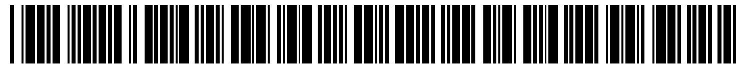



(19) **United States**

(12) **Patent Application Publication**

**Moran et al.**

(10) **Pub. No.: US 2019/0015528 A1**

(43) **Pub. Date: Jan. 17, 2019**

(54) **ENGINEERED MICROBIAL POPULATION**

(71) Applicant: **Board of Regents, The University of Texas System**, Austin, TX (US)

(72) Inventors: **Nancy Moran**, Austin, TX (US); **Jeffrey Barrick**, Austin, TX (US); **Sean Leonard**, Austin, TX (US)

(21) Appl. No.: **16/029,686**

(22) Filed: **Jul. 9, 2018**

**Related U.S. Application Data**

(60) Provisional application No. 62/529,754, filed on Jul. 7, 2017.

**Publication Classification**

(51) **Int. Cl.**

| | |
|---|---|
| ***A61K 48/00*** | (2006.01) |
| ***A61K 9/00*** | (2006.01) |
| ***A23K 50/90*** | (2006.01) |
| ***C12N 15/74*** | (2006.01) |
| ***C12N 1/20*** | (2006.01) |

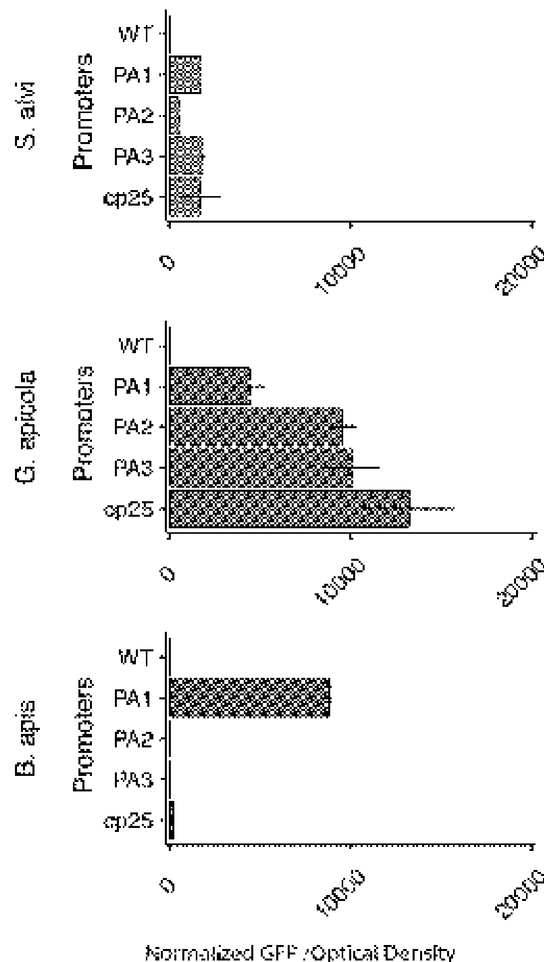
(52) **U.S. Cl.**

CPC ........ ***A61K 48/005*** (2013.01); ***A61K 9/0053*** (2013.01); ***A23K 50/90*** (2016.05); *C12N 2310/141* (2013.01); ***C12N 1/20*** (2013.01); *A01K 2227/706* (2013.01); *C12N 2310/20* (2017.05); ***C12N 15/746*** (2013.01)

(57) **ABSTRACT**

Provided herein are genetically engineered bacteria that are native to a host insect microbiome. Further provided are methods of inducing RNA interference in an insect, such as a bee, by administering the genetically engineered bacteria.

**Specification includes a Sequence Listing.**



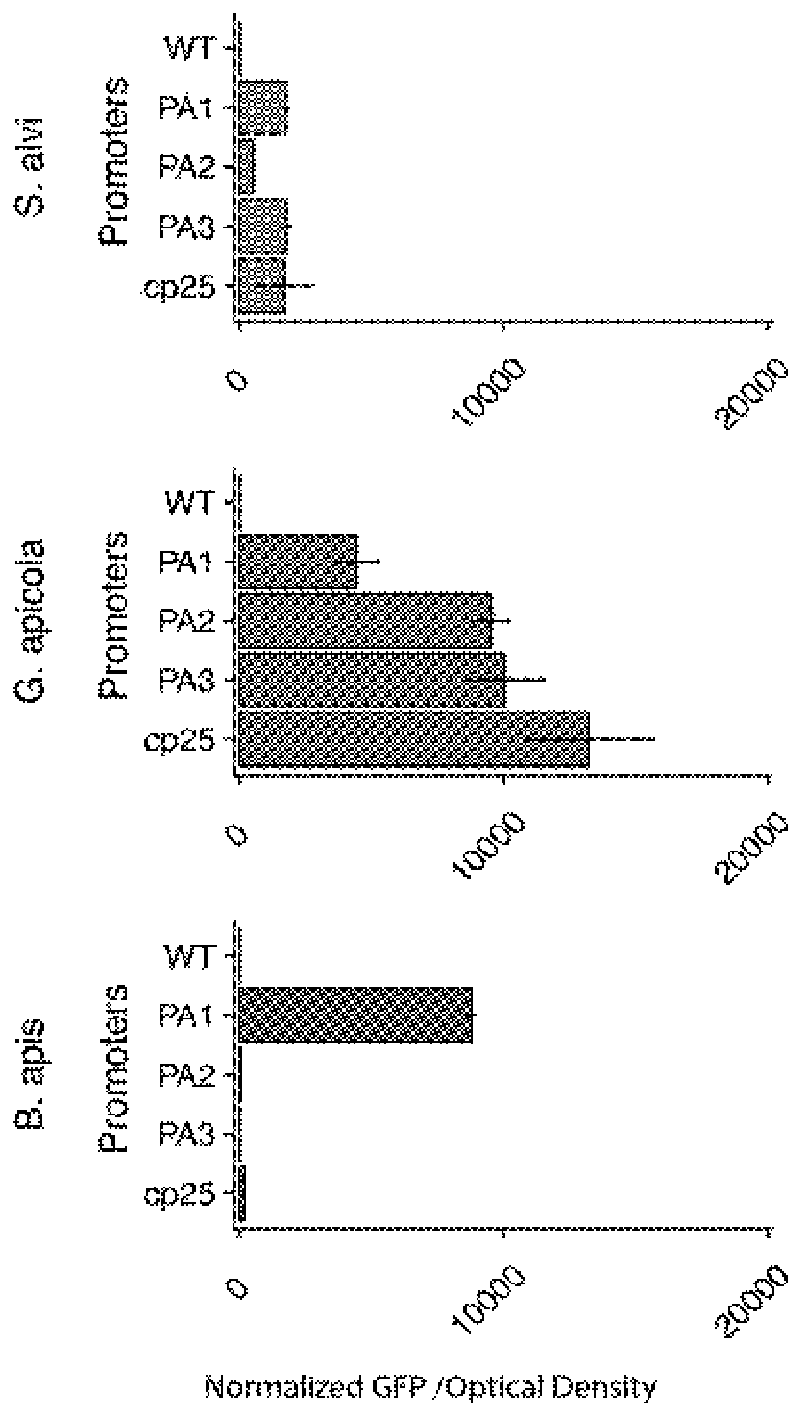



Figure 1

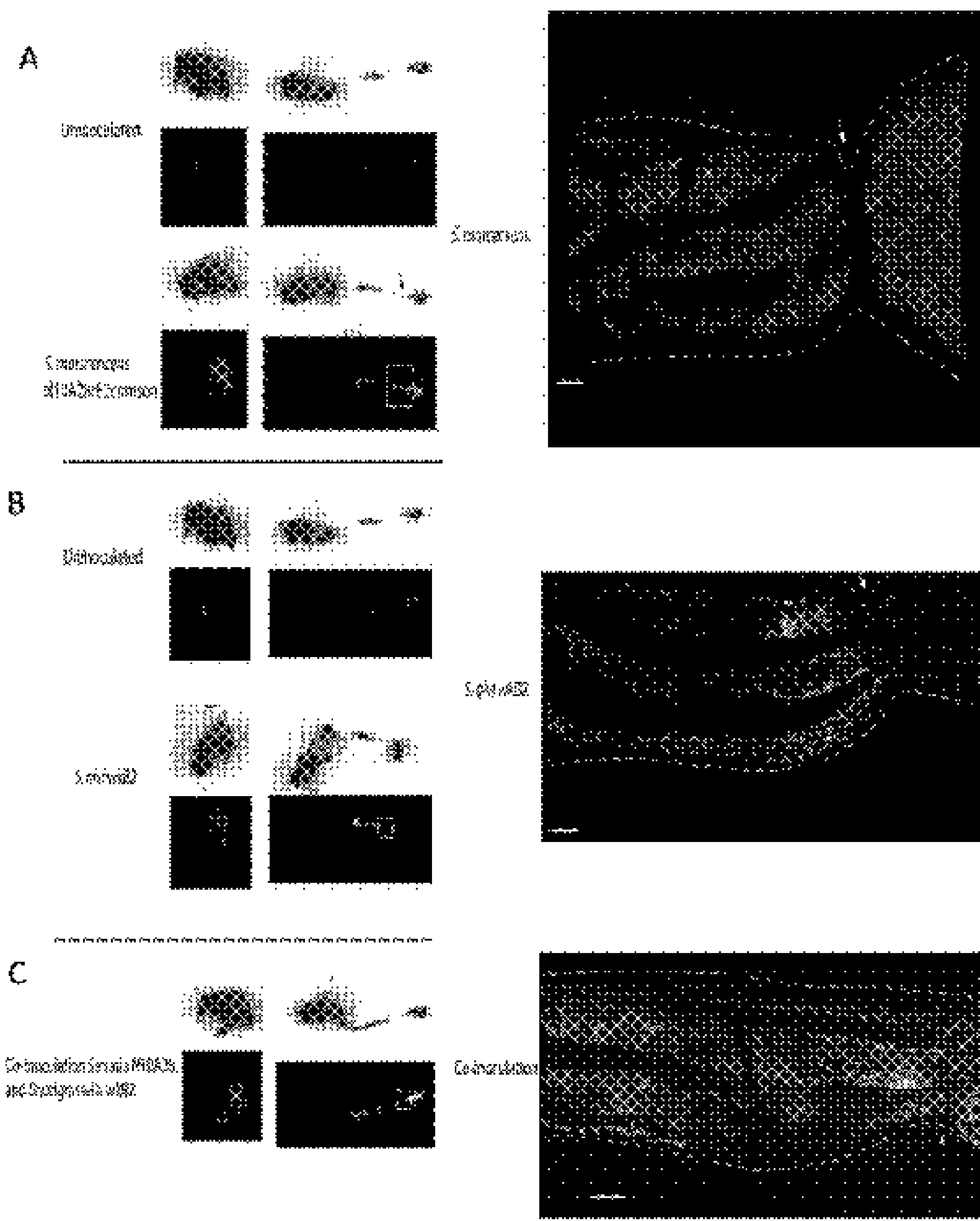



Figure 2A-Figure 2C

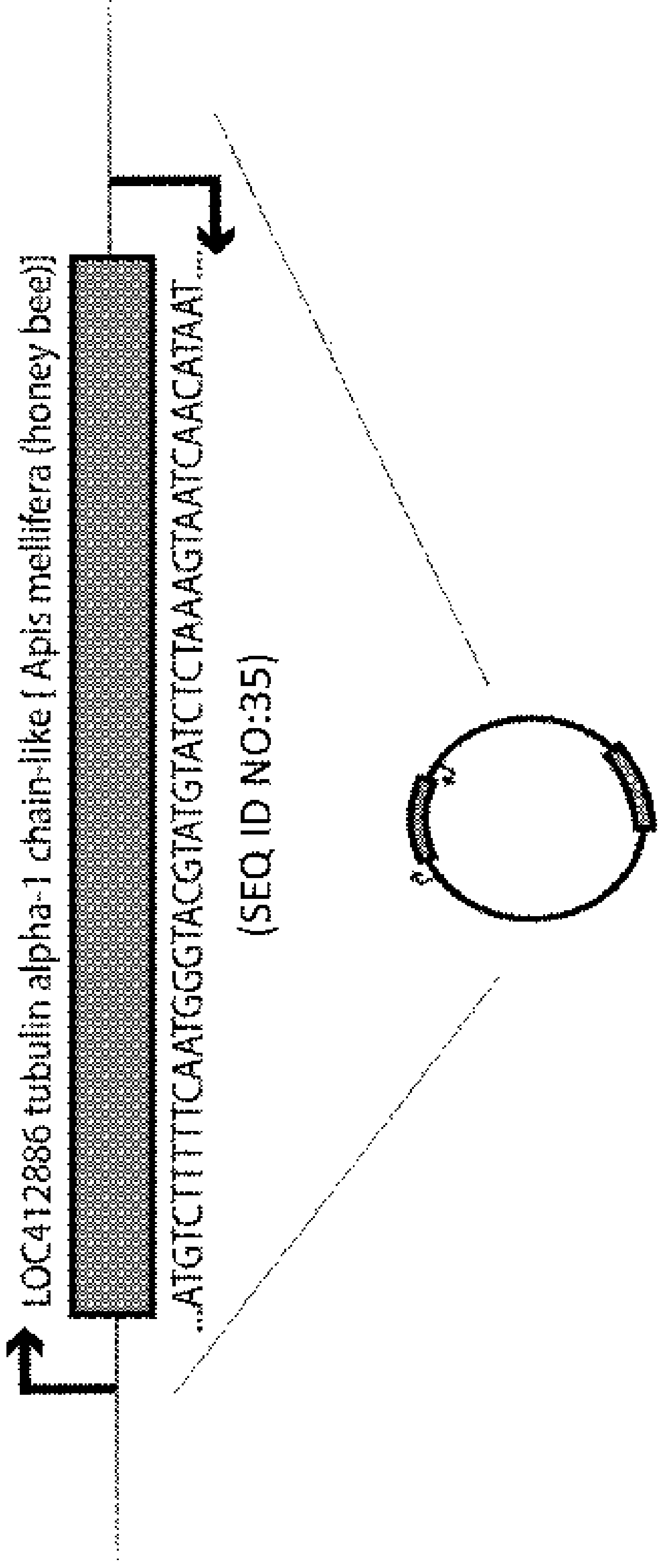



Figure 3

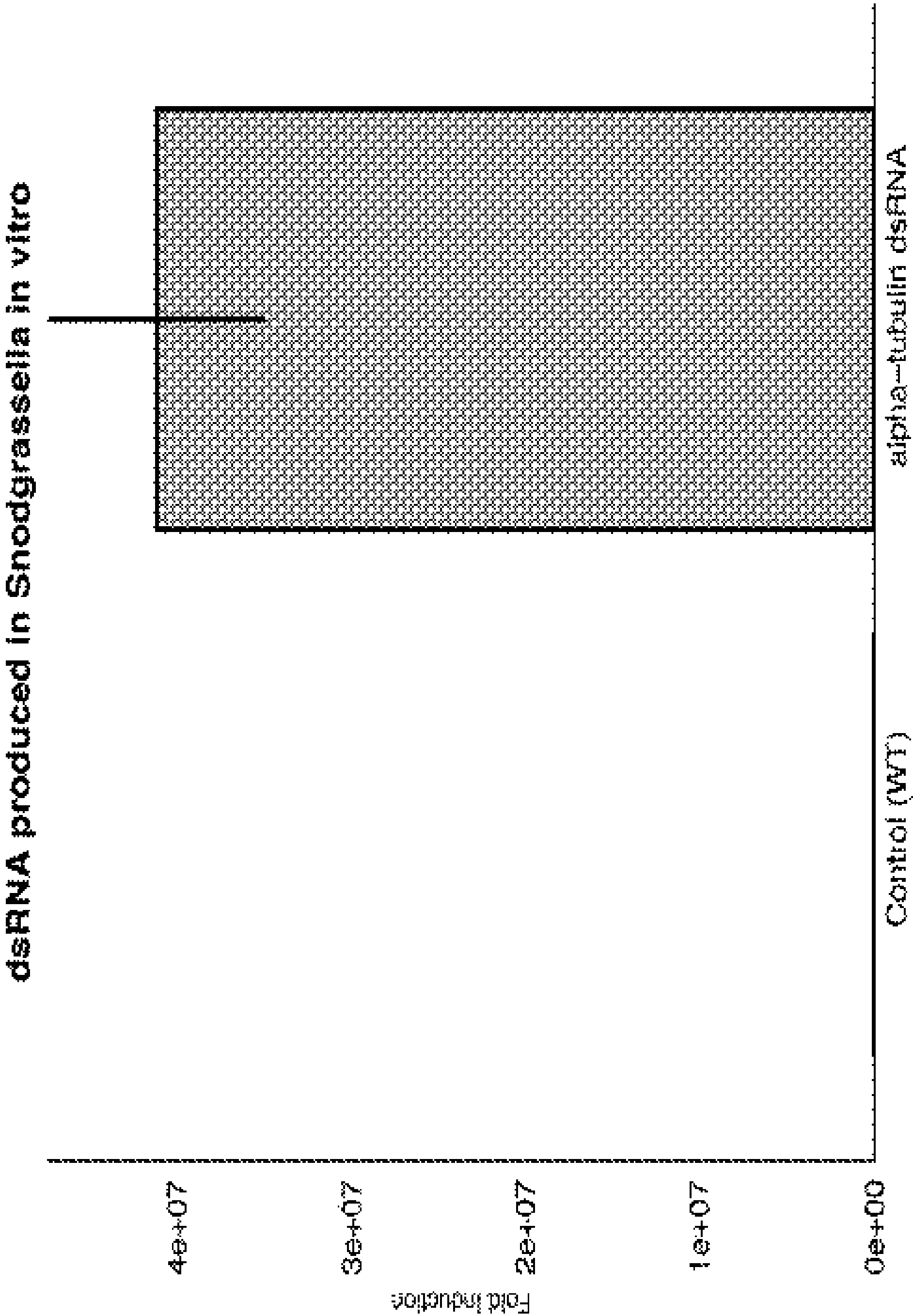



Figure 4

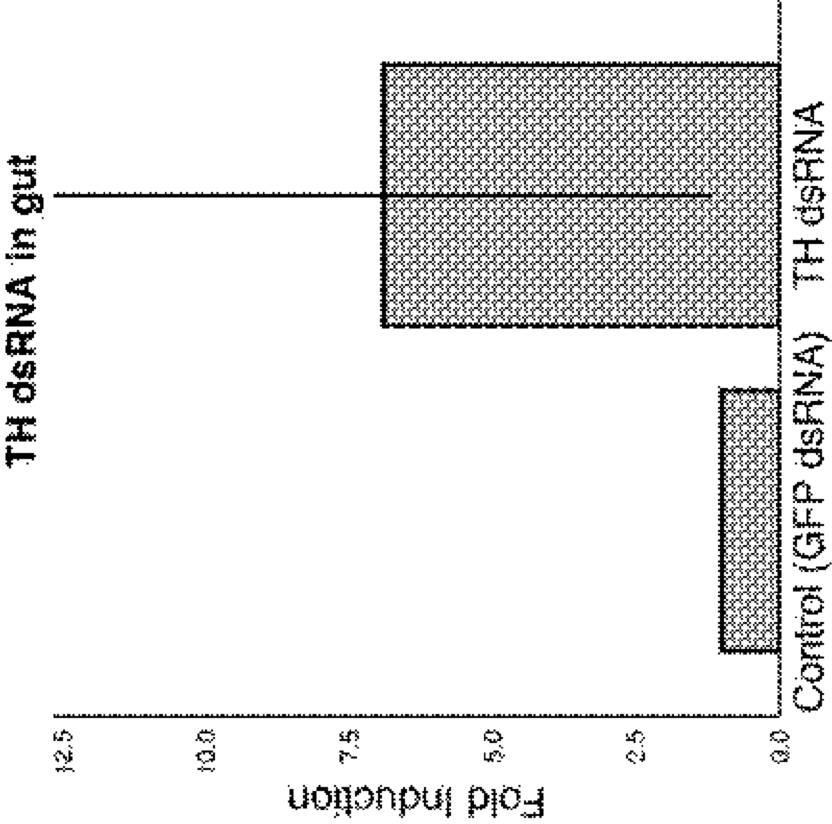



Figure 5

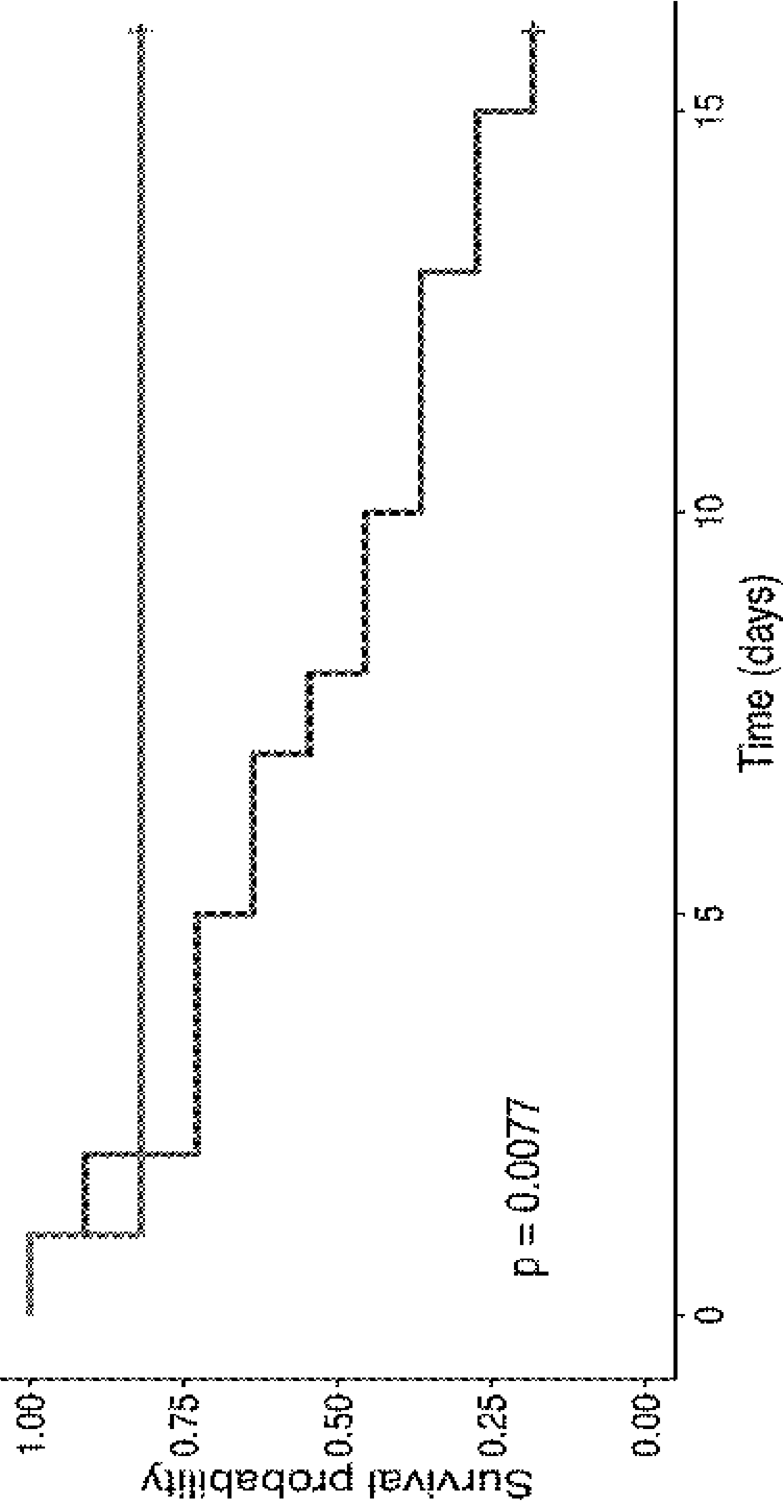



Figure 6

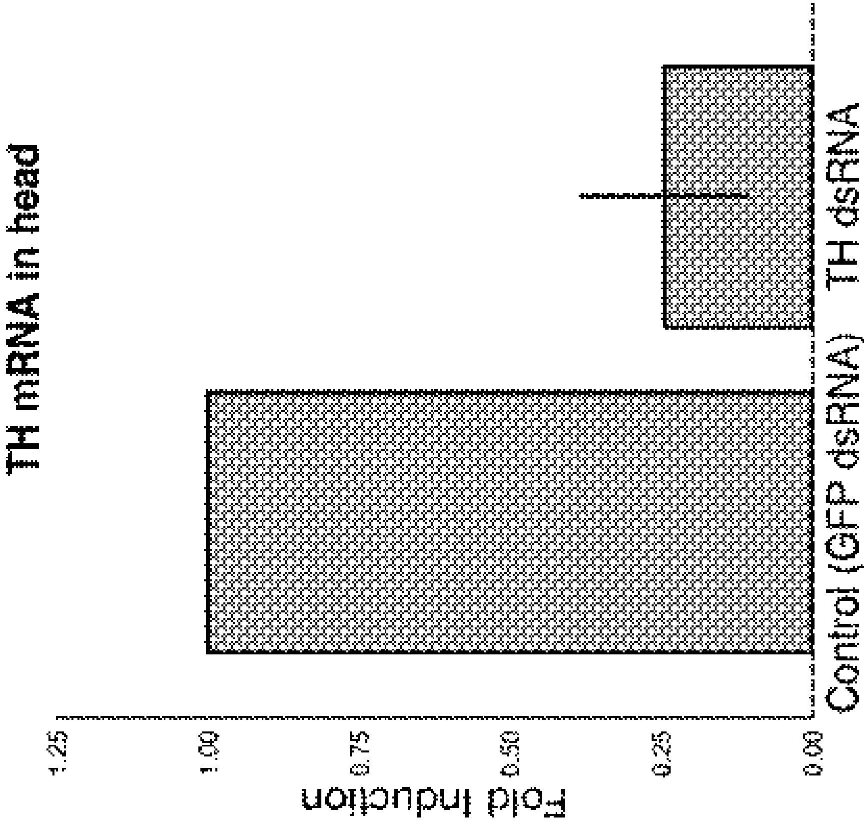



Figure 7

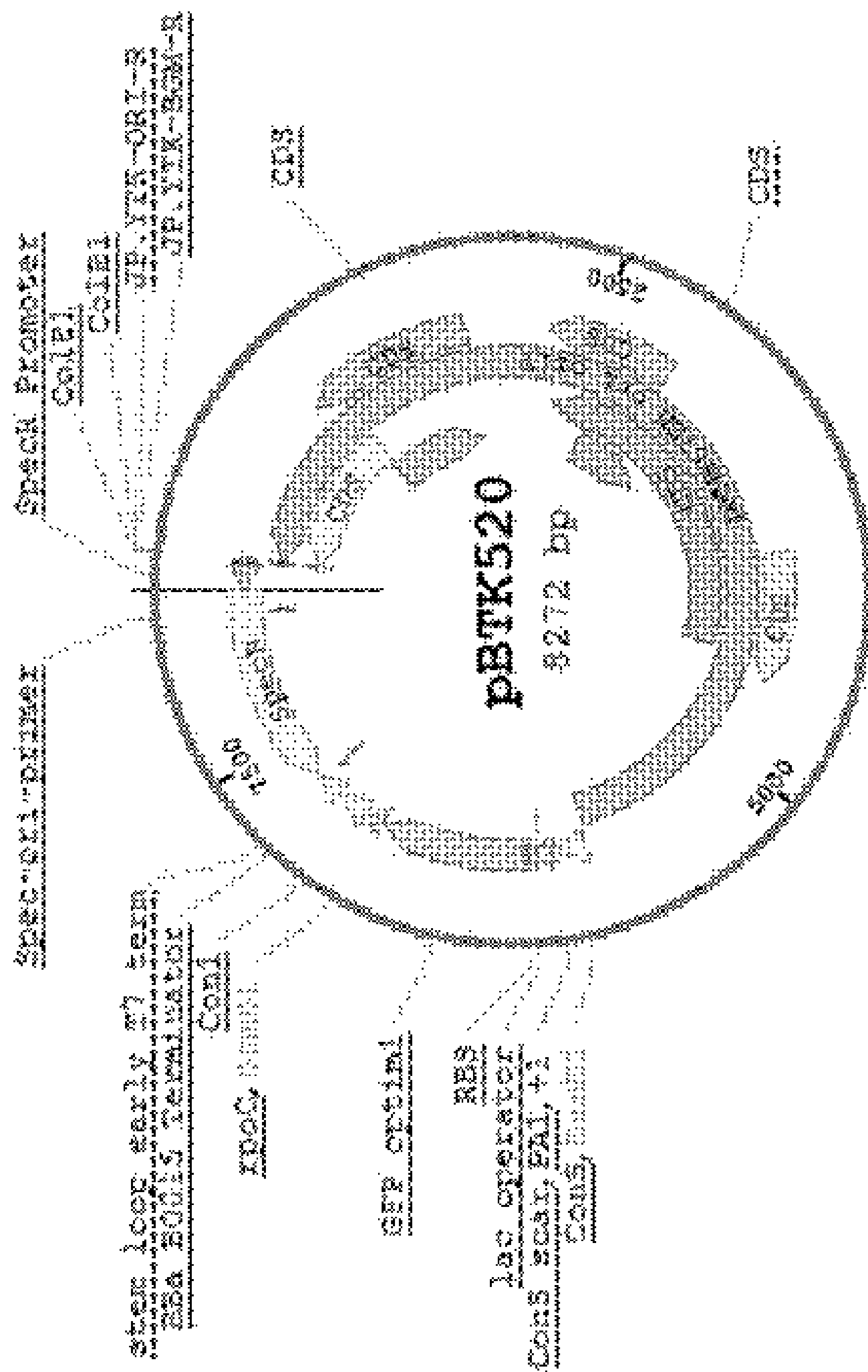



Figure 8A

Figure 8B

Figure 8C

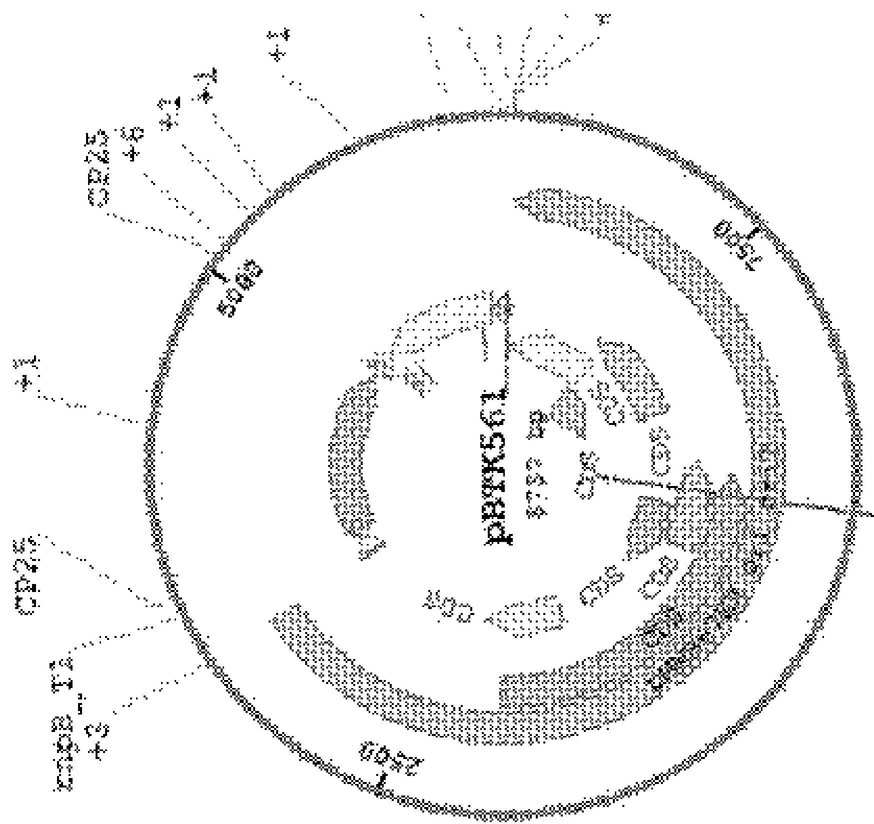


Figure 8D

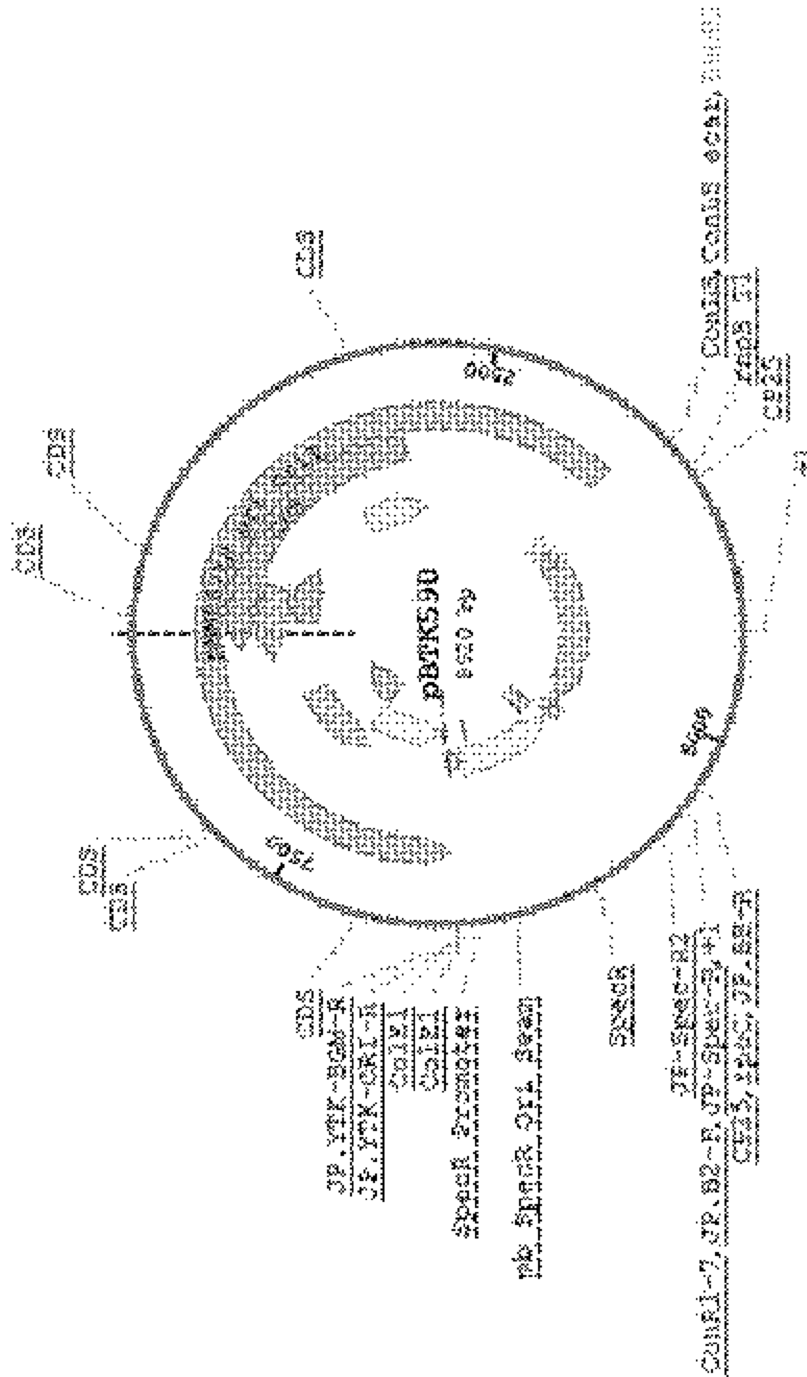



Figure 8E

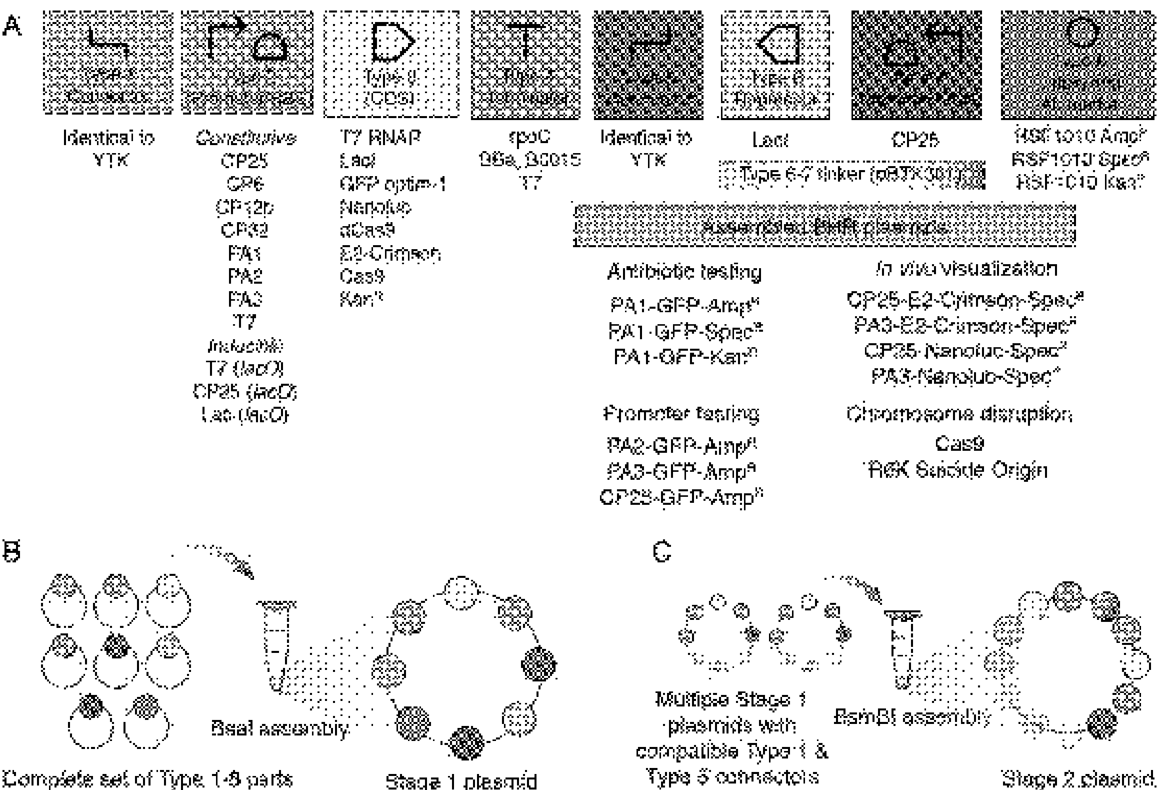



Figure 9A-Figure 9C

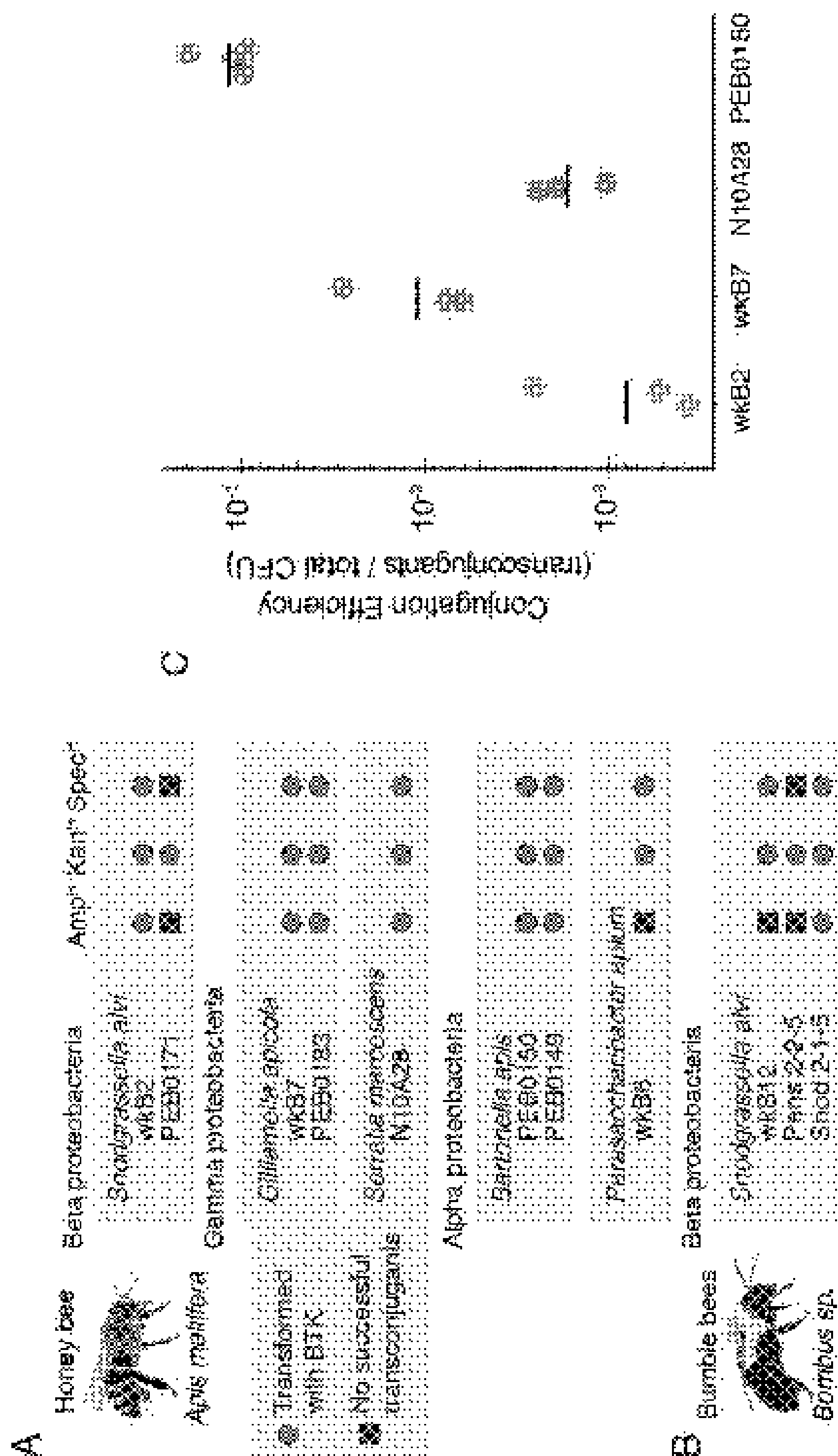



Figure 10A-Figure 10C

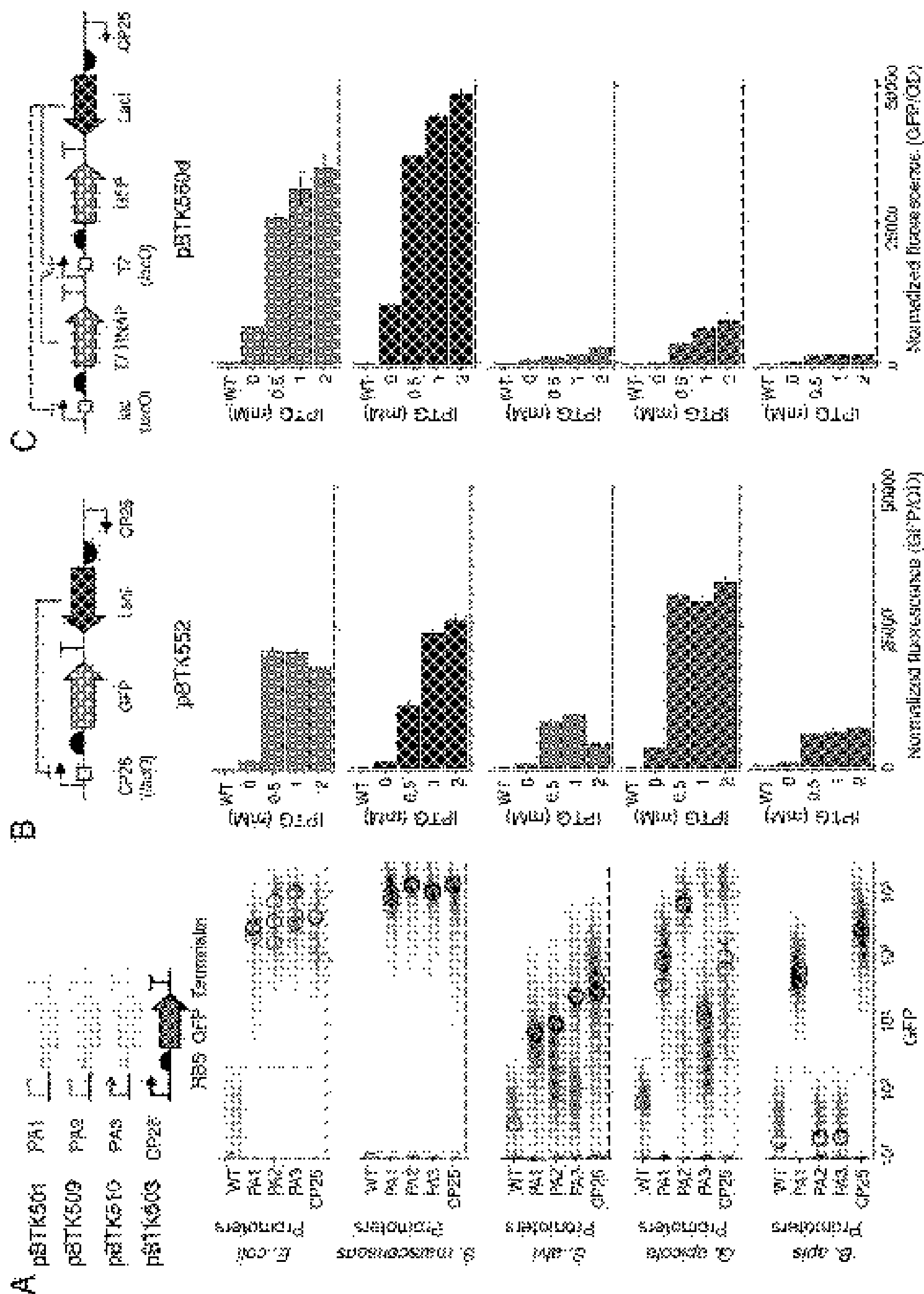



Figure 11A-Figure 11C

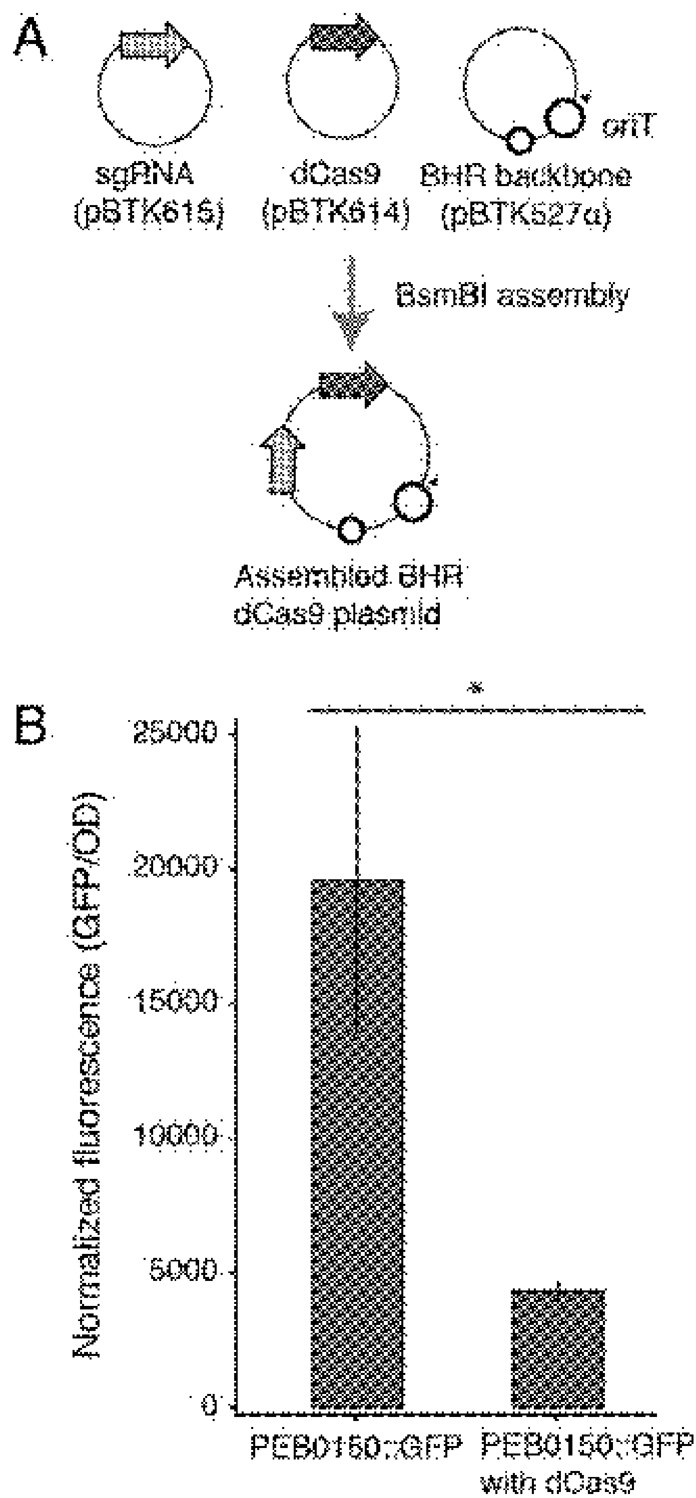



Figure 12A-Figure 12B

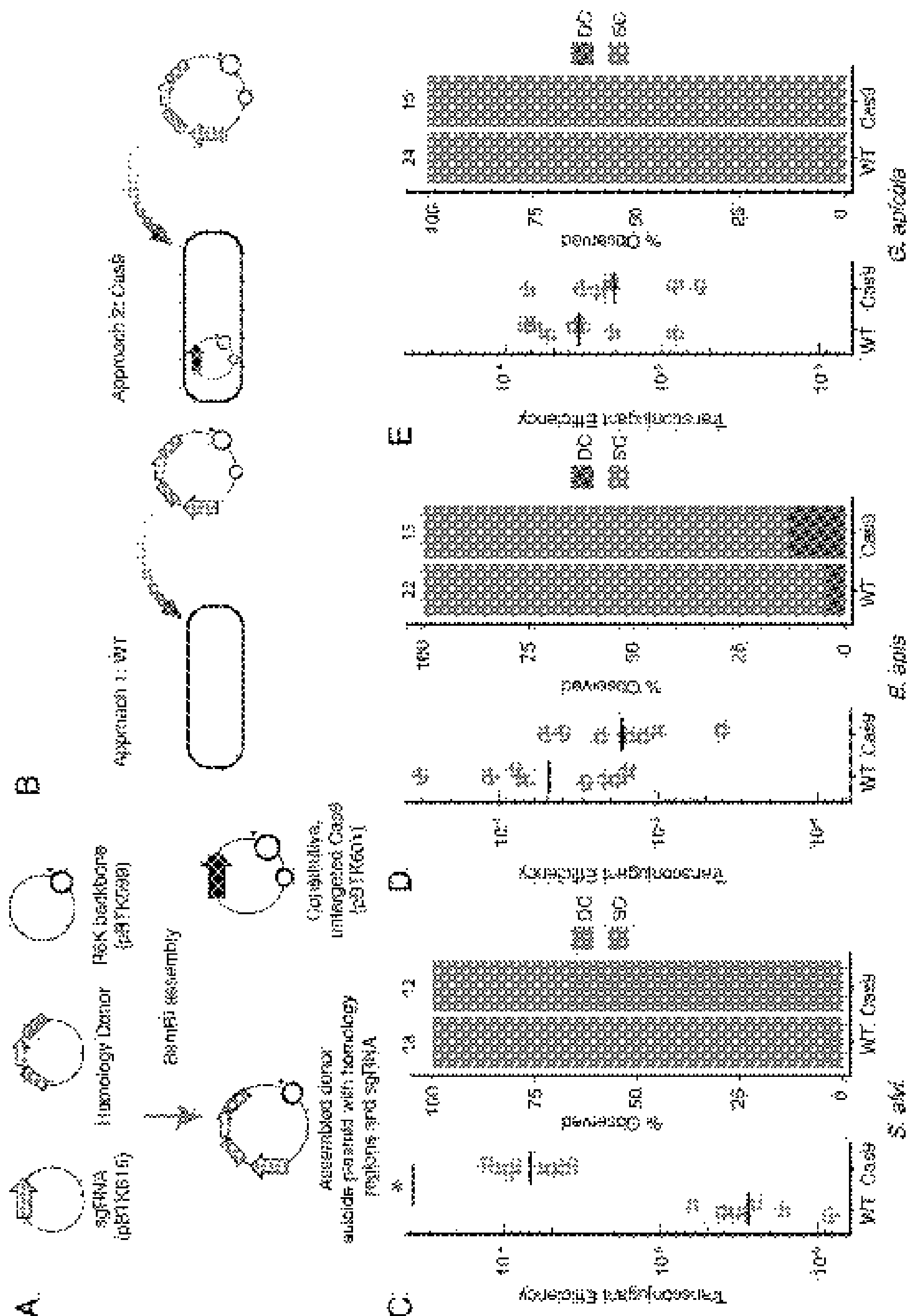



Figure 13A-Figure 13E

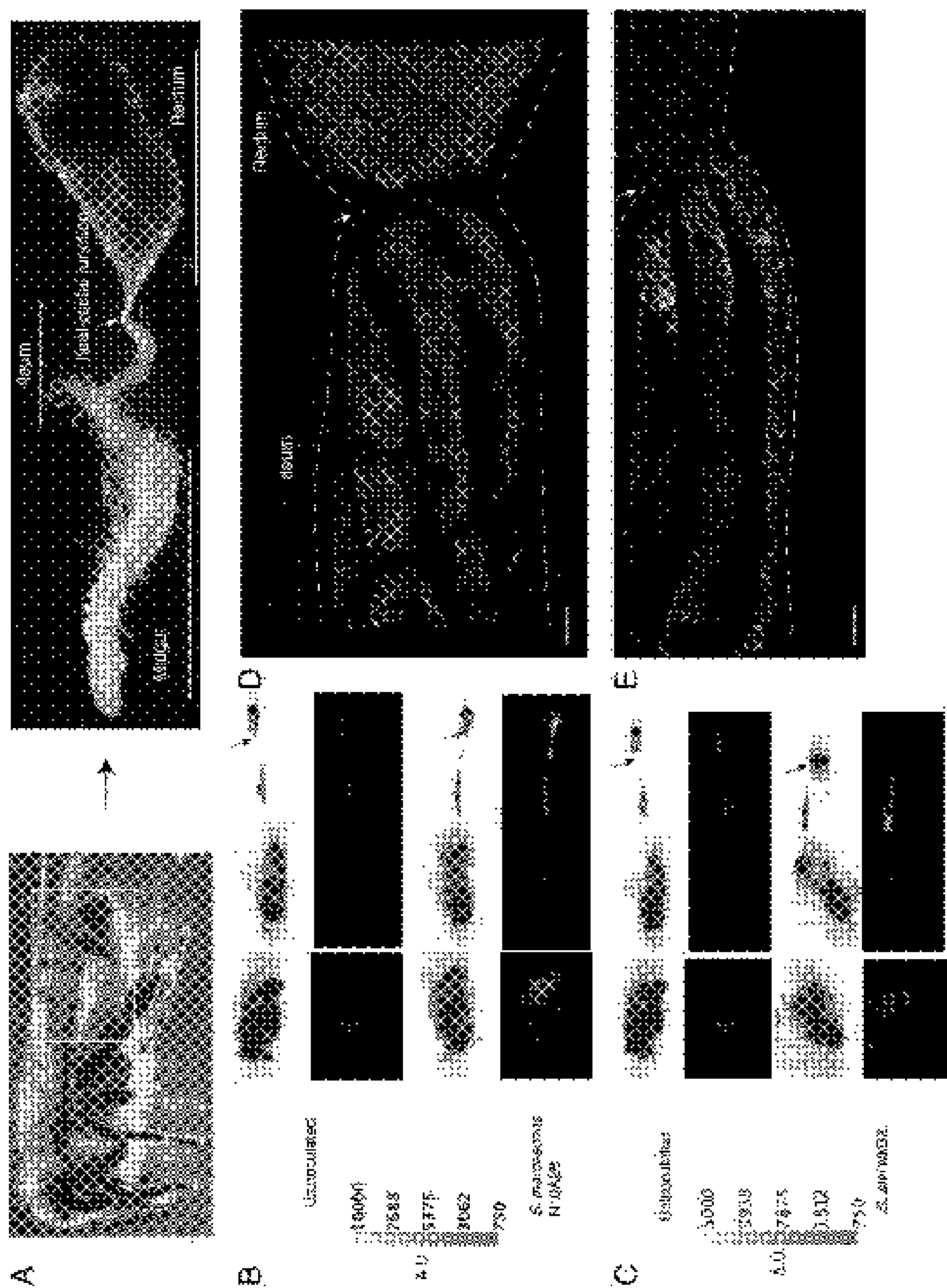


Figure 14A-Figure 14E

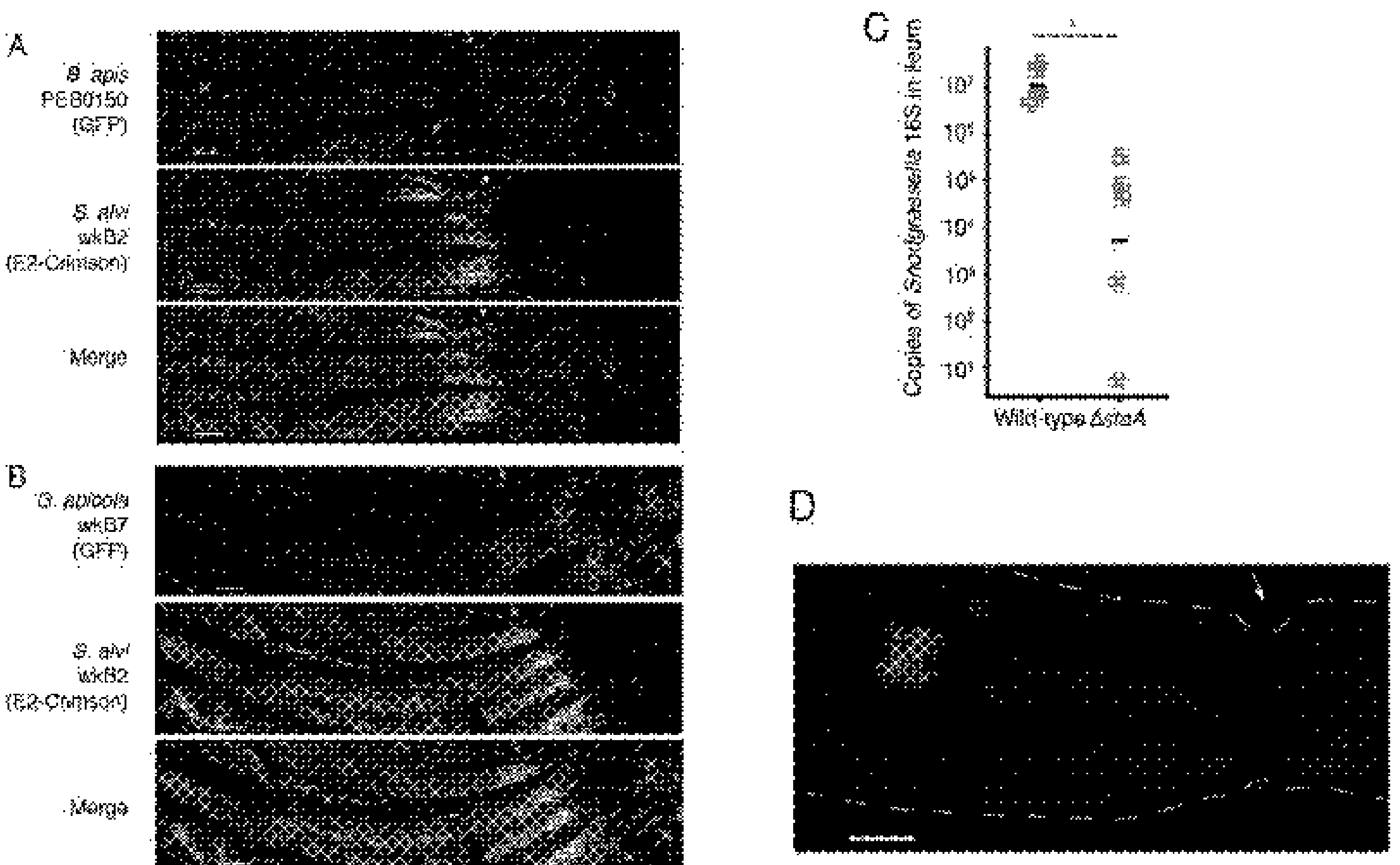



Figure 15A-Figure 15D

Broad-host-range Plasmid Screen in Bee Gut Bacteria

| **Plasmid** | RP4 | pTD-C_sfYFP | pAKgfp | pBMTBX-3 | pMMB67EH | pBMTBX-2 |
|---|---|---|---|---|---|---|
| **Origin** | RP4 | mini-rp4 | pBBR1 | pBBR1 | RSF1010 | pBBR1 |
| **Antibiotic Marker** | Amp, Kan | Strep, Spec | Amp | Cam | Amp | Kan |
| ***Snodgrassella alvi*** | - | - | + | - | + | + |
| ***Gilliamella apicola*** | - | - | - | - | + | - |

"+" indicates transconjugants were successfully passaged on selective media, and plasmid was re-isolated and able to re-transform *E. coli* DH5alpha cells.

Figure 16

Bacterial Strains

| Species and strain | Source | ID |
|---|---|---|
| *E. coli* MFDpir | [1] | N/A |
| *E. coli* DH5alpha | Thermo-Fisher | CAT# 11319-019 |
| *E. coli* EC100D | Lucigen | CAT# ECP09500 |
| *Snodgrassella alvi* wkB2 | [2] | ATCC BAA-2449 |
| *Snodgrassella alvi* PEB0171 | This study | N/A |
| *Snodgrassella alvi* wkB12 | [2] | N/A |
| *Snodgrassella alvi* Snod 2-1-5 | This study | N/A |
| *Snodgrassella alvi* PENS 2-2-5 | This study | N/A |
| *Gilliamella apicola* wkB7 | [3] | N/A |
| *Gilliamella apicola* PEB0183 | This study | N/A |
| *Bartonella apis* PEB0150 | [4] | N/A |
| *Bartonella apis* PEB0149 | [4] | N/A |
| *Parasaccharibacter apium* wkB8 | This study | N/A |
| *Serratia marcescens* N10A28 | This study | N/A |

Figure 17

| part | type | 5'-site | 3'-site | description | marker | origin | Part Source |
|---|---|---|---|---|---|---|---|
| | | | | **Entry Vector** | | | |
| pBTK001 | entry vector | | | entry vector for generating new parts | CamR | p15A | 1 |
| | | | | **pYTK001 from Yeast Toolkit is a suitable entry vector** | | | |
| | | | | **Type 1 - Connectors** | | | |
| | | | | **Use Connector sequences from Yeast Toolkit** | | | |
| | | | | **Type 2 - Promoters** | | | |
| | *Constitutive* | | | | | | |
| pBTK102 | promoter | 2 | 2 | T7 + RBS | CamR | ColE1 | 2 |
| pBTK107 | promoter | 2 | 2 | CP25 + RBS | CamR | ColE1 | 3 |
| pBTK110 | promoter | 2 | 2 | CP6 + RBS | CamR | ColE1 | 3 |
| pBTK112 | promoter | 2 | 2 | CP12b + RBS | CamR | ColE1 | 2 |
| pBTK113 | promoter | 2 | 2 | CP22 + RBS | CamR | ColE1 | 2 |
| pBTK119 | promoter | 2 | 2 | PA1 + RBS | CamR | p15A | 4 |
| pBTK120 | promoter | 2 | 2 | PA2 + RBS | CamR | p15A | 4 |
| pBTK121 | promoter | 2 | 2 | PA3 + RBS | CamR | p15A | 4 |
| | *Inducible* | | | | | | |
| pBTK103 | promoter | 2 | 2 | Lac lacO + RBS | CamR | ColE1 | 5 |
| pBTK124 | promoter | 2 | 2 | CP25 (lacO) | CamR | ColE1 | |
| pBTK130 | promoter | 2 | 2 | T7 (lacO) | CamR | ColE1 | |
| | | | | **Type 3 - Coding Sequences** | | | |
| pBTK200 | coding_sequence | 3 | 3 | T7 RNA Polymerase | CamR | ColE1 | 2 |
| pBTK203 | coding_sequence | 3 | 3 | LacI repressor | CamR | ColE1 | 6 |
| pBTK205 | coding_sequence | 3 | 3 | GFP optim-1 | CamR | ColE1 | 7 |
| pBTK206 | coding_sequence | 3 | 3 | [illegible] | CamR | ColE1 | 8 |
| pBTK209 | coding_sequence | 3 | 3 | dCas9 | CamR | ColE1 | 9 |
| pBTK224 | coding_sequence | 3 | 3 | E2-Crimson | CamR | ColE1 | 10 |
| pBTK229 | For homologous recombination | 3 | 3 | Kanamycin Resistance | CamR | ColE1 | 2 |
| | | | | **Type 4 - Terminators** | | | |
| pBTK300 | terminator | 4 | 4 | [illegible] | CamR | ColE1 | 11 |
| pBTK301 | terminator | 6 | 7 | BBa_B0015 Terminator | CamR | ColE1 | 12 |
| pBTK305 | terminator | 4 | 4 | T7 | CamR | ColE1 | 2 |
| | | | | **Type 5 - Connectors** | | | |
| | | | | **Use Connector sequences from Yeast Toolkit** | | | |
| | | | | **Type 6 - Repressors (reverse CDS)** | | | |
| pBTK211 | | 6 | 6 | LacI Reverse | CamR | ColE1 | 13 |
| | | | | **Type VII - Reverse promoters** | | | |
| pBTK136 | promoter | 7 | 7 | CP25 + RBS Reverse | CamR | ColE1 | 3 |
| | | | | **Type 8 - Origins and Antibiotics** | | | |
| pBTK401 | origin_marker | 8 | 8 | mRFP1 dropout | AmpR | RSF1010 | 14 |
| pBTK402 | origin_marker | 8 | 8 | mRFP1 dropout | KanR | RSF1010 | 14 |
| pBTK403 | origin_marker | 8 | 8 | mRFP1 dropout | SpecR | RSF1010 | 14 |

Figure 18

| Vectors for Stage 2 Assembly | | | | | | | |
|---|---|---|---|---|---|---|---|
| pBTK527 | Stage 1 | ConLS' | ConRE' | SPACE | KanR | RSF1010 | |
| pBTK527s | Stage 1 | ConLS' | ConRE' | SPACE | SpecR | RSF1010 | |
| pBTK599 | for homologous recombination | 6 | 8 | R6K-sfGFP | AmpR | R6k | 15 |
| pBTK599s | for homologous recombination | 6 | 8 | R6K-sfGFP | SpecR | R6k | 15 |

| Assembled Plasmids | | | | | | |
|---|---|---|---|---|---|---|
| pBTK501 | Stage 1 | ConLS' | ConRE | PA1 GFP optim1 | AmpR | RSF1010 |
| pBTK503 | Stage 1 | ConLS' | ConRE | CP25 GFP optim1 | AmpR | RSF1010 |
| pBTK509 | Stage 1 | ConLS' | ConRE | PA2 GFP optim1 | AmpR | RSF1010 |
| pBTK510 | Stage 1 | ConLS' | ConRE | PA3 GFP optim1 | AmpR | RSF1010 |
| pBTK519 | Stage 1 | ConLS | ConR1 | PA1 GFP optim1 | KanR | RSF1010 |
| pBTK520 | Stage 1 | ConLS | ConR1 | PA1 GFP optim1 | SpecR | RSF1010 |
| pBTK549 | Stage 1 | ConLS' | ConRE | T7 RNAP + CP25 LacI reverse | SpecR | RSF1010 |
| pBTK541 | Stage 1 | ConLS | ConRE | T7 lacO GFP optim1 | AmpR | RSF1010 |
| pBTK550d | Stage 2 | NA | NA | pBTK549 + pBTK541 | SpecR | RSF1010 |
| pBTK552 | Stage 1 | ConLS | ConR1 | CP25(lacO) inducible GFP | SpecR | RSF1010 |
| pBTK553 | Stage 1 | ConLS | ConR1 | CP25 NanoLuc luciferase | SpecR | RSF1010 |
| pBTK554 | Stage 1 | ConLS | ConR1 | PA3 NanoLuc luciferase | SpecR | RSF1010 |
| pBTK569 | Stage 2 | NA | NA | CP25 E2-crimson | SpecR | RSF1010 |
| pBTK570 | Stage 2 | NA | NA | PA3 E2-crimson | SpecR | RSF1010 |
| pBTK601 | Stage 2 | NA | NA | Cas9 (no sgRNA) | SpecR | RSF1010 |
| pBTK614 | Stage 1 | ConL1 | ConRE | dCas9 (no sgRNA) | AmpR | ColE1 |
| pBTK615 | Stage 1 | ConLS | ConR2 | sgRNA (gfp) | AmpR | ColE1 |
| pBTK619 | Stage 1 | ConLS | ConRE | CP25b GFP optim1 | KanR | RSF1010 |
| pBTK620 | Stage 1 | ConLS | ConRE | CP32 GFP optim1 | KanR | RSF1010 |
| pBTK621 | Stage 1 | ConLS | ConRE | CP6 GFP optim1 | KanR | RSF1010 |

Figure 18 (Continued)

# Oligonucleotides Used in the Study

| ID | Use | Source | Sequence (5' to 3') | |
|---|---|---|---|---|
| wkB2_KO_1 | Validate wkB2::staA mutant, upstream junction F | This Study | TAGCAACAGCGACAGAACCG | SEQ ID NO:13 |
| wkB2_KO_2 | Validate wkB2::staA mutant, upstream junction R | This Study | CTCAGGCGCAATCACAAATG | SEQ ID NO:14 |
| wkB2_KO_3 | Validate wkB2::staA mutant, downstream junction F | This Study | AATGCTGTTTTCCCGGGAT | SEQ ID NO:15 |
| wkB2_KO_4 | Validate wkB2::staA mutant, downstream junction R | This Study | CTGGAAATTCCTGCTGCCAG | SEQ ID NO:16 |
| wkB7_KO_1 | Validate wkB7::sacA mutant, upstream junction F | This Study | TCGCTTGTTCACAGCGATAGA | SEQ ID NO:17 |
| wkB7_KO_2 | Validate wkB7::sacA mutant, upstream junction R | This Study | GCTTTTGCCATTGTCACCGG | SEQ ID NO:18 |
| wkB7_KO_3 | Validate wkB7::sacA mutant, downstream junction F | This Study | ATTGGCAACGCTACCTTTGC | SEQ ID NO:19 |
| wkB7_KO_4 | Validate wkB7::sacA mutant, downstream junction R | This Study | GTCGTCGTATTTGTCAGGAGCA | SEQ ID NO:20 |
| d150_KO_1 | Validate PEB0150::carG mutant, upstream junction F | This Study | TGTTCAGGACGTGAATACACCTTAT | SEQ ID NO:21 |
| d150_KO_2 | Validate PEB0150::carG mutant, upstream junction R | This Study | AAATTGCAGTTTGATTTGATGCTGG | SEQ ID NO:22 |
| d150_KO_3 | Validate PEB0150::carG mutant, downstream junction F | This Study | GCGAGCCCATTTATACCCATATAAA | SEQ ID NO:23 |
| d150_KO_4 | Validate PEB0150::carG mutant, downstream junction R | This Study | TTCAAGGTTCCATTTGCCTTTTTCA | SEQ ID NO:24 |
| KO_5 | Amplify backbone of suicide plasmid F | This Study | ACCTGTTGATAGTACGTAGTAAGCTC | SEQ ID NO:25 |
| KO_6 | Amplify backbone of suicide plasmid R | This Study | TCAAAATTCCTTTGAGAGCCTCTAAG | SEQ ID NO:26 |
| pMMB_F | Validate replication of pMMB67EH / BTK plasmids F | This Study | TCATGTCATCCAGGTCAAACTC | SEQ ID NO:27 |
| pMMB_R | Validate replication of pMMB67EH / BTK plasmids R | This Study | TCAAGAAATCCAGAGCCATCAA | SEQ ID NO:28 |
| Beta-1009-qtF | 16S betaproteobacteria qPCR primer F, Beta-1009-qtF | * | CTTAGAGATAGGAGAGTG | SEQ ID NO:29 |
| Beta-1115-qtR | 16S betaproteobacteria qPCR primer R, Beta-1115-qtR | * | TAATGATGGCAACTAATGACAA | SEQ ID NO:30 |
| dCas9_sg | sgRNA sequence targeting gfp for dCas9 | This Study | CCCAATGCTCGTCCAATTAGACC | SEQ ID NO:31 |
| wkB2_KO_sg | Cas9 sgRNA targeting S. alvi wkB2 staA | This Study | GAATTCGAAGCATAAGAATT | SEQ ID NO:32 |
| wkB7_KO_sg | Cas9 sgRNA targeting S. apicola wkB7 sacA | This Study | TTCCACTTAATACGTGCATC | SEQ ID NO:33 |
| d150_KO_sg | Cas9 sgRNA targeting B. apis PEB0150 carG | This Study | ACACCATGTCCATGCGAAAA | SEQ ID NO:34 |

Figure 19

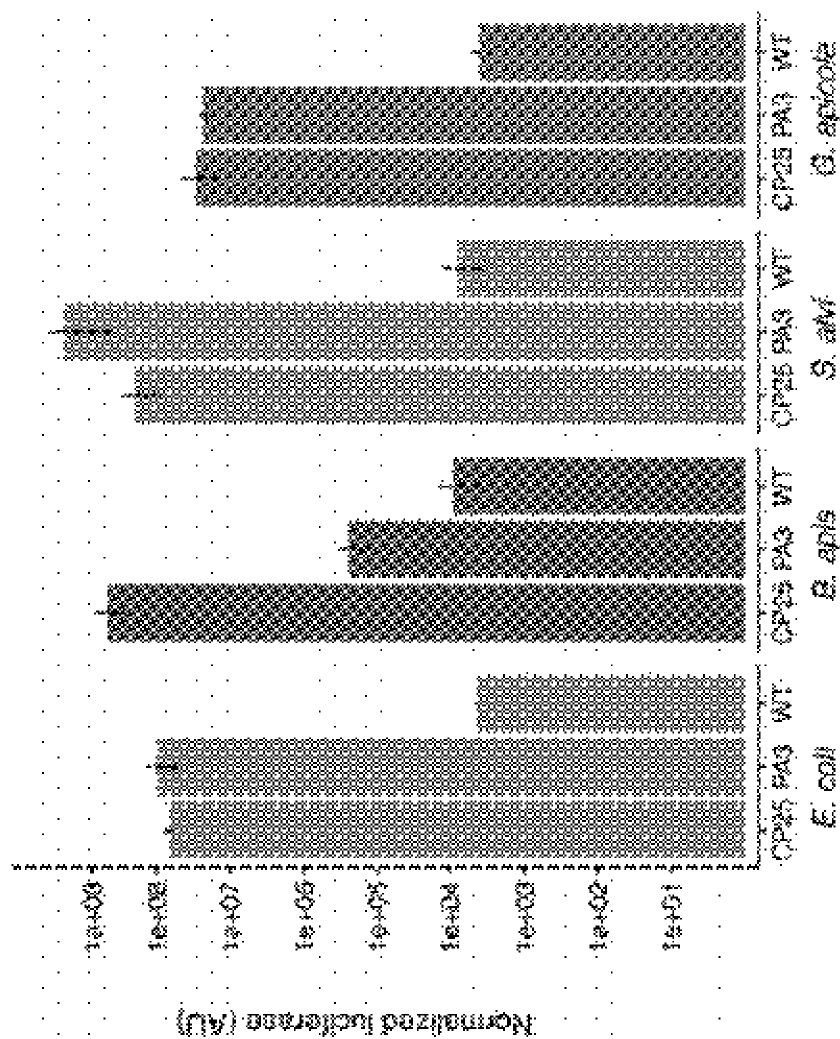



Figure 20

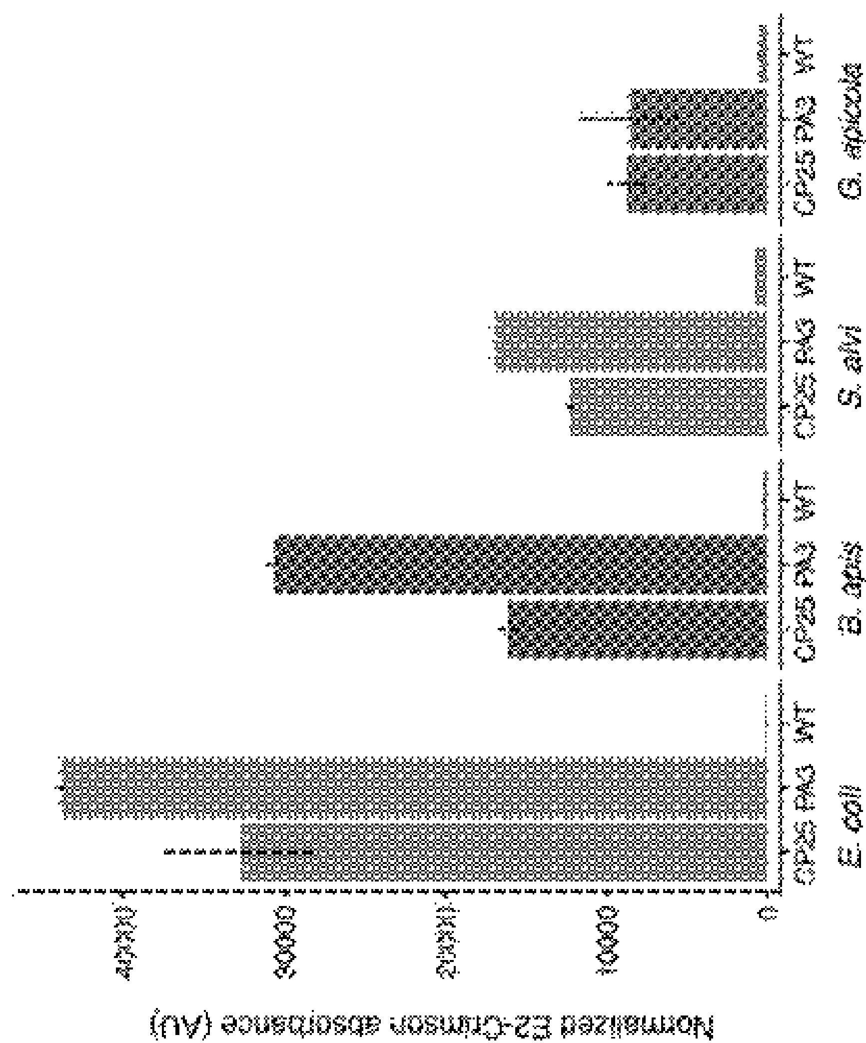



Figure 21

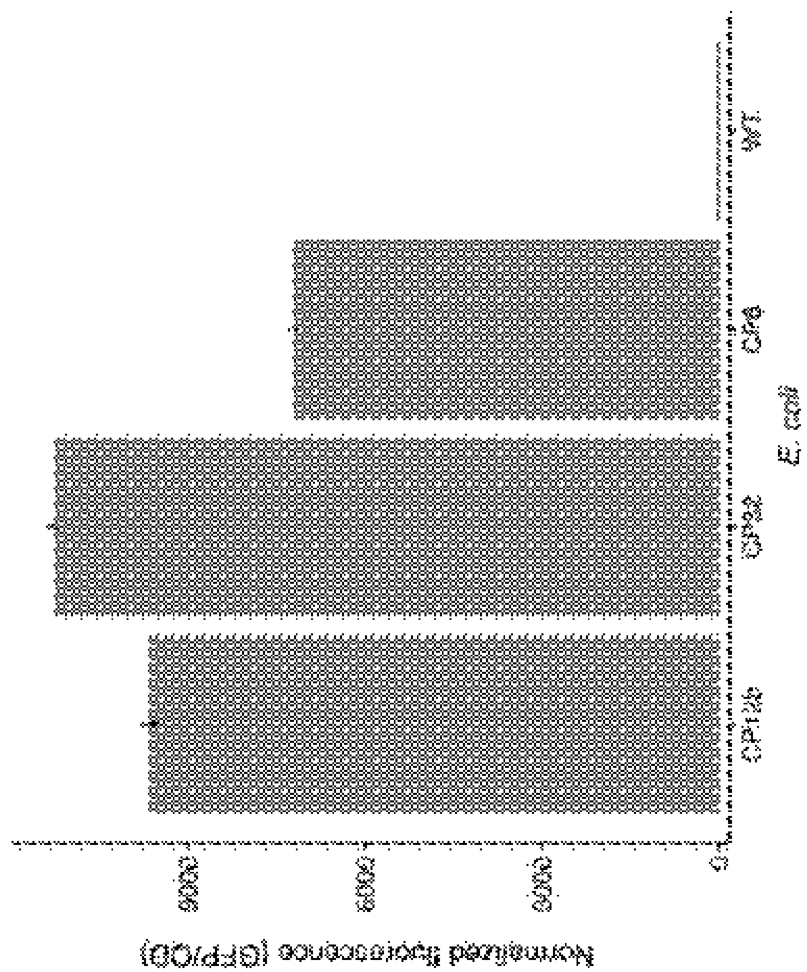



Figure 22

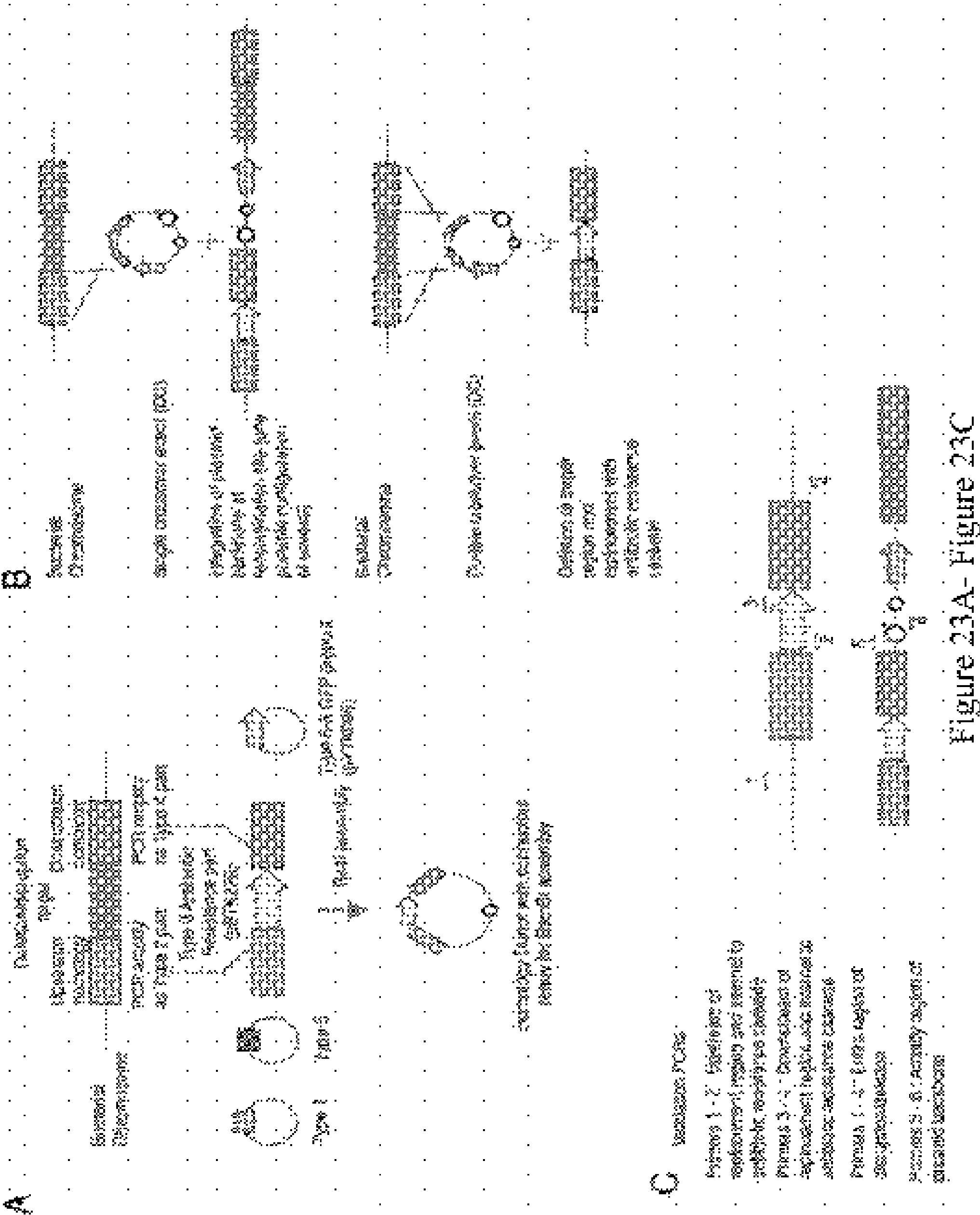


Figure 23A- Figure 23C

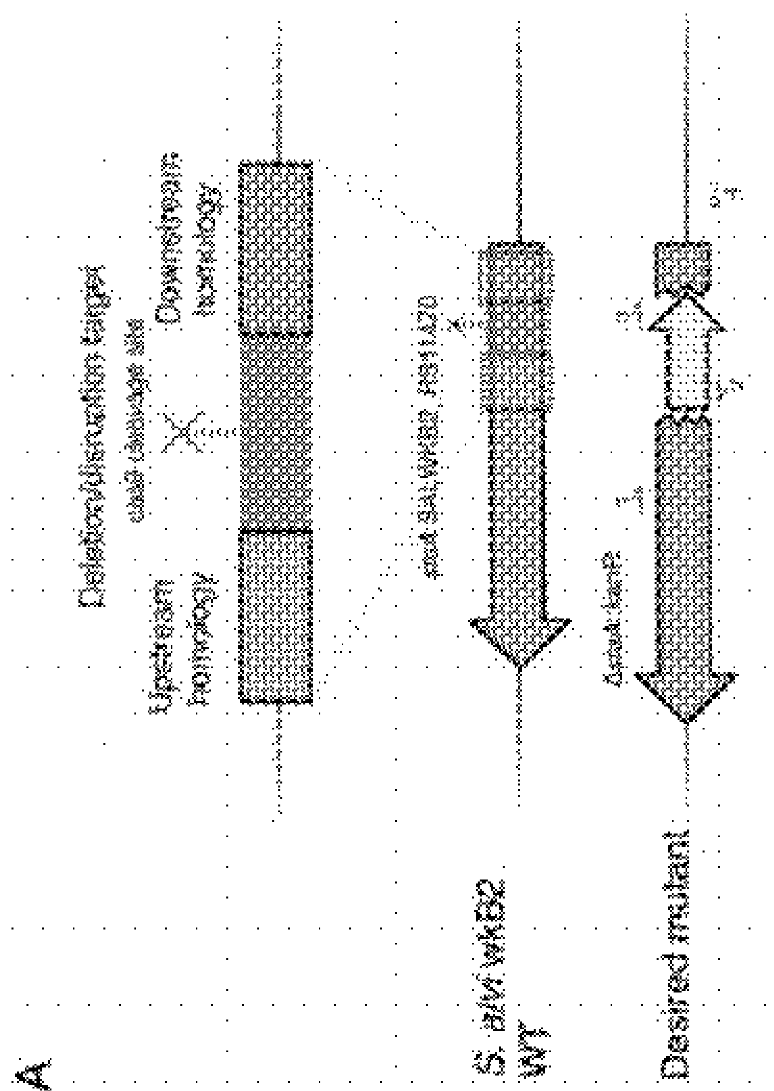



Figure 24A

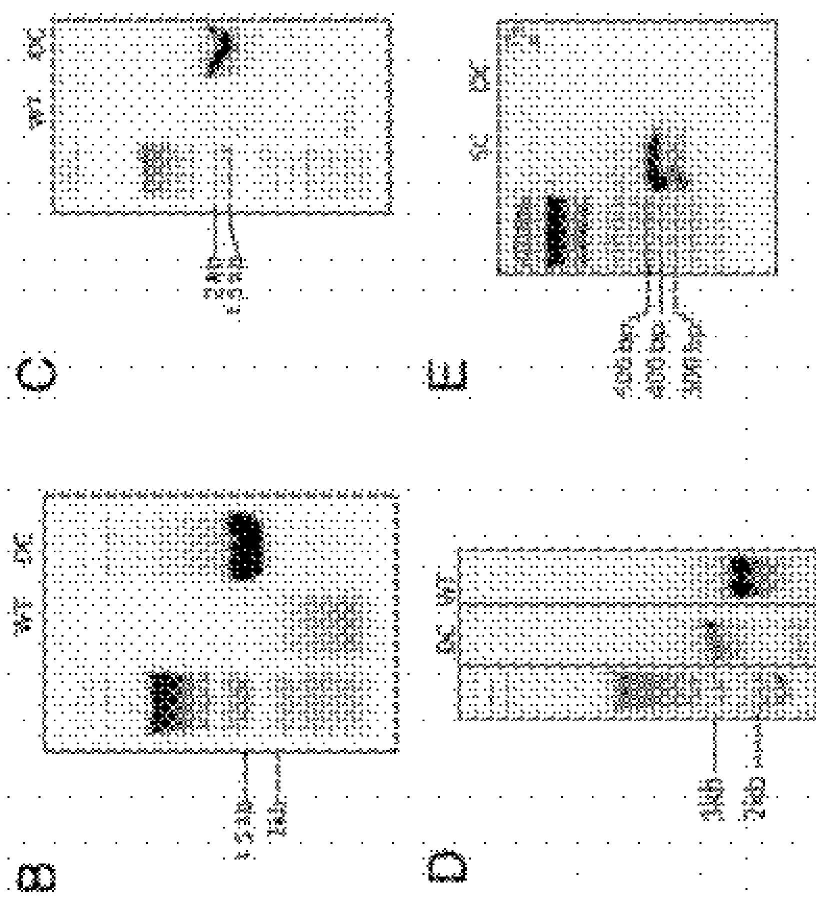



Figure 24B- Figure 24E

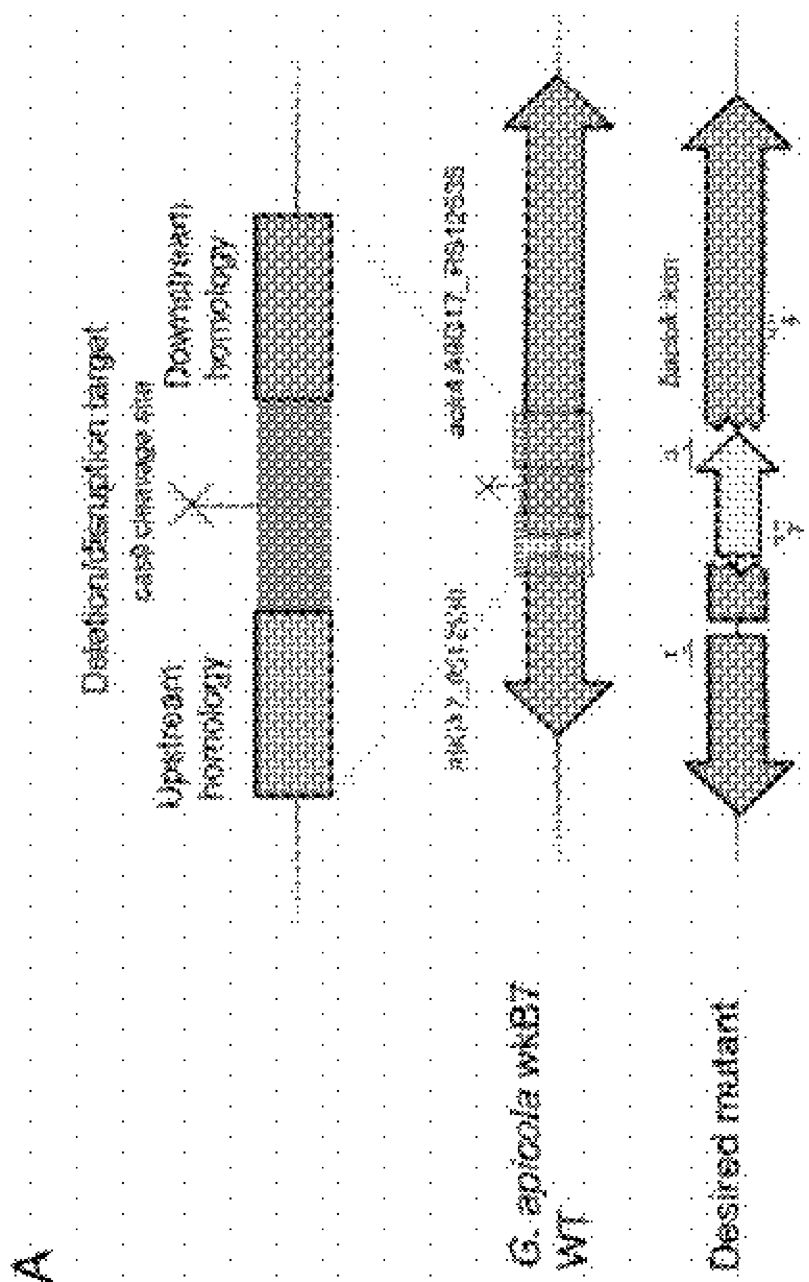



Figure 25A

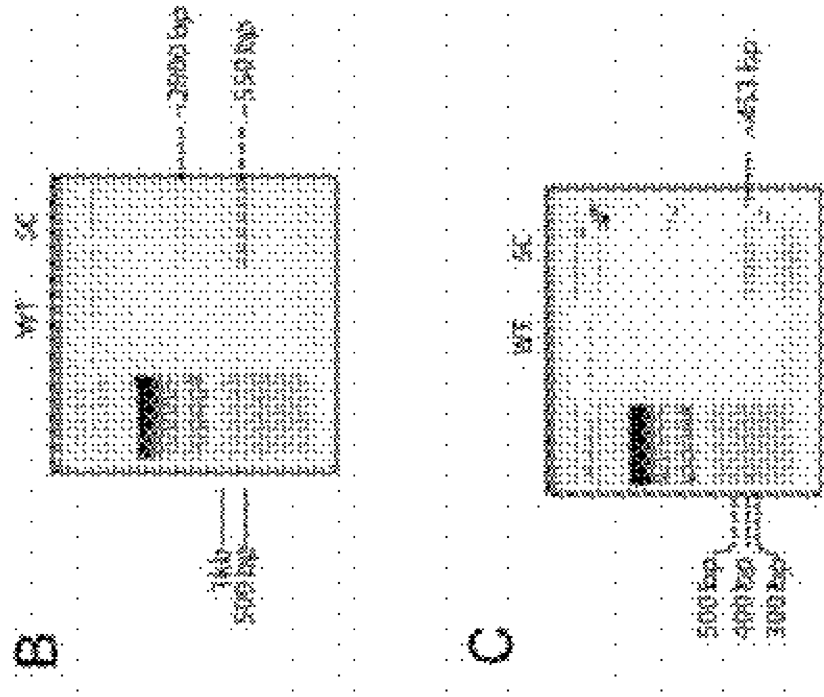



Figure 25B- Figure 25C

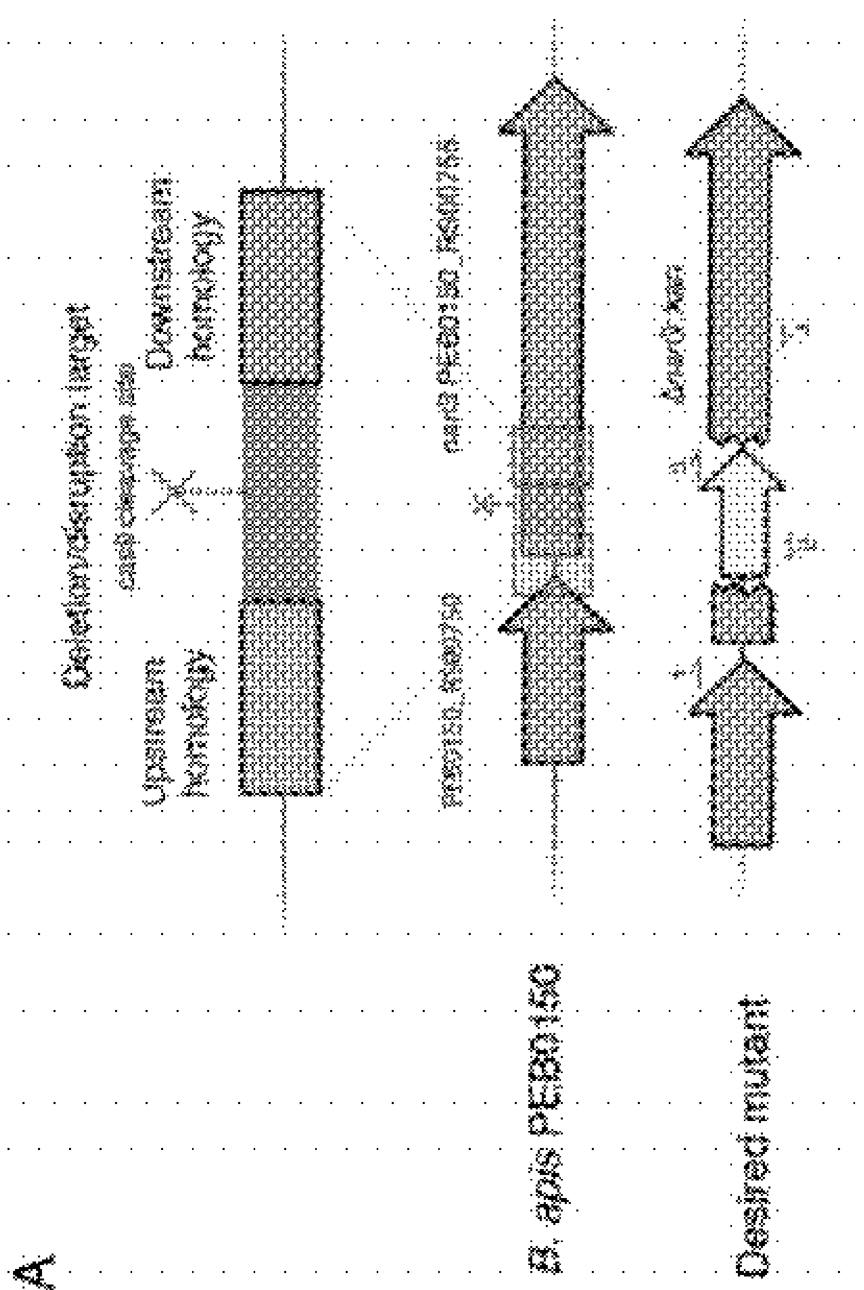



Figure 26A

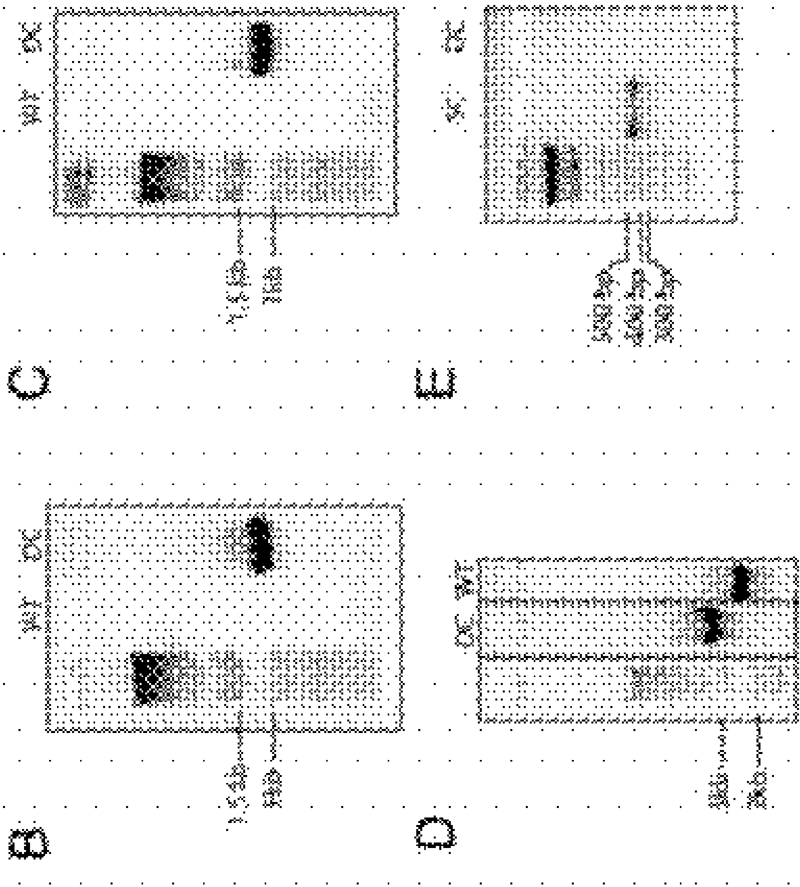


Figure 26B- Figure 26E

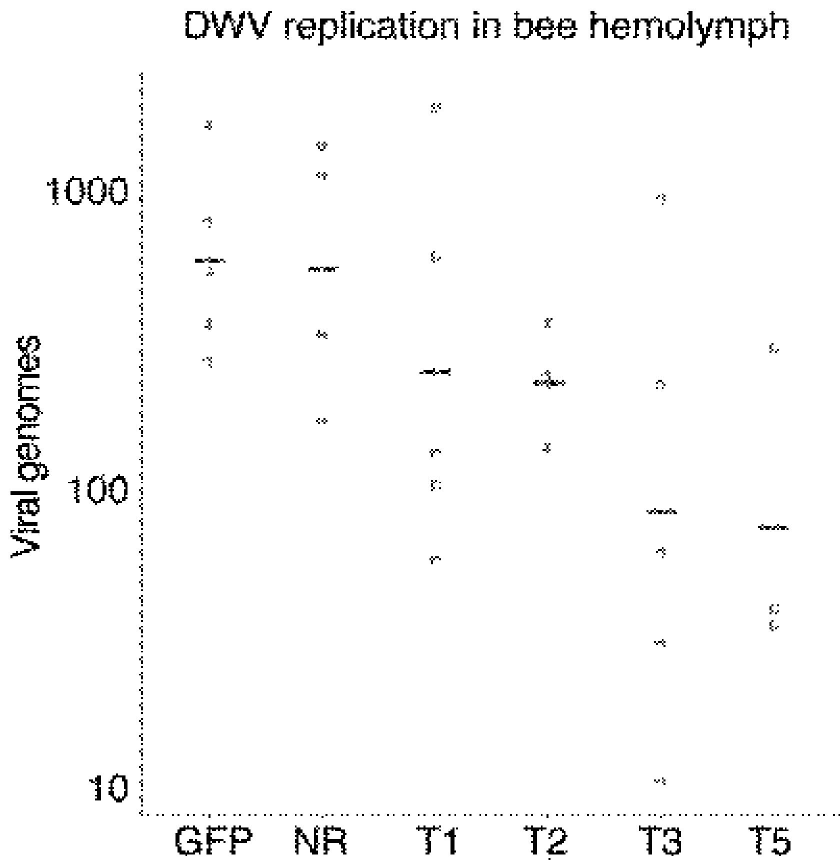



Figure 27

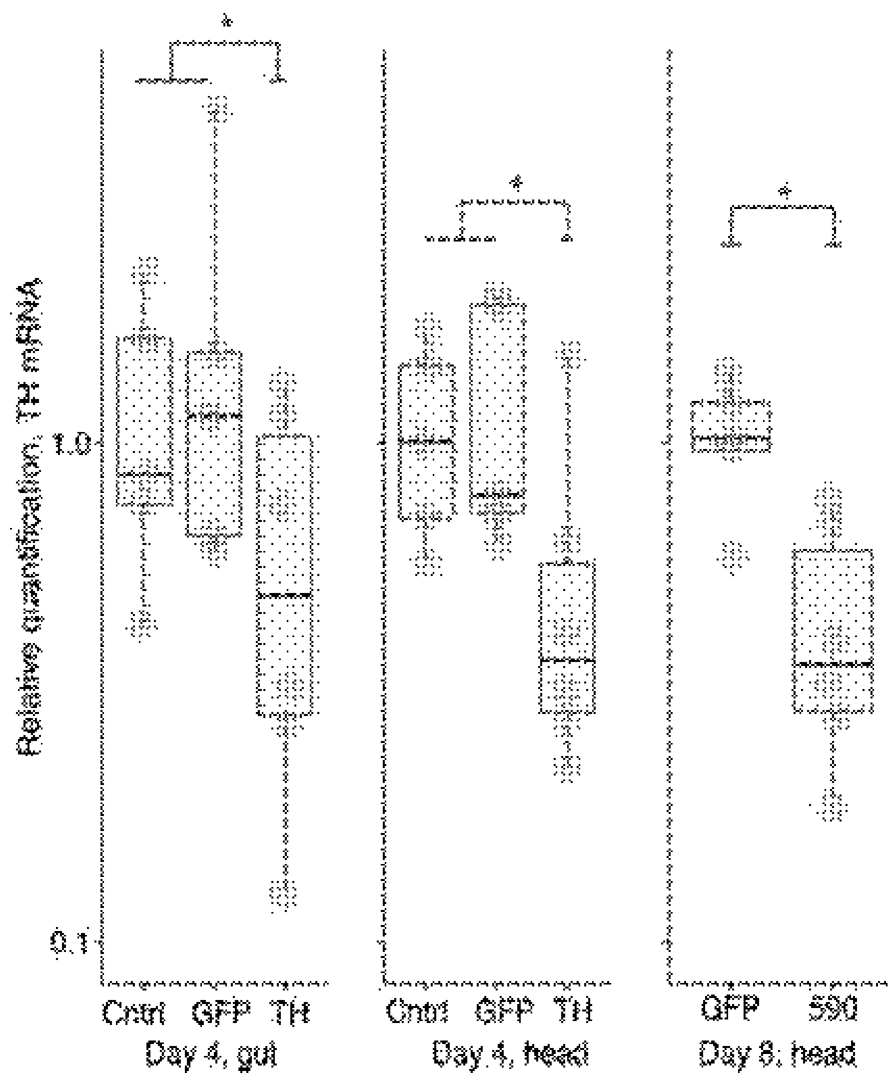



Figure 28

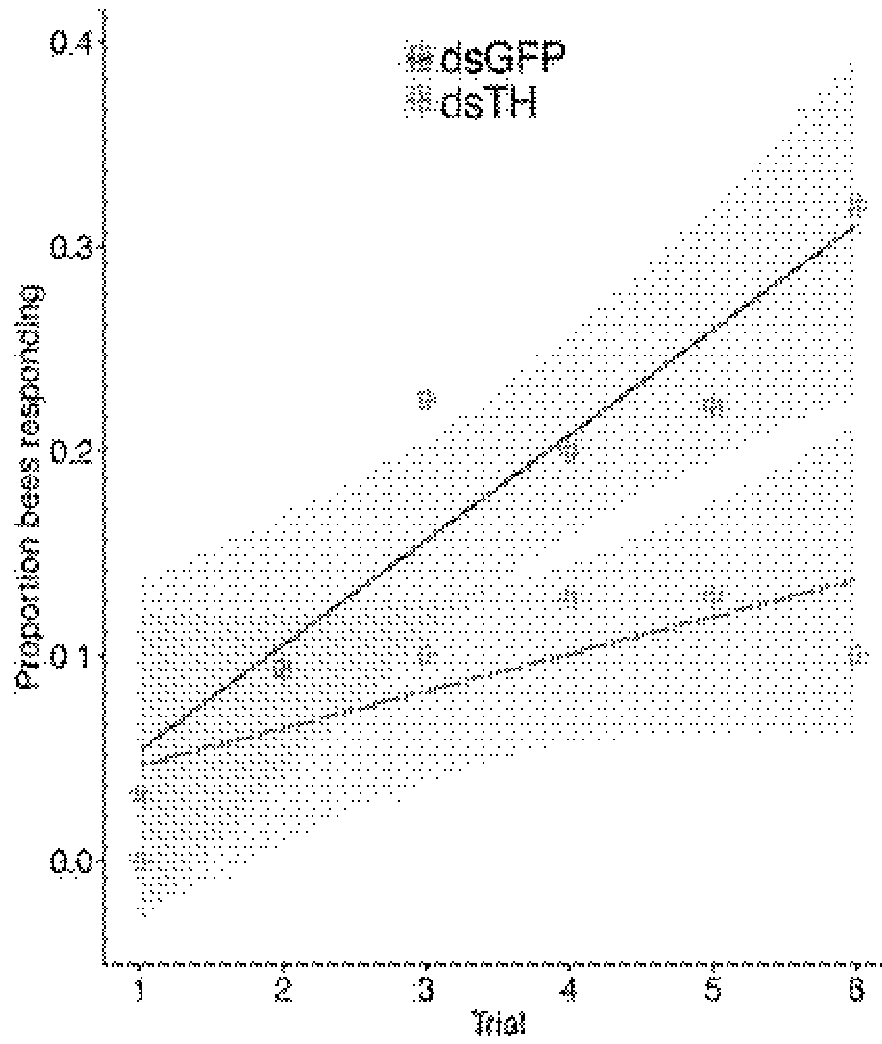



Figure 29

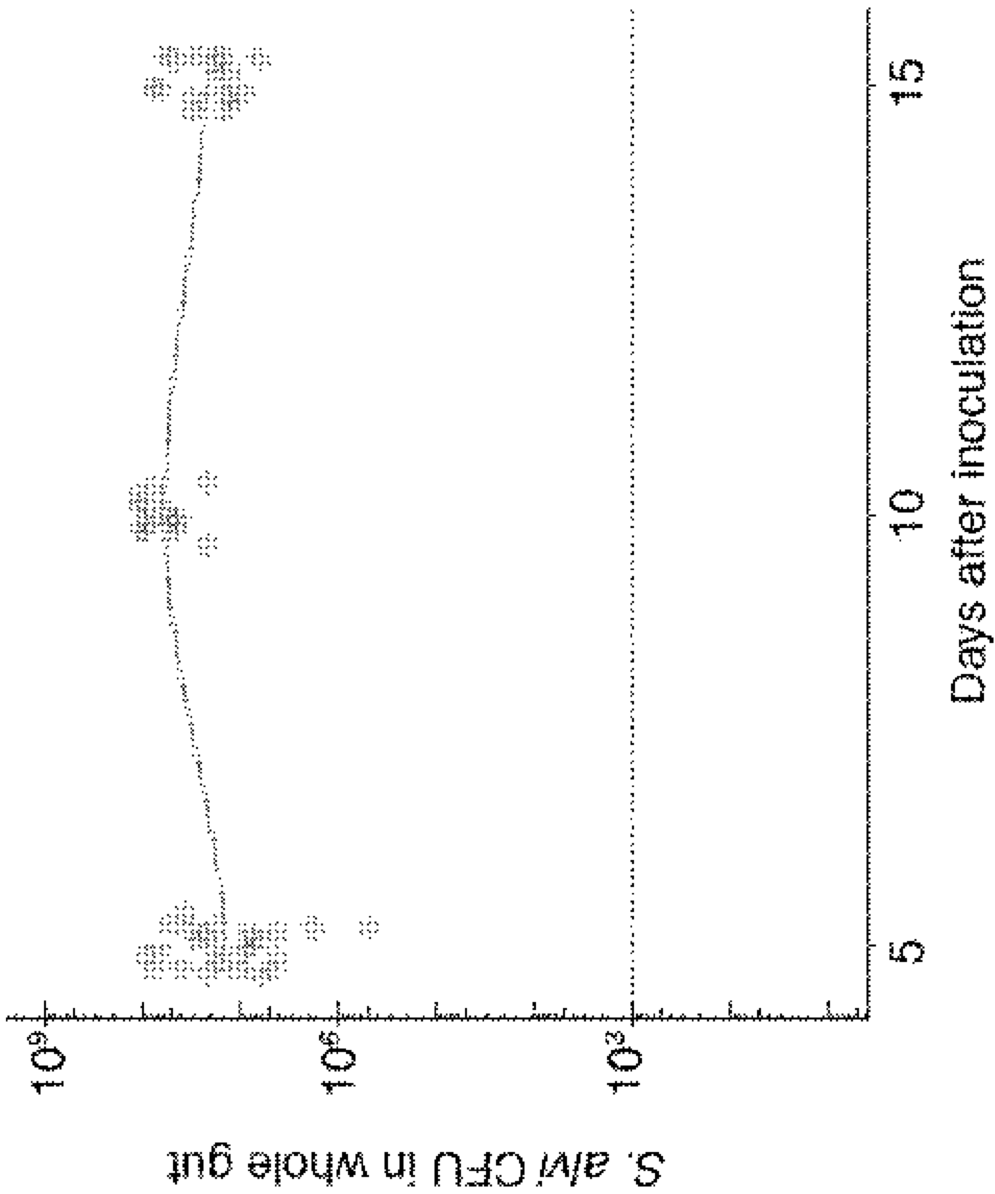



Figure 30

# ENGINEERED MICROBIAL POPULATION

## CROSS REFERENCE TO RELATED APPLICATIONS

**[0001]** This application claims priority to U.S. Provisional Application No. 62/529,754, filed Jul. 7, 2017 which is hereby incorporated by reference herein in its entirety.

## STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

**[0002]** This invention was made with government support under Grant No. HR0011-15-C-0095 awarded by the Defense Advanced Research Projects Agency. The government has certain rights in the invention.

## FIELD OF THE INVENTION

**[0003]** The present invention relates generally to the field of molecular biology. More particularly, it concerns microbiome engineering.

## BACKGROUND OF THE INVENTION

**[0004]** Most animals have symbiotic microorganisms living in their bodies. Most often these are bacteria that are specialized to live in particular animal species. Honey bees (*Apis mellifera*) are important agricultural pollinators. Unfortunately, recent years have seen substantial bee colony losses (e.g., Colony Collapse Disorder), due to a myriad of complex causes. Some of the most significant causes are bee viral pathogens and *Varroa* mite infestations.

**[0005]** RNA interference (RNAi) is a powerful technique to specifically downregulate gene expression in many eukaryotes by providing dsRNA with sequence identity to eukaryotic genes. RNAi has been shown to function in honey bees previously, and has been used to knock down bee genes to investigate their function and also improve bee health by lowering pathogen burden (e.g., viral and *varroa* mite). Injection or feeding of double stranded RNA (dsRNA) has been shown to trigger RNAi in honey bees, leading to molecular cascades that can degrade invading viral RNA and kill pests (e.g., *Varroa* mites). However, production of large quantities of dsRNA in the laboratory is expensive and the dsRNA itself is unstable and rapidly degrades in the environment. Additionally, injecting dsRNA into bees is traumatic and feeding them dsRNA is unreliable. Thus, there is an unmet need for improved methods of inducing RNAi in bees.

## SUMMARY OF THE INVENTION

**[0006]** In one embodiment, the invention relates to a microbial composition comprising one or more bacteria genetically engineered to express at least one heterologous nucleic acid, wherein the one or more bacteria are native to the microbiome of a host insect.

**[0007]** In one embodiment, the host insect is selected from the group consisting of a honey bee and a bumble bee.

**[0008]** In one embodiment, the one or more bacteria is *Snodgrassella alvi, Bartonella apis, Gilliamella apicola, Serratia marcescens, Parasaccharibacter apium*, or *Lactobacillus* sp.

**[0009]** In one embodiment, the composition comprises 2, 3, 4, or 5 bacterial species.

**[0010]** In one embodiment, the one or more bacteria express at least two heterologous nucleic acids.

**[0011]** In one embodiment, the heterologous nucleic acid encodes a polypeptide that improves the health of a host insect.

**[0012]** In one embodiment, the heterologous nucleic acid encodes a pesticide degrading polypeptide or a cytochrome.

**[0013]** In one embodiment, the heterologous nucleic acid is an inhibitory nucleic acid. In one embodiment, the inhibitory nucleic acid is selected from the group consisting of an antisense DNA, dsRNA, siRNA, shRNA, sgRNA and a miRNA.

**[0014]** In one embodiment, the heterologous nucleic acid is incorporated into a broad host range plasmid.

**[0015]** In one embodiment, the broad host range plasmid comprises at least one regulatory sequence selected from the group consisting of an RSF1010 origin of replication, a PA1 promoter sequence, a PA2 promoter sequence, a PA3 promoter sequence, a cp25 promoter sequence, and a detectable marker.

**[0016]** In one embodiment, the composition is a bee-ingestible composition.

**[0017]** In one embodiment, the bacteria are present as a live suspension. In one embodiment, the bacteria are present as a lyophilized powder. In one embodiment, the composition is in solid form. In one embodiment, the composition is in liquid form. In one embodiment, the composition comprises protein. In one embodiment, the composition comprises pollen. In one embodiment, the composition is a sucrose solution. In one embodiment, the composition is a corn syrup solution. In one embodiment, the composition comprises a carbohydrate or sugar supplement.

**[0018]** In one embodiment, the invention relates to an insect comprising a microbial composition comprising one or more bacteria genetically engineered to express at least one heterologous nucleic acid, wherein the one or more bacteria are native to the microbiome of a host insect.

**[0019]** In one embodiment, the insect is a bee, a honey bee, a forager, a hive bee, a pupae, an adult bee, and a bee colony parasite.

**[0020]** In one embodiment, the invention relates to a method for producing a microbial composition comprising one or more bacteria genetically engineered to express at least one heterologous nucleic acid, wherein the one or more bacteria are native to the microbiome of a host insect, comprising transfecting said bacterial species with an expression cassette comprising at least one heterologous nucleic acid.

**[0021]** In one embodiment, the invention relates to a method for downregulating expression of a target gene product, comprising administering an effective amount of a microbial composition comprising one or more bacteria genetically engineered to express at least one heterologous nucleic acid, wherein the one or more bacteria are native to the microbiome of a host insect to the host insect, wherein said bacteria express an inhibitor of said target gene product.

**[0022]** In one embodiment, the target gene product is a gene from an organism selected from the group consisting of a pathogen, a parasite, a virus, a mite, Acute Bee Paralysis Virus (ABPV), Kashmir Bee Virus (KBV), Israeli Acute Paralysis Virus (IAPV), *Nosema ceranae*, Deformed Wing Virus, and *Varroa destructor* mite.

**[0023]** In one embodiment, the insect is selected from the group consisting of a bee, a honey bee, a forager, a hive bee, a pupae, an adult bee, and a bee colony parasite.

**[0024]** In one embodiment, the target gene is selected from the group consisting of TOM70, TIM22, TOM40, Imp2, mitochondrial Hsp70, ATM1-ABC transporter proteins, Frataxin, Ferredoxin, ERV1, ferredoxin, NADPH oxidoreductase [FNR], pyruvate dehydrogenase α subunit, pyruvate dehydrogenase β subunit, mitochondrial glycerol-3-phosphate dehydrogenase (mtG3PDH), manganese-containing superoxide dismutase (MnSOD), DNAJ (Hsp70 interacting), Iron Sulfur cluster ISU1, Cystein desulfurase Nsf1, NAR1, RLI1, ATPase subunit A, RNA polymerase I, RNA polymerase III, Inhibitor of apoptosis (IAP), and FAS apoptotic.

**[0025]** In one embodiment, the invention relates to a method for modulating expression of a target gene product in an insect, comprising administering an effective amount of a microbial composition comprising one or more bacteria genetically engineered to express at least one heterologous nucleic acid to said insect

**[0026]** In one embodiment, the insect is selected from the group consisting of a bee, a honey bee, a forager, a hive bee, a pupae, an adult bee, and a bee colony parasite.

**[0027]** In one embodiment, the invention relates to a method for reducing the susceptibility of a bee to a disease or disorder selected from the group consisting of Colony Collapse Disorder (CCD) and infection, comprising administering an effective amount of a microbial composition comprising one or more bacteria genetically engineered to express at least one heterologous nucleic acid, wherein the one or more bacteria are native to the microbiome of a host insect, to said bee, wherein said bacteria express an inhibitor of a pathogen or parasite specific gene product.

**[0028]** In one embodiment, the pathogen or parasite is selected from the group consisting of Acute Bee Paralysis Virus (ABPV), Kashmir Bee Virus (KBV), Israeli Acute Paralysis Virus (IAPV), *Nosema ceranae*, *Nosema ceranu*, *Nosema apis*, Deformed Wing Virus, and *Varroa destructor* mite.

**[0029]** In one embodiment, the bacteria express at least two non-contiguous dsRNAs downregulating expression of a pathogen or parasite specific gene product.

**[0030]** In one embodiment, the invention relates to a method for reducing the susceptibility of a bee colony to infestation by the Small Hive Beetle, comprising administering an effective amount of a microbial composition comprising one or more bacteria genetically engineered to express at least one heterologous nucleic acid to hive components, wherein said bacteria express an inhibitor of a Small Hive Beetle specific gene product.

**[0031]** In one embodiment, the bacteria express at least two non-contiguous dsRNAs downregulating expression of a Small Hive Beetle specific gene product.

**[0032]** In one embodiment, the invention relates to a method for expression of a heterologous nucleic acid sequence in a bee, the method comprising administering to the bee at least one modified *Snodgrassella alvi* bacterium comprising an expression plasmid for expression of the heterologous nucleic acid sequence

**[0033]** In one embodiment, the heterologous nucleic acid sequence encodes a molecule selected from the group consisting of a protein, a peptide, an inhibitory RNA, a dsRNA, a siRNA, a shRNA, a sgRNA and a miRNA.

## BRIEF DESCRIPTION OF THE DRAWINGS

**[0034]** The following drawings form part of the present specification and are included to further demonstrate certain aspects of the present invention. The invention may be better understood by reference to one or more of these drawings in combination with the detailed description of specific embodiments presented herein.

**[0035]** FIG. **1**: In vitro fluorescent protein expression in bee gut microbiome bacteria.

**[0036]** FIGS. **2A-2C**: Fluorescent imaging of engineered bee gut bacteria colonizing the bee gut microbiome.

**[0037]** FIG. **3**: Schematic of dsRNA-expressing broad-host-range construct.

**[0038]** FIG. **4**: In vitro production of dsRNA by engineered *S. alvi*.

**[0039]** FIG. **5**: In vivo production of dsRNA in the bee gut.

**[0040]** FIG. **6**. dsRNA against essential honey bee gene alpha-tubulin increases mortality.

**[0041]** FIG. **7**: Quantification of tyrosine hydrolase mRNA in the head of bees.

**[0042]** FIGS. **8A-8D**: Schematics depicting various plasmids including pBTK520, pBTK570, pBTK562, and pBTK561

**[0043]** FIGS. **9A-9C**. Depicts design of the bee microbiome toolkit (BTK) and schematic assembly. FIG. **9A** depicts the BTK designed for Golden Gate assembly according to a scheme with eight-part types compatible with the yeast toolkit (YTK) (Lee M E et al., 2015, ACS Synth Biol., 4:975-986). Parts of each type generated in this study are shown in the top panel. Type 1-5 and Type 8 parts are defined as in the YTK except that Type 3 open reading frames include the stop codon. Type 6 and 7 parts are either replaced with a linker part or used to incorporate a reverse reading frame encoding a transcriptional regulator for inducible expression of the main Type 3 open-reading frame and its promoter, respectively, during Stage 1 assembly, so that costly or toxic genes can be repressed while they are assembled into transcriptional units. FIG. **9B** Schematic of Stage 1 (BsaI) assembly. Plasmid parts are shown, but PCR products with appropriate overhangs can be substituted. FIG. **9c** Schematic Stage 2 (BsmBI) assembly. Compatible Stage 2 connectors are described in the YTK documentation.

**[0044]** FIGS. **10A-10B**: Depicts the bee microbiome toolkit (BTK) functions in diverse bee-associated bacteria. FIG. **10A** depicts replication of the BTK backbone and the function of three antibiotic resistance cassettes were tested in eight honey bee-associated bacterial strains as described in the Methods. At least one antibiotic resistance cassette functioned in each strain, and the kanamycin cassette functioned in all eight strains. FIG. **10B** depicts replication of the BTK backbone and the function of three antibiotic resistance cassettes were tested in three bumble bee-associated bacterial strains. Again, the plasmid with kanamycin resistance was maintained in all three bacteria. FIG. **10C** depicts conjugation frequency in four bee gut-associated strains. Black bars are the geometric mean and each point is an independent conjugation. Conjugation in *B. apis* is the most efficient, with conjugation efficiency approximately 10%

**[0045]** FIGS. **11A-11C**: Depicts constitutive and inducible control of in vitro gene expression in bee gut bacteria. FIG. **11A** depicts flow cytometry results of GFP fluorescence from four broad-host-range promoters in each of four honey bee-associated bacterial strains and an *E. coli* control. One representative fluorescence distribution for each promoter is

shown, with the medians from three biological replicates plotted as open circles. Spotted grey line indicates maximum detected fluorescence in wildtype cells. Median fluorescent values were calculated from cells more fluorescent than wildtype. FIG. **11B** depicts GFP fluorescence from a designed CP25 (lacO) promoter at different levels of IPTG-induction, measured in four BGM strains and an *E. coli* control. All tested species are responsive to IPTG induction, and *G. apicola* shows the highest expression across all strains. FIG. 11C depicts GFP fluorescence from a T7 (lacO) promoter at different levels of IPTG-induction of T7 RNAP expression in the same four BGM strains. Schematics in FIG. **11A** through FIG. **11C** show the design of tested constructs using Synthetic Biology Open Language (SBOL) standard glyphs (Galdzicki M et al., 2014. Nat Biotechnol., 32:545-550). Error bars are standard deviations (n=3).

**[0046]** FIGS. **12A-12B**: Depicts dCas9 gene silencing in *Bartonella apis*. FIG. **12A** depicts schematic assembly of dCas9 plasmids for gene suppression. FIG. **12B** depicts fluorescence from chromosomally integrated GFP in PEB0150 in the presence and absence of dCas9 and sgRNA targeting GFP. Background fluorescence of wild-type PEB0150 was subtracted. GFP fluorescence decreased in presence of dCas9 targeting GFP (p=0.004, Kruskal-Wallis rank sum test). Error bars are 95% confidence intervals (n=4).

**[0047]** FIGS. **13A-13E**: Depicts Cas9 assisted gene disruption in species from the bee gut microbiota. FIG. **13A** depicts schematic assembly of R6K-based suicide plasmids. Assembly strategy and validation primers are described in FIG. **23**. FIG. **13B** depicts the two tested approaches for gene disruption. The suicide plasmids were introduced into either wild-type bacteria or bacteria possessing the constitutively active Cas9 (pBTK601). FIG. 13C depicts transconjugation frequency and percent of desired mutants in *B. apis*, in the presence and absence of Cas9. The Cas9 plasmid did not increase the efficiency of genome modification. Numbers above each bar indicate the number of clones evaluated. FIG. **13D** depicts transconjugation frequency and proportion of desired mutants in *S. alvi*. *S. alvi* wkB2 showed increased efficiency of genome modification in the presence of the Cas9 plasmid (p=0.0007, Kruskal-Wallis rank sum test). FIG. **13E** depicts transconjugation frequency and proportion of desired mutants in *G. apicola*. Each point in FIG. **13C** and FIG. **13E** is from an independent conjugation experiment. Bars in FIG. **13C** and FIG. **13E** represent the geometric mean of estimated transconjugation efficiencies.

**[0048]** FIGS. 14A-**14E**: Depicts visualization of engineered bacteria in the honey bee gut. FIG. **14A** depicts intact honey bee worker and dissection of honey bee gut showing brightfield microscopy of midgut, ileum, and rectum. FIG. **14B** depicts fluorescent imaging of whole bee (left) and dissected bee (right) 5 days after inoculation with *S. marcescens* N10A28 expressing E2-Crimson (plasmid pBTK570). Control bee is uninoculated. Color corresponds to pixel fluorescence intensity. Engineered *S. marcescens* N10A28 is present in the midgut, ileum, and rectum. FIG. **14C** Similar to FIG. **14B**, with *S. alvi* wkB2 expressing E2-Crimson as inoculum. Control bee is identical to FIG. **14B**, but different fluorescent intensity scales are used for comparison between bees inoculated with *S. alvi* and *S. marcescens*. Engineered *S. alvi* wkB2 is visibly fluorescent in the midgut and ileum. FIG. 14D depicts confocal imaging of partial ileum and rectum in bees inoculated with *S. marcescens* N10A28 expressing E2-Crimson (red). As in FIG. **14B**, *S. marcescens* can be seen robustly colonizing throughout the ileum and rectum. FIG. 14E Similar to FIG. 14D, with *S. alvi* wkB2 expressing E2-Crimson (green). *Snodgrassella alvi* wkB2 colonizes the ileum, but not the rectum. Scale bars in FIG. **14D** and FIG. **14E** are 100 m. Images are representative of multiple bees inspected (n=3-5 per condition) for (FIG. **14B** through FIG. **14E**). White and black arrows correspond to the ileum-rectum junction across images (FIG. 14A through FIG. **14E**).

**[0049]** FIGS. **15A-15D**: Depicts visible co-inoculation of the bee gut with species from the bee gut microbiota and the role of staA in colonization. FIG. **15A** depicts the ileum-rectum junction imaged by confocal fluorescence microscopy 5 days after co-inoculating *B. apis* PEB0150 and *S. alvi* wkB2. When co-inoculated, *B. apis* and *S. alvi* are co-located in the ileum, but only *B. apis* colonizes the rectum. FIG. **15B**, similar to FIG. 15A, depicts images taken 5 days after co-inoculation of *G. apicola* wkB7 and *S. alvi* wkB2. As in FIG. **15A**, *S. alvi* remains restricted to the ileum, while *G. apicola* is present in both ileum and rectum. Scale bars are 100 m. Images are representative of multiple bees inspected (n=3 per condition). FIG. **15C** depicts the number of *S. alvi* 16S ribosomal DNA copies 5 days after inoculating newly emerged worker bees with the *S. alvi* WT or ΔstaA mutant, based on quantitative PCR. Horizontal bars represent means per condition (n=5). The ΔstaA has a significant colonization defect compared to WT (p=2.7×10-6, Kruskall-Wallis rank sum test). FIG. **15D** depicts the ileums of bees inoculated with *S. alvi* wkB2ΔstaA expressing E2-Crimson (pBTK570) were imaged 5 days after colonization. Mutants achieved lower colonization levels than did *S. alvi* WT (see FIG. **6E**). Localized colonization was typical of multiple ileums inspected (n=3). Scale bar is 100 μm.

**[0050]** FIG. **16** depicts broad-host-range plasmid screen in bee gut bacteria.

**[0051]** FIG. 17 depicts the bacterial strains used in this study.

**[0052]** FIG. **18** depicts BTK plasmids.

**[0053]** FIG. **19** depicts oligonucleotides used in the study.

**[0054]** FIG. **20** depicts expression of nanoluc in BGM strains. Bee gut microbiota strains were conjugated with pBTK563 (CP25-Nanoluc) or pBTK564 (PA3-Nanoluc) and luciferase activity measured per manufacturer's instructions. Luciferase is plotted on a log 10 scale. Bars are the mean from three biological replicates and error bars are standard deviation.

**[0055]** FIG. **21** depicts expression of E2-Crimson in BGM strains. Bee gut microbiota strains with pBTK569 (CP25-E2-Crimson) or pBTK570 (PA3-E2-Crimson) were measured for expression of E2-Crimson. Bars are the mean from three biological replicates and error bars are standard deviation. E2-Crimson fluorescence expression was measured as GFP except using excitation/emission wavelengths of 611/646 nm.

**[0056]** FIG. **22** depicts validation of additional BTK promoters in *E. coli*. Promoter parts CP12b, CP32, and CP6 were used to build GFP expression plasmids (pBTK619, pBTK620, pBTK621) and tested in *E. coli* MFDpir. Bars are the mean from three biological replicates and error bars are standard deviation.

**[0057]** FIGS. **23A-23C**: Depicts construction of replacement cassette plasmids. FIG. **23A** depicts schematic assem-

bly of homology donor plasmids using BTK Golden Gate assembly. Type 2, 3, and 4 overhangs are repurposed for the upstream homology, antibiotic resistance, and downstream homology respectively. Combined with parts Type 1, 5, and pYTK095 (Type 6-8), the donor plasmid is then ready for further BsmBI assembly as shown in FIG. **12**. FIG. **23B** depicts possible outcomes from homologous recombination crossing-over events. Single-crossover mutants are more common and occur with integration of the suicide plasmid backbone. Double-crossover mutants are the desired mutant, where only the antibiotic resistance cassette is retained in the chromosome of the target strain. FIG. **23C** depicts sets of PCR primers used to verify mutants.

**[0058]** FIGS. **24A-24E**: Depicts schematic and validation of staA disruption in *S. alvi*. FIG. **24A** depicts schematic of staA disruption in *S. alvi*. FIG. **24B** depicts DNA gel with products from the PCR of upstream junctions using primers 1-2. The double-crossover mutant (DC) shows a band at the expected size of 1583 bp, wild type (WT) shows no amplification. FIG. **24C** depicts DNA gel with products from the PCR of downstream junctions using primers 3-4. DC shows a band at the expected size of 1587 bp, WT shows no amplification. FIG. **24D** depicts DNA gel with products from the PCR of the entire region using primers 1-4. DC shows an increase in size of approximately 880 bp compared to WT. Lanes rearranged from original gel for clarity. FIG. **24E** depicts DNA gel with products from the PCR of plasmid backbone. A single crossover mutant (SC) shows expected amplification at 453 bp, DC shows no amplification. Black bars indicate image cropping for clarity.

**[0059]** FIGS. **25A-25C**: Depicts schematic and validation of ackA disruption in *G. apicola*. FIG. **25A** depicts schematic of planned ackA disruption in *G. apicola*. FIG. **25B** depicts DNA gel showing products from the PCR of upstream junctions using primers 1-2. Forty clones were tested, and two distinct single cross-over events (SC) were observed: a smaller than expected band at ~550 bp and a second, larger band at ~2800 bp (often in the same clone). Neither of these is the expected size of 1501 bp. This may represent a duplication or other complex recombination event. No downstream junctions were effectively amplified using primers 3-4 (data not shown). PCR using primers 1-4 showed no change in size between wild-type (WT) and SC mutants, indicating the SC event likely did not disrupt ackA function (data not shown). FIG. **25C** depicts DNA gel showing product from the PCR using primers KO_5 and KO_6. Expected product at 453 bp shows retention of the suicide backbone in the SC mutants.

**[0060]** FIGS. **26A-26E**: Depicts schematic and validation of narG disruption in *B. apis*. FIG. **26A** depicts schematic of planned narG disruption in *B. apis*. FIG. **26B** depicts DNA gel with products from the PCR of upstream junctions using primers 1-2. The double-crossover mutant (DC) shows a band at the expected size of 1169 bp, wild type (WT) shows no amplification. FIG. **26C** depicts DNA gel with products from the PCR of downstream junctions using primers 3-4. DC shows a band at the expected size of 1309 bp, WT shows no amplification. FIG. **26D** depicts DNA gel with products from the PCR of the entire region using primers 1-4. DC shows an increase in size of approximately 757 bp compared to WT. Lanes rearranged from original gel for clarity. FIG. **26E** depicts DNA gel with products from the PCR of plasmid backbone. A single crossover mutant (SC) shows expected amplification at 453 bp. DC shows no amplification. Black bars indicate image cropping for clarity.

**[0061]** FIG. **27**: Depicts exemplary experimental results of DWV replication in bee hemolymph.

**[0062]** FIG. **28**: Depicts exemplary experimental results demonstrating the relative quantification of TH mRNA.

**[0063]** FIG. **29**: Depicts exemplary experimental results demonstrating the proportion of bees responding to dsTH or dsGFP.

**[0064]** FIG. **30**: Depicts exemplary experimental results demonstrating the proportion of *S. alvi* CFU persisting in the gut of inoculated bees at different times.

## DETAILED DESCRIPTION

**[0065]** In some embodiments, the present disclosure provides methods and compositions concerning microbiome engineering, such as in the gut microbiome of arthropods. In certain embodiments, one or more species of bacteria that are naturally present in the arthropod gut are genetically engineered. These methods may be used to improve arthropod health in a variety of ways. For example, these engineered bacterial species may then be introduced to an insect population, such as a bee population, to directly produce one or more compounds in the insect gut. In some instances, the engineered bacterial species are used to directly express double-stranded RNA (dsRNA) in an insect. In some embodiments, the microbe-produced dsRNA induces RNA interference (RNAi) against a host insect gene product. In some embodiments, the microbe-produced dsRNA induces RNA interference (RNAi) against a gene product of a pathogen or parasite of a host insect.

**[0066]** A broad host range plasmid (e.g., comprising a RSF 1010 origin of replication, such as from the pMMB67EH plasmid) may be used to introduce genes into one or more arthropod gut microbiome strains including, but not limited to, *Snodgrassella alvi*, *Bartonella apis*, *Gilliamella apicola*, *Serratia* sp., *Parasaccharibacter apium*, and *Lactobacillus* sp.

**[0067]** In one embodiment, the invention relates to the use of paired, convergent transcriptional promoters on a plasmid to generate dsRNA constructs against any target gene product or series of targets. In this embodiment, no purification of the dsRNA is required, as the genetically engineered bacteria continually produce the desired dsRNA and the dsRNA is directly taken up through the gut of the arthropod. Because these are modified natural arthropod gut microbiome species, they can be maintained over the arthropod's lifetime. In one embodiment, they may be naturally transferred to newly emerged arthropod (e.g., newly emerged bees in a hive) from adults. Thus, the present methods may reduce the need for frequent and costly reapplication.

**[0068]** The present studies showed the production of dsRNA directly in the bee gut by engineering natural members of their microbiota. By producing dsRNA directly in the honey bee gut, the present methods can provide a continual and cost-effective supply of dsRNA to silence bee or beepest genes. In one example, the efficacy of the present methods was demonstrated by using *Snodgrassella alvi* to produce a dsRNA that targets an essential honey bee cytoskeletal gene. Application of this engineered bacterium to bees resulted in increased bee mortality compared to one encoding an off-target dsRNA control. Further aspects may comprise additional methods to increase dsRNA expression,

such as knock out of bacterial RNaseII or RNase III or using T7 RNA polymerase to produce dsRNA.

**[0069]** Thus, in one embodiment, the present methods may be used to produce and administer dsRNA by genetically engineering bacteria that are native to the bee gut microbiome. The methods may be useful in reducing bee mortality from RNA viruses (e.g., Israeli acute paralysis virus, Deformed wing virus) and for *Varroa* mite and Small Hive Beetle control as well as other microbial and insect pathogens. As bacterial production can be inexpensive, the present methods be useful to small and large scale apiaries. Further embodiments may concern bumble bee production, an industry that is important in providing pollination for many fruit and vegetable crops.

**[0070]** In some cases, Colony Collapse Disorder (CCD) of honeybees can be due to *Varroa* mite infections. *Varroa* mites are suspected of acting as vectors for a number of honey bee pathogens, including Deformed Wing Virus (DWV), Kashmir Bee Virus (KBV), Acute Bee Paralysis Virus (ABPV) and Black Queen Cell Virus (BQCV), and may weaken the immune systems of their hosts, leaving them vulnerable to infections. Accordingly, certain embodiments of the present disclosure provide methods of reducing the susceptibility of honeybees to Colony Collapse Disorder (CCD). Thus, in some aspects, the dsRNA of the present methods and compositions may target a gene product of picorna-like viruses (e.g., Acute Bee Paralysis Virus (ABPV), Kashmir Bee Virus (KBV), and Israeli Acute Paralysis Virus (IAPV)), *Nosema* parasite, Deformed Wing Virus, and/or a *Varroa destructor* mite, or Small Hive Beetle. The inhibitory nucleic acid molecules (e.g., dsRNA) may be complementary to the mRNA of the target gene product. For example, the inhibitory nucleic acid molecule can be complementary to region of a mRNA comprising at least 10, 15, 20, 21, 22, 23, 24, or 25 contiguous nucleotides. In some aspects, the inhibitory nucleic acid molecule can be complementary to region of a mRNA comprising at least 50, 100, 150, 200, 300 400 or 500 contiguous nucleotides or essentially the entire mRNA. In some aspects a dsRNA comprise at least 10, 15, 20, 21, 22, 23, 24, or 25 complementary nucleotides. For example, the dsRNA produced by the engineered bacteria may comprise a sequence complementary to *Varroa destructor* mite mRNA of and capable of inducing degradation of the *Varroa destructor*-specific mRNA. The target mRNA sequences for use in the present methods include those described in, but are not limited to, U.S. Patent Publication No. 20140371298 or 20150133532; both incorporated herein by reference in their entirety.

**[0071]** In some embodiments, there is provided an expression cassette encoding a dsRNA under the control of one or more broad host range promoters, wherein the dsRNA is complementary to a target gene product of Acute Bee Paralysis Virus (ABPV), Kashmir Bee Virus (KBV), Israeli Acute Paralysis Virus (IAPV), *Nosema* parasite, Deformed Wing Virus, *Varroa destructor* mite, or Small Hive Beetle. The expression cassette may encode for 2, 3, 4, 5, 6, 7, 8, 9, 10 or more inhibitory nucleic acids, such as dsRNAs. In some aspects, the inhibitory nucleic acids may target the same virus, parasite or mite. In other aspects, the inhibitory nucleic acids target mRNAs encoding different gene products, such as a dsRNA to a gene product of the *Varroa destructor* mite.

**[0072]** In further embodiments, there are provided microbial compositions comprising one or more insect microbiome bacteria engineered to express one or more heterologous nucleic acids, such as inhibitory nucleic acids. The inhibitory nucleic acid may be dsRNA, siRNA, shRNA, and/or miRNA. The dsRNAs may also be precursor miRNAs. In some cases, the inhibitory nucleic acids may be at least 15 base pairs in length, such as at 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 base pairs in length. In further aspects, the inhibitory nucleic acids may be complementary to 50, 100 or more based pairs of a target mRNA. Exemplary insect microbiome species include, but are not limited to, *Snodgrassella alvi*, *Bartonella apis*, *Gilliamella apicola*, *Serratia marcescens*, *Parasaccharibacter apium* and *Lactobacillus* sp. The microbial composition may comprise 2, 3, 4, or 5 bacterial species. Each engineered bacterium may express 1, 2, 3, 4, 5, or more heterologous nucleic acids. The nucleic acids may be pathogen, pest or parasite-specific, such as parasites which may infect insects, particularly bees. Exemplary pathogens, pests and parasites include, but are not limited to, viruses, microbial pathogens, Acute Bee Paralysis Virus (ABPV), Kashmir Bee Virus (KBV), Israeli Acute Paralysis Virus (IAPV), Deformed Wing Virus, parasites such as *Nosema ceranae*, mites such as the *Varroa destructor* mite and pests such as the Small Hive Beetle (*Aethina tumida*).

**[0073]** The microbial composition may be formulated for insect ingestion, particularly as an ingestible composition (e.g., oral formulations). The composition may be solid or liquid (e.g. a sucrose solution or corn syrup solution). The composition may comprise additives or excipients, such as a carbohydrate or sugar supplement.

**[0074]** The microbial composition may be formulated for application to an insect pest food source, for example bee hive components.

**[0075]** The engineered bacteria or microbial compositions provided herein may be used to reduce insect susceptibility to infections, such as viral infections. They may also be used to down-regulated expression of a target gene product, such as by inducing RNAi in the insect.

**[0076]** In addition, the same bacterial species including *S. alvi* and *G. apicola* live in both honey bees (species in the genus *Apis*) and bumble bees (species in the genus *Bombus*), and the broad host range plasmid replicated in diverse bacterial hosts. Thus, the present methods may be applied to other bees or other species of insects, such as to induce similar RNAi effects. Additionally, the present methods allow for easy targeting of bee-specific genes to silence, and thus may be used for research.

Definitions

**[0077]** Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which the invention pertains. Although any methods and materials similar or equivalent to those described herein can be used in the practice for testing of the present invention, the preferred materials and methods are described herein. In describing and claiming the present invention, the following terminology will be used.

**[0078]** It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting.

**[0079]** As used herein the specification, "a" or "an" may mean one or more. As used herein in the claim(s), when used

in conjunction with the word "comprising," the words "a" or "an" may mean one or more than one.

**[0080]** The use of the term "or" in the claims is used to mean "and/or" unless explicitly indicated to refer to alternatives only or the alternatives are mutually exclusive, although the disclosure supports a definition that refers to only alternatives and "and/or." As used herein "another" may mean at least a second or more.

**[0081]** Throughout this application, the term "about" is used to indicate that a value includes the inherent variation of error for the device, the method being employed to determine the value, or the variation that exists among the study subjects.

**[0082]** As used herein, "essentially free," in terms of a specified component, is used herein to mean that none of the specified component has been purposefully formulated into a composition and/or is present only as a contaminant or in trace amounts. The total amount of the specified component resulting from any unintended contamination of a composition is therefore well below 0.01%. Most preferred is a composition in which no amount of the specified component can be detected with standard analytical methods. "Genetically engineered bacteria" refers to bacterial cells that replicate a heterologous nucleic acid, or express a polypeptide encoded by a heterologous nucleic acid.

**[0083]** "Heterologous nucleic acid" is one that originates from a source foreign to the particular host cell, or, if from the same source, is modified from its original form.

**[0084]** As used herein, an "instructional material" includes a publication, a recording, a diagram, or any other medium of expression which can be used to communicate the usefulness of a compound, composition, vector, or system of the invention in the kit. Optionally, or alternately, the instructional material can describe one or more methods of modulating expression of a gene product using a compound, composition, vector, or system of the invention in the kit. The instructional material of the kit of the invention can, for example, be affixed to a container which contains the identified compound, composition, vector, or delivery system of the invention or be shipped together with a container which contains the identified compound, composition, vector, or system. Alternatively, the instructional material can be shipped separately from the container with the intention that the instructional material and the kit be used cooperatively by the recipient.

**[0085]** "Promoter" is a nucleic acid sequence that acts as a signal sequence necessary to initiate transcription of a gene.

**[0086]** As used herein, the term "bee" is defined as any of several winged, hairy-bodied, usually stinging insects of the superfamily Apoidea in the order Hymenoptera, including both solitary and social species and characterized by sucking and chewing mouthparts for gathering nectar and pollen. Exemplary bee species include, but are not limited to species in the genera *Apis*, *Bombus*, *Trigona*, *Osmia* and the like. In one embodiment, bees include, but are not limited to bumblebees (*Bombus terrestris*, *Bombus impatiens*, or other *Bombus* species) and honeybees (*Apis mellifera* or *Apis cerana*).

**[0087]** As used herein, the term "colony" is defined as a population of dozens to typically several tens of thousand honeybees that cooperate in nest building, food collection, and brood rearing. A colony normally has a single queen, the remainder of the bees being either "workers" (females) or "drones" (males). The social structure of the colony is maintained by the queen and workers and depends on an effective system of communication. Division of labor within the worker caste primarily depends on the age of the bee but varies with the needs of the colony. Reproduction and colony strength depend on the queen, the quantity of food stores, and the size of the worker force. Honeybees can also be subdivided into the categories of "hive bees", usually for the first part of a workers lifetime, during which the "hive bee" performs tasks within the hive, and "forager bee", during the latter part of the bee's lifetime, during which the "forager" locates and collects pollen and nectar from outside the hive, and brings the nectar or pollen into the hive for consumption and storage.

**[0088]** As used herein, the term "susceptibility" is defined as the ability of a bee or bee colony to become infested or infected by and/or support proliferation of a pathogen, including, but not limited to, degree of infection, severity of symptoms, infectivity to other individuals (contagion), and the like. Susceptibility can be assessed, for example, by monitoring infectivity, presence of symptoms, such as, but not limited to, hunger, vitality, flight range, etc, presence of pathogenic organisms, mortality or time course of a disease in an individual bee or bee population following a challenge with the pathogen.

**[0089]** As used herein, the terms "bee disease" or "bee colony disease" are defined as undesirable changes in the behavior, physiology, morphology, reproductive fitness, economic value, viability, honey production, pollination capability, resistance to infection and/or infestation of a bee, a population of bees and/or a bee colony, directly or indirectly resulting from contact with a parasite or a parasite-infected bee or other organism.

**[0090]** As used herein, the term "downregulating expression" is defined as causing, directly or indirectly, reduction in the transcription of a desired gene, reduction in the amount, stability or translatability of transcription products (e.g. RNA) of said gene, reduction in translation of the polypeptide(s) encoded by the desired gene and/or reduction in the amount, stability, or alteration of biochemical function of the polypeptides encoded by the desired gene, so as to reduce the amount or function of the gene products. Downregulating expression of a gene RNA can be monitored, for example, by direct detection of gene transcripts (for example, by PCR), by detection of polypeptide(s) encoded by the gene or bee RNA (for example, by Western blot or immunoprecipitation), by detection of biological activity of polypeptides encode by the gene (for example, catalytic activity, ligand binding, and the like), or by monitoring changes in a cell or organism resulting from reduction in expression of a desired gene or RNA.

**[0091]** "Measuring" or "measurement," or alternatively "detecting" or "detection," means assessing the presence, absence, quantity or amount (which can be an effective amount) of either a given substance within a clinical or subject-derived sample, including the derivation of qualitative or quantitative concentration levels of such substances, or otherwise evaluating the values or categorization of a subject's clinical parameters.

**[0092]** The term "operably linked", as used in reference to a regulatory sequence and a structural nucleotide sequence, means that the regulatory sequence causes regulated expression of the linked structural nucleotide sequence. "Regulatory sequences" or "control elements" refer to nucleotide

sequences located upstream, within, or downstream of a structural nucleotide sequence, and which influence the timing and level or amount of transcription, RNA processing or stability, or translation of the associated structural nucleotide sequence. Regulatory sequences may include promoters, translation leader sequences, introns, enhancers, stem-loop structures, repressor binding sequences, termination sequences, pausing sequences, polyadenylation recognition sequences, and the like.

**[0093]** As used herein, the term "RNA silencing agent" refers to an RNA which is capable of inhibiting or "silencing" the expression of a target gene. In certain embodiments, the RNA silencing agent is capable of preventing complete processing (e.g, the full translation and/or expression) of an mRNA molecule through a post-transcriptional silencing mechanism.

**[0094]** RNA silencing agents include noncoding RNA molecules, for example RNA duplexes comprising paired strands, as well as precursor RNAs from which such small non-coding RNAs can be generated. Exemplary RNA silencing agents include dsRNAs such as siRNAs, miRNAs and shRNAs. In one embodiment, the RNA silencing agent is capable of inducing RNA interference. In another embodiment, the RNA silencing agent is capable of mediating translational repression.

**[0095]** RNA interference commonly refers to the process of sequence-specific post-transcriptional gene silencing in animals mediated by short interfering RNAs (siRNAs). The corresponding process in plants is commonly referred to as post-transcriptional gene silencing or RNA silencing and is also referred to as quelling in fungi. The process of post-transcriptional gene silencing is thought to be an evolutionarily-conserved cellular defense mechanism used to prevent the expression of foreign genes and is commonly shared by diverse eukaryotic species, including plants and animals. Such protection from foreign gene expression may have evolved in response to the production of double-stranded RNAs (dsRNAs) derived from viral infection or from the random integration of transposon elements into a host genome via a cellular response that specifically destroys homologous single-stranded RNA or viral genomic RNA.

**[0096]** The presence of long dsRNAs in cells stimulates the activity of a ribonuclease III enzyme referred to as dicer. Dicer is involved in the processing of the dsRNA into short pieces of dsRNA known as short interfering RNAs (siRNAs). Short interfering RNAs derived from dicer activity are typically about 21 to about 23 nucleotides in length and comprise about 19 base pair duplexes. The RNAi response also features an endonuclease complex, commonly referred to as an RNA-induced silencing complex (RISC), which mediates cleavage of single-stranded RNA having sequence complementary to the antisense strand of the siRNA duplex. Cleavage of the target RNA takes place in the middle of the region complementary to the antisense strand of the siRNA duplex.

**[0097]** The term "siRNA" refers to small inhibitory RNA duplexes (generally between 18-30 basepairs, between 19 and 25 basepairs) that induce the RNA interference (RNAi) pathway. Typically, siRNAs are chemically synthesized as 21mers with a central 19 bp duplex region and symmetric 2-base 3'-overhangs on the termini, although it has been recently described that chemically synthesized RNA duplexes of 25-30 base length can have as much as a 100-fold increase in potency compared with 21mers at the same location. The observed increased potency obtained using longer RNAs in triggering RNAi is theorized to result from providing Dicer with a substrate (27mer) instead of a product (21mer) and that this improves the rate or efficiency of entry of the siRNA duplex into RISC.

**[0098]** The term "shRNA", as used herein, refers to an RNA agent having a stem-loop structure, comprising a first and second region of complementary sequence, the degree of complementarity and orientation of the regions being sufficient such that base pairing occurs between the regions, the first and second regions being joined by a loop region, the loop resulting from a lack of base pairing between nucleotides (or nucleotide analogs) within the loop region. The number of nucleotides in the loop is a number between and including 3 to 23, or 5 to 15, or 7 to 13, or 4 to 9, or 9 to 11. Some of the nucleotides in the loop can be involved in base-pair interactions with other nucleotides in the loop.

**[0099]** The terms "subject," "individual," and the like are used interchangeably herein, and refer to any animal, or cells thereof whether in vitro or in situ, amenable to the methods described herein. In certain non-limiting embodiments, subject or individual is a bee.

**[0100]** "Sample" or "biological sample" as used herein means a biological material isolated from a subject. The biological sample may comprise cellular and/or non-cellular material obtained from the subject. One example of a biological sample is a tissue sample.

**[0101]** According to one embodiment of the present disclosure, the nucleic acid in the expression cassette is capable of causing cleavage and/or degradation of a target polynucleotide sequence. As used herein, the phrases "target" or "target polynucleotide sequence" refer to any sequence present in a target cell (e.g., viral or microbial pathogen), whether naturally occurring sequence or a heterologous sequence present due to an intracellular or extracellular pathogenic infection or a disease, which polynucleotide sequence has a function that is desired to be reduced or inhibited. The target sequence may be a coding sequence, that is, it is translated to express a protein or a functional fragment thereof. Alternatively, the target sequence may be non-coding, but may have a regulatory function, or it may be without any known function.

**[0102]** As used herein, the term "treating" includes abrogating, substantially inhibiting, slowing or reversing the progression of a condition, substantially ameliorating clinical or aesthetical symptoms of a condition or substantially preventing the appearance of clinical or aesthetical symptoms of a condition.

**[0103]** The term "gene" is intended to include any target sequence intended to be "silenced", whether or not transcribed and/or translated, including regulatory sequences, such as promoters, enhancers and other non-coding sequences.

**[0104]** Ranges: throughout this disclosure, various aspects of the invention can be presented in a range format. It should be understood that the description in range format is merely for convenience and brevity and should not be construed as an inflexible limitation on the scope of the invention. Accordingly, the description of a range should be considered to have specifically disclosed all the possible subranges as well as individual numerical values within that range. For example, description of a range such as from 1 to 6 should be considered to have specifically disclosed subranges such as from 1 to 3, from 1 to 4, from 1 to 5, from 2 to 4, from

2 to 6, from 3 to 6 etc., as well as individual numbers within that range, for example, 1, 2, 2.7, 3, 4, 5, 5.3, and 6. This applies regardless of the breadth of the range.

## DESCRIPTION

**[0105]** In one embodiment, the present invention relates to a method for the production of an encoded molecule in an arthropod in vivo, comprising administering a genetically engineered bacterium to the arthropod, wherein the genetically engineered bacterium comprises a nucleic acid for expression of the gene encoded molecule. Therefore, in various embodiments the invention relates to nucleic acid molecules for expression of an encoded molecule, engineered bacteria comprising the nucleic acid molecules, arthropods comprising the engineered bacteria, and methods of use of the nucleic acid molecules and engineered bacteria to treat or prevent a disease or disorder of the arthropod. In one embodiment, the invention provides compositions and methods of using engineered bacteria to deliver nucleic acid molecules to insects for therapeutic purposes. In one embodiment, the invention provides compositions and methods of using engineered bacteria to deliver nucleic acid molecules to insects for purposes of modulating at least one of the level or activity of a gene, the level or activity of a protein or a behavior of the insect. Exemplary behaviors that can be modified according to the invention include, but are not limited to, foraging behavior and aggression. In one embodiment, the invention provides compositions and methods of using engineered bacteria to deliver nucleic acid molecules to insects for purposes of killing or otherwise impairing the insect. In one embodiment, the invention provides compositions and methods of using engineered bacteria to deliver nucleic acid molecules to insects to harm or to kill a pathogen or parasite of the insect through expression of a nucleic acid molecule that is harmful for the pathogen or parasite of the insect.

**[0106]** In some embodiments, the invention relates to methods of use of the engineered bacteria to treat or prevent a disease or disorder of a bee or colony of bees. In one embodiment, the disease or disorder is associated with a bee or bee colony parasite or pathogen

### Compositions

**[0107]** In one embodiment, the invention provides nucleic acid molecules for expression of a heterologous protein or polypeptide. In certain aspects, the heterologous polypeptide is a polypeptide that improves the health of a host insect. In one embodiment, the heterologous polypeptide is a pesticide-degrading polypeptide or a cytochrome.

**[0108]** In one embodiment, the invention provides nucleic acid molecules for expression of a modulator of a gene or protein of interest. Consequently, in the context of the present invention, the term "nucleic acid molecule" refers to a DNA or RNA molecule. DNA is understood as double-stranded deoxyribonucleic acid molecules. These may be linear or circular. Exemplary RNA molecules that can be encoded by the nucleic acid molecules of the invention, include, but are not limited to RNAi, dsRNA, siRNA, shRNA, miRNA and short guide RNA (sgRNA). In one embodiment, the encoded RNA molecule functions as an inhibitor of a target RNA. In one embodiment, the target RNA is an RNA of a pathogen or other organism that may harm the host insect. In one embodiment, the target RNA is an RNA of the host insect.

**[0109]** In specific aspects, the modulator is an inhibitory nucleic acid. In particular aspects, the inhibitory nucleic acid comprises an antisense DNA. In some aspects, the inhibitory nucleic acid is a dsRNA, siRNA, shRNA, sgRNA, or miRNA. The nucleic acid may be at least 15 base pairs in length, such as 19 to 25 base pairs in length. In other aspects, the nucleic acid is longer in length, such as at least 30 base pairs.

**[0110]** An RNAi construct is understood as a double-stranded RNA that works in accordance with the principle of an interfering RNA. It is known in the state of the art how RNAi constructs have to be designed that degrade target RNA (Voorhoeve et al. (2003). "Knockdown stands up". Trends Biotechnol. 21 (1): 2-4; Henschel A, Buchholz F, Habermann B (2004) "DEQOR: a web-based tool for the design and quality control of siRNAs". Nucleic Acids Res 32 (Web Sever issue): W113-20). In this context, it relates to genes that encode for double-stranded RNA. In particular, the RNAi construct may encode for a dsRNA (double-stranded RNA) against an organism that harms the insect. In certain cases, the inhibitory nucleic acid can be targeted to an insect gene, such a bee gene. For example, insect genes can be targeted to improve the health or alter the development of a host insect. For example, vitellogenin can be targeted to alter foraging behavior, and thus modulate pollinator effectiveness.

**[0111]** The nucleic acid may be (i) plasmid DNA, (ii) linear DNA, (iii) circular DNA, (iv) single stranded DNA, (v) RNA, (vi) non natural DNA like molecules or (vii) a hybrid formed out of any of these molecules. These molecules are known in the art (see Sambrook and Russel, 2001). In one embodiment, the nucleic acid molecule is a broad-host-range vector. In one embodiment, the nucleic acid molecule is integrated into the bacterial genomic DNA.

### Broad Host Vector

**[0112]** In one embodiment, at least one nucleic acid molecule to be expressed is cloned into a broad-host-range vector backbone. Therefore, in one embodiment, the invention provides broad host vectors for use in the generation of genetically modified bacteria for expression of a gene product. The broad host vectors of the invention may comprise at least one of: a broad-host-range promoter, an antibiotic resistance cassette, an oriT sequence for delivery into recipient cells via conjugation, a ribosome binding site, an origin of replication, a bacterial terminator, at least one nucleic acid sequence for expression, and a sequence that encodes for a detectable marker.

**[0113]** Promoters are operably linked with those encoding sequences of which they initiate transcription. The encoding sequences may encode for RNA constructs, peptides or proteins. Exemplary promoters are promoters that are active in a broad range of cells including, but not limited to, CMV promoter, TK promoter, PA1, PA2, and PA3 from bacteriophage T7, cp12b, cp18, cp32, cp6, and cp25 promoters, and SV40. In one embodiment, the promoter is inducible. An exemplary inducible promoter includes a modified CP25 promoter with lacO sites (cp25 (lacO); SEQ ID NO: 12).

**[0114]** Generally, a promoter of the invention is understood to enable the transcription of DNA sequences in genetically engineered bacteria. The particular promoter

sequence will depend on the bacteria in which the gene product is to be expressed. In this context, a number of DNAs or vectors could be cloned that present a combination of different promoters having the gene sequence encoding for the desired gene product or a reporter such as GFP.

**[0115]** In exemplary embodiments, the broad host range vector comprises the RSF 1010 origin of replication and further comprises at least one heterologous nucleic acid sequence under the control of one or more (e.g., 2 or 3) broad host range promoters, such as PA1, PA2, PA3, cp12b, cp18, cp32, cp6, cp25 and cp25 (lacO). In one embodiment, the nucleic acid molecule of the invention comprises paired, convergent transcriptional promoters on a broad range vector. Exemplary pairs of transcriptional promoters that can be included on a nucleic acid molecule (e.g., for the generation of dsRNA molecules) include, but are not limited to at least two of PA1, PA2, PA3, cp12b, cp18, cp32, cp6, cp25 and cp25 (lacO). Exemplary, paired convergent transcriptional promoters include, but are not limited to, paired convergent cp25 promoters, paired convergent cp6 promoters, paired convergent cp18 promoters and paired convergent cp25 (lacO) promoters.

**[0116]** In one embodiment, the origin of replication is selected based on the bacterial species to be engineered to express the broad range vector. Origins of replication that can be included in the broad-range-vector of the invention include, but are not limited to, the RSF1010 plasmid origin of replication, the pBBR1 plasmid origin of replication, the RK2 plasmid origin of replication, the RP4 plasmid origin of replication and the pAMß1 plasmid origin of replication.

**[0117]** Furthermore, the nucleic acid molecules of the invention may also comprise enhancers. Enhancers may enhance the transcription initiation of a promoter, while they are not directly attached to the sequence of the promoter or the encoding sequence. Enhancers may be located far away or even on another DNA. Preferably, they are located on the same DNA as the promoter and the encoding sequence.

Toolkit for Modular Assembly of Broad-Host-Range Plasmids

**[0118]** In one embodiment, the invention provides one or more plasmids comprising genetic parts for the modular construction of vectors of the invention. In one embodiment, one or more plasmids comprises at least one of a broad-host-range promoter, an antibiotic resistance cassette, an oriT sequence for delivery into recipient cells via conjugation, a ribosome binding site, an origin of replication, a bacterial terminator, at least one nucleic acid sequence for expression, and a sequence that encodes for a detectable marker.

**[0119]** In one embodiment, the toolkit comprises a plasmid containing a promoter sequence for use in generating a vector of the invention. In one embodiment, the plasmid further comprises a ribosome binding site (RBS). Promoter sequences that may be included in a toolkit plasmid of the invention include, but are not limited to, SEQ ID NO: 12, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO: 41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, and SEQ ID NO:45. Exemplary plasmids comprising a promoter sequence for use in generating a vector of the invention include, but are not limited to, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58 and SEQ ID NO:59.

**[0120]** In one embodiment, the toolkit comprises a plasmid containing a reverse promoter sequence for use in generating a vector of the invention. In one embodiment, the plasmid further comprises a ribosome binding site (RBS). An exemplary reverse promoter sequence that may be included in a toolkit plasmid of the invention is set forth in SEQ ID NO:46. An exemplary plasmid comprising a reverse promoter sequence for use in generating a vector of the invention is SEQ ID NO:60.

**[0121]** In one embodiment, the toolkit comprises a plasmid containing a coding sequence for use in generating a vector of the invention. Coding sequences that may be included in a toolkit plasmid of the invention include, but are not limited to, a coding sequence encoding a sgRNA, dsRNA, siRNA, shRNA, miRNA, protein, or peptide. In one embodiment, the coding sequence encodes dCas9. An exemplary dCas9 coding sequence that may be included in a toolkit plasmid of the invention is set forth in SEQ ID NO:47. An exemplary plasmid comprising a coding sequence encoding dCas9 for use in generating a vector of the invention is SEQ ID NO:65. Exemplary plasmids comprising coding sequences for T7 RNA Polymerase, LacI repressor, GFP optim-1, Nanoluc, E2-Crimson and Kanamycin Resistance are provided as SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:67 and SEQ ID NO:68 respectively.

**[0122]** In one embodiment, the toolkit comprises a plasmid containing a terminator sequence for use in generating a vector of the invention. Terminator sequences that may be included in a toolkit plasmid of the invention include, but are not limited to, rpoC, BBa_B0015 and T7 terminators. Exemplary terminator coding sequences that may be included in a toolkit plasmid of the invention are set forth in SEQ ID NO:48, SEQ ID NO:49 and SEQ ID NO:50. Exemplary plasmids comprising a terminator sequence for use in generating a vector of the invention include, but are not limited to, SEQ ID NO:69, SEQ ID NO:70 and SEQ ID NO:71.

**[0123]** In one embodiment, the toolkit comprises a plasmid containing a repressor sequence for use in generating a vector of the invention. An exemplary plasmid comprising a repressor sequence for use in generating a vector of the invention is SEQ ID NO:66.

**[0124]** In one embodiment, the toolkit comprises a plasmid containing an origin of replication sequence for use in generating a vector of the invention. Origin of replication sequences that may be included in a toolkit plasmid of the invention include, but are not limited to, RSF1010 and R6k. Exemplary plasmids comprising an origin of replication sequence for use in generating a vector of the invention include, but are not limited to, SEQ ID NO:401, SEQ ID NO:402, SEQ ID NO:403 and SEQ ID NO: 11.

Inducible System

**[0125]** In one embodiment, the invention provides an inducible system for inducible expression of a gene product in bacteria. In one embodiment, the inducible system of the invention comprises lacI driven by the CP25 promoter, and a cp25 (lacO) promoter (SEQ ID NO: 12) operably linked to a nucleic acid sequence to be expressed. In one embodiment, the inducible system comprises lacI driven by the CP25 promoter, T7 RNAP under control of the inducible lac promoter, and a T7 promoter with lacO sites operably linked to a nucleic acid sequence to be expressed. In one embodi-

ment, the inducible system is on a single plasmid or vector. Alternatively, the inducible plasmid system may be on two or more plasmids or vectors.

Vectors for Integration into the Bacterial Genome

**[0126]** In one embodiment, the vector of the invention is capable of integrating into the genome of a target bacterial species. In one embodiment, the vector may comprise homology arms flanking at least one of a broad-host-range promoter, an antibiotic resistance cassette, a ribosome binding site, a bacterial terminator, a nucleic acid sequence for expression, and a sequence that encodes for a detectable marker. In one embodiment, a homology arm comprises at least 100 bp, at least 200 bp, at least 300 bp, at least 400 bp, at least 500 bp, at least 600 bp, at least 700 bp, at least 800 bp, at least 900 bp, at least 1000 bp, at least 1100 bp, at least 1200 bp, at least 1300 bp, at least 1400 bp, or at least 1500 bp having at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% sequence identity to a target genomic DNA sequence. Alternatively the vector that is capable of integrating into the genome of a target bacterial species may be integrated using any method that is known in the art, including, but not limited to, lambda/red recombination, and may therefore comprises one or more elements allowing for the integration of the vector or selection or counter selection of bacteria comprising the integrated vector.

**[0127]** In one embodiment, the vector for integration comprises an origin of replication of a suicide plasmid that prevents plasmid replication in the target bacterial species. An exemplary suicide plasmid origin that can be used on a vector for integration includes, but is not limited to, the R6K plasmid origin of replication.

Nucleic Acid Sequences for Expression

**[0128]** In one embodiment, the broad-host-range vector of the invention comprises at least one nucleic acid sequence to be expressed. In one embodiment, the nucleic acid sequence expresses a RNA, protein or peptide. In one embodiment the nucleic acid sequence encodes a therapeutic molecule. For example, in various embodiments, the nucleic acid sequence encodes a pesticide degrading polypeptide or a cytochrome, or encodes an enzyme in a biosynthetic pathway producing neuro-active molecules that affect behaviors such as those determining aggression or pollinator effectiveness.

**[0129]** In one embodiment, the nucleic acid sequence to be expressed is a modulator of (e.g., an inhibitor or activator) of a target gene or gene product. In various embodiments, the present invention includes compositions for modulating the level or activity of a target gene or gene product in a subject, a cell, a tissue, or an organ in need thereof. In various embodiments, the compositions of the invention modulate the amount or activity of a target polypeptide, the amount or activity of a target protein, the amount or activity of a target mRNA. In one embodiment, the compositions of the invention modulate the amount or activity of a target polypeptide, the amount or activity of a target protein, the amount or activity of a target mRNA in an insect. In one embodiment, the compositions of the invention modulate the amount or activity of a target polypeptide, the amount or activity of a target protein, the amount or activity of a target mRNA in a pathogen, or parasite of an insect.

Activators

**[0130]** In one embodiment, the nucleic acid molecules and engineered bacteria of the invention can be used to activate or increase the levels of expression or activity of a target gene or protein in a subject. Therefore, in one embodiment, the invention relates to nucleic acid molecules encoding a protein, a recombinant polypeptide, an active polypeptide fragment, or combinations thereof, which function to increase, or activate, the expression or activity of a target gene or protein. It will be understood by one skilled in the art, based upon the disclosure provided herein, that an increase in the level of a target gene or protein encompasses the increase in gene expression, including transcription, translation, or both. The skilled artisan will also appreciate, once armed with the teachings of the present invention, that an increase in the level of a protein includes an increase in protein activity (e.g., enzymatic activity, receptor binding activity, etc.). Thus, increasing the level or activity of a target gene or protein includes, but is not limited to, increasing the amount of polypeptide, increasing transcription, translation, or both, of a nucleic acid encoding a target polypeptide; and it also includes increasing any activity of a target polypeptide as well.

**[0131]** Examples of host proteins for which it would be beneficial to have increased expression using the methods of the invention include, but are not limited to, immune system proteins (e.g., Dorsal, argonaut (ago2), relish, and proteins in the toll and imd pathways) conferring greater disease resistance or enzymes in pathways producing neuro-active molecules that affect behaviors such as those determining aggression or pollinator effectiveness.

Inhibitors

**[0132]** In various embodiments, the modulator of the invention comprises an inhibitor of a target gene or protein. In one embodiment, the inhibitor of the invention decreases the amount of polypeptide, the amount of mRNA, the amount of activity, or a combination thereof, of the target gene or protein. Exemplary target genes that can be inhibited according to the methods of the invention include, but are not limited to, TOM70, TIM22, TOM40, Imp2, mitochondrial Hsp70, ATM1-ABC transporter proteins, Frataxin, Ferredoxin, ERV1, ferredoxin, NADPH oxido-reductase [FNR], pyruvate dehydrogenase α subunit, pyruvate dehydrogenase β subunit, mitochondrial glycerol-3-phosphate dehydrogenase (mtG3PDH), manganese-containing superoxide dismutase (MnSOD), DNAJ (Hsp70 interacting), Iron Sulfur cluster ISU1, Cystein desulfurase Nsf1, NAR1, RLI1, ATPase subunit A, RNA polymerase I, RNA polymerase III, Inhibitor of apoptosis (IAP), and FAS apoptotic.

**[0133]** It will be understood by one skilled in the art, based upon the disclosure provided herein, that a decrease in the level of a target gene or protein encompasses the decrease in the expression, including transcription, translation, or both. The skilled artisan will also appreciate, once armed with the teachings of the present invention, that a decrease in the level of a target protein includes a decrease in the activity of the target protein. Thus, decrease in the level or activity of a target protein includes, but is not limited to, decreasing the amount of polypeptide, and decreasing transcription, translation, or both, of a nucleic acid encoding a target protein, and it also includes decreasing any activity of the target protein as well.

**[0134]** In one embodiment, the invention provides a generic concept for inhibiting an essential gene of a virus or parasite of a bee. In one embodiment, the inhibitor is selected from the group consisting of a small interfering RNA (siRNA), a microRNA, an antisense nucleic acid, and a ribozyme.

**[0135]** One skilled in the art will appreciate, based on the disclosure provided herein, that one way to decrease the mRNA and/or protein levels of a target protein in a cell is by reducing or inhibiting expression of the nucleic acid encoding the target protein. Thus, the protein level can be decreased using a molecule or compound that inhibits or reduces gene expression such as, for example, dsRNA, siRNA, an antisense molecule or a ribozyme. However, the invention should not be limited to these examples.

**[0136]** In one embodiment, siRNA is used to decrease the level of a target gene or protein. RNA interference (RNAi) is a phenomenon in which the introduction of double-stranded RNA (dsRNA) into a diverse range of organisms and cell types causes degradation of the complementary mRNA. In the cell, long dsRNAs are cleaved into short 21-25 nucleotide small interfering RNAs, or siRNAs, by a ribonuclease known as Dicer. The siRNAs subsequently assemble with protein components into an RNA-induced silencing complex (RISC), unwinding in the process. Activated RISC then binds to complementary transcript by base pairing interactions between the siRNA antisense strand and the mRNA. The bound mRNA is cleaved and sequence specific degradation of mRNA results in gene silencing. See, for example, U.S. Pat. No. 6,506,559; Fire et al., 1998, Nature 391(19):306-311; Timmons et al., 1998, Nature 395:854; Montgomery et al., 1998, TIG 14 (7):255-258; David R. Engelke, Ed., RNA Interference (RNAi) Nuts & Bolts of RNAi Technology, DNA Press, Eagleville, PA (2003); and Gregory J. Hannon, Ed., RNAi A Guide to Gene Silencing, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2003). Soutschek et al. (2004, Nature 432:173-178) describe a chemical modification to siRNAs that aids in intravenous systemic delivery. Optimizing siRNAs involves consideration of overall G/C content, C/T content at the termini, Tm and the nucleotide content of the 3' overhang. See, for instance, Schwartz et al., 2003, Cell, 115:199-208 and Khvorova et al., 2003, Cell 115:209-216. Therefore, the present invention also includes methods of decreasing protein levels using RNAi technology.

**[0137]** In other related aspects, the invention includes a nucleic acid encoding an inhibitor, wherein an inhibitor such as a dsRNA, siRNA or antisense molecule, inhibits a target gene, operably linked to a nucleic acid comprising a promoter/regulatory sequence such that the nucleic acid is preferably capable of directing expression of the dsRNA, siRNA or antisense molecule encoded by the nucleic acid. Thus, the invention encompasses expression vectors and methods for the introduction of exogenous DNA into cells with concomitant expression of the exogenous DNA in the cells such as those described, for example, in Sambrook et al. (2012, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, New York) and as described elsewhere herein.

**[0138]** In another aspect, the invention includes a vector comprising an siRNA or antisense polynucleotide. The dsRNA, siRNA or antisense polynucleotide can be cloned into a number of types of vectors as described elsewhere herein. For expression of the dsRNA, siRNA or antisense polynucleotide, at least one module in each promoter functions to position the start site for RNA synthesis.

**[0139]** In order to assess the expression of the dsRNA, siRNA or antisense polynucleotide, the expression vector to be introduced into a cell can also contain either a selectable marker gene or a reporter gene or both to facilitate identification and selection of expressing cells from the population of cells sought to be transfected or infected through viral vectors. In other embodiments, the selectable marker may be carried on a separate piece of DNA and used in a co-transfection procedure. Both selectable markers and reporter genes may be flanked with appropriate regulatory sequences to enable expression in the host cells. Useful selectable markers are known in the art and include, for example, antibiotic-resistance genes, such as neomycin resistance and the like.

**[0140]** In one embodiment of the invention, an antisense nucleic acid sequence which is expressed by a plasmid vector is used to inhibit a target gene or protein. The antisense expressing vector is used to transfect a bacterial cell to generate a genetically engineered bacteria of the invention.

**[0141]** Antisense molecules and their use for inhibiting gene expression are well known in the art (see, e.g., Cohen, 1989, In: Oligodeoxyribonucleotides, Antisense Inhibitors of Gene Expression, CRC Press). Antisense nucleic acids are DNA or RNA molecules that are complementary, as that term is defined elsewhere herein, to at least a portion of a specific mRNA molecule (Weintraub, 1990, Scientific American 262:40). In the cell, antisense nucleic acids hybridize to the corresponding mRNA, forming a double-stranded molecule thereby inhibiting the translation of genes.

**[0142]** The use of antisense methods to inhibit the translation of genes is known in the art, and is described, for example, in Marcus-Sakura (1988, Anal. Biochem. 172: 289). Such antisense molecules may be provided to the cell via genetic expression using DNA encoding the antisense molecule as taught by Inoue, 1993, U.S. Pat. No. 5,190,931.

**[0143]** Alternatively, antisense molecules of the invention may be made synthetically and then provided to the cell. Antisense oligomers of between about 10 to about 30, and more preferably about 15 nucleotides, are preferred, since they are easily synthesized and introduced into a target cell. Synthetic antisense molecules contemplated by the invention include oligonucleotide derivatives known in the art which have improved biological activity compared to unmodified oligonucleotides (see U.S. Pat. No. 5,023,243).

**[0144]** Compositions and methods for the synthesis and expression of antisense nucleic acids are as described elsewhere herein.

**[0145]** Ribozymes and their use for inhibiting gene expression are also well known in the art (see, e.g., Cech et al., 1992, J. Biol. Chem. 267:17479-17482; Hampel et al., 1989, Biochemistry 28:4929-4933; Eckstein et al., International Publication No. WO 92/07065; Altman et al., U.S. Pat. No. 5,168,053). Ribozymes are RNA molecules possessing the ability to specifically cleave other single-stranded RNA in a manner analogous to DNA restriction endonucleases. Through the modification of nucleotide sequences encoding these RNAs, molecules can be engineered to recognize specific nucleotide sequences in an RNA molecule and

cleave it (Cech, 1988, J. Amer. Med. Assn. 260:3030). A major advantage of this approach is the fact that ribozymes are sequence-specific.

**[0146]** There are two basic types of ribozymes, namely, tetrahymena-type (Hasselhoff, 1988, Nature 334:585) and hammerhead-type. Tetrahymena-type ribozymes recognize sequences which are four bases in length, while hammerhead-type ribozymes recognize base sequences 11-18 bases in length. The longer the sequence, the greater the likelihood that the sequence will occur exclusively in the target mRNA species. Consequently, hammerhead-type ribozymes are preferable to tetrahymena-type ribozymes for inactivating specific mRNA species, and 18-base recognition sequences are preferable to shorter recognition sequences which may occur randomly within various unrelated mRNA molecules.

Modified Cas9 System

**[0147]** In one embodiment, the invention provides a modified CRISPR system for expression of a sgRNA in bacteria, wherein one or more components of the CRISPR system is encoded on a broad-host-range vector of the invention. In one embodiment, the CRISPR system of the invention comprises an endonuclease enzyme (e.g., Cas9) which binds to a target nucleic acid sequence via a sgRNA, to modulate the target sequence. In one embodiment, the CRISPR system of the invention can be used to modulate a target sequence in an insect. In one embodiment, the CRISPR system of the invention can be used to modulate a target sequence in a parasite or pathogen of an insect.

**[0148]** In some embodiments, the Cas9 enzyme comprises catalytically dead Cas9 (dCas9) or a homolog, an ortholog or mimic thereof. Orthologs of Cas9 may be from a genus which includes but is not limited to *Corynebacter, Sutterella, Legionella, Treponema, Filifactor, Eubacterium, Streptococcus, Lactobacillus, Mycoplasma, Bacteroides, Flavivola, Flavobacterium, Sphaerochaeta, Azospirillum, Gluconacetobacter, Neisseria, Roseburia, Parvibaculum, Staphylococcus, Nitratifractor, Mycoplasma* and *Campylobacter*. Catalytically dead Cas9 mimics include, but are not limited to, proteins or peptides which are capable of interaction with an sgRNA to target the CRISPR fusion construct to a site of interest. Catalytically dead or inactive Cas9, and homologs, orthologs or mimics thereof are referred to herein collectively as "dCas9."

**[0149]** In one embodiment, dCas9 lacks cleavage or nickase activity. In one embodiment dCas9, or an ortholog thereof, has a diminished nuclease activity of at least 60%, 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96, 97%, 98%, 99% or 100% as compared with a wild-type Cas9 enzyme or ortholog. In one embodiment, the CRISPR fusion constructs comprising dCas9 serve as DNA-binding proteins with very little or no catalytic activity. In one embodiment, a dCas9 comprises one or more mutations in its catalytic domain which disrupt or inactivate the nuclease activity of the Cas9 enzyme.

**[0150]** Guide RNAs

**[0151]** The present system may be used with any short guide RNA (sgRNA.) In embodiments of the invention the terms guide sequence and guide RNA are used interchangeably. In general, a guide sequence is any polynucleotide sequence having sufficient complementarity with a target polynucleotide sequence to hybridize with the target sequence and direct sequence-specific binding of a CRISPR complex to the target sequence, in some embodiments, the degree of complementarity between a guide sequence and its corresponding target sequence, when optimally aligned using a suitable alignment algorithm, is about or more than about 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97.5%, 99%, or more. Optimal alignment may be determined with the use of any suitable algorithm for aligning sequences, non-limiting example of which include the Smith-Waterman algorithm, the Needleman-Wunsch algorithm, algorithms based on the Burrows-Wheeler Transform (e.g. the Burrows Wheeler Aligner), ClustalW, Clustal X, BLAT, Novoalign, ELAND (Illumina, San Diego, Calif.), SOAP, and Maq. In some embodiments, a guide sequence is about or more than about 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 75, or more nucleotides in length. In some embodiments, a guide sequence is less than about 75, 50, 45, 40, 35, 30, 25, 20, 15, 12, or fewer nucleotides in length. In some embodiments, the guide sequence is 10-30 nucleotides long. The ability of a guide sequence to direct sequence-specific binding of a CRISPR complex to a target sequence may be assessed by any suitable assay. For example, cleavage of a target polynucleotide sequence may be evaluated in a test tube by providing the target sequence, components of a CRISPR complex, including the guide sequence to be tested and a control guide sequence different from the test guide sequence, and comparing binding or rate of cleavage at the target sequence between the test and control guide sequence reactions. Other assays are possible, and will occur to those skilled in the art. A guide sequence may be selected to target any target sequence. In one embodiment, the sgRNAs for use in the system of the invention hybridize to a target sequence in an insect administered the genetically engineered bacteria. In one embodiment, the sgRNAs for use in the system of the invention hybridize to a target sequence in a pathogen or a pest of an insect administered the genetically engineered bacteria.

Constructs Encoding Multiple Gene Products

**[0152]** In one embodiment, the compositions of the invention encode at least 1, at least 2, at least 3, at least 4, at least 5 or more than 5 gene products. In one embodiments, multiple gene products may be encoded on a single nucleic acid molecule. In an alternative embodiment, multiple gene products may be encoded on multiple nucleic acid molecules, for example, at least 1, at least 2, at least 3, at least 4, at least 5 or more than 5 gene products may be encoded on at least 1, at least 2, at least 3, at least 4, at least 5 or more than 5 nucleic acid molecules. In one embodiment, a single bacterial cell comprises multiple nucleic acid molecules encoding multiple gene products of the invention.

**[0153]** In one embodiment, the compositions of the invention comprise multiple genetically modified bacteria or populations thereof wherein each genetically modified bacteria or populations comprises at least one nucleic acid molecule encoding at least one inhibitor of the invention. Therefore, in various embodiments, the invention provides a composition comprising at least 1, at least 2, at least 3, at least 4, at least 5 or more than 5 genetically modified bacteria or populations comprising nucleic acid molecules encoding for at least 1, at least 2, at least 3, at least 4, at least 5 or more than 5 gene products of the invention. In one embodiment, multiple genetically modified bacterial popu-

lations are of the same bacterial species. In one embodiment, multiple genetically modified bacterial populations are of different bacterial species.

**[0154]** In one embodiment, multiple nucleic acid molecules encode multiple gene products of the same target gene or protein. In an alternative embodiment, multiple nucleic acid molecules encode gene products of the multiple target genes or proteins. Multiple target genes or proteins may be multiple targets in a single pathogen or parasite, or multiple targets in multiple pathogens or parasite. Therefore, in one embodiment, the invention provides a composition comprising multiple genetically encoded bacteria comprising multiple nucleic acid constructs for inhibiting at least 1, at least 2, at least 3, at least 4, at least 5 or more than 5 target genes or proteins. In one embodiment, the invention provides a composition comprising multiple genetically encoded bacteria comprising multiple nucleic acid constructs for inhibiting or modulating at least 1, at least 2, at least 3, at least 4, at least 5 or more than 5 target pathogens or parasites.

Genetically Modified Bacteria

**[0155]** In one embodiment, the invention provides genetically modified bacteria comprising a nucleic acid molecule for expression of a heterologous nucleic acid sequence. In one embodiment the present disclosure provides microbial compositions comprising one or more bacteria genetically engineered to express a heterologous nucleic acid. In one embodiment, the one or more bacteria are native to the microbiome of an arthropod. In one embodiment, the one or more bacteria are native to the microbiome of a bee. In some aspects, the one or more bacteria are *Snodgrassella alvi*, *Bartonella apis*, *Gilliamella apicola*, *Serratia marcescens* or *Parasaccharibacter apium*. In one embodiment, the one or more bacteria are bacterial species that can live in bee hive materials. In one embodiment, *Parasaccharibacter apium* or *Lactobacillus* sp. In some aspects, the composition comprises 2, 3, 4, or 5 genetically engineered bacterial species.

**[0156]** In one embodiment, one or more nucleic acid molecule for expression of a heterologous nucleic acid sequence is contained in the bacterial cell. In one embodiment, one or more nucleic acid molecule for expression of a heterologous nucleic acid sequence is integrated into the genome of said one or more bacteria. In certain aspects, the one or more bacteria express at least two heterologous nucleic acids.

Host Organisms

**[0157]** The methods of the present invention can be applied to all arthropods. The compositions and methods of the invention can be used for all purposes where the production of a gene encoded molecule in an arthropod is suitable and/or desirable. This includes the use of the method of the invention to generate a host arthropod, where at least one genetically modified bacteria for production of a gene encoded molecule has been administered to the host arthropod.

**[0158]** In one embodiment of the present invention, the arthropod may be, but is not limited, to an insect, an arachnid, a crustacean and a myriapodum. In one embodiment, the arthropod is an insect. Insects that can serve as hosts according to the methods of the invention include, but are not limited to, a holometabolic insect, a hemimetabolic insect, and an insect from the order Hymenoptera, Coleoptera or Orthoptera. Exemplary insects include, but are not limited to, a honey bee (*Apis mellifera*) and a small hive beetle (*Aethina tumida*). In another embodiment of the present invention, the arthropod is an arachnid, for example, a mite.

**[0159]** In some embodiments, the insect is a bee. As used herein, the term "bee" refers to both an adult bee and pupal cells thereof. According to one embodiment, the bee is in a hive. An adult bee is defined as any of several winged, hairy-bodied, usually stinging insects of the superfamily Apoidea in the order Hymenoptera, including both solitary and social species and characterized by sucking and chewing mouthparts for gathering nectar and pollen. Exemplary bee species include, but are not limited to, *Apis, Bombus, Trigona, Osmia* and the like. In one embodiment, bees include, but are not limited to bumblebees (*Bombus terrestris*), honeybees (*Apis mellifera*) (including foragers and hive bees) and *Apis* cerana.

**[0160]** According to one embodiment, the bee is part of a colony. The term "colony" refers to a population of bees comprising dozens to typically several tens of thousand bees that cooperate in nest building, food collection, and brood rearing. A colony normally has a single queen, the remainder of the bees being either "workers" (females) or "drones" (males). The social structure of the colony is maintained by the queen and workers and depends on an effective system of communication. Division of labor within the worker caste primarily depends on the age of the bee but varies with the needs of the colony. Reproduction and colony strength depend on the queen, the quantity of food stores, and the size of the worker force. Honeybees can also be subdivided into the categories of "hive bees", usually for the first part of a workers lifetime, during which the "hive bee" performs tasks within the hive, and "forager bee", during the latter part of the bee's lifetime, during which the "forager" locates and collects pollen and nectar from outside the hive, and brings the nectar or pollen into the hive for consumption and storage.

Formulations

**[0161]** In some embodiments, the composition is a bee-ingestible composition. In certain aspects, the bacteria are present as a live suspension or a lyophilized powder. The composition may be in solid form or liquid form, such as a sucrose solution or a corn syrup solution. In some aspects, the composition comprises protein and/or pollen. In additional aspects, the composition further comprises a carbohydrate or sugar supplement or other additive or excipient as described elsewhere herein.

Methods

**[0162]** The present invention also relates to the use of a genetically engineered bacteria comprising a nucleic acid encoding a molecule for use in a method for protecting an arthropod against pest influences, parasites or ectoparasites or for use in a method for treating an arthropod disease in an arthropod, wherein the genetically engineered bacteria is fed to the arthropod or is administered orally. All embodiments described above with respect to the method of the invention also apply to said use of the invention.

**[0163]** In one embodiment, there is provided a method for producing a genetically engineered bacterial species accord-

ing to the composition of the embodiments (e.g., engineered bacteria) comprising transfecting said bacterial species with an expression cassette comprising at least one heterologous nucleic acid.

**[0164]** In one embodiment, there is provided a method for downregulating expression of a target gene product in an insect pathogen or parasite comprising administering an effective amount of one or more genetically engineered bacteria to a host insect, wherein said bacteria express an inhibitor of the target gene product. In some aspects, the target gene product is a pathogen or parasite-specific gene product, such as a virus- or mite-specific gene product. Parasites and pathogens that can be targeted using the methods of the invention include, but are not limited to, Acute Bee Paralysis Virus (ABPV), Kashmir Bee Virus (KBV), Israeli Acute Paralysis Virus (IAPV), *Nosema ceranae*, deformed wing virus, or *Varroa destructor* mite and other parasites or pathogens which infect insects.

**[0165]** In one embodiment, there is provided a method for downregulating expression of a target gene product in a target insect. In one embodiment, the method comprises administering an effective amount of one or more genetically engineered bacteria to a food source of a target insect, wherein said bacteria express a dsRNA complementary to a gene product of the target insect. In one embodiment, the target insect is a pest insect. In one embodiment, the target insect is a Small Hive Beetle, and a food source of the target insect is a bee hive.

**[0166]** In some aspects, the one or more genetically engineered bacteria are comprised in a composition of the embodiments described herein. In particular aspects, the one or more genetically engineered bacteria are selected from the group consisting of *Snodgrassella alvi*, *Bartonella apis*, *Gilliamella apicola*, *Serratia marcescens*, *Parasaccharibacter apium*, and *Lactobacillus* sp.

**[0167]** In one embodiment, there is provided a method for expression of a heterologous nucleic acid sequence in an arthropod. In one embodiment, the method comprises administering to the arthropod a composition comprising one or more genetically modified bacteria, wherein the one or more genetically modified bacteria comprise a vector for expression of a heterologous nucleic acid sequence. In one embodiment, one or more bacteria is native to the host arthropod microbiome and as such is able to persist in the gut of the arthropod for at least one day, at least two days, at least three days, at least 4 days, at least 5 days, at least 6 days, at least 7 days, at least 8 days, at least 9 days, at least 10 days, at least 11 days, at least 12 days, at least 13 days, at least 14 days, at least 15 days, or for more than 15 days. In one embodiment, the one or more bacteria persists in the gut of the arthropod for the duration of the arthropod's lifespan.

**[0168]** In a further embodiment, there is provided a method for reducing the susceptibility of a bee to a disease or disorder associated with a pathogen or parasite. In one embodiment, the disease or disorder is Colony Collapse Disorder (CCD). Therefore, in one embodiment, the invention provides a method of treating or preventing CCD comprising administering an effective amount of one or more genetically engineered bacteria to a bee, or bee hive component, wherein said bacteria express a dsRNA complementary to a pathogen or parasite-specific gene product.

**[0169]** In another embodiment, there is provided method for reducing the susceptibility of a bee to an infection comprising feeding the bee an effective amount of one or more genetically engineered bacteria to said bee, wherein said bacteria express a dsRNA complementary to a pathogen or parasite-specific gene product.

**[0170]** In some aspects of the above embodiments, the pathogen or parasite is Acute Bee Paralysis Virus (ABPV), Kashmir Bee Virus (KBV), Israeli Acute Paralysis Virus (IAPV), *Nosema ceranae*, deformed wing virus, *Varroa destructor* mite, or Small Hive Beetle. In certain aspects, the *Nosema ceranae* is *N. ceranae* or *N. apis*. In specific aspects, the pathogen or parasite-specific gene product is an mRNA encoding a *Nosema* mitosomal protein. In some aspects, the *Nosema* mitosomal protein is selected from the group consisting of TOM70, TIM22, TOM40, Imp2, mitochondrial Hsp70, ATM1-ABC transporter proteins, Frataxin, Ferredoxin, ERV1, ferredoxin, NADPH oxido-reductase [FNR], pyruvate dehydrogenase a subunit, pyruvate dehydrogenase 3 subunit, mitochondrial glycerol-3-phosphate dehydrogenase (mtG3PDH), manganese-containing superoxide dismutase (MnSOD), DNAJ (Hsp70 interacting), Iron Sulfur cluster ISU1, Cystein desulfurase Nfs1, NAR1 and RLI1. In other aspects, the parasite-specific gene product is ATPase subunit A, RNA polymerase I, RNA polymerase III, Inhibitor of apoptosis (IAP), or FAS apoptotic. In certain aspects, the bacteria express at least two non-contiguous dsRNAs targeting a gene product of at least one of Acute Bee Paralysis Virus (ABPV), Kashmir Bee Virus (KBV), Israeli Acute Paralysis Virus (IAPV), *Nosema ceranae*, deformed wing virus, *Varroa destructor* mite, and/or Small Hive Beetle. In certain aspects, the bacteria express at least two non-contiguous dsRNAs targeting a gene product of at least two of Acute Bee Paralysis Virus (ABPV), Kashmir Bee Virus (KBV), Israeli Acute Paralysis Virus (IAPV), *Nosema ceranae*, deformed wing virus, *Varroa destructor* mite, and/or Small Hive Beetle.

**[0171]** In one embodiment, the agents of the present invention are used to prevent the *Varroa destructor* mite from living as a parasite on the bee, or larvae thereof. In one embodiment, the agents of the present invention are used to prevent the Small Hive Beetle from living as a parasite on the bee colony. Therefore, in one embodiment, the agents of the present invention are capable of down-regulating expression of a gene product of a *Varroa destructor* mite or Small Hive Beetle.

**[0172]** In one embodiment, the agents of the present invention are used to prevent a viral infection of the bee, or larvae thereof. Therefore, in one embodiment, the agents of the present invention are capable of down-regulating expression of a gene product of a viral gene.

**[0173]** As used herein, the phrase "gene product" refers to an RNA molecule or a protein. According to one embodiment, the targeted gene product is one which is essential for parasite viability, parasite reproduction, viral replication or viral pathogenicity. Down-regulation of such a gene product would typically result in killing of a parasite, extermination of the parasite population or reduced viral pathogenicity.

**[0174]** It will be appreciated that whilst the agents of the present invention are capable of downregulating expression of a gene product of a parasite or pathogen, it is preferable that they downregulate to a lesser extent expression of the gene product in other animals, such as the bee. Accordingly, the agents of the present invention must be able to distinguish between the target gene and the bee gene, down-regulating the former to a greater extent than the latter.

According to another embodiment the agents of the present invention do not down-regulate the bee gene whatsoever. This may be effected by targeting a gene that is present or expressed in the target pathogen or parasite and not in the bee. Alternatively, the agents of the present invention may be targeted to parasite-specific sequences of a gene that is expressed both in the target parasite and in the bee.

**[0175]** According to one embodiment the agents of the present invention target segments of genes that are at least 100 bases long and do not carry any sequence longer than 19 bases that is entirely homologous to any bee-genome sequence or human-genome sequence.

**[0176]** Downregulating expression of a gene product can be monitored, for example, by direct detection of gene transcripts (for example, by PCR), by detection of polypeptide(s) encoded by the gene or bee pathogen RNA (for example, by Western blot or immunoprecipitation), by detection of biological activity of polypeptides encode by the gene (for example, catalytic activity, ligand binding, and the like), or by monitoring changes in the parasite or pathogen (for example, reduced proliferation, reduced virulence, reduced motility etc) and by testing bee infectivity/pathogenicity.

**[0177]** Downregulation of a target gene product can be effected on the genomic and/or the transcript level using a variety of agents which interfere with transcription and/or translation (e.g., RNA silencing agents, Ribozyme, DNAzyme and antisense).

**[0178]** According to one embodiment, the agent which down-regulates expression of a target gene product is a polynucleotide agent, such as an RNA silencing agent. According to this embodiment, the polynucleotide agent is greater than 15 base pairs in length.

**[0179]** As used herein, the phrase "RNA silencing" refers to a group of regulatory mechanisms (e.g. RNA interference (RNAi), transcriptional gene silencing (TGS), post-transcriptional gene silencing (PTGS), quelling, co-suppression, and translational repression) mediated by RNA molecules which result in the inhibition or "silencing" of the expression of a corresponding protein-coding gene or bee pathogen RNA sequence. RNA silencing has been observed in many types of organisms, including plants, animals, and fungi.

**[0180]** In one embodiment of the present disclosure, synthesis of RNA silencing agents suitable for use with the present invention can be effected as follows. First, the pathogen polypeptide mRNA or other target sequence is scanned downstream of the AUG start codon for AA dinucleotide sequences. Occurrence of each AA and the 3' adjacent 19 nucleotides is recorded as potential siRNA target sites. Preferably, siRNA target sites are selected from the open reading frame, as untranslated regions (UTRs) are richer in regulatory protein binding sites. UTR-binding proteins and/or translation initiation complexes may interfere with binding of the siRNA endonuclease complex. It will be appreciated though, that siRNAs directed at untranslated regions may also be effective, as demonstrated for GAPDH wherein siRNA directed at the 5' UTR mediated about 90% decrease in cellular GAPDH mRNA and completely abolished protein level (see Ambion® technical library 91/912 at the Ambion® website).

**[0181]** Second, potential target sites are compared to an appropriate genomic database (e.g., human, mouse, rat, insect, etc.) using any sequence alignment software, such as the BLAST software available from the NCBI server of the NIH. Putative target sites which exhibit significant homology to other coding sequences are filtered out. For example, host (e.g. bee) target sites can be filtered out in this manner.

**[0182]** Qualifying target sequences are selected as template for siRNA synthesis. Preferred sequences are those including low G/C content as these have proven to be more effective in mediating gene silencing as compared to those with G/C content higher than 55%. Several target sites are preferably selected along the length of the target gene or sequence for evaluation. For better evaluation of the selected siRNAs, a negative control is preferably used in conjunction. Negative control siRNA preferably include the same nucleotide composition as the siRNAs but lack significant homology to the genome. Thus, a scrambled nucleotide sequence of the siRNA is preferably used, provided it does not display any significant homology to host gene sequences, or non-relevant pathogen target sequences.

**[0183]** It will be appreciated that the dsRNA sequences target RNA transcripts complementary to DNA sequences of the targeted gene which are expressed in the parasite (transcribed into RNA), and that the actual complementation taking place in the RNAi pathways occurs following reduction of the dsRNA to smaller fragments by the RNAi enzymes.

**[0184]** It will be appreciated that since *Varroa* mites use their mouths to puncture the bee exoskeleton and feed on the bee's hemolymph, the present invention contemplates delivering the polynucleotide agents of the present invention to the bees via the genetically engineered bacteria, whereby they become presented in the bee's hemolymph thereby becoming available to the mite. Thus, according to another embodiment, the nucleic acid agents are delivered indirectly to the mites (e.g. via the bee). In this embodiment, the promoter of the nucleic acid construct is typically operational in bee gut bacterial cells

**[0185]** In one embodiment said nucleic acid encodes an siRNA or an antisense RNA. In this case, the siRNA or antisense RNA molecule is produced in the arthropod by feeding the arthropod with the genetically engineered bacteria. In this context, the nucleic acid is preferably a vector encoding the siRNA and further containing promoter sequences enabling the production of the siRNA.

**[0186]** In one embodiment, of the present invention, the genetically engineered bacteria is fed to the arthropod. In principle, every feed can be used that is accepted by the arthropod to be fed. This includes any kind material that is consumed orally by the arthropods, independent on whether it is natural feed, agricultural feed or laboratory feed and independent on whether it is consumed naturally or is administered by means of technical devices or is taken up casually. In one embodiment, the feed that is used to induce the production of the gene encoded molecules in the arthropods is either a liquid feed comprising the genetically engineered bacteria, a dry feed mixed with a solution comprising the genetically engineered bacteria or a dry feed comprising the genetically engineered bacteria in any of these formulations.

**[0187]** As detailed herein, bee feeding is common practice amongst bee-keepers, for providing both nutritional and other, for example, supplemental needs. Bees typically feed on honey and pollen, but have been known to ingest non-natural feeds as well. Bees can be fed various foodstuffs including, but not limited to Wheast (a dairy yeast grown on cottage cheese), soybean flour, yeast (e.g. brewer's yeast,

torula yeast) and yeast products products-fed singly or in combination and soybean flour fed as a dry mix or moist cake inside the hive or as a dry mix in open feeders outside the hive. Also useful is sugar, or a sugar syrup. The addition of 10 to 12 percent pollen to a supplement fed to bees improves palatability. The addition of 25 to 30 percent pollen improves the quality and quantity of essential nutrients that are required by bees for vital activity.

**[0188]** Cane or beet sugar, isomerized corn syrup, and type-50 sugar syrup are satisfactory substitutes for honey in the natural diet of honey bees. The last two can be supplied only as a liquid to bees.

**[0189]** Liquid feed can be supplied to bees inside the hive by, for example, any of the following methods: friction-top pail, combs within the brood chamber, division board feeder, boardman feeder, etc. Dry sugar may be fed by placing a pound or two on the inverted inner cover. A supply of water must be available to bees at all times. In one embodiment, pan or trays in which floating supports-such as wood chips, cork, or plastic sponge—are present are envisaged. Detailed descriptions of supplemental feeds for bees can be found in, for example, USDA publication by Standifer, et al 1977, entitled "Supplemental Feeding of Honey Bee Colonies" (USDA, Agriculture Information Bulletin No. 413).

**[0190]** All the bees in a hive are potentially susceptible to the pathogenic diseases detailed herein. Thus, according to some embodiments, the bees can be honeybees, forager bees, hive bees and the like.

**[0191]** Also provided is a method for reducing the susceptibility of a bee to a disease caused by pathogens, the method effected by feeding the bee on an effective amount of a nucleic acid or nucleic acid construct comprising a nucleic acid agent downregulating expression of a polypeptide of the bee pathogen and/or causing cleavage and/or degradation of a bee pathogen RNA. Methods for reducing the susceptibility of a bee colony or bee-hive to bee pathogens by feeding oligonucleotides and/or polynucleotides are envisaged. Thus, in some embodiments, the present invention can be used to benefit any numbers of bees, from a few in the hive, to the entire bee population within a hive and its surrounding area. It will be appreciated, that in addition to feeding of oligonucleotides and/or polynucleotides for reduction of the bee pathogen infection and infestation, enforcement of proper sanitation (for example, refraining from reuse of infested hives) can augment the effectiveness of treatment and prevention of infections.

Kits

**[0192]** The invention also includes a kit comprising a composition of the invention, for example, a kit comprising one or more plasmid comprising a genetic part for the modular construction of a vector for use in generating a genetically modified bacterium of the invention or a genetically modified bacterium comprising a nucleic acid molecule for expression of a heterologous gene product. In one embodiment, the kit may also comprise instructional material which describes, for instance, methods of generation of a vector of the invention, methods of generating a genetically engineered bacterium of the invention, or methods of administering a genetically engineered bacterium of the invention to a target arthropod.

I. EXAMPLES

**[0193]** The following examples are included to demonstrate preferred embodiments of the invention. It should be appreciated by those of skill in the art that the techniques disclosed in the examples which follow represent techniques discovered by the inventor to function well in the practice of the invention, and thus can be considered to constitute preferred modes for its practice. However, those of skill in the art should, in light of the present disclosure, appreciate that many changes can be made in the specific embodiments which are disclosed and still obtain a like or similar result without departing from the spirit and scope of the invention.

Example 1 Microbiome Engineering in Honey Bee Gut

**[0194]** Several bacterial species including *Snodgrassella alvi*, *Bartonella apis*, *Gilliamella apicola*, and *Parasaccharibacter apium* from the bee gut were engineered to express heterologous genes. A broad host range plasmid pMMB67EH was used to express the heterologous gene. It was then confirmed that these modified bacteria can reestablish and be maintained in bees.

**[0195]** A variety of broad host range promoters were tested for their ability to express GFP in vitro in the various bacterial species. The promoters that were tested included PA1, PA2, PA3, and cp25 (FIG. **1**). For *S. alvi*, all of the promoters tested resulted in similar expression of GFP. PA2, PA3, and cp25 resulted in higher GFP expression as compared to PA1 in *G. apicola*. Finally, PA1 was observed to have the highest expression of GFP in *B. apis*.

**[0196]** To determine if the modified bacteria could continue to function in the bee gut, the engineered strains *Snodgrassella alvi* and *Serratia marcescens* producing fluorescent proteins were inoculated into germ-free bees and the guts were fluorescently imaged 5 days later (FIG. **2**). The control uninoculated bees showed no fluorescent signal. On the other hand, bees inoculated with either the *S. marcescens* or *S. alvi* had expression of the fluorescent proteins. In addition, co-inoculation with both fluorescent strains resulted in expression of both fluorescent proteins.

**[0197]** In order to test the ability of the engineered bacterial species to produce dsRNA in honey bees, the broad host range backbone plasmid and the knowledge of promoter strength in bee gut species were used to produce a dsRNA-producing construct (FIG. **3**). The construct has flanking broad range promoters driving RNA transcription of the tubulin alpha-1 chain-like gene (*Apis mellifera* (honey bee)). The *Snodgrassella alvi* wkB2 strain was used to produce dsRNA using the construct and was then grown in cultures in vitro for 48 hours. RNA was extracted, transcribed to cDNA, and qPCR was performed, normalizing to bacterial 16S copies (FIG. **4**). The bacteria was found to produce dsRNA in vitro for alpha tubulin.

**[0198]** Next, it was determined whether the bacteria would produce dsRNA in vivo in bees. Bees were inoculated with *S. alvi* expressing Tyrosine Hydroxylase (TH) dsRNA or a GFP dsRNA control. After 6 days, guts were homogenized and RNA was extracted. Quantitative RT-PCR was used to assess the relative amount of TH RNA present in the gut compared to an actin control (FIG. **5**).

Example 2 Induction of RNAi in Honey Bees

**[0199]** The broad-host-range backbone was used to build two dsRNA-producing constructs to test in vivo. In one construct, a dsRNA cassette targeting alpha-tubulin, an essential bee gene, was used. Downregulation of this essen-

tial gene by induction of an RNAi response should result in increased mortality. Newly emerged workers were colonized with *S. alvi* wkB2 containing dsRNA constructs targeting alpha-tubulin or GFP-control. Survival was monitored daily (FIG. 6). The bees with dsRNA constructs targeting alpha tubulin were observed to have increased mortality. Thus, the microbe-produced dsRNA can induces RNAi in vivo in honey bees.

**[0200]** The second construct that was tested had dsRNA targeting of tyrosine hydroxylase (TH). Bees were inoculated with bacteria containing the construct targeting TH and the ability to induce a systemic RNAi response and down regulated gene transcription in other body sites was determined. FIG. 7 shows that the dsRNA construct was able to induce RNAi and downregulated expression of TH in the bee head. Therefore, the present methods can be used to engineer bacteria which are introduced to bees and induce expression of a gene or induce RNAi against a target gene. This method may be broadly useful for downregulating honey bee genes for research as well as industrial uses to improve bee health and lower the impact of bee pathogens.

Example 3—Materials and Methods

**[0201]** The plasmids used for the above experiments were all derived from the rsf1010 origin from the pMMB67EH broad host range *E. coli* plasmid. The pBTK520 plasmid was used to express GFP fluorescent protein and had spectinomycin resistance (FIG. 8A). The pBTK570 (FIG. 8B) plasmid expressing E2-Crimson fluorescent protein was used to validate gene expression in vivo. The pBTK562 plasmid (FIG. 8C) expressing dsRNA against GFP was used as control in dsRNA experiments. The pBTK561 plasmid (FIG. 8D) expressing dsRNA against bee alpha tubulin was used to show in vivo killing of bees and expression of dsRNA in vitro. Finally, the pBTK590 plasmid (FIG. 8E) expressing dsRNA against Tyrosine Hydroxlyase was used to confirm gene knock down in the head in vivo.

Example 4—Genetic Engineering of Bee Gut Microbiome Bacteria with a Toolkit for Modular Assembly of Broad-Host-Range Plasmids

**[0202]** A preliminary screen with a variety of broad-host-range plasmids was performed with different replication origins (RP4, pBBR1, RSF1010) and antibiotic resistance markers (kanamycin, ampicillin, chloramphenicol, spectinomycin) for their ability to be transferred by conjugation and stably maintained in two bacterial species, *S. alvi* and *G. apicola*, which are both abundant in the honey bee gut (FIG. 16). Plasmid pMMB67EH, a synthetic plasmid containing an RSF1010 origin (Furste J P, et al., 1986, Gene, 48:119-131), was the most versatile: it replicated in both species. Plasmids containing an RSF1010 origin are known to be extremely broad-hostrange (BHR) because they encode multiple ORFs that make them less dependent on the presence of specific proteins in a host cell for replication (Jain A et al., 2013, FEMS Microbiol Lett., 348:87-96; Meyer R. et al., 2009, Plasmid, 62:57-70). Additionally, they contain a promiscuous origin of transfer (oriT) that enables one-way transfer of the plasmid to a recipient cell from a donor cell encoding a conjugation apparatus in trans on the chromosome, such as *E. coli* MFDpir (Ferrieres L. et al., 2010, J Bacteriol., 192:6418-6427).

