



US 20250263679A1

(19) United States

(12) Patent Application Publication

NØRHOLM et al.

(10) Pub. No.: US 2025/0263679 A1

(43) Pub. Date: Aug. 21, 2025

(54) MUTANT RNASE E FOR ENHANCING RECOMBINANT PROTEIN EXPRESSION

(71) Applicant: DANMARKS TEKNISKE UNIVERSITET, Kongens Lyngby (DK)

(72) Inventors: Morten Helge Hauberg NØRHOLM, Kongens Lyngby (DK); Sophia Andrea Helga HEYDE, Kongens Lyngby (DK)

(21) Appl. No.: 18/563,026

(22) PCT Filed: Jun. 2, 2022

(86) PCT No.: PCT/EP2022/065066

§ 371 (c)(1),

(2) Date: Nov. 21, 2023

(30) Foreign Application Priority Data

Jun. 2, 2021 (EP) 21177466.6

Publication Classification

(51) Int. Cl.

C12N 9/22 (2006.01)

C07K 14/245 (2006.01)

C12N 9/12 (2006.01)

C12N 9/80 (2006.01)

C12N 15/70 (2006.01)

(52) U.S. Cl.

CPC C12N 9/22 (2013.01); C07K 14/245 (2013.01); C12N 9/1247 (2013.01); C12N 9/80 (2013.01); C12N 15/70 (2013.01); C12Y 207/07006 (2013.01); C12Y 301/26012 (2013.01); C12Y 305/01028 (2013.01); C07K 2319/60 (2013.01)

(57)

ABSTRACT

The invention provides a microbial host cell for enhanced recombinant expression of a target protein, said host cell comprising a mutant RNase E enzyme to be coexpressed with a target gene of interest. The invention further provides a method of enhancing recombinant protein expression using said microbial host cell. The method is particularly useful for the expression of proteins that are otherwise difficult to express in traditional expression systems, such as proteins which are toxic to the host cell. The invention further provides an auxiliary plasmid comprising a *rne** gene encoding a mutant RNase E enzyme and a LysS gene encoding T7 lysozyme.

Specification includes a Sequence Listing.

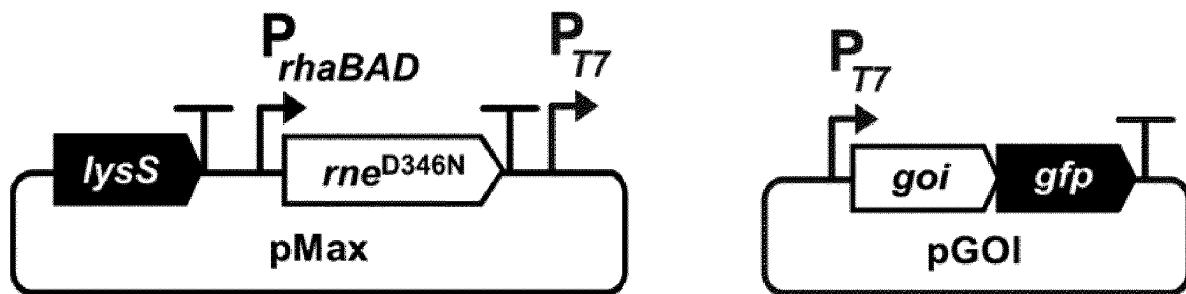


Figure 1A

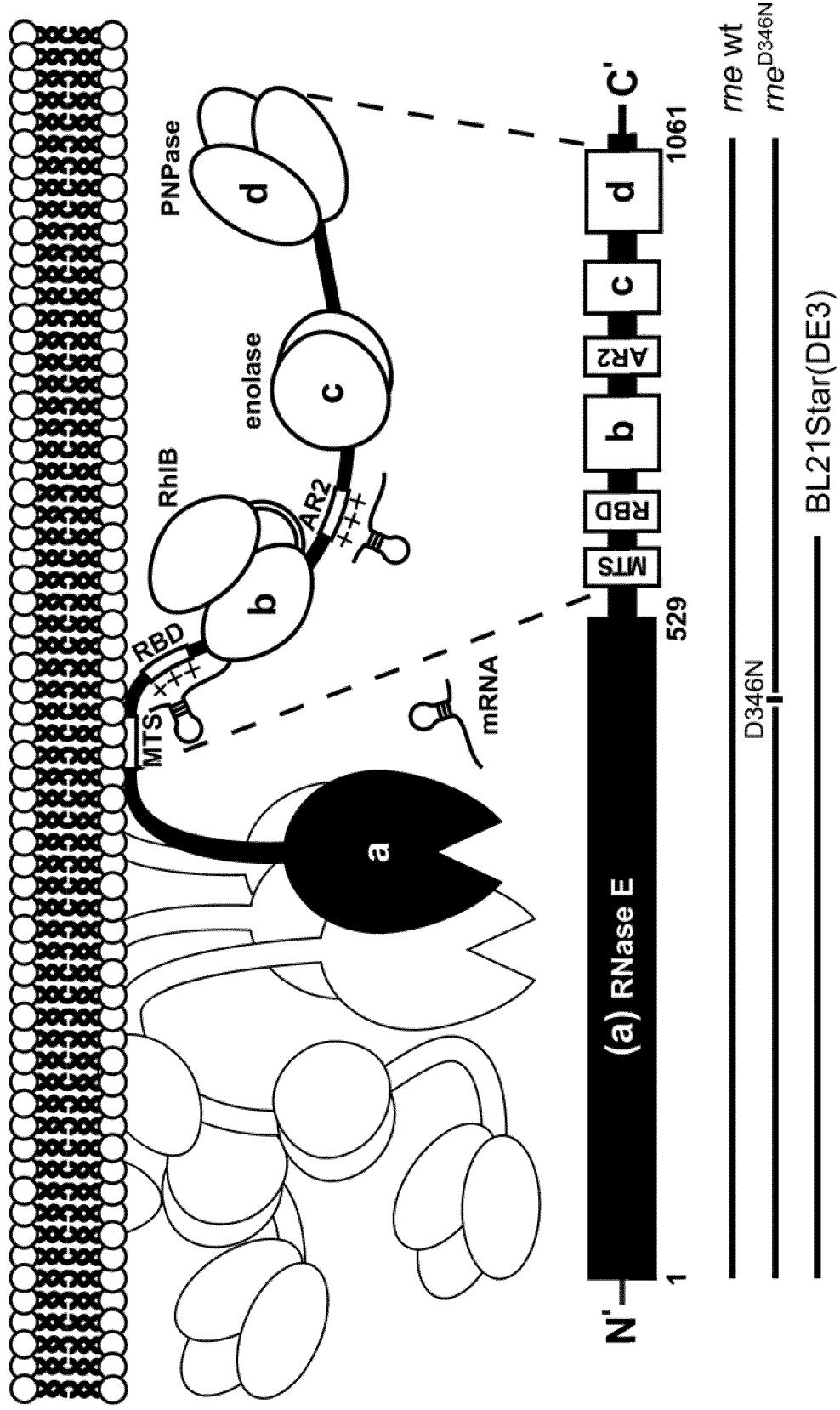


Figure 1B

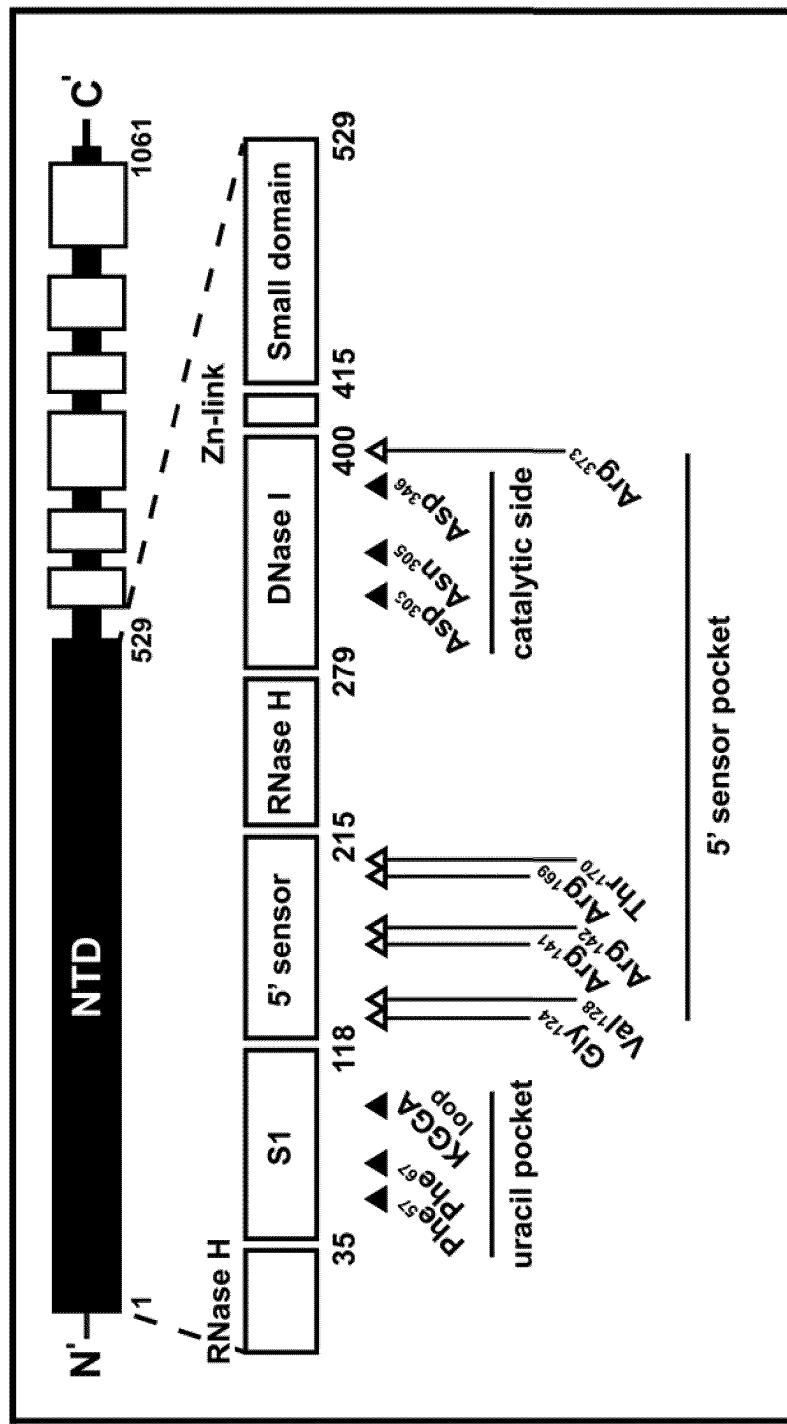


Figure 2A

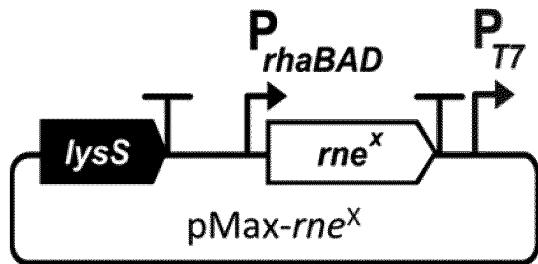


Figure 2B

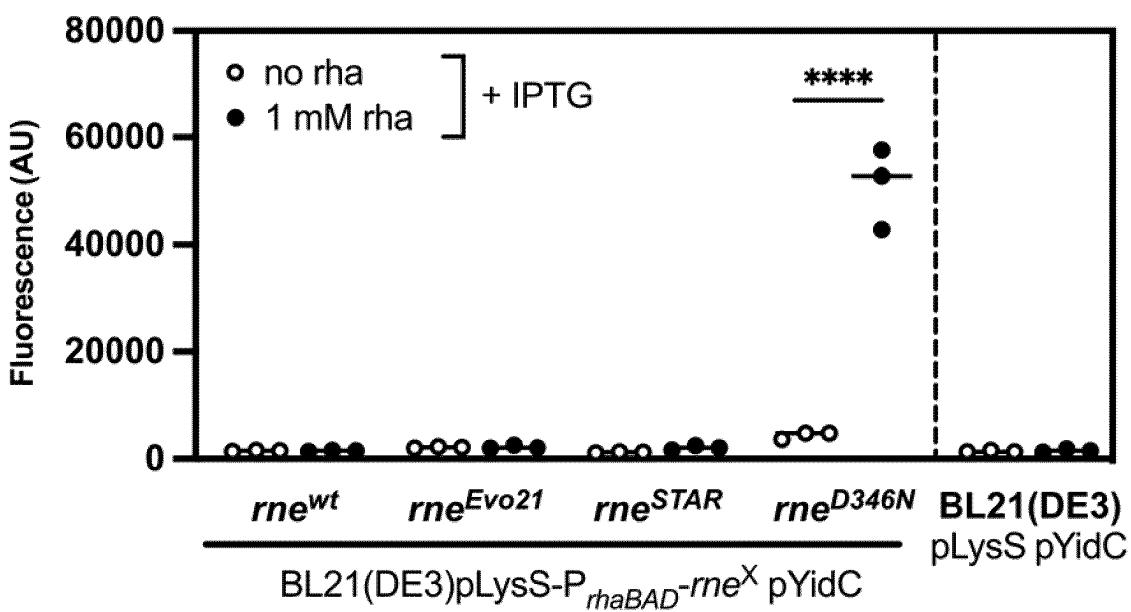


Figure 3A

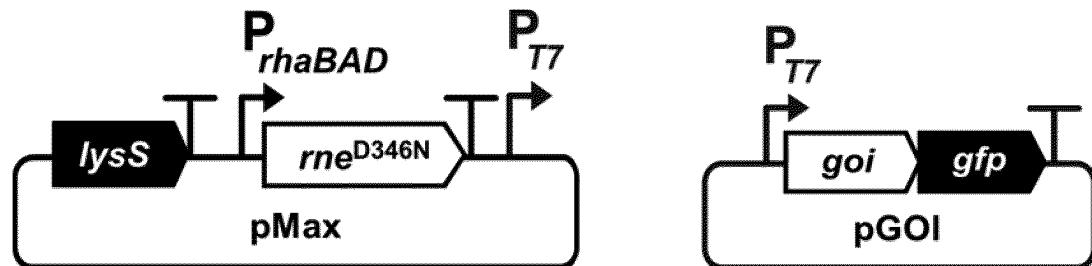


Figure 3B

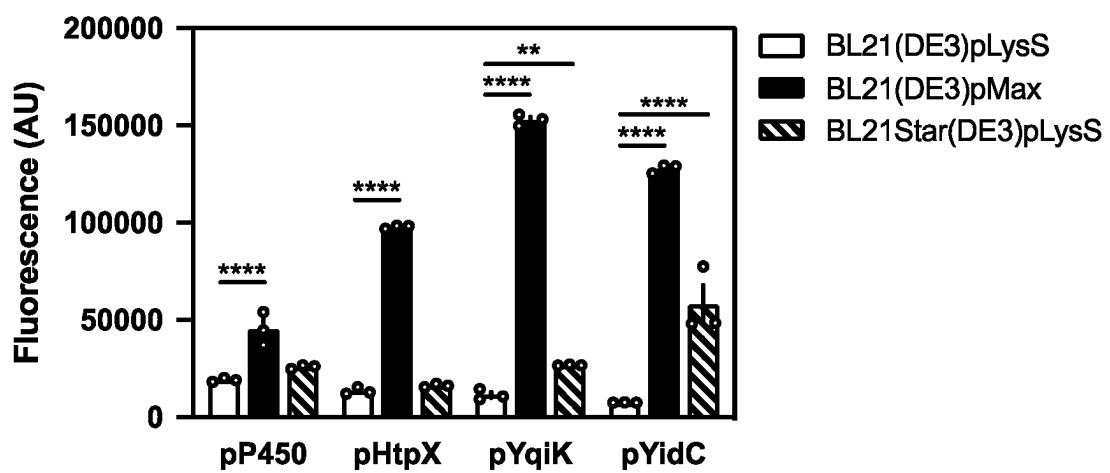


Figure 4

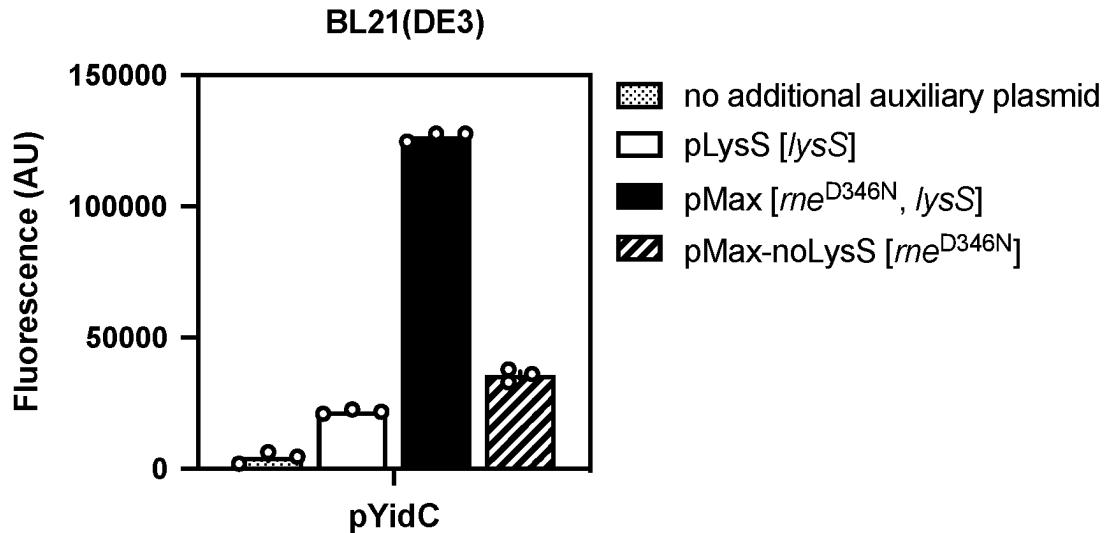


Figure 5

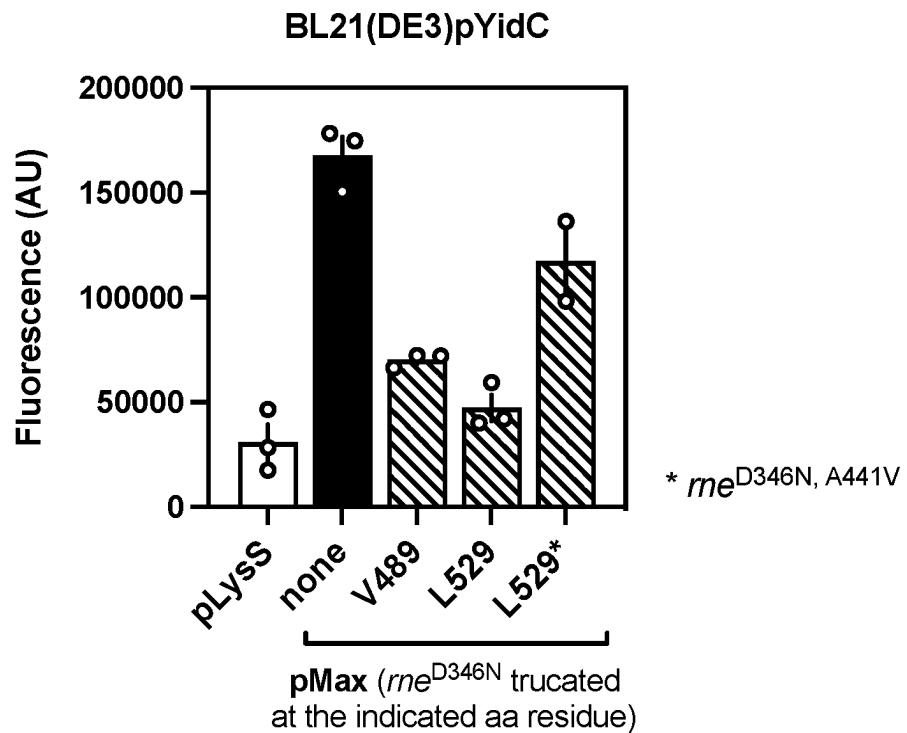


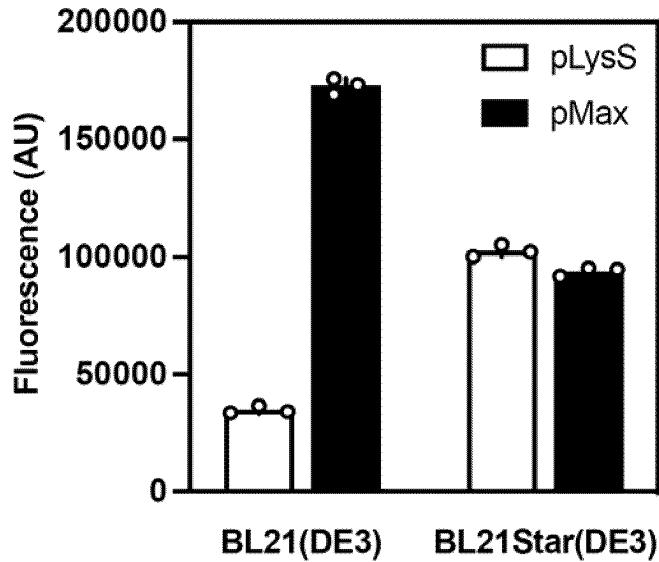
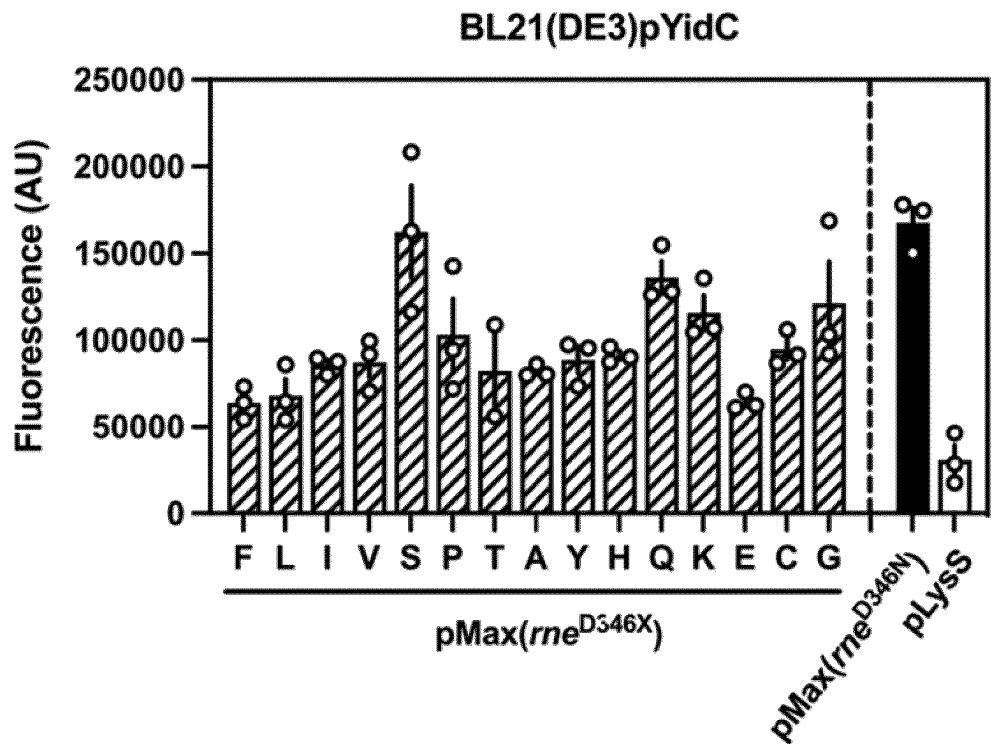
Figure 6**Figure 7**

Figure 8

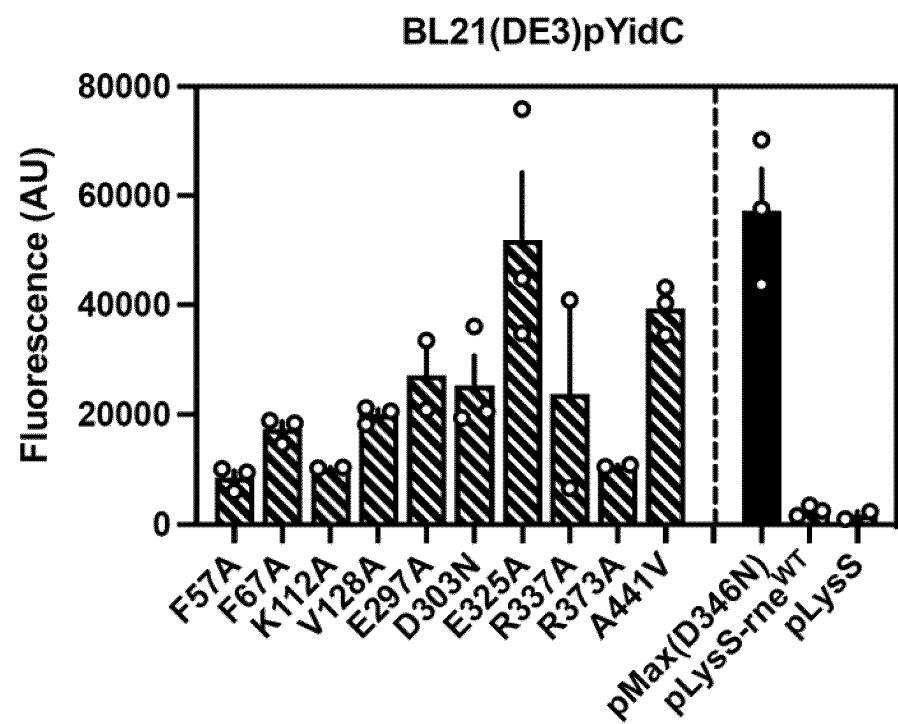


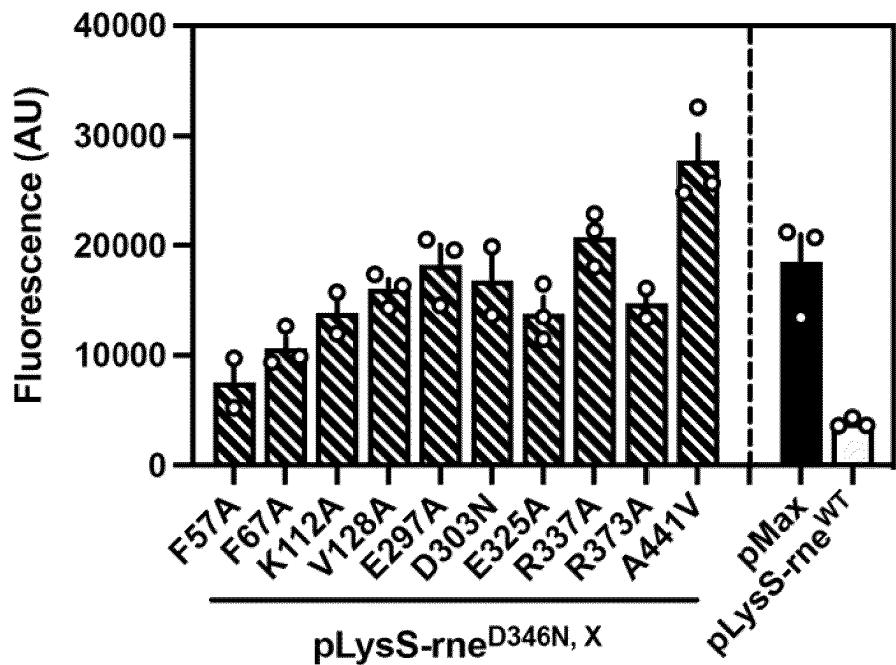
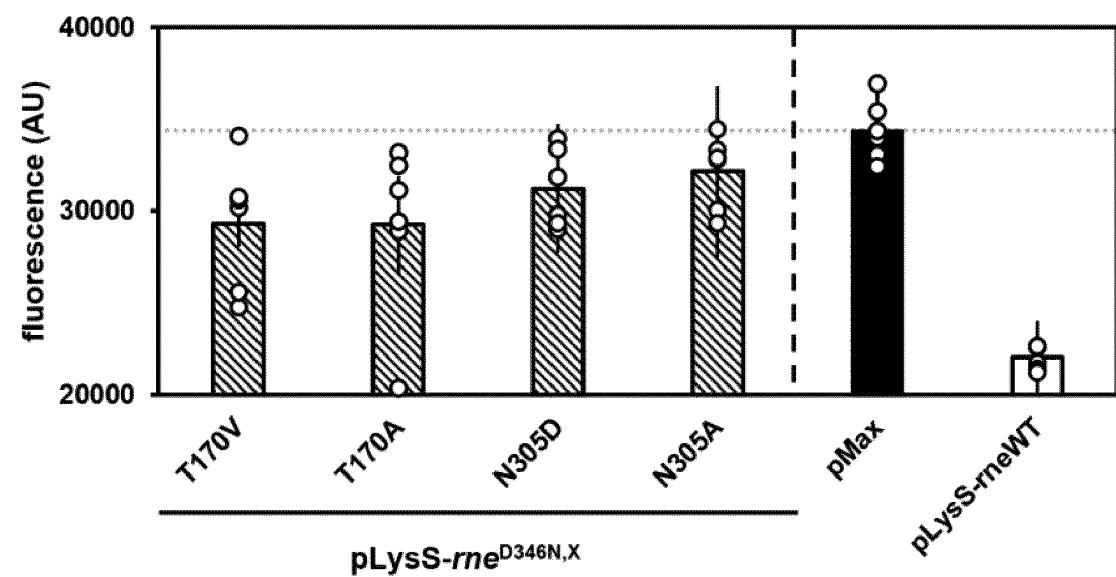
Figure 9**A****BL21(DE3)pYidC****B**

Figure 10

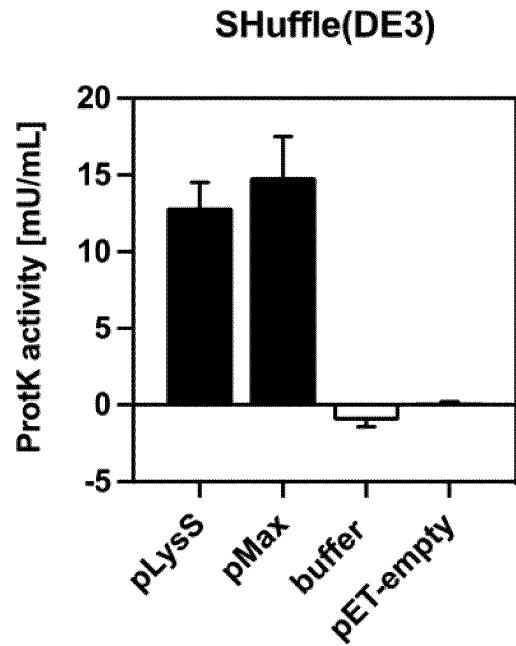
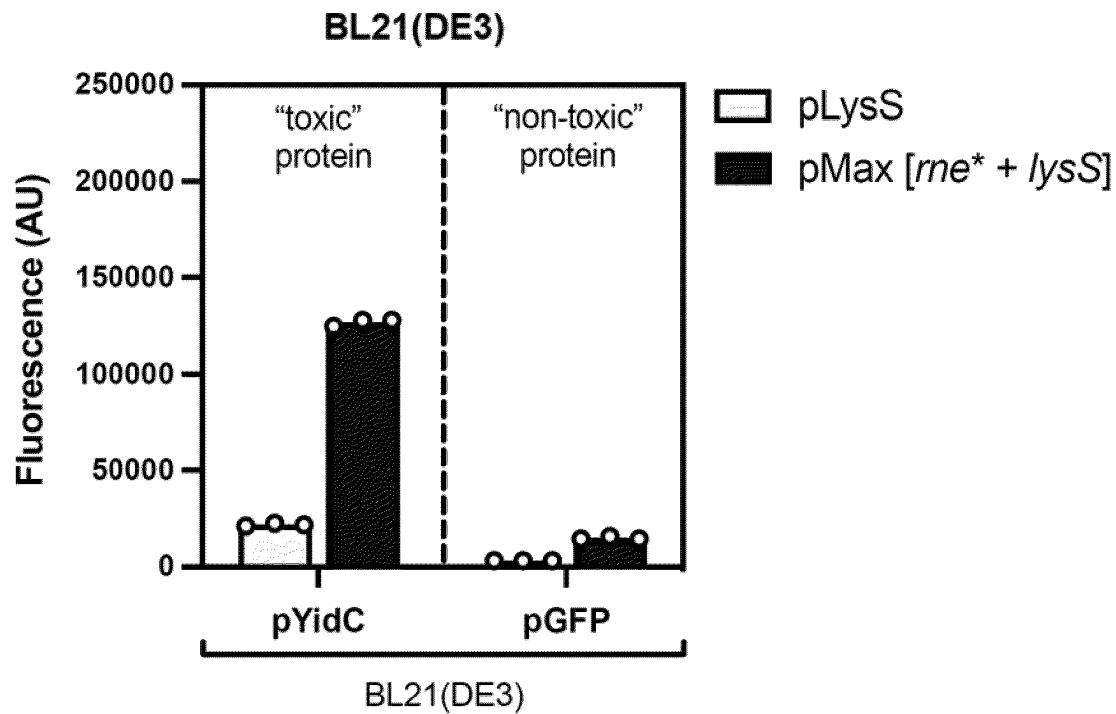


Figure 11



MUTANT RNASE E FOR ENHANCING RECOMBINANT PROTEIN EXPRESSION

FIELD OF THE INVENTION

[0001] The Invention provides a method of enhancing recombinant protein expression in microbial organisms by providing a mutant ribonuclease E (RNase E) gene, such as on an auxiliary plasmid, to be coexpressed with a target gene of interest. The method is particularly useful in the expression of proteins which are otherwise difficult to express in traditional expression systems.

BACKGROUND

[0002] Recombinant production of proteins enables biochemical and structural studies, and the pET vectors hosted in *Escherichia coli* strain BL21(DE3) are the most popular approach in research laboratories. Fast growth, high cell density, inexpensive culturing, availability of a wide variety of pET expression vectors combined with the detailed knowledge of *E. coli*'s genetics, physiology, and metabolism make it the preferred laboratory workhorse. However, the bacteria often show impaired growth and fitness loss when being used for protein production—a problem that is highly gene-specific, and clear guiding principles for gene and cell factory optimization are still lacking.

[0003] DE3 strains carry the T7 RNA polymerase (RNAP) gene under the control of the lacUV5 promoter, a stronger version of the native lacZ promoter inducible with IPTG. This allows the expression of any gene of interest under the control of a T7 promoter. Different strains have been isolated in which the toxicity of membrane protein production is reduced, leading to improved production yields: the “Walker strains” C41(DE3) and C43(DE3) (Miroux et al 1996), evolved in the late 1990s to tolerate over-production of membrane proteins, the more recently characterised derivative “mutant56” (Baumgarten et al 2017) evolved for higher production of the toxic membrane protein YidC, and the strains C44(DE3) and C45(DE3) (Angius et al 2018) evolved similarly. In all these cases, gene expression induced by IPTG inhibited colony formation on agar plates before mutations occurred and tolerance was achieved mainly due to reduced T7 RNAP activity, either via lacI mutations (Kwon et al 2015), or via promoter modifications, point mutations, or truncations in the T7 RNAP gene.

[0004] RNase E is an essential membrane-associated enzyme in *E. coli*. Involved in the maturation of both ribosomal RNA and tRNA, as well as total mRNA decay, and mediates the assembly of a multi-enzyme complex referred to as the “RNA degradosome” (FIG. 1A). It has previously been shown that only the N-terminal half of RNase E (amino acids 1-529 in *E. coli* RNase E), accommodating the active catalytic domain, is essential for cell growth, and the C-terminal non-catalytic region is mostly disordered and known to function as a scaffold mediating the association of the enzymes polynucleotide phosphorylase (PNPase), ATP-dependent RNA helicase (Rhlb) and endonuclease.

[0005] *E. coli* RNase E (EcRNase E) is the founding member of Type I RNase Es, found in betaproteobacteria, gammaproteobacteria and cyanobacteria, and has been extensively characterised. It is a large protein containing 1061 amino acids and can be divided into two domains. The N-terminal domain (NTD) is responsible for the endoribonuclease activity and the C-terminal domain (CTD) forms

the structural scaffold for an RNA-degrading multienzyme complex, the degradosome. The catalytic NTD consists of five subdomains (illustrated in FIG. 1B): an RNase H domain, an S1 domain, a 5' sensor, a deoxyribonuclease (DNase) I domain and a small domain (Mardle et al 2019). It is a homotetramer formed by interactions between the small domains. The catalytic site is located in the DNase I domain and harbours a hydrated magnesium ion, coordinated by two aspartic acids, Asp₃₀₃, positioned by asparagine Asn₃₀₅, and Asp₃₄₆, that is essential for the hydrolytic cleavage of the RNA substrate. EcRNase E cleaves single-stranded A/U-rich regions and has a strong preference for a 5' monophosphate. Specificity for uracil at the +2 position relative to the cleavage site is defined by the uracil pocket in the S1 domain. This pocket is comprised of Phe₆₇, positioned by Phe₅₇, and the Lys₁₁₂-Gly₁₁₃-Ala₁₁₄-Alanis₁₁₅ (KGAA) loop. In addition, recognition of a 5' monophosphorylated substrate requires the 5' sensor pocket, formed by amino acids Gly₁₂₄, Val₁₂₈, Arg₁₄₁, Arg₁₄₂, Arg₁₆₉, Thr₁₇₀ and Arg₃₇₃ with Val₁₂₈, Arg₁₆₉ and Thr₁₇₀ playing the critical roles in 5' monophosphate detection. The binding of the RNA substrate by the 5' sensor is predicted to induce a significant structural conformational change that helps to correctly position the RNA substrate for cleavage.

[0006] A truncation of the rne locus, rne131, resulting in an RNase E polypeptide lacking its non-catalytic region (while retaining amino acid residues 1-584) was isolated in a screen for suppressors of a temperature-sensitive allele of the mukB gene (Kido et al 1996). A later study showed that in strains such as BL21(DE3), introducing the rne131 truncation caused a bulk stabilisation of mRNA degradation, including mRNA produced by T7 RNAP (Lopez et al 1999). The rne131 truncation was engineered into the commercially available BL21Star (DE3) with the rationale that stabilising bulk mRNA would result in increased protein production. However, the commercial strain also comes with a note suggesting that it might be unsuitable for the over-expression of toxic genes.

[0007] Membrane proteins (MPs) are important drug targets and play essential roles in basic cellular mechanisms. In both prokaryotes and eukaryotes, 20-30% of all genes encode membrane proteins. MPs are involved in fundamental mechanisms such as transport of nutrients and signal molecules, response to environmental changes, membrane stability, maintenance of the redox potential, defence, and energy conversion. With natural abundances often too low to isolate sufficient material for *In vitro* studies, structural and biochemical investigations are limited by our ability to produce and purify MPs recombinantly in a functional state. MPs are also notoriously known for causing burden in expression systems and are hence considered toxic proteins.

SUMMARY OF THE INVENTION

[0008] In a first aspect, the invention provides a prokaryotic microbial host cell for recombinant expression of a target protein, said cell comprising

[0009] a. a gene encoding an enzyme having endoribonuclease activity (E.C. 3.1.26), said enzyme preferably having at least 75% amino acid sequence identity to SEQ ID NO. 2, wherein said gene is on the genome of said cell,

[0010] b. a first recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 75% sequence identity

with SEQ ID NO. 2, and wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substitutions which results in the mutant RNase E having decreased activity compared to the RNase E of SEQ ID NO. 2, and

[0011] c. a second recombinant gene encoding said target protein,

[0012] wherein expression of said target protein is enhanced compared to a cell lacking said first recombinant gene.

[0013] In a second aspect, the invention provides a prokaryotic vector comprising

[0014] a. a gene encoding a mutant RNase E (E.C. 3.1.26.12) having at least 75% amino acid sequence identity to SEQ ID NO. 2, wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substitutions which results in the mutant RNase E having decreased activity compared to the RNase E of SEQ ID NO. 2, and

[0015] b. a gene encoding a T7 lysozyme (E.C. 3.5.1.28)

[0016] In a third aspect, the invention provides the use of the prokaryotic vector according to the second aspect of the invention for enhancing expression of a recombinant gene encoding a target protein in a host prokaryotic microbial cell.

[0017] In a fourth aspect, the invention provides a method for the production of a target protein comprising culturing in a suitable culture medium, a host cell according to the first aspect of the invention, optionally inducing expression of said target protein, followed by isolation and purification of the expressed target protein by a well-known technique.

[0018] In a sixth aspect, the invention provides a method for producing a prokaryotic recombinant microbial cell having enhanced expression of a recombinant target protein,

[0019] a. providing a prokaryotic microbial cell comprising (i) a gene encoding an enzyme having endoribonuclease activity (E.C. 3.1.26), said enzyme preferably having at least 75% amino acid sequence identity to SEQ ID NO. 2, wherein said RNase gene is on the genome of said cell, and (II) a gene encoding said recombinant target protein,

[0020] b. transforming said microbial cell with a prokaryotic vector comprising a recombinant gene encoding a mutant RNase E having at least 75% amino acid sequence identity to SEQ ID NO. 2, and wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substitutions which results in the mutant RNase E having decreased activity compared to the RNase E of SEQ ID NO. 2,

[0021] wherein expression of said target protein is enhanced compared to a cell not transformed with said prokaryotic vector.

DESCRIPTION OF THE INVENTION

Definitions and Abbreviations

[0022] “Amino acid residue substitution” at a specific position means substitution with any amino acid different from the native amino acid residue that is present at that specific position.

[0023] Conservative vs nonconservative amino acid substitution: A conservative amino acid substitution replaces an amino acid with another amino acid that is similar in size

and chemical properties such that the substitution has no or only minor effect on protein structure and function; meanwhile a nonconservative amino acid substitution replaces an amino acid with another amino acid that is dissimilar and thereby is likely to affect structure and function of the protein.

[0024] rne* means a mutant rne gene comprising one or more mutations resulting in one or more amino acid substitutions in the catalytic N-terminal domain of RNase E enzyme and thereby modulates the endoribonuclease activity (E.C. 3.1.26) of the RNase E enzyme, compared to the non-mutated version of the enzyme.

[0025] Standard one or three letter abbreviations for amino acids are used herein.

DESCRIPTION OF FIGURES

[0026] FIG. 1A: Illustration of the *E. coli* RNA degradosome assembled around RNase E. N- and C-terminal domain of the membrane-bound essential endonuclease RNase E ([a] black) and the localisation of associated enzymes Rhlb [b], enolase [c], and PNPase [d] along the C-terminal non-catalytic scaffolding region are displayed. The mutation of the rne gene present on plasmid pMax (rne^{D346N}) as well as the truncation of the rne gene in BL21Star (DE3) are indicated.

[0027] FIG. 15: Important elements of the catalytically active N-terminal domain of *E. coli* RNase E. The location of the different elements along the N-terminal domain (NTD) of *E. coli* RNase E as well as important amino acid residues forming the uracil pocket, the 5' sensor pocket, and the catalytic side in the mature protein are indicated.

[0028] FIG. 2A: Plasmid illustration. Schematic illustration of plasmids pMax-me^X.

[0029] FIG. 25: Effect of rne mutations on protein production. Co-expression of me variants on auxiliary plasmid pMax-me^X (illustrated) along with yidC-gfp expression vector pYidC in BL21(DE3). RNase E production is under the control of a rhamnose promotor. YidC-GFP expression is induced via IPTG. The measured fluorescence is proportional to YidC-GFP expression under the different conditions.

[0030] FIG. 3A: Plasmid Illustrations. Schematic illustration of plasmids pMax and vectors used for co-expression of GFP fusion proteins: pYidC, p450, pHtpX and pYqiK. Gene names are abbreviated as gene of interest (GOI).

[0031] FIG. 3B: Effect of rne mutations on the expression of challenging/toxic proteins. Heterologous production of a variety of “challenging/toxic” GFP-fusion proteins in BL21Star (DE3) harbouring a pLysS plasmid compared to expression of the same genes in BL21(DE3) when co-expressing either the auxiliary plasmid pLysS or pMax. The measured fluorescence is proportional to expression of the GFP-fusion proteins.

[0032] FIG. 4: Effect of LysS and rne^{D346N} N co-expression on the expression of challenging/toxic proteins. Heterologous production of “challenging/toxic” GFP-fusion protein YidC-GFP in BL21(DE3) without co-expression of an auxiliary plasmid and when co-expressing either auxiliary plasmid pLysS (lysS), pMax (lysS, rne^{D346N}) or pMax not harbouring a lysS gene (pMax-noLysS: rne^{D345N}, no lysS). The measured fluorescence is proportional to YidC-GFP expression under the different conditions.

[0033] FIG. 5: Effect of a truncation of the rne^{D346N} gene present on pMax on protein co-expression. Co-expression of

pLysS, pMax (not truncated ("none")), and C-terminally truncated rne^{D346} variants on pMax (truncated at either position V489 or L529) and expression vector pYidC. The measured fluorescence is proportional to YidC-GFP expression under the different conditions.

[0034] FIG. 6: pMax co-expression in BL21Star (DE3). Co-expression of auxiliary plasmid pLysS or pMax with expression vector pYidC in BL21(DE3) and BL21Star (DE3). The measured fluorescence is proportional to YidC-GFP expression under the different conditions.

[0035] FIG. 7: Effect of different amino acid substitutions in rne position D346 on the performance of pMax. Co-expression of rne^{D346X} variants on pMax (specific amino acid residue changes are indicated in one-letter-code below each bar), pMax (harbouring rne^{D346N}) or pLysS along with expression vector pYidC. The measured fluorescence is proportional to YidC-GFP expression under the different conditions.

[0036] FIG. 8: Effect of different amino acid substitutions across rne on the performance of pMax. Co-expression of rne* variants on pMax (specific amino acid residue changes and positions are indicated in one-letter-code below each bar), pMax (harbouring rne^{D346N}), pLysS-rneWT or pLysS along with expression vector pYidC. The measured fluorescence is proportional to YidC-GFP expression under the different conditions.

[0037] FIG. 9: Effect of additional amino acid substitution across rne on the performance of pMax. (A) and (B) Co-expression of rne^{D346N, X} variants on pMax (additional amino acid residue change and position is indicated in one-letter-code below each bar), pMax (harbouring rne^{D345N}), pLysS-rneWT or pLysS along with expression vector pYidC. The measured fluorescence is proportional to YidC-GFP expression under the different conditions.

[0038] FIG. 10: Effect of pMax co-expression on Proteinase K production in Shuffle (DE3). Activity of Proteinase K expressed in the cytosol in SHuffle (DE3) co-expressing either plasmid pLysS or pMax. Negative control samples are pure buffer and expression from a pET-empty vector that carries no Proteinase K gene.

[0039] FIG. 11: Effect of lysS and rne^{346N} co-expression on the expression of challenging/toxic proteins (YidC) and non-toxic proteins (GFP). Heterologous production of GFP-fusion protein YidC-GFP and GFP in BL21(DE3) when co-expressing either auxiliary plasmid pLysS (lysS) or pMax (lysS, rneD346N). The measured fluorescence is proportional to YidC-GFP expression under the different conditions.

DETAILED DESCRIPTION OF THE INVENTION

[0040] The present invention concerns a method of enhancing recombinant protein expression in microbial organisms by providing a mutant RNase E enzyme on an auxiliary plasmid.

I. Microbial Host Cell for Recombinant Gene Expression

[0041] In one aspect, the present invention provides a microbial host cell for recombinant expression of a target protein.

[0042] The microbial host cell of the invention comprises

[0043] a. a gene on the genome of said cell encoding an enzyme having endoribonuclease activity (E.C. 3.1. 26),

[0044] b. a recombinant gene encoding a target protein, and

[0045] c. a recombinant gene (rne*) encoding a mutant RNase E enzyme, and

[0046] wherein expression of said target protein is enhanced compared to a cell lacking said recombinant gene (rne*) encoding said mutant RNase E enzyme.

[0047] RNase activity is generally considered an important activity of a microbial cell. In some microorganisms (e.g. *E. coli*), knocking out RNase genes leads to non-viable cells. RNase activity in a host cell affects recombinant protein expression. The present invention provides a means for regulating RNase activity by providing an RNase E mutant on an auxiliary plasmid. The activity of the RNase E mutant is modulated through one or more point mutations in the N-terminal catalytic region of the RNase E enzyme. Without wishing to be bound by theory, it is believed that the mutant RNase E enzyme thereby acts as a competitive inhibitor of the native RNases, as the mutant RNase E is still able to engage with RNA, leading to less degradation of RNA due to its decreased activity. As disclosed in the background section, the active RNase E degradosome is a tetramer of the individually expressed RNase E enzymes; the homotetramer primarily being formed by interactions between the small domains of the RNase Es. In the microbial host cell of the invention, the homotetramer may comprise 1, 2, 3, or 4 mutant RNase E enzymes (the remaining units being native RNase E enzyme). Without wishing to be bound by theory, further to the above, it is believed that the one or more mutant RNase E enzyme in the homotetramer degradosome are dominant over the native RNase E enzyme and thereby decrease the overall activity of the degradosome complex. Expression levels of the mutant RNase E can be regulated by selecting suitable promoters and RBS sequences, and can thereby be tuned to outcompete native RNases.

I.i Host Cell

[0048] In one embodiment, the host cell is a prokaryotic microbial host cell, such as *Escherichia coli*, *Pseudomonas putida*, *Bacillus subtilis* and *Bacillus licheniformis*. In one embodiment, the prokaryotic microbial host cell is selected from gram-negative bacteria, such as *Escherichia coli* and *Pseudomonas putida*. In one embodiment, the prokaryotic host cell is selected from gram-positive bacteria, such as *Bacillus subtilis*. In a preferred embodiment, the prokaryotic microbial host cell is selected from *E. coli* and *B. subtilis*. In a further preferred embodiment, the host cell is selected from *E. coli* strain BL21, BL21(DE3), and BL21Star (DE3), or K-12 MG1655; preferably strain BL21(DE3).

[0049] In another embodiment, the host cell is a eukaryotic microbial host cell. In one embodiment, the eukaryotic microbial host cell is selected from *Saccharomyces cerevisiae*, *Aspergillus niger*, and *Pichia pastoris*.

[0050] In yet another embodiment, the host cell is a mammalian cell, in one embodiment, the mammalian cell is selected from CHO and HEK cell lines.

I.ii Genomic Gene Encoding Endoribonuclease Activity

[0051] In one embodiment, the microbial host cell of the invention comprises a gene on its genome encoding an enzyme having endoribonuclease activity (E.C. 3.1.26).

[0052] In one embodiment, the prokaryotic microbial host cell of the invention comprises a gene encoding an enzyme having endoribonuclease activity (E.C. 3.1.26) on its genome. In one embodiment, said enzyme having endoribonuclease activity (E.C. 3.1.26) has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity to SEQ ID NO: 2, 4, 6, or 8.

[0053] In one embodiment, the prokaryotic microbial host cell of the invention comprises a gene encoding an RNase E enzyme (E.C. 3.1.26.12) on its genome. In one embodiment, the RNase E enzyme encoded by a gene on the genome of the host cell has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity to SEQ ID NO: 2 or 4.

[0054] In general, the sequence of the catalytic region of RNase E is highly conserved among Gram-negative bacteria, but some properties of the catalytic domain may be species-specific. For gram-negative bacteria, in one embodiment, the RNase E enzyme encoded by a gene on the genome of the host cell has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% amino acid sequence identity to a RNase E native to said prokaryotic host cell.

[0055] For example, where the prokaryotic host cell is *E. coli*, the host cell comprises a gene on its genome encoding a RNase E (E.C. 3.1.26.12) having at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity to SEQ ID NO: 2. For example, where the prokaryotic host cell is *Pseudomonas putida*, the host cell comprises a gene on its genome encoding a RNase E (E.C. 3.1.26.12) having at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity to SEQ ID NO: 4.

[0056] Preferably, the RNase E gene on the genome of the microbial host cell is full length, compared to the native RNase E of said host cell, i.e. not truncated, and thereby comprises both its N-terminal catalytic domain and its C-terminal non-catalytic domain. For illustrations of *E. coli* rne domains, see FIG. 1.

[0057] Gram-positive bacteria do not possess any native homologs of RNase E. For example in *B. subtilis*, RNA degradation is done via RNases J1/J2 and Y which form a complex with additional enzymes similar to the degradosome in *E. coli* mediated via RNase E. Whereas RNase J2 seems to possess a similar function and structure as RNase E but no homology, *B. subtilis* RNase Y can be functionally replaced by *E. coli* RNase E. Full-length RNase E almost completely restores wild type growth of an rne null mutant. RNase E (*E. coli*) and RNase Y (*B. subtilis*) require a Mg²⁺ ion to function and are involved in the initiation of mRNA decay. Although the amino acid sequence of RNase Y shows a low identity to that of RNase E, they share the same function as endo-ribonucleases with relaxed sequence specificity. Additionally, a degradosome-like complex centred around RNase Y has been proposed.

[0058] For gram-positive bacteria, in one embodiment, the genome of the host cell comprises a gene encoding an

enzyme having endo-ribonuclease activity (E.C. 3.1.26). In one embodiment, such enzyme has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% amino acid sequence identity to a RNase Y enzyme native to said gram-negative host cell. For example, where the prokaryotic host cell is *Bacillus subtilis*, the host cell comprises a gene on its genome encoding a RNase Y having at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity to SEQ ID NO: 6. For example, where the prokaryotic host cell is *Bacillus licheniformis*, the host cell comprises a gene on its genome encoding a RNase Y having at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity to SEQ ID NO: 8. Fungal RNases such as RNases of *Saccharomyces*, *Pichia*, and *Aspergilli* species mostly belong to the RNase T1 family, that do not share sequence identity as such with the *E. coli* RNase E. Major RNases important in *S. cerevisiae*, *Pichia*, and *Aspergilli* include major cytoplasmic deadenylase CCR4 (gene expression regulation and poly-A shortening/mRNA decay), Pan2/Pan3 complex (mRNA deadenylase), and especially Rpb4/Rpb7 (cytosolic mRNA decay).

[0059] In one embodiment, the eukaryotic microbial host cell of the invention comprises one or more gene(s) encoding RNase(s) belonging to the RNase T1 family on its genome. In one embodiment, the one or more RNases encoded on the genome of the eukaryotic host cell has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% amino acid sequence identity to a RNase native to said eukaryotic host cell.

[0060] For example, where the eukaryotic host cell is *Saccharomyces cerevisiae*, the host cell comprises a gene encoding major cytoplasmic deadenylase CCR4, Pan2/Pan3 complex and/or Rpb4/Rpb7 on its genome on its genome. In one embodiment, said CCR4 has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity to SEQ ID NO: 10, said Pan2 has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity to SEQ ID NO: 12, said Pan3 has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity to SEQ ID NO: 14, said Rpb4 has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity to SEQ ID NO: 16, and said Rpb7 has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity to SEQ ID NO: 18.

[0061] For example, where the eukaryotic host cell is *Pichia pastoris*, the host cell comprises a gene encoding major cytoplasmic deadenylase CCR4, Pan2/Pan3 complex and/or Rpb4/Rpb7 on its genome on its genome. In one embodiment, said CCR4 has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity to SEQ ID NO: 19, said Pan2 has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity to SEQ ID NO: 20, said Pan3 has at least 70, 71, 73, 74, 75, 76, 77, 78,

79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity to SEQ ID NO: 21, said Rpb4 has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity to SEQ ID NO: 22, and said Rpb7 has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity to SEQ ID NO: 23, [0062] For example, where the eukaryotic host cell is *Aspergillus niger*, the host cell comprises a gene encoding major cytoplasmic deadenylase CCR4, Pan2/Pan3 complex and/or Rpb4/Rpb7 on its genome on its genome. In one embodiment, said CCR4 has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity to SEQ ID NO: 24, said Pan2 has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity to SEQ ID NO: 25, said Pan3 has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity to SEQ ID NO: 26, said Rpb4 has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity to SEQ ID NO: 27, and said Rpb7 has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity to SEQ ID NO: 28,

I.iii Recombinant Gene Encoding Target Protein

[0063] The microbial cell of the invention comprises one or more recombinant gene(s) encoding one or more target protein(s). The term protein may refer to any peptide, polypeptide or protein.

[0064] The recombinant gene may be provided on a plasmid or incorporated into the genome of the microbial host cell. Such plasmid comprising the recombinant gene encoding the target protein shall further comprise commonly known functionalities for maintenance of the plasmid in the cell, such as an origin of replication—as recognized by a person skilled in the art. The number of copies of the recombinant gene in the microbial cell may be regulated by placing the gene on a plasmid with a high or low copy number.

[0065] The recombinant gene encoding the target protein may be inducible or constitutively expressed. In a preferred embodiment, expression of the recombinant gene is regulated by an inducible promoter. Any inducible promoter may be used. The induction of protein expression in bacteria is well known in the art. In one embodiment of the present invention, the induction of protein expression is for example made by the addition of isopropyl- β -D-thiogalactopyranoside (IPTG), as specified below.

[0066] In one embodiment, the microbial host cell comprises a recombinant gene encoding a target protein and a gene encoding a T7 RNA polymerase (E.C. 2.7.7.6), wherein expression of said recombinant gene is regulated by an inducible T7 promoter.

[0067] In expression of the recombinant gene encoding the target protein regulated by such T7 system, the T7 RNA polymerase (RNAP) may be under the control of an inducible promoter, such as a lac promoter inducible by IPTG. Further, the T7 promoter regulating the expression of the recombinant gene may be regulated by the function of a lac

operon (hence inducible by IPTG), for example by means of a lac repressor and lac operator.

[0068] In a further embodiment, the microbial host cell further comprises a gene encoding a T7 lysozyme (E.C. 3.5.1.28). T7 lysozyme is a natural inhibitor of the T7 RNA polymerase. In one preferred embodiment, the microbial host cell comprises a gene encoding a T7 RNA polymerase (E.C. 2.7.7.6), a gene encoding a T7 lysozyme (E.C. 3.5.1.28), and a recombinant gene encoding a target protein, wherein expression of said recombinant gene is regulated by an inducible T7 promoter.

[0069] The T7 lysozyme is especially useful in systems, where inducible expression of the T7 RNA polymerase is “leaky”. The “DE3” strains of *E. coli* comprise such leaky T7 RNA polymerase expression, as it is regulated by the IPTG inducible lacUV5 promoter, but which is known to be leaky and thereby allow for some basal expression of the T7 RNA polymerase.

[0070] In one embodiment, the microbial host cell is an *E. coli* strain comprising the T7 system, such as the “DE3” strains, where the recombinant gene encoding the target protein is provided on a pET vector, such as pET28a+ (SEQ ID NO. 46), preferably in combination with expression of a gene encoding T7 lysozyme (E.C. 3.5.1.28), such as LysS (SEQ ID NO. 36) provided on a second plasmid—e.g. pMax (SEQ ID NO. 38) of the present invention, which comprises pLysS SEQ ID No. 29) as the backbone.

[0071] In yet another embodiment, expression of the recombinant gene encoding the target protein is regulated by a promoter which is recognized by the host's native polymerase, this promoter is preferably an inducible promoter.

[0072] In one embodiment, the promoter for regulating the expression of the recombinant gene encoding the target protein is selected from rhaBAD promoter (SEQ ID NO 41), araBAD promoter (SEQ ID NO 88), Ptrc promotor (SEQ ID NO 89), Ptet promoter (SEQ ID NO 90), Ptac promoter (SEQ ID NO 91), and PL promoter (SEQ ID NO 92).

[0073] In one embodiment, said promoter is native to the host cell.

I.iv Recombinant Gene (Rne*) Encoding a Mutant RNase E

[0074] The microbial host cell of the invention comprises a recombinant gene rne* encoding a mutant RNase E enzyme.

[0075] In one embodiment, the microbial cell comprises a mutant RNase E enzyme having one or more amino acid residue substitutions, which facilitates improved expression of a target protein, preferably a toxic protein such as YidC (SEQ ID NO. 68), compared to expression of said target protein in a parent cell (from which the microbial cell was derived) lacking expression of the mutant RNase E.

[0076] In one preferred embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 2, wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substitutions, preferably a dominant substitution, wherein said one or more amino acid residue substitutions facilitates improved expression of a target protein, preferably a toxic protein such as YidC (SEQ ID NO. 68), compared to expression of said target protein in

a parent cell (from which the microbial cell was derived) lacking expression of the mutant RNase E.

[0077] A person skilled in the art will recognize that such amino acid residue substitution which facilitates improved expression of a target protein may be identified by expressing a candidate mutant RNase E comprising a candidate amino acid substitution in the microbial cell together with the target protein, preferably a toxic protein such as YidC (SEQ ID NO. 68), and confirming improved expression of YidC compared to a parent cell (from which the microbial cell was derived) lacking expression of the candidate mutant RNase E.

[0078] In one embodiment, the microbial host cell of the invention comprises a vector comprising a recombinant gene *rne** encoding a mutant RNase E enzyme. The vector may be a prokaryotic or eukaryotic plasmid, comprising an origin of replication suitable for replication in the respective prokaryotic or eukaryotic host cell, independently of the chromosome.

[0079] In one embodiment, the microbial host cell of the invention comprises a recombinant gene *rne** encoding a mutant RNase E enzyme on its genome.

[0080] As mentioned previously, an essential function of RNase E is mRNA degradation; the N-terminal domain (NTD) of RNase E is responsible for this endoribonuclease activity. The *rne** gene comprises one or more point mutations resulting in one or more amino acid substitutions in the catalytic N-terminal domain of the RNase E enzyme, thereby modulating the activity of the RNase E enzyme, compared to the non-mutated version of the enzyme.

[0081] As mentioned previously, the catalytic N-terminal domain of RNase E is highly conserved. In one embodiment, the recombinant *rne** gene encoding mutant RNase E comprises an amino acid substitution in the DNase I domain. In one embodiment, the amino acid substitution affects metal ion chelation in the RNase E enzyme.

[0082] Table 1 provides a list of residues of *E. coli* and *P. putida* RNase E relevant for the present invention and describes the function in *E. coli* RNase E when mutated. In a preferred embodiment, the mutant RNase E encoded by *rne** comprises one or more amino acid substitution(s) at one or more positions relative to the positions listed in table 1. A person skilled in the art would know how to perform sequence alignment of homologous RNase E sequences from different organisms to identify these specific positions with reference to the information provided in table 1.

TABLE 1

Relevant residues of <i>E. coli</i> RNase E (SEQ ID NO. 2) and <i>P. putida</i> (SEQ ID NO. 4)				
Position in <i>E. coli</i>	Position in <i>P. putida</i>	Location (FIG. 1B)	Predicted function of mutant forms of <i>E. coli</i> RNase E	Ref.
D346	D343	DNase I	Metal ion chelation - activity decrease	1
E297	E294	DNase I	Metal ion chelation (similar to D346N) - active	1
D303	D300	DNase I	Metal ion chelation (same as for D346N); activity decrease, nearly inactive	1
N305	N302	DNase I	Activity decrease (supports D303 through hydrogen bonding)	1, 2
E325	E322	DNase I	Metal ion chelation (similar to D346N) - active	1
R337	R334	DNase I	Metal ion chelation (similar to D346N); highly conserved	1
D349	D346	DNase I	Metal ion chelation (similar to D346N) - active	1
V128	V126	5' sensor	Removes enhancement of cleavage seen for substrate with 5' monophosphate; highly conserved; inactive	2
R169	R167	5' sensor	Removes enhancement of cleavage seen for substrate with 5' monophosphate	1, 2
T170	T168	5' sensor	Inactive	2
F57	F57	S1	50-fold activity decrease	2
F67	F67	S1	50-fold activity decrease	2
K112	K110	S1	50-fold activity decrease	2
G124	G122	5' sensor	5' sensor/RNA binding, possibly decreased function or inactive	4
R141	R139	5' sensor	5' sensor/RNA binding, possibly decreased function or inactive	4
R142	R140	5' sensor	5' sensor/RNA binding, possibly decreased function or inactive	4
R373	R370	DNase I	5' sensor/RNA binding, possibly decreased function or inactive	4
A441	A438	Small domain (C-terminal)	unknown	This study

Ref:

1 Garey et al 2009,

2 Callaghan et al 2005,

3 Kim et al 2014,

4 Mardle et al 2019.

[0083] Mutations at residues D346, E297, D303, N305, E325, R337, D349, V128, R169, T170, F57, F67, K112, G124, R141, R142, or R373 are based on prior art predicted to decrease the activity of RNase E.

[0084] Residues F57, F67, K112, D303, and D346 in *E. coli* RNase E are active site residues (Garey et al 2009).

[0085] Residues D346, E297, D303, N305, E325, R337, and D349 are important for metal ion chelation. Specifically residues D346, E297, D303, E325, R337, and D349 within the DNase I domain are important for metal ion chelation (Garey et al 2009), while residue N305 supports D303 through hydrogen bonding (Callaghan et al 2005).

[0086] Residues G124, V128, R141, R142, R169, T170, R373, F57, F67, K112 are important for RNA recognition (5' sensor domain and other domains). Several residues in the S1 domain could contribute to RNA binding, but only three residues: F57, F67, and K112 provide obvious contacts to the substrate (Garey et al 2009). Therefore, specifically residues F57, F67, K112 are important, as they are in contact with RNA (S1 domain). Residues E297, D303, N305, E325, R337, D346, D349, R373 are important DNase I-like domain residues.

[0087] In one embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 2, and wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substitutions which results in the mutant RNase E having decreased activity compared to the RNase E of SEQ ID NO. 2.

[0088] In one embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 2, wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substitutions which results in the mutant RNase E having decreased activity compared to the RNase E of SEQ ID NO. 2, and wherein said one or more amino acid residue substitutions facilitates improved expression of a target protein, preferably a toxic protein such as YidC (SEQ ID NO. 68), compared to expression of said target protein in a parent cell (from which the microbial cell was derived) lacking expression of the mutant RNase E.

[0089] In one embodiment, the activity of the mutant RNase E is decreased by at least 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, or 100% compared to the RNase E of SEQ ID NO. 2 when expressed in a host cell of the invention. In one embodiment, the activity of the mutant RNase E is decreased at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 fold, such as even 20, 40, 60, 80, 100 fold, such as even 200, 300, 400, 500, 600, 700, 800, 900 or 1000 fold compared to the RNase E of SEQ ID NO. 2 when expressed in a host cell of the invention.

[0090] In one embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100%

sequence identity with SEQ ID NO. 2, and wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substitutions at one or more positions selected from D346, E297, D303, N305, E325, R337, D349, V128, R169, T170, F57, F67, K112, G124, R141, R142, R373, and A441 relative to SEQ ID NO. 2

[0091] In one embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 2, and wherein the amino acid sequence of said mutant RNase E has an amino acid residue substitution which reduces the metal ion chelation ability of the enzyme. In a preferred embodiment, said metal ion chelation ability is reduced by substituting an amino acid at a position selected from D346, E297, D303, N305, E325, R337, and D349 relative to SEQ ID NO. 2.

[0092] In a preferred embodiment, the amino acid residue substitutions mentioned above are non-conservative substitutions.

[0093] In one preferred embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 2, and wherein the amino acid sequence of said mutant RNase E has an amino acid residue substitution which results in the mutant RNase E having reduced metal ion chelation ability compared to the RNase E of SEQ ID NO. 2. In a preferred embodiment, said metal ion chelation ability is reduced by substituting an amino acid at a position selected from D346, E297, D303, E325, R337, and D349 relative to SEQ ID NO. 2.

[0094] In one embodiment, the metal ion chelation ability of the mutant RNase E is decreased by at least 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, or 100% compared to the RNase E of SEQ ID NO. 2 when expressed in a host cell of the invention. In one embodiment, the metal ion chelation ability of the mutant RNase E is decreased at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 fold, such as even 20, 40, 60, 80, 100 fold, such as even 200, 300, 400, 500, 600, 700, 800, 900 or 1000 fold compared to the RNase E of SEQ ID NO. 2 when expressed in a host cell of the invention.

[0095] In a preferred embodiment, the microbial cell comprising a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 2, and wherein the amino acid sequence of said mutant RNase E has a single amino acid residue substitution at position D346 relative to SEQ ID NO. 2, wherein said amino acid is substituted to any amino acid other than aspartate. The essential aspartate residue in position 346 in the so-called DNase I subdomain of the native RNase E is involved in chelating an essential Mg⁺² ion. The aspartate residue is predicted to act as a general base to activate the attacking water essential for the catalytic activity of the enzyme (Callaghan et al 2005). The replacement of Asp-346 with the polar amino acid Asn was

previously shown to decrease RNA cleavage by about 25-fold (Callaghan et al 2005).

[0096] In yet a preferred embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 2, and wherein the amino acid sequence of said mutant RNase E has a single amino acid residue substitution at position D346 relative to SEQ ID NO. 2, wherein said amino acid is substituted to asparagine.

[0097] In one embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 2, and wherein the amino acid sequence of said mutant RNase E has an amino acid residue substitution which reduces the RNA recognition ability of the enzyme. In one such embodiment, said RNA recognition ability is reduced by substituting an amino acid at a position selected from G124, V128, R141, R142, R169, T170, R373, F57, F67, and K112 relative to SEQ ID NO. 2.

[0098] In one embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 2, and wherein the amino acid sequence of said mutant RNase E has an amino acid residue substitution which modifies an RNA contact point of the enzyme. In one such embodiment, said RNA contact point is modified by substituting an amino acid at a position selected from F57, F67, and K112 relative to SEQ ID NO. 2.

[0099] In one preferred embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 2, and wherein the amino acid sequence of said mutant RNase E has an amino acid residue substitution which results in the mutant RNase E having a modified RNA contact point compared to the RNase E of SEQ ID NO. 2. In a preferred embodiment, said substituting for modifying the RNA contact point is an amino acid at a position selected from F57, F67, and K112 relative to SEQ ID NO. 2.

[0100] In one embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 2, and wherein the amino acid sequence of said mutant RNase E has an amino acid residue substitution in the 5' sensor pocket of the enzyme, preferably the pocket 'anchors'. In one such embodiment, said pocket 'anchor' is modified by substituting an amino acid at a position selected from V128, and R373 relative to SEQ ID NO. 2.

[0101] In yet a preferred embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E

has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 2, and wherein the amino acid sequence of said mutant RNase E has an amino acid residue substitution at position A441 relative to SEQ ID NO. 2.

[0102] In one embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 4, and wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substitutions which results in the mutant RNase E having decreased activity compared to the RNase E of SEQ ID NO. 4.

[0103] In one embodiment, the activity of the mutant RNase E is decreased by at least 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, or 100% compared to the RNase E of SEQ ID NO. 4 when expressed in a host cell of the invention. In one embodiment, the activity of the mutant RNase E is decreased at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 fold, such as even 20, 40, 60, 80, 100 fold, such as even 200, 300, 400, 500, 600, 700, 800, 900 or 1000 fold compared to the RNase E of SEQ ID NO. 4 when expressed in a host cell of the invention.

[0104] In one embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 4, and wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substitutions at one or more positions selected from D343, E294, D300, N302, E322, R334, D346, V126, R167, T168, F57, F67, K110, G122, R139, R140, R370, and A438 relative to SEQ ID NO. 4.

[0105] In one embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 4, and wherein the amino acid sequence of said mutant RNase E has an amino acid residue substitution which reduces the metal ion chelation ability of the enzyme. In a preferred embodiment, said metal ion chelation ability is reduced by substituting an amino acid at a position selected from D343, E294, D300, N302, E322, R334, and D346 relative to SEQ ID NO. 4.

[0106] In one preferred embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 4, and wherein the amino acid sequence of said mutant RNase E has an amino acid residue substitution which results in the mutant RNase E having reduced metal ion chelation ability compared to the RNase E of SEQ ID NO. 4. In a preferred embodiment, said metal ion chelation ability is reduced by

substituting an amino acid at a position selected from D343, E294, D300, E322, R334, and D346 relative to SEQ ID NO. 4.

[0107] In one embodiment, the metal ion chelation ability of the mutant RNase E is decreased by at least 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, or 100% compared to the RNase E of SEQ ID NO. 4 when expressed in a host cell of the invention. In one embodiment, the metal ion chelation ability of the mutant RNase E is decreased at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 fold, such as even 20, 40, 60, 80, 100 fold, such as even 200, 300, 400, 500, 600, 700, 800, 900 or 1000 fold compared to the RNase E of SEQ ID NO. 4 when expressed in a host cell of the invention.

[0108] In a preferred embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 4, and wherein the amino acid sequence of said mutant RNase E has a single amino acid residue substitution at position D343 relative to SEQ ID NO. 4, wherein said amino acid is substituted to any amino acid other than aspartate.

[0109] In yet a preferred embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 4, and wherein the amino acid sequence of said mutant RNase E has a single amino acid residue substitution at position D343 relative to SEQ ID NO. 4, wherein said amino acid is substituted to asparagine.

[0110] In one embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 4, and wherein the amino acid sequence of said mutant RNase E has an amino acid residue substitution which reduces the RNA recognition ability of the enzyme. In one such embodiment, said RNA recognition ability is reduced by substituting an amino acid at a position selected from G122, V126, R139, R140, R167, T168, R370, F57, F67, and K110 relative to SEQ ID NO. 4.

[0111] In one embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 4, and wherein the amino acid sequence of said mutant RNase E has an amino acid residue substitution which modifies an RNA contact point of the enzyme. In one such embodiment, said RNA contact point is modified by substituting an amino acid at a position selected from F57, F67, and K110 relative to SEQ ID NO. 4.

[0112] In one preferred embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 4, and wherein the amino acid sequence of said mutant RNase E has an amino acid residue substitution which results in the mutant RNase E having a modified RNA contact point compared to the RNase E of SEQ ID NO. 2. In a preferred embodiment, said substituting for modifying the RNA contact point is an amino acid at a position selected from F57, F67, and K110 relative to SEQ ID NO. 4.

or 100% sequence identity with SEQ ID NO. 4, and wherein the amino acid sequence of said mutant RNase E has an amino acid residue substitution which results in the mutant RNase E having a modified RNA contact point compared to the RNase E of SEQ ID NO. 2. In a preferred embodiment, said substituting for modifying the RNA contact point is an amino acid at a position selected from F57, F67, and K110 relative to SEQ ID NO. 4.

[0113] In one embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 4, and wherein the amino acid sequence of said mutant RNase E has an amino acid residue substitution in the 5' sensor pocket of the enzyme, preferably the pocket 'anchors'. In one such embodiment, said pocket 'anchor' is modified by substituting an amino acid at a position selected from V126, and R370 relative to SEQ ID NO. 4.

[0114] In yet a preferred embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 4, and wherein the amino acid sequence of said mutant RNase E has an amino acid residue substitution at position A438 relative to SEQ ID NO. 4.

[0115] As mentioned previously, though not being homologous, *B. subtilis* RNase Y can be functionally replaced by *E. coli* RNase E.

[0116] In one embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase Y, wherein the amino acid sequence of said mutant RNase Y has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 6, and wherein the amino acid sequence of said mutant RNase Y has one or more amino acid residue substitutions which results in the mutant RNase Y having decreased activity compared to the RNase Y of SEQ ID NO. 6.

[0117] In one embodiment, the activity of the mutant RNase Y is decreased by at least 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, or 100% compared to the RNase E of SEQ ID NO. 6 when expressed in a host cell of the invention. In one embodiment, the activity of the mutant RNase Y is decreased at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 fold, such as even 20, 40, 60, 80, 100 fold, such as even 200, 300, 400, 500, 600, 700, 800, 900 or 1000 fold compared to the RNase E of SEQ ID NO. 6 when expressed in a host cell of the invention.

[0118] No amino acid positions for *B. subtilis* RNase Y can be found which are homologue to amino acid positions for *E. coli* RNase E listed in table 1, however, the His-Asp doublet conserved in HD domain proteins (such as RNase Y) in amino acid position H368 and D369 of *B. subtilis* RNase Y is similarly involved in metal chelation as D346 in *E. coli* RNase E. RNase Y mutants comprising amino acid substitutions H368A or D369A show lower activity/cleavage than RNase Y wild type (Shahbabian et al 2009).

[0119] In one preferred embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase Y has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 6, and wherein the amino acid sequence of said mutant RNase Y has an amino acid residue substitution which results in the mutant RNase Y having reduced metal ion chelation ability compared to the RNase Y of SEQ ID NO. 6.

[0120] In one embodiment, the metal ion chelation ability of the mutant RNase Y is reduced by at least 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, or 100% compared to the RNase E of SEQ ID NO. 6 when expressed in a host cell of the invention. In one embodiment, the metal ion chelation ability of the mutant RNase Y is reduced at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 fold, such as even 20, 40, 60, 80, 100 fold, such as even 200, 300, 400, 500, 600, 700, 800, 900 or 1000 fold compared to the RNase E of SEQ ID NO. 6 when expressed in a host cell of the invention.

[0121] In one embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase Y, wherein the amino acid sequence of said mutant RNase Y has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 6 and wherein the amino acid sequence of said mutant RNase Y has an amino acid residue substitution at a position selected from H368 and D369 relative to SEQ ID NO. 6.

[0122] In one embodiment, the microbial cell comprises a recombinant gene encoding a mutant RNase Y, wherein the amino acid sequence of said mutant RNase Y has at least 70, 71, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO. 6, and wherein the amino acid sequence of said mutant RNase Y has an amino acid residue substitution selected from H368A and D369A relative to SEQ ID NO. 6.

[0123] In one embodiment, the decreased activity, reduced metal ion chelation ability, modified RNA contact point, etc. for the mutant RNase E (or RNase Y) described above is relative to an RNase E (or RNase Y) having SEQ ID NOs referred to herein. In another embodiment, the decreased activity, reduced metal ion chelation ability, modified RNA contact point, etc. for the mutant RNase E (or RNase Y) described above is relative to the same identical RNase E (or RNase Y) enzyme, but lacking the specific mutation. In yet another embodiment, the decreased activity, reduced metal ion chelation ability, modified RNA contact point, etc. for the mutant RNase E (or RNase Y) described above is relative to the activity of the endoribonuclease enzyme encoded on the genome of the microbial host cell.

[0124] In a further embodiment, in addition to the rne* gene comprising one or more point mutations as disclosed herein resulting in one or more amino acid substitutions in the catalytic N-terminal domain of RNase E enzyme modulating the activity of the RNase E enzyme compared to the non-mutated version of the enzyme, said host cell further comprises a gene encoding T7 lysozyme (LysS) (E.C. 3.5.1.28). In one embodiment, plasmid pLyS (SEQ ID NO. 29, see table 3 for further specifications) serves as the backbone for a vector comprising the recombinant rne* gene.

[0125] Expression of said mutant RNase E may be constitutive or regulated by an inducible promoter, such as the rhaBAD promoter (SEQ ID NO 41), araBAD promoter (SEQ ID NO 88), Ptrc promoter (SEQ ID NO 89), T7 promoter, (SEQ ID NO 47), Ptet promoter (SEQ ID NO 90), Ptac promoter (SEQ ID NO 91), PL promoter (SEQ ID NO 92).

[0126] Expression of said mutant RNase E may further be regulated by optimizing the translational strength of the rne* gene. In bacteria, translational strength is defined by the Shine-Dalgarno/ribosome binding site (RBS) sequence directly upstream of the start codon, while in eukaryotic cells translation initiation regions/Kozak elements can be used to modify translational strength. RBSs in *E. coli* conferring a broad range of translational strengths can be found in the literature, e.g. Bonde et al, 2016. A skilled person in the art can optimize the translational strength of the gene by constructing variants of the ribosomal binding site and testing which variant performs better.

[0127] In a preferred embodiment, the host cell of the invention comprises plasmid pMax (SEQ ID NO. 38) comprising the recombinant rne* gene encoding a mutant RNase E.

II. Auxiliary Plasmid Comprising Recombinant Gene Encoding Mutant RNase E

[0128] In one aspect, the invention provides a plasmid comprising a recombinant rne* gene encoding a mutant RNase E and a lysS gene encoding T7 lysozyme (E.C. 3.5.1.28), as described herein.

[0129] Any of the above mentioned favorable rne mutations may be exploited on a plasmid comprising (a) a gene encoding the mutant rne enzyme as well as (b) a gene encoding a T7 lysozyme.

[0130] The plasmid may be prokaryotic or eukaryotic, comprising an origin of replication suitable for replication in the respective prokaryotic or eukaryotic host cell, independently of the chromosome.

[0131] In one embodiment, the plasmid comprises

[0132] a. a gene encoding a mutant RNase E having at least 75% amino acid sequence identity to SEQ ID NO. 2, wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substitutions which results in the mutant RNase E having decreased activity compared to the RNase E of SEQ ID NO. 2, and

[0133] b. a gene encoding a T7 lysozyme (E.C. 3.5.1.28)

[0134] In a preferred embodiment, the plasmid comprises

[0135] a. a gene encoding a mutant RNase E (E.C. 3.1.26.12) having at least 75% amino acid sequence identity to SEQ ID NO. 2, wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substitutions at one or more positions selected from D346, E297, D303, N305, E325, R337, D349, V128, R169, T170, F57, F67, K112, G124, R141, R142, R373, and A441 relative to SEQ ID NO. 2, and

[0136] b. a gene encoding a T7 lysozyme (E.C. 3.5.1.28)

[0137] In a further preferred embodiment, the plasmid the amino acid residue at position 346 of said mutant RNase E relative to SEQ ID NO. 2 is asparagine.

- [0138] In one embodiment, the plasmid comprises
- [0139] a. a gene encoding a mutant RNase E (E.C. 3.1.26.12) having at least 75% amino acid sequence identity to SEQ ID NO. 2, wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substitutions at one or more positions in the DNase I-like domain, and
- [0140] b. a gene encoding a T7 lysozyme (E.C. 3.5.1. 28)
- [0141] In one embodiment, the plasmid comprises
- [0142] a. a gene encoding a mutant RNase E (E.C. 3.1.26.12) having at least 75% amino acid sequence identity to SEQ ID NO. 2, wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substitutions which results in a mutant RNase E having reduced metal ion chelation ability, and
- [0143] b. a gene encoding a T7 lysozyme (E.C. 3.5.1. 28)
- [0144] In one embodiment, the plasmid comprises
- [0145] a. a gene encoding a mutant RNase E (E.C. 3.1.26.12) having at least 75% amino acid sequence identity to SEQ ID NO. 2, wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substitutions which results in a mutant RNase E having a modified RNA contact point, and
- [0146] b. a gene encoding a T7 lysozyme (E.C. 3.5.1. 28)
- [0147] In one embodiment, the plasmid comprises
- [0148] a. a gene encoding a mutant RNase E (E.C. 3.1.26.12) having at least 75% amino acid sequence identity to SEQ ID NO. 2, wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substitutions at one or more positions in the 5' sensor pocket, preferably the pocket 'anchors', and
- [0149] b. a gene encoding a T7 lysozyme (E.C. 3.5.1. 28)
- [0150] In one embodiment, plasmid pLyS (SEQ ID NO. 29) serves as the backbone for the plasmid comprising the recombinant rne* gene. In one preferred embodiment, the plasmid comprising a recombinant rne* gene encoding a mutant RNase E of the invention is pMax (SEQ ID NO. 38).
- [0151] Preparation of plasmid comprising the mutant rne gene may be done by any suitable cloning technique known by a person skilled in the art.

III Target Protein(s) of the Invention

[0152] The microbial host cell of the present invention expressing an auxiliary plasmid comprising mutant rne* as described herein has the advantage of providing a more robust and high yielding process for the production of target proteins, compared to industrially common used strains derived from *E. coli* BL21(DE3), e.g. BL21(DE)pLysS or BL21Star (DE3).

[0153] The present invention is particularly suitable for the expression of proteins that cause a burden and negatively affect the fitness of the cell in other commonly used expression systems. A commonly used term for such proteins is 'toxic proteins'. Some proteins may be difficult to express due to their size, complexity in folding, aggregation issue, etc. or their expression may cause problems with resource competition for other essential proteins or genes in the cell.

[0154] A person skilled in the art will recognize that the burden refers to product-specific metabolic toxicity which a microbial host cell genetically engineered to synthesize the target protein experiences during production of the product under production conditions, and which results in a fitness cost that can be quantified by measuring the percent reduction in the maximum exponential growth rate of the cell (along the growth curve) during production of the product under production conditions as compared to a parent microbial cell devoid or incapable of said production when grown under comparable production conditions.

[0155] Expression of the target protein is significantly enhanced when expressed in a host cell of the invention, compared to expression of said protein in another host cell lacking the rne* gene as disclosed herein.

[0156] In one embodiment, the target protein of the invention is a protein the expression of which is enhanced by at least 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, or 100% when expressed in a host cell of the invention, compared to expression of said target protein in the same host cell lacking the rne* gene or the rne* in combination with LysS, as disclosed herein.

[0157] The target protein of the invention is a protein the expression of which is enhanced at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 fold, such as even 20, 40, 60, 80, 100 fold, such as even 200, 300, 400, 500, 600, 700, 800, 900 or 1000 fold when expressed in a host cell of the invention, compared to expression of said target protein in the same host cell lacking the rne* gene or the rne* in combination with LysS, as disclosed herein.

[0158] In one embodiment, the target protein is selected from membrane proteins, antibody-like proteins, industrial enzymes such as carbohydrases, proteases, and lipases; and pharmaceutical proteins such as peptides, hormones, and proteins for vaccine development. The present invention is widely applicable of a range of different proteins which have proven difficult to express using other common expression systems. As mentioned in the background section, expression of membrane proteins is considered 'toxic' for the host cell, and they are notoriously known for causing burden in expression systems. Hence, in a preferred embodiment, the target protein is a membrane protein.

[0159] In one embodiment, the target protein is a soluble protein; in another embodiment the target protein is an insoluble protein. In one embodiment the target protein is a secreted protein; in another embodiment the target protein is a non-secreted protein. In one embodiment the target protein is of eukaryotic origin; in another embodiment the target protein is of prokaryotic origin. In one embodiment the target protein is native to the host cell; in another embodiment the target protein is non-native to the host cell.

[0160] One aspect of the invention concerns the use of a prokaryotic vector of the invention, as disclosed here, for enhancing the expression of a recombinant gene encoding a target protein in a prokaryotic host cell.

IV Methods of Preparing a Microbial Host Cell of the Invention for Expression of a Target Protein

[0161] Bacterial transformation may be referred to as a stable genetic change brought about by taking up DNA, and competence refers to the state of being able to take up exogenous DNA. Some bacteria are naturally capable of

taking up DNA under laboratory conditions and such species carry sets of genes specifying machinery for bringing DNA across the cell's membrane or membranes, while others have to be induced by laboratory procedures in which cells are passively made permeable to DNA, using conditions that do not normally occur in nature. Chilling cells in the presence of divalent cations such as Ca²⁺ (in CaCl₂) prepares the cell walls to become permeable to plasmid DNA. Cells are incubated on ice with the DNA and then briefly heat-shocked (e.g. 42° C. for 30-120 seconds), which causes the DNA to enter the cell and is a well-known method in the art [Sambrook et al., A Laboratory Manual (1989) CSH]. Electroporation is another way to make cells take up DNA. To persist and be stably maintained in the cell, a plasmid DNA molecule must contain an origin of replication, which allows it to be replicated in the cell independently of the chromosome.

[0162] In one aspect, the invention provides a method for producing a recombinant microbial cell having enhanced expression of a recombinant target protein.

[0163] In one embodiment, the method for producing a recombinant microbial cell having enhanced expression of a recombinant target protein comprises the steps

[0164] a. providing a microbial cell comprising

[0165] i. a gene on the genome of said cell encoding an enzyme having endoribonuclease activity (E.C. 3.1.26), and

[0166] ii. a gene encoding said recombinant target protein,

[0167] b. transforming said microbial cell with a prokaryotic vector comprising a recombinant gene encoding a mutant RNase E (E.C. 3.1.26.12) having at least 75% amino acid sequence identity to SEQ ID NO. 2, and wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substitutions at one or more positions which results in a mutant RNase E having decreased activity compared to the RNase E of SEQ ID NO. 2,

[0168] wherein expression of said target protein is enhanced compared to a cell not transformed with said prokaryotic vector.

[0169] In one embodiment, in the method for producing a prokaryotic recombinant microbial cell having enhanced expression of a recombinant target protein, said microbial cell further comprises (iii) a gene encoding a T7 RNA polymerase (E.C. 2.7.7.6), and either said microbial cell or said prokaryotic vector comprises a gene encoding a T7 lysozyme (E.C. 3.5.1.28), and expression of the said recombinant gene encoding said target protein is regulated by an inducible T7 promoter

[0170] In one preferred embodiment, the method for producing a prokaryotic recombinant microbial cell having enhanced expression of a recombinant target protein comprises the steps

[0171] a. providing an *E. coli* BL21(DE3) strain comprising a gene encoding said recombinant target protein,

[0172] b. transforming said microbial cell with pMax

[0173] In another embodiment, the method for producing a recombinant microbial cell having enhanced expression of a target protein comprises the steps

[0174] a. providing a microbial cell comprising on its genome

[0175] i. a gene on the genome of said cell encoding an enzyme having endoribonuclease activity (E.C. 3.1.26) and

[0176] ii. a first recombinant gene encoding a mutant RNase E (E.C. 3.1.26.12) having at least 75% amino acid sequence identity to SEQ ID NO. 2, and wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substitutions which results in a mutant RNase E having decreased activity compared to the RNase E of SEQ ID NO. 2,

[0177] b. transforming said microbial cell with a vector comprising a second recombinant gene encoding said target protein,

[0178] wherein expression of said target protein is enhanced compared to a cell lacking said first recombinant gene.

V Method for the Production of a Target Protein

[0179] In one aspect, the present invention provides a method for enhancing recombinant protein expression of a target protein, comprising the steps of

[0180] a. providing a microbial host cell of the invention as described herein,

[0181] b. culturing said host cell in a suitable culture medium, such as a medium that supports growth of said host cell,

[0182] c. optionally inducing expression of the recombinant gene encoding the target protein, and

[0183] d. optionally isolating and purifying the expressed target protein by well-known techniques.

VI Method for Screening for Dominant RNase E Mutations

[0184] The microbial cell of the Invention comprises wild type RNase E gene on its genome as well as a mutant RNase E on a plasmid. Hence, the cell produces both wild type and mutant RNase E. For enhancing expression of a target protein, the mutation in the RNase E enzyme presumably facilitates a dominance over the wild type RNase E.

[0185] An example of a method for screening for dominant mutations within the RNase E gene that provide enhanced expression of a recombinant target protein is provided below.

[0186] To screen for dominant mutations within the RNase E gene, a person skilled in the art may use a microbial cell comprising:

[0187] i. a gene on the genome of said cell encoding an RNase enzyme (E.C. 3.1.26), and

[0188] ii. a gene encoding a recombinant YidC-GFP fusion protein (=target protein use in the screening)

[0189] The skilled person should then transform said microbial cell with a prokaryotic vector comprising a recombinant gene encoding a mutant RNase E (that is aimed to be tested/screened) having at least 75% amino acid sequence identity to SEQ ID NO. 2, and wherein the amino acid sequence of said mutant RNase E has an amino acid residue substitution at position X, wherein position x is the residue aimed to be tested/screened. If the tested mutant RNase E exhibits dominance over the genetically encoded wildtype RNase E, fluorescence of the GFP coupled YidC-fusion protein can be observed. If the investigated RNase E muta-

tion is not dominant over the wildtype RNase E enzyme no fluorescence (or only very little) is observed. The same prokaryotic vector, but lacking the mutant RNase E, may be used as negative control. A dominant RNase mutant will facilitate enhanced fluorescence compared to the negative control.

[0190] In one embodiment, the prokaryotic microbial host cell of the invention comprises

[0191] A. a gene encoding an enzyme having endoribonuclease activity (E.C. 3.1.26), wherein said gene is on the genome of said cell,

[0192] B. a first recombinant gene encoding a mutant RNase E, and

[0193] C. a second recombinant gene encoding a target protein,

wherein the mutant RNase E enzyme has one or more amino acid residue substitutions which facilitates enhanced expression of the target protein, wherein said amino acid residue substitution is identified and selected by a screening method comprising the steps of

[0194] A. expressing the target protein together with a candidate mutant RNase E comprising a candidate amino acid residue substitution in the host cell,

[0195] B. expressing the target protein in a parent cell (from which the host cell was derived) lacking expression of the candidate mutant RNase E,

[0196] C. comparing expression levels of the target protein in (a) and (b), and identifying one or more candidate(s) which facilitate enhanced expression of said target protein.

[0197] The candidate mutant RNase may be selected randomly, such as in screening a large library of candidate

mutant RNases, or be specific predicted mutations based on knowledge of protein structure, etc.

VII Measuring Target Protein Expression

[0198] The present invention provides a method for enhancing recombinant protein expression as well as a microbial host cell having enhanced expression of a target protein. The enhanced expression is relative to an otherwise identical cell, but which does not comprise a mutant rne* gene, as disclosed herein.

[0199] An increase in expression of a target protein may be measured by direct measurement of the amount of the target protein if an assay for such direct measurement of said target protein exists. Alternatively, the expression of the target protein may be measured by fusion of the target protein to a GFP for fluorescence detection. The protein, optionally fused to GFP, may be his-tagged for purification purposes.

Examples

[0200] Bacterial strains: Bacterial strains used in the examples are identified in Table 2. *Escherichia coli*/strain TOP10 (Thermo Fischer Scientific) was used for DNA manipulations, such as plasmid engineering and amplifications. Different strains of *E. coli* BL21 were used for the expression of genes of interest.

[0201] Bacterial strains were grown aerobically at either 37 or 30° C. in Luria-Bertani (LB) broth or agar, supplemented with 50 µg/ml kanamycin, 25 µg/ml chloramphenicol or 100 µg/ml ampicillin depending on the resistance marker of the plasmid used.

TABLE 2

Bacterial strains of the examples		
Bacterial strain names	Genetic features	Source
<i>E. coli</i> TOP10	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(Str ^R) endA1 λ ⁻	Invitrogen
<i>E. coli</i> BL21	F- ompT gal dcm lon hsdS _B (r _B ⁻ m _B ⁻) [malB ⁺] _{K-12} (λ ^S)	Invitrogen
<i>E. coli</i> BL21(DE3)	F- ompT gal dcm lon hsdS _B (r _B ⁻ m _B ⁻) λ(DE3) [lacI lacUV5-T7p07 ind1 sam7 nin5] [malB ⁺] _{K-12} (λ ^S) an <i>E. coli</i> B strain with DE3, a λ prophage carrying the T7 RNA polymerase gene and lacI ^q Transformed plasmids containing T7 promoter-driven expression are repressed until IPTG induction of T7 RNA polymerase from a lac promoter.	Invitrogen
<i>E. coli</i> BL21Star(DE3)	F- ompT gal dcm lon hsdS _B (r _B ⁻ m _B ⁻) λ(DE3) [lacI lacUV5-T7p07 ind1 sam7 nin5] [malB ⁺] _{K-12} (λ ^S) rne131 The strain carries a mutated rne gene (rne131) which encodes a truncated RNase E enzyme	Invitrogen
<i>E. coli</i> Evo21(DE3)	F- ompT gal dcm lon hsdS _B (r _B ⁻ m _B ⁻) [malB ⁺] _{K-12} (λ ^S) The strain carries a mutated rne gene which encodes a truncated RNase E enzyme	This study

[0202] Plasmids: Plasmids used in the examples are identified in Table 3. The construction of plasmids and transformation into the relevant microbial host cell were performed using standard molecular biology techniques recognized and practised without difficulty by a person skilled in the art, such as the techniques described in Sambrook et al. [A Laboratory Manual (1989) CSH].

[0203] Plate reader experiments. For growth and protein production assays, strains were cultured overnight in 5 mL LB liquid growth medium. Dilutions of 1:50 were grown aerobically for 24 hours in 96-well plates at 37° C. and 200 rpm using Gas Permeable Adhesive Seal (Thermo Fisher Scientific, Waltham, MA, USA) to avoid evaporation. 1 mM IPTG was added at OD₆₀₀=0.3. Growth (absorbance at 600

TABLE 3

Plasmids and genes of the examples		
Plasmid/Vector name	Genetic features	Source
pLysS (SEQ ID No.: 29)	CmR (SEQ ID No.: 30) cat promoter (SEQ ID No.: 32) P15A ori (SEQ ID No.: 33) Φ3.8 promoter (SEQ ID No.: 34) 3.5 (T7 lysozyme) (SEQ ID No.: 35) tet promoter (SEQ ID No.: 37)	Studier, 1991
pMax (SEQ ID No.: 38)	pLysS backbone rne(D346N) (SEQ ID No.: 39) rhaBAD promoter (SEQ ID No.: 41) BCD (SEQ ID No.: 42) rmc terminator (SEQ ID No.: 43) T7 terminator (SEQ ID No.: 44)	This study
pMax-noLysS (SEQ ID No.: 45)	CmR (SEQ ID No.: 30) cat promoter (SEQ ID No.: 32) P15A ori (SEQ ID No.: 33) tet promoter (SEQ ID No.: 37) rne(D346N) (SEQ ID No.: 39) rhaBAD promoter (SEQ ID No.: 41) BCD (SEQ ID No.: 42) rmc terminator (SEQ ID No.: 44)	This study
pET28a+ (SEQ ID No.: 46)	T7 promoter (SEQ ID No.: 47) Lac operator (SEQ ID No.: 48) f1 ori (SEQ ID No.: 49) KanR (SEQ ID No.: 50) pBR322_ori(SEQ ID No.: 52) bom (SEQ ID No.: 53) rop (SEQ ID No.: 54) lacI promoter (SEQ ID No.: 55) lacI (SEQ ID No.: 56)	Vendor EMD Biosciences addgene.org/vector-database/2565/
pP450 (SEQ ID No.: 58)	pET28a+ backbone Membrane protein P450-GFP-His8 (SEQ ID No.: 59)	Vazquez-Albacete et al., 2016
pHtpX (SEQ ID No.: 61)	pET28a+ backbone Membrane protein HtpX-GFP-His8 (SEQ ID No.: 62)	Daley et al., 2005
pYqik (SEQ ID No.: 64)	pET28a+ backbone Membrane protein YqiK-GFP-His8 (SEQ ID No.: 65)	Daley et al., 2005
pYidC (SEQ ID No.: 67)	pET28a+ backbone Membrane protein YidC-GFP-His8 (SEQ ID No.: 68) hp6 (SEQ ID No.: 70) AmpR (SEQ ID No.: 71)	This study
pMax-truncV489 (SEQ ID No.: 73)	pMax backbone rne(D346N) truncated at V489 (SEQ ID No.: 74)	This study
pMax-truncL529 (SEQ ID No.: 76)	pMax backbone rne(D346N) truncated at L529 (SEQ ID No.: 77)	This study
pMax-D346X (SEQ ID No.: 79)	pMax backbone rne(D346X) (SEQ ID No.: 80)	This study
pMax-rne ^{WT} (SEQ ID No.: 81)	pMax backbone rne wildtype (SEQ ID No.: 2)	This study
pMax-rne ^{STAR} (SEQ ID No.: 82)	pMax backbone rne131 present in BL21Star(DE3) (SEQ ID No.: 83)	This study BL21Star(DE3): Kido et al., 1996
pMax-rne ^{Evo21} (SEQ ID No.: 85)	pMax backbone rne version present in Evo21(DE3) (SEQ ID No.: 86)	This study

nm) and fluorescence (GFP: excitation at 485 nm, emission at 528 nm) was measured in 20 min intervals while continuous shaking using a Synergy H1 plate reader (BioTek Instruments, Winooski, VT, USA). All measurements were performed in triplicate.

[0204] Statistical analysis. All experiments were performed in triplicate. Error bars and significance values were calculated using the program PRISM. Error bars indicated represent the average squared deviation from the mean (SD). A one-way ANOVA with Dunnett's multiple comparison test was employed to evaluate differences in expression levels of recombinant protein between Evo21(DE3) and other expression hosts. P values <0.05 were accepted as statistically significant. The different significance levels indicated as stars in figures correspond to p-value <0.05 (*), p<0.01 (**), p<0.001 (***) and p<0.0001 (****).

Example 1: Rne Point Mutation on Auxiliary Plasmid Increases Protein Production

[0205] As discussed previously, truncation of the rne locus, such as rne131 in the commercially available BL21Star (DE3) resulting in the RNase E polypeptide lacking its non-catalytic region (while retaining amino acid residues 1-584), causes a bulk stabilization of mRNA degradation, including mRNA produced by T7RNAP (Lopez et al 1999).

[0206] Similarly, the present inventors, while evolving a BL21(DE3) strain by tailored evolution to overcome the challenges of protein production toxicity, have identified a truncated me mutant highly efficient in protein production. This strain was named Evo21(DE3) and comprises a truncation of the encoded 1061-residue *E. coli* endoribonuclease RNase E after amino acid 702 and therefore a polypeptide lacking the last 359 residues of its C-terminus (see FIG. 1A). Evo21(DE3) was isolated from outgrowing bacterial colonies on week-old agar plates. This combined with the knowledge that dominant rne mutants have previously been observed (Briegel et al 2006) implied that different rne variants could be studied by simple co-expression from a plasmid in the presence of the wildtype me on the genome.

[0207] Membrane protein YidC was produced as a C-terminal GFP-His8 fusion protein from a pET28a+ derived expression vector (pYidC: SEQ ID NO.: 67), as shown in Table 3 and described by Drew et al 2006. The expression vector comprises a lacI gene, under the control of its native lacI promoter, encoding the lac repressor that binds to a lacO (operator) upstream of the target YidC gene and blocks its expression. When the target gene is operably linked to a T7 promoter (as in pET28a+), then the expression of the target gene is first induced upon the addition of IPTG.

[0208] Plasmid pLysS (SEQ ID NO. 29) comprising a gene encoding T7 lysozyme (SEQ ID NO. 36) was utilised to limit basal T7 RNAP expression (Studier et al 1991). Plasmid pLysS was chosen to function as a backbone for co-expression of different rne variants to avoid the cellular burden of having three plasmids present simultaneously. All rne genes expressed on the pLysS backbone were cloned seamlessly in between a lysS terminator and a T7 ϕ3.8 promoters controlling T7 lysS expression flanked by additional terminator sequences: T7 terminator (upstream) and rnc terminator (downstream).

[0209] To compare different variant of the rne gene at different expression levels, the me variants were cloned in front of the rhamnose-inducible rhaBAD promoter on the

pLysS plasmid backbone—see illustration pMAX-rne^X in FIG. 2A, where rne^X is either a full-length rne (rne^{WT}) and truncated rne (rne^{STAR} and rne^{Evo21}) were successfully cloned. Further, a spontaneous mutant: rne^{D345N} was included in the study.

[0210] rne^{WT} encodes full-length *E. coli* RNase E (SEQ ID NO.: 2). Truncated rne^{STAR} encodes truncated *E. coli* RNase E polypeptide amino acids 1-584 (SEQ ID NO.: 84). Truncated rne^{Evo21} encodes truncated RNase E polypeptide amino acids 1-702 (SEQ ID NO.: 87).

[0211] Mutant rne^{D346N} encodes full-length RNase E polypeptide comprising amino acid substitution D346N (SEQ ID NO.: 40).

[0212] *E. coli* BL21(DE3) cells co-expressing pLysS and pYidC fail to express detectable levels of the YidC-GFP fusion protein. When the pLysS plasmid was substituted with a pLysS plasmid expressing the full-length rne^{WT} gene or the truncated rne^{STAR} or rne^{Evo21} genes neither had a positive effect on YidC-GFP expression level. However, the mutant rne^{D346N} was found to significantly enhance YidC-GFP production (see FIG. 2B).

Example 2: Rne Point Mutation on Auxiliary Plasmid Increases Protein Production Compared to Genomic Rne Truncation

[0213] Expression of the rne^{D346N} mutant gene on a plasmid (pMax) in *E. coli* strain BL21(DE3) comprising the native rne gene of the genome was tested further to establish that its positive effect on heterologous protein production levels was not limited to the expression of YidC as illustrated in example 1, but that the effect of the rne^{D346N} gene in pMax is more versatile.

[0214] pMax (SEQ ID NO.: 38) comprises the pLysS as the backbone, with the rne^{D346N} mutant gene operably linked to the rhabAD promoter on the pLysS plasmid backbone, see illustration in FIG. 3A. The LysS encoding T7 lysozyme helps limit basal T7 RNA polymerase expression.

[0215] Membrane proteins P450, HtpX, YqiK, and YidC were produced as C-terminal GFP-His8 fusion proteins from a pET28a+ derived expression vector (pP450: SEQ ID NO.: 58, pHtpX: SEQ ID NO.: 61, pYqiK: SEQ ID NO.: 64, and pYidC: SEQ ID NO.: 67), see illustration in FIG. 3A.

[0216] The effect of co-expression of pMax (comprising the rne^{D346N} mutant) compared to the pLysS control, in BL21(DE3) strains producing membrane proteins P450, HtpX, YqiK, and YidC from expression vectors pP450, pHtpX, pYqiK, and pYidC can be seen in FIG. 3B. The data clearly showed that co-expression of the rne (D346N) mutant significantly improved protein production and that the effect was not gene-specific.

[0217] This was further compared to the expression of the proteins in strain BL21Star (DE3)pLysS (comprising genomic truncated rne gene and expressing plasmid pLysS). It was found that BL21(DE3)pMax comprising the rne^{D346N} mutant provided on the pLysS plasmid could increase protein production even further than the commercially available solution *E. coli* host: BL21Star (DE3)pLysS (see FIG. 3B).

[0218] This provides a simple tool, in the form of an auxiliary plasmid: pMax, that can be transformed into other strains, to improve protein production titers.

Example 3: Synergy by Co-Expression of Rne Point Mutation and LysS

[0219] pMax harbours both a lysS gene encoding T7 lysozyme and a rne^{D346N} gene encoding an RNase E enzyme having an aspartate to asparagine substitution of amino acid residue 346. The T7 lysozyme—commonly expressed via the pLysS plasmid—counteracts the inherent leakiness of the T7 promotor when controlling expression of a GOI (gene of interest) upon IPTG induction in any *E. coli* strain harbouring a genetically integrated gene encoding T7 RNAP (T7 RNA polymerase). The presence of the T7 RNAP gene in such strains is annotated as “DE3”, e.g. In BL21 (DE3). To investigate whether the combination of the lysS and rne^{D346N} gene is essential to obtain the increase in protein production observed when co-expressing plasmid pMax with a pET/T7 expression vector such as demonstrated in example 2—expression levels of YidC were compared when co-expressing plasmids having either the lysS gene individually (pLysS: SEQ ID NO.: 29) or the rne^{D346N} gene individually (pMax-noLysS: SEQ ID NO.: 45) or harbouring both lysS and rne^{D346N} (pMax: SEQ ID NO.: 38) with a pET/T7 expression vector driving the expression of yidC membrane protein (pYidC: SEQ ID NO.: 67).

[0220] As seen in FIG. 4, pMax-noLys (rne^{D346N} alone) provides a statistically significant improvement in protein expression compared to pLysS (lysS alone), while pMax (the combination of lysS and rne^{D346N}) provides an even greater improvement compared to both pMax-noLys and pLysS, and thereby demonstrates a true synergy between rne^{D346N} and LysS.

Example 4: Rne Truncation and Point Mutation Combined on Auxiliary Plasmid

[0221] The effect of expressing an RNase enzyme having a combination of the amino acid mutation D346N along with truncation of the C-terminal domain of the enzyme was investigated. Two genes encoding different truncated rne^{D346N} versions of RNase E harboured on pMax only comprising the N-terminal catalytic half of the enzyme were prepared to explore if such truncated versions of the full rne^{D346N} gene on plasmid pMax would confer the same “protein production-enhancing effect” as observed with the single rne^{D346N} mutant on pMax. The two truncated rne versions encode amino acid residues 1-489 (pMax_truncV489: SEQ ID NO.:75) and 1-529 (pMax_truncL529: SEQ ID NO.:78) of the full rne, respectively, while still harbouring mutation D346N. The truncated rne^{D346N} versions harboured on pMax were co-expressed in BL21(DE3) with the pET/T7 expression vector driving the expression of yidC membrane protein (pYidC: SEQ ID NO.: 67). As seen in FIG. 5, the pMax carrying the truncated rne^{D346N} versions perform better than simply pLysS, but the non-truncated rne^{D346N} outperforms the truncated versions.

[0222] A spontaneous mutant of truncL529-rne^{D346N} at position A441V arose during the experiment. This mutant had elevated protein production closer to pMax (full-length rne^{D346N} gene) levels (see FIG. 5).

Example 5: Rne Genomic Truncation Combined with Point Mutation on Auxiliary Plasmid

[0223] BL21Star comprises genomic mutant rne131 which encodes a truncated RNase E lacking the C-terminal region. pMax was co-expressed with pYidC in both *E. coli*

BL21(DE3) and BL21Star (DE3) to examine the effect of having a full length vs a truncated RNase E expressed by the host cell genome in combination with the mutant RNase E encoded by rne*. FIG. 6 clearly shows that the expression of YidC in BL21(DE3) is significantly greater than expression in BL21Star (DE3). It was further shown that the expression of YidC when co-expressed with pMax is in fact no improvement over simply co-expression with pLysS.

Example 6: Role of Different Amino Acids Substitution: Rne^{D346X}

[0224] The importance of the specific nature of the amino acid residue replacing the native aspartate at position 346 in rne encoded RNase E was investigated. For this, a site-directed mutagenesis library was created comprising a range of possible amino acid substitutions in position D346. For this, plasmids similar to pMax were created in which the aspartate encoded within the rne gene in amino acid position 346 was exchanged to amino acids Phe, Leu, Ile, Val, Ser, Pro, Thr, Ala, Tyr, His, Gin, Lys, Glu, Cys and Gly. This pMax (rne^{D346X}) mutant library was co-expressed with pYidC in BL21(DE3) and expression of YidC was compared to co-expression of pLysS with pYidC. As seen in FIG. 7, all tested rne^{D346X} substitutions on pMax outperform pLysS and enhance expression of YidC, from which can be concluded that other amino acid substitutions in the same location D346 in the rne gene lead to the same effect described for the D346N mutation in pMax.

Example 7: Various Rne Mutations on Auxiliary Plasmids Increase Protein Production

[0225] Site-directed mutagenesis (see Table 4) was employed to prepare different substitutions in the RNase E enzyme predicted herein to alter RNase E functionality (see Table 1). Similar to example 6, a site-directed mutagenesis library was created comprising the amino acid substitutions specified in Table 4. This second pMax (rne^X) mutant library was co-expressed with pYidC in BL21(DE3) and expression of YidC was compared to co-expression of pMax (harbouring rne^{D346N}) with pYidC, co-expression of pLysS with pYidC as well as co-expression of pLysS-rneWT with pYidC. As seen in FIG. 8, several rne^X substitutions on pMax outperform pLysS and pLysS-rneWT, and enhance expression of the target protein YidC. Hence, amino acid substitutions in several different locations in the rne gene lead to the same effect described for the D346N mutation in pMax.

[0226] The data support that mutations in the DNase I like domain, such as exemplified by E297A, D303N, E325A, R337A, D346N, and R373A in FIG. 8, are favourable in regard to modulating the activity of the mutant RNase E, leading to enhanced expression of the target protein YidC.

[0227] The data also support that mutations which effect the metal iron chelating ability of the RNase E enzyme, such as exemplified by E297A, D303N, E325A, R337A, and D346N in FIG. 8, are favourable in regard to modulating the activity of the mutant RNase E, leading to enhanced expression of the target protein YidC.

[0228] The data further support that mutations which modify an RNA contact point of the RNase E enzyme, such as exemplified by F57A, F67A, K112A in FIG. 8, are

favourable in regard to modulating the activity of the mutant RNase E, leading to enhanced expression of the target protein YidC.

[0229] The data further support that mutations in the 5' sensor pocket, such as the pocket anchors, such as exemplified by V128A and R373A in FIG. 8, are favourable in regard to modulating the activity of the mutant RNase E, leading to enhanced expression of the target protein YidC.

[0230] Finally, the data further support that a mutation at A441 (FIG. 8) is favourable in regard to modulating the activity of the mutant RNase E, leading to enhanced expression of the target protein YidC.

TABLE 4

RNase E substitutions explored		
Position	Substitution	Codon change
Glu-297	Ala	GCG
Asp-303	Asn	GCG
Glu-325	Ala	GCG
Val-128	Ala	GCG
Arg-337	Ala	GCG
Phe57	Ala	GCG
Phe67	Ala	GCG
Lys112	Ala	GCG
R373	Ala	GCG
A441	Val	ACC

¹Garey et al 2009,

²Callaghan et al 2005,

³Kim et al 2014.

Example 8: Double Rne Mutation on Auxiliary Plasmid

[0231] Site-directed mutagenesis (similar to Example 7) was employed to prepare double substitution in the RNase E enzyme: rne^{D346N, X}. These double mutants were co-expressed with pYidC in BL21(DE3), and expression of YidC was compared to co-expression of pMax (harbouring rne^{D346N}) with pYidC, and co-expression of pLysS-rneWT with pYidC. As seen in FIG. 9, a slight improved effect was obtained for some of the double mutants.

Example 9: Improving Translation Initiation Region to Regulate Efficiency of Translation

[0232] Since both RNase E activity (altered via mutations) and the expression level of the enzyme itself (altered via expression optimisation) is predicted to be of importance for the underlying mechanism of pMax, a TIR (Translation Initiation Region) library is created using site-directed random mutagenesis. The efficiency of translation initiation is dependent on the nucleotide sequence of the TIR, which comprises the Shine-Dalgarno (SD) sequence and the regions up- and downstream of the SD and is often the rate-limiting step when it comes to protein production in bacteria. Nucleotide changes within this region can affect RNase E production levels greatly as they affect mRNA secondary structure and binding of the ribosome. The TIR ahead of the rne gene on pMax is randomized to obtain pMax versions expressing rne at different levels, which can potentially further enhance protein production titers using pMax. The creation of the TIR library, moreover, enables the isolation of pMax plasmid variants in which rne* expression levels can be precisely tuned via the already existing rhamnose-inducible promoter (PrhaBAD). For the creation of the

TIR library, the six nucleotides upstream of the ATG codon of the BCD located downstream of the PrhaBAD promoter controlling expression of the rne* gene on plasmid pMax are randomised (NNNNNN). Additionally, the 2nd and 3rd codon of the BCD will be replaced by AARGCN. The pMax mutant TIR library is then co-transformed into BL21 (DE3) along with the pYidC expression vector, and colonies are plated on LB agarose plates containing different concentrations of the inducer L-rhamnose (0.1-5 mM). Cells producing high amounts of the YidC-GFP fusion protein will be identified via green fluorescence visible under UV light and will be isolated and sequenced. This way, individual TIR sequences with optimal rne expression levels are identified based on their stimulatory effect on YidC protein production. Additionally, colonies that show YidC expression on 1 mM rha concentration plates will be restreaked on an agarose plate dilution series ranging from 0 to 5 mM rhamnose to identify TIRs that allow tight tunability of rne* expression on pMax. Such candidates will express high YidC levels on agarose plates containing 5 mM rhamnose and will not show YidC expression when plated on 0 mM rha, respectively.

Example 10: Effect of Plasmid Copy Number

[0233] The pLysS plasmid backbone currently used as the basis for pMax is replaced by other plasmid backbones propagated by different origins of replication (ori). To this end, rne* and lysS gene along with their respective regulatory elements (respective promoter and terminator elements) are transferred onto alternative vector backbones maintained in the cell with a) high copy number (e.g. plasmids carrying ori pColE1/F1 (300-500 copies) or pUC derivative pMB1 (500-700 copies)) or b) low copy number (e.g. plasmids harbouring ori pSC101 (5 copies), RK2 (4-7), p15A (10 copies), R6k (15-20 copies), ColE1 (15-20 copies), pMB1 (15-20 copies) or pBR322 (15-20 copies)). The plasmid backbone pMax variants are co-transformed along with pYidC into BL21(DE3), and optimal backbones are identified based on their stimulatory effect on YidC protein production. It is expected, that the low copy plasmid backbone pMax variants will outperform high copy vector versions and pLysS similarly to pMax enhancing expression of the target protein YidC, from which will be concluded that the plasmid backbone copy number is a limiting factor for an optimal tuning of RNase E* levels in the production host cell.

Example 11: Protein Production in Other DE3-Strains

[0234] It was investigated whether pMax has a positive effect on protein production in another DE3 strain (other than BL21(DE3))—hence another strain that performs T7 RNAP-dependent gene expression. Effect of pMax co-expression on Proteinase K production (activity expressed in cytosol) in *E. coli* Shuffle (DE3) was compared to co-expression using pLysS or pET-empty vector (FIG. 10), finding a positive effect for pMax.

Example 12: Protein Production in Non DE3 Strains

[0235] It is investigated whether pMax or the co-expression of solely rne (D346N) (pMax-nlysS) has a positive effect on protein production in non DE3 strains—hence strains that do not perform T7 RNAP-dependent gene

expression. Expression levels and growth of *E. coli* BL21 (DE3) and *E. coli* non-DE3 strains: BL21 and K12 MG1655 are compared, when co-expressing pMax and a GOI such as yidC under the control of non-T7-dependent promoter systems such as the rhamnose-inducible PrhaBAD, the arabinose-Inducible promoter ParaBAD, the tryptophan-inducible promoter Ptrc, the tetracycline-Inducible promoter Ptet, the IPTG-inducible promotor (Ptac) and the heat-Inducible promoter PL.

Example 13: Expression of Challenging/Toxic Proteins (YldC) Vs Non-Toxic Proteins (GFP)

[0236] The heterologous production of GFP-fusion protein YidC-GFP and GFP in BL21(DE3) was investigated when co-expressing either auxiliary plasmid pLysS (lysS) or pMax (lysS, mD346N). As seen in FIG. 11, pMax facilitates an improvement in expression compared to pLysS.

REFERENCES

- [0237] Angius, F.; Illoia, O.; Amrani, A.; Suisse, A.; Rosset, L.; Legrand, A.; Abou-Hamdan, A.; Uzan, M.; Zito, F.; Miroux, B. A Novel Regulation Mechanism of the T7 RNA Polymerase Based Expression System Improves Overproduction and Folding of Membrane Proteins. *Sci. Rep.* 2018, 8 (1), 1-11. doi.org/10.1038/s41598-018-26668-y.
- [0238] Baumgarten, T.; Schlegel, S.; Wagner, S.; Löw, M.; Eriksson, J.; Bonde, I.; Herrgård, M. J.; Heipieper, H. J.; Nørholm, M. H. H.; Slotboom, D. J.; et al. Isolation and Characterization of the *E. Coli* Membrane Protein Production Strain Mutant56(DE3). *Sci. Rep.* 2017, 7 (March), 1-14. doi.org/10.1038/srep45089.
- [0239] Bonde, M. T., Pedersen, M., Klausen, M. S., Jensen, S. I., Wulff, T., Harrison, S., et al. (2016). Predictable tuning of protein expression in bacteria. *Nat. Methods* 13. doi:10.1038/nmeth.3727.
- [0240] Briegel, K. J.; Baker, A.; Jain, C. Identification and Analysis of *Escherichia Coli* Ribonuclease E Dominant-Negative Mutants. *Genetics* 2006, 172 (1), 7-15. doi.org/10.1534/genetics.105.048553.
- [0241] Callaghan, A. J.; Marcaida, M. J.; Stead, J. A.; McDowall, K. J.; Scott, W. G.; Luisi, B. F. Structure of *Escherichia Coli* RNase E Catalytic Domain and Implications for RNA Turnover. *Nature* 2005, 437 (7062), 1187-1191. doi.org/10.1038/nature04084.
- [0242] Daley et al: Global topology analysis of the *Escherichia coli* inner membrane proteome. *Science* Vol 308, Issue 5726, 27 May 2005 DOI: 10.1126/science.1109730
- [0243] Drew, D.; Lerch, M.; Kunji, E.; Slotboom, D. J.; de Gier, J. W. Optimization of Membrane Protein Overexpression and Purification Using GFP Fusions. *Nat. Methods* 2006, 3 (4), 303-313. doi.org/10.1038/nmeth0406-303.
- [0244] Fellmeier et al.: Green fluorescent protein functions as a reporter from protein localization in *Escherichia coli*. American Society for Microbiology. Journal of Bacteriology. 2000 DOI: 10.1128/JB.182.14.4068-4076.2000
- [0245] Garrey, S. M.; Blech, M.; Riffell, J. L.; Hankins, J. S.; Stickney, L. M.; Diver, M.; Hsu, Y. H. R.; Kunanithy, V.; Mackie, G. A. Substrate Binding and Active Site Residues in RNases E and G: Role of the 5'-Sensor. *J. Biol. Chem.* 2009, 284 (46), 31843-31850. doi.org/10.1074/jbc.M109.063263.
- [0246] Kido, M.; Yamanaka, K.; Mitani, T.; Niki, H.; Ogura, T.; Hiraga, S. RNase E Polypeptides Lacking a Carboxyl-Terminal Half Suppress a MukB Mutation in *Escherichia Coli*. *J. Bacteriol.* 1996, 178 (13), 3917-3925. doi.org/10.1128/jb.178.13.3917-3925.1996.
- [0247] Kim, D. et al. Modulation of RNase E activity by alternative RNA binding sites. *PLoS One* 9, (2014).
- [0248] Kwon, S. K.; Kim, S. K.; Lee, D. H.; Kim, J. F. Comparative Genomics and Experimental Evolution of *Escherichia Coli* BL21(DE3) Strains Reveal the Landscape of Toxicity Escape from Membrane Protein Over-production. *Sci. Rep.* 2015, 5 (October), 16076. doi.org/10.1038/srep16076.
- [0249] Lopez, P. J.; Marchand, I.; Joyce, S. A.; Dreyfus, M. The C-Terminal Half of RNase E, Which Organizes the *Escherichia Coli* Degradosome, Participates in mRNA Degradation but Not RRNA Processing in Vivo. *Mol. Microbiol.* 1999, 33 (1), 188-199. doi.org/10.1046/j.1365-2958.1999.01465.x.
- [0250] Mardle et al: A structural and biochemical comparison of Ribonuclease E homologues from pathogenic bacteria highlights species-specific properties. *Scientific Reports* volume 9, Article number: 7952 (2019). doi.org/10.1038/s41598-019-44385-y
- [0251] Miroux & Walker: Over-production of proteins in *Escherichia coli* mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels, 289-298 (1996).
- [0252] Shahabian et al: RNase Y, a novel endoribonuclease, initiates riboswitch turnover in *Bacillus subtilis*. *EMBO J* (2009)28:3523-3533 doi.org/10.1038/emboj.2009.283
- [0253] Studier, F. W. Use of Bacteriophage T7 Lysozyme to Improve an Inducible T7 Expression System. *J. Mol. Biol.* 1991, 219 (1), 37-44. doi.org/10.1016/0022-2836(91)90855-Z.
- [0254] Vazquez-Albacete et al.: An expression tag toolbox for microbial production of membrane bound plant cytochromes P450. *Biotechnology and Bioengineering* Vol 144, Iss 4, pp 751-760. April 2017. doi: 10.1002/bit.26203

ITEMS OF THE INVENTION

- [0255] 1. A prokaryotic microbial host cell for recombinant expression of a target protein, said cell comprising
- [0256] a. a gene encoding an enzyme having endoribonuclease activity (E.C. 3.1.26), said enzyme preferably having at least 75% amino acid sequence identity to SEQ ID NO. 2, wherein said gene is on the genome of said cell,
- [0257] b. a first recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 75% sequence identity with SEQ ID NO. 2, and wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substitutions at one or more positions selected from D346, E297, D303, N305, E325, R337, D349, V128, R169, T170, F57, F67, K112, G124, R141, R142, R373, and A441 relative to SEQ ID NO. 2, and

- [0258] c. a second recombinant gene encoding said target protein,
- [0259] wherein expression of said target protein is enhanced compared to a cell lacking said first recombinant gene.
- [0260] 2. A host cell according to item 1, wherein said mutant RNase E has one or more amino acid residue substitutions at one or more positions selected from D346, E297, D303, N305, E325, R337, and D349.
- [0261] 3. A host cell according to items 1 or 2, wherein said amino acid residue substitution at position 346 of said mutant RNase E relative to SEQ ID NO. 2 is asparagine.
- [0262] 4. A host cell according to any of items 1-3, wherein said cell further comprises a first prokaryotic vector, and wherein said first recombinant gene encoding said mutant RNase E is comprised on said first prokaryotic vector.
- [0263] 5. A host cell according to any of items 1-4, wherein said cell further comprises
- [0264] d. a gene encoding a T7 RNA polymerase (E.C. 2.7.7.6),
- [0265] e. optionally a gene encoding a T7 lysozyme (E.C. 3.5.1.28) and wherein expression of said second recombinant gene is regulated by an Inducible T7 promoter.
- [0266] 6. A host cell according to item 5, wherein said gene encoding said T7 lysozyme is located on the first prokaryotic vector, and wherein said second recombinant gene encoding said target gene is located on a second prokaryotic vector.
- [0267] 7. A host cell according to any one of items 1-4, wherein expression of said second recombinant gene is regulated by an Inducible promoter selected from rhabAD promoter, arabaAD promoter, Ptrc promoter, Ptet promoter, Ptac promoter, and PL promoter.
- [0268] 8. A host cell according to any one of items 1-7, wherein said target protein is a protein the expression of which is enhanced by at least 10% compared to expression of said protein in the same host cell lacking said first recombinant gene.
- [0269] 9. A host cell according to any one of items 1-8, wherein said cell is selected from *E. coli*, *Bacillus subtilis*, *Bacillus licheniformis*, and *Pseudomonas putida*.
- [0270] 10. A prokaryotic vector comprising
- [0271] a. a gene encoding a mutant RNase E (E.C. 3.1.26.12) having at least 75% amino acid sequence identity to SEQ ID NO. 2, wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substitutions at one or more posi-
- tions selected from D346, E297, D303, N305, E325, R337, D349, V128, R169, T170, F57, F67, K112, G124, R141, R142, R373, and A441 relative to SEQ ID NO. 2, and
- [0272] b. a gene encoding a T7 lysozyme (E.C. 3.5.1.28)
- [0273] 11. A prokaryotic vector according to item 10, wherein amino acid residue at position 346 of said mutant RNase E relative to SEQ ID NO. 2 is asparagine.
- [0274] 12. Use of the prokaryotic vector according to item 10 or 11 for enhancing expression of a recombinant gene encoding a target protein in a host prokaryotic microbial cell.
- [0275] 13. A method for the production of a target protein comprising culturing in a suitable culture medium, a host cell according to any one of items 1-9, optionally inducing expression of said target protein, followed by Isolation and purification of the expressed target protein.
- [0276] 14. A method for producing a prokaryotic recombinant microbial cell having enhanced expression of a recombinant target protein,
- [0277] a. providing a prokaryotic microbial cell comprising (I) a gene encoding an enzyme having endoribonuclease activity (E.C. 3.1.26), said enzyme preferably having at least 75% amino acid sequence identity to SEQ ID NO. 2, wherein said RNase gene is on the genome of said cell, and (ii) a gene encoding said recombinant target protein,
- [0278] b. transforming said microbial cell with a prokaryotic vector comprising a recombinant gene encoding a mutant RNase E (E.C. 3.1.26.12) having at least 75% amino acid sequence identity to SEQ ID NO. 2, and wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substitutions at one or more positions selected from D346, E297, D303, N305, E325, R337, D349, V128, R169, T170, F57, F67, K112, G124, R141, R142, R373, and A441 relative to SEQ ID NO. 2,
- [0279] wherein expression of said target protein is enhanced compared to a cell not transformed with said prokaryotic vector.
- [0280] 15. Method for producing a prokaryotic recombinant microbial cell having enhanced expression of a recombinant target protein according to item 14, wherein said microbial cell further comprises (III) a gene encoding a T7 RNA polymerase (E.C. 2.7.7.6), wherein expression of the said recombinant gene encoding said target protein is regulated by a T7 promoter, and wherein said microbial cell or said prokaryotic vector optionally comprises a gene encoding a T7 lysozyme (E.C. 3.5.1.28).

 SEQUENCE LISTING

```

<160> NUMBER OF SEQ ID NOS: 92

<210> SEQ ID NO 1
<211> LENGTH: 3186
<212> TYPE: DNA
<213> ORGANISM: E. coli
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)...(3186)
<223> OTHER INFORMATION: E. coli rne (wild type) encoding RNase E

<400> SEQUENCE: 1

```

-continued

atg aaa aga atg tta atc aac gca act cag cag gaa gag ttg cgc gtt Met Lys Arg Met Leu Ile Asn Ala Thr Gln Gln Glu Glu Leu Arg Val 1 5 10 15	48
gcc ctt gta gat ggg cag cgt ctg tat gac ctg gat atc gaa agt cca Ala Leu Val Asp Gly Gln Arg Leu Tyr Asp Leu Asp Ile Glu Ser Pro 20 25 30	96
ggg cac gag cag aaa aag gca aac atc tac aaa ggt aaa atc acc cgc Gly His Glu Gln Lys Lys Ala Asn Ile Tyr Lys Gly Lys Ile Thr Arg 35 40 45	144
att gaa ccg agt ctg gaa gct gct ttt gtt gat tac ggc gct gaa cgt Ile Glu Pro Ser Leu Glu Ala Ala Phe Val Asp Tyr Gly Ala Glu Arg 50 55 60	192
cac ggt ttc ctc cca cta aaa gaa att gcc cgc gaa tat ttc cct gct His Gly Phe Leu Pro Leu Lys Glu Ile Ala Arg Glu Tyr Phe Pro Ala 65 70 75 80	240
aac tac agt gct cat ggt cgt ccc aac att aaa gat gtg ttg cgt gaa Asn Tyr Ser Ala His Gly Arg Pro Asn Ile Lys Asp Val Leu Arg Glu 85 90 95	288
ggt cag gaa gtc att gtt cag atc gat aaa gaa gag cgc ggc aac aaa Gly Gln Glu Val Ile Val Gln Ile Asp Lys Glu Glu Arg Gly Asn Lys 100 105 110	336
ggc gcg gca tta acc acc ttt atc agt ctg gcg ggt agc tat ctg gtt Gly Ala Ala Leu Thr Thr Phe Ile Ser Leu Ala Gly Ser Tyr Leu Val 115 120 125	384
ctg atg ccg aac aac ccg cgc gcg ggt ggc att tct cgc cgt atc gaa Leu Met Pro Asn Asn Pro Arg Ala Gly Gly Ile Ser Arg Arg Ile Glu 130 135 140	432
ggc gac gac cgt acc gaa tta aaa gaa gca ctg gca agc ctt gaa ctg Gly Asp Asp Arg Thr Glu Leu Lys Glu Ala Leu Ala Ser Leu Glu Leu 145 150 155 160	480
ccg gaa ggc atg ggg ctt atc gtg cgc acc gct ggc gtc ggc aaa tct Pro Glu Gly Met Gly Leu Ile Val Arg Thr Ala Gly Val Gly Lys Ser 165 170 175	528
gct gag gcg ctg caa tgg gat tta agc ttc cgt ctg aaa cac tgg gaa Ala Glu Ala Leu Gln Trp Asp Leu Ser Phe Arg Leu Lys His Trp Glu 180 185 190	576
gcc atc aaa aaa gcc gct gaa agc cgc ccg gcc ttc ctg att cat Ala Ile Lys Lys Ala Ala Glu Ser Arg Pro Ala Pro Phe Leu Ile His 195 200 205	624
cag gag agc aac gta atc gtt cgc gca ttc cgc gat tac tta cgt cag Gln Glu Ser Asn Val Ile Val Arg Ala Phe Arg Asp Tyr Leu Arg Gln 210 215 220	672
gac atc ggc gaa atc ctt atc gat aac ccg aaa gtg ctc gaa ctg gca Asp Ile Gly Glu Ile Leu Ile Asp Asn Pro Lys Val Leu Glu Leu Ala 225 230 235 240	720
cgt cag cat atc gct gca tta ggt cgc ccg gat ttc agc agc aaa atc Arg Gln His Ile Ala Ala Leu Gly Arg Pro Asp Phe Ser Ser Lys Ile 245 250 255	768
aaa ctg tac acc ggc gag atc ccg ctg ttc agc cac tac cag atc gag Lys Leu Tyr Thr Gly Glu Ile Pro Leu Phe Ser His Tyr Gln Ile Glu 260 265 270	816
tca cag atc gag tcc gcc ttc cag cgt gaa gtt cgt ctg ccg tct ggt Ser Gln Ile Glu Ser Ala Phe Gln Arg Glu Val Arg Leu Pro Ser Gly 275 280 285	864
ggg tcc att gtt atc gac agc acc gaa ggc tta acg gcc atc gac atc Gly Ser Ile Val Ile Asp Ser Thr Glu Ala Leu Ile Asp Ile 290 295 300	912

-continued

aac tcc gca cgc gcg acc cgc ggc gat atc gaa gaa acc gcg ttt Asn Ser Ala Arg Ala Thr Arg Gly Gly Asp Ile Glu Glu Thr Ala Phe 305 310 315 320	960
aac act aac ctc gaa gct gcc gat gag att gct cgt cag ctg cgc ctg Asn Thr Asn Leu Glu Ala Ala Asp Glu Ile Ala Arg Gln Leu Arg Leu 325 330 335	1008
cgt gac ctc ggc ggc ctg att gtt atc gac ttc atc gac atg acg cca Arg Asp Leu Gly Gly Leu Ile Val Ile Asp Phe Ile Asp Met Thr Pro 340 345 350	1056
gta cgc cac cag cgt gcg gta gaa aac cgt ctg cgt gaa gcg gtg cgt Val Arg His Gln Arg Ala Val Glu Asn Arg Leu Arg Glu Ala Val Arg 355 360 365	1104
cag gac cgt gcg cgt att caa atc agc cat att tct cgc ttt ggc ctg Gln Asp Arg Ala Arg Ile Gln Ile Ser His Ile Ser Arg Phe Gly Leu 370 375 380	1152
ctg gaa atg tcc cgt cag cgc ctg agc cca tca ctg ggt gaa tcc agc Leu Glu Met Ser Arg Gln Arg Leu Ser Pro Ser Leu Gly Glu Ser Ser 385 390 395 400	1200
cat cac gtc tgc ccg cgc tcc ggt acc ggt acc gtg cgt gac aac His His Val Cys Pro Arg Cys Ser Gly Thr Gly Thr Val Arg Asp Asn 405 410 415	1248
gaa tcg ctg tcg ctc tct att ctg cgt ctg atc gaa gaa gaa gcg ctg Glu Ser Leu Ser Leu Ser Ile Leu Arg Leu Ile Glu Glu Ala Leu 420 425 430	1296
aaa gag aac acc cag gaa gtt cac gcc att gtt cct gtg cca atc gct Lys Glu Asn Thr Gln Glu Val His Ala Ile Val Pro Val Pro Ile Ala 435 440 445	1344
tct tat ctg ctg aat gaa aaa cgt tct gcg gtg aat gcc att gaa acg Ser Tyr Leu Leu Asn Glu Lys Arg Ser Ala Val Asn Ala Ile Glu Thr 450 455 460	1392
cgt cag gac ggt gtt cgc tgc gtg att gtg cca aac gat cag atg gaa Arg Gln Asp Gly Val Arg Cys Val Ile Val Pro Asn Asp Gln Met Glu 465 470 475 480	1440
acc ccg cac tac cac gtg ctg cgc gtg cgt aaa ggg gaa gaa acg ccg Thr Pro His Tyr His Val Leu Arg Val Arg Lys Gly Glu Glu Thr Pro 485 490 495	1488
acc tta agc tac atg ctg ccg aag ctg cat gaa gaa gcg atg gcg ctg Thr Leu Ser Tyr Met Leu Pro Lys Leu His Glu Glu Ala Met Ala Leu 500 505 510	1536
ccg tct gaa gaa gag ttc gct gaa cgt aag cgt ccg gaa caa cct gcg Pro Ser Glu Glu Glu Phe Ala Glu Arg Lys Arg Pro Glu Gln Pro Ala 515 520 525	1584
ctg gca acc ttt gcc atg ccg gat gtg ccg cca gcg cca acc cca gct Leu Ala Thr Phe Ala Met Pro Asp Val Pro Pro Ala Pro Thr Pro Ala 530 535 540	1632
gaa cct gcc gcg cct gtc gta gcc cca gca cct aaa tct gca ccg gca Glu Pro Ala Ala Pro Val Val Ala Pro Ala Pro Lys Ser Ala Pro Ala 545 550 555 560	1680
aca cca gcc gct cct gcc caa cct ggg ctg ttg agc cgc ttc ttc ggc Thr Pro Ala Ala Pro Ala Gln Pro Gly Leu Leu Ser Arg Phe Phe Gly 565 570 575	1728
gca ctg aaa gcg ctg ttc agc ggt ggt gaa gaa acc aaa ccg tcc gag Ala Leu Lys Ala Leu Phe Ser Gly Gly Glu Glu Thr Lys Pro Ser Glu 580 585 590	1776
caa cca aca ccg aaa gca gaa gcg aaa ccg gaa cgt caa cag gat cgt Gln Pro Thr Pro Lys Ala Glu Ala Lys Pro Glu Arg Gln Asp Arg 595 600 605	1824

-continued

cgc aag cct cgt cag aac aac cgc cgt gac cgt aat gag cgc cgc gac Arg Lys Pro Arg Gln Asn Asn Arg Arg Asp Arg Asn Glu Arg Arg Asp 610 615 620	1872
acc cgt agt gaa cgt act gaa ggc agc gat aat cgc gaa gaa aac cgt Thr Arg Ser Glu Arg Thr Glu Gly Ser Asp Asn Arg Glu Glu Asn Arg 625 630 635 640	1920
cgt aat cgt cgc cag gca cag cag cag act gcc gag acg cgt gag agc Arg Asn Arg Arg Gln Ala Gln Gln Thr Ala Glu Thr Arg Glu Ser 645 650 655	1968
cgt cag cag gtt gag gta acg gaa aaa gcg cgt acc acc gac gag cag Arg Gln Gln Val Glu Val Thr Glu Lys Ala Arg Thr Thr Asp Glu Gln 660 665 670	2016
caa gcg ccg cgt cgt gaa cgt agc cgc cgc cgt aat gat gat aaa cgt Gln Ala Pro Arg Arg Glu Arg Ser Arg Arg Asn Asp Asp Lys Arg 675 680 685	2064
cag gcg caa caa gaa gcg aag gcg ctg aat gtt gaa gag caa tct gtt Gln Ala Gln Gln Glu Ala Lys Ala Leu Asn Val Glu Glu Gln Ser Val 690 695 700	2112
cag gaa acc gaa cag gaa cgt gta cgt ccg gtt cag ccg cgt cgt Gln Glu Thr Glu Gln Glu Arg Val Arg Pro Val Gln Pro Arg Arg 705 710 715 720	2160
aaa cag cgt cag ctc aat cag aaa gtg cgt tac gag caa agc gta gcc Lys Gln Arg Gln Leu Asn Gln Lys Val Arg Tyr Glu Gln Ser Val Ala 725 730 735	2208
gaa gaa gcg gta gtc gca ccg gtg gtt gaa gaa act gtc gct gcc gaa Glu Glu Ala Val Val Ala Pro Val Val Glu Glu Thr Val Ala Ala Glu 740 745 750	2256
cca att gtt cag gaa gcg cca gct cca cgc aca gaa ctg gtg aaa gtc Pro Ile Val Gln Glu Ala Pro Ala Pro Arg Thr Glu Leu Val Lys Val 755 760 765	2304
ccg ctg cca gtc gta gcg caa act gca cca gaa cag caa gaa gag aac Pro Leu Pro Val Val Ala Gln Thr Ala Pro Glu Gln Gln Glu Glu Asn 770 775 780	2352
aat gct gat aac cgt gac aac ggt ggc atg ccg cgt cgt tct cgc cgc Asn Ala Asp Asn Arg Asp Asn Gly Gly Met Pro Arg Arg Ser Arg Arg 785 790 795 800	2400
tcg cct cgt cac ctg cgc gta agt ggt cag cgt cgt cgt cgc tat cgt Ser Pro Arg His Leu Arg Val Ser Gly Gln Arg Arg Arg Tyr Arg 805 810 815	2448
gac gag cgt tat cca acc cag tcg cca atg ccg ttg acc gta gcg tgc Asp Glu Arg Tyr Pro Thr Gln Ser Pro Met Pro Leu Thr Val Ala Cys 820 825 830	2496
gcg tct ccg gaa ctg gcc tct ggc aaa gtc tgg atc cgc tat cca att Ala Ser Pro Glu Leu Ala Ser Gly Lys Val Trp Ile Arg Tyr Pro Ile 835 840 845	2544
gta cgt ccg caa gat gta cag gtt gaa gag cag cgc gaa cag gaa gaa Val Arg Pro Gln Asp Val Gln Val Glu Glu Gln Arg Glu Gln Glu Glu 850 855 860	2592
gta cat gtg cag ccg atg gtg act gag gtc cct gtc gcc gcc gct atc Val His Val Gln Pro Met Val Thr Glu Val Pro Val Ala Ala Ala Ile 865 870 875 880	2640
gaa ccg gtt gtt agc gcg cca gtt gtt gaa gaa gtg gcc ggt gtc gta Glu Pro Val Val Ser Ala Pro Val Val Glu Glu Val Ala Gly Val Val 885 890 895	2688
gaa gcc ccc gtt cag gtt gcc gaa ccg caa ccg gaa gtg gtt gaa acg Glu Ala Pro Val Gln Val Ala Glu Pro Gln Pro Glu Val Val Glu Thr 900 905 910	2736

-continued

acg cat cct gaa gtg att gct gcc gcg gta act gaa cag ccg cag gtg	2784
Thr His Pro Glu Val Ile Ala Ala Ala Val Thr Glu Gln Pro Gln Val	
915 920 925	
att acc gag tct gat gtt gcc gta gcc cag gaa gtt gca gaa caa gca	2832
Ile Thr Glu Ser Asp Val Ala Val Ala Gln Glu Val Ala Glu Gln Ala	
930 935 940	
gaa ccg gtg gtt gaa ccg cag gaa gag acg gca gat att gaa gaa gtt	2880
Glu Pro Val Val Glu Pro Gln Glu Glu Thr Ala Asp Ile Glu Glu Val	
945 950 955 960	
gtc gaa act gct gag gtt gta gtt gct gaa cct gaa gtt gtc gct caa	2928
Val Glu Thr Ala Glu Val Val Ala Glu Pro Glu Val Val Ala Gln	
965 970 975	
cct gcc gcg cca gta gtc gct gaa gtc gca gca gaa gtt gaa acg gta	2976
Pro Ala Ala Pro Val Val Ala Glu Val Ala Ala Glu Val Glu Thr Val	
980 985 990	
gct gcg gtc gaa cct gag gtc acc gtt gag cat aac cac gct acc gcg	3024
Ala Ala Val Glu Pro Glu Val Thr Val Glu His Asn His Ala Thr Ala	
995 1000 1005	
cca atg acg cgc gct cca gca ccg gaa tat gtt ccg gag gca ccg	3069
Pro Met Thr Arg Ala Pro Ala Pro Glu Tyr Val Pro Glu Ala Pro	
1010 1015 1020	
cgt cac agt gac tgg cag cgc cct act ttt gcc ttc gaa ggt aaa	3114
Arg His Ser Asp Trp Gln Arg Pro Thr Phe Ala Phe Glu Gly Lys	
1025 1030 1035	
ggt gcc gca ggt ggt cat acg gca aca cat cat gcc tct gct gct	3159
Gly Ala Ala Gly Gly His Thr Ala Thr His His Ala Ser Ala Ala	
1040 1045 1050	
cct gcg cgt ccg caa cct gtt gag taa	3186
Pro Ala Arg Pro Gln Pro Val Glu	
1055 1060	
<210> SEQ ID NO 2	
<211> LENGTH: 1061	
<212> TYPE: PRT	
<213> ORGANISM: E. coli	
<400> SEQUENCE: 2	
Met Lys Arg Met Leu Ile Asn Ala Thr Gln Gln Glu Glu Leu Arg Val	
1 5 10 15	
Ala Leu Val Asp Gly Gln Arg Leu Tyr Asp Leu Asp Ile Glu Ser Pro	
20 25 30	
Gly His Glu Gln Lys Lys Ala Asn Ile Tyr Lys Gly Lys Ile Thr Arg	
35 40 45	
Ile Glu Pro Ser Leu Glu Ala Ala Phe Val Asp Tyr Gly Ala Glu Arg	
50 55 60	
His Gly Phe Leu Pro Leu Lys Glu Ile Ala Arg Glu Tyr Phe Pro Ala	
65 70 75 80	
Asn Tyr Ser Ala His Gly Arg Pro Asn Ile Lys Asp Val Leu Arg Glu	
85 90 95	
Gly Gln Glu Val Ile Val Gln Ile Asp Lys Glu Glu Arg Gly Asn Lys	
100 105 110	
Gly Ala Ala Leu Thr Thr Phe Ile Ser Leu Ala Gly Ser Tyr Leu Val	
115 120 125	
Leu Met Pro Asn Asn Pro Arg Ala Gly Gly Ile Ser Arg Arg Ile Glu	
130 135 140	

-continued

Gly	Asp	Asp	Arg	Thr	Glu	Leu	Lys	Glu	Ala	Leu	Ala	Ser	Leu	Glu	Leu
145					150				155						160
Pro	Glu	Gly	Met	Gly	Leu	Ile	Val	Arg	Thr	Ala	Gly	Val	Gly	Lys	Ser
	165					170									175
Ala	Glu	Ala	Leu	Gln	Trp	Asp	Leu	Ser	Phe	Arg	Leu	Lys	His	Trp	Glu
	180						185								190
Ala	Ile	Lys	Lys	Ala	Ala	Glu	Ser	Arg	Pro	Ala	Pro	Phe	Leu	Ile	His
	195					200									205
Gln	Glu	Ser	Asn	Val	Ile	Val	Arg	Ala	Phe	Arg	Asp	Tyr	Leu	Arg	Gln
	210					215									220
Asp	Ile	Gly	Glu	Ile	Leu	Ile	Asp	Asn	Pro	Lys	Val	Leu	Glu	Leu	Ala
	225					230				235					240
Arg	Gln	His	Ile	Ala	Ala	Leu	Gly	Arg	Pro	Asp	Phe	Ser	Ser	Lys	Ile
	245					250				255					255
Lys	Leu	Tyr	Thr	Gly	Glu	Ile	Pro	Leu	Phe	Ser	His	Tyr	Gln	Ile	Glu
	260					265									270
Ser	Gln	Ile	Glu	Ser	Ala	Phe	Gln	Arg	Glu	Val	Arg	Leu	Pro	Ser	Gly
	275					280				285					
Gly	Ser	Ile	Val	Ile	Asp	Ser	Thr	Glu	Ala	Leu	Thr	Ala	Ile	Asp	Ile
	290					295				300					
Asn	Ser	Ala	Arg	Ala	Thr	Arg	Gly	Gly	Asp	Ile	Glu	Glu	Thr	Ala	Phe
	305					310				315					320
Asn	Thr	Asn	Leu	Glu	Ala	Ala	Asp	Glu	Ile	Ala	Arg	Gln	Leu	Arg	Leu
	325					330				335					
Arg	Asp	Leu	Gly	Gly	Leu	Ile	Val	Ile	Asp	Phe	Ile	Asp	Met	Thr	Pro
	340					345				350					
Val	Arg	His	Gln	Arg	Ala	Val	Glu	Asn	Arg	Leu	Arg	Glu	Ala	Val	Arg
	355					360				365					
Gln	Asp	Arg	Ala	Arg	Ile	Gln	Ile	Ser	His	Ile	Ser	Arg	Phe	Gly	Leu
	370					375				380					
Leu	Glu	Met	Ser	Arg	Gln	Arg	Leu	Ser	Pro	Ser	Leu	Gly	Glu	Ser	Ser
	385					390				395					400
His	His	Val	Cys	Pro	Arg	Cys	Ser	Gly	Thr	Gly	Thr	Val	Arg	Asp	Asn
	405					410				415					
Glu	Ser	Leu	Ser	Leu	Ser	Ile	Leu	Arg	Leu	Ile	Glu	Glu	Glu	Ala	Leu
	420					425				430					
Lys	Glu	Asn	Thr	Gln	Glu	Val	His	Ala	Ile	Val	Pro	Val	Pro	Ile	Ala
	435					440				445					
Ser	Tyr	Leu	Leu	Asn	Glu	Lys	Arg	Ser	Ala	Val	Asn	Ala	Ile	Glu	Thr
	450					455				460					
Arg	Gln	Asp	Gly	Val	Arg	Cys	Val	Ile	Val	Pro	Asn	Asp	Gln	Met	Glu
	465					470				475					480
Thr	Pro	His	Tyr	His	Val	Leu	Arg	Val	Lys	Gly	Glu	Glu	Thr	Pro	
	485					490				495					
Thr	Leu	Ser	Tyr	Met	Leu	Pro	Lys	Leu	His	Glu	Glu	Ala	Met	Ala	Leu
	500					505				510					
Pro	Ser	Glu	Glu	Glu	Phe	Ala	Glu	Arg	Lys	Arg	Pro	Glu	Gln	Pro	Ala
	515					520				525					
Leu	Ala	Thr	Phe	Ala	Met	Pro	Asp	Val	Pro	Pro	Ala	Pro	Thr	Pro	Ala
	530					535				540					
Glu	Pro	Ala	Ala	Pro	Val	Val	Ala	Pro	Ala	Pro	Lys	Ser	Ala	Pro	Ala

-continued

545	550	555	560
Thr Pro Ala Ala Pro Ala Gln Pro Gly Leu Leu Ser Arg Phe Phe Gly			
565	570	575	
Ala Leu Lys Ala Leu Phe Ser Gly Gly Glu Glu Thr Lys Pro Ser Glu			
580	585	590	
Gln Pro Thr Pro Lys Ala Glu Ala Lys Pro Glu Arg Gln Gln Asp Arg			
595	600	605	
Arg Lys Pro Arg Gln Asn Asn Arg Arg Asp Arg Asn Glu Arg Arg Asp			
610	615	620	
Thr Arg Ser Glu Arg Thr Glu Gly Ser Asp Asn Arg Glu Glu Asn Arg			
625	630	635	640
Arg Asn Arg Arg Gln Ala Gln Gln Thr Ala Glu Thr Arg Glu Ser			
645	650	655	
Arg Gln Gln Val Glu Val Thr Glu Lys Ala Arg Thr Thr Asp Glu Gln			
660	665	670	
Gln Ala Pro Arg Arg Glu Arg Ser Arg Arg Asn Asp Asp Lys Arg			
675	680	685	
Gln Ala Gln Gln Glu Ala Lys Ala Leu Asn Val Glu Glu Gln Ser Val			
690	695	700	
Gln Glu Thr Glu Gln Glu Arg Val Arg Pro Val Gln Pro Arg Arg			
705	710	715	720
Lys Gln Arg Gln Leu Asn Gln Lys Val Arg Tyr Glu Gln Ser Val Ala			
725	730	735	
Glu Glu Ala Val Val Ala Pro Val Val Glu Glu Thr Val Ala Ala Glu			
740	745	750	
Pro Ile Val Gln Glu Ala Pro Ala Pro Arg Thr Glu Leu Val Lys Val			
755	760	765	
Pro Leu Pro Val Val Ala Gln Thr Ala Pro Glu Gln Gln Glu Glu Asn			
770	775	780	
Asn Ala Asp Asn Arg Asp Asn Gly Gly Met Pro Arg Arg Ser Arg Arg			
785	790	795	800
Ser Pro Arg His Leu Arg Val Ser Gly Gln Arg Arg Arg Arg Tyr Arg			
805	810	815	
Asp Glu Arg Tyr Pro Thr Gln Ser Pro Met Pro Leu Thr Val Ala Cys			
820	825	830	
Ala Ser Pro Glu Leu Ala Ser Gly Lys Val Trp Ile Arg Tyr Pro Ile			
835	840	845	
Val Arg Pro Gln Asp Val Gln Val Glu Glu Gln Arg Glu Gln Glu Glu			
850	855	860	
Val His Val Gln Pro Met Val Thr Glu Val Pro Val Ala Ala Ala Ile			
865	870	875	880
Glu Pro Val Val Ser Ala Pro Val Val Glu Glu Val Ala Gly Val Val			
885	890	895	
Glu Ala Pro Val Gln Val Ala Glu Pro Gln Pro Glu Val Val Glu Thr			
900	905	910	
Thr His Pro Glu Val Ile Ala Ala Val Thr Glu Gln Pro Gln Val			
915	920	925	
Ile Thr Glu Ser Asp Val Ala Val Ala Gln Glu Val Ala Glu Gln Ala			
930	935	940	
Glu Pro Val Val Glu Pro Gln Glu Glu Thr Ala Asp Ile Glu Glu Val			
945	950	955	960

-continued

```

Val Glu Thr Ala Glu Val Val Val Ala Glu Pro Glu Val Val Ala Gln
965          970          975

Pro Ala Ala Pro Val Val Ala Glu Val Ala Ala Glu Val Glu Thr Val
980          985          990

Ala Ala Val Glu Pro Glu Val Thr Val Glu His Asn His Ala Thr Ala
995          1000         1005

Pro Met Thr Arg Ala Pro Ala Pro Glu Tyr Val Pro Glu Ala Pro
1010         1015         1020

Arg His Ser Asp Trp Gln Arg Pro Thr Phe Ala Phe Glu Gly Lys
1025         1030         1035

Gly Ala Ala Gly Gly His Thr Ala Thr His His Ala Ser Ala Ala
1040         1045         1050

Pro Ala Arg Pro Gln Pro Val Glu
1055         1060

```

```

<210> SEQ ID NO 3
<211> LENGTH: 3240
<212> TYPE: DNA
<213> ORGANISM: Pseudomonas putida
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)..(3240)
<223> OTHER INFORMATION: P. putida rne (wild type) encoding RNase E

```

<400> SEQUENCE: 3

atg aaa aga atg ctg att aac gca act caa ccc gaa gag ttg cgt gta	48
Met Lys Arg Met Leu Ile Asn Ala Thr Gln Pro Glu Glu Leu Arg Val	
1 5 10 15	
 gcc ctg gtg gac ggc caa cgc ctc tac gac ctg gac atc gag tcc ggc	96
Ala Leu Val Asp Gly Gln Arg Leu Tyr Asp Leu Asp Ile Glu Ser Gly	
20 25 30	
 gca cgt gag cag aaa aag gcc aac atc tac aaa ggc aag atc acc cgc	144
Ala Arg Glu Gln Lys Lys Ala Asn Ile Tyr Lys Gly Lys Ile Thr Arg	
35 40 45	
 atc gaa ccc agc ctc gaa gcc gcc ttc gtc gac ttc ggt tcc gaa cgt	192
Ile Glu Pro Ser Leu Glu Ala Ala Phe Val Asp Phe Gly Ser Glu Arg	
50 55 60	
 cac ggc ttc ctg ccg ctg aaa gaa atc tcc cgc gaa tac ttc aag aaa	240
His Gly Phe Leu Pro Leu Lys Glu Ile Ser Arg Glu Tyr Phe Lys Lys	
65 70 75 80	
 gcc ccc gaa ggc cgg gtg aac atc aag gaa gtg ctc agc gaa ggc cag	288
Ala Pro Glu Gly Arg Val Asn Ile Lys Glu Val Leu Ser Glu Gly Gln	
85 90 95	
 gaa gtc atc gtc cag gtc gag aag gaa gag cgc ggc aac aaa ggc gcc	336
Glu Val Ile Val Gln Val Glu Lys Glu Glu Arg Gly Asn Lys Gly Ala	
100 105 110	
 gcc ctc acc acc ttc atc agc ctg gct ggc cgc tac ctg gtg ctg atg	384
Ala Leu Thr Thr Phe Ile Ser Leu Ala Gly Arg Tyr Leu Val Leu Met	
115 120 125	
 ccc aac aac ccg cgt gct ggc ggc atc tcc cgc cgc atc gaa ggc gaa	432
Pro Asn Asn Pro Arg Ala Gly Gly Ile Ser Arg Arg Ile Glu Gly Glu	
130 135 140	
 gag cgc aac gaa ctg cgc gaa gcc ctg aac ggc ctg acc gtg ccg ggc	480
Glu Arg Asn Glu Leu Arg Glu Ala Leu Asn Gly Leu Thr Val Pro Gly	
145 150 155 160	
 gac atg ggc ctg atc gtg cgc act gcc ggc ctt ggc cgc agc agc gaa	528
Asp Met Gly Leu Ile Val Arg Thr Ala Gly Leu Gly Arg Ser Ser Glu	

-continued

165	170	175	
gaa atg cag tgg gac ctc gac tac ctg ctg cag ctg tgg acc gcc atc Glu Met Gln Trp Asp Leu Asp Tyr Leu Leu Gln Leu Trp Thr Ala Ile 180 185 190			576
aag gaa gca tcc ctg gac cgc gcc gcg cca ttc ctg atc tac cag gaa Lys Glu Ala Ser Leu Asp Arg Ala Ala Pro Phe Leu Ile Tyr Gln Glu 195 200 205			624
agc aac gtc atc atc cgc gcc atc cgc gac tac ctg cgc cag gac atc Ser Asn Val Ile Ile Arg Ala Ile Arg Asp Tyr Leu Arg Gln Asp Ile 210 215 220			672
ggt gaa gtg ctg atc gac agc atc gac gcc cag gaa gag gcc ctg acc Gly Glu Val Leu Ile Asp Ser Ile Asp Ala Gln Glu Glu Ala Leu Thr 225 230 235 240			720
tta atc cgc cag gtg atg ccg cag tac gac gac aag gtg aaa ctg tac Phe Ile Arg Gln Val Met Pro Gln Tyr Ala Ser Lys Val Lys Leu Tyr 245 250 255			768
gaa gac agc gta ccg ctg ttc aac cgc ttc cag atc gaa agc cag atc Glu Asp Ser Val Pro Leu Phe Asn Arg Phe Gln Ile Glu Ser Gln Ile 260 265 270			816
gaa acc gcc ttc cag cgc gtg gtc gac ctg ccg tcc ggt ggt tcg atc Glu Thr Ala Phe Gln Arg Val Val Asp Leu Pro Ser Gly Gly Ser Ile 275 280 285			864
gtg atc gac ccg acc gaa gcc ctg gtg tct atc gac atc aac tcg gcg Val Ile Asp Pro Thr Glu Ala Leu Val Ser Ile Asp Ile Asn Ser Ala 290 295 300			912
cgc gcc acc aaa ggc agc gat atc gaa gaa acc gcc ctg cag acc aac Arg Ala Thr Lys Gly Ser Asp Ile Glu Glu Thr Ala Leu Gln Thr Asn 305 310 315 320			960
ctg gaa gcg gag gaa atc gcc cgc cag ctg cgc ctg cgt gac atc Leu Glu Ala Ala Glu Glu Ile Ala Arg Gln Leu Arg Leu Arg Asp Ile 325 330 335			1008
ggc ggc ctg atc gtg atc gac ttc atc gac atg acc ccg gcg aaa aac Gly Gly Leu Ile Val Ile Asp Phe Ile Asp Met Thr Pro Ala Lys Asn 340 345 350			1056
cag cgc gcc gtt gaa gaa cgc gtg cgc gaa tgc ctg gaa gcg gac cgt Gln Arg Ala Val Glu Glu Arg Val Arg Glu Cys Leu Glu Ala Asp Arg 355 360 365			1104
gcc cgc gtg cag gtt ggc cgc atc tcg cgc ttc ggc ctg ctg gaa atg Ala Arg Val Gln Val Gly Arg Ile Ser Arg Phe Gly Leu Leu Glu Met 370 375 380			1152
tcc cgt cag cgc ctg cgc cca tcg ctg ggc gaa agc agc ggc atc gtc Ser Arg Gln Arg Leu Arg Pro Ser Leu Gly Glu Ser Ser Gly Ile Val 385 390 395 400			1200
tgc cca cgc tgc tcc ggc acc ggc atc atc cgt gac gtg gag tcg ctg Cys Pro Arg Cys Ser Gly Thr Gly Ile Ile Arg Asp Val Glu Ser Leu 405 410 415			1248
tcg ctg gcc atc ctg cgc ctg atc gaa gaa gaa gcc ctg aag gac cgc Ser Leu Ala Ile Leu Arg Leu Ile Glu Glu Ala Leu Lys Asp Arg 420 425 430			1296
act gcc gaa gtc cgt gca caa gtg cca atc ccg gtg gcc gcg ttc ctg Thr Ala Glu Val Arg Ala Gln Val Pro Ile Pro Val Ala Ala Phe Leu 435 440 445			1344
ctc aac gag aag cgc aac tcg atc acc aag atc gaa ctg cgc acc cgt Leu Asn Glu Lys Arg Asn Ser Ile Thr Lys Ile Glu Leu Arg Thr Arg 450 455 460			1392
gcg cgc atc atc atc ctg ccg aac gat cac ctg gaa act ccg cac ttc Ala Arg Ile Ile Ile Leu Pro Asn Asp His Leu Glu Thr Pro His Phe			1440

-continued

465	470	475	480		
gaa gtc cag cgc ctg cgc gac gac aac ccg gaa gtg ctg aac aac cag Glu Val Gln Arg Leu Arg Asp Asp Asn Pro Glu Val Leu Asn Asn Gln	485	490	495	1488	
tcc agc tac gag atc gcc acc acc gaa gcc gaa gag gca ccg cag cag Ser Ser Tyr Glu Ile Ala Thr Thr Glu Ala Glu Glu Ala Pro Gln Gln	500	505	510	1536	
acc gcc acc cgc acc ctg gtt cgc cag gaa gca gcg gtc aag acc gcc Thr Ala Thr Arg Thr Leu Val Arg Gln Glu Ala Ala Val Lys Thr Ala	515	520	525	1584	
ccg gcc cgc acc aac gcg ccg gta ccg gtt gct gaa gag cct cag gct Pro Ala Arg Thr Asn Ala Pro Val Pro Val Ala Glu Glu Pro Gln Ala	530	535	540	1632	
gtt gcc ccg gcc ccg gcc ccg agc gcc ccg gag cca agc ctg ttc Val Ala Pro Ala Ala Pro Ala Pro Ser Ala Pro Glu Pro Ser Leu Phe	545	550	555	560	1680
aaa ggc ctg gtg aag tcg ctg gtc agc ctg ttc gcc ggc aag gac gaa Lys Gly Leu Val Lys Ser Leu Val Ser Leu Phe Ala Gly Lys Asp Glu	565	570	575	1728	
cct gca gct gca cct gtc gtc gcc gag aaa ccg gcc gag cgc Pro Ala Ala Ala Pro Val Val Ala Glu Lys Pro Ala Ala Glu Arg	580	585	590	1776	
agc ccg cgc aac gag gag cgc cgc aac ggc cgt caa cag agc cgc aac Ser Pro Arg Asn Glu Glu Arg Arg Asn Gly Arg Gln Gln Ser Arg Asn	595	600	605	1824	
ccg aac ggc cgc cgc gac gaa gag cgc aag ccg cgt gaa gag cgt gcc Arg Asn Gly Arg Arg Asp Glu Glu Arg Lys Pro Arg Glu Glu Arg Ala	610	615	620	1872	
gag cgc gcc ccg cgc gaa gag cgc cag cca cgc gag gag cgt gca ccg Glu Arg Ala Pro Arg Glu Glu Arg Gln Pro Arg Glu Glu Arg Ala Pro	625	630	635	640	1920
cgc gaa gag cgt gca cca cgc gaa gag cgc gcc cca cgc gaa gag cgt Arg Glu Glu Arg Ala Pro Arg Glu Glu Arg Ala Pro Arg Glu Glu Arg	645	650	655	1968	
gca cca cgc cag cca cgc gaa gac cgc cgc agc aac cgc ggc gaa gag Ala Pro Arg Gln Pro Arg Glu Asp Arg Arg Ser Asn Arg Gly Glu Glu	660	665	670	2016	
cgc gtg cgc gaa ctg cgt gag ccg ctg gat gcc acc ccg cca gcc gaa Arg Val Arg Glu Leu Arg Glu Pro Leu Asp Ala Thr Pro Pro Ala Glu	675	680	685	2064	
cgc gaa gag cgc cag cca cgt gaa gag cgt gta gcc cgt gaa gaa cgc Arg Glu Glu Arg Gln Pro Arg Glu Glu Arg Val Ala Arg Glu Glu Arg	690	695	700	2112	
gcc cca cgt gaa gag cgt gca cct cgc gaa gaa cgt gct ccg cgt gaa Ala Pro Arg Glu Glu Arg Ala Pro Arg Glu Glu Arg Ala Pro Arg Glu	705	710	715	720	2160
gag cgc gca cct cgc gaa gaa cgt gct ccg cgt gaa gag cgt gct cct Glu Arg Ala Pro Arg Glu Glu Arg Ala Pro Arg Glu Glu Arg Ala Pro	725	730	735	2208	
cgc gaa gaa cgt gct ccg cgt gaa gag cgc gca cct cgc gaa gaa cgt Arg Glu Glu Arg Ala Pro Arg Glu Glu Arg Ala Pro Arg Glu Glu Arg	740	745	750	2256	
ccg cca cgc gaa gaa cgc gcc ccg cgc cca cca cgc gaa gag cgc cag Ala Pro Arg Glu Glu Arg Ala Pro Arg Pro Pro Arg Glu Glu Arg Gln	755	760	765	2304	
cca cgc gta gcc gaa gaa gcg gcc gag cag gct gcc gaa ctg gcc gaa Pro Arg Val Ala Glu Glu Ala Ala Glu Gln Ala Ala Glu Leu Ala Glu				2352	

-continued

770	775	780	
gag caa ctg ccg aac gaa gag ctg ctg caa gac gaa cag gaa ggc acc Glu Gln Leu Pro Asn Glu Glu Leu Leu Gln Asp Glu Gln Glu Gly Thr 785 790 795 800			2400
gat ggc gag cgt ccg cgc cgc tcc cgt ggc cag cgt cgt cgc agc Asp Gly Glu Arg Pro Arg Arg Ser Arg Gly Gln Arg Arg Arg Ser 805 810 815			2448
aac cgt cgt gag cgc cag cgc aac gcc aat ggc gag ctg atc gac ggt Asn Arg Arg Glu Arg Gln Arg Asn Ala Asn Gly Glu Leu Ile Asp Gly 820 825 830			2496
ggc gaa gag gaa ggc agc gaa gag cag cca caa cag cac cag gcc Gly Glu Glu Glu Gly Ser Glu Glu Gln Pro Gln Gln His Gln Ala 835 840 845			2544
acc gag ctg ggt gcc gaa ctg gcc ggc ctg gca gtg act gcc gct Thr Glu Leu Gly Ala Glu Leu Ala Ala Gly Leu Ala Val Thr Ala Ala 850 855 860			2592
gtt gcc agc agc aac atc agc gcc gac gcc gaa gcc cag gcc aac cag Val Ala Ser Ser Asn Ile Ser Ala Asp Ala Glu Ala Gln Ala Asn Gln 865 870 875 880			2640
cag gcc gaa ctg gcc acc gcc gaa atc gct gcc gca gca gag acc gac Gln Ala Glu Leu Ala Thr Ala Glu Ile Ala Ala Ala Ala Glu Thr Asp 885 890 895			2688
aac agc cac gcc gct cag ccg gtc gag aag gct gaa aag gtc gag ccg Asn Ser His Ala Ala Gln Pro Val Glu Lys Ala Glu Lys Val Glu Pro 900 905 910			2736
gtt gaa gcc gtc gcc aag gct gaa gac gtt gcc gta gcc cca gtg gtg Val Glu Ala Val Ala Lys Ala Glu Asp Val Ala Val Ala Pro Val Val 915 920 925			2784
gag cag cct gtc agc gag ccg gtt gtc gtg gcc gag gtc act gcc gag Glu Gln Pro Val Ser Glu Pro Val Val Val Ala Glu Val Thr Ala Glu 930 935 940			2832
ccg gta gtc gaa gtc gct ccg caa cac ctg gtt gaa gaa ggc cct gct Pro Val Val Glu Val Ala Pro Gln His Leu Val Glu Glu Ala Pro Ala 945 950 955 960			2880
gcc gaa ccg gta atc gtc gcc gaa gcc cct gtt gaa acg cct gct gtc Ala Glu Pro Val Ile Val Ala Glu Ala Pro Val Glu Thr Pro Ala Val 965 970 975			2928
gaa gcg ggc gaa atc gaa aaa gct ccg gcc gtg gtt gaa acc gct ccg Glu Ala Gly Ile Glu Lys Ala Pro Ala Val Val Glu Thr Ala Pro 980 985 990			2976
gtt gcc gag cag cct gca ccg gtt gtc gaa gcc cag cca gaa gta gtc Val Ala Glu Gln Pro Ala Pro Val Val Glu Ala Gln Pro Glu Val Val 995 1000 1005			3024
gca gag cct gca cct gtc gtc gag cct gcc ccg gta gag gca Ala Glu Pro Ala Pro Val Val Glu Pro Ala Pro Val Glu Ala 1010 1015 1020			3069
gaa gct gcc acc gtc atg ctg gcc aac ggc cgt gcg ccg aac gac Glu Ala Ala Thr Val Met Leu Ala Asn Gly Arg Ala Pro Asn Asp 1025 1030 1035			3114
ccg cgt gaa gtg cgc cgc cgc aag cgt gag gcc gag gcc gcc gcc Pro Arg Glu Val Arg Arg Lys Arg Glu Ala Glu Ala Ala Ala 1040 1045 1050			3159
aaa gcc gcg cag gaa gct gct gca gcg acc gag gag cca gcc ctg Lys Ala Ala Gln Glu Ala Ala Ala Ala Thr Glu Glu Pro Ala Leu 1055 1060 1065			3204
gaa gcc gcc gat gag cac aag cct cat cac agt tga Glu Ala Ala Asp Glu His Lys Pro His His Ser			3240

-continued

1070

1075

```

<210> SEQ ID NO 4
<211> LENGTH: 1079
<212> TYPE: PRT
<213> ORGANISM: Pseudomonas putida

<400> SEQUENCE: 4

Met Lys Arg Met Leu Ile Asn Ala Thr Gln Pro Glu Glu Leu Arg Val
1           5          10          15

Ala Leu Val Asp Gly Gln Arg Leu Tyr Asp Leu Asp Ile Glu Ser Gly
20          25          30

Ala Arg Glu Gln Lys Lys Ala Asn Ile Tyr Lys Gly Lys Ile Thr Arg
35          40          45

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

His Gly Phe Leu Pro Leu Lys Glu Ile Ser Arg Glu Tyr Phe Lys Lys
65          70          75          80

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

Glu Val Ile Val Gln Val Glu Lys Glu Glu Arg Gly Asn Lys Gly Ala
100         105         110

Ala Leu Thr Thr Phe Ile Ser Leu Ala Gly Arg Tyr Leu Val Leu Met
115         120         125

Pro Asn Asn Pro Arg Ala Gly Gly Ile Ser Arg Arg Ile Glu Gly Glu
130         135         140

Glu Arg Asn Glu Leu Arg Glu Ala Leu Asn Gly Leu Thr Val Pro Gly
145         150         155         160

Asp Met Gly Leu Ile Val Arg Thr Ala Gly Leu Gly Arg Ser Ser Glu
165         170         175

Glu Met Gln Trp Asp Leu Asp Tyr Leu Leu Gln Leu Trp Thr Ala Ile
180         185         190

Lys Glu Ala Ser Leu Asp Arg Ala Ala Pro Phe Leu Ile Tyr Gln Glu
195         200         205

Ser Asn Val Ile Ile Arg Ala Ile Arg Asp Tyr Leu Arg Gln Asp Ile
210         215         220

Gly Glu Val Leu Ile Asp Ser Ile Asp Ala Gln Glu Glu Ala Leu Thr
225         230         235         240

Phe Ile Arg Gln Val Met Pro Gln Tyr Ala Ser Lys Val Lys Leu Tyr
245         250         255

Glu Asp Ser Val Pro Leu Phe Asn Arg Phe Gln Ile Glu Ser Gln Ile
260         265         270

Glu Thr Ala Phe Gln Arg Val Val Asp Leu Pro Ser Gly Gly Ser Ile
275         280         285

Val Ile Asp Pro Thr Glu Ala Leu Val Ser Ile Asp Ile Asn Ser Ala
290         295         300

Arg Ala Thr Lys Gly Ser Asp Ile Glu Glu Thr Ala Leu Gln Thr Asn
305         310         315         320

Leu Glu Ala Ala Glu Glu Ile Ala Arg Gln Leu Arg Leu Arg Asp Ile
325         330         335

Gly Gly Leu Ile Val Ile Asp Phe Ile Asp Met Thr Pro Ala Lys Asn
340         345         350

```

-continued

Gln	Arg	Ala	Val	Glu	Glu	Arg	Val	Arg	Glu	Cys	Leu	Glu	Ala	Asp	Arg
355				360							365				
Ala	Arg	Val	Gln	Val	Gly	Arg	Ile	Ser	Arg	Phe	Gly	Leu	Leu	Glu	Met
370				375							380				
Ser	Arg	Gln	Arg	Leu	Arg	Pro	Ser	Leu	Gly	Glu	Ser	Ser	Gly	Ile	Val
385				390					395				400		
Cys	Pro	Arg	Cys	Ser	Gly	Thr	Gly	Ile	Ile	Arg	Asp	Val	Glu	Ser	Leu
	405				410				415						
Ser	Leu	Ala	Ile	Leu	Arg	Leu	Ile	Glu	Glu	Ala	Leu	Lys	Asp	Arg	
	420				425				430						
Thr	Ala	Glu	Val	Arg	Ala	Gln	Val	Pro	Ile	Pro	Val	Ala	Ala	Phe	Leu
	435				440				445						
Leu	Asn	Glu	Lys	Arg	Asn	Ser	Ile	Thr	Lys	Ile	Glu	Leu	Arg	Thr	Arg
	450				455				460						
Ala	Arg	Ile	Ile	Ile	Leu	Pro	Asn	Asp	His	Leu	Glu	Thr	Pro	His	Phe
	465				470				475				480		
Glu	Val	Gln	Arg	Leu	Arg	Asp	Asp	Asn	Pro	Glu	Val	Leu	Asn	Asn	Gln
	485				490				495						
Ser	Ser	Tyr	Glu	Ile	Ala	Thr	Thr	Glu	Ala	Glu	Ala	Pro	Gln	Gln	
	500				505				510						
Thr	Ala	Thr	Arg	Thr	Leu	Val	Arg	Gln	Glu	Ala	Ala	Val	Lys	Thr	Ala
	515				520				525						
Pro	Ala	Arg	Thr	Asn	Ala	Pro	Val	Pro	Val	Ala	Glu	Glu	Pro	Gln	Ala
	530				535				540						
Val	Ala	Pro	Ala	Ala	Pro	Ala	Pro	Ser	Ala	Pro	Glu	Pro	Ser	Leu	Phe
	545				550				555				560		
Lys	Gly	Leu	Val	Lys	Ser	Leu	Val	Ser	Leu	Phe	Ala	Gly	Lys	Asp	Glu
	565				570				575						
Pro	Ala	Ala	Ala	Ala	Pro	Val	Val	Ala	Glu	Lys	Pro	Ala	Ala	Glu	Arg
	580				585				590						
Ser	Pro	Arg	Asn	Glu	Glu	Arg	Arg	Asn	Gly	Arg	Gln	Gln	Ser	Arg	Asn
	595				600				605						
Arg	Asn	Gly	Arg	Arg	Asp	Glu	Glu	Arg	Lys	Pro	Arg	Glu	Arg	Ala	
	610				615				620						
Glu	Arg	Ala	Pro	Arg	Glu	Glu	Arg	Gln	Pro	Arg	Glu	Glu	Arg	Ala	Pro
	625				630				635				640		
Arg	Glu	Glu	Arg	Ala	Pro	Arg	Glu	Glu	Arg	Ala	Pro	Arg	Glu	Arg	
	645				650				655						
Ala	Pro	Arg	Gln	Pro	Arg	Glu	Asp	Arg	Ser	Asn	Arg	Gly	Glu	Glu	
	660				665				670						
Arg	Val	Arg	Glu	Leu	Arg	Glu	Pro	Leu	Asp	Ala	Thr	Pro	Pro	Ala	Glu
	675				680				685						
Arg	Glu	Glu	Arg	Gln	Pro	Arg	Glu	Glu	Arg	Val	Ala	Arg	Glu	Glu	Arg
	690				695				700						
Ala	Pro	Arg	Glu	Glu	Arg	Ala	Pro	Arg	Glu	Glu	Arg	Ala	Pro	Arg	Glu
	705				710				715				720		
Glu	Arg	Ala	Pro	Arg	Glu	Glu	Arg	Ala	Pro	Arg	Glu	Arg	Ala	Pro	
	725				730				735						
Arg	Glu	Glu	Arg	Ala	Pro	Arg	Glu	Glu	Arg	Ala	Pro	Arg	Glu	Glu	Arg
	740				745				750						
Ala	Pro	Arg	Glu	Glu	Arg	Ala	Pro	Arg	Pro	Pro	Arg	Glu	Arg	Gln	

-continued

755	760	765
Pro Arg Val Ala Glu Glu Ala Ala Glu Gln Ala Ala Glu		
770	775	780
Glu Gln Leu Pro Asn Glu Glu Leu Leu Gln Asp Glu Gln Glu Gly Thr		
785	790	795
Asp Gly Gly Arg Pro Arg Arg Ser Arg Gly Gln Arg Arg Arg Ser		
805	810	815
Asn Arg Arg Glu Arg Gln Arg Asn Ala Asn Gly Glu Leu Ile Asp Gly		
820	825	830
Gly Glu Glu Glu Gly Ser Glu Glu Gln Pro Gln Gln His Gln Ala		
835	840	845
Thr Glu Leu Gly Ala Glu Leu Ala Ala Gly Leu Ala Val Thr Ala Ala		
850	855	860
Val Ala Ser Ser Asn Ile Ser Ala Asp Ala Glu Ala Gln Ala Asn Gln		
865	870	875
Gln Ala Glu Leu Ala Thr Ala Glu Ile Ala Ala Ala Glu Thr Asp		
885	890	895
Asn Ser His Ala Ala Gln Pro Val Glu Lys Ala Glu Lys Val Glu Pro		
900	905	910
Val Glu Ala Val Ala Lys Ala Glu Asp Val Ala Val Ala Pro Val Val		
915	920	925
Glu Gln Pro Val Ser Glu Pro Val Val Val Ala Glu Val Thr Ala Glu		
930	935	940
Pro Val Val Glu Val Ala Pro Gln His Leu Val Glu Ala Pro Ala		
945	950	955
Ala Glu Pro Val Ile Val Ala Glu Ala Pro Val Glu Thr Pro Ala Val		
965	970	975
Glu Ala Gly Glu Ile Glu Lys Ala Pro Ala Val Val Glu Thr Ala Pro		
980	985	990
Val Ala Glu Gln Pro Ala Pro Val Val Glu Ala Gln Pro Glu Val Val		
995	1000	1005
Ala Glu Pro Ala Pro Val Val Val Glu Pro Ala Pro Val Glu Ala		
1010	1015	1020
Glu Ala Ala Thr Val Met Leu Ala Asn Gly Arg Ala Pro Asn Asp		
1025	1030	1035
Pro Arg Glu Val Arg Arg Lys Arg Glu Ala Glu Ala Ala Ala		
1040	1045	1050
Lys Ala Ala Gln Glu Ala Ala Ala Ala Thr Glu Glu Pro Ala Leu		
1055	1060	1065
Glu Ala Ala Asp Glu His Lys Pro His His Ser		
1070	1075	

```

<210> SEQ ID NO 5
<211> LENGTH: 1563
<212> TYPE: DNA
<213> ORGANISM: Bacillus subtilis
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)..(1563)
<223> OTHER INFORMATION: B. subtilis rny (wild type) encoding RNase Y

<400> SEQUENCE: 5

```

atg acc cca att atg atg gtt ctc atc tcc att ttg ctg att cta ctc 48
Met Thr Pro Ile Met Met Val Leu Ile Ser Ile Leu Ile Leu Leu

