lys Tyr Val Asp
355

<210> 11
<211> 263
<212> PRT
<213> Artificial Sequence

<220>
<223> NanR

<400> 11

Met Gly Leu Met Asn Ala Phe Asp Ser Gln Thr Glu Asp Ser Ser Pro
1 5 10 15

Ala Ile Gly Arg Asn Leu Arg Ser Arg Pro Leu Ala Arg Lys Lys Leu
20 25 30

Ser Glu Met Val Glu Glu Glu Leu Glu Gln Met Ile Arg Arg Arg Glu
35 40 45

Phe Gly Glu Gly Glu Gln Leu Pro Ser Glu Arg Glu Leu Met Ala Phe
50 55 60

Phe Asn Val Gly Arg Pro Ser Val Arg Glu Ala Leu Ala Ala Leu Lys
65 70 75 80

Arg Lys Gly Leu Val Gln Ile Asn Asn Gly Glu Arg Ala Arg Val Ser
85 90 95

Arg Pro Ser Ala Asp Thr Ile Ile Gly Glu Leu Ser Gly Met Ala Lys
100 105 110

Asp Phe Leu Ser His Pro Gly Gly Ile Ala His Phe Glu Gln Leu Arg
115 120 125

Leu Phe Phe Glu Ser Ser Leu Val Arg Tyr Ala Ala Glu His Ala Thr
130 135 140

Asp Glu Gln Ile Asp Leu Leu Ala Lys Ala Leu Glu Ile Asn Ser Gln
145 150 155 160

Ser Leu Asp Asn Asn Ala Ala Phe Ile Arg Ser Asp Val Asp Phe His
165 170 175

Arg Val Leu Ala Glu Ile Pro Gly Asn Pro Ile Phe Met Ala Ile His
180 185 190

Val Ala Leu Leu Asp Trp Leu Ile Ala Ala Arg Pro Thr Val Thr Asp
195 200 205

Gln Ala Leu His Glu His Asn Asn Val Ser Tyr Gln Gln His Ile Ala
210 215 220

Ile Val Asp Ala Ile Arg Arg His Asp Pro Asp Glu Ala Asp Arg Ala
225 230 235 240

Leu Gln Ser His Leu Asn Ser Val Ser Ala Thr Trp His Ala Phe Gly
245 250 255

Gln Thr Thr Asn Lys Lys Lys
260

<210> 12
<211> 512
<212> PRT
<213> Artificial Sequence

<220>
<223> CadC

<400> 12

Met Gln Gln Pro Val Val Arg Val Gly Glu Trp Leu Val Thr Pro Ser
1 5 10 15

Ile Asn Gln Ile Ser Arg Asn Gly Arg Gln Leu Thr Leu Glu Pro Arg
20 25 30

Leu Ile Asp Leu Leu Val Phe Phe Ala Gln His Ser Gly Glu Val Leu
35 40 45

Ser Arg Asp Glu Leu Ile Asp Asn Val Trp Lys Arg Ser Ile Val Thr
50 55 60

Asn His Val Val Thr Gln Ser Ile Ser Glu Leu Arg Lys Ser Leu Lys
65 70 75 80

Asp Asn Asp Glu Asp Ser Pro Val Tyr Ile Ala Thr Val Pro Lys Arg
85 90 95

Gly Tyr Lys Leu Met Val Pro Val Ile Trp Tyr Ser Glu Glu Glu Gly
100 105 110

Glu Glu Ile Met Leu Ser Ser Pro Pro Pro Ile Pro Glu Ala Val Pro
115 120 125

Ala Thr Asp Ser Pro Ser His Ser Leu Asn Ile Gln Asn Thr Ala Thr
130 135 140

Pro Pro Glu Gln Ser Pro Val Lys Ser Lys Arg Phe Thr Thr Phe Trp
145 150 155 160

Val Trp Phe Phe Phe Leu Leu Ser Leu Gly Ile Cys Val Ala Leu Val
165 170 175

Ala Phe Ser Ser Leu Asp Thr Arg Leu Pro Met Ser Lys Ser Arg Ile
180 185 190

Leu Leu Asn Pro Arg Asp Ile Asp Ile Asn Met Val Asn Lys Ser Cys
195 200 205

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