**[0203]** Because of these characteristics of pMMB67EH, it was decided to create a toolkit of genetic parts for hierarchical and combinatorial assembly into its RSF1010-derived backbone for testing in additional species (FIG. 9A). These BTK parts are compatible with the Golden Gate cloning scheme used by the Yeast Toolkit (YTK) (Lee M E, 2015, ACS Synth Biol., 4:975-986), and connector parts from the YTK are required for BTK assembly. BTK parts are classified into eight types defined by the specific flanking overhangs generated by type IIS restriction enzyme cleavage. Entry vectors containing any complete set of parts labeled 1-8 can be combined via one BsaI Golden Gate assembly reaction into a complete plasmid (FIG. 9B). This assembly creates a Stage 1 plasmid comprising parts 2-4 flanked by assembly connector parts 1 and 5 with vector backbone components in parts 6-8. Transcriptional units from multiple Stage 1 plasmids that have matching sets of connector parts can be further composed into one vector by Stage 2 assembly using BsmBI (FIG. 9C).

**[0204]** To create BTK vector backbone parts, the high-copy number bacterial ColE1 origin of YTK part 8 plasmids ($Amp^R$, $Kan^R$, $Spec^R$) was replaced with the RSF1010 origin from pMMB67EH. These backbones retain the oriT for delivery into recipient cells via conjugation, which is useful for genetically modifying bacterial species and strains lacking established chemical or electrical transformation techniques. In the original YTK, Type 6 and 7 parts encode a yeast marker and a yeast origin, respectively. Type 6 part overhangs was repurposed for flanking a DNA sequence encoding an optional additional CDS (such as a repressor) in the reverse orientation relative to the Type 3 part CDS, and Type 7 part overhangs for incorporating an optional reverse promoter for driving expression of the part 6 CDS. A combined Type 6-7 linker (pBTK301) can be used in lieu of these parts to create constructs lacking this extra reverse gene. These vectors were used to construct a variety of plasmids containing a single fluorescent protein driven by a broad-host-range promoter. To build more complex assemblies, such as those with T7 RNA polymerase driving inducible expression of GFP, multiple vectors from BsaI Stage 1 assembly in a Stage 2 BsmBI assembly were combined. Notable components of the BTK that expand on the set of genetic parts available in the YTK for compatible assembly include: 3 BHR plasmids with different antibiotic resistance cassettes and oriT as Type 8 origin parts for Stage 1 assembly; 2 BHR plasmids ready for Stage 2 assembly ($Spec^R$, $Kan^R$); 11 bacterial promoter/RBS combinations as Type 2 parts; 6 new CDSs including E2-Crimson (Strack R L et al., 2009, Biochemistry, 48:8279-8281) and Nanoluc (Hall M P et al., 2012, ACS Chem Biol., 7:1848-1857) for in vivo visualization as Type 3 parts; 3 bacterial terminators as Type 4 parts; 1 transcriptional repressor (LacI) as a Type 6 part; 2 R6K-origin plasmid backbones to assemble suicide plasmids for gene disruption or chromosomal modification and pre-assembled plasmids with BHR promoters for immediate testing in new bacterial strains.

**[0205]** In summary, BTK, the first Golden Gate toolkit designed specifically for the combinatorial assembly of broad-host-range plasmids was built, with the aim of expanding synthetic biology into diverse bacteria native to non-laboratory environments. In this study, BTK was applied to modify bacteria found in the honey bee gut microbiome. These species are typical of many other bacteria in natural microbial communities of interest: they have

only been cultured recently, are phylogenetically diverse, and have few or no established genetic tools. Fundamental BTK components needed for genetic modification was validated, including antibiotic selection markers, conjugation procedures, and promoters to express proteins under constitutive or inducible control. BGM strains engineered with the BTK colonize the guts of newly emerged bees, and fluorescent in vivo imaging revealed a characteristic spatial distribution of each species in the gut.

**[0206]** The species engineered are within the Proteobacteria, a diverse Gram-negative phylum that is a common component of animal- and plant-associated communities. Although it has not yet being tested more broadly, BTK components should also be useful for genetically modifying other bacteria native to other natural communities. The core of the BTK is the RSF1010 plasmid origin, which is known to replicate in diverse bacterial lineages including Cyanobacteria, *Agrobacterium*, and others (Jain A et el., 2013, FEMS Microbiol Lett., 348:87-96; Meyer R et al., 2009, Plasmid, 62:57-71; Taton A et al., 2014, Nucleic Acids Res., 42:gku673-e136; Clewell D B et al., 1974, J Bacteriol., 117:283-289). The BTK also includes promoters that have previously been shown to function in both Gram-negative and Gram-positive bacteria (Jensen P R et al., 1998, Appl Environ Microbiol., 64:82-87). The E2-Crimson reporter gene fluoresces at far-red excitation wavelengths, which is ideal for in vivo imaging of bacteria through tissue in host-associated systems (Strack R L et al., 2009, Biochemistry, 48:8279-8281). While broadhost-range plasmids have already been extensively used to study newly isolated bacteria in the past (Clewell D B et al., 1974, J Bacteriol., 117:283-289), the combinatorial nature of this new toolkit makes it possible to test multiple antibiotic resistance markers and promoters, which are more difficult to replace in plasmids that rely on classical cloning approaches.

**[0207]** Standard part definitions enable researchers to customize a toolkit by adding new capabilities for their own applications, as it was done with re-using parts from the yeast toolkit (YTK) (Lee M E et al., 2015, ACS Synth Biol., 4:975-986). Separating antibiotic resistance cassettes and replication origins into different subparts and adding to the library of choices available for each function allows more combinations of antibiotics and origins to be tested when first working with a new species. Gram-positive origins, such as pAMβ1 (Clewell D B et al., 1974, J Bacteriol., 117:283-289), would be especially useful for the manipulation of other common host-associated phyla such as Firmicutes and Actinobacteria (Robinson C J et al., 2010, Microbiol Mol Biol Rev., 74:453-476). Further validation of some BTK parts, such as the dCas9 and Cas9 systems, is needed to conclude that they will function reliably across diverse species. Other established broad-host-range tools—such as Tn7-transposon integration (Choi K H et al., 2005, Nat Meth., 2:443-448), Group II intron-based gene disruption (Enyeart P J et al., 2013, Mol Syst Biol., 9:685-685), and emerging CRISPR methods for targeted mutagenesis (Banno S et al., 2018, Nat Microbiol., 339:819)—could also be incorporated into the BTK-compatible Golden Gate framework in the future.

**[0208]** Application of the BTK to engineering bee gut bacteria enables new approaches to microbiome research in these insect species that are important pollinators and model systems for studying social behavior and learning. For example, gene disruption combined with fluorescent visualization of bacterial cells in living bees can be used to improve understanding of the molecular basis of host-microbe and microbe-microbe interactions and their relevance to host health. The BTK can also be used to implement and test biotechnological approaches for mitigating threats to bee health. For example, it could be used to engineer commensal gut bacteria to degrade pesticides or suppress pathogen populations (i.e., paratransgenesis) (Rangberg A et al., 2012, Integr Comp Biol., 52:89-99). These efforts could one day profoundly affect the health of the bee colonies that sustain modern agriculture.

**[0209]** The materials and methods employed in these experiments are now described.

**[0210]** Bacterial Culture

**[0211]** A complete list of bacterial strains used in this work and their sources is available as FIG. **17**. Unless otherwise specified, bacterial strains *S. alvi* wkB2, *G. apicola* wkB7, *Parasaccharibacter apium* wkB6, *B. apis* PEB0150, *G. apicola* PEB0183, *B. apis* PEB0149, and *S. alvi* PEB0171, *S. marcescens* N10A28 were grown on Columbia agar supplemented with 5% sterile sheep's blood (B-COL) and incubated at 35° C. in a 5% CO2 atmosphere as static cultures. *E. coli* were cultured at 37° C. with orbital shaking at 225 rpm over a 1-inch diameter. *E. coli* MFDpir was grown in LB supplemented with 0.3 mM diaminopimelic acid (DAP). *E. coli* EC100D and *E. coli* DH5α were grown in LB.

**[0212]** For antibiotic selection, the following concentrations were used: ampicillin (100 μg/mL *E. coli*, 30 μg/mL *S. alvi*, 30 μg/mL *G. apicola*, 30 μg/mL *B. apis*, 300 μg/mL *S. marcescens*), kanamycin (50 μg/mL *E. coli*, 20 μg/mL *S. alvi*, 20 μg/mL *G. apicola*, 20 μg/mL *B. apis*), spectinomycin (60 μg/mL *E. coli*, 30 μg/mL for *S. alvi*, 30 μg/mL *G. apicola*, 30 μg/mL *B. apis*, 30 μg/mL *P. apium*, 180 μg/mL *S. marcescens*).

**[0213]** Biparental Conjugation

**[0214]** MFDpir with mobilizable plasmid ("donor strain") was grown overnight, shaking in LB with appropriate selective antibiotics and DAP (0.3 mM) supplementation. Recipient strains (wkB2, wkB7, PEB0150, PEB0183, PEB0171, N10A28, wkB6, wkB12, Snod 2-15, Pens 2-2-5) were grown overnight on solid media. Recipient and donor strains were washed in 1 mL PBS, spun down (1006×g for 5 minutes), and resuspended with 1 mL of PBS. These two suspensions were mixed in a 9:1 OD ratio of recipient:donor and spotted (without filter) onto a B-COL plate supplemented with 0.3 mM DAP. Conjugations proceeded overnight (~12-14 hours) and were scraped from the plate into PBS the next morning. Conjugation mixtures were again gently spun down (1006×g) and washed twice in PBS to remove residual DAP. Approximately 100 μl of this mixture (and 1:10, 1:100 dilutions) was plated onto selective antibiotic plates and incubated 2-3 days to obtain transconjugant colonies. Transconjugants were passaged again on selective media and confirmed by PCR amplification of a plasmid sequence and visible fluorescence, when appropriate. For the initial broad-host-range plasmid screen, transconjugants were further verified by plasmid reisolation and electroporation into *E. coli* DH5α cells. To determine conjugation efficiency, mating mixtures were serially diluted and plated on selective and non-selective plates. Conjugation frequency was calculated as the number of fluorescent transconjugant CFUs on selective plates per total CFUs on non-selective plates

**[0215]** BTK Construction

**[0216]** Construction of the BTK backbone was carried out with Gibson assembly (Gibson D G et al., 2009, Nat Meth., 6:343-345) following established protocols. New part plasmids were constructed using a previously published BsmBI assembly protocol for the yeast toolkit (YTK) (LEE M E et al., 2015, ACS Synth Biol., 4:975-986) with inserts synthesized as doublestranded DNA gBlocks (IDT). New parts were cloned into the pYTK001 entry vector. The BTK kit uses the entry vector plasmid, connector parts (Type 1 and Type 5), and part sequence overhangs of the YTK. In contrast to the YTK, Type 3 parts of the BTK include a stop codon, as the Type 4 terminators do not include a stop codon. A list of BTK plasmids is available in FIG. **18**.

**[0217]** Measuring BGM GFP In Vitro

**[0218]** To measure fluorescence, 50 μL of ~0.2 OD bacterial cultures were pooled on B-COL agar plates and incubated for 48 hours. Cells were scraped into PBS and then loaded into wells of a 96 well plate to measure fluorescent excitation using a Tecan Spark 10M multimode microplate reader at excitation/emission wavelengths of 485/535. Fluorescent readings were corrected with blank values, and then normalized by OD. Gain was set manually and consistent throughout experiments.

**[0219]** Flow Cytometry Analysis of GFP Expression

**[0220]** As with plate reader measurements, 50 μL of ~0.2 OD bacterial culture was pulled onto BCOL agar plates and incubated for 48 hours. Bacteria were scraped into PBS, washed, and then gently spun down (1006×g for 5 minutes). Cells were resuspended vigorously to disrupt any biofilm, and then diluted to ~0.1 OD in HPLC-grade water. Cells were counterstained with SYTO 17 red nucleic acid stain (Thermo-Fisher), and then ran samples on a BD LSR-Fortessa SORP Flow Cytometer. Data were acquired with FACSDiva v6.1.3, and then analyzed with FlowJo v10.4.2. All samples were run under identical conditions. GFP-A voltage was consistent throughout experiments. Nonfluorescent controls were used to determine forward-scatter, side-scatter, and APC-A (counterstain) gates that were then set individually for each species.

**[0221]** Tn7-Transposition in *B. apis*

**[0222]** For the chromosomal insertion of gfp into *B. apis*, a tri-parental mating was performed with *B. apis*, *E. coli* MFDpir with pTNS2, and *E. coli* MFDpir with pTN7-PA1-gfp-kan in an 8:1:1 ratio. Conjugation proceeded for 12 hours, and transformed *B. apis* was selected with kanamycin as in biparental conjugation.

**[0223]** Broad-host-range dCas9 plasmids are created by BsmBI assembly of 3 parts plasmids containing: (1) the sgRNA transcriptional unit (pBTK615), (2) the dCas9 transcriptional unit (pBTK614), and (3) the broad-host-range backbone with ConLE and ConRE connector sequences (pBTK527a). To repress gfp expression in *B. apis*, the gfp nontemplate strand was targeted by using the N20 sequence: 5'-CGTCTAATTCCACGAGGATT (SEQ ID NO:7). The sgRNA plasmid can be retargeted using MEGAWHOP cloning (Miyazaki K et al., 2011, Meth Enz., 498:399-406). Briefly, in MEGAWHOP cloning a double-stranded PCR product containing the sequence change to be introduced, but otherwise identical to a portion of the plasmid, is used as a "megaprimer" to re-amplify the whole plasmid in a second PCR reaction. Because the sgRNA targeting sequence is short, it is possible to include a new target sequence flanked by 20 bp of homology to the plasmid on either side in one of the primers used in the initial PCR reaction to generate the megaprimer. The fully assembled CRISPRi plasmid (pBTK618) was conjugated into *B. apis* with chromosomally integrated PA1-gfp, and GFP fluorescence was measured as above.

**[0224]** Chromosomal Disruption using Cas9 and Homologous Recombination

**[0225]** Plasmid pBTK601 contains Cas9 driven by the kanamycin resistance gene promoter on the broad-host-range backbone. This plasmid was conjugated into *S. alvi* wkB2, *G. apicola* wkB7, and *B. apis* PEB0150 and maintained with spectinomycin. The CP25-driven sgRNA is on plasmid pBTK615 and can be retargeted using MEGA-WHOP cloning (Miyazaki K et al., 2011, Meth Enz., 498:399-406). A full description of homology donor plasmid assembly is available in FIG. **23**. Briefly, a genomic homology segment upstream of the gene of interest to disrupt or replace is amplified with Type 2 part overhangs, and a downstream genomic homology segment is amplified as a Type 4 part. Upstream homology, antibiotic resistance cassette (Type 3), and downstream homology are combined in a single BsaI reaction with ConLE and ConRE to form a Stage 1 assembly of the replacement cassette. The final BsmBI assembly includes: (1) the sgRNA plasmid, (2) the replacement cassette plasmid, and (3) pBTK599 (R6K suicide plasmid backbone). This final assembly must be transformed into pir$^+$ strains, such as EC100D or MFDpir.

**[0226]** Efficiency of Chromosomal Disruption with and without Cas9

**[0227]** Recipient BGM strains (wkB2, wkB2::pBTK601, wkB7, wkB7::pBTK601, PEB0150, PEB0150::pBTK601) were grown on B-COL plates for 48 hours prior to conjugation. Donor *E. coli* strains were grown in liquid culture overnight prior to conjugation. Donor and recipients were washed in PBS and mixed in a 1:9 ratio (by OD), and 100 μL was plated on B-COL+0.3 mM DAP media for overnight conjugation. After 14 hours, the entire conjugation mixtures were scraped into PBS and washed twice to remove residual DAP, and dilutions were plated on selective agar plates (B-COL+Kanamycin 20 μg/mL) and nonselective agar plates (B-COL). Efficiency of gene disruption was calculated as (# of transconjugant cells)/(# of total cells). To identify single-crossover and double-crossover mutants, a series of PCR reactions were conducted as described in FIG. **23**. Briefly, transconjugants were screened for the appropriate upstream and downstream junctions with colony PCR. Potential double-crossover mutants were then further screened for the size of the disrupted region, and loss of the suicide plasmid backbone.

**[0228]** Laboratory Care of Honey Bees

**[0229]** Microbiota-free bees were obtained and raised using methods described previously (Kwong W K et al., 2014, Proc Natl Acad Sci USA., 111:11509-11514). Briefly, pupae were pulled under sterile conditions from brood combs obtained from outdoor hives. These pupae emerged in a sterile incubator (becoming newly emerged adult workers) and were then sorted into individual cup cages for further development in the laboratory. Prior to inoculation, newly emerged workers were allowed to feed on sterile irradiated pollen (Betterbee) and 50% sucrose solution ad libitum. For any individual experiment, all pupae were obtained from the same hive. When raised in this manner, *Apis mellifera* workers remain uncolonized by core BGM bacteria species and show very low levels of environmental

bacteria in their guts (Powell J E et al., 2014. Appl Environ Microbiol., 80:7378-7387). It is critical to pull the pupae from frames at an early stage. before the mouthparts have hardened, as later pupal stages will begin to ingest hive material and may be colonized

**[0230]** Mono- and Co-Innoculation of Engineered BGM into Honey Bees

**[0231]** After obtaining newly emerged workers, bees were chilled at 4° C. for 30 minutes and then coated in sugar syrup containing resuspended bacterial inoculum, transferred to cup cages, and allowed to groom each other. The inoculum generally contained 200 μL of OD ~0.1 bacterial suspension combined with 800 μL of 1:1 sucrose:water solution. Approximately 30 μL of this solution per bee was used for inoculations (corresponding to $10^7$ bacteria per bee to ensure robust inoculation). Plate counts of the inoculum were used to confirm concentrations.

**[0232]** In Vivo Imaging of Bacterial Burden Using E2-Crimson

**[0233]** To visualize in vivo expression of E2-Crimson in living bees. a Syngene G:Box Chemi XX6 gel doc system was used. Bees were chilled on ice for 30 min to minimize movement. then imaged using manufactuers recommended instructions for far-red fluorescent probe visualization: "Red LED" light source and "Filter 705M" emission filter. All bees were imaged under identical conditions: 5 minutes exposure time for whole bee and 30 seconds for bees with dissected guts. Images were saved as TIFF files for further analysis in FIJI (Schindelin J et al., 2012, Nat Meth., 9:676-682). In FIJI. fluorescence intensity was mapped to the "mplmagma" scale. A representative bee for each condition is shown. No further image manipulation was performed. Different scales are used for comparing fluorescent *S. marcescens* and fluorescent BGM species due to the increased fluorescent protein production and titer of *S. marcescens*.

**[0234]** Confocal Fluorescence Microscopy

**[0235]** Fluorescent images were obtained at the UT ICMB Microscopy core on a Zeiss 710 Laser Scanning Confocal microscope. Bees were chilled and then dissected to expose rectum. ileum, and midgut. Without puncturing the gut, the entire gut compartment was transferred to an Ibidi μ-Dish 35 mm (CAT #81156) and then placed on the microscope. Images were taken with a 20× objective and tiled using Zeiss software. Z-stack 2-channel fluorescent images were taken and combined using Imaris software. Intensity on individual channels was false colored to correspond to species-specific coloring. Display intensity of individual channels was scaled linearly to aid in visualization of different species, but no further transformations or background reduction was used.

**[0236]** qPCR to Assess Colonization of StaA Mutant

**[0237]** Absolute quantification of 16S rRNA gene copies specific to *S. alvi* was performed as described previously (Powell J E et al., 2016, Proc Natl Acad Sci USA., 113: 13887-13892). Cohorts of newly emerged bees were hand-fed with equal amounts (~$10^4$ CFU/bee) of either wild-type *S. alvi* or the staA mutant. Control bees were maintained identically but remained uninoculated. After five days, five bees from each group were dissected and DNA was isolated from individual bee guts using the cetyltrimethylammonium bromide (CTAB) extraction method outlined previously (Powell J E et al., 2014. Apl Environ Microbiol., 80:7378-7387). After extraction, *S. alvi*-specific primers were used for quantitative PCR and absolute quantification based on 10-fold dilution of the target sequence in a pGEM-T plasmid vector. Reactions were run in triplicate.

**[0238]** Quantification and Statistical Analysis

**[0239]** All data processing and statistical analyses were done in R. Kruskal-Wallis rank sum tests were used to assess significance in the dCas9 gene repression experiment and the Cas9-assisted genome modification experiments.

**[0240]** The results of the experiments are now described.

**[0241]** BTK Plasmids Function in Diverse Bacterial Species Found in the Bee Gut

**[0242]** It was sought to explore the host range of the RSF1010 origin used as a basis for the BTK in the context of a larger set of bee-associated bacterial strains. Simultaneously, it was needed to identify antibiotic resistance genes able to function in each bacterial strain. To do so, three BTK plasmids were constructed. each with a different antibiotic resistance marker and encoding GFP driven by the PA1 promoter: pBTK501 ($Amp^R$). pBTK519 ($Kan^R$). pBTK520 ($Spec^R$). Biparental matings were performed between *E. coli* MFDpir donors containing each plasmid and bee gut-associated strains (see Methods). Stable transconjugants were obtained for all of the Gram-negative strains tested with at least one of these three plasmids. as verified by further passaging on antibiotic-containing media, PCR amplification of plasmid sequences, and GFP expression (FIG. **10**A). Successfully transformed bacterial species include Alpha-, Beta-, and Gammaproteobacteria and strains isolated from different bee species (*A. mellifera, Bombus terrestris. Bombus impatiens*, and *Bombus pensylvanicus*). Several of the bacterial species (*S. alvi. G. apicola. B. apis*, and *Parasaccharibacter apium*) are phylogenetically distant from any established model organisms and have no previously reported genetic tools. Transfer of the BTK plasmids was efficient, with $>10^{-3}$ transconjugants per CFU for four diverse bacterial species (FIG. **10**B)

**[0243]** Identifying Functional Promoters in BGM Species

**[0244]** While some sequence features of transcriptional promoters are conserved across bacterial species, there is no guarantee that promoters designed to function in model organisms will function effectively in new bacterial isolates from a natural community of interest (Whitaker W R et al., 2017, Cell, 169:538-538). The BTK includes BHR promoters and RBS combinations as Type 2 parts that can be used to build plasmids to identify functional sequences for driving protein expression in new bacterial hosts. The function of the BHR promoters PA1 (pBTK501). PA2 (pBTK509), PA3 (pBTK510), and CP25 (pBTK503) were compared in *S. alvi* wkB2. *G. apicola* wkB7. *B. apis* PEB0150. and *S. marcescens* N10A28, all isolated from honey bee gut communities. Promoters PA1. PA2. PA3 are strong early promoters from bacteriophage T7 (Siebenlist U., 1979, Nucleic Acids Res., 6:1895-1907). The synthetic CP25 promoter was originally designed to function in *Lactococcus* strains (Jensen P R et al., 1998, Appl Environ Microbiol., 64:82-87). and the BTK includes other promoters from this series.

**[0245]** Using flow cytometry, fluorescent protein expression from these promoters (FIG. **10**A) was characterized. These promoters display significant variability in activity across strains when they are all tested with the same RBS. As expected, the promoter-RBS pairs function most strongly in *S. marcescens*, which is most closely related to *E. coli*. In the other BGM strains. expression was weaker. but there was a signal above background for most promoters that were tested with this RBS. Fluorescence is generally lower in *S.*

*alvi*, *G. apicola*, and *B. apis* than it is in *E. coli*. In *E. coli* the distributions of fluorescent per cell for the PA2, PA3, and CP25 promoters are noticeably bimodal. This may be an intrinsic property of the promoter or due to the accumulation of "broken" plasmids with mutations that inactivate burdensome GFP expression (Sleight S C et al., 2010, J Biol Eng., 4:12). In BGM strains, with the exception of CP25 in *G. apicola*, these distributions are unimodal, indicating consistent fluorescent expression across single cells. PA3 expression was strong in *S. alvi*, and this observation was used to design a constitutive E2-Crimson-expressing plasmid (pBTK570) to test expression in vivo, as described in later sections. Validation of additional parts (E2-Crimson, NanoLuc, and other CP-series promoters) is available in FIG. 20, FIG. **21**, FIG. **22**.

**[0246]** Inducible Gene Expression in BGM Species

**[0247]** Induction systems are required for the temporal control of gene expression, and are useful for testing the functional roles of microbes in gut environments (Lim B et al., 2017, Cell, 169:547-558). Two lacI induction systems were tested: one simple system composed of a modified CP25 promoter with lacO sites and a more complex system that uses T7 RNA polymerase (T7 RNAP). IPTG-induction of these systems was tested in *E. coli* MFDpir, *S. alvi* wkB2, *G. apicola* wkB7, *B. apis* PEB0150, and *S. marcescens* N10A28. The simple system (pBTK552) showed robust induction of GFP in all strains tested (FIG. **11B**). Interestingly, *G. apicola* GFP expression with this system surpassed that of *E. coli* and *S. marcescens*.

**[0248]** For the T7 RNAP system, two transcriptional unit plasmids (pBTK549d, pBTK541) were built, one bearing lacI driven by the CP25 promoter and T7 RNAP under control of the inducible lac promoter and the other with GFP expressed from a T7 promoter with lacO sites, and combined them into a composite plasmid (pBTK550d), (FIG. **11C**). Expression was strong in *S. marcescens* N10A28 and *E. coli* MFDpir, with maximal GFP expression after induction surpassing the simpler system in which lacI directly regulates GFP expression. However, in *G. apicola* wkB7, *S. alvi* wkB2, and *B. apis* PEB0150, a weaker induction of GFP was seen compared to the simpler system. The cause of this weak expression is unknown. It may be due to poor transcription from the lac promoter driving T7 RNAP or to an intrinsic incompatibility between T7 RNAP and the intracellular environment in the BGM species tested. In all strains, the inducible T7 RNAP construct showed appreciable background expression when not induced.

**[0249]** CRISPRi Repression of Chromosomal Gene Expression in *Bartonella apis*

**[0250]** BTK was used to suppress gene activity in a BGM bacterium. Catalytic mutants of Cas9 (dCas9) have been used to reduce transcription of target genes, an approach termed CRISPR interference (CRISPRi), in diverse mammalian and bacterial systems (Barrangou R et al., 2017, Nat Microbiol., 2:1-9). To expand this approach to new non-model bacterial species, a modified dCas9 system was established in which targeting is achieved by a BTK part encoding a small guide RNA (sgRNA) (FIG. **12A**) (Bikard D et al., 2013, Nucleic Acids Res., 41:7429-7437). To test the system, the sgRNA was targeted to a PA1-driven GFP gene in *B. apis* PEB0150, which was inserted into the chromosome using Tn7-based integration (Choi K H et al., 2005, Nat Meth., 2:443-448). GFP expression was significantly reduced in the presence of a sgRNA targeted to the GFP sequence (FIG. **12B**). Coupled with the induction system, this ability to repress a target gene enables functional studies of essential genes that cannot be disrupted entirely.

**[0251]** Cas9-Assisted Gene Disruption in the BGM

**[0252]** Gene disruption is an important tool for establishing gene function and for studying interactions between genes. After identifying functional antibiotic cassettes in the earlier plasmid-replication screen, it was attempted to use homologous recombination to disrupt chromosomal genes in the BGM strains. To improve the efficacy of targeted gene disruption, a two-step approach was implemented based on using Cas9 cleavage for chromosomal modifications (Barrangou R et al., 2017, Nat Microbiol., 2:1-9) (see Methods). In step one, Cas9 is introduced into a cell on the BTK backbone (pBTK601) without any targeting sgRNA. In step two, a second round of conjugation is used to deliver a suicide plasmid with the replacement cassette (~1000 bp homology flanking a functional antibiotic resistance gene) and the sgRNA targeting the desired chromosomal location. The suicide plasmid is made with Golden Gate assembly using repurposed Type 2-4 overhangs and an R6K origin of replication (FIG. **13A**, FIG. **13B**). The sgRNA can be retargeted using MEGAWHOP cloning (Miyazaki K, 2011, Meth Enz., 498:399-406) (see Methods). A detailed description of suicide plasmid assembly and validation of mutants is shown in FIG. **23**. It was expected that Cas9 cleavage might facilitate recombination into the chromosome and that it would also select against single-crossover integrations, in which the suicide plasmid backbone is incorporated into the chromosome, because they preserve the cleavage site, whereas double-crossover integrations result in replacement of the targeted gene sequence with just the antibiotic resistance cassette and delete the cleavage site.

**[0253]** To test the utility of this scheme, it was attempted to generate gene disruptions in three BGM species. In *S. alvi* wkB2 staA (SALWKB2_RS11470) was targeted, an adhesion gene previously implicated in a genome-wide screen as important for gut colonization (FIG. **24**) (Powell J E et al., 2016, Proc Natl Acad Sci USA, 113:13887-13892). In *G. apicola* wkB7 acetate kinase ackA was targeted (A9G17_RS12535) (FIG. **25**), and in *B. apis* PEB0150 nitrate reductase narG (PEB0150_RS00755) was targeted (FIG. **26**). Homology regions were designed to be internal to each coding sequence, so that even single-crossover events would disrupt gene function. For *S. alvi* wkB2, the multi-step system showed higher efficiency compared to basic homologous recombination not using Cas9. In the presence of Cas9, wkB2 mutants were obtained more frequently and were more often double-crossover mutants (FIG. **13D**). In contrast, *B. apis* PEB0150 showed relatively high gene disruption efficiency even in the absence of Cas9, and the Cas9 system had little effect on improving the number of double-crossover mutants (FIG. **13C**). In *G. apicola* wkB7 the Cas9 was also not helpful, and no double-crossover mutants were obtained (FIG. **13E**). The *G. apicola* wkB7 mutants isolated showed irregular PCR amplification at the expected junctions (FIG. **25**), indicating that it could not effectively disrupt ackA, perhaps because it is an essential gene in this species. These experiments validated this general approach to gene disruption in multiple BGM species.

**[0254]** Engineered Strains Colonize Bees and can be Directly Visualized in the Ileum

**[0255]** Next the ability of engineered BGM strains to colonize newly emerged worker bees removed from the hive before they acquire a normal microbiota was tested. Previously, *S. alvi* and *G. apicola* have been visualized in bees using fluorescent in situ hybridization (Martinson V G et al., 2012, Apl Environ Microbiol., 78:2830-2840). However, this technique can only be used at one time point because it requires sacrificing the bee. In contrast, fluorescent reporter strains can be used to non-destructively estimate bacterial abundance and observe how bacterial community structure changes over time in live bees. Previous studies have examined how *S. alvi* colonizes the honey bee gut (Kwong W K et al., 2014, Proc Natl Acad Sci USA., 111:11509-11514; Engel P et al., 2013, J of Api Res., 52:1-24), but colonization by *G. apicola, B. apis*, and *S. marcescens* has not been investigated (Raymann K et al., 2017, PLoS Biol., 15:e2001861; Burritt N L et al., 2016, PLoS One., 11:e0167752). Newly emerged workers with ~$10^4$ CFU per bee of either *S. marcescens* N10A28 or *S. alvi* wkB2 was inoculated, each carrying a constitutively expressed E2-Crimson fluorescent protein (pBTK570). After 5 days, bees from each group were dissected and their guts were examined (FIG. **14**). Fluorescent bacteria were successfully imaged directly in guts without preparation or fixation, preserving natural community structure. *S. marcescens* N10A28 shows robust colonization in all gut compartments, while other species show spatially restricted colonization. As previously reported, *S. alvi* wkB2 robustly colonizes the ileum, with little colonization in the midgut and rectum. Additionally, co-inoculations with *S. alvi* wkB2 and either *B. apis* PEB0150 or *G. apicola* wkB7 engineered were performed to express GFP (pBTK520). The guts of colonized bees were again dissected and fluorescently imaged in vivo co-colonization of these defined communities (FIG. **15**A, FIG. **15**B). While both *B. apis* and *G. apicola* are found in the ileum colocated with *S. alvi*, they also colonize the rectum, in contrast to *S. alvi*.

**[0256]** *Snodgrassella* StaA Contributes to Gut Colonization In Vivo

**[0257]** Finally, it was sought to validate the usefulness of the BTK for disrupting specific genes in BGM species in order to investigate their function. StaA belongs to a family of YadA-like adhesion proteins important for colonization and pathogenicity in multiple host-associated species (Linke D et al., 2006, Trends Microbiol., 14:264-270). These trimeric autotransporter proteins localize to the bacterial membrane and form "lollipop" structures that allow bacteria to adhere to epithelial cells (Ribet D et al., 2015, Microbes Infect., 17:173-183; Tahir El Y et al., 2001, Int J Med Microbiol., 291:209-218). Orthologs of these genes are found in multiple *S. alvi* genomes, including those from honey bee- and bumble bee-associated strains (Kwong W K et al., 2014, Proc Natl Acad Sci USA., 111:11509-11514). In the previous works screening a transposon mutant library identified staA (SALWKB2_RS11470) as necessary for the fitness of *S. alvi* during gut colonization (Powell J E et al., 2016, Proc Natl Acad Sci USA., 113:13887-13892). However, it was not able to isolate a mutant from the library with a transposon disrupting staA, and thus it could not fully characterize and validate the role of this gene. Using the BTK, a ΔstaA mutant was generated in *S. alvi* wkB2 (as described above). The ΔstaA mutant and a wild-type control was labeled with a BTK plasmid expressing E2-Crimson (pBTK570) to assess the effects of disrupting this gene in the context of the bee gut. The wkB2 ΔstaA mutant shows reduced colonization efficacy compared to a wild-type control, as measured by qPCR of *S. alvi* 16S rRNA gene copies (FIG. **15**C). The colonization pattern of this mutant in terms of its localization within the gut (FIG. **15**D) is distinct from that of wild-type *S. alvi* (FIG. **14**E). After 6 days, the mutant does not form the contiguous, robust colonization of the ileum wall seen for the wild type strain. Instead, colonization is apparently restricted to small patches, while the majority of the ileum remains uncolonized.

### Example 5. Engineering Bee Guts with Symbiont-Produced dsRNA

**[0258]** These experiments provide data to demonstrate the use of engineered gut bacteria (*Snodgrassella alvi*) to manipulate bee gene expression, behavior, and immune response.

**[0259]** Engineered Symbionts Reduce Viral Replication in Bees.

**[0260]** Emerged bees were engineered with gut bacteria harboring one of six plasmids: (1) "NR"—this is an empty plasmid control, expressing no double-stranded RNA; (2) "dsGFP"—this is a dsRNA control, expressing GFP dsRNA; (3) "DWV T1"—dsRNA with DWV target 1, (4) "DWV T2"—dsRNA with DWV target 2; (5) "DWV T3"—dsRNA with DWV target 3, (6) "DWV T5"—dsRNA with DWV target 5.

**[0261]** After 5 days, bees from each condition were hand fed 5.7e6 viral particles of DWV.

**[0262]** After 48 hours, the fat bodies and hemolymph of the bees was dissected, and RNA was isolated from this tissue. cDNA was generated and then absolute qPCR quantification was performed to determine viral genomes (FIG. **27**).

**[0263]** Engineered Symbionts Reduce Expression of Target Genes Across Body Tissues.

**[0264]** Emerged bees were engineered with gut bacteria harboring one of three plasmids: (1) "NR"—an empty plasmid control, expressing no double-stranded RNA; (2) "dsGFP"-a dsRNA control, expressing GFP dsRNA; (3) "dsTH"—dsRNA targeting Tyrosine Hydroxylase (TH).

**[0265]** At two time points, Day 4 and Day 8, RNA was isolated from the gut and head. cDNA was generated, and relative quantitation was performed to determine expression (FIG. **28**).

**[0266]** Engineered Gut Bacteria Alter Host Behavior.

**[0267]** Behavioral assays were performed on emerged bees engineered with gut bacteria harboring one of three plasmids: (1) "NR"—an empty plasmid control, expressing no double-stranded RNA; (2) "dsGFP"-a dsRNA control, expressing GFP dsRNA; (3) "dsTH"—dsRNA targeting Tyrosine Hydroxylase (TH).

**[0268]** This behavioral assay measured aversive learning, the ability of bees to learn to associate a scent with a shock. Bees colonized with bacteria expressing dsTH learned significantly more slowly than bees colonized with control bacteria (dsGFP) (FIG. **29**).

**[0269]** Suppression of *Varroa* with Engineered Gut Bacteria.

**[0270]** *Varroa* mites are the major bee pest, and suppressing them is a promising application of dsRNA technology. Emerged bees are engineered with gut bacteria expressing at

least one dsRNA with sequence specific to essential *Varroa* genes. Targets are present individually or as a composite of multiple targeting sequences. This dsRNA moves into the bee hemolymph and when *Varroa* mites feed, they will induce a lethal self-targeting RNAi response. This is different from the previous described RNAi work because the RNAi response is in the *varroa* mite, not the bee.

**[0271]** Experimental Design: Colonize bees with gut bacteria expressing *Varroa* targeting constructs (dsVarroa), capture live *Varroa* mites from hives, expose defined number of *varroa* mites to bees, and measure mortality of *varroa* mites fed on dsVarroa bees. Compare to control NR and dsGFP colonized bees.

**[0272]** Suppression of *Nosema* with Engineered Gut Bacteria.

**[0273]** *Nosema* is another pathogen of bees. Suppression of *Nosema* in bees is approached in two ways. (1) use symbiont-mediated RNAi to alter and strengthen the bee immune response, and (2) trigger lethal RNAi in *nosema*

**[0274]** Experimental Design: Colonize bees with gut bacteria expression lethal *nosema* RNAi or immune-strengthening RNAi, expose bees to a defined number of *nosema* spores, after 48-72 hours, count the number of *nosema* spores. Compare to control bees with no RNAi or control dsGFP RNAi

**[0275]** Suppression of Small Hive Beetle with Engineered Gut Bacteria

**[0276]** The small hive beetle is another major pest in beekeeping. Unlike DWV, *Varroa*, and *Nosema*, however, it does not attack bees directly, but instead it feeds on hive components. Therefore, to target this insect with bacterially induced RNAi, the approach is different. For this a different bacterial species that can live in hive materials, such as *Parasaccharibacter apium* or *Lactobacillus* sp. is used. These bacteria are transformed to produce dsRNA identical to an essential hive beetle gene. When these bacteria are consumed by the hive beetle, they would kill the beetles or impair their reproduction.

## Example 6: Persistence of Engineered Bacteria in Host

**[0277]** Two separate cohorts of bees were obtained from different hives and allowed to emerge in lab. Within 24 hours after emergence, bees were inoculated with ~1e04 CFU of *S. alvi* wkB2 engineered to constitutively express GFP (plasmid pBTK520) and separated into individual cups with 5-10 bees. Bees were fed sucrose solution supplemented with 60 ug/mL spectinomycin. Every five days, five to eight bees from each cohort were dissected and their entire gut contents plated on selective media with spectinomycin to estimate the CFU of engineered *S. alvi* remaining in their gut.

**[0278]** Across all three time points, no significant difference existed between time points or hives in this experiment. Over 99% of colonies were fluorescent in each sample, indicating the genetic device remained functional (FIG. 30). This shows that engineered *S. alvi* wkB2 persistently colonizes and functions in the bee gut despite different bee genetic backgrounds. Bees were fed spectinomycin to maintain the plasmid, but other strategies could be used to ensure maintenance and function of the device in engineered strains (ie, chromosomal incorporation).

**[0279]** Engineered *S. alvi* persists and functions over the measured lifetime of the bee in laboratory environments. Without being bound by theory, it is expected that the engineered gut bacteria would persist over the lifetime in bees in their natural environment, as well.

## Example 7: Sequences

**[0280]**

| SEQ ID NO | Description |
|---|---|
| SEQ ID NO: 1 | pMMB67EH |
| SEQ ID NO: 2 | pBTK520 |
| SEQ ID NO: 3 | pBTK570 |
| SEQ ID NO: 4 | pBTK502 |
| SEQ ID NO: 5 | pBTK501 |
| SEQ ID NO: 6 | pBTK599 |
| SEQ ID NO: 8 | pBTK401 |
| SEQ ID NO: 9 | pBTK402 |
| SEQ ID NO: 10 | pBTK403 |
| SEQ ID NO: 11 | pBTK599 |
| SEQ ID NO: 12 | CP25 (lacO) promoter |
| SEQ ID NO: 26 | pBTK001 |
| SEQ ID NO: 27 | T7 promoter + RBS |
| SEQ ID NO: 28 | Lac lacO inducible promoter + RBS |
| SEQ ID NO: 29 | CP25 promoter + RBS |
| SEQ ID NO: 40 | CP6 promoter + RBS |
| SEQ ID NO: 41 | CP12 promoter + RBS |
| SEQ ID NO: 42 | CP32 promoter + RBS |
| SEQ ID NO: 43 | PA1 promoter + RBS |
| SEQ ID NO: 44 | PA2 promoter + RBS |
| SEQ ID NO: 45 | PA3 promoter + RBS |
| SEQ ID NO: 46 | CP25 + RBS reverse |
| SEQ ID NO: 47 | dCas9 |
| SEQ ID NO: 48 | rpoC terminator |
| SEQ ID NO: 49 | Bba_B0015 terminator |
| SEQ ID NO: 50 | T7 terminator |
| SEQ ID NO: 51 | pBTK103 |
| SEQ ID NO: 52 | pBTK105 |
| SEQ ID NO: 53 | pBTK107 |
| SEQ ID NO: 54 | pBTK110 |
| SEQ ID NO: 55 | pBTK112 |
| SEQ ID NO: 56 | pBTK113 |
| SEQ ID NO: 57 | pBTK119 |
| SEQ ID NO: 58 | pBTK120 |
| SEQ ID NO: 59 | pBTK121 |
| SEQ ID NO: 60 | pBTK138 |
| SEQ ID NO: 61 | pBTK200 |
| SEQ ID NO: 62 | pBTK202 |
| SEQ ID NO: 63 | pBTK205 |
| SEQ ID NO: 64 | pBTK206 |
| SEQ ID NO: 65 | pBTK209 |
| SEQ ID NO: 66 | pBTK211 |
| SEQ ID NO: 67 | pBTK224 |
| SEQ ID NO: 68 | pBTK229 |
| SEQ ID NO: 69 | pBTK300 |
| SEQ ID NO: 70 | pBTK301 |
| SEQ ID NO: 71 | pBTK305 |
| SEQ ID NO: 72 | pBTK527 |
| SEQ ID NO: 73 | pBTK570alpha |
| SEQ ID NO: 74 | pBTK599a |
| SEQ ID NO: 75 | pBTK501 |
| SEQ ID NO: 76 | pBTK503 |
| SEQ ID NO: 77 | pBTK509 |
| SEQ ID NO: 78 | pBTK512 |
| SEQ ID NO: 79 | pBTK519 |
| SEQ ID NO: 80 | pBTK525 |
| SEQ ID NO: 81 | pBTK529 |
| SEQ ID NO: 82 | pBTK541 |
| SEQ ID NO: 83 | pBTK550a |
| SEQ ID NO: 84 | pBTK552 |
| SEQ ID NO: 85 | pBTK560 |
| SEQ ID NO: 86 | pBTK564 |
| SEQ ID NO: 87 | pBTK569 |
| SEQ ID NO: 88 | pBTK579 |
| SEQ ID NO: 89 | pBTK601 |
| SEQ ID NO: 90 | pBTK614 |
| SEQ ID NO: 91 | pBTK615 |
| SEQ ID NO: 92 | pBTK619 |

-continued

| SEQ ID NO | Description |
| --- | --- |
| SEQ ID NO: 93 | sBFK6ZL |
| SEQ ID NO: 94 | sBFK6ZI |

**[0281]** All of the methods disclosed and claimed herein can be made and executed without undue experimentation in light of the present disclosure. While the compositions and methods of this invention have been described in terms of preferred embodiments, it will be apparent to those of skill in the art that variations may be applied to the methods and in the steps or in the sequence of steps of the method described herein without departing from the concept, spirit and scope of the invention. More specifically, it will be apparent that certain agents which are both chemically and physiologically related may be substituted for the agents described herein while the same or similar results would be achieved. All such similar substitutes and modifications apparent to those skilled in the art are deemed to be within the spirit, scope and concept of the invention as defined by the appended claims.

SEQUENCE LISTING

```
<160> NUMBER OF SEQ ID NOS: 94

<210> SEQ ID NO 1
<211> LENGTH: 8629
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized.

<400> SEQUENCE: 1

agcttggctg ttttggcgga tgagagaaga ttttcagcct gatacagatt aaatcagaac      60

gcagaagcgg tctgataaaa cagaatttgc ctggcggcag tagcgcggtg gtcccacctg     120

accccatgcc gaactcagaa gtgaaacgcc gtagcgccga tggtagtgtg gggtctcccc     180

atgcgagagt agggaactgc caggcatcaa ataaaacgaa aggctcagtc gaaagactgg     240

gcctttcgtt ttatctgttg tttgtcggtg aacgctctcc tgagtaggac aaatccgccg     300

ggagcggatt tgaacgttgc gaagcaacgg cccggagggt ggcgggcagg acgcccgcca     360

taaactgcca ggcatcaaat taagcagaag gccatcctga cggatggcct ttttgcgttt     420

ctacaaactc ttttgtttat ttttctaaat acattcaaat atgtatccgc tcatgagaca     480

ataaccctga taaatgcttc aataatattg aaaaaggaag agtatgagta ttcaacattt     540

ccgtgtcgcc cttattccct tttttgcggc attttgcctt cctgtttttg ctcacccaga     600

aacgctggtg aaagtaaaag atgctgaaga tcagttgggt gcacgagtgg gttacatcga     660

actggatctc aacagcggta agatccttga gagttttcgc cccgaagaac gttttccaat     720

gatgagcact tttaaagttc tgctatgtgg cgcggtatta tcccgtgttg acgccgggca     780

agagcaactc ggtcgccgca tacactattc tcagaatgac ttggttgagt actcaccagt     840

cacagaaaag catcttacgg atggcatgac agtaagagaa ttatgcagtg ctgccataac     900

catgagtgat aacactgcgg ccaacttact tctgacaacg atcggaggac cgaaggagct     960

aaccgctttt ttgcacaaca tgggggatca tgtaactcgc cttgatcgtt gggaaccgga    1020

gctgaatgaa gccataccaa acgacgagcg tgacaccacg atgcctgtag caatggcaac    1080

aacgttgcgc aaactattaa ctggcgaact acttactcta gcttcccggc aacaattaat    1140

agactggatg gaggcggata aagttgcagg accacttctg cgctcggccc ttccggctgg    1200

ctggtttatt gctgataaat ctggagccgg tgagcgtggg tctcgcggta tcattgcagc    1260

actggggcca gatggtaagc cctcccgtat cgtagttatc tacacgacgg ggagtcaggc    1320

aactatggat gaacgaaata gacagatcgc tgagataggt gcctcactga ttaagcattg    1380

gtaactgtca gaccaagttt actcatatat actttagatt gatttctgaa agcgaccagg    1440

tgatcggcgt ggcaagacta gcagcgaaac cgtagaaaga catgatccag ccgaacgcat    1500
```

-continued

```
tggagaactc attcaaattc aagttgacca tagcccggca attccttcca atgcctgccc    1560

atcagcgccg agcccgacgg gtcccactcg tcccggatcc gatcgagccg ggttggatcg    1620

ggtcggtggc tctggtaacg acgagtatgc cgatcccggc tggccgtcct ggccgccaca    1680

tgaggcatgc tccgcgtcct tgcaatactg tgtttacata cagtctatcg cttagcggaa    1740

agttctttta ccctcagccg aaatgcctgc cgttgctaga cattgccagc cagtgcccgt    1800

cactcccgta ctaactgtca cgaacccctg caataactgt cacgcccccc tgcaatcact    1860

gtcacgatcc catgcaatca atgtcaagcc accaaaactg aacccccagc agggccgggg    1920

gatggcgggc tgttggacca atccatccct gattatctca gactaatccc ctaggcgagg    1980

ttatcagcgc ccttgtgggc cgctgctgcc cttgcccaat atgcccggcc agaggccgga    2040

tagctggtcc attcgctgcc ctaggctaca cacggcccca ccgccgcgcg gcaggcggaa    2100

aggcgggcca agcccgctca acccccaccc aaaccccgca gaaatacgct ggagcgcttt    2160

tagccgattc agcggactta accccaaccc gacgggtggc ggcgcgtgtg cagcccagcc    2220

gggcctgtcc aggtcgatca ttcagcaacg atcaccattc tggcgaggcg gacgacagcc    2280

caaggcgcgc tcgtcgtcgc gttcaaggta cgcatccatt gccgccatga gccgatcctc    2340

cggccactcg ctgcagttca ccttggccaa aatcatggcc cccaccagca ccttgcgcct    2400

tgttccgttc ttgcgctctt gctgctgttc ccttgcccgc tcccgctgaa tttcggcatt    2460

gattcgcgct cgttgttctt cgagcttggc cagccgatcc gccgccttgt tgctcccctt    2520

aacccatttg aaacccaatc gttcacgtgc tgtcacgtcg gatcccatgg aggcacagcg    2580

gaggcaatcc agcccatccc ttgtagggga gggcgcaatc aacggttctt attcgagacc    2640

ctggcctaac ggccacccct cgggcggtgc gctcgccgag ggccattgca tggagccgaa    2700

aagcaaaagc aacagcgagg cagcatggcg atttatcacc ttacggcgaa aacgggcagc    2760

aggtcgggcg gcccatcggc cagggccaag gccgactaca tcccagcgca aggcaagtat    2820

gcccgcgaca tggatgaagt cttgcacgcc gaatccgggc acatgccgga gttcgtcgag    2880

aggcccagcg aatcctggga tgctgcagcc atgtacgacc gagccaatgg gaggcctgtc    2940

aaggaggtcg aatttgaccc gacggcagcg atgcccatcg accagcagcc ggcgcaggcg    3000

tccgagttcg cccagcaccc gacggtgccc gagcgcctgc cgtatacgct ggccatccat    3060

gccggtggcg gcgagaaccc gcaccgccac ctgatgatcc ccgagcggat caatgacggc    3120

atcgagcggc ccgccgctca gtggttcaag cggtacaacg gcaagacccc ggagaagggc    3180

ggggcacaga agaccgaagc gctcaagccc aaggcatggc ttgagcagac ccgcgaggca    3240

tgggcagacc atgccaaccg ggcctcagcg agggcaggcc aagacgaacg aattgaaacc    3300

agcacaattg agggcaaggg aatcgagcgc atgcccggtg ttaccctggg gacgaaagtg    3360

gtggagatgg aaggccgggg catccgcacc gacccggcag acgtggccct gaacatcgac    3420

accgccaacg cccagatcat cgacctacag gaataccggc aggcaataga ccatgaacgc    3480

aatcgacaga gtgaagaaac cccagaggca caacgagtta gcggagcaga tcgaaccgct    3540

ggcccagagc atggcgacac tggccgacga agcccgccag gtcatgagcc agacaagcaa    3600

ggccagcagc gagcaggcgg aggcaggcct gacagcaacg agccagaccg gggcggcatg    3660

ggtggagccg gccaaagagc tcgcggaggc agcgcccgag gtgagcagcg ccgcgcagag    3720

agcccggagc gagtcgaggg ggtggcacgg gacgcaatgg atacccgtgc tgctggattc    3780
```

-continued

aatgatgccc aaggcggtgc tgctgatcgc atcgccgatc ttgctcgacc tgcagcaact 3840

gacaaacgcg gacggatcga tatggctgcg attggcggcc agatgaagcc agcaaggact 3900

ttgcaggcca taggccgaca gctcaaggcc atgggctgtg agcgcctcga tatcggcgtc 3960

agggacgcac ccacgggcca gatgatgaac cgggaatggc cagccgccga agtgctccag 4020

aacacgccat ggctcaagcg gatgaatgcc caggccaatg acgtgcatat caggcccgcc 4080

gagcaggagc ggcatggtct ggtgctggtg gacgacctca gcgagcttga cctggatgac 4140

atgaaagccg agggcaggga gcctgcactg gtacgcgaca aacgccagcc gacctatacg 4200

gacccggtca aggtcgacga agccgcaggc ggtgaacttc gggggcagct tgcccggacg 4260

ctggccagcg agtacgacgc cgacccggcc agcgccgaca gccgccacta tggccgcttg 4320

gcggccttca ccaaccgcaa ggacaagcac accacccgcg ccggccatca gccgcgggtg 4380

ctgctgcgtg aatccaaggg caagaccgcc accgctggcc cggcgctggt gcagcaggct 4440

ggccagacga tagacaaggc aacgcggccg caggacaacg aacgcaggct ggccaacatc 4500

gaccgacccg agcggcagcc tagccgcccc aggcgcaacg agctggcacg gtaccgcagc 4560

gagatggccg ggctggtcaa gcgcctcggc catgacctca gcaagcgcga ctttatcgcc 4620

gcgcagaagc tggccagccg ggggccgcag tgcgagcgaa tcggcaaggc catggccgag 4680

gccagccccg cgctggcaga gcgcaagccc ggccacgaac cggattacat cgagcgcacc 4740

gtcagcaagg tcatgggtct gcccagcgtc cagcttgcgc gggccgagct ggcacgggca 4800

aaggcaaccc gaccgagcgc aatggaacgc ggcggcaacg atttcagcct gtacgcattg 4860

agtcggtacc aacgcatgcc atacctatgc tactcaagca cagcacgggg ttttacggcc 4920

tacgaaaaaa gcgcctcagg gtcggtctac ctgatcaaaa gtgacaaggg ctattcgttg 4980

ccgcgtggcc ttggctatac gtcaaacaag gccgagcctg gccgcctttc agtcgctgat 5040

atggccagcc ttaaccttga cggctgcacc ttgtccttgt tccgcgaaga caagcctttc 5100

ggccccggca agtttctcgg tgactgatat gaaagaccaa aaggacaagc agaccggcga 5160

aatgctggcc agcccagccg atgtacgcca agcgcgatcc gccgagcgcc tgcagggacc 5220

agggatgcgc cagcgcaacg tatggctgcc cgccgacgca tacgacggcc tgcgcgagtg 5280

cctggaagaa ctcagagcgg cgcaggcggc gggtagtgac ccgcgcagcg cctaaccacc 5340

aactgcctgc aaaggaggca atcaatggct acccataagc ctatcaatat tctggaggcg 5400

ttcgcagcag cgccgcccac gctggactac gtttggccca acatggtggc cggtacggtc 5460

ggggcgctgg tgtcgcccgg tggtgccggt aaatccatgc tggccctgca actggccgca 5520

cagatcgccg gagggccgga tatgctgggc gtgggcgacc tgcccaaagg accggcgatc 5580

tacctgcccg cagcagaccc gaccaacgcc attcacccac gcctgcaagc cattgggcag 5640

caccgcagcg ccgaggaacg gcaagccgtg gctgacggcc tgctgatcca gccgctgatc 5700

ggcagcctgc ccaacatcat ggccccggag tggttcgacg gcctcaagcg cgccgccgag 5760

ggccgccgcc tgatggtgct ggacacgctg cgccggttcc acatcgagga agaaaacgcc 5820

agcggcccca tggccaggcc catcggtcgc atggagccca tcgccgccga taccgggtgc 5880

tatcacgtgc tactgcaccc tgcccagccg ggcgccgacc tgctggagcg aggcgcaccg 5940

cagcaggcca gccggggcag ctcggcactg gtcgataaca tccgctggca gtcctacctg 6000

tagagcatga aacgcgacga ggccgaggca tgggcgtgcg aagacgaacc gagccggttc 6060

-continued

ttcgccagct tagggtgcag acaggcaccc tatggcgacc agttcgatgc taggcggttc 6120

aggcggcatg aaggcggggc gatcaagccc gacgtgatgg agaggcagcg accgagcacg 6180

ggggtgcccc gtggtgaagc ctaagaacaa gcacagcctc agcoacgtcc ggcacgaccc 6240

ggcgcactgc ctggccccgg gcctgctcgg tgccctcaag cgggcgcagc gcaagcgcag 6300

caagctggac gtgacgtatg actacggcga cggcaagcgg atcgagttca gcggcccgga 6360

gccgctgggc gctgatgatc tgcgcatcct gcaagggctg gtggccatgg ctgggcctaa 6420

tggcctagtg attggcccgg aacccaagac agccggggga aggcagatcc ggctgctact 6480

ggcacccacg tgggaggccg tacccgatga atgccatgtg gtcgaaggtc gctacgggcc 6540

gctggcaaag gaaatcgggg cagaggtcga tagtggtggg gcgctcaagc acatacagga 6600

ctgcatcgag cgcctttgga aggtatccat catcgcccag aatggccgca aggggcaggg 6660

gtttcggctg ctgtcggagt acgccagcga cgaggcggac gggcgcctgt acgtggccct 6720

gaacccattg atcgcgcagg ccgtcatggg tggcgggcag catgcgcgcc taagcatggc 6780

cgcggtgcgg gagcggaccg gcgcaacgcc ccgcccgatg ccccagcggc tgtgcggctg 6840

gatcgacccc ggcaaaaccg gcaaggcttc catagataca ttgtgcggct atgtcgggcc 6900

gtcagaggcc agtggttcga ccatgcgcaa gcgccgcaag cgggtgcgcg aggcgctgcc 6960

ggagctggtc gcgctgggct ggacggtaac cgagttcgcg gcgggcaagt acgacatcac 7020

ccggcccaag gcggcaggct gacccccccc actctattgt aaacaagaca tttttatctt 7080

ttatattcaa tggcttattt tcctgctaat tggtaataaa atgaaaaatt atctgctaag 7140

atatagatta aattcattta gatcaattgc ttcctgaggg atgcgtaaat gctcgataag 7200

cgccttcgcc ggcttcgctt ccagatgtac gcttctctgc tcccgaacgc cagcaagacg 7260

tagcccagcg cgtcggccag cttgcaattc gcgctaactt acattaattg cgttgcgctc 7320

actgcccgct ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg 7380

cgcggggaga ggcggtttgc gtattgggcg ccagggtggt ttttcttttc accagtgaga 7440

cgggcaacag ctgattgccc ttcaccgcct ggccctgaga gagttgcagc aagcggtcca 7500

cgtggtttgc cccagcaggc gaaaatcctg tttgatggtg gttaacggcg ggatataaca 7560

tgagctgtct tcggtatcgt cgtatcccac taccgagata tccgcaccaa cgcgcagccc 7620

ggactcggta atggcgcgca ttgcgcccag cgccatctga tcgttggcaa ccagcatcgc 7680

agtgggaacg atgccctcat tcagcatttg catggtttgt tgaaaaccgg acatggcact 7740

ccagtcgcct tcccgttccg ctatcggctg aatttgattg cgagtgagat atttatgcca 7800

gccagccaga cgcagacgcg ccgagacaga acttaatggg cccgctaaca gcgcgatttg 7860

ctggtgaccc aatgcgacca gatgctccac gcccagtcgc gtaccgtctt catgggagaa 7920

aataatactg ttgatgggtg tctggtcaga gacatcaaga aataacgccg gaacattagt 7980

gcaggcagct tccacagcaa tggcatcctg gtcatccagc ggatagttaa tgatcagccc 8040

actgacgcgt tgcgcgagaa gattgtgcac cgccgcttta caggcttcga cgccgcttcg 8100

ttctaccatc gacaccacca cgctggcacc cagttgatcg gcgcgagatt taatcgccgc 8160

gacaatttgc gacggcgcgt gcagggccag actggaggtg gcaacgccaa tcagcaacga 8220

ctgtttgccc gccagttgtt gtgccacgcg gttgggaatg taattcagct ccgccatcgc 8280

cgcttccact ttttcccgcg ttttcgcaga aacgtggctg gcctggttca ccacgcggga 8340

-continued

```
aagggtatga tacgagaacc aggcaacatc tgcgacatcg tataacgtta atggcttacc    6400

attcaaaacc atgcattgcc tatcaacggg gagcaatacc gcaaaaagcc gacaggtttt    6460

gcaccattgg atggtgtcaa cgtaaatgcc gcttcgcctt cgcgcgcgaa ttgcaagctg    6520

atccgggctt atcgactgca cggtgcacca atgcttctgg cgtcaggcag ccatcggaag    6580

ctgtggtatg gctgtgcagg tcgtaaatca ctgcataatt cgtgtcgctc aaggcgcact    6640

cccgttctgg ataatgtttt ttgcgccgac atcataacgg ttctggcaaa tattctgaaa    6700

tgagctgttg acaattaatc atcggctcgt ataatgtgtg gaattgtgag cggataacaa    6760

tttcacacag gaaacagacc tagagctagg tacccgggga tactctagag tcgacctgca    6820

ggcatgca                                                             6828
```

<210> SEQ ID NO 2
<211> LENGTH: 8272
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized.

<400> SEQUENCE: 2

```
acgctacttg cattacagct tacgaaccga acaggcttat gtccactggg ttcgtgcctt      60

catccgtttc cacggtgtgc gtcacccggc aaccttgggc agcagcgaag tcgaggcatt     120

tctgtcctgg ctggcgcatg ccaaaatccc ttaacgtgag tcagcctgcc gccttgggcc     180

gggtgatgtc gtactggccc gacgcgaacc aggtaaacgt aacgcccagc gagaccagat     240

aaggcaacgc atcgcgaacc agctagaggc gatttgcgac ggtcgaaacc atggcctatg     300

acggccagac atagccgcac aaggtatcta tggaagcctt gccggttttg ccggggtcga     360

tccagccaca cagccgctgg tgcagcaggc gggcggtttc gctgtccagc gcccgcacct     420

cgtccatgct gatgcgcaca tgctggccgc caccaatgac ggcctgcgcg atcaaggggt     480

tcagggccac gtacaggcgc ccgtccgcct cgtcgctggc gtactccgac agcagccgaa     540

acccctgccg cttgcggcca ttctgggcga tgctggatcc cttcaaaggc agctcgatgc     600

agtcctgtac gtgcctgcgc gacccaacac tatcgaactc tgccccgatt tactttgacc     660

gcgcccgata gctacctttg accacatggc attcagcggt gacggcctcc cacttgggtt     720

ccaggaacag ccggagctgc cgtccgcctt cggtcttggg ttccgggcca agcactaggc     780

cattaggccc agccatggcc accagccctt gcaggatgcg cagatcatca gcgcccagcg     840

gctccgggcc gctgaactcg atccgcttgc cgtcgccgta gtcatacgtc acgtccagct     900

tgctgagctc gagctagccc agctagaggg aacggaaacg gacggcggcc agccagtgcg     960

aagggtagtg aaggaagtgg atgcggatgc gattgctatc aggctcaacc aagggcaacc    1020

cccttgctct tgcgctgcct ctccagcacg gcgggcttga gcaccccgcc gtcatgccgc    1080

ctgaaccacc gatcagcgaa cggtgcgcca tagttggcct tgctcacacc gaagcggacg    1140

aagaaccggc gctggtcgtc gtccacaccc cattcctcgg cctcgggcgc ggtcatgctc    1200

gacaggtagg actgccagcg gatgttatcg accagtaccg agctgcccgg gctggcctgc    1260

tgctggtcgc atgcgaacct aatggcagcg aactggatgg aatggtgacg gaccacgatc    1320

gagcaccggg tatccgcggc gatggcctcc atgcgaccga tgacctgggc catgggccgg    1380

atggcgtttc attcctagct gtggaacggg agctgcgtgc aacgcaaact aaggcggagg    1440
```

continued

```
aacncggcgg agcgcttgcg gacgacgacc cactccgggg acctgatgtt gggcaggatg    1500
aagatcagcg gatggatccg caggccgtca gaccaggatc gacgacaatc ggcgccgagg    1560
tgccccccaa gggcgtgcag gcggtgatga atggcggtgg gcgggtcttc ggcggggcagg   1620
tagatcaccg ggcccgtggg cagttcgccc accgccagca gatccggccc gccggcaatc    1680
tgtgcggcca gttgcagggc cagcatggat ttacccggcac caccgggcga caccagcgcc   1740
ccgaccgtac cggccaccat gttgggcaaa acgtagtcca gcggtggcgg cgctgctgcg    1800
aacgcatcca gactattgcc aggcccatgg gtagcccatg attgccttcct ttgcaggacg   1860
ttggcggtta ggcgctggcg gggtcaatcc cacggcaatg agccgctatg agttcctacc    1920
ggcactcggc cagcgcctgc tattcgtcgc cggtcagcca gaacctgcgc tgacgcatcc    1980
ctttggcgtt catgcgctcg gcatatcggc cttggcgtac agcgtcaggg ctggccagca    2040
ggtcgccggt ctgcttgtcc ttttggtctt tcatatcgat caccgagaaa cttgccgggg    2100
aagaaaggcc tgtcctagcg gacaaaggcc aaggcgaagc cgtcaaggtt aaggccggcc    2160
atcccagcga ctgcaaagcg gccggcctcg gcctgctttg acgtataaac aacgccaacg    2220
ggcaaccaat agccttgtc acttctgatc aggtagaccg accctgaagc gctttcttcg     2280
tatcccataa aaccccttc tgtgcgtgag tactcatagt ataacaggcg tgagtaccaa     2340
cgcaagcact acatgctgaa atctggcccg ccccgtccca tgcctcgctg gcggggtgcc   2400
ggtgcccgtg ccagctcggc ccgcgcaagc tggacgctgg gcagaccat gacctcgctg     2460
aaggcgagcc agctgtactc agctccgtgg aagggcttgc gatcggaacg agctggcatg   2520
gaccggccca tggccttgcc gattccatcg gaccgaggc aacggctggc cagcccatgc     2580
gcggcgataa agtcgcactt gctgacgtca tgacgcaagc gcttgaccag cccggccatc   2640
tcgctgcggt actcgtcgag cgccgtgcgc cggtggcggc taagctgccg ctcggggcagt   2700
tcgaggctgg ccagcctgcg ggccttctcc tgctgccgct gggcctgctc gatctgctgg   2760
ccagcctgct gcacccagcgc cgggccagcg gtggcggtct tgccctggga ttcacgcagc  2820
agcacccacg gatgatacc gggcgcggtg gtgtgcttgc cattgcggtt ggtgaagccc    2880
gacaagagcc catcgtggcg gatgccggcg atggccgggc aggcgacgtc atcgccggcc   2940
agcgcccggc caatctgccc ccgaagttca ccgcctgcgg cgtcgccccac cttgacccat  3000
gcccgatagc tcttcgggcc ggtttccacc accagcgcag gctcccggcc ctcggcttcc   3060
atgtcatccca ggtcaaactc gctgaggtcg tcccaccagca ccagaccatg ccgctcctgc 3120
tcggcgggcc tgatatacac gtcattgccc tgggcattca tccgcttgag ccatggcgtg   3180
ttcaggagca attcaggagc tgacccattcc aggttcatca tatggcaggt gggtgcgtcc  3240
atgacgacga tatcgaacgc atcccagccc atggccttga gatgccggcc tatggcatgc   3300
aaagtcctgt cgttcttcat cgggccacca agcgcagcca gatcgagccg tcctcggttg   3360
tcagtggcgt caggtcgagc aagagcaacg atgcgatcag cagcaccacc gtaggcatca   3420
tggaagccag catcacggtt agccatagct tccagtgcca ccccccgcgac gcgctccggg  3480
cgctctgcgc ggcgctgctc accgcggcg ctacctcccg caacctcttg gccagctcca    3540
aacacgacgc aactgtatgg agctggcatc tatgccaatc agccgcatgc gactcgatgg   3600
cctgcttggc ctggctcatg accgcccggg cttcgtcggc cagtgtcgcc atgcctctggg  3660
accgcggttc gatccgatcc gatcactcgc tgatgcatcc ggatccattc actcgtcgc    3720
```

-continued

```
ttgcgttcaa ggtcaattgc atcccggtaa taatgtaaga agatgatatg ggcgaaggcg    3780
gtgacgatga taaggggacc gtatgcaagg taggaaggga tgccccggcc ttccaaatcc    3840
accacgttgg gcgccaggtg aacaccgggc aggcgctcga tgcgctgcgc ctcaagtgtt    3900
ctgaggtcaa tgcgggcgtc gtggccagcc cgctcaaatg ccgggatggc atggacggcc    3960
catgccctgc gggtctgctc aagccatgcc ttgggcttga gcgcttcggt cttctgtgcc    4020
ccgcccttct ccggggtctt gccgttgtac cgcttgaacc actgagcggc gggccgctcg    4080
atgccgtcaa tgatcagctc ggagatcatc aggtggaaga gaggggtatc gaagccaacg    4140
gaaaggatgg aaagcgtata aggcaaggcg taggcaaagg taaggagatg ggcgaaatcg    4200
gacgccagcg ccttctgctg gtcgacggtc agctcgacgg gcagggcaaa ttcgacctcc    4260
ttgaacagcc gcccattggc gcgttcatac aggtcggcag catcccagta gtcggcgggc    4320
cgcttgacga actccggcat gtgcccggat tcggcgtgca agcattcatc catgtcgcgg    4380
gaaaaattgc attgaagctg gatgaagtcg gactaggacc tggccgattg gacgccagac    4440
atgcagaagg ttttcgaaga aaggaggtca atcgccatgc tgccacgatg ttgcaattgc    4500
ttttcggctc catgcaatgg cgctcggaga gcgcaccgcc cgaaggctgg ccgttaggcc    4560
agttactcga agagaaaccg gtaagtagcc cctcccctac aaagtagggt cgggattgcc    4620
gccgctgtgc ctccatgata gcctacgaga cagcacatta acaatggggt gtcaagatgg    4680
ttaaggggag caacaaggcg gcggatcggc tggccaagct cgaagaacaa cgagcgcgaa    4740
tacatgacga aattcagcgg gagcaggcaa gggaacagca gaaagagagc aagaacgaac    4800
aaaggagcaa ggtgctggtg ggggccatga ttttggaaaa ggtgaaaagc agcgagtggc    4860
cggaggatcg gctcatggcg gcaacggatg cgtaccttga acgcgaccac gacggcgcct    4920
tgttcggtcc gccgccacgc cagaaggatg agcgcgcctg aatgatcgac cgagacaggc    4980
cctgcggggc tgcaacagcg ccccccaccc tcgggtaggg ggaaaggccg ctaaagcggc    5040
taaaagcgct ccagcgtatt tctgcggggt ttggtgtggg gtttagcggg ctttgcccgc    5100
atttcacccc gacgcgaagc ggtgcggcgg tgtgaagacc agcgcagaga atagacaagc    5160
tatccggccc atggcagggc atatagggca agggcagaag agccccaaac gggcgctgat    5220
aaccgcgccc agtggattat tcttagataa tcatggatgg atttttccaa caccccgcca    5280
gcccccgccc ctgctgggtt tgcaggtttg ggggcgtgac agttattgca ggggttcgtg    5340
acagttattg caggggggcg tgacagttat tgcaggggtt cgtgacagtt agtacgggag    5400
tgacgggcac tggctggcaa tgtctagcaa cggcaggcat ttcggctgag ggtaaaagaa    5460
atttcagcta agcgatagac tgtaagtaaa aaccgtattg aacggaagcg gaccatgaat    5520
aatgagggcg aacggaaggc aagccgggaa agggaaaatg gtcgaaaacc gagccaaagc    5580
cccgagcaaa ccctactcta tcagatcgtt gacgagtatt acccggcatt cgctgcgctt    5640
atggcagagc agggaaagga attgccgggc tatgagcaac gggaatttga agaatttctc    5700
caatgcgggc ggctggagca tggcattcta cgggttcgct gcgagtcttg ccacgccgag    5760
cacctggtcg ctttcagaaa tcaatctaaa gtatatatga gtaaacttgg tctgacaggc    5820
aactgaattc gaatctagac ggtagagcca aaacaagagg gtacaagaac agatctaaag    5880
gaccatctga atcatgcgcg gatgacacga actcacgacg gcgatcacag acattaaccc    5940
aaagtaaaga aactgagaca aagtggaaaa tagtcgaaaa aacgacttac tgaaacggcc    6000
```

continued

```
gataattgca gacgaacgtc atccaaacga gtatcgactc aaagcctacc ctctaggatc    6060

cttacagcca tagcgaggga aacggcgagg aattgcgagc ggctaaacct tacccaaagc    6120

aacagaggag atattacata tgagtaaagg agaagagctt ttcacaggag ttgtcccaat    6180

cctcgtggaa ttagacggtg atgttaatgg gcacaagttc tctgtcagtg gagagggtga    6240

aggcgatgca acatatggaa agcttgaccc taaatttatt tgcaccacgg gtaaactacc    6300

tgttccatgg ccaacacttg tcactacttt cgggtatggt gttcaatgct ttgcgagata    6360

cccagatcat atgaaacagc atgacttttt caagagtgcc atgcccgaag gttatgtaca    6420

ggaaagaact atatttttca aagatgacgg gaactacaag acacgtgctg aagtcaagtt    6480

tgaaggtgat acccttgtta atagaatcga gttaaaaggt attgatttta aggaagatgg    6540

aaatatttta ggacacaaac tggaatataa ctataactca cataatgtgt acatcatggc    6600

agacaagcaa aagaacggta tcaaggttaa cttcaagatc agacacaaca ttgaggatgg    6660

aagcgttcaa ctagcagacc attatcaaca aaatactcca attggcgatg gccctgtcct    6720

tttaccagac aaccattacc tgtccactca atctgccctt tcgaaagatc ccaacgaaaa    6780

gcgcgaccac atggtccttc ttgagtttgt aacggctgct gggattacac acggcatgga    6840

tgaactatac aaataaatgc gtaatcgtta atcggcaaat aacgtaaaaa cccgcttcgg    6900

cgggtttttt tatgggggga gtttagggaa agagcatttg tcagcatgca atgagacgac    6960

ggggtcatca cggctcatca tgcgccaaac aaatgtgtgc aatacacgct cggatgactg    7020

aatgatgacc gacctgactg gggacagccg atccaacatc gactgtcagc gacgcagacc    7080

aacgaaagcc aacggaagtc aactagtgca atgcagtaca aacggcatcc aatcaaaagc    7140

aaggctcagt cgaaagactg ggcctttcgt tttatctgtt gtttgtcggt gaacgctctc    7200

tactagagtc acactggctc accttcgggt gggcctttct gcgtttatac cgagcggccg    7260

cttattggcc gactaccttg gtgatctcgc ctttcacgta gtggacaaat tcttccaact    7320

gatctgcgcg cgaggccaag cgatcttctt cttgtccaag ataagcctgt ctagcttcaa    7380

gtatgacggg ctgatactgg gccggcaggc gctccattgc ccagtcggca gcgacatcct    7440

tcggcgcgat tttgccggtt actgcgctgt accaaatgcg ggacaacgta agcactacat    7500

ttcgctcatc gccagcccag tcgggcggcg agttccatag cgttaaggtt tcatttagcg    7560

cctcaaatag atcctgttca ggaaccggat caaagagttc ctccgccgct ggacctacca    7620

aggcaacgct atgttctctt gcttttgtca gcaagatagc cagatcaatg tcgatcgtgg    7680

ctggctcgaa gatacctgca agaatgtcat tgcgctgcca ttctccaaat tgcagttcgc    7740

gcttagctgg ataacgccac ggaatgatgt cgtcgtgcac aacaatggtg acttctacag    7800

cgcggagaat ctcgctctct ccaggggaag ccgaagtttc caaaaggtcg ttgatcaaag    7860

ctcgccgcgt tgtttcatca agccttacgg tcaccgtaac cagcaaatca atatcactgt    7920

gtggcttcag gccgccatcc actgcggagc cgtacaaatg tacggccagc aacgtcggtt    7980

cgagatggcg ctcgatgacg ccaactacct ctgatagttg agtcgatact tcggcgatca    8040

ccgcttccct cataatgttt aactttgttt tagggcgact gcccgctgcg gtaacatcgt    8100

tgctgctcca taacatcaaa catcgaccca cggcgtaacg cgcttgctgc ttggatgccc    8160

gaggcataga ctgtacccca aaaaaacagt cataacaagc catgaaaacc gccactgcgc    8220

cgttaccacc gctgcgttcg gtcaaggttc tggaccagtt gcgtgagcgc at            8272
```

continued

<210> SEQ ID NO 3
<211> LENGTH: 7999
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized,

<400> SEQUENCE: 3

tccgcttcgt ggccgggctt gcgctctgcc agcgctgggc tggcctcggc catggccttg 60

aagatttcct aggcaatgcg gacccggctg gcctgcttcc gagcgggcgt aacgtcgacc 120

ttgctgaggt aatgaacgta gagcttgacc agcccggaca tatcgctgcg gtacccgtcc 180

agcgccgtgc gccggtggcg gctaagctgc cgctcgggca gttcgaggct gggcagcctg 240

cggcccttcc cctgctgccg ctgggcctgc tcgatctgcc gggcagcctg ctgcaccagc 300

gccgggcccg cggtggcggt cttgcccttg gattcacgca gcagcacccc cggctgatcc 360

aagcgcaggc tggtgtgctc gtcccggcgc ttggcgaagc aagccaagcg gacctagtgg 420

aggcagtcgc agctggacgc gtcgccgtcg tactcgatgc caagcgtacg ggccatatgc 480

cccgaagttc cacggcctgc gggctcggcc acctcgaccc atgcctgata gttcctcggg 540

ctggcttcca ctaccagggc aggctcccgg ccctcggctt tcatgtcatc caggtcaaac 600

tcgctgaggt cgtccaccag caccagacca tgccgctcct gctcggcggg cctgatatcc 660

acgtcattgc cctgggcatt catcccgctg agccatggcg tgttctggag cactacggcg 720

gatgaaactc aacggttcct aatctggccg gtggctgagc aactgaagcc gatactgaag 780

agctcaatgc aactggactc gagcagtcgg aatatggacc gaacagtact gtcgcctatc 840

atcggcgcac caagcgcagc cagatcgagc cgtccctcgg tgtcagtggc gtcaggtcga 900

gcaagagcaa cgatgcgatc agcagcacca ccgtagccat catggaagcc agcatcacgg 960

ttagccatag cttccagtgc caccccggcg acgcgctccg ggcgctctgc gcggcgctgc 1020

tcaccctggc ggctaccttc cgcaactctt tggccagctc cacccatgcc gcccctgtct 1080

ggcgctgggc tttcagcacc tacgccgccc gagcctagcc ggcctgcttg gtctggatcc 1140

tgacctgccg ggcttagtcg gacggtgtcg aactgctatg ggccagaggt tagatctgct 1200

ccgctaactc gttgatgccc ctggatttcc tcactctgtc gattgcgttc atggtctatt 1260

gcctcccggt attcctgtaa gtcgatgatc tgggcgttgg cggtgtcgat gttcagggcc 1320

acgtctgccc ggtcggtgcg gatgccccgg ccttccatct ccaccacgtt cggccccagg 1380

tgaacaccgg gcaggcgctc gatgccctgc gcctcaagtg ttctgtggtc aatgcggcgg 1440

tagtggaccg aacgctatca tgcccggttg gactggtagg aacctgactc gagggtatgc 1500

tacagaactg aattgggctc gagcgcttcg gtctactgtg aaccgcactt atccggggtc 1560

ttgccgttgt accgcttgaa ccactgagcg gcgggccgct cgatgccgtc attgatccgc 1620

tcggagatca tcaggtggca gtgccgcttc tcgccgccac cggcatggat ggccagcgta 1680

tacggcaggc gctcggcacc ggtcaggtgc tgggcgaact cggacgcccg cgccttctgc 1740

tggtcgaggg tcagctcgac cggcaggcca aattcgacct ccttgaacag ccgcccattg 1800

gagcgttcct aacggtaggc agcctcaccg tagtcggagg gacgctagtc gacctcaggc 1860

atgtgcccgg attccgcgtg caagacttca tccatgtcgc gggcatactt gccttcgcgc 1920

tggatgtagt aggccttggc aatggcagct tggccgaacg aactgctgcc ggttttagcc 1980

-continued

```
gtcaggtgcc aactcgcccc gatgcctcgc tgttgctttc gctttcaggc tcccggcact    2040
ggccctcgga gagcgcaccg cccgaagggc ggccgtcagg cccgcctctc gacgagcacc    2100
cggtaagtgc gccctcccct acaaagtagg gtcgggattg ccgccgctgt gcctccatga    2160
tagcctacga gacagcacat taacaatggg gtgtcaagat ggttaagggg agcaacaagg    2220
cggcggatcg gctggcccag ctcgaagcac aacgagcgcg aatcaatgcc gaaattcagc    2280
gggagcgggc aagggaaccg cagcaagcgc gcaagaacga aacaaggcgc aaggtgctgg    2340
tgggggcccc gatttcggcc aaggcgacca gccgcgcgtg gccggaggct cggcccatgg    2400
cggcaatgga tgcgtcactc gaccgcgccc aagcccgcgc cttgcccggt ctgccgcccc    2460
gccagaagga tgagccgggc tgaatgatcg accgagacag gccctgcggg gctgcacacg    2520
cgcccccacc cttcgggtag ggggaaaggc cgctaaagcg gctaaaagcg ctccagcgta    2580
tttctgcggg gtttggtgtg gggttcagcg ggcttcgccc gccttccccc ctgccgcgca    2640
gaggcggggc ggtgcgtcgc ctcgcgccgc gcctagcaca gctcccaggc ctctggccgg    2700
gcccatcggg cccggcccgc agcgcccccc aagggcgctg ctcccgcagc ctcgccgctt    2760
attcctagat aatcatggat ggatttttcc aacaccccgc cagcccccgc ccctgctggg    2820
tttgcaggtc tggggcgctg acagccattg cagggcttcg tgacagttat tgcagggggg    2880
cgtgacagtt attgcagggg ttcgtgaccg ttcgtacggg agtgacgggc actggctggc    2940
aatgtctagc aacggcaggc atttcggctg agggtaaacg aacttcccgc taagcgatcg    3000
cctgcatgta cacccagtcc tgccaggccg cggccacagc ctcccgtggc ggccaggccg    3060
gccagcaggg ctcgggatcc tggtcgttcc ccgcgccacc gacccgcgcc cacccctctc    3120
tatcagatcg ttgacgagta ttacccggca ttcgctgcgc ttatggcaga gcagggaaag    3180
gaattgccgg gctatgtgca acgggaattt gaagaatttc tccaatgcgg gcggctggag    3240
catggctttc tacgggttcg ctgcgagtct tgcccacgcc agcaccctgg cgcttccaga    3300
aatcaatcta aagtatatct gagtaaactt ggtctgacag gcccctgaat tcgcatctag    3360
ctggcccgcc cccccccagc cggtccccgc cccgaccccc aggccccctc gcctccatgcg    3420
cggacgcccc gacccccgca cgggcgctcc agccactccc cccccgtccc gcccccgccc    3480
caacgtggca attcgtcgca atacgctctc actgaactgg ccgataattg cagacgaacg    3540
ggtgaaacaa aacggttgac aacatgaagt aaacacggta cgatgtacca catgaaacga    3600
cagtgaagca atacagaaac agaggagata ttacatatgg atagcactga gaacgtcatc    3660
aagcccttca tgcgcttcaa ggtgcacatg gagggctccg tgaacggcca cgagttcgag    3720
atcgagggcg agggcgaggg ccgcccctac gagggcaccc agaccgccaa gctgaaggtg    3780
accaagggcg gccccctgcc cttcgcctgg gacatcctgt ccccccagtt ccagtacggc    3840
tccaaggcgt acgtgaagca ccccgccgac atccccgact acaagaagct gtccttcccc    3900
gagggcttca agtgggagcg cgtgatgaac ttcgaggacg gcggcgtggt gaccgtgacc    3960
caggactcct ccctgcagga cggcaccctc atctacaagg tgaagttcat cggcgtgaac    4020
ttcccctccg acggccccgt aatgcagaag aagaccatgg gctgggaggc ctccaccgag    4080
cgcctgtacc cccgcgacgg cgtgctgaag ggcgagaccc acaaggccct gaagctgaag    4140
gacggcggcc actacctgtg tgagttcaag tccatctaca tggccaagaa gcccgtgaag    4200
ctgcccggct actactacgt ggacgccaag ctggacatca cctcccacaa cgaggactac    4260
```

continued

```
aacgtggtgg agcagtacga gagcgcagcg gacggcaacc aactgctacc gactcaaggt    4320
atggacgccc tgtacaagca agacgaattg tacctcagta atcgcctatc agccaatacc    4380
gtaaaaaccc gcttcggcgg gtttttttat ggggggagtt tagggaaaga gcatttgtca    4440
gctggatcag cagatgggaa ccggatactc ctcagcctcc agagtagcca atgagacgcc    4500
gagcggccgc ttattgcccg actaccttgg tgatctcgcc tttcacgtag tggacaaatt    4560
cttccaactg atctgcgcgc gaggccaagc gatcttcttc ttgtccaaga taagcctgtc    4620
tagcttcaag tatgacgggc tgatactggg ccggcaggcg ctccattgcc cagtcggcag    4680
cgacatcctt cggcgcgatt ttgccggtta ctgcgctgta ccaaatgcgg gacaacgtaa    4740
gcactacatt tcgctcatcg ccagcccagt cgggcggcga gttccatagc gttaaggttt    4800
catttagcgc ctcaaataga tcctgttcag gaaccggatc aaagagttcc tccgccgctg    4860
gacctaccaa ggcaacgcta tgttctcttg cttttgtcag caagatagcc agatcaatgt    4920
cgatcgtggc tggctcgaag atacctgcaa gaatgtcatt gcgctgccat tctccaaatt    4980
gcagttcgcg cttagctgga taacgccacg gaatgatgtc gtcgtgcaca acaatggtga    5040
cttctacagc gcggagaatc tcgctctctc caggggaagc cgaagtttcc aaaaggtcgt    5100
tgatcaaagc tcgccgcgtt gtttcatcaa gccttacggt caccgtaacc agcaaatcaa    5160
tatcactgtg tggcttcagg ccgccatcca ctgcggagcc gtacaaatgt acggccagca    5220
acgtcggttc gagatggcgc tcgatgacgc caactacctc tgatagttga gtcgatactt    5280
cggcgatcac cgcttccctc atgatgttta actttgtttt agggcgactg ccctgctgcg    5340
taacatcgtt gctgctccat aacatcaaac atcgacccac ggcgtaacgc gcttgctgct    5400
tggatgcccg aggcatagac tgtaccccaa aaaaacagtc ataacaagcc atgaaaaccg    5460
ccactgcgcc gttaccaccg ctgcgttcgg tcaaggttct ggaccagttg cgtgagcgca    5520
tacgctactt gcattacagc ttacgaaccg aacaggctta tgtccactgg gttcgtgcct    5580
tcatccgttt ccacggtgtg cgtcacccgg caaccttggg cagcagcgaa gtcgaggcat    5640
ttctgtactg gatggtactg aacaaaatcc attaacgtga gtctgcctgc agccttgggc    5700
agggtgatgc agtacttgcc agccgcgctc taggttaacg tacctgcctg agcgacaagc    5760
tccggcaacg cctcgcgcac ccgcttgcgg cgcttgcgca tggtcgaacc actggcctct    5820
gacggccaga catagccgca caaggtatct atggaagcct tgccggtttt gccggggtcg    5880
atccagcccc acgccgcctg gtgcagcagg cgggcggttt cgctgtcccg cgcccgcacc    5940
tcgtccatgc tgatgcgcac atgctggccg ccacccatga cggcctgcgc gatcaagggg    6000
ttcagggcca agtacaggcg aacgtcagcc tagtcgatgg agtacttagt cagcagacgc    6060
aaccccatgc gatttgaggc attctgggcg atgctggata aatttccaag gagctcgatg    6120
cagtcctgta tgtgcttgag cgccccaccc ctatcgaacc ctgccccgat ttcccttgcc    6180
agcgccccat agctacctta gacaacatgg cattcagcgg tgacggcctc ccactttggt    6240
tccaggaaca gcccggagct gcgtccgccc tcgggcttgg gttccgggcc aagcactagg    6300
ccattaggcc cagccatggc caccagccct tgcaggatgc gcagatcatc agcgcccagc    6360
ggctccgggc cgctgaactc gatccgcttg ccgtcgccgt agtcatacgt cacgtccagc    6420
ttgctgcgct tgcgctcgcc ccgcttgagg gcacggaaca ggccgggggc cagacagtgc    6480
gccgggtcgt gccggacgtg gctgacgctg tgctcgttcg taggcttacc ccggcggacc    6540
```

-continued

```
aaccctgctc ttgcgatgcc tatccagccc ggcggggttg agcaccaagc agtcatgacg    6600

aatgaaaccc agctcagcga aaggtgagcc atagttggcc ttgctcaacc agcagcggcc    6660

gaagaacggg cgctggtcgt cgtccacacc ccattcctcg gcctcggcgc tggtcatgct    6720

cgacaggtag gactgccagc ggatgttatc gaccagtacc gagctgcccc ggctggcctg    6780

ctgctggtcg cctgcgccca tcatggccgc gcccttgctg gcatggtgca ggaacacgat    6840

agagcaaccg gtatcggcgg cgatggcctc catgcgaccg atgacctggg ccatggggcc    6900

gatggcgttc tattcataga tgtggaaccg gagcagcgtg taccgcaact tacggcgggg    6960

gaccaaggcg gagcgcttga ggccgtagca accctcaggg gacatgatgt tgggcaggct    7020

gccgatcagc ggctggatca gcaggccgtc agcgacggct tgccgttcct cggcgctgag    7080

gtgcgcccca agggcgtgca ggcggtgatg aatggcggtg ggcgggtctt cggcgggcag    7140

gtcgatcacc gggccggtgg gcagctcgcc cacctccagc agatccggcc cgcctgcaat    7200

atgatcggcc agttgcaggg aacgcatgga tttaccggca aaaccgggcg aaaccagagc    7260

aacgaaagta aaggcaacca tgttggcaaa aacgtagtcc agcggtggcg gagcagatgc    7320

gaacgcctcc agaatattga taggcttatg ggtagccatt gattgcctcc tttgcaggca    7380

gttggtggtt aggcgctggc ggggtcacta ccccgccccc gcgccgctct gagttcttcc    7440

aggcactcgc gcagcgcctc gtattcgtcg tcggtcagcc agaacttgcg ctgacgcatc    7500

cctttggcct tcatgcgctc ggcatatcgc gcttggcgta cagcgtcagg gctggccagc    7560

aggtcgaacg tatgcttgtc atttaggtac ttcataaacg taaccgagca aattgcaggg    7620

gaagaaaggc ttgtcttcgc ggaacaagga aaaggtgaag aagtcaaggt taaggctggc    7680

catatcaggc actgaaaagc ggccagcctc ggcctcgttt gacgtataac caaagccacc    7740

gggcaaccaa tagcccttgc cactttgatc caggtagacc gacccgaaag cgcttttttc    7800

gtatttccta aaaccccctt ctgtgcgtga gtacttcatg tataacaggc gtgagtaccc    7860

acgcaagcac tacatgctga aatctggccc gccccctgtc atgccacgct ggcggggtgc    7920

aggtgaacga gaaagatcgg aacgcgaaag atggacgatg ggcagaaccc tgaccatgct    7980

gacggtgcgc tagaagtaa                                                 7999
```

<210> SEQ ID NO 4
<211> LENGTH: 8101
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized.

<400> SEQUENCE: 4

```
atgaatgtaa taagcttcga ggccgggctc gagcatcgcc agcgctgggc tggcctaggc      60

catggccttg ccgatttcct cggcactgcg gccccggctg gccagcttct gcgcggcgat     120

aaagtcgcac ttgctgaggt catgaccgaa gcgcttgacc agcccgcgca tctcgctgcg     180

gtacttgtcc agcgccgtgc gccggtggcg gctaagctgc cgctcgggca gttcgaggct     240

ggccagcctg cgggccttcc cctgctgccg ctgggcctgc tcgatctgct gggcagcctg     300

atgcaaatgc gacggcaccg aggtgcaggc attgccattg gattcaagca gaagcaaaca     360

cggctgataa ccggcgcggg tggtgcgctt gtccttgcgg ttggtgaagc ccgccaagcg     420

gacaaagtgg aggcagtcgg agctgcacgg gtcggcgtcg tactcgatgg aaagcgtacg     480
```

-continued

```
ggcaatatgc aaccgaagta aaccgcatgc gggcacgggc aactagaacc atgccagata     540

gttcatacgg atggattcca ataccagggc aggcacaagg aactcggatt taatgaaatc     600

caggtcaaac tcgctgaggt cgtcgcacag caccagacca tgccgctcct gctcggcggg     660

cctgatatac acgtcattgc cctgggcatt catccgcttg agcgatgggg tgttctggag     720

cacttcggcg gctgaccatt cccggttcat catctggccg gtgggtgcgt ccctgacgcc     780

gatatcgaag cgctcacagc ccatggcctt gagcatgtcg cctatggcct gcaaagtcct     840

gtcgatattc atcggaaccc aacgcgaagc aagaacgagc agtacaaggt tgtcagtggc     900

gtcaggtcga gaaagagcaa agatgcgatc agcaagcaaa aagtaggaat aatggaagcc     960

agcatcacgg ttagccatag cttccagtgc caccccgcgc acgcgctccg ggcgctctgc    1020

gcggcgctgc tcacctcggc ggctacctcc cgcaactctt tggccagctc caccatgccc    1080

gccccgtcat ggcgctgggc tttcagcccc tccgccgcca gcgccacgct ggccagcttg    1140

gtcaggatca tgaactgcag ggataagtag gaaagaatag aaatgctatg ggcaagaggt    1200

tagaaatgca aagcaaactc gttgaagccc atggaattca taaacatgtc gattgcgttc    1260

atggcctatt gcgtcccggc attccagtaa gtcgacgatc tgggcgttgg cggtgccgat    1320

gttcagggcc acgtctgccc ggtcggtgcg gatgccccgg ccttccatcc ccagcacgtt    1380

cggccccagg tgaacaccgg gcaggcgctc gatgccctgc gcctcaagtg ttctgaggtc    1440

aatgcgggcg tcgtggccag cccgcatcaa tgcccggttg gcatggtcgg cccatgcctc    1500

gaggatatgc taaagaaatg aattggacta gagagattag gtataatgtg aaaagaaatt    1560

ataagggatc ttgacgttga aaagcatgaa aaatagagag gaggaagcat agatgaagtc    1620

attgatccgc tcggagatca tcaggtggca gtgcccgttc tcgccgccac cggcatggat    1680

ggccagcgta taccgcaggc gctcgccacc ggtcagctgc tgggcgaact cggacgccag    1740

cgccttctgc tggtcgaggg tcagcacgac cggcagggca aattcgaact ccttgaacag    1800

ccgcccattg gcgcgttcat acaggacggc agcatcccag tagtcggcgg gccgcacgac    1860

gaataaaggc atgtgaaagg attggaagtg aaagacttca taaaagtagc gggcaaaatt    1920

gaataagcgc tggaagtagc aggcaatggc aatggaagaa tggccgaacg aaatgctgcc    1980

ggtcctcgcc gtaacgtgat aaatcgccat gctgcctcgc tgttgctttt gctttccggc    2040

tccatgcaat ggccctcgga gagcgcaccg cccgaagggc ggcccctagg ccagcctctc    2100

gaagagaaac cggtaagtgc gccctcccct acaaagtagg gtcgggattg ccgccgctgt    2160

gcctccatga tagcctacga gacagcacat taacaatggg gtgtcaagat ggttaagggg    2220

agcaaaaagg aggaggatag gatggaaaag atagaagaac aaagagagcg aataaatgcc    2280

gaaataaagc gggaggaggc aagggaaaag aagcaagagc gaaagaaaga aaacaggagc    2340

aaggtgctgg tgggggccat gatttaggcc aaggtgaaca gcagcgagtg gccggaggat    2400

cggcccatgg cggcaatgga tgcgcacctt gaaccgcacc acgaccgcgc cttgctcggt    2460

ctgccgcccc gcccgaagga tgagccgggc tgaatgatcg accgagacag gcccagcggg    2520

gctgcacacg cgcccccacc cttcgcgtag ggggaaaggc cgctaaagcg gctaaaagcg    2580

ataaagaata tttaagaggg gtttggtgtg gggtaaagag ggataagacc gaataaaacc    2640

ctgccgcgca gcggcggggc ggtgcgtagc ctagcgcagc gaatagacca gctaccggcc    2700

atcaggaagg gaatattggg aaaggaaagc agagcaaaac aagggcgatg ataaacgagc    2760
```

-continued

```
atagtggatc attcatagtc aatcatggtc ggatctttcc aactcccagc cagccccagc    2820
acctgatggg tttgcaggtc tggggggtgg aacgttattg aaggggttcg tgccagttat    2880
tgcagggggg cgtgacagtt attgcagggg ttcgtgacag ttagtacggg agtgacgggc    2940
actggctggc aatgtctagc aacggcaggc atttcggctg agggtaaaag aactttccgc    3000
taagcgatag actgtatgta aacacagtat tgcaaggacg cggaacatgc ctcatgtggc    3060
ggccaggacg gccagccggg atcgggatac tggtcgttac cagagccacc gacccgagca    3120
aacccttctc tatcagatcg ttgacgagta ttacccggca ttcgctgcgc ttatggcagc    3180
gacggcaacg gacttgacgg gatacgtgca acgggaattc gacgaattta taccatgcgg    3240
gcggctggag catggctttc tacgggttcg ctgcgagtcc tgccacgccg agcacctggt    3300
cgccctcaga aatcaatcta aagtatatat gagtaaactt ggtctgacag gcgcctgaat    3360
tcgcatctag actgatgcga cgtggtagag cccacaacag ccggtacaag caacgatctc    3420
caggcaactc tgcccaatgc gaggacgcca agcaccaacg acggcgatcc cagccattcc    3480
cacacagtac agccaatgcg accccgtggc aattcgtagc aatccaacgg accgccattc    3540
atccctctgc ccctccaaaa gcaaaaacgc gccgaagcgg gtttttacgt aaatcaggtg    3600
aaactgaccg actttggcag tttattcttg acatgtagtg aggggcctgg tataatcaca    3660
tagtactgtt tatgagtcaa ggagaagagc ttttcactgg agttgtccca atcctcgtgg    3720
aattagacgg tgatgttaat gggcacaagt tctctgtcag tggagagggt gaaggcgacg    3780
caacatatgg aaagctgacc cttaaattta tttgcactac gggtaaacta cctgttccat    3840
ggccaacact ggtcactacc ttcggctatg gggttcagtg ctttgcgagc tacccagatc    3900
acatgaaaca gcacgacttt ttcaagagtg caatgcccga aggctatgta caggagagaa    3960
ccatcttttt taaggatgac gggaactata agacacgcgc cgaagtgaag ttcgagggtg    4020
atacccttgt taatagaatc gagttaaagg gtattgactt taaggaagat ggaaacattt    4080
taggccacaa actggaatat aactataact cccatagtgt gtacattatg gccgacaagc    4140
aaaagaacgg tatcaaggtc aacttcaaga ttagacacaa cattgaggat ggaagcgttc    4200
aactagcaga ccattatcaa caaaacaccc caattggaga tgggcctgtg ctgttaccag    4260
acaaccatta cctgtccaca caatctgccc tttcgaaaga tcccaacgaa aagcgcgacc    4320
acatggtcct tcttgagttt gtcacggctg ctgggattac acacggcatg gatgaactat    4380
acaaataaat ccaacagtac tatgtgatta taccagcccc ctcactacat gtcaagaata    4440
aactgcccaa ggtaatcgtt aatccgcaaa taacgtaaaa acccgcttcg gcgggttttt    4500
ttatggggga agtttaggga aagagcattc gtcagctgcg tatcaagacg ttcaagagct    4560
gttacgatcc aaccggactc aagacgatgc gaggcgatgg agggtaaaac aagacgggcc    4620
gctcatttgc cgactcacct ggtgatctcg ccttacacgt agtggacaaa ttctaccaac    4680
tgatctgcgc gcgaggccaa gcgatcttct tcttgtccaa gataagcctg tctagcttca    4740
agtatgacgg gctgatactg ggccggcagg cgctccattg cccagtcggc agcgacatcc    4800
ttcggcgcga ttttgccggt tactgcgctg taccaaatgc gggacaacgt aagcactaca    4860
tttcgctcat cgccagccca gtcgggcggc gagttccata gcgttaaggt ttcatttagc    4920
gcctcaaata gatcctgttc aggaaccgga tcaaagagtt cctccgccgc tggacctacc    4980
aaggcaacgc tatgttctct tgcttttgtc agcaagatag ccagatcaat gtcgatcgtg    5040
```

-continued

```
gatggatcga agctaacggc aagcacgtca ttgcgctgcc attcacaacc ttgcagttcg    5100
agcccagctg gatcaagcca aggcacgctg tagtcgtgca aacaaatggt gactacttac    5160
gcgcggagaa tctcgctctc tccagggaaa gccgaagttt ccaaaaggtc gttgatcaaa    5220
gctcgccgcg ttgtttcatc aagccttacg gtcaccgtaa ccagcaaatc aatatcactg    5280
tgtggcttca ggccgccatc cactgcggag ccgtacaaat gtacggccag caacgtcggt    5340
tcgagatggc gctcgatgac gccaactacc tctgatagtt gagtcgatac ttcggcgatc    5400
accgcttccc tactaatgtt tatccttgtt ttacggagcc tgccctgatg agttacatcg    5460
ttgcagatcc atttcatcta aactcgaccc aaggcgtaac gagcttgatg attggatgcc    5520
cgaggcatag actgtacccc aaaaaaacag tcataacaag ccatgaaaac cgccactgcg    5580
ccgttaccac cgctgcgttc ggtcaaggtt ctggaccagt tgcgtgagcg catacgctac    5640
ttgcattaca gcttacgaac cgaacaggct tatgtccact gggtcgtgcc cttcatccgt    5700
ttcacggtgt tgcgtaaccc ggctaccttg ggtagcagcg aagtcgaggc atttctgtcc    5760
tggctggtca tgtccaaccc aacttaacgt gagtcagacc gacgccttgg gacggttgat    5820
gtcgtacttg ccggccgcga actcggttac cgtccagccc agcgcgaccg gctccggcaa    5880
cgccccgcgc accggcttgc ggcgcctgcg catggtcgaa ccacggcgct ctgacggcca    5940
gacatagccg cacaaggtat ctatggaagc cttgccggtt ttgccggggt cgatccagcc    6000
acacagccgc tggtgcagca ggcggggcgg ttcgcatgcc agcgccggca cctcgtccat    6060
gatgatgcgc aactgatggc agcaacaacc gacggcatgc gagaacaagg ggttcagggc    6120
aacgtaacgg agcccgtccg aatcgtagcc ggcgtaatcc gacggcagcc gaaacccatg    6180
ccgcctgcgg ccattctggg cgatgatgga taccttccaa aggcgctcga tgcagtcctg    6240
tatgtgcttg agcgccccac cactatcgac ctctgccccg atttcctttg ccagcgcccg    6300
atagctaacc ttgaccaacc ggcattcagc ggtgacggcc tcccacttgg gttccaggaa    6360
cagccggagc tgccgtccgc cttcggtctt gggttccggg ccaagcacta ggccattagg    6420
aacagaactg gaaccaagcc attgcaggac gagcagatca taacgcaacc gaggcaacgg    6480
gacgctgaac tagcaagcat tgccgcagcc gtagtcatcc gtacggtaac gattgctgcg    6540
cttgcgctcg ccccgcttga gggcacggaa caggccgggg gccagacagt gcgccgggtc    6600
gtgccggacg tggccgaggc tgtgcctgtt cttaggcttc accacggggc acccccttgc    6660
tcttgcgctg cctcttcagc acggcgggct tgagcacccc gccgtcatgc cgcctgaacc    6720
accgatcagc gaacggtgcg ccatagttgg ccttgctcac accgaagcgg acgaagaacc    6780
ggcgctggtc gtcgtaacca aacaattcct aggcctaggc gatgctaatg atcgacaggt    6840
aggactgcca gaggatgtta tagtccagta aagagctgcc aaggctggcc tgctgctggt    6900
cgccatgcgc catcatggcc gcgcccttgc tggcatggtg cagggaacag atagagcacc    6960
cggtatcggc gggcatggcc tccatcgcac cgatgacctg ggccatgggg ccgcctgcgt    7020
tttcttcctc gatgtggaac cggcgcagcg tgtccagcac catcaggcgg cggccctcgg    7080
cggcgcgctt gaggccgtcg aagcactccg gggccatgat gttgggcagg ctgccgatca    7140
gacgctggac aagcaggccg taagccaacg attgccgttc atcggcgatg aggtgcgacc    7200
caagggcgtg caggcggtga tgaatgcggg tgggcgggtc ttcggcgggc aggtagatca    7260
aaggcaaggc gggcagttcg aacaccatca gaagataacg aacgcctgca atctgcaggc    7320
```

-continued

```
aaggttgccg ggccagactg gatttaaggg aaccacaggg agcaacaagc gacccgaacg    7380

taccggaccc aatgctgggc aaaacgtagc aacgcggtgg aggcgctgct gagaacgact    7440

ccagaatatt gataggctta tgggtagcca ttgattgcct cctttgcagg cagttggtgg    7500

ttaggcgctg gcgggctcac taccccggcc ctgcgccgcc ctgagttctt ccaggcactc    7560

gcgcagcgcc tcgtattcgt cgtcggtcag ccagaacttg cgctgacgca tccctttggc    7620

cttcatgcgc tcggcatatc gcgcttggcg tacagcgtca gggcttggca gcaggttcgc    7680

ggtcagattg taatattggt atttcatatc agtcaacagg aaactggacg gggccgaacg    7740

gattgtattc gaggaaactg gactaggtgc agccgtaacg gttaaggatg gactatactg    7800

cgactgaaaa gcggccagcc tcggccttgc ttgacgtata accaaagcca ccgggcaacc    7860

aatagcccct gtcacttttg atcaggtaga ccgaccctga agcgcttttt tcgtattcca    7920

taaaaccccc ttctgtgcgt gagtactcat agtataacag gcgtgagtac caacgcaagc    7980

actacatgct gactttaggc aagccactgc aactgcatgg atggcggggt gacggtgacc    8040

gtgccagctc ggcccgagca agctggacgc tgggcagacc catgacattg atgacggtgc    8100

g                                                                    8101
```

<210> SEQ ID NO 5
<211> LENGTH: 6737
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized,

<400> SEQUENCE: 5

```
taatccgctt cgtggccggg cttgcgctct gccagcgctg ggctggcctc ggccatggcc      60

ttgccgattt cctcggcact gcggcccggg ctggccagct tctgcgcggc gataaagtcg     120

cacttgctga ggtcatgacc gaagcgcttg acccagccgg ccatcttgct gcggtactcg     180

tccagcgccg tgcgccggtg gcggcttagc tgccgctcgg gcagttcgag gctggccagc     240

atgcgggcct tatcctgctg aagctgggcc tgctcgatct gatggccagc atgctgaacc     300

agcgcagggc cagcggtggc ggtcttgccc ttggattatc gatgcagatc aaacggatgc     360

taaccggcgc gggtcgtgtg cttgtccttg cggttcgtga agcccgccaa gcggccatag     420

tggccgctgt cggcgctggc cgggtcggcg tcgtactcgc tggccagcgt ccgggcaatc     480

tgcccccgaa gttcaccgcc tgcggcgtcg gcccacttga cccatgcctg atagttcttc     540

gggcttggtt ccactaccag ggcaggctcc cggccctcgg ctttcatgtc atccaggtca     600

aactcgatga ggtcgttatc cagcaccaga aactgcagct aatgctaggc gggcctgatc     660

taaacgtatt tgccctgggc attcatcagc ttgcgcaatg gagtgctatg gagcacttcg     720

gcggctgacc attcccggtt catcatctgg ccggtgggtg cgtccctgac gccgatatcg     780

aagcgctcac agcccatggc cttgagctgc cggcctatgg cctgcaaagt cctgtcgttc     840

ttcatcgggc cacccagcgc agccagatcg agccgtcctc ggttgtcagt ggcgtcaggt     900

cgagcaagag caaccatgcc atcagcagca ccaccgtagg catcatggaa gccagcatca     960

aggttagcca tagcttactg tgccaccccc gagacgagct aagggcgatc tgcgcggagc    1020

tgctcacctc ggcggctacc tcccgcaact ctttggccag ctccacccat gccgcccctg    1080

tatggcgctg ggcttatagc aactccgccg aatgcgaatc gatggcatgc ttggtctggc    1140
```

continued

```
tacagaactg aagggattcg taggccagtg tagccatgca atgggcaagc ggttcgatct    1200
gatccgatca atcgatgatg aatcaggata tattcaatca gtcgatgcag ttcaaggtat    1260
attgcctccc ggtattcctg taagtcgatg atctggcgct tggcggtgtc gatgctcagg    1320
gccacgtctg ccggtcggtc gcggatgccc cggccttcca tctccaccac gttcggcccc    1380
aggtgaacac cgggcaggcg ctcgatgccc tgcgcctcaa gtgttctgtg gtcaatgcgg    1440
gcgtcgtggc cagcccgctc taatgcccgg ttggcatggt cggcccatgc ctcgcgggtc    1500
tgctcaagcc atgccttggg attgagagct taggtcttct gtgcccagcc attctcaggg    1560
gtcttgacgt tgttcagctt gatccactga gaggcgggcc gatcgatgcc gtctttgatc    1620
cgctcggaga tcatcaggtg gcagtgcggg ttctcgccgc cacgggcatg gatggccagc    1680
gtatacggca ggcgctcggc acggtcaggt tgctggcgca actcggacgc cagcgccttc    1740
tgctggtcga gggtcagctc gacccgcagg gcaaattcga cctccttgaa cagccgcccc    1800
ttggcgagtc aatccaggtc ggcatgcatc cagtagtagg agggccgatc gacgaaatcc    1860
ggcatgtgcc aggcctaggc gtgcaagacc tactccatgc agcgggaatc attgccttcg    1920
cgctggatgc agtcggcctt ggcgctggcc gattggccgc ccgacctgct gccggctttc    1980
gccgcaaggc gataaatcgc catgctgcct cgctgctgct tttgctttcc ggctccatgc    2040
aatggccctc ggagagcgca ccgcccgaag ggtggccgtt aggccagttt ctcgaagaga    2100
aacccgtaag tgcgccctcc cctacaaagt agggtcggga ttgccgccgc tgtgcctcca    2160
tgataagaca agagaaagca aattaaaaaa ggggagtaaa gatggtaaag gggagcaaac    2220
aggcagagga taggctggcc aagcacgaag aacaacgagc gagaataact gacgaaattc    2280
agcggagcgg ggcaagggaa cagcagcaag agcgcaagaa cgaaacaagg cgcaaggtgc    2340
tggcgggggc catgattttg gccaaggtga acagcagcga gtggccggag gatcggctca    2400
tggcggcaat ggatgcgtac cttgaacgcg acccacgacc cgccttgttc ggtcagccgc    2460
cacgccagca ggatgagccg ggctgaatga ttgaccgaga caggccctgc ggggctgcac    2520
aagcagaaac aaccataggg taggggaaaa ggaagataaa gaggcaaaac gagaacaagc    2580
gtaaatatgc ggggatggca gtggggttta gaggggtttg aaagccttta aaaacgaagc    2640
gcagcggtgg ggcggtgtgc agcctagcgc agcgaataga ccagcatacc ggcctctggc    2700
cgggcatatt gggcaagggc agcagcgccc cacaaggggc ctgataaccg cgcctagtgg    2760
attattctta gatcatcatg gatggatttt tcccaacacc cgccagcccc cgcccctgct    2820
gggtttgcag gtttgggggc gtgacagtta ttgcaggggt tcgtgacagt tattgcaggg    2880
gggcagtgac agttatgcag gggttcgtga cagttagtac gggaggtgac gggcacggct    2940
ggcaatgtac agcacaggca ggcacatcgg atgaggggta aagaacttta agctaagaga    3000
tagactgtat gtaaacaacg tattgcaagg acgcggaaca tgcctcatgt ggcggccagg    3060
acggccagcc gggatcggga tactggtcgt taccagagcc accgacccga gcaaaccctt    3120
ctctatccga tcgttgacga gtattacccg gcattcgctg cgcttatggc agagcaggga    3180
aaggaattgc cgggctatgt gcaacgggaa tttgaagaat ttctccaatg cgggcggctg    3240
gagcatggct ttctaagggt tagctgcgag tattgcaacg aagagcaact ggtcgcttta    3300
agaaatcaat ctaaagtata tatgagtaaa cttggtctga caggcccctg aattcgcatc    3360
tagactgatg agtcgtggta gagccaaaaa aagccggtac aagcaagcat atccaggacc    3420
```

continued

```
atctgaatca tgcgcggatg aatcgaactc aagacggaga taccagaact tatcccaatg   3480
tacagaatat gagacaacga ggcaatagga agcaataata aggacagtat ttatactttc   3540
tgcccctcca aaagcaaaaa cccgccgaag cgggcttta cgtaaatcag gtgaaactga   3600
ccgactttgg cagtttattc ttgacatgta gtgaggggc tggtataatc acatagtact   3660
gtttatgcgt gaatgtatct caatcccacg tggccaggct ggagtccaaa ttggtaatgc   3720
ctgttgggaa ttatattgtc tggaacatgg catccaacct gatggtcaga tgcccatcag   3780
aataaaaatc ggaggaggag atgataagtt aactacttta ttatatgaat aaggagaagg   3840
aataacaagt aaaagagcgg ttttcataga attagagaat aaagaagtag aagaggttag   3900
cacaggcact tatcgtcaac tcttccatcc agaacaactg atcacgggca aagaggatgc   3960
agcaaacaat tatgcccgtg gtcattacac aattggcaag gaaattgttg accttgtctt   4020
ggatcgtatt cgtaaacttg cagatcaatg tactgggatc caaggattcc tcatcttcca   4080
ttaaattgga ggtggaactg gatatggtta aaaatatatg ttgataggat gtatgtatgt   4140
ggaatatgga aagtagtaaa aattggaata agcaatttat aaagcaaaaa aagtaaaaaa   4200
tgctgtggtc gaaacttata attcaattct cacaacccat actactttag aacattccga   4260
ttgtgcattt atggtcgaca atgaggctat ttatgatatc tgtcgccgta acttggacat   4320
cgaaagaccg acgtatacca acttgaaccg actgattggt caaatcgttt cttctatcac   4380
tgcctcactt cgtttcgatg gcgcccttaa tgttgatctg acggaattcc aaactaattt   4440
agtaaaataa aaaagaattc atttaaaata ggttaaataa gagaaagtaa tttaagatga   4500
aataagagta aaagaaaaac tttatgtagc ggaaaataag aatgaatgat tagaacaggc   4560
taatcagatg gtaaaatgcg atgctcgtca tggaaaatat atggcgtgtt gtatgctgta   4620
cagaggagat gttgtacgta aggatgttaa tgcagccatt gcaactatca agactaagcg   4680
caccattcaa tttgtcgatt ggtgccaaac tggattcaaa gttggcatta attatcaacc   4740
acctactgtt gtacccagtg gtgatcttgc taaagtacaa cgagctgttt gcatgttatc   4800
aaataataat gataaagaag aagaaaggga aagaaatgaa aataaatttg atttaatgta   4860
agaaaagaga gaaatagtaa aatggaaagt aggtgagggt atggaggaag gtgagtatac   4920
cgaagcacgt gaagatcttg ctgctttgga gaaggattat gaagagcttg gtatggattc   4980
cgcagaaggt gagggagagg gagctgaaga atactaaatc caacagtact atgtgattat   5040
accagccccc tcactacatg tcaagaataa actgccaaag gtaatcgtta atcccgcaat   5100
aacgtaaaaa cccgcttcgg cgggtttttt tatgggggga gtttagggaa agagcatttg   5160
taatatgcgt ataaagaagt taagtagatg ttaagaaaac aaaggaaata agagaatgcg   5220
aggcgatgga gggtaaaaac agagcggcag attaaatgac gactacattg gtgaaataga   5280
ctttcacgta gtggacaaat tcttccaact gatctgcgcg cgaggccaag cgatcttctt   5340
cttgtccaag ataagcctgt ctagcttcaa gtatgacggg ctgatactgg gccggcaggc   5400
gctccattgc ccagtcggca gcgacatcct tcggcgcgat tttgccggtt actgcgctgt   5460
accaaatgcg ggacaacgta agcactacat ttcgctcatc gccagcccag tcgggcggcg   5520
agttaaatag agttaaggta tattataggg aataaaatag atagagttat ggaaacggat   5580
caaagagttc ctccgccgct ggacctacca aggcaacgct atgttctctt gcttttgtca   5640
gaaagatagc aagaaaaatg tagaacgtgg atggatagaa gataacagaa agaaagtaat   5700
```

-continued

```
tgcgctgcca ttatcaacct tgcagttcgc gattagatgg atatcgaatc ggaatgatgt    5760

agtcgtgcac aacaatggtg aattcaacag agcggagaat atcgcatatt aacgggaacg    5820

ccgaagtttc caaaaggtcg ttgatcaaag ctcgccgcgt tgtttcatca agccttacgg    5880

tcaccgtaac cagcaaatca atatcactgt gtggcttcag gccgccatcc actgcggagc    5940

cgtacaaatg tacggccagc aacgtcggtt cgagatggcg ctcgatgacg ccaactacct    6000

ctgatagttg agtcgatact tcggcgatca ccgcttccct catgatgttt aactttgttt    6060

tagggcgact gccctgctgc gtaacatcgt tgctgctcca taacatcaaa catcgaccca    6120

aggcgtaacg agcttgctgc ttggatgccc gaggcataga ctgtacccca aaaaaacagt    6180

cataacaagc catgaaaacc gccactgcgc cgttaccacc gctgcgttcg gtcaaggttc    6240

tggaccagtt gcgtgagcgc atacgctact tgcattacag cttacgaacc gaacaggctt    6300

atgtccactg ggttcgtgcc ttcatccgtt tccacggtgt gcgtcacccg gcaaccttgg    6360

gtagcagcga agtcgaggca tttctgtcct ggctggtaac gaccaaaatc aattaaagtg    6420

agtcagactg aagccttggg aagggtgatg tagtacttgc aagccgagcc atcggttacc    6480

gtccagccca gcgcgaccag ctccggcaac gcctcgcgca cccgcttgcg gcgcttgcgc    6540

atggtcgaac cactggcctc tgacggccag acatagccgc acaaggtatc tatggaagcc    6600

ttgccggttt tgccggggtc gatccagcca cacagccgct ggtgcagcag gcgggcggtt    6660

tcgctgtcca gcgcccgcac ctcgtccatg ctgatgcgca catgctggcc gccacccatg    6720

aaggcatgcg agatcaaggg gttcagggcc aagtacaggc gaccgtaagc atcgacgatg    6780

gagtaatcag aaagcagccg aaccccatgc agctagaggc aattctgggc gatgatggat    6840

accttccaaa ggcgctcgat gcagtcctgt atgtgcttga gcgccccacc actatcgacc    6900

tctgccccga tttcctttgc cagcgcccga tagctacctt tgaccacatg gcattcagcg    6960

gtgacggcct cccacttggg ttccaggaac agccggagct gccgtccgcc ttcggtcttg    7020

ggttccgggc caagcactag gccattaggc ccagccatgg ccacccagcc ttgcaggatg    7080

agcagatcct aagcgaaccg aggctcaggg aagctgaact agatccgatt gaacgacgag    7140

tagtcatcag tcacgtaacg attgctgcgc ttgcgctagc aacgcatgcg ggcacggaac    7200

aggccggggg ccagacagtg cgccgggtcg tgccggacgt ggctgaggct gtgcttgttc    7260

ttaggcttca ccacggggca cccgcttgct cttgcgctgc ctctccagca cggcgggctt    7320

gagcaccccg ccgtcatgcc gcctgaacca ccgatcagcg aacggtgcgc catagttggc    7380

cttgctcaca ccgaagcgga cgaagaaccg gcgctggtcg tcgtccacac cccattcctc    7440

ggcctcgggc atggtcatgc tagtcaggta ggactgaacg aggacgttat agtccagtcc    7500

agcgctgccc aggctggcct gatgctggtc gactgcgacc atcctggacg agcccatgct    7560

ggcatggtgc aggaacacga tagagcaccc ggtatcggcg gcgatggcct ccatgcgacc    7620

gatgacctgg gccatggggc cgctggcgtt ttcttcctcg atgtggaacc ggcgcagcgt    7680

gtccagcacc atcaggcggc ggccctcggc ggcgcgcttg aggccgtcga accactccgg    7740

ggccatgatg ttgggcaggc tgccgatcag cggctggatc agcaggccgt cagccacggc    7800

ttgccgttcc taggcgatga ggtgcgcccc aagggcgtgc aggcggtgat gaatggcggt    7860

gggcgggtcc tcggcgggca ggtagatcac cgggccggtg ggcagttcgc ccacctccag    7920

aagatcaggc aagcctgcca tatgcgaggc aagtcgaagg gaacgcatgg atttacaggc    7980
```

continued

---

aacaccgggc gacccagcg accggaccgc accggcaacc atgttgggcc aaccgcagtc 8040

aagcggtggc ggcgctgctg agcccgacte cagcatattg atacgcttat gggtagacat 8100

tgatcgcctc ctttgcaggc agttggtggc taggcgctgg cggggtcact accacccgcac 8160

tgcgccgctc tgagttcttc caggcactgg cgcagcgcgt cgtattcgtc gtcggtcagc 8220

cagaacttgc gctgacgcat cccttcggcc ttcatgcgct cggcatatcg cgcttggcgt 8280

acagcgtcag ggctggccag caggtcgccg gtctgcttgt ccttcaggtc tttcatatca 8340

gtcaccgcga aacttgacgc ggccgaacgc attgacttgc aggtacaagg aacaggtgac 8400

gacgtaacgc ttacggatgg aactataagc gactgaaacg aggccagact aggcattgtt 8460

tgacgtataa ccaaagccac cgggcaacca atagcccttg tcacttttga tcaggtagac 8520

cgaccctgaa gcgctttttt cgtattccat aaaacccccc tctgtgcgtg agtactcata 8580

gtataacagg cgtgagtacc aacgcaagca ctacatgctg aaatctggcc cgcccctgtc 8640

catgcatcgc tggccgggtg aaggtgaccg tgacaggtcg gaacgcgatc gatggaagct 8700

gggcagaccc atgcacttgc tgacggtgcg atcgatg 8730

<210> SEQ ID NO 6
<211> LENGTH: 5820
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized.

<400> SEQUENCE: 6

tacccagctc agtggcaggg attgcgatcc gaccgcgatg ggctggactc gggcatggcc 60

ttgccgattc cctcggcact gcggcccggg ctggccagcc tctgcgcggc gataaagtcg 120

cacctgctga ggtcatgacc gaagcgcttg accagcccgg ccatctcgcc gcgggtactcg 180

tccagcgccg tgcgccgggtg gcggcatagc tgcccgctgg gcagttcgag gctggccagc 240

ctgcgggcct tctcctgctg ccgcatggcc tgctcgatct gctggccagc ctgcatgcacc 300

agcgccgggc cagcggtggc ggtcctgccc ttggattacc gacgcagacc aaccggatga 360

taccgggcgc gggtggtgtg attgccattg aggttggtga agcccgaacc gaggccatcg 420

tggcggctgc cggcgctggc cgggtcggcg tcgtactcgc tggccagcgt ccgggcaatc 480

tgcccccgaa gttcaccgcc tgcggcgtcg gccaccttga cccatgcctg atagttcttc 540

gggcatggtt cccatacccg ggcaggatcc cggccctcgg ctttcatgtc atcccaggtca 600

aacatcgctga ggtcgtcccac cagcaccaga ccatgccgct cctgcatggc gggcccgata 660

tacacgtacc tgccctgggc attcatccgc ttgacgcatg gagtgctatg gagcacttcg 720

gaggctgcac attccaggtc aatcaatcgg aaggcgggtg agtccctgcc gacgatatcg 780

aagcgctccc agcccatggc cttgagctgc cggcccatgg cctgcaaggt cctgccgttc 840

ttcatcgggc caccaagcgc agccagatcg agcggtcctc ggttgtcagt ggcgtcaggt 900

cgagcaagag caacgatgcg atcagcagca ccacccgtagg catcatggaa gccagcatca 960

cggtcagcca tagcctccag tgccacccac gcgacgcgct ccgggcgctc tgcgcggcgc 1020

tgcccacctc ggcggatcac taccgcaccc atttggcacg atccacaact gacgccactg 1080

tctggcgctg ggctttcagc cactccgccg cctgcggcctc gctggcctgc ttggtctggc 1140

taccgactg aagggattcg tagggcagtg tagccatgcc atgggcaagc ggttcgatct 1200

-continued

```
gatccgatca atcgatgatg aatcaggatc tattcaatcc gtcgaatgcg ttcacggtct    1260
attgcatccc ggtcatactg tacgacgatg atctgggagc tggcggtgtc gatgccaagg    1320
gccacgtctg ccaggtcggc gcggatgccc cggccttcca tctccaccac gttcggcccc    1380
aggtgaacac cgggcaggcg ctcgatgccc tgcgcctcaa gtgttctgtg gtcaatgcgg    1440
gcgtcgtggc cagcccgctc taatgcccgg ttggcatggt cggcccatgc ctcgcgggtc    1500
tgctcaagcc atgccttggg cttgagcgct tcggtcttct gtgccccgcc cttctccggg    1560
gtcctgacgc tgtacagctc gatccactga gaggcgggcc gatcgatgcc gtcctcgatc    1620
agcacggcga tactcaggtg gatgtcgggg ttctcgaagc acccggcatg gatggcaagc    1680
gtatacggca ggcgctcggc accggtcagg tgctggcgga actcggacgc cagcgccttc    1740
tgctggtcga gggtcagctc gaccggcagg gcaaattcga cctccttgaa cagccgccca    1800
ttggcgcgtt catacaggtc ggcagcatcc cagtagtcgg cgggccgctc gacgaactcc    1860
ggcacgtgcc aggcctaggc gtgcaagacc tactccatgc agcgggcatc attgccttcg    1920
agcacgatgc agtcggactc ggccccggcc gatttgcagc cagacctgct gacggctttc    1980
gccgcaaggc gataaatcgc catgcggccc cgctgctgcc tttgccttta ggctccatgc    2040
aatggcccgc ggagagcgca ccgcccgaag ggtggccgtc aggccagttt ctcgaagaga    2100
aacccgtaag tgcgccctcc cctacaacgt agggtcggga ttgccgccgc tgtgccttca    2160
tgatagccta cgagacagca cattaacaat ggggtgtcaa gatggttaag gggagcaaca    2220
aggcggagga taggctggcc aagcacgacg aacaacgagc gagcatacat gacgaaattc    2280
agcgggagcg ggctagggca aagcagacag agcgcaagca agcaacaagg agcaaggtgc    2340
tggtcggggc catgattttg gccaaggtga acagcagcga gtggccggag gatcggctca    2400
tggccgcaat ggatgcgtac cttgaacgcg accacgaccg cgccctgttc ggtctgccgc    2460
cacgccagca ggatgagccg ggctgaatga tcgaccgaga caggccctgc ggggctgcac    2520
acgcgccccc accctcgggg taggggggaa ggccgctaaa gcggctaaaa gcgctccagc    2580
gtacatatgc ggggatagga gtgggattta gagggcattg aacgccattc aacaggacgc    2640
gacgcggtgg ggcggtgtga agccaagcgc agcgaataga aacgcatacc ggcccatggc    2700
cgggcatatt gggcaagggc agcagcgccc cacaaggggg ctgataaccg cgcctagtgg    2760
attattctta gataatcatg gatggatttt tccaacaccc cgccagcccc cgcccctgct    2820
gggtttgcag gtttgggggc gtgacagtta ttgcaggggt tcgtgacagt tattgcaggg    2880
gggcgtgaca gttattgcag gggttcgtga cagttagtac gggagtgacg ggcactggct    2940
ggcaatgtct agcaaggcaa ggcaatacgg atgaggctca aagtacattc agctaagaga    3000
tagactgtcc gtcaacaacg tattgcaagg aagcggaaca tgcctcatgt ggcggccagg    3060
acggccagcc gggatcggga tactggtcgt taccagagcc accgacccga gcaaaccctt    3120
ctctatcaga tcgttgacga gtattacccg gcattcgctg cgcttatggc agagcaggga    3180
aaggaattgc cgggctatgt gcaacgggaa tttgaagaat ttctcccaat cgggcggctg    3240
gagcatggcc ttctacgggc tcgctgcgag tcttgccacg ccgagcacct ggtcgctttc    3300
agcaataccc atcaagtata tatgagtaca attggcatga aaggccaatg aattcgaatc    3360
tagatggtag agccacaaac agccggtaca agcaacgatc tccaggacca tctgaatcat    3420
gagcggatga aacgaatcaa agacggagcc aacgaaatac aaacaaaagt aacgacaatg    3480
```

continued

```
agacaaagtg gaaattagtc gaaataacga atacatgaac tggccgataa ttgcagaaga   3540
aaggaaagtc attcatattt atgcccatcc aaaagcaaca aaccgcagaa gagggatttt   3600
acgtaaatca ggtgaaactg accgactttg gcagtttatt cttgacatgt agtgaggggg   3660
ctggtataat cacatagtac tgtttatgat ggctgtagca gcggcgcaga agaaccgcga   3720
gatgatcgca atccagaaat cttacagcat cgagaatgga tatccggcga ggcgacgatc   3780
tctcgtggac gatgctcgga ttgaaactta ggtcgatcag caaacgaagc agagcgtttt   3840
ggaagaagca agtcaaaggg agaatgaagc tggactgaac gaggaggaag tggtgctggc   3900
aaagaagata gagcaatgcc aggatgagca gaacacggtg caaacaagag agctagtgct   3960
gcgcctgaga gagggatcgg gttcactcgc gaggatcctg aagacgatcg agaacttcaa   4020
ggggacggtg accacgtggg agtcccggcc gtccaagaag gagggcctcc agttcgacgt   4080
gctcgtccag gtcgacatga cgaggcaata cttgcatcca ctgcacagga atctccgccc   4140
gagcaaagcc atggaaggcg taacccagca agcggaaaac taggacagca taaaggatcc   4200
atggataccg aggcaagcga aagacctcga aactagaaac aacccgatga aaacaataga   4260
gcctgaccta gacatgaatc atcctggatt cgcggacaag gagtaccgcg cccgtagaaa   4320
gttcatcgcc gaaattgcat tcgcctacag gtacggtgac gcgataccca ccgtcccgta   4380
cacggaaacg gaaaccgaga cttggacccg ggtgatcaac accctcgtcg atctggtgcc   4440
gaaacacgca tgcgcggaat atcgaagaaa ttttaagaag atgcaggagg agaagatatt   4500
agaggaaaac agtaaaaatc agctgcagga ggtgagagag ttataggaga agaacaaagg   4560
gttcaagttg aggccagcgg aaggactgca aacgtcaagg gaatacttgt agaggatggc   4620
gttcagagtg ttccagagca ctcagtacat tcgtcacatc aagagcccgt atcatacccc   4680
ggaaccagat tgcatccacg agctgttggg acacatgccg ttgttggccg atccaagctt   4740
cgccccaatc tcccaggaaa tcggcctcgc ctcgctcggt gcctccgacg aggaaatcga   4800
gaaactgtcc acgaattaca ggttcaccga cgaactcggc ctctgcaagg aggggccgga   4860
tgtgaaagcc taggatgccg gactaatgtc agccaaaggg gagcaaattc aacgcattac   4920
gagcggaaaa tgcgagaatc gacccatcga gaagaaatcc aacgccgttc aaaagtaaca   4980
ggatcaagat taccaacgga tctacctcgt cgcggatagc ttcgaggatg ccaaagagaa   5040
attccgccgt tgggtgtcca ccatgagccg acctttgagg gttaggtacg atccgtacac   5100
gcaagcgggt gagaacctgg acagcgtgga caggatggac aatcagatgg ctcaagtgaa   5160
cacggaaatg acgcatctca cgaacgctga caacaagctg aagacatcgt tcgcgtaaat   5220
aaaacagtac tatgagatta taccagaccc ataactaaaa gtaaagaata aactgcaaaa   5280
ggtaataagt aatccgaaaa tacagtaaaa aaccgcttcg gagggatttt ttatgagggg   5340
agttcaggga aagagcattt gtcagctgga tcagcagatg ggaacccgga actcctcagc   5400
ctccagagta gccaatgaga cgccgagcgg ccgcctattt gccgactacc ttggtgatct   5460
cgccttccac gtagtggaca aattcttcca actgatctgc gcgcgaggcc aagcgatctt   5520
cttcttgtcc aagataagcc tgtctagctt caagtatgac gggctgatac tgggccggca   5580
ggcgctaaat tgcccagtcg gaagcgaaat aattcggagc gatttagaag gtaacagagc   5640
tgtaccaaat gcgggacaac gtaagcacta catttcgctc atcgccagcc cagtcgggcg   5700
gagagttcaa tagcgttaag gtttcattta gagcaaaaaa tagaacatgt taaggaaacg   5760
```

-continued

```
gatcaaagag ttcctcagcc gatggaacta accaggaacc gatacgttct attgcatttg    5820
tacgcaagca agccagatca atgtcgatcg tggcaggatc gacgataccc gacagaatgt    5880
cattgcgctg ccattctcca aattgcagtt cgcgcttagc tggataacgc cacggaatga    5940
tgtcgtcgtg cacaacaatg gtgacttcta cagcgcggag aatctcgctc tctccagggg    6000
aagccgaagt ttccaaaagg tcgttgatca aagctcgccg cgttgtttca tcaagcctta    6060
cggtcaccgt aaccagcaaa tcaatatcac tgtgtggctt caggccgcca tccactgcgg    6120
agccgtacaa atgtacggcc agcaacgtcg gttcgagatg gcgctcgatg acgccaacta    6180
cctctgatag ttgagtcgat acttcggcga tcaccgcttc cctcatgatg tttaactttg    6240
ttttagggcg actgccctgc tgcgtaacat cgttgctgct ccataacatc aaacatcgac    6300
ccacggcgta acgcgcttgc tgcttggatg cccgaggcat agactgtacc ccaaaaaaac    6360
agtcataaca agccatgaaa accgccactg cgccgttacc accgctgcgt tcggtcaagg    6420
ttctggacca gttgcgtgag cgcatacgct acttgcatta cagcttacga accgaacagg    6480
cttatgtcca ctgggttcgt gccttcatcc gtttccacgg tgtgcgtcac ccggcaacct    6540
tgggcagcag cgaagtcgag gcatttctgt cctggctggc catgaccaaa atcccttaac    6600
gtgagtcagc ctgcgcgctt gggcgcggtg atgtcgtact tgcccgcccg gaactcggtt    6660
accgtccagc ccagcgcgac cagctccggc aacgcctcgc gcacccgctt gcggcgcttg    6720
cgcatggtcg aaccactggc ctctgacggc cagacatagc cgcacaaggt atctatggaa    6780
gccttgccgg ttttgccggg gtcgatccag cccacagcc gatggtgcag caggcgggcg    6840
gttccgatgc caagcgaccg caccccgtcc atgcctgatg gaccatgatg gacgccaacc    6900
atgacggcct gcgccgatca ggggttcagg gccacgtaca ggcgcccgtc cgcctcgtcg    6960
ctggcgtact ccgacagcag ccgaaacccc tgccgcttgc ggccattctg ggcgatgatg    7020
gataccttcc aaaggcgctc gatgcagtcc tgtatgtgct tgagcgcccc accactatcg    7080
acctctgccc cgatttcctt tgccagcgcc cgatagctac ctttgaccac atggcattca    7140
gcggtgacgg cctcccactt gggttccagg aacagccgga gctgccgtcc gccttcggtc    7200
ttgggttccg ggccaagcac taggccatta ggcccagcca tggccaccag cccttgcagg    7260
atgcgcagat catcagcgcc cagcggctcc gggccgctga actcgatccg cttgccgtcg    7320
ccgtagtcat acgtcacgtc cagcttgctg cgcttgcgct cgcccgcctt gagggcacgg    7380
aacaggccgg gggccagaca gtgcgccggg tcgtgccgga cgtggctgag gctgtgcttg    7440
ttcttaggct tcaccacggg gcaccccctt gctcttgcgc tgcccctcca gcacggcggg    7500
cttgagatcc aagccgtacc gaagcctgca aacccgatca gagcacggtg agccatagtt    7560
ggccctgatc aaacccgagc ggacggagca aaggccatgg tagtcgtacc aaccccattc    7620
ctcggcctcg gcgctggtca tgctcgacag gtaggactgc cagcgggatg ttatcgaccg    7680
taccgagctg ccccggctgg cctgctgctg gtcgcctgcc ccatcatggc ccgcgccctt    7740
gctggcatgg tgcaggaaca cgatagagca cccggtatcg gcggcgatgg cctccatgcg    7800
accgatgacc tgggccatgg ggccgctggc gttttcttcc tcgatgtgga accggcgcag    7860
agtgtccagc accctcggcc ggcggcaatc ggcggcgagc ttgagggagt agcaccaatc    7920
cggggccatg atgttgggca ggctgccgat cagcggctgg atcagcaggc cgtcagccac    7980
ggcctgacgc tactcggcgc tgaggtgcgc cacctgggcg tgcccggggt gatgaatggc    8040
```

-continued

```
ggtgggaggg tatttcgggg gaaggtagcc aacggggacg gtgggcagtt agcccaaatc    8100

aagcagatcc ggcccgactg aactcagtgc ggccagttgc aggggcagcc tggccttacc    8160

ggcaccacgg ggcgacaaca gcgccccgac cgtaccggcc accatgttgg gcaaaacgta    8220

gtccagcggt ggcggcgctg ctgcgaacgc ctccagaata ttgataggct tatgggtagc    8280

cattgattgc ctcctttgca ggcagttggt ggttaggcgc tggcggggtc actacccccg    8340

ccctgcgccg ctctgagttc ttccaggcac tcgcgcagcg cctcgtattc gtcgtcggtc    8400

agccagaacc tgcgctgacg aatccctttg gacttcatgc gatcggaatc tagcgcttgg    8460

agtacagcga caggggtggc aagcaggtcg caggcctgca tgtccctttg gtctccaatc    8520

tcagtcaccg agaaacttgc cggggccgaa aggcctgtcc tcgccgaaca aggacaaggt    8580

gcagccgtca aggttaaggc tggccatatc agcgactgaa aagcggccag cctcggcctt    8640

gtttgacgta tcaccaacgc cacgggccca ccaatagccc ttgtcacttt tgatcaggta    8700

gaccgaacct gaagcgcttc tttcgcattc aatcaaaacc aattcagtgc gtgacgaatc    8760

atcgcatcac aggcgtgacg accaacgcca gaacaaactg atgcaatctg gacccgcact    8820

gtccatgcct cgctgcgggg gtgccggtgc ccgtgccagc tcggcccgcg caagctggac    8880

gctggccaga ccgatgacct tgctgacggc gcgctcgatg                          8920
```

<210> SEQ ID NO 7
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized,

<400> SEQUENCE: 7

```
cgtctaattc cacgaggatt                                                  20
```

<210> SEQ ID NO 8
<211> LENGTH: 7703
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pRK40 plasmid
      backbone sequence, circular

<400> SEQUENCE: 8

```
ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt      60

tgcctgactc cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag     120

tgctgcaatg ataccgcgag acccacgctc accggctcca gatttatcag caataaacca     180

gccagccgga agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc     240

tattaattgt tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt     300

tgttgccatt gctacaggca tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag     360

ctccggttcc caacgatcaa ggcgagttac atgatccccc atgttgtgca aaaaagcggt     420

tagctccttc ggtcctccga tcgttgtcag aagtaagttg gccgcagtgt tatcactcat     480

ggttatggca gcactgcata attctcttac tgtcatgcca tccgtaagat gcttttctgt     540

gactggtgag tactcaacca agtcattctg agaatagtgt atgcggcgac cgagttgctc     600

ttgcccggcg tcaatacggg ataataccgc gccacatagc agaactttaa aagtgctcat     660
```

-continued

```
aattggatta agttcttcgg gggaaatat atctaggatc ttaccgatgt tgagattaag      720

ttcgatgtta aaaactagtg aaccaaactg atattcagca tatttaattt taaacagagt      780

ttctgggtga gcaaaaacag gaaggcaaaa tgcagcaaaa aagggaataa gggcgacacg      840

gaaatgttga atactcatac tcttcctttt tcaatattat tgaagcattt atcagggtta      900

ttgtctcatg agcggataca tatttgaatg tatttagaaa aataaacaaa taggggttcc      960

gcgcacattt ccccgaaaag tgccacctga cgtgaccaaa atcccttaac gtgagttcgc     1020

atgccgaatt gggccgggtg atgtcgttcc tgcccgaagc gattcgggtt aacgttaagc     1080

aatgcgagtc aagctaaggc aacgcctcgc gattccgatt gaggcgattg agtttggtcg     1140

aaacactggc ctctgacggc cagacatagc cgcacaaggt atctatggaa gccttgccgg     1200

ttttgccggg gtcgatccag ccacacagcc gctggtgcag caggcgggcg gtttcgctgt     1260

ccagcgcccg caccttgtcc atgctgatgc gcacatgctg gccgccaccc atgacggcct     1320

gcgcgatcaa ggggttcagg gccacgtaca ggcgcccgtc cgcctcgtcg ctggcgtact     1380

aagacagatg aagtaaaccc tgccgcttgc ggccattatg ggcgatgatg gataccttcc     1440

aaatggagtc gatgcagtcc tgtattgtgt tgaggcaacc aaaaccatcg aactcagacc     1500

cgatttcctt tgccagcgcc cgatagctac ctttgaccac atggcattca gcggtgacgg     1560

cctcccactt gggttccagg aacagccgga gctgccgtcc gccttcggtc ttgggttccg     1620

ggccaagcac taggccatta ggcccagcca tggccaccag cccttgcagg atgcgcagat     1680

catcagcgcc cagcggctcc gggccgctga actcgatccg cttgccgtcg ccgtagtcat     1740

aagtcaagtc aagcatgctg agttgagcat agccccgatt gagggcaagg aacatggaag     1800

gggccagtca gtcgtcaggg tagtgcagga agtggctgag gatgagattg ttctaaggat     1860

tcaccacggg gcacccattt gctcttgcgc tgcctctcca gcacggcggg cttgagcacc     1920

cgcccgtcat gccgcctgaa ccaccgatca gcgaacggtg cgccatagtt ggccttgctc     1980

acaccgaagc ggacgaagaa ccggcgctgg tcgtcgtcca caccccattc ctcggcctcg     2040

gcgctggtca tgctcgacag gtaggactgc cagcggatgt tatcgaccag taccgagctg     2100

accccgatgg aatgatgatg gtagcatgag aaactaatgg aagagaattt gatggaatgg     2160

tgcaggaaca cgatagagca cccggtatcg gcggcgatgg cctccatgcg accgatgacc     2220

tgggccatgg ggccgctggc gtttacttcc tcgatgtgga accggcgcag cgtgtccagc     2280

accatcaggc ggcggccctc ggcggcgcgc ttgaggccgt cgaaccactc cggggccatg     2340

atgttgggca ggctgccgat cagcggctgg atcagcaggc cgtcagccac ggcttgccgt     2400

taatcgggcg tgatgtgagc aattagggcg tgattgggat gatgaatggc ggtgattggg     2460

tattcgggcg gattgtagta aacgtggcag gtagccaatc aagcagatac ggacccgaat     2520

ggcccgaatg aatctgtgcc ggccagttgc agggccagca tggtattacc ggataccaag     2580

ggcgacacca gcgccccgac cgtaccggcc accatgttgg gcaaaacgta gtccagcggt     2640

ggcgtaaatg atgcgaacgc atcaagaata ttgataggat tatggtagcg aattgattgc     2700

ctccttgcca ggcagttggt ggttaggcgc tggcgggggc actacccccg ccctgcgccg     2760

atatgagttc ttcaaggatc tagagcagag aatagtattc gtagtcggtc agcaagaact     2820

tgcgctgacg catccctttg gccttcatgc gctcggcata tcgcgcttgg cgtacagcgt     2880

cagggctggc cagcaggtcg ccggtctgct tgtccttttg gtcttccata tcagtcacgg     2940
```

-continued

```
agcaaattgc aggggaagca aggcaagtca tagcggaaca aggacaaggt gaagccgtca   3000
aggaaaaggc tggccatatc agcgactgca aagcggaaag aatcggaatt gtttgaagta   3060
taaccaaagc cacgggcaa ccaatagccc ttgtcacttt tgatcaggta gacgacccct   3120
gaagcgcttt ttgcgtattc cataaaaccc ccttctgtgc gtgagtactc atagtataac   3180
aggcgtgcgt acccacgcca gcacctacatg ctgaaatctg gcccgcccct gtccatgcct   3240
cgctggcggg gtgccggtgc ccgtgccagc tcggcccgcg cagcctggac gctgggcaga   3300
aacatgaccc tgcagaaggc gagcacgatg taatccgata agtggcaggg attgcgatat   3360
gacaagagca ggatggaatc ggccatggcc ttgccgattc aatcggaact gaggccaagg   3420
ctggccagct tctgcgcggc gataaagtcg cacttgctga ggtcatgacc gaagcgcttg   3480
accagcccgg ccatctcgct gcggtactcg tccagcgccg tgcgccggtg gcggctaagc   3540
tgcccgctcgg gcagttcgag gctggccagc ctgcgggcct tctcctgctg ccgctgggcc   3600
tgctcgatct gctggccagc ctgctgcacc agcgccgggc cagcggtggc ggtcttgccc   3660
ttggattcac gcagcagcac ccacggctga taaccggcgc gggtggtgtg cttgtccttg   3720
aggaaggtga agccgacaca gaggccactg tggggatgca aggcgatggc aggggcgggcg   3780
tcgtactcgc tggccagcgc ccgggcaatc tgccccgaa gttcaccgcc tgcggcgtcg   3840
gccacctga ccacgccgtg atagttcttc gggcgcgttt ccacgaccag ggcaggctcc   3900
cggccctcgg ctttcatgtc atccaggtca aactcgctga ggtcgtcccac cagcaccaga   3960
ccatgccgct cctgctcggc gggcctgata tcaccgtcat tgcccctgggc attcatccgc   4020
ttgagaactg gagtgttcatg gagcacttcg gaggcgagcc attcccggtt catcatcagg   4080
acggcgggtg agtcccatgcc gacgatatcg aagcgctacc agcccatggc attgagatgt   4140
cggcctatgg cctgcaaagt cctgccgttc ttcatcgggc caccaagcgc agccagatcg   4200
agccgtcgtc ggttgtcagc ggcgccaggc cgagcaagag caacgatgcg atcagcagca   4260
ccaccgtagg catcatggaa gccagcatca cggttagcca tagcttccag tgccaccccc   4320
gcgacgcgct ccgggcgctc tgcgcggcgc tgctcacctc ggcggctacc tcccgcaact   4380
attaggaaag atcaacaccc gacgcccatg tatggcgatg ggcttcaagc aactccgacg   4440
cctgcgcgtc gctggcctgc ttggtctggc tcatgacctg ccgggcttcg tcggccagtg   4500
tcgccatgcc ctgggccagc ggttcgatcc gctcgcgctaa ctcgctgatg cctctgcatt   4560
tcttcatctt gtcgattgcg ttcatggtct attgcctccc ggtattcctg taagtcgatg   4620
atctgggcgt tggcggtgtc gatgttcagg gccacgtctg cccggtcggt gcggatgccc   4680
aggccttcca tatccaaccc gttcggaacc aggtgaaacc agggcaggcg atcgatgacc   4740
tgcgcataca gtgtaatgtg gtcaacgcgg gagtcgtggc aagcccgatc taatgcaagg   4800
ttggcatggc aggcaatgcc atcgcgggtc tgatcaagcc atgccatggg attgagagat   4860
tcggtcttcc gtgccccgcc cttctccggg gtcttgccgc tgtaccgctt gaaccactga   4920
gagacgggcc gatcgatgcc gtcaatgatc agatcgagaa tattcaggtg gaagagaggg   4980
ttctcgccgc cacggcatg gatggccagc gtatacggca ggcgctcggc accggtcagg   5040
tgcatgggca aatcggacgc aagcgcattc tgcatggtag gggtcagatc gaaggtaagg   5100
gcaaattcga cctccttgaa cagccgccca ttggcgcgtt catacaggtc ggcagcatcc   5160
cagtagtcgg cgggccgatc gacgaactcc ggcatgtgcc cggattcggc gtgcaagact   5220
```

-continued

tacccaatgc agcggaccta attgccttcg agctggatgc agtcggactt ggccccggcc 5280

gatcggacgc aagcccatgc gacgggtttc gacgcaaggc gatcaatcgc catgccgcct 5340

cgccgttgcc tttgcttttc ggctccatgc aatggccctc ggagagcgca ccgcccgaag 5400

ggtggccgtc aggccagttc ctcgaagaga aacccgtaag tgcgccctcc cctacaaagt 5460

agggctggga ttgccgccgc tgtgcctcca tgatagccta cgagacagca cattaacaat 5520

ggggctgtca gatggttaag gggagcaaca aggcggcgga tcggctggcc aagctcgaag 5580

aacaaagcgc gagcatcccc gacgaaattc agcgggagcg ggcccaggcc cagccagcag 5640

agcgcaagca agccaaaagg agccaggtgc tggtgggggc catgaccttg gaccaggtgc 5700

acagcagcga gtggccggag gatcggctca tggcgccaat ggatgcgtac cttgaacgcg 5760

accacgaccg cgccttgttc ggtcagccgc cacgccagaa ggatgagccg ggctgaatga 5820

tcgaccgcga cagccccagc ggggctgccc acgcgccccc accctccggg tagggggaaa 5880

ggccgctaaa gcggctaaaa gcgctccagc gtatttctgc ggggcttggt gtgggcttta 5940

gagggatttg accgcatttc accctgacgc gacgcggtgg ggcggcgtgt agccctagac 6000

agcgaatcga accgctatcc ggccctcggc agggcatatc gggcaagggc agccgcgacc 6060

cacaagggcg ctgataaccg cgcctagtgg attatcctta gataatcatg gatggatttt 6120

tccaacaccc cgccagcccc cgccctgcct gggtttgcag gtttgggggc gtgacagtta 6180

ttgcaggggc tcgtgacagc tattgcaggg gggcgcgaca gttatcgcag gggtccgtga 6240

cagctagtac gggagtgacg ggcacctggc ggcaatgtcc agcaacggca ggcaattcgg 6300

atgagggtca aagcaatttc agctaagcga tagcccgtcc gtcaacaccg tattgcaagg 6360

aagcggacca tgcccaatgc gggcgccagg aaggccagcc gggaccgggc tactggtagt 6420

taccagagcc accgaccgga gcaaacccct ctctatcaga tcgttgacga gtattacccg 6480

gcattcgctg cgcttatggc agagcaggga aaggaattgc cgggctatgt gcaacgggaa 6540

tttgaagaat ttctccaatg cgggcggctg gagcatggct ttctacgggt tcgctgcgag 6600

tcttgccacg ccgagccacc tggtcgcttc agaatccaat ctaaagtata tatgagtaaa 6660

attggtatga caggccacta gagcccagtc catcccagtg atcgagattg aaatcccatc 6720

cagtgataga gatactgagc acggatctga aagaggagaa aggatctatg gcgagtagcg 6780

aagacgttat caaagagttc atgcgcttca aagttcgtat ggaaggttcc gttaacggtc 6840

acgagttcga aatcgaaggt gaaggtgaag gtcgtccgta cgaaggtact cagaccgcta 6900

aactgaaagt taccaaaggt ggtccgctgc cgttcgcttg ggacatcctg tccccgcagt 6960

tccagtacgg ttccaaagcc tacgtcaaac acccggctga catcccggac tacctgaaac 7020

tgtccttccc ggaaggtttc aaatgggaac gtgttatgaa cttcgaagac ggtggtgttg 7080

ttaccgttac ccaggactcc tccctgcaag acggtgagtt catctacaaa gttaaactgc 7140

gtggtactaa cttcccgtcc gacggtccgg ttatgcagaa aaaaaccatg ggttgggaag 7200

cttccaccga acgtatgtac ccggaagacg gtgccctgaa aggtgaaatc aaaatgcgtc 7260

tgaaactgaa agacggtggt cactacgacg ctgaagttaa aaccacctac atggctaaaa 7320

aaccggttca gctgccgggt gcttacaaaa ccgacatcaa actggacatc acctcccaca 7380

acgaagacta caccatcgtt gaacagtacg aacgtgctga aggtcgtcac tccaccggtg 7440

cttaataagg atctccaggc atcaaataaa acgaaaggct cagtcgaaag actgggcctt 7500

continued

```
tagctttatc tgttgtttgc aggtgaacgc tctcctactc agtcacaatg gatcacattc   7560

gggtgggcct ttctgcgttt atacgtaggc atccccgagc ggccgcgatt atcaaaaagg   7620

atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta aagtatatat   7680

gagtaaactt ggtctgacag tta                                           7703
```

<210> SEQ ID NO 9
<211> LENGTH: 7608
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pEK402 plasmid
      backbone sequence, circular

<400> SEQUENCE: 9

```
ttagaaaaac tcatcgagca tcaaatgaaa ctgcaattta ttcatatcag gattatcaat     60

accatatttt tgaaaaagcc gtttctgtaa tgaaggagaa aactcaccga ggcagttcca    120

taggatggca agatcctggt atcggtctgc gattccgact cgtccaacat caatacaacc    180

tattaatttc ccctcgtcaa aaataaggtt atcaagtgag aaatcaccat gagtgacgac    240

tgaatccggt gagaatggca aaagcttatg catttctttc cagacttgtt caacaggcca    300

gccattacgc tcgtcatcaa aatcactcgc atcaaccaaa ccgttattca ttcgtgattg    360

cgcctgagcg agacgaaata cgcgatcgct gttaaaagga caattacaaa caggaatcga    420

atgcaaccgg cgcaggaaca ctgccagcgc atcaacaata ttttcacctg aatcaggata    480

ttcttctaat acctggaatg ctgttttccc ggggatcgca gtggtgagta accatgcatc    540

atcaggagta cggataaaat gcttgatggt cggaagaggc ataaattccg tcagccagtt    600

tagtctgacc atctcatctg taacatcatt ggcaacgcta cctttgccat gtttcagaaa    660

caactctggc gcatcgggct tcccatacaa tcgatagatt gtcgcacctg attgcccgac    720

attatcgcga gcccatttat acccatataa atcagcatcc atgttggaat ttaatcgcgg    780

cctcgagcaa gacgtttccc gttgaatatg gctcataaca ccccttgtat tactgtttat    840

gtaagcagac agttttattg ttcatgatga tatattttta tcttgtgcaa tgtaacatca    900

gagattttga gacacaacgt ggctttgttg aataaatcga acttttgctg agttgaaggc    960

tcagtcatga ccaaaatccc ttaacgtgag tcagcctgcc gcctggggcc gggtgatgtc   1020

gtaccctgcc gccgcgaact cggttaccgt ccagcccagc gcgaccagct ccggcaacgc   1080

ctcgcgcacc cgctggcggc gcttgcgcat ggtcgaacca ctggcctctg acggccagac   1140

atagccgcac aaggtatcta tggaagcctt gccggttttg ccggggtcga tccagccaca   1200

cagccgctgg tgcagcaggc gggcggtttc gctgtccagc gcccgcacct cgtccatgct   1260

gatgcgcaca tgctggccgc cacccatgac ggcctgcgcg atcaaggggt tcagggccac   1320

gtacaggcgc ccgtccgcct cgtcgctggc gtactccgac agcagccgaa acccctgccg   1380

cttgcggcca ttctgggcga tgatggatac cttccaaagg cgctcgatgc agtcctgtat   1440

gtgcttgagc gccccaccac tatcgacctc tgccccgatt tcctttgcca gcgcccgata   1500

gctacctttg accacatggc attcagcggt gacggcctcc cacttgggtt ccaggaacag   1560

ccggagctgc cgtccgcctt cggtcttggg ttccgggcca agcactaggc cattaggccc   1620

agccatggcc accagccctt gcaggatgcg cagatcatca gcgcccagcg gctccgggcc   1680

gctgaactcg atgcgcttgc cgtcgccgta gtcatacgtc acgtccagct tgctgcgctt   1740
```

-continued

```
gagctagccc agctagaggg aacggaacag gacggaggcc agtcagtgcg aagggtagtg    1800
aaggaagtgg atgaggatga gattgatata aggctaaacc aagggggacc aactagatat    1860
tggcctggcc ctccagcacg gcgggcttga gcacccegcc gtcatgccgc ctgaaccacc    1920
gatcagcgaa cggtgcgcca tagttggcct tgctcacacc gaagcggacg aagaaccggc    1980
gctggtcgtc gtccacaccc cattcctcgg cctcggcgct ggtcatgctc gacaggtagg    2040
actgccagcg gatgatatcg accagtaccg agctgccccg gctggcctgc tgctggtcgc    2100
atgcgaacct aatggaagcg aactagatgg aatggtgacg gatcacgata gagcacaagg    2160
tatcggaggc gatggaatcc atgcgaacga tgacctgggc aatgggaccg atggcgtttt    2220
cttcctcgat gtggaacggc cgcagcgtgc ccagcaccat caggcggcgg ccctcggcgg    2280
cgcgcttgag gccgtcgaac cactccgggg ccatgatgtt gggcaggctg ccgatcagcg    2340
gctggatccg caggccgtca gcccacggct gccgttcctc ggcgctgagg tgcgccccaa    2400
gggcgtgcag gaggagatga atggcggtgg gaggggatac ggcggggagg tagaacaacg    2460
ggccggtggg aagtaagccc aactccagca gatccggacc gactgcaatc tgtgcggacc    2520
gttgcagggc cagcatggat ttaccgcgac caccggcgca caccagcgcc ccgaccgtac    2580
cggccaccat gttgggcaaa acgtagtcca gcggtgcggc cgctgctgcg aacgcctcca    2640
gaacattgac aggcctatgg gtagccattg attgcctcct ttgcaggcag ttggcggtta    2700
ggcgctggcg gggtcactac ccccgcccta cgccgctcta agttcctcca ggcacacgcg    2760
aagcgaatag tattcgtagc aggtcagcca gacctagagc tgacgcatcc atttggaatt    2820
aatgcgatag gaatatagcg attggcgtac agcgacaggg atggccagca ggtagccggt    2880
ctgcctgtcc ttttcgtctt tcatatcagt caccgagaaa cttgccgggg ccgaaaggct    2940
tgtccttcgc gaacaaggac aaggtgcagc cgtcaaggtt aaggctggcc atatcagcga    3000
ctgaaaagcg gcccgccctc gccttgtttg acgtataacc aaagccaccg ggcaaccaat    3060
agcccttgtc acttttgatc aggtagaccg acccagaagc gcttttttcg tattccatca    3120
aacccaattc tgtgcgtgag tactcatagc ataacaggcg tgagtaaaaa agcaagaact    3180
aaaagatgaa atatggaacg aacccgtcca tgcctcgatg gagggatgcc ggtgccagtg    3240
ccagctcggc ccgcgcaagc tggacgctgg gcagacccat gacctcgctg acggtgcgct    3300
cgatgtaatc cgcttcgtgg ccgggcttgc gctccgccag cgctggcctg gcctcggcca    3360
tggccttgcc gatttcctcg gcactgcggc ccggcctggc cagctcctgc gcggcgatca    3420
agtcgcactt gctgaggtca tgacccgagc gcttgaccag cccggccatc tcgcctgcgt    3480
aatcgtaacg agcagtgcgc aggtaaaggc taagcaagag atcggaaagt tagaagatgg    3540
aaagcatgcg ggcctatacc tgatgcagca gggcctgatc gatcagatgg aaagcctgat    3600
gcaccagcgc cgggccagcg gtggcggtcc tgccctcgga ttcacgcagc agcaccccag    3660
gctgataacc ggcgcgggtg gtgtgcttgc ccttgcggtt ggtgaagccc gccaagcggc    3720
catagtggcg gctgtcggcg ctggccgggt cggcgtcgta ctcgctggcc agcgtccggg    3780
caatctgccc ccgaagttca ccgccagcgg cgtcgggcac cttgacccat gcctgatagt    3840
tatacgggca ggttaaaaca aacaaggcag gatcccggcc atcagcttta atgtcataca    3900
ggtcaaactc gctgaggtcg tccaccagca ccagaccatg ccgctcctgc tcggcggccc    3960
tgaaatacac gtaactgccc tgggcattca taagcatgag aaatggagtg ttatggagca    4020
```

-continued

```
attcggaggc tgaccattcc aggttcatca tatggcaggc gggtgcgtcc atgacgaagc    4080
tatcgaagcg atcacagccc atggccttga gatgacggcc tatggcatgc aacgacatgt    4140
cgttcttcat cgggccacca agcgcagcca gatcgagccg tcctcggttg tcagtggcgt    4200
caggtcgagc aagagcaacg atgcgatcag cagcaccacc gtaggcatca tggaagccag    4260
catcacggtt agccatagct tccagtgcca cccccgcgac gcgctccggg cgctctgcgc    4320
ggcgctgctc acctcggcgg ctacctcccg caactctttg gccagctcca cccatgccgc    4380
acctgtatgg agctgggctc tacgccactc agccgcatgc gactcgatgg aatgcatggt    4440
atggctactg aactgaaggg attcgtaggc aagtgtagcc atgcatgggc aacgcggttc    4500
gatcagctgc gctaactcgc tgatgcctcc ggatctcttc actcagtcga ttgcgctcat    4560
ggtctattgc ctcccggtat tcctgcaagc cgatgatctg ggcgctggcg gtgtcgatgt    4620
tcagggccac gtctgcccgg tcggtgcgga tgcccccggc ttccatctcc accacgttcg    4680
gacccaggtg aaccacgggc aggcgctcga tgccctgagc ctccagtgtt atgtggtacc    4740
tgcggcagtc gtggccagcc agctccactg accggtggcc atggtcggcc aatgcctagc    4800
gggtctgctc aagccatgcc ttgggcttga gcgcctcggc cttccgtgcc ccgcccttct    4860
ccgggatctc gccgctgtac cgcttgaacc actgagcggc gggcgcgctg atgcgctcat    4920
tgatccgctc ggagatcatc aggtggcagt gcgggttctc gccgccaccg gcatggatgg    4980
ccagcgtata cggcaggcgc tcggcaccgg tcaggtgctg ggcgaactcg gacgccagcg    5040
aatactgctg gtcgagggtc agctcgaccg gacgggaaca ttcgacatcc ttgcacagcc    5100
gaccattggc gagttcatcc aggtcggccg aatcccagta gtcggcgggc agctcgaagc    5160
actccggcac gtgcccggac tcggcgtgca agacctcatc catgctgcgg gcatacttgc    5220
cttcgcgctg gatgcagtcg gccttggccc tggccgattg gccgcccgac ctgccgccgg    5280
tttacgccgt aaggtgatca atcgccatgc tgccacgctg ttgcatttgc ttttcggctc    5340
catgcaatgg ccctcggaga gcgcaccgcc cgaagggtgg ccgttaggcc agtttctcga    5400
agacgaaacg gtacgtgcgc aatcccatcc aaagcagggc aggcattgcc gacgcagtgc    5460
atccatgcta gactaagcga aagcacatta aaaacggggc gtcaagatgg ttaacgggcg    5520
caacaaggcg gcggatcggc tggccaagcc cgaagaacaa cgagcgcgaa tcaatgccga    5580
aattcagcgc gagcgggcaa gggaacagca gcaagagcgc aagaacgaaa caaggcgcaa    5640
ggtgctggtg ggggccatga ttttggccca ggtgaacagc agcgagtggc cggaggatcg    5700
gctcatggcg gcaatggatg cgtaccttga acgcgaccac gacggcgcct tgttcggtct    5760
gacgcaacgc aagcaggatg agcccgcctg aatgatacgc agcgacaggc aatgccgggc    5820
tgcacaagcg aacccaacct taggctaggc ggacagcacg atcaagcggc tacaagagct    5880
ccagcgtatt tctgcggggt ttggtgtggc gtttagcggg ctttgcccgc ctttcccccт    5940
gccgcgcagc ggtgcggcgg tgtgcagccc agcgcagcga atagaccagc tatccggcct    6000
ctggccgggc atattgggca agggcagcag cgccccacca gggcgctgat aaccgcgcct    6060
agtggattac tcttagataa tcatggatgg atttttccaa caccccgcca gcccccgccc    6120
atgctgggtc tgcacgtttg ggggcagtgc agttactgca ggggtcagtg aacgtcattg    6180
caggcgggcg tgacagttac tgcaggggtc cgtgacagtc agtacgggag tgacgggcac    6240
tggctggcta tgtctagcta aggcaagcct ttcggctgac ggtcaaagcc atttccgatc    6300
```

-continued

```
agcgatagcc tgtacgtaaa aaccgtattg aacggaagcg gaacaatgct aatgacgagg   6360

aacggaaggc aagccgggaa agggaaactg gtacgaacca gagccaaaga aacgagaaac   6420

cccttctcta tcagatcgtt gacgagtatt accgggcatt cgctgcgctt atggcagagc   6480

agggaaagga attgccgggc tatgtgcaac gggaatttga agaatttctc caatgcgggc   6540

ggctggagca tggcttctta cgggttcgct gcgagttctg ccacgccgag cacctggtcg   6600

cttacagaaa tcaattcaaa gtatatatga gtaaacttgg tctgacaggc ccctggagac   6660

aagacaataa aagtgataga gattgaactc aataacagtg atagagatac tgagcaaggc   6720

tatgaaagcg gagaaaggaa atatggagcg tagcgaagac gttaacaaag agttcatgcg   6780

tttcaaagtt cgtatggaag gttccgttaa cggtcacgag ttcgaaatcg aaggtgaagg   6840

tgaaggtcgt ccgtacgaag gtacccagac cgctaaactg aaagttacca aaggtggtcc   6900

gctgccgttc gcttgggaca tcctgtcccc gcagttccag tacggttcca aagcttacgt   6960

taaacacccg gctgacatcc cggactacct gaaactgtcc ttcccggaag gtttcaaatg   7020

ggaacgtgtt atgaacttcg aagacggtgg tgttgttacc gttacccagg actcctccct   7080

gcaagacggt gagttcatct acaaagttaa actgcgtggt actaacttcc cgtccgacgg   7140

tccggttatg cagaaaaaaa ccatgggttg ggaagcttcc accgaacgta tgtacccgga   7200

agacggtgct ctgaaaggtg aaatcaaaat gcgtctgaaa ctgaaagacg gtggtcacta   7260

cgacgctgaa gttaaaacca cctacatggc taaaaaaccg gttcagctgc cgggtgctta   7320

caaaaaagac atcaaactgg acatcacctc ccacaacgaa gactacacca tcgttgaaca   7380

gtacgaacgt gctgaaggtc gtcactccac cggtgcttaa tacggtatcc aaggcatacc   7440

ataaaacgaa aggctcagtc gaaagactgg gcctttcgtt ttatctgttg tttgtcggtg   7500

aacgctctct actagagtca cactggctca ccttcgggtg ggcctttctg cgtttataag   7560

tggctctcac cgacgcgccg cgtgttacca ccaattaacc aattctga                7608
```

```
<210> SEQ ID NO 10
<211> LENGTH: 7762
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pDIK403 plasmid
      backbone sequence, circular

<400> SEQUENCE: 10
```

```
ttatttgccg actaccttgg tgatctcgcc tttcacgtag tggacaaatt cttccaactg     60

atctgcgcgc gaggccaagc gatcttcttc ttgtccaaga taagcctgtc tagcttcaag    120

tatgacgggc tgatactggg ccggcaggcg ctccattgcc cagtcggcag cgacatcctt    180

cggcgcgatt ttgccggtta ctgcgctgta ccaaatgcgg gacaacgtaa gcactacatt    240

tcgctcatcg ccagcccagt cgggcggcga gttccatagc gttaaggttt catttagcgc    300

ctcaaataga tcctgttcag gaaccggatc aaagagttcc tccgccgctg gacctaccaa    360

ggcaacgcta tgttctcttg cttttgtcag caagatagcc agatcaatgt cgatcgtggc    420

tggctcgaag atacctgcaa gaatgtcatt gcgctgccat tctccaaatt gcagttcgcg    480

cttagctgga taacgccacg gaatgatgtc gtcgtgcaca acaatggtga cttctacagc    540

gcggagaatc tcgctctctc caggggaagc cgaagtttcc aaaaggtcgt tgatcaaagc    600
```

-continued

```
tagccgagtc gtttcatcca gacttaaggc aaccgtaacc agccaatacc tatcactgtg     660
tggcctaagc aagccatcca atgcggagcc gtccaaatgc aaggccagcc aagtcggttc     720
gagatgggcg tcgatgacgc caacctacct tgatagttga gtcgatactt cggcgatcac     780
cgcttccctc ataacgttta actttgtttt agggcgactg ccctgctgcg taacatcgtt     840
gctgctccat aacatcaaac atcgacccac ggcgtaacgc gcttgctgct tggatgcccg     900
aggcatagac tgtaccccca aaaacacgtc ataccagcc atgaaaaccg ccactgcgcc     960
gttacaaccg atgcgttcgg taacggttcc ggcccagttg cgtgagagcc tacgctactt    1020
gacctaacgc ttacgaaccg aacaggatta tgtccaatgg gttcgcgact tactccgttt    1080
ccacggtgtg cgtcacccgg caaccttggg tagcagcgaa gtcgaggcat ttctgtcctg    1140
gctggtcatg accaaaatcc cttaacgtga gtcagcctgc cgccctgggc cgggcgatgt    1200
cgtacttgcc cgcccgcaac tcggttaccg tcccgcccag cgcgaccagc tccggcaacg    1260
cctcgcgcac ccgcttgcgg cgcttgcgca tggtcgaacc actggcctct gacggccaga    1320
aataggagca aacggtatcc atggaagccc tgccggtttc gacgggtcg atccagaacc    1380
aacgcagctg gtgcagaacg aggggcgttc agctgcaacg agcccgaacc tagtccatgc    1440
tgatgcgcac atgctggccg ccaccatga cggcctgcgc gatcaagggg ttcagggcca    1500
cgtacaggcg cccgtccgcc tcgtcgctgg cgtactccga cagcagccga aacccctgcc    1560
gcttgcggcc attctgggcg atgatggata ccttcccaag gcgcttcgatg cagtcctgta    1620
tgtgcttgag cgccccacca ctatcgacct ctgcccgat ttccttgcc agcgcccgat    1680
agccaactc gacccaatgg aattcagcgg tgacggcatc aacctcgggt taccggaacc    1740
gaccgagctg aagtcagccc taggcctgg gttccgggcc aagcactagg aactcaggcc    1800
cagccatggc caccagcccc tgcaggatgc gcagatcatc agcgcccagc ggctccgggc    1860
cgccgaactc gatccgcttg ccgtcgccgc agtcatacgc cacgtccagc ttgccgcgct    1920
tgcgctcgcc ccgcttgagg gcacggcaca ggccggggc cagacagtgc gccgggtcgt    1980
gccggacgtg gctgaggctg tgcttgttct taggcttcac cacggggcac cccctgctc    2040
ttgcgatgcc tatccagccc gggcggattg agcaccaagc agtcatgcag aatgaaaaac    2100
cgatcagcga acggtgcgcc atagttggcc ttgcccacac cgaagcggac gaagaaccgg    2160
cgccggtcgc cgtcccacac ccattcctcg gcctcggcgc tggtcatgcc cgacaggtag    2220
gacttgccagc ggatgttatc gaccagtacc gagcttgccc ggctggcctg ctgcttggtcg    2280
cctgcgccca tcatggccgc gcccttgctg gcatggtgca ggaacacgat agagcacccg    2340
gtaccggcgg agatggactc aatgcgaccg atgaccatgg aactgggcc gatggcgttt    2400
tatacatcga tgtggaaccg gagcagagtg taccgcaaca taaggcggcg gaccacggcg    2460
gagcgattga ggccgtagca aaccacaggg gacccgatgc tgggcaggct gacgataagc    2520
ggccggatca gcaggccgtc agccacggcc tgccgtccc cggcgctgag gtgccgccca    2580
aggccgtgca ggccgtgatg aatgccggtg ggcggtatc aggcggcacg gtagataacc    2640
gggccggtgg gcagttcgcc caccttccagc agatccggcc cgccttgcaat ctgtgcggcc    2700
agtagaaggg aaagcatgga tttacccggca aaaccgggcg aaacccagagc aacgacagtc    2760
ccggccaccca tgttggcgaa aacgtagtcc agcggtggcg gcgcttgctgc gaacgcctcc    2820
agaatattga taggcttatg ggtagccatt gattgcctcc tttgcagcca gttggcggtt    2880
```

-continued

```
aggcgatggc ggggtaacta aacccgaccc gagccgatcc gagttcttcc aggcactagc    2940
gatgcgactc gtattagtcg taggtcagcc agtacatgcg atgacgaatc aatttggact    3000
tcatgcgctc ggcatatcgc gcttggcgta cagcgtcagg gctggccagc aggtcgccgg    3060
tctgcttgtc cttttggtct ttcatatcag tcaccgagaa acttgccggg gccgaaaggc    3120
ttgtcttcgc ggaacaagga caaggtgcag ccgtcaaggt taaggctggc catatcagcg    3180
actgaaaagc ggccagcctc ggccttgttt gacgtatcac caaagccacc gggcaaccaa    3240
tagccattgc aactattgca aaggtagacc gacccagacg agctattttc gtataccatc    3300
aatcccacta atgtgagtga gtacccatcg tattacaggc gtgcgtaacc aagcaagacc    3360
tacatgctga aatccggccc gcccctgtcc atgcctcgct ggcggcgtgc cggtgcccgt    3420
gccagctcgg cccgcgcaag ctggacgctg ggcagaccca tgacctcgct gacggtgcgc    3480
tcgatgtaat ccgcttcgtg gccgggcttg cgctctgcca gcgcttggct ggcctcggcc    3540
atggccttgc cgatttcctc ggcactgcgg ccccggctgg ccagcttctg cgcggcgata    3600
aagtcgaccc tgctgaggtc atgcccgacg agctggacca gacggaccct atagctgagg    3660
taccgtacca gagccgtgcg aaggtggcgg atacgctgcc gatcggacag ttcgatgatg    3720
gccagcctgc gggccttctc ctgccgccgc tgggcctgcc cgatccgctg gccagcctgc    3780
tgcaccagcc ccgggccagc ggtggcggtc ttgcccttgg attcacgcag cagcacccac    3840
ggctgatacc cggcgcgggt ggtgtgcttg tcctggcggt tggtgaagcc cgccaagcgg    3900
cccatagtgg cggctgtggc gctggccggg tcggcgtcgt actcgctggc cagcgtccgg    3960
gataatgccc aacgaagttc aacgcctgcg gagtcggaca aattgaaccc tgcccgatcg    4020
ttccaagggc tggttatccc taccaggcca ggctcccggc actcggcttt aatgccatcc    4080
aggtcaaact cgctgaggtc gtcgaccagc accagaccat gccgcccctg ctcggcgggc    4140
ctgatataca cgtcattgcc ctgggcattc atccgcttga gccatgcgct gttctcgagc    4200
actacggcgg ctgaccgatc ccggttcatc atctggccgg tgggtgcgtc cctgacgccg    4260
atatcgaagc gctcacagcc catggccttg agctgtcggc ctatggcctg caaagtcctg    4320
tagtcattca taggcaaccc aagcgcagcc agatcgagcc gtcctcggtt gtcagtggcg    4380
tcaggtcgag caagagcaac gatgcgatca gcagcaccac cgtaggcatc atggaagcca    4440
gcatcacggt tagccatagc ttccagtgcc acccccgcga cgcgctccgg gcgctctgcg    4500
cggcgctgct cacctcggcg gctacctccc gcaactcttt ggccagctcc acccatgccg    4560
cccctgtctg gcgctgggct ttcagcccat ccgccgcctg cgccacgctg gcctgcttgg    4620
tatggatacc gacccgacgg gattcgtcgg acagcgtagc aatgccatgg gacgcgggtt    4680
agcacgtctc agctaactcg ttgacgactc tggacctatc aactcagtcg attgcgttcc    4740
tggcctattg aatccaggta ttcctgtacg tagccgatcc gggcgatggc ggtgacgatg    4800
ttcagggcca cgtccgcccg gtcggcgcgg atgccccggc cttccatctc cacaacgttc    4860
ggcccaaggt gatcaaaggg aaggcgatcg atgccctgcg aatcaagtgt tatgacgtac    4920
atgcggcgct cgtggccagc ccgctctaat gcccggttgg catggtcggc ccatgcctcg    4980
agggtatgca aacgcaatgc attggcattg agcgcatacg tattcagtgc aacgccattc    5040
tccggggtct tgccgttgta ccgcttgaac cactgagcgg cgggccgctc gatgccgtca    5100
ttgatccgct cggagatcat caggtcgcag tgcgggttct cgccgccacc ggcatcgatg    5160
```

-continued

```
gacagagtcc aaggcaggcg atcggcaccg gtccggtgcc gggcgaaatc ggacgccagc     5220
gacccatgcc ggtcgagggc cagcccgacc ggcccggacc attcgaaatc attgaacagc     5280
cgcccattgg cgcgctcata caggccggca gcatcccagc agtcggcggg ccgcccgacg     5340
aaccccggca tgtgcccgga ttcggcgtgc aagaccccac ccatgccgcg ggcacacttg     5400
cctccgcgcc ggatgtagtc ggccccggcc ctggccgatc ggccgcccga cctgcccgcg     5460
gtttccgccg taaggtgata aatcgccatg ctgccccgcc gttgccttcg ctttccggct     5520
aaccgaacgc gacccaggcg agcgcaacgc cagcagggtg gacgccaggc cagtcccacg     5580
aagagaaccc ggtcagtgcg cactccacta caccgcaggg taggcattgc agccgctgtg     5640
cctccatgac agccaacgac acagcacatc aacaacgggg tgtcaagatg gttaaggggа     5700
gcaacaaggc ggcggatcgg ctggccaagc tcgaagaaca acgagccgca atcaacgccg     5760
aaattcagcg ggagcgggca agggaacagc agcaagagcg caagaacgaa acaagggcca     5820
aggccgctgg cggggcccatg atttcggcca aggtgaacag cagcgagtgg ccggaggatc     5880
ggccccatggc ggccatggcc gagtacattg aacgcgacca agcccgagcc ttgtccggtc     5940
tgcccgaacg aacgaaggcc gagccgggcc gactgatcga cagcgaaagg aactgcgggg     6000
ctgcacacgc gcccccaccc ttcggctagc gggaaaggcc gctaaagcgg ctaaaagcgc     6060
tccagcgtac ttctgcgggg tttggctgtg ggttcagcgg gctttgcccg cctttccccc     6120
tgccgcgcag cggtggggcg gtgtgcagcc tagcgcagcg aatagaccag ctatccggcc     6180
tctggccggg catactgggc aagggcagca gcgccccaca agggcgctga taaccgcgcc     6240
tagccggatt attcaagcta atcccggatc gatttcatca aaccccagcc agcccccagcc     6300
aatgctgggc ttgcaggttc ggggcggtga cagtcattgc aggggctagt gaccgctatt     6360
gcaggggggc gtgacagtta ttgcaggggc tcgtgacagc tagtacggga gtgacgggca     6420
ctggctggca atgtctagca acggcaggca tttcggctga gggtaaaaga actttccgct     6480
aagcgataga ctgtatgtaa acacagtatt gcaaggacgc ggaacatgcc tcatgcggcg     6540
gccaggacgg ccagccggga tcgggatacc ggtcgctacc agagccaccg acccgagcaa     6600
aacccatatc atccgatcgc tgccgagtcc taccccgacc tagccgagct tatggcagcg     6660
caggcaaagg aattgccggg ctatgcgcaa cgggaatttg aagaatttct ccaatgcggg     6720
cggccggagc atggctttcc acggcctcgc tgcgagtctc gccacgccga gcagccggtc     6780
gctttcagca atccatctca agtcatatcg agtaaacttg gtctgacagg cccctcgaga     6840
ccagtcccta tcagtcatcg agattgcact ccctatcagt gatcgagatc ctgagcacgg     6900
atccgaacga ggcgaaagga tatccggcga gtcgccgaga cgttatcaac gagtccatgc     6960
gttccaacgc tagtatggca ggttccgtta acggccaaga gttcgaaatc gacggcgacg     7020
gtgaaggtcg taagcaagca ggtaccaaga acgccaaacc gaacgctacc aacgccggtc     7080
cgccgccgtc cgcttgggac atccgcgtcc cgcagctcca gtacgcttcc aaagcctacg     7140
ttcaaaaccc ggctgaactc aaggactacc tgccactgtc attcccggcc ggttccaact     7200
gggaacgtgc tatgaacttc gaagacggtg gtgttgttac cgttacccag gactcctccc     7260
tgcaagacgc tgccgtaatc taccaagtta aactgcgtgc tactaactac aagtccgacg     7320
gtccggttac gcagaaaaaa accatgggtt gggaagcttc caccgaacgt atgtacccgg     7380
aagacggtgc tctgaaaggc gaaatcaaaa tgcgcctgaa actgaaagac ggtggtcact     7440
```

-continued

```
aagacgatga agttaaatcc aactacatgg attaaaaacc ggttcagatg aagggtgatt    7500

aataaaacga aatcaaactg gacttcacct aactcaaaga agtctaaacc atcgttgatc    7560

agtacgaacg tgctgaaggt cgtcactcca ccggtgctta ataaggatct ccaggcatca    7620

aataaaacga aaggctcagt cgaaagactg ggcctttcgt tttatctgtt gtttgtcggt    7680

gaacgctctc tactagagtc acactggctc accttcgggt gggcctttct gcgtttatca    7740

gtcggtctca ccgagcggcc gc                                              7762


&lt;210&gt; SEQ ID NO 11
&lt;211&gt; LENGTH: 2799
&lt;212&gt; TYPE: DNA
&lt;213&gt; ORGANISM: Artificial Sequence
&lt;220&gt; FEATURE:
&lt;223&gt; OTHER INFORMATION: Chemically Synthesized, pDIK399 backbone
      sequence, circular

&lt;400&gt; SEQUENCE: 11

aggggatcca attccgtgct aggtgggctg aactacatgg ttggctggt ttacctcagcc      60

atccgattgc actcatctgc taagcgggcg gtagcgggcc agcctcgcag agcaggattc     120

ccgttgagca ccgccaggtg cgaataaggg acagtgaaga aggaacaccc gctcgcgggt     180

gggcctactt cacctatcct gcccggctga cgccgttgga tacaccaagg aaagtctaca     240

cgaacccttt ggcaaaatcc tgtatatcgt gcgaaaaagg atggatatac cgaaaaaatc     300

gctataatga ccccgaagca gggttatgca gcggaaaacg gaatggatcc ggccacgatg     360

agtccgggcg agcggatctg aagctcagca gttcaaactg ttgctagtac gtactaagct     420

atcatgtttc aagtactaag atctcatgtt tacgtacata agctcaaatg ttttacgatc     480

taaacctcca tggctaacgt actaagctct catggctaac gtactaagct ctcatgtttc     540

acgtactaag ctctcatgtt tgaacaataa aattaatata aatcagcaac ttaaatagcc     600

tctaaggttt taagttttat aagaaaaaaa agaatatata aggcttttaa agcttttaag     660

gtttaacggt tgtggacaac aagccaggga tgtaacgcac tgagaagccc ttagagcctc     720

tcaaagcatt tttgagtgct aatggaactt ttacgggatg aactggggat taccctaatc     780

agcggtggcg gacgctatcg aactagtgga tacccgggcc tgcaggaatt cgctaactag     840

cttatcgata ccgtcgacct cgagggggggg cccggtaccc aggacgcgtc gaattaattc     900

cgctagcttc acgctgccgc aagcactcag ggcgcaaggg ctgctaaagg aagcggaaca     960

cgtagaaagc cagtccgcag aaacggtgct gaccccggat gaatgtcagc tactgggcta    1020

tctggacaag ggaaaacgca agcgcaaaga gaaagcaggt agcttgcagt gggatttatca    1080

aactggattc tacctagctc actttaactt taccaatgca gtttcaaatc aatctaaagt    1140

atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgaggc acctatctca    1200

gcgatctgtc tatttcgttc atccatagtt gcctgactcc ccgtcgtgta gataactacg    1260

atacgggagg gcttaccatc tggccccagt gctgcaatga taccgcggga cccacgctca    1320

ccggctccag atttatcagc aataaaccag ccagccggaa gggccgagcg cagaagtggt    1380

cctgcaactt tatccgcctc catccagtct attaattgtt gccgggaagc tagagtaagt    1440

agttcgccag ttaatagttt gcgcaacgtt gttgccattg ctacaggcat cgtggtgtca    1500

cgctcgtcgt ttggtatggc ttcattcagc tccggttccc aacgatcaag gcgagttaca    1560

tgatccccca tgttgtgcaa aaaagcggtt agctccttcg gtcctccgat cgttgtcaga    1620
```

-continued

```
agtaagttgg aagcagtgtc atccccactg gttccggacg aactgcatcc ttctcctact   1680

gtcacgaccc aagtaagctg atttactgtg actggcgagc actcaaaccc gtccccatgc   1740

gaacagtgta tgcggcgacc gagttgctct tgcccggcgt caatacggga taataccgcg   1800

ccacatagca gaactttaaa agtgctcatc attggaaaac gttcttcggg gcgaaaactc   1860

tcaaggatct taccgctgtt gagatccagt tcgatgtaac ccactcgtgc acccaactga   1920

tcttcagcat cttttacttt caccagcgtt tctgggtgag caaaaacagg aaggcaaaat   1980

gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa tactcatact cttccttttt   2040

caatattatt gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt   2100

atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgcc   2160

tgatgagacg atgaaagtga aacgtgattt catgcgtcat tttgaacatt ttgtaaatct   2220

tatttaatca tgtgtgcggc aattcacatt tcatttatga atgtttcttt aacatcgcgg   2280

caaccaacga aacggcaggc taggatctta gatcctagcg aacgacgagc aatcctagct   2340

gagcaaaggc gaacgagatg tcacaggtgc agtccctatc ctggcggacc tggcaggtgc   2400

tgtcaacggc cataagtttt ccgtgcgtgg cgagggtgaa ggtgacgcaa ctaatggtaa   2460

actgacgctg aagttcatct gtactactgg taaactgccg gttccttggc cgactctggt   2520

aacgacgctg acttatggtg ttcagtgctt tgctcgttat ccggaccata tgaagcagca   2580

tgacttcttc aagtccgcca tgcccgaagg ctatgtgcag gaacgcacga tttccttcaa   2640

ggacgacggc aactacaaca agcgcgcgga agtgaaattc gaaggcgata ccctggtaaa   2700

ccgcattgag ctgaaaggca ttgacttcaa agacgacggc aatatcctgg gacacaagct   2760

ggaatacaac tttaacagcc acaatgttta catcaccgcc gataaacaaa aaaatggcat   2820

taaagcgaat tttaaaattc gccacaacgt ggaggatggc agcgtgcagc tggctgatca   2880

ctaccagcaa aacactccaa tcggtgatgg tcctgttctg ctgccagaca atcactatct   2940

gagcacgcaa agcgttctgt ctaaagatcc gaacgagaaa cgcgatcata tggttctgct   3000

ggagttcgta accgcagcgg gcatcacgca tggtatggat gaactgtaca aatgaccagg   3060

aatcaaataa aacgaaaggc tcagtcgaaa gactgggcct ttcgttttat ctgttgtttg   3120

tcggtgaacg ctctctacta gagtcacact ggctcacctt cgggtgggcc tttctgcgtt   3180

tataagtccg tctcaagca                                                3199
```

<210> SEQ ID NO 12
<211> LENGTH: 109
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, CP25 lacO sequence

<400> SEQUENCE: 12

```
ctttggcagt ttattcttga catgtagtga gggggctggt ataatcacat agtactgttg     60

gaattgtgag cggataacaa ttcccatacg aaacagagga gatattaac                109
```

<210> SEQ ID NO 13
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized

continued

<400> SEQUENCE: 13

tacccaatgc gacggaaccc 20

<210> SEQ ID NO 14
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized

<400> SEQUENCE: 14

atcaggagca atcacgactg 20

<210> SEQ ID NO 15
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized

<400> SEQUENCE: 15

aatgctgttt tcccggggat 20

<210> SEQ ID NO 16
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized

<400> SEQUENCE: 16

ctggaaattg ctgctgccag 20

<210> SEQ ID NO 17
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized

<400> SEQUENCE: 17

tcgcttgttc acagcgatag a 21

<210> SEQ ID NO 18
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized

<400> SEQUENCE: 18

gcttttgcca ttctcaccgg 20

<210> SEQ ID NO 19
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized

<400> SEQUENCE: 19

attggaacag atacctttgc 20

<210> SEQ ID NO 20
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized

<400> SEQUENCE: 20

gtcgtggtat tgtcaggagc a                                              21

<210> SEQ ID NO 21
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized

<400> SEQUENCE: 21

tgttcaggtc gtcaataccc attcc                                          25

<210> SEQ ID NO 22
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized

<400> SEQUENCE: 22

aacccgaagc ttccattgcc gatcg                                          25

<210> SEQ ID NO 23
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized

<400> SEQUENCE: 23

gagagaccct ttataccccc atccc                                          25

<210> SEQ ID NO 24
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized

<400> SEQUENCE: 24

ttcaaggttc aattcgactc tttca                                          25

<210> SEQ ID NO 25
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized

<400> SEQUENCE: 25

accctttgcc agttcgttcc aagctc                                         26

<210> SEQ ID NO 26
<211> LENGTH: 26

continued

<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized

<400> SEQUENCE: 26

tcaaaattgc tttgagaggc tctaag 26

<210> SEQ ID NO 27
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized

<400> SEQUENCE: 27

tcatgtcatc caggtcaaac tc 22

<210> SEQ ID NO 28
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesised

<400> SEQUENCE: 28

tgaagaaatc cagaggcatc aa 22

<210> SEQ ID NO 29
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesised

<400> SEQUENCE: 29

cttagagata ggagagtg 18

<210> SEQ ID NO 30
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesised

<400> SEQUENCE: 30

taatgatggc aactaatgac aa 22

<210> SEQ ID NO 31
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized

<400> SEQUENCE: 31

cccaatcctc gtggaattag acg 23

<210> SEQ ID NO 32
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized

-continued

<400> SEQUENCE: 32

gacttagacg aatcagactc 20

<210> SEQ ID NO 33
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized

<400> SEQUENCE: 33

ttccatttca tacgagactc 20

<210> SEQ ID NO 34
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized

<400> SEQUENCE: 34

acaccatgtc catgcgaaaa 20

<210> SEQ ID NO 35
<211> LENGTH: 49
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized

<400> SEQUENCE: 35

atgtcttttt caatgggtac gtatgtatct ctaaagtaat caacataat 49

<210> SEQ ID NO 36
<211> LENGTH: 2481
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: chemically synthesized, pKD3

<400> SEQUENCE: 36

agcgagctcg atatcaaatt acgccccgcc ctgccactca tcgcagtact gttgtaattc 60

attaagcatt ctgccgacat ggaagccatc acaaacggca tgatgaacct gaatcgccag 120

cggcatcagc accttgtcgc cttgcgtata atatttgccc atggtgaaaa cgggggcgaa 180

gaagttgtcc atattggcca cgtttaaatc aaaactggtg aaactcaccc agggattggc 240

tgagacgaga aacatattct caataaaccc tttagggaaa taggccaggt tttcaccgta 300

acacgccaca tcttgcgaat atatgtgtag aaactgccgg aaatcgtcgt ggtattcact 360

ccagagcgat gaaaacgttt cagtttgctc atggaaaacg gtgtaacaag ggtgaacact 420

atcccatatc accagctcac cgtctttcat tgccatacga aattccggat gagcattcat 480

caggcgggca agaatgtgaa taaaggccgg ataaaacttg tgcttatttt tctttacggt 540

ctttaaaaag gccgtaatat ccagctgaac ggtctggtta taggtacatt gagcaactga 600

ctgaaatgcc tcaaaatgtt ctttacgatg ccattgggat atatcaacgg tggtatatcc 660

agtgattttt ttctccattt tagcttcctt agctcctgaa aatctcgata actcaaaaaa 720

tacgcccggt agtgatctta tttcattatg gtgaaagttg gaacctctta cgtgccgatc 780

continued

```
aacctaacct aggccaagcc aacccgacca gaatcaagca atatgaagct aaccccagtg     840

gtggcgaacc aagacaggcc tataaagata aaaggcgttc aaccccggcg gatcccatgt     900

gcgctctcct gttcctgcct ttcggtttac cggtgtcatt ccgctgttat ggccgcgttt     960

gtctcattcc acgcctgaca ctcagttccg ggtaggcagt tcgctccaag ctggactgta    1020

tgcacgaacc ccccgttcag tccgaccgct gcgccttatc cggtaactat cgtcttgagt    1080

ccaacccgga aagacatgca aaagcaccac tggcagcagc cactggtaat tgatttagag    1140

gagttagtct tgaagtcatg cgccggttaa ggctaaactg aaaggacaag ttttggtgac    1200

tgcgctcctc caagccagtt acctcggttc aaagagttgg tagctcagag aaccttcgaa    1260

aaaccgccct gcaaggcggt ttttcgtttt tcagagcaag agattacgcg cagaccaaaa    1320

cgatctcaac ccctgattct gtggataacc gtagtcggcc agacggaaag tgaaacgtga    1380

tttcatgcgt cattttgaac atttggtaaa tcttatttaa taatggtgtc ggccattcac    1440

atttaattta tgtttgtttt attaacatcg aggcaatcca agttacggcc ggttcggatc    1500

ttagctacta gagtaagtgg agttatacta gatgcgtata ggcgaagagc tgttcaatgg    1560

tgtcgtcccc attcctggtg aactggatgg tgatgtcaac ggtcataagt tttccgtgcg    1620

tggcgagggc gaaggtgacg caactaatgg taaactgacg ctgaagttca tctgtactac    1680

tggtaaactg ccggttcctt ggccgactct ggtaacgacg ctgacttatg gtgttcagtg    1740

ctttgctcgt tatccggacc atatgaagca gcatgacttc ttcaagtccg ccatgccgga    1800

aggctatgtg caggaacgca cgatttcctt taaggatgac ggcacgtaca aaacgcgtgc    1860

ggaagtgaaa tttgaaggcg ataccctggt aaaccgcatt gagctgaaag gcattgactt    1920

taaagaggac ggcaatatcc tgggccataa gctggaatac aattttaaca gccacaatgt    1980

ttacatcacc gccgataaac aaaaaaatgg cattaaagcg aattttaaaa ttcgccacaa    2040

cgtggaggat ggcagcgtgc agctggctga tcactaccag caaaacactc caatcggtga    2100

tggtcctgtt ctgctgccag acaatcacta tctgagcacg caaagcgttc tgtctaaaga    2160

tccgaacgag aaacgcgatc atatggttct gctggagttc gtaaccgcag cgggcatcac    2220

gcatggtatg gatgaactgt acaaatgacc aggcattaaa taaaacgaaa ggctcagtcg    2280

aaagactggg cctttcgttt tatctgttgt ttgtcggtga acgctctcta ctagagtcac    2340

actggctcac cttcgggtgg gcctttctgc gtttatacgt ctctgaccag accaataaaa    2400

aacgccccgc ggcaaccgag cgttctgaac aaatccagat ggagttctga ggtcattact    2460

ggatctatca acaggagtcc a                                              2481
```

<210> SEQ ID NO 37
<211> LENGTH: 52
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically synthesized, T7 + RBS promoter from pET41b

<400> SEQUENCE: 37

```
aaggtaatac gactcactat agggcaacga aaccgaggcg atataccata tg              52
```

<210> SEQ ID NO 38
<211> LENGTH: 64
<212> TYPE: DNA

-continued

```
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, lac lacO + RBS from
      pBTK103

<400> SEQUENCE: 38

aacgtttaca ctttatgctt ccggctcgta tgttgtgtgg aattgtgagc ggataacaat      60

ttcaaggagg agaaattaac tatg                                             84


<210> SEQ ID NO 39
<211> LENGTH: 92
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, cp25 + RBS from pBTK107

<400> SEQUENCE: 39

aacgctttgg cagtttattc ttgacatgta gtgagggggc tggtataatc acatagtact      60

gttatacaga aacagaggag atattaacta tg                                    92


<210> SEQ ID NO 40
<211> LENGTH: 93
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, CP6 + RBS from pBTK110

<400> SEQUENCE: 40

aacgcatgtg ggagtttatt cttgacaccg atattttcgg atgctataat aactgagtac      60

tgttatacag aaacagagga gatattaact atg                                   93


<210> SEQ ID NO 41
<211> LENGTH: 93
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, cp25 + RBS from
      pBTK112

<400> SEQUENCE: 41

aacgcatata caagtttatt cttgacacta gtcggccaaa atgatataat acctgagtac      60

tgttatacag aaacagagga gatattacat atg                                   93


<210> SEQ ID NO 42
<211> LENGTH: 93
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, cp32 + RBS from pBTK113

<400> SEQUENCE: 42

aacgcatacg ggagtttatt cttgacatat tgccggtgtg ttggtataat aacttagtac      60

tgttatacag aaacagagga gatattaact atg                                   93


<210> SEQ ID NO 43
<211> LENGTH: 103
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, PA1 + RBS from pBTK119
```

-continued

<400> SEQUENCE: 43

aacgttatca aaacgagtaa tgaccttaag tataacataa aggataactt aagccatagc 60

gagggacagg gcgaatacag aaacagagga gatattacat atg 103

<210> SEQ ID NO 44
<211> LENGTH: 104
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, P42 + RBS from pBTK120

<400> SEQUENCE: 44

aacgcacgaa aaacaggtat tgacaacatg aagtaacatg cagtaagata caaatcgcta 60

ggtaacacta gcagcataca gaaacagagg agatattaca tatg 104

<210> SEQ ID NO 45
<211> LENGTH: 103
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, P43 + RBS from pBTK121

<400> SEQUENCE: 45

aacgggtgaa acaaaacggt tgacaacatg aagtaaacac ggtacgatgt accacatgaa 60

acgacagtga gtcaatacag aaacagagga gatattacat atg 103

<210> SEQ ID NO 46
<211> LENGTH: 102
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, CP25 + RBS reverse
      promoter from pBTK103

<400> SEQUENCE: 46

gagtacctct cctgtatctt gtgtacttgc gtaccgtaat atctcctctg tttctgtata 60

aacgtactaa gtgattatac aagcccactc actacatgtc aagtataacc tgccaaagcc 120

ga 102

<210> SEQ ID NO 47
<211> LENGTH: 4115
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, dCas9 from pBTK209

<400> SEQUENCE: 47

tatgatggat aagaaatact caataggctt agctatcggc acaaatagcg tcggatgggc 60

ggtgatcact gatgaatata aggttccgtc taaaaagttc aaggttctgg gaaatacaga 120

ccgccacagt atcaaaaaaa atcttatagg ggctctttta tttgacagtg gagagacagc 180

ggaagcgact cgtctcaaac ggacagctcg tagaaggtat acacgtcgga agaatcgtat 240

ttgttatcta caggagattt tttcaaatga gatggcgaaa gtagatgata gtttctttca 300

tcgacttgaa gagtcttttt tggtggaaga agacaagaag catgaacgtc atcctatttt 360

tggaaatata gtagatgaag ttgcttatca tgagaaatat ccaactatct atcatctgcg 420

aaaaaaattg gtagattcta ctgataaagc ggatttgcgc ttaatctatt tggccttagc 480

-continued

```
gcatatgatt aagtttcgtg gtcatttttt gattgaggga gatttaaatc ctgataatag     540

tgatgtggac aaactattta tccagttggt acaaacctac aatcaattat ttgaagaaaa     600

ccctattaac gcaagtggag tagatgctaa agcgattctt tctgcacgat tgagtaaatc     660

aagacgatta gaaaatctca ttgctcagct ccccggtgag aagaaaaatg gcttatttgg     720

gaatctcatt gctttgtcat tgggtttgac ccctaatttt aaatcaaatt ttgatttggc     780

agaagatgct aaattacagc tttcaaaaga tacttacgat gatgatttag ataatttatt     840

ggcgcaaatt ggagatcaat atgctgattt gtttttggca gctaagaatt tatcagatgc     900

tattttactt tcagatatcc taagagtaaa tactgaaata actaaggctc ccctatcagc     960

ttcaatgatt aaacgctacg atgaacatca tcaagacttg actcttttaa aagctttagt    1020

tcgacaacaa cttccagaaa agtataaaga aatctttttt gatcaatcaa aaaacggata    1080

tgcaggttat attgatgggg gagctagcca agaagaattt tataaattta tcaaaccaat    1140

tttagaaaaa atggatggta ctgaggaatt attggtgaaa ctaaatcgtg aagatttgct    1200

gcgcaagcaa cggacctttg acaacggctc tattccccat caaattcact tgggtgagct    1260

gcatgctatt ttgagaagac aagaagactt ttatccattt ttaaaagaca atcgtgagaa    1320

gattgaaaaa atcttgactt ttcgaattcc ttattatgtt ggtccattgg cgcgtggcaa    1380

tagtcgtttt gcatggatga ctcggaagtc tgaagaaaca attaccccat ggaattttga    1440

agaagttgtc gataaaggtg cttcagctca atcatttatt gaacgcatga caaactttga    1500

taaaaatctt ccaaatgaaa aagtactacc aaaacatagt ttgctttatg agtattttac    1560

ggtttataac gaattgacaa aggtcaaata tgttactgaa ggaatgcgaa aaccagcatt    1620

tctttcaggt gaacagaaga aagccattgt tgatttactc ttcaaaacaa atcgaaaagt    1680

aaccgttaag caattaaaag aagattattt caaaaaaata gaatgttttg atagtgttga    1740

aatttcagga gttgaagata gatttaatgc ttcattaggt acctaccatg atttgctaaa    1800

aattattaaa gataaagatt ttttggataa tgaagaaaat gaagatatct tagaggatat    1860

tgttttaaca ttgaccttat ttgaagatag ggagatgatt gaggaaagac ttaaaacata    1920

tgctcacctc tttgatgata aggtgatgaa acagcttaaa cgtcgccgtt atactggttg    1980

gggacgtttg tctcgaaaat tgattaatgg tattagggat aagcaatctg gcaaaacaat    2040

attagatttt ttgaaatcag atggttttgc caatcgcaat tttatgcagc tgatccatga    2100

tgatagtttg acatttaaag aagacattca aaaagcacaa gtgtctggac aaggcgatag    2160

tttacatgaa catattgcaa atttagctgg tagccctgct attaaaaaag gtattttaca    2220

gactgtaaaa gttgttgatg aattggtcaa agtaatgggg cggcataagc cagaaaatat    2280

cgttattgaa atggcacgtg aaaatcagac aactcaaaag ggccagaaaa attcgcgaga    2340

gcgtatgaaa cgaatcgaag aaggtatcaa agaattagga agtcagattc ttaaagagca    2400

tcctgttgaa aatactcaat tgcaaaatga aaagctctat ctctattatc tccaaaatgg    2460

aagagacatg tatgtggacc aagaattaga tattaatcgt ttaagtgatt atgatgtcga    2520

tcacattgtt ccacaaagtt tccttaaaga cgattcaata gacaataagg tcttaacgcg    2580

ttctgataaa aatcgtggta aatcggataa cgttccaagt gaagaagtag tcaaaaagat    2640

gaaaaactat tggagacaac ttctaaacgc caagttaatc actcaacgta agtttgataa    2700

tttaacgaaa gctgaacgtg gaggtttgag tgaacttgat aaagctggtt ttatcaaacg    2760
```

-continued

```
aaaatggtta gaaactagcc aaatcaatca gcatgtggca aaatttttgg atacgtcgat   2820

gaaaaataaa tacgatgaaa atgaaaaaca tatttgtagg gttaaagtga ttacctaaac   2880

atctaaatta gtttctgact tccgaaaaga tttccaattc tataaagtac gtgagattaa   2940

caattaccat catgcccatg atgcgtatct aaatgccgtc gttggaactg ctttgattaa   3000

gaaaaatcca aaacttgaat cggagtttgt ctatggtgat tataaagttt atgatgttcg   3060

taaaatgatt gctaagtctg agcaagaaat aggcaaagca accgcaaaat atttctttta   3120

atctaatatc atgtaattct taaaaaatga aattacaatt gaaaatggtg agattcgaaa   3180

aagcaatcta atcgaaacta atggggaaac tggagaaatt gtctgggata aagggcgaga   3240

ttttgccaca gtgcgcaaag tattgtccat gcgccaagtc aatattgtca agaaaacaga   3300

agtacagaca gggggattct ccaaggagtc aatttaccca aaaagaaatt cggacaagct   3360

tattgctcgt aaaaaagact gggatccaaa aaaatatggt ggttttgata gtccaacggt   3420

agcttattca gtcctagtgg ttgctaaggt ggaaaaaggg aaatcgaaga agttaaaatc   3480

agttaaagag ttactaggga tcacaattat ggaaagaagt tcctttgaaa aaaatccgat   3540

tgacttttta gaagctaaag gatataagga agttaaaaaa gacttaatca ttaaactacc   3600

taaatatagt cttttgagtt tagaaaacgg tcgtaaacgg atgctggcta gtgccggaga   3660

attacaaaaa ggaaatgagc tggctctgcc aagcaaatat gtgaattttt tatatttagc   3720

tagtcattat gaaaagttga agggtagtcc agaagataac gaacaaaaac aattgtttgt   3780

ggagcagcat aagcattatt tagatgagat tattgagcaa atcagtgaat tttctaagcg   3840

tgttatttta gcagatgcca atttagataa agttcttagt gcatataaca aacatagaga   3900

caaaccaata cgtgaacaag cagaaaatat tattcattta tttacgttga cgaatcttgg   3960

agctcccgct gcttttaaat atttcgatac aacaattgat cgtaaacgat atacgtctac   4020

aaagaagttt ttagatgcca ctcttatcca tcaatccatc actggtcttt atgaaacacg   4080

cattgatttg agtcagctag gaggtgactg aatcc                               4115
```

<210> SEQ ID NO 18
<211> LENGTH: 91
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, rpoC terminator from pBTK300

<400> SEQUENCE: 18

```
atccgtaatc gttcatacgc aactaaagta aaaaccagcc taggctggtt ttttcatggg     60

gggagtttag ggaaagagca tttgtcagca g                                    91
```

<210> SEQ ID NO 19
<211> LENGTH: 137
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, BBa_B0015 terminator from pBTK301

<400> SEQUENCE: 19

```
tacaccaggc atcaaataaa acgaaaggct cagtcgaaag actgggcctt tcgttttatc     60

tgttgtttgt cggtgaacgc tctctactag agtcacactg gctcaccttc gggtgggcct    120
```

continued

ttctgagttc atccgga 137

<210> SEQ ID NO 50
<211> LENGTH: 62
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, T7 terminator from
      pBTK305

<400> SEQUENCE: 50

atccaaacta gcataacccc ttggggcctc taaacgggtc ttgaggggtt ttttgctggc 60

tg 62

<210> SEQ ID NO 51
<211> LENGTH: 1710
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pBTK102

<400> SEQUENCE: 51

tcaccgtctt tcattgccat acgaaattcc ggatgagcat tcatcaggcg ggcaagaatg 60

tgaataaagg ccggataaaa cttgtgctta tttttcttta cggtctttaa aaaggccgta 120

atatccagct gaacggtctg gttataggta cattgagcaa ctgactgaaa tgcctcaaaa 180

tgttctttac gatgccattg ggatatatca acggtggtat atccagtgat ttttttctcc 240

attttagctt ccttagctcc tgaaaatctc gataactcaa aaaatacgcc cggtagtgat 300

cttatttcat tatggtgaaa gttggaacct cttacgtgcc gatcaatctc tgaccaaaat 360

cccttaacgt gagttttcgt tccactgagc gtcagacccc gtagaaaaga tcaaaggatc 420

ttcttgagat cctttttttc tgcgcgtaat ctgctgcttg caaacaaaaa aaccaccgct 480

accagcggtg gtttgtttgc cggatcaaga gctaccaact ctttttccga aggtaactgg 540

cttcagcaga gcgcagatac caaatactgt tcttctagtg tagccgtagt taggccacca 600

cttcaagaac tctgtagcac cgcctacata cctcgctctg ctaatcctgt taccagtggc 660

tgctgccagt ggcgataagt cgtgtcttac cgggttggac tcaagacgat agttaccgga 720

taaggcgcag cggtcgggct gaacgggggg ttcgtgcaca cagcccagct tggagcgaac 780

gacctacacc gaactgagat acctacagcg tgagctatga gaaagcgcca cgcttcccga 840

agggagaaag gcggacaggt atccggtaag cggcagggtc ggaacaggag agcgcacgag 900

ggagcttcca gggggaaacg cctggtatct ttatagtcct gtcgggtttc gccacctctg 960

acttgagcgt cgatttttgt gatgctcgtc aggggggcgg agcctatgga aaaacgccag 1020

caacgcggcc tttttacggt tcctggcctt ttgctggcct tttgctcaca tgttctttcc 1080

tgcgttatcc cctgattctg tggataaccg tagtcggtcc caaacgtaat acgactcact 1140

ataggatcca gatccagcgg agatattcca tatgcgagcc caggccaatc aattacgacc 1200

ggcggcaacc gagcgttctg aaccaaatcc agatggagtt ctgaggtcat tactggatct 1260

tacacaggag taccagagag atcgatatca aattacgccc cgccctgcca ctcatcgcag 1320

tactgttgta attcattaag cattctgccg acatggaagc catcacaaac ggcatgatga 1380

acctgaatcg ccagcggcat cagcaccttg tcgccttgcg tataatattt gcccatggtg 1440

continued

---

aaaacggggg cgaagaagtt gtccatattg gccacgttta aatcaaaact ggtgaaactc 1500

acccagggat tggctgagac gaaaaacata ttctcaataa accctttagg gaaataggcc 1560

aggttttcac cgtaacacgc cacatcttgc gaatatatgt gtagaaactg ccggaaatcg 1620

tcgtggtatt cactccagag cgatgaaaac gtttcagttt gctcatggaa aacggtgtaa 1680

caagggtgaa cactatccca tatcaccagc 1710

<210> SEQ ID NO 52
<211> LENGTH: 1742
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pDIK103

<400> SEQUENCE: 52

tttagcttcc ttagctcctg aaaatctcga taactcaaaa aatacgcccg gtagtgatct 60

tatttcatta tggtgaaagt tggaacctct tacgtgcccg atcaatcatg accaaaatcc 120

cttaacgtga gttttcgttc cactgagcgt cagaccccgt agaaaagatc aaaggatctt 180

cttgagatcc tttttttctg cgcgtaatct gctgcttgca aacaaaaaaa ccaccgctac 240

cagcggtggt ttgtttgccg gatcaagagc taccaactct ttttccgaag gtaactggct 300

tcagcagagc gcagatacca aatactgttc ttctagtgta gccgtagtta ggccaccact 360

tcaagaactc tgtagcaccg cctacatacc tcgctctgct aatcctgtta ccagtggctg 420

ctgccagtgg cgataagtcg tgtcttaccg ggttggactc aagacgatag ttaccggata 480

aggcgcagcg gtcgggctga acggggggtt cgtgcacaca gcccagcttg gagcgaacga 540

cctacaccga actgagatac ctacagcgtg agctatgaga aagcgccacg cttcccgaag 600

ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg aacaggagag cgcacgaggg 660

agcttccagg gggaaacgcc tggtatcttt atagtcctgt cgggtttcgc cacctctgac 720

ttgagcgtcg atttttgtga tgctcgtcag gggggcggag cctatggaaa aacgccagca 780

acgcggcctt tttacggttc ctggcctttt gctggccttt tgctcacatg ttctttcctg 840

agttatcccc tgattctgtg gataaccgta gtcggtatca aacgcttacc atttacgatt 900

ccggctcgta tgttgtgtgg aattgtgagc ggataacaat tttaaagagg agaaattaac 960

tatgtgagac cagaccaata aaaaacgccc ggcggcaacc gagcgttctg aacaaatcca 1020

gatggagttc tgaggtcatt actggatcta tcaacaggag tccaagcgag ctcgatatca 1080

aattacgccc cgccctgcca ctcatcgcag tactgttgta attcattaag cattctgccg 1140

acatggaagc catcacaaac ggcatgatga acctgaatcg ccagcggcat cagcaccttg 1200

tcgccttgcg tataatattt gcccatggtg aaaacggggg cgaagaagtt gtccatattg 1260

gccacgttta aatcaaaact ggtgaaactc acccagggat tggctgagac gaaaaacata 1320

ttctcaataa accctttagg gaaataggcc aggttttcac cgtaacacgc cacatcttgc 1380

gaatatatgt gtagaaactg ccggaaatcg tcgtggtatt cactccagag cgatgaaaac 1440

gtttcagttt gctcatggaa aacggtgtaa caagggtgaa cactatccca tatcaccagc 1500

tcaccgtctt tcattgccat acgaaattcc ggatgagcat tcatcaggcg ggcaagaatg 1560

tgaataaagg ccggataaaa cttgtgctta tttttcttta cggtctttaa aaaggccgta 1620

atatccagct gaacggtctg gttataggta cattgagcaa ctgactgaaa tgcctcaaaa 1680

-continued

tgttctttac gatgccattg ggatatatca acggtggtat atccagtgat ttttttctcc 1740

at 1742

<210> SEQ ID NO 53
<211> LENGTH: 1750
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pBTK107

<400> SEQUENCE: 53

tactgacccc aatccattca agtgagtttt agttccactg agcgtcagac cccgtagaaa 60

agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa 120

aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca actctttttc 180

cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgttcttcta gtgtagccgt 240

agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaatcc 300

tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac 360

gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca 420

gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta tgagaaagcg 480

ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag 540

gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt 600

ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcagggggg cggagcctat 660

ggaaaaacgc cagcaacgcg gcctttttac ggttcctggc cttttgctgg ccttttgctc 720

acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtagtagg tatcaaaagc 780

tttggcagtt tattcttgac atgtagtgag ggggctggta taatcacata gtactgttat 840

acagaaacag aggagatatt acatatgtga gaccagacca ataaaaaacg cccggcggca 900

accgagcgtt ctgaacaaat ccagatggag ttctgaggtc attactggat ctatcaacag 960

gagtccaagc gagctcgata tcaaattacg ccccgccctg ccactcatcg cagtactgtt 1020

gtaattcatt aagcattctg ccgacatgga agccatcaca aacggcatga tgaacctgaa 1080

tcgccagcgg catcagcacc ttgtcgcctt gcgtataata tttgcccatg gtgaaaacgg 1140

gggcgaagaa gttgtccata ttggccacgt ttaaatcaaa actggtgaaa ctcacccagg 1200

gattggctga gacgaaaaac atattctcaa taaacccttt agggaaatag gccaggtttt 1260

caccgtaaca cgccacatct tgcgaatata tgtgtagaaa ctgccggaaa tcgtcgtggt 1320

attcactcca gagcgatgaa aacgtttcag tttgctcatg gaaaacggtg taacaagggt 1380

gaacactatc ccatatcacc agctcaccgt ctttcattgc catacggaat tccggatgag 1440

cattcatcag gcgggcaaga atgtgaataa aggccggata aaacttgtgc ttatttttct 1500

ttacggtctt taaaaaggcc gtaatatcca gctgaacggt ctggttatag gtacattgag 1560

caactgactg aaatgcctca aaatgttctt tacgatgcca ttgggatata tcaacggtgg 1620

tatatccagt gatttttttc tccattttag cttccttagc tcctgaaaat ctcgataact 1680

caaaaaatac gcccggtagt gatcttattt cattatggtg aaagttggaa cctcttacgt 1740

gcccgatcaa 1750

<210> SEQ ID NO 54

continued

<211> LENGTH: 1751
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesised, pDTK110

<400> SEQUENCE: 54

ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca gcttggagcg 60

aacgacctac accgaactga gatacctaca gcgtgagcta tgagaaagcg ccacgcttcc 120

cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag gagagcgcac 180

gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt ttcgccacct 240

ctgacttgag cgtcgatttt tgtgatgctc gtcagggggg cggagcctat ggaaaaacgc 300

cagcaacgcg gcctttttac ggttcctggc cttttgctgg ccttttgctc acatgttctt 360

tcctgcgtta tcccctgatt ctgtggataa ccgtagtcgg tctcaaagcg atgtgggagt 420

ttattcttga cacagatatt tccggatgat ataataactg agtactgtta tacagaaaca 480

gaggagatat taccatgtgg agcccagccc aatcaaaaac gcccggaggc aaacgagagt 540

tatgaaacca taccgatgga gttccgaggc catcactgga tatccaaacc ggacgccaag 600

cgagctcgat atcaaattac gccccgcccc gccactcatc gcagtactgt tgtaattcat 660

taagcattct gccgacatgg aagccatcac aaacggcatg atgaacctga atcgccagcg 720

gcatcagcac cttgtcgcct tgcgtataat atttgcccat ggtgaaaacg ggggcgaaga 780

agttgtccat attggccacg tttaaatcaa aactggtgaa actcacccag ggattggctg 840

agacgaaaaa catattctca ataaaccctt tagggaaata ggccaggttt tcaccgtaac 900

acgccacatc ttgcgaatat atgtgtagaa actgccggaa atcgtcgtgg tattcactcc 960

agagcgatga aaacgtttca gtttgctcat ggaaaacggt gtaacaaggg tgaacactat 1020

cccatatcac cagctcaccg tctttcattg ccatacgaaa ttccggatga gcattcatca 1080

ggcgggcaag aatgtgaata aaggccggat aaaacttgtg cttatttttc tttacggtct 1140

ttaaaaaggc cgtaatatcc agctgaacgg tctggttata ggtacattga gcaactgact 1200

gaaatgcctc aaaatgttct ttacgatgcc attgggatat atcaacggtg gtatatccag 1260

tgattttttt ctccatttta gcttccttag ctcctgaaaa tctcgataac tcaaaaaata 1320

cgcccggtag tgatcttatt tcattatggt gaaagttgga acctcttacg tgccgatcaa 1380

atcatgacca aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa 1440

aagatcaaag gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca 1500

aaaaaaccac cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt 1560

ccgaaggtaa ctggcttcag cagagcgcag ataccaaata ctgttcttct agtgtagccg 1620

tagttaggcc accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaatc 1680

ctgttaccag tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga 1740

cgatagttac c 1751

<210> SEQ ID NO 55
<211> LENGTH: 1751
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesised, pDTK110

-continued

<400> SEQUENCE: 55

```
tcttgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac cccgtagaaa      60
agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa     120
aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca actctttttc     180
cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgttcttcta gtgtagccgt     240
agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaatcc     300
tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac     360
gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca     420
gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta tgagaaagcg     480
ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag     540
gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt     600
ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcagggggg cggagcctat     660
ggaaaaacgc cagcaacgcg gcctttttac ggttcctggc cttttgctgg ccttttgctc     720
acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtagtagg tatcaaaagc     780
atatacaagt ttattcttga cactagtggc ccaaaatgat ataatacctg agtactgtta     840
tacagaaaca gaggagatat tacatatgtg agaacagaac aataaaaaac gcccggcggc     900
aaccgagcgt tctgaacaaa tccagatgga gttctgaggt cattactgga tctatcaaca     960
ggagtccaag cgagctcgat atcaaattac gccccgccct gccactcatc gcagtactgt    1020
tgtaattcat taagcattct gccgacatgg aagccatcac aaacggcatg atgaacctga    1080
atcgccagcg gcatcagcac cttgtcgcct tgcgtataat atttgcccat ggtgaaaacg    1140
ggggcgaaga agttgtccat attggccacg tttaaatcaa aactggtgaa actcacccag    1200
ggattggctg agacgaaaaa catattctca ataaaccctt tagggaaata ggccaggttt    1260
tcaccgtaac acgccacatc ttgcgaatat atgtgtagaa actgccggaa atcgtcgtgg    1320
tattcactcc agagcgatga aaacgtttca gtttgctcat ggaaaacggt gtaacaaggg    1380
tgaacactat cccatatcac cagctcaccg tctttcattg ccatacgaat ttccggatga    1440
gcattcatca ggcgggcaag aatgtgaata aaggccggat aaaacttgtg cttatttttc    1500
tttacggtct ttaaaaaggc cgtaatatcc agctgaacgg tctggttata ggtacattga    1560
gcaactgact gaaatgcctc aaaatgttct ttacgatgcc attgggatat atcaacggtg    1620
gtatatccag tgattttttt ctccatttta gcttccttag ctcctgaaaa tctcgataac    1680
tcaaaaaata cgcccggtag tgatcttatt tcattatggt gaaagttgga acctcttacg    1740
tgcccgatca a                                                          1751
```

<210> SEQ ID NO 56
<211> LENGTH: 1751
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized. pBTK113

<400> SEQUENCE: 56

```
tcttgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac cccgtagaaa      60
agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa     120
```

-continued

```
aataaaaacc gataaaagcg gtggtttgtt tgccggatca agagctacca actctttttc     180
cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgttcttcta gtgtagccgt     240
agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaatcc     300
tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac     360
gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca     420
gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta tgagaaagcg     480
ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag     540
gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt     600
ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcagggggg cggagcctat     660
ggaaaaacgc cagcaacgcg gcctttttac ggttcctggc cttttgctgg ccttttgctc     720
acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtagtcgg tctcaaacgc     780
atacgggagt ttattcttga catattgccg gtgtgttggt ataataactt agtaccgtta     840
tacagaatca gaggagatct taccatgtgc agcccagacc aatcaaaatc gaccggaggc     900
aaccgagcgc tatgaaacca taccgatgga gttcagaggc aattactggc tatttcaacc     960
ggagtccaag cgagctcgat atcaaattac gccccgccct gccactcatc gcagtactgt    1020
tgtaattcat taagcattct gccgacatgg aagccatcac aaacggcatg atgaacctga    1080
atcgccagcg gcatcagcac cttgtcgcct tgcgtataat atttgcccat ggtgaaaacg    1140
ggggcgaaga agttgtccat attggccacg tttaaatcaa aactggtgaa actcacccag    1200
ggattggctg agacgaaaaa catattctca ataaaccctt tagggaaata ggccaggttt    1260
tcaccgtaac acgccacatc ttgcgaatat atgtgtagaa actgccggaa atcgtcgtgg    1320
tattcactcc agagcgatga aaacgtttca gtttgctcat ggaaaacggt gtaacaaggg    1380
tgaacactat cccatatcac cagctcaccg tctttcattg ccatacgaaa ttccggatga    1440
gcattcatca ggcgggcaag aatgtgaata aaggccggat aaaacttgtg cttatttttc    1500
tttacggtct ttaaaaaggc cgtaatatcc agctgaacgg tctggttata ggtacattga    1560
gcaactgact gaaatgcctc aaaatgttct ttacgatgcc attgggatat atcaacggtg    1620
gtatatccag tgattttttt ctccatttta gcttccttag ctcctgaaaa tctcgataac    1680
tcaaaaaata cgcccggtag tgatcttatt tcattatggt gaaagttgga acctcttacg    1740
tgcccgatca a                                                         1751
```

<210> SEQ ID NO 57
<211> LENGTH: 1761
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pRK119

<400> SEQUENCE: 57

```
aacaaaaccg aactgagcta aatacagcgt gagctatgag aaagcgccac gcttcccgaa      60
gggagaaagg cggacaggta tccggtaagc ggcagggtcg gaacaggaga gcgcacgagg     120
gagcttccag ggggaaacgc ctggtatctt tatagtcctg tcgggtttcg ccacctctga     180
cttgagcgtc gatttttgtg atgctcgtca ggggggcgga gcctatggaa aaacgccagc     240
aacgcggcct ttttacggtt cctggccttt tgctggcctt ttgctcacat gttctttcct     300
```

-continued

```
gagttatccc atgcttatgt ggataaacgt agtcggtatc aaacggtatc aaacagagta     360

ttgacttaaa gtctaaccta taggatactt acagccatcg agagggacac ggcgaatacc     420

gaaacagagg agatattaca tatgtgagac cagaccaata aaaaacgccc ggcggcaacc     480

gagcgtttcg aacaaatcca gatggagttc tgaggtcatt actggatcta tcaacaggag     540

tccaagcgag ctcgatatca aattacgccc cgccctgcca ctcatcgcag tactgttgta     600

attcattaag cattctgccg acatggaagc catcacaaac ggcatgatga acctgaatcg     660

ccagcggcat cagcaccttg tcgccttgcg tataatattt gcccatggtg aaaacggggg     720

cgaagaagtt gtccatattg gccacgttta aatcaaaact ggtgaaactc acccagggat     780

tggctgagac gaaaaacata ttctcaataa accctttagg gaaataggcc aggttttcac     840

cgtaacacgc cacatcttgc gaatatatgt gtagaaactg ccggaaatcg tcgtggtatt     900

cactccagag cgatgaaaac gtttcagttt gctcatggaa aacggtgtaa caagggtgaa     960

cactatccca tatcaccagc tcaccgtctt tcattgccat acgaaattcc ggatgagcat    1020

tcatcaggcg ggcaagaatg tgaataaagg ccggataaaa cttgtgctta tttttcttta    1080

cggtctttaa aaaggccgta atatccagct gaacggtctg gttataggta cattgagcaa    1140

ctgactgaaa tgcctcaaaa tgttctttac gatgccattg ggatatatca acggtggtat    1200

atccagtgat ttttttctcc attttagctt ccttagctcc tgaaaatctc gataactcaa    1260

aaaatacgcc cggtagtgat cttatttcat tatggtgaaa gttggaacct cttacgtgcc    1320

cgatcaatca tgaccaaaat cccttaacgt gagttttcgt tccactgagc gtcagacccc    1380

gtagaaaaga tcaaaggatc ttcttgagat cctttttttc tgcgcgtaat ctgctgcttg    1440

caaacaaaaa aaccaccgct accagcggtg gtttgtttgc cggatcaaga gctaccaact    1500

ctttttccga aggtaactgg cttcagcaga gcgcagatac caaatactgt tcttctagtg    1560

tagccgtagt taggccacca cttcaagaac tctgtagcac cgcctacata cctcgctctg    1620

ctaatcctgt taccagtggc tgctgccagt ggcgataagt cgtgtcttac cgggttggac    1680

tcaagacgat agttaccgga taaggcgcag cggtcgggct gaacgggggg ttcgtgcaca    1740

cagcccagct tggagcgaac g                                               1761
```

<210> SEQ ID NO 50
<211> LENGTH: 1762
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized. pBTK120

<400> SEQUENCE: 58

```
cagcccagct tggagcgaac gacctacacc gaactgagat acctacagcg tgagctatga       60

gaaagcgcca cgcttcccga agggagaaag gcggacaggt atccggtaag cggcagggtc      120

ggaacaggag agcgcacgag ggagcttcca gggggaaacg cctggtatct ttatagtcct      180

gtcgggtttc gccacctctg acttgagcgt cgatttttgt gatgctcgtc aggggggcgg      240

agcctatgga aaaacgccag caacgcggcc tttttacggt tcctggcctt ttgctggcct      300

tttgctcaca tgttctttcc tgcgttatcc cctgattctg tggataaccg tagtcggtat      360

caaacgcacg aaaaacaggt attgacaaca tgaagtaaca tgcagtaaga tacaaatcgc      420

taggtaacac tagcagcata cagaaacaga ggagatatta catatgtgag accagaccaa      480
```

-continued

```
ataaaaacgc ccggcggcaa ccgagcgttc tgaacaaatc cagatggagt tctgaggtca     540
ttactggatc tatcaacagg agtccaagcg agctcgatat caaattacgc cccgccctgc     600
cactcatcgc agtactgttg taattcatta agcattctgc cgacatggaa gccatcacaa     660
acggcatgat gaacctgaat cgccagcggc atcagcacct tgtcgccttg cgtataatat     720
ttgcccatgg tgaaaacggg ggcgaagaag ttgtccatat tggccacgtt taaatcaaaa     780
ctggtgaaac tcacccaggg attggctgag acgaaaaaca tattctcaat aaacccttta     840
gggaaatagg ccaggttttc accgtaacac gccacatctt gcgaatatat gtgtagaaac     900
tgccggaaat cgtcgtggta ttcactccag agcgatgaaa acgtttcagt ttgctcatgg     960
aaaacggtgt aacaagggtg aacactatcc catatcacca gctcaccgtc tttcattgcc    1020
atacgaaatt ccggatgagc attcatcagg cgggcaagaa tgtgaataaa ggccggataa    1080
aacttgtgct tatttttctt tacggtcttt aaaaaggccg taatatccag ctgaacggtc    1140
tggttatagg tacattgagc aactgactga aatgcctcaa aatgttcttt acgatgccat    1200
tgggatatat caacggtggt atatccagtg atttttttct ccattttagc ttccttagct    1260
cctgaaaatc tcgataactc aaaaaatacg cccggtagtg atcttatttc attatggtga    1320
aagttggaac ctcttacgtg ccgatcaact catgaccaaa atcccttaac gtgagttttc    1380
gttccactga gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag atcctttttt    1440
tctgcgcgta atctgctgct tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt    1500
gccggatcaa gagctaccaa ctctttttcc gaaggtaact ggcttcagca gagcgcagat    1560
accaaatact gttcttctag tgtagccgta gttaggccac cacttcaaga actctgtagc    1620
accgcctaca tacctcgctc tgctaatcct gttaccagtg gctgctgcca gtggcgataa    1680
gtcgtgtctt accgggttgg actcaagacg atagttaccg gataaggcgc agcggtcggg    1740
ctgaacgggg ggttcgtgca ca                                             1762
```

<210> SEQ ID NO 59
<211> LENGTH: 1701
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pDTK121

<400> SEQUENCE: 59

```
ctgagatacc tacagcgtga gctatgagaa agcgccacgc ttcccgaagg gagaaaggcg      60
gacaggtatc cggtaagcgg cagggtcgga acaggagagc gcacgaggga gcttccaggg     120
ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc acctctgact tgagcgtcga     180
tttttgtgat gctcgtcagg ggggcggagc ctatggaaaa acgccagcaa cgcggccttt     240
ttacggttcc tggccttttg ctggcctttt gctcacatgt tctttcctgc gttatcccct     300
gattctgtgg ataaccgtag tcggtctcaa acgggtgaaa caaaacggtt gacaacatga     360
agtaaaaacg gtacgatgta ccacatgaaa cgacagtgag tcaccaaaga aaccgaggga     420
atattaacta tgtgagacca gaccaataaa aaacgcccgg cggcaaccga gcgttctgaa     480
aaaaaaaaga tggagttctg aggtcattac tggatctatc aacaggagtc caagcgagct     540
cgatatcaaa ttacgccccg ccctgccact catcgcagta ctgttgtaat tcattaagca     600
ttctgccgac atggaagcca tcacaaacgg catgatgaac ctgaatcgcc agcggcatca     660
```

continued

```
gatcctgtc gactcgagta tactacttgc aattggtgca aacgggggcg aagcagttgt     720

aatcattggc aacgcttcca tcaccaatgg tgcaactaac aatgggattg gatgaaaagc     780

aaaacatatt ctcaataaac cctttaggga aatagcccag gttttcaccg taacacgcca     840

catcttgcga atatatgtgt agaaactgcc ggaaatcgtc gtggtattca ctccagagcg     900

atgaaaacgt ttcagtttgc tcatggaaaa cggtgtaaca agggtgaaca ctatcccata     960

tcaccagctc accgtctttc attgccatac gaaattccgg atgagcattc atcaggcggg    1020

caagaatgtg aataaaggcc ggataaaact tgtgcttatt tttctttacg gtctttaaaa    1080

aggccgtaat atccagctga acggtctggt tataggtaca ttgagcaact gactgaaatg    1140

cctcaaaatg ttctttacga tgccattggg atatatcaac ggtggtatat ccagtgattt    1200

ttttctccat tttagcttcc ttagctcctg aaaatctcga taactcaaaa aatacgcccg    1260

gtagtgatct tatttcatta tggtgaaagt tggaacctct tacgtgcccg atcaatcatg    1320

accaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccgt agaaaagatc    1380

aaaggatctt cttgagatcc tttttttctg cgcgtaatct gctgcttgca aacaaaaaaa    1440

accaccgcta ccagcggtgg tttgtttgcc ggatcaagag ctaccaactc tttttccgaa    1500

gtaactggct tcagcagagc gcagatacca aatactgttc ttctagtgta gccgtagtta    1560

ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct aatcctgtta    1620

ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttggactc aagacgatag    1680

ttaccggata aggcgcagcg gtcgggctga acggggggtt cgtgcacaca gcccagcttg    1740

gagcgaacga cctacaccga a                                               1761
```

<210> SEQ ID NO 60
<211> LENGTH: 1700
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pBTK138

<400> SEQUENCE: 60

```
tactgaacca aatcccttca agtgagtttt agttccactg agcgtcagac cccgtagaaa      60

agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa     120

aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca actctttttc     180

cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgttcttcta gtgtagccgt     240

agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaatcc     300

tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac     360

gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca     420

gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta tgagaaagcg     480

ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag     540

gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt     600

ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcagggggg cggagcctat     660

ggaaaaacgc cagcaacgcg gcctttttac ggttcctggc cttttgctgg ccttttgctc     720

acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtattacc gcctttgagt     780

gagctgatac cgctcgccgc agccgaacga ccgagcgcag cgagtcagtg agcgaggaag     840
```

-continued

```
aatatgtgtt tattccagcc aactcaattc atgtcaagta tattttgact aagccgatgt      900

gacccagacc attaaaattg aacggcggca aacgagagtt atgaacaaat aacgatgggg      960

ttctgaggtc attactggat ctatcaacag gagtccaagc gagctcgata tcaaattacg     1020

ccccgccctg ccactcatcg cagtactgtt gtaattcatt aagcattctg ccgacatgga     1080

agccatcaca aacggcatga tgaacctgaa tcgccagcgg catcagcacc ttgtcgcctt     1140

gcgtataata tttgcccatg gtgaaaacgg gggcgaagaa gttgtccata ttggccacgt     1200

ttaaatcaaa actggtgaaa ctcacccagg gattggctga gacgaaaaac atattctcaa     1260

taaacccttt agggaaatag gccaggtttt caccgtaaca cgccacatct tgcgaatata     1320

tgtgtagaaa ctgccggaaa tcgtcgtggt attcactcca gagcgatgaa aacgtttcag     1380

tttgctcatg gaaaacggtg taacaagggt gaacactatc ccatatcacc agctcaccgt     1440

ctttcattgc catacggaat tccggatgag cattcatcag gcgggcaaga atgtgaataa     1500

aggccggata aaacttgtgc ttatttttct ttacggtctt taaaaaggcc gtaatatcca     1560

gctgaacggt ctggttatag gtacattgag caactgactg aaatgcctca aaatgttctt     1620

tacgatgcca ttgggatata tcaacggtgg tatatccagt gatttttttc tccattttag     1680

cttccttagc tcctgaaaat ctcgataact caaaaaatac gcccggtagt gatcttattt     1740

cattatggtg aaagttggaa cctcttacgt gcccgatcaa                           1780
```

<210> SEQ ID NO 61
<211> LENGTH: 4315
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pR 620

<400> SEQUENCE: 61

```
tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac cccgtagaaa       60

agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa      120

aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca actctttttc      180

cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgttcttcta gtgtagccgt      240

agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaatcc      300

tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac      360

gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca      420

gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta tgagaaagcg      480

ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag      540

gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt      600

ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcagggggg cggagcctat      660

ggaaaaacgc cagcaacgcg gcctttttac ggttcctggc cttttgctgg ccttttgctc      720

acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtattacc gcctttgagt      780

acacgattca catcgctaag aacgacttct ctgacatcga actggctgct atcccgttca      840

aacctatggc tgcccattcc ggtgagcgtc tagctcgaga aacgctggcc attgagcatg      900

agtcctacga gatgggtgaa gcacgcttcc gcaagatgtt tgagcgtcaa cttaaagctg      960

gtgaggttgc ggataacgct gccgccaagc ctctcatcac taccctactc cctaagatga     1020
```

continued

```
ttgcaagcct aatcgaatgg tttgaggacg tgccagctca gagcggaacg agcccgaacg    1080
actaccagtc aatgcaagca atccagacgg aagccgtagc gtccatcacc attcagaccc    1140
ctccggcttg cctaaccagc gctgacaata caaccgttca ggctgtagca agcgcaatcg    1200
gtcggcgcac tgaggacgag gctcgcttcg gtcgcatccg tgacctgaag gctaagcact    1260
tcaagaaaaa cgttgaggaa caactcaaca agcgcgtagg gcacgctcac aagcaagcat    1320
ttatgcaagt tgtcgaggct gacatgctct ctaagggtct actcggtggc gaggcgtggt    1380
attcgtggca tacggaagcc tatcccactg taggagtacg atgcatacag atgcccattg    1440
agtcaaacgg aatggttagc ttccacagcc aacccgatgg agtagtaggt cacgactatg    1500
agactatcga actcgcaccc gaatacgctg aggctatcgc aacccgtgca ggtgcgctgg    1560
ctggcatctc tccgatgttc caacccgcgc tagttcctcc taagccgtgg actggcatta    1620
ctggtggtgg ctattgggct aacggtcgtc gtccactggc gctggtgcgt actcacagta    1680
agaaagcact gatgcgctac gaagacgttt acatgcctga ggtgtacaaa gcgattaaca    1740
ttgcgcaaca accgcaatgg aaaatcaaca agcaagtact agcggtagcc aacgcaatca    1800
accagtggca gactcgtccg gtcgaggcca tcccgcagca tgcgcgtgca gaacccaagc    1860
tgaaaccgga agacatcgac atgaatcctg aggctctcac cgcgtggaaa cgtgctgccg    1920
ctgctgtgta ccgcaaggac aaggctcgca agtctcgccg tatcagcctt gagttcatgc    1980
ttgagcaagc caataagttt gctaaccata aggccatctg gttcccttac aacatggact    2040
ggcgcggtcg tgtttacgct gtgtcaatgt tcaacccgca aggtaacgat atgaccaaag    2100
gactgcttac gctggcgaaa ggtaaaccaa tcggtaagga aggttactac tggctgaaaa    2160
tccacggtgc aaactgtgcg ggtgtcgata aggttccgtt ccctgagcgc atcaagttca    2220
ttgaggaaaa ccacgagaac atcatggctt gcgctaagtc tccactggag aacacttggt    2280
gggctgagca agattctccg ttctgcttcc ttgcgttctg ctttgagtac gctggggtac    2340
agcaccacgg cctgagctat aactgctccc ttccgctggc gtttgacggg tcttgctctg    2400
gcatccagca cttctccgcg atgctccgag atgaggtagg tggtcgcgcg gttaacttgc    2460
ttcctagtga aaccgttcag gacatctacg ggattgttgc taagaaagtc aacgagattc    2520
tacaagcaga cgcaatcaat gggaccgata acgaagtagt taccgtgacc gatgagaaca    2580
ctggtgaaat ctctgagaaa gtcaagctgg gcactaaggc actggctggt caatggctgg    2640
cttacggtgt tactcgcagt gtgactaagc gttcagtcat gacgctggct tacgggtcca    2700
aagagttcgg cttccgtcaa caagtgctgg aagataccat tcagccagct attgattccg    2760
gcaagggtcc gatgttcact caggccgctc aggctgctgg atacatggct aagctgattt    2820
gggaatctgt gagcgtgacg gtggtagctg cggttgaagc aatgaactgg cttaagtctg    2880
ctgctaagct gctggctgct gaggtcaaag ataagaagac tggagagatt cttcgcaagc    2940
gttgcgctgt gcattgggta actcctgatg gtttccctgt gtggcaggaa tacaagaagc    3000
ctattcagac gcgcttgaac ctgatgttcc tcggtcagtt ccgcttacag cctaccatta    3060
acaccaacaa agatagcgag attgatgcac acaaacagga gtctggtatc gctcctaact    3120
ttgtacacag ccaagacggt agccacctta gtacgactgt agtgtgggca cacgagaagt    3180
acggaatcga atctttgcac ctgattcacg actccttcgg taccattccg gctgacgctg    3240
cgaacctgtt caaagcagtg cgcgaaacta tggttgacac atatgagtct tgtgatgtac    3300
```

-continued

```
tggctgattt ctacgaccag ttcgctgacc agttgcaaga gtctcaattg gaccaaatgc   3360

aagcaattcc ggctaaaggt aactggaccc taagtgacct attcgagtcg gacttcgagt   3420

tcgcgtaaat cctgagacca gaccaataaa aaacgcccgg cggcaaccga gcgttctgaa   3480

caaatccaga tggagttctg aggtcattac tggatctatc aacaggagtc caagcgagct   3540

cgatatcaaa ttacgccccg ccctgccact catcgcagta ctgttgtaat tcattaagca   3600

ttctgccgac atggaagcca tcacaaacgg catgatgaac ctgaatcgcc agcggcatca   3660

gcaccttgtc gccttgcgta taatatttgc ccatggtgaa aacgggggcg aagaagttgt   3720

ccatattggc cacgtttaaa tcaaaactgg tgaaactcac ccagggattg gctgagacga   3780

aaaacatatt ctcaataaac cctttaggga aataggccag gttttcaccg taacacgcca   3840

catcttgcga atatatgtgt agaaactgcc ggaaatcgtc gtggtattca ctccagagcg   3900

atgaaaacgt ttcagtttgc tcatggaaaa cggtgtaaca agggtgaaca ctatcccata   3960

tcaccagctc accgtctttc attgccatac gaaattccgg atgagcattc atcaggcggg   4020

caagaatgtg aataaaggcc ggataaaact tgtgcttatt tttctttacg gtctttaaaa   4080

aggccgtaat atccagctga acggtctggt tataggtaca ttgagcaact gactgaaatg   4140

cctcaaaatg ttctttacga tgccattggg atatatcaac ggtggtatat ccagtgattt   4200

ttttctccat tttagcttcc ttagctcctg aaaatctcga taactcaaaa aatacgcccg   4260

gtagtgatct tatttcatta tggtgaaagt tggaacctct tacgtgcccg atcaa        4315
```

<210> SEQ ID NO 62
<211> LENGTH: 2746
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pDIK203

<400> SEQUENCE: 62

```
tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac cccgtagaaa     60

agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa    120

aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca actctttttc    180

cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgttcttcta gtgtagccgt    240

agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaatcc    300

tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac    360

gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca    420

gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta tgagaaagcg    480

ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag    540

gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt    600

ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcagggggg cggagcctat    660

ggaaaaacgc cagcaacgcg gcctttttac ggttcctggc cttttgctgg ccttttgctc    720

acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtagtcgg tctcatatga    780

aaccagtaac gttatacgat gtcgcagagt atgccggtgt ctcttatcag accgtttccc    840

gcgtggtgaa ccaggccagc cacgtttctg cgaaaacgcg ggaaaaagtg gaagcggcga    900

tggcggagct gaattacatt cccaaccgcg tggcacaaca actggcgggc aaacagtcgt    960
```

continued

```
tgcagattgg agttgaaacc taaagaatgg aactgcaagc gacgacgaac attgacgagg   1020
agaaaaaatc tagcgaagaa aaacagggtg aaagcgtggc ggtgacgatg gtagaaagaa   1080
gcggcgtcga agcccgtaaa gcggccgtgc acaatcttcc cgcgcaacgc gtcagtgggc   1140
tgatcattaa ctatccgctg gatgaccagg atgccattgc tgtggaagct gcctgcacta   1200
atgttccggc gttatttctt gatgtctctg accagacacc catcaacagt attattttct   1260
cccatgaaga cggtacgcga ctgggcgtgg agcatctggt cgcattgggt caccagcaaa   1320
tcgcgctgtt agcgggccca ttaagttctg tctcggcgcg tctgcgtctg gctggctggc   1380
ataaatatct cactcgcaat caaattcagc cgatagcgga acgggaaggc gactggagtg   1440
ccatgtccgg ttttcaacaa accatgcaaa tgctgaatga gggcatcgtt cccactgcga   1500
tgctggttgc caacgatcag atggcgctgg gcgcaatgcg cgccattacc gagtccgggc   1560
tgcgcgttgg tgcggatatc tcggtagtgg gatacgacga taccgaagac agctcatgtt   1620
atatcccgcc gttaaccacc atcaaacagg attttcgcct gctggggcaa accagcgtgg   1680
accgcttgct gcaactctct cagggccagg cggtgaaggg caatcagctg ttgcccgtct   1740
cactggtgaa aagaaaaacc accctggcgc ccaatacgca aaccgcctct ccccgcgcgt   1800
tggccgattc attaatgcag ctggcacgac aggtttcccg actggaaagc gggcagtgaa   1860
tcctgagacc agaccaataa aaaacgcccg gcggcaaccg agcgttctga acaaatccag   1920
atggagttct gaggtcatta ctggatctat caacaggagt ccaagcgagc tcgatatcaa   1980
attacgcccc gccctgccac tcatcgcagt actgttgtaa ttcattaagc attctgccga   2040
catggaagcc atcacaaacg gcatgatgaa cctgaatcgc cagcggcatc agcaccttgt   2100
cgccttgcgt ataatatttg cccatggtga aaacgggggc gaagaagttg tccatattgg   2160
ccacgtttaa atcaaaactg gtgaaactca cccagggatt ggctgagacg aaaaacatat   2220
tctcaataaa ccctttaggg aaataggcca ggttttcacc gtaacacgcc acatcttgcg   2280
aatatatgtg tagaaactgc cggaaatcgt cgtggtattc actccagagc gatgaaaacg   2340
tttcagtttg ctcatggaaa acggtgtaac aagggtgaac actatcccat atcaccagct   2400
caccgtcttt cattgccata cgaaattccg gatgagcatt catcaggcgg gcaagaatgt   2460
gaataaaggc cggataaaac ttgtgcttat ttttctttac ggtctttaaa aaggccgtaa   2520
tatccagctg aacggtctgg ttataggtac attgagcaac tgactgaaat gcctcaaaat   2580
gttctttacg atgccattgg gatatatcaa cggtggtata tccagtgatt tttttctcca   2640
ttttagcttc cttagctcct gaaaatctcg ataactcaaa aaatacgccc ggtagtgatc   2700
ttatttcatt atggtgaaag ttggaacctc ttacgtgccg atcaa                   2746
```

<210> SEQ ID NO 63
<211> LENGTH: 2386
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pK 6205

<400> SEQUENCE: 63

```
ttctgaaaaa aatcaattca agtgagtttt agttccactg agcgtcagac cccgtagaaa     60
agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa    120
aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca actctttttc    180
```

-continued

```
cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgttcttcta gtgtagccgt     240
agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaatcc     300
tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac     360
gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca     420
gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta tgagaaagcg     480
ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag     540
gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt     600
ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcagggggg cggagcctat     660
ggaaaaacgc cagcaacgcg gcctttttac ggttcctggc cttttgctgg ccttttgctc     720
acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtagtcgg tctcatatga     780
gtaaaggaga agagcttttc accggagttg tcccaatcct cgtggaattg gacggtgatg     840
ttaatgggca caagttctct gtcagtggag agggtgaagg cgacgcaaca tatggcaagc     900
tgacccttaa atttatttgc accactggta aactacctgt tccatggcca acactggtca     960
ctacgttcgg gtatggggtt cagtgctttg cgcgctaccc agatcacatg aaacagcacg    1020
actttttcaa gagtgccatg cccgaaggct atgtacagga gagaaccatc tttttcaagg    1080
atgacggcaa ctataagaca cgcgccgaag tgaagttcga gggtgatacc cttgttaata    1140
gaatcgagtt aaagggtatt gacttcaagg aagatggaaa tattttaggc cacaaactgg    1200
aatataacta taactcccat aatgtgtaca ttatggccga caagcaaaag aacggtatca    1260
aggttaactt caagattaga cacaacattg aggatggaag cgttcaactc gccgaccatt    1320
accaacaaaa caccccaatc ggcgatggcc ctgtgctgtt accagacaac cattacctgt    1380
ccactcaatc tgtcctttcg aaagatccca acgaaaagcg cgaccacatg gtccttcttg    1440
agtttgtcac ggctgctggg attacacacg gcatggatga actatacaaa taaatcctga    1500
gaccagacca ataaaaaacg cccggcggca accgagcgtt ctgaacaaat ccagatggag    1560
ttctgaggtc attactggat ctatcaacag gagtccaagc gagctcgata tcaaattacg    1620
ccccgccctg ccactcatcg cagtactgtt gtaattcatt aagcattctg ccgacatgga    1680
agccatcaca aacggcatga tgaacctgaa tcgccagcgg catcagcacc ttgtcgcctt    1740
gcgtataata tttgcccatg gtgaaaacgg gggcgaagaa gttgtccata ttggccacgt    1800
ttaaatcaaa actggtgaaa ctcacccagg gattggctga gacgaaaaac atattctcaa    1860
taaacccttt agggaaatag gccaggtttt caccgtaaca cgccacatct tgcgaatata    1920
tgtgtagaaa ctgccggaaa tcgtcgtggt attcactcca gagcgatgaa aacgtttcag    1980
tttgctcatg gaaaacggtg taacaagggt gaacactatc ccatatcacc agctcaccgt    2040
ctttcattgc catacgaaat tccggatgag cattcatcag gcgggcaaga atgtgaataa    2100
aggccggata aaacttgtgc ttatttttct ttacggtctt taaaaaggcc gtaatatcca    2160
gctgaacggt ctggttatag gtacattgag caactgactg aaatgcctca aaatgttctt    2220
tacgatgcca ttgggatata tcaacggtgg tatatccagt gatttttttc tccattttag    2280
cttccttagc tcctgaaaat ctcgataact caaaaaatac gcccggtagt gatcttattt    2340
cattatggtg aaagttggaa cctcttacgt gcccgatcaa                          2380
```

<210> SEQ ID NO 64

-continued

<211> LENGTH: 2179
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pDIK206

<400> SEQUENCE: 64

tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac cccgtagaaa 60

agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa 120

aaaaaaaccc gctaccagcg gtggtttgtt tgccggatca agagctacca actctttttc 180

cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgttcttcta gtgtagccgt 240

agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaatcc 300

tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac 360

gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca 420

gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta tgagaaagcg 480

ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag 540

gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt 600

ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcagggggg cggagcctat 660

ggaaaaacgc cagcaacgcg gcctttttac ggttcctggc cttttgctgg ccttttgctc 720

acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtagtcgg tctcatatgg 780

tcttcacact cgaagatttc gttggggact ggcgacagac agccggctac aacctggacc 840

aagtcattga aacggaggtt gtgtccagtt tgtttcagta tatcggggtg taagtaaatc 900

agttcaaccg gattgtactg agcggtgcaa atgggctgca gatcgaaatc aatgacatcc 960

tcccgtatga aggtctgagc ggcgaccaaa tgggccagat cgaaaaaatt tttaaggtgg 1020

tgtaccctgc ggatgatcat cactttaagg tgatcctgca ctatggcaca ctggtaatcg 1080

acggggttac gccgaacatg atcgactatt tcggacggcc gtatgaaggc atcgccgtgt 1140

tcgacggcaa aaagatcact gtaacaggga ccctgtggaa cggcaacaaa attatcgacg 1200

agcgcctgat caacccagac ggctccctgc tgttccgagt aaccatcaac ggagtgaccg 1260

gctggcggct gtgcgaacgc attctggcgt aaatcctgag acgagaccaa taaaaaacgc 1320

ccggcggcaa ccgagcgttc tgaacaaatc cagatggagt tctgaggtca ttactggatc 1380

tatcaacagg agtccaagcg agctcgatat caaattacgc cccgccctgc cactcatcgc 1440

agtactgttg taattcatta agcattctgc cgacatggaa gccatcacaa acggcatgat 1500

gaacctgaat cgccagcggc atcagcacct tgtcgccttg cgtataatat ttgcccatgg 1560

tgaaaacggg ggcgaagaag ttgtccatat tggccacgtt taaatcaaaa ctggtgaaac 1620

tcacccaggg attggctgag acgaaaaaca tattctcaat aaaccctttag ggaaatagg 1680

ccaggttttc accgtaacac gccacatctt gcgaatatat gtgtagaaac tgccggaaat 1740

cgtcgtggta ttcactccag agcgatgaaa acgtttcagt ttgctcatgg aaaacggtgt 1800

aacaagggtg aacactatcc catatcacca gctcaccgtc tttcattgcc atacgaaatt 1860

ccggatgagc attcatcagg cgggcaagaa tgtgaataaa ggccggataa aacttgtgct 1920

tatttttctt tacggtcttt aaaaaggccg taatatccag ctgaacggtc tggttatagg 1980

tacattgagc aactgactga aatgcctcaa aatgttcttt acgatgccat tgggatatat 2040

continued

```
aaacggtggc atatccagtg atttttttct ccattttagc ttccttagct cctgaaaatc    2100

tagataactc aaccaatacg ccggtcagtg atatcatttc atttcggtga aagtttggac    2160

ctcttacgtg cccgatcaa                                                  2179
```

<210> SEQ ID NO 65
<211> LENGTH: 3773
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pR 6209

<400> SEQUENCE: 65

```
tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac cccgtagaaa      60

agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa     120

aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca actctttttc     180

cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgttcttcta gtgtagccgt     240

agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaatcc     300

tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac     360

gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca     420

gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta tgagaaagcg     480

ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag     540

gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt     600

ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcagggggg cggagcctat     660

ggaaaaacgc cagcaacgcg gcctttttac ggttcctggc cttttgctgg ccttttgctc     720

acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtagtcgg tctcatatga     780

tggataagaa atactcaata ggcttagcta tcggcacaaa tagcgtcgga tgggcggtga     840

tcactgatga atataaggtt ccgtctaaaa agttcaaggt tctgggaaat acagaccgcc     900

acagtatcaa aaaaaatctt ataggggctc ttttatttga cagtggagag acagcggaag     960

cgactcgtct caaacggaca gctcgtagaa ggtatacacg tcggaagaat cgtatttgtt    1020

atctacagga gattttttca aatgagatgg cgaaagtaga tgatagtttc tttcatcgac    1080

ttgaagagtc ttttttggtg gaagaagaca agaagcatga acgtcatcct atttttggaa    1140

atatagtaga tgaagttgct tatcatgaga aatatccaac tatctatcat ctgcgaaaaa    1200

aattggtaga ttctactgat aaagcggatt tgcgcttaat ctatttggcc ttagcgcata    1260

tgattaagtt tcgtggtcat tttttgattg agggagattt aaatcctgat aatagtgatg    1320

tggacaaact atttatccag ttggtacaaa cctacaatca attatttgaa gaaaacccta    1380

ttaacgcaag tggagtagat gctaaagcga ttctttctgc acgattgagt aaatcaagac    1440

gattagaaaa tctcattgct cagctccccg gtgagaagaa aaatggctta tttgggaatc    1500

tcattgcttt gtcattgggt ttgaccccta attttaaatc aaattttgat ttggcagaag    1560

atgctaaatt acagctttca aaagatactt acgatgatga tttagataat ttattggcgc    1620

aaattggaga tcaatatgct gatttgtttt tggcagctaa gaatttatca gatgctattt    1680

tactttcaga tatcctaaga gtaaatactg aaataactaa ggctccccta tcagcttcaa    1740

tgattaaacg ctacgatgaa catcatcaag acttgactct tttaaaagct ttagttcgac    1800
```

-continued

aacaattacc agaaaagtat aaagaaattt tctttgatca atcaaaaaac ggatatgcag 1860

gttatattga tgggggagca agccaagaag aatttataaa atttataaaa ccaattttag 1920

aaaaaatgga tggtactgag gaattattgg tgaaactaaa tcgtgaagat ttgctgcgca 1980

agcaacggac ctttgacaac ggctctattc cccatcaaat tcacttgggt gagctgcatg 2040

ctattttgag aagacaagaa gacttttatc catttttaaa agacaatcgt gagaagattg 2100

aaaaaatctt gacttttcga attccttatt atgttggtcc attggcgcgt ggcaatagtc 2160

gttttgcatg gatgactcgg aagtctgaag aaacaattac accatggaat tttgaagaag 2220

ttgtcgataa aggtgcttca gctcaatcat ttattgaacg catgacaaac tttgataaaa 2280

atcttccaaa tgaaaaagta ctaccaaaac atagtttgct ttatgagtat tttacggttt 2340

ataacgaatt gacaaaggtc aaatatgtta ctgaaggaat gcgaaaacca gcatttcttt 2400

caggtgaaca gaagaaagcc attgttgatt tactcttcaa aacaaatcga aaagtaaccg 2460

ttaagcaatt aaaagaagat tatttcaaaa aaatagaatg ttttgatagt gttgaaattt 2520

caggagttga agatagattt aatgcttcat taggtaccta ccatgatttg ctaaaaatta 2580

ttaaagataa agattttttg gataatgaag aaaatgaaga tatcttagag gatattgttt 2640

taacattgac cttatttgaa gatagggaga tgattgagga aagacttaaa acatatgctc 2700

acctctttga tgataaggtg atgaaacagc ttaaacgtcg ccgttatact ggttggggac 2760

gtttgtctcg aaaattgatt aatggtatta gggataagca atctggcaaa acaatattag 2820

attttttgaa atcagatggt tttgccaatc gcaattttat gcagctgatc catgatgata 2880

gtttgacatt taaagaagac attcaaaaag cacaagtgtc tggacaaggc gatagtttac 2940

atgaacatat tgcaaattta gctggtagcc ctgctattaa aaaaggtatt ttacagactg 3000

taaaagttgt tgatgaattg gtcaaagtaa tggggcggca taagccagaa aatatcgtta 3060

ttgaaatggc acgtgaaaat cagacaactc aaaagggcca gaaaaattcg cgagagcgta 3120

tgaaacgaat cgaagaaggt atcaaagaat taggaagtca gattcttaaa gagcatcctg 3180

ttgaaaatac tcaattgcaa aatgaaaagc tctatctcta ttatctccaa aatggaagag 3240

acatgtatgt ggaccaagaa ttagatatta atcgtttaag tgattatgat gtcgatcaca 3300

ttgttccaca aagtttcctt aaagacgatt caatagacaa taaggtctta acgcgttctg 3360

ataaaaatcg tggtaaatcg gataacgttc caagtgaaga agtagtcaaa aagatgaaaa 3420

actattggag acaacttcta aacgccaagt taatcactca acgtaagttt gataatttaa 3480

cgaaagctga acgtggaggt ttgagtgaac ttgataaagc tggttttatc aaacgccaat 3540

tggttgaaac tcgccaaatc actaagcatg tggcacaaat tttggatagt cgcatgaata 3600

ctaaatacga tgaaaatgat aaacttattc gagaggttaa agtgattacc ttaaaatcta 3660

aattagtttc tgacttccga aaagatttcc aattctataa agtacgtgag attaacaatt 3720

accatcatgc ccatgatgcg tatctaaatg ccgtcgttgg aactgctttg attaagaaat 3780

atccaaaact tgaatcggag tttgtctatg gtgattataa agtttatgat gttcgtaaaa 3840

tgattgctaa gtctgagcaa gaaataggca aagcaaccgc aaaatatttc ttttactcta 3900

atatcatgaa cttcttcaaa acagaaatta cacttgcaaa tggagagatt cgcaaacgcc 3960

ctctaatcga aactaatggg gaaactggag aaattgtctg ggataaaggg cgagattttg 4020

ccacagtgcg caaagtattg tccatgcccc aagtcaatat tgtcaagaaa acagaagtac 4080

-continued

```
agtcaggcgg attctaaacg gagtcaattt taccaaaaag aattcgggtc aagcttattg    4140

atcgtaatca agtctgggtt aataaaaatt atggtggttt tgctagtacc aaggtagatt    4200

attcagtcct agtggttgct aaggttgaaa aagggaaatc gaagaagtta aaatccgtta    4260

aagagttact agggatcaca attatggaaa gaagttcctt tgaaaaaaat ccgattgact    4320

tttaagaagc taaaggatat aaggaagtta aaaaagaatt aatcattaaa ctaccttaat    4380

atagtttttt tgagttagaa aacggttgta aacggatgct ggctagtgcc ggagaattcc    4440

aataaggtta tgatgctggt atgccaagca aattatgtaa ttttttatat ttagctagtc    4500

attatgatta gttgaagggt agtccagtag attacgaaca aatacaattg tttgttgagc    4560

agcataagca ttatttagat gagattattg agcaaatcag tgaattttct aagcgtgtta    4620

ttttagcaga tgccaattta gataaagttc ttagtgcata taacaaacat agagacaaac    4680

cattacgtga acaagcagaa aatattattc atttattcac gttgacgaat cttggagctc    4740

ccgctgcttt taaatatttt gatacaacaa ttgatcgtaa acgatatacg tctacaaaag    4800

aagttttaga tgccactctt atccatcaat ccatcaatgg tatttatgct aatcgcattg    4860

atttgagtca gattggaggt gactgaatcc tgagacaaga aaatcaaaat aagcccgggg    4920

gcaaccgagc gtttcgaaca aatgcagatg gagttctgag gtcattactg gatctatcaa    4980

caggagtcca agcgagctcg atatcaaatt acgccccgcc ctgccactca tcgcagtact    5040

gttgtaattc attaagcatt ctgccgacat ggaagccatc acaaacggca tgatgaacct    5100

gaatcgccag cggcatcagc accttgtcgc cttgcgtata atatttgccc atggtgaaaa    5160

cgggggcgaa gaagttgtcc atattggcca cgtttaaatc aaaactggtg aaactcaccc    5220

agggattggc tgagacgaaa aacatattct caataaaccc tttagggaaa taggccaggt    5280

tttcaccgta acacgccaca tcttgcgaat atatgtgtag aaactgccgg aaatcgtcgt    5340

ggtattcact ccagagcgat gaaaacgttt cagtttgctc atggaaaacg gtgtaacaag    5400

ggtgaacact atcccatatc accagctcac cgtctttcat tgccatacgg aattccggat    5460

gagcattcat caggcgggca agaatgtgaa taaaggccgg ataaaacttg tgcttatttt    5520

tctttacggt ctttaaaaag gccgtaatat ccagctgaac ggtctggtta taggtacatt    5580

gagcaactga ctgaaatgcc tcaaaatgtt ctttacgatg ccattgggat atatcaacgg    5640

tggtatatcc agtgattttt ttctccattt tagcttcctt agctcctgaa aatctcgata    5700

actcaaaaaa tacgcccggt agtgatctta tttcattatg gtgaaagttg gaacctctta    5760

cgtgccgatc aac                                                      5773
```

```
<210> SEQ ID NO 66
<211> LENGTH: 2749
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pDTK211

<400> SEQUENCE: 66

tttagcttcc ttagctcctg aaaatctcga taactcaaaa aatacgcccg gtagtgatct      60

tatttcatta tggtgaaagt tggaacctct tacgtgccga tcaacgtctc attttcgcca     120

cttaacgtga gttttcgttc cactgagcgt cagaccccgt agaaaagatc aaaggatctt     180

cttgagatcc tttttttctg cgcgtaatct gctgcttgca aacaaaaaaa ccaccgctac     240
```

-continued

```
aagcggtggt ttgtttgccg gatcaagagc taccaactct ttttccgaag gtaactggct     300
tcagcagagc gcagatacca aatactgttc ttctagtgta gccgtagtta ggccaccact     360
tcaagaactc tgtagcaccg cctacatacc tcgctctgct aatcctgtta ccagtggctg     420
ctgccagtgg cgataagtcg tgtcttaccg ggttggactc aagacgatag ttaccggata     480
aggcgcagcg gtcgggctga acggggggtt cgtgcacaca gcccagcttg gagcgaacga     540
cctacaccga actgagatac ctacagcgtg agctatgaga aagcgccacg cttcccgaag     600
ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg aacaggagag cgcacgaggg     660
agcttccagg gggaaacgcc tggtatcttt atagtcctgt cgggtttcgc cacctctgac     720
ttgagcgtcg atttttgtga tgctcgtcag gggggcggag cctatggaaa aacgccagca     780
acgcggcctt tttacggttc ctggcctttt gctggccttt tgctcacatg ttctttcctg     840
cgttatcccc tgattctgtg gataaccgta gtcggtctca tacatcactg cccgctttcc     900
agtcgggaaa cctgtcgtgc cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg     960
gtttgcgtat tgggcgccag ggtggttttt cttttcacca gtgagacggg caacagctga    1020
ttgcccttca ccgcctggcc ctgagagagt tgcagcaagc ggtccacgct ggtttgcccc    1080
agcaggcgaa aatcctgttt gatggtggtt aacggcggga tataacatga gctgtcttcg    1140
gtatcgtcgt atcccactac cgagatatcc gcaccaacgc gcagcccgga ctcggtaatg    1200
gcgcgcattg cgcccagcgc catctgatcg ttggcaacca gcatcgcagt gggaacgatg    1260
ccctcattca gcatttgcat ggtttgttga aaaccggaca tggcactcca gtcgccttcc    1320
cgttccgcta tcggctgaat ttgattgcga gtgagatatt tatgccagcc agccagacgc    1380
agacgcgccg agacagaact taatgggccc gctaacagcg cgatttgctg gtgacccaat    1440
gcgaccagat gctccacgcc cagtcgcgta ccgtcttcat gggagaaaat aatactgttg    1500
atgggtgtct ggtcagagac atcaagaaat aacgccggaa cattagtgca ggcagcttcc    1560
acagcaatgg catcctggtc atccagcgga tagttaatga tcagcccact gacgcgttgc    1620
gcgagaagat tgtgcaccgc cgctttacag gcttcgacgc cgcttcgttc taccatcgac    1680
accaccacgc tggcacccag ttgatcggcg cgagatttaa tcgccgcgac aatttgcgac    1740
ggcgcgtgca gggccagact ggaggtggca acgccaatca gcaacgactg tttgcccgcc    1800
agttgttgtg ccacgcggtt gggaatgtaa ttcagctccg ccatcgccgc ttccactttt    1860
tcccgcgttt tcgcagaaac gtggctggcc tggttcacca cgcgggaaac ggtctgataa    1920
gagacaccgg catactctgc gacatcgtat aacgttactg gtttcatgcg ttgagacccg    1980
aacaatacca aaacgcaggc ggcaacaagg agttctgatc aaatccagat ggaggaatga    2040
ggtcattacc ggatcatcca aaaggagtcc aagcgagatc gataacaaat tacgccccgc    2100
cctgccactc atcgcagtac tgttgtaatt cattaagcat tctgccgaca tggaagccat    2160
cacaaacggc atgatgaacc tgaatcgcca gcggcatcag caccttgtcg ccttgcgtat    2220
aatatttgcc catggtgaaa acgggggcga agaagttgtc catattggcc acgtttaaat    2280
caaaactggt gaaactcacc cagggattgg ctgagacgaa aaacatattc tcaataaacc    2340
ctttagggaa ataggccagg ttttcaccgt aacacgccac atcttgcgaa tatatgtgta    2400
gaaactgccg gaaatcgtcg tggtattcac tccagagcga tgaaaacgtt tcagtttgct    2460
catggaaaac ggtgtaacaa gggtgaacac tatcccatat caccagctca ccgtctttca    2520
```

-continued

ttgccatacg gaactccgga tgagcattca tcaggcgggc aagaatgtga ataaaggccg 2580

gataaaactt gtgcttattt ttctttacgg tctttaaaaa ggccgtaata tccagctgaa 2640

cggtctggtt ataggtacat tgagcaactg actgaaatgc ctcaaaatgt tctttacgat 2700

gccattggga tatatcaacg gtggtatatc cagtgatttt tttctccat 2749

<210> SEQ ID NO 67
<211> LENGTH: 2380
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pEK204

<400> SEQUENCE: 67

tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac cccgtagaaa 60

agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa 120

aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca actctttttc 180

cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgttcttcta gtgtagccgt 240

agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaatcc 300

tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac 360

gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca 420

gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta tgagaaagcg 480

ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag 540

gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt 600

ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcagggggg cggagcctat 660

ggaaaaacgc cagcaacgcg gcctttttac ggttcctggc cttttgctgg ccttttgctc 720

acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtagtcgg tctcatatgg 780

atagcactga gaacgtcatc aagcccttca tgcgcttcaa ggtgcacatg gagggctccg 840

tgaacggcca cgagttcgag atcgagggcg tgggcgaggg cgagccctac gagggcaccc 900

agaccgccaa gctgaaggtg accaagggcg gccccctgcc cttcgcctgg gacatcctgt 960

cccctcagtt cttctacggc tccaaggcgt acatcaagca ccccgccgac atccccgact 1020

acttcaagca gtccttcccc gagggcttca agtgggagcg cgtgatgaac ttcgaggacg 1080

gcggcgtggt gaccgtgacc caggactcct ccctgcagga cggcaccctc atctacaacg 1140

tgaagttcat cggcgtgaac ttccccctcg acggccccgt aatgcagaag aagaccatgg 1200

gatgggagcc atccaatgcg agcaactacc aacgcgaagg agtgcagatg ggcgagaacc 1260

aactggagct gaagctgaag ggcggcggcc actacgatgc tgaggtcaag taaaactacc 1320

tggccaagaa gaaagtgcag atgcccggcg aactctaagc ggactacaag atcgacatca 1380

cctcccacaa cgaggactac accgtcgtgg agcagtacga gcgcgccgag gccgcccacc 1440

accggttcca gactcaaggt atggacgaac tgtacaagca agacgaattg taaaacatga 1500

gaccagacca ataaaaaacg cccggcggca accgagcgtt ctgaacaaat ccagatggag 1560

ttctgaggtc attactggat ctatcaacag gagtccaagc gagctcgata tcaaattacg 1620

ccccgccctg ccactcatcg cagtactgtt gtaattcatt aagcattctg ccgacatgga 1680

agccatcaca aacggcatga tgaacctgaa tcgccagcgg catcagcacc ttgtcgcctt 1740

-continued

```
gcgtataata tttgcccatg gtgaaaacgg gggcgaagaa gttgtccata ttggccacgt    1800

ttaaatcaaa actggtgaaa ctcacccagg gattggctga gacgaaaaac atattctcaa    1860

taaacccttt agggaaatag gccaggtttt caccgtaaca cgccacatct tgcgaatata    1920

tgtgtagaaa ctgccggaaa tcgtcgtggt attcactcca gagcgatgaa aacgtttcag    1980

tttgctcatg gaaaacggtg taacaagggt gaacactatc ccatatcacc agctcaccgt    2040

ctttcattgc catacggaat tccggatgag cattcatcag gcgggcaaga atgtgaataa    2100

aggccggata aaacttgtgc ttatttttct ttacggtctt taaaaaggcc gtaatatcca    2160

gctgaacggt ctggttatag gtacattgag caactgactg aaatgcctca aaatgttctt    2220

tacgatgcca ttgggatata tcaacggtgg tatatccagt gatttttttc tccattttag    2280

cttccttagc tcctgaaaat ctcgataact caaaaaatac gcccggtagt gatcttattt    2340

cattatggtg aaagttggaa cctcttacgt gcccgatcaa                          2380
```

<210> SEQ ID NO 68
<211> LENGTH: 2673
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pDTK229

<400> SEQUENCE: 68

```
atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca cacagcccag      60

cttggagcga acgacctaca ccgaactgag atacctacag cgtgagctat gagaaagcgc     120

cacgcttccc gaagggagaa aggcggacag gtatccggta agcggcaggg tcggaacagg     180

agagcgcacg agggagcttc cagggggaaa cgcctggtat ctttatagtc ctgtcgggtt     240

tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc ggagcctatg     300

gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc ttttgctggc cttttgctca     360

catgttcttt cctgcgttat cccctgattc tgtggataac cgtattaccg cctttgagtg     420

taccgagcgg ccgcgtgtta caaccaatta accaattctg attagaaaaa ctcatcgagc     480

atcaaatgaa actgcaattt attcatatca ggattatcaa taccatattt ttgaaaaagc     540

cgtttctgta atgaaggaga aaactcaccg aggcagttcc ataggatggc aagatcctgg     600

tatcggtctg cgattccgac tcgtccaaca tcaatacaac ctattaattt cccctcgtca     660

aaaataaggt tatcaagtga gaaatcacca tgagtgacga ctgaatccgg tgagaatggc     720

aaaagtttat gcatttcttt ccagacttgt tcaacaggcc agccattacg ctcgtcatca     780

aaatcactcg catcaaccaa accgttattc attcgtgatt gcgcctgagc gagacgaaat     840

acgcgatcgc tgttaaaagg acaattacaa acaggaatcg aatgcaaccg gcgcaggaac     900

actgccagcg catcaacaat attttcacct gaatcaggat attcttctaa tacctggaat     960

gctgttttcc cggggatcgc agtggtgagt aaccatgcat catcaggagt acggataaaa    1020

tgcttgatgg tcggaagagg cataaattcc gtcagccagt ttagtctgac catctcatct    1080

gtaacatcat tggcaacgct acctttgcca tgtttcagaa acaactctgg cgcatcgggc    1140

ttcccataca atcgatagat tgtcgcacct gattgcccga cattatcgcg agcccattta    1200

tacccatata aatcagcatc catgttggaa tttaatcgcg gcctggagca agacgtttcc    1260

cgttgaatat ggctcataac accccttgta ttactgttta tgtaagcaga cagttttatt    1320
```

-continued

```
gttcatgatg atatattttt atcttgtgca atgtaacatc agagattttg agacacaacg    1380
tggctttgtt gaataaatcg aacttttgct gagttgaagg atcagatcat gagtccagtc    1440
caataaaaaa cgcccggcgg caaccgagcg ttctgaacaa atccagatgg agttctgagg    1500
tcattactgg atctatcaac aggagtccaa gcgagctcga tatcaaatta cgccccgccc    1560
tgccactcat cgcagtactg ttgtaattca ttaagcattc tgccgacatg gaagccatca    1620
caaacggcat gatgaacctg aatcgccagc ggcatcagca ccttgtcgcc ttgcgtataa    1680
tatttgccca tggtgaaaac gggggcgaag aagttgtcca tattggccac gtttaaatca    1740
aaactggtga aactcaccca gggattggct gagacgaaaa acatattctc aataaaccct    1800
ttagggaaat aggccaggtt ttcaccgtaa cacgccacat cttgcgaata tatgtgtaga    1860
aactgccgga aatcgtcgtg gtattcactc cagagcgatg aaaacgtttc agtttgctca    1920
tggaaaacgg tgtaacaagg gtgaacacta tcccatatca ccagctcacc gtctttcatt    1980
gccatacgga attccggatg agcattcatc aggcgggcaa gaatgtgaat aaaggccgga    2040
taaaacttgt gcttattttt ctttacggtc tttaaaaagg ccgtaatatc cagctgaacg    2100
gtctggttat aggtacattg agcaactgac tgaaatgcct caaaatgttc tttacgatgc    2160
cattgggata tatcaacggt ggtatatcca gtgatttttt tctccatttt agcttcctta    2220
gctcctgaaa atctcgataa ctcaaaaaat acgcccggta gtgatcttat ttcattatgg    2280
tgaaagttgg aacctcttac gtgccgatca aatcatgacc aaaatccctt aacgtgagtt    2340
ttcgttccac tgagcgtcag accccgtaga aaagatcaaa ggatcttctt gagatccttt    2400
ttttctgcgc gtaatctgct gcttgcaaac aaaaaaacca ccgctaccag cggtggtttg    2460
tttgccggat caagagctac caactctttt tccgaaggta actggcttca gcagagcgca    2520
gataccaaat actgttcttc tagtgtagcc gtagttaggc caccacttca agaactctgt    2580
agcaccgcct acatacctcg ctctgctaat cctgttacca gtggctgctg ccagtggcga    2640
taagtcgtgt cttaccgggt tggactcaag acg                                 2673
```

<210> SEQ ID NO 69
<211> LENGTH: 1749
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pDIK200

<400> SEQUENCE: 69

```
tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac cccgtagaaa      60
agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa     120
aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca actctttttc     180
cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgttcttcta gtgtagccgt     240
agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaatcc     300
tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac     360
gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca     420
gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta tgagaaagcg     480
ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag     540
gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt     600
```

-continued

```
ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcagggggg cggagcctat     660
ggaaaaacgc cagcaacgcg gcctttttac ggttcctggc cttttgctgg ccttttgctc     720
acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtagtcgg tctcaatccg     780
taatcgttaa tccgcaaata acgtaaaaac ccgcttcggc gggttttttt atggggggag     840
tttagggaaa gagcatttgt cagctgtgag accgagacca taaaaaacgc ccggcggcaa     900
ccgagcgttc tgaacaaatc cagatggagt tctgaggtca ttactggatc tatcaacagg     960
agtccaagcg agctcgatat caaattacgc cccgccctgc cactcatcgc agtactgttg    1020
taattcatta agcattctgc cgacatggaa gccatcacaa acggcatgat gaacctgaat    1080
cgccagcggc atcagcacct tgtcgccttg cgtataatat ttgcccatgg tgaaaacggg    1140
ggcgaagaag ttgtccatat tggccacgtt taaatcaaaa ctggtgaaac tcacccaggg    1200
attggctgag acgaaaaaca tattctcaat aaacccttta gggaaatagg ccaggttttc    1260
accgtaacac gccacatctt gcgaatatat gtgtagaaac tgccggaaat cgtcgtggta    1320
ttcactccag agcgatgaaa acgtttcagt ttgctcatgg aaaacggtgt aacaagggtg    1380
aacactatcc catatcacca gctcaccgtc tttcattgcc atacggaatt ccggatgagc    1440
attcatcagg cgggcaagaa tgtgaataaa ggccggataa aacttgtgct tatttttctt    1500
tacggtcttt aaaaaggccg taatatccag ctgaacggtc tggttatagg tacattgagc    1560
aactgactga aatgcctcaa aatgttcttt acgatgccat tgggatatat caacggtggt    1620
atatccagtg atttttttct ccattttagc ttccttagct cctgaaaatc tcgataactc    1680
aaaaaatacg cccggtagtg atcttatttc attatggtga aagttggaac ctcttacgtg    1740
ccgatcaac                                                            1749
```

<210> SEQ ID NO 70
<211> LENGTH: 1795
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pBTK201

<400> SEQUENCE: 70

```
tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac cccgtagaaa       60
agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa      120
aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca actctttttc      180
cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgttcttcta gtgtagccgt      240
agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaatcc      300
tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac      360
gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca      420
gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta tgagaaagcg      480
ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag      540
gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt      600
ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcagggggg cggagcctat      660
ggaaaaacgc cagcaacgcg gcctttttac ggttcctggc cttttgctgg ccttttgctc      720
acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtagtcgg tctcatacac      780
```

continued

```
aaggcatcca atccaacgca aggcccagtc gaccgaatgg gacttccgtt ttatccgttg     840
tttgccgggtg aacgctatcc actcgagtca aactggatca acttcgggtg ggcccctatg     900
cgtttatacc gatgagacca gaccaataaa aaaccccggc cggcaaccga gcgttctgaa     960
caaatccaga tggagttctg aggtcattac tggatctatc aacaggagtc caagcgagct    1020
cgatatcaaa ttacgccccg ccctgccact catcgcagta ctgttgtaat tcattaagca    1080
ttctgccgac atggaagcca tcacaaacgg catgatgaac ctgaatcgcc agcggcatca    1140
gcaccttgtc gccttgcgta taatatttgc ccatggtgaa aacgggggcg aagaagttgt    1200
ccatattggc cacgtttaaa tcaaaactgg tgaaactcac ccagggattg gctgagacga    1260
aaaacatatt ctcaataaac cctttaggga aataggccag gttttcaccg taacacgcca    1320
catcttgcga atatatgtgt agaaactgcc ggaaatcgtc gtggtattca ctccagagcg    1380
atgaaaacgt ttcagtttgc tcatggaaaa cggtgtaaca agggtgaaca ctatcccata    1440
tcaccagctc accgtctttc attgccatac gaaattccgg atgagcattc atcaggcggg    1500
caagaatgtg aataaaggcc ggataaaact tgtgcttatt tttctttacg gtctttaaaa    1560
aggccgtaat atccagctga acggtctggt tataggtaca ttgagcaact gactgaaatg    1620
cctcaaaatg ttctttacga tgccattggg atatatcaac ggtggtatat ccagtgattt    1680
ttttctccat tttagcttcc ttagctcctg aaaatctcga taactcaaaa aatacgcccg    1740
gtagtgatct tatttcatta tggtgaaagt tggaacctct tacgtgcccg atcaa          1795
```

<210> SEQ ID NO 71
<211> LENGTH: 1720
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesised, pDTK005

<400> SEQUENCE: 71

```
ggtcgggctg aacggggggt tcgtgcacac agcccagctt ggagcgaacg acctacaccg      60
aactgagata cctacagcgt gagctatgag aaagcgccac gcttcccgaa gggagaaagg     120
aggacaggta tccggtaagc ggcagggtcg gaacaggaga gcgcacgagg gagcttccag     180
ggggaaacgc ctggtatctt tatagtcctg tcgggtttcg ccacctctga cttgagcgtc     240
gatttttgtg atgctcgtca ggggggcgga gcctatggaa aaacgccagc aacgcggcct     300
ttttacggtt cctggccttt tgctggcctt ttgctcacat gttctttcct gcgttatccc     360
ctgattctgt ggataaccgt attaccgcct ttgagtgagc tgataccgct cgccgcagcc     420
gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga agagcgccca atacgcaaac     480
cgcctctccc cgcgcgttgg ccgattcatt aatgcagctg gcacgacagg tttcccgact     540
ggaaagcggg cagtgagcgc aacgcaatta atgtgagtta gctcactcat taggcaccc      600
actgttgtaa ttcattaagc attctgccga catggaagcc atcacaaacg gcatgatgaa     660
cctgaatcgc cagcggcatc agcaccttgt cgccttgcgt ataatatttg cccatggtga     720
aaacgggggc gaagaagttg tccatattgg ccacgtttaa atcaaaactg gtgaaactca     780
cccagggatt ggctgagacg aaaaacatat tctcaataaa ccctttaggg aaataggcca     840
ggttttcacc gtaacacgcc acatcttgcg aatatatgtg tagaaactgc cggaaatcgt     900
cgtggtattc actccagagc gatgaaaacg tttcagtttg ctcatggaaa acggtgtaac     960
```

-continued

```
aagggtgaac actatcccat atcaccagct caccgtcttt cattgccata cgaaattccg    1020

gatgagcatt catcaggcgg gcaagaatgt gaataaaggc cggataaaac ttgtgcttat    1080

ttttctttac ggtctttaaa aaggccgtaa tatccagctg aacggtctgg ttataggtac    1140

attgagcaac tgactgaaat gcctcaaaat gttctttacg atgccattgg gatatatcaa    1200

cggtggtata tccagtgatt tttttctcca ttttagcttc cttagctcct gaaaatctcg    1260

ataactcaaa aaatacgccc ggtagtgatc ttatttcatt atggtgaaag ttggaacctc    1320

ttacgtgccc gatcaatcat gaccaaaatc ccttaacgtg agttttcgtt ccactgagcg    1380

tcagaccccg tagaaaagat caaaggatct tcttgagatc ctttttttct gcgcgtaatc    1440

tgctgcttgc aaacaaaaaa accaccgcta ccagcggtgg tttgtttgcc ggatcaagag    1500

ctaccaactc tttttccgaa ggtaactggc ttcagcagag cgcagatacc aaatactgtt    1560

cttctagtgt agccgtagtt aggccaccac ttcaagaact ctgtagcacc gcctacatac    1620

ctcgctctgc taatcctgtt accagtggct gctgccagtg gcgataagtc gtgtcttacc    1680

gggttggact caagacgata gttaccggat aaggcgcagc                          1720
```

<210> SEQ ID NO 72
<211> LENGTH: 7536
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized. pBTK327

<400> SEQUENCE: 72

```
ggctggaccg aatgcggggc ttatcctgct gacgctgggc atgctcgatc tgctggaccg      60

aatgctgacc aagcgaaggg aaacgcgtgg aggtcatgcc attggattcc agcaggcagc     120

ccgacggctg ataaccggcg cgggtggtgc gcttgtcctt gcggttggtg aagcccgcca     180

agcgcccata gtggcggctg tcggcgctgg ccgggtcggc gtcgtactcg ctggccagcg     240

tccgggcaat ctgcccccga agttcaccgc ctgcggcgtc ggccaccttg acccatgcct     300

gatagttctt cgggctggtt tccactacca gggcaggctc ccggccctcg gctttcatgt     360

aatccaggtc aaccgagctg aggtcgtcca accgcaaccg aacctgaagc tactgctagg     420

cgggcctgac atacacgtca ttgccctggg cattcatccg cttgagccat ggcgtgttct     480

ggagcacttc ggcggctgac cattcccggt tcatcatctg gccggtggtg gcgtccctga     540

cgccgatatc gaagcgctca cagcccatgg ccttgagctg tcggcctatg gcctgcaaag     600

tcctgtcgtt cttcatcggg ccacccagcg cagccagatc gagccgtcct cggttgtcag     660

tggcgtaccg tagagaacga gaaacgatgc gatcagaagc aaaaccgtag gaatcatggc     720

agccagaatc aaggtaagcc atagcataca gtgccaaacc agcgacgagc taccgggcgt     780

atgcgagggc atgctaaccc aggcggatac ataccgaacc tatttggacc gatccaaacc     840

tgccgccccc gtctggcgcc gggccttcag ccacccggcc gcctgcgccc cgctggcctg     900

attggtatgg ataccgaccc gaagggattc gtaggccaag gtagccatgc tatggcaacg     960

cggttcgatc tgctccgcta actcgttgat gcctctggat ttcttcactc tgtcgattgc    1020

gttcatggtc tattgaatcc aggtatacct gtacgtagcc gatccgggcg ttggccgtgt    1080

cgatgttccg ggccacgtct gccccgtcgg tgcggatgcc ccggccttcc atctccacca    1140

cgttcggccc caggtgaaca ccgggcaggc gctcgatgcc ctgcgcctca agtgctctgt    1200
```

continued

```
ggtcaatgcg gggcaagtgg aacgccagcc ataccgaacg gttggcatgg tagggccatg   1260
aatcgagggc atgccaacgc aatgccttgg gatttgagac ttcggcattc tgtgccaagc   1320
ccttctcggc ggtccttgcg ttgtaccgcc tgaaccactg agcggcgggc cgctcgatgc   1380
cgtcattgac ccgcccggac atcatcaggc ggcagccgcg gttcccgccg ccacccgcat   1440
ggacggccag cgtacacggc aggcgctcgg cacccggtcg gtgcacggcg aactcggacg   1500
ccagcgccta ctgcaggtcg agggacagca cgacccgccg ggcaaattcg aactccttga   1560
aacgccgccc attggcgcgc taatacaggc aggcagcatc aacgcagtcg gagggcagct   1620
agacgaactc aggcatgtgc aaggactcgg agtgcaagcc ttcaccaatg tagcgggact   1680
acttgccttc gcgctggatg tagtcggcct tggccctggc cgattgcccg cccgacctgc   1740
tgccggtttt cgccgtaagg tgataaatcg ccatgctgcc tcgccgttgc ttttgctttt   1800
cggcaccatg caatggccca cggcgagcgc accgcccgaa gggtggccgt taggccagtt   1860
tctcgaagcg aaacccgtaa gtgcgccctc ccctacaacg tagggacggg attgccgccg   1920
atgacgatcc atgacagccc aagacgaagc aaataaaaca tggggcgtcc agatggttac   1980
ggggcgaacc aagggcgcgg atcggctggc aacgccagca gacccaacgc agcgaatacc   2040
tgccgaaatt cagccggagc gggcaaggga acagcagcaa gagccgaaga acgaaacaag   2100
gcgcaaggtc ctggcggggc ccatgatttt ggccaaggtc aacagcagcg agtggccgga   2160
ggacacggct atggcggcca tggacgcgta ccttgaacgc gacccacgac gcgccctgtt   2220
cggactgccg ccacgccaga aggacggacc gggcaggaat atcgacccga acaggccctg   2280
agggggtgca aacgcgaccc aacccatcgg gtacggggca aggccgatcc agcggctacc   2340
agcgctaacg agtacttctg aggggattgg tgtgggattc agcgggcttt gaccgcattt   2400
cccactgccg cgcagcggtg gggcggtgtg tagccagccg cagccaatag accagctatc   2460
cggcctctgg ccgggcatac tgggcaaggg cagcagcgcc ccacaagggc gctgatcaac   2520
gcgcctagtg gattattcta agataatcaa ggatggatta ttccaacacc ccgccagccc   2580
ccgcccctgc tgggattgca ggttaggggg cgtgacagta attgcagggg ttcgagacag   2640
ttacagacgg ggggcgtgac agttactgca ggggacagtg aacgacagtc agggagtgcc   2700
gggcactggc tggcaatgtc tagcaacggc aggcatttcg gctgagggta aaagaacttt   2760
ccgctaaggc atagactgta tgtaaacaca gtattgcaag gacgccgaac atgcctcatg   2820
tggcggccag gacggccagc cgggaacggg atacaggtcg ttacccagac cacccgaccg   2880
agcaaaccca tctcaatcag atcgaagacg agtaaaaccc ggcaaacgct gcgcaaatgg   2940
aagacaaggc aaacgaattg aagggctatg tgcaacggga atttgaagcc tttcacaact   3000
gagggaggca ggagcatggc tttcaagggc ttcgccgaga gtctgaaacc gacgagaacc   3060
tggccgattc aagcaatcca tatcaagtca atatgagtca aattggtatg aacggcaact   3120
gaattcgcat ctagactgat gagacgtggc agagccacaa acagccggta caagcaacga   3180
tatccaggac aatccgaatc atgcgcggca gaccggaacc aacgacgggc atcacagacc   3240
ttaaccccaa gtacagaccc tgcgacaacg tggcaattcg tcgcaatacc acgtaacggg   3300
attccaatcc aatcgattcg agatagtacc agcccgtgcc tgcccgaact tatcgcgagt   3360
tataacccat ttatacccat ataaatcagc atccatgttg gaatttaatc gcgtgagcct   3420
atgcaacggg cttcccatac aatcgattag agattgtcac cgcacctgat tgcccgacat   3480
```

-continued

```
tatcgagagc tattaaactt tttataccct attaattagc atcaatgttg gacttttatc    3540
gagagagcca tacgtaatcg ttattccgca aattacgtta aatcccgatt aggcgggttt    3600
ttttatgggg ggagtttagg gaaagagcat ttgtcagctg gaaatctgct cgtcagtggt    3660
gctcacactg acgaatcatg tacagatcat acggatgact gcctgggcga tcacaactaa    3720
gcaagaacgc cggaaccagc gccggcgaac accacctgca atatggcata ttcacaacgt    3780
ccacgtctca agccgttaca gagactgttc gaacccactg tgcacctgcg tacaccaggc    3840
atcaaatata acgtaaggct aagtcgaacg aatgggaatc tagtattatc tgttgattgt    3900
aggtgaaagc tatctaatag agtcacactg gatcacattc gggtggcact ttatgcgttt    3960
ataccgagcg gccgcgtgtt acaaccaatt aaccaattct gattagaaaa actcatcgag    4020
catcaaatga aactgcaatt tattcatatc aggattatca ataccatatt tttgaaaaag    4080
ccgtttctgt aatgaaggag aaaactcacc gaggcagttc cataggatgg caagatcctg    4140
gtatcggtct gcgattccga ctcgtccaac atcaatacaa cctattaatt tcccctcgtc    4200
aaaaataagg ttatcaagtg agaaatcacc atgagtgacg actgaatccg gtgagaatgg    4260
caaaagctta tgcatttctt tccagacttg ttcaacaggc cagccattac gctcgtcatc    4320
aaaatcactc gcatcaacca aaccgttatt cattcgtgat tgcgcctgag cgagacgaaa    4380
tacgcgatcg ctgttaaaag gacaattaca aacaggaatc gaatgcaacc ggcgcaggaa    4440
cactgccagc gcatcaacaa tattttcacc tgaatcagga tattcttcta atacctggaa    4500
tgctgttttc ccggggatcg cagtggtgag taaccatgca tcatcaggag tacggataaa    4560
atgcttgatg gtcggaagag gcataaattc cgtcagccag tttagtctga ccatctcatc    4620
tgtaacatca ttggcaacgc tacctttgcc atgtttcaga aacaactctg gcgcatcggg    4680
cttcccatac aatcgataga ttgtcgcacc tgattgcccg acattatcgc gagcccattt    4740
atacccatat aaatcagcat ccatgttgga atttaatcgc ggcctcgagc aagacgtttc    4800
ccgttgaata tggctcataa caccccttgt attactgttt atgtaagcag acagttttat    4860
tgttcatgat gatatatttt tatcttgtgc aatgtaacat cagagatttt gagacacaac    4920
gtggctttgt tgaataaatc gaacttttgc tgagttgaag gatcagatca cgcatcttcc    4980
ccttaacgtg agttcagctg ccgccttggg ccgggtgatg tcgtacttgc ccgccgcgaa    5040
ctcggttacc gtccagccca gcgcgaccag ctccggcaac gcctcgcgca cccgcttgcg    5100
gcgcttgcgc atggtcgaac cactggcctc tgacggccag acatagccgc acaaggtatc    5160
tatggaagcc ttgccggttt tgccggggtc gatccagcca cacagccgct ggtgcagcag    5220
gcgggcggtt tagctgtcca gaggccgatc atcgtccatg atgatgagct aatgctggcc    5280
gacaccaatg aaggcatgcg agatcaaggg gttcagggcc aagtacaggc gaacgtaagc    5340
atcgtagctg gagtaatcag aaagcagacg aaacccatgc agctggaggc aattcagggc    5400
gatgatggat accttccaaa ggcgctcgat gcagtcctgt atgtgcttga gcgccccacc    5460
aatatagtcc tatgcaacga tttcctttgc aagcgcaaga tagctaactt tgacccaatg    5520
gcattcagcg gtgacggcct ccccacttgg ttccaggaac agccggagct gccgtcccgc    5580
ttcggtattg ggttcagggc aatgcaatag gactttaggc aatgccatgg aaaccagacc    5640
ttgcaggatg cgcagatcat cagcgcccag cggctccggg ccgctgaact cgatccgctt    5700
gcccgccgcc tagtcatacc tcaccgccag cttgctgcgc ttgcgctcgc cccgcttgag    5760
```

continued

```
ggcacggatc aggccggggg aaagacagtg agcccggtcg tgcccggagt ggctgaggct    5820

gtgcatgttc ttaggattca aaacggggca aaccctgcat attgcgatgc atatccagca    5880

cggcgggctt gagcaccccg ccgtcatgcc gcctgaacca ccgatcagcg aacggtgcgc    5940

catagttggc cttgctcaca ccgaagcgga cgaagaaccg gcgctggtcg tcgtccacac    6000

cccattcctc ggcctcggcg ctggtcatgc tcgacaggta ggactgccag cggatgttat    6060

cgaccagtac cgagctgccc cggctggcct gctgctggtc gcctgcgccc atcatggccg    6120

agccattgca ggcatggtgc aggtacacga tagagcaacc ggtatcgggc gagatggaat    6180

aaatgagacc gatgaaatgg gaaatggggc agatggagta ttatacatcg atgtggaacc    6240

ggcgcagcgc gtccagcacc atcaggcggc ggccctcggc ggcgccgttg aggccgtcga    6300

accactccgg ggccatgatg ttgggcaggc tgccgatcag cggctcgatc agcaggccgt    6360

cagccacggc ttgccgttcc tcggcgctga ggtgcgcccc aagggcgtgc aggcggtgat    6420

gaatggcggt gggcgggtct tcggcgggca ggtcgatcac cgggccggtg ggcagttcgc    6480

aaacctaacg aagataaggc aagcatgcaa tatgatagga aagttgaagg gaaagcatgg    6540

atttaaaggc aaaacagggc gaaaccagcg aaacgaaagt aaacgcaaac atgtagggca    6600

aaacgtagtc cagccgtggc ggcgctgctg cgaacgcctc cagaatattg ataggcttat    6660

gggtagccat tgattgcctc ctttgcaggc agttgcgtgg taggcgctgg cggggtcact    6720

acccccgccc tgcgccgctc tgagttcttc caggcactcg cgcaggcgcc cgtattcgtc    6780

gtcggtcagc cagaacttgc gctgacgcat ccctttggcc ttcatgcgct cggcatatcg    6840

agcataggcg aatgcgtacg ggctggaacg aaggacgacg gtatgcttgt aattaaggtc    6900

tttcatatca gtaacagaga aactagaacg ggacgaaacg attgacttcg aggaacaagg    6960

acaaggtgca gccgtcaagg ttaaggctgg ccatatcagc gactgaaaag cggccagcct    7020

cggccttgtt tgacgtataa ccaaagccac cgggcaacca atagcccttg tcactttgga    7080

tcaggtagac cgaccctgaa gcgctttttt cgtattccat aaaacccccc tctgtgcgtg    7140

agtactcata gtataacagg cgtgagtacc aacgcaagca ctacatgctg aaatctggcc    7200

agcccatgtc aatgcatcgc tggcgggtgc aaggtgaacg tgcaagatcg gaacgcgaac    7260

gctggacgct gggcagaccc atgaccttgc tgacggtgcg ctcgatgtaa tccgcttcgt    7320

ggccgggctt gcgctctgcc agcgctgggc tggcctcggc catggccttg ccgatttcct    7380

cggcactgcg gccccggctg gccagcttct gcgcggcgat aaagtcgcac ttgctgaggt    7440

catgaccgaa gcgcttgacc agcccggcca tctcgctgcg gtactcgtcc agcgccgtgc    7500

gcacgtggcg gataagatgc agctcgggca gttcga                              7536
```

<210> SEQ ID NO 73
<211> LENGTH: 7430
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pRK5CValpha

<400> SEQUENCE: 73

```
ggggcagaaa ggctagtcta agcggaaaca ggacaaggtg aagccgtaaa ggttaaggct      60

ggccatatca gcgactgaaa agcggccagc ctcggccttg tttgacgtat aaccaaagcc     120

accgggcaac caatagccct tgtcactttt gatcagctag accgaccctg aaggcctttt     180
```

-continued

```
ttcgtattcc atcaaacccc attcgtgcgc tgcgcaatca tagtataacc ggcgtcagtc    240
aaaacgaacg aactaaatgc tgaaaaatgg aacgccaatg taaacgaatc gatggcgggg    300
tgccggtgcc cgtgccagcc cggccgcggc aagccggacg ctgggcagac ccatgacctt    360
gctgacggtg cgctcgatgc aatccgcttc gtggccgggc ttgcgctctg ccagcgctgg    420
gctggcctcg gcccaggccc tgccgatttc ctcggcactg cggccccggc tggccagctt    480
ctgcgcggcg atcaagtcgc acttgctgac gtccacgacg aagcgcttga ccagcccggc    540
aatcaagatg aggtaatcga aaagcgaaga gagcaggtgc aggcaaagat gaagcaaggg    600
aagaaagacg atggcaagcc tgcgggactc atccagatgc agatggaact gatcgatatg    660
ctggccagcc tgctgcacca gcgccgggcc agcggcgggc gtcttgccct tggattcacg    720
cagcagcacc cacggctgat aaccggcgcg ggtggcgtgc ttgtccttgc ggttggtgaa    780
gcccgcccag cggccatcgc ggcggctgtc ggcgccggcc gggtcggcgt cgtaccagct    840
ggccagcgtc cgggcaatcc gcccccgaag ttcaccgccc gcggcgtcgg ccaccctgac    900
aaacgaatga tagttattcg ggctggtttc aactacaacg gaacggtaac ggaaccaggc    960
tttcatgtca taaacgtcaa aatagctgac gtagaaaacc agaaccagac aatgccgatc   1020
ctgctcggcg ggcctgatat acaggtcatt gcgctgggca ttcatccgct tgagccatgg   1080
cgtgctctgc agcacttcgc cggctgacca ttcacggttc atcatctggc cggtggtgcc   1140
gtccctgacg ccgaaatcga agcgcatcca gcccaaggcc ttgagctgtc ggcccatggc   1200
ctgcaaagtc ctgtcgttcc tcatcgggcc accaagcgca gcccgatcga gccgacctcg   1260
gttgaaagtg gagtcaggtc gagcaagagc aaagaagaga taagcagaac aaacgaaggc   1320
atcaaggaag aaagcataac ggttagaaca agctaaagaa gaaaaaaaag agaacgcatc   1380
cggccgctcc gcgccgcgcc gctcacctcg gcggctacgc ccggcaactc tttggccagc   1440
tccacccatg ccgcccctgc ctggcgctgg gctttcagcc actcgcgcgc ctgcgcctcg   1500
ctggcctgcc tggtctggcc catgacctgc cgggcctcgc cggccagtgt cgccatgctc   1560
tgggccagcg gttcgatctg ctccgctaac tcgttgatgc ctctggattt cttcactctg   1620
tagaatgcga tactggtata ttgacaaacg gtatacatga aagtcgatga tatgggagtt   1680
ggcggtgtcg atgttcaggg ccacgtctgc ccggtcggtg cggatgcccc ggcgctccat   1740
ctccaccacg ttcggcccca ggtgaacacc gggcaggcgc tcgatgccct gcgcctcaag   1800
tgttcctgtg tcaacgcggg cgtcgcggcc agcccgctcc aatgcccggt tggcacggtc   1860
ggcccatgcc tcgcgggtcc gctcaagcca tgcccagggc ttgagcgctt cggtcctctg   1920
tgcccagccc ttctcagggg tattgcagta gtaccgattg aaacactgag aggcgggaag   1980
atcgatgcag taatagatcc gatcggagaa aatcaggtgg aagtgcgggt tatcgcagcc   2040
aacggaatgg atggcaagcg tataggcaag gagcacggca aaggacaggt gatgggagaa   2100
ctcggacggc agcgccttcc gctggccgag ggtcagctcg acgggcaggg caaattcgac   2160
atccatgaac agccgaaccc tgggcggttc ataaaggtag gaagcataac agtacgaggc   2220
gggccgctcg acgcactccg gcatgcgccc ggattcggcg tgccagcatt catccatgtc   2280
gaggcaatac ttgacttagc gatggatgta gtaggcattg gaaacggaag attggcagac   2340
cgacctgctg ccggctttcg ccgtaaggtg atcaacggcc atgcagcctc gctgccgctt   2400
ttgcctttcg gctccatgca atggccctcg gagagcgcac cgcccgaagg gtggccgtta   2460
```

-continued

```
ggccagtttc tagtagagta aacggtaagt gagccctacc attcaaagtc gggtcgggct    2520
tgccgtagcc gtgcctacct gatagcatcc gagccagacc attcacaatg gggtgtcacg    2580
atggttaagg ggagcaacaa ggcggcggat cggctggcca agctcgaaga acaacgagcg    2640
cgaatcaatg ccgaaattca gcgggagcgg gcaaggaaac agcagcaaga gcgcaagaac    2700
gaaacaaggc gcaaggtgct ggtggggggc atgattttgg ccaaggtgaa cagcagcgag    2760
tggccggagg atcggctcat ggcggcaatg gatgcgtacc ttgaacgcga ccacgaccgc    2820
gacctgttcg gtctgcagcc acgccagctc gatgagcagc gctgaatgct agtccgagcc    2880
aggcccatgc gggctgatca agcgccacca acctacgggc aggggcaatg gacgctaatg    2940
cggctaaaag cgctccagcg tatttctgcg gggtttggtg tggggtttag cgggctttgc    3000
ccgcctttcc ccctgccgcg cagcggtggg gcggtgtgta gcctagcgca gcgaatagac    3060
cagctatccg gcctctggcc gggcatattg ggcaagggca gcagcgcccc acaagggcgc    3120
tgatagccgc gcctagtgga ttatttctta ataatcatgg atggattttt ccaacacccc    3180
gccagcaccc gacccagctg ggtttgcagc tttgcggcgc tgccagttct tgccgcggtt    3240
agtgcaagtc attgcagggc ggcgtgcacg ttatcgcagc ggttcgtgcc agttagtacg    3300
ggagtgacgg gcactggctg gcaatgtcta gcaacggcag gcatttcggc tgagggtaaa    3360
agaactttcc gctaagcgat agactgtatg taaacacagt attgcaagga cgcggaacat    3420
gcctcatgtg gcgggcagga cggccagccg ggatcgggat actggtcgtt acccgagcca    3480
ccgacccgag caaacccttc tctatcagat cgttgacgag tattacccgg cattcgctgc    3540
gattatggca gagcagggta aggtattgcc gggctatgtg aacgcggact ttgtagaatt    3600
tatccaatgc gggcggatgg agcttgcgtt tatacgggtt agctgcgagt attgccaagc    3660
cgagcacctg gtcgctttca gaaatcaatc taaagtatat atgagtaaac ttggtctgac    3720
aggcccctga attcgcatct agactgatga gacgtggtag agcgacaaac agccggtaca    3780
agcaacgatc tcccggacca tctgaatcat gcgcggatga cacgaactca cgacggcgat    3840
cacagcacat aacccacagt acagacactg cgacaacgtg gcaattcgtc gcaatacaac    3900
gttacgggct taccatacca tagtttagtg attgtcatgc aacctgattg accgacattc    3960
tcgcgagtta taacccattt atacccatat aaatcagcat ccatgttgga atttaatcgc    4020
gtgagcctat gtaacgggct tcccatacaa tcgattagag attgtcatcg cacctgattg    4080
cccgacatta tcgcgagtta taacccattt atacccatat aaatcagcat ccatgttgga    4140
atttaatcgc gtgagcccatc cgtaatcgtt aatccgcaaa taacgtaaaa acccgcttcg    4200
gaggttttta ttatgggggc agtttaggga aagaccattc gtcagctgcg tatcaagacg    4260
ttcaagagcc gttacgaacc aaccggaccc aagacaatgc gaggcgatgg agggcaaacc    4320
aagagaggcc gattatttgc agcctaactc ggtgatatcg aatttcaagt agtggaaacc    4380
ttcttccaac tgatctgcgc gcgaggccaa gcgatcttct tcttgtccaa gataagcctg    4440
tatagattca agttcgacgc gatgatactg ggcccgaagc agctccattg aaccgtaggc    4500
agcgacatcc ttcggcgcga ttttgccggt tactgcgctg taccaaatgc gggacaacgt    4560
aagcaatcca tttcgatacc agccagacca gtcgggaggc gagttcaatc gagttaaggt    4620
ttcatttagc gcctcaaata gatcctgttc aggaaccgga tcaaagagtt cctccgccgc    4680
tggacctacc aaggcaacgc tatgttctct tgcttttgtc agcaagatag ccagatcaat    4740
```

-continued

```
gtcgatagtg gatggatcga agatacaagc aagtatgtca ttgcgctgcc attcacaatc   4800
ttgcagttcg agcttagctg gataacgcca aggtatgatg tagtcgtgcc aacaaatggt   4860
gacttctaca gcgcggagaa tctcgctctc tccaggggaa gccgaagttt ccaaaaggtc   4920
gttgatcaaa gctcgcccgc ttgtttcatc aagccttacg gtcaccgtaa ccagcaaatc   4980
aatatcactg tgtggcttca ggccgccatc cactgcggag ccgtacaaat gtacggccag   5040
caacgtcggt tcgagatggc gctcgatgac gccaactacc tctgatagtt gagtcgatac   5100
ttggcgatca ccgcttccc tcataatgtc tacctcgtc ttaggggcc tgccctgatg   5160
agtaacatcg ttgctgatcc atcccatcca actcgaccc aaggcgtacc gagcctgatg   5220
cttggatgcc cgaggcatag actgtaccc aaaaaaacag tcataacaag ccatgaaaac   5280
cgccactgcg ccgttaccac cgctgcgttc ggtcaaggtt ctggaccagt tgcgtgagcg   5340
catacgctac ttgcattaca gcttacgaac cgaacaggct tatgtccact gggttcgtgc   5400
cttcatccgt ttccacggtg tgcgtcaccc ggcaaccttg ggtagcagcg aagtcgaggc   5460
atttctgtcc tggctggtca tgcccaacct acctcaagc gagtcagcct gacgccttgg   5520
gacgggtgcc gtcgcacttg acagcgcgga actcggttac agtcagccc agcgcgacc    5580
gctccggcaa cgcctcgcgc accgcttgc gggcgctgcg catggtcgaa ccactcgcct   5640
ctgacggcca gacataggcc cacaaggtat ctatggaagc cttgccggtt ttgccggggt   5700
cgatccagcc acacagccgc tggtgcagca ggcgggcggt ttcgcctgcc agcgcccgca   5760
cctcgtccat gctgatgcgc acatgctggc cgccacccat gacggcctgc gcgatcaagg   5820
ggttcagggc cacgtacagg agcccgtccg actcgtagct ggcgtactcc gaccgcagcc   5880
gacaccactg cagctgcgg cactcctggg agctgatgga tacccccacc aggcgctagc   5940
tgcagtcctg tatgctgctt agcgcccgac cactatcgac ctctgcccgg atttcctttg   6000
ccagcgcccg atagctaccc ttgaccacat ggcattcagc ggtgacggcc tcccacttgg   6060
gttccaggaa cagccggagc tgccgtccgc cttcggtctt gggttccggg ccaagcacta   6120
ggccattagg cccagcccatg gccacccagcc cttgcaggat gcgcagatca tcagcgccca   6180
gacgctacgc gacgctgccc tagcccagcc tgccgtagcc gtagcccatc gtcccgtccc   6240
gcttgctgcg cttgcgctcg ccccgcttga gggcacggaa caggccgggg gccagacagt   6300
gcgccgggtc gtgccggacg tggccgaggc tgtgcctgtt cttaggcttc accacggggc   6360
acccccttgc tcttgcgctg cctctccagc acggcgggct tgagcacccc gccgtcatgc   6420
cgcctgaacc accgatcagc gaacggtgcg ccatagttgg ccttgctcac accgaagcgg   6480
acgaagaccc ggcgctggtc gtcgtccaca accccatcct aggcctaggc gatggctatg   6540
atcgacaggt aggcctgccca gcggatgtta tagcccagta acgacgctgcc acggcctggcc   6600
tgcctgatgg agcccgagcc aatcatggcc gagcccttgc tggcatggtg aaggaaaacg   6660
atagagcacc cggtatcggc ggcgatggcc tccatgcgac cgatgacctg ggccatgggg   6720
aagctggcgt tttcgtactc gatgcggacc aggcgcagcg tgtccagacc aatcagtagg   6780
cggccctcgg cggcgcgctt gaggccgtcg aacccactccg gggccatgat gttgggcagg   6840
atgccgatca gaggctggct aagcaggccg tacgccaagg attgccgttc atcggcgatg   6900
aggtgcgccc caagggcgtg caggcggtga tgaatggcgg tgggcgggtc ttcggcgggc   6960
aggtagatca ccgggccggt gggcagttcg cccacctcca gcagatccgg cccgcctgca   7020
```

continued

```
atctgtgcgg aaagctgcag ggccagactg gatttacagg aaccacaggg agtaacaagc    7080

gacccgaccg taccggaccc aatgctgggc aaaacgtagt aaagcggtgg aggcgctgct    7140

gcgaacggcc ccagaatatt gataggctta tgggtagcca ttgattgcct cctttgcagg    7200

cagttggtgg ttaggcggtg gcggggtcac taccccggcc ctggcccgct ctgagttctt    7260

ccaggcaatc gcgcagcgcc tcgtattcgt cgtcggtcag ccagaacttg cgctgacgca    7320

tccctttggc cttcatgcgc tcggcatatc gcgcttggcg tacagcgtca gggctggcca    7380

gcaggtagcc ggtctgcttg tacttctggc ctttcatatc agtcaccgcg aacctcgcc     7439
```

<210> SEQ ID NO 74
<211> LENGTH: 9304
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized. pBTK399c

<400> SEQUENCE: 74

```
aagaaaagtg aaacctgcct gatgagacga tgctagtgca aagtgattta atgcgttatt      60

ttgaaattta tgtaaatata atttaatatt gtgtgcggca attcacattt aattaatgct     120

tgttttctta acatcgcggc aaatcaagaa acggcaggtt cggatcttag ctactagaga     180

aagaggagaa atactagatg cgtaaaggcg aagagctgtt cactggtgtc gtccctattc     240

tggtggaact ggatggtgat gtcaacggtc ataagttttc cgtgcgtggc gagggtgaag     300

gtgacgcaac taatggtaaa ctgacgctga agttcatctg tactactggt aaactgccgg     360

ttccttggcc gactctggta acgacgctga cttatggtgt tcagtgcttt gctcgttatc     420

cggaccatat gaagcagcat gacttcttca agtccgccat gccggaaggc tatgtgcagg     480

aacgcacgat ttcctttaag gatgacggca cgtacaaaac gcgtgcggaa gtgaaatttg     540

aaggcgatac cctggtaaac cgcattgagc tgaaaggcat tgactttaaa gaggacggca     600

atatcctggg ccataagctg gaatacaatt ttaacagcca caatgtttac atcaccgccg     660

ataaacaaaa aaatggcatt aaagcgaatt ttaaaattcg ccacaacgtg gaggatggca     720

gcgtgcagct ggctgatcac taccagcaaa acactccaat aggtgatggt cctgttctgc     780

tgccagacaa tcactatctg agcacgcaaa gcgttctgtc taaagatccg aacgagaaac     840

gcgatcatat ggttctgctg gagttcgtaa ccgcagcggg catcacgcat ggtatggatg     900

aactgtacaa atgacctagg ataaaataaa acgaaaggct cagtcgaaag actgggcctt     960

tcgttttatc tgttgtttgt cggtgaacgc tctctactag agtcacactg gctcaccttc    1020

gggtgggcct ttctgcgttt atacgtacgt atacagaaag ggggattaat taacgtgatg    1080

gtgggctgcc attactggtt ggcttggttt aatcagcatt aagcttgccc tactctgtta    1140

agccggaggt agccggaccg aatcgcagcg aaggattacc gttgagcacc gcacggtgcg    1200

aataagggac agtgaagaag gaacacccgc tcgcgggtgg gcctacttca cctatcctgc    1260

acggctgacc aagttggcta accccggcta agtccaaccg aaccctttgg accacaactg    1320

tatatcgtgc gaaaaaggat ggatatcccg aaaaaatcgc tataatgacc ccgaagcagg    1380

gttatgcagc ggaaaaagga attgatacgg aaacgatgcg taggcgctag aggatctgat    1440

gatcagcagt tcaacctgtt gatagtacgt actaagctct catgtttcac gtactaagct    1500

ctcatgttta acgtactaag ctctcatgtt taacgaacta aaccctcatg gctaacgtac    1560
```

-continued

```
tatgctatca tggctaacgt aataagatct aatgcttatc gtactaagct atcctgtttg      1620

aaccatacta ttacctatca tatgcaacta aactagactc tacggtttta agttttatac      1680

gaaaaaaaag aatatataag gcttttaaag cttttaaggt ttaacggttg tggacaacaa      1740

gccagggatg taacgcactg agaagccctt agagcctctc aaagcaattt tgagtgacac      1800

aggaacactt aacggctgac atgggaattc ccctccaccg cggtggcggc cgctctagaa      1860

ctagtggatc ccccgggctg caggaattcg atatcaagct tatcgatacc gtcgacctcg      1920

agggggggcc aggtaccgag gaacgcgtca attaattaag ataggataac gatgccgacc      1980

gacctaaggg agcctagggt gataaggtaa gaggaaaacg tagtaaggcc gtccgcagtc      2040

acggcgctga ccccggatga atgtcagcta ctgggctatc tggacaaggg aaaacgcaag      2100

cgcaaagaga aagcaggtag cttgcagtgg gattatcaaa aaggatcttc acctagatcc      2160

ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg      2220

acagttaccg agcggccgct tatttgccga ctaccttggt gatctcgcct ttcacgtagt      2280

ggacaaattc ttccaactga tctgcgcgcg aggccaagcg atcttcttct tgtccaagat      2340

aagcctgtct agcttcaagt atgacgggct gatactgggc cggcaggcgc tccattgccc      2400

agtcggcagc gacatccttc ggcgcgattt tgccggttac tgcgctgtac caaatgcggg      2460

acaacgtaag cactacattt cgctcatcgc cagcccagtc gggcggcgag ttccatagcg      2520

ttaaggtttc atttagcgcc tcaaatagat cctgttcagg aaccggatca aagagttcct      2580

ccgccgctgg acctaccaag gcaacgctat gttctcttgc ttttgtcagc aagatagcca      2640

gatcaatgtc gatcgtggct ggctcgaaga tacctgcaag aatgtcattg cgctgccatt      2700

ctccaaattg cagttcgcgc ttagctggat aacgccacgg aatgatgtcg tcgtgcacaa      2760

caatggtgac ttctacagcg cggagaatct cgctctctcc aggggaagcc gaagtttcca      2820

aaaggtcgtt gatcaaagct cgccgcgttg tttcatcaag ccttacggtc accgtaacca      2880

gcaaatcaat atcactgtgt ggcttcaggc cgccatccac tgcggagccg tacaaatgta      2940

cggccagcaa cgtcggttcg agatggcgct cgatgacgcc aactacctct gatagttgag      3000

tcgatacttc ggcgatcacc gcttccctca tgatgtttaa ctttgtttta gggcgactgc      3060

cctgctgcgt aacatcgttg ctgctccata acatcaaaca tcgacccacg gcgtaacgcg      3120

cttgctgctt ggatgcccga ggcatagact gtaccccaaa aaaacagtca taacaagcca      3180

tgaaaaccgc cactgcgccg ttaccaccgc tgcgttcggt caaggttctg gaccagttgc      3240

gtgagcgcat acgctacttg cattacagct tacgaaccga acaggcttat gtccactggg      3300

ttcgtgcctt catccgtttc cacggtgtgc gtcacccggc aaccttgggc agcagcgaag      3360

tcgaggcatt tctgtcctgg ctgg                                             3384
```

<210> SEQ ID NO 75
<211> LENGTH: 6001
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized. pBTK301

<400> SEQUENCE: 75

```
actcttcctt tttcaatatt attgaagcat ttatcagggt tattgtctca tgagcggata        60

catatttgaa tgtatttaga aaaataaaca aataggggtt ccgcgcacat ttccccgaaa       120
```

-continued

```
agtgccaccc gtcccgaccc aactccctta aagtgagtca gactgccgcc ttgggcaggg     180
tgcggtagta cttgccagcc gagcactcgg ttcccgtcca gcccagggcg acccgctccg     240
gcaacgcctc gcggaccggc ttgcggcgct tgcgcatggc cgaaccactg gcctcggacg     300
gccagacata gccgcacaag gtatcgatgg aagccctgcc ggttttgccg gggtcgatcc     360
agccacaccg ccgcaggtgc agcagggggg cggtaacgca gtcccagcgc cgcaccacgt     420
ccatgctgat gcgcacatgc tggccgcccc ccatgacggc ctgcgcgatc aaggggttca     480
gggccaagta cagggcgccg tacgcctcgc agctgggata atccgacagc agccgaaccc     540
aatgccgctc gaggccattc tgggcgatga tggccaacta aacaaggagc tagccgaagt     600
cctgcatgtg cttgagcgcc ccacactatc cgacccctgc cccgatttcc tttgccagcg     660
cccgatagcc accttgaccc acatggcatt cagcggtgac ggccccccac ttgggctcca     720
ggaacagccg gagcatgccg ccgccatcgg tcttgggttc cgggccaagc actaggccat     780
taggccgcgc catggcgacc agcccatgca ggatgcgcag atcatcagcg cccagcggct     840
aagggaagcc gaccaagatc agctcgaccg agcggtagtc atacgcaacg taacgcttgc     900
tgcgcttgcg atcgcaacgc ttgacggccc ggcccaggcc ggggccaagc aagtgcgacg     960
ggtcgtgccg gacgcggccg aggccgtgcc tgttcctagg cttcaccacg gggcaccccc    1020
ttgcgcttgc gctgcctctc cagcacggcg ggccggagca cccgccgtcc atgccgcctg    1080
aaccaccgat cagcgaacgg tgcgccatcg ttggccttgc tcacaccgaa gcggacgaag    1140
aaccggcgct ggtcgtcgtc cacaccccat tcctcggcct cggcgctggt catgctcgac    1200
aggcaggccc gacggaggcc gttaccgacc agtaccgagc tgccccggct ggcccgatgc    1260
tggccgaccg agcccatccc ggccgcgccc ttgccggacc ggtgcaggcc aacgatagcg    1320
cacccggtac cggcggcgat ggcccccatg cgaccgatga cctggcccat ggggccgctg    1380
gcgccttcct cctccatgtg gaacggcggc agcgcgtcca gcaccatcag gcggcggccc    1440
tcggcggcgc gcttgaggcc gtcgaaccac tccggggcca tgatgctggg caggcctgcg    1500
atcagcggct ggatcagcag gccgtcagcc acggcctgcc gttcctcggc gctgaggtgc    1560
gacccaaggg agtgcaggcg gtgccgactg gaggccggcg ggtcccaggc gggcaggtag    1620
atcaccgggc cggtggcgag ttcgcccacc tccagcagac ccggcccgcc tgcaacctgt    1680
gcggccagtc gcaggcgcag catggattta ccggcaccac cgggcgacac cagcgcccgg    1740
accgtaccgg ccacccatgt gggcaaaacg tagtccagcg gtggcggcgc tgctgcgaac    1800
gccaccagca tattgatcgg cttatgggta gccattgatt gcctcctttg caggcagttg    1860
gtggctaggc gatggcgggg tacccacccc agccccgagc agctccgagt tatttcaggc    1920
aatcggcccg agcccagtcc tagtcgtcgg tacgccagca attgcggtgc agccccaatt    1980
tggccttacc gagccaggca tatcgcgctc ggcgcaaagc gtccggcatg gaccgcaggt    2040
cgccggtctg cttgcccttc tggtccttca tatcagtcac cgagaaactt gccgggcccg    2100
aacggattgc attcggggca aacggaaccg gtgcagcagc aacggtaacg gatggccatc    2160
tcagcgactg aaaagcggcc agcctcggcc ttgttcgacg tataaccaaa gccaccgggc    2220
aaccaatagc aattgtaccc tttgatacgg tagcccgacc atgcagcgct ttttccgtct    2280
tccataaaac ccccttctgt gcgtgagtac tcatagtata acaggcgtga gtaccaacgc    2340
aagcactaca tgctgaaatc tggcccgccc ctgtccatgc ctcgctgccg gggtgccggt    2400
```

-continued

```
gaccgtgcca gatcggaccg agccagatgg aagcagggca gacccatgcc attgcagacg    2460
gtgcgatcga tgtcatcagc ttcgaggccg ggctagagca atgccagagc tgggcaggcc    2520
tcggccatgg ccttgccgat ttcctcggca ctgcggcccc ggctggccag cttctgcgcg    2580
gcgataaagt cgcacttgct gaggtcatga ccgaagcgct tgacaagccc ggccatctcg    2640
ctgcggtacc cgtccagcgc cgtgcgccgg tggcggtcaa gctgccgctc gggcagttcg    2700
aggctggcca gcctgcgggc cttctcctgc tgccgctggg cctgctcgat ctgctggcca    2760
gcctgatgca accgcgacgg gaccgcggtg gaggcctggc actggaatca accgaagagc    2820
accaaggccc gatcaaaggc gaggctggtg tgctcgtact tgcggctggt gaagcccagc    2880
aagcggccat agtggcggct gtcggcgctg gcgggctcgg cgtcgtactc gctggccagc    2940
gtccgggcaa tctgcccccg aagttcaccg cctgccgcgt cggccacctt gacccatgcc    3000
tgatagttct tgggctggt ttccactacc agggcaggct cccggccctc ggcttctatg    3060
tcatccaggt caaactcgct gaggtcgtcc accagcacca gaccatgccg ctcctgctcg    3120
gagggaatga tatccaagtc attgccatgg accacatacc gattgagacc tggcgcgttc    3180
tggagaaacc aggcggatga aacttcaagg ttaccacatc ggcagctggg tgcgtccatg    3240
acgccgatat cgaagcgctc acagcccatg gcctcgagct gtcggcctat ggcctgcaaa    3300
gtcctgtcgt tcttcatcgg gcgaccaagc gcagccagat cgagccgtcc tcggtcgtca    3360
gtggcgtccg gtcgagcaag agcaacgatg cgatcagcag caccaccgta ggcatcatgg    3420
aagccagcat cacggttagc catagcttcc agtgccaccc ccgcgacgcg ctccgggcgc    3480
tatgcgaggc gatgctaccc taggcggcta actcccgaca atctcggccc agctccaacc    3540
atgccgaacc tgtcaggcgc tgggcattca gaccacaagc agccagagcc tagcaggaat    3600
gcttcgtctg gctcatgaac tgccgggctt cgtcggccag tgtcgccatg ctctggcgca    3660
gcggctcgat ctgctccgct aactcgttga tgccctcgga tttctccacc ctgtcgattg    3720
cgttcatggt ctatcgcctc ccggtattcc tgtaagtcga tgatctgggc gttggcggtg    3780
tcgatgttca gggccacgtc tgcccggtcg gtgcggatgc cccggccttc catctccacc    3840
aagcaaggcc aaaggtgacc aacgggaagg agctcgatgc aatgcgaatc aagtgatatg    3900
tggccaatgc gggcgtcgtg gcgagccgcc tctaatgcac ggttggcatg gtcggcccat    3960
gcctcgcggg tctgctcaag ccatgccttg ggcttgagcg cttcggtctt ctgtgcccag    4020
cccttctccg gggtcttgcc gttgtaccgc ttgaaccgct gagcggcggg ccgctcgatg    4080
ccgtcattga tccgctcgga gatcatcagg tggcagtgcg ggttctcgcc gccaccggca    4140
tggatggcca gagtatacgg caggcgatcg gaccggtcag ggtgcagggc gaccacggac    4200
gacagagccc tatgctggtc gaggctaagc tagaccggca gggcaaattc gaccacattg    4260
aacagaagcc aattggagcg ttccaaacgg taggcagcat aaccgaagtc ggacggaagc    4320
tcgacgaacc ccggcatgtg cccggattcg gcgtgcaaga cttcatccat gtcgcgggca    4380
taccgaactc agcgctggca gtagacggcc ttggccatgg aagaaggccc gaccgaaatg    4440
ctgccggttt tcgccgtaag gtgataaatc gccatgctgc ctcgctgttg ctttcgcttt    4500
taggctaacc gaaccggccc taggagagcg aacggcaaga agggcagacg ttaggcaagt    4560
ttctcgaaga gaaaccggta agtgcgccct ccccfacaaa gtagggtcgg gattgccgcc    4620
gctgcgcctc catgatagcc tacgagacag cacattaaca atggggtgtc aagatggtta    4680
```

-continued

```
aggggagcca aacggaggcg gatcggatgg aaacgctaga agcccaaagc gagcgaatcc    4740
atgccgaccc taacgcgggc agggcaaggg aaccgcagca agacgcgaag aacgaaaccc    4800
ggcgcaaggc gctggtgggg gccatcgatt tggccaaggc gaacagcagc gagtggccgg    4860
aggatcggcc catggcggca atggatggcc accttgaacg cgaccacgac cgcgccttgt    4920
tcggtctgcc gcccacgccg aaggatcgac cgggctgaat gatcgaccga gacaggcccc    4980
gcggggctgc acaccgcgcc ccacccttcg ggtaggggga aaggccgcta aagcggctaa    5040
aagcgatcca gagtatttcc gagggcttta gtgtcggtcc tagcgcgatt tgcccgaatt    5100
tacccatgcc gagcagaggc ggggcggtgc gtcgcctagc gaagcgaatc gaccagatct    5160
ccggcctctg gccggccata ttgggcaagg gcagcagcgc cccacaaggg cgctgataac    5220
cgcgcctagc ggattattcc tagataatca tggatcgatt tttccaacac cccgccagcc    5280
cccgcccctg ctgggtttgc aggttcgggg gcgtgacagc tattgcaggg gttcgcgaca    5340
gttatcgcag gggggcgtga cagttattgc aggggctcgc gacagctagt acgggagtga    5400
agggcaatgc atggcaatgc ataccaacgg aaggcatttc ggctgagggt aaccgaactt    5460
taggctaagc gatagaatgc atgtaaacca agtacggaca ggacggcggc aatgcctact    5520
gtggcggcca ggacggccag ccgggatcgg gatactggtc gttaccagag ccaccgaccc    5580
gagcaaaccc ttctctatca gatcgttgac gagtattacc cggcattcgc tggcgctatg    5640
gcagagcagg gaaaggaatt gccgggctat gtgcaacggg aattcgaaga atttctccaa    5700
tgcggggggc tggagcatgg ctttctacgg gttcgctgcg agtcctgcca cgccgagcac    5760
atggcagcta tacgaatcca atctaaagta tatccgagta aactcggtct gaccggaccc    5820
tgcaatagca tatcgaatga tgcgacgtgg tagcgccaca aacccgcggt aaccgcaacg    5880
atcctcagga ccatctgaat catgccggga tgacacgaac tcacgacggc gatcacagac    5940
attaacccac agtacagaca ctgccacaac gtggcaattc gtcgcaatac aacgctatca    6000
aaaagagtat tgacctaaag tctaacctat aggatactta cagccatcga gagggacacg    6060
gcgaatacag aaacagagga gatattacat atgagtaaag gagaagagct tttcacagga    6120
gttgtcccaa ttcttgttga attagatggt gatgttaatg ggcacaagtt ctctgtcagt    6180
ggagagggtg aaggcgatgc aacatatggc aagctgaccc ttaaatttat ttgcaccacg    6240
ggtaaactac ctgttccatg gccaacactg gtcactacgc tcgggtatgg ggttcagtgc    6300
tttgcgcgct acccagatca catgaaacag cacgactttt tcaagagtgc aatgcccgaa    6360
ggctatgtac aggagagaac catcttcttt aaggatgacg gcaactacaa gacacgtgcc    6420
gaagtcaagt ttgagggtga tacccttgtt aatagaatcg agttaaaagg tattgatttt    6480
aaggaagatg gaaacattct aggacacaaa ctggaatata actataactc acacaatgtg    6540
tacatcatgg cagacaaaca aaagaatggc atcaaggtta acttcaagat tagacacaac    6600
attgaggatg gaagcgttca actagccgac cattaccaac aaaacactcc aattggcgat    6660
ggccctgtgc tgttaccaga caaccattac ctgtccacac aatctgccct ttcgaaagat    6720
cccaacgaaa agcgcgacca catggtcctt cttgagtttg taacggctgc tgggattaca    6780
cacggcatgg atgaactata caaataactc agtaatagtc aatcggaacc taacgtaacc    6840
acccgcttcg gcgggttttt ttatgggggg agtttaggga aagagcattt gtcagctgga    6900
aatctgctgg tcagtggtgc tcacactgac gaatcatgta cagatcatac cgatgactgc    6960
```

-continued

```
atggcgactc aaacctaagc aagtcagccg gatccagagc aggcgaaatc cactgcatct    7020

atggcatatc aaaccagtcc aagtcaaacg aagttaaaga gatgttaagt aaacctagtg    7080

cactgcagta cacaaaaaac cactcaagac ccgtttagag gccacaaggg gttatgctaa    7140

tgctgcttta agaaccactt tcacatttaa gttgtttttc taatccgcat atgatcaatt    7200

caaggccgaa taagaaggct ggctctgcac cttggtgatc aaataattcg atagcttgtc    7260

gtaataatgg cggcatacta ctcagtagta ggtgttccct ttcttcttta gcgacttgat    7320

gatcctgatc ttccaatacg caacctaaag taaaatgccc cactgcgctg agtgcatatc    7380

atgcattctc tagtgaaaaa cattgcttgc ataaaaaggc taattgattt tagagagttt    7440

catactgttt ttctgtaggc cgtgtacgta aatgtacttt tgctccatcg cgatgactta    7500

gtaaagcaca tctaaaactt ttagcgttat tacgtaaaaa atcttgccag ctttcccctt    7560

ctaaagggca aaagtgagta tggtgcctat ctaacatctc aatggctaag gcgtcgagca    7620

aagcccgctt attttttaca tgccaataca atgtaggctg ctctacacct agcttctggg    7680

cgagtttacg ggttgttaaa ccttcgattc cgacctcatt aagcagctct aatgcgctgt    7740

taatcacttt acttttatct aatctagaca tcattaattc ctaatttttg ttgacacttt    7800

tagatctctc ccctatagtg agtcgtatta atttccgagc ggccgcgatt atcaaaaagg    7860

atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta aagtatatat    7920

gagtaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc    7980

tgtctatttc gttcatccat agttgcctga ctccccgtcg tgtagataac tacgatacgg    8040

gagggcttac catctggccc cagtgctgca atgataccgc gagacccacg ctcaccggct    8100

ccagatttat cagcaataaa ccagccagcc ggaagggccg agcgcagaag tggtcctgca    8160

actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt aagtagttcg    8220

ccagttaata gtttgcgcaa cgttgttgcc attgctacag gcatcgtggt gtcacgctcg    8280

tcgtttggta tggcttcatt cagctccggt tcccaacgat caaggcgagt tacatgatcc    8340

cccatgttgt gcaaaaaagc ggttagctcc ttcggtcctc cgatcgttgt cagaagtaag    8400

ttggccgcag tgttatcact catggttatg gcagcactgc ataattctct tactgtcatg    8460

ccatccgtaa gatgcttttc tgtgactggt gagtactcaa ccaagtcatt ctgagaatag    8520

tgtatgcggc gaccgagttg ctcttgcccg gcgtcaatac gggataatac cgcgccacat    8580

agcagaactt taaaagtgct catcattgga aaacgttctt cggggcgaaa actctcaagg    8640

atcttaccgc tgttgagatc cagttcgatg taacccactc gtgcacccaa ctgatcttca    8700

gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca aaatgccgca    8760

aaaaagggaa taagggcgac acggaaatgt tgaatactca t                         8801
```

<210> SEQ ID NO 76
<211> LENGTH: 6790
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pBTK303

<400> SEQUENCE: 76

```
actcttcctt tttcaatatt attgaagcat ttatcagggt tattgtctca tgagcggata      60

catatttgaa tgtatttaga aaaataaaca aataggggtt ccgcgcacat ttccccgaaa     120
```

-continued

```
agtgccaccc gtcccgaccc aactccctta aagtgagtca gactgccgcc ttgggccggg     180
tgcagtagta attgccagcc gagcactcgg ttaccgtcca gacccgagcg aaccgctacg     240
gcaacgcctc gcgcaccggc ttgccgcgct tgcgcatggt cgaaccactg gcctctgacg     300
gccagacata gccgcacaag gtatctatgg aagccttgcc ggtttcgccg gggtcgatcc     360
agccacaccg ccgcgggtgc agccggcggg cggttccgcc gtcccagcgc cgcacccgct     420
ccatgctgat gcgcacatgc tggccgccac ccatgacggc ctgcgcgatc aaggggttca     480
gggccaagta aaggcgaccg taagcctagc agctggcgta atccgaaagc agccgaaacc     540
aatgccgcta gaggcaattc tgggcgatga tggctaacta aaccaggagc tagccgaagt     600
cctgtatgtg cttgagcgcc ccaccactat cgacctctgc ccggatttcc tttgccagcg     660
cccgatagct acctttgacc acatggcatt cagcggtgac ggcctcccac ttgggctcca     720
ggaacagccg gagctgccgt ccgcctccgg tcttgggttc cgggccaagc actaggccat     780
taggccccgc catggccacc agccctgcca ggatgcgcag atcatcagcg ccccagcggt     840
aaggcaagcc gacccagctc agctcgacgc agccgcagtc atccgcaacg taccgcttgc     900
tgcgcttgcg atcgcaacgc ttgcccgacc ggcccaggcc ggggccaagc aagtgcgacg     960
ggtcgtgccg gacgcggctg aggcggtgct tgttcctagg cttcaccacg gggcaccccc    1020
ttgctcttgc gctgcctctc cagcacggcg ggcctgagca cccggccgtc atgccgcctg    1080
aaccaccgat cagcgaacgg tgcgccatcg ttggccttgc tcacaccgaa gcggacgaag    1140
aacggcgcct ggtcgtcgtc cacaccccat tcctcggcct cggcgctggt catgctcgac    1200
aggcaggcac gaccgaggca gttacgcacc agtaccgagc tgccccggct ggcccgctgc    1260
tggccgactg agcccatcac ggccgcgccc ttgccggcac ggtgcaggca caagatagcg    1320
caccggtact cggcggcgat ggcctccatg cgaccgatga cctgggccat ggggccgctg    1380
gcgctttctt cctcgatgtg gaaccggcgc agcgcgtcca gcaccatcag gcggccggcc    1440
tcggcggcgc gcttgaggcc gtcgaacccc tccggggcca tgatgctggg caggctgccg    1500
atcagcggct ggatcagcag gccgtcagcc acggcttgcc gttcctcggc gctgaggtgc    1560
gacccaaggg agtgcaggcg gtgacgactg gaggcgggcg ggtcccaggc gggcacgtcg    1620
atcaccgggc cggtggcgag ttcgcccacc tccagcagat ccggcccgcc tgcaatctgt    1680
gcggccagtt gcaggcgcag catggattta ccggcaccac cgggcgacac cagcgcccgg    1740
accgtaccgg ccacccatgt ggcgcaaacg tagtccagcg gtggcggcgc tgctgcgaac    1800
gcctccagca tattgatcgg cttatgggta gccattgatt gcctcctttg caggcagttg    1860
gtgcctaggc gatgcagggc taccaaaccc agccctgagc agctctgagt tattccaggc    1920
aatcgagccg agcctagtcc tagtcgtcgg taagccagca attgcgctgc agcctcaatt    1980
tggccttacc gagccaggca tatcgcgctc ggcgcaaagc gtacggcatg gaccgcaggt    2040
cgccggtctg cttgcccttt tggtcctcca tatcagtcac cgagaaactt gccggggccg    2100
aacgcttgca attcgaggca aacggaaacg gtgcagaagc aacggtaacg gatggcaatc    2160
tcagcgactg aaaagcggcc agcctcggcc ttgttcgacg tataaccaaa gccaccgggc    2220
aaccaatcgc aattgtaccc tttgatcagg tagcccgacc atgcagagct ttttccgtct    2280
tccataaaac ccccttctgt gcgtgagtac tcatagtata acaggcgtga gtacccacgc    2340
aagcactaca tgctgaaatc tggcccgccc ctgtccatgc ctcgctggcg gggtgccggt    2400
```

continued

```
gaccgtgcca gatcggaccg agcaagatgg aagcaggcca gacccatgac attgcagacg    2460
gtgcgatcga tgtaatcagc ttcgacggcg ggctagagca atgccagagc tgggcaggcc    2520
tcggccatgg ccttgccgat ttcctcggca ctgcggcccc ggctggccag cttctgcgcg    2580
gcgataaagt cgcacttgct gaggtcatga ccgaagcgct tgagcagccc ggccatctcg    2640
ctgcggtacc cgtccagcgc cgtgcgccgg tggcggctca gctgccgctc gggcagttcg    2700
aggcaggcca gcctgcgggc cttctcctgc tgccgctggg cctgctcgat ctgctggcca    2760
gcctgatgca aacgcgacgg gcaagcgggt gaggtcttgc aattggattc aagcagaagc    2820
aacccaggct gataaacggc gaggtaggtg tgctagtacc tgcggtaggt gacgcccagc    2880
aagcggcgat agtggcggct gtcggcgctg gcgggctcgg cgtcgtactc gctggccagc    2940
gtccgggcaa tctgcccccg aagttcaccg cctgcggcgt cggccacctt gacccatgcc    3000
tgatagttct tcgggctggt ttccactacc agggcaggct cccggccctc ggcttccatg    3060
tcatccaggt caaactcgct gaggtcgtcc accagcacca gaccatgccg ctcctgctcg    3120
gagggaatga tatacaagtc attgccatgg gattacatcc gattgagact tggcagtgtc    3180
tggagaactc aggcagatga aactacaagg ttactcatac ggcaggtggg tgcgtcaatg    3240
acgccgatat cgaagcgctc acagcccatg gccttgagct gtcggcctat ggcctgcaaa    3300
gtcctgtcgt tcttcatcgg gccaccaagc gcagccagat cgagccgtcc tcggttgtca    3360
gtggcgtccg gtcgagcaag agcaacgatg cgatcagcag caccaccgta ggcatcatgg    3420
aagccagcat cacggttagc catagcttcc agtgccaccc ccgcgacgcg ctccgggcgc    3480
tatgcgaggc gatgctatcc taggcagcta aatcccgata atctaggccc agctccaacc    3540
atgccgaccc tgtcaggcgc tgggcattca gcaacaaagc agcctgagcc tagcaggact    3600
gcttggtctg gctcatgacc tgcccggctt cgtcggccag tgtcgccatg ctctgggcca    3660
gcggttcgat ctgctccgct aactcgttga tgcctctgga tttcttcact ctgtcgattg    3720
cgttcatggt ctattgcctc ccggtattcc tgtaagtcga tgatctgggc gttggcgctg    3780
tcgatgttca gggccacgtc tgccccgtcg gtgcggatgc cccggccttc catctccacc    3840
aagataggcc aaacggtgac aacgggaagg agctcgatgc aatgcgaatc aagtgatatg    3900
tggtcaatgc gggcgtcgtg gccagccggc tctaatgccc ggttggcatg gtcggcccat    3960
gcctcgcggg tctgctcaag ccatgccttg ggcttgagcg cttcggtctt ctgtgccccg    4020
cccttctccg gggtcttgcc gttgtaccgc ttgaaccact gagcggcggg ccgcacgatg    4080
ccgtcattga tccgctcgga gatcatcagg tggcagtgcg ggttctcgcc gcccaccgca    4140
tggatggcca gagtatacgg aaggcgatcg gatacggtca ggtgctgggc gatctcggac    4200
gacagagcct tatgctggtc gagggtaagc tagaccggca gggcaaattc gacctcattg    4260
aacagaagcc aattggagcg ttactaacgg taggcagatt aaacgtagtc ggcgggaagc    4320
tcgacgaact ccggcatgtg cccggattcg gcgtgcaaga cttcatccat gtcgcgggca    4380
tacctgactc agcgctggtt gtagtcggcc ttggccatgg aagattggcc gacgcaactg    4440
ctgccggttt tcgccgtaag gtgatcaatc gccatgctgc ctcgctgttg ctttggcttt    4500
taggctacct gacttggccc taggagagcg aacggcaaga aggggggacg ttaggcaagt    4560
ttctcgaaga gaaaccggta agtgcgccct ccccgacaaa gtagggtcgg gattgccgcc    4620
gctgtgcctc catgatagcc tacgagacag cacattaaca atgggctgtc aagatcgtta    4680
```

-continued

```
aggggagcca aacggaggcg gatcggatgg aaaagcagga agcacaagca gagcgaatcc    4740
atgccgaaac taacgcgggc agggcaaggg aaccgcagca agagcgcaag aacgaaaaac    4800
gggcgcaagg gctggctggg gccacgattt tggccaaggc gaacagcagc gagtggccgg    4860
aggatcggcc catggcggca atggatggcc accttgaacg cgaccacgac cgcgccttgt    4920
tcggtctgcc gcccacgccg aaggatggac cgggctgaat gatcgaccga gacaggccct    4980
gcggggctgc acacgcgccc ccacccttcg ggtagggggа aaggccgcta aagcggctaa    5040
aagcgatcca gagtatttca gaggggtttg gtgtacggta tagcggcatt tgcccgcatt    5100
tacccatgcc gagcagaggc ggggcggtgc gtagccatgc gaacgcaatc gacccagatt    5160
ccggcctctg gccggcgcata ttgggcaagg gcagcagcgc cccacaaggg cgctgataac    5220
cgcgcctagt ggattattct tagataatca tggatggatt tttcaacacc ccggccagcc    5280
cccgcccctg ctgggtttgc aggtacgggg gcgtgacagt tattgcaggg gttcgtgacc    5340
gttatcgcag ggggggcgtga cagtcattgc aggggtcgtc gaccgtcagt acggggagtga    5400
agggcaatgg atggcaatgc atagcaacgg aaggcatttc ggctgagggt aaacgaactt    5460
taagctaagc gatagaatgc atgtaaaccc agtaacgaca ggacagcggc aatgcctaat    5520
gtggcggcca ggaccgccag ccggggatcgg gatactggtc gttaccagag ccaccgaccc    5580
gagcaaaccc ttctctatca gatcgttgac gagtattacc cggcattcgc tgcgcttatg    5640
gcagagcagg gaacggaact gccgggctat gtgcaacggg aattcgaagc atttctccaa    5700
tgcgggcggc tggagcatgg ctttctacgg gttcgctgcg agtcctgcca cgccgagcac    5760
atggacagct tacgaaatca atctaagta tatacgagta aactagtat gacagacaac    5820
tgacatagca tatagaatga tgacgactgg tagacgacca aacagcaggt aaacgcaacg    5880
atcacagga ccatctgaat catgccgcgga tgacacgaac tcaccacggc gatcacagac    5940
attaacccac agtacagaca ctgccacaac gtggcaattc gtcgcaatac aacgccttgg    6000
cagcctattc ttgacatgta gtgaggggc tggtatcaatc acatagtact gttatacaga    6060
aacagaggag atatcacata tgagtaaagg agaagagctt ttcacaggag ttgtcccaat    6120
aatcgtggca ttacgaggtg atgtaaatgg gaccaagttc tatgacagtg gagaggtgac    6180
aggcgacgca acatatggca agctgaccct taaatttatt tgcaccacgg gtaaactacc    6240
tgttccatgg ccaacactgg tcactacgtt cgggtatggg gttcagtgct ttgcgcgcta    6300
cccagatcac atgaaacagc acgacttttt caagagtgca atgcccgaag gctatgtaca    6360
ggagagaacc atcttttta aggatgacgg caactataag acacgcgccg aagtgaagtt    6420
agaggatgca aacccatgtta atagaatcga gttaaagggc attgacttta aggaagatgg    6480
aaacaatttta ggccaaaacc tggaatataa ctataactcc aataacgtgt aactaatggc    6540
agacaagcaa aagaacggta tcaaggttaa cttcaagatc agacacaacc ttgaggatgg    6600
aagcgttcaa ctagccgacc attaccaaca aaacaccca attggcgatg ggccgtgct    6660
gttacaagac aaccattacc tgtccactca atctgcacta tagaaagatc aaacgaaac    6720
gcgcgaccac atggtccttc ttgagttgt caccgctgct gggattacac acggcatgga    6780
tgaactatac aaataaatcc gtaacgtta atccgcaac aaagtaaac aaagcctagg    6840
cgggtttttt tatgggggga gtttagggaa agagcatttg tcagcatgaa atctgctcgt    6900
cagtcgtgct cacactgacg aatcatgtac agatcatacc gatgactgcc tggcgactca    6960
```

-continued

```
aaactaagca agtcagacgg aaacaagagc gggcaaaacc aatgcatatt tggcaatatt      7020

aaccagtaca agtcaaacgc agttacagtg atgttaagta aaactagtga aatgcagtac      7080

acaaaaaacc cctcaagacc cgtttagagg ccccaagggg ttatgctaat gctgctttaa      7140

gacccacttt cacatttaag ttgtttttct aatccgcata tgatcaattc aaggccgaat      7200

aagaaggctg gctctgcacc ttggtgatca aataattcga tagcttgtcg taataatggc      7260

ggcatactat cagtagtagg tgtttccctt tcttctttag cgacttgatg ctcttgatct      7320

tccaatacgc aacctaaagt aaaatgcccc aatgcgatga gtgcaatatc tgcctttatt      7380

agtgaaaatc attgttggca taataaggct aattgattta agtgagtttc attcagtttt      7440

tctgtaggcc gtgtacctaa atgtactttt gctccatcgc gatgacttag taaagcacat      7500

ctaaaacttt tagcgttatt acgtaaaaaa tcttgccagc tttccccttc taaagggcaa      7560

aagtgagtat ggtgcctatc taacatctca atggctaagg cgtcgagcaa agcccgctta      7620

ttttttacat gccaatacaa tgtaggctgc tctacaccta gcttctgggc gagtttacgg      7680

gttgttaaac attgcattcc gacctcatta agcagctcta atgcgctgtt aatcacttta      7740

attttatcta aacgagaccc gagtatatcc aattctaaaa aattatttat agatctatcc      7800

cctatagtga gtcgtattaa tttcgcgagc gccgcgatta tcaaaaagga tcttcaccta      7860

gatcctttta aattaaaaat gaagttttaa atcaatctaa agtatatatg agtaaacttg      7920

gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg      7980

ttcatccata gttgcctgac tccccgtcgt gtagataact acgatacggg agggcttacc      8040

atctggcccc agtgctgcaa tgataccgcg agacccacgc tcaccggctc cagatttatc      8100

agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc      8160

ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag      8220

tttgcgcaac gttgttgcca ttgctacagg catcgtggtg tcacgctcgt cgtttggtat      8280

ggcttcattc agctccggtt cccaacgatc aaggcgagtt acatgatccc ccatgttgtg      8340

caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc agaagtaagt tggccgcagt      8400

gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag      8460

atgcttttct gtgactggtg agtactcaac caagtcattc tgagaatagt gtatgcggcg      8520

accgagttgc tcttgcccgg cgtcaatacg ggataatacc gcgccacata gcagaacttt      8580

aaaagtgctc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct      8640

gttgagatcc agttcgatgt aacccactcg tgcacccaac tgatcttcag catcttttac      8700

tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat      8760

aagggcgaca cggaaatgtt gaatactcat                                       8790
```

<210> SEQ ID NO 77
<211> LENGTH: 6002
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized. pBTK309

<400> SEQUENCE: 77

```
actcttcctt tttcaatatt attgaagcat ttatcagggt tattgtctca tgagcggata        60

catatttgaa tgtatttaga aaaataaaca aataggggtt ccgcgcacat ttccccgaaa       120
```

-continued

```
agtgccaccc gtcccgaccc aactccatta aagtgagtca gactgccgcc ttgggccggg     180
tgccgtagta cttgccagcc gagcactcgg ttcccgtcca gcccagagcg acccgctacg     240
gcaacgcctc gcgcacccgc ttgcggcgcc tgcgcatggc cgaaccactg gcctccgacg     300
gccagacata gccgcacaag gtatcctatg aagccctgcc ggtttcgccg gggtcgatcc     360
agccacaccg ccgcaggtgc agcaggcggg cggttccgca gtcccagcgc cgcacctcgt     420
ccatgctgat gcgcacatgc tggccgccac ccatgacggc ctgcgcgatc aaggggttca     480
gggccacgta caggcgaccg taggcctcgc agctggcgta ctcggacagc agccgaaacc     540
actgccgctc gaggccattc tgggcgatga tggcccactc acccaggcgc tagccgcagt     600
cctgcatgtg cttgagcgcc ccaccactac cgacccctgc cccgatttcc tttgccagcg     660
cccgatagcc acctttgacc acatggcatt cagccgtgac ggcctcccac ttgggctcca     720
ggcacagccg gagcagccgc ccgccctcgg tcttgggttc cgggccaagc actaggccat     780
taggcccagc catggccacc agcccctgca ggatgcgcag atcatcagcg ccccagcggt     840
caggccagcc gacccagctc agctcgacgc agccgcagtc atccgccacg tacccgctgc     900
tgcgctgcgc atcgccacgc ttgacggccc ggcccaggcc ggggccaagc cagtgcgcag     960
ggtcgtgccg gacgcggctg aggccgtgcc tgttcctagg cttcaccacg gggcaccccc    1020
ttgctcttgc gctgcctctc cagcacggcg ggcctgagca cccgccgtcc atgcagcctg    1080
aaccaccgac cagcgaacgg tgcgccatcg ttggccttgc tcacaccgaa gcggacgaag    1140
aaccggcgca ggtcgtcgtc cacaccccat tcctcggcca cggcgctggt catgcacgac    1200
aggcaggccc gacggaggcc gttcccgccc agtcccgagc tgccccggct ggcccgctgc    1260
tggccgcctg agcccatccc ggccgcgccc ttgccggccc ggtgcaggcc cacgacagcg    1320
cagccggtac cggcggcgac ggcctccatg cgaccgatga cctggcccat ggggccgctg    1380
gcgccttctc cctcgatgtg gaaccggcgc agcgcgtcca gcaccatcag gcggccgccc    1440
tcggcggcgc gcttgaggcc gtcgaaccac tccggggcca tgatgctggg caggcagccg    1500
atcagcggca ggatcagcag gccgacagcc acggcatgcc gttcctcggc gctgaggtgc    1560
gaccccaggg agtgcaggcg gtgccgcctg gaggccggcg ggtcccaggc gggcaggtcg    1620
atcaccgggc cggtcggcag ttcgcccacc tccagcagac ccggcccgcc tgcaatctgt    1680
gcggccagtt gcaggcgcag catggattta ccggcaccac cgggcgacac cagcgcccgg    1740
accgcaccgg ccacccatgt gggcaaaacg tagtccagcg gtggcggcgc tgctgcgaac    1800
gcccccagca tattgatagg cttacgggta gccatcgatc gcctcctttg caggcagttg    1860
gtggctaggc gatggcgggg taccccaccc agccccgagc agctccgagt tattccaggc    1920
aatcgagccg agcccagtcc tagtcgtcgg tacgccagca attgcgctgc agcccccatt    1980
tggccttccc gagccaggca tatcgcgctc ggcgccaagc gtccggcatg gacccgcagt    2040
cgccggtctg cttgcccttt tggtcctcca tatcagtcac cgagaaactt gccggggccg    2100
aaccgcttgc cttcgcggca caacgcaccg gtgcagcacg caacgctacg gatggccatc    2160
tcagcgactg aaaagcggcc agcctcggcc ttgttcgacg tataaccaaa gccaccgggc    2220
aaccaatcgc aattgtaccc tttgatacgg tagcccgacc atgcagagct ttttccgtct    2280
tccatcaaac ccccttctgt gcgtgagtac tcatagtata acaggcgtga gtaccaacgc    2340
aagcactaca tgctgaaatc tggcccgccc ctgtccatgc ctcgctggcg gggtgccggt    2400
```

-continued

```
gaccgtgcca gatcggaccg agcgagatgg aagcgcggca gacccatgcc attgccgacg    2460
gtgcgctcga tgtcatcagc ttcgccgccg ggctcgcgcc atgccagcgc tgggccggcc    2520
tcggccatgg ccttgccgat ttcctcggca ctgcggcccc ggctggccag cttctgcgcg    2580
gcgatcaagc cgcacttgct gaggtcatga ccgaagcgct tgacgagccc ggccatctcg    2640
ctgcggtacc cgtccagcgc cgtgcgccgg tggcggtcaa gctgccgctc gggcagttcg    2700
aggcaggcca gcctgcgggc cttcacctgc tgccgctggg cctgcacgat ctgcaggcca    2760
gacggatgcc aacgcgaccg gaccgcggtg gaggccttgc actttggatt cagcagcagc    2820
aaccaaggcc gatcaaaggc gaggccggtg tgcttgtacc tgcggctggt gacgccagcc    2880
aagcggcgat agtggcggct gtcggcgctg gcgggctcgg cgtcgcactc gctggccagc    2940
gtcgggcgaa tctgccccgg aagttcaccg cctgcggcgt cggcgacctt gacccatgcc    3000
tgatagttct tcgggctggt ttcccactcc aggggcaggt ccgggccctc ggcttccatg    3060
tcatcgaggt caaactcgct gaggtcgtcc accagcacca gaccatgccg ctcctgctcg    3120
gaggcaatga tatccaagtc attgccatgc gattccatcc gattgagccc tggcgcgttc    3180
tggagcaact cggccgatga aactccacgc ttcaccatcc ggcccgtggg tgcgccaatg    3240
acgccgatac cgaagcgctc acagcccatg gcgttgagcc gtcggcctat ggcctgcaaa    3300
gtcgcgtcgc tcttcatcgc gccacccagc gcagccagat cgagccgtcc tcggccgtca    3360
gtggcgtcag gtcgagcaag agcaacgatg cgatcagcag caccaccgta ggcatcatgg    3420
aagccagcat cacggttagc catagcttcc agtgccaccc ccgcgacgcg ctccgggcgc    3480
tatgcgaggc gatgctaccc taggcggcta actcccgaca atctcggccc agctccaacc    3540
atgccgaccc tgtccggcgc tgggcattca gacccaaagc agcccgagcc tagccgcact    3600
gcttcgtctg gctcatgacc tgccgggctt cgtcggccag tgtcgccatg ctctggcgca    3660
gcggctcgat ctgctccgct aactcgttga tgcctctgga tttcttcact ctgtcgattg    3720
cgttcatggt ctattgcctc ccggtattcc tgtaagtcga tgatcagggc gttggcggtg    3780
tcgatgttca gggccacgtc tgccccgtcg gtgcggatgc cccggccttc catctccacc    3840
aagcccggcc aaccgtgacc aacggcaacg agctcgatgc aatgcgcatc aagtgctatg    3900
tggtcaatgc gggcgtcgtg gcgagccggc tctaatgcgc ggttggcatg gtcggcccat    3960
gcgtcgcggg tctgctcaag ccatgccttg ggcttgagcg cttcggtctt ctgtgcccgg    4020
cccttctccg gggtcttgcc gttgtaccgc ttgaaccact gagcggcggg ccgcacgatg    4080
ccgtcattga tccgctcgga gatcatcagg tggcagtgcg ggttcacgcc gcccacggca    4140
tggatggcca gagtatacgg aaggcgatcg gaaccggtca ggtgcaggga gaaccgggac    4200
gacagagccc tatgctggtc gagggcaagc tagaccggca gggcaaattc gacccacttg    4260
aacagaagcc aattggagcg ttcacaaagg taggcagacc aaacgcagtc ggcgggaagc    4320
tcgacgaacc ccggcatgtg cccggattcg gcgtgcaaga cttcatccat gtcgcgggca    4380
tacccgactc agcgctggcc gtagccggcc ttggccatgg aagcccggcc gaccgaaatg    4440
ctgccggttt tcgccgtaag gtgataaatc gccatgctgc ctcgctgttg cttttgcttt    4500
taggctaccc gacccggccc taggagagcg aacagcaaga aggggcgacg ttaggcaagt    4560
ttcacgaaga gaaaccggta agtgcgccct cccctacaaa gtagggtcgg gattgccgcc    4620
gctgcgcctc catgatagcc tacgagacag cacattaaca atgggcgtgc aagatcgtta    4680
```

-continued

```
aggggagcta caaggaggcg gatcggatgg caacgctaga agcccaacgc gagcggatcc    4740
atgccgaccc taagcgggag agggcaaggg aaccgcagca agagcggaag aacgaaaacc    4800
gggcaaggc gctggtgggg gccatgattc tggccaaggc gaacagcagc gagtggccgg    4860
aggatcggcc catggcggca atggatggcc accttgaacc cgaccacgac cgcgccttgt    4920
tcggtctgcc gcccacgccg aaggatgagc cgggctgaat gatcgaccga gacaggccct    4980
gcggggctgc acacgcgccc cccacccttc ggtagggggg aaggccgcta aagcggctaa    5040
aagcgatcca gagtatttcc gagggctttg gtgtcggtc tagccggatt tgcccgactt    5100
tacccatgcc gagcagaggc ggggcggtgc gtagcctagc gacgcgaatc gaccagatct    5160
ccgcccgtcg gccggcgata ttggccaagc gcagcagcgc cccacaaggg cgctgataac    5220
cgcgcctagc ggattattcc tagataatca tggatcgatt tttccaacac cccgccagcc    5280
cccgcccctg ctgggtttgc aggttcgggg gcgtgacagt tattgcaggg gttcgtgaca    5340
gttattgcag gggggcgtga cagttattgc aggggttcgt gacagttagt acgggagtga    5400
agggcaatgc atggcaatgc atagcaacgc aaggcatttc ggctgagggt aaccgaactt    5460
taagctaagc gatagaatgc atgtaaaccc agtacagaca ggacgcgggc aatgcctact    5520
gtggcggcca ggacggccag ccgggatcgg gatactggtc gttaccagag ccaccgaccc    5580
gagcaaaccc ttctctatca gatcgttgac gagtattacc cggcattcgc tgcgcttatg    5640
gcagagcagg gaacggaatt gccgggctat gtgcaacggg aattcgaagc atttctccaa    5700
tgcgggcggc tggagcatgg cctttctacgg gttcgctgcg agtcttgcca cgccgagcac    5760
atggtagctc taagaatcca atctaagta tatctgagta aactcggtct gacctgcacc    5820
tgcatagcca tatggatga tgcgacgtgg tagacgcaca aaccgcaggt aacctgcaag    5880
atcctcagga ccatctgaat catgcgcgga tgacacgaac tcacgacggc gatcacagac    5940
attaacccac agtacagaca ctgcgacaac gtggcaattc gtcgcaatac aacgcacgaa    6000
aaacaggtat tgacaacatg aagtaacatg cagtaagata caaatcgcta ggtaacacta    6060
gcagcatcca gaaacagagg agatattcca tatgagtaaa ggagaagagc ttttcacagg    6120
agtagtacca atcctagtgg aattagacgg tgatgttacc gggcacaagt tatctgtatg    6180
tggagagggc gaaggcgacc caacatatgg caagctgacc cttaaattta tttgcaccac    6240
gggcaaacta cctgctccat ggccaacacc ggtcactacc ttcgggtatg gggttcagtg    6300
ctttgcgcgc taccccagatc acatgaaaca gcacgacttt ttcaagagtg ccatgcccga    6360
aggctatgta caggagagaa ccatcttttt taaggatgac ggcaactata agacacgcgc    6420
agaagtgaag ttcgagggtg ataccctggc taatagaatc gagttaaagg gtattgactt    6480
taaggaagac ggaaatattc taggccacaa actggaatac aactataact ccccaaatgt    6540
gtacattatg gccgacaagc aaaagaacgg tatcaaggtc aactcaaagc tacgacacaa    6600
cattgaggat ggaagcgttc aactagccga ccattaccaa caaaacaccc caattggcga    6660
tgggcctgtg atgttacctg accaccatta actgccaacc aactcagccc tttcgaaagc    6720
tcccaacgaa aagcgcgacc acatggtcct tcttgagttt gtcaccgctg ctgggattac    6780
accccgactg gatgaactcc acccataccc aagtaatagc tactccgacc atacagtacc    6840
aacccgcttc ggcgggtttt tttatggggg gagtttaggg aaagagcatt tgtcagctgg    6900
aaatctgctc gtcagtggtg ctcacactga cgaatcatgc acagatcata ccgatcactg    6960
```

-continued

```
aatggagtct aactaatacg aatgacagcc ggatccagcg caggcggaac aaactggata    7020

tatggcatct aactaaagtc aacgacttca gaagtaaacg agatgttacg aaccactagt    7080

gcactgcagt acacaaaaaa ccccccaaga cccgtttaga ggccccaagg ggttatgcta    7140

atgctgcttt aagacccact ttcacattta agttgttttt ctaatccgca tatgatcaat    7200

tcaaggccga ataagaaggc tggctctgca ccttggtgat caaataattc gatagcttgt    7260

cgtaataatg gcggcatact atcagtagta ggtgttccct tttcttcttt agcgacttga    7320

tgctcttgat cttccaatac gcaacctaaa gtaaaatgcc ccacagcgct gagtgcatat    7380

aatgcattct ctagtgaaaa accttgttgg cataaaaagg ctaattgatt ttcgagagtt    7440

tcatactgtt tttctgtagg ccgtgtacct aaatgtactt ttgctccatc gcgatgactt    7500

agtaaagcac atctaaaact tttagcgtta ttacgtaaaa aatcttgcca gctttcccct    7560

tctaaagggc aaaagtgagt atggtgccta tctaacatct caatggctaa ggcgtcgagc    7620

aaagccgcct tattttttac atgccaatac aatgtaggct gctctacacc tagcttctgg    7680

gcgagtttac gggttgttaa accttcgatt ccgacctcat taagcagctc taatgcgctg    7740

ttaatcactt tactttatct aatctagagc atcattatct cacctatcac aaccttcttt    7800

ctagatctct cccctatagt gagtcgtatt aatttccgag cggccgcgat tatcaaaaag    7860

gatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata    7920

tgagtaaact tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat    7980

ctgtctattt cgttcatcca tagttgcctg actccccgtc gtgtagataa ctacgatacg    8040

ggagggctta ccatctggcc ccagtgctgc aatgataccg cgagacccac gctcaccggc    8100

tccagattta tcagcaataa accagccagc cggaagggcc gagcgcagaa gtggtcctgc    8160

aactttatcc gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc    8220

gccagttaat agtttgcgca acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc    8280

gtcgtttggt atggcttcat tcagctccgg ttcccaacga tcaaggcgag ttacatgatc    8340

ccccatgttg tgcaaaaaag cggttagctc cttcggtcct ccgatcgttg tcagaagtaa    8400

gttggccgca gtgttatcac tcatggttat ggcagcactg cataattctc ttactgtcat    8460

gccatccgta agatgctttt ctgtgactgg tgagtactca accaagtcat tctgagaata    8520

gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaata cgggataata ccgcgccaca    8580

tagcagaact ttaaaagtgc tcatcattgg aaaacgttct tcggggcgaa aactctcaag    8640

gatcttaccg ctgttgagat ccagttcgat gtaacccact cgtgcaccca actgatcttc    8700

agcatctttt actttcacca gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc    8760

aaaaaaggga ataagggcga cacggaaatg ttgaatactc at                      8802
```

<210> SEQ ID NO 78
<211> LENGTH: 8301
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized. pBTK310

<400> SEQUENCE: 78

```
actcttcctt tttcaatatt attgaagcat ttatcagggt tattgtctca tgagcggata      60

catatttgaa tgtatttaga aaaataaaca aataggggtt ccgcgcacat ttccccgaaa     120
```

-continued

```
agtgccaccc gtccggaccg aactccctta aagtgagtca gactgccgcc ttgggccggg    180
tgccgtcgta cttgccagcc gagcactcgg ttcccgtcca gcccagcgcg accagctccg    240
gcaacgcctc gcgcaccggc ttgcggcgcc tgcgcatggc cgaaccactg gcctccgacg    300
gccagacata gccgcacaag gtatcgatgg aagccctgcc ggtttgccgg gggtcgatcc    360
agccacgccg ccgcaggtgc agcaggcggg cggtaacgca gtcccagcgc cgcaccccgt    420
ccatgctgat gcgcacatgc tggccgcccc ccatgacggc ctgcgcgatc aaggggttca    480
gggccacgta caggcgcccg tccgcctcgc cgctggcgta ctccgacagc agccgaaccc    540
cctgcagctc gaggccattc tgggcgatga tggccaactc caccaggagc tagccgcagt    600
cctgcatgtg cttgagcgcc ccaccactac cgaccctgc cccgatttcc tttgccagcg    660
cccgatagcc acctttgacc acatggcatt cagcggtgac ggcctcccac ttgggctcca    720
ggcacagccg gagcatgccg accgcctcgg tcttgggttc cgggccaagc actaggccat    780
taggcccagc catggcccac agcccatgca ggatgcgcag atcatcagcg cccagcggct    840
caggccagcc gacctagctc agctcgccgc agccgtagtc ctccgccacg tcccgcttgc    900
tgcgcttgcg ctcgccaagc ttgccggcac ggcccaggcc ggggcccagc cagtgcgcag    960
ggtcgtgccg gacgcggcgg aggccgtgcc tgttcctagg cttcaccacg gggcaccccc   1020
ttgcccttgc gctgcctctc cagcacggcg ggcctgagca ccccgccgtc atgccgcctg   1080
aaccacccga cagcgaacgg tgcgccatcg ttggccttgc tccacaccaa gcggacgaag   1140
aacccgcgca ggtcgtcgtc cacacccca tccccggcca cggcgctggt catgcacgac   1200
aggcaggcca gcccgcggcc gttccgcacc agtcccgagc tgccccggct ggcccgctgc   1260
tggccgcctg agcccatccc ggcccgcccc ttgccggccc ggtgcaggcc cacgacagcg   1320
caccggtacc cggcggcgac ggcctccatg cgaccgatga cctggccat ggggccgctg   1380
gcgcctttcc cctcgatgtg gaaccgcggc agcgcgtcca gcaccatcag gcggccgccc   1440
tcggcggcgc gcttgaggcc gtcgaaccac tccggggcca tgatgctggg cagggcagcg   1500
atcagcggca ggatcagcag gccgtcagcc acggcatgcc gttccacggc gctgaggtgc   1560
gaccccaggg agtgcaggcg gtgccgcctg gaggccggcg ggtcccaggc gggcacgtcg   1620
atcaccgggc cggtgcgcag ttcgcccacc tccagcagac ccggcccgcc tgcaatctgt   1680
gcggccagtc gcaggcgcag catggattta ccggcaccac cgggcgacac cagcgcccg   1740
accgtaccgg ccacccatgtt gggcaaaacg tagtccagcg gtggcggcgc tgctgcgaac   1800
gccaccgaca tattgatcgg cttacgggta gccaccgatc gcctcctttg cagggcagttg  1860
gtgcctaggc gatgccgggc tacctacccc agcccgagc agctccgagt tattccaggc   1920
aatcgcgccg agcccagtcc tagtcgtcgg tacgccagca attgcgcatgc agccccaatt  1980
tggcctcccc gagccaggca tatcgcgcc gggcgcaagc gtccgggatg gaccgcaggt   2040
cgccggtctg cttgcccttc tggtccttca tatcagtcac cgagaaactt gccgggcgcg   2100
aacggattgc attcgaggca caaggaaacg gtgcagcagc caacggtacg gatggcaatc   2160
tcagcgactg aaaagcggcc agccacggcc ttgtacgacg tataaccaaa gccaccgggc   2220
aacccaatagc aattgtaccc tttgatcagg tagcccgacc atgaagagct ttttcgtat   2280
tccatcaaac ccccatctgt gcgtgagtac tcatagtata acaggcgtga gtacccacgc   2340
aagcactaca tgctgaaatc tggcccgccc ctgtccatgc ctcgctgggg gggtgccggt   2400
```

-continued

```
gaccgtgcca gatcggaccg agccagctgg aagctgggca gacccatgcc attgctgacg    2460
gtgcgatcga tgtcatcagc ttcgccgccg ggctcgagct atgccagagc tgggctggcc    2520
tcggccatgg ccttgccgat ttcctcggca ctgcggcccc ggctggccag cttctgcgcg    2580
gcgataaagt cgcacttgct gaggtcatga ccgaagcgct tgaccagccc ggccatctcg    2640
ctgcggtacc cgtccagcgc cgtgcgccgg tggcggctca gctgccgctc gggcagttcg    2700
aggcaggcca gcctgcgggc cttctcctgc tgccgctggg ccctgcacgat ctgcaggcca   2760
gacagatgca aacgcgaacg gaccgcggtg gaggacttgc aattggattc aagcagaagc    2820
aacgaaggct gatcaacggc gaggatggtg tgctcgtact tgcggatggt gaagccagcc    2880
aagcggcgat agtggcggct gtcggcgctg gcgggctcgg cgtcgtactc gctggccagc    2940
gtccgggcaa tctgccccgg aagttcaccg cctgcggcgt cggccacctt gacccatgcc    3000
tgacagttca tcgggctggc ttccactacc agggcaggca ccgggccctc ggctaccatg    3060
tcaacgaggc caaactcgca gaggacgtcc accagcacca gacccagccg ctcccgctcg    3120
gagggaatga tatccaagtc attgcccatg gaatccatcc gattgagcac tggcgagttc    3180
tggagaactc aggcggatga aactccacgg ttcaccatcc ggccggtggg tgcgacaatg    3240
acgccgatac cgaagcgctc acagcccatg gcgttgagct gtcggcctat ggcctgcaaa    3300
gtcctgtcgt tcttcatcgg gccaccaagc gcagccagat cgagccgtcc tcggttgtca    3360
gtggcgtccg gtcgagccag agccacgatg cgatcagccg caccaccgta ggcaacatgg    3420
aagccagcac caggttagc catagcttcc agtgccaccc ccgcgacgcg ctccgggcgc     3480
tatgcgaggc gatgctaccc taggcggcta aatcccgaca atcttggccc agctccaacc    3540
atgccgaccc tgtcaggcgc tgggcattca gaacctaagc agcctgagcc tagctggact    3600
gcttggtctg gctcatgacc tgccggcctt cgtcggccag tgtcgccatg ctctggcgca    3660
gcggttcgat ctgctccgct aactcgttga tgcctctgga tttcttcact ctgtcgattg    3720
cgttcatggt ctattgcctc ccggtattcc tgtaagtcga tgatcagggc gttggcggtg    3780
tcgatgttca gggccacgtc tgccccgtcg gtgcggatgc cccggccttc catctccacc    3840
aagttaggcc aaatgtgaac aacggaatgg agctcgatgc aatgcgaatc aagtgctatg    3900
tggtcaatgc gggcgtcgtg gccagcccgc tctaatgccc ggttggcatg gtcggcccat    3960
gcctcgcggg tctgctcaag ccatgccttg ggcttgagcg cttcggtctt ctgtgcccgg    4020
cccttctccg gggtcttgcc gttgtaccgc ttgaaccact gagcggcggg ccgctcgatg    4080
ccgtcattga tccgctcgga gatcatcagg tggcagtgcg ggttctcgcc gcccacggca    4140
tggatggcca gagtataccg aaggcgatcg gaaccggtca ggtgctgggc gaactcggac    4200
gacagagcct tatgctggtc gaggggaagc tagaccggca gggcaaattc gacctcattg    4260
aacagaagcc aattggagcg ttcataacgg taggcagatt aaacgtagtc ggcgggaagc    4320
tcgacgaact ccggcatgtg cccggattcg gcgtgcaaga cttcatccat gtcgcgggca    4380
tacttgactc agcgctggct gtagtcggcc ttggccatgg aagattggcc gacggaactg    4440
ctgccggttt tcgccgtaag gtgataaatc gccatgctgc ctcgctgttg ctttggcttt    4500
taggctactt gattgggccc taggagagcg aacggcaaga agggtggacg ttaggcaagt    4560
ttctcgaaga gaaaccggta agtgcgccct cccctacaaa gtagggtcgg gattgccgcc    4620
gctgtgcctc catgatagcc tacgagacag cacattaaca atgggctgtc aagatcgtta    4680
```

-continued

```
aggggagcca aacggaggcg gatcggatgg caacgctaga agcccaaagc gagcgaatcc    4740

atgccgaccc tacgcgggcg agggcaaggg caccgcagca agagcgcaag aacgaaaccc    4800

ggcgcaaggc gctggtgggc gccatgattt tggccaaggc gaacagcagc gagtggccgg    4860

aggatcggcc catggcggca atggatggcc acctcgaacg cgaccacgac cgcgccttgt    4920

tcggtctgcc gcccacgccg aaggatgagc cgggctgaat gatcgaccga gacaggccct    4980

gcggggctgc accacgcgcc ccacccttcg ggtagggggа aaggccgcta aagcggctaa    5040

aagcgcacca gagtatttcc gcgggctttg gtgtcgggtc tagcgggatt tgcccgcatt    5100

tacccatgcc gagcagaggc ggggcggtgc gtcgcctagc gaccgcgatc gaccagatct    5160

ccggcctctg gcggggcata ttgggcaagg gcagcagcgc cccacaaggg cgctgataac    5220

cgcgcctagc ggatcattcc tagataatca tggatcgatt tttccaacac ccggccagcc    5280

cccgcccctg ctgggtttgc aggttcgggg gcgtgacagc tattgcaggg gttcgtgaca    5340

gttattgcag gggggcgtga cagttattgc aggggttcgt gacagttagt acgggagtga    5400

agggcaatgg atggcaatgc atcgcaacgg caggcatttc ggctgagggt caccgaactt    5460

tagcttaagc gatcgaatgc atgtaaaccc agtcccgaca ggaccgcggc catgcctact    5520

gtggcggcca ggaccgccag ccgggatcgg gatactggtc gttaccagag ccacggaccc    5580

gagcaaaccc ttctctatca gatcgctgac gagtattacc cggcattcgc tgcgcctatg    5640

gcagagcagg gaaaggaatt gccgggctat gtgcaacggg aatttgaaga atttctccaa    5700

tgcgggcggc tggagcatgg ctttctacgg gttcgctgcg agtcttgcca cgccgagcac    5760

atggcagctc tacgaaatca atctaaagta tatccgagta aacttggtct gaccggcacc    5820

tgcattagca tatcgaatga tgcgacgtgg tagagccaca aaccgcaggt caccgcaacg    5880

atcgccagga ccatctgaat catgcgcgga tgacacgaac tcacgacggc gatcacagac    5940

attaacccac agtacagaca ctgcgacaac gtggcaattc gtcgcaatac aacgcgtgaa    6000

acaaaacggt tgacaacatg aagtaaacac ggtacgatgt accacatgaa acgacagtga    6060

gtcaatacag aaacagagga gatattacat atgagtaaag gagaagagct tttcacagga    6120

gttgtcccaa ttcttgttga attagatggt gatgttaatg ggcacaagtt ctctgtcagt    6180

ggagagggtg aaggcgatgc aacatatggc aagctgaccc ttaaatttat ttgcaccacg    6240

ggtaaactac ctgttccatg gccaacactg gtcactacgc tcgggtatgg ggttcagtgc    6300

tttgcgcgct acccagatca catgaaacag cacgactttt tcaagagtgc aatgcccgaa    6360

ggctatgtac aggagagaac catcttcttt aaggatgacg gcaactacaa gacacgcgcc    6420

gaagtgaagc tagagggtga taccctcgtc aatagaatcg agttaaaggg tattgatttt    6480

aaggaagatg gaaacattct aggccacaaa atggaatata actacaactc ccataatgtg    6540

tacatcatgg cagacaagca aaagaacggc atcaaggtta acttcaagat cagacacaac    6600

attgaggatg gaagcgttca actagccgac cattaccaac aaaacacccc aattggcgat    6660

gggcctgtgc tgttaccaga caaccattac ctgtccactc aatctgccct ttcgaaagat    6720

cccaacgaaa agcgcgacca catggtcctt cttgagtttg tcacggctgc tgggattaca    6780

catggcatgg atgaactata caaataactc agtcaacgtc aatccgcacc tacgtaaccc    6840

acccgcttcg gcgggttttt ttatgggggg agtttaggga aagagcattt gtcagctgga    6900

aatcggctcg tcagtggtgc tcacactgac gaatcatgta cagatcatac cgatgactgc    6960
```

-continued

```
atggcgactc aaaacatagc aagacagccg gatccagagc aggcgaaacc aactgcatat    7020

atggcatatc aaaacagtcc aagtcaaacg aagtaaaaga gatgaaaaga aaaactagtg    7080

cactgcagta cacaaaaaac ccctcaagac ccgtttagag gccccaaggg gttatgctaa    7140

tgctgcttta agacccactt tcacatttaa gttgtttttc taatccgcat atgatcaatt    7200

caaggccgaa taagaaggct ggctctgcac cttggtgatc aaataattcg atagcttgtc    7260

gtaataatgg cggcatacta tcagtagtag gtgttttccct ttcttcttta gcgacttgat    7320

gatcttgatc ttccaatacg caacctaaag taaaatgccc cactgcgctg agtgcatatc    7380

atgcattctc tagtgaaaaa aattgttggc ataaaaaggc taattgattt tcgagagttt    7440

catactgttt ttctgtaggc cgtgtacctа aatgtacttt tgctccatcg cgatgactta    7500

gtaaagcaca tctaaaactt ttagcgttat tacgtaaaaa atcttgccag ctttcccctt    7560

ctaaagggca aaagtgagta tggtgcctat ctaacatctc aatggctaag gcgtcgagca    7620

aagcccgctt atttttttaca tgccaataca atgtaggctg ctctacacct agcttctggg    7680

cgagtttacg ggttgttaaa ccttcgattc cgacctcatt aagcagctct aatgcgctgt    7740

taatcacttt acttttatct aatctagaca tcagtatatc tactctacta aactaattc    7800

tagatctctc cctatagtga agtcgtatta atttccgagc ggccgcgatt atcaaaaagg    7860

atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta aagtatatat    7920

gagtaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc    7980

tgtctatttc gttcatccat agttgcctga ctccccgtcg tgtagataac tacgatacgg    8040

gagggcttac catctggccc cagtgctgca atgataccgc gagacccacg ctcaccggct    8100

ccagatttat cagcaataaa ccagccagcc ggaagggccg agcgcagaag tggtcctgca    8160

actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt aagtagttcg    8220

ccagttaata gtttgcgcaa cgttgttgcc attgctacag gcatcgtggt gtcacgctcg    8280

tcgtttggta tggcttcatt cagctccggt tcccaacgat caaggcgagt tacatgatcc    8340

cccatgttgt gcaaaaaagc ggttagctcc ttcggtcctc cgatcgttgt cagaagtaag    8400

ttggccgcag tgttatcact catggttatg gcagcactgc ataattctct tactgtcatg    8460

ccatccgtaa gatgcttttc tgtgactggt gagtactcaa ccaagtcatt ctgagaatag    8520

tgtatgcggc gaccgagttg ctcttgcccg gcgtcaatac gggataatac cgcgccacat    8580

agcagaactt taaaagtgct catcattgga aaacgttctt cggggcgaaa actctcaagg    8640

atcttaccgc tgttgagatc cagttcgatg taacccactc gtgcacccaa ctgatcttca    8700

gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca aaatgccgca    8760

aaaaagggaa taagggcgac acggaaatgt tgaatactca t                        8801
```

```
<210> SEQ ID NO 79
<211> LENGTH: 8110
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized. pBTK319

<400> SEQUENCE: 79

aacaccccct gtattactgt ttatgtaagc agacagtttt attgttcatg atgatatatt      60

tttatcttgt gcaatgtaac atcagagatt ttgagacaca acgtggcttt gttgaataaa     120
```

-continued

```
tagaactttc gatgagttga aggctcagtc atgaccaaca taccttaacg tgagtcagcc      180
tgccgaattg ggccgggtga tgtcgtactc gaccgcagcg aactcggttc aagtccagcc      240
cagcgcgacc agctccggca acgcctcgcg cacccgcttg cggcgcttgc gcatggtcga      300
accactggcc tctgacggcc agacatagcc gcacaaggta tctatcgaag ccttgccggt      360
tttgccgggg tcgatccagc cacccagccg ctggtgcagc aggcgggcgg tttcgctgtc      420
cagcgcccgc accttgtcca tgctgatgcg caccagctgg ccgccacccc tgacggcctg      480
agcgataacg gggtcaaggg aaacgtaacg gagcccgtcc gactcgtagc tgggcgaatc      540
agtcagaagc agtaaaacct gacgcatgcg gacttcatgg gagatgatgg atacctactc      600
aaggcgctcg atgcagtcct gtatgcgctt gagcgcccca ccactatcga cctcgcccgc      660
gatttccttt gccagcgccc gatagctacc tttgaccaca tggcattcag cggtgacggc      720
ctcccacttg ggttccagga acagccggag ctgccgtccg ccttcggtct tgggttccgg      780
gccaagcact aggccattcg gcccagccat ggccaccagc ccttgcagga tgcgcagatc      840
atcagagccc agcggatccg ggccgctgca atcgataagc ttgccgtagc agtagtaatc      900
agtcaagtcc agctagatgc gattgcgatc gacccgattg agggcaaggc aaaggcaggg      960
ggccagacag tggcccgggt cgtgccggac gtggctgagg ctgtgcttgt tcttaggctt     1020
cacacggggg cacccccttg ctcttgcgct gcctctccag caggcgggcc ttgagcaccc     1080
cgccgtcatg ccgcctgaac caccgatccg cgaacggtgc gccatagttg gcctggctca     1140
caccgaagcg gacgaagaac cggcgctggt cgtcgtccac accccattcc tcggcctcgg     1200
agctggtcct gatcgaacgg taggactgcc agcggatgtt atcgaccagt accgagatgc     1260
aacggatggc atgctgatgg tagcctgcgc aactcatggc agcgccattg atggcatggt     1320
gcaggaacac gatagagcac ccggtatcgg cggcgatggc ctccatgcga ccgatgacct     1380
gggccatggg gccgctggcg ttttcctcct cgatgtggaa ccggcgcagc gtgtccagca     1440
ccatcaggcg gcggcccacg gcggcgcgct tgaggccgtc gaaccactcc ggggccatga     1500
tgttgggcag gctgccgatc agcggctgga tcagcaggcc gtcagccacg gcttgccgtt     1560
aatcggagct gaggtgagcc aaacggcgct gatggcggtg atgatggccg gtggcggggt     1620
cttcgcgggg caggtagatc accggcgcgg tgggcagttc gcccacctcc agcagatccg     1680
gcccgcctgc aatccgtgcg gccagttgca gggccagcat ggatttaccg gcaccaccgg     1740
gcgacaaccg cgccccgacc gtaccggcca ccatgatggg caaaacgtcg tccagcggtg     1800
gcggcgctgc tgcgaacgcc tcccgaatat tgatagggtt atgggtagcc attgattgcc     1860
taccatgcag gacgatggtg gttaggagct ggcgggtcca ataccaagca aatgcgaagc     1920
tatgagttcc taccggaacc agcgcagcgc atcgcattcg tagtcggtcc gaccggaatt     1980
gagctgacgc atccattggc aattcatgcg atcggcatcc agcgctggcc gtacagagtc     2040
aggcctggcc agcaggtcgc cggtcctgct gtcctttggc tctttcatat cagtcaccga     2100
gatacttgcc ggggcagtaa ggctagtatc agcggaaaaa ggtcaaggtg aagccgtaac     2160
ggttaaggct ggccatatca gcgactgaaa agcggccagc ctcggccttg tttgacgtat     2220
aaccaaagcc aacggaaacc aactagacct tgtcactttc gatcaggtag aacgacaatg     2280
aagcgctttt ttcgtattcc ataaaacccc cttctgtgcg tgagtactca tagtataaca     2340
ggcgtgagta ccaacgcaag cactacatgc tgaaatctgg cccgcccctg tccatgcctc     2400
```

-continued

```
gatggagggg tgccggtgcc agtgccagcc aggcccgagc aagcaggacg atgggcagac    2460
aaacgaacta gatgaaggtg agctcgatga aatccgattc gtggccgggc ttgcgctatg    2520
ccagcgctgg gctggcctgg gccatggccc tgccgatttc ctggccactg cggccccggc    2580
tggccagctt ctgccgggcc ataaagtcgc acttgctgag gtcatgaccg aagcgcttga    2640
ccagcccggc catcaggctg cggtactcga ccagcgccga gcgccggtgg cggccaagct    2700
gccgctgggg cagttcgagg ctggccagcc tgcggggcct ctcctgctgc cgctggggct    2760
gatcgatctg atggcaagcc tgctgcacca gagccgggcc agcggtggcg gtctcgaact    2820
tggatcaacg aagcagaacc aacggctgca aacccgagcg ggtggcgtgc ttgtccttgc    2880
ggttcgtgaa gcccgccaag cggccatagc ggcggctgtc ggcgctggcc gggtcggcgt    2940
cgtactcgcc ggccagcgtc cgggcaatcc gcccccgaag ttcaccgcct gcggcgtcgg    3000
ccaccttgac ccatgcctga cagttcttcg ggctggtttc cactaccagg gcaggctccc    3060
ggccctcggc tttcatgtca tcccaggtca actcgctgag gtcgtccacc agcaccagac    3120
aatgcagatc atgctaggcg ggcctgatcc aaacgtaata gaccagggca ttactcagct    3180
tgagcaatgg agtgctatgg agcactacgg aggctgaaca ttcccggttc atcactgggc    3240
cggtcggtgc gtcactgacg ccgatatcga agcgctcaca gcccatggcc ttgagctgtc    3300
ggcctatggc ctgcaaagtc ctgtcgttcc tcatcgggcc accaagccga gccagatcga    3360
gccgtcctcg gttgtcagtg gcgtcaggtc gagcaagagc aacgatgcga tcagcagcac    3420
caccgtaggc atcatggaag ccagcatcac ggttagccat agcttccagt gccacccccg    3480
agacgagatc agggcgatcc gagcggagcc gatcaccacg gaggctaact aacgcaaatc    3540
tttggaaagc taaacaaatg aagcccatgc atggcgatgg gatttcagcc aatccgaagc    3600
ctgcgcctcg ctggcctgcc tggtctggcc catgacctgc cgggcctcgt cggccagtgt    3660
cgccatgctc tgggccagcg gttcgatctg ctcggctaac tcgttgatgc ctctggattt    3720
cttcactctg tcgattgcgt tcatggtcta ttgcctcccg gtattcctgt aagtcgatga    3780
tctgggcgtt ggcggtgtcg atgttcaggg ccacgtctgc ccggtcggtg cggatgcccc    3840
ggcctaacct atcaaaaacg ttcggcaaca ggtgaaaacc gggcaagagc tagacgaact    3900
gcgcctcaag tgttctgtgg tcaatgcggg cgtcgctggc agcccgctcc aatgcccggt    3960
tggcatggtc ggcccatgcc tcgccggtcc gctcaagcca tgcccgggcc ttgagcgctt    4020
cggtcttctg tgcccggccc ttctccgggg tcttgccgtt gtaccgcttg aaccactgag    4080
cggcggggcg ctcgatgccg tcattgatcc gctcggagat catcaggtgg cagtgcgggt    4140
tatcgaagcc aacggaatgg atggccagcg tatacggaag gagcacggca aaggacaggt    4200
gatgggagca atcggaagcc agcgccttca gatggaagag ggtcagatcg aacggcaggg    4260
aacaataagc atccatgaac agccgcaaca tggcgcgttc atcaaggtcg gaagcataac    4320
agtagtcggc gggccgctcg acgaactccg gcatgcgccc ggattcgggg tgcaagactt    4380
aatccatgtc gaggcaatac ttgccatcgc gatggatgta gtcggcattg gaccacgacg    4440
attggccgcc cgacctgctg ccggcttccg ccgtaaggtg atcaatcgcc atgcagcctc    4500
gatgatgcta ttgcatttcg gatccatgca atggccatcg gagacgcaac agcccgaagg    4560
gtggccgtta ggccagtttc tcgaagagca accggcaagc gcgccctccc ctacaaagta    4620
gggtcgggat tgccgccgcc gtgcctccat gatagcctac gagacagcac attaacaatg    4680
```

-continued

```
gggcgtaccg atggctacgg ggcgcaacca ggcggcggcc cggcccgacc agctcgaagc   4740
accacgagcg agcccaactg cagcaattca gagggagagg gacccggacc agccgcaagc   4800
gcgcaagaac gaaacaaggc gcaaggtgcc ggtggcggcc atgatttggc ccaaggtgaa   4860
cagcagcgag tggccggagg atcggctcat ggcggcaatg gatgccgtac ttgaacgcga   4920
ccacgaccgc gcctcgttcg gtctgccgcc acgccagaag gatgagccgg gctgaatgat   4980
cgacccgagc aggccctgcg gggctgcaca cgcgcccccc ccctccgggt aggggggaag   5040
gacgctaccg aggccaaccg agctccagcg tattcctgcg gggtccggtg tggggcttcg   5100
agggctttgc cagcctttcc cactgcagcg cagcggtggg gaggccgtgt gcctagagcc   5160
gcgaatagac cagccatccg gcctccggcc gggcatattg ggcaagggca gcagcgcccc   5220
acaagggcgc tgataacggc gcctagtgga ttatccttag ataatcatgg atggattttt   5280
ccaacacccc gcccgccccc gcccctgctg ggttggcagg tttggggggg tgacagttat   5340
tgcagggggt cgtgacagtt attgcagggg ggcgtgacag ttattgcagg ggttcgtgac   5400
agtcagtacg ggagtgacgg gacctggctg gacccgtata gaccgggacg gacctcaggc   5460
tgacggtaca agccctttcc gatcagcgcc agcccgtatg taccccaagt attgcaaggc   5520
cgcggaacat gcctcatgtg gcggccagga cggccagccg ggatcgggat actggccgtt   5580
accagagcca ccgacccgag caaacccttc tctatcagat cgttgacgag tattacccgg   5640
catccgctgc gcttatggca gagcagggaa aggaatgccc gggctatgtg caacgggaat   5700
ttgaagaatt tctccaatgc gggcggctgg agcatggctt tctacgggtt cgctgcgagt   5760
attgccaagc agcgcaactg gtcgccttca gacctcaatc tacctcatct atgacgcacc   5820
ttggcatgac aggccactga attcgcatcc agatggtaga gacccaaacc gacggtaacc   5880
gcaacgatcc ccaggaccat ctgaatcatg cgcggatgac acgaactcac gacggcgatc   5940
acagacatta acccacagta cagacactgc gacaacgtgg caatccgtcg caataccgtc   6000
tcactgaact ggccgataat tgccgacgca cgttatccaa aagacgattg acttaaagtc   6060
taccctatcg gatacttaca gccatcgaga gggacacggc gaggaattgt gagcggataa   6120
caccccaata cagcaaacga ggcgatatta cataccagta aaggagcagc gattcccacc   6180
ggagctgtcc caatcctgcc ggaattagac ggtgatctta atgggcacaa gttctctgtc   6240
agtggagagg gtgaagggga cgcaacatac ggcaagctga cccttaaatt tatttgcacc   6300
acgggtaaac taccctgttc atggccaaca ctggtcacta cgttcgggta tggggttcag   6360
tgctttgcgc gctaccccga tcacatgaaa cagcacgact ttttcaagag tgccatgccc   6420
gacggctatg taccggagcg caccatcttc tttaaggatg acggcaacta taagacacgc   6480
gccgaagtga agttcgaggg tgataccctc gttaatagca tcgagctaaa gggtatcgac   6540
tttaaggacg atggaaatat tttaggacac aaactggact ataactataa atcccatact   6600
gtgtacatta tggccgacaa gcaaaagaac ggtatcaagg ttaacttcaa gatcagacac   6660
aacattgagg atggaagcgt tcaactagcc gaccattacc aacaaaaaac accaattggc   6720
gatggccctg tgctgttacc agacaaccat tacctgtcca ctcaatctgc cctttcgaaa   6780
gatcccaacg aaaagagcga ccacatggtc cttcttgagt ttgtaacggc tgctgggatt   6840
acacacggca tggatgaact atacaaataa atccgtaatc gttaatccgc aaataacgta   6900
aaaacccgct tcggcgggtt ttttatgggg ggagtttagg ggaaagagca tttgtcagct   6960
```

-continued

```
gaaaatgaga agaaggggtc atcaaggatc atcaagagcc aacaaaatgt gtgaaataca     7020

agcacggatg aatgcatgaa gaaagcaatg aatgggaaca gaagaaaaac ataagcatgt     7080

gagagaagca gacacccgac agatcaaggc agttaactag tgcactgcag tacaccaggc     7140

atcaaataaa acgaaaggct cagtcgaaag actgggcctt tcgttttatc tgttgtttgt     7200

cggtgaacgc tctcctgagt aggacaaatc cgccgggagc ggatttgaac gttgcgaagc     7260

aacggcccgg agggtggcgg gcaggacgcc cgccataaac tgccaggcat caaattaagc     7320

agaaggccat cctgacggat ggcctttttg cgtttctaca aactcttttg tttatttttc     7380

taaatacatt caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa     7440

tattgaaaaa ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt     7500

gcggcatttt gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct     7560

gaagatcagt tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaagatc     7620

cttgagagtt ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta     7680

tgtggcgcgg tattatcccg tattgacgcc gggcaagagc aactcggtcg ccgcatacac     7740

tattctcaga atgacttggt tgagtactca ccagtcacag aaaagcatct tacggatggc     7800

atgacagtaa gagaattatg cagtgctgcc ataaccatga gtgataacac tgcggccaac     7860

ttacttctga caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg     7920

gatcatgtaa ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac     7980

gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaactggc     8040

gaactactta ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt     8100

gcaggaccac ttctgcgc                                                   8118
```

<210> SEQ ID NO 80
<211> LENGTH: 8272
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized. pBTK320

<400> SEQUENCE: 80

```
acgctacttg cattacagct tacgaaccga acaggcttat gtccactggg ttcgtgcctt       60

catccgtttc cacggtgtgc gtcacccggc aaccttgggc agcagcgaag tcgaggcatt      120

tctgtcctgg ctggcgaatg agcgcaagtt tcggctcgcg gctcgagcgt ttcgggatgc      180

cccgcattac gccgggatat aaggctcgcg gatccgaggc gaatgagcgc gagcgccagt      240

aaggcaaagc ataacgaacc agctaaaggc gatttgcgaa ggtagaaacc atggcatatg      300

aaggcaagac ataacgaacc aaggaatcta tggaagaacc gaaggctttg aaggggtaga      360

tacagaaaca aagccgatgg tgaacgaggc gggcgcttta gatgaaaagc gaaagcaaat      420

cgtcaatgct gatgcgcaca tgctggcagc caccatgaac ggcctgcgcg atcaagggct      480

taaggaaaac gtaaaggagc aagtacgacc agtagctggc gtaacaagac agcaagcaac      540

acccctgccg cttgcggcca ttctgggcga tgatggatac cttccaaagg cgctcgatgc      600

agtcatgtaa gtgaatgagc gaaacaaacc tatagaaatc tgaaacgatt taataagaac      660

gcgcccgata gctacctttg accacatggc attcagcggt gacggcctcc cactgggttc      720

ccaggaacag ccggagctgc cgtccgcctt cggtcttggg ttccgggcca agcactaggc      780
```

-continued

```
aattaggccc agccatggcc aaccgccatt gaaggatgcg aagaacatcc gagcccagcg     840
gatccgggcc gatgaactcg atccgcttgc agtcgccgta gtcctaagtc aagtccagct     900
tgccgcgctt gcgctcgccc cgcttgaggg caggaacagc gccgggggcc agacagtgcg     960
ccgggtcgtg ccggacgtgg ctgaggctgc gcttgctctt aggcttcacc acggggcacc    1020
ccctggctct tgcgctgcct ctccagcacg gcgggcttga gcaccccgcc gtcctgccgc    1080
ctgaaccacc gatcagcgca cggtgcgcca tagttggcct tgctcacacc gaagcggacg    1140
aagaaaaggc gatggtagtc gtccacaccc aattccatgg aatcggggct ggtcaagatc    1200
gacaggtagg aatgcaagcg gatgcaatcg aaccgtaacg agctgcaaag gatggcatgc    1260
tgccgctcgc ctgcgcccat catggccgcg cccttgctgg catggtgcag gaacacgata    1320
gagcaccggg tatcggcggc gatggcctcc atgcgaccga tgaccaggga catgggcccg    1380
ctggccgttt cttcctcgat gtggaaccgg cgcaggcgtg ccagcaccat caggccggcg    1440
ccctcggcgg cgcgcttgag gccgtcgaac cactccgggg ccatgatgtt gggcaggctg    1500
aagataagcg gatggatcag aaggccgtca gacacggatt gacgtaaatc ggcgctgagg    1560
tgcgcaaaca gggcgtgccg gaggtgatga atggcggtgg gaggggattc ggcgggaagg    1620
tagatcaccg ggccggtggg cagttcgccc acctccagca gatccggccc gcctgcaatc    1680
tgtgcgggca gttgcagggc cagcatggat ttacccgcac caccggcgca caccagcgcc    1740
ccgaccgtac cggccaccat gttgggcaaa acgtagtcca gcggtggcgg cgctgctgcg    1800
aacgccacca gaatattgat aggcttatgg gtagccattg attgccatct ttgcaggcag    1860
ttggtggtta ggcgctggcg gggtcaatac aaccgcaatg agcagctatg agttcttacc    1920
ggcactagcg aagcgaatcg tattcgtacg aggtcagcca gacctggagc tgacgcatcc    1980
ctttgcgctt catgcgctcg gcatatcgcg cttggcgtac agcgtcaggg ctggccagca    2040
ggtcgccggt ctgcttgtcc ttttggtctt tcatatcagt caccgagaaa cttgccgggg    2100
ccgaaaggct tgtcttcgcg gaacaaggac aaggtgcagc cgtcaaggtt aaggctggcc    2160
atatcagcga ctgaaaagcg gccagcctcg gcctggtttg acgtataacc aaagccaccg    2220
ggcaaaaacc agcccttgtc aatttcgatc aggtagaacg aacctgaagc gatttcttcg    2280
tattccataa aacccccttc tgtgcgtgag tactcatagc ataacaggcg tgagtaccaa    2340
cgcaagcacc acatgctgaa atctggcccg ccccgtccca tgcctcgctg gcggggtgcc    2400
ggtgccggtg ccagctgggc ccgcgcaagc tggacgctgg gcagacccat gacctggctg    2460
acggtgcgct cgatgtaatc cgcttcgtgg ccgggcttgc gctctgcccg cgctgggctg    2520
gactcggcca tggccttgcc gatttcatcg gaactgaggc aacggctggc aagcttatgc    2580
gaggcgatca agtcgaactt gatgaggtca tgaccgaagc gatttgaaag aacggcaatc    2640
tagctgaggt aatcgtaacg agccgtgcgc aggtggaggc taagctgaag atcgggaagt    2700
tcgagcctgg ccagcctgcg ggcctcctcc tgctgccgcc gggcctgctc gatcagctgg    2760
aaagcatgct gaaccagcgc aggggcagcg gtggcggtct tgccatggct ttacggaagc    2820
agcacccacg gctgataacc ggcgcgggtg gtgtgcttgt ccttgcggtt ggtgaagccc    2880
gaaaagaggc aatagtggcg gatgacggcg atggccgggt aggcgtagtc atcgctggcc    2940
agcgtccggg caatctgccc ccgaagttca ccgcctgcgg cgtcggccac cttgacccat    3000
gcctgatagt tcttcgggct ggtttccact acgagggcag gctccggccc ctcggctttc    3060
```

-continued

```
atgccatcca ggtcaaactc gatgacgtcg tacaccagca aacgaccatg aagctcatgc    3120
taggcgggcc tgctatacac gtcaccgccc tgggcattca taagcatgcg aactggcgtg    3180
ttctggagca cttcggcggc tgaccattcc cggttcatca tctggccggt gggtgcgtcc    3240
ctgacgccga tatccgaagc ctcacagccc atggccttga gctgccggcc tatggcctgc    3300
aaagtcctgt cgttcttcat cgggccaccc agcgcagcca gatcgagccg tcctcggttg    3360
tcagtggcgt caggtcgagc aagagcaacg atgcgatcag cagcaccacc gtaggcatca    3420
tggaagcacg aatcaaggtc agcccatagc taccgtgcca acaccgagcc gagctcaggg    3480
agcctgcggc ggcgctgctc aactcggcgg ctaccaaacg aaccctcttg gaacgctacc    3540
cccatgccgc ccctgtctgg cgctggcgcc tcagccactc cgccgcctgc gcctcgctgg    3600
cctgcttggc ctggctcatg acctgccggg cttcgctggc cagtgctgcc atgcgctggg    3660
ccagcggttc gatctgctcc gctcactcgt tgatgccctt ggatttcttc actctgttga    3720
ttgcgttcat ggtctattgc ctccccggta tcctgtaagt cgatgatctg ggcgttggcg    3780
gtgccgatgc tacggcaccc gtctgccagc tagggcggga tgcccggccc ttccatatcc    3840
aacaacgttc gacccaggtg aacccggggc aggcgctaga tgcccagagc ctccagtgtt    3900
ctgctgtcaa tgcgggcgtc gtggccagcc cgctctaatg cccggctggc atggtcggcc    3960
catgcctcgc gggtctgctc aagccatgcc ttggccttga gcgctccggt cttcgctgcc    4020
ccgcccttct ccggggtctt gccgttgtac cgcttgaacc actgagcggc gggccgctcg    4080
atgccgtcat tgatccgctc ggcgatcatc aggtggcagt gcgggttctc gccgccaccg    4140
gacctgatgc aacgcgtata aggcaggcgc taggcaacgc taacgcgatg ggcgaaatcg    4200
gacgccagcg aattctgctg gtcgacggtc agctcgaacg gacgggcaaa ttcgacatcc    4260
ttgaacagcc gcccattggc gcgttcatac aggtcggcag catcccagta gtcggcgggc    4320
cgctcgacga actccggcat gtgcccggat tcggcgtgca agacctcatc catgtcgcgg    4380
gcatacttgc cttcgcgctg gatgtagtcg gcctgggccc tggccgattg gccgcccgac    4440
ctgctgccgg ttttcgccgt aaggtgatca atcgccatgc tgcctcgctg ttgcttttgc    4500
tttccggctc aatgcaatgc aactcggcga gagcacagcc agcccggtgg aagtcaggcc    4560
agtttctcga agagaaaccg gtaagtcgcc cctcccctac aaagtagggt cgggattgcc    4620
gccgctgtgc ctccatgata gcctacgaga cagcacatta acaatggggt gtcaagatgg    4680
ttaaggggag caacaaggcg gcggatcggc tggccaagct cgaagaacaa cgagcgcgaa    4740
tcaatgccga aattcagcgg gagcgggcaa gggaacagca gcaagagcgc aagaacgaaa    4800
aacgcagcaa ggtgctggtg ggggccatga ttttgcaaaa ggtgaaaagc agcgagtggc    4860
aggacgatcg gatcatggcg gaccccgatg agtaccttga aagcgaaacc gacccgcgct    4920
tgtacggtcc gacgcaacgc aagcacgatg agcccgcatg aatgatacgc agagacaggc    4980
cctgcggggc tgcacacgcc ccccacccct tcgggtaggg ggaaagcccg ctaaagcggc    5040
tacaagcagc aacgcgtatc tatgcggggc ttggcgtggg gtttagcggg atttgcaagc    5100
ctttccccct gccgcgcagc ggtggggcgg tgtgtagcct agcgcagcga atagaccagc    5160
tatccggccc atggcagggc atatcgggca agggcagacg agcccccaaa gggcgctgct    5220
aacccgccct agtggattat tcttagatat tcatggatgg atttttccaa caccccgcca    5280
gcccccgccc ctgccgggtt tgcaggtttg ggggcgtgac agttactgca ggggctcgtg    5340
```

-continued

```
aacgtatttg caggggggcg tgcaagttcc tgccgggtcc cgtgacagtt agtacgggcg    5400
tgacgggcac tggccggcca tgtccagcca cggcacgcac ttcggctgcg ggtcaaagcc    5460
ctttccgcta agcgatagac tgtatgtaaa cacagtattg caaggacgcg gaacatgcct    5520
catgtggcgg ccaggacggc cagccggatc cgggatactg gtcgttacca gagccaccga    5580
cccgagcaaa cccttctcta tcagatcgtt gacgagtatt acccggcatt cgctgcgctt    5640
atggcagagc agggaaagga attgccgggc tatgtgcaac gggaattgaa agaatttctc    5700
caccgcgggc ggctggagca tggccttcta cgggtcagcc gcgcgtcttg cacggcagcg    5760
caccgggtcg ctttcagcca tccctcttca gtctatatga gtcccattgg tatgacaggc    5820
ccccgaattc gcatctagat ggtagagcca caaacagccg gtacaagcaa cgatctccag    5880
gaccatctga atcatgcggc gatgacacga actcacgacg gcgatcacag acattaaccc    5940
acagtacaga cactgcgaca acgtggcaat tcgtcgcaat accgtctcac tgaactggcc    6000
gataattgca gacgaacgtt atcaaacaga gtattgactt aaagtctaac ctataggata    6060
cttacagcca tagtgaggga cacggcgcgg cattgtgagc ggctaaaact tacctaaagc    6120
cacagaggcg atattaacta tgagtaacgg agccgagctc ttacaaggcg ttgtccaact    6180
cctcgtggaa ttagacggtg atgttaatgg gcacaagttc tctgtcagtg gagagggtga    6240
aggcgacgca acatatggca agctgaccct taaatttatt tgcaccacgg gtaaactacc    6300
tgttccatgg ccaacactgg tcactacgtt cgggtatggg gttcagtgct ttgcgcgcta    6360
cccagatcac atgaaacagc acgacttttt caagagtgca atgcccgaag gctatgtacc    6420
ggagagaccc atctttttta aggacgacgg caactacaag acacgcgacg aagtgaagtt    6480
cgagggtgac accctcgtta atcgaatcga gttaaagggt attgacttta aggaagatgg    6540
aaatatttta ggacacaaac tggaatataa ctataactcc cataatgtgt acattatggc    6600
cgacaagcaa aagaacggta tcaaggttaa cttcaagatc agacacaaca ttgaggatgg    6660
aagcgttcaa ctagccgacc attaccaaca aaacacccca attggcgatg ggcctgtgct    6720
gttaccagac aaccattacc tgtccactca atctgccctt tcgaaagatc ccaacgaaaa    6780
gagcgaaccc atggtacttc ttgagtttgt caccgctgcc gggattacac caggcatggc    6840
tgaactatac aaataaatgc gtaatcgtta atcagcaaat aacgtaaaaa cccgcttcgg    6900
cgggtttttt tatgggggga gtttagggaa agagcatttg tcagcatgca atgagacgac    6960
ggggtcatca cggctcatca tgcgcccaac aaatgtgtgc aatacacgct cggatgactg    7020
catgatgacc gcactgactg gggacagccg atccacctca gcctgtgaga gaagcagacc    7080
aacgaaagcc caaggaagtc aactagtgca atgcagtaca aacggcatcc aataaaaagc    7140
aaggctaccc agcaagactg ggcccttcgc ttttactgtc gtttgtcggt gaacgctctc    7200
tcctgagtag gacccggctc aacttcgggc gggcccttcc gcgtttcatc agtgcgggcg    7260
cttatttgcc gactaccttg gtgatctcgc ctttcacgta gtggacaaat tcttccaact    7320
gatctgcgcg cgaggccaag cgatcttctt cttgtccaag ataagcctgt ctagcttcaa    7380
gtatgacggg ctgatactgg gccggcaggc gctccattgc ccagtcggca gcgacatcct    7440
tcggcgcgat tttgccggtt actgcgctgt accaaatgcg ggacaacgta agcactacat    7500
ttcgctcatc gccagcccag tcgggcggcg agttccatag cgttaaggtt tcatttagcg    7560
cctcaaatag atcctgttca ggaaccggat caaagagttc ctccgccgct ggacctacca    7620
```

-continued

```
aggcaaagcc atgttatatt gatttagtca gaaagatagc aagaacaatg tagaacgtgg    7680

atggctagca gatacaagca agaagtaaca tgcgcagcca ttatccaaat tgcaagtagc    7740

gcttagctgg ataacgccac ggaatgatgt cgtcgtgcac aacaatggtg acttctacag    7800

cgcggagaat ctcgctctct ccaggggaag ccgaagtttc caaaaggtcg ttgatcaaag    7860

ctcgccgcgt tgtttcatca agccttacgg tcaccgtaac cagcaaatca atatcactgt    7920

gtggcttcag gccgccatcc actgcggagc cgtacaaatg tacggccagc aacgtcggtt    7980

agagatggcg ctcgatgacg ccaactacct ctgatagttg agtcgatact taggcgatca    8040

ccgcttccct catgatgttt aactcctgtt aagggcgact gcccggatgc gtaccatagt    8100

tgctgctcca taacatcaaa catcgaccca cggcgtaacg cgcttgctgc ttggatgccc    8160

gaggcataga ctgtacccca aaaaaacagt cataacaagc catgaaaacc gccactgcgc    8220

cgttaccacc gctgcgttcg gtcaaggttc tggaccagtt gcgtgagcgc at            8272
```

```
<210> SEQ ID NO 81
<211> LENGTH: 11299
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pDIK349

<400> SEQUENCE: 81
```

```
gcccgcgcca gctggacgct gggcagaccc atgaccttgc tgacggtgcg ctcgatgtaa      60

tccgcttcgt ggccgggctt gcgctctgcc agcgctgggc tggcctcggc catggccttg     120

ccgatttcct cggccatgcg gccccggctg gccagcttct gcgcggcgat aaagtcgccc     180

ttgctgaggt catgaacgca gagcttgccc agcccggcca tatcgctgcg gtaccgtacc     240

agcgccgtgc gccggtggcg gctaagctgc cgctcgggca gttcgaggct gggcagcctg     300

cgggccttcc cctgctgccg ctgggcctgc tcgatctgct ggccagcctg ctgcaccagc     360

gccgggccag cggtggcggt cttgcccttg gattcacgca gcagcacccc cggctgatca     420

ccggcgcggg tggtgtgctt gtccttgcgg ttggtgaagc ccgccaagcg gccatagtgg     480

aggctgtcgg agctggaggg gtaggcgtcg taatcgatgg aaagcgtacg ggccatatgc     540

ccccgaagtt caccgcctgc gggcgtcggc acctcgaccc atgcctgata gttcttcggg     600

ctggtttcca ctaccagggc aggctcccgg ccctcggctt tcatgtcatc caggtcaaac     660

tcgctgaggt cgtccaccag caccagacca tgccgctcct gctcggcggg cctgatatac     720

acgtcattgc cctgggcatt catccgcttg agcccatggc gtgttctggg cactccggcg     780

gatgaaactc aacggtacct catctggccg gtgggcgagc aactgaagcc gatcacgacg     840

agcccaacgc aactggactc gagcagtcgg catcaggacc gaacagtact gtcgccattc     900

atcgggaccc aacgcgaagc aagaacgagc agtcccaggc tgtcagtggc gtcaggtagc     960

gcaagagcaa cgatgcgatc agcagcacca ccgtaggcat catggaagcc agcatcacgg    1020

ttagcaatag attccagtgc caccccaggc aagcgctacg ggcgcaatgc gaggcgatgc    1080

tcaccctcgg cgctaccctc cgccactctt tggccagctc cacccatgcc gcccctgtct    1140

ggcgctgggc tttcagaacc tacgccgacc gagcacagcc ggcccgattg gtctggatcc    1200

tgaccctgcc gggcttcgtc gcccagtgtc ccatgctctg ggccagcggt tcgatctgct    1260

ccgctaactc gttgatgcct ctggatttcc tcactctgtc gattgcgttc atggtctatt    1320
```

-continued

```
gacccaaggc attcctgtca gtcgatgatc tgggcgttgg aggtgtagat gttcaggggc    1380
aagatctgcc ggtcggtgcg gatgcccagg acttccatcc aaaccaagtt aggccccagg    1440
tgaacacggg gcaggcggtc gatgccctgc gcgtcaagtg ttctgtggtc aatgcggcgg    1500
tcgtggcgag ccggctctaa tgcgcggttg gcatggtcgg ccgatgcctc gcgggtctgc    1560
tcaagccatg ccttgggctt gagcgcttcg gtcttctgtg ccccgccctt ctccggggtc    1620
ttgccgttgt accgcttgaa cccctgagcg gcgggccgct cgatgccgtc attgatccgc    1680
taggagatca tacggtggca gtgcgcgttc tagccgaacc aggcatggat ggccagcgta    1740
tacggaaggc gatcggaacc ggtcaggtgc tgggcgaacc aggacgaacg agcccaatgc    1800
tggtcgaggg tcagctcgac cggcaggcga aattcgaccc ccttgaacag ccgcccattg    1860
gcgcgttcat acaggtcggc agcatcccag tagtcggcgg gccgctcgac gaacccggcc    1920
atgtgcccgg attcggcgtg caagacttca tccatgtcgc gggcatactt gccttcgcgc    1980
tggatgtcgt cggccttggc cctggccgat tggccgcccg acctgctgcc ggtttccgcc    2040
gtcaggtgcc aactcgaccc gatgcctcgc tgttgcttta gatttcaggc tacccgaact    2100
ggccctagga gagcgaaccg aacgaagggc ggccgtaggc aacgactatc gacgagaacc    2160
cggcaagtgc gcgctccgcc acaaagtagg gtcgggattg ccgccgctgt gcctccatga    2220
tagcctacga gacagcacac taacaatggg gtgtcaagac ggttaagggg agcaacaagg    2280
cggcggatcg gctggccaag ctcgaagaac aacgagcgcg aatcaatgcc gaacttcagc    2340
gggagcgggc aagggaacag cagcaagagc gcaagaacga aacaaggcgc aaggtgctgg    2400
tgggcgaccc gatttaggcc aaggcgaccc gacgcgagtg gacggaggat aggcccatgg    2460
aggcaatgga tgcgcaactc gaccgcgacc aagcccgagc attgcccggt atgccgaacc    2520
gccagaagga tgagccgggc tgaacgatcg accgagacag gcgcggcggg gctgcacacg    2580
cgcccccacc cttcgggtag ggggaaaggc cgctaaagcg gctaaaagcg ctccagcgta    2640
tttctgcggg gtttggtgtg gggtttagcg ggcttcgccc gccttccccc ctgccgcgca    2700
gcggtggggc ggtgtgtagc ctagcgcagc gaatagacca gctatccggc ctctggccgg    2760
gaccatcggc aacgcgaagc agcgcccacc aagggcgatg atcaccgagc atcgccgatt    2820
attcctagac aatcatggac ggatctttcc aacaccccgc cagcccccgc ccctgctggg    2880
tttgcaggtc tgggcgcgtg acagccattg caggggttcg tgacagttat tgcagggggg    2940
cgtgacagtt attgcagggg ttcgtgacag ttagtacggg agtgacgggc actggctggc    3000
aatgtctagc aacggcaggc atttcggctg agggtaaacg aacttccggc taagcgatcg    3060
actgcatgta aaccaagtcc tgccaggccg aggcacatgc atccctggcc ggccaggacg    3120
gacagcaggg atcgggatcc tggtcgttcc cagacgcaac gacccgagcc aacccctatc    3180
tatcagatcg ttgaccagta ttacccggca ttcgccgagc ttatggcagc gacgggaacg    3240
gaacagcggg gctacgtgca acgggaattc gaagaatttc tccaacgcgg gcggcaggag    3300
aatgcatttc taccggttcg atgcgagtcc tgccacgacg agcacctggt agcacccagc    3360
aatcaatcta aagtatatat gagtaaactt ggtctgacag gccccagaat tcgcatctag    3420
aacgggtatc gatcagtacc actcaagcca aactgcaacc aaccgcaacg ggatccaact    3480
acaatcgatt agagattgtc atcgcacctg attgcccgac attatcgcga gttataaccc    3540
atttatacac atataaatca gcatccatgc tggaatttaa tcgcgcgagc ctatgaacac    3600
```

-continued

```
gattaacatc gctaagaacg acttctctga catcgaactg gctgctatcc cgttcaacac    3660
tctggctgac cattacggtg agcgtttagc tcgcgaacag ttggcccttg agcatgagtc    3720
ttacgagatg ggtgaagcac gcttccgcaa gatgtttgag cgtcaactta aagctggtga    3780
ggttgcggat aacgctgccg ccaagcctct catcactacc ctactcccta agatgattgc    3840
acgcatcaac gactggtttg aggaagtgaa agctaagcgc ggcaagcgcc cgacagcctt    3900
ccagttcctg caagaaatca agccggaagc cgtagcgtac atcaccatta agaccactct    3960
ggcttgccta accagtgctg acaatacaac cgttcaggct gtagcaagcg caatcggtcg    4020
ggccattgag gacgaggctc gcttcggtcg tatccgtgac cttgaagcta agcacttcaa    4080
gaaaaacgtt gaggaacaac tcaacaagcg cgtagggcac gtctacaaga aagcatttat    4140
gcaagttgtc gaggctgaca tgctctctaa gggtctactc ggtggcgagg cgtggtcttc    4200
gtggcataag gaagactcta ttcatgtagg agtacgctgc atcgagatgc tcattgagtc    4260
aaccggaatg gttagcttac accgccaaaa tgctggcgta gtaggtcaag actctgagac    4320
tatcgaactc gcacctgaat acgctgaggc tatcgcaacc cgtgcaggtg cgctggctgg    4380
catctctccg atgttccaac cttgcgtagt tcctcctaag ccgtggactg gcattactgg    4440
tggtggctat tgggctaacg gtcgtcgtcc tctggcgctg gtgcgtactc acagtaagaa    4500
agcactgatg cgctacgaag acgtttacat gcctgaggtg tacaaagcga ttaacattgc    4560
gcaaaacacc gcatggaaaa tcaacaagaa agtcctagcg gtcgccaacg taatcaccaa    4620
gtggaagcat tgtccggtcg aggacatccc tgcgattgag cgtgaagaac tcccgatgaa    4680
accggaagac atcgacatga atcctgaggc tctcaccgcg tggaaacgtg ctgccgctgc    4740
tgtgtaccgc aaggacaagg ctcgcaagtc tcgccgtatc agccttgagt tcatgcttga    4800
gcaagccaat aagtttgcta accataaggc catctggttc ccttacaaca tggactggcg    4860
cggtcgtgtt tacgctgtgt caatgttcaa cccgcaaggt aacgatatga ccaaaggact    4920
gcttacgctg gcgaaaggta aaccaatcgg taaggaaggt tactactggc tgaaaatcca    4980
cggtgcaaac tgtgcgggtg tcgataaggt tccgttccct gagcgcatca agttcattga    5040
ggaaaaccac gagaacatca tggcttgcgc taagtctcca ctggagaaca cttggtgggc    5100
tgagcaagat tctccgttct gcttccttgc gttctgcttt gagtacgctg gggtacagca    5160
ccacggcctg agctataact gctcccttcc gctggcgttt gacgggtctt gctctggcat    5220
ccagcacttc tccgcgatgc tccgagatga ggtaggtggt cgcgcggtta acttgcttcc    5280
tagtgaaacc gttcaggaca tctacgggat tgttgctaag aaagtcaacg agattctaca    5340
agcagacgca atcaatggga ccgataacga agtagttacc gtgaccgatg agaacactgg    5400
tgaaatctct gagaaagtca agctgggcac taaggcactg gctggtcaat ggctggctta    5460
cggtgttact cgcagtgtga ctaagcgttc agtcatgacg ctggcttacg ggtccaaaga    5520
gttcggcttc cgtcaacaag tgctggaaga taccattcag ccagctattg attccggcaa    5580
gggtctgatg ttcactcagc cgaatcaggc tgctggatac atggctaagc tgatttggga    5640
atctgtgagc gtgacggtgg tagctgcggt tgaagcaatg aactggctta agtctgctgc    5700
taagctgctg gctgctgagg tcaaagataa gaagactgga gagattcttc gcaagcgttg    5760
cgctgtgcat tgggtaactc ctgatggttt ccctgtgtgg caggaataca agaagcctat    5820
tcagacgcgc ttgaacctga tgttcctcgg tcagttccgc ttacagccta ccattaacac    5880
```

continued

```
aatcaaagct agcgagattg atgcacacca aacggagtcc ggtaacgatc ataccattgt    5940

aatcagacca gacggtagcc aactacgtca gactgaagtg tgggcaaacg agaagtaagg    6000

aatcgaatct tttgcactga ttcacgactc cttcggtacc attccggctg acgctgcgaa    6060

cctgttcaaa gcagtgcgcg aaactatggt tgacacatat gagtcctgtg atgtactggc    6120

tgatttctac gaccagttcg ctgaccagtt gcacgagtct caattggaca aaatgccagc    6180

acttccggct aaaggtaact tgaacctccg tgacatctta gagtcggact tcgcgttcgc    6240

gtaacgccgc aatcgttacc aagcaattca agtaaaaacc agctacggcg ggttatttta    6300

tggcgggacc ttcggcaatc agccattgtc agctgctgac gagacggatc agcacatggg    6360

aacccgctcc tcagcctcca gagtagcgtc tcaagcagtt acagagatgt tacgaacccc    6420

caggacatcc gagaatgcga gcggatggag ggtacaaccg ggagagtgtt caccgacaaa    6480

caacagatca aacaaaaggc ccagtcttcc gactgagcct tttgttttat ttgatgtctg    6540

gcagttcccc gccgctatct gatcatcaaa taataatttc ttcatgttga aaatctccaa    6600

aacaaaaggc taccaaagga gactaaacta gtatcggtta atacgcttgc ttttacatac    6660

atgccagcta taccgtaggg aacccagtcg tgccagatgc attaacgaat aggccaaacgc    6720

gcgggcagag gcggtttgcg tattgggcgc caggctggtt tttcttttca ccagtgacac    6780

cggcaacagc tgattgcccc tcacgccctg gccctgagag agttgcagca agcggtccac    6840

gctggtttgc cccagcaggc gaaaatcctg tttgatggtg gttaacggcg ggatataaca    6900

tgagctgtct tcggtatcgt cgtatcccac taccgagata tccgcaccaa cgcgcagccc    6960

ggactaggta atggcgagca ttgcgcccag agccatatga tagttaggat aaagcatagc    7020

agtggaaacg atgccatcca tatgcatttg aatggattga tgaaaaaagg aactggaact    7080

ccagtcgccc tcccgttccg ctatcggctg aatttgattg cgagtgagat atttatgcca    7140

gccagccaga cgcagacgcg ccgagacaga acttaatggg cccgctaaca gcgcgatttg    7200

ctggtgaccc aatgcgacca gatgctccac gcccagtcgc gtaccgtctt catgggagaa    7260

aataatactg ttgatgggtg tctggtcaga gacatcaaga aataacgccg gaacattagt    7320

gaaggcagct taaacagcaa tggcatactg gtaatcaagc ggatagttaa tgatcagccc    7380

actgacgcgc tgcgcgagaa gattgtgcac cgccgcttta caggcttcga cgccgcttcg    7440

ttctaccatc gacaccacca cgctggcacc cagttgatcg gcgcgagatt taatcgccgc    7500

gacaatttgc gacggcgcgt gcagggccag actggaggtg gcaacgccaa tcagcaacga    7560

ctgtttgccc gccagttgtt gtgccacgcg gttgggaatg taattcagct ccgccatcgc    7620

agcaaaacaa ttttcaagcg ttttcgcaga aacgtggatg gactggttca aaacgcgggc    7680

aacggtatga tacgagacac aggcaaaacc tgcgacatcg tataacgtta atggcataat    7740

gagcaaatca aatgcatctc gtgtacttgc gtaacgtaac atatccatag tttacgtatc    7800

acagcactat gtgattatac cagcccactc actacatgtc aagaataaac tgccaaagcc    7860

gagcggacgc tttttgcccg aataccttgc tgatctagcc tttcacgtag tggacaaatt    7920

cttccaactg atctgcgcgc gaggccaagc gatcttcttc ttgtccaaga taagcctgtc    7980

tagcttcaag tatgacgggc tgatactggg ccggcaggcg atccatgcc aagtcggacg     8040

cgacatcctt cggcgcgatt ttgccggtta ctgcgctgta ccaaatgcgg gacaacgtaa    8100

gcactacatt tcgctcatcg ccagcccagt cgggcggcga gttccatagc gttaaggttt    8160
```

-continued

```
aatttagcgc atcaaatcga tactgttacg gacccggatc aacgagttcc tacgccgatg    8220

gacctaacca ggccaagcta tgttcaattg atttgttacg aacgatagcc agatcaatgt    8280

cgatcgtggc tggctcgaag atacccgcaa gaatgtcatt gcgctgccat tctccaaatt    8340

gcagttcgcg cttagctgga taacgccacg gaatgatgtc gtcgtgcaca acaatggtga    8400

cttctacagc gcggagaatc tcgctctctc caggggaagc cgaagtttcc aaaaggtcgt    8460

tgatcaaagc tcgccgcgtt gtttcatcaa gccttacggt caccgtaacc agcaaatcaa    8520

tatcactgtg tggcttcagg ccgccatcca ctgcggagcc gtacaaatgt acggccagca    8580

acgtcggttc gagatggcgc tcgatgacgc caactacctc tgatagttga gtcgatactt    8640

cggcgatcac cgcttccctc ataatgttta actttgtttt agggcgactg ccctgctgcg    8700

taacatcgtt gctgctccat aacatcaaac atcgacccac ggcgtaacgc gcttgctgct    8760

tggatgcccg aggcatagac tgtaccccaa aaaaacagtc ataacaagcc atgaaaaccg    8820

ccactgcgcc gttaccaccg ctgcgttcgg tcaaggttct ggaccagttg cgtgagcgca    8880

tacgctactt gcattacagc ttacgaaccg aacaggctta tgtccactgg gttcgtgcct    8940

tcatccgttt catcggtgtg cgtcaccagg caaccttggg tagcagcgca gtcgaggcat    9000

ttctgtcctg gctggtcatg acgaaaatgc cttaacgtga gtcagcctgc cgcctggggc    9060

cgggtgatgt cgtacttgcc cgccgcgaac tcggttacgg tccagcccag cgcgaccagc    9120

tccggcaacg cctcgcgcac ccgcttgcgg cgcttgcgca tggtcgaacc actggcctct    9180

gacggccaga catagccgca caggtatctt atggaagcct tgccggtttt gccggggtcg    9240

atccagacac aatgcagctg gtgcagcagg aggggcggtt agctgttatg agcccgaacc    9300

tagtccatgc tgatgcgcac atgctggccg aacccatga aggcctgcgc gatcaagggg    9360

ttcaggccca cgtacaggcg cccgtccgcc tcgtcgctgc cgtactccga cagcagccga    9420

aaccccctgcc gcttgcggcc attctggcgc atgatggata ccttccaaag gcgctcgatg    9480

cagtcctgta tgtgcttgag cgccccacca ctatcgacct ctgcccgcat ttcctttgcc    9540

agcgcccgct agctaccttt gacccacatgg cattcagcgg tgacggcctc ccacttgggt    9600

tacaggatca gacggagctg cagtccgccc tagggtcttg gttcagggcc aagcactagg    9660

ccattaggcc cagccatggc cacgagccgt tgcaggatgc gcagatcatc agcgcccagc    9720

ggctccgggc cgctgaactc gatccgcttg ccgtcgccgt agtcatacgt cacgtccagc    9780

ttgcttgcgct tgcgctcgcc ccgcttgcgg gcacggaaca ggccgggggc cagacagtgc    9840

gccgggtcgt gccggacgtg gctgaggctg tgcttgttct taggcttcac cacggggcac    9900

caccatgctc ttgcgatgcc tatccagccc ggcggtattg agctcccagc agtcatgacg    9960

aatgaacccc agctcagcga acggtgagcc atcgttggcc ttgctcaccc agccgcggtc    10020

gacgaaacgg agctggtcgt agtccacccc aactcactgg gactcggagc tggtcatgct    10080

cgacaggtag gactgccagc ggatgttatc gaccagtacc gagctgcccc ggctggcctg    10140

atgctggtcg aatgcgacca tactggacgc gacctcgatg gactggtgct ggctcaagct    10200

agagcacccg gtatcggcgg cgatggcctc catgcgaccg atgacctggg ccatggggcc    10260

gatgggtttc tattcatcga tgtggaaccg gagcagcgtg taccgcacct tacggcgggg    10320

gccctcggcg gcgcgcttga ggccgtcgaa ccactccggg gccatgatgt tgggcaggct    10380

gccgatcagc ggctggatca gcaggccgtc agcgacggct tgccgttcct cggcgctgag    10440
```

continued

```
gtgcgaacca agggcgtgca ggcggtgatg aatgacgggtg ggcgggtatt aggcgggacg   10500

gtagataccc gggccggtgg gacgacagcc aacccccagc agatccggcc agcccgaact   10560

ctgctgcggc agttgcaggg ccagcatgga tttaccggca ccaccgggcg acaccagcgc   10620

cccgaccgta ccggccacca tgttggcaa aacgtagtcc agcggtgggg gcgctgctgc   10680

gaacgcctcc agaacattga taggcatatg ggtagccatt gattgcctcc tttgcaggca   10740

gttggtggtt aggcgctggc gggcacata ccccgccct gcgccgctct gagttcttcc   10800

aggcaatcgc gaacgcactc gtatacgtcg taggacagcc agaccatgcg atgacgaatc   10860

aataaggcca taatgagctc ggccaatcgc gattggagta cagcgaaagg gatggccagc   10920

aggtcgcggg tctgcttgtc cttttggtct ttcatatcag tcaccgagaa acttgccggg   10980

gccgaaaggc ttgtcttcgc ggaacaagga caaggtgcag ccgtcaaggt taaggctggc   11040

catatcagcg actgaaagc ggccagccte ggcctgtt gacgtatac caaagccacc   11100

gggcaaccca tagcccttg cactttgat caggtagacc gaccctgaag cgcttttttc   11160

gtaaaaaata aaaccaacta atgtgcgtga gtaacgatag tataacaggc gtgagtaaca   11220

aagcaagcac taacagatga aatcagcccc gaaccagtac atgcaaagat gggcgggtgc   11280

cggtgcccgt gccagctcg                                                 11299
```

<210> SEQ ID NO 82
<211> LENGTH: 6039
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pR K54'

<400> SEQUENCE: 82

```
gcactgcggc cccggctggc cagcttctgc gcggccataa agtcggactt gctgaggtca      60

tgaccgaagc gcttgaccag cccggccatc tcgctgcggt actcgtccag cgcctgcgcc     120

cggtggcggc taagctgccg ctcgggcagt tcgaggctgg ccagcctgcg ggccttctcc     180

tgctgccgct gggcctgctc gatctgctgg ccagcctgct gcacccagcgc cgggccagcg     240

gtggcggtca tgcccttgga ttcacgaagc agcaccaacg gatgataacc gggcgcggtg     300

gtgtgcttgt ccttgcggtt ggtgaagccc gccaagcggc catagtggcg gctgtcgggg     360

ctggccgggt cggcgtcgta ctcgctggcc agcgtccggg caatctgccc ccgaagttca     420

ccgcctgcgg cgtcggccac cttgaccat gcctgatagt tcttcgggct ggtttccact     480

accagggcag gctcccggcc ctcggctttc atgtcatcca ggtcaaactc gctgaggtcg     540

tcaacaagca aacgaaactg aagcacatgc taggcaggcc tgatataaac gtaaatgacc     600

tgggcattca taagcttgag aactggatgc ttatgagaca attcagaggc tgaccattcc     660

aggaaaatca tatgaaaggc gggtacgtac atgacgaaga tatcgaagcg atacaagacc     720

atggccttga gctgtcggcc tatggcctgc aaagtcctgt cgttctcat cgggccacca      780

agcacagcca gatcgagccg tactcagttg taagagagc aaggacgagc aagaacaacg      840

atgcgatccg cagcaccacc gtaggcatca tggaagccag catcacggtt agccatagct     900

taaagtgcca aaaccgagac gagacaaggc agatcagagc gggcagatc aaatcagagg      960

ctaccctcccg caacttcttg gccagctcca ccca tgccgc cccttgtcgg cgctgggctt   1020

tcagccactc cgccgcctgc gcctcgctgg cctgcttggc ctggctcatg acctgccggg    1080
```

-continued

```
attcgtaggc aagtgtagcc atgcctaggg aaagcggttc gatctgatcc gattactagt    1140
tgattgaatc ggatttattc aatctgtcga ttgcgttatt ggtctattgc atcccggtat    1200
tcctgtaagt cgatgatctg gggcttggcg gtgtcgatgt tcagggccac gtctgcccgg    1260
tcggtgcgga tgcccggcgc ttccatctcc accacgttcg gcgccaggtg aacaccgggc    1320
aggcgctcga tgccctgcgc ctccagtgtt ctgtggtcca tgcgggcgtc gtggccagcc    1380
cgcttcaatg cccggttggc atggtcggcc catgcctcgc gggtcttgtc aagccatgcc    1440
ttgggattga gagcttaggt attctgtgcc aagcccttct aagggtattt gacgttgtac    1500
agcttgatcc aatgagaggc gggccgatcg atgccgtatt tgatccgatc ggtgattatc    1560
aggtggcagt gcgggttctc gccgccaccg gcatggatgg ccagcgtata cggcaggcgc    1620
tcggcaccgg tcaggtgctg ggcgaactcg gacgccagcg ccttctgctg gtcgaggtcc    1680
agcttgaccg gcagggcaaa ttcgacctcc ttgaacagcc gcccattggc gcgttcatcc    1740
aggttggcag catcccagta gtcggcgggc cgctcgacga actccggcat gtgcccggat    1800
taggcgtgca agtcttactc aatgtcgcgg gattacttgc attggcgatg gatgtagtcg    1860
gacttggccc tggccgattg gacgccagtc atgctgaagg ttttcgaagt aaggtgatcc    1920
atcgccatgc tgcctcgctg ttgcttttgc ttttcggctc catgcaatgg ccctcggaga    1980
gcgcaccgcc cgaaggtggc ccgttaggcc agtttctcga agagaaacgg gtaagtgcgc    2040
cctcccctac aaagtagggt cgggattgcc gccgctgtgc ctccatgata gcctacgaga    2100
cagcacatta acaatggggt gtccagatgg ttaagggggg caacaaggcg gcggattggc    2160
tggccaagct agttgaacta agtgcgtgta ttattgaaga aattcagagg gagcggattt    2220
gggaatagca gattgagcgc aagtacgata aatggcgata ggtgctggtg ggggccatgc    2280
ttttggcgaa ggtgaacagc agcgagtggc cggaggatcg gctcatggcg gcaatggatg    2340
cgtaccttga acgcgacgac gacggcgcct tgttcggtct gccgccacgc cagaaggatg    2400
agcccggctg aatgatcgac cgagacaggc cctgcggggc tgcacacgcg cccccaccct    2460
tcgggtaggg ggaaaggccg ctaaagcggc taaagcgctt ccagcgtatt tctgcggggt    2520
ttggtgtggg gtttagaggg atttgcaagc atttcctatt gacgcgtagc ggtgggtagg    2580
tgtgtagcct agcgcagcga atagaccagc tatccggcct ctggccgggc atattgggca    2640
agggcagcag cgcgccacaa gggcgctgat aaccgcgcct agtggattat tcttagataa    2700
tcatggatgg atttttccca cacccccgcca gcccccgccc ctgctgggtt tgcaggtttg    2760
ggggcgtgac agttattgca ggggttcgtg acagttattg caggggggcg tgacagttat    2820
tgcaggggta agtgaaagta agtacgggag tgatggaatc tggctggatt tgtctagatt    2880
aggcaggctt ttcggatgag ggtaaaagta atttccgata agcgatagtc tgtatgtatt    2940
aacagtattg aatggaagcg gaccatgcct aatgatgagg aatggaaggc aagccgggat    3000
cgggatactg gtcgttacca gagccacgga cccgagcaaa cccttctcta tcagatcgtt    3060
gacgagtatt aaccggaatt agctgcgctt atggcagagc agggaaaggt attgccgggc    3120
tatgtgcaac gggaatttga agaatttctc caatgcgggc ggctggagca tggctttcta    3180
agggtagctt gagagtattg aatggcagcg aacctggtcg atttcagatt tatttctatt    3240
gtatatatga gtaaacttgg tctgacaggc ccctgaattc gcatctagat ggtagagcca    3300
caaacagccg gtacaagcaa cgatctccag gaccatctga atcatgcgcg gatgacacga    3360
```

-continued

```
aatcaagcag gagcaaacag aacttaaccc aaagcaaaga aactgcgacc aagtggcact    3420
tagccgaccc aacgcatcac tgacctggcc gatcaatgca gacgaacgtc ggtcagtttc    3480
accggattta cgtaaaaacc cgcttcggcg ggtttttgct tttggagggg cagaaagatg    3540
aatgactgtc aaattaatac gactcactat agggagagga attgtgagcg gataacaatt    3600
ccatacagaa acagaggaga tatatgagta aaggagaaga gcttttcact ggagttgtcc    3660
caatcctcgt ggaattagac ggtgatgtta atgggcacaa gttctctgtc agtggagagg    3720
gtgaaggcga agcaacatac ggcaagctga cccttaaatt tatttgcacc actgggaaac    3780
tacctgttcc atggccaaca cttgtcacta cgttcgggta tggggtacag tgcttcgcgc    3840
gctacccaga tcacatgaaa cagcacgact ttttcaagag tgccatgccc gaaggctatg    3900
tacaggagag aaccatcttt tttaaggatg acggcaacta taagacacgc gccgaagtga    3960
agttcgaggg tgataccctt gttaatagaa tcgagttaaa gggtattgac tttaaggaag    4020
atggaaacat tttaggccac aaactggaat ataactataa ctcccataat gtgtacatta    4080
tggccgacaa gcaaaagaac ggtatcaagg ttaacttcaa gatcagacac aacatcgagg    4140
atggaagcgt tcaactagcc gaccattacc aacaaaacac accaattggc gatggccctg    4200
tgctgttacc agacaaccat tacctgtcca ctcaatctgc cctttcgaaa gatcccaacg    4260
aaaagcgcga ccacatggtc cttcttgagt ttgtcacggc tgctgggatt acacacggca    4320
tggatgaact atacaaataa atccctgcta acaaagcccg aaaggaagct gagttggctg    4380
ctgccaccgc tgagcaataa ctagcataac cccttggggc ctctaaacgg gtcttgaggg    4440
gttttttgct gaaaggagga actatatgct gaccccaatg aaaaccaacc aatagttagg    4500
atcacaagga aaaaagaatg agaagccagg aggcggagct tatcaaaaag gatcttcacc    4560
tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact    4620
tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt    4680
cgttcatcca tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta    4740
ccatctggcc ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta    4800
tcagcaataa accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc    4860
gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat    4920
agtttgcgca acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt    4980
atggcttcat tcagctccgg ttcccaacga tcaaggcgag ttacatgatc ccccatgttg    5040
tgcaaaaaag cggttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca    5100
gtgttatcac tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta    5160
agatgctttt ctgtgactgg tgagtactca accaagtcat tctgagaata gtgtatgcgg    5220
cgaccgagtt gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact    5280
ttaaaagtgc tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg    5340
ctgttgagat ccagttcgat gtaacccact cgtgcaccca actgatcttc agcatctttt    5400
actttcacca gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga    5460
ataagggcga cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc    5520
atttatcagg gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa    5580
caaatagggg ttccgcgcac atttccccga aaagtgccac ctgtcatgac caaaatccct    5640
```

-continued

```
tacggtgagt aagcctgccg aattggccgg ggtgatgtcg tactggacgg aagcgaactc    5700
ggttaaagtc aagccaagcg agaccagctc aggcaaagcc tagcgcaacc gattgcgggg    5760
cttgcgcatg gtcgaaccac tggcctctga cggccagaca tagccgcaca aggtatctat    5820
ggaagccttg ccggttttgc cggggtcgat ccagccacac agcggctggt gcagcagggg    5880
ggcggtttcg ctgtccagcg cccgcacctc gtccatgctg atgcgcacat gctggccgcc    5940
acccatgacg gcctgcgcga tcaaggggtt cagggccacg tacaggcgcc cgtccgcctc    6000
gtcgctgggc tactcagcca gacgccgcaa aacctgaagc ttgcggaact tatgggagct    6060
gatggatacc ttccaaaggc gatcgatgca gtcctgtatg tgcttgagcg accaacaact    6120
atcgacctcc gccccgattt ccttcgccag cgcccgatag ctacctttga ccacacggca    6180
ttcagcggtg acggcctccc acttggcttc caggaacagc cggagctgcc gtccgccttc    6240
ggtcctgggt tccgggccca gccctaggcc attagggcca gccatggcca ccagcccttg    6300
caggatgcgc agatcatccg cgcccagcgg ctccgggccg ctgaactcga tccgcctgcc    6360
gtcgcagtcg tactaagtca agtccagctc gatgcgcttg agctcgaacc gattgagggc    6420
aaggaaacgg aaggcggcca gaccgtgcgc agggtcgtgc aggacgtggc tgcagctgtg    6480
cttgctctta ggctccacca cggggcaccc ccttgctctc gcgccgcctc tccagcacgg    6540
cggccttgag caccccggcg tcatgccgcc tgaaccaccg atcagcgaac ggtgcgccat    6600
agttggcctt gctcacaccg aagcggacga agaaccggcg ctggtcgtcg tccacacccc    6660
attcctcggc ctcggcgctg gtcatgctcg acaggtagga ctgccagcgg atgttatcga    6720
aacgtaacga gatgcaacgg atggcctgcc gatggtagcc tgcgccaatc atggccgcgc    6780
aatagatggc atggtgacgg aacccgatcg agcacccggc atcgccgggc atggcctacc    6840
tgcgaccgat gacctgggcc atgggcccgc tggcgttttc ttcctcgatg tggaaccggc    6900
gcagcgtgtc cagcaccatc aggcgcgggc cctcgccggc gcgcctgagg ccgtcgaacc    6960
actccggggc catgatgttg ggcaggctgc cgatcagcgg ctggatcagc aggccgtcag    7020
ccacggcttg ccgttcctcg gcgctgaggt gcgccccaag ggcgtgcagg cggtgatgaa    7080
tggcggtggc aggggattcg gaggccaggc agatcaaagg gacggtgggc agttcgaacc    7140
cctccagcag atcggcgccg cctgcaatcc gtgcggccag ttgcaggcgc agcacggatt    7200
tacggcacgc acgggcgacc accagcgccc cgaccgtacc ggccaccatg ttgggcaaaa    7260
cgtagtccag cggtggcggc gctgctgcga acgcctccag aatattgata ggcttatggg    7320
tagccattga ttgcctcctt tgcaggccgt tggtggttcg gcgctggcgg ggtcactacc    7380
aacgcaatgc gacgctatga gttcatcacg gaacacgagc agcgcctagt attcgtagtc    7440
ggtcagcacg aactagagct gacgcatccc tttggcattc atgcgctagg aataacgagc    7500
ttggcgtaca gagtcagggc tggccagcag gtcgccggtc tgctagtact tttggtattt    7560
catatcagtc accgagaaac ttgcccgggc cgaaagcctt gtcttcgcgg aacaagcaca    7620
agggcaagcc gtccaggtta aggcatgcca tatcagagcc tgcaaagagg aacgcctagg    7680
cctcgtttga cgtacaacca aagccaccgg gcaacccata gcccttgtca ctttcgatca    7740
ggtagaacga aactgaagcg attttatcgc attccataca aaccccttct gtgcgcgagt    7800
actcatagta taacaggcgt gagtaccaac gcaagcacta catgcctgaa tctggcccgc    7860
ccgcgtccat gccccgctgg cgggcgcccg gtgcccgtgc cagccgggcc cgcgcaagct    7920
```

-continued

```
ggacgatggg aagacaaatg aactagatga aggtgcgatc gatgtaatcc gattcgtggc    7980

agggcttgcg atatgaaagc gatgggatgg aatcggaaaa ggcctagacg atttcctag     8039


<210> SEQ ID NO 83
<211> LENGTH: 12068
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized. pBTK350d

<400> SEQUENCE: 83

atggaaactg aggccaaggc tggccagctc atgcgcggcg ataaagtagc aattgctgag      60

gtcatgaccg aagcgcttga ccagcccggc catctcgctg cggtactcgt ccagcgccgt     120

gcgccggtgg cggctaagct gccgctcggg cagttcgagg ctggccagcc tgcggggctt     180

ctcctgctgc cgctgggcct gctcgatctg ctggccagcc tgctgcaccc gcgccgggcc     240

agcggtggcg gtcttgccct tggattcacg cagcagcacc cacggctgat aaccggcgcg     300

ggtggtgtgc ttgtccttgc ggttggtgaa gcccgcaagc aggccatagt ggcggctgtc     360

ggcgctggcc gggtcggcga agtactcgcc ggccaggatc agggcaatct gaccccgacg     420

ttcaccgccc gcggcgtcgg ccaccttgac ccatgcctga tagttcttcg ggctggtttc     480

cactaccagg gcaggctccc ggccctcggc tttcatgtca tccaggtcaa actcgctgag     540

gtcgttcacc agcaccagac catgccgctc ctgctcgggc ggcctgatat ccacgtcatt     600

gccctgggca ttcatccgct tgagccatgg cgtgttctgg agcacttcgg cggctgaccc     660

ttcccggttc atcaaatggc aggtggtgcg gtccccgacg aagaaatagc agcgcaaacc     720

gacccatggc ttgacgatgc ggccaatggc atgcaaagtc atgtcgttct taatcgggcc     780

accaagcgca gccagatcga gccgtcctcg gttgtcagtg gcgtcaggtc gagcaagagc     840

aacgatgcga tcagcagcac caccgtaggc atcatggaag ccagcatcac ggttagccat     900

agcttccagt gccacccccg cgacgcgctc cgggcgctct gcgcggcgct gctcacctcg     960

gcggctaccc cccgcaactc tttggccagc tccacccatg ccgcccctgt ctggcgctgg    1020

gataaagccc aatccgaagc atgcgcatcg atggcctgca tggtcaggct aatgacatgc    1080

cgggcttcgc cggccagtgc cgccatgctc tggcccagcg gttcgatctg ctcggctaac    1140

tcgccgatgc ctcgcgattc cttcactctg tcgatacgcc tcatggtcta ttgcctcccg    1200

gtattcctgt aagtcgatga tctgggcgtt ggcggtgtcg atgttcaggg ccacgtctgc    1260

ccggtcggtg cggatgcccc ggccttccat ctccaccacg ttcggcccca ggtgaacacc    1320

gggcaggcgc tagacgaccc gagcctaacc tgttcagtgg taaccgaggg agtcgaggcc    1380

agcccgatca aatgcaaggc tggcaaggtc ggcccatgcc tagcggtgct gatcaagacc    1440

tgccatgggc ttgacgagca aggtcatctg tgccccgacc ttctccgggg tattgcagtt    1500

gtaccgcttg aacaactgag cggcggcgcg ctcgatgccg tcattgatcc gctcggagat    1560

aatcaggtgg aagtgaaggc tatcgcagcc aacggcatgg atggccagcg tataccgacg    1620

gcgctcggca ccggtcaggt gctgggcgaa ctcggacgcc agcgccttct gctggtcgag    1680

ggtcagatcg aacggaaggg aaaaaaagcc atccaagacc agccgcaaat tggcgcgttc    1740

atccaggtcg gcagcatccc agtagtcggc gggccgctcg acgaactccg gcatgtgccc    1800

ggattcggcg tgcaagactc catccatgtc gcgggcatac ttgcccttcg gctggatgta    1860
```

continued

```
gtcggcattg gaccgggccg attggccagcc agcccgatg aaggcchtcg aagtaaggtg    1920
attaatagcc atgccgactc gatgccgctc ttgccttcg gatccatgcc atggcccatcg    1980
gagagcgcac cgcccgaagg gtggccgtta ggccagtttc tcgaagagaa acoggtaagt    2040
gcgccctccc ctacaaagta gggtcgggat tgccgccgcc gtgccccat gatagcctac    2100
gagacagcac attaacaatg gggtgtcaag atggttaagg ggagcaacaa ggcggcggat    2160
cggctggcca agctcgaaga acaacgagcg cgaatcaatg ccgaaattca gcgggagcgg    2220
gacaggggac agcggacga gagcaagcac gaccaaggc gacaggtgct ggtgcgggcc    2280
atgatttggc aacaggtgca cagcagcagc tggccggagg atcggctact ggcggcaatg    2340
gatgcgtacc ttgaacggga ccacgacggc gcctcgttcg gtctgccgcc acgccagaag    2400
gatgagcggg gctgaatgat cgaccgagac aggccctgcg gggccgcaca cgcgcccca    2460
ccctacgggt aggggaaag gccgctaaag cggctaaaag cgctccagcg tatttctgcg    2520
gggtttggtg tggggtttag cgggctttgc ccgcctttcc ccctgccgcg cagcggtggg    2580
gaggcgtgta gactagagca gagcatagcc cagcatacg gactcaggcc gggcatattg    2640
ggcaaggggca gacgcgaccc aacaggcgc tgataaaagc gactagtggc ttatactttg    2700
ataatcatgg atggattttc ccaacacccc gccagccccc gccccgctg ggtttgcagg    2760
tttgcgggcg tgacagttac tgcagggggtt cgtgacagtt attgcagggg ggcgtgacag    2820
ttattgcagg ggttcgtgac agttagtacg ggagtgacgg gcactggctg gcaatgtcta    2880
gcaacggcag gcattcggc tgagggtaaa agaactttcc gctaagcgat agactgtatg    2940
tatacaatgc attgcaagga agcggaacct gactcatgtg gaggccaggc aggccagacg    3000
ggacggggac actggtagtc accagagcca aagacccagc caacccattc tataccagat    3060
cgtcgacgag tattaccgg cattcgctgc gcttatgca gagcagggaa aggaatgcc    3120
gggctatgtg caacgggaat ttgaagaatt tctccaatgc gggcggctgg agcatggctt    3180
tctacgggtt cgctgcgagt cttgccacgc cgagcacctg gtcgctttca gaaatcaatc    3240
taaagtatat atgagtaaac ttggtctgac aggcccctga attcgcatct agacagggtc    3300
ttgctatgtc aacctaacga aaacctgccc atcaacgttc aaccctatg attccggatc    3360
gtacgttgtg tggaattgtg agcggataac aatttcaaag aggagaaatt aactatgaac    3420
acgattaaca tcgctaagaa cgacttctcc gacatcgaac tggctgctat cccgttcaac    3480
actctggctg accattacgg tgagcgttta gctcgcgaac agttggccct tgagcatgag    3540
tcttacgaga tgggtgaagc acgcttccgc aagatgtttg agcgtcaact taaagctggt    3600
gaggctgcgg ataacgatgc agccaagccc atcatcacta ccctactccc taagatgatt    3660
gacacgatca acgactggtc tgaggaagtg aaagctaagc gaggcaagcg accgacagcc    3720
ttccagttcc tgcaagaaac aaagccggaa gccgtagcgt acatcaccat taagaccact    3780
ctggcttgcc taaccagtgc tgacaataca accgttcagg ctgtagcaag cgcaatcggt    3840
agggccattg aggacgaggc tagcttaggc agtatccgtg acctggaagc taagcacttc    3900
aagaaaaacg ttgaggaaca actcaacaag cgcgtaggc acgtctacaa gaaagcattt    3960
atgcaagttg tagaggatga aatgctatct aagggtctac taggtggcga ggcgtggtct    4020
tcgtggcata aggaagactc tattcatgta ggagtacgct gcatcgagat gctcattgag    4080
tcaaccggaa tggttagcct acaccgccaa aatgctggcg tagtaggtca agaactctgag    4140
```

-continued

```
aatatagatc tagcaaatga atacgctgag gatatcgata aacgtgtagg tgcgctggct    4200
ggcaatatct agatgttcca aactgagata gttcctaata agccgtggcc tggcaatact    4260
ggtggtggct attgcgctaa cggtcgtcgt cctctggcgc tggtgcgtac tcacagtaag    4320
aaagcactga tggcgtacga agacgtttac atgcctgagg tgtacaaagc gattaacatt    4380
gcgcaaaaca ccgcatggaa aatcaacaag aaagtcctag cggtcgccaa cgtaatcacc    4440
aagtggaagc attgtccggt cgaggacatc cctgcgattg agcgtgaagc actcccgatg    4500
aatccggatg aatcgaccta gaatcctgag gatatcaacg agtggaaacg tgctgcagct    4560
gatgtgtacc gatatgacta ggctcgatag tatcgcagta tacgccttgc gttcatgatt    4620
gagcaagcca ataagtttgc taaccataag gccatctggc tcccttacaa catggactgg    4680
cgcggtcgtg tttacgctgc gtcaatgttc aacccgcaag gtaacgatat gaccaaagga    4740
ctgcttacgc tggcgaaagg taaaccaatc ggtaaggaag gttactactg gctgaaaatc    4800
cacggtgcaa actgtgcggg tgtcgataag gttccgttcc ctgagcgcat caaggtcatt    4860
gaggaaaacc aagagaacat aatggcttgc gataagtatc aactggagca aactaggtgg    4920
gatgagatag attcaaagtc atgcataata gagtatcgca ttgagtaagc tggggtaatg    4980
cacaacggcc tgagctataa ctgctccctt ccgctgcgct ttgacgggtc ttgctctggc    5040
atccagcact tctccgcgat gctccgagat gaggtaggtg gtcgcgcggt taacttgctt    5100
cctagtgaaa ccgttcagga catctacggg attgttgcta agaaagtcaa cgagattcta    5160
caagcagacg caatcaatgg gaccgataac gaagtagtta ccgtgaccga tgagaacact    5220
ggtgaaatct atgagaaagt caagctgggc actaaggcac tggctggtca atggctggct    5280
tacggtgtta ctcgcagtgt gactaagcgt tcagtcatga cgctggctta cgggtccaaa    5340
gagttcggct tccgtcaaca agtgctggaa gataccattc agccagctat tgattccggc    5400
aagggtctga tgttcactca gccgaatcag gctgctggat acatggctaa gctgatttgg    5460
gaatctgtga gcgtgacggt ggtagctgcg gttgaagcaa tgaactggct taagtctgct    5520
gctaagctgc tggctgctga ggtcaaagat aagaagactg gagagattct tcgcaagcgt    5580
tgcgctgtgc attgggtaac tcctgatggt ttccctgtgt ggcaggaata caagaagcct    5640
attcagacgc gcttgaacct gatgttcctc ggtcagttcc gcttacagcc taccattaac    5700
accaacaaag atagcgagat tgatgcacac aaacaggagt ctggtatcgc tcctaacttt    5760
gtacacagcc aagacggtag ccaccttcgt aagactgtag tgtgggcaca cgagaagtac    5820
ggaatcgaat cttttgcact gattcacgac tccttcggta ccattccggc tgacgctgcg    5880
aacctgttca aagcagtgcg cgaaactatg gttgacacat atgagtcttg tgatgtactg    5940
gctgatttct acgaccagtt cgctgaccag ttgcacgagt ctcaattgga caaaatgcca    6000
gcacttccgg ctaaaggtaa cttgaacctc cgtgacatct tagagtcgga cttcgcgttc    6060
gcgtaaatcc gtaatcgtta atccgcaaat aacgtaaaaa cccgcttcgg cgggtttttt    6120
tatgggggga gtttagggaa agagcatttg tcagcagatg aactggacga tactggagca    6180
cgaacgtcgg tcagttcaac ctgatttacg taaaaacccg cttcggcggg ttttttgctt    6240
tggaggggca gaaagatgaa tgactgtcca atttatagga atcacatatg ggagaggact    6300
tgtgagcgga taacaattcc atacagaaac agaggagata tatgagtaaa ggagaagagc    6360
ttttcacagg agttgtccca atcctcgtgg aattagacgg tgatgttaat gggcacaagt    6420
```

continued

```
tatcgtcacg tgggagggga gacggcgacg aacaatatgg aacgcagacc atttaattta   6480
tttgcaaacc gggtaaacta aatgcaacca ggccaaaacc ggtcactacg ttcgggtatg   6540
gggttcagtg ctttgcggcc tacccagatc acatgaaaca gcacgacttt ttcaagagtg   6600
caatgcccga aggctatgta caggagagaa ccatcttttt taaggatgac ggcaactata   6660
agacacgcgc cgaagtgaag ttcgagggtg ataccctggt taatagaatc gagttaaagg   6720
gtattgactt taaggaagat ggaaatattt taggccacaa actggaatat aactataact   6780
aacataatgc gtacattatg gacgacaagc aaacgaacgg tatcaaggtt aactacaagc   6840
tacgaaacca aattgaggca ggaagcgttc aactagaaga aacttaaaca aacaacaacc   6900
caattgggga tgggcctgtg ctgttaccag acaaccatta cctgtccact caatctgccc   6960
tttcgaaaga tcccaacgaa aagcgcgacc acatggtcct tcttgagttt gtcacggctg   7020
ctgggattac acacggcatg gatgaactat acaaataaat ccctgctaac aaagcccgaa   7080
aggaagctga gttggctgct gccaccgctg agcaataact agcataaccc cttggggcct   7140
ctaaacgggt cttgaggggt tttttgctga aaggaggaac tatatccgga tatcccgcaa   7200
aacccaacca tagtcaggca accaaggacc aaagcagtta aaggcatgtt aaggaacacc   7260
aggacatccg agaatgcgag gcgatggagg gtacaaccgg gagagtgttc acggacaaac   7320
aacagataaa acaaaaggcc cagtcttccg actgagcctt ttgttttatt tgatgcctgg   7380
cagttccccg ccgctatctg atcaatcaaa aataattcct tcatgatgaa aatctccaaa   7440
aaaaaaggct cccaaaggag ccttcaattg tatcggttta tcagcatgct tttacatcac   7500
tgcccgattc aaagcaggga aacccagtcg gaaagctgca ttaacgaatc ggccaacgcg   7560
aggggagagg aggtatgcgt attgggcgcc agggtggttt ttctttcacc agtgagaacg   7620
ggcaacagct gattgccctt caccgcctgg ccctgagaga gttgcagcaa gcggtccacg   7680
ctggtttgcc ccagcaggcg aaaatcctgt ttgatggtgg ttaacggcgg gatataacat   7740
gagctgtctt cggtatcgtc gtatcccact accgagatat ccgcaccaac gcgcagcccg   7800
gactcggtaa tggcgcgcat tgcgcccagc gccatctgat cgttggcaac cagcatcgca   7860
gtgggaacga tgccctcatt cagcatttgc atggtttgtt gaaaaccgga catggcactc   7920
cagtcgcctt cccgttccgc tatcggctga atttgattgc gagtgagata tttatgccag   7980
ccagccagac gcagacgcgc cgagacagaa cttaatgggc ccgctaacag cgcgatttgc   8040
tggtgaccca atgcgaccag atgctccacg cccagtcgcg taccgtcttc atgggagaaa   8100
ataatactgt tgatgggtgt ctggtcagag acatcaagaa ataacgccgg aacattagtg   8160
caggcagctt ccacagcaat ggcatcctgg tcatccagcg gatagttaat gatcagccca   8220
ctgacgcgtt gcgcgagaag attgtgcacc gccgctttac aggcttcgac gccgcttcgt   8280
tctaccatcg acaccaccac gctggcaccc agttgatcgg cgcgagattt aatcgccgcg   8340
acaatttgcg acggcgcgtg cagggccaga ctggaggtgg caacgccaat cagcaacgac   8400
tgtttgcccg ccagttgttg tgccacgcgg ttgggaatgt aattcagctc cgccatcgcc   8460
gcttccactt tttcccgcgt tttcgcagaa acgtggctgg cctggttcac cacgcgggaa   8520
acggtctgat aagagacacc ggcatactct gcgacatcgt ataacgttac tggtttcatg   8580
agtacctctc ctgtatcttg tgtacctgcg taccgtaata tctcctctgt ttctgtatac   8640
cagtactatg tgattataac agcccctcaa ctacatgtca agaataaact gccaaagccg   8700
```

-continued

```
agcggaagcc tatttgacga atacctggcc gatctcgacc ttcccgtagt ggccaaattc   8760
ttccaactga tatgcgagcg aggccaagcg atcttcttcc tgtccaagct aagcctgtct   8820
agcttcaagt atgacgggct gatactgggc cggcaggcgc tccattgccc agtcggcagc   8880
gacatccttc ggcgcgattt tgccggttac tgcgctgtac caaatgcggg acaacgtaag   8940
cactacattt cgctcatcgc cagcccagtc gggcggcgag ttccatagcg ttaaggtttc   9000
atttagcgcc tcaaatagat cctgttcagg aaccggatca aagagttcct ccgccgctgg   9060
acctaccaag gcaacgctat gttctcttgc ttttgtcagc aagatagcca gatcaatgtc   9120
gatcgtggct ggctcgaaga tacctgcaag aatgtcattg cgctgccatt ctccaaattg   9180
cagttcgcgc ttagctggat aacgccacgg aatgatgtcg tcgtgcacaa caatggtgac   9240
ttctacagcg cggagaatct cgctctctcc aggggaagcc gaagtttcca aaaggtcgtt   9300
gatcaaagct cgccgcgttg tttcatcaag ccttacggtc accgtaacca gcaaatcaat   9360
atcactgtgt ggcttcaggc cgccatccac tgcggagccg tacaaatgta cggccagcaa   9420
cgtcggttcg agatggcgct cgatgacgcc aactacctct gatagttgag tcgatacttc   9480
ggcgatcacc gcttccctca tgatgtttaa ctttgtttta gggcgactgc cctgctgcgt   9540
aacatcgttg ctgctccata acatcaaaca tcgacccacg gcgtaacgcg cttgctgctt   9600
ggatgcccga ggcatagact gtaccccaaa aaaacagtca taacaagcca tgaaaaccgc   9660
cactgcgccg ttaccaccgc tgcgttcggt caaggttctg gaccagttgc gtgagcgcat   9720
acgctacttg cattacagct tacgaaccga acaggcttat gtccactggg ttcgtgcctt   9780
catccgtttc cacggtgtgc gtcacccggc aaccttgggc agcagcgaag tcgaggcatt   9840
tctgtcctgg ctggcgaacg agcgcaaggt ttcggtctcc acgcatcgtc aggcattggc   9900
gggcgatgtc gtactgtccc gccgcgaacc cggttaccgc ccagcccagc gcgaccagcc   9960
ccggcaacgc ctcgccgcac cgcttgcggc gcttgcgcat ggtcgaacca ctggcctctg  10020
acggccagac atagccgcac aaggtatcta tggaagcctt gccggttttg ccggggtcga  10080
tccagccaca cagccgctgg tgcagcaggc gggcggtttc gctgtccagc gcccgcacct  10140
cgtccatgct gatgcgcaca tgctggcagc cacccatgac ggcctgcgcg atcaaggggt  10200
tcagggccac gtacaggcgc ccgtccgcct cgtcgctggc gtactccgac agcagccgaa  10260
acccctgccg cttgcggcca ttctggcgca tgatggatac cttccaaagg cgctcgatgc  10320
agtcctgtat gtgcttgagc gccccaccac tatcgacctc tgccccgatt tccttcgcca  10380
gcgcccgata gctacctttg accacatggc attcagcggt gacggcctcc cacttgggtt  10440
ccaggcaccg caggagatgc agtccgcctc aggtcttggc ttccggcacc agcaccaggc  10500
catcaggccc agccatggcc acaagccatc gacggatgcg cagatcatcc gcgcccagcg  10560
gatccgggcc gatgaactcg atccgcttgc cgtcgccgta gtccaagtca acgtccagct  10620
tgccgcgctc gcgctcgccc cgcttgaggg caccgaacac gcccgggggc agacagtgcg  10680
caggctagtg caggaagtgg atgaggatgc gattgctatc aggcttcacc aaggggcacc  10740
ccattggtct tgcgctgcct ctccagcacg gcgggcttga gcacccggcc gtcatgccgc  10800
atgaaaaccc gatcagcgca aggtgcgcca tagtcggcct tgctcaaccc gacgcggacg  10860
aagaaccggc gctggtcgtc gtccacaccc cattcctcgg cctcggcgct ggtcatgctc  10920
gacaggtagg actgccagcg gatgttatcg accagtaccg agctgccccg gctggcctgc  10980
```

-continued

```
tgctggtcgc atgcgaacct catggcagcg aactggatgg catggtgatg gatcacgatc  11040

gagcaacggc tatcggaggc gatggcatcc atgcgaacga tgacctgggc catggggacg  11100

ctggcgtttt cttgctcgat gtggaacggc cgcagcgtgc ccagcaccat caggcggcgg  11160

cccccggcgg cgcgcttgag gccgtcgaac cactccgggc ccatgatgtt gggcaggctg  11220

ccgatcagcg gctggatccg caggccgtca gccacggctt gccgttcctc ggcgctgagg  11280

tgcgccccca gggcgtgccg gcggtgatga atggcggtgg gcgggtcttc ggcgggcagg  11340

tagatcaccg ggccggtggc cagttcgccc acctccagcc gatccggccc gcctgcaatc  11400

tgtgcggcca gttgcagggc cagcatggcc ttcccggatc caccgggagc cacccagacc  11460

ccgaccgtac cggccaccat gttgggcaaa acgtagtcca gcggtggcgg cgctgctgcg  11520

aacgcctcca gaatattgat aggcttatgg gtagccattg attgcctcct ttgcaggcag  11580

ttggtggtta ggcgctggcg gggtcactcc ccccgccctg cgccgctctg agttcttcca  11640

ggcactggcg cagcgccctg tattcgtcgt cggtcagcca gaacttgcgc tgacgcatcc  11700

cttcggactc catgcgatcg gactatcgcg cttgccgtcc agcgtcaggg ctggccagcc  11760

ggtcgcaggc ctgcctgtcc ttttggtctc tactatcagc caccgagacc cttgccgggg  11820

ccgaaaggcc tgtccttcgc gaacaaggac aaggtgcagc cgtcaaggtt aaggctggcc  11880

atatcagcga ctgaaaagcg gccagcctcg gccttgtttg acgtataacc aaagccaccg  11940

ggcaacccat agcccttgtc acttttgatc aggtagaccg accctgaagc gcttttttcg  12000

tatttcccta aacccccttc tgtgcgtgcg tactcatagt ataacaggcg tgagtaccaa  12060

cgcaagcacc aactgctgca ctctggacca cacctgtcca tgcctcgctg gcgggctgcc  12120

ggtgccagtg caagctaggc cagcgcacgc tggacgatgg gacgacccct gacctcgctg  12180

acggtgcgcc cgatgtaatc cgcttcgtgg ccgggcttgc gctccgccag cgctggcctg  12240

gcctcggcca tggccttgcc gatttc                                        12266
```

<210> SEQ ID NO 81
<211> LENGTH: 9492
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesised, pDIK352

<400> SEQUENCE: 81

```
ctgaaatctg gcccgcccct gtccatgcct cgctggcggg gtgccggtgc ccgtgccagc     60

tcggcccgcg caagctggac gctgggcaga ccccatgacc tgctgacggt gcgctcgatg    120

tacccagctc agtggcaggc ctggcgctcc gaccgcgatg ggctggcctc ggccatggcc    180

ttgccgattc catcggacct gaggccacgg ctggccagcc tatgcgcggc gatcaagtcg    240

caccggatga ggtcatgacc gacgcgattg accagcaagg accttcagct gaggtactcg    300

tccagcgccg tgcgccggtg gcggctaagc tgccgctcgg gcagttcgag gctggccagc    360

ctgcgggccc tatcctgctg cagcagggcc tgctcgatcc gatggccagc ctgcagaacc    420

agcgccgggc cagcggtggc ggtcttgccc ttggattcac gcagcagcac ccacggctga    480

tacccgggcg gggtcgtgtg cttgccattg aggtcggtga agcccgaccc gaggcgatcg    540

tggcggctgt cggcgctggc cgggtcggcg tcgtactcgc tggccagcgt ccgggcaatc    600

tgcccccgaa gttcaccgcc tgcggcgtcg gccaccttga cccatgcctg atagttcttc    660
```

-continued

```
gggcgggttc aaacgaacgg ggcaggatcc aggcccaagg atttcatgtc atccaggtcc     720
aacccgatga ggtcgtaccc cagcaacaga aaatgcagcc aatgccaggc gggcccgatc     780
tacacgtcac tgccctgggc attcatccgc ttgagccatg gcgtgctctg gagcacttcg     840
gcggctgacc attcccggtc catcatctgg ccggtcggtg cgtccctgac gccgatatcg     900
aagcgctcac agcccatggc cttgagctgc cggcctatgg cctgcaaagt cctgccgttc     960
ttcatcgggc caccaagcgc agccagatcg agccgtcctc ggttgtcagt ggcgtcaggt    1020
agaccaagag aaacgatgcg atcagcagca aaaccgtagg aatcaaggca gcacgcatcc    1080
aggccagcca tagcatacag tgccaccccc gagacgagca aagggcgatc tgcgcggagc    1140
tgcccacctc ggcggctacc tcccgcaacc ctttggccag ctccacccat gccgcccctg    1200
tctggcgcgc ggccttcagc cactccgccg cctgcgcctc gctggcctgc ttggcctggc    1260
tcatgaccta ccgggcttcg tcggccagtg tcgccatgct ctgggccagc ggttcgatct    1320
gctccgctca ctcgatgatg cctcaggatc tcttcactct gtcgaatgcg ttcatggtct    1380
attgcatccc ggtcatactg tacgacgatg atctggagca tggcggtgtc gatgccaagg    1440
gcaacgtatg aacggtaggc gaggatgccc aggcctacca tatccaaacc gttcggaacc    1500
aggtgaacac cgggcaggcg ctcgatgccc tgcgcctcaa gtgttctgtg gtcaatgcgg    1560
gcgtcgtggc cagcccgctc taatgccggc ttggcatggt cggcccatgc ctcgcgggtc    1620
tgctcaagcc atgccttggg cttgagcgct tcggtcttct gtgccccgcc cttctccggg    1680
gtcttgccgt tgtaccgctt gaaccactga gcggcgggcc gctcgatgcc gtcattgatc    1740
agcacggaga tactcaggtg gaacgacggg ttatcgaagc aaacggaatg gatggcaagc    1800
gtacaaggca ggcgctaggc aacggtaagg tgctggagga aatcggaagc aagcgcattc    1860
tgccggtcga gggtcagctc gacccgcagg gcaaattcga cctccttgaa cagccgccca    1920
ttggcgcgtt catacaggtc ggcagcatcc cagtagtcgg cgggccgctc gacgaactcc    1980
ggcatgtgcc cggattcggc gtgcaagcat tcatccatgt cgcgggcata cttgccttcg    2040
cgctggatgt agtcggcctt ggccctggcc gattggccgc ccgacctgct gccggttttc    2100
gacgcaaggc gatcaatcgc aatgccgccc agcatgatgc tttgcttttc ggctccatgc    2160
aatggccctc ggagagcgca ccgcccgaag ggtggccgtt aggccagttt ctcgaagaga    2220
aacgcgtaag tgcgccctcc cctacaaagt agggtcggga ttgccgccgc tgtgcctcca    2280
tgatagccta cgagacagca cattaaccat ggggatgtca gatggatcag gggagcaacc    2340
aggcggcgga tcggctggcc aagcatcgag aacaacgagc gcgcaatcat gccgaaattc    2400
agcgagagcg ggccagggca cagcagcaag agcgcaagca agcaacaagg agccaggtgc    2460
tggcaggggc aatgatttag gaccaggtga aacgcagaga gtggccggag gatcggatcc    2520
tggcggaaca ggatgagtac attgaaacgc aaaacgaacg agccatgttc ggtcagaagc    2580
cacgccagaa ggatgagccg ggctgaatga tcgaccgaga caggccctgc ggggctgcac    2640
aagcgaaccc aacccatggg tagggagaca ggcagctaca gaggctaaac gagcacaagc    2700
gtcattctgc ggggattggt gtgggggtta gcggggcttg cccgcctttc ccccagccgc    2760
gacgcggtgc gggagtgtgc agcccagcgc agcgaataga aacgcaatcc ggccacaggc    2820
cgggcatcat gggcaagggc agcagcgccc cacaagggcg ctgacaaccg cgcccagtgg    2880
attattctta gataatcatg gatggatttt tccaacaccc cgcccagccc cgccctgct     2940
```

-continued

```
gggcatgcag gtttgggggc gtgccagtta ttgcaggggt tagtgacagt tattgcaggg    3000
gggcgtgcca gttcctgcag gggttcgtga cagtcagtcc gggaggacag ggcccaggct    3060
ggcaatgtcc agcaacggca ggcatttcgg ctgaggttaa aagaactttc cgctaagcga    3120
tagactgtac gtaaacacag tattgcaagg acgcggaaca tgcctcatgt ggcggccagg    3180
acggccagcc gggatcggga tactggtcgt taccagagcc accgacccga gcaaaccctt    3240
ctctatcaga tcgttgacga gtattaccca gcattcgctg cgcttatggc agagcaggga    3300
aaggaattgc cgggctatgc gacccgggca tttgaagacc ttctccaatg cggggcggtg    3360
gagcatggcc ttctaagggc tagccgcacg tattgcaacg acgacgcact ggtcgcttta    3420
agaaatcaac ctaaagtata tatgagtaaa cttggtctga caggcccctg aattcgcatc    3480
tagactgatg agacgtggta gagcacacaa cagccggtac aagcaacgat ctccaggacc    3540
atctggatca tgcgcggatg acacgaactc acgacggcga tcacagacat taacccacag    3600
tacagacact gcgacaacgt ggcaatcgta cgccatacca cgtcggtcag tttcacctga    3660
tttacgtaca aacccgattc ggcgggtttt tgcttttgga ggggcagacc gatgaatgcc    3720
tgtcctttgg cagtttattc ttgccatgta gtgcgggggc tggtataatc acctagtact    3780
gttggaattg tgagcggata acaattccac acagaaacag aggagatatt acatatgagt    3840
aaaggagaag agcttttcac aggagttgtc ccaattcttg ttgaattaga cggtgatgtt    3900
aatgggcaca agttctctgt cagtggagag ggtgaaggcg acgcaacata tggcaagctg    3960
accctcaaat ttatctgcac cacgggtaaa ctacctgttc catggccaac actggtcact    4020
actttcgggt atggtgttca atgctttgcg agatacccag atcatatgaa acagcatgac    4080
tttttcaaga gtgccatgcc cgaaggttat gtacaggaaa gaactatatt tttcaaagat    4140
gacgggaact acaagacacg tgctgaagtc aagtttgagg gtgataccct tgttaataga    4200
atcgagttaa agggtattga ctttaaggaa gatggaaata ttttaggcca caaactggaa    4260
tataactata actcccataa tgtgtacatt atggccgaca agcaaaagaa cggtatcaag    4320
gttaacttca agatcagaca caacattgag gatggaagcg ttcaactagc cgaccattac    4380
caacaaaata ctccaattgg cgatggccct gtccttttac cagacaacca ttacctgtcc    4440
actcaatctg ccctttcgaa agatcccaac gaaaagcgcg accacatggt ccttcttgag    4500
tttgtcaccg ctgctgggat tacacacggc atggatgaac tatacaaata aatgcgtaat    4560
cgttaatccg caaataacgt aaaaacccgc ttcggcgggt ttttttatgg ggggagttta    4620
gggaaagagc atttgtcagc tgcgtctcca gcagttaacg agatgttacg aacccccagg    4680
acctcagcga atgcgaggcg atggacggta caccgggcag agtgctcacc gacctaaacc    4740
agcaaaacca aacgggaacg tatttccgac gagccttttg ttttacttgc tgtcaggacg    4800
ttcccagccg ctcttcgctc ctcaactccc aatttcttca tgttgaaact ctcaaaaacc    4860
aaaggctcca aaaggagcct ttaattgtat cggtttatca gcttgctttt acatcactgc    4920
cagcttccca gtccggaccc ctgtcggccc agctgcatta atgaatcggc caacgcgcgg    4980
ggagaggcgg tttgcgtatt gggcgccagg gtggtttttc ttttcaccag tgacacgggc    5040
aacagctgat tgcccttcac cgcctggccc tgagagagtt gcagcaagcg gtccacgctg    5100
gtttgcccca gcaggcgaaa atcctgtttg atggtggtta acggcgggat ataacatgag    5160
ctgtcttcgg tatcgtcgta tcccactacc gagatatccg caccaacgcg cagcccggac    5220
```

continued

```
taggtaatgg agcgcattgc gacccagacc atatgatagc tggcaaaacg aatcgcagtg    5280
ggaacgatgc aatcattcag aatttgaatg gtttgatgca aacaggaact ggcaccaacg    5340
tcgccttccc gttccgctac cggccgaatt tgattgcgag tgagatattt atgccagcca    5400
gccagacgca gacgcgccga gacagaactt aatgggcccg ctaacagcgc gatttgctgg    5460
tgacccaatg cgaccagatg ctccacgccc agtcgcgtac cgtcttcatg ggagaaaata    5520
atactgttga tgggtgtctg gtcagagaca tcaagaaata acgccggaac attagtgcag    5580
gcagcttcca cagcaatggc atcctggtca tccagcggat agttaatgat cagcccactg    5640
acgcgttgcg cgagaagatt gtgcaccgcc gctttacagg cttcgacgcc gcttcgttct    5700
accatcgaca ccaccacgct ggcacccagt tgatcggcgc gagatttaat cgccgcgaca    5760
atttgcgacg gcgcgtgcag ggccagactg gaggtggcaa cgccaatcag caacgactgt    5820
ttgcccgcca gttgttgtgc cacgcggttg ggaatgtaat tcagctccgc catcgccgct    5880
tccacttttt cccgcgtttt cgcagaaacg tggctggcct ggttcaccac gcgggaaacg    5940
gtctgataag agacaccggc atactctgcg acatcgtata acgttactgg tttcatgagt    6000
aaccctaatg tatcatgtgc aattgcgtac agtcatatca aatcagtttc tgtcaaaacg    6060
tactatgtga ttataccagc ccgctcacta catgtcaaga ataaactgcc aaagccgagc    6120
ggccgcttat ttgccgacta ccttggtgat ctcgcctttc acgtagtgga caaattcttc    6180
caactgatct gcgcgcgagg ccaagcgatc ttcttcttgt ccaagataag cctgtctagc    6240
ttcaagtatg acgggctgat actgggccgg caggcgctcc attgcccagt cggcagcgac    6300
atccttcggc gcgattttgc cggttactgc gctgtaccaa atgcgggaca acgtaagcac    6360
tacatttcgc tcatcgccag cccagtcggg cggcgagttc catagcgtta aggtttcatt    6420
tagcgcctca aatagatcct gttcaggaac cggatcaaag agttcctccg ccgctggacc    6480
taccaaggca acgctatgtt ctcttgcttt tgtcagcaag atagccagat caatgtcgat    6540
cgtggctggc tcgaagatac ctgcaagaat gtcattgcgc tgccattctc caaattgcag    6600
ttcgcgctta gctggataac gccacggaat gatgtcgtcg tgcacaacaa tggtgacttc    6660
tacagcgcgg agaatctcgc tctctccagg ggaagccgaa gtttccaaaa ggtcgttgat    6720
caaagctcgc cgcgttgttt catcaagcct tacggtcacc gtaaccagca aatcaatatc    6780
actgtgtggc ttcaggccgc catccactgc ggagccgtac aaatgtacgg ccagcaacgt    6840
cggttcgaga tggcgctcga tgacgccaac tacctctgat agttgagtcg atacttcggc    6900
gatcaccgct tccctcatga tgtttaactt tgttttaggg cgacttgccc tgctgcgtaa    6960
atcgttgctg ctccattaca tcacacatcg acccacggcg taacgcgctt gctgcttgga    7020
tgcccgaggc atagactgta ccccaaaaaa acagtcataa caagccatga aaaccgccac    7080
tgcgccgtta ccaccgctgc gttcggtcaa ggttctggac cagttgcgtg agcgcatacg    7140
ctacttgcat tacagcttac gaaccgaacc ggcttatgtc cactgggttc gtgccttcat    7200
ccgtttccac ggtgtgcgtc acccggcaac cttgggcagc agcgaagtcg aggcatttct    7260
gtcctggctg gtcatgacca aaatccctta acgtgagtca gcctgccgcc ttgggccggg    7320
tgcgtcagta cttgccagcc gagtactcgg ttcccgtcca gcccagacgc acccgctacg    7380
gcaacgcctc gcgcacccgc ttgcggcgct tgcgcatggt cgaaccactg gcctctgacg    7440
gccagacata gccgcacaag gtatctatgg aagccttgcc ggttttgccg gggtcgatcc    7500
```

-continued

```
agccaaacag aagctggtgc agcaggaggg aggtaaagct gtccagagcc agcaccctgt    7560

aaaagatgct gagcaaatgc tggccgacac aaatgaaggc atgcgcgatc aagggcttca    7620

gggccacgta caggcgcccg tccgcctcgt cgctgggcta ctccgacagc agcccaaacc    7680

cctgccgctt gcggccattc tgggcgatga tggataccct ccaaagcgcc tcgatgcagt    7740

cctgcatgtg cttgagcgcc ccaccactac cgaccactgc cccgatttcc tttgccagcg    7800

ccggatagct accttgaccc acatggcatt cagcggtgac ggcctcccac ttgggttcca    7860

ggaacagccg gagctgccgt ccgcctcggt cttgggttcc agggccaagc actagggcct    7920

taggccaagc catggcaacc agcccatgca ggatgcgacg atcatcagcg ccacgcggct    7980

ccgggccgcc gaactcgatc cgcttgccgc cgccgtagtc atacgtcacg tccagcttgc    8040

tgcgcttgcg ctcgccccgc ttgagggcac ggaacaggcc gggggccaga cagtgcgccg    8100

ggtcgtgccg gacgtggctg aggctgtgct tgttcttagg cttcaccacg gggcaccccc    8160

ttgctcttgc gctgcctctc cagcacggcg ggcttgagca ccccgccgtc atgccgcctg    8220

aaccaaagct aagcgaacgg tgcgccatcg ttggccttgc tatcaacgct gcggacgacg    8280

aacgggagct ggtcgtcgtc aaccccaccc tactcggcct aggcgctggt cctgctagcc    8340

aggtaggact gccagcggat gttatcgacc agtaccgagc tgcccgggct ggcctgctgc    8400

tggtcgcctg cgcccatcat ggccgcgccc ttgctggcat ggtgcaggaa caggatagag    8460

caccgggtac cggcggcgac ggcctccatg cgaccgatga cctgggccat ggggccgctg    8520

gcgcttctct cctcgatgtg gaaccggcgc agcgtgtcca gcaccatcag gcggcggccc    8580

taggcggcgc gcttgaggcc gtcgaaccac taccgggcca tgatgatggg caggctgacg    8640

atcagaggct ggatcagccg gacgtcagcc acggcatgcc gttcctaggc gatgaggtgc    8700

gcccaaggcg cgtgcaggcg gtgatgaatg gcggtgggcg ggtgttcggc gggcaggtag    8760

atcaccgggc cggtgggcag ttcgcccacc tccagcagat ccggcccgcc tgcaatctgt    8820

gcggccagtt gcagggccag catggattta ccggcacgac cgggcgcacc cagcgccccg    8880

accgtaccgg ccaccatgtt gggcaaaacg tagtccagcg gtggcggcgc tgctgcgaac    8940

gactcaagca tattgatcgg atttcgggta gcaccagatc gactccttcg aaggcagttg    9000

gtggctaggc gctggcgggg tcactacccc cgcccgcgcc cgcctcgagt tctttcaggc    9060

actcgcgcag cgcctcgtat tcgtcgtcgc tcagccagaa cttgcgctga cgcatccctt    9120

tggccttcat gcgctcggca tatcgcgctt ggcgtacagc gtcagggctg gccagcaggt    9180

cgccggtctg cttgtccttt tggtctttca tatcagtcac cgagaaactt gccggggccg    9240

aacggattgc attcgaggca aacggaaacg gtgcagaagc aaacggtaag gatggcaatc    9300

taacgcactg aaacgaggcc agccacggcc ttgtacgacg tataacaaac gcaccgggcc    9360

aaccaatagc aattgtaact tttgatacgg tagacggacc atgcagagct ttttacgtat    9420

tccataaaac ccccctctgt gcgtgagtac tcatagtata acaggcgtga gtaccaacgc    9480

aagcaataca tg                                                        9492
```

<210> SEQ ID NO 65
<211> LENGTH: 7787
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pDTK563

continued

<400> SEQUENCE: 85

```
taatgaacaa aatccattca agtgagtcag aatgccgaca tgggccgggt gatgacgtcc     60
ttgcccgccg cgaactcggc tacggtccag ccgagcgcga ccagctcggg caacgcctcg    120
cgcacccgcc tgcggcgctt gcgcatggtc gaacgactgg cctcgacggc ccagacatag    180
ccgcacaagg tatcgatgga agccttgccg gtttcgccgg ggtcgatcca gcccacaagc    240
cgctggtgca gcaggcgggc ggtttcgctg tcccagcgcc gcacctcgtc catgctgatg    300
agcacatgca ggccgaaccc aatgacggcc tgcgcgatca aggggtcacg ggccacgtcc    360
aggcgaacga aagcctagtc gatggcgtcc tccgacagca gacgaaaccc atgccgattg    420
cggccattcc gggcgatgat ggataccttc caaaggcgcc cgatgcagtc ctgtacgtgc    480
ttgagcgccc caccactatc gacctctgcc ccgatttccc ttgccagcgc ccgatagcta    540
cctttgacca catggcattc agcggtgacg gcctcccact tgggttccag gaacagccgg    600
agctgccgtc cgccttcggt cttgggttcc gggccaagca ctaggccatt aggcccagcc    660
atggcaacca gaccatgcac gatgcgaaga tcatcagagc aaagcgatcc agggccgatg    720
aacacgatcc gattgaagtc gacgaagtca tacgacaagc aaagcatgct gagcacgagc    780
tcgcccgccc tgaggcgacg gaacaggccg ggggccagac agtgcgccgg gtcgtgccgg    840
acgtggctga ggctgtgctt gttcctaggc ttcaccacgg ggcacccccct tgctcttgcg    900
ctgcctctcc agcaacggcg gcttgagcac cccgccgtca tgccgcctga accaccgatc    960
agcgaacggt gcgccatcgt tggcctgctc cacaccgaag cggacgaaga acggcgcctg   1020
gtcgtagtcc aaccccaatc aatcggactc ggcgctggtc atgctcgacc ggtcggactg   1080
aacgcggatg ttatcgaccc gtcccgagcc gaccccgatg gactgctgct ggtcgcatgc   1140
gcccatcatg gccgcgcccc tgctggcatg gtgcaggaac acgatagagc accgggtatc   1200
ggcggcgatg gcctccatgc gacccatgac ctgggccatg gggccgctgg cgtttctctc   1260
ctcgatgtgg aacccggcga gcgtgtcccg cacccatcagg cggcgggcct cggcgggcgg   1320
cttgaggccg tcgaacccct ccgggggcct gatgttgggc aggctgccga tcagcggctg   1380
gatcagatgg aagtcagcca aggcatgccg ttcctcggcg atgatgtgcg aaccaagggc   1440
gtgcaggcgg tgatgaatgg cggtggcggg gtcttcggcg ggcaggtaga tcaccggcgc   1500
ggtggcagct tcgcccaccc ccagcagatc cggcccgccc gcaatctgtg cggccagttg   1560
cagggccagc atggatttac cggcacccac gggcgacacc agcgcccgga ccgtaccggc   1620
caccatgttg ggcaaaacgt agtccagcgg tggcgggcgt gctgcgaacg cctccagaat   1680
attgataggc ttatgggtag aaataggttg aatcctttgc aggcagttgg tggtcaggcg   1740
atggcgggga aactaaaccc gaccatagcc gatcatagta attccaggca atcgcgaagc   1800
gaatcgtata agtagtaggc aagccagaac ttgcgctgac gaatcccttt ggcctcaatg   1860
cgctcggcat atcgcgcttg gcgtacagcg tcagggctgg ccagcaggtc gccggtctgc   1920
ttgtcatttc ggtcattcat atcatcaacc gagcaaattg aagggcaaga aaggcatgtc   1980
ttcgcggaac aaggacaagg tgcagccgtc aaggttaagg ctggccatat cagcgactga   2040
aatgcgccca gactcggccc tgtttgaacg atcacccaag aaacagggcc aacatagcc    2100
cttgtcactt ttgatcaggt agacccgacc tgaagcgctt ttttcgtatt ccataaaacc   2160
ccctctgtgc cgtgagtacc catagtataa caggcgtgag taccaacgca agcactacat   2220
```

-continued

```
gatgaaatcc ggcccgaccc tgtccatgcc tagccggagg ggtgccggtg accgcgaccg   2280

atgggaacgc gaacgatgga agctgggccg acccatgacc ttgccgaagg tgcgccagct   2340

gtaatccgcc tcgtggcggg gcttggcgtc tgccagcgcc gggcggcgct cggccatggc   2400

cttggcgatc tcctcggcac tgcggccgcg gctggccagc ttctggcggg cgataaagtc   2460

gcacctgctg aggtcatgac cgaagcgctt gacccagccg gccatctcgc tgcggtactc   2520

gtccagcgcc gtgcgccggt ggcggctaag ctgccgctcg ggcagttcga ggctggccag   2580

aatgcgggcc ttctcatgcc gacgccgggc atgccgcatc tgctggaacg aatgccgacc   2640

aagcgaaggg aacgcggtgg aggtcatgcc attggattca agccgcagcc accccggatg   2700

ataaccggcg cgggcggtgc gcttgtcctt gcggtcggtg aagcccgcca agcggccata   2760

gtggcggcgg tcggcgctgg ccgggtcggc gtcgtactcg ctggccagcg tccgggcaat   2820

ctgccccgga agttcaccgc ctgcggcgtc ggccaccttg acccatgcct gatagttctt   2880

cgggctggtt tccactacca gggcaggctc ccggccctcg gctttcatgt catccaggtc   2940

aaccaagctg aggtcgtcca accgcaaccg aaccggaagc tactgctagg aggggctgct   3000

atccaagtca ttgccatggg aattcatccg attgagaacc ggcggttcat ggcgcaattc   3060

ggcggctgac cattcccggc tcatcatccg gccggcgggc gcgtccctga cgccgatatc   3120

gaagcgctca cagcccatgg ccttgagccg tcggcctatg gcctgcaaag tcctgccgtt   3180

cttcatcggg ccaccaagcg cagccagatc gagccgtcct cggttgtcag tggcggtcgg   3240

tcgagcaaga gcaaccgatg gatcagcagc accaccgtag gcatcatgga agccagcatc   3300

aaggttagcc atcgcttcca gtgccacccc agcgacgagc taggggagct atgcgcgccg   3360

atgccaaccc aggcggatcc atcccgaccc tatttggaca gatccaaccc tgccgcaact   3420

gtcgggcgcc gggccttcag ccacccggcc gcctgcgccc cgctggcctg cttggcctgg   3480

ctcatgaccc gccgggcttc gtcggccagc gtcgccatgc tctggggcag cggttcgatc   3540

tgcatcggct aatcgttgat gcctcaggat ttcttcatcc tgtcgattgc gttcatggtc   3600

tattgcctcc cggtattcct gtaagtcgat gatcaggccg ttggcggtgt cgatgttcag   3660

ggccaagtcc gacccgtccg tgcggatgcc aaggccttcc atatccaacc agttcggacc   3720

caggcgaaca ccgggcaggc gctcgatgcc ctgcgcctca agtgcctcgt ggtcaatgcg   3780

ggcgccgtgg ccagcccgcc ctaatgcccg gttggcatgg tcggcccatg cctcgcgggt   3840

ctgcatcagc catgccttgg gcttgagcgc ttcggtcttc tgtgccccgc ccttcatcgg   3900

ggtcatgccg ttgtaccgct tgaaccactg agcggcgggc cgctcgatgc cgtcattgat   3960

aagccaggcg atcccaggcc ggcacgcgcg gttcccgacg accccggaat ggatggaacg   4020

agtacaaggc aggcgatcgg aacccgtccg gtgccgggcg aactcggacg aacgcgaatt   4080

atgccggtcg aggggaagcc agcccggccg ggccaattcg aactccttgc aacgccgacc   4140

attggcgcgc tcatacaggc cggcagcatc ccagcagtcg gcgggccgct cgacgaactc   4200

aggcatgtgc aaggattcgg agtgcaagcc ttccccaatg tagcgggact aattgcattc   4260

gcgcatggat gtagtcggcc tggccctggc cgattggccg cccgacctgc tgccggtttt   4320

agccgtacgg tgataaatcg aaatgctgcc tagccgttgc ttttgctttt aggcccaatg   4380

caatggcccc cggcgagcgc accgcccgaa gggtggcccg taggccagtt tctcgaagcg   4440

aaaccggtaa gtgcgcccgc ccctacaaag taggggcggg attgccgccg ctgtgcctcc   4500
```

-continued

```
atgatagccc aagagaatgc aaattaacca tggggagtca agatggttac ggggagaaac    4560
aaggcgggcg atcggatggc aaagcaagca gaacaaagcg agcgaatacc tgccgaaatt    4620
cagccggagc gggcaaggga acagcagcaa gagcgcaaga acgaaacaag gcgcaaggtg    4680
ctggcggggc ccatgatttt ggccaaggtc aacagcagcc agtggccgga ggatcggctc    4740
atggcggcaa tggatgcgta ccttgaacgc gacccacgac gcgccttgtt cggtctgccg    4800
ccacgcccga aggatgagcc gggctgaatg atcgaccgag acaggccctg cggggctgca    4860
aacgcgaccc aacccttcgc gtacgcggca aggccgatca agcggctacc agcgctaacg    4920
agtactttcg aggggtttgc tgtgccgttc agcggcattc gaccgcattt aaccctgacg    4980
cgcagcggtc gggcgctgtc tagcctagcc cagccaatac accagctatc cggcctctgg    5040
ccgggcatac tgggcaaggc cagcagcgcc ccacaagggc gctgatacac gcgcctagtg    5100
gattattctt agataatcat ggatggattt ttccaacacc ccgccagccc ccgcccctgc    5160
tgggtttgca ggtttggggg cgtgacagtt attgcagggg ttcgtgacag ttattgcagg    5220
ggggcgtgac agttattgca ggggtcgtga caggtagtca aggcagtgac gggcactggc    5280
tggcaatgtc tagcaacggc aggcatttcg gctgacggta aacgaacttt aagctaagcg    5340
atagactgta tgtaaacaca gtattgcaag gacgcggaac atgcctcatg tggcggccag    5400
gacggccagc cgggatcggg atactggtcg ttaccagagc caccgacccg agcaaaccct    5460
tctctatccg atcgttgacg agtattaccc ggcattcgct gcgcttatgg cagagcaggg    5520
aacggaattg ccgggctatg tgccaacgga atttgaagaa tttctccaat gcggcggct    5580
ggtgcatggc tttctaaggg ttcgctgcga gtctcgaacc gacgagtacc tggtcgcttt    5640
aagaaatcca tatcaagtcc atatgagtca actggtatcg aacggcaact gacccgaact    5700
ctagatggta gagccacaaa cagccggtac aagcaacgac ctccaggacc atctgaatca    5760
tgcgcggatc acacgaactc acgacggcga tcacagacac taacccacag tacagacact    5820
gcgacaacgt ggccattcgt cgccataccg tctcactgaa ctggccgata attgcagacg    5880
aacgctttgg cagtttattc ttgacatgta gtgagggggc tggtataatc acatagtact    5940
gttatacaga aaccgaggcg atattaacta tggtcctacc actcgaagct ttcgccgggg    6000
actggcgaca gacagccggc tacaacctgg accaagtccc tgaacaggga ggtgcgtcca    6060
gtttgtttca gaatctcggg gtgtccgtaa ctccgatcca aaggacagtc ctgagcggtg    6120
aaacaggcct gaagatcgac atccacgtca tcatcccgta tgaaggtctg agcggcgacc    6180
aacaggccca gatcgaaaca atttccaagg tggtgcaccc tgtggatgat catcacttta    6240
aggccatccc gaactatggc aaactggtca tcgacgggcc tacgccgatc atgcccgact    6300
attccggacg gacgcatgca ggcaacgccg tgttcgaacg caacaagatc actgcaaacg    6360
ggacccatgc gacaggaacc aacaccatcg aagagcgacc gatcaaaacc gacggctacc    6420
tgccgttccg agtaaccatc aacggagtga ccggccggcc gctgccgcaa cgcattctgg    6480
agtaaatccg tactcgttca tacgcaacta acgtaaaacc aagcccaggc gggtcctttt    6540
atggggggag tttaggcaaa gagcattgct cagctggatc agcagatggg aaccggatac    6600
taccaagccc aacgagtagc aactgagacg aagagcggcc gattacttgc agacctaact    6660
ggtgatctcg ccttacacgt agtggacaaa ttcttccaac tgatcagcgc gcgaggccaa    6720
gcgatcttcc tcttgtccaa gataagcctg tctagcttca agtatgacgg gctgatactg    6780
```

-continued

```
ggccggaagg agctcaattg aaacgaaggc agcgacatcc ttcggcgagc ttttgcaggt     6840
taccgagctg taccaaatgc gggacaacgc aagcactacc tttcgctact agccagaacc     6900
gtcgggcggc gagttccata gcgttaaggt ttcatttagc gcctcaaata gatcctgttc     6960
aggaaccgga tcaaagagtt cctccgccgc tggacctacc aaggcaacgc tatgttctct     7020
tgcttttgtc agcaagatag ccagatcaat gtcgatcgtg gctggctcga agatacctgc     7080
aagaatgtca ttgcgctgcc attctccaaa ttgcagttcg cgcttagctg gataacgcca     7140
cggaatgatg tcgtcgtgca caacaatggt gacttctaca gcgcggagaa tctcgctctc     7200
tccaggggaa gccgaagttt ccaaaaggtc gttgatcaaa gctcgccgcg ttgtttcatc     7260
aagccttacg gtcaccgtaa ccagcaaatc aatatcactg tgtggcttca ggccgccatc     7320
cactgcggag ccgtacaaat gtacggccag caacgtcggt tcgagatggc gctcgatgac     7380
gccaactacc tctgatagtt gagtcgatac ttcggcgatc accgcttccc tcataatgtt     7440
taactttgtt ttagggcgac tgcccctgct gcgtaacatc gttgctgctc ataacatcaa     7500
acatcgaccc acggcgtaac gcgcttgctg cttggatgcc cgaggcatag actgtacccc     7560
aaaaaaacag tcataacaag ccatgaaaac cgccactgcg ccgttaccac cgctgcgttc     7620
ggtcaaggtt ctggaccagt tgcgtgagcg catacgctac ttgcattaca gcttacgaac     7680
cgaacaggct tatgtccact gggttcgtgc cttcatccgt ttccacggtg tgcgtcaccc     7740
ggcaaccttg ggtagcagcg aagtcgaggc atttctgtcc tggctgg                  7787
```

<210> SEQ ID NO 86
<211> LENGTH: 7798
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pDIK564

<400> SEQUENCE: 86

```
tcgatgtaat ccgctccgtg gccgggcttg cgctctgcca gcgctgggct ggcctcggcc       60
atggccttgc cgattcctc ggcactgcgg ccccggctgg ccagcttctg cgcggcgata      120
aagtcgaccc tgctgaggtc atgaccgatg agctcgaaca gacaggaact ctcgctgagg     180
tactcgtcca gcgcgtgcgc ccggtcgcgg ctaagctgcc gctcgggcag ttcgagcctg     240
gccagcctgc gggccttctc ctgctcgcgc tgggcctgcc cgatcctgctg gccagcctgc   300
tgcaccagcg ccggggccagc ggtggcggtc ttgccctttgg attcacgcag cagcaccccac  360
ggctgataac cggcgcgggt ggtgtgcttg tcctcgcggt tggtgaagcc cgccaagcgg     420
acctagtggc ggctgtaggc gatggcaggc taggcgtagc actcgctggc cagcgtcagg     480
gacctctgcc caggaagttc accgcctgcg gcgtcggcca actggaccct tgcccgatcg     540
ttcttgggcc tggtctaccc taccaggcca ggatccaggc catcggcttt catgtcatcc     600
aggtcaaact cgctgaggtc gtccaccagc accagaccat gccgctcctg ctcggcgggc     660
ctgatatcca agtcattgcc ctgggcattc atccgcttga gacctggcgt gttcaggagc     720
acttcggcgg ctgaccattc ccggttcatc atctggccgg tgggtgcgtc cctgacgccg     780
atctcgatgc gatcaaagcc aatggcattg agctgaaggc atatggaatg aaacgaaatg     840
tcgttcttca tcgggccacc aagcgcagcc agatcgagcc gtcctcggtt gtcagtggcg     900
tcaggtcgag caagagcaac gatgcgatca gcagcaccac cgtaggcatc atggaagcca     960
```

continued

```
gattcaaggc tagccatagc ttccagtgcc acccccgaga agcgctaagg gagcacgcag   1020
aggcgatgcc aaccaaggcg gataccacac gaacctattc ggccagatcc acccatgacg   1080
cccctgtctg gcgctgggct ttcagccact ccgcgcgctg cgcctcgctg gcctgcttgg   1140
tctggctcat gaccgcgcgg gcttcgtcgg ccagcgtcgc catgctctgg gccagcggtt   1200
cgatctgctc cgctaactcg ttgatgcctc tggatttctt cactctgtcg attgcgttca   1260
tggtctattg cctcccggta ttcctgtaag tcgatgatct gggcgttggc ggtgtcgatg   1320
ttcagggcca agtcagaccg gtcggcgcgc atgcccaggc attccatctc accaacgttc   1380
ggccccaggc gatcaacggc caggcgatcg atgccctgcg catcaagtgt tatgccgtcc   1440
atgcgggcgc cgtggccagc ccgcctcaat gccgggttgg catggtcggc ccatgcctcg   1500
cgggctcgct caagccatgc cttgggcttg agcgcttcgg tcttcgtgcc cccgccctta   1560
tccggggtct tgccgttgta ccgcttgaac cactgagcgg cgggccgctc gatgccgtca   1620
ttgatccgct cggagatcat caggtggccg tgcgggttct cgccgccacc ggcatggatg   1680
gacagcagtc aaggcaggcg atcggcaccg gtcagctgcc gggcgaactc ggccgccagc   1740
gacctatgcc ggtcgagggc cagcacgacc ggcaggcaca attcgaactc attgaacagc   1800
cgcccattgg cgcgttcata caggtcggca gcatcccagt agtcggcggg ccgctcgacg   1860
aactccggca tgtgcccgga ttcggcgtgc aagacctcat ccatgtcgcg ggcatacttg   1920
ccttcgcgct ggatgtagtc ggccttggcc ctggccgatt ggccgcccga cctgcttgcg   1980
gtttccgccg tcaggtgata aatcgccatg ctgccttcgc gttgcttttg cttttcggct   2040
acctgcactg gacctaggcg agcgcaacgc acgcaggtgc gacgctaggc cagtcccacg   2100
aagagcaacc ggtcagtgcg cactccccta aaccgtaggc tagggattgc agccgctgtg   2160
cctccatgat agcctacgag acagcacatt aacaatgggc tgtcaagatg gttaagggga   2220
gcaacaaggc ggcggatcgg ctggccaagc tcgaagaaca acgagcgcga atcaatgccg   2280
aaattcagcg ggagcgggca agggaacagc agcaagagcg caagaacgaa acaaggcgca   2340
aggtgctggt gggggccatg atttcggcca aggtgaacag cagcgagtgg ccggaggatc   2400
ggcccatggc ggccatggcc gagtacattg aacgcgaaca agaccgagcc ttgtacggtc   2460
tgccgccacg ccagaaggat gagccgggcc gaatgatcga ccgagacagg ccctgcgggg   2520
ctgcacaggc gcccccaccc ttcgggtagg gggaaaggcc gctaaagcgg ctaaaagcgc   2580
tccagcgtat ttctgcgggg tttggtgtgg ggtttagcgg gctttgcccg ccttttcccc   2640
tgccgcgcag cggtggggcg gtgtgtagcc tagcgcagcg aatcgacccg ctatccggcc   2700
tatggaaggc aatacagggc aagggcagca gagcccaaca agggcgatga taccgcagcc   2760
tagcggatta ttcttagcta atcacggatg gattactaca acacccagcc agccccagcc   2820
aatgctgggc ttgcaggttc ggggccgtga aagtcattgc aggggctagt gacacgtatt   2880
gcaggggggc gtgacagtta ttgcaggggc tcgtgacagc tagtacggga gtgacgggca   2940
atggctggca atgtctagca aaggcaggca tttcggatga gggtaaaagc aatttcagat   3000
aagcgatcga ctgtatgtaa acacagtatt gcaaggaagc ggaacatgcc tcatgtggcg   3060
gacaggacgc acagcaggga taggcatacc ggtcgctacc agagccaacg aaccgagacc   3120
acccttctct atcagatcgt tgacgagtat tacccggcat tcgctgcgct tatggcagcg   3180
caggcaaagc aattgccggc ctatgcgcaa cgggaatttg aagaatttct ccaatgcggg   3240
```

continued

```
aggctggagc atggctttct aagggttggc tgcgagtatt gacctggagt gatcctggtc   3300
gattttagta atcaatatta agtaaatatg agtaaatttg gtatgaaagg aaactgaatt   3360
cgcatctaga tggtagagcc acaaacagcc ggtacaagca acgatctcca ggaccatctg   3420
aatcatgcgc ggatgacacg aagtcacgac gggcatcaca gacattaacc cacagtacag   3480
acaactgcac aacgtggcca ttcgtcgcca tacgttctca ctgaactggc cgataattgc   3540
agacgaacgg gtgaaacaaa acggttgaca acatgaagta aacacggtac gatgtaccac   3600
atgaaaagtc agtgagtata taactgaata gaggagattt taactatggt attcacactc   3660
gatgatttcg ttggggactg gagacagtca gacggctaca aactggaact agtacctgct   3720
caggcaggtg tgtccagttt gtttcagaat ctcgggtgtt ccgtaactcc gatccaaagg   3780
attgtcctga gcggtgaaaa tgggctgaag atcgacatcc atgtcatcat cccgtatgaa   3840
ggtcagagcg gcgacccaat gggccagatc gaaaaaattt ttaaggtggt gtaccctgtg   3900
gatgatcatc acttaagggt gatcctgcac tatggcacac tggtaatcga cgggggtacg   3960
atgaaaatga tagtctattt aggtcggccg tatgaaggca tagccgtgtt agtcggaatt   4020
aagattactg tacaagggtc aatgtcggtc ggctacaata ttatcgaagt gagcctgatc   4080
aacccgaccg gctccctgct gtttcgagta accatcaacg gagtgaccgg ctggcggctg   4140
tgcgaacgca ttctggcgta aatccgtaat cgttaatccg caaataacgt aaaaacccgc   4200
ttcggcgggt ttttttatgg ggggagttta gggaaagagc atttgtcagc tggatcagca   4260
gatgggaacc ggatactcct cagcctcccg agtagccaat gagacgccga gcggccgctt   4320
atttgaagtc taccttggtg atatcgactc tatcgtagtg gactaattat tactactgct   4380
atgctagcga ggccaagcga tattctatct gtccaagata agcctgtatc gattcaagtc   4440
tgacgggctg atactgggcc ggcaggcgct ccattgccca gtcggcagcg acatccttcg   4500
gcgcgatttt gccggttact gcgctgtacc aaatgcggga caacgtaagc actacatttc   4560
gctcatcgcc agcccagtcg ggcggcgagt tccatagcgt taaggtttca tttagcgcct   4620
caaatagatc ctgttcagga accggatcaa agagttcctc cgccgctgga cctaccaagg   4680
aatcgatatg ttatcttgct tttgtcagca agatagatcg atcaatgtcg atcgttgatg   4740
gctcgaagat accggcaaga atgtcattgc gctgccattc tccaaattgc agttcgcgct   4800
tagctggata acgccacgga atgatgtcgt cgtgcacaac aatggtgact tctacagcgc   4860
ggagaatctc gctctctcca ggggaagccg aagttttcca aaggtcgttg atcaaagctc   4920
gccgcgttgt ttcatcaagc cttacggtca ccgtaaccag caaatcaata tcactgtgtg   4980
gattcaggcc gactttatct gaggatacgt aattatgtac ggccagtatc gtcggttagt   5040
gatggtgatc gatgaagcca actaccctcg atagttgagt agtactttcg gagttcatcg   5100
attccatctt aatgtttatc tttgtttctg ggcgactgcc atgctgagtc aattcgttgc   5160
tgctccataa catcaaacat cgacccacgg cgtaacgcgc ttgctgcttg gatgcccgag   5220
gattagactg taccctatta aatcagtctt aactagattt gattacagcc aatgcgtagt   5280
taccaccgct gcgttcggtc aaggttctgg accagttgcg tgagcgcata cgctacttgc   5340
attacagctt aagtaatgta aaggctatcg taattcgggt tagtgcattc atccgttcca   5400
acggtgtgcg tcacccggca accttgggta gcagcgaagt cgaggcattt ctgtcctggc   5460
tggtcatgac caaaatccct taacgtgagt cagcctgccg ccttggccgg ggtgatgtcg   5520
```

continued

```
tacctgaccg aagcgaactc ggttacagtc aagcccagcg agcccagatc aggcaaagcc    5580
tagcgaaccc gattgagggg attgcgactg gtcgaaaaac tggcccatgc aggccagacc    5640
tagccgcaca aggtatctat ggaagccttg ccggtttgcc cggggtcgat ccagccacac    5700
agccgctggt gcagcaggcg gggggtttcg ctgtccagcg cccgcacctc gtccatgctg    5760
atgcgcacat gctggccgcc acccatgacg gcctgcgcga tcaaggggtt caggggcacg    5820
tacaggcgcc cgtccgcctc gtcgctggcg tactccgaca gcagccgaaa cccctgccgc    5880
ttgcggaccc tatggcagcc gatggatccc ttcaaaaggc gatcgatgcc gtcccgtatg    5940
tgccggagcg acccaaaccc atcgacatcc gacccgattc aatttgcaag agcccgatcg    6000
ctacctttga ccacatggca ttcagcggtg acggcctccc acttggcttc caggaacagc    6060
cggagctgcc gtccgccttc ggtcctgggt tccggggcaa gcacgaggcc attaggccca    6120
gccatggcca ccagcccttg caggatgcgc agatcatccg cgcccagcgg ctccggggcg    6180
ctgaactcga tccgcttgcc gtcgccgtag tcatacgtca cgtccagctt gctgcgcttg    6240
agcacgaacc gattgagggc aaggaaacgc aagggcgacc gacccgcagc agggtcgtgc    6300
aggacgtggc tgcagcatgc attgcaatta ggctcccaca aggggcaacc aattgcgatt    6360
gcgctgcctc tccagcacgg cgggcctgag cacccgcccg tcatgccgcc tgaaccaccg    6420
atcagcgaac ggtgcgccat agttggcctt gctcacaccg aagccgacga agaaccgccg    6480
ctggtcgtcg tcccacaccc attcctcggc ctcggcgctg gtcatgctcg acaggtagga    6540
ctgccagcgg atgttatcga ccagtaccga gctgccccgg ctggcctgct gctggtcgcc    6600
tgcgcaaatc atggcagcgc aattgctggc atggcgaagc aacccgatcg agcaccaggt    6660
atcgcaggcg atggcatcca tgcgaccgcc gacctgggcc atgggcaagc tggcgttttc    6720
ttcctcgatg tggaaccggc gcagcgtgtc cagcaccatc aggccgcggc cctcgcggcc    6780
gcgcctgagg ccgtcgaacc actcgcgggc catgatgttg ggcaggctgc cgatcagcgg    6840
ctggatcagc aggccgtcag ccacggcttg ccgttcctcg gcgctgaggt gcgcccccag    6900
ggcgtgcagg cggtgatgaa tggcggtggg cgggtcttcg gcgggcaggt agatcaccgg    6960
gacggtgggc agttcgaccc aatccagccg atccggcacg aatgcaatct gtgcggcacg    7020
ttgcagggcc agcatggatt taccggcacc accgggcgac accagcgccc cgaccgtacc    7080
ggccaccatg ttgggcaaaa cgtagtccag cggtggcggc gctgctgcga acgcctccag    7140
aataatgata ggcttatggg tagccattga ttgcctcctt tgcaggcagt tggtggttag    7200
gcgctggcgg ggtcactccc cccgccctgc gccgctctga gttcttcccg gcactcgcgc    7260
agcgcatcgc attcgtagtc ggtcagcacg aacttgcgcc gacgcatccc tttggcattc    7320
atgcgatcgc aatcccgcgc ttggcgtcca gagtcagggc tggccagccg gtcgccggtc    7380
tgcccgtccc tttggtattc aatcccagtc accgagcaac ttgccggggc agccagcatt    7440
gtcctcgcgg aacaaggaca aggtgcagcc gtcaaggtta aggccgccca tatcagcgac    7500
tgcaaagcgc aacgcatcgc aattgcttga agtcccacca aagccaacgg gacccccatc    7560
gcccctgtca cttttgatca ggtagaccga ccctgaagcg cttttttcgt atttccataa    7620
accccattcc gtgcgtgcgc aatcatagta tacaggacgc gagtacaacc gacggcaatc    7680
catgctgaaa tctggcccgc ccctgtccat gcctcgctgg cggggtgccg gtgcccgtgc    7740
cagctcggcc cgcgcaagct ggacgctggg cagacccatg accttgctga cggtgcgc      7790
```

-continued

<210> SEQ ID NO 87
<211> LENGTH: 7988
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pDIK569

<400> SEQUENCE: 87

```
tcgatgtaat ccgcttcgtg gccgggcttg cgctctgcca gcgctgggct ggcctcggcc      60
atggccttgc cgatttcctc ggcactgcgg aaacggatgg aaagcattcg agcggcgatc     120
aagtcgatcc tgctgaggtc atgacccgat agcttgacca gacgggaact atagctgagg     180
taatcgtcca gcgacgtggc ccggtcggcg ctaagctgac gctcgggcag ttcgaggctg     240
gccagcctgc gggacttctc ctgctgcagc tgggcctgct cgatctgctg gccagcctgc     300
tgcaccagcg ccgggccagc ggtggcggtc ttgcccttgg attcacgcag cagcacccac     360
ggctgataac cggcgcgggt ggtgtgcttg tccttgcggt tggtgaagcc cgccaagcgg     420
aacttgtggc ggctgtaggc gatggcaggg taggcgtagt aatcgctggc aagcgtcagg     480
gatcatggcc aacgaagttc aacgcctgcg gagtcggaca aattgaaccc tgcctgatcg     540
ttctccgggc tggttcgcac tacgagggca ggctcccggc cctgggcttt catgccatcc     600
aggtcaaacc cgctgaggtc gtcgaccagc accagaccat gcagctcctg ctcggcgggc     660
ctgatatcca cgtcattgcc ctgggcattc atccgcttga gccatggcgt gttctggagc     720
acttcggcgg ctgaccattc ccggttcatc atctggccgg tgggtgcgtc cctgacgccg     780
atctcgatgc gatcaaagcc aatggcattg agctgcaggc ctatggcatg aaacgtcatg     840
tagccattca taggcaaccc aagcgcagcc agctcgagcc gtcctcggtt gtcgctggcg     900
tcaggtcgag caagagcaac gatgcgatca gcagcaccac cgtaggcatc atggaagcca     960
gcatcacggt tagccatagc ttccagtgcc acccccgcga cgcgctccgg gcgctctgcg    1020
cggcgctgct caccttggcg gctacctccc gcaacttctt ggccagctcc acccatgccg    1080
cccctgtctg gcgctgggct ttcagccact ccgccgcctg cgcctcgctg gcctgcttgg    1140
tctggatcct gacctgacgg gattcgtcgg aacgcgtagc aatgcctcgg gacacgggtt    1200
cgatctgctc cgctaactcg ttgatgcctc tggatttctt cactctgtcg attgcgttca    1260
tggtctattg cctcccggta ttcctgtaag tcgatgatct gggcgttggc ggtgtcgatg    1320
ttcagggcca cgtctgcccg gtcggtgcgg atgccccggc cttccatctc caccacgttc    1380
ggccccaggt gaacaccggg caggcgctcg atgccctgcg cctcaagtgt tctgtggtca    1440
atgcgggcgt agtggaaagc aagcatcatt gacaggttgg aatggtaggc aattgcatcg    1500
agggttatgc taagcaatgc attgggattg agcgcttagg tattcgtgca aacgccattc    1560
tacgggtctt tgcagttgta aagcttgatc aactgagagg agggccgatc gatgccgtct    1620
ttgatccgct cggagatcat caggtcgcag tgcgggttct cgccgccacc ggcatcgatg    1680
gacagagtct aaggcaggcg atcggcaacg gtcatgtgct gggcgaactc ggacgcaagc    1740
gccttctgct ggtcgagggt cagctcgacc ggcagggcca attcgacctc cttgaacagc    1800
agcccattgg agcgttacta aaggtcggca gattccaagt agtcggaggg aagctcgacg    1860
aactccggca tgtgcccgga ttcggcgtgc aagacttcat ccatgtcgcg ggcatacttg    1920
ccttcgcgct ggatgtagtc ggcgttggcc ctggccgatt ggcggccgga cctgctgccg    1980
```

-continued

```
gttttagccg tacggtgcta aatcgcaatg atgcctagcc gttgcctttg atttacggct    2040
aaacgaaatg gaccctggcg agcgcaacgc aagcaaggtg gacgttaggc aagtaaatcg    2100
aagagaaacc ggtaagtgcg ccctccccta caaagtaggg tcgggattgc cgccgctgtg    2160
cctccatgat agcctacgag acagcacatt aacaatgggg tgtcaagatg gttaagggga    2220
gcaacaaggc ggcggatcgg ctggccaagc tcgaagaaca acgagcgcga atcaatgccg    2280
aaattcagcg ggagcgggca agggaacagc agcaagagcg caagaacgaa acaaggcgca    2340
aggctgatgg tggggcaatg atttcggcca aggtgaaatg aagcgagtgg aaggaggatc    2400
ggctcatggc ggctatggct gagtacattg aacgcgaaca agtacgcgcc ttgttcggtc    2460
tgccgccacg ccagaaggat gagccgggct gaatgatcga ccgagacagg ccctgcgggg    2520
ctgcacacgc gccccaccc ttcgggtagg gggaaaggcc gctaaagcgg ctaaaagcgc    2580
tccagcgtat ttctgcgggg tttggtgtgg ggtttagcgg gcttggcccg ccttcccccc    2640
tgccgcgcag cggtggggcg gtgtgtagcc tagcgcagcg aatagaccag ctatccggcc    2700
tatggtaggg aatacttggg aagggcagca gagccccaca agggcgatgc tacacgcgcc    2760
tagctgatta ttatcagcta atacaggatg gatttttcaa aaacccagcc agccccagcc    2820
cctgctgggc ttgcaggttt gggggcgtga cagttattgc aggggttcgt gacagttatt    2880
gcaggggggc gtgacagtta ttgcaggggc tcgtgacagc tagtacggga gtgacgggca    2940
ctggctggca atgtctagca acggcaggca tttcggctga gggtaaaaga acttaccgct    3000
aagccgatga ctgtatgtaa acaccagtat gcaaggaagc ggaacatgcc tcatgtggcg    3060
gacaggacgg aacgcaggga tagggataca ggtcgttacc agacgccacg aaccgagaca    3120
aaccttatcc atccgatcga tgcagagtac taccgggata tagctgagct tatggcagag    3180
cagggaaagg aattgccggg ctatgtgcaa cgggaatttg aagaatttct ccaatgcggg    3240
cggctggagc atggcttttt acgggttcgc tgcgagtctt gccacgccga gcacctggtc    3300
gctttcagaa atcaatctaa agtatatatg agtaaacttg gtctgacagg ccccggaatt    3360
cgcatctaga tggtagagcc aaaaacagcc ggtacaagca acgatctcca ggaccatctg    3420
aatcatgcgc ggatgaaacg aactcaagcc ggcgataaca gacataaacc aaacgtaacg    3480
acactgcgac aacgtggcaa ttcgtcgcaa tacggtctca ctgaactggc cgataattgc    3540
agacgaacgc tttggcagtt tattcttgac atgtagtgag gggggtggta taatcacata    3600
gtactgttat acagaaacag aggagatatt acatatggat agcactgaga acgtcatcaa    3660
ggagttcatg cgcttcaagg tgcacatgga gggctccgtg aacggccacg agttcgagat    3720
cgagggcgag ggcgagggcc gcccctacga gggcacccag accgccaagc tgaaggtgac    3780
caagggtggc cccctgccct tcgcctggga catcctgtcc cctcagttca tgtacggctc    3840
caaggcctac gtgaagcacc ccgccgacat ccccgactac ttgaagctgt ccttccccga    3900
gggcttcaag tgggagcgcg tgatgaactt cgaggacggc ggcgtggtga ccgtgaccca    3960
ggactcctcc ctgcaggacg gcgagttcat ctacaaggtg aagctgcgcg gcaccaactt    4020
cccctccgac ggccccgtaa tgcagaagaa gaccatgggc tgggaggcct cctccgagcg    4080
gatgtacccc gaggacggcg ccctgaaggg cgagatcaag cagaggctga agctgaagga    4140
cggcggccac tacgacgctg aggtcaagac cacctacaag gccaagaagc ccgtgcagct    4200
gcccggcgcc tacaacgtca acatcaagtt ggacatcacc tcccacaacg aggactacac    4260
```

-continued

agtggtgggg aagtaagagc gagccgaggc aagccaaaac atgttccaga atcacggtat 4320

ggacggattg taacagaacg aagaatgta aatccgtaac agttaatacg aaccaaaagt 4380

aaaaacccgc ttcggcgggt tttttatggg gggagtttta gggaaagagc atttgtcagc 4440

tggatcagca gatggcaacc ggatactcct cagcctccag agtagccaat gagacgccga 4500

gcggccgctt atttgccgac taccttggtg atctcgcctt tcacgtagtg gacaaattct 4560

tccaactgat ctgcgcgcga ggccaagcga tcttcttctt gtccaagata agcctgtcta 4620

gattcaagta tgacgggctg atactgggcc ggcaggcgct ccattgccca gtcggcagcg 4680

acatccttcg gcgcgatttt gccggttact gcgctgtacc aaatgcggga caacgtaagc 4740

actacatttc gctcatcgcc agcccagtcg ggcggcgagt tccatagcgt taaggtttca 4800

tttagcgcct caaatagatc ctgttcagga accggatcaa agagttcctc cgccgctgga 4860

cctaccaagg caacgctatg ttctcttgct tttgtcagca agatagccag atcaatgtcg 4920

atcgtggctg gctcgaagat acctgcaaga atgtcattgc gctgccattc tccaaattgc 4980

agttcgcgct tagctggata acgccacgga atgatgtcgt cgtgcacaac aatggtgact 5040

tctacagcgc ggagaatctc gctctctcca ggggaagccg aagtttccaa aaggtcgttg 5100

atcaaagctc gccgcgttgt ttcatcaagc cttacggtca ccgtaaccag caaatcaata 5160

tcactgtgtg gcttcaggcc gccatccact gcggagccgt acaaatgtac ggccagcaac 5220

gtcggttcga gatggcgctc gatgacgcca actacctctg atagttgagt cgatacttcg 5280

gcgatcaccg cttccctcat aatgtttaac tttgtttag ggcgactgcc ctgctgcgta 5340

acatcgttgc tgctccatta acatcaaacc tcgacccacg gcgtaacgcg cttgctgctt 5400

gatgcccgag gcatagactg taccccaaaa aaacagtcat aacaagccat gaaaaccgcc 5460

actgcgccgt taccaccgct gcgttcggtc aaggttctgg accagttgcg tgagcgcata 5520

cgctacttgc attacagctt acgaaccgaa caggcttatg tccactgggt tcgtgccttc 5580

atccgtttcc acggtgtgcg tcacccggca acctgggca gcagcgaagt cgaggcattt 5640

ctgtcctggc tggtcatgac caaaatcccc tcaccgtgag tagcctgccg ccttgggccg 5700

ggtgatgtcg taatgcaacg aagcgaactc ggttacagtc aagcccagcg agtacagatc 5760

cggcaacgcc tcgcgcaacc gcttgcggcg cttgcgcatg gtcgaaccac tggcctctga 5820

cggccagaca tagccgcaca aggtatctat ggaagccttg ccggtttgcc cggggtcgat 5880

ccagccacac agccgctggt gcagcaggcg ggcggtttcg ctgtccagcg cccgcacctc 5940

gtccatgctg atgcgcacat gctggccgcc acccatgacg gcctgcgcga tcaaggggtt 6000

aagggaaacg taacggagcc agtccgcctc gtcgctgggc tactccgacc gacgccgacc 6060

aacctgacgc ttgcggacct tatggcgact gatggatacc ttccaaaggc gatcgatgcc 6120

gtcctgtatg tgctgcagcg accaacacct atcgacatcc gcaacgattt accttgcacg 6180

cgcccgatag ctaccttga ccacatggca ttcagcggtg acggcctccc acttgggttc 6240

caggcaacgc aggacatgcc gtccgcattc ggtctgggct taaggcaacc gaccaaggcc 6300

attaggccca gccatggcca ccagcccttg caggatgcgc agatcatcag cgcccagcgg 6360

atccgggccg atgcaatcga taagcatgcc gtcgccgtcg tcatacgtcc agtccagatt 6420

gctgcgcttg cgctcgcccc gcttgagggc acggaacagg ccgggggcca gacagtgcgc 6480

cgggtcgtgc cggacgtggc tgaggctgtg cttgttctta ggctcacca cggggcaccc 6540

-continued

```
aatgatatt gagatgaatc taacgcaagg agggcatgag aaaaacgaag taatgcagcc    6600

tgaacaaacg atcagagaac ggtgcgaaca agttggaata gatcacaacg aagcggaaga    6660

agaaccggcg ctggtcgtcg tccacacccc attcctcggc ctcggcgctg gtcatgctcg    6720

acaggtagga ctgccagcgg atgttatcga ccagtaccga gctgccccgg ctggcctgct    6780

gctggtcgcc tgcgcccatc atggccgcgc ccttgctggc atggtgcagg aacacgatag    6840

agcacccggt atcggcggcg atggcctcca tgcgaccgat gacctgggcc atggggccgc    6900

tggcgttttc ttcctcgatg tggaacaggc gaagcgtgtc aagcaacatc aggcgaggc    6960

aatcggaggc gagcatgagg aagtcgaacc aatccgggc aatgatgttg ggcaggatgc    7020

cgatcagcgg ctggatcagc aggccgtcag ccacggcttg ccgttcctcg gcgctgaggt    7080

gcgccccaag ggcgtgcagg cggtgatgaa tggcggtggg cgggtcttcg gcgggcaggt    7140

agatcaccgg gccggtgggc agttcgccca cctccagcag atccggcccg cctgcaatct    7200

gtgcggccag ttgcagggcc agcatggatt taccggcacc accgggcgac accagcgccc    7260

cgaccgtacc ggccaccatg ttgggcaaaa cgtagtaaag cggtggggc gatgcagc    7320

aagcctacg aatatgata ggctatggg tagccattga ttgcctaatt tgcaggaagt    7380

tggcggttag gcgctggcgg ggtcactacc cccgccctgc gccgctctga gttcttccag    7440

gcactcgcgc agcgcctcgc attcgtcgtc ggtcagccag aactcgccgt gacgcatccc    7500

tttggccttc atgcgctcgg catatcgcgc ttggcgtaca gcgtcagggc tggccagcag    7560

gtcgccggtc tgcttgtcct tttggtcttt catatcagtc accgagaaac ttgccggggc    7620

agaaaggctt gtattagcgg aaacaggaca aggtgcagcc gtaaaggtta aggcaggaaa    7680

tatcagaga tgaaagcgg aaagcatcgg aattgattga agtaaaaaa aagcaaagg    7740

gcaaccaata gcccctgtca cttttgatca ggtagaccga cccgaaagcg cttttttcgt    7800

attccataaa acccccttcc gtgcgcgagc actcatagta taacaggcgt gagtaccaac    7860

gcaagcacta catgctgaaa tctggcccgc ccctgtccat gcctcgctgg cggggtgccg    7920

gtgcccgtgc cagctcggcc cgcgcaagct ggacgctggg cagacccatg accttgctga    7980

aggtgagc                                                             7988
```

<210> SEQ ID NO 88
<211> LENGTH: 7999
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized. pBTK370

<400> SEQUENCE: 88

```
taagcttcga ggccgggctt gagcatgcc agcgcagggc tggcctaggc catggcattg      60

aagattcca aggcaatgcg gaacccggtg gaaagcttca gagcggagat aaagacgaac     120

ttgctgaggc catgaccgaa gcgcttgacc agcccggcca tctcgctgcg gtactcgtcc     180

agcgcagtgc gaacgtggcg gataagatgc agatcgggca gttcgaggat ggcaagaatg     240

cgggccttct cctgctgccg ctgggcctgc tcgatctgct ggccagcctg ctgcaccagc     300

gaagggaacg aggtgaaggc attgccattg gattcaagca gaagcaaaca aggcagatac     360

ccggcgcggg tggtgtgctt gtccttgcgg ttggtgaagc ccgccaagcg gccatagtgg     420

cggctgtcgg cgctgcgcgg gtcggcgtcg tactcgctgg ccagcgtccg ggcaatctgc     480
```

-continued

```
aaccgaagtc aaccgaatgc gggcacgggc aacttgaacc atgccagatc gttccaaggg     540
atggcttcca atcccagggc aggcacaggc aactcggatc tactgcaatc aaggacaatc     600
tcgctgaggc cgtcacacag cacaagacca tgccgctccc gctcggcggg cctgatatac     660
acgtcattgc cctgggcatt catcgcttg  agccatgggc tgttctggag cacttcgggg     720
gctgaccatc ccgggttcat catctggccg gtgggtgcgt ccctgacgcc gatatcgaag     780
cgctcacagc ccatggcctc gagctgtcgg cctatggcct gcaaagtcct gtcgttcttc     840
atcggaacac aaagcgaagc aagaacgagc agtcacaggc tgtcagtggc gtcaggtaga     900
gaaagagcaa agatgagatc agcagcacca aagtaggaac aatggaagcc agcaacaagg     960
ttagccatag cttccagtgc cacccccgcc acgcgctccg ggcgctctgc gcggcgctgc    1020
tcacctcggc ggctacctcc cgcaactctt tggccagctc caccatgccg gcccctgtct    1080
ggcgctgggc tttcagcccc tccgccgcct gcgcctcgct ggcctgcttg gtctggctca    1140
tgacctgccg ggcttcgtcg gccagtgtcg ccatgctctg ggccagcggt tcgatctgct    1200
aagcaaactc gttgatgcca atggattcat tacactgtc  gattgcgttc atggacatt     1260
gaacaaggc attcctgtca gtcgatgatc tgggcgttgg aggtgaagat gttcaaggcc     1320
acgtctgccc ggtcggtgcc gatgcccggc ccttccatcc ccacacgtt  cggcccagg     1380
tgaacacggc gcaggcgctc gatgccctgc gcctcaagtc ttctcggtcg aatgcggcgg    1440
tcgtggccag cccgctctca tgcccggttg gcatggtcgg cccatgcctc gcgggtctgc    1500
tcaagccatg ccttgggctt gagcgcttcg gtcttctgtg ccccgccctt ctccggggtc    1560
ttgccgttga aacgcttgaa aaacagagcg gagggcagcc agatgcagtc attgacaagc    1620
taggagatca taaggtggca gtgcggttc  tagcagaaac aggcaaggat ggcaagagtc    1680
tacggcaggc gctcggcacc ggtcaggtgc tgggcgaacc cggacgccag cgccctctgc    1740
tggccgaggg tcagctcgac cggcaggca  aattcgaccc ccttgaacag ccgccccattg   1800
gcgcgttcat acaggtcggc agcatcccag tagtcggcgg gccgcacgac gaacaccggc    1860
atgacgccgg attcggcgtg caagacttca tccatgtcgc gggcatactt gcctacgcgc    1920
tggacgtacc aggcattggc aatggcagcc tggccgaacg aactgctgcc ggttccagcc    1980
gtaacgtgat aaatcgccat gctgcctcgc tgttgctttt gcttcacggc tccatgcaat    2040
ggccctcgga gagcgcaccg cccgaaggcc ggccgctagg ccagcctctc gaagagaaac    2100
cggtaagtgc gccctcccct acaaagtagg gtcgggattg ccgccgctgt gcctccatga    2160
tagcctacga gacagcacat tacaatggg  gtgtcaagat ggttaagggg agcaacaagg    2220
aggcggatcg gatggaaacg atcgaagacc aacgagacgc aatcaatgcc gacccaagcc    2280
gggagagggc aagggaaccg cagcaagagc gaccgaacga aaccaggagc aaggcgatgg    2340
tgggggaacc gattctggcc aagggcaaca gacgcgagtg gacggaggat aggcacatgg    2400
cggcaatgga tgcgtacctt gaacgcgacc acgaccgcgc cttgtcaggt ctgccgccac    2460
gacagaagga tgagcagggc tgacagatcg aacgagaacg gaccagaggg gatgcaaacg    2520
cgcccccacc cttcgggtag ggggaaaggc cgctaaagcg gctaaaagcg ctcccagcgta   2580
tttcagaggg gtttggtgtg gggtccagcg ggatcggacc gacttaaacc atgacgagcc    2640
gcggtggggc ggtgcgtagc ctagcgcagc gaatagacca gctatccggc ctctggccgg    2700
gcatattggg caaggcagc  agcgcccac  aagggcgctg ataaccgcgc ctagtcgatt    2760
```

-continued

```
attcttagcc aatcatggcc ggctccttcc aacccccagc cagccccagc caatgctggg   2820

tttgcaggtc tggggcagtg aacgccattg cagggcttcg tgccagttct tgcccggggg   2880

cgtgacagtc attgcagggg ttcgcgacag ttagcacggg agtgacgggc actggctggc   2940

aatgcctagc aacggcaggc atttcggctg agggcaaaag aactcccgcc taagcgatag   3000

actgcatgta aaccacagtc tgccaggacg cggaacatgc ctcacgtggc ggccaggacg   3060

gccagccggg atcgggatac tggtcgttac cagagccacc gacccgagca aacccctctc   3120

tatcagatcg ttgccgagta ttcccggcca ttcgccgagc ttctggcagc gaccgggaag   3180

gaccggaccg gatcagtgca aagggcattc gacgaatttc tcccatgagg gaggccggcg   3240

catggctttc tacggcttcg ctgccagtcc tgccacgccg agcacctggt cgcccccaga   3300

aatcaatcta aagtatatac gagtaaactc ggtccgacag gcccccgaat tcgcatctag   3360

atggcagagc caccaacagc cggtacagac aacgatctcc aggaccatct gaatcatgcg   3420

cggatgacac gaactcacga cggcgatcac agacattaac ccacagtaca gacactgcga   3480

caccgtggca attcgtagca atcccgtctc actgaaatgg cagccaattg cagccgaacg   3540

ggtgaaacca aacggttgcc aacccgaacg aaccacggta agctgcaacc aatgaaaagc   3600

cagcgagtca atacagaaac agaggagata ttacacatgg atagcactga gaacggcatc   3660

aagcccttca tgcgcttcaa ggtgcacatg gagggctccg tgaacggcca cgagttcgag   3720

atcgagggcg tgggcgaggg caagccctac gagggcaccc agaccgccaa gctgcaagtg   3780

accaagggcg gccccctgcc cttcgcctgg gacatcctgt ccccccagtt cttctacggc   3840

tccaaggcgc aactcaagca aacggcagcc atcccggacc aactcaagcc gtcccccacc   3900

gagggcttca agtgggagcg agtgatgaac ttcgaggacg gaggcgtggt gaccgtgacc   3960

caggactcct ccctgcagga cggcaccctc atctaccacg tgaagttcat cggcgtgaac   4020

ttcccctccg acggccccgt aatgcagaag aagactctgg gctgggagcc ctccactgag   4080

cgcaactacc ccgcgcacgg cgtgctgaag ggcgagaacc acatgggcgt gaagctgaag   4140

ggcggcggcc actacctgtg tgagttcaag ttcatctaca tggccaagaa gcccgtgaag   4200

atgccaggcc aacctcacgc ggacctaacg atcgacatca aatccaaccc agcggaatcc   4260

accgtggtgg agcagtacga gcgcgccgag gcccgccacc acctgttcca gactcacggt   4320

atggacgaat tgtacaagca cgacgaattg taaatccgta atcgttaatc cgcaaataac   4380

gtaaaaaccc gcttcggcgg gttttttat ggggggagtt tagggaaaga gcatttgtca   4440

gctggatccg cagatgggaa ccggatactc ctcagcctcc agagtagcca atgagacgcc   4500

gagcggacgc ttcttgcccg aatcccttgc tgctcccgcc tttcacgtcg tggccaaatt   4560

atttccaacg atctgcgcgc gagggccagc gatccccttc ttgtccaagc tacgcctgtc   4620

tagcctaccg tatgacgggc tgctactggg caggcagggg atcccatgcc cagtccgacg   4680

cgacatcctt cggcgcgatt ttgccggtta ctgcgctgta ccaaatgcgg gacaacgtaa   4740

gcactacact tagcctatcg acagccacgc agggcgcaga gttccatagc gttaaggttt   4800

catttagcgc ctcaaatcga tcctgttcag gaaccggatc aaagagttcc tccgccgctg   4860

gacctaccaa ggcaacgcta tgttctcttg cttttgtcag caagatagcc agatcaatgt   4920

cgatcgtggc tggctcgaag atacctgcaa gaatgtcatt gcgctgccat tctccaaatt   4980

gcagttcgcg cttagctgga taacgccacg gaatgatgtc gtcgtgcaca acaatggtga   5040
```

-continued

```
attctaatgc gaggagactc tagctctctc aaggggaagc agtagttcca aacaggtagt    5100
tgaacaaagc tagccgagta gtttcatcta gactaaggta aaccgtaacc agctaatacc    5160
tatcactgtg tggcttcagg ccgccatcca ctgcggagcc gtacaaatgt acggccagca    5220
acgtcggttc gagatggcgc tcgatgacgc caactacctc tgatagttga gtcgatactt    5280
cggcgatcac cgcttccctc atactgttta actttgtttt agggcgactg ccctgctgcg    5340
taacatcgtt gctgctccat aacatcaaac atcgacccac ggcgtaacgc gcttgctgct    5400
tggatgcccg aggcatagac tgtacccaca aaccacagtc ataccaagcc atgaaaaacg    5460
aaactgagcc gttaccaccg atgcgttcgg taaaggttca ggaccagttg agtgagagca    5520
tacgctactt gcattacagc ttacgaacgg aacaggctta tgtccactgg gttcgtgcct    5580
tcatccgttt ccacggtgtg cgtcaccggg caaccttggg tagcagcgaa gtcgaggcat    5640
ttctgtcctg gctggtcatg accaaaatcc cttaacgtga gtcagcctgc cgccttgggc    5700
cgggtgatgt cgtacttgcc cgccgcgaac tcggttaccg tccagcccag cgcgaccagc    5760
tacggaacgg aatcgagcac aagcatgcgg agctagagca tggtcgaacc aatggcatat    5820
gacggaacga aatcgaagca aacggtatca atggaagcat tgccggtttt gacggggtcg    5880
atccagccac acagccgctg gtgcagcagg cgggcggttt cgctgtccag cgcccgcacc    5940
tcgtccatgc tgatgcgcac atgctgcccg ccacccatga cggcctgcgc gatcaagggg    6000
ttcagggcca cgtacaggcg cccgttcgcc tcgtcgctgg cgtactccga cagcagccga    6060
aacccctgcc gcttgcggcc attctgggcg atgatggata ccttccaaag gcgctcgatg    6120
aagtcatgta tgtgcttgag agccccacca atatcgaact atgcccagat ttcctatgcc    6180
agcgcaagca agctaactta gaccacatgg aattcagagg tgacggactc aaactagggt    6240
tccaggaaca gccggagctg ccgtccgcct tcggtcttgg gttccgggcc aagcactagg    6300
ccattaggcc cagccatggc caccagccct tgcaggatgc gcagatcatc agcgcccagc    6360
ggctccgggc cgctgaactc gatccgcttg ccgtcgccgt agtcatacgt cacgtccagc    6420
ttgctgcgct tgcgctcgcc ccgcttgagg gcacgggaca ggccgggggc cagacagtgc    6480
gacggtacga gacggaagtg gatgaggatg tgatagttca taggctaaca aacggggaac    6540
cccgctgctc ttgcgctgcc tctccagcac ggcgggcttg agcacccgcc cgtcatgccg    6600
cctgaaccac cgatcagcga acggtgcgcc atagttggcc ttgctcacac cgaagcggac    6660
gaagaaccgg cgctggtcgt cgtccacacc ccattcctcg gcctcggcgc tggtcatgct    6720
cgacaggtcg gactgcccgc ggatgttatc gacccagtac gagctgcccc ggctggcctg    6780
atgctggtag aatgcgaaca taatggaagc gaaccagatg gaatggtgca ggaacaagat    6840
agaagcaacg gtatcggcgg agatggaatc aatgcgaacg atgaactggg aaatggggcc    6900
gatggagtta tattcataga tgtggaaacg gagcaagatg taaagcaaca taaggcgcgg    6960
gcgctcggcg gcgcgcttga ggcgctcgaa ccactcgggc gcgatgatgt tgggcaggct    7020
gaagataagc ggatggatca gaaggcagtc agccacggca tgacgtaact aggcgatgag    7080
gtgcgcccca agggcgtgca ggcggtgatg aatggcggtg ggcgggtctt cggcgggccg    7140
gtagataacc gggccggtgg gaagtaagcc aaactcaagc agatcgggca agcctgaact    7200
ctgtgcggcc agttgcaggg ccagcatgga tttaccggca ccaccgggcg accccagcgc    7260
cccgaccgta ccggccacca tgttgggcaa aacgtagtcc agcggtggcg gcgctgctgc    7320
```

-continued

```
gatcgaatcc agatcattga taggcatatg ggtagcaatc gattgcatcc tttgcaggcc    7380

gttggtggtc agggcgatgg ggggtcacta accccgaacc gagccgatat gagttcttcc    7440

aggcactcgc gcagcgcctc gtattcgtcg tcggtcagcc agaacttgcg ctgacgcatc    7500

cctttggcct tcatgcgctc ggcatatcgc gcttggcgta cagcgtcagg gctggccagc    7560

aggtcgcccg tctgcttgtc cttttggtct ttcatatcag tcacccgaga acttgccggg    7620

gccgaaaggc ttgtcttcgc ggaacaagga caaggtgcag ccgtcaaggt taaggctggc    7680

catatccgcg aatgaaatgc ggccaggctc ggcctcgttc gacgtcatcc ccaccgcacc    7740

gggcaaccca tagccattgc aactttgcct caggcagacc gacccagatg agcttcttcc    7800

gtattccata aaacccccgt ctgtgcgtga gtactcatag tataacaggc gtgagtacca    7860

acgcaagcac tacatgctga aatctggccc gcccctgtcc atgcctcgct ggcggggtgc    7920

cggtgcccgt gccagctcgg cccgcgcaag ctggacgctg ggcagaccca tgaccttgct    7980

gacggtgcgc tcgatgtaa                                                 7999
```

<210> SEQ ID NO 89
<211> LENGTH: 12760
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesised, pDIK601

<400> SEQUENCE: 89

```
ggggccgaaa ggcttgtctt cgcggaacaa ggacaaggtg cagccgtcaa ggttaaggct      60

ggccatatcc gcgactgcca agcggcccgc atcggcactg tttgacgtat accccaagcc     120

accggcaccc caccgacccc tgtcactttc gatcaggtcg accgaccatg aagcgctttt     180

ttcgtattcc ataaaacccc cttctgtgcg tgagtactca tagtataaca ggcgtgagta     240

ccaacgcaag cactacatgc tgaaatctgg cccgcccctg tccatgcctc gctggcgggg     300

tgccggtgcc cgtgccagct cggcccgcgc aagctggacg ctgggcagac ccatgacctt     360

gctgacggtg cgctcgatgt aatccgcttc gtggccgggc ttgcgctctg ccagcgctgg     420

gatggcatcg gacctggccc tgcccgattc atcggcaatg aggcaccagc tggccagatt     480

ctgcgcggcg ataaagtcgc acttgctgag gtcatgaccg aagcgcttga ccagcccggc     540

catctcgctg cggtactcgt ccagcgccgt gcgccggtgg cggctaagct gccgctcggg     600

cagttcgagg ctggccagcc tgcgggcctt ctcctgctgc cgctgggcct gctcgatctg     660

ctggccagcc tgctgcacca gcgccgggcc agcggtggcg gtcttgccct tggattcacg     720

cagcagcacc cacggctgcc aacccgcgcg ggtggctgtc ttgtccttgc ggttggtgcc     780

gaccgccctg aggccatcgc gggcgctgtc gggcgctggc gggtcggagt agtcctcgct     840

ggccagcgtc agggcaatcc gacccgcgcg ttcaccgccc gagggcgtgg accccgtgcc     900

ccatgcctga tagttcttcg ggctggtttc cactaccagg gcaggctccc ggccctcggc     960

tttcatgtca tcaaggtcaa actcgctgcg gtcgtccacc agcaccagac catgccgctc    1020

ctgcttgggc ggcctgatat acacgtcatt gccctgggca ttcatccgct tgagcccatg    1080

cgtgtatgcg agcacttcgg aggcagcccc ttcccggttc atcacctggc cggtgcggtc    1140

gtccctgacg ccgatatcga agcgctcaca gcccatggcc ttgagctgtc ggcctatggc    1200

ctgcaaagtc ctgtcgttct tcatcgggcc accaagcgca gccagatcga gccgtcctcg    1260
```

-continued

```
gttgtcagtg gagtcaggtc gagcaagagc aacgatgaga tatgcagatc aacggtaggc    1320
atcatggatg aatgcatcac ggttagacct agctacaagt gacccccacg agaggcgatc    1380
cgggcgctct gcgcggcgct gctcacctcg gcggctacct cccgcaactc tttggccagc    1440
tccacccatg ccgccctgct ctggcgctgg gctttcagcc actccgccgc ctgcgcctcg    1500
ctggcctgct tggtctggct catgacctgc cgggcttcgt cggccagtgt cgccatgctc    1560
tgggccagcg gttcgatctg ctccgctaac tcgttgatgc ctctggattt cttcactctg    1620
tagattgcgt tactggtata ttgcctaccg gtatacatgt aagtcgatgc tatgggagtt    1680
ggcggtgtcg atgttaaggg aaacgtatgc aaggacggtg aggatgaccc ggcctaaact    1740
ctccaccacg ttcggcccca ggtgaacacc gggcaggcgc tcgatgccct gcgcctcaag    1800
tgttctgtgg tcaatgcggg cgtcgtggcc agcccgctct aatgcccggt tggcatggtc    1860
ggcccatgcc tcgcgggtct gctcaagcca tgccttgggc ttgagcgctt cggtcttctg    1920
tgccccgccc ttctccgggg tcttgccgtt gtacccgctg aaccactgag cggcgggccg    1980
atcgatgccg tactcgatcc gatcggagct catcaggtgc aagtgcgggt tatcgcagcc    2040
aacggaatgc atggcaagcg tatacggccg gagctcggca aaggtcaggt gatgggagtc    2100
ctcggacgcc agcgccttct gctggtcgag ggtcagctcg accggcaggg caaattcgac    2160
ctccttgaac agccgcccat tggcgcgttc atacaggtcg gcagcatccc agtagtcggc    2220
gggccgctcg acgaactccg gcatgatgcc ggattcggcg tgcaagactt catccatgtc    2280
gcggggcatc ttgccttcgc gctggatgta gtcggccttg gcccatgccg attggccgcc    2340
agtcctgctg aaggttttcg aagtaaggtg atcctaagcc atgctgactc gatgctgatt    2400
ttgcttttcg gatccatgca atggccatcg gagaacgatc agcccgaagg gtggccgtta    2460
ggccagtttc tcgaagagaa accggtaagt gcgccctccc ctacaaagta gggtcgggat    2520
tgccgccgct gtgcctccat gatagcctac gagacagcac attaacaatg gggtgtcaag    2580
atggttaagg ggagcaacca ggcggcggat cggcatggca agctcgaaga acaacgagcg    2640
cgaatcaatg ccgaaattca gcgggagcgg gcaagggaac agcagcaaga gcgcaagaac    2700
gatacaaggc gaacggtgct ggtgcgggcc atgccttggc aaacggtgcc aagcaagacg    2760
tggccggagg atcggctcat ggcggcaatg gatgcgtacc ttgaacgcga ccacgaccgc    2820
gccttgttcg gtctgccgcc acgccagaag gatgagccgg gctgaatgat cgacgcgagc    2880
aggccctgcg gggcatgcac cgcgccccca cccttcgggt aggggaaacg gccgctaacg    2940
cggctaaacg cgctccagcg tatttctgcg gggtttggtg tggggtttcg cgggctttgc    3000
aagcctttcc aactgaagcg aagcggtggg gaggagtgta gactagagca gagaataggc    3060
aagctatccg gactctggcc gggcaatatg ggctagggca gatgcgaacc aaacgggagc    3120
tgctaaaagc gactagtgga tttatacttg atcctcatgg atggattttt aaacacaacc    3180
gccagccccc gcccctgctg ggtttgcagg tttgcgggcg tgacagttat tgcaggggtt    3240
agtgaaagtc attgcagggg ggcgatgacg ttatagaagg ggttcgtgac agttagtacg    3300
gggatgacgg gcactggctg gcaatgtcta gcaacggcag gcattacggc tgagggtaaa    3360
agaactttcc gataagagct agacggtatg taaacaaagc attgcaaggc agcggaaact    3420
gcctcatgtg gcggccagga cggccagccg ggatcgggat actggtcgtt accagagcca    3480
ccgacccgag caaacccttc tctatcagat cgttgacgag tattacccgg cattcgctgc    3540
```

continued

```
gattatggca gagcagggta aggtatgcc gggctatgtg aatcgggact ttgcagaatt    3600

tatccaatgc gggcggatgg agcttggctt tatacgggtt agctgcgagt attgccaagc    3660

cgagcacctg gtcgctttca gaaatcaatc taaagtatat atgagtaaac ttggtctgac    3720

aggcccctga attcgcatct agacttgaact gggcgataat tgcagacgaa cggacagtca    3780

ttcatcttc tgcccctcca aaagcaaaaa cccgcccgaag cgggatttta cgtaaatccg    3840

gtgaaactga ccgatattac gactcactat aggaactaat ttagtaccct atccaagaat    3900

tatttacca ttggctaccc atgcgcatgc aatttcagcc gaccaagcgt atccgaact    3960

aattacgta gaggaaatca agcctgattg attcgaacg gattattagc tggtaaaatg    4020

cgatcctcgc catggaaaat atatggcgtg ttgtatgctg tacagaggag atgttgtacc    4080

taaggatgtt aatgcagcca ttgcaactat caagactaag cgcaccattc aatttgtcga    4140

ttggtgccca actggattca aagttggcat taattatcaa ccaccgactg ttgtaccagg    4200

tggtgatctt gctaaagtac aacgagctgt ttgcatgtta tccaatacta ctgctatcgc    4260

agtagattgg gactgattg accttaactt tgattaatg taagctaagc gagcattgt     4320

aatctggtac gtcggtgagg gtatggagga aggtgagttc tacgaagacc gatcccatct    4380

agtgagtcgc attagtaatc gttaatccgc aaataacgta aaaacccgct tcggcgggtt    4440

ttttcatggg gggagtttag ggaaagagca tttgtcagct ggctgccaac caaaccagat    4500

gtcaacaacg ctacaacgct gatcctcca ctcagcaaaa gttcgattta ttcaacaaag    4560

ccacgttgtg tctcaaaatc tctgatgtta cattgcacaa gataaaaata tatcatcatg    4620

aacaataacta atgtctgctt aactaaacag tactacaagg ggtgctatgg atacgaatta    4680

atcaataggc ttagctatcg gaccaaatcg agtcgcatgg gaggcgatcc atgccgaata    4740

taaggttcgg tctaaaaagc tcaaggttct gggaaataca gaccgccaca gtatcaaaaa    4800

aaatcttata ggggctcttt tatttgacag tggagagaca gcggaagcga ctcgtctaaa    4860

acggacagct cgtacaaggt atacacgtcg gaagaatcgt atttgttatc tacaggagat    4920

ttttttcaaat gagatggcga aagtagatga tagttttttt catcgacttg aagagtcttt    4980

tttggtggca gacgaacga agcatgcacg tactcctatc tttggaaatc tagtagatgc    5040

agttgcttac catgagaaat atccaactac ctatcatctg cgaaaaaaat tggtagattc    5100

tactgataaa gcggatttgc gcttaatcta tttggcctta gcgcatatga ttaagtttcg    5160

tggtcatttt ttgattgagg gagatttaaa tcctgataat agtgatgtgg acaaactatt    5220

tatccagttg gtacaaacct acaatcaatt atttgaagaa aacccctatta acgcaagtgg    5280

agtagatgcc aaagcgattc tttctgcacg attgagtaca tacagacgat tagcaaatct    5340

aattgctcag ctccccggtg agaagaaaca tggcttattt gggaatctca ttgctttgtc    5400

attgggtttg acccctaatt ttaaatcaaa ttttgatttg gcagaagatg ctaaattaca    5460

gctttcaaaa gatacttacg atgatgattt agataattta ttggcgcaaa ttggagatca    5520

atatgctgat ttgtttttgg cagctaagaa tttatcagat gctatttac tttcagatat    5580

cctaagagta aatactgaaa taactaaggc tcccctatca gcttcaatga ttaaacgcta    5640

agatgaacat aatcaagacc tgactcttt aaaagcttta gttcgacaac aacttccaga    5700

aaagtataaa gaaatctttt ttgatcaatc aaaaaacgga tatgcaggtt atattgatgg    5760

gggagctagc caagaagaat tttataaatt tatcaaacca atttagaaa aaatggatgg    5820
```

-continued

```
tactgaggta tttttggtga aactaaatcg tgatgatttg atgcgcaagc aacggaactt      5880
tgacaaaggc tatcttcccc atcaaattca attgggtgag atgcatgata ttttgagatg      5940
acaagaagac ttttatccat ttttaaaaga caatcgtgag aagattgaaa aaatcctgac      6000
ttttcgaatt ccttattatg ttggtccatt gggcgcgtgg aatagtcgtt ttgcatggat      6060
gactcgggag tctgaagaaa caattacccc atggaatttt gaagaagttg tcgataaagg      6120
tgcttcagct caatcattta ttgaacgcat gacaaacttt gataaaaatc ttccaaatga      6180
aaaagtacta cccaaacata gtttgcttta tgagtatttt acggtttata acgaattgac      6240
aaaggtaaca tatgttactg aaggaatgcg aaaaccagca tttctttcag gtgaacagaa      6300
gaaagccatt gttgatttac tcttcaaaac aaatcgaaaa gtaaccgtta agcaattaaa      6360
agaagattat ttcaaaaaaa tagaatgttt tgatagtgtt gaaatttcag gagttgaaga      6420
tagatttaat gcttcattag gtacctacca tgatttgcta aaaattatta aagataaaga      6480
ttttttggat aatgaagaaa atgaagatat cttagaggat attgttttaa cattgacctt      6540
atttgaagat agggagatga ttgaggaaag acttaaaaca tatgctcacc tctttgatga      6600
taaggtgatg aaacagctta aacgtcgccg ttatactggt tggggacgtt tgtctcgaaa      6660
attgattaat ggtattaggg ataagcaatc tggcaaaaca atattagatt ttttgaaatc      6720
agatggtttt gccaatcgca atttcatgca gctgatccat gatgatagtt tgacatttaa      6780
agaagacatt caaaaagcac aagtgtctgg acaaggcgat agtttacatg aacatattgc      6840
aaatttagct ggtagccctg ctattaaaaa aggtattttg cagactgtaa aagttgttga      6900
tgaattggtc aaagtaatgg ggcggcataa gccagaaaat atcgttattg aaatggcacg      6960
tgaaaatcag acaactcaaa agggccagaa aaattcgcga gagcgtatga aacgaatcga      7020
agaaggtatc aaagaattag gaagtcagat tcttaaagag catcctgttg aaaatactca      7080
attgcaaaat gaaaagctct atctctatta tctccaaaat ggaagagaca tgtatgtgga      7140
ccaagaatta gatattaatc gtttaagtga ttatgatgtc gatcacattg ttccacaaag      7200
tttccttaaa gacgattcaa tagacaataa ggtcttaacg cgttctgata aaaatcgtgg      7260
taaatcggat aacgttccaa gtgaagaagt agtcaaaaag atgaaaaact attggagaca      7320
acttctaaac gccaagttaa tcactcaacg taagtttgat aatttaacga aagctgaacg      7380
tggaggtttg agtgaacttg ataaagctgg ttttatcaaa cgccaattgg ttgaaactcg      7440
ccaaatcact aagcatgtgg cacaaatttt ggatagtcgc atgaatacta aatacgatga      7500
aaatgataaa cttattcgag aggttaaagt gattacctta aaatctaaat tagtttctga      7560
cttccgaaaa gatttccaat tctataaagt acgtgagatt aacaattacc atcatgccca      7620
tgatgcgtat ctaaatgccg tcgttggaac tgctttgatt aagaaatatc caaaacttga      7680
atcggagttt gtctatggtg attataaagt ttatgatgtt cgtaaaatga ttgctaagtc      7740
tgagcaagaa ataggcaaag caaccgcaaa atatttcttt tactctaata tcatgaactt      7800
cttcaaaaca gaaattacac ttgcaaatgg agagattcgc aaacgccctc taatcgaaac      7860
taatggggaa actggagaaa ttgtctggga taaagggcga gattttgcca cagtgcgcaa      7920
agtattgtcc atgccccaag tcaatattgt caagaaaaca gaagtacaga caggcggatt      7980
ctccaaggag tcaattttac caaaaagaaa ttcggacaag cttattgctc gtaaaaaaga      8040
ctgggatcca aaaaaatatg gtggttttga tagtccaacg gtagcttatt cagtcctagt      8100
```

-continued

```
ggtcgatccg gtggaaaccg ggccatcgca gacgccaaca tacgccaacg agttactagg     8160
gatcaacatc atggaaagca gttcccttga aaccaatccg attgactttt tagcagatcc     8220
aggatataag gaagttaaaa aagacttaat cattaaacta cctaaatata gtctttttga     8280
gttagaaaac ggtcgtaaac ggatgctggc tagtgccgga gaattacaaa aaggaaatga     8340
gctggctctg ccaagcaaat atgtgaattt tttatattta gctagtcatt atgaaaagtt     8400
gaagggtagt ccagaagata acgaacaaaa acaattgttt gtggagcagc ataagcatta     8460
tttagatgcg attcctgcgc aactcagtga atttcctacg agtgcctatt tagcagatgc     8520
acccctagcc aacgctatta gtgcatatca caaccataga gacccaaccc tacgcgaacc     8580
agcagaaaat attattcatt tatttacgtt gacgaatctt ggagctcccg ctgcttttaa     8640
atattttgat acaacaattg atcgtaaacg atatacgtct acaaaagaag ttttagatgc     8700
cactcttatc catcaatcca tcactggtct ttatgaaaca cgcattgatt tgagtcagct     8760
aggaggtgac tgaatccaaa gcaagctgat aaacccgatc aattaaaggc tccttttgga     8820
gaccttttcc ttggagattc tacccatgca gaccccatta tttgacgatc agctagcggc     8880
ggggaactgc aagccatcca atccaaccca aggctcagtc ggccgactgg gactcctgtt     8940
ttatctgttg tttgtcggtg aacactctcc cggctgagca gttacagaga tgttacgaac     9000
ccccaggaca tccgagaatg cgaggcgatg gagggtacaa cccgagcggc cgcttatttg     9060
ccgactaact tggtgatctc gcctttcacg tagtggacaa attctttcaa ctgatctgcg     9120
cgcgaggcca agcgatcttc ttcttgtcca agataagcct gtctagcttc aagtatgacg     9180
ggctgatact gggccggcag gcgctccatt gcccagtcgg cagcgacatc attcggcgcg     9240
attctgccgg ttactgcgct gtaccaaatg cgggacaacg taagcactac atttcgctca     9300
tcgccagccc agtcgggcgg cgagttccat agcgttaagg tttcatttag cgcctcaaat     9360
agatcctgtt caggaaccgg atcaaagagt tcctccgccg ctggacctac caaggcaacg     9420
ctatgttctc ttgcttttgt cagcaagata gccagatcaa tgtcgatcgt ggctggctcg     9480
aagatacccg caagaatgtc attgcgctgc cattctccaa attgcagttc gcgcttagct     9540
ggataaagcc acggaatgat gtcgtcgtgc acaacaatgg tgacttctac agcgcggaga     9600
atctcgctct ctccagggga agccgaagtt tccaaaaggt cgttgatcaa agctcgccgc     9660
gttgtttcat caagccttac ggtcaccgta accagcaaat caatatcact gtgtggcttc     9720
aggccgccat ccactgcgga gccgtacaaa tgtacggcca gcaacgtcgg ttcgagatgg     9780
cgctcgatga cgccaactac ctctgatagt tgagtcgata cttcggcgat caccgcttcc     9840
ctcatgatgt ttaactttgt tttagggcga ctgccctgct gcgtaacatc gttgctgctc     9900
cataaaatca aacctagccc caacgcgtca agcgctgcct gattggatgc cagcggcatc     9960
gaccgtaccc caccaaacca gtcctaacca gcacgcaaca cagcagtgcc gacgccaacc    10020
ccgctgcgtt cggtcaaggt tctggaccag ttgcctgagc gcatacgcta cttgcattac    10080
agcctaacga aagcaaaggc ttatgtaccc tgggctagtg cattcatccg tttccacggt    10140
gtgcgtcacc cggcaatcct gggtagcagc gaagtcgagg catttctgtc ctggctggtc    10200
atgaccaaca tacccatacg tgagtcagcc tgccgcattg ggccgggtgc tgtcgtaatt    10260
gcccgccgcg aactcggtta ccgtcaagcc cagcgcgacc agctccggca acgcctcgcg    10320
cacccgcttg cggcgcttgc gcatggtcga accactggcc tctgacggcc agacatagcc    10380
```

-continued

```
gatcaaggta tatctggatc atttgcaggc tttgccgggc tagttcaagc aatcagacag   10440
atggcgatgc aggcgggcgg tttcgctgtc aagcgcaagc aactcgtaat tgctgatgcg   10500
cacatgctgg ccgccaccca tgacggcctg cgcgatcaag gggttcaggg ccaccgacag   10560
gcgcccgtcc gcctcgtcgc tggcgcactc cgacagcagc cgaaacccct gccgcctgcg   10620
gccattctgg gcgatgatgg ataccttcca aaggcgctcg atgcagtcct gtatgtgctt   10680
gagcgcccca ccactatcga cctctgcccc gattttcttt gccagcgccc gatagctacc   10740
tttgatctca tggcattcag aggtgatggc atcccacttg ggttccaggc aatgcccgcg   10800
atgccgtccg aattcggtct tgggtttcgg gacaagaacc aggccattag gacaagaact   10860
ggccaccagc ccttgcagga tgcgcagatc atcagcgccc aggggctccg ggccgctgaa   10920
ctcgatccgc ttgccgtcgc cgtagtcata cgtcacgtcc agcttgctgc gcttgcgctc   10980
gccccgcttg agggcacgga acaggccggg ggccagaccg tgcgccgggt cgtgccggac   11040
gtggctgagg ctgtgcttgt tcttaggctt caccacgggg cacccccttg ctcttgcgct   11100
gaatctaatc aacggagggc ttgatgcacc agccgtaatg aagcatgatc aacgatatag   11160
agtacggtgc gaattagttg gactagtcaa aacgaagcag gacgaagatc aggcgctggt   11220
cgtcgtccac acgccattcc tcggcctcgg cgctggtcat gctcgacagg taggactgcc   11280
agcggatgtt atcgaccagt accgagctgc cccggctggc ctgctgctgg tcgcctgcgc   11340
ccatcatggc cgcgcccttg ctggcatggt gcaggaacac gatagagcac ccggtatcgg   11400
cggcgatggc ctccatgcga ccgatgacct gggccatggg gccgctggcg ttttcttcct   11460
agcagtggca aaggcgatgc gtgtccagca aaatcaggcg gacggcctcg gaggcgagct   11520
tgcggaagtc gatccaatcc ggggccatga tgttgggacg gatgccgatc agcggctggc   11580
tcagcaggcc gtcagccacg gcttgccgtt cctcggcgct gaggtgcgcc ccaagggcgt   11640
gcaggcggtg atgaatggcg gtgggcgggt cttcggcggg caggtagatc accggcccgg   11700
tgggcagttc gcccacctcc agcagatccg gcccgcctgc aatctgtgcg gccagttgca   11760
gggccagcat ggattttacc gcaccaccgg gcgacaccag cgccccgacc gtaccggcca   11820
aactgttggg caacaagtag taacgcggtg gaggcgatgc tgcgaaagcc taacgaatat   11880
tgataggctt atgggtagcc attgattgcc tcctttgcag gcagttggtg gttaggcgct   11940
gggggggtca ctaccccggc cctgcgccgc tctgagttct tccaggcact cgcgcagcgc   12000
ctcgtattcg tcgtcggtca gccagaactt gcgctgacgc atccccttgg ccttcatgcg   12060
ctcggcatat cgcgcttggc gtacagcgtc agggctggcc agcaggtcgc cggtctgctt   12120
gtcctttggc tatttaatct aagtcaacga gaatcatgcc                         12160
```

<210> SEQ ID NO 90
<211> LENGTH: 8600
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pR K614

<400> SEQUENCE: 90

```
gatttagcct atagaaaggg gtatttacgg atattcatgc gacaaaaaat tgtgtggaat      60
acacgctcgg atgactgcat gatgaccgca ctgactgggg acagcagatc cacctaagcc     120
tgtgagagaa gcagacaccc gacagatcaa ggcagttacg tctcaccaac caaaccagat     180
```

continued

```
gtcaaaaccg atcaaagcca gatcctacaa atcagcaaaa gttcgatttc ttcaacaacg     240
aaacgttgtg tatcaaaatc tatgaagtta aattgcaaaa gataaaaatc tatcaacatg     300
aacaataaaa ctgtctgctt acataaacag taatacaagg ggtgttatgg ataagaaata     360
ctcaataggc ttagctatcg gcacaaatag cgtcggatgg gcggtgatca ctgatgaata     420
taaggttccg tctaaaaagt tcaaggttct gggaaataca gaccgccaca gtatcaaaaa     480
aaatcttata ggggctcttt tatttgacag tggagagaca gcggaagcga ctcgtctcaa     540
acggacagct cgtagaaggt atacacgtcg gaagaatcgt atttgtatct acaggagatt     600
ttttcaaatc gagatggcga aagtagatga tagtttcttt catcgacttg aagagtcttt     660
tttggtggaa gaagacaaga agcatgaacg tcatcctatt tttggaaata tagtagatga     720
agttgcttat catgagaaat atccaactat ctatcatctg cgaaaaaaat tggtagattc     780
tactgataaa gcggatttgc gcttaatcta tttggcctta gcgcatatga ttaagtttcg     840
tggtcatttt ttgattgagg gagatttaaa tcctgataat agtgatgtgg acaaactatt     900
tatccagttg gtacaaacct acaatcaatt atttgaagaa aaccctatta acgcaagtgg     960
agtagatgct aaagcgattc tttctgcacg attgagtaaa tcaagacgat tagaaaatct    1020
cattgctcag ctccccggtg agaagaaaaa tggcttattt gggaatctca ttgctttgtc    1080
attgggtttg acccctaatt ttaaatcaaa ttttgatttg gcagaagatg ctaaattaca    1140
gctttcaaaa gatacttacg atgatgattt agataattta ttggcgcaaa ttggagatca    1200
atatgctgat ttgtttttgg cagctaagaa tttatcagat gctattttac tttcagatat    1260
cctaagagta aatactgaaa taactaaggc tcccctatca gcttcaatga ttaaacgcta    1320
cgatgaacat catcaagact tgactctttt aaaagcttta gttcgacaac aacttccaga    1380
aaagtataaa gaaatctttt ttgatcaatc aaaaaacgga tatgcaggtt atattgatgg    1440
gggagctagc caagaagaat tttataaatt tatcaaacca attttagaaa aaatggatgg    1500
tactgaggaa ttattggtga aactaaatcg tgaagatttg ctgcgcaagc aacggacctt    1560
tgacaacggc tctattcccc atcaaattca cttgggtgag ctgcatgcta ttttgagaag    1620
acaagaagac ttttatccat ttttaaaaga caatcgtgag aagattgaaa aaatcttgac    1680
ttttcgaatt ccttattatg ttggtccatt ggcgcgtggc aatagtcgtt ttgcatggat    1740
gactcggaag tctgaagaaa caattacccc atggaatttt gaagaagttg tcgataaagg    1800
tgcttcagct caatcattta ttgaacgcat gacaaacttt gataaaaatc ttccaaatga    1860
aaaagtacta ccaaaacata gtttgcttta tgagtatttt acggtttata acgaattgac    1920
aaaggtcaaa tatgttactg aaggaatgcg aaaaccagca tttctttcag gtgaacagaa    1980
gaaagccatt gttgatttac tcttcaaaac aaatcgaaaa gtaaccgtta agcaattaaa    2040
agaagattat ttcaaaaaaa tagaatgttt tgatagtgtt gaaatttcag gagttgaaga    2100
tagatttaat gcttcattag gtacctacca tgatttgcta aaaattatta aagataaaga    2160
ttttttggat aatgaagaaa atgaagatat cttagaggat attgttttaa cattgacctt    2220
atttgaagat agggagatga ttgaggaaag acttaaaaca tatgctcacc tctttgatga    2280
taaggtgatg aaacagctta aacgtcgccg ttatactggt tggggacgtt tgtctcgaaa    2340
attgattaat ggtattaggg ataagcaatc tggcaaaaca atattagatt ttttgaaatc    2400
agatggtttt gccaatcgca atttatgcag ctgatccatg atgatagttt gacatttaaa    2460
```

continued

```
agaagaattc aaaaaagcac aagtgaatgg aaaaggagca agtttaaatg aaaaaattgc    2520
aaaaatagca ggtagaaatg atataaaaaa aggtatttta aagaaagtaa aagtagttgc    2580
tgaattggtc aaagtaatgg gggggcataa gccagaaaat atcgttattg aaatggcacg    2640
tgaaaatcag acaactcaaa agggccagaa aaattcgcga gagcgtatga aacgaatcga    2700
agaaggtatc aaagaattag gaagtcagat tcttaaagag catcctgttg aaaatactca    2760
attgcaaaat gaaaagctct atctctatta tctccaaaat ggaagagaca tgtatgtgga    2820
acaagaatta gatattaatc gtttaagtga ttatgatgtc gatcacattg ttccacaaag    2880
tttccttaaa gacgattcaa tagacaataa ggtcttaacg cgttctgata aaaaccgtgg    2940
taaatcggat aacgttccaa gtgaagaagt agtcaaaaag atgaaaaact attggagaca    3000
acttctaaac gccaagttaa tcactcaacg taagtttgat aatttaacga aagctgaacg    3060
tggaggtttg agtgaacttg ataaagctgg ttttatcaaa cgccaattgg ttgaaactcg    3120
ccaaatcact aagcatgtgg cacaaatttt ggatagtcgc atgaatacta aatacgatga    3180
aaatgataaa cttattcgag aggttaaagt gattacctta aaatctaaat tagtttctga    3240
cttccgaaaa gatttccaat tctataaagt acgtgagatt aacaattacc atcatgccca    3300
tgatgcgtat ctaaatgccg tcgttggaac tgctttgatt aagaaatatc caaaacttga    3360
atcggagttt gtctatggtg attataaagt ttatgatgtt cgtaaaatga ttgctaagtc    3420
tgagcaagaa ataggcaaag caaccgcaaa atatttcttt tactctaata tcatgaactt    3480
cttcaaaaca gaaattacac ttgcaaatgg agagattcgc aaacgccctc taatcgaaac    3540
taatggggaa actggagaaa ttgtctggga taaagggcga gattttgcca cagtgcgcaa    3600
agtattgtcc atgccccaag tcaatattgt caagaaaaca gaagtacaga caggcggatt    3660
ctccaaggag tcaattttac caaaaagaaa ttcggacaag cttattgctc gtaaaaaaga    3720
ctgggatcca aaaaaatatg gtggttttga tagtccaacg gtagcttatt cagtcctagt    3780
ggttgctaag gtggaaaaag ggaaatcgaa gaagttaaaa tccgttaaag agttactagg    3840
gatcacaatt atggaaagaa gttcctttga aaaaaatccg attgactttt tagaagctaa    3900
aggatataag gaagttaaaa aagacttaat cattaaacta cctaaatata gtctttttga    3960
gttagaaaac ggtcgtaaac ggatgctggc tagtgccgga gaattacaaa aaggaaatga    4020
gctggctctg ccaagcaaat atgtgaattt tttatattta gctagtcatt atgaaaagtt    4080
gaagggtagt ccagaagata acgaacaaaa acaattgttt gtggagcagc ataagcatta    4140
tttagatgag attattgagc aaatcagtga attttctaag cgtgttattt tagcagatgc    4200
caatttagat aaagttctta gtgcatataa caaacataga gacaaaccaa tacgtgaaca    4260
agcagaaaat attattcatt tatttacgtt gacgaatctt ggagctcccg ctgcttttaa    4320
atattttgat acaacaattg atcgtaaacg atatacgtct acaaaagaag ttttagatgc    4380
cactcttatc catcaatcca tcactggtct ttatgaaaca cgcattgatt tgagtcagct    4440
aggaggtgac tgatcttcta gaaagctgat aaaccgatac aattaaaggc tccttttgga    4500
gccttttttt ttggagattt tcaacatgaa gaaattatta tttgatgatc agatagcggc    4560
ggggaaatgc aagaaataaa ataaaaaaaa aggcaaagtc ggaagaatgg gcatatagtt    4620
ttatctgttg tttgtcggtg aacactctcc cggctgagca tgagacggaa atctgctcgt    4680
cagtggtgct cacactgacg aatcatgtac agatcataac gatgactgcc tggcgactca    4740
```

continued

```
aacttaagca agtcagacgg aaccagagcc ggcgaacacc actgcatatc tggcatatcc   4800

aatcagtcca actagtgcac tgcagtaacg aggccgagcc tatcaaaaag gatcttcacc   4860

tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact   4920

tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt   4980

cgttcatcca tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta   5040

ccatctggcc ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta   5100

tcagcaataa accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc   5160

gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat   5220

agtttgcgca acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt   5280

atggcttcat tcagctccgg ttcccaacga tcaaggcgag ttacatgatc ccccatgttg   5340

tgcaaaaaag cggttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca   5400

gtgttatcac tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta   5460

agatgctttt ctgtgactgg tgagtactca accaagtcat tctgagaata gtgtatgcgg   5520

cgaccgagtt gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact   5580

ttaaaagtgc tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg   5640

ctgttgagat ccagttcgat gtaacccact cgtgcaccca actgatcttc agcatctttt   5700

actttcacca gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga   5760

ataagggcga cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc   5820

atttatcagg gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa   5880

caaatagggg ttccgcgcac atttccccga aaagtgccac ctgtcatgac caaaatccct   5940

taacgtgagt tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa aggatcttct   6000

tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca   6060

gcggtggttt gtttgccgga tcaagagcta ccaactcttt ttccgaaggt aactggcttc   6120

agcagagcgc agataccaaa tactgttctt ctagtgtagc cgtagttagg ccaccacttc   6180

aagaactctg tagcaccgcc tacatacctc gctctgctaa tcctgttacc agtggctgct   6240

gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt accggataag   6300

gcgcagcggt cgggctgaac ggggggttcg tgcacacagc ccagcttgga gcgaacgacc   6360

tacaccgaac tgagatacct acagcgtgag ctatgagaaa gcgccacgct tcccgaaggg   6420

agaaaggcgg acaggtatcc ggtaagcggc agggtcggaa caggagagcg cacgagggag   6480

cttccagggg gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca cctctgactt   6540

gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa cgccagcaac   6600

gcggcctttt tacggttcct ggccttttgc tggccttttg ctcacatgtt ctttcctgcg   6660

ttatcccctg attctgtgga taaccgtgcg gcggcccct                          6689
```

<210> SEQ ID NO 91
<211> LENGTH: 2507
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized. pBTK515

<400> SEQUENCE: 91

-continued

```
gacttagcct atcgatggta gagccactca cagccggtcc aagcaacgct atccaggacc     60
atctgaatca tgcgcggctg accagaactc acgacggaga ttccagacct taccccaacg    120
tacagacact gcgacaacgt ggcaattcgt cgcaatacgg tctcactgaa ctggccgata    180
attgcagacg aacgctttgg cagtttattc ttgacatgta gtgaggggcc tggtataatc    240
accatacgtc aattccacga ggattgtttt agagctagaa atagcaagtt aaaataaggc    300
tagtccgtta tcaacttgaa aaagtggcac cgagtcggtg ctttttttga attctctaga    360
gtcgaactgc agctgattag atctattacc ctgttatccc tactcgagtt catgtgaagc    420
tccattagca tccgtaatcg tttttcagca aattacgtca aatcccgatt aggcctgttt    480
ttttatgggg ggagtttagg gaaagagcat ttgtcagctg gctgccaatg agacgacggg    540
gtcatcaggg ctcatcatgc gccaaacaaa tgtgtgcaat acacgctcgg atgactgcat    600
gatgaccgca ctgactgggg acagcagatc caccctaagc tgtgagagaa gcagacaccc    660
gacagatcaa ggcagttaac tagtgcactg cagtacagcg gccgcgatta tcaaaaagga    720
tcttcaccta gatcctttta aattaaaaat gaagttttaa atcaatctaa agtatatatg    780
agtaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct    840
gtctatttcg ttcatccata gttgcctgac tccccgtcgt gtagataact acgatacggg    900
agggcttacc atctggcccc agtgctgcaa tgataccgcg agacccacgc tcaccggctc    960
cagatttatc agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa   1020
ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc   1080
cagttaatag tttgcgcaac gttgttgcca ttgctacagg catcgtggtg tcacgctcgt   1140
cgtttggtat ggcttcattc agctccggtt cccaacgatc aaggcgagtt acatgatccc   1200
ccatgttgtg caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc agaagtaagt   1260
tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc   1320
catccgtaag atgcttttct gtgactggtg agtactcaac caagtcattc tgagaatagt   1380
gtatgcggcg accgagttgc tcttgcccgg cgtcaatacg ggataatacc gcgccacata   1440
gcagaacttt aaaagtgctc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga   1500
tcttaccgct gttgagatcc agttcgatgt aacccactcg tgcacccaac tgatcttcag   1560
catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa   1620
aaaagggaat aagggcgaca cggaaatgtt gaatactcat actcttcctt tttcaatatt   1680
attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga   1740
aaaataaaca aataggggtt ccgcgcacat ttccccgaaa agtgccacct gtcctgaccc   1800
aaaccattta aagtgagttt tagtcaacct gagcgcaaga accagtagcc aagctcaacg   1860
gatcctattg agctcatttt tttctgagcg tactctgatg ctgcaaacca aaccaaaacc   1920
cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt ccgaaggtaa   1980
ctggcttcag cagagcgcag ataccaaata ctgttcttct agtgtagccg tagttaggcc   2040
accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaatc ctgttaccag   2100
tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga cgatagttac   2160
cggataaggc gcagcggtcg ggctgaacgg ggggttcgtg cacacagccc agcttggagc   2220
gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc gccacgcttc   2280
```

continued

```
aagaagggag aaaggaggac aggtatccgg taagcggcag ggtcggaaca ggagagcgca    2340

agcgggagct tccaggggga aacgcctggt atctttatag tcctgtcggg tttcgccacc    2400

tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta tggaaaaacg    2460

ccagcaacgc ggccttttta cggttcctgg ccttttgctg gccttttgct cacatgttct    2520

ttcctgcgtt atcccctgat tctgtggata accgtgcggc cgcccct                  2567
```

<210> SEQ ID NO 92
<211> LENGTH: 8054
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized, pDIK619

<400> SEQUENCE: 92

```
acagatcata ccgatgactg cctggcgact cacaactaag caagacagcc ggaaccagcg      60

ccggcgaaca ccactgcata tatggcatat cacaacagtc caactagtgc actgcagtcc     120

aacaggaatc aactaaaacg aatggctacg tagcaagacc gggcctttcg ttttatcagt     180

tgtttgtcgg tgaacgatcc atcctagagc aaacctggcc aacctcaggg tgggcctttc     240

tgcgtttata ccgagcggac gcgtgttaca accaattaac caattctgat tagaaaaact     300

catcgagcat caaatgaaac tgcaatttat tcatatcagg attatcaata ccatattttt     360

gaaaaagccg tttctgtaat gaaggagaaa actcaccgag gcagttccat aggatggcaa     420

gatcctggta tcggtctgcg attccgactc gtccaacatc aatacaacct attaatttcc     480

cctcgtcaaa aataaggtta tcaagtgaga aatcaccatg agtgacgact gaatccggtg     540

agaatggcaa aagcttatgc atttctttcc agacttgttc aacaggccag ccattacgct     600

cgtcatcaaa atcactcgca tcaaccaaac cgttattcat tcgtgattgc gcctgagcga     660

gacgaaatac gcgatcgctg ttaaaaggac aattacaaac aggaatcgaa tgcaaccggc     720

gcaggaacac tgccagcgca tcaacaatat tttcacctga atcaggatat tcttctaata     780

cctggaatgc tgttttcccg gggatcgcag tggtgagtaa ccatgcatca tcaggagtac     840

ggataaaatg cttgatggtc ggaagaggca taaattccgt cagccagttt agtctgacca     900

tctcatctgt aacatcattg gcaacgctac ctttgccatg tttcagaaac aactctggcg     960

catcgggctt cccatacaat cgatagattg tcgcacctga ttgcccgaca ttatcgcgag    1020

cccatttata cccatataaa tcagcatcca tgttggaatt taatcgcggc ctggagcaag    1080

acgtttcccg ttgaatatgg ctcataacac cccttgtatt actgtttatg taagcagaca    1140

gttttattgt tcatgatgat atatttttat cttgtgcaat gtaacatcag agattttgag    1200

aaacaaagtg gatttgttga ataaatagaa atttagatga gttgaaggat cagtcatgcc    1260

aaaaatcacc tacagtgacg aagcctgccg aatttggacg ggtgacgtcg tactagaacg    1320

ccgcgaactc ggttaccgtc cagcccagcg cgaacagctc cggcaacgcc tcggcgaccg    1380

gatcgaggcg attgcgcttg gtcgaaccac tggcctatga aggccagact tagccgaacc    1440

aggtatctca ggaagccttg ccggtcttgc cggggtcgat ccagcccacc agcccgctggt   1500

gaagcaggcg ggcggtttcg atgtccaggc aacgcaactc gtccatgatg atgcgcaact    1560

gctggccgcc acccatgacg gcctgcgcga tcaaggggtt cagggccacg tacaggcgcc    1620

cgtccgcctc gtcgctgggg tactccgaca gcagccgaaa cccctgccgc ttgcggccat    1680
```

-continued

```
tatgggagtt gatggatacc ttcaaaaggc gatcgatgca gtacctgtag tgcttgagcg    1740
aaccaaaact atcgaactct gaccggattt aatttgaaag agccggatag ataccttgct    1800
ccacatggca ttcagcggtg acggcctccc acttggcttc caggaacagc cggagctgcc    1860
gtccgccttc ggtcctgggt tccgggccaa gcacgaggcc attaggccca gccatggcca    1920
ccagcccttg caggatgcgc agatcatccg cgcccagcgg ctccgggccg ctgaactcga    1980
tccgcttgcc gtcgccgtcg tcatacgtca cgtccagctt gctgcgcttg cgctcgcccc    2040
gattgagggc aaggaaaagg aagggggccc gaccgtgagc agggacgtgc aggacgtggc    2100
tgagggatgt attgatatta ggcttcaccc agggggcacc aattgctatt gagctgaatc    2160
tccagcaggc cgggcttgag caccccgccg tcatgccgcc tgaaccaccg atcagcgaac    2220
ggtgcgccat agttggcctt gctcacaccg aagcggacga agaaccgggg ctggtcgtcg    2280
tccacacccc attcctcggc ctcggcgctg gtcatgctcg acaggtaggc ctgccagcgg    2340
atgttatcga ccagtaccga gctgccccgg ctggcctgct gctggtcgcc tgcgcccatc    2400
atggccagcc aattgatggc atggatgagg aaccggatag agcacaaggt atcggcgggc    2460
atggcataca tgcgaaagct gaccttgggc atggggaagc tggcgtttta ttcctcgatg    2520
tggaaccggc gcagcgtgtc cagcaccatc aggcggcggc cctcggcggc gcgcttgagg    2580
ccgtcgaacc actccggggc catgatgttg ggcaggctgc cgatcagcgg ctggatcagc    2640
aggccgtcag ccacggcttg ccgttcctcg gcgctgaggt gcgccccaag ggcgtgcagg    2700
cggtgatgaa tggcggtggg cgggtcttcg gcgggcaggt agatcaccgg gccggtgggc    2760
agttcgaccc aatccagcag atccggaccg aatgcaatcc gtgcggaaag ttgcaggccc    2820
agcctggatt taccggaccc aacggggagc aaacgcgacc agaccgtacc ggccaacatg    2880
ttgggcaaaa cgtagtccag cggtggcggc gctgcctcga acgcctccag aatattgata    2940
ggcctatggg tagccattga ttgcctcctc tgcaggcagc tggtggttag gcgctggcgg    3000
ggtcactacc cccgccctgc gccgcttcga gttcttcctg gcactcgcgc agcgccctgt    3060
attcgtcgtc ggtcagcccg aacttgcgct gacgcatccc tttggccttc atgcgctcgg    3120
aatatagcgc ttggcgtaca gagtcagggc tggccagaag gtagccggta tgcttgtact    3180
tttggtcttc catatcagtc accgagaaac ttgccggggc cgaaagcctt gtcttcgcgg    3240
aacaaggaca aggtgcagcc gtcaagctta aggctggcca tatcagcgac tgaaaagcgg    3300
ccagcctcgg ccttgtttga cgtcataccc aagccaccgg gcaacccaat gcccttgtca    3360
cttttgatca ggtcgaccga ccctgaagcg ctttttctgt attccataaa acccccttct    3420
gtgcgtgagt aatcatagta tacaaggcgt gagtacaaac gaaagcaata aatgctgacc    3480
tatggaacgc aactgtaacc gactcgatgg agggctgacg gtgcccgtgc aagctcggcc    3540
agagcaagct ggacgatggg aagacccatg aactcgatga aggtgcgatc gatgcaatcc    3600
gcttcgtggc cgggcttggc ctctgccagc gctgggctgg cctcggccat ggcctcgccg    3660
atttcatcgg aactgaggcc aaggctggcc agcttctgcg aggcgataac gtcgcaattg    3720
ctgaggtcat gaccgaagcg cttgaccagc ccggccatct cgctgcggta ctcgtccagc    3780
gaactgagcc ggtgcaggct aagctgaagc tagggcagtc agaggctggc aagcctgagg    3840
gccttctcct gctgccgctg ggcctgctcg atctgctggc cagcctgctg caccagcgcc    3900
gggccagcgg tggcggtctt gcccttggat tcacgcagca gcacccacgg ctgataacgg    3960
```

-continued

```
gagcgggtgg tgtgcttgtc attgcggttg gtgcagcacg cacggcggcc atcgcggagg    4020
atgccgggcg tggccgggtc gggcacgtcc tagccggcca gagtccgggc aatccgcacc    4080
cgaagttcac cgcccgcggc gtcggccacc ttgacccatg cctgatagtt cttgcggctg    4140
gtttccacta ccaggcaggc ctcccggccc tcggctttca tgtcatccag gtcaaactcg    4200
ctgaggtcgt cccaccgcac cagaccatgc cgctcctgct cggcgggcct gatatacacg    4260
tcattgccct gggcattcat ccgcttgagc catggcgtgt tctggagcac ttcggcggct    4320
gacccatccc ggttcatccc atggccggtg ggtgcgtccc tgccgccgct atcgaagcgc    4380
tcccagcccc tggccttgcg atgtcggccc atggcctgcc aagtcctgtc gttccccatc    4440
gggccaccaa gcgcagccag atcgagccgc cctcggttgc cagtggcgtc aggtcgagca    4500
agagcaacga tgcgatcagc agcaccaccg taggcatcat ggaagccagc atcacggtta    4560
gccatagctt ccagtgccac ccccgcgacg cgctccgggc gctctgcgcg gcgctgctca    4620
cctcggcggc taccttccgc aactccttgg ccagcttcac ccatgccgcc cctgtctggc    4680
gatggcattc cagccactcc gacgcctgcg catcgctggc atgcctggtc tggcccatgc    4740
catgccgggc ttcgccggcc agtgccgccc tgctccgggc cagcggttcg atctgctccg    4800
ctaactcgtt gatgcctctg gatttcttca ctctgtcgat tgcgttcatg gtctattgcc    4860
tcccggtatt cctgcaagtc gatgatctgg gcgttggcgg tgtcgatgtt caggcccacg    4920
tctgcccggt cggtgcggat gccccggcct tccatctcca ccacgttcgg ccccaggtga    4980
accccgggca ggcgctcgat gccctgcgcc tcaagtgttc tgtggtcaat gcgggcgtcg    5040
tggccagccc gatctaatgc caggttggca tggtcggccc atgcctagcg ggtctgatcc    5100
agccatgccc tgggcttgcg agctccggtc ttctgtgccc agcccctatc cggggctatg    5160
ccgttgtacc gcttgaacca ctgagcggcg ggccgctcga tgccgtcatt gatcgcgtcg    5220
gagatcatca ggtggcagtg cgggttctcg ccgccaccgg catggatggc cagcgtatac    5280
ggcaggcgct cggcaccggt cagggtgctg gcgaactcgg acgccagcgc cttctgctgg    5340
tcgagggtca gctcgaccgg cagggcaaat tcgaccttct tgaacagccg cccattggcg    5400
agttcatcca ggtcggatgc atcccagtcg taggcggccc gatcgcagcc atccggcatg    5460
tgcccggatt cggcgtgcaa gacttcatcc atgtcgcggg catacttgcc ttcgcgctgg    5520
atgtagtcgg ccttggccct ggcggattgg ccgcccgacc tgctgccggt tttcgccgta    5580
aggtgatcaa tcgccatgct gcctcgctgt tgcttttgct tttcggctcc atgcaatggc    5640
cctcggagcg cgcaccgccc gaagggtggc cgttaggcca gtttctcgaa gagcaaccgg    5700
tacgtgagcc atcccatcca aagtagggtc gggattgacg cagcctgccc tcccggatcg    5760
catacgagcc agcccattca cactggggtg tacggatggt tacgggagcc acccaggagg    5820
aggatcggct ggccaagctc gacgaacccc gagcgcgact cactgccgcc attcagaggg    5880
agccggcaag ggaacagcag caagagcgca agaacgaaac aaggcgcaag gtgctggtgg    5940
gggccatgct tttggcccgc gtgcacagca gagcgtggcc ggcggatagg atcctgcagg    6000
caatggatgc gtaccttgaa cgcgaccacg accgcgcctt gttcggtctg ccgccacgcc    6060
agcaggatga gacgggatga atgctcgccc gagccaggcc atgcgggcct gcccacgagc    6120
ccccaccctt cgggtagggg gaaaggccgc taaagcggct aaaagcgctc cagcgtattt    6180
ctgcgggttc tggtgtgggg tttagcgggc tttgcccgcc tttccccctg ccgcgcagcg    6240
```

-continued

```
gtgggagggt gtgtagacta gagcagagca tagacccagc atccggaatc tggccgggct    6300

tatcgggcta gggcagaagc gacccaaacg ggcgctgata aacgcgaatc gtggactatt    6360

cttagataat catggatgga tttttccaac accccgccag cccccgcccc tgctgcgttt    6420

gcaggtttgg gggcgtgaca gttattgcag gggttcgtga cagttattgc aggggggcgt    6480

gacagttatt gcaggggttc gtgacagtta gtacgggagt gacgggcact ggctggcaat    6540

gtctagcaac ggcaggcatt tcggctgagg gtaaaagaac tttccgctaa gcgatagact    6600

gtatgtaacc accgtattgc aaggacgcgg aacctgaatc atgtgggggc aaggacggcc    6660

agcccggatc gggccaatgg tagttaaccg agccaacagc cagagcaaac aattcaatat    6720

cagatcgttg acgagtatta cccggcattc gctgcgctta tggcagagca gggaaaggaa    6780

ttgccgggct atgtgcaacg ggaatttgaa gaatttctgc aatgcgggcg gctggagcat    6840

ggcattctac gggttcgctg cgagtcttgc cacgccgagc accctggtgc tttcagaaat    6900

caatcctaag tatatatgag taaacttggt ctgacaggcc cctgaattcg catctagatg    6960

gtagagaccc aaccagaccg tacctagacc gatctcaagg accctctgct tactgcgagg    7020

atgacaagca atcccgaccg agatcaacga aattaaacca aagtacagcc aatgcgcacc    7080

cgtggcaatt cgtcgcaata ccgtctcact gaactggcgc ataattgcag acgaacgcat    7140

atacaagttt atttctgaca ctagtcggcc aaaatgatat aatacctgag tactgttata    7200

cagaaacaga ggagatatta catatgagta aaggagaaga gcttttcact ggagttgtcc    7260

caattcttgt ggaattagac ggtgatgtta atgggcacaa gttctctgtc agtggagagg    7320

gtgaaggcga agcaacatac ggaaaactta cccttaaatt tatttgcact actggaaaac    7380

taccagttcc atggccaaca ctggtcacta cgttcgggta tggggttcag tgcttcgcgc    7440

gctacccaga tcacatgaaa cagcacgact ttttcaagag tgccatgccc gaaggctatg    7500

tacaggagag aaccatcttt tttaaggatg acggcaacta taagacacgc gccgaagtga    7560

agttcgaggg tgataccctt gttaatagaa tcgagttaaa gggtattgac tttaaggaag    7620

atggaaatat tttaggccac aaactggaat ataactataa ctcccataat gtgtacatta    7680

tggccgacaa gcaaaagaac ggtatcaagg ttaacttcaa gatcagacac aacattgagg    7740

atggaagcgt tcaactagcc gaccattacc aacaaaacac cccaattggc gatggccctg    7800

tgctgttacc agacaaccat tacctgtcca ctcaatctgc cctttcgaaa gatcccaacg    7860

aaaagcgcga ccacatggtc cttcttgagt ttgtcacggc tgctgggatt acacacggca    7920

tggatgaact atacaaataa atcctaacta gcataacccc ttggggcctc taaacgggtc    7980

ttgaggggtt ttttgctggc tgagcatgag aaggaaatct gatcgtaagt ggtgcttacc    8040

atgacgatcc atgt                                                     8054
```

<210> SEQ ID NO 93
<211> LENGTH: 8054
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Chemically Synthesized. pBTK820

<400> SEQUENCE: 93

```
actgcctggc gactcacaac taagcaagac agccggaacc agcgccggcg aacaccactg      60

catatatggc atatcacaac agtccaacta gtgcactgca gtacaccagg catcaaataa     120
```

-continued

```
aacgaaaggc tacgtagcta gactggcccc ttcgccttcc atgtcgtttg taggcgaacg     180
atccctacta gagtcaaccc ggctcaacta cggggcgggc tttccgcgtt tatcccgagc     240
ggccgcgtgc taaaaccaat taaccaattc tgattagaaa aactcatcga gcatcaaatg     300
aaactgcaat ttattcatat caggattatc aataccatat ttttgaaaaa gccgtttctg     360
taatgaagga gaaaactcac cgaggcagtt ccataggatg gcaagatcct ggtatcggtc     420
tgcgattccg actcgtccaa catcaatcca acctattaat ttcccctcgt caaaaataag     480
gttatcaagt gagaaatcac catgagtgac gactgaatcc ggtgagaatg gcaaaagctt     540
atgcatttct ttccagactt gttcaacagg ccagccatta cgctcgtcat caaaatcact     600
cgcatcaacc aaaccgttat tcattcgtga ttgcgcctga gcgagacgaa atacgcgatc     660
gctgttaaaa ggacaattac aaacaggaat cgaatgcaac cggcgcagga acactgccag     720
cgcatcaaca atattttcac ctgaatcagg atattcttct aatacctgga atgctgtttt     780
cccggggatc gcagtggtga gtaaccatgc atcatcagga gtacggataa aatgcttgat     840
ggtcggaaga ggcataaatt ccgtcagcca gtttagtctg accatctcat ctgtaacatc     900
attggcaacg ctacctttgc catgtttcag aaacaactct ggcgcatcgg gcttcccata     960
caatcgatag attgtcgcac ctgattgccc gacattatcg cgagcccatt tatacccata    1020
taaatcagca tccatgttgg aatttaatcg cggcctcgag caagacgttt cccgttgaat    1080
atggctcata acaccccttg tattactgtt tatgtaagca gacagtttta ttgttcatga    1140
tgatatattt ttatcttgtg caatgtaaca tcagagattt tgagacacaa cgtggctttg    1200
ttgaataccc agcccttttg ctgagttgaa ggatcagtca tgcccaaact aactaaacgt    1260
gagccagccc gacgcattgg gacgggtgcc gtcgcacttg cccgccgagc actcggttcc    1320
cgtccagccc agcgcgaccg gctccggcaa cgcctcgcgc accggcttgc gggcgctgcg    1380
catggtcgaa ccactggccc ctgacggcca gacatagccg cacaaggtat ctatggaagc    1440
cttgccggtt ttgccggggt cgatccagcc acacagcccc tggtgcagca ggcgggcggt    1500
ttcgctgtcc agcgcccgca cctcgtccat gctgatgcgc acatgctggc cgccacccat    1560
gacggcatgc gagcccaagg ggttcagggc caacgccagg agcccgtacg actcgccgct    1620
ggcgtactcc gacagcagcc gaaacccctg ccgcttgcgg ccattctggg cgatgatgga    1680
taccttccaa aggcgctcga tgcagtcctg tatgtgcttg agcgccccac cactatcgac    1740
ctctgccccg atttcctttg ccagcgcccg atagctacct ttgaccacat ggcattcagc    1800
ggtgacggcc tcccacttgg gttccaggaa cagccggagc tgccgtccgc cttcggtctt    1860
gggttccggg ccaagcacta ggccattagg cccagccatg gccaccagcc cttgcaggat    1920
gcgcagatca tcagcgccca gcggctccgg gccgctgaac tcgatccgct tgccgtcgcc    1980
gtagtcatac gtcacgtcca gcttgctgcg cttgcgctcg cccgcttgcc gggcacggcc    2040
caggccgggg gccagacagt gcgccgggtc gtgccggacg tggccgaggc tgtgcctgtt    2100
cttagcattc cacccggggc caccccttgc tattgcgatg cctcccagc aaggccggct    2160
tgagcacccc gccgtcatgc cgcctgaacc accgatcagc gaacggtgcg ccatagttgg    2220
cctaccattc aagcagaccc ggcgctggtc gtcgccaacc cacccatcct ccccgaacct    2280
cggcctcggc gctggtcatg ctcgacaggt aggactggcc gcggatgtta tcgaccagta    2340
ccgagctgcc ccggctggcc tgctgctggt cgcctgcgcc catcatggcc gcgcccttgc    2400
```

-continued

```
tggcatggtg aagggaaacg atcgagaacc aggtataggc gggatggcc tacctgagcc   2460
agacgaactg ggccatgggg aagcatgcgt tttcatcatc gatgctgacc aggcgcagcg   2520
tgtccagcac catcaggcgg cggccctcgg cggcgcgctt gaggccgtcg aaccactccg   2580
gggccatgat gttggcgcagg ctgccgatca gcggctggat cagcagcgcg tcagccacgg   2640
cttgccgttc ctcggcgctg aggtgcgccc caagggcgtg caggcggtga tgaatggcgg   2700
tgggcgggtc ttcggcgggc aggtagatca ccgggccggt gggcagttcg ccaccctcca   2760
gacgatacgg aacgcatgca atctgcgcgg aacgatgacg ggcaagcatg gatttaacgg   2820
aaccaacggg agccaaacgc gaccgaccg taccggaacc aatgctgggc aaccgtagt    2880
ccagcggtgg cggcgctgct gcgaacgcct ccagaatatt gataggctta tgggtagcca   2940
ttgattgcct cctttgcagg cagttcgtgg ttaggcgctg gcggggtcac taccccgcc    3000
ctgcgccgct ctgagttctt ccaggcactc gcgcagcgcc tcgtattcgt cgtcggtccg   3060
ccgagacttg cgctgacgca tcccttggc cttcatgcgc tcggcatatc gcgcttggcg    3120
taaagagtca gggctggcca gacggtagcc ggtcagattg tactcatggt atttcatatc   3180
agtcaaagac aacctggccg gggccgaacg gattgtcatc gaggaaaacg gacaaggtgc   3240
agccgtcaag gttaaggcgc gcgatatcag cgacgaaaa gcggccagcc tcggccttgt    3300
ttgacgtata acgaaagcca ccgggcaacc aatagccctt gtcactttg atcaggtaga    3360
ccgaccctga agcgctttt tcgtattcca taaaaccccc ttctgtgcgt gagtactcat    3420
agtataaacg gcgtgagtac caacgcaagc actacatgct gaaatctggc ccgccccctgt  3480
aaacgaatcg atggcggggc gacggcgccc gtgccagatc ggcccgagcc agctggaagc   3540
tgggcagacc aatgaacttg atgacggtgc gatcgatgta atccgcttcg tggccgggct   3600
tgcgctctgc cagcgctggg ctggcctcgg ccatggcctt gccgattccc tcggcactgc   3660
ggcccggcc ggccagcttc tgcgcggcga taaagtcgca cttgctgagg tcatgaccga    3720
agcgcttgac cagcccggcc atctcgctgc ggtacacgtc cagcgccgtg cgccggtggc   3780
ggctaagctg ccgctcgggc agttcgaggc tggccagcct gcgggccttc tcctgctgcc   3840
gatggaatcg atcgatatgc tggccagccc gatgcaaacg agccaggacc gaggtagagg   3900
tcttgccctt ggattcacgc agcagcaccc acggctgata acggccgcgg gtggtgtgct   3960
tgtccttgcg gttggtgaag cccgccaagc ggccatagtg gcggctgtcg gcgctgcccg   4020
ggtcggcgtc gtactcgctg gccagcgtcc gggcaatctg ccccgaagt tcacccgcctg   4080
cggcgtcggc caccctgacc catgcctgat agttcatcgg gctggattcc actacccggg   4140
aaggctaacg gacctaggca ttacagtaca aacggtaaca atcgcagagg tagtccaacc   4200
gacccagaca atgccgatcc tgatcagcgc gactgatata aacgacattg aactggaact   4260
taaacagatc gagccatggc gtgtacatgga gacactaggc ggatgaaact taccggttac   4320
tcatctggcc ggtggctgcc tcgctgacgc cgatatcgaa gcgctcacag cccatggcct   4380
tgacgatgac gactatggcc tgacaagtcc tgtcgctatc aatcaggaca aacagcgacg   4440
ccgatcgcag ccgtcctcgg ttgtcagtgg cgtcaggtcg agcaagagca acgatgcgat   4500
aagcagaacc aacgtaggca tactgaacgc aagcatacag gttagcaatc gattccagtg   4560
ccacccccgc gacgcgctcc gggcgctctg cgcggcgctg ctcaccctcgg cggctaccctc 4620
ccgcaactcc ttggccagcc ccacccatgc cgccctgtc tggcgctggg ctttcagcca    4680
```

-continued

```
atccgacgcc tgcgcatcgc tggccagctc ggtcgggatc atgcactgcc gggcaacgtc    4740
ggccagtgtc gaccagatcc gggccagcgg ttcgatatgc tacgcaaact agttgatgcc    4800
tctggatttc ttcactctgc cgattgcgtt catggtctac tgcctccggg tattcctgta    4860
agtcgatgac ctggccgttg gcggcgtcga tgttcagggc cacgtctgcc cggtcggtgc    4920
ggatgccccg gcctgccatc tccaccacgt tcggccccag gtgaacaccg ggcaggcgct    4980
cgatgccctg cgcctcaagt gttcgtggtc caatgcgggc gtcgtggcca gcccgctcta    5040
atgccaggtc ggcaaggtcg gacccagccc agcgggtatg atccagaact gactagggct    5100
tgaacgattc ggtcatatgc gacccgaccc tatcagggga attgccgttg taccgcttgc    5160
accactgagc ggcggcgcgc tcgatgccgc cattgatccg ctggagatcc atcaggtggc    5220
agtgcgggtc ctcgccgcca ccggcatgga tggccagcgc atacggcagg cgctcggcac    5280
cggtcaggtg ctgggcgaac tcggacgcca gcgcctcctg ctggtcgagg gtcagctcga    5340
ccggcagggc aaatccgacc tcctcgaaca gccgcccatc ggcgcgttca tacaggtcgg    5400
aagcatacca gtacgaggcg ggccgctcga agacacaagc aatgcgaacg gattcggagt    5460
gacagaattc atccatgtcg agggcataca tgccacagcg atggacgtag taggccttgg    5520
cccggcccga ttggccgccc gacccgctgc cggtcctcgc cgtaacgtga taaatcgcca    5580
tgcggcctcg ctgttgcttc tgccttcggc ctccatgcaa tggccctcgg agagcggcac    5640
gcccgaaggg tggccgttag gccagcttct cgaagagaaa ccggtaagtg cgcccctccc    5700
tacaaagtag ggtcgggatt gccgccgctg tgcctccatg atagcctacg agacagcaca    5760
ttaacaatgc ggtgcacaga tggtcaaggc gagcaaacac gaggcggatc ggctgcaacc    5820
gatcgaagca aaacgagcgc gaatcaatgc agacaatacc aggggacggg aaagggaacc    5880
gcagcaagag cgcaagaacg aaacaaggcg caaggcgctg gtggggcgca tgatcctggc    5940
caaggtgaac agcagcgagc ggcgcgagga tcggcccatg gcggcaatgg atgcgcacct    6000
tgaacgcgac cacgaccgcg ccttgctcgg tctgccgcca cgccagaagg atgagccggg    6060
ctgaatgatc gaccgagaca ggccctgcgg ggctgcacac gcgcccccac ccttcgggta    6120
gggggaacgc aagcaaacgc ggctaaacgc gatccagagc atttcagagg ggttcggtgt    6180
ggggcttagc gggccttgcc cgcccctccc cctgccgcgc agcggcgggg cggtgcgtag    6240
cctagcgcag cgaacagacc agctacccgg cctcggcgcg ggcacattgg gcaaggcgag    6300
cagcgcccca caggggcgct gatcaccgcg cctagtggat tattcttaga taatcatgga    6360
tggatttttc cacacccccg ccagcccccg cccctgctgg gtttgcaggt ttgggggcgt    6420
gacagttatc gaacgggttc gtgacagtta ttgcacgggc gagtgacagt tattgcaggg    6480
gttcgtgcaa gttagtacgc gagtgaaggc aactggatgc aactgcatag aaacggaagg    6540
aataaaggcc gagggtacca gaacattccg atacgcgata gactgcatgt aacaacagta    6600
ttgcaaggac gcggaacatg cctcatgtgg cggccaggac ggccagccgg gatcgggata    6660
atggcagtta aacgagaacc agaaccgacc aaaacctacc atatcagatc gttgacgagt    6720
attacccggc attcgctgcg cttatggccg agcagggaaa ggaattgccg ggctatgtgc    6780
aacgggaatc tgaacgaatc atccaatgcg gggcgctgga gaatggattt ataacggttc    6840
gctgcgagtc ttgccacgcc gagcacctgg tcgcattcag aaatcaatct aaagtatata    6900
tgagcaaacc tggtctgaca ggcccctgaa ttcgcatcta gatggcagag ccacaaacag    6960
```

continued

```
aaggtaaatg aaacgatctc aaggacactc tgaatcatgc gaggatgaac agaatcaacg    7020

aacgcgatca aagacattca aaaacagtac agaaactgcg aaaacgtggc aattcgtagc    7080

aataccgtct cactgaactg gcagataatt gcagacgaac gcatacggga gtttattctt    7140

gacatattgc cggtgtgttg gtataataac ttagtactgt tatacagaaa cagaggagat    7200

attacatatg agtaaaggag aagaactttt cactggagtt gtcccaattc ttgttgaatt    7260

agatggtgat gttaatgggc acaaattttc tgtcagtgga gagggtgaag gtgatgcaac    7320

atacggaaaa cttaccctta aatttatttg cactactgga aaactacctg ttccatggcc    7380

aacacttgtc actactttcg gttatggtgt tcaatgcttt gcgagatacc cagatcatat    7440

gaaacagcac gactttttca agagtgccat gcccgaaggc tatgtacagg agagaaccat    7500

ctttttaaag gatgacggca actataagac acgcgccgaa gtgaagttcg agggcgatac    7560

cctggttaat agaatcgagt taaagggtat tgacttcaag gaagatggca acattctggg    7620

ccacaaactg gaatataact ataactccca taatgtgtac attatggccg acaagcaaaa    7680

gaatggcatc aaggttaact tcaagatccg ccacaacatt gaggatggaa gcgttcaact    7740

agcagaccat tatcaacaaa atactccaat tggcgatggc cctgtccttt taccagacaa    7800

ccattacctg tccactcaat ctgccctttc gaaagatccc aacgaaaagc gcgaccacat    7860

ggtccttctt gagtttgtca cggctgctgg gattacacac ggcatggatg aactatacaa    7920

ataaatccta actagcatag ccccttgggg cctctaaacg ggtcttgagg ggttttttgc    7980

tggctgagca tgagacggaa atctgctcgt cagtggtgct cacactgacg aatcatgtac    8040

agttcattcc gatg                                                      8054
```

<210> SEQ ID NO 94
<211> LENGTH 8054
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION. Chemically Synthesized. pBTK621

<400> SEQUENCE: 94

```
gaggcagcgt gttcaaacca attaacaaat tatgaataga aaaactaatc gagcatcaaa      60

tgaaactgca atttattcat atcaggatta tcaataccat atttttgaaa aagccgtttc     120

tgtaatgaag gagaaaactc accgaggcag ttccatagga tggcaagatc ctggtatcgg     180

tctgcgattc cgactcgtcc aacatcaata caacctatta atttcccctc gtcaaaaata     240

aggttatcaa gtgagaaatc accatgagtg acgactgaat ccggtgagaa tggcaaaagc     300

ttatgcattt ctttccagac ttgttcaaca ggccagccat tacgctcgtc atcaaaatca     360

ctcgcatcaa ccaaaccgtt attcattcgt gattgcgcct gagcgagacg aaatacgcga     420

tcgctgttaa aaggacaatt acaaacagga atcgaatgca accggcgcag gaacactgcc     480

agcgcatcaa caatattttc acctgaatca ggatattctt ctaatacctg gaatgctgtt     540

ttcccgggga tcgcagtggt gagtaaccat gcatcatcag gagtacggat aaaatgcttg     600

atggtcggaa gaggcataaa ttccgtcagc cagtttagtc tgaccatctc atctgtaaca     660

tcattggcaa cgctaccttt gccatgtttc agaaacaact ctggcgcatc gggcttccca     720

tacaatcgat agattgtcgc acctgattgc ccgacattat cgcgagccca tttataccca     780

tataaatcag catccatgtt ggaatttaat cgcggcctcg agcaagacgt ttcccgttga     840
```

-continued

```
atttggatca tatcaaccct tgtcccactg tttctgtatg aagccagttt tattgttact     900
gatgatatcc ttttatattg tgccatgtca aatcagagcc tttgagaccc aaagcggatt     960
tgttgaataa atcgaacttt tgctgagttg aaggatcagt catgaccaaa atcccttaac    1020
gtgagtcagc ctgccgcctt gggccggtgt atgtcgtact tgcccgccgc gaactcggtt    1080
accgtcccgc ccagcgcgac cagctccggc aacgcctcgc gcacccgctt gcggcgcttg    1140
cgcatggtcg aaccactggc ctctgacggc cagcatcagc cgcacaaggt atctatggaa    1200
gacctgacgg ttttgacggg gtcgatcccg cccaacagcc gatggtgacg cagggcgggc    1260
gttccgatgc accgcgaccg caccccgtcc atgccgatgc gatcatgatg gacgccaccc    1320
atgacgggcc gcgccatcaa ggggttcagg gccacgtaca gggcgccgtc cgccccgtcg    1380
ctggcgtacc ccgacagcag ccgaaacccc tgccgcttgc ggccattctg ggcgacgatg    1440
gatacctccc aaaggcgctc gatgcagtcc tgtatgtgct tgagcgcccc accaccatcg    1500
acccctgccc cgatttcctt tgccagcgcc cgatagctac ctttgaccac atggcattca    1560
gaggtgacgg aatccaactt gggttccagg aaccgcagga gatgccgtcc gacttcggtc    1620
ttggtttccg ggccaagccc taggccatta ggcccagccc tggccacccg aacttgcagg    1680
atgccgcagc catcagcgcc cagcggctcc gggccgctga actcgatccg cttgccgtcg    1740
ccgtagtcac acgtcacgtc cagcttgctg cgcttgcgcc cgcccgcctt gagggcacgg    1800
aacaggccgg gggccagaca gtgcgccggg tcgtgccgga cgtggctgag gctgtgcttg    1860
ttcttaggct tcacccaggg gcacccccct gctcttgcgc tgcctctccc gcacggcggg    1920
attgagaccc aagccgtccc gacgcctgca aacccgatca gagcacggtg agccatagtt    1980
ggccctgctc acccgcacgc ggccgcagca caggcgatgg tagtcgtccc aaccccattc    2040
ctcggcctcg gcgctggtca tgctcgacag gtaggactgc cagcggatgt tatcgaccag    2100
taccgagctg ccccggctgg cctgctgctg gtcgcctgcg cccatcatgg ccgcgccctt    2160
gctggcatgg tgcaggaaca cgatagagca cccggtatcg gcggcgatgg cctccatgcg    2220
accgatgacc tgggccatgg ggccgctggc gtttgcttcc tcgatgtgga accggcgcag    2280
agtgcaccgc aaccccaggc gggcgcactc gggcgcgagc ttgcggcagt agcaccactc    2340
cgggcccatg atgttgggca ggctgccgat cagcggctgg atcagcaggc cgtcagccac    2400
ggcttgccgc tcctcggcgc tgaggtgcgc cccaaggggc tgcaggcggt gatgaatggc    2460
ggtgggcggg tcttcggcgg gcaggtagat cacggggccg gtgggcagtt cgcccacctc    2520
cagcagatcc ggcccgcctg caatctgtgc ggccagttgc agggccagca tggattcacc    2580
ggcacaaccg gggaaaccca gagcccagcc agtacgggcc aacccgttgg gacaaaagtc    2640
gtccagaggc gggcgagctg atgcgaacgc atccagaata ttgccaggct tatgcgtagc    2700
aatcgattgc atcccttgca ggccgttggc ggttacgagc tggcgggtgc aatacccacg    2760
cccgcgcgcg ctctgagttc ttccaggcac tcggcagcgg cctcgtattc gtcgtcggtc    2820
agccagaccc tgcgctgacg aatccctttg gactccatgc gatcgcaatc tagcgcttgg    2880
cgtacagcgt cagggctggc cagcaggtcg ccggtctgct tgtccttttg gtcttccata    2940
tacgtaaccg agccaattgc aggggccgca aggcctgtcc tagcggaacc aggccaaggt    3000
gcagccgtca aggttaaggc tggccatatc agcgactgaa aagcggcccg cctcggcctt    3060
gtttgacgta taaccaaagc caccggccaa ccaatagccc ttgtcacttt tgatcaggta    3120
```

continued

```
gaccgaccct gatgcgattt tttcgtattc aataaaaacc aatttcgtgc gtgcgtaatc    3180
atcgtatacc aggcgtgcgt accaacgcaa gaccaaaatg atgaaatatg gaccgcaact    3240
gtccatgcct cgctgcgggg gtgccggtgc ccgtgccagc tcggcccggg caagctggac    3300
gctggccaga cccatgacct tgctgacggc gcgctcgatg taatccgctt cgtggccggg    3360
cttgcgctct gcccgcgctg ggctggcctc ggccatggcc ttgccgattt cctcggcact    3420
gcggccccgg ctggccagct tctgcgcggc gatcaagtcg cactcgctga ggtcatgacc    3480
gacgcgattg accagcaccg aactcatgcc gaggtactcg tccagcgaag tgcgccggtg    3540
gaggctacgc tgccgatcgg gaagttagcg gatggccagc atgccgccct tatccagatg    3600
ccgctgggcc tgctcgatct gctggccagc ctgctgcacc agcgccgggc cagcggtggc    3660
ggtcctgccc ttggattcac gcagcagcac ccacggctga taaccggcgc gggtggtgtg    3720
cttgtccttg cggttggtga agcccgccaa gcggccatcg tggcggctgt cggcgctggc    3780
cgggtcggcg tcgtactcgc tggccagcgt ccgggcaatc tgcccccgaa gttcaccgcc    3840
tgcggcgtcg gccaccttga ccaccgactg atcgtactcc gggccggttt caaccaaacg    3900
ggcaggatcc aggccatcgg atttcatgtc atccaggtca aactcgatgc ggtcgtaacc    3960
cagcaccaga ccatgccgct cctgctcggc gggcctgata tacacgtcat tgcccctggc    4020
attcatccgc ttgagccatg gcgtgctctg gagcacttcg gcggctgacc attcccggtt    4080
catcatctgg ccggtggtgg cgtccctgac gccgatatcg aagcgctcac agcccatggc    4140
cttgagctgt cggcctatgg cctgcaagct cctgtcgttc ttcatcgggc caccaagcgc    4200
agccagatcg agccgtactc ggttgtaagt gggcgcaggc agcgcaagcg caacgatgcg    4260
atcagcagca aaaccgtcgg aatcatggca gaccgcatca aggtcagacc tagccaacag    4320
tgccaccccc gcgacgcgct ccgggcgctc tgcgcggcgc tgctcacctc gggggctacc    4380
tcccgcaact ctttggccag ctccacccat gccgccccgg tctggcgctg ggctttcagc    4440
cactccgccg cctgcgcctc gctggcctgc ttggtctggc tcatgacctg ccgggcttcg    4500
tcggccagtg tcgccatgct ctgggccagc ggttcgatct gctccgctaa ctcgttgatg    4560
aatcaggatc tattcaatcc gtcgatcgcg ttacctggtc attgcctacc ggtaccaatg    4620
taagccgatg atccggcgct tggccgtgtc gatgctcagg gccacgtctg cccggccggt    4680
gcggatgccc cggccttcca tctccacccc gttcggcccc aggtgaacac cgggcaggcg    4740
ctcgatgccc tgcgcctcaa gtgttctgtg gtcaatgcgg gcgtcgtggc cagcccgctc    4800
tcatgcccgg ttggcatggt cggcccatgc ctcgcgggtc tgctcaagcc atgccttggg    4860
attgagagcc taggcattcc gtgcccagcc attctcaggg gtatcgcagt tgtaccgatt    4920
gatccaatga gaggcgggcc gatcgatgcc gtacctgatc cgctcggagc tactcaggtg    4980
gacgtgaggg ttatcgacgc caccggcatg gatggccagc gtatacggcc ggcgctaggc    5040
accggtcagg tgctggcgga actcggacgc cagcgccttc tgctggtcga gggtcagctc    5100
gaccggaatg gacaattcga aatcctgcaa cagccgaaca ttggcgcgtt aatccaggtc    5160
ggcagcatcc cagtagtcgg cgggccgctc gacgaactcc ggcatgtgcc cggattcggc    5220
gtgcaagccc tactcaatgc agcgggcata attgccttcg agctggatgt agtcggcatt    5280
ggccctggcc gattggccgc ccgacctgct gccggttttc gccgtaaggt gataaatcgc    5340
catgctgcct cgctgttgct tttgcttttc ggctccatgc aatggccctc ggagagcgca    5400
```

-continued

```
aagcccagac ggtggaagtt aggccagttt atcgaagaga aacccgtacg tgcgcccatc   5460
aatacaaagt aggggaggga ttgccgacgc tgtgcctaca tgataggata agagacagca   5520
cattaacaat ggggtgtcaa gatggttaag gggagcaaca aggcggcgga tcggctggcc   5580
aagctcgaag aacaacgagc gcgaatcaat gccgaaattc agcgggagcg ggcaagggaa   5640
cagcagcaag agcgcaagca cgaaacaagg cgcaaggtgc tggtgggggc catgattttg   5700
gccaaggtga acagcagcga gtggccggag gatcggctca tggcggcaat ggatgcgtac   5760
attgaaagcg aacacgaacg agcattgttc ggtcagaagc aacgcaagat ggatgagacg   5820
ggcagaatga tagacagaga aaggccatgc ggggcagaac aagcgcaacc aaacataggg   5880
tagggggaaa ggcggctaaa gcggctaaaa gcgctccagc gtatttctgc ggggtttggt   5940
gtgggtttta gcgggctttg cccgcctttc cccctgccgc gcagcggtgg ggcggtgtgt   6000
agccaagcgc agcgaatcga ccagcaatcc ggccactggc cgggcatatt gggcaagggc   6060
agcagcgccc cacaagggcg ctgataaccg cgcctagtgg atttatttta gataatcatg   6120
gatggatttt tacaaaaccc agcaagaccc agcccatgct gggtatgatg gtttgggggc   6180
gtgacagtta ttgcaggggt tagtgaatgt tatttgcagg gggcgtgacc gttatggatg   6240
gggttcgtga cagttagtac gggagtgacg gggactgggt gggaatgtct aggaacggga   6300
gggatttcgg ctgagggtaa aagaactttc cggtaagcga tagactgtat gtaaacacag   6360
tattgcaagg acgcggaaca tgcctcatgt ggcggccagg acggccagcc gggatcggga   6420
tacaggtcgt taccagagcc accgacccga gcaaaccctt ctctatcaga tcgttgacga   6480
gtattaacgg gactaagctg agctaatggc agagcaggga aaggaatgca agggctatgt   6540
gataacggga tttgaaggat ttctccaatg agggcagatg gagcaagggt ttctacgggt   6600
tcgctgcgag tcttgccacg ccgagcaagt ggtcgctttc agaaatcaat ctaaagtata   6660
tatgagtaaa cttggtctga caggccactg aattcgcatc tagatggtag agccacaaac   6720
agccggtaca agcaacgatc tccaggacca tctgaatcat gcgcggatga cacgaactca   6780
cgacggcgat cacagacatt aaccccacgt accgacagtg cgacaacgtg gcaattcgtc   6840
gatataacgt atactgatac tggccgatta ttgcagaaga aagcatgtgg gagttaattc   6900
ttgacacaga tatttccgga tgatataata actgagtact gttatacaga aacagaggag   6960
atattacata tgagtaaagg agaagaactt ttcactggag ttgtcccaat tcttgttgaa   7020
ttagatggtg atgttaatgg gcacaaattt tctgtcagtg gagagggtga aggtgatgca   7080
acatacggaa aacttaccct taaatttatt tgcactactg gaaaactacc tgttccatgg   7140
ccaacacttg tcactacttt ctcttatggt gttcaatgct tttcaagata cccagatcat   7200
atgaaacagc atgacttttt caagagtgcc atgcccgaag gttatgtaca ggaaagaact   7260
atatttttca aagatgacgg gaactacaag acacgtgctg aagtcaagtt tgaaggtgat   7320
acccttgtta atagaatcga gttaaagggt attgacttta aggaagatgg aaatatttta   7380
ggacacaaac tggaatataa ctataactcc cataatgtgt acattatggc agacaagcaa   7440
aagaacggta tcaaggttaa cttcaagatc agacacaaca ttgaggatgg aagcgttcaa   7500
ctagcagacc attatcaaca aaatactcca attggcgatg gccctgtcct tttaccagac   7560
aaccattacc tgtccactca atctgccctt tcgaaagatc ccaacgaaaa gcgcgaccac   7620
atggtccttc ttgagtttgt aacggctgct gggattacac acggcatgga tgaactatac   7680
```

-continued

```
aattaaatcc tatctagctt aaccccttgg ggccttctta aggggcttga ggggcttttt   7740
gatggctgcg aatgagacgg aattctgctc gtcggtggtg atctcactga agtttcatgt   7800
acagatcata ccgatgactg cctggcgact cacaactaag caagacagcc ggaaccagcg   7860
ccggcgaaca ccactgcata tatggcatat cacaacagtc caactagtgc actgcagtac   7920
accaggcatc aaataaaacg aaaggctcag tcgaaagact gggcctttcg ttttatctgt   7980
tgtttgtcgg tgaacgctct ctactagagt cacactggct caccttcggg tgggcctttc   8040
tgcgtttata aagc                                                     8054
```

What is claimed is:

**1**. A microbial composition comprising one or more bacteria genetically engineered to express at least one heterologous nucleic acid, wherein the one or more bacteria are native to the microbiome of a host insect.

**2**. The composition of claim **1**, wherein the host insect is selected from the group consisting of a honey bee and a bumble bee.

**3**. The composition of claim **1**, wherein the one or more bacteria are selected from the group consisting of *Snodgrassella alvi*, *Bartonella apis*, *Gilliamella apicola*, *Serratia marcescens*, *Parasaccharibacter apium*, and *Lactobacillus* sp.

**4**. The composition of claim **1**, wherein the composition comprises 2, 3, 4, or 5 bacterial species.

**5**. The composition of claim **1**, wherein the one or more bacteria express at least two heterologous nucleic acids.

**6**. The composition of claim 1, wherein the heterologous nucleic acid encodes a polypeptide that improves the health of a host insect.

**7**. The composition of claim **6**, wherein the heterologous nucleic acid encodes a pesticide degrading polypeptide or a cytochrome.

**8**. The composition of claim 1, wherein the heterologous nucleic acid is an inhibitory nucleic acid.

**9**. The composition of claim **8**, wherein the inhibitory nucleic acid is selected from the group consisting of an antisense DNA, dsRNA, siRNA, shRNA, sgRNA and a miRNA.

**10**. The composition of claim 1, wherein at least one heterologous nucleic acid is a broad host range plasmid.

**11**. The composition of claim **10**, wherein the broad host range plasmid comprises at least one regulatory sequence selected from the group consisting of an RSF1010 origin of replication, a PA1 promoter sequence, a PA2 promoter sequence, a PA3 promoter sequence, a cp25 promoter sequence, and a detectable marker.

**12**. The composition of claim 1, wherein the composition is a bee-ingestible composition.

**13**. The composition of claim **12**, wherein the composition comprises at least one selected from the group consisting of: the bacteria are present as a live suspension, the bacteria are present as a lyophilized powder, the composition is in solid form, the composition is in liquid form, the composition comprises protein, the composition comprises pollen, the composition is a sucrose solution, and the composition is a corn syrup solution.

**14**. The composition of claim **13**, further comprising a carbohydrate or sugar supplement.

**15**. An insect comprising the composition of claim 1.

**16**. The insect of claim **15**, wherein the insect is a honey bee.

**17**. A method for producing composition of claim 1, comprising transfecting said bacterial species with an expression cassette comprising at least one heterologous nucleic acid.

**18**. A method for downregulating expression of a target gene product, comprising administering an effective amount of a composition of claim **1** to an insect, wherein said bacteria express an inhibitor of said target gene product.

**19**. The method of claim **18**, wherein the target gene product is a gene from an organism selected from the group consisting of a pathogen, a parasite, a virus, a mite, Acute Bee Paralysis Virus (ABPV), Kashmir Bee Virus (KBV), Israeli Acute Paralysis Virus (IAPV), *Nosema ceranae*, Deformed Wing Virus, *Varroa destructor* mite, and Small Hive Beetle.

**20**. The method of claim **18**, wherein the insect is selected from the group consisting of a bee, a honey bee, a forager, a hive bee, a pupae, an adult bee, and a bee colony parasite.

**21**. The method of claim **18**, wherein the target gene is selected from the group consisting of TOM70, TIM22, TOM40, Imp2, mitochondrial Hsp70, ATM1-ABC transporter proteins, Frataxin, Ferredoxin, ERV1, ferredoxin, NADPH oxido-reductase [FNR], pyruvate dehydrogenase α subunit, pyruvate dehydrogenase β subunit, mitochondrial glycerol-3-phosphate dehydrogenase (mtG3PDH), manganese-containing superoxide dismutase (MnSOD), DNAJ (Hsp70 interacting), Iron Sulfur cluster ISU1, Cystein desulfurase Nfl1, NAR1, RLI1, ATPase subunit A, RNA polymerase I, RNA polymerase III, Inhibitor of apoptosis (IAP), and FAS apoptotic.

**22**. A method for reducing the susceptibility of a bee to a disease or disorder selected from the group consisting of Colony Collapse Disorder (CCD) and infection, comprising administering an effective amount of a composition of claim 1 to said bee, wherein said bacteria express an inhibitor of a pathogen or parasite specific gene product.

**23**. The method of claim **22**, wherein the pathogen or parasite is selected from the group consisting of Acute Bee Paralysis Virus (ABPV), Kashmir Bee Virus (KBV), Israeli Acute Paralysis Virus (IAPV), *Nosema ceranae*, *Nosema cerana*, *Nosema apis*, Deformed Wing Virus, and *Varroa destructor* mite.

**24**. The method of claim **22**, wherein the bacteria express at least two non-contiguous dsRNAs downregulating expression of a pathogen or parasite specific gene product.

**25**. A method for reducing the susceptibility of a bee colony to infestation by the Small Hive Beetle, comprising administering an effective amount of a composition of claim **1** to hive components, wherein said bacteria express an inhibitor of a Small Hive Beetle specific gene product.

**26**. The method of claim **25**, wherein the bacteria express at least two non-contiguous dsRNAs downregulating expression of a Small Hive Beetle specific gene product.

**27**. A method for expression of a heterologous nucleic acid sequence in a bee, the method comprising administering to the bee at least one modified *Snodgrassella alvi* bacterium comprising an expression plasmid for expression of the heterologous nucleic acid sequence.

**28**. The method of claim **27**, wherein the heterologous nucleic acid sequence encodes a molecule selected from the group consisting of a protein, a peptide, an antisense DNA, an inhibitory RNA, a dsRNA, a siRNA, a shRNA, a sgRNA and a miRNA.

* * * * *