-continued

1	5	10	15	
ggt tta gtt gtt ggc tac ttt gtt cgt aaa acc att gcc gaa gcg aaa Gly Leu Val Val Gly Tyr Phe Val Arg Lys Thr Ile Ala Glu Ala Lys 20 25 30				96
att gcg ggc gca cgc ggt gca gcc gag caa att ctt gaa gat gcg aag Ile Ala Gly Ala Arg Gly Ala Ala Glu Gln Ile Leu Glu Asp Ala Lys 35 40 45				144
cgt gat gct gaa gca ctg aaa aaa gaa gct ctg ctt gaa gca aag gat Arg Asp Ala Glu Ala Leu Lys Lys Glu Ala Leu Leu Glu Ala Lys Asp 50 55 60				192
gaa atc cac aaa ctt cga ata gat gct gaa cag gaa gtt cgt gaa aga Glu Ile His Lys Leu Arg Ile Asp Ala Glu Gln Glu Val Arg Glu Arg 65 70 75 80				240
cga aat gag ctt caa aaa caa gaa aac cgt tta ctc caa aag gag gaa Arg Asn Glu Leu Gln Lys Gln Glu Asn Arg Leu Leu Gln Lys Glu Glu 85 90 95				288
aac ctt gat cgc aaa cat gag gga att gat aaa cgg gaa gcg atg ttg Asn Leu Asp Arg Lys His Glu Gly Ile Asp Lys Arg Glu Ala Met Leu 100 105 110				336
gag aag aaa gat cat tct ctg aat gaa cga caa caa cat att gaa gag Glu Lys Asp His Ser Leu Asn Glu Arg Gln Gln His Ile Glu Glu 115 120 125				384
atg gaa agc aaa gtg gat gag atg att cgt atg cag cag tct gag ttg Met Glu Ser Lys Val Asp Glu Met Ile Arg Met Gln Gln Ser Glu Leu 130 135 140				432
gaa cga att tcg agt ctg act cgt gac gaa gcg aaa caa atc att ctt Glu Arg Ile Ser Ser Leu Thr Arg Asp Glu Ala Lys Gln Ile Ile Leu 145 150 155 160				480
gag cgg gtt gaa aac gag ctt tca cat gac atc gcc atc atg aca aaa Glu Arg Val Glu Asn Glu Leu Ser His Asp Ile Ala Ile Met Thr Lys 165 170 175				528
gaa act gaa aac cgt gcg aaa gaa gag gcg gat aag aaa gcg aaa aac Glu Thr Glu Asn Arg Ala Lys Glu Ala Asp Lys Lys Ala Lys Asn 180 185 190				576
att ctt tca ctc gcc tta cag cgc tgc gca gcg gac cac gtt gcc gaa Ile Leu Ser Leu Ala Leu Gln Arg Cys Ala Ala Asp His Val Ala Glu 195 200 205				624
aca acg gta tca gtt gtc aat ctt cca aat gat gag atg aaa gga cgt Thr Thr Val Ser Val Val Asn Leu Pro Asn Asp Glu Met Lys Gly Arg 210 215 220				672
atc atc gga cgg gaa ggg cgt aac att cgt acg ctt gaa acg ctg aca Ile Ile Gly Arg Glu Gly Arg Asn Ile Arg Thr Leu Glu Thr Leu Thr 225 230 235 240				720
gga att gag ctg att att gat gat acg cct gaa gct gtc att ctt tcc Gly Ile Asp Leu Ile Ile Asp Asp Thr Pro Glu Ala Val Ile Leu Ser 245 250 255				768
gga ttt gat ccg atc aga cgt gag aca gcc aga att gct ctt gat aaa Gly Phe Asp Pro Ile Arg Arg Glu Thr Ala Arg Ile Ala Leu Asp Lys 260 265 270				816
ctc gtt cag gat ggc cgt att cat ccg gca cgg att gaa gaa atg gtt Leu Val Gln Asp Gly Arg Ile His Pro Ala Arg Ile Glu Glu Met Val 275 280 285				864
gaa aaa tct cgc cgc gag gtc gat gac tat att cgt gag atg ggt gag Glu Lys Ser Arg Arg Glu Val Asp Asp Tyr Ile Arg Glu Met Gly Glu 290 295 300				912
caa acg aca ttt gag gtt ggc gtt cat ggc ctc cac cca gat ctc atc Gln Thr Thr Phe Glu Val Gly Val His Gly Leu His Pro Asp Leu Ile				960

-continued

305	310	315	320	
aag atc ctc ggc cgc tta aag ttc cgt aca agc tac ggt caa aat gtg Lys Ile Leu Gly Arg Leu Lys Phe Arg Thr Ser Tyr Gly Gln Asn Val 325 330 335				1008
ctt aag cat tcc atg gaa gtc gca ttc ttg gcc ggt cta atg gca tcg Leu Lys His Ser Met Glu Val Ala Phe Leu Ala Gly Leu Met Ala Ser 340 345 350				1056
gag ctt ggt gaa gac gca aag ctt gct aaa cgt gcg ggt ctt ctt cac Glu Leu Gly Glu Asp Ala Lys Leu Ala Lys Arg Ala Gly Leu Leu His 355 360 365				1104
gac atc ggg aaa gca att gac cat gaa gta gaa gga agc cac gtt gag Asp Ile Gly Lys Ala Ile Asp His Glu Val Glu Gly Ser His Val Glu 370 375 380				1152
atc ggg gta gag ctt gcg acc aaa tat aaa gag cac cca gtc gtg att Ile Gly Val Glu Leu Ala Thr Lys Tyr Lys Glu His Pro Val Val Ile 385 390 395 400				1200
aac agt att gca tca cac cac ggg gac gag gag ccg act tcc att att Asn Ser Ile Ala Ser His His Gly Asp Glu Glu Pro Thr Ser Ile Ile 405 410 415				1248
gct gta ctg gta gct gca gca gat gcg ctt tcc gct gca aga cct ggc Ala Val Leu Val Ala Ala Asp Ala Leu Ser Ala Ala Arg Pro Gly 420 425 430				1296
gca aga agt gag acg ctc gag aat tat att cga aga ctt gaa aaa ctt Ala Arg Ser Glu Thr Leu Glu Asn Tyr Ile Arg Arg Leu Glu Lys Leu 435 440 445				1344
gaa gaa att tct gag tcc tac gaa ggt gtt gaa aaa tca ttt gcc att Glu Glu Ile Ser Glu Ser Tyr Glu Gly Val Glu Lys Ser Phe Ala Ile 450 455 460				1392
cag gct gga cgc gaa gtg cga att atg gtg aag ccg gat tca att aat Gln Ala Gly Arg Glu Val Arg Ile Met Val Lys Pro Asp Ser Ile Asn 465 470 475 480				1440
gat ctt gag gct cat cga ctg gcg cga gat atc ccg aag cga att gag Asp Leu Glu Ala His Arg Leu Ala Arg Asp Ile Arg Lys Arg Ile Glu 485 490 495				1488
gac gag ctc gat tat cca ggt cat att aag gtt aca gta atc aga gag Asp Glu Leu Asp Tyr Pro Gly His Ile Lys Val Thr Val Ile Arg Glu 500 505 510				1536
act cga gcc gta gag tat gca aaa taa Thr Arg Ala Val Glu Tyr Ala Lys 515 520				1563
 <210> SEQ ID NO 6				
<211> LENGTH: 520				
<212> TYPE: PRT				
<213> ORGANISM: Bacillus subtilis				
<400> SEQUENCE: 6				
Met Thr Pro Ile Met Met Val Leu Ile Ser Ile Leu Leu Ile Leu Leu 1 5 10 15				
Gly Leu Val Val Gly Tyr Phe Val Arg Lys Thr Ile Ala Glu Ala Lys 20 25 30				
Ile Ala Gly Ala Arg Gly Ala Ala Glu Gln Ile Leu Glu Asp Ala Lys 35 40 45				
Arg Asp Ala Glu Ala Leu Lys Lys Glu Ala Leu Leu Glu Ala Lys Asp 50 55 60				
Glu Ile His Lys Leu Arg Ile Asp Ala Glu Gln Glu Val Arg Glu Arg 65 70 75 80				

-continued

Arg Asn Glu Leu Gln Lys Gln Glu Asn Arg Leu Leu Gln Lys Glu Glu			
85	90	95	
Asn Leu Asp Arg Lys His Glu Gly Ile Asp Lys Arg Glu Ala Met Leu			
100	105	110	
Glu Lys Asp His Ser Leu Asn Glu Arg Gln Gln His Ile Glu Glu			
115	120	125	
Met Glu Ser Lys Val Asp Glu Met Ile Arg Met Gln Gln Ser Glu Leu			
130	135	140	
Glu Arg Ile Ser Ser Leu Thr Arg Asp Glu Ala Lys Gln Ile Ile Leu			
145	150	155	160
Glu Arg Val Glu Asn Glu Leu Ser His Asp Ile Ala Ile Met Thr Lys			
165	170	175	
Glu Thr Glu Asn Arg Ala Lys Glu Ala Asp Lys Lys Ala Lys Asn			
180	185	190	
Ile Leu Ser Leu Ala Leu Gln Arg Cys Ala Ala Asp His Val Ala Glu			
195	200	205	
Thr Thr Val Ser Val Val Asn Leu Pro Asn Asp Glu Met Lys Gly Arg			
210	215	220	
Ile Ile Gly Arg Glu Gly Arg Asn Ile Arg Thr Leu Glu Thr Leu Thr			
225	230	235	240
Gly Ile Asp Leu Ile Ile Asp Asp Thr Pro Glu Ala Val Ile Leu Ser			
245	250	255	
Gly Phe Asp Pro Ile Arg Arg Glu Thr Ala Arg Ile Ala Leu Asp Lys			
260	265	270	
Leu Val Gln Asp Gly Arg Ile His Pro Ala Arg Ile Glu Glu Met Val			
275	280	285	
Glu Lys Ser Arg Arg Glu Val Asp Asp Tyr Ile Arg Glu Met Gly Glu			
290	295	300	
Gln Thr Thr Phe Glu Val Gly Val His Gly Leu His Pro Asp Leu Ile			
305	310	315	320
Lys Ile Leu Gly Arg Leu Lys Phe Arg Thr Ser Tyr Gly Gln Asn Val			
325	330	335	
Leu Lys His Ser Met Glu Val Ala Phe Leu Ala Gly Leu Met Ala Ser			
340	345	350	
Glu Leu Gly Glu Asp Ala Lys Leu Ala Lys Arg Ala Gly Leu Leu His			
355	360	365	
Asp Ile Gly Lys Ala Ile Asp His Glu Val Glu Gly Ser His Val Glu			
370	375	380	
Ile Gly Val Glu Leu Ala Thr Lys Tyr Lys Glu His Pro Val Val Ile			
385	390	395	400
Asn Ser Ile Ala Ser His His Gly Asp Glu Glu Pro Thr Ser Ile Ile			
405	410	415	
Ala Val Leu Val Ala Ala Ala Asp Ala Leu Ser Ala Ala Arg Pro Gly			
420	425	430	
Ala Arg Ser Glu Thr Leu Glu Asn Tyr Ile Arg Arg Leu Glu Lys Leu			
435	440	445	
Glu Glu Ile Ser Glu Ser Tyr Glu Gly Val Glu Lys Ser Phe Ala Ile			
450	455	460	
Gln Ala Gly Arg Glu Val Arg Ile Met Val Lys Pro Asp Ser Ile Asn			
465	470	475	480

-continued

Asp	Leu	Glu	Ala	His	Arg	Leu	Ala	Arg	Asp	Ile	Arg	Lys	Arg	Ile	Glu
485						490						495			

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

Thr	Arg	Ala	Val	Glu	Tyr	Ala	Lys	
515						520		

<210> SEQ ID NO 7

<211> LENGTH: 1560

<212> TYPE: DNA

<213> ORGANISM: *Bacillus licheniformis*

<220> FEATURE:

<221> NAME/KEY: CDS

<222> LOCATION: (1)..(1560)

<223> OTHER INFORMATION: *Bacillus licheniformis* rny encoding RNase Y

<400> SEQUENCE: 7

atg	agt	cct	tta	gca	att	ctc	atc	tcc	att	ttt	ctg	agc	cta	tcc	tgt	48
Met	Ser	Pro	Leu	Ala	Ile	Leu	Ile	Ser	Ile	Leu	Leu	Ser	Leu	Phe	Cys	
1									10					15		

tta	gtt	ggc	tac	tat	gtt	cgt	aaa	atc	att	gcc	gaa	gca	aaa	att	96	
Leu	Val	Val	Gly	Tyr	Tyr	Val	Arg	Lys	Ile	Ile	Ala	Glu	Ala	Lys	Ile	
									25					30		

tca	ggt	gcg	cga	aat	gca	gcc	gaa	caa	att	ctt	gga	gac	gca	aag	cgg	144
Ser	Gly	Ala	Arg	Asn	Ala	Ala	Glu	Gln	Ile	Leu	Gly	Asp	Ala	Lys	Arg	
									35			40		45		

gat	gct	gaa	gcg	ttg	aaa	aaa	gaa	gcc	ctt	ctt	gaa	gca	aag	gac	gag	192
Asp	Ala	Glu	Ala	Leu	Lys	Lys	Glu	Ala	Leu	Leu	Glu	Ala	Lys	Asp	Glu	
									50			55		60		

att	cat	acg	ctt	cggt	ata	gaa	gct	gaa	caa	gaa	gtt	cgt	gaa	aga	cga	240
Ile	His	Thr	Leu	Arg	Ile	Glu	Ala	Glu	Gln	Glu	Val	Arg	Glu	Arg	Arg	
									65			70		75		80

aat	gag	ctt	caa	aaa	caa	gaa	aac	cgt	tta	ctt	caa	aag	gaa	gag	aac	288
Asn	Glu	Leu	Gln	Lys	Gln	Glu	Asn	Arg	Leu	Leu	Gln	Lys	Glu	Glu	Asn	
									85			90		95		

ctt	gac	cga	aaa	gat	gaa	tca	tta	gat	aaa	cgg	gaa	gcg	atg	ttt	gag	336
Leu	Asp	Arg	Lys	Asp	Glu	Ser	Leu	Asp	Lys	Arg	Glu	Ala	Met	Leu	Glu	
									100			105		110		

aag	aaa	gat	cat	tct	ctg	aat	gaa	cga	caa	caa	cat	att	gaa	gag	atg	384
Lys	Lys	Asp	His	Ser	Leu	Asn	Glu	Arg	Gln	Gln	His	Ile	Glu	Glu	Met	
									115			120		125		

gaa	agc	aaa	gtg	gat	gaa	atg	att	cgt	atg	cag	cag	tca	gag	ttg	gag	432
Glu	Ser	Lys	Val	Asp	Glu	Met	Ile	Arg	Met	Gln	Gln	Ser	Glu	Leu	Glu	
									130			135		140		

cgt	att	tca	agt	ctg	acg	cga	gat	gaa	gcg	aag	caa	atc	att	ctg	gaa	480
Arg	Ile	Ser	Ser	Leu	Thr	Arg	Asp	Glu	Ala	Lys	Gln	Ile	Ile	Leu	Glu	
									145			150		155		160

cgg	gtt	gaa	aac	gag	ctt	tcc	cat	gac	atc	gcg	atc	atg	atg	aaa	gaa	528
Arg	Val	Glu	Asn	Glu	Leu	Ser	His	Asp	Ile	Ala	Ile	Met	Met	Lys	Glu	
									165			170		175		

agt	gaa	aat	cga	gcg	aaa	gaa	gag	gct	gat	aaa	aag	gcg	aaa	aat	att	576
Ser	Glu	Asn	Arg	Ala	Lys	Glu	Glu	Ala	Asp	Lys	Ala	Lys	Asn	Ile		
									180			185		190		

ctt	tca	ttg	gct	tta	cag	cgc	tgt	gcg	gca	gat	cat	gtg	gcc	gag	aca	624
Leu	Ser	Leu	Ala	Leu	Gln	Arg	Cys	Ala	Ala	Asp	His	Val	Ala	Glu	Thr	
									195			200		205		

acg	gtt	tca	gtt	gtc	aac	ctt	cca	aat	gat	gag	atg	aaa	ggc	cgc	att	672
Thr	Val	Ser	Val	Val	Asn	Leu	Pro	Asn	Asp	Glu	Met	Lys	Gly	Arg	Ile	
									210			215		220		

-continued

atc gga cgt gaa gga cgt aat atc cgt aca ttg gaa aca ttg acg ggg Ile Gly Arg Glu Gly Arg Asn Ile Arg Thr Leu Glu Thr Leu Thr Gly 225 230 235 240	720
atc gat ctg atc atc gac gat acg ccg gaa gcc gtt atc ctt tca gga Ile Asp Leu Ile Ile Asp Asp Thr Pro Glu Ala Val Ile Leu Ser Gly 245 250 255	768
ttt gat ccg atc agg cgt gaa aca gcc agg atc gct ctt gat aaa ctc Phe Asp Pro Ile Arg Arg Glu Thr Ala Arg Ile Ala Leu Asp Lys Leu 260 265 270	816
gtt cag gat ggc cgc att cat cct gcc aga atc gaa gaa atg gtt gag Val Gln Asp Gly Arg Ile His Pro Ala Arg Ile Glu Glu Met Val Glu 275 280 285	864
aaa tcc cgc cgt gaa gtg gat gat tac atc cgc gaa atg ggt gaa cag Lys Ser Arg Arg Glu Val Asp Asp Tyr Ile Arg Glu Met Gly Glu Gln 290 295 300	912
acg acg ttt gaa gtt gga gtt cac ggt ctt cat ccc gat tta atc aaa Thr Thr Phe Glu Val Gly Val His Gly Leu His Pro Asp Leu Ile Lys 305 310 315 320	960
atc ctc ggc ttg aag ttc aga acg tac gga cag aat gta ctg Ile Leu Gly Arg Leu Lys Phe Arg Thr Ser Tyr Gly Gln Asn Val Leu 325 330 335	1008
aaa cat tca atg gaa gtg gcg ttt ctg aca ggc ttg atg gct tca gag Lys His Ser Met Glu Val Ala Phe Leu Thr Gly Leu Met Ala Ser Glu 340 345 350	1056
ctc gga gaa gac gta acg ctc gca aaa agg gca gga ctt ctt cat gat Leu Gly Glu Asp Val Thr Leu Ala Lys Arg Ala Gly Leu Leu His Asp 355 360 365	1104
atc ggg aaa gcg att gac cat gag gtg gaa gga agc cac gtt gaa atc Ile Gly Lys Ala Ile Asp His Glu Val Glu Gly Ser His Val Glu Ile 370 375 380	1152
ggt gtg gaa ctg gcc acc aag tat aaa gag cat cca gtt gtc atc aac Gly Val Glu Leu Ala Thr Lys Tyr Lys Glu His Pro Val Val Ile Asn 385 390 395 400	1200
agt atc gca tct cac cac gga gat cag gag ccg act tcc atc atc gcc Ser Ile Ala Ser His His Gly Asp Gln Glu Pro Thr Ser Ile Ile Ala 405 410 415	1248
gtg ctc gtg gcc gca gcc gat gcg ctg tct gcc gca aga cca ggc gca Val Leu Val Ala Ala Ala Asp Ala Leu Ser Ala Ala Arg Pro Gly Ala 420 425 430	1296
aga agc gaa acg ctc gaa aac tat att cgc agg ctt gaa aaa ctg gaa Arg Ser Glu Thr Leu Glu Asn Tyr Ile Arg Arg Leu Glu Lys Leu Glu 435 440 445	1344
gaa atc tcc gaa tca tac gaa ggt gtt gaa aaa tca ttc gcg ata caa Glu Ile Ser Glu Ser Tyr Glu Gly Val Glu Lys Ser Phe Ala Ile Gln 450 455 460	1392
gcg gga cga gag gtg cgg atc atg gtc aaa cct gac tca atc aat gat Ala Gly Arg Glu Val Arg Ile Met Val Lys Pro Asp Ser Ile Asn Asp 465 470 475 480	1440
ctt gaa gcc cat cgt ttg gcg cgt gat att cga aaa cga att gag gac Leu Glu Ala His Arg Leu Ala Arg Asp Ile Arg Lys Arg Ile Glu Asp 485 490 495	1488
gag ctt gac tac cct gga cac atc aag gta acc gta atc agg gaa acg Glu Leu Asp Tyr Pro Gly His Ile Lys Val Thr Val Ile Arg Glu Thr 500 505 510	1536
aga gcc gtt gaa tac gca aaa taa Arg Ala Val Glu Tyr Ala Lys 515	1560

-continued

<210> SEQ ID NO 8
<211> LENGTH: 519
<212> TYPE: PRT
<213> ORGANISM: *Bacillus licheniformis*

<400> SEQUENCE: 8

Met Ser Pro Leu Ala Ile Leu Ile Ser Ile Leu Leu Ser Leu Phe Cys
1 5 10 15

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

Ser Gly Ala Arg Asn Ala Ala Glu Gln Ile Leu Gly Asp Ala Lys Arg
35 40 45

Asp Ala Glu Ala Leu Lys Lys Glu Ala Leu Leu Glu Ala Lys Asp Glu
50 55 60

Ile His Thr Leu Arg Ile Glu Ala Glu Gln Glu Val Arg Glu Arg Arg
65 70 75 80

Asn Glu Leu Gln Lys Gln Glu Asn Arg Leu Leu Gln Lys Glu Glu Asn
85 90 95

Leu Asp Arg Lys Asp Glu Ser Leu Asp Lys Arg Glu Ala Met Leu Glu
100 105 110

Lys Lys Asp His Ser Leu Asn Glu Arg Gln Gln His Ile Glu Glu Met
115 120 125

Glu Ser Lys Val Asp Glu Met Ile Arg Met Gln Gln Ser Glu Leu Glu
130 135 140

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

Arg Val Glu Asn Glu Leu Ser His Asp Ile Ala Ile Met Met Lys Glu
165 170 175

Ser Glu Asn Arg Ala Lys Glu Glu Ala Asp Lys Lys Ala Lys Asn Ile
180 185 190

Leu Ser Leu Ala Leu Gln Arg Cys Ala Ala Asp His Val Ala Glu Thr
195 200 205

Thr Val Ser Val Val Asn Leu Pro Asn Asp Glu Met Lys Gly Arg Ile
210 215 220

Ile Gly Arg Glu Gly Arg Asn Ile Arg Thr Leu Glu Thr Leu Thr Gly
225 230 235 240

Ile Asp Leu Ile Ile Asp Asp Thr Pro Glu Ala Val Ile Leu Ser Gly
245 250 255

Phe Asp Pro Ile Arg Arg Glu Thr Ala Arg Ile Ala Leu Asp Lys Leu
260 265 270

Val Gln Asp Gly Arg Ile His Pro Ala Arg Ile Glu Glu Met Val Glu
275 280 285

Lys Ser Arg Arg Glu Val Asp Asp Tyr Ile Arg Glu Met Gly Glu Gln
290 295 300

Thr Thr Phe Glu Val Gly Val His Gly Leu His Pro Asp Leu Ile Lys
305 310 315 320

Ile Leu Gly Arg Leu Lys Phe Arg Thr Ser Tyr Gly Gln Asn Val Leu
325 330 335

Lys His Ser Met Glu Val Ala Phe Leu Thr Gly Leu Met Ala Ser Glu
340 345 350

Leu Gly Glu Asp Val Thr Leu Ala Lys Arg Ala Gly Leu Leu His Asp

-continued

355	360	365	
Ile Gly Lys Ala Ile Asp His Glu Val Glu Gly Ser His Val Glu Ile			
370	375	380	
Gly Val Glu Leu Ala Thr Lys Tyr Lys Glu His Pro Val Val Ile Asn			
385	390	395	400
Ser Ile Ala Ser His His Gly Asp Gln Glu Pro Thr Ser Ile Ile Ala			
405	410	415	
Val Leu Val Ala Ala Ala Asp Ala Leu Ser Ala Ala Arg Pro Gly Ala			
420	425	430	
Arg Ser Glu Thr Leu Glu Asn Tyr Ile Arg Arg Leu Glu Lys Leu Glu			
435	440	445	
Glu Ile Ser Glu Ser Tyr Glu Gly Val Glu Lys Ser Phe Ala Ile Gln			
450	455	460	
Ala Gly Arg Glu Val Arg Ile Met Val Lys Pro Asp Ser Ile Asn Asp			
465	470	475	480
Leu Glu Ala His Arg Leu Ala Arg Asp Ile Arg Lys Arg Ile Glu Asp			
485	490	495	
Glu Leu Asp Tyr Pro Gly His Ile Lys Val Thr Val Ile Arg Glu Thr			
500	505	510	
Arg Ala Val Glu Tyr Ala Lys			
515			

```

<210> SEQ ID NO 9
<211> LENGTH: 2514
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cereviciae
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)..(2514)
<223> OTHER INFORMATION: CCR4

```

<400> SEQUENCE: 9

atg aac gac cct tct tta cta ggc tac cct aat gtg ggc ccg cag cag	48
Met Asn Asp Pro Ser Leu Leu Gly Tyr Pro Asn Val Gly Pro Gln Gln	
1 5 10 15	
caa cag caa cag cag caa cag cag cat gcc ggg cta cta gga aag ggt	96
Gln Gln Gln Gln Gln Gln His Ala Gly Leu Leu Gly Lys Gly	
20 25 30	
aca cca aac gct cta cag cag cta cac atg aac cag ctc acc ggg	144
Thr Pro Asn Ala Leu Gln Gln Leu His Met Asn Gln Leu Thr Gly	
35 40 45	
ata cct cca ccg gga ctc atg aac aac agc gat gta cac act tcc agc	192
Ile Pro Pro Pro Gly Leu Met Asn Asn Ser Asp Val His Thr Ser Ser	
50 55 60	
aac aac aac tcg cgc cag ttg ctc gac caa ttg gct aat ggg aac gcg	240
Asn Asn Asn Ser Arg Gln Leu Leu Asp Gln Leu Ala Asn Gly Asn Ala	
65 70 75 80	
aac atg ctc aat atg aac atg gat aac aac aat aac aac	288
Asn Met Leu Asn Met Asn Met Asp Asn Asn Asn Asn Asn Asn Asn	
85 90 95	
aac aac aac aac aac ggc ggt gga agc ggg gtg atg atg aat	336
Asn Asn Asn Asn Asn Asn Gly Gly Ser Gly Val Met Met Asn	
100 105 110	
gcc tcc aca gcc gca gta aat tct ata ggg atg gtt cct acc gtt ggt	384
Ala Ser Thr Ala Ala Val Asn Ser Ile Gly Met Val Pro Thr Val Gly	
115 120 125	

-continued

acg cct gtg aac atc aat gtg aac gcc agt aac cct tta tta cac ccg Thr Pro Val Asn Ile Asn Val Asn Ala Ser Asn Pro Leu Leu His Pro 130 135 140	432
cac ttg gat gac cct tct tta cta aac aat ccg atc tgg aag ctt caa His Leu Asp Asp Pro Ser Leu Leu Asn Asn Pro Ile Trp Lys Leu Gln 145 150 155 160	480
ttg cat ctg gct gca gtg tcc gca caa tct ctg ggg caa ccc aac att Leu His Leu Ala Ala Val Ser Ala Gln Ser Leu Gly Gln Pro Asn Ile 165 170 175	528
tat gct aga caa aac gct atg aag aag tat ttg gct aca caa cag gct Tyr Ala Arg Gln Asn Ala Met Lys Lys Tyr Leu Ala Thr Gln Gln Ala 180 185 190	576
cag caa gct cag caa caa gcg cag caa cag gca cag cag cag gtc ccg Gln Gln Ala Gln Gln Ala Gln Gln Gln Ala Gln Gln Gln Val Pro 195 200 205	624
ggc cca ttt ggc ccc gga cct cag gct gca cca cca gct ttg cag ccc Gly Pro Phe Gly Pro Gly Pro Gln Ala Ala Pro Pro Ala Leu Gln Pro 210 215 220	672
acc gat ttc cag caa tct cac att gca gaa gcc tcc aaa tca ctg gta Thr Asp Phe Gln Gln Ser His Ile Ala Glu Ala Ser Lys Ser Leu Val 225 230 235 240	720
gac tgc aca aag caa gcc ttg atg gaa atg gcc gac act ctc acc gac Asp Cys Thr Lys Gln Ala Leu Met Glu Met Ala Asp Thr Leu Thr Asp 245 250 255	768
agc aag aca gca aag aaa caa caa cct acg gga gat agc act ccc tca Ser Lys Thr Ala Lys Lys Gln Gln Pro Thr Gly Asp Ser Thr Pro Ser 260 265 270	816
ggc acg gca act aac agt gca gtt tct aca cca ttg act ccc aag ata Gly Thr Ala Thr Asn Ser Ala Val Ser Thr Pro Leu Thr Pro Lys Ile 275 280 285	864
gag ctg ttt gct aat ggc aag gac gaa gcc aac cag gcg ctc tta caa Glu Leu Phe Ala Asn Gly Lys Asp Glu Ala Asn Gln Ala Leu Leu Gln 290 295 300	912
cac aag aaa ctg tct cag tac agc atc gac gaa gat gac gac att gaa His Lys Lys Leu Ser Gln Tyr Ser Ile Asp Glu Asp Asp Asp Ile Glu 305 310 315 320	960
aac aga atg gtc atg ccc aag gac tcg aaa tac gac gac caa tta tgg Asn Arg Met Val Met Pro Lys Asp Ser Lys Tyr Asp Asp Gln Leu Trp 325 330 335	1008
cac gcg cta gat ttg tcc aac ttg caa atc ttc aat atc agc gcc aac His Ala Leu Asp Leu Ser Asn Leu Gln Ile Phe Asn Ile Ser Ala Asn 340 345 350	1056
atc ttc aag tac gat ttt cta acg aga cta tat ttg aat ggc aat agc Ile Phe Lys Tyr Asp Phe Leu Thr Arg Leu Tyr Leu Asn Gly Asn Ser 355 360 365	1104
ctc acg gaa ctg cca gcg gag atc aag aac cta agc aac cta cgc gtt Leu Thr Glu Leu Pro Ala Glu Ile Lys Asn Leu Ser Asn Leu Arg Val 370 375 380	1152
ttg gac ctg tcg cat aat agg tta aca tct cta cca gcg gaa cta ggc Leu Asp Leu Ser His Asn Arg Leu Thr Ser Leu Pro Ala Glu Leu Gly 385 390 395 400	1200
tca tgt ttc caa ttg aaa tac ttc tac ttt ttt gat aac atg gtc acc Ser Cys Phe Gln Leu Lys Tyr Phe Tyr Phe Asp Asn Met Val Thr 405 410 415	1248
aca tta cca tgg gag ttt ggg aac ctg tgt aac ctt cag ttt ctt ggt Thr Leu Pro Trp Glu Phe Gly Asn Leu Cys Asn Leu Gln Phe Leu Gly 420 425 430	1296

-continued

gta gaa gga aac ccc tta gaa aaa cag ttt tta aag atc ctt aca gaa Val Glu Gly Asn Pro Leu Glu Lys Gln Phe Leu Lys Ile Leu Thr Glu 435 440 445	1344
aaa tct gtc acg gga ttg att ttc tac cta agg gac aac aga cca gag Lys Ser Val Thr Gly Leu Ile Phe Tyr Leu Arg Asp Asn Arg Pro Glu 450 455 460	1392
att ccc tta ccg cat gaa cgt agg ttc atc gaa atc aat acc gat ggg Ile Pro Leu Pro His Glu Arg Arg Phe Ile Glu Ile Asn Thr Asp Gly 465 470 475 480	1440
gaa cca cag agg gag tac gat tct ttg cag caa tcc act gag cat ctg Glu Pro Gln Arg Glu Tyr Asp Ser Leu Gln Gln Ser Thr Glu His Leu 485 490 495	1488
gcc acc gat tta gcc aag agg acg ttc acc gtc tta tcc tac aac acc Ala Thr Asp Leu Ala Lys Arg Thr Phe Thr Val Leu Ser Tyr Asn Thr 500 505 510	1536
tta tgt caa cac tat gcc acc cca aaa atg tac cgt tac aca ccg tcg Leu Cys Gln His Tyr Ala Thr Pro Lys Met Tyr Arg Tyr Thr Pro Ser 515 520 525	1584
tgg gcg tta agt tgg gat tac agg cgc aat aaa tta aag gag cag att Trp Ala Leu Ser Trp Asp Tyr Arg Arg Asn Lys Leu Lys Glu Gln Ile 530 535 540	1632
ctc tcg tac gac agt gat ctg ttg tgt tta caa gaa gtg gag tct aag Leu Ser Tyr Asp Ser Asp Leu Leu Cys Leu Gln Glu Val Glu Ser Lys 545 550 555 560	1680
act ttt gaa gag tat tgg gtg ccc cta ttg gac aag cac ggt tat aca Thr Phe Glu Tyr Trp Val Pro Leu Leu Asp Lys His Gly Tyr Thr 565 570 575	1728
ggc atc ttc cat gca aag gca aga gcc aag acc atg cat tcc aag gac Gly Ile Phe His Ala Lys Ala Arg Ala Lys Thr Met His Ser Lys Asp 580 585 590	1776
tcc aag aaa gtg gac ggg tgt tgc att ttt ttc aaa agg gac caa ttc Ser Lys Lys Val Asp Gly Cys Cys Ile Phe Phe Lys Arg Asp Gln Phe 595 600 605	1824
aag ttg atc acc aaa gac gcc atg gat ttc agc ggt gct tgg atg aaa Lys Leu Ile Thr Lys Asp Ala Met Asp Phe Ser Gly Ala Trp Met Lys 610 615 620	1872
cac aag aag ttc caa aga act gaa gat tat tta aac cgt gca atg aac His Lys Lys Phe Gln Arg Thr Glu Asp Tyr Leu Asn Arg Ala Met Asn 625 630 635 640	1920
aaa gac aac gtt gca ctg ttc tta aag cta caa cac att cct agt ggc Lys Asp Asn Val Ala Leu Phe Leu Lys Leu Gln His Ile Pro Ser Gly 645 650 655	1968
gac acc ata tgg gcg gtc acc acg cat ttg cac tgg gat cca aaa ttt Asp Thr Ile Trp Ala Val Thr His Leu His Trp Asp Pro Lys Phe 660 665 670	2016
aat gat gtc aag aca ttc caa gta ggt gtc ctg tta gat cat ctg gaa Asn Asp Val Lys Thr Phe Gln Val Gly Val Leu Leu Asp His Leu Glu 675 680 685	2064
acg ctg cta aag gag gag aca tcg cac aat ttt aga cag gac att aag Thr Leu Leu Lys Glu Glu Thr Ser His Asn Phe Arg Gln Asp Ile Lys 690 695 700	2112
aaa ttt cct gtg ctc att tgt ggt gac ttc aat tca tac atc aac tcc Lys Phe Pro Val Leu Ile Cys Gly Asp Phe Asn Ser Tyr Ile Asn Ser 705 710 715 720	2160
gcc gta tac gaa ttg ata aat aca ggc cgt gtc caa ata cat caa gag Ala Val Tyr Glu Leu Ile Asn Thr Gly Arg Val Gln Ile His Gln Glu 725 730 735	2208

-continued

gga aat ggt aga gat ttc ggt tac atg tcg gag aaa aat ttc tca cat Gly Asn Gly Arg Asp Phe Gly Tyr Met Ser Glu Lys Asn Phe Ser His 740 745 750	2256
aac ttg gct ctc aaa tca agt tat aat tgt atc gga gaa cta cca ttc Asn Leu Ala Leu Lys Ser Ser Tyr Asn Cys Ile Gly Glu Leu Pro Phe 755 760 765	2304
acc aat ttc aca ccg tca ttc aca gat gtt atc gac tat ata tgg ttt Thr Asn Phe Thr Pro Ser Phe Thr Asp Val Ile Asp Tyr Ile Trp Phe 770 775 780	2352
tct aca cat gct cta agg gtg cgt ggg cta ttg ggt gaa gtg gac cct Ser Thr His Ala Leu Arg Val Arg Gly Leu Leu Gly Glu Val Asp Pro 785 790 795 800	2400
gaa tac gtg agt aag ttt atc ggg ttc ccc aac gac aaa ttc ccc agt Glu Tyr Val Ser Lys Phe Ile Gly Phe Pro Asn Asp Lys Phe Pro Ser 805 810 815	2448
gac cat ata cca ttg tta gca aga ttt gaa ttt atg aag aca aac aca Asp His Ile Pro Leu Leu Ala Arg Phe Glu Phe Met Lys Thr Asn Thr 820 825 830	2496
ggc agt aag aaa gta taa Gly Ser Lys Lys Val 835	2514

<210> SEQ ID NO 10

<211> LENGTH: 837

<212> TYPE: PRT

<213> ORGANISM: *Saccharomyces cereviciae*

<400> SEQUENCE: 10

Met Asn Asp Pro Ser Leu Leu Gly Tyr Pro Asn Val Gly Pro Gln Gln 1 5 10 15
Gln Gln Gln Gln Gln Gln His Ala Gly Leu Leu Gly Lys Gly 20 25 30
Thr Pro Asn Ala Leu Gln Gln Leu His Met Asn Gln Leu Thr Gly 35 40 45
Ile Pro Pro Pro Gly Leu Met Asn Asn Ser Asp Val His Thr Ser Ser 50 55 60
Asn Asn Asn Ser Arg Gln Leu Leu Asp Gln Leu Ala Asn Gly Asn Ala 65 70 75 80
Asn Met Leu Asn Met Asn Met Asp Asn Asn Asn Asn Asn Asn Asn 85 90 95
Asn Asn Asn Asn Asn Asn Gly Gly Ser Gly Val Met Met Asn 100 105 110
Ala Ser Thr Ala Ala Val Asn Ser Ile Gly Met Val Pro Thr Val Gly 115 120 125
Thr Pro Val Asn Ile Asn Val Asn Ala Ser Asn Pro Leu Leu His Pro 130 135 140
His Leu Asp Asp Pro Ser Leu Leu Asn Asn Pro Ile Trp Lys Leu Gln 145 150 155 160
Leu His Leu Ala Ala Val Ser Ala Gln Ser Leu Gly Gln Pro Asn Ile 165 170 175
Tyr Ala Arg Gln Asn Ala Met Lys Lys Tyr Leu Ala Thr Gln Gln Ala 180 185 190
Gln Gln Ala Gln Gln Ala Gln Gln Gln Ala Gln Gln Gln Val Pro 195 200 205
Gly Pro Phe Gly Pro Gly Pro Gln Ala Ala Pro Pro Ala Leu Gln Pro

-continued

210	215	220
Thr Asp Phe Gln Gln Ser His Ile Ala Glu Ala Ser Lys Ser Leu Val		
225	230	235
Asp Cys Thr Lys Gln Ala Leu Met Glu Met Ala Asp Thr Leu Thr Asp		
245	250	255
Ser Lys Thr Ala Lys Lys Gln Gln Pro Thr Gly Asp Ser Thr Pro Ser		
260	265	270
Gly Thr Ala Thr Asn Ser Ala Val Ser Thr Pro Leu Thr Pro Lys Ile		
275	280	285
Glu Leu Phe Ala Asn Gly Lys Asp Glu Ala Asn Gln Ala Leu Leu Gln		
290	295	300
His Lys Lys Leu Ser Gln Tyr Ser Ile Asp Glu Asp Asp Asp Ile Glu		
305	310	315
Asn Arg Met Val Met Pro Lys Asp Ser Lys Tyr Asp Asp Gln Leu Trp		
325	330	335
His Ala Leu Asp Leu Ser Asn Leu Gln Ile Phe Asn Ile Ser Ala Asn		
340	345	350
Ile Phe Lys Tyr Asp Phe Leu Thr Arg Leu Tyr Leu Asn Gly Asn Ser		
355	360	365
Leu Thr Glu Leu Pro Ala Glu Ile Lys Asn Leu Ser Asn Leu Arg Val		
370	375	380
Leu Asp Leu Ser His Asn Arg Leu Thr Ser Leu Pro Ala Glu Leu Gly		
385	390	395
Ser Cys Phe Gln Leu Lys Tyr Phe Tyr Phe Phe Asp Asn Met Val Thr		
405	410	415
Thr Leu Pro Trp Glu Phe Gly Asn Leu Cys Asn Leu Gln Phe Leu Gly		
420	425	430
Val Glu Gly Asn Pro Leu Glu Lys Gln Phe Leu Lys Ile Leu Thr Glu		
435	440	445
Lys Ser Val Thr Gly Leu Ile Phe Tyr Leu Arg Asp Asn Arg Pro Glu		
450	455	460
Ile Pro Leu Pro His Glu Arg Arg Phe Ile Glu Ile Asn Thr Asp Gly		
465	470	475
Glu Pro Gln Arg Glu Tyr Asp Ser Leu Gln Gln Ser Thr Glu His Leu		
485	490	495
Ala Thr Asp Leu Ala Lys Arg Thr Phe Thr Val Leu Ser Tyr Asn Thr		
500	505	510
Leu Cys Gln His Tyr Ala Thr Pro Lys Met Tyr Arg Tyr Thr Pro Ser		
515	520	525
Trp Ala Leu Ser Trp Asp Tyr Arg Arg Asn Lys Leu Lys Glu Gln Ile		
530	535	540
Leu Ser Tyr Asp Ser Asp Leu Leu Cys Leu Gln Glu Val Glu Ser Lys		
545	550	555
Thr Phe Glu Glu Tyr Trp Val Pro Leu Leu Asp Lys His Gly Tyr Thr		
565	570	575
Gly Ile Phe His Ala Lys Ala Arg Ala Lys Thr Met His Ser Lys Asp		
580	585	590
Ser Lys Lys Val Asp Gly Cys Cys Ile Phe Phe Lys Arg Asp Gln Phe		
595	600	605
Lys Leu Ile Thr Lys Asp Ala Met Asp Phe Ser Gly Ala Trp Met Lys		
610	615	620

-continued

His Lys Lys Phe Gln Arg Thr Glu Asp Tyr Leu Asn Arg Ala Met Asn
 625 630 635 640
 Lys Asp Asn Val Ala Leu Phe Leu Lys Leu Gln His Ile Pro Ser Gly
 645 650 655
 Asp Thr Ile Trp Ala Val Thr Thr His Leu His Trp Asp Pro Lys Phe
 660 665 670
 Asn Asp Val Lys Thr Phe Gln Val Gly Val Leu Leu Asp His Leu Glu
 675 680 685
 Thr Leu Leu Lys Glu Glu Thr Ser His Asn Phe Arg Gln Asp Ile Lys
 690 695 700
 Lys Phe Pro Val Leu Ile Cys Gly Asp Phe Asn Ser Tyr Ile Asn Ser
 705 710 715 720
 Ala Val Tyr Glu Leu Ile Asn Thr Gly Arg Val Gln Ile His Gln Glu
 725 730 735
 Gly Asn Gly Arg Asp Phe Gly Tyr Met Ser Glu Lys Asn Phe Ser His
 740 745 750
 Asn Leu Ala Leu Lys Ser Ser Tyr Asn Cys Ile Gly Glu Leu Pro Phe
 755 760 765
 Thr Asn Phe Thr Pro Ser Phe Thr Asp Val Ile Asp Tyr Ile Trp Phe
 770 775 780
 Ser Thr His Ala Leu Arg Val Arg Gly Leu Leu Gly Glu Val Asp Pro
 785 790 795 800
 Glu Tyr Val Ser Lys Phe Ile Gly Phe Pro Asn Asp Lys Phe Pro Ser
 805 810 815
 Asp His Ile Pro Leu Leu Ala Arg Phe Glu Phe Met Lys Thr Asn Thr
 820 825 830
 Gly Ser Lys Lys Val
 835

<210> SEQ ID NO 11
 <211> LENGTH: 3348
 <212> TYPE: DNA
 <213> ORGANISM: *Saccharomyces cereviciae*
 <220> FEATURE:
 <221> NAME/KEY: CDS
 <222> LOCATION: (1)..(3348)
 <223> OTHER INFORMATION: Pan2

<400> SEQUENCE: 11

atg aat aat tgg caa cat ttc ttc aac aat cca gtt gat ctt tcg gaa	48
Met Asn Asn Trp Gln His Phe Phe Asn Asn Pro Val Asp Leu Ser Glu	
1 5 10 15	
cat ttg aag aag cca tac ttt cgc ttc gat aat agg gat aag gaa att	96
His Leu Lys Lys Pro Tyr Phe Arg Phe Asp Asn Arg Asp Lys Glu Ile	
20 25 30	
aca gcg att agc ttc gat gag aag gca aac tta att tgg agt gga gac	144
Thr Ala Ile Ser Phe Asp Glu Lys Ala Asn Leu Ile Trp Ser Gly Asp	
35 40 45	
agc tat ggt tgc att tgc tca tat gat cca act ttt caa ctt tat aca	192
Ser Tyr Gly Cys Ile Ser Ser Tyr Asp Pro Thr Phe Gln Leu Tyr Thr	
50 55 60	
aga tat agg ggc cac ata ggt gga aat tcc gtg aag gat att ctc agt	240
Arg Tyr Arg Gly His Ile Gly Gly Asn Ser Val Lys Asp Ile Leu Ser	
65 70 75 80	
cat cgg gat ggt att tta tct att agt gaa gat tcc tta cac ttt gct	288

-continued

His Arg Asp Gly Ile Leu Ser Ile Ser Glu Asp Ser Leu His Phe Ala		
85	90	95
aat aga aga ggt gtt act aaa ttg aac ctc act agc att gat att gct		336
Asn Arg Arg Gly Val Thr Lys Leu Asn Leu Thr Ser Ile Asp Ile Ala		
100	105	110
gca ttt agc gaa ttg aac act atg tgc tat tct cct cat tca ctg aaa		384
Ala Phe Ser Glu Leu Asn Thr Met Cys Tyr Ser Pro His Ser Leu Lys		
115	120	125
aac aat atc tac tgt ggt gac aac aca aat tgg gga att gcg tcc		432
Asn Asn Ile Tyr Cys Gly Gly Asp Asn Thr Asn Trp Gly Ile Ala Ser		
130	135	140
att gac ttg aac aga ggt tgc tta gat tcc ctc ttg aat tac tca tct		480
Ile Asp Leu Asn Arg Gly Cys Leu Asp Ser Leu Leu Asn Tyr Ser Ser		
145	150	155
aaa gtg aag tta atg tgc tct aat aat aaa gtt ttg tct atc gga aga		528
Lys Val Lys Leu Met Cys Ser Asn Asn Lys Val Leu Ser Ile Gly Arg		
165	170	175
caa aca ggg act gtg gat ttg cta gat cca aca tcg aat cgt act atc		576
Gln Thr Gly Thr Val Asp Leu Leu Asp Pro Thr Ser Asn Arg Thr Ile		
180	185	190
aaa tca ttt aat gca cac tct gca tcc ata tcc gct atg gat tta cgg		624
Lys Ser Phe Asn Ala His Ser Ala Ser Ile Ser Ala Met Asp Leu Arg		
195	200	205
gat aac acc ttg gtt aca gta ggg aag tcc aaa aga ttt tat aac tta		672
Asp Asn Thr Leu Val Thr Val Gly Lys Ser Lys Arg Phe Tyr Asn Leu		
210	215	220
tac gct gac cca ttt gtg aat gtt tac gac ttg aga aca atg cgt caa		720
Tyr Ala Asp Pro Phe Val Asn Val Tyr Asp Leu Arg Thr Met Arg Gln		
225	230	235
ctc cct ctt tcc ttt tct aaa gga aca act atg gga tct gga ggc		768
Leu Pro Pro Val Ser Phe Ser Lys Gly Thr Thr Met Gly Ser Gly Gly		
245	250	255
gca gat ttt gtt caa tta cat cct ttg ctt cct act gtt atg atc gtc		816
Ala Asp Phe Val Gln Leu His Pro Leu Leu Pro Thr Val Met Ile Val		
260	265	270
gcc tca agt tct ggt tgc ttt gat ttc atc gac ctt tcc aat cca act		864
Ala Ser Ser Ser Gly Ser Phe Asp Phe Ile Asp Leu Ser Asn Pro Thr		
275	280	285
tta aga aca caa tat gtt cat cct tgc cag tgc att aaa aag tta tgt		912
Leu Arg Thr Gln Tyr Val His Pro Cys Gln Ser Ile Lys Lys Leu Cys		
290	295	300
ttg tcc ccc aat ggt gac gta ttg ggt ata cta gaa gct gat aat cac		960
Leu Ser Pro Asn Gly Asp Val Leu Gly Ile Leu Glu Ala Asp Asn His		
305	310	315
cta gat aca tgg aga aga tca tca aac aac atg gga atg ttt acc aat		1008
Leu Asp Thr Trp Arg Arg Ser Ser Asn Asn Met Gly Met Phe Thr Asn		
325	330	335
acc cct gaa atg cta gca tat cct gat tat ttt aat gac att acc tct		1056
Thr Pro Glu Met Leu Ala Tyr Pro Asp Tyr Phe Asn Asp Ile Thr Ser		
340	345	350
gac ggc cca ata tct gtc gac gat gaa aca tat cca ttg agt tct gtg		1104
Asp Gly Pro Ile Ser Val Asp Asp Glu Thr Tyr Pro Leu Ser Ser Val		
355	360	365
ggg atg ccg tac tat ctt gat aaa ctt ttg tct gca tgg ccc cct gta		1152
Gly Met Pro Tyr Tyr Leu Asp Lys Leu Leu Ser Ala Trp Pro Pro Val		
370	375	380
gtg ttt aaa agt gaa ggt acc ata ccg caa tta aca ggt aag tca ccc		1200

-continued

Val Phe Lys Ser Glu Gly Thr Ile Pro Gln Leu Thr Gly Lys Ser Pro 385 390 395 400		
tta cca tcg agc ggc aaa tta aaa agt aac ctt gct gtg atc tcg agc Leu Pro Ser Ser Gly Lys Leu Lys Ser Asn Leu Ala Val Ile Ser Ser 405 410 415		1248
caa aat gag aag ttg agc aca caa gaa ttt cct ttg tta aga tat gat Gln Asn Glu Lys Leu Ser Thr Gln Glu Phe Pro Leu Leu Arg Tyr Asp 420 425 430		1296
cgc acc aaa tac ggt atg aga aat gct ata cca gat tac gtt tgt cta Arg Thr Lys Tyr Gly Met Arg Asn Ala Ile Pro Asp Tyr Val Cys Leu 435 440 445		1344
aga gat ata agg aaa cag ata aca agc ggt tta gaa acc agc gat ata Arg Asp Ile Arg Lys Gln Ile Thr Ser Gly Leu Glu Thr Ser Asp Ile 450 455 460		1392
cag aca tat acc tca atc aac aag tac gaa gta ccc cct gca tac agt Gln Thr Tyr Thr Ser Ile Asn Lys Tyr Glu Val Pro Pro Ala Tyr Ser 465 470 475 480		1440
aga ctt cca ctg aca tca ggt aga ttt ggt act gat aat ttt gat ttt Arg Leu Pro Leu Thr Ser Gly Arg Phe Gly Thr Asp Asn Phe Asp Phe 485 490 495		1488
acg ccc ttt aat aac act gag tat tca gga ttg gat cca gat gtt gat Thr Pro Phe Asn Asn Thr Glu Tyr Ser Gly Leu Asp Pro Asp Val Asp 500 505 510		1536
aat cac tac aca aat gct atc ata caa ttg tat cgc ttt att cca gaa Asn His Tyr Thr Asn Ala Ile Ile Gln Leu Tyr Arg Phe Ile Pro Glu 515 520 525		1584
atg ttt aat ttc gtt gtt ggg tgg tgg aaa gac gag aat ttt gaa aca Met Phe Asn Phe Val Val Gly Cys Leu Lys Asp Glu Asn Phe Glu Thr 530 535 540		1632
acg ttg cta act gat cta ggc tac ctc ttt gac atg atg gaa aga tca Thr Leu Leu Thr Asp Leu Gly Tyr Leu Phe Asp Met Met Glu Arg Ser 545 550 555 560		1680
cat gga aaa ata tgt agt tct tcc aat ttt cag gcg tca ttg aaa tcc His Gly Lys Ile Cys Ser Ser Asn Phe Gln Ala Ser Leu Lys Ser 565 570 575		1728
tta act gat aaa aga caa tta gaa aac ggt gaa cca caa gaa cat tta Leu Thr Asp Lys Arg Gln Leu Glu Asn Gly Glu Pro Gln Glu His Leu 580 585 590		1776
gaa gag tat tta gaa tcg ctg tgc ata agg gaa agt atc gag gat ttt Glu Glu Tyr Leu Glu Ser Leu Cys Ile Arg Glu Ser Ile Glu Asp Phe 595 600 605		1824
aat tct tct gaa agt att aaa cgc aat atg cct caa aaa ttt aac aga Asn Ser Ser Glu Ser Ile Lys Arg Asn Met Pro Gln Lys Phe Asn Arg 610 615 620		1872
ttc ctg ctc tcg caa ctt att aaa gag gaa gcg cag aca gtc aac cat Phe Leu Leu Ser Gln Leu Ile Lys Glu Ala Gln Thr Val Asn His 625 630 635 640		1920
aat atc acc cta aat caa tgc ttt ggt ttg gaa acg gaa ata cga aca Asn Ile Thr Leu Asn Gln Cys Phe Gly Leu Glu Thr Glu Ile Arg Thr 645 650 655		1968
gag tgt agc tgt gat cac tac gac act acc gtc aaa ctt cta ccc tcc Glu Cys Ser Cys Asp His Tyr Asp Thr Thr Val Lys Leu Leu Pro Ser 660 665 670		2016
tta tca ata tca gga atc aac aaa acc gta atc aaa caa ttg aac aag Leu Ser Ile Ser Gly Ile Asn Lys Thr Val Ile Lys Gln Leu Asn Lys 675 680 685		2064
aaa agc aat gga cag aat att ttg cct tat att gaa tat gcc atg aaa		2112

-continued

Lys Ser Asn Gly Gln Asn Ile Leu Pro Tyr Ile Glu Tyr Ala Met Lys 690 695 700			
aat gta acc caa aag aac agt att tgc cca acc tgc ggc aaa acc gaa Asn Val Thr Gln Lys Asn Ser Ile Cys Pro Thr Cys Gly Lys Thr Glu 705 710 715 720		2160	
act atc acc cag gag tgt act gtc aag aat tta cct tca gtg ttg tca Thr Ile Thr Gln Glu Cys Thr Val Lys Asn Leu Pro Ser Val Leu Ser 725 730 735		2208	
tta gaa tta tca cta tta gat acc gaa ttt tcc aat ata agg tcg tcg Leu Glu Leu Ser Leu Leu Asp Thr Glu Phe Ser Asn Ile Arg Ser Ser 740 745 750		2256	
aaa aac tgg tta act agt gaa ttt tat gga agc atc att aaa aac aag Lys Asn Trp Leu Thr Ser Glu Phe Tyr Gly Ser Ile Ile Lys Asn Lys 755 760 765		2304	
gca gtt cta aga tcg acg gcg tcc gaa ttg aag ggc aca agc cac ata Ala Val Leu Arg Ser Thr Ala Ser Glu Leu Lys Gly Thr Ser His Ile 770 775 780		2352	
ttt aaa tac gaa ttg aat ggt tac gtg gct aaa atc act gat aac aat Phe Lys Tyr Glu Leu Asn Gly Tyr Val Ala Lys Ile Thr Asp Asn Asn 785 790 795 800		2400	
aac gag acg cgt cta gta aca tat gtc aaa aaa tat aat cca aaa gag Asn Glu Thr Arg Leu Val Thr Tyr Val Lys Lys Tyr Asn Pro Lys Glu 805 810 815		2448	
aat tgc ttc aag tgg ctc atg ttt aat gat tat ttg gtt gtt gag ata Asn Cys Phe Lys Trp Leu Met Phe Asn Asp Tyr Leu Val Val Glu Ile 820 825 830		2496	
aca gag gaa gag gcg ctt aaa atg aca tac cct tgg aaa aca cca gaa Thr Glu Glu Ala Leu Lys Met Thr Tyr Pro Trp Lys Thr Pro Glu 835 840 845		2544	
att atc ata tat tgt gat gcg gaa gaa tta cga aaa cct ttc ttt tct Ile Ile Ile Tyr Cys Asp Ala Glu Leu Arg Lys Pro Phe Phe Ser 850 855 860		2592	
gtt gat acg tat tcc atc aac tat gac ata ctt ttc cgt gat tat ttc Val Asp Thr Tyr Ser Ile Asn Tyr Asp Ile Leu Phe Arg Asp Tyr Phe 865 870 875 880		2640	
gca aac gga ata aga gat act gca aga cgt gaa tat aag tta tta aca Ala Asn Gly Ile Arg Asp Thr Ala Arg Arg Glu Tyr Lys Leu Thr 885 890 895		2688	
cat gat gag gca cct aaa tct gga acc ttg gtt gcc att gat gcc gaa His Asp Glu Ala Pro Lys Ser Gly Thr Leu Val Ala Ile Asp Ala Glu 900 905 910		2736	
ttt gtc tca tta caa agt gaa cta tgt gaa atc gat cat caa gga atc Phe Val Ser Leu Gln Ser Glu Leu Cys Glu Ile Asp His Gln Gly Ile 915 920 925		2784	
aga agt att att cga cct aaa aga act gct ttg gcc aga ata tcc att Arg Ser Ile Ile Arg Pro Lys Arg Thr Ala Leu Ala Arg Ile Ser Ile 930 935 940		2832	
att aga ggc gaa gaa gga gaa ctg tat gga gta cca ttt gtc gat gat Ile Arg Gly Glu Glu Gly Glu Leu Tyr Gly Val Pro Phe Val Asp Asp 945 950 955 960		2880	
tat gtg gta aac acg aac cac ata gaa gac tat ttg aca aga tat agt Tyr Val Val Asn Thr Asn His Ile Glu Asp Tyr Leu Thr Arg Tyr Ser 965 970 975		2928	
ggg att ctt cct ggt gac ttg gac cct gaa aag agt acc aaa agg ctt Gly Ile Leu Pro Gly Asp Leu Asp Pro Glu Lys Ser Thr Lys Arg Leu 980 985 990		2976	
gtg aga aga aac gtt gta tat cga aaa gtc tgg ctt tta atg cag ctc		3024	

-continued

Val Arg Arg Asn Val Val Tyr Arg Lys Val Trp Leu Leu Met Gln Leu			
995	1000	1005	
gga tgc gta ttt gtt ggt cat ggt ttg aat aat gac ttc aaa cac		3069	
Gly Cys Val Phe Val Gly His Gly Leu Asn Asn Asp Phe Lys His			
1010	1015	1020	
att aat att aat gtc cca aga aac caa att cgc gac act gcc ata		3114	
Ile Asn Ile Asn Val Pro Arg Asn Gln Ile Arg Asp Thr Ala Ile			
1025	1030	1035	
tat ttt cta caa gga aag aga tat ctt tca ttg cgt tat ctg gca		3159	
Tyr Phe Leu Gln Gly Lys Arg Tyr Leu Ser Leu Arg Tyr Leu Ala			
1040	1045	1050	
tat gtg ttg tta gga atg aat atc caa gag gga aat cac gat tca		3204	
Tyr Val Leu Leu Gly Met Asn Ile Gln Glu Gly Asn His Asp Ser			
1055	1060	1065	
att gaa gat gcc cat act gcc ttg att ctt tac aaa aaa tat ctc		3249	
Ile Glu Asp Ala His Thr Ala Leu Ile Leu Tyr Lys Lys Tyr Leu			
1070	1075	1080	
cac ctg aaa gaa aaa gct atc ttg gag aaa gta ctg aac agc gtg		3294	
His Leu Lys Glu Lys Ala Ile Phe Glu Lys Val Leu Asn Ser Val			
1085	1090	1095	
tac gaa gaa gga aga gcc cat aat ttc aaa gtt cca gaa act tca		3339	
Tyr Glu Glu Gly Arg Ala His Asn Phe Lys Val Pro Glu Thr Ser			
1100	1105	1110	
aag gga taa		3348	
Lys Gly			
1115			

<210> SEQ ID NO 12

<211> LENGTH: 1115

<212> TYPE: PRT

<213> ORGANISM: *Saccharomyces cereviciae*

<400> SEQUENCE: 12

Met Asn Asn Trp Gln His Phe Phe Asn Asn Pro Val Asp Leu Ser Glu				
1	5	10	15	
His Leu Lys Lys Pro Tyr Phe Arg Phe Asp Asn Arg Asp Lys Glu Ile				
20	25	30		
Thr Ala Ile Ser Phe Asp Glu Lys Ala Asn Leu Ile Trp Ser Gly Asp				
35	40	45		
Ser Tyr Gly Cys Ile Ser Ser Tyr Asp Pro Thr Phe Gln Leu Tyr Thr				
50	55	60		
Arg Tyr Arg Gly His Ile Gly Gly Asn Ser Val Lys Asp Ile Leu Ser				
65	70	75	80	
His Arg Asp Gly Ile Leu Ser Ile Ser Glu Asp Ser Leu His Phe Ala				
85	90	95		
Asn Arg Arg Gly Val Thr Lys Leu Asn Leu Thr Ser Ile Asp Ile Ala				
100	105	110		
Ala Phe Ser Glu Leu Asn Thr Met Cys Tyr Ser Pro His Ser Leu Lys				
115	120	125		
Asn Asn Ile Tyr Cys Gly Gly Asp Asn Thr Asn Trp Gly Ile Ala Ser				
130	135	140		
Ile Asp Leu Asn Arg Gly Cys Leu Asp Ser Leu Leu Asn Tyr Ser Ser				
145	150	155	160	
Lys Val Lys Leu Met Cys Ser Asn Asn Lys Val Leu Ser Ile Gly Arg				
165	170	175		

-continued

Gln	Thr	Gly	Thr	Val	Asp	Leu	Leu	Asp	Pro	Thr	Ser	Asn	Arg	Thr	Ile
180															190
Lys	Ser	Phe	Asn	Ala	His	Ser	Ala	Ser	Ile	Ser	Ala	Met	Asp	Leu	Arg
195															205
Asp	Asn	Thr	Leu	Val	Thr	Val	Gly	Lys	Ser	Lys	Arg	Phe	Tyr	Asn	Leu
210															220
Tyr	Ala	Asp	Pro	Phe	Val	Asn	Val	Tyr	Asp	Leu	Arg	Thr	Met	Arg	Gln
225															240
Leu	Pro	Pro	Val	Ser	Phe	Ser	Lys	Gly	Thr	Thr	Met	Gly	Ser	Gly	Gly
245															255
Ala	Asp	Phe	Val	Gln	Leu	His	Pro	Leu	Leu	Pro	Thr	Val	Met	Ile	Val
260															270
Ala	Ser	Ser	Ser	Gly	Ser	Phe	Asp	Phe	Ile	Asp	Leu	Ser	Asn	Pro	Thr
275															285
Leu	Arg	Thr	Gln	Tyr	Val	His	Pro	Cys	Gln	Ser	Ile	Lys	Lys	Leu	Cys
290															300
Leu	Ser	Pro	Asn	Gly	Asp	Val	Leu	Gly	Ile	Leu	Glu	Ala	Asp	Asn	His
305															320
Leu	Asp	Thr	Trp	Arg	Arg	Ser	Ser	Asn	Asn	Met	Gly	Met	Phe	Thr	Asn
325															335
Thr	Pro	Glu	Met	Leu	Ala	Tyr	Pro	Asp	Tyr	Phe	Asn	Asp	Ile	Thr	Ser
340															350
Asp	Gly	Pro	Ile	Ser	Val	Asp	Asp	Glu	Thr	Tyr	Pro	Leu	Ser	Ser	Val
355															365
Gly	Met	Pro	Tyr	Tyr	Leu	Asp	Lys	Leu	Leu	Ser	Ala	Trp	Pro	Pro	Val
370															380
Val	Phe	Lys	Ser	Glu	Gly	Thr	Ile	Pro	Gln	Leu	Thr	Gly	Lys	Ser	Pro
385															400
Leu	Pro	Ser	Ser	Gly	Lys	Leu	Lys	Ser	Asn	Leu	Ala	Val	Ile	Ser	Ser
405															415
Gln	Asn	Glu	Lys	Leu	Ser	Thr	Gln	Glu	Phe	Pro	Leu	Leu	Arg	Tyr	Asp
420															430
Arg	Thr	Lys	Tyr	Gly	Met	Arg	Asn	Ala	Ile	Pro	Asp	Tyr	Val	Cys	Leu
435															445
Arg	Asp	Ile	Arg	Lys	Gln	Ile	Thr	Ser	Gly	Leu	Glu	Thr	Ser	Asp	Ile
450															460
Gln	Thr	Tyr	Thr	Ser	Ile	Asn	Lys	Tyr	Glu	Val	Pro	Pro	Ala	Tyr	Ser
465															480
Arg	Leu	Pro	Leu	Thr	Ser	Gly	Arg	Phe	Gly	Thr	Asp	Asn	Phe	Asp	Phe
485															495
Thr	Pro	Phe	Asn	Asn	Thr	Glu	Tyr	Ser	Gly	Leu	Asp	Pro	Asp	Val	Asp
500															510
Asn	His	Tyr	Thr	Asn	Ala	Ile	Ile	Gln	Leu	Tyr	Arg	Phe	Ile	Pro	Glu
515															525
Met	Phe	Asn	Phe	Val	Val	Gly	Cys	Leu	Lys	Asp	Glu	Asn	Phe	Glu	Thr
530															540
Thr	Leu	Leu	Thr	Asp	Leu	Gly	Tyr	Leu	Phe	Asp	Met	Met	Glu	Arg	Ser
545															560
His	Gly	Lys	Ile	Cys	Ser	Ser	Ser	Asn	Phe	Gln	Ala	Ser	Leu	Lys	Ser
565															575
Leu	Thr	Asp	Lys	Arg	Gln	Leu	Glu	Asn	Gly	Glu	Pro	Gln	Glu	His	Leu

-continued

580	585	590	
Glu Glu Tyr Leu Glu Ser Leu Cys Ile Arg Glu Ser Ile Glu Asp Phe			
595	600	605	
Asn Ser Ser Glu Ser Ile Lys Arg Asn Met Pro Gln Lys Phe Asn Arg			
610	615	620	
Phe Leu Leu Ser Gln Leu Ile Lys Glu Glu Ala Gln Thr Val Asn His			
625	630	635	640
Asn Ile Thr Leu Asn Gln Cys Phe Gly Leu Glu Thr Glu Ile Arg Thr			
645	650	655	
Glu Cys Ser Cys Asp His Tyr Asp Thr Thr Val Lys Leu Leu Pro Ser			
660	665	670	
Leu Ser Ile Ser Gly Ile Asn Lys Thr Val Ile Lys Gln Leu Asn Lys			
675	680	685	
Lys Ser Asn Gly Gln Asn Ile Leu Pro Tyr Ile Glu Tyr Ala Met Lys			
690	695	700	
Asn Val Thr Gln Lys Asn Ser Ile Cys Pro Thr Cys Gly Lys Thr Glu			
705	710	715	720
Thr Ile Thr Gln Glu Cys Thr Val Lys Asn Leu Pro Ser Val Leu Ser			
725	730	735	
Leu Glu Leu Ser Leu Leu Asp Thr Glu Phe Ser Asn Ile Arg Ser Ser			
740	745	750	
Lys Asn Trp Leu Thr Ser Glu Phe Tyr Gly Ser Ile Ile Lys Asn Lys			
755	760	765	
Ala Val Leu Arg Ser Thr Ala Ser Glu Leu Lys Gly Thr Ser His Ile			
770	775	780	
Phe Lys Tyr Glu Leu Asn Gly Tyr Val Ala Lys Ile Thr Asp Asn Asn			
785	790	795	800
Asn Glu Thr Arg Leu Val Thr Tyr Val Lys Lys Tyr Asn Pro Lys Glu			
805	810	815	
Asn Cys Phe Lys Trp Leu Met Phe Asn Asp Tyr Leu Val Val Glu Ile			
820	825	830	
Thr Glu Glu Ala Leu Lys Met Thr Tyr Pro Trp Lys Thr Pro Glu			
835	840	845	
Ile Ile Ile Tyr Cys Asp Ala Glu Glu Leu Arg Lys Pro Phe Phe Ser			
850	855	860	
Val Asp Thr Tyr Ser Ile Asn Tyr Asp Ile Leu Phe Arg Asp Tyr Phe			
865	870	875	880
Ala Asn Gly Ile Arg Asp Thr Ala Arg Arg Glu Tyr Lys Leu Leu Thr			
885	890	895	
His Asp Glu Ala Pro Lys Ser Gly Thr Leu Val Ala Ile Asp Ala Glu			
900	905	910	
Phe Val Ser Leu Gln Ser Glu Leu Cys Glu Ile Asp His Gln Gly Ile			
915	920	925	
Arg Ser Ile Ile Arg Pro Lys Arg Thr Ala Leu Ala Arg Ile Ser Ile			
930	935	940	
Ile Arg Gly Glu Glu Gly Glu Leu Tyr Gly Val Pro Phe Val Asp Asp			
945	950	955	960
Tyr Val Val Asn Thr Asn His Ile Glu Asp Tyr Leu Thr Arg Tyr Ser			
965	970	975	
Gly Ile Leu Pro Gly Asp Leu Asp Pro Glu Lys Ser Thr Lys Arg Leu			
980	985	990	

-continued

```

Val Arg Arg Asn Val Val Tyr Arg Lys Val Trp Leu Leu Met Gln Leu
995 1000 1005

Gly Cys Val Phe Val Gly His Gly Leu Asn Asn Asp Phe Lys His
1010 1015 1020

Ile Asn Ile Asn Val Pro Arg Asn Gln Ile Arg Asp Thr Ala Ile
1025 1030 1035

Tyr Phe Leu Gln Gly Lys Arg Tyr Leu Ser Leu Arg Tyr Leu Ala
1040 1045 1050

Tyr Val Leu Leu Gly Met Asn Ile Gln Glu Gly Asn His Asp Ser
1055 1060 1065

Ile Glu Asp Ala His Thr Ala Leu Ile Leu Tyr Lys Lys Tyr Leu
1070 1075 1080

His Leu Lys Glu Lys Ala Ile Phe Glu Lys Val Leu Asn Ser Val
1085 1090 1095

Tyr Glu Glu Gly Arg Ala His Asn Phe Lys Val Pro Glu Thr Ser
1100 1105 1110

Lys Gly
1115

```

```

<210> SEQ ID NO: 13
<211> LENGTH: 2040
<212> TYPE: DNA
<213> ORGANISM: Saccharomyces cereviciae
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)..(2040)
<223> OTHER INFORMATION: Pan3

```

<400> SEQUENCE: 13

atg gac aaa atc aat cct gat tgg gcg aag gat att ccg tgc aga aat	48
Met Asp Lys Ile Asn Pro Asp Trp Ala Lys Asp Ile Pro Cys Arg Asn	
1 5 10 15	
atc act att tat ggc tac tgc aaa aag gag aaa gaa ggt tgc cct ttc	96
Ile Thr Ile Tyr Gly Tyr Cys Lys Glu Lys Glu Gly Cys Pro Phe	
20 25 30	
aaa cac agc gat aac act acc gct act acc ata aat gac gtt cct cct	144
Lys His Ser Asp Asn Thr Thr Ala Thr Ile Asn Asp Val Pro Pro	
35 40 45	
cca ata gat gtg ggt gag gct aca act ccg acc atg aca tca gtt cct	192
Pro Ile Asp Val Gly Glu Ala Thr Thr Pro Thr Met Thr Ser Val Pro	
50 55 60	
aag ttc aat gct aaa gta tcc gca agt ttc act ccg atg aca gtc ggt	240
Lys Phe Asn Ala Lys Val Ser Ala Ser Phe Thr Pro Met Thr Val Gly	
65 70 75 80	
agt gac tcc tta acc act gtg acg aat acc acc tcc gct gct aca aat	288
Ser Asp Ser Leu Thr Thr Val Thr Asn Thr Ser Ala Ala Thr Asn	
85 90 95	
gct act ggc aat atc gcc atg gca gct acc tct gct act gct tct aca	336
Ala Thr Gly Asn Ile Ala Met Ala Ala Thr Ser Ala Thr Ala Ser Thr	
100 105 110	
gtt aat ccg atg att aat ccg ata gtt aat agc tcg tta gtg aat aac	384
Val Asn Pro Met Ile Asn Pro Ile Val Asn Ser Ser Leu Val Asn Asn	
115 120 125	
aat aac aat aat agt aat ata agc ata tca ata cca act acc gct tcg	432
Asn Asn Asn Ser Asn Ile Ser Ile Ser Ile Pro Thr Thr Ala Ser	
130 135 140	

-continued

agt tcc aat tac gac ccc ttc aat gcc ccc att ttc act ccg tct tca Ser Ser Asn Tyr Asp Pro Phe Asn Ala Pro Ile Phe Thr Pro Ser Ser 145 150 155 160	480
acc tcc tca att cac act aat gca aat gca cat tct ttt cca ttt ccc Thr Ser Ser Ile His Thr Asn Ala Asn Ala His Ser Phe Pro Phe Pro 165 170 175	528
tcc att gca aat tct ggt ggc ata aat ata aac gcc act gat gat aat Ser Ile Ala Asn Ser Gly Gly Ile Asn Ile Asn Ala Thr Asp Asp Asn 180 185 190	576
agt aac aat atg agt atg gct aat aat gtg cca cct cct atg caa ccg Ser Asn Asn Met Ser Met Ala Asn Asn Val Pro Pro Pro Met Gln Pro 195 200 205	624
cca ccc ata gag agt agt aat ctt aag tac cca cgt att tat ccg cct Pro Pro Ile Glu Ser Ser Asn Leu Lys Tyr Pro Arg Ile Tyr Pro Pro 210 215 220	672
cct cac agt ctt cta cag tat cac cta tat gca cct gaa cag cca tca Pro His Ser Leu Leu Gln Tyr His Leu Tyr Ala Pro Glu Gln Pro Ser 225 230 235 240	720
tca ttg aaa tca tta tta aag cct aat gaa agg tct gca gat cag ctt Ser Leu Lys Ser Leu Leu Lys Pro Asn Glu Arg Ser Ala Asp Gln Leu 245 250 255	768
ttc att cca aac aat att aga gaa gat tta acc aag aaa aac tta tcg Phe Ile Pro Asn Asn Ile Arg Glu Asp Leu Thr Lys Lys Asn Leu Ser 260 265 270	816
att ttg cag gtt ttc ccc tct tca ggt aaa gtt ata cca agt att gta Ile Leu Gln Val Phe Pro Ser Ser Gly Lys Val Ile Pro Ser Ile Val 275 280 285	864
caa gat tat ttt aat ttg gtt cca ttg aac ttc aat aat aac gat ttt Gln Asp Tyr Phe Asn Leu Val Pro Leu Asn Phe Asn Asn Asn Asp Phe 290 295 300	912
tta aat aaa act acg ctc ttc aaa gtt ttt tcc aat tat gac ggt aaa Leu Asn Lys Thr Thr Leu Phe Lys Val Phe Ser Asn Tyr Asp Gly Lys 305 310 315 320	960
gcc tac gtt ttg aag agg ctt cct aac atc gat aag tca atg aat cca Ala Tyr Val Leu Lys Arg Leu Pro Asn Ile Asp Lys Ser Met Asn Pro 325 330 335	1008
aac aaa ata tcc aaa ata tat cag ata tgg tca aaa att aat tgt aca Asn Lys Ile Ser Lys Ile Tyr Gln Ile Trp Ser Lys Ile Asn Cys Thr 340 345 350	1056
aat ttg ata aag ttt agg gac att ttt caa act act aaa ttt ggt gat Asn Leu Ile Lys Phe Arg Asp Ile Phe Gln Thr Thr Lys Phe Gly Asp 355 360 365	1104
ctg tct att tgt ttg gtc ttt gac tac tac cca aac tcg cta tct ttg Leu Ser Ile Cys Leu Val Phe Asp Tyr Tyr Pro Asn Ser Leu Ser Leu 370 375 380	1152
tat gat tac cac ttt gtt aat ttc cct aag ttt cca ata acg aat aat Tyr Asp Tyr His Phe Val Asn Phe Pro Lys Phe Pro Ile Thr Asn Asn 385 390 395 400	1200
tat tta tgg ata tat tta gtt caa ctc acc aat gta ata aac tct atc Tyr Leu Trp Ile Tyr Leu Val Gln Leu Thr Asn Val Ile Asn Ser Ile 405 410 415	1248
cat tca caa aac ttg agt att ggc aat aca tta aac tgg aga aaa gtt His Ser Gln Asn Leu Ser Ile Gly Asn Thr Leu Asn Trp Arg Lys Val 420 425 430	1296
ttg att act ggg gac cca ggg aga atc aag tta tca cac tgc aat ttt Leu Ile Thr Gly Asp Pro Gly Arg Ile Lys Leu Ser His Cys Asn Phe 435 440 445	1344

-continued

atg gac ctt ttg ttc aat gat gat act gat acc gta gta tct tcc ggc Met Asp Leu Leu Phe Asn Asp Asp Thr Asp Thr Val Val Ser Ser Gly 450 455 460	1392
gga agt acc ata gag gga caa caa cag cta gac tac aaa tat tta gga Gly Ser Thr Ile Glu Gly Gln Gln Leu Asp Tyr Lys Tyr Leu Gly 465 470 475 480	1440
gag cta tta ttt aac cta tcc att aat att gaa aac tct aat aac aac Glu Leu Leu Phe Asn Leu Ser Ile Asn Ile Glu Asn Ser Asn Asn Asn 485 490 495	1488
act gcc cct aaa gaa tat cga ttg gag gaa ata acc cct caa tca att Thr Ala Pro Lys Glu Tyr Arg Leu Glu Glu Ile Thr Pro Gln Ser Ile 500 505 510	1536
gat gac atg aga cag atc gat gat aag ttc aag gat gta ctc aag tat Asp Asp Met Arg Gln Ile Asp Asp Lys Phe Lys Asp Val Leu Lys Tyr 515 520 525	1584
ctg ata tca gac aac ggc gat tcc aaa aag agc att cat gat ctt act Leu Ile Ser Asp Asn Gly Asp Ser Lys Lys Ser Ile His Asp Leu Thr 530 535 540	1632
agt cac ttt tat gat aag atg ttc atg gtc ctg gaa tcg tca caa acc Ser His Phe Tyr Asp Lys Met Phe Met Val Leu Glu Ser Ser Gln Thr 545 550 555 560	1680
tat aca gaa tac atg gag tct gtc tta tca aga gaa cta gaa aat ggc Tyr Thr Glu Tyr Met Glu Ser Val Leu Ser Arg Glu Leu Glu Asn Gly 565 570 575	1728
aga tta ttt agg ctg gtc aac aag cta aat tgc att ttt ggt aga atc Arg Leu Phe Arg Leu Val Asn Lys Leu Asn Cys Ile Phe Gly Arg Ile 580 585 590	1776
gaa tca aga ata gac ata aat tgg tcc gaa tct ggg act aaa ttc ccc Glu Ser Arg Ile Asp Ile Asn Trp Ser Glu Ser Gly Thr Lys Phe Pro 595 600 605	1824
att ata cta ttt tat gac tac gta ttc cat caa gtg gat tcg aat ggg Ile Ile Leu Phe Tyr Asp Tyr Val Phe His Gln Val Asp Ser Asn Gly 610 615 620	1872
aaa cca ata atg gat tta act cat gtc cta aga tgt ttg aac aaa tta Lys Pro Ile Met Asp Leu Thr His Val Leu Arg Cys Leu Asn Lys Leu 625 630 635 640	1920
gac gct ggt att caa gaa aag tta atg ttg gta acg ccc gat gag tta Asp Ala Gly Ile Gln Glu Lys Leu Met Leu Val Thr Pro Asp Glu Leu 645 650 655	1968
aac tgt att att ata tcc tat aag gag ttg aag gac ttg ata gaa tcc Asn Cys Ile Ile Ile Ser Tyr Lys Glu Leu Lys Asp Leu Ile Glu Ser 660 665 670	2016
acc ttt cga tcc atc acc caa taa Thr Phe Arg Ser Ile Thr Gln 675	2040

<210> SEQ ID NO 14

<211> LENGTH: 679

<212> TYPE: PRT

<213> ORGANISM: *Saccharomyces cereviciae*

<400> SEQUENCE: 14

Met Asp Lys Ile Asn Pro Asp Trp Ala Lys Asp Ile Pro Cys Arg Asn
1 5 10 15

Ile Thr Ile Tyr Gly Tyr Cys Lys Lys Glu Lys Glu Gly Cys Pro Phe
20 25 30

Lys His Ser Asp Asn Thr Thr Ala Thr Ile Asn Asp Val Pro Pro
35 40 45

-continued

Pro Ile Asp Val Gly Glu Ala Thr Thr Pro Thr Met Thr Ser Val Pro
 50 55 60
 Lys Phe Asn Ala Lys Val Ser Ala Ser Phe Thr Pro Met Thr Val Gly
 65 70 75 80
 Ser Asp Ser Leu Thr Thr Val Thr Asn Thr Thr Ser Ala Ala Thr Asn
 85 90 95
 Ala Thr Gly Asn Ile Ala Met Ala Ala Thr Ser Ala Thr Ala Ser Thr
 100 105 110
 Val Asn Pro Met Ile Asn Pro Ile Val Asn Ser Ser Leu Val Asn Asn
 115 120 125
 Asn Asn Asn Ser Asn Ile Ser Ile Ser Ile Pro Thr Thr Ala Ser
 130 135 140
 Ser Ser Asn Tyr Asp Pro Phe Asn Ala Pro Ile Phe Thr Pro Ser Ser
 145 150 155 160
 Thr Ser Ser Ile His Thr Asn Ala Asn Ala His Ser Phe Pro Phe Pro
 165 170 175
 Ser Ile Ala Asn Ser Gly Gly Ile Asn Ile Asn Ala Thr Asp Asp Asn
 180 185 190
 Ser Asn Asn Met Ser Met Ala Asn Asn Val Pro Pro Pro Met Gln Pro
 195 200 205
 Pro Pro Ile Glu Ser Ser Asn Leu Lys Tyr Pro Arg Ile Tyr Pro Pro
 210 215 220
 Pro His Ser Leu Leu Gln Tyr His Leu Tyr Ala Pro Glu Gln Pro Ser
 225 230 235 240
 Ser Leu Lys Ser Leu Leu Lys Pro Asn Glu Arg Ser Ala Asp Gln Leu
 245 250 255
 Phe Ile Pro Asn Asn Ile Arg Glu Asp Leu Thr Lys Lys Asn Leu Ser
 260 265 270
 Ile Leu Gln Val Phe Pro Ser Ser Gly Lys Val Ile Pro Ser Ile Val
 275 280 285
 Gln Asp Tyr Phe Asn Leu Val Pro Leu Asn Phe Asn Asn Asn Asp Phe
 290 295 300
 Leu Asn Lys Thr Thr Leu Phe Lys Val Phe Ser Asn Tyr Asp Gly Lys
 305 310 315 320
 Ala Tyr Val Leu Lys Arg Leu Pro Asn Ile Asp Lys Ser Met Asn Pro
 325 330 335
 Asn Lys Ile Ser Lys Ile Tyr Gln Ile Trp Ser Lys Ile Asn Cys Thr
 340 345 350
 Asn Leu Ile Lys Phe Arg Asp Ile Phe Gln Thr Thr Lys Phe Gly Asp
 355 360 365
 Leu Ser Ile Cys Leu Val Phe Asp Tyr Tyr Pro Asn Ser Leu Ser Leu
 370 375 380
 Tyr Asp Tyr His Phe Val Asn Phe Pro Lys Phe Pro Ile Thr Asn Asn
 385 390 395 400
 Tyr Leu Trp Ile Tyr Leu Val Gln Leu Thr Asn Val Ile Asn Ser Ile
 405 410 415
 His Ser Gln Asn Leu Ser Ile Gly Asn Thr Leu Asn Trp Arg Lys Val
 420 425 430
 Leu Ile Thr Gly Asp Pro Gly Arg Ile Lys Leu Ser His Cys Asn Phe
 435 440 445

-continued

Met	Asp	Leu	Leu	Phe	Asn	Asp	Asp	Thr	Asp	Thr	Val	Val	Ser	Ser	Gly
450				455							460				
Gly	Ser	Thr	Ile	Glu	Gly	Gln	Gln	Leu	Asp	Tyr	Lys	Tyr	Leu	Gly	
465				470				475					480		
Glu	Leu	Leu	Phe	Asn	Leu	Ser	Ile	Asn	Ile	Glu	Asn	Ser	Asn	Asn	
							485		490			495			
Thr	Ala	Pro	Lys	Glu	Tyr	Arg	Leu	Glu	Glu	Ile	Thr	Pro	Gln	Ser	Ile
							500		505			510			
Asp	Asp	Met	Arg	Gln	Ile	Asp	Asp	Lys	Phe	Lys	Asp	Val	Leu	Lys	Tyr
							515		520			525			
Leu	Ile	Ser	Asp	Asn	Gly	Asp	Ser	Lys	Lys	Ser	Ile	His	Asp	Leu	Thr
							530		535			540			
Ser	His	Phe	Tyr	Asp	Lys	Met	Phe	Met	Val	Leu	Glu	Ser	Ser	Gln	Thr
							545		550			555		560	
Tyr	Thr	Glu	Tyr	Met	Glu	Ser	Val	Leu	Ser	Arg	Glu	Leu	Glu	Asn	Gly
							565		570			575			
Arg	Leu	Phe	Arg	Leu	Val	Asn	Lys	Leu	Asn	Cys	Ile	Phe	Gly	Arg	Ile
							580		585			590			
Glu	Ser	Arg	Ile	Asp	Ile	Asn	Trp	Ser	Glu	Ser	Gly	Thr	Lys	Phe	Pro
							595		600			605			
Ile	Ile	Leu	Phe	Tyr	Asp	Tyr	Val	Phe	His	Gln	Val	Asp	Ser	Asn	Gly
							610		615			620			
Lys	Pro	Ile	Met	Asp	Leu	Thr	His	Val	Leu	Arg	Cys	Leu	Asn	Lys	Leu
							625		630			635		640	
Asp	Ala	Gly	Ile	Gln	Glu	Lys	Leu	Met	Leu	Val	Thr	Pro	Asp	Glu	Leu
							645		650			655			
Asn	Cys	Ile	Ile	Ile	Ser	Tyr	Lys	Glu	Leu	Lys	Asp	Leu	Ile	Glu	Ser
							660		665			670			
Thr	Phe	Arg	Ser	Ile	Thr	Gln									
							675								

<210> SEQ ID NO 15																
<211> LENGTH: 666																
<212> TYPE: DNA																
<213> ORGANISM: <i>Saccharomyces cereviciae</i>																
<220> FEATURE:																
<221> NAME/KEY: CDS																
<222> LOCATION: (1)..(666)																
<223> OTHER INFORMATION: Rpb4																
<400> SEQUENCE: 15																
atg	aat	gtt	tct	aca	tca	acc	ttt	caa	aca	aga	cgg	aga	aga	ttg	aag	
Met	Asn	Val	Ser	Thr	Ser	Thr	Phe	Gln	Thr	Arg	Arg	Arg	Arg	Leu	Lys	
1							5		10			15				
aaa	gtg	gag	gaa	gaa	aat	gca	gct	act	cta	caa	ctg	ggc	cag	gaa		
Lys	Val	Glu	Glu	Glu	Asn	Ala	Ala	Thr	Leu	Gln	Leu	Gly	Gln	Glu		
							20		25			30				
ttc	cag	ctg	aaa	cag	ata	aat	cat	cag	ggt	gaa	gag	gag	gaa	ttg	att	
Phe	Gln	Leu	Lys	Gln	Ile	Asn	His	Gln	Gly	Glu	Glu	Glu	Glu	Leu	Ile	
							35		40			45				
gcc	ttg	aac	cta	agt	gaa	gcc	agg	tta	gta	atc	aaa	gaa	gct	ctt	gta	
Ala	Leu	Asn	Leu	Ser	Glu	Ala	Arg	Leu	Val	Ile	Lys	Glu	Ala	Leu	Val	
							50		55			60				
gaa	cgt	agg	aga	gca	ttt	aaa	aga	tcg	caa	aaa	aaa	cac	aag	aag		
Glu	Arg	Arg	Arg	Ala	Phe	Lys	Arg	Ser	Gln	Lys	Lys	His	Lys	Lys		
							65		70			75		80		

-continued

cat ttg aag cac gaa aac gcc aat gat gaa act acg gca gta gag gat His Leu Lys His Glu Asn Ala Asn Asp Glu Thr Thr Ala Val Glu Asp 85 90 95	288
gaa gat gat gat ctg gat gaa gat gac gtc aac gct gat gat gat gat Glu Asp Asp Asp Leu Asp Glu Asp Asp Val Asn Ala Asp Asp Asp Asp 100 105 110	336
ttt atg cat tct gaa act agg gag aag gag ttc gag tct atc gac gtt Phe Met His Ser Glu Thr Arg Glu Lys Glu Leu Glu Ser Ile Asp Val 115 120 125	384
ctg tta gaa cag aca acg gga gga aat aat aaa gat ttg aaa aat acc Leu Leu Glu Gln Thr Thr Gly Gly Asn Asn Lys Asp Leu Lys Asn Thr 130 135 140	432
atg cag tat tta aca aat ttc tcc cga ttt aga gac caa gaa acc gtc Met Gln Tyr Leu Thr Asn Phe Ser Arg Phe Arg Asp Gln Glu Thr Val 145 150 155 160	480
ggg gca gtt ata cag ctt ctg aaa agc act ggg tta cat cct ttt gaa Gly Ala Val Ile Gln Leu Leu Lys Ser Thr Gly Leu His Pro Phe Glu 165 170 175	528
gtg gcg caa cta ggt tct ttg gcc tgt gac aca gct gat gaa gca aag Val Ala Gln Leu Gly Ser Leu Ala Cys Asp Thr Ala Asp Glu Ala Lys 180 185 190	576
act tta att cca agc tta aac aat aaa ata tca gac gat gag ttg gaa Thr Leu Ile Pro Ser Leu Asn Asn Lys Ile Ser Asp Asp Glu Leu Glu 195 200 205	624
agg ata cta aag gaa ttg tca aac cta gaa aca ctc tat taa Arg Ile Leu Lys Glu Leu Ser Asn Leu Glu Thr Leu Tyr 210 215 220	666

<210> SEQ ID NO 16

<211> LENGTH: 221

<212> TYPE: PRT

<213> ORGANISM: *Saccharomyces cereviciae*

<400> SEQUENCE: 16

Met Asn Val Ser Thr Ser Thr Phe Gln Thr Arg Arg Arg Arg Leu Lys 1 5 10 15
Lys Val Glu Glu Glu Asn Ala Ala Thr Leu Gln Leu Gly Gln Glu 20 25 30
Phe Gln Leu Lys Gln Ile Asn His Gln Gly Glu Glu Glu Leu Ile 35 40 45
Ala Leu Asn Leu Ser Glu Ala Arg Leu Val Ile Lys Glu Ala Leu Val 50 55 60
Glu Arg Arg Ala Phe Lys Arg Ser Gln Lys Lys His Lys Lys Lys 65 70 75 80
His Leu Lys His Glu Asn Ala Asn Asp Glu Thr Thr Ala Val Glu Asp 85 90 95
Glu Asp Asp Asp Leu Asp Glu Asp Asp Val Asn Ala Asp Asp Asp Asp 100 105 110
Phe Met His Ser Glu Thr Arg Glu Lys Glu Leu Glu Ser Ile Asp Val 115 120 125
Leu Leu Glu Gln Thr Thr Gly Gly Asn Asn Lys Asp Leu Lys Asn Thr 130 135 140
Met Gln Tyr Leu Thr Asn Phe Ser Arg Phe Arg Asp Gln Glu Thr Val 145 150 155 160
Gly Ala Val Ile Gln Leu Leu Lys Ser Thr Gly Leu His Pro Phe Glu

-continued

165	170	175
Val Ala Gln Leu Gly Ser Leu Ala Cys Asp Thr Ala Asp Glu Ala Lys 180	185	190
Thr Leu Ile Pro Ser Leu Asn Asn Lys Ile Ser Asp Asp Glu Leu Glu 195	200	205
Arg Ile Leu Lys Glu Leu Ser Asn Leu Glu Thr Leu Tyr 210	215	220

<210> SEQ ID NO 17

<211> LENGTH: 516

<212> TYPE: DNA

<213> ORGANISM: *Saccharomyces cereviciae*

<220> FEATURE:

<221> NAME/KEY: CDS

<222> LOCATION: (1)..(516)

<223> OTHER INFORMATION: Rpb7

<400> SEQUENCE: 17

atg ttt ttt att aaa gac ctt tcg ctt aat att acc ctt cat ccg tcc Met Phe Phe Ile Lys Asp Leu Ser Leu Asn Ile Thr Leu His Pro Ser 1 5 10 15	48
ttt ttc ggt cct cga atg aag caa tat cta aag aca aaa cta ttg gaa Phe Phe Gly Pro Arg Met Lys Gln Tyr Leu Lys Thr Lys Leu Leu Glu 20 25 30	96
gag gtt gaa ggt tcg tgt acg ggt aaa ttc gga tat att ctt tgt gtc Glu Val Glu Gly Ser Cys Thr Gly Lys Phe Gly Tyr Ile Leu Cys Val 35 40 45	144
cta gac tat gat aat ata gat att caa cgt ggg aga ata ttg ccc aca Leu Asp Tyr Asp Asn Ile Asp Ile Gln Arg Gly Arg Ile Leu Pro Thr 50 55 60	192
gat ggg tcc gcc gag ttc aac gtg aaa tat aga gct gta gtt ttc aaa Asp Gly Ser Ala Glu Phe Asn Val Lys Tyr Arg Ala Val Val Phe Lys 65 70 75 80	240
cca ttt aaa ggg gaa gta gtg gac ggc aca gtc gtt tca tgt tct cag Pro Phe Lys Gly Glu Val Val Asp Gly Thr Val Val Ser Cys Ser Gln 85 90 95	288
cac ggg ttc gaa gtg caa gta ggt cca atg aaa gta ttt gtg aca aag His Gly Phe Glu Val Gln Val Gly Pro Met Lys Val Phe Val Thr Lys 100 105 110	336
cat ctg atg cct caa gat tta acc ttt aat gcg ggt tca aac cca cca His Leu Met Pro Gln Asp Leu Thr Phe Asn Ala Gly Ser Asn Pro Pro 115 120 125	384
tca tac caa agt tcc gag gat gtc atc acc ata aaa agt aga att aga Ser Tyr Gln Ser Ser Glu Asp Val Ile Thr Ile Lys Ser Arg Ile Arg 130 135 140	432
gtt aaa att gaa ggt tgt atc agt caa gtg agt tct att cac gca atc Val Lys Ile Glu Gly Cys Ile Ser Gln Val Ser Ser Ile His Ala Ile 145 150 155 160	480
ggg agt atc aaa gaa gat tat ttg ggt gct att taa Gly Ser Ile Lys Glu Asp Tyr Leu Gly Ala Ile 165 170	516

<210> SEQ ID NO 18

<211> LENGTH: 171

<212> TYPE: PRT

<213> ORGANISM: *Saccharomyces cereviciae*

<400> SEQUENCE: 18

Met Phe Phe Ile Lys Asp Leu Ser Leu Asn Ile Thr Leu His Pro Ser

-continued

1	5	10	15												
Phe	Phe	Gly	Pro	Arg	Met	Lys	Gln	Tyr	Leu	Lys	Thr	Lys	Leu	Leu	Glu
20	25														
30															
Glu	Val	Glu	Gly	Ser	Cys	Thr	Gly	Lys	Phe	Gly	Tyr	Ile	Leu	Cys	Val
35	40														
45															
Leu	Asp	Tyr	Asp	Asn	Ile	Asp	Ile	Gln	Arg	Gly	Arg	Ile	Leu	Pro	Thr
50	55														
55															
Asp	Gly	Ser	Ala	Glu	Phe	Asn	Val	Lys	Tyr	Arg	Ala	Val	Val	Phe	Lys
65	70														
75															
80															
Pro	Phe	Lys	Gly	Glu	Val	Val	Asp	Gly	Thr	Val	Val	Ser	Cys	Ser	Gln
85	90														
95															
His	Gly	Phe	Glu	Val	Gln	Val	Gly	Pro	Met	Lys	Val	Phe	Val	Thr	Lys
100	105														
110															
His	Leu	Met	Pro	Gln	Asp	Leu	Thr	Phe	Asn	Ala	Gly	Ser	Asn	Pro	Pro
115	120														
125															
Ser	Tyr	Gln	Ser	Ser	Glu	Asp	Val	Ile	Thr	Ile	Lys	Ser	Arg	Ile	Arg
130	135														
140															
Val	Lys	Ile	Glu	Gly	Cys	Ile	Ser	Gln	Val	Ser	Ser	Ile	His	Ala	Ile
145	150														
155															
Gly	Ser	Ile	Lys	Glu	Asp	Tyr	Leu	Gly	Ala	Ile					
165	170														

<210> SEQ_ID NO 19
<211> LENGTH: 743
<212> TYPE: PRT
<213> ORGANISM: Pichia pastoris
<220> FEATURE:
<221> NAME/KEY: peptide
<222> LOCATION: (1)..(743)
<223> OTHER INFORMATION: CCR4
<400> SEQUENCE: 19
Met Asn Thr Ala Ser Gln Tyr Lys Ala Gln Gly Asn Gln Gly Ile
1 5 10 15
Met Arg His Ala Pro Gln Pro Gln Gln Thr Gln Ile Pro Ser Gln Tyr
20 25 30
Leu Leu Gln Leu Gln Gly Ser His Ala Thr Asn Gln Pro Gln Thr
35 40 45
Asp Ser Ser Asn Pro Ser Leu Leu Val Gln Ala Leu Phe Gln Gly Pro
50 55 60
Gln Ser His Gln Gly Gln Arg Ile Gly Asn Gln Gln Met Gln Pro His
65 70 75 80
Val Val Leu Gly Ser Gln Thr Leu Gln Ser Ser Ser Val Val Pro Pro
85 90 95
Gly Gln Asn Gln Pro Thr Ser Gln Ser Pro His Tyr Asp Thr Ile Gln
100 105 110
Val Glu Asp Pro Thr Thr Phe His Trp Gln His Gln Val Gln Leu Val
115 120 125
Gln Met Ser Arg Lys Cys Asn Gln Pro His Phe Tyr Ala Arg His Ala
130 135 140
Ala Val Ser Ser Arg Lys Leu Leu Gly Lys Asn Gly Asn Asn Asn Met
145 150 155 160
Pro Gly Val Asp Gly Ser Gln Gly Asn Gln Pro Pro Asn Leu Leu Glu
165 170 175

-continued

Ala Thr Lys Thr Leu Leu Met Ser Thr Asp Pro Thr Ala Glu Gln Ala
 180 185 190
 Glu Ser Gly Glu Asn Lys Ser Ala Thr Asp Pro Leu Leu Trp Phe Lys
 195 200 205
 Lys Leu Ser Asn Asp Thr Gln Asp Glu Asp Glu Glu Ile Asp Ala Ser
 210 215 220
 Thr Leu Lys Asp Pro Thr Arg Gln Leu Trp Arg Ala Leu Asp Leu Ser
 225 230 235 240
 Gly Gln Gln Leu Leu His Leu Ser Glu Lys Leu Phe Arg Tyr Asp Phe
 245 250 255
 Leu Thr Lys Leu Tyr Leu Asn Gly Asn Gly Leu Thr Glu Leu Pro Ser
 260 265 270
 Ser Ile Arg Gln Leu Lys Ser Leu Thr Val Leu Asp Val Ser Gln Asn
 275 280 285
 Leu Leu Ser Ser Phe Pro Pro Glu Leu Gly Ile Leu Phe Asn Leu Arg
 290 295 300
 Tyr Ile Tyr Ala Phe Asp Asn Arg Leu Thr Asp Ile Pro Phe Glu Phe
 305 310 315 320
 Gly Asn Leu Tyr Glu Leu Glu Phe Leu Gly Ile Glu Gly Asn Val Asn
 325 330 335
 Met Asn Pro Glu Tyr Val Asn Ile Leu Ala Lys Arg Gly Ser Arg Gly
 340 345 350
 Leu Thr Ile His Leu Arg Asp Asn Ala Pro Arg Pro Thr Pro Pro Lys
 355 360 365
 Ser Arg Gln Trp Ile Tyr Phe Ser Asn Asp Gly Glu Ile Ile Glu Glu
 370 375 380
 Gln Glu Tyr Arg Gln Gln Gln Thr Glu Asp Asp Ile Val Asn Thr Phe
 385 390 395 400
 Thr Met Met Thr Tyr Asn Thr Leu Cys Gln His Tyr Ala Thr Lys Lys
 405 410 415
 Met Tyr Arg Tyr Thr Pro Ser Trp Ala Leu Asp Trp Asp Tyr Arg Arg
 420 425 430
 Glu Arg Leu Lys Glu Gln Ile Leu Asp Leu Gln Thr Asp Ile Ile Cys
 435 440 445
 Leu Gln Glu Val Glu His Lys Thr Phe Asp Asp Phe Trp Gln Pro Ile
 450 455 460
 Met Leu Ser His Gly Tyr Lys Gly Ile Phe His Val Lys Ser Arg Ala
 465 470 475 480
 Lys Thr Met Lys Glu Ser Ser Ala Tyr Lys Val Asp Gly Cys Ala Thr
 485 490 495
 Phe Tyr Arg Thr Ser Lys Phe Gln Ala Val Glu Arg Lys His Phe Glu
 500 505 510
 Tyr Gly Arg Ile Ala Met Ser Gln Asp Lys Phe Lys Lys Thr Glu Asp
 515 520 525
 Leu Phe Asn Arg Phe Leu Asn Lys Asp Asn Ile Ala Ser Val Leu Ile
 530 535 540
 Leu Glu His Ile Pro Ser Gly Asn Lys Leu Val Val Ala Asn Thr His
 545 550 555 560
 Leu His Trp Asp Pro Glu Phe Asn Asp Val Lys Thr Met Gln Val Gly
 565 570 575

-continued

Val	Leu	Leu	Asp	Glu	Leu	Gln	Ala	Val	Ile	Arg	Lys	His	Leu	Ser	Pro
580															590
Lys Asp Ile Thr Lys Val Pro Leu Leu Ile Cys Gly Asp Phe Asn Ser															
595															605
Lys Val His Ser Ala Val Tyr Gln Leu Phe Ser Gln Gly Thr Val Asp															
610															620
Lys His Glu Asp Ile Ile Gly Arg Asp Tyr Gly Lys Phe Thr Glu Glu															
625															640
Gly Phe Arg His Pro Phe His Leu Gln Ser Ser Tyr Asp Ser Ile Gly															
645															655
Glu Leu Pro Tyr Thr Asn Val Ser Pro Thr Phe Thr Asp Val Ile Asp															
660															670
Tyr Ile Trp Tyr Ser Thr Pro Ser Leu Ser Val Lys Gly Val Leu Gly															
675															685
Gln Val Asp Pro Asp Tyr Ser Lys Asn Ile Ile Gly Phe Pro Asn Ala															
690															700
Asp Phe Pro Ser Asp His Ile Pro Leu Leu Ser Thr Phe Met Phe Lys															
705															720
Lys Ser Ser Ala Pro Arg Pro Asp Thr Arg Val Asp Phe Arg Ser Asp															
725															735
Phe Arg Gly Ser Arg Lys Thr															
740															

<210>	SEQ_ID_NO	20													
<211>	LENGTH:	1118													
<212>	TYPE:	PRT													
<213>	ORGANISM:	Pichia pastoris													
<220>	FEATURE:														
<221>	NAME/KEY:	peptide													
<222>	LOCATION:	(1)..(1118)													
<223>	OTHER INFORMATION:	Pan2													
<400> SEQUENCE: 20															
Met	Glu	Gly	Trp	Asn	Glu	Ile	Asn	Arg	Val	Pro	Cys	Met	Arg	Pro	Ser
1						5			10			15			
Phe Arg Ile Ser Asp Pro Thr Arg Gln Ser Val Pro Ser Ala Leu Leu															
						20			25			30			
Phe Asp Asp Ser His Asp Leu Thr Trp Ile Gly Ser Glu Asp Gly Phe															
						35			40			45			
Val Lys Ser Leu Ala Asp Gln Met Leu Thr Pro Tyr Thr Ser Phe Arg															
						50			55			60			
Cys His Ser Ser Lys Val Leu Gln Leu Leu Asn Asn Lys Arg Gly Ile															
						65			70			75			80
Leu Ser Leu Ser Glu Asn Ser Ile Lys Leu Thr Ser Arg Thr Gly Leu															
						85			90			95			
Cys Arg Met Asn Leu His Asp Ile Ser Val Asn His Ser Arg Ser Met															
						100			105			110			
Ala Tyr Thr Ser Asn Thr Glu Asn Glu Val Leu Leu Gly Gly Phe Gln															
						115			120			125			
Gly Lys Leu Val Lys Leu Asn Leu Met Arg Gly Glu Ile Ser Asp Thr															
						130			135			140			
Ile Pro Tyr Asp Ala Pro Val Tyr Thr Met Ser Ala Asn Leu Ser Gln															
						145			150			155			160
Ile Cys Leu Gly Arg Val Asp Gly Thr Val Asp Ile Leu Asp Pro Arg															

-continued

165	170	175
Ser Asn Asp Ile Ile Ala Ser Phe Ser Gly His Met Arg Thr Leu Ser		
180	185	190
Ser Met Asp Cys Arg Gly Asn Thr Leu Ile Thr Thr Gly Tyr Ser Val		
195	200	205
Arg Asn Gly Thr Phe Tyr Ala Asp Pro Ile Ala Asn Leu Tyr Asp Leu		
210	215	220
Arg Thr Lys Ser Leu Met Pro Pro Val Thr Phe Pro Ala Gly Ala Ser		
225	230	235
Phe Val Lys Leu His Pro Lys Leu Pro Asn Val Ala Ile Leu Ala Ser		
245	250	255
Ser Ala Gly Leu Ile His Phe Val Asn Leu Tyr Ser Pro Met Asn Val		
260	265	270
Ser Leu Tyr Gln Ala Asp Val Gly Ser Tyr Met Asn Asn Phe Glu Val		
275	280	285
Ser Pro Ser Gly Asp Phe Met Gly Phe Thr Asp Ser Phe Gln Asn Val		
290	295	300
His Leu Trp Ser Asn Ala Pro Asp Leu Ser Ser Ala His Ile Pro Asn		
305	310	315
Leu Ala Ser Ser Leu Glu Gln Ala Thr Met Leu Pro Pro Pro Ala Asn		
325	330	335
Ser Asp Ile Ile Pro Ala Asp Ser Asp Ser Thr Val Pro Leu Ser Ser		
340	345	350
Ile Gly Met Pro Tyr Tyr Asp Gln Pro Leu Leu Ser Asn Trp Pro Phe		
355	360	365
Asp Met Lys Phe Ser Leu Gly His Val Pro Lys Lys Ile Asn Pro Glu		
370	375	380
Leu Ile Asn Ser Ser Lys Ser Val Val Glu Asn Tyr Asn Ala His Ala		
385	390	395
Phe Gly Gln Asn Asn Leu Pro Ser Gly Met Gly Val Asn Met Asn Ile		
405	410	415
Pro Gly Val Phe Lys Thr Asp Thr Asp Thr Arg Ile Ala Val Tyr Ser		
420	425	430
Arg Glu Lys Tyr Gly Pro Arg Asn Val Ala Gln Pro Tyr Tyr Arg Leu		
435	440	445
His Asp Arg Ser Asn Gln Lys Asn Lys Phe Val Pro Lys Phe Ile Ser		
450	455	460
Glu Arg Leu Asp Val Asp Glu Lys Ser Ser Ser Thr Ala Glu Asn Thr		
465	470	475
Asn Ser Asp Phe Lys Asp Thr Leu Tyr Glu Gln Leu Phe Asp Cys Lys		
485	490	495
Pro His Ser Ser Gln Glu Val Pro Tyr Cys Phe Thr Lys Leu Asp Ile		
500	505	510
His Tyr Ser Lys Phe Gly Val His Asp Phe Asp Phe Glu Phe Phe Asn		
515	520	525
Lys Thr Asn Leu Ser Gly Leu Glu Ser His Val Gln Ser Ala Tyr Cys		
530	535	540
Asn Ser Leu Leu Gln Leu Tyr Arg Phe Ala Pro Leu Met Phe Asn Tyr		
545	550	555
Ala Val Gly Ser Leu Ala Glu Glu Ser Val Asp Asn Asn Ser Leu Leu		
565	570	575

-continued

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

 His Ile Ala Pro Ser Asn Phe Gln Lys Met Leu Ser Lys Ile Pro Glu
 595 600 605

 Ala Ser Arg Ile Leu Asp Ile Tyr Asn Asn Gly Tyr Asn Arg Asp Glu
 610 615 620

 His Ser Gln Arg Val Leu Ile Gln Ser Phe Asn Asn Phe Leu Leu Glu
 625 630 635 640

 Arg Ile Ser Leu Asp Glu Met Asn Gln Ala Lys Ser Asn Ala Pro His
 645 650 655

 Ile Phe Asn Thr Ile Met Gly Ile Pro Val Glu Cys Glu Phe Ile Gly
 660 665 670

 Val Gly Cys Gly Leu Arg Lys Ile Gly Asn Thr Thr Leu Tyr Ser Leu
 675 680 685

 Asp Val Arg His Pro Lys Thr Asn Asn Val Val Leu Asn Lys Lys Leu
 690 695 700

 Asn His Ser Ile Ile Pro Tyr Ile Glu Leu Ala Leu Ser Arg Asn Leu
 705 710 715 720

 Ser Val Asn Met Ala Cys Asp Asn Cys Asn Glu Thr His Ala Phe Asp
 725 730 735

 Thr Leu Tyr Thr Val Lys Asp Leu Pro Pro Ile Leu Ser Leu Asn Leu
 740 745 750

 Glu Leu Thr Asn Gln Glu Leu Ser Glu Leu Lys Ala Ser Lys Thr Asn
 755 760 765

 Trp Leu Ala Ser Glu Phe Tyr Ala Ser Met Asn Lys Gly Arg Ile Cys
 770 775 780

 Leu Lys Ser Ile Thr Thr Gly Phe Arg His Leu Lys Tyr Glu Leu Leu
 785 790 795 800

 Gly Tyr Val Ala Gln Val Thr Asp Lys Glu Gly Asn Ser Asn Leu Val
 805 810 815

 Thr Phe Val Lys Val Gly Lys Asp Glu Trp Phe Leu Phe Asn Asp Phe
 820 825 830

 Leu Val Met Pro Ile Ser Glu His Glu Val Leu Asn Leu Asn Tyr Trp
 835 840 845

 Trp Lys Lys Pro Val Ile Val Ile Tyr Lys Asn Ser Glu Ser Thr Asn
 850 855 860

 Met Phe Asp Tyr Glu Gly Trp Arg Gln Asn Leu Asn Gln Asp Ile Leu
 865 870 875 880

 Tyr Arg Asp His Phe Ser Arg Gly Thr Arg Glu Gly Lys Ile Ile Glu
 885 890 895

 Tyr Glu Leu Leu Thr Lys Glu Glu Ala Pro Gln Pro Gly Thr Leu Val
 900 905 910

 Ala Ile Asp Ala Glu Phe Val Val Ile Glu Pro Glu Leu Val Glu Phe
 915 920 925

 Asn Ser Asp Gly Thr Lys Lys Val Ile Arg Pro Leu Lys Asn Ser Leu
 930 935 940

 Ala Arg Val Ser Val Leu Arg Gly Asp Thr Gly Pro Lys Glu Gly Ile
 945 950 955 960

 Pro Phe Ile Asp Asp Tyr Val Ile Ile Glu Glu Pro Ile Asn Asp Tyr
 965 970 975

-continued

Leu	Thr	Ser	Trp	Ser	Gly	Ile	Glu	Pro	Asp	Asp	Leu	Asn	Leu	Glu	Lys
980						985								990	
<hr/>															
Ser	Lys	Arg	Ser	Leu	Thr	Thr	Leu	Gln	Ala	Val	Tyr	Arg	Lys	Leu	Trp
995							1000							1005	
<hr/>															
Leu	Leu	Leu	Asn	Leu	Gly	Cys	Ile	Phe	Val	Gly	His	Gly	Leu	Ile	
1010						1015							1020		
<hr/>															
Asn	Asp	Phe	Arg	Thr	Ile	Asn	Leu	Ser	Val	Pro	Lys	Gln	Gln	Val	
1025						1030							1035		
<hr/>															
Arg	Asp	Thr	Ala	Glu	Leu	Tyr	Phe	Leu	Lys	Lys	Glu	Lys	Arg	Lys	
1040						1045							1050		
<hr/>															
Leu	Ser	Leu	Lys	Phe	Leu	Thr	Tyr	Ala	Val	Leu	Arg	Arg	Glu	Val	
1055						1060							1065		
<hr/>															
Gln	Lys	Gly	Asn	His	Asp	Ser	Ile	Glu	Asp	Ala	Lys	Ala	Ala	Leu	
1070						1075							1080		
<hr/>															
Met	Leu	Tyr	Arg	Lys	Tyr	Ile	Gln	Leu	Asn	Asn	Thr	Gly	Glu	Leu	
1085						1090							1095		
<hr/>															
Gln	His	Thr	Leu	Glu	Glu	Val	Tyr	Met	Glu	Gly	Gln	Met	Leu	Ser	
1100						1105							1110		
<hr/>															
Phe	Lys	Val	Pro	Thr											
															1115

<210> SEQ ID NO 21
<211> LENGTH: 614
<212> TYPE: PRT
<213> ORGANISM: *Pichia pastoris*
<220> FEATURE:
<221> NAME/KEY: peptide
<222> LOCATION: (1)..(614)
<223> OTHER INFORMATION: Pan3

<400> SEQUENCE: 21

Met	Ser	Ala	Asn	Ser	Asn	Ser	Thr	Asn	Thr	Glu	Asn	His	Asp	Trp	Ala
1							5		10					15	
<hr/>															
Lys	Asp	Ile	Asp	Cys	Lys	Asn	Val	Phe	Ile	His	Gly	Tyr	Cys	Lys	Phe
							20		25					30	
<hr/>															
Glu	Asn	Lys	Gly	Cys	Tyr	Phe	Lys	His	Pro	Ser	Asp	Asp	Ser	Ile	Glu
							35		40					45	
<hr/>															
Lys	Asn	Lys	Glu	Thr	Lys	Glu	Thr	Lys	Ser	Ser	Ala	Pro	Leu	Ser	Ser
							50		55					60	
<hr/>															
Gly	Thr	Glu	Gly	Ser	Ser	Ile	Arg	Lys	Lys	Phe	Ser	Phe	Glu	Ala	
							65		70					80	
<hr/>															
Pro	Ser	Phe	Thr	Pro	Ser	Gly	Ser	Val	Ala	Ser	Leu	Thr	Asn	Lys	Phe
							85		90					95	
<hr/>															
Ser	Thr	Met	Ser	Pro	Lys	Leu	Asp	Glu	Ile	Pro	Thr	Phe	Val	Pro	Thr
							100		105					110	
<hr/>															
Ser	Gln	Gln	Thr	Pro	Lys	Lys	Pro	Gln	Gln	Val	Ser	Ala	Glu	Ser	Asn
							115		120					125	
<hr/>															
Asp	Tyr	Phe	Ala	Pro	Pro	Asn	Ser	Met	Phe	Thr	Pro	Ser	Val	Pro	Ser
							130		135					140	
<hr/>															
Asp	Asn	Ser	Gln	Pro	Gly	Leu	Pro	Leu	Ser	Ala	Gln	Phe	Pro	Gln	His
							145		150					160	
<hr/>															
Gly	Thr	Thr	Ala	Pro	His	Asn	Val	Val	Pro	Glu	Met	Phe	Tyr	Ala	Gln
							165		170					175	
<hr/>															
Ala	Thr	Gln	Tyr	Pro	Leu	Asn	Tyr	Asn	Leu	Tyr	Ala	Pro	Pro	Pro	

-continued

180	185	190
Pro His Phe Asn Leu His Leu Asn Pro Asn Glu Arg Thr Val Ser Ser		
195	200	205
Phe Phe Ile Asp Asp Asn Leu Arg Glu Ser Leu Gln Lys Arg Asn Glu		
210	215	220
Ala Cys Leu Gln Thr Phe Ser Asn Pro Ser Ile Pro Asp Ile Val Gly		
225	230	235
240		
Val Tyr His Ser Leu Val Pro Leu Asn Asn Phe Asp Asn Asn Ser		
245	250	255
Ala Arg Tyr Gly Ala Val Ser His Met Tyr Lys Ala Thr Ser Asn Lys		
260	265	270
Asp Ala Arg Leu Tyr Ala Leu Arg Arg Ile Glu Asn Val Asn Ile Thr		
275	280	285
Asp Lys Gln Ala Phe Lys Thr Ile Lys Ala Trp Ser Ser Ile Glu Asn		
290	295	300
Ser Asn Ile Val Lys Val His Glu Ala Phe Thr Thr Thr Val Phe Gly		
305	310	315
320		
Gly Asn Ser Leu Val Val Ala Tyr Asp Phe Tyr Gly Asn Ala Lys Thr		
325	330	335
Leu Ile Glu Ile His Phe Gln Asn Gln Pro Glu Leu Ile Thr Glu Gln		
340	345	350
His Leu Trp Ser Tyr Leu Ile Gln Leu Val Asn Ala Leu Asn Glu Val		
355	360	365
His Asp Lys Gly Leu Ala Val Arg Ser Ile Asp Leu Ser Lys Val Ile		
370	375	380
Val Thr Asn Lys Asn Arg Ile Lys Leu Ser Gly Cys Gly Ile Val Asp		
385	390	395
400		
Ile Leu Gln His Glu Asn Thr Glu Asp Ile Ala Gln Leu Gln Lys Lys		
405	410	415
Asp Leu Glu Leu Leu Ala Lys Leu Leu Tyr Asp Leu Ser Ile Thr Ser		
420	425	430
Ile Tyr Gly Ser Val Ser Phe Asp Asp Lys Thr Gln Asp Gln Ile Ile		
435	440	445
Asp His Leu Lys Phe Ser Asp Asp Tyr Lys Ala Thr Leu Lys Tyr Leu		
450	455	460
Ile Ser Ala Asp Phe Asn Leu Lys Glu Ile Gln Thr Arg Ile Ala Pro		
465	470	475
480		
Arg Leu Leu Asp Val Ile Asp Gly Leu Gln Asn Ser Asn Asp Phe Ile		
485	490	495
Glu Ser Gln Leu Ser Thr Glu Leu Glu Asn Ala Arg Leu Val Arg Leu		
500	505	510
Met Ser Lys Leu Ser Phe Leu His Gly Arg Pro Glu His Glu Gly Asp		
515	520	525
Pro Asn Trp Ser Glu Ser Ser Ala Asn Tyr Pro Leu Thr Leu Phe Val		
530	535	540
Asp Tyr Val Tyr His Gln Val Asp Glu Arg Gly Tyr Pro Val Val Asp		
545	550	555
560		
Leu Ala His Val Ile Thr Cys Leu Asn Lys Leu Asp Ala Gly Ile Glu		
565	570	575
Glu Arg Ile Leu Leu Val Ser Lys Asp Glu Lys Asn Cys Ile Ile Ile		
580	585	590

-continued

Ser Tyr Lys Glu Leu Lys Thr Leu Ile Glu Glu Gly Phe Asn Glu Leu
595 600 605

Arg Met Arg Lys Asp Ile
610

<210> SEQ ID NO 22

<211> LENGTH: 186

<212> TYPE: PRT

<213> ORGANISM: *Pichia pastoris*

<220> FEATURE:

<221> NAME/KEY: peptide

<222> LOCATION: (1)..(186)

<223> OTHER INFORMATION: Rpb4

<400> SEQUENCE: 22

Met Asn Val Ser Thr Ser Thr Val Gly Ala Arg Arg Arg Arg Ala Lys
1 5 10 15

Gln Gln Val Asp Asp Glu Glu Asn Ala Thr Leu Leu Arg Leu Gly Pro
20 25 30

Glu Phe Ala Leu Lys Gln Tyr Asp His Asp Gly Asn Glu His Asp Leu
35 40 45

Ile Ala Leu Ser Leu Ser Glu Ser Arg Leu Leu Ile Arg Glu Ala Leu
50 55 60

Lys Ala Arg Ser Arg Ala Arg Asn Gly Gly Val Asp Ile Glu Ser Ser
65 70 75 80

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

Val Ala Asn Gly Val Val Lys Lys Thr Leu Asp Tyr Leu Asn Thr Phe
100 105 110

Ala Arg Phe Lys Asp Glu Glu Thr Cys Thr Ala Val Asp Gln Leu Leu
115 120 125

His Asn Ser Ser Asp Cys Ser Val Leu His Pro Phe Glu Ile Ala Gln
130 135 140

Leu Ser Ser Leu Gly Cys Glu Asp Val Asp Glu Ala Ile Thr Leu Ile
145 150 155 160

Pro Ser Leu Ala Ala Lys Lys Glu Val Asn Leu Gln Arg Ile Leu Asp
165 170 175

Glu Leu Asn Arg Leu Glu Asp Pro Tyr Lys
180 185

<210> SEQ ID NO 23

<211> LENGTH: 171

<212> TYPE: PRT

<213> ORGANISM: *Pichia pastoris*

<220> FEATURE:

<221> NAME/KEY: peptide

<222> LOCATION: (1)..(171)

<223> OTHER INFORMATION: Rpb7

<400> SEQUENCE: 23

Met Phe Phe Leu Lys Asp Leu Ser Leu Ile Leu Thr Leu His Pro Ser
1 5 10 15

Tyr Phe Gly Pro Gln Met Asn Gln Tyr Leu Arg Glu Lys Leu Leu Thr
20 25 30

Asp Val Glu Gly Thr Cys Thr Gly Gln Phe Gly Tyr Ile Val Thr Val
35 40 45

-continued

Leu	Asp	Gly	Met	Asn	Ile	Asp	Val	Gly	Lys	Gly	Arg	Ile	Ile	Pro	Gly	
50									60							
Ser	Gly	Ser	Ala	Glu	Phe	Glu	Val	Lys	Tyr	Arg	Ala	Val	Val	Trp	Lys	
65									75						80	
Pro	Phe	Lys	Gly	Glu	Val	Val	Asp	Ala	Ile	Val	Ser	Asn	Val	Ser	Pro	
									85			90			95	
Ile	Gly	Phe	Ala	Asp	Val	Gly	Pro	Leu	Asn	Val	Phe	Val	Ser	Thr		
								100			105			110		
Arg	Leu	Ile	Pro	Asp	Asn	Leu	Val	Tyr	Asn	Pro	Ser	Asn	Ser	Pro	Pro	
								115			120			125		
Ala	Tyr	Met	Ser	Asn	Asp	Glu	Leu	Ile	Thr	Lys	Gly	Ser	Lys	Val	Arg	
								130			135			140		
Leu	Lys	Val	Val	Gly	Thr	Arg	Thr	Asp	Val	Asn	Glu	Ile	Tyr	Ala	Ile	
								145			150			155		160
Gly	Ser	Ile	Lys	Glu	Asp	Phe	Leu	Gly	Ala	Ile						
								165			170					

<210> SEQ ID NO 24																
<211> LENGTH: 656																
<212> TYPE: PRT																
<213> ORGANISM: Aspergillus niger																
<220> FEATURE:																
<221> NAME/KEY: peptide																
<222> LOCATION: (1) .. (656)																
<223> OTHER INFORMATION: CCR4																
<400> SEQUENCE: 24																
Met	Asn	Gly	Gly	Gln	Ala	His	Gln	Arg	Phe	Gly	Met	Gln	Ile	Pro	Lys	
1									5			10			15	
Phe	Gln	Ser	Gln	Ser	His	His	Pro	His	Pro	Ala	Gln	Gln	Ala	His	His	
									20			25			30	
His	Ala	His	His	Asn	Gln	Ala	Ser	His	Ser	Ile	Asn	His	Gln	His	Asn	
									35			40			45	
Phe	Ser	Ser	Gly	Ala	Leu	Ala	Ala	Ala	Thr	Pro	His	Phe	Thr	Pro	Gly	
									50			55			60	
Pro	Leu	Gln	Asn	Gly	Ala	His	Val	Asn	Val	Asp	Glu	Asp	Ile	Asp	Glu	
									65			70			80	
Thr	Met	Asn	Glu	His	Trp	Gln	Gln	Leu	Gln	Leu	Ala	Ala	Glu	Ser		
									85			90			95	
Arg	Gln	Ala	Ser	Ser	Pro	His	Tyr	Tyr	Ala	Arg	Thr	Val	Ala	Gln	Gln	
									100			105			110	
Thr	Lys	Gly	Ile	Gln	Ile	Ala	Pro	Ser	Gln	Pro	Glu	Ser	Gln	Glu	Asn	
									115			120			125	
Gly	Ser	Gly	Asp	Arg	Asn	Gly	Leu	Val	Lys	Ser	Lys	Pro	Ala	Pro	Arg	
									130			135			140	
Gln	Gly	Trp	His	Ala	Leu	Asp	Phe	Gly	Gly	Gln	Gly	Leu	Arg	Ala	Leu	
									145			150			160	
Ala	Thr	Ser	Leu	Phe	His	Tyr	Thr	Phe	Leu	Glu	Lys	Leu	Tyr	Leu	Asn	
									165			170			175	
His	Asn	Lys	Leu	Lys	Thr	Leu	Pro	Pro	Ala	Ile	Gly	Gln	Leu	Arg	Lys	
									180			185			190	
Leu	Thr	His	Leu	Asp	Leu	Ser	Ser	Asn	Asp	Ile	Ser	Glu	Leu	Pro	Glu	
									195			200			205	
Glu	Ile	Gly	Met	Leu	Thr	Ser	Leu	Lys	Gln	Leu	Leu	Leu	Phe	Asp	Asn	

-continued

210	215	220
Asn Ile Arg Thr Leu Pro Phe Glu Met Gly Tyr Leu Tyr Arg Leu Glu		
225	230	235
Met Leu Gly Ile Glu Gly Asn Pro Leu Asn Asp Val Leu Lys Ser Gln		
245	250	255
Ile Ile Lys Glu Gly Thr Lys Ala Leu Val Arg Tyr Leu Arg Glu Glu		
260	265	270
Met Pro Val His Leu Pro Pro Asp Arg Asp Trp Ile Ile Leu Asp		
275	280	285
Glu Thr Ala Ser Ser Ser Asn Ser Pro Thr Glu Lys Ile Thr Val Leu		
290	295	300
Ser His Asn Ala Leu Cys Asp Ser Ser Ala Thr Pro Ser His Phe Gly		
305	310	315
Tyr Thr Pro Ser Arg Val Leu Ser Trp Glu Phe Arg Arg Glu Leu Ile		
325	330	335
Leu Ser Glu Leu Arg Ser His Asp Ser Asp Ile Ile Cys Leu Gln Glu		
340	345	350
Ile Asp Gln Gly Ser Tyr Asn Gly Phe Phe Arg Glu Gln Leu Ala Tyr		
355	360	365
Asn Asp Tyr Lys Gly Val Tyr Trp Pro Arg Gly Arg Ala Met Gly Met		
370	375	380
Gln Glu Glu Ala Lys Ser Val Asp Gly Cys Ala Thr Phe Phe Lys		
385	390	395
Gly Ser Lys Phe Ile Leu Leu Asp Lys Gln Met Ile Asn Phe Gly Gln		
405	410	415
Thr Ala Val Arg Arg Pro Asp Ala Lys Gly Gln Asp Asp Ile Tyr Asn		
420	425	430
Arg Leu Trp Gln Lys Asp His Ile Ala Val Val Ile Phe Leu Glu Asn		
435	440	445
Arg Leu Thr Gly Ser Arg Phe Ile Val Val Asn Ala His Leu Tyr Trp		
450	455	460
Asp Pro Ala Phe Lys Asp Val Lys Leu Ile Gln Thr Ala Ile Leu Met		
465	470	475
Glu Glu Ile Thr Lys Leu Ser Glu Lys Tyr Ala Lys Phe Pro Pro Cys		
485	490	495
Thr Asp Lys Thr Ala Phe Arg Phe Ser Glu Ala Glu Val Glu Tyr Ala		
500	505	510
Ser Gly Asp Gln Ile Pro Leu Phe Met Cys Gly Asp Phe Asn Ser Ala		
515	520	525
Pro Gly Ser Ala Ala Tyr Asn Leu Val Ala His Gly Arg Leu Thr Glu		
530	535	540
Ser His Pro Asp Leu Glu Lys Arg Leu Tyr Gly Asn Leu Ser Arg Val		
545	550	555
Gly Met Thr His Pro Phe Lys Leu Lys Ser Ala Tyr Asn Ser Ile Gly		
565	570	575
Glu Leu Ser Phe Thr Asn Tyr Thr Pro Asp Phe Lys Asp Ile Leu Asp		
580	585	590
Tyr Ile Trp Tyr Thr Ser Asn Thr Leu His Val Ser Ala Leu Leu Gly		
595	600	605
Glu Val Asp Lys Glu Tyr Leu Gln Lys Val Pro Gly Phe Pro Asn Phe		
610	615	620

-continued

His Phe Pro Ser Asp His Val Ala Leu Phe Ala Glu Phe Thr Val Lys
625 630 635 640

Gly Lys Lys Gly Lys Val Val Glu Ala Asp Phe Gly Pro Gln Arg Asn
645 650 655

<210> SEQ ID NO 25

<211> LENGTH: 1150

<212> TYPE: PRT

<213> ORGANISM: Aspergillus niger

<220> FEATURE:

<221> NAME/KEY: peptide

<222> LOCATION: (1)..(1150)

<223> OTHER INFORMATION: Pan2

<400> SEQUENCE: 25

Met Glu Ala Asp Trp Asp Glu Leu Ser Arg Ile Pro Val Pro Ala Pro
1 5 10 15

Ser Val His Ala Leu Pro Thr Ile Ala Thr Ala Ile Ala Phe Asp Asp
20 25 30

Val Met Glu Leu Leu Trp Gly Arg Ile Thr Ser Phe Phe Gly Pro Glu
35 40 45

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

Arg Gln Ile Ile Phe His Asp Arg Gly Val Ile Ser Leu Ser Pro Lys
65 70 75 80

Ser Val His Met Ile Thr Arg Arg Gly Leu Thr Gln Trp His Ile Ala
85 90 95

His Glu Glu Met Thr Asp Leu Arg Cys Met Ser Phe Thr Ala Gln Thr
100 105 110

Asn Arg Ile Ile Val Ala Gly Cys Gln Lys Ser Met Phe Thr Ile Asp
115 120 125

Ile Asp Lys Gly Ile Ile Ile Asp Lys Leu His Thr Glu Tyr Asn Tyr
130 135 140

Thr Ile Met Lys Lys Ser Arg Tyr Leu Cys Ala Ala Thr Asp Thr Gly
145 150 155 160

Ser Val Asn Ala Leu Ser Leu Asn Asp Phe Ser Val Val Lys Ser Trp
165 170 175

Lys Ala His Gly Thr Ala Val Asn Asp Met Asp Ala Arg Asn Asp Leu
180 185 190

Leu Val Thr Cys Gly Phe Ser Val Arg His Leu Gly Ser Pro Ile Val
195 200 205

Asp Pro Leu Ala Asn Val Tyr Asp Leu Lys Thr Leu Ser Pro Leu Pro
210 215 220

Pro Ile Pro Phe His Ala Gly Ala Ala Tyr Val Arg Met His Pro Lys
225 230 235 240

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

Val Asp Leu Met Asn Pro Asn Ala Ile Asn Leu Arg Gln Ala Asn Val
260 265 270

Ser Phe Met Leu Gly Ile Asp Leu Ser Pro Ser Gly Glu Ala Leu Ala
275 280 285

Ile Asn Asp Ala Glu Cys Ala Ile His Leu Trp Gly Ser Pro Ala Lys
290 295 300

-continued

Val His Phe Asn Glu Met Ser Lys Glu Ala Glu Phe Gly Asp Val Ala		
305	310	315
320		
Pro Arg Pro Pro Thr Leu Asp Trp Ser Pro Glu Thr Pro Leu Ser Met		
325	330	335
Ile Gly Met Pro Tyr Tyr His Glu Arg Leu Phe Ser Ala Trp Pro Ser		
340	345	350
His Leu Val Phe Glu Val Gly Ser Pro Pro Pro Gln Val Asp Gln Ala		
355	360	365
Leu Ile Pro Tyr Leu Arg Pro Ala Glu Leu Gly His His Ala Pro Asn		
370	375	380
Pro Lys Lys Thr Arg Arg Tyr Gln Val Glu Asn Thr Arg Ala Leu Ala		
385	390	395
400		
Ser Ala Glu Pro Ala Leu Ile Ala Pro Lys Phe Leu Ser Glu Lys Ala		
405	410	415
Arg Glu Gln Asn Lys Ala Lys Ser Glu Gly Ala Ile Ser Asp Ala Ala		
420	425	430
Glu Ala Leu Ala Gly Ala Lys Ile Asn Gly Glu Thr Asp Asp Asp Pro		
435	440	445
Leu Leu Lys Tyr Ser Asn Val Glu Ile Lys Tyr Ser Arg Phe Gly Val		
450	455	460
Asp Asp Phe Asp Phe Arg Phe Tyr Asn Gln Thr Thr Phe Ser Gly Leu		
465	470	475
480		
Glu Thr His Ile Ala Asn Ser Phe Thr Asn Ala Leu Leu Gln Leu Phe		
485	490	495
Lys Phe Ile Pro Tyr Ile Arg Asn Val Ala Leu His His Ala Ala Ser		
500	505	510
Ser Cys Ile Phe Glu Thr Cys Leu Leu Cys Glu Met Gly Tyr Leu Phe		
515	520	525
Asp Met Leu Glu Lys Ala Ser Gly Gln Asn Cys Gln Ala Thr Asn Leu		
530	535	540
Leu Lys Thr Phe Ser Ser Tyr Arg Glu Ala Ser Asn Leu Gly Leu Phe		
545	550	555
560		
Glu Glu Asn Leu Thr Asn Lys Ser Leu Ser Ala Ala Ile Gln Ala Val		
565	570	575
Asn Arg Phe Phe Leu Gly Gln Ile Ser His Asp Phe Arg Met Ile Ser		
580	585	590
Pro Ser Ser Asp Asp Leu Asp His Arg Leu Ala Thr Val Ala Ser Glu		
595	600	605
Ser Ile Arg Cys Met Phe Cys Gln Asn Glu Ile Val Arg Pro Gly Asn		
610	615	620
Ser Leu Val Asn Glu Leu Asn Tyr Pro Ala Ile Asp Ile Lys Gln Ala		
625	630	635
640		
Arg Arg Asn Pro Ala Phe Arg Phe Ser Asn Ile Leu Arg Ala Ser Ile		
645	650	655
Glu Arg Glu Ala Gln Asn Arg Gly Trp Cys Asn Tyr Cys Arg Arg Tyr		
660	665	670
Gln Gln Val Ala Ile Arg Lys Ser Val His Arg Met Pro Gln Val Leu		
675	680	685
Met Leu Asn Ala Ala Leu Thr Asn Pro Ile Cys Arg Arg Leu Trp Ala		
690	695	700
Ile Pro Gly Trp Leu Pro Glu Glu Val Gly Ile Val Ile Glu Gly Gly		

-continued

705	710	715	720
Gln Ile Leu Cys Phe Glu Gly Glu Asp	Leu Lys Leu Arg Val Gln Ala		
725	730	735	
Lys Met Pro Gly Leu Val Val Tyr Asp	Leu Val Gly Leu Val Cys Glu		
740	745	750	
Ile Asp Ile Pro Glu His Gln Lys Ala His	Leu Val Ser Phe Ile Asn		
755	760	765	
Val Ser Ile Ser Ser Arg Glu Pro Glu Thr Lys Asn Lys Trp His Leu			
770	775	780	
Phe Asn Asp Phe Leu Val Thr Glu Val Asp	Lys Glu Glu Ala Leu Arg		
785	790	795	800
Phe Asn Gln Pro Trp Lys Ile Pro Cys Val	Leu Ala Tyr Gln Val Gln		
805	810	815	
Asp Gly Arg His Ala Met Asp Asp Thr Trp Lys Asp Ala Leu Asp Thr			
820	825	830	
Thr Leu Leu Phe Arg Asp Trp Ser Leu Asn Gly Gly Arg Pro Val Glu			
835	840	845	
Ser Arg Val Thr Leu Ser Glu Glu Glu Lys Pro Thr Pro Gly Thr Pro			
850	855	860	
Val Ala Leu Asp Thr Glu Phe Val Asp Leu Glu Lys Ala Glu Ile Asp			
865	870	875	880
Val Lys Ala Asp Gly Ser Gln Glu Ile Val Arg Pro Ser Lys Ser Gly			
885	890	895	
Leu Ala Arg Val Ser Val Leu Arg Gly Ser Gly Ile Arg Glu Gly Val			
900	905	910	
Pro Phe Ile Asp Asp Tyr Ile Thr Ile Lys Glu Asn Ile Val Asp Tyr			
915	920	925	
Val Thr Gln Tyr Ser Gly Ile Lys Pro Gly Asp Leu Asp Pro Arg Val			
930	935	940	
Ser Gln His Asn Leu Val Pro Leu Lys Val Ala Tyr Lys Lys Leu Trp			
945	950	955	960
Leu Leu Leu Asn Leu Gly Cys Val Phe Val Gly His Gly Leu Ala Ser			
965	970	975	
Asp Phe Arg Lys Val Asn Ile Gln Val Pro Lys Ser Gln Thr Val Asp			
980	985	990	
Thr Gln Tyr Leu Phe Phe His Pro Gly Lys Asn Arg Arg Leu Ser Leu			
995	1000	1005	
Arg Tyr Leu Ala Trp Ala Val Phe Lys Glu Tyr Ile Gln Glu Glu			
1010	1015	1020	
Pro Ala Asp Asp Ser Gln Gly His Asp Ser Ile Glu Asp Ala Arg			
1025	1030	1035	
Met Ala Leu Arg Leu Trp Lys Lys Phe Lys Glu Tyr Glu Asp Ala			
1040	1045	1050	
Gly Ile Val Ser Gln Ile Leu Glu Glu Ile Phe Arg Glu Gly Ser			
1055	1060	1065	
Lys Leu Gly Phe Arg Pro Pro Arg Asn Gly Val Pro Thr Val			
1070	1075	1080	
Leu Ser Arg Pro Gly Thr Ala Val Thr Met Gln Asn Asn Ser Gly			
1085	1090	1095	
Arg Asn Thr Pro Ser Thr Ser Asp Val Ala Gly Ala Ala Ala Ser			
1100	1105	1110	

-continued

Ala Pro Ala Thr Pro Arg Gln Ala Phe Arg Arg Ser Ile Ala Leu
1115 1120 1125

Thr Pro Ser Asn Gly Ser Phe Ala Gly Pro Gly Thr Gly Asp Phe
1130 1135 1140

Phe Ser Gly Ser Pro Leu Lys
1145 1150

<210> SEQ ID NO 26

<211> LENGTH: 642

<212> TYPE: PRT

<213> ORGANISM: Aspergillus niger

<220> FEATURE:

<221> NAME/KEY: peptide

<222> LOCATION: (1)..(642)

<223> OTHER INFORMATION: Pan3

<400> SEQUENCE: 26

Met Ala Thr Thr Gly Lys Ser Ala Thr Leu Glu Asp Ala Arg His Gly
1 5 10 15

Thr Gly Ser Pro Lys Met Lys Gly Arg Gly Cys Ala Phe Asn His Asp
20 25 30

Pro His Lys Val Asn Ser Gly Tyr Gln Ser Asp Ser Asn Lys Lys Arg
35 40 45

Leu Asn Val Asp Ser Pro Ser Phe Thr Pro Ser Ile Leu Ser Ser Asn
50 55 60

Gly Ser Ser Pro Thr Ser Gln Ser Ala Thr Met Lys Lys Met Ala Thr
65 70 75 80

Ile Ser Pro Lys Ala Ala Ser Ala Ala Pro Phe Gln Pro Arg Ser Ile
85 90 95

Ser Ser Arg Ser Asn Ser Ser Thr Pro Thr Thr Arg Pro Gly Thr Met
100 105 110

Thr Pro Asp Trp Ser Val Ala Glu Val Gln Glu Phe Val Pro Gln Gly
115 120 125

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

Val Pro Ser Thr Ser Ala Phe Asp Pro Phe Val Thr Ala Pro Asn Pro
145 150 155 160

Leu Ser Ala Ala Asn Ala Val Gly Pro Val Gln Ala Asn Pro Phe Ser
165 170 175

His Asp Thr Ala Ala Ala Leu Asn Gly Ala Ala Phe Phe Ala Asn Gln
180 185 190

Ser Gly Phe Gln Gln Pro Val Gln Tyr His Met Tyr Ala Pro Ile Gly
195 200 205

Pro His Ser Gln Asn Thr Leu Gly Tyr Gln Arg Asn Val His Asp Leu
210 215 220

Phe Leu Pro Asn Asp Leu Arg Glu Glu Met Gln Lys Lys Ala Ala Ala
225 230 235 240

Thr Leu Gln Thr Leu Pro Asn Thr Gln Leu Pro Ala Gln Val Asp Tyr
245 250 255

Phe His Ser Leu Val Pro Leu Asp Leu Asn His Gln Lys Asn Ala Thr
260 265 270

Ile Phe Gly Phe Pro Ser Trp Val Tyr Lys Ala Gln Ser Ser Lys Asp
275 280 285

-continued

Gly Asn Phe Tyr Ala Leu Arg Arg Leu Glu Gly Phe Arg Leu Thr Asn
290 295 300

Glu Lys Ala Ile Arg Ser Val Gln Ala Trp Lys Arg Val Cys Asn Gly
305 310 315 320

Ser Val Val Thr Val His Asp Ala Phe Thr Ser Arg Ser Phe Gln Asp
325 330 335

Ser Ser Leu Ile Phe Val Thr Asp Tyr His Pro Leu Ser Lys Thr Leu
340 345 350

Ala Glu Gln His Leu Gly Ala Gly Asn Arg Phe Gln Gly Arg His Asn
355 360 365

Thr His Ile Pro Glu Gln Val Leu Trp Gly Tyr Met Thr Gln Ile Ala
370 375 380

Asn Ala Leu Lys Ala Ile His Ala Ser Gln Leu Ala Ala Arg Ile Ile
385 390 395 400

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

Ala Cys Ala Ile Met Asp Val Val Gln Phe Asp Thr Gln Arg Ser Leu
420 425 430

Ala Glu Leu Gln Arg Gln Asp Leu Val Asn Phe Gly Gln Leu Ile Val
435 440 445

Thr Leu Gly Ala Asn Gln Pro Asn Val Met His Asn Pro Thr Lys Ala
450 455 460

Met Glu His Phe Thr Arg Ala Tyr Thr Ala Gln Leu Lys Asn Ser Val
465 470 475 480

Phe Trp Leu Leu Asn Gly Leu Gln Lys Asp Gln Glu Arg Asn Ile Asp
485 490 495

Ile Phe Ile Thr Gly Ile Ser Ser Thr Leu Met Ser Thr Phe Asp Ser
500 505 510

Ala Leu His Leu Asp Asp Gln Leu Thr Ser Asp Leu Ser Arg Glu Leu
515 520 525

Glu Asn Gly Arg Leu Val Arg Leu Met Thr Lys Leu Asn Phe Val Asn
530 535 540

Glu Arg Pro Glu Tyr Glu His Asp Arg Gln Trp Ser Glu Asn Gly Glu
545 550 555 560

Arg Tyr Phe Leu Lys Ile Phe Arg Asp Tyr Val Phe His Gln Val Asp
565 570 575

Ala Gln Gly Asp Pro Val Val Asp Leu Gly His Val Leu Met Cys Leu
580 585 590

Asn Lys Leu Asp Ala Gly Thr Asp Glu Lys Ile Thr Leu Ile Ser Arg
595 600 605

Asp Glu Gln Ser Cys Phe Val Val Ser Tyr Lys Glu Leu Lys Lys Ala
610 615 620

Leu Glu Ser Ser Phe Gln Ala Leu Leu Lys Pro Ser Ala Ser Arg Arg
625 630 635 640

Leu His

```

<210> SEQ ID NO 27
<211> LENGTH: 148
<212> TYPE: PRT
<213> ORGANISM: Aspergillus niger
<220> FEATURE:
<221> NAME/KEY: peptide
<222> LOCATION: (1)..(148)

```

-continued

<223> OTHER INFORMATION: Rpb4

<400> SEQUENCE: 27

Met Ser Val Gln Leu Pro Pro Ala Thr His Arg Lys Arg Thr Leu Pro
1 5 10 15

Gln Gly Glu Leu Glu Ala Ala Ser Thr Leu Lys Leu Gly Ala Asp Gln
20 25 30

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

Val Leu Glu Asn Lys Arg Arg Gly Gly Lys Lys Tyr Glu Glu Pro Glu
50 55 60

Asn Leu Thr Lys Thr Leu Asp Tyr Leu Glu Val Phe Ala Arg Phe Lys
65 70 75 80

Asp Glu Glu Asn Ile Lys Ala Val Glu Arg Leu Leu Asn Ser His Thr
85 90 95

Glu Leu Glu Met Phe Glu Arg Ser Gln Leu Gly Ser Leu Cys Cys Asp
100 105 110

Asn Ala Glu Ala Lys Ser Leu Ile Pro Ser Leu Gln His Lys Ile
115 120 125

Ser Asp Gly Asp Leu Gln Glu Leu Leu Asp Glu Leu Thr Lys Leu Arg
130 135 140

Asn Phe Thr Glu
145

<210> SEQ ID NO 28

<211> LENGTH: 167

<212> TYPE: PRT

<213> ORGANISM: Aspergillus niger

<220> FEATURE:

<221> NAME/KEY: peptide

<222> LOCATION: (1) .. (167)

<223> OTHER INFORMATION: Rpb7

<400> SEQUENCE: 28

Met Phe Phe Leu Lys Glu Glu Val Lys Val Ile Thr Leu His Pro Ser
1 5 10 15

Tyr Phe Gly Pro Asn Met Arg Glu Tyr Leu Ile Asn Arg Leu Asn Glu
20 25 30

Glu Glu Glu Gly Arg Cys Thr Gly Asp His Phe Val Ile Cys Val Met
35 40 45

Asp Met Val Asp Ile Gly Glu Gly Arg Val Leu Pro Gly Ser Gly Gln
50 55 60

Ala Glu Tyr Thr Ile Lys Tyr Arg Ala Ile Ile Trp Lys Pro Phe Arg
65 70 75 80

Gly Glu Thr Val Asp Ala Ile Val Thr Ser Val Lys Pro Thr Gly Ile
85 90 95

Phe Thr Leu Ala Gly Pro Leu Ser Val Phe Ile Ala Arg Lys Asn Ile
100 105 110

Pro Ser Asp Ile Lys Trp Glu Pro Asn Thr Val Pro Pro Gln Tyr Thr
115 120 125

Asp His Ala Asp Gln Val Ile Glu Lys Gly Thr Ser Leu Arg Leu Lys
130 135 140

Ile Leu Gly Val Lys Pro Asp Val Ala Ala Ile Asn Ala Ile Gly Thr
145 150 155 160

-continued

Ile Lys Glu Asp Tyr Leu Gly
165

```

<210> SEQ ID NO 29
<211> LENGTH: 4886
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)..(4886)
<223> OTHER INFORMATION: Plasmid pLyss

<400> SEQUENCE: 29

gaattccgga tgagcattca tcaggcggo aagaatgtga ataaaggccg gataaaactt      60
gtgcttattt ttctttacgg tctttaaaaa ggccgtataa tccagctgaa cggtctggtt      120
ataggcatat tgagcaactg actgaaatgc ctcaaaaatgt tctttacgt gccattggga      180
tataatcaacg gtggtatatac cagtgatttt tttctccatt ttagcttcct tagctcctga      240
aaatctcgat aactcaaaaa atacgcccgg tagtgatctt atttcattat ggtgaaagtt      300
ggaaccttctt acgtgccat caacgtctca ttttcgcca aagttggccc agggcttccc      360
ggtatcaaca gggacaccag gatttattta ttctgcgaag tgatcttcgg tcacaggat      420
ttattcggcg caaagtgcgt cgggtgatgc tgccaaactt ctgatttagt gtatgtatgt      480
gtttttgagg tgctccagtg gcttctgtt ctatcagctg tccctctgt tcagctactg      540
acggggtgtt gcgtaacggc aaaagcaccc cgccacatca gcgcctagccg agtgtataact      600
ggcttactat gttggactg atgagggtgt cagtgaagt cttcatgtgg caggagaaaa      660
aaggctgcac cggtgctca gcagaatatg tgatacagga tatattccgc ttctcgctc      720
actgactcgc tacgctcggt cgttcgactg cggcgagccg aaatggctt cgaacggggc      780
ggagatttcc tggaaagatgc caggaagata cttAACAGGG aagtggagg gcccggcaa      840
agccgtttt ccataaggctc cggccccctg acaagcatca cggaaatctga cgctcaaattc      900
agtgggtggcg aaacccgaca ggactataaa gataccaggc gttccctg gcggctccct      960
cgtgcgtct cctgttccctg cctttcgggtt taccgggttc attcggctgt tatggccgcg      1020
tttgtctcat tccacgcctg acactcagtt ccgggttaggc agttcgctcc aagctggact      1080
gtatgcacga acccccccgtt cagtcgcacc gctgcgcctt atccggtaac tattgtcttgc      1140
agtccaaaccc ggaaagacat gcaaaagcac cactggcagc agccactggt aattgattna      1200
gaggagtttag tcttgaagtc atgcggcggt taaggctaaa ctgaaaggac aagttttgg      1260
gactgcgtc ctccaagccca gttacctcggt ttcaaaagagt tggttagctca gagaacccctc      1320
aaaaaaacccgc cctgcaaggc ggtttttccg ttttcagagc aagagattac ggcgcagacca      1380
aaacgatctc aagaagatca tcttattaaat cagataaaat atttcttagat ttcaatgtca      1440
tttatctctt caaatgttagc acctgaagtc agccccatac gatataagtt gtaattctca      1500
tggggacag cttatcatcg ataagcttta atgcggtagt ttatcacagt taaattgcta      1560
acgcagtcag gcaccgtgtt tgaaatctaa caatgcgtc atcgtcatcc tcggcaccgt      1620
caccctggat gctgtaggca taggottgggt tatgcggta ctgcccccc tcttgcggga      1680
tacgcgtccat tccgacagca tcgcccgtc ctatggcgtg ctgctagccg tataatgcgtt      1740
gatgcaattt ctatgcgcac ccgttctcggt agcactgtcc gaccgcttgc gccggccccc      1800

```

-continued

agtccctgctc	gttgcgtac	ttggagccac	tatcgactac	gcatcgatgg	cgaccacacc	1860
cgtccctgtgg	atccggccca	ttggctgcct	cccacacttg	gatatgcctc	ctcgaggcct	1920
tatagaattt	tttataagac	ttgctgcatta	tttgacctcc	aatgcgaaca	aaggaaacc	1980
gtgtgtgtct	cccttttagt	agttcaatta	attatccacg	gtcagaagt	accagttcgt	2040
tcttcctcca	ccaacgctta	aggtcgaacg	aagggcaacg	cttcggcgcc	acctcatgat	2100
gggcgcgaag	accagegcct	tcgtacttag	ccagcagtgt	gacaaggagt	gagcgaaggg	2160
attgcatttgc	ggctggcgta	aagttacgt	cgaacttacc	tttategtcg	ataccaccaa	2220
caaggcagac	gccgatagag	tttggttgt	aacccttagc	gtgagagcct	acagccatct	2280
catctcgatcc	tgcctccaca	gtaccgtctc	gcttgatgat	aaagtggtat	cccacatcga	2340
gccaaccctg	ctctttgtgc	cactggcgaa	tctcacggac	accaacatc	tgacttggct	2400
tggtagccg	gcagtgaaca	aagattgcgt	cagtagattc	acgttggta	aactgtacac	2460
gagccattat	ttctttccctc	ctttcccttt	taatctatca	aaggggaccc	ggatcctcta	2520
cgcggacgc	atcggtggccg	gcatcaccgg	cgcacagg	gcgggttgctg	gcgcctata	2580
cgcgcacatc	accgatgggg	aagatcgccc	tcgcccacttc	gggctcatga	gctgttgc	2640
cggcgtgggt	atgggtggcag	gccccgtggc	cggggactg	ttggggcgcc	tctccttgca	2700
tgcaccattc	cttgcggcg	cggtgctcaa	cggcctcaac	ctactactgg	gctgtttcct	2760
aatgcaggag	tcgcataagg	gagagcgtcg	accgatgccc	ttgagagcct	tcaacccagt	2820
cagtccttc	cgggtggcgc	ggggcatgac	tatcgtcgc	gcacttatga	ctgtcttctt	2880
tatcatgcaa	ctcgtaggac	agggtggccg	agcgctctgg	gtcattttcg	gcgaggaccc	2940
cttcgctgg	agcgcgacga	tgatggcct	gtcgcttgc	gtattcgaa	tcttcacgc	3000
cctcgctcaa	gccttcgtca	ctggtcccgc	caccaaactg	tccggcgaga	agcaggccat	3060
tatcgccggc	atggcgccg	acgcgttgg	ctacgttctg	ctggcggtcg	cgacgcgagg	3120
ctggatggcc	ttccccattt	tgattttct	cgcttccggc	ggcatcgaaa	tgcccgctt	3180
gcaggccatg	ctgtccaggc	aggtagatga	cgaccatca	ggacagcttc	aaggatcgct	3240
cgcggctt	accagctaa	cttcgatcac	tggaccgtcg	atcgtcacgg	cgatttatgc	3300
cgcctcgccg	agcacatgga	acgggttggc	atggattgt	ggcgccgccc	tataccttgt	3360
ctgcctcccc	gcgttgcgtc	gccccgtgc	gagccggcc	acctcgaccc	gaatggaa	3420
cggccgcacc	tcgctaaccgg	attcaccact	ccaagaattt	gagccaatca	attcttgcgg	3480
agaactgtga	atgcgaaac	caacccttgg	cagaacat	ccatcgctc	cgccatctcc	3540
agcagccgca	cgcggcgcat	ctcgggcago	gttgggtct	ggccacgggt	gcgcgtatc	3600
gtgtcttgt	cggttgggac	ccggcttaggc	tgggggggtt	gccttactgg	tttagcagaat	3660
gaatcaccga	tacgcgagcg	aacgtgaago	gactgtcgct	gaaaaacgtc	tgcgaccctga	3720
gcaacaacat	aatgggtctt	cggttccgt	gtttcgtaaa	gtctggaaac	gcggaaagtcc	3780
cctacgtgt	gttgaagtt	cccgcaacag	agagtggaa	caaccgggt	taccacgata	3840
ctatgactga	gagtcaacgc	catgagcgcc	ctcattttctt	attctgagtt	acaacagtcc	3900
gcaccgcgtgt	ccggtagctc	cttcgggtgg	gcgcggggca	tgactatcg	cgccgcactt	3960
atgactgtct	tctttatcat	gcaactcgta	ggacaggtgc	cgccagcgcc	caacagtccc	4020
cggccacgg	ggcctgccac	cataccacg	ccgaaacaag	cgccctgcac	cattatgttc	4080

-continued

cggatctgca tcgcaggatg ctgctggcta ccctgtggaa cacctacatc tgtattaacg	4140
aagcgctaac cggttttatac aggcctctggg aggcaaaaata aatgatcata tcgtcaatta	4200
ttacctccac ggggagagcc tgagcaaact ggccctcaggc atttgagaag cacacggtca	4260
cactgctcc ggtgtcaat aaaccggtaa accagcaata gacataagcg gctatttaac	4320
gaccctgccc tgaaccgacg accgggtcga atttgcttc gaatttctgc cattcatcg	4380
cttattatca cttattcagg cgtagcacca ggcgttaag ggcaccaata actgcctaa	4440
aaaaattacg ccccgccctg ccactcatacg cagtaactgtt gtaattcatt aagcattctg	4500
ccgacatgga agccatcaca gacggcatga tgaacctgaa tcgccagccg catcagcacc	4560
ttgtcgccctt gcgtataata tttgcccattg gtggaaacgg gggcgaagaa gttgtccata	4620
ttggccacgt ttaaatcaaa actgggtgaaa ctcacccagg gattggctga gacgaaaaac	4680
atattctcaa taaacccttt agggaaatag gccagggttt caccgtaaaca cgccacatct	4740
tgcgaatata tggtagaaa ctgcggaaa tcgtcggtt attcactcca gagcgatgaa	4800
aacgtttcag ttgcctcatg gaaaacgggtg taacaagggtt gaacactatc ccataatcacc	4860
agctcaccgt ctttcatatgc catacg	4886

<210> SEQ ID NO 30
<211> LENGTH: 660
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)..(660)
<223> OTHER INFORMATION: Gene encoding chloramphenicol
acetyltransferase, which confers resistance to chloramphenicol

<400> SEQUENCE: 30

atg gag aaa atc act gga tat acc acc gtt gat ata tcc caa tgg	48
Met Glu Lys Lys Ile Thr Gly Tyr Thr Val Asp Ile Ser Gln Trp	
1 5 10 15	
cat cgt aaa gaa cat ttt gag gca ttt cag tca gtt gct caa tgt acc	96
His Arg Lys Glu His Phe Glu Ala Phe Gln Ser Val Ala Gln Cys Thr	
20 25 30	
tat aac cag acc gtt cag ctg gat att acg gcc ttt tta aag acc gta	144
Tyr Asn Gln Thr Val Gln Leu Asp Ile Thr Ala Phe Leu Lys Thr Val	
35 40 45	
aag aaa aat aag cac aag ttt tat ccg gcc ttt att cac att ctt gcc	192
Lys Lys Asn Lys His Lys Phe Tyr Pro Ala Phe Ile His Ile Leu Ala	
50 55 60	
cgc ctg atg aat gct cat ccg gaa ttc cgt atg gca atg aaa gac ggt	240
Arg Leu Met Asn Ala His Pro Glu Phe Arg Met Ala Met Lys Asp Gly	
65 70 75 80	
gag ctg gtg ata tgg gat agt gtt cac cct tgt tac acc gtt ttc cat	288
Glu Leu Val Ile Trp Asp Ser Val His Pro Cys Tyr Thr Val Phe His	
85 90 95	
gag caa act gaa acg ttt tca tcg ctc tgg agt gaa tac cac gac gat	336
Glu Gln Thr Glu Thr Phe Ser Ser Leu Trp Ser Glu Tyr His Asp Asp	
100 105 110	
ttc cgg cag ttt cta cac ata tat tcg caa gat gtc gcg tgt tac ggt	384
Phe Arg Gln Phe Leu His Ile Tyr Ser Gln Asp Val Ala Cys Tyr Gly	
115 120 125	

-continued

gaa aac ctg gcc tat ttc cct aaa ggg ttt att gag aat atg ttt ttc Glu Asn Leu Ala Tyr Phe Pro Lys Gly Phe Ile Glu Asn Met Phe Phe 130 135 140	432
gtc tca gcc aat ccc tgg gtg agt ttc acc agt ttt gat tta aac gtg Val Ser Ala Asn Pro Trp Val Ser Phe Thr Ser Phe Asp Leu Asn Val 145 150 155 160	480
gcc aat atg gac aac ttc ttc gcc ccc gtt ttc acc atg ggc aaa tat Ala Asn Met Asp Asn Phe Phe Ala Pro Val Phe Thr Met Gly Lys Tyr 165 170 175	528
tat acg caa ggc gac aag gtg ctg atg ccg ctg gcg att cag gtt cat Tyr Thr Gln Gly Asp Lys Val Leu Met Pro Leu Ala Ile Gln Val His 180 185 190	576
cat gcc gtc tgt gat ggc ttc cat gtc ggc aga atg ctt aat gaa tta His Ala Val Cys Asp Gly Phe His Val Gly Arg Met Leu Asn Glu Leu 195 200 205	624
caa cag tac tgc gat gag tgg cag ggc ggg gcg taa Gln Gln Tyr Cys Asp Glu Trp Gln Gly Gly Ala 210 215	660

<210> SEQ ID NO: 31

<211> LENGTH: 219

<212> TYPE: PRT

<213> ORGANISM: artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 31

Met Glu Lys Lys Ile Thr Gly Tyr Thr Thr Val Asp Ile Ser Gln Trp 1 5 10 15
--

His Arg Lys Glu His Phe Glu Ala Phe Gln Ser Val Ala Gln Cys Thr 20 25 30

Tyr Asn Gln Thr Val Gln Leu Asp Ile Thr Ala Phe Leu Lys Thr Val 35 40 45

Lys Lys Asn Lys His Lys Phe Tyr Pro Ala Phe Ile His Ile Leu Ala 50 55 60

Arg Leu Met Asn Ala His Pro Glu Phe Arg Met Ala Met Lys Asp Gly 65 70 75 80
--

Glu Leu Val Ile Trp Asp Ser Val His Pro Cys Tyr Thr Val Phe His 85 90 95

Glu Gln Thr Glu Thr Phe Ser Ser Leu Trp Ser Glu Tyr His Asp Asp 100 105 110
--

Phe Arg Gln Phe Leu His Ile Tyr Ser Gln Asp Val Ala Cys Tyr Gly 115 120 125
--

Glu Asn Leu Ala Tyr Phe Pro Lys Gly Phe Ile Glu Asn Met Phe Phe 130 135 140
--

Val Ser Ala Asn Pro Trp Val Ser Phe Thr Ser Phe Asp Leu Asn Val 145 150 155 160
--

Ala Asn Met Asp Asn Phe Phe Ala Pro Val Phe Thr Met Gly Lys Tyr 165 170 175
--

Tyr Thr Gln Gly Asp Lys Val Leu Met Pro Leu Ala Ile Gln Val His 180 185 190
--

His Ala Val Cys Asp Gly Phe His Val Gly Arg Met Leu Asn Glu Leu 195 200 205
--

Gln Gln Tyr Cys Asp Glu Trp Gln Gly Gly Ala 210 215
--

-continued

```

<210> SEQ ID NO 32
<211> LENGTH: 103
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: promoter
<222> LOCATION: (1)..(103)
<223> OTHER INFORMATION: cat promoter

<400> SEQUENCE: 32

```

tgatcgac gtaagagggtt ccaacttca ccataatgaa ataagatcac taccggcgt	60
attttttagt ttatcgagat tttcaggago taaggaagct aaa	103

```

<210> SEQ ID NO 33
<211> LENGTH: 545
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: rep_origin
<222> LOCATION: (1)..(545)
<223> OTHER INFORMATION: P15A ori

<400> SEQUENCE: 33

```

ttgagatcgt ttttgtctgc gcgtaatctc ttgtctgaa aacgaaaaaaaa ccgccttgca	60
gggcgggttt tcgaagggttc tctgagctac caactctttg aaccgaggta actggctgg	120
aggagcgcag tcacaaaaac ttgtccttgc agtttagcct taaccggcgc atgacttcaa	180
gactaactcc tctaaatcaa ttaccagtgg ctgctgccag tggtgcttt gcatgtctt	240
ccgggttggc ctcaagacga tagttaccgg ataaggcgcgac ggggtggac tgaacggggg	300
gttcgtgtcat acagtccagc ttggagcgcgaa ctgccttacc ggaactgagt gtcaggcggt	360
gaatgagaca aacgcggcca taacagcggta atgacacccgg taaaccggaa ggcaggaaca	420
ggagagcgcgca cgagggagcc gccaggggaa acgcctggta tctttatagt cctgtcggt	480
ttcgccacca ctgatttgag cgtcagattt cgtatgtttt gtcagggggg cggagccat	540
ggaaa	545

```

<210> SEQ ID NO 34
<211> LENGTH: 34
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: promoter
<222> LOCATION: (1)..(34)
<223> OTHER INFORMATION: Phi 3.8 promoter

<400> SEQUENCE: 34

```

taattaattt aactcaactaa agggagacca cago	34
--	----

```

<210> SEQ ID NO 35
<211> LENGTH: 456
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:

```

-continued

<221> NAME/KEY: CDS
 <222> LOCATION: (1)..(456)
 <223> OTHER INFORMATION: T7 lysozyme

<400> SEQUENCE: 35

atg gct cgt gta cag ttt aaa caa cgt gaa tct act gac gca atc ttt	48
Met Ala Arg Val Gln Phe Lys Gln Arg Glu Ser Thr Asp Ala Ile Phe	
1 5 10 15	
gtt cac tgc tcg gct acc aag cca agt cag aat gtt ggt gtc cgt gag	96
Val His Cys Ser Ala Thr Lys Pro Ser Gln Asn Val Gly Val Arg Glu	
20 25 30	
att cgc cag tgg cac aaa gag cag ggt tgg ctc gat gtg gga tac cac	144
Ile Arg Gln Trp His Lys Glu Gln Gly Trp Leu Asp Val Gly Tyr His	
35 40 45	
ttt atc atc aag cga gac ggt act gtg gag gca gga cga gat gag atg	192
Phe Ile Ile Lys Arg Asp Gly Thr Val Glu Ala Gly Arg Asp Glu Met	
50 55 60	
gct gta ggc tct cac gct aag ggt tac aac cac aac tct atc ggc gtc	240
Ala Val Gly Ser His Ala Lys Gly Tyr Asn His Asn Ser Ile Gly Val	
65 70 75 80	
tgc ctt gtt ggt atc gac gat aaa ggt aag ttc gac gct aac ttt	288
Cys Leu Val Gly Ile Asp Asp Lys Gly Lys Phe Asp Ala Asn Phe	
85 90 95	
acg cca gcc caa atg caa tcc ctt cgc tca ctg ctt gtc aca ctg ctg	336
Thr Pro Ala Gln Met Gln Ser Leu Arg Ser Leu Leu Val Thr Leu Leu	
100 105 110	
gct aag tac gaa ggc gct gtg ctt cgc gcc cat cat gag gtg gcg ccg	384
Ala Lys Tyr Glu Gly Ala Val Leu Arg Ala His His Glu Val Ala Pro	
115 120 125	
aag gct tgc cct tcg ttc gac ctt aag cgt tgg tgg gag aag aac gaa	432
Lys Ala Cys Pro Ser Phe Asp Leu Lys Arg Trp Glu Lys Asn Glu	
130 135 140	
ctg gtc act tct gac cgt gga taa	456
Leu Val Thr Ser Asp Arg Gly	
145 150	

<210> SEQ ID NO 36

<211> LENGTH: 151
 <212> TYPE: PRT
 <213> ORGANISM: artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 36

Met Ala Arg Val Gln Phe Lys Gln Arg Glu Ser Thr Asp Ala Ile Phe	
1 5 10 15	
Val His Cys Ser Ala Thr Lys Pro Ser Gln Asn Val Gly Val Arg Glu	
20 25 30	
Ile Arg Gln Trp His Lys Glu Gln Gly Trp Leu Asp Val Gly Tyr His	
35 40 45	
Phe Ile Ile Lys Arg Asp Gly Thr Val Glu Ala Gly Arg Asp Glu Met	
50 55 60	
Ala Val Gly Ser His Ala Lys Gly Tyr Asn His Asn Ser Ile Gly Val	
65 70 75 80	
Cys Leu Val Gly Gly Ile Asp Asp Lys Gly Lys Phe Asp Ala Asn Phe	
85 90 95	
Thr Pro Ala Gln Met Gln Ser Leu Arg Ser Leu Leu Val Thr Leu Leu	
100 105 110	

-continued

Ala Lys Tyr Glu Gly Ala Val Leu Arg Ala His His Glu Val Ala Pro
115 120 125

Lys Ala Cys Pro Ser Phe Asp Leu Lys Arg Trp Trp Glu Lys Asn Glu
130 135 140

Leu Val Thr Ser Asp Arg Gly
145 150

<210> SEQ ID NO 37
<211> LENGTH: 29
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: promoter
<222> LOCATION: (1)..(29)
<223> OTHER INFORMATION: tet promoter

<400> SEQUENCE: 37

ttgacagctt atcatcgata agcttaat 29

<210> SEQ ID NO 38
<211> LENGTH: 8447
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)..(8447)
<223> OTHER INFORMATION: Plasmid pMax

<400> SEQUENCE: 38

cgtatggcaa tgaaagacgg tgagctggtg atatggata gtgttcaccc ttgttacacc 60
gttttccatg agcaaactga aacgtttca tcgctctgga gtgaataccca cgacgattc 120
ccgcagtttc tacacatata ttcgcaagat gtggcgtgt acggtaaaa cctggcttat 180
ttccctaaag ggtttattga gaatatgttt ttctgtctcag ccaateccctg ggtgagttc 240
accagtttg atttaaacgt ggccaatatg gacaacttct tcggccccgt tttcaccatg 300
ggcaaatatt atacgcaagg cgacaagggtg ctgatgccgc tggcgattca gtttcatcat 360
ggcgctgttgc atggcttcca tgcggcaga atgcttaatg aattacaaca gtactgcgtat 420
gagtggcagg gcggggcgta attttttaa ggcagttatt ggtgcctta aacgccttgt 480
gctacgcctg aataagtgtat aataagcggta tgaatggcag aaattcgaaa gcaaattcga 540
cccggtcgtc ggttcagggc agggtcgtta aatagccgt tatgtctatt gctggttac 600
cggtttatttgc actaccggaa gcagttgtgac cgtgtgttcc tcaaattgcgtt gaggccagtt 660
tgctcaggct ctccccgtgg aggtataat tgacgatatg atcattttt ctgcctccca 720
gagcctgata aaaacggtta ggcgttcgtt aatacagatg taggtgttcc acagggttagc 780
cagcagcatc ctgcgtatgca gatcggaaac ataatggtgc agggcgottt tttcggcgtt 840
ggtatggtgg caggccccgt ggccggggga ctgttggcg ctgcccgcac ctgtcctacg 900
agttgcgtatg taaagaagac agtcataagt gcccgcac tagtcatgcc ccgcgcac 960
cggaggagc taccggacag cggtgccgac tggtaact cagaataaga aatgaggccg 1020
ctcatggcgt tgacttcag tcatagtatc gtggtatcac cgggtggttc cactctgtt 1080

-continued

tgcgggcaac ttcagcagca cgttagggac ttccgcgtt ccagactta cgaaacacgg	1140
aaaccgaaga ccattcatgt tgttgctcg gtcgcagacg ttttcagca gcagtgcctt	1200
cacgttcgtc cgcgtatcg tgattcattc tgctaaccag taaggcaacc cccgcagcct	1260
agccgggtcc tcaacgcacag gagcacgatc atgcgcaccc gtggccagga cccaaacgctg	1320
cccgagatgc gcccgcgtcgc gctgctggag atggcggacg cgatggatat gttctgcca	1380
gggttggtt ggcgcattcac agtttcgc aagaattgtat tggctccaat tcttggatgt	1440
gtgaatccgt tagcgagggtg cgcgcggctt ccattcaggt cgaggtggcc cggctccatg	1500
caccgcgcacg caacgcgggg aggccagacaa ggtatagggc ggccgcctaca atccatgc	1560
acccgttcca tggctcgcc gaggoggcat aaatcgccgt gacgatcagc ggtccagtga	1620
tgcgaagttag gctggtaaga gcccgcagcg atcccttgaag ctgtccctga tggctgtcat	1680
ctacctgcctt ggacagcatg gcctgcaacg cgggcatecc gatgcgcgcg gaaggagaa	1740
gaatcataat ggggaaggcc atccagcctc gctgcgcgaa cgcgcacaa acgtacccca	1800
gcccgcgtggc cgccatgcgcg ggcataatgg cctgcttctc ggcgaaacgt ttggcggcgg	1860
gaccagtgcg gaaggcttga gcgaggcgt gcaagattcc gaataccgcg acgcacaggc	1920
cgatcatcg tgcgcctccag cggaaagcggt cctgcgcgaa aatgacccag agcgcgtgc	1980
gcacctgtcc tacgagttgc atgataaaga agacagtcat aagtgcggcg acgatagtca	2040
tgcggccgcg ccaccggaaag gagctgactg ggttgaaggc tctcaaggc atcggtcgac	2100
gtctccctt atgcgacttcc tgcattagga acaagccag tagtaggtt aggccgttga	2160
gcaccgcgcg cgcaaggaaat ggtgcgtca aggagatggc gcccaacagt cccccggca	2220
cggggcctgc caccataccc acgcgcacaa aagcgctcat gagccgcgaa tggcgagccc	2280
gatctccccc atcggtgatc tcggcgatata gggccgcacg aaccgcaccc gtggcgcgg	2340
tgcggccgcg cacatcggtt cccgggttgc ggtatccggg ccccttgcgat agataaaaa	2400
ggaaaggagg aaagaataaa tggctcggtt acagttaaa caacgtgaat ctactgacgc	2460
aatctttgtt cactgtcg cttccaaagcc aatgcacaaat gttgggttcc gtgcgttc	2520
ccagtgccac aaagagcagg gttggctcg tgcggatatac cactttatca tcaagcgaga	2580
cggtactgtg gaggcaggac gagatgagat ggctgttaggc tctcaacgcg aagggttacaa	2640
ccacaactct atcggtgtct gccttgcgtt tggtatcgac gataaaggta agttcgacgc	2700
taactttacg ccagccaaa tgcataatccct tcgctactg ctgtcacac tgctggctaa	2760
gtacgaaggc gctgtgttcc tgcgcgcatac tgagggtggcg ccgaaggctt gcccctcggtt	2820
cgacccataag cttgggtggg agaagaacga actggtcact tctgcacgc gataaaccgc	2880
tgcggccgcg cttgggtggg cccgtcgacg cggcgcgcgc tgaggatacc tcagccacca	2940
aaattcgcgca aattgtgaac atcatcactg tcatcttcc tcgggttgcata atggccatt	3000
ttccctgtcag taacgagaag gtcgcgaatt caggcgctt ttagacttgt cgtaatgaaa	3120
ttctttttaa gaaggagact atatatgaaa gcaatttcg tactgaaaca tcttaatcat	3180
gctaaggagg ttttctaattt aaaagaatgt taatcaacgc aactcagcag gaagatgtgc	3240
gcccgcgttgc tgcggccctt tgcgttgttgc atgacctgga tatcgaaagt ccaggccacg	3300
agcagaaaaaa ggccaaacatc tacaaggta aaatcaccgc cattgaaccc agtctggaaag	3360

-continued

ctgcttttgt tgattacggc gctgaacgto acggtttcccccactaaaaaa gaaattgccc	3420
gcgaatattt ccctgctaactacagtgctc atggtcgtcc caacattaaa gatgtgtgc	3480
gtgaagggtca ggaagtcatgtttagatcg ataaagaaga ggcggcaac aaaggcgcgg	3540
cattaaccac ctttatcgt ctggcgggta gctatctgg tctgatgccg aacaaccgc	3600
gcccgggtgg catttcgcg cgtatcgaag ggcacgaccg taccgaattaa aagaagcac	3660
tggcaaggct tgaaactgccg gaaggcatgg ggcttategt ggcacccgt ggcgtcgca	3720
aatctgtga ggcgctgaa tgggattaa gcttcgtct gaaacactgg gaagccatca	3780
aaaaagccgc tgaaagccgc ccggccccgt tcctgattca tcaggagagc aacgtaatcg	3840
ttcgcgcatt ccgcgattac ttacgtcagg acatcggcga aatccttatac gataaccga	3900
aagtgcgtca actggcacgt cagcatatcg ctgcattagg tgcggccgtt ttcagcagca	3960
aaatcaaact gtacacccgcg gagatcccgc tgttcagccca ctaccagatc gagtacacaga	4020
tgcagtcgcg cttccagcgt gaagttcgto tgcgtctgg tggttccatt gttatcgaca	4080
gcaccgaagc gttaacggcc atcgacatca actccgcacg cgcgacccgc ggccggcgata	4140
tgcgaaagaaac cgcggttaac actaacctcg aagctgcgta tgagattgtc cgtagctgc	4200
gcgtgcgtga cctcgccgcg ctgattgtta tcaacttcat cgacatgacg ccagtagcc	4260
accagcgtgc ggtagaaaaac cgtagcgtg aagcggtgcg tcaggaccgt gcgcgtattc	4320
aaatcagccca tatttcgcg tttggcctgc tgaaatgtc ccgtcagcgc ctgagcccat	4380
cactgggtga atccagccat cacgtctgcg cgcgtctgc cggtagccgt accgtcgtg	4440
acaacgaatc gctgtcgctc tctattctgc gtctgatcga agaagaagcgc ctgaaagaga	4500
acacccagga agttcacgcc attgttcctg tgccaatcgc ttcttatctg ctgaatgaaa	4560
aacgttctgc ggtgaatgcc attgaaacgc gtcaggacgg tgttcgtgc gtgattgtc	4620
caaacgatca gatggaaacc ccgcactacc acgtgtcgcc cgtgcgtaaa gggaaagaaa	4680
cgccgacccctt aagctacatcg ctgcgcgaa tgcgtgcgaa agcgatggcg ctgcgtctg	4740
aagaagagtt cgctgaacgt aagegtccgg aacaacctgc gctggcaacc tttgcctgc	4800
cggatgtgccg cccagegcca accccagctg aacctgcgcg cgcgtcgta gccccagcac	4860
ctaaatctgc accggcaaca ccagccgctc ctgccccacc tgggtgttg agccgcttct	4920
tccggcgact gaaagcgctg ttccgggtg gtgaagaaac caaaccgtcc gagcaaccaa	4980
caccgaaagc agaagcgaaa ccggAACGTC aacaggatcg tcgcaagcct cgtcagaaca	5040
accggccgtga ccgtaatgag cgccgcgaca cccgtatgtg acgtactgaa ggcagcgata	5100
atcgcaaga aaccgtcgat aatcgctcc aggacacgca gcagactgcc gagacgcgtg	5160
agagccgtca cgagggttag gtaacggaaa aagcgctac caccgacgag cagcaagcgc	5220
cgcgtcgctga acgttagccgc cgccgtatag atgataaaccg tcaggcgcaaa caagaagcga	5280
aggcgctgaa tggtgaagag caatctgttc agggaaaccga acaggaagaa cgtgtacgtc	5340
cggttcagcc cgctcgtaaa cagcgctcago tcaatcgaaat gtcgtctac gagcaagcgc	5400
tagccgaaga agcggtagtc gcacccgggtgg ttggaaac tgcgtcgcc gaccaattt	5460
ttcaggaaagc gccagctcca cgcacagaac tggtaaaat cccgctgcgc gtcgtacgc	5520
aaaactgcacc agaacagcaa gaagagaaca atgctgataa ccgtgacaaac ggtggcatgc	5580
cgcgtcgcttc tcgcgcgtcg cctcgctacc tgcgcgtaaat tggtcagcgt cgtcgctgc	5640

-continued

atcgtgacga	gcgttatcca	accaggatcg	caatgccgtt	gaccgttagcg	tgcgcgtctc	5700
cggaaactggc	ctctggcaaa	gtctggatcc	gctatccaat	tgtacgtccg	caagatgtac	5760
aggttgaaga	gcagcgcgaa	caggaagaag	tacatgtgca	gccgatggtg	actgagggtcc	5820
ctgtcgccgc	cgctatcgaa	ccggttgtta	gcgcgcagg	tgttgaagaa	gtggccggtg	5880
tcgtagaagc	ccccgttcag	gttgcgcgaa	cgcaaccgg	agtgggtgaa	acgacgcata	5940
ctgaagtgtat	tgctgcccgc	gttaactgaa	agccgcagg	gattaccgag	tctgtatgtt	6000
ccgtageccca	ggaagttgca	gaacaaggag	aaccgggtgg	tgaaccgcag	gaagagacgg	6060
cagatattga	agaagttgtc	gaaactgctg	aggttgttagt	tgctgaacct	gaagttgtt	6120
ctcaaccctgc	cgcgccagta	gtcgcgtgaa	tcgcagcaga	agttgaaacg	gtagctgcgg	6180
tcgaacctga	ggtcaccgtt	gagcataacc	acgctaccgc	gcacatgacg	cgcgcgtccag	6240
caccggata	tgttccggag	gcacccgcgc	acagtgaactg	gcagcgcocct	actttgcct	6300
tcgaaggtaa	aggtgcgcga	ggtggtcata	cgccaacaca	tcatgcctct	gctgctcctg	6360
cgcgtccgca	acctgtttag	taaagggtac	ctagaaatca	tccttagcga	aagctaagga	6420
tttttttat	ctgttaattaa	ttgaactcac	taaaggggaga	ccacaggggt	ttccctttgt	6480
tcgcattgg	ggtcaaataa	tgcgcaagtc	ttataaaca	ttctataagg	ctccgaggag	6540
gcataatccaa	gtgtgggagg	cagccaatgg	gcggatcca	caggacgggt	gtggtcgcca	6600
tgatcgcgta	gtcgatagtg	gctccaaatg	gcgaaggcg	caggactgg	cgccggccaa	6660
agcggtegga	cagtgcgtcc	agaacgggtg	cgcataaaaa	ttgcataac	gcataatgcg	6720
ctagcagcac	gccatagtga	ctggcgatgc	tgtcggatg	gacgataatcc	cgcaagaggc	6780
ccggcagtagc	cgccataacc	aaggctatgc	ctacagcata	cagggtgacg	gtggcgagga	6840
tgacgatgag	cgcatgttta	gatttcatac	acgggcctg	actgcgttag	caatttaact	6900
gtgataaaact	accgcattaa	agcttatcga	tgataagtcg	tcaaacatga	gaattacaac	6960
ttatatacgta	tggggctgac	ttcaggtgt	acatttgaag	agataaaatg	cactgaaatc	7020
tagaaatatt	ttatctgatt	aataagatga	tcttctttag	atcgttttgg	tctgcgcgt	7080
atctcttgc	ctgaaaacga	aaaaaccgc	ttgcaggggc	gtttttcgaa	ggttctctga	7140
gctaccaact	ctttgaaccg	aggtaactgg	cttggaggag	cgcagtcacc	aaaacttgc	7200
ctttcagttt	agccttaacc	ggcgcgtac	ttcaagacta	actcctctaa	atcaattacc	7260
agtggctgct	gccagtggt	ctttgcgt	tcttccggg	ttggactcaa	gacgatagtt	7320
accggataag	gcccggcggt	cggactgaa	ggggggttcg	tgcatacagt	ccagcttgg	7380
gcgaactgcc	tacccggaa	tgagtgtcg	gcgtggatg	agacaaacgc	ggccataaca	7440
gcggaaatgac	accggtaaac	cgaaaggcg	gaacaggaga	gcgcacgagg	gagccgcag	7500
gggaaacgcc	tggtatctt	atagtcctgt	cgggtttcgc	caccactgat	ttgagcgtca	7560
gatttcgtga	tgcttgcag	ggggcggag	cctatggaaa	aacggcttg	ccgcggccct	7620
ctcaactcc	tgttaagtat	cttcgtggca	tcttccagg	aatctcogcc	ccgttgcata	7680
gccatttccg	ctcgccgcag	tcgaacgacc	gagcgttagcg	agtcaatgt	cgaggaagcg	7740
gaatatatcc	tgtatcacat	attctgtga	cgcacccgt	cagccctttt	tctcctgcca	7800
catgaagcac	ttcactgaca	ccctcatcg	tgccaaacata	gtaagccagt	atacactccg	7860
ctagcgcgt	tgccggcg	tgctttgtcc	gttacgcacc	accccgctcg	tagctgaaca	7920

-continued

ggagggacag ctgatagaaa cagaagccac tggagcacct caaaaacacc atcatacact	7980
aaatcagtaa gttggcagca tcacccgacg cacttgcgc cgaataaata cctgtacgg	8040
aagatcaacct cgccagaataa ataaatccctgt gtgtccctgt tgataccggg aagccctggg	8100
ccaactttg gcgaaaatga gacgttgatc ggcacgtaag aggttccaac tttcaccata	8160
atgaaataag atcaactaccg ggcgtatccc tttagtatttc gagatttca ggagctaagg	8220
aagctaaaaat ggagaaaaaa atcaactggat ataccaccgt tgatataatcc caatggcatc	8280
gtaaagaaca ttttggggca tttcagtcag ttgctcaatg tacctataac cagaccgttc	8340
agctggatat tacggccttt taaaagaccg taaaagaaaaaa taagcacaag ttttatccgg	8400
cctttattca cattctgcc cgccgtatga atgctcatcc ggaattc	8447

<210> SEQ ID NO 39
<211> LENGTH: 3186
<212> TYPE: DNA
<213> ORGANISM: E. coli
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)..(3186)
<223> OTHER INFORMATION: rne(D346N) encoding mutant RNase E of E.coli comprising a single point mutation: D346N

<400> SEQUENCE: 39

atg aaa aga atg tta atc aac gca act cag cag gaa gag ttg cgc gtt	48
Met Lys Arg Met Leu Ile Asn Ala Thr Gln Gln Glu Glu Leu Arg Val	
1 5 10 15	
gcc ctt gta gat ggg cag cgt ctg tat gac ctg gat atc gaa agt cca	96
Ala Leu Val Asp Gly Gln Arg Leu Tyr Asp Leu Asp Ile Glu Ser Pro	
20 25 30	
ggg cac gag cag aaa aag gca aac atc tac aaa ggt aaa atc acc cgc	144
Gly His Glu Gln Lys Lys Ala Asn Ile Tyr Lys Gly Lys Ile Thr Arg	
35 40 45	
att gaa ccg agt ctg gaa gct ttt gtt gat tac ggc gct gaa cgt	192
Ile Glu Pro Ser Leu Glu Ala Ala Phe Val Asp Tyr Gly Ala Glu Arg	
50 55 60	
cac ggt ttc ctc cca cta aaa gaa att gcc cgc gaa tat ttc cct gct	240
His Gly Phe Leu Pro Leu Lys Glu Ile Ala Arg Glu Tyr Phe Pro Ala	
65 70 75 80	
aac tac agt gct cat ggt cgt ccc aac att aaa gat gtg ttg cgt gaa	288
Asn Tyr Ser Ala His Gly Arg Pro Asn Ile Lys Asp Val Leu Arg Glu	
85 90 95	
ggg cag gaa gtc att gtt cag atc gat aaa gaa gag cgc ggc aac aaa	336
Gly Gln Glu Val Ile Val Gln Ile Asp Lys Glu Glu Arg Gly Asn Lys	
100 105 110	
ggc gcg gca tta acc acc ttt atc agt ctg gcg ggt agc tat ctg gtt	384
Gly Ala Ala Leu Thr Thr Phe Ile Ser Leu Ala Gly Ser Tyr Leu Val	
115 120 125	
ctg atg ccg aac aac ccg cgc ggc ggt ggc att tct cgc cgt atc gaa	432
Leu Met Pro Asn Asn Pro Arg Ala Gly Gly Ile Ser Arg Arg Ile Glu	
130 135 140	
ggc gac gac cgt acc gaa tta aaa gaa gca ctg gca agc ctt gaa ctg	480
Gly Asp Asp Arg Thr Glu Leu Lys Glu Ala Leu Ala Ser Leu Glu Leu	
145 150 155 160	
ccg gaa ggc atg ggg ctt atc gtg cgc acc gct ggc gtc ggc aaa tct	528
Pro Glu Gly Met Gly Leu Ile Val Arg Thr Ala Gly Val Gly Lys Ser	
165 170 175	
gct gag ggc ctg caa tgg gat tta agc ttc cgt ctg aaa cac tgg gaa	576

-continued

Ala Glu Ala Leu Gln Trp Asp Leu Ser Phe Arg Leu Lys His Trp Glu			
180	185	190	
gcc atc aaa aaa gcc gct gaa agc cgc ccg gcc ccg ttc ctg att cat		624	
Ala Ile Lys Lys Ala Ala Glu Ser Arg Pro Ala Pro Phe Leu Ile His			
195	200	205	
cag gag agc aac gta atc gtt cgc gca ttc cgc gat tac tta cgt cag		672	
Gln Glu Ser Asn Val Ile Val Arg Ala Phe Arg Asp Tyr Leu Arg Gln			
210	215	220	
gac atc ggc gaa atc ctt atc gat aac ccg aaa gtg ctc gaa ctg gca		720	
Asp Ile Gly Glu Ile Leu Ile Asp Asn Pro Lys Val Leu Glu Leu Ala			
225	230	235	240
cgt cag cat atc gct gca tta ggt cgc ccg gat ttc agc agc aaa atc		768	
Arg Gln His Ile Ala Ala Leu Gly Arg Pro Asp Phe Ser Ser Lys Ile			
245	250	255	
aaa ctg tac acc ggc gag atc ccg ctg ttc agc cac tac cag atc gag		816	
Lys Leu Tyr Thr Gly Glu Ile Pro Leu Phe Ser His Tyr Gln Ile Glu			
260	265	270	
tca cag atc gag tcc gcc ttc cag cgt gaa gtt cgt ctg ccg tct ggt		864	
Ser Gln Ile Glu Ser Ala Phe Gln Arg Glu Val Arg Leu Pro Ser Gly			
275	280	285	
ggg tcc att gtt atc gac agc acc gaa gcg tta acg gcc atc gac atc		912	
Gly Ser Ile Val Ile Asp Ser Thr Glu Ala Leu Thr Ala Ile Asp Ile			
290	295	300	
aac tcc gca cgc gcg acc cgc ggc ggc gat atc gaa gaa acc gcg ttt		960	
Asn Ser Ala Arg Ala Thr Arg Gly Gly Asp Ile Glu Glu Thr Ala Phe			
305	310	315	320
aac act aac ctc gaa gct gcc gat gag att gct cgt cag ctg cgc ctg		1008	
Asn Thr Asn Leu Glu Ala Ala Asp Glu Ile Ala Arg Gln Leu Arg Leu			
325	330	335	
cgt gac ctc ggc ctg att gtt atc aac ttc atc gac atg acg cca		1056	
Arg Asp Leu Gly Leu Ile Val Ile Asn Phe Ile Asp Met Thr Pro			
340	345	350	
gta cgc cac cag cgt gcg gta gaa aac cgt ctg cgt gaa gcg gtt cgt		1104	
Val Arg His Gln Arg Ala Val Glu Asn Arg Leu Arg Glu Ala Val Arg			
355	360	365	
cag gag cgt gcg cgt att caa atc agc cat att tct cgc ttt ggc ctg		1152	
Gln Asp Arg Ala Arg Ile Gln Ile Ser His Ile Ser Arg Phe Gly Leu			
370	375	380	
ctg gaa atg tcc cgt cag cgc ctg agc cca tca ctg ggt gaa tcc agc		1200	
Leu Glu Met Ser Arg Gln Arg Leu Ser Pro Ser Leu Gly Glu Ser Ser			
385	390	395	400
cat cac gtc tgc ccg cgc tgc tcc ggt acc ggt acc gtg cgt gac aac		1248	
His His Val Cys Pro Arg Cys Ser Gly Thr Gly Thr Val Arg Asp Asn			
405	410	415	
gaa tcg ctg tcg ctc tct att ctg cgt ctg atc gaa gaa gaa gcg ctg		1296	
Glu Ser Leu Ser Leu Ser Ile Leu Arg Leu Ile Glu Glu Ala Leu			
420	425	430	
aaa gag aac acc cag gaa gtt cac gcc att gtt cct gtg cca atc gct		1344	
Lys Glu Asn Thr Gln Glu Val His Ala Ile Val Pro Val Pro Ile Ala			
435	440	445	
tct tat ctg ctg aat gaa aaa cgt tct gcg gtg aat gcc att gaa acg		1392	
Ser Tyr Leu Leu Asn Glu Lys Arg Ser Ala Val Asn Ala Ile Glu Thr			
450	455	460	
cgt cag gac ggt gtt cgc tgc gtg att gtg cca aac gat cag atg gaa		1440	
Arg Gln Asp Gly Val Arg Cys Val Ile Val Pro Asn Asp Gln Met Glu			
465	470	475	480
acc ccg cac tac cac gtg ctg cgc gtg cgt aaa ggg gaa gaa acg ccg		1488	

-continued

Thr Pro His Tyr His Val Leu Arg Val Arg Lys Gly Glu Glu Thr Pro		
485	490	495
acc tta agc tac atg ctg ccg aag ctg cat gaa gaa gcg atg gcg ctg		1536
Thr Leu Ser Tyr Met Leu Pro Lys Leu His Glu Glu Ala Met Ala Leu		
500	505	510
ccg tct gaa gaa gag ttc gct gaa cgt aag cgt ccg gaa caa cct gcg		1584
Pro Ser Glu Glu Glu Phe Ala Glu Arg Lys Arg Pro Glu Gln Pro Ala		
515	520	525
ctg gca acc ttt gcc atg ccg gat gtg ccg cca gcg cca acc cca gct		1632
Leu Ala Thr Phe Ala Met Pro Asp Val Pro Pro Ala Pro Thr Pro Ala		
530	535	540
gaa cct gcc gcg cct gtc gta gcc cca gca cct aaa tct gca ccg gca		1680
Glu Pro Ala Ala Pro Val Val Ala Pro Ala Pro Lys Ser Ala Pro Ala		
545	550	555
560		
aca cca gcc gct cct gcc caa cct ggg ctg ttg agc cgc ttc ttc ggc		1728
Thr Pro Ala Ala Pro Ala Gln Pro Gly Leu Leu Ser Arg Phe Phe Gly		
565	570	575
gca ctg aaa gcg ctg ttc agc ggt ggt gaa gaa acc aaa ccg tcc gag		1776
Ala Leu Lys Ala Leu Phe Ser Gly Gly Glu Glu Thr Lys Pro Ser Glu		
580	585	590
caa cca aca ccg aaa gca gaa gcg aaa ccg gaa cgt caa cag gat cgt		1824
Gln Pro Thr Pro Lys Ala Glu Ala Lys Pro Glu Arg Gln Gln Asp Arg		
595	600	605
cgc aag cct cgt cag aac aac cgc cgt gac cgt aat gag cgc cgc gac		1872
Arg Lys Pro Arg Gln Asn Asn Arg Arg Asp Arg Asn Glu Arg Arg Asp		
610	615	620
acc cgt agt gaa cgt act gaa ggc agc gat aat cgc gaa gaa aac cgt		1920
Thr Arg Ser Glu Arg Thr Glu Gly Ser Asp Asn Arg Glu Asn Arg		
625	630	635
640		
cgt aat cgt cgc cag gca cag cag act gcc gag acg cgt gag agc		1968
Arg Asn Arg Arg Gln Ala Gln Gln Thr Ala Glu Thr Arg Glu Ser		
645	650	655
cgt cag cag gtt gag gta acg gaa aaa gcg cgt acc acc gac gag cag		2016
Arg Gln Gln Val Glu Val Thr Glu Lys Ala Arg Thr Thr Asp Glu Gln		
660	665	670
caa gcg ccg cgt gaa cgt acg cgc cgc cgt aat gat gat aaa cgt		2064
Gln Ala Pro Arg Arg Glu Arg Ser Arg Arg Arg Asn Asp Asp Lys Arg		
675	680	685
cag gcg caa caa gaa gcg aag gcg ctg aat gtt gaa gag caa tct gtt		2112
Gln Ala Gln Gln Glu Ala Lys Ala Leu Asn Val Glu Glu Gln Ser Val		
690	695	700
cag gaa acc gaa cag gaa gaa cgt gta cgt ccg gtt cag ccg cgt cgt		2160
Gln Glu Thr Glu Gln Glu Glu Arg Val Arg Pro Val Gln Pro Arg Arg		
705	710	715
720		
aaa cag cgt cag ctc aat cag aaa gtg cgt tac gag caa agc gta gcc		2208
Lys Gln Arg Gln Leu Asn Gln Lys Val Arg Tyr Glu Gln Ser Val Ala		
725	730	735
gaa gaa gcg gta gtc gca ccg gtg gtt gaa gaa act gtc gct gcc gaa		2256
Glu Glu Ala Val Val Ala Pro Val Val Glu Glu Thr Val Ala Ala Glu		
740	745	750
cca att gtt cag gaa gcg cca gct cca cgc aca gaa ctg gtg aaa gtc		2304
Pro Ile Val Gln Glu Ala Pro Ala Pro Arg Thr Glu Leu Val Lys Val		
755	760	765
ccg ctg cca gtc gta gcg caa act gca cca gaa cag caa gaa gag aac		2352
Pro Leu Pro Val Val Ala Gln Thr Ala Pro Glu Gln Gln Glu Glu Asn		
770	775	780
aat gct gat aac cgt gac aac ggt ggc atg ccg cgt cgt tct cgc cgc		2400

-continued

Asn Ala Asp Asn Arg Asp Asn Gly Gly Met Pro Arg Arg Ser Arg Arg 785 790 795 800	
tcg cct cgt cac ctg cgc gta agt ggt cag cgt cgt cgt cgc tat cgt Ser Pro Arg His Leu Arg Val Ser Gly Gln Arg Arg Arg Arg Tyr Arg 805 810 815	2448
gac gag cgt tat cca acc cag tcg cca atg ccg ttg acc gta gcg tgc Asp Glu Arg Tyr Pro Thr Gln Ser Pro Met Pro Leu Thr Val Ala Cys 820 825 830	2496
gcg tct ccg gaa ctg gcc tct ggc aaa gtc tgg atc cgc tat cca att Ala Ser Pro Glu Leu Ala Ser Gly Lys Val Trp Ile Arg Tyr Pro Ile 835 840 845	2544
gta cgt ccg caa gat gta cag gtt gaa gag cag cgc gaa cag gaa gaa Val Arg Pro Gln Asp Val Gln Val Glu Glu Gln Arg Glu Gln Glu Glu 850 855 860	2592
gta cat gtg cag ccg atg gtg act gag gtc cct gtc gcc gcc gct atc Val His Val Gln Pro Met Val Thr Glu Val Pro Val Ala Ala Ala Ile 865 870 875 880	2640
gaa ccg gtt gtt agc gcg cca gtt gtt gaa gaa gtc gcc ggt gtc gta Glu Pro Val Val Ser Ala Pro Val Val Glu Glu Val Ala Gly Val Val 885 890 895	2688
gaa gcc ccc gtt cag gtt gcc gaa ccg caa ccg gaa gtg gtt gaa acg Glu Ala Pro Val Gln Val Ala Glu Pro Gln Pro Glu Val Val Glu Thr 900 905 910	2736
acg cat cct gaa gtg att gct gcc gcg gta act gaa cag ccg cag gtg Thr His Pro Glu Val Ile Ala Ala Val Thr Glu Gln Pro Gln Val 915 920 925	2784
att acc gag tct gat gtt gcc gta gcc cag gaa gtt gca gaa caa gca Ile Thr Glu Ser Asp Val Ala Val Ala Gln Glu Val Ala Glu Gln Ala 930 935 940	2832
gaa ccg gtg gtt gaa ccg cag gaa gag acg gca gat att gaa gaa gtt Glu Pro Val Val Glu Pro Gln Glu Glu Thr Ala Asp Ile Glu Glu Val 945 950 955 960	2880
gtc gaa act gct gag gtt gta gtt gct gaa cct gaa gtt gtt gct caa Val Glu Thr Ala Glu Val Val Ala Glu Pro Glu Val Val Ala Gln 965 970 975	2928
cct gcc gcg cca gta gtc gct gaa gtc gca gca gaa gtt gaa acg gta Pro Ala Ala Pro Val Val Ala Glu Val Ala Ala Glu Val Glu Thr Val 980 985 990	2976
gct gcg gtc gaa cct gag gtc acc gtt gag cat aac cac gct acc gcg Ala Ala Val Glu Pro Glu Val Thr Val Glu His Asn His Ala Thr Ala 995 1000 1005	3024
cca atg acg cgc gct cca gca ccg gaa tat gtt ccg gag gca ccg Pro Met Thr Arg Ala Pro Ala Pro Glu Tyr Val Pro Glu Ala Pro 1010 1015 1020	3069
cgt cac agt gac tgg cag cgc cct act ttt gcc ttc gaa ggt aaa Arg His Ser Asp Trp Gln Arg Pro Thr Phe Ala Phe Glu Gly Lys 1025 1030 1035	3114
ggt gcc gca ggt ggt cat acg gca aca cat cat gcc tct gct gct Gly Ala Ala Gly Gly His Thr Ala Thr His His Ala Ser Ala Ala 1040 1045 1050	3159
cct gcg cgt ccg caa cct gtt gag taa Pro Ala Arg Pro Gln Pro Val Glu 1055 1060	3186

<210> SEQ ID NO 40

<211> LENGTH: 1061

<212> TYPE: PRT

<213> ORGANISM: E. coli

-continued

<400> SEQUENCE: 40

```

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

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

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

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

His Gly Phe Leu Pro Leu Lys Glu Ile Ala Arg Glu Tyr Phe Pro Ala
65          70          75          80

Asn Tyr Ser Ala His Gly Arg Pro Asn Ile Lys Asp Val Leu Arg Glu
85          90          95

Gly Gln Glu Val Ile Val Gln Ile Asp Lys Glu Glu Arg Gly Asn Lys
100         105         110

Gly Ala Ala Leu Thr Thr Phe Ile Ser Leu Ala Gly Ser Tyr Leu Val
115         120         125

Leu Met Pro Asn Asn Pro Arg Ala Gly Gly Ile Ser Arg Arg Ile Glu
130         135         140

Gly Asp Asp Arg Thr Glu Leu Lys Glu Ala Leu Ala Ser Leu Glu Leu
145         150         155         160

Pro Glu Gly Met Gly Leu Ile Val Arg Thr Ala Gly Val Gly Lys Ser
165         170         175

Ala Glu Ala Leu Gln Trp Asp Leu Ser Phe Arg Leu Lys His Trp Glu
180         185         190

Ala Ile Lys Lys Ala Ala Glu Ser Arg Pro Ala Pro Phe Leu Ile His
195         200         205

Gln Glu Ser Asn Val Ile Val Arg Ala Phe Arg Asp Tyr Leu Arg Gln
210         215         220

Asp Ile Gly Glu Ile Leu Ile Asp Asn Pro Lys Val Leu Glu Leu Ala
225         230         235         240

Arg Gln His Ile Ala Ala Leu Gly Arg Pro Asp Phe Ser Ser Lys Ile
245         250         255

Lys Leu Tyr Thr Gly Glu Ile Pro Leu Phe Ser His Tyr Gln Ile Glu
260         265         270

Ser Gln Ile Glu Ser Ala Phe Gln Arg Glu Val Arg Leu Pro Ser Gly
275         280         285

Gly Ser Ile Val Ile Asp Ser Thr Glu Ala Leu Thr Ala Ile Asp Ile
290         295         300

Asn Ser Ala Arg Ala Thr Arg Gly Gly Asp Ile Glu Glu Thr Ala Phe
305         310         315         320

Asn Thr Asn Leu Glu Ala Ala Asp Glu Ile Ala Arg Gln Leu Arg Leu
325         330         335

Arg Asp Leu Gly Gly Leu Ile Val Ile Asn Phe Ile Asp Met Thr Pro
340         345         350

Val Arg His Gln Arg Ala Val Glu Asn Arg Leu Arg Glu Ala Val Arg
355         360         365

Gln Asp Arg Ala Arg Ile Gln Ile Ser His Ile Ser Arg Phe Gly Leu
370         375         380

Leu Glu Met Ser Arg Gln Arg Leu Ser Pro Ser Leu Gly Glu Ser Ser

```

-continued

385	390	395	400
His His Val Cys Pro Arg Cys Ser Gly Thr Gly Thr Val Arg Asp Asn			
405	410	415	
Glu Ser Leu Ser Leu Ser Ile Leu Arg Leu Ile Glu Glu Ala Leu			
420	425	430	
Lys Glu Asn Thr Gln Glu Val His Ala Ile Val Pro Val Pro Ile Ala			
435	440	445	
Ser Tyr Leu Leu Asn Glu Lys Arg Ser Ala Val Asn Ala Ile Glu Thr			
450	455	460	
Arg Gln Asp Gly Val Arg Cys Val Ile Val Pro Asn Asp Gln Met Glu			
465	470	475	480
Thr Pro His Tyr His Val Leu Arg Val Arg Lys Gly Glu Glu Thr Pro			
485	490	495	
Thr Leu Ser Tyr Met Leu Pro Lys Leu His Glu Glu Ala Met Ala Leu			
500	505	510	
Pro Ser Glu Glu Glu Phe Ala Glu Arg Lys Arg Pro Glu Gln Pro Ala			
515	520	525	
Leu Ala Thr Phe Ala Met Pro Asp Val Pro Pro Ala Pro Thr Pro Ala			
530	535	540	
Glu Pro Ala Ala Pro Val Val Ala Pro Ala Pro Lys Ser Ala Pro Ala			
545	550	555	560
Thr Pro Ala Ala Pro Ala Gln Pro Gly Leu Leu Ser Arg Phe Phe Gly			
565	570	575	
Ala Leu Lys Ala Leu Phe Ser Gly Gly Glu Glu Thr Lys Pro Ser Glu			
580	585	590	
Gln Pro Thr Pro Lys Ala Glu Ala Lys Pro Glu Arg Gln Gln Asp Arg			
595	600	605	
Arg Lys Pro Arg Gln Asn Asn Arg Arg Asp Arg Asn Glu Arg Arg Asp			
610	615	620	
Thr Arg Ser Glu Arg Thr Glu Gly Ser Asp Asn Arg Glu Asn Arg			
625	630	635	640
Arg Asn Arg Arg Gln Ala Gln Gln Thr Ala Glu Thr Arg Glu Ser			
645	650	655	
Arg Gln Gln Val Glu Val Thr Glu Lys Ala Arg Thr Thr Asp Glu Gln			
660	665	670	
Gln Ala Pro Arg Arg Glu Arg Ser Arg Arg Arg Asn Asp Asp Lys Arg			
675	680	685	
Gln Ala Gln Gln Glu Ala Lys Ala Leu Asn Val Glu Glu Gln Ser Val			
690	695	700	
Gln Glu Thr Glu Gln Glu Glu Arg Val Arg Pro Val Gln Pro Arg Arg			
705	710	715	720
Lys Gln Arg Gln Leu Asn Gln Lys Val Arg Tyr Glu Gln Ser Val Ala			
725	730	735	
Glu Glu Ala Val Val Ala Pro Val Val Glu Glu Thr Val Ala Ala Glu			
740	745	750	
Pro Ile Val Gln Glu Ala Pro Ala Pro Arg Thr Glu Leu Val Lys Val			
755	760	765	
Pro Leu Pro Val Val Ala Gln Thr Ala Pro Glu Gln Gln Glu Glu Asn			
770	775	780	
Asn Ala Asp Asn Arg Asp Asn Gly Gly Met Pro Arg Arg Ser Arg Arg			
785	790	795	800

-continued

```

Ser Pro Arg His Leu Arg Val Ser Gly Gln Arg Arg Arg Arg Tyr Arg
805           810           815

Asp Glu Arg Tyr Pro Thr Gln Ser Pro Met Pro Leu Thr Val Ala Cys
820           825           830

Ala Ser Pro Glu Leu Ala Ser Gly Lys Val Trp Ile Arg Tyr Pro Ile
835           840           845

Val Arg Pro Gln Asp Val Gln Val Glu Glu Gln Arg Glu Gln Glu Glu
850           855           860

Val His Val Gln Pro Met Val Thr Glu Val Pro Val Ala Ala Ala Ile
865           870           875           880

Glu Pro Val Val Ser Ala Pro Val Val Glu Glu Val Ala Gly Val Val
885           890           895

Glu Ala Pro Val Gln Val Ala Glu Pro Gln Pro Glu Val Val Glu Thr
900           905           910

Thr His Pro Glu Val Ile Ala Ala Ala Val Thr Glu Gln Pro Gln Val
915           920           925

Ile Thr Glu Ser Asp Val Ala Val Ala Gln Glu Val Ala Glu Gln Ala
930           935           940

Glu Pro Val Val Glu Pro Gln Glu Glu Thr Ala Asp Ile Glu Glu Val
945           950           955           960

Val Glu Thr Ala Glu Val Val Ala Glu Pro Glu Val Val Ala Gln
965           970           975

Pro Ala Ala Pro Val Val Ala Glu Val Ala Ala Glu Val Glu Thr Val
980           985           990

Ala Ala Val Glu Pro Glu Val Thr Val Glu His Asn His Ala Thr Ala
995           1000          1005

Pro Met Thr Arg Ala Pro Ala Pro Glu Tyr Val Pro Glu Ala Pro
1010          1015          1020

Arg His Ser Asp Trp Gln Arg Pro Thr Phe Ala Phe Glu Gly Lys
1025          1030          1035

Gly Ala Ala Gly Gly His Thr Ala Thr His His Ala Ser Ala Ala
1040          1045          1050

Pro Ala Arg Pro Gln Pro Val Glu
1055          1060

```

```

<210> SEQ ID NO 41
<211> LENGTH: 119
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: promoter
<222> LOCATION: (1)..(119)
<223> OTHER INFORMATION: rhaBAD promoter

<400> SEQUENCE: 41

caccacaatt cagcaaattt gtaacatcat cacgttcatc tttccctggg tgccaatggc      60
ccattttcct gtcagtaacg agaaggctcg agaattcaggc gcttttaga ctggtcgt          119

```

```

<210> SEQ ID NO 42
<211> LENGTH: 54
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:

```

-continued

```

<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)..(53)
<223> OTHER INFORMATION: BCD

<400> SEQUENCE: 42
atgaaagcaa tttcgtact gaaacatctt aatcatgcta aggagggttt ctaa 54

<210> SEQ ID NO 43
<211> LENGTH: 41
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: terminator
<222> LOCATION: (1)..(41)
<223> OTHER INFORMATION: rrnc terminator

<400> SEQUENCE: 43
agaaatcatc cttagcgaaa gctaaggatt ttttttatct g 41

<210> SEQ ID NO 44
<211> LENGTH: 48
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: terminator
<222> LOCATION: (1)..(48)
<223> OTHER INFORMATION: T7 terminator

<400> SEQUENCE: 44
ctagcataac cccttggggc ctctaaacgg gtcttgaggg gttttttg 48

<210> SEQ ID NO 45
<211> LENGTH: 7855
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)..(7855)
<223> OTHER INFORMATION: Plasmid pMax-noLyss

<400> SEQUENCE: 45
cgtatggcaa tgaaagacgg tgagctggtg atatggata gtgttcaccc ttgttacacc 60
gttttccatg agcaaaactga aacgtttca tcgctctgga gtgaataccca cgacgattc 120
ccgcagttc tacacatata ttgcgaagat gtggcgtgtt acggtgaaaa cctggctat 180
ttccctaaag gtttattga gaatatgttt ttcgtctcag ccaatccctg ggtgagttc 240
accagtttg atttaaacgt ggccaatatg gacaacttct tcgccccgt tttcaccatg 300
ggcaaatatt atacgcaagg cgacaagggt ctgatgccgc tggcgattca ggttcatcat 360
cccgcttgt atggcttcca tgtcggcaga atgcttaatg aattacaaca gtactgcgt 420
gagtggcagg gcggggcgta attttttaa ggcagttatt ggtgcctta aacgcctggt 480
gctacgcctg aataagtgtat aataagcggta tgaatggcag aaattcgaaa gcaaattcga 540
cccggtcgtc ggttcaggggc agggtcgtta aatagccgct tatgtctatt gctggttac 600

```

-continued

cggtttattt actaccggaa gcagtgtgac cgtgtgcttc tcaaattgcct gaggccagtt 660
tgctcaggct ctccccgtgg aggtataat tgacgatag atcattttt ctgcctccca 720
gagcctgata aaaacggta gcgcttcgtt aatacagatg taggtgttcc acagggttagc 780
cagcagcatc ctgcgtatgca gatccggaaac ataatgggtc acggcgttt tttccggctg 840
ggtatggtgg caggccccgt ggccggggga ctgttggcg ctgcccggcac ctgtcctacg 900
agttgcatga taaaagaagac agtctatagt gcgccgacga tagtcatgcc ccgcgcac 960
cggaaggagc taccggacag cggtgcggac tggtaact cagaataaga aatgaggccg 1020
ctcatggcgt tgactctcag tcatacttgc tggatcatc cgggtggttc cactctctgt 1080
tgccggcaac ttccagcagca cgttagggac ttccgcgtt ccagacttta cgaaacacgg 1140
aaaccgaaga ccattcatgt tggtgctcg gtcgcagacg ttttgcagca gcagtcgctt 1200
cacgttgcgt cgcgtatcgg tgattcatc tgcataaccag taaggcaacc ccgcagccct 1260
agccgggtcc tcaacgacag gaggcagatc atgcgcaccc gtggccagga cccaaacgctg 1320
cccgagatgc gccgcgtgac gctgctggag atggcggacg cgatggatat gttctgcca 1380
gggttggttt ggcattcac agtttcgc aagaattgtat tggcttcaat tcttggatgt 1440
gtgaatecgt tagcggaggtg ccgcgggtt ccattcaggt cgagggtggcc cggctccatg 1500
caccgcgacg caacgggggg aggccagacaa ggtataggcc ggcgcctaca atccatgcca 1560
acccgttcca tgcgtcgcc gagggggcat aaatcgccgt gacgatcagc ggtccagtga 1620
tcgaagttag gctggtaaga gcccggagcg atccattgaag ctgtccctga tggctgtcat 1680
ctacctgcgt ggacagcatg gcgtgcacg cgggcacccc gatgcgcgcg gaaggagaa 1740
gaatcataat ggggaaggcc atccagccctc ggcgcgcgaa cgccagcaag acgtacccca 1800
gcccgtggc cgccatgcgc ggcataatgg cctgcgttcc ggcggaaacgt ttggccggg 1860
gaccagtgcg gaagggttgc gcgaggccgt gcaagattcc gataaccgcg agcgcacggc 1920
cgatcatcg tgcgtcccg cgaaagccgt cctgcgcgaa aatgaccccg agcgcgtgcgc 1980
gcacccgtgc tacggatgc atgataaaga agacagtcat aagtgcggcg acgatagtca 2040
tgccccggc ccacccggaaag gagctgactg ggttgaaggc tctcaaggcc atcggtcgac 2100
gtctccctt atgcgactcc tgcattagga agcagcccg tagtaggttggccgggttga 2160
gcaccgcgcg cgcaaggaat ggtgcgtca aggagatggc gcccaacagt ccccccggcca 2220
cgggggcctgc caccataccc acggccaaac aagcgctcat gagcccgaaag tggcgagccc 2280
gatcttcccc atcggtgtat tcggcgatata gggccgcacg aaccgcaccc tggccggccgg 2340
tgatgcggc cacgtgcgt ccggcgtaga ggatccgggt ttaattaaagc ggccgcggcg 2400
cgcccgctga cgccgcgcg gctgaggata cctcaggccac cacaatttcg caaattgtga 2460
acatcatcac gttcatctt ccctgggtgc caatggccca ttttctgtc agtaacgaga 2520
aggtgcgcgaa ttcaaggccgt ttttagactg gtcgtatgaa aattctttt aagaaggaga 2580
ctatataatgc aagcaatttt cgtactgaaa catcttaatc atgctaagga ggtttctaa 2640
tgaaaagaat gttaatcaac gcaactcagc aggaagagtt gcgcgttgcc cttgttagatg 2700
ggcagcgtct gtatgacccgt gatatcgaaa gtccaggccgca cgagcagaaaa aaggccaaaca 2760
tctacaaagg taaaatcacc cgcatgtgaaac cgagtctggaa agctgtttt gttgattacg 2820
gcgcgtgcgaaac tcacgggttc ctccccactaa aagaaattgc cgcgcgaaat ttcctgtca 2880

-continued

actacagtgc tcatggcgt cccaacatta aagatgtgtt ggcgtgaaggta caggaagtca	2940
ttgttcagat cgataaagaa gagcgccgca acaaaggcgcc ggcattaaacc acctttatca	3000
gtctgggggg tagcttatctg gttctgatgc cgaacaacccc gcgcgcgggtt ggcatttctc	3060
gccgtatcga aggcgacgac cgttaccgaat taaaagaagc actggcaagc cttgaactgc	3120
cggaggcat ggggcttatac gtgcgcaccc ctggcgtcgcc caaatctgtt gaggcgctgc	3180
aatgggattt aagcttcgtt ctgaaacact gggaaaggcat caaaaaagcc gctgaaagcc	3240
gccccggcccc gttcctgatt catcaggaga gcaacgtaat cgttcgcgc ttcccgatt	3300
acttacgtca ggacatcgcc gaaatcctta tcgataaccc gaaagtgcgc gaactggcac	3360
gtcagcataat cgctgcattt ggtcgcccg atttcagcag caaaaatcaaa ctgtacacccg	3420
gcgagatccc gctgttcagc cactaccaga tcgagtcaca gatcgagtcc gccttccagc	3480
gtgaagttcg tctggcgctt ggtggttcca ttgttatcgca cagcacccaa gcggttaacgg	3540
ccatcgacat caactccgca cgccgcgaccc gcggccggcga tatcgaagaa accgcgttta	3600
acactaacct cgaagctgcc gatgagattt ctgcgtcagct ggcgcctgcgtt gacccgcgc	3660
gcgtgattgt tatcaacttc atcgacatga cgccagtagc ccaccagcgtt gcggttagaaa	3720
acgcgtctcg tgaagcggtt cgtcaggacc gtgcgcgtat tcaaatcagc catatttctc	3780
gccttggct gctggaaatg tcccgctcgc gcctgagccccc atcaactgggtt gaatccagcc	3840
atcacgtctc cccgcgtgc tccggtaccg gtaccgtgc tgacaacccaa tcgctgtcgc	3900
tctctattct gcgtctgate gaagaagaag cgctgaaaga gaacacccag gaagttcagc	3960
ccattgttcc tgcgtccaaatc gcttcttatac tgctgaatga aaaacgttgc gcggtgaatg	4020
ccattgaaac gcgtcaggacc ggtgttcgtt gcgtgattgtt gccaacgcgtt cagatggaaa	4080
ccccgcacta ccacgtgtc cgcgtcgta aaggggaa aacgcgcacc ttaagctaca	4140
tgcgtccaa gctgcgtgaa gaagcgatgg cgctgcccgtc tgaagaagag ttgcgtgaac	4200
gtaaagcgcc ggaacaacctt cgctggcaat ccttgccat gcccgttgc cggccagcgc	4260
caacccccagc tgaacctgcc gcgcctgtcg tagccccagc acctaaatctt gcacccggca	4320
caccagccgc tcctgeccaa cctgggtgt tgagccgtt ctccggcgc tcgaaagcgc	4380
tgttcagccgg tggtaagaa accaaaccgtt ccgagcaacc aacaccggaaa gcagaagcga	4440
aaccggaaacg tcaacaggat cgctgcgtc ctcgtcagaa caaccgcgtt gaccgtatg	4500
agcgccgcga cacccgtagt gaacgtactg aaggcagcga taatcgccaa gaaaaccgtt	4560
gtaatcgctcg ccaggcacag cagcagactg ccgagacccgt tgagagccgtt cagcagggtt	4620
aggtaacggaa aaaacgcgtt accaccgcgtt agcagcaacgc gcccgttgcgtt gaacgttagcc	4680
gcccgcgtt aaatgtataaa cgctcaggccgaa aacaagaagc gaaggcgctt aatgttgc	4740
agcaaatctgtt tcaggaaacc gaaacggaaatc aacgtgttgcgtt ccgcgtcgta	4800
aacagcgtca gctcaatcgtt aaagtgcgtt acgagcaacgc cgtagccgaa gaagcggtt	4860
tcgcacccgtt ggttgcgtt actgtcgctg ccgaaccaat tgcgtccggaa gcccgttgc	4920
cacgcacaga actgggtgaaa gtccggctgc cagtcgttgc gcaaaactgc ccagaacacgc	4980
aagaagagaa caatgtgtt aaccgtgttcc acgggtggcat gcccgttgcgtt tctcgccgtt	5040
ccgcgttgc tcccgccgtt agtggtcgtt gtcgtcgccgtt ctatcgatc gaggcggtt	5100
caacccagtc gccaatggccg ttgaccgtt cgtgcgcgtt tccggaaactg gcccgttgc	5160

-continued

aagtctggat ccgctatcca attgtacgtc cgcaagatgt acagggtgaa gagcagcg	5220
aacaggaaga agtacatgtg cagccgatgg tgactgaggt ccctgtcgcc gcccgtatcg	5280
aaccggtgt tagcgccca gttgttgaag aagtggccgg tgcgttagaa gccccgttc	5340
aggttgcga accgcaaccg gaagtgggtt aaacgacgca tcctgaagtg attgtgccg	5400
cggtaactga acagccgcag gtgattaccg agtctgtatgt tgccgtagcc caggaagttg	5460
cagaacaagc agaaccgggtg gttgaaccgc aggaagagac ggcagatatt gaagaagttg	5520
tcgaaactgc tgagggtgta gttgtgaac ctgaagttgt tgctcaacct gccgcgcag	5580
tagtcgtga agtgcgacga gaagttgaaa cggtagctgc ggtcgaaacct gaggtcacccg	5640
ttgagcataa ccacgttacc gcgcataatga cggcgctcc agcaccggaa tatgtccgg	5700
aggcacccgcg tcacagtgac tggcagcgccttacttttgc cttcgaaggt aaaggtgccg	5760
cagggtgtca tacggcaaca catcatgcct ctgtgtcc tgcgcgtccg caacctgttg	5820
agtaaagggtt acctagaaat catcottago gaaagctaag gatTTTTT atctgggtt	5880
ccctttgttc gcattggagg tcaaataatgcgcaagtcttataaacaatt ctataaggct	5940
ccgaggaggc atatccaagt gtggaggca gccaatgggc cggatccaca ggacgggtgt	6000
ggtcgcctatg atcgcgtatg ctagatgtgc tccaaatgtc gaagcggcgcg ggactggcg	6060
ggggccaaag cggtcggaca gtgtccgag aacgggtgcg catagaaatt gcatcaacgc	6120
atatagegtc acgacgtacgc catagtgact ggatgtctgc tccgtatggc cgatatccgc	6180
caagaggccc ggcagtaccg gcataaccaa gcctatgcct acagcatcca gggtgacgg	6240
ggcgaggatg acgtgagcg cattgttaga ttccatcacatgcgtcc tgcgttagca	6300
attnaactgt gataaactac cgcattaaatg cttatcgatg ataagctgtc aaacatgaga	6360
attacaactt atatcgatg gggctgactt caggtgtatc atttgaagag ataaattgca	6420
ctgaaatcta gaaatattttt atctgattaa taagatgatc ttcttgatgcgatcgttgc	6480
tgcgcgtatc ctcttgcctt gaaaacgaaa aaaccgcctt gcagggcggt ttttgcgaa	6540
ttctctgagc taccacttgc ttgaaaccgag gtaactggct tggaggagcg cagtcaccaa	6600
aacttgcctt ttcagtttgcgtt ccttaaccgg cgcgtactt caagactaacatccat	6660
caattaccag tggctgtgc cagttgtgtt tttgcgtatgc tttccgggtt ggactcaaga	6720
cgtatgttac cggataaggc gcagcggcgtq gactgaacgg ggggttcgtg catabactcc	6780
agcttggagc gaaactgccta cccggaaactg agtgcgttgcgtt gttggatgag acaaacgcgg	6840
ccataaacgc ggaatgacac cggtaaaccg aaaggcagga acaggagagc gcacgaggaa	6900
ggccgcaggaa gaaacgcctt gttttttttt agtgcgttgcgtt gttttgcctt ccactgttt	6960
gagcgtcaga ttctgtatgc tttgtcagggg gggggggcgttatggaaaaa cggcttgcctt	7020
ggggccctctt cacttccctt ttaagtatgc tccgtatgc tttccaggaaa tctccgcctt	7080
tttcgtatgc cattttccgtt cggccgttgcgtt gttttttttt agtgcgttgcgtt gttttgcctt	7140
aggaagcgaa atatatccgt tttttttttt tttttttttt tttttttttt tttttttttt	7200
tcttgcacatc tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt	7260
acactccgtt acgcgtatgt tccggccgtt tttttgcgtt tttttttttt tttttttttt	7320
gctgaacagg agggacagct gatagaaaca gaagccactg gggccatca aaaacaccat	7380
catacactaa atcagtaagt tggcagcatac acccgacgca ctttgcgcgg aataaatacc	7440

-continued

tgtgacggaa gatcaacttcg cagaataaaat aaatccttgtt gtccctgttg ataccggaa	7500
gccctggggcc aacttttggc gaaaatgaga cgttgategg cacgttaagag gttccaactt	7560
tcaccataat gaaataagat cactaccggg cgtatTTTT gagttatcgA gatTTcagg	7620
agctaaggaa gctaaaatgg agaaaaaaat cactggatat accaccgttg atatatccca	7680
atggcatcgta aaagaacatt ttgaggcatt tcagttagtt gctcaatgtaa cctataacc	7740
gaccgtttagt ctggatataa cggccTTTT aaagaccgtt aagaaaaata agcacaagtt	7800
ttatccggcc tttattcaca ttcttgcccg cctgtatggat gtcatccgg aattc	7855

```
<210> SEQ ID NO 46
<211> LENGTH: 5144
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)...(5144)
```

<400> SEQUENCE: 46

taatacact cactataggg gaattgtgag cggataacaa ttcccctcta gaaataattt
tgtttaact taagaaggag accagttct agcataaccc ctggggcct ctaaacgggt 60
cttgaggggg ttttgctga aaggagaac tatatccgga ttggcgaatg ggacgcgcc 120
tgtagcggcg cattaagcgc ggcgggtgtg gtggtaacgc gcagcgtgac cgctacactt
ggcagcgc 300
cc tagcgcgc tccttcgc ttcttcctt cctttcgc cacttcgc
ggcttcccc 360
gtcaagctct aaatcgaaaa ctcctttag gttccgatt tagtgctta
ccgcaccc 420
cc acttgattag ggtgatggg cacgtagtgg gccatcgccc
tgatagacgg 480
ttttcgccc tttgacggtt gagtccacgt tcttaatag tggactcttgc
ttccaaactg 540
gaacaacact caaccctatc tcggctattt ctttgcattt ataagggatt
tttgcgatt 600
cgccctatc gttaaaaat gagctgatt aacaaaaatt taacgcgaat
tttacaaa 660
tattaacggtt tacaatttca ggtggcactt ttggggaaa tgcgcggaa
accctatattt 720
gtttatTTT ctaaatacat tcaaataatgt atccgcctcat gaattaattc
tttagaaaaac 780
tcatcgagca tcaaataatgaa ctgcaatttca ttcatatcg gattatcaat
accatatttt 840
tgaaaaagcc gtttctgtaa tgaaggagaa aactcaccga ggcagttcca
taggatggca 900
agatcctggt atcggcttc gattccgact cgtccaaacat caatacaacc
tattaatttc 960
ccctcgtaa aaataaggtt atcaagttagg aaatcaccat gagtgacgac
tgaatccgg 1020
gagaatggca aaagtttatg catttttc cagactgtt caacaggcca
gcattacgc 1080
tgcgtcatcaa aatcactcgatc atcaacccaa cgcgttattca ttgcgttattg
cgccctgagcg 1140
agacgaaata cgcgcgtgcgtt gttaaaagga caattacaaa caggaatcgatc
atgcacccgg 1200
cgcaggaaaca ctgcccgcgc atcaacaata tttcacctg aatcaggata
ttcttctaat 1260
accttggaaatg ctgtttccc gggatcgca gtggtagtta accatgcattc
atcaggagta 1320
cggataaaaat gtttgcattt cggaaagggc ataaattccg tcaggcattt
tagtctgacc 1380
atctcatctg taacatcatt ggcaacgta cctttgcattt gtttcaaaaa
caactctggc 1440
gcacgggct tcccatacaa tcgatagatt gtgcgcaccc attggccgac

-continued

attatcgca gcccatttat acccatataa atcagcatcc atgttggaat ttaatcgccg	1500
ccttagagcaa gacgtttccc gttgaatatg gtcataaca ccccttgtat tactgtttat	1560
gtaaagcac agttttatg ttcatgacca aaatcccta acgtgagtt tcgttccact	1620
gagcgtcaga ccccgtagaa aagatcaaag gatcttctg agatccttt tttctgcgcg	1680
taatctgctg cttgcaaaca aaaaaccac cgctaccagc ggtggttgt ttgcccgtc	1740
aagagctacc aactttttt ccgaaggtaa ctggcttcag cagagcgcag ataccaaata	1800
ctgtccttct agtgtagccg tagttaggcc accacttcaa gaactctgta gcaccgccta	1860
catacctcgc tctgctaatac ctgttaccag tggctgtgc cagtggcgat aagtctgtc	1920
ttaccgggtt ggactcaaga cgatagttac cggtataaggc gcagcggtcg ggctgaacgg	1980
ggggttcgtg cacacagccc agcttggago gaacgcacca caccgaactg agataacctac	2040
agegtgagct atgagaaagc gccaacgttc cccaaggggag aaaggcggac aggtatccgg	2100
taagcggcag ggtcggaaaca ggagagcgcgca cgagggagct tccagggggaa aacgcctgg	2160
atctttatacg tccctgtcggt tttcgccacc tctgacttga gctcgattt ttgtgtatgc	2220
cgtcagggggg gggggggctta tgaaaaaacg ccagcaacgc ggcctttta cggttctgg	2280
ccttttgcgt gccttttgcgt cacatgttct ttcctgcgtt atccccgtat tctgtggata	2340
accgttattac cgcctttgag tgagctgata ccgcgcgcg cagccgaacg accgagcgcga	2400
gegagtcagt gagcggaggaa gcggaaagagc gcctgatgcg gtatttctc cttacgcata	2460
tgtgcggat ttcacaccgc atatatggtg cactctcgtt acaatctgct ctgtatccgc	2520
atagttaaacg cagtatacac tccgtatcg ctacgtgact gggtcatggc tgcccccga	2580
cacccgccaa cacccgctga cgcgcctga cgggcttgc tgctccggc atccgcttac	2640
agacaagctg tgaccgtctc cggagactgc atgtgtcaga gtttttacc gtcatcaccg	2700
aaacgcgcga ggcagtcgc gttaagctca tcagcgtggc cgtgaagcga ttacagatg	2760
tctgcctgtt catccgcgtc cagtcgtt agtttctcca gaagcgttaa tgtctggctt	2820
ctgataaaacg gggccatgtt aaggcgggtt tttccgtt tggtcactga tgccctccgt	2880
taagggggat ttctgttcat ggggttaatg ataccgatga aacgagagag gatgtcactcg	2940
atacgggtta ctgatgatga acatgcccg ttactggaaac gttgtgaggg taaacaactg	3000
gccccatggta tgccggggaa ccagaaaaaa atcactcagg gtcaatgcca ggcgttcgtt	3060
aatacagatg taggtgttcc acagggttagc cagcagacatc ctgcgtatgc gatccggAAC	3120
ataatgggtc agggcgctga ctccgcgtt tccagacttt acgaaacacg gaaaccgaaag	3180
accattcatg ttgttgctca gggtcgac gttttgcagc agcagtcgtt tcacgttcgc	3240
tcgcgtatcg gtgattcatt ctgctaaaccat gtaaggcaac cccgcaccc tagccgggtc	3300
ctcaacgaca ggagcacgtat cgtgcgcacc cgtggggccg ccatgcggc gataatggcc	3360
tggttctcgcc cggaaacgttt ggtggggggaa ccagtgacca aggcttgcgc gggggcgtgc	3420
aagattccga ataccgcaag cgacaggccg atcatcgatc cgctccagcg aaagcggtcc	3480
tcgcccggaaa tgacccagag cgctgcccgc acctgttccat cgagttgcgtt gataaagaag	3540
acagtctataa gtgcggcgac gatagtcatg ccccgccccc accggaaagga gctgactggg	3600
ttgaaggctc tcaagggcat cggtcagat cccgtgcctt aatgagtgag ctaacttaca	3660
ttaattcgct tgcgctcaact gccegttttc cagtcggggaa acctgtcgat ccagtcgtat	3720

-continued

taatgaatcg	gcacaacgcgc	ggggagaggg	ggttgcgt	ttggggccca	gggtggttt	3780
tctttcacc	agtgagacgg	gcaacagctg	attgcccttc	accgcctggc	cctgagagag	3840
ttgcagcaag	cggtccacgc	tggttgccc	cagcaggcga	aaatccgtt	tgatgggt	3900
taacgggggg	atataacatg	agctgtcttc	ggtatcgctg	tatcccacta	ccgagatatc	3960
cgcaccaacg	cgcagccgg	actcggtaat	ggcgccgtt	gcccggcgt	ccatctgatc	4020
gttggcaacc	agcatecgag	tgggaacgt	gccctcatc	agcattgca	tggttgtt	4080
aaaaccggac	atggcactcc	agtcgccttc	ccgttccgt	atcggtgaa	tttgattgcg	4140
agtgagatat	ttatgccagc	cagccagacg	cagacgcgcc	gagacagaac	ttaatggcc	4200
cgcataacagc	cgcatttgct	ggtgacccaa	tgcgaccaga	tgctccacgc	ccagtcgcgt	4260
accgtcttca	tgggagaaaa	taatactgtt	gatgggtgtc	tggtcagaga	catcaagaaa	4320
taacggccgg	acattagtgc	aggcagcttc	cacagcaatg	gcatccgtt	catccagcgg	4380
atagttaatg	atcageccac	tgacgcgtt	cgcgagaaga	ttgtgcaccc	ccgctttaca	4440
gggttcgacg	ccgcttcgtt	ctaccatcg	caccaccacg	ctggcaccca	gttgcgttgc	4500
gcgagattta	atcgccgcga	caatttgcga	cggcgctgtc	agggccagac	tggaggtggc	4560
aacgcataatc	gcacacgact	gtttccccgc	cagttgtgt	gacacgoggt	tggaaatgta	4620
attcagctcc	gcacatgcgcg	cttccacttt	ttccgcgtt	ttcgcagaaa	cgtggctggc	4680
ctgggttacc	acgcggggaa	cggctcgata	agagacaccc	gcataactctg	cgacatcgta	4740
taacgttact	ggtttacat	tcaccaccc	gaattgactc	tcttcgggc	gttatcatgc	4800
cataccgcga	aaggtttgc	gccattcgat	ggtgtccggg	atctcgacgc	tctcccttat	4860
gcgactctg	cattaggaag	cagccagta	gttaggtttag	gccgttgagc	accgeccgg	4920
caaggaatgg	tgcgtgcga	gagatggcgc	ccaacagtcc	ccggccacacg	gggcctgcca	4980
ccataaccac	gcccggaa	gcccgtatga	gcccgaagt	gcccggccga	tcttccccat	5040
cggtgtatgc	ggcgatata	gcccacccgt	ggcgcgggt	atgcggccca		5100
cgatgcgtcc	ggcgtagagg	atcgagatct	cgatccccgc	aaat		5144

```

<210> SEQ ID NO 47
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: promoter
<222> LOCATION: (1)..(19)
<223> OTHER INFORMATION: T7 promoter

<400> SEQUENCE: 47

```

taatacgtact	cactatagg	19
-------------	-----------	----

```

<210> SEQ ID NO 48
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)..(25)
<223> OTHER INFORMATION: lac operator

```

-continued

<400> SEQUENCE: 48	
ggaattgtga gcgataaca attcc	25
<210> SEQ ID NO 49	
<211> LENGTH: 456	
<212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: synthetic	
<220> FEATURE:	
<221> NAME/KEY: rep_origin	
<222> LOCATION: (1)..(456)	
<223> OTHER INFORMATION: f1 ori	
<400> SEQUENCE: 49	
acgcgcctg tagcggcgca ttaagcgccg cgggtgtggt ggttacgcgc agcgtgaccg	60
ctacacttgc cagcgccta gcgcgcgc tc ttccgtttt ctcccttcc ttctcgcca	120
cgttcgcgg ctcccccgtt caagctctaa atcgggggtt cccttaggg ttccgattta	180
gtgttttacg gcacctcgac cccaaaaaac ttgatttaggg tgatggttca cgtagtggc	240
catcgccctg atagacgggtt ttccgcctt tgacgttggt gtccacgttc tttaatagt	300
gactcttggt ccaaacttggt acaacactca accctatctc ggtctattct ttgttattat	360
aagggtttt gccgatttcg gcctatttgtt taatggatgtttaa gctgatttaa caaaaattta	420
acgcgaaattt taacaaaata ttaacgttta caattt	456
<210> SEQ ID NO 50	
<211> LENGTH: 816	
<212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: synthetic	
<220> FEATURE:	
<221> NAME/KEY: CDS	
<222> LOCATION: (1)..(816)	
<223> OTHER INFORMATION: KanR	
<400> SEQUENCE: 50	
atg agc cat att caa cgg gaa acg tct tgc tct agg cgg cga tta aat Met Ser His Ile Gln Arg Glu Thr Ser Cys Ser Arg Pro Arg Leu Asn 1 5 10 15	48
tcc aac atg gat gct gat tta tat ggg tat aaa tgg gct cgc gat aat Ser Asn Met Asp Ala Asp Leu Tyr Gly Tyr Lys Trp Ala Arg Asp Asn 20 25 30	96
gtc ggg caa tca ggt gcg aca atc tat cga ttg tat ggg aag ccc gat Val Gly Gln Ser Gly Ala Thr Ile Tyr Arg Leu Tyr Gly Lys Pro Asp 35 40 45	144
gcg cca gag ttg ttt ctg aaa cat ggc aaa ggt agc gtt gcc aat gat Ala Pro Glu Leu Phe Leu Lys His Gly Lys Ser Val Ala Asn Asp 50 55 60	192
gtt aca gat gag atg gtc aga cta aac tgg ctg acg gaa ttt atg cct Val Thr Asp Glu Met Val Arg Leu Asn Trp Leu Thr Glu Phe Met Pro 65 70 75 80	240
ctt ccg acc atc aag cat ttt atc cgt act cct gat gat gca tgg tta Leu Pro Thr Ile Lys His Phe Ile Arg Pro Asp Asp Ala Trp Leu 85 90 95	288
ctc acc act gcg atc ccc ggg aaa aca gca ttc cag gta tta gaa gaa Leu Thr Thr Ala Ile Pro Gly Lys Thr Ala Phe Gln Val Leu Glu Glu 100 105 110	336

-continued

tat cct gat tca ggt gaa aat att gtt gat gcg ctg gca gtg ttc ctg Tyr Pro Asp Ser Gly Glu Asn Ile Val Asp Ala Leu Ala Val Phe Leu 115 120 125	384
cgc cgg ttg cat tcg att cct gtt tgt aat tgt cct ttt aac agc gat Arg Arg Leu His Ser Ile Pro Val Cys Asn Cys Pro Phe Asn Ser Asp 130 135 140	432
cgc gta ttt cgt ctc gct cag gcg caa tca cga atg aat aac ggt ttg Arg Val Phe Arg Leu Ala Gln Ala Gln Ser Arg Met Asn Asn Gly Leu 145 150 155 160	480
gtt gat gcg agt gat ttt gat gac gag cgt aat ggc tgg cct gtt gaa Val Asp Ala Ser Asp Phe Asp Asp Glu Arg Asn Gly Trp Pro Val Glu 165 170 175	528
caa gtc tgg aaa gaa atg cat aaa ctt ttg cca ttc tca ccg gat tca Gln Val Trp Lys Glu Met His Lys Leu Leu Pro Phe Ser Pro Asp Ser 180 185 190	576
gtc gtc act cat ggt gat ttc tca ctt gat aac ctt att ttt gac gag Val Val Thr His Gly Asp Phe Ser Leu Asp Asn Leu Ile Phe Asp Glu 195 200 205	624
ggg aaa tta ata ggt tgt att gat gtt gga cga gtc gga atc gca gac Gly Lys Leu Ile Gly Cys Ile Asp Val Gly Arg Val Gly Ile Ala Asp 210 215 220	672
cga tac cag gat ctt gcc atc cta tgg aac tgc ctc ggt gag ttt tct Arg Tyr Gln Asp Leu Ala Ile Leu Trp Asn Cys Leu Gly Glu Phe Ser 225 230 235 240	720
cct tca tta cag aaa cgg ctt ttt caa aaa tat ggt att gat aat cct Pro Ser Leu Gln Lys Arg Leu Phe Gln Lys Tyr Gly Ile Asp Asn Pro 245 250 255	768
gat atg aat aaa ttg cag ttt cat ttg atg ctc gat gag ttt ttc taa Asp Met Asn Lys Leu Gln Phe His Leu Met Leu Asp Glu Phe Phe 260 265 270	816

<210> SEQ ID NO 51

<211> LENGTH: 271

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 51

Met Ser His Ile Gln Arg Glu Thr Ser Cys Ser Arg Pro Arg Leu Asn 1 5 10 15
Ser Asn Met Asp Ala Asp Leu Tyr Gly Tyr Lys Trp Ala Arg Asp Asn 20 25 30
Val Gly Gln Ser Gly Ala Thr Ile Tyr Arg Leu Tyr Gly Lys Pro Asp 35 40 45
Ala Pro Glu Leu Phe Leu Lys His Gly Lys Ser Val Ala Asn Asp 50 55 60
Val Thr Asp Glu Met Val Arg Leu Asn Trp Leu Thr Glu Phe Met Pro 65 70 75 80
Leu Pro Thr Ile Lys His Phe Ile Arg Thr Pro Asp Asp Ala Trp Leu 85 90 95
Leu Thr Thr Ala Ile Pro Gly Lys Thr Ala Phe Gln Val Leu Glu Glu 100 105 110
Tyr Pro Asp Ser Gly Glu Asn Ile Val Asp Ala Leu Ala Val Phe Leu 115 120 125
Arg Arg Leu His Ser Ile Pro Val Cys Asn Cys Pro Phe Asn Ser Asp 130 135 140

-continued

Arg Val Phe Arg Leu Ala Gln Ala Gln Ser Arg Met Asn Asn Gly Leu
145 150 155 160

Val Asp Ala Ser Asp Phe Asp Asp Glu Arg Asn Gly Trp Pro Val Glu
165 170 175

Gln Val Trp Lys Glu Met His Lys Leu Leu Pro Phe Ser Pro Asp Ser
180 185 190

Val Val Thr His Gly Asp Phe Ser Leu Asp Asn Leu Ile Phe Asp Glu
195 200 205

Gly Lys Leu Ile Gly Cys Ile Asp Val Gly Arg Val Gly Ile Ala Asp
210 215 220

Arg Tyr Gln Asp Leu Ala Ile Leu Trp Asn Cys Leu Gly Glu Phe Ser
225 230 235 240

Pro Ser Leu Gln Lys Arg Leu Phe Gln Lys Tyr Gly Ile Asp Asn Pro
245 250 255

Asp Met Asn Lys Leu Gln Phe His Leu Met Leu Asp Glu Phe Phe
260 265 270

<210> SEQ ID NO 52

<211> LENGTH: 589

<212> TYPE: DNA

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthetic

<220> FEATURE:

<221> NAME/KEY: rep_origin

<222> LOCATION: (1)..(589)

<223> OTHER INFORMATION: pBR322_ori

<400> SEQUENCE: 52

tttagatcct tttttctgc gcgtaatctg ctgcttgcaa aaaaaaaaaac caccgctacc 60

agcggtgggtt tgtttgcgg atcaagagct accaacttctt ttccgaaagg taactggctt 120

cagcagagcg cagataccaa atactgtcct tctagtgttag ccgttagtttag gccaccactt 180

caagaactct gtagcaccgc ctatacacct cgctctgcta atccctgttac cagtggctgc 240

tgcgcgtggc gataagtgcgt gtcttaccgg gttggactca agacgatagt taccggataa 300

ggcgccagcg tcggggctgaa cgggggggttc gtgcacacag cccagttgg agcgaacgac 360

ctacaccgaa ctgagatacc tacagcgtga gctatgagaa agcgccacgc ttcccgaagg 420

gaaaaaggcg gacaggtatc cggtaagcgg cagggtcggaa acaggagagc gcacgaggga 480

gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc acctctgact 540

tgagcgtcga tttttgtgat gtcgtcagg ggggcggagc ctatggaaa 589

<210> SEQ ID NO 53

<211> LENGTH: 143

<212> TYPE: DNA

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: synthetic

<220> FEATURE:

<221> NAME/KEY: misc_feature

<222> LOCATION: (1)..(143)

<223> OTHER INFORMATION: bom

<400> SEQUENCE: 53

cctgatcggtt tattttctcc ttacgcatct gtgcggattt tcacaccgca tataatggtc 60

actctcagta caatctgctc tgatgccgca tagttaagcc agtatacact ccgctatgc 120

-continued

tacgtgactg ggtcatggct gcg	143
<210> SEQ ID NO 54	
<211> LENGTH: 192	
<212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: synthetic	
<220> FEATURE:	
<221> NAME/KEY: misc_feature	
<222> LOCATION: (1)..(192)	
<223> OTHER INFORMATION: rop	
<400> SEQUENCE: 54	
gtgacccaaac aggaaaaaac cgcccttaac atggcccgct ttatcagaag ccagacatta	60
acgcttctgg agaaaactcaa cgagctggac gcggatgaac aggcagacat ctgtgaatcg	120
cttcacgacc acgctgatga gcttaccgc agctgcctcg cgcgttcgg tcatgacggt	180
gaaaacctct ga	192
<210> SEQ ID NO 55	
<211> LENGTH: 78	
<212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: synthetic	
<220> FEATURE:	
<221> NAME/KEY: promoter	
<222> LOCATION: (1)..(78)	
<223> OTHER INFORMATION: lacI promoter	
<400> SEQUENCE: 55	
gacaccatcg aatggcgcaa aaccttcgc ggtatggcat gatacgccc ggaagagagt	60
caattcaggg tggtaat	78
<210> SEQ ID NO 56	
<211> LENGTH: 1083	
<212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: synthetic	
<220> FEATURE:	
<221> NAME/KEY: CDS	
<222> LOCATION: (1)..(1083)	
<223> OTHER INFORMATION: lacI encoding lac repressor	
<400> SEQUENCE: 56	
gtg aaa cca gta acg tta tac gat gtc gca gag tat gcc ggt gtc tct	48
Val Lys Pro Val Thr Leu Tyr Asp Val Ala Glu Tyr Ala Gly Val Ser	
1 5 10 15	
tat cag acc gtt tcc cgc gtg gtg aac cag gcc agc cac gtt tct gcg	96
Tyr Gln Thr Val Ser Arg Val Val Asn Gln Ala Ser His Val Ser Ala	
20 25 30	
aaa acg cgg gaa aaa gtg gaa gcg gcg atg gcg gag ctg aat tac att	144
Lys Thr Arg Glu Lys Val Ala Ala Met Ala Glu Leu Asn Tyr Ile	
35 40 45	
ccc aac cgc gtg gca caa caa ctg gcg ggc aaa cag tcg ttg ctg att	192
Pro Asn Arg Val Ala Gln Gln Leu Ala Gly Lys Gln Ser Leu Leu Ile	
50 55 60	
ggc gtt gcc acc tcc agt ctg gcc ctg cac gcg ccg tcg caa att gtc	240
Gly Val Ala Thr Ser Ser Leu Ala Leu His Ala Pro Ser Gln Ile Val	
65 70 75 80	

-continued

gcg gcg att aaa tct cgc gcc gat caa ctg ggt gcc agc gtg gtg Ala Ala Ile Lys Ser Arg Ala Asp Gln Leu Gly Ala Ser Val Val Val 85 90 95	288
tcg atg gta gaa cga agc ggc gtc gaa gcc tgt aaa gcg gcg gtg cac Ser Met Val Glu Arg Ser Gly Val Glu Ala Cys Lys Ala Val His 100 105 110	336
aat ctt ctc gcg caa cgc gtc agt ggg ctg atc att aac tat ccg ctg Asn Leu Leu Ala Gln Arg Val Ser Gly Leu Ile Ile Asn Tyr Pro Leu 115 120 125	384
gat gac cag gat gcc att gct gtg gaa gct gcc tgc act aat gtt ccg Asp Asp Gln Asp Ala Ile Ala Val Glu Ala Ala Cys Thr Asn Val Pro 130 135 140	432
gcg tta ttt ctt gat gtc tct gac cag aca ccc atc aac agt att att Ala Leu Phe Leu Asp Val Ser Asp Gln Thr Pro Ile Asn Ser Ile Ile 145 150 155 160	480
ttc tcc cat gaa gac ggt acg cga ctg ggc gtg gag cat ctg gtc gca Phe Ser His Glu Asp Gly Thr Arg Leu Gly Val Glu His Leu Val Ala 165 170 175	528
ttg ggt cac cag caa atc gcg ctg tta gcg ggc cca tta agt tct gtc Leu Gly His Gln Gln Ile Ala Leu Leu Ala Gly Pro Leu Ser Ser Val 180 185 190	576
tcg gcg cgt ctg cgt gct ggc tgg cat aaa tat ctc act cgc aat Ser Ala Arg Leu Arg Leu Ala Gly Trp His Lys Tyr Leu Thr Arg Asn 195 200 205	624
caa att cag ccg ata gcg gaa cgg gaa ggc gac tgg agt gcc atg tcc Gln Ile Gln Pro Ile Ala Glu Arg Glu Gly Asp Trp Ser Ala Met Ser 210 215 220	672
ggg ttt caa caa acc atg caa atg ctg aat gag ggc atc gtt ccc act Gly Phe Gln Gln Thr Met Gln Met Leu Asn Glu Gly Ile Val Pro Thr 225 230 235 240	720
gcg atg ctg gtt gcc aac gat cag atg gcg ctg ggc gca atg cgc gcc Ala Met Leu Val Ala Asn Asp Gln Met Ala Leu Gly Ala Met Arg Ala 245 250 255	768
att acc gag tcc ggg ctg cgc gtt ggt gcg gat atc tcg gta gtg gga Ile Thr Glu Ser Gly Leu Arg Val Gly Ala Asp Ile Ser Val Val Gly 260 265 270	816
tac gac gat acc gaa gac agc tca tgt tat atc ccg ccg tta acc acc Tyr Asp Asp Thr Glu Asp Ser Ser Cys Tyr Ile Pro Pro Leu Thr Thr 275 280 285	864
atc aaa cag gat ttt cgc ctg ctg ggg caa acc agc gtg gac cgc ttg Ile Lys Gln Asp Phe Arg Leu Leu Gly Gln Thr Ser Val Asp Arg Leu 290 295 300	912
ctg caa ctc tct cag ggc cag gcg gtg aag ggc aat cag ctg ttg ccc Leu Gln Leu Ser Gln Gly Gln Ala Val Lys Gly Asn Gln Leu Leu Pro 305 310 315 320	960
gtc tca ctg gtg aaa aga aaa acc acc ctg gcg ccc aat acg caa acc Val Ser Leu Val Lys Arg Lys Thr Thr Leu Ala Pro Asn Thr Gln Thr 325 330 335	1008
gcc tct ccc cgc gcg ttg gcc gat tca tta atg cag ctg gca cga cag Ala Ser Pro Arg Ala Leu Ala Asp Ser Leu Met Gln Leu Ala Arg Gln 340 345 350	1056
gtt tcc cga ctg gaa agc ggg cag tga Val Ser Arg Leu Glu Ser Gly Gln 355 360	1083

-continued

<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 57

Val Lys Pro Val Thr Leu Tyr Asp Val Ala Glu Tyr Ala Gly Val Ser
1 5 10 15

Tyr Gln Thr Val Ser Arg Val Val Asn Gln Ala Ser His Val Ser Ala
20 25 30

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

Pro Asn Arg Val Ala Gln Gln Leu Ala Gly Lys Gln Ser Leu Leu Ile
50 55 60

Gly Val Ala Thr Ser Ser Leu Ala Leu His Ala Pro Ser Gln Ile Val
65 70 75 80

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

Ser Met Val Glu Arg Ser Gly Val Glu Ala Cys Lys Ala Val His
100 105 110

Asn Leu Leu Ala Gln Arg Val Ser Gly Leu Ile Ile Asn Tyr Pro Leu
115 120 125

Asp Asp Gln Asp Ala Ile Ala Val Glu Ala Ala Cys Thr Asn Val Pro
130 135 140

Ala Leu Phe Leu Asp Val Ser Asp Gln Thr Pro Ile Asn Ser Ile Ile
145 150 155 160

Phe Ser His Glu Asp Gly Thr Arg Leu Gly Val Glu His Leu Val Ala
165 170 175

Leu Gly His Gln Gln Ile Ala Leu Leu Ala Gly Pro Leu Ser Ser Val
180 185 190

Ser Ala Arg Leu Arg Leu Ala Gly Trp His Lys Tyr Leu Thr Arg Asn
195 200 205

Gln Ile Gln Pro Ile Ala Glu Arg Glu Gly Asp Trp Ser Ala Met Ser
210 215 220

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

Ala Met Leu Val Ala Asn Asp Gln Met Ala Leu Gly Ala Met Arg Ala
245 250 255

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

Tyr Asp Asp Thr Glu Asp Ser Ser Cys Tyr Ile Pro Pro Leu Thr Thr
275 280 285

Ile Lys Gln Asp Phe Arg Leu Leu Gly Gln Thr Ser Val Asp Arg Leu
290 295 300

Leu Gln Leu Ser Gln Gly Gln Ala Val Lys Gly Asn Gln Leu Leu Pro
305 310 315 320

Val Ser Leu Val Lys Arg Lys Thr Thr Leu Ala Pro Asn Thr Gln Thr
325 330 335

Ala Ser Pro Arg Ala Leu Ala Asp Ser Leu Met Gln Leu Ala Arg Gln
340 345 350

Val Ser Arg Leu Glu Ser Gly Gln
355 360

-continued

```
<210> SEQ ID NO 58
<211> LENGTH: 7772
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)..(7772)
<223> OTHER INFORMATION: Plasmid pP450

<400> SEQUENCE: 58

taatacgaact cactataggg gaattgttag cggtataacaa ttccccctcta gaaataattt 60
tgtttaactt taagaaggag actcgaccat ggaattatca caagtttgta caaaaaagga 120
ggctggcgcc ggaaccaatt cagtcgactg gatccaagaa ggagatataa ccccctcagc 180
agcgatggcg acaatggagg tagaggccgc ggccgccacg gtgctggccg cgccttgct 240
gttccctccg cgcatactca aactgctgtt attcgttagtg acgctcttgtt acctggcccg 300
agccctgagg cggccacgca aaagcaccac caagtgcagc agcacaacgt ggcgcctcgcc 360
cccgccggc gttggcaacc cggcgctccc accgggtccc gtgcccgtggc ccgtcgccgg 420
caaacctggcg gagatgctgc tgaacaagec ggcattccgc tggtatccacc agatgtatgcg 480
cgagatgggc acggacatcg cctgcgtcaa gcttggcgcc gtccacgtcg tgtccatcac 540
ctggccggat atcgcgcggg aggtgctcg gaagcaggac gcacacttca tatcccgccc 600
gctcacccctc gcctccgaga cgttcagcgg cgggtacccgg aacgcgcgtgc tctgcctca 660
ccggcgaccag tggaagaaga tgcgcgcgt cctcacccctc gagatcatct gcccgtccgg 720
ccacgcctgg ctccacgaca agcgcaccga cgaggccgac aacctcaccc gctaegtcata 780
caaacctcgcc accaaagccg ccacccggca cgtgcgcgtc gacgtcaggc acgtcgctcg 840
tcactattgc ggcaacgtt tccgcgcct catgttcaac aggcgctact tcggcgagcc 900
ccaggctgac ggccgtccgg ggccgatggaa ggtgctgtat atggacgcgg tggtcacctc 960
cctcggeetc ctctacgcct tctgcgtctc cgactacccctc ccctggctgc gggcctcgat 1020
cctcgaeggc cacgagaaga tcgtcaaggaa ggttaacgtg ggggtgaaca ggctccacga 1080
cacggctatc gacgaccgggt ggaggcagtg gaagagcggc gagcggcagg agatggagga 1140
cttccttggat gtgctcatca ctctcaaggaa cggccaggccg aacccgcgtc tgaccatcg 1200
ggaggtcaaa ggcgcgtcac aggacatcac gttcgccgcgt gtggacaacc cgtcgaaacgc 1260
cgtggagtg ggcgcgtccg agatggtaa caacccggag gtgtatggcga aggcgatggaa 1320
ggaggtgtac cgcgcgtcg gacggggagag gctagtgacg gagtcggacaca ttccgaagct 1380
caactacgtg aaggcctgca tccggggaggg tttccgtctc caccgggtgg cgccttcaa 1440
cgtgccccac gtcgcgcgtc cgcacaccac catgcgcggc taccgcgttc ccaaggccg 1500
ccacgtgatc ctgagccgca cggggctggg ccgcaccccg cgcgtgtggg acgagccct 1560
gcgcgttctac cgggaccgac acctcgccac cgcgcgtcc gacgtcgccgc tcaccgagaa 1620
cgacctgcgg ttcatctcc tcagcaccgg ccgcgcggc tgcacgcgtcc cgtcgctcg 1680
caccggccatg agcgtcatgc tcttcggaaag gtcctgcgtc ggggtcacct ggagcaagcc 1740
cgccgggggtg gaggccgtgg acctcagcga gtccaaagacg gacacccatca tggccacccc 1800
gctgggtgtc gacgctgagc ccaggctgac ggcgcaccc tcaccgttca tctccatcg 1860
```

-continued

atcgccagct gagggagaaa acctgtactt ccagggtcaa ttcagcaag gagaagaact 1920
ttcactgga gttgtccaa ttcttgtga attagatgt gatgttaatg ggcacaatt 1980
ttctgtcagt ggagagggtg aaggtatgc tacatacggaa aactcaccc ttaaaatttat 2040
ttgcactact ggaaaactac ctgttccatg gccaacactt gtcactactc tgaccatgg 2100
tgttcaatgc ttttccgtt atccggatca catgaaacgg catgacttt tcaagagtgc 2160
catgcccga ggttatgtac aggaacgcac tatatcttc aaagatgacg ggaactacaa 2220
gacgcgtgct gaagtcaagt ttgaagggtga tacccttgtt aatcgatcg agttaaagg 2280
tattgatttt aaagaagatg gaaacattct cgacacaaa ctagagtaca actataactc 2340
acacaatgtat tacatcacgg cagacaaaca aaagaatgga atcaaagcta acttcaaaat 2400
tcgcgcacaac attgaagatg gttccgttca actagcagac cattatcaac aaaatactcc 2460
aattggcgtat gggccgttcc ttttaccaga caaccattac ctgtcgacac aatctgcct 2520
ttcgaaagat cccaaacgaaa agcgtgacca catggtcctt cttgagttt taactgtgc 2580
tgggattaca catggcatgg atgagctcta caaaaagttt gggcccatc atcatcacca 2640
ccaccaccac tgagatccgg ctgctaaaca agccccaaaag gaagctgagt tggctgtgc 2700
caccgctgag caataacttag cataaccctt tggggcctctt aaacgggtct tgaggggttt 2760
tttgctgaaa ggaggaacta tatccggatt ggcaatggg acgcgcctg tagggcgca 2820
ttaagcgccgg cgggtgtggt ggttacgcgc agegtgaccc ctacacttgc cagcgcccta 2880
gegcgcgcctt cttecgctt cttecccttc ttctcgccca cgttcgccgg ctteccccgt 2940
caagctctaa atcgccggctt cccttaggg ttccgattta gtgccttacg gcacctcgac 3000
ccaaaaaaac ttgatttaggg ttaggttca cgtatgggc catggccctg atagacggtt 3060
tttcgcctt tgacgttggat gtcacgttca tttaatagtg gactcttgtt ccaaactgg 3120
acaacactca accctatctc ggttatattttt tttgattttt aagggattttt gccgatttcg 3180
gectattggt taaaaaatgat gctgatttaa caaaaattta acgcgaattt taacaaaata 3240
ttaacgttta caatttcagg tggcactttt cggggaaatg tgccggaaac ccctatttgt 3300
ttattttctt aaatacatc aaatatgtat ccgctcatga attaattctt agaaaaactc 3360
atcgagcatc aaatgaaact gcaattttt catatcaggat ttatcaatac catatttttg 3420
aaaaagccgt ttctgtatg aaggagaaaaa ctcaccgagg cagttccata ggttgcac 3480
atcctggat cggctcgca ttccgactcg tccaaacatca atacaaccta ttaatttccc 3540
ctcgtcaaaa ataagggttat caagtggaaa atcaccatgat gtcacgttca aatccggat 3600
gaatggcaaa agtttatgca ttctttcca gacttgttca acaggccagc cattacgctc 3660
gtcatcaaaa tcactcgcat caacaaacc gttatttattt cgtgatttgc cctgagcgag 3720
acgaaatacg cgatcgctgt taaaaggaca attacaaaca ggaatcgaaat gcaaccggcg 3780
cagggacact gccagcgcat caacaatattt ttccacgttca tcaggatattt cttctataac 3840
ctggaaatgtt gtttccggg ggatcgcat ggtgatgttca catgcatcat caggagtgac 3900
gataaaatgc ttgatggatc gaagaggcat aaattccgtc agccagttt gtcgtaccat 3960
ctcatctgtat acatcattgg caacgttacc tttgcatgtt ttcagaaaca actctggcgc 4020
atcgccgttcc ccatataatc gatagattgtt cgcacgtat tgcccgacat tatcgac 4080
ccatattataat ccatataat cagcatccat gttggattt aatcgccggcc tagagcaaga 4140

-continued

cgtttccgt tgaatatggc tcataaacacc ccttgttatta ctgtttatgt aagcagacag	4200
ttttattgtt catgaccaaaa atcccttaac gtgagtttc gttccactga gcgtcagacc	4260
ccgttagaaaa gatcaaagga tcttctttagat atccctttttt tctgcgcgta atctgctgct	4320
tgcaaaacaaa aaaaccaccg ctaccagcgg tggtttggtt gccggatcaa gagctaccaa	4380
ctcttttcc gaaggtaact ggcttcagca gagcgcagat accaaataact gtccttctag	4440
tgttagccgt a ttaggccac cacttcaaga actctgttagc accgcctaca tacctcgctc	4500
tgctaattcc tttaccatgt gctgtgcga gtggcgataa gtcgtgtctt accgggttgg	4560
actcaagacg atagttaccg gataaggcgc a ggggtcggg ctgaacgggg ggttcgtgca	4620
cacagccccg cttggggcga acgacctaca ccgaactgag atacctacag cgtgagctat	4680
gagaaagcgc cacgcttccc gaagggagaa aggccggacag gtatccggta agcggcaggg	4740
tcggAACAGG agagcgcacg agggagcttc cagggggaaa cgcctggat ctttatagtc	4800
ctgtcggtt tcgcccaccc tgacttgago gtcgatTTT gtgtatgtcg tcagggggc	4860
ggagccatag gaaaaacgccc agcaacgcgg ccttttacg gttccctggcc ttttgcggc	4920
cttttgcgtca catgtttttt cctgogttat cccctgatcc tggataac cgtattacgg	4980
cctttgatgtg acgtgatacc gctgcgcga gccaacgcac cgagcgcagc gagtcaatgt	5040
gggggggggg ggaagagcgc ctgatgcggt attttctct tacgcatctg tgccgttattt	5100
cacaccgcat atatggtgca ctctcgtac aatctgtct gatgcgcgat agttaagcca	5160
gtatacactc cgctatcgct acgtgactgg gtcgtggctg cgcccegaca cccgccaaca	5220
cccgctgacg cgccctgacg ggcttgcgtc ctccggcat ccgcctacag acaagctgt	5280
accgtctccg ggagctgcat gtgtcagagg tttcacggt catcacggaa acgcgcgagg	5340
cagctgcccgg aaagctcatc acgctggctg tgaagcgtt accatgttc tgccgttca	5400
tccgcgtcca gtcgttgag tttccaga agcgttaatg tctggcttct gataaaggcg	5460
cccatgttaa gggcggttt ttccgtttt gtcactgtat cctccgtta agggggattt	5520
ctgttcatgg gggtaatgtat accgatgaaa cgagagagga tgctcacgt acgggttact	5580
gatgtgaac atgcccgggtt actggaaacgt tgtgagggtt aacaactggc ggtatggatg	5640
cgccgggacc agagaaaaat cactcagggt caatgccgc gttcgatc tacagatgt	5700
gggtttccac agggtagcca gcagcatctt gcatgcgcata tccggaaat aatggtcag	5760
ggcgctgact tccgcgtttc cagactttac gaaacacggaa aaccgaagac cattcatgtt	5820
gttgctcagg tcgcagacgt tttgcagcag cagtcgttcc acgttcgcctc gctatcggt	5880
gattcattct gtcgttccatc aaggcaaccc cgccagccata gcccggcttca caacgcacgg	5940
agcacgtatca tgcgcaccccg tggggccgcg atgcggcgta taatggccctg cttctcgcc	6000
aaacgtttgg tggcgggacc agtgacgaa gcttgagcga gggcggtc gattccgaaat	6060
accgcacgcg acaggccgat catcgctcgcc ctccagccgaa agcggcttc gcccggaaat	6120
acccagacgc ctgcggcac ctgttccatcg agttgcgtat taaagaagac agtcatatagt	6180
gccccgcacgca tagtcatgcc ccgcgcggccac cggaaaggagc tgactgggtt gaaggcttc	6240
aaggccatcg gtcgagatcc cgggtgcctaa tgagtgcgtt aacttacatt aattgcgtt	6300
cgctcactgc cccgtttccca gtcggggaaac ctgtcgtgcc agctgcattat atgaatcgcc	6360
caacgcgcgg ggagaggcgg tttgcgtatt gggcgccagg gtggttttc ttttaccag	6420

-continued

tgagacgggc aacagctgat tgcccttcac cgccctggccc tgagagaggat gcagaacgcg	6480
gtccacgctg gtttgcggca gcaggcgaaa atcctgtttg atgggtggta acggcgggat	6540
ataacatgag ctgttctcg tatactcgta tcccactacc gagatatccg caccaacgcg	6600
cagccccggac tcggtaatgg cgccgcattgc gcccagcgcc atctgatcgt tggcaaccag	6660
catcgcgtg ggaacgtgc cctcattcag catttgcgtg gtttgtgaa aaccggacat	6720
ggcactccag tcgccttccc gttccgtat cggctgaatt tgattgcgag tgagatattt	6780
atgccagcca gccagacgca gacgcgcccga gacagaacctt aatggggcccg ctaacagcgc	6840
gatttgcgtgg tgacccaatg cgaccagatg ctccacgcgc agtcgcgtac cgtcttcatg	6900
ggagaaaaata atactgttga tgggtgtctg gtcagagaca tcaagaataa acgcccggAAC	6960
attagtgcag gcagcttcca cagcaatggc atcctggtca tccagcggat agttaatgtat	7020
cagccccactg acgcgtgtcg cgagaagatt gtgcaccgcg gtttacagg ctgcacgcgc	7080
gcttcgttct accatcgaca ccaccacgtt ggcacccagt tgatcggcgc gagatttaat	7140
cgccgcgaca atttgcgacg ggcgcgtgcag ggccagactg gaggtggccaa cgccaatcag	7200
caacgactgt ttgcccggca gttgttgtgc caagcgggttg ggaatgtaat tcagctccgc	7260
catcgcccgct tccactttt ccccggtttt cgccagaaacg tggctggcct ggttcaccac	7320
gcggggaaacg gtctgataag agacacccggc atactctgcg acatcgata acgttactgg	7380
tttcacatc accaccctga attgactctc ttccggggc tatcatgcata taccgcgaaa	7440
ggttttgcgc catcgatgg tgcgtggat ctgcacgttc tcccttatgc gactctgcgaa	7500
ttaggaagca gcccaggtagt aggttggggc cggtggacac cgccgcgcgca aggaatggtg	7560
catgcaagga gatggggccc aacagtcccc cggccacggg gcctgcaccat accccacgc	7620
cggaaacaacgc gtcgtatggc ccggaaatggc gagcccgatc ttccccatcg gtgtgtcg	7680
cgtatataaggc gcacggcaacc gcacccgtgg cgccgggtgat gcccggccacg atgcgtccgg	7740
cgttagaggat cgagatctcg atcccgcgaa at	7772

<210> SEQ ID NO 59
 <211> LENGTH: 2469
 <212> TYPE: DNA
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: synthetic
 <220> FEATURE:
 <221> NAME/KEY: CDS
 <222> LOCATION: (1)..(2469)
 <223> OTHER INFORMATION: Membrane protein P450(SbCYP79A1)-GFP-His8

<400> SEQUENCE: 59

atg gcg aca atg gag gta gag gcc gcg gcc acg gtg ctg gcc gcg	48
Met Ala Thr Met Glu Val Glu Ala Ala Ala Ala Thr Val Leu Ala Ala	
1 5 10 15	
ccc ttg ctg tcc tcc gcg atc ctc aaa ctg ctg cta ttc gta gtg	96
Pro Leu Leu Ser Ser Ser Ala Ile Leu Lys Leu Leu Leu Phe Val Val	
20 25 30	
acg ctc tcg tac ctg gcc cga ctg agg cca cgc aaa agc acc	144
Thr Leu Ser Tyr Leu Ala Arg Ala Leu Arg Arg Pro Arg Lys Ser Thr	
35 40 45	
acc aag tgc agc agc aca acg tgc gcc tcg ccc ccg gcc ggc gtt ggc	192
Thr Lys Cys Ser Ser Thr Thr Cys Ala Ser Pro Pro Ala Gly Val Gly	
50 55 60	