Asp Leu Val Ala Thr Ser Leu Asn Thr Phe Ser Thr Phe Met Val His
225 230 235 240

Asp Lys Ile Asn Tyr Asn Ile Asp Glu Pro Ser Ser Ser Gly Lys Thr
245 250 255

Leu Ser Ile Ala Phe Val Asn Gln Cys Gln Tyr Arg Ala Gln Gln Cys
260 265 270

Phe Met Ser Ile Lys Leu Val Asp Asn Ala Asp Gly Ser Thr Met Leu
275 280 285

Asp Lys Arg Tyr Val Ile Thr Asn Gly Asn Gln Leu Ala Ile Gln Asn
290 295 300

Asp Leu Leu Glu Ser Leu Ser Lys Ala Leu Asn Gln Pro Trp Pro Gln
305 310 315 320

Arg Met Gln Glu Thr Leu Gln Lys Ile Leu Pro His Arg Gly Ala Leu
325 330 335

Leu Thr Asn Phe Tyr Gln Ala His Asp Tyr Leu Leu His Gly Asp Asp
340 345 350

Lys Ser Leu Asn Arg Ala Ser Glu Leu Leu Gly Glu Ile Val Gln Ser
355 360 365

Ser Pro Glu Phe Thr Tyr Ala Arg Ala Glu Lys Ala Leu Val Asp Ile
370 375 380

Val Arg His Ser Gln His Pro Leu Asp Glu Lys Gln Leu Ala Ala Leu
385 390 395 400

Asn Thr Glu Ile Asp Asn Ile Val Thr Leu Pro Glu Leu Asn Asn Leu
405 410 415

Ser Ile Ile Tyr Gln Ile Lys Ala Val Ser Ala Leu Val Lys Gly Lys
420 425 430

Thr Asp Glu Ser Tyr Gln Ala Ile Asn Thr Gly Ile Asp Leu Glu Met
435 440 445

Ser Trp Leu Asn Tyr Val Leu Leu Gly Lys Val Tyr Glu Met Lys Gly
450 455 460

Met Asn Arg Glu Ala Ala Asp Ala Tyr Leu Thr Ala Phe Asn Leu Arg
465 470 475 480

Pro Gly Ala Asn Thr Leu Tyr Trp Ile Glu Asn Gly Ile Phe Gln Thr
485 490 495

Ser Val Pro Tyr Val Val Pro Tyr Leu Asp Lys Phe Leu Ala Ser Glu
500 505 510

<210> 13
<211> 329
<212> PRT
<213> Artificial Sequence

<220>
<223> Cbh

<400> 13

Met Cys Thr Gly Leu Ala Leu Glu Thr Lys Asp Gly Leu His Leu Phe
1 5 10 15

Gly Arg Asn Met Asp Ile Glu Tyr Ser Phe Asn Gln Ser Ile Ile Phe
20 25 30

Ile Pro Arg Asn Phe Lys Cys Val Asn Lys Ser Asn Lys Lys Glu Leu
35 40 45

Thr Thr Lys Tyr Ala Val Leu Gly Met Gly Thr Ile Phe Asp Asp Tyr
50 55 60

Pro Thr Phe Ala Asp Gly Met Asn Glu Lys Gly Leu Gly Cys Ala Gly
65 70 75 80

Leu Asn Phe Pro Val Tyr Val Ser Tyr Ser Lys Glu Asp Ile Glu Gly
85 90 95

Lys Thr Asn Ile Pro Val Tyr Asn Phe Leu Leu Trp Val Leu Ala Asn
100 105 110

Phe Ser Ser Val Glu Glu Val Lys Glu Ala Leu Lys Asn Ala Asn Ile
115 120 125

Val Asp Ile Pro Ile Ser Glu Asn Ile Pro Asn Thr Thr Leu His Trp
130 135 140

Met Ile Ser Asp Ile Thr Gly Lys Ser Ile Val Val Glu Gln Thr Lys
145 150 155 160

Glu Lys Leu Asn Val Phe Asp Asn Asn Ile Gly Val Leu Thr Asn Ser
165 170 175

Pro Thr Phe Asp Trp His Val Ala Asn Leu Asn Gln Tyr Val Gly Leu
180 185 190

Arg Tyr Asn Gln Val Pro Glu Phe Lys Leu Gly Asp Gln Ser Leu Thr
195 200 205

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

Pro Ala Ser Arg Phe Ile Arg Val Ala Phe Leu Arg Asp Ala Met Ile
 225 230 235 240

Lys Asn Asp Lys Asp Ser Ile Asp Leu Ile Glu Phe Phe His Ile Leu
 245 250 255

Asn Asn Val Ala Met Val Arg Gly Ser Thr Arg Thr Val Glu Glu Lys
 260 265 270

Ser Asp Leu Thr Gln Tyr Thr Ser Cys Met Cys Leu Glu Lys Gly Ile
 275 280 285

Tyr Tyr Tyr Asn Thr Tyr Glu Asn Asn Gln Ile Asn Ala Ile Asp Met
 290 295 300

Asn Lys Glu Asn Leu Asp Gly Asn Glu Ile Lys Thr Tyr Lys Tyr Asn
 305 310 315 320

Lys Thr Leu Ser Ile Asn His Val Asn
 325

<210> 14
 <211> 1539
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> CadC ORF

<400> 14
 atgcaacaac ctgtagttcg cgttggcgaa tggcttgta ctccgtccat aaaccaaatt 60
 agccgcaatg ggcgtcaact tacccttgag ccgagattaa tcgatcttct ggttttcttt 120
 gctcaacaca gtggcgaagt acttagcagg gatgaactta tcgataatgt ctggaagaga 180
 agtattgtca ccaatcacgt tgtgacgcag agtatctcag aactacgtaa gtcattaataa 240

gataatgatg aagatagtcc tgtctatatc gctactgtac caaagcgcgg ctataaatta	300
atggtgccgg ttatctggta cagcgaagaa gagggagagg aaataatgct atcttcgcct	360
ccccctatac cagaggcggg tcctgccaca gattctccct cccacagtct taacattcaa	420
aacaccgcaa cgccacctga acaatcccca gttaaaagca aacgattcac taccttttgg	480
gtatggtttt ttttcctgtt gtcgttaggt atctgtgtag cactggtagc gttttcaagt	540
cttgatacac gtcttcctat gagcaaatcg cgtattttgc tcaatccacg cgatattgac	600
attaatatgg taaataaaaag ttgtaacagc tggagttccc cgtatcagct ctcttacgcg	660
ataggcgtgg gtgatttggg ggcgacatca cttaacacct tctccacctt tatggtgcat	720
gacaaaatca actacaacat tgatgaaccg agcagttccg gtaaaacatt atctattgcg	780
tttgttaatc agtgccaata ccgtgctcaa caatgcttta tgtcgataaa attggtagac	840
aatgcagatg gttcaaccat gctggataaa cgttatgtca tcactaacgg taatcagctg	900
gcgattcaaa atgattttact ggagagttta tcaaaagcgt taaaccaacc gtggccacaa	960
cgaatgcagg agacgctcca gaaaattttg ccgcatcgtg gtgcgttatt aactaatttt	1020
tatcaggcac atgattatth actgcatggc gatgataaat cattgaaccg tgccagtga	1080
ttattaggtg agattgttca atcatcccca gaattttacct acgcgagagc agaaaaagca	1140
ttagttgata tcgtgcgcca ttctcaacat cttttagatg aaaaacaatt agcagcactg	1200
aacacagaaa tagataacat tgttacactg ccggaattga acaacctgtc cattatatat	1260
caaataaaag cggtcagtgc tctggtaaaa ggtaaaacag atgagtctta ccaggcgata	1320
aatactggca ttgatcttga aatgtcctgg ctaaattatg tgttgcttgg caaggtttat	1380
gaaatgaagg ggatgaaccg ggaagcagct gatgcatatc tcaccgcctt taatttacgc	1440
ccaggggcaa acacccttta ctggattgaa aatggtatat tccagacttc tgttccttat	1500
gttgtacctt atctcgacaa atttcttgct tcagaataa	1539