

WO 2006/047589 A2

Attorney Docket No.0359.210WO/15686WO02

IMPROVED ALANINE 2,3-AMINOMUTASES
AND RELATED POLYNUCLEOTIDES

FIELD OF THE INVENTION

[01] The present invention is related to the field of enzymology, and particularly to the field of alanine 2,3-aminomutase (AAM) enzymology. More specifically, the present invention is directed to alanine 2,3-aminomutase polypeptides having improved enzymatic activity (*i.e.*, high substrate turnover) and stability, and to polynucleotides sequences encoding for the improved alanine 2,3-aminomutase polypeptides. The present invention is useful because the alanine 2,3-aminomutase polypeptides can be coupled to other enzymes to produce synthetic organic chemicals, such as pantothenic acid or 3-hydroxypropionic acid in high yields.

BACKGROUND OF THE INVENTION

[02] Organic chemicals such as organic acids, esters, and polyols can be used to synthesize plastic materials and other products. To meet the increasing demand for organic chemicals, more efficient and cost-effective production methods are being developed which utilize raw materials based on carbohydrates rather than hydrocarbons. For example, certain bacteria have been used to produce large quantities of lactic acid used in the production of polylactic acid.

[03] 3-hydroxypropionic acid (3-HP) is an organic acid. Several chemical synthesis routes have been described to produce 3-HP, and biocatalytic routes have also been disclosed (WO 01/16346 to Suthers et al.). 3-HP has utility for specialty synthesis and can be converted to commercially important intermediates by known methods in the chemical industry, *e.g.*, acrylic acid by dehydration, malonic acid by oxidation, esters by esterification reactions with alcohols, and 1,3-propanediol by reduction.

[04] The compound 3-HP can be produced biocatalytically from PEP or pyruvate, through a key beta-alanine intermediate (FIG. 1). Beta-alanine can be synthesized in

-2-

cells from carnosine, beta-alanyl arginine, beta-alanyl lysine, uracil via 5,6-dihydrouracil and N-carbamoyl-beta-alanine, N-acetyl-beta-alanine, anserine, or aspartate. However, these routes are commercially unviable because they require rare precursors or starting compounds that are more valuable than 3-HP. Therefore, production of 3-HP using biocatalytic routes would be more efficient if alpha-alanine could be converted to beta-alanine directly (FIG. 1). Unfortunately, a naturally occurring enzyme that inter-converts alpha-alanine to beta-alanine has not yet been identified. It would be advantageous if enzymatic activities that carry out the conversion of alpha-alanine to beta-alanine were identified, such as an alanine 2,3-aminomutase. Accordingly, it is one object of the present invention to identify enzymes with improved alanine 2-3-aminomutase activity.

[05] Lysine 2,3-aminomutase (KAM), which catalyzes the anaerobic interconversion of lysine to beta-lysine, was first described by Barker in *Clostridium* SB4 (now *C. subterminale*) catalyzing the first step in the fermentation of lysine. KAM has been purified from *C. subterminale*, the gene cloned and expressed in *E. coli*. See e.g., U.S. Pat. 6,248,874, which issued on June 19, 2001 to Frey *et al.*, the whole of which is hereby incorporated herein by reference. The specific activity of purified KAM from *C. subterminale* SB4 cells has been reported as 30-40 units/mg (Lieder *et al.*, Biochemistry 37:2578 (1998)), where a unit is defined as μ moles lysine/min. The corresponding purified recombinantly produced KAM had equivalent enzyme activity (34.5 ± 1.6 μ moles lysine/min/mg protein). See U.S. Patent Application Publication No. 2003/0113882 A1, which published on June 19, 2003 to Frey *et al.*, the whole of which is incorporated herein by reference.

[06] Based upon the sequence of the KAM from *C. subterminale*, KAM genes have been annotated in the genomes of other organisms. However, in most cases, the enzymatic activities of the polypeptides encoded by these genes have not been confirmed. Exceptions are the *B. subtilis* gene (Chen, D., Ruzicka, F.J., and Frey, P.A. (2000) Biochem. J. 348:539-549)), and the *Porphyromonas gingivalis* and *F. nucleatum* genes. The *B. subtilis* KAM, encoded by the *yodO* gene, is more resistant to O₂ than the *C. subterminale* KAM, but it is markedly less active. As reported by Frey, the *B. subtilis* KAM has a specific activity of only 0.62 U/mg.

-3-

[07] *C. subterminale* SB4 KAM has been reported to have some cross-reactivity with L-alanine, converting it into beta-alanine. See U.S. Patent Application Publication No. 2003/0113882 A1. WO 03/062173 and WO 02/42418 disclose the first reports of AAM activity based upon modification of *kam* genes. In these applications, the synthetic *aam* genes had AAM activity as detected by the complementation of a Δ panD *E. coli* strain. However, because alanine is not the natural substrate for this enzyme, the activity for this conversion is substantially less than the activity for conversion of lysine — its natural substrate. The AAM activity of a variant of *B. subtilis* KAM that also had AAM activity at approximately 0.001 U/mg. It is an object of the present invention to provide polynucleotides encoding a polypeptide having substantially enhanced AAM activity over that found in the wild-type enzymes.

SUMMARY OF THE INVENTION

[08] The present invention has multiple aspects. In one aspect, the present invention is directed to polypeptides that catalyze the reaction of FIG. 1. In one embodiment of this first aspect, the present invention is directed to a polypeptide having alanine 2,3-aminomutase (AAM) activity, preferably as measured by the assay of Example 8, and,

(a) having a polypeptide selected from the group consisting of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48 and 51;

(b) having an amino acid sequence which has at least 98% homology, with the amino acid sequence selected from the group consisting of SEQ ID NO: 2, 22, 28, 32, and 36;

(c) having an amino acid sequence which has at least 99% homology, with the amino acid sequence selected from the group consisting of SEQ ID NO: 4, 6, 8, 12, 16, 24, 26, 30, 34 and 40;

(d) being a polypeptide encoded by a nucleic acid sequence which hybridizes under high stringency conditions with either (i) the nucleotide sequence of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 41, 43, 45, 47 or 49; (ii) a subsequence of (i) of at least 100 nucleotides, or (iii) a complementary strand of (i) or (ii) (J. Sambrook, E. F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning, A Laboratory Manual*, 2d edition, Cold Spring Harbor, N.Y.); or

(e) being a variant of the polypeptide of (c) comprising a substitution, deletion, and/or insertion of one to six amino acids therefrom and having AAM activity from about 1 to about 30 μ M β -alanine produced /hour 1 cell OD at pH 7.0-7.6, 25°C.

[09] Collectively, the polypeptides of (b) and (c) above are referred to herein as "homologous polypeptides." For purposes of the present invention, the degree of homology between two amino acid sequences is expressed as "percent homology," "percent identity," "% identity," "percent identical," and "% identical" are used interchangeably herein to refer to the percent amino acid sequence identity that is obtained by ClustalW analysis (version W 1.8 available from European Bioinformatics Institute, Cambridge, UK), counting the number of identical matches in the alignment and dividing such number of identical matches by the length of the

-5-

reference sequence, and using the following default ClustalW parameters to achieve slow/accurate pairwise optimal alignments – Gap Open Penalty:10; Gap Extension Penalty:0.10; Protein weight matrix: Gonnet series; DNA weight matrix: IUB; Toggle Slow/Fast pairwise alignments = SLOW or FULL Alignment.

[10] In one embodiment, the present invention is also directed to an AAM polypeptide as described herein in isolated and purified form.

[11] In another embodiment, the present invention is directed to an AAM polypeptide as described herein in lyophilized form.

[12] In yet another embodiment, the present invention is directed to a composition comprising an AAM polypeptide as described herein and a suitable carrier, typically a buffer solution, more typically an aqueous buffer solution having a pH between 6.0 and 8.0. The composition may also be in a lyophilized form.

[13] The novel AAM polypeptides of the present invention have significantly enhanced AAM activity relative to the wild-type KAM polypeptides from which they are ultimately derived. By significantly enhanced AAM activity is meant that the AAM polypeptide of the present invention has an AAM activity within the range of about 1 to about 32 μ M β -alanine produced/hour 1 cell OD (units), preferably from about 10 to about 32 units, more preferably from about 20 to about 32 units; most preferably from about 25 to about 32 units.

[14] Preferred AAM polypeptides of the present invention have an amino acid sequences of SEQ ID NOs: 2, 6, 12, 16, 20, 24, 28, 30, 32, 34, 38, 44, 46 or 48; more preferably they have an amino acid sequence of SEQ ID NOs: 6, 12, 28, 34, 46 or 48; most preferably, they have an amino acid sequence of SEQ ID NOs: 28 or 34.

[15] One of the grandparent molecules is the KAM of *Bacillus subtilis*, which had no detectible AAM activity. The DNA encoding this grandparent molecule was modified as described in WO 03/062173, entitled “Alanine 2,3-aminomutase,” to produce a polypeptide having a detectible alanine 2,3-aminomutase activity.

[16] In the present application, the applicants utilized as one parent molecule a polynucleotide sequence of SEQ ID NO: 58, which encoded the 471 residue polypeptide of SEQ ID NO: 59 and which exhibited an AAM activity of

approximately .001 U/mg (units/ mg of cell mass). The molecule of SEQ ID NO: 59 differs from the wild-type *B. subtilis* KAM, which had no detectible AAM activity, by having the following four (4) amino acid substitutions: L103M, M136V, Y140H and D339H.

[17] In yet another embodiment, the present invention is directed to a polypeptide having from about 1 to about 32 units of AAM activity and typically varying from the polypeptide of SEQ ID NO: 59 by 1-7 amino acid residues, more typically by 1-6 amino acid residues, even more typically by 1-5 amino acid residues, and most typically by 1-4 amino acid residues.

[18] In its second aspect, the present invention is directed to a polynucleotide sequence that encodes for the correspondingly referenced AAM polypeptide. Given the degeneracy of the genetic code, the present invention is also directed to any polynucleotide that encodes for the above referenced AAM polypeptides of the present invention. In another preferred embodiment, the present invention is directed to certain specific polynucleotides of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47 and 49 that encode for the novel AAM polypeptides of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48 and 51, respectively. Preferred polynucleotides encode for a polypeptide of SEQ ID NO: 2, 6, 12, 16, 20, 24, 28, 30, 32, 34, 38, 44, 46 or 48; more preferably they encode a polypeptide of SEQ ID NO: 6, 12, 28, 34, 46 or 48; most preferably, they have a polypeptide of sequence of SEQ ID NO: 28 or 34.

[19] In a third aspect, the present invention is directed to a nucleic acid construct, a vector, or a host cell comprising a polynucleotide sequence encoding an AAM polypeptide of the present invention operatively linked to a promoter.

[20] In a fourth aspect, the present invention is directed to a method of making an AAM polypeptide of the present invention comprising (a) cultivating a host cell transformed with a nucleic acid sequence encoding an AAM polypeptide of the present invention under conditions suitable for production of the polypeptide; and (b) providing glucose to the cultivated host cells under conditions suitable for the production of β -alanine. The β -alanine may be optionally recovered from the cells.

-7-

[21] In a fifth aspect, the present invention is directed to a method of producing b-alanine comprising (a) cultivating a host cell transformed with a nucleic acid sequence encoding an AAM polypeptide of the present invention under conditions suitable for production of the polypeptide; and (b) providing glucose to the cultivated host cells under conditions suitable for the production of b-alanine. The b-alanine may be optionally recovered from the cells.

BRIEF DESCRIPTION OF SEVERAL VIEWS OF THE DRAWINGS

[22] FIG. 1 shows the reversible reaction between alpha-alanine (*i.e.*, L-alanine or 2-aminopropionic acid) and beta-alanine (3-aminopropionic acid) that is catalyzed by alanine 2,3-aminomutase.

[23] FIG. 2 is a pathway for 3-hydroxypropionate (3-HP) synthesis from alpha-alanine, via beta-alanine as an intermediate.

[24] FIG. 3 is a 4036 bp expression vector (pCK110900-I Bla) of the present invention comprising a P15A origin of replication (P15A ori), a lacI repressor, a CAP binding site, a lac promoter (lac), a T7 ribosomal binding site (T7g10 RBS), and a chloramphenicol resistance gene (camR).

[25] FIGS. 4A-4J in combination provide an alignment chart of the amino acid sequences of four parental polypeptides that were used to produce the AAM of the present invention. The parental polypeptides were non-naturally occurring and derived in part from the KAM of *Clostridium stricklandii* (SEQ ID NO: 53), *Porphyromonas gingivalis* (SEQ ID NO: 55), *Fusobacterium nucleatum* (SEQ ID NO: 57), and *Bacillus subtilis* (SEQ ID NO: 59), respectively. The sequences of two wild-type KAM are disclosed in SEQ ID NOS: 60 (P GI2529467_G8_AAB81159.1_) and 61 (P GI2634361_EMB_CAB13860.1_). A consensus sequence is also provided as SEQ ID NO: 62).

[26] The foregoing summary, as well as the following detailed description of certain embodiments of the present invention, will be better understood when read in conjunction with the appended drawings. For the purpose of illustrating the invention, there is shown in the drawings, certain embodiments. It should be understood, however, that the present invention is not limited to the arrangements and instrumentality shown in the attached drawings.

DETAILED DESCRIPTION OF THE INVENTION

[27] The present invention has multiple aspects. In one aspect, the present invention is directed to a polypeptide having alanine 2,3-aminomutase (AAM) activity, preferably as measured by the assay of Example 8, and

(a) having a polypeptide selected from the group consisting of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48 and 51;

(b) having an amino acid sequence which has at least 98% homology, with the amino acid sequence selected from the group consisting of SEQ ID NO: 2, 22, 28, 32, and 36;

(c) having an amino acid sequence which has at least 99% homology, with the amino acid sequence selected from the group consisting of SEQ ID NO: 4, 6, 8, 12, 16, 24, 26, 30, 34 and 40;

(d) being a polypeptide encoded by a nucleic acid sequence which hybridizes under high stringency conditions with either (i) the nucleotide sequence of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 41, 43, 45, 47 or 49; (ii) a subsequence of (i) of at least 100 nucleotides, or (iii) a complementary strand of (i) or (ii) (J. Sambrook, E. F. Fritsch, and T. Maniatis, 1989, Molecular Cloning, A Laboratory Manual, 2d edition, Cold Spring Harbor, N.Y.); or

(e) being a variant of the polypeptide of (d) comprising a substitution, deletion, and/or insertion of one to six amino acids therefrom and having AAM activity from about 1 to about 30 μ M β -alanine produced /hour 1 cell OD at pH 7.0-7.6, 25°C.

[28] Collectively, the polypeptides of (b) and (c) above are referred to herein as "homologous polypeptides." For purposes of the present invention, the degree of homology between two amino acid sequences is expressed as "percent homology," "percent identity," "% identity," "percent identical," and "% identical" are used interchangeably herein to refer to the percent amino acid sequence identity that is obtained by ClustalW analysis (version W 1.8 available from European Bioinformatics Institute, Cambridge, UK), counting the number of identical matches in the alignment and dividing such number of identical matches by the length of the reference sequence, and using the following default ClustalW parameters to achieve slow/accurate pairwise optimal alignments – Gap Open Penalty:10; Gap Extension

-10-

Penalty:0.10; Protein weight matrix: Gonnet series; DNA weight matrix: IUB;
Toggle Slow/Fast pairwise alignments = SLOW or FULL Alignment.

[29] AAM polypeptides are sensitive to oxygen and are preferably maintained and used in an oxygen deficient environment. If the AAM polypeptide becomes inactivated due to exposure to oxygen, it can be activated by anaerobic incubation with a sulfhydryl compound for one hour at 37°C in accordance with the method described in Chirpich, et al., Journal Biol. Chem., 245(7): 1778-1789 (1970), which is incorporated herein by reference in its entirety. AAM polypeptides of the present invention are preferably utilized in whole cell form (i.e., as a whole cell transformed with an AAM polynucleotide that is used under conditions such that the encoded AAM polypeptide is expressed in the cell) or alternatively, both isolated and utilized under anoxic conditions. AAM polypeptides of the present invention may be isolated, and optionally purified, under anaerobic conditions (e.g., under a nitrogen atmosphere) in accordance with the method described in Petrovich, et al., Journal Biol. Chem., 266(12):7656-7660 (1991), which describes the isolation and purification of lysine-2,3-aminomutase and which is incorporated herein by reference in its entirety. As used herein, the term "anoxic" refers to oxygen deficient. The AAM polypeptides in whole cell form or as isolated enzymes may be lyophilized. In yet another embodiment, the present invention is directed to a composition comprising an AAM polypeptide as described herein (e.g., in whole cell form or as an isolated polypeptide) and a suitable carrier, typically a buffer, more typically an aqueous buffer solution having a pH from about 6.0 to about 8.0. It is also within the scope of the present invention that the aqueous buffered composition be lyophilized to provide a composition in a lyophilized form, wherein the composition is reconstituted by the addition of an aqueous based composition.

[30] In one embodiment, the present invention is also directed to an AAM polypeptide as described herein in isolated and purified form.

[31] In another embodiment, the present invention is directed to an AAM polypeptide as described herein in lyophilized form. Lyophilization is performed using standard lyophilization equipment. Typically, a solution containing the polypeptide is dispensed in an appropriate sized vial, frozen and placed under reduced

pressure to cause the water to evaporate, leaving the lyophilized (freeze-dried) polypeptide behind. Prior to use, the lyophilized polypeptide is reconstituted with distilled water or an appropriate buffer solution.

[32] In yet another embodiment, the present invention is directed to a composition comprising an AAM polypeptide as described herein and a suitable carrier, typically a buffer solution, more typically an aqueous buffer solution having a pH between 6.0 and 8.0. The composition may also be in a lyophilized form.

[33] The novel AAM polypeptides of the present invention have significantly enhanced AAM activity relative to the wild-type KAM polypeptides from which they are ultimately derived. By significantly enhanced AAM activity is meant that the AAM polypeptide of the present invention has an AAM activity within the range of about 1 to about 32 μM β -alanine produced/hour 1 cell OD (units), preferably from about 10 to about 32 units, more preferably from about 20 to about 32 units; most preferably from about 25 to about 32 units.

[34] Table 1 provides a chart showing the AAM activities of the various AAM polypeptides of the present invention, identified by their clone number and SEQ ID NO. In Table 1, the $\text{OD}_{600\text{nm}}$ is reported at harvest after 5 hours ($t=5$) of incubation. Table 1 also reports the total μM of β -alanine produced after 5 hours per 1 cell OD. Finally, the last column of Table 1 reports the rate of β -alanine (μM) produced/hr /1 cell OD.

Table 1

Seq. ID No.	Harvest OD _{600nm} t= 5	uM β -alanine produced at t=5/1 cell OD	Rate of β -alanine(uM) produced /hr 1 Cell OD
34	1.0	159.7	31.9
10	3.7	31.7	6.3
38	4.0	54.9	11.0
20	3.0	73.4	14.7
14	3.7	33.5	7.7
22	2.2	4.8	1.0
42	5.0	17.5	3.5
26	3.7	23.9	4.8
18	4.7	19.3	3.9
44	2.9	64.4	12.9
51	3.7	35.0	7.0
36	3.0	29.8	6.0
48	1.1	110.1	22.0
12	4.7	17.8	3.6
4	3.7	22.4	4.5
16	1.0	136.0	19.4
24	1.4	94.7	18.9
46	1.7	107.6	20.7
28	1.5	148.0	29.2
40	1.4	14.6	2.9
32	1.6	93.2	13.6
2	1.5	87.5	17.5
30	2.7	72.6	14.3
6	1.7	125.7	23.0

[35] Preferred AAM polypeptides of the present invention have an amino acid sequences of SEQ ID NOs: 2, 6, 12, 16, 20, 24, 28, 30, 32, 34, 38, 44, 46 or 48; more preferably they have an amino acid sequence of SEQ ID NOs: 6, 12, 28, 34, 46 or 48; most preferably, they have an amino acid sequence of SEQ ID NOs: 28 or 34.

[36] The ultimate grandparent molecule is the KAM of *Bacillus subtilis*, which had no detectible AAM activity. The DNA encoding this grandparent molecule was modified as described in WO 03/062173, entitled "Alanine 2,3-aminomutase," to produce a polypeptide having a detectible alanine 2,3-aminomutase activity.

[37] In the present application, the applicants utilized as one parent molecule a polynucleotide of SEQ ID NO: 58, which encoded the 471 residue polypeptide of SEQ ID NO: 59 and which exhibited an AAM activity of approximately .001 U/mg (units/ mg of cell mass). The molecule of SEQ ID NO: 59 differs from the wild-type *B. subtilis* KAM (SEQ ID NO: 60), which had no detectible AAM activity, by having the following four (4) amino acid substitutions: L103M, M136V, Y140H and D339H.

[38] Other grandparent molecules utilized as starting materials in the present invention were the DNA sequences from other microorganisms (e.g., *Porphyromonas gingivalis*, *Fusobacterium nucleatum*, and *Clostridium sticklandii*) that encoded a KAM polypeptide. These DNA sequences were modified using standard techniques to introduce point substitutions that ultimately produced a KAM polypeptide that also had a detectible cross-reactivity with α -alanine. One such parent molecule that was derived from *Porphyromonas gingivalis* is the polynucleotide of SEQ ID NO: 54 which encodes the 416 residue polypeptide of SEQ ID NO: 55. The parental polypeptide of SEQ ID NO: 55 differs from the wild-type *Porphyromonas gingivalis* KAM by having the following seven (7) amino acid substitutions: N19Y, E30K, L53P, H85Q, I192V, D331G, and M342T. Another such parent molecule that was derived from *F. nucleatum* is the polynucleotide of SEQ ID NO: 56 which encodes the 425 residue polypeptide of SEQ ID NO: 57.

[39] Yet another parent polynucleotide was derived by modification of the polynucleotide in *C. stricklandii* that encodes KAM. The resulting parental polynucleotide, which has a detectable cross-reactivity with α -alanine, is the polynucleotide of SEQ ID NO: 52 which encodes the 416 residue polypeptide of SEQ ID NO: 53.

[40] The above described parental polypeptides of SEQ ID NOs: 53, 55, 57 and 58 are compared in the alignment chart of FIG. 4. From the alignment chart, it can be seen that the KAMs from *P. gingivalis*, *C. stricklandii*, and *F. nucleatum* are truncated at the N-terminus and at the C-terminus relative to the KAM from *B. subtilis*, while between the four species, about 40% of the residue positions in the central portion of the KAM polypeptide are conserved. Based upon the truncated species in the alignment chart of FIG. 4, it can be inferred that the first 8 amino acid residues at the

-14-

N-terminus of SEQ ID NO: 58 and the last 40 residues at the C-terminus of SEQ ID NO: 58 are not necessary for KAM activity, or the AAM activity that is derived therefrom. In FIG. 4, there is also provided a consensus sequence.

[41] The AAM polypeptide molecules of the present invention with their enhanced AAM activity were made by applying directed evolution techniques to the above-described parental molecules. These techniques are described in further detail herein.

[42] In yet another aspect, the present invention is directed to AAM polypeptides that have enhanced activity in coupled reactions.

[43] In another embodiment, the present invention is directed to an AAM a polypeptide encoded by a nucleic acid sequence which hybridizes under high stringency conditions with either (i) the nucleotide sequence of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 41, 43, 45, 47 or 49; (ii) a subsequence of (i) of at least 100 nucleotides, or (iii) a complementary strand of (i) or (ii) (J. Sambrook, E. F. Fritsch, and T. Maniatis, 1989, Molecular Cloning, A Laboratory Manual, 2d edition, Cold Spring Harbor, N.Y.). For polynucleotides of at least 100 nucleotides in length, low to very high stringency conditions are defined as prehybridization and hybridization at 42°C in 5x SSPE, 0.3% SDS, 200 µg/ml sheared and denatured salmon sperm DNA, and either 25% formamide for low stringencies, 35% formamide for medium and medium-high stringencies, or 50% formamide for high and very high stringencies, following standard Southern blotting procedures.

[44] For polynucleotides of at least 100 nucleotides in length, the carrier material is finally washed three times each for 15 minutes using 2x SSC, 0.2% SDS at least at 50°C (low stringency), at least at 55°C (medium stringency), at least at 60°C. (medium-high stringency), at least at 65°C (high stringency), and at least at 70°C. (very high stringency).

[45] In another embodiment, the present invention is directed to a variant of the polypeptide of (d) comprising a substitution, deletion, and/or insertion of one to six amino acids therefrom and having AAM activity from about 1 to about 30 µM β-alanine produced /hour 1 cell OD at pH 7.0-7.6, 25°C, such as determined by the method of Example 8. Preferably, amino acid changes are of a minor nature, that is

-15-

conservative amino acid substitutions that do not significantly affect the folding and/or activity of the protein; small deletions, typically of one to six amino acids; small amino- or carboxyl-terminal extensions; a small linker peptide; or a small extension that facilitates purification by changing net charge or another function, such as a poly-histidine tract, an antigenic epitope or a binding domain.

[46] Examples of conservative substitutions are within the group of basic amino acids (arginine, lysine and histidine), acidic amino acids (glutamic acid and aspartic acid), polar amino acids (glutamine and asparagine), hydrophobic amino acids (leucine, isoleucine and valine), aromatic amino acids (phenylalanine, tryptophan and tyrosine), and small amino acids (glycine, alanine, serine, threonine, proline, cysteine and methionine). Amino acid substitutions, which do not generally alter the specific activity are known in the art and are described, for example, by H. Neurath and R. L. Hill, 1979, In, The Proteins, Academic Press, New York. The most commonly occurring exchanges are Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Val, Ala/Glu, and Asp/Gly as well as these in reverse.

[47] In another embodiment, the present invention is directed to a fragment of (a), (b) or (c), as described above in the first paragraph of the Detailed Description, that has from about 1 to about 30 μ M β -alanine produced /hour 1 cell OD at pH 7.0-7.6, 25°C, such as determined by the method of Example 8. By the term "fragment" is meant that the polypeptide has a deletion of 1 to 8 amino acid residues from the N-terminus or 1-40 residues from the C-terminus, or both. Preferably, the deletion is 1 to 20 residues from the C-terminus, more preferably, the deletion is 1 to 10 residues from the C-terminus.

Polynucleotides

[48] In its second aspect, the present invention is directed to a polynucleotide sequence that encodes for an AAM polypeptide of the present invention. Given the degeneracy of the genetic code, the present invention is also directed to any polynucleotide that encodes for the above referenced AAM polypeptides of the present invention. In its second aspect, the present invention is directed to a

-16-

polynucleotide sequence that encodes for the correspondingly referenced AAM polypeptide. Given the degeneracy of the genetic code, the present invention is also directed to any polynucleotide that encodes for the above referenced AAM polypeptides of the present invention. In a preferred embodiment, the present invention is directed to certain specific polynucleotides of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47 and 49 that encode for the novel AAM polypeptides of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48 and 51, respectively. Preferred polynucleotides encode for a polypeptide of SEQ ID NO: 2, 6, 12, 16, 20, 24, 28, 30, 32, 34, 38, 44, 46 or 48; more preferably they encode a polypeptide of SEQ ID NO: 6, 12, 28, 34, 46 or 48; most preferably, they have a polypeptide of sequence of SEQ ID NO: 28 or 34.

[49] To make the improved AAM polypeptides of the present invention, one starts with one or more wild-type polynucleotides that encode a KAM polypeptide. The term "wild-type" polynucleotide means that the nucleic acid fragment does not comprise any mutations from the form isolated from nature. The term "wild-type" protein means that the protein will be active at a level of activity found in nature and typically will comprise the amino acid sequence as found in nature. Thus, the term "wild type" or "grand-parent sequence" indicates a starting or reference sequence prior to a manipulation of the invention.

[50] Suitable sources of wild-type KAM as a starting material to be improved is readily identified by screening genomic libraries for the KAM activity. A particularly suitable source of KAM is the *yodO* gene of *Bacillus sp.* bacteria as found in nature. Using the published KAM gene sequences for *B. subtilis* (e.g., WO 03 0623173 A2), primers for amplification of the genes from their respective gene libraries were created using conventional techniques. One such technique for isolating the KAM of *B. subtilis* is disclosed in Chen *et al.*, "A novel lysine 2,3-aminomutase encoded by the *yodO* gene of *Bacillus subtilis*: characterization on observation of organic radical intermediates," Biochem J. 348:539-549 (2000), which is incorporated herein by reference.

[51] The starting polynucleotides of SEQ ID NOs: 52, 54, 56 and 58 were obtained using the techniques disclosed in WO 03 0623173 A2 which is incorporated herein by reference for the disclosure of those techniques as recited in the examples therein. Specifically, WO 03 0623173 A2 discloses a *B. subtilis* wild-type lysine 2,3-aminomutase (KAM), and a mutated form thereof, which encodes an alanine 2,3-aminomutase (AAM). In addition, WO 03 0623173 A2 also discloses a *P. gingivalis* wild-type lysine 2,3-aminomutase (KAM) and a mutated form thereof, which encodes an alanine 2,3-aminomutase (AAM).

[52] Beginning with the polynucleotide of SEQ ID NO: 58, a non-naturally occurring and mutated and/or evolved enzyme, having unknown AAM activity is generated using any one of the well-known mutagenesis or directed evolution methods. See, e.g., Ling, et al., "Approaches to DNA mutagenesis: an overview," Anal. Biochem., 254(2):157-78 (1997); Dale, et al., "Oligonucleotide-directed random mutagenesis using the phosphorothioate method," Methods Mol. Biol., 57:369-74 (1996); Smith, "In vitro mutagenesis," Ann. Rev. Genet., 19:423-462 (1985); Botstein, et al., "Strategies and applications of in vitro mutagenesis," Science, 229:1193-1201 (1985); Carter, "Site-directed mutagenesis," Biochem. J., 237:1-7 (1986); Kramer, et al., "Point Mismatch Repair," Cell, 38:879-887 (1984); Wells, et al., "Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites," Gene, 34:315-323 (1985); Minshull, et al., "Protein evolution by molecular breeding," Current Opinion in Chemical Biology, 3:284-290 (1999); Christians, et al., "Directed evolution of thymidine kinase for AZT phosphorylation using DNA family shuffling," Nature Biotechnology, 17:259-264 (1999); Crameri, et al., "DNA shuffling of a family of genes from diverse species accelerates directed evolution," Nature, 391:288-291; Crameri, et al., "Molecular evolution of an arsenate detoxification pathway by DNA shuffling," Nature Biotechnology, 15:436-438 (1997); Zhang, et al., "Directed evolution of an effective fucosidase from a galactosidase by DNA shuffling and screening," Proceedings of the National Academy of Sciences, U.S.A., 94:454-4509; Crameri, et al., "Improved green fluorescent protein by molecular evolution using DNA shuffling," Nature Biotechnology, 14:315-319 (1996); Stemmer, "Rapid evolution of a protein in vitro by DNA shuffling," Nature, 370:389-391 (1994); Stemmer, "DNA shuffling by

random fragmentation and reassembly: *In vitro* recombination for molecular evolution," Proceedings of the National Academy of Sciences, U.S.A., 91:10747-10751 (1994); WO 95/22625; WO 97/0078; WO 97/35966; WO 98/27230; WO 00/42651; WO 01/75767 and U.S. Pat. 6,537,746 which issued to Arnold, *et al.* on March 25, 2003 and is entitled "Method for creating polynucleotide and polypeptide sequences."

[53] Any of these methods can be applied to generate AAM polynucleotides. To maximize any diversity, several of the above-described techniques can be used sequentially. Typically, a library of shuffled polynucleotides is created by one mutagenic or evolutionary technique and their expression products are screened to find the polypeptides having the highest AAM activity. Then, a second mutagenic or evolutionary technique is applied to polynucleotides encoding the most active polypeptides to create a second library, which in turn is screened for AAM activity by the same technique. The process of mutating and screening can be repeated as many times as needed, including the insertion of point mutations, to arrive at a polynucleotide that encodes a polypeptide with the desired activity, thermostability, or cofactor preference.

[54] Alternatively, polynucleotides and oligonucleotides of the invention can be prepared by standard solid-phase methods, according to known synthetic methods. Typically, fragments of up to about 100 bases are individually synthesized, then joined (e.g., by enzymatic or chemical ligation methods, or polymerase mediated methods) to form essentially any desired continuous sequence. For example, polynucleotides and oligonucleotides of the invention can be prepared by chemical synthesis using, e.g., the classical phosphoramidite method described by *Beaucage et al.* (1981) *Tetrahedron Letters* 22:1859-69, or the method described by *Matthes et al.* (1984) *EMBO J.* 3:801-05, e.g., as it is typically practiced in automated synthetic methods. According to the phosphoramidite method, oligonucleotides are synthesized, e.g., in an automatic DNA synthesizer, purified, annealed, ligated and cloned in appropriate vectors.

[55] In addition, essentially any nucleic acid can be custom ordered from any of a variety of commercial sources, such as The Midland Certified Reagent Company,

Midland, TX, The Great American Gene Company, Ramona, CA, ExpressGen Inc., Chicago, IL, Operon Technologies Inc., Alameda, CA, all of which have internet web sites, and many others. Similarly, peptides and antibodies can be custom ordered from any of a variety of sources, such as PeptidoGenic, HTI Bio-products, Inc., BMA Biomedicals Ltd. (U.K.), Bio.Synthesis, Inc., and many others.

[56] Polynucleotides may also be synthesized by well-known techniques as described in the technical literature. See, e.g., *Carruthers et al.*, *Cold Spring Harbor Symp. Quant. Biol.* 47:411-418 (1982), and *Adams et al.*, *J. Am. Chem. Soc.* 105:661 (1983). Double stranded DNA fragments may then be obtained either by synthesizing the complementary strand and annealing the strands together under appropriate conditions, or by adding the complementary strand using DNA polymerase with an appropriate primer sequence.

[57] General texts which describe molecular biological techniques useful herein, including mutagenesis, include Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology, volume 152 Academic Press, Inc., San Diego, CA ("Berger"); *Sambrook et al.*, Molecular Cloning - A Laboratory Manual (2nd Ed.), volumes 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989 ("Sambrook"); and Current Protocols in Molecular Biology, F.M. Ausubel et al., eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc. (supplemented through 2000) ("Ausubel"). Examples of techniques sufficient to direct persons of skill through *in vitro* amplification methods, including the polymerase chain reaction (PCR) the ligase chain reaction (LCR), Q β -replicase amplification and other RNA polymerase mediated techniques (e.g., NASBA) are found in Berger, Sambrook, and Ausubel, as well as *Mullis et al.*, (1987) U.S. Patent No. 4,683,202; PCR Protocols A Guided to Methods and Applications (*Innis et al.*, eds.) Academic Press Inc. San Diego, CA (1990); Arnheim & Levinson (October 1, 1990) Chemical and Engineering News 36-47; The Journal Of NIH Research (1991) 3:81-94; *Kwoh et al.* (1989) Proc. Natl. Acad. Sci. USA 86:1173; *Guatelli et al.* (1990) Proc. Natl. Acad. Sci. USA 87:1874; *Lomell et al.* (1989) J. Clin. Chem. 35:1826; *Landegren et al.*, (1988) Science 241:1077-1080; Van Brunt (1990) Biotechnology 8:291-294; Wu and Wallace, (1989) Gene 4:560; *Barringer et*

al. (1990) Gene 89:117, and Sooknanan and Malek (1995) Biotechnology 13:563-564. Improved methods of cloning *in vitro* amplified nucleic acids are described in *Wallace et al.*, U.S. Pat. No. 5,426,039. Improved methods of amplifying large nucleic acids by PCR are summarized in *Cheng et al.* (1994) Nature 369:684-685 and the references therein, in which PCR amplicons of up to 40kb are generated. One of skill will appreciate that essentially any RNA can be converted into a double stranded DNA suitable for restriction digestion, PCR expansion and sequencing using reverse transcriptase and a polymerase. *See*, Ausubel, Sambrook and Berger, *all supra*.

[58] It will be appreciated by those skilled in the art due to the degeneracy of the genetic code, a multitude of nucleotide sequences encoding AAM polypeptides of the invention may be produced, some of which bear substantial identity to the nucleic acid sequences explicitly disclosed herein. It is also within the scope of the present invention that the polynucleotides encoding the AAM polypeptides of the present invention may be codon optimized for optimal production from the host organism selected for expression. Those having ordinary skill in the art will recognize that tables and other references providing codon preference information for a wide range of organisms are readily available. *See e.g.*, Henaut and Danchin, "*Escherichia coli* and *Salmonella*," Neidhardt, et al. Eds., ASM Press, Washington D.C., p. 2047-2066 (1996).

[59] It is to be noted that expression in *E. coli* is different than in other organisms. For example, in the present invention, the codon (tgg) encodes Trp (W) for residue position 31 in the parent polypeptide of SEQ ID NO: 59. However, the corresponding codon for residue position 31 is "tga" in each of the progeny polynucleotides of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, and 47 encoding for the AAM polypeptides of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, and 48, respectively. One skilled in the art recognizes that the codon "tga" is usually a stop (nonsense) codon. However, in the present expression system used in the Δ panD *E. coli* strain, and under the selection conditions imposed, this codon is read through by the *E. coli* as a sense codon and is expressed, presumably as Trp (W). Others have reported that "tga" is the weakest stop codon for *E. coli* and that it is often read through as a sense codon

for Trp (W) in high expression. See e.g., Parker, J., "Errors and Alternatives in Reading the universal Genetic Code," Microbiological Reviews, 53(3): 273-298 (1989); Roth, J., "UGA Nonsense Mutations in *Salmonella typhimurium*," J. of Bacteriology, 102(2):467-475 (1970); and McBeath, G. and Kast, P., "UGA Read-Through Artifacts—When Popular Gene Expression Systems Need a Patch," BioTechniques, 24:789-794 (May 1998), which are incorporated herein by reference. Hence for expression in non-*E. coli* systems, it would be advantageous to alter the codon (tga) at residue position 31 to "tgg" which is the universal sense codon for Trp (W).

[60] In SEQ ID NO: 49, the codon encoding for residue 72 is "tag" which is read as a stop codon. However, two fragments are produced. The first fragment, having residues 1-71 of SEQ ID NO: 50, does not have any detectable AAM activity. The second fragment that is produced begins with residue 73 (Val) instead of the usual Met. This second fragment has 399 residues (SEQ ID NO: 51) and does have significant AAM activity (see Table 2) based upon the assay of Example 8. Thus, the first 72 residues at the N-terminus of the AAM polypeptide (based upon the consensus sequence or the parental KAM sequence from *B. subtilis*) are not absolutely necessary for AAM activity.

[61] In the present case, several round No. 1 libraries were created by applying a variety of mutagenic techniques to the polynucleotides of SEQ ID NOs: 52, 54, 56 and 58.

[62] In its third aspect, the present invention is directed to an expression vector and to a host cell comprising a polynucleotide of the present invention operatively linked to a control sequence. To obtain expression of the variant gene encoding an AAM polypeptide, the variant gene was first operatively linked to one or more heterologous regulatory sequences that control gene expression to create a nucleic acid construct, such as an expression vector or expression cassette. Thereafter, the resulting nucleic acid construct, such as an expression vector or expression cassette, was inserted into an appropriate host cell for ultimate expression of the AAM polypeptide encoded by the shuffled gene. A "nucleic acid construct" is defined herein as a nucleic acid molecule, either single- or double-stranded, which is isolated from a naturally

occurring gene or which has been modified to contain segments of nucleic acid combined and juxtaposed in a manner that would not otherwise exist in nature. Thus, in one aspect, the present invention is directed to a nucleic acid construct comprising a polynucleotide encoding an AAM polypeptide of the present invention.

[63] The term “nucleic acid construct” is synonymous with the term “expression cassette” when the nucleic acid construct contains all the control sequences required for expression of a coding sequence of the present invention. The term “coding sequence” is defined herein as a nucleic acid sequence, which directly specifies the amino acid sequence of its protein product. A coding sequence can include, but is not limited to, DNA, cDNA, and recombinant nucleic acid sequences.

[64] An isolated polynucleotide encoding an AAM polypeptide of the present invention may be manipulated in a variety of ways to provide for expression of the polypeptide. Manipulation of the isolated polynucleotide prior to its insertion into a vector may be desirable or necessary depending on the expression vector. The techniques for modifying polynucleotides and nucleic acid sequences utilizing recombinant DNA methods are well known in the art.

[65] The term “control sequence” is defined herein to include all components, which are necessary or advantageous for the expression of a polypeptide of the present invention. Each control sequence may be native or foreign to the nucleic acid sequence encoding the polypeptide. Such control sequences include, but are not limited to, a leader, polyadenylation sequence, propeptide sequence, promoter, signal peptide sequence, and transcription terminator. At a minimum, the control sequences include a promoter, and transcriptional and translational stop signals. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding region of the nucleic acid sequence encoding a polypeptide.

[66] The term “operably linked” is defined herein as a configuration in which a control sequence is appropriately placed at a position relative to the coding sequence of the DNA sequence such that the control sequence directs the expression of a polypeptide.

[67] The control sequence may be an appropriate promoter sequence. The "promoter sequence" is a relatively short nucleic acid sequence that is recognized by a host cell for expression of the longer coding region that follows. The promoter sequence contains transcriptional control sequences, which mediate the expression of the polypeptide. The promoter may be any nucleic acid sequence which shows transcriptional activity in the host cell of choice including mutant, truncated, and hybrid promoters, and may be obtained from genes encoding extracellular or intracellular polypeptides either homologous or heterologous to the host cell.

[68] For bacterial host cells, suitable promoters for directing the transcription of the nucleic acid constructs of the present invention, include the promoters obtained from the *E. coli* lac operon, *Streptomyces coelicolor* agarase gene (dagA), *Bacillus subtilis* levansucrase gene (sacB), *Bacillus licheniformis* alpha-amylase gene (amyL), *Bacillus stearothermophilus* maltogenic amylase gene (amyM), *Bacillus amyloliquefaciens* alpha-amylase gene (amyQ), *Bacillus licheniformis* penicillinase gene (penP), *Bacillus subtilis* xylA and xylB genes, and prokaryotic beta-lactamase gene (Villa-Kamaroff et al., 1978, Proceedings of the National Academy of Sciences USA 75: 3727-3731), as well as the tac promoter (DeBoer et al., 1983, Proceedings of the National Academy of Sciences USA 80: 21-25). Further promoters are described in "Useful proteins from recombinant bacteria" in Scientific American, 1980, 242: 74-94; and in Sambrook et al., 1989, *supra*.

[69] For filamentous fungal host cells, suitable promoters for directing the transcription of the nucleic acid constructs of the present invention include promoters obtained from the genes for *Aspergillus oryzae* TAKA amylase, *Rhizomucor miehei* aspartic proteinase, *Aspergillus niger* neutral alpha-amylase, *Aspergillus niger* acid stable alpha-amylase, *Aspergillus niger* or *Aspergillus awamori* glucoamylase (glaA), *Rhizomucor miehei* lipase, *Aspergillus oryzae* alkaline protease, *Aspergillus oryzae* triose phosphate isomerase, *Aspergillus nidulans* acetamidase, and *Fusarium oxysporum* trypsin-like protease (WO 96/00787), as well as the NA2-tpi promoter (a hybrid of the promoters from the genes for *Aspergillus niger* neutral alpha-amylase and *Aspergillus oryzae* triose phosphate isomerase), and mutant, truncated, and hybrid promoters thereof.

[70] In a yeast host, useful promoters are obtained from the genes for *Saccharomyces cerevisiae* enolase (ENO-1), *Saccharomyces cerevisiae* galactokinase (GAL1), *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH2/GAP), and *Saccharomyces cerevisiae* 3-phosphoglycerate kinase. Other useful promoters for yeast host cells are described by Romanos et al., 1992, *Yeast* 8:423-488.

[71] The control sequence may also be a suitable transcription terminator sequence, a sequence recognized by a host cell to terminate transcription. The terminator sequence is operably linked to the 3' terminus of the nucleic acid sequence encoding the polypeptide. Any terminator, which is functional in the host cell of choice, may be used in the present invention.

[72] Preferred terminators for filamentous fungal host cells are obtained from the genes for *Aspergillus oryzae* TAKA amylase, *Aspergillus niger* glucoamylase, *Aspergillus nidulans* anthranilate synthase, *Aspergillus niger* alpha-glucosidase, and *Fusarium oxysporum* trypsin-like protease.

[73] Preferred terminators for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* enolase, *Saccharomyces cerevisiae* cytochrome C (CYC1), and *Saccharomyces cerevisiae* glyceraldehyde-3-phosphate dehydrogenase. Other useful terminators for yeast host cells are described by Romanos et al., 1992, *supra*.

[74] The control sequence may also be a suitable leader sequence, a nontranslated region of an mRNA which is important for translation by the host cell. The leader sequence is operably linked to the 5' terminus of the nucleic acid sequence encoding the polypeptide. Any leader sequence that is functional in the host cell of choice may be used in the present invention. Preferred leaders for filamentous fungal host cells are obtained from the genes for *Aspergillus oryzae* TAKA amylase and *Aspergillus nidulans* triose phosphate isomerase. Suitable leaders for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* enolase (ENO-1), *Saccharomyces cerevisiae* 3-phosphoglycerate kinase, *Saccharomyces cerevisiae* alpha-factor, and *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH2/GAP).

[75] The control sequence may also be a polyadenylation sequence, a sequence operably linked to the 3' terminus of the nucleic acid sequence and which, when transcribed, is recognized by the host cell as a signal to add polyadenosine residues to transcribed mRNA. Any polyadenylation sequence that is functional in the host cell of choice may be used in the present invention. Preferred polyadenylation sequences for filamentous fungal host cells are obtained from the genes for *Aspergillus oryzae* TAKA amylase, *Aspergillus niger* glucoamylase, *Aspergillus nidulans* anthranilate synthase, *Fusarium oxysporum* trypsin-like protease, and *Aspergillus niger* alpha-glucosidase. Useful polyadenylation sequences for yeast host cells are described by Guo and Sherman, 1995, Molecular Cellular Biology 15: 5983-5990.

[76] The control sequence may also be a signal peptide coding region that codes for an amino acid sequence linked to the amino terminus of a polypeptide and directs the encoded polypeptide into the cell's secretory pathway. The 5' end of the coding sequence of the nucleic acid sequence may inherently contain a signal peptide coding region naturally linked in translation reading frame with the segment of the coding region that encodes the secreted polypeptide. Alternatively, the 5' end of the coding sequence may contain a signal peptide coding region that is foreign to the coding sequence. The foreign signal peptide coding region may be required where the coding sequence does not naturally contain a signal peptide coding region.

[77] Alternatively, the foreign signal peptide coding region may simply replace the natural signal peptide coding region in order to enhance secretion of the polypeptide. However, any signal peptide coding region that directs the expressed polypeptide into the secretory pathway of a host cell of choice may be used in the present invention.

[78] Effective signal peptide coding regions for bacterial host cells are the signal peptide coding regions obtained from the genes for *Bacillus* NCIB 11837 maltogenic amylase, *Bacillus stearothermophilus* alpha-amylase, *Bacillus licheniformis* subtilisin, *Bacillus licheniformis* beta-lactamase, *Bacillus stearothermophilus* neutral proteases (nprT, nprS, nprM), and *Bacillus subtilis* prsA. Further signal peptides are described by Simonen and Palva, 1993, Microbiological Reviews 57: 109-137.

[79] Effective signal peptide coding regions for filamentous fungal host cells are the signal peptide coding regions obtained from the genes for *Aspergillus oryzae*

TAKA amylase, *Aspergillus niger* neutral amylase, *Aspergillus niger* glucoamylase, *Rhizomucor miehei* aspartic proteinase, *Humicola insolens* cellulase, and *Humicola lanuginosa* lipase.

[80] Useful signal peptides for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* alpha-factor and *Saccharomyces cerevisiae* invertase. Other useful signal peptide coding regions are described by Romanos *et al.*, 1992, *supra*.

[81] The control sequence may also be a propeptide coding region that codes for an amino acid sequence positioned at the amino terminus of a polypeptide. The resultant polypeptide is known as a proenzyme or propolypeptide (or a zymogen in some cases). A propolypeptide is generally inactive and can be converted to a mature active polypeptide by catalytic or autocatalytic cleavage of the propeptide from the propolypeptide. The propeptide coding region may be obtained from the genes for *Bacillus subtilis* alkaline protease (aprE), *Bacillus subtilis* neutral protease (nprT), *Saccharomyces cerevisiae* alpha-factor, *Rhizomucor miehei* aspartic proteinase, and *Myceliophthora thermophila* lactase (WO 95/33836).

[82] Where both signal peptide and propeptide regions are present at the amino terminus of a polypeptide, the propeptide region is positioned next to the amino terminus of a polypeptide and the signal peptide region is positioned next to the amino terminus of the propeptide region.

[83] It may also be desirable to add regulatory sequences, which allow the regulation of the expression of the polypeptide relative to the growth of the host cell. Examples of regulatory systems are those which cause the expression of the gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. In prokaryotic host cells, suitable regulatory sequences include the lac, tac, and trp operator systems. In yeast host cells, suitable regulatory systems include the ADH2 system or GAL1 system. In filamentous fungi, suitable regulatory sequences include the TAKA alpha-amylase promoter, *Aspergillus niger* glucoamylase promoter, and *Aspergillus oryzae* glucoamylase promoter.

-27-

[84] Other examples of regulatory sequences are those which allow for gene amplification. In eukaryotic systems, these include the dihydrofolate reductase gene, which is amplified in the presence of methotrexate, and the metallothionein genes, which are amplified with heavy metals. In these cases, the nucleic acid sequence encoding the AAM polypeptide of the present invention would be operably linked with the regulatory sequence.

Expression Vectors

[85] In another aspect, the present invention is also directed to a recombinant expression vector comprising a polynucleotide of the present invention (which encodes an AAM polypeptide of the present invention), and one or more expression regulating regions. An expression regulating region includes a promoter, a terminator, a replication origin, etc., depending on the type of hosts into which they are to be introduced. The various nucleic acid and control sequences described above may be joined together to produce a recombinant expression vector which may include one or more convenient restriction sites to allow for insertion or substitution of the nucleic acid sequence encoding the polypeptide at such sites. Alternatively, the nucleic acid sequence of the present invention may be expressed by inserting the nucleic acid sequence or a nucleic acid construct comprising the sequence into an appropriate vector for expression. In creating the expression vector, the coding sequence is located in the vector so that the coding sequence is operably linked with the appropriate control sequences for expression.

[86] The recombinant expression vector may be any vector (e.g., a plasmid or virus), which can be conveniently subjected to recombinant DNA procedures and can bring about the expression of the polynucleotide sequence. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vectors may be linear or closed circular plasmids.

[87] The expression vector may be an autonomously replicating vector, *i.e.*, a vector that, exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, *e.g.*, a plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The vector may contain

-28-

any means for assuring self-replication. Alternatively, the vector may be one which, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. Furthermore, a single vector or plasmid or two or more vectors or plasmids which together contain the total DNA to be introduced into the genome of the host cell, or a transposon may be used.

[88] The expression vector of the present invention preferably contains one or more selectable markers, which permit easy selection of transformed cells. A selectable marker is a gene the product of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, and the like. Examples of bacterial selectable markers are the *dal* genes from *Bacillus subtilis* or *Bacillus licheniformis*, or markers, which confer antibiotic resistance such as ampicillin, kanamycin, chloramphenicol (Example 1) or tetracycline resistance. Suitable markers for yeast host cells are ADE2, HIS3, LEU2, LYS2, MET3, TRP1, and URA3.

[89] Selectable markers for use in a filamentous fungal host cell include, but are not limited to, *amdS* (acetamidase), *argB* (ornithine carbamoyltransferase), *bar* (phosphinothricin acetyltransferase), *hph* (hygromycin phosphotransferase), *niaD* (nitrate reductase), *pyrG* (orotidine-5'-phosphate decarboxylase), (sulfate adenylyltransferase), and *trpC* (anthranilate synthase), as well as equivalents thereof. Preferred for use in an *Aspergillus* cell are the *amdS* and *pyrG* genes of *Aspergillus nidulans* or *Aspergillus oryzae* and the *bar* gene of *Streptomyces hygroscopicus*.

[90] The vectors of the present invention preferably contain an element(s) that permits integration of the vector into the host cell's genome or autonomous replication of the vector in the cell independent of the genome. For integration into the host cell genome, the vector may rely on the nucleic acid sequence encoding the polypeptide or any other element of the vector for integration of the vector into the genome by homologous or nonhomologous recombination.

[91] Alternatively, the vector may contain additional nucleic acid sequences for directing integration by homologous recombination into the genome of the host cell. The additional nucleic acid sequences enable the vector to be integrated into the host cell genome at a precise location(s) in the chromosome(s). To increase the likelihood

of integration at a precise location, the integrational elements should preferably contain a sufficient number of nucleic acids, such as 100 to 10,000 base pairs, preferably 400 to 10,000 base pairs, and most preferably 800 to 10,000 base pairs, which are highly homologous with the corresponding target sequence to enhance the probability of homologous recombination. The integrational elements may be any sequence that is homologous with the target sequence in the genome of the host cell. Furthermore, the integrational elements may be non-encoding or encoding nucleic acid sequences. On the other hand, the vector may be integrated into the genome of the host cell by non-homologous recombination.

[92] For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question. Examples of bacterial origins of replication are P15A, pSC101, pMB1 and ColE1. Origins of replication of plasmids pBR322 (which has a pMB1 origin of replication), pUC19 (which has a ColE1 origin of replication), pACYC177 and pACYC184 (which have a P15A origin of replication), permit replication in *E. coli*; origins of replication for plasmids pUB110, pE194, pTA1060, or pAM.beta.1 permit replication in *Bacillus*. Examples of origins of replication for use in a yeast host cell are the 2 micron origin of replication, ARS1, ARS4, the combination of ARS1 and CEN3, and the combination of ARS4 and CEN6. The origin of replication may be one having a mutation which makes its functioning temperature-sensitive in the host cell (see, e.g., Ehrlich, 1978, Proceedings of the National Academy of Sciences USA 75: 1433).

[93] More than one copy of a nucleic acid sequence of the present invention may be inserted into the host cell to increase production of the gene product. An increase in the copy number of the nucleic acid sequence can be obtained by integrating at least one additional copy of the sequence into the host cell genome or by including an amplifiable selectable marker gene with the nucleic acid sequence where cells containing amplified copies of the selectable marker gene, and thereby additional copies of the nucleic acid sequence, can be selected for by cultivating the cells in the presence of the appropriate selectable agent.

[94] The procedures used to ligate the elements described above to construct the recombinant nucleic acid construct and expression vectors of the present invention are

-30-

well known to one skilled in the art (see, e.g., J. Sambrook, E. F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning, A Laboratory Manual*, 2d edition, Cold Spring Harbor, N.Y.).

[95] Many of the expression vectors for use in the present invention are commercially available. Suitable commercial expression vectors include p3xFLAGTM™ expression vectors from Sigma-Aldrich Chemicals, St. Louis MO., which includes a CMV promoter and hGH polyadenylation site for expression in mammalian host cells and a pBR322 origin of replication and ampicillin resistance markers for amplification in *E. coli*. Other suitable expression vectors are pBluescriptII SK(-) and pBK-CMV, which are commercially available from Stratagene, LaJolla CA, and plasmids that are derived from pBR322 (Gibco BRL), pUC (Gibco BRL), pREP4, pCEP4 (Invitrogen) or pPoly (Lathe et al., 1987, *Gene* 57, 193-201).

[96] Example 6 herein discloses the use of the expression vector pCK110900-I Bla, as shown in the vector map of FIG. 3.

Host Cells

[97] Host cells for use in expressing the expression vectors of the present invention include but are not limited to, bacterial cells, such as *E. coli*, *Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells (e.g., *Saccharomyces cerevisiae* or *Pichia pastoris* (ATCC Accession No. 201178)); insect cells such as *Drosophila* S2 and *Spodoptera Sf9* cells; animal cells such as CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are well known in the art.

[98] By way of example, *Escherichia coli* W3110 was transformed by an expression vector for expressing the shuffled genes of the present invention. The expression vector was created by operatively linking a variant gene of the present invention to the *lac* promoter under control of the *lacI* repressor gene. The expression vector also contained the P15A origin of replication and the chloramphenicol resistance gene. The transformed *Escherichia coli* W3110 was cultured under appropriate culture medium containing chloramphenicol such that only transformed *E*

coli cells that expressed the expression vector survived. See e.g., Example 1. Purification

[99] Once the AAM polypeptides were expressed by the variant genes in *E. coli*, the polypeptides were purified from the cells and or the culture medium using any one or more of the well known techniques for protein purification, including lysozyme treatment, sonication, filtration, salting, ultra-centrifugation, affinity chromatography, and the like under strict anoxic conditions. Suitable solutions for high efficiency extraction of proteins from bacteria, such as *E. coli*, are commercially available under the trade name CellLytic B™ from Sigma-Aldrich of St. Louis MO. A suitable process for purifying AAM polypeptides sufficiently from cell lysate for applications in a chemical process is disclosed in the references: Chirpich, T. P. et al., J. Biol. Chem., 1970, 245, 1778-1789; and Petrovich, R. M. et al., J. Biol. Chem., 1991, 266, 7656-7660, both of which are incorporated herein by reference.

Screening

[100] After several rounds of directed evolution were performed, the resulting libraries of exemplary AAM polypeptides were screened. Screening for transformed cells that express a polypeptide having AAM activity is, in general, a two-step process. First, one physically separates the cells and then determines which cells do and do not possess a desired property. Selection is a form of screening in which identification and physical separation are achieved simultaneously by expression of a selection marker, which, in some genetic circumstances, allows cells expressing the marker to survive while other cells die (or vice versa). Exemplary screening markers include luciferase, β -galactosidase, and green fluorescent protein. Selection markers include drug and toxin resistance genes, such as resistance to chloramphenicol, ampicillin and the like. Although spontaneous selection can and does occur in the course of natural evolution, in the present methods selection is performed by man.

[101] The AAM polynucleotides generated by the mutagenesis or directed evolution method are screened in accordance with the protocol described in Example 8 to identify those having enhanced activity that are suitable for inclusion as an improved AAM polypeptide of the present invention. In the process of Example 8, the

screening of clones from the expression libraries for enhanced AAM activity was performed by measuring the conversion of α -alanine to β -alanine using liquid chromatography and mass spectrometry. Based upon the screening results, the AAM polypeptides of the present invention are listed in Table 2 below along with their residue changes and enhanced AAM activity relative to one parental AAM polypeptide, *i.e.*, the polypeptide of SEQ ID NO: 59.

Table 2

Seq. ID No.	Residue changes relative to parent SEQ ID NO: 59	Rate of β -alanine(μ M) produced /hr 1 Cell OD
34	I177L, I227M, G308R, I408L, F416S, D447G	31.9
10	I298V, G308R, F416S, D447G	6.3
38	D125N, I177L, T210S,	11.0
20	K2E, I307L,	14.7
14	K13E, L17R, L197P, I200T, M281V, F310S, F416S, D447G	7.7
22	Y72H, L118P, R145L, I220V, F240L, S250P, R311C, F416S, D447G	1.0
42	K19R, T99S, G308R, F416S, D447G	3.5
26	N80K, G308R, E319G, R325G, Q350R	4.8
18	Q32R, S74P, S113T, L118P, G308R, F416S, D447G	3.9
44	D79E, G308R, S329P, F393S, F414S, D445G, L453S,	12.9
51 (fragment)	A73V, G308R, Y331N, F416S, D447G	7.0
36	D79E, S93P, N132D, M281I, G308R, Y331N, F416S, D447G	6.0
48	K2E, M76I, D79E, T131A, L203P, G308R, Y331C, F416S, D447G	22.0
12	R38G, C134G, C141R, L203P, I280T, G308R, F416S, D447G	3.6
4	2KE, I220V, N237D, G308R, D360G, K361R, F416S, D447G	4.5

-33-

16	K13E, L17R, L197P, I200T, M281V, G308R, F310S, F416S, D447G	19.4
24	E23D, L43S, D124G, Y137H, K156E, G308R, D411G, F416S, D447G	18.9
46	W18R, M76I, D79E, V90A, M152T, I163T, S178P, V215G, G308R, V354A, F416S, D447G	20.7
28	E22G, Y71C, S74P, H108R, D187G, I244V, G308R, E396G, F416S, D447G, F454S	29.2
40	Y137H, G308R, D411G, F416S, D422V, D447G	2.9
32	H35R, D79E, K98T, T99S, N132S, S135P, E204G, K230R, G308R, F416S, D447G	13.6
2	W235R, S250P, C254R, D276G, G308R, Y380C, I381T, F416S, K440E, D447G	17.5
30	Q32R, N67S, H140R, G308R, F416S, D447G	14.3
6	E24G, M96I, E109G, G308R, F416S, D447G	23.0
8	G308R, S329P, F416S, D447G, L455S	14.7

[102] In Table 2 above, it is seen that the AAM polypeptides of the present invention have from 2 to 11 residue differences than their parent polypeptide of SEQ ID NO: 59, and very significant AAM activity as evidenced by the production of β -alanine in the assay of Example 8. In comparison, β -alanine was not detected for SEQ ID NO: 59 under the assay conditions used to test the AAM variants. However, some β -alanine production for parental SEQ ID NO: 59 was detected in a qualitative growth based complementation assay.

[103] Referring to Table 2 above, two preferred residue changes for the AAM polypeptides of the present invention relative to the parental sequence of SEQ ID NO: 59 are G308R and F416S. In those AAM polypeptides of the present invention that are at least 447 residues long, an additional preferred residue change is D447G relative to the parental sequence of SEQ ID NO: 59. Additional suitable residue

changes are G308K, F416M and D447L, A, I or V. Thus, in one aspect, the present invention is directed to an AAM polypeptide having at least 5 amino acid residue changes, typically 5-11 residue changes, relative to SEQ ID NO: 59 or a truncated fragment thereof as taught herein, the residue changes including from 1 to 3 residue changes selected from the group consisting of G308R, G308K, F416S, F416M, D447G, D447L, D447A, D447I and D447V.

[104] Based upon the AAM activity in Table 2, an especially preferred AAM polypeptide of the present invention is a polypeptide having 95% sequence homology with the polypeptide of SEQ ID NO: 34, more preferably 98% homology, most preferably 99% homology.

[105] The parental polypeptides of SEQ ID NOs: 53, 55 and 57 demonstrate that the residues 1-8 at the N-terminus and residues 434-473 at the C-terminus are not necessary for KAM or AAM activity. Likewise, the polypeptide fragment of SEQ ID NO: 51, which is a 399 residue expression product, discloses that the first 72 amino acids at the N-terminus relative to the parental clone of SEQ ID NO: 59 are not necessary for AAM activity. (See Table 2) Thus, it is also within the scope of the present invention that the polypeptides described herein include fragments thereof that lack from 1 to 72 residues from their N-terminus relative to the parental sequence of SEQ ID NO: 59, typically from 1 to 40 residues, more typically from 1-20 residues, most typically from 1 to 11 residues. It is also within the scope of the present invention that the above described N-terminal truncation be utilized in combination with a C-terminal truncation as described elsewhere herein.

[106] Only a very few ($\leq 0.5\%$) of the mutations to the parental *B. subtilis* KAM (SEQ ID NO: 59) backbone were found to be beneficial. Specifically, for every 1000 clones screened, there occurred only 3-5 single point or double point mutations that were beneficial. In fact, some of the mutations were found to be detrimental.

[107] The first of the following two sets of sequences provides the sequence of the wild type *B. subtilis* lysine 2,3-aminomutase (KAM) polypeptides of the prior art, as deposited (GI_2529467_GB_AAB81159.1). This sequence (SEQ ID NO: 60) was not used as a parent sequence but is provided only for purposes of comparison.

-35-

MKNK WYKPKRHWKEIELWKDVPEEKWNDWLWQLTHT
 VRTLDD LKKVINLTEDEEEGVRISTKTIPLNITPYASYL
 MDPDNPRCPVRMQSVPLSEEMHKTKYDLEDPLHEDED
 SRVPGLTHRYPDRLVFLVTNQCSMYCRYCTRRRFSGQI
 GMGVPPKKQLDAAIAYIRETPEIRDCLISGGDGLLINDQI
 LEYILKELRSIPHLEVIRIGTRAPVVFPQRITDHLCEILK
 KYHPVWLNTHFNNTSIEMTEESVEACEKLVNAGVPVGN
 QAVVLAGINDSVPIMKKLMHDLVKIRVRPYYIYQCDLS
 EGIGHFRAPVSKGLEIIEGLRGHTSGYAVPTFVVDAPGG
 GKGIALQPNYVLSQSPDKVILRNFEFVITSYPEPENYIP
 NQADAYFESVFPETADKKEPIGLSAIFADKEVSFTPEN
 DRIKRREAYIANPEHETLKDRRERRRDQLKEKKFLAQK
 KQKETECGGDSS

[108] The second sequence in the set indicates the diversity of the AAM polypeptides of the present invention relative to the known wild-type *B. subtilis* KAM sequence by designating with the letter "X" followed by the residue number those residues in the Applicants' AAM polypeptides that differ from those of wild-type *B. subtilis* KAM sequence:

MX₂NK WYKPKRHWX₁₃EIEX₁₇WX₁₉DVPX₂₃X₂₄KWNDWLW
 X₃₂LTX₃₅TVX₃₈TLDDX₄₃KKVINLTEDEEEGVRISTKTIPL
 X₆₇ITPX₇₁X₇₂X₇₃X₇₄LMDPX₇₉X₈₀PRCPVRMQSVPLX₉₃EE X₉₆H
 X₉₈X₉₉KYDLEDPLX₁₀₈X₁₀₉DEDSX₁₁₄VPGX₁₁₈THRYPX₁₂₄RVLF
 LVTX₁₃₂QX₁₃₄X₁₃₅X₁₃₆X₁₃₇CRX₁₄₀X₁₄₁TRRX₁₄₅FSGQIGMGVP
 X₁₅₆KQLDAAIAYIRETPEIRDCLISGGDGLLINX₁₈₇QILEYI
 LKEX₁₉₇RSX₂₀₀PHX₂₀₃X₂₀₄VIRIGTRAPVVFPQRITDHX₂₂₄CEI
 LKX₂₃₀X₂₃₁HPVX₂₃₅LX₂₃₇THX₂₄₀NTSIEMTEEX₂₅₀VEAX₂₅₄EKL
 VNAGVPVGNQAVVLAGINX₂₇₆SVPX₂₈₀X₂₈₁KKLMHDLVKI
 RVRPYYIYQCDLSEGX₃₀₇X₃₀₈HX₃₁₀X₃₁₁APVSKGLX₃₁₉IIEGL
 RGHTX₃₂₉GX₃₃₁AVPTFVVX₃₃₉APGGGGKIALX₃₅₀PNYVLSQ
 SPX₃₆₀KVILRNFEFVITSYPEPENX₃₈₀X₃₈₁PNQADAYFESV
 X₃₉₃PX₃₉₅TADKKEPIGLSAX₄₀₈FAX₄₁₁KEVSX₄₁₆TPENVX₄₂₂RI
 KRREAYIANPEHETLX₄₄₀DRREX₄₄₅RX₄₄₇QLKEKKX₄₅₄X₄₅₅A
 QQKKQKETECGGDSS

The diversity of changes at various residue positions for the AAM polypeptides of the present invention are shown to the right of the arrow in Table 2 below and relative amino acid residues of wild-type KAM of *B. subtilis* (GI_2529467_GB_AAB81159.1) (SEQ ID NO: 60) which are shown to the left of the arrow:

Table 3

X ₂	K → E
X ₁₃ :	K → E
X ₁₇ :	L → R
X ₁₉ :	K → R
X ₂₃ :	E → D, G
X ₂₄ :	E → G
X ₃₂ :	Q → R,
X ₃₅ :	H → R
X ₃₈ :	R → G
X ₄₃ :	L → S
X ₆₇ :	N → S
X ₇₁ :	Y → C
X ₇₂ :	Y → H, W
X ₇₃ :	A → V
X ₇₄ :	S → P
X ₇₉ :	D → E
X ₈₀ :	N → K
X ₉₃ :	S → P
X ₉₆ :	M → I
X ₉₈ :	K → T
X ₉₉ :	T → S
X ₁₀₈ :	H → R
X ₁₀₉ :	E → G
X ₁₁₄ :	R → P
X ₁₁₈ :	L → P
X ₁₂₄ :	D → N
X ₁₃₂ :	N → D, S
X ₁₃₄ :	C → G
X ₁₃₅ :	S → P
X ₁₃₆ :	M → V
X ₁₃₇ :	Y → H
X ₁₄₀ :	Y → H
X ₁₄₁ :	C → R
X ₁₄₅ :	R → L
X ₁₅₆ :	K → E
X ₁₈₇ :	D → G
X ₁₉₇ :	L → P
X ₂₀₀ :	I → T
X ₂₀₃ :	L → P
X ₂₀₄ :	E → G
X ₂₂₄ :	L → P
X ₂₃₀ :	K → R
X ₂₃₁ :	Y → H
X ₂₃₅ :	W → R
X ₂₃₇ :	N → D

-37-

X ₂₄₀ :	F→L
X ₂₅₀ :	S→P
X ₂₅₄ :	C→Y, R
X ₂₇₆ :	D→G
X ₂₈₀ :	I→T
X ₂₈₁ :	M→I, V
X ₃₀₇ :	I→L
X ₃₀₈ :	G→R
X ₃₁₀ :	F→S
X ₃₁₁ :	R→C
X ₃₁₉ :	E→G
X ₃₂₉ :	S→P
X ₃₃₁ :	Y→N
X ₃₃₉ :	D→H
X ₃₅₀ :	Q→R
X ₃₆₀ :	D→G
X ₃₆₁ :	K→R
X ₃₈₀ :	Y→C
X ₃₈₁ :	I→T
X ₃₉₃ :	F→S
X ₃₉₅ :	E→G
X ₄₀₈ :	I→L
X ₄₁₁ :	D→G
X ₄₁₆ :	F→S
X ₄₂₂ :	D→V
X ₄₄₀ :	K→E
X ₄₄₅ :	R→K
X ₄₄₇ :	D→G
X ₄₅₄ :	F→S
X ₄₅₅ :	L→S

[109] In a fourth aspect, the present invention is directed to a method of making an AAM a nucleic polypeptide of the present invention comprising (a) cultivating a host cell transformed with a nucleic acid sequence encoding an AAM polypeptide of the present invention under conditions suitable for production of the polypeptide; and (b) providing glucose to the cultivated host cells under conditions suitable for the production of β -alanine. The β -alanine may be optionally recovered from the cells.

Example 1: Transformation protocol for *aam* libraries/ $\Delta panD$ strain

[110] A mutant *E. coli* strain - $\Delta panD$, derived from BW25113 which is described in Datsenko, K.A. and Wanner, B.L., Proc. Natl. Acad. Sci. USA 97:6640-6645 (2000)

-38-

was used as the host strain for screening of the *aam* gene libraries. The protocol used to make the deletion is detailed in Example 4 of Cargill patent application WO 03/062173.

[111] Chemical competent *E. coli* Δ *panD* was removed from -80°C frozen storage and thawed. Thereafter, it was kept on ice until used. An aliquot (100 μl per transformation) was transferred into a sterile 1.5ml centrifuge tube. A KCM (5X) salt solution was added until the concentration in the aliquot was 1X. KCM consists of 700 mM KCl; 10 mM morpholinopropanesulphonic acid (MOPS) adjusted to pH 5.8. 1-5 μl of the ligation mixture was added to the cells. The cells containing the ligation mixture were first incubated on ice for 30 minutes. The cells were heat shocked at 42°C for 1 min, and subsequently incubated on ice for 2 minutes. 500 μl of SOC (Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, 1st Ed., pp. A.2 and A.3, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY) was added to the cells, and the cells were incubated at 37°C for 1 hour with agitation. The cells were then centrifuged at 5000 rpm for 3 minutes, and the SOC was removed. The cell pellet was re-suspended in 500 μl of M9 selection medium ((Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, 1st Ed., pp. A.2 and A.3, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY) and incubated at 30°C for 2-4 hours with agitation. The cells were then plated onto M9 minimal agar medium supplemented with 1% mannose, 20 μM iron citrate, 5.0 g/l α -alanine, 0.1mM isopropyl- β -D-thiogalactoside (IPTG) (Sigma Chemical Corp., St. Louis, MO), 50mM MOPS, 25mM bicarbonate, and 30 $\mu\text{g/ml}$ chloramphenicol. The plated cells were incubated at 30°C for 3 days or until colonies were of sufficient size to be picked using the Q-BOTTM robot colony picker (Genetix USA, Inc, Boston MA).

[112] In Round 2 of the transformation, the above procedure was followed except that the incubation temperature of the last two incubations in the procedure was increased to 37°C , and M9 minimal selection medium was not supplemented with α -alanine (0 g/L α -alanine).

A. Alternate Transformation protocol for *aam* libraries/ Δ *panD* KI*fldA* strain

[113] A mutant *E. coli* strain Δ *panD*, derived from BW25113 which is described in Datsenko, K.A. and Wanner, B.L., *Proc. Natl. Acad. Sci. USA* 97:6640-6645 (2000) is used as the host strain for screening of the *aam* gene libraries. The protocol used to make the deletion is detailed in Example 4 of International patent publication WO 03/062173. Optimally, a strain additionally having an increased expression of the flavodoxin (*fldA*) gene was used as the host strain for screening of the *aam* gene libraries, since increased flavodoxin enhances aminomutase activity when produced in *E. coli*. See USSN _____, by Cargill, Inc. (Liao, et al), filed October 14, 2005, entitled "Increasing the Activity of Radical S-Adenosyl Methionine (SAM) Enzymes" describes the production of β -alanine from cells that express AAM and overexpress flavodoxin at Examples 1-4, and these examples are incorporated herein by reference. This same application, USSN _____, by Cargill, Inc. (Liao, et al.) filed October 14, 2005, describes in Example 4 (incorporated herein) the construction of a strain of *E. coli* in which an artificial $P_{lac/ara}$ hybrid promoter was placed immediately upstream of the *fldA* gene. Strains carrying the artificial promoter before the *fldA* gene are designated KI*fldA*, where KI refers to "knock-in").

[114] Competent cells of *E. coli* Δ *panD* KI*fldA* are prepared either chemically or electrochemically using standard protocols. Competent *E. coli* Δ *panD* KI*fldA* was removed from -80°C frozen storage and thawed. Thereafter, it was kept on ice until used. An aliquot (100 μ l per transformation) was transferred into a sterile 1.5ml centrifuge tube. A KCM (5X) salt solution was added until the concentration in the aliquot was 1X. KCM consists of 700 mM KCl; 10 mM morpholinopropanesulphonic acid (MOPS) adjusted to pH 5.8. 1-5 μ l of the ligation mixture was added to the cells. The cells containing the ligation mixture were first incubated on ice for 30 minutes. The cells were heat shocked at 42°C for 1 min, and subsequently incubated on ice for 2 minutes. 500 μ l of SOC (Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, 1st Ed., pp. A.2 and A.3, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY) was added to the cells, and the cells were incubated at 37°C for 1 hour with agitation. The cells

-40-

were then centrifuged at 5000 rpm for 3 minutes, and the SOC was removed. Pellets were subsequently resuspended in a medium appropriate for either the complementation assay (Example 3) or the biotransformation assay (Example 4).

Example 2: Cloning of *aam* genes into pCK110900 series vectors

[115] The strategy employed for cloning the alanine aminomutase genes into an inducible expression system involved the isolation of the *aam* gene by PCR and cloning of the PCR fragment into the *SfiI* restriction sites downstream from a mutant *lac* promoter/operator system. Initially, PCR primers were designed to contain a nucleotide sequence that is specific to the 5' and 3' ends of the *aam* gene, as well as the Shine-Delgarno sequence of the ribosome-binding site, and the unique *SfiI* restriction sites. The gene was then amplified from a template, purified and digested with the restriction endonuclease *SfiI*. The restricted PCR fragment was purified using the QIAquick PCR purification kit (Qiagen), and cloned into the *SfiI* sites of the expression vector pCK110900-I Bla of FIG. 3 under the control of a *lac* promoter and *lacI* repressor gene. The expression vector also contained the P15a origin of replication and the chloramphenicol resistance gene. Shuffled *aam* gene libraries were cloned by the same method. Several clones were found that expressed an active alanine 2,3-aminomutase (as per the method of Example 8) and the synthetic genes were sequenced. A polynucleotide sequence designated BSAAM (SEQ ID NO: 58) - was used as a starting material for all further mutations and shuffling. BSAAM (SEQ ID NO: 58) has approximately 99.2% nucleotide identity with the wild-type *Bacillus subtilis* lysine aminomutase (GenBank Accession No. H10329).

Example 3: Screening via the Tier 2a growth assay

Tier 2a growth Assay

[116] The growth assay identifies variants capable of generating the essential metabolite AcetylCoA via β -alanine produced by AAM variants in the *E. coli* Δ *panD* host strain. Growth is therefore a function of CoA production, and indirectly of AAM activity.

-41-

A. Procedure

[117] AAM active clones from the tier 1 complementation assay were picked with a QBOT™ robot colony picker (Genetix USA, Inc., Boston MA) and inoculated into a 96-well master plate. The inoculums were grown in the 96 well master plate on a buffered minimal selection media ($\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$ 12.8g/L; KH_2PO_4 3g/L; NaCl 0.5g/L; NH_4Cl 1g/L; MgSO_4 2mM; CaCl_2 0.04mM; mannose 2%; IPTG 1mM; ferric citrate 20 μM ; chloramphenicol 30 $\mu\text{g/ml}$; MOPS pH 7, 50mM; and sodium bicarbonate pH 9, 25mM) (hereinafter “MSM”) to which was added 0.1uM β -alanine and 0.5g/L α -alanine. Plates were covered using AirPore™ microporous tape (Qiagen, Inc.) and incubated at 25°C, 250 rpm, 85% humidity until cultures reached saturation, at which time glycerol was added to the cultures to a final concentration of 20-30%, and the plates stored at -80°C.

[118] Samples from a frozen master plate were arrayed into an “inoculum” plate containing buffered minimal selection media (MSM), as described above, further containing 0.5g/L α -alanine. The inoculum plates were covered with AirPore™ microporous tape (Qiagen, Inc.) and incubated at 25°C, 250 rpm, 85% humidity until cultures reached saturation.

[119] 15 μl from the inoculum plate was inoculated into a 96-well “assay” plate containing 185 μl of fresh MSM with 0.5g/L α -alanine. The assay plates were covered with AirPore™ microporous tape (Qiagen, Inc.) and a lid, and incubated at 25°C, 85% humidity, 250rpm. Measurements of OD at 600nm were made at discrete times for a period of approximately (~) 40hours.

B. Data Analysis

[120] Since library variants exhibit unique growth profiles, it was preferable to calculate and compare growth rates (slopes) at three (3) different growth phases (early, mid and late) to identify all potentially improved variants. Clones that exhibit three (3) standard deviations above the plate average in any of the three (3) phases were designated as potentially improved variants and were retested in tier 2b for comparative ranking.

Example 4: Screening via the Tier 2b growth assay

[121] The stringency of the growth screen is increased in Tier 2b by excluding α -alanine (the substrate for AAM) from the medium. Under these conditions, the cell relies on internal/cellular pools of α -alanine to serve as a substrate for AAM, and subsequently, for cell growth. AAM variants capable of utilizing low, intracellular pools of α -alanine might represent low K_M variants.

A. Procedure

[122] Samples from a frozen master plate were arrayed into an "inoculum" plate containing buffered minimal selection media (MSM), as described above, further containing 0.5g/L α -alanine. The inoculum plates were covered using AirPore™ microporous tape and incubated at 25°C, 250 rpm, 85% humidity until cultures reached growth saturation.

[123] A TECAN™ Robotic Sample Processor (Columbus, Ohio) was used to remove 10 μ l of inoculum from each Tier 2a variant from the inoculum plates and seed it in replicates of 8 into each of the following:

96-well Assay plate containing 190 μ l of fresh MSM, 0.5g/L α -alanine.

96-well Assay plate containing 190 μ l of fresh MSM, containing no α -alanine.

The Assay plates were covered with AirPore™ microporous tape and a lid and grown at 25°C, 85% humidity, 250rpm. Samples were collected at time points for approximately 3-4 days and the OD_{600nm} was measured for each sample.

B. Tier 2b Data Analysis

[124] Variants were ranked by the following 3 criteria:

- i) Growth ratio equal to a final culture OD₆₀₀ on medium without α -alanine/final culture OD_{600nm} on medium containing α -alanine;
- ii) Final culture OD₆₀₀; and
- iii) Initial growth rates (in phase 1, from approximately 0-20 hour)

Clones with final culture OD_{600nm} > 0.7 were retained.

-43-

Clones were then ranked based on the growth ratio of criteria (i). Any clones with a $OD_{600nm} > 0.7$ were retained. However, clones that did not meet the above two criteria, but had a very good initial growth rate (iii) were also selected for further evaluation.

Example 5: Screening via Tier 2c- PCR analysis

The PCR screen identifies variants that contain the correct size gene in the expression vector prior to further screening for function. It excludes unstable gene variants that may have undergone deletions/truncations during the screening process.

A. Procedure

Potentially improved variants from frozen master plates were inoculated into a 96-microwell plate containing LB media with 1% glucose and 30 µg/mL chloramphenicol. Cultures were grown at 25°C, 250 rpm, 85% humidity in plates covered with AirPore™ microporous tape (Qiagen, Inc.) until cultures reached saturation, approximately 2 days. 10 µL of the culture was transferred to a PCR plate and boiled at 99°C for 10 minutes to disrupt the cells. Thereafter, 90 µL of the following PCR Master Mix was added to the disrupted cells:

PCR Master Mix:

10 µL	10X Taq Polymerase Buffer (QIAGEN, Valencia CA)
4 µL	25 mM MgCl ₂
2 µL	10 mM dNTPs
1.25 µL	20 µM primer – B _{forward} (specific for BsAAM gene)
1.25 µL	20 µM primer – B _{reverse} (specific for BsAAM gene)
1 µL	5U/µL Taq polymerase (QIAGEN)
70.5 µL	Sterile water
90 µL	Total volume

The *Bacillus* specific primers used in the PCR reaction are as follows:

-44-

B-forward:

5'ccagcctggccataaggagatatacatatgaaaaacaaatggtataaac 3' SEQ ID NO: 63

B-reverse:

5' atggtgatggtgatggtggccagtttggccttatgaagaatcccctccgc 3' SEQ ID NO: 64

The amplification reaction was run for 30 cycles. The first cycle was run at 94°C for 1 minute. Thereafter, the extension procedure was performed for 29 cycles: 94.0°C for 1 minute; 55.0°C for 30 seconds; and 72.0°C for 1 minute. The final extension was performed at 72.0°C for 5 minutes. The products of the PCR reactions were analyzed by gel-electrophoresis on a 0.8% agarose gel.

Example 6: Growth of AAM variants for β -alanine production (50 ml scale).**Cell selection method for identifying AAM activity.**

[125] To identify genes encoding polypeptides that can perform the alanine 2,3-aminomutase reaction, an efficient screen or selection for the desired activity is needed. Therefore, a selection method was developed by recognizing that *E. coli* uses beta-alanine for the synthesis of pantothenic acid, which in turn is a component of coenzyme A (CoA) and of acyl carrier protein (ACP). CoA and ACP are the predominant acyl group carriers in living organisms, and are essential for growth.

[126] In *E. coli*, the primary route to beta-alanine is from aspartate in a reaction catalyzed by aspartate decarboxylase (E.C. 4. 1. 1.1 1), which is encoded by the *panD* gene. A functional deletion mutation of *panD* (shown as $\Delta panD$) results in beta-alanine auxotrophy and growth inhibition, which can be alleviated by the exogenous addition of pantothenate or beta-alanine, or by the production of beta-alanine from another source.

[127] Strain description: *E. coli* $\Delta panD$ host (derived from BW25113, described in Datsenko, K.A. and Wanner, B.L., Proc. Natl. Acad. Sci. USA 97:6640-6645 (2000)), transformed with pCK110900-I Bla vector (low promoter strength resulting from mutated lac promoter sequence). The inoculum culture was grown in buffered minimal selection medium (MSM): M9 salts, pH 7.0-7.4, 50mM MOPs, pH 7.0, 25

-45-

mM sodium bicarbonate, pH 9.0, 1mM isopropyl- β -D-thiogalactoside (IPTG), 30 μ g/ml chloramphenicol, 0.1g/L alanine, 5uM pyridoxine HCl, and 20uM ferric citrate. A 1:20 dilution of inoculum was used to inoculate 50ml of MSM medium described above. Cultures were incubated at 25°C, 250 rpm for approximately 3 days or until the culture reaches OD_{600nm} ~1. Then, α -alanine was added to the medium to a final concentration of 300 mM, and pantothenate was added to about 300uM. Incubation of the supplemented medium continued at 25°C, 250 rpm. Samples were removed from the medium for analysis at time points from t= 0 through t=5 hours following the addition of α -alanine.

Example 7: Method for extracting cells for β -alanine detection

[128] Cells from the cultures of Example 6 were harvested by centrifugation of the cultures. The supernatant (spent media) was decanted and saved for further analysis (below). The cell pellets were washed with water. Pellets may be stored at -80°C for future extraction. The 50ml cell pellets (OD ~ 4.0) were re-suspended completely in a test tube in 0.9 ml water. The extraction volume for each sample was adjusted to this proportion according to the harvest OD₆₀₀. An equal volume of methanol (-20°C) and 200 μ L of micro-glass beads was added and the mixture vortexed vigorously. Tubes containing the mixtures were placed on dry ice/EtOH, or in a -80°C freezer, for about 30 min. The frozen contents in the tube were thawed at room temperature and vortexed vigorously again, and centrifuged at maximum speed for about 10 minutes. The supernatants were filtered using 0.2–0.45 micron filter plates, or syringe filters.

[129] The spent medium was filtered using a 0.2-0.45 micron filter plate or syringe filter. The filtered spent medium was diluted 1:10 in -20°C methanol/water (final methanol concentration 50%).

[130] The β -alanine content of cell extract and spent media was analyzed by LC/MS/MS (Example 8).

For spent medium sample, the first minute was diverted to waste. The β -alanine peak arrived at approximately 2.0 minutes.

-46-

The assay can be scaled to 2ml, if only the spent media is analyzed.

Example 8: Assay for β -alanine (LC/MS/MS)

[131] β -alanine was determined using a combination of liquid chromatography and mass spectrometry. Suitable analytes were the cell extracts and spent media as prepared in Example 7.

[132] The liquid chromatography (LC) phase was performed using an ASTEC CHIROBIOTICTM T 4.6 cm x 50 mm chiral LC column (Advanced Separation Technologies, Inc., Whippany, N.J. USA). The mobile phase consisted of two solutions: A: 0.25% aqueous acetic acid; and B: 0.25% (v/v) acetic acid in methanol. The elution was isocratic @ 0.6ml/minute.

[133] The mass spectrometer (MS) analysis was performed on a Micromass Ultima Triple Quad mass spectrometer, using the following tune parameters:

Capillary: 3.5 kV; cone: 20 V; hex 1: 15 V; aperture: 1.0V; source temp: 100°C; desolvation temp: 350°C; cone gas: 40 L/hr; desolvation gas: 500 L/h; low mass resolution(Q1): 12; high mass resolution (Q1): 12; ion energy (Q1): 0.1; collision cell entrance: -5; collision energy: 14; exit: 1; low mass resolution (Q2): 15 high mass resolution (Q2): 15; ion energy (Q2): 3.0; multiplier: 650 V.

MS Method**Alanine transitions**

Analyte	Parent Ion (m/z)	Daughter Ion (m/z)	Dwell Time (sec)
α -alanine	90	44.7	0.1
β -alanine	90	30.7	0.1
α -lysine	147	84.5	0.1
β -lysine	147	70.5	0.1

The inter-channel delay was 0.1 seconds.

-47-

CLAIMS

WHAT IS CLAIMED IS:

1. A polypeptide having alanine 2,3-aminomutase activity (hereinafter an "AAM polypeptide") and
 - (a) having an amino acid sequence selected from the group consisting of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48 and 51;
 - (b) having an amino acid sequence which has at least 98% homology with the amino acid sequence selected from the group consisting of SEQ ID NO: 2, 22, 28, 32, and 36;
 - (c) having an amino acid sequence which has at least 99% homology with the amino acid sequence selected from the group consisting of SEQ ID NO: 4, 6, 8, 12, 16, 24, 26, 30, 34 and 40;
 - (d) being a polypeptide encoded by a nucleic acid sequence which hybridizes under high stringency conditions with either (i) the nucleotide sequence of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 41, 43, 45, 47 or 49; (ii) a subsequence of (i) of at least 100 nucleotides, or (iii) a complementary strand of (i) or (ii); or
 - (e) being a variant of the polypeptide of (d) comprising a substitution, deletion, and/or insertion of one to six amino acids therefrom and having AAM activity from about 1 to about 30 μ M β -alanine produced /hour 1 cell OD at pH 7.0-7.6, 25°C.
2. The polypeptide of claim 1 having an amino acid sequence selected from the group consisting of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48 and 51.
3. The polypeptide of claim 1 having an amino acid sequence which has at least 98% homology with the amino acid sequence selected from the group consisting of SEQ ID NO: 2, 22, 28, 32, and 36.

-48-

4. The polypeptide of claim 1 having an amino acid sequence which has at least 99% homology with the amino acid sequence selected from the group consisting of SEQ ID NO: 4, 6, 8, 12, 16, 24, 26, 30, 34 and 40.

5. The polypeptide of claim 1 being a polypeptide encoded by a nucleic acid sequence which hybridizes under high stringency conditions with either (i) the nucleotide sequence of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 41, 43, 45, 47 or 49; (ii) a subsequence of (i) of at least 100 nucleotides, or (iii) a complementary strand of (i) or (ii)

6. The polypeptide of claim 1 being a variant of the polypeptide of (d) comprising a substitution, deletion, and/or insertion of one to six amino acids therefrom and having AAM activity from about 1 to about 30 μ M β -alanine produced /hour 1/cell OD at pH 7.0-7.6, 25°C.

7. An AAM polypeptide having an amino acid sequence of SEQ ID NO: 2, 6, 12, 16, 20, 24, 28, 30, 32, 34, 38, 44, 46 or 48.

8. The AAM polypeptide of claim 7 having an amino acid sequence of SEQ ID NO: 6, 12, 28, 34, 46 or 48.

9. The AAM polypeptide of claim 8 having an amino acid sequence of SEQ ID NO: 28 or 34.

10. A polynucleotide encoding an AAM polypeptide of claim 1.

11. A polynucleotide encoding a polypeptide having AAM activity, said polynucleotide having SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 41, 43, 45, 47 or 49.

12. An isolated and purified polynucleotide which encodes a polypeptide of claim 1.

13. An expression vector comprising a polynucleotide of claim 10 or 11 operatively linked to a promoter.

14. A host cell transformed to express a polynucleotide of claim 10.
15. A method of making an AAM polypeptide of claim 1, comprising (a) cultivating a host cell comprising a nucleic acid construct comprising a nucleic acid sequence encoding the AAM polypeptide under conditions suitable for production of the polypeptide; and (b) recovering the AAM polypeptide.
16. An AAM polypeptide of claim 1 in lyophilized form.
17. A composition comprising a polypeptide of claim 1 in a buffered medium.
18. An AAM polypeptide having from 5 to 11 amino acid residue changes relative to SEQ ID NO: 59 or a fragment thereof, the residue changes including from 1 to 3 residue changes selected from the group consisting of G308R, G308K, F416S, F416M, D447G, D447L, D447A, D447I and D447V.

1/8

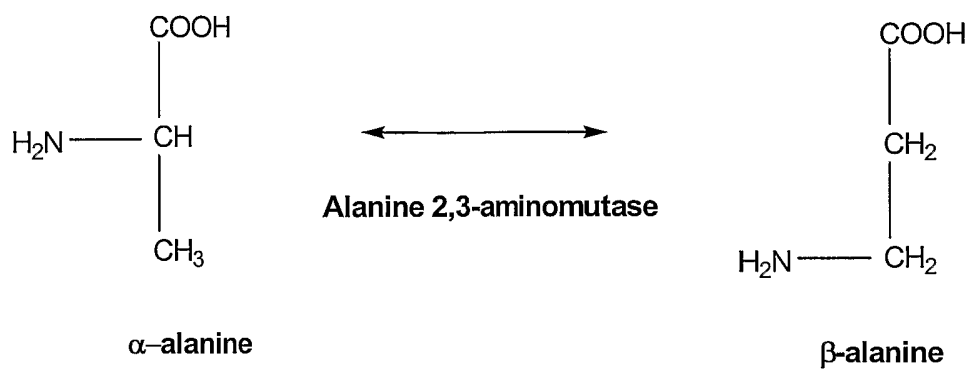
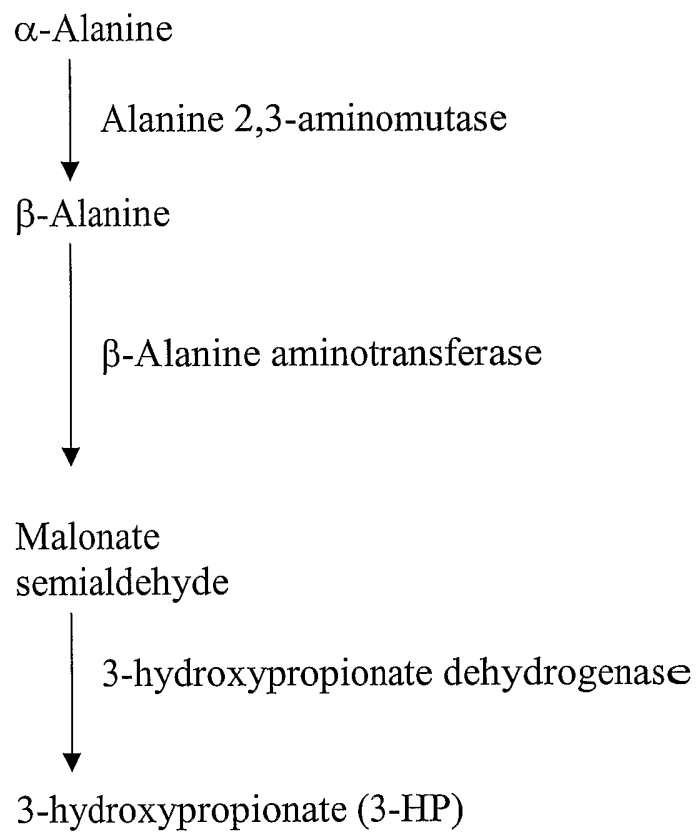


FIG. 1

2/8

**FIG. 2**

3/8

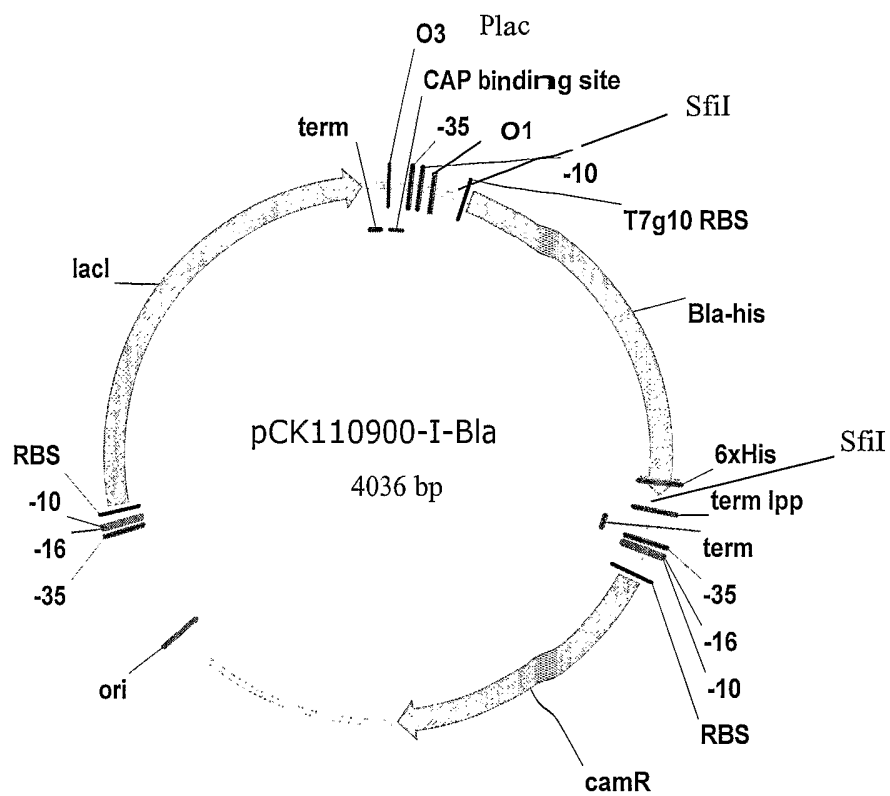


FIG. 3

4/8

SEQ ID NO:

1

50

P_GI2529467_G8_AAB81159.1_ 60 (1)
 MKNKWKYKPRHWKEIELWKDVPEEKWNDWLWQLTHTVRTLDDLKVINLT
 P_GI2634361_EMB_CAB13860.1_ 61 (1)
 MKNKWKYKPRHWKEIELWKDVPEEKWNDWLWQLTHTVRTLDDLKVINLT
 P_S00701550 59 (1)
 MKNKWKYKPRHWKEIELWKDVPEEKWNDWLWQLTHTVRTLDDLKVINLT
 P_S00701551 53 (1) -----
 MSLKDKFFETHVSQEDWNDWKWQVRNRIKTVEELKKYIPLT
 P_S00701552 55 (1) -----
 MAESRRKYFEPDVITDEQWYDWHWOVLNRIKTLTDLKKYVTLT
 P_S01032894 57 (1) -----
 MNTVNTRKKFFPNVITDEEWNDWTWQVKNRLKSVEDELEKYVDLS
 Consensus 62 (1)
 MKNKWKYKPRHWKEIELWKDVPEEKWNDWLWQLTHTVRTLDDLKVINLT

FIG. 4A

51

100

P_GI2529467_G8_AAB81159.1_ (51)
 EDEEEGVRISTKTIPLNITPYASLMDPDNPRCPVRMQSVPLSEEMHKTK
 P_GI2634361_EMB_CAB13860.1_ (51)
 EDEEEGVRISTKTIPLNITPYASLMDPDNPRCPVRMQSVPLSEEMHKTK
 P_S00701550 (51)
 EDEEEGVRISTKTIPLNITPYASLMDPDNPRCPVRMQSVPLSEEMHKTK
 P_S00701551 (41)
 PEEEEGVKRCLDTLRMAITPYYSLIDVENPNDPVRKQAVPLSLELHRAA
 P_S00701552 (43)
 AEEEEGVKESPKVLRMAITPYYSLIDPENPNCPVRKQAITQOELVRAP
 P_S01032894 (44)
 EEEETEGVVRTLETLRMAITPFYFSLIDLNSDRCPVRKQAITIRELHOSD
 Consensus (51)
 EDEEEGVRISTKTIPLNITPYASLMDPDNPRCPVRMQSVPLSEEMHKTK

FIG. 4B

P_GI2529467_G8_AAB81159.1

P_GI2634361_EMB_CAB13860.1

P_S00701550

P_S00701551

P_S00701552

P_S01032894

Consensus

101

(101)

YDLEDPLHEDEDSRPVGLTHRYPDRLVFLVTNQCSMYCRYCTRRRFSGQI

(101)

YDLEDPLHEDEDSPVPGLTTHRYPDRLVFLVTNQCSMYCRYCTRRRFSGQI

(101)

YDMEPLHEDEDSPVPGLTTHRYPDRLVFLVTNQCSVYCRHCTRRRFSGQI

(91)

SDMEDPLHEDEGDSVPVGLTHRYPDRLVLLMTDQCSVYCRHCTRRRFAGRT

(93)

EDQVDEPSEDEDSPVPGLTTHRYPDRLVFLITDKCSMYCRHCTRRRFAGQK

(94)

ADMLDPLHEDEDSPVPGLTTHRYPDRLVLLITDMCSVYCRHCTRRRFAGSS

(101)

YDMEDPLHEDEDSPVPGLTTHRYPDRLVFLVTNQCSVYCRHCTRRRFSGQI

150

FIG. 4C

P_GI2529467_G8_AAB81159.1

P_GI2634361_EMB_CAB13860.1

P_S00701550

P_S00701551

P_S00701552

P_S01032894

Consensus

151

(151)

GMGVPKKQLDAAIAYIRETPEIRDCILISGGDGLLINDQILEYILKELRSI

(151)

GMGVPKKQLDAAIAYIRETPEIRDCILISGGDGLLINDQILEYILKELRSI

(151)

GMGVPKKQLDAAIAYIRETPEIRDCILISGGDGLLINDQILEYILKELRSI

(141)

DSAVDTKQIDAAIEYIKNTPOVRVLLSGGDALLISDEKIEYIIRRIREI

(143)

DASSPSEIRDRCIDYIANTPTVRDVLISGGDALLVSDERIEYILKRIREV

(144)

DGAMPMDRIDKATIEYIAKTPQVRVLLSGGDALLVSNKKLESIQKLPAI

(151)

GMGVPKKQLDAAIAYIRETPEIRDCILISGGDGLLINDQILEYILKELRSI

200

FIG. 4D

P_GI2529467_G8_AAB81159.1_

P_GI2634361_EMB_CAB13860.1_

P_S00701550

P_S00701551

P_S00701552

P_S01032894

Consensus

(201)

(201)

(201)

(191)

(193)

(194)

(201)

PHLEVIRIGTRAPVVFQRI

PHLEVIRIGTRAPVVFQRI

PHLEVIRIGTRAPVVFQRI

PHVEVIRIGSRVPVMPQRI

PHVEVIRIGSRTPVVLQRI

PHVEVIRIGSRTPVVLQRI

PHVEVIRIGTRAPVVFQRI

ITDHLCEILKKYHPVWLNT

ITDHLCEILKKYHPVWLNT

ITDHLCEILKKYHPVWLNT

ITPELVSMILKKYHPVWLNT

ITPQLVDMILKKYHPVWLNT

ITPELCNMLKKYHPVWLNT

ITDHLCEILKKYHPVWLNT

THFNFTSI

THFNFTSI

THFNFTSI

THFNHPNEI

THFNHPNEV

THFNHPNEV

THFNFTSI

EMTEES

EMTEES

EMTEES

EMTEES

EMTEEA

EMTEEA

EMTEES

FIG. 4E

P_GI2529467_G8_AAB81159.1_

P_GI2634361_EMB_CAB13860.1_

P_S00701550

P_S00701551

P_S00701552

P_S01032894

Consensus

(251)

(251)

(251)

(241)

(243)

(244)

(251)

VEACEKLVNAGVPVGNQAVVLAG

VEACEKLVNAGVPVGNQAVVLAG

VEACEKLVNAGVPVGNQAVVLAG

KRACELLADAGIPLGNQSVLLAG

VEACERMANAGIPIGNQTVLLRG

KKACEMLIADAGVPIGNQTVLLRG

VEACEKLVNAGVPVGNQAVVLAG

INDSVPI

INDSVPI

INDSVPI

INDSVPI

INDSVPI

INDSVPI

INDSVPI

IMKKIMHDL

IMKKIMHDL

IMKKIMHDL

IMHDKI

IMHDKI

IMHDKI

IMKKIMHDL

VKIRVRPYY

VKIRVRPYY

VKIRVRPYY

VKIRVRPYY

VKIRVRPYY

VKIRVRPYY

VKIRVRPYY

FIG. 4F

P_GI2529467_G8_AAB81159.1_
P_GI2634361_EMB_CAB13860.1_
P_S00701550
P_S00701551
P_S00701552
P_S01032894
Consensus

401

(401) EPIGLSAIFADKKEVSFTPENVDRIKRREAYIANPEHETLKDRREKRDQLK
(401) EPIGLSAIFADKKEVSFTPENVDRIKRREAYIANPEHETLKDRREKRDQLK
(401) EPIGLSAIFADKKEVSFTPENVDRIKRREAYIANPEHETLKDRREKRDQLK
(386) HKVGAGLLNGETATLEDEGLERKQGHH-----
(386) HKEGVAALSGGQQLAIEPSDLARKKRKFDKN-----
(389) EISGVYMLDEGLEMSLEFBSHLARHERNKKRAEAEGKK-----
(401) EPIGLSAIFADKKEVSFTPENVDRIKRREAYIANPEHETLKDRREKRGQLK

450

FIG. 4I

P_GI2529467_G8_AAB81159.1_
P_GI2634361_EMB_CAB13860.1_
P_S00701550
P_S00701551
P_S00701552
P_S01032894
Consensus

451

(451) EKKFLAQKKQKETECGGDSS-
(451) EKKFLAQKKQKETECGGDSS-
(451) EKKFLAQKKQKETECGGDSS-
(415) -----
(417) -----
(426) -----
(451) EKKFLAQKKQKETECGGDSS

471

FIG. 4J

-1-

SEQUENCE LISTING

<110> Chatterjee, Ranjini
 Chen, Michelle
 Louie, Susan
 Mitchell, Ken
 Fox, Richard

<120> Improved Alanine 2,3-Aminomutases and Related Polynucleotides

<130> 0359.210WO/15686WO02

<160> 64

<170> PatentIn version 3.3

<210> 1
 <211> 1416
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 1
 atgaaaaaca aatggtataa accgaaacgg cattggaagg agatcgagct atggaaggac 60
 gttccggaag agaaatggaa cgattggcctt tgacagctga cgcacactgt aagaacgtta 120
 gatgatttaa agaaagtcac taatctgacc gaggatgaag aggaaggcgt ccgtatttct 180
 accaaaaCga tccccttaaa tattacacct tactatgctt ctttaatgga ccccgacaat 240
 ccgagatgcc cggtacgcat gcagtctgtg ccgctttctg aagaaatgca caaaacaaaa 300
 tacgataTgg aagaccgct tcatgaggat gaagattcac cggtaaccgg tctgacacac 360
 cgctatCCg accgtgtgct gtttcttgtc acgaatcaat gttccgtgta ctgccgccac 420
 tgcacacgcc ggcgccttttc cggacaaatc ggaatgggcg tccccaaaaa acagcttgat 480
 gctgcaattg cttatatccg ggaaacaccc gaaatccgcg attgtttaat ttcaggcggt 540
 gatgggcTgc tcatcaacga ccaaatttta gaatatattt taaaagagct gcgcagcatt 600
 ccgcacTgg aagtcacccg catcggaaca cgtgctcccg tcgtctttcc gcagcgcatt 660
 accgatcatc tgtgcgagat attgaaaaaa taccatccgg tccggctgaa caccattttt 720
 aacacaagca tcgaaatgac agaagaaccc gttgaggcac gtgaaaagct ggtgaacgcg 780
 ggagtgcCgg tcggaaatca ggctgtcgta ttagcaggta ttaatggctc ggttccaatt 840
 atgaaaaagc tcatgcatga cttggtaaaa atcagagctc gtccttatta tatttaccaa 900
 tgtgatCtgt cagaaggaat aaggcatttc cgtgctcctg tttccaaagg tttggagatc 960
 attgaagggc tgagaggtca tacctcaggc tatgcgggtc ctacctttgt cgttcacgca 1020

-2-

```

ccaggcggag ggggtaaaat cgccctgcag ccgaactatg tctgtctca aagtcccgac 1080
aaagtgatct taagaaattt tgaagggtgtg attacgtcat atccggaacc agagaattgt 1140
accccccaatc aggacagcgc ctatttttgag tccgttttcc ctgaaaccgc tgacaaaaag 1200
gagccgatcg ggctgagtgc cttttttgct gacaaagaag ttctgtctac acccgaaaat 1260
gtagacagaa tcaaacggcg tgaggcatac atcgcaaadc cggagcatga aacattagaa 1320
gatcggcggtg agaaaagagg tcagctcaaa gaaaagaaat ttttggcgca gcagaaaaaa 1380
cagaaagaga ctgaatgcgg aggggattct tcataa 1416

```

<210> 2
 <211> 471
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 2

```

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
1           5           10           15

```

```

Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
          20           25           30

```

```

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

```

```

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

```

```

Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Ser Leu Met Asp Pro Asp Asn
65           70           75           80

```

```

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
          85           90           95

```

```

His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu His Glu Asp Glu Asp
          100          105          110

```

```

Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
          115          120          125

```

-3-

Leu Val Thr Asn Gln Cys Ser Val Tyr Cys Arg His Cys Thr Arg Arg
 130 135 140

Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
 145 150 155 160

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
 165 170 175

Ile Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr
 180 185 190

Ile Leu Lys Glu Leu Arg Ser Ile Pro His Leu Glu Val Ile Arg Ile
 195 200 205

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

Cys Glu Ile Leu Lys Lys Tyr His Pro Val Arg Leu Asn Thr His Phe
 225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Pro Val Glu Ala Arg Glu Lys
 245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
 260 265 270

Gly Ile Asn Gly Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
 275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
 290 295 300

Glu Gly Ile Arg His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
 305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
 325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
 340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

-4-

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Cys Thr Pro Asn Gln
 370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
 385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Ser
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

Asn Pro Glu His Glu Thr Leu Glu Asp Arg Arg Glu Lys Arg Gly Gln
 435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 3
 <211> 1416
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 3
 atggaaaaca aatggtataa accgaaacgg cattggaagg agatcgagtt atggaaggac 60
 gttccggaag agaaatggaa cgattggctt tgacagctga cacacactgt aagaacgtta 120
 gatgatttaa agaaagtcac taatctgacc gaggatgaag aggraaggcgt ccgtatttct 180
 accaaaacga tccccttaaa tattacacct tactatgctt ctttaatgga ccccgacaat 240
 ccgagatgcc cggtacgcat gcagtctgtg ccgctttctg aagaaatgca caaaacaaaa 300
 tacgatatgg aagaccgct tcatgaggat gaagattcac cgggtgcccgg tctgacacac 360
 cgctatcccg accgtgtgct gtttcttctg acgaatcagt gttccgtgta ctgccgccac 420
 tgcacacgcc ggcgcttttc cggacaaatc ggaatgggcg tcccaaaaa acagcttgat 480
 gctgcaattg cttatatccg ggaaacaccc gaaatccgcg attgtttaat ttcaggcggc 540
 gatgggctgc tcatcaacga ccaaatttta gaatatattt taaaagagct gcgcagcatt 600

-5-

```

ccgcatctgg aagtcacccg catcggaaca cgtgctcccg tcgtctttcc gcagcgcgtt 660
accgatcatc tgtgcgagat attgaaaaaa tatcatcccg tctggctgga caccattttt 720
aacacaagca tcgaaatgac agaagaatcc gttgaggcat gtgaaaagct ggtgaacgcg 780
ggagtgccgg tcggaaatca ggctgtcgta ttacgaggta ttaatgattc ggttccaatt 840
atgaaaaagc tcatgcatga cttggtaaaa atcagagtcc gtccttatta tatttaccaa 900
tgtgatctgt cagaaggaat aaggcatttc cgtgctcctg tttccaaagg tttggagatc 960
attgaagggc tgagaggtca tacctcaggc tatgcggttc ctacctttgt cgttcacgca 1020
ccaggcggag gaggtaaaat cgcctgcag ccgaactatg tcctgtctca aagtcttggc 1080
agagtgatct taagaaattt tgaagggtg attacgtcat acccggaacc agagaattat 1140
atccccaatc aggcagacgc ctattttgag tcggttttcc ctgaaaccgc tgacaaaaag 1200
gagccgatcg ggctgagtgc cttttttgct gacaaagaag tttcgtctac acctgaaaat 1260
gtagacagaa tcaaacggcg tgaggcatac atcgcaaadc cggagcatga aacattaaaa 1320
gatcggcggtg agaaaagagg tcagctcaaa gaaaagaaat ttttggcgca gcagaaaaaa 1380
cagaaagaga ctgaatgcgg aggggattct tcataa 1416

```

<210> 4
 <211> 471
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 4

```

Met Glu Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
1           5           10           15

```

```

Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
           20           25           30

```

```

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

```

```

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

```

```

Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Ser Leu Met Asp Pro Asp Asn
65           70           75           80

```

-6-

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
85 90 95

His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu His Glu Asp Glu Asp
100 105 110

Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
115 120 125

Leu Val Thr Asn Gln Cys Ser Val Tyr Cys Arg His Cys Thr Arg Arg
130 135 140

Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
145 150 155 160

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
165 170 175

Ile Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr
180 185 190

Ile Leu Lys Glu Leu Arg Ser Ile Pro His Leu Glu Val Ile Arg Ile
195 200 205

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

Cys Glu Ile Leu Lys Lys Tyr His Pro Val Trp Leu Asp Thr His Phe
225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
290 295 300

-7-

Glu Gly Ile Arg His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
 305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
 325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
 340 345 350

Tyr Val Leu Ser Gln Ser Pro Gly Arg Val Ile Leu Arg Asn Phe Glu
 355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
 385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Ser
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
 435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 5

<211> 1416

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 5

atgaaaaaca aatggtataa accgaaacgg CATTGGAAGG agatcgagtt atggaaggac 60

gttcCGGAAG ggaaatggaa cgattggctt Tgacagctga cacacactgt aagaacgtta 120

gatgatttaa agaaagtcac taatctgacc Gaggatgaag aggaaggcgt ccgtatttct 180

-8-

```

accaaaacga tccccttaaa tattacacct tactatgctt ctttaatgga ccccgacaat 24 0
ccgagatgcc cggtagcat gcagtctgtg ccgctttctg aagaaataca caaaacaaaa 30 0
tacgatatgg aagacccgct tcatggggat gaagactcac cggtagccgg tctgacacac 36 0
cgctatcccg accgtgtgct gtttcttgtc acgaatcaat gttctgtgta ctgccgccac 42 0
tgcacacgcc ggcgcttttc cggacaaatc ggaatgggcg tccccaaaaa acagcttgat 48 0
gctgcaattg cttatatccg ggaaacaccc gaaatccgcg attgtttaat ttcaggcggg 54 0
gatgggctgc tcatcaacga ccaaatttta gaatatattt taaaagagct gcgcagcatt 60 0
ccgcatctgg aagtcacccg catcggaaca cgtgcccccg tcgtctttcc gcagcgcatt 66 0
accgatcatc tgtgcgagat attgaaaaaa tatcatccgg tctggctgaa caccattttt 72 0
aacacaagca tcgaaatgac agaagaatcc gttgaggcat gtgaaaagct ggtgaacgcg 78 0
ggagtgccgg tcggaaatca ggctgtcgta ttagcaggta ttaatgattc ggtccaatt 84 0
atgaaaaagc tcatgcatga cttggtaaaa atcagagtcc gtccttatta tatttaccaa 90 0
tgtgatctgt cagaaggaat aaggcatttc cgtgctcctg tttccaaagg tttggagatc 96 0
attgaagggc tgagaggtca tacctcaggc tatgcggttc ctacctttgt cgttcacgca 102 0
ccaggcggag gaggtaaaaat cgccctgcag ccgaactatg tcctgtctca aagtcctgac 108 0
aaagtgatct taagaaattt tgaagggtgtg attacgtcat atccggaacc agagaattat 114 0
atccccaatc aggcagacgc ctattttgag tccgttttcc ctgaaaccgc tgacaaaaag 120 0
gagccgatcg ggctgagtgc catttttgct gacaaagaag tttcgtctac acctgaaaat 126 0
gtagacagaa tcaaacggcg tgaggcatac atcgcaaatc cggagcatga aacattaaaa 132 0
gatcggcgtg agaaaagagg tcagctcaaa gaaaagaaat ttttggcgca gcagaaaaaa 138 0
cagaaagaga ctgaatgcgg aggggattct tcataa 144 0

```

<210> 6

<211> 471

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 6

```

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
1           5           10          15

```

-9-

Leu Trp Lys Asp Val Pro Glu Gly Lys Trp Asn Asp Trp Leu Trp Gln
 20 25 30

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

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

Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Ser Leu Met Asp Pro Asp Asn
 65 70 75 80

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Ile
 85 90 95

His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu His Gly Asp Glu Asp
 100 105 110

Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
 115 120 125

Leu Val Thr Asn Gln Cys Ser Val Tyr Cys Arg His Cys Thr Arg Arg
 130 135 140

Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
 145 150 155 160

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
 165 170 175

Ile Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr
 180 185 190

Ile Leu Lys Glu Leu Arg Ser Ile Pro His Leu Glu Val Ile Arg Ile
 195 200 205

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

Cys Glu Ile Leu Lys Lys Tyr His Pro Val Trp Leu Asn Thr His Phe
 225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
 245 250 255

-10-

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
290 295 300

Glu Gly Ile Arg His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Ser
405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
450 455 460

Glu Cys Gly Gly Asp Ser Ser
465 470

-11-

<210> 7
 <211> 1416
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 7
 atgaaaaaca aatggtat~~t~~aa accgaaacgg cattggaagg agatcgagtt atggaaggac 60
 gttccggaag agaaatg~~g~~aa cgattggc~~t~~t tgacagctga cacacactgt aagaacgtta 120
 gatgatttaa agaaagt~~c~~at taatctgacc gaggatgaag aggaaggcgt ccgtatttct 180
 accaaaacga tcccctt~~a~~aa tattacacca tactatgcga gcttaatgga tccagaaaaac 240
 ccacgttgtc cggtacg~~c~~at gcagtctgtg ccgctttccg aagaaatgca caaaacaaaa 300
 tacgatatgg aagaccc~~g~~ct tcatgaggat gaagattcac cggtagccgg tctgacacac 360
 cgctatcccg accgtgt~~g~~ct gtttcttgtc acgaatcaat gttccgtgta ctgccgccac 420
 tgcacacgcc ggcgctt~~t~~tc cggacaaatc ggaatggg~~c~~g tccccaaaa acagcttgat 480
 gctgcaattg cttatat~~c~~cg ggaaacaccc gaaatccg~~c~~g attgtttaat ttcaggcgggt 540
 gatgggctgc tcatcaa~~c~~ga ccaaatttta gaatatattt taaaagagct gcgcagcatt 600
 ccgcatctgg aagtc~~a~~tc~~c~~g catcggaaca cgtgctcccg t~~c~~gtctttcc gcagcgcatt 660
 accgatcatc cgtgcga~~g~~at attgaaaaaa tatcatccgg tctggctgaa caccattttt 720
 aacacaagca tcgaaat~~g~~ac agaagaatcc gttgaggcat gtgaaaagct ggtgaacgcg 780
 ggagtgccgg tcggaat~~t~~ca ggctgtcgta ttagcaggta ttaatgattc ggttccaatt 840
 atgaaaaagc tcatgca~~t~~ga cttggtaaaa atcagagtcc gtccttatta tatttaccaa 900
 tgtgatctgt cagaagg~~a~~at aaggcatttc cgtgctcctg tctccaaagg tttggagatc 960
 attgaagggc tgagagg~~t~~ca tacc~~c~~caggc tatgcggttc ctacctttgt cgttcacgca 1020
 ccaggcgggag gaggtaa~~a~~at cgccttcgag ccgaactatg t~~c~~ctgtctca aagtcctgac 1080
 aaagtgatct taagaaa~~t~~tt tgaaggtgtg attacgtcat atccggaacc agagaattat 1140
 atccccaatc aggcaga~~c~~gc ctatttttgag tccgtttccc ctgaaaccgc tgacaaaaag 1200
 gagccgatcg ggctgag~~t~~gc cattttt~~g~~ct gacaaagaag tttcgtctac acctgaaaat 1260
 gtagacagaa tcaaacg~~g~~c~~g~~ tgaggcctac atcgcaaatc cggagcatga aacattaaaa 1320
 gatcggcgtg agaaaag~~a~~gg tcagctcaaa gaaaagaaat tttcggcgca gcagaaaaaa 1380
 cagaaagaga ctgaatg~~c~~gg aggggattct tcataa 1416

-12-

<210> 8
 <211> 471
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 8

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
 1 5 10 15

Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
 20 25 30

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

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

Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Ser Leu Met Asp Pro Glu Asn
 65 70 75 80

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
 85 90 95

His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu His Glu Asp Glu Asp
 100 105 110

Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
 115 120 125

Leu Val Thr Asn Gln Cys Ser Val Tyr Cys Arg His Cys Thr Arg Arg
 130 135 140

Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
 145 150 155 160

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
 165 170 175

Ile Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr
 180 185 190

-13-

Ile Leu Lys Glu Leu Arg Ser Ile Pro His Leu Glu Val Ile Arg Ile
 195 200 205

Gly Thr Arg Ala Pro Val Val Phe Pro Gln Arg Ile Thr Asp His Pro
 210 215 220

Cys Glu Ile Leu Lys Lys Tyr His Pro Val Trp Leu Asn Thr His Phe
 225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
 245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
 260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
 275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
 290 295 300

Glu Gly Ile Arg His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
 305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Pro Gly Tyr Ala Val Pro Thr Phe
 325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
 340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Ser Pro Glu Thr Ala Asp Lys Lys
 385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Ser
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

-14-

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
 435 440 445

Leu Lys Glu Lys Lys Phe Ser Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 9
 <211> 1416
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 9
 atgaaaaaca aatggtataa accgaaacgg cattggaagg agatcgagtt atggaaggac 60
 gttccggaag agaaatggaa cgattggcct tgacagctga cacacactgt aagaacgtta 120
 gatgatttaa agaaagtcac taatctgacc gaggatgaag aggaaggcgt ccgtatttct 180
 accaaaacga tccccttaaa tattacacct tactatgctt ctttaatgga ccccgacaat 240
 ccgagatgcc cgggtgcgcat gcagtctgtg ccgctttctg aagaaatgca caaaacaaaa 300
 tacgatatgg aagaccgcgt tcatgaggat gaagattcac cggtaaccgg tctgacacac 360
 cgctatcccg accgtgtgct gtttcttctg acgaatcaat gttccgtgta ctgccgccac 420
 tgcacacgcc ggcgcttttc cggacaaatc ggaatgggcg tccccaaaa acagcttgat 480
 gctgcaattg cttatatccg ggaaacaccc gaaatccgcg attgtttaat ttccaggcgg 540
 gatgggctgc tcatcaacga ccaaatttta gaatatattt taaaagagct gcgcagcatt 600
 ccgcatctgg aagtcacccg catcggaaca cgtgctcccg tcgtctttcc gcagcgcatt 660
 accgatcatc tgtgcgagat attgaaaaaa tatcatccgg tctggctgaa caaccatttt 720
 aacacaagca tcgaaatgac agaagaatcc gttgaggcat gtgaaaagct ggtgaacgcg 780
 ggagtgcccg tcggaaatca ggctgtcgta ttagcaggta ttaatgattc ggttccaatt 840
 atgaaaaagc tcatgcatga cttggtaaaa atcagagtcg gtccttatta tgtttaccaa 900
 tgtgatctgt cagaaggaat aaggcatttc cgtgctcctg tttccaaagg ttgggagatc 960
 attgaagggc tgagagggtca tacctcaggc tatgcggttc ctacctttgt cgttcacgca 1020
 ccaggcggag gaggtaaaaat cgccctgcag ccgaactatg tcctgtctca aagtcctgac 1080

-15-

aaagtgatct taagaaatTT tgaaggTgtg attacgtcat atccggaacc agagaattat 1140
 atccccaatc aggcagacgc ctatTTtgag tccgtTTtcc ctgaaaccgc tgacaaaaag 1200
 gagccgatcg ggctgagtgc cattTTtgct gacaaagaag tttcgtctac acctgaaaat 1260
 gtagacagaa tcaaacggcg tgaggcatac atcgcaaatc cggagcatga aacattaataa 1320
 gatcggcgTg agaaaagagg tcagctcaaa gaaaagaaat ttttggcgca gcagaaaaaa 1380
 cagaaagaga ctgaatgcgg aggggattct tcataa 1416

<210> 10
 <211> 471
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 10

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
1 5 10 15

Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
20 25 30

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

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

Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Ser Leu Met Asp Pro Asp Asn
65 70 75 80

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
85 90 95

His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu His Glu Asp Glu Asp
100 105 110

Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
115 120 125

Leu Val Thr Asn Gln Cys Ser Val Tyr Cys Arg His Cys Thr Arg Arg
130 135 140

-16-

Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
 145 150 155 160

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
 165 170 175

Ile Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr
 180 185 190

Ile Leu Lys Glu Leu Arg Ser Ile Pro His Leu Glu Val Ile Arg Ile
 195 200 205

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

Cys Glu Ile Leu Lys Lys Tyr His Pro Val Trp Leu Asn Thr His Phe
 225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
 245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
 260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
 275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Val Tyr Gln Cys Asp Leu Ser
 290 295 300

Glu Gly Ile Arg His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
 305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
 325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
 340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

-17-

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
 385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Ser
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
 435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 11
 <211> 1416
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 11
 atgaaaaaca aatggtataa accgaaacg cattggaagg agatcgagtt atggaaggac 60
 gtcccggaag agaaatggaa cgattggctt tgacagctga cacacactgt aggaacgtta 120
 gatgatttaa agaaagtcac caatctgacc gaggatgaag aggaaggcgt ccgtatttct 180
 accaaaacga tccccttaaa tattacacct tactatgctt ctttaatgga ccccgacaat 240
 ccgagatgcc cggtagcat gcagtctgtg ccgctttctg aagaaatgca caaaacaaaa 300
 tacgatatgg aagaccgct tcatgaggat gaagattcac cggtaccggt tctgacacac 360
 cgctatcccg accgtgtgct gtttcttgtc acgaatcaag gttccgtgta ctgccgccac 420
 cgcacacgcc ggcgcttttc cggacaaatc ggaatgggag tccccaaaa acagcttgat 480
 gctgcaattg cttatatccg ggaaacaccc gaaatccgag attgtttaat ttcaggcggt 540
 gatgggctgc tcatcaacga ccaaatttta gaatatattt taaaagagct gcgcagcatt 600
 ccgcatccgg aagtcacccg catcggaaca cgtgctcccg tcgtcttccc gcagcgcatt 660

-18-

```

accgatcatc tgtgcgagat attgaaaaaa tatcatccgg tctggctgaa caccatttt 720
aacacaagca tcgaaatgac agaagaatcc gttgaggcat gtgaaaagct ggtgaacgcg 780
ggagtgccgg tcggaaatca ggctgtcgta ttagcaggta ttaatgattc ggttccaact 840
atgaaaaagc tcatgcatga cttggtaaaa atcagagtcc gtccttatta tatttaccaa 900
tgtgatctgt cagaaggaat aaggcatttc cgtgctcctg tttccaaagg tttggagatc 960
attgaagggc tgagaggcca tacctcaggc tatgogggttc ctacctttgt cgttcacgca 1020
ccaggcggag gaggtaaaaat cgccctgcag ccgaactatg tcctgtctca aagtcctgac 1080
aaagtgatct taagaaatct tgaagggtgtg attacgtcat atccggaacc agagaattat 1140
atccccaatc aggcagacgc ctattttgag tccgttttcc ctgaaaccgc tgacaaaaag 1200
gagccgatcg ggctgagtgc catttttgct gacaaagaag tttcgtctac acctgaaaat 1260
gtagacagaa tcaaacggcg tgaggcatac atcgcaaatc cggagcatga aacattaaaa 1320
gatcggcgtg agaaaagagg tcagctcaaa gaaaagaaat ttttggcgca gcagaaaaaa 1380
cagaaagaga ctgaatgcgg aggggattct tcataa 1416

```

<210> 12
 <211> 471
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 12

```

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
1           5           10           15

```

```

Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
          20           25           30

```

```

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

```

```

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

```

```

Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Ser Leu Met Asp Pro Asp Asn
65           70           75           80

```

-19-

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
85 90 95

His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu His Glu Asp Glu Asp
100 105 110

Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
115 120 125

Leu Val Thr Asn Gln Gly Ser Val Tyr Cys Arg His Arg Thr Arg Arg
130 135 140

Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
145 150 155 160

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
165 170 175

Ile Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr
180 185 190

Ile Leu Lys Glu Leu Arg Ser Ile Pro His Pro Glu Val Ile Arg Ile
195 200 205

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

Cys Glu Ile Leu Lys Lys Tyr His Pro Val Trp Leu Asn Thr His Phe
225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
260 265 270

Gly Ile Asn Asp Ser Val Pro Thr Met Lys Lys Leu Met His Asp Leu
275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
290 295 300

Glu Gly Ile Arg His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
305 310 315 320

-20-

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
 325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
 340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
 385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Ser
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
 435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 13

<211> 1416

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 13

atgaaaaaca aatggtataa accgaaacgg cattgggagg agatcgagcg atggaaggac 60

gttccggaag agaaatggaa cgattggcct tgacagctga cacacactgt aagaacgtta 120

gatgatttaa agaaagtcac taatctgacc gaggatgaag aggaaggcgt ccgtatttct 180

acaaaaacga tccccttaaa tattacacct tactatgctt ccttaatgga ccccgacaat 240

-21-

```

ccgagatgcc cggtagcat gcagtctgtg ccgctttctg aagaaatgca caaaacaaaa 300
tacgatatgg aagaccgct tcatgaggat gaagattcac cggtagccgg tctgacacac 360
cgctatcccg accgtgtgct gtttcttgtc acgaatcaat gttccgtgta ctgccgccac 420
tgcacacgcc ggcgcttttc cggacaaatc gggatgggcg tccccaaaaa acagcttgat 480
gctgcaattg cttatatccg ggaaacaccc gaaatccgcg attgtttaat ttcaggcggt 540
gatgggctgc tcatcaacga ccaaatttta gaatatattt taaaagagcc gcgcagcact 600
ccgcatctgg aagtcacccg catcggaaca cgtgctcccg tcgtctttcc gcagcgcatt 660
accgatcatc tgtgcgagat attgaaaaaa tatcatccgg tctggctgaa caccattttt 720
aacacaagca tcgaaatgac agaagaatcc gttgaggcat gtgaaaagct ggtgaacgcg 780
ggagtgccgg tcggaaatca ggctgtcgta ttagcaggta ttaatgattc ggttccaatt 840
gtgaaaaagc tcatgcatga ctgggtaaaa atcagagtcc gtccctatta tatttaccaa 900
tgtgatctgt cagaaggaat aaggcattcc cgtgctcctg tttccaaagg tttggagatc 960
attgaagggc tgagaggatca tacctcaggc tatgcgggtc ctacctttgt cgttcacgca 1020
ccaggcggag gaggtaaaaat cgccctgcag ccgaactatg tcctgtctca aagtctgac 1080
aaagtgatct taagaaattt tgaagggtgtg attacgtcat atccggaacc agagaattat 1140
atccccaatc aggagacgc ctattttgag tccgttttcc ctgaaaccgc tgacaaaaag 1200
gagccgatcg ggctgagtgc cttttttgct gacaaagaag tttcgtctac acctgaaaat 1260
gtagacagaa tcaaacggcg tgaggcatat atcgcaaatc cggagcatga aacattaaaa 1320
gatcggcgtg agaaaagagg tcagctcaaa gaaaagaaat ttttggcgca gcagaaaaaa 1380
cagaaagaga ctgaatgcgg aggggattct tcataa 1416

```

<210> 14

<211> 471

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 14

```

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Glu Glu Ile Glu
1           5           10           15

```

```

Arg Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
          20           25           30

```

-22-

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

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

Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Ser Leu Met Asp Pro Asp Asn
 65 70 75 80

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
 85 90 95

His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu His Glu Asp Glu Asp
 100 105 110

Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
 115 120 125

Leu Val Thr Asn Gln Cys Ser Val Tyr Cys Arg His Cys Thr Arg Arg
 130 135 140

Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
 145 150 155 160

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
 165 170 175

Ile Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr
 180 185 190

Ile Leu Lys Glu Pro Arg Ser Thr Pro His Leu Glu Val Ile Arg Ile
 195 200 205

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

Cys Glu Ile Leu Lys Lys Tyr His Pro Val Trp Leu Asn Thr His Phe
 225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
 245 250 255

-23-

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
 260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Val Lys Lys Leu Met His Asp Leu
 275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
 290 295 300

Glu Gly Ile Arg His Ser Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
 305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
 325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
 340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
 385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Ser
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
 435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 15

<211> 1416

-24-

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 15

```

atgaaaaaca aatggtataa accgaaacgg cattgggagg agatcgagcg atggaaggac      60
gttccggaag agaaatggaa cgattggcct tgacagctga cacacactgt aagaacgtta      120
gatgatttaa agaaagtcac taatctgacc gaggatgaag aggaaggcgt ccgtatttct      180
acaaaaacga tccccctaaa tattacacct tactatgctt ccttaatgga ccccgacaat      240
ccgagatgcc cggtagcat gcagctctgtg ccgctttctg aagaaatgca caaaacaaaa      300
tacgatatgg aagaccgct tcatgaggat gaagattcac cggtagccgg tctgacacac      360
cgctatcccg accgtgtgct gttcttctgc acgaatcaat gttccgtgta ctgccgccac      420
tgcacacgcc ggcgcttttc cggaacaaatc gggatgggcg tccccaaaaa acagcttgat      480
gctgcaattg cttatatccg ggaaacaccc gaaatccgcg attgtttaat ttcaggcggg      540
gatgggctgc tcatcaacga ccaaatctta gaatatattt taaaagagcc ggcgagcact      600
ccgcatctgg aagtcacccg catcgggaaca cgtgctcccg tcgtctttcc gcagcgcatt      660
accgatcatc tgtgcgagat attgaaaaaa tatcatccgg tctggctgaa caccattttt      720
aacacaagca tcgaaatgac agaagaatcc gttgaggcat gtgaaaagct ggtgaacgcg      780
ggagtgccgg tcggaaatca ggctgtcgta ttagcaggta ttaatgattc ggttccaatt      840
gtgaaaaagc tcatgcatga cttggtaaaa atcagagtcc gtccttatta tattttaccaa      900
tgtgatctgt cagaaggaat aaggcattcc cgtgctcctg tttccaaagg tttggagatc      960
attgaagggc tgagaggcca taccctcaggc tatgcgggtc ctacctttgt cgttcacgca     1020
ccaggcggag gaggtaaaat cgccctgcag ccgaactatg tcctgtctca aagtcctgac     1080
aaagtgatct taagaaattt tgaaagggtgtg attacgtcat atccggaacc agagaattat     1140
atccccaatc aggagacgc ctatcttgag tccgttttcc ctgaaaccgc tgacaaaaag     1200
gagccgatcg ggctgagtgc cattcttgct gacaaagaag tttcgtctac acctgaaaat     1260
gtagacagaa tcaaacggcg tgaggcatac atcgcaaatc cggagcatga aacattaaaa     1320
gatcggcgtg agaaaagagg tcagctcaaa gaaaagaaat ttttggcgca gcagaaaaaa     1380
cagaaagaga ctgaatgcgg aggggattct tcataa                                1446

```

<210> 16

<211> 471

-25-

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 16

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Glu Glu Ile Glu
 1 5 10 15

Arg Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
 20 25 30

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

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

Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Ser Leu Met Asp Pro Asp Asn
 65 70 75 80

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
 85 90 95

His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu His Glu Asp Glu Asp
 100 105 110

Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
 115 120 125

Leu Val Thr Asn Gln Cys Ser Val Tyr Cys Arg His Cys Thr Arg Arg
 130 135 140

Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
 145 150 155 160

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
 165 170 175

Ile Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr
 180 185 190

Ile Leu Lys Glu Pro Arg Ser Thr Pro His Leu Glu Val Ile Arg Ile
 195 200 205

-26-

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

Cys Glu Ile Leu Lys Lys Tyr His Pro Val Trp Leu Asn Thr His Phe
 225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
 245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
 260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Val Lys Lys Leu Met His Asp Leu
 275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
 290 295 300

Glu Gly Ile Arg His Ser Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
 305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
 325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
 340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
 385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Ser
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

-27-

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
 435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 17

<211> 1416

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 17

```

atgaaaaaca aatggtataa accgaaacgg cattggaagg agatcgagtt atggaaggac      60
gttccggaag agaaatggaa cgattggctt tgacggctga cacacactgt aagaacgtta      120
gatgatttaa agaaagtcac taatctgacc gaggatgaag aggaaggcgt ccgtatttct      180
acaaaaacga tccccttaaa tattacacct tactatgctc ctttaatgga cccgcacaaat      240
ccgagatgcc cggtacgcat gcagtctgtg ccgctttccg aagaaatgca caaaacaaaa      300
tacgatatgg aagaccgcgt tcatgaggat gaagatacac cggtagccgg tccgacacac      360
cgctatcccc accgtgtgct gtttcttgct acgaatcaat gctccgtgta ctgccgcacac      420
tgcacacgcc ggcgcttttc cggacaaatc ggaatgggcg tccccaaaaa acagcttgat      480
gctgcaattg cttatatccg ggaaacaccc gaaatccgcg attgtttaat ttcaggcggg      540
gatgggctgc tcatcaacga ccaaatttta gaatatattt taaaagagct gcgcagcatt      600
ccgcatctgg aagtcacccg catcggaaca cgtgctcccg tcgtctttcc gcagcgcatt      660
accgatcatc tgtgcgagat attgaaaaaa tatcatccgg tctggctgaa caccatcttt      720
aacacaagca tcgaaatgac agaagaatcc gttgaggcat gtgaaaagct ggtgaacgcg      780
ggagtgccgg tcggaaatca ggctgtcgta ttagcaggta ttaatgattc ggttccaatt      840
atgaaaaagc tcatgcatga cttggtaaaa atcagagtcc gtccttatta tatttaacaa      900
tgtgatctgt cagaaggaat aaggcatttc cgtgctcctg tttccaaagg tttggagatc      960
attgaagggc tgagaggtca tacctcaggc tatgcggttc ctacctttgt cgttcacgca     1020
ccaggcgga gaggtaaaat cgccctgcag ccgaactatg tcctgtctca aagtcctgac     1080
aaagtgatct taagaaatth tgaaggtgtg attacgtcat atccggaacc agagaattat     1140

```

-28-

```

atccccaatc aggcagacgc ctattttgag tccgttttcc ctgaaaccgc tgacaaaag 1200
gagccgatcg ggctgagtgc catttttgct gacaaagaag ttctgtctac acctgaaaat 1260
gtagacagaa tcaaacggcg tgaggcatatc atcgcaaatac cggagcatga aacattaata 1320
gatcggcggtg agaaaagagg tcagctcaaa gaaaagaaat ttttggcgca gcagaaaaaa 1380
cagaaagaga ctgaatgcgg aggggattct tcataa 1416

```

<210> 18

<211> 471

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 18

```

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
1           5           10           15

```

```

Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Arg
          20           25           30

```

```

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

```

```

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

```

```

Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Pro Leu Met Asp Pro Asp Asn
65           70           75           80

```

```

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
          85           90           95

```

```

His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu His Glu Asp Glu Asp
          100          105          110

```

```

Thr Pro Val Pro Gly Pro Thr His Arg Tyr Pro Asp Arg Val Leu Phe
          115          120          125

```

```

Leu Val Thr Asn Gln Cys Ser Val Tyr Cys Arg His Cys Thr Arg Arg
          130          135          140

```

-29-

Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
 145 150 155 160

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
 165 170 175

Ile Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr
 180 185 190

Ile Leu Lys Glu Leu Arg Ser Ile Pro His Leu Glu Val Ile Arg Ile
 195 200 205

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

Cys Glu Ile Leu Lys Lys Tyr His Pro Val Trp Leu Asn Thr His Phe
 225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
 245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
 260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
 275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
 290 295 300

Glu Gly Ile Arg His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
 305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
 325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
 340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

-30-

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Ser
405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
450 455 460

Glu Cys Gly Gly Asp Ser Ser
465 470

<210> 19

<211> 1416

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 19

```

atggaaaaca aatggtataa accgaaacgg cattggaagg agatcgagtt atggaaggac      60
gttccggaag agaaatggaa cgattggctt tgacagctga cacacactgt aagaacgtta      120
gatgatttaa agaaagtcac taatctgacc gaggatgaag aggaaggcgt ccgtatttct      180
acaaaaacga tccccttaaa tattacacct tactatgctt ctttaatgga cccgcacaa      240
ccgagatgcc cggtagcat gcagtctgtg ccgctttctg aagaaatgca caaaacaaaa      300
tacgatatgg aagaccgcgt tcatgaggat gaagattcac cggtagccgg tctgacacac      360
cgctatcccg accgtgtgct gtttcttgct acgaatcaat gttccgtgta ctgccgccac      420
tgacacagcc ggcgcttttc cggacaaatc ggaatgggcg tccccaaaaa acagcttgat      480
gctgcaattg cttatatccg ggaaacaccc gaaatccgcg attgtttaat ttcaggcggg      540
gatgggctgc tcatcaacga ccaaatttta gaatatattt taaaagagct gcgcagcatt      600
ccgcatctgg aagtcacccg catcggaaca cgtgctcccg tcgtctttcc gcagcgatt      660
accgatcatc tgtgcgagat attgaaaaaa tatcatccgg tctggctgaa caccatttt      720

```

-31-

```

aacacaagca tcgaaatgac agaagaatcc gttgaggcat gtgaaaagct ggtgaacgcg      780
ggagtgccgg tcggaaatca ggctgtcgta ttagcaggta ttaatgattc ggttccaatt      840
atgaaaaagc tcatgcatga cttggtaaaa atcagagtcc gtccttatta tatttaccaa      900
tgtgatctgt ctgagggctt ggggcatttc cgtgctcctg tttccaaagg tttggagatc      960
attgaagggc tgagagggtca tacctcaggc tatgcggttc ctacctttgt cgttcacgca    1020
ccaggcggag gaggtaaaat cgcctgcag ccgaactatg tcctgtcaca aagtcctgac    1080
aaagtgatct taagaaatth tgaagggtgtg attacgtcat atccggaacc agagaattat    1140
atccccaatc aggcagacgc ctattttgag tccgttttcc ctgaaaccgc tgacaaaag      1200
gagccgatcg ggctgagtgc catttttgct gacaaagaag tttcgtttac acctgaaat      1260
gtagacagaa tcaaacggcg tgaggcatat atcgcaaata cggagcatga aacattaata      1320
gatcggcgtg agaaaagaga tcagctcaaa gaaaagaaat ttttggcgca gcagaaaata      1380
cagaaagaga ctgaatgcgg aggggattct tcataa                                1416

```

<210> 20

<211> 471

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 20

```

Met Glu Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
1              5              10              15

```

```

Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
20              25              30

```

```

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

```

```

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

```

```

Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Ser Leu Met Asp Pro Asp Asn
65              70              75              80

```

```

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
85              90              95

```

-32-

His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu His Glu Asp Glu Asp
 100 105 110

Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
 115 120 125

Leu Val Thr Asn Gln Cys Ser Val Tyr Cys Arg His Cys Thr Arg Arg
 130 135 140

Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
 145 150 155 160

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
 165 170 175

Ile Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr
 180 185 190

Ile Leu Lys Glu Leu Arg Ser Ile Pro His Leu Glu Val Ile Arg Ile
 195 200 205

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

Cys Glu Ile Leu Lys Lys Tyr His Pro Val Trp Leu Asn Thr His Phe
 225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
 245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
 260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
 275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
 290 295 300

Glu Gly Leu Gly His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
 305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Phe
405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Asp Gln
435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
450 455 460

Glu Cys Gly Gly Asp Ser Ser
465 470

```
<210> 21
<211> 1416
<212> DNA
<213> Artificial Sequence
```

<220>
<223> Synthetic Construct

```
<400> 21
atgaaaaaca aatggtataa accgaaacgg cattggaagg agatcgagtt atggaaggac      60
gtcccggaag agaaatggaa cgattggcctt tgacagctga cacacactgt aagaacgtta      120
gatgatttaa agaaagtcac taatctgacc gaggatgagg aggaaggcgt ccgtatttct      180
acaaaaacga tccccttaaa tattacacct taccatgctt ctttaatgga ccccgacaat      240
ccgagatgcc cggtacgcat gcagtcctgtg ccgctttctg aagaaatgca caaaacaaaa      300
```

-34-

```

tacgacatgg aagacccgct tcatgaggat gaagattcac cggtagccgg tcgacacac 360
cgctatcccg accgtgtgct gtttcttgtc acgaatcaat gttccgtgta ctgccgccac 420
tgcacacgcc ggctcttttc cggacaaatc ggaatgggag tccccaaaaa acagcttgat 480
gctgcaattg cttatatccg ggaaacaccc gaaatccgag attgtttaat ttgaggcggg 540
gatgggctgc tcatcaacga ccaaatttta gaatatattt taaaagagct gcgcagcatt 600
ccgcatctgg aagtcacccg catcggaaca cgtgctcccg tcgtctttcc gcagcgcgtt 660
accgatcatc tgtgcgagat attgaaaaaa tatcatccgg tctggctgaa caccatctt 720
aacacaagca tcgaaatgac agaagaaccc gttgaggcat gtgaaaagct ggtgaacgag 780
ggagtgccgg tcggaaatca ggctgtcgta ttagcgggta ttaatgattc ggttccaatt 840
atgaaaaagc tcatgcatga cttggtaaaa atcagagtcc gtccttatta tattttacca 900
tgtgatctgt cagaaggaat aaggcatttc tgtgtctctg tttccaaagg ttggagatc 960
attgaagggc tgagagggtca tacctcaggc tatgcgggtc ctacctttgt cgttcacgca 1020
ccaggcggag gaggtaaaat cgccctgcag ccgaactatg tcctgtctca aagtctgac 1080
aaagtgatct taagaaattt tgaagggtgtg attacgtcat atccggagcc agagaattat 1140
atccccaatc aggagacgc ctatttttgag tccgttttcc ctgaaaccgc tgacaaaaag 1200
gagccgatcg ggctgagtgc cattttttgct gacaaagaag tttcgtctac acctgaaaat 1260
gtagacagaa tcaaacggcg tgaggcatac atcgcaaatc cggagcatga aacattaaaa 1320
gatcggcgtg agaaaagagg tcagctcaaa gaaaagaaat ttttggcgca gcagaaaaaa 1380
cagaaagaga ctgaatgcgg aggggattct tcataa 1416

```

<210> 22
 <211> 471
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 22

```

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
1           5           10           15

```

```

Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
          20           25           30

```

-35-

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

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

Pro Leu Asn Ile Thr Pro Tyr His Ala Ser Leu Met Asp Pro Asp Asn
 65 70 75 80

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
 85 90 95

His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu His Glu Asp Glu Asp
 100 105 110

Ser Pro Val Pro Gly Pro Thr His Arg Tyr Pro Asp Arg Val Leu Phe
 115 120 125

Leu Val Thr Asn Gln Cys Ser Val Tyr Cys Arg His Cys Thr Arg Arg
 130 135 140

Leu Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
 145 150 155 160

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
 165 170 175

Ile Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr
 180 185 190

Ile Leu Lys Glu Leu Arg Ser Ile Pro His Leu Glu Val Ile Arg Ile
 195 200 205

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

Cys Glu Ile Leu Lys Lys Tyr His Pro Val Trp Leu Asn Thr His Leu
 225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Pro Val Glu Ala Cys Glu Lys
 245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
 260 265 270

-36-

Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
 275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
 290 295 300

Glu Gly Ile Arg His Phe Cys Ala Pro Val Ser Lys Gly Leu Glu Ile
 305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
 325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
 340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
 385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Ser
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
 435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 23
 <211> 1416
 <212> DNA
 <213> Artificial Sequence

-37-

<220>

<223> Synthetic Construct

<400> 23

```

atgaaaaaca aatggtataa accgaaacgg cattggaagg agatcgagtt atggaaagac      60
gttccggacg aaaagtggaa cgattggctt tgacagctga cacacactgt aagaacgtta      120
gatgattcaa agaaagtcac taatctgacc gaggatgaag aggaaggcgt ccgtatttct      180
acaaaaacga tccccctaaa tattacacct tactatgctt ctttaatgga ccccgacaat      240
ccgagatgcc cggtagcat gcagtctgtg ccactttctg aagaaatgca caaaacaaaa      300
tacgatatgg aagaccgct tcatgaggat gaagattcac cggtagccgg tctgacacac      360
cgctatcccg gccgtgtgct gtttcttctc acgaatcaat gttccgtgca ctgccgccac      420
tgcacacgcc ggcgcttttc cggacaaatc ggaatgggag tccccgaaa acagcttgat      480
gctgcaattg cttatatccg ggaaacaccc gaaatccgag attgtttaat ttcaggcggg      540
gatgggctgc tcatcaacga ccaaatttta gaatatattt taaaagagct gcgcagcatt      600
ccgcatctgg aagtcacccg catcggaaca cgtgctcccg tcgtctttcc gcagcgcatt      660
accgatcatc tgtgagagat attgaaaaaa tatcatccgg tctggctgaa caccattttt      720
aacacaagca tcgaaatgac agaagaatcc gttgaggcat gtgaaaagct ggtgaacgcg      780
ggagtgccgg tcggaaatca ggctgtcgta ttagcaggta ttaatgatcc gggtccaatt      840
atgaaaaagc tcatgcatga cttggtaaaa atcagagctc gtccttatta tattttaccaa      900
tgtgatctgt cagaaggaat aaggcatttc cgtgctcctg tttccaaagg tttggagatc      960
attgaagggc tgagaggtca tacctcaggc tatgcggctc ctacctttgt cgttcacgca     1020
ccaggcggag gaggtaaaat cgccctgcag ccgaactatg tcctgtctca aagtcctgac     1080
aaagtgatct taagaaattt tgaagggtg attacgtcat atccggaaCc agagaattat     1140
atccccaatc aggacagcgc ctattttgag tccgttttcc ctgaaaccgc tgacaaaaag     1200
gagccgatcg ggctgagtgc cttttttgct ggcaaagaag tttcgtctac acctgaaaaat     1260
gtagacagaa tcaaacggcg tgaggcatac atcgcaaatc cggagcatga aacattaaaa     1320
gatcggcgtg agaaaagagg tcagctcaaa gaaaagaaat ttttggcgCa gcagaaaaaa     1380
cagaaagaga ctgaatgcgg aggggattct tcataa                                1416

```

<210> 24

<211> 471

<212> PRT

<213> Artificial Sequence

-38-

<220>

<223> Synthetic Construct

<400> 24

Met	Lys	Asn	Lys	Trp	Tyr	Lys	Pro	Lys	Arg	His	Trp	Lys	Glu	Ile	Glu
1				5					10					15	

Leu	Trp	Lys	Asp	Val	Pro	Asp	Glu	Lys	Trp	Asn	Asp	Trp	Leu	Trp	Gln
			20					25					30		

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

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

Pro	Leu	Asn	Ile	Thr	Pro	Tyr	Tyr	Ala	Ser	Leu	Met	Asp	Pro	Asp	Asn
65					70					75					80

Pro	Arg	Cys	Pro	Val	Arg	Met	Gln	Ser	Val	Pro	Leu	Ser	Glu	Glu	Met
			85						90					95	

His	Lys	Thr	Lys	Tyr	Asp	Met	Glu	Asp	Pro	Leu	His	Glu	Asp	Glu	Asp
			100					105					110		

Ser	Pro	Val	Pro	Gly	Leu	Thr	His	Arg	Tyr	Pro	Gly	Arg	Val	Leu	Phe
		115					120					125			

Leu	Val	Thr	Asn	Gln	Cys	Ser	Val	His	Cys	Arg	His	Cys	Thr	Arg	Arg
	130						135					140			

Arg	Phe	Ser	Gly	Gln	Ile	Gly	Met	Gly	Val	Pro	Glu	Lys	Gln	Leu	Asp
145					150					155					160

Ala	Ala	Ile	Ala	Tyr	Ile	Arg	Glu	Thr	Pro	Glu	Ile	Arg	Asp	Cys	Leu
				165					170					175	

Ile	Ser	Gly	Gly	Asp	Gly	Leu	Leu	Ile	Asn	Asp	Gln	Ile	Leu	Glu	Tyr
			180					185					190		

Ile	Leu	Lys	Glu	Leu	Arg	Ser	Ile	Pro	His	Leu	Glu	Val	Ile	Arg	Ile
		195					200					205			

-39-

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

Cys Glu Ile Leu Lys Lys Tyr His Pro Val Trp Leu Asn Thr His Phe
 225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
 245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
 260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
 275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
 290 295 300

Glu Gly Ile Arg His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
 305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
 325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
 340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
 385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Gly Lys Glu Val Ser Ser
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
 435 440 445

-40-

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 25
 <211> 1416
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 25
 atgaaaaaca aatggtataa accgaaacgg cattggaagg agatcgagtt atggaaggac 60
 gttccggaag agaaatggaa cgattggcct tgacagctga cacacactgt aagaacgcttg 120
 gatgatttaa agaaagtcac taacctgacc gaggatgaag aggaaggcgt ccgtatatttct 180
 accaaaacga tccccttaaa tattacacct tactatgctt ctttaatgga ccccgacaaa 240
 ccgagatgcc cggtagcat gcagtctgtg ccgctttctg aagaaatgca caaacaacaaa 300
 tacgatattg aagaccgct tcatgaggat gaagattcac cggtagccgg tctgacacac 360
 cgctatcccg accgtgtgct gtttcttctg acgaatcaat gttccgtgta ctgcccgcac 420
 tgcacacgcc ggcgcttttc cggacaaatc ggaatgggcg tccccaaaaa acagcttgat 480
 gctgcaattg cttatatccg ggaaacaccc gaaatccgag attgtttaat ttcaggcggg 540
 gatgggctgc tcatcaacga ccaaatttta gaatatattt taaaagagct gcgcagcatt 600
 ccgcatctgg aagtcacccg catcggaaca cgtgctcccg tcgtctttcc gcagcgcatt 660
 accgatcatc tgtgagagat attgaaaaaa tatcatccgg tctggctgaa caccattttt 720
 aacacaagca tcgaaatgac agaagaatcc gttgaggcat gtgaaaagct ggtgaacgcg 780
 ggagtgccgg tcggaaatca ggctgtcgta ttagcaggta ttaatgattc ggttccaatt 840
 atgaaaaagc tcatgcatga cttggtaaaa atcagagtcg gtccttatta tat ttaccaa 900
 tgtgacctgt cagaaggaat aaggcatttc cgtgctcctg tttccaaagg tttggggatc 960
 attgaagggc tgggagggtca tacctcaggc tatgcgggtc ctacctttgt cgttcacgca 1020
 ccaggcggag gaggtaaaat cgccctgcgg ccgaactatg tcctgtctca aagtcctgac 1080
 aaagtgatct taagaaattt tgaagggtg attacgtcat atccggaacc agagaattat 1140
 atccccaatc aggcagacgc ctattttgag tccgttttcc ctgaaaccgc tga caagaag 1200

-41-

gagccgatcg ggctgagtg cttttttgct gacaaagaag tttcgtctac acctgaaaaat 1260
 gtagacagaa tcaaacggcg tgaggcatac atcgcaaadc cggagcatga aacattaaaa 1320
 gatcggcgtg agaaaagagg tcagctcaaa gaaaagaaat ttttggcgca gcagaaaaaaa 1380
 cagaaagaga ctgaatgcgg aggggattct tcataa 1416

<210> 26
 <211> 471
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 26

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
 1 5 10 15

Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
 20 25 30

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

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

Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Ser Leu Met Asp Pro Asp Lys
 65 70 75 80

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
 85 90 95

His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu His Glu Asp Glu Asp
 100 105 110

Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
 115 120 125

Leu Val Thr Asn Gln Cys Ser Val Tyr Cys Arg His Cys Thr Arg Arg
 130 135 140

Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
 145 150 155 160

-42-

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
 165 170 175

Ile Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr
 180 185 190

Ile Leu Lys Glu Leu Arg Ser Ile Pro His Leu Glu Val Ile Arg Ile
 195 200 205

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

Cys Glu Ile Leu Lys Lys Tyr His Pro Val Trp Leu Asn Thr His Phe
 225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
 245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
 260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
 275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
 290 295 300

Glu Gly Ile Arg His Phe Arg Ala Pro Val Ser Lys Gly Leu Gly Ile
 305 310 315 320

Ile Glu Gly Leu Gly Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
 325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Arg Pro Asn
 340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

-43-

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Ser
405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
450 455 460

Glu Cys Gly Gly Asp Ser Ser
465 470

<210> 27

<211> 1416

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 27

```

atgaaaaaca aatggtataa accgaaacgg cattggaagg agatcgagtt atggaaggac      60
gttccgggag agaaatggaa cgattggctt tgacagctga cacacactgt aagaacgtta      120
gatgatttaa agaaagtcac taatctgacc gaggatgaag aggaaggcgt ccgtatttct      180
acaaaaacga tccccttaaa tattacacct tgctatgctc ctttaatgga ccccgacaac      240
ccgagatgcc cgttacgcat gcagtctgtg ccgctttctg aagaaatgca caaaacaaaa      300
tacgatatgg aagaccgct tcgtgaggat gaagattcac cgttaccgg tctgacacac      360
cgctatcccg accgtgtgct gtttcttgtc acgaatcaat gttccgtgta ctgccgccac      420
tgcacacgcc ggcgcttttc cggacaaatc ggaatgggcg tccccaaaaa acagcttgat      480
gctgcaattg cttatatccg ggaaacaccc .gaaatccgcg attgtttaat ttcaggcggg      540
gatgggctgc tcatcaacgg ccaaatttta gaatatatct taaaagagct gcgcagcatt      600
ccgcatctgg aagtcacccg catcgggaaca cgtgctcccg tcgtctttcc gcagcgcatt      660
accgatcatc tgtgcgagat attgaaaaaa tatcatccgg tctggctgaa caccattttt      720
aacacaagcg tcgaaatgac agaagaatcc gttgaggcat gtgaaaagct ggtgaacgcg      780

```

-44-

```

ggagtgccgg tcggaaatca ggctgtogta ttagcaggta ttaatgattc ggttccaatt      840
atgaaaaaagc tcatgcatga cttggtaaaa atcagagtcc gtccttatta tatttaccaa      900
tgtgatctgt cagaaggaat aaggcatttc cgtgctcctg tttccaaagg tttggagatc      960
attgaagggc tgagaggtca tacctcaggc tatgcggttc ctacctttgt cgttcacgca     1020
ccaggcggag ggggtaaaat cgccctgcag ccgaactatg tcctgtctca aagtccctgac     1080
aaagtaatct taagaaatct tgaagggtg attacgtcat atccggaacc agagaattat     1140
atccccaatc aggcagacgc ctattttgag tccgttttcc ctggaaccgc tgacaaaaag     1200
gagccgatcg ggctgagtgc catttttgct gacaaagaag tttcgtctac acctgaaaat     1260
gtagacagaa tcaaacggcg tgaggcatac atcgcaaadc cggagcatga aacattaaaa     1320
gatcggcggtg agaaaagagg tcagctcaaa gaaaagaaat ctttggcgca gcagaaaaaa     1380
cagaaagaga ctgaatgcgg aggggattct tcataa                                1416

```

<210> 28
 <211> 471
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 28

```

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
1           5           10           15

```

```

Leu Trp Lys Asp Val Pro Gly Glu Lys Trp Asn Asp Trp Leu Trp Gln
          20           25           30

```

```

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

```

```

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

```

```

Pro Leu Asn Ile Thr Pro Cys Tyr Ala Pro Leu Met Asp Pro Asp Asn
65           70           75           80

```

```

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
          85           90           95

```

-45-

His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu Arg Glu Asp Glu Asp
 100 105 110

Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
 115 120 125

Leu Val Thr Asn Gln Cys Ser Val Tyr Cys Arg His Cys Thr Arg Arg
 130 135 140

Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
 145 150 155 160

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
 165 170 175

Ile Ser Gly Gly Asp Gly Leu Leu Ile Asn Gly Gln Ile Leu Glu Tyr
 180 185 190

Ile Leu Lys Glu Leu Arg Ser Ile Pro His Leu Glu Val Ile Arg Ile
 195 200 205

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

Cys Glu Ile Leu Lys Lys Tyr His Pro Val Trp Leu Asn Thr His Phe
 225 230 235 240

Asn Thr Ser Val Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
 245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
 260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
 275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
 290 295 300

Glu Gly Ile Arg His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
 305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
 325 330 335

-46-

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
 340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Gly Thr Ala Asp Lys Lys
 385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Ser
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
 435 440 445

Leu Lys Glu Lys Lys Ser Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 29

<211> 1416

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 29

atgaaaaaca aatggtataa accgaaacgg cattggaagg agatcgagtt atggaaggac 60

gttccggaag agaaatggaa cgattggctt tgacggctga cacacactgt aagaacgtta 120

gatgatttaa agaaagtcac taatctgacc gaggatgaag aggaaggcgt ccgtatttct 180

acaaaaacga tccccttaag tattacacct tactatgctt ctttaatgga ccccgacaat 240

ccgagatgcc cggtagcat gcagtctgtg ccgctttctg aggaaatgca caaaacaaaa 300

tacgatatgg aagaccgct tcatgaggat gaagattcac cggtagccgg tctgacacac 360

-47-

```

cgctatcccg accgtgtgct gtttcttgct acgaatcaat gttccgtgta Ctgccgccgc 420
tgcacacgcc ggcgcttttc cggacagatc ggaatgggcg tccccaaaaa acagcttgat 480
gctgcaattg cttatatccg ggaaacaccc gaaatccgcg attgtttaat Ttcaggcggt 540
gatgggctgc tcatcaacga ccaaatttta gaatatattt taaaagagct Gcgcagcatt 600
ccgcatctgg aagtcatccg catcggaaca cgtgctcccg tcgtctttcc Gcagcgcatt 660
accgatcatc tgtgcgagat attgaaaaaa tatcatccgg tctggctgaa Caccattttt 720
aacacaagca tcgaaatgac agaagaatcc gttgaggcat gtgaaaagct Ggtgaacgcg 780
ggagtgccgg tcggaaatca ggctgtcgta ttagcaggta ttaatgattc Ggttccaatt 840
atgaaaaagc tcatgcatga cctggtaaaa atcagagtcc gtccttatta Tatttaccaa 900
tgtgatctgt cagaaggaat acggcatttc cgtgctcctg tttccaaagg Tttggagatc 960
attgaagggc tgagaggtca tacctcaggc tatgcggttc ctacctttgt Cgttcacgca 1020
ccaggcggag gaggtaaaat cgccctgcag ccgaactatg tcctgtctca Aagtccctgac 1080
aaagtgatct taagaaattt tgaagggtg attacgtcat atccggaacc Agagaattat 1140
atccccaatc aggcagacgc ctattttgag tccgttttcc ctgaaaccgc Tgacaaaaag 1200
gagccgatcg ggctgagtgc cttttttgct gacaaagaag tttcgtctac Acctgaaaat 1260
gtagacagaa tcaaacggcg tgaggcatat atcgcaaatc cggagcatga Aacattaaaa 1320
gatcggcggtg agaaaagagg tcagctcaaa gaaaagaaat ttttggcgca Gcagaaaaaa 1380
cagaaagaga ctgaatgcgg aggggattct tcataa 1416

```

<210> 30

<211> 471

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 30

```

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
1           5           10           15

```

```

Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Arg
          20           25           30

```

```

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

```

-48-

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

Pro Leu Ser Ile Thr Pro Tyr Tyr Ala Ser Leu Met Asp Pro Asp Asn
 65 70 75 80

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
 85 90 95

His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu His Glu Asp Glu Asp
 100 105 110

Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
 115 120 125

Leu Val Thr Asn Gln Cys Ser Val Tyr Cys Arg Arg Cys Thr Arg Arg
 130 135 140

Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
 145 150 155 160

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
 165 170 175

Ile Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr
 180 185 190

Ile Leu Lys Glu Leu Arg Ser Ile Pro His Leu Glu Val Ile Arg Ile
 195 200 205

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

Cys Glu Ile Leu Lys Lys Tyr His Pro Val Trp Leu Asn Thr His Phe
 225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
 245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
 260 265 270

-49-

Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
 275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
 290 295 300

Glu Gly Ile Arg His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
 305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
 325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
 340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
 385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Ser
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
 435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 31

<211> 1416

<212> DNA

<213> Artificial Sequence

<220>

-50-

<223> Synthetic Construct

<400> 31

```

atgaaaaaca aatggtataa accgaaacgg cattggaagg agatcgagtt atggaaggac      60
gttccggaag agaaatggaa cgattggctt tgacagctga cacgcactgt aagaacgttta      120
gatgatttaa agaaagtcac taatctgacc gaggatgaag aggaaggcgt ccgtatttct      180
acaaaaacga tccccttaaa tattacacct tactatgcga gcttaatgga tccagaaaac      240
ccacgttgtc cggtagcat gcagtctgtg ccgctttctg aagaaatgca cacaagcaaa      300
tatgacatgg aagatccgct tcatgaggat gaagattcac cggtagccgg tctgacacac      360
cgctatcccg accgtgtgct gtttcttgct acgagtcaat gtcccggtga ctgccgcac      420
tgcacacgcc ggcgcttttc cggacaaatc ggaatggcg tccccaaaa acagcttgat      480
gctgcaattg cttatatccg ggaaacaccc gaaatccgag attgtttaat ttcaggcgg      540
gatgggctgc tcatcaacga ccaaatttta gaatatattt taaaagagct gcgcagcatt      600
ccgcatctgg gagtcatccg catcggaaca cgtgctcccg tcgtctttcc gcagcgcat      660
accgatcatc tgtgcgagat attgaaaaga tatcatccgg tctggctgaa caccatttt      720
aacacaagca tcgaaatgac agaagaatcc gttgaggcat gtgaaaagct ggtgaacgcg      780
ggagtgcgag tcggaaatca ggctgtcgta ttagcaggta ttaatgattc ggttccaatt      840
atgaaaaagc tcatgcatga cttggtaaaa atcagagtcc gtccttatta tatttaccaa      900
tgtgatctgt cagaaggaat aaggcatttc cgtgctcctg tttccaaagg tttggagatc      960
attgaagggc tgagagggtca tacctcaggc tatgcggttc ctacctttgt cgttcacgca      1020
ccaggcggag gaggtaaaat cgcctgcag ccgaactatg tcctgtctca aagtcctgac      1080
aaagtgatct taagaaattt tgaagggtgtg attacgtcat atccggaacc agagaattat      1140
atccccaatc aggcagacgc ctattttgag tccgttttcc ctgaaaccgc tgacaaaag      1200
gagccgatcg ggctgagtgc ctttttgct gacaaagaag tttcgtctac acctgaaaat      1260
gtagacagaa tcaaacggcg tgaggcatat atcgcaaatc cggagcatga aacattaaa      1320
gatcggcggtg agaaaagagg tcagctcaaa gaaaagaaat ttttggcgca gcagaaaaa      1380
cagaaagaga ctgaatgcgg aggggattct tcataa      1416

```

<210> 32

<211> 471

<212> PRT

<213> Artificial Sequence

<220>

-51-

<223> Synthetic Construct

<400> 32

Met	Lys	Asn	Lys	Trp	Tyr	Lys	Pro	Lys	Arg	His	Trp	Lys	Glu	Ile	Glu
1				5					10					15	

Leu	Trp	Lys	Asp	Val	Pro	Glu	Glu	Lys	Trp	Asn	Asp	Trp	Leu	Trp	Gln
			20					25					30		

Leu	Thr	Arg	Thr	Val	Arg	Thr	Leu	Asp	Asp	Leu	Lys	Lys	Val	Ile	Asn
		35					40					45			

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

Pro	Leu	Asn	Ile	Thr	Pro	Tyr	Tyr	Ala	Ser	Leu	Met	Asp	Pro	Glu	Asn
65					70					75					80

Pro	Arg	Cys	Pro	Val	Arg	Met	Gln	Ser	Val	Pro	Leu	Ser	Glu	Glu	Met
				85					90					95	

His	Thr	Ser	Lys	Tyr	Asp	Met	Glu	Asp	Pro	Leu	His	Glu	Asp	Glu	Asp
			100					105					110		

Ser	Pro	Val	Pro	Gly	Leu	Thr	His	Arg	Tyr	Pro	Asp	Arg	Val	Leu	Phe
		115					120					125			

Leu	Val	Thr	Ser	Gln	Cys	Pro	Val	Tyr	Cys	Arg	His	Cys	Thr	Arg	Arg
	130					135					140				

Arg	Phe	Ser	Gly	Gln	Ile	Gly	Met	Gly	Val	Pro	Lys	Lys	Gln	Leu	Asp
145					150					155					160

Ala	Ala	Ile	Ala	Tyr	Ile	Arg	Glu	Thr	Pro	Glu	Ile	Arg	Asp	Cys	Leu
				165					170					175	

Ile	Ser	Gly	Gly	Asp	Gly	Leu	Leu	Ile	Asn	Asp	Gln	Ile	Leu	Glu	Tyr
			180					185					190		

Ile	Leu	Lys	Glu	Leu	Arg	Ser	Ile	Pro	His	Leu	Gly	Val	Ile	Arg	Ile
		195					200					205			

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

-52-

Cys Glu Ile Leu Lys Arg Tyr His Pro Val Trp Leu Asn Thr His Phe
 225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
 245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
 260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
 275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
 290 295 300

Glu Gly Ile Arg His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
 305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
 325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
 340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
 385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Ser
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
 435 440 445

-53-

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 33
 <211> 1416
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 33
 atgaaaaaca aatggtataa accgaaacgg cattggaagg agatcgagtt atggaaggac 60
 gttccggaag agaaatggaa cgattggcct tgacagctga cacacactgt aagaacgtta 120
 gatgatttaa agaaagtcac taatctgacc gaggatgaag aggaaaggcgt ccgtattttct 180
 accaaaacga tccccttaaa tattacacct tactatgctt ctttaatgga ccccgacaat 240
 ccgagatgcc cggtagcat gcagtctgtg ccgctttctg aagaaatgca caaaacaaaa 300
 tacgatatgg aagaccgct tcatgaggat gaagattcac cggtaaccgg tctgacacac 360
 cgctatcccg accgtgtgct gtttcttgtc acgaatcaat gttccgtgta ctgccgccac 420
 tgcacacgcc ggcgcttttc cggacaaatc ggaatgggcg tccccaaaaa acagcttgat 480
 gctgcaattg cttatatccg ggaaacaccc gaaatccgcg actgtctgtt gtctggcggt 540
 gatgggctgc tcatcaacga ccaaatttta gaatatattt taaaagagct gcgcagcatt 600
 ccgcatctgg aagtcacccg catcggaaca cgtgctcccg tcgtctttcc gcagcgcatt 660
 accgatcacc tgtgcgagat gttaaaaaaa tatcatccgg tctggctgaa caccattttt 720
 aacacaagca tcgaaatgac agaagaatcc gttgaggcat gtgaaagct ggtgaacgcg 780
 ggagtgccgg tcggaaatca ggctgtcgta ttagcaggta ttaatgattc ggttccaatt 840
 atgaaaaagc tcatgcatga cttggtaaaa atcagagtcc gtccctatta tatttaccac 900
 tgtgatctgt cagaaggaat aaggcatttc cgtgctcctg tttccaaagg tttggagatc 960
 attgaagggc tgagaggtca tacctcaggc tatgcgggtc ctacccttctg cgttcacgca 1020
 ccaggcggag gaggtaaaat cgcctgcag ccgaactatg tctgtctca aagtcctgac 1080
 aaagtgatct taagaaattt tgaagggtgtg attacgtcat atccggaacc agagaattat 1140
 atccccaatc aggcagacgc ctattttgag tccgttttcc ctgaaaccgc tgacaaaaag 1200
 gagccgatcg ggctgagtgc gctgtttgct gacaaagaag tttcgtctac acctgaaaat 1260

-54-

gtagacagaa tcaaacggcg tgaggcatac atcgcaaadc cggagcatga aacattaaaa 1320
gatcggcggtg agaaaagagg tcagctcaaa gaaaagaaat ttttggcgca gcagaaaaaa 1380
cagaaagaga ctgaatgcgg aggggattct tcataa 1416

<210> 34
<211> 471
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 34

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
1 5 10 15

Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
20 25 30

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

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

Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Ser Leu Met Asp Pro Asp Asn
65 70 75 80

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
85 90 95

His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu His Glu Asp Glu Asp
100 105 110

Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
115 120 125

Leu Val Thr Asn Gln Cys Ser Val Tyr Cys Arg His Cys Thr Arg Arg
130 135 140

Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
145 150 155 160

-55-

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
165 170 175

Leu Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr
180 185 190

Ile Leu Lys Glu Leu Arg Ser Ile Pro His Leu Glu Val Ile Arg Ile
195 200 205

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

Cys Glu Met Leu Lys Lys Tyr His Pro Val Trp Leu Asn Thr His Phe
225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
290 295 300

Glu Gly Ile Arg His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
385 390 395 400

-56-

Glu Pro Ile Gly Leu Ser Ala Leu Phe Ala Asp Lys Glu Val Ser Ser
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
 435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 35

<211> 1416

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 35

```

atgaaaaaca aatggtataa accgaaacgg cattggaagg agatcgagtt atggaaggac      60
gttccggaag agaaatggaa cgattggctt tgacagctga cacacactgt aagaacgtta      120
gatgatttaa agaaagtcac taatctgacc gaggatgaag aggaaggcgt ccgtatttct      180
acaaaaacga tccccttaaa tatcacacct tactatgcga gcttaatgga tccagaaaac      240
ccacgttgtc cggtagcat gcagtctgtg ccgcttcctg aagaaatgca caaaacaaaa      300
tacgatatgg aagaccgct tcatgaggat gaagattcac cggtagccgg tctgacacac      360
cgctatcccg accgtgtgct gtttcttgct acggatcaat gttccgtgta ctgccgccac      420
cgcacacgcc ggcgcttctc cggacaaatc ggaatgggcg tccccgaaaa acagcttgat      480
gctgcaattg cttacatccg ggaaacaccc gaaatccgcg attgtttaat ttcaggcgggt      540
gatgggctgc tcatcaacga ccaaatttta gaatatattt taaaagagct gcgcagcatt      600
ccgcatctgg aagtcacccg catcggaaca cgtgctcccg tcgtctttcc gcagcgcatt      660
accgatcatc tgtgcgagat attgaaaaaa catcatcccg tctggctgaa caccattttt      720
aacacaagca tcgaaatgac agaagaatcc gttgaggcat atgaaaagct ggtgaacgcg      780
ggagtgccgg tcggaaatca ggctgttgta ttagcaggta ttaatgattc ggttccaatt      840

```

-57-

```

ataaaaaagc tcatgcatga cttggtaaaa atcagagtcc gtccttatta tatttaccaa    900
tgtgacctgt cagaaggaat aaggcatttc cgtgctcctg ttcccaaagg tttggagatc    960
attgaagggc tgagaggtca tacctcaggc tatgcgggtc ctacctttgt cgttcacgca   1020
ccaggcggag gaggtaaaat cgcctgcag ccgaactatg tcctgtctca aagtccctgac   1080
aaagtgatct taagaaattt tgaagggtg attacgtcat atccggaacc agagaattat   1140
atccccaatc aggcagacgc ctattttgag tccgttttcc ctgaaaccgc tgacaaaaag   1200
gagccgatcg ggctgagtgc catttttgct gacaaagaag ttctgtctac acctgaaaat   1260
gtagacagaa tcaaacggcg tgaggcatac atcgcaaadc cggagcatga aacattaaaa   1320
gatcggcgtg agaaaagagg tcagctcaaa gaaaagaaat ttttggcgca gcagaaaaaa   1380
cagaaagaga ctgaatgcgg aggggattct tcataa                                1416

```

<210> 36
 <211> 471
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 36

```

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
1          5          10          15

```

```

Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
          20          25          30

```

```

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

```

```

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

```

```

Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Ser Leu Met Asp Pro Glu Asn
65          70          75          80

```

```

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Pro Glu Glu Met
          85          90          95

```

```

His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu His Glu Asp Glu Asp
          100          105          110

```

-58-

Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
115 120 125

Leu Val Thr Asp Gln Cys Ser Val Tyr Cys Arg His Arg Thr Arg Arg
130 135 140

Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Glu Lys Gln Leu Asp
145 150 155 160

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
165 170 175

Ile Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr
180 185 190

Ile Leu Lys Glu Leu Arg Ser Ile Pro His Leu Glu Val Ile Arg Ile
195 200 205

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

Cys Glu Ile Leu Lys Lys His His Pro Val Trp Leu Asn Thr His Phe
225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Tyr Glu Lys
245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Ile Lys Lys Leu Met His Asp Leu
275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
290 295 300

Glu Gly Ile Arg His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
325 330 335

-59-

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Ser
405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
450 455 460

Glu Cys Gly Gly Asp Ser Ser
465 470

<210> 37

<211> 1416

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 37

atgaaaaaca aatggtataa accgaaacgg cattggaagg agatcgagtt atggaaggac	60
gttcggaag agaaatggaa cgattggctt tgacagctga cacacactgt aagaaCgtta	120
gatgatttaa agaaagtcac taatctgacc gaggatgaag aggaaggcgt ccgtatttct	180
acaaaaacga tccccttaaa tattacacct tactatgctt cttaaatgga ccccgacaat	240
ccgagatgcc cggtagcat gcagtctgtg ccgctttctg aagaaatgca caaaaCaaaa	300
tacgatatgg aagacccgct tcatgaggat gaagattcac cggtagccgg totgaCacac	360
cgctatccca accgtgtgct gtttcttctc acgaatcaat gttccgtgta ctgccgccac	420

-60-

```

tgcacacgcc ggcgcttttc cggacaaatc ggaatgggag tccccaaaaa acagcttgat 480
gctgcaattg cttatatccg ggaaacaccc gaaatccgag actgtctgtt gtctggcggt 540
gatgggctgc tcatcaacga ccaaatttta gaatatatct taaaagagct gcgcagcatt 600
ccgcactctg aagtcatttc tatcggttct cgtgcgccag tcgtctttcc gcagcgcat 660
accgatcatc tgtgcgagat attgaaaaaa tatcatccgg tctggctgaa caccattttt 720
aacacaagca tcgaaatgac agaagaatcc gttgaggcat gtgaaaagct ggtgaacgcg 780
ggagtgcgag tcggaaatca ggctgtcgta ttagcaggta ttaatgatct ggttccaatt 840
atgaaaaagc tcatgcatga cttggtaaaa atcagagtcc gtccttatta tattttacaa 900
tgtgatctgt cagaaggaat agggcatttc cgtgctcctg tttccaaagg tttggagatc 960
attgaagggc tgagagggtca tacctcaggc tatgcggttc ctacctttgt cgttcacgca 1020
ccaggcggag gaggtaaaaat cggcctgcag ccgaactatg tcctgtcaca aagtcctgac 1080
aaagtgatct taagaaattt tgaagggtgtg attacgtcat atccggaacc agagaattat 1140
atccccaatc aggcagacgc ctattttgag tccgttttcc ctgaaacgc tgacaaaaag 1200
gagccgatcg ggctgagtgc cttttttgct gacaaagaag tttcgtttac acctgaaaat 1260
gtagacagaa tcaaacggcg tgaggcatac atcgcaaatc cggagcatga aacattaaaa 1320
gatcggcggt agaaaagaga tcagctcaaa gaaaagaaat ttttggcgca gcagaaaaaa 1380
cagaaagaga ctgaatgcgg aggggattct tcataa 1416

```

<210> 38
 <211> 471
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 38

```

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
.1          5          10          15

```

```

Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
          20          25          30

```

```

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

```

-61-

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

Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Ser Leu Met Asp Pro Asp Asn
 65 70 75 80

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
 85 90 95

His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu His Glu Asp Glu Asp
 100 105 110

Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asn Arg Val Leu Phe
 115 120 125

Leu Val Thr Asn Gln Cys Ser Val Tyr Cys Arg His Cys Thr Arg Arg
 130 135 140

Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
 145 150 155 160

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
 165 170 175

Leu Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr
 180 185 190

Ile Leu Lys Glu Leu Arg Ser Ile Pro His Leu Glu Val Ile Arg Ile
 195 200 205

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

Cys Glu Ile Leu Lys Lys Tyr His Pro Val Trp Leu Asn Thr His Phe
 225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
 245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
 260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
 275 280 285

-62-

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
 290 295 300

Glu Gly Ile Gly His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
 305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
 325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
 340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
 385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Phe
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Asp Gln
 435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 39

<211> 1416

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

-63-

<400> 39

```

atgaaaaaca aatggtataa accgaaacgg cattggaagg agatcgagtt atggaaggac      60
gttccggaag agaaatggaa cgattggctt tgacagctga cacacactgt aagaacgtta      120
gatgatttaa agaaagtcac taatctgacc gaggatgaag aggaaggcgt ccgtatttct      180
acaaaaacga tccccttaaa tattacacct tactatgctt ctttaatgga ccccgacaat      240
ccgagatgcc cggtagcat gcagtctgtg ccgctttctg aagaaatgca caaaacaaaa      300
tacgatatgg aagaccgct tcatgaggat gaagattcac cggtagccgg tctgacacac      360
cgctatcccg accgtgtgct gtttctgtc acgaatcaat gttccgtgca ctgccgccac      420
tgacacgcc gccgcttttc cggacaaatc ggaatggcg tccccaaaa acagcttgat      480
gctgcaattg cttatatccg ggaaacaccc gaaatccgct attgtttaat ttcaggcgg      540
gatgggctgc tcatcaacga ccaaatttta gaatatattt taaaagagct gcgcagcatt      600
ccgcacctgg aagtcacccg catcggaaca cgtgctcccg tcgtctttcc gcagcgcatt      660
accgatcatc tgtgcgagat attgaaaaaa tatcatccgg tctggctgaa caccatttt      720
aacacaagca tcgaaatgac agaagaatcc gttgaggcat gtgaaaagct ggtgaacgcg      780
ggagtgccgg tcggaaatca ggctgtcgta ttagcaggta ttaatgattc ggttccaatt      840
atgaaaaagc tcatgcatga cttggtaaaa atcagagtcc gtccttatta tatttaccac      900
tgtgatctgt cagaaggaat aaggcatttc cgtgctcctg tttccaaagg tttggagatc      960
attgaagggc tgagagggtca tacctcaggc tatgcggttc ctacctttgt cgttcacgca      1020
ccaggcggag gtggtaaaaat cgcctgcag ccgaactatg tcctgtctca aagtcctgac      1080
aaagtgatct taagaaattt tgaagggtgtg attacgtcat atccggaacc agagaattat      1140
atccccaatc aggcagacgc ctattttgag tccgttttcc ctgaaaccgc tgacaaaaag      1200
gagccgatcg ggctgagtgc catttttgct ggcaaagaag tttcgtctac acctgaaaat      1260
gtagtcagaa tcaaacggcg tgaggcatac atcgcaaatc cggagcatga aacattaaaa      1320
gatcggcgtg agaaaagagg tcagctcaaa gaaaagaaat ttttggcgca gcagaaaaaa      1380
cagaaagaga ctgaatgcgg aggggattct tcataa                                1416

```

<210> 40

<211> 471

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

-64-

<400> 40

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
1 5 10 15

Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
20 25 30

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

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

Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Ser Leu Met Asp Pro Asp Asn
65 70 75 80

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
85 90 95

His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu His Glu Asp Glu Asp
100 105 110

Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
115 120 125

Leu Val Thr Asn Gln Cys Ser Val His Cys Arg His Cys Thr Arg Arg
130 135 140

Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
145 150 155 160

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
165 170 175

Ile Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr
180 185 190

Ile Leu Lys Glu Leu Arg Ser Ile Pro His Leu Glu Val Ile Arg Ile
195 200 205

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

-65-

Cys Glu Ile Leu Lys Lys Tyr His Pro Val Trp Leu Asn Thr His Phe
 225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
 245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
 260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
 275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
 290 295 300

Glu Gly Ile Arg His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
 305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
 325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
 340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
 385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Gly Lys Glu Val Ser Ser
 405 410 415

Thr Pro Glu Asn Val Val Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
 435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

-66-

Glu Cys Gly Gly Asp Ser Ser
465 470

<210> 41
<211> 1416
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 41
atgaaaaaca aatggtataa accgaaacgg cattggaagg agatcgagtt atggagggac 60
gtcccggaag agaaatggaa cgattggcctt tgacagctga cacacactgt aagaacgtta 120
gatgatttaa agaaagtcac taatctgacc gaggatgraag aggaaggcgt ccgtatttct 180
accaaaacga tccccttaaa tattacacct tactatgctt ctttaatgga ccccgacaat 240
ccgaggtgcc cggtagcat gcagtctgtg cactgtctg aggaatgca caaaagcaaa 300
tatgacatgg aagatccgct tcatgaggat gaagattcac cggtagccgg tctgacacac 360
cgctatcccg accgtgtgct gtttcttgtc acgaatcaat gttccgtgta ctgccgccac 420
tgacacgcc ggcgcttttc cggacaaatc ggaatgggag tccccaaaaa acagcttgat 480
gctgcaattg cttatatccg ggaaacaccc gaaatccgag attgtttaat ttcaggcggg 540
gatgggctgc tcatcaacga ccaaatttta gaatatattt taaaagagct gcgcagcatt 600
ccgcatctgg aagtcacccg catcggaaca cgtgctcccg tctcttttcc gcagcgcatt 660
accgatcatc tgtgcgagat attgaaaaaa tatcatcccg tctggctgaa caccattttt 720
aacacaagca tcgaaatgac agaagaatcc gttgagcgtc gtgaaaagct ggtgaacgcg 780
ggagtgccgg tcggaaatca ggctgtcgta ttagcaggta ttaatgattc ggttccaatt 840
atgaaaaagc tcatgcatga cttggtaaaa atcagagctc gtccttatta tatttaccaa 900
tgtgatctgt cagaaggaat aaggcatttc cgtgctcctg ttcccaaagg tttggagatc 960
attgaagggc tgagaggtca tacctcaggc tatgcgggtc ctacctttgt cgttcacgca 1020
ccgggaggag gaggtaaaat cgccctgcag ccgaac tatg tcctgtctca aagtcctgac 1080
aaagtgatct taagaaattt tgaagggtgtg attacgtcat atccggaacc agagaattat 1140
atccccaatc aggcagacgc ctattttgag tccgttttcc ctgaaaccgc tgacaaaaag 1200
gagccgatcg ggctgagtg cttttttgct gacaaaagag tttcgtctac acctgaaaat 1260
gtagacagaa tcaaacggcg tgaggcgtag atcgcaaatc cggagcatga aacattaaaa 1320

-67-

gatcggcgctg agaaaagagg tcagctcaaa gaaaagaaat ttttggcgca gcagaaaaaa 1380
 cagaaagaga ctgaatgcgg aggggattct tcataa 1416

<210> 42
 <211> 471
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 42

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
 1 5 10 15

Leu Trp Arg Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
 20 25 30

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

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

Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Ser Leu Met Asp Pro Asp Asn
 65 70 75 80

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
 85 90 95

His Lys Ser Lys Tyr Asp Met Glu Asp Pro Leu His Glu Asp Glu Asp
 100 105 110

Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
 115 120 125

Leu Val Thr Asn Gln Cys Ser Val Tyr Cys Arg His Cys Thr Arg Arg
 130 135 140

Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
 145 150 155 160

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
 165 170 175

-68-

Ile Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr
 180 185 190

Ile Leu Lys Glu Leu Arg Ser Ile Pro His Leu Glu Val Ile Arg Ile
 195 200 205

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

Cys Glu Ile Leu Lys Lys Tyr His Pro Val Trp Leu Asn Thr His Phe
 225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
 245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
 260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
 275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
 290 295 300

Glu Gly Ile Arg His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
 305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
 325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
 340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
 385 390 395 400

-69-

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Ser
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
 435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 43
 <211> 1416
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 43
 atgaaaaaca aatggtataa accgaaacgg cattggaagg agatcgagtt atggaaggac 60
 gttccggaag agaaatggaa cgattggcctt tgacagctga cacacactgt aagaacgtta 120
 gatgatttaa agaaagtcac taatctgacc gaggatgaag aggaaggcgt ccgtattttct 180
 accaaaacga tccccttaaa tattacacca tactatgcga gcttaatgga tccagaaaaac 240
 ccacgttgctc cggtacgcat gcagtcctgtg ccgctttccg aagaaatgca caaaacaaaa 300
 tacgatatgg aagaccgct tcatgaggat gaagattcac cggtaaccgg tctgacacac 360
 cgctatcccg accgtgtgct gtttcttgct acgaatcaat gttccgtgta ctgccgccac 420
 tgcacacgcc ggcgcttttc cggacaaatc ggaatgggcg tccccaaaaa acagcttgat 480
 gctgcaattg cttatatccg ggaaacaccc gaaatccgag attgtttaat ttcaggcgggt 540
 gatgggctgc tcatcaacga ccaaatttta gaatatattt taaaagagct gcgcagcatt 600
 ccgcacatcg aagtcacccg catcggaaca cgtgctcccg tcgtctttcc gcagcgcatt 660
 accgatcatc cgtgcgagat attgaaaaaa tatcatccgg tctggctgaa caccattttt 720
 aacacaagca tcgaaatgac agaagaatcc gttgaggcat gtgaaaagct ggtgaacgcg 780
 ggagtgcggg tcggaaatca ggctgtcgta ttagcaggta ttaatgattc ggttccaatt 840
 atgaaaaagc tcatgcatga cttggtaaaa atcagagtcc gtccttatta tattttacaa 900

-70-

```

tgtgatctgt cagaaggaat aaggcatttc cgtgctcctg tctccaaagg tttggagatc   960
attgaagggc tgagaggtca taccccaggc tatgcggttc ctacctttgt cgttcacgca  1020
ccaggcggag gaggtaaaat cgccttcag ccgaactatg tcctgtctca aagtcctgac   1080
aaagtgatct taagaaattt tgaaggtgtg attacgtcat atccggaac agagaattat   1140
atccccaatc aggcagacgc ctattttgag tccgtttccc ctgaaaccg tgacaaaaag   1200
gagccgatcg ggctgagtgc catttttgct gacaaagaag tttcgtcta acctgaaaat   1260
gtagacagaa tcaaacggcg tgaggcctac atcgcaaadc cggagcatga aacattaaaa   1320
gatcggcgtg agaaaagagg tcagctcaaa gaaaagaaat tttcggcgca gcagaaaaaa   1380
cagaaagaga ctgaatgcgg aggggattct tcataa                               1416

```

<210> 44
 <211> 471
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 44

```

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
1           5           10           15

```

```

Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
          20           25           30

```

```

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

```

```

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

```

```

Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Ser Leu Met Asp Pro Glu Asn
65           70           75           80

```

```

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
          85           90           95

```

```

His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu His Glu Asp Glu Asp
100           105           110

```

-71-

Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
 115 120 125

Leu Val Thr Asn Gln Cys Ser Val Tyr Cys Arg His Cys Thr Arg Arg
 130 135 140

Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
 145 150 155 160

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
 165 170 175

Ile Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr
 180 185 190

Ile Leu Lys Glu Leu Arg Ser Ile Pro His Leu Glu Val Ile Arg Ile
 195 200 205

Gly Thr Arg Ala Pro Val Val Phe Pro Gln Arg Ile Thr Asp His Pro
 210 215 220

Cys Glu Ile Leu Lys Lys Tyr His Pro Val Trp Leu Asn Thr His Phe
 225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
 245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
 260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
 275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
 290 295 300

Glu Gly Ile Arg His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
 305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Pro Gly Tyr Ala Val Pro Thr Phe
 325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
 340 345 350

-72-

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Ser Pro Glu Thr Ala Asp Lys Lys
 385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Ser
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
 435 440 445

Leu Lys Glu Lys Lys Phe Ser Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 45

<211> 1416

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 45

```

atgaaaaaca aatggtataa accgaaacgg cattggaagg agatcgagtt acggaaggac      60
gttccggaag agaaatggaa cgattggctt tgacagctga cgcacactgt aagaacgtta      120
gatgatttaa agaaagtcac taatctgacc gaggatgaag aggaaggcgt Ccgtatttct      180
accaaaacga tccccttaaa tattacacct tactatgcga gcttaattga tccagaaaac      240
ccacgttgct cggtacgcat gcagtctgct cgcgtgtctg aagaaatgca Caaaacaaaa      300
tacgatatgg aagaccgct tcatgaggat gaagattcac cggtaccggg tctgacacac      360
cgctatcccg accgtgtgct gtttcttgct acgaatcaat gttccgtgta Ctgccgccac      420
tgcacacgcc ggcgcttttc cggacaaatc ggaacgggctg tccccaaaaa acagcttgat      480

```

-73-

```

gctgcaactg cttatatccg ggaaacaccc gaaatccgcg attgtttaat tccaggcggt      540
gatgggctgc tcatcaacga ccaaatttta ggatatattt taaaagagct gcgcagcatt      600
ccgcatctgg aagtcacccg catcggaaca cgtgcccccg tcggctttcc gcagcgcat      660
accgatcatc tgtgcgagat attgaaaaaa tatcatccgg tctggctgaa caccatttt      720
aacacaagca tcgaaatgac agaagaatcc gttgaggcat gtgaaaagct ggtgaacgcg      780
ggagtgcggg tcggaaatca ggctgtcgta ttagcaggta ttaatgattc ggttccaatt      840
atgaaaaagc tcatgcatga cttggtaaaa atcagagtcc gtccttatta tatttacCaa      900
tgtgatctgt cagaaggaat aaggcatttc cgtgctcctg tttccaaagg tttggagatc      960
attgaagggc tgagagggtca tacctcaggc tatgcgggtc ctacctttgt cgttcacgca     1020
ccaggcggag gaggtaaaaat cgccctgcag ccgaactatg ccctgtctca aagtcctgac     1080
aaagtgatct taagaaattht tgaagggtgtg attacgtcat atccggaacc agagaattat     1140
atccccaatc aggcagacgc ctatthttgag tccgtthttcc ctgaaaccgc tgacaaaag      1200
gagccgatcg ggctgagtgc catthtttgc t gacaaagaag tttcgtctac acctgaaaat     1260
gtagacagaa tcaaacggcg tgaggcatatc atcgcaaatc cggagcatga aacattaaaa     1320
gatcggcgtg agaaaagagg tcagctcaaa gaaaagaaat ttttggcgca gcagaaaaaa     1380
cagaaagaga ctgaatgcgg aggggattct tcataa                                1416

```

<210> 46
 <211> 471
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 46

```

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
1           5           10           15

```

```

Leu Arg Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
          20           25           30

```

```

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

```

```

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

```

-74-

Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Ser Leu Ile Asp Pro Glu Asn
65 70 75 80

Pro Arg Cys Pro Val Arg Met Gln Ser Ala Pro Leu Ser Glu Glu Met
85 90 95

His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu His Glu Asp Glu Asp
100 105 110

Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
115 120 125

Leu Val Thr Asn Gln Cys Ser Val Tyr Cys Arg His Cys Thr Arg Arg
130 135 140

Arg Phe Ser Gly Gln Ile Gly Thr Gly Val Pro Lys Lys Gln Leu Asp
145 150 155 160

Ala Ala Thr Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
165 170 175

Ile Pro Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Gly Tyr
180 185 190

Ile Leu Lys Glu Leu Arg Ser Ile Pro His Leu Glu Val Ile Arg Ile
195 200 205

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

Cys Glu Ile Leu Lys Lys Tyr His Pro Val Trp Leu Asn Thr His Phe
225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
275 280 285

-75-

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
 290 295 300

Glu Gly Ile Arg His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
 305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
 325 330 335

Val Val His Ala Pro Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
 340 345 350

Tyr Ala Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
 385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Ser
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
 435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 47
 <211> 1416
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 47
 atggaaaaca aatggtataa accgaaacgg cattggaagg agatcgagtt atggaaggac 60

-76-

gttccggaag agaaatggaa cgattggctt tgacagctga cacacactgt aagaacgtta 120
gatgatTTaa agaaagtcac taatctgacc gaggatgaag aggaaggcgt ccgtatttct 180
acccaaacga tccccttaaa tattacacct tactatgcga gcttaattga tccagaaaac 240
ccacgttgct cggtacgcat gcagtctgtg ccgctttccg aagaaatgca caaaacaaaa 300
tacgatatgg aagatccgct tcatgaggat gaagattcac cggtacccgg cctgacacac 360
cgctatcccg accgtgtgct gtttcttgct gcgaatcaat gttccgtgta ctgccgccac 420
tgcacacgcc ggcgcttttc cggacaaatc ggaatgggcg tccccaaaa acagcttgat 480
gctgcaattg cttatatccg ggaaacaccc gaaatccgcg attgtttaat ttcaggcggt 540
gatgggctgc tcatcaacga ccaaatttta gaatatattt taaaagagct gcgcagcatt 600
ccgcatccgg aagtcacccg catcggaaca cgtgcccccg tcgtctttcc gcagcgcatt 660
accgatcatc tgtgcgagat attgaaaaaa tatcatccgg tctggctgaa caccattttt 720
aacacaagca tcgaaatgac agaagaatcc gttgaggcat gtgaaaagct ggtgaacgcg 780
ggagtgccgg tcggaaatca ggctgtcgta ttagcaggta ttaatgattc ggttccaatt 840
atgaaaaagc tcatgcatga cttggtaaaa atcagagtcc gtcttatta tatttaccaa 900
tgtgatctgt cagaaggaat aaggcatttc cgtgcccctg tttccaaagg tttggagatc 960
attgaagggc tgagagggtca tacctcaggc tgtgcggttc ctacctttgt cgttcacgca 1020
ccaggcggag gaggtaaaat cgcctgcag ccgaactatg tctgtctca aagtcctgac 1080
aaagtgatct taagaaattt tgaagggtgtg attacgtcat atccggaacc agagaattat 1140
atccccaacc aggcagacgc ctattttgag tccgttttcc ctgaaaccgc tgacaaaaag 1200
gagccgatcg ggctgagtgc ctttttgct gacaaagaag tttcgtctac acctgaaaat 1260
gtagacagaa tcaaacggcg tgaggcatat atcgcaaatc cggagcatga aacattaaaa 1320
gatcggcgtg agaaaagggg tcagctcaaa gaaaagaaat ttttggcgca gcagaaaaaa 1380
cagaaagaga ctgaatgcgg aggggattct tcataa 1416

<210> 48

<211> 471

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 48

-77-

Met	Glu	Asn	Lys	Trp	Tyr	Lys	Pro	Lys	Arg	His	Trp	Lys	Glu	Ile	Glu	
1				5					10					15		
Leu	Trp	Lys	Asp	Val	Pro	Glu	Glu	Lys	Trp	Asn	Asp	Trp	Leu	Trp	Gln	
			20					25					30			
Leu	Thr	His	Thr	Val	Arg	Thr	Leu	Asp	Asp	Leu	Lys	Lys	Val	Ile	Asn	
		35					40					45				
Leu	Thr	Glu	Asp	Glu	Glu	Glu	Gly	Val	Arg	Ile	Ser	Thr	Lys	Thr	Ile	
	50					55					60					
Pro	Leu	Asn	Ile	Thr	Pro	Tyr	Tyr	Ala	Ser	Leu	Ile	Asp	Pro	Glu	Asn	
65					70					75					80	
Pro	Arg	Cys	Pro	Val	Arg	Met	Gln	Ser	Val	Pro	Leu	Ser	Glu	Glu	Met	
				85					90						95	
His	Lys	Thr	Lys	Tyr	Asp	Met	Glu	Asp	Pro	Leu	His	Glu	Asp	Glu	Asp	
			100					105						110		
Ser	Pro	Val	Pro	Gly	Leu	Thr	His	Arg	Tyr	Pro	Asp	Arg	Val	Leu	Phe	
		115					120					125				
Leu	Val	Ala	Asn	Gln	Cys	Ser	Val	Tyr	Cys	Arg	His	Cys	Thr	Arg	Arg	
	130					135					140					
Arg	Phe	Ser	Gly	Gln	Ile	Gly	Met	Gly	Val	Pro	Lys	Lys	Gln	Leu	Asp	
145					150					155					160	
Ala	Ala	Ile	Ala	Tyr	Ile	Arg	Glu	Thr	Pro	Glu	Ile	Arg	Asp	Cys	Leu	
				165					170					175		
Ile	Ser	Gly	Gly	Asp	Gly	Leu	Leu	Ile	Asn	Asp	Gln	Ile	Leu	Glu	Tyr	
			180					185					190			
Ile	Leu	Lys	Glu	Leu	Arg	Ser	Ile	Pro	His	Pro	Glu	Val	Ile	Arg	Ile	
		195					200					205				
Gly	Thr	Arg	Ala	Pro	Val	Val	Phe	Pro	Gln	Arg	Ile	Thr	Asp	His	Leu	
	210					215						220				
Cys	Glu	Ile	Leu	Lys	Lys	Tyr	His	Pro	Val	Trp	Leu	Asn	Thr	His	Phe	
225					230					235					240	

-78-

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
 245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
 260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
 275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
 290 295 300

Glu Gly Ile Arg His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
 305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Cys Ala Val Pro Thr Phe
 325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
 340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
 385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Ser
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
 435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

-79-

Glu Cys Gly Gly Asp Ser Ser
465 470

<210> 49

<211> 1416

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 49

```

atgaaaaaca aatggtataa accgaaacgg cattggaagg agatcgagtt atggaaggac      60
gttccggaag agaaatggaa cgattggctt tgacagctga cacacactgt aagaacgtta      120
gatgatttaa agaaagtcac taatctgacc gaggatgaag aggaaggcgt ccgtatttct      180
acaaaaacga tccccttaaa tattacacct tactaggttt ctttaatgga ccccgacaat      240
ccgagatgcc cggtagcat gcagtctgtg ccactgtctg aagaaatgca caaaacaaaa      300
tacgatatgg aagaccgct tcatgaggat gaagattcac cggtagccgg tctgacacac      360
cgctatcccg accgtgtgct gtttcttgtc acgaatcaat gttccgtgta ctgccgccac      420
tgcacacgcc ggcgcttttc cggacaaatc ggaatgggcg tccccaaaaa acagcttgat      480
gctgcaattg cttatatccg ggaaacaccc gaaatccgcg attgtttaat ttcaggcggc      540
gatgggctgc tcatcaacga ccaaatttta gaatatattt taaaagagct gcgcagcatt      600
ccgcacatcg aagtcacccg catcggaaca cgtgctcccg tcgtctttcc gcagcgcatt      660
accgatcatc tgtgcgagat attgaaaaaa tatcatcccg tctggctgaa caccattttt      720
aacacaagca tcgaaatgac agaagaatcc gttgaggcat gtgaaaagct ggtgaacgcg      780
ggagtgcccg tcggaaatca ggctgtcgta ttagcaggta ttaatgattc ggttccaatt      840
atgaaaaagc tcatgcatga cttggtaaaa atcagagtc cgtccttatta tatttaccac      900
tgtgatctgt cagaaggaat aaggcatttc cgtgctcctg tttccaaagg tttggagatc      960
attgaagggc tgagagggtc cacctcaggc aatgcggttc ccacctttgt cgttcacgca     1020
ccaggcggag gaggtaaaaat cgccctgcag ccgaactatg tcctgtctca aagtcctgac     1080
aaagtgatct taagaaatct tgaagggtgtg attacgtcat atccggaacc agagaattat     1140
atccccaatc aggcagacgc ctattttgag tccgttttcc ctgaaaccgc tgacaaaaag     1200
gagccgatcg ggctgagtgc catttttgct gacaaagaag tttcgtctac acctgaaaat     1260
gtagacagaa tcaaaccggc tgaggcatac atcgcaaatc cggagcatga aacattaaaa     1320
gatcggcgtg agaaaagagg tcagctcaaa gaaaagaaat ttttggcgca gcagaaaaaa     1380

```

-80-

cagaaagaga ctgaatgcgg aggggattct tcataa

1416

<210> 50
 <211> 71
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 50

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
 1 5 10 15

Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
 20 25 30

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

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

Pro Leu Asn Ile Thr Pro Tyr
 65 70

<210> 51
 <211> 399
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 51

Val Ser Leu Met Asp Pro Asp Asn Pro Arg Cys Pro Val Arg Met Gln
 1 5 10 15

Ser Val Pro Leu Ser Glu Glu Met His Lys Thr Lys Tyr Asp Met Glu
 20 25 30

Asp Pro Leu His Glu Asp Glu Asp Ser Pro Val Pro Gly Leu Thr His
 35 40 45

Arg Tyr Pro Asp Arg Val Leu Phe Leu Val Thr Asn Gln Cys Ser Val
 50 55 60

-81-

Tyr Cys Arg His Cys Thr Arg Arg Arg Phe Ser Gly Gln Ile Gly Met
65 70 75 80

Gly Val Pro Lys Lys Gln Leu Asp Ala Ala Ile Ala Tyr Ile Arg Glu
85 90 95

Thr Pro Glu Ile Arg Asp Cys Leu Ile Ser Gly Gly Asp Gly Leu Leu
100 105 110

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

Pro His Leu Glu Val Ile Arg Ile Gly Thr Arg Ala Pro Val Val Phe
130 135 140

Pro Gln Arg Ile Thr Asp His Leu Cys Glu Ile Leu Lys Lys Tyr His
145 150 155 160

Pro Val Trp Leu Asn Thr His Phe Asn Thr Ser Ile Glu Met Thr Glu
165 170 175

Glu Ser Val Glu Ala Cys Glu Lys Leu Val Asn Ala Gly Val Pro Val
180 185 190

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

Met Lys Lys Leu Met His Asp Leu Val Lys Ile Arg Val Arg Pro Tyr
210 215 220

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

Pro Val Ser Lys Gly Leu Glu Ile Ile Glu Gly Leu Arg Gly His Thr
245 250 255

Ser Gly Asn Ala Val Pro Thr Phe Val Val His Ala Pro Gly Gly Gly
260 265 270

Gly Lys Ile Ala Leu Gln Pro Asn Tyr Val Leu Ser Gln Ser Pro Asp
275 280 285

-82-

Lys Val Ile Leu Arg Asn Phe Glu Gly Val Ile Thr Ser Tyr Pro Glu
 290 295 300

Pro Glu Asn Tyr Ile Pro Asn Gln Ala Asp Ala Tyr Phe Glu Ser Val
 305 310 315 320

Phe Pro Glu Thr Ala Asp Lys Lys Glu Pro Ile Gly Leu Ser Ala Ile
 325 330 335

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

Lys Arg Arg Glu Ala Tyr Ile Ala Asn Pro Glu His Glu Thr Leu Lys
 355 360 365

Asp Arg Arg Glu Lys Arg Gly Gln Leu Lys Glu Lys Lys Phe Leu Ala
 370 375 380

Gln Gln Lys Lys Gln Lys Glu Thr Glu Cys Gly Gly Asp Ser Ser
 385 390 395

<210> 52
 <211> 1245
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<220>
 <221> misc_feature
 <223> This parental sequence is a modification of the wild-type KAM of
 Clostridium stricklandii

<220>
 <221> CDS
 <222> (1)..(1245)

<400> 52
 atg agt tta aag gat aag ttt ttt aca cat gta agc caa gaa gat tgg 48
 Met Ser Leu Lys Asp Lys Phe Phe Thr His Val Ser Gln Glu Asp Trp
 1 5 10 15
 aat gat tgg aaa tgg caa gta aga aat cgt ata aag act gtt gaa gaa 96
 Asn Asp Trp Lys Trp Gln Val Arg Asn Arg Ile Lys Thr Val Glu Glu
 20 25 30
 ctt aaa aaa tat att cca ctt act cca gaa gaa gaa gaa ggg gta aaa 144
 Leu Lys Lys Tyr Ile Pro Leu Thr Pro Glu Glu Glu Glu Gly Val Lys
 35 40 45

-83-

cgc tgt ctt gat aca tta cgt atg gct att act cca tac tat cta tcg	192
Arg Cys Leu Asp Thr Leu Arg Met Ala Ile Thr Pro Tyr Tyr Leu Ser	
50 55 60	
cta att gat gta gaa aat cca aat gac cct gta aga aag caa gct gta	240
Leu Ile Asp Val Glu Asn Pro Asn Asp Pro Val Arg Lys Gln Ala Val	
65 70 75 80	
cct ctt tct tta gag ctg cat cgc gca gcg tct gat atg gaa gac cca	288
Pro Leu Ser Leu Glu Leu His Arg Ala Ala Ser Asp Met Glu Asp Pro	
85 90 95	
ctt cat gaa gat gga gat tct cca gtt cca gga ctt aca cat cgc tat	336
Leu His Glu Asp Gly Asp Ser Pro Val Pro Gly Leu Thr His Arg Tyr	
100 105 110	
cct gat cgc gtt ctt ctt tta atg act gat caa tgt tca gta tac tgc	384
Pro Asp Arg Val Leu Leu Leu Met Thr Asp Gln Cys Ser Val Tyr Cys	
115 120 125	
cgc cac tgt act cgt aga cgc ttc gct ggt cga aca gat tct gct gtt	432
Arg His Cys Thr Arg Arg Arg Phe Ala Gly Arg Thr Asp Ser Ala Val	
130 135 140	
gat acg aag caa ata gat gct gcg att gaa tat atc aaa aat act cca	480
Asp Thr Lys Gln Ile Asp Ala Ala Ile Glu Tyr Ile Lys Asn Thr Pro	
145 150 155 160	
caa gta aga gac gtt cta ctt tca gga gga gat gct cta tta atc tca	528
Gln Val Arg Asp Val Leu Leu Ser Gly Gly Asp Ala Leu Leu Ile Ser	
165 170 175	
gat gaa aag ctt gag tac aca atc aga aga ctt cgt gaa ata cca cac	576
Asp Glu Lys Leu Glu Tyr Thr Ile Arg Arg Leu Arg Glu Ile Pro His	
180 185 190	
gtt gag gtt att cgt att gga tca cgt gta cca gtt gta atg cca caa	624
Val Glu Val Ile Arg Ile Gly Ser Arg Val Pro Val Val Met Pro Gln	
195 200 205	
cgt att aca cca gaa cta gtt tct atg ctt aaa aag tat cat cca gta	672
Arg Ile Thr Pro Glu Leu Val Ser Met Leu Lys Lys Tyr His Pro Val	
210 215 220	
tgg tta aat aca cac ttc aac cat cct aat gaa att act gaa gag tct	720
Trp Leu Asn Thr His Phe Asn His Pro Asn Glu Ile Thr Glu Glu Ser	
225 230 235 240	
aaa cgt gca tgt gag tta ctt gct gat gca ggt att cct ctt gga aat	768
Lys Arg Ala Cys Glu Leu Leu Ala Asp Ala Gly Ile Pro Leu Gly Asn	
245 250 255	
caa agt gtg ctt ctt gca ggt gta aat gat tgc atg cac gtt atg aaa	816
Gln Ser Val Leu Leu Ala Gly Val Asn Asp Cys Met His Val Met Lys	
260 265 270	
aaa cta gta aat gac tta gtt aaa ata cgc gta cgt cct tac tat att	864

-84-

Lys Leu Val Asn Asp Leu Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile
 275 280 285
 tat caa tgt gac ctt tca gtt gga att gag cac ttt cgc act cca gtt 912
 Tyr Gln Cys Asp Leu Ser Val Gly Ile Glu His Phe Arg Thr Pro Val
 290 295 300
 gca aag gga ata gaa ata att gaa ggc tta aga gga cat act tca gga 960
 Ala Lys Gly Ile Glu Ile Ile Glu Gly Leu Arg Gly His Thr Ser Gly
 305 310 315 320
 tac tgc gtt cct aca ttt gtt gtg cat gca cct ggt ggt gga gga aaa 1008
 Tyr Cys Val Pro Thr Phe Val Val His Ala Pro Gly Gly Gly Gly Lys
 325 330 335
 act cca gtt atg cca aac tat gtt att tca caa aat cac aat aaa gtt 1056
 Thr Pro Val Met Pro Asn Tyr Val Ile Ser Gln Asn His Asn Lys Val
 340 345 350
 att tta cgt aac ttt gaa ggt gta att aca act tac gat gag cct gat 1104
 Ile Leu Arg Asn Phe Glu Gly Val Ile Thr Thr Tyr Asp Glu Pro Asp
 355 360 365
 cat tat act ttc cac tgt gac tgt gat gta tgc act gga aaa aca aat 1152
 His Tyr Thr Phe His Cys Asp Cys Asp Val Cys Thr Gly Lys Thr Asn
 370 375 380
 gtt cat aag gtt gga gta gct gga ctt cta aat gga gag aca gcg aca 1200
 Val His Lys Val Gly Val Ala Gly Leu Leu Asn Gly Glu Thr Ala Thr
 385 390 395 400
 ctt gaa cct gag ggt ttg gaa aga aaa caa aga gga cat cac taa 1245
 Leu Glu Pro Glu Gly Leu Glu Arg Lys Gln Arg Gly His His
 405 410

<210> 53

<211> 414

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 53

Met Ser Leu Lys Asp Lys Phe Phe Thr His Val Ser Gln Glu Asp Trp
 1 5 10 15

Asn Asp Trp Lys Trp Gln Val Arg Asn Arg Ile Lys Thr Val Glu Glu
 20 25 30

Leu Lys Lys Tyr Ile Pro Leu Thr Pro Glu Glu Glu Glu Gly Val Lys
 35 40 45

-85-

Arg Cys Leu Asp Thr Leu Arg Met Ala Ile Thr Pro Tyr Tyr Leu Ser
 50 55 60

Leu Ile Asp Val Glu Asn Pro Asn Asp Pro Val Arg Lys Gln Ala Val
 65 70 75 80

Pro Leu Ser Leu Glu Leu His Arg Ala Ala Ser Asp Met Glu Asp Pro
 85 90 95

Leu His Glu Asp Gly Asp Ser Pro Val Pro Gly Leu Thr His Arg Tyr
 100 105 110

Pro Asp Arg Val Leu Leu Leu Met Thr Asp Gln Cys Ser Val Tyr Cys
 115 120 125

Arg His Cys Thr Arg Arg Arg Phe Ala Gly Arg Thr Asp Ser Ala Val
 130 135 140

Asp Thr Lys Gln Ile Asp Ala Ala Ile Glu Tyr Ile Lys Asn Thr Pro
 145 150 155 160

Gln Val Arg Asp Val Leu Leu Ser Gly Gly Asp Ala Leu Leu Ile Ser
 165 170 175

Asp Glu Lys Leu Glu Tyr Thr Ile Arg Arg Leu Arg Glu Ile Pro His
 180 185 190

Val Glu Val Ile Arg Ile Gly Ser Arg Val Pro Val Val Met Pro Gln
 195 200 205

Arg Ile Thr Pro Glu Leu Val Ser Met Leu Lys Lys Tyr His Pro Val
 210 215 220

Trp Leu Asn Thr His Phe Asn His Pro Asn Glu Ile Thr Glu Glu Ser
 225 230 235 240

Lys Arg Ala Cys Glu Leu Leu Ala Asp Ala Gly Ile Pro Leu Gly Asn
 245 250 255

Gln Ser Val Leu Leu Ala Gly Val Asn Asp Cys Met His Val Met Lys
 260 265 270

Lys Leu Val Asn Asp Leu Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile
 275 280 285

-86-

Tyr Gln Cys Asp Leu Ser Val Gly Ile Glu His Phe Arg Thr Pro Val
290 295 300

Ala Lys Gly Ile Glu Ile Ile Glu Gly Leu Arg Gly His Thr Ser Gly
305 310 315 320

Tyr Cys Val Pro Thr Phe Val Val His Ala Pro Gly Gly Gly Gly Lys
325 330 335

Thr Pro Val Met Pro Asn Tyr Val Ile Ser Gln Asn His Asn Lys Val
340 345 350

Ile Leu Arg Asn Phe Glu Gly Val Ile Thr Thr Tyr Asp Glu Pro Asp
355 360 365

His Tyr Thr Phe His Cys Asp Cys Asp Val Cys Thr Gly Lys Thr Asn
370 375 380

Val His Lys Val Gly Val Ala Gly Leu Leu Asn Gly Glu Thr Ala Thr
385 390 395 400

Leu Glu Pro Glu Gly Leu Glu Arg Lys Gln Arg Gly His His
405 410

<210> 54
<211> 1251
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<220>
<221> CDS
<222> (1)..(1251)

<400> 54
atg gca gaa agt cgt aga aag tat tat ttc cct gat gtc acc gat gag 48
Met Ala Glu Ser Arg Arg Lys Tyr Tyr Phe Pro Asp Val Thr Asp Glu
1 5 10 15
caa tgg tac gac tgg cat tgg cag gtc ctc aat cga att aag acg ctc 96
Gln Trp Tyr Asp Trp His Trp Gln Val Leu Asn Arg Ile Lys Thr Leu
20 25 30
gac cag ctg aaa aag tac gtt aca ctc acc gct gaa gaa gaa gag gga 144
Asp Gln Leu Lys Lys Tyr Val Thr Leu Thr Ala Glu Glu Glu Glu Gly

-87-

35	40	45	
gta aaa gaa tcg ccc Val Lys Glu Ser Pro 50	aaa gta ctc cga atg gct atc Lys Val Leu Arg Met Ala 55	aca cct tat tat Thr Pro Tyr Tyr 60	192
ttg agt ttg ata gac Leu Ser Leu Ile Asp 65	ccc gag aat cct aat Pro Glu Asn Pro Asn 70	tgt ccg att cgt aaa caa Cys Pro Ile Arg Lys Gln 75	240
gcc att cct act caa Ala Ile Pro Thr 85	cag gaa ctg gta cgt Gln Glu Leu Val Arg 90	gct cct gaa gat cag gta Ala Pro Glu Asp Gln Val 95	288
gac cca ctt agt gaa Asp Pro Leu Ser Glu 100	gat gaa gat tcg ccc Asp Glu Asp Ser Pro 105	gta ccc gga ctg act cat Val Pro Gly Leu Thr His 110	336
cgt tat ccg gat cgt Arg Tyr Pro Asp Arg 115	gta ttg ttc ctt atc Val Leu Phe Leu Ile 120	acg gac aaa tgt tcg atg Thr Asp Lys Cys Ser Met 125	384
tac tgt cgt cat tgt Tyr Cys Arg His Cys 130	act cgc cgt cgc ttc Thr Arg Arg Arg Phe 135	gca gga cag aaa gat gct Ala Gly Gln Lys Asp Ala 140	432
tct tct cct tct gag Ser Ser Pro Ser Glu 145	cgc atc gat cga tgc Arg Ile Asp Arg Cys 150	att gac tat ata gcc aat Ile Asp Tyr Ile Ala Asn 155	480
aca ccg aca gtc cgc Thr Pro Thr Val Arg 165	gat gtt ttg cta tcg Asp Val Leu Leu Ser 170	gga ggc gat gcc ctc ctt Gly Gly Asp Ala Leu Leu 175	528
gtc agc gac gaa cgc Val Ser Asp Glu Arg 180	ttg gaa tac ata ttg Leu Glu Tyr Ile Leu 185	aag cgt ctg cgc gaa gta Lys Arg Leu Arg Glu Val 190	576
cct cat gtg gag att Pro His Val Glu Ile 195	gtt cgt ata gga agc Val Arg Ile Gly Ser 200	cgt acg ccg gta gtc ctc Arg Thr Pro Val Val Leu 205	624
cct cag cgt ata acg Pro Gln Arg Ile Thr 210	cct caa ttg gtg gat Pro Gln Leu Val Asp 215	atg ctc aaa aaa tat cat Met Leu Lys Lys Tyr His 220	672
ccg gtg tgg ctg aac Pro Val Trp Leu Asn 225	act cac ttc aac cac Thr His Phe Asn His 230	ccg aat gaa gtt acc gaa Pro Asn Glu Val Thr Glu 235	720
gaa gca gtg gag gct Glu Ala Val Glu Ala 245	tgt gaa aga atg gcc Cys Glu Arg Met Ala 250	aat gcc ggt att ccg ttg Asn Ala Gly Ile Pro Leu 255	768
ggc aac caa acg gtt Gly Asn Gln Thr Val 260	tta ttg cgt gga atc Leu Leu Arg Gly Ile 265	aat gat tgt aca cat gtg Asn Asp Cys Thr His Val 270	816

-88-

```

atg aag aga ttg gta cat ttg ctg gta aag atg cgt gtg cgt cct tac      864
Met Lys Arg Leu Val His Leu Leu Val Lys Met Arg Val Arg Pro Tyr
      275                      280                      285

tat ata tat gta tgc gat ctt tcg ctt gga ata ggt cat ttc cgc acg      912
Tyr Ile Tyr Val Cys Asp Leu Ser Leu Gly Ile Gly His Phe Arg Thr
      290                      295                      300

ccg gta tct aaa gga atc gaa att atc gaa aat ttg cgc gga cac acc      960
Pro Val Ser Lys Gly Ile Glu Ile Ile Glu Asn Leu Arg Gly His Thr
      305                      310                      315                      320

tcg ggc tat gca gtt cct acc ttt gtg gta ggt gct ccg ggg ggt ggt      1008
Ser Gly Tyr Ala Val Pro Thr Phe Val Val Gly Ala Pro Gly Gly Gly
      325                      330                      335

ggt aag ata cct gta acg ccg aac tat gtt gta tct cag tcc cca cga      1056
Gly Lys Ile Pro Val Thr Pro Asn Tyr Val Val Ser Gln Ser Pro Arg
      340                      345                      350

cat gtg gtt ctt cgc aat tat gaa ggt gtt atc aca acc tat acg gag      1104
His Val Val Leu Arg Asn Tyr Glu Gly Val Ile Thr Thr Tyr Thr Glu
      355                      360                      365

ccg gag aat tat cat gag gag tgc gat tgt gag gac tgt cga gcc ggt      1152
Pro Glu Asn Tyr His Glu Glu Cys Asp Cys Glu Asp Cys Arg Ala Gly
      370                      375                      380

aag cat aaa gag ggt gta gct gca ctt tcc gga ggt cag cag ttg gct      1200
Lys His Lys Glu Gly Val Ala Ala Leu Ser Gly Gly Gln Gln Leu Ala
      385                      390                      395                      400

atc gag cct tcc gac tta gct cgc aaa aaa cgc aag ttt gat aag aac      1248
Ile Glu Pro Ser Asp Leu Ala Arg Lys Lys Arg Lys Phe Asp Lys Asn
      405                      410                      415

taa                                                                    1251

```

```

<210> 55
<211> 416
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Synthetic Construct

```

```

<400> 55

```

```

Met Ala Glu Ser Arg Arg Lys Tyr Tyr Phe Pro Asp Val Thr Asp Glu
1          5          10          15

```

```

Gln Trp Tyr Asp Trp His Trp Gln Val Leu Asn Arg Ile Lys Thr Leu
20          25          30

```

```

Asp Gln Leu Lys Lys Tyr Val Thr Leu Thr Ala Glu Glu Glu Gly

```

-89-

35	40	45
Val Lys Glu Ser Pro Lys Val Leu Arg Met Ala Ile Thr Pro Tyr Tyr		
50	55	60
Leu Ser Leu Ile Asp Pro Glu Asn Pro Asn Cys Pro Ile Arg Lys Gln		
65	70	75
Ala Ile Pro Thr Gln Gln Glu Leu Val Arg Ala Pro Glu Asp Gln Val		
	85	90
Asp Pro Leu Ser Glu Asp Glu Asp Ser Pro Val Pro Gly Leu Thr His		
	100	105
Arg Tyr Pro Asp Arg Val Leu Phe Leu Ile Thr Asp Lys Cys Ser Met		
	115	120
Tyr Cys Arg His Cys Thr Arg Arg Arg Phe Ala Gly Gln Lys Asp Ala		
	130	135
Ser Ser Pro Ser Glu Arg Ile Asp Arg Cys Ile Asp Tyr Ile Ala Asn		
145	150	155
Thr Pro Thr Val Arg Asp Val Leu Leu Ser Gly Gly Asp Ala Leu Leu		
	165	170
Val Ser Asp Glu Arg Leu Glu Tyr Ile Leu Lys Arg Leu Arg Glu Val		
	180	185
Pro His Val Glu Ile Val Arg Ile Gly Ser Arg Thr Pro Val Val Leu		
	195	200
Pro Gln Arg Ile Thr Pro Gln Leu Val Asp Met Leu Lys Lys Tyr His		
	210	215
Pro Val Trp Leu Asn Thr His Phe Asn His Pro Asn Glu Val Thr Glu		
225	230	235
Glu Ala Val Glu Ala Cys Glu Arg Met Ala Asn Ala Gly Ile Pro Leu		
	245	250
Gly Asn Gln Thr Val Leu Leu Arg Gly Ile Asn Asp Cys Thr His Val		
	260	270

-90-

Met Lys Arg Leu Val His Leu Leu Val Lys Met Arg Val Arg Pro Tyr
 275 280 285

Tyr Ile Tyr Val Cys Asp Leu Ser Leu Gly Ile Gly His Phe Arg Thr
 290 295 300

Pro Val Ser Lys Gly Ile Glu Ile Ile Glu Asn Leu Arg Gly His Thr
 305 310 315 320

Ser Gly Tyr Ala Val Pro Thr Phe Val Val Gly Ala Pro Gly Gly Gly
 325 330 335

Gly Lys Ile Pro Val Thr Pro Asn Tyr Val Val Ser Gln Ser Pro Arg
 340 345 350

His Val Val Leu Arg Asn Tyr Glu Gly Val Ile Thr Thr Tyr Thr Glu
 355 360 365

Pro Glu Asn Tyr His Glu Glu Cys Asp Cys Glu Asp Cys Arg Ala Gly
 370 375 380

Lys His Lys Glu Gly Val Ala Ala Leu Ser Gly Gly Gln Gln Leu Ala
 385 390 395 400

Ile Glu Pro Ser Asp Leu Ala Arg Lys Lys Arg Lys Phe Asp Lys Asn
 405 410 415

<210> 56
 <211> 1278
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<220>
 <221> CDS
 <222> (1)..(1278)

<400> 56
 atg aat aca gtt aat act cgt aaa aaa ttt ttc cca aat gta act gat 48
 Met Asn Thr Val Asn Thr Arg Lys Lys Phe Phe Pro Asn Val Thr Asp
 1 5 10 15
 gaa gaa tgg aat gat tgg aca tgg caa gta aaa aac cgc ctt aaa agt 96
 Glu Glu Trp Asn Asp Trp Thr Trp Gln Val Lys Asn Arg Leu Lys Ser
 20 25 30

-91-

gtt gaa gat tta gaa aaa tat gtt gat tta agt gaa gaa gaa aca gaa Val Glu Asp Leu Glu Lys Tyr Val Asp Leu Ser Glu Glu Glu Thr Glu 35 40 45	144
ggg gtt gta cgc act ctt gaa act tta cgt atg gca atc act cca ttt Gly Val Val Arg Thr Leu Glu Thr Leu Arg Met Ala Ile Thr Pro Phe 50 55 60	192
tac ttc tca ttg ata gat ttg aat agt gat cgc tgc cca ata cgt aag Tyr Phe Ser Leu Ile Asp Leu Asn Ser Asp Arg Cys Pro Ile Arg Lys 65 70 75 80	240
caa gct ata cct act ata cga gaa ata cat caa tct gat gct gat atg Gln Ala Ile Pro Thr Ile Arg Glu Ile His Gln Ser Asp Ala Asp Met 85 90 95	288
ttg gat cct cta cat gaa gat gaa gac tct cca gta cca gga tta act Leu Asp Pro Leu His Glu Asp Glu Asp Ser Pro Val Pro Gly Leu Thr 100 105 110	336
cat cgc tat cca gat cgt gtt tta ctt cta ata aca gac atg tgt tct His Arg Tyr Pro Asp Arg Val Leu Leu Leu Ile Thr Asp Met Cys Ser 115 120 125	384
gta tac tgt cgc cac tgc act cgt cgc aga ttt gct ggg tca agt gat Val Tyr Cys Arg His Cys Thr Arg Arg Arg Phe Ala Gly Ser Ser Asp 130 135 140	432
ggg gct atg cct atg gat aga att gac aaa gca ata gaa tat att gca Gly Ala Met Pro Met Asp Arg Ile Asp Lys Ala Ile Glu Tyr Ile Ala 145 150 155 160	480
aaa act cca caa gta agg gat gta ttg tta tca gga gga gat gca ctt Lys Thr Pro Gln Val Arg Asp Val Leu Leu Ser Gly Gly Asp Ala Leu 165 170 175	528
cta gtt tct aat aaa aaa tta gaa agc ata atc caa aaa cta cgc gca Leu Val Ser Asn Lys Lys Leu Glu Ser Ile Ile Gln Lys Leu Arg Ala 180 185 190	576
ata cct cat gtt gaa ata atc aga ata gga agt cgt aca cca gtt gtt Ile Pro His Val Glu Ile Ile Arg Ile Gly Ser Arg Thr Pro Val Val 195 200 205	624
tta cct caa aga att act cct gaa tta tgt aat atg tta aag aaa tat Leu Pro Gln Arg Ile Thr Pro Glu Leu Cys Asn Met Leu Lys Lys Tyr 210 215 220	672
cat cca att tgg atg aat act cat ttt aac cac cct caa gaa gta acg His Pro Ile Trp Met Asn Thr His Phe Asn His Pro Gln Glu Val Thr 225 230 235 240	720
cca gaa gct aaa aaa gct tgt gaa atg ttg gca gat gca gga gtt cca Pro Glu Ala Lys Lys Ala Cys Glu Met Leu Ala Asp Ala Gly Val Pro 245 250 255	768
tta gga aat caa act gta cta tta aga gga ata aat gac agt gta cct	816

-92-

Leu Gly Asn Gln Thr Val Leu Leu Arg Gly Ile Asn Asp Ser Val Pro	
260 265 270	
gta atg aaa agg tta gta cat gat tta gta atg atg cgt gta cgc cct	864
Val Met Lys Arg Leu Val His Asp Leu Val Met Met Arg Val Arg Pro	
275 280 285	
tat tat att tac caa tgt gac tta tct atg gga ctc gaa cac ttc cgc	912
Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser Met Gly Leu Glu His Phe Arg	
290 295 300	
aca cca gtt tct aaa ggt ata gaa att att gaa gga tta cgt gga cat	960
Thr Pro Val Ser Lys Gly Ile Glu Ile Ile Glu Gly Leu Arg Gly His	
305 310 315 320	
aca tct gga tat gca gta cca aca ttt gtt gtg cat gca cct ggt ggt	1008
Thr Ser Gly Tyr Ala Val Pro Thr Phe Val Val His Ala Pro Gly Gly	
325 330 335	
gga gga aaa act cca gta atg cct caa tat gta att tct caa tct cct	1056
Gly Gly Lys Thr Pro Val Met Pro Gln Tyr Val Ile Ser Gln Ser Pro	
340 345 350	
cat cgt gta gtt tta cgc aac ttt gaa gga gtt ata aca act tat aca	1104
His Arg Val Val Leu Arg Asn Phe Glu Gly Val Ile Thr Thr Tyr Thr	
355 360 365	
gaa cca gaa aat tat aca cat gaa cct tgt tat gat gaa gaa aaa ttt	1152
Glu Pro Glu Asn Tyr Thr His Glu Pro Cys Tyr Asp Glu Glu Lys Phe	
370 375 380	
gaa aaa atg tat gaa ata agt gga gtt tat atg cta gat gaa gga tta	1200
Glu Lys Met Tyr Glu Ile Ser Gly Val Tyr Met Leu Asp Glu Gly Leu	
385 390 395 400	
gaa atg tca cta gaa cct agc cac tta gca cgt cat gaa cgc aat aaa	1248
Glu Met Ser Leu Glu Pro Ser His Leu Ala Arg His Glu Arg Asn Lys	
405 410 415	
aag aga gca gaa gct gaa ggg aaa aaa taa	1278
Lys Arg Ala Glu Ala Glu Gly Lys Lys	
420 425	
<210> 57	
<211> 425	
<212> PRT	
<213> Artificial Sequence	
<220>	
<223> Synthetic Construct	
<400> 57	
Met Asn Thr Val Asn Thr Arg Lys Lys Phe Phe Pro Asn Val Thr Asp	
1 5 10 15	

-93-

Glu Glu Trp Asn Asp Trp Thr Trp Gln Val Lys Asn Arg Leu Lys Ser
20 25 30

Val Glu Asp Leu Glu Lys Tyr Val Asp Leu Ser Glu Glu Glu Thr Glu
35 40 45

Gly Val Val Arg Thr Leu Glu Thr Leu Arg Met Ala Ile Thr Pro Phe
50 55 60

Tyr Phe Ser Leu Ile Asp Leu Asn Ser Asp Arg Cys Pro Ile Arg Lys
65 70 75 80

Gln Ala Ile Pro Thr Ile Arg Glu Ile His Gln Ser Asp Ala Asp Met
85 90 95

Leu Asp Pro Leu His Glu Asp Glu Asp Ser Pro Val Pro Gly Leu Thr
100 105 110

His Arg Tyr Pro Asp Arg Val Leu Leu Leu Ile Thr Asp Met Cys Ser
115 120 125

Val Tyr Cys Arg His Cys Thr Arg Arg Arg Phe Ala Gly Ser Ser Asp
130 135 140

Gly Ala Met Pro Met Asp Arg Ile Asp Lys Ala Ile Glu Tyr Ile Ala
145 150 155 160

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

Leu Val Ser Asn Lys Lys Leu Glu Ser Ile Ile Gln Lys Leu Arg Ala
180 185 190

Ile Pro His Val Glu Ile Ile Arg Ile Gly Ser Arg Thr Pro Val Val
195 200 205

Leu Pro Gln Arg Ile Thr Pro Glu Leu Cys Asn Met Leu Lys Lys Tyr
210 215 220

His Pro Ile Trp Met Asn Thr His Phe Asn His Pro Gln Glu Val Thr
225 230 235 240

Pro Glu Ala Lys Lys Ala Cys Glu Met Leu Ala Asp Ala Gly Val Pro
245 250 255

-94-

Leu Gly Asn Gln Thr Val Leu Leu Arg Gly Ile Asn Asp Ser Val Pro
 260 265 270

Val Met Lys Arg Leu Val His Asp Leu Val Met Met Arg Val Arg Pro
 275 280 285

Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser Met Gly Leu Glu His Phe Arg
 290 295 300

Thr Pro Val Ser Lys Gly Ile Glu Ile Ile Glu Gly Leu Arg Gly His
 305 310 315 320

Thr Ser Gly Tyr Ala Val Pro Thr Phe Val Val His Ala Pro Gly Gly
 325 330 335

Gly Gly Lys Thr Pro Val Met Pro Gln Tyr Val Ile Ser Gln Ser Pro
 340 345 350

His Arg Val Val Leu Arg Asn Phe Glu Gly Val Ile Thr Thr Tyr Thr
 355 360 365

Glu Pro Glu Asn Tyr Thr His Glu Pro Cys Tyr Asp Glu Glu Lys Phe
 370 375 380

Glu Lys Met Tyr Glu Ile Ser Gly Val Tyr Met Leu Asp Glu Gly Leu
 385 390 395 400

Glu Met Ser Leu Glu Pro Ser His Leu Ala Arg His Glu Arg Asn Lys
 405 410 415

Lys Arg Ala Glu Ala Glu Gly Lys Lys
 420 425

<210> 58
 <211> 1416
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<220>
 <221> CDS
 <222> (1)..(1416)

-95-

<400> 58
 atg aaa aac aaa tgg tat aaa ccg aaa cgg cat tgg aag gag atc gag 48
 Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
 1 5 10 15

tta tgg aag gac gtt ccg gaa gag aaa tgg aac gat tgg ctt tgg cag 96
 Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
 20 25 30

ctg aca cac act gta aga acg tta gat gat tta aag aaa gtc att aat 144
 Leu Thr His Thr Val Arg Thr Leu Asp Asp Leu Lys Lys Val Ile Asn
 35 40 45

ctg acc gag gat gaa gag gaa ggc gtc cgt att tct acc aaa acg atc 192
 Leu Thr Glu Asp Glu Glu Glu Gly Val Arg Ile Ser Thr Lys Thr Ile
 50 55 60

ccc tta aat att aca cct tac tat gct tct tta atg gac ccc gac aat 240
 Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Ser Leu Met Asp Pro Asp Asn
 65 70 75 80

ccg aga tgc ccg gta cgc atg cag tct gtg ccg ctt tct gaa gaa atg 288
 Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
 85 90 95

cac aaa aca aaa tac gat atg gaa gac ccg ctt cat gag gat gaa gat 336
 His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu His Glu Asp Glu Asp
 100 105 110

tca ccg gta ccc ggt ctg aca cac cgc tat ccc gac cgt gtg ctg ttt 384
 Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
 115 120 125

ctt gtc acg aat caa tgt tcc gtg tac tgc cgc cac tgc aca cgc cgg 432
 Leu Val Thr Asn Gln Cys Ser Val Tyr Cys Arg His Cys Thr Arg Arg
 130 135 140

cgc ttt tcc gga caa atc gga atg ggc gtc ccc aaa aaa cag ctt gat 480
 Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
 145 150 155 160

gct gca att gct tat atc cgg gaa aca ccc gaa atc cgc gat tgt tta 528
 Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
 165 170 175

att tca ggc ggt gat ggg ctg ctc atc aac gac caa att tta gaa tat 576
 Ile Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr
 180 185 190

att tta aaa gag ctg cgc agc att ccg cat ctg gaa gtc atc cgc atc 624
 Ile Leu Lys Glu Leu Arg Ser Ile Pro His Leu Glu Val Ile Arg Ile
 195 200 205

gga aca cgt gct ccc gtc gtc ttt ccg cag cgc att acc gat cat ctg 672
 Gly Thr Arg Ala Pro Val Val Phe Pro Gln Arg Ile Thr Asp His Leu
 210 215 220

-96-

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

-97-

450 455 460
 gaa tgc gga ggg gat tct tca taa 1416
 Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 59
 <211> 471
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 59

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
 1 5 10 15

Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
 20 25 30

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

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

Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Ser Leu Met Asp Pro Asp Asn
 65 70 75 80

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
 85 90 95

His Lys Thr Lys Tyr Asp Met Glu Asp Pro Leu His Glu Asp Glu Asp
 100 105 110

Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
 115 120 125

Leu Val Thr Asn Gln Cys Ser Val Tyr Cys Arg His Cys Thr Arg Arg
 130 135 140

Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
 145 150 155 160

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu

-98-

				165						170						175
Ile	Ser	Gly	Gly	Asp	Gly	Leu	Leu	Ile	Asn	Asp	Gln	Ile	Leu	Glu	Tyr	
			180					185					190			
Ile	Leu	Lys	Glu	Leu	Arg	Ser	Ile	Pro	His	Leu	Glu	Val	Ile	Arg	Ile	
		195					200					205				
Gly	Thr	Arg	Ala	Pro	Val	Val	Phe	Pro	Gln	Arg	Ile	Thr	Asp	His	Leu	
	210					215					220					
Cys	Glu	Ile	Leu	Lys	Lys	Tyr	His	Pro	Val	Trp	Leu	Asn	Thr	His	Phe	
225					230					235					240	
Asn	Thr	Ser	Ile	Glu	Met	Thr	Glu	Glu	Ser	Val	Glu	Ala	Cys	Glu	Lys	
				245					250					255		
Leu	Val	Asn	Ala	Gly	Val	Pro	Val	Gly	Asn	Gln	Ala	Val	Val	Leu	Ala	
			260					265					270			
Gly	Ile	Asn	Asp	Ser	Val	Pro	Ile	Met	Lys	Lys	Leu	Met	His	Asp	Leu	
		275					280					285				
Val	Lys	Ile	Arg	Val	Arg	Pro	Tyr	Tyr	Ile	Tyr	Gln	Cys	Asp	Leu	Ser	
	290					295					300					
Glu	Gly	Ile	Gly	His	Phe	Arg	Ala	Pro	Val	Ser	Lys	Gly	Leu	Glu	Ile	
305					310					315					320	
Ile	Glu	Gly	Leu	Arg	Gly	His	Thr	Ser	Gly	Tyr	Ala	Val	Pro	Thr	Phe	
				325					330					335		
Val	Val	His	Ala	Pro	Gly	Gly	Gly	Gly	Lys	Ile	Ala	Leu	Gln	Pro	Asn	
			340					345					350			
Tyr	Val	Leu	Ser	Gln	Ser	Pro	Asp	Lys	Val	Ile	Leu	Arg	Asn	Phe	Glu	
		355					360					365				
Gly	Val	Ile	Thr	Ser	Tyr	Pro	Glu	Pro	Glu	Asn	Tyr	Ile	Pro	Asn	Gln	
	370					375					380					
Ala	Asp	Ala	Tyr	Phe	Glu	Ser	Val	Phe	Pro	Glu	Thr	Ala	Asp	Lys	Lys	
385					390					395					400	

-99-

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Phe
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Asp Gln
 435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 60

<211> 471

<212> PRT

<213> lysine 2,3-aminomutase from *Bacillus subtilis*

<400> 60

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
 1 5 10 15

Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
 20 25 30

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

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

Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Ser Leu Met Asp Pro Asp Asn
 65 70 75 80

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
 85 90 95

His Lys Thr Lys Tyr Asp Leu Glu Asp Pro Leu His Glu Asp Glu Asp
 100 105 110

Ser Arg Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
 115 120 125

-100 -

Leu	Val	Thr	Asn	Gln	Cys	Ser	Met	Tyr	Cys	Arg	Tyr	Cys	Thr	Arg	Arg
130						135					140				
Arg	Phe	Ser	Gly	Gln	Ile	Gly	Met	Gly	Val	Pro	Lys	Lys	Gln	Leu	Asp
145					150					155					160
Ala	Ala	Ile	Ala	Tyr	Ile	Arg	Glu	Thr	Pro	Glu	Ile	Arg	Asp	Cys	Leu
				165					170					175	
Ile	Ser	Gly	Gly	Asp	Gly	Leu	Leu	Ile	Asn	Asp	Gln	Ile	Leu	Glu	Tyr
			180					185					190		
Ile	Leu	Lys	Glu	Leu	Arg	Ser	Ile	Pro	His	Leu	Glu	Val	Ile	Arg	Ile
		195					200					205			
Gly	Thr	Arg	Ala	Pro	Val	Val	Phe	Pro	Gln	Arg	Ile	Thr	Asp	His	Leu
	210					215					220				
Cys	Glu	Ile	Leu	Lys	Lys	Tyr	His	Pro	Val	Trp	Leu	Asn	Thr	His	Phe
225					230					235					240
Asn	Thr	Ser	Ile	Glu	Met	Thr	Glu	Glu	Ser	Val	Glu	Ala	Cys	Glu	Lys
				245					250					255	
Leu	Val	Asn	Ala	Gly	Val	Pro	Val	Gly	Asn	Gln	Ala	Val	Val	Leu	Ala
			260					265					270		
Gly	Ile	Asn	Asp	Ser	Val	Pro	Ile	Met	Lys	Lys	Leu	Met	His	Asp	Leu
		275					280					285			
Val	Lys	Ile	Arg	Val	Arg	Pro	Tyr	Tyr	Ile	Tyr	Gln	Cys	Asp	Leu	Ser
	290					295					300				
Glu	Gly	Ile	Gly	His	Phe	Arg	Ala	Pro	Val	Ser	Lys	Gly	Leu	Glu	Ile
305					310					315					320
Ile	Glu	Gly	Leu	Arg	Gly	His	Thr	Ser	Gly	Tyr	Ala	Val	Pro	Thr	Phe
				325					330					335	
Val	Val	Asp	Ala	Pro	Gly	Gly	Gly	Gly	Lys	Ile	Ala	Leu	Gln	Pro	Asn
			340					345					350		

-101-

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
 385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Phe
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Arg Arg Asp Gln
 435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 61

<211> 471

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct

<400> 61

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
 1 5 10 15

Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln
 20 25 30

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

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

-102-

Pro Leu Asn Ile Thr Pro Tyr Tyr Ala Ser Leu Met Asp Pro Asp Asn
65 70 75 80

Pro Arg Cys Pro Val Arg Met Gln Ser Val Pro Leu Ser Glu Glu Met
85 90 95

His Lys Thr Lys Tyr Asp Leu Glu Asp Pro Leu His Glu Asp Glu Asp
100 105 110

Ser Pro Val Pro Gly Leu Thr His Arg Tyr Pro Asp Arg Val Leu Phe
115 120 125

Leu Val Thr Asn Gln Cys Ser Met Tyr Cys Arg Tyr Cys Thr Arg Arg
130 135 140

Arg Phe Ser Gly Gln Ile Gly Met Gly Val Pro Lys Lys Gln Leu Asp
145 150 155 160

Ala Ala Ile Ala Tyr Ile Arg Glu Thr Pro Glu Ile Arg Asp Cys Leu
165 170 175

Ile Ser Gly Gly Asp Gly Leu Leu Ile Asn Asp Gln Ile Leu Glu Tyr
180 185 190

Ile Leu Lys Glu Leu Arg Ser Ile Pro His Leu Glu Val Ile Arg Ile
195 200 205

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

Cys Glu Ile Leu Lys Lys Tyr His Pro Val Trp Leu Asn Thr His Phe
225 230 235 240

Asn Thr Ser Ile Glu Met Thr Glu Glu Ser Val Glu Ala Cys Glu Lys
245 250 255

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
290 295 300

-103-

Glu Gly Ile Gly His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
 305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
 325 330 335

Val Val Asp Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
 340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
 385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Phe
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Asp Gln
 435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 62

<211> 471

<212> PRT

<213> Artifical Sequence

<400> 62

Met Lys Asn Lys Trp Tyr Lys Pro Lys Arg His Trp Lys Glu Ile Glu
 1 5 10 15

Leu Trp Lys Asp Val Pro Glu Glu Lys Trp Asn Asp Trp Leu Trp Gln

-104-

20					25					30					
Leu	Thr	His	Thr	Val	Arg	Thr	Leu	Asp	Asp	Leu	Lys	Lys	Val	Ile	Asn
		35					40					45			
Leu	Thr	Glu	Asp	Glu	Glu	Glu	Gly	Val	Arg	Ile	Ser	Thr	Lys	Thr	Ile
		50					55					60			
Pro	Leu	Asn	Ile	Thr	Pro	Tyr	Tyr	Ala	Ser	Leu	Met	Asp	Pro	Asp	Asn
		65					70					75			80
Pro	Arg	Cys	Pro	Val	Arg	Met	Gln	Ser	Val	Pro	Leu	Ser	Glu	Glu	Met
				85					90					95	
His	Lys	Thr	Lys	Tyr	Asp	Met	Glu	Asp	Pro	Leu	His	Glu	Asp	Glu	Asp
			100					105					110		
Ser	Pro	Val	Pro	Gly	Leu	Thr	His	Arg	Tyr	Pro	Asp	Arg	Val	Leu	Phe
			115					120					125		
Leu	Val	Thr	Asn	Gln	Cys	Ser	Val	Tyr	Cys	Arg	His	Cys	Thr	Arg	Arg
			130					135					140		
Arg	Phe	Ser	Gly	Gln	Ile	Gly	Met	Gly	Val	Pro	Lys	Lys	Gln	Leu	Asp
							150						155		160
Ala	Ala	Ile	Ala	Tyr	Ile	Arg	Glu	Thr	Pro	Glu	Ile	Arg	Asp	Cys	Leu
				165					170					175	
Ile	Ser	Gly	Gly	Asp	Gly	Leu	Leu	Ile	Asn	Asp	Gln	Ile	Leu	Glu	Tyr
			180					185					190		
Ile	Leu	Lys	Glu	Leu	Arg	Ser	Ile	Pro	His	Leu	Glu	Val	Ile	Arg	Ile
			195					200					205		
Gly	Thr	Arg	Ala	Pro	Val	Val	Phe	Pro	Gln	Arg	Ile	Thr	Asp	His	Leu
							215						220		
Cys	Glu	Ile	Leu	Lys	Lys	Tyr	His	Pro	Val	Trp	Leu	Asn	Thr	His	Phe
							230						235		240
Asn	Thr	Ser	Ile	Glu	Met	Thr	Glu	Glu	Ser	Val	Glu	Ala	Cys	Glu	Lys
				245					250					255	

-105-

Leu Val Asn Ala Gly Val Pro Val Gly Asn Gln Ala Val Val Leu Ala
 260 265 270

Gly Ile Asn Asp Ser Val Pro Ile Met Lys Lys Leu Met His Asp Leu
 275 280 285

Val Lys Ile Arg Val Arg Pro Tyr Tyr Ile Tyr Gln Cys Asp Leu Ser
 290 295 300

Glu Gly Ile Arg His Phe Arg Ala Pro Val Ser Lys Gly Leu Glu Ile
 305 310 315 320

Ile Glu Gly Leu Arg Gly His Thr Ser Gly Tyr Ala Val Pro Thr Phe
 325 330 335

Val Val His Ala Pro Gly Gly Gly Gly Lys Ile Ala Leu Gln Pro Asn
 340 345 350

Tyr Val Leu Ser Gln Ser Pro Asp Lys Val Ile Leu Arg Asn Phe Glu
 355 360 365

Gly Val Ile Thr Ser Tyr Pro Glu Pro Glu Asn Tyr Ile Pro Asn Gln
 370 375 380

Ala Asp Ala Tyr Phe Glu Ser Val Phe Pro Glu Thr Ala Asp Lys Lys
 385 390 395 400

Glu Pro Ile Gly Leu Ser Ala Ile Phe Ala Asp Lys Glu Val Ser Ser
 405 410 415

Thr Pro Glu Asn Val Asp Arg Ile Lys Arg Arg Glu Ala Tyr Ile Ala
 420 425 430

Asn Pro Glu His Glu Thr Leu Lys Asp Arg Arg Glu Lys Arg Gly Gln
 435 440 445

Leu Lys Glu Lys Lys Phe Leu Ala Gln Gln Lys Lys Gln Lys Glu Thr
 450 455 460

Glu Cys Gly Gly Asp Ser Ser
 465 470

<210> 63

-106-

<211> 49
<212> DNA
<213> artificial sequence

<220>
<223> Bacillus specific primer

<220>
<221> misc_feature
<223> Forward primer

<400> 63
ccagcctggc cataaggaga tatacatatg aaaaacaaat ggtataaac

49

<210> 64
<211> 50
<212> DNA
<213> artificial sequence

<220>
<223> Bacillus specific primer

<220>
<221> misc_feature
<223> Reverse primer

<400> 64
atggtgatgg tgatggtggc cagtttggcc ttatgaagaa tcccctccgc

50