-continued

aac ccg ccg ctc cca ccg ggt ccc gtg ccg tgg ccc gtc gtc ggc aac Asn Pro Pro Leu Pro Pro Gly Pro Val Pro Trp Pro Val Val Gly Asn 65 70 75 80	240
ctg ccg gag atg ctg ctg aac aag ccg gca ttc ccg tgg atc cac cag Leu Pro Glu Met Leu Leu Asn Lys Pro Ala Phe Arg Trp Ile His Gln 85 90 95	288
atg atg ccg gag atg ggc acg gac atc gcc tgc gtc aag ctt ggc ggc Met Met Arg Glu Met Gly Thr Asp Ile Ala Cys Val Lys Leu Gly Gly 100 105 110	336
gtc cac gtc gtg tcc atc acc tgc ccg gag atc gcg ccg gag gtg ctc Val His Val Val Ser Ile Thr Cys Pro Glu Ile Ala Arg Glu Val Leu 115 120 125	384
cgg aag cag gac gcc aac ttc ata tcc ccg ctc acc ttc gcc tcc Arg Lys Gln Asp Ala Asn Phe Ile Ser Arg Pro Leu Thr Phe Ala Ser 130 135 140	432
gag acg ttc agc ggc ggg tac ccg aac gcc gtg ctc tgc ccc tac ggc Glu Thr Phe Ser Gly Gly Tyr Arg Asn Ala Val Leu Ser Pro Tyr Gly 145 150 155 160	480
gac cag tgg aag aag atg ccg ccg gtc ctc acc tcc gag atc atc tgc Asp Gln Trp Lys Lys Met Arg Arg Val Leu Thr Ser Glu Ile Ile Cys 165 170 175	528
ccg tcc ccg cac gcc tgg ctc cac gac aag ccg acc gac gag gcc gag Pro Ser Arg His Ala Trp Leu His Asp Lys Arg Thr Asp Glu Ala Asp 180 185 190	576
aac ctc acc ccg tac gtc tac aac ctc gcc acc aaa gcc gcc acc ggc Asn Leu Thr Arg Tyr Val Tyr Asn Leu Ala Thr Lys Ala Ala Thr Gly 195 200 205	624
gac gtc gcc gtc gac gtc agg cac gtc gct cgt cac tat tgc ggc aac Asp Val Ala Val Asp Val Arg His Val Ala Arg His Tyr Cys Gly Asn 210 215 220	672
gtt atc ccg ccg ctc atg ttc aac agg ccg tac ttc ggc gag ccc cag Val Ile Arg Arg Leu Met Phe Asn Arg Arg Tyr Phe Gly Glu Pro Gln 225 230 235 240	720
gct gac ggc ggt ccg ggg ccg atg gag gtg ctg cat atg gac gcc gtg Ala Asp Gly Gly Pro Gly Pro Met Glu Val Leu His Met Asp Ala Val 245 250 255	768
tcc acc tcc ctc ggc ctc ctc tac gcc ttc tgc gtc tcc gac tac ctc Phe Thr Ser Leu Gly Leu Leu Tyr Ala Phe Cys Val Ser Asp Tyr Leu 260 265 270	816
ccc tgg ctg ccg ggc ctc gac ctc gac ggc cac gag aag atc gtc aag Pro Trp Leu Arg Gly Leu Asp Leu Asp Gly His Glu Lys Ile Val Lys 275 280 285	864
gag gct aac gtg gcg gtg aac agg ctc cac gac acg gtc atc gac gag Glu Ala Asn Val Ala Val Asn Arg Leu His Asp Thr Val Ile Asp Asp 290 295 300	912
cgg tgg agg cag tgg aag agc ggc gag ccg cag gag atg gag gac ttc Arg Trp Arg Gln Trp Lys Ser Gly Glu Arg Gln Glu Met Glu Asp Phe 305 310 315 320	960
ctg gat gtg ctc atc act ctc aag gac gcc cag ggc aac ccg ctg ctg Leu Asp Val Leu Ile Thr Leu Lys Asp Ala Gln Gly Asn Pro Leu Leu 325 330 335	1008
acc atc gag gag gtc aaa gcg cag tca cag gac atc acg ttc ggc gcg Thr Ile Glu Glu Val Lys Ala Gln Ser Gln Asp Ile Thr Phe Ala Ala 340 345 350	1056
gtg gac aac ccg tcg aac gcc gtg gag tgg gcg ctg gca gag atg gtg Val Asp Asn Pro Ser Asn Ala Val Glu Trp Ala Leu Ala Glu Met Val 355 360 365	1104

-continued

aac aac ccg gag gtg atg gcg aag gcg atg gag gag ctg gac cgc gtc Asn Asn Pro Glu Val Met Ala Lys Ala Met Glu Glu Leu Asp Arg Val 370 375 380	1152
gtc gga ccg gag agg cta gtg cag gag tcg gac att ccg aag ctc aac Val Gly Arg Glu Arg Leu Val Gln Glu Ser Asp Ile Pro Lys Leu Asn 385 390 395 400	1200
tac gtg aag gcc tgc atc cgg gag gct ttc cgt ctg cac ccg gtg gcg Tyr Val Lys Ala Cys Ile Arg Glu Ala Phe Arg Leu His Pro Val Ala 405 410 415	1248
ccc ttc aac gtg ccc cac gtc gcg ctc gcc gac acc acc atc gcc ggc Pro Phe Asn Val Pro His Val Ala Leu Ala Asp Thr Thr Ile Ala Gly 420 425 430	1296
tac cgc gtt ccc aag ggc agc cac gtg atc ctg agc cgc acg ggg ctg Tyr Arg Val Pro Lys Gly Ser His Val Ile Leu Ser Arg Thr Gly Leu 435 440 445	1344
ggc cgc aac ccg cgc gtg tgg gac gag ccc ctg cgc ttc tac ccg gac Gly Arg Asn Pro Arg Val Trp Asp Glu Pro Leu Arg Phe Tyr Pro Asp 450 455 460	1392
cga cac ctc gcc acc gcc gcg tcc gac gtc gcg ctc acc gag aac gac Arg His Leu Ala Thr Ala Ala Ser Asp Val Ala Leu Thr Glu Asn Asp 465 470 475 480	1440
ctg cgg ttc atc tcc ttc agc acc ggc cgc cgc ggc tgc atc gcc gcg Leu Arg Phe Ile Ser Phe Ser Thr Gly Arg Arg Gly Cys Ile Ala Ala 485 490 495	1488
tcg ctc ggc acc gcc atg agc gtc atg ctc ttc gga agg ctc ctg cag Ser Leu Gly Thr Ala Met Ser Val Met Leu Phe Gly Arg Leu Leu Gln 500 505 510	1536
ggg ttc acc tgg agc aag ccc gcc ggg gtg gag gcc gtg gac ctc agc Gly Phe Thr Trp Ser Lys Pro Ala Gly Val Glu Ala Val Asp Leu Ser 515 520 525	1584
gag tcc aag agc gac acc ttc atg gcc acc ccg ctg gtg ctg cac gct Glu Ser Lys Ser Asp Thr Phe Met Ala Thr Pro Leu Val Leu His Ala 530 535 540	1632
gag ccc agg ctg ccg gcg cac ctc tac ccg tcc atc tcc atc gga tcg Glu Pro Arg Leu Pro Ala His Leu Tyr Pro Ser Ile Ser Ile Gly Ser 545 550 555 560	1680
cca gct gag gga gaa aac ctg tac ttc cag ggt caa ttc agc aaa gga Pro Ala Glu Gly Glu Asn Leu Tyr Phe Gln Gly Gln Phe Ser Lys Gly 565 570 575	1728
gaa gaa ctt ttc act gga gtt gtc cca att ctt gtt gaa tta gat ggt Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu Val Glu Leu Asp Gly 580 585 590	1776
gat gtt aat ggg cac aaa ttt tct gtc agt gga gag ggt gaa ggt gat Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp 595 600 605	1824
gct aca tac gga aaa ctc acc ctt aaa ttt att tgc act act gga aaa Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys 610 615 620	1872
cta cct gtt cca tgg cca aca ctt gtc act act ctg acc tat ggt gtt Leu Pro Val Pro Trp Pro Thr Leu Val Thr Leu Thr Tyr Gly Val 625 630 635 640	1920
caa tgc ttt tcc cgt tat ccg gat cac atg aaa ccg cat gac ttt ttc Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Arg His Asp Phe Phe 645 650 655	1968
aag agt gcc atg ccc gaa ggt tat gta cag gaa cgc act ata tct ttc Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Ser Phe 660 665 670	2016

-continued

aaa gat gac ggg aac tac aag acg cgt gct gaa gtc aag ttt gaa ggt Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu Gly 675 680 685	2064
gat acc ctt gtt aat cgt atc gag tta aag ggt att gat ttt aaa gaa Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu 690 695 700	2112
gat gga aac att ctc gga cac aaa cta gag tac aac tat aac tca cac Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr Asn Ser His 705 710 715 720	2160
aat gta tac atc acg gca gac aaa caa aag aat gga atc aaa gct aac Asn Val Tyr Ile Thr Ala Asp Lys Gln Lys Asn Gly Ile Lys Ala Asn 725 730 735	2208
ttc aaa att cgc cac aac att gaa gat ggt tcc gtt caa cta gca gac Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val Gln Leu Ala Asp 740 745 750	2256
cat tat caa caa aat act cca att ggc gat ggc cct gtc ctt tta cca His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro Val Leu Leu Pro 755 760 765	2304
gac aac cat tac ctg tcg aca caa tct gcc ctt tcg aaa gat ccc aac Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn 770 775 780	2352
gaa aag cgt gac cac atg gtc ctt ctt gag ttt gta act gct gct ggg Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala Gly 785 790 795 800	2400
att aca cat ggc atg gat gag ctc tac aaa aag ctt gcg gcc cat cat Ile Thr His Gly Met Asp Glu Leu Tyr Lys Lys Leu Ala Ala His His 805 810 815	2448
cat cac cac cac cac tga His His His His His 820	2469

<210> SEQ ID NO 60

<211> LENGTH: 822

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 60

Met Ala Thr Met Glu Val Glu Ala Ala Ala Ala Thr Val Leu Ala Ala 1 5 10 15
--

Pro Leu Leu Ser Ser Ser Ala Ile Leu Lys Leu Leu Leu Phe Val Val 20 25 30

Thr Leu Ser Tyr Leu Ala Arg Ala Leu Arg Arg Pro Arg Lys Ser Thr 35 40 45

Thr Lys Cys Ser Ser Thr Thr Cys Ala Ser Pro Pro Ala Gly Val Gly 50 55 60

Asn Pro Pro Leu Pro Pro Gly Pro Val Pro Trp Pro Val Val Gly Asn 65 70 75 80
--

Leu Pro Glu Met Leu Leu Asn Lys Pro Ala Phe Arg Trp Ile His Gln 85 90 95

Met Met Arg Glu Met Gly Thr Asp Ile Ala Cys Val Lys Leu Gly Gly 100 105 110
--

Val His Val Val Ser Ile Thr Cys Pro Glu Ile Ala Arg Glu Val Leu 115 120 125
--

Arg Lys Gln Asp Ala Asn Phe Ile Ser Arg Pro Leu Thr Phe Ala Ser 130 135 140
--

-continued

Glu Thr Phe Ser Gly Gly Tyr Arg Asn Ala Val Leu Ser Pro Tyr Gly
 145 150 155 160
 Asp Gln Trp Lys Lys Met Arg Arg Val Leu Thr Ser Glu Ile Ile Cys
 165 170 175
 Pro Ser Arg His Ala Trp Leu His Asp Lys Arg Thr Asp Glu Ala Asp
 180 185 190
 Asn Leu Thr Arg Tyr Val Tyr Asn Leu Ala Thr Lys Ala Ala Thr Gly
 195 200 205
 Asp Val Ala Val Asp Val Arg His Val Ala Arg His Tyr Cys Gly Asn
 210 215 220
 Val Ile Arg Arg Leu Met Phe Asn Arg Arg Tyr Phe Gly Glu Pro Gln
 225 230 235 240
 Ala Asp Gly Gly Pro Gly Pro Met Glu Val Leu His Met Asp Ala Val
 245 250 255
 Phe Thr Ser Leu Gly Leu Leu Tyr Ala Phe Cys Val Ser Asp Tyr Leu
 260 265 270
 Pro Trp Leu Arg Gly Leu Asp Leu Asp Gly His Glu Lys Ile Val Lys
 275 280 285
 Glu Ala Asn Val Ala Val Asn Arg Leu His Asp Thr Val Ile Asp Asp
 290 295 300
 Arg Trp Arg Gln Trp Lys Ser Gly Glu Arg Gln Glu Met Glu Asp Phe
 305 310 315 320
 Leu Asp Val Leu Ile Thr Leu Lys Asp Ala Gln Gly Asn Pro Leu Leu
 325 330 335
 Thr Ile Glu Glu Val Lys Ala Gln Ser Gln Asp Ile Thr Phe Ala Ala
 340 345 350
 Val Asp Asn Pro Ser Asn Ala Val Glu Trp Ala Leu Ala Glu Met Val
 355 360 365
 Asn Asn Pro Glu Val Met Ala Lys Ala Met Glu Glu Leu Asp Arg Val
 370 375 380
 Val Gly Arg Glu Arg Leu Val Gln Glu Ser Asp Ile Pro Lys Leu Asn
 385 390 395 400
 Tyr Val Lys Ala Cys Ile Arg Glu Ala Phe Arg Leu His Pro Val Ala
 405 410 415
 Pro Phe Asn Val Pro His Val Ala Leu Ala Asp Thr Thr Ile Ala Gly
 420 425 430
 Tyr Arg Val Pro Lys Gly Ser His Val Ile Leu Ser Arg Thr Gly Leu
 435 440 445
 Gly Arg Asn Pro Arg Val Trp Asp Glu Pro Leu Arg Phe Tyr Pro Asp
 450 455 460
 Arg His Leu Ala Thr Ala Ala Ser Asp Val Ala Leu Thr Glu Asn Asp
 465 470 475 480
 Leu Arg Phe Ile Ser Phe Ser Thr Gly Arg Arg Gly Cys Ile Ala Ala
 485 490 495
 Ser Leu Gly Thr Ala Met Ser Val Met Leu Phe Gly Arg Leu Leu Gln
 500 505 510
 Gly Phe Thr Trp Ser Lys Pro Ala Gly Val Glu Ala Val Asp Leu Ser
 515 520 525
 Glu Ser Lys Ser Asp Thr Phe Met Ala Thr Pro Leu Val Leu His Ala
 530 535 540

-continued

Glu	Pro	Arg	Leu	Pro	Ala	His	Leu	Tyr	Pro	Ser	Ile	Ser	Ile	Gly	Ser
545															560
Pro	Ala	Glu	Gly	Glu	Asn	Leu	Tyr	Phe	Gln	Gly	Gln	Phe	Ser	Lys	Gly
															575
Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu	Val	Glu	Leu	Asp	Gly
															590
Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly	Glu	Gly	Glu	Gly	Asp
															605
Ala	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	Cys	Thr	Thr	Gly	Lys
															620
Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr	Leu	Thr	Tyr	Gly	Val
															640
Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys	Arg	His	Asp	Phe	Phe
															655
Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Glu	Arg	Thr	Ile	Ser	Phe
															670
Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu	Val	Lys	Phe	Glu	Gly
															685
Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly	Ile	Asp	Phe	Lys	Glu
															700
Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr	Asn	Tyr	Asn	Ser	His
															720
Asn	Val	Tyr	Ile	Thr	Ala	Asp	Lys	Gln	Lys	Asn	Gly	Ile	Lys	Ala	Asn
															735
Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser	Val	Gln	Leu	Ala	Asp
															750
His	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Gly	Asp	Gly	Pro	Val	Leu	Leu	Pro
															765
Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu	Ser	Lys	Asp	Pro	Asn
															780
Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe	Val	Thr	Ala	Ala	Gly
															800
Ile	Thr	His	Gly	Met	Asp	Glu	Leu	Tyr	Lys	Lys	Leu	Ala	Ala	His	His
															815
His	His	His	His	His	His										
															820

```

<210> SEQ ID NO 61
<211> LENGTH: 6878
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)..(6878)
<223> OTHER INFORMATION: Plasmid pHtpX

<400> SEQUENCE: 61

taatacact cactataggg gaatttgtag cggtataacaa ttccccctcta gaaataattt
tgtttaacctt taagaaggag accagtttat gatgcgaatc gcgctttcc tgctaacgaa 60
cctggccgta atggtcgttt tcgggctgtt actgagcctg acagggatac agtcgagcag 120
cgttcagggg ctgatgatca tggccttgcgtt gttcggtttt ggtggttcct tcgtttcgct 180

```

-continued

tctgatgtcc	aatggatgg	cattacgata	tgttggcgaa	gaagtgtatcg	agcaaccgcg	300
taacgaaagg	gaacgttggc	tggtaataac	tgttagcaacc	caggctcgat	aggcgaaaat	360
cgctatgcgc	caagtggcta	tctaccatgc	gccggacatac	aacgcgttttgc	caaccgggt	420
gcgcgcgtat	gcctctctgg	tttgtgtcag	caccggtttg	ctgcagaaca	tgagccccgaa	480
tgaagccgag	gcggtaatttgc	ctcacaaat	cagccacatac	gcataatgggt	atatggtcac	540
catgaacgtgc	attcaggcg	tttgtgaacac	cttcgttatac	tttatttccc	gtattctggc	600
gcagcttgc	gcgggttttgc	ttggcgaaaa	tcgtgatgaa	ggtaagaaaa	gcaacggcaa	660
cccgctgatc	tactttgcgg	ttgcaacggt	tctggaaactg	gtgtttggta	ttctggcgag	720
cattatcacc	atgtggttct	cgcgcatcg	tgaattccat	gtgtatgcgg	gttccggaaaa	780
actgggttgg	cgcgagaaaa	tgattgccgc	gctgcagcg	ctgaaaacca	gctatgaacc	840
gcaagaagca	accagcatga	tggctctctg	cattaacggt	aagtgcataat	cgctcagtga	900
gttgttcatg	acccacccgc	cgctggataa	acgaattgaa	gtctcgatgt	cggtgtaaaa	960
cctgaagtcg	gtacctggat	ccgaaaacct	gtacttccag	ggtaattca	gcaaaaggaga	1020
agaacttttc	actggagtttgc	tcccaatttgc	tgttgaatttgc	gtgtgtatgc	ttaatgggca	1080
caaattttct	gtcagttggag	agggtgaagg	tgtatgtaca	tacggaaaac	tcacccttaa	1140
atttatgtgc	actactggaa	aactacctgt	tccatggcca	acacttgcata	ctactctgac	1200
ctatgggttgc	aatgtttttgc	cccggttatacc	ggatcacatg	aaacggcatg	acttttcaa	1260
gagtgccatg	cccgaaagggtt	atgtacagga	acgcactata	tcttcaaaag	atgacgggaa	1320
ctacaagacg	cgtgctgaag	tcaagtttgc	aggtgatacc	cttgcataatc	gtatcgagtt	1380
aaagggtatt	gatttaaag	aagatggaaa	cattctcgaa	cacaaacttag	agtacaacta	1440
taactcacac	aatgtataca	tcacggcaga	caaacaaaa	aatggaaatca	aagctaactt	1500
caaaatttgc	cacaacatttgc	aagatgggttgc	cgttcaacta	gcagaccatt	atcaacaaaa	1560
tactccaatttgc	ggcgatggcc	ctgttccccc	accagacaac	cattacctgt	cgacacaatc	1620
tgcctttcg	aaagatccca	acgaaaageg	tgaccacatg	gtccttctgc	agtttgcac	1680
tgctgctggg	attacacatg	gcatggatgc	gctctacaaa	aagcttgcgg	cccatcatca	1740
tcaccaccac	caccacgtgc	atccggctgc	taacaaagcc	cgaaaggaaag	ctgagttggc	1800
tgctgccacc	gctgatgcata	aactacgtata	accccttggg	gcctctaaac	gggtcttgag	1860
gggtttttgc	ctgaaaggag	gaactatatac	cgatttggcg	aatgggacgc	gccctgtatc	1920
ggcgcatata	gcgcggcgaa	tgtgggtgg	acgcgcagcg	tgaccgctac	acttgcacgc	1980
gcccttagcgc	ccgctcccttgc	cgcttccccc	ccttccttgc	tcgcccacgtt	cgccggcttgc	2040
ccccgtcaatgc	ctctaaatgc	ggggctcccttgc	ttagggttgc	gatttgtatgc	tttacggcac	2100
ctcgacccca	aaaaacttgc	ttagggttgc	gttccacgttgc	gtggggccatc	gccctgtatc	2160
acggtttttc	gccctttgcac	gttggatgc	acgttcttgc	atagtggact	cttgcataatc	2220
actggaaacaa	cactcaaccc	tatctcggtc	tattttttgc	atttataagg	gattttgcgg	2280
atttcggttgc	atgggttataa	aatgtatgtgc	atttacaaa	atttacgtgc	gaattttac	2340
aaaaatattaa	cgtttacaat	ttcaggtggc	acttttcggg	gaaatgtgcg	cgaaacccct	2400
atttgtttat	ttttctaaat	acattcaat	atgtatccgc	tcatgtatgc	atttgcataatc	2460
aaactcatcg	agcatcaaat	gaaactgcac	tttattcata	tcaggattat	caataccata	2520

-continued

ttttgaaaa agccgttct gtaatgaagg agaaaactca ccgaggcagt tccataggat	2580
ggcaagatcc tggtatcggt ctgcgattcc gactcgtcca acatcaatac aacctattaa	2640
tttcccctcg tcaaaaataa ggatatcaag tgagaaaatca ccatgagtga cgactgaatc	2700
cggtgagaat ggcaaaaagg ttatgcatttc ttccagact tgttcaacag gccagccatt	2760
acgctcgatca tcaaaaatcac tcgcatcaac caaacgtta ttcatcgat attgcgcctg	2820
agcgagacga aatacgcgtat cgctgttaaa aggacaatta caaacaggaa tcgaatgcaa	2880
ccggcgcagg aacactgcca ggcgcataac aatattttca cctgaatcag gatatttttc	2940
taataacctgg aatgcgtttt tccccggat cgcagtggtg agtaaccatg catcatcagg	3000
agtacggata aaatgcgtga tggtcgaaag aggataaat tccgtcagcc agtttagtct	3060
gaccatctca tctgttaacat cattggcaac gctaccttg ccatgtttca gaaacaactc	3120
tggcgcatacg ggcttccat acaatcgata gattgtcgca cctgattgcc cgacattatc	3180
gcgagccat ttataccat ataaatcago atccatgttg gaatttaatc gggcctaga	3240
gcaagacgtt tcccggtgaa tatggctcat aacacccctt gtattactgt ttatgtaa	3300
agacagttt attgttcatg accaaaatcc cttaacgtga gtttgcgtc cactgagcgt	3360
cagacccctg agaaaaagtc aaaggatctt cttgagatcc ttttttctg cgctaatct	3420
gctgcttgca aacaaaaaaaaa ccacgcgtac cagcgggtgg ttttttgccg gatcaagagc	3480
taccaactct tttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc	3540
ttcttagtgta gccgtagtttta ggccaccact tcaagaactc tgttagcaccc cctacatacc	3600
tcgctctgt aatctgtta ccagtggctg ctgccagtg cgataagtgc tgcttaccg	3660
ggttggactc aagacgatag ttacggata aggcgcagcg gtcgggctga acggggggtt	3720
cgtgcacaca gcccgatctg gagegaacga cctacacccg actgagatac ctacgcgtg	3780
agctatgaga aagcgcacg cttcccgaaag ggagaaaaggc ggacaggatcc cccgtaaagcg	3840
gcagggtcg aacaggagag cgcacgagg agttccagg gggaaacgcc tggtatctt	3900
atagtcctgt cgggttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgatcg	3960
ggggggggag cctatggaaa aacgcgcacg acgcggccctt tttacgggtc ctggcccttt	4020
gctggccctt tgctcacatg ttcttcctg cgttateccc tgattctgtg gataaccgtt	4080
ttaccgcctt tgagttagct gataccgcctc gccgcageccg aacgaccgag cgacgcgt	4140
cagtgagcga ggaagcggaa gagcgcctga tgccgtattt tctccttacg catctgtgcg	4200
gtatccaca ccgcataatg ggtgcactct cagtacaatc tgctctgtatccgcataat	4260
aagcccgat atactcccgat atcgatcgat gactgggtca tggctgcgcg ccgcaccccg	4320
ccaaaccccg ctgacgcgcctc ctgacgggct tgcgtcgatcc cggcatccgc ttacagacaa	4380
gctgtgaccc tctccggag ctgcgtgtt cagaggatcc caccgtcatc accgaaacgc	4440
gcggaggcgc tgccgtaaag ctcacatcgat tggtcgatggaa ggcattcaca gatgtctgc	4500
tgttcatccg cgtccgcgtc gtttagtttcc tccagaagcg ttaatgtctg gcttcgtata	4560
aagcggggca tgtaaaggc ggtttttcc tggtcgatggaa ctgatgcctc cgtgtaaagg	4620
ggatccatgt tcatgggggtt aatgataccg atgaaacgcg agaggatgtt cacgataccg	4680
gttactgtatg atgaacatgc cccgttactg gaaacgttgcg agggtaaaca actggcgat	4740
tggatgeggc gggaccagag aaaaatcaact cagggtcaat gccagegctt cgttaataca	4800

-continued

gatgttagtg ttccacaggg tagccagcg catcctgcga tgcagatccg gaacataatg	4860
gtgcagggcg ctgacttccg cgtttccaga ctttacgaaa cacggaaacc gaagaccatt	4920
catgttgttgc ctcaggatcg agacgttttgc cagcagcagt cgcttcacgt tcgctcgctg	4980
atcggtgatt cattctgcta accagtaagg caaccccgcc agcctagccg ggtcctcaac	5040
gacaggagca cgatcatgcg caccctgtgg gccgcatgc cggcgataat ggcctgttc	5100
tgcggaaac gtttggtggc gggaccagtgc aegaaggctt gagcggaggc gtgcaagatt	5160
ccgaataccg caagcgacag gcccgtatgc gtcgcgtcc agcgaaagcg gtcctcgccg	5220
aaaatgaccc agagcgctgc cggcacctgt cctacgagtt gcatgataaa gaagacagtc	5280
ataagtgcgg cgacgatagt catgccccgc gcccacccggaa aggagctgac tgggttgaag	5340
gctctcaagg gcateggctcg agatcccgtt gcctaatgag tgagctaact tacattaatt	5400
cggttgcgtc cactgccccgc tttccagtcg gggaaacctgt cgtgccagct gcattaatga	5460
atcggccaac ggcgggggag aggccgtttt cgtattgggc gccagggtgg tttttttttt	5520
caccagttag acgggcaaca gctgattgc cttcacccgc tggccctgag agagttgcag	5580
caagcggtcc acgctgggtt gccccagcag gcgaaaatcc tgggtatgg tggtaacgg	5640
cgggatataa catgagctgt cttegggtatc gtcgtatccc actacccgaga tatccgcacc	5700
aacgcgcagc cgggactcgg taatggcgeg cattgcgecc agcgccatct gatcggttgc	5760
aaccagcatc gcagtggaa cgatgcccattt attcagcatt tgcattgggtt gttaaaacc	5820
ggacatggca ctccagtcgc ctteccgttc cgtatccggc tgaatttgat tgcgagttag	5880
atatttatgc cagccagcca gacgcagacg cggcagaca gaacttaatg ggccgcataa	5940
cagcgcgatt tgctggtgac ccaatgcgcac cagatgtcc acgcccagtc gcgtaccgtc	6000
ttcatgggag aaaataatac tggatgggg tgtctggta gagacatcaa gaaataacgc	6060
cggAACATTA GTGCAGGGCAG CTTCCACAGO ATGGCATCC TGGTCATCCA GCGGATAGTT	6120
aatgatcagc ccactgacgc gttgcgcag aagattgtgc accgcgcgtt tacaggcttc	6180
gacgcccgtt cgttctacca tcgacaccac cagctggca cccagttgtat cggcgcgaga	6240
tttaategcg cgcacaattt gcgcacggcgc gtgcaggggc agactggagg tggcaacgcc	6300
aatcagcaac gactgtttgc ccgcacgttg ttgtgccacg cgggtggaa tggtaattcag	6360
ctccgcaccatc cccgcgttcca cttttcccg cgttttgcga gaaacgtggc tggcctgggtt	6420
caccacgcgg gaaacggctt gataagagac accggcatac tctgcacat cgtataacgt	6480
tactggtttc acattcacca ccctgaatttgc actctttccg gggcgatc atgcataacc	6540
gcgaaagggtt ttgcgcattt cgatgggtgc cgggatctcg acgtctccc ttatgcact	6600
cctgcattag gaagcagccc agtagtaggt tgaggccgtt gaggcaccgc gccgcaagga	6660
atgggtgcatg caaggagatg ggcggccaca gttcccccggc cacggggccct gccaccatac	6720
ccacgcggaa acaagcgctc atgagcccga agtggcgagc ccgatcttcc ccatcggtga	6780
tgtcgccatg ataggcgcca gcaaccgcac ctgtggcgcc ggtgatgcgc gccacgatgc	6840
gtccggcgta gaggatcgag atctcgatcc cgcgaaat	6878

<210> SEQ ID NO 62

<211> LENGTH: 1671

<212> TYPE: DNA

<213> ORGANISM: Artificial sequence

<220> FEATURE:

-continued

```

<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)..(1671)
<223> OTHER INFORMATION: Membrane protein HtpX-GFP-His8

<400> SEQUENCE: 62

atg atg cga atc gcg ctc ttc ctg cta acg aac ctg gcc gta atg gtc      48
Met Met Arg Ile Ala Leu Phe Leu Leu Thr Asn Leu Ala Val Met Val
1           5          10          15

gtt ttc ggg ctg gta ctg agc ctg aca ggg ata cag tcg agc agc gtt      96
Val Phe Gly Leu Val Leu Ser Leu Thr Gly Ile Gln Ser Ser Ser Val
20          25          30

cag ggg ctg atg atc atg gcc ttg ctg ttc ggt ttt ggt ggt tcc ttc      144
Gln Gly Leu Met Ile Met Ala Leu Leu Phe Gly Phe Gly Ser Phe
35          40          45

gtt tcg ctt ctg atg tcc aaa tgg atg gca tta cga tct gtt ggc ggg      192
Val Ser Leu Leu Met Ser Lys Trp Met Ala Leu Arg Ser Val Gly Gly
50          55          60

gaa gtg atc gag caa ccg cgt aac aac agg gaa cgt tgg ctg gtc aat      240
Glu Val Ile Glu Gln Pro Arg Asn Glu Arg Glu Arg Trp Leu Val Asn
65          70          75          80

act gta gca acc cag gct cgt cag gcg ggg atc gct atg ccg caa gtc      288
Thr Val Ala Thr Gln Ala Arg Gln Ala Gly Ile Ala Met Pro Gln Val
85          90          95

gct atc tac cat gcg ccg gac atc aac gct ttt gca acc ggt gcg cgc      336
Ala Ile Tyr His Ala Pro Asp Ile Asn Ala Phe Ala Thr Gly Ala Arg
100         105         110

cgt gat gcc tct ctg gtt gtc agc acc ggt ttg ctg cag aac atg      384
Arg Asp Ala Ser Leu Val Ala Val Ser Thr Gly Leu Leu Gln Asn Met
115         120         125

agc ccg gat gaa gcc gag gcg gta att gct cac gaa atc agc cac atc      432
Ser Pro Asp Glu Ala Glu Ala Val Ile Ala His Glu Ile Ser His Ile
130         135         140

gcc aat ggt gat atg gtc acc atg acg ctg att cag ggc gtg gtc aac      480
Ala Asn Gly Asp Met Val Thr Met Thr Leu Ile Gln Gly Val Val Asn
145         150         155         160

acc ttc gtt atc ttt att tcc cgt att ctg gcg cag ctt gcc gcg ggt      528
Thr Phe Val Ile Phe Ile Ser Arg Ile Leu Ala Gln Leu Ala Ala Gly
165         170         175

ttt atg ggc gga aat cgt gat gaa ggt gaa gag agc aac ggc aac ccg      576
Phe Met Gly Gly Asn Arg Asp Glu Gly Glu Ser Asn Gly Asn Pro
180         185         190

ctg atc tac ttt gcg gtt gca acg gtt ctg gaa ctg gtc ttt ggt att      624
Leu Ile Tyr Phe Ala Val Ala Thr Val Leu Glu Leu Val Phe Gly Ile
195         200         205

ctg gcg agc att atc acc atg tgg ttc tcg cgt cat cgt gaa ttc cat      672
Leu Ala Ser Ile Ile Thr Met Trp Phe Ser Arg His Arg Glu Phe His
210         215         220

gct gat gcc ggt tcg gca aaa ctg gtt ggt cgc gag aaa atg att gcc      720
Ala Asp Ala Gly Ser Ala Lys Leu Val Gly Arg Glu Lys Met Ile Ala
225         230         235         240

gcg ctg cag ccg ctg aaa acc agc tat gaa ccg caa gaa gca acc agc      768
Ala Leu Gln Arg Leu Lys Thr Ser Tyr Glu Pro Gln Glu Ala Thr Ser
245         250         255

atg atg gct ctc tgc att aac ggt aag tcg aaa tcg ctc agt gag ttg      816
Met Met Ala Leu Cys Ile Asn Gly Lys Ser Lys Ser Leu Ser Glu Leu
260         265         270

```

-continued

ttc atg acc cac ccg ccg ctg gat aaa cga att gaa gct ctg cgt acg Phe Met Thr His Pro Pro Leu Asp Lys Arg Ile Glu Ala Leu Arg Thr 275 280 285	864
ggt gaa tac ctg aag tcg gta cct gga tcc gaa aac ctg tac ttc cag Gly Glu Tyr Leu Lys Ser Val Pro Gly Ser Glu Asn Leu Tyr Phe Gln 290 295 300	912
ggt caa ttc agc aaa gga gaa gaa ctt ttc act gga gtt gtc cca att Gly Gln Phe Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile 305 310 315 320	960
ctt gtt gaa tta gat ggt gat gtt aat ggg cac aaa ttt tct gtc agt Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser 325 330 335	1008
gga gag ggt gaa ggt gat gct aca tac gga aaa ctc acc ctt aaa ttt Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe 340 345 350	1056
att tgc act act gga aaa cta cct gtt cca tgg cca aca ctt gtc act Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr 355 360 365	1104
act ctg acc tat ggt gtt caa tgc ttt tcc cgt tat ccg gat cac atg Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met 370 375 380	1152
aaa cgg cat gac ttt ttc aag agt gcc atg ccc gaa ggt tat gta cag Lys Arg His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln 385 390 395 400	1200
gaa cgc act ata tct ttc aaa gat gac ggg aac tac aag acg cgt gct Glu Arg Thr Ile Ser Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala 405 410 415	1248
gaa gtc aag ttt gaa ggt gat acc ctt gtt aat cgt atc gag tta aag Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys 420 425 430	1296
ggt att gat ttt aaa gaa gat gga aac att ctc gga cac aaa cta gag Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu 435 440 445	1344
tac aac tat aac tca cac aat gta tac atc acg gca gac aaa caa aag Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Thr Ala Asp Lys Gln Lys 450 455 460	1392
aat gga atc aaa gct aac ttc aaa att cgc cac aac att gaa gat ggt Asn Gly Ile Lys Ala Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly 465 470 475 480	1440
tcc gtt caa cta gca gac cat tat caa caa aat act cca att ggc gat Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp 485 490 495	1488
ggc cct gtc ctt tta cca gac aac cat tac ctg tcg aca caa tct gcc Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala 500 505 510	1536
ctt tcg aaa gat ccc aac gaa aag cgt gac cac atg gtc ctt ctt gag Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu 515 520 525	1584
ttt gta act gct gct ggg att aca cat ggc atg gat gag ctc tac aaa Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys 530 535 540	1632
aag ctt gcg gcc cat cat cac cac cac cac cac tga Lys Leu Ala Ala His His His His His His His 545 550 555	1671

<210> SEQ ID NO 63

<211> LENGTH: 556

<212> TYPE: PRT

-continued

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 63

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

Val Phe Gly Leu Val Leu Ser Leu Thr Gly Ile Gln Ser Ser Ser Val
20 25 30

Gln Gly Leu Met Ile Met Ala Leu Leu Phe Gly Phe Gly Gly Ser Phe
35 40 45

Val Ser Leu Leu Met Ser Lys Trp Met Ala Leu Arg Ser Val Gly Gly
50 55 60

Glu Val Ile Glu Gln Pro Arg Asn Glu Arg Glu Arg Trp Leu Val Asn
65 70 75 80

Thr Val Ala Thr Gln Ala Arg Gln Ala Gly Ile Ala Met Pro Gln Val
85 90 95

Ala Ile Tyr His Ala Pro Asp Ile Asn Ala Phe Ala Thr Gly Ala Arg
100 105 110

Arg Asp Ala Ser Leu Val Ala Val Ser Thr Gly Leu Leu Gln Asn Met
115 120 125

Ser Pro Asp Glu Ala Glu Ala Val Ile Ala His Glu Ile Ser His Ile
130 135 140

Ala Asn Gly Asp Met Val Thr Met Thr Leu Ile Gln Gly Val Val Asn
145 150 155 160

Thr Phe Val Ile Phe Ile Ser Arg Ile Leu Ala Gln Leu Ala Ala Gly
165 170 175

Phe Met Gly Gly Asn Arg Asp Glu Gly Glu Glu Ser Asn Gly Asn Pro
180 185 190

Leu Ile Tyr Phe Ala Val Ala Thr Val Leu Glu Leu Val Phe Gly Ile
195 200 205

Leu Ala Ser Ile Ile Thr Met Trp Phe Ser Arg His Arg Glu Phe His
210 215 220

Ala Asp Ala Gly Ser Ala Lys Leu Val Gly Arg Glu Lys Met Ile Ala
225 230 235 240

Ala Leu Gln Arg Leu Lys Thr Ser Tyr Glu Pro Gln Glu Ala Thr Ser
245 250 255

Met Met Ala Leu Cys Ile Asn Gly Lys Ser Lys Ser Leu Ser Glu Leu
260 265 270

Phe Met Thr His Pro Pro Leu Asp Lys Arg Ile Glu Ala Leu Arg Thr
275 280 285

Gly Glu Tyr Leu Lys Ser Val Pro Gly Ser Glu Asn Leu Tyr Phe Gln
290 295 300

Gly Gln Phe Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile
305 310 315 320

Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser
325 330 335

Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe
340 345 350

Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr
355 360 365

Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met

-continued

370	375	380
Lys Arg His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln		
385	390	395
Glu Arg Thr Ile Ser Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala		
405	410	415
Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys		
420	425	430
Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu		
435	440	445
Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Thr Ala Asp Lys Gln Lys		
450	455	460
Asn Gly Ile Lys Ala Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly		
465	470	475
Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp		
485	490	495
Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala		
500	505	510
Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu		
515	520	525
Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys		
530	535	540
Lys Leu Ala Ala His		
545	550	555

```

<210> SEQ ID NO 64
<211> LENGTH: 7658
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)..(7658)
<223> OTHER INFORMATION: Plasmid pYqiK

<400> SEQUENCE: 64

taatacgaact cactataggg gaattgttag cggtataacaa ttccccctcta gaaataat 60
tgtttaacctt taagaaggag accagtttat ggatgatatt gttaattctg tgccctcctg 120
gatgtttacc gcaattatttgc cctgtatgcatt tctgtttattt attggaattt ttttcgc 180
gctctatcgat cgcgcattcg cagagcaagc ttttgttcgt actgggttag gtgggcaaaa 240
agtggtaatgc agccgggtggcg caatcgatgtat gccgatctttt catgaaataaa tccccatcaa 300
tatgaataact ctgaagctgg aagtcaagccg ctcaaccattt gatagccctga ttacgaaaga 360
tcgtatgcgcgt tcgtatgttag tagtgcgtttt ctttgtgcgg gtaaaaacctt cagtagaagg 420
gattgccacc gctgccaga cgctggggca acgcacccctg tcgcctgaag acttacgtat 480
gttgggtgaa gataaatttgc tgcgtgcctt ccgtgcaca gctgcgc 540
ttagttacag gatacccgcg agaactttgtt gcagggggttgc caaaatacag tggcagaaga 600
cctgtcgaaa aacggtctgg aactggagag cgtttcaattt accaacttta accagaccc 660
gaaaacacat ttcaatccga acaatgcctt tgacgccc ggtttaacca aactgactca 720
ggaaacacatcgcc gccaacgtaa cgaagttgaa caggatgttag aagttgcgg 780
gcgtgaaaaa aatcgcgatcg cgtatcgca caagctggag attgaacacgc aagaagcg 840

```

-continued

tatgacgott	gagcaggagc	agcagggtaa	aacccgtact	gccgaacaga	atgcacgtat	900
tgcggtttt	gaagctgaac	gtcgtcgtga	agcagagcag	acacgaattc	tggctgaacg	960
acagattcag	gaaacagaaa	tcgaccgcga	acaggccgtc	cgctcaagaa	agggtgaagc	1020
tgaacgtgaa	gttcgcatta	aagagatcga	acagcagcag	gtcacccgaaa	tcgctaacc	1080
gacgaaatcg	atcgctattt	ccgccaatac	ggaacaacag	tcccaggcag	aagcgcgtgc	1140
taatctcgca	cttgcagaag	cggtaagcgc	ccaacaaaac	gtagaaacca	ctcgccagac	1200
tgccgaagcc	gatcgtgcta	aacaagttgc	cctaatcgct	gccgcgcagg	atgcagaaac	1260
caaagcgggt	gaaactgaccg	tgccggcga	agcagaaaaaa	gaagccgcag	agatgcaggc	1320
ggcggctatc	gtttagtttag	ccgaagctac	acgtaaaaag	ggtctggcgg	aagcagaagc	1380
acaacgtgcg	ctgaacgtat	ctatcaacgt	actttctgtat	gaacaaacca	gccttaaatt	1440
caaactggcc	ttgttgcagg	cgtgcctgc	ggtaatagag	aaatccgttg	agccgatgaa	1500
gtcaatcgac	ggtatcaaga	ttattcagg	cgatggctcg	aatcgtggcg	gcccgtcggg	1560
tgatgcaaac	acgggtaatg	tggggggcgg	aaacctggcg	gaacaagcat	tatcagccgc	1620
tctctcttac	cgcacacagg	caccgctgtat	tgactcccttgc	ctcaatgaaa	ttggcgtttc	1680
aggcggctca	ctggcggcat	tgacttcacc	cttaacctca	acaactcccg	tcgaagaaaa	1740
agcagagtcg	gtacctggat	ccgaaaaacct	gtacttccag	ggtcaattca	gcaaggaga	1800
agaacttttc	actggagttt	tcccaattct	tgttgaattt	gatggtgatg	ttaatggca	1860
caaattttct	gtcagtggag	agggtgaagg	tgatgctaca	tacggaaaac	tcacccttaa	1920
atttatttgc	actactggaa	aactacctgt	tccatggcca	acacttgtca	ctactctgac	1980
ctatgggttt	caatgctttt	cccggttatcc	ggatcacatg	aaacggcatg	acttttcaa	2040
gagtgccatg	cccgaaagggtt	atgtacagga	acgcactata	tcttcaaag	atgacgggaa	2100
ctacaagacg	cgtgctgaag	tcaagtttga	aggtgataacc	tttgttaatc	gtatcgagtt	2160
aaagggtatt	gattttaaag	aagatggaaa	cattctcgga	cacaaaactag	agtacaacta	2220
taactcacac	aatgtataaca	tcacggcaga	caaacaaaag	aatggaatca	aagctaactt	2280
caaaattcgc	cacaacattt	aagatggttc	cgttcaacta	gcagaccatt	atcaacaaaa	2340
tactccaatt	ggcgatggcc	ctgtcccttt	accagacaac	cattacctgt	cgacacaatc	2400
tgccctttcg	aaagatccca	acgaaaaacgc	tgaccacatg	gtcccttcttg	agttttaac	2460
tgctgctggg	attacacatg	gcatggatga	gctctacaaa	aagcttgcgg	ccccatcatca	2520
tcaccaccac	caccactgag	atccggctgc	taacaaagcc	cgaaaggaaag	ctgagttggc	2580
tgctgccacc	gctgagcaat	aactagcata	acccttggg	gcctctaaac	gggtctttag	2640
gggttttttg	ctgaaaggag	gaactatatc	cgatttggcg	aatgggacgc	gcctgttagc	2700
ggcgcattaa	gogcggcggg	tgtggtgggt	acgogcagcg	tgaccgctac	acttgcacg	2760
gccttagcgc	ccgctccctt	cgcttcttc	ccttccttc	tcgcccacgtt	cgccggctt	2820
ccccgtcaag	ctctaaatcg	ggggctccct	tttagggttcc	gattttagtgc	tttacggcac	2880
ctcgacccca	aaaaacttga	tttagggtat	ggttcacgt	gtggccatc	gcctgtatag	2940
acggttttc	gcctttgac	gttggagtcc	acgttcttta	atagtggact	cttggccaa	3000
actggaaaca	cactcaaccc	tatctcggtc	tattcttttgc	atttataagg	gattttgcgc	3060
atttcggcct	attggtaaa	aatgagctg	atthaacaaa	aatthaacgc	gaattttaac	3120

-continued

aaaatattaa	cgtttacaat	ttcaggtggc	actttcggg	gaaatgtgcg	cggaaccct	3180
atttgttat	ttttctaaat	acattcaaat	atgtatccgc	tcatgaatta	attcttagaa	3240
aaactcatcg	agcatcaaat	gaaactgcaa	tttattcata	tcaggattat	caataccata	3300
ttttgaaaa	agccgtttct	gtaatgaagg	agaaaactca	ccgaggcagt	tccataggat	3360
ggcaagatcc	tggtatcggt	ctgcgatcc	gactcgtcca	acatcaatac	aacctattaa	3420
tttccccctcg	tcaaaaataa	ggttatcaag	tgagaaatca	ccatgagtga	cgactgaatc	3480
cggtgagaat	ggcaaaagtt	tatgcatttc	tttccagact	tgttcaacag	gccagccatt	3540
acgctcgatca	tcaaaaatcac	tcgcatcaac	caaaccgtta	ttcattcgta	attgcgcctg	3600
agcgagacga	aatacgcgat	cgctgttaaa	aggacaatta	caaacaggaa	tcgaatgcaa	3660
ccggcgcagg	aacactgccca	gkgcatcaac	aatatttca	cctgaatcag	gatatttttc	3720
taatacctgg	aatgctgttt	tcccgggat	cgcagtggtt	agtaaccatg	catcatcagg	3780
agtagggata	aatgcttga	tggtcggaaag	aggcataaaat	tccgtcagcc	agtttagtct	3840
gaccatctca	tctgtAACAT	cattggcaac	gctaccttgc	ccatgtttca	gaaacaactc	3900
tggcgcatcg	ggcttcccatt	acaatcgata	gattgtcgca	cctgattgcc	cgacattatc	3960
gcgagcccat	ttatacccat	ataaaatcago	atccatgttg	gaatttaatc	gcccctaga	4020
gcaagacgtt	tcccgttggaa	tatggctcat	aacacccctt	gtattactgt	ttatgtaa	4080
agacagtttt	attgttcatg	acccaaatcc	cttaacgtga	gttttcgttc	cactgagcgt	4140
cagacccctgt	agaaaagatc	aaaggatctt	cttgagatcc	ttttttctg	cgcgtaatct	4200
gctgcttgc	aacaaaaaaaaa	ccaccgctac	cagcgggtgg	ttgtttgcgg	gatcaagagc	4260
taccaactct	ttttccgaag	gtaactggct	tcagcagagc	gcagatacca	aatactgtcc	4320
ttcttagtgc	gccgttagtta	ggccaccact	tcaagaactc	tgtagcaccc	cctacataacc	4380
tcgtctgtct	aatcctgtta	ccagtggctg	ctgccagtgg	cgataagtgc	tgtcttaccg	4440
ggttggactc	aagacgatag	ttaccggata	aggcgcagcg	gtcgggctga	acgggggggtt	4500
cgtgcacaca	gcccacccct	gagcgaacga	cctacacccga	actgagatac	ctacagcgt	4560
agctatgaga	aagcgcacacg	cttcccgaaag	ggagaaaaggc	ggacaggtat	ccggtaagcg	4620
gcaggggtcg	aacaggagag	cgcacgaggg	agttccagg	gggaaacgcc	tggtatctt	4680
atagtccctgt	cgggtttcgc	cacccctgtac	tttagcgtcg	atttttgtga	tgctcgtcag	4740
ggggggcggag	cctatggaaa	aacgcgcagca	acgcggccctt	tttacggttc	ctggcccttt	4800
gctggccctt	tgctcACATG	ttctttctgt	cgttatcccc	tgattctgtg	gataaccgt	4860
ttaccgcctt	tgagtgagct	gataccgcctc	ccgcgcagccg	aacgacccgg	cgcagcgt	4920
cagtgagcga	ggaagcggaa	gagcgcctga	tgcggtattt	tcccttacg	cattctgtcg	4980
gtatTTcaca	ccgcataata	ggtgactct	cagtacaatc	tgctctgtat	ccgcataatc	5040
aagcccgat	acactcccgct	atcgctacgt	gactgggtca	tggctgcgcc	ccgacacccg	5100
ccaaacacccg	ctgacgcgccc	ctgacgggct	tgtctgctcc	ccgcataccgc	ttacagacaa	5160
gctgtgaccg	tctccggag	ctgcgtgtgt	cagaggtttt	caccgtcatc	accgaaacgc	5220
gccccccgc	tgccgtaaag	ctcatcageg	tggtcgtgaa	gcatgtccaca	gatgtctgcc	5280
tgttcatccg	cgtccagctc	gttgagtttc	tccagaagcg	ttaatgtctg	gcttctgata	5340
aagcggggcca	tgttaaggc	ggtttttcc	tgtttgtca	ctgatgcctc	cgtgtaaagg	5400

-continued

ggatttctgt tcatgggggt aatgataccg atgaaacgag agaggatgt cacgatacg	5460
tttactgtatc atgaacatgc ccggtaactg gaacgttgcg aggtaaaca actggcgta	5520
tggatgcgc gggaccagag aaaaatcaact cagggtcaat gccagcgctt cgtaataca	5580
gtatgttgtt ttccacaggg tagecagcag catcctgca tgcaatccg gaacataatg	5640
gtgcaggcg ctgacttccg cgtttccaga ctttacgaaa cacggaaacc gaagaccatt	5700
catgttgtt ctcaggatcg agacgtttt cagcagcagt cgcttcacgt tcgctcgct	5760
atcggtgatt cattctgcta accagtaagg caaccccgcc agcctagccg ggctctcaac	5820
gacaggagca cgatcatgcg caccgtggg gcccgtatgc cggcgataat ggctcgctc	5880
tcgcccggaaac gtttggggc gggaccagtg acgaaggctt gagcgagggc gtgcaagatt	5940
ccgaataccg caagcgacag gccgatcato gtcgcgcctcc agcgaaagcg gtccctcgcc	6000
aaaaatgaccc agagcgctgc cggcacctgt cctacgagt gcatgataaa gaagacagtc	6060
ataaagtgcgg cgacgatagt catgccccgcg gcccacccggaggagctgac tgggttgaag	6120
gctctcaagg gcatcggtcg agatcccggt gcctaatacg ttagactaact tacattaatt	6180
gcgttgcgt cactgcccgc tttccagtcg gaaacctgt cgtgccagct gcattaatga	6240
atcggccaac ggcggggggag aggccgtttt cgtattggc gccagggtgg tttttttttt	6300
caccagttag acgggcaaca gctgattgcg cttcaccgcg tggccctgag agagttgcag	6360
caagcggtcc acgctgggtt gccccagcag gcgaaaatcc tgggtatgg tggtaacgg	6420
cgggatataa catgagctgt ctccggatc gtcgtatccc actaccgaga tatccgcacc	6480
aacgcgcagc cccgactcgg taatggcgcg cattgcgcgc agcgccatct gatcggtggc	6540
aaccagcatc cgatgggaa cgatgcctc attcagcatt tgcatgggtt gttaaaacc	6600
ggacatggca ctccagtcgc ctccggatc cgctatcgcc tgaatttgc tgcgagttag	6660
atatttatgc cagccagcca gacgcagacg cgccgagaca gaacttaatg gcccgcetaa	6720
cagcgcgatt tgctgggtac ccaatgcgc acagatgcctcc acgcccagtc gcttacgc	6780
ttcatgggag aaaataatac tggatgggg tggatggta gagacatcaa gaaataacgc	6840
cggaacatta gtgcaggcag cttccacago aatggcatcc tggatccgcgatggatgtt	6900
aatgatcagc ccactgacgc gttgcgcgag aagattgtgc accggccgtt tacaggctc	6960
gacgcgcgtt cgttctacca tcgacaccac cacgtggca cccagttatcgccgagaa	7020
ttaatcgcc ggcacaattt ggcacggcgc gtgcaggcc agactggagg tggcaacgc	7080
aatcagcaac gactgtttgc cccggatgtt tggatccgcgatggatgtt gttaaatccat	7140
ctccggccatc ggcgttcca cttttcccg cttttcgca gaaacgtggc tggccctgggtt	7200
caccacgggg gaaacgggtct gataagagac accggccatc tctgcgcacat cgtataacgt	7260
tactggtttc acattcacca ccctgaattt actctcttc gggcgctatc atgcataacc	7320
gcgaaagggtt ttgcgcatt cgatgggtc cggatctcg acgctctcc ttatgcgact	7380
cctgcattag gaagcgccccc agtagtaggt tgaggccgtt gagcaccgcg gccgcaagga	7440
atgggtgcattg caaggagatg ggcggcaaca gttccggccgac cggggccgtt gcccacccat	7500
ccacgcggaa acaagcgctc atgagccgcg agtggcgagc cggatctcc ccatacggtga	7560
tgtcgccatg ataggcgcca gcaaccgcac ctgtggcgcc ggtgatgcgcg gccacgatgc	7620
gtccggcgta gaggatcgag atctcgatcc cgcgaaat	7658

-continued

```

<210> SEQ ID NO 65
<211> LENGTH: 2451
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)..(2451)
<223> OTHER INFORMATION: Membrane protein YqiK-GFP-His8

<400> SEQUENCE: 65

atg gat gat att gtt aat tct gtg ccc tcc tgg atg ttt acc gca att      48
Met Asp Asp Ile Val Asn Ser Val Pro Ser Trp Met Phe Thr Ala Ile
1           5           10          15

att gcc gta tgc att ctg ttt att att gga att att ttc gcc agg ctc      96
Ile Ala Val Cys Ile Leu Phe Ile Ile Gly Ile Ile Phe Ala Arg Leu
20          25          30

tat cgt cgc gct tcg gca gag caa gct ttt gtt cgt act ggt tta ggt      144
Tyr Arg Arg Ala Ser Ala Glu Gln Ala Phe Val Arg Thr Gly Leu Gly
35          40          45

ggg caa aaa gtg gta atg agc ggt ggc gca atc gtc atg ccg atc ttt      192
Gly Gln Lys Val Val Met Ser Gly Gly Ala Ile Val Met Pro Ile Phe
50          55          60

cat gaa ata atc ccc atc aat atg aat act ctg aag ctg gaa gtc agc      240
His Glu Ile Ile Pro Ile Asn Met Asn Thr Leu Lys Leu Glu Val Ser
65          70          75          80

cgc tca acc att gat agc ctg att acg aaa gat cgt atg cgc gtc gat      288
Arg Ser Thr Ile Asp Ser Leu Ile Thr Lys Asp Arg Met Arg Val Asp
85          90          95

gta gta gtc gct ttc ttt gtg cgg gta aaa cct tca gta gaa ggg att      336
Val Val Val Ala Phe Phe Val Arg Val Lys Pro Ser Val Glu Gly Ile
100         105         110

gcc acc gct gcc cag acg ctg ggg caa cgc acc ctg tcg cct gaa gac      384
Ala Thr Ala Ala Gln Thr Leu Gly Gln Arg Thr Leu Ser Pro Glu Asp
115         120         125

tta cgt atg ttg gtt gaa gat aaa ttt gtc gat gcc ctc cgt gca aca      432
Leu Arg Met Leu Val Glu Asp Lys Phe Val Asp Ala Leu Arg Ala Thr
130         135         140

gct gcg caa atg acc atg cat gag tta cag gat acc cgc gag aac ttt      480
Ala Ala Gln Met Thr His Glu Leu Gln Asp Thr Arg Glu Asn Phe
145         150         155         160

gtg cag ggg gtg caa aat aca gtg gca gaa gac ctg tcg aaa aac ggt      528
Val Gln Gly Val Gln Asn Thr Val Ala Glu Asp Leu Ser Lys Asn Gly
165         170         175

ctg gaa ctg gag agc gtt tca ctt acc aac ttt aac cag acc tcg aaa      576
Leu Glu Leu Glu Ser Val Ser Leu Thr Asn Phe Asn Gln Thr Ser Lys
180         185         190

gaa cat ttc aat ccg aac aat gcc ttt gac gcc gaa ggt tta acc aaa      624
Glu His Phe Asn Pro Asn Asn Ala Phe Asp Ala Glu Gly Leu Thr Lys
195         200         205

ctg act cag gaa aca gag cgc cgt cgc gaa cgt aac gaa gtt gaa      672
Leu Thr Gln Glu Thr Glu Arg Arg Arg Glu Arg Asn Glu Val Glu
210         215         220

cag gat gta gaa gtt gcg gtg cgt gaa aaa aat cgc gat gcg cta tcg      720
Gln Asp Val Glu Val Ala Val Arg Glu Lys Asn Arg Asp Ala Leu Ser
225         230         235         240

cgc aag ctg gag att gaa cag caa gaa gcg ttt atg acg ctt gag cag      768

```

-continued

Arg Lys Leu Glu Ile Glu Gln Gln Glu Ala Phe Met Thr Leu Glu Gln		
245	250	255
gag cag cag gtt aaa acc cgt act gcc gaa cag aat gca cgt att gcg		816
Glu Gln Gln Val Lys Thr Arg Thr Ala Glu Gln Asn Ala Arg Ile Ala		
260	265	270
gct ttt gaa gct gaa cgt cgt cgt gaa gca gag cag aca cga att ctg		864
Ala Phe Glu Ala Glu Arg Arg Glu Ala Glu Gln Thr Arg Ile Leu		
275	280	285
gct gaa cga cag att cag gaa aca gaa atc gac cgc gaa cag gcc gtc		912
Ala Glu Arg Gln Ile Gln Glu Thr Glu Ile Asp Arg Glu Gln Ala Val		
290	295	300
cgc tca aga aag gtt gaa gct gaa cgt gaa gtt cgc att aaa gag atc		960
Arg Ser Arg Lys Val Glu Ala Glu Arg Glu Val Arg Ile Lys Glu Ile		
305	310	315
gaa cag cag cag gtc acc gaa atc gct aac cag acg aaa tcg atc gct		1008
Glu Gln Gln Val Thr Glu Ile Ala Asn Gln Thr Lys Ser Ile Ala		
325	330	335
att gcc gcc aaa tcg gaa caa cag tcc cag gca gaa gcg cgt gct aat		1056
Ile Ala Ala Lys Ser Glu Gln Gln Ser Gln Ala Glu Ala Arg Ala Asn		
340	345	350
ctc gca ctt gca gaa gcg gta agc gcc caa caa aac gta gaa acc act		1104
Leu Ala Leu Ala Glu Ala Val Ser Ala Gln Gln Asn Val Glu Thr Thr		
355	360	365
cgc cag act gcc gaa gcc gat cgt gct aaa caa gtt gcc cta atc gct		1152
Arg Gln Thr Ala Glu Ala Asp Arg Ala Lys Gln Val Ala Leu Ile Ala		
370	375	380
gcc gcg cag gat gca gaa acc aaa gcg gtt gaa ctg acc gtg cgg gcg		1200
Ala Ala Gln Asp Ala Glu Thr Lys Ala Val Glu Leu Thr Val Arg Ala		
385	390	395
400		
aaa gca gaa aaa gaa gcc gca gag atg cag gcg gcg gct atc gtt gag		1248
Lys Ala Glu Lys Glu Ala Ala Glu Met Gln Ala Ala Ala Ile Val Glu		
405	410	415
tta gcc gaa gct aca cgt aaa aag ggt ctg gcg gaa gca gaa gca caa		1296
Leu Ala Glu Ala Thr Arg Lys Lys Gly Leu Ala Glu Ala Glu Ala Gln		
420	425	430
cgt gcg ctg aac gat gct atc aac gta ctt tct gat gaa caa acc agc		1344
Arg Ala Leu Asn Asp Ala Ile Asn Val Leu Ser Asp Glu Gln Thr Ser		
435	440	445
ctt aaa ttc aaa ctg gcc ttg ttg cag gcg ctg cct gcg gta ata gag		1392
Leu Lys Phe Lys Leu Ala Leu Leu Gln Ala Pro Ala Val Ile Glu		
450	455	460
aaa tcc gtt gag ccg atg aag tca atc gac ggt atc aag att att cag		1440
Lys Ser Val Glu Pro Met Lys Ser Ile Asp Gly Ile Lys Ile Ile Gln		
465	470	475
480		
gtc gat ggt ctg aat cgt ggc ggc gct gcg ggt gat gca aac acg ggt		1488
Val Asp Gly Leu Asn Arg Gly Ala Ala Gly Asp Ala Asn Thr Gly		
485	490	495
aat gtg ggg ggc gga aac ctg gcg gaa caa gca tta tca gcc gct ctc		1536
Asn Val Gly Gly Asn Leu Ala Glu Gln Ala Leu Ser Ala Ala Leu		
500	505	510
tct tac cgc aca cag gca ccg ctg att gac tcc ttg ctc aat gaa att		1584
Ser Tyr Arg Thr Gln Ala Pro Leu Ile Asp Ser Leu Leu Asn Glu Ile		
515	520	525
ggc gtt tca ggc ggc tca ctg gcg gca ttg act tca ccc tta acc tca		1632
Gly Val Ser Gly Gly Ser Leu Ala Ala Leu Thr Ser Pro Leu Thr Ser		
530	535	540
aca act ccc gtc gaa gaa aaa gca gag tcg gta cct gga tcc gaa aac		1680

-continued

Thr Thr Pro Val Glu Glu Lys Ala Glu Ser Val Pro Gly Ser Glu Asn 545 550 555 560	
ctg tac ttc cag ggt caa ttc agc aaa gga gaa gaa ctt ttc act gga Leu Tyr Phe Gln Gly Gln Phe Ser Lys Gly Glu Glu Leu Phe Thr Gly 565 570 575	1728
gtt gtc cca att ctt gtt gaa tta gat ggt gat gtt aat ggg cac aaa Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys 580 585 590	1776
ttt tct gtc agt gga gag ggt gaa ggt gat gct aca tac gga aaa ctc Phe Ser Val Ser Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu 595 600 605	1824
acc ctt aaa ttt att tgc act act gga aaa cta cct gtt cca tgg cca Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro 610 615 620	1872
aca ctt gtc act act ctg acc tat ggt gtt caa tgc ttt tcc cgt tat Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr 625 630 635 640	1920
ccg gat cac atg aaa cgg cat gac ttt ttc aag agt gcc atg ccc gaa Pro Asp His Met Lys Arg His Asp Phe Phe Lys Ser Ala Met Pro Glu 645 650 655	1968
ggt tat gta cag gaa cgc act ata tct ttc aaa gat gac ggg aac tac Gly Tyr Val Gln Glu Arg Thr Ile Ser Phe Lys Asp Asp Gly Asn Tyr 660 665 670	2016
aag acg cgt gct gaa gtc aag ttt gaa ggt gat acc ctt gtt aat cgt Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg 675 680 685	2064
atc gag tta aag ggt att gat ttt aaa gaa gat gga aac att ctc gga Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly 690 695 700	2112
cac aaa cta gag tac aac tat aac tca cac aat gta tac atc acg gca His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Thr Ala 705 710 715 720	2160
gac aaa caa aag aat gga atc aaa gct aac ttc aaa att cgc cac aac Asp Lys Gln Lys Asn Gly Ile Lys Ala Asn Phe Lys Ile Arg His Asn 725 730 735	2208
att gaa gat ggt tcc gtt caa cta gca gac cat tat caa caa aat act Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr 740 745 750	2256
cca att ggc gat ggc cct gtc ctt tta cca gac aac cat tac ctg tcg Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser 755 760 765	2304
aca caa tct gcc ctt tcg aaa gat ccc aac gaa aag cgt gac cac atg Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met 770 775 780	2352
gtc ctt ctt gag ttt gta act gct gct ggg att aca cat ggc atg gat Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp 785 790 795 800	2400
gag ctc tac aaa aag ctt gcg gcc cat cat cat cac cac cac cac Glu Leu Tyr Lys Leu Ala Ala His His His His His His His His His 805 810 815	2448
tga	2451

```

<210> SEQ ID NO 66
<211> LENGTH: 816
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

```

-continued

<400> SEQUENCE: 66

```

Met Asp Asp Ile Val Asn Ser Val Pro Ser Trp Met Phe Thr Ala Ile
1           5          10          15

Ile Ala Val Cys Ile Leu Phe Ile Ile Gly Ile Ile Phe Ala Arg Leu
20          25          30

Tyr Arg Arg Ala Ser Ala Glu Gln Ala Phe Val Arg Thr Gly Leu Gly
35          40          45

Gly Gln Lys Val Val Met Ser Gly Gly Ala Ile Val Met Pro Ile Phe
50          55          60

His Glu Ile Ile Pro Ile Asn Met Asn Thr Leu Lys Leu Glu Val Ser
65          70          75          80

Arg Ser Thr Ile Asp Ser Leu Ile Thr Lys Asp Arg Met Arg Val Asp
85          90          95

Val Val Val Ala Phe Phe Val Arg Val Lys Pro Ser Val Glu Gly Ile
100         105         110

Ala Thr Ala Ala Gln Thr Leu Gly Gln Arg Thr Leu Ser Pro Glu Asp
115         120         125

Leu Arg Met Leu Val Glu Asp Lys Phe Val Asp Ala Leu Arg Ala Thr
130         135         140

Ala Ala Gln Met Thr Met His Glu Leu Gln Asp Thr Arg Glu Asn Phe
145         150         155         160

Val Gln Gly Val Gln Asn Thr Val Ala Glu Asp Leu Ser Lys Asn Gly
165         170         175

Leu Glu Leu Glu Ser Val Ser Leu Thr Asn Phe Asn Gln Thr Ser Lys
180         185         190

Glu His Phe Asn Pro Asn Asn Ala Phe Asp Ala Glu Gly Leu Thr Lys
195         200         205

Leu Thr Gln Glu Thr Glu Arg Arg Arg Glu Arg Asn Glu Val Glu
210         215         220

Gln Asp Val Glu Val Ala Val Arg Glu Lys Asn Arg Asp Ala Leu Ser
225         230         235         240

Arg Lys Leu Glu Ile Glu Gln Gln Glu Ala Phe Met Thr Leu Glu Gln
245         250         255

Glu Gln Gln Val Lys Thr Arg Thr Ala Glu Gln Asn Ala Arg Ile Ala
260         265         270

Ala Phe Glu Ala Glu Arg Arg Glu Ala Glu Gln Thr Arg Ile Leu
275         280         285

Ala Glu Arg Gln Ile Gln Glu Thr Glu Ile Asp Arg Glu Gln Ala Val
290         295         300

Arg Ser Arg Lys Val Glu Ala Glu Arg Glu Val Arg Ile Lys Glu Ile
305         310         315         320

Glu Gln Gln Gln Val Thr Glu Ile Ala Asn Gln Thr Lys Ser Ile Ala
325         330         335

Ile Ala Ala Lys Ser Glu Gln Gln Ser Gln Ala Glu Ala Arg Ala Asn
340         345         350

Leu Ala Leu Ala Glu Ala Val Ser Ala Gln Gln Asn Val Glu Thr Thr
355         360         365

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

Ala Ala Gln Asp Ala Glu Thr Lys Ala Val Glu Leu Thr Val Arg Ala
385         390         395         400

```

-continued

Lys Ala Glu Lys Glu Ala Ala Glu Met Gln Ala Ala Ala Ile Val Glu
405 410 415

Leu Ala Glu Ala Thr Arg Lys Lys Gly Leu Ala Glu Ala Glu Ala Gln
420 425 430

Arg Ala Leu Asn Asp Ala Ile Asn Val Leu Ser Asp Glu Gln Thr Ser
435 440 445

Leu Lys Phe Lys Leu Ala Leu Leu Gln Ala Leu Pro Ala Val Ile Glu
450 455 460

Lys Ser Val Glu Pro Met Lys Ser Ile Asp Gly Ile Lys Ile Ile Gln
465 470 475 480

Val Asp Gly Leu Asn Arg Gly Gly Ala Ala Gly Asp Ala Asn Thr Gly
485 490 495

Asn Val Gly Gly Asn Leu Ala Glu Gln Ala Leu Ser Ala Ala Leu
500 505 510

Ser Tyr Arg Thr Gln Ala Pro Leu Ile Asp Ser Leu Leu Asn Glu Ile
515 520 525

Gly Val Ser Gly Gly Ser Leu Ala Ala Leu Thr Ser Pro Leu Thr Ser
530 535 540

Thr Thr Pro Val Glu Glu Lys Ala Glu Ser Val Pro Gly Ser Glu Asn
545 550 555 560

Leu Tyr Phe Gln Gly Gln Phe Ser Lys Gly Glu Glu Leu Phe Thr Gly
565 570 575

Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys
580 585 590

Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu
595 600 605

Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro
610 615 620

Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr
625 630 635 640

Pro Asp His Met Lys Arg His Asp Phe Phe Lys Ser Ala Met Pro Glu
645 650 655

Gly Tyr Val Gln Glu Arg Thr Ile Ser Phe Lys Asp Asp Gly Asn Tyr
660 665 670

Lys Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg
675 680 685

Ile Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly
690 695 700

His Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Thr Ala
705 710 715 720

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

Ile Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr
740 745 750

Pro Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser
755 760 765

Thr Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met
770 775 780

Val Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp
785 790 795 800

-continued

Glu	Leu	Tyr	Lys	Lys	Leu	Ala	Ala	His						
805					810									815

```

<210> SEQ ID NO 67
<211> LENGTH: 8471
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)..(8471)
<223> OTHER INFORMATION: Plasmid pYidC

<400> SEQUENCE: 67

taataacgact cactataggg gaattgttag cggtataacaa ttccccctcta gaaataattt      60
tgtttaacctt taagaaggag accagtttat ggattcgcaa cgcaatcttt tagtcatcgc      120
tttgctgttc gtgtctttca tgatctggca agcctggag caggataaaa acccgcaacc      180
tcaggcccaa cagaccacgc agacaacgcac caccgcagcg ggttagcgccg ccgaccagg      240
cgtaccggcc agtggccagg ggaaaactgtat ctcggtaaag accgacgtgc ttgatctgac      300
catcaacacc cgtggtggtg atgttgagca agctctgtct cctgcttacc cgaaagagct      360
gaactctacc cagccgttcc agctgttggaa aacttcacccg cagtttattt atcaggcaca      420
gagcggctcg accggctcgat atggcccggaa taacccgggt aacggcccgcc gtccgctgta      480
taacggttcaa aaagacgctt atgtgctggc tgaaggtaaa aacgaactgc aggtgccgat      540
gacgtataacc gacgcggcag gcaacacgtt taccaaaacg tttgtcttga aacgtggta      600
ttacgctgtc aacgtcaact acaacgtgcgaa acgcgtggc gagaaaaccgc tggaaatctc      660
ctcggttggc cagttgaagc aatccatcac tctgccaccg catctcgata ccggaagcag      720
caacttcgcgatcgtccaccc tccgtggcgc ggcgtactcc acgcctgacg agaagtatga      780
gaaatacaag ttcgatacca ttggcgataa cgaaaacctg aacatctttt cgaaagggtgg      840
ttgggtggcg atgctgcaac agtatttcgcg acggcgctgg atcccgatca acgacggatc      900
caacaacttc tataccgcta atctgggtaa cggcatcgcc gctatcggtt ataaatctca      960
ggcggtactg gttcagccctg gtcagactgg cgcgatgaa acgaccctgtt ggggtggccc      1020
ggaaatccag gacaaaatgg cagctgtgc tccgcacccg gatctgaccc ttgattacgg      1080
ttgggttggc ttcatctctc agccgctgtt caaaactgtctt aatggatcc atagcttgt      1140
gggttaactgg ggcttctcca ttatcatcat cacctttatc gttcgtggca tcatgtaccc      1200
gctgaccaaa ggcgtactaca cctccatggc gaagatcgatc atgttgcacg cgaagattca      1260
ggcaatgcgt gaggcgctgg gcgtacgaa acagcgtatc agccaggaaa tggatggcgct      1320
gtacaaaatgtt gagaaggatc acccgctggg cggctgttcc cccgtgttgc tccagatgcc      1380
aatcttcgtt ggcgttgcgtt acatgtgtatc ggggttccgtt gaaactgcgtc aggccaccgtt      1440
tgcactgtgg atccacgacc tgcgtggcaca ggaccctgtt tacatcctgc cgatcctgtat      1500
ggcgtaacg atggttctca ttcaaaaatgtt gtcgcccggacc acagtgaccc acccgatgca      1560
gcagaagatc atgaccttta tgccggcat cttcacccgtt ttcttccgtt ggttcccgcc      1620
aggctgtggc ctgtactata tcgtcagca cctggtaacc attattcagc agcagctgtat      1680
ttaccgttgtt ctggaaaaac gtggcctgca tagccgcgag aagaaaaat ccgaattcag      1740
caaaggagaa gaaactttca ctggagttgtt cccaaatttctt gttgaatttag atggtgatgt      1800

```

-continued

taatggcac	aaattttctg	tcagtggaga	gggtgaagg	gatgctacat	acggaaaact	1860
caccctaaa	tttatttgca	ctactggaaa	actacctgtt	ccatggccaa	cacttgtcac	1920
tactctgacc	tatggtgttc	aatgctttc	ccgttatccg	gatcacatga	aacggcatga	1980
cttttcaag	agtgccatgc	ccgaaggta	tgtacaggaa	cgcactatat	ctttcaaaga	2040
tgacgggaaac	tacaagacgc	gtgctgaagt	caagttgaa	ggtgataccc	ttgttaatcg	2100
tatcgagtt	aagggtattt	atttaaaga	agatggaaac	attctcgac	acaaaactaga	2160
gtacaactat	aactcacaca	atgtatacat	cacggcagac	aaacaaaaga	atggaatcaa	2220
agctaacttc	aaaattcgcc	acaacattga	agatggttc	gttcaactag	cagaccatta	2280
tcaacaaaat	actccaattt	gcatggccc	tgccttttta	ccagacaacc	attacctgtc	2340
gacacaatct	gcccttcga	aagatccaa	cgaaaagcgt	gaccacatgg	tccttcttga	2400
gtttgtact	gtctgtggg	ttacacatgg	catggatgag	ctctacaat	tgggtaccc	2460
cgagcaccac	caccaccacc	actgacatag	gaggtcctcc	tatgtcaatt	caacatttc	2520
gtgtcgccct	tattcccttt	tttgcggcat	tttgcctcc	tgtttttgtc	caccagaaa	2580
cgctggtgaa	agtaaaagat	gctgaagatc	agttgggtgc	acgagtgggt	tacatcgAAC	2640
tggatctcaa	cagcggtaaag	atccttgaga	gttttcgccc	cgaagaacgt	tttccaaatga	2700
tgagcacttt	taaagtctg	ctatgtggcg	cggtattatc	ccgtgttgc	gccgggcaag	2760
agcaactcg	tcgcccata	cactattctc	agaatgactt	ggttgagtac	tcaccagtca	2820
cagaaaagca	tcttacggat	ggcatgacag	taagagaatt	atgcagtgt	gccataacca	2880
tgagtgtataa	cactcgcc	aacttacttc	tgacaacatg	cgaggacccg	aaggagctaa	2940
ccgccttttt	gcacaacatg	ggggatcatg	taactcgcc	tgatcggtgg	gaaccggagc	3000
tgaatgaagc	cataccaaac	gacgagcgt	acaccacatg	gcctgcagca	atggcaacaa	3060
cgttgcgcaa	actattaact	ggcgaactac	ttactctage	ttcccgccaa	caattaatag	3120
actggatgga	ggcggataaa	gttgcaggac	cacttctgc	ctcgccctt	ccggctggct	3180
ggtttattgc	tgataaatct	ggagccggtg	agcgtgggtc	tgcgggtatc	attgcagcac	3240
tgggggccaga	tggttaagccc	tcccgatcg	tagttatcta	cacgacgggg	agtcaggcaa	3300
ctatggatga	acgaaataga	cagatcgctg	agatagggtc	ctcaactgatt	aagcatttgt	3360
aactaacaaa	ccccgaaagg	aagctgagtt	ggctgctgc	accgctgagc	aataacttagc	3420
ataacccctt	ggggcctcta	aacgggtctt	gaggggtttt	ttgctgaaag	gaggaactat	3480
atccggatttgc	gcgaatggga	cgcgcctgt	agcggcgc	taagcgcggc	gggtgtggtg	3540
gttacgcgca	gcgtgaccgc	tacacttgc	agcgccctag	cgcccgctcc	tttcgccttc	3600
ttcccttctt	ttctcgccac	gttgcggc	tttccccgtc	aagctctaaa	tcgggggctc	3660
cctttagggt	tccgatttag	tgctttacgg	cacctcgacc	ccaaaaaact	tgattagggt	3720
gatgggttcac	gtatggggcc	atcgccctga	tagacggttt	ttcgccctt	gacgtggag	3780
tccacgttct	ttaatagtgg	actcttgc	caaactggaa	caacactcaa	ccctatctcg	3840
gtctattctt	ttgatttata	agggattttgc	ccgatttcg	cctattgggt	aaaaaatgag	3900
ctgatataac	aaaaatttaa	cgcgaatttt	aacaaaat	taacgtttac	aatttcaggt	3960
ggcacttttc	ggggaaatgt	gcgcggaaacc	cctatttgt	tatTTTCTA	aatacatca	4020
aatatgtatac	cgctcatgaa	ttaattctta	aaaaactca	tcgagcatca	aatgaaactg	4080

-continued

caatttattc atatcaggat tatcaatacc atattttga aaaagccgtt tctgtaatga	4140
aggagaaaac tcaccgaggc agttccatag gatggcaaga tcctggatc ggtctgcgtat	4200
tccgactcgta ccaacatcaa tacaacctat taattcccc tcgtcaaaaa taaggttatc	4260
aagttagaaaa tcaccatgag tgacgactga atccggtag aatggcaaaa gtttatgcat	4320
ttctttccag acttgttcaa caggccagcc attacgctcg tcatcaaaat cactcgcatc	4380
aacccaaaccg ttatttcattc gtgattgcgc ctgagcggaa cgaaatacgc gatcgctgtt	4440
aaaaggacaa ttacaaacag gaatcgaatg caaccggcgc aggaacactg ccagcgcata	4500
aacaatattt tcacctgaat caggatattc ttctaatacc tggaaatgctg tttcccccgg	4560
gatcgcaatgt gtgagtaacc atgcatacatc aggagtaacgg ataaaatgct tggatggctgg	4620
aagaggcata aattccgtca gccagtttag tctgaccatc tcatctgtaa catcattggc	4680
aacgcttaccc ttggcatgtt tcagaaacaa ctctggcgc tcgggcttcc catacaatcg	4740
atagattgtc gcacctgattt gcccgcattt atcgcgcggcc catttatacc catataaatac	4800
agcatccatg ttggaaattt atcgcggcct agagcaagac gtttcccggtt gaatatggct	4860
cataacaccc cttgttattac tgtttatgtt agcagacagt tttattgttc atgaccaaaa	4920
tcccttaacg tgatgtttcg ttccactgag cgtcagaccc cgtagaaaaag atcaaaggat	4980
tttcttgaga tccctttttt ctgcgcgtaa tctgctgtt gcaaacaaaa aaaccacgc	5040
taccagcggt ggtttgcggcc cccggatcaag agctaccaac tcttttcccg aaggttaactg	5100
gcttcagcag agcgcagata ccaaataactg tccttcttagt gtagccgtt taggcccacc	5160
acttcaagaa ctctgttagca ccgcctacat acctcgctct gctaattcttgc ttaccagtgg	5220
ctgctgccag tggcgataag tcgtgtctt ccgggttggg ctcaagacga tagttaccgg	5280
ataaggcga ggggtcgccc tgaacgggggg gttcgtgcac acagcccacg ttggagcgaa	5340
cgacctacac cgaactgaga tacctacago gtgagctatg agaaagcgcc acgcttcccg	5400
aaggggagaaa ggcggacagg tatccggtaa cggggcagggt cggAACAGGA gagcgcacga	5460
ggggagcttcc agggggaaac gcctggatc ttatagtc tgctgggtt cgccacctct	5520
gacttggcg tcgatgtttt tgatgtctgt caggggggcc gggctatgg aaaaacgcca	5580
gcaacgcggc ctttttacgg ttccctggcct tttgctggcc tttgtctcac atgttcttc	5640
ctgcgttatac ccctgttatttctt gtggataacc gtattaccgc ctttggatgtt gctgtatcc	5700
ctcgccgcag ccgaacgacc gagegcageg agtcagtggag cggaaagcg gaagagcgcc	5760
tgatgcggta ttttcttctt acgcatactgt gcggttatttcc acaccgcata tatggtgcac	5820
tctcagtaca atctgtctg atgcgcata gttaaagccag tatacactcc gctatcgcta	5880
cgtgactggg tcatggctgc gccccacac ccggcaacac ccgctgacgc gcccgtacgg	5940
gcttggctgc tcccgccatc cgcttacaga caagctgtt gcttcccg gaggctgcgtat	6000
tgtcagaggt ttccatggc atcaccggaa cggccggaggc agctgggtt aagctcatca	6060
gcgtgggtgtt gaaaggatcc acagatgtctt gcttggatgtt ccgggtccag ctcgttggat	6120
ttctccagaa gcttaatgtt ctgggttctt gttaaagccggg ccatgttaag ggcgggtttt	6180
tcctgtttgg tcaactgtatgc ctccgtgtaa gggggatttc tggatgttgg ggtaatgata	6240
ccgatgaaac gagagaggat gctcacgata cgggttactg atgatgaaca tgcccggtt	6300
ctggaaacgtt gtgagggtaa acaactggcg gttatggatgc ggcgggacca gagaaaaatc	6360

-continued

actcagggtc aatgccagcg cttegttaat acagatgtag gtgttccaca gggtagccag	6420
cacgatcctg ccatgcagat ccgaaacata atggcgcagg ggcgtgactt ccgcgttcc	6480
agactttacg aaacacggaa accgaagaco attcatgtt ttcgtcaggt cgccagacgtt	6540
ttgcagcagc agtcgttca cgttcgtcg cgtatcggtt attcattctg ctaaccagta	6600
aggcaacccc gccagectag ccgggtcctc aacgacagga gcacgatcat ggcacccgt	6660
ggggccgcca tgccggcgat aatggcctgc ttctcgccg aacgtttgtt ggccgggacca	6720
gtgacgaagg cttgagcag ggcgtgcaag attccgaata ccgcaagcga caggccgatc	6780
atcgctcgcc tccagcgaaa gccgtcctcg ccgaaaatcg cccagagcgc tgccggcacc	6840
tgtcctacga gttgcatgat aaagaagaca gtcataagtg cggcgacgt agtcatgccc	6900
cgcccccacc ggaaggagct gactgggtt aaggctctca agggcatcg tcgagatccc	6960
ggtgccataat gagtgagcta acttacatta attgcgttgc gtcactgcc cgcttcag	7020
tcgggaaacc tgcgtgcca gctgcattaa tgaatcgccg aacgcgcggg gagaggcggt	7080
ttgcgtatttggcggccaggg tgggtttttt tttcaccagt gagacggcaca acagctgatt	7140
gccccttacc gcctggccct gagagagttt cagcaagcgg tccacgctgg tttgccccag	7200
caggcgaaaa tcctgttga tgggtgtttaa cggcgggata taacatgagc tgccttcgg	7260
atcgctgtat cccactaccg agatatccgc accaacgcgc agcccgact cggtaatggc	7320
gcgcattgcg cccagcgcca tctgtatcggtt ggcaaccaggc atcgcgatgg gaacgtatgcc	7380
ctcattcagc atttgcattgg tttgttggaaa accggacatg gcactccagt cgccctcccg	7440
ttccgcatacggcgttgcattt gattgcgttgcgtt gagatattta tgccagccag ccagacgcag	7500
acgcgcgcgacg acagaactta atggggccgc taacagcgcg atttgcgttgcgtt gacccaatgc	7560
gaccagatgc tccacgccc gtcgcgtacc gtcttcattgg gagaataaa tactgttgcgtt	7620
gggtgtctgg tcagagacat caagaataaa cgccggaaaca ttagtgcagg cagttccac	7680
agcaatggca tcctgttcat ccagcgata gttaatgtatc agcccaactga cgcgttgcgc	7740
gagaagatttgcacccggccttttacggggttgcgttccatccatcgacac	7800
caccacgctg gcacccagttt gatcgccgcg agatttaatc gccgcgacaa tttgcacgg	7860
cgcgtgcagg gccagacttgg aggtggcaac gccaatcagc aacgactgtt tgccgcac	7920
tttgttgcc acgcgggttgg gaatgttattt cagctccgcg atcgccgcgtt ccacttttc	7980
ccgcgttttc gcagaaacgt ggctggcctg gtttgcgttgcgtt ccacccatcgaa	8040
gacacccggca tactctgcga catgtataaa ctttgcgttgcgtt ccacccatcgaa	8100
ttgactcttc tccggggcgtt atcatgcattt accgcgttgcgtt gtttgcgttgcgtt attcgatgtt	8160
gtccgggatc tcgacgttgcgtt ccacccatcgaa tactctgcgttgcgtt ccacccatcgaa	8220
ggttgaggcc gtttgcgttgcgtt ccacccatcgaa tactctgcgttgcgtt ccacccatcgaa	8280
acagtcccccc ggcacccgggg ctcgttgcgtt ccacccatcgaa tactctgcgttgcgtt ccacccatcgaa	8340
cgaaggtggcg agcccgatct tcccccattcggttgcgttgcgtt ccacccatcgaa	8400
cacctgtggc gccgggttgcgtt ccacccatcgaa tactctgcgttgcgtt ccacccatcgaa	8460
tcccgccaaa t	8471

-continued

```

<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)..(2397)
<223> OTHER INFORMATION: Membrane protein YidC-GFP-His8

<400> SEQUENCE: 68

atg gat tcg caa cgc aat ctt tta gtc atc gct ttg ctg ttc gtg tct      48
Met Asp Ser Gln Arg Asn Leu Leu Val Ile Ala Leu Leu Phe Val Ser
1          5           10          15

ttc atg atc tgg caa gcc tgg gag cag gat aaa aac ccg caa cct cag      96
Phe Met Ile Trp Gln Ala Trp Glu Gln Asp Lys Asn Pro Gln Pro Gln
20         25           30

gcc caa cag acc acg cag aca acg acc acc gca gcg ggt agc gcc gcc      144
Ala Gln Gln Thr Thr Gln Thr Thr Ala Ala Gly Ser Ala Ala
35         40           45

gac cag ggc gta ccg gcc agt ggc cag ggg aaa ctg atc tcg gtt aag      192
Asp Gln Gly Val Pro Ala Ser Gly Gln Gly Lys Leu Ile Ser Val Lys
50         55           60

acc gac gtg ctt gat ctg acc atc aac acc cgt ggt ggt gat gtt gag      240
Thr Asp Val Leu Asp Leu Thr Ile Asn Thr Arg Gly Gly Asp Val Glu
65         70           75           80

caa gct ctg ctc gct tac ccg aaa gag ctg aac tct acc cag ccg      288
Gln Ala Leu Leu Pro Ala Tyr Pro Lys Glu Leu Asn Ser Thr Gln Pro
85         90           95

ttc cag ctg ttg gaa act tca ccg cag ttt att tat cag gca cag agc      336
Phe Gln Leu Leu Glu Thr Ser Pro Gln Phe Ile Tyr Gln Ala Gln Ser
100        105          110

ggc ctg acc ggt cgt gat ggc ccg gat aac ccg gct aac ggc ccg cgt      384
Gly Leu Thr Gly Arg Asp Gly Pro Asp Asn Pro Ala Asn Gly Pro Arg
115        120          125

ccg ctg tat aac gtt gaa aaa gac gct tat gtg ctg gct gaa ggt caa      432
Pro Leu Tyr Asn Val Glu Lys Asp Ala Tyr Val Leu Ala Glu Gly Gln
130        135          140

aac gaa ctg cag gtg ccg atg acg tat acc gac ggc gca ggc aac acg      480
Asn Glu Leu Gln Val Pro Met Thr Tyr Thr Asp Ala Ala Gly Asn Thr
145        150          155          160

ttt acc aaa acg ttt gtc ctg aaa cgt ggt gat tac gct gtc aac gtc      528
Phe Thr Lys Thr Phe Val Leu Lys Arg Gly Asp Tyr Ala Val Asn Val
165        170          175

aac tac aac gtg cag aac gct ggc gag aaa ccg ctg gaa atc tcc tcc      576
Asn Tyr Asn Val Gln Asn Ala Gly Glu Lys Pro Leu Glu Ile Ser Ser
180        185          190

ttt ggt cag ttg aag caa tcc atc act ctg cca ccg cat ctc gat acc      624
Phe Gly Gln Leu Lys Gln Ser Ile Thr Leu Pro Pro His Leu Asp Thr
195        200          205

ggc agc agc aac ttc gca ctg cac acc ttc cgt ggc gcg ggc tac tcc      672
Gly Ser Ser Asn Phe Ala Leu His Thr Phe Arg Gly Ala Ala Tyr Ser
210        215          220

acg cct gac gag aag tat gag aaa tac aag ttc gat acc att gcc gat      720
Thr Pro Asp Glu Lys Tyr Glu Lys Tyr Lys Phe Asp Thr Ile Ala Asp
225        230          235          240

aac gaa aac ctg aac atc tct tcg aaa ggt ggt tgg gtg gcg atg ctg      768
Asn Glu Asn Leu Asn Ile Ser Ser Lys Gly Gly Trp Val Ala Met Leu
245        250          255

caa cag tat ttc gcg acg gcg tgg atc ccg cat aac gac ggt acc aac      816

```

-continued

Gln Gln Tyr Phe Ala Thr Ala Trp Ile Pro His Asn Asp Gly Thr Asn 260 265 270	
aac ttc tat acc gct aat ctg ggt aac ggc atc gcc gct atc ggc tat Asn Phe Tyr Thr Ala Asn Leu Gly Asn Gly Ile Ala Ala Ile Gly Tyr 275 280 285	864
aaa tct cag ccg gta ctg gtt cag cct ggt cag act ggc gcg atg aac Lys Ser Gln Pro Val Leu Val Gln Pro Gly Gln Thr Gly Ala Met Asn 290 295 300	912
agc acc ctg tgg gtt ggc ccg gaa atc cag gac aaa atg gca gct gtt Ser Thr Leu Trp Val Gly Pro Glu Ile Gln Asp Lys Met Ala Ala Val 305 310 315 320	960
gct ccg cac ctg gat ctg acc gtt gat tac ggt tgg ttg tgg ttc atc Ala Pro His Leu Asp Leu Thr Val Asp Tyr Gly Trp Leu Trp Phe Ile 325 330 335	1008
tct cag ccg ctg ttc aaa ctg ctg aaa tgg atc cat agc ttt gtg ggt Ser Gln Pro Leu Phe Lys Leu Leu Lys Trp Ile His Ser Phe Val Gly 340 345 350	1056
aac tgg ggc ttc tcc att atc atc acc ttt atc gtt cgt ggc atc Asn Trp Gly Phe Ser Ile Ile Ile Thr Phe Ile Val Arg Gly Ile 355 360 365	1104
atg tac ccg ctg acc aaa gcg cag tac acc tcc atg gcg aag atg cgt Met Tyr Pro Leu Thr Lys Ala Gln Tyr Thr Ser Met Ala Lys Met Arg 370 375 380	1152
atg ttg cag ccg aag att cag gca atg cgt gag cgt ctg ggc gat gac Met Leu Gln Pro Lys Ile Gln Ala Met Arg Glu Arg Leu Gly Asp Asp 385 390 395 400	1200
aaa cag cgt atc agc cag gaa atg atg gcg ctg tac aaa gct gag aag Lys Gln Arg Ile Ser Gln Glu Met Met Ala Leu Tyr Lys Ala Glu Lys 405 410 415	1248
gtt aac ccg ctg ggc ggc tgc ttc ccg ctg ctg atc cag atg cca atc Val Asn Pro Leu Gly Gly Cys Phe Pro Leu Leu Ile Gln Met Pro Ile 420 425 430	1296
ttc ctg gcg ttg tac tac atg ctg atg ggt tcc gtt gaa ctg cgt cag Phe Leu Ala Leu Tyr Tyr Met Leu Met Gly Ser Val Glu Leu Arg Gln 435 440 445	1344
gca ccg ttt gca ctg tgg atc cac gac ctg tcg gca cag gac ccg tac Ala Pro Phe Ala Leu Trp Ile His Asp Leu Ser Ala Gln Asp Pro Tyr 450 455 460	1392
tac atc ctg ccg atc ctg atg ggc gta acg atg ttc ttc att cag aag Tyr Ile Leu Pro Ile Leu Met Gly Val Thr Met Phe Phe Ile Gln Lys 465 470 475 480	1440
atg tcg ccg acc aca gtg acc gac ccg atg cag cag aag atc atg acc Met Ser Pro Thr Thr Val Thr Asp Pro Met Gln Gln Lys Ile Met Thr 485 490 495	1488
ttt atg ccg gtc atc ttc acc gtg ttc ttc ctg tgg ttc ccg tca ggt Phe Met Pro Val Ile Phe Thr Val Phe Phe Leu Trp Phe Pro Ser Gly 500 505 510	1536
ctg gtg ctg tac tat atc gtc agc aac ctg gta acc att att cag cag Leu Val Leu Tyr Tyr Ile Val Ser Asn Leu Val Thr Ile Ile Gln Gln 515 520 525	1584
cag ctg att tac cgt ggt ctg gaa aaa cgt ggc ctg cat agc cgc gag Gln Leu Ile Tyr Arg Gly Leu Glu Lys Arg Gly Leu His Ser Arg Glu 530 535 540	1632
aag aaa aaa tcc gaa ttc agc aaa gga gaa gaa ctt ttc act gga gtt Lys Lys Lys Ser Glu Phe Ser Lys Gly Glu Glu Leu Phe Thr Gly Val 545 550 555 560	1680
gtc cca att ctt gtt gaa tta gat ggt gat gtt aat ggg cac aaa ttt	1728

-continued

Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe 565 570 575	
tct gtc agt gga gag ggt gaa ggt gat gct aca tac gga aaa ctc acc Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr 580 585 590	1776
ctt aaa ttt att tgc act act gga aaa cta cct gtt cca tgg cca aca Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr 595 600 605	1824
ctt gtc act act ctg acc tat ggt gtt caa tgc ttt tcc cgt tat ccg Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro 610 615 620	1872
gat cac atg aaa cgg cat gac ttt ttc aag agt gcc atg ccc gaa ggt Asp His Met Lys Arg His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly 625 630 635 640	1920
tat gta cag gaa cgc act ata tct ttc aaa gat gac ggg aac tac aag Tyr Val Gln Glu Arg Thr Ile Ser Phe Lys Asp Asp Gly Asn Tyr Lys 645 650 655	1968
acg cgt gct gaa gtc aag ttt gaa ggt gat acc ctt gtt aat cgt atc Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile 660 665 670	2016
gag tta aag ggt att gat ttt aaa gaa gat gga aac att ctc gga cac Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His 675 680 685	2064
aaa cta gag tac aac tat aac tca cac aat gta tac atc acg gca gac Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Thr Ala Asp 690 695 700	2112
aaa caa aag aat gga atc aaa gct aac ttc aaa att cgc cac aac att Lys Gln Lys Asn Gly Ile Lys Ala Asn Phe Lys Ile Arg His Asn Ile 705 710 715 720	2160
gaa gat ggt tcc gtt caa cta gca gac cat tat caa caa aat act cca Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro 725 730 735	2208
att ggc gat ggc cct gtc ctt tta cca gac aac cat tac ctg tcg aca Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr 740 745 750	2256
caa tct gcc ctt tcg aaa gat ccc aac gaa aag cgt gac cac atg gtc Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val 755 760 765	2304
ctt ctt gag ttt gta act gct gct ggg att aca cat ggc atg gat gag Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu 770 775 780	2352
ctc tac aaa ttg ggt acc ctc gag cac cac cac cac cac tga Leu Tyr Lys Leu Gly Thr Leu Glu His His His His His His His 785 790 795	2397

<210> SEQ ID NO 69

<211> LENGTH: 798

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 69

Met Asp Ser Gln Arg Asn Leu Leu Val Ile Ala Leu Leu Phe Val Ser 1 5 10 15
--

Phe Met Ile Trp Gln Ala Trp Glu Gln Asp Lys Asn Pro Gln Pro Gln 20 25 30

Ala Gln Gln Thr Thr Gln Thr Thr Ala Ala Gly Ser Ala Ala

-continued

35	40	45
Asp Gln Gly Val Pro Ala Ser Gly Gln Gly Lys Leu Ile Ser Val Lys		
50	55	60
Thr Asp Val Leu Asp Leu Thr Ile Asn Thr Arg Gly Gly Asp Val Glu		
65	70	75
Gln Ala Leu Leu Pro Ala Tyr Pro Lys Glu Leu Asn Ser Thr Gln Pro		
85	90	95
Phe Gln Leu Leu Glu Thr Ser Pro Gln Phe Ile Tyr Gln Ala Gln Ser		
100	105	110
Gly Leu Thr Gly Arg Asp Gly Pro Asp Asn Pro Ala Asn Gly Pro Arg		
115	120	125
Pro Leu Tyr Asn Val Glu Lys Asp Ala Tyr Val Leu Ala Glu Gly Gln		
130	135	140
Asn Glu Leu Gln Val Pro Met Thr Tyr Thr Asp Ala Ala Gly Asn Thr		
145	150	155
Phe Thr Lys Thr Phe Val Leu Lys Arg Gly Asp Tyr Ala Val Asn Val		
165	170	175
Asn Tyr Asn Val Gln Asn Ala Gly Glu Lys Pro Leu Glu Ile Ser Ser		
180	185	190
Phe Gly Gln Leu Lys Gln Ser Ile Thr Leu Pro Pro His Leu Asp Thr		
195	200	205
Gly Ser Ser Asn Phe Ala Leu His Thr Phe Arg Gly Ala Ala Tyr Ser		
210	215	220
Thr Pro Asp Glu Lys Tyr Glu Lys Tyr Lys Phe Asp Thr Ile Ala Asp		
225	230	235
240		
Asn Glu Asn Leu Asn Ile Ser Ser Lys Gly Gly Trp Val Ala Met Leu		
245	250	255
Gln Gln Tyr Phe Ala Thr Ala Trp Ile Pro His Asn Asp Gly Thr Asn		
260	265	270
Asn Phe Tyr Thr Ala Asn Leu Gly Asn Gly Ile Ala Ala Ile Gly Tyr		
275	280	285
Lys Ser Gln Pro Val Leu Val Gln Pro Gly Gln Thr Gly Ala Met Asn		
290	295	300
Ser Thr Leu Trp Val Gly Pro Glu Ile Gln Asp Lys Met Ala Ala Val		
305	310	315
320		
Ala Pro His Leu Asp Leu Thr Val Asp Tyr Gly Trp Leu Trp Phe Ile		
325	330	335
Ser Gln Pro Leu Phe Lys Leu Leu Lys Trp Ile His Ser Phe Val Gly		
340	345	350
Asn Trp Gly Phe Ser Ile Ile Ile Thr Phe Ile Val Arg Gly Ile		
355	360	365
Met Tyr Pro Leu Thr Lys Ala Gln Tyr Thr Ser Met Ala Lys Met Arg		
370	375	380
Met Leu Gln Pro Lys Ile Gln Ala Met Arg Glu Arg Leu Gly Asp Asp		
385	390	395
400		
Lys Gln Arg Ile Ser Gln Glu Met Met Ala Leu Tyr Lys Ala Glu Lys		
405	410	415
Val Asn Pro Leu Gly Gly Cys Phe Pro Leu Leu Ile Gln Met Pro Ile		
420	425	430
Phe Leu Ala Leu Tyr Tyr Met Leu Met Gly Ser Val Glu Leu Arg Gln		
435	440	445

-continued

Ala Pro Phe Ala Leu Trp Ile His Asp Leu Ser Ala Gln Asp Pro Tyr
 450 455 460
 Tyr Ile Leu Pro Ile Leu Met Gly Val Thr Met Phe Phe Ile Gln Lys
 465 470 475 480
 Met Ser Pro Thr Thr Val Thr Asp Pro Met Gln Gln Lys Ile Met Thr
 485 490 495
 Phe Met Pro Val Ile Phe Thr Val Phe Phe Leu Trp Phe Pro Ser Gly
 500 505 510
 Leu Val Leu Tyr Tyr Ile Val Ser Asn Leu Val Thr Ile Ile Gln Gln
 515 520 525
 Gln Leu Ile Tyr Arg Gly Leu Glu Lys Arg Gly Leu His Ser Arg Glu
 530 535 540
 Lys Lys Lys Ser Glu Phe Ser Lys Gly Glu Glu Leu Phe Thr Gly Val
 545 550 555 560
 Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe
 565 570 575
 Ser Val Ser Gly Glu Gly Glu Asp Ala Thr Tyr Gly Lys Leu Thr
 580 585 590
 Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr
 595 600 605
 Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro
 610 615 620
 Asp His Met Lys Arg His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly
 625 630 635 640
 Tyr Val Gln Glu Arg Thr Ile Ser Phe Lys Asp Asp Gly Asn Tyr Lys
 645 650 655
 Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile
 660 665 670
 Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His
 675 680 685
 Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Thr Ala Asp
 690 695 700
 Lys Gln Lys Asn Gly Ile Lys Ala Asn Phe Lys Ile Arg His Asn Ile
 705 710 715 720
 Glu Asp Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro
 725 730 735
 Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr
 740 745 750
 Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val
 755 760 765
 Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu
 770 775 780
 Leu Tyr Lys Leu Gly Thr Leu Glu His His His His His His
 785 790 795

```

<210> SEQ ID NO 70
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: misc_feature
  
```

-continued

<222> LOCATION: (1)..(25)
<223> OTHER INFORMATION: hp6

<400> SEQUENCE: 70

tgcacatagga ggtccctccta tgtca	25
-------------------------------	----

<210> SEQ ID NO 71
<211> LENGTH: 861
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)..(861)
<223> OTHER INFORMATION: AmpR

<400> SEQUENCE: 71

atg tca att caa cat ttc cgt gtc gcc ctt att ccc ttt ttt gcg gca Met Ser Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Ala Ala 1 5 10 15	48
---	----

ttt tgc ctt cct gtt ttt gct cac cca gaa acg ctg gtg aaa gta aaa Phe Cys Leu Pro Val Phe Ala His Pro Glu Thr Leu Val Lys Val Lys 20 25 30	96
--	----

gat gct gaa gat cag ttg ggt gca cga gtg ggt tac atc gaa ctg gat Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp 35 40 45	144
--	-----

ctc aac agc ggt aag atc ctt gag agt ttt cgc ccc gaa gaa cgt ttt Leu Asn Ser Gly Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe 50 55 60	192
--	-----

cca atg atg agc act ttt aaa gtt ctg cta tgt ggc gcg gta tta tcc Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser 65 70 75 80	240
---	-----

cgt gtt gac gcc ggg caa gag caa ctc ggt cgc cgc ata cac tat tct Arg Val Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser 85 90 95	288
--	-----

cag aat gac ttg gtt gag tac tca cca gtc aca gaa aag cat ctt acg Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr 100 105 110	336
---	-----

gat ggc atg aca gta aga gaa tta tgc agt gct gcc ata acc atg agt Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser 115 120 125	384
---	-----

gat aac act gcg gcc aac tta ctt ctg aca acg atc gga gga ccg aag Asp Asn Thr Ala Ala Asn Leu Leu Thr Thr Ile Gly Gly Pro Lys 130 135 140	432
---	-----

gag cta acc gct ttt ttg cac aac atg ggg gat cat gta act cgc ctt Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu 145 150 155 160	480
---	-----

gat cgt tgg gaa ccg gag ctg aat gaa gcc ata cca aac gac gag cgt Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg 165 170 175	528
---	-----

gac acc acg atg cct gca gca atg gca aca acg ttg cgc aaa cta tta Asp Thr Thr Met Pro Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu 180 185 190	576
---	-----

act ggc gaa cta ctt act cta gct tcc cgg caa caa tta ata gac tgg Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp 195 200 205	624
---	-----

atg gag gcg gat aaa gtt gca gga cca ctt ctg cgc tcg gcc ctt ccg Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro 210 215 220	672
---	-----

-continued

```

gct ggc tgg ttt att gct gat aaa tct gga gcc ggt gag cgt ggg tct      720
Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser
225          230           235           240

cgc ggt atc att gca gca ctg ggg cca gat ggt aag ccc tcc cgt atc      768
Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile
245          250           255           255

gta gtt atc tac acg acg ggg agt cag gca act atg gat gaa cga aat      816
Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn
260          265           270           270

aga cag atc gct gag ata ggt gcc tca ctg att aag cat tgg taa      861
Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp
275          280           285           285

<210> SEQ ID NO 72
<211> LENGTH: 286
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 72

Met Ser Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala
1          5           10           15

Phe Cys Leu Pro Val Phe Ala His Pro Glu Thr Leu Val Lys Val Lys
20         25           30           30

Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp
35         40           45           45

Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe
50         55           60           60

Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser
65         70           75           80

Arg Val Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser
85         90           95           95

Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr
100        105          110          110

Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser
115        120          125          125

Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys
130        135          140          140

Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu
145        150          155          160

Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg
165        170          175          175

Asp Thr Thr Met Pro Ala Ala Met Ala Thr Thr Leu Arg Lys Leu Leu
180        185          190          190

Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp
195        200          205          205

Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro
210        215          220          220

Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser
225        230          235          240

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

Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn

```

-continued

260	265	270
Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp		
275	280	285
<210> SEQ ID NO 73		
<211> LENGTH: 6728		
<212> TYPE: DNA		
<213> ORGANISM: Artificial sequence		
<220> FEATURE:		
<223> OTHER INFORMATION: synthetic		
<220> FEATURE:		
<221> NAME/KEY: misc_feature		
<222> LOCATION: (1)..(6728)		
<223> OTHER INFORMATION: Plasmid pMax-truncV489		
<400> SEQUENCE: 73		
cgtatggcaa tgaaaagacgg tgagctggtg atatgggata gtgttcaccc ttgttacacc	60	
gttttccatg agcaaaactga aacgtttca tcgcctcggta gtgaataccca cgacgatttc	120	
cggcagtttc tacacatata ttcgcaagat gtggcggtt acggtaaaaa cctggcctat	180	
ttcccttaag ggttttatga gaatatgtt ttgcgtctcg ccaatccctg ggtgatgttc	240	
accagttttg atttaaacgt ggccaatatg gacaacttct tcgccccgtt tttcaccatg	300	
ggcaaatatt atacgcaagg cgacaagggtg ctgatgcccgc tggcgattca ggttcatcat	360	
gccgtctgtg atggcttcca tgtcggcaga atgcttaatg aattacaaca gtactgcgt	420	
gagtggcagg gcggggcgta attttttaa ggtagtttatt ggtgcctta aacgcctgg	480	
gctacgcctg aataagtgtat aataagcggta tgaatggcag aaattcggaaa gcaaattcga	540	
cccggtcgtc ggttcagggc agggtcgtta aatagccgt tatgtctatt gctggttac	600	
cggttttatgg actaccggaa gcagtgtgac cgtgtgcctc tcaaattgcgtt gaggccagtt	660	
tgctcaggat ctccccgtgg aggtataataat tgacgatatg atcattattt ctgcctccca	720	
gagcctgata aaaacggttt gcgcttcgtt aatacagatg taggtgttcc acagggttagc	780	
cagcagcattt ctgcgtatgca gatccggaaac ataatggcgtt agggcgcttg ttccggcggt	840	
ggtagatgtgg cagggcccggtt ggcggggggta ctgttggcgctt ctgcggcac ctgtccatc	900	
agttgcgtatgta taaagaagac agtcataatgtt gcccgcacgatg tagtcatgcc ccgcgcocac	960	
cgaaaggagc taccggacacag cgggtcggtac tgggtatcact cagaataaga aatggggccg	1020	
ctcatggcggtt tgactctcgtt tcatacgatgtc gtggatcactc cgggtgggttc cactcttgt	1080	
tgcggggcaac ttccagcagca cgttagggac ttccgcgtt ccagacttta cgaaacacgg	1140	
aaacccgaaga ccattcatgt tggctcgtt gtcgcagacg ttttgcagca gcagtcgtt	1200	
cacgttgcgtt cgcgtatcggt tgattcatc tgcataaccag taaggcaacc cccgcaggctt	1260	
agccgggtcc tcaacgcacag gagcacgtatc atgcgcaccc gtggccagga cccaaacgcgt	1320	
cccgagatgc gccgcgtgcg gctgtggag atggcgacg cgtatggatat gttctgcca	1380	
gggttggttt ggcgcattcac agtttcgcgc aagaattgtat tggctccat tcttggatgt	1440	
gtgaatccgt tagcgaggtt ccgcggctt ccattcaggat cggatggcc cggccatcg	1500	
caccgcgcacg caacgcgggg aggacacaa ggtataggcc ggcgcctaca atccatgc	1560	
acccgttcca tggctcgcc gaggccgcataa atacgcgtt gacgtatcgc ggtccagtgt	1620	
tgcgaagtttag gctggtaaga gcccgcacgc atccttgcgtt ctgtccctga tggctcgtcat	1680	
ctacccgttcca ggcacgcgtt ggcgcattcac cgggcgttcc gatgcgcgcgaa gaaaggagaa	1740	

-continued

gaatcataat	gggaaaggcc	atccagcctc	gcgtcgcaa	cgcaccaag	acgttagcca	1800
gcccgttggc	cgccatgcgg	gcgataatgg	cctgcttc	ccgaaacgt	ttggggcg	1860
gaccagtgtac	gaagggttga	gcgaggcg	gcaagattcc	gaataccgca	agcgacaggc	1920
cgatcatcg	cgcgctccag	cgaaagcggt	cctcgccgaa	aatgacccag	agcgctcg	1980
gcacctgtcc	tacgagttgc	atgataaaga	agacagtc	aagtgcggcg	acgatagtca	2040
tgccccgcgc	ccacccgaaag	gagctgactg	ggttgaaggc	tctcaaggc	atcggtcgac	2100
gtctccctt	atgcgactcc	tgcattagga	agcagcccag	tagtaggtt	aggccgttga	2160
gcaccgcgc	cgcaaggaat	ggtgcgtca	aggagatggc	gcccaacagt	ccccggcca	2220
cggggctgc	caccataccc	acgcccggaa	aagcgctcat	gagccgaaag	tggcgagccc	2280
gatctcccc	atcggtgtat	tcggcgat	aggcgccagc	aaccgcac	gtggcgccgg	2340
tgatgcggc	cacgatcg	ccggcgtaga	ggatccgggt	cccccttgc	agattaaaaa	2400
ggaaaggagg	aaagaataa	tggctcggt	acagttaaa	caacgtgaat	ctactgacgc	2460
aatctttgtt	cactgctcg	ctaccaagcc	aagtca	gttgggttcc	gtgagatcg	2520
ccagtggcac	aaagagcagg	gttggctcg	tgtggatac	cactttatca	tcaagcgaga	2580
cggtactgtg	gaggcgaggac	gagatgagat	ggctgttaggc	tctcacgc	agggttacaa	2640
ccacaactct	atcggtgtct	gccttgg	ttgtatcgac	gataaaggta	agttcgacgc	2700
taactttacg	ccagccaaa	tgcaatccct	tcgctcact	cttgcacac	tgctggctaa	2760
gtacgaaggc	gctgtgttcc	gcccacatca	tgagggtggcg	ccgaaggctt	gcccttcgtt	2820
cgacctaag	cggtgggggg	agaagaacga	actggta	tctgaccgt	gataaacccgc	2880
tgagcaataa	ctagcataac	cccttggggc	ctctaaacgg	gtcttgagg	gtttttgtt	2940
aattaagcgg	ccgcggcg	cccgctcgac	cgcgccggc	tgaggatacc	tcagccacca	3000
caattcagca	aattgtgaac	atcatcacgt	tcatcttcc	ctgggttgc	atggcccatt	3060
ttccctgtcg	taacgagaag	gtcgcgaa	caggcg	tttagactgt	cgtatgaaa	3120
ttcttttaa	gaaggagact	atatatgaaa	gcaattttcg	tactgaaaca	tcttaatcat	3180
gctaaggagg	ttttctaatg	aaaagaatgt	taatcaacgc	aactcagcag	gaagatgtc	3240
gcgttgcct	tgttagatgg	cagcgtctgt	atgacctgg	tatcgaaagt	ccagggcacg	3300
agcagaaaaaa	ggcaaaacatc	tacaaaggta	aaatcacccg	cattgaac	agtctggaa	3360
ctgctttgt	tgattacggc	gctgaacgtc	acggtttct	cccactaaa	gaaattgccc	3420
gcaaatattt	ccctgttaac	tacagtgtc	atggcg	caacattaaa	gatgtgttgc	3480
gtgaaggta	ggaagtcatt	gttcagatcg	ataaagaaga	gcccggca	aaaggcg	3540
cattaaccac	ctttatcgt	ctgggggt	gctatctgt	tctgatgc	aacaacccgc	3600
gcgcgggtgg	catttctcg	cgtatcgaa	gcgacgacc	taccgaat	aaagaagcac	3660
tggcaaggct	tgaactcg	gaaggcatgg	ggcttatcg	gcccac	ggcg	3720
aatctgtga	ggcgctgca	tgggatttaa	gcttccgt	gaaacactgg	gaagccatca	3780
aaaaagccgc	tgaaagccgc	ccggccccgt	tcctgattca	tcaggagac	aacgtatcg	3840
ttcgccgatt	ccgcgttac	ttacgtcagg	acatcg	aatccttac	gataacccga	3900
aagtgcgtca	actggcacgt	cagcatatcg	ctgcattag	tcgcccggat	ttcagcagca	3960
aaatcaaact	gtacacccgc	gagatcccgc	tgttcagcca	ctaccagatc	gagtcacaga	4020

-continued

tcgagtcgc cttccagcgt gaagttcgto tgecgctgg tggttccatt gttatcgaca	4080
gcaccgaagc gttaacggcc atcgacatca actccgcacg cgcgaccgcg ggccggcgata	4140
tcgaagaaac cgcggttaac actaacctcg aagctgccga tgagattgtc cgtagctgc	4200
gcctgcgtga cctcgccgc ctgattgttca aacttcat cgacatgacg ccagtacgcc	4260
accagcgtgc ggtagaaaac cgtagtgcgt aagcggtgcg tcaggaccgt gcgcgtattc	4320
aaatcagcca tatttctcgc tttggcctgc tgaaatgtc ccgtcagcgc ctgagcccat	4380
cactgggtga atccagccat cacgtctgcg cgcgtctgc cggtagccgt accgtgcgtg	4440
acaacgaatc gctgtcgctc tctattctgc gtctgatcga agaagaagcg ctgaaagaga	4500
acaccccagga agttcacgac attgttcctg tgccaatcgc ttcttatctg ctgaatgaaa	4560
aacgttctgc ggtgaatgcc attgaaacgcg gtcaggacgg tgtagtgcgtc gtgattgtc	4620
caaacgatca gatggaaacc ccgcactacc acgtgctgcg cgtgagggtt cctagaaatc	4680
atcccttagcg aaagctaagg attttttttta tctgttaatta attgaactca ctaaaggag	4740
accacagcggg tttcccttttgc ttgcattgg aggtcaaaata atgcgcagaat cttataaaca	4800
attctataag gctccgagga ggcataatcca agtgtggag gcagccatg ggccggatcc	4860
acaggacggg tgtggtcgatc atgatcgtcgt agtgcataatg ggctccaaatg agcgaagcga	4920
gcaggactgg gccccggcca aagcggtcg acagtgcctc gagaacgggt gcgcataaaaa	4980
attgcataatc cgcatatagc gctagcagca cgccatagtg actggcgatg ctgtcgaaat	5040
ggacgatatac ccgcataaggccc cccggcagta ccggcataac caagcctatg cctacagcat	5100
ccagggtgac ggtggccgagg atgacgatga ggcattgtt agatttcata cacgggtgcct	5160
gactgcgtta gcaatttaac tgtgataaac taccgcattaa aagcttatcg atgataagct	5220
gtcaaacatg agaattacaa cttatatcg atggggctga cttaggtgc tacattgaa	5280
gagataaatt gcaactgaaat cttagaaatat tttatctgat taataagatg atcttctgaa	5340
gatcgttttgc gtctgcgtc aatcttgcgc tctgaaaacg aaaaaaccgc cttgcagggc	5400
ggttttcga aggttctctg agctaccaac tctttgaacc gaggttaactg gcttggagga	5460
gcgcagtcac caaaacttgt ccttcgtt tagccttaac cggcgcatga cttcaagact	5520
aactcctcta aatcaattac cagtggtcg tgcgtggat gctttgcatt gtcttcgg	5580
gttggactca agacgatagt taccggataa ggccgcggcc tcggactgaa cgggggggttc	5640
gtgcatacag tccagcttgc agcgaactgc ctacccggaa ctgagtgca ggcgtggaaat	5700
gagacaaacg cggccataac agcggaaatga caccggtaaa ccgaaaggca ggaacaggag	5760
agcgcacgag ggagccgcca gggaaacgc ctggatctt tatagtccctg tcgggtttcg	5820
ccaccactga tttagcgatc agatctcgatc atgctgtca gggggggggaa gcctatggaa	5880
aaacggctt gccgcggccc tctcaatcc ctgttaagta tcttcgtgc atcttcagg	5940
aaatctccgc cccgttcgtatc agccatttcc gctcgccgcg gtcgaacgcg cgagcgtac	6000
gagtcgtga gcgaggaagc ggaatataatc ctgtatcaca tattctgctg acgcaccgg	6060
gcagcctttt ttcttcgtcc acatgaagca ctcaactgac accctcatca gtgccaacat	6120
agtaageccag tatacactcc gctagcgctg atgtccggcg gtgcgtttgc cgtagcgcac	6180
caccccgatca gtagctgaac aggagggaca gctgatagaa acagaagccca ctggagcacc	6240
tcaaaaaacac catcatacac taaatcgtatc agtggcgcg acatccgcgac gcactttgcg	6300

-continued

ccgaataaat	acctgtgacg	gaagatcaact	tcgcagaataa	aataaatcct	ggtgtccctg	6360											
ttgataccgg	gaagccctgg	gccaaactttt	ggcgaaaatg	agacgttgat	cggcacgtaa	6420											
gagggttcaa	ctttccccat	aatgaaataa	gatcaactacc	gggcgtattt	ttttagttat	6480											
cgagatttc	aggagctaaag	gaagctaaaa	tggagaaaaaa	aatcaactgga	tataaccacgg	6540											
ttgatatac	ccaatggcat	cgtaaaagaac	attttgggc	atttcagtca	gttgcataat	6600											
gtacctataa	ccagaccgtt	cagctggata	ttacggcctt	tttaaagacc	gtaaagaaaa	6660											
ataaggcaca	gttttatccg	gcctttattc	acattcttc	ccgcctgatg	aatgctcatc	6720											
cggaattc						6728											
 <210> SEQ ID NO 74																	
<211> LENGTH: 1467																	
<212> TYPE: DNA																	
<213> ORGANISM: Artificial sequence																	
<220> FEATURE:																	
<223> OTHER INFORMATION: synthetic																	
<220> FEATURE:																	
<221> NAME/KEY: CDS																	
<222> LOCATION: (1)..(1467)																	
<223> OTHER INFORMATION: rneD346N truncated at V489																	
 <400> SEQUENCE: 74																	
atg	aaa	aga	atg	tta	atc	aac	gca	act	cag	cag	gaa	gag	ttg	cgc	gtt	48	
Met		Lys	Arg	Met	Leu	Ile	Asn	Ala	Thr	Gln	Gln	Glu	Glu	Leu	Arg	Val	
1				5					10					15			
gcc	ctt	gtt	gat	ggg	cag	cgt	ctg	tat	gac	ctg	gat	atc	gaa	agt	cca	96	
Ala																	
Leu	Val	Asp	Gly	Gln	Arg	Leu	Tyr	Asp	Leu	Asp	Ile	Glu	Ser	Pro			
20									25					30			
ggg	cac	gag	cag	aaa	aag	gca	aac	atc	tac	aaa	ggt	aaa	atc	acc	cgc	144	
Gly																	
His	Glu	Gln	Lys	Lys	Ala	Asn	Ile	Tyr	Lys	Gly	Lys	Ile	Thr	Arg			
35									40					45			
att	gaa	ccg	agt	ctg	gaa	gct	gtt	gat	tac	ggc	gct	gaa	cgt		192		
Ile																	
Glu	Pro	Ser	Leu	Glu	Ala	Ala	Phe	Val	Asp	Tyr	Gly	Ala	Glu	Arg			
50									55					60			
cac	ggt	tcc	ctc	cca	cta	aaa	gaa	att	gcc	cgc	gaa	tat	ttc	cct	gct	240	
His	Gly	Phe	Leu	Pro	Leu	Lys	Glu	Ile	Ala	Arg	Glu	Tyr	Phe	Pro	Ala		
65									70					75		80	
aac	tac	agt	ctt	ccc	ttt	aaa	gat	gtt	ttg	cgt	gaa					288	
Asn	Tyr	Ser	Ala	His	Gly	Arg	Pro	Asn	Ile	Lys	Asp	Val	Leu	Arg	Glu		
85									90					95			
ggt	cag	gaa	gtc	att	gtt	cag	atc	gat	aaa	gaa	gag	cgc	ggc	aac	aaa	336	
Gly	Gln	Glu	Val	Ile	Val	Gln	Ile	Asp	Lys	Glu	Glu	Arg	Gly	Asn	Lys		
100									105					110			
ggc	gcg	gca	tta	acc	acc	ttt	atc	agt	ctg	gcg	ggt	agc	tat	ctg	gtt	384	
Gly	Ala	Ala	Leu	Thr	Thr	Phe	Ile	Ser	Leu	Ala	Gly	Ser	Tyr	Leu	Val		
115									120					125			
ctg	atg	ccg	aac	aac	ccg	cgc	ggt	ggc	att	tct	cgc	cgt	atc	gaa		432	
Leu	Met	Pro	Asn	Asn	Pro	Arg	Ala	Gly	Gly	Ile	Ser	Arg	Arg	Ile	Glu		
130									135					140			
ggc	gac	gac	cgt	acc	gaa	tta	aaa	gaa	gca	ctg	gca	agc	ctt	gaa	ctg	480	
Gly	Asp	Asp	Arg	Thr	Glu	Leu	Lys	Glu	Ala	Leu	Ala	Ser	Leu	Glu	Leu		
145									150					155		160	
ccg	gaa	ggc	atg	ggg	ctt	atc	gtg	cgc	acc	gct	ggc	gtc	ggc	aaa	tct	528	
Pro	Glu	Gly	Met	Gly	Leu	Ile	Val	Arg	Thr	Ala	Gly	Val	Gly	Lys	Ser		
165									170					175			
gct	gag	gag	ctg	caa	tgg	gat	tta	agc	tcc	cgt	ctg	aaa	cac	tgg	gaa	576	

-continued

Ala Glu Ala Leu Gln Trp Asp Leu Ser Phe Arg Leu Lys His Trp Glu			
180	185	190	
gcc atc aaa aaa gcc gct gaa agc cgc ccg gcc ccg ttc ctg att cat		624	
Ala Ile Lys Lys Ala Ala Glu Ser Arg Pro Ala Pro Phe Leu Ile His			
195	200	205	
cag gag agc aac gta atc gtt cgc gca ttc cgc gat tac tta cgt cag		672	
Gln Glu Ser Asn Val Ile Val Arg Ala Phe Arg Asp Tyr Leu Arg Gln			
210	215	220	
gac atc ggc gaa atc ctt atc gat aac ccg aaa gtg ctc gaa ctg gca		720	
Asp Ile Gly Glu Ile Leu Ile Asp Asn Pro Lys Val Leu Glu Leu Ala			
225	230	235	240
cgt cag cat atc gct gca tta ggt cgc ccg gat ttc agc agc aaa atc		768	
Arg Gln His Ile Ala Ala Leu Gly Arg Pro Asp Phe Ser Ser Lys Ile			
245	250	255	
aaa ctg tac acc ggc gag atc ccg ctg ttc agc cac tac cag atc gag		816	
Lys Leu Tyr Thr Gly Glu Ile Pro Leu Phe Ser His Tyr Gln Ile Glu			
260	265	270	
tca cag atc gag tcc gcc ttc cag cgt gaa gtt cgt ctg ccg tct ggt		864	
Ser Gln Ile Glu Ser Ala Phe Gln Arg Glu Val Arg Leu Pro Ser Gly			
275	280	285	
ggg tcc att gtt atc gac agc acc gaa gcg tta acg gcc atc gac atc		912	
Gly Ser Ile Val Ile Asp Ser Thr Glu Ala Leu Thr Ala Ile Asp Ile			
290	295	300	
aac tcc gca cgc gcg acc cgc ggc ggc gat atc gaa gaa acc gcg ttt		960	
Asn Ser Ala Arg Ala Thr Arg Gly Gly Asp Ile Glu Glu Thr Ala Phe			
305	310	315	320
aac act aac ctc gaa gct gcc gat gag att gct cgt cag ctg cgc ctg		1008	
Asn Thr Asn Leu Glu Ala Ala Asp Glu Ile Ala Arg Gln Leu Arg Leu			
325	330	335	
cgt gac ctc ggc ctg att gtt atc aac ttc atc gac atg acg cca		1056	
Arg Asp Leu Gly Leu Ile Val Ile Asn Phe Ile Asp Met Thr Pro			
340	345	350	
gta cgc cac cag cgt gcg gta gaa aac cgt ctg cgt gaa gcg gtt cgt		1104	
Val Arg His Gln Arg Ala Val Glu Asn Arg Leu Arg Glu Ala Val Arg			
355	360	365	
cag gag cgt gcg cgt att caa atc agc cat att tct cgc ttt ggc ctg		1152	
Gln Asp Arg Ala Arg Ile Gln Ile Ser His Ile Ser Arg Phe Gly Leu			
370	375	380	
ctg gaa atg tcc cgt cag cgc ctg agc cca tca ctg ggt gaa tcc agc		1200	
Leu Glu Met Ser Arg Gln Arg Leu Ser Pro Ser Leu Gly Glu Ser Ser			
385	390	395	400
cat cac gtc tgc ccg cgc tgc tcc ggt acc ggt acc gtg cgt gac aac		1248	
His His Val Cys Pro Arg Cys Ser Gly Thr Gly Thr Val Arg Asp Asn			
405	410	415	
gaa tcg ctg tcg ctc tct att ctg cgt ctg atc gaa gaa gaa gcg ctg		1296	
Glu Ser Leu Ser Leu Ser Ile Leu Arg Leu Ile Glu Glu Ala Leu			
420	425	430	
aaa gag aac acc cag gaa gtt cac gcc att gtt cct gtg cca atc gct		1344	
Lys Glu Asn Thr Gln Glu Val His Ala Ile Val Pro Val Pro Ile Ala			
435	440	445	
tct tat ctg ctg aat gaa aaa cgt tct gcg gtg aat gcc att gaa acg		1392	
Ser Tyr Leu Leu Asn Glu Lys Arg Ser Ala Val Asn Ala Ile Glu Thr			
450	455	460	
cgt cag gac ggt gtt cgc tgc gtg att gtg cca aac gat cag atg gaa		1440	
Arg Gln Asp Gly Val Arg Cys Val Ile Val Pro Asn Asp Gln Met Glu			
465	470	475	480
acc ccg cac tac cac gtg ctg cgc gtg		1467	

-continued

Thr Pro His Tyr His Val Leu Arg Val
485

<210> SEQ ID NO 75
<211> LENGTH: 489
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 75

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

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

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

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

His Gly Phe Leu Pro Leu Lys Glu Ile Ala Arg Glu Tyr Phe Pro Ala
65 70 75 80

Asn Tyr Ser Ala His Gly Arg Pro Asn Ile Lys Asp Val Leu Arg Glu
85 90 95

Gly Gln Glu Val Ile Val Gln Ile Asp Lys Glu Glu Arg Gly Asn Lys
100 105 110

Gly Ala Ala Leu Thr Thr Phe Ile Ser Leu Ala Gly Ser Tyr Leu Val
115 120 125

Leu Met Pro Asn Asn Pro Arg Ala Gly Ile Ser Arg Arg Ile Glu
130 135 140

Gly Asp Asp Arg Thr Glu Leu Lys Glu Ala Leu Ala Ser Leu Glu Leu
145 150 155 160

Pro Glu Gly Met Gly Leu Ile Val Arg Thr Ala Gly Val Gly Lys Ser
165 170 175

Ala Glu Ala Leu Gln Trp Asp Leu Ser Phe Arg Leu Lys His Trp Glu
180 185 190

Ala Ile Lys Lys Ala Ala Glu Ser Arg Pro Ala Pro Phe Leu Ile His
195 200 205

Gln Glu Ser Asn Val Ile Val Arg Ala Phe Arg Asp Tyr Leu Arg Gln
210 215 220

Asp Ile Gly Glu Ile Leu Ile Asp Asn Pro Lys Val Leu Glu Leu Ala
225 230 235 240

Arg Gln His Ile Ala Ala Leu Gly Arg Pro Asp Phe Ser Ser Lys Ile
245 250 255

Lys Leu Tyr Thr Gly Glu Ile Pro Leu Phe Ser His Tyr Gln Ile Glu
260 265 270

Ser Gln Ile Glu Ser Ala Phe Gln Arg Glu Val Arg Leu Pro Ser Gly
275 280 285

Gly Ser Ile Val Ile Asp Ser Thr Glu Ala Leu Thr Ala Ile Asp Ile
290 295 300

Asn Ser Ala Arg Ala Thr Arg Gly Gly Asp Ile Glu Glu Thr Ala Phe
305 310 315 320

Asn Thr Asn Leu Glu Ala Ala Asp Glu Ile Ala Arg Gln Leu Arg Leu
325 330 335

-continued

Arg	Asp	Leu	Gly	Gly	Leu	Ile	Val	Ile	Asn	Phe	Ile	Asp	Met	Thr	Pro	
340																350
Val	Arg	His	Gln	Arg	Ala	Val	Glu	Asn	Arg	Leu	Arg	Glu	Ala	Val	Arg	
355																365
Gln	Asp	Arg	Ala	Arg	Ile	Gln	Ile	Ser	His	Ile	Ser	Arg	Phe	Gly	Leu	
370																380
Leu	Glu	Met	Ser	Arg	Gln	Arg	Leu	Ser	Pro	Ser	Leu	Gly	Glu	Ser	Ser	
385																400
His	His	Val	Cys	Pro	Arg	Cys	Ser	Gly	Thr	Gly	Thr	Val	Arg	Asp	Asn	
405																415
Glu	Ser	Leu	Ser	Leu	Ser	Ile	Leu	Arg	Leu	Ile	Glu	Glu	Ala	Leu		
420																430
Lys	Glu	Asn	Thr	Gln	Glu	Val	His	Ala	Ile	Val	Pro	Val	Pro	Ile	Ala	
435																445
Ser	Tyr	Leu	Leu	Asn	Glu	Lys	Arg	Ser	Ala	Val	Asn	Ala	Ile	Glu	Thr	
450																460
Arg	Gln	Asp	Gly	Val	Arg	Cys	Val	Ile	Val	Pro	Asn	Asp	Gln	Met	Glu	
465																480
Thr	Pro	His	Tyr	His	Val	Leu	Arg	Val								
																485

```

<210> SEQ ID NO 76
<211> LENGTH: 6848
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)..(6848)
<223> OTHER INFORMATION: Plasmid pMax-truncL529

<400> SEQUENCE: 76

cgtatggcaa tgaaagacgg tgagctggtg atatggata gtgttcaccc ttgttacacc 60
gttttccatg agcaaactga aacgtttca tcgctctgga gtgaataccca cgacgattc 120
ccgcagttc tacacatata ttcgcaagat gtggcgtgtt acggtaaaaa cctggcttat 180
ttccctaaag ggtttattga gaatatgtt ttcgtcttag ccaatccctg ggtgagttc 240
accagtttg atttaaacgt ggccaatatg gacaacttct tcgccccctg tttcaccatg 300
ggcaaataatt atacgcaagg cgacaagggtg ctgatgccgc tggcgattca gttcatcat 360
gccgtctgtg atggcttcca tgcggcaga atgcttaatg aattacaaca gtactcgat 420
gagtggcagg gggggcgta attttttaa ggcagttatt ggtgcctta aacgccttgt 480
gctacgcctg aataagtgtat aataagcggaa tgaatggcag aaattcgaaa gcaaattcga 540
cccggtcgta ggttcaggggc agggtcgtta aatagccgt tatgtctatt gctggttac 600
cggtttattg actaccggaa gcagtgtgac cgtgtgttc tcaaattgcct gaggccagtt 660
tggtcagggt ctccccgtgg aggtataat tgacgatatg atcattttt ctgcctccca 720
gagcctgata aaaacggta ggcgttcgtt aatacagatg taggtgttcc acagggttagc 780
cagcagcatc ctgcgtatgca gatccggaaat ataatggtc agggcgctg tttcggcg 840
ggtatggtg caggccccgt ggccggggga ctgttggcg ctgcccgcac ctgtcctacg 900
agttgcatga taaagaagac agtcataagt gcggcgacga tagtcatgcc ccgcgcccac 960

```

-continued

cggaaggagc taccggacag cggtgcggac tttttaact cagaataaga aatgaggccg	1020
ctcatggcgt tgacttcag tcatagtata gtggtatcac cggttggttc cactcttgt	1080
tgcgggcaac ttccagcagca cgttagggac ttccgcgtt ccagacttta cgaaacacgg	1140
aaaccgaaga ccattcatgt ttttgcgtac gtcgcagacg ttttgcagca gcagtcgctt	1200
cacgttcgtc cgcgtatcgg tgattcattc tgctaaccag taaggcaacc cggccagcct	1260
agccgggtcc tcaacgacag gageacgatc atgcgcaccc gtggccagga cccaacgctg	1320
cccgagatgc gccgcgtgcg gctgtggag atggcggacg cgtatggatat gttctgcca	1380
gggttggttt ggcattcac agtttccgc aagaattgtat tggctccaat tcttggagt	1440
gtgaatccgt tagcgagggtg ccgcgggttt ccattcaggt cgagggtggcc cggctccatg	1500
cacccgcgacg caacgcgggg aggccagacaa ggtatagggc ggcgcctaca atccatgcca	1560
acccgttcca tgtgtcgcc gagggggcat aaatcgccgt gacgatcagc ggtccagtga	1620
tcgaagttag gctggtaaga gccgcggagcg atccattgaag ctgtccctga tggctgtcat	1680
ctacctgcgtt ggacacgcgt gcctgcacacg cgggcacccc gatgcgcgcg gaagcgagaa	1740
gaatcataat ggggaaggcc atccagccctc gctgcgcgaa cgccagcaag acgtacgcca	1800
gccccgtggc cgccatgcgc gcgataatgg cctgcgttcc gccgaaacgt ttgggtggcg	1860
gaccagtgtac gagggttgc gcgagggtgt gcaagattcc gaataccgca agcgacaggc	1920
cgatcatcggt cgcgctccag cgaaaacggcgt cctgcgcgaa aatgacccag acgcgtgccg	1980
gcacccgttcc tacgagttgc atgataaaga agacagtcat aagtgcggcg acgatagtca	2040
tgcggccgcg ccacccggaaag gagctgactg ggttgaaggc tctcaaggcc atcggtcgac	2100
gctctccctt atgcgactcc tgcatttagga acaagccag tagtaggttgc aggccgttgc	2160
gcaccgcgcg cgcacccaaat ggtgcgtca aggagatggc gccaacagt ccccccggcca	2220
cggggccgtc caccataccc acggccaaac aagcgctcat gagccgcgaa tggcgagccc	2280
gatcttcccc atcggtgtat tggcgatat aggccgcgcg aaccgcaccc tggcgccgg	2340
tgtgcggc cacgtgcgtt ccggcgtaga ggatccgggtt ccccttgcgtt agattaaaa	2400
ggaaaggagg aaagaaataa tggctcggtt acagttttaaa caacgtgaat ctactgacgc	2460
aatctttgtt cactgtcggtt ctaccaagecc aagtcagaat gttgggttcc gtgagatcc	2520
ccagtggcgc aaagagcagg gttggctcgat tggggatacc cactttatca tcaacgcgaga	2580
cggtaactgtt gaggcaggac gagatgagat ggctgttaggc tctcacgcata agggttacaa	2640
ccacaactct atcggtgttcc gccttgggg tggatcgac gataaaggta agttcgacgc	2700
taactttacg ccagcccaa tgcaatccct tcgcgtcactt ctgttcacac tgctggctaa	2760
gtacgaaggc gctgtgttcc ggcgcgcata tgagggtggcc cggaaaggctt gcccggctt	2820
cgaccccttaag cggttgggg agaagaacgaa actgggtact tctgaccgtt gataaaccgc	2880
tgagcaataa ctacgtataac cccttgggg ctctaaacgg gtcttgagggtt gttttttgtt	2940
aattaagcgg cggcgccgcg cccgtcgacg cggcgccggc tgaggatacc tcaacgcacca	3000
caattcagca aatttgcgtt acatcacgt tcatcttccttcc ctgggtgcac atggccatt	3060
ttcctgtcg taacgagaag gtcgcgttcc caggcgctttt tagactgtt cgtaatgaaa	3120
ttctttttaa gaaggagact atatatgaaa gcaatttcg tactgaaaca tcttaatcat	3180
gctaaggagg ttttctaatg aaaagaatgt taatcaacgc aactcagcag gaagagttgc	3240

-continued

gcgttgcctt	tgtatgggg	cagcgctgt	atgacacctgga	tatcgaaagt	ccaggcaca	3300
agcagaaaaa	ggcaaaccatc	tacaaaggta	aatcaccccg	cattgaaccg	agtctggaa	3360
ctgcttttgt	tgattacggc	gctgaacgtc	acggtttcc	cccactaaa	gaaattgccc	3420
gcgaatattt	ccctgcta	ac tacagtgc	atggtcgtcc	caacattaa	gatgtgtgc	3480
gtgaaggta	ggaagtcat	gtttagatcg	ataaagaaga	gcgcggca	aaaggcgcgg	3540
cattaaccac	ctttatca	ctgggggta	gctatctgg	tctgatgc	aacaaccgc	3600
gcgcgggtgg	catttcgc	cgtatcgaa	gcgcaccc	taccgaat	aaagaagcac	3660
tggcaagect	tgaactgccc	gaaggcatgg	ggcttacgt	gcgcaccc	ggcgtcgca	3720
aatctgtga	ggcgctgaa	tgggattaa	gcttcgt	gaaacactgg	gaagccatca	3780
aaaaagccgc	tgaaagccgc	ccggccccgt	tcctgattca	tcaggagagc	aacgtatcg	3840
ttcgcgccatt	ccgcgattac	ttacgtcagg	acatcgccgaa	aatccttata	gataaccgc	3900
aagtgtcgaa	actggcactgt	cagcatatcg	ctgcattagg	tgcggcgat	ttcagcagca	3960
aaatcaaact	gtacacccgc	gagatcccgc	tgttgcgca	ctaccagatc	gagtacacaga	4020
tcgagtcgc	cttccagcgt	gaagttcg	tgcgtctgg	tggttccatt	gttatcgaca	4080
gcaccgaagc	gttaacggcc	atcgacatca	actccgcac	cgcgacccgc	ggcggcgata	4140
tcaagaaac	cgcgat	actaacctcg	aagctgccc	ttagattgt	cgtcagctgc	4200
gcctgcgt	cctcgccgc	ctgattgtt	tcaacttcat	cgacatgac	ccagtacgcc	4260
accagcgtgc	ggtagaaaac	cgtctgcgt	aagcggtgc	tcaggac	gcgcgtattc	4320
aaatcagcca	tatttcgc	tttggcctgc	tggaaatgtc	ccgtcagcgc	ctgagccat	4380
cactgggtga	atccagccat	cacgtctgc	cgcgtgetc	cggtacc	accgtcg	4440
acaacgaatc	gtcgctgc	tctattctgc	gtctgatcg	agaagaagcg	ctgaaagaga	4500
acacccagga	atgtcacgc	attgttctgc	tgcataatcg	ttcttatctg	ctgaatgaaa	4560
aacgttctgc	ggtgaatg	ccgcataatcg	gtcaggac	tgttgcgtc	gtgattgtc	4620
caaacgatca	gatggaaacc	ccgcactacc	acgtgctgc	cgtgcgtaaa	gggaaagaaa	4680
cgccgac	ttt aagctatcg	ctgcgcg	tgcataatcg	agcgtggcg	ctgcgtctg	4740
aagaagagtt	cgctgacgt	aagegtccgg	aacaacctgc	gctgagg	tta	4800
atccttagc	aaagctatcg	attttttta	tctgtat	attgaaactca	ctaaaggag	4860
accacagcc	tttcccttgc	ttcgattgg	aggtaata	atgcgtca	cttataaaca	4920
attctataag	gtcccgagga	ggcatatca	agtgtggag	gcagccat	ggccggatcc	4980
acaggacgg	tgtggcg	atgatcg	agtcgat	ggctca	agtcgac	5040
gcaggactgg	ggggggcc	aagcggtcg	acgtgctc	gagaacgggt	gcgcata	5100
attgcata	ccgcata	gctagcagca	cgcata	cgat	ctgtcgaa	5160
ggacgat	ccgc	aaagg	ccggcag	ccggcata	caagcctat	5220
ccagggtgac	ggtgc	ggagg	atgacgat	gcgcatt	atgat	5280
gactgcgtt	gcaat	ttaaac	tgtgata	taacc	cgat	5340
gtcaaaatc	agaat	ccat	ttatcg	atgggg	cttgcagg	5400
gagataaaat	gactgaaat	ctagaaat	tttatcgat	taataa	agat	5460
gatcg	tttg	gtctgcgcgt	aatcttgc	tctgaaaac	aaaaaaccgc	5520

-continued

gggttttcga	aggttctctg	agctaccAAC	tcttGAACC	gaggtaACTG	gttggagGA	5580
gcccAGTCAC	caaaaCTTGT	cTTTCAGTT	tagCCTTAAC	cgGCATG	CTTCAGACT	5640
aactccTCTA	aatcaATTAC	cAGTGGCTGC	TGCCAGTGTG	GTTCATG	GTCTTCCGG	5700
gttggactCA	agacgatAGT	TACCGATAA	GGCGCAGCG	TCGGACTGAA	CGGGGGGTTC	5760
gtgcatacAG	TCCAGCTTG	AGCAGACTGC	CTACCCGGAA	CTGAGTGTCA	GGCGTGGAA	5820
gagacaAAACG	CGGCCATAAC	AGCGGAATGA	ACCGGTAAA	CGAAAGGCA	GGAAACAGGAG	5880
agcgcacGAG	GGAGCCGCA	GGGGAAACGC	CTGGTATCTT	TATAGTCCTG	TGGGTTTCG	5940
ccaccactGA	TTTGAGCGTC	AGATTCGTG	ATGCTTGTCA	GGGGGGCGGA	GCCTATGGAA	6000
aaacggTTT	GGCGGGCCC	TCTCCTTC	CTGTAGTA	TCTTCCTGC	ATCTCCAGG	6060
aaatCTCCGC	CCCGTTCGTA	AGCCATTCC	GTCGCCGCA	GTGAAACGAC	CGAGCGTAGC	6120
gagtcaGTG	GCAGGAAAGC	GGAAATATATC	CTGTATACAA	TATTCTGTG	ACGCACCGGT	6180
gcagcTTT	TTCTCTGCC	ACATGAAGCA	CTTCACTGAC	ACCCTCATCA	GTGCCAACAT	6240
agtaAGCCAG	TATACACTCC	GCTAGCGTC	ATGTCGGCG	GTGCTTTC	CGTTACGCAC	6300
cacccCGTC	GTAGCTGAAC	AGGAGGGACA	GCTGATAGAA	ACAGAACCCA	CTGGAGCACC	6360
tcaaaaACAC	CATCATAACAC	TAATCAGTA	AGTGGCAGC	ATCACCCGAC	GCACCTTGC	6420
ccgaataAAAT	ACCTGTGACG	GAAGATCACT	TCGCAGAAATA	AATAAATCCT	GGTGTCCCTG	6480
ttgataccGG	GAAGCCCTGG	GCCAACTTT	GGCGAAAATG	AGACGTTGAT	CGGCACGTAA	6540
gaggTTCCAA	CTTTCCACAT	AATGAAATAA	GATCACTACC	GGGCGTATT	TTTGAGTT	6600
cgagatTTTC	AGGAGCTAAG	GAAGCTAAAA	TGGAGAAAAAA	AATCACTGGA	TATACCAACG	6660
ttgatataTC	CCAATGGCAT	CgtAAAGAAC	ATTTGAGGC	ATTCAGTCA	GTGCTCAAT	6720
gtacctaAA	CCAGACC GTT	CAGCTGGATA	TTACGGCCTT	TTAAAGACC	GTAAAGAAAA	6780
ataagcacAA	GTTTTATCCG	GCCTTATT	ACATTCTGC	CGCCTGTG	AATGTCATC	6840
ccqaattc						6848

```
<210> SEQ ID NO 77
<211> LENGTH: 1587
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: CDS
```

<223> OTHER INFORMATION

<400> SEQUENCE: 77
atg aaa aga atg tta atc aac gca act cag cag gaa gag ttg cgc gtt 48
Met Lys Arg Met Leu Ile Asn Ala Thr Gln Gln Glu Glu Leu Arg Val

```
gcc ctt gta gat ggg cag cgt ctg tat gac ctg gat atc gaa agt cca 96  
Ala Leu Val Asp Gly Gln Arg Leu Tyr Asp Leu Asp Ile Glu Ser Pro  
20 35 50 65 80 95
```

ggg cac gag cag aaa aag gca aac atc tac aaa ggt aaa atc acc cgc 144
 Gly His Glu Gln Lys Lys Ala Asn Ile Tyr Lys Gly Lys Ile Thr Arg
 25 40 45

```

att gaa ccg agt ctg gaa gct gct ttt gtt gat tac ggc gct gaa cgt      192
Ile Glu Pro Ser Leu Ala Ala Phe Val Asp Tyr Gly Ala Glu Arg
      50          55          60

```

-continued

cac ggt ttc ctc cca cta aaa gaa att gcc cgcc gaa tat ttc cct gct		240
His Gly Phe Leu Pro Leu Lys Glu Ile Ala Arg Glu Tyr Phe Pro Ala		
65 70 75 80		
aac tac agt cat ggt cgt ccc aac att aaa gat gtg ttg cgt gaa		288
Asn Tyr Ser Ala His Gly Arg Pro Asn Ile Lys Asp Val Leu Arg Glu		
85 90 95		
ggc gac gac gtc att gtt cag atc gat aaa gaa gag cgcc ggc aac aaa		336
Gly Gln Glu Val Ile Val Gln Ile Asp Lys Glu Glu Arg Gly Asn Lys		
100 105 110		
ggc gcg gca tta acc acc ttt atc agt ctg gcg ggt agc tat ctg gtt		384
Gly Ala Ala Leu Thr Thr Phe Ile Ser Leu Ala Gly Ser Tyr Leu Val		
115 120 125		
ctg atg ccg aac aac ccg cgcc ggc att tct cgc cgt atc gaa		432
Leu Met Pro Asn Asn Pro Arg Ala Gly Gly Ile Ser Arg Arg Ile Glu		
130 135 140		
ggc gac gac cgt acc gaa tta aaa gaa gca ctg gca agc ctt gaa ctg		480
Gly Asp Asp Arg Thr Glu Leu Lys Glu Ala Leu Ala Ser Leu Glu Leu		
145 150 155 160		
ccg gaa ggc atg ggg ctt atc gtg ccg acc gct ggc gtc ggc aaa tct		528
Pro Glu Gly Met Gly Leu Ile Val Arg Thr Ala Gly Val Gly Lys Ser		
165 170 175		
gct gag gcg ctg caa tgg gat tta agc ttc cgt ctg aaa cac tgg gaa		576
Ala Glu Ala Leu Gln Trp Asp Leu Ser Phe Arg Leu Lys His Trp Glu		
180 185 190		
gcc atc aaa aaa gcc gct gaa agc ccg ccg gcc ttc ctg att cat		624
Ala Ile Lys Lys Ala Ala Glu Ser Arg Pro Ala Pro Phe Leu Ile His		
195 200 205		
cag gag agc aac gta atc gtt ccg gca ttc ccg gat tac tta cgt cag		672
Gln Glu Ser Asn Val Ile Val Arg Ala Phe Arg Asp Tyr Leu Arg Gln		
210 215 220		
gac atc ggc gaa atc ctt atc gat aac ccg aaa gtg ctc gaa ctg gca		720
Asp Ile Gly Glu Ile Leu Ile Asp Asn Pro Lys Val Leu Glu Leu Ala		
225 230 235 240		
cgt cag cat atc gct gca tta ggt ccg ccg gat ttc agc agc aaa atc		768
Arg Gln His Ile Ala Ala Leu Gly Arg Pro Asp Phe Ser Ser Lys Ile		
245 250 255		
aaa ctg tac acc ggc gag atc ccg ctg ttc agc cac tac cag atc gag		816
Lys Leu Tyr Thr Gly Glu Ile Pro Leu Phe Ser His Tyr Gln Ile Glu		
260 265 270		
tca cag atc gag tcc gcc ttc cag cgt gaa gtt cgt ctg ccg tct ggt		864
Ser Gln Ile Glu Ser Ala Phe Gln Arg Glu Val Arg Leu Pro Ser Gly		
275 280 285		
ggc tcc att gtt atc gac agc acc gaa ggc tta acg gcc atc gac atc		912
Gly Ser Ile Val Ile Asp Ser Thr Glu Ala Leu Thr Ala Ile Asp Ile		
290 295 300		
aac tcc gca ccg ggc acc ccg ggc ggc gat atc gaa gaa acc ggc ttt		960
Asn Ser Ala Arg Ala Thr Arg Gly Gly Asp Ile Glu Glu Thr Ala Phe		
305 310 315 320		
aac act aac ctc gaa gct gcc gat gag att gct cgt cag ctg ccg ctg		1008
Asn Thr Asn Leu Glu Ala Ala Asp Glu Ile Ala Arg Gln Leu Arg Leu		
325 330 335		
cgt gac ctc ggc ggc ctg att gtt atc aac ttc atc gac atg acg cca		1056
Arg Asp Leu Gly Gly Leu Ile Val Ile Asn Phe Ile Asp Met Thr Pro		
340 345 350		
gta ccg cac cag cgt gca gaa aac cgt ctg cgt gaa ggc gtg cgt		1104
Val Arg His Gln Arg Ala Val Glu Asn Arg Leu Arg Glu Ala Val Arg		
355 360 365		

-continued

cag gac cgt gcg cgt att caa atc agc cat att tct cgc ttt ggc ctg Gln Asp Arg Ala Arg Ile Gln Ile Ser His Ile Ser Arg Phe Gly Leu 370 375 380	1152
ctg gaa atg tcc cgt cag cgc ctg agc cca tca ctg ggt gaa tcc agc Leu Glu Met Ser Arg Gln Arg Leu Ser Pro Ser Leu Gly Glu Ser Ser 385 390 395 400	1200
cat cac gtc tgc ccg cgc tgc tcc ggt acc ggt acc gtg cgt gac aac His His Val Cys Pro Arg Cys Ser Gly Thr Gly Thr Val Arg Asp Asn 405 410 415	1248
gaa tcg ctg tcg ctc tct att ctg cgt ctg atc gaa gaa gaa gcg ctg Glu Ser Leu Ser Leu Ser Ile Leu Arg Leu Ile Glu Glu Ala Leu 420 425 430	1296
aaa gag aac acc cag gaa gtt cac gcc att gtt cct gtg cca atc gct Lys Glu Asn Thr Gln Glu Val His Ala Ile Val Pro Val Pro Ile Ala 435 440 445	1344
tct tat ctg ctg aat gaa aaa cgt tct gcg gtg aat gcc att gaa acg Ser Tyr Leu Leu Asn Glu Lys Arg Ser Ala Val Asn Ala Ile Glu Thr 450 455 460	1392
cgt cag gac ggt gtt cgc tgc gtg att gtg cca aac gat cag atg gaa Arg Gln Asp Gly Val Arg Cys Val Ile Val Pro Asn Asp Gln Met Glu 465 470 475 480	1440
acc ccg cac tac cac gtg ctg cgc gtg cgt aaa ggg gaa gaa acg ccg Thr Pro His Tyr His Val Leu Arg Val Arg Lys Gly Glu Glu Thr Pro 485 490 495	1488
acc tta agc tac atg ctg ccg aag ctg cat gaa gaa gcg atg gcg ctg Thr Leu Ser Tyr Met Leu Pro Lys Leu His Glu Glu Ala Met Ala Leu 500 505 510	1536
ccg tct gaa gaa gag ttc gct gaa cgt aag cgt ccg gaa caa cct gcg Pro Ser Glu Glu Glu Phe Ala Glu Arg Lys Arg Pro Glu Gln Pro Ala 515 520 525	1584
ctg Leu	1587
<210> SEQ ID NO 78 <211> LENGTH: 529 <212> TYPE: PRT <213> ORGANISM: Artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Synthetic Construct	
<400> SEQUENCE: 78	
Met Lys Arg Met Leu Ile Asn Ala Thr Gln Gln Glu Leu Arg Val 1 5 10 15	
Ala Leu Val Asp Gly Gln Arg Leu Tyr Asp Leu Asp Ile Glu Ser Pro 20 25 30	
Gly His Glu Gln Lys Lys Ala Asn Ile Tyr Lys Gly Lys Ile Thr Arg 35 40 45	
Ile Glu Pro Ser Leu Glu Ala Ala Phe Val Asp Tyr Gly Ala Glu Arg 50 55 60	
His Gly Phe Leu Pro Leu Lys Glu Ile Ala Arg Glu Tyr Phe Pro Ala 65 70 75 80	
Asn Tyr Ser Ala His Gly Arg Pro Asn Ile Lys Asp Val Leu Arg Glu 85 90 95	
Gly Gln Glu Val Ile Val Gln Ile Asp Lys Glu Glu Arg Gly Asn Lys 100 105 110	
Gly Ala Ala Leu Thr Thr Phe Ile Ser Leu Ala Gly Ser Tyr Leu Val 115 120 125	

-continued

Leu	Met	Pro	Asn	Asn	Pro	Arg	Ala	Gly	Gly	Ile	Ser	Arg	Arg	Ile	Glu
130															140
Gly	Asp	Asp	Arg	Thr	Glu	Leu	Lys	Glu	Ala	Leu	Ala	Ser	Leu	Glu	Leu
145															160
Pro	Glu	Gly	Met	Gly	Leu	Ile	Val	Arg	Thr	Ala	Gly	Val	Gly	Lys	Ser
															165
															170
															175
Ala	Glu	Ala	Leu	Gln	Trp	Asp	Leu	Ser	Phe	Arg	Leu	Lys	His	Trp	Glu
															180
															185
															190
Ala	Ile	Lys	Lys	Ala	Ala	Glu	Ser	Arg	Pro	Ala	Pro	Phe	Leu	Ile	His
															195
															200
															205
Gln	Glu	Ser	Asn	Val	Ile	Val	Arg	Ala	Phe	Arg	Asp	Tyr	Leu	Arg	Gln
															210
															215
															220
Asp	Ile	Gly	Glu	Ile	Leu	Ile	Asp	Asn	Pro	Lys	Val	Leu	Glu	Leu	Ala
															225
															230
															235
															240
Arg	Gln	His	Ile	Ala	Ala	Leu	Gly	Arg	Pro	Asp	Phe	Ser	Ser	Lys	Ile
															245
															250
															255
Lys	Leu	Tyr	Thr	Gly	Glu	Ile	Pro	Leu	Phe	Ser	His	Tyr	Gln	Ile	Glu
															260
															265
															270
Ser	Gln	Ile	Glu	Ser	Ala	Phe	Gln	Arg	Glu	Val	Arg	Leu	Pro	Ser	Gly
															275
															280
															285
Gly	Ser	Ile	Val	Ile	Asp	Ser	Thr	Glu	Ala	Leu	Thr	Ala	Ile	Asp	Ile
															290
															295
															300
Asn	Ser	Ala	Arg	Ala	Thr	Arg	Gly	Gly	Asp	Ile	Glu	Glu	Thr	Ala	Phe
															305
															310
															315
															320
Asn	Thr	Asn	Leu	Glu	Ala	Ala	Asp	Glu	Ile	Ala	Arg	Gln	Leu	Arg	Leu
															325
															330
															335
Arg	Asp	Leu	Gly	Gly	Leu	Ile	Val	Ile	Asn	Phe	Ile	Asp	Met	Thr	Pro
															340
															345
															350
Val	Arg	His	Gln	Arg	Ala	Val	Glu	Asn	Arg	Leu	Arg	Glu	Ala	Val	Arg
															355
															360
															365
Gln	Asp	Arg	Ala	Arg	Ile	Gln	Ile	Ser	His	Ile	Ser	Arg	Phe	Gly	Leu
															370
															375
															380
Leu	Glu	Met	Ser	Arg	Gln	Arg	Leu	Ser	Pro	Ser	Leu	Gly	Ser	Ser	
															385
															390
															395
															400
His	His	Val	Cys	Pro	Arg	Cys	Ser	Gly	Thr	Gly	Thr	Val	Arg	Asp	Asn
															405
															410
															415
Glu	Ser	Leu	Ser	Leu	Ser	Ile	Leu	Arg	Leu	Ile	Glu	Glu	Ala	Leu	
															420
															425
															430
Lys	Glu	Asn	Thr	Gln	Glu	Val	His	Ala	Ile	Val	Pro	Val	Pro	Ile	Ala
															435
															440
															445
Ser	Tyr	Leu	Leu	Asn	Glu	Lys	Arg	Ser	Ala	Val	Asn	Ala	Ile	Glu	Thr
															450
															455
															460
Arg	Gln	Asp	Gly	Val	Arg	Cys	Val	Ile	Val	Pro	Asn	Asp	Gln	Met	Glu
															465
															470
															475
															480
Thr	Pro	His	Tyr	His	Val	Leu	Arg	Val	Arg	Lys	Gly	Glu	Glu	Thr	Pro
															485
															490
															495
Thr	Leu	Ser	Tyr	Met	Leu	Pro	Lys	Leu	His	Glu	Glu	Ala	Met	Ala	Leu
															500
															505
															510
Pro	Ser	Glu	Glu	Glu	Phe	Ala	Glu	Arg	Lys	Arg	Pro	Glu	Gln	Pro	Ala
															515
															520
															525

Leu

-continued

```
<210> SEQ ID NO 79
<211> LENGTH: 8447
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)..(8447)
<223> OTHER INFORMATION: Plasmid pMax-D346X, wherein NNN may be TTT,
CTG, AAT, GTG, AGC, CCG, ACC, GCG, TAT, CAT, CAG, AAA, GAA,
TGC, or GGC

<400> SEQUENCE: 79

cgtatggcaa tgaaaagacgg tgagctggtg atatggata gtgttcaccc ttgttacacc      60
gttttccatg agcaaactga aacgtttca tcgcctctgg gtgaataccca cgacgattc      120
cggcagttc tacacatata ttgcagaatgt gtggcggtt acggtaaaaa cctggccat      180
ttccctaaag ggtttattga gaatatgttt ttcgtctcag ccaatccctg ggttagttc      240
accagtttg atttaaacgt ggccaatatg gacaacttct tcggcccccgt tttcaccatg      300
ggcaaatatt atacgcaagg cgacaagggt ctgatgccgc tggcgattca ggttcatcat      360
gccgtctgtg atggcttcca tgcggcaga atgcttaatg aattacaaca gtactgcgat      420
gagtggcagg gcggggcgta attttttaa ggcagttatt ggtgcctta aacgccttgt      480
gctacgcctg aataagtgtat aataagcggta tgaatggcag aaattcggaaa gcaaattcga      540
cccggtcgct gggtcagggc agggtcgtt aatagccgt tatgtctatt gctggttac      600
cggtttattg actaccggaa gcagtggtac cgtgtgttcc tcaaattgcgtt gaggccagtt      660
tgctcaggct ctccccgtgg aggtataat tgacgatatg atcattttt ctgcctccca      720
gagcctgata aaaacgggtt ggcgttcgtt aatacagatg taggtgttcc acagggttagc      780
cagcagcatc ctgcgtatgc gatccggaa ataatggtgc agggcgctt tttcgccgtg      840
ggtatggtgg caggccccgt ggccggggga ctgttggcg ctgcggcac ctgtcctacg      900
agttgcgtatg taaagaagac agtcataatg gcccgcacgt tagtcatgcc ccgcgcacac      960
cggaaggagc taccggacag cgggtggac tggtaact cagaataaga aatggggccg      1020
ctcatggcgatg tgactctcag tcatagatgc gtggatcac cggttgggtc cactctctgt      1080
tgccggcaac ttccagcgtca cgttagggac ttccgcgtt ccagacttta cgaaacacgg      1140
aaaccggaa ccattcatgt tggtaact gtcgcacgt tttgcacgc gcagtcgtt      1200
cacgttcgtc cgccgtatcgg tgattcatc tgcataaccgg taaggcaacc ccgcgcaccc      1260
agccgggtcc tcaacgcacag gagcacgtc atgcgcaccc gtggccagga cccaaacgcgt      1320
cccgagatgc ggcgtgtcg gctgtggag atggcggacg cgatggatat gttctgcacca      1380
gggttgggtt ggcattcac agttccgc aagaattgtat tggctccat tcttggagtg      1440
gtgaatccgt tagcgaggtt ccattcgatg cgagggtggcc cggctccatg      1500
caccggcgtc caacgcgggg aggcaacaa ggtatagggc ggcgcctaca atccatgcca      1560
acccgttcca tgcgtcgcc gagggggcat aaatgcgcgt gacgtacgc ggtccagtga      1620
tgcgtgttag gctggtaaga gcccggatcg atccgttgcgat ctgtccctga tggtcgtcat      1680
ctacactgcgt ggacagcatg gcctgcacgc cggccatccc gatgcgcgcg gaagcgagaa      1740
gaatcataat gggaaaggcc atccagccctc ggcgtcgccaa cgccagcaag acgtgcacca      1800
```

-continued

ggcgcgtggc	cgccatgccg	gctgataatgg	cctgcttctc	gcccggaaacgt	ttgggtggcg	1860
gaccagtgc	gaagggttga	gctggggcgt	gcaagattcc	gaataccgca	agcgacaggc	1920
cgatcatcg	cgcgctccag	cgaaagcggt	cctcgccgaa	aatgaccag	agcgctggc	1980
gcacctgtcc	tacgagttgc	atgataaaga	agacagtcat	aagtggcg	acgatagtca	2040
tgccccgcgc	ccaccggaaag	gagctgactg	ggttgaaggc	tctcaagggc	atcggtcgac	2100
gctctccctt	atgcgactcc	tgcatttagga	agcagcccg	tagtaggtg	aggccgttga	2160
gcaccgcgc	cgcaaggaat	ggtgcatgc	aggagatggc	gcccacagt	ccccggcca	2220
cggggcgtgc	caccataccc	acgcccggaaac	aacgcgtcat	gagccggaaag	tggcgagccc	2280
gatctcccc	atcggtgatg	tcggcgatat	aggcgccagc	aaccgcac	gtggcgccgg	2340
tgatgcccgc	cacgatcggt	ccggcgtaga	ggatccgggt	cccccttgat	agattaaaaa	2400
ggaaaggagg	aaagaataa	tggctcggt	acagttaaa	caacgtgaat	ctactgacgc	2460
aatctttgtt	cactgctcg	ctaccaagec	aagtcaaat	gttgggttcc	gtgagattcg	2520
ccagtggcac	aaagagcagg	gttggctcga	tgtggataac	cactttatca	tcaagcgaga	2580
cggtaactgtg	gaggcaggac	gagatgagat	ggctgttaggc	tctcacgcta	agggttacaa	2640
ccacaactct	atcggcgct	gccttgggg	tggtatcgac	gataaaggta	agttcgacgc	2700
taactttacg	ccagccaaa	tgcaatccct	tcgctcaact	cttgcacac	tgctggctaa	2760
gtacgaaggc	gctgtgttcc	gcccggatca	tgaggtggcg	ccgaaggctt	gcccttcgtt	2820
cgaccttaag	cgttgggtgg	agaagaacga	actggtaact	tctgaccgtg	gataaaccgc	2880
tgagcaataa	ctagcataac	cccttggggc	ctctaaacgg	gtcttgggg	gtttttgtt	2940
aattaagcgg	ccgcggcg	cccgctcgac	cgcgccggc	tgaggatacc	tcagccacca	3000
caattcagca	aattgtgaac	atcatcacgt	tcatcttcc	ctgggttgc	atggcccatt	3060
ttctgtcag	taacgagaag	gtcgcaatt	caggcgctt	ttagactggt	cgtaatgaaa	3120
ttcttttaa	aaaggagact	atatatgaaa	gcaatttcg	tactgaaaca	tcttaatcat	3180
gtaaggagg	ttttctaatg	aaaagaatgt	taatcaacgc	aactcagcag	gaagagtgc	3240
gcgttgcct	tgtatgtgg	cagcgctgt	atgacctgga	tatcgaaagt	ccaggccacg	3300
agcagaaaaa	ggcaaaccatc	tacaaaggta	aaatcacccg	cattgaaccg	agtctggaaag	3360
ctgcttttgt	tgattacggc	gctgaacgtc	acggtttcc	cccactaaaa	gaaattgccc	3420
gcgaatat	ccctgcta	tacagtgc	atggcgttcc	caacattaaa	gatgtgtgc	3480
gtgaagggtca	ggaagtccatt	gttcagatcg	ataaagaaga	gcccggcaac	aaaggccgg	3540
cattaaaccac	ctttatcagt	ctgggggt	gctatctgg	tctgatgc	acaacccgc	3600
gcgcgggtgg	catttcgc	cgtatcgaa	gcgcacgacc	taccgaat	aaagaagcac	3660
tggcaaggct	tgaactgccc	gaaggcatgg	ggcttatcg	gcccggct	ggcgtcg	3720
aatctgtca	ggcgctgca	ttggattaa	gcttcgtt	gaaacactgg	gaagccatca	3780
aaaaagccgc	tgaaagccgc	ccggccccgt	tcctgattca	tacggagagc	aacgtaatcg	3840
ttcgcgcat	ccgcgattac	ttacgtcagg	acatcgccg	aatccttac	gataaccgc	3900
aatgtgc	actggc	cacgtatcg	ctgcattagg	tgcggcgat	ttcagcagca	3960
aaatcaaact	gtacaccggc	gagatcccgc	tgttcagcca	ctaccagatc	gagtcacaga	4020
tcgagtcgc	cttccagcgt	gaagttcg	tgccgtctgg	tggttccatt	gttatcgaca	4080

-continued

gcaccgaaac	gttaacggcc	atcgacatca	actccgcacg	cgcgacccgc	ggggcgata	4140
tgcgatgttt	aaac	actaacctcg	aagctgccg	ttagattgt	cgtcagctgc	4200
gcctgcgtga	cctcgccgc	ctgattgtt	tcnnnttcat	cgacatgacg	ccagtacgcc	4260
accagcgtgc	ggtagaaaac	cgtctgcgt	aaggcggtgc	tcaaggaccgt	gcgcgtattc	4320
aaatcagcca	tatttctcgc	tttggcctgc	tggaaatgtc	ccgtcagcgc	ctgagccat	4380
cactgggtga	atccagccat	cacgtctgc	cgcgcgtc	cggttaccgg	accgtgcgt	4440
acaacgaatc	gctgtcgctc	tctattctgc	gtctgatcg	agaagaagcg	ctgaaagaga	4500
acacccagga	atgttccctg	tgccaatcg	ttcttatctg	ctgaatgaaa		4560
aacgttctgc	ggtgaatgcc	attgaaacgc	gtcaggacgg	tgttcgtgc	gtgattgtc	4620
caaacgatca	gatggaaacc	ccgcactacc	acgtgcgtgc	cgtgcgtaaa	gggaaagaaa	4680
cggccacctt	aagctacatg	ctgcccga	tgcatgaaga	agcgatggcg	ctgcccgtctg	4740
aagaagagtt	cgctgaacgt	aagcgccgg	aaacaacctgc	gctggcaacc	tttgcctatgc	4800
cggatgtgcc	ccagcgc	ccccagctg	aacctgcgc	ccctgtcgta	gccccagcac	4860
ctaaatctgc	acccgcaaca	ccagccgctc	ctgcccacc	tgggctgttgc	agccgcttct	4920
tcggcgcact	gaaagcgctg	ttcagcggtg	gtgaagaaac	caaaccgtcc	gagcaaccaa	4980
caccgaaagc	agaagcgaaa	ccggAACGTC	aacaggatcg	tgcgaagcct	cgtcagaaca	5040
accgcgcgtga	ccgtatgag	cgccgcgaca	cccgtagtga	acgtactgaa	ggcagcgtata	5100
atcgcgaaga	aaaccgtcgt	aatcgctgc	aggcacagca	gcagactgcc	gagacgcgt	5160
agagccgtca	cgagggtgag	gtAACGGAAA	aagcgctac	caccgcgag	cagcaagcgc	5220
cgcgtcgta	acgtagccgc	cgccgtatg	atgataaacc	tcaaggcgca	caagaagcga	5280
aggcgctgaa	tgttgaagag	caatctgttc	aggaaaccga	acaggaagaa	cgtgtacgtc	5340
cggttcagcc	cgctcgtaaa	cagcgctago	tcaatcagaa	agtgcgttac	gagcaagcg	5400
tagccgaaga	agcggtagtc	gcaccgggtgg	ttgaagaaac	tgcgtcgcc	gaaccaatttgc	5460
ttcaggaagc	ccagactcca	cgcacagaac	tggtggaaatgt	cccgctgcca	gtcgtagcgc	5520
aaactgcacc	agaacacgaa	gaagagaaca	atgctgataa	ccgtgacaac	ggtggcatgc	5580
cgcgtcggtc	tcggccgtcg	cctcgctacc	tgcgcgtaa	tggtcagcgt	cgtcgctcg	5640
atcgtgcgca	gcgttatcca	acccagtcgc	caatgccgtt	gaccgttagcg	tgcgcgtctc	5700
cggaaactggc	ctctggcaaa	gtctggatcc	gctatcca	tgtacgtccg	caagatgtac	5760
aggttgaaga	gcagcgcgaa	caggaagaag	tacatgtgc	ccgcgtgg	actgagggtcc	5820
ctgtcgccgc	cgctatcgaa	ccgggttgtt	gcgcgcagg	tgttgaagaa	gtggccgggt	5880
tcgttagaagg	ccccgttcag	gttgcgaaac	cgcaaccgg	agtgggtgaa	acgacgcata	5940
ctgaagtgtat	tgctgcgcgc	gttaactgaa	agccgcagg	gattaccgag	tctgtatgttgc	6000
ccgtagccca	ggaagttgca	gaacaaggcg	aaccgggtgg	tgaaccgcag	gaagagacgg	6060
cagatattga	agaagttgtc	gaaactgtcg	aggttgtatgt	tgctgaacct	gaagttgttgc	6120
ctcaacactgc	cgcgcaggta	gtcgctgaag	tgcgcgcaga	agttgaaacg	gtagctgcgg	6180
tgcgatgttt	ggtcaccgtt	gagcataacc	acgttaccgc	cccaatgacg	cgcgctccag	6240
caccggaaata	tgttccggag	gcaccgcgtc	acagtgcactg	gcagcgcct	actttgcct	6300
tcgaaggtaa	aggtgcgc	ggtggtcata	cgccaacaca	tcatgcctct	gctgtcctg	6360

-continued

cgcgtccgca acctgtttag taaaggttac ctagaaatca tccttagcga aagctaagga	6420
tttttttat ctgttaattaa ttgactcac taaaggggaga ccacagcggt ttcccttgt	6480
tcgcattgga ggtcaataa tgccaaagtc ttataaaca ttctataagg ctccgaggag	6540
gcataatecaa gtgtgggagg cagccatgg gccggatcca caggacgggt gtggtcgcca	6600
tgatcgcgta gtcgatagt gctccaagta gcgaagcgag caggactggg cggcggccaa	6660
agcggtegga cagtgtccg agaacgggtg cgcatagaaa ttgcatcaac gcatatagcg	6720
ctagcagcac gccatagtga ctggcgatgc tgtcggaatg gacgatatcc cgcaagaggc	6780
ccggcagtagc cggcataacc aagcctatgc ctacagcattc cagggtgacg gtgccgagga	6840
tgacgatgag cgcattgtta gattcatac acgggtgcctg actgcgttag caatttaact	6900
gtgataaact accgcattaa agcttatacga tgataagctg tcaaacatga gaattacaac	6960
ttatatacgta tggggctgac ttcaagggtct acatggaaatg agataaaatg cactgaaatc	7020
tagaaatatt ttatctgatt aataagatga tcttcggatg atcggtttgg tctgcgcgta	7080
atctcttgct ctgaaaacga aaaaaccgc ttgcaggggcg gttttcgaa ggttctctga	7140
gctaccactt ctttgaaccg aggtaactgg cttggaggag cgcatgtacc aaaacttgc	7200
ctttcagttt agccttaacc ggcgcgtac ttcaagacta actcctctaa atcaattacc	7260
agtggctgtg cccagtggtg cttttgcattg tctttccggg ttggactcaa gacgatagtt	7320
accggataag ggcgcagggt cggactgaac ggggggttcg tgcatacagt ccagcttgg	7380
gcgaactgcc tacccggAAC tgagtgtcag gcgtggaatg agacaaacgc ggccataaca	7440
gcggaaatgac accggtaaac cggaaaggcg gaacaggaga gcgcacgagg gagccgcag	7500
gggaaacgcc tggatcttt atagtcctgt cgggtttcgc caccactgat ttgagcgtca	7560
gatttcgtga tgcttgcag gggggcggag cctatggaaa aacggctttg ccgcggccct	7620
ctcaactccc tggatcttat cttctggca tcttcaggaa aatctccgc ccgttcgtaa	7680
gcatttcgc ctcgcgcag tcgaacgacc gagcgttagcg agtcagttag cgaggaaagcg	7740
gaatataatcc tggatcttat attctgtca cgcaccggc cagcctttt tctctgcca	7800
catgaagcac ttcaactgaca ccctcatcg tgccaaacata gtaagccagt atacactccg	7860
ctagcgtgtg tggccggcgg tgctttgcg gttacgcacc accccgtcag tagctgaaca	7920
ggagggacac ctgatagaaa cagaagccac tggagcaccat caaaaacacc atcataact	7980
aaatcgtaa gttggcagca tcacccgacg cactttgcgc cgaataaata cctgtgacgg	8040
aaagatcactt cgcagaataa ataaatcctgt gtgtccctgt tgataaccggg aagccctggg	8100
ccaaacttttgcg gacggaaatga gacgttgatc ggcacgtaaag aggttccaac tttcaccata	8160
atgaaataag atcaactaccg ggcgtatccc tttagtttgcg gagatccatc ggagctaagg	8220
aagctaaaat ggagaaaaaa atcaactggat ataccacccgt tgatataatcc caatggcatc	8280
gtaaagaaca ttttgaggca ttgcgtatcc ttgcgtcaatg tacctataac cagaccgttc	8340
agctggatatacgttccctt taaaagaccg taaaagaaaaaa taagcacaag ttttatccgg	8400
cctttattca cattcttgcg cgcctgtatca atgcgtatcc ggaattc	8447

<210> SEQ ID NO 80

<211> LENGTH: 1061

<212> TYPE: PRT

<213> ORGANISM: E. coli

-continued

<220> FEATURE:
 <221> NAME/KEY: PEPTIDE
 <222> LOCATION: (1)..(1061)
 <223> OTHER INFORMATION: Translation of rneD346X, wherien X is Phe(F),
 Leu(L), Ile(I), Val(V), Ser(S), Pro(P), Thr(T), Ala(A), Tyr(Y),
 His(H), Gln(Q), Lys(K), Glu(E), Cyc(C), or Gly(G).

<400> SEQUENCE: 80

Met	Lys	Arg	Met	Leu	Ile	Asn	Ala	Thr	Gln	Gln	Glu	Glu	Leu	Arg	Val
1				5				10			15				
Ala	Leu	Val	Asp	Gly	Gln	Arg	Leu	Tyr	Asp	Leu	Asp	Ile	Glu	Ser	Pro
	20					25			30						
Gly	His	Glu	Gln	Lys	Lys	Ala	Asn	Ile	Tyr	Lys	Gly	Lys	Ile	Thr	Arg
	35					40			45						
Ile	Glu	Pro	Ser	Leu	Glu	Ala	Ala	Phe	Val	Asp	Tyr	Gly	Ala	Glu	Arg
	50				55			60							
His	Gly	Phe	Leu	Pro	Leu	Lys	Glu	Ile	Ala	Arg	Glu	Tyr	Phe	Pro	Ala
	65				70			75		80					
Asn	Tyr	Ser	Ala	His	Gly	Arg	Pro	Asn	Ile	Lys	Asp	Val	Leu	Arg	Glu
	85					90				95					
Gly	Gln	Glu	Val	Ile	Val	Gln	Ile	Asp	Lys	Glu	Glu	Arg	Gly	Asn	Lys
	100					105			110						
Gly	Ala	Ala	Leu	Thr	Thr	Phe	Ile	Ser	Leu	Ala	Gly	Ser	Tyr	Leu	Val
	115				120			125							
Leu	Met	Pro	Asn	Asn	Pro	Arg	Ala	Gly	Gly	Ile	Ser	Arg	Arg	Ile	Glu
	130				135			140							
Gly	Asp	Asp	Arg	Thr	Glu	Leu	Lys	Glu	Ala	Leu	Ala	Ser	Leu	Glu	Leu
	145				150			155			160				
Pro	Glu	Gly	Met	Gly	Leu	Ile	Val	Arg	Thr	Ala	Gly	Val	Gly	Lys	Ser
	165				170			175							
Ala	Glu	Ala	Leu	Gln	Trp	Asp	Leu	Ser	Phe	Arg	Leu	Lys	His	Trp	Glu
	180					185			190						
Ala	Ile	Lys	Lys	Ala	Ala	Glu	Ser	Arg	Pro	Ala	Pro	Phe	Leu	Ile	His
	195				200			205							
Gln	Glu	Ser	Asn	Val	Ile	Val	Arg	Ala	Phe	Arg	Asp	Tyr	Leu	Arg	Gln
	210				215			220							
Asp	Ile	Gly	Glu	Ile	Leu	Ile	Asp	Asn	Pro	Lys	Val	Leu	Glu	Leu	Ala
	225				230			235			240				
Arg	Gln	His	Ile	Ala	Ala	Leu	Gly	Arg	Pro	Asp	Phe	Ser	Ser	Lys	Ile
	245				250			255							
Lys	Leu	Tyr	Thr	Gly	Glu	Ile	Pro	Leu	Phe	Ser	His	Tyr	Gln	Ile	Glu
	260				265			270							
Ser	Gln	Ile	Glu	Ser	Ala	Phe	Gln	Arg	Glu	Val	Arg	Leu	Pro	Ser	Gly
	275				280			285							
Gly	Ser	Ile	Val	Ile	Asp	Ser	Thr	Glu	Ala	Leu	Thr	Ala	Ile	Asp	Ile
	290				295			300							
Asn	Ser	Ala	Arg	Ala	Thr	Arg	Gly	Gly	Asp	Ile	Glu	Glu	Thr	Ala	Phe
	305				310			315			320				
Asn	Thr	Asn	Leu	Glu	Ala	Ala	Asp	Glu	Ile	Ala	Arg	Gln	Leu	Arg	Leu
	325				330			335							
Arg	Asp	Leu	Gly	Gly	Leu	Ile	Val	Ile	Xaa	Phe	Ile	Asp	Met	Thr	Pro
	340				345			350							
Val	Arg	His	Gln	Arg	Ala	Val	Glu	Asn	Arg	Leu	Arg	Glu	Ala	Val	Arg

-continued

355	360	365
Gln Asp Arg Ala Arg Ile Gln Ile Ser His Ile Ser Arg Phe Gly Leu		
370	375	380
Leu Glu Met Ser Arg Gln Arg Leu Ser Pro Ser Leu Gly Glu Ser Ser		
385	390	395
His His Val Cys Pro Arg Cys Ser Gly Thr Gly Thr Val Arg Asp Asn		
405	410	415
Glu Ser Leu Ser Leu Ser Ile Leu Arg Leu Ile Glu Glu Ala Leu		
420	425	430
Lys Glu Asn Thr Gln Glu Val His Ala Ile Val Pro Val Pro Ile Ala		
435	440	445
Ser Tyr Leu Leu Asn Glu Lys Arg Ser Ala Val Asn Ala Ile Glu Thr		
450	455	460
Arg Gln Asp Gly Val Arg Cys Val Ile Val Pro Asn Asp Gln Met Glu		
465	470	475
480		
Thr Pro His Tyr His Val Leu Arg Val Arg Lys Gly Glu Thr Pro		
485	490	495
Thr Leu Ser Tyr Met Leu Pro Lys Leu His Glu Glu Ala Met Ala Leu		
500	505	510
Pro Ser Glu Glu Glu Phe Ala Glu Arg Lys Arg Pro Glu Gln Pro Ala		
515	520	525
Leu Ala Thr Phe Ala Met Pro Asp Val Pro Pro Ala Pro Thr Pro Ala		
530	535	540
Glu Pro Ala Ala Pro Val Val Ala Pro Ala Pro Lys Ser Ala Pro Ala		
545	550	555
560		
Thr Pro Ala Ala Pro Ala Gln Pro Gly Leu Leu Ser Arg Phe Phe Gly		
565	570	575
Ala Leu Lys Ala Leu Phe Ser Gly Gly Glu Glu Thr Lys Pro Ser Glu		
580	585	590
Gln Pro Thr Pro Lys Ala Glu Ala Lys Pro Glu Arg Gln Gln Asp Arg		
595	600	605
Arg Lys Pro Arg Gln Asn Asn Arg Arg Asp Arg Asn Glu Arg Arg Asp		
610	615	620
Thr Arg Ser Glu Arg Thr Glu Gly Ser Asp Asn Arg Glu Asn Arg		
625	630	635
640		
Arg Asn Arg Arg Gln Ala Gln Gln Gln Thr Ala Glu Thr Arg Glu Ser		
645	650	655
Arg Gln Gln Val Glu Val Thr Glu Lys Ala Arg Thr Thr Asp Glu Gln		
660	665	670
Gln Ala Pro Arg Arg Glu Arg Ser Arg Arg Asn Asp Asp Lys Arg		
675	680	685
Gln Ala Gln Gln Glu Ala Lys Ala Leu Asn Val Glu Glu Gln Ser Val		
690	695	700
Gln Glu Thr Glu Gln Glu Glu Arg Val Arg Pro Val Gln Pro Arg Arg		
705	710	715
720		
Lys Gln Arg Gln Leu Asn Gln Lys Val Arg Tyr Glu Gln Ser Val Ala		
725	730	735
Glu Glu Ala Val Val Ala Pro Val Val Glu Glu Thr Val Ala Ala Glu		
740	745	750
Pro Ile Val Gln Glu Ala Pro Ala Pro Arg Thr Glu Leu Val Lys Val		
755	760	765

-continued

```

Pro Leu Pro Val Val Ala Gln Thr Ala Pro Glu Gln Gln Glu Glu Asn
770           775           780

Asn Ala Asp Asn Arg Asp Asn Gly Gly Met Pro Arg Arg Ser Arg Arg
785           790           795           800

Ser Pro Arg His Leu Arg Val Ser Gly Gln Arg Arg Arg Arg Tyr Arg
805           810           815

Asp Glu Arg Tyr Pro Thr Gln Ser Pro Met Pro Leu Thr Val Ala Cys
820           825           830

Ala Ser Pro Glu Leu Ala Ser Gly Lys Val Trp Ile Arg Tyr Pro Ile
835           840           845

Val Arg Pro Gln Asp Val Gln Val Glu Glu Gln Arg Glu Gln Glu Glu
850           855           860

Val His Val Gln Pro Met Val Thr Glu Val Pro Val Ala Ala Ala Ile
865           870           875           880

Glu Pro Val Val Ser Ala Pro Val Val Glu Glu Val Ala Gly Val Val
885           890           895

Glu Ala Pro Val Gln Val Ala Glu Pro Gln Pro Glu Val Val Glu Thr
900           905           910

Thr His Pro Glu Val Ile Ala Ala Ala Val Thr Glu Gln Pro Gln Val
915           920           925

Ile Thr Glu Ser Asp Val Ala Val Ala Gln Glu Val Ala Glu Gln Ala
930           935           940

Glu Pro Val Val Glu Pro Gln Glu Glu Thr Ala Asp Ile Glu Glu Val
945           950           955           960

Val Glu Thr Ala Glu Val Val Ala Glu Pro Glu Val Val Ala Gln
965           970           975

Pro Ala Ala Pro Val Val Ala Glu Val Ala Ala Glu Val Glu Thr Val
980           985           990

Ala Ala Val Glu Pro Glu Val Thr Val Glu His Asn His Ala Thr Ala
995           1000          1005

Pro Met Thr Arg Ala Pro Ala Pro Glu Tyr Val Pro Glu Ala Pro
1010          1015          1020

Arg His Ser Asp Trp Gln Arg Pro Thr Phe Ala Phe Glu Gly Lys
1025          1030          1035

Gly Ala Ala Gly Gly His Thr Ala Thr His His Ala Ser Ala Ala
1040          1045          1050

Pro Ala Arg Pro Gln Pro Val Glu
1055          1060

```

```

<210> SEQ ID NO 81
<211> LENGTH: 8447
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)..(8447)
<223> OTHER INFORMATION: pMax-rnewT

<400> SEQUENCE: 81

cgtatggcaa tgaaagacgg tgagctggtg atatggata gtgttcaccc ttgttacacc      60
gttttccatg agcaaactga aacgtttca tcgctctgga gtgaataccca cgacgattc     120

```

-continued

cggcagtttc tacacatata ttgcgaatgt gtggcggtt acggtaaaa cctggctat	180
ttccctaaag ggtttattga gaatatgtt ttcgtctca gcaatccctg ggtgatgttc	240
accagtttg atttaaacgt ggccaatatg gacaacttct tcgccccgt tttcaccatg	300
ggcaaataatt atacgcaagg cgacaagggt ctgatgccgc tggcgattca gggtcatcat	360
gccgtctgtg atggcttcca tgcggcaga atgcttaatg aattacaaca gtactgcgt	420
gagtggcagg gccccggta attttttaa ggagtttatt ggtgcccta aacgcctgg	480
gctacgcctg aataagtgtat aataagcggta tgaatggcag aaattcggaaa gcaaattcga	540
cccggtcgct gggtcagggc agggtcgta aatagccgt tatgtctatt gctggttac	600
cgggttattg actaccggaa gcagtgtgac cgtgtgttcc tcaaatgcct gagggcagtt	660
tgcgtcaggct ctccccgtgg aggtataat tgacgatatg atcatttatt ctgcctccca	720
gagcctgata aaaacggta ggcgttcgtt aatacagatg taggtgttcc acagggttagc	780
cagcagcatc ctgcgtatgca gatccggaaataatggtgc agggcgottt tttcgccgtg	840
ggtatggtgg caggccccgtt ggccgggggat ctgttggggc ctgcggcac ctgttctacg	900
agttgcgtga taaagaagac agtcataatgt gcccgcacgat tagtcatgcc ccgcgcacac	960
cggaaggagc taccggacag cgggtcggtac tggtaact cagaataaga aatggggcc	1020
ctcatggcgt tgactctca gtcatagtatc gtggatcac cgggtgggtc cactcttgt	1080
tgcggcaac ttccagcgtca cgtaggggac ttccgcgtt ccagacttta cgaaacacgg	1140
aaaccgaaga ccattcatgt tggatgtcgt gtcgcacgat ttttgcacgac gcaatgcgtt	1200
cacgttgcgt cgcgtatcggt gatccatcc tgcataaccg taaggcaacc ccgcgcac	1260
agccgggtcc tcaacgcacag gagcacgatc atgcgcaccc gtggccagga cccaaacgt	1320
cccgagatgc ggcgcgtgcg gctgtggag atggcggacg cgtggatat gttctgcacaa	1380
gggttgggtt ggcattcac agttccgcgaa aagaattgtat tggctccat tcttggagtg	1440
gtgaatccgt tagcgagggtg ccgcggctt ccattcaggat cgagggtggcc cggctccatg	1500
caccgcgacg caacgcgggg aggccacaa ggtataggc ggcgcctaca atccatgca	1560
acccgttcca tgtgtcgcc gagggcggcat aaatgcggat gacgatcgcg ggtccagtga	1620
tgcgtttag gtcgttataa gcccgcacgc atccgttgcgat ctgtccctga tggcgatcat	1680
ctacctgcgtt ggacagcatg gcgtgcacgc cggccatccc gatgcggccg gaagcgacaa	1740
gaatcataat ggggaaggcc atccagccctc ggcgtcgacaa cggccacaa acgtacccca	1800
gcccgcgtggc cgcgcgtgcg gcgataatgg cctgcgttcc gcccggatgtt ttggcgccg	1860
gaccagtgcgac gaaaggcttgc ggcggggcgt gcaagatcc gaaatccgcg acgcacaggc	1920
cgatcatcgatc cgcgcgtccatcg cggaaacgcgtt cctgcggccaa aatgcacccacg acgcgtgcgc	1980
gcacctgtcc tacgagttgc atgataaaga agacagtcat aatgcggccg acgtatgtca	2040
tgcggccgcgc ccacccggaaag gagctgactg ggttgaaggc tctcaaggcc atcggtcgac	2100
gtctccctt atgcgtactcc tgcatttagga agcagccacg tagtaggttgc gggcgatgt	2160
gcacccgcgcgc cgcacggaaat ggtgcgtca aggatggc gcccacacgt ccccccggca	2220
cgccgcgtgc caccatacc acggccaaac aacgcgtcat gagccggaaat tggcgacggcc	2280
gatcttcccc atcggtatgc tgcgtatgcgatc aggcgcacgc aacccgcaccc gtggcgccgg	2340
tgatgcggcc acgcgtgcgtt ccggcgatgc ggtatccgggtt ccccttgcgtt agataaaaaa	2400

-continued

ggaaaggagg aaagaataa tggctcggt acagttaaa caacgtaat ctactgacgc	2460
aatcttggta cactgctcggt ctaccaagcc aagtcaaat gttgggtgtcc gttagatcg	2520
ccagtggcac aaagagcagg gttggctga tgtgggatac cactttatca tcaagcgaga	2580
cggtaactgtg gaggcaggac gagatgagat ggctgttaggc tctcacgcta agggttacaa	2640
ccacaactct atccggctct gccttgggtgg tggtatcgac gataaaggta agttcgacgc	2700
taactttacg ccagccccaaa tgcaatccct tgcgtcaactg cttgtcacac tgctggctaa	2760
gtacgaaggc gctgtgtttc gcgcggcatca tgaggtggcg ccgaaggcgtt gcccttcgtt	2820
cgaccttaag cggtgggtggg agaagaacga actggtaact tctgaccgtg gataaaccgc	2880
tgagcaataa ctacgataac cccttggggc ctctaaacgg gtcttgggg gttttttgtt	2940
aattaagcgg ccgcggcgcg cccgtcgacg cgcgccggc tgaggatacc tcagccacca	3000
caattcagca aattgtgaac atcatcacgt tcatctttcc ctgggttgcac atggcccatt	3060
ttoctgtcaag taacgagaag gtcgogaatt caggcgcttt ttagactggt cgtaatgaaa	3120
ttctttttaa gaaggagact atatatgaaa gcaattttcg tactgaaaca tcttaatcat	3180
gtctaggagg ttttctaattg aaaagaatgt taatcaacgc aactcagcag gaagagtgc	3240
gegtgtccct ttagatggg cagcgctgt atgacctggta tatcgaaagt ccaggccacg	3300
agcagaaaaa ggcaaacatc tacaaggta aaatcaccccg cattgaaccg agtctggaaag	3360
ctgttttgt tgattacggc gctgaacgtc acggttccct cccactaaaa gaaattggcc	3420
gcgaatattt ccctgtcaac tacagtgcct atggctgtcc caacattaaa gatgtgtgc	3480
gtgaagggtca ggaagtccatt gttcagatcg ataaagaaga ggcggcaac aaaggcgg	3540
cattaaccac ctttatcagt ctgggggtta gctatctgg tctgatgcgg aacaaccgc	3600
gcgcgggtgg catttctcg cgtatcgaa ggcacgaccg taccgaattt aaagaagcac	3660
tggcaagccct tgaactgccg gaaggcatgg ggcttatcg ggcgcaccgc ggcgtcgca	3720
aatctgtca ggcgtgcaaa tgggatttaa gcttcgtct gaaacactgg gaagccatca	3780
aaaaageccgc tgaaagccgc ccggcccg tccgtattca tcaggagagc aacgtaatcg	3840
ttcgccgattt ccgcgattac ttacgtcagg acatcgccgaa aatccttatac gataaccgc	3900
aagtgcgtca actggcactgt cagcatatcg ctgcattagg tgcgcggat ttcagcagca	3960
aaatcaaact gtacacccgc gagatcccgc tgttcagcca ctaccagatc gagtcacacg	4020
tcgagtcgc cttccagcgt gaagttcgct tgccgtctgg tggttccatt gttatcgaca	4080
gcaccgaagc gttaacggcc atcgacatca actccgcacg cgccgacccgc ggccggata	4140
tcgaagaaac cgccgttaac actaacctcg aagctgccc tgagattgtc cgtagctgc	4200
gcctgcgtga cctcgccgc ctgattgtta tgcacttcat cgacatcgacg ccagtacgccc	4260
accagcgtgc ggtagaaaac cgtctgcgtg aagcggtgcg tcaggaccgt ggcgtattc	4320
aaatcagcca tatttctcg tttggccgtc tggaaatgtc ccgtcagcgc ctgagcccat	4380
cactgggtga atccagccat cacgtctgcg cgcgtctgcg cggtagccgtt accgtcgctg	4440
acaacgaatc gctgtcgctc tctattctgc gtctgatcgaa agaagaagcg ctgaaagaga	4500
acacccagga atttcactgcg attttccctg tgccaaatcg tctttatctg ctgaaatgaaa	4560
aacgttctgc ggtgaatgcc attgaaacgc gtcaggacgg tgttcgctgc gtgattgtgc	4620
caaacgatca gatggaaacc ccgcactacc acgtgctgcg cgtgcgtaaa ggggaagaaa	4680

-continued

cgccgacatt aagctacatg ctgccgaago tgcataaga agcgatggcg ctggctctg	4740
aagaagagtt cgctgaacgt aagcgccgg aacaacctgc gctggcaacc ttggccatgc	4800
cggatgtgcc gccagcgcca acccccagtc aacctgcccgc gcctgtcgta gccccagcac	4860
ctaaatctgc accggcaaca ccagccgctc ctgcccacc tgggctgtt agccgcttct	4920
tcggcgeact gaaagcgctg ttccagcggtg gtgaagaaac caaacgtcc gagcaaccaa	4980
caccgaaaagc agaagcgaaa ccggaaacgtc aacaggatcg tgcgaagct cgtcagaaca	5040
accggccgtga ccgtaatgag cgccgcgaca cccgtatgtg acgtactgaa ggcagcgata	5100
atcgcgaaaga aaaccgtcgt aatcgctgcc aggcacagca gcagactgcc gagacgcgtg	5160
agagccgtca gcagggttag gtaacggaaa aagcgctac caccgacgag cagcaagcgc	5220
cgcgtcgtga acgtagccgc cgccgtaatg atgataaactg tcaggcgcaaa caagaagcga	5280
aggcgctgaa tggtgaagag caatctgttc aggaaaccgaa acaggaagaa cgtgtacgtc	5340
cggttcagcc gcgtcgtaaa cagcgctcgc tcaatcagaa agtgcgttac gagcaaagcg	5400
tagccgaaga agcggtagtc gcacccgtgg ttgaagaaac tgcgtcgcc gaaccaattg	5460
ttaaggaaacgccagtcgc acgcacagaac tgggtaaagt cccgctgcgc gtcgtacgc	5520
aaactgcacc agaacagcaa gaagagaaca atgctgataa ccgtgacaac ggtggcatgc	5580
cgcgtcgttc tgcggctcg cctcgctacc tgcgtcgtaag tggtcagcgt cgtcgctc	5640
atcggtacgc gcgttatcca acccgatcgca caatccgtt gaccgtacgc tgcgtcgctc	5700
cggaaactggc ctctggcaaa gtctggatcc gctatccat tgtacgtccg caagatgtac	5760
aggttgaaga gcagcgcgaa caggaagaag tacatgtgc gccgatggtg actgaggatcc	5820
ctgtcgccgc cgctatcgaa ccgggttta gcgcgcagg tggtgaagaa gtggccgtg	5880
tcgttagaagc ccccggttag gttgcgaaac cgcaacccggaa agtgggttaga acgacgcattc	5940
ctgaagtgtat tgctgcgcg gtaactgaaac agccgcaggattaccggag tctgtatgttgc	6000
ccgtageccca ggaagttgca gaacaaggcg aaccgggtgg tgaacccgcg gaagagacgg	6060
cagatattga agaagttgtc gaaactgctg aggtttagt tgctgaacct gaagtttttg	6120
ctcaacatgc cgccgcagta gtcgtgaag tgcgcgcgaa agttgaaacgc gtagetgcgg	6180
tcgaacatgc ggtcaccgtt gacataacc acgttaccgc gcaatgacgc cgctccag	6240
cacccgaaata tgttccggag gcacccgcgc acagtgtact gcagcgccct actttgcct	6300
tcgaaggtaa aggtgccgca ggtggtcata cggcaacaca tcatgcctct gctgtcctg	6360
cgcgtccgca acctgttag taaagggttac ctggaaatca tccttagcga aagctaaggaa	6420
tttttttat ctgttaattaa ttgaaactcataa taaaggggaga ccacagcggtt ttccctttgt	6480
tcgcattggaa ggtcaataaa tgcgcgcgaa ttataaaccat ttctataagg ctccgaggag	6540
gcataatccaa gtgtgggagg cagccaatgg gccggatcca caggacgggt gtggcgccca	6600
tgatcgctgt gtcgtatgt gctccaaatgta gcaagcgag caggactggg cggccggccaa	6660
agcggtcgaa cagtgtccg agaacgggtg cgcataaaaaa ttgcgtcaac gcatatagcg	6720
ctagcagcac gccatagtgatgc ctggcgatgc tgcgtgttgc gacgtatcc cgcaagaggc	6780
ccggcgtac cggcataacc aagccatgc ctacagcatc cagggtgcacg gtggcgaggaa	6840
tgacgtatgtg cgcattgtta gatttcatac acgggtgcctg actgcgttag caatattaact	6900
tgatgataact accgcattaa agcttatacgatgc tgataagctg tcaaacatga gaattacaac	6960

-continued

ttatatcgta tggggctgac ttcaggtgct acatggaaag agataaaatgg cactgaaatc	7020
tagaaaattt ttatctgatt aataagatga ttttcttgag atcgaaaaatgg tctgcgcgta	7080
atctcttgcg ctgaaaacgaa aaaaacccgc ttgcaggcg gttttcgaa ggttctctga	7140
gctaccaact ctttgaaccg aggtaactgg ctggaggag cgcaatcacc aaaacttgtc	7200
cttcagttt agcctaacc ggcgcattgac ttcaagacta actcctctaa atcaattacc	7260
agtggctgct gccagtggtg ctttgcattt ttttccggg ttggactcaa gacgatagtt	7320
accggataag ggcgcgggtt cggaactgaaac ggggggttcg tgcatcagt ccagttgga	7380
gcgaactgcc tacccggAAC tgagtgtcag gcgtggatag agacaaacgc ggcataaca	7440
gcggaaatgac accggtaaac cgaaaggcag gaacaggaga ggcgcacggag gagccgcag	7500
gggaaacgcc tggtatcttt atagtcctgt cgggtttcgc caccactgat ttgagcgtca	7560
gatttcgtga tgcttgtagt gggggcgag cctatggaaa aacggcttg ccgcggccct	7620
ctcaacttccc tgtaagtat cttctggca tttccagga aatctccggcc cgggtcgtaa	7680
gcattttccg ctgcggcag tcgaacgacc gacgttagc agtcgtgag cgaggaaacg	7740
gaatataatcc tggatcacat attctgtca cgcacccgtt cggctttt tctctgcca	7800
catgaagcac ttcaactgaca ccctcatcg tgccaaatata gtaagccagt atacactccg	7860
ctagcgtga tgccggcg tgctttgcg gttacgcacc accccgtcag tagctgaaca	7920
ggagggacac ctgatagaaa cagaagccac tggagccat caaaaacacc atcatacact	7980
aaatcgtaa gttggcagca tcacccgacg cacttgcgc cgaataataa cctgtgacgg	8040
aaatcgtttt cgcagaataa ataaatccgt gtgtccctgt tgataccggg aagccctggg	8100
ccaaactttt gcaaaatga gacgttgatc ggcacgtaaag aggttccaac tttcaccata	8160
atgaaataag atcaactaccg ggcgtatTTT tttagtttgcg gagatTTCA ggagctaagg	8220
aagctaaaat ggagaaaaaa atcaactggat ataccacgt tgatataatcc caatggcatc	8280
gtaaagaaca ttttggggca ttctgtcgat ttgtcaatg tacctataac cagaccgttc	8340
agctggatat tacggcttt taaagacgg taaagaaaaa taagcacaag ttttacccgg	8400
cctttattca cattcttgcc cgcctgtatca atgctcatcc ggaattc	8447

```

<210> SEQ ID NO 82
<211> LENGTH: 7112
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)..(7112)
<223> OTHER INFORMATION: pMax-rneSTAR

```

```

<400> SEQUENCE: 82

cgtatggcaa taaaagacgg tgagctgggt atatggata gtgttacccc ttgttacacc      60
gttttccatg aacaaactga aacgtttca tcgtctggaa gtgaataacca cgacgatttc     120
cgccagtttacacatata ttgcgtatgtt acgggtaaaaa cctggccat      180
ttccctaaag gttttattga gaatatgttt ttgtctcgat ccaatccctg ggtgagttc     240
accagtttgcgtttaaaacgt ggcaatgtt gacaacttct tcggccccgtt ttccatcg      300
ggcaaatattt atacgcaagg cgacaagggtt ctgtatggccg tggcgattca ggttcatcat    360

```

-continued

ggcgctgtg atggcttcca tgcggcaga atgcttaatg aattacaaca gtactgcgat	420
gagtggcagg gcgccccgtt attttttaa ggcaattttt ggtccctta aacgccttgt	480
gctacgcctg aataagtgtat aataagcgga tgaatggcag aaattcgaaa gcaaattcga	540
cccggtcgtc gggtcagggc agggtcgta aatagccgt tatgtctatt gctggttac	600
cggtttattt actaccggaa gcagtgtgac cgtgtcgctt tcaaattgcctt gaggccagtt	660
tgcgtcaggct ctccccgtgg aggtataat tgcgtatgcgat atcattttt ctgcctccca	720
gaggcctgata aaaacgggta gcgttcgtt aatacagatg taggtgttcc acagggttagc	780
cagcagcatc ctgcgtatgca gatccggaaac ataatggtgc agggcgcttg ttccggcggt	840
ggtatggtgg caggccccgtt ggccggggga ctgttggggc ctgcggcac ctgttctacg	900
agttgcgtatgta taaagaagac agtcataatg gcggcgacga tagtcatgcc ccgcgcac	960
cggaaaggagc taccggacag cggtgcggac tggtaactt cagaataaga aatgaggccg	1020
ctcatggcgt tgacttcgt tcatagtatc gtggatcac cggttgggtt cacttcgt	1080
tgccggcaac ttccagcgtca cgtagggac ttccgcgtt ccagacttta cgaaacacgg	1140
aaacccgaaga ccattcatgt tggctcgtac gtccgcacgc ttttgcgcga gcaatcgctt	1200
cacgttgcgt cgcgtatcgg tgattcatc tgcataaccg taaggcaacc ccgcgcac	1260
agccgggtcc tcaacgcacag gagcacgtatcgtgcacccc gtggccagga cccaaacgcgt	1320
cccgagatgc gccgcgtgcg gctgtggag atggcggacgc cgatggatat gttctgcac	1380
gggttggttt ggcattcac agtttcgcg aagaattgtat tggctcaat tcttggagt	1440
gtgaatccgt tagcgagggtg ccgcggcgtt ccattcagggt cgagggtggcc cggctccatg	1500
caccgcgacg caacgcgggg aggccacaa ggtataggc ggcgcctaca atccatgca	1560
acccgttcca tgcgtcgcc gagggggcat aaatgcgcgt gacgtacgc ggtccagtga	1620
tgcgttgcgtat gctggtaaga gcccgcacgc atccttgcgtat ctgtccgtat tggctgtcat	1680
ctacctgcgtt ggacacgcgt gcgtcgtatcgc cgggcacccc gatgcgcgcg gaagecgagaa	1740
gaatcataat gggaaaggcc atccagcctc gcgtcgcgaa cgccagcaag acgtaccca	1800
gcccgcgttgcg cgcgtatcgc ggcataatgg cctgttgcgtat gcccggaaacgtt tgggtggcg	1860
gaccagtcgtac gaagggttgc gcgaggccgt gcaagattcc gaataccgcg acgcacaggc	1920
cgatcatcgatc cgcgtatccatc cgcgtatccatc gatgcgcgcgaa aatgcgcgcg acgtaccca	1980
gcacctgtcc tacgagttgc atgataaaga agacagtcat aatgcgcgcg acgtacgtca	2040
tgcgtccgcgc caccatcacc acggccaaac aacgcgtatc gatgcgcgcg acgtacgtca	2100
gctctccctt atgcgtactcc tgcatttagga agcagccatc tagtaggttgcg aggccgttgc	2160
gcaccgcgcgc cgcgtatccatc gatgcgcgcg acgtacgtca aatgcgcgcg acgtacgtca	2220
cggggcctgc caccatcacc acggccaaac aacgcgtatc gatgcgcgcg acgtacgtca	2280
gatcttcccc atcggtgtat tgcgtatccatc gatgcgcgcg acgtacgtca aatgcgcgcg	2340
tgcgtccgcgc caccatcacc acggccaaac aacgcgtatc gatgcgcgcg acgtacgtca	2400
ggaaaggagg aaagaataa tggctcggtt acagttttaaa caacgtaat ctactgcgc	2460
aatctttgtt cactgctcggtt ctaccaagcc aagtcgtatc gatgcgcgcg acgtacgtca	2520
ccagggtggccac aaagacgcgg gttggctcgatc tgcgtccgcgc caccatcacc tcaacgcgc	2580
cggtactgtg gaggcaggac gagatgagat ggcgttaggc tgcgtatc gatgcgcgcg acgtacgtca	2640

-continued

ccacaactct atccgcgtct gccttgg tggtatcgac gataaaggta agttcgcacgc	2700
taactttacg ccagccaaa tgcaatccct tcgctcactg cttgtcacac tgctggctaa	2760
gtacgaaggc gctgtgttcc gcccgcattca tgaggtggc ccgaaggcctt gcccttcgtt	2820
cgacctaag cgttggggg agaagaacga actggtaact tctgaccgtg gataaaccgc	2880
tgagcaataa ctagcataac cccttggggc ctctaaacggc gtcttgaggg gtttttggtt	2940
aattaagcgg cccggccgcg cccgtcgacg cggcgccgc tgaggatacc tcagccacca	3000
caatttcgca aatttgtgaac atcatcacgt tcatctttcc ctggttgcca atggcccat	3060
ttectgtcag taacgagaag gtcgcaattt caggcgcttt ttagactggt cgtaatgaaa	3120
ttcttttaa gaaggagact atatatgaaa gcaatttcg tactgaaaca tcttaatcat	3180
gctaaggagg ttttctaattg aaaagaatgt taatcaacgc aactcagcag gaagagtgc	3240
gcgttgcctt tgtatgtgg cagcgctgt atgacctggta tatcgaaagt ccaggccacg	3300
agcagaaaaa ggcaaaccatc tacaaaggta aaatcaccccg cattgaaccg agtctggaa	3360
ctgttttgt tgattacggc gctgaacgc acggtttccctt cccactaaaaa gaaattggcc	3420
gcaaatattt cctgtcaac tacagtgcct atggtcgtcc caacattaaa gatgtgtgc	3480
gtgaaggta ggaagtcatt gttcagatcg ataaagaaga ggcggcaac aaaggcgccgg	3540
cattaaccac ctttatcagt ctgggggta gctatctggt tctgatgccc aacaaccgc	3600
gcgcgggtgg cattttcgcg cgtatcgaa ggcacgaccg taccgaattta aaagaagcac	3660
tggcaaggct tgaactgcgcg gaaggcatgg ggettatacg ggcacccgctt ggcgtcgca	3720
aatctgtga ggcgtgcaaa tgggatttaa gcttcgtct gaaacactgg gaagccatca	3780
aaaaagccgc tgaaagccgc cggccccgtt tcctgattca tcaggagagc aacgtaatcg	3840
ttcgccgattt cccgcatttac ttacgtcagg acatcgccgaa aatccttatac gataaccgc	3900
aagtgcgtca actggcacgt cagcatatcg ctgcatttgcg tgcggccgtt ttcagcagca	3960
aaatcaaact gtacaccggc gagatcccgc tgttcagecca ctaccagatc gagtcacaga	4020
tcgagtcgc cttccagcgtt gaagttcgcc tggccgtctgg tggttccattt gttatcgaca	4080
gcaccgaaggc gttaacggcc atcgacatca actccgcacg cgcacccgc ggcggcgata	4140
tcgaagaaac cgcgtttaac actaacctcg aagctgcgcg tgagattgtc cgtagctgc	4200
gcctgcgtga cctcgccgc ctgattgtta tcgacttcat cgacatgcgc ccagtagcgc	4260
accagcgtgc ggtagaaaac cgtctgcgtg aagcggtgcg tcaggaccgt ggcgtattc	4320
aaatcagcca tattttcgc ttggcctgc tgaaaatgtc cggcgtcgcc ctgagcccat	4380
cactgggtga atccagtcac cacgtttgcg cgcgttgcgc tggtactggc accgtgcgtg	4440
acaacgaatc gtcgtcgctc tctattctgc gtctgatcgca agaagaagcg ctgaaagaga	4500
acacccagga agttcaccgc attgttcctg tgccaaatcgcc ttcttacctg ctgaaatgaaa	4560
aacgttctgc ggtaaatgcc attgaaactc gtcaggacgg tgcgtcgctgt gtaattgtgc	4620
caaacgatca gatggaaacc cccgcactacc acgtgcgtcg cgtgcgtaaa gggaaagaaa	4680
cccccaaccctt aagctacatcg ctgcccggc tgcatgaaga agcgtggcg ctgcccgt	4740
aagaagaggc cgctgaacgt aagcgccggc aacaaccgtc gtcggcaacc tttgcgtatc	4800
cggatgtgcc ccctgcgcca acggccagctg aacctgcgcg ccctgttgta gtcggccac	4860
cgaaagctgc accggcaaca ccaggcagctc ctgcacaacc tgggtgttg agccgcttct	4920

-continued

tcggcgcact gaaagcgctg ttcagcggtg gaagaaacca aaccgaccga gcaaccagca 4980
ccgaaagcg aaggcgaacc ggaacgtcaa caggatcg tcagaacaac 5040
cgccgtgaag gttacctaga aatcatcctt agcgaaagct aaggattttt ttatctgt 5100
attaattgaa ctcactaaag ggagaccaca gcggttccc ttgttcgca ttggaggctca 5160
aataatgogc aagtcttata aacaattcta taaggctcg aggaggcata tccaagtgt 5220
ggaggcagcc aatggggccgg atccacagga cgggtgtgg cgccatgatc gcgtgatcg 5280
tagtggctcc aagtagcgaa gcgagcagga ctggggccggc gcggacagtg 5340
ctccgagaac gggtgccat agaaaattgca tcaacgcata tagcgctagc agcacgc 5400
agtgactggc gatgctgtcg gaatggacga tatccgcaa gaggccggc agtaccggc 5460
taaccaagcc tatgcctaca gcatccagg tgacggtgcc gaggatgacg atgagcgcat 5520
tgttagattt catacaccggt gcctgactgc gttagcaatt taactgtgat aaactaccgc 5580
attaaagctt atcgatgata agctgtcaaa catgagaatt acaacttata tcgtatgggg 5640
ctgacttcag gtgctacatt tgaagagata aattgcactg aaatctagaa atattttatc 5700
tgattaataa gatgatctc ttgagatcgt tttggctctgc gcgtaatctc ttgctctgaa 5760
aacgaaaaaa ccgccttgca gggcggttt tcgaagggtc tctgagctac caactcttg 5820
aaccgaggtt actggcttgg aggagcgcag tcacccaaac ttgtccttc agtttagcct 5880
taaccggcgc atgacttcaa gactaactcc tctaaatcaa ttaccagtgg ctgctgccag 5940
tggtgctttt gcatgtcttt ccgggttgga ctcaagacga tagttaccgg ataaggcgca 6000
gcggctggc tgaacggggg gttcgatgc acagtccagc ttggagcgaa ctgccttaccc 6060
ggaactgagt gtcaggcggt gaatgagaca aacgcggcca taacagcgga atgacaccgg 6120
taaaaccgaaa ggcaggaaaca ggagagcgca cgagggagcc gcaaggggaa acgcctggta 6180
tctttatagt cctgtcggtt ttgcacacca ctgatttgcg cgtcagattt cgtatgttt 6240
gtcagggggggg eggagectat gggaaaacgg ctttgcgcgc gccccttcac ttccctgtta 6300
agtatcttc tggcatctc cagggaaatct ccgcggcgtt cgtaagccat ttccgtcg 6360
cgccatcgaa cgaccgagcg tagcgatgc gtgagcgagg aagcggaaata tatcctgtat 6420
cacatattct gctgaegcac cggtgcagec tttttctcc tgccacatga agcacttcac 6480
tgacaccctc atcagtgcca acatagtaag ccagttatata ctccgttagc gctgtatgtcc 6540
ggcggtgttt ttgcggttac gcaccacccc gtcagtagctt gaacaggagg gacagctgtat 6600
agaaacagaaa gcaactggag cacctcaaaa acaccatcat acactaaatc agtaatgtgg 6660
cagcatcacc cgacgcaccc ttgcggcgtt aaataccctgt gacggaaagat cacttcgc 6720
aataaaataaa ttctgggtgtc cctgttgata ccgggaagcc ctggggcaac ttttggcgaa 6780
aatgagacgt tgatcgac gtaagaggtt ccaacttca ccataatgaa ataagatcac 6840
tacccggcgtt attttttgcg ttatcgagat ttccaggagc taaggaagct aaaatggaga 6900
aaaaaaatcac tggatatacc accgttgata tttcccaatg gcatcgtaaa gaacattttg 6960
aggcattca gtcagttgtt caatgtaccc ataaccagac cggtcagctg gatattacgg 7020
cctttttaaa gaccgtaaag aaaaataagc acaagtttta tccggccctt attcacattc 7080
ttggccgcact gatgaatgtt catccggaaat tc 7112

-continued

```

<211> LENGTH: 1851
<212> TYPE: DNA
<213> ORGANISM: E. coli
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)..(1851)
<223> OTHER INFORMATION: rneSTAR encoding a truncated RNase E

<400> SEQUENCE: 83

atg aaa aga atg tta atc aac gca act cag cag gaa gag ttg cgc gtt      48
Met Lys Arg Met Leu Ile Asn Ala Thr Gln Gln Glu Glu Leu Arg Val
1           5          10          15

gcc ctt gta gat ggg cag cgt ctg tat gac ctg gat atc gaa agt cca      96
Ala Leu Val Asp Gly Gln Arg Leu Tyr Asp Leu Asp Ile Glu Ser Pro
20          25          30

ggg cac gag cag aaa aag gca aac atc tac aaa ggt aaa atc acc cgc      144
Gly His Glu Gln Lys Lys Ala Asn Ile Tyr Lys Gly Lys Ile Thr Arg
35          40          45

att gaa ccg agt ctg gaa gct gtt gat tac ggc gct gaa cgt      192
Ile Glu Pro Ser Leu Glu Ala Ala Phe Val Asp Tyr Gly Ala Glu Arg
50          55          60

cac ggt ttc ctc cca cta aaa gaa att gcc cgc gaa tat ttc cct gct      240
His Gly Phe Leu Pro Leu Lys Glu Ile Ala Arg Glu Tyr Phe Pro Ala
65          70          75          80

aac tac agt gct cat ggt cgt ccc aac att aaa gat gtg ttg cgt gaa      288
Asn Tyr Ser Ala His Gly Arg Pro Asn Ile Lys Asp Val Leu Arg Glu
85          90          95

ggt cag gaa gtc att gtt cag atc gat aaa gaa gag cgc ggc aac aaa      336
Gly Gln Glu Val Ile Val Gln Ile Asp Lys Glu Glu Arg Gly Asn Lys
100         105         110

ggc gcg gca tta acc acc ttt atc agt ctg gcg ggt agc tat ctg gtt      384
Gly Ala Ala Leu Thr Thr Phe Ile Ser Leu Ala Gly Ser Tyr Leu Val
115         120         125

ctg atg ccg aac aac ccg cgc ggc ggt ggc att tct ccg cgt atc gaa      432
Leu Met Pro Asn Asn Pro Arg Ala Gly Gly Ile Ser Arg Arg Ile Glu
130         135         140

ggc gac gac cgt acc gaa tta aaa gaa gca ctg gca agc ctt gaa ctg      480
Gly Asp Asp Arg Thr Glu Leu Lys Glu Ala Leu Ala Ser Leu Glu Leu
145         150         155         160

ccg gaa ggc atg ggg ctt atc gtg cgc acc gct ggc gtc ggc aaa tct      528
Pro Glu Gly Met Gly Leu Ile Val Arg Thr Ala Gly Val Gly Lys Ser
165         170         175

gct gag ggc ctg caa tgg gat tta agc ttc cgt ctg aaa cac tgg gaa      576
Ala Glu Ala Leu Gln Trp Asp Leu Ser Phe Arg Leu Lys His Trp Glu
180         185         190

gcc atc aaa aaa gcc gct gaa agc agc ccg gcc ccg ttc ctg att cat      624
Ala Ile Lys Ala Ala Glu Ser Arg Pro Ala Pro Phe Leu Ile His
195         200         205

cag gag agc aac gta atc gtt cgc gca ttc cgc gat tac tta cgt cag      672
Gln Glu Ser Asn Val Ile Val Arg Ala Phe Arg Asp Tyr Leu Arg Gln
210         215         220

gac atc ggc gaa atc ctt atc gat aac ccg aaa gtc ctc gaa ctg gca      720
Asp Ile Gly Glu Ile Leu Ile Asp Asn Pro Lys Val Leu Glu Leu Ala
225         230         235         240

cgt cag cat atc gct gca tta ggt cgc ccg gat ttc agc agc aaa atc      768
Arg Gln His Ile Ala Ala Leu Gly Arg Pro Asp Phe Ser Ser Lys Ile
245         250         255

aaa ctg tac acc ggc gag atc ccg ctg ttc agc cac tac cag atc gag      816
Lys Leu Tyr Thr Gly Glu Ile Pro Leu Phe Ser His Tyr Gln Ile Glu

```

-continued

260	265	270	
tca cag atc gag tcc gcc ttc cag cgt gaa gtt cgt ctg ccg tct ggt Ser Gln Ile Glu Ser Ala Phe Gln Arg Glu Val Arg Leu Pro Ser Gly 275 280 285			864
ggt tcc att gtt atc gac agc acc gaa gcg tta acg gcc atc gac atc Gly Ser Ile Val Ile Asp Ser Thr Glu Ala Leu Thr Ala Ile Asp Ile 290 295 300			912
aac tcc gca cgc gcg acc cgc ggc ggc gat atc gaa gaa acc gcg ttt Asn Ser Ala Arg Ala Thr Arg Gly Asp Ile Glu Glu Thr Ala Phe 305 310 315 320			960
aac act aac ctc gaa gct gcc gat gag att gct cgt cag ctg cgc ctg Asn Thr Asn Leu Glu Ala Ala Asp Glu Ile Ala Arg Gln Leu Arg Leu 325 330 335			1008
cgt gac ctc ggc ctg att gtt atc gac ttc atc gac atg acg cca Arg Asp Leu Gly Leu Ile Val Ile Asp Phe Ile Asp Met Thr Pro 340 345 350			1056
gta cgc cac cag cgt gcg gta gaa aac cgt ctg cgt gaa gcg gtg cgt Val Arg His Gln Arg Ala Val Glu Asn Arg Leu Arg Glu Ala Val Arg 355 360 365			1104
cag gac cgt gcg cgt att caa atc agc cat att tct cgc ttt ggc ctg Gln Asp Arg Ala Arg Ile Gln Ile Ser His Ile Ser Arg Phe Gly Leu 370 375 380			1152
ctg gaa atg tcc cgt cag cgc ctg agc cca tca ctg ggt gaa tcc agt Leu Glu Met Ser Arg Gln Arg Leu Ser Pro Ser Leu Gly Glu Ser Ser 385 390 395 400			1200
cat cac gtt tgt ccg cgt tgt tct ggt act ggc acc gtg cgt gac aac His His Val Cys Pro Arg Cys Ser Gly Thr Gly Thr Val Arg Asp Asn 405 410 415			1248
gaa tcg ctg tcg ctc tct att ctg cgt ctg atc gaa gaa gaa gcg ctg Glu Ser Leu Ser Leu Ser Ile Leu Arg Leu Ile Glu Glu Ala Leu 420 425 430			1296
aaa gag aac acc cag gaa gtt cac gcc att gtt cct gtg cca atc gct Lys Glu Asn Thr Gln Glu Val His Ala Ile Val Pro Val Pro Ile Ala 435 440 445			1344
tct tac ctg aat gaa aaa cgt tct gcg gta aat gcc att gaa act Ser Tyr Leu Leu Asn Glu Lys Arg Ser Ala Val Asn Ala Ile Glu Thr 450 455 460			1392
cgt cag gac ggt gtg cgc tgt gta att gtg cca aac gat cag atg gaa Arg Gln Asp Gly Val Arg Cys Val Ile Val Pro Asn Asp Gln Met Glu 465 470 475 480			1440
acc ccg cac tac cac gtg ctg cgc gtg cgt aaa ggg gaa gaa acc cca Thr Pro His Tyr His Val Leu Arg Val Lys Gly Glu Glu Thr Pro 485 490 495			1488
acc tta agc tac atg ctg ccg aag ctg cat gaa gaa gcg atg gcg ctg Thr Leu Ser Tyr Met Leu Pro Lys Leu His Glu Glu Ala Met Ala Leu 500 505 510			1536
ccg tct gaa gaa gag ttc gct gaa cgt aag cgt ccg gaa caa cct gcg Pro Ser Glu Glu Glu Phe Ala Glu Arg Lys Arg Pro Glu Gln Pro Ala 515 520 525			1584
ctg gca acc ttt gcc atg ccg gat gtg ccg cct gcg cca acg cca gct Leu Ala Thr Phe Ala Met Pro Asp Val Pro Pro Ala Pro Thr Pro Ala 530 535 540			1632
gaa cct gcc gcg cct gtt gta gct cca gca ccg aaa gct gca ccg gca Glu Pro Ala Ala Pro Val Val Ala Pro Ala Pro Lys Ala Ala Pro Ala 545 550 555 560			1680
aca cca gca gct cct gca caa cct ggg ctg ttg agc cgc ttc ttc ggc Thr Pro Ala Ala Pro Ala Gln Pro Gly Leu Leu Ser Arg Phe Phe Gly			1728

-continued

565	570	575	
gca ctg aaa gcg ctg ttc agc ggt gga aga aac caa acc gac cga gca Ala Leu Lys Ala Leu Phe Ser Gly Gly Arg Asn Gln Thr Asp Arg Ala 580 585 590			1776
acc agc acc gaa agc aga agc gaa acc gga acg tca aca gga tcg tcg Thr Ser Thr Glu Ser Arg Ser Glu Thr Gly Thr Ser Thr Gly Ser Ser 595 600 605			1824
caa gcc tcg tca gaa caa ccg ccg tga Gln Ala Ser Ser Glu Gln Pro Pro 610 615			1851
 <210> SEQ ID NO 84 <211> LENGTH: 616 <212> TYPE: PRT <213> ORGANISM: E. coli			
 <400> SEQUENCE: 84			
Met Lys Arg Met Leu Ile Asn Ala Thr Gln Gln Glu Glu Leu Arg Val 1 5 10 15			
Ala Leu Val Asp Gly Gln Arg Leu Tyr Asp Leu Asp Ile Glu Ser Pro 20 25 30			
Gly His Glu Gln Lys Lys Ala Asn Ile Tyr Lys Gly Lys Ile Thr Arg 35 40 45			
Ile Glu Pro Ser Leu Glu Ala Ala Phe Val Asp Tyr Gly Ala Glu Arg 50 55 60			
His Gly Phe Leu Pro Leu Lys Glu Ile Ala Arg Glu Tyr Phe Pro Ala 65 70 75 80			
Asn Tyr Ser Ala His Gly Arg Pro Asn Ile Lys Asp Val Leu Arg Glu 85 90 95			
Gly Gln Glu Val Ile Val Gln Ile Asp Lys Glu Glu Arg Gly Asn Lys 100 105 110			
Gly Ala Ala Leu Thr Thr Phe Ile Ser Leu Ala Gly Ser Tyr Leu Val 115 120 125			
Leu Met Pro Asn Asn Pro Arg Ala Gly Ile Ser Arg Arg Ile Glu 130 135 140			
Gly Asp Asp Arg Thr Glu Leu Lys Glu Ala Leu Ala Ser Leu Glu Leu 145 150 155 160			
Pro Glu Gly Met Gly Leu Ile Val Arg Thr Ala Gly Val Gly Lys Ser 165 170 175			
Ala Glu Ala Leu Gln Trp Asp Leu Ser Phe Arg Leu Lys His Trp Glu 180 185 190			
Ala Ile Lys Lys Ala Ala Glu Ser Arg Pro Ala Pro Phe Leu Ile His 195 200 205			
Gln Glu Ser Asn Val Ile Val Arg Ala Phe Arg Asp Tyr Leu Arg Gln 210 215 220			
Asp Ile Gly Glu Ile Leu Ile Asp Asn Pro Lys Val Leu Glu Leu Ala 225 230 235 240			
Arg Gln His Ile Ala Ala Leu Gly Arg Pro Asp Phe Ser Ser Lys Ile 245 250 255			
Lys Leu Tyr Thr Gly Glu Ile Pro Leu Phe Ser His Tyr Gln Ile Glu 260 265 270			
Ser Gln Ile Glu Ser Ala Phe Gln Arg Glu Val Arg Leu Pro Ser Gly 275 280 285			

-continued

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

```

<210> SEQ ID NO 85
<211> LENGTH: 9224
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)..(9224)
<223> OTHER INFORMATION: Plasmid pMac-rneEvo21

<400> SEQUENCE: 85

```

-continued

cgtatggcaa	tgaaaagacgg	tgagctggtg	atatggata	gtgttcaccc	ttgttacacc	60
gtttccatg	accaaactga	aacgtttca	tcgcctcgg	gtgaatacca	cgacgattc	120
cggcagttc	tacacatata	ttcgcgaagat	gtggcgttgt	acggtgaaaa	cctggccat	180
ttccctaaag	ggtttattga	gaatatgtt	ttcgctctcg	ccaateccctg	ggtgagttc	240
accagtttg	attnaacgt	ggccaatatg	gacaacttct	tcgccccctg	tttcaccatg	300
ggccaatatt	atacgcaagg	cgacaagggt	ctgatgccc	tggcgattca	ggttcatcat	360
gccgtctgt	atggcttcca	tgtcgccaga	atgcttaatg	aattacaaca	gtactgcgt	420
gagtggcagg	gcggggcgta	attttttaa	ggcagttatt	ggtgcccta	aacgcctgg	480
gctacgcctg	aataagtgt	aataagcgg	tgaatggcag	aaatcgaaa	gcaaattcga	540
cccggtcg	ggttcagggc	agggtcgta	aatagccgt	tatgtctatt	gctggttac	600
cggtttattg	actaccggaa	gcagtgtgac	cgtgtgcctc	tcaaatgcct	gaggccagtt	660
tgctcaggct	ctccccgtgg	aggtaataat	tgacgatatg	atcatttatt	ctgcctccca	720
gagcctgata	aaaacggta	gchgctcg	aatacagatg	taggtgttcc	acagggttagc	780
cagcagcatc	ctgcgatgca	gatccggaa	ataatggtgc	agggcgctt	tttcggcg	840
ggtatggtgg	caggccccgt	ggccggggga	ctgttggcg	ctgcccgcac	ctgtcctacg	900
agttgcgtga	taaagaagac	agtataagt	gcccgcacga	tagtcatgcc	ccgcgcac	960
cggaaaggagc	tacccggacag	cggcggac	tgttgtaact	cagaataaga	aatgaggccg	1020
ctcatggcg	tgactctcg	tcatagtac	gtggatcac	cggttgg	tactctgt	1080
tgccggcaac	ttcagcagca	cgtagggac	ttccgcgtt	ccagacttta	cgaaacacgg	1140
aaaccgaaga	ccattcatgt	tgttgctcg	gtcgacacg	ttttgcagca	gcagtcgc	1200
cacgttcg	cgcgtatcg	tgattcattc	tgctaaacc	taaggcaacc	ccgcac	1260
agccgggtcc	tcaacgacag	gagcacgatc	atgcgcaccc	gtggccagga	cccaacgc	1320
cccgagatgc	gcccgcgtcg	gctgtggag	atggcggac	cgatggat	gttctgc	1380
gggttgg	gcccattcac	agtttcccg	aagaattgt	tggctcaat	tcttggagtg	1440
gtgaatccgt	tagcgagg	ccgcggc	ccattcgat	cgaggtggcc	cggtccat	1500
caccgcgac	caacgcggg	aggcagacaa	ggtataaggc	ggcgcctaca	atccatgcca	1560
acccgttcca	tgtgcgc	gaggcggcat	aaatcgcc	gacgatc	ggtcc	1620
tcaagttag	gtcgtaaga	gccgcgac	atccttgaag	ctgtccctg	tggcgtcat	1680
ctacctgc	ggacagcatg	gcctgc	ccatccc	gatgcgc	gaagc	1740
gaatcataat	ggggaaaggcc	atccagcc	gcgtcg	cgccac	acgtac	1800
gcgcgtcg	gcgc	catgc	gac	ccgaaacgt	ttggtggcg	1860
gaccagtgac	gaaggctg	gcgaggcgt	gcaagattc	gaatacc	gcgacaggc	1920
cgcgc	atcgact	cggcgtcc	cgaaagcgt	cctgc	ccg	1980
gcacctgtcc	tacgagttc	atgataaaga	agacagt	cat	aa	2040
tgc	ccacccgg	gaag	ctgact	ggc	atcggtcg	2100
gc	atcgact	ccat	tgcat	ggc	aggcgttg	2160
gcacccgc	cgcaaggaat	ggtgc	atgc	aggatggc	gcccac	2220
cg	gggc	caccat	acgc	ccaaac	gagccgaa	2280

-continued

gatcttcccc atcgggtatg tcggcgatat aggcgccagc aaccgcacct gtggcgccgg	2340
tgatgcggc cacgtcggt cccggtaga ggatccgggt cccctttatg agattaaaaaa	2400
ggaaaggagg aaagaataa tggctcggt acagtttaaa caacgtaat ctactgacgc	2460
aatctttgtt cactgtcggt ctaccaagec aagtcaaat gttgggttcc gtgagatcg	2520
ccagtggtcac aaagagcagg gttggctcga tgtggataac cactttatca tcaagcgaga	2580
cggtaactgtg gaggcaggac gagatgagat ggctgttaggc tctcaegcta agggttacaa	2640
ccacaactct atcggcgctc gccttgggt tggtatcgac gataaaggta agttcgacgc	2700
taactttacg ccagccaaa tgcaatccct tcgctcaactg cttgtcacac tgctggctaa	2760
gtacgaaggc gctgtgttc gcgcggcatca tgaggtggcg ccgaaggctt gcccttcgtt	2820
cgacccttaag cggtgggtgg agaagaacga actggtaact tctgaccgtg gataaaccgc	2880
tgagcaataa ctagcataac cccttggggc ctctaaacgg gtcttgaggg gttttttgtt	2940
aattaagcgg ccgcggcgcc cccgtcgacg cgcgcgcggc tgaggatacc tcagccacca	3000
caattcagca aattgtgaac atcatcacgt tcatcttcc ctgggtgcca atggcccatt	3060
ttcctgtcag taacgagaag gtcgcgaatt caggcgctt ttagactggt cgtaatgaaa	3120
ttctttttaa aaaggagact atatatgaaa gcaatttcg tactgaaaca tcttaatcat	3180
gctaaggagg ttttctaataa aaaagaatgt taatcaacgc aactcagcag gaagagtgc	3240
gcgttgcctc ttagatggg cagcgctgt atgacctggg tatcgaagat ccaggggcagc	3300
agcagaaaaa ggcaaacatc tacaaaggta aaatcacccg cattgaaccg agtctggaaag	3360
ctgcttttgt tgattacggc gctgaacgc acgggttcccccactaaa gaaattgccc	3420
gcgaatattt ccctgctaact tacagtgc tc atgggtcgcc caacattaaa gatgtgtgc	3480
gtgaagggtca ggaagtcatt gttcagatcg ataaagaaga gcgcggcaac aaaggcgccg	3540
cattaaaccac ctttatcagt ctggcggtta gctatctggg tctgtatgcg aacaaccgc	3600
gcgcgggtgg catttttcgc cgtatcgaa ggcgcacccg taccgaattaa aaagaagcac	3660
tggcaaggct tgaactgccc gaaggcatgg ggcttatacg tgcacccgct ggcgtcgca	3720
aatctgtga ggccgtgcaaa tggttttaa gcttccgttca gaaacactgg gaagccatca	3780
aaaaagccgc tgaaagccgc ccggccccgt tcctgattca tcaggagagc aacgtaatcg	3840
ttcgcgcatt ccgcgttac ttacgtcagg acatcgccgaa atcccttatac gataaccgc	3900
aagtgcgtca actggcacgt cagcatatcg ctgcattagg tgcacccggat ttcagcagca	3960
aaatcaaact gtacaccggc gagatcccgc tggtcagccca ctaccagatc gagtcacaga	4020
tgcgtccgc ctccagcggt gaagttcgcc tgcgtctgg tggttccatt gttatcgaca	4080
gcaccgaagc gttaacggcc atcgacatca actccgcacgc cgcgcacccgc ggccggcgata	4140
tcgaagaaac ccgcgttaac actaacctcg aagctgcggta tgagattgtc cgtacgtgc	4200
gcctgcgtga cctcgccggc ctgattgtta tgcgttcatcg cgcacatgcg ccagtcacgc	4260
accagcggtgc ggtagaaaaac cgtctcggt aagcggtgcg tcaggaccgt ggcgtattc	4320
aaatcagccca tatttctcg tttggctgc tgaaatgtc ccgtcagcgc ctgagccat	4380
cactgggtga atccagccat cacgtctgc cgcgtctgc cggtaaccggat accgtcggt	4440
acaacgaatc gctgtcgctc tctattctgc gtctgatcgaa agaagaagcg ctgaaagaga	4500
acacccagga agttcacgccc attgttccgt tgccaaatcg tccttatctg ctgaaatgaaa	4560

-continued

aacgttctgc	ggtgaatgcc	attgaaacgc	gtcaggacgg	tgttcgtgc	gtgattgtc	4620
caaacgtatca	gatggaaacc	ccgactacc	acgtgctg	cgtgcgtaaa	gggaaagaaa	4680
cgccgaccc	aagctacatg	ctgccaago	tgcataaga	agcgatggcg	ctgcccgt	4740
aagaagagg	cgctgaacgt	aagegtccgg	aacaacctgc	gctggcaacc	tttgcacatgc	4800
cggatgtgcc	gccagcgcca	accccagctg	aacctgccc	gcctgtcgta	gccccagcac	4860
ctaaatctgc	accggcaaca	ccagccgctc	ctgcccacc	tgggctgtg	agccgcttct	4920
tccggcgact	gaaagcgctg	ttcagcggt	gtgaagaaac	caaaccgtcc	gagcaaccaa	4980
caccgaaagc	agaagcgaaa	ccggaacgtc	aacaggatcg	tgcataagc	cgtcagaaca	5040
accggccgtga	ccgtaatgag	cgccgcgaca	cccgtatgt	acgtactgaa	ggcagcgata	5100
atcgcgaaga	aaaccgtcgt	aatcgctccc	aggcacagca	gcagactgcc	gagacgcgt	5160
agagccgtca	cgagggttag	gtaacggaaa	aagcgctac	caccgacgag	cagcaagcgc	5220
cgcgtcgtga	acgtagccgc	cgccgtatg	atgataaaacg	tccaggcgca	caagaagcga	5280
aggcgctgaa	tgttgaagag	caaggtaatg	actccaaactt	attgatagt	ttttatgttc	5340
agataatgcc	cgatgacttt	gtcatgcagc	tccaccgatt	ttgagaacga	cagcgacttc	5400
cgtcccagcc	gtgccagggt	ctgcctcaga	ttcaggttat	gccgctcaat	tccgtcgta	5460
tatcgcttgc	tgattacgt	cagtttccc	ttcaggcggg	attcatacag	cgccagcc	5520
tccgtcatcc	atatcaccac	gtcaaagggt	gacagcgagc	tcataagacg	ccccagcg	5580
gccatagtgc	gttcaccgaa	tacgtgcgca	acaaccgtct	tccggagact	gtcatacgc	5640
taaaacagcc	agcgctggcg	cgatttagcc	ccgacatagc	cccactgtc	gtccatttcc	5700
gcmcagacga	tgacgtca	gccggctgt	atgcgcgagg	ttaccgactg	cgccctgagt	5760
tttttaagt	acgtaaaatc	gtgttggagc	caacgccc	aatgcgggct	gttgcggc	5820
atccaaacgcc	attcatggcc	atataatga	ttttctggtg	cgtaccgggt	tgagaagcgg	5880
tgtaagtgaa	ctgcagttgc	catgtttac	ggcagtgaga	gcagagatag	cgtgtatgc	5940
cgccgggtct	tttgcgttta	cgcaccaccc	cgtcaatgc	tgaacaggag	ggacagctga	6000
tagaaacaga	agccactgga	gcaccta	aaacccatca	tacactaaat	cgtaaatgt	6060
cgagcatcac	cgaagagcaa	tctgttcagg	aaaccgaaca	ggaagaacgt	gtacgtccgg	6120
ttcagccgc	tcgtaaacag	cgtcaatgc	atcagaaatg	cggttacgag	caaagcgt	6180
ccgaagaagc	ggtagtcgc	ccgggtgg	aagaaactgt	cgtgcggaa	ccaattttcc	6240
aggaagcgcc	agctccacgc	acagaactgg	tgaaagtccc	gtgcggatc	gtacgc	6300
ctgcaccaga	acagcaagaa	gagaacaatg	ctgataacc	tgacaacgg	ggcatgcgc	6360
gtcggttctc	cgcgtcgct	cgtcacctgc	cgtaatgt	tccgtatgc	cgtcgatc	6420
gtgacgacgc	ttatccaacc	cagtgcggaa	tgccgttgac	cgtacgtgc	cggttccgg	6480
aactggcctc	tggcaagtc	tggatccgt	atccaaattg	acgtccgca	gatgtacagg	6540
ttgaagagca	gogcgaacag	gaagaagtc	atgtgcagcc	gatggtgact	gaggccctg	6600
tcggccgcgc	tatcgaaacc	gttgttagcg	cgccagttgt	tgaagaatg	gccgggtcg	6660
tagaaagcccc	cgttcagg	gccgaaccgc	aaccggaaatg	ggttgaaacg	acgcatactg	6720
aagtgattgc	tgccgcggta	actgaacago	cgcaggatgt	taccgactt	gatgttgc	6780
tagcccaagga	agttgcagaa	caagcagaac	cggtggt	accgcaggaa	gagacggcag	6840

-continued

atattgaaga	agttgtcgaa	actgctgagg	ttgttagttgc	tgaacctgaa	gttggcgc	6900
aacctgccc	gccagtagtc	gctgaagtcg	cagcagaagt	tgaaacggta	gctgggtcg	6960
aacctgagg	caccgttgag	cataaccacg	ctaccgcgc	aatgacgcgc	gctccagcac	7020
cggaatatgt	tccggaggca	ccgegtacaca	gtgactggca	gcccctact	tttgcctcg	7080
aaggtaaagg	tgccgcaggt	ggtcatacgg	caacacatca	tgcctctgct	gctctgcgc	7140
gtccgcaacc	tgttgagtaa	aggttaccta	gaaatcatcc	ttagcgaag	ctaaggatt	7200
tttttatctg	taattaattt	aactcactaa	agggagacca	cagcggttc	ccttgctcg	7260
cattggaggt	caaataatgc	gcaagtctta	taaacaattc	tataaggctc	cgaggaggca	7320
tatccaagtg	tgggaggcag	ccaatggcc	ggatccacag	gacgggtgt	gtgcgcattga	7380
tcgcgtagtc	gatagtggct	ccaaagttagcg	aagcgagcag	gactgggcgg	cggccaaagc	7440
ggtcggacag	tgctccgaga	acgggtgcgc	atagaaatttgc	catcaacgc	tatagcgcta	7500
gcagcacgcc	atagtgactg	gcatgtctgt	cggaatggac	gatacccgc	aagaggcccc	7560
gcagtacccg	cataaccaag	cctatgccta	cagcatccag	ggtgacgggt	ccgaggatga	7620
cgatgagcgc	attgttagat	ttcatacacy	gtgcctgact	gcgttagcaa	tttaacttg	7680
ataaaactacc	gcattaaagc	ttatcgatga	taagctgtca	aacatgagaa	ttacaactta	7740
tatcgatgg	ggctgacttc	aggtgctaca	tttgaagaga	taaattgcac	tgaaatctag	7800
aaatatttta	tctgattaat	aagatgatct	tcttgagatc	gtttgggtct	gcgcgtatc	7860
tcttgctctg	aaaacgaaaa	aaccgccttg	cagggcggtt	ttcgaaggt	tctctgagct	7920
accaacttctt	tgaaccgagg	taactggctt	ggaggagcgc	agtcaccaaa	acttgtcctt	7980
tcaagtttagc	cttaaccggc	gcatgacttc	aagactaact	cctctaaatc	aattaccagt	8040
ggctgctgcc	agtgggtctt	ttgcatgtct	ttccgggttg	gactcaagac	gatagttacc	8100
ggataaggcg	cagcggtcgg	actgaacggg	gggttcgtgc	atacagtcca	gcttggagcg	8160
aactgcctac	ccggaaactga	gtgtcaggcg	tggaatgaga	caaacgcgc	cataacagcg	8220
gaatgacacc	ggtaaacccg	aaggcaggaa	caggagagcg	cacgagggag	ccgcccagggg	8280
aaacgcctgg	tatctttata	gtcctgtcgg	gttgcaccc	cactgatgg	agcgtcagat	8340
ttcgtgatgc	ttgtcagggg	ggcggagcct	atggaaaaac	ggcttgcgc	cggccctctc	8400
acttccctgt	taagtatctt	cctggcatct	tccaggaaat	ctccgcggcc	ttcgtaaagcc	8460
atttccgctc	gccgcagtcg	aacgaccgag	cgtagcgagt	cagtgagcga	ggaagcgaa	8520
tatatccctgt	atcacatatt	ctgctgacgc	accgggtcgac	cctttttct	cctgcacat	8580
gaagcacttc	actgacaccc	tcatcagtgc	caacatagta	agccagtata	cactccgcta	8640
gcgcgtatgt	cggcggtgc	ttttgcgtt	acgcaccaccc	ccgtcagtag	ctgaacagga	8700
gggacagctg	atagaaacag	aagccactgg	agcacctcaa	aaacaccatc	atacactaaa	8760
tcaagtaagtt	ggcagcatca	ccgcacgcac	tttgcgcgc	ataaataacct	gtgacggaa	8820
atcacttcgc	agaataaata	aatcctggtg	tccctgtga	taccggaaag	ccctggccca	8880
acttttggcg	aaaatgagac	gttgatcgcc	acgtaagagg	ttccaacttt	caccataatg	8940
aaataagatc	actaccgggc	gtattttttg	agttatcgag	attttcagga	gctaaggaaag	9000
ctaaaatgga	gaaaaaaaaatc	actggatata	ccaccgttg	tatatccaa	tggcatcgta	9060
aagaacattt	tgaggcattt	cagtcagtt	ctcaatgtac	ctataaccag	accgttcagc	9120

-continued

tggatattac ggcctttta aagaccgtaa agaaaaataa gcacaaggtt tatccggcct	9180
ttattcacat tcttgcggc ctgatgaatg ctcatccggaa attc	9224
<210> SEQ ID NO 86	
<211> LENGTH: 2187	
<212> TYPE: DNA	
<213> ORGANISM: Artificial sequence	
<220> FEATURE:	
<223> OTHER INFORMATION: synthetic	
<220> FEATURE:	
<221> NAME/KEY: CDS	
<222> LOCATION: (1)..(2187)	
<223> OTHER INFORMATION: rneEvo21 encoding truncated RNase E	
<400> SEQUENCE: 86	
atg aaa aga atg tta atc aac gca act cag cag gaa gag ttg cgc gtt	48
Met Lys Arg Met Leu Ile Asn Ala Thr Gln Gln Glu Glu Leu Arg Val	
1 5 10 15	
gcc ctt gta gat ggg cag cgt ctg tat gac ctg gat atc gaa agt cca	96
Ala Leu Val Asp Gly Gln Arg Leu Tyr Asp Leu Asp Ile Glu Ser Pro	
20 25 30	
ggg cac gag cag aaa aag gca aac atc tac aaa ggt aaa atc acc cgc	144
Gly His Glu Gln Lys Lys Ala Asn Ile Tyr Lys Gly Lys Ile Thr Arg	
35 40 45	
att gaa ccg agt ctg gaa gct gct ttt gtt gat tac ggc gct gaa cgt	192
Ile Glu Pro Ser Leu Glu Ala Ala Phe Val Asp Tyr Gly Ala Glu Arg	
50 55 60	
cac ggt ttc ctc cca cta aaa gaa att gcc cgc gaa tat ttc cct gct	240
His Gly Phe Leu Pro Leu Lys Glu Ile Ala Arg Glu Tyr Phe Pro Ala	
65 70 75 80	
aac tac agt gct cat ggt cgt ccc aac att aaa gat gtg ttg cgt gaa	288
Asn Tyr Ser Ala His Gly Arg Pro Asn Ile Lys Asp Val Leu Arg Glu	
85 90 95	
ggt cag gaa gtc att gtt cag atc gat aaa gaa gag cgc ggc aac aaa	336
Gly Gln Glu Val Ile Val Gln Ile Asp Lys Glu Glu Arg Gly Asn Lys	
100 105 110	
ggc gcg gca tta acc acc ttt atc agt ctg gcg ggt agc tat ctg gtt	384
Gly Ala Ala Leu Thr Thr Phe Ile Ser Leu Ala Gly Ser Tyr Leu Val	
115 120 125	
ctg atg ccg aac acc ccc cgc gcg ggt ggc att tct cgc cgt atc gaa	432
Leu Met Pro Asn Asn Pro Arg Ala Gly Gly Ile Ser Arg Arg Ile Glu	
130 135 140	
ggc gac gac cgt acc gaa tta aaa gaa gca ctg gca agc ctt gaa ctg	480
Gly Asp Asp Arg Thr Glu Leu Lys Glu Ala Leu Ala Ser Leu Glu Leu	
145 150 155 160	
ccg gaa ggc atg ggg ctt atc gtg cgc acc gct ggc gtc ggc aaa tct	528
Pro Glu Gly Met Gly Leu Ile Val Arg Thr Ala Gly Val Gly Lys Ser	
165 170 175	
gct gag ggc ctg caa tgg gat tta agc ttc cgt ctg aaa cac tgg gaa	576
Ala Glu Ala Leu Gln Trp Asp Leu Ser Phe Arg Leu Lys His Trp Glu	
180 185 190	
gcc atc aaa aaa gcc gct gaa agc agc cgc ccg gcc ttc ctg att cat	624
Ala Ile Lys Lys Ala Ala Glu Ser Arg Pro Ala Pro Phe Leu Ile His	
195 200 205	
cag gag agc aac gta atc gtt cgc gca ttc cgc gat tac tta cgt cag	672
Gln Glu Ser Asn Val Ile Val Arg Ala Phe Arg Asp Tyr Leu Arg Gln	
210 215 220	
gac atc ggc gaa atc ctt atc gat aac ccg aaa gtg ctc gaa ctg gca	720

-continued

Asp Ile Gly Glu Ile Leu Ile Asp Asn Pro Lys Val Leu Glu Leu Ala 225 230 235 240	
cgt cag cat atc gct gca tta ggt cgc ccg gat ttc agc agc aaa atc Arg Gln His Ile Ala Ala Leu Gly Arg Pro Asp Phe Ser Ser Lys Ile 245 250 255	768
aaa ctg tac acc ggc gag atc ccg ctg ttc agc cac tac cag atc gag Lys Leu Tyr Thr Gly Glu Ile Pro Leu Phe Ser His Tyr Gln Ile Glu 260 265 270	816
tca cag atc gag tcc gcc ttc cag cgt gaa gtt cgt ctg ccg tct ggt Ser Gln Ile Glu Ser Ala Phe Gln Arg Glu Val Arg Leu Pro Ser Gly 275 280 285	864
ggt tcc att gtt atc gac agc acc gaa gcg tta acg gcc atc gac atc Gly Ser Ile Val Ile Asp Ser Thr Glu Ala Leu Thr Ala Ile Asp Ile 290 295 300	912
aac tcc gca cgc gcg acc cgc ggc gat atc gaa gaa acc gcg ttt Asn Ser Ala Arg Ala Thr Arg Gly Gly Asp Ile Glu Glu Thr Ala Phe 305 310 315 320	960
aac act aac ctc gaa gct gcc gat gag att gct cgt cag ctg cgc ctg Asn Thr Asn Leu Glu Ala Ala Asp Glu Ile Ala Arg Gln Leu Arg Leu 325 330 335	1008
cgt gac ctc ggc ggc ctg att gtt atc gac ttc atc gac atg acg cca Arg Asp Leu Gly Leu Ile Val Ile Asp Phe Ile Asp Met Thr Pro 340 345 350	1056
gta cgc cac cag cgt gcg gta gaa aac cgt ctg cgt gaa gcg gtg cgt Val Arg His Gln Arg Ala Val Glu Asn Arg Leu Arg Glu Ala Val Arg 355 360 365	1104
cag gac cgt gcg cgt att caa atc agc cat att tct cgc ttt ggc ctg Gln Asp Arg Ala Arg Ile Gln Ile Ser His Ile Ser Arg Phe Gly Leu 370 375 380	1152
ctg gaa atg tcc cgt cag cgc ctg agc cca tca ctg ggt gaa tcc agc Leu Glu Met Ser Arg Gln Arg Leu Ser Pro Ser Leu Gly Glu Ser Ser 385 390 395 400	1200
cat cac gtc tgc ccg cgc tgc tcc ggt acc ggt acc gtg cgt gac aac His His Val Cys Pro Arg Cys Ser Gly Thr Gly Thr Val Arg Asp Asn 405 410 415	1248
gaa tcg ctg tcg ctc tct att ctg cgt ctg atc gaa gaa gaa gcg ctg Glu Ser Leu Ser Leu Ser Ile Leu Arg Leu Ile Glu Glu Ala Leu 420 425 430	1296
aaa gag aac acc cag gaa gtt cac gcc att gtt cct gtg cca atc gct Lys Glu Asn Thr Gln Glu Val His Ala Ile Val Pro Val Pro Ile Ala 435 440 445	1344
tct tat ctg ctg aat gaa aaa cgt tct gcg gtg aat gcc att gaa acg Ser Tyr Leu Leu Asn Glu Lys Arg Ser Ala Val Asn Ala Ile Glu Thr 450 455 460	1392
cgt cag gac ggt gtt cgc tgc gtg att gtg cca aac gat cag atg gaa Arg Gln Asp Gly Val Arg Cys Val Ile Val Pro Asn Asp Gln Met Glu 465 470 475 480	1440
acc ccg cac tac cac gtg ctg cgc gtg cgt aaa ggg gaa gaa acg ccg Thr Pro His Val Arg Val Lys Gly Glu Glu Ala Thr Pro 485 490 495	1488
acc tta agc tac atg ctg ccg aag ctg cat gaa gaa gcg atg gcg ctg Thr Leu Ser Tyr Met Leu Pro Lys Leu His Glu Glu Ala Met Ala Leu 500 505 510	1536
ccg tct gaa gaa gag ttc gct gaa cgt aag cgt ccg gaa caa cct gcg Pro Ser Glu Glu Glu Phe Ala Glu Arg Lys Arg Pro Glu Gln Pro Ala 515 520 525	1584
ctg gca acc ttt gcc atg ccg gat gtg ccg cca gcg cca acc cca gct	1632

-continued

Leu Ala Thr Phe Ala Met Pro Asp Val Pro Pro Ala Pro Thr Pro Ala			
530	535	540	
gaa cct gcc gcg cct gtc gta gcc cca gca cct aaa tct gca ccg gca			1680
Glu Pro Ala Ala Pro Val Val Ala Pro Ala Pro Lys Ser Ala Pro Ala			
545	550	555	560
aca cca gcc gct cct gcc caa cct ggg ctg ttg agc ccg ttc ttc ggc			1728
Thr Pro Ala Ala Pro Ala Gln Pro Gly Leu Leu Ser Arg Phe Phe Gly			
565	570	575	
gca ctg aaa gcg ctg ttc agc ggt ggt gaa gaa acc aaa ccg tcc gag			1776
Ala Leu Lys Ala Leu Phe Ser Gly Gly Glu Glu Thr Lys Pro Ser Glu			
580	585	590	
caa cca aca ccg aaa gca gaa gcg aaa ccg gaa cgt caa cag gat cgt			1824
Gln Pro Thr Pro Lys Ala Glu Ala Lys Pro Glu Arg Gln Gln Asp Arg			
595	600	605	
cgc aag cct cgt cag aac aac cgc cgt gac cgt aat gag cgc cgc gac			1872
Arg Lys Pro Arg Gln Asn Asn Arg Arg Asp Arg Asn Glu Arg Arg Asp			
610	615	620	
acc cgt agt gaa cgt act gaa ggc agc gat aat cgc gaa gaa aac cgt			1920
Thr Arg Ser Glu Arg Thr Glu Gly Ser Asp Asn Arg Glu Glu Asn Arg			
625	630	635	640
cgt aat cgt cgc cag gca cag cag act gcc gag acg cgt gag agc			1968
Arg Asn Arg Arg Gln Ala Gln Gln Thr Ala Glu Thr Arg Glu Ser			
645	650	655	
cgt cag cag gtt gag gta acg gaa aaa gcg cgt acc acc gac gag cag			2016
Arg Gln Gln Val Glu Val Thr Glu Lys Ala Arg Thr Thr Asp Glu Gln			
660	665	670	
caa gcg ccg cgt cgt gaa cgt agc cgc cgc cgt aat gat gat aaa cgt			2064
Gln Ala Pro Arg Arg Glu Arg Ser Arg Arg Asn Asp Asp Lys Arg			
675	680	685	
cag gcg caa caa gaa gcg aag gcg ctg aat gtt gaa gag caa ggt aat			2112
Gln Ala Gln Gln Glu Ala Lys Ala Leu Asn Val Glu Glu Gln Gly Asn			
690	695	700	
gac tcc aac tta ttg ata gtg ttt tat gtt cag ata atg ccc gat gac			2160
Asp Ser Asn Leu Leu Ile Val Phe Tyr Val Gln Ile Met Pro Asp Asp			
705	710	715	720
ttt gtc atg cag ctc cac cga ttt tga			2187
Phe Val Met Gln Leu His Arg Phe			
725			

<210> SEQ ID NO 87

<211> LENGTH: 728

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic Construct

<400> SEQUENCE: 87

Met Lys Arg Met Leu Ile Asn Ala Thr Gln Gln Glu Glu Leu Arg Val			
1	5	10	15
Ala Leu Val Asp Gly Gln Arg Leu Tyr Asp Leu Asp Ile Glu Ser Pro			
20	25	30	
Gly His Glu Gln Lys Lys Ala Asn Ile Tyr Lys Gly Lys Ile Thr Arg			
35	40	45	
Ile Glu Pro Ser Leu Glu Ala Ala Phe Val Asp Tyr Gly Ala Glu Arg			
50	55	60	
His Gly Phe Leu Pro Leu Lys Glu Ile Ala Arg Glu Tyr Phe Pro Ala			
65	70	75	80

-continued

Asn	Tyr	Ser	Ala	His	Gly	Arg	Pro	Asn	Ile	Lys	Asp	Val	Leu	Arg	Glu	
85								90						95		
Gly	Gln	Glu	Val	Ile	Val	Gln	Ile	Asp	Lys	Glu	Glu	Arg	Gly	Asn	Lys	
100								105						110		
Gly	Ala	Ala	Leu	Thr	Thr	Phe	Ile	Ser	Leu	Ala	Gly	Ser	Tyr	Leu	Val	
115								120						125		
Leu	Met	Pro	Asn	Asn	Pro	Arg	Ala	Gly	Gly	Ile	Ser	Arg	Arg	Ile	Glu	
130								135						140		
Gly	Asp	Asp	Arg	Thr	Glu	Leu	Lys	Glu	Ala	Leu	Ala	Ser	Leu	Glu	Leu	
145								150						155		160
Pro	Glu	Gly	Met	Gly	Leu	Ile	Val	Arg	Thr	Ala	Gly	Val	Gly	Lys	Ser	
165								170						175		
Ala	Glu	Ala	Leu	Gln	Trp	Asp	Leu	Ser	Phe	Arg	Leu	Lys	His	Trp	Glu	
180								185						190		
Ala	Ile	Lys	Lys	Ala	Ala	Glu	Ser	Arg	Pro	Ala	Pro	Phe	Leu	Ile	His	
195								200						205		
Gln	Glu	Ser	Asn	Val	Ile	Val	Arg	Ala	Phe	Arg	Asp	Tyr	Leu	Arg	Gln	
210								215						220		
Asp	Ile	Gly	Glu	Ile	Leu	Ile	Asp	Asn	Pro	Lys	Val	Leu	Glu	Leu	Ala	
225								230						235		240
Arg	Gln	His	Ile	Ala	Ala	Leu	Gly	Arg	Pro	Asp	Phe	Ser	Ser	Lys	Ile	
245								250						255		
Lys	Leu	Tyr	Thr	Gly	Glu	Ile	Pro	Leu	Phe	Ser	His	Tyr	Gln	Ile	Glu	
260								265						270		
Ser	Gln	Ile	Glu	Ser	Ala	Phe	Gln	Arg	Glu	Val	Arg	Leu	Pro	Ser	Gly	
275								280						285		
Gly	Ser	Ile	Val	Ile	Asp	Ser	Thr	Glu	Ala	Leu	Thr	Ala	Ile	Asp	Ile	
290								295						300		
Asn	Ser	Ala	Arg	Ala	Thr	Arg	Gly	Gly	Asp	Ile	Glu	Glu	Thr	Ala	Phe	
305								310						315		320
Asn	Thr	Asn	Leu	Glu	Ala	Ala	Asp	Glu	Ile	Ala	Arg	Gln	Leu	Arg	Leu	
325								330						335		
Arg	Asp	Leu	Gly	Gly	Leu	Ile	Val	Ile	Asp	Phe	Ile	Asp	Met	Thr	Pro	
340								345						350		
Val	Arg	His	Gln	Arg	Ala	Val	Glu	Asn	Arg	Leu	Arg	Glu	Ala	Val	Arg	
355								360						365		
Gln	Asp	Arg	Ala	Arg	Ile	Gln	Ile	Ser	His	Ile	Ser	Arg	Phe	Gly	Leu	
370								375						380		
Leu	Glu	Met	Ser	Arg	Gln	Arg	Leu	Ser	Pro	Ser	Leu	Gly	Glu	Ser	Ser	
385								390						395		400
His	His	Val	Cys	Pro	Arg	Cys	Ser	Gly	Thr	Gly	Thr	Val	Arg	Asp	Asn	
405								410						415		
Glu	Ser	Leu	Ser	Leu	Ser	Ile	Leu	Arg	Leu	Ile	Glu	Glu	Glu	Ala	Leu	
420								425						430		
Lys	Glu	Asn	Thr	Gln	Glu	Val	His	Ala	Ile	Val	Pro	Val	Pro	Ile	Ala	
435								440						445		
Ser	Tyr	Leu	Leu	Asn	Glu	Lys	Arg	Ser	Ala	Val	Asn	Ala	Ile	Glu	Thr	
450								455						460		
Arg	Gln	Asp	Gly	Val	Arg	Cys	Val	Ile	Val	Pro	Asn	Asp	Gln	Met	Glu	
465								470						475		480
Thr	Pro	His	Tyr	His	Val	Leu	Arg	Val	Arg	Lys	Gly	Glu	Glu	Thr	Pro	

-continued

485	490	495
Thr Leu Ser Tyr Met Leu Pro Lys Leu His Glu Glu Ala Met Ala Leu		
500	505	510
Pro Ser Glu Glu Glu Phe Ala Glu Arg Lys Arg Pro Glu Gln Pro Ala		
515	520	525
Leu Ala Thr Phe Ala Met Pro Asp Val Pro Pro Ala Pro Thr Pro Ala		
530	535	540
Glu Pro Ala Ala Pro Val Val Ala Pro Ala Pro Lys Ser Ala Pro Ala		
545	550	555
560		
Thr Pro Ala Ala Pro Ala Gln Pro Gly Leu Leu Ser Arg Phe Phe Gly		
565	570	575
Ala Leu Lys Ala Leu Phe Ser Gly Gly Glu Glu Thr Lys Pro Ser Glu		
580	585	590
Gln Pro Thr Pro Lys Ala Glu Ala Lys Pro Glu Arg Gln Gln Asp Arg		
595	600	605
Arg Lys Pro Arg Gln Asn Asn Arg Arg Asp Arg Asn Glu Arg Arg Asp		
610	615	620
Thr Arg Ser Glu Arg Thr Glu Gly Ser Asp Asn Arg Glu Glu Asn Arg		
625	630	635
640		
Arg Asn Arg Arg Gln Ala Gln Gln Thr Ala Glu Thr Arg Glu Ser		
645	650	655
Arg Gln Gln Val Glu Val Thr Glu Lys Ala Arg Thr Thr Asp Glu Gln		
660	665	670
Gln Ala Pro Arg Arg Glu Arg Ser Arg Arg Arg Asn Asp Asp Lys Arg		
675	680	685
Gln Ala Gln Gln Glu Ala Lys Ala Leu Asn Val Glu Glu Gln Gly Asn		
690	695	700
Asp Ser Asn Leu Leu Ile Val Phe Tyr Val Gln Ile Met Pro Asp Asp		
705	710	715
720		
Phe Val Met Gln Leu His Arg Phe		
725		

```

<210> SEQ ID NO 88
<211> LENGTH: 295
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: promoter
<222> LOCATION: (1)..(295)
<223> OTHER INFORMATION: araBAD promoter

<400> SEQUENCE: 88

aaacccaattg tccatattgc atcagacatt gccgtcactg cgtctttac tggctcttct    60
cgctaaccaa accggtaacc ccgcttatta aaagcattct gtaacaaagc gggaccaaag    120
ccatgacaaa aacgcgtaac aaaagtgtct ataatcacgg cagaaaagtc cacattgatt    180
atttgcacgg cgtcacactt tgctatgcc tagcattttt atccataaga ttagcggatc    240
ctacctgacg cttttatcg caactctcta ctgtttctcc atacccgttt ttttg    295

```

```

<210> SEQ ID NO 89
<211> LENGTH: 54
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence

```

-continued

```

<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: promoter
<222> LOCATION: (1)..(54)
<223> OTHER INFORMATION: Ptet promoter

<400> SEQUENCE: 89

ttgacaatta atcatccggc tcgtataatg gaatttgtag cggtataacaa ttcc      54

<210> SEQ ID NO 90
<211> LENGTH: 82
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: promoter
<222> LOCATION: (1)..(82)
<223> OTHER INFORMATION: Ptet promoter

<400> SEQUENCE: 90

tcatgttta cagcttatca tcgatttagct ttaatgcgggt agtttatcac agttaaattt      60
ctaacgcagt caggcaccgt gt                                         82

<210> SEQ ID NO 91
<211> LENGTH: 29
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: promoter
<222> LOCATION: (1)..(29)
<223> OTHER INFORMATION: Ptac promoter

<400> SEQUENCE: 91

ttgacaatta atcatcggtc cggtataatg                                         29

<210> SEQ ID NO 92
<211> LENGTH: 51
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: synthetic
<220> FEATURE:
<221> NAME/KEY: promoter
<222> LOCATION: (1)..(51)
<223> OTHER INFORMATION: PL promoter

<400> SEQUENCE: 92

ggcggtgttg acataaatac cactggcggt gatactgagc acatcagcag g      51

```

1. A prokaryotic microbial host cell for recombinant expression of a target protein, said cell comprising
 - A. a gene encoding an enzyme having endoribonuclease activity (E.C. 3.1.26), wherein said gene is on the genome of said cell,
 - B. a first recombinant gene encoding a mutant RNase E, wherein the amino acid sequence of said mutant RNase E has at least 75% sequence identity with SEQ ID NO. 2, and wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substi-

- tution(s) which results in the mutant RNase E having decreased activity compared to the RNase E of SEQ ID NO. 2, and
- C. a second recombinant gene encoding said target protein, wherein expression of said target protein is enhanced compared to a cell lacking said first recombinant gene.
2. The prokaryotic microbial host cell according to claim 1, wherein the target protein is a toxic protein such as wherein the target protein is a membrane protein.
3. (canceled)

4. The prokaryotic microbial host cell according to claim **1**, wherein the enzyme having ribonuclease activity (E.C. 3.1.26), encoded by the gene on the genome, is native to the host cell.

5. The prokaryotic microbial host cell according to claim **1**, wherein the enzyme having ribonuclease activity (E.C. 3.1.26), encoded by the gene on the genome, is an RNase E enzyme having at least 75% sequence identity with SEQ ID NO. 2.

6. The prokaryotic microbial host cell according to claim **1**, wherein the one or more amino acid residue substitutions is at one or more positions selected from D346, E297, D303, N305, E325, R337, D349, V128, R169, T170, F57, F67, K112, G124, R141, R142, or R373 relative to SEQ ID NO. 2.

7. The prokaryotic microbial host cell according to claim **1**, wherein the one or more amino acid residue substitutions is at one or more positions in the DNase I-like domain, such as one or more positions selected from E297, D303, N305, E325, R337, D346, D349, and R373.

8. (canceled)

9. The prokaryotic microbial host cell according to claim **1**, wherein the one or more amino acid residue substitutions results in a mutant RNase E having reduced metal ion chelation ability compared to the RNase E of SEQ ID NO. 2, such as one or more positions selected from D346, E297, D303, E325, R337, and D349 relative to SEQ ID NO. 2.

10. (canceled)

11. The prokaryotic microbial host cell according to claim **1**, wherein the one or more amino acid residue substitutions results in a mutant RNase E having a modified RNA contact point compared to the RNase E of SEQ ID NO. 2, such as one or more positions selected from F57, F67, and K112.

12. (canceled)

13. The prokaryotic microbial host cell according to claim **1**, wherein the one or more amino acid residue substitutions is in the 5' sensor pocket, preferably the pocket 'anchors', such as one or more amino acid residue substitutions is at positions V128 and/or R373.

14. (canceled)

15. The prokaryotic microbial host cell according to claim **1**, wherein the amino acid residue substitution is A441.

16. The prokaryotic microbial host cell according to claim **1**, wherein the amino acid residue substitution facilitates the enhanced expression of said target protein, and wherein said amino acid residue substitution is identified and selected by a screening method comprising the steps of

- A. expressing the target protein together with a candidate mutant RNase E comprising a candidate amino acid residue substitution in the host cell,
- B. expressing the target protein in a parent cell (from which the host cell was derived) lacking expression of the candidate mutant RNase E,
- C. comparing expression levels of the target protein in (a) and (b), and identifying one or more candidate(s) which facilitate enhanced expression of said target protein.

17. The prokaryotic microbial host cell according to claim **1**, wherein said cell further comprises a first prokaryotic vector, and wherein said first recombinant gene encoding said mutant RNase E is comprised on said first prokaryotic vector.

18. The prokaryotic microbial host cell according to claim **1**, wherein said cell further comprises

D. a gene encoding a T7 RNA polymerase (E.C. 2.7.7.6),
E. optionally a gene encoding a T7 lysozyme (E.C. 3.5.1.28)

and wherein expression of said second recombinant gene is regulated by an inducible T7 promoter.

19. The prokaryotic microbial host cell according to claim **18**, wherein said gene encoding said T7 lysozyme is located on the first prokaryotic vector, and wherein said second recombinant gene encoding said target gene is located on a second prokaryotic vector.

20. The prokaryotic microbial host cell according to claim **1**, wherein expression of said second recombinant gene is regulated by an inducible promoter selected from rhaBAD promoter, araBAD promoter, Ptrc promotor, Ptet promoter, Ptac promoter, and PL promoter.

21. The prokaryotic microbial host cell according to claim **1**, wherein said target protein is a protein the expression of which is enhanced by at least 10% compared to expression of said protein in the same host cell lacking said first recombinant gene.

22. The prokaryotic microbial host cell according to claim **1**, wherein said cell is selected from *E. coli*, *Bacillus subtilis*, *Bacillus licheniformis*, and *Pseudomonas putida*.

23. A prokaryotic vector comprising

A. a gene encoding a mutant RNase E (E.C. 3.1.26.12) having at least 75% amino acid sequence identity to SEQ ID NO. 2, wherein the amino acid sequence of said mutant RNase E has one or more amino acid residue substitution(s) which results in the mutant RNase E having decreased activity compared to the RNase E of SEQ ID NO. 2, and

B. a gene encoding a T7 lysozyme (E.C. 3.5.1.28)

24. The prokaryotic vector according to claim **23**, wherein the one or more amino acid residue substitutions (i) is at one or more positions in the DNase I-like domain, (ii) results in a mutant RNase E having reduced metal ion chelation ability compared to the RNase E of SEQ ID NO. 2. (iii) results in a mutant RNase E having a modified RNA contact point compared to the RNase E of SEQ ID NO. 2, or (iv) is in the 5' sensor pocket, such as the pocket 'anchors'.

25. (canceled)

26. (canceled)

27. (canceled)

28. (canceled)

29. A method for the production of a target protein, comprising culturing in a suitable culture medium, a prokaryotic microbial host cell according to claim **1**, expressing said target protein, and optionally isolating the expressed target protein.

* * * * *