



(51) International Patent Classification:

C12N 9/22 (2006.01)

C12N 9/78 (2006.01)

A61K 38/43 (2006.01)

C12N 15/62 (2006.01)

C07K 14/31 (2006.01)

TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(21) International Application Number:

PCT/US2021/041809

(22) International Filing Date:

15 July 2021 (15.07.2021)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/052,175

15 July 2020 (15.07.2020)

US

(71) Applicant: LIFEEDIT THERAPEUTICS, INC.

[US/US]; 104 T. W. Alexander Drive, Building 20, Research Triangle Park, North Carolina 27709 (US).

(72) Inventors: BOWEN, Tyson D.;

137 Waltons Creek Road,

Morrisville, North Carolina 27560 (US). CRAWLEY,

Alexandra Briner; 1220 Scott Pl., Cary, North Carolina

27511 (US). ELICH, Tedd D.; 5 Alumwood Place,

Durham, North Carolina 27705 (US).

(74) Agent: SHERK, Andrea B.;

Womble Bond Dickinson

(US) LLP, P.O. Box 7037, Atlanta, Georgia 30357-0037

(US).

(81) Designated States (unless otherwise indicated, for every

kind of national protection available): AE, AG, AL, AM,

AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ,

CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO,

DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,

HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN,

KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD,

ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO,

NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW,

SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN,

TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every

kind of regional protection available): ARIPO (BW, GH,

GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ,

UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,

TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,

EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,

MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,

Declarations under Rule 4.17:

— as to applicant's entitlement to apply for and be granted a

patent (Rule 4.17(ii))

— as to the applicant's entitlement to claim the priority of the

earlier application (Rule 4.17(iii))

— of inventorship (Rule 4.17(iv))

Published:

— with international search report (Art. 21(3))

— before the expiration of the time limit for amending the

claims and to be republished in the event of receipt of

amendments (Rule 48.2(h))

— with sequence listing part of description (Rule 5.2(a))

(54) Title: URACIL STABILIZING PROTEINS AND ACTIVE FRAGMENTS AND VARIANTS THEREOF AND METHODS OF USE

(57) Abstract: Compositions and methods comprising uracil stabilizing polypeptides for targeted editing of nucleic acids are provided. Compositions comprise uracil stabilizing polypeptides. Also provided are fusion proteins comprising i) a DNA-binding polypeptide; ii) a deaminase; and iii) a uracil stabilizing polypeptide (USP). The fusion proteins include RNA-guided nucleases fused to deaminases and further fused to a USP, optionally in complex with guide RNAs. Compositions also include nucleic acid molecules encoding the USPs or the fusion proteins. Vectors and host cells comprising the nucleic acid molecules encoding the USPs or the fusion proteins are also provided.



URACIL STABILIZING PROTEINS AND ACTIVE FRAGMENTS AND VARIANTS THEREOF AND METHODS OF USE

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority to U.S. Provisional Application No. 63/052,175, filed July 15, 2020, which is incorporated by reference herein in its entirety.

STATEMENT REGARDING THE SEQUENCE LISTING

The Sequence Listing associated with this application is provided in ASCII format in lieu of a paper copy, and is hereby incorporated by reference into the specification. The ASCII copy named L103438_1220WO_0106_5_SL.txt is 658,586 bytes in size, was created on July 14, 2021, and is being submitted electronically via EFS-Web.

FIELD OF THE INVENTION

The present invention relates to the field of molecular biology and gene editing.

BACKGROUND OF THE INVENTION

Targeted genome editing or modification is rapidly becoming an important tool for basic and applied research. Initial methods involved engineering nucleases such as meganucleases, zinc finger fusion proteins or TALENs, requiring the generation of chimeric nucleases with engineered, programmable, sequence-specific DNA-binding domains specific for each particular target sequence. RNA-guided nucleases (RGNs), such as the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-associated (Cas) proteins of the CRISPR-Cas bacterial system, allow for the targeting of specific sequences by complexing the nucleases with guide RNA that specifically hybridizes with a particular target sequence. Producing target-specific guide RNAs is less costly and more efficient than generating chimeric nucleases for each target sequence. Such RNA-guided nucleases can be used to edit genomes through the introduction of a sequence-specific, double-stranded break that is repaired via error-prone non-homologous end-joining (NHEJ) to introduce a mutation at a specific genomic location.

Additionally, RGNs are useful for targeted DNA editing approaches. Targeted editing of nucleic acid sequences, for example targeted cleavage, to allow for introduction of a specific modification into genomic DNA, enables a highly nuanced approach to studying gene function and gene expression. Such targeted editing also may be deployed for targeting genetic diseases in humans or for introducing agronomically beneficial mutations in the genomes of crop plants. The development of genome editing tools provides new approaches to gene editing-based mammalian therapeutics and agrobiotechnology.

BRIEF SUMMARY OF THE INVENTION

Compositions and methods for modifying a target DNA molecule are provided. The compositions find use in modifying a target DNA molecule of interest. Compositions provided comprise uracil stabilizing polypeptides. Also provided are fusion proteins comprising a DNA-binding polypeptide, a deaminase polypeptide, and a uracil stabilizing polypeptide. Compositions provided also include nucleic acid molecules encoding the uracil stabilizing polypeptides or the fusion proteins, and vectors and host cells comprising the nucleic acid molecules. The methods disclosed herein are drawn to binding a target sequence of interest within a target DNA molecule of interest and modifying the target DNA molecule of interest.

DETAILED DESCRIPTION

Many modifications and other embodiments of the inventions set forth herein will come to mind to one skilled in the art to which these inventions pertain having the benefit of the teachings presented in the foregoing descriptions. Therefore, it is to be understood that the inventions are not to be limited to the specific embodiments disclosed and that modifications and other embodiments are intended to be included within the scope of the appended claims. Although specific terms are employed herein, they are used in a generic and descriptive sense only and not for purposes of limitation.

I. Overview

This disclosure provides uracil stabilizing polypeptides (USPs), which stabilize uracil residues in a DNA molecule, and nucleic acid molecules encoding the same.

Targeted nucleobase editing, also referred to as base editing, was developed by Komor *et al.* in 2016 using a cytosine deaminase (rAPOBEC1) operably linked to a modified RNA guided nuclease (*Sp*Cas9) (*Nature* 533: 420-424). In the system described by Komor *et al.*, the guide RNA guides the rAPOBEC1-Cas9 fusion protein to the target DNA sequence, where the rAPOBEC1 deaminates a target cytosine (C) to a uracil (U), which has the base-pairing properties of thymine (T). Using this system, targeted C>T mutations could be introduced into a DNA molecule.

A major drawback for base editing using the rAPOBEC1-Cas9 fusion *in vivo* was that cellular Uracil DNA Glycosylase (UDG) recognized the U:G heteroduplex DNA and catalyzed the removal of uracil from the DNA to leave an abasic site, thereby initiating base-excision repair with a reversion of the U:G pair to a C:G pair as the most common outcome, although indel (insertion or deletion) formation was also observed. By incorporating a Uracil DNA Glycosylase Inhibitor (UGI) onto the rAPOBEC1-Cas9 fusion protein, the uracil stayed present long enough for replication to occur and introduce the desired C>T mutation.

The present invention finds that by stabilizing the uracil created by the deaminated cytosine, the creation of the abasic site can be prevented and the desired C>T mutation is more likely to be introduced.

This was achieved by the identification of Uracil Stabilizing Proteins (also referred to as Uracil Stabilizing Polypeptides, or USPs).

In some embodiments, the USP is provided as part of a fusion protein that comprises a DNA-binding polypeptide, a deaminase polypeptide, and a uracil stabilizing polypeptide. In some embodiments, the DNA-binding polypeptide is or is derived from a meganuclease, zinc finger fusion protein, or TALEN. In some embodiments, the DNA-binding polypeptide is an RNA-guided nuclease, such as a Cas9 polypeptide, that binds to a guide RNA (also referred to as gRNA), which, in turn, binds a target nucleic acid sequence via strand hybridization. In other embodiments, the USP is provided alone.

In some embodiments, the deaminase polypeptide may be a deaminase domain that can deaminate a nucleobase, such as, for example, cytidine. The deamination of a nucleobase by a deaminase can lead to a point mutation at the respective residue, which is referred to herein as "nucleic acid editing", or "base editing". Fusion proteins comprising an RNA-guided nuclease (RGN) polypeptide and a deaminase polypeptide can thus be used for the targeted editing of nucleic acid sequences.

Such fusion proteins are useful for targeted editing of DNA *in vitro*, *e.g.*, for the generation of mutant cells. These mutant cells may be in plants or animals. Such fusion proteins may also be useful for the introduction of targeted mutations, *e.g.*, for the correction of genetic defects in mammalian cells *ex vivo*, *e.g.*, in cells obtained from a subject that are subsequently re-introduced into the same or another subject; and for the introduction of targeted mutations *in vivo*, *e.g.*, the correction of genetic defects or the introduction of deactivating mutations in disease-associated genes in a mammalian subject. Such fusion proteins may also be useful for the introduction of targeted mutations in plant cells, *e.g.*, for the introduction of beneficial or agronomically valuable traits or alleles.

The terms "protein," "peptide," and "polypeptide" are used interchangeably herein, and refer to a polymer of amino acid residues linked together by peptide (amide) bonds. The terms refer to a protein, peptide, or polypeptide of any size, structure, or function. Typically, a protein, peptide, or polypeptide will be at least three amino acids long. A protein, peptide, or polypeptide may refer to an individual protein or a collection of proteins. One or more of the amino acids in a protein, peptide, or polypeptide may be modified, for example, by the addition of a chemical entity such as a carbohydrate group, a hydroxyl group, a phosphate group, a farnesyl group, an isofarnesyl group, a fatty acid group, a linker for conjugation, functionalization, or other modification, etc. A protein, peptide, or polypeptide may also be a single molecule or may be a multi-molecular complex. A protein, peptide, or polypeptide may be just a fragment of a naturally occurring protein or peptide. A protein, peptide, or polypeptide may be naturally occurring, recombinant, or synthetic, or any combination thereof.

Any of the proteins provided herein may be produced by any method known in the art. For example, the proteins provided herein may be produced via recombinant protein expression and purification, which is especially suited for fusion proteins comprising a peptide linker. Methods for recombinant protein expression and purification are well known, and include those described by Green and Sambrook, *Molecular*

Cloning: A Laboratory Manual (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)), the entire contents of which are incorporated herein by reference.

II. *Uracil Stabilizing Proteins*

Novel uracil-stabilizing polypeptides (USPs) are presently disclosed and set forth as SEQ ID NOs: 1-16. The USPs described herein are useful in applications where stabilizing a uracil in a DNA molecule is desired.

As used herein, the terms “uracil stabilizing protein,” “uracil stabilizing polypeptide,” and “USPs” refer to a polypeptide having uracil stabilizing activity. As used herein, the term “uracil stabilizing activity” refers to the ability of a molecule (e.g., polypeptide) to increase the mutation rate of at least one cytidine, deoxycytidine, or cytosine to a thymidine, deoxythymidine, or thymine in a nucleic acid molecule by a deaminase compared to the mutation rate by the deaminase in the absence of the molecule (e.g., uracil stabilizing polypeptide). Without being bound by a theory or mechanism of action, it is believed that the presently disclosed USPs may function by maintaining the presence of uracil in single-stranded DNA that has been generated through the deamination of a cytidine, deoxycytidine, or cytosine base for a sufficient period of time to allow for replication to occur and introduce the desired C>T mutation. Uracil stabilizing activity may occur through inhibition of uracil DNA glycosylase, the base excision repair pathway, or mismatch repair mechanisms.

In some embodiments, the presently disclosed USPs or active variants or fragments thereof that retain uracil stabilizing activity increase the mutation rate of at least one cytidine, deoxycytidine, or cytosine to a thymidine, deoxythymidine, or thymine in a nucleic acid molecule by a deaminase by at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 100%, at least 150%, at least 200%, or more compared to the mutation rate by a deaminase in the absence of the USP. Conversely, the mutation rate of at least one cytidine, deoxycytidine, or cytosine to any nucleobase other than thymidine, deoxythymidine, or thymine (i.e., guanosine, deoxyguanosine, guanine, adenosine, deoxyadenosine, adenine) in a nucleic acid molecule by a deaminase is reduced by the presently disclosed USPs or active variants or fragments thereof by at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or more compared to the mutation rate by a deaminase in the absence of the USP. An increase or decrease in the mutation rate of a cytidine, deoxycytidine, or cytosine to another nucleobase can be measured by comparing the rate of mutation of a particular deaminase to a particular nucleobase in the presence or absence of the USP. In those embodiments wherein the deaminase has been targeted to a specific region of a nucleic acid molecule via fusion with a DNA-binding polypeptide, the mutation rate of cytidines, deoxycytidines, or cytosines within or adjacent to the target sequence to which

the DNA-binding polypeptide binds can be measured using any method known in the art, including polymerase chain reaction (PCR), restriction fragment length polymorphism (RFLP), or DNA sequencing.

The presently disclosed novel USPs or active variants or fragments thereof that retain uracil stabilizing activity may be introduced into the cell as part of a deaminase-DNA-binding polypeptide fusion, and/or may be co-expressed with a DNA-binding polypeptide-deaminase fusion or with a DNA-binding polypeptide-deaminase-USP fusion, to increase the efficiency of introducing the desired C>T mutation in a target DNA molecule. The presently disclosed USPs retaining uracil stabilizing activity have the amino acid sequence of any of SEQ ID NOs: 1-16 or a variant or fragment thereof. In some embodiments, the USP has an amino acid sequence having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identity to the amino acid sequence of any of SEQ ID NOs: 1-16. In particular embodiments, the USP comprises an amino acid sequence having at least 80% sequence identity to any one of SEQ ID NOs: 1, 2, 4, 5, and 7-15. In other embodiments, the USP comprises an amino acid sequence having at least 81% sequence identity to SEQ ID NO: 3 or 16. In still other embodiments, the USP comprises an amino acid sequence having at least 82% sequence identity to SEQ ID NO: 6.

III. Fusion Proteins

Some aspects of this disclosure provide fusion proteins that comprise a DNA-binding polypeptide and a deaminase polypeptide, and in some embodiments, a USP polypeptide. Such fusion proteins are useful for targeted editing of DNA *in vitro*, *ex vivo*, or *in vivo*.

The term "fusion protein" as used herein refers to a hybrid polypeptide which comprises protein domains from at least two different proteins. A fusion protein may comprise different domains, for example, a DNA-binding domain and a deaminase. In some embodiments, a fusion protein is in a complex with, or is in association with, a nucleic acid, *e.g.*, RNA.

The deaminase polypeptide comprises a deaminase domain that can deaminate a nucleobase, such as, for example, cytidine. The deamination of a nucleobase by a deaminase can lead to a point mutation at the respective residue, which is referred to herein as "nucleic acid editing" or "base editing". Fusion proteins comprising an RGN polypeptide variant or domain and a deaminase domain can thus be used for the targeted editing of nucleic acid sequences. In some embodiments, a deaminase comprises an amino acid sequence at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or more identical to any one of SEQ ID NOs 47, 48 and 76-94. In some embodiments, a deaminase comprises an amino acid sequence at least 80% identical to any one of SEQ ID NOs 47, 48 and 76-94. In some embodiments, a deaminase comprises an

amino acid sequence at least 85% identical to any one of SEQ ID NOs 47, 48 and 76-94. In some embodiments, a deaminase comprises an amino acid sequence at least 90% identical to any one of SEQ ID NOs 47, 48 and 76-94. In some embodiments, a deaminase comprises an amino acid sequence at least 95% identical to any one of SEQ ID NOs 47, 48 and 76-94. In other embodiments, a deaminase comprises an amid acid sequence at least 99% identical to any one of SEQ ID NOs 47, 48 and 76-94. In some specific embodiments, a deaminase comprises an amino acid sequence as set forth in any one of SEQ ID NOs 47, 48 and 76-94.

The presently disclosed fusion proteins comprise a DNA-binding polypeptide. As used herein, the term "DNA-binding polypeptide" refers to any polypeptide which is capable of binding to DNA. In certain embodiments, the DNA-binding polypeptide portion of the presently disclosed fusion proteins binds to double-stranded DNA. In particular embodiments, the DNA-binding polypeptide binds to DNA in a sequence-specific manner. As used herein, the terms "sequence-specific" or "sequence-specific manner" refer to the selective interaction with a specific nucleotide sequence.

Two polynucleotide sequences can be considered to be substantially complementary when the two sequences hybridize to each other under stringent conditions. Likewise, a DNA-binding polypeptide is considered to bind to a particular target sequence in a sequence-specific manner if the DNA-binding polypeptide binds to its sequence under stringent conditions. By "stringent conditions" or "stringent hybridization conditions" is intended conditions under which the two polynucleotide sequences (or the polypeptide binds to its specific target sequence) will bind to each other to a detectably greater degree than to other sequences (*e.g.*, at least 2-fold over background). Stringent conditions are sequence-dependent and will be different in different circumstances. Typically, stringent conditions will be those in which the salt concentration is less than about 1.5 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3, and the temperature is at least about 30°C for short sequences (*e.g.*, 10 to 50 nucleotides) and at least about 60°C for long sequences (*e.g.*, greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Exemplary low stringency conditions include hybridization with a buffer solution of 30 to 35% formamide, 1 M NaCl, 1% SDS (sodium dodecyl sulfate) at 37°C, and a wash in 1X to 2X SSC (20X SSC = 3.0 M NaCl/0.3 M trisodium citrate) at 50 to 55°C. Exemplary moderate stringency conditions include hybridization in 40 to 45% formamide, 1.0 M NaCl, 1% SDS at 37°C, and a wash in 0.5X to 1X SSC at 55 to 60°C. Exemplary high stringency conditions include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1X SSC at 60 to 65°C. Optionally, wash buffers may comprise about 0.1% to about 1% SDS. Duration of hybridization is generally less than about 24 hours, usually about 4 to about 12 hours. The duration of the wash time will be at least a length of time sufficient to reach equilibrium.

The T_m is the temperature (under defined ionic strength and pH) at which 50% of a complementary target sequence hybridizes to a perfectly matched sequence. For DNA-DNA hybrids, the T_m can be approximated from the equation of Meinkoth and Wahl (1984) *Anal. Biochem.* 138:267-284: $T_m = 81.5^\circ\text{C} + 16.6 (\log M) + 0.41 (\%GC) - 0.61 (\% \text{ form}) - 500/L$; where M is the molarity of monovalent cations, %GC is the

percentage of guanosine and cytosine nucleotides in the DNA, % form is the percentage of formamide in the hybridization solution, and L is the length of the hybrid in base pairs. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence and its complement at a defined ionic strength and pH. However, severely stringent conditions can utilize a hybridization and/or wash at 1, 2, 3, or 4°C lower than the thermal melting point (T_m); moderately stringent conditions can utilize a hybridization and/or wash at 6, 7, 8, 9, or 10°C lower than the thermal melting point (T_m); low stringency conditions can utilize a hybridization and/or wash at 11, 12, 13, 14, 15, or 20°C lower than the thermal melting point (T_m). Using the equation, hybridization and wash compositions, and desired T_m, those of ordinary skill will understand that variations in the stringency of hybridization and/or wash solutions are inherently described. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology—Hybridization with Nucleic Acid Probes*, Part I, Chapter 2 (Elsevier, New York); and Ausubel et al., eds. (1995) *Current Protocols in Molecular Biology*, Chapter 2 (Greene Publishing and Wiley-Interscience, New York). See Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York).

In certain embodiments, the sequence-specific DNA-binding polypeptide is an RNA-guided, DNA-binding polypeptide (RGDBP). As used herein, the terms “RNA-guided, DNA-binding polypeptide” and “RGDBP” refer to polypeptides capable of binding to DNA through the hybridization of an associated RNA molecule with the target DNA sequence.

In some embodiments, the DNA-binding polypeptide of the fusion protein is a nuclease, such as a sequence-specific nuclease. As used herein, the term “nuclease” refers to an enzyme that catalyzes the cleavage of phosphodiester bonds between nucleotides in a nucleic acid molecule. In some embodiments, the DNA-binding polypeptide is an endonuclease, which is capable of cleaving phosphodiester bonds between nucleotides within a nucleic acid molecule, whereas in other embodiments, the DNA-binding polypeptide is an exonuclease that is capable of cleaving the nucleotides at either end (5' or 3') of a nucleic acid molecule. In some embodiments, the sequence-specific nuclease is selected from the group consisting of a meganuclease, a zinc finger nuclease, a TAL-effector DNA binding domain-nuclease fusion protein (TALEN), and an RNA-guided nuclease (RGN) or variants thereof wherein the nuclease activity has been reduced or inhibited.

As used herein, the term “meganuclease” or “homing endonuclease” refers to endonucleases that bind a recognition site within double-stranded DNA that is 12 to 40 bp in length. Non-limiting examples of meganucleases are those that belong to the LAGLIDADG family that comprise the conserved amino acid motif LAGLIDADG (SEQ ID NO: 75). The term “meganuclease” can refer to a dimeric or single-chain meganuclease.

As used herein, the term “zinc finger nuclease” or “ZFN” refers to a chimeric protein comprising a zinc finger DNA-binding domain and a nuclease domain.

As used herein, the term “TAL-effector DNA binding domain-nuclease fusion protein” or “TALEN” refers to a chimeric protein comprising a TAL effector DNA-binding domain and a nuclease domain.

As used herein, the term “RNA-guided nuclease” or “RGN” refers to an RNA-guided, DNA-binding polypeptide that has nuclease activity. RGNs are considered “RNA-guided” because guide RNAs form a complex with the RNA-guided nucleases to direct the RNA-guided nuclease to bind to a target sequence and in some embodiments, introduce a single-stranded or double-stranded break at the target sequence.

Non-limiting examples of RGNs useful in the presently disclosed compositions and methods include those disclosed in Publication Nos. WO 2020/139783, WO 2019/236566, WO 2021/030344, WO/2021/138247, and Application Nos. PCT/US2021/028843 and PCT/US2021/031794, filed April 23, 2021 and May 11, 2021, respectively, each of which is herein incorporated by reference in its entirety. In some embodiments, a presently disclosed fusion protein comprises an RGN comprising an amino acid sequence at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or more identical to any one of SEQ ID NOs: 40 and 95-142. In some embodiments, a presently disclosed fusion protein comprises an RGN having an amino acid sequence at least 80% identical to any one of SEQ ID NOs: 40 and 95-142. In some embodiments, a presently disclosed fusion protein comprises an RGN having an amino acid sequence at least 85% identical to any one of SEQ ID NOs: 40 and 95-142. In some embodiments, a presently disclosed fusion protein comprises an RGN having an amino acid sequence at least 90% identical to any one of SEQ ID NOs: 40 and 95-142. In some embodiments, a presently disclosed fusion protein comprises an RGN having an amino acid sequence at least 95% identical to any one of SEQ ID NOs: 40 and 95-142. In other embodiments, a presently disclosed fusion protein comprises an RGN having an amino acid sequence at least 99% identical to any one of SEQ ID NOs: 40 and 95-142. In some specific embodiments, a presently disclosed fusion protein comprises an RGN having an amino acid sequence as set forth in any one of SEQ ID NOs: 40 and 95-142.

According to the present invention, an RGN protein that has been mutated to become nuclease-inactive or “dead”, such as for example dCas9, is herein referred to as an RNA-guided, DNA-binding polypeptide. One exemplary suitable nuclease-inactive Cas9 domain is the D10A/H840A Cas9 domain mutant (see, *e.g.*, Qi et al., *Cell*. 2013; 152(5): 1173-83, the entire contents of which are incorporated herein by reference). Additionally, suitable nuclease-inactive Cas9 domains of other known RNA guided nucleases (RGNs) can be determined (for example, a nuclease-inactive variant of the RGN APG08290.1 disclosed in U.S. Patent Publication No. 2019/0367949, the entire contents of which are incorporated herein by reference herein).

The term “RGN polypeptide” encompasses RGN polypeptides that only cleave a single strand of a target nucleotide sequence, which is referred to herein as a nickase. Such RGNs have a single functioning nuclease domain. RGN nickases can be naturally-occurring nickases or can be RGN proteins that naturally cleave both strands of a double-stranded nucleic acid molecule that has been mutated within additional

nuclease domains such that the nuclease activity of these mutated domains is reduced or eliminated, to become a nickase. In some embodiments, the nickase RGN of the fusion protein comprises a D10A mutation (for example nAPG07433.1 (SEQ ID NO: 41)) which renders the RGN capable of cleaving only the non-base edited, target strand (the strand which comprises the PAM and is base paired to a gRNA) of a nucleic acid duplex. In some embodiments, the nickase RGN of the fusion protein comprises a D10A mutation or an equivalent mutation thereof in any one of SEQ ID NOs: 40 and 95-142. In some embodiments, the nickase RGN of the fusion protein comprises a H840A mutation, which renders the RGN capable of cleaving only the base-edited, non-target strand (the strand which does not comprise the PAM and is not base paired to a gRNA) of a nucleic acid duplex. A nickase RGN comprising an H840A mutation, or an equivalent mutation, has an inactivated HNH domain. A nickase RGN comprising a D10A mutation, or an equivalent mutation, has an inactivated RuvC domain. The deaminase acts on the non-target strand. A nickase comprising a D10A mutation, or an equivalent mutation, has an inactive RuvC nuclease domain and is not able to cleave the non-targeted strand of the DNA, i.e., the strand where base editing is desired.

Other additional exemplary suitable nuclease inactive Cas9 domains include, but are not limited to, D10A/D839A/H840A, and D10A/D839A/H840A/N863A mutant domains (See, *e.g.*, Mali et al., *Nature Biotechnology*. 2013; 31(9): 833-838, the entire contents of which are incorporated herein by reference). Additional suitable RGN proteins mutated to be nickases will be apparent to those of skill in the art based on this disclosure and knowledge in the field (such as for example the RGNs disclosed in PCT Publication No. WO 2019/236566) and are within the scope of this disclosure.

In some embodiments the RGN nickase retaining nickase activity comprises an amino acid sequence that is at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to nAPG07433.1 (SEQ ID NO: 41).

Any method known in the art for introducing mutations into an amino acid sequence, such as PCR-mediated mutagenesis and site-directed mutagenesis, can be used for generating nickases or nuclease-dead RGNs. See, *e.g.*, U.S. Publ. No. 2014/0068797 and U.S. Pat. No. 9,790,490; each of which is incorporated by reference in its entirety. RNA-guided nucleases (RGNs) allow for the targeted manipulation of a single site within a genome and are useful in the context of gene targeting for therapeutic and research applications. In a variety of organisms, including mammals, RNA-guided nucleases have been used for genome engineering by stimulating either non-homologous end joining or homologous recombination. RGNs include CRISPR-Cas proteins, which are RNA-guided nucleases directed to the target sequence by a guide RNA (gRNA) as part of a Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) RNA-guided nuclease system, or active variants or fragments thereof.

Some aspects of this disclosure provide fusion proteins that comprise an RNA-guided DNA-binding polypeptide, a deaminase polypeptide, and a USP. In some embodiments, the RNA-guided DNA-binding polypeptide is an RNA-guided nuclease. In further embodiments, the RNA-guided nuclease is a naturally-occurring CRISPR-Cas protein or an active variant or fragment thereof. CRISPR-Cas systems are classified

into Class I or Class II systems. Class II systems comprise a single effector nuclease and include Types II, V, and VI. Each class is subdivided into types (Types I, II, III, IV, V, VI), with some types further divided into subtypes (e.g., Type II-A, Type II-B, Type II-C, Type V-A, Type V-B).

In certain embodiments, the CRISPR-Cas protein is a naturally-occurring Type II CRISPR-Cas protein or an active variant or fragment thereof. As used herein, the term “Type II CRISPR-Cas protein,” “Type II CRISPR-Cas effector protein,” or “Cas9” refers to a CRISPR-Cas effector protein that requires a trans-activating RNA (tracrRNA) and comprises two nuclease domains (RuvC and HNH), each of which is responsible for cleaving a single strand of a double-stranded DNA molecule.

In other embodiments, the CRISPR-Cas protein is a naturally-occurring Type V CRISPR-Cas protein or an active variant or fragment thereof. As used herein, the term “Type V CRISPR-Cas protein,” “Type V CRISPR-Cas effector protein,” or “Cas12” refers to a CRISPR-Cas effector protein that cleaves dsDNA and comprises a single RuvC nuclease domain or a split-RuvC nuclease domain and lacks an HNH domain (Zetsche et al 2015, *Cell* doi:10.1016/j.cell.2015.09.038; Shmakov et al 2017, *Nat Rev Microbiol* doi:10.1038/nrmicro.2016.184; Yan et al 2018, *Science* doi:10.1126/science.aav7271; Harrington et al 2018, *Science* doi:10.1126/science.aav4294). It is to be noted that Cas12a is also referred to as Cpf1, and does not require a tracrRNA, although other Type V CRISPR-Cas proteins, such as Cas12b, do require a tracrRNA. Most Type V effectors can also target ssDNA (single-stranded DNA), often without a PAM requirement (Zetsche et al 2015; Yan et al 2018; Harrington et al 2018). The term “Type V CRISPR-Cas protein” encompasses the unique RGNs comprising split RuvC nuclease domains, such as those disclosed in U.S. Provisional Appl. No. 62/955,014 filed December 30, 2019, the contents of which are incorporated by reference in its entirety.

In still other embodiments, the CRISPR-Cas protein is a naturally-occurring Type VI CRISPR-Cas protein or an active variant or fragment thereof. As used herein, the term “Type VI CRISPR-Cas protein,” “Type VI CRISPR-Cas effector protein,” or “Cas13” refers to a CRISPR-Cas effector proteins that do not require a tracrRNA and comprise two HEPN domains that cleave RNA. The term “guide RNA” refers to a nucleotide sequence having sufficient complementarity with a target nucleotide sequence to hybridize with the target sequence and direct sequence-specific binding of an associated RGN to the target nucleotide sequence. For CRISPR-Cas RGNs, the respective guide RNA is one or more RNA molecules (generally, one or two), that can bind to the RGN and guide the RGN to bind to a particular target nucleotide sequence, and in those instances wherein the RGN has nickase or nuclease activity, also cleave the target nucleotide sequence. A guide RNA comprises a CRISPR RNA (crRNA) and in some embodiments, a trans-activating CRISPR RNA (tracrRNA).

A CRISPR RNA comprises a spacer sequence and a CRISPR repeat sequence. The “spacer sequence” is the nucleotide sequence that directly hybridizes with the target nucleotide sequence of interest. The spacer sequence is engineered to be fully or partially complementary with the target sequence of interest. In various embodiments, the spacer sequence can comprise from about 8 nucleotides to about 30 nucleotides, or more. For example, the spacer sequence can be about 8, about 9, about 10, about 11, about

12, about 13, about 14, about 15, about 16, about 17, about 18, about 19, about 20, about 21, about 22, about 23, about 24, about 25, about 26, about 27, about 28, about 29, about 30, or more nucleotides in length. In some embodiments, the spacer sequence is about 10 to about 26 nucleotides in length, or about 12 to about 30 nucleotides in length. In particular embodiments, the spacer sequence is about 30 nucleotides in length.

5 In some embodiments, the degree of complementarity between a spacer sequence and its corresponding target sequence, when optimally aligned using a suitable alignment algorithm, is about or more than about 50%, about 60%, about 70%, about 75%, about 80%, about 81%, about 82%, about 83%, about 84%, about 85%, about 86%, about 87%, about 88%, about 89%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or more. In particular embodiments, the
10 spacer sequence is free of secondary structure, which can be predicted using any suitable polynucleotide folding algorithm known in the art, including but not limited to mFold (see, *e.g.*, Zuker and Stiegler (1981) *Nucleic Acids Res.* 9:133-148) and RNAfold (see, *e.g.*, Gruber et al. (2008) *Cell* 106(1):23-24).

The CRISPR RNA repeat sequence comprises a nucleotide sequence that forms a structure, either on its own or in concert with a hybridized tracrRNA, that is recognized by the RGN molecule. In various
15 embodiments, the CRISPR RNA repeat sequence can comprise from about 8 nucleotides to about 30 nucleotides, or more. For example, the CRISPR repeat sequence can be about 8, about 9, about 10, about 11, about 12, about 13, about 14, about 15, about 16, about 17, about 18, about 19, about 20, about 21, about 22, about 23, about 24, about 25, about 26, about 27, about 28, about 29, about 30, or more nucleotides in length. In some embodiments, the degree of complementarity between a CRISPR repeat sequence and its
20 corresponding tracrRNA sequence, when optimally aligned using a suitable alignment algorithm, is about or more than about 50%, about 60%, about 70%, about 75%, about 80%, about 81%, about 82%, about 83%, about 84%, about 85%, about 86%, about 87%, about 88%, about 89%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or more.

In some embodiments, the guide RNA further comprises a tracrRNA molecule. A trans-activating
25 CRISPR RNA or tracrRNA molecule comprises a nucleotide sequence comprising a region that has sufficient complementarity to hybridize to a CRISPR repeat sequence of a crRNA, which is referred to herein as the anti-repeat region. In some embodiments, the tracrRNA molecule further comprises a region with secondary structure (*e.g.*, stem-loop) or forms secondary structure upon hybridizing with its corresponding crRNA. In particular embodiments, the region of the tracrRNA that is fully or partially
30 complementary to a CRISPR repeat sequence is at the 5' end of the molecule and the 3' end of the tracrRNA comprises secondary structure. This region of secondary structure generally comprises several hairpin structures, including the nexus hairpin, which is found adjacent to the anti-repeat sequence. There are often terminal hairpins at the 3' end of the tracrRNA that can vary in structure and number, but often comprise a GC-rich Rho-independent transcriptional terminator hairpin followed by a string of Us at the 3' end. See,
35 for example, Briner et al. (2014) *Molecular Cell* 56:333-339, Briner and Barrangou (2016) *Cold Spring Harb Protoc*; doi: 10.1101/pdb.top090902, and U.S. Publication No. 2017/0275648, each of which is herein incorporated by reference in its entirety.

In various embodiments, the anti-repeat region of the tracrRNA that is fully or partially complementary to the CRISPR repeat sequence comprises from about 6 nucleotides to about 30 nucleotides, or more. For example, the region of base pairing between the tracrRNA anti-repeat sequence and the CRISPR repeat sequence can be about 6, about 7, about 8, about 9, about 10, about 11, about 12, about 13, about 14, about 15, about 16, about 17, about 18, about 19, about 20, about 21, about 22, about 23, about 24, about 25, about 26, about 27, about 28, about 29, about 30, or more nucleotides in length. In particular embodiments, the anti-repeat region of the tracrRNA that is fully or partially complementary to a CRISPR repeat sequence is about 10 nucleotides in length. In some embodiments, the degree of complementarity between a CRISPR repeat sequence and its corresponding tracrRNA anti-repeat sequence, when optimally aligned using a suitable alignment algorithm, is about or more than about 50%, about 60%, about 70%, about 75%, about 80%, about 81%, about 82%, about 83%, about 84%, about 85%, about 86%, about 87%, about 88%, about 89%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or more.

In various embodiments, the entire tracrRNA can comprise from about 60 nucleotides to more than about 210 nucleotides. For example, the tracrRNA can be about 60, about 65, about 70, about 75, about 80, about 85, about 90, about 95, about 100, about 105, about 110, about 115, about 120, about 125, about 130, about 135, about 140, about 150, about 160, about 170, about 180, about 190, about 200, about 210 or more nucleotides in length. In particular embodiments, the tracrRNA is about 100 to about 201 nucleotides in length, including about 95, about 96, about 97, about 98, about 99, about 100, about 105, about 106, about 107, about 108, about 109, and about 110 nucleotides in length.

Guide RNAs form a complex with the RNA-guided nucleases to direct the RNA-guided nuclease to bind to a target sequence and introduce a single-stranded or double-stranded break at the target sequence. After the target sequence has been cleaved, the break can be repaired such that the DNA sequence of the target sequence is modified during the repair process. Provided herein are methods for using mutant variants of RNA-guided nucleases, which are either nuclease inactive or nickases, which are linked to deaminases to modify a target sequence in the DNA of host cells. The mutant variants of RNA-guided nucleases in which the nuclease activity is inactivated or significantly reduced may be referred to as RNA-guided, DNA-binding polypeptides, as the polypeptides are capable of binding to, but not necessarily cleaving, a target sequence. RNA-guided nucleases only capable of cleaving a single strand of a double-stranded nucleic acid molecule are referred to herein as nickases.

A target nucleotide sequence is bound by an RGN and hybridizes with the guide RNA associated with the RGN. The target sequence can then be subsequently cleaved by the RGN if the polypeptide possesses nuclease activity, which encompasses activity as a nickase.

The guide RNA can be a single guide RNA or a dual-guide RNA system. A single guide RNA comprises the crRNA and optionally tracrRNA on a single molecule of RNA, whereas a dual-guide RNA system comprises a crRNA and a tracrRNA present on two distinct RNA molecules, hybridized to one another through at least a portion of the CRISPR repeat sequence of the crRNA and at least a portion of the

tracrRNA, which may be fully or partially complementary to the CRISPR repeat sequence of the crRNA. In some of those embodiments wherein the guide RNA is a single guide RNA, the crRNA and optionally tracrRNA are separated by a linker nucleotide sequence.

In general, the linker nucleotide sequence is one that does not include complementary bases in order to avoid the formation of secondary structure within or comprising nucleotides of the linker nucleotide sequence. In some embodiments, the linker nucleotide sequence between the crRNA and tracrRNA is at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, or more nucleotides in length. In particular embodiments, the linker nucleotide sequence of a single guide RNA is at least 4 nucleotides in length.

In certain embodiments, the guide RNA can be introduced into a target cell, organelle, or embryo as an RNA molecule. The guide RNA can be transcribed *in vitro* or chemically synthesized. In other embodiments, a nucleotide sequence encoding the guide RNA is introduced into the cell, organelle, or embryo. In some of these embodiments, the nucleotide sequence encoding the guide RNA is operably linked to a promoter (*e.g.*, an RNA polymerase III promoter). The promoter can be a native promoter or heterologous to the guide RNA-encoding nucleotide sequence.

In various embodiments, the guide RNA can be introduced into a target cell, organelle, or embryo as a ribonucleoprotein complex, as described herein, wherein the guide RNA is bound to an RNA-guided nuclease polypeptide.

The guide RNA directs an associated RNA-guided nuclease to a particular target nucleotide sequence of interest through hybridization of the guide RNA to the target nucleotide sequence. A target nucleotide sequence can comprise DNA, RNA, or a combination of both and can be single-stranded or double-stranded. A target nucleotide sequence can be genomic DNA (*i.e.*, chromosomal DNA), plasmid DNA, or an RNA molecule (*e.g.*, messenger RNA, ribosomal RNA, transfer RNA, micro RNA, small interfering RNA). The target nucleotide sequence can be bound (and in some embodiments, cleaved) by an RNA-guided nuclease *in vitro* or in a cell. The chromosomal sequence targeted by the RGN can be a nuclear, plastid or mitochondrial chromosomal sequence. In some embodiments, the target nucleotide sequence is unique in the target genome.

In some embodiments, the target nucleotide sequence is adjacent to a protospacer adjacent motif (PAM). A PAM is generally within about 1 to about 10 nucleotides from the target nucleotide sequence, including about 1, about 2, about 3, about 4, about 5, about 6, about 7, about 8, about 9, or about 10 nucleotides from the target nucleotide sequence. The PAM can be 5' or 3' of the target sequence. In some embodiments, the PAM is 3' of the target sequence. Generally, the PAM is a consensus sequence of about 2-6 nucleotides, but in particular embodiments, can be 1, 2, 3, 4, 5, 6, 7, 8, 9, or more nucleotides in length.

Upon recognizing its corresponding PAM sequence, the RGN can cleave the target nucleotide sequence at a specific cleavage site. As used herein, a cleavage site is made up of the two particular nucleotides within a target nucleotide sequence between which the nucleotide sequence is cleaved by an RGN. The cleavage site can comprise the 1st and 2nd, 2nd and 3rd, 3rd and 4th, 4th and 5th, 5th and 6th, 7th and

8th, or 8th and 9th nucleotides from the PAM in either the 5' or 3' direction. As RGNs can cleave a target nucleotide sequence resulting in staggered ends, in some embodiments, the cleavage site is defined based on the distance of the two nucleotides from the PAM on the positive (+) strand of the polynucleotide and the distance of the two nucleotides from the PAM on the negative (-) strand of the polynucleotide.

5 RGNs can be used to deliver a fused polypeptide, polynucleotide, or small molecule payload to a particular genomic location. In some embodiments, a nuclease-inactive or a nickase RGN is operably linked to a deaminase and also to a USP of the invention.

As used herein, the term "deaminase" or "deaminase polypeptide" refers to a polypeptide that catalyzes a deamination reaction. The deaminase may be a naturally-occurring deaminase enzyme or an active fragment or variant thereof. In some embodiments, the deaminase is a cytidine deaminase, catalyzing the hydrolytic deamination of cytidine or deoxycytidine to uracil or deoxyuracil, respectively. Cytidine deaminases may work on either DNA or RNA, and typically operate on single-stranded nucleic acid molecules. In preferred embodiments, an RGN which has nickase activity on the target strand nicks the target strand, while the complementary, non-target strand is modified by the deaminase. Cellular DNA-
10 repair machinery may repair the nicked, target strand using the modified non-target strand as a template, thereby introducing a mutation in the DNA.

In some embodiments, the cytidine deaminase is an apolipoprotein B mRNA-editing complex (APOBEC) family deaminase. In some of these embodiments, the deaminase is an APOBEC1 family deaminase. In some embodiments, the cytidine deaminase is an activation-induced cytidine deaminase (AID). In some embodiments, the deaminase is an ACF1/ASE deaminase. In certain embodiments, the
20 deaminase is an adenosine deaminase. In some of these embodiments, the deaminase is an ADAT family deaminase. Additional suitable deaminase enzymes or domains will be apparent to the skilled artisan based on this disclosure.

One exemplary suitable type of deaminase polypeptides are cytidine deaminases, for example, of the APOBEC family. The apolipoprotein B mRNA editing complex (APOBEC) family of cytosine deaminase
25 enzymes encompasses eleven proteins that serve to initiate mutagenesis in a controlled and beneficial manner (Conticello *et al.*, 2008. *Genome Biology*, 9(6): 229). One family member, activation-induced cytidine deaminase (AID), is responsible for the maturation of antibodies by converting cytosines in ssDNA to uracils in a transcription-dependent, strand-biased fashion (Reynaud *et al.*, 2003. *Nature Immunology*,
30 4(7): 631-638). The apolipoprotein B editing complex 3 (APOBEC3) enzyme provides protection to human cells against a certain HIV-1 strain via the deamination of cytosines in reverse-transcribed viral ssDNA (Bhagwat *et al.*, 2004. *DNA Repair (Amst)*, 3(1): 85-9). These proteins all require a Zn²⁺-coordinating motif (HisX- Glu-X₂₃₋₂₆-Pro-Cys-X₂₋₄-Cys) and bound water molecule for catalytic activity. The Glu residue acts to activate the water molecule to a zinc hydroxide for nucleophilic attack in the deamination reaction. Each
35 family member preferentially deaminates at its own particular "hotspot", ranging from WRC (wherein W is A or T and R is A or G) for hAID, to TTC for hAPOBEC3F (Navaratnam *et al.*, 2006. *Intl J Hematol* 83(3): 195-200). A recent crystal structure of the catalytic domain of APOBEC3G revealed a secondary structure

comprised of a five-stranded β -sheet core flanked by six α -helices, which is believed to be conserved across the entire family (Holden *et al.*, 2008. *Nature* 456(7218): 121-124). The active center loops have been shown to be responsible for both ssDNA binding and in determining "hotspot" identity (Chelico *et al.*, 2009. *J Biol Chem* 284(41): 27761-27765). Overexpression of these enzymes has been linked to genomic instability and cancer, thus highlighting the importance of sequence-specific targeting (Pham *et al.*, 2005. *Biochem* 44(8): 2703-2715). In some embodiments, the deaminase polypeptide may be a deaminase polypeptide that can deaminate a cytidine to yield a uracil. The deamination of a nucleobase by a deaminase can lead to a point mutation at the respective residue, thereby modifying the DNA molecule. This act of modification is also referred to herein as nucleic acid editing, or base editing. Fusion proteins comprising a Cas9 variant or domain, a deaminase domain, and a USP domain can thus be used for the targeted editing of nucleic acid sequences.

In some embodiments, a nuclease inactive RGN or nickase RGN fused to a deaminase and an USP of the invention can be targeted to a particular location of a nucleic acid molecule (i.e., target nucleic acid molecule), which in some embodiments is a particular genomic locus, to alter the expression of a desired sequence. In some embodiments, the binding of a fusion protein to a target sequence results in deamination of a nucleotide base, resulting in conversion from one nucleotide base to another. In some embodiments, the binding of this fusion protein to a target sequence results in deamination of a nucleotide base adjacent to the target sequence. The nucleotide base adjacent to the target sequence that is deaminated and mutated using the presently disclosed compositions and methods may be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 base pairs from the 5' or 3' end of the target sequence (bound by the gRNA) within the target nucleic acid molecule. Some aspects of this disclosure provide fusion proteins comprising (i) a nuclease-inactive or nickase RGN polypeptide; (ii) a deaminase polypeptide; and (iii) a uracil stabilizing polypeptide.

The instant disclosure provides fusion proteins of various configurations. In some embodiments, the deaminase polypeptide is fused to the N-terminus of the RGN polypeptide. In some embodiments, the deaminase polypeptide is fused to the C-terminus of the RGN polypeptide.

In some embodiments, the USP domain, deaminase domain, and RNA-guided, DNA-binding polypeptide are fused to each other via a linker. Various linker lengths and flexibilities between the three functional domains of the fusion protein can be employed (*e.g.*, ranging from very flexible linkers of the form (GGGGS)_n and (G)_n to more rigid linkers of the form (EAAAK)_n and (XP)_n in order to achieve the optimal length for deaminase activity for the specific applications. The term "linker," as used herein, refers to a chemical group or a molecule linking two molecules or moieties, *e.g.*, a binding domain and a cleavage domain of a nuclease. In some embodiments, a linker joins an RNA guided nuclease and a deaminase. In some embodiments, a linker joins a dCas9 and a deaminase. Typically, the linker is positioned between, or flanked by, two groups, molecules, or other moieties and connected to each one via a covalent bond, thus connecting the two. In some embodiments, the linker is an amino acid or a plurality of amino acids (*e.g.*, a peptide or protein). In some embodiments, the linker is an organic molecule, group, polymer, or chemical

moiety. In some embodiments, the linker is 5-100 amino acids in length, for example, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30-35, 35-40, 40-45, 45-50, 50-60, 60-70, 70-80, 80-90, 90-100, 100-150, or 150-200 amino acids in length. Longer or shorter linkers are also contemplated.

In some embodiments, the linker comprises a (GGGGS)_n, a (G)_n, an (EAAAK)_n, or an (XP)_n motif, or a combination of any of these, wherein n is independently an integer between 1 and 30. In some embodiments, n is independently 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30, or, if more than one linker or more than one linker motif is present, any combination thereof. Additional suitable linker motifs and linker configurations will be apparent to those of skill in the art. In some embodiments, suitable linker motifs and configurations include those described in Chen *et al.*, 2013 (*Adv Drug Deliv Rev.* 65(10):1357-69, the entire contents of which are incorporated herein by reference). Additional suitable linker sequences will be apparent to those of skill in the art.

In some embodiments, the general architecture of exemplary fusion proteins provided herein comprises the structure: [NH₂]-[deaminase]-[RGN polypeptide]-[USP]-[COOH]; [NH₂]-[USP]-[deaminase]-[RGN polypeptide]-[COOH]; [NH₂]-[USP]-[RGN polypeptide]-[deaminase]-[COOH]; [NH₂]-[RGN polypeptide]-[deaminase]-[USP]-[COOH]; [NH₂]-[RGN polypeptide]-[USP polypeptide]-[deaminase polypeptide]-[COOH]; or [NH₂]-[deaminase polypeptide]-[USP polypeptide]-[RGN polypeptide]-[COOH], wherein NH₂ is the N-terminus of the fusion protein, and COOH is the C-terminus of the fusion protein.

Some aspects of this disclosure provide deaminase-RGN-USP fusion proteins, deaminase-nuclease inactive RGN-USP fusion proteins and deaminase-nickase RGN-USP fusion proteins, with increased C→T nucleobase editing efficiency as compared to a similar fusion protein that does not comprise a USP domain.

In some embodiments, the fusion protein comprises the structure: [NH₂]-[deaminase]-[nuclease-inactive RGN]-[USP]-[COOH]; [NH₂]-[deaminase polypeptide]-[USP]-[nuclease-inactive RGN]-[COOH]; [NH₂]-[USP]-[deaminase]-[nuclease-inactive RGN]-[COOH]; [NH₂]-[USP]-[nuclease-inactive RGN]-[deaminase]-[COOH]; [NH₂]-[nuclease-inactive RGN]-[deaminase]-[USP]-[COOH]; or [NH₂]-[nuclease-inactive RGN]-[USP]-[deaminase]-[COOH]. It should be understood that “nuclease-inactive RGN” represents any RGN, including any CRISPR-Cas protein, which has been mutated to be nuclease-inactive. It should also be understood that “USP” represents one or more USP polypeptides.

In other embodiments, the fusion protein comprises the structure: [NH₂]-[deaminase]-[RGN nickase]-[USP]-[COOH]; [NH₂]-[deaminase]-[USP]-[RGN nickase]-[COOH]; [NH₂]-[USP]-[deaminase]-[RGN nickase]-[COOH]; [NH₂]-[USP]-[RGN nickase]-[deaminase]-[COOH]; [NH₂]-[RGN nickase]-[deaminase]-[USP]-[COOH]; or [NH₂]-[RGN nickase]-[USP]-[deaminase]-[COOH]. It should be understood that “RGN nickase” represents any RGN, including any CRISPR-Cas protein, which has been mutated to be active as a nickase. It should also be understood that “USP” represents one or more USP polypeptides.

In some embodiments, the fusion protein comprises a cytidine deaminase having at least 80% sequence identity to any one of SEQ ID NOs: 47, 48 and 76-94, an RGN (or nickase thereof) having at least

80% sequence identity to any one of SEQ ID NOs: 40, 41, 95-142, and a USP having at least 80% sequence identity to any one of SEQ ID NOs: 1-16.

In some embodiments, the fusion protein comprises a cytidine deaminase having at least 85% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, an RGN (or nickase thereof) having at least 85% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, and a USP having at least 85% sequence identity to any one of SEQ ID NOs: 1-16.

In some embodiments, the fusion protein comprises a cytidine deaminase having at least 90% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, an RGN (or nickase thereof) having at least 90% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, and a USP having at least 90% sequence identity to any one of SEQ ID NOs: 1-16.

In some embodiments, the fusion protein comprises a cytidine deaminase having at least 95% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, an RGN (or nickase thereof) having at least 95% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, and a USP having at least 95% sequence identity to any one of SEQ ID NOs: 1-16.

In some embodiments, the fusion protein comprises a cytidine deaminase having the amino acid sequence set forth in any one of SEQ ID NOs: 47, 48, and 76-94, an RGN (or nickase thereof) having the amino acid sequence set forth in any one of SEQ ID NOs: 40, 41 and 95-142, and a USP having the amino acid sequence set forth in any one of SEQ ID NOs: 1-16.

In some embodiments, the "-" used in the general architecture above indicates the presence of an optional linker sequence. In some embodiments, the fusion proteins provided herein do not comprise a linker sequence. In some embodiments, at least one of the optional linker sequences are present.

Other exemplary features that may be present are localization sequences, such as nuclear localization sequences, cytoplasmic localization sequences, export sequences, such as nuclear export sequences, or other localization sequences, as well as sequence tags that are useful for solubilization, purification or detection of the fusion proteins. Suitable localization signal sequences and sequences of protein tags that are provided herein, and include, but are not limited to, biotin carboxylase carrier protein (BCCP) tags, myc-tags, calmodulin-tags, FLAG-tags, hemagglutinin (HA)-tags, polyhistidine tags, also referred to as histidine tags or His-tags, maltose binding protein (MBP)-tags, nus-tags, glutathione-S-transferase (GST)-tags, green fluorescent protein (GFP)-tags, thioredoxin- tags, S-tags, Softags (*e.g.*, Softag 1, Softag 3), streptags, biotin ligase tags, FAsH tags, V5 tags, and SBP-tags. Additional suitable sequences will be apparent to those of skill in the art.

In certain embodiments, the presently disclosed fusion proteins comprise at least one cell-penetrating domain that facilitates cellular uptake of the fusion protein. Cell-penetrating domains are known in the art and generally comprise stretches of positively charged amino acid residues (*i.e.*, polycationic cell-penetrating domains), alternating polar amino acid residues and non-polar amino acid residues (*i.e.*, amphipathic cell-penetrating domains), or hydrophobic amino acid residues (*i.e.*, hydrophobic cell-penetrating domains) (*see, e.g.*, Milletti F. (2012) *Drug Discov Today* 17:850-860). A non-limiting example

of a cell-penetrating domain is the trans-activating transcriptional activator (TAT) from the human immunodeficiency virus 1.

In some embodiments, USPs or fusion proteins provided herein further comprise a nuclear localization sequence (NLS). The nuclear localization signal, plastid localization signal, mitochondrial localization signal, dual-targeting localization signal, and/or cell-penetrating domain can be located at the amino-terminus (N-terminus), the carboxyl-terminus (C-terminus), or in an internal location of the fusion protein.

In some embodiments, the NLS is fused to the N-terminus of the fusion protein or USP. In some embodiments, the NLS is fused to the C-terminus of the fusion protein or USP. In some embodiments, the NLS is fused to the N-terminus of the USP of the fusion protein. In some embodiments, the NLS is fused to the C-terminus of the USP of the fusion protein. In some embodiments, the NLS is fused to the N-terminus of the RGN polypeptide of the fusion protein. In some embodiments, the NLS is fused to the C-terminus of the RGN polypeptide of the fusion protein. In some embodiments, the NLS is fused to the N-terminus of the deaminase polypeptide of the fusion protein. In some embodiments, the NLS is fused to the C-terminus of the deaminase polypeptide of the fusion protein. In some embodiments, the NLS is fused to the fusion protein or UPS via one or more linkers. In some embodiments, the NLS is fused to the fusion protein or UPS without a linker. In some embodiments, the NLS comprises an amino acid sequence of any one of the NLS sequences provided or referenced herein. In some embodiments, the NLS comprises an amino acid sequence as set forth in SEQ ID NO: 42 or SEQ ID NO: 45.

In some embodiments, fusion proteins as provided herein comprise the full-length sequence of a uracil stabilizing protein, *e.g.*, any one of SEQ ID NO: 1-16. In other embodiments, however, fusion proteins as provided herein do not comprise a full-length sequence of a USP, but only a fragment thereof. For example, in some embodiments, a fusion protein provided herein further comprises an RNA-guided, DNA-binding domain, a deaminase domain, and an active fragment of a USP.

In some embodiments, a fusion protein of the invention comprises an RGN, a deaminase, and a USP, wherein the USP has an amino acid sequence of at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical to any of SEQ ID NO: 1-16. Examples of such fusion proteins are described in the Examples section here.

In some embodiments, the fusion protein comprises one USP polypeptide. In some embodiments, the fusion protein comprises at least two USP polypeptides, operably linked either directly or via a linker. In some embodiments, the fusion protein comprises one USP polypeptide, and a second USP polypeptide is co-expressed with the fusion protein.

Another embodiment of the invention is a ribonucleoprotein complex comprising the fusion protein and the guide RNA, either as a single guide or as a dual guide RNA (collectively referred to as gRNA).

IV. Nucleotides Encoding Uracil Stabilizing Polypeptides, Fusion Proteins, and/or gRNA

The present disclosure provides polynucleotides encoding the presently disclosed uracil stabilizing polypeptides (SEQ ID NOs: 17-32). The present disclosure further provides polynucleotides encoding for fusion proteins which comprise a deaminase and DNA-binding polypeptide, for example a meganuclease, a zinc finger fusion protein, or a TALEN. The present disclosure further provides polynucleotides encoding for fusion proteins which comprise a USP, a deaminase domain, and an RNA-guided, DNA-binding polypeptide. Such RNA-guided, DNA-binding polypeptide may be an RGN or RGN variant. The protein variant may be nuclease-inactive or a nickase. The RGN may be a CRISPR-Cas protein or active variant or fragment thereof. SEQ ID NOs: 40 and 41 are non-limiting examples of an RGN and a nickase RGN variant, respectively. Examples of CRISPR-Cas nucleases are well-known in the art, and similar corresponding mutations can create mutant variants which are also nickases or are nuclease inactive.

An embodiment of the invention provides a polynucleotide encoding a fusion protein which comprises an RGN, a deaminase, and a USP described herein (SEQ ID NO: 1-16, or a variant thereof). In some embodiments, a second polynucleotide encodes the guide RNA required by the RGN for targeting to the nucleotide sequence of interest. In other embodiments, the guide RNA and the fusion protein are encoded by the same polynucleotide.

The use of the term "polynucleotide" is not intended to limit the present disclosure to polynucleotides comprising DNA. Those of ordinary skill in the art will recognize that polynucleotides can comprise ribonucleotides (RNA) and combinations of ribonucleotides and deoxyribonucleotides. Such deoxyribonucleotides and ribonucleotides include both naturally occurring molecules and synthetic analogues. The polynucleotides disclosed herein also encompass all forms of sequences including, but not limited to, single-stranded forms, double-stranded forms, stem-and-loop structures, and the like.

An embodiment of the invention is a nucleic acid molecule comprising a sequence at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or is 100% identical to any of SEQ ID NOs: 17-32, wherein the nucleic acid molecule encodes a USP having uracil stabilizing activity. The nucleic acid molecule may further comprise a heterologous promoter or terminator. The nucleic acid molecule may encode a fusion protein, where the encoded USP is operably linked to a DNA-binding polypeptide, and/or a deaminase. In some embodiments, the nucleic acid molecule encodes a fusion protein, where the encoded USP is operably linked to an RGN and/or a deaminase.

Nucleic acid molecules comprising a polynucleotide which encodes a USP of the invention can be codon optimized for expression in an organism of interest. A "codon-optimized" coding sequence is a polynucleotide coding sequence having its frequency of codon usage designed to mimic the frequency of preferred codon usage or transcription conditions of a particular host cell. Expression in the particular host cell or organism is enhanced as a result of the alteration of one or more codons at the nucleic acid level such that the translated amino acid sequence is not changed. Nucleic acid molecules can be codon optimized, either wholly or in part. Codon tables and other references providing preference information for a wide range of organisms are available in the art (*see, e.g.,* Campbell and Gowri (1990) *Plant Physiol.* 92:1-11 for

a discussion of plant-preferred codon usage). Methods are available in the art for synthesizing plant-preferred genes. See, for example, U.S. Patent Nos. 5,380,831, and 5,436,391, and Murray *et al.* (1989) *Nucleic Acids Res.* 17:477-498, herein incorporated by reference.

Polynucleotides encoding the USPs, fusion proteins, and/or gRNAs described herein can be provided in expression cassettes for *in vitro* expression or expression in a cell, organelle, embryo, or organism of interest. The cassette will include 5' and 3' regulatory sequences operably linked to a polynucleotide encoding a USP and/or a fusion protein comprising a USP, an RNA-guided DNA-binding polypeptide and a deaminase, and/or gRNA provided herein that allows for expression of the polynucleotide. The cassette may additionally contain at least one additional gene or genetic element to be cotransformed into the organism. Where additional genes or elements are included, the components are operably linked. The term "operably linked" is intended to mean a functional linkage between two or more elements. For example, an operable linkage between a promoter and a coding region of interest (*e.g.*, region coding for a USP, deaminase, RNA-guided DNA-binding polypeptide, and/or gRNA) is a functional link that allows for expression of the coding region of interest. Operably linked elements may be contiguous or non-contiguous. When used to refer to the joining of two protein coding regions, by operably linked is intended that the coding regions are in the same reading frame. Alternatively, the additional gene(s) or element(s) can be provided on multiple expression cassettes. For example, the nucleotide sequence encoding a presently disclosed uracil stabilizing polypeptide, either alone or as a component of a fusion protein, can be present on one expression cassette, whereas the nucleotide sequence encoding a gRNA can be on a separate expression cassette. Another example may have the nucleotide sequence encoding a presently disclosed USP alone on a first expression cassette, a second expression cassette encoding a fusion protein comprising a USP, and a nucleotide sequence encoding a gRNA on third expression cassette. Such an expression cassette is provided with a plurality of restriction sites and/or recombination sites for insertion of the polynucleotides to be under the transcriptional regulation of the regulatory regions. Expression cassettes which comprise a selectable marker gene may also be present.

The expression cassette will include in the 5'-3' direction of transcription, a transcriptional (and, in some embodiments, translational) initiation region (*i.e.*, a promoter), a USP-encoding polynucleotide of the invention, and a transcriptional (and in some embodiments, translational) termination region (*i.e.*, termination region) functional in the organism of interest. The promoters of the invention are capable of directing or driving expression of a coding sequence in a host cell. The regulatory regions (*e.g.*, promoters, transcriptional regulatory regions, and translational termination regions) may be endogenous or heterologous to the host cell or to each other. As used herein, "heterologous" in reference to a sequence is a sequence that originates from a foreign species, or, if from the same species, is substantially modified from its native form in composition and/or genomic locus by deliberate human intervention. As used herein, a chimeric gene comprises a coding sequence operably linked to a transcription initiation region that is heterologous to the coding sequence.

Convenient termination regions are available from the Ti-plasmid of *A. tumefaciens*, such as the octopine synthase and nopaline synthase termination regions. See also Guerineau *et al.* (1991) *Mol. Gen. Genet.* 262:141-144; Proudfoot (1991) *Cell* 64:671-674; Sanfacon *et al.* (1991) *Genes Dev.* 5:141-149; Mogen *et al.* (1990) *Plant Cell* 2:1261-1272; Munroe *et al.* (1990) *Gene* 91:151-158; Ballas *et al.* (1989) *Nucleic Acids Res.* 17:7891-7903; and Joshi *et al.* (1987) *Nucleic Acids Res.* 15:9627-9639.

Additional regulatory signals include, but are not limited to, transcriptional initiation start sites, operators, activators, enhancers, other regulatory elements, ribosomal binding sites, an initiation codon, termination signals, and the like. See, for example, U.S. Pat. Nos. 5,039,523 and 4,853,331; EPO 0480762A2; Sambrook *et al.* (1992) *Molecular Cloning: A Laboratory Manual*, ed. Maniatis *et al.* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.), hereinafter "Sambrook 11"; Davis *et al.*, eds. (1980) *Advanced Bacterial Genetics* (Cold Spring Harbor Laboratory Press), Cold Spring Harbor, N.Y., and the references cited therein.

In preparing the expression cassette, the various DNA fragments may be manipulated, so as to provide for the DNA sequences in the proper orientation and, as appropriate, in the proper reading frame. Toward this end, adapters or linkers may be employed to join the DNA fragments or other manipulations may be involved to provide for convenient restriction sites, removal of superfluous DNA, removal of restriction sites, or the like. For this purpose, *in vitro* mutagenesis, primer repair, restriction, annealing, resubstitutions, *e.g.*, transitions and transversions, may be involved.

A number of promoters can be used in the practice of the invention. The promoters can be selected based on the desired outcome. The nucleic acids can be combined with constitutive, inducible, growth stage-specific, cell type-specific, tissue-preferred, tissue-specific, or other promoters for expression in the organism of interest. See, for example, promoters set forth in WO 99/43838 and in US Patent Nos: 8,575,425; 7,790,846; 8,147,856; 8,586,832; 7,772,369; 7,534,939; 6,072,050; 5,659,026; 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463; 5,608,142; and 6,177,611; herein incorporated by reference.

For expression in plants, constitutive promoters also include CaMV 35S promoter (Odell *et al.* (1985) *Nature* 313:810-812); rice actin (McElroy *et al.* (1990) *Plant Cell* 2:163-171); ubiquitin (Christensen *et al.* (1989) *Plant Mol. Biol.* 12:619-632 and Christensen *et al.* (1992) *Plant Mol. Biol.* 18:675-689); pEMU (Last *et al.* (1991) *Theor. Appl. Genet.* 81:581-588); and MAS (Velten *et al.* (1984) *EMBO J.* 3:2723-2730).

Examples of inducible promoters are the Adh1 promoter which is inducible by hypoxia or cold stress, the Hsp70 promoter which is inducible by heat stress, the PPDK promoter and the pepcarboxylase promoter which are both inducible by light. Also useful are promoters which are chemically inducible, such as the In2-2 promoter which is safener induced (U.S. Pat. No. 5,364,780), the Axig1 promoter which is auxin induced and tapetum specific but also active in callus (PCT US01/22169), the steroid-responsive promoters (see, for example, the ERE promoter which is estrogen induced, and the glucocorticoid-inducible promoter in Schena *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:10421-10425 and McNellis *et al.* (1998) *Plant J.* 14(2):247-257) and tetracycline-inducible and tetracycline-repressible promoters (see, for example,

Gatz *et al.* (1991) *Mol. Gen. Genet.* 227:229-237, and U.S. Pat. Nos. 5,814,618 and 5,789,156), herein incorporated by reference.

Tissue-specific or tissue-preferred promoters can be utilized to target expression of an expression construct within a particular tissue. In certain embodiments, the tissue-specific or tissue-preferred promoters are active in plant tissue. Examples of promoters under developmental control in plants include promoters that initiate transcription preferentially in certain tissues, such as leaves, roots, fruit, seeds, or flowers. A "tissue specific" promoter is a promoter that initiates transcription only in certain tissues. Unlike constitutive expression of genes, tissue-specific expression is the result of several interacting levels of gene regulation. As such, promoters from homologous or closely related plant species can be preferable to use to achieve efficient and reliable expression of transgenes in particular tissues. In some embodiments, the expression comprises a tissue-preferred promoter. A "tissue preferred" promoter is a promoter that initiates transcription preferentially, but not necessarily entirely or solely in certain tissues.

In some embodiments, the nucleic acid molecules encoding a USP described herein comprise a cell type-specific promoter. A "cell type specific" promoter is a promoter that primarily drives expression in certain cell types in one or more organs. Some examples of plant cells in which cell type specific promoters functional in plants may be primarily active include, for example, BETL cells, vascular cells in roots, leaves, stalk cells, and stem cells. The nucleic acid molecules can also include cell type preferred promoters. A "cell type preferred" promoter is a promoter that primarily drives expression mostly, but not necessarily entirely or solely in certain cell types in one or more organs. Some examples of plant cells in which cell type preferred promoters functional in plants may be preferentially active include, for example, BETL cells, vascular cells in roots, leaves, stalk cells, and stem cells.

The nucleic acid sequences encoding the USPs, fusion proteins, and/or gRNAs can be operably linked to a promoter sequence that is recognized by a phage RNA polymerase for example, for *in vitro* mRNA synthesis. In such embodiments, the *in vitro*-transcribed RNA can be purified for use in the methods described herein. For example, the promoter sequence can be a T7, T3, or SP6 promoter sequence or a variation of a T7, T3, or SP6 promoter sequence. In such embodiments, the expressed protein and/or RNAs can be purified for use in the methods of genome modification described herein.

In certain embodiments, the polynucleotide encoding the USP, fusion protein, and/or gRNA also can be linked to a polyadenylation signal (*e.g.*, SV40 polyA signal and other signals functional in plants) and/or at least one transcriptional termination sequence. Additionally, the sequence encoding the deaminase or fusion protein also can be linked to sequence(s) encoding at least one nuclear localization signal, at least one cell-penetrating domain, and/or at least one signal peptide capable of trafficking proteins to particular subcellular locations, as described elsewhere herein.

The polynucleotide encoding the USP, fusion protein, and/or gRNA can be present in a vector or multiple vectors. A "vector" refers to a polynucleotide composition for transferring, delivering, or introducing a nucleic acid into a host cell. Suitable vectors include plasmid vectors, phagemids, cosmids, artificial/mini-chromosomes, transposons, and viral vectors (*e.g.*, lentiviral vectors, adeno-associated viral

vectors, baculoviral vector). The vector can comprise additional expression control sequences (*e.g.*, enhancer sequences, Kozak sequences, polyadenylation sequences, transcriptional termination sequences), selectable marker sequences (*e.g.*, antibiotic resistance genes), origins of replication, and the like. Additional information can be found in "Current Protocols in Molecular Biology" Ausubel *et al.*, John Wiley & Sons, New York, 2003 or "Molecular Cloning: A Laboratory Manual" Sambrook & Russell, Cold Spring Harbor Press, Cold Spring Harbor, N.Y., 3rd edition, 2001.

The vector can also comprise a selectable marker gene for the selection of transformed cells. Selectable marker genes are utilized for the selection of transformed cells or tissues. Marker genes include genes encoding antibiotic resistance, such as those encoding neomycin phosphotransferase II (NEO) and hygromycin phosphotransferase (HPT), as well as genes conferring resistance to herbicidal compounds, such as glufosinate ammonium, bromoxynil, imidazolinones, and 2,4-dichlorophenoxyacetate (2,4-D).

In some embodiments, the expression cassette or vector comprising the sequence encoding a fusion protein comprising an RNA-guided DNA-binding polypeptide, such as an RGN, can further comprise a sequence encoding a gRNA. The sequence(s) encoding the gRNA can be operably linked to at least one transcriptional control sequence for expression of the gRNA in the organism or host cell of interest. For example, the polynucleotide encoding the gRNA can be operably linked to a promoter sequence that is recognized by RNA polymerase III (Pol III). Examples of suitable Pol III promoters include, but are not limited to, mammalian U6, U3, H1, and 7SL RNA promoters and rice U6 and U3 promoters.

As indicated, expression constructs comprising nucleotide sequences encoding the USPs, fusion proteins, and/or gRNAs can be used to transform organisms of interest. Methods for transformation involve introducing a nucleotide construct into an organism of interest. By "introducing" is intended to introduce the nucleotide construct to the host cell in such a manner that the construct gains access to the interior of the host cell. The methods of the invention do not require a particular method for introducing a nucleotide construct to a host organism, only that the nucleotide construct gains access to the interior of at least one cell of the host organism. The host cell can be a eukaryotic or prokaryotic cell. In particular embodiments, the eukaryotic host cell is a plant cell, a mammalian cell, or an insect cell. Methods for introducing nucleotide constructs into plants and other host cells are known in the art including, but not limited to, stable transformation methods, transient transformation methods, and virus-mediated methods.

The methods result in a transformed organism, such as a plant, including whole plants, as well as plant organs (*e.g.*, leaves, stems, roots, etc.), seeds, plant cells, propagules, embryos and progeny of the same. Plant cells can be differentiated or undifferentiated (*e.g.* callus, suspension culture cells, protoplasts, leaf cells, root cells, phloem cells, pollen).

"Transgenic organisms" or "transformed organisms" or "stably transformed" organisms or cells or tissues refers to organisms that have incorporated or integrated a polynucleotide encoding a deaminase of the invention. It is recognized that other exogenous or endogenous nucleic acid sequences or DNA fragments may also be incorporated into the host cell. *Agrobacterium*-and biolistic-mediated transformation remain the two predominantly employed approaches for transformation of plant cells. However, transformation of a

host cell may be performed by infection, transfection, microinjection, electroporation, microprojection, biolistics or particle bombardment, electroporation, silica/carbon fibers, ultrasound mediated, PEG mediated, calcium phosphate co-precipitation, polycation DMSO technique, DEAE dextran procedure, and viral mediated, liposome mediated and the like. Viral-mediated introduction of a polynucleotide encoding a deaminase, fusion protein, and/or gRNA includes retroviral, lentiviral, adenoviral, and adeno-associated viral mediated introduction and expression, as well as the use of Caulimoviruses (*e.g.*, cauliflower mosaic virus), Geminiviruses (*e.g.*, bean golden yellow mosaic virus or maize streak virus), and RNA plant viruses (*e.g.*, tobacco mosaic virus).

Transformation protocols as well as protocols for introducing polypeptides or polynucleotide sequences into plants may vary depending on the type of host cell (*e.g.*, monocot or dicot plant cell) targeted for transformation. Methods for transformation are known in the art and include those set forth in US Patent Nos: 8,575,425; 7,692,068; 8,802,934; 7,541,517; each of which is herein incorporated by reference. See, also, Rakoczy-Trojanowska, M. (2002) *Cell Mol Biol Lett.* 7:849-858; Jones *et al.* (2005) *Plant Methods* 1:5; Rivera *et al.* (2012) *Physics of Life Reviews* 9:308-345; Bartlett *et al.* (2008) *Plant Methods* 4:1-12; Bates, G.W. (1999) *Methods in Molecular Biology* 111:359-366; Binns and Thomashow (1988) *Annual Reviews in Microbiology* 42:575-606; Christou, P. (1992) *The Plant Journal* 2:275-281; Christou, P. (1995) *Euphytica* 85:13-27; Tzfira *et al.* (2004) *TRENDS in Genetics* 20:375-383; Yao *et al.* (2006) *Journal of Experimental Botany* 57:3737-3746; Zupan and Zambryski (1995) *Plant Physiology* 107:1041-1047; Jones *et al.* (2005) *Plant Methods* 1:5;

Transformation may result in stable or transient incorporation of the nucleic acid into the cell. "Stable transformation" is intended to mean that the nucleotide construct introduced into a host cell integrates into the genome of the host cell and is capable of being inherited by the progeny thereof. "Transient transformation" is intended to mean that a polynucleotide is introduced into the host cell and does not integrate into the genome of the host cell.

Methods for transformation of chloroplasts are known in the art. See, for example, Svab *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87:8526-8530; Svab and Maliga (1993) *Proc. Natl. Acad. Sci. USA* 90:913-917; Svab and Maliga (1993) *EMBO J.* 12:601-606. The method relies on particle gun delivery of DNA containing a selectable marker and targeting of the DNA to the plastid genome through homologous recombination. Additionally, plastid transformation can be accomplished by transactivation of a silent plastid-borne transgene by tissue-preferred expression of a nuclear-encoded and plastid-directed RNA polymerase. Such a system has been reported in McBride *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:7301-7305.

The cells that have been transformed may be grown into a transgenic organism, such as a plant, in accordance with conventional ways. See, for example, McCormick *et al.* (1986) *Plant Cell Reports* 5:81-84. These plants may then be grown, and either pollinated with the same transformed strain or different strains, and the resulting hybrid having constitutive expression of the desired phenotypic characteristic identified. Two or more generations may be grown to ensure that expression of the desired phenotypic characteristic is

stably maintained and inherited and then seeds harvested to ensure expression of the desired phenotypic characteristic has been achieved. In this manner, the present invention provides transformed seed (also referred to as "transgenic seed") having a nucleotide construct of the invention, for example, an expression cassette of the invention, stably incorporated into their genome.

Alternatively, cells that have been transformed may be introduced into an organism. These cells could have originated from the organism, wherein the cells are transformed in an *ex vivo* approach.

The sequences provided herein may be used for transformation of any plant species, including, but not limited to, monocots and dicots. Examples of plants of interest include, but are not limited to, corn (maize), sorghum, wheat, sunflower, tomato, crucifers, peppers, potato, cotton, rice, soybean, sugarbeet, sugarcane, tobacco, barley, and oilseed rape, Brassica sp., alfalfa, rye, millet, safflower, peanuts, sweet potato, cassava, coffee, coconut, pineapple, citrus trees, cocoa, tea, banana, avocado, fig, guava, mango, olive, papaya, cashew, macadamia, almond, oats, vegetables, ornamentals, and conifers.

Vegetables include, but are not limited to, tomatoes, lettuce, green beans, lima beans, peas, and members of the genus *Curcumis* such as cucumber, cantaloupe, and musk melon. Ornamentals include, but are not limited to, azalea, hydrangea, hibiscus, roses, tulips, daffodils, petunias, carnation, poinsettia, and chrysanthemum. Preferably, plants of the present invention are crop plants (for example, maize, sorghum, wheat, sunflower, tomato, crucifers, peppers, potato, cotton, rice, soybean, sugarbeet, sugarcane, tobacco, barley, oilseed rape, etc.).

As used herein, the term plant includes plant cells, plant protoplasts, plant cell tissue cultures from which plants can be regenerated, plant calli, plant clumps, and plant cells that are intact in plants or parts of plants such as embryos, pollen, ovules, seeds, leaves, flowers, branches, fruit, kernels, ears, cobs, husks, stalks, roots, root tips, anthers, and the like. Grain is intended to mean the mature seed produced by commercial growers for purposes other than growing or reproducing the species. Progeny, variants, and mutants of the regenerated plants are also included within the scope of the invention, provided that these parts comprise the introduced polynucleotides. Further provided is a processed plant product or byproduct that retains the sequences disclosed herein, including for example, soymeal.

The polynucleotides encoding the USPs, fusion proteins, and/or gRNAs can be used to transform any eukaryotic species, including but not limited to animals (*e.g.*, mammals, insects, fish, birds, and reptiles), fungi, amoeba, algae, and yeast. The polynucleotides encoding the USPs, fusion proteins, and/or gRNAs can also be used to transform any prokaryotic species, including but not limited to, archaea and bacteria (*e.g.*, *Bacillus* spp., *Klebsiella* spp., *Streptomyces* spp., *Rhizobium* spp., *Escherichia* spp., *Pseudomonas* spp., *Salmonella* spp., *Shigella* spp., *Vibrio* spp., *Yersinia* spp., *Mycoplasma* spp., *Agrobacterium* spp., and *Lactobacillus* spp.).

Conventional viral and non-viral based gene transfer methods can be used to introduce nucleic acids in mammalian cells or target tissues. Such methods can be used to administer nucleic acids encoding a fusion protein of the invention and optionally a gRNA to cells in culture, or in a host organism. Non-viral vector delivery systems include DNA plasmids, RNA (*e.g.* a transcript of a vector described herein), naked

nucleic acid, and nucleic acid complexed with a delivery vehicle, such as a liposome. Viral vector delivery systems include DNA and RNA viruses, which have either episomal or integrated genomes after delivery to the cell. Non-limiting examples include vectors utilizing Caulimoviruses (*e.g.*, cauliflower mosaic virus), Geminiviruses (*e.g.*, bean golden yellow mosaic virus or maize streak virus), and RNA plant viruses (*e.g.*, tobacco mosaic virus). For a review of gene therapy procedures, see Anderson, *Science* 256: 808- 813 (1992); Nabel & Feigner, *TIBTECH* 11:211-217 (1993); Mitani & Caskey, *TIBTECH* 11:162-166 (1993); Dillon, *TIBTECH* 11:167-175 (1993); Miller, *Nature* 357:455-460 (1992); Van Brunt, *Biotechnology* 6(10): 1149-1154 (1988); Vigne, *Restorative Neurology and Neuroscience* 8:35-36 (1995); Kremer & Perricaudet, *British Medical Bulletin* 51(1):31-44 (1995); Haddada et al., in *Current Topics in Microbiology and Immunology*, Doerfler and Bohm (eds) (1995); and Yu et al., *Gene Therapy* 1:13-26 (1994).

Methods of non-viral delivery of nucleic acids include lipofection, *Agrobacterium*-mediated transformation, nucleofection, microinjection, biolistics, virosomes, liposomes, immunoliposomes, polycation or lipid: nucleic acid conjugates, naked DNA, artificial virions, and agent-enhanced uptake of DNA. Lipofection is described in *e.g.*, U.S. Pat. Nos. 5,049,386, 4,946,787; and 4,897,355) and lipofection reagents are sold commercially (*e.g.*, Transfectam™ and Lipofectin™). Cationic and neutral lipids that are suitable for efficient receptor-recognition lipofection of polynucleotides include those of Feigner, WO 91/17424; WO 91/16024. Delivery can be to cells (*e.g.* in vitro or ex vivo administration) or target tissues (*e.g.* in vivo administration). The preparation of lipid:nucleic acid complexes, including targeted liposomes such as immunolipid complexes, is well known to one of skill in the art (see, *e.g.*, Crystal, *Science* 270:404-410 (1995); Blaese et al., *Cancer Gene Ther.* 2:291- 297 (1995); Behr et al., *Bioconjugate Chem.* 5:382-389 (1994); Remy et al., *Bioconjugate Chem.* 5:647-654 (1994); Gao et al., *Gene Therapy* 2:710-722 (1995); Ahmad et al., *Cancer Res.* 52:4817-4820 (1992); U.S. Pat. Nos. 4,186,183, 4,217,344, 4,235,871, 4,261,975, 4,485,054, 4,501,728, 4,774,085, 4,837,028, and 4,946,787).

The use of RNA or DNA viral based systems for the delivery of nucleic acids takes advantage of highly evolved processes for targeting a virus to specific cells in the body and trafficking the viral payload to the nucleus. Viral vectors can be administered directly to patients (in vivo) or they can be used to treat cells in vitro, and the modified cells may optionally be administered to patients (ex vivo). Conventional viral based systems could include retroviral, lentivirus, adenoviral, adeno-associated and herpes simplex virus vectors for gene transfer. Integration in the host genome is possible with the retrovirus, lentivirus, and adeno-associated virus gene transfer methods, often resulting in long term expression of the inserted transgene. Additionally, high transduction efficiencies have been observed in many different cell types and target tissues.

The tropism of a retrovirus can be altered by incorporating foreign envelope proteins, expanding the potential target population of target cells. Lentiviral vectors are retroviral vectors that are able to transduce or infect non-dividing cells and typically produce high viral titers. Selection of a retroviral gene transfer system would therefore depend on the target tissue. Retroviral vectors are comprised of cis-acting long terminal repeats with packaging capacity for up to 6-10 kb of foreign sequence. The minimum cis-acting

LTRs are sufficient for replication and packaging of the vectors, which are then used to integrate the therapeutic gene into the target cell to provide permanent transgene expression. Widely used retroviral vectors include those based upon murine leukemia virus (MuLV), gibbon ape leukemia virus (GaLV), Simian Immuno deficiency virus (SIV), human immuno deficiency virus (HIV), and combinations thereof (see, e.g., Buchscher et al., *J. Virol.* 66:2731-2739 (1992); Johann et al., *J. Virol.* 66:1635-1640 (1992); Sommnnerfelt et al., *Virol.* 176:58-59 (1990); Wilson et al., *J. Virol.* 63:2374-2378 (1989); Miller et al., *J. Virol.* 65:2220-2224 (1991); PCT/US94/05700).

In applications where transient expression is preferred, adenoviral based systems may be used. Adenoviral based vectors are capable of very high transduction efficiency in many cell types and do not require cell division. With such vectors, high titer and levels of expression have been obtained. This vector can be produced in large quantities in a relatively simple system. Adeno-associated virus ("AAV") vectors may also be used to transduce cells with target nucleic acids, e.g., in the in vitro production of nucleic acids and peptides, and for in vivo and ex vivo gene therapy procedures (see, e.g., West et al., *Virology* 160:38-47 (1987); U.S. Pat. No. 4,797,368; WO 93/24641; Katin, *Human Gene Therapy* 5:793-801 (1994); Muzyczka, *J. Clin. Invest.* 94:1351 (1994). Construction of recombinant AAV vectors are described in a number of publications, including U.S. Pat. No. 5,173,414; Tratschin et al., *Mol. Cell. Biol.* 5:3251-3260 (1985); Tratschin, et al., *Mol. Cell. Biol.* 4:2072-2081 (1984); Hermonat & Muzyczka, *PNAS* 81:6466-6470 (1984); and Samulski et al., *J. Virol.* 63:03822-3828 (1989). Packaging cells are typically used to form virus particles that are capable of infecting a host cell. Such cells include 293 cells, which package adenovirus, and ψ J2 cells or PA317 cells, which package retrovirus.

Viral vectors used in gene therapy are usually generated by producing a cell line that packages a nucleic acid vector into a viral particle. The vectors typically contain the minimal viral sequences required for packaging and subsequent integration into a host, other viral sequences being replaced by an expression cassette for the polynucleotide(s) to be expressed. The missing viral functions are typically supplied in trans by the packaging cell line. For example, AAV vectors used in gene therapy typically only possess ITR sequences from the AAV genome which are required for packaging and integration into the host genome. Viral DNA is packaged in a cell line, which contains a helper plasmid encoding the other AAV genes, namely rep and cap, but lacking ITR sequences.

The cell line may also be infected with adenovirus as a helper. The helper virus promotes replication of the AAV vector and expression of AAV genes from the helper plasmid. The helper plasmid is not packaged in significant amounts due to a lack of ITR sequences. Contamination with adenovirus can be reduced by, e.g., heat treatment to which adenovirus is more sensitive than AAV. Additional methods for the delivery of nucleic acids to cells are known to those skilled in the art. See, for example, US20030087817, incorporated herein by reference.

In some embodiments, a host cell is transiently or non-transiently transfected with one or more vectors described herein. In some embodiments, a cell is transfected as it naturally occurs in a subject. In some embodiments, a cell that is transfected is taken from a subject. In some embodiments, the cell is

derived from cells taken from a subject, such as a cell line. In some embodiments, the cell or cell line is prokaryotic. In other embodiments, the cell or cell line is eukaryotic. In further embodiments, the cell or cell line is derived from insect, avian, plant, or fungal species. In some embodiments, the cell or cell line may be mammalian, such as for example human, monkey, mouse, cow, swine, goat, hamster, rat, cat, or dog.

5 A wide variety of cell lines for tissue culture are known in the art. Examples of cell lines include, but are not limited to, C8161, CCRF-CEM, MOLT, mIMCD-3, NHDF, HeLaS3, Huh1, Huh4, Huh7, HUVEC, HASMC, HEK_n, HEK_a, MiaPaCell, Panel, PC-3, TFI, CTLL-2, CIR, Rat6, CVI, RPTE, AIO, T24, 182, A375, ARH-77, Calu1, SW480, SW620, SKOV3, SK-UT, CaCo2, P388D1, SEM-K2, WEHI- 231, HB56, TIB55, Jurkat, *145.01*, LRMB, Bcl-1, BC-3, IC21, DLD2, Raw264.7, NRK, NRK-52E, MRC5, MEF, Hep
10 G2, HeLa B, HeLa T4. COS, COS-1, COS-6, COS-M6A, BS-C-1 monkey kidney epithelial, BALB/3T3 mouse embryo fibroblast, 3T3 Swiss, 3T3-L1, 132-d5 human fetal fibroblasts; 10.1 mouse fibroblasts, 293-T, 3T3, 721, 9L, A2780, A2780ADR, A2780cis, A172, A20, A253, A431, A-549, ALC, B16, B35, BCP-1 cells, BEAS-2B, bEnd.3, BHK-21, BR 293, BxPC3, C3H-10T1/2, C6/36, Cal-27, CHO, CHO-7, CHO-IR, CHO-K1, CHO-K2, CHO-T, CHO Dhfr^{-/-}, COR-L23, COR-L23/CPR, COR-L235010, CORL23/ R23,
15 COS-7, COV-434, CML T1, CMT, CT26, D17, DH82, DU145, DuCaP, EL4, EM2, EM3, EMT6/AR1, EMT6/AR10.0, FM3, H1299, H69, HB54, HB55, HCA2, HEK-293, HeLa, Hepalclc7, HL-60, HMEC, HT-29, Jurkat, *1Y* cells, K562 cells, Ku812, KCL22, KGI, KYOI, LNCap, Ma-Mel 1-48, MC-38, MCF-7, MCF-10A, MDA-MB-231, MDA-MB-468, MDA-MB-435, MDCKII, MDCKII, MOR/ 0.2R, MONO-MAC 6, MTD-1A, MyEnd, NCI-H69/CPR, NCI-H69/LX10, NCI-H69/LX20, NCI-H69/LX4, NIH-3T3, NALM-1,
20 NW-145, OPCN/OPCT cell lines, Peer, PNT-1A/ PNT 2, RenCa, RIN-5F, RMA/RMAS, Saos-2 cells, Sf-9, SkBr3, T2, T-47D, T84, THP1 cell line, U373, U87, U937, VCaP, Vero cells, WM39, WT-49, X63, YAC-1, YAR, and transgenic varieties thereof. Cell lines are available from a variety of sources known to those with skill in the art (see, e.g., the American Type Culture Collection (ATCC) (Manassas, Va.)).

In some embodiments, a cell transfected with one or more vectors described herein is used to
25 establish a new cell line comprising one or more vector-derived sequences. In some embodiments, a cell transiently transfected with a fusion protein of the invention and optionally a gRNA, or with a ribonucleoprotein complex of the invention, and modified through the activity of fusion protein or ribonucleoprotein complex, is used to establish a new cell line comprising cells containing the modification but lacking any other exogenous sequence. In some embodiments, cells transiently or non-transiently
30 transfected with one or more vectors described herein, or cell lines derived from such cells are used in assessing one or more test compounds.

In some embodiments, one or more vectors described herein are used to produce a non-human transgenic animal or transgenic plant. In some embodiments, the transgenic animal is an insect. In further
embodiments, the insect is an insect pest, such as a mosquito or tick. In other embodiments, the insect is a
35 plant pest, such as a corn rootworm or a fall armyworm. In some embodiments, the transgenic animal is a bird, such as a chicken, turkey, goose, or duck. In some embodiments, the transgenic animal is a mammal, such as a human, mouse, rat, hamster, monkey, ape, rabbit, swine, cow, horse, goat, sheep, cat, or dog.

V. *Variants and Fragments of Polypeptides and Polynucleotides*

The present disclosure provides active variants and fragments of naturally-occurring (*i.e.*, wild-type) uracil stabilizing polypeptides, the amino acid sequence of which are set forth as SEQ ID NO: 1-16, and polynucleotides encoding the same.

While the activity of a variant or fragment may be altered compared to the polynucleotide or polypeptide of interest, the variant and fragment should retain the functionality of the polynucleotide or polypeptide of interest. For example, a variant or fragment may have increased activity, decreased activity, different spectrum of activity or any other alteration in activity when compared to the polynucleotide or polypeptide of interest.

Fragments and variants of naturally-occurring USPs, such as those disclosed herein, will retain activity such that if they are part of a fusion protein further comprising a deaminase or a fragment thereof and/or a DNA-binding polypeptide or a fragment thereof, said fusion protein will exhibit increased C→T nucleobase editing efficiency as compared to a similar fusion protein that does not comprise a USP domain.

The term "fragment" refers to a portion of a polynucleotide or polypeptide sequence of the invention. "Fragments" or "biologically active portions" include polynucleotides comprising a sufficient number of contiguous nucleotides to retain the biological activity (*i.e.*, deaminase activity on nucleic acids). "Fragments" or "biologically active portions" include polypeptides comprising a sufficient number of contiguous amino acid residues to retain the biological activity. Fragments of the USPs include those that are shorter than the full-length sequences due to the use of an alternate downstream start site. A biologically active portion of a USP can be a polypeptide that comprises, for example, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, or more contiguous amino acid residues of any of SEQ ID NOs: 1-16, or a variant thereof. Such biologically active portions can be prepared by recombinant techniques and evaluated for activity.

In general, "variants" is intended to mean substantially similar sequences. For polynucleotides, a variant comprises a deletion and/or addition of one or more nucleotides at one or more internal sites within the native polynucleotide and/or a substitution of one or more nucleotides at one or more sites in the native polynucleotide. As used herein, a "native" or "wild type" polynucleotide or polypeptide comprises a naturally occurring nucleotide sequence or amino acid sequence, respectively. For polynucleotides, conservative variants include those sequences that, because of the degeneracy of the genetic code, encode the native amino acid sequence of the gene of interest. Naturally occurring allelic variants such as these can be identified with the use of well-known molecular biology techniques, as, for example, with polymerase chain reaction (PCR) and hybridization techniques as outlined below. Variant polynucleotides also include synthetically derived polynucleotides, such as those generated, for example, by using site-directed mutagenesis but which still encode the polypeptide or the polynucleotide of interest. Generally, variants of a particular polynucleotide disclosed herein will have at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more sequence identity to that

particular polynucleotide as determined by sequence alignment programs and parameters described elsewhere herein.

Variants of a particular polynucleotide disclosed herein (*i.e.*, the reference polynucleotide) can also be evaluated by comparison of the percent sequence identity between the polypeptide encoded by a variant polynucleotide and the polypeptide encoded by the reference polynucleotide. Percent sequence identity between any two polypeptides can be calculated using sequence alignment programs and parameters described elsewhere herein. Where any given pair of polynucleotides disclosed herein is evaluated by comparison of the percent sequence identity shared by the two polypeptides they encode, the percent sequence identity between the two encoded polypeptides is at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more sequence identity.

In particular embodiments, the presently disclosed polynucleotides encode a USP comprising an amino acid sequence having at least 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or greater identity to an amino acid sequence of any of SEQ ID NOs: 1-16.

A biologically active variant of a uracil stabilizing polypeptide of the invention may differ by as few as about 1-15 amino acid residues, as few as about 1-10, such as about 6-10, as few as 5, as few as 4, as few as 3, as few as 2, or as few as 1 amino acid residue. In specific embodiments, the polypeptides can comprise an N-terminal or a C-terminal truncation, which can comprise at least a deletion of 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, amino acids or more from either the N or C terminus of the polypeptide.

It is recognized that modifications may be made to the USPs provided herein creating variant proteins and polynucleotides. Changes designed by man may be introduced through the application of site-directed mutagenesis techniques. Alternatively, native, as yet-unknown or as yet unidentified polynucleotides and/or polypeptides structurally and/or functionally-related to the sequences disclosed herein may also be identified that fall within the scope of the present invention. Conservative amino acid substitutions may be made in nonconserved regions that do not alter the function of the uracil stabilizing polypeptide. Alternatively, modifications may be made that improve the activity of the uracil stabilizing polypeptide.

Variant polynucleotides and proteins also encompass sequences and proteins derived from a mutagenic and recombinogenic procedure such as DNA shuffling. With such a procedure, one or more different USPs disclosed herein (e.g., SEQ ID NO: 1-16) is manipulated to create a new USP possessing the desired properties. In this manner, libraries of recombinant polynucleotides are generated from a population of related sequence polynucleotides comprising sequence regions that have substantial sequence identity and can be homologously recombined *in vitro* or *in vivo*. For example, using this approach, sequence motifs encoding a domain of interest may be shuffled between the USP sequences provided herein and other subsequently identified USP genes to obtain a new gene coding for a protein with an improved property of interest, such as an increased K_m in the case of an enzyme. Strategies for such DNA shuffling are known in the art. See, for example, Stemmer (1994) *Proc. Natl. Acad. Sci. USA* 91:10747-10751; Stemmer (1994)

Nature 370:389-391; Cramer *et al.* (1997) *Nature Biotech.* 15:436-438; Moore *et al.* (1997) *J. Mol. Biol.* 272:336-347; Zhang *et al.* (1997) *Proc. Natl. Acad. Sci. USA* 94:4504-4509; Cramer *et al.* (1998) *Nature* 391:288-291; and U.S. Patent Nos. 5,605,793 and 5,837,458. A "shuffled" nucleic acid is a nucleic acid produced by a shuffling procedure such as any shuffling procedure set forth herein. Shuffled nucleic acids are produced by recombining (physically or virtually) two or more nucleic acids (or character strings), for example in an artificial, and optionally recursive, fashion. Generally, one or more screening steps are used in shuffling processes to identify nucleic acids of interest; this screening step can be performed before or after any recombination step. In some (but not all) shuffling embodiments, it is desirable to perform multiple rounds of recombination prior to selection to increase the diversity of the pool to be screened. The overall process of recombination and selection are optionally repeated recursively. Depending on context, shuffling can refer to an overall process of recombination and selection, or, alternately, can simply refer to the recombinational portions of the overall process.

As used herein, "sequence identity" or "identity" in the context of two polynucleotides or polypeptide sequences makes reference to the residues in the two sequences that are the same when aligned for maximum correspondence over a specified comparison window. When percentage of sequence identity is used in reference to proteins it is recognized that residue positions which are not identical often differ by conservative amino acid substitutions, where amino acid residues are substituted for other amino acid residues with similar chemical properties (*e.g.*, charge or hydrophobicity) and therefore do not change the functional properties of the molecule. When sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Sequences that differ by such conservative substitutions are said to have "sequence similarity" or "similarity". Means for making this adjustment are well known to those of skill in the art. Typically, this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated, *e.g.*, as implemented in the program PC/GENE (Intelligenetics, Mountain View, California).

As used herein, "percentage of sequence identity" means the value determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison, and multiplying the result by 100 to yield the percentage of sequence identity.

Unless otherwise stated, sequence identity/similarity values provided herein refer to the value obtained using GAP Version 10 using the following parameters: % identity and % similarity for a nucleotide

sequence using GAP Weight of 50 and Length Weight of 3, and the nwsgapdna.cmp scoring matrix; % identity and % similarity for an amino acid sequence using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix; or any equivalent program thereof. By "equivalent program" is intended any sequence comparison program that, for any two sequences in question, generates an alignment having identical nucleotide or amino acid residue matches and an identical percent sequence identity when compared to the corresponding alignment generated by GAP Version 10.

Two sequences are "optimally aligned" when they are aligned for similarity scoring using a defined amino acid substitution matrix (*e.g.*, BLOSUM62), gap existence penalty and gap extension penalty so as to arrive at the highest score possible for that pair of sequences. Amino acid substitution matrices and their use in quantifying the similarity between two sequences are well-known in the art and described, *e.g.*, in Dayhoff et al. (1978) "A model of evolutionary change in proteins." In "Atlas of Protein Sequence and Structure," Vol. 5, Suppl. 3 (ed. M. O. Dayhoff), pp. 345-352. Natl. Biomed. Res. Found., Washington, D.C. and Henikoff et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:10915-10919. The BLOSUM62 matrix is often used as a default scoring substitution matrix in sequence alignment protocols. The gap existence penalty is imposed for the introduction of a single amino acid gap in one of the aligned sequences, and the gap extension penalty is imposed for each additional empty amino acid position inserted into an already opened gap. The alignment is defined by the amino acids positions of each sequence at which the alignment begins and ends, and optionally by the insertion of a gap or multiple gaps in one or both sequences, so as to arrive at the highest possible score. While optimal alignment and scoring can be accomplished manually, the process is facilitated by the use of a computer-implemented alignment algorithm, *e.g.*, gapped BLAST 2.0, described in Altschul et al. (1997) *Nucleic Acids Res.* 25:3389-3402, and made available to the public at the National Center for Biotechnology Information Website (www.ncbi.nlm.nih.gov). Optimal alignments, including multiple alignments, can be prepared using, *e.g.*, PSI-BLAST, available through www.ncbi.nlm.nih.gov and described by Altschul et al. (1997) *Nucleic Acids Res.* 25:3389-3402.

With respect to an amino acid sequence that is optimally aligned with a reference sequence, an amino acid residue "corresponds to" the position in the reference sequence with which the residue is paired in the alignment. The "position" is denoted by a number that sequentially identifies each amino acid in the reference sequence based on its position relative to the N-terminus. Owing to deletions, insertion, truncations, fusions, etc., that must be taken into account when determining an optimal alignment, in general the amino acid residue number in a test sequence as determined by simply counting from the N-terminal will not necessarily be the same as the number of its corresponding position in the reference sequence. For example, in a case where there is a deletion in an aligned test sequence, there will be no amino acid that corresponds to a position in the reference sequence at the site of deletion. Where there is an insertion in an aligned reference sequence, that insertion will not correspond to any amino acid position in the reference sequence. In the case of truncations or fusions there can be stretches of amino acids in either the reference or aligned sequence that do not correspond to any amino acid in the corresponding sequence.

VI. Antibodies

Antibodies to the USPs, fusion proteins, or ribonucleoproteins comprising the USPs of the present invention, including those having the amino acid sequence set forth as SEQ ID NOs: 1-16 or active variants or fragments thereof, are also encompassed. Methods for producing antibodies are well known in the art (see, for example, Harlow and Lane (1988) *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.; and U.S. Pat. No. 4,196,265). These antibodies can be used in kits for the detection and isolation of USPs or fusion proteins or ribonucleoproteins comprising USPs described herein. Thus, this disclosure provides kits comprising antibodies that specifically bind to the polypeptides or ribonucleoproteins described herein, including, for example, polypeptides comprising a sequence of at least 85% identity to any of SEQ ID NOs: 1-16.

VII. Systems and Ribonucleoprotein Complexes for Binding a Target Sequence of Interest and Methods of Making the Same

The present disclosure provides a system which targets to a nucleic acid sequence and modifies a target nucleic acid sequence. In some embodiments, an RNA-guided, DNA-binding polypeptide, such as an RGN, and the gRNA are responsible for targeting the ribonucleoprotein complex to a nucleic acid sequence of interest; the deaminase polypeptide is responsible for modifying the targeted nucleic acid sequence from C>U; the uracil stabilizing polypeptide allows the uracil to persist in the DNA molecule so that the desired DNA repair occurs, thereby introducing the C>T mutation. The guide RNA hybridizes to the target sequence of interest and also forms a complex with the RNA-guided, DNA-binding polypeptide, thereby directing the RNA-guided, DNA-binding polypeptide to bind to the target sequence. The RNA-guided, DNA-binding polypeptide is one domain of a 3-domain fusion protein; the second domain is a deaminase, and the third domain is a USP described herein. In some embodiments, the RNA-guided, DNA-binding polypeptide is an RGN, such as a Cas9. Other examples of RNA-guided, DNA-binding polypeptides include RGNs such as those described in U.S. Patent Application Publication No. 2019/0367949 (herein incorporated in its entirety by reference). In some embodiments, the RNA-guided, DNA-binding polypeptide is a Type II CRISPR-Cas polypeptide, or an active variant or fragment thereof. In some embodiments, the RNA-guided, DNA-binding polypeptide is a Type V CRISPR-Cas polypeptide, or an active variant or fragment thereof. In other embodiments, the RNA-guided, DNA-binding polypeptide is a Type VI CRISPR-Cas polypeptide. In other embodiments, the DNA-binding domain of the fusion protein does not require an RNA guide, such as a zinc finger nuclease, TALEN, or meganuclease polypeptide. In some of these embodiments, the nuclease activity of each has been inactivated. In further embodiments, the RNA-guided, DNA-binding polypeptide comprises an amino acid sequence of an RGN, such as an amino acid sequence having at least 80% sequence identity to APG07433.1 (SEQ ID NO: 40) or an active variant or fragment thereof such as nickase APG07433.1 (SEQ ID NO: 41). In further embodiments, the RNA-guided, DNA-binding polypeptide comprises an amino acid sequence of an RGN, such as an amino acid

sequence having at least 85% sequence identity to APG07433.1 (SEQ ID NO: 40) or an active variant or fragment thereof such as nickase APG07433.1 (SEQ ID NO: 41). In further embodiments, the RNA-guided, DNA-binding polypeptide comprises an amino acid sequence of an RGN, such as an amino acid sequence having at least 90% sequence identity to APG07433.1 (SEQ ID NO: 40) or an active variant or fragment thereof such as nickase APG07433.1 (SEQ ID NO: 41). In further embodiments, the RNA-guided, DNA-binding polypeptide comprises an amino acid sequence of an RGN, such as an amino acid sequence having at least 95% sequence identity to APG07433.1 (SEQ ID NO: 40) or an active variant or fragment thereof such as nickase APG07433.1 (SEQ ID NO: 41). In further embodiments, the RNA-guided, DNA-binding polypeptide comprises an amino acid sequence of an RGN, such as APG07433.1 (SEQ ID NO: 40) or an active variant or fragment thereof such as nickase APG07433.1 (SEQ ID NO: 41).

The system for binding a target sequence of interest provided herein can be a ribonucleoprotein complex, which is at least one molecule of an RNA bound to at least one protein. The ribonucleoprotein complexes provided herein comprise at least one guide RNA as the RNA component and a fusion protein comprising a deaminase, a USP of the invention, and an RNA-guided, DNA-binding polypeptide as the protein component. The ribonucleoprotein complex can be purified from a cell or organism that has been transformed with polynucleotides that encode the fusion protein and a guide RNA and cultured under conditions to allow for the expression of the fusion protein and guide RNA. Thus, methods are provided for making a USP, a fusion protein, or a fusion protein ribonucleoprotein complex. Such methods comprise culturing a cell comprising a nucleotide sequence encoding a USP, a fusion protein, and in some embodiments a nucleotide sequence encoding a guide RNA, under conditions in which the USP or fusion protein (and in some embodiments, the guide RNA) is expressed. The USP, fusion protein, or fusion ribonucleoprotein can then be purified from a lysate of the cultured cells.

Methods for purifying a USP, fusion protein, or fusion ribonucleoprotein complex from a lysate of a biological sample are known in the art (*e.g.*, size exclusion and/or affinity chromatography, 2D-PAGE, HPLC, reversed-phase chromatography, immunoprecipitation). In particular methods, the USP or fusion protein is recombinantly produced and comprises a purification tag to aid in its purification, including but not limited to, glutathione-S-transferase (GST), chitin binding protein (CBP), maltose binding protein, thioredoxin (TRX), poly(NANP), tandem affinity purification (TAP) tag, myc, AcV5, AU1, AU5, E, ECS, E2, FLAG, HA, nus, Softag 1, Softag 3, Strep, SBP, Glu-Glu, HSV, KT3, S, S1, T7, V5, VSV-G, 6xHis, biotin carboxyl carrier protein (BCCP), and calmodulin. Generally, the tagged USP, fusion protein, or fusion ribonucleoprotein complex is purified using immunoprecipitation or other similar methods known in the art.

An "isolated" or "purified" polypeptide, or biologically active portion thereof, is substantially or essentially free from components that normally accompany or interact with the polypeptide as found in its naturally occurring environment. Thus, an isolated or purified polypeptide is substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. A protein that is substantially free of

cellular material includes preparations of protein having less than about 30%, 20%, 10%, 5%, or 1% (by dry weight) of contaminating protein. When the protein of the invention or biologically active portion thereof is recombinantly produced, optimally culture medium represents less than about 30%, 20%, 10%, 5%, or 1% (by dry weight) of chemical precursors or non-protein-of-interest chemicals.

5

VIII. Methods of Modifying a Target Sequence

The present disclosure provides methods for modifying a target nucleic acid molecule (e.g., target DNA molecule) of interest. The methods include delivering a system comprising at least one guide RNA or a polynucleotide encoding the same, and at least one fusion protein comprising a USP of the invention, a deaminase, and an RNA-guided, DNA-binding polypeptide or a polynucleotide encoding the same to the target sequence or a cell, organelle, or embryo comprising the target sequence. In some of these embodiments, the fusion protein comprises any one of the amino acid sequences of SEQ ID NOs: 1-16, or an active variant or fragment thereof.

In some embodiments, the methods comprise contacting a DNA molecule with (a) a fusion protein comprising a USP, a deaminase, and an RNA-guided, DNA-binding polypeptide, such as for example a nuclease-inactive or a nickase Cas9 domain; and (b) a gRNA targeting the fusion protein of (a) to a target nucleotide sequence of the DNA molecule; wherein the DNA molecule is contacted with the fusion protein and the gRNA in an amount effective and under conditions suitable for the deamination of a nucleotide base. In some embodiments, the target DNA molecule comprises a sequence associated with a disease or disorder, and wherein the deamination of the nucleotide base results in a sequence that is not associated with a disease or disorder. In some embodiments, the disease or disorder affects animals. In further embodiments, the disease or disorder affects mammals, such as humans, cows, horses, dogs, cats, goats, sheep, swine, monkeys, rats, mice, or hamsters. In some embodiments, the target DNA sequence resides in an allele of a crop plant, wherein the particular allele of the trait of interest results in a plant of lesser agronomic value. The deamination of the nucleotide base results in an allele that improves the trait and increases the agronomic value of the plant.

In some embodiments, the desired mutation comprises a T→C point mutation associated with a disease or disorder, and wherein the deamination of the mutant C base results in a sequence that is not associated with a disease or disorder. In some embodiments, the deamination corrects a point mutation in the sequence associated with the disease or disorder.

In some embodiments, the sequence associated with the disease or disorder encodes a protein, and wherein the deamination introduces a stop codon into the sequence associated with the disease or disorder, resulting in a truncation of the encoded protein. In some embodiments, the contacting is performed *in vivo* in a subject susceptible to having, having, or diagnosed with the disease or disorder. In some embodiments, the disease or disorder is a disease associated with a point mutation, or a single-base mutation, in the

genome. In some embodiments, the disease is a genetic disease, a cancer, a metabolic disease, or a lysosomal storage disease.

Thus, the presently disclosed compositions and methods can be used for the treatment of a disease or a disorder associated with a sequence (i.e., the sequence is causal for the disease or disorder or causal for symptoms associated with the disease or disorder) that is mutated in order to treat the disease or disorder or the reduction of symptoms associated with the disease or disorder. As used herein, the term “treat” or “treatment” refers to the administration of a pharmaceutical composition disclosed herein comprising a USP or a fusion protein, to a subject having a disease or disorder. Treatment can be prophylactic by preventing the onset of symptoms associated with a disease or disorder in a subject susceptible to the disease or disorder (e.g., genetically predisposed). Desirable effects of treatment include, but are not limited to, preventing occurrence or recurrence of disease, alleviation of symptoms, diminishment of any direct or indirect pathological consequences of the disease, decreasing the rate of disease progression, amelioration or palliation of the disease state, and remission or improved prognosis.

A pharmaceutical composition is a composition that is employed to prevent, reduce in intensity, cure or otherwise treat a disease or disorder that comprises an active ingredient (i.e., a USP or fusion protein or nucleic acid molecule encoding the same) and a pharmaceutically acceptable carrier. A pharmaceutically acceptable carrier refers to one or more compatible solid or liquid filler, diluents or encapsulating substances which are suitable for administration to a human or other vertebrate animal. In some embodiments, the pharmaceutical composition comprises a pharmaceutically acceptable carrier that is non-naturally occurring.

Pharmaceutical compositions used in the presently disclosed methods can be formulated with suitable carriers, excipients, and other agents that provide suitable transfer, delivery, tolerance, and the like. A multitude of appropriate formulations are known to those skilled in the art. Non-limiting examples include a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. Administered intravenously, particular carriers are physiological saline or phosphate buffered saline (PBS). Pharmaceutical compositions for oral or parenteral use may be prepared into dosage forms in a unit dose suited to fit a dose of the active ingredients. Such dosage forms in a unit dose include, for example, tablets, pills, capsules, injections (ampoules), suppositories, etc. These compositions also may contain adjuvants including preservative agents, wetting agents, emulsifying agents, and dispersing agents. Prevention of the action of microorganisms may be ensured by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, and the like. It also may be desirable to include isotonic agents, for example, sugars, sodium chloride and the like. Prolonged absorption of the injectable pharmaceutical form may be brought about by the use of agents delaying absorption, for example, aluminum monostearate and gelatin.

Pharmaceutical compositions comprising the USP or fusion proteins or nucleic acid molecules encoding the same or cells comprising the same can be administered to a subject via any route, such as orally, buccally, parenterally, topically, by inhalation or insufflation (i.e., through the mouth or through the nose), or rectally. Administering can be performed, for example, once, a plurality of times, and/or over one or more extended periods.

An effective amount of a pharmaceutical composition of the invention is any amount that is effective to achieve its purpose (e.g., prevention of or recovery from, including partial recovery, or prevention or slowing of disorder or disease caused by a specific sequence). The effective amount, usually expressed in mg/kg can be determined by routine methods during pre-clinical and clinical trials by those of skill in the art.

In those embodiments wherein the method comprises delivering a polynucleotide encoding a guide RNA and/or a fusion protein, the cell or embryo can then be cultured under conditions in which the guide RNA and/or fusion protein are expressed. In various embodiments, the method comprises contacting a target sequence with a ribonucleoprotein complex comprising a gRNA and a fusion protein (which comprises a USP of the invention, a deaminase, and an RNA-guided DNA-binding polypeptide). In certain embodiments, the method comprises introducing into a cell, organelle, or embryo comprising a target sequence a ribonucleoprotein complex of the invention. The ribonucleoprotein complex of the invention can be one that has been purified from a biological sample, recombinantly produced and subsequently purified, or *in vitro*-assembled as described herein. In those embodiments wherein the ribonucleoprotein complex that is contacted with the target sequence or a cell organelle, or embryo has been assembled *in vitro*, the method can further comprise the *in vitro* assembly of the complex prior to contact with the target sequence, cell, organelle, or embryo.

A purified or *in vitro* assembled ribonucleoprotein complex of the invention can be introduced into a cell, organelle, or embryo using any method known in the art, including, but not limited to electroporation. Alternatively, a fusion protein comprising a USP of the invention, a deaminase, and a RNA-guided, DNA-binding polypeptide, and a polynucleotide encoding or comprising the guide RNA can be introduced into a cell, organelle, or embryo using any method known in the art (e.g., electroporation).

Upon delivery to or contact with the target sequence or cell, organelle, or embryo comprising the target sequence, the guide RNA directs the fusion protein to bind to the target sequence in a sequence-specific manner. The target sequence can subsequently be modified via the deaminase domain and the USP domain of the fusion protein. In some embodiments, the binding of this fusion protein to a target sequence results in modification of a nucleotide adjacent to the target sequence. The nucleotide base adjacent to the target sequence may be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 base pairs from the 5' or 3' end of the target sequence. A fusion protein comprising a USP of the invention, a deaminase, and a RNA-guided, DNA-binding polypeptide can introduce targeted C>T mutations with greater efficiency compared to a fusion protein which comprises a deaminase and an RNA-guided, DNA-binding polypeptide alone.

Methods to measure binding of the fusion protein to a target sequence are known in the art and include chromatin immunoprecipitation assays, gel mobility shift assays, DNA pull-down assays, reporter assays, microplate capture and detection assays. Likewise, methods to measure cleavage or modification of a target sequence are known in the art and include *in vitro* or *in vivo* cleavage assays wherein cleavage is confirmed using PCR, sequencing, or gel electrophoresis, with or without the attachment of an appropriate label (e.g., radioisotope, fluorescent substance) to the target sequence to facilitate detection of degradation products. Alternatively, the nicking triggered exponential amplification reaction (NTEXPAR) assay can be used (see, e.g., Zhang et al. (2016) *Chem. Sci.* 7:4951-4957). *In vivo* cleavage can be evaluated using the Surveyor assay (Guschin et al. (2010) *Methods Mol Biol* 649:247-256).

In some embodiments, the methods involve the use of a RNA-binding, DNA-guided domain, as part of the fusion protein, complexed with more than one guide RNA. The more than one guide RNA can target different regions of a single gene or can target multiple genes. This multiple targeting enables the deaminase domain of the fusion protein to modify nucleic acids, thereby introducing multiple mutations in the target nucleic acid molecule (e.g., genome) of interest. The USP domain of the fusion protein increases the efficacy of introduction of the desired mutations.

In those embodiments wherein the method involves the use of an RNA-guided nuclease (RGN), such as a nickase RGN (i.e., is only able to cleave a single strand of a double-stranded polynucleotide, for example nAPG07433.1 (SEQ ID NO: 41)), the method can comprise introducing two different RGNs or RGN variants that target identical or overlapping target sequences and cleave different strands of the polynucleotide. For example, an RGN nickase that only cleaves the positive (+) strand of a double-stranded polynucleotide can be introduced along with a second RGN nickase that only cleaves the negative (-) strand of a double-stranded polynucleotide. Alternatively, two different fusion proteins may be provided, where each fusion protein comprises a different RGN with a different PAM recognition sequence, so that a greater diversity of nucleotide sequences may be targeted for mutation.

One of ordinary skill in the art will appreciate that any of the presently disclosed methods can be used to target a single target sequence or multiple target sequences. Thus, methods comprise the use of a fusion protein comprising a single RNA-guided, DNA-binding polypeptide in combination with multiple, distinct guide RNAs, which can target multiple, distinct sequences within a single gene and/or multiple genes. The deaminase domain of the fusion protein would then introduce mutations at each of the targeted sequences. The USP domain of the fusion protein increases the efficacy of introduction of the desired mutations. Also encompassed herein are methods wherein multiple, distinct guide RNAs are introduced in combination with multiple, distinct RNA-guided, DNA binding polypeptides. Such RNA-guided, DNA-binding polypeptides may be multiple RGN or RGN variants. These guide RNAs and guide RNA/fusion protein systems can target multiple, distinct sequences within a single gene and/or multiple genes.

IX. Cells Comprising a Polynucleotide Genetic Modification

Some embodiments provide methods for using the fusion proteins provided herein. In some embodiments, the fusion protein is used to introduce a point mutation into a target nucleic acid molecule by deaminating a target nucleobase, *e.g.*, a C residue. In some embodiments, the deamination of the target nucleobase results in the correction of a genetic defect, *e.g.*, in the correction of a point mutation that leads to a loss of function in a gene product. In some embodiments, the genetic defect is associated with a disease or disorder, *e.g.*, a lysosomal storage disorder or a metabolic disease, such as, for example, type I diabetes. In some embodiments, the methods provided herein are used to introduce a deactivating point mutation into a gene or allele that encodes a gene product that is associated with a disease or disorder. For example, in some embodiments, methods are provided herein that employ a fusion protein to introduce a deactivating point mutation into an oncogene (*e.g.*, in the treatment of a proliferative disease). A deactivating mutation may, in some embodiments, generate a premature stop codon in a coding sequence, which results in the expression of a truncated gene product, *e.g.*, a truncated protein lacking the function of the full-length protein. In some embodiments, the purpose of the methods provide herein is to restore the function of a dysfunctional gene via genome editing. The fusion proteins provided herein can be validated for gene editing-based human therapeutics in vitro, *e.g.*, by correcting a disease associated mutation in human cell culture. It will be understood by the skilled artisan that the fusion proteins provided herein, *e.g.*, the fusion proteins comprising a RNA-guided, DNA-binding domain, a deaminase domain, and a USP of the invention can be used to correct any single point C>T mutation. Deamination of the mutant C to U leads to a correction of the mutation.

In some embodiments, a fusion protein comprising an RNA-guided, DNA-binding domain, a deaminase domain, and a USP of the invention may be used for generating mutations in a targeted gene or targeted region of a gene of interest. In some embodiments, a fusion protein of the invention may be used for saturate mutagenesis of a targeted gene or region of a targeted gene of interest followed by high-throughput forward genetic screening to identify novel mutations and/or phenotypes. In other embodiments, a fusion protein described herein may be used for generating mutations in a targeted genomic location, which may or may not comprise coding DNA sequence. Libraries of cell lines generated by the targeted mutagenesis described above may also be useful for study of gene function or gene expression.

Fusion proteins of the invention may also be used to efficiently generate knock-out (KO) lines, including entire libraries of KO lines, through targeted insertion of nonsense mutations. Fusion proteins comprising a RNA-guided, DNA-binding domain, a deaminase domain, and a USP of the invention can convert three codons (CAA, CAG, and CGA) into STOP codons (TAG, TAA, or TGA) if targeted to the coding DNA strain, and can convert TGG into a STOP codon if targeted to the non-coding DNA strain. Billon et al (2017, *Mol Cell* 67: 1068-1079; incorporated by reference herein) provide a database of over 3.4 million guide RNAs in eight eukaryotic species useful for generation STOP codons. Additionally, Kuscu et al (2017, *Nature Methods* 14(7): 710-714; incorporated by reference herein) identified ~260,000 unique i-stop sgRNAs in the human genome which can target nearly 17,000 genes to introduce early stop codons. In some embodiments, the KO lines are eukaryotic cells. In other embodiments, the KO lines are prokaryotic

cells. In some embodiments, the KO lines generated using a fusion protein of the invention are human cell lines. In other embodiments, the KO lines are mammalian cell lines, for example mouse, rat, monkey, cat, dog, cow, pig, sheep, or horse cell lines. In other embodiments, the KO lines are avian cells. In other
 5 other embodiments, the KO lines are insect cells. In other embodiments, the KO lines are microbial cells. In still other embodiments, the KO lines are plant cells. In further embodiments, the KO lines are Arabidopsis, soybean, maize, cotton, tomato, potato, or bean cells. In further embodiments, the cell lines are plant seeds.

In some embodiments, a fusion protein provided herein may be useful in therapeutic genome editing. For example, a fusion protein comprising a RNA-guided, DNA-binding domain, a deaminase domain, and a USP of the invention may be used to generate targeted nonsense mutations of *PCSK9*
 10 (proprotein convertase subtilisin/kexin type 9). *PCSK9* is involved in lipoprotein homeostasis, and agents which block PCSK9 can lower low-density lipoprotein particle (LDL) concentrations in the blood. Naturally occurring nonsense variants in *PCSK9* in individuals result in substantially reduced blood cholesterol levels and reduced risk in coronary heart disease (Cohen et al. (2006) *N Engl J Med* 354: 1264-1272). Chadwick et al. (2017, *Arterioscler Thromb Vasc Biol* 37: 1741-1747; incorporated by reference
 15 herein) have found that they can successfully introduce targeted C>T mutations into the *PCSK9* gene, thereby reducing PCSK9 protein levels and plasma cholesterol levels in mice. In some embodiments, a similar approach may be taken with a fusion protein of the invention. Further, the skilled artisan will understand that the instantly disclosed fusion proteins can be used to correct other point mutations and mutations associated with other cancers and with diseases other than cancer including other proliferative
 20 diseases.

Provided herein are cells and organisms comprising a target nucleic acid molecule of interest that has been modified using a process mediated by a fusion protein, optionally with a gRNA as described herein. In some of these embodiments, the fusion protein comprises a USP comprising an amino acid sequence of any of SEQ ID NOs: 1-16, or an active variant or fragment thereof. In some embodiments, the
 25 fusion protein comprises a USP comprising an amino acid sequence at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to any of SEQ ID NOs: 1-16. In some embodiments, the fusion protein further comprises a deaminase and a RNA-guided, DNA-binding polypeptide. In further embodiments, the fusion protein comprises a deaminase and an RGN or a variant thereof, such as an amino
 30 acid sequence having at least 80% sequence identity to APG07433.1 (SEQ ID NO: 40) or its nickase variant nAPG07433.1 (SEQ ID NO: 41). In further embodiments, the fusion protein comprises a deaminase and an RGN or a variant thereof, such as an amino acid sequence having at least 85% sequence identity to APG07433.1 (SEQ ID NO: 40) or its nickase variant nAPG07433.1 (SEQ ID NO: 41). In further embodiments, the fusion protein comprises a deaminase and an RGN or a variant thereof, such as an amino
 35 acid sequence having at least 90% sequence identity to APG07433.1 (SEQ ID NO: 40) or its nickase variant nAPG07433.1 (SEQ ID NO: 41). In further embodiments, the fusion protein comprises a deaminase and an RGN or a variant thereof, such as an amino acid sequence having at least 95% sequence identity to

APG07433.1 (SEQ ID NO: 40) or its nickase variant nAPG07433.1 (SEQ ID NO: 41). In further embodiments, the fusion protein comprises a deaminase and an RGN or a variant thereof, such as APG07433.1 (SEQ ID NO: 40) or its nickase variant nAPG07433.1 (SEQ ID NO: 41). In other embodiments, the fusion protein comprises a deaminase and a Cas9 or a variant thereof, such as for example dCas9 or nickase Cas9. In some embodiments, the fusion protein comprises a nuclease-inactive or nickase variant of a Type II CRISPR-Cas polypeptide. In other embodiments, the fusion protein comprises a nuclease-inactive or nickase variant of a Type V CRISPR-Cas polypeptide. In still other embodiments, the fusion protein comprises a nuclease-inactive or nickase variant of a Type VI CRISPR-Cas polypeptide.

The modified cells can be eukaryotic (*e.g.*, mammalian, plant, insect, avian cell) or prokaryotic. Also provided are organelles and embryos comprising at least one nucleotide sequence that has been modified by a process utilizing a fusion protein as described herein. The genetically modified cells, organisms, organelles, and embryos can be heterozygous or homozygous for the modified nucleotide sequence.

The mutation(s) introduced by the deaminase domain of the fusion protein can result in altered expression (up-regulation or down-regulation), inactivation, or the expression of an altered protein product or an integrated sequence. In those instances wherein the mutation(s) results in either the inactivation of a gene or the expression of a non-functional protein product, the genetically modified cell, organism, organelle, or embryo is referred to as a “knock out”. The knock out phenotype can be the result of a deletion mutation (*i.e.*, deletion of at least one nucleotide), an insertion mutation (*i.e.*, insertion of at least one nucleotide), or a nonsense mutation (*i.e.*, substitution of at least one nucleotide such that a stop codon is introduced).

In other embodiments, the mutation(s) introduced by the deaminase domain of the fusion protein results in the production of a variant protein product. The expressed variant protein product can have at least one amino acid substitution and/or the addition or deletion of at least one amino acid. The variant protein product can exhibit modified characteristics or activities when compared to the wild-type protein, including but not limited to altered enzymatic activity or substrate specificity.

In yet other embodiments, the mutation(s) introduced by the deaminase domain of the fusion protein can result in an altered expression pattern of a protein. As a non-limiting example, mutation(s) in the regulatory regions controlling the expression of a protein product can result in the overexpression or downregulation of the protein product or an altered tissue or temporal expression pattern.

Some aspects of this disclosure provide kits comprising a fusion protein comprising an RNA-guided, DNA-binding polypeptide, such as an RGN polypeptide, for example a nuclease-inactive Cas9 domain, and a deaminase of the invention, and, optionally, a linker positioned between the Cas9 domain and the deaminase. In addition, in some embodiments, the kit comprises suitable reagents, buffers, and/or instructions for using the fusion protein, *e.g.*, for *in vitro* or *in vivo* DNA or RNA editing. In some embodiments, the kit comprises instructions regarding the design and use of suitable gRNAs for targeted editing of a nucleic acid sequence.

X: Additional Applications of USPs

USPs described herein also possess utility beyond genomic base editing. In general, USPs are useful in applications where stabilizing a uracil nucleobase in a DNA molecule is desired. Through a natural process or by the hand of man, a uracil may be introduced into genomic DNA by DNA damage caused by reactive oxygen species, ionizing radiation, and/or alkylating agents. Studies on the mechanisms of DNA repair, such as the base excision repair (BER) pathway, or studies which measure DNA repair capacity, may use a USP of the invention to inhibit repair of the uracil.

Additionally, USPs described herein may be useful for the treatment of various cancers. For example, fluoropyrimidines including 5-fluorouracil (5-FU) and its deoxyribonucleoside metabolite 5-fluorodeoxyuridine (5-FdU) have been widely used in the treatment of various solid tumors, including colorectal cancer. 5-FdU is active through the inhibition of thymidylate synthase, which consequently introduces uracil and 5-FU incorporation into the genome of the cell. As described above, base repair enzymes such as UDG recognize uracil nucleobases in the genomic DNA and remove them. Yan et al (2016; *Oncotarget* 7 (37): 59299-59313) found that UDG depleted cells were arrested and displayed sustained DNA damage following 5-FdU treatment, indicating that UDG's actions in removal of uracil and 5-FU played a role in the effectiveness of the 5-FdU treatment of the tumor. Delivery of a USP of the invention in combination with fluoropyrimidines may enhance the effectiveness of this treatment of tumors. Thus, pharmaceutical compositions comprising a presently disclosed USP and a fluoropyrimidine are provided, along with methods of treating a cancer with effective amounts of such pharmaceutical compositions. Fluoropyrimidines are a class of anti-cancer antimetabolites that includes capecitabine, capecitabine, doxifluridine, fluorouracil, 5-fluorodeoxyuridine, and tegafur.

The article "a" and "an" are used herein to refer to one or more than one (*i.e.*, to at least one) of the grammatical object of the article. By way of example, "a polypeptide" means one or more polypeptides.

All publications and patent applications mentioned in the specification are indicative of the level of those skilled in the art to which this disclosure pertains. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended claims.

Non-limiting embodiments include:

1. An isolated polypeptide comprising an amino acid sequence having:
 - a) at least 80% sequence identity to any one of SEQ ID NOs: 1, 2, 4, 5, and 7-15;

b) at least 81% sequence identity to SEQ ID NO: 3 or 16; or

c) at least 82% sequence identity to SEQ ID NO: 6;

wherein said polypeptide has uracil stabilizing activity and wherein said polypeptide further comprises a heterologous amino acid sequence.

2. The isolated polypeptide of embodiment 1, wherein the amino acid sequence has at least 85% sequence identity to any one of SEQ ID NOs: 1-16.

3. The isolated polypeptide of embodiment 1, wherein the amino acid sequence has at least 90% sequence identity to any one of SEQ ID NOs: 1-16.

4. The isolated polypeptide of embodiment 1, wherein the amino acid sequence has at least 95% sequence identity to any one of SEQ ID NOs: 1-16.

5. The isolated polypeptide of embodiment 1, wherein the amino acid sequence has 100% sequence identity to any one of SEQ ID NOs: 1-16.

6. The isolated polypeptide of embodiment 1, wherein the polypeptide has the sequence of any one of SEQ ID NOs: 33-39.

7. A pharmaceutical composition comprising a non-naturally occurring pharmaceutically acceptable carrier and a polypeptide comprising an amino acid sequence having:

a) at least 80% sequence identity to any one of SEQ ID NOs: 1, 2, 4, 5, and 7-15;

b) at least 81% sequence identity to SEQ ID NO: 3 or 16; or

c) at least 82% sequence identity to SEQ ID NO: 6;

wherein said polypeptide has uracil stabilizing activity.

8. The pharmaceutical composition of embodiment 7, wherein the polypeptide comprises an amino acid sequence having at least 85% sequence identity to any one of SEQ ID NOs: 1-16.

9. The pharmaceutical composition of embodiment 7, wherein the polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1-16.

10. The pharmaceutical composition of embodiment 7, wherein the polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 1-16.

11. The pharmaceutical composition of embodiment 7, wherein the polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1-16.

12. A pharmaceutical composition comprising a non-naturally occurring pharmaceutically acceptable carrier and a nucleic acid molecule comprising a polynucleotide encoding a polypeptide comprising an amino acid sequence having:

a) at least 80% sequence identity to any one of SEQ ID NOs: 1, 2, 4, 5, and 7-15;

b) at least 81% sequence identity to SEQ ID NO: 3 or 16; or

c) at least 82% sequence identity to SEQ ID NO: 6;

wherein said polypeptide has uracil stabilizing activity.

13. The pharmaceutical composition of embodiment 12, wherein the polypeptide comprises an amino acid sequence having at least 85% sequence identity to any one of SEQ ID NOs: 1-16.

14. The pharmaceutical composition of embodiment 12, wherein the polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1-16.

15. The pharmaceutical composition of embodiment 12, wherein the polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 1-16.

5 16. The pharmaceutical composition of embodiment 12, wherein the polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1-16.

17. The pharmaceutical composition of embodiment 7 or 12, wherein the polypeptide has the sequence of any one of SEQ ID NOs: 33-39.

10 18. The pharmaceutical composition of any one of embodiments 7-17, further comprising a fluoropyrimidine.

19. A nucleic acid molecule comprising a polynucleotide encoding a polypeptide comprising an amino acid sequence having:

a) at least 80% sequence identity to any one of SEQ ID NOs: 1, 2, 4, 5, and 7-15;

b) at least 81% sequence identity to SEQ ID NO: 3 or 16; or

15 c) at least 82% sequence identity to SEQ ID NO: 6;

wherein said polypeptide has uracil stabilizing activity; and

wherein said nucleic acid molecule further comprises a heterologous promoter operably linked to said polynucleotide.

20 20. The nucleic acid molecule of embodiment 19, wherein the polypeptide comprises an amino acid sequence having at least 85% sequence identity to any one of SEQ ID NOs: 1-16.

21. The nucleic acid molecule of embodiment 19, wherein the polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1-16.

22. The nucleic acid molecule of embodiment 19, wherein the polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 1-16.

25 23. The nucleic acid molecule of embodiment 19, wherein the polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1-16.

24. The nucleic acid molecule of embodiment 19, wherein the polypeptide has the sequence of any one of SEQ ID NOs: 33-39.

30 25. A composition comprising a fluoropyrimidine and a polypeptide comprising an amino acid sequence having:

a) at least 80% sequence identity to any one of SEQ ID NOs: 1, 2, 4, 5, and 7-15;

b) at least 81% sequence identity to SEQ ID NO: 3 or 16; or

c) at least 82% sequence identity to SEQ ID NO: 6;

wherein said polypeptide has uracil stabilizing activity.

35 26. The composition of embodiment 25, wherein the polypeptide comprises an amino acid sequence having at least 85% sequence identity to any one of SEQ ID NOs: 1-16.

27. The composition of embodiment 25, wherein the polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1-16.

28. The composition of embodiment 25, wherein the polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 1-16.

29. The composition of embodiment 25, wherein the polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1-16.

30. A composition comprising a fluoropyrimidine and a nucleic acid molecule encoding a polypeptide comprising an amino acid sequence having:

a) at least 80% sequence identity to any one of SEQ ID NOs: 1, 2, 4, 5, and 7-15;

b) at least 81% sequence identity to SEQ ID NO: 3 or 16; or

c) at least 82% sequence identity to SEQ ID NO: 6;

wherein said polypeptide has uracil stabilizing activity.

31. The composition of embodiment 30, wherein the polypeptide comprises an amino acid sequence having at least 85% sequence identity to any one of SEQ ID NOs: 1-16.

32. The composition of embodiment 30, wherein the polypeptide comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 1-16.

33. The composition of embodiment 30, wherein the polypeptide comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 1-16.

34. The composition of embodiment 30, wherein the polypeptide comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 1-16.

35. The composition of embodiment 25 or 30, wherein the polypeptide has the sequence of any one of SEQ ID NOs: 33-39.

36. A fusion protein comprising: (i) a DNA-binding polypeptide; (ii) a deaminase; and (iii) at least one uracil stabilizing polypeptide (USP) having at least 80% sequence identity to any one of SEQ ID NOs: 1-16.

37. The fusion protein of embodiment 36, wherein at least one USP has at least 85% sequence identity to any one of SEQ ID NOs: 1-16.

38. The fusion protein of embodiment 36, wherein at least one USP has at least 90% sequence identity to any one of SEQ ID NOs: 1-16.

39. The fusion protein of embodiment 36, wherein at least one USP has at least 95% sequence identity to any one of SEQ ID NOs: 1-16.

40. The fusion protein of embodiment 36, wherein at least one USP has 100% sequence identity to any one of SEQ ID NOs: 1-16.

41. The fusion protein of embodiment 36, wherein the USP has the sequence of any one of SEQ ID NOs: 33-39.

42. The fusion protein of embodiment 36 or 41, wherein the deaminase is a cytidine deaminase.

43. The fusion protein of embodiment 42, wherein the cytidine deaminase is an activation-induced cytidine deaminase (AID) or a member of the apolipoprotein B mRNA-editing complex (APOBEC) family of deaminases.

44. The fusion protein of embodiment 43, wherein the cytidine deaminase comprises an amino acid sequence having at least 80% sequence identity to any one of SEQ ID NOs: 47, 48 and 76-94.

45. The fusion protein of embodiment 43, wherein the cytidine deaminase comprises an amino acid sequence having at least 85% sequence identity to any one of SEQ ID NOs: 47, 48 and 76-94.

46. The fusion protein of embodiment 43, wherein the cytidine deaminase comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 47, 48 and 76-94.

47. The fusion protein of embodiment 43, wherein the cytidine deaminase comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 47, 48 and 76-94.

48. The fusion protein of embodiment 43, wherein the cytidine deaminase comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 47, 48 and 76-94.

49. The fusion protein of any one of embodiments 36-44, wherein the DNA-binding polypeptide is a meganuclease, zinc finger fusion protein, or a TALEN.

50. The fusion protein of any one of embodiments 36-44, wherein the DNA-binding polypeptide is an RNA-guided, DNA-binding polypeptide.

51. The fusion protein of embodiment 50, wherein the RNA-guided, DNA-binding polypeptide is an RNA-guided nuclease polypeptide (RGN).

52. The fusion protein of embodiment 51, wherein the RGN is a Type II CRISPR-Cas polypeptide.

53. The fusion protein of embodiment 51, wherein the RGN is a Type V CRISPR-Cas polypeptide.

54. The fusion protein of embodiment 51, wherein the RGN comprises an amino acid sequence having at least 80% sequence identity to any one of SEQ ID NOs: 40 and 95-142.

55. The fusion protein of embodiment 51, wherein the RGN comprises an amino acid sequence having at least 85% sequence identity to any one of SEQ ID NOs: 40 and 95-142.

56. The fusion protein of embodiment 51, wherein the RGN comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 40 and 95-142.

57. The fusion protein of embodiment 51, wherein the RGN comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 40 and 95-142.

58. The fusion protein of embodiment 51, wherein the RGN comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 40 and 95-142.

59. The fusion protein of any one of embodiments 51-58, wherein the RGN is an RGN nickase.

60. The fusion protein of any one of embodiments 51-59, wherein the fusion protein comprises an RGN, a cytidine deaminase, and a USP.

61. The fusion protein of embodiment 60, wherein the fusion protein comprises an RGN having at least 80% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, a cytidine deaminase having at least 80% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and a USP having at least 80% sequence identity to any one of SEQ ID NOs: 1-16.

5 62. The fusion protein of embodiment 60, wherein the fusion protein comprises an RGN having at least 85% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, a cytidine deaminase having at least 85% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and a USP having at least 85% sequence identity to any one of SEQ ID NOs: 1-16.

10 63. The fusion protein of embodiment 60, wherein the fusion protein comprises an RGN having at least 90% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, a cytidine deaminase having at least 90% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and a USP having at least 90% sequence identity to any one of SEQ ID NOs: 1-16.

15 64. The fusion protein of embodiment 60, wherein the fusion protein comprises an RGN having at least 95% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, a cytidine deaminase having at least 95% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and a USP having at least 95% sequence identity to any one of SEQ ID NOs: 1-16.

20 65. The fusion protein of embodiment 60, wherein the fusion protein comprises an RGN having 100% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, a cytidine deaminase having 100% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and a USP having 100% sequence identity to any one of SEQ ID NOs: 1-16.

66. The fusion protein of any of embodiments 36-66, wherein the fusion protein further comprises at least one nuclear localization signal (NLS).

25 67. A nucleic acid molecule comprising a polynucleotide encoding a fusion protein comprising: (i) a DNA-binding polypeptide; (ii) a deaminase; and (iii) at least one uracil stabilizing polypeptide (USP), wherein the USP is encoded by a nucleotide sequence that:

- a) has at least 80% sequence identity to any one of SEQ ID NOs: 17-32,
- b) is set forth in any one of SEQ ID NOs: 17-32,
- c) encodes an amino acid sequence at least 80% identical to SEQ ID NOs: 1-16 and further possesses the sequence of any one of SEQ ID NOs: 33-39,
- 30 d) encodes an amino acid sequence at least 80% identical to an amino acid sequence set forth in any one of SEQ ID NOs: 1-16, or
- e) encodes an amino acid sequence set forth in any one of SEQ ID NOs: 1-16.

68. The nucleic acid molecule of embodiment 67, wherein the USP is encoded by a nucleotide sequence that:

- 35 a) has at least 85% sequence identity to any one of SEQ ID NOs: 17-32,
- b) is set forth in any one of SEQ ID NOs: 17-32,

c) encodes an amino acid sequence at least 85% identical to SEQ ID NOs: 1-16 and further possesses the sequence of any one of SEQ ID NOs: 33-39,

d) encodes an amino acid sequence at least 85% identical to an amino acid sequence set forth in any one of SEQ ID NOs: 1-16, or

5 e) encodes an amino acid sequence set forth in any one of SEQ ID NOs: 1-16.

69. The nucleic acid molecule of embodiment 67, wherein the USP is encoded by a nucleotide sequence that:

a) has at least 90% sequence identity to any one of SEQ ID NOs: 17-32,

b) is set forth in any one of SEQ ID NOs: 17-32,

10 c) encodes an amino acid sequence at least 90% identical to SEQ ID NOs: 1-16 and further possesses the sequence of any one of SEQ ID NOs: 33-39,

d) encodes an amino acid sequence at least 90% identical to an amino acid sequence set forth in any one of SEQ ID NOs: 1-16, or

e) encodes an amino acid sequence set forth in any one of SEQ ID NOs: 1-16.

15 70. The nucleic acid molecule of embodiment 67, wherein the USP is encoded by a nucleotide sequence that:

a) has at least 95% sequence identity to any one of SEQ ID NOs: 17-32,

b) is set forth in any one of SEQ ID NOs: 17-32,

20 c) encodes an amino acid sequence at least 95% identical to SEQ ID NOs: 1-16 and further possesses the sequence of any one of SEQ ID NOs: 33-39,

d) encodes an amino acid sequence at least 95% identical to an amino acid sequence set forth in any one of SEQ ID NOs: 1-16, or

e) encodes an amino acid sequence set forth in any one of SEQ ID NOs: 1-16.

25 71. The nucleic acid molecule of embodiment 67, wherein the USP is encoded by a nucleotide sequence that:

a) is set forth in any one of SEQ ID NOs: 17-32,

b) encodes an amino acid sequence 100% identical to SEQ ID NOs: 1-16 and further possesses the sequence of any one of SEQ ID NOs: 33-39, or

c) encodes an amino acid sequence set forth in any one of SEQ ID NOs: 1-16.

30 72. The nucleic acid molecule of embodiment 67, wherein the deaminase is a cytidine deaminase.

73. The nucleic acid molecule of embodiment 72, wherein the cytidine deaminase is an activation-induced cytidine deaminase (AID) or a member of the apolipoprotein B mRNA-editing complex (APOBEC) family of deaminases.

35 74. The nucleic acid molecule of embodiment 73, wherein the cytidine deaminase comprises an amino acid sequence having at least 80% sequence identity to any one of SEQ ID NOs: 47, 48 and 76-94.

75. The nucleic acid molecule of embodiment 73, wherein the cytidine deaminase comprises an amino acid sequence having at least 85% sequence identity to any one of SEQ ID NOs: 47, 48 and 76-94.

76. The nucleic acid molecule of embodiment 73, wherein the cytidine deaminase comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 47, 48 and 76-94.

5 77. The nucleic acid molecule of embodiment 73, wherein the cytidine deaminase comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 47, 48 and 76-94.

78. The nucleic acid molecule of embodiment 73, wherein the cytidine deaminase comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 47, 48 and 76-94.

10 79. The nucleic acid molecule of any one of embodiments 67-78, wherein the DNA-binding polypeptide is a meganuclease, zinc finger fusion protein, or a TALEN.

80. The nucleic acid molecule of any one of embodiments 67-78, wherein the DNA-binding polypeptide is an RNA-guided, DNA-binding polypeptide.

81. The nucleic acid molecule of embodiment 80, wherein the RNA-guided, DNA-binding polypeptide is an RNA-guided nuclease polypeptide (RGN).

15 82. The nucleic acid molecule of embodiment 81, wherein the RGN is a Type II CRISPR-Cas polypeptide.

83. The nucleic acid molecule of embodiment 81, wherein the RGN is a Type V CRISPR-Cas polypeptide.

20 84. The nucleic acid molecule of embodiment 81, wherein the RGN comprises an amino acid sequence having at least 80% sequence identity to any one of SEQ ID NOs: 40 and 95-142.

85. The nucleic acid molecule of embodiment 82, wherein the RGN comprises an amino acid sequence having at least 85% sequence identity to any one of SEQ ID NOs: 40 and 95-142.

86. The nucleic acid molecule of embodiment 81, wherein the RGN comprises an amino acid sequence having at least 90% sequence identity to any one of SEQ ID NOs: 40 and 95-142.

25 87. The nucleic acid molecule of embodiment 81, wherein the RGN comprises an amino acid sequence having at least 95% sequence identity to any one of SEQ ID NOs: 40 and 95-142.

88. The nucleic acid molecule of embodiment 81, wherein the RGN comprises an amino acid sequence having 100% sequence identity to any one of SEQ ID NOs: 40 and 95-142.

30 89. The nucleic acid molecule of any one of embodiments 81-88, wherein the RGN is an RGN nickase.

90. The nucleic acid molecule of any one of embodiments 81-89, wherein the fusion protein comprises an RGN, a cytidine deaminase, and a USP.

35 91. The nucleic acid molecule of embodiment 90, wherein the fusion protein comprises an RGN having at least 80% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, a cytidine deaminase having at least 80% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and a USP having at least 80% sequence identity to any one of SEQ ID NOs: 1-16.

92. The nucleic acid molecule of embodiment 90, wherein the fusion protein comprises an RGN having at least 85% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, a cytidine deaminase having at least 85% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and a USP having at least 85% sequence identity to any one of SEQ ID NOs: 1-16.

5 93. The nucleic acid molecule of embodiment 90, wherein the fusion protein comprises an RGN having at least 90% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, a cytidine deaminase having at least 90% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and a USP having at least 90% sequence identity to any one of SEQ ID NOs: 1-16.

10 94. The nucleic acid molecule of embodiment 90, wherein the fusion protein comprises an RGN having at least 95% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, a cytidine deaminase having at least 95% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and a USP having at least 95% sequence identity to any one of SEQ ID NOs: 1-16.

15 95. The nucleic acid molecule of embodiment 90, wherein the fusion protein comprises an RGN having 100% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, a cytidine deaminase having 100% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and a USP having 100% sequence identity to any one of SEQ ID NOs: 1-16.

96. The nucleic acid molecule of any of embodiments 67-95, wherein the polynucleotide encoding the fusion protein is operably linked at its 5' end to a heterologous promoter.

20 97. The nucleic acid molecule of any of embodiments 67-95, wherein the polynucleotide encoding the fusion protein is operably linked at its 3' end to a heterologous terminator.

98. The nucleic acid molecule of any of embodiments 67-97, wherein the fusion protein comprises one or more nuclear localization signals.

99. The nucleic acid molecule of any of embodiments 67-98, wherein the fusion protein is codon optimized for expression in a eukaryotic cell.

25 100. The nucleic acid molecule of any of embodiments 67-99, wherein the fusion protein is codon optimized for expression in a prokaryotic cell.

101. A vector comprising the nucleic acid molecule of any one of embodiments 67-100.

102. A vector comprising the nucleic acid molecule of any one of embodiments 81-95, further comprising at least one nucleotide sequence encoding a guide RNA (gRNA) capable of hybridizing to a target sequence.

30 103. The vector of embodiment 102, wherein the gRNA is a single guide RNA.

104. The vector of embodiment 102, wherein the gRNA is a dual guide RNA.

105. A cell comprising the fusion protein of any of embodiments 36-66.

106. A cell comprising the fusion protein of any one of embodiments 51-66, wherein the cell further comprises a guide RNA.

107. A cell comprising the nucleic acid molecule of any of embodiments 67-100.

108. A cell comprising the vector of any of embodiments 101 through 104.

109. The cell of any one of embodiments 105-108, wherein the cell is a prokaryotic cell.

110. The cell of any one of embodiments 105-108, wherein the cell is a eukaryotic cell.

111. The cell of embodiment 110, wherein the cell is an insect, avian, or mammalian cell.

112. The cell of embodiment 110, wherein the cell is a plant or fungal cell.

5 113. A pharmaceutical composition comprising a pharmaceutically acceptable carrier and the nucleic acid molecule of any one of embodiments 19-24 and 67-100, the composition of any one of embodiments 25-35, the fusion protein of any one of embodiments 36-66, the vector of any one of embodiments 101-104, or the cell of any one of embodiments 105-111.

10 114. A method for making a fusion protein comprising culturing the cell of any one of embodiments 105-112 under conditions in which the fusion protein is expressed.

115. A method for making a fusion protein comprising introducing into a cell the nucleic acid molecule of any of embodiments 67-100 or a vector of any one of embodiments 101-104 and culturing the cell under conditions in which the fusion protein is expressed.

116. The method of embodiment 114 or 115, further comprising purifying said fusion protein.

15 117. A method for making an RGN fusion ribonucleoprotein complex, comprising introducing into a cell the nucleic acid molecule of any one of embodiments 81-95 and a nucleic acid molecule comprising an expression cassette encoding for a guide RNA, or the vector of any of embodiments 102-104, and culturing the cell under conditions in which the fusion protein and the gRNA are expressed and form an RGN fusion ribonucleoprotein complex.

20 118. The method of embodiment 117, further comprising purifying said RGN fusion ribonucleoprotein complex.

119. A system for modifying a target DNA molecule comprising a target DNA sequence, said system comprising:

25 a) a fusion protein comprising an RNA-guided nuclease polypeptide (RGN), a cytidine deaminase, and at least one uracil stabilizing polypeptide (USP), wherein the USP is at least 80% identical to any one of SEQ ID NOs: 1-16, or a nucleotide sequence encoding said fusion protein; and

b) one or more guide RNAs capable of hybridizing to said target DNA sequence or one or more nucleotide sequences encoding the one or more guide RNAs (gRNAs);

30 wherein said nucleotide sequences encoding the one or more guide RNAs and encoding the fusion protein are each operably linked to a promoter heterologous to said nucleotide sequence; and

wherein the one or more guide RNAs are capable of forming a complex with the fusion protein in order to direct said fusion protein to bind to said target DNA sequence and modify the target DNA molecule.

35 120. The system of embodiment 119, wherein the USP is at least 85% identical to any one of SEQ ID NOs: 1-16.

121. The system of embodiment 119, wherein the USP is at least 90% identical to any one of SEQ ID NOs: 1-16.

122. The system of embodiment 119, wherein the USP is at least 95% identical to any one of SEQ ID NOs: 1-16.

123. The system of embodiment 119, wherein the USP is 100% identical to any one of SEQ ID NOs: 1-16.

5 124. The system of embodiment 119, wherein the USP comprises the sequence set forth in any one of SEQ ID NOs: 33-39.

125. The system of any one of embodiments 119-124, wherein the target DNA sequence is located adjacent to a protospacer adjacent motif (PAM) that is recognized by the RGN.

10 126. The system of any one of embodiments 119-125, wherein the target DNA molecule is within a cell.

127. The system of embodiment 126, wherein the cell is a eukaryotic cell.

128. The system of embodiment 127, wherein the eukaryotic cell is a plant cell.

129. The system of embodiment 127, wherein the eukaryotic cell is a mammalian cell.

130. The system of embodiment 127, wherein the eukaryotic cell is an insect cell.

15 131. The system of embodiment 126, wherein the cell is a prokaryotic cell.

132. The system of any one of embodiments 119-131, wherein the RGN of the fusion protein is a Type II CRISPR-Cas polypeptide.

133. The system of any one of embodiments 119-131, wherein the RGN of the fusion protein is a Type V CRISPR-Cas polypeptide.

20 134. The system of any one of embodiments 119-131, wherein the RGN of the fusion protein is at least 80% identical to any one of SEQ ID NOs: 40 and 95-142.

135. The system of any one of embodiments 119-131, wherein the RGN of the fusion protein is at least 85% identical to any one of SEQ ID NOs: 40 and 95-142.

25 136. The system of any one of embodiments 119-131, wherein the RGN of the fusion protein is 90% identical to any one of SEQ ID NOs: 40 and 95-142.

137. The system of any one of embodiments 119-131, wherein the RGN of the fusion protein is 95% identical to any one of SEQ ID NOs: 40 and 95-142.

138. The system of any one of embodiments 119-131, wherein the RGN of the fusion protein is 100% identical to any one of SEQ ID NOs: 40 and 95-142.

30 139. The system of any one of embodiments 119-138, wherein the RGN of the fusion protein is an RGN nickase.

140. The system of any of embodiments 119-139, wherein the cytidine deaminase is at least 80% identical to any one of SEQ ID NOs: 47, 48 and 76-94.

35 141. The system of any of embodiments 119-139, wherein the cytidine deaminase is at least 85% identical to any one of SEQ ID NOs: 47, 48 and 76-94.

142. The system of any of embodiments 119-139, wherein the cytidine deaminase is at least 90% identical to any one of SEQ ID NOs: 47, 48 and 76-94.

143. The system of any of embodiments 119-139, wherein the cytidine deaminase is at least 95% identical to any one of SEQ ID NOs: 47, 48 and 76-94.

144. The system of any of embodiments 119-139, wherein the cytidine deaminase is 100% identical to any one of SEQ ID NOs: 47, 48 and 76-94.

5 145. The system of any of embodiments 119-144, wherein the fusion protein comprises a RGN having at least 80% sequence identity to any one of SEQ ID NOs: 40, 41, 95 and 142, a cytidine deaminase having at least 80% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and a USP having at least 80% sequence identity to any one of SEQ ID NOs: 1-16.

10 146. The system of any of embodiments 119-144, wherein the fusion protein comprises a RGN having at least 85% sequence identity to any one of SEQ ID NOs: 40, 41, 95 and 142, a cytidine deaminase having at least 85% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and a USP having at least 85% sequence identity to any one of SEQ ID NOs: 1-16.

15 147. The system of any of embodiments 119-144, wherein the fusion protein comprises a RGN having at least 90% sequence identity to any one of SEQ ID NOs: 40, 41, 95 and 142, a cytidine deaminase having at least 90% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and a USP having at least 90% sequence identity to any one of SEQ ID NOs: 1-16.

20 148. The system of any of embodiments 119-144, wherein the fusion protein comprises a RGN having at least 95% sequence identity to any one of SEQ ID NOs: 40, 41, 95 and 142, a cytidine deaminase having at least 95% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and a USP having at least 95% sequence identity to any one of SEQ ID NOs: 1-16.

149. The system of any of embodiments 119-144, wherein the fusion protein comprises a RGN having 100% sequence identity to any one of SEQ ID NOs: 40, 41, 95 and 142, a cytidine deaminase having 100% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and a USP having 100% sequence identity to any one of SEQ ID NOs: 1-16.

25 150. The system of any of embodiments 119-149, wherein the fusion protein comprises one or more nuclear localization signals.

151. The system of any of embodiments 119-150, wherein the fusion protein is codon optimized for expression in a eukaryotic cell.

30 152. The system of any of embodiments 119-151, wherein nucleotide sequences encoding the one or more guide RNAs and the nucleotide sequence encoding a fusion protein are located on one vector.

153. A method for modifying a target DNA molecule comprising a target DNA sequence, said method comprising delivering a system according to any one of embodiments 119-152 to said target DNA molecule or a cell comprising the target DNA molecule.

35 154. The method of embodiment 153, wherein said modified target DNA molecule comprises a C>T mutation of at least one nucleotide within the target DNA molecule.

155. The method of embodiment 153, wherein said modified target DNA molecule comprises a C>T mutation of at least one nucleotide within the target DNA sequence.

156. A method for modifying a target DNA molecule comprising a target sequence comprising:

a) assembling an RGN-deaminase-USP ribonucleotide complex *in vitro* by combining:

i) one or more guide RNAs capable of hybridizing to the target DNA sequence; and

ii) a fusion protein comprising an RNA-guided nuclease polypeptide (RGN), a cytidine

deaminase, and at least one uracil stabilizing polypeptide (USP), wherein the USP is at least 80% identical to any one of SEQ ID NOs: 1-16;

under conditions suitable for formation of the RGN-deaminase-USP ribonucleotide complex; and

b) contacting said target DNA molecule or a cell comprising said target DNA molecule with the *in vitro*-assembled RGN-deaminase-USP ribonucleotide complex;

wherein the one or more guide RNAs hybridize to the target DNA sequence, thereby directing said fusion protein to bind to said target DNA sequence and modification of the target DNA molecule occurs.

157. The method of embodiment 156, wherein the USP is at least 85% identical to any one of SEQ ID NOs: 1-16.

158. The method of embodiment 156, wherein the USP is at least 90% identical to any one of SEQ ID NOs: 1-16.

159. The method of embodiment 156, wherein the USP is at least 95% identical to any one of SEQ ID NOs: 1-16.

160. The method of embodiment 156, wherein the USP is 100% identical to any one of SEQ ID NOs: 1-16.

161. The method of embodiment 156, wherein the USP comprises the sequence of any one of SEQ ID NOs: 33-39.

162. The method of any one of embodiments 156-161, wherein said modified target DNA molecule comprises a C>T mutation of at least one nucleotide within the target DNA molecule.

163. The method of any one of embodiments 156-161, wherein said modified target DNA molecule comprises a C>T mutation of at least one nucleotide within the target DNA sequence.

164. The method of any one of embodiments 156-163, wherein the RGN of the fusion protein is a Type II CRISPR-Cas polypeptide.

165. The method of any of embodiments 156-163, wherein the RGN of the fusion protein is a Type V CRISPR-Cas polypeptide.

166. The method of any of embodiments 156-163, wherein the RGN of the fusion protein is at least 80% identical to any one of SEQ ID NOs: 40 and 95-142.

167. The method of any of embodiments 156-163, wherein the RGN of the fusion protein is at least 85% identical to any one of SEQ ID NOs: 40 and 95-142.

168. The method of any of embodiments 156-163, wherein the RGN of the fusion protein is at least 90% identical to any one of SEQ ID NOs: 40 and 95-142.

169. The method of any of embodiments 156-163, wherein the RGN of the fusion protein is at least 95% identical to any one of SEQ ID NOs: 40 and 95-142.

170. The method of any of embodiments 156-163, wherein the RGN of the fusion protein is 100% identical to any one of SEQ ID NOs: 40 and 95-142.

171. The method of any of embodiments 156-170, wherein the RGN of the fusion protein is an RGN nickase.

5 172. The method of any of embodiments 156-171, wherein the cytidine deaminase is at least 80% identical to any one of SEQ ID NOs: 47, 48 and 76-94.

173. The method of any of embodiments 156-171, wherein the cytidine deaminase is at least 85% identical to any one of SEQ ID NOs: 47, 48 and 76-94.

10 174. The method of any of embodiments 156-171, wherein the cytidine deaminase is at least 90% identical to any one of SEQ ID NOs: 47, 48 and 76-94.

175. The method of any of embodiments 156-171, wherein the cytidine deaminase is at least 95% identical to any one of SEQ ID NOs: 47, 48 and 76-94.

176. The method of any of embodiments 156-171, wherein the cytidine deaminase is 100% identical to any one of SEQ ID NOs: 47, 48 and 76-94.

15 177. The method of any one of embodiments 156-176, wherein the fusion protein comprises an RGN having at least 80% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, a cytidine deaminase having at least 80% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and a USP having at least 80% sequence identity to any one of SEQ ID NOs: 1-16.

20 178. The method of any one of embodiments 156-176, wherein the fusion protein comprises an RGN having at least 85% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, a cytidine deaminase having at least 85% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and a USP having at least 85% sequence identity to any one of SEQ ID NOs: 1-16.

25 179. The method of any one of embodiments 156-176, wherein the fusion protein comprises an RGN having at least 90% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, a cytidine deaminase having at least 90% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and a USP having at least 90% sequence identity to any one of SEQ ID NOs: 1-16.

30 180. The method of any one of embodiments 156-176, wherein the fusion protein comprises an RGN having at least 95% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, a cytidine deaminase having at least 95% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and a USP having at least 95% sequence identity to any one of SEQ ID NOs: 1-16.

181. The method of any one of embodiments 156-176, wherein the fusion protein comprises an RGN having 100% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, a cytidine deaminase having 100% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and a USP having 100% sequence identity to any one of SEQ ID NOs: 1-16.

35 182. The method of any of embodiments 156-181, wherein the fusion protein comprises one or more nuclear localization signals.

183. The method of any of embodiments 156-182, wherein the fusion protein is codon optimized for expression in a eukaryotic cell.

184. The method of any of embodiments 156-183, wherein said target DNA sequence is located adjacent to a protospacer adjacent motif (PAM).

185. The method of any of embodiments 156-184, wherein the target DNA molecule is within a cell.

186. The method of embodiment 185, wherein the cell is a eukaryotic cell.

187. The method of embodiment 186, wherein the eukaryotic cell is a plant cell.

188. The method of embodiment 186, wherein the eukaryotic cell is a mammalian cell.

189. The method of embodiment 186, wherein the eukaryotic cell is an insect cell.

190. The method of embodiment 185, wherein the cell is a prokaryotic cell.

191. The method of any one of embodiments 185-190, further comprising selecting a cell comprising said modified DNA molecule.

192. A cell comprising a modified target DNA sequence according to the method of embodiment 191.

193. The cell of embodiment 192, wherein the cell is a eukaryotic cell.

194. The cell of embodiment 193, wherein the eukaryotic cell is a plant cell.

195. A plant comprising the cell of embodiment 194.

196. A seed comprising the cell of embodiment 194.

197. The cell of embodiment 193, wherein the eukaryotic cell is a mammalian cell.

198. The cell of embodiment 193, wherein the eukaryotic cell is an insect cell.

199. The cell of embodiment 192, wherein the cell is a prokaryotic cell.

200. A method for producing a genetically modified cell with a correction in a causal mutation for a genetically inherited disease, the method comprising introducing into the cell:

a) a fusion protein comprising an RNA-guided nuclease polypeptide (RGN), a cytidine deaminase, and at least one uracil stabilizing polypeptide (USP), wherein the USP is at least 80% identical to any one of SEQ ID NOs: 1-16, or a polynucleotide encoding said fusion protein, wherein said polynucleotide encoding the fusion protein is operably linked to a promoter to enable expression of the fusion protein in the cell; and

b) one or more guide RNAs (gRNA) capable of hybridizing to a target DNA sequence, or a polynucleotide encoding said gRNA, wherein said polynucleotide encoding the gRNA is operably linked to a promoter to enable expression of the gRNA in the cell;

whereby the fusion protein and gRNA target to the genomic location of the causal mutation and modify the genomic sequence to remove the causal mutation.

201. The method of embodiment 200, wherein the USP is at least 85% identical to any one of SEQ ID NOs: 1-16.

202. The method of embodiment 200, wherein the USP is at least 90% identical to any one of SEQ ID NOs: 1-16.

203. The method of embodiment 200, wherein the USP is at least 95% identical to any one of SEQ ID NOs: 1-16.

204. The method of embodiment 200, wherein the USP is 100% identical to any one of SEQ ID NOs: 1-16.

205. The method of embodiment 200, wherein said RGN of the fusion protein is a nickase.

206. The method of embodiment 200 or 205, wherein the USP comprises the sequence of any one of SEQ ID NOs: 33-39.

207. The method of any of embodiments 200-206, wherein the genome modification comprises introducing a C>T mutation of at least one nucleotide within the target DNA sequence.

208. The method of any of embodiments 200-207, wherein the cell is an animal cell.

209. The method of embodiment 208, wherein the animal cell is a mammalian cell.

210. The method of embodiment 209, wherein the cell is derived from a dog, cat, mouse, rat, rabbit, horse, sheep, goat, cow, pig, or human.

211. The method of any of embodiments 200-210, wherein the correction of the causal mutation comprises introducing a stop codon.

212. A composition comprising:

- a) a fusion protein comprising: (i) a DNA-binding polypeptide; and (ii) a deaminase; or a nucleic acid molecule encoding the fusion protein; and
- b) a uracil stabilizing polypeptide (USP) having at least 80% sequence identity to any one of SEQ ID NOs: 1-16; or a nucleic acid molecule encoding the USP.

213. The composition of embodiment 212, wherein the USP has at least 85% sequence identity to any one of SEQ ID NOs: 1-16.

214. The composition of embodiment 212, wherein the USP has at least 90% sequence identity to any one of SEQ ID NOs: 1-16.

215. The composition of embodiment 212, wherein the USP has at least 95% sequence identity to any one of SEQ ID NOs: 1-16.

216. The composition of embodiment 212, wherein the USP has 100% sequence identity to any one of SEQ ID NOs: 1-16.

217. The composition of embodiment 212, wherein the fusion protein further comprises a uracil stabilizing polypeptide (USP) having at least 80% sequence identity to any one of SEQ ID NOs: 1-16.

218. The composition of embodiment 212, wherein the fusion protein further comprises a uracil stabilizing polypeptide (USP) having at least 85% sequence identity to any one of SEQ ID NOs: 1-16.

219. The composition of embodiment 212, wherein the fusion protein further comprises a uracil stabilizing polypeptide (USP) having at least 90% sequence identity to any one of SEQ ID NOs: 1-16.

220. The composition of embodiment 212, wherein the fusion protein further comprises a uracil stabilizing polypeptide (USP) having at least 95% sequence identity to any one of SEQ ID NOs: 1-16.

221. The composition of embodiment 212, wherein the fusion protein further comprises a uracil stabilizing polypeptide (USP) having 100% sequence identity to any one of SEQ ID NOs: 1-16.

5 222. The composition of embodiment 212, wherein the DNA-binding polypeptide is a meganuclease, zinc finger fusion protein, or a TALEN.

223. The composition of embodiment 212, wherein the DNA-binding polypeptide is an RNA-guided, DNA-binding polypeptide.

10 224. The composition of embodiment 223, wherein the RNA-guided, DNA-binding polypeptide is an RNA-guided nuclease polypeptide (RGN).

225. The composition of embodiment 224, wherein the RGN is an RGN nickase.

226. A vector comprising a nucleic acid molecule encoding a fusion protein and a nucleic acid molecule encoding a uracil stabilizing polypeptide (USP), wherein said fusion protein comprises a DNA-binding polypeptide and a deaminase, and wherein said USP has at least 80% sequence identity to any one
15 of SEQ ID NOs: 1-16.

227. The vector of embodiment 226, said USP has at least 85% sequence identity to any one of SEQ ID NOs: 1-16.

228. The vector of embodiment 226, said USP has at least 90% sequence identity to any one of SEQ ID NOs: 1-16.

20 229. The vector of embodiment 226, said USP has at least 95% sequence identity to any one of SEQ ID NOs: 1-16.

230. The vector of embodiment 226, said USP has 100% sequence identity to any one of SEQ ID NOs: 1-16.

25 231. The vector of embodiment 226, wherein the fusion protein further comprises a uracil stabilizing polypeptide (USP) having at least 80% sequence identity to any one of SEQ ID NOs: 1-16.

232. The vector of embodiment 226, wherein the fusion protein further comprises a uracil stabilizing polypeptide (USP) having at least 85% sequence identity to any one of SEQ ID NOs: 1-16.

233. The vector of embodiment 226, wherein the fusion protein further comprises a uracil stabilizing polypeptide (USP) having at least 90% sequence identity to any one of SEQ ID NOs: 1-16.

30 234. The vector of embodiment 226, wherein the fusion protein further comprises a uracil stabilizing polypeptide (USP) having at least 95% sequence identity to any one of SEQ ID NOs: 1-16.

235. The vector of embodiment 226, wherein the fusion protein further comprises a uracil stabilizing polypeptide (USP) having 100% sequence identity to any one of SEQ ID NOs: 1-16.

35 236. The vector of embodiment 226, wherein the DNA-binding polypeptide is a meganuclease, zinc finger fusion protein, or a TALEN.

237. The vector of embodiment 226, wherein the DNA-binding polypeptide is an RNA-guided, DNA-binding polypeptide.

238. The vector of embodiment 237, wherein the RNA-guided, DNA-binding polypeptide is an RNA-guided nuclease polypeptide (RGN).

239. The vector of embodiment 238, wherein the RGN is an RGN nickase.

240. A cell comprising the vector of any one of embodiments 226-239.

241. A cell comprising:

a) a fusion protein comprising: (i) a DNA-binding polypeptide; and (ii) a deaminase; or a nucleic acid molecule encoding the fusion protein; and

b) a uracil stabilizing polypeptide (USP) having at least 80% sequence identity to any one of SEQ ID NOs: 1-16; or a nucleic acid molecule encoding the USP.

242. The cell of embodiment 241, wherein the USP has at least 85% sequence identity to any one of SEQ ID NOs: 1-16.

243. The cell of embodiment 241, wherein the USP has at least 90% sequence identity to any one of SEQ ID NOs: 1-16.

244. The cell of embodiment 241, wherein the USP has at least 95% sequence identity to any one of SEQ ID NOs: 1-16.

245. The cell of embodiment 241, wherein the USP has 100% sequence identity to any one of SEQ ID NOs: 1-16.

246. The cell of embodiment 241, wherein the fusion protein further comprises a uracil stabilizing polypeptide (USP) having at least 80% sequence identity to any one of SEQ ID NOs: 1-16.

247. The cell of embodiment 241, wherein the fusion protein further comprises a uracil stabilizing polypeptide (USP) having at least 85% sequence identity to any one of SEQ ID NOs: 1-16.

248. The cell of embodiment 241, wherein the fusion protein further comprises a uracil stabilizing polypeptide (USP) having at least 90% sequence identity to any one of SEQ ID NOs: 1-16.

249. The cell of embodiment 241, wherein the fusion protein further comprises a uracil stabilizing polypeptide (USP) having at least 95% sequence identity to any one of SEQ ID NOs: 1-16.

250. The cell of embodiment 241, wherein the fusion protein further comprises a uracil stabilizing polypeptide (USP) having 100% sequence identity to any one of SEQ ID NOs: 1-16.

251. The cell of embodiment 241, wherein the DNA-binding polypeptide is a meganuclease, zinc finger fusion protein, or a TALEN.

252. The cell of embodiment 241, wherein the DNA-binding polypeptide is an RNA-guided, DNA-binding polypeptide.

253. The cell of embodiment 252, wherein the RNA-guided, DNA-binding polypeptide is an RNA-guided nuclease polypeptide (RGN).

254. The cell of embodiment 253, wherein the RGN is an RGN nickase.

The following examples are offered by way of illustration and not by way of limitation.

EXPERIMENTAL

Example 1: Uracil Stabilizing Polypeptides

Amino acid sequences for the Uracil Stabilizing Polypeptides (USPs) of the invention are provided as SEQ ID NOs: 1-16, as shown in Table 1. All USPs disclosed are from *Staphylococcus* spp and range from 112 to 116 amino acids in length. The polypeptides of the USPs described herein possess 16.4% to 20.7% negative charges, with an expected pI of 3.86-4.55, except for APG05963 which has an expected pI of 5.26.

USPs APG06351, APG03399, APG04638, APG09242, APG02463, APG04080, APG01791, APG04001, and APG03327 share a unique consensus C-terminus sequence of “KEGGNDHE” (SEQ ID NO: 33). USPs APG05198 and APG05756 share a unique C-terminus sequence of “EKENYNNE” (SEQ ID NO: 34). APG05963 possesses a unique C-terminus sequence of “EKEKHKNE” (SEQ ID NO: 35); APG06702 possesses a unique C-terminus sequence of “DKGDDNHD” (SEQ ID NO: 36); APG05316 possesses a unique C-terminus sequence of “QKGGQ” (SEQ ID NO: 37); APG09230 possesses a unique C-terminus sequence of “KGENKYE” (SEQ ID NO: 38); and APG04100 possesses a unique C-terminus sequence of “KQGENNHE” (SEQ ID NO: 39).

Table 1: Uracil stabilizing polypeptides

USP ID	SEQ ID NO
APG03399	1
APG06702	2
APG05198	3
APG04638	4
APG06351	5
APG05963	6
APG09230	7
APG04100	8
APG04001	9
APG02463	10
APG09242	11
APG04080	12
APG05316	13
APG03327	14
APG01791	15
APG05756	16

Example 2: USP fusion proteins exhibit increased base editing activity in mammalian cells

Residues predicted to deactivate the RuvC domain of the RGN APG07433.1 (SEQ ID NO: 40; PCT publication WO 2019/236566, incorporated by reference herein) were identified and the RGN was modified to a nickase variant (nAPG07433.1; SEQ ID NO: 41). Fusion proteins comprising a cytidine deaminase, namely APG09980 (SEQ ID NO: 47; see PCT/US2019/068079, incorporated by reference herein) or APG07386CTD (SEQ ID NO: 48; see PCT/US2019/068079), were produced. Of the USPs in Table 1, three were selected for assaying for activity in fusion proteins, namely APG03399, APG06702, and APG05198. Deaminase, USP, and nRGN nucleotide sequences codon optimized for expression were synthesized as fusion proteins with an N-terminal nuclear localization tag and cloned into the pTwist CMV (Twist Biosciences) expression plasmid. A fusion protein lacking a USP of the invention comprises, starting at the amino terminus, the SV40 NLS (SEQ ID NO: 42) operably linked at the C-terminal end to 3X FLAG Tag (SEQ ID NO: 43), operably linked at the C-terminal end to a deaminase, operably linked at the C-terminal end to a peptide linker (SEQ ID NO: 44), operably linked at the C-terminal end to the nRGN (for example, nAPG07433.1, which is SEQ ID NO: 41), finally operably linked at the C-terminal end to the nucleoplasmin NLS (SEQ ID NO: 45). A fusion protein comprising a USP of the invention comprises, starting at the amino terminus, the SV40 NLS (SEQ ID NO: 42) operably linked at the C-terminal end to 3X FLAG Tag (SEQ ID NO: 43), operably linked at the C-terminal end to a deaminase, operably linked at the C-terminal end to a peptide linker (SEQ ID NO: 44), operably linked at the C-terminal end to the nRGN (for example, SEQ ID NO: 41), operably linked at the C-terminal end to a second linker sequence (SEQ ID NO: 46), operably linked at the C-terminal end to a USP of the invention, finally operably linked at the C-terminal end to the nucleoplasmin NLS (SEQ ID NO: 45). Table 2 shows the fusion proteins produced and tested for activity. All fusion proteins comprise at least one NLS and a 3X FLAG Tag, as described above.

Table 2: Fusion proteins assayed for C>T Editing

Fusion protein	SEQ ID NO
APG09980-nAPG07433.1	49
APG09980-nAPG07433.1-APG03399	50
APG09980-nAPG07433.1-APG06702	51
APG09980-nAPG07433.1-APG05198	52
APG07386-CTD-nAPG07433.1	53
APG07386-CTD-nAPG07433.1-APG03399	54
APG07386-CTD-nAPG07433.1-APG06702	55
APG07386-CTD-nAPG07433.1-APG05198	56

Expression plasmids comprising an expression cassette encoding for a sgRNA were also produced. Human genomic target sequences and the sgRNA sequences for guiding the fusion proteins to the genomic targets are indicated in Table 3. The genomic loci for each target sequence is also indicated.

Table 3: guide RNA sequences

sgRNA ID	Target sequence	sgRNA sequence	Genomic locus
SGN000143	57	63	69
SGN000169	58	64	70
SGN000173	59	65	71
SGN000929	60	66	72
SGN000930	61	67	73
SGN001101	62	68	74

500 ng of plasmid comprising an expression cassette comprising a coding sequence for a fusion protein shown in Table 2 and 500 ng of plasmid comprising an expression cassette encoding for an sgRNA shown in Table 3 were co-transfected into HEK293FT cells at 75-90% confluency in 24-well plates using Lipofectamine 2000 reagent (Life Technologies). Cells were then incubated at 37° C for 72 h. Following incubation, genomic DNA was then extracted using NucleoSpin 96 Tissue (Macherey-Nagel) following the manufacturer's protocol. The genomic region flanking the targeted genomic site was PCR amplified and products were purified using ZR-96 DNA Clean and Concentrator (Zymo Research) following the manufacturer's protocol. The purified PCR products were then sent for Next Generation Sequencing on Illumina MiSeq (2x250). Results were analyzed for indel formation or specific cytosine mutation out to +30 nucleotides, where the last nucleotide at the 3' end of the target sequence described in Table 3 is +1, and wherein the +30 nucleotide is 29 nucleotides upstream or 5' from the +1 nucleotide in the target sequence set forth in SEQ ID NOs: 57-62 (the target sequences of SEQ ID NOs: 57-62 are indicated as lower-case text within the genomic locus sequences set forth in SEQ ID NOs: 69-74, respectively).

Tables 4 through 15 show cytidine base editing for each combination of a fusion protein from Table 2 and a guide RNA from Table 3. Fusion proteins are identified by their SEQ ID NO. The numbering of cytidines (Cs) in Tables 4-15 are as in the preceding paragraph wherein the last nucleotide at the 3' end of the target sequence is +1 and the numbering proceeds in the 3' to 5' direction in the target sequence and corresponding genomic locus sequence. Interestingly, when comparing the activity of a deaminase-nRGN fusion protein with that of a corresponding deaminase-nRGN plus a USP, using the same guide RNA, the conversion of a cytidine to the desired thymidine is higher, with less conversion to an adenosine or guanosine.

Table 4: C>N Editing Rate using deaminase APG09980 and guide SGN000143

Fusion Protein (SEQ ID NO)		C10	C15	C20	C21	C22	C25
49	A	2.8	0.1	0	0.2	0.6	0.1
	G	1.1	0.1	0.1	0.1	0.2	0
	T	11	0.4	0.2	0.3	2	0.7
50	A	0	0	0	0	0	0
	G	0	0	0	0	0	0
	T	23.5	1.4	1.7	2	4.6	0.7

Fusion Protein (SEQ ID NO)		C10	C15	C20	C21	C22	C25
51	A	0.1	0	0	0	0	0
	G	0	0	0	0	0	0
	T	16.3	1.2	1.5	1.6	3.1	0.6
52	A	0.2	0	0	0	0	0
	G	0	0	0	0	0	0
	T	21.4	1.6	1.1	1.4	2.5	0.7

The results of Table 4 show that the rate of C>T editing at position C10 increased in samples with a USP.

5 **Table 5: C>N Editing Rate using deaminase APG07386 and guide SGN000143**

Fusion Protein (SEQ ID NO.)		C10	C15	C20	C21	C22	C25
53	A	0.1	0.4	0	0.5	1	0
	G	0	0.8	0	0	0.1	0
	T	0.5	4	0.3	0.7	3.8	0
54	A	0	0	0	0	0.1	0
	G	0	0	0	0	0	0
	T	1	4.5	1	1.7	3.1	0.1
55	A	0	0.1	0	0	0.1	0
	G	0	0	0	0	0	0
	T	1.6	9.7	2.5	3.2	5.8	0.2
56	A	0	0.1	0	0	0	0
	G	0	0	0	0	0	0
	T	1.9	10.2	2.6	3.1	7	0.2

The results of Table 5 show that the rate of C>T editing at position C15 increased in samples with a USP.

10 **Table 6: C>N Editing Rate using deaminase APG09980 and guide SGN000169**

Fusion Protein (SEQ ID NO.)		C9	C13	C15	C18	C20	C23
49	A	0.5	0.8	3.6	0.3	1.3	1.3
	G	1.5	1.8	8.3	0	0.2	0.1
	T	1.5	23.9	16.3	0.3	8.6	4.9
50	A	0	0	0.6	0	0.1	0.2
	G	0.2	0	2.5	0	0	0
	T	7	38.1	35.8	0.5	16.6	13.4
51	A	0.2	0.3	1.3	0.2	0.3	0.4
	G	0.6	0.6	2.1	0.1	0.1	0

Fusion Protein (SEQ ID NO.)		C9	C13	C15	C18	C20	C23
	T	9.2	42.4	40.8	0.8	20.4	16.6
52	A	0.8	1	2.4	0.7	0.6	1.5
	G	0.5	1.7	3.7	0.5	0.5	0.2
	T	8.5	37.8	35.7	1.2	17.1	14.9

The results of Table 6 show that the rate of C>T formation at positions C13 and C15 increased with the addition of a USP.

5 **Table 7: C>N Editing Rate using deaminase APG07386 and guide SGN000169**

Fusion Protein (SEQ ID NO.)		C9	C13	C15	C18	C20	C23
53	A	0.8	0.7	2	0	0.5	0.3
	G	5.4	0.8	10.2	0.1	0.1	0
	T	3.5	5.1	11.2	0.3	2.6	2.4
54	A	0.1	0.5	0.8	0	0.3	0.2
	G	0.7	0.2	2.5	0.2	0.1	0
	T	8.5	15.9	24.8	1.5	6.4	7.1
55	A	0.3	0.5	1.3	0.4	0.2	0.4
	G	2	0.5	3.5	0.1	0.2	0.1
	T	14.3	22.7	37.1	2	9.2	7.9
56	A	0.1	0.1	0.7	0	0.2	0
	G	0.7	0	2.9	0	0	0
	T	14.7	24	36.4	1.3	8.4	8.9

The results of Table 7 show that the rate of C>T formation at positions C13 and C15 increased with the addition of a USP.

10 **Table 8: C>N Editing Rate using deaminase APG09980 and guide SGN000930**

Fusion Protein (SEQ ID NO.)		C17	C19	C22
49	A	0.7	0.4	1.4
	G	34.5	0.4	0
	T	2.6	2.4	2.2
50	A	0.3	0.1	0.3
	G	8.6	0.3	0
	T	33	2.8	3.2
51	A	0.2	0.1	0.2
	G	5.1	0.1	0.1
	T	35.7	3.6	3.3

Fusion Protein (SEQ ID NO.)		C17	C19	C22
52	A	0.4	0.3	0.9
	G	11.2	0	0
	T	23.7	2.5	2.5

The results of Table 8 show that the rate of C>G editing at position C17 decreased in all samples with a USP in favor of C>T changes.

5

Table 9: C>N Editing Rate using deaminase APG07386 and guide SGN000930

Fusion Protein (SEQ ID NO.)		C17	C19	C22
53	A	0.2	0	0.4
	G	14.7	0	0.4
	T	1.1	0.2	2.2
54	A	0	0	0
	G	1.4	0	0
	T	9.1	0.4	3.7
55	A	0.4	0.1	0.3
	G	2.6	0.2	0.2
	T	11.8	0.9	4.5
56	A	0.1	0.2	0
	G	2.7	0	0.1
	T	17.4	0.8	6.7

The results of Table 9 show that the C>G editing at position C17 decreased in all samples with a USP in favor of C>T changes.

10

Table 10: C>N Editing Rate using deaminase APG09980 and guide SGN00173

Fusion Protein (SEQ ID NO.)		C1	C2	C3	C4	C7	C8	C10	C11	C17	C20
49	A	0	0	0.1	0.3	3.1	6.9	0.3	0	3.4	0.1
	G	0	0.1	0	0.2	0.3	1	0.1	0.4	3.4	0.1
	T	0.1	0	0	2.5	11.1	12.3	1.7	0.2	12.1	1.7
50	A	0	0.1	0	0.1	0.9	2.9	0	0.4	1.5	0
	G	0.1	0.1	0	0	0.5	0.8	0.4	0.3	1.9	0.2
	T	0.1	0.1	0.3	14	39	39.3	28.7	13	31.1	12
51	A	0	0	0	0	1.2	1.6	0	0	0.9	0
	G	0	0	0	0	0	1.4	0.1	0	2.2	0.1
	T	0	0	0.3	10.4	33.5	33.4	23.8	10.9	26.8	10.3

Fusion Protein (SEQ ID NO.)		C1	C2	C3	C4	C7	C8	C10	C11	C17	C20
52	A	0	0	0	0	1.8	1.2	0.3	0	1.5	0
	G	0	0	0	0	0	0.8	0	0.5	2.7	0
	T	0	0	0.6	10.5	33.1	36.8	26.7	13.2	28.6	9.6

The results of Table 10 show that the rate of C>T formation at positions C4, C7, C8, C10, C11, C17 and C20 increased with the addition of a USP.

5 **Table 11: C>N Editing Rate using deaminase APG07386 and guide SGN00173**

Fusion Protein (SEQ ID NO.)		C1	C2	C3	C4	C7	C8	C10	C11	C17	C20
53	A	0	0	0	0.1	0.4	1.1	0	2.5	2.8	0
	G	0	0	0	0	0.1	0.3	0	0.5	2	0
	T	0	0	0	0.5	1.4	2.5	0	7	11.3	0.7
54	A	0	0	0	0	0	0.1	0	0.7	0.8	0
	G	0	0	0	0	0	0	0	0	0.6	0
	T	0	0	0	1.6	6.1	8.6	2.2	13.3	23	1.8
55	A	0	0	0	0	0.2	0	0	1	0.9	0
	G	0	0	0	0	0	0.3	0.1	0.1	0.5	0
	T	0	0	0.2	2.2	9.1	11.2	3.2	13	21.3	2
56	A	0	0	0	0	0	0	0	0.9	0.8	0
	G	0	0	0	0	0	0.1	0	0.2	0.5	0
	T	0	0	0.1	2.7	12.3	15	4.3	18.3	27.1	1.9

The results of Table 11 show that the rate of C>T formation at positions C7, C8, C11 and C17 increased with the addition of a USP.

10 **Table 12: C>N Editing Rate using deaminase APG09980 and guide SGN000929**

Fusion Protein (SEQ ID NO.)		C6	C9	C13	C23
49	A	0.2	1.1	0.2	5.2
	G	0.2	1.5	1.4	20.1
	T	1.2	5.2	1	7.4
50	A	0	0.1	0	1.3
	G	0	0.1	0	3.3
	T	2.2	9.4	3.2	30.8
51	A	0	0.2	0	0.8
	G	0	0.1	0	1.6
	T	1.5	9.2	3.5	34

Fusion Protein (SEQ ID NO.)		C6	C9	C13	C23
52	A	0.5	0.6	0	1.8
	G	0.1	0.5	0	6.5
	T	1.6	8.6	2.7	22.3

The results of Table 12 show that the rate of C>T formation at position C23 increased with the addition of a USP.

5

Table 13: C>N Editing Rate using deaminase APG07386 and guide SGN000929

Fusion Protein (SEQ ID NO.)		C6	C9	C13	C23
53	A	1.5	0	0	1.5
	G	0.6	0	0.1	5.2
	T	5	0.3	0.1	2.1
54	A	0	0	0	0.1
	G	0	0	0	0.1
	T	4.2	0.3	0	4.6
55	A	0.2	0	0.1	0
	G	0	0	0	1.4
	T	10.4	0.9	0.3	8.2
56	A	0.1	0	0	0.2
	G	0.2	0	0	0.5
	T	10.8	0.6	0.4	8.5

The results of Table 13 show that the rate of C>T formation at positions C6 and C23 increased with the addition of a USP.

10

Table 14: C>N Editing Rate using deaminase APG09980 and guide SGN001101

Fusion Protein (SEQ ID NO.)		C2	C6	C7	C18
49	A	0.1	1.8	0.4	3.3
	G	0	2.8	0.1	18.5
	T	0.1	6.5	1.8	5.3
50	A	0.2	0.3	0.2	0.6
	G	0.1	0.1	0.2	2.1
	T	0.1	9.1	0.7	25.4
51	A	0.2	0.1	0.1	0.3
	G	0.1	0.1	0	2.4
	T	0.1	7.9	0.7	21.8
52	A	0.2	0.9	0.2	1.2
	G	0	0.8	0	4.8

Fusion Protein (SEQ ID NO.)		C2	C6	C7	C18
	T	0.2	7.7	1.3	19.4

The results of Table 14 show that the rate of C>T formation at position C18 increased with the addition of a USP.

5 **Table 15: C>N Editing Rate using deaminase APG07386 and guide SGN001101**

Fusion Protein (SEQ ID NO.)		C2	C6	C7	C18
53	A	0.1	0	4	2.2
	G	0	0	2.1	11
	T	0	0	10.4	3.9
54	A	0.1	0	0.1	0
	G	0	0	0.1	0.6
	T	0	1.7	8.5	6.7
55	A	0.2	0.1	0.6	0.3
	G	0	0	0.5	1.4
	T	0	1.7	10.7	9.2
56	A	0.1	0	0.3	0.2
	G	0	0	0.4	1.4
	T	0	1.9	12.1	9.8

The results of Table 15 show that the rate of C>T formation at position C18 increased with the addition of a USP. The rate of C>G conversion was decreased at position C18 with the addition of a USP.

Tables 16 and 17 show the rate of indel formation for each fusion protein/guide combination tested.

10 The fusion protein is indicated by SEQ ID NO. The data indicates that the fusion proteins comprising a USP described herein decreased the rate of indel formation at all target genomic locations tested.

Table 16: Insertion and Deletion Rate with APG09980 and USPs

sgRNA ID	SEQ ID NO: 49	SEQ ID NO: 50	SEQ ID NO: 51	SEQ ID NO: 52
SGN000143	2.78	0	0	0.62
SGN000169	18.02	1.62	3.89	9.1
SGN000173	27.46	7.86	7.17	7.48
SGN000929	3.49	0.2	0.65	2.36
SGN000930	6.03	0.81	0.65	1.65
SGN001101	2.5	1.66	0.92	1.86

Table 17: Insertion and Deletion Rate with APG07386-CTD and USPs

sgRNA ID	SEQ ID NO: 53	SEQ ID NO: 54	SEQ ID NO: 55	SEQ ID NO: 56
SGN000143	0.13	0.05	0	0.06
SGN000169	5.6	1.38	3.11	0.32
SGN000173	13.8	0.99	3.49	2.57
SGN000929	0.5	0	0.3	0.16
SGN000930	1.7	0	0.23	0.31
SGN001101	1.78	0	0.48	0.39

Example 3: Testing different delivery formats

To determine if the base editors are capable of delivery in different formats, mRNA delivery was tested with primary T-cells. Purified CD3⁺ T-cells or PBMCs were thawed, activated using CD3/CD28 beads (ThermoFisher) for 3 days, then nucleofected using the Lonza 4D-Nucleofector X unit and Nucleocuvette strips. The P3 Primary Cell kit was used for both mRNA and RNP delivery. Cells were transfected using the EO-115 and EH-115 programs for mRNA and RNP delivery respectively. Cells were cultured in CTS OpTimizer T cell expansion medium (ThermoFisher) containing IL-2, IL-7, and IL-15 (Miltenyi Biotec) for 4 days post nucleofection before being harvested using a Nucleospin Tissue genomic DNA isolation kit (Machery Nagel).

Amplicons surrounding the editing sites were generated by PCR and subjected to NGS sequencing using the Illumina Nexterra platform using 2x250bp paired end sequencing. The estimated base editing rate was determined by calculating the overall substitution rate for each sample. The average and number of samples for each guide tested are shown in Tables 18 and 19 below.

APG09980-nAPG07433.1-APG03399 and APG05840-nAPG07433.1-APG03399 when delivered by mRNA show high rates of base editing as several targets. There are very low rates of indel formation despite the high substitution rate, due to the incorporation of USP2 in the base editing construct.

Table 18: Average base editing rate for APG09980-nAPG07433.1-APG03399

Fusion Construct	Gene Name	SGN	Average % Substitutions	Average % Indels	N
APG09980-nAPG07433.1-APG03399	Gene 1	SGN000754	23.32917436	1.128931	6
APG09980-nAPG07433.1-APG03399	Gene 1	SGN000755	59.37254849	7.1037823	4
APG09980-nAPG07433.1-APG03399	Gene 2	SGN001061	13.60100568	0.4214674	3
APG09980-nAPG07433.1-APG03399	Gene 2	SGN001062	26.9304354	4.3225871	4
APG09980-nAPG07433.1-APG03399	Gene 2	SGN001063	75.27761104	0.8163273	4
APG09980-nAPG07433.1-APG03399	Gene 2	SGN001064	72.94658862	1.0468487	3

Table 19: Average base editing rate for APG05840-nAPG07433.1-APG03399

Fusion Construct	Gene Name	SGN	Average % Substitutions	Average % Indels	N
APG05840-nAPG07433.1-APG03399	Gene 1	SGN000754	57.7775198	5.14624384	4
APG05840-nAPG07433.1-APG03399	Gene 1	SGN000755	68.352455	4.98538891	3
APG05840-nAPG07433.1-APG03399	Gene 2	SGN001061	14.6830209	0	2
APG05840-nAPG07433.1-APG03399	Gene 2	SGN001062	39.7312597	2.9887885	4
APG05840-nAPG07433.1-APG03399	Gene 2	SGN001063	70.4564399	0.25727852	4
APG05840-nAPG07433.1-APG03399	Gene 2	SGN001064	53.2112842	1.98008536	3

THAT WHICH IS CLAIMED:

1. An isolated polypeptide comprising an amino acid sequence having:
 - a) at least 80% sequence identity to any one of SEQ ID NOs: 1, 2, 4, 5, and 7-15;
 - b) at least 81% sequence identity to SEQ ID NO: 3 or 16; or
 - c) at least 82% sequence identity to SEQ ID NO: 6;

wherein said polypeptide has uracil stabilizing activity and wherein said polypeptide further comprises a heterologous amino acid sequence.

2. The isolated polypeptide of claim 1, wherein the polypeptide has the sequence of any one of SEQ ID NOs: 33-39.

3. A pharmaceutical composition comprising a non-naturally occurring pharmaceutically acceptable carrier and a polypeptide comprising an amino acid sequence having:

- a) at least 80% sequence identity to any one of SEQ ID NOs: 1, 2, 4, 5, and 7-15;
- b) at least 81% sequence identity to SEQ ID NO: 3 or 16; or
- c) at least 82% sequence identity to SEQ ID NO: 6;

wherein said polypeptide has uracil stabilizing activity.

4. A pharmaceutical composition comprising a non-naturally occurring pharmaceutically acceptable carrier and a nucleic acid molecule comprising a polynucleotide encoding a polypeptide comprising an amino acid sequence having:

- a) at least 80% sequence identity to any one of SEQ ID NOs: 1, 2, 4, 5, and 7-15;
- b) at least 81% sequence identity to SEQ ID NO: 3 or 16; or
- c) at least 82% sequence identity to SEQ ID NO: 6;

wherein said polypeptide has uracil stabilizing activity.

5. The pharmaceutical composition of claim 3 or 4, wherein the polypeptide has the sequence of any one of SEQ ID NOs: 33-39.

6. The pharmaceutical composition of any one of claims 3-5, further comprising a fluoropyrimidine.

7. A nucleic acid molecule comprising a polynucleotide encoding a polypeptide comprising an amino acid sequence having:

- a) at least 80% sequence identity to any one of SEQ ID NOs: 1, 2, 4, 5, and 7-15;
- b) at least 81% sequence identity to SEQ ID NO: 3 or 16; or
- c) at least 82% sequence identity to SEQ ID NO: 6;

wherein said polypeptide has uracil stabilizing activity; and

wherein said nucleic acid molecule further comprises a heterologous promoter operably linked to said polynucleotide.

8. The nucleic acid molecule of claim 7, wherein the polypeptide has the sequence of any one of SEQ ID NOs: 33-39.

9. A composition comprising a fluoropyrimidine and a polypeptide comprising an amino acid sequence having:

- a) at least 80% sequence identity to any one of SEQ ID NOs: 1, 2, 4, 5, and 7-15;
- b) at least 81% sequence identity to SEQ ID NO: 3 or 16; or
- c) at least 82% sequence identity to SEQ ID NO: 6;

wherein said polypeptide has uracil stabilizing activity.

10. A composition comprising a fluoropyrimidine and a nucleic acid molecule encoding a polypeptide comprising an amino acid sequence having:

- a) at least 80% sequence identity to any one of SEQ ID NOs: 1, 2, 4, 5, and 7-15;
- b) at least 81% sequence identity to SEQ ID NO: 3 or 16; or
- c) at least 82% sequence identity to SEQ ID NO: 6;

wherein said polypeptide has uracil stabilizing activity.

11. The composition of claim 9 or 10, wherein the polypeptide has the sequence of any one of SEQ ID NOs: 33-39.

12. A fusion protein comprising: (i) a DNA-binding polypeptide; (ii) a deaminase; and (iii) at least one uracil stabilizing polypeptide (USP) having at least 80% sequence identity to any one of SEQ ID NOs: 1-16.

13. The fusion protein of claim 12, wherein the USP has the sequence of any one of SEQ ID NOs: 33-39.

14. The fusion protein of claim 12 or 13, wherein the deaminase is a cytidine deaminase.

15. The fusion protein of claim 14, wherein the cytidine deaminase is an activation-induced cytidine deaminase (AID) or a member of the apolipoprotein B mRNA-editing complex (APOBEC) family of deaminases.

16. The fusion protein of claim 15, wherein the cytidine deaminase comprises an amino acid sequence having at least 80% sequence identity to any one of SEQ ID NOs: 47, 48 and 76-94.

17. The fusion protein of any one of claims 12-16, wherein the DNA-binding polypeptide is a meganuclease, zinc finger fusion protein, or a TALEN.

18. The fusion protein of any one of claims 12-16, wherein the DNA-binding polypeptide is an RNA-guided, DNA-binding polypeptide.

19. The fusion protein of claim 18, wherein the RNA-guided, DNA-binding polypeptide is an RNA-guided nuclease polypeptide (RGN).

20. The fusion protein of claim 19, wherein the RGN is a Type II CRISPR-Cas polypeptide.

21. The fusion protein of claim 19, wherein the RGN is a Type V CRISPR-Cas polypeptide.

22. The fusion protein of claim 19, wherein the RGN comprises an amino acid sequence having at least 80% sequence identity to any one of SEQ ID NOs: 40 and 95-142.

23. The fusion protein of any one of claims 19-22, wherein the RGN is an RGN nickase.

24. The fusion protein of any one of claims 19-23, wherein the fusion protein comprises an RGN, a cytidine deaminase, and a USP.

25. The fusion protein of claim 24, wherein said RGN has at least 80% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, said cytidine deaminase has at least 80% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and said USP has at least 80% sequence identity to any one of SEQ ID NOs: 1-16.

26. The fusion protein of any of claims 12-25, wherein the fusion protein further comprises at least one nuclear localization signal (NLS).

27. A nucleic acid molecule comprising a polynucleotide encoding a fusion protein comprising: (i) a DNA-binding polypeptide; (ii) a deaminase; and (iii) at least one uracil stabilizing polypeptide (USP), wherein the USP is encoded by a nucleotide sequence that:

a) has at least 80% sequence identity to any one of SEQ ID NOs: 17-32,

b) is set forth in any one of SEQ ID NOs: 17-32,

15 c) encodes an amino acid sequence at least 80% identical to SEQ ID NOs: 1-16 and further possesses the sequence of any one of SEQ ID NOs: 33-39,

d) encodes an amino acid sequence at least 80% identical to an amino acid sequence set forth in any one of SEQ ID NOs: 1-16, or

e) encodes an amino acid sequence set forth in any one of SEQ ID NOs: 1-16.

20 28. The nucleic acid molecule of claim 27, wherein the deaminase is a cytidine deaminase.

29. The nucleic acid molecule of claim 28, wherein the cytidine deaminase is an activation-induced cytidine deaminase (AID) or a member of the apolipoprotein B mRNA-editing complex (APOBEC) family of deaminases.

30 30. The nucleic acid molecule of claim 29, wherein the cytidine deaminase comprises an amino acid sequence having at least 80% sequence identity to any one of SEQ ID NOs: 47, 48 and 76-94.

31. The nucleic acid molecule of any one of claims 27-30, wherein the DNA-binding polypeptide is a meganuclease, zinc finger fusion protein, or a TALEN.

32. The nucleic acid molecule of any one of claims 27-30, wherein the DNA-binding polypeptide is an RNA-guided, DNA-binding polypeptide.

30 33. The nucleic acid molecule of claim 32, wherein the RNA-guided, DNA-binding polypeptide is an RNA-guided nuclease polypeptide (RGN).

34. The nucleic acid molecule of claim 33, wherein the RGN is a Type II CRISPR-Cas polypeptide.

35 35. The nucleic acid molecule of claim 33, wherein the RGN is a Type V CRISPR-Cas polypeptide.

36. The nucleic acid molecule of claim 33, wherein the RGN comprises an amino acid sequence having at least 80% sequence identity to any one of SEQ ID NOs: 40 and 95-142.

37. The nucleic acid molecule of any one of claims 33-36, wherein the RGN is an RGN nickase.

38. The nucleic acid molecule of any one of claims 33-36, wherein the fusion protein comprises an RGN, a cytidine deaminase, and a USP.

39. The nucleic acid molecule of claim 38, wherein said RGN has at least 80% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, said cytidine deaminase has at least 80% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and said USP has at least 80% sequence identity to any one of SEQ ID NOs: 1-16.

40. The nucleic acid molecule of any of claims 27-39, wherein the polynucleotide encoding the fusion protein is operably linked at its 5' end to a heterologous promoter.

41. The nucleic acid molecule of any of claims 27-39, wherein the polynucleotide encoding the fusion protein is operably linked at its 3' end to a heterologous terminator.

42. The nucleic acid molecule of any of claims 29-41, wherein the fusion protein comprises one or more nuclear localization signals.

43. The nucleic acid molecule of any of claims 27-42, wherein the fusion protein is codon optimized for expression in a eukaryotic cell.

44. The nucleic acid molecule of any of claims 27-43, wherein the fusion protein is codon optimized for expression in a prokaryotic cell.

45. A vector comprising the nucleic acid molecule of any one of claims 27-44.

46. A vector comprising the nucleic acid molecule of any one of claims 33-39, further comprising at least one nucleotide sequence encoding a guide RNA (gRNA) capable of hybridizing to a target sequence.

47. The vector of claim 46, wherein the gRNA is a single guide RNA.

48. The vector of claim 46, wherein the gRNA is a dual guide RNA.

49. A cell comprising the fusion protein of any of claims 12-26.

50. A cell comprising the fusion protein of any one of claims 18-25, wherein the cell further comprises a guide RNA.

51. A cell comprising the nucleic acid molecule of any of claims 27-44.

52. A cell comprising the vector of any of claims 45 through 48.

53. The cell of any one of claims 49-52, wherein the cell is a prokaryotic cell.

54. The cell of any one of claims 49-52, wherein the cell is a eukaryotic cell.

55. The cell of claim 54, wherein the cell is an insect, avian, or mammalian cell.

56. The cell of claim 54, wherein the cell is a plant or fungal cell.

57. A pharmaceutical composition comprising a pharmaceutically acceptable carrier and the nucleic acid molecule of any one of claims 7, 8, 27-44, the composition of any one of claims 9-11, the fusion protein of any one of claims 12-26, the vector of any one of claims 45-48, or the cell of any one of claims 49-55.

58. A method for making a fusion protein comprising culturing the cell of any one of claims 49-56 under conditions in which the fusion protein is expressed.

59. A method for making a fusion protein comprising introducing into a cell the nucleic acid molecule of any of claims 27-44 or a vector of any one of claims 45-48 and culturing the cell under conditions in which the fusion protein is expressed.

60. The method of claim 58 or 59, further comprising purifying said fusion protein.

61. A method for making an RGN fusion ribonucleoprotein complex, comprising introducing into a cell the nucleic acid molecule of any one of claims 33-39 and a nucleic acid molecule comprising an expression cassette encoding for a guide RNA, or the vector of any of claims 46-48, and culturing the cell under conditions in which the fusion protein and the gRNA are expressed and form an RGN fusion ribonucleoprotein complex.

62. The method of claim 61, further comprising purifying said RGN fusion ribonucleoprotein complex.

63. A system for modifying a target DNA molecule comprising a target DNA sequence, said system comprising:

a) a fusion protein comprising an RNA-guided nuclease polypeptide (RGN), a cytidine deaminase, and at least one uracil stabilizing polypeptide (USP), wherein the USP is at least 80% identical to any one of SEQ ID NOs: 1-16, or a nucleotide sequence encoding said fusion protein; and

b) one or more guide RNAs capable of hybridizing to said target DNA sequence or one or more nucleotide sequences encoding the one or more guide RNAs (gRNAs);

wherein said nucleotide sequences encoding the one or more guide RNAs and encoding the fusion protein are each operably linked to a promoter heterologous to said nucleotide sequence;

and

wherein the one or more guide RNAs are capable of forming a complex with the fusion protein in order to direct said fusion protein to bind to said target DNA sequence and modify the target DNA molecule.

64. The system of claim 63, wherein the target DNA sequence is located adjacent to a protospacer adjacent motif (PAM) that is recognized by the RGN.

65. The system of claim 63 or 64, wherein the target DNA molecule is within a cell.

66. The system of claim 65, wherein the cell is a eukaryotic cell.

67. The system of claim 66, wherein the eukaryotic cell is a plant cell.

68. The system of claim 66, wherein the eukaryotic cell is a mammalian cell.

69. The system of claim 66, wherein the eukaryotic cell is an insect cell.

70. The system of claim 65, wherein the cell is a prokaryotic cell.

71. The system of any one of claims 63-70, wherein the RGN of the fusion protein is a Type II CRISPR-Cas polypeptide.

72. The system of any one of claims 63-70, wherein the RGN of the fusion protein is a Type V CRISPR-Cas polypeptide.

73. The system of any one of claims 63-70, wherein the RGN of the fusion protein is at least 80% identical to any one of SEQ ID NOs: 40 and 95-142.

74. The system of any one of claims 63-73, wherein the RGN of the fusion protein is an RGN nickase.

75. The system of any of claims 63-74, wherein the cytidine deaminase is at least 80% identical to any one of SEQ ID NOs: 47, 48 and 76-94.

76. The system of any of claims 63-75, wherein the USP comprises the sequence of any one of SEQ ID NOs: 33-39.

77. The system of any one of claims 63-76, wherein the RGN has at least 80% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, the cytidine deaminase has at least 80% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and the USP has at least 80% sequence identity to any one of SEQ ID NOs: 1-16.

78. The system of any of claims 63-77, wherein the fusion protein comprises one or more nuclear localization signals.

79. The system of any of claims 63-78, wherein the fusion protein is codon optimized for expression in a eukaryotic cell.

80. The system of any of claims 63-79, wherein nucleotide sequences encoding the one or more guide RNAs and the nucleotide sequence encoding a fusion protein are located on one vector.

81. A method for modifying a target DNA molecule comprising a target DNA sequence, said method comprising delivering a system according to any one of claims 63-80 to said target DNA molecule or a cell comprising the target DNA molecule.

82. The method of claim 81, wherein said modified target DNA molecule comprises a C>T mutation of at least one nucleotide within the target DNA molecule.

83. The method of claim 81, wherein said modified target DNA molecule comprises a C>T mutation of at least one nucleotide within the target DNA sequence.

84. A method for modifying a target DNA molecule comprising a target sequence comprising:

a) assembling an RGN-deaminase-USP ribonucleotide complex *in vitro* by combining:

i) one or more guide RNAs capable of hybridizing to the target DNA sequence; and

ii) a fusion protein comprising an RNA-guided nuclease polypeptide (RGN), a cytidine deaminase, and at least one uracil stabilizing polypeptide (USP), wherein the USP is at least 80% identical to any one of SEQ ID NOs: 1-16;

under conditions suitable for formation of the RGN-deaminase-USP ribonucleotide complex; and

b) contacting said target DNA molecule or a cell comprising said target DNA molecule with the *in vitro*-assembled RGN-deaminase-USP ribonucleotide complex;

wherein the one or more guide RNAs hybridize to the target DNA sequence, thereby directing said fusion protein to bind to said target DNA sequence and modification of the target DNA molecule occurs.

85. The method of claim 84, wherein said modified target DNA molecule comprises a C>T mutation of at least one nucleotide within the target DNA molecule.

86. The method of claim 84, wherein said modified target DNA molecule comprises a C>T mutation of at least one nucleotide within the target DNA sequence.

5 87. The method of any one of claims 84-86, wherein the RGN of the fusion protein is a Type II CRISPR-Cas polypeptide.

88. The method of any of claims 84-86, wherein the RGN of the fusion protein is a Type V CRISPR-Cas polypeptide.

10 89. The method of any of claims 84-86, wherein the RGN of the fusion protein is at least 80% identical to any one of SEQ ID NOs: 40 and 95-142.

90. The method of any of claims 84-89, wherein the RGN of the fusion protein is an RGN nickase.

91. The method of any of claims 84-90, wherein the cytidine deaminase is at least 80% identical to any one of SEQ ID NOs: 47, 48 and 76-94.

15 92. The method of any of claims 84-91, wherein the USP comprises the sequence of any one of SEQ ID NOs: 33-39.

20 93. The method of any of claims 84-92, wherein the RGN has at least 80% sequence identity to any one of SEQ ID NOs: 40, 41, and 95-142, the cytidine deaminase has at least 80% sequence identity to any one of SEQ ID NOs: 47, 48, and 76-94, and the USP has at least 80% sequence identity to any one of SEQ ID NOs: 1-16.

94. The method of any of claims 84-93, wherein the fusion protein comprises one or more nuclear localization signals.

95. The method of any of claims 84-94, wherein the fusion protein is codon optimized for expression in a eukaryotic cell.

25 96. The method of any of claims 84-95, wherein said target DNA sequence is located adjacent to a protospacer adjacent motif (PAM).

97. The method of any of claims 84-96, wherein the target DNA molecule is within a cell.

98. The method of claim 97, wherein the cell is a eukaryotic cell.

99. The method of claim 98, wherein the eukaryotic cell is a plant cell.

30 100. The method of claim 98, wherein the eukaryotic cell is a mammalian cell.

101. The method of claim 98, wherein the eukaryotic cell is an insect cell.

102. The method of claim 97, wherein the cell is a prokaryotic cell.

103. The method of any one of claims 97-102, further comprising selecting a cell comprising said modified DNA molecule.

35 104. A cell comprising a modified target DNA sequence according to the method of claim 103.

105. The cell of claim 104, wherein the cell is a eukaryotic cell.

106. The cell of claim 105, wherein the eukaryotic cell is a plant cell.

107. A plant comprising the cell of claim 106.

108. A seed comprising the cell of claim 106.

109. The cell of claim 105, wherein the eukaryotic cell is a mammalian cell.

110. The cell of claim 105, wherein the eukaryotic cell is an insect cell.

5 111. The cell of claim 104, wherein the cell is a prokaryotic cell.

112. A method for producing a genetically modified cell with a correction in a causal mutation for a genetically inherited disease, the method comprising introducing into the cell:

10 a) a fusion protein comprising an RNA-guided nuclease polypeptide (RGN), a cytidine deaminase, and at least one uracil stabilizing polypeptide (USP), wherein the USP is at least 80% identical to any one of SEQ ID NOs: 1-16, or a polynucleotide encoding said fusion protein, wherein said polynucleotide encoding the fusion protein is operably linked to a promoter to enable expression of the fusion protein in the cell; and

15 b) one or more guide RNAs (gRNA) capable of hybridizing to a target DNA sequence, or a polynucleotide encoding said gRNA, wherein said polynucleotide encoding the gRNA is operably linked to a promoter to enable expression of the gRNA in the cell;

whereby the fusion protein and gRNA target to the genomic location of the causal mutation and modify the genomic sequence to remove the causal mutation.

113. The method of claim 112, wherein said RGN of the fusion protein is a nickase.

20 114. The method of claim 112 or 113, wherein the USP comprises the sequence of any one of SEQ ID NOs: 33-39.

115. The method of any of claims 112-114, wherein the genome modification comprises introducing a C>T mutation of at least one nucleotide within the target DNA sequence.

116. The method of any of claims 112-115, wherein the cell is an animal cell.

117. The method of claim 116, wherein the animal cell is a mammalian cell.

25 118. The method of claim 117, wherein the cell is derived from a dog, cat, mouse, rat, rabbit, horse, sheep, goat, cow, pig, or human.

119. The method of any of claims 112-118, wherein the correction of the causal mutation comprises introducing a stop codon.

120. A composition comprising:

30 a) a fusion protein comprising: (i) a DNA-binding polypeptide; and (ii) a deaminase; or a nucleic acid molecule encoding the fusion protein; and

b) a uracil stabilizing polypeptide (USP) having at least 80% sequence identity to any one of SEQ ID NOs: 1-16; or a nucleic acid molecule encoding the USP.

35 121. The composition of claim 120, wherein the fusion protein further comprises a uracil stabilizing polypeptide (USP) having at least 80% sequence identity to any one of SEQ ID NOs: 1-16.

122. The composition of claim 120 or 121, wherein the DNA-binding polypeptide is a meganuclease, zinc finger fusion protein, or a TALEN.

123. The composition of claim 120 or 121, wherein the DNA-binding polypeptide is an RNA-guided, DNA-binding polypeptide.

124. The composition of claim 123, wherein the RNA-guided, DNA-binding polypeptide is an RNA-guided nuclease polypeptide (RGN).

5 125. The composition of claim 124, wherein the RGN is an RGN nickase.

126. A vector comprising a nucleic acid molecule encoding a fusion protein and a nucleic acid molecule encoding a uracil stabilizing polypeptide (USP), wherein said fusion protein comprises a DNA-binding polypeptide and a deaminase, and wherein said USP has at least 80% sequence identity to any one of SEQ ID NOs: 1-16.

10 127. The vector of claim 126, wherein the fusion protein further comprises a uracil stabilizing polypeptide (USP) having at least 80% sequence identity to any one of SEQ ID NOs: 1-16.

128. The vector of claim 126 or 127, wherein the DNA-binding polypeptide is a meganuclease, zinc finger fusion protein, or a TALEN.

129. The vector of claim 126 or 127, wherein the DNA-binding polypeptide is an RNA-guided,
15 DNA-binding polypeptide.

130. The vector of claim 129, wherein the RNA-guided, DNA-binding polypeptide is an RNA-guided nuclease polypeptide (RGN).

131. The vector of claim 130, wherein the RGN is an RGN nickase.

132. A cell comprising the vector of any one of claims 126-131.

20 133. A cell comprising:

a) a fusion protein comprising: (i) a DNA-binding polypeptide; and (ii) a deaminase; or a nucleic acid molecule encoding the fusion protein; and

b) a uracil stabilizing polypeptide (USP) having at least 80% sequence identity to any one of SEQ ID NOs: 1-16; or a nucleic acid molecule encoding the USP.

25 134. The cell of claim 133, wherein the fusion protein further comprises a uracil stabilizing polypeptide (USP) having at least 80% sequence identity to any one of SEQ ID NOs: 1-16.

135. The cell of claim 134, wherein the DNA-binding polypeptide is a meganuclease, zinc finger fusion protein, or a TALEN.

136. The cell of claim 134, wherein the DNA-binding polypeptide is an RNA-guided, DNA-
30 binding polypeptide.

137. The cell of claim 136, wherein the RNA-guided, DNA-binding polypeptide is an RNA-guided nuclease polypeptide (RGN).

138. The cell of claim 137, wherein the RGN is an RGN nickase.

SEQUENCE LISTING

<110> LifeEDIT, LLC

<120> URACIL STABILIZING PROTEINS AND ACTIVE FRAGMENTS AND VARIANTS THEREOF AND METHODS OF USE

<130> L103438 1220WO (0106.5)

<140>

<141>

<150> 63/052,175

<151> 2020-07-15

<160> 142

<170> PatentIn version 3.5

<210> 1

<211> 116

<212> PRT

<213> Staphylococcus spp

<223> APG03399

<400> 1

Met Lys Thr Thr Thr Gln Glu Leu Lys Gln Tyr Ile Thr Arg Leu Phe
1 5 10 15

Gln Leu Ser Asn Asn Glu Thr Trp Glu Cys Glu Ala Leu Glu Glu Ala
20 25 30

Ala Glu Asn Ile Leu Pro Glu Arg Phe Ile Asn Asn Ser Leu Leu Ala
35 40 45

His Leu Thr Leu Asn Thr Tyr Thr Tyr Tyr Asn Asp Glu Leu His Glu
50 55 60

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

Ile Gly Tyr Leu Asp His Phe Asp Met Asp Phe Leu Tyr Leu Thr Asp
85 90 95

Thr Lys Asn Thr Ile Ile Asp Glu Arg His Leu Leu Lys Gln Gly Glu
100 105 110

Asn Asn His Glu
115

<210> 2
<211> 116
<212> PRT
<213> Staphylococcus spp

<223> APG06702

<400> 2

Met Thr Ile Glu Asn Glu Leu Lys Glu Tyr Ile Thr Asp Leu Phe Glu
1 5 10 15

Leu Pro Lys Asp Glu Gln Trp Leu Cys Glu Ala Val Asp Glu Val Ala
20 25 30

Asp Asn Ile Leu Pro Asp Glu Tyr Val Thr Leu Gly Pro Leu Arg Asn
35 40 45

Lys Thr Leu Glu Thr Phe Thr Tyr Tyr Ser Asn Thr Leu His Glu Arg
50 55 60

Asn Ile Tyr Pro Phe Ile Leu Tyr Tyr Gln Asn Arg Leu Ile Ala Ile
65 70 75 80

Gly Tyr Ile Asp Glu Asn Tyr Glu Met Asp Phe Met Tyr Leu His Asn
85 90 95

Thr Ala Ile Thr Leu Ile Asp Gln Arg Tyr Leu Ile Asp Lys Gly Asp
100 105 110

Asp Asn His Asp
115

<210> 3
<211> 116
<212> PRT
<213> Staphylococcus spp

<223> APG05198

<400> 3

Met Thr Leu Glu Lys Gln Leu Lys Glu Tyr Ile Thr Asn Leu Phe Asn
1 5 10 15

Leu Pro Lys Asp Glu Lys Trp Glu Cys Glu Ser Ile Glu Glu Val Ala
20 25 30

Asp Asn Ile Leu Pro Asp Gln Tyr Ile Arg Leu Gly Pro Leu Thr Asn
35 40 45

Lys Ile Leu His Thr Tyr Thr Tyr Tyr Ser Asp Thr Leu His Glu Arg
50 55 60

His Ile Tyr Pro Phe Ile Leu Tyr Tyr Gln Lys Gln Leu Ile Ala Ile
65 70 75 80

Gly Tyr Ile Asp Glu Thr Asn Asp Met Asp Phe Leu Tyr Leu His Asn
85 90 95

Thr Val Met Pro Leu Leu Asp Gln Arg His Leu Leu Glu Lys Glu Asn
100 105 110

Tyr Asn Asn Glu
115

<210> 4

<211> 116

<212> PRT

<213> Staphylococcus spp

<223> APG04638

<400> 4

Met Lys Thr Thr Thr Gln Glu Leu Lys Gln Tyr Ile Thr Arg Leu Phe
1 5 10 15

Gln Leu Ser Asn Asn Glu Thr Trp Glu Cys Glu Ala Leu Glu Glu Ala
20 25 30

Ala Glu Asn Ile Leu Pro Glu Arg Phe Ile Asn Asp Ser Pro Leu Ala
35 40 45

His Leu Thr Leu Glu Thr Tyr Thr Tyr Tyr Asn Asn Glu Leu His Asp
50 55 60

Leu Ser Ile Tyr Pro Phe Leu Met Tyr Ala Asn Asn Gln Leu Ile Ser
65 70 75 80

Val Gly Tyr Leu Asn His Phe Asp Met Asp Phe Leu Tyr Leu Thr Asp
85 90 95

Thr Lys Asn Thr Ile Ile Asp Glu Arg His Leu Leu Arg Glu Gly Gly
100 105 110

Thr His His Glu
115

<210> 5
<211> 116
<212> PRT
<213> Staphylococcus spp

<223> APG06351

<400> 5

Met Lys Thr Ile Thr Gln Glu Leu Lys Gln Tyr Ile Thr His Leu Phe
1 5 10 15

Gln Leu Ser Asn Asn Glu Ala Trp Glu Cys Glu Ala Leu Glu Glu Ala
20 25 30

Ala Glu Asn Ile Leu Pro Glu Arg Phe Ile Asn Asp Ser Pro Leu Val
35 40 45

His Leu Thr Leu Glu Thr Tyr Thr Tyr Tyr Asn Asp Glu Leu His Glu
50 55 60

Leu Ser Ile Tyr Pro Phe Leu Met Tyr Ala Asn Asn Gln Leu Ile Ser
65 70 75 80

Val Gly Tyr Leu Asp His Phe Asp Met Asp Phe Leu Tyr Leu Thr Asp
85 90 95

Thr Gln Asn Ile Ile Ile Asp Glu Arg His Leu Leu Lys Gln Gly Gly
100 105 110

Gln Asp His Glu
115

<210> 6
<211> 116
<212> PRT
<213> Staphylococcus spp

<223> APG05963

<400> 6

Met Thr Leu Glu Gln Gln Leu Lys His Tyr Ile Thr Asn Leu Phe Lys
1 5 10 15

Leu Pro Lys Asp Glu Lys Trp Glu Cys Glu Ser Ile Glu Glu Val Ala
20 25 30

Asp Asn Ile Leu Pro Asp Gln Tyr Val Arg Leu Gly Pro Leu Thr Asn
35 40 45

Lys Ile Leu His Thr Tyr Thr Tyr Tyr Ser Asp Thr Leu His Glu Arg
50 55 60

His Ile Tyr Pro Phe Ile Leu Tyr Tyr Gln Lys Gln Leu Ile Ala Ile
65 70 75 80

Gly Tyr Ile Asp Glu Thr Asn Asp Met Asp Phe Leu Tyr Leu His Asn
85 90 95

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

His Lys Asn Glu
115

<210> 7
<211> 115
<212> PRT
<213> Staphylococcus spp

<223> APG09230

<400> 7

Met Lys Thr Gln Ile Asn Glu Leu Lys Ser Tyr Ile Thr Asn Leu Phe
1 5 10 15

Gln Leu Thr Ser Gln Glu Asn Trp Glu Cys Glu Ala Leu Asn Glu Ser
20 25 30

Ala Asp Asn Ile Leu Pro Glu Arg Phe Leu Asn Asp Ser Pro Leu Ser
35 40 45

Asn Arg Ile Leu Glu Thr Tyr Thr Tyr Tyr Asn Asn Glu Leu His Glu
50 55 60

Leu Ser Ile Tyr Pro Phe Leu Met Tyr Tyr Gln Asn Gln Leu Ile Ser
65 70 75 80

Ile Gly Tyr Leu Asp His Tyr Asp Met Asp Phe Leu Tyr Leu Asn Asp
85 90 95

Thr Gln Thr Thr Ile Ile Asp Glu Arg His Leu Leu Lys Gly Glu Asn
100 105 110

Lys Tyr Glu
115

<210> 8

<211> 116

<212> PRT

<213> Staphylococcus spp

<223> APG04100

<400> 8

Met Lys Thr Thr Thr Gln Glu Leu Lys Gln Tyr Ile Thr Arg Leu Phe
1 5 10 15

Gln Leu Ser Asn Glu Glu Ser Trp Glu Cys Glu Val Leu Asp Glu Val
20 25 30

Ala Glu Asn Ile Leu Pro Pro Arg Phe Val Asp Gly Ser Pro Leu Thr

35

40

45

His Leu Thr Leu Glu Thr Tyr Thr Tyr Tyr Asn Asn Glu Leu His Asp
 50 55 60

Leu Ser Ile Tyr Pro Phe Leu Met Tyr Ala Asn Asn Gln Leu Ile Ser
 65 70 75 80

Val Gly Tyr Leu Asp His Phe Asp Met Asp Phe Leu Tyr Leu Thr Asp
 85 90 95

Thr Lys Asn Thr Ile Ile Asp Glu Arg His Leu Leu Lys Gln Gly Glu
 100 105 110

Asn Asn His Glu
 115

<210> 9

<211> 116

<212> PRT

<213> Staphylococcus spp

<223> APG04001

<400> 9

Met Lys Thr Thr Thr Gln Glu Leu Lys Gln Tyr Ile Thr Arg Leu Phe
 1 5 10 15

Gln Leu Ser Asn Asn Glu Thr Trp Glu Cys Glu Ala Leu Glu Glu Ala
 20 25 30

Ala Glu Asn Ile Leu Pro Thr Arg Phe Val Asp His Thr Pro Leu Ala
 35 40 45

His Leu Thr Leu Asp Thr Tyr Thr Tyr Tyr Asn Asp Glu Leu His Glu
 50 55 60

Leu Ser Ile Tyr Pro Phe Leu Met Tyr Ala Asn Asn Gln Leu Ile Ser
 65 70 75 80

Ile Gly Tyr Leu Asp His Phe Asp Met Asp Phe Leu Tyr Leu Thr Asp
 85 90 95

Thr Lys Asn Ile Ile Ile Asp Glu Arg His Leu Leu Arg Glu Gly Gly
100 105 110

Asn Asp Arg Glu
115

<210> 10
<211> 116
<212> PRT
<213> Staphylococcus spp

<223> APG02463

<400> 10

Met Gln Thr Thr Thr Gln Glu Leu Lys Gln Tyr Ile Thr Arg Leu Phe
1 5 10 15

Gln Leu Ser Asn Asn Glu Thr Trp Glu Cys Glu Ala Leu Glu Glu Ala
20 25 30

Ala Glu Asn Ile Leu Pro Glu Arg Phe Ile Asn Gly Ser Pro Leu Ala
35 40 45

His Leu Thr Leu Glu Thr Tyr Thr Tyr Tyr Asn Asp Glu Leu His Glu
50 55 60

Leu Ser Ile Tyr Pro Phe Leu Met Tyr Ala Asn Asn Gln Leu Ile Ser
65 70 75 80

Ile Gly Tyr Leu Asp His Phe Asp Met Asp Phe Leu Tyr Leu Thr Asp
85 90 95

Thr Lys Asn Thr Ile Ile Asp Glu Arg His Leu Leu Lys Glu Gly Gly
100 105 110

Asn Asn His Glu
115

<210> 11
<211> 116
<212> PRT

<213> Staphylococcus spp

<223> APG09242

<400> 11

Met Lys Thr Thr Thr Gln Glu Leu Lys Gln Tyr Ile Thr Arg Leu Phe
1 5 10 15

Gln Leu Ser Asn Asn Glu Thr Trp Glu Cys Glu Ala Leu Glu Asp Thr
20 25 30

Ala Glu Asn Ile Leu Pro Glu Cys Phe Ile Asn Asp Thr Pro Leu Ala
35 40 45

His Leu Thr Leu Glu Thr Tyr Thr Tyr Tyr Asn Asp Glu Leu His Glu
50 55 60

Leu Ser Ile Tyr Pro Phe Leu Met Tyr Ala Asn Asn Gln Leu Ile Ser
65 70 75 80

Val Gly Tyr Leu Asp His Phe Asp Met Asp Phe Leu Tyr Leu Thr Asp
85 90 95

Thr Lys Asn Thr Ile Ile Asp Glu Arg His Leu Leu Lys Glu Gly Gly
100 105 110

Asn Asp His Glu
115

<210> 12

<211> 116

<212> PRT

<213> Staphylococcus spp

<223> APG04080

<400> 12

Met Lys Thr Thr Thr Gln Glu Leu Lys Gln Tyr Ile Thr Arg Leu Phe
1 5 10 15

Gln Leu Ser Asn Asn Glu Thr Trp Glu Cys Glu Ala Leu Glu Glu Ala
20 25 30

Ala Ala Asn Ile Leu Pro Glu Arg Phe Ile Asn Asn Ser Pro Leu Ala
35 40 45

His Leu Thr Leu Glu Thr Tyr Thr Tyr Tyr Asn Asp Glu Leu His Glu
50 55 60

Leu Ser Ile Tyr Pro Phe Leu Met Tyr Ala Asn Asn Gln Leu Ile Ser
65 70 75 80

Val Gly Tyr Leu Asp His Phe Asp Met Asp Phe Leu Tyr Leu Thr Asp
85 90 95

Thr Lys Asn Thr Ile Ile Asp Glu Arg His Leu Leu Lys Glu Gly Gly
100 105 110

Asn Asn His Glu
115

<210> 13
<211> 112
<212> PRT
<213> Staphylococcus spp

<223> APG05316

<400> 13

Met Thr Leu Ser Lys Gln Leu Lys Thr Tyr Ile Thr Glu Arg Phe Lys
1 5 10 15

Leu Asn Tyr Gln Glu Thr Trp Ala Cys Glu Thr Ile Asp Ala Val Ala
20 25 30

Glu Asp Val Leu Pro Glu Lys Tyr Ile Lys Asn Ser Pro Leu Glu His
35 40 45

Lys Ile Leu Asn Thr Phe Thr Tyr Tyr Asn Asp Glu Leu His Glu Ile
50 55 60

Ser Ile Tyr Pro Phe Leu Cys Tyr Leu Asp Lys Glu Leu Val Ala Ile
65 70 75 80

Gly Tyr Leu Asp Asn Phe Asp Leu Asp Phe Ile Phe Leu Asn Asp Thr
85 90 95

His Gln Val Ile Ile Asp Glu Arg Tyr Leu Leu Gln Lys Gly Gly Gln
100 105 110

<210> 14

<211> 113

<212> PRT

<213> Staphylococcus spp

<223> APG03327

<400> 14

Met Lys Thr Thr Thr Gln Glu Leu Lys Gln Tyr Ile Thr Arg Leu Phe
1 5 10 15

Gln Leu Ser Asn Asn Glu Thr Trp Glu Cys Glu Ala Leu Glu Asp Ala
20 25 30

Ala Glu Asn Ile Leu Pro Thr Arg Phe Val Asp His Thr Pro Leu Ala
35 40 45

His Leu Thr Leu Glu Thr Tyr Thr Tyr Tyr Asn Asp Glu Leu His Glu
50 55 60

Leu Ser Ile Tyr Pro Phe Leu Ile Tyr Ala Asn Asn Gln Leu Ile Ser
65 70 75 80

Ile Gly Tyr Leu Asp His Phe Asp Met Asp Phe Leu Tyr Leu Thr Asp
85 90 95

Thr Lys Asn Thr Ile Ile Asp Glu Arg His Leu Leu Gln Lys Gly Arg
100 105 110

Glu

<210> 15

<211> 116

<212> PRT

<213> Staphylococcus spp

<223> APG01791

<400> 15

Met Lys Thr Thr Thr Gln Glu Leu Lys Gln Tyr Ile Thr Arg Leu Phe
1 5 10 15

Gln Leu Ser Asn Asn Glu Thr Trp Glu Cys Glu Ala Leu Glu Glu Ala
20 25 30

Ala Glu Asn Ile Leu Pro Thr Arg Phe Val Gly His Thr Pro Leu Ala
35 40 45

His Leu Thr Leu Glu Thr Tyr Thr Tyr Tyr Asn Asp Glu Leu His Glu
50 55 60

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

Ile Gly Tyr Leu Asp His Phe Asp Met Asp Phe Leu Tyr Leu Thr Asp
85 90 95

Thr Lys Asn Thr Leu Ile Asp Glu Arg His Leu Leu Lys Gln Gly Gly
100 105 110

Asn Asp His Glu
115

<210> 16

<211> 116

<212> PRT

<213> Staphylococcus spp

<223> APG05756

<400> 16

Met Thr Leu Glu Gln Gln Leu Lys His Tyr Ile Thr Asn Leu Phe Ser
1 5 10 15

Leu Pro Gln Asp Glu Lys Trp Glu Cys Glu Ser Ile Glu Glu Val Ala
20 25 30

Asp Asn Ile Leu Pro Asp Gln Tyr Val Arg Leu Gly Pro Leu Thr Asn

35

40

45

Lys Ile Leu His Thr Tyr Thr Tyr Tyr Ser Asp Lys Leu His Glu Arg
 50 55 60

His Ile Tyr Pro Phe Ile Leu Tyr Tyr Gln Lys Gln Leu Ile Ala Ile
 65 70 75 80

Gly Tyr Ile Asp Glu Ile Asn Asp Met Asp Phe Leu Tyr Leu His Asn
 85 90 95

Thr Val Met Pro Leu Leu Asp Gln Arg His Leu Leu Glu Lys Glu Asn
 100 105 110

Tyr Asn Asn Glu
 115

<210> 17
 <211> 348
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> APG03399 mammalian codon optimized sequence

<400> 17
 atgaagacca ccaccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac 60
 aacgagacct gggagtgcga ggccctggag gaggccgccg agaacatcct gcccgagagg 120
 ttcatcaaca acagcctgct ggcccacctg accctgaaca cctacaccta ctacaacgac 180
 gagctgcacg agctgagcat ctacccttc ctgatgtacg ccaacgacca gctgatcagc 240
 atcggctacc tggaccactt cgacatggac ttcctgtacc tgaccgacac caagaacacc 300
 atcatcgacg agaggcacct gctgaagcag ggcgagaaca accacgag 348

<210> 18
 <211> 348
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> APG06702 mammalian codon optimized sequence

<400> 18

atgaccatcg agaacgagct gaaggagtac atcaccgacc tgttcgagct gcccgaaggac	60
gagcagtggc tgtgcgaggc cgtggacgag gtggccgaca acatcctgcc cgacgagtac	120
gtgaccctgg gccccctgag gaacaagacc ctggagacct tcacctacta cagcaacacc	180
ctgcacgaga ggaacatcta ccccttcatc ctgtactacc agaacaggct gatcgccatc	240
ggctacatcg acgagaacta cgagatggac ttcattgtacc tgcacaacac cgccatcacc	300
ctgatcgacc agaggtacct gatcgacaag ggcgacgaca accacgac	348

<210> 19
 <211> 348
 <212> DNA
 <213> Artificial Sequence

<220>

<223> APG05198 mammalian codon optimized sequence

<400> 19	
atgaccctgg agaagcagct gaaggagtac atcaccaacc tgttcaacct gcccgaaggac	60
gagaagtggg agtgcgagag catcgaggag gtggccgaca acatcctgcc cgaccagtac	120
atcaggctgg gccccctgac caacaagatc ctgcacacct acacctacta cagcgacacc	180
ctgcacgaga ggcacatcta ccccttcatc ctgtactacc agaagcagct gatcgccatc	240
ggctacatcg acgagaccaa cgacatggac ttcctgtacc tgcacaacac cgtgatgccc	300
ctgctggacc agaggcacct gctggagaag gagaactaca acaacgag	348

<210> 20
 <211> 348
 <212> DNA
 <213> Artificial Sequence

<220>

<223> APG04638 mammalian codon optimized sequence

<400> 20	
atgaagacca ccaccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac	60
aacgagacct gggagtgcga ggccctggag gaggccgccg agaacatcct gcccgagagg	120
ttcatcaacg acagccccct ggcccacctg accctggaga cctacaccta ctacaacaac	180
gagctgcacg acctgagcat ctacccttc ctgatgtacg ccaacaacca gctgatcagc	240
gtgggctacc tgaaccactt cgacatggac ttcctgtacc tgaccgacac caagaacacc	300

atcatcgacg agaggcacct gctgagggag ggcggcacc accacgag 348

<210> 21
 <211> 348
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> APG06351 mammalian codon optimized sequence

<400> 21
 atgaagacca tcaccagga gctgaagcag tacatcacc accgtgtcca gctgagcaac 60
 aacgaggcct gggagtgcga ggccttgag gaggccgccg agaacatcct gcccgagagg 120
 ttcataacg acagccccct ggtgcacctg accctggaga cctacaccta ctacaacgac 180
 gagctgcacg agctgagcat ctacccttc ctgatgtacg ccaacaacca gctgatcagc 240
 gtgggctacc tggaccactt cgacatggac ttcctgtacc tgaccgacac ccagaacatc 300
 atcatcgacg agaggcacct gctgaagcag ggcggccagg accacgag 348

<210> 22
 <211> 348
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> APG05963 mammalian codon optimized sequence

<400> 22
 atgaccctgg agcagcagct gaagcactac atcaccaacc tgttcaagct gccaaggac 60
 gagaagtggg agtgcgagag catcgaggag gtggccgaca acatcctgcc cgaccagtac 120
 gtgaggctgg gccccctgac caacaagatc ctgcacacct acacctacta cagcgacacc 180
 ctgcacgaga ggcacatcta ccccttcac ctgtactacc agaagcagct gatcgccatc 240
 ggctacatcg acgagaccaa cgacatggac ttcctgtacc tgcacaacac cgtgatgccc 300
 ctgctggacc agaggcacct gctggagaag gagaagcaca agaacgag 348

<210> 23
 <211> 345
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> APG09230 mammalian codon optimized sequence

<400> 23
atgaagaccc agatcaacga gctgaagagc tacatcacca acctgttcca gctgaccagc 60
caggagaact gggagtgcga ggccctgaac gagagcgccg acaacatcct gcccagagg 120
ttcctgaacg acagccccct gagcaacagg atcctggaga cctacaccta ctacaacaac 180
gagctgcacg agctgagcat ctacccttc ctgatgtact accagaacca gctgatcagc 240
atcggctacc tggaccacta cgacatggac ttcctgtacc tgaacgacac ccagaccacc 300
atcatcgacg agaggcacct gctgaagggc gagaacaagt acgag 345

<210> 24
<211> 348
<212> DNA
<213> Artificial Sequence

<220>
<223> APG04100 mammalian codon optimized sequence

<400> 24
atgaagacca ccaccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac 60
gaggagagct gggagtgcga ggtgctggac gaggtggccg agaacatcct gccccagg 120
ttcgtggacg gcagccccct gaccacctg accctggaga cctacaccta ctacaacaac 180
gagctgcacg acctgagcat ctacccttc ctgatgtacg ccaacaacca gctgatcagc 240
gtgggctacc tggaccactt cgacatggac ttcctgtacc tgaccgacac caagaacacc 300
atcatcgacg agaggcacct gctgaagcag ggcgagaaca accacgag 348

<210> 25
<211> 348
<212> DNA
<213> Artificial Sequence

<220>
<223> APG04001 mammalian codon optimized sequence

<400> 25
atgaagacca ccaccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac 60
aacgagacct gggagtgcga ggccctggag gaggccgccg agaacatcct gccaccagg 120
ttcgtggacc acacccccct ggccacctg accctggaca cctacaccta ctacaacgac 180
gagctgcacg agctgagcat ctacccttc ctgatgtacg ccaacaacca gctgatcagc 240

atcggctacc tggaccactt cgacatggac ttcctgtacc tgaccgacac caagaacatc	300
atcatcgacg agaggcacct gctgagggag ggcggcaacg acaggag	348

<210> 26
 <211> 348
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> APG02463 mammalian codon optimized sequence

<400> 26	
atgcagacca ccaccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac	60
aacgagacct gggagtgcga ggccctggag gaggccgccg agaacatcct gcccgagagg	120
ttcatcaacg gcagccccct ggcccacctg accctggaga cctacaccta ctacaacgac	180
gagctgcacg agctgagcat ctacccttc ctgatgtacg ccaacaacca gctgatcagc	240
atcggctacc tggaccactt cgacatggac ttcctgtacc tgaccgacac caagaacacc	300
atcatcgacg agaggcacct gctgaaggag ggcggcaaca accacgag	348

<210> 27
 <211> 348
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> APG09242 mammalian codon optimized sequence

<400> 27	
atgaagacca ccaccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac	60
aacgagacct gggagtgcga ggccctggag gacaccgccg agaacatcct gcccgagtgc	120
ttcatcaacg acacccccct ggcccacctg accctggaga cctacaccta ctacaacgac	180
gagctgcacg agctgagcat ctacccttc ctgatgtacg ccaacaacca gctgatcagc	240
gtgggctacc tggaccactt cgacatggac ttcctgtacc tgaccgacac caagaacacc	300
atcatcgacg agaggcacct gctgaaggag ggcggcaacg accacgag	348

<210> 28
 <211> 348
 <212> DNA
 <213> Artificial Sequence

<220>

<223> APG04080 mammalian codon optimized sequence

<400> 28

atgaagacca ccacccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac	60
aacgagacct gggagtgcga ggccctggag gaggccgccg ccaacatcct gcccagagg	120
ttcatcaaca acagccccct ggcccacctg accctggaga cctacaccta ctacaacgac	180
gagctgcacg agctgagcat ctacccttc ctgatgtacg ccaacaacca gctgatcagc	240
gtgggctacc tggaccactt cgacatggac ttctgtacc tgaccgacac caagaacacc	300
atcatcgacg agaggcacct gctgaaggag ggcggcaaca accacgag	348

<210> 29

<211> 336

<212> DNA

<213> Artificial Sequence

<220>

<223> APG05316 mammalian codon optimized sequence

<400> 29

atgacctga gcaagcagct gaagacctac atcaccgaga ggttcaagct gaactaccag	60
gagacctggg cctgcgagac catcgacgcc gtggccgagg acgtgctgcc cgagaagtac	120
atcaagaaca gccccctgga gcacaagatc ctgaacacct tcacctacta caacgacgag	180
ctgcacgaga tcagcatcta ccccttcctg tgctacctgg acaaggagct ggtggccatc	240
ggctacctgg acaacttcga cctggacttc atcttcctga acgacacca ccaggtgatc	300
atcgacgaga ggtacctgct gcagaagggc ggccag	336

<210> 30

<211> 339

<212> DNA

<213> Artificial Sequence

<220>

<223> APG03327 mammalian codon optimized sequence

<400> 30

atgaagacca ccacccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac	60
aacgagacct gggagtgcga ggccctggag gacgccgccg agaacatcct gccaccagg	120
ttctgggacc acacccccct ggcccacctg accctggaga cctacaccta ctacaacgac	180

gagctgcacg agctgagcat ctacccttc ctgatctacg ccaacaacca gctgatcagc	240
atcggctacc tggaccactt cgacatggac ttctgtacc tgaccgacac caagaacacc	300
atcatcgacg agaggcacct gctgcagaag ggcagggag	339

<210> 31
 <211> 348
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> APG01791 mammalian codon optimized sequence

<400> 31	
atgaagacca ccaccagga gctgaagcag tacatcacca ggctgttcca gctgagcaac	60
aacgagacct gggagtgcga ggccctggag gaggccgccg agaacatcct gccaccagg	120
ttcgtgggcc acacccccct ggcccacctg accctggaga cctacaccta ctacaacgac	180
gagctgcacg agctgagcat ctacccttc ctgatgtacg ccaacgacca gctgatcagc	240
atcggctacc tggaccactt cgacatggac ttctgtacc tgaccgacac caagaacacc	300
ctgatcgacg agaggcacct gctgaagcag ggcggcaacg accacgag	348

<210> 32
 <211> 348
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> APG05756 mammalian codon optimized sequence

<400> 32	
atgacctgg agcagcagct gaagcactac atcaccaacc tgttcagcct gccccaggac	60
gagaagtggg agtgcgagag catcgaggag gtggccgaca acatcctgcc cgaccagtac	120
gtgaggctgg gccccctgac caacaagatc ctgcacacct acacctacta cagcgacaag	180
ctgcacgaga ggcacatcta ccccttcac ctgtactacc agaagcagct gatcgccatc	240
ggctacatcg acgagatcaa cgacatggac ttctgtacc tgcacaacac cgtgatgccc	300
ctgctggacc agaggcacct gctggagaag gagaactaca acaacgag	348

<210> 33
 <211> 8
 <212> PRT

<213> Artificial Sequence

<220>

<223> Shared consensus C-terminus sequence

<400> 33

Lys Glu Gly Gly Asn Asp His Glu
1 5

<210> 34

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> APG05756 and APG05198 unique C-term seq

<400> 34

Glu Lys Glu Asn Tyr Asn Asn Glu
1 5

<210> 35

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> APG05963 unique C-term seq

<400> 35

Glu Lys Glu Lys His Lys Asn Glu
1 5

<210> 36

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> APG06702 unique C-terminal seq

<400> 36

Asp Lys Gly Asp Asp Asn His Asp
1 5

<210> 37

<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> APG05316 unique C-terminal seq

<400> 37

Gln Lys Gly Gly Gln
1 5

<210> 38
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> APG09230 unique C-term seq

<400> 38

Lys Gly Glu Asn Lys Tyr Glu
1 5

<210> 39
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> APG04100 unique C-term seq

<400> 39

Lys Gln Gly Glu Asn Asn His Glu
1 5

<210> 40
<211> 1071
<212> PRT
<213> Artificial Sequence

<220>
<223> RGN APG07433.1

<400> 40

Met Arg Glu Leu Asp Tyr Arg Ile Gly Leu Asp Ile Gly Thr Asn Ser
1 5 10 15

Ile Gly Trp Gly Val Ile Glu Leu Ser Trp Asn Lys Asp Arg Glu Arg
20 25 30

Tyr Glu Lys Val Arg Ile Val Asp Gln Gly Val Arg Met Phe Asp Arg
35 40 45

Ala Glu Met Pro Lys Thr Gly Ala Ser Leu Ala Glu Pro Arg Arg Ile
50 55 60

Ala Arg Ser Ser Arg Arg Arg Leu Asn Arg Lys Ser Gln Arg Lys Lys
65 70 75 80

Asn Ile Arg Asn Leu Leu Val Gln His Gly Val Ile Thr Gln Glu Glu
85 90 95

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

Ile Arg Leu Asp Gly Leu Asp Arg Leu Leu Asn His Phe Glu Trp Ala
115 120 125

Arg Leu Leu Ile His Leu Ala Gln Arg Arg Gly Phe Lys Ser Asn Arg
130 135 140

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

Gln Leu Asn Glu Lys Arg Leu Ser Leu Tyr Arg Thr Val Gly Glu Met
165 170 175

Trp Met Lys Asp Pro Asp Phe Ser Lys Tyr Asp Arg Lys Arg Asn Ser
180 185 190

Pro Asn Glu Tyr Val Phe Ser Val Ser Arg Ala Glu Leu Glu Lys Glu
195 200 205

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

Ser Lys Asp Leu Gln Glu Thr Tyr Leu Gln Ile Trp Thr His Gln Leu

225		230		235		240									
Pro	Phe	Ala	Ser	Gly	Asn	Ala	Ile	Leu	Asn	Lys	Val	Gly	Tyr	Cys	Ser
				245					250					255	
Leu	Leu	Lys	Gly	Lys	Glu	Arg	Arg	Ile	Pro	Lys	Ala	Thr	Tyr	Thr	Phe
			260					265					270		
Gln	Tyr	Phe	Ser	Ala	Leu	Asp	Gln	Val	Asn	Arg	Thr	Arg	Leu	Gly	Pro
		275					280					285			
Asp	Phe	Gln	Pro	Phe	Thr	Lys	Glu	Gln	Arg	Glu	Ile	Ile	Leu	Asn	Asn
	290					295					300				
Met	Phe	Gln	Arg	Thr	Asp	Tyr	Tyr	Lys	Lys	Lys	Thr	Ile	Pro	Glu	Val
305					310					315					320
Thr	Tyr	Tyr	Asp	Ile	Arg	Lys	Trp	Leu	Glu	Leu	Asp	Glu	Thr	Ile	Gln
				325					330					335	
Phe	Lys	Gly	Leu	Asn	Tyr	Asp	Pro	Asn	Glu	Glu	Leu	Lys	Lys	Ile	Glu
			340					345					350		
Lys	Lys	Pro	Phe	Ile	Asn	Leu	Lys	Ala	Phe	Tyr	Glu	Ile	Asn	Lys	Val
		355					360					365			
Val	Ala	Asn	Tyr	Ser	Glu	Arg	Thr	Asn	Glu	Thr	Phe	Ser	Thr	Leu	Asp
	370					375					380				
Tyr	Asp	Gly	Ile	Gly	Tyr	Ala	Leu	Thr	Val	Tyr	Lys	Thr	Asp	Lys	Asp
385					390					395					400
Ile	Arg	Ser	Tyr	Leu	Lys	Ser	Ser	His	Asn	Leu	Pro	Lys	Arg	Cys	Tyr
				405					410					415	
Asp	Asp	Gln	Leu	Ile	Glu	Glu	Leu	Leu	Ser	Leu	Ser	Tyr	Thr	Lys	Phe
			420					425					430		
Gly	His	Leu	Ser	Leu	Lys	Ala	Ile	Asn	His	Val	Leu	Ser	Ile	Met	Gln
		435					440					445			

Lys Gly Asn Thr Tyr Lys Glu Ala Val Asp Gln Leu Gly Tyr Asp Thr
450 455 460

Ser Gly Leu Lys Lys Glu Lys Arg Ser Lys Phe Leu Pro Pro Ile Ser
465 470 475 480

Asp Glu Ile Thr Asn Pro Ile Val Lys Arg Ala Leu Thr Gln Ala Arg
485 490 495

Lys Val Val Asn Ala Ile Ile Arg Arg His Gly Ser Pro His Ser Val
500 505 510

His Ile Glu Leu Ala Arg Glu Leu Ser Lys Asn His Asp Glu Arg Thr
515 520 525

Lys Ile Val Ser Ala Gln Asp Glu Asn Tyr Lys Lys Asn Lys Gly Ala
530 535 540

Ile Ser Ile Leu Ser Glu His Gly Ile Leu Asn Pro Thr Gly Tyr Asp
545 550 555 560

Ile Val Arg Tyr Lys Leu Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr
565 570 575

Ser Leu Lys Glu Ile Pro Ala Asp Thr Phe Phe Asn Glu Leu Lys Lys
580 585 590

Glu Arg Asn Gly Ala Pro Ile Leu Glu Val Asp His Ile Leu Pro Tyr
595 600 605

Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Ser
610 615 620

Asp Glu Asn Arg Lys Lys Gly Asn Arg Ile Pro Tyr Thr Tyr Phe Leu
625 630 635 640

Glu Thr Asn Lys Asp Trp Glu Ala Phe Glu Arg Tyr Val Arg Ser Asn
645 650 655

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

660

665

670

Leu Pro Arg Glu Ser Glu Leu Ile Lys Glu Arg His Leu Asn Asp Thr
 675 680 685

Arg Tyr Ala Ser Thr Phe Leu Lys Asn Phe Ile Glu Gln Asn Leu Gln
 690 695 700

Phe Lys Glu Ala Glu Asp Asn Pro Arg Lys Arg Arg Val Gln Thr Val
 705 710 715 720

Asn Gly Val Ile Thr Ala His Phe Arg Lys Arg Trp Gly Leu Glu Lys
 725 730 735

Asp Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
 740 745 750

Ala Cys Thr Asp His His Met Val Thr Arg Val Thr Glu Tyr Tyr Gln
 755 760 765

Ile Lys Glu Ser Asn Lys Ser Val Lys Lys Pro Tyr Phe Pro Met Pro
 770 775 780

Trp Glu Gly Phe Arg Asp Glu Leu Leu Ser His Leu Ala Ser Gln Pro
 785 790 795 800

Ile Ala Lys Lys Ile Ser Glu Glu Leu Lys Ala Gly Tyr Gln Ser Leu
 805 810 815

Asp Tyr Ile Phe Val Ser Arg Met Pro Lys Arg Ser Ile Thr Gly Ala
 820 825 830

Ala His Lys Gln Thr Ile Met Arg Lys Gly Gly Ile Asp Lys Lys Gly
 835 840 845

Lys Thr Ile Ile Ile Glu Arg Leu His Leu Lys Asp Ile Lys Phe Asp
 850 855 860

Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Met Ala Thr
 865 870 875 880

Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Glu His Gly Lys Asn Ser Lys
885 890 895

Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Thr Gly
900 905 910

Asn Leu Ile Lys Arg Val Lys Val Glu Gly Gln Ala Lys Ser Phe Val
915 920 925

Arg Glu Val Asn Gly Gly Val Ala Gln Asn Gly Asp Leu Val Arg Val
930 935 940

Asp Leu Phe Glu Lys Asp Asp Lys Tyr Tyr Met Val Pro Ile Tyr Val
945 950 955 960

Pro Asp Thr Val Cys Ser Glu Leu Pro Lys Lys Val Val Ala Ser Ser
965 970 975

Lys Gly Tyr Glu Gln Trp Leu Thr Leu Asp Asn Ser Phe Thr Phe Lys
980 985 990

Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asp Glu
995 1000 1005

Asp Arg Phe Leu Tyr Phe Gly Thr Leu Asp Ile Asp Ser Asp Arg
1010 1015 1020

Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Lys Asn Glu Tyr
1025 1030 1035

Arg Tyr Ser Leu Lys Thr Ile Glu Asp Leu Glu Lys Tyr Glu Val
1040 1045 1050

Gly Val Leu Gly Asp Leu Arg Leu Val Arg Lys Glu Thr Arg Arg
1055 1060 1065

Asn Phe His
1070

<210> 41

<211> 1071
<212> PRT
<213> Artificial Sequence

<220>

<223> nRGN APG07433.1

<400> 41

Met Arg Glu Leu Asp Tyr Arg Ile Gly Leu Ala Ile Gly Thr Asn Ser
1 5 10 15

Ile Gly Trp Gly Val Ile Glu Leu Ser Trp Asn Lys Asp Arg Glu Arg
20 25 30

Tyr Glu Lys Val Arg Ile Val Asp Gln Gly Val Arg Met Phe Asp Arg
35 40 45

Ala Glu Met Pro Lys Thr Gly Ala Ser Leu Ala Glu Pro Arg Arg Ile
50 55 60

Ala Arg Ser Ser Arg Arg Arg Leu Asn Arg Lys Ser Gln Arg Lys Lys
65 70 75 80

Asn Ile Arg Asn Leu Leu Val Gln His Gly Val Ile Thr Gln Glu Glu
85 90 95

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

Ile Arg Leu Asp Gly Leu Asp Arg Leu Leu Asn His Phe Glu Trp Ala
115 120 125

Arg Leu Leu Ile His Leu Ala Gln Arg Arg Gly Phe Lys Ser Asn Arg
130 135 140

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

Gln Leu Asn Glu Lys Arg Leu Ser Leu Tyr Arg Thr Val Gly Glu Met
165 170 175

Trp Met Lys Asp Pro Asp Phe Ser Lys Tyr Asp Arg Lys Arg Asn Ser

180

185

190

Pro Asn Glu Tyr Val Phe Ser Val Ser Arg Ala Glu Leu Glu Lys Glu
 195 200 205

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

Ser Lys Asp Leu Gln Glu Thr Tyr Leu Gln Ile Trp Thr His Gln Leu
 225 230 235 240

Pro Phe Ala Ser Gly Asn Ala Ile Leu Asn Lys Val Gly Tyr Cys Ser
 245 250 255

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

Gln Tyr Phe Ser Ala Leu Asp Gln Val Asn Arg Thr Arg Leu Gly Pro
 275 280 285

Asp Phe Gln Pro Phe Thr Lys Glu Gln Arg Glu Ile Ile Leu Asn Asn
 290 295 300

Met Phe Gln Arg Thr Asp Tyr Tyr Lys Lys Lys Thr Ile Pro Glu Val
 305 310 315 320

Thr Tyr Tyr Asp Ile Arg Lys Trp Leu Glu Leu Asp Glu Thr Ile Gln
 325 330 335

Phe Lys Gly Leu Asn Tyr Asp Pro Asn Glu Glu Leu Lys Lys Ile Glu
 340 345 350

Lys Lys Pro Phe Ile Asn Leu Lys Ala Phe Tyr Glu Ile Asn Lys Val
 355 360 365

Val Ala Asn Tyr Ser Glu Arg Thr Asn Glu Thr Phe Ser Thr Leu Asp
 370 375 380

Tyr Asp Gly Ile Gly Tyr Ala Leu Thr Val Tyr Lys Thr Asp Lys Asp
 385 390 395 400

Ile Arg Ser Tyr Leu Lys Ser Ser His Asn Leu Pro Lys Arg Cys Tyr
405 410 415

Asp Asp Gln Leu Ile Glu Glu Leu Leu Ser Leu Ser Tyr Thr Lys Phe
420 425 430

Gly His Leu Ser Leu Lys Ala Ile Asn His Val Leu Ser Ile Met Gln
435 440 445

Lys Gly Asn Thr Tyr Lys Glu Ala Val Asp Gln Leu Gly Tyr Asp Thr
450 455 460

Ser Gly Leu Lys Lys Glu Lys Arg Ser Lys Phe Leu Pro Pro Ile Ser
465 470 475 480

Asp Glu Ile Thr Asn Pro Ile Val Lys Arg Ala Leu Thr Gln Ala Arg
485 490 495

Lys Val Val Asn Ala Ile Ile Arg Arg His Gly Ser Pro His Ser Val
500 505 510

His Ile Glu Leu Ala Arg Glu Leu Ser Lys Asn His Asp Glu Arg Thr
515 520 525

Lys Ile Val Ser Ala Gln Asp Glu Asn Tyr Lys Lys Asn Lys Gly Ala
530 535 540

Ile Ser Ile Leu Ser Glu His Gly Ile Leu Asn Pro Thr Gly Tyr Asp
545 550 555 560

Ile Val Arg Tyr Lys Leu Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr
565 570 575

Ser Leu Lys Glu Ile Pro Ala Asp Thr Phe Phe Asn Glu Leu Lys Lys
580 585 590

Glu Arg Asn Gly Ala Pro Ile Leu Glu Val Asp His Ile Leu Pro Tyr
595 600 605

Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Ser

610

615

620

Asp Glu Asn Arg Lys Lys Gly Asn Arg Ile Pro Tyr Thr Tyr Phe Leu
 625 630 635 640

Glu Thr Asn Lys Asp Trp Glu Ala Phe Glu Arg Tyr Val Arg Ser Asn
 645 650 655

Lys Phe Phe Ser Lys Lys Lys Arg Glu Tyr Leu Leu Lys Arg Ala Tyr
 660 665 670

Leu Pro Arg Glu Ser Glu Leu Ile Lys Glu Arg His Leu Asn Asp Thr
 675 680 685

Arg Tyr Ala Ser Thr Phe Leu Lys Asn Phe Ile Glu Gln Asn Leu Gln
 690 695 700

Phe Lys Glu Ala Glu Asp Asn Pro Arg Lys Arg Arg Val Gln Thr Val
 705 710 715 720

Asn Gly Val Ile Thr Ala His Phe Arg Lys Arg Trp Gly Leu Glu Lys
 725 730 735

Asp Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
 740 745 750

Ala Cys Thr Asp His His Met Val Thr Arg Val Thr Glu Tyr Tyr Gln
 755 760 765

Ile Lys Glu Ser Asn Lys Ser Val Lys Lys Pro Tyr Phe Pro Met Pro
 770 775 780

Trp Glu Gly Phe Arg Asp Glu Leu Leu Ser His Leu Ala Ser Gln Pro
 785 790 795 800

Ile Ala Lys Lys Ile Ser Glu Glu Leu Lys Ala Gly Tyr Gln Ser Leu
 805 810 815

Asp Tyr Ile Phe Val Ser Arg Met Pro Lys Arg Ser Ile Thr Gly Ala
 820 825 830

Ala His Lys Gln Thr Ile Met Arg Lys Gly Gly Ile Asp Lys Lys Gly
835 840 845

Lys Thr Ile Ile Ile Glu Arg Leu His Leu Lys Asp Ile Lys Phe Asp
850 855 860

Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Met Ala Thr
865 870 875 880

Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Glu His Gly Lys Asn Ser Lys
885 890 895

Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Thr Gly
900 905 910

Asn Leu Ile Lys Arg Val Lys Val Glu Gly Gln Ala Lys Ser Phe Val
915 920 925

Arg Glu Val Asn Gly Gly Val Ala Gln Asn Gly Asp Leu Val Arg Val
930 935 940

Asp Leu Phe Glu Lys Asp Asp Lys Tyr Tyr Met Val Pro Ile Tyr Val
945 950 955 960

Pro Asp Thr Val Cys Ser Glu Leu Pro Lys Lys Val Val Ala Ser Ser
965 970 975

Lys Gly Tyr Glu Gln Trp Leu Thr Leu Asp Asn Ser Phe Thr Phe Lys
980 985 990

Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asp Glu
995 1000 1005

Asp Arg Phe Leu Tyr Phe Gly Thr Leu Asp Ile Asp Ser Asp Arg
1010 1015 1020

Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Lys Asn Glu Tyr
1025 1030 1035

Arg Tyr Ser Leu Lys Thr Ile Glu Asp Leu Glu Lys Tyr Glu Val

1040

1045

1050

Gly Val Leu Gly Asp Leu Arg Leu Val Arg Lys Glu Thr Arg Arg
1055 1060 1065

Asn Phe His
1070

<210> 42
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> SV40 NLS

<400> 42

Pro Lys Lys Lys Arg Lys Val
1 5

<210> 43
<211> 22
<212> PRT
<213> Artificial Sequence

<220>
<223> 3X FLAG tag

<400> 43

Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp Tyr
1 5 10 15

Lys Asp Asp Asp Asp Lys
20

<210> 44
<211> 32
<212> PRT
<213> Artificial Sequence

<220>
<223> Peptide linker 1

<400> 44

Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly Ser Glu Thr Pro Gly Thr

1 5 10 15

Ser Glu Ser Ala Thr Pro Glu Ser Ser Gly Gly Ser Ser Gly Gly Ser
 20 25 30

<210> 45
<211> 16
<212> PRT
<213> Artificial Sequence

<220>
<223> Nucleoplasmin NLS

<400> 45

Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys
1 5 10 15

<210> 46
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Peptide linker 2

<400> 46

Ser Gly Gly Ser Gly Gly Ser Gly Gly Ser
1 5 10

<210> 47
<211> 230
<212> PRT
<213> Staphylococcus spp

<223> APG09980_cytidine deaminase

<400> 47

Met Ala Ala Gly Pro Ala Pro Glu Ala Arg Ser Leu Met Asp Glu Gln
1 5 10 15

Thr Phe Leu Asp Asn Phe Asn Asn Leu Lys Tyr Pro Arg Lys Thr Tyr
 20 25 30

Leu Cys Tyr Glu Val Glu Leu Leu Val Gly Glu Asn His Ile Pro Leu
 35 40 45

Asp Asp Tyr Lys Gly Phe Val His Asn Glu Gly Phe Asp Met Gly Leu
50 55 60

Glu Arg Cys His Ala Glu Leu Ile Phe Leu Glu Arg Met Ala Ser Trp
65 70 75 80

Asn Leu Asp Thr Glu Leu Arg Tyr Arg Ile Thr Val Phe Ile Ser Trp
85 90 95

Ser Pro Cys Pro Glu Cys Ala Asp Glu Leu Val Lys Phe Leu Arg Glu
100 105 110

Asn Arg His Val Asn Leu Arg Ile Phe Ala Ala Arg Ile Tyr Asp Trp
115 120 125

Tyr Gln Gly Tyr Glu Ala Gly Leu Arg Ala Leu Lys Ala Ala Gly Ala
130 135 140

Glu Val Ala Met Met Thr Leu His Glu Phe Glu Tyr Cys Trp Asn Asn
145 150 155 160

Phe Val Asp His Gln Gln Asp Glu Asp Thr Pro Phe Pro Pro Trp Asp
165 170 175

Asn Leu Val Ala Arg Ser Glu Glu Leu Ser Gln Arg Leu Glu Gly Ile
180 185 190

Leu Gln Pro Ser Val Leu Val Phe Cys Trp Pro Ser Gln Val Ser Val
195 200 205

Thr Ala Ala His Ser Asp Ile Met Ser Gln Ala Ser Arg Ala Trp Glu
210 215 220

Lys Arg Arg Asp Pro Pro
225 230

<210> 48
<211> 182
<212> PRT
<213> Artificial Sequence

<220>

<223> APG07386-CTD_cytidine deaminase

<400> 48

Met Glu Ala Thr Tyr Pro His Ile Phe Tyr Phe His Phe Lys Asn Leu
1 5 10 15

Arg Lys Ala Tyr Gly Arg Asn Glu Thr Trp Leu Cys Phe Thr Met Glu
20 25 30

Ile Ile Lys Gln His Ser Thr Val Phe Trp Glu Thr Gly Val Phe Arg
35 40 45

Asn Gln Val Tyr Pro Glu Ser Leu Cys His Ala Glu Arg Cys Phe Leu
50 55 60

Ser Trp Phe Cys Glu Asp Ile Leu Ser Pro Asn Thr Asp Tyr Arg Val
65 70 75 80

Thr Trp Tyr Thr Ser Trp Ser Pro Cys Leu Asp Cys Ala Gly Glu Val
85 90 95

Ala Glu Phe Leu Ala Arg His Ser Asn Val Lys Leu Ala Ile Phe Ala
100 105 110

Ala Arg Leu Tyr Tyr Phe Trp Asp Pro His Tyr Gln Gln Gly Leu Arg
115 120 125

Ser Leu Ser Glu Lys Gly Ala Ser Val Glu Ile Met Gly Tyr Lys Asp
130 135 140

Phe Lys Tyr Cys Trp Glu Asn Phe Val Tyr Asn Gly Asp Glu Pro Phe
145 150 155 160

Lys Pro Trp Lys Gly Leu Lys Tyr Asn Phe Leu Phe Leu Asp Ser Lys
165 170 175

Leu Gln Glu Ile Leu Gln
180

<210> 49
<211> 4158
<212> DNA
<213> Artificial Sequence

<220>
<223> APG09980_nAPG07433.1

<400> 49
atggccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat 60
cacgatatcg actacaagga cgacgatgat aagatggccg ctggacctgc ccctgaagcc 120
agaagcctga tggacgagca gaccttcctg gataacttca acaacctgaa gtaccccaga 180
aagacctacc tgtgctacga ggtggaactg ctggtgggcg agaaccacat ccctctggac 240
gactacaagg gcttcgtgca caacgagggc tttgatatgg gcctggaacg gtgccacgcc 300
gagctgatct tcctggaaag aatggcctct tggaatctgg acaccgagct gagataccgg 360
atcaccgtgt tcatctcttg gtcccatgt cctgagtgcg ccgacgagct ggttaagttc 420
ctgcgggaaa atagacacgt gaacctgaga atctttgccg ctagaatcta cgactggtat 480
cagggctacg aggccggcct gcgggccctg aaagccgccg gagccgaagt ggccatgatg 540
accctgcacg agttcgagta ctgctggaac aacttcgtgg accaccagca ggacgaagat 600
acacctttcc cccctggga caacctggtc gccagatccg aggaactgag ccagagactg 660
gaaggcatcc tgcagcctag cgtgctggtg ttctgctggc ctagccaagt gtccgtgacc 720
gccgcccaca gcgacatcat gagccaggcc tctagagcct gggagaagcg gcgggacccc 780
cccagcggcg gatctagcgg cggctctagc ggcagtgaaa cccctggtac aagcgagtct 840
gccacacctg agagcagcgg cggaagcagc ggaggaagca tgagagagct ggactacaga 900
attggcctgg ccatcggcac caacagcatc ggatggggcg tgatcgagct gtcctggaac 960
aaagaccggg agagatacga gaaggtcaga atcgtggatc aaggcgtgag aatgttcgac 1020
agagccgaga tgcccaagac aggcgccagc ttagctgaac ccagaagaat cgccagatcc 1080
agcagacgga gactgaatcg caagtcccag agaaagaaaa acatccggaa cctgctggtg 1140
caacacggcg tgatcacaca ggaggaactg gatagcctgt accccctgag caaaaagagc 1200
atggacatct ggggcattcg gctcgacggc ctggacagac tcctcaatca tttcgagtgg 1260
gccagactgc tgatccacct ggctcagaga cggggcttta agtccaacag aaagagtga 1320
ctgaaagata cagagacagg caaggtgctg agcagcatcc aactgaacga gaaacggctg 1380

agcttgata	gaaccgtggg	cgagatgtgg	atgaaggacc	ccgacttctc	taaatacgat	1440
aggaagagaa	atagcccca	cgaatacgtg	ttcagcgtgt	ctagagccga	gctggaaaag	1500
gaaatcgtga	ccctgttcgc	cgcccagcgg	agattccaga	gcccttacgc	cagcaaagat	1560
ctgcaggaga	catatctgca	gatctggacc	caccaactgc	ctttcgccag	cggcaatgcc	1620
atcctgaaca	aggtcggata	ctgctccctg	ttgaaaggca	aagaaagaag	gattcccaag	1680
gctacataca	ccttccaata	cttctctgct	ctggaccagg	tgaatcggac	cagactggga	1740
cctgatttcc	agcccttcac	caaggagcaa	cgggaaatta	tcttgaacaa	catgttccag	1800
aggacagatt	actacaagaa	gaaaaccatc	cccgagggtga	cctactatga	catacgggaag	1860
tggctggaat	tggacgagac	aattcagttc	aagggcctga	actacgaccc	taacgaggaa	1920
ctgaagaaga	tcgagaagaa	gccttttatc	aatctgaagg	ccttctacga	gatcaacaag	1980
gtggtggcca	actacagcga	aagaaccaac	gagaccttct	ccaccctgga	ctacgacggc	2040
atcggctacg	ccctgaccgt	gtacaaaacc	gacaaggata	tccgcagcta	cctgaagagc	2100
agtcacaacc	tacctaagag	atgctacgac	gaccaactga	tcgaggaact	gctgagcctg	2160
agctacacaa	agttcggcca	cctgtccctg	aaagccatca	accacgtgct	gtctatcatg	2220
cagaagggca	atacctacaa	ggaagccgtg	gaccaactgg	gctacgacac	cagcggcctt	2280
aagaaggaga	agaggtccaa	gttcctgcca	cctatttctg	atgaaatcac	gaatccaatc	2340
gtgaaaaggg	ccctgaccca	ggccagaaaa	gtggtgaacg	ccataattag	aagacacgga	2400
tctcctcact	ccgtgcacat	cgagctggcc	agagagctga	gcaagaacca	cgacgagcgg	2460
acaaagatcg	tcagcgccca	ggatgaaaac	tacaagaaaa	acaagggcgc	tatcagcatc	2520
ctgtctgagc	acggcatcct	gaaccctaca	ggctacgaca	tcgtgagata	caaactgtgg	2580
aaggagcagg	gcgaacgggtg	cgcctacagc	ctgaaggaaa	tccctgccga	tacatttttc	2640
aacgagctga	agaaggaacg	caacggcgcc	cctatccttg	aagtggacca	catcctgccc	2700
tacagccagt	ccttcacgca	ctcctaccac	aacaagggtcc	tgggtgtacag	cgacgaaaac	2760
cggaaaaagg	gcaacagaat	cccttatacc	tacttcctgg	aaaccaacaa	ggattggggag	2820
gcctttgagc	ggtacgtgcg	gagcaacaaa	ttcttctcca	agaaaaagcg	agagtacctt	2880
ctgaagcggg	cttatctgcc	tagagaatct	gagctgatca	aagaacgcca	cctgaacgac	2940
accagatacg	cctctacctt	cctgaagaac	ttcatcgagc	agaacctgca	gttcaaggaa	3000

gccgaggaca accccagaaa aagacgggtg caaacctga acggcggttat caccgcccac	3060
ttcagaaagc ggtggggcct ggagaaggac cggcaggaga catacctcca tcacgctatg	3120
gacgccatca tcgtggcttg tacagaccac cacatgggtca ccagagtgac cgagtactat	3180
cagatcaagg aaagcaacaa gagcgtgaag aagccctatt ttcctatgcc ttgggaaggc	3240
ttccgggacg agctgctgag ccacttggct tctcagccta tcgccaagaa aatcagcgag	3300
gaactgaagg cgggctacca gagcctggac tacatcttcg tgtccagaat gcctaagaga	3360
agcattacag gcgctgctca taagcagacc atcatgcgga agggaggaat tgacaagaag	3420
ggcaaaacaa tcatcatcga acggctgcac ctgaaggata tcaagttcga cgagaacgga	3480
gatttcaaga tgggtgggcaa ggaacaggac atggccacat acgaagctat taaacagaga	3540
tacctggagc acggcaagaa tagcaagaag gccttcgaga cccctctgta caagcccagc	3600
aaaaagggca caggtaacct gatcaagcgg gtgaagggtg aaggacaggc caagagcttt	3660
gtgaggggaag tgaacggcgg agtggcccaa aatggcgatc tggttagagt tgatttgttt	3720
gagaaggatg ataagtacta catggtcccc atctacgtgc cagacaccgt gtgtagcgag	3780
ctgccccaaa aggtggctgc cagctctaag ggctatgagc agtggctgac actggataac	3840
agcttcacct ttaagttcag cctgtaccct tatgatctgg tgcggctggc caagggagat	3900
gaggatcggc tcctgtactt tggcaccctg gacatcgaca gcgacagact taacttcaag	3960
gacgtgaaca agccaagcaa gaagaacgag taccggtaca gcttgaaaac catcgaggac	4020
ttggagaagt acgaggtggg cgtgctgggc gatctaagac tgggtccggaa ggaaactcga	4080
agaaacttcc acagcggcgg aagcaaaaga cctgccgcta caaagaaggc cggccaggcc	4140
aagaaaaaga agtaatga	4158

<210> 50

<211> 4530

<212> DNA

<213> Artificial Sequence

<220>

<223> APG09980_nAPG07433.1_APG03399

<400> 50

atgcccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat	60

cacgatatcg actacaagga cgacgatgat aagatggccg ctggacctgc ccctgaagcc	120

agaagcctga tggacgagca gaccttcctg gataacttca acaacctgaa gtaccccaga	180
aagacctacc tgtgctacga ggtggaactg ctggtgggcg agaaccacat ccctctggac	240
gactacaagg gcttcgtgca caacgagggc tttgatatgg gccttgaacg gtgccacgcc	300
gagctgatct tcctggaaag aatggcctct tggaatctgg acaccgagct gagataccgg	360
atcaccgtgt tcatctcttg gtcccatgt cctgagtgcg ccgacgagct ggttaagttc	420
ctgcgggaaa atagacacgt gaacctgaga atctttgccg ctagaatcta cgactgggtat	480
cagggtacg aggccggcct gcgggccctg aaagccgccg gagccgaagt ggccatgatg	540
accctgcacg agttcgagta ctgctggaac aacttcgtgg accaccagca ggacgaagat	600
acacctttcc cccctggga caacctggtc gccagatccg aggaactgag ccagagactg	660
gaaggcatcc tgcagcctag cgtgctggtg ttctgctggc ctagccaagt gtccgtgacc	720
gccgcccaca gcgacatcat gagccaggcc tctagagcct gggagaagcg gcgggacccc	780
cccagcggcg gatctagcgg cggctctagc ggcagtgaac cccctggtac aagcgagtct	840
gccacacctg agagcagcgg cggaagcagc ggaggaagca tgagagagct ggactacaga	900
attggcctgg ccatcggcac caacagcatc ggatggggcg tgatcgagct gtcctggaac	960
aaagaccggg agagatacga gaaggtcaga atcgtggatc aaggcgtgag aatgttcgac	1020
agagccgaga tgcccaagac aggcgccagc ttagctgaac ccagaagaat cgccagatcc	1080
agcagacgga gactgaatcg caagtcccag agaaagaaaa acatccggaa cctgctggtg	1140
caacacggcg tgatcacaca ggaggaactg gatagcctgt accccctgag caaaaagagc	1200
atggacatct ggggcattcg gctcgacggc ctggacagac tcctcaatca tttcgagtgg	1260
gccagactgc tgatccacct ggctcagaga cggggcttta agtccaacag aaagagtga	1320
ctgaaagata cagagacagg caaggtgctg agcagcatcc aactgaacga gaaacggctg	1380
agcttgata gaaccgtggg cgagatgtgg atgaaggacc ccgacttctc taaatacgat	1440
aggaagagaa atagcccaa cgaatacgtg ttcagcgtgt ctagagccga gctggaaaag	1500
gaaatcgtga ccctgttcgc cgcccagcgg agattccaga gcccttacgc cagcaaagat	1560
ctgcaggaga catatctgca gatctggacc caccaactgc ctttcgccag cggcaatgcc	1620
atcctgaaca aggtcggata ctgctccctg ttgaaaggca aagaaagaag gattcccaag	1680
gctacataca ctttccaata cttctctgct ctggaccagg tgaatcggac cagactggga	1740

cctgatttcc agcccttcac caaggagcaa cgggaaatta tcttgaacaa catgttccag	1800
aggacagatt actacaagaa gaaaaccatc cccgaggtga cctactatga catacgggaag	1860
tggctggaat tggacgagac aattcagttc aagggcctga actacgaccc taacgaggaa	1920
ctgaagaaga tcgagaagaa gcctttttatc aatctgaagg ccttctacga gatcaacaag	1980
gtggtggcca actacagcga aagaaccaac gagaccttct ccaccctgga ctacgacggc	2040
atcggctacg ccctgaccgt gtacaaaacc gacaaggata tccgcagcta cctgaagagc	2100
agtcacaacc tacctaagag atgctacgac gaccaactga tcgaggaact gctgagcctg	2160
agctacacaa agttcggcca cctgtccctg aaagccatca accacgtgct gtctatcatg	2220
cagaagggca atacctacaa ggaagccgtg gaccaactgg gctacgacac cagcggcctt	2280
aagaaggaga agaggtccaa gttcctgcca cctatttctg atgaaatcac gaatccaatc	2340
gtgaaaaggg ccctgaccca ggccagaaaa gtggtgaacg ccataattag aagacacgga	2400
tctcctcact ccgtgcacat cgagctggcc agagagctga gcaagaacca cgacgagcgg	2460
acaaagatcg tcagcgccca ggatgaaaac tacaagaaaa acaagggcgc tatcagcatc	2520
ctgtctgagc acggcatcct gaaccctaca ggctacgaca tcgtgagata caaactgtgg	2580
aaggagcagg gcgaacggtg cgcctacagc ctgaaggaaa tccctgccga tacatttttc	2640
aacgagctga agaaggaacg caacggcgcc cctatccttg aagtggacca catcctgccc	2700
tacagccagt cttcatcga ctctaccac aacaagggtc tggtgtacag cgacgaaaac	2760
cggaaaaagg gcaacagaat cccttatacc tacttcctgg aaaccaacaa ggattgggag	2820
gcctttgagc ggtacgtgcg gagcaacaaa ttcttctcca agaaaaagcg agagtacctt	2880
ctgaagcggg cttatctgcc tagagaatct gagctgatca aagaacgcca cctgaacgac	2940
accagatacg cctctacctt cctgaagaac ttcacgagc agaacctgca gttcaaggaa	3000
gccgaggaca accccagaaa aagacgggtg caaacctga acggcgttat caccgcccac	3060
ttcagaaagc ggtggggcct ggagaaggac cggcaggaga catacctcca tcacgctatg	3120
gacgccatca tcgtggcttg tacagaccac cacatggtca ccagagtgaac cgagtactat	3180
cagatcaagg aaagcaacaa gagcgtgaag aagccctatt ttcctatgcc ttgggaaggc	3240
ttccgggacg agctgctgag ccacttggtc tctcagccta tcgccaagaa aatcagcgag	3300
gaactgaagg ccggctacca gagcctggac tacatcttcg tgtccagaat gcctaagaga	3360

agcattacag gcgctgctca taagcagacc atcatgcgga agggaggaat tgacaagaag	3420
ggcaaaacaa tcatcatcga acggctgcac ctgaaggata tcaagttcga cgagaacgga	3480
gatttcaaga tgggtgggcaa ggaacaggac atggccacat acgaagctat taaacagaga	3540
tacctggagc acggcaagaa tagcaagaag gccttcgaga cccctctgta caagcccagc	3600
aaaaagggca caggtaacct gatcaagcgg gtgaaggtgg aaggacaggc caagagcttt	3660
gtgaggggaag tgaacggcgg agtggcccaa aatggcgatc tggttagagt tgatttgttt	3720
gagaaggatg ataagtacta catggtcccc atctacgtgc cagacaccgt gtgtagcgag	3780
ctgccccaaa aggtggtcgc cagctctaag ggctatgagc agtggctgac actggataac	3840
agcttcacct ttaagttcag cctgtaccct tatgatctgg tgcggctggt caagggagat	3900
gaggatcggg tcctgtactt tggcaccctg gacatcgaca gcgacagact taacttcaag	3960
gacgtgaaca agccaagcaa gaagaacgag taccggtaca gcttgaaaac catcgaggac	4020
ttggagaagt acgaggtggg cgtgctgggc gatctaagac tgggtccggaa ggaaactcga	4080
agaaacttcc acagtggagg cagcggcggg tctggaggca gcatgaaaac caccacacag	4140
gagctgaagc agtacatcac cagactgttc caactgagca acaacgagac atgggagtg	4200
gaggccctgg aagaggccgc cgagaacatc ctgcctgaga gattcatcaa caacagcctg	4260
ctggcccacc tgacactgaa cacctacacc tactacaacg atgagctgca cgagctttct	4320
atctaccctt tcctgatgta cgccaacgac caactgatca gcatcggcta cctggaccac	4380
ttcgacatgg acttccttta tctgaccgac accaagaaca ccatcatcga cgagcggcac	4440
ctgctgaaac agggcgagaa caaccacgag agcggcggaa gcaaaagacc tgccgctaca	4500
aagaaggccg gccaggccaa gaaaaagaag	4530

<210> 51

<211> 4530

<212> DNA

<213> Artificial Sequence

<220>

<223> APG09980_nAPG07433.1_APG06702

<400> 51

atgcccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat	60

cacgatatcg actacaagga cgacgatgat aagatggccg ctggacctgc ccctgaagcc	120

agaagcctga tggacgagca gaccttcctg gataacttca acaacctgaa gtaccccaga	180
aagacctacc tgtgctacga ggtggaactg ctggtgggcg agaaccacat ccctctggac	240
gactacaagg gcttcgtgca caacgagggc tttgatatgg gccttgaacg gtgccacgcc	300
gagctgatct tcctggaaag aatggcctct tggaatctgg acaccgagct gagataccgg	360
atcaccgtgt tcatctcttg gtcccatgt cctgagtgcg ccgacgagct ggttaagttc	420
ctgcgggaaa atagacacgt gaacctgaga atctttgccg ctagaatcta cgactgggtat	480
cagggtacg aggccggcct gcgggccctg aaagccgccg gagccgaagt ggccatgatg	540
accctgcacg agttcgagta ctgctggaac aacttcgtgg accaccagca ggacgaagat	600
acacctttcc cccctggga caacctggtc gccagatccg aggaactgag ccagagactg	660
gaaggcatcc tgcagcctag cgtgctggtg ttctgctggc ctagccaagt gtccgtgacc	720
gccgcccaca gcgacatcat gagccaggcc tctagagcct gggagaagcg gcgggacccc	780
cccagcggcg gatctagcgg cggctctagc ggcagtgaac cccctggtac aagcgagtct	840
gccacacctg agagcagcgg cggaagcagc ggaggaagca tgagagagct ggactacaga	900
attggcctgg ccatcggcac caacagcatc ggatggggcg tgatcgagct gtcctggaac	960
aaagaccggg agagatacga gaaggtcaga atcgtggatc aaggcgtgag aatgttcgac	1020
agagccgaga tgcccaagac aggcgccagc ttagctgaac ccagaagaat cgccagatcc	1080
agcagacgga gactgaatcg caagtcccag agaaagaaaa acatccggaa cctgctggtg	1140
caacacggcg tgatcacaca ggaggaactg gatagcctgt accccctgag caaaaagagc	1200
atggacatct ggggcattcg gctcgacggc ctggacagac tcctcaatca tttcgagtgg	1260
gccagactgc tgatccacct ggctcagaga cggggcttta agtccaacag aaagagtga	1320
ctgaaagata cagagacagg caaggtgctg agcagcatcc aactgaacga gaaacggctg	1380
agcttgata gaaccgtggg cgagatgtgg atgaaggacc ccgacttctc taaatacgat	1440
aggaagagaa atagcccaa cgaatacgtg ttcagcgtgt ctagagccga gctggaaaag	1500
gaaatcgtga ccctgttcgc cgcccagcgg agattccaga gcccttacgc cagcaaagat	1560
ctgcaggaga catatctgca gatctggacc caccaactgc ctttcgccag cggcaatgcc	1620
atcctgaaca aggtcggata ctgctccctg ttgaaaggca aagaaagaag gattcccaag	1680
gctacataca ctttccaata cttctctgct ctggaccagg tgaatcggac cagactggga	1740

cctgatttcc agcccttcac caaggagcaa cgggaaatta tcttgaacaa catgttccag	1800
aggacagatt actacaagaa gaaaaccatc cccgaggtga cctactatga catacgggaag	1860
tggctggaat tggacgagac aattcagttc aagggcctga actacgaccc taacgaggaa	1920
ctgaagaaga tcgagaagaa gcctttttatc aatctgaagg ccttctacga gatcaacaag	1980
gtggtggcca actacagcga aagaaccaac gagaccttct ccaccctgga ctacgacggc	2040
atcggctacg ccctgaccgt gtacaaaacc gacaaggata tccgcagcta cctgaagagc	2100
agtcacaacc tacctaagag atgctacgac gaccaactga tcgaggaact gctgagcctg	2160
agctacacaa agttcggcca cctgtccctg aaagccatca accacgtgct gtctatcatg	2220
cagaagggca atacctacaa ggaagccgtg gaccaactgg gctacgacac cagcggcctt	2280
aagaaggaga agaggtccaa gttcctgcca cctatttctg atgaaatcac gaatccaatc	2340
gtgaaaaggg ccctgaccca ggccagaaaa gtggtgaacg ccataattag aagacacgga	2400
tctcctcact ccgtgcacat cgagctggcc agagagctga gcaagaacca cgacgagcgg	2460
acaaagatcg tcagcgccca ggatgaaaac tacaagaaaa acaagggcgc tatcagcatc	2520
ctgtctgagc acggcatcct gaaccctaca ggctacgaca tcgtgagata caaactgtgg	2580
aaggagcagg gcgaacggtg cgcctacagc ctgaaggaaa tccctgccga tacatttttc	2640
aacgagctga agaaggaacg caacggcgcc cctatccttg aagtggacca catcctgccc	2700
tacagccagt cttcatcga ctctaccac aacaagggtc tggtgtacag cgacgaaaac	2760
cggaaaaagg gcaacagaat cccttatacc tacttcctgg aaaccaacaa ggattgggag	2820
gcctttgagc ggtacgtgcg gagcaacaaa ttcttctcca agaaaaagcg agagtacctt	2880
ctgaagcggg cttatctgcc tagagaatct gagctgatca aagaacgcca cctgaacgac	2940
accagatacg cctctacctt cctgaagaac ttcacgagc agaacctgca gttcaaggaa	3000
gccgaggaca accccagaaa aagacgggtg caaacctga acggcgttat caccgcccac	3060
ttcagaaagc ggtggggcct ggagaaggac cggcaggaga catacctcca tcacgctatg	3120
gacgccatca tcgtggcttg tacagaccac cacatggtca ccagagtgaac cgagtactat	3180
cagatcaagg aaagcaacaa gagcgtgaag aagccctatt ttcctatgcc ttgggaaggc	3240
ttccgggacg agctgctgag ccacttggtc tctcagccta tcgccaagaa aatcagcgag	3300
gaactgaagg ccggctacca gagcctggac tacatcttcg tgtccagaat gcctaagaga	3360

agcattacag gcgctgctca taagcagacc atcatgcgga agggaggaat tgacaagaag	3420
ggcaaaacaa tcatcatcga acggctgcac ctgaaggata tcaagttcga cgagaacgga	3480
gatttcaaga tgggtgggcaa ggaacaggac atggccacat acgaagctat taaacagaga	3540
tacctggagc acggcaagaa tagcaagaag gccttcgaga cccctctgta caagcccagc	3600
aaaaagggca caggtaacct gatcaagcgg gtgaagggtg aaggacaggc caagagcttt	3660
gtgaggggaag tgaacggcgg agtggcccaa aatggcgatc tggttagagt tgatttgttt	3720
gagaaggatg ataagtacta catggtcccc atctacgtgc cagacaccgt gtgtagcgag	3780
ctgccccaaa aggtggctgc cagctctaag ggctatgagc agtggctgac actggataac	3840
agcttcacct ttaagttcag cctgtaccct tatgatctgg tgcggctggt caaggagat	3900
gaggatcggg tcctgtactt tggcaccctg gacatcgaca gcgacagact taacttcaag	3960
gacgtgaaca agccaagcaa gaagaacgag taccggtaca gcttgaaaac catcgaggac	4020
ttggagaagt acgaggtggg cgtgctgggc gatctaagac tgggtccgga ggaaactcga	4080
agaaacttcc acagtggagg cagcggcgga tctggaggca gcatgaccat cgagaacgag	4140
ctgaaggaat acatcactga tctgttcgag ctgccaagg acgagcagtg gctgtgcgag	4200
gccgtggacg aggtggccga taacatcctg cctgatgagt acgtgacctt gggccctctg	4260
agaaacaaga ccctggaaac cttcacctac tacagcaaca cactgcacga gcggaacatc	4320
tacccttca tcctgtacta ccagaacaga ctgatcgcca taggatatat cgacgaaaac	4380
tacgagatgg acttcatgta cctgcacaac accgccatca ccctgatcga ccagagatac	4440
ctgattgata agggcgacga caaccacgac agcggcgga gcaaaagacc tgccgctaca	4500
aagaaggccg gccaggccaa gaaaaagaag	4530

<210> 52

<211> 4530

<212> DNA

<213> Artificial Sequence

<220>

<223> APG09980_nAPG07433.1_APG05198

<400> 52

atgcccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat	60

cacgatatcg actacaagga cgacgatgat aagatggccg ctggacctgc ccctgaagcc	120

agaagcctga tggacgagca gaccttcctg gataacttca acaacctgaa gtaccccaga	180
aagacctacc tgtgctacga ggtggaactg ctggtgggcg agaaccacat ccctctggac	240
gactacaagg gcttcgtgca caacgagggc tttgatatgg gccttgaacg gtgccacgcc	300
gagctgatct tcctggaaag aatggcctct tggaatctgg acaccgagct gagataccgg	360
atcaccgtgt tcatctcttg gtcccatgt cctgagtgcg ccgacgagct ggttaagttc	420
ctgcgggaaa atagacacgt gaacctgaga atctttgccg ctagaatcta cgactgggtat	480
cagggtacg aggccggcct gcgggccctg aaagccgccg gagccgaagt ggccatgatg	540
accctgcacg agttcgagta ctgctggaac aacttcgtgg accaccagca ggacgaagat	600
acacctttcc cccctggga caacctggtc gccagatccg aggaactgag ccagagactg	660
gaaggcatcc tgcagcctag cgtgctggtg ttctgctggc ctagccaagt gtccgtgacc	720
gccgcccaca gcgacatcat gagccaggcc tctagagcct gggagaagcg gcgggacccc	780
cccagcggcg gatctagcgg cggctctagc ggcagtgaac cccctggtac aagcgagtct	840
gccacacctg agagcagcgg cggaagcagc ggaggaagca tgagagagct ggactacaga	900
attggcctgg ccatcggcac caacagcatc ggatggggcg tgatcgagct gtcctggaac	960
aaagaccggg agagatacga gaaggtcaga atcgtggatc aaggcgtgag aatgttcgac	1020
agagccgaga tgcccaagac aggcgccagc ttagctgaac ccagaagaat cgccagatcc	1080
agcagacgga gactgaatcg caagtcccag agaaagaaaa acatccggaa cctgctggtg	1140
caacacggcg tgatcacaca ggaggaactg gatagcctgt accccctgag caaaaagagc	1200
atggacatct ggggcattcg gctcgacggc ctggacagac tcctcaatca tttcgagtgg	1260
gccagactgc tgatccacct ggctcagaga cggggcttta agtccaacag aaagagtga	1320
ctgaaagata cagagacagg caaggtgctg agcagcatcc aactgaacga gaaacggctg	1380
agcttgata gaaccgtggg cgagatgtgg atgaaggacc ccgacttctc taaatacgat	1440
aggaagagaa atagcccaa cgaatacgtg ttcagcgtgt ctagagccga gctggaaaag	1500
gaaatcgtga ccctgttcgc cgcccagcgg agattccaga gcccttacgc cagcaaagat	1560
ctgcaggaga catatctgca gatctggacc caccaactgc ctttcgccag cggcaatgcc	1620
atcctgaaca aggtcggata ctgctccctg ttgaaaggca aagaaagaag gattcccaag	1680
gctacataca ctttccaata cttctctgct ctggaccagg tgaatcggac cagactggga	1740

cctgatttcc agcccttcac caaggagcaa cgggaaatta tcttgaacaa catgttccag	1800
aggacagatt actacaagaa gaaaaccatc cccgaggtga cctactatga catacgggaag	1860
tggctggaat tggacgagac aattcagttc aagggcctga actacgaccc taacgaggaa	1920
ctgaagaaga tcgagaagaa gcctttttatc aatctgaagg ccttctacga gatcaacaag	1980
gtggtggcca actacagcga aagaaccaac gagaccttct ccaccctgga ctacgacggc	2040
atcggctacg ccctgaccgt gtacaaaacc gacaaggata tccgcagcta cctgaagagc	2100
agtcacaacc tacctaagag atgctacgac gaccaactga tcgaggaact gctgagcctg	2160
agctacacaa agttcggcca cctgtccctg aaagccatca accacgtgct gtctatcatg	2220
cagaagggca atacctacaa ggaagccgtg gaccaactgg gctacgacac cagcggcctt	2280
aagaaggaga agaggtccaa gttcctgcca cctatttctg atgaaatcac gaatccaatc	2340
gtgaaaaggg ccctgaccca ggccagaaaa gtggtgaacg ccataattag aagacacgga	2400
tctcctcact ccgtgcacat cgagctggcc agagagctga gcaagaacca cgacgagcgg	2460
acaaagatcg tcagcgccca ggatgaaaac tacaagaaaa acaagggcgc tatcagcatc	2520
ctgtctgagc acggcatcct gaaccctaca ggctacgaca tcgtgagata caaactgtgg	2580
aaggagcagg gcgaacggtg cgcctacagc ctgaaggaaa tccctgccga tacatttttc	2640
aacgagctga agaaggaacg caacggcgcc cctatccttg aagtggacca catcctgccc	2700
tacagccagt cttcatcga ctctaccac aacaagggtcc tgggtgtacag cgacgaaaac	2760
cggaaaaagg gcaacagaat cccttatacc tacttcctgg aaaccaacaa ggattgggag	2820
gcctttgagc ggtacgtgcg gagcaacaaa ttcttctcca agaaaaagcg agagtacctt	2880
ctgaagcggg cttatctgcc tagagaatct gagctgatca aagaacgcca cctgaacgac	2940
accagatacg cctctacctt cctgaagaac ttcacgagc agaacctgca gttcaaggaa	3000
gccgaggaca accccagaaa aagacgggtg caaacctga acggcgttat caccgcccac	3060
ttcagaaagc ggtggggcct ggagaaggac cggcaggaga catacctcca tcacgctatg	3120
gacgccatca tcgtggcttg tacagaccac cacatggtca ccagagtgaac cgagtactat	3180
cagatcaagg aaagcaacaa gagcgtgaag aagccctatt ttcctatgcc ttgggaaggc	3240
ttccgggacg agctgctgag ccacttggtt tctcagccta tcgccaagaa aatcagcgag	3300
gaactgaagg ccggctacca gagcctggac tacatcttcg tgtccagaat gcctaagaga	3360

agcattacag gcgctgctca taagcagacc atcatgcgga agggaggaat tgacaagaag	3420
ggcaaaacaa tcatcatcga acggctgcac ctgaaggata tcaagttcga cgagaacgga	3480
gatttcaaga tgggtgggcaa ggaacaggac atggccacat acgaagctat taaacagaga	3540
tacctggagc acggcaagaa tagcaagaag gccttcgaga cccctctgta caagcccagc	3600
aaaaagggca caggtaacct gatcaagcgg gtgaagggtg aaggacaggc caagagcttt	3660
gtgaggggaag tgaacggcgg agtggcccaa aatggcgatc tggttagagt tgatttgttt	3720
gagaaggatg ataagtacta catggtcccc atctacgtgc cagacaccgt gtgtagcgag	3780
ctgccccaaa aggtggctgc cagctctaag ggctatgagc agtggctgac actggataac	3840
agcttcacct ttaagttcag cctgtaccct tatgatctgg tgcggctggt caaggagat	3900
gaggatcggg tcctgtactt tggcaccctg gacatcgaca gcgacagact taacttcaag	3960
gacgtgaaca agccaagcaa gaagaacgag taccggtaca gcttgaaaac catcgaggac	4020
ttggagaagt acgaggtggg cgtgctgggc gatctaagac tgggtccggaa ggaaactcga	4080
agaaacttcc acagtggagg cagcggcgga tctggaggca gcatgaccct ggaaaagcaa	4140
ctgaaggaat acatcaccaa cctgttcaac ctgccaagg acgagaagtg ggagtgcgag	4200
agcatcgagg aagtggccga caacatcctg cctgatcagt acatcagact gggccctctg	4260
accaacaaga tcctgcatac atacacctac tacagcgaca ccctgcacga gcggcacatc	4320
tacccttca tcctttatta ccagaaacaa ctgatcgcca tcggctacat cgacgagaca	4380
aacgacatgg acttcctgta cctgcacaac accgtgatgc ctctgctgga ccagagacac	4440
ctgctggaaa aagagaacta caacaacgag agcggcgga gcaaaagacc tgccgctaca	4500
aagaaggccg gccaggccaa gaaaaagaag	4530

<210> 53

<211> 4014

<212> DNA

<213> Artificial Sequence

<220>

<223> APG07386-CTD_nAPG07433.1

<400> 53

atgcccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat	60

cacgatatcg actacaagga cgacgatgat aagatggagg ccacctaccc ccacatcttc	120

tacttccact tcaagaacct gcggaaggcc tacggcagaa acgaaacctg gctgtgtttc	180
acaatggaaa tcatcaagca gcacagcacc gtgttctggg agacaggcgt gttcagaaat	240
caagtgtacc ctgagagcct gtgccacgcc gaaagatgct ttctgagctg gttctgtgaa	300
gatatcctga gccctaacac cgactaccgg gtgacatggc acacaagctg gtcccatgt	360
ctggattgcg ccggcgaggt ggccgagttc ctggccaggc attctaactg gaagctggct	420
atctttgccg ctagactgta ctatttctgg gacctcact accagcaggg cctgagaagc	480
ctgagcgaga agggagccag cgtggaaatc atgggctaca aggacttcaa atactgctgg	540
gaaaactttg tgtacaacgg cgacgagcca ttcaagccct ggaagggcct gaagtacaac	600
ttcctgttcc tggacagcaa gctgcaggag attctgcaga gcggcggatc tagcggcggc	660
tctagcggca gtgaaacccc tgggtacaagc gagtctgcca cacctgagag cagcggcggg	720
agcagcggag gaagcatgag agagctggac tacagaattg gcctggccat cggcaccaac	780
agcatcggat ggggcgtgat cgagctgtcc tggaacaaag accgggagag atacgagaag	840
gtcagaatcg tggatcaagg cgtgagaatg ttcgacagag ccgagatgcc caagacaggc	900
gccagcttag ctgaaccag aagaatcgcc agatccagca gacggagact gaatcgcaag	960
tcccagagaa agaaaaacat ccggaacctg ctggtgcaac acggcgtgat cacacaggag	1020
gaactggata gcctgtaccc cctgagcaaa aagagcatgg acatctgggg cattcggctc	1080
gacggcctgg acagactcct caatcatttc gagtgggcca gactgctgat ccacctggct	1140
cagagacggg gctttaagtc caacagaaag agtgaactga aagatacaga gacaggcaag	1200
gtgctgagca gcatccaact gaacgagaaa cggctgagct tgtatagaac cgtgggcgag	1260
atgtggatga aggacccga cttctctaaa tacgatagga agagaaatag cccaacgaa	1320
tacgtgttca gcgtgtctag agccgagctg gaaaaggaaa tcgtgaccct gttcgccgcc	1380
cagcggagat tccagagccc ttacgccagc aaagatctgc aggagacata tctgcagatc	1440
tggaccacc aactgccttt cgccagcggc aatgccatcc tgaacaaggc cggatactgc	1500
tccctgttga aaggcaaaga aagaaggatt cccaaggcta catacacctt ccaatacttc	1560
tctgctctgg accaggtgaa tcggaccaga ctgggacctg atttccagcc cttaccaag	1620
gagcaacggg aaattatctt gaacaacatg ttccagagga cagattacta caagaagaaa	1680
accatccccg aggtgacctt ctatgacata cggaagtggc tggaattgga cgagacaatt	1740

cagttcaagg gcctgaacta cgaccctaac gaggaactga agaagatcga gaagaagcct	1800
tttatcaatc tgaaggcctt ctacgagatc aacaaggtgg tggccaacta cagcgaaaga	1860
accaacgaga ctttctccac cctggactac gacggcatcg gctacgccct gaccgtgtac	1920
aaaaccgaca aggatatccg cagctacctg aagagcagtc acaacctacc taagagatgc	1980
tacgacgacc aactgatcga ggaactgctg agcctgagct acacaaagtt cggccacctg	2040
tccctgaaag ccatcaacca cgtgctgtct atcatgcaga agggcaatac ctacaaggaa	2100
gccgtggacc aactgggcta cgacaccagc ggccttaaga aggagaagag gtccaagttc	2160
ctgccaccta tttctgatga aatcacgaat ccaatcgtga aaagggccct gacccaggcc	2220
agaaaagtgg tgaacgcat aattagaaga cacggatctc ctactccgt gcacatcgag	2280
ctggccagag agctgagcaa gaaccacgac gagcggacaa agatcgtcag cgcccaggat	2340
gaaaactaca agaaaaacaa gggcgctatc agcatcctgt ctgagcacgg catcctgaac	2400
cctacaggct acgacatcgt gagatacaaa ctgtggaagg agcagggcga acggtgcgcc	2460
tacagcctga aggaaatccc tgccgataca tttttcaacg agctgaagaa ggaacgcaac	2520
ggcgccccta tccttgaagt ggaccacatc ctgccctaca gccagtcctt catcgactcc	2580
taccacaaca aggtcctggt gtacagcgac gaaaaccgga aaaagggcaa cagaatccct	2640
tatacctact tcctggaac caacaaggat tgggaggcct ttgagcggta cgtgcggagc	2700
aacaaattct tctccaagaa aaagcgagag taccttctga agcgggctta tctgcctaga	2760
gaatctgagc tgatcaaaga acgccacctg aacgacacca gatacgctc taccttctg	2820
aagaacttca tcgagcagaa cctgcagttc aaggaagccg aggacaaccc cagaaaaaga	2880
cgggtgcaaa ccgtgaacgg cgttatcacc gccacttca gaaagcggtg gggcctggag	2940
aaggaccggc aggagacata cctccatcac gctatggacg ccatcatcgt ggcttgtaca	3000
gaccaccaca tggtcaccag agtgaccgag tactatcaga tcaaggaaag caacaagagc	3060
gtgaagaagc cctattttcc tatgccttgg gaaggcttcc gggacgagct gctgagccac	3120
ttggcttctc agcctatcgc caagaaaatc agcgaggaac tgaaggccgg ctaccagagc	3180
ctggactaca tcttcgtgtc cagaatgcct aagagaagca ttacaggcgc tgctcataag	3240
cagaccatca tgcggaaggg aggaattgac aagaagggca aaacaatcat catcgaacgg	3300
ctgcacctga aggatatcaa gttcgacgag aacggagatt tcaagatggt gggcaaggaa	3360

caggacatgg ccacatacga agctattaaa cagagatacc tggagcacgg caagaatagc	3420
aagaaggcct tcgagacccc tctgtacaag cccagcaaaa agggcacagg taacctgac	3480
aagcgggtga aggtggaagg acaggccaag agctttgtga ggggaagtga cggcggagtg	3540
gccccaaatg gcgatctggt tagagttgat ttgtttgaga aggatgataa gtactacatg	3600
gtcccatct acgtgccaga caccgtgtgt agcgagctgc caaaaaagggt ggtcgccagc	3660
tctaagggt atgagcagtg gctgacactg gataacagct tcacctttaa gttcagcctg	3720
tacccttatg atctggtgcg gctggtcaag ggagatgagg atcggttcct gtactttggc	3780
accctggaca tcgacagcga cagacttaac ttcaaggacg tgaacaagcc aagcaagaag	3840
aacgagtacc ggtacagctt gaaaaccatc gaggacttgg agaagtacga ggtgggcgtg	3900
ctgggcgatc taagactggt ccggaaggaa actcgaagaa acttccacag cggcgggaagc	3960
aaaagacctg ccgctacaaa gaaggccggc caggccaaga aaaagaagta atga	4014

<210> 54

<211> 4386

<212> DNA

<213> Artificial Sequence

<220>

<223> APG07386-CTD_nAPG07433.1_APG03399

<400> 54

atggccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat	60
cacgatatcg actacaagga cgacgatgat aagatggagg ccacctaccc ccacatcttc	120
tacttccact tcaagaacct gcggaaggcc tacggcagaa acgaaacctg gctgtgtttc	180
acaatggaaa tcatcaagca gcacagcacc gtgttctggg agacaggcgt gttcagaaat	240
caagtgtacc ctgagagcct gtgccacgcc gaaagatgct ttctgagctg gttctgtgaa	300
gatatcctga gccctaacac cgactaccgg gtgacatggt acacaagctg gtcccatgt	360
ctggattgcg ccggcgaggt ggccgagttc ctggccaggc attctaactg gaagctggct	420
atctttgccg ctagactgta ctatttctgg gaccctcact accagcaggg cctgagaagc	480
ctgagcgaga agggagccag cgtggaaatc atgggctaca aggacttcaa atactgctgg	540
gaaaactttg tgtacaacgg cgacgagcca ttcaagccct ggaagggcct gaagtacaac	600
ttcctgttcc tggacagcaa gctgcaggag attctgcaga gcggcggatc tagcggcggc	660

tctagcggca	gtgaaacccc	tggtacaagc	gagtctgcc	cacctgagag	cagcggcgga	720
agcagcggag	gaagcatgag	agagctggac	tacagaattg	gcctggccat	cggcaccaac	780
agcatcggat	ggggcgtgat	cgagctgtcc	tggaacaaag	accgggagag	atacgagaag	840
gtcagaatcg	tggatcaagg	cgtgagaatg	ttcgacagag	ccgagatgcc	caagacaggc	900
gccagcttag	ctgaaccag	aagaatcgcc	agatccagca	gacggagact	gaatcgcaag	960
tcccagagaa	agaaaaacat	ccggaacctg	ctggtgcaac	acggcgtgat	cacacaggag	1020
gaactggata	gcctgtaccc	cctgagcaaa	aagagcatgg	acatctgggg	cattcggctc	1080
gacggcctgg	acagactcct	caatcatttc	gagtggggcca	gactgctgat	ccacctggct	1140
cagagacggg	gctttaagtc	caacagaaaag	agtgaactga	aagatacaga	gacaggcaag	1200
gtgctgagca	gcatccaact	gaacgagaaa	cggctgagct	tgtatagaac	cgtgggcgag	1260
atgtggatga	aggaccccga	cttctctaaa	tacgatagga	agagaaatag	ccccaacgaa	1320
tacgtgttca	gcgtgtctag	agccgagctg	gaaaaggaaa	tcgtgaccct	gttcgccgcc	1380
cagcggagat	tccagagccc	ttacgccagc	aaagatctgc	aggagacata	tctgcagatc	1440
tggaccacc	aactgccttt	cgccagcggc	aatgccatcc	tgaacaaggt	cggatactgc	1500
tccctgttga	aaggcaaaga	aagaaggatt	ccaaggcta	catacacctt	ccaatacttc	1560
tctgctctgg	accaggtgaa	tcggaccaga	ctgggacctg	atttccagcc	cttcaccaag	1620
gagcaacggg	aaattatctt	gaacaacatg	ttccagagga	cagattacta	caagaagaaa	1680
accatccccg	aggtgacct	ctatgacata	cggaagtggc	tggaattgga	cgagacaatt	1740
cagttcaagg	gcctgaacta	cgaccctaac	gaggaactga	agaagatcga	gaagaagcct	1800
tttatcaatc	tgaaggcctt	ctacgagatc	aacaagggtg	tggccaacta	cagcgaaaga	1860
accaacgaga	ccttctccac	cctggactac	gacggcatcg	gctacgccct	gaccgtgtac	1920
aaaaccgaca	aggatatccg	cagctacctg	aagagcagtc	acaacctacc	taagagatgc	1980
tacgacgacc	aactgatcga	ggaactgctg	agcctgagct	acacaaagtt	cggccacctg	2040
tccctgaaag	ccatcaacca	cgtgctgtct	atcatgcaga	agggaatac	ctacaaggaa	2100
gccgtggacc	aactgggcta	cgacaccagc	ggccttaaga	aggagaagag	gtccaagttc	2160
ctgccaccta	tttctgatga	aatcacgaat	ccaatcgtga	aaagggccct	gaccaggcc	2220
agaaaagtgg	tgaacgcat	aattagaaga	cacggatctc	ctcactccgt	gcacatcgag	2280

ctggccagag agctgagcaa gaaccacgac gagcggacaa agatcgtcag cgcccaggat	2340
gaaaactaca agaaaaacaa gggcgctatc agcatcctgt ctgagcacgg catcctgaac	2400
cctacaggct acgacatcgt gagatacaaa ctgtggaagg agcagggcga acggtgcgcc	2460
tacagcctga aggaaatccc tgccgataca tttttcaacg agctgaagaa ggaacgcaac	2520
ggcgccccta tccttgaagt ggaccacatc ctgccctaca gccagtcctt catcgactcc	2580
taccacaaca aggtcctggg gtacagcgac gaaaaccgga aaaagggcaa cagaatccct	2640
tatacctact tcctggaac caacaaggat tgggaggcct ttgagcggta cgtgcggagc	2700
aacaaattct tctccaagaa aaagcgagag taccttctga agcgggctta tctgcctaga	2760
gaatctgagc tgatcaaaga acgccacctg aacgacacca gatacgctc taccttcctg	2820
aagaacttca tcgagcagaa cctgcagttc aaggaagccg aggacaaccc cagaaaaaga	2880
cgggtgcaaa ccgtgaacgg cgttatcacc gcccacttca gaaagcggtg gggcctggag	2940
aaggaccggc aggagacata cctccatcac gctatggacg ccatcatcgt ggcttgtaca	3000
gaccaccaca tggtcaccag agtgaccgag tactatcaga tcaaggaaag caacaagagc	3060
gtgaagaagc cctattttcc tatgccttgg gaaggcttcc gggacgagct gctgagccac	3120
ttggcttctc agcctatcgc caagaaaatc agcgaggaac tgaaggccgg ctaccagagc	3180
ctggactaca tcttcgtgtc cagaatgcct aagagaagca ttacaggcgc tgctcataag	3240
cagaccatca tgcggaaggg aggaattgac aagaagggca aaacaatcat catcgaacgg	3300
ctgcacctga aggatatcaa gttcgacgag aacggagatt tcaagatggg gggcaaggaa	3360
caggacatgg ccacatacga agctattaaa cagagatacc tggagcacgg caagaatagc	3420
aagaaggcct tcgagacccc tctgtacaag ccagcaaaa agggcacagg taacctgatc	3480
aagcgggtga aggtggaagg acaggccaag agctttgtga ggggaagtga cggcggagtg	3540
gccccaaatg gcgatctggg tagagttgat ttgtttgaga aggatgataa gtactacatg	3600
gtcccatct acgtgccaga caccgtgtgt agcgagctgc ccaaaaaggt ggtcgccagc	3660
tctaagggct atgagcagtg gctgacactg gataacagct tcacctttaa gttcagcctg	3720
tacccttatg atctgggtgcg gctgggtcaag ggagatgagg atcggttcct gtactttggc	3780
accctggaca tcgacagcga cagacttaac ttcaaggacg tgaacaagcc aagcaagaag	3840
aacgagtacc ggtacagctt gaaaaccatc gaggacttgg agaagtacga ggtgggcgtg	3900

ctgggcatc taagactggt ccggaaggaa actcgaagaa acttccacag tggaggcagc	3960
ggcggatctg gaggcagcat gaaaaccacc acacaggagc tgaagcagta catcaccaga	4020
ctgttccaac tgagcaaaa cgagacatgg gaggcagagg ccctggaaga ggccgcccag	4080
aacatcctgc ctgagagatt catcaacaac agcctgctgg cccacctgac actgaacacc	4140
tacacctact acaacgatga gctgcacgag ctttctatct accccttcct gatgtacgcc	4200
aacgaccaac tgatcagcat cggctacctg gaccacttcg acatggactt cttttatctg	4260
accgacacca agaaccat catcgacgag cggcacctgc tgaaacaggg cgagaacaac	4320
cacgagagcg gcggaagcaa aagacctgcc gctacaaaga aggccggcca ggccaagaaa	4380
aagaag	4386

<210> 55
 <211> 4386
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> APG07386-CTD_nAPG07433.1_APG06702

<400> 55	
atggccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat	60
cacgatatcg actacaagga cgacgatgat aagatggagg ccacctaccc ccacatcttc	120
tactttccact tcaagaacct gcggaaggcc tacggcagaa acgaaacctg gctgtgtttc	180
acaatggaaa tcatcaagca gcacagcacc gtgttctggg agacaggcgt gttcagaaat	240
caagtgtacc ctgagagcct gtgccacgcc gaaagatgct ttctgagctg gttctgtgaa	300
gatatcctga gccctaacac cgactaccgg gtgacatggt acacaagctg gtcccatgt	360
ctggattgcg ccggcgaggt ggccgagttc ctggccaggc attctaactg gaagctggct	420
atctttgccg ctagactgta ctatttctgg gaccctcact accagcaggg cctgagaagc	480
ctgagcgaga agggagccag cgtggaaatc atgggctaca aggacttcaa atactgctgg	540
gaaaactttg tgtacaacgg cgacgagcca ttcaagccct ggaagggcct gaagtacaac	600
ttcctgttcc tggacagcaa gctgcaggag attctgcaga gcggcggatc tagcggcggc	660
tctagcggca gtgaaacccc tggtacaagc gaggctgcca cacctgagag cagcggcggg	720
agcagcggag gaagcatgag agagctggac tacagaattg gcctggccat cggcaccaac	780

agcatcggat	ggggcgtgat	cgagctgtcc	tggaacaaag	accgggagag	atacgagaag	840
gtcagaatcg	tggtatcaagg	cgtgagaatg	ttcgacagag	ccgagatgcc	caagacaggc	900
gccagcttag	ctgaaccag	aagaatcgcc	agatccagca	gacggagact	gaatcgcaag	960
tcccagagaa	agaaaaacat	ccggaacctg	ctggtgcaac	acggcgtgat	cacacaggag	1020
gaactggata	gcctgtaccc	cctgagcaaa	aagagcatgg	acatctgggg	cattcggctc	1080
gacggcctgg	acagactcct	caatcatttc	gagtgggcca	gactgctgat	ccacctggct	1140
cagagacggg	gctttaagtc	caacagaaaag	agtgaactga	aagatacaga	gacaggcaag	1200
gtgctgagca	gcatccaact	gaacgagaaa	cggctgagct	tgtatagaac	cgtgggagag	1260
atgtggatga	aggaccccg	cttctctaaa	tacgatagga	agagaaatag	ccccaacgaa	1320
tacgtgttca	gcgtgtctag	agccgagctg	gaaaaggaaa	tcgtgaccct	gttcgccgcc	1380
cagcggagat	tccagagccc	ttacgccagc	aaagatctgc	aggagacata	tctgcagatc	1440
tggacccacc	aactgccttt	cgccagcggc	aatgccatcc	tgaacaaggt	cggatactgc	1500
tccctgttga	aaggcaaaga	aagaaggatt	cccaaggcta	catacacctt	ccaatacttc	1560
tctgctctgg	accaggtgaa	tcggaccaga	ctgggacctg	atttccagcc	cttcaccaag	1620
gagcaacggg	aaattatctt	gaacaacatg	ttccagagga	cagattacta	caagaagaaa	1680
accatccccg	aggtgaccta	ctatgacata	cggaagtggc	tggaattgga	cgagacaatt	1740
cagttcaagg	gcctgaacta	cgaccctaac	gaggaactga	agaagatcga	gaagaagcct	1800
tttatcaatc	tgaaggcctt	ctacgagatc	aacaaggctg	tggccaacta	cagcgaaaga	1860
accaacgaga	ccttctccac	cctggactac	gacggcatcg	gctacgccct	gaccgtgtac	1920
aaaaccgaca	aggatatccg	cagctacctg	aagagcagtc	acaacctacc	taagagatgc	1980
tacgacgacc	aactgatcga	ggaactgctg	agcctgagct	acacaaagtt	cggccacctg	2040
tccctgaaag	ccatcaacca	cgtgctgtct	atcatgcaga	agggaataac	ctacaaggaa	2100
gccgtggacc	aactgggcta	cgacaccagc	ggccttaaga	aggagaagag	gtccaagttc	2160
ctgccaccta	tttctgatga	aatcacgaat	ccaatcgtga	aaagggccct	gaccagggcc	2220
agaaaagtgg	tgaacgcat	aattagaaga	cacggatctc	ctcactccgt	gcacatcgag	2280
ctggccagag	agctgagcaa	gaaccacgac	gagcggacaa	agatcgtcag	cgcccaggat	2340
gaaaactaca	agaaaaacaa	gggcgctatc	agcatcctgt	ctgagcacgg	catcctgaac	2400

cctacaggct acgacatcgt gagatacaaa ctgtggaagg agcagggcga acggtgcgcc	2460
tacagcctga aggaaatccc tgccgataca tttttcaacg agctgaagaa ggaacgcaac	2520
ggcgccccta tccttgaagt ggaccacatc ctgccctaca gccagtcctt catcgactcc	2580
taccacaaca aggtcctgggt gtacagcgac gaaaaccgga aaaagggcaa cagaatccct	2640
tatacctact tcctggaaac caacaaggat tgggaggcct ttgagcggta cgtgcggagc	2700
aacaaattct tctccaagaa aaagcgagag taccttctga agcgggctta tctgcctaga	2760
gaatctgagc tgatcaaaga acgccacctg aacgacacca gatacgctc taccttctg	2820
aagaacttca tcgagcagaa cctgcagttc aaggaagccg aggacaaccc cagaaaaaga	2880
cgggtgcaaa ccgtgaacgg cgttatcacc gcccacttca gaaagcggtg gggcctggag	2940
aaggaccggc aggagacata cctccatcac gctatggacg ccatcatcgt ggcttgtaca	3000
gaccaccaca tggtcaccag agtgaccgag tactatcaga tcaaggaaag caacaagagc	3060
gtgaagaagc cctattttcc tatgccttgg gaaggcttcc gggacgagct gctgagccac	3120
ttggcttctc agcctatcgc caagaaaatc agcgaggaac tgaaggccgg ctaccagagc	3180
ctggactaca tcttcgtgtc cagaatgcct aagagaagca ttacaggcgc tgctcataag	3240
cagaccatca tgcggaaggg aggaattgac aagaagggca aaacaatcat catcgaacgg	3300
ctgcacctga aggatatcaa gttcgacgag aacggagatt tcaagatgggt gggcaaggaa	3360
caggacatgg ccacatacga agctattaaa cagagatacc tggagcacgg caagaatagc	3420
aagaaggcct tcgagacccc tctgtacaag cccagcaaaa agggcacagg taacctgatc	3480
aagcgggtga aggtggaagg acaggccaag agctttgtga gggaagtga cggcggagt	3540
gccccaaatg gcgatctgggt tagagttgat ttgtttgaga aggatgataa gtactacatg	3600
gtcccatct acgtgccaga caccgtgtgt agcgagctgc ccaaaaaggt ggtcgccagc	3660
tctaagggt atgagcagtg gctgacactg gataacagct tcacctttaa gttcagcctg	3720
tacccttatg atctggtgcg gctgggtcaag ggagatgagg atcggttcct gtactttggc	3780
accctggaca tcgacagcga cagacttaac ttcaaggacg tgaacaagcc aagcaagaag	3840
aacgagtacc ggtacagctt gaaaaccatc gaggacttgg agaagtacga ggtgggcgtg	3900
ctgggcgatc taagactgggt ccggaaggaa actcgaagaa acttccacag tggaggcagc	3960
ggcggatctg gaggcagcat gaccatcgag aacgagctga aggaatacat cactgatctg	4020

ttcgagctgc ccaaggacga gcagtggctg tgcgaggccg tggacgaggt ggccgataac	4080
atcctgcctg atgagtagct gaccttgggc cctctgagaa acaagaccct ggaaaccttc	4140
acctactaca gcaacacact gcacgagcgg aacatctacc ccttcacacct gtactaccag	4200
aacagactga tcgccatagg atatatcgac gaaaactacg agatggactt catgtacctg	4260
cacaacaccg ccatcacctt gatcgaccag agatacctga ttgataaggg cgacgacaac	4320
cacgacagcg gcggaagcaa aagacctgcc gctacaaaga aggccggcca ggccaagaaa	4380
aagaag	4386

<210> 56
 <211> 4386
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> APG07386-CTD_nAPG07433.1_APG05198

<400> 56	
atggccccta agaagaaaag aaaggtggac tacaaggacc acgacggcga ctacaaagat	60
cacgatatcg actacaagga cgacgatgat aagatggagg ccacctaccc ccacatcttc	120
tactttccact tcaagaacct gcggaaggcc tacggcagaa acgaaacctg gctgtgtttc	180
acaatggaaa tcatcaagca gcacagcacc gtgttctggg agacaggcgt gttcagaaat	240
caagtgtacc ctgagagcct gtgccacgcc gaaagatgct ttctgagctg gttctgtgaa	300
gatatcctga gccctaacac cgactaccgg gtgacatggg acacaagctg gtccccatgt	360
ctggattgcg ccggcgaggt ggccgagttc ctggccaggc attctaactg gaagctggct	420
atctttgccg ctagactgta ctatttctgg gacctcact accagcaggg cctgagaagc	480
ctgagcgaga agggagccag cgtggaaatc atgggctaca aggacttcaa atactgctgg	540
gaaaactttg tgtacaacgg cgacgagcca ttcaagccct ggaagggcct gaagtacaac	600
ttcctgttcc tggacagcaa gctgcaggag attctgcaga gcggcgatc tagcggcggc	660
tctagcggca gtgaaacccc tgggtacaagc gagtctgcca cacctgagag cagcggcgga	720
agcagcggag gaagcatgag agagctggac tacagaattg gcctggccat cggcaccaac	780
agcatcggat ggggcgtgat cgagctgtcc tggaacaaag accgggagag atacgagaag	840
gtcagaatcg tggatcaagg cgtgagaatg ttcgacagag ccgagatgcc caagacaggc	900

gccagcttag	ctgaaccag	aagaatcgcc	agatccagca	gacggagact	gaatcgcaag	960
tcccagagaa	agaaaaacat	ccggaacctg	ctggtgcaac	acggcgtgat	cacacaggag	1020
gaactggata	gcctgtaccc	cctgagcaaa	aagagcatgg	acatctgggg	cattcggctc	1080
gacggcctgg	acagactcct	caatcatttc	gagtgggcca	gactgctgat	ccacctggct	1140
cagagacggg	gctttaagtc	caacagaaaag	agtgaactga	aagatacaga	gacaggcaag	1200
gtgctgagca	gcatccaact	gaacgagaaa	cggctgagct	tgtatagaac	cgtgggcgag	1260
atgtggatga	aggaccccga	cttctctaaa	tacgatagga	agagaaatag	ccccaacgaa	1320
tacgtgttca	gcgtgtctag	agccgagctg	gaaaaggaaa	tcgtgaccct	gttcgccgcc	1380
cagcggagat	tccagagccc	ttacgccagc	aaagatctgc	aggagacata	tctgcagatc	1440
tggaccacc	aactgccttt	cgccagcggc	aatgccatcc	tgaacaaggt	cggatactgc	1500
tccctgttga	aaggcaaaga	aagaaggatt	cccaaggcta	catacacctt	ccaatacttc	1560
tctgctctgg	accaggtgaa	tcggaccaga	ctgggacctg	atttccagcc	cttcaccaag	1620
gagcaacggg	aaattatctt	gaacaacatg	ttccagagga	cagattacta	caagaagaaa	1680
accatccccg	aggtgacctt	ctatgacata	cggaagtggc	tggaattgga	cgagacaatt	1740
cagttcaagg	gcctgaacta	cgaccctaac	gaggaactga	agaagatcga	gaagaagcct	1800
tttatcaatc	tgaaggcctt	ctacgagatc	aacaaggctg	tggccaacta	cagcgaaaga	1860
accaacgaga	ccttctccac	cctggactac	gacggcatcg	gctacgccct	gaccgtgtac	1920
aaaaccgaca	aggatatccg	cagctacctg	aagagcagtc	acaacctacc	taagagatgc	1980
tacgacgacc	aactgatcga	ggaactgctg	agcctgagct	acacaaagtt	cggccacctg	2040
tccctgaaag	ccatcaacca	cgtgctgtct	atcatgcaga	agggaataac	ctacaaggaa	2100
gccgtggacc	aactgggcta	cgacaccagc	ggccttaaga	aggagaagag	gtccaagtgc	2160
ctgccacctt	tttctgatga	aatcacgaat	ccaatcgtga	aaagggccct	gaccagggcc	2220
agaaaagtgg	tgaacgccat	aattagaaga	cacggatctc	ctcactccgt	gcacatcgag	2280
ctggccagag	agctgagcaa	gaaccacgac	gagcggacaa	agatcgtcag	cgcccaggat	2340
gaaaactaca	agaaaaacaa	gggcgctatc	agcatcctgt	ctgagcacgg	catcctgaac	2400
cctacaggct	acgacatcgt	gagatacaaa	ctgtggaagg	agcagggcga	acggtgcgcc	2460
tacagcctga	aggaaatccc	tgccgataca	tttttcaacg	agctgaagaa	ggaacgcaac	2520

ggcgccccta	tccttgaagt	ggaccacatc	ctgccctaca	gccagtcctt	catcgactcc	2580
taccacaaca	aggtcctggt	gtacagcgac	gaaaaccgga	aaaaggggcaa	cagaatccct	2640
tatacctact	tcctggaaac	caacaaggat	tgggaggcct	ttgagcggta	cgtgcggagc	2700
aacaaattct	tctccaagaa	aaagcgagag	taccttctga	agcgggctta	tctgcctaga	2760
gaatctgagc	tgatcaaaga	acgccacctg	aacgacacca	gatacgctc	taccttcctg	2820
aagaacttca	tcgagcagaa	cctgcagttc	aaggaagccg	aggacaaccc	cagaaaaaga	2880
cgggtgcaaa	ccgtgaacgg	cgttatcacc	gcccacttca	gaaagcggtg	gggcctggag	2940
aaggaccggc	aggagacata	cctccatcac	gctatggacg	ccatcatcgt	ggcttgtaca	3000
gaccaccaca	tggtcaccag	agtgaccgag	tactatcaga	tcaaggaaag	caacaagagc	3060
gtgaagaagc	cctattttcc	tatgccttgg	gaaggcttcc	gggacgagct	gctgagccac	3120
ttggcttctc	agcctatcgc	caagaaaatc	agcgaggaac	tgaaggccgg	ctaccagagc	3180
ctggactaca	tcttcgtgtc	cagaatgcct	aagagaagca	ttacaggcgc	tgctcataag	3240
cagaccatca	tgcggaaggg	aggaattgac	aagaagggca	aaacaatcat	catcgaacgg	3300
ctgcacctga	aggatatcaa	gttcgacgag	aacggagatt	tcaagatggg	gggcaaggaa	3360
caggacatgg	ccacatacga	agctatttaa	cagagatacc	tggagcacgg	caagaatagc	3420
aagaaggcct	tcgagacccc	tctgtacaag	cccagcaaaa	agggcacagg	taacctgac	3480
aagcgggtga	aggtggaagg	acaggccaag	agctttgtga	gggaagtga	cggcggagtg	3540
gccccaaatg	gcgatctggt	tagagttgat	ttgtttgaga	aggatgataa	gtactacatg	3600
gtcccatct	acgtgccaga	caccgtgtgt	agcgagctgc	ccaaaagggt	ggtcgccagc	3660
tctaagggt	atgagcagtg	gctgacactg	gataacagct	tcacctttaa	gttcagcctg	3720
tacccttatg	atctggtgcg	gctggtcaag	ggagatgagg	atcggttcct	gtactttggc	3780
accctggaca	tcgacagcga	cagacttaac	ttcaaggacg	tgaacaagcc	aagcaagaag	3840
aacgagtacc	ggtacagctt	gaaaaccatc	gaggacttgg	agaagtacga	ggtgggcgtg	3900
ctgggcgatc	taagactggt	ccggaaggaa	actcgaagaa	acttccacag	tggaggcagc	3960
ggcggatctg	gaggcagcat	gaccctggaa	aagcaactga	aggaatacat	caccaacctg	4020
ttcaacctgc	ccaaggacga	gaagtgggag	tgcgagagca	tcgaggaagt	ggccgacaac	4080
atcctgcctg	atcagtacat	cagactgggc	cctctgacca	acaagatcct	gcatacatac	4140

acctactaca gcgacaccct gcacgagcgg cacatctacc ctttcacctt ttattaccag	4200
aaacaactga tcgccatcgg ctacatcgac gagacaaacg acatggactt cctgtacctg	4260
cacaacaccg tgatgcctct gctggaccag agacacctgc tggaaaaaga gaactacaac	4320
aacgagagcg gcggaagcaa aagacctgcc gctacaaaga aggccggcca ggccaagaaa	4380
aagaag	4386

<210> 57
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> SGN000143 target sequence

<400> 57	
catggcagta cattagagca	20

<210> 58
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> SGN000169 target sequence

<400> 58	
cacatctcga gcaagacgtt	20

<210> 59
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> SGN000173 target sequence

<400> 59	
cttctatagc ctccttcccc	20

<210> 60
 <211> 25
 <212> DNA
 <213> Artificial Sequence

<220>

<223> SGN000929 target sequence

<400> 60

atctggaggg aacttacagc atatg

25

<210> 61

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> SGN000930 target sequence

<400> 61

gaacaactca aatggaaatg aatat

25

<210> 62

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> SGN001101 target sequence

<400> 62

atattttctt taatggtgcc aggca

25

<210> 63

<211> 130

<212> RNA

<213> Artificial Sequence

<220>

<223> SGN000143 sgRNA sequence

<400> 63

cauggcagua cauuagagca gucauaguuc cauuaaagcc aaaaguggcu uugauguuuc

60

uaugauaagg guuucgaccc guggcgucgg ggauccgucg ccgauugaaa ugggcuucuc

120

cccauuuauu

130

<210> 64

<211> 130

<212> RNA

<213> Artificial Sequence

<220>

<223> SGN000169 sgRNA sequence

<400> 64
cacaucucga gcaagacguu gucauaguuc cauuaaagcc aaaaguggcu uugauguuuc 60
uagauaagg guuucgaccc guggcgucgg ggauccgug ccgauugaaa ugggcuucuc 120
cccauuuauu 130

<210> 65
<211> 130
<212> RNA
<213> Artificial Sequence

<220>
<223> SGN000173 sgRNA sequence

<400> 65
cuucuauagc cuccuucucc gucauaguuc cauuaaagcc aaaaguggcu uugauguuuc 60
uagauaagg guuucgaccc guggcgucgg ggauccgug ccgauugaaa ugggcuucuc 120
cccauuuauu 130

<210> 66
<211> 135
<212> RNA
<213> Artificial Sequence

<220>
<223> SGN000929 sgRNA sequence

<400> 66
aucuggaggg aacuuacagc auauggucau aguuccauga aagccaaaag uggcuuugau 60
guuucuauga uaagggguuc ggcccguggc gucggggau ggcugcccau uccgaugggc 120
uucucccca uuuuu 135

<210> 67
<211> 135
<212> RNA
<213> Artificial Sequence

<220>
<223> SGN000930 sgRNA sequence

<400> 67
gaacaacuca aauggaaaug aaauaugucau aguuccauga aagccaaaag uggcuuugau 60
guuucuauga uaagggguuc ggcccguggc gucggggau ggcugcccau uccgaugggc 120
uucucccca uuuuu 135

<210> 68
<211> 130
<212> RNA
<213> Artificial Sequence

<220>
<223> SGN001101 sgRNA sequence

<400> 68
auauuuuucu uaauggugcc aggcaguc au aguuuccauua aagccaaaag uggcuuugau 60
guuucuauga uaaggguuuc gacccguggc gucggggaucc gccugcccau ugaaaugggc 120
uucuccccau 130

<210> 69
<211> 200
<212> DNA
<213> Artificial Sequence

<220>
<223> SGN000143 genomic locus

<400> 69
gaagtcagca atcttcagct ctcccttgag ccctaagagc agattttctg gctttatgtc 60
tctgtgaatc accttcttcc catggcagta cattagagca tctgccaact cctccatgat 120
ctgggagggg caacgacagg caatcagaca aggatggacc tccagctaca agcagcacac 180
cctcagcctc gagcccccta 200

<210> 70
<211> 200
<212> DNA
<213> Artificial Sequence

<220>
<223> SGN000169 genomic locus

<400> 70
tccagcaggt cagcaaagaa tttatagccc cccttgagca cacagagggc tacaatgtga 60
tggcctccca tctccttcat cacatctcga gcaagacgtt cagtcctaca gaaataaaat 120
caggaattta atagaaagtt tcatacatta aactttataa caaacacctc ttagtcatta 180
aacttcaca ccaacctggg 200

<210> 71
<211> 200
<212> DNA
<213> Artificial Sequence

<220>
<223> SGN000173 genomic locus

<400> 71
gtataactgg ccagccttta atacattgct gcttagagtc aaagcatgta ctagagttg 60
gtatgattta tcttttttgggt cttctatagc ctccttcccc atcccatca gtcttaatca 120
gtcttggttac gttatgacta atctttgggg atttgtgcag aatgttattt tagataagca 180
aaaacgagca aaatagggga 200

<210> 72
<211> 200
<212> DNA
<213> Artificial Sequence

<220>
<223> SGN000929 genomic locus

<400> 72
gactctcca cccagacct cctagttctt gggcaggtga acaactacag cttttccatg 60
cctccttggt gagagatctg gagggaaactt acagcatatg agcacctgct agggagagcc 120
cctggccctg gcaggaccga tacctcggct cctggacacc atttctattt actcaatata 180
cactgtttgt acaaagtaca 200

<210> 73
<211> 200
<212> DNA
<213> Artificial Sequence

<220>
<223> SGN000930 genomic locus

<400> 73
tttatcatca caatacctcg ggtccagatg ggccccttag tgagttccgc taccacagga 60
aggatgactc aataagaaca actcaaattg aaatgaatat gggcccaggg tgggtgtgag 120
gagtggcagg ggattcttga gagacaggac cattgcccac agcctatgtg agtactgttg 180
cttgtcttga aagaaggcaa 200

<210> 74
<211> 200
<212> DNA
<213> Artificial Sequence

<220>
<223> SGN001101 genomic locus

<400> 74
tttcttacct cttctagttg gcatgctttg atgacgcttc tgtatctata ttcacatag 60
gaaacaccaa agatgatatt ttctttaatg gtgccaggca taatccagga aaactgagaa 120
cagaatgaaa ttcttcact gtgcttaatt ttaccctctg aaggctccag ttctccata 180
atcaccatta gaagtgaagt 200

<210> 75
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> meganuclease family motif

<400> 75

Leu Ala Gly Leu Ile Asp Ala Asp Gly
1 5

<210> 76
<211> 182
<212> PRT
<213> Artificial Sequence

<220>
<223> Truncated APG09980_cystidine deaminase

<400> 76

Met Asp Glu Gln Thr Phe Leu Asp Asn Phe Asn Asn Leu Lys Tyr Pro
1 5 10 15

Arg Lys Thr Tyr Leu Cys Tyr Glu Val Glu Leu Leu Val Gly Glu Asn
20 25 30

His Ile Pro Leu Asp Asp Tyr Lys Gly Phe Val His Asn Glu Gly Phe
35 40 45

Asp Met Gly Leu Glu Arg Cys His Ala Glu Leu Ile Phe Leu Glu Arg
50 55 60

Met Ala Ser Trp Asn Leu Asp Thr Glu Leu Arg Tyr Arg Ile Thr Val
65 70 75 80

Phe Ile Ser Trp Ser Pro Cys Pro Glu Cys Ala Asp Glu Leu Val Lys
85 90 95

Phe Leu Arg Glu Asn Arg His Val Asn Leu Arg Ile Phe Ala Ala Arg
100 105 110

Ile Tyr Asp Trp Tyr Gln Gly Tyr Glu Ala Gly Leu Arg Ala Leu Lys
115 120 125

Ala Ala Gly Ala Glu Val Ala Met Met Thr Leu His Glu Phe Glu Tyr
130 135 140

Cys Trp Asn Asn Phe Val Asp His Gln Gln Asp Glu Asp Thr Pro Phe
145 150 155 160

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

Leu Glu Gly Ile Leu Gln
180

<210> 77
<211> 201
<212> PRT
<213> Artificial Sequence

<220>
<223> APG05840_cytidine deaminase

<400> 77

Met Glu Ala Ser Pro Ala Ser Arg Pro Arg Pro Leu Met Asp Pro His
1 5 10 15

Met Phe Thr Gly Asn Phe Thr Asn Asn Pro Arg Val Phe Gly Leu His
20 25 30

Gln Thr Tyr Leu Cys Tyr Glu Val Lys Arg Gln Gly Pro Asp Gly Thr
35 40 45

Arg Asp Leu Met Asn Glu Gln Arg Asp Phe Leu Cys Asn Gln Ala Lys
50 55 60

Asn His Phe Ser Gly Ser Glu Asp His His Ala Glu Arg Cys Phe Leu
65 70 75 80

Asp Arg Ile Pro Ser Trp Gln Leu Asp Pro Ala Gln Thr Tyr Arg Val
85 90 95

Thr Cys Phe Ile Ser Trp Ser Pro Cys Phe Ser Cys Ala Gln Glu Val
100 105 110

Ala Glu Phe Leu His Glu Asn Pro His Val Asn Leu Arg Ile Phe Ala
115 120 125

Ala Arg Ile Tyr Asp Tyr Leu Pro Arg Tyr Glu Glu Gly Leu Gln Met
130 135 140

Leu Gln Asn Ala Gly Ala Gln Val Ser Ile Met Thr Ser Glu Glu Phe
145 150 155 160

Gly His Cys Trp Asp Thr Phe Val Asp Arg Gln Gly His Pro Phe Gln
165 170 175

Pro Trp Glu Gly Leu Asp Glu His Ser Gln Ala Leu Ser Gly Arg Leu
180 185 190

Gln Ala Ile Leu Gln Asn Gln Gly Asn
195 200

<210> 78

<211> 185

<212> PRT

<213> Artificial Sequence

<220>

<223> Truncated APG05840_cytidine deaminase

<400> 78

Met Asp Pro His Met Phe Thr Gly Asn Phe Thr Asn Asn Pro Arg Val
1 5 10 15

Phe Gly Leu His Gln Thr Tyr Leu Cys Tyr Glu Val Lys Arg Gln Gly
20 25 30

Pro Asp Gly Thr Arg Asp Leu Met Asn Glu Gln Arg Asp Phe Leu Cys
35 40 45

Asn Gln Ala Lys Asn His Phe Ser Gly Ser Glu Asp His His Ala Glu
50 55 60

Arg Cys Phe Leu Asp Arg Ile Pro Ser Trp Gln Leu Asp Pro Ala Gln
65 70 75 80

Thr Tyr Arg Val Thr Cys Phe Ile Ser Trp Ser Pro Cys Phe Ser Cys
85 90 95

Ala Gln Glu Val Ala Glu Phe Leu His Glu Asn Pro His Val Asn Leu
100 105 110

Arg Ile Phe Ala Ala Arg Ile Tyr Asp Tyr Leu Pro Arg Tyr Glu Glu
115 120 125

Gly Leu Gln Met Leu Gln Asn Ala Gly Ala Gln Val Ser Ile Met Thr
130 135 140

Ser Glu Glu Phe Gly His Cys Trp Asp Thr Phe Val Asp Arg Gln Gly
145 150 155 160

His Pro Phe Gln Pro Trp Glu Gly Leu Asp Glu His Ser Gln Ala Leu
165 170 175

Ser Gly Arg Leu Gln Ala Ile Leu Gln
180 185

<210> 79

<211> 192

<212> PRT

<213> Artificial Sequence

<220>

<223> APG00868_cytidine deaminase

<400> 79

Met Glu Pro Trp Arg Pro Ser Pro Arg Asn Pro Met Asp Arg Ile Asp
1 5 10 15

Pro Asn Thr Phe Arg Phe His Phe Pro Asn Leu Leu Tyr Ala Ser Gly
20 25 30

Arg Lys Leu Cys Tyr Leu Cys Phe Gln Val Glu Thr Gly Asp Tyr Phe
35 40 45

Ser Cys Asp Asp Ser Asp Arg Gly Val Phe Arg Asn Lys Val His Pro
50 55 60

Trp Ala Arg Cys His Ala Glu Gln Cys Phe Leu Ser Trp Phe Arg Asp
65 70 75 80

Gln Tyr Pro Cys Arg Asp Glu Tyr Tyr Asn Val Thr Trp Phe Leu Ser
85 90 95

Trp Ser Pro Cys Pro Thr Cys Ala Glu Glu Val Val Glu Phe Leu Glu
100 105 110

Glu Tyr Arg Asn Leu Thr Leu Ser Ile Phe Thr Ser Arg Leu Tyr Tyr
115 120 125

Phe Tyr His Pro Asn Tyr Gln Gln Gly Leu Arg Lys Leu Trp Asp Ala
130 135 140

Gly Val Gln Leu Asp Ile Met Ser Cys Asp Asp Phe Glu His Cys Trp
145 150 155 160

Asp Asn Phe Val Asp His Lys Gly Met Arg Phe Gln Arg Arg Asn Leu
165 170 175

Leu Lys Asp Tyr Asp Phe Leu Ala Ala Glu Leu Gln Glu Ile Leu Arg
180 185 190

<210> 80

<211> 178

<212> PRT
<213> Artificial Sequence

<220>
<223> Truncated APG00868_cytidine deaminase

<400> 80

Met Asp Pro Asn Thr Phe Arg Phe His Phe Pro Asn Leu Leu Tyr Ala
1 5 10 15

Ser Gly Arg Lys Leu Cys Tyr Leu Cys Phe Gln Val Glu Thr Gly Asp
20 25 30

Tyr Phe Ser Cys Asp Asp Ser Asp Arg Gly Val Phe Arg Asn Lys Val
35 40 45

His Pro Trp Ala Arg Cys His Ala Glu Gln Cys Phe Leu Ser Trp Phe
50 55 60

Arg Asp Gln Tyr Pro Cys Arg Asp Glu Tyr Tyr Asn Val Thr Trp Phe
65 70 75 80

Leu Ser Trp Ser Pro Cys Pro Thr Cys Ala Glu Glu Val Val Glu Phe
85 90 95

Leu Glu Glu Tyr Arg Asn Leu Thr Leu Ser Ile Phe Thr Ser Arg Leu
100 105 110

Tyr Tyr Phe Tyr His Pro Asn Tyr Gln Gln Gly Leu Arg Lys Leu Trp
115 120 125

Asp Ala Gly Val Gln Leu Asp Ile Met Ser Cys Asp Asp Phe Glu His
130 135 140

Cys Trp Asp Asn Phe Val Asp His Lys Gly Met Arg Phe Gln Arg Arg
145 150 155 160

Asn Leu Leu Lys Asp Tyr Asp Phe Leu Ala Ala Glu Leu Gln Glu Ile
165 170 175

Leu Arg

<210> 81
<211> 152
<212> PRT
<213> Artificial Sequence

<220>
<223> APG30125_cytidine deaminase

<400> 81

Met Ile Asn Leu Val Gln Thr Pro Tyr Asp Leu Asn Ser Gly Tyr Pro
1 5 10 15

Ile Val Arg Arg Thr Leu Glu Asp Lys Lys Lys Leu Val Lys His Glu
20 25 30

Gly Phe Gly Pro Glu Ser Cys Cys Ala Thr Ile Glu Tyr Thr Leu Arg
35 40 45

Gly Asn Ser Arg Tyr Ala Phe Gly Asn Ser Gln Met Gln Val Glu Met
50 55 60

Pro Pro Asn Ile Tyr Ala His Asn Trp Val Lys Leu His Gly Glu Met
65 70 75 80

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

Ser Ala Val Leu Pro Ile Thr Ser Ala Tyr Ile Glu Leu Arg Pro Cys
100 105 110

Glu Ala Ser Cys Leu Pro Ala Leu His Asn Met Leu Pro Asp Asn Ile
115 120 125

Thr Val Tyr Phe Ser Phe Leu His Pro Thr Gln Val Asp Gln Trp Lys
130 135 140

Gln Ser Ala Arg Ala Leu Cys Ala
145 150

<210> 82
<211> 152

<212> PRT
<213> Artificial Sequence

<220>
<223> APG30126_cytidine deaminase

<400> 82

Met Pro Val Leu Ile Lys Val Pro Tyr Asp Ile Asn Ser Ala Asn Gly
1 5 10 15

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

Asp Asn Gly Gly Ile Thr Gly Ile Gly Ala Gly Ser Cys Cys Ser Phe
35 40 45

Val Val Tyr Met Lys His Gly Gly Asp Val Asp Asn Val Phe Gly Asn
50 55 60

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

Ala Cys Ala His Gly Glu Leu Thr Ala Leu Trp Asn Ala Ile Ala Asp
85 90 95

Glu Pro Ser Ile Pro Thr Ile Leu Ala Met Tyr Ile Glu Met Ser Pro
100 105 110

Cys Thr Lys Cys Gln Ser Ala Leu Asp Asn Leu Leu Gln Pro Gly Gln
115 120 125

Glu Ile Tyr Tyr Ser Phe Asp His Pro Gly Glu Val Lys Ala Trp Gln
130 135 140

Ala Ala Ala Lys His Leu Cys Ala
145 150

<210> 83
<211> 148
<212> PRT
<213> Artificial Sequence

<220>

<223> APG30127_cytidine deaminsae

<400> 83

Met Ile Lys Thr Glu Tyr Ser Leu Ser Ser Gly Tyr Pro Ile Val Arg
1 5 10 15

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

Pro Glu Ser Cys Cys Ala Val Val Gln Tyr Arg Leu Arg Gly Asn Ile
35 40 45

Arg Tyr Ala Phe Gly Asn Ser Arg Met Gln Ile Ser Met Pro Pro Gly
50 55 60

Ile Tyr Thr His Asn Trp Val Arg Leu His Gly Glu Met Ala Ala Leu
65 70 75 80

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

Pro Ile Thr Ala Ala Tyr Ile Glu Leu Arg Pro Cys Glu Ala Asn Cys
100 105 110

Met Gln Ala Leu Arg Asn Ile Leu Pro Glu Asp Ala Cys Val Tyr Tyr
115 120 125

Ser Phe Asp His Pro Thr Gln Val Asp Glu Trp Lys Leu Arg Ala Asn
130 135 140

Glu Leu Cys Ser
145

<210> 84

<211> 147

<212> PRT

<213> Artificial Sequence

<220>

<223> APG30128_cytidine deaminase

<400> 84

Met Ile Lys Thr Glu Tyr Thr Leu Arg Ser Gly Tyr Pro Ile Val Arg
1 5 10 15

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

Pro Glu Ser Cys Cys Ala Val Val Glu Tyr Arg Leu Arg Gly Asn Ile
35 40 45

Arg Tyr Ala Phe Gly Asn Ser Arg Met Gln Val Ser Val Pro Pro Gly
50 55 60

Ile Tyr Thr Asn Asn Trp Val Arg Leu His Gly Glu Met Ala Ala Leu
65 70 75 80

Val Ala Ala Ile Glu Arg Ile Glu Arg Phe Ser Ser Asp Asp Val Ile
85 90 95

Pro Ile Thr Ala Ala Tyr Ile Glu Leu Arg Pro Cys Glu Ala Asn Cys
100 105 110

Met Gln Ala Leu His Asn Ile Leu Pro Glu Asn Ala Asn Val Tyr Tyr
115 120 125

Ser Phe Asp His Pro Thr Gln Val Glu Glu Trp Lys Leu Arg Ala His
130 135 140

Glu Leu Cys
145

<210> 85

<211> 147

<212> PRT

<213> Artificial Sequence

<220>

<223> APG30129_cytidine deaminase

<400> 85

Met Ile Lys Thr Glu Tyr Ala Leu Asn Ser Gly Tyr Pro Ile Val Arg
1 5 10 15

Arg Thr Leu Glu Asp Lys Lys Lys Leu Val Lys Gln Pro Gly Phe Gly
20 25 30

Pro Glu Ser Cys Cys Ala Val Val Glu Tyr Arg Leu Arg Gly Asn Ile
35 40 45

Arg Tyr Ala Phe Gly Asn Ser Gln Met His Val Ser Met Pro Pro Gly
50 55 60

Ile Tyr Thr His Asn Trp Val Lys Leu His Gly Glu Met Ala Ala Leu
65 70 75 80

Val Ala Ala Ile Asn Arg Ile Glu Arg Phe Ser Thr Asp Asp Val Ile
85 90 95

Pro Ile Thr Gly Ala Tyr Ile Glu Leu Arg Pro Cys Glu Ala Asn Cys
100 105 110

Leu Gln Ala Leu His Asn Ile Leu Pro Glu Asp Ala Asn Val Tyr Tyr
115 120 125

Ser Phe Asp His Pro Thr Gln Leu Asp Glu Trp Lys Leu Arg Ala His
130 135 140

Glu Leu Cys
145

<210> 86
<211> 147
<212> PRT
<213> Artificial Sequence

<220>
<223> APG30130_cytidine deaminase

<400> 86

Met Ile Gln Leu Val Gln Ala Glu Tyr Ser Ile Lys Ser Gly Tyr Pro
1 5 10 15

Ile Val Arg Arg Thr Leu Glu Asp Lys Lys Lys Leu Ile Glu Lys Pro
20 25 30

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

Gly Ser Thr Arg Tyr Ala Phe Gly Asn Ser Gln Met Lys Met Glu Met
50 55 60

Pro Pro Asp Ile Tyr Thr His Asn Trp Val Lys Leu His Ala Glu Met
65 70 75 80

Ala Ala Leu Val Ala Ala Ile Arg Arg Ile Glu Arg Phe Asp Ala Asp
85 90 95

Lys Glu Gln Val Pro Ile Thr Asn Val Tyr Ile Glu Leu Arg Pro Cys
100 105 110

Glu Ala Asn Cys Met Gln Ala Leu Gln Asn Ile Leu Pro Asp Gly Thr
115 120 125

Thr Val Tyr Tyr Ser Phe Leu His Pro Thr Glu Val Glu Glu Trp Lys
130 135 140

Arg Ser Ala
145

<210> 87
<211> 235
<212> PRT
<213> Artificial Sequence

<220>
<223> APG05241_cytidine deaminase

<400> 87

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

Arg Ile Lys Pro Trp Glu Phe Val Ala Phe Phe Asn Pro Gln Glu Leu
20 25 30

Arg Lys Glu Thr Cys Leu Leu Tyr Glu Ile Lys Trp Gly Asn Gln Asn
35 40 45

Ile Trp Arg His Ser Asn Gln Asn Thr Ser Gln His Ala Glu Ile Asn
50 55 60

Phe Met Glu Lys Phe Thr Ala Glu Arg His Phe Asn Ser Ser Val Arg
65 70 75 80

Cys Ser Ile Thr Trp Phe Leu Ser Trp Ser Pro Cys Trp Glu Cys Ser
85 90 95

Lys Ala Ile Arg Lys Phe Leu Asp His Tyr Pro Asn Val Thr Leu Ala
100 105 110

Ile Phe Ile Ser Arg Leu Tyr Trp His Met Asp Gln Gln His Arg Gln
115 120 125

Gly Leu Lys Glu Leu Val His Ser Gly Val Thr Ile Gln Ile Met Ser
130 135 140

Tyr Ser Glu Tyr His Tyr Cys Trp Arg Asn Phe Val Asp Tyr Pro Gln
145 150 155 160

Gly Glu Glu Asp Tyr Trp Pro Lys Tyr Pro Tyr Leu Trp Ile Met Leu
165 170 175

Tyr Val Leu Glu Leu His Cys Ile Ile Leu Gly Leu Pro Pro Cys Leu
180 185 190

Lys Ile Ser Gly Ser His Ser Asn Gln Leu Ala Leu Phe Ser Leu Asp
195 200 205

Leu Gln Asp Cys His Tyr Gln Lys Ile Pro Tyr Asn Val Leu Val Ala
210 215 220

Thr Gly Leu Val Gln Pro Phe Val Thr Trp Arg
225 230 235

<210> 88

<211> 262

<212> PRT

<213> Artificial Sequence

<220>

<223> APG07280_cytidine deaminase

<400> 88

Met His His Ser Ala Arg Leu Pro Pro Asn Cys Ile Val Ser Arg Tyr
1 5 10 15

Ala Asn Ala Pro Trp Thr Val Leu Pro Leu Pro Leu Pro Pro Thr Glu
20 25 30

Ala Pro Ala Thr Gly Asp Asp Thr Leu Arg Arg Arg Ile Glu Pro Trp
35 40 45

Glu Phe Glu Ala Phe Phe Asn Pro Gln Glu Leu Arg Arg Glu Ala Cys
50 55 60

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

Ala Lys Asn Thr Val Asp Ser His Val Glu Val Asn Phe Ile Gln Asn
85 90 95

Leu Thr Ala Gly Arg Tyr Cys Arg Pro Ser Thr Arg Cys Ser Ile Leu
100 105 110

Trp Phe Leu Ser Trp Ser Pro Cys Ser Ser Cys Ser Lys Ala Ile Arg
115 120 125

Leu Phe Leu Ser Gln His Pro Gly Val Ser Leu Val Ile Tyr Val Ala
130 135 140

Arg Leu Phe Gln His Met Asp Pro Gln Asn Arg Gln Gly Leu Arg Glu
145 150 155 160

Leu Ile His Ser Gly Val Thr Ile Gln Val Met Arg Pro Gln Glu Tyr
165 170 175

Asp Tyr Cys Trp Lys Asn Phe Val Asn Tyr Pro Pro Gly Gln Glu Glu
180 185 190

His Trp Pro Arg Tyr Pro Val Gln Cys Met Thr Leu Tyr Asn Leu Glu
195 200 205

Leu Tyr Cys Ile Ile His Asn Leu Pro Pro Cys Val Arg Ile Ser Lys
210 215 220

Gln Arg Gln Ser Gln Leu Ala Phe Phe Ser Leu Gly Leu Glu Asn Val
225 230 235 240

His Tyr Gln Arg Ile Pro Pro Pro Leu Leu Leu Leu Thr Gly Leu Val
245 250 255

Phe Val Phe Pro Trp Lys
260

<210> 89
<211> 211
<212> PRT
<213> Artificial Sequence

<220>
<223> APG09260_cytidine deaminase

<400> 89

Met Glu Asp Phe His Cys Val Cys Val Tyr Cys Pro Leu Gln Val Ser
1 5 10 15

Val Met Ser Asp Tyr Gln Ala Ala Ser Ser Ser Pro Asp Cys Pro Gln
20 25 30

Cys Pro Gln His Ile Leu Thr Gly Ala Glu Ala Ala Val Ser Tyr Ser
35 40 45

Gln Phe Cys Glu Thr Phe Gly Phe His Met Gly Pro Ser Gly Ala Arg
50 55 60

Ala Leu Leu Leu Phe Tyr Glu Leu Trp Gly Pro Ser Gly Thr Leu Val
65 70 75 80

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

Tyr Ser Asn Tyr Met Pro Cys Gln Glu Cys Ser Gln Thr Leu Ile Ser
100 105 110

Phe Leu Leu Arg Tyr Pro Trp Val Arg Leu Asp Leu Leu Phe Ser Gln
115 120 125

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

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

Ala Trp Gly His Leu Leu Arg Cys Phe Val Arg Asp Val Pro Pro Ser
165 170 175

Ala Leu Gln Leu Pro Leu Leu Pro Glu Arg Val Glu Ala Asp Arg Val
180 185 190

Asn Ala Ile His Ile Ser Ala Thr Thr Gly Ile Gly Pro Ala Phe Trp
195 200 205

Thr Ser His
210

<210> 90
<211> 234
<212> PRT
<213> Artificial Sequence

<220>
<223> APG08360_cytidine deaminase

<400> 90

Met Asp Pro Gln Arg Leu Arg Gln Trp Pro Gly Pro Gly Pro Ala Ser
1 5 10 15

Arg Gly Gly Tyr Gly Gln Arg Pro Arg Ile Arg Asn Pro Glu Glu Trp
20 25 30

Phe His Glu Leu Ser Pro Arg Thr Phe Ser Phe His Phe Arg Asn Leu
35 40 45

Arg Phe Ala Ser Gly Arg Asn Arg Ser Tyr Ile Cys Cys Gln Val Glu
50 55 60

Gly Lys Asn Cys Phe Phe Gln Gly Ile Phe Gln Asn Gln Lys Lys Arg
65 70 75 80

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

Gln Asn Gln Cys Tyr Arg Ile Ile Cys Tyr Val Thr Trp Ser Pro Cys
100 105 110

His Asn Cys Ala Lys Glu Leu Val Asp Phe Ile Ser Asn Arg His His
115 120 125

Leu Ser Leu Gln Leu Phe Ala Ser Arg Leu Tyr Phe His Trp Val Arg
130 135 140

Cys Tyr Gln Arg Gly Leu Gln Arg Leu Gln Ala Asn Arg Val Ser Val
145 150 155 160

Ala Val Met Lys Gly Pro Glu Phe Lys Asp Cys Trp Glu Lys Phe Val
165 170 175

Asp His Gln Gly Glu Ser Phe Pro Ser Trp Glu Lys Leu Glu Gln Tyr
180 185 190

Ser Glu Ser Ile Ser Arg Arg Leu Ser Arg Ile Leu Arg Phe Ala Asn
195 200 205

Gln Asn Asn Leu Glu Asp Ser Phe Arg Asp Leu Arg Leu Gly Ser Pro
210 215 220

Ser Pro Ser Ser Ser Arg Ser Asp Ser Arg
225 230

<210> 91
<211> 395
<212> PRT
<213> Artificial Sequence

<220>
<223> APG07386_cytidine deaminase

<400> 91

Met Asn Pro Gln Ile Arg Asn Pro Met Glu Gly Met Asp Arg His Ala
1 5 10 15

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

Thr Trp Leu Cys Tyr Glu Val Lys Ile Arg Lys Asp Pro Ser Lys Leu
35 40 45

Pro Trp Asp Thr Gly Val Phe Arg Gly Gln Val Arg Pro Lys Leu Gln
50 55 60

Ser Asn Arg Arg Tyr Glu Leu Ser Asn Trp Glu Cys Arg Lys His Val
65 70 75 80

Tyr Phe Gln Pro Gln Tyr His Ala Glu Met Cys Phe Leu Ser Trp Phe
85 90 95

Cys Gly Asn Gln Leu Pro Ala His Lys Arg Phe Gln Ile Thr Trp Phe
100 105 110

Val Ser Trp Thr Pro Cys Pro Asp Cys Val Ala Lys Val Thr Glu Phe
115 120 125

Leu Ala Glu His Pro Asn Val Thr Leu Thr Ile Ser Val Ala Arg Leu
130 135 140

Tyr Tyr Tyr Arg Gly Lys Asp Trp Arg Arg Ala Leu Cys Arg Leu His
145 150 155 160

Gln Ala Gly Ala Arg Val Lys Ile Met Asp Tyr Glu Glu Phe Ala Tyr
165 170 175

Cys Trp Glu Asn Phe Val Tyr Asn Glu Gly Gln Ser Phe Met Pro Trp
180 185 190

Asp Lys Phe Asp Asp Asn Tyr Ala Phe Leu His His Lys Leu Lys Glu
195 200 205

Ile Leu Arg Asn Pro Met Glu Ala Thr Tyr Pro His Ile Phe Tyr Phe
210 215 220

His Phe Lys Asn Leu Arg Lys Ala Tyr Gly Arg Asn Glu Thr Trp Leu
225 230 235 240

Cys Phe Thr Met Glu Ile Ile Lys Gln His Ser Thr Val Phe Trp Glu
245 250 255

Thr Gly Val Phe Arg Asn Gln Val Tyr Pro Glu Ser Leu Cys His Ala
260 265 270

Glu Arg Cys Phe Leu Ser Trp Phe Cys Glu Asp Ile Leu Ser Pro Asn
275 280 285

Thr Asp Tyr Arg Val Thr Trp Tyr Thr Ser Trp Ser Pro Cys Leu Asp
290 295 300

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

Leu Ala Ile Phe Ala Ala Arg Leu Tyr Tyr Phe Trp Asp Pro His Tyr
325 330 335

Gln Gln Gly Leu Arg Ser Leu Ser Glu Lys Gly Ala Ser Val Glu Ile
340 345 350

Met Gly Tyr Lys Asp Phe Lys Tyr Cys Trp Glu Asn Phe Val Tyr Asn
355 360 365

Gly Asp Glu Pro Phe Lys Pro Trp Lys Gly Leu Lys Tyr Asn Phe Leu
370 375 380

Phe Leu Asp Ser Lys Leu Gln Glu Ile Leu Gln
385 390 395

<210> 92

<211> 213

<212> PRT

<213> Artificial Sequence

<220>

<223> APG07386-NTD_cytidine deaminase

<400> 92

Met Asn Pro Gln Ile Arg Asn Pro Met Glu Gly Met Asp Arg His Ala
1 5 10 15

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

Thr Trp Leu Cys Tyr Glu Val Lys Ile Arg Lys Asp Pro Ser Lys Leu
35 40 45

Pro Trp Asp Thr Gly Val Phe Arg Gly Gln Val Arg Pro Lys Leu Gln
50 55 60

Ser Asn Arg Arg Tyr Glu Leu Ser Asn Trp Glu Cys Arg Lys His Val
65 70 75 80

Tyr Phe Gln Pro Gln Tyr His Ala Glu Met Cys Phe Leu Ser Trp Phe
85 90 95

Cys Gly Asn Gln Leu Pro Ala His Lys Arg Phe Gln Ile Thr Trp Phe
100 105 110

Val Ser Trp Thr Pro Cys Pro Asp Cys Val Ala Lys Val Thr Glu Phe
115 120 125

Leu Ala Glu His Pro Asn Val Thr Leu Thr Ile Ser Val Ala Arg Leu
130 135 140

Tyr Tyr Tyr Arg Gly Lys Asp Trp Arg Arg Ala Leu Cys Arg Leu His
145 150 155 160

Gln Ala Gly Ala Arg Val Lys Ile Met Asp Tyr Glu Glu Phe Ala Tyr
165 170 175

Cys Trp Glu Asn Phe Val Tyr Asn Glu Gly Gln Ser Phe Met Pro Trp
180 185 190

Asp Lys Phe Asp Asp Asn Tyr Ala Phe Leu His His Lys Leu Lys Glu
195 200 205

Ile Leu Arg Asn Pro
210

<210> 93
<211> 267
<212> PRT
<213> Artificial Sequence

<220>
<223> APG09688_cytidine deaminase

<400> 93

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

His Glu Leu His Asn Ser Arg Gly Arg Ser Pro Glu Ser Asp Thr Met
20 25 30

Ala Ser Glu Lys Gly Pro Ser Asn Lys Asp Tyr Thr Leu Arg Arg Arg
35 40 45

Ile Glu Pro Trp Glu Phe Glu Val Phe Phe Asp Pro Gln Glu Leu Arg
50 55 60

Lys Glu Ala Cys Leu Leu Tyr Glu Ile Lys Trp Gly Ala Ser Ser Lys
65 70 75 80

Thr Trp Arg Ser Ser Gly Lys Asn Thr Thr Asn His Val Glu Val Asn
85 90 95

Phe Leu Glu Lys Leu Thr Ser Glu Gly Arg Leu Gly Pro Ser Thr Cys
100 105 110

Cys Ser Ile Thr Trp Phe Leu Ser Trp Ser Pro Cys Trp Glu Cys Ser
115 120 125

Thr Ala Ile Arg Glu Phe Leu Ser Gln His Pro Gly Val Thr Leu Val
130 135 140

Ile Phe Val Ala Arg Leu Phe Gln His Met Asp Arg Arg Asn Arg Gln
145 150 155 160

Gly Leu Lys Asp Leu Val Thr Ser Gly Val Thr Val Gln Val Met Ser
165 170 175

Val Ser Glu Tyr Cys Tyr Cys Trp Glu Asn Phe Val Asn Tyr Pro Pro
180 185 190

Gly Lys Ala Ala Gln Trp Pro Arg Tyr Pro Pro Arg Trp Met Leu Met
195 200 205

Tyr Ala Leu Glu Leu Tyr Cys Ile Ile Leu Gly Leu Pro Pro Cys Leu
210 215 220

Lys Ile Ser Arg Arg His Gln Lys Gln Leu Thr Phe Phe Ser Leu Thr
225 230 235 240

Pro Gln Tyr Cys His Tyr Lys Met Ile Pro Pro Tyr Ile Leu Leu Ala
245 250 255

Thr Gly Leu Leu Gln Pro Ser Val Pro Trp Arg
260 265

<210> 94
<211> 373
<212> PRT
<213> Artificial Sequence

<220>
<223> APG02316_cytidine deaminase

<400> 94

Met Ile Glu Asn Pro Ala Leu Lys Ile Ser Lys Arg Asp Arg Ile Lys
1 5 10 15

Ile Ser Asn Leu Lys Asp Leu Lys Asn Ala Leu Lys Asn Glu Asn Tyr
20 25 30

Asn Ile Lys Ala Asn Asp Lys Glu Lys Phe Lys Asp Glu Val Lys Lys
35 40 45

Val Phe Asn Ile Asn Asp Asp Ile Phe Glu Arg Leu Asp Lys Cys Leu
50 55 60

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

Tyr Ile Lys Lys Ile Met Ile Phe Glu Asp Lys His Glu Ile Ile Cys
85 90 95

Glu Lys Leu Lys Arg Ile Lys Lys Leu Tyr Ile Asn Arg Glu Glu Tyr
100 105 110

Glu Arg Glu Lys Ser Thr Arg Asp Asn Val Glu His Ile Ile Glu Val
115 120 125

Ile Glu Lys Thr Lys Glu Asn Val Ser Arg Lys Ile Ser Leu Glu Glu
130 135 140

Leu Glu Arg Leu Glu Ile Leu Glu Glu Glu Leu Glu Asp Lys Tyr Leu
145 150 155 160

Phe Ala Lys Asp Ile Glu Phe Leu Lys Lys Met Ile Leu Gly Asn Cys
165 170 175

Lys Asn Val Ile Glu Ser Tyr Asn Glu Lys Thr Lys Ile Lys Thr Leu
180 185 190

Lys Met Lys Ile Pro Lys Glu Ile Asn Tyr Ser Tyr Ile Lys Ala Lys
195 200 205

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

Met Asn Arg Leu Ile Lys Ser Ile Asp Lys Tyr Ile Glu His Tyr Lys
225 230 235 240

Asp Asp Ile Phe Asn Ile Asn Gln Ser Leu Ala Leu Gln Asp Ser Ile
245 250 255

Asn Ile Ala Leu Ala Thr Phe Asp Asn Lys Glu Phe Lys Ala Ile Ser
260 265 270

Gly Lys Asn Asp Ile Glu Asp Tyr Cys Lys Val Ile Pro Ile Glu Lys
275 280 285

Ser Arg Phe Lys Ser Arg Lys Val Asn Lys Leu Gly Glu Leu Gly Ile
290 295 300

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

His Glu Lys Ile Lys Gln Lys Ile Leu Lys Asp Arg Gly Asn Leu Thr
325 330 335

Leu Tyr Thr Lys Trp Glu Pro Cys Pro Ser Cys Tyr Phe Val Ile Ser
340 345 350

Gln Phe Cys Glu Lys Tyr Pro Asn Ile Lys Val Glu Val Lys Tyr Asn
355 360 365

Lys Lys Tyr Gly Glu
370

<210> 95
<211> 1060
<212> PRT
<213> Bacillus sp.

<223> APG00969_RGN

<400> 95

Met Lys Gln Thr Ser Tyr Ser Leu Gly Leu Asp Ile Gly Ile Ala Ser
1 5 10 15

Val Gly Tyr Gly Leu Ile Asp Asn Asp Glu Asn Ile Val Asp Ala Gly
20 25 30

Val Arg Leu Phe Pro Glu Ala Asn Ser Lys Asn Asn Asp Gly Arg Arg
35 40 45

Lys Ser Arg Gly Ser Lys Arg Leu Ile Arg Arg Lys Arg His Arg Ile
50 55 60

Ala Arg Val Lys His Leu Leu Lys Glu Ser Gly Ile Asp Val Ser Tyr

65					70						75					80
Glu	Asn	Thr	Val	Leu	Thr	Asn	Pro	Tyr	Glu	Ile	Arg	Cys	Lys	Gly	Leu	
				85					90					95		
Thr	Leu	Pro	Leu	Thr	Asn	His	Glu	Leu	Ser	Ile	Ala	Leu	Leu	His	Leu	
			100					105					110			
Ala	Lys	Arg	Arg	Gly	Val	His	Asn	Val	Lys	Ser	Met	Asp	Gln	Glu	Lys	
		115					120					125				
Val	Lys	Gly	Asn	Glu	Leu	Ser	Thr	Lys	Glu	Gln	Leu	Thr	Ile	Asn	Asp	
	130					135					140					
Asn	Leu	Leu	Glu	Glu	Lys	Phe	Ile	Cys	Glu	Leu	Gln	Leu	Glu	Arg	Leu	
145					150					155					160	
Asn	Lys	Glu	Gly	Ile	Val	Arg	Ser	His	Ser	Asn	Arg	Phe	Lys	Thr	Ala	
				165					170					175		
Asp	Tyr	Ile	Lys	Glu	Ile	Lys	Asn	Leu	Leu	Glu	Thr	Gln	Ala	Lys	Gln	
			180					185					190			
Asn	Thr	Leu	Val	Thr	Glu	Glu	Phe	Ile	Glu	Lys	Tyr	Ile	Glu	Ile	Phe	
		195					200					205				
Ser	Gly	Arg	Arg	Lys	Tyr	Tyr	Glu	Gly	Pro	Gly	Gly	Glu	Ser	Lys	Tyr	
	210					215					220					
Gly	Trp	Lys	Gly	Asp	Ile	Glu	Lys	Trp	Tyr	Glu	Gly	Leu	Met	Gly	Lys	
225					230					235					240	
Cys	Thr	Tyr	Phe	Pro	Lys	Glu	Leu	Arg	Cys	Val	Lys	His	Ala	Tyr	Ser	
				245					250					255		
Ala	Ala	Leu	Phe	Asn	Leu	Leu	Asn	Asp	Leu	Asn	Asn	Leu	Ser	Ile	Asn	
			260					265					270			
Arg	Glu	Glu	Asp	Thr	Lys	Leu	Ser	Gln	Tyr	Glu	Lys	Glu	Gln	Ile	Ile	
	275						280					285				

Glu Lys Ile Phe Lys Val Arg Lys Thr Pro Thr Leu Thr Gln Ile Ala
290 295 300

Lys Leu Leu Lys Val Asp Pro Thr Asn Ile Lys Gly Phe Arg Thr Lys
305 310 315 320

Ala Asn Gly Thr Pro Asp Phe Leu Ser Ile Lys Ile Tyr His Asp Leu
325 330 335

Lys Gly Ile Ile Asp Asp Lys Gln Leu Leu Asp Asp Ile Ala Phe Leu
340 345 350

Asp Asn Val Ala Gln Ile Leu Thr Val Trp Gln Asp Ser Gln Ser Ile
355 360 365

Gln Glu Lys Leu Lys Thr Leu Asn Lys Asn Leu Asp Asp Lys Thr Ile
370 375 380

Lys Glu Ile Ser Glu Leu Lys Lys Tyr Thr Gln Thr His Ser Leu Ser
385 390 395 400

Leu Lys Leu Ile Asn Val Leu Leu Pro Glu Leu Trp Glu Thr Thr Lys
405 410 415

Asn Gln Met Thr Ile Leu Ser Glu Leu Lys Leu Lys Pro Arg Lys Ile
420 425 430

Asp Leu His Asn Cys Asn Glu Ile Pro Val Asn Met Ile Asn Asp Leu
435 440 445

Ile Val Ser Pro Val Val Arg Arg Ser Leu Thr Gln Ser Ile Glu Met
450 455 460

Ile Asn Gln Ile Ile Lys Asp Tyr Gly His Pro Arg Glu Ile Val Ile
465 470 475 480

Glu Leu Ala Arg Glu Lys Asn Ser Glu Glu Lys Lys Asn Phe Ile Lys
485 490 495

Ser Leu Asn Glu Lys Asn Lys Gln Ile Asn Asp Glu Val Ile Glu Lys

500

505

510

Leu Asn Ala Ser Asn His Arg Asp Asn Lys Gly Met Phe Asn Lys Val
 515 520 525

Lys Leu Trp Ile Leu Gln Asp Gly His Cys Leu Tyr Ser Leu Lys Pro
 530 535 540

Ile Arg Leu Glu Asp Leu Leu Asn Asn Pro Asn His Tyr Glu Ile Asp
 545 550 555 560

His Ile Ile Pro Lys Ser Val Ser Phe Asp Asp Ser Met Ser Asn Lys
 565 570 575

Val Leu Val Tyr Gln Ile Glu Asn Ser Lys Lys Gly Asn Arg Thr Pro
 580 585 590

Tyr Gln Tyr Leu Thr Ser Ala Asp Lys Thr Ile Thr Tyr Glu Lys Phe
 595 600 605

Lys Ala Asn Ile Thr Gln Leu Ala Lys Ser Asn His Lys Ile Ser Lys
 610 615 620

Lys Lys Leu Asp Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe His
 625 630 635 640

Ile Lys Lys Glu Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala
 645 650 655

Thr Arg Ser Leu Ile Asn Leu Leu Lys Tyr Tyr Phe Ser Glu Lys Asp
 660 665 670

Ile Asn Val Lys Val Lys Ser Ile Asn Gly Ser Phe Thr Asp Tyr Leu
 675 680 685

Arg Lys Leu Trp Asn Phe Pro Lys Asp Arg Glu Phe Tyr His Lys His
 690 695 700

His Ala Glu Asp Ala Leu Ile Ile Ala Met Ala Asn Lys Ile Phe Thr
 705 710 715 720

Thr Arg Lys Ile Phe Lys Glu Gln Asn Ser Val Phe Ser Asp Glu Gln
725 730 735

Ile Leu Asp Gly Glu Val Thr Asn Ile Leu Ser Asp Asp Gln Phe Gln
740 745 750

Ala Glu Phe Thr Glu Lys Phe Tyr Lys Val Gln Ala Ile Lys Lys Tyr
755 760 765

Asp Lys Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Gln
770 775 780

Leu Phe Asp Asp Thr Leu Tyr Ser Thr Arg Glu Phe Glu Gly Glu Glu
785 790 795 800

Tyr Tyr Ile Gly Lys Ile Lys Asp Ile Tyr Asn Leu Lys Asp Lys Arg
805 810 815

Leu Lys Lys Ile Phe Thr Lys Ser Pro Glu Lys Ile Leu Met Tyr Gln
820 825 830

His Asp Ser Gln Thr Phe Lys Lys Leu Lys Gln Ile Met Arg Ser Tyr
835 840 845

Glu Asp Glu Val Asn Pro Leu Ala Lys Tyr His Lys Glu Thr Gly Glu
850 855 860

Tyr Leu Arg Lys Glu Cys Lys Lys Gly Asn Gly Pro Ile Val Lys Ser
865 870 875 880

Leu Lys Tyr Arg Val Thr Lys Leu Gly Val His Lys Asp Ile Thr His
885 890 895

Lys Tyr Glu Asn Ser Lys Asn Lys Val Val Ile Leu Ser Leu Lys Pro
900 905 910

Phe Arg Met Asp Val Phe Lys Glu Asn Gly Val Tyr Lys Phe Ile Thr
915 920 925

Ile Arg Tyr Cys Asp Leu Lys Glu Thr Val Asn Ser Tyr Thr Ile Ser

930

935

940

Glu His Leu Tyr Lys Ala Lys Leu Lys Ala Lys Asp Ile Lys Ser Met
 945 950 955 960

Asp Ser Phe Lys Trp Ser Phe Tyr Lys Asn Asp Leu Leu Glu Tyr Asn
 965 970 975

Gly Glu Leu Cys Thr Phe Lys Gly Val Asn Asp Asp Lys Lys Asn Lys
 980 985 990

Ile Glu Val Asn Trp Val Glu Lys Asn Phe Ala Ile Tyr Ala Glu Lys
 995 1000 1005

Lys Asn Leu Lys Ser Lys Gln Leu Val Lys Ser Ile Thr Lys Ser
 1010 1015 1020

Thr Val Lys Ser Leu Leu Lys Tyr Thr Thr Asp Ile Leu Gly Asn
 1025 1030 1035

Arg Tyr Pro Val Arg Asn Glu Lys Leu Lys Leu Met Ile Arg Lys
 1040 1045 1050

Gln Thr Phe Arg Gly Asp Leu
 1055 1060

<210> 96
 <211> 1090
 <212> PRT
 <213> Rhizobium sp.

<223> APG03128_RGN

<400> 96

Met Asn Asn Leu Leu Tyr Ser Phe Asp Ile Gly Thr Asn Ser Ile Gly
 1 5 10 15

Trp Cys Val Phe Ala Leu Asp Glu Val Gly Asp Pro Cys Arg Ile Val
 20 25 30

Asp Leu Gly Ala Arg Ile Tyr Ala Asp Gly Arg Asp Pro Gln Thr Lys
 35 40 45

Thr Ser Leu Ala Val Ala Arg Arg Glu Ala Arg Ala Met Ser Arg Arg
50 55 60

Arg Asp Arg Ser Leu Arg Arg Arg Lys Ala Thr Leu Arg Thr Met Ile
65 70 75 80

Glu Tyr Gly Leu Met Pro Ala His Lys Val Glu Gln Glu Thr Leu Leu
85 90 95

Arg Lys Thr Gly Asp Arg Glu Gly Gly Asp Glu Gly Phe Asn Pro Tyr
100 105 110

Ala Leu Arg Ala Arg Ala Leu Gly Glu Lys Leu Pro Pro Phe Tyr Ile
115 120 125

Gly Arg Ala Leu Phe His Leu Gly Gln Arg Arg Gly Phe Lys Ser Asn
130 135 140

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

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

Gly Ala Trp Leu Ala Met Arg Arg Ala Asp Gly His Pro Val Arg Leu
180 185 190

Arg Ala Gly Ser Glu Val Phe Asp Ala Glu Gly Tyr Ala Phe Tyr Pro
195 200 205

Glu Arg Ser Leu Leu Glu Asp Glu Phe Arg Gln Ile Trp Thr Ala Gln
210 215 220

Ala Val His His Pro Gln Leu Leu Thr Ser Glu Arg Arg Ala His Leu
225 230 235 240

Phe Arg Val Met Phe Tyr Gln Arg Pro Leu Lys Lys Pro Leu Val Gly
245 250 255

Arg Cys Ser Phe Asn Pro Ala Glu Ala Arg Leu Ser Arg Ala His Pro
260 265 270

Leu Phe Gln Glu Phe Arg Leu Tyr Lys Glu Val Asn Asp Leu Glu Val
275 280 285

Val Leu Pro Asp Gln Ser His Arg Lys Leu Thr Leu Asp Glu Arg Asn
290 295 300

Ala Leu Val Ala Lys Leu Lys Ser Ser Arg Lys Ala Ser Phe Ser Val
305 310 315 320

Leu Arg Arg Thr Leu Lys Leu Thr Pro Asp Leu Ala Phe Asn Lys Glu
325 330 335

Ser Glu Ala Arg Lys Asp Leu Leu Gly Asp Glu Ile Asn Ser Ala Leu
340 345 350

Ala Asp Ala Lys Met Phe Gly Val Arg Trp Gly Gly Phe Pro Arg Ala
355 360 365

Arg Gln Trp Glu Ile Ile Thr His Leu Lys Glu Glu Glu Asn Pro Ala
370 375 380

Arg Leu Ser Asp Trp Leu Lys Ser Glu Phe Gly Leu Asp Asp Glu Arg
385 390 395 400

Val Val Ala Ile Ala Asn Ile Ala Leu Pro Glu Gly Tyr Gly Arg Leu
405 410 415

Gly Glu Thr Ala Leu Ala Ser Met Leu Glu Glu Met Lys Thr Ala Val
420 425 430

Ile Pro Glu Ser Glu Ala Ala Lys Arg Cys Gly Tyr Asp His Ser Asn
435 440 445

Leu Ala Lys Glu Gln Asp Glu Gly Leu Asp Ile Leu Pro Glu Tyr Gln
450 455 460

Glu Ile Leu Glu Arg His Ile Pro Pro Gly Thr Gly Asp Pro Asp Asp
465 470 475 480

Ile Tyr Asp Ile Arg Lys Gly Arg Ile Thr Asn Pro Thr Val His Ile
485 490 495

Gly Leu Asn Gln Leu Arg Arg Val Val Asn Ala Leu Ile Lys Arg His
500 505 510

Gly Lys Pro Arg His Ile Val Val Glu Leu Ala Arg Asp Leu Gln Leu
515 520 525

Ser Glu Lys Gln Lys Ala Asp Val Asn Arg Arg Ile Ala Lys Asn Thr
530 535 540

Arg Glu Ala Glu Gly Arg Ser Gln Lys Leu Ile Glu Met Gly Gln Leu
545 550 555 560

Asp Thr Gly Tyr Asn Arg Leu Leu Leu Lys Leu Trp Glu Glu Leu Asn
565 570 575

Gln Asp Lys Pro Glu Asp Arg Val Cys Ile Tyr Ser Gly Lys Pro Ile
580 585 590

Gly Ile Ala Met Leu Phe Ser Gly Glu Val Asp Ile Asp His Ile Leu
595 600 605

Pro Trp Ser Lys Thr Leu Asp Asp Ser Gln Ala Asn Lys Leu Leu Cys
610 615 620

Leu Lys Ser Ala Asn Arg Gln Lys Arg Asn Arg Thr Pro Ser Asp Val
625 630 635 640

Pro Glu Trp Arg Asp Cys Tyr Glu Glu Val Leu Ala Arg Ala Ala Arg
645 650 655

Leu Pro Lys Asn Lys Arg Trp Arg Phe Ala Ala Asp Ala Met Gln Gln
660 665 670

Phe Glu Ala Glu Gly Gly Phe Leu Ala Arg Gln Leu Thr Asp Thr Gln
675 680 685

Tyr Leu Ser Arg Met Ala Phe Glu Tyr Leu Ser Ala Leu Phe Pro Ser
690 695 700

Glu Glu Ala Asp Lys Trp Gly Glu Leu Arg Gln Arg Lys Arg Val His
705 710 715 720

Val Val Pro Gly Arg Leu Thr Glu Leu Leu Arg Arg Asn Trp Gly Leu
725 730 735

Asn Thr Leu Leu Pro Asp His Asn Leu Gly Glu Met Ala Gln Glu Lys
740 745 750

Asn Arg Lys Asp His Arg His His Ala Ile Asp Ala Ala Val Ile Gly
755 760 765

Val Thr Ser Arg Ser Leu Leu Gln Arg Met Ser Gly Ala Ala Ala Arg
770 775 780

Leu Asp Asp Val Ala Phe Asp Asp Leu Val Arg Thr Val Val Lys Asp
785 790 795 800

Asn Pro Pro Trp Pro Gly Phe Arg Glu Glu Leu Leu Gly Cys Ile Asn
805 810 815

Arg Val Thr Val Ser His Lys Pro Asp His Gly Thr Val Ser Arg Ala
820 825 830

Ala Tyr Ala Gln Gly Lys Gly Gln Thr Ala Gly Lys Leu His Asn Asp
835 840 845

Thr Ala Tyr Gly Ile Thr Gly Leu Gln Asp Met Lys Gly Ser Pro Leu
850 855 860

Val Val Arg Arg Lys Pro Phe Met Ala Leu Glu Val Lys Asp Ile Ala
865 870 875 880

Ser Ile Arg Asp Thr Glu Leu Gln Ser Ala Leu Tyr Asp Ala Ile Gly
885 890 895

Ser Leu Thr Glu Lys Lys Ala Leu Gln Glu Ala Leu Val Arg Phe Arg
900 905 910

Asp Arg His Pro Gln Phe Lys Gly Ile Arg Arg Val Arg Val Leu Glu
915 920 925

Ala Leu Ser Val Ile Pro Ile Val Asp Asn Asn Gly Lys Ala Tyr Lys
930 935 940

Gly Tyr Lys Gly Asp Ala Asn Tyr Arg Tyr Glu Val Trp Glu Thr Leu
945 950 955 960

Asp Gly Arg Trp His Thr Glu Val Val Ser Met Phe Asp Ala His Gln
965 970 975

Pro Gly Trp Gln Ser Pro Phe His Arg Gln His Pro Ala Ala Arg Arg
980 985 990

Val Leu Lys Leu Gln Gln Asn Asp Met Val Ala Tyr Glu His Pro Ala
995 1000 1005

Asp Gly Tyr Thr Ile Ala Arg Val Val Lys Phe Ser Asn Asp Lys
1010 1015 1020

Arg Ile Tyr Phe Ala Ser His Arg Glu Ser Gly Ser Leu Lys Ala
1025 1030 1035

Arg Asp Ala Asp Lys Gly Asp Pro Phe Thr Tyr Phe Ala Lys Ala
1040 1045 1050

Asn Asn Gly Leu Arg Asp Ile Lys Cys Arg Gln Val Arg Ile Asp
1055 1060 1065

Ala Ala Gly Arg Val Phe Asp Pro Gly Pro Gln Asp Arg Glu Ala
1070 1075 1080

Arg Gln Ser Lys Gly Arg Ala
1085 1090

<210> 97
<211> 1108
<212> PRT
<213> Brevibacillus sp.

<223> APG09748_RGN

<400> 97

Met Ala Ile Arg Ser Ile Lys Leu Lys Leu Lys Thr Arg Thr Gly Pro
1 5 10 15

Glu Ala Gln Asn Leu Arg Lys Gly Ile Trp Arg Thr His Arg Leu Leu
20 25 30

Asn Glu Gly Val Ala Tyr Tyr Met Lys Met Leu Leu Leu Phe Arg Gln
35 40 45

Glu Ser Thr Gly Gly Gln Thr Lys Lys Glu Leu Gln Glu Glu Leu Val
50 55 60

Arg His Ile Arg Glu Gln Gln Gln Lys Asn Arg Ala Asp Lys Asn Thr
65 70 75 80

Gln Ala Leu Pro Leu Asp Lys Ala Phe Ala Ala Leu Arg Gln Leu Tyr
85 90 95

Glu Leu Leu Val Pro Ser Ser Ile Gly Gln Ser Gly Asp Ala Gln Ile
100 105 110

Ile Ser Arg Lys Phe Leu Ser Pro Leu Val Asp Pro Asn Ser Glu Gly
115 120 125

Gly Lys Gly Thr Ser Lys Ala Gly Ala Lys Pro Thr Trp Gln Lys Lys
130 135 140

Lys Glu Ala Asn Asp Pro Thr Trp Glu Gln Asp Tyr Glu Lys Trp Lys
145 150 155 160

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

Glu Tyr Gly Ile Arg Pro Ile Phe Pro Leu Tyr Thr Asn Thr Val Ala
180 185 190

Asp Ile Ala Trp Leu Pro Leu Gln Ser Asn Gln Phe Val Arg Thr Trp

195

200

205

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

Ser Trp Asn Lys Arg Val Gln Glu Glu Tyr Ser Lys Leu Gln Glu Lys
 225 230 235 240

Met Thr Gln Leu Asn Glu Gln Leu Glu Gly Gly Gln Glu Trp Ile Ser
 245 250 255

Leu Leu Glu Gln Tyr Glu Glu Gln Arg Glu Gln Glu Leu Ile Glu Asn
 260 265 270

Met Thr Ala Ala Asn Asp Lys Tyr Arg Ile Thr Lys Arg Gln Met Lys
 275 280 285

Gly Trp Asn Glu Leu Tyr Glu Gln Trp Ser Thr Val Leu Pro Asn Ala
 290 295 300

Ser His Glu Gln Tyr Arg Glu Ala Leu Lys Arg Val Gln Gln Arg Leu
 305 310 315 320

Arg Gly Arg Phe Gly Asp Ala His Phe Phe Gln Tyr Leu Met Lys Glu
 325 330 335

Glu His His Leu Ile Trp Lys Gly Asn Pro Gln Arg Ile His Tyr Phe
 340 345 350

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

Ala Thr Met Thr Leu Pro Asp Ala Arg Lys His Pro Leu Trp Val Arg
 370 375 380

Phe Asp Ala Arg Gly Gly Asn Leu Gln Asp Tyr Tyr Leu Thr Ala Glu
 385 390 395 400

Ala Asp Asn Pro Arg Ser Arg Arg Phe Val Thr Phe Ser Gln Leu Ile
 405 410 415

Trp Pro Asn Glu Ser Gly Trp Met Glu Lys Gln Asp Val Glu Val Glu
420 425 430

Leu Ala Leu Ser Lys Gln Phe Tyr Gln Gln Val Thr Leu Gln Lys Asn
435 440 445

Asp Lys Gly Lys Gln Glu Ile Glu Phe Lys Asp Lys Gly Ser Gly Ser
450 455 460

Thr Phe Ser Gly His Leu Gly Gly Ala Lys Leu Gln Leu Glu Arg Gly
465 470 475 480

Asp Leu Glu Lys Glu Glu Lys Asp Phe Glu Gly Gly Glu Ile Gly Ser
485 490 495

Val Tyr Leu Asn Ile Val Ile Asp Phe Glu Pro Leu Gln Glu Val Lys
500 505 510

Asn Gly Arg Leu Gln Ser Pro Tyr Gly Gln Val Leu Gln Leu Val Arg
515 520 525

Arg Pro Asn Glu Phe Pro Lys Val Thr Thr Tyr Lys Ser Glu Glu Leu
530 535 540

Val Glu Trp Ile Lys Ser Ser Thr Lys Asp Ser Ala Gly Val Glu Ser
545 550 555 560

Leu Glu Ser Gly Phe Arg Val Met Ser Ile Asp Leu Gly Leu Arg Thr
565 570 575

Ala Ala Ala Thr Ser Ile Phe Ser Val Glu Glu Ser Asn Asp Ala Asn
580 585 590

Ala Ala Gly Phe Ser Tyr Trp Ile Glu Gly Thr Pro Leu Val Ala Val
595 600 605

His Lys Arg Ser Tyr Met Leu Lys Leu Pro Gly Glu Gln Val Glu Lys
610 615 620

Gln Val Arg Glu Lys Arg Asp Glu Arg Gln Asp Gln Gln Arg Arg Val

625		630		635		640
Arg Phe Gln Ile	Arg Ile Leu Ser	Gln Val Ile Arg	Met Ala Lys Lys			
	645		650			655
Gln Asn Arg Glu	Arg Ala Asp Glu	Leu Asp His Leu	Ser Gln Ala Leu			
	660		665			670
Glu Lys Gln Lys	Ser Leu Leu Asp	Gln Thr Asp Arg	Thr Phe Trp Asn			
	675		680			685
Gly Ile Val Cys	Asp Leu Thr Asp	Ala Leu Arg Glu	Lys Glu Gly Gly			
	690		695			700
Trp Glu Gln Ala	Val Val Gln Ile	His Arg Lys Ala	Glu Glu His Val			
	705		710			715
Gly Lys Val Val	Gln Ala Trp Arg	Lys Arg Phe Asp	Ala Asp Glu Arg			
	725		730			735
Lys Gly Ile Ala	Gly Leu Ser Met	Trp Ser Ile Glu	Glu Leu Asp Ser			
	740		745			750
Leu Arg Lys Leu	Leu Ile Ser Trp	Ser Arg Arg Thr	Arg Asn Pro Arg			
	755		760			765
Glu Ile Asn Cys	Phe Glu Gln Gly	His Thr Ser His	Gln Arg Leu Leu			
	770		775			780
Thr His Ile Gln	Asn Val Lys Glu	Asp Arg Leu Lys	Gln Leu Ser His			
	785		790			795
Ala Ile Val Met	Thr Ala Leu Gly	Tyr Val Tyr Asp	Glu Lys Lys Leu			
	805		810			815
Glu Trp Phe Ala	Lys Tyr Pro Ala	Cys Gln Val Ile	Leu Phe Glu Asn			
	820		825			830
Leu Ser Gln Tyr	Arg Ser Asn Met	Asp Arg Ser Thr	Lys Glu Asn Ser			
	835		840			845

Thr Leu Met Lys Trp Ala His Arg Ser Ile Pro Lys Tyr Val His Met
850 855 860

Gln Ala Glu Pro Tyr Gly Ile Gln Ile Gly Asp Val Arg Ala Glu Tyr
865 870 875 880

Ser Ser Arg Phe His Ala Lys Thr Gly Thr Pro Gly Ile Arg Cys Lys
885 890 895

Met Val Ser Gly His Asp Leu Gln Gly Arg Arg Phe Glu Asn Leu Gln
900 905 910

Lys Arg Leu Ile Ser Glu Gln Phe Leu Thr Glu Glu Gln Val Lys Gln
915 920 925

Leu Arg Pro Gly Asp Ile Val Pro Asp Asp Ser Gly Glu Trp Phe Met
930 935 940

Thr Leu Ser Asp Gly Ser Glu Gly Lys Glu Val Val Phe Leu Gln Ala
945 950 955 960

Asp Ile Asn Ala Ala Gln Asn Leu Gln Lys Arg Phe Trp Gln Arg Tyr
965 970 975

Asn Glu Leu Phe Lys Val Ser Cys Arg Val Leu Ile Arg Gly Glu Glu
980 985 990

Glu Tyr Leu Ile Pro Lys Thr Lys Ser Val Gln Ala Lys Leu Gly Lys
995 1000 1005

Gly Leu Phe Val Lys Lys Thr Asp Thr Val Met Lys Asp Val Tyr
1010 1015 1020

Val Trp Asp Ser Gln Ala Lys Leu Lys Gly Lys Thr Thr Phe Thr
1025 1030 1035

Glu Glu Ser Glu Ser Pro Glu Gln Leu Glu Asp Phe Gln Glu Ile
1040 1045 1050

Ile Glu Glu Ala Glu Glu Ala Lys Gly Thr Tyr Arg Thr Leu Phe

1055

1060

1065

Arg Asp Pro Ser Gly Val Phe Phe Pro Glu Phe Val Trp Ser Thr
 1070 1075 1080

Gln Lys Asp Phe Trp Ser Glu Val Lys Arg Arg Leu Tyr Gly Lys
 1085 1090 1095

Leu Arg Glu Arg Phe Leu Met Lys Thr Arg
 1100 1105

<210> 98

<211> 1429

<212> PRT

<213> Chryseobacterium sp.

<223> APG00771_RGN

<400> 98

Met Thr Lys Asn Ile Leu Gly Leu Asp Leu Gly Thr Asn Ser Ile Gly
 1 5 10 15

Trp Ala Leu Ile Lys Gln Asp Phe Glu Asn Lys Gln Gly Glu Ile Leu
 20 25 30

Gly Met Gly Ser Arg Ile Ile Pro Met Asp Ala Gly Asp Met Gly Lys
 35 40 45

Phe Ala Glu Gly Ala Thr Val Ser Lys Thr Ala Asp Arg Thr Lys Tyr
 50 55 60

Arg Ser Val Arg Arg Leu Arg Glu Arg Phe Leu Leu Arg Arg Glu Arg
 65 70 75 80

Leu His Arg Val Leu His Leu Leu Asn Phe Leu Pro Gln His Tyr Ala
 85 90 95

Ser Gln Ile Asp Phe Glu Lys Arg Phe Gly Lys Phe Lys Val Glu Thr
 100 105 110

Glu Pro Lys Leu Ala Trp Lys Asn Ile Glu Gly Lys Phe Ser Phe Leu
 115 120 125

Phe Gln Asn Ser Phe Asn Glu Met Leu Glu Asp Phe Lys Ala Asn Glu
130 135 140

Gln Asp Leu Lys Ile Pro Tyr Asp Trp Thr Ile Tyr Phe Leu Arg Lys
145 150 155 160

Lys Ala Leu Ser Gln Lys Ile Glu Lys Glu Glu Leu Ala Trp Ile Leu
165 170 175

Leu Asn Phe Asn Gln Lys Arg Gly Tyr Tyr Gln Leu Arg Gly Glu Glu
180 185 190

Glu Glu Glu Asn Pro Asn Lys Leu Val Glu Phe Tyr Ser Leu Lys Val
195 200 205

Val Asp Val Leu Ala Asp Glu Pro Gln Lys Gly Lys Ser Glu Ile Trp
210 215 220

Tyr Ser Leu Val Leu Lys Asn Gly Trp Ile Tyr Arg Arg Ala Ser Lys
225 230 235 240

Ile Pro Leu Phe Glu Trp Lys Asp Lys Thr Arg Asp Phe Ile Val Thr
245 250 255

Thr Asp Leu Asn Asp Asp Gly Ser Ile Lys Lys Asp Lys Asp Gly Asn
260 265 270

Glu Lys Arg Ser Phe Arg Ala Pro Gly Glu Asn Asp Trp Thr Leu Val
275 280 285

Lys Lys Lys Thr Glu Gln Glu Ile Glu Gln Ser His Lys Thr Val Gly
290 295 300

Thr Tyr Ile Tyr Glu Thr Leu Leu Leu Asn Pro Lys Gln Lys Ile Lys
305 310 315 320

Gly Lys Leu Val Arg Thr Ile Glu Arg Lys Phe Tyr Lys Asp Glu Leu
325 330 335

Lys Gln Ile Leu Glu Lys Gln Lys Glu Phe His Gln Glu Leu Gln Asn
340 345 350

Asp Asp Leu Tyr Asn Asp Cys Ile Arg Glu Leu Tyr Arg Asn Asn Glu
355 360 365

Ala His Gln Leu Thr Leu Ser Lys Lys Asp Phe Val His Leu Leu Met
370 375 380

Asp Asp Leu Ile Phe Tyr Gln Arg Pro Leu Arg Ser Gln Lys Ser Ser
385 390 395 400

Ile Ser Asn Cys Thr Leu Glu Phe Arg Lys Tyr Lys Asp Glu Asn Gly
405 410 415

Ile Glu His Thr Gln Tyr Leu Lys Ala Ile Pro Lys Ser Asn Pro Tyr
420 425 430

Tyr Gln Glu Phe Arg Leu Trp Gln Trp Ile Tyr Asn Leu Asn Ile Tyr
435 440 445

Arg Lys Asp Asp Glu Val Asn Val Thr Lys Asp Phe Leu Asn Thr Thr
450 455 460

Lys Glu Phe Glu Asn Leu Phe Glu Phe Leu Asn Ser Arg Lys Glu Val
465 470 475 480

Asp Gln Lys Ala Leu Leu Lys Tyr Phe Lys Leu Asn Glu Lys Thr His
485 490 495

Arg Trp Asn Phe Val Glu Asp Lys Lys Tyr Pro Cys Asn Glu Thr Lys
500 505 510

Thr Met Ile Ser Ser Arg Leu Asp Lys Val Glu Asn Ile Ser Asp Asp
515 520 525

Phe Leu Thr Arg Asn Ile Glu Gln Lys Ile Trp His Ile Ile Tyr Ser
530 535 540

Val Asn Asp Lys Ile Glu Tyr Glu Lys Ala Leu Lys Ser Phe Ala Arg
545 550 555 560

Lys His His Leu Asp Glu Thr Ser Phe Phe Glu Ala Phe Arg Lys Phe
565 570 575

Pro Pro Phe Lys Ser Glu Tyr Gly Ser Phe Ser Glu Lys Ala Ile Lys
580 585 590

Lys Leu Leu Pro Leu Met Arg Leu Gly Lys Tyr Trp Asn Tyr Ala Asp
595 600 605

Ile Asp Lys Tyr Ser Arg Glu Arg Ile Gln Lys Ile Ile Thr Gly Glu
610 615 620

Tyr Asp Glu Asn Ile Lys Asp Lys Ile Arg Glu Lys Ala Ile His Leu
625 630 635 640

Thr Lys Glu Asn Asp Phe Gln Gly Leu Gln Leu Trp Leu Ala Gln Tyr
645 650 655

Leu Val Tyr Gly Arg His Ser Glu Ala Ser Met Ile Gly Lys Trp Asn
660 665 670

Ser Ala Asp Asp Leu Glu Val Phe Leu Lys Asp Phe Lys Gln His Ser
675 680 685

Leu Arg Asn Pro Ile Val Glu Gln Val Ile Thr Glu Thr Leu Arg Val
690 695 700

Val Lys Asp Ile Trp Leu Lys Tyr Gly Asn Gly Thr Lys Asp Phe Phe
705 710 715 720

Asn Glu Ile His Ile Glu Leu Gly Arg Glu Met Lys Leu Pro Ala Asp
725 730 735

Asp Arg Lys Lys Leu Thr Ser Gln Ile Ser Glu Asn Glu Asn Thr Asn
740 745 750

Leu Arg Ile Lys Ala Leu Leu Ala Glu Met Ile Asn Asp His Ser Val
755 760 765

Glu Asn Val Arg Pro Phe Ser Pro Met Gln Gln Glu Ile Leu Lys Ile
770 775 780

Tyr Glu Asp Gly Val Leu Asn Ser Asp Ile Glu Ile Glu Asp Asp Ile
785 790 795 800

Leu Lys Ile Ser Lys Thr Ala Gln Pro Ser Pro Ser Asp Leu Lys Arg
805 810 815

Tyr Lys Leu Trp Leu Glu Gln Lys Tyr Lys Ser Pro Tyr Thr Gly Gln
820 825 830

Ile Ile Pro Leu Asn Lys Leu Phe Thr Pro Glu Tyr Glu Ile Glu His
835 840 845

Ile Ile Pro Gln Ser Arg Tyr Phe Asp Asp Ser Phe Ser Asn Lys Val
850 855 860

Ile Cys Glu Ser Ala Val Asn Lys Leu Lys Asp Asn Tyr Ile Gly Leu
865 870 875 880

Glu Phe Ile Lys Lys Leu Gly Gly Thr Val Val Glu Cys Gly Asn Gly
885 890 895

Lys Ser Val Thr Val Leu Lys Glu Asp Gln Tyr Glu Asp Phe Val Lys
900 905 910

Lys His Tyr Ala Asn Asn Arg Gly Lys Arg Asn Lys Leu Leu Met Glu
915 920 925

Asp Ile Pro Glu Lys Met Ile Glu Arg Gln Leu Asn Asp Thr Arg Tyr
930 935 940

Ile Ser Lys Tyr Ile Ser Gly Val Leu Ser Asn Ile Val Arg Val Glu
945 950 955 960

Asp Gly Ser Asp Glu Gly Ile Asn Ser Lys Asn Ile Val Pro Gly Asn
965 970 975

Gly Lys Ile Thr Ser Gln Leu Lys Gln Asp Trp Gly Leu Asn Asp Val
980 985 990

Trp Asn Asp Leu Ile Leu Pro Arg Phe Glu Arg Met Asn Gln Leu Thr
995 1000 1005

Asn Ser Thr Asp Phe Thr Ala Trp Asn Glu Asn Tyr Gln Lys Phe
1010 1015 1020

Leu Pro Thr Val Pro Ile Glu Tyr Ser Lys Gly Phe Ser Lys Lys
1025 1030 1035

Arg Ile Asp His Arg His His Ala Leu Asp Ala Leu Val Ile Ala
1040 1045 1050

Cys Ala Thr Lys Asp His Val Asn Leu Leu Asn Asn Gln Ser Ala
1055 1060 1065

Lys Ser Asp Thr Lys Arg Tyr Asp Leu Lys Lys Lys Leu Met Lys
1070 1075 1080

Phe Glu Lys Val Val Tyr His His Thr Gln Thr Gly Glu Lys Ile
1085 1090 1095

Glu Arg Glu Ile Pro Arg Gln Phe Leu Lys Pro Trp Glu Asn Phe
1100 1105 1110

Thr Val Asp Ala Lys His Asn Leu Glu Ser Ile Ile Val Ser Phe
1115 1120 1125

Lys Gln Asn Leu Arg Val Ile Asn Lys Ala Thr Asn Tyr Tyr Glu
1130 1135 1140

Lys Tyr Val Glu Lys Asp Gly Thr Lys Asn Lys Glu Arg Val Glu
1145 1150 1155

Gln Thr Gly Thr Asn Trp Ala Ile Arg Lys Pro Met His Lys Asp
1160 1165 1170

Thr Val Ser Gly Lys Val Asp Leu Pro Trp Val Lys Val Pro Lys
1175 1180 1185

Gly Lys Ile Leu Thr Ala Thr Arg Lys Ser Leu Asp Thr Ser Phe
1190 1195 1200

Asp Leu Lys Ser Ile Gly Ser Ile Thr Asp Thr Gly Ile Gln Lys
1205 1210 1215

Ile Leu Lys Asn Tyr Leu Ala Phe Lys Asp Gly Asn Pro Glu Met
1220 1225 1230

Ala Phe Ser Pro Glu Gly Ile Asp Asp Leu Asn Lys Asn Ile Glu
1235 1240 1245

Lys Tyr Asn Asp Gly Lys Pro His Gln Pro Ile Asn Lys Val Arg
1250 1255 1260

Val Phe Glu Leu Gly Ser Lys Phe Gln Val Gly Gln Thr Gly Asn
1265 1270 1275

Lys Lys Asp Lys Tyr Val Glu Ala Ala Lys Gly Thr Asn Leu Phe
1280 1285 1290

Phe Ala Val Tyr Glu Asp Glu Lys Gly Lys Arg Ser Tyr Glu Thr
1295 1300 1305

Ile Pro Leu Asn Glu Val Ile Glu Arg Gln Lys Gln Gly Leu Ser
1310 1315 1320

Val Val Asp Leu Lys Asp Thr Asn Asp Phe Tyr Leu Cys Pro Asn
1325 1330 1335

Asp Leu Val Tyr Ile Pro Ser Gly Asp Glu Leu Glu Asn Val Asn
1340 1345 1350

Asn Val Asp Phe Lys Asp Ile Lys Lys Asp Lys Asn Glu Arg Ile
1355 1360 1365

Tyr Lys Val Val Ser Phe Ser Gly Ser Gln Ile Phe Phe Val Arg
1370 1375 1380

Gln Asp Ile Ala Thr Phe Ile Val Asn Lys Ala Glu Leu Ser Thr
1385 1390 1395

Leu Asn Lys Met Glu Arg Ala Ile Asp Gly Ser Met Ile Lys Glu
1400 1405 1410

Asn Cys Ile Lys Leu Asn Ile Asp Arg Leu Gly Asn Ile Ser Lys
1415 1420 1425

Ala

<210> 99
<211> 1112
<212> PRT
<213> Bacillus sp.

<223> APG02789_RGN

<400> 99

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

Asp Ser Ile Tyr Leu Arg Lys Ala Leu Trp Arg Thr His Gln Leu Ile
20 25 30

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

Glu Ala Ile Gly Asp Lys Thr Lys Glu Ala Tyr Gln Ala Glu Leu Ile
50 55 60

Asn Ile Ile Arg Asn Gln Gln Arg Asn Asn Gly Ser Ser Glu Glu His
65 70 75 80

Gly Ser Asp Gln Glu Ile Leu Ala Leu Leu Arg Gln Leu Tyr Glu Leu
85 90 95

Ile Ile Pro Ser Ser Ile Gly Glu Ser Gly Asp Ala Asn Gln Leu Gly
100 105 110

Asn Lys Phe Leu Tyr Pro Leu Val Asp Pro Asn Ser Gln Ser Gly Lys
115 120 125

Gly Thr Ser Asn Ala Gly Arg Lys Pro Arg Trp Lys Arg Leu Lys Glu
130 135 140

Glu Gly Asn Pro Asp Trp Glu Leu Glu Lys Lys Lys Asp Glu Glu Arg
145 150 155 160

Lys Ala Lys Asp Pro Thr Val Lys Ile Phe Asp Asn Leu Asn Lys Tyr
165 170 175

Gly Leu Leu Pro Leu Phe Pro Leu Phe Thr Asn Ile Gln Lys Asp Ile
180 185 190

Glu Trp Leu Pro Leu Gly Lys Arg Gln Ser Val Arg Lys Trp Asp Lys
195 200 205

Asp Met Phe Ile Gln Ala Ile Glu Arg Leu Leu Ser Trp Glu Ser Trp
210 215 220

Asn Arg Arg Val Ala Asp Glu Tyr Lys Gln Leu Lys Glu Lys Thr Glu
225 230 235 240

Ser Tyr Tyr Lys Glu His Leu Thr Gly Gly Glu Glu Trp Ile Glu Lys
245 250 255

Ile Arg Lys Phe Glu Lys Glu Arg Asn Met Glu Leu Glu Lys Asn Ala
260 265 270

Phe Ala Pro Asn Asp Gly Tyr Phe Ile Thr Ser Arg Gln Ile Arg Gly
275 280 285

Trp Asp Arg Val Tyr Glu Lys Trp Ser Lys Leu Pro Glu Ser Ala Ser
290 295 300

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

Glu Gly Phe Gly Asp Pro Lys Val Phe Ser Phe Leu Ala Asn Arg Glu
325 330 335

Asn Arg Asp Ile Trp Arg Gly His Ser Glu Arg Ile Tyr His Ile Ala

340

345

350

Ala Tyr Asn Gly Leu Gln Lys Lys Leu Ser Arg Thr Lys Glu Gln Ala
 355 360 365

Thr Phe Thr Leu Pro Asp Ala Ile Glu His Pro Leu Trp Ile Arg Tyr
 370 375 380

Glu Ser Pro Gly Gly Thr Asn Leu Asn Leu Phe Lys Leu Glu Glu Lys
 385 390 395 400

Gln Lys Lys Asn Tyr Tyr Val Thr Leu Ser Lys Ile Ile Trp Pro Ser
 405 410 415

Glu Glu Lys Trp Ile Glu Lys Glu Asn Ile Glu Ile Pro Leu Ala Pro
 420 425 430

Ser Ile Gln Phe Asn Arg Gln Ile Lys Leu Lys Gln His Val Lys Gly
 435 440 445

Lys Gln Glu Ile Ser Phe Ser Asp Tyr Ser Ser Arg Ile Ser Leu Asp
 450 455 460

Gly Val Leu Gly Gly Ser Arg Ile Gln Phe Asn Arg Lys Tyr Ile Lys
 465 470 475 480

Asn His Lys Glu Leu Leu Gly Glu Gly Asp Ile Gly Pro Val Phe Phe
 485 490 495

Asn Leu Val Val Asp Val Ala Pro Leu Gln Glu Thr Arg Asn Gly Arg
 500 505 510

Leu Gln Ser Pro Ile Gly Lys Ala Leu Lys Val Ile Ser Ser Asp Phe
 515 520 525

Ser Lys Val Ile Asp Tyr Lys Pro Lys Glu Leu Met Asp Trp Met Asn
 530 535 540

Thr Gly Ser Ala Ser Asn Ser Phe Gly Val Ala Ser Leu Leu Glu Gly
 545 550 555 560

Met Arg Val Met Ser Ile Asp Met Gly Gln Arg Thr Ser Ala Ser Val
565 570 575

Ser Ile Phe Glu Val Val Lys Glu Leu Pro Lys Asp Gln Glu Gln Lys
580 585 590

Leu Phe Tyr Ser Ile Asn Asp Thr Glu Leu Phe Ala Ile His Lys Arg
595 600 605

Ser Phe Leu Leu Asn Leu Pro Gly Glu Val Val Thr Lys Asn Asn Lys
610 615 620

Gln Gln Arg Gln Glu Arg Arg Lys Lys Arg Gln Phe Val Arg Ser Gln
625 630 635 640

Ile Arg Met Leu Ala Asn Val Leu Arg Leu Glu Thr Lys Lys Thr Pro
645 650 655

Asp Glu Arg Lys Lys Ala Ile His Lys Leu Met Glu Ile Val Gln Ser
660 665 670

Tyr Asp Ser Trp Thr Ala Ser Gln Lys Glu Val Trp Glu Lys Glu Leu
675 680 685

Asn Leu Leu Thr Asn Met Ala Ala Phe Asn Asp Glu Ile Trp Lys Glu
690 695 700

Ser Leu Val Glu Leu His His Arg Ile Glu Pro Tyr Val Gly Gln Ile
705 710 715 720

Val Ser Lys Trp Arg Lys Gly Leu Ser Glu Gly Arg Lys Asn Leu Ala
725 730 735

Gly Ile Ser Met Trp Asn Ile Asp Glu Leu Glu Asp Thr Arg Arg Leu
740 745 750

Leu Ile Ser Trp Ser Lys Arg Ser Arg Thr Pro Gly Glu Ala Asn Arg
755 760 765

Ile Glu Thr Asp Glu Pro Phe Gly Ser Ser Leu Leu Gln His Ile Gln

770

775

780

Asn Val Lys Asp Asp Arg Leu Lys Gln Met Ala Asn Leu Ile Ile Met
785 790 795 800

Thr Ala Leu Gly Phe Lys Tyr Asp Lys Glu Glu Lys Asp Arg Tyr Lys
805 810 815

Arg Trp Lys Glu Thr Tyr Pro Ala Cys Gln Ile Ile Leu Phe Glu Asn
820 825 830

Leu Asn Arg Tyr Leu Phe Asn Leu Asp Arg Ser Arg Arg Glu Asn Ser
835 840 845

Arg Leu Met Lys Trp Ala His Arg Ser Ile Pro Arg Thr Val Ser Met
850 855 860

Gln Gly Glu Met Phe Gly Leu Gln Val Gly Asp Val Arg Ser Glu Tyr
865 870 875 880

Ser Ser Arg Phe His Ala Lys Thr Gly Ala Pro Gly Ile Arg Cys His
885 890 895

Ala Leu Thr Glu Glu Asp Leu Lys Ala Gly Ser Asn Thr Leu Lys Arg
900 905 910

Leu Ile Glu Asp Gly Phe Ile Asn Glu Ser Glu Leu Ala Tyr Leu Lys
915 920 925

Lys Gly Asp Ile Ile Pro Ser Gln Gly Gly Glu Leu Phe Val Thr Leu
930 935 940

Ser Lys Arg Tyr Lys Lys Asp Ser Asp Asn Asn Glu Leu Thr Val Ile
945 950 955 960

His Ala Asp Ile Asn Ala Ala Gln Asn Leu Gln Lys Arg Phe Trp Gln
965 970 975

Gln Asn Ser Glu Val Tyr Arg Val Pro Cys Gln Leu Ala Arg Met Gly
980 985 990

Glu Asp Lys Leu Tyr Ile Pro Lys Ser Gln Thr Glu Thr Ile Lys Lys
995 1000 1005

Tyr Phe Gly Lys Gly Ser Phe Val Lys Asn Asn Thr Glu Gln Glu
1010 1015 1020

Val Tyr Lys Trp Glu Lys Ser Glu Lys Met Lys Ile Lys Thr Asp
1025 1030 1035

Thr Thr Phe Asp Leu Gln Asp Leu Asp Gly Phe Glu Asp Ile Ser
1040 1045 1050

Lys Thr Ile Glu Leu Ala Gln Glu Gln Gln Lys Lys Tyr Leu Thr
1055 1060 1065

Met Phe Arg Asp Pro Ser Gly Tyr Phe Phe Asn Asn Glu Thr Trp
1070 1075 1080

Arg Pro Gln Lys Glu Tyr Trp Ser Ile Val Asn Asn Ile Ile Lys
1085 1090 1095

Ser Cys Leu Lys Lys Lys Ile Leu Ser Asn Lys Val Glu Leu
1100 1105 1110

<210> 100

<211> 1070

<212> PRT

<213> Bacillus sp.

<223> APG05083.1_RGN

<400> 100

Met Arg Glu Leu Asp Tyr Arg Ile Gly Leu Asp Ile Gly Thr Asn Ser
1 5 10 15

Ile Gly Trp Gly Ile Ile Glu Leu Ser Trp Asn Lys Asp Arg Glu Gln
20 25 30

Tyr Glu Lys Ala Arg Ile Val Asp Lys Gly Val Arg Met Phe Asp Lys
35 40 45

Ala Glu Ile Pro Lys Thr Gly Ala Ser Leu Ala Glu Pro Arg Arg Ile
50 55 60

Ala Arg Ser Ser Arg Arg Arg Leu Asn Arg Lys Ser Gln Arg Lys Lys
65 70 75 80

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

Leu Ala Ser Leu Tyr Pro Leu Thr Lys Ser Ser Met Asp Ile Trp Asp
100 105 110

Ile Arg Leu Asp Gly Leu Asp Arg Leu Leu Asp Arg Phe Glu Trp Thr
115 120 125

Arg Leu Leu Ile His Leu Ala Gln Arg Arg Gly Phe Lys Ser Asn Arg
130 135 140

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

Gln Ala Asn Glu Lys Arg Leu Ser Leu Tyr Arg Thr Val Gly Glu Met
165 170 175

Trp Met Lys Asn Glu Asp Phe Ser Lys Tyr Asp Lys Arg Arg Asn Ser
180 185 190

Ser Asn Glu Tyr Val Phe Ser Val Ser Arg Ala Asp Leu Glu Lys Glu
195 200 205

Ile Val Thr Leu Phe Glu Ala Gln Arg Lys Phe Gln Ser Ser Tyr Ala
210 215 220

Ser Ala Asp Leu Gln Lys Thr Tyr Leu Gln Ile Trp Ala His Gln Leu
225 230 235 240

Pro Phe Ala Ser Gly Asn Ala Ile Val Asn Lys Val Gly Tyr Cys Ser
245 250 255

Leu Leu Lys Gly Lys Glu Lys Arg Val Pro Lys Ala Thr Tyr Thr Phe
260 265 270

Gln Tyr Phe Ser Thr Leu Asp Gln Ile Asn Arg Thr Arg Leu Gly Pro
275 280 285

Asn Phe Gln Pro Phe Thr Lys Glu Gln Arg Asp Val Ile Leu Asp Glu
290 295 300

Met Phe Asn Arg Thr Asp Tyr Tyr Lys Lys Lys Thr Ile Pro Glu Val
305 310 315 320

Thr Tyr Tyr Asp Ile Arg Lys Trp Leu Ala Leu Asp Glu Thr Ile Gln
325 330 335

Phe Lys Gly Leu Thr Tyr Asp Pro Asn Glu Glu Leu Lys Lys Ile Glu
340 345 350

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

Val Thr Asn Tyr Ala Lys Lys Thr Asn Glu Ala Phe Ser Thr Leu Asp
370 375 380

Tyr Asp Thr Phe Ala Tyr Ala Leu Thr Val Tyr Lys Thr Asp Lys Asp
385 390 395 400

Ile Arg Ser Tyr Leu Lys Lys Ser Asn Asn Leu Ser Lys Cys Cys Tyr
405 410 415

Asp Asp Gln Leu Ile Glu Glu Leu Leu Thr Leu Ser Tyr Thr Lys Phe
420 425 430

Gly His Leu Ser Phe Lys Ala Ile Asn His Val Leu Pro Ile Met Gln
435 440 445

Glu Gly Arg Thr Tyr Gln Glu Ala Ile His Gln Leu Gly Tyr Asp Ala
450 455 460

Thr Asn Leu Lys Lys Glu Asn Arg Ser Met Phe Leu Pro Leu Phe Pro
465 470 475 480

Asp Glu Ile Thr Asn Pro Ile Val Lys Arg Ala Leu Thr Gln Ala Arg
485 490 495

Lys Val Val Asn Ala Ile Ile Arg Arg Tyr Gly Ser Pro Asn Ser Val
500 505 510

His Ile Glu Leu Ala Arg Glu Leu Ser Lys Ser His Asp Glu Arg Thr
515 520 525

Lys Ile Met Lys Ala His Asp Glu Asn Tyr Lys Lys Asn Lys Gly Ala
530 535 540

Ile Ser Ile Leu Ile Glu Asn Gly Ile Leu Asn Pro Thr Gly Tyr Asp
545 550 555 560

Ile Val Arg Tyr Lys Leu Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr
565 570 575

Ser Leu Lys Gln Ile Pro Ala Asn Thr Phe Phe Asn Glu Met Lys Lys
580 585 590

Glu Arg Ser Gly Ser Pro Val Leu Glu Ile Asp His Ile Leu Pro Tyr
595 600 605

Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Gly
610 615 620

Asp Glu Asn Gln Lys Lys Gly Asn Arg Ile Pro Tyr Thr Tyr Phe Leu
625 630 635 640

Glu Gly Asn Lys Asp Trp Glu Ser Phe Glu Ser Tyr Val Arg Leu Asn
645 650 655

Ser Phe Phe Ser Lys Lys Lys Arg Gly Tyr Leu Leu Lys Lys Ala Tyr
660 665 670

Leu Pro Arg Glu Ser Asn Met Ile Lys Glu Arg His Leu Asn Asp Thr
675 680 685

Arg Tyr Ala Ser Ser Tyr Leu Lys Asn Phe Ile Glu Lys Asn Leu Lys
690 695 700

Phe Lys Glu Val Glu Gly Ser Thr Arg Lys Lys His Val Gln Thr Val
705 710 715 720

Asn Gly Ile Ile Thr Ala His Leu Arg Lys Arg Trp Gly Leu Glu Lys
725 730 735

Asp Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
740 745 750

Ala Cys Thr Asp His His Met Val Thr Lys Val Thr Glu Tyr Tyr Gln
755 760 765

Ile Lys Glu Ser Asn Lys Ser Ile Arg Lys Pro Tyr Phe Pro Leu Pro
770 775 780

Trp Val Gly Phe Arg Glu Glu Ile Leu Ser His Leu Ala Arg Gln Pro
785 790 795 800

Ile Ala Arg Lys Ile Ser Glu Glu Leu Lys Ile Gly Tyr Gln Ser Phe
805 810 815

Asp Tyr Ile Leu Val Ser Arg Met Pro Lys Arg Ser Val Thr Gly Ala
820 825 830

Ala His Glu Gln Thr Ile Met Lys Lys Gly Gly Ile Asp Lys Lys Gly
835 840 845

Lys Thr Ile Ile Ile Lys Arg Val Tyr Leu Lys Asp Ile Lys Phe Asp
850 855 860

Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Leu Ala Thr
865 870 875 880

Tyr Glu Ala Ile Lys Gln Arg Tyr Ile Glu Tyr Gly Lys Glu Ser Lys
885 890 895

Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Lys Gly
900 905 910

Asn Leu Ile Lys Lys Ile Lys Val Glu Val Gln Thr Lys Ser Phe Val
915 920 925

Arg Glu Val Asn Gly Gly Val Ala Gln Asn Gly Asp Leu Val Arg Val
930 935 940

Asp Leu Phe Glu Lys Asp Asn Arg Tyr Tyr Met Ile Pro Ile Tyr Val
945 950 955 960

Met Asp Thr Val His Ser Glu Leu Pro Asn Lys Ala Val Thr Ser Ser
965 970 975

Lys Gly Tyr Glu Gln Trp Leu Thr Ile Asp Asn Ser Phe Thr Phe Lys
980 985 990

Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asn Glu
995 1000 1005

Asp Arg Phe Leu Tyr Phe Ser Thr Leu Asp Ile Asn Ser Asp Arg
1010 1015 1020

Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Gln Ala Glu Asn
1025 1030 1035

Arg Tyr Ser Leu Lys Thr Ile Glu Asn Leu Glu Lys Tyr Glu Val
1040 1045 1050

Gly Val Leu Gly Asp Leu Arg Phe Val Arg Gln Glu Ile Arg Lys
1055 1060 1065

Asn Phe
1070

<210> 101
<211> 1071
<212> PRT
<213> Bacillus sp.

<223> APG07433.1_RGN

<400> 101

Met Arg Glu Leu Asp Tyr Arg Ile Gly Leu Asp Ile Gly Thr Asn Ser

1	5	10	15
Ile Gly Trp Gly Val Ile Glu Leu Ser Trp Asn Lys Asp Arg Glu Arg	20	25	30
Tyr Glu Lys Val Arg Ile Val Asp Gln Gly Val Arg Met Phe Asp Arg	35	40	45
Ala Glu Met Pro Lys Thr Gly Ala Ser Leu Ala Glu Pro Arg Arg Ile	50	55	60
Ala Arg Ser Ser Arg Arg Arg Leu Asn Arg Lys Ser Gln Arg Lys Lys	65	70	75
Asn Ile Arg Asn Leu Leu Val Gln His Gly Val Ile Thr Gln Glu Glu	85	90	95
Leu Asp Ser Leu Tyr Pro Leu Ser Lys Lys Ser Met Asp Ile Trp Gly	100	105	110
Ile Arg Leu Asp Gly Leu Asp Arg Leu Leu Asn His Phe Glu Trp Ala	115	120	125
Arg Leu Leu Ile His Leu Ala Gln Arg Arg Gly Phe Lys Ser Asn Arg	130	135	140
Lys Ser Glu Leu Lys Asp Thr Glu Thr Gly Lys Val Leu Ser Ser Ile	145	150	155
Gln Leu Asn Glu Lys Arg Leu Ser Leu Tyr Arg Thr Val Gly Glu Met	165	170	175
Trp Met Lys Asp Pro Asp Phe Ser Lys Tyr Asp Arg Lys Arg Asn Ser	180	185	190
Pro Asn Glu Tyr Val Phe Ser Val Ser Arg Ala Glu Leu Glu Lys Glu	195	200	205
Ile Val Thr Leu Phe Ala Ala Gln Arg Arg Phe Gln Ser Pro Tyr Ala	210	215	220

Ser Lys Asp Leu Gln Glu Thr Tyr Leu Gln Ile Trp Thr His Gln Leu
225 230 235 240

Pro Phe Ala Ser Gly Asn Ala Ile Leu Asn Lys Val Gly Tyr Cys Ser
245 250 255

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

Gln Tyr Phe Ser Ala Leu Asp Gln Val Asn Arg Thr Arg Leu Gly Pro
275 280 285

Asp Phe Gln Pro Phe Thr Lys Glu Gln Arg Glu Ile Ile Leu Asn Asn
290 295 300

Met Phe Gln Arg Thr Asp Tyr Tyr Lys Lys Lys Thr Ile Pro Glu Val
305 310 315 320

Thr Tyr Tyr Asp Ile Arg Lys Trp Leu Glu Leu Asp Glu Thr Ile Gln
325 330 335

Phe Lys Gly Leu Asn Tyr Asp Pro Asn Glu Glu Leu Lys Lys Ile Glu
340 345 350

Lys Lys Pro Phe Ile Asn Leu Lys Ala Phe Tyr Glu Ile Asn Lys Val
355 360 365

Val Ala Asn Tyr Ser Glu Arg Thr Asn Glu Thr Phe Ser Thr Leu Asp
370 375 380

Tyr Asp Gly Ile Gly Tyr Ala Leu Thr Val Tyr Lys Thr Asp Lys Asp
385 390 395 400

Ile Arg Ser Tyr Leu Lys Ser Ser His Asn Leu Pro Lys Arg Cys Tyr
405 410 415

Asp Asp Gln Leu Ile Glu Glu Leu Leu Ser Leu Ser Tyr Thr Lys Phe
420 425 430

Gly His Leu Ser Leu Lys Ala Ile Asn His Val Leu Ser Ile Met Gln

435

440

445

Lys Gly Asn Thr Tyr Lys Glu Ala Val Asp Gln Leu Gly Tyr Asp Thr
 450 455 460

Ser Gly Leu Lys Lys Glu Lys Arg Ser Lys Phe Leu Pro Pro Ile Ser
 465 470 475 480

Asp Glu Ile Thr Asn Pro Ile Val Lys Arg Ala Leu Thr Gln Ala Arg
 485 490 495

Lys Val Val Asn Ala Ile Ile Arg Arg His Gly Ser Pro His Ser Val
 500 505 510

His Ile Glu Leu Ala Arg Glu Leu Ser Lys Asn His Asp Glu Arg Thr
 515 520 525

Lys Ile Val Ser Ala Gln Asp Glu Asn Tyr Lys Lys Asn Lys Gly Ala
 530 535 540

Ile Ser Ile Leu Ser Glu His Gly Ile Leu Asn Pro Thr Gly Tyr Asp
 545 550 555 560

Ile Val Arg Tyr Lys Leu Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr
 565 570 575

Ser Leu Lys Glu Ile Pro Ala Asp Thr Phe Phe Asn Glu Leu Lys Lys
 580 585 590

Glu Arg Asn Gly Ala Pro Ile Leu Glu Val Asp His Ile Leu Pro Tyr
 595 600 605

Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Ser
 610 615 620

Asp Glu Asn Arg Lys Lys Gly Asn Arg Ile Pro Tyr Thr Tyr Phe Leu
 625 630 635 640

Glu Thr Asn Lys Asp Trp Glu Ala Phe Glu Arg Tyr Val Arg Ser Asn
 645 650 655

Lys Phe Phe Ser Lys Lys Lys Arg Glu Tyr Leu Leu Lys Arg Ala Tyr
660 665 670

Leu Pro Arg Glu Ser Glu Leu Ile Lys Glu Arg His Leu Asn Asp Thr
675 680 685

Arg Tyr Ala Ser Thr Phe Leu Lys Asn Phe Ile Glu Gln Asn Leu Gln
690 695 700

Phe Lys Glu Ala Glu Asp Asn Pro Arg Lys Arg Arg Val Gln Thr Val
705 710 715 720

Asn Gly Val Ile Thr Ala His Phe Arg Lys Arg Trp Gly Leu Glu Lys
725 730 735

Asp Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
740 745 750

Ala Cys Thr Asp His His Met Val Thr Arg Val Thr Glu Tyr Tyr Gln
755 760 765

Ile Lys Glu Ser Asn Lys Ser Val Lys Lys Pro Tyr Phe Pro Met Pro
770 775 780

Trp Glu Gly Phe Arg Asp Glu Leu Leu Ser His Leu Ala Ser Gln Pro
785 790 795 800

Ile Ala Lys Lys Ile Ser Glu Glu Leu Lys Ala Gly Tyr Gln Ser Leu
805 810 815

Asp Tyr Ile Phe Val Ser Arg Met Pro Lys Arg Ser Ile Thr Gly Ala
820 825 830

Ala His Lys Gln Thr Ile Met Arg Lys Gly Gly Ile Asp Lys Lys Gly
835 840 845

Lys Thr Ile Ile Ile Glu Arg Leu His Leu Lys Asp Ile Lys Phe Asp
850 855 860

Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Met Ala Thr

[illegible]

<210> 102
<211> 1071
<212> PRT
<213> Bacillus sp.

<223> APG07513.1_RGN

<400> 102

Met Arg Glu Leu Asp Tyr Arg Ile Gly Leu Asp Ile Gly Thr Asn Ser
1 5 10 15

Ile Gly Trp Gly Val Ile Glu Leu Ser Trp Asn Lys Asp Arg Glu Gln
20 25 30

Tyr Glu Lys Thr Arg Ile Val Asp Lys Gly Val Arg Met Phe Asp Lys
35 40 45

Ala Glu Ile Pro Lys Thr Gly Ala Ser Leu Ala Glu Pro Arg Arg Ile
50 55 60

Ala Arg Ser Ser Arg Arg Arg Leu Asn Arg Lys Ser Gln Arg Lys Lys
65 70 75 80

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

Leu Thr Ser Leu Tyr Pro Leu Ser Lys Ser Ser Met Asp Ile Trp Asp
100 105 110

Ile Arg Leu Asp Gly Leu Asp Arg Leu Leu Asp Arg Phe Glu Trp Ala
115 120 125

Arg Leu Leu Ile His Leu Ala Gln Arg Arg Gly Phe Lys Ser Asn Arg
130 135 140

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

Gln Val Asn Glu Lys Arg Leu Ser Leu Tyr Arg Thr Val Gly Glu Met
165 170 175

Trp Met Lys Asn Ala Asp Cys Ser Lys Tyr Gly Lys Arg Arg Asn Ser
180 185 190

Pro Asn Glu Tyr Val Phe Ser Val Ser Arg Ala Asp Leu Glu Lys Glu
195 200 205

Ile Val Thr Leu Phe Glu Ala Gln Arg Lys Phe His Ser Ser Tyr Ala
210 215 220

Ser Val Asp Leu Gln Lys Thr Tyr Ile Gln Ile Trp Ala His Gln Leu
225 230 235 240

Pro Phe Ala Ser Gly Asn Ala Ile Val Asn Lys Val Gly Tyr Cys Ser
245 250 255

Leu Leu Lys Gly Lys Glu Lys Arg Val Pro Lys Ala Thr Tyr Thr Phe
260 265 270

Gln Tyr Phe Asn Thr Leu Asp Gln Ile Asn Arg Thr Arg Leu Gly Pro
275 280 285

Asn Phe Gln Pro Phe Thr Lys Glu Gln Arg Asp Ile Ile Leu Asp Lys
290 295 300

Met Phe Gln Arg Thr Asp Tyr Tyr Lys Lys Lys Thr Ile Pro Glu Val
305 310 315 320

Thr Tyr Tyr Asp Ile Arg Lys Trp Leu Ala Leu Asp Glu Thr Ile Gln
325 330 335

Phe Lys Gly Leu Thr Tyr Asp Pro Asn Glu Glu Leu Lys Lys Ile Glu
340 345 350

Met Lys Pro Phe Ile Asn Leu Lys Pro Phe Tyr Glu Ile Lys Lys Val
355 360 365

Val Thr Asn Tyr Ala Lys Lys Thr Asn Glu Val Phe Ser Ala Leu Asp
370 375 380

Tyr Asp Thr Val Ala Tyr Ala Leu Thr Val Tyr Lys Thr Asp Lys Asp
385 390 395 400

Ile Arg Ser Tyr Leu Lys Arg Ser Asn Asn Leu Ser Lys Arg Cys Tyr
405 410 415

Asp Asp Gln Leu Ile Glu Glu Leu Leu Thr Leu Ser Tyr Thr Lys Phe
420 425 430

Gly His Leu Ser Phe Lys Ala Ile Asn His Val Leu Pro Ile Met Gln
435 440 445

Glu Gly Arg Thr Tyr Gln Glu Ala Ile His Gln Leu Gly Tyr Asp Thr
450 455 460

Thr Asn Leu Lys Lys Glu Asn Arg Ser Met Phe Leu Pro Ile Ile Pro
465 470 475 480

Asp Glu Ile Thr Asn Pro Ile Val Lys Arg Ala Leu Thr Gln Ala Arg
485 490 495

Lys Val Val Asn Ala Ile Ile Arg Arg Tyr Gly Ser Pro Asn Ser Val
500 505 510

His Ile Glu Leu Ala Arg Glu Leu Ser Lys Ser His Asp Glu Arg Lys
515 520 525

Lys Ile Met Thr Ala His Asp Glu Asn Tyr Lys Lys Asn Lys Gly Ala
530 535 540

Val Ser Ile Leu Ile Asp Asn Gly Ile Leu Asn Pro Thr Gly Tyr Asp
545 550 555 560

Ile Val Arg Tyr Lys Leu Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr
565 570 575

Ser Leu Lys Lys Ile Pro Ala Asn Thr Phe Phe Asn Glu Leu Lys Lys
580 585 590

Glu Arg Ser Gly Pro Pro Val Leu Glu Val Asp His Ile Leu Pro Tyr
595 600 605

Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Gly
610 615 620

Asp Glu Asn Gln Lys Lys Gly Asn Arg Ile Pro Tyr Thr Phe Phe Ser
625 630 635 640

Glu Glu Asp Lys Glu Trp Glu Ser Phe Glu Ser Tyr Val Arg Ser Asn
645 650 655

Ser Phe Phe Ser Lys Lys Lys Arg Gly Tyr Leu Leu Lys Lys Ala Tyr
660 665 670

Leu Pro Arg Glu Ser Asn Leu Ile Lys Glu Arg His Leu Asn Asp Thr
675 680 685

Arg Tyr Ala Ser Ser Tyr Leu Lys Asn Phe Ile Glu Lys Asn Leu Lys
690 695 700

Phe Lys Glu Ala Val Gly Ile Thr Arg Lys Lys Tyr Val Gln Thr Val
705 710 715 720

Asn Gly Val Ile Thr Ala His Leu Arg Lys Arg Trp Gly Leu Glu Lys
725 730 735

Asp Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
740 745 750

Ala Cys Thr Asp His His Met Val Thr Lys Val Thr Glu Tyr Tyr Gln
755 760 765

Ile Lys Glu Gly Asn Lys Ser Ile Lys Lys Pro Tyr Phe Pro Leu Pro
770 775 780

Trp Met Gly Phe Arg Glu Glu Ile Leu Ser His Leu Glu Ser Gln Pro
785 790 795 800

Ile Ala Arg Lys Ile Ser Glu Glu Leu Lys Ile Gly Tyr Gln Ser Pro
805 810 815

Asp Tyr Ile Leu Val Ser Arg Met Pro Lys Arg Ser Val Thr Gly Ser
820 825 830

Ala His Asp Gln Thr Val Met Lys Lys Gly Asp Ile Asp Lys Lys Gly
835 840 845

Lys Thr Ile Ile Ile Lys Arg Val His Leu Lys Asp Ile Lys Phe Asp
850 855 860

Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Leu Ala Thr
865 870 875 880

Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Glu Tyr Arg Lys Glu Ser Lys
885 890 895

Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Lys Gly
900 905 910

Asn Leu Ile Lys Lys Ile Lys Val Glu Val Gln Thr Lys Ser Phe Val
915 920 925

Arg Glu Ile Asn Gly Gly Val Ala Gln Asn Gly Asp Leu Val Arg Val
930 935 940

Asp Leu Phe Glu Lys Asp Asn Arg Tyr Tyr Met Val Pro Ile Tyr Val
945 950 955 960

Val Asp Thr Val Arg Ser Glu Leu Pro Asn Lys Ala Val Thr Ser Ser
965 970 975

Lys Gly Tyr Glu Gln Trp Leu Ser Ile Asp Asn Ser Phe Thr Phe Lys
980 985 990

Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asp Glu
995 1000 1005

Asp Arg Phe Leu Tyr Phe Ser Thr Leu Asp Ile Asn Ser Asp Arg
1010 1015 1020

Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Gln Ala Glu Tyr
1025 1030 1035

Arg Tyr Ser Leu Lys Thr Ile Glu Asn Leu Glu Lys Tyr Glu Ile
1040 1045 1050

Gly Val Leu Gly Asp Leu Arg Leu Val Arg Gln Glu Thr Arg Lys
1055 1060 1065

Ile Phe Lys
1070

<210> 103

<211> 1072

<212> PRT

<213> Bacillus sp.

<223> APG08290.1_RGN

<400> 103

Met Ser Glu Leu Asp Tyr Arg Ile Gly Leu Asp Ile Gly Thr Asn Ser
1 5 10 15

Ile Gly Trp Gly Val Ile Glu Leu Phe Trp Asn Lys Asp Arg Glu Arg
20 25 30

Tyr Glu Lys Val Arg Ile Val Asp Lys Gly Val Arg Met Phe Asp Lys
35 40 45

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

Ala Arg Ser Ser Arg Arg Arg Leu Asn Arg Lys Ser Gln Arg Lys Lys
65 70 75 80

Glu Ile Arg Asn Leu Leu Val Gln His Gly Met Ile Thr Gln Glu Glu
85 90 95

Leu Asp Leu Leu Tyr Pro Leu Ser Lys Lys Ser Ile Asp Ile Trp Asp
100 105 110

Ile Arg Leu Asp Gly Leu Asp Arg Leu Leu Asn His Leu Glu Trp Ala
115 120 125

Arg Leu Leu Ile His Leu Ala Gln Arg Arg Gly Phe Lys Ser Asn Arg

130

135

140

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

Gln Val Asn Glu Lys Arg Leu Phe Leu Tyr Arg Thr Val Gly Glu Met
 165 170 175

Trp Ile Lys Asp Ala Glu Phe Ser Lys Tyr Asp Arg Arg Arg Asn Ser
 180 185 190

Pro Asn Glu Tyr Val Phe Ser Val Ser Arg Ala Asp Leu Glu Lys Glu
 195 200 205

Ile Val Thr Leu Phe Glu Ala Gln Arg Lys Phe Gln Ser Ser Tyr Ala
 210 215 220

Ser Lys Asn Leu Gln Glu Thr Tyr Leu Gln Ile Trp Ala His Gln Leu
 225 230 235 240

Pro Phe Ala Ser Gly Asn Ala Ile Leu Asn Lys Val Gly Tyr Cys Ser
 245 250 255

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

Gln Tyr Phe Ser Ala Leu Asp Gln Val Asn Arg Thr Arg Leu Gly Pro
 275 280 285

Asp Phe Gln Pro Phe Thr Gln Glu Gln Lys Glu Ile Ile Leu Asp Lys
 290 295 300

Met Phe Gln Arg Thr Asp Tyr Tyr Lys Lys Lys Thr Ile Pro Glu Val
 305 310 315 320

Ser Tyr Tyr Asp Ile Arg Lys Trp Leu Glu Leu Asp Glu Thr Ile Gln
 325 330 335

Phe Lys Gly Leu Asn Tyr Asp Pro Asn Glu Glu Leu Lys Lys Ile Glu
 340 345 350

Lys Lys Pro Phe Ile Asn Leu Lys Ala Phe Tyr Glu Ile Lys Lys Val
355 360 365

Val Ala Asn Tyr Ala Glu Arg Thr Asn Glu Ala Phe Ser Thr Leu Asp
370 375 380

Tyr Asp Ala Ile Ala Tyr Ala Leu Thr Val Tyr Lys Thr Asp Lys Asp
385 390 395 400

Ile Arg Ser Tyr Leu Lys Lys Ser Asn Asn Leu Ser Lys Arg Cys Tyr
405 410 415

Asp Asp Gln Leu Ile Glu Glu Leu Phe Thr Leu Ser Tyr Thr Lys Phe
420 425 430

Gly His Leu Ser Phe Lys Ala Ile Asn His Val Leu Pro Ile Met Gln
435 440 445

Glu Gly Arg Thr Tyr Gln Glu Ala Ile His Gln Leu Gly Tyr Asp Thr
450 455 460

Thr Asn Leu Lys Lys Glu Asn Arg Ser Met Phe Leu Pro Leu Ile Pro
465 470 475 480

Asp Glu Ile Thr Asn Pro Ile Val Lys Arg Ala Ile Thr Gln Ala Arg
485 490 495

Lys Val Val Asn Ala Ile Ile Arg Arg Tyr Gly Ser Pro Asn Ser Val
500 505 510

His Ile Glu Leu Ala Arg Glu Leu Ser Lys Ser His Asp Glu Arg Lys
515 520 525

Lys Ile Met Thr Ala His Asp Glu Asn Tyr Lys Lys Asn Lys Gly Ala
530 535 540

Ile Ser Ile Leu Ile Glu Asn Gly Ile Leu Asn Pro Thr Gly Tyr Asp
545 550 555 560

Ile Val Arg Tyr Lys Leu Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr

565

570

575

Ser Leu Lys Glu Ile Pro Pro Asp Thr Phe Phe Asn Glu Leu Lys Lys
 580 585 590

Glu Arg Asn Gly Ser Pro Ile Leu Glu Val Asp His Ile Leu Pro Tyr
 595 600 605

Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Ser
 610 615 620

Asp Glu Asn Arg Asn Lys Gly Asn Arg Ile Pro Tyr Thr Tyr Phe Leu
 625 630 635 640

Glu Thr Asn Lys Asp Trp Glu Ala Phe Glu Arg Tyr Val Arg Ser Asn
 645 650 655

Lys Leu Phe Ser Lys Lys Lys Arg Glu Tyr Leu Leu Lys Lys Thr Tyr
 660 665 670

Leu Pro Arg Glu Ser Glu Leu Ile Lys Glu Arg His Leu Asn Asp Thr
 675 680 685

Arg Tyr Ala Ser Thr Phe Leu Lys Asn Phe Ile Glu Gln Asn Leu Gln
 690 695 700

Phe Lys Glu Val Glu Val Asn Leu Arg Lys Lys Arg Val Gln Thr Val
 705 710 715 720

Asn Gly Val Ile Thr Ala His Leu Arg Lys Arg Trp Gly Leu Glu Lys
 725 730 735

Asn Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
 740 745 750

Ala Cys Thr Asp His His Met Val Thr Arg Ile Thr Glu Tyr Tyr Gln
 755 760 765

Ile Lys Glu Ser Asn Lys Ser Val Lys Lys Pro Tyr Phe Pro Met Pro
 770 775 780

Trp Glu Gly Phe Arg Asp Glu Leu Leu Ser His Leu Ala Ser Gln Pro
785 790 795 800

Ile Ala Lys Lys Ile Ser Glu Glu Leu Lys Ala Gly Tyr Gln Ser Ser
805 810 815

Asp Tyr Ile Phe Val Ser Arg Met Pro Lys Arg Ser Val Thr Gly Ala
820 825 830

Ala His Asp Gln Thr Ile Arg Arg Lys Gly Gly Ile Asp Lys Lys Gly
835 840 845

Lys Thr Ile Ile Ile Lys Arg Val Arg Leu Lys Asp Ile Lys Phe Asp
850 855 860

Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Leu Ala Thr
865 870 875 880

Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Glu His Arg Lys Asn Ser Lys
885 890 895

Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Thr Gly
900 905 910

Asn Leu Ile Lys Arg Val Lys Ile Glu Gly Gln Thr Lys Ala Phe Val
915 920 925

Arg Glu Val Asn Gly Gly Val Ala Gln Asn Ser Asp Leu Val Arg Val
930 935 940

Asp Leu Phe Glu Lys Asp Asp Lys Tyr Tyr Met Val Pro Ile Tyr Val
945 950 955 960

Pro Asp Thr Val Cys Ser Glu Leu Pro Lys Lys Val Val Lys Ser Gly
965 970 975

Lys Gly Tyr Glu Gln Trp Leu Thr Leu Asp Asn Ser Phe Thr Phe Lys
980 985 990

Ser Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asn Glu

995

1000

1005

Asp Arg Phe Leu Tyr Phe Gly Thr Leu Asp Ile Asp Ser Asp Arg
1010 1015 1020

Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Gln Asn Glu Tyr
1025 1030 1035

Arg Tyr Ser Leu Lys Thr Ile Glu Asn Leu Glu Lys Tyr Glu Val
1040 1045 1050

Gly Val Leu Gly Asp Leu Arg Leu Val Lys Gln Glu Thr Arg Arg
1055 1060 1065

Ile Phe Asn Arg
1070

<210> 104
<211> 1337
<212> PRT
<213> Enterococcus sp.

<223> APG05459.1_RGN

<400> 104

Met Lys Lys Asp Tyr Val Ile Gly Leu Asp Ile Gly Thr Asn Ser Val
1 5 10 15

Gly Trp Ala Val Met Thr Glu Asp Tyr Gln Leu Val Lys Lys Lys Met
20 25 30

Pro Ile Tyr Gly Asn Thr Glu Lys Lys Lys Ile Lys Lys Asn Phe Trp
35 40 45

Gly Val Arg Leu Phe Glu Glu Gly His Thr Ala Glu Asp Arg Arg Leu
50 55 60

Lys Arg Thr Ala Arg Arg Arg Ile Ser Arg Arg Arg Asn Arg Leu Arg
65 70 75 80

Tyr Leu Gln Ala Phe Phe Glu Glu Ala Met Thr Ala Leu Asp Glu Asn
85 90 95

Phe Phe Ala Arg Leu Gln Glu Ser Phe Leu Val Pro Glu Asp Lys Lys
100 105 110

Trp His Arg His Pro Ile Phe Ala Lys Leu Glu Asp Glu Val Ala Tyr
115 120 125

His Glu Thr Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Ala Asp
130 135 140

Ser Ser Glu Gln Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145 150 155 160

Ile Val Lys Tyr Arg Gly His Phe Leu Ile Glu Gly Lys Leu Ser Thr
165 170 175

Glu Asn Ile Ser Val Lys Glu Gln Phe Gln Gln Phe Met Ile Ile Tyr
180 185 190

Asn Gln Thr Phe Val Asn Gly Glu Ser Arg Leu Val Ser Ala Pro Leu
195 200 205

Pro Glu Ser Val Leu Ile Glu Glu Glu Leu Thr Glu Lys Ala Ser Arg
210 215 220

Thr Lys Lys Ser Glu Lys Val Leu Gln Gln Phe Pro Gln Glu Lys Ala
225 230 235 240

Asn Gly Leu Phe Gly Gln Phe Leu Lys Leu Met Val Gly Asn Lys Ala
245 250 255

Asp Phe Lys Lys Val Phe Gly Leu Glu Glu Glu Ala Lys Ile Thr Tyr
260 265 270

Ala Ser Glu Ser Tyr Glu Glu Asp Leu Glu Gly Ile Leu Ala Lys Val
275 280 285

Gly Asp Glu Tyr Ser Asp Val Phe Leu Ala Ala Lys Asn Val Tyr Asp
290 295 300

Ala Val Glu Leu Ser Thr Ile Leu Ala Asp Ser Asp Lys Lys Ser His
305 310 315 320

Ala Lys Leu Ser Ser Ser Met Ile Val Arg Phe Thr Glu His Gln Glu
325 330 335

Asp Leu Lys Lys Phe Lys Arg Phe Ile Arg Glu Asn Cys Pro Asp Glu
340 345 350

Tyr Asp Asn Leu Phe Lys Asn Glu Gln Lys Asp Gly Tyr Ala Gly Tyr
355 360 365

Ile Ala His Ala Gly Lys Val Ser Gln Leu Lys Phe Tyr Gln Tyr Val
370 375 380

Lys Lys Ile Ile Gln Asp Ile Ala Gly Ala Glu Tyr Phe Leu Glu Lys
385 390 395 400

Ile Ala Gln Glu Asn Phe Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly
405 410 415

Val Ile Pro His Gln Ile His Leu Ala Glu Leu Gln Ala Ile Ile His
420 425 430

Arg Gln Ala Ala Tyr Tyr Pro Phe Leu Lys Glu Asn Gln Glu Lys Ile
435 440 445

Glu Gln Leu Val Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ser
450 455 460

Lys Gly Asp Ala Ser Thr Phe Ala Trp Leu Lys Arg Gln Ser Glu Glu
465 470 475 480

Pro Ile Arg Pro Trp Asn Leu Gln Glu Thr Val Asp Leu Asp Gln Ser
485 490 495

Ala Thr Ala Phe Ile Glu Arg Met Thr Asn Phe Asp Thr Tyr Leu Pro
500 505 510

Ser Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Lys Phe Met
515 520 525

Val Phe Asn Glu Leu Thr Lys Ile Ser Tyr Thr Asp Asp Arg Gly Ile
530 535 540

Lys Ala Asn Phe Ser Gly Lys Glu Lys Glu Lys Ile Phe Asp Tyr Leu
545 550 555 560

Phe Lys Thr Arg Arg Lys Val Lys Lys Lys Asp Ile Ile Gln Phe Tyr
565 570 575

Arg Asn Glu Tyr Asn Thr Glu Ile Val Thr Leu Ser Gly Leu Glu Glu
580 585 590

Asp Gln Phe Asn Ala Ser Phe Ser Thr Tyr Gln Asp Leu Leu Lys Cys
595 600 605

Gly Leu Thr Arg Ala Glu Leu Asp His Pro Asp Asn Ala Glu Lys Leu
610 615 620

Glu Asp Ile Ile Lys Ile Leu Thr Ile Phe Glu Asp Arg Gln Arg Ile
625 630 635 640

Arg Thr Gln Leu Ser Thr Phe Lys Gly Gln Phe Ser Ala Glu Val Leu
645 650 655

Lys Lys Leu Glu Arg Lys His Tyr Thr Gly Trp Gly Arg Leu Ser Lys
660 665 670

Lys Leu Ile Asn Gly Ile Tyr Asp Lys Glu Ser Gly Lys Thr Ile Leu
675 680 685

Asp Tyr Leu Ile Lys Asp Asp Gly Val Ser Lys His Tyr Asn Arg Asn
690 695 700

Phe Met Gln Leu Ile Asn Asp Ser Gln Leu Ser Phe Lys Asn Ala Ile
705 710 715 720

Gln Lys Ala Gln Ser Ser Glu His Glu Glu Thr Leu Ser Glu Thr Val
725 730 735

Asn Glu Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Tyr Gln Ser
740 745 750

Leu Lys Ile Val Asp Glu Leu Val Ala Ile Met Gly Tyr Ala Pro Lys
755 760 765

Arg Ile Val Val Glu Met Ala Arg Glu Asn Gln Thr Thr Ser Thr Gly
770 775 780

Lys Arg Arg Ser Ile Gln Arg Leu Lys Ile Val Glu Lys Ala Met Ala
785 790 795 800

Glu Ile Gly Ser Asn Leu Leu Lys Glu Gln Pro Thr Thr Asn Glu Gln
805 810 815

Leu Arg Asp Thr Arg Leu Phe Leu Tyr Tyr Met Gln Asn Gly Lys Asp
820 825 830

Met Tyr Thr Gly Asp Glu Leu Ser Leu His Arg Leu Ser His Tyr Asp
835 840 845

Ile Asp His Ile Ile Pro Gln Ser Phe Met Lys Asp Asp Ser Leu Asp
850 855 860

Asn Leu Val Leu Val Gly Ser Thr Glu Asn Arg Gly Lys Ser Asp Asp
865 870 875 880

Val Pro Ser Lys Glu Val Val Lys Asp Met Lys Ala Tyr Trp Glu Lys
885 890 895

Leu Tyr Ala Ala Gly Leu Ile Ser Gln Arg Lys Phe Gln Arg Leu Thr
900 905 910

Lys Gly Glu Gln Gly Gly Leu Thr Leu Glu Asp Lys Ala His Phe Ile
915 920 925

Gln Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys Asn Val Ala Gly
930 935 940

Ile Leu Asp Gln Arg Tyr Asn Ala Asn Ser Lys Glu Lys Lys Val Gln
945 950 955 960

Ile Ile Thr Leu Lys Ala Ser Leu Thr Ser Gln Phe Arg Ser Ile Phe
965 970 975

Gly Leu Tyr Lys Val Arg Glu Val Asn Asp Tyr His His Gly Gln Asp
980 985 990

Ala Tyr Leu Asn Cys Val Val Ala Thr Thr Leu Leu Lys Val Tyr Pro
995 1000 1005

Asn Leu Ala Pro Glu Phe Val Tyr Gly Glu Tyr Pro Lys Phe Gln
1010 1015 1020

Ala Phe Lys Glu Asn Lys Ala Thr Ala Lys Thr Ile Ile Tyr Thr
1025 1030 1035

Asn Leu Met Arg Phe Phe Thr Glu Asp Glu Pro Arg Phe Met Lys
1040 1045 1050

Asp Gly Glu Ile Leu Trp Ser Asn Ser Tyr Leu Lys Asn Ile Lys
1055 1060 1065

Lys Glu Leu Asn Tyr His Gln Met Asn Ile Val Lys Lys Val Glu
1070 1075 1080

Val Gln Lys Gly Gly Phe Ser Lys Glu Ser Ile Lys Pro Lys Gly
1085 1090 1095

Pro Ser Asn Lys Leu Ile Pro Val Lys Asn Gly Leu Asp Pro Gln
1100 1105 1110

Lys Tyr Gly Gly Phe Asp Ser Pro Val Val Ala Tyr Thr Val Leu
1115 1120 1125

Phe Thr His Glu Lys Gly Lys Lys Pro Leu Ile Lys Gln Glu Ile
1130 1135 1140

Leu Gly Ile Thr Ile Met Glu Lys Thr Arg Phe Glu Gln Asn Pro
1145 1150 1155

Ile Leu Phe Leu Glu Glu Lys Gly Phe Leu Arg Pro Arg Val Leu
1160 1165 1170

Met Lys Leu Pro Lys Tyr Thr Leu Tyr Glu Phe Pro Glu Gly Arg
1175 1180 1185

Arg Arg Leu Leu Ala Ser Ala Lys Glu Ala Gln Lys Gly Asn Gln
1190 1195 1200

Met Val Leu Pro Glu His Leu Leu Thr Leu Leu Tyr His Ala Lys
1205 1210 1215

Gln Cys Leu Leu Pro Asn Gln Ser Glu Ser Leu Ala Tyr Val Glu
1220 1225 1230

Gln His Gln Pro Glu Phe Gln Glu Ile Leu Glu Arg Val Val Asp
1235 1240 1245

Phe Ala Glu Val His Thr Leu Ala Lys Ser Lys Val Gln Gln Ile
1250 1255 1260

Val Lys Leu Phe Glu Ala Asn Gln Thr Ala Asp Val Lys Glu Ile
1265 1270 1275

Ala Ala Ser Phe Ile Gln Leu Met Gln Phe Asn Ala Met Gly Ala
1280 1285 1290

Pro Ser Thr Phe Lys Phe Phe Gln Lys Asp Ile Glu Arg Ala Arg
1295 1300 1305

Tyr Thr Ser Ile Lys Glu Ile Phe Asp Ala Thr Ile Ile Tyr Gln
1310 1315 1320

Ser Thr Thr Gly Leu Tyr Glu Thr Arg Arg Lys Val Val Asp
1325 1330 1335

<210> 105

<211> 1436

<212> PRT

<213> Enterococcus sp.

<223> APG04583.1_RGN

<400> 105

Met Ala Lys Asn Ile Leu Gly Leu Asp Leu Gly Thr Asn Ser Ile Gly
1 5 10 15

Trp Ala Leu Val Gln Gln Asp Phe Glu Asn Lys Glu Gly Asn Ile Leu
20 25 30

Gly Met Gly Ser Arg Ile Ile Pro Met Ser Gln Asp Ile Leu Gly Glu
35 40 45

Phe Gly Lys Gly Asn Ser Ile Ser Gln Thr Ala Glu Arg Thr Gly Tyr
50 55 60

Arg Gly Val Arg Arg Leu Arg Glu Arg His Leu Leu Arg Arg Glu Arg
65 70 75 80

Leu His Arg Val Leu His Leu Leu Gly Phe Leu Pro Lys His Tyr Asp
85 90 95

Glu Lys Ile Asp Phe Thr Gln Arg Phe Gly Lys Phe Ile Asn Gln Ala
100 105 110

Glu Pro Lys Leu Ala Phe Asp Ser Glu Phe Leu Phe Lys Asp Ser Phe
115 120 125

His Glu Met Leu Ala Asp Phe Lys Gln Asn Gln Pro Glu Phe Leu Lys
130 135 140

Asp Lys Asn Gly Glu Asp Cys Leu Val Pro Tyr Asp Trp Thr Ile Tyr
145 150 155 160

Tyr Leu Arg Lys Lys Ala Leu Thr Gln Lys Ile Glu Lys Tyr Glu Leu
165 170 175

Ala Trp Leu Ile Leu Asn Phe Asn Gln Lys Arg Gly Tyr Tyr Gln Leu
180 185 190

Arg Gly Glu Glu Glu Lys Glu Asn Pro Asn Lys Leu Val Gly Phe His
195 200 205

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

Gly Glu Thr Trp Tyr Ser Leu His Leu Glu Asn Gly Trp Val Tyr Arg
225 230 235 240

Arg Ser Ser Lys Ile Ser Leu Ala Asp Trp Lys Asp Lys Val Arg Asp
245 250 255

Phe Ile Val Thr Thr Asp Leu Asn Asp Asp Gly Ser Glu Lys Leu Asp
260 265 270

Lys Asp Gly Ile Val Lys Arg Ser Phe Arg Ala Pro Ser Ala Asp Asp
275 280 285

Trp Thr Leu Leu Lys Lys Lys Thr Glu Lys Asp Ile Asp Asn Ser Asn
290 295 300

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

Gln Lys Ile Lys Gly Lys Met Val Arg Thr Ile Glu Arg Lys Phe Tyr
325 330 335

Lys Gln Glu Leu Glu Gln Ile Leu Lys Thr Gln Lys Glu Phe His Ser
340 345 350

Glu Leu Gln Ser Glu Asn Leu Leu Gln Asp Cys Val Arg Glu Leu Tyr
355 360 365

Arg Asn Asn Glu Gln His Gln Gln Met Leu Glu Ala Lys Asp Phe Val
370 375 380

His Leu Phe Leu Asn Asp Ile Ile Phe Tyr Gln Arg Pro Leu Arg Ser
385 390 395 400

Gln Lys Ser Ser Ile Ser Asn Cys Thr Leu Glu Phe Arg Lys Ser Lys
405 410 415

Asn Glu Asn Gly Glu Glu Val Ile His Arg Leu Lys Val Ile Ala Lys

420

425

430

Ser Asn Pro Tyr Tyr Gln Glu Phe Arg Leu Leu Gln Trp Val Gln Asn
 435 440 445

Leu Ala Ile Tyr Thr Lys Asp Asp Asp Lys Asn Val Thr Asn Glu Phe
 450 455 460

Leu Lys Ser Thr Gln Asp Trp Glu Asp Leu Leu Arg Trp Leu His Ser
 465 470 475 480

Lys Lys Glu Ile Lys Gln Asp Ala Leu Ile Lys Phe Leu Ile Glu Lys
 485 490 495

Lys Gly Leu Lys Gly Lys Ala Leu Thr Ile Glu Val Ala Lys Tyr Arg
 500 505 510

Trp Asn Tyr Val Gln Asp Lys Asp Tyr Pro Gly Asn Glu Thr Arg Tyr
 515 520 525

Leu Ile Gln Ser Arg Leu Asp Lys Val Glu Tyr Ala Pro Lys Asp Phe
 530 535 540

Leu Thr Tyr Glu Asn Glu Met Ala Leu Trp His Ile Ile Tyr Ser Ile
 545 550 555 560

Asn Asp Lys Ile Glu Tyr Glu Lys Ala Leu Lys Ser Phe Ala Asn Lys
 565 570 575

Lys Gly Leu Asp Glu Val Thr Phe Val Glu Ala Phe Lys Lys Phe Pro
 580 585 590

Pro Phe Lys Ser Asp Tyr Gly Ser Phe Ser Glu Lys Ala Ile Lys Lys
 595 600 605

Leu Leu Pro Leu Met Arg Phe Gly Thr Gln Trp Asn Trp Asp Asn Ile
 610 615 620

Asp Gln Asn Ser Lys Glu Arg Ile Gly Lys Ile Leu Thr Gly Glu Tyr
 625 630 635 640

Asp Glu Asn Ile Lys Gly Arg Val Arg Glu Lys Ala Arg His Leu Asn
645 650 655

Ser Glu Thr Asp Phe Gln Ala Leu Pro Leu Trp Leu Ala Gln Tyr Val
660 665 670

Val Tyr Gly Arg His Ser Glu Ala Asp Ile Ala Gly Lys Trp Asn Ser
675 680 685

Val Asp Asp Leu Lys Gln Phe Leu Asp Asp Phe Lys Gln His Ser Leu
690 695 700

Arg Asn Pro Ile Val Glu Gln Val Ile Thr Glu Thr Leu Arg Ala Val
705 710 715 720

Lys Asp Ile Trp Asn Phe Tyr Gly Lys Gly Ala Lys Asp Phe Phe Ser
725 730 735

Glu Ile His Ile Glu Leu Gly Arg Glu Met Lys Asn Thr Ala Asp Glu
740 745 750

Arg Lys Arg Ile Thr Thr Met Val Thr Asp Asn Glu Asn Thr Asn Leu
755 760 765

Arg Ile Lys Ala Leu Leu Ala Glu Met Ala Leu Asp Gln Asn Val Asp
770 775 780

Asn Val Arg Pro Tyr Ser Pro Met Gln Gln Glu Ile Leu Lys Ile Tyr
785 790 795 800

Glu Glu Gly Val Leu Asn Ala Glu Glu Asn Ile Asp Asp Asp Ile Leu
805 810 815

Lys Ile Ser Lys Thr Ala Gln Pro Ser Ala Thr Asp Leu Lys Arg Tyr
820 825 830

Lys Leu Trp Leu Glu Gln Lys Tyr Arg Ser Pro Tyr Thr Gly Gln Met
835 840 845

Ile Pro Leu Asn Lys Leu Phe Thr Pro Glu Tyr Glu Ile Glu His Ile

850

855

860

Ile Pro Gln Ser Arg Tyr Phe Asp Asp Ser Met Ser Asn Lys Val Ile
 865 870 875 880

Cys Glu Ala Ala Val Asn Lys Leu Lys Asp Asn Gln Ile Gly Leu Val
 885 890 895

Phe Ile Lys Asn His His Gly Glu Val Val Asp Phe Gly Met Gly Lys
 900 905 910

Gln Val Lys Ile Leu Glu Val Ser Asp Tyr Glu Asp Phe Val Lys Gln
 915 920 925

Asn Tyr Asn Lys Asn Arg Gly Lys Arg Asn Lys Leu Leu Leu Glu Asp
 930 935 940

Ile Pro Glu Lys Met Ile Glu Arg Gln Leu Asn Asp Thr Arg Tyr Ile
 945 950 955 960

Ser Lys Tyr Ile Thr Gln Val Leu Ser Asn Ile Val Arg Asp Asp Lys
 965 970 975

Glu Gly Ser Lys Asp Asp Gly Val Asn Ser Lys Asn Ile Val Pro Gly
 980 985 990

Asn Gly Lys Ile Thr Thr Arg Leu Lys Gln Asp Trp Gly Leu Asn Asp
 995 1000 1005

Val Trp Asn Asp Leu Val Leu Pro Arg Phe Glu Arg Met Asn Thr
 1010 1015 1020

Leu Thr Asn Ser Asn Asp Phe Thr Ser Lys Asn Thr His Gly Lys
 1025 1030 1035

Thr Ile Pro Thr Val Pro Ile Glu Leu Ser Lys Gly Phe Ser Lys
 1040 1045 1050

Lys Arg Ile Asp His Arg His His Ala Met Asp Ala Leu Val Ile
 1055 1060 1065

Ala Cys Ala Thr Arg Asp His Val Asn Leu Leu Asn Asn Glu Ser
1070 1075 1080

Ser Lys Ser Asp Thr Lys Arg Tyr Asp Leu Asn Arg Lys Leu Arg
1085 1090 1095

Lys Tyr Glu Lys Val Ala Tyr Asn Asp Pro Lys Thr Gly Glu Arg
1100 1105 1110

Ile Glu Lys Glu Val Pro Lys Asp Phe Ile Lys Pro Trp Glu Thr
1115 1120 1125

Phe Thr Glu Asp Thr Arg Thr Leu Leu Glu Asn Ile Val Ile Ser
1130 1135 1140

Phe Lys Gln Asn Leu Arg Val Ile Asn Lys Ala Thr Asn Tyr Tyr
1145 1150 1155

Glu Lys Ile Glu Asn Gly Lys Lys Val Lys Val Glu Gln Lys Gly
1160 1165 1170

Ile Asn Trp Ala Val Arg Lys Ala Leu His Lys Glu Thr Val Ser
1175 1180 1185

Gly Gln Val His Leu Asp Arg Ile Lys Val Ala Lys Gly Lys Ile
1190 1195 1200

Leu Thr Ala Thr Arg Lys Thr Leu Asp Ala Ser Phe Asn Glu Lys
1205 1210 1215

Thr Ile Glu Ser Ile Thr Asp Thr Gly Ile Gln Lys Ile Leu Leu
1220 1225 1230

Asn Tyr Leu Lys Ser Lys Asp Asn Asn Pro Glu Val Ala Phe Ser
1235 1240 1245

Pro Glu Gly Ile Glu Glu Leu Asn Lys Asn Ile Arg Leu Tyr Asn
1250 1255 1260

Asp Gly Lys Ala His Gln Pro Ile Leu Lys Val Arg Val Phe Glu

1265

1270

1275

Gln Gly Ser Lys Phe Thr Leu Gly Glu Thr Gly Asn Lys Thr Thr
 1280 1285 1290

Lys Phe Val Glu Ala Ala Lys Gly Thr Asn Leu Phe Phe Gly Ile
 1295 1300 1305

Tyr Glu Asp Lys Ser Gly Lys Arg Ser Tyr Glu Thr Ile Pro Leu
 1310 1315 1320

Asn Ile Val Ile Glu Arg Gln Lys Gln Gly Leu Gln Ala Val Pro
 1325 1330 1335

Glu Thr Asn Glu Lys Gly Lys Gln Leu Leu Phe Thr Leu Ser Pro
 1340 1345 1350

Asn Asp Leu Val Tyr Val Pro Glu Glu Gly Val Phe Asp Glu Asn
 1355 1360 1365

Asn Ile Lys Val Asp Arg Ile Tyr Lys Val Val Ser Phe Ser Thr
 1370 1375 1380

Tyr Gln Cys Phe Phe Val Arg Asn Asp Val Ser Thr Ser Val Val
 1385 1390 1395

Asn Lys Val Glu Tyr Ser Ala Leu Asn Lys Met Glu Lys Ser Ile
 1400 1405 1410

Asp Asn Ile Met Ile Lys Glu Asn Cys Val Lys Leu Asn Val Asp
 1415 1420 1425

Arg Leu Gly Lys Ile Ser Lys Ala
 1430 1435

<210> 106

<211> 1456

<212> PRT

<213> Enterococcus sp.

<223> APG01688.1_RGN

<400> 106

Met Met Ile Lys Asn Ile Leu Gly Leu Asp Leu Gly Thr Asn Ser Ile
1 5 10 15

Gly Trp Ala Leu Ile Lys Gln Asp Phe Glu Asn Lys His Gly Glu Ile
20 25 30

Leu Gly Met Gly Ser Arg Ile Ile Pro Met Ser Gln Asp Ile Leu Gly
35 40 45

Asp Phe Gly Lys Gly Asn Ser Ile Ser Gln Thr Ala Asp Arg Thr Lys
50 55 60

Tyr Arg Ser Val Arg Arg Leu Arg Glu Arg Phe Leu Leu Arg Arg Glu
65 70 75 80

Arg Leu His Arg Val Leu His Leu Leu Asn Phe Leu Pro Gln His Tyr
85 90 95

Ala Ser Gln Ile Asp Phe Glu Lys Lys Phe Gly Lys Phe Lys Ser Glu
100 105 110

Thr Glu Pro Lys Leu Ala Trp Glu Asn Trp Gly Gly Lys Phe Ser Phe
115 120 125

Leu Phe Gln Asn Ser Phe Asn Glu Met Leu Glu Asp Phe Lys Ala Ala
130 135 140

Gly Gln Gly Leu Lys Ile Pro Tyr Asp Trp Thr Ile Tyr Tyr Leu Arg
145 150 155 160

Lys Lys Ala Leu Ser Gln Lys Ile Glu Lys Glu Glu Leu Ala Trp Ile
165 170 175

Leu Leu Asn Phe Asn Gln Lys Arg Gly Tyr Tyr Gln Leu Arg Gly Glu
180 185 190

Glu Glu Glu Glu Asn Pro Asn Lys Leu Val Glu Phe Tyr Ser Leu Lys
195 200 205

Ile Val Asp Val Val Ala Asp Glu Pro Gln Lys Gly Lys Ser Asp Ile
210 215 220

Trp Tyr Ser Leu Ile Leu Glu Asn Gly Trp Val Tyr Arg Arg Ala Ser
225 230 235 240

Lys Ile Pro Leu Phe Asp Trp Lys Asp Lys Thr Arg Asp Phe Ile Val
245 250 255

Thr Thr Asp Leu Asn Asp Asp Arg Ser Val Lys Thr Asp Lys Glu Gly
260 265 270

Asn Glu Lys Arg Ser Phe Arg Ala Pro Ser Glu Asn Asp Trp Thr Leu
275 280 285

Val Lys Lys Lys Thr Glu Gln Glu Ile Asp Gln Ser His Lys Thr Val
290 295 300

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

Lys Gly Lys Leu Val Arg Thr Ile Glu Arg Lys Phe Tyr Lys Asp Glu
325 330 335

Leu Lys Gln Ile Leu Glu Lys Gln Lys Glu Phe His Gln Glu Leu Lys
340 345 350

Asn Asp Asp Leu Tyr Asn Asp Cys Ile Arg Glu Leu Tyr Arg Asn Asn
355 360 365

Glu Ala His Gln Leu Thr Leu Ser Lys Lys Asp Phe Val His Leu Leu
370 375 380

Met Asp Asp Leu Ile Phe Tyr Gln Arg Pro Leu Arg Ser Gln Lys Ser
385 390 395 400

Ser Ile Ser Asn Cys Thr Leu Glu Phe Arg Lys Tyr Lys Asp Glu Asn
405 410 415

Gly Ile Glu His Thr Gln Tyr Leu Lys Ala Ile Pro Lys Ser Asn Pro
420 425 430

Tyr Tyr Gln Glu Phe Arg Leu Trp Gln Trp Met Tyr Asn Leu Asn Ile
435 440 445

Tyr Arg Lys Asp Asp Glu Ala Asn Val Thr Lys Glu Phe Leu Asn Thr
450 455 460

Asn Lys Asp Phe Glu Ser Leu Phe Glu Phe Leu Asn Asn Arg Lys Glu
465 470 475 480

Ile Glu Gln Lys Pro Leu Ile Lys Phe Leu Leu Glu Gln Lys Asp Ile
485 490 495

Asn Lys Lys Leu Leu Asn Ala Glu Ala Glu Lys Tyr Arg Trp Asn Tyr
500 505 510

Val Glu Asp Lys Lys Tyr Pro Cys Asn Glu Thr Lys Thr Met Ile Ser
515 520 525

Ser Arg Leu Asp Lys Val Glu Asn Ile Ser Asp Asp Phe Leu Thr Arg
530 535 540

Asp Ile Glu Gln Lys Ile Trp His Ile Ile Tyr Ser Val Asn Asp Lys
545 550 555 560

Ile Glu Tyr Glu Lys Ala Leu Lys Ser Phe Ala Thr Arg Asn Asp Leu
565 570 575

Asp Glu Asn Ser Phe Ile Glu Ala Phe Lys Lys Phe Ser Pro Phe Lys
580 585 590

Ser Glu Tyr Gly Ser Phe Ser Glu Lys Ala Ile Lys Lys Leu Leu Pro
595 600 605

Leu Met Arg Leu Gly Lys Tyr Trp Tyr Glu Asp Glu Ile Val Lys His
610 615 620

Ser Asp Ile Tyr Phe Lys Asn Ile Glu Asn Leu Leu Gly Asp Phe Ser
625 630 635 640

Asn Arg Asp Lys Lys Ile Ser Glu Glu Asp Lys Glu Lys Trp Asn Lys
645 650 655

Ser Ile Asn Leu Lys Leu Gln Glu Glu Leu Lys Asp Phe Gln Thr Ala
660 665 670

Glu Ile Asp Leu Phe Gln Gly Leu Arg Leu His Ile Ala Gln Tyr Leu
675 680 685

Val Tyr Gly Arg His Ser Glu Ala Ser Met Ile Gly Lys Trp Asn Ser
690 695 700

Ala Glu Asp Leu Glu Glu Phe Leu Lys Asp Phe Lys Gln His Ser Leu
705 710 715 720

Arg Asn Pro Ile Val Glu Gln Val Ile Thr Glu Thr Leu Arg Val Val
725 730 735

Lys Asp Ile Trp Leu Lys Tyr Gly Asn Gly Ala Lys Asp Phe Phe Asn
740 745 750

Glu Ile His Ile Glu Leu Gly Arg Glu Met Lys Leu Pro Ala Asp Asp
755 760 765

Arg Lys Lys Leu Thr Asn Gln Ile Ser Glu Asn Glu Asn Thr Asn Phe
770 775 780

Arg Ile Lys Ala Leu Leu Ala Glu Met Met Asn Asp Ser Ser Val Glu
785 790 795 800

Asn Val Arg Pro Phe Ser Pro Met Gln Gln Glu Ile Leu Lys Ile Tyr
805 810 815

Glu Asp Asp Val Leu Lys Ser Asp Ile Glu Ile Glu Asp Asp Ile Leu
820 825 830

Lys Ile Ser Lys Thr Ala Gln Pro Ser Pro Ser Asp Leu Lys Arg Tyr
835 840 845

Lys Leu Trp Leu Glu Gln Lys Tyr Lys Ser Pro Tyr Thr Gly Gln Ile
850 855 860

Ile Pro Leu Asn Lys Leu Phe Thr Pro Glu Tyr Glu Ile Glu His Ile
865 870 875 880

Ile Pro Gln Ser Arg Tyr Phe Asp Asp Ser Phe Ser Asn Lys Val Ile
885 890 895

Cys Glu Ser Ala Val Asn Lys Leu Lys Asp Asn Tyr Ile Gly Leu Glu
900 905 910

Phe Ile Lys Gln Phe Gly Gly Thr Ile Ile Glu Leu Gly Phe Gly Lys
915 920 925

Ser Ile Lys Val Phe Glu Thr Lys Glu Tyr Glu Asp Phe Val Lys Lys
930 935 940

His Tyr Ala Asn Asn Gln Gly Lys Arg Asn Lys Leu Leu Met Glu Asp
945 950 955 960

Ile Pro Glu Lys Met Ile Glu Arg Gln Met Asn Asp Thr Arg Tyr Ile
965 970 975

Ser Lys Tyr Ile Ser Gly Val Leu Ser Asn Ile Val Arg Val Glu Asp
980 985 990

Gly Ser Asp Glu Gly Val Asn Ser Lys Asn Ile Val Pro Gly Asn Gly
995 1000 1005

Lys Ile Thr Thr Gln Leu Lys Gln Asp Trp Gly Leu Asn Asp Val
1010 1015 1020

Trp Asn Asp Leu Ile Leu Pro Arg Phe Glu Arg Met Asn Gln Leu
1025 1030 1035

Thr Asn Ser Lys Val Phe Thr Ala Trp Asn Glu Asn Tyr Gln Lys
1040 1045 1050

Phe Leu Pro Thr Val Pro Ile Glu Tyr Ser Lys Gly Phe Ser Lys
1055 1060 1065

Lys Arg Ile Asp His Arg His His Ala Leu Asp Ala Leu Val Ile
1070 1075 1080

Ala Cys Ala Thr Lys Asp His Val Asn Leu Leu Asn Asn Gln Ser
1085 1090 1095

Ala Lys Ser Asp Thr Lys Arg Tyr Asp Leu Lys Lys Lys Ser Met
1100 1105 1110

Lys Phe Glu Lys Val Val Tyr Asn Asp Ala Lys Thr Gly Glu Lys
1115 1120 1125

Ile Glu Arg Glu Val Pro Lys Gln Phe Leu Lys Pro Trp Glu Asn
1130 1135 1140

Phe Thr Leu Asp Val Lys His Asn Leu Glu Thr Ile Ile Val Ser
1145 1150 1155

Phe Lys Gln Asn Leu Arg Val Ile Asn Lys Ala Thr Asn Tyr Tyr
1160 1165 1170

Glu Lys Tyr Val Glu Lys Asp Gly Thr Lys Asn Lys Glu Arg Val
1175 1180 1185

Glu Gln Thr Gly Thr Asn Trp Ala Ile Arg Lys Pro Met His Lys
1190 1195 1200

Asp Thr Val Ser Gly Lys Val Asp Leu Pro Trp Val Lys Val Pro
1205 1210 1215

Lys Gly Lys Ile Leu Thr Ala Thr Arg Lys Ser Leu Asp Ser Ser
1220 1225 1230

Phe Asp Leu Lys Ser Ile Gly Ser Ile Thr Asp Thr Gly Ile Gln
1235 1240 1245

Lys Ile Leu Lys Asn Tyr Leu Ala Phe Lys Asp Gly Asn Pro Glu
1250 1255 1260

Leu Ala Phe Ser Pro Glu Gly Ile Asp Asp Leu Asn Lys Asn Ile
1265 1270 1275

Glu Lys Tyr Asn Asp Gly Lys Pro His Gln Pro Ile Asn Lys Val
1280 1285 1290

Arg Val Phe Glu Leu Gly Ser Lys Phe Gln Val Gly Gln Ser Gly
1295 1300 1305

Asn Lys Lys Asp Lys Tyr Val Glu Ala Ala Lys Gly Thr Asn Leu
1310 1315 1320

Phe Phe Ala Val Tyr Glu Asp Glu Lys Gly Lys Arg Asn Tyr Glu
1325 1330 1335

Thr Ile Pro Leu Asn Glu Val Ile Glu Arg Gln Lys Gln Gly Leu
1340 1345 1350

Ser Val Val Asp Leu Lys Gly Thr Asn Asp Phe Tyr Leu Cys Pro
1355 1360 1365

Asn Asp Phe Val Tyr Ile Pro Ser Gly Asp Glu Leu Glu Asn Ile
1370 1375 1380

Asn Asn Val Asp Phe Lys Asp Ile Lys Lys Glu Ile Asn Glu Arg
1385 1390 1395

Ile Tyr Lys Val Val Ser Phe Thr Gly Asn Arg Leu Ser Cys Ile
1400 1405 1410

Pro Tyr Met Val Ala Thr Thr Ile Val Asn Lys Leu Glu Phe Thr
1415 1420 1425

Gln Leu Asn Lys Ile Glu Phe Thr Lys Glu Lys Glu Ile Cys Ile
1430 1435 1440

Lys Leu Asn Val Asp Arg Leu Gly Asn Ile Ser Lys Ala
1445 1450 1455

<210> 107
<211> 1071
<212> PRT
<213> Bacillus sp.

<223> APG05733.1_RGN

<400> 107

Met Arg Glu Leu Asp Tyr Arg Ile Gly Leu Asp Ile Gly Thr Asn Ser
1 5 10 15

Ile Gly Trp Gly Val Ile Glu Leu Ser Trp Asn Lys Asp Arg Glu Gln
20 25 30

Tyr Glu Lys Thr Arg Ile Val Asp Lys Gly Val Arg Met Phe Asp Lys
35 40 45

Ala Glu Ile Pro Lys Thr Gly Ala Ser Leu Ala Glu Pro Arg Arg Ile
50 55 60

Ala Arg Ser Ser Arg Arg Arg Leu Asn Arg Lys Ser Gln Arg Lys Lys
65 70 75 80

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

Leu Thr Ser Leu Tyr Pro Leu Ser Lys Ser Ser Met Asp Ile Trp Asp
100 105 110

Ile Arg Leu Asp Gly Leu Asp Arg Leu Leu Asp Arg Phe Glu Trp Ala
115 120 125

Arg Leu Leu Ile His Leu Ala Gln Arg Arg Gly Phe Lys Ser Asn Arg
130 135 140

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

Gln Val Asn Glu Lys Arg Leu Ser Leu Tyr Arg Thr Val Gly Glu Met
165 170 175

Trp Met Lys Asn Ala Asp Cys Ser Lys Tyr Gly Lys Arg Arg Asn Ser
180 185 190

Pro Asn Glu Tyr Val Phe Ser Val Ser Arg Ala Asp Leu Glu Lys Glu

195

200

205

Ile Val Thr Leu Phe Glu Ala Gln Arg Lys Phe His Ser Ser Tyr Ala
 210 215 220

Ser Val Asp Leu Gln Lys Thr Tyr Ile Gln Ile Trp Ala His Gln Leu
 225 230 235 240

Pro Phe Ala Ser Gly Asn Ala Ile Val Asn Lys Val Gly Tyr Cys Ser
 245 250 255

Leu Leu Lys Gly Lys Glu Lys Arg Val Pro Lys Ala Thr Tyr Thr Phe
 260 265 270

Gln Tyr Phe Asn Thr Leu Asp Gln Ile Asn Arg Thr Arg Leu Gly Pro
 275 280 285

Asn Phe Gln Pro Phe Thr Lys Glu Gln Arg Asp Ile Ile Leu Asp Lys
 290 295 300

Met Phe Gln Arg Thr Asp Tyr Tyr Lys Lys Lys Thr Ile Pro Glu Val
 305 310 315 320

Thr Tyr Tyr Asp Ile Arg Lys Trp Leu Ala Leu Asp Glu Thr Ile Gln
 325 330 335

Phe Lys Gly Leu Thr Tyr Asp Pro Asn Glu Glu Leu Lys Lys Ile Glu
 340 345 350

Met Lys Pro Phe Ile Asn Leu Lys Pro Phe Tyr Glu Ile Lys Lys Val
 355 360 365

Val Thr Asn Tyr Ala Lys Lys Thr Asn Glu Val Phe Ser Ala Leu Asp
 370 375 380

Tyr Asp Thr Val Ala Tyr Ala Leu Thr Val Tyr Lys Thr Asp Lys Asp
 385 390 395 400

Ile Arg Ser Tyr Leu Lys Arg Ser Asn Asn Leu Ser Lys Arg Cys Tyr
 405 410 415

Asp Asp Gln Leu Ile Glu Glu Leu Leu Thr Leu Ser Tyr Thr Lys Phe
420 425 430

Gly His Leu Ser Phe Lys Ala Ile Asn His Val Leu Pro Ile Met Gln
435 440 445

Glu Gly Arg Thr Tyr Gln Glu Ala Ile His Gln Leu Gly Tyr Asp Thr
450 455 460

Thr Asn Leu Lys Lys Glu Asn Arg Ser Met Phe Leu Pro Ile Ile Pro
465 470 475 480

Asp Glu Ile Thr Asn Pro Ile Val Lys Arg Ala Leu Thr Gln Ala Arg
485 490 495

Lys Val Val Asn Ala Ile Ile Arg Arg Tyr Gly Ser Pro Asn Ser Val
500 505 510

His Ile Glu Leu Ala Arg Glu Leu Ser Lys Ser His Asp Glu Arg Lys
515 520 525

Lys Ile Met Thr Ala His Asp Glu Asn Tyr Lys Lys Asn Lys Gly Ala
530 535 540

Ile Ser Ile Leu Ile Glu Asn Gly Ile Leu Asn Pro Thr Gly Tyr Asp
545 550 555 560

Ile Val Arg Tyr Lys Leu Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr
565 570 575

Ser Leu Lys Lys Ile Pro Ala Asn Thr Phe Phe Asn Glu Leu Lys Lys
580 585 590

Glu Arg Ser Gly Pro Pro Val Leu Glu Val Asp His Ile Leu Pro Tyr
595 600 605

Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Gly
610 615 620

Asp Glu Asn Gln Lys Lys Gly Asn Arg Ile Pro Tyr Thr Phe Phe Ser

625		630		635		640
Glu Glu Asp Lys	Glu Trp Glu Ser Phe	Glu Ser Tyr Val Arg Ser Asn				
	645	650			655	
Ser Phe Phe Ser	Lys Lys Lys Arg Gly Tyr Leu Leu Lys	Lys Ala Tyr				
	660	665		670		
Leu Pro Arg Glu Ser Asn Leu Ile Lys Glu Arg His Leu Asn Asp Thr						
	675	680		685		
Arg Tyr Ala Ser Ser Tyr Leu Lys Asn Phe Ile Glu Lys Asn Leu Lys						
	690	695		700		
Phe Lys Glu Ala Val Gly Ile Thr Arg Lys Lys Tyr Val Gln Thr Val						
705	710	715		720		
Asn Gly Val Ile Thr Ala His Leu Arg Lys Arg Trp Gly Leu Glu Lys						
	725	730		735		
Asp Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val						
	740	745		750		
Ala Cys Thr Asp His His Met Val Thr Lys Val Thr Glu Tyr Tyr Gln						
	755	760		765		
Ile Lys Glu Gly Asn Lys Ser Ile Lys Lys Pro Tyr Phe Pro Leu Pro						
	770	775		780		
Trp Met Gly Phe Arg Glu Glu Ile Leu Ser His Leu Glu Ser Gln Pro						
785	790	795		800		
Ile Ala Arg Lys Ile Ser Glu Glu Leu Lys Ile Gly Tyr Gln Ser Ser						
	805	810		815		
Asp Tyr Ile Leu Val Ser Arg Met Pro Lys Arg Ser Val Thr Gly Ser						
	820	825		830		
Ala His Asp Gln Thr Val Met Lys Lys Gly Gly Ile Asp Lys Lys Gly						
	835	840		845		

Lys Thr Ile Ile Ile Lys Arg Val His Leu Lys Asp Ile Lys Phe Asp
850 855 860

Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Leu Ala Thr
865 870 875 880

Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Glu Tyr Arg Lys Lys Ser Lys
885 890 895

Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Lys Gly
900 905 910

Asn Leu Ile Lys Lys Ile Lys Val Glu Val Gln Thr Lys Ser Phe Val
915 920 925

Arg Glu Ile Asn Gly Gly Val Ala Gln Asn Gly Asp Leu Val Arg Val
930 935 940

Asp Leu Phe Glu Lys Asp Asn Arg Tyr Tyr Met Val Pro Ile Tyr Val
945 950 955 960

Val Asp Thr Val Arg Ser Glu Leu Pro Asn Lys Ala Val Thr Ser Ser
965 970 975

Lys Gly Tyr Glu Gln Trp Leu Ser Ile Asp Asn Ser Phe Thr Phe Lys
980 985 990

Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asp Glu
995 1000 1005

Asp Arg Phe Leu Tyr Phe Ser Thr Leu Asp Ile Asn Ser Asp Arg
1010 1015 1020

Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Gln Ala Glu Tyr
1025 1030 1035

Arg Tyr Ser Leu Lys Thr Ile Glu Asn Leu Glu Lys Tyr Glu Ile
1040 1045 1050

Gly Val Leu Gly Asp Leu Arg Leu Val Arg Gln Glu Thr Arg Lys

1055

1060

1065

Ile Phe Lys
1070

<210> 108
<211> 1403
<212> PRT
<213> Chryseobacterium sp.

<223> APG06207.1_RGN

<400> 108

Met Val Glu Arg Ile Leu Gly Leu Asp Leu Gly Val Ser Ser Ile Gly
1 5 10 15

Trp Ala Leu Val Glu Glu Asp Val Glu Asn Pro Glu Asn Asn Lys Ile
20 25 30

Ile Lys Leu Gly Val Arg Val Asn Pro Leu Thr Ile Asp Glu Lys Thr
35 40 45

Asn Phe Glu Lys Gly Lys Ser Ile Thr Thr Asn Ala Gly Arg Thr Ser
50 55 60

Ala Arg Ser Ala Arg Arg Asn Leu Gln Arg Phe Lys Leu Arg Arg Lys
65 70 75 80

Ser Leu Trp Glu Val Leu Ile Asn Tyr Lys Ile Ile Lys Asn Asp Thr
85 90 95

Val Leu Ala Glu Ile Gly Lys Asn Ser Thr Phe Gln Thr Gln Tyr Leu
100 105 110

Arg Ala Lys Ala Ala Lys Glu Lys Ile Glu Leu Asp Glu Leu Ala Arg
115 120 125

Val Leu Phe Leu Ile Asn Lys Lys Arg Gly Tyr Lys Ser Ser Arg Lys
130 135 140

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

Lys Glu Leu Tyr Glu Lys Asn Tyr Thr Pro Gly Gln Tyr Val Tyr Asp
165 170 175

Leu Leu Lys Asn Gly Lys Lys Gln Ile Pro Asp Phe Tyr Arg Ser Asp
180 185 190

Leu Gln Asn Glu Leu Asp Arg Val Trp Glu Phe Gln Lys Gln Tyr His
195 200 205

Pro Asp Ile Leu Thr Asp Glu Tyr Lys Lys Glu Leu Glu Gly Lys Gly
210 215 220

Gln Arg Ala Thr Ala Ala Ile Phe Leu Asn Arg Tyr Lys Ile Tyr Thr
225 230 235 240

Ala Asp Asn Lys Gly Thr Arg Glu Glu Lys Arg Phe Gln Ala Tyr Gln
245 250 255

Trp Arg Val Glu Ala Leu Ser Gln Pro Leu Asp Ile Lys Glu Val Ala
260 265 270

Tyr Val Ile Thr Glu Ile Asn Asn Asn Leu Asn Asn Ser Ser Gly Tyr
275 280 285

Leu Gly Ala Ile Ser Asp Arg Ser Lys Lys Leu Tyr Phe Asn Gly Gln
290 295 300

Thr Ile Gly Glu Tyr Leu Tyr Glu Glu Leu Ser Leu Asn Pro His Ala
305 310 315 320

Arg Leu Lys Asn Gln Val Phe Tyr Arg Gln Asp Tyr Gln Asp Glu Phe
325 330 335

Glu Lys Ile Trp Glu Thr Gln Ala Lys Phe His Ser Val Leu Thr Thr
340 345 350

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

Leu Lys Ser Gln Lys Gly Leu Ile Ser Phe Cys Glu Phe Glu Ser Lys
370 375 380

Glu Ile Val Ile Glu Glu Asn Gly Lys Lys Lys Arg Lys Lys Val Gly
385 390 395 400

Leu Lys Val Val Pro Lys Ser Ser Pro Leu Phe Gln Glu Phe Lys Ile
405 410 415

Trp Gln Val Leu Asn Asn Ile Glu Ile Lys Asn Glu Asp Gly Arg Arg
420 425 430

Tyr Leu Asp Gln Glu Glu Lys Glu Leu Leu Phe Asn Glu Leu Asn Cys
435 440 445

Lys Gly Asn Leu Thr Ser Gln Lys Cys Leu Glu Phe Leu Gly Phe Lys
450 455 460

Ser Lys Glu Thr Lys Ile Asn Phe Lys Val Ile Glu Gly Asn Arg Thr
465 470 475 480

Asn Glu Lys Leu Tyr Asp Ala Phe Leu Lys Ile Leu Glu Leu Lys Gly
485 490 495

Tyr Asn Val Phe Glu Leu Leu Lys Ile Lys Glu Asp Arg Asp Glu Ala
500 505 510

Lys Leu Ser Glu Leu Lys Ala Ser Ala Asp Glu Ile Lys Arg Met Val
515 520 525

Lys Glu Ile Phe Asn Ala Asn Leu Ile Asn Thr Ser Ile Leu Asp Phe
530 535 540

Asn Ala Glu Leu Glu Gly Lys Asp Phe Glu Asn Gln Ala Ser Tyr Gln
545 550 555 560

Phe Trp His Leu Ile Tyr Ser Tyr Glu Gly Asp Asn Ser Pro Ser Gly
565 570 575

Asn Glu Lys Leu Tyr Glu Leu Leu Glu Asn Lys Phe Gly Phe Lys Lys
580 585 590

Glu His Ser Lys Ile Leu Ser Thr Ile Val Phe Pro Gln Glu Tyr Gly
595 600 605

Ser Leu Ser Ala Lys Ala Met Arg Arg Ile Tyr Pro Tyr Ile Lys Asp
610 615 620

Asn Lys Tyr Ser Glu Ala Cys Leu Leu Ala Gly Tyr Asn His Ser Lys
625 630 635 640

Gln Ser Leu Thr Lys Glu Glu Leu Glu Asn Arg Lys Leu Lys Glu Gln
645 650 655

Leu Glu Val Leu Pro Lys Asn Ser Leu Arg Asn Pro Val Val Glu Lys
660 665 670

Ile Leu Asn Gln Met Ile Asn Leu Val Asn Ala Ile Ile Ser Glu His
675 680 685

Gly Lys Leu Asp Gly Val Arg Ile Glu Leu Ala Arg Glu Leu Lys Lys
690 695 700

Ser Ala Gln Glu Arg Glu Asp Met Thr Lys Ser Ile Asn Glu Ala Thr
705 710 715 720

Ile Leu His Gln Lys Tyr Ala Thr Val Leu Arg Gln Glu Phe Gly Val
725 730 735

Val Asn Pro Ser Arg Asn Asp Ile Ile Arg Tyr Lys Leu Tyr Met Glu
740 745 750

Leu Ala Gly Asn Gly Tyr Lys Asp Leu Tyr Thr Asn Val Lys Ile Glu
755 760 765

Lys Glu Asn Ile Phe Thr Asp Lys Tyr Asp Ile Asp His Ile Ile Pro
770 775 780

Gln Ser Arg Phe Phe Asp Asp Ser Phe Ser Asn Lys Val Leu Val Pro
785 790 795 800

Arg Gln Ala Asn Leu Asp Lys Gly Asn Leu Thr Gly Tyr Asp Phe Met
805 810 815

Ser Asn Lys Gly Ala Glu Arg Glu Glu His Phe Leu Asn Val Ile Lys
820 825 830

Asp Leu Leu Glu Asn Gly Ser Ile Ser Lys Ala Lys Tyr Glu Lys Leu
835 840 845

Lys Lys Lys Gly Ile Glu Ile Gly Asp Gly Phe Ile Glu Arg Asp Leu
850 855 860

Arg Asp Thr Gln Tyr Ile Ala Arg Lys Ala Lys Glu Ile Leu Phe Glu
865 870 875 880

Ile Thr Asn Ser Val Ile Ser Thr Ser Gly Arg Ile Thr Asp Lys Leu
885 890 895

Arg Glu Asp Trp Asp Leu Val Asn Thr Met Lys Glu Leu Asn Leu Asp
900 905 910

Lys Tyr Arg Arg Leu Gly Leu Thr Glu Ile Val Ile Asn Ser Lys Gly
915 920 925

Glu Glu Lys Glu Val Ile Pro Asp Trp Thr Lys Arg Asn Asp His Arg
930 935 940

His His Ala Met Asp Ala Leu Thr Val Ala Phe Thr Thr Arg Asn His
945 950 955 960

Ile Gln Tyr Leu Asn Tyr Leu Asn Ala Arg Lys Asp Glu Lys His Lys
965 970 975

Glu His Gln Asn Ile Tyr Ala Ile Glu Asn Ile Ile Thr Glu Ile Ile
980 985 990

Glu Lys Lys Asn Gly Ser Lys Ser Arg Arg Phe Lys Thr Pro Met Asn
995 1000 1005

Arg Phe Arg His Glu Ala Lys Gln His Leu Lys Glu Val Leu Val
1010 1015 1020

Ser His Lys Thr Lys Asn Lys Val Val Thr Ser Asn Ile Asn Lys
1025 1030 1035

Thr Lys Lys Lys Arg Gly Ile His Lys Lys Ala Glu Leu Thr Pro
1040 1045 1050

Arg Gly Gln Leu His Lys Glu Thr Val Tyr Gly Ser Lys Lys Phe
1055 1060 1065

Leu Gln Ser Arg Glu Glu Lys Val Ser Ala Lys Phe Asp Tyr Glu
1070 1075 1080

Thr Ile Leu Met Val Val Asn Pro Ile His Arg Asn Ala Leu Leu
1085 1090 1095

Asn Arg Leu Lys Glu Tyr Gly Asn Asp Pro Lys Lys Ala Phe Ser
1100 1105 1110

Gly Lys Asn Ala Val Asn Lys Thr Pro Val Tyr Leu Ser Asp Ser
1115 1120 1125

Lys Ile Glu Val Leu Pro Glu Lys Val Thr Leu Ser Trp Phe Glu
1130 1135 1140

Thr Gly Tyr Thr Ile Arg Lys Ala Val Thr Pro Asp Asn Phe Lys
1145 1150 1155

Asp Tyr Lys Asn Leu Glu Lys Ile Thr Asp Met Gly Ile Lys Lys
1160 1165 1170

Ile Leu Lys Glu Arg Leu Asp Gln Phe Lys Gly Asn Ala Lys Glu
1175 1180 1185

Ala Phe Ser Asn Leu Asp Lys Glu Pro Ile Trp Leu Asn Glu Glu
1190 1195 1200

Lys Gly Ile Ala Ile Lys Thr Val Thr Ile Thr Gly Val Ser Asn
1205 1210 1215

Ala Glu Ser Leu His Phe Lys Lys Asp His Leu Gly Lys Glu Ile
1220 1225 1230

Leu Asp Glu Asn Gly Asn Lys Ile Pro Val Asp Phe Val Ser Thr
1235 1240 1245

Gly Asn Asn His His Leu Ala Ile Tyr Leu Asn Glu Glu Glu Lys
1250 1255 1260

Leu Asp Asp Lys Met Val Thr Phe Tyr Glu Ala Val Leu Arg Val
1265 1270 1275

Asn Gln Gly Leu Pro Val Ile Asp Lys Asn Tyr Asn Arg Glu Lys
1280 1285 1290

Gly Tyr Lys Phe Leu Met Thr Leu Lys Gln Asn Glu Met Phe Val
1295 1300 1305

Phe Pro Asn Glu Glu Phe Asp Pro Ser Glu Ile Asp Leu Leu Asp
1310 1315 1320

Glu Lys Asn Leu Glu Gln Ile Ser Gln Asn Leu Phe Arg Val Gln
1325 1330 1335

Lys Ile Ser Lys Val Gly Tyr Gly Asn Ser Phe Ile Arg Asp Phe
1340 1345 1350

Val Phe Arg His His Leu Glu Thr Thr Val Glu Asp Arg Lys Glu
1355 1360 1365

Leu Arg Asn Thr Thr Tyr Ile Gln Leu Lys Ser Leu Glu Gly Leu
1370 1375 1380

Arg Asn Ile Val Lys Val Arg Leu Asn His Leu Gly Lys Ile Val
1385 1390 1395

Gln Ile Gly Glu Tyr
1400

<210> 109

<211> 1404

<212> PRT
<213> Sphingobacterium sp.

<223> APG01647.1_RGN

<400> 109

Met Met Lys Arg Ile Leu Gly Leu Asp Leu Gly Thr Ser Ser Ile Gly
1 5 10 15

Trp Ala Tyr Val Leu Glu Ala Gln Asp Glu Asn Glu Ile Ser Gln Ile
20 25 30

Lys Lys Leu Gly Val Arg Val Asn Pro Leu Thr Thr Asp Glu Gln Leu
35 40 45

Asn Phe Glu Lys Gly Lys Pro Ile Thr Thr Asn Ala Gly Arg Thr Leu
50 55 60

Ala Arg Ser Ala Arg Arg Asn Leu Gln Arg Tyr Lys Leu Arg Arg Glu
65 70 75 80

Asn Leu Ile Glu Val Leu Lys Arg Glu Lys Trp Ile Asp Asp Lys Ser
85 90 95

Ile Leu Ala Glu Asn Gly Asn Lys Ser Thr Phe Glu Thr Tyr Ala Leu
100 105 110

Arg Ala Ala Ala Ala Ser Lys Glu Ile Lys Leu Glu Glu Leu Ala Arg
115 120 125

Val Leu Leu Met Ile Asn Lys Lys Arg Gly Tyr Lys Ser Ser Arg Lys
130 135 140

Val Asn Lys Glu Asp Asp Gly Gln Leu Val Asp Gly Met Glu Val Ala
145 150 155 160

Lys Lys Leu Tyr His Glu Asn Leu Thr Pro Gly Glu Phe Val Tyr Arg
165 170 175

Leu Leu Lys Glu Gly Lys Lys Lys Ile Pro Asp Phe Tyr Arg Ser Asp
180 185 190

Leu Gln Arg Glu Leu Asp Leu Ile Trp Asp Phe Gln Phe Lys Tyr Tyr
195 200 205

Pro Asp Leu Leu Thr Gln Asp Leu Lys Glu Glu Leu Lys Gly Lys Gly
210 215 220

Leu Arg Ala Thr Ser Ala Ile Phe Trp Ser Lys Tyr Lys Phe Asn Thr
225 230 235 240

Ala Glu Asn Lys Gly Thr Arg Glu Glu Lys Lys Ile Gln Ala Tyr Gln
245 250 255

Trp Arg Lys Asp Gly Leu Thr Lys Gln Leu Glu Gln Glu Glu Val Ala
260 265 270

Tyr Ile Ile Ala Glu Ile Asn Gly Gln Leu Ala Asn Ser Ser Gly Tyr
275 280 285

Leu Gly Ala Ile Ser Asp Arg Ser Lys Glu Leu Tyr Phe Asn Lys Gln
290 295 300

Thr Val Gly Gln Phe Leu Phe Asn Gln Leu Gln His Asn Thr His Ala
305 310 315 320

Arg Val Lys Asn Gln Val Phe Tyr Arg Gln Asp Tyr Leu Asp Glu Phe
325 330 335

Glu Lys Ile Trp Ser Val Gln Ser Gln Tyr His Gln Gly Leu Thr Glu
340 345 350

Gln Leu Lys Glu Glu Ile Arg Asp Ile Ile Ile Phe Tyr Gln Arg Lys
355 360 365

Leu Lys Ser Gln Lys Gly Leu Ile Ser Phe Cys Glu Phe Glu Gln His
370 375 380

Glu Val Val Ile Asn Gly Lys Asn Lys Ile Val Gly Leu Arg Val Ala
385 390 395 400

Pro Lys Ser Ser Pro Ile Phe Gln Glu Phe Lys Ile Trp Gln Gln Ile

405

410

415

Asn Asn Val Lys Leu Arg Asn Lys Ile Thr Lys Glu Val Tyr Asn Leu
 420 425 430

Ala Glu Glu Gln Lys Arg Ser Leu Phe Glu Thr Leu Asn Leu Lys Gly
 435 440 445

Lys Leu Ser Ala Tyr Gln Val Leu Asn Leu Ile Glu Val Lys Pro Lys
 450 455 460

Glu Trp Glu Leu Asn Tyr Thr Glu Leu Glu Gly Asn Ser Thr Asn Leu
 465 470 475 480

Ala Leu Tyr Asn Ala Tyr Leu Asp Ile Leu Asp Ile Glu Gly Tyr Asp
 485 490 495

Val Arg Ser Glu Leu Lys Ile Lys Leu Asn Lys Asp Glu Ile Thr Leu
 500 505 510

Ser Asp Leu Asp Ile Pro Val Ser Glu Ile Lys Gly Met Ile Arg Arg
 515 520 525

Ile Phe Gln His Leu Gly Ile Asn Thr Ser Ile Leu Asp Phe Asn Ala
 530 535 540

Gln Met Gln Gly Asp Asp Phe Glu Lys Gln Leu Ser Tyr Gln Leu Trp
 545 550 555 560

His Leu Leu Tyr Ser Tyr Glu Glu Asp Asn Ser Lys Thr Gly Met Glu
 565 570 575

Arg Leu Tyr Asn Gln Leu Asn Leu Lys Phe Gly Phe Ser Leu Asp Gln
 580 585 590

Ala Lys Ile Phe Gly Lys Ile Ala Leu Gln Glu Asp Tyr Gly Asn Leu
 595 600 605

Ser Thr Lys Ala Ile Arg Lys Ile Tyr Pro Tyr Ile Gln Asp Val Glu
 610 615 620

Tyr Ser Val Ala Cys Lys Gln Ala Gly Tyr Asn His Ser Lys Ser Ser
625 630 635 640

Leu Thr Lys Glu Gln Leu Thr Asn Arg Val Leu Lys Asp His Leu Asp
645 650 655

Ile Leu Pro Lys Asn Ser Leu Arg Asn Pro Val Val Glu Lys Ile Leu
660 665 670

Asn Gln Met Val Asn Val Val Asn Thr Leu Ile Glu Thr Glu Asn Asp
675 680 685

Lys Leu Ile Lys Glu Gly Lys Asn Ala Asp Phe Arg Phe Asp Glu Ile
690 695 700

Arg Ile Glu Leu Ala Arg Glu Leu Lys Lys Asn Ala Lys Glu Arg Glu
705 710 715 720

Glu Leu Thr Lys Ala Ile Asn Thr Ser Lys Ser Glu His Glu Lys Ile
725 730 735

Ile Lys Ile Leu Gln Thr Glu Asp Gly Ile Lys Asn Pro Thr Arg Asn
740 745 750

Asp Ile Ile Arg Phe Lys Leu Tyr Gln Glu Leu Lys Asn Asn Gly Tyr
755 760 765

Lys Asn Leu Tyr Thr Asn Glu Tyr Ile Gln Arg Lys Asp Leu Phe Thr
770 775 780

Asn Val Tyr Asp Ile Asp His Ile Ile Pro Gln Ser Arg Leu Phe Asp
785 790 795 800

Asp Ser Phe Ser Asn Lys Val Leu Val Pro Arg Asn Ile Asn Ile Glu
805 810 815

Lys Gly Asn Gln Thr Ala Phe Asp Tyr Val His Ala Lys Phe Gly Glu
820 825 830

Asp Gly Ile Glu Ala Tyr Glu Ala Arg Val Glu Arg Leu Phe Asn Leu

835

840

845

Lys Glu Glu Gly Val Ser Arg Ser Lys Tyr Lys Lys Leu Leu Met Arg
 850 855 860

Gly Leu Asp Ile Gly Glu Gly Phe Ile Glu Arg Asp Leu Arg Asp Ser
 865 870 875 880

Gln Tyr Ile Ala Lys Lys Ala Lys Ala Met Leu Phe Glu Ile Ala Pro
 885 890 895

Ser Val Ile Ser Thr Ser Gly Ser Val Thr Asp Arg Leu Arg Glu Asp
 900 905 910

Trp Gly Leu Val Ser Val Met Lys Glu Leu Asn Leu Pro Lys Phe Lys
 915 920 925

Ala Val Gly Leu Thr Glu Tyr Leu Glu Thr Lys Asp Gly Asn Arg Lys
 930 935 940

Glu Val Ile Lys Asp Trp Ser Lys Arg Asn Asp His Arg His His Ala
 945 950 955 960

Met Asp Ala Leu Thr Val Ala Phe Thr Lys His Ser His Ile His Tyr
 965 970 975

Leu Asn His Leu Asn Ala Arg Lys Asn Glu Lys Ser Glu Phe Phe Ser
 980 985 990

Thr Ile Lys Ala Ile Glu Ile Lys Glu Thr Tyr Val Glu Arg Asp Asp
 995 1000 1005

Leu Gly Asn Arg Lys Arg Leu Phe Lys Glu Pro Ile Pro His Phe
 1010 1015 1020

Arg Ser Ile Ala Lys Glu His Leu Glu Ser Val Leu Val Ser His
 1025 1030 1035

Lys Ala Lys Asn Lys Val Val Thr Lys Asn Arg Asn Lys Ile Asp
 1040 1045 1050

Gly Lys Lys Gln Ala Gln Glu Val Leu Thr Pro Arg Gly Gln Leu
1055 1060 1065

His Lys Glu Ser Ile Tyr Gly Lys Ile Leu Gln Tyr Ala Ser Lys
1070 1075 1080

Glu Glu Lys Ile Ser Ala Lys Phe Asp Ala Ala Thr Ile Ala Met
1085 1090 1095

Val Ser Asn Pro Arg Tyr Arg Ser Ala Leu Leu Asp Arg Leu Ser
1100 1105 1110

Ala Tyr Gly Asn Asp Pro Lys Lys Ala Phe Thr Gly Lys Asn Ser
1115 1120 1125

Pro Asn Lys Ser Pro Ile Tyr Leu Asp Gln Glu Lys Gln Leu Ala
1130 1135 1140

Val Pro Glu Lys Val Lys Leu Val Trp Leu Glu Glu Asp Tyr Thr
1145 1150 1155

Ile Arg Lys Glu Ile Gly Pro Asp Leu Lys Ile Glu Lys Val Ile
1160 1165 1170

Asp Gln Gly Val Lys Arg Ile Leu Glu Gln Arg Leu Thr Glu Tyr
1175 1180 1185

Gly Gly Asp Ser Lys Lys Ala Phe Ser Asp Leu Asp Asn Asn Pro
1190 1195 1200

Ile Trp Leu Asn Lys Glu Lys Gly Ile Ala Ile Lys Arg Val Thr
1205 1210 1215

Ile Ser Gly Val Lys Asn Ala Glu Ala Leu His Ile Lys Lys Asp
1220 1225 1230

His Leu Gly Gln Val Leu Lys Asp Lys Asp Gly Lys Pro Met Pro
1235 1240 1245

Val Asp Phe Val Ser Thr Gly Asn Asn His His Val Ala Ile Tyr

1250

1255

1260

Glu Asp Asp Asn Gly Ala Leu Gln Glu Arg Ile Val Pro Phe Tyr
 1265 1270 1275

Glu Ala Val Ala Arg Val Asn Asn Gly Phe Pro Ile Ile Asp Lys
 1280 1285 1290

His Tyr Asn Gln His Leu Gly Trp Lys Phe Leu Phe Ser Met Lys
 1295 1300 1305

Gln Asn Glu Met Phe Leu Phe Pro Ser Glu Glu Phe Asp Pro Lys
 1310 1315 1320

Gln Ile Asn Leu Leu Asp Arg Ser Asn Tyr Arg Leu Ile Ser Lys
 1325 1330 1335

Asn Leu Phe Arg Val Gln Lys Phe Gly Glu Leu Ser Lys Ser Gly
 1340 1345 1350

Phe Trp Phe Arg His His Leu Glu Thr Ser Val Asp Val Asn Lys
 1355 1360 1365

Ser Leu Arg Ser Ile Ser Tyr Phe Asp Phe Tyr Ser Lys Asp Phe
 1370 1375 1380

Met Lys Asn Ile Ala Lys Val Arg Ile Asn His Leu Gly Glu Ile
 1385 1390 1395

Ile His Val Gly Glu Tyr
 1400

<210> 110

<211> 1453

<212> PRT

<213> Chryseobacterium sp.

<223> APG08032.1_RGN

<400> 110

Met Met Ile Lys Asn Ile Leu Gly Leu Asp Leu Gly Thr Asn Ser Ile
 1 5 10 15

Gly Trp Ala Leu Ile Lys Gln Asp Phe Glu Asn Lys Tyr Gly Glu Ile
20 25 30

Leu Gly Met Gly Ser Arg Ile Ile Pro Met Asp Gln Lys Val Leu Gly
35 40 45

Glu Phe Gly Ser Gly Asn Ser Val Ser Gln Thr Ala Asp Arg Thr Lys
50 55 60

Tyr Arg Gly Val Arg Arg Leu Arg Glu Arg Phe Leu Leu Arg Arg Glu
65 70 75 80

Arg Leu His Arg Val Leu Asn Val Leu Asn Phe Leu Pro Ile His Tyr
85 90 95

Ala Ser Gln Ile Asp Phe Glu Lys Arg Phe Gly Lys Phe Arg Glu Glu
100 105 110

Ile Glu Pro Lys Leu Val Tyr Asn Asn Asp Gly Phe Ile Phe Lys Asn
115 120 125

Ser Phe Glu Glu Met Leu Ser Asp Phe Arg Lys His Gln Pro Gln Ile
130 135 140

Leu Glu Asn Asp Lys Lys Ile Pro Tyr Asp Trp Thr Ile Tyr Tyr Leu
145 150 155 160

Arg Lys Lys Ala Leu Thr Gln Lys Ile Glu Lys Glu Glu Leu Ala Trp
165 170 175

Ile Leu Leu Asn Phe Asn Gln Lys Arg Gly Tyr Tyr Gln Leu Arg Gly
180 185 190

Glu Asp Ile Glu Glu Glu Lys Asp Lys Thr Phe Val Arg Leu Lys Val
195 200 205

Glu Lys Ile Val Asp Ser Gly Asp Asn Val Lys Gly Lys Ile Leu Tyr
210 215 220

Asp Val Tyr Phe Asp Asn Gly Trp Lys Tyr Asp Lys Gln Ile Val Lys
225 230 235 240

Thr Glu Asp Trp Ile Glu Arg Val Lys Glu Phe Ile Val Thr Glu Ser
245 250 255

Phe Leu Lys Asn Gly Asp Ile Lys Arg Thr Tyr Lys Ala Val Asp Ser
260 265 270

Glu Lys Asp Trp Ile Ala Ile Lys Thr Lys Thr Glu Gln Glu Ile Glu
275 280 285

Lys Ser His Lys Thr Val Gly Thr Tyr Ile Tyr Glu Thr Ile Leu His
290 295 300

Asn Pro Lys Gln Lys Ile Lys Gly Lys Leu Val Arg Thr Ile Glu Arg
305 310 315 320

Lys Phe Tyr Lys Asp Glu Leu Arg Gln Ile Leu Glu Lys Gln Lys Glu
325 330 335

Phe His Pro Glu Leu Gln Asn Asp Asp Leu Tyr Ser Asp Cys Val Arg
340 345 350

Glu Leu Tyr Arg Asn Asn Glu Ala His Gln Leu Thr Leu Ser Lys Lys
355 360 365

Asp Phe Val His Leu Phe Ile Glu Asp Ile Ile Phe Tyr Gln Arg Pro
370 375 380

Leu Arg Ser Gln Lys Ser Ser Ile Ser Asn Cys Thr Leu Glu Tyr Arg
385 390 395 400

Arg Tyr Lys Asp Lys His Gly Val Glu His Ile Gln Tyr Leu Lys Ala
405 410 415

Ile Pro Lys Ser Asn Pro Tyr Tyr Gln Glu Phe Arg Ile Trp Gln Trp
420 425 430

Met Tyr Asn Leu Asn Ile Tyr Lys Arg Asp Asp Asp Thr Asn Val Thr
435 440 445

Lys Glu Phe Leu Asn Ser Thr Glu Asp Phe Glu Asn Leu Phe Asp Phe
450 455 460

Leu Asn Asn Arg Lys Glu Val Glu Gln Lys Ala Leu Leu Lys His Phe
465 470 475 480

Lys Leu Asn Glu Lys Thr His Arg Trp Asn Phe Val Glu Asn Lys Lys
485 490 495

Tyr Pro Cys Asn Glu Thr Lys Thr Met Ile Ser Ser Arg Leu Asp Lys
500 505 510

Val Glu Asn Ile Ser Glu Asn Phe Leu Thr Ser Glu Ile Glu His Lys
515 520 525

Ile Trp His Ile Ile Tyr Ser Val Asn Asp Lys Ile Glu Tyr Glu Lys
530 535 540

Ala Leu Lys Ser Phe Ala Asn Lys Asn Asn Leu Asp Glu Val Ser Phe
545 550 555 560

Phe Glu Ala Phe Lys Lys Phe Pro Pro Phe Lys Asn Glu Tyr Gly Ser
565 570 575

Phe Ser Glu Lys Ala Ile Lys Lys Leu Leu Pro Leu Met Arg Val Gly
580 585 590

Lys Tyr Trp Asp Glu Asp Glu Ile Val Lys Asn Ser Asp Thr Tyr Phe
595 600 605

Lys Asn Ile Glu Asp Leu Leu Gly Asn Ile Ala Arg Lys Glu Glu Asn
610 615 620

Ile Ser Asp Asp Asp Arg Lys Lys Tyr Asn Lys Thr Val Asn Leu Lys
625 630 635 640

Leu Arg Glu Glu Leu Glu Ile Phe Gln Asp Ala Glu Ile Ala Ser Phe
645 650 655

Gln Lys Leu Arg Leu His Ile Ala Gln Tyr Leu Val Tyr Gly Arg His
660 665 670

Ser Glu Ala Ser Ile Ile Gly Lys Trp Asn Ser Ala Asp Asp Leu Glu
675 680 685

Glu Phe Leu Lys Glu Phe Lys Gln His Ser Leu Arg Asn Pro Ile Val
690 695 700

Glu Gln Val Ile Thr Glu Thr Leu Arg Val Val Lys Asp Ile Trp Met
705 710 715 720

Lys Tyr Gly Asn Gly Ala Lys Asp Phe Phe Asn Glu Ile His Ile Glu
725 730 735

Leu Gly Arg Glu Met Lys Leu Pro Ala Glu Asp Arg Lys Lys Leu Thr
740 745 750

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

Leu Ala Glu Met Met Asn Asp Ser Ser Val Glu Asn Val Arg Pro Phe
770 775 780

Ser Pro Met Gln Gln Glu Ile Leu Lys Ile Tyr Glu Asp Gly Val Leu
785 790 795 800

Asn Ser Gly Ile Glu Ile Glu Asp Glu Tyr Leu Lys Ile Ser Lys Thr
805 810 815

Ala Gln Pro Ser Ser Ser Asp Leu Lys Arg Tyr Lys Leu Trp Leu Glu
820 825 830

Gln Lys Tyr Lys Ser Pro Tyr Thr Gly Gln Ile Ile Pro Leu Asn Lys
835 840 845

Leu Phe Thr Pro Glu Tyr Glu Ile Glu His Ile Ile Pro Lys Ser Arg
850 855 860

Tyr Phe Asp Asp Ser Phe Ser Asn Lys Ile Ile Cys Glu Ser Ala Val
865 870 875 880

Asn Lys Leu Lys Asp Asn Tyr Ile Gly Leu Gly Phe Ile Lys Gln Phe
885 890 895

Ala Gly Ala Ile Val Glu Leu Gly Phe Gly Lys Asn Val Lys Val Phe
900 905 910

Glu Ile Glu Glu Tyr Glu Asp Phe Val Lys Lys His Tyr Ala Asn Asn
915 920 925

Arg Gly Lys Arg Asn Lys Leu Leu Leu Glu Glu Ile Pro Glu Lys Met
930 935 940

Ile Glu Arg Gln Leu Asn Asp Thr Arg His Ile Ser Lys Tyr Ile Ser
945 950 955 960

Ser Val Leu Ser Asn Ile Val Arg Val Glu Asp Gly Thr Asp Glu Gly
965 970 975

Val Asn Ser Lys Asn Ile Val Pro Gly Asn Gly Lys Ile Thr Ile Gln
980 985 990

Leu Lys Gln Asp Trp Gly Leu Asn Asp Val Trp Asn Asp Leu Ile Leu
995 1000 1005

Pro Arg Phe Glu Arg Met Asn Gln Leu Thr Asn Ser Thr Asp Phe
1010 1015 1020

Thr Ala Trp Asn Lys Asn His Gln Lys Tyr Leu Pro Thr Val Pro
1025 1030 1035

Ile Glu Phe Ser Lys Gly Phe Ser Lys Lys Arg Ile Asp His Arg
1040 1045 1050

His His Ala Leu Asp Ala Leu Val Ile Ala Cys Thr Thr Lys Asp
1055 1060 1065

His Val Asn Leu Leu Asn Asn Gln Ser Ala Lys Ser Asp Thr Lys
1070 1075 1080

Arg Tyr Asp Leu Lys Lys Lys Leu Met Lys Phe Glu Lys Val Val
1085 1090 1095

Tyr Asn His Thr Gln Thr Gly Glu Lys Ile Glu Arg Asp Ile Pro
1100 1105 1110

Lys Gln Phe Leu Lys Pro Trp Glu Thr Phe Thr Ile Asp Ala Lys
1115 1120 1125

Asn Arg Leu Glu Thr Ile Ile Val Ser Phe Lys Gln Asn Leu Arg
1130 1135 1140

Val Ile Asn Lys Ala Thr Asn His Tyr Glu Lys Tyr Val Glu Lys
1145 1150 1155

Glu Gly Val Lys Met Lys Glu Arg Val Glu Gln Thr Gly Thr Asn
1160 1165 1170

Trp Ala Ile Arg Lys Ser Leu His Glu Glu Thr Val Ser Gly Lys
1175 1180 1185

Ile Thr Leu Ser Trp Val Glu Thr Ser Lys Gly Glu Phe Ile Thr
1190 1195 1200

Ala Thr Arg Lys Pro Leu Asp Ser Ser Phe Thr Phe Glu Lys Ile
1205 1210 1215

Asn Lys Ile Thr Asp Thr Gly Ile Gln Lys Ile Leu Asn Asn Tyr
1220 1225 1230

Leu Glu Ala Lys Asp Asn Asn Pro Glu Leu Ala Phe Ser Ala Glu
1235 1240 1245

Gly Ile Glu Asp Leu Asn Lys Asn Ile Glu Lys Tyr Asn Asp Gly
1250 1255 1260

Lys Ser His Gln Pro Ile Ile Lys Ile Arg Leu Tyr Glu Lys Gly
1265 1270 1275

Lys Gly Arg Phe Val Leu Gly Gln Thr Gly Asn Lys Ile Asn Lys
1280 1285 1290

Tyr Val Gln Gly Ser Pro Asn Leu Phe Phe Ala Ile Tyr Lys Asp
1295 1300 1305

Glu Asn Gly Lys Lys Ile Phe Glu Ser Ile Arg Leu Asp Ile Val
1310 1315 1320

Ile Glu Arg Leu Lys Gln Gly Leu Gln Ala Ile Pro Glu Thr Asn
1325 1330 1335

Gln Asn Gly Val Ser Leu Tyr Gln Ser Leu Ser Pro Leu Asp Leu
1340 1345 1350

Val Tyr Ile Pro Thr Glu Tyr Glu Leu Glu Ser Pro His Ile Leu
1355 1360 1365

Asp Phe Ser Lys Leu Asn Lys Thr Gln Ile Thr Arg Leu Tyr Asn
1370 1375 1380

Thr Asn Asp Phe Ser Gly Val Thr Ala Tyr Phe Ser Gln Asn Ser
1385 1390 1395

Phe Ala Lys His Ile Tyr Pro Lys Glu Met Asp Leu Ser Trp Asn
1400 1405 1410

Glu Lys Lys Gln Lys Leu Ser Gly Ser Phe Asp Ser Lys Thr Ala
1415 1420 1425

Ser Tyr Asn Asn Thr Ser Ile Lys Asp Ile Phe Ile Lys Val Lys
1430 1435 1440

Val Asp Arg Leu Gly Asn Ile Ser Lys Ala
1445 1450

<210> 111

<211> 1093

<212> PRT

<213> Brucella sp.

<223> APG05712.1_RGN

<400> 111

Met Lys Thr Leu Gly Ile Asp Leu Gly Thr Ala Ser Ile Gly Trp Ala
1 5 10 15

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

Gln Ser Glu Met Ala Gly Arg Asp Pro Gln Ser Lys Ala Ser Leu Ala
35 40 45

Val Ala Arg Arg Asp Ala Arg Gly Ala Arg Arg Arg Arg Asp Arg Tyr
50 55 60

Leu Lys Arg Arg Arg Arg Leu Leu Asp Leu Leu Thr Glu His Gly Leu
65 70 75 80

Ala Pro Ala Asp Glu Lys Ser Arg Lys Ala Leu Val Arg Glu Tyr Glu
85 90 95

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

Arg Ala Leu Asp Glu Ala Leu Thr Pro Tyr Glu Ile Gly Arg Ala Ile
115 120 125

Phe Gln Leu Asn Gln Arg Arg Gly Phe Lys Ser Asn Arg Lys Thr Asp
130 135 140

Ser Asn Asp Pro Glu Gln Gly Lys Ile Ala Thr Ala Ile His Val Leu
145 150 155 160

Asp Ala Lys Met Asp Glu Asp Lys Ala Arg Thr Phe Gly Glu Trp Leu
165 170 175

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

Asp Gly Asp Ser Tyr Asp Phe Tyr Pro Ser Arg Ala Ala Leu Glu Arg
195 200 205

Glu Phe Asp Arg Leu Met Ala Ser Gln Lys Arg Phe His Pro Asp Leu

210

215

220

Leu Asn Ser Ser Val Ile Asp Asp Ile Arg Lys Val Val Phe His Gln
 225 230 235 240

Arg Pro Leu Lys Pro Val Gln Pro Gly Lys Cys Ser Tyr Asn His Arg
 245 250 255

Glu Ser Arg Leu Pro Lys Ala His Pro Leu Phe Gln Lys Phe Arg Leu
 260 265 270

Leu Lys Glu Val Asn Glu Leu Glu Ile Val Gly Glu Asp Gln Arg Tyr
 275 280 285

Val Lys Leu Thr Pro Ala Gln Arg Asp Val Leu Thr Leu Ala Leu Arg
 290 295 300

Thr Gly Leu Thr Lys Gln Gly Arg Leu Pro Phe Ser Lys Leu Arg Ser
 305 310 315 320

Ile Leu Lys Leu Gly Lys Glu Val Arg Phe Asn Lys Glu Lys Asp Asn
 325 330 335

Arg Thr Asp Leu Glu Gly Asp Val Ile His Phe Arg Val Ser Arg Pro
 340 345 350

Asp Cys Phe Gly Asn Arg Trp Ala Ala Met Pro Val Glu Glu Gln Ala
 355 360 365

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

Asp Trp Leu Lys Asn Glu Ala Gly Leu Asp Glu Ala His Ala Arg Ala
 385 390 395 400

Val Ala Asp Thr Pro Val Pro Asp Gly Phe Gly Arg Met Gly Pro Ser
 405 410 415

Ala Leu Ser Ala Leu Ala Asp Ala Met Glu His Glu Ile Asp Ala Gln
 420 425 430

Gly Phe Val Ile Thr Glu Ala Glu Ala Ala Lys Arg Val Tyr Gly Arg
435 440 445

Thr Asn Ser Glu Ala Asp Pro Gly Arg Lys Gly Val Asp Gln Leu Pro
450 455 460

Lys Tyr Gln Glu Val Leu Gln Arg His Ile Pro Pro Gly Thr Gly Glu
465 470 475 480

Pro Asp Asp Pro Tyr Asp Glu Tyr Met Gly Arg Ile Thr Asn Pro Thr
485 490 495

Val His Ile Ala Leu Asn Gln Leu Arg Arg Leu Val Asn Ala Leu Ile
500 505 510

Arg Lys Tyr Gly Lys Pro Asn Lys Ile Ala Ile Glu Val Gly Arg Glu
515 520 525

Leu Lys Leu Asn Glu Lys Gln Arg Asn Glu Val Asn Arg Glu Ile Gly
530 535 540

Gln Asn Thr Arg Ala Ala Met Ala Arg Gly Gln Gln Leu Val Glu Ile
545 550 555 560

Phe Lys Gln Pro Asn Thr Gly Tyr Asn Arg Leu Arg Leu Glu Leu Trp
565 570 575

Glu Asp Leu Asn Arg Glu Gln Pro Leu Lys Arg Leu Cys Thr Tyr Cys
580 585 590

Gly Lys Ala Ile Ala Ala His Met Leu Phe Asn Gly Glu Thr Asp Ile
595 600 605

Asp His Ile Leu Pro Tyr Ser Lys Thr Leu Asp Asp Ser Lys Ala Asn
610 615 620

Arg Leu Leu Cys Cys Thr Pro Cys Asn Arg Glu Lys Lys Asn Tyr Ala
625 630 635 640

Pro Ala Asn Val Leu Gln Trp Arg Asp His Tyr Gly Glu Ile Leu Ala

645

650

655

Arg Ala Thr Ala Leu Pro Lys Asn Lys Gln Trp Arg Phe Ala Glu Asp
 660 665 670

Ala Met Thr Arg Tyr Glu Ala Glu Gly Gly Phe Leu Ala Arg Gln Leu
 675 680 685

Thr Asp Met Gln Tyr Ile Ser Arg Leu Ala Leu Thr Tyr Leu Ala His
 690 695 700

Leu Tyr Asp Tyr Glu Glu Pro Asp Leu Asp Gly Val Tyr Lys Arg His
 705 710 715 720

Asp Arg Val Arg Ala Leu Pro Gly Arg Met Thr Glu Met Leu Arg Arg
 725 730 735

Gln Trp Ala Leu Asn Glu Leu Leu His Gly His Asn Leu Ala Gly Gly
 740 745 750

Asp Gly Ala Lys Glu Lys Asn Arg Leu Asp His Arg His His Ala Ile
 755 760 765

Asp Ala Ile Val Ile Ala Cys Thr Ser Gln Ser Leu Ile Asn Arg Leu
 770 775 780

Ser Thr Ala Ala Gly Glu Ala Glu Glu Arg Gly Ala Ala Arg Val Val
 785 790 795 800

Glu Arg Ile Asp Pro Pro Trp Pro Ser Phe Arg Glu Asp Val Arg Glu
 805 810 815

Ala Val Asn Ala Ile Val Val Ser His Lys Pro Asp His Gly Thr Ala
 820 825 830

Ser Arg Ser Gly Tyr Asp Lys Gly Arg Gly Gln Thr Ala Gly Lys Leu
 835 840 845

His Asn Asp Thr Ala Tyr Gly Glu Thr Gly Glu Lys Asp His Asn Gly
 850 855 860

Asn Asn Leu Val Val Arg Arg Ile Ala Ile Ser Asp Ile Lys Arg Ser
865 870 875 880

Ala Asp Ile Met Lys Ile Arg Thr Asn Ala His Gly His Ser Glu Leu
885 890 895

Arg Asp Arg Leu Tyr Glu Ala Thr Arg Asp Leu Glu Gly Lys Ala Phe
900 905 910

Glu Gln Ala Val Thr Ala Phe Val Lys His Asp Ala Lys Phe Lys Gly
915 920 925

Ile Arg His Val Arg Val Thr Glu Val Gln Asn Pro Val Trp Ile Thr
930 935 940

His Gly Gly Gly Lys Tyr Lys Lys Gly Tyr Leu Pro Gly Gly Asn Asp
945 950 955 960

Arg Phe Asp Val Trp Glu Leu Pro Asp Gly Lys Trp Asp Ala Glu Val
965 970 975

Val Thr Thr Phe Asp Ala His Arg Pro Asp Phe Thr Pro Arg Met Arg
980 985 990

Ile Glu His His Asn Ala Arg Lys Ile Met Ser Leu Lys Lys Gly Asp
995 1000 1005

Met Ile Ala Tyr Asp Asp Pro Asp Ser Gly Lys Arg Val Ile Ala
1010 1015 1020

Ile Val Arg Lys Phe Asp Gln Arg Asn Lys Gln Leu Tyr Leu Asp
1025 1030 1035

Pro His Asn Glu Ala Gly Asn Leu Asp Gln Arg Glu Lys Glu Lys
1040 1045 1050

Thr Tyr Lys Pro Leu Arg Pro Met Pro Asn Pro Leu Lys Lys Tyr
1055 1060 1065

Arg Pro Arg Gln Val Arg Val Asp Glu Ile Gly Gln Val Phe Asp

1070

1075

1080

Pro Gly Pro Trp Trp Glu Lys Arg Ser Asp
 1085 1090

<210> 112
 <211> 1054
 <212> PRT
 <213> Staphylococcus sp.

<223> APG01658.1_RGN

<400> 112

Met Lys Glu Lys Tyr Ile Leu Gly Leu Asp Leu Gly Ile Thr Ser Val
 1 5 10 15

Gly Tyr Gly Ile Ile Asn Phe Glu Thr Lys Lys Ile Ile Asp Ala Gly
 20 25 30

Val Arg Leu Phe Pro Glu Ala Asn Val Asp Asn Asn Glu Gly Arg Arg
 35 40 45

Ser Lys Arg Gly Ser Arg Arg Leu Lys Arg Arg Arg Ile His Arg Leu
 50 55 60

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

Gln Ile Pro Thr Ser Asn Asn Pro Tyr Gln Ile Arg Val Lys Gly Leu
 85 90 95

Ser Glu Ile Leu Ser Lys Asp Glu Leu Ala Ile Ala Leu Leu His Leu
 100 105 110

Ala Lys Arg Arg Gly Ile His Asn Ile Asn Val Ser Ser Glu Asp Glu
 115 120 125

Asp Ala Ser Asn Glu Leu Ser Thr Lys Glu Gln Ile Asn Arg Asn Asn
 130 135 140

Lys Leu Leu Lys Asp Lys Tyr Val Cys Glu Val Gln Leu Gln Arg Leu
 145 150 155 160

Lys Glu Gly Gln Ile Arg Gly Glu Lys Asn Arg Phe Lys Thr Thr Asp
165 170 175

Ile Leu Lys Glu Ile Asp Gln Leu Leu Lys Val Gln Lys Asp Tyr His
180 185 190

Asn Leu Asp Ile Asp Phe Ile Asn Gln Tyr Lys Glu Ile Val Glu Thr
195 200 205

Arg Arg Glu Tyr Phe Glu Gly Pro Gly Gln Gly Ser Pro Phe Gly Trp
210 215 220

Asn Gly Asp Leu Lys Lys Trp Tyr Glu Met Leu Met Gly His Cys Thr
225 230 235 240

Tyr Phe Pro Gln Glu Leu Arg Ser Val Lys Tyr Ala Tyr Ser Ala Asp
245 250 255

Leu Phe Asn Ala Leu Asn Asp Leu Asn Asn Leu Ile Ile Gln Arg Asp
260 265 270

Asn Ser Glu Lys Leu Glu Tyr His Glu Lys Tyr His Ile Ile Glu Asn
275 280 285

Val Phe Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu
290 295 300

Ile Gly Val Asn Pro Glu Asp Ile Lys Gly Tyr Arg Ile Thr Lys Ser
305 310 315 320

Gly Thr Pro Gln Phe Thr Glu Phe Lys Leu Tyr His Asp Leu Lys Ser
325 330 335

Ile Val Phe Asp Lys Ser Ile Leu Glu Asn Glu Ala Ile Leu Asp Gln
340 345 350

Ile Ala Glu Ile Leu Thr Ile Tyr Gln Asp Glu Gln Ser Ile Lys Glu
355 360 365

Glu Leu Asn Lys Leu Pro Glu Ile Leu Asn Glu Gln Asp Lys Ala Glu
370 375 380

Ile Ala Lys Leu Ile Gly Tyr Asn Gly Thr His Arg Leu Ser Leu Lys
385 390 395 400

Cys Ile His Leu Ile Asn Glu Glu Leu Trp Gln Thr Ser Arg Asn Gln
405 410 415

Met Glu Ile Phe Asn Tyr Leu Asn Ile Lys Pro Asn Lys Val Asp Leu
420 425 430

Ser Glu Gln Asn Lys Ile Pro Lys Asp Met Val Asn Asp Phe Ile Leu
435 440 445

Ser Pro Val Val Lys Arg Thr Phe Ile Gln Ser Ile Asn Val Ile Asn
450 455 460

Lys Val Ile Glu Lys Tyr Gly Ile Pro Glu Asp Ile Ile Ile Glu Leu
465 470 475 480

Ala Arg Glu Asn Asn Ser Asp Asp Arg Lys Lys Phe Ile Asn Asn Leu
485 490 495

Gln Lys Lys Asn Glu Ala Thr Arg Lys Arg Ile Asn Glu Ile Ile Gly
500 505 510

Gln Thr Gly Asn Gln Asn Ala Lys Arg Ile Val Glu Lys Ile Arg Leu
515 520 525

His Asp Gln Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ser Ile Ala
530 535 540

Leu Met Asp Leu Leu Asn Asn Pro Gln Asn Tyr Glu Val Asp His Ile
545 550 555 560

Ile Pro Arg Ser Val Ala Phe Asp Asn Ser Ile His Asn Lys Val Leu
565 570 575

Val Lys Gln Ile Glu Asn Ser Lys Lys Gly Asn Arg Thr Pro Tyr Gln
580 585 590

Tyr Leu Asn Ser Ser Asp Ala Lys Leu Ser Tyr Asn Gln Phe Lys Gln
595 600 605

His Ile Leu Asn Leu Ser Lys Ser Lys Asp Arg Ile Ser Lys Lys Lys
610 615 620

Lys Asp Tyr Leu Leu Glu Glu Arg Asp Ile Asn Lys Phe Glu Val Gln
625 630 635 640

Lys Glu Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg
645 650 655

Glu Leu Thr Ser Tyr Leu Lys Ala Tyr Phe Ser Ala Asn Asn Met Asp
660 665 670

Val Lys Val Lys Thr Ile Asn Gly Ser Phe Thr Asn His Leu Arg Lys
675 680 685

Val Trp Arg Phe Asp Lys Tyr Arg Asn His Gly Tyr Lys His His Ala
690 695 700

Glu Asp Ala Leu Ile Ile Ala Asn Ala Asp Phe Leu Phe Lys Glu Asn
705 710 715 720

Lys Lys Leu Gln Asn Ala Asn Lys Ile Leu Glu Lys Pro Thr Ile Glu
725 730 735

Asn Asn Thr Lys Lys Val Thr Val Glu Lys Glu Glu Asp Tyr Asn Asn
740 745 750

Val Phe Glu Thr Pro Lys Leu Val Glu Asp Ile Lys Gln Tyr Arg Asp
755 760 765

Tyr Lys Phe Ser His Arg Val Asp Lys Lys Pro Asn Arg Gln Leu Ile
770 775 780

Asn Asp Thr Leu Tyr Ser Thr Arg Met Lys Asp Glu His Asp Tyr Ile
785 790 795 800

Val Gln Thr Ile Thr Asp Ile Tyr Gly Lys Asp Asn Thr Asn Leu Lys
805 810 815

Lys Gln Phe Asn Lys Asn Pro Glu Lys Phe Leu Met Tyr Gln Asn Asp
820 825 830

Pro Lys Thr Phe Glu Lys Leu Ser Ile Ile Met Lys Gln Tyr Ser Asp
835 840 845

Glu Lys Asn Pro Leu Ala Lys Tyr Tyr Glu Glu Thr Gly Glu Tyr Leu
850 855 860

Thr Lys Tyr Ser Lys Lys Asn Asn Gly Pro Ile Val Lys Lys Ile Lys
865 870 875 880

Leu Leu Gly Asn Lys Val Gly Asn His Leu Asp Val Thr Asn Lys Tyr
885 890 895

Glu Asn Ser Thr Lys Lys Leu Val Lys Leu Ser Ile Lys Asn Tyr Arg
900 905 910

Phe Asp Val Tyr Leu Thr Glu Lys Gly Tyr Lys Phe Val Thr Ile Ala
915 920 925

Tyr Leu Asn Val Phe Lys Lys Asp Asn Tyr Tyr Tyr Ile Pro Lys Asp
930 935 940

Lys Tyr Gln Glu Leu Lys Glu Lys Lys Lys Ile Lys Asp Thr Asp Gln
945 950 955 960

Phe Ile Ala Ser Phe Tyr Lys Asn Asp Leu Ile Lys Leu Asn Gly Asp
965 970 975

Leu Tyr Lys Ile Ile Gly Val Asn Ser Asp Asp Arg Asn Ile Ile Glu
980 985 990

Leu Asp Tyr Tyr Asp Ile Lys Tyr Lys Asp Tyr Cys Glu Ile Asn Asn
995 1000 1005

Ile Lys Gly Glu Pro Arg Ile Lys Lys Thr Ile Gly Lys Lys Thr
1010 1015 1020

Glu Ser Ile Glu Lys Phe Thr Thr Asp Val Leu Gly Asn Leu Tyr
1025 1030 1035

Leu His Ser Thr Glu Lys Ala Pro Gln Leu Ile Phe Lys Arg Gly
1040 1045 1050

Leu

<210> 113
<211> 1076
<212> PRT
<213> Bacillus sp.

<223> APG06498.1_RGN

<400> 113

Met Ser Glu Leu Asp Tyr Arg Ile Gly Leu Asp Ile Gly Thr Asn Ser
1 5 10 15

Ile Gly Trp Gly Val Ile Glu Leu Phe Trp Asn Lys Asp Arg Glu Arg
20 25 30

Tyr Glu Lys Val Arg Ile Val Asp Lys Gly Val Arg Met Phe Asp Lys
35 40 45

Ala Glu Ile Pro Lys Thr Gly Ala Ser Leu Ala Glu Pro Arg Arg Ile
50 55 60

Ala Arg Ser Ser Arg Arg Arg Leu Asn Arg Lys Ser Gln Arg Lys Lys
65 70 75 80

Glu Ile Arg Asn Leu Leu Val Gln His Gly Val Ile Thr Gln Glu Glu
85 90 95

Leu Asp Leu Leu Tyr Pro Leu Ser Lys Lys Ser Met Asp Ile Trp Asp
100 105 110

Ile Arg Leu Asp Gly Leu Asp Arg Leu Leu Asn His Leu Glu Trp Thr
115 120 125

Arg Leu Leu Ile His Leu Ala Gln Arg Arg Gly Phe Lys Ser Asn Arg
130 135 140

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

Gln Val Asn Glu Lys Arg Leu Ser Leu Tyr Arg Thr Val Gly Glu Met
165 170 175

Trp Ile Lys Asp Ala Glu Phe Ser Lys Tyr Asp Arg Arg Arg Asn Ser
180 185 190

Pro Asn Glu Tyr Val Phe Ser Val Ser Arg Ala Asp Leu Glu Lys Glu
195 200 205

Ile Val Thr Leu Phe Glu Ala Gln Arg Lys Phe Gln Ser Ser Tyr Ser
210 215 220

Ser Lys Asn Leu Gln Glu Thr Tyr Leu Gln Ile Trp Ala His Gln Leu
225 230 235 240

Pro Phe Ala Ser Gly Asn Ala Ile Leu Asn Lys Val Gly Tyr Cys Ser
245 250 255

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

Gln Tyr Phe Ser Ala Leu Asp Gln Val Asn Arg Thr Arg Leu Gly Pro
275 280 285

Asp Phe Gln Pro Phe Thr Gln Glu Gln Lys Glu Ile Ile Leu Asp Lys
290 295 300

Met Phe Gln Arg Thr Asp Tyr Tyr Lys Lys Lys Thr Ile Pro Glu Val
305 310 315 320

Ser Tyr Tyr Asp Ile Arg Lys Trp Leu Glu Leu Asp Glu Thr Ile Gln
325 330 335

Phe Lys Gly Leu Asn Tyr Asp Pro Asn Glu Glu Leu Lys Lys Ile Glu

340

345

350

Lys Lys Pro Phe Ile Asn Leu Lys Ala Phe Tyr Glu Ile Lys Lys Val
 355 360 365

Val Ala Asn Tyr Ala Glu Arg Thr Asn Glu Ala Phe Ser Thr Leu Asp
 370 375 380

Tyr Asp Ala Ile Ala Tyr Ala Leu Thr Val Tyr Lys Thr Asp Lys Asp
 385 390 395 400

Ile Arg Ser Tyr Leu Lys Lys Ser Asn Asn Leu Ser Lys Arg Cys Tyr
 405 410 415

Asp Asp Gln Leu Ile Glu Glu Leu Phe Thr Leu Ser Tyr Thr Lys Phe
 420 425 430

Gly His Leu Ser Phe Lys Ala Ile Asn Arg Val Leu Pro Ile Met Gln
 435 440 445

Glu Gly Arg Thr Tyr Gln Glu Ala Ile Gln Gln Leu Gly Tyr Asp Thr
 450 455 460

Thr Asn Leu Lys Lys Glu Asn Arg Ser Met Phe Leu Pro Leu Ile Pro
 465 470 475 480

Asp Glu Ile Thr Asn Pro Ile Val Lys Arg Ala Leu Thr Gln Ala Arg
 485 490 495

Lys Val Val Asn Ala Ile Ile Arg Arg Tyr Gly Ser Pro Asn Ser Val
 500 505 510

His Ile Glu Leu Ala Arg Glu Leu Ser Lys Ser His Asp Glu Arg Lys
 515 520 525

Lys Ile Met Thr Ala His Asp Glu Asn Tyr Lys Lys Asn Lys Gly Ala
 530 535 540

Ile Ser Ile Leu Ile Glu Asn Gly Ile Leu Asn Pro Thr Gly Tyr Asp
 545 550 555 560

Ile Val Arg Tyr Lys Leu Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr
565 570 575

Ser Leu Lys Glu Ile Pro Pro Asp Thr Phe Phe Asn Glu Leu Lys Lys
580 585 590

Glu Arg Asn Gly Ser Ser Ile Leu Glu Val Asp His Ile Leu Pro Tyr
595 600 605

Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Ser
610 615 620

Asp Glu Asn Arg Asn Lys Gly Asn Arg Ile Pro Tyr Thr Tyr Phe Leu
625 630 635 640

Glu Thr Asn Lys Asp Trp Glu Ala Phe Glu Arg Tyr Val Arg Ser Asn
645 650 655

Lys Leu Phe Ser Lys Lys Lys Arg Glu Tyr Leu Leu Lys Lys Ala Tyr
660 665 670

Leu Pro Arg Glu Ser Glu Leu Ile Lys Glu Arg His Leu Asn Asp Thr
675 680 685

Arg Tyr Ala Ser Thr Phe Leu Lys Asn Phe Ile Glu Gln Asn Leu Gln
690 695 700

Phe Lys Glu Val Glu Val Asn Leu Arg Lys Lys Arg Val Gln Thr Val
705 710 715 720

Asn Gly Val Ile Thr Ala His Leu Arg Lys Arg Trp Gly Leu Glu Lys
725 730 735

Asn Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
740 745 750

Ala Cys Thr Asp His His Met Val Thr Arg Ile Thr Glu Tyr Tyr Gln
755 760 765

Ile Lys Glu Ser Asn Lys Ser Val Lys Lys Pro Tyr Phe Pro Met Pro

770

775

780

Trp Glu Gly Phe Arg Asp Glu Leu Leu Ser His Leu Ala Ser Gln Pro
 785 790 795 800

Ile Ala Lys Lys Ile Ser Glu Glu Leu Lys Ala Gly Tyr Gln Ser Ser
 805 810 815

Asp Tyr Ile Phe Val Ser Arg Met Pro Lys Arg Ser Val Thr Gly Ala
 820 825 830

Ala His Asn Gln Met Ile Arg Arg Lys Gly Gly Ile Asp Lys Lys Gly
 835 840 845

Lys Thr Ile Ile Ile Lys Arg Val Arg Leu Lys Asp Ile Lys Phe Asp
 850 855 860

Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Leu Ala Thr
 865 870 875 880

Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Asp His Gly Lys Asn Leu Lys
 885 890 895

Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Thr Gly
 900 905 910

Asn Leu Ile Lys Arg Val Lys Ile Glu Gly Gln Thr Lys Ala Phe Val
 915 920 925

Arg Glu Val Asn Gly Gly Val Ala Gln Asn Ser Asp Leu Val Arg Val
 930 935 940

Asp Leu Phe Glu Lys Asp Asp Lys Tyr Tyr Met Val Pro Ile Tyr Val
 945 950 955 960

Pro Asp Thr Val Cys Ser Glu Leu Pro Lys Lys Val Val Lys Ser Gly
 965 970 975

Lys Gly Tyr Glu Gln Trp Leu Thr Leu Asp Asn Ser Phe Thr Phe Lys
 980 985 990

Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asp Glu
995 1000 1005

Asp Arg Phe Leu Tyr Phe Gly Thr Leu Asp Ile Asp Ser Asp Arg
1010 1015 1020

Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Gln Asn Glu Tyr
1025 1030 1035

Arg Tyr Ser Leu Lys Thr Ile Glu Asp Leu Glu Lys Tyr Glu Val
1040 1045 1050

Gly Val Leu Gly Asp Leu Arg Leu Val Arg Lys Glu Thr Arg Arg
1055 1060 1065

Asn Phe His Glu Ile Lys Ile Asn
1070 1075

<210> 114
<211> 1108
<212> PRT
<213> Brevibacillus sp.

<223> APG09106.1_RNG

<400> 114

Met Ala Ile Arg Ser Ile Lys Leu Lys Leu Lys Thr Arg Thr Gly Pro
1 5 10 15

Glu Ala Gln Asn Leu Arg Lys Gly Ile Trp Arg Thr His Arg Leu Leu
20 25 30

Asn Glu Gly Val Ala Tyr Tyr Met Lys Met Leu Leu Leu Phe Arg Gln
35 40 45

Glu Ser Thr Gly Gly Gln Thr Lys Lys Glu Leu Gln Glu Glu Leu Val
50 55 60

Arg His Ile Arg Glu Gln Gln Gln Lys Asn Arg Ala Asp Lys Asn Thr
65 70 75 80

Gln Ala Leu Pro Leu Asp Lys Ala Phe Ala Ala Leu Arg Gln Leu Tyr
85 90 95

Glu Leu Leu Val Pro Ser Ser Ile Gly Gln Ser Gly Asp Ala Gln Ile
100 105 110

Ile Ser Arg Lys Phe Leu Ser Pro Leu Val Asp Pro Asn Ser Glu Gly
115 120 125

Gly Lys Gly Thr Ser Lys Ala Gly Ala Lys Pro Thr Trp Gln Lys Lys
130 135 140

Lys Glu Ala Asn Asp Pro Thr Trp Glu Gln Asp Tyr Glu Lys Trp Lys
145 150 155 160

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

Glu Tyr Gly Ile Arg Pro Ile Phe Pro Leu Tyr Thr Asn Thr Val Ala
180 185 190

Asp Ile Ala Trp Leu Pro Leu Gln Ser Asn Gln Phe Val Arg Thr Trp
195 200 205

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

Ser Trp Asn Lys Arg Val Gln Glu Glu Tyr Ser Lys Leu Gln Glu Lys
225 230 235 240

Met Thr Gln Leu Asn Glu Gln Leu Glu Gly Gly Gln Glu Trp Ile Ser
245 250 255

Leu Leu Glu Gln Tyr Glu Glu Gln Arg Glu Gln Glu Leu Ile Glu Asn
260 265 270

Met Thr Ala Ala Asn Asp Lys Tyr Arg Ile Thr Lys Arg Gln Met Lys
275 280 285

Gly Trp Asn Glu Leu Tyr Glu Gln Trp Ser Thr Val Leu Pro Asn Ala
290 295 300

Ser His Glu Gln Tyr Arg Glu Ala Leu Lys Arg Val Gln Gln Arg Leu
305 310 315 320

Arg Gly Arg Phe Gly Asp Ala His Phe Phe Gln Tyr Leu Met Lys Glu
325 330 335

Glu His His Leu Ile Trp Lys Gly Asn Pro Gln Arg Ile His Tyr Phe
340 345 350

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

Ala Thr Met Thr Leu Pro Asp Ala Arg Lys His Pro Leu Trp Val Arg
370 375 380

Phe Asp Ala Arg Gly Gly Asn Leu Gln Asp Tyr Tyr Leu Thr Ala Glu
385 390 395 400

Ala Asp Asn Pro Arg Ser Arg Arg Phe Val Thr Phe Ser Gln Leu Ile
405 410 415

Trp Pro Asn Glu Ser Gly Trp Met Glu Lys Gln Asp Val Glu Val Glu
420 425 430

Leu Ala Leu Ser Lys Gln Phe Tyr Gln Gln Val Thr Leu Gln Lys Asn
435 440 445

Asp Lys Gly Lys Gln Glu Ile Glu Phe Lys Asp Lys Gly Ser Gly Ser
450 455 460

Thr Phe Ser Gly His Leu Gly Gly Ala Lys Leu Gln Leu Glu Arg Gly
465 470 475 480

Asp Leu Glu Lys Glu Glu Lys Asp Phe Glu Gly Gly Glu Ile Gly Ser
485 490 495

Val Tyr Leu Asn Ile Val Ile Asp Phe Glu Pro Leu Gln Glu Val Lys
500 505 510

Asn Gly Arg Leu Gln Ser Pro Tyr Gly Gln Val Leu Gln Leu Val Arg
515 520 525

Arg Pro Asn Glu Phe Pro Lys Val Thr Thr Tyr Lys Ser Glu Glu Leu
530 535 540

Val Glu Trp Met Lys Ala Ser Gln Asn His Ser Ser Gly Val Glu Ser
545 550 555 560

Leu Glu Ser Gly Phe Arg Val Met Ser Ile Asp Leu Gly Leu Arg Thr
565 570 575

Ala Ala Ala Thr Ser Ile Phe Ser Val Glu Glu Ser Asn Asp Ala Asn
580 585 590

Ala Ala Gly Phe Ser Tyr Trp Ile Glu Gly Thr Pro Leu Val Ala Val
595 600 605

His Lys Arg Ser Tyr Met Leu Lys Leu Pro Gly Glu Gln Val Glu Lys
610 615 620

Gln Val Arg Glu Lys Arg Asp Glu Arg Gln Asp Gln Gln Arg Arg Val
625 630 635 640

Arg Phe Gln Ile Arg Ile Leu Ser Gln Val Ile Arg Met Ala Lys Lys
645 650 655

Gln Asn Arg Glu Arg Ala Asp Glu Leu Asp His Leu Ser Gln Ala Leu
660 665 670

Glu Lys Gln Lys Ser Leu Leu Asp Gln Thr Asp Arg Thr Phe Trp Asn
675 680 685

Gly Ile Val Cys Asp Leu Thr Asp Ala Leu Arg Glu Lys Glu Gly Gly
690 695 700

Trp Glu Gln Ala Val Val Gln Ile His Arg Lys Ala Glu Glu His Val
705 710 715 720

Gly Lys Val Val Gln Ala Trp Arg Lys Arg Phe Asp Ala Asp Glu Arg
725 730 735

Lys Gly Ile Ala Gly Leu Ser Met Trp Ser Ile Glu Glu Leu Asp Ser
740 745 750

Leu Arg Lys Leu Leu Ile Ser Trp Ser Arg Arg Thr Arg Asn Pro Gln
755 760 765

Glu Ile Asn Arg Phe Glu Gln Gly His Thr Ser His Gln Arg Leu Leu
770 775 780

Thr His Ile Gln Asn Val Lys Glu Asp Arg Leu Lys Gln Leu Ser His
785 790 795 800

Ala Ile Val Met Thr Ala Leu Gly Tyr Val Tyr Asp Glu Lys Lys Leu
805 810 815

Glu Trp Phe Ala Lys Tyr Pro Ala Cys Gln Val Ile Leu Phe Glu Asn
820 825 830

Leu Ser Gln Tyr Arg Ser His Met Asp Arg Ser Thr Lys Glu Asn Ser
835 840 845

Thr Leu Met Lys Trp Ala His Arg Ser Ile Pro Lys Tyr Val His Met
850 855 860

Gln Ala Glu Pro Tyr Gly Ile Gln Ile Gly Asp Val Arg Ala Glu Tyr
865 870 875 880

Ser Ser Arg Phe His Ala Lys Thr Gly Thr Pro Gly Ile Arg Cys Lys
885 890 895

Met Val Lys Gly Gln Glu Leu Gln Gly Lys Arg Phe Glu Asn Leu Gln
900 905 910

Lys Arg Leu Val Ser Glu Gln Phe Leu Thr Glu Glu Gln Val Lys Gln
915 920 925

Leu Arg Pro Gly Asp Ile Val Pro Asp Asp Ser Gly Glu Trp Phe Met
930 935 940

Thr Leu Ser Asp Gly Ser Glu Gly Lys Glu Val Val Phe Leu Gln Ala
945 950 955 960

Asp Ile Asn Ala Ala Gln Asn Leu Gln Lys Arg Phe Trp Gln Arg Tyr
965 970 975

Asn Glu Leu Phe Lys Val Ser Cys Arg Val Leu Ile Arg Gly Glu Glu
980 985 990

Glu Tyr Leu Ile Pro Lys Ala Lys Ser Val Gln Ala Lys Leu Gly Lys
995 1000 1005

Gly Leu Phe Val Lys Lys Thr Asp Thr Val Met Lys Asp Val Tyr
1010 1015 1020

Val Trp Asp Ser Gln Ala Lys Leu Lys Gly Lys Thr Thr Phe Thr
1025 1030 1035

Glu Glu Ser Glu Ser Pro Glu Gln Leu Glu Asp Phe Gln Glu Ile
1040 1045 1050

Ile Glu Glu Ala Glu Glu Ala Lys Gly Thr Tyr Arg Thr Leu Phe
1055 1060 1065

Arg Asp Pro Ser Gly Val Phe Phe Pro Glu Phe Val Trp Asn Thr
1070 1075 1080

Gln Lys Asp Phe Trp Ser Glu Val Lys Arg Arg Leu Tyr Gly Lys
1085 1090 1095

Leu Arg Glu Arg Phe Leu Met Lys Thr Arg
1100 1105

<210> 115
<211> 1150
<212> PRT
<213> Enterococcus sp.

<223> APG09882.1_RGN

<400> 115

Met Tyr Ser Ile Gly Leu Asp Leu Gly Ile Ser Ser Val Gly Trp Ser

1	5	10	15
Val Ile Asp Glu Arg Thr Gly Asn Val Ile Asp Leu Gly Ile Arg Leu	20	25	30
Phe Ser Ala Lys Asn Ser Glu Lys Asn Leu Glu Arg Arg Thr Asn Arg	35	40	45
Gly Gly Arg Arg Leu Ile Arg Arg Lys Thr Asn Arg Leu Lys Asp Ala	50	55	60
Lys Lys Ile Leu Ala Ala Val Gly Phe Tyr Glu Asp Lys Ser Leu Lys	65	70	75
Asn Ser Cys Pro Tyr Gln Leu Arg Val Lys Gly Leu Thr Glu Pro Leu	85	90	95
Ser Lys Gly Glu Ile Tyr Lys Val Thr Leu His Ile Leu Lys Lys Arg	100	105	110
Gly Ile Ser Tyr Leu Asp Glu Asp Asp Thr Glu Ala Ala Lys Glu Ser	115	120	125
Gln Asp Tyr Lys Glu Gln Val Arg Glu Asn Ala Gln Leu Leu Thr Lys	130	135	140
Tyr Thr Pro Gly Gln Ile Gln Leu Gln Arg Leu Lys Glu Asn Asn Arg	145	150	155
Val Lys Thr Gly Ile Asn Ala Gln Gly Asn Tyr Gln Leu Asn Val Phe	165	170	175
Lys Val Ser Ala Tyr Ala Asn Glu Leu Ala Thr Ile Leu Lys Thr Gln	180	185	190
Gln Ala Phe Tyr Pro Asn Glu Leu Thr Asp Asp Trp Ile Ala Leu Phe	195	200	205
Val Gln Pro Gly Ile Ala Glu Glu Ala Gly Leu Ile Tyr Arg Lys Arg	210	215	220

Pro Tyr Tyr His Gly Pro Gly Asn Glu Ala Asn Asn Ser Pro Tyr Gly
225 230 235 240

Arg Trp Ser Asp Phe Lys Lys Thr Gly Gln Pro Ala Thr Asn Ile Phe
245 250 255

Asp Lys Leu Ile Gly Lys Asp Phe Gln Gly Glu Leu Arg Ala Ser Gly
260 265 270

Leu Ser Leu Ser Ala Gln Gln Tyr Asn Leu Leu Asn Asp Leu Thr Asn
275 280 285

Leu Lys Ile Asp Gly Glu Ile Pro Leu Ser Pro Glu Gln Lys Glu Tyr
290 295 300

Ile Leu Ala Glu Leu Met Thr Lys Glu Phe Thr Arg Phe Gly Val Asn
305 310 315 320

Asp Val Val Lys Leu Leu Gly Val Lys Lys Glu Arg Leu Ser Gly Trp
325 330 335

Arg Leu Asp Lys Lys Gly Lys Pro Glu Ile His Thr Leu Lys Gly Tyr
340 345 350

Arg Asn Trp Arg Lys Ile Phe Ala Glu Ser Gly Ile Asp Leu Ala Thr
355 360 365

Leu Pro Thr Glu Thr Ile Asp Cys Leu Ala Lys Val Leu Thr Leu Asn
370 375 380

Thr Glu Arg Glu Gly Ile Glu Asn Thr Leu Ala Phe Glu Leu Ser Glu
385 390 395 400

Leu Ala Glu Ser Val Lys Leu Leu Val Leu Asp Arg Tyr Lys Glu Leu
405 410 415

Ser Gln Ser Val Ser Thr Gln Ala Trp His Arg Phe Ser Leu Lys Thr
420 425 430

Leu His Leu Leu Ile Pro Glu Leu Met Asn Ala Thr Ser Glu Gln Asn

435

440

445

Thr Leu Leu Glu Gln Phe Gln Leu Lys Ser Asp Val Arg Lys Arg Tyr
 450 455 460

Ser Glu Tyr Lys Lys Leu Pro Thr Lys Asp Val Leu Thr Glu Ile Tyr
 465 470 475 480

Asn Pro Thr Val Asn Lys Thr Val Ser Gln Ala Phe Lys Val Ile Asp
 485 490 495

Ala Leu Leu Val Lys Tyr Gly Lys Glu Gln Ile Arg Tyr Ile Thr Ile
 500 505 510

Glu Met Pro Arg Asp Asp Asn Glu Glu Asp Glu Lys Lys Arg Ile Lys
 515 520 525

Glu Leu His Ala Lys Asn Ser Gln Arg Lys Asn Asp Ser Gln Ser Tyr
 530 535 540

Phe Met Gln Lys Ser Gly Trp Ser Gln Glu Lys Phe Gln Thr Thr Ile
 545 550 555 560

Gln Lys Asn Arg Arg Phe Leu Ala Lys Leu Leu Tyr Tyr Tyr Glu Gln
 565 570 575

Asp Gly Ile Cys Ala Tyr Thr Gly Leu Ser Ile Ser Pro Glu Leu Leu
 580 585 590

Val Ser Asp Ser Thr Glu Ile Asp His Ile Ile Pro Ile Ser Ile Ser
 595 600 605

Leu Asp Asp Ser Ile Asn Asn Lys Val Leu Val Leu Ser Lys Ala Asn
 610 615 620

Gln Val Lys Gly Gln Gln Thr Pro Tyr Asp Ala Trp Met Asp Gly Ser
 625 630 635 640

Phe Lys Lys Ile Asn Gly Lys Phe Ser Asn Trp Asp Asp Tyr Gln Lys
 645 650 655

Trp Val Glu Ser Cys His Phe Ser His Lys Lys Glu Asn Asn Leu Leu
660 665 670

Glu Thr Arg Asn Ile Phe Asp Ser Glu Gln Val Glu Lys Phe Leu Ala
675 680 685

Arg Asn Leu Asn Asp Thr Arg Tyr Ala Ser Arg Leu Val Leu Asn Thr
690 695 700

Leu Gln Ser Phe Phe Ala Asn Gln Glu Thr Lys Val Arg Val Val Asn
705 710 715 720

Gly Ser Phe Thr His Thr Leu Arg Lys Lys Trp Gly Ala Asp Leu Asp
725 730 735

Lys Thr Arg Glu Thr His His His His Ala Val Asp Ala Thr Leu Cys
740 745 750

Ala Val Thr Pro Phe Val Lys Val Ser Arg Tyr His Tyr Ala Val Lys
755 760 765

Glu Glu Thr Gly Glu Lys Val Met Arg Glu Ile Asp Phe Glu Thr Gly
770 775 780

Glu Ile Val Asp Glu Met Ser Tyr Arg Glu Phe Lys Lys Ser Lys Lys
785 790 795 800

Tyr Glu Arg Lys Thr Tyr Gln Val Lys Trp Pro Asn Phe Arg Glu Gln
805 810 815

Leu Lys Pro Val Asn Leu His Pro Arg Ile Lys Phe Ser His Gln Val
820 825 830

Asp Arg Lys Ala Asn Arg Lys Leu Ser Asp Ala Thr Ile Tyr Ser Val
835 840 845

Arg Glu Lys Thr Glu Val Lys Thr Leu Lys Ser Gly Lys Gln Lys Ile
850 855 860

Thr Thr Asp Glu Tyr Thr Ile Gly Lys Ile Lys Asp Ile Tyr Thr Val

865		870		875		880
Asp Gly Trp Glu	Ala Phe Lys Lys Lys	Gln Asp Lys Leu Leu Met Lys				
	885	890				895
Asp Leu Asp Glu Lys Thr Tyr Glu Arg Leu Leu Ser Ile Ala Glu Thr						
	900	905				910
Thr Pro Asp Phe Gln Glu Val Glu Glu Lys Asn Gly Lys Val Lys Arg						
	915	920				925
Val Lys Arg Ser Pro Phe Ala Val Tyr Cys Glu Glu Asn Asp Ile Pro						
	930	935				940
Ala Ile Arg Lys Tyr Ala Lys Lys Asn Asn Gly Pro Leu Ile Arg Ser						
945	950	955				960
Leu Lys Tyr Tyr Asp Gly Lys Leu Asn Lys His Ile Asn Ile Thr Lys						
	965	970				975
Asp Ser Gln Gly Arg Pro Val Glu Lys Thr Lys Asn Gly Arg Lys Val						
	980	985				990
Thr Leu Gln Ser Leu Lys Pro Tyr Arg Tyr Asp Ile Tyr Gln Asp Leu						
	995	1000				1005
Glu Thr Lys Ala Tyr Tyr Thr Val Gln Leu Tyr Tyr Ser Asp Leu						
1010	1015					1020
Arg Phe Val Glu Gly Lys Tyr Gly Ile Thr Glu Lys Glu Tyr Met						
1025	1030					1035
Lys Lys Val Ala Glu Gln Thr Lys Gly Gln Val Val Arg Phe Cys						
1040	1045					1050
Phe Ser Leu Gln Lys Asn Asp Gly Leu Glu Ile Glu Trp Lys Asp						
1055	1060					1065
Ser Gln Cys Tyr Asp Val Arg Phe Tyr Asn Phe Gln Ser Ala Asn						
1070	1075					1080

Ser Ile Asn Phe Lys Gly Leu Glu Gln Glu Met Met Pro Ala Glu
1085 1090 1095

Asn Gln Phe Lys Gln Lys Pro Tyr Asn Asn Gly Ala Ile Asn Leu
1100 1105 1110

Asn Ile Ala Lys Tyr Gly Lys Glu Gly Lys Lys Leu Arg Lys Phe
1115 1120 1125

Asn Thr Asp Ile Leu Gly Lys Lys His Tyr Leu Tyr Tyr Glu Lys
1130 1135 1140

Glu Pro Lys Asn Ile Ile Lys
1145 1150

<210> 116
<211> 1408
<212> PRT
<213> Sphingobacterium sp.

<223> APG02675.1_RGN

<400> 116

Met Lys Arg Ile Leu Gly Leu Asp Leu Gly Thr Ser Ser Ile Gly Trp
1 5 10 15

Ala Tyr Val Leu Glu Ala Gln Asp Glu Asn Glu Ile Ser Lys Ile Glu
20 25 30

Lys Leu Gly Val Arg Val Asn Pro Leu Thr Thr Asp Glu Gln Leu Asn
35 40 45

Phe Glu Lys Gly Lys Pro Ile Thr Thr Asn Ala Gly Arg Thr Leu Ala
50 55 60

Arg Ser Ala Arg Arg Asn Leu Gln Arg Phe Lys Leu Arg Arg Asn Asn
65 70 75 80

Leu Ile Glu Val Leu Lys Arg Glu Lys Trp Ile Asp Asp Lys Ser Ile
85 90 95

Leu Ala Glu Asn Gly Asn Lys Ser Thr Phe Glu Thr Tyr Ala Leu Arg
100 105 110

Ala Ala Ala Ala Ser Lys Glu Ile Lys Leu Glu Glu Leu Ala Arg Val
115 120 125

Leu Leu Met Ile Asn Lys Lys Arg Gly Tyr Lys Ser Ser Arg Lys Val
130 135 140

Ile Lys Glu Asp Asp Gly Gln Leu Val Asp Gly Met Glu Val Ala Lys
145 150 155 160

Lys Leu Tyr His Glu Asn Leu Thr Pro Gly Glu Phe Val Cys Arg Leu
165 170 175

Leu Lys Glu Gly Lys Lys Lys Leu Pro Asp Phe Tyr Arg Ser Asp Leu
180 185 190

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

Glu Leu Leu Thr Arg Asp Phe Lys Glu Glu Leu Lys Gly Lys Gly Leu
210 215 220

Arg Ala Thr Ser Ala Ile Phe Trp Ser Lys Tyr Lys Phe Asn Thr Ala
225 230 235 240

Glu Asn Lys Gly Thr Lys Glu Glu Lys Lys Ile Gln Ala Tyr Gln Trp
245 250 255

Arg Lys Asp Gly Leu Thr Lys Gln Leu Ala Lys Glu Glu Val Ala Tyr
260 265 270

Ile Ile Ala Glu Ile Asn Gly Gln Leu Ala Asn Ser Ser Gly Tyr Leu
275 280 285

Gly Ala Ile Ser Asp Arg Ser Lys Glu Leu Tyr Phe Asn Lys Gln Thr
290 295 300

Val Gly Gln Phe Leu Phe Asn Gln Leu Gln Gln Asn Thr His Ala Arg
305 310 315 320

Leu Lys Asn Gln Val Phe Tyr Arg Gln Asp Tyr Leu Asp Glu Phe Glu
325 330 335

Lys Ile Trp Cys Ile Gln Ser Gln Phe His Gln Ala Leu Thr Glu Gln
340 345 350

Leu Lys Glu Glu Ile Arg Asp Ile Ile Ile Phe Tyr Gln Arg Lys Leu
355 360 365

Lys Ser Gln Lys Gly Leu Ile Ser Phe Cys Glu Phe Glu Gln His Glu
370 375 380

Val Val Val Asn Gly Lys Asn Lys Val Val Gly Leu Arg Val Ala Pro
385 390 395 400

Lys Ser Ser Pro Ile Phe Gln Glu Phe Lys Ile Trp Gln Gln Leu Asn
405 410 415

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

Glu Glu Gln Lys Lys Leu Leu Phe Glu Thr Leu Asn Leu Lys Gly Lys
435 440 445

Leu Ser Ala Tyr Gln Val Leu Ser Leu Ile Glu Gly Lys Pro Lys Glu
450 455 460

Trp Glu Leu Asn Tyr Thr Glu Leu Glu Gly Asn Ser Thr Asn Ser Ala
465 470 475 480

Leu Tyr Asn Ala Tyr Leu Asn Ile Leu Asp Ile Glu Gly Tyr Asp Val
485 490 495

Arg Ser Glu Leu Lys Ile Lys Leu Asn Lys Asp Glu Ile Thr Leu Ser
500 505 510

Asp Leu Asp Ile Pro Val Ser Glu Ile Lys Asp Met Ile Arg Arg Ile
515 520 525

Phe Gln His Leu Gly Ile Asn Thr Ser Ile Leu Asp Phe Asp Ala Gln
530 535 540

Met Gln Gly Asp Asp Phe Glu Lys Gln Leu Ser Tyr Gln Leu Trp His
545 550 555 560

Leu Leu Tyr Ser Tyr Glu Glu Asp Asp Ser Lys Thr Gly Met Asp Arg
565 570 575

Leu Tyr Asn Gln Leu Asn Leu Lys Phe Gly Phe Ser Leu Asp Gln Ala
580 585 590

Lys Phe Ile Gly Lys Ile Ala Leu Gln Asp Asp Tyr Gly Asn Leu Ser
595 600 605

Thr Lys Ala Ile Arg Gln Ile Tyr Pro Tyr Ile Gln Asp Ala Glu Tyr
610 615 620

Ser Ala Ala Cys Lys Leu Ala Gly Tyr Asn His Ser Lys Ser Ser Leu
625 630 635 640

Thr Lys Glu Gln Leu Thr Asn Arg Val Leu Lys Asp His Leu Asp Ile
645 650 655

Leu Pro Lys Asn Ser Leu Arg Asn Pro Val Val Glu Lys Ile Leu Asn
660 665 670

Gln Met Val Asn Val Val Asn Thr Leu Ile Asp Thr Glu Asn Asp Lys
675 680 685

Leu Ile Lys Glu Gly Lys Asn Ala Asp Phe Arg Phe Asp Glu Ile Arg
690 695 700

Ile Glu Leu Ala Arg Glu Leu Lys Lys Asn Ala Lys Glu Arg Glu Glu
705 710 715 720

Leu Thr Lys Ala Ile Asn Thr Ser Lys Glu Gln His Glu Lys Ile Ile
725 730 735

Lys Ile Leu Gln Thr Glu Asp Gly Ile Lys Asn Pro Thr Arg Asn Asp
740 745 750

Ile Ile Arg Phe Lys Leu Tyr Gln Glu Leu Lys Asn Asn Gly Tyr Lys
755 760 765

Asn Leu Tyr Thr Asn Glu Tyr Ile Gln Arg Lys Asp Leu Phe Thr Asn
770 775 780

Ile Tyr Asp Ile Asp His Ile Val Pro Gln Ser Arg Leu Phe Asp Asp
785 790 795 800

Ser Phe Ser Asn Lys Val Leu Val Pro Arg Asn Ile Asn Ile Glu Lys
805 810 815

Gly Asn Gln Thr Ala Phe Asp Tyr Val Arg Thr Lys Phe Gly Glu Asp
820 825 830

Gly Ile Glu Ala Tyr Glu Ala Arg Val Glu Arg Leu Phe Asn Leu Lys
835 840 845

Glu Glu Gly Val Ser Arg Ser Lys Tyr Lys Lys Leu Leu Met Arg Gly
850 855 860

Ser Asp Ile Gly Glu Gly Phe Ile Glu Arg Asp Leu Arg Asp Ser Gln
865 870 875 880

Tyr Ile Ala Lys Lys Ala Lys Ala Met Leu Phe Glu Ile Thr Pro Ser
885 890 895

Val Ile Ser Thr Ser Gly Ser Val Thr Asp Arg Leu Arg Glu Asp Trp
900 905 910

Gly Leu Val Ser Val Met Lys Glu Leu Asn Leu Pro Lys Phe Lys Ala
915 920 925

Val Gly Leu Thr Glu Tyr Leu Glu Thr Lys Asp Gly Asn Arg Lys Glu
930 935 940

Val Ile Lys Asp Trp Ser Lys Arg Asn Asp His Arg His His Ala Met
945 950 955 960

Asp Ala Leu Thr Val Ala Phe Thr Lys His Ser His Ile Gln Tyr Leu
965 970 975

Asn His Leu Asn Ala Arg Lys Asn Glu Lys Ser Glu Phe Phe Ser Ser
980 985 990

Ile Lys Ala Ile Glu Ile Lys Glu Thr Tyr Val Glu Arg Asp Asp Leu
995 1000 1005

Gly Asn Arg Lys Arg Leu Phe Lys Glu Pro Ile Ala His Phe Arg
1010 1015 1020

Thr Ile Val Lys Glu His Leu Glu Arg Val Leu Val Ser His Lys
1025 1030 1035

Ala Lys Asn Lys Val Val Thr Lys Asn Arg Asn Lys Ile Asp Gly
1040 1045 1050

Lys Lys Glu Ala Gln Glu Val Leu Thr Pro Arg Gly Gln Leu His
1055 1060 1065

Lys Glu Thr Ile Tyr Gly Lys Ile Leu Gln Tyr Ala Ser Lys Glu
1070 1075 1080

Glu Lys Ile Ser Ala Lys Phe Asp Ala Ala Thr Ile Ala Met Val
1085 1090 1095

Ser Asn Pro Arg Tyr Arg Ser Ala Leu Leu Asp Arg Leu Ile Glu
1100 1105 1110

Tyr Gly Asn Asp Pro Lys Lys Ala Phe Thr Gly Lys Asn Ser Pro
1115 1120 1125

Ser Lys Ser Pro Ile Tyr Leu Asp Gln Gln Arg Gln Leu Ala Val
1130 1135 1140

Pro Glu Lys Val Lys Leu Val Trp Leu Glu Ala Asp Tyr Thr Ile
1145 1150 1155

Arg Lys Glu Ile Gly Pro Asp Leu Lys Ile Glu Lys Val Ile Asp
1160 1165 1170

Gln Gly Val Lys Arg Ile Leu Lys Gln Arg Leu Ser Val Tyr Asp
1175 1180 1185

Gly Asp Thr Lys Lys Ala Phe Ser Asp Leu Asp Asn Asn Pro Ile
1190 1195 1200

Trp Leu Asn Glu Glu Lys Gly Ile Ala Ile Lys Arg Val Thr Ile
1205 1210 1215

Ser Gly Val Lys Asn Ala Glu Ala Leu His Thr Lys Lys Asp His
1220 1225 1230

Leu Gly Gln Val Leu Glu Asp Lys Asp Gly Asn Thr Met Pro Val
1235 1240 1245

Asp Phe Val Ser Thr Gly Asn Asn His His Val Ala Ile Tyr Glu
1250 1255 1260

Asp Ser Glu Gly Asn Leu Tyr Asp Arg Val Val Ser Phe Tyr Glu
1265 1270 1275

Ala Val Glu Arg Ile Asn Gln Asn Leu Pro Ile Val Asp Arg His
1280 1285 1290

Tyr Lys Gln Glu Glu Gly Trp Lys Phe Leu Phe Ser Met Lys Gln
1295 1300 1305

Asn Glu Leu Phe Val Phe Pro Ser Ile Asn Phe Asp Pro Lys Glu
1310 1315 1320

Ile Asp Leu Leu Asp Lys Lys Asn Tyr Lys Glu Val Ser Lys His
1325 1330 1335

Leu Phe Arg Met Gln Lys Phe Ser Lys Val Glu Tyr Gly Asn Ser
1340 1345 1350

Ala Val Arg Asp Tyr Val Phe Arg His His Leu Glu Thr Ser Ile
1355 1360 1365

Ile Asp Thr Lys Glu Leu Arg Asp Ile Ala Tyr Lys Val Phe Lys
1370 1375 1380

Ser Ile Gly Glu Phe Lys Ser Phe Val Lys Ile Arg Thr Asn His
1385 1390 1395

Leu Gly Glu Ile Val His Leu Gly Glu Tyr
1400 1405

<210> 117
<211> 1369
<212> PRT
<213> Chryseobacterium sp.

<223> APG01405.1_RGN

<400> 117

Ala Arg Thr Asp Tyr Arg Gly Ile Arg Arg Leu Arg Glu Arg Phe Leu
1 5 10 15

Leu Arg Arg Glu Arg Leu His Arg Ile Leu Asn Val Leu Asn Phe Leu
20 25 30

Pro Glu His Tyr Ala Ser Gln Ile Asp Phe Asp Lys Arg Phe Gly Lys
35 40 45

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

Lys Phe Ser Phe Leu Phe Gln Thr Ser Phe Asn Glu Met Leu Glu Asp
65 70 75 80

Phe Lys Ala His Gly Gln Asp Leu Lys Val Pro Tyr Asp Trp Thr Ile
85 90 95

Tyr Tyr Leu Arg Lys Lys Ala Leu Ser Gln Lys Ile Glu Lys Glu Glu
100 105 110

Leu Ala Trp Ile Leu Leu Asn Phe Asn Gln Lys Arg Gly Tyr Tyr Gln
115 120 125

Leu Arg Gly Glu Glu Glu Glu Glu Asn Pro Asn Lys Leu Ile Glu Phe

130

135

140

Tyr Ser Leu Lys Ile Ile Asp Val Leu Ala Asp Glu Pro Gln Lys Gly
 145 150 155 160

Lys Ser Asp Ile Trp Tyr Ser Leu Val Leu Glu Asn Gly Trp Ile Tyr
 165 170 175

Arg Arg Ser Ser Lys Thr Ser Leu Leu Asp Trp Lys Asp Lys Ile Arg
 180 185 190

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

Asp Lys Glu Gly Asn Glu Lys Arg Ser Phe Arg Ala Pro Gly Glu Asn
 210 215 220

Asp Trp Thr Leu Val Lys Thr Lys Thr Glu Gln Glu Ile Asp Arg Ser
 225 230 235 240

Arg Lys Thr Val Gly Thr Tyr Ile Tyr Glu Thr Leu Leu Gln Asn Pro
 245 250 255

Lys Gln Lys Ile Lys Gly Lys Leu Val Arg Thr Ile Glu Arg Lys Phe
 260 265 270

Tyr Lys Glu Glu Leu Lys Gln Ile Leu Glu Lys Gln Lys Glu Phe His
 275 280 285

Tyr Glu Leu Gln Ser Asp Asp Leu Tyr Asp Asp Cys Ile Arg Glu Leu
 290 295 300

Tyr Arg Asn Asn Glu Ala His Gln Leu Thr Leu Ser Lys Lys Asp Phe
 305 310 315 320

Val His Leu Phe Met Glu Asp Ile Ile Phe Tyr Gln Arg Pro Leu Lys
 325 330 335

Ser Gln Lys Ser Ser Ile Ser Asn Cys Thr Leu Glu Phe Arg Lys Tyr
 340 345 350

Lys Asp Glu Asn Gly Val Glu His Thr Gln Phe Leu Lys Ala Ile Pro
355 360 365

Lys Ser Asn Pro Tyr Tyr Gln Glu Phe Arg Ile Trp Gln Trp Ile Phe
370 375 380

Asn Leu Asn Ile Tyr Lys Lys Asp Asn Asp Asp Asn Val Thr Lys Glu
385 390 395 400

Phe Leu Ser Thr Thr Glu Asp Phe Glu Asn Leu Phe Glu Phe Leu Asn
405 410 415

Asn Arg Lys Glu Ile Asp Gln Lys Ala Leu Leu Lys His Phe Lys Leu
420 425 430

Asn Glu Lys Thr His Arg Trp Lys Tyr Val Glu Asp Lys Lys Tyr Pro
435 440 445

Cys Asn Glu Thr Lys Thr Met Ile Ser Glu Arg Leu Lys Lys Val Glu
450 455 460

Asn Ile Ser Asp Asp Phe Leu Thr Arg Gly Met Glu Gln Lys Ile Trp
465 470 475 480

His Ile Ile Tyr Ser Val Asn Asp Lys Thr Glu Tyr Glu Lys Ala Leu
485 490 495

Lys Ser Phe Ala Glu Lys Asn Asn Leu Asp Glu Asn Ser Phe Phe Glu
500 505 510

Ala Phe Arg Lys Phe Pro Pro Phe Lys Ser Glu Tyr Gly Ser Phe Ser
515 520 525

Glu Lys Ala Ile Lys Lys Leu Leu Pro Leu Met Arg Leu Gly Lys Tyr
530 535 540

Trp Ser Tyr Ala Asn Ile Asp Leu Tyr Ser Lys Asn Arg Ile Gln Lys
545 550 555 560

Ile Ile Thr Gly Glu Phe Asp Glu Asn Ile Lys Asp Arg Val Arg Glu

565

570

575

Lys Ala Ile His Leu Thr Ala Glu Asn Asp Phe Gln Gly Leu Gln Leu
 580 585 590

Trp Leu Ala Gln Tyr Ile Val Tyr Gly Arg His Ser Glu Ala Thr Met
 595 600 605

Ile Gly Lys Trp Asn Ser Ala Asp Asp Leu Glu Glu Phe Leu Lys Glu
 610 615 620

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

Glu Thr Leu Arg Val Val Lys Asp Ile Trp Leu Lys Tyr Gly Asn Gly
 645 650 655

Ala Lys Asp Phe Phe Asn Glu Val His Ile Glu Leu Gly Arg Glu Met
 660 665 670

Lys Gln Thr Lys Asp Glu Arg Ala Asn Ala Thr Lys Thr Ile Thr Glu
 675 680 685

Asn Glu Asn Thr Asn Leu Arg Ile Lys Ala Leu Leu Ala Glu Met Met
 690 695 700

Asn Asp His Ser Val Glu Asn Val Arg Pro Tyr Ser Pro Met Gln Gln
 705 710 715 720

Glu Ile Leu Lys Ile Tyr Glu Asp Gly Ile Leu Lys Ser Asp Ile Glu
 725 730 735

Ile Asp Asp Asp Ile Leu Lys Ile Ser Lys Thr Ala Gln Pro Ser Ser
 740 745 750

Ser Asp Leu Lys Arg Tyr Lys Leu Trp Leu Glu Gln Lys Tyr Lys Ser
 755 760 765

Pro Tyr Thr Gly Gln Ile Ile Pro Leu Asn Lys Leu Phe Thr Pro Glu
 770 775 780

Tyr Glu Ile Glu His Ile Ile Pro Gln Ser Arg Tyr Phe Asp Asp Ser
785 790 795 800

Leu Ser Asn Lys Ile Ile Cys Glu Ser Ala Val Asn Lys Leu Lys Asp
805 810 815

Asn Tyr Ile Gly Leu Gly Phe Ile Lys Gln Phe Gly Gly Thr Ile Ile
820 825 830

Glu Cys Gly Phe Gly Lys Arg Val Lys Val Phe Lys Ala Glu Glu Tyr
835 840 845

Glu Glu Phe Val Lys Lys His Tyr Ala Asn Asn Arg Gly Lys Arg Asn
850 855 860

Lys Leu Leu Leu Glu Asp Ile Pro Glu Lys Met Ile Glu Arg Gln Met
865 870 875 880

Asn Asp Thr Arg His Ile Ser Lys Tyr Ile Ser Gly Ile Leu Ser His
885 890 895

Ile Val Arg Val Glu Asp Gly Thr Asp Glu Gly Val Asn Ser Lys Asn
900 905 910

Val Ile Pro Gly Asn Gly Lys Ile Thr Thr Gln Leu Arg Gln Asp Trp
915 920 925

Gly Leu Asn Asp Val Trp Asn Glu Leu Val Leu Pro Arg Phe Glu Arg
930 935 940

Met Asn Gln Leu Thr Asn Ser Thr Asp Phe Thr Ser Trp Asn Glu Asn
945 950 955 960

His Gln Lys Tyr Leu Pro Thr Val Pro Val Glu Phe Ser Lys Gly Phe
965 970 975

Ser Lys Lys Arg Ile Asp His Arg His His Ala Leu Asp Ala Leu Val
980 985 990

Ile Ala Cys Ala Thr Lys Asp His Val Asn Leu Leu Asn Asn Gln Ser

995					1000					1005				
Ala	Lys	Ser	Asp	Thr	Lys	Arg	Tyr	Asp	Leu	Lys	Lys	Lys	Leu	Met
	1010					1015					1020			
Lys	Phe	Glu	Lys	Thr	Val	Tyr	Lys	Asp	Pro	Gln	Thr	Glu	Lys	Arg
	1025					1030					1035			
Ile	Glu	Arg	Glu	Val	Pro	Lys	Tyr	Phe	Leu	Lys	Pro	Trp	Glu	Thr
	1040					1045					1050			
Phe	Thr	Val	Asp	Ala	Lys	Asn	Lys	Leu	Glu	Thr	Ile	Ile	Val	Ser
	1055					1060					1065			
Phe	Lys	Gln	Asn	Leu	Arg	Val	Ile	Asn	Lys	Ala	Thr	Asn	Tyr	Tyr
	1070					1075					1080			
Glu	Lys	Tyr	Val	Asp	Lys	Asp	Gly	Val	Lys	Thr	Lys	Glu	Arg	Val
	1085					1090					1095			
Glu	Gln	Thr	Gly	Thr	Asn	Trp	Ala	Ile	Arg	Lys	Pro	Met	His	Lys
	1100					1105					1110			
Glu	Thr	Val	Ser	Gly	Lys	Ile	Asp	Leu	Pro	Trp	Val	Lys	Val	Pro
	1115					1120					1125			
Lys	Gly	Lys	Ile	Leu	Thr	Ala	Thr	Arg	Lys	Ser	Leu	Asp	Thr	Ser
	1130					1135					1140			
Phe	Asp	Leu	Lys	Ala	Ile	Ala	Ser	Ile	Thr	Asp	Thr	Gly	Ile	Gln
	1145					1150					1155			
Lys	Ile	Leu	Lys	Asn	Tyr	Leu	Glu	Phe	Lys	Glu	Ser	Pro	Glu	Leu
	1160					1165					1170			
Ala	Phe	Ser	Pro	Glu	Gly	Ile	Glu	Asp	Met	Asn	Lys	Asn	Ile	Lys
	1175					1180					1185			
Lys	Tyr	Asn	Gly	Gly	Lys	Pro	His	Gln	Pro	Ile	Ser	Lys	Val	Arg
	1190					1195					1200			

Val Phe Glu Leu Gly Ser Lys Phe Gln Val Gly Gln Thr Gly Asn
1205 1210 1215

Lys Lys Asp Lys Tyr Val Glu Ala Ala Lys Gly Thr Asn Leu Phe
1220 1225 1230

Phe Ala Ile Tyr Glu Asp Arg Lys Gly Lys Arg Ser Tyr Glu Thr
1235 1240 1245

Ile Pro Leu Asn Glu Val Ile Glu Arg Gln Lys Gln Gly Leu Ser
1250 1255 1260

Val Val Asp Leu Lys Asn Ile Asn Asp Phe Phe Leu Cys Pro Asn
1265 1270 1275

Asp Leu Val Tyr Ile Pro Ser Gly Asp Glu Leu Glu Asn Gly Gly
1280 1285 1290

Ser Ile Glu Ile Lys Asn Ile Thr Lys Glu Lys Ser Glu Arg Phe
1295 1300 1305

Tyr Lys Val Val Ser Phe Ser Gly Ser Gln Ile Phe Phe Val Arg
1310 1315 1320

His Asp Ile Ala Val Ser Ile Val Asn Lys Gly Glu Phe Ser Thr
1325 1330 1335

Leu Asn Lys Met Glu Arg Ala Ile Asp Gly Ser Met Val Lys Glu
1340 1345 1350

Ser Cys Ile Lys Leu Lys Ile Asp Arg Leu Gly Asn Val Leu Lys
1355 1360 1365

Ala

<210> 118

<211> 1440

<212> PRT

<213> Chryseobacterium sp.

<223> APG06250.1_RGN

<400> 118

Met Lys Asn Ile Leu Gly Leu Asp Leu Gly Thr Asn Ser Ile Gly Trp
1 5 10 15

Ala Leu Val Lys Gln Asp Phe Lys Asn Lys Gln Gly Glu Ile Leu Gly
20 25 30

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

Gly Lys Gly Asn Ser Val Ser Gln Thr Ala Glu Arg Thr Lys Tyr Arg
50 55 60

Ser Ala Arg Arg Leu Arg Glu Arg Phe Leu Leu Arg Arg Glu Arg Leu
65 70 75 80

His Arg Val Leu Asn Ile Leu Asn Phe Leu Pro Glu His Tyr Ala Ser
85 90 95

Gln Ile Asp Phe Glu Lys Arg Phe Gly Lys Phe Lys Val Glu Thr Glu
100 105 110

Pro Lys Leu Ala Trp Lys Asn Ile Glu Gly Gln Phe Ser Phe Leu Phe
115 120 125

Gln Asn Ser Phe Asn Glu Met Leu Glu Asp Phe Lys Ala Asn Ala Gln
130 135 140

Asp Leu Arg Ile Pro Tyr Asp Trp Thr Ile Tyr Tyr Leu Arg Lys Lys
145 150 155 160

Ala Leu Ser Gln Lys Ile Glu Lys Glu Glu Leu Val Trp Ile Leu Leu
165 170 175

Asn Phe Asn Gln Lys Arg Gly Tyr Tyr Gln Leu Arg Gly Glu Glu Glu
180 185 190

Glu Glu Asn Pro Asn Lys Leu Val Glu Leu Tyr Ser Leu Lys Val Val
195 200 205

Asp Val Phe Val Asp Glu Pro Gln Lys Gly Lys Ser Asp Ile Trp Tyr
210 215 220

Ser Leu Val Leu Glu Asn Gly Trp Ile Tyr Arg Arg Ala Ser Lys Ile
225 230 235 240

Ser Leu Phe Asp Trp Lys Asp Lys Ile Arg Asp Phe Ile Val Thr Thr
245 250 255

Asp Leu Asn Asp Asp Gly Ser Val Lys Thr Asp Lys Asp Gly Asn Glu
260 265 270

Lys Arg Ser Phe Arg Ala Pro Gly Glu Asn Asp Trp Thr Leu Val Lys
275 280 285

Lys Lys Thr Glu Gln Glu Ile Glu Gln Ser His Lys Thr Val Gly Met
290 295 300

Tyr Ile Tyr Glu Thr Leu Arg Ala Asn Pro Lys Gln Lys Ile Lys Gly
305 310 315 320

Lys Leu Val Arg Thr Ile Glu Arg Lys Phe Tyr Lys Glu Glu Leu Arg
325 330 335

Gln Ile Leu Glu Lys Gln Lys Glu Phe His Gln Glu Leu Gln Asn Asp
340 345 350

Asp Leu Tyr Asn Asp Cys Ile Arg Glu Leu Tyr Arg Asn Asn Glu Ala
355 360 365

His Gln Leu Thr Leu Ser Lys Lys Asp Phe Val His Leu Phe Ile Glu
370 375 380

Asp Ile Val Phe Tyr Gln Arg Pro Leu Arg Ser Gln Lys Ser Ser Ile
385 390 395 400

Ser Asn Cys Thr Leu Glu Phe Arg Lys Tyr Lys Asp Glu Asn Gly Ala
405 410 415

Glu His Thr Gln Tyr Leu Lys Ala Ile Pro Lys Ser Asn Pro Tyr Tyr
420 425 430

Gln Glu Phe Arg Leu Trp Gln Trp Ile Phe Asn Leu Asn Leu Tyr Lys
435 440 445

Lys Asp Asn Asp Glu Asn Val Ile Lys Glu Phe Leu Thr Thr Thr Gln
450 455 460

Asp Val Glu Asn Leu Phe Glu Phe Leu Asn Asn Arg Lys Glu Ile Asp
465 470 475 480

Gln Lys Ala Leu Leu Lys His Phe Lys Leu Ser Glu Lys Thr His Arg
485 490 495

Trp Asn Phe Val Glu Asp Lys Lys Tyr Pro Cys Asn Glu Thr Lys Thr
500 505 510

Met Ile Ala Thr Arg Leu Glu Lys Val Glu Asn Ile Ser Asp Asp Phe
515 520 525

Leu Thr Arg Glu Ile Glu Gln Lys Ile Trp His Ile Ile Tyr Ser Val
530 535 540

Asn Asp Lys Ile Glu Tyr Glu Lys Ala Leu Lys Ser Phe Ala Leu Lys
545 550 555 560

His Asn Leu Asp Glu Asn Ser Phe Phe Glu Ala Phe Arg Lys Phe Pro
565 570 575

Pro Phe Lys Ser Glu Tyr Gly Ser Phe Ser Glu Lys Ala Ile Lys Lys
580 585 590

Leu Leu Pro Leu Met Arg Leu Gly Lys Tyr Trp Asp Tyr Ala Asn Ile
595 600 605

Asp Gln Phe Ser Lys Gly Arg Ile Gln Lys Ile Ile Asn Gly Glu Tyr
610 615 620

Asp Glu Asn Ile Lys Asp Arg Val Arg Glu Lys Ala Val His Leu Thr
625 630 635 640

Ser Glu Asn Asp Phe Gln Gly Leu Gln Leu Trp Leu Ala Gln Tyr Ile
645 650 655

Val Tyr Gly Arg His Ser Glu Ala Ser Ile Ala Gly Lys Trp Asn Ser
660 665 670

Ala Asp Asp Leu Glu Glu Phe Leu Lys Asp Phe Lys His His Ser Leu
675 680 685

Arg Asn Pro Ile Val Glu Gln Val Ile Thr Glu Thr Leu Arg Val Val
690 695 700

Lys Asp Ile Trp Leu Lys Tyr Gly Lys Gly Ala Lys Asp Phe Phe Asn
705 710 715 720

Glu Ile His Ile Glu Leu Gly Arg Glu Met Lys Leu Pro Ala Asp Asp
725 730 735

Arg Lys Lys Leu Thr Asn Gln Ile Thr Glu Asn Glu Asn Thr Asn Leu
740 745 750

Arg Ile Lys Ala Leu Leu Ala Glu Met Met Asn Asp His Gly Val Glu
755 760 765

Asn Val Arg Pro Phe Ser Pro Gln Gln Gln Glu Ile Leu Lys Ile Tyr
770 775 780

Glu Asp Gly Val Leu Asn Ser Asp Ile Glu Ile Glu Asp Glu Tyr Leu
785 790 795 800

Lys Ile Ser Lys Thr Ala Gln Pro Ser Pro Ser Asp Leu Lys Arg Tyr
805 810 815

Lys Leu Trp Leu Glu Gln Lys Tyr Lys Ser Pro Tyr Thr Gly Gln Ile
820 825 830

Ile Pro Leu Asn Lys Leu Phe Thr Pro Glu Tyr Glu Ile Glu His Ile
835 840 845

Ile Pro Gln Ser Arg Tyr Phe Asp Asp Ser Phe Ser Asn Lys Ile Ile
850 855 860

Cys Glu Ser Ala Val Asn Lys Leu Lys Asp Asn Tyr Ile Gly Leu Gly
865 870 875 880

Phe Ile Lys Gln Phe Gly Gly Thr Ile Ile Glu Leu Gly Phe Gly Lys
885 890 895

Ser Val Lys Val Phe Asp Ile Asp Glu Tyr Glu Asp Phe Val Lys Lys
900 905 910

His Tyr Ala Asn Asn Arg Ser Lys Arg Asn Lys Leu Leu Leu Glu Asp
915 920 925

Ile Pro Glu Lys Met Ile Glu Arg Gln Ile Asn Asp Thr Arg Tyr Ile
930 935 940

Ser Lys Tyr Ile Ser Gly Ile Leu Ser Asn Ile Val Arg Leu Glu Asp
945 950 955 960

Gly Ser Asp Glu Gly Ile Asn Ser Lys Asn Ile Val Pro Gly Asn Gly
965 970 975

Lys Ile Thr Thr Gln Leu Lys Lys Asp Trp Gly Leu Asn Asp Val Trp
980 985 990

Asn Asp Leu Ile Leu Pro Arg Phe Lys Arg Met Asn Gln Leu Thr Asn
995 1000 1005

Ser Thr Asp Phe Ile Ala Trp Asn Glu Asn Phe Gln Lys Phe Leu
1010 1015 1020

Pro Thr Val Pro Ile Glu Tyr Ser Lys Gly Phe Ser Lys Lys Arg
1025 1030 1035

Ile Asp His Arg His His Ala Leu Asp Ala Leu Val Ile Ala Cys
1040 1045 1050

Ala Thr Lys Asp His Val Asn Leu Leu Asn Asn Gln Ser Ala Lys
1055 1060 1065

Ser Glu Thr Lys Arg Tyr Asp Leu Lys Lys Lys Leu Met Lys Phe
1070 1075 1080

Glu Arg Val Val Tyr Leu His Thr Gln Thr Gly Glu Lys Ile Glu
1085 1090 1095

Arg Glu Val Pro Lys His Phe Leu Lys Pro Trp Glu Asn Phe Thr
1100 1105 1110

Val Asp Val Lys His Asn Leu Asp Thr Val Ile Val Ser Phe Lys
1115 1120 1125

Gln Asn Leu Arg Val Ile Asn Lys Ala Thr Asn Tyr Tyr Glu Lys
1130 1135 1140

Tyr Ala Glu Asn Gly Gly Ala Arg Asn Lys Ala Arg Val Glu Gln
1145 1150 1155

Lys Gly Ile Asn Trp Ala Ile Arg Lys Pro Met His Lys Asp Thr
1160 1165 1170

Val Ser Gly Lys Val Asp Leu Pro Trp Val Lys Val Pro Lys Gly
1175 1180 1185

Lys Ile Leu Thr Ala Thr Arg Lys Ser Leu Asp Thr Ser Phe Asp
1190 1195 1200

Leu Lys Ser Ile Gly Ser Ile Thr Asp Thr Gly Ile Gln Lys Ile
1205 1210 1215

Leu Lys Asn Tyr Leu Met Leu Lys Asp Gly Asn Pro Glu Leu Ala
1220 1225 1230

Phe Ser Pro Glu Gly Ile Glu Asp Leu Asn Lys Ser Ile Glu Lys
1235 1240 1245

Tyr Asn Asp Gly Lys Pro His Gln Pro Ile Asn Lys Val Arg Val
1250 1255 1260

Phe Glu Leu Gly Ser Lys Phe Gln Ile Gly Gln Ala Gly Asn Lys
1265 1270 1275

Lys Asp Lys Tyr Val Glu Ala Ala Lys Gly Thr Asn Leu Phe Phe
1280 1285 1290

Ala Val Tyr Glu Asp Lys Asn Gly Lys Arg Ser Tyr Glu Thr Ile
1295 1300 1305

Pro Leu Asn Glu Val Ile Glu Arg Gln Lys Gln Gly Leu Pro Val
1310 1315 1320

Val Asp Leu Arg Ser Thr Asn Asp Phe Tyr Leu Cys Pro Asn Asp
1325 1330 1335

Leu Val Tyr Ile Leu Ser Asp Asp Glu Arg Gly Asn Met Thr Asn
1340 1345 1350

Asn Asp Phe Glu Asn Leu Ser Asp Glu Gln Val Lys Arg Ile Tyr
1355 1360 1365

Lys Phe Val Ser Cys Thr Gly Gly Glu Gly His Phe Ile Pro Tyr
1370 1375 1380

Ala Asn Ala Thr Glu Ile Ile Lys Asn Glu Asn Gly Thr Asn Ser
1385 1390 1395

Lys Ser Glu Arg Met Gln Asn Phe Tyr Asp Gly Ser Cys Met Leu
1400 1405 1410

Asp Lys Asn Ser Lys Pro Ile Met Ile Lys Glu Asn Cys Val Lys
1415 1420 1425

Leu Lys Val Asp Arg Leu Gly Asn Ile Ser Lys Ile
1430 1435 1440

<210> 119

<211> 1071

<212> PRT

<213> Bacillus sp.

<223> APG06877.1_RGN

<400> 119

Met Arg Glu Leu Asp Tyr Arg Ile Gly Leu Asp Ile Gly Thr Asn Ser
1 5 10 15

Ile Gly Trp Gly Val Ile Glu Leu Ser Trp Asn Lys Asp Arg Glu Gln
20 25 30

Tyr Glu Lys Thr Arg Ile Val Asp Lys Gly Val Arg Met Phe Asp Lys
35 40 45

Ala Glu Ile Pro Lys Thr Gly Ala Ser Leu Ala Glu Pro Arg Arg Ile
50 55 60

Ala Arg Ser Ser Arg Arg Arg Leu Asn Arg Lys Ser Gln Arg Lys Lys
65 70 75 80

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

Leu Thr Ser Leu Tyr Pro Leu Ser Lys Ser Ser Met Asp Ile Trp Asp
100 105 110

Ile Arg Leu Asp Gly Leu Asp Arg Leu Leu Asp Arg Phe Glu Trp Ala
115 120 125

Arg Leu Leu Ile His Leu Ala Gln Arg Arg Gly Phe Lys Ser Asn Arg
130 135 140

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

Gln Val Asn Glu Lys Arg Leu Ser Leu Tyr Arg Thr Val Gly Glu Met
165 170 175

Trp Met Lys Asn Ala Asp Phe Ser Lys Tyr Gly Lys Arg Arg Asn Ser
180 185 190

Pro Asn Glu Tyr Val Phe Ser Val Ser Arg Ala Asp Leu Glu Lys Glu
195 200 205

Ile Val Thr Leu Phe Glu Ala Gln Arg Lys Phe His Ser Ser Tyr Ala
210 215 220

Ser Val Asp Leu Gln Lys Thr Tyr Ile Gln Ile Trp Ala His Gln Leu
225 230 235 240

Pro Phe Ala Ser Gly Asn Ala Ile Val Asn Lys Val Gly Tyr Cys Ser
245 250 255

Leu Leu Lys Gly Lys Glu Lys Arg Val Pro Lys Ala Thr Tyr Thr Phe
260 265 270

Gln Tyr Phe Asn Thr Leu Asp Gln Ile Asn Arg Thr Arg Leu Gly Pro
275 280 285

Asn Phe Gln Pro Phe Thr Lys Glu Gln Arg Asp Ile Ile Leu Asp Lys
290 295 300

Met Phe Gln Arg Thr Asp Tyr Tyr Lys Lys Lys Thr Ile Pro Glu Val
305 310 315 320

Thr Tyr Tyr Asp Ile Arg Lys Trp Leu Ala Leu Asp Glu Thr Ile Gln
325 330 335

Phe Lys Gly Leu Thr Tyr Asp Pro Asn Glu Glu Leu Lys Lys Ile Glu
340 345 350

Met Lys Pro Phe Ile Asn Leu Lys Pro Phe Tyr Glu Ile Lys Lys Val
355 360 365

Val Thr Asn Tyr Ala Lys Lys Thr Asn Glu Val Phe Ser Ala Leu Asp
370 375 380

Tyr Asp Thr Val Ala Tyr Ala Leu Thr Val Tyr Lys Thr Asp Lys Asp
385 390 395 400

Ile Arg Ser Tyr Leu Lys Arg Ser Asn Asn Leu Ser Lys Arg Cys Tyr
405 410 415

Asp Asp Gln Leu Ile Glu Glu Leu Leu Thr Leu Ser Tyr Thr Lys Phe

420

425

430

Gly His Leu Ser Phe Lys Ala Ile Asn His Val Leu Pro Ile Met Gln
 435 440 445

Glu Gly Arg Thr Tyr Gln Glu Ala Ile His Gln Leu Gly Tyr Asp Thr
 450 455 460

Thr Asn Leu Lys Lys Glu Asn Arg Ser Met Phe Leu Pro Ile Ile Pro
 465 470 475 480

Asp Glu Ile Thr Asn Pro Ile Val Lys Arg Ala Leu Thr Gln Ala Arg
 485 490 495

Lys Val Val Asn Ala Ile Ile Arg Arg Tyr Gly Ser Pro Asn Ser Val
 500 505 510

His Ile Glu Leu Ala Arg Glu Leu Ser Lys Ser His Asp Glu Arg Lys
 515 520 525

Lys Ile Met Thr Ala His Asp Glu Asn Tyr Lys Lys Asn Lys Gly Ala
 530 535 540

Ile Ser Ile Leu Ile Glu Asn Gly Ile Leu Asn Pro Thr Gly Tyr Asp
 545 550 555 560

Ile Val Arg Tyr Lys Leu Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr
 565 570 575

Ser Leu Lys Lys Ile Ser Ala Asn Thr Phe Phe Asn Glu Leu Lys Lys
 580 585 590

Glu Arg Ser Gly Pro Pro Val Leu Glu Val Asp His Ile Leu Pro Tyr
 595 600 605

Ser Gln Ser Phe Ile Asp Ser Tyr His Asn Lys Val Leu Val Tyr Gly
 610 615 620

Asp Glu Asn Gln Lys Lys Gly Asn Arg Ile Pro Tyr Thr Phe Phe Ser
 625 630 635 640

Glu Glu Asp Lys Glu Trp Glu Ser Phe Glu Ser Tyr Val Arg Ser Asn
645 650 655

Ser Phe Phe Ser Lys Lys Lys Arg Gly Tyr Leu Leu Lys Lys Ala Tyr
660 665 670

Leu Pro Arg Glu Ser Asn Leu Ile Lys Glu Arg His Leu Asn Asp Thr
675 680 685

Arg Tyr Ala Ser Ser Tyr Leu Lys Asn Phe Ile Glu Lys Asn Leu Lys
690 695 700

Phe Lys Glu Ala Val Gly Ile Thr Arg Lys Lys Tyr Val Gln Thr Val
705 710 715 720

Asn Gly Val Ile Thr Ala His Leu Arg Lys Arg Trp Gly Leu Glu Lys
725 730 735

Asp Arg Gln Glu Thr Tyr Leu His His Ala Met Asp Ala Ile Ile Val
740 745 750

Ala Cys Thr Asp His His Met Val Thr Lys Val Thr Glu Tyr Tyr Gln
755 760 765

Ile Lys Glu Gly Asn Lys Ser Ile Lys Lys Pro Tyr Phe Pro Leu Pro
770 775 780

Trp Met Gly Phe Arg Glu Glu Ile Leu Ser His Leu Glu Ser Gln Pro
785 790 795 800

Ile Ala Arg Lys Ile Ser Glu Glu Leu Lys Ile Gly Tyr Gln Ser Ser
805 810 815

Asp Tyr Ile Leu Val Ser Arg Met Pro Lys Arg Ser Val Thr Gly Ser
820 825 830

Ala His Asp Gln Thr Val Met Lys Lys Gly Gly Ile Asp Lys Lys Gly
835 840 845

Lys Thr Ile Ile Ile Lys Arg Val His Leu Lys Asp Ile Lys Phe Asp

850

855

860

Glu Asn Gly Asp Phe Lys Met Val Gly Lys Glu Gln Asp Leu Ala Thr
 865 870 875 880

Tyr Glu Ala Ile Lys Gln Arg Tyr Leu Glu Tyr Arg Lys Glu Ser Lys
 885 890 895

Lys Ala Phe Glu Thr Pro Leu Tyr Lys Pro Ser Lys Lys Gly Lys Gly
 900 905 910

Asn Leu Ile Lys Lys Ile Lys Val Glu Val Gln Thr Lys Ser Phe Val
 915 920 925

Arg Glu Ile Asn Gly Gly Val Ala Gln Asn Gly Asp Leu Val Arg Val
 930 935 940

Asp Leu Phe Glu Lys Asp Asn Arg Tyr Tyr Met Val Pro Ile Tyr Val
 945 950 955 960

Val Asp Thr Val Arg Ser Glu Leu Pro Asn Lys Ala Val Thr Ser Ser
 965 970 975

Lys Gly Tyr Glu Gln Trp Leu Ser Ile Asp Asn Ser Phe Thr Phe Lys
 980 985 990

Phe Ser Leu Tyr Pro Tyr Asp Leu Val Arg Leu Val Lys Gly Asp Glu
 995 1000 1005

Asp Arg Phe Leu Tyr Phe Ser Thr Leu Asp Ile Asn Ser Asp Arg
 1010 1015 1020

Leu Asn Phe Lys Asp Val Asn Lys Pro Ser Lys Gln Ala Glu Tyr
 1025 1030 1035

Arg Tyr Ser Leu Lys Thr Ile Glu Asn Leu Glu Lys Tyr Glu Ile
 1040 1045 1050

Gly Val Leu Gly Asp Leu Arg Leu Val Arg Gln Glu Thr Arg Lys
 1055 1060 1065

Ile Phe Lys
1070

<210> 120
<211> 1091
<212> PRT
<213> Rhizobium sp.

<223> APG09053.1_RGN

<400> 120

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

Gly Phe Cys Val Phe Ala Leu Asp Glu Asn Gly Glu Pro Tyr Arg Ile
20 25 30

Ile Asp Ile Gly Thr Arg Ile Tyr Ala Asp Gly Arg Asp Pro Gln Ser
35 40 45

Lys Ala Ser Leu Ala Val Ala Arg Arg Glu Ala Arg Ala Met Ser Arg
50 55 60

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

Thr Glu Tyr Gly Leu Met Pro Ala Glu Ala Ser Ala Arg Gln Ala Leu
85 90 95

Ile Ala Glu Thr Ser Asp Arg Asn Gly Ala Ala Gly Glu Ala Ser Ala
100 105 110

Asn Pro Tyr Ala Leu Arg Ala Met Ala Leu Asp Glu Lys Leu Pro Leu
115 120 125

Tyr Trp Ile Gly Arg Val Leu Phe His Leu Asn Gln Arg Arg Gly Phe
130 135 140

Lys Ser Asn Arg Lys Thr Asp Arg Lys Asp Asn Glu Lys Gly Val Ile
145 150 155 160

Ala Leu Gly Ile Gly Glu Leu Arg Ala Ala Met His Lys Thr Lys Ala
165 170 175

Arg Thr Tyr Gly Glu Trp Leu Ala Ala Arg Arg Glu Asp Gly His Val
180 185 190

Thr Arg Leu Arg Ala Gly Ser Asp Ala Phe Glu Gly Asp Gly Tyr Ala
195 200 205

Phe Tyr Pro Glu Arg Ser Leu Leu Glu Ala Glu Phe Arg Glu Ile Trp
210 215 220

Ala Arg Gln Ser Ser Phe Tyr Pro Asp Val Leu Thr Glu Ala Arg Gly
225 230 235 240

Asp His Leu Phe Gln Val Met Phe Tyr Gln Arg Pro Leu Lys Lys Pro
245 250 255

Arg Val Gly Lys Cys Ala Phe Asn Pro Leu Glu Glu Arg Leu Ala Lys
260 265 270

Ala His Pro Leu Phe Gln Glu Phe Arg Leu Tyr Lys Glu Val Asn Glu
275 280 285

Leu Glu Leu Val Leu Pro Asp His Ser His Lys Lys Leu Asp Ile Glu
290 295 300

Gln Arg Asn Ala Leu Val Ile Leu Leu Arg Ala Asn Arg Glu Val Thr
305 310 315 320

Phe Lys Lys Leu Arg Thr Ala Leu Lys Leu Ala Pro Asp Ile Val Phe
325 330 335

Asn Lys Glu Ser Glu Ser Arg Glu Lys Leu Arg Gly Asp Glu Val His
340 345 350

Ser Val Leu Ala Asp Lys Lys Arg Phe Gly Pro Gln Trp Gly Ala Leu
355 360 365

Ser Arg Glu Arg Gln Trp Gln Ile Ile Gln Thr Leu Lys Asp Glu Glu
370 375 380

Asn Pro Glu Lys Leu Phe Gly Trp Leu Lys Glu Glu Phe Gly Phe Glu
385 390 395 400

Gly Glu Lys Ala Asp Ala Ile Ala Asn Ala Pro Leu Pro Glu Gly Tyr
405 410 415

Gly Arg Leu Gly Glu Thr Ala Leu Ser Ser Met Leu Asp Glu Met Lys
420 425 430

Ala Ala Val Ile Pro Glu Ala Glu Ala Ala Lys Asn Cys Gly Tyr Asp
435 440 445

His Ser Lys Leu Gly Glu Asp Arg Glu Glu Gly Glu Ala Phe Leu Pro
450 455 460

Pro Tyr Gln Glu Ile Leu Ser Arg Gln Ile Pro Pro Gly Thr Asn Asp
465 470 475 480

Pro Ala Asp Ile Tyr Asp Ile Arg Met Gly Arg Phe Thr Asn Pro Thr
485 490 495

Val His Ile Gly Leu Asn Gln Leu Arg Arg Val Val Asn Ala Leu Ile
500 505 510

Ala Arg His Gly Lys Pro Gln Phe Val Ser Leu Glu Leu Ala Arg Asp
515 520 525

Leu Gln Leu Ser Glu Lys Gln Lys Ala Glu Ala Asn Arg Ala Ile Ala
530 535 540

Lys Asn Thr Arg Glu Ala Ala Ala Arg Ser Lys Lys Leu Gly Glu Met
545 550 555 560

Gly Gln Leu Asp Thr Gly Tyr Asn Arg Leu Leu Leu Lys Leu Trp Glu
565 570 575

Glu Leu Asn Pro Ser Lys Pro Glu Asp Arg Val Cys Ile Tyr Ser Gly
580 585 590

Lys Pro Ile Gly Ile Asp Met Leu Phe Ser Ala Glu Val Asp Val Asp
595 600 605

His Ile Leu Pro Trp Ser Lys Thr Leu Asp Asp Ser Gln Ala Asn Lys
610 615 620

Leu Leu Cys Leu Lys Ser Ala Asn Arg Gln Lys Arg Asn Arg Ala Pro
625 630 635 640

Ala Asp Val Pro Glu Trp Arg Asp Arg Tyr Glu Glu Ile Leu Ala Arg
645 650 655

Ala Ala Arg Leu Ala Pro Asn Lys Arg Arg Arg Phe Ala Pro Asn Ala
660 665 670

Met Ala Glu Phe Glu Ala Glu Gly Gly Phe Leu Ala Arg Gln Leu Thr
675 680 685

Asp Thr Gln Tyr Leu Ser Arg Met Ala Arg Glu Tyr Leu Ser Cys Leu
690 695 700

Phe Pro Ser Ala Glu Ile Asp Arg Arg Gly Glu Arg Lys Lys Lys Ile
705 710 715 720

Arg Val Ile Val Ser Pro Gly Arg Leu Thr Glu Met Leu Arg Arg Asn
725 730 735

Trp Gly Leu Asn Asn Ile Leu Pro Asp His Asn Leu Gly Glu Met Thr
740 745 750

Gln Ala Lys Asn Arg Lys Asp His Arg His His Ala Ile Asp Ala Ala
755 760 765

Val Val Gly Val Thr Thr Arg Ser Leu Leu Gln Arg Ile Ala Thr Ala
770 775 780

Ala Gly Arg Leu Asp Glu Ala Asp Phe Glu Asn Leu Val Arg Lys Met
785 790 795 800

Val Ser Glu Asn Pro Pro Trp Pro Thr Phe Arg Glu Glu Leu Gln Ala
805 810 815

Ala Val Asn Gly Ile Val Val Ser His Lys Pro Asp His Gly Thr Val
820 825 830

Ser Arg Lys Gly Tyr Ala Glu Gly Lys Gly Arg Thr Ala Gly Arg Leu
835 840 845

His Asn Asp Thr Ala Tyr Gly Phe Gly Thr Asp Ala Ser Gly Asn Pro
850 855 860

Val Ala Val Arg Arg Lys Leu Phe Thr Ser Leu Glu Ala Lys Asp Ile
865 870 875 880

Pro Met Ile Arg Asp Pro Glu Leu Gln Ala Glu Leu Tyr Ala Ala Ile
885 890 895

Asp Gly Leu Asp Asp Arg Arg Ala Leu Gln Asp Ala Leu Leu Ser Phe
900 905 910

Arg Arg Thr His Pro Lys Phe Lys Gly Ile Arg Arg Val Arg Met Ala
915 920 925

Glu Thr Leu Ser Phe Ile Pro Ile Arg Asp Ser Gln Gly Asn Val Tyr
930 935 940

Lys Gly Tyr Lys Gly Asp Ala Asn Tyr Arg Tyr Asp Val Trp Glu Thr
945 950 955 960

Leu Asp Gly Lys Trp His Ala Asp Val Val Thr Met Phe Asn Ala His
965 970 975

Gln Pro Asp Trp Arg Ser Pro Val His Gln Glu His Pro Thr Ala Arg
980 985 990

Arg Val Leu Arg Leu Gln Gln Asn Asp Met Val Ala Tyr Glu His Pro
995 1000 1005

Asn Asp Gly Tyr Thr Ile Ala Arg Val Val Lys Phe Asn Thr Ala
1010 1015 1020

Gly Ile Val Tyr Phe Ala Ser His Arg Glu Ser Gly Ser Leu Lys
1025 1030 1035

Ala Arg Asp Ala Asp Lys Gln Asp Pro Phe Lys Tyr Phe Ser Lys
1040 1045 1050

Ser Ala Ala Gly Leu Lys Asp Ile Gln Cys Arg Gln Ile Arg Ile
1055 1060 1065

Asp Ala Ala Gly Arg Val Phe Asp Pro Gly Pro Gln Asp Arg Ala
1070 1075 1080

Ser Lys Ser Thr Arg Lys Thr Asn
1085 1090

<210> 121
<211> 1401
<212> PRT
<213> Myroides sp.

<223> APG04293.1_RGN

<400> 121

Met Lys Arg Ile Leu Gly Ile Asp Leu Gly Thr Thr Ser Ile Gly Trp
1 5 10 15

Ala Tyr Val His Glu Ala Glu Asn Glu Asn Glu Ile Ser Ser Ile Glu
20 25 30

Lys Leu Gly Val Arg Val Asn Pro Leu Thr Val Asp Glu Gln Ile Asn
35 40 45

Phe Glu Lys Gly Lys Pro Ile Thr Thr Asn Ala Thr Arg Thr Leu Phe
50 55 60

Arg Ser Ala Arg Arg Asn Leu Gln Arg Phe Lys Leu Arg Arg Lys His
65 70 75 80

Leu Ile Glu Ile Leu Lys His Asn Asn Trp Ile Ser Lys Gln Thr Leu
85 90 95

Leu Thr Glu Asn Gly Asn His Thr Thr Phe Glu Thr Ile Ser Leu Arg

100					105					110					
Asp	Lys	Ala	Ala	Thr	Gln	Lys	Val	Thr	Leu	Glu	Glu	Leu	Ala	Arg	Val
		115					120					125			
Leu	Leu	Met	Ile	Asn	Lys	Lys	Arg	Gly	Tyr	Lys	Ser	Ser	Arg	Lys	Val
	130					135					140				
Asn	Asn	Thr	Glu	Glu	Gly	Gln	Leu	Ile	Asp	Gly	Met	Glu	Val	Ala	Lys
145					150					155					160
Lys	Leu	Tyr	His	Glu	Gln	Leu	Thr	Pro	Gly	Gln	Leu	Ser	Leu	Gln	Leu
				165					170					175	
Ile	Lys	Gln	Gly	Val	Lys	Arg	Leu	Pro	Asp	Tyr	Tyr	Arg	Ser	Asp	Leu
			180					185					190		
Asn	His	Glu	Leu	Asp	Leu	Ile	Trp	Ala	Phe	Gln	Gln	Lys	Phe	Tyr	Pro
		195					200					205			
His	Ile	Leu	Thr	Ser	Glu	Phe	Lys	Glu	Leu	Leu	Lys	Gly	Lys	Gly	Leu
	210					215					220				
Arg	Ala	Thr	Ser	Asn	Gln	Phe	Trp	Thr	Asn	Tyr	Asn	Phe	Asn	Thr	Ala
225					230					235					240
Asp	Asn	Lys	Gly	Thr	Arg	Glu	Glu	Lys	Lys	Leu	Gln	Asn	Tyr	Gln	Trp
				245					250					255	
Arg	Ser	Asp	Ala	Leu	Ile	Lys	Gln	Leu	Asp	Lys	Glu	Ile	Val	Ala	Tyr
			260					265					270		
Val	Ile	Ser	Glu	Ile	Asn	Gly	Gln	Ile	Asn	Asn	Ser	Ser	Gly	Tyr	Leu
		275					280					285			
Gly	Ala	Ile	Ser	Asp	Arg	Ser	Lys	Glu	Leu	Tyr	Phe	Asn	Gln	Leu	Thr
	290					295					300				
Val	Gly	Gln	Tyr	Leu	His	Arg	Gln	Leu	Ile	Ala	Asn	Pro	His	Thr	Arg
305					310					315					320

Leu Lys Ser Gln Val Phe Tyr Arg Gln Asp Tyr Leu Asp Glu Phe Glu
325 330 335

Lys Ile Trp Glu Thr Gln Ser Lys Phe Tyr Pro Lys Glu Leu Thr Thr
340 345 350

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

Leu Lys Ser Gln Lys Gly Leu Ile Ser Phe Cys Glu Phe Glu Gln Glu
370 375 380

Gln Lys Ile Ile Asn Gly Lys Thr Lys Thr Ile Gly His Arg Val Ile
385 390 395 400

Pro Lys Ser Ser Pro Leu Phe Gln Glu Phe Lys Ile Trp Gln Gln Leu
405 410 415

His Asn Val Val Leu Arg Asn Lys Lys Thr Asn Val Ile Thr Pro Leu
420 425 430

Ala Glu Glu Gln Lys Ser His Leu Phe Glu Glu Leu Asn Leu Lys Gly
435 440 445

Lys Leu Ser Ser Thr Gln Ile Leu Lys Leu Ile Glu Asp Lys Pro Lys
450 455 460

Asp Trp Glu Leu Asn Tyr Ser Glu Leu Glu Gly Asn Asn Thr Asn Lys
465 470 475 480

Ala Leu Tyr Asn Ala Tyr Leu Asp Ile Leu Asp Leu Glu Gly Tyr Asp
485 490 495

Ile Arg Asp Glu Leu Lys Ile Lys Leu Asn Lys Asp Asp Ile Glu Leu
500 505 510

Ser Asp Leu Asp Val Asn Ala Ser Glu Ile Lys Glu Met Ile Phe Ser
515 520 525

Ile Phe Lys His Leu Gly Ile Asn Thr Glu Ile Leu Glu Phe Asp Ala

530

535

540

Thr Leu Lys Asp Lys Ala Phe Glu Lys Gln Ala Ser Tyr Gln Leu Trp
 545 550 555 560

His Leu Leu Tyr Ser Tyr Glu Glu Asp Asn Ser Pro Thr Gly Leu Asp
 565 570 575

Arg Leu His Asn Leu Leu Gln Lys Lys Phe Asn Phe Thr Leu Asp Gln
 580 585 590

Ala Lys Leu Val Gly Asn Val Leu Phe Gln Asp Asp Tyr Gly Asn Leu
 595 600 605

Ser Ser Lys Ala Ile Lys Asn Ile Phe Pro Tyr Ile Thr Asp Asn Asn
 610 615 620

Tyr Ser Thr Ala Cys Glu Leu Ala Gly Tyr Lys His Ser Lys His Ser
 625 630 635 640

Leu Thr Lys Glu Glu Asn Glu Lys Arg Glu Leu Lys Ser Arg Leu Thr
 645 650 655

Ile Leu Thr Lys Asn Ser Leu Arg Asn Pro Val Val Glu Lys Ile Leu
 660 665 670

Asn Gln Met Ile Asn Val Ile Asn Thr Leu Ile Glu Ser Glu Asn Asp
 675 680 685

Lys Leu Val Ala Gln Gly Lys Glu Pro Asn Phe Gln Phe Asp Glu Ile
 690 695 700

Arg Ile Glu Leu Ala Arg Glu Leu Lys Lys Asn Ala Lys Glu Arg Glu
 705 710 715 720

Glu Leu Thr Lys Ser Met Thr Gln Gly Lys Thr Asn His Glu Lys Ile
 725 730 735

Ile Lys Ile Leu Gln Lys Glu Asp Gly Ile Lys Asn Pro Thr Arg Asn
 740 745 750

Asp Ile Thr Arg Phe Lys Leu Tyr Ser Glu Leu Lys Asn Asn Gly Tyr
755 760 765

Lys Asp Leu Tyr Thr Asn Glu Tyr Ile Glu Arg Lys Asp Ile Phe Ser
770 775 780

Lys Glu Tyr Asp Ile Glu His Ile Ile Pro Gln Ser Lys Leu Phe Asp
785 790 795 800

Asp Ser Phe Ser Asn Lys Thr Leu Val Arg Arg Asn Val Asn Leu Lys
805 810 815

Lys Gly Asn Gln Thr Ala Tyr Asp Phe Ile Leu Ser Glu Tyr Gly Gln
820 825 830

Glu Lys Ala Asn Glu Phe Glu Thr Arg Ile Thr Asn Leu Tyr Ser Leu
835 840 845

Gly Lys Asp Glu Gly Ile Ser Lys Ser Lys Tyr Lys Lys Leu Leu Met
850 855 860

Gln Glu Ser Glu Ile Gly Lys Gly Phe Ile Glu Arg Asp Leu Arg Glu
865 870 875 880

Thr Gln Tyr Ile Ala Lys Lys Ala Lys Ser Leu Leu Leu Glu Ile Thr
885 890 895

Arg Asn Val Val Ser Thr Ser Gly Gly Ile Thr Asp Arg Leu Arg Glu
900 905 910

Asp Trp Gly Leu Val Asn Val Met Lys Glu Leu Asn Leu Lys Lys Phe
915 920 925

Arg Asp Ala Gly Leu Thr Glu Phe Val Glu Met Lys Asp Gly Asn Lys
930 935 940

Lys Glu Val Ile Ile Asp Trp Thr Lys Arg Asn Asp His Arg His His
945 950 955 960

Ala Met Asp Ala Leu Thr Ile Ala Phe Thr Lys His Asn His Ile Gln

965

970

975

Tyr Leu Asn His Leu Asn Ala Arg Lys Asn Thr Lys Asp Glu Leu His
 980 985 990

Thr Asn Ile Ile Ala Ile Glu Lys Lys Glu Thr Thr Ile Ile Thr Asp
 995 1000 1005

Glu Arg Gly Asn Lys Lys Arg Ile Phe Lys Glu Pro Ile Pro Asn
 1010 1015 1020

Phe Arg Asn Ile Ala Lys Ser His Leu Glu Ala Ile Leu Val Ser
 1025 1030 1035

His Lys Ala Lys Asn Lys Val Val Thr Lys Asn Ile Asn Lys Ile
 1040 1045 1050

Ser Gly Lys Lys Ile Gly Gln Gln Thr Leu Thr Pro Arg Gly Gln
 1055 1060 1065

Leu His Lys Glu Thr Val Tyr Gly Lys Ile Arg Gln Tyr Val Ser
 1070 1075 1080

Lys Glu Glu Lys Val Gly Pro Lys Phe Thr Lys Glu Val Ile Glu
 1085 1090 1095

Leu Val Ser Asn Pro Thr Tyr Arg Asp Leu Leu Leu Lys Arg Leu
 1100 1105 1110

Gln Glu Asn Asn Asn Asp Pro Lys Lys Ala Phe Gly Gly Lys Asn
 1115 1120 1125

Ala Leu Ser Lys Ser Pro Ile Tyr Ile Asn Leu Glu Lys Asn Ile
 1130 1135 1140

Ile Val Pro Glu Val Val Lys Leu Val Trp Leu Glu Asp Asp Tyr
 1145 1150 1155

Thr Ile Arg Lys Glu Ile Ser Pro Glu Leu Lys Ile Asp Lys Val
 1160 1165 1170

Val Asp Glu Gly Val Lys Arg Ile Leu Gln Asn Arg Leu Ala Leu
1175 1180 1185

Tyr Asn Gly Asp Ala Lys Ile Ala Phe Ser Asp Leu Asp Asn Asn
1190 1195 1200

Pro Ile Trp Leu Ser Lys Asp Lys Gly Ile Thr Ile Lys Arg Val
1205 1210 1215

Thr Leu Ser Gly Val Lys Asn Ala Thr Ala Leu His Thr Lys Lys
1220 1225 1230

Asp His Leu Gly Gln Pro Ile Leu Asp His Lys Gly Gln Glu Ile
1235 1240 1245

Pro Val Asp Phe Ile Ser Thr Gly Asn Asn His His Val Ala Ile
1250 1255 1260

Tyr Glu Asp Lys Asn Gly Lys Leu Gln Glu Ser Val Val Ser His
1265 1270 1275

Phe Glu Ala Val Glu Arg Val Asn Gln Gln Leu Pro Ile Ile Asp
1280 1285 1290

Lys Thr Phe Asn Gln His Leu Gly Trp Lys Phe Leu Phe Thr Met
1295 1300 1305

Lys Gln Asn Glu Leu Phe Leu Phe Pro Ser Asp Asp Phe Asn Pro
1310 1315 1320

Asn Glu Ile Asp Leu Phe Asp Thr Arg Asn Tyr Ala Val Ile Ser
1325 1330 1335

Lys His Leu Phe Arg Ala Gln Lys Leu Ser Thr Lys Asp Tyr Ser
1340 1345 1350

Phe Arg His His Leu Glu Thr Asn Val Glu Asp Ile Lys Asp Leu
1355 1360 1365

Lys Thr Ile Thr Trp Arg Arg Glu Gly Leu Asn Gly Ile Gln Asn

1370

1375

1380

Ile Val Lys Ile Arg Thr Asn His Leu Gly Glu Ile Val His Ile
 1385 1390 1395

Gly Glu Tyr
 1400

<210> 122
 <211> 1432
 <212> PRT
 <213> Chryseobacterium sp.

<223> APG01308.1_RGN

<400> 122

Met Ala Lys Asn Ile Leu Gly Leu Asp Leu Gly Thr Asn Ser Ile Gly
 1 5 10 15

Trp Ala Leu Ile Asn Gln Asp Phe Glu Asn Lys Gln Gly Lys Ile Leu
 20 25 30

Gly Met Gly Ser Arg Ile Ile Pro Met Ser Gln Asp Ile Leu Gly Asp
 35 40 45

Phe Gly Lys Gly Asn Ser Val Ser Gln Thr Ala Ala Arg Thr Gly Tyr
 50 55 60

Arg Gly Val Arg Arg Leu Arg Glu Arg Phe Leu Leu Arg Arg Glu Arg
 65 70 75 80

Leu His Arg Val Leu Asn Ile Ile Asn Phe Leu Pro Glu His Tyr Ala
 85 90 95

Ser Gln Ile Asp Phe Glu Lys Arg Phe Gly Lys Phe Lys Glu Glu Thr
 100 105 110

Glu Pro Lys Leu Ala Tyr Asn Lys Asp Gly Phe Val Phe Lys Asp Ser
 115 120 125

Phe Glu Glu Met Leu Ala Asp Phe Lys Asn Tyr Gln Pro Gln Leu Leu
 130 135 140

Glu Asn Asp Lys Lys Ile Pro Tyr Asp Trp Thr Ile Tyr Tyr Leu Arg
145 150 155 160

Lys Lys Ala Leu Ser Gln Lys Ile Glu Lys Glu Glu Leu Ala Trp Ile
165 170 175

Leu Leu Asn Phe Asn Gln Lys Arg Gly Tyr Tyr Gln Leu Arg Gly Glu
180 185 190

Asp Phe Glu Glu Glu Lys Asp Lys Met Phe Val Arg Leu Lys Val Glu
195 200 205

Lys Ile Ile Asp Ser Gly Asp Asn Val Lys Gly Lys Ile Leu Tyr Asp
210 215 220

Val Tyr Phe Glu Asn Gly Trp Lys Tyr Asp Lys Gln Ile Val Lys Thr
225 230 235 240

Glu Asp Trp Ile Glu Arg Val Lys Glu Phe Ile Val Thr Glu Ser Phe
245 250 255

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

Lys Asp Trp Ile Ala Ile Lys Thr Lys Thr Glu Gln Glu Ile Asp Lys
275 280 285

Ser His Lys Thr Val Gly Val Tyr Ile Tyr Glu Thr Leu Leu Gln Asn
290 295 300

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

Phe Tyr Lys Asp Glu Leu Lys Gln Ile Leu Glu Lys Gln Lys Glu Phe
325 330 335

His Gln Glu Leu Lys Asn Asp Asp Leu Tyr Asn Asp Cys Val Arg Glu
340 345 350

Leu Tyr Arg Asn Asn Glu Ala His Gln Leu Thr Leu Ser Lys Lys Glu
355 360 365

Phe Val His Leu Leu Met Glu Asp Ile Ile Phe Tyr Gln Arg Pro Leu
370 375 380

Arg Ser Gln Lys Ser Ser Ile Ser Asn Cys Ser Leu Glu Phe Arg Lys
385 390 395 400

Tyr Lys Asp Glu Asn Gly Val Glu His Ile Gln Tyr Leu Lys Ala Val
405 410 415

Pro Lys Ser Asn Pro Tyr Tyr Gln Glu Phe Arg Ile Trp Gln Trp Ile
420 425 430

Phe Asn Leu Asn Ile Tyr Lys Arg Asp Asp Glu Gln Asn Pro Val Thr
435 440 445

Thr Glu Phe Leu Asn Thr Thr Thr Asp Ile Glu Asn Leu Phe Glu Phe
450 455 460

Leu Asn Asn Arg Lys Glu Val Asp Gln Lys Ala Leu Leu Lys His Phe
465 470 475 480

Lys Leu Asn Glu Lys Thr His Arg Trp Lys Tyr Val Glu Asp Lys Lys
485 490 495

Tyr Pro Cys Asn Glu Thr Lys Ser Met Ile Ser Glu Arg Leu Lys Lys
500 505 510

Val Glu Asn Ile Ala Asn Asp Phe Leu Thr Arg Glu Ile Glu Gln Lys
515 520 525

Ile Trp His Ile Ile Tyr Ser Val Asn Asp Lys Ile Glu Tyr Glu Lys
530 535 540

Ala Leu Lys Ser Phe Ala Lys Lys Asn Asn Leu Asp Glu Asn Ser Phe
545 550 555 560

Phe Glu Ala Phe Lys Lys Phe Pro Pro Phe Lys Ser Glu Tyr Gly Ser
565 570 575

Phe Ser Glu Lys Ala Ile Lys Lys Leu Leu Pro Leu Met Arg Leu Gly
580 585 590

Lys Tyr Trp Asn Cys Glu Asn Ile Ser Asp Asp Ser Lys Glu Arg Ile
595 600 605

Gln Lys Ile Ile Asn Gly Glu Tyr Asp Glu Asn Ile Lys Asp Arg Val
610 615 620

Arg Glu Lys Ala Ile His Leu Thr Ser Glu Asn Asn Phe Gln Gly Leu
625 630 635 640

Gln Leu Trp Leu Ala Gln Tyr Val Val Tyr Asp Arg His Ser Glu Ala
645 650 655

Ser Met Ile Gly Lys Trp Asn Ser Ala Asp Asp Leu Glu Glu Phe Leu
660 665 670

Lys Glu Phe Lys Gln His Ser Leu Arg Asn Pro Ile Val Glu Gln Val
675 680 685

Ile Thr Glu Thr Leu Arg Val Val Lys Asp Ile Trp Leu Lys Tyr Gly
690 695 700

Asn Gly Ala Lys Asp Phe Phe Asn Glu Ile His Ile Glu Leu Gly Arg
705 710 715 720

Glu Met Lys Gln Thr Lys Asp Glu Arg Leu Asp Ala Thr Lys Arg Ile
725 730 735

Thr Glu Asn Glu Asn Thr Asn Leu Arg Ile Lys Ala Leu Leu Ala Glu
740 745 750

Met Met Asn Asp Asn Ser Val Glu Asn Val Arg Pro Tyr Ser Pro Met
755 760 765

Gln Gln Glu Ile Leu Lys Ile Tyr Glu Asp Gly Val Leu Asn Ser Gly
770 775 780

Ile Glu Ile Glu Asp Glu Tyr Leu Lys Ile Ser Lys Thr Ala Gln Pro
785 790 795 800

Ser Pro Ser Asp Leu Lys Arg Tyr Lys Leu Trp Leu Glu Gln Arg Tyr
805 810 815

Lys Ser Pro Tyr Thr Gly Gln Ile Ile Pro Leu Asn Lys Leu Phe Thr
820 825 830

Pro Glu Tyr Glu Ile Glu His Ile Ile Pro Gln Ser Arg Tyr Phe Asp
835 840 845

Asp Ser Phe Ser Asn Lys Ile Ile Cys Glu Ser Ala Val Asn Lys Leu
850 855 860

Lys Asp Asn Tyr Ile Gly Leu Gly Phe Ile Lys Gln Phe Gly Gly Thr
865 870 875 880

Ile Val Glu Cys Gly Leu Gly Lys Asn Val Lys Val Phe Glu Val Asn
885 890 895

Glu Tyr Glu Asp Phe Val Lys Lys His Tyr Ala Asn Asn Arg Gly Lys
900 905 910

Arg Asn Lys Leu Leu Leu Glu Glu Ile Pro Glu Lys Met Ile Glu Arg
915 920 925

Gln Leu Asn Asp Thr Arg His Ile Ser Lys Tyr Ile Ser Gly Val Leu
930 935 940

Ser Asn Ile Val Arg Val Glu Asp Gly Ser Asp Glu Gly Val Asn Ser
945 950 955 960

Lys Asn Ile Val Pro Gly Asn Gly Lys Ile Thr Thr Gln Leu Lys Gln
965 970 975

Asp Trp Gly Leu Asn Asp Val Trp Asn Asp Leu Ile Leu Pro Arg Phe
980 985 990

Glu Arg Met Asn Gln Leu Thr Asn Ser Thr Asp Phe Thr Ala Trp Asn
995 1000 1005

Glu Asn His Gln Lys Phe Leu Pro Thr Val Pro Ile Glu Phe Ser
1010 1015 1020

Lys Gly Phe Ser Lys Lys Arg Ile Asp His Arg His His Ala Leu
1025 1030 1035

Asp Ala Leu Val Ile Ala Cys Ala Thr Lys Asp His Ile Asn Leu
1040 1045 1050

Leu Asn Asn Gln Ser Ala Lys Ser Asp Thr Lys Arg Tyr Asp Leu
1055 1060 1065

Lys Lys Lys Leu Met Lys Phe Glu Lys Gly Val Tyr Asn His Pro
1070 1075 1080

Gln Thr Gly Glu Arg Ile Gln Arg Asp Val Pro Lys Gln Phe Leu
1085 1090 1095

Lys Pro Trp Glu Ser Phe Thr Ile Asp Ala Lys Asn Asn Leu Asp
1100 1105 1110

Lys Ile Ile Ile Ser Phe Lys Gln Asn Leu Arg Val Ile Asn Lys
1115 1120 1125

Ala Thr Asn Tyr Tyr Glu Lys Tyr Val Glu Lys Asn Gly Ile Lys
1130 1135 1140

Thr Lys Glu Arg Val Glu Gln Thr Gly Thr Asn Trp Ala Ile Arg
1145 1150 1155

Lys Pro Met His Lys Glu Thr Val Ser Gly Ile Val Asn Leu Pro
1160 1165 1170

Trp Val Lys Val Pro Lys Gly Lys Ile Leu Thr Ala Thr Arg Lys
1175 1180 1185

Ser Leu Asp Thr Thr Phe Asp Leu Lys Ser Ile Asn Ser Ile Thr
1190 1195 1200

Asp Thr Gly Ile Gln Lys Ile Leu Arg Asn Tyr Leu Glu Phe Lys
1205 1210 1215

Gly Ser Pro Glu Leu Ala Phe Ser Pro Glu Gly Ile Glu Asp Met
1220 1225 1230

Asn Lys Asn Ile Glu Lys Tyr Asn Asp Gly Lys Leu His Gln Pro
1235 1240 1245

Ile Asn Lys Val Arg Val Phe Glu Leu Gly Ser Lys Phe Gln Val
1250 1255 1260

Gly Gln Thr Gly Asn Lys Lys Asp Lys Tyr Val Glu Ala Ala Lys
1265 1270 1275

Gly Thr Asn Leu Phe Phe Ala Val Tyr Glu Asp Lys Asn Gly Lys
1280 1285 1290

Arg Asn Tyr Glu Thr Ile Pro Leu Asn Ile Val Ile Glu Arg Gln
1295 1300 1305

Lys Gln Gly Leu Leu Ala Cys Pro Glu Ser Asn Glu Lys Gly Glu
1310 1315 1320

Lys Leu Leu Phe Gln Leu Ser Pro Asn Asp Phe Val Tyr Leu Thr
1325 1330 1335

Thr Glu Glu Glu Asn Asp Asn Ala Thr Pro Ile Asn Phe Ser Leu
1340 1345 1350

Leu Ser Lys Glu Gln Ile Asn Asn Leu Tyr Lys Ile Val Ser Phe
1355 1360 1365

Thr Gly Asn Arg Leu Tyr Gly Ile Pro Ile Cys Val Ala Thr Thr
1370 1375 1380

Ile Val Asn Lys Ala Glu Tyr Thr Gln Leu Asn Lys Ile Glu Phe
1385 1390 1395

Thr Lys Glu Lys Asp Leu Leu Leu Lys Leu Asn Val Asn Arg Leu
1400 1405 1410

Gly Asp Val Lys Thr Phe Thr Ala Asn Asp Ile Arg Lys Ile Phe
1415 1420 1425

Asn Arg Gln Glu
1430

<210> 123
<211> 1092
<212> PRT
<213> Bacillus sp.

<223> APG06646.1_RGN

<400> 123

Met Arg Tyr Ser Ile Gly Leu Asp Ile Gly Thr Thr Ser Ile Gly Asn
1 5 10 15

Ala Val Ile Asn Lys Asp Leu Gln Arg Phe Glu His Leu Gly Val Arg
20 25 30

Ile Phe Asp Ala Ala Glu Asn Pro Lys Asp Gly Ser Ser Leu Ser Ala
35 40 45

Pro Arg Arg Leu Ala Arg Ser Ser Arg Arg Arg Leu Arg Arg Arg Lys
50 55 60

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

Thr Lys Thr Gln Val Lys Asn Leu Tyr Asn Ser Lys Asn Ile Asn Leu
85 90 95

Asp Ile Trp Asp Ile Arg Val Ser Gly Ile Asp Arg Lys Leu Phe Asn
100 105 110

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

Lys Ser Asn Arg Lys Ser Glu Leu Lys Glu Asp Asp Asn Gly Ala Ile
130 135 140

Leu Ser Ala Val Lys Glu Asn Arg Glu Leu Met Asp Glu Lys Gly Tyr
145 150 155 160

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

Lys Arg Asn Lys Gly Gly Asp Tyr Ser His Val Val Ala Arg Ser Asp
180 185 190

Ile Glu Asn Glu Ile Cys Leu Leu Phe Gln Lys Gln Arg Glu Tyr Gly
195 200 205

His Pro Phe Ala Thr Glu Glu Asn Glu Glu Ala Phe Leu Ser Ile Trp
210 215 220

Ser Ser Gln Arg Pro Phe Ser Thr Lys Asp Asp Ile Val Lys Lys Ile
225 230 235 240

Gly Asn Cys Thr Leu Glu Pro Lys Glu Lys Arg Ala Pro Lys Ser Thr
245 250 255

Tyr Thr Phe Glu Arg Phe Arg Ala Leu Asp Lys Leu Asn Arg Leu Arg
260 265 270

Ile Leu Ser Thr Thr Ala Pro Ser Arg Pro Leu Thr Asn Glu Glu Arg
275 280 285

Lys Ser Ile Leu Ser Ser Leu Phe Ser Lys Lys Glu Val Lys Tyr Lys
290 295 300

Glu Leu Arg Lys Leu Leu Lys Leu Thr Asp Asp Gln Arg Phe Asn Glu
305 310 315 320

Ile Tyr Tyr Ser Pro Asp Glu Thr Ile Glu Lys Thr Glu Asn Arg Thr
325 330 335

Phe Leu Ser Leu Glu Ser Gln Tyr Lys Ile Lys Lys Ile Ile Glu Lys
340 345 350

Thr Glu Ser Lys Asn Met Gln Ser Ser Tyr His Pro Ile Asp Tyr Asp

355

360

365

Thr Ile Gly Tyr Ala Leu Thr Val Phe Lys Asp Asp Lys Asp Ile Gln
 370 375 380

His Tyr Leu Gln Asn Ser Tyr Ile Asp Ser Lys Gly Lys Ala Ile Pro
 385 390 395 400

Asn Met Ala Asn Arg Glu Tyr Asn Leu Glu Leu Ile Glu Glu Leu Leu
 405 410 415

Gly Leu Ser Phe Ala Lys Phe Gly His Leu Ser Leu Lys Ala Leu Asn
 420 425 430

Asn Ile Leu Pro Tyr Met Glu Glu Gly Glu Pro Tyr His Ile Ala Cys
 435 440 445

Glu Met Ala Ser Tyr Gln Phe Ser Gln Arg Leu Ser Lys Glu Lys His
 450 455 460

Arg Leu Leu Pro Pro Ile Pro Val Asp Glu Ile Pro Asn Pro Val Val
 465 470 475 480

Val Arg Ala Leu Thr Gln Val Arg Lys Val Leu Asn Ser Ile Ile Lys
 485 490 495

Lys Tyr Gly Pro Pro Ser Asp Ile Tyr Ile Glu Leu Ala Arg Glu Met
 500 505 510

Ser Lys Pro Phe Lys Glu Arg Lys Ser Leu Glu Arg Glu Phe Asn Glu
 515 520 525

Asn Arg Gln Ile Asn Glu Lys Ala Lys Ala His Ile Ser Glu Leu Tyr
 530 535 540

Arg Ile Pro Asn Asp Pro Arg Pro His Asp Ile Leu Lys Phe Lys Leu
 545 550 555 560

Trp Asn Glu Gln Asn Gly Ile Cys Pro Tyr Ser Leu Lys Pro Ile Ser
 565 570 575

Ile Glu Tyr Leu Phe Asn Ile Gly Tyr Ala Glu Val Asp His Ile Ile
580 585 590

Pro Tyr Ser Arg Ser Phe Asp Asp Ser Asn Gly Asn Lys Val Leu Val
595 600 605

Leu Thr Arg Glu Asn Gln Asn Lys Leu Asn Arg Thr Pro Tyr Glu Trp
610 615 620

Phe Gly His Glu Glu Asn Arg Trp Glu Asp Phe Val Ser Phe Ile Arg
625 630 635 640

Thr Met Lys Val Gly Lys Lys Lys Lys Asn Met Leu Leu Lys Lys Asn
645 650 655

Phe Asp Glu Glu Gln Glu Glu Gln Ile Leu Ser Arg Asn Leu Asn Asp
660 665 670

Thr Arg Tyr Ile Thr Arg Tyr Ile Lys Ser Phe Ile Glu Asp Asn Leu
675 680 685

Glu Phe Arg Thr Glu Glu Asn Lys Glu Gln Tyr Val His Thr Val Asn
690 695 700

Gly Ala Tyr Thr Ser Leu Met Arg Lys Arg Trp Gly Leu Asn Lys Asp
705 710 715 720

Arg Arg Gly Asn Asp Leu His His Ala Val Asp Ala Ala Ile Ile Ala
725 730 735

Val Ser Leu Pro Phe Lys Asn Lys Val Asn Ala Tyr Phe Lys Arg Gln
740 745 750

Glu Thr Gly Leu Ser Lys Leu Leu Asn Asn Lys Lys Asp Ile Phe Pro
755 760 765

Glu Pro Trp Arg Asn Phe Ile Lys Glu Leu Glu Ala Arg Met Ile Gln
770 775 780

Asp Pro Glu Lys Met Lys Arg Ala Leu Glu Ser Leu Glu Leu Glu Thr

785					790					795					800
Tyr	Gly	Glu	Ile	Phe	Leu	Asn	Lys	Leu	Lys	Pro	Ile	Phe	Val	Ser	Arg
				805					810					815	
Met	Pro	Lys	His	Ser	Ile	Lys	Gly	Gln	Ile	His	Glu	Glu	Thr	Ile	Arg
			820					825					830		
Arg	Val	Arg	Gly	Phe	Thr	Glu	Glu	Gly	Phe	Leu	Val	Thr	Val	Lys	Lys
		835					840					845			
Thr	Arg	Leu	Asp	Gln	Ile	Pro	Phe	Asp	Lys	Asn	Gly	Asp	Phe	Pro	Met
	850					855					860				
Tyr	Gly	Lys	Glu	Thr	Asp	Ile	Lys	Thr	Tyr	Met	Ala	Ile	Lys	Gln	Arg
865					870					875					880
Tyr	Leu	Glu	Tyr	Gly	Gln	Asp	Lys	Gln	Lys	Ala	Phe	Ala	Val	Pro	Leu
				885					890					895	
Arg	Lys	Pro	Ser	Lys	Asn	Pro	Lys	Asn	Ala	Pro	Ile	Val	Arg	Ser	Val
			900					905					910		
Lys	Ile	Glu	Gly	Lys	Ala	Asn	Arg	Val	Val	Met	Leu	Asp	Asp	Lys	Ala
		915					920					925			
Ala	Ala	Asp	Asn	Ala	Ser	Ile	Val	Arg	Thr	Glu	Val	Phe	Arg	His	Lys
	930					935					940				
Lys	Thr	Gly	Glu	Tyr	Tyr	Leu	Thr	Pro	Val	Tyr	Val	Ala	Asp	Ile	Leu
945					950					955					960
Ser	Asn	Lys	Ile	Pro	Asp	Arg	Leu	Ile	Thr	Ile	Lys	Lys	Ser	Tyr	Ser
				965					970					975	
Asp	Trp	Asp	Arg	Ile	Thr	Asp	Glu	His	Glu	Tyr	Leu	Phe	Ser	Leu	Tyr
			980					985					990		
Asn	Asn	Asp	Leu	Val	Lys	Ile	Ile	Leu	Pro	Lys	Glu	Lys	Glu	Thr	Lys
		995					1000					1005			

Lys Tyr Thr Gly Gly Asn His Leu Trp Gln Glu Gly Phe Phe Tyr
1010 1015 1020

Phe Lys Gly Val Asp Ser Ser Asn Ala Gly Ile Lys Ile Ile Asn
1025 1030 1035

His Leu Asn Ser Phe Glu Ala Arg Ile Gly Thr Lys Arg Leu Ile
1040 1045 1050

Ala Phe Glu Lys Tyr Gln Val Asn Pro Leu Gly Glu Ile Asn Lys
1055 1060 1065

Val His Gly Glu Lys Arg Pro Gly Glu Leu Leu Asn Lys Glu Glu
1070 1075 1080

Ile Lys Glu Asn Arg Lys Asn Ile Ser
1085 1090

<210> 124
<211> 1735
<212> PRT
<213> Pedobacter sp.

<223> APG06622_RGN

<400> 124

Met Thr Lys Asn Ile Leu Gly Leu Asp Leu Gly Thr Asn Ser Ile Gly
1 5 10 15

Trp Ala Leu Ile Lys Ile Asp His Glu Asn Lys Thr Leu Lys Ile Ile
20 25 30

Ala Leu Gly Ser Arg Ile Ile Pro Met Asp Gly Gln Glu Phe Ser Lys
35 40 45

Phe Asn Ala Gly Gln Lys Ile Glu Ser Ala Ala Gly Asn Arg Thr Ala
50 55 60

Leu His Arg Ala Arg Ile Thr Lys Glu Arg Tyr Leu Leu Arg Arg Asp
65 70 75 80

Arg Leu His Leu Val Leu Asn Leu Leu Glu Ala Leu Pro Asn His Tyr
85 90 95

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

Lys Glu Gly Lys Glu Pro Lys Met Ala Tyr Leu Pro Thr Lys Asn Leu
115 120 125

Glu Asn Lys Phe Asp Phe Tyr Phe Glu Glu Ala Phe Asn Glu Met Ile
130 135 140

Glu Asp Leu Gln Lys Val Asn Pro Glu Ile Lys Asn Glu Lys Lys Lys
145 150 155 160

Arg Val Pro Lys Asp Trp Thr Ile Tyr Phe Leu Arg Asn Lys Ala Met
165 170 175

Ser Gln Lys Val Ser Leu Glu Glu Leu Ala Trp Val Leu Leu Ser Tyr
180 185 190

Asn Gln Lys Arg Gly Thr Asp Ser Glu Glu Ile Glu Asn Glu Asp Glu
195 200 205

Lys Ser Glu Ile Lys Glu Gln Ile Asp Leu Lys Val Ile Lys Ala Glu
210 215 220

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

Ala Glu Asn Phe Lys Tyr Lys Glu Tyr Thr Thr Lys Gln Leu Thr Phe
245 250 255

Glu Asp Asp Val Lys Glu Val Thr Lys Ile Leu Lys Leu Asp Glu Ala
260 265 270

Gly Asn Ile Ile Asn Glu Lys Thr Ala Tyr Gln Ile Ser Asp Leu Tyr
275 280 285

Asn Leu Thr Leu Ala Asn Ile Glu His Thr Asp Ile Asn Asp Lys Lys
290 295 300

Ala Lys His Lys Tyr Asp Phe Ile Tyr Thr Asn Gly Trp Cys Thr Glu
305 310 315 320

Lys Lys Lys Glu Asn Phe Asp Val Thr Tyr Gln Lys Leu Glu Lys Gln
325 330 335

Ile Lys Glu Glu Ser Lys Val Ile Thr Glu Leu Phe Val Val Ser Asn
340 345 350

Lys Tyr Asp Phe Ala Gly Lys Pro Glu Ser Ile Ile Pro Ser Ile Lys
355 360 365

Leu Pro Asp Phe Asn Ser Glu Gly Ser Lys Asp Trp Thr Leu Leu Lys
370 375 380

Lys Lys Thr Glu Lys Glu Ile Val Lys Ser Glu Lys Thr Leu Ser Gly
385 390 395 400

Trp Ile Tyr Asp Val Leu Lys Asn Asp Val Asn Thr Gly Glu Arg Thr
405 410 415

Lys Ile Ile Gly Gly Leu Phe Gln Thr Ile Asp Arg Lys Phe Tyr Arg
420 425 430

Glu Glu Leu Tyr Gln Ile Ile Asn Thr Gln Arg Lys Phe His Thr Ser
435 440 445

Thr Leu Asp Asp Lys Asp Thr Phe Glu Lys Cys Val Lys Leu Leu Tyr
450 455 460

Pro His Asn Lys Asp His Ser Lys Thr Leu Leu Glu Asn Lys Gln Ala
465 470 475 480

Leu Thr Asn Leu Leu Ile Glu Asp Ile Leu Leu Tyr Gln Arg Pro Leu
485 490 495

Lys Ser Lys Lys Ser Glu Ile Ser Asp Cys Lys Tyr Glu Ile Asp Tyr
500 505 510

Trp Lys Glu Asp Val Asn Ile Ser Thr Gly Glu Val Val Glu Ile Pro
515 520 525

Val Tyr Lys Lys Ala Ile Phe Ala Ser His Pro Leu Phe Gln Glu Phe
530 535 540

Arg Ile Trp Asp Lys Ile His Asn Ile Lys Leu Ile Gln Leu Glu Ser
545 550 555 560

Lys Asp Ala Asp Gly Asn Thr Gln Thr Asn Val Asp Ile Thr Ser Gln
565 570 575

Tyr Phe Lys Pro Glu Asn Tyr Arg Ala Leu Phe Asn His Phe Asn Gly
580 585 590

Arg Ser Thr Val Ala Ile Asp Asp Phe Leu Lys Phe Cys Glu Glu Gln
595 600 605

Phe Lys Leu Gly Ile Gly Lys Lys Gly Glu Arg Lys Ile Leu Trp Asn
610 615 620

Tyr Pro Val Glu Glu Glu Phe Lys Gly Asn Glu Thr Arg Lys Gly Phe
625 630 635 640

Glu Asn Arg Phe Arg Lys Cys Gly Phe Asp Gly Phe Asp Asp Phe Met
645 650 655

Thr Gln Gln Lys Glu Ile Glu Leu Trp His Tyr Leu Tyr Ser Val Ser
660 665 670

Ser Ala Glu Arg Lys Lys Val Ser Gln Val Lys Asp Ser Ser Asp Pro
675 680 685

Lys Phe Gly Lys Thr Gly Ile Trp Asn Phe Phe Thr Lys Tyr Phe Lys
690 695 700

Asp Glu Asn Ile Ser Leu Glu Thr Leu Glu Lys Leu Cys Arg Asp Phe
705 710 715 720

Glu Thr Tyr Pro Lys Phe Thr Ser Lys Tyr Ala Ser Tyr Ser Val Lys
725 730 735

Ala Leu Asn Lys Met Leu Ala Val Met Arg Val Gly Lys Asn Phe Leu
740 745 750

Thr Ala Glu Ser Val Asn Glu Lys Trp Gln Glu Lys Tyr Ile Glu Arg
755 760 765

Ala Asp Thr Ile Ile Lys Lys Asn Asn Glu Ile Asp Trp Ser Ala Glu
770 775 780

Ser Ile Asp Leu Thr Asp Ile Ile Leu Asn Asp Val Asn Val Lys Asn
785 790 795 800

Gly Glu Ile Pro Phe Pro Lys Gly Leu Phe Asn Thr Phe Lys Ser Phe
805 810 815

Glu Thr Val Asp Asp Phe Glu Phe Leu Asn Leu Thr Gln Ala Ser Tyr
820 825 830

Phe Ile Tyr Gly Arg His Ser Glu Leu Ala Gln Ala Lys Tyr Trp Asn
835 840 845

Ser Pro Thr Lys Ile Arg Glu Glu Ile Thr Lys Glu Leu Lys His His
850 855 860

Ser Leu Asn Asn Pro Thr Ala Glu Lys Val Leu Lys Glu Thr Leu Lys
865 870 875 880

Val Val Ala Asp Ile Trp Glu Ile Tyr Gly Asn Ser Glu Glu Lys Phe
885 890 895

Phe Asn Glu Ile His Leu Glu Val Ala Arg Glu Leu Gln Lys Ser Asn
900 905 910

Gln Glu Lys Lys Asp Ile Ile Glu Arg Gln Lys Asn Asn Arg Ala Glu
915 920 925

Asn Asn Arg Ile Arg Asn Ile Leu Glu Glu Phe Leu Ser Ser Ser Pro
930 935 940

Tyr Asn Ala Lys Arg Gly Asn Gln Asp His Phe Glu Arg Leu Lys Ile
945 950 955 960

Val Glu Asp Gly Ala Lys Val Arg Ser Tyr Ser Asp Lys Glu Phe Tyr
965 970 975

Ser Asn Asn Lys Asp Lys Phe Ser Lys Ile Asp Ile Glu Val Ile Leu
980 985 990

Lys Lys Lys Lys Ile Thr Gln Ser Asp Phe Glu Lys Tyr Lys Leu Trp
995 1000 1005

Ile Glu Gln Gly Tyr Lys Ser Pro Tyr Thr Asp Gln Ile Ile Lys
1010 1015 1020

Leu Thr Asp Leu Phe Asp Gly Asn Lys Tyr Asn Val Asp His Ile
1025 1030 1035

Phe Ser Arg Ala Ala Ile Thr Asn Asp Ser Leu Asn Asn Lys Val
1040 1045 1050

Val Cys Glu Ala Val Ile Asn Arg Phe Lys Ser Tyr Lys Thr Gly
1055 1060 1065

Arg Glu Phe Ile Ala Gln Phe Gly Gly Lys Thr His Thr Ile Phe
1070 1075 1080

Asp Asp Lys Leu Asn Lys Asn Val Thr Phe Ser Leu Ile Asp Glu
1085 1090 1095

Asp Ala Tyr Val Ser Leu Val Lys Ser Gln Phe Lys Asp Ala Lys
1100 1105 1110

Arg Leu Ile Leu Leu Ser Lys Glu Val Pro Lys Ala Phe Thr Glu
1115 1120 1125

Ser Gln Leu Asn Asn Ala Lys Tyr Ile Ala Arg Lys Ala Met Glu
1130 1135 1140

Leu Leu Ser His Ile Val Arg Glu Glu Gly Glu Val Glu Phe Arg
1145 1150 1155

Ser Lys Asn Val Leu Pro Val Ser Gly Ala Ile Thr Asp Arg Leu
1160 1165 1170

Lys Lys Glu Trp Lys Ile Asn Gln Ile Trp Thr Glu Leu Leu Lys
1175 1180 1185

Pro Arg Phe Glu Arg Leu Asn Lys Ile His Asn Ser Glu Ala Phe
1190 1195 1200

Gly Ser Tyr Ala Ile Ser Lys Ser Gly His Gln Tyr Phe Asp Ile
1205 1210 1215

Asn Thr Lys Tyr Ile Leu Gly Glu Asn Asp Lys Phe Asp Leu Lys
1220 1225 1230

Arg Leu Asp His Arg His His Ala Leu Asp Ala Leu Ile Ile Ala
1235 1240 1245

Leu Cys Thr Asp Asn His Val Gln Tyr Phe Asn Asn Ile Asn Ser
1250 1255 1260

Gly Ile Thr Asn Lys Lys Lys Gly Lys Met Glu Ala Ile Lys Lys
1265 1270 1275

Gln Arg Ala Gly Ile Lys Arg Gln Ile Met Tyr Ser Glu Lys Asp
1280 1285 1290

Lys Glu Asn Pro Asn Glu Lys Val Trp Arg Tyr Met Leu Pro Gly
1295 1300 1305

Ser Phe Arg Asn Lys Glu Ser Val Asn Asn Glu Lys Asp Ser Val
1310 1315 1320

Val Asp Leu Ser Trp Glu Asn Asn Tyr Ser Ala Lys Glu Ser Asn
1325 1330 1335

Asp Tyr Arg Lys Thr Ile Leu Glu Cys Leu Asp Asn Cys Ile Val
1340 1345 1350

Ser Phe Lys Asn Asp Phe Lys Ile Val Ser Lys Ser Ser Asn Lys
1355 1360 1365

Tyr Glu Ser Tyr Tyr Asp Glu Asn Gly Asn Leu Arg Leu Asp Ser
1370 1375 1380

Asn Gly Lys Pro Gln Lys Glu Phe Ile Ser Gln Gln Asn Asn Asn
1385 1390 1395

Gln Lys His Trp Ser Val Arg Lys Pro Leu His Thr Asp Asn Pro
1400 1405 1410

Ser Ala Glu Ile Thr Leu Gln Phe Asp Arg Leu Lys Leu Ile Asp
1415 1420 1425

Asn Leu Gly Lys Val Asp Leu Ile Ile Glu Glu Lys Ile Lys Asp
1430 1435 1440

Ala Val Lys Glu Ile Leu Gly Met His Asn Asn Lys Ile Gly Lys
1445 1450 1455

Ala Glu Ser Tyr Leu Lys Lys Asn Pro Leu Ile Leu Glu Glu Lys
1460 1465 1470

Lys Ile Glu Tyr Thr Asp Phe Lys Ile Ser Asn Lys Lys Tyr Arg
1475 1480 1485

Lys Arg Gln Pro Ile Thr Asn Leu Ser Val Arg Thr Gly Thr Gly
1490 1495 1500

Ala Leu Pro Thr Tyr Asp Ser Val Val Asp Arg Ile Asn Lys Val
1505 1510 1515

Ser Asp Phe Val Leu Arg Asn Glu Leu Leu Gln His Leu Lys Ile
1520 1525 1530

Asn Asp Asn Asp Ile Asp Lys Ala Phe Ser Ile Glu Gly Ile Asn
1535 1540 1545

Glu Phe Asn Ser Asn Arg Lys Ile Pro Val Phe Arg Leu Pro Ile
1550 1555 1560

Ser Glu Ser Ala Thr Leu Lys Phe Lys Ile Gly His Lys Asn Ser
1565 1570 1575

Asn Ser His Lys Trp Val Glu Thr Gly Gly Asn Phe His Phe Arg
1580 1585 1590

Ile Ala Asp Glu Glu Asn Ser Arg Asp Tyr Glu Thr Ile Pro Leu
1595 1600 1605

Arg Ile Ala Ile Glu Met Glu Lys Asp Asn Ile Ile Arg Arg Gln
1610 1615 1620

Ile Val Asp Asn Glu Asn Glu Ile Val Leu Ser Pro Asn Asp Leu
1625 1630 1635

Val Tyr Val Pro Leu Glu Lys Asn Glu Ile Val Asp Phe Asn Asn
1640 1645 1650

Leu Thr Ile Gln Gln Ile Lys Arg Ile Tyr Lys Phe Val Phe Ser
1655 1660 1665

Thr Lys Lys Glu Ala His Phe Ile Pro Tyr Ser Ser Ser Ala Glu
1670 1675 1680

Ile Ile Lys Asn Glu Asn Gly Thr Asn Gly Tyr Ser Glu Arg Thr
1685 1690 1695

Lys Asp Phe Phe Asp Gly Asn Val Leu Gln Asp Asp Lys Gly Lys
1700 1705 1710

Pro Ile Gln Val Lys Ser Val Cys Trp Lys Leu Lys Val Asp Arg
1715 1720 1725

Leu Gly Asn Ile Ser Lys Ala
1730 1735

<210> 125
<211> 1224
<212> PRT
<213> Chitinophaga sp.

<223> APG02787_RGN

<400> 125

Met Lys Lys Leu Ile Val Gly Leu Asp Leu Gly Val Asn Ser Ile Gly
1 5 10 15

Trp Ala Val Val Gly Asp Asp Asp Ser Asn Arg Gln Ile Leu Gly Met
20 25 30

Gly Ser Arg Ile Ile Pro Leu Ser Pro Asp Asp Asn Asn Glu Phe Val
35 40 45

Thr Gly Asn Ala Ile Ser Lys Asn Ala Lys Arg Thr Glu Lys Arg Thr
50 55 60

Gln Arg Lys Gly Tyr Asp Arg Tyr Gln Gln Arg Arg Arg Lys Leu Val
65 70 75 80

Ala Phe Leu Glu Arg His Ser Met Met Val Glu Lys Glu Leu Met Cys
85 90 95

Leu Pro Pro Leu Glu Leu Trp Ala Leu Arg Asn Lys Gly Ile Thr Gln
100 105 110

Ala Leu Ser Ser Arg Glu Leu Gly Arg Val Leu Tyr His Leu Asn Gln
115 120 125

Arg Arg Gly Tyr Arg Ser Ser Arg Ser Glu Glu Ser Ser Asp Gly Lys
130 135 140

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

Lys Thr Arg Gly Ile Thr Ile Gly Gln His Phe Tyr Gln Lys Leu Gln
165 170 175

Thr Ala His Asn Arg Asn Glu Pro Tyr Arg Ile Lys Gln Gln Val Phe
180 185 190

Pro Arg Gln Ala Tyr Ile Asp Glu Phe Asp Ala Ile Ile Asn Gln Gln

195

200

205

Lys Asn His His Ser Ser Leu Thr Ala Asp Glu Ile Ala Phe Leu Arg
 210 215 220

Asp Glu Ile Ile Tyr Met Gln Arg Pro Leu Lys Ser Gln Lys Gly Leu
 225 230 235 240

Val Ser Ile Cys Glu Phe Glu Gly Gln Tyr Arg Lys Asn Ala Lys Asp
 245 250 255

Ala Glu Val Phe Val Gly Pro Lys Val Ala Pro Arg Ser Ser Pro Leu
 260 265 270

Ala Gln Val Cys Lys Ile Trp Glu Ser Ile Asn Asn Ile Thr Phe Thr
 275 280 285

Asn Lys Arg Arg Glu Pro Tyr Pro Ile Asn Thr Asp Gln Lys Leu Arg
 290 295 300

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

Phe Phe Ile Leu Gly Ile Lys Pro Gly Asp Gly Tyr Thr Gly Asn Lys
 325 330 335

Gln Thr Ala Lys Gly Ile Gln Gly Asn Leu Thr Asn Ala Ala Ile Ser
 340 345 350

Thr Ile Ile Pro Glu Arg Lys Asp Leu Leu Asn Phe Glu Leu Lys Ile
 355 360 365

Val Val Asp Glu Glu Asp Ala His Leu Ile Asn Arg Ser Asn Gly Glu
 370 375 380

Ile Ile Ser Thr Val Gln Lys Lys Thr Val Ser Ala Asp Tyr Glu Arg
 385 390 395 400

Ser Pro Phe Phe Arg Leu Trp His Val Ile Tyr Ser Ile Ser Asp Ile
 405 410 415

Asp Lys Cys Ile Glu Thr Leu Met Glu Lys Phe Gly Leu Thr Lys Glu
420 425 430

Gln Ala Val Ala Leu Ser Arg Leu Asp Phe Arg Lys Ala Gly Phe Ala
435 440 445

Asn Lys Ser Ala Arg Ala Met Arg Lys Ile Leu Pro Tyr Leu Met Asp
450 455 460

Gly His Ile Phe Ser Asp Ala Met Ser Leu Ala Gly Tyr Asn His Ser
465 470 475 480

Ala Ser Leu Thr His Glu Glu Asn Leu Lys Arg Gln Leu Lys Gly Ser
485 490 495

Leu Pro Leu Leu Lys Lys Asn Ser Leu Arg Gln Pro Val Val Glu Lys
500 505 510

Ile Leu Asn Gln Met Ile Gly Val Met Asn Ala Ile Ile Lys Glu Trp
515 520 525

Gly Arg Pro Asp Glu Ile Arg Val Glu Leu Ala Arg Glu Leu Lys Gln
530 535 540

Ser Ala Glu Glu Arg Asn Lys Ala Phe Ala Ala Leu Ser Lys Arg Glu
545 550 555 560

Arg Glu Gly Glu Ile Ile Thr Lys Arg Ile Glu Ala Glu Phe Arg Gln
565 570 575

Phe Gly Ile Arg Ala Thr Arg Asn Asn Ile Leu Lys Trp Arg Leu Phe
580 585 590

His Glu Ile Ser Asn Asp Glu Ala Arg Leu Asn Ala Thr Cys Val Tyr
595 600 605

Cys Gly His Pro Phe Gly Ile Thr Asp Ala Leu Arg Gly Ser Ser Val
610 615 620

Asp Ile Glu His Ile Ile Pro Lys Ser Lys Leu Phe Asp Asp Ser Gln

625		630		635		640									
Gly	Asn	Lys	Thr	Leu	Val	His	Arg	Arg	Cys	Asn	Glu	Asp	Lys	Gly	Asp
				645					650					655	
Arg	Thr	Ala	Phe	Asp	Phe	Met	Ser	Ala	Lys	Gly	Glu	Ser	Val	Leu	Ala
			660					665					670		
Ser	Tyr	Ile	Glu	Thr	Val	Asp	Arg	Leu	Tyr	Gln	Ser	Asn	Val	Ile	Ser
		675					680					685			
Arg	His	Lys	Arg	Asp	Arg	Leu	Leu	Thr	Thr	Gly	Asp	Asp	Ile	Pro	Lys
	690					695					700				
Asp	Phe	Ile	Glu	Arg	Gln	Leu	Arg	Glu	Thr	Gln	Tyr	Ile	Ala	Arg	Lys
705					710					715					720
Ala	Lys	Glu	Ile	Leu	Glu	Gln	Ile	Cys	Tyr	Asn	Val	Trp	Ser	Thr	Ser
				725				730						735	
Gly	Gln	Ile	Thr	Glu	Tyr	Leu	Arg	Arg	Ile	Trp	Gly	Trp	Asn	Asp	Val
			740					745					750		
Leu	Met	Asn	Leu	Gln	Leu	Pro	Val	Tyr	Arg	Asp	Leu	Gly	Leu	Thr	Glu
		755					760					765			
Trp	Lys	Glu	Tyr	Thr	Asp	Arg	Asp	Gly	Asn	Leu	His	Arg	Gln	Glu	Val
	770					775					780				
Ile	Lys	Asp	Trp	Ser	Lys	Arg	Asn	Asp	His	Arg	His	His	Ala	Ile	Asp
785					790				795						800
Ala	Leu	Thr	Ile	Ala	Cys	Thr	Arg	Gln	Gly	Phe	Ile	Gln	Arg	Ile	Asn
				805					810					815	
Arg	Leu	Asn	Ala	Thr	Asn	Asn	Arg	Glu	Glu	Met	Arg	Ala	Ser	Leu	Asn
			820					825					830		
Gly	Asn	Tyr	Asp	Pro	Lys	Lys	Asn	Leu	Leu	Glu	Asn	Tyr	Ile	Phe	Ala
		835					840					845			

Leu Lys Pro Phe Ser Thr Lys Gln Val Glu Val Ala Ala Ala Gln Ile
850 855 860

Leu Ile Ser Phe Lys Pro Gly Lys Arg Val Ala Thr Leu Ser Lys Leu
865 870 875 880

Lys Ala Lys Gly Lys Asn Ser Gly Thr Gly Val Ile Val Pro Arg Gly
885 890 895

Pro Leu Ser Glu Glu Ser Val Tyr Gly Arg Val Arg Leu Met Asp Ser
900 905 910

Lys Lys Gln Pro Val Arg Tyr Leu Phe Glu His Pro Glu Leu Ile Phe
915 920 925

Lys Pro Tyr Ile Arg Val Leu Val Glu Glu Arg Leu Ala Ser Tyr Ser
930 935 940

Gln Asp Leu Lys Lys Ala Leu Ala Ser Leu Lys Lys Asp Pro Ile Tyr
945 950 955 960

Leu Asp Gly Thr Lys Thr Ile Ile Leu Glu Trp Gly Ser Cys Tyr Lys
965 970 975

Asp Glu Val Val Ile Lys Tyr Pro Leu Thr Asp Ile Lys Ala Lys Asp
980 985 990

Thr Gln Phe Ile Ile Asp Gly His Ile Arg Ala Ile Val Asn Glu Arg
995 1000 1005

Leu Ser Gln Tyr Lys Gly Asn Glu Lys Lys Ala Phe Glu His Pro
1010 1015 1020

Leu Tyr Ser Asp Ala Ala Cys Thr Lys Glu Ile Arg Ser Val Arg
1025 1030 1035

Cys Leu Thr Gly Leu Thr Ala Val Val Ser Val Lys Lys Asp Glu
1040 1045 1050

Ser Gly Arg Glu Ile Gly Phe Val Lys Pro Gly Asn Asn His His

1055

1060

1065

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

Cys Thr Phe Trp His Ala Val Glu Arg Lys Lys Trp Gly Met Pro
 1085 1090 1095

Ala Val Ile Thr Gln Pro Gly Val Ile Trp Asp Asn Ile Leu Gln
 1100 1105 1110

Lys Gln Asp Asn Phe Pro Gln Ala Phe Leu Glu Lys Leu Pro Leu
 1115 1120 1125

Glu Ala Trp Thr Phe Asp Lys Ser Leu Gln Gln Asn Glu Met Phe
 1130 1135 1140

Ile Val Gly Leu Thr Ile Glu Glu Val Gln Thr Cys Ile Asp Ser
 1145 1150 1155

Phe Asp Tyr Ala Arg Leu Ser Lys His Leu Tyr Arg Ile Gln Lys
 1160 1165 1170

Leu Ser Ser Thr Gly Tyr Lys Ile Val Leu Arg His His Leu Ala
 1175 1180 1185

Thr Thr Ile Glu His Thr Arg Asp Met Ile Gly Ile Ala Ser Pro
 1190 1195 1200

Ser Ser Phe Thr Gly Ile Lys Val Tyr Val Ser Lys Thr Gly Lys
 1205 1210 1215

Ile Ser Met Pro His Asp
 1220

<210> 126

<211> 1457

<212> PRT

<213> Mucilaginibacter sp.

<223> APG06248_RGN

<400> 126

Met Lys Asn Ile Leu Gly Leu Asp Leu Gly Thr Asn Ser Ile Gly Trp
1 5 10 15

Ala Leu Ile Gln Ser Asp Phe Asp Asn Lys Glu Gly Lys Ile Leu Gly
20 25 30

Thr Gly Ser Arg Ile Ile Pro Met Ser Gln Asp Ile Lys Asp Glu Tyr
35 40 45

Gly Lys Gly Asn Ser Ile Ser Gln Thr Ala Asp Arg Thr Arg Phe Arg
50 55 60

Ser Leu Arg Arg Leu Arg Glu Arg His Leu Leu Arg Arg Glu Arg Leu
65 70 75 80

His Arg Ile Leu Asn Ile Leu Gly Phe Leu Pro Glu His Tyr Ala Asp
85 90 95

Gln Ile Asp Phe Lys Asn Arg Leu Gly Gln Tyr Leu Pro Glu Lys Glu
100 105 110

Pro Lys Leu Ala Tyr Arg Asn Asn Glu Phe Ile Phe Lys Ala Ser Phe
115 120 125

Glu Glu Met Leu Ile Asp Phe Arg Glu Gln Gln Pro Glu Met Leu Gln
130 135 140

Gly Asp Lys Lys Val Pro Tyr Asp Trp Thr Ile Tyr Tyr Leu Arg Lys
145 150 155 160

Lys Ala Leu Thr Gln Lys Val Glu Lys Ala Glu Leu Ala Trp Leu Ile
165 170 175

Leu Asn Phe Asn Gln Lys Arg Gly Tyr Tyr Gln Leu Arg Gly Glu Glu
180 185 190

Gln Gln Glu Asn Pro Asn Lys Lys Val Glu Phe His Ser Leu Lys Val
195 200 205

Val Asp Val Ile Ala Asp Glu Thr Gln Arg Gly Lys Ala Glu Thr Trp
210 215 220

Tyr Ser Leu Ile Leu Glu Asn Gly Trp Val Tyr Arg Arg Ser Ser Lys
225 230 235 240

Thr Pro Leu Asp Asn Trp Arg Asp Lys Gln Arg Asp Phe Ile Val Thr
245 250 255

Thr Asp Leu Asn Asp Asp Gly Ser Ile Lys Lys Asp Lys Asp Gly Asn
260 265 270

Ile Lys Arg Ser Phe Arg Ala Pro Asp Glu Asn Asp Trp Thr Leu Leu
275 280 285

Lys Lys Lys Thr Glu Gln Asp Ile Asp His Ala Asn Lys Thr Ala Gly
290 295 300

Ala Tyr Ile Tyr Asp His Leu Leu Val Ala Pro Asp Gln Lys Ile Lys
305 310 315 320

Gly Lys Leu Val Arg Thr Ile Glu Arg Arg Phe Tyr Lys Thr Glu Leu
325 330 335

Lys Ala Ile Leu Ala Lys Gln Gln Asn Phe His Ala Glu Leu Gln Asn
340 345 350

Lys Asp Leu Phe Leu Asp Cys Val Arg Glu Leu Tyr Arg His Asn Leu
355 360 365

Ser His Leu Ser Asn Leu Asn Gln Lys Asp Thr Ile His Leu Leu Thr
370 375 380

Glu Asp Ile Ile Phe Tyr Gln Arg Pro Leu Arg Ser Gln Lys Ser Ser
385 390 395 400

Ile Gly Asn Cys Pro Leu Glu Phe Arg Arg Tyr Gln Asn Lys Glu Gly
405 410 415

Gly Asp Gln Thr Glu Tyr Leu Lys Val Ile Pro Lys Ser His Pro Leu
420 425 430

Tyr Gln Glu Phe Arg Leu Trp Gln Trp Ile Tyr Asn Leu Gly Ile Tyr
435 440 445

Gln Lys Glu Asp Asp Thr Asn Val Thr Ala Asp Leu Ile Lys Gly Ile
450 455 460

Gln Gly Tyr Gln Asp Leu Phe Asp Phe Leu Asn Ser Lys Lys Glu Ile
465 470 475 480

Lys Gln Asp Thr Leu Ile Lys Phe Leu Leu Ser Gln Thr Gly Leu Lys
485 490 495

Gly Lys Glu Leu Asn Lys Ala Ala Leu Thr Leu Arg Trp Asn Phe Val
500 505 510

Glu Asp Lys Thr Tyr Pro Cys Asn Glu Thr Gly Thr Glu Ile Arg Ala
515 520 525

Arg Leu Ala Lys Val Ala Tyr Ile Pro Glu Asn Phe Leu Thr Ala Glu
530 535 540

Lys Ala Leu Ser Leu Trp His Ile Ile Tyr Ser Val Asn Asp Lys Lys
545 550 555 560

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

Glu Glu Thr Phe Phe Ile Asn Phe Gln Lys Phe Pro Pro Phe Gln Ser
580 585 590

Glu Tyr Gly Ala Phe Ser Glu Lys Ala Ile Lys Lys Leu Leu Pro Leu
595 600 605

Ile Arg Leu Gly Lys Tyr Trp His Trp Glu Ala Ile Asp Pro Lys Ser
610 615 620

Gln Glu Arg Ile Glu Lys Ile Leu Ser Gly Glu Ala Asp Gln Asn Ile
625 630 635 640

Lys Glu Arg Ile Arg Glu Lys Ala Met His Leu Thr Arg Glu Glu His
645 650 655

Phe Gln Gly Leu Gln Leu Trp Leu Ala Asn Tyr Ile Val Tyr Asp Arg
660 665 670

His Ala Glu Ala Ala Asn Pro Asp Lys Trp Ser Thr Val Ser Asp Leu
675 680 685

Glu Asp Tyr Leu Lys Glu Leu Lys Gln His Ser Leu Arg Asn Pro Ile
690 695 700

Val Glu Gln Val Val Thr Glu Thr Leu Arg Val Val Lys Asp Ile Trp
705 710 715 720

Leu Thr Tyr Gly Asn Gly Ala Lys Asn Phe Phe Asp Glu Ile His Ile
725 730 735

Glu Leu Ala Arg Glu Met Lys Asn Thr Met Glu Glu Arg Lys Arg Leu
740 745 750

Thr Ser Ile Val Thr Asp Asn Glu Asn Thr Asn Leu Arg Ile Lys Ala
755 760 765

Leu Leu Val Glu Leu Phe Glu Asp Gly Asn Ile Glu Asn Val Arg Pro
770 775 780

Tyr Ser Pro Ile Gln His Glu Ala Leu Lys Ile Phe Glu Asp Gly Ile
785 790 795 800

Leu Asn Ser Gly Ile Glu Ile Asp Asp Asp Ile Leu Lys Ile Ser Lys
805 810 815

Thr Pro Gln Pro Thr Ser Thr Glu Leu Lys Arg Tyr Lys Leu Trp Leu
820 825 830

Glu Gln Lys Tyr Arg Ser Pro Tyr Thr Gly Asn Val Ile Pro Leu Asn
835 840 845

Arg Leu Phe Thr Thr Asp Tyr Glu Ile Glu His Ile Ile Pro Gln Ser
850 855 860

Arg Phe Phe Asp Asp Ser Phe Ser Asn Lys Ile Ile Cys Glu Ser Ala
865 870 875 880

Ile Asn Lys Leu Lys Asp Lys Gln Leu Gly Leu Glu Phe Ile Lys Asn
885 890 895

His Ala Gly Thr Ile Val Glu Leu Gly Leu Gly Lys Thr Ala Arg Val
900 905 910

Phe Asp Glu Glu Arg Tyr Gln Asp Phe Val Lys Ala Asn Tyr Thr Lys
915 920 925

Asn Pro Gly Lys Arg Ser Lys Leu Met Leu Glu Glu Ile Pro Asp Lys
930 935 940

Met Ile Ala Arg Gln Leu Asn Asp Thr Arg Tyr Ile Ser Lys Phe Ile
945 950 955 960

Thr Ser Val Leu Ser Asn Ile Val Arg Ser Glu Thr Lys Asp Asp Gly
965 970 975

Val Asn Ser Lys Asn Val Leu Leu Cys Asn Gly Arg Ile Thr Ser Val
980 985 990

Leu Arg Gln Asp Trp Gly Met Asn Asp Ile Trp Asn Glu Leu Ile Leu
995 1000 1005

Pro Arg Phe Glu Arg Leu Asn Gln Ile Thr Asp Ser Thr Val Phe
1010 1015 1020

Thr Ala Trp Asn Glu Asn His Gln Lys Phe Leu Pro Thr Val Pro
1025 1030 1035

Met Gly Phe Ser Lys Gly Phe Gln Arg Lys Arg Ile Asp His Arg
1040 1045 1050

His His Ala Met Asp Ala Leu Ile Ile Ala Ser Cys Thr Arg Asp
1055 1060 1065

His Ile Asn Leu Leu Asn Asn Glu Tyr Ala Lys Ser Ser Lys Arg
1070 1075 1080

Phe Asp Leu Asn Arg Lys Leu Arg Lys Phe Glu Lys Val Ala Tyr
1085 1090 1095

Lys Ser Pro Val Asn Gly Gln Gln Ile Asp Lys Asp Ile Pro Thr
1100 1105 1110

Val Phe Ile Lys Pro Trp Gln Arg Phe Thr Ile Asp Ala Lys Cys
1115 1120 1125

Ser Leu Glu Asp Ile Ile Val Ser Phe Lys Gln Asn Leu Arg Val
1130 1135 1140

Ile Asn Lys Thr Thr Asn Tyr Phe Asp Gly Tyr Lys Asp Glu His
1145 1150 1155

Gly Ile Val Asn Phe Asp Asn Lys Gly Ala Pro Val Lys Gln Gln
1160 1165 1170

Arg Lys Gln Thr Lys Gly Asp Ser Trp Ala Ile Arg Lys Pro Met
1175 1180 1185

His Lys Asp Phe Val Tyr Gly Lys Val Asp Leu Lys His Ile Lys
1190 1195 1200

Ile Pro Lys Gly Lys Phe Ala Thr Ala Ile Arg Lys Ser Leu Asp
1205 1210 1215

Thr Thr Phe Asp Gln Lys Lys Ile Gln Thr Ile Thr Asp Thr Gly
1220 1225 1230

Val Gln Lys Ile Leu Ile Asn Tyr Leu Gln Ser Lys Asp Asn Lys
1235 1240 1245

Ala Glu Leu Ala Phe Ser Pro Glu Gly Ile Glu Asp Met Asn Gln
1250 1255 1260

Asn Ile Arg Ala Tyr Asn Asp Gly Lys Leu His Gln Pro Ile Lys
1265 1270 1275

Lys Val Arg Val Phe Glu Ile Gly Ser Arg Phe Pro Leu Gly Glu
1280 1285 1290

Thr Gly Asn Lys Lys Thr Lys Tyr Val Glu Thr Ala Lys Gly Thr
1295 1300 1305

Asn Leu Phe Phe Ala Ile Tyr Gln His Lys Asn Lys Arg Asn Tyr
1310 1315 1320

Glu Thr Val Pro Leu Asn Thr Val Ile Glu Arg Gln Lys Gln His
1325 1330 1335

Leu Ser Pro Val Pro Thr Glu Asp Asp Lys Gly Gly Thr Leu Leu
1340 1345 1350

Phe Tyr Leu Ser Pro Asn Asp Leu Val Tyr Val Pro Thr Pro Asp
1355 1360 1365

Glu Leu Gly Asn Asn Leu Gly Gly Leu Asp Ile Arg Asn Leu Asn
1370 1375 1380

Lys Glu Gln Arg Arg Arg Ile Tyr Lys Val Glu Lys Ala Ser Gly
1385 1390 1395

Val Glu Cys Tyr Phe Ile Arg Gln Asp Ile Ala Tyr Leu Ile Lys
1400 1405 1410

Ser Tyr Asp Ala Thr Thr Lys Ile Gly Glu Leu Glu Ser Gln Asn
1415 1420 1425

Lys Leu Gln Thr Thr Met Ser Ser Asp Arg Leu Lys Ile Ala Asp
1430 1435 1440

Val Cys Ile Lys Leu Lys Ile Asp Arg Leu Gly Ile Ser Arg
1445 1450 1455

<210> 127
<211> 1662
<212> PRT
<213> Acidovorax sp.

<223> APG06007_RGN

<400> 127

Met Lys Lys Ile Leu Gly Leu Asp Ile Gly Thr Thr Ser Ile Gly Tyr
1 5 10 15

Ser Ile Ile Lys Val Ser Asp Glu Lys Asn Glu Asn Ile Ile Ser Asp
20 25 30

Lys Val Ala Ala Ser Asp Ile Asn Asn Asp Arg Ile Gly Ile Phe Glu
35 40 45

Glu Ala Ile Gly Val Arg Ile Ile Asn Gln Asp Thr Asp Arg Phe Asp
50 55 60

Arg Gly Leu Lys Leu Asn Asp Pro Lys Gly Ser Thr Leu Thr Pro Thr
65 70 75 80

Ala Asn Arg Arg Lys Phe Arg Gly Val Arg Arg Met Arg Asn Arg Tyr
85 90 95

Lys Leu Arg Arg Asn Lys Leu Tyr Lys Val Leu Ser Ile Leu Gly Leu
100 105 110

Thr Pro Asp Glu Ser Tyr Phe Thr Asn Lys Lys Gly Lys Arg Gly Lys
115 120 125

Asp Asn Asp Ile Gly Lys Ala Leu Tyr Gly Leu Arg Glu Arg Ala Ala
130 135 140

Arg Glu Lys Ile Ser Leu Ser Glu Leu Ala Arg Val Leu Leu His Leu
145 150 155 160

Asn Gln Leu Arg Gly Tyr Ser Ser Asp Arg Phe Ala Asn Asp Glu Glu
165 170 175

Thr Lys Phe Asp Tyr Tyr Leu Thr Glu Val Val Glu Ala Asp Trp Glu
180 185 190

Asn Arg Thr Ala Val Phe Glu Lys Gly Ser Thr Glu Asp Val Lys Tyr

195

200

205

Tyr Gln Phe Lys Ala Lys Leu Arg Phe Val Glu Pro Tyr Val Ile Val
 210 215 220

Glu Asp Asp Glu Lys Thr Ile Leu Tyr Glu Val Glu Gly Tyr Ile Phe
 225 230 235 240

Lys Lys Glu Leu Asp Phe Glu Val Gly Asp Phe Ile Thr Ile Lys Ala
 245 250 255

Pro Glu Phe Lys Gln Asn Lys Lys Gly Lys Val Val Ile Glu Glu Tyr
 260 265 270

Tyr Lys Ile Thr Pro Thr Lys Pro Asp Ala Thr Asp Trp Lys Tyr Lys
 275 280 285

Tyr Gln Thr Leu His Lys Asn Leu Ser Asp Trp Cys Lys Ala Gly Gly
 290 295 300

Thr Val Gly Ser Tyr Phe Tyr Lys Asn Phe Tyr Glu Leu Thr Thr Leu
 305 310 315 320

Ser Arg Ile Arg Asn Asn Val Val Asn Arg Glu Trp Tyr Glu Asn Glu
 325 330 335

Phe Asp Ala Ile Trp Thr Leu Gln Tyr Glu Phe Tyr Lys Asp Phe Phe
 340 345 350

Glu Lys Ile Asp Val Ser Glu Leu Val Lys Gly Val Phe Lys Asp Tyr
 355 360 365

Ser Ala Val Leu Asn Asp Val Ile Ser Lys Ser Gly Asn Lys Ala Gln
 370 375 380

Leu Gln Thr Leu Ile Lys Asp Lys Ile Ile Phe Tyr Gln Arg Pro Trp
 385 390 395 400

Gln Gln Ala Lys Ser Lys Gly Arg Cys Glu Phe Glu Lys Ile Lys Val
 405 410 415

Lys Arg Lys Phe Ala Ile Lys Gly Thr Gly Thr Lys Glu Leu Arg Glu
420 425 430

Glu His Ile Gly Arg Lys Val Val Pro Arg Ser His Pro Ile Phe Gln
435 440 445

Asp Phe Lys Val Trp Gln Gln Ile Asn Asn Val Arg Leu Ile Trp Asn
450 455 460

Thr Ala Asp Glu Lys Ile Asn Leu Phe Glu Asn Ala Glu Leu Ser Glu
465 470 475 480

Lys Leu Thr Gly Lys Thr Ile Glu Lys Val Lys Glu Leu Leu Phe Lys
485 490 495

Lys Leu Gln Gly Ser Lys Thr Leu Ser Trp Arg Thr Phe Val Lys Glu
500 505 510

Glu Leu Gly Leu Asp Thr Tyr Asp Glu Ile Glu Glu Lys Arg Lys Ser
515 520 525

Lys Val Ser Lys Lys Gln Gly Leu Asp Ile Glu Thr Gly Glu Ile Val
530 535 540

Asp Ser Phe Phe Ser Val Asn Phe Arg Lys Arg Lys Arg Asp Gly Thr
545 550 555 560

Phe Asp Asp Ile Lys Leu Lys Gly Asn Thr Thr Lys Ala Glu Leu Gln
565 570 575

Asn Ile Leu Lys Glu Lys Arg Ala Glu Trp Phe Asp Thr Leu His Asn
580 585 590

Gln Asn Lys Lys Val Thr Asn Leu Gln Leu Leu Trp Glu Ile Ile Tyr
595 600 605

Asp Ile Thr Asn Ser Asp Ala Thr Lys Val Ser Ser Ile Ile Lys Lys
610 615 620

His Phe Asp Phe Asp Glu Ser Val Cys Leu Thr Leu Ala Lys Val Gln

625		630		635		640									
Phe	Asp	Asp	Ala	Gly	Met	Ala	Asn	Leu	Ser	Ala	Lys	Ala	Ile	Arg	Asn
				645					650					655	
Leu	Leu	Pro	Leu	Met	Ser	Asp	Gly	Ala	Asn	Tyr	Ser	Lys	Ile	Thr	Lys
			660					665					670		
Ala	Lys	Val	Asp	Glu	Leu	Ile	Ser	Leu	Asn	Thr	Ser	Glu	Thr	Asp	Val
		675					680					685			
Asp	Ala	Asp	Asp	Lys	Leu	Asn	Gly	Leu	Lys	Asp	Phe	Val	Thr	Asp	Arg
	690					695					700				
Lys	Ala	Arg	Ile	Arg	Leu	Ser	Lys	Phe	Ser	Asn	Glu	Asn	Asp	Phe	Lys
705					710					715					720
Tyr	Leu	Asn	Tyr	Trp	Glu	Ala	Ala	Ala	Ile	Val	Tyr	Gly	Ser	His	Ser
				725					730					735	
Ser	Lys	Lys	Ser	Ser	Glu	Thr	Lys	Ile	Met	Glu	Arg	Val	Lys	Asn	His
			740					745					750		
Ser	Met	Asn	Asn	Pro	Val	Val	Glu	Lys	Ile	Val	Asn	Glu	Ala	Ile	Ser
		755					760					765			
Ile	Val	Asn	Leu	Leu	Lys	Glu	Lys	Tyr	Gly	Phe	Asp	Glu	Ile	Arg	Ile
	770					775					780				
Glu	Leu	Ser	Arg	Glu	Leu	Arg	Ala	Ser	Ala	Asp	Glu	Arg	Gln	Gln	Met
785					790					795					800
Trp	Glu	Gly	Met	Ser	Asn	Gly	Ala	Ala	Lys	Asn	Glu	Trp	Ala	Lys	Lys
				805					810					815	
Met	Leu	Arg	Glu	Ile	Lys	Gln	Ala	Met	Leu	Asp	Asn	Asn	Arg	Asp	Ile
			820					825					830		
Ser	Tyr	Leu	Asp	Thr	Glu	Ile	Ser	Asn	Ser	Asn	Leu	Asp	Lys	Ile	Lys
		835					840					845			

Ile Ile Glu Asp Val Val Lys Tyr Gln Lys Pro Val Glu Phe Lys Thr
850 855 860

Lys Glu Lys Glu Tyr Lys Leu Thr Glu Pro Thr Lys Ala Glu Val Thr
865 870 875 880

Lys Tyr Leu Leu Trp Leu Glu Gln Asn Phe Lys Cys Pro Tyr Thr Asn
885 890 895

Gln Pro Ile Pro Leu Thr Asp Val Phe Ala Arg Gly Lys Val Val Glu
900 905 910

Ile Glu His Ile Ile Pro Arg Glu Arg Tyr Tyr Ser Asn Ala Tyr Ala
915 920 925

Asn Lys Val Ile Thr Trp Arg Glu Val Asn Gln Glu Lys Gly Asn Lys
930 935 940

Thr Ala Tyr Glu Phe Ile Val Ser Lys Arg Asn Pro Ala Gly Thr Val
945 950 955 960

Gln Val Gly Lys Lys Glu Ile Ser Leu Val Ser Ala Ser Gln Trp Glu
965 970 975

Asn His Val Lys Glu Met Phe Pro Lys Ser Gly Lys Arg Thr Asn Leu
980 985 990

Leu Arg Lys Glu Ile Pro Glu Asp Pro Ile Glu Arg Thr Leu Lys Glu
995 1000 1005

Asn Gln Tyr Ile Asn Lys Lys Leu Lys Glu Lys Leu Gly Glu Leu
1010 1015 1020

Val Gly Ser Asp Lys Val Trp Val Thr Ser Gly Ala Ile Thr Asp
1025 1030 1035

Met Leu Arg Glu Lys Trp Lys Leu Asn Asp Ile Trp Lys Glu Leu
1040 1045 1050

Val Gln Glu Arg Phe Leu Lys Phe Asn Asn Gly Glu Val Lys Asp

1055	1060	1065
Arg Leu Asn Leu Lys Tyr Thr 1070	Thr Lys Tyr Tyr Thr 1075	Asn Asn Lys Thr 1080
Asn Gln Tyr Glu Asp Val 1085	Glu Lys Phe Ala Gly 1090	Met Ser Lys Arg 1095
Ile Asp His Arg His His 1100	Ala Leu Asp Ala Ile 1105	Ile Ile Ala Cys 1110
Thr Lys Gln Ser His Ile 1115	Gln Tyr Ile Asn Asn 1120	Leu Asn Ala Ile 1125
Asn Ser Ala Asp Ile Glu 1130	Asn Asp Asp Asn Lys 1135	Lys Thr Lys Tyr 1140
Lys Gly Ile Lys Thr Asp 1145	Val Cys Ile Gly Asn 1150	Ser Ser Ser Lys 1155
Phe Lys Phe Pro Trp Asp 1160	Glu Asn Lys Phe Ile 1165	Pro Glu Val Lys 1170
Gln Ala Leu Gln Asn Ile 1175	Leu Ile Ser His Lys 1180	Asn Ser Asn Val 1185
Leu Ile Ser Pro Ser Lys 1190	His Arg Asn Asn Lys 1195	Asp Ile Asn Ser 1200
Gly Lys Val Ala Ser Val 1205	Arg Gly Glu Leu His 1210	Leu Gln Thr Asn 1215
Tyr Ala Lys Lys Lys Tyr 1220	Phe Glu Asn Glu Lys 1225	Ile Asp Ile Val 1230
Lys Leu Ile Pro Leu Leu 1235	Phe Lys Lys Lys Phe 1240	Glu Asn Gln Asn 1245
Gln Thr Leu Val Arg Phe 1250	Lys Ser Phe Asp Glu 1255	Ile Ile Lys Glu 1260

Ala Val Leu Lys Glu Lys Tyr Gln Asn Val Leu Ile Ser Leu Phe
1265 1270 1275

Lys Asp Tyr Asp Ser Glu Lys Leu Ser Thr Gln Thr Thr Lys Asp
1280 1285 1290

Phe Ser Lys Lys Ile Leu Lys Arg Ile Glu Asp Glu Lys Leu Phe
1295 1300 1305

Ile His Glu Lys Ser Gly Glu Lys Met Thr Trp Leu Ser Thr Phe
1310 1315 1320

Thr Glu Lys Asn Ala Ser Val Arg Pro Met Gly Leu Ala Met Asp
1325 1330 1335

Leu Asn Ser Glu Lys Glu Ile Lys Asp Ile Ser Asn Pro Arg Leu
1340 1345 1350

Arg Arg Ile Gly Gln Tyr Arg Leu Asn Tyr Ile Asn Lys Lys Ile
1355 1360 1365

Ile Glu Ile Lys Ala Leu Lys Leu Pro Phe Lys Glu Glu Glu Asn
1370 1375 1380

Asn Arg Ile Arg Glu Val Arg Glu Ile Pro Leu Tyr Asn Asn Ala
1385 1390 1395

Ile Tyr Glu Leu Arg Ile Lys Gln Lys Asp Gly Lys Tyr His Trp
1400 1405 1410

Val Glu Ile Lys Asp Val Gln His Ser Asp Leu Glu Asn Ile Glu
1415 1420 1425

Tyr Ala Lys Gln Glu Thr Thr Ile Ala Val Lys Asn Lys Ile Lys
1430 1435 1440

Asp Phe Ser Leu Asp Glu Leu Lys Glu Arg Tyr Ile Ser Asn Pro
1445 1450 1455

Ile Phe Leu Ser Glu Lys Pro Thr Pro Val Lys Lys Ala Arg Gln

1460

1465

1470

Lys Ser Phe Phe Gln Asp Leu Tyr Gln Leu Asn Asp Gly Arg Tyr
 1475 1480 1485

Val Tyr Ser Arg Asp Val Phe Met Thr Tyr Val Phe Val Pro Glu
 1490 1495 1500

Asp Ser Asn Arg Lys Asp Ala Lys Arg Glu Leu Glu Phe Leu Lys
 1505 1510 1515

Phe Phe Asp Ala Val Lys Ile Val Thr Leu Glu Lys Pro Glu Lys
 1520 1525 1530

Ile Asp Tyr Gln Lys Leu Ile Lys Lys Glu Asp His Lys Leu Leu
 1535 1540 1545

Phe Thr Leu Ser Lys Asn Asp Ile Val Tyr Ile Pro Glu Glu Ile
 1550 1555 1560

Leu Thr Asp Glu Gln Ile Ala Glu Ile Asn Trp Asn Asp Lys Arg
 1565 1570 1575

Asp Ile Val Pro Lys Leu Phe Val Val Lys Glu Met Thr Pro Ser
 1580 1585 1590

Arg Asn Glu Ile Val Phe Gln His Leu Tyr Lys Ser Asp Ser Ile
 1595 1600 1605

Arg Val Asn Gly Asp Glu Ala Lys Asn Ile Leu Asn Thr Glu Lys
 1610 1615 1620

Leu Glu Glu Gln Ile Lys Tyr Gly Asp Thr Asn Met Trp Lys Arg
 1625 1630 1635

Cys Ile Lys Val Asp Ile Asp Lys Leu Gly Thr Glu Ile Lys Pro
 1640 1645 1650

Tyr Trp Glu Asn Asn Lys Ile Val Leu
 1655 1660

<210> 128
<211> 1075
<212> PRT
<213> Bacillus sp.

<223> APG02874_RGN

<400> 128

Met Asp Tyr Arg Ile Gly Leu Asp Ile Gly Thr Asn Ser Val Gly Trp
1 5 10 15

Gly Val Ile Arg Met Gln Tyr Asn Asp Leu Leu Glu Arg Tyr Glu Lys
20 25 30

Gln Gly Ile Val Asp Ala Gly Val Arg Met Phe Asp Lys Ala Glu Ile
35 40 45

Pro Lys Thr Gly Ala Ser Leu Ala Ala Pro Arg Arg Glu Ala Arg Ser
50 55 60

Ser Arg Lys Arg Leu Arg Arg Arg Ser Gln Arg Lys Lys Ala Ile Arg
65 70 75 80

Tyr Leu Ile Ile Leu His Glu Leu Leu Ser Glu Asp Gln Leu Gln Gln
85 90 95

Leu Tyr Pro Leu Gln Lys Glu Ser Ile Asp Ile Trp Asp Ile Arg Leu
100 105 110

Glu Ser Leu Asp Arg Val Leu Asn Gln Thr Glu Trp Ala Arg Leu Leu
115 120 125

Ile His Leu Gly Gln Arg Arg Gly Phe Lys Ser Asn Arg Lys Ser Glu
130 135 140

Tyr Lys Glu Glu Glu Thr Gly Lys Val Leu Val Asn Ile Gln Glu Asn
145 150 155 160

Gln Lys Ala Leu Gln Gln Tyr Arg Thr Val Ala Glu Met Trp Met Lys
165 170 175

Asp Ala Lys Phe Ser Gln Phe Asp Lys Lys Arg Asn Ala Pro Gly Glu
180 185 190

Tyr Leu Phe Ser Val Ser Arg Tyr Asp Leu Gln Asn Glu Ile Gln Thr
195 200 205

Ile Phe Glu Ser Gln Arg Asn Leu Gly Ser Pro Tyr Ala Thr Thr Asp
210 215 220

Leu Glu Glu Glu Tyr Ile Ser Val Trp Ser Lys Gln Leu Pro Phe Ala
225 230 235 240

Ser Gly Glu Asp Ile Leu Lys Lys Val Gly Tyr Cys Ser Leu Glu Arg
245 250 255

Lys Glu Arg Arg Ile Pro Lys Ala Thr Tyr Thr Phe Gln Tyr Phe Ile
260 265 270

Leu Leu Asp Gly Leu Asn Lys Ile Arg Val Gly Ala Asp Ala Arg Gln
275 280 285

Leu Thr Thr Glu Glu Arg Ser Ile Leu Ile Glu Lys Ile Leu Thr Arg
290 295 300

Asn Asp Tyr Phe Ser Lys Lys Ser Leu Pro Asp Leu Thr Tyr Gly Asp
305 310 315 320

Leu Arg Lys Trp Leu Ala Leu Asp Ser Ser Leu Leu Phe Lys Asp Leu
325 330 335

Thr Tyr Asp Pro Asn Glu Lys Val Ser Lys Val Glu Lys Lys Val Phe
340 345 350

Ala Asn Leu Lys Ser Leu Tyr Glu Leu Arg Lys Val Val Glu Thr Ser
355 360 365

Ser Asn Gly Gln Ser Phe Glu Arg Ile Asp Tyr Asp Thr Phe Gly Tyr
370 375 380

Ala Leu Thr Val Tyr Lys Thr Asp Glu Asp Ile Arg Arg Tyr Leu Lys
385 390 395 400

Lys Pro Ser Asn Leu Ala His Lys Val Tyr Glu Glu Asp Leu Ile Glu
405 410 415

Arg Leu Leu Ser Leu Ser Tyr Glu Lys Phe Gly His Leu Ser Leu Lys
420 425 430

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

Glu Ala Val Gln Ala Ala Gly Tyr Asp Ile Thr Gly Leu Gln Ser Ala
450 455 460

Lys Arg Thr Leu Leu Leu Pro Thr Ile Pro Ala Asp Ile Thr Asn Pro
465 470 475 480

Val Val Arg Arg Ala Leu Thr Gln Thr Arg Lys Val Val Asn Ala Ile
485 490 495

Ile Lys Arg Tyr Gly Ser Pro Leu Thr Ile His Ile Glu Leu Ala Arg
500 505 510

Glu Leu Ser Lys Asp His Gln Glu Arg Gln Gln Ile Leu Lys Asn Gln
515 520 525

Gln Ala Asn Tyr Glu Arg Asn Lys Gly Ala Ile Ser Val Leu Val Glu
530 535 540

Asn Gly Ile Leu Asn Pro Thr Gly Tyr Asp Ile Val Arg Tyr Lys Leu
545 550 555 560

Trp Lys Glu Gln Gly Glu Arg Cys Ala Tyr Ser Leu Lys Val Ile Pro
565 570 575

Pro Ser Val Phe Phe Gly Glu Leu Lys Arg Glu Arg Asn His Ile Pro
580 585 590

Thr Leu Glu Val Asp His Ile Leu Pro Tyr Ser Gln Ser Phe Met Asp
595 600 605

Asn Tyr His Asn Lys Val Leu Val Tyr Ser Asp Glu Asn Gln Lys Lys
610 615 620

Gly Asn Arg Ile Pro Ser Gln Tyr Leu Met Ala Val Ser Pro Glu Lys
625 630 635 640

Trp Thr Lys Phe Glu Glu Tyr Val Ser Ser Asn Lys Gly Phe Ser Lys
645 650 655

Gln Lys Arg Gln Tyr Leu Leu Lys Lys Asp Tyr Ser Ser Arg Glu Arg
660 665 670

Asp Ile Val Lys Glu Arg His Leu Val Asp Thr Arg Tyr Ala Thr Arg
675 680 685

Phe Met Lys Asn Phe Ile Glu Lys Thr Leu Leu Phe Lys Glu Ser Arg
690 695 700

Ser Ser Ser Leu Thr Lys Arg Val Gln Thr Val Asn Gly Lys Ile Thr
705 710 715 720

Ala His Phe Arg Lys Arg Trp Gly Leu Glu Lys Glu Arg Thr Glu Thr
725 730 735

Tyr Leu His His Ala Leu Asp Ala Ile Val Val Ala Cys Thr Asp Gln
740 745 750

His Met Val Thr Arg Val Thr Glu Tyr Tyr Gln Glu Arg Glu Arg Asn
755 760 765

Val Arg Gln Pro Tyr Phe Pro Trp Pro Trp Glu Gly Phe Arg Asp Glu
770 775 780

Leu Leu Asn Tyr Leu Asn Arg Gln Pro Asp Ser Leu Glu Ile Ser Gln
785 790 795 800

Ser Ile Lys Gln Asn Leu Phe Leu Arg Asp Tyr Leu Met Val Ser Arg
805 810 815

Met Pro Lys His Ser Val Thr Gly Met Ser His Lys Gln Thr Ile Arg
820 825 830

Ser Lys Ala Cys Val Asp Glu Lys Gly Lys Val Lys Thr Thr Lys Arg
835 840 845

Val Leu Leu Gln Asp Ile Lys Phe Asp Ser Asn Gly Asp Phe Lys Met
850 855 860

Ile Gly Lys Glu Gln Asp Met Ala Thr Tyr Asn Ala Ile Lys Glu Arg
865 870 875 880

Phe Leu Ser Ser Gly Gly Asn Val Lys Ser Ala Phe Thr Glu Pro Leu
885 890 895

His Lys Pro Ser Lys Lys Gly Lys Gly Asn Val Ile Arg Lys Val Thr
900 905 910

Val Glu Asp Gln Thr Lys Ser Phe Val Arg Asp Val Asn Gly Gly Val
915 920 925

Ala Glu Asn Gly Ser Leu Val Arg Ile Asp Leu Phe Glu Lys Glu Gly
930 935 940

Met Tyr Tyr Val Val Pro Ile Tyr Val Gly Asp Thr Val Leu Lys Gln
945 950 955 960

Leu Pro Ile Lys Ile Ser Ala Ser Asn Arg Gly Tyr Tyr Lys Trp Val
965 970 975

Glu Leu Asp Glu Asn Tyr Arg Phe Lys Phe Ser Leu Tyr Ala His Asp
980 985 990

Leu Ile Arg Ile Val Thr Lys Asp Glu Asp Arg Phe Phe Tyr Phe Ser
995 1000 1005

Tyr Phe Asp Asn Asp Gly Asn Arg Ile Lys Phe Leu Asn Ile Asn
1010 1015 1020

Ser Pro Thr Gly Lys Asn Glu Asn Arg Tyr Gly Val Arg Ser Ile
1025 1030 1035

Lys Leu Ile Glu Lys Tyr Asn Val Ser Thr Leu Gly Asp Ile Ser
1040 1045 1050

Leu Val Lys Ser Glu Glu Arg Gln Leu Phe Gln Gly Met Lys Lys
1055 1060 1065

Lys Val Leu Asn Ile Ile Asn
1070 1075

<210> 129
<211> 1068
<212> PRT
<213> Bacillus sp.

<223> APG03850_RGN

<400> 129

Met Lys Tyr Val Leu Gly Leu Asp Ile Gly Ile Ala Ser Cys Gly Trp
1 5 10 15

Ala Val Ile Asn Gln Glu Lys His Arg Ile Glu Asp Leu Gly Val Arg
20 25 30

Ile Phe Asp Lys Ala Glu Asn Pro Lys Asp Gly Lys Ser Leu Ala Thr
35 40 45

Pro Arg Arg Asp Ala Arg Ser Thr Arg Arg Thr Leu Arg Arg Lys Lys
50 55 60

His Arg Met Gln Arg Ile Lys Ile Leu Leu Val Lys His Gly Leu Leu
65 70 75 80

Ser Lys Thr Glu Ile Asp His Leu Tyr Glu Ser Ala Thr Glu Ile Asp
85 90 95

Val Trp Tyr Leu Arg Leu Asn Ala Leu Glu Arg Arg Leu Asn Pro Lys
100 105 110

Glu Phe Ala Arg Val Leu Ile His Leu Ala Lys Arg Arg Gly Phe Lys
115 120 125

Ser Asn Arg Lys Glu Thr Thr Leu Ser Glu Asn Gly Gln Ile Leu Glu

130

135

140

Asn Ile Ser Glu Asn Leu Gln Ile Met Glu Gln Gln Asn Tyr Arg Thr
 145 150 155 160

Val Gly Glu Met Ile Leu Lys Asp Lys Lys Phe Glu Asn His Lys Arg
 165 170 175

Asn Lys Asp Gly Thr Tyr Ile Gly Thr Val Thr Arg Gln Gln Leu Lys
 180 185 190

Glu Glu Ile Gln Met Ile Phe Asn Ala Gln Arg Leu Tyr Lys Asn Asp
 195 200 205

Tyr Ala Thr Glu Glu Phe Glu Ser Ser Tyr Leu Glu Ile Trp Ala Ser
 210 215 220

Gln Arg Pro Tyr Ala Ser Lys Asp Gln Ile Glu Lys Met Ile Gly Tyr
 225 230 235 240

Cys Thr Leu Glu Pro Lys Glu Lys Arg Val Pro Lys Ala Ser Trp Ser
 245 250 255

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

Asn Lys Asp Arg Ile Glu Glu Leu Ser Phe Glu Glu Lys Asn Gln Ile
 275 280 285

Met Asn Leu Ala Leu Glu Lys Ser Ile Val Lys Tyr Ile Asp Ile Arg
 290 295 300

Lys Leu Leu Ser Ile Pro Asn Glu Phe His Phe Asn Asn Leu Leu Tyr
 305 310 315 320

Ser Ala Asp Thr Val Asp Thr Ala Val Glu Asn Lys Lys Cys Ile Glu
 325 330 335

Phe Lys Glu Tyr His Ser Ile Asn Lys Leu Tyr Lys Gln Ile Tyr Gly
 340 345 350

Lys Ser Val Pro Asn Leu Leu Pro Ile Asp Tyr Asp Thr Ile Ala Cys
355 360 365

Gly Leu Thr Ile Phe Lys Asp Asp Lys Asp Ile Leu Ala Tyr Leu Gln
370 375 380

Asn Lys Tyr Val Asn Ala Lys Gly Lys Pro Ile Ser Asn Leu Ala Lys
385 390 395 400

Lys Thr Tyr Asp Asp Thr Phe Ile Gln Ala Leu Leu Thr Leu Asn Phe
405 410 415

Ser Lys Met Gly His Leu Ser Phe Lys Ala Leu Lys Asn Ile Ile Pro
420 425 430

Phe Leu Glu Glu Gly Leu Ser Tyr Asp Lys Ala Cys Glu Lys Ala Gly
435 440 445

Tyr Asn Phe Lys Gly Thr Ser His Ala Glu Lys Thr Lys Tyr Leu Pro
450 455 460

Val Ile Pro Gln Asn Thr Asn Pro Val Val His Arg Ala Leu Ser Gln
465 470 475 480

Thr Arg Lys Val Ile Asn Ala Ile Ile Lys Lys Tyr Gly Ser Pro Ser
485 490 495

Ala Ile His Ile Glu Thr Ala Arg Glu Leu Ser Lys Thr Phe Gln Glu
500 505 510

Arg Lys Glu Ile Asp Ser Met Tyr Gln Asp Asn Ser Lys Lys Asn Glu
515 520 525

His Ala Ile His Lys Leu Lys Glu Leu Gly Leu Ile Asn Pro Ser Gly
530 535 540

Ile Asn Ile Val Lys Phe Lys Leu Trp Asn Glu Gln Asp Gly Arg Cys
545 550 555 560

Met Tyr Ser Gly Lys Tyr Ile Glu Pro His Arg Leu Phe Glu Glu Gly

565

570

575

Tyr Thr Glu Val Asp His Ile Leu Pro Tyr Ser Arg Ser Leu Asp Asp
 580 585 590

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

Gly Asn Lys Thr Pro Tyr Glu Tyr Ile Gly Lys Thr Ser Ile Trp His
 610 615 620

Glu Phe Glu Thr Arg Val Gln Ser Asn Lys Arg Ile Asn Lys Lys Lys
 625 630 635 640

Gln Gln Lys Leu Leu Leu Gln Tyr Phe Ser Tyr Thr Arg Glu Gln Glu
 645 650 655

Phe Ile Lys Arg Asn Leu Asn Asp Thr Arg Tyr Ala Thr Ile Tyr Leu
 660 665 670

Ser Thr Leu Ile Gln Gln His Leu Ile Phe Ser Glu Ser Ser Arg Lys
 675 680 685

Lys Lys Val His Thr Val Ser Gly Ile Ile Thr Ser His Leu Arg Ser
 690 695 700

Arg Trp Gly Phe Asn Lys Asp Arg Lys Glu Gly His Ile His His Ala
 705 710 715 720

Leu Asp Ala Val Ile Val Ala Val Thr Ser Asp His Met Ile Gln Arg
 725 730 735

Val Thr Lys Tyr Tyr Lys Leu Lys Glu Leu Asn Arg Asn Leu Gln Ala
 740 745 750

Lys Arg Met Gln Phe Pro Glu Pro Trp Glu Gly Phe Arg Leu Glu Leu
 755 760 765

Glu Ala Arg Ile Ser Pro Asn Thr Gln Gln Tyr Leu Lys Arg Ile Leu
 770 775 780

Phe Lys Asn Tyr Ala Asp Val Asn Leu Ser Glu Ile Lys Pro Ile Phe
785 790 795 800

Val Ser Arg Met Pro Lys Arg Ser Ile Thr Gly Glu Leu His Gln Glu
805 810 815

Thr Ile Arg Lys Leu Ile Gly Tyr Asn Glu Lys Gly Lys Val Leu Thr
820 825 830

Ala Ile Lys Thr Lys Leu Glu Asp Ile Pro Phe Asp Ala Asn Gly Asp
835 840 845

Phe Pro Met Tyr Gly Lys Glu Thr Asp Leu Tyr Thr Tyr Asn Ala Ile
850 855 860

Lys Glu Arg Tyr Leu Ser His Lys Lys Asp Lys Arg Lys Ser Phe Gln
865 870 875 880

Asp Pro Leu Tyr Lys Pro Thr Lys Ser Gly Glu Ile Gly Pro Leu Ile
885 890 895

Lys Ser Ile Lys Ile Met Asp Thr Arg Thr Ile Val Asn Pro Val Asn
900 905 910

Gln Gly Lys Gly Val Val Tyr Asn Ser Lys Ile Ala Arg Thr Asp Val
915 920 925

Phe Lys Lys Asp Glu Lys Tyr Tyr Leu Ile Pro Ile Tyr Thr Ile Asp
930 935 940

Leu Leu Lys Asn Ile Leu Pro Gln Lys Ala Ile Thr Ala Gly Lys Gly
945 950 955 960

Tyr Glu Asp Trp Thr Thr Ile Asp Pro Ser Phe Thr Phe Leu Phe Ser
965 970 975

Leu Phe Pro Asn Asp Leu Ile Gln Ile Val Pro Ser Lys Asn Lys Thr
980 985 990

Ile Lys Ala Arg Thr Thr Val Ser Lys Lys Glu Val Leu Leu Pro Ser

995

1000

1005

Leu Thr Gly Tyr Phe Lys Gly Val His Ser Gly Thr Ala Gly Ile
 1010 1015 1020

Thr Val Glu Thr His Asp Gly Ser Val Ile Ala Asn Val Gly Ser
 1025 1030 1035

Lys Gln Leu Leu Leu Phe Glu Lys Tyr Gln Val Asp Val Leu Gly
 1040 1045 1050

His Tyr Thr Lys Ile Lys Glu Glu Lys Arg Ile Gly Met Val Ile
 1055 1060 1065

<210> 130

<211> 1081

<212> PRT

<213> Bacillus sp.

<223> APG07553_RGN

<400> 130

Met Gln Tyr Val Leu Gly Leu Asp Ile Gly Ile Ala Ser Cys Gly Trp
 1 5 10 15

Ala Val Ile Asn Gln Glu Lys Glu Arg Ile Glu Asp Leu Gly Val Arg
 20 25 30

Ile Phe Asp Lys Ala Glu Asn Pro Lys Asp Gly Lys Ser Leu Ala Ala
 35 40 45

Pro Arg Arg Asp Ala Arg Ser Thr Arg Arg Thr Leu Arg Arg Lys Lys
 50 55 60

His Arg Met Gln Arg Ile Lys Ile Leu Leu Val Lys His Gly Leu Leu
 65 70 75 80

Ser Gln Thr Glu Leu Asp His Leu Tyr Glu Ser Ala Thr Glu Val Asp
 85 90 95

Val Trp Asn Leu Arg Leu Asp Ala Ile Glu Arg Lys Leu Asn Pro Lys
 100 105 110

Glu Phe Thr Arg Val Leu Ile His Leu Ala Lys Arg Arg Gly Phe Lys
115 120 125

Ser Asn Ser Lys Glu Thr Thr Leu Ser Glu Asn Gly Gln Ile Leu Glu
130 135 140

Ser Ile Ser Glu Asn Gln Gln Ile Met Glu Gln Lys Asn Tyr Lys Thr
145 150 155 160

Ile Gly Glu Met Ile Leu Lys Asp Lys Lys Phe Glu Asn His Lys Arg
165 170 175

Asn Lys Asp Gly Thr Tyr Ile Gly Thr Val Thr Arg Gln Gln Leu Gln
180 185 190

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

Tyr Ala Thr Lys Glu Phe Glu Ser Ser Tyr Leu Glu Ile Trp Ala Ser
210 215 220

Gln Arg Pro Tyr Ala Ser Lys Asp Gln Ile Glu Lys Met Ile Gly Tyr
225 230 235 240

Cys Thr Leu Glu Gln Lys Glu Lys Arg Val Pro Arg Ala Ser Trp Ser
245 250 255

Phe Gln Tyr Phe Val Ala Leu His Thr Ile Asn Asn Leu Arg Leu Ile
260 265 270

Ser Lys Asp Arg Ile Glu Glu Leu Ser Phe Lys Glu Lys Lys Gln Ile
275 280 285

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

Lys Leu Leu Ser Ile Pro Asn Glu Leu His Phe Asn Ser Leu Leu Tyr
305 310 315 320

Ser Ala Asp Thr Val Asp Thr Thr Val Glu Asn Arg Lys Cys Ile Glu
325 330 335

Leu Lys Glu Tyr His Ser Ile Asn Lys Val Tyr Lys Gln Ile Tyr Gly
340 345 350

Lys Asn Ala Leu Asn Leu Leu Pro Ile Asp Tyr Asp Thr Ile Ala Tyr
355 360 365

Gly Leu Thr Ile Phe Lys Asp Asp Lys Asp Ile Leu Glu His Leu Lys
370 375 380

Asn Lys Tyr Val Asn Ala Lys Gly Lys Pro Ile Asn Asn Leu Ala Lys
385 390 395 400

Lys Thr Tyr Asp Asp Thr Phe Ile Gln Ala Leu Leu Thr Leu Asn Phe
405 410 415

Ser Lys Met Gly His Leu Ser Phe Lys Ala Leu Lys Asn Ile Ile Pro
420 425 430

Phe Leu Glu Glu Gly Leu Ser Tyr Asp Lys Ala Cys Glu Lys Ala Gly
435 440 445

Tyr Asn Phe Lys Gly Thr Ser Tyr Thr Glu Gln Thr Lys Tyr Leu Pro
450 455 460

Val Ile Pro Gln Asn Thr Asn Pro Val Val His Arg Ala Leu Ser Gln
465 470 475 480

Thr Lys Lys Val Ile Asn Ala Ile Ile Lys Lys Tyr Gly Ser Pro Asn
485 490 495

Ala Ile His Ile Glu Thr Ala Arg Glu Leu Ser Lys Thr Phe Gln Glu
500 505 510

Arg Lys Glu Ile Asp Ser Met Tyr Gln Asp Asn Ser Lys Lys Asn Glu
515 520 525

His Ala Ile His Lys Leu Lys Glu Leu Gly Asn Ile Asn Pro Ser Gly
530 535 540

Ile Asn Ile Val Lys Phe Lys Leu Trp Asn Glu Gln Asp Gly Lys Cys
545 550 555 560

Met Tyr Ser Gly Lys Tyr Ile Glu Pro His Arg Leu Phe Glu Glu Gly
565 570 575

Tyr Thr Glu Val Asp His Ile Leu Pro Tyr Ser Arg Ser Leu Asp Asp
580 585 590

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

Gly Asn Lys Thr Pro Tyr Glu Tyr Met Gly Asn Thr Ser Ile Trp Asp
610 615 620

Glu Tyr Glu Ile Arg Val Gln Ser Asn Lys Lys Ile Asn Lys Lys Lys
625 630 635 640

Gln Gln Lys Leu Leu Leu Gln His Phe Ser Tyr Ala Arg Glu Gln Glu
645 650 655

Phe Ile Glu Arg Asn Leu Asn Asp Thr Arg Tyr Ala Thr Ile Tyr Leu
660 665 670

Ser Ser Leu Ile Gln Gln His Leu Ile Phe Ser Glu Ser Ser Arg Lys
675 680 685

Lys Lys Val His Thr Val Ser Gly Ile Ile Thr Ser His Leu Arg Ser
690 695 700

Arg Trp Gly Phe Asn Lys Asp Arg Lys Glu Gly His Ile His His Ala
705 710 715 720

Leu Asp Ala Val Ile Val Ala Val Thr Ser Asp His Met Ile Gln Arg
725 730 735

Val Thr Lys Tyr Tyr Lys Leu Lys Glu Leu Asn Arg Asn Leu Gln Ala
740 745 750

Lys Arg Met Pro Phe Pro Glu Pro Trp Glu Gly Phe Arg Leu Glu Leu
755 760 765

Glu Ala Arg Ile Ser Pro Asn Thr Gln Gln Tyr Leu Lys Gly Leu Arg
770 775 780

Phe Lys Asn Tyr Ala Asp Val Asn Leu Cys Glu Ile Lys Pro Ile Phe
785 790 795 800

Val Ser Lys Met Pro Lys Arg Ser Ile Thr Gly Glu Leu His Gln Glu
805 810 815

Thr Ile Arg Lys Phe Ile Gly Tyr Asn Glu Lys Gly Lys Val Leu Thr
820 825 830

Ala Ile Lys Thr Lys Leu Glu Asp Ile Pro Phe Asp Ala Asn Gly Asp
835 840 845

Phe Pro Met Tyr Gly Lys Glu Thr Asp Leu Tyr Thr Tyr Asn Ser Ile
850 855 860

Lys Glu Arg Tyr Leu Ser His Lys Lys Asp Lys Arg Lys Ser Phe Gln
865 870 875 880

Glu Pro Leu Tyr Lys Pro Thr Lys Ser Gly Gly Ile Gly Pro Leu Ile
885 890 895

Lys Ser Ile Lys Ile Met Asp Thr Arg Thr Ile Val Asn Pro Val Asn
900 905 910

Gln Gly Lys Gly Val Val Tyr Asn Ser Lys Ile Ala Arg Thr Asp Val
915 920 925

Phe Lys Lys Asp Asp Lys Tyr Tyr Leu Ile Pro Ile Tyr Thr Ile Asp
930 935 940

Leu Met Lys Asn Ile Leu Pro Gln Lys Ala Ile Thr Ala Gly Lys Gly
945 950 955 960

Tyr Glu Asp Trp Ile Thr Ile Asp His Ser Phe Thr Phe Leu Phe Ser
965 970 975

Leu Phe Pro Asn Asp Leu Ile Lys Ile Val Pro Ser Lys Asn Lys Glu
980 985 990

Ile Lys Ala Arg Ser Thr Ser Ser Lys Lys Glu Ile Leu Leu Pro Ser
995 1000 1005

Leu Ile Gly Tyr Phe Lys Ser Val His Ser Gly Thr Ala Gly Ile
1010 1015 1020

Thr Val Glu Ser His Asp Gly Arg Phe Ile Ala Asn Val Gly Ser
1025 1030 1035

Lys Gln Leu Leu Leu Phe Glu Lys Tyr Gln Val Asp Val Leu Gly
1040 1045 1050

His Tyr Thr Lys Ile Lys Glu Glu Lys Arg Ile Gly Met Ala Thr
1055 1060 1065

Cys Asn Asp Asn Lys Lys Ser Thr Ala Phe Gly Ser Leu
1070 1075 1080

<210> 131
<211> 1404
<212> PRT
<213> Chryseobacterium sp.

<223> APG03031_RGN

<400> 131

Met Lys Lys Thr Leu Gly Ile Asp Leu Gly Thr Asn Ser Ile Gly Trp
1 5 10 15

Val Ile Arg Asn Ala Asp Leu Leu Asp Asn Gln Ile Glu Asp Phe Gly
20 25 30

Val Ile Thr Phe Glu Lys Gly Val Ala Ser Glu Lys Gly Asn Glu Phe
35 40 45

Pro Lys Val Gln Lys Arg Thr Glu Ser Arg Gly Lys Arg Arg Asn Tyr
50 55 60

Gln Ser Glu Lys Tyr Arg Lys Phe Ala Leu Leu Ala Phe Leu Ile Gln
65 70 75 80

Lys Asn Met Cys Pro Leu Thr Met Asp Glu Phe Cys Asn Trp Lys Asp
85 90 95

Tyr Arg Lys Gly Lys Lys Arg Glu Tyr Pro Gln Ser Glu Glu Phe Leu
100 105 110

Asn Trp Ile Arg Phe Asp Phe Asn Gly Asp Gly Leu Pro Asp Phe His
115 120 125

Leu Phe Gly Lys Glu Lys Asp Asp Ser Tyr Tyr Val Phe Arg Ala Phe
130 135 140

Ala Ala Asp Glu Asn His Lys Asn Val Phe Asp Asn Asn Pro His Ile
145 150 155 160

Leu Gly Arg Val Leu Tyr Gln Met Val Gln Arg Arg Gly Phe Lys Gly
165 170 175

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

Gly Thr Lys Gly Arg Asn Asp Ile Glu Glu Tyr Ile Glu Lys Tyr Lys
195 200 205

Thr Leu Gly Ala Ala Leu Tyr Tyr His Gln Lys Asp Phe Gly Gly Arg
210 215 220

Ile Lys Gln Arg Tyr Asn Leu Arg Lys Asp Asn Glu Asn Glu Leu Lys
225 230 235 240

Glu Ile Cys Lys Ile His Gln Leu Ser Gln Glu Asp Tyr Glu Lys Leu
245 250 255

Trp Lys Ala Ile Ile Trp Gln Arg Pro Leu Arg Thr Gln Lys Gly Leu
260 265 270

Val Gly Asn Cys Ile Tyr Glu Lys Asn Lys Lys Arg Val Ala Val Ser

275

280

285

His Pro Leu Tyr Glu Glu Tyr Arg Thr Trp Val Phe Ile Asn Asn Leu
 290 295 300

Asn Ile Ile Pro Pro Phe Gly Glu Glu Lys Gln Asn Tyr Ile Ser Glu
 305 310 315 320

Lys Ile Tyr Pro Leu Phe Tyr Lys Ser Lys Pro Asp Phe Glu Leu Ser
 325 330 335

Asp Ile Asp Lys Gln Leu Lys Lys Asp Gly Ala Lys Arg Thr Ser Lys
 340 345 350

His Asn Asp Lys Thr Lys Val Leu Ser Val Lys Leu Leu Lys Gln Phe
 355 360 365

Gln Asp Val Phe Gly Glu Asp Trp Lys Gln Lys Leu Gln Trp Asp Ile
 370 375 380

Asn Gly Asp Arg Asn Ser Gln Pro Glu Lys Lys Glu Lys Ser Leu Tyr
 385 390 395 400

Thr Phe Glu Asp Leu Trp His Ile Leu Leu Thr Phe Asp Gly Gln Glu
 405 410 415

Asn Leu Lys Lys Phe Ala Leu Glu Lys Leu Asn Leu Asp Asp Glu Lys
 420 425 430

Ala Ile Lys Phe Ser Lys Ile Lys Leu Gln Gln Gly Tyr Ala Thr Leu
 435 440 445

Ser Ile Ser Ala Ile Lys Lys Ile Leu Pro Tyr Leu Gln Arg Gly Phe
 450 455 460

Leu Tyr Ser His Ala Val Tyr Met Ala Asn Leu Tyr Lys Met Leu Gly
 465 470 475 480

Glu Thr Thr Ile Ser Glu Thr Leu Thr Asn His Phe Ser Asp Glu Ile
 485 490 495

Leu Lys Ile Leu Glu Asn Asp Asp Phe Asn Arg Lys Asn Ser Ala Ile
500 505 510

Ile Asn Ser Leu Ile Ser Glu Met Leu Asn Asp Asp Asn Arg Tyr Tyr
515 520 525

Ile Ala Asp Asn Arg Asn Leu Asp Lys Thr Glu Glu Gln Ile Val Leu
530 535 540

Lys Ser Ile Thr Asp Asn Tyr Gly Leu Glu Thr Trp Gly Asn Phe Asp
545 550 555 560

Glu Glu Lys Lys Met Glu Ile Ile Asp Phe Val Ser Ser Lys Tyr Leu
565 570 575

Glu Phe Leu Lys Lys Thr Tyr Thr Glu Lys Gly Asn Leu Phe Ile Gln
580 585 590

Pro Glu Arg Ile His Asp Lys Ile Phe Asn Phe Leu Lys Glu Thr Tyr
595 600 605

His Val Ser Glu Asp Lys Ile Lys Tyr Leu Trp His Pro Ser Glu Gln
610 615 620

Glu Asn Tyr Ala Lys Ala Asp Glu Tyr Phe Glu Tyr Ser Ile Gly Ser
625 630 635 640

Lys Lys Tyr Tyr Val Lys Glu Glu Ser Thr Gln Asn Phe Leu Ser Arg
645 650 655

Asn Pro Asn Ala Glu Phe Asp Gly Arg Gln Leu Lys Leu Leu Gly Asn
660 665 670

Pro Glu Pro Leu Ser Lys Gly Phe Lys Asn Pro Met Ala Leu Lys Ser
675 680 685

Leu His Lys Leu Lys Asn Leu Val Asn Tyr Leu Leu Gln Thr Gly Lys
690 695 700

Ile Asp Glu Gly Thr Arg Val Val Ile Glu Ile Ala Arg Glu Leu Asn

705		710		715		720									
Asp	Lys	Asn	Lys	Arg	Lys	Ala	Ile	Glu	Asn	Trp	Gln	Lys	Asp	Arg	Glu
				725					730					735	
Lys	Glu	Asn	Glu	Ala	Phe	Arg	Lys	Glu	Ile	Glu	Val	Tyr	Lys	Glu	Gln
			740					745					750		
Phe	Pro	Asn	Ile	Asn	Leu	Ile	Asp	Glu	Asn	Thr	Leu	Ile	Arg	Lys	Ile
		755					760					765			
Arg	Leu	Trp	His	Glu	Gln	Asn	Lys	Ile	Cys	Leu	Tyr	Thr	Gly	Lys	Thr
	770					775					780				
Ile	Pro	Phe	Ser	Glu	Leu	Ile	Ile	Gly	Asn	Lys	Tyr	Asp	Phe	Glu	His
785					790					795					800
Thr	Ile	Pro	Ala	Ser	Ile	Ser	Phe	Asp	Asn	Glu	Leu	Lys	Asn	Leu	Thr
			805					810						815	
Ile	Ser	Asp	Ser	Ser	Tyr	Asn	Arg	Leu	Tyr	Lys	Gly	Lys	Lys	Phe	Pro
		820						825					830		
Thr	Gln	Leu	Ser	Asn	Tyr	Asp	Ser	Glu	Gln	Thr	Ile	Asn	Gly	Glu	Ser
	835						840					845			
Cys	Asn	Ser	Ile	Leu	Arg	Asn	Ile	Glu	Phe	Ile	Phe	Gly	Glu	Arg	Ile
	850					855					860				
Val	Glu	Tyr	Lys	Glu	Ile	Lys	Gly	Lys	Thr	Glu	Lys	Ile	Val	Lys	Trp
865					870					875					880
Lys	Lys	Ile	Asp	Glu	Leu	Glu	Lys	Gln	Phe	Asp	Glu	Trp	Lys	Lys	Lys
			885						890					895	
Ala	Ser	Tyr	Ala	Ser	Thr	Lys	Glu	Ile	Lys	Asp	Asn	Cys	Ile	Val	Lys
			900					905					910		
Tyr	His	Thr	Ile	Lys	Met	Asp	Leu	Asp	Tyr	Leu	Lys	Ser	Lys	Leu	Ala
	915						920					925			

Thr Phe Thr Ile Thr Glu Tyr Lys Ala Gly Trp Arg Asn Ser Gln Ile
930 935 940

Arg Asp Thr Gln Ile Ile Thr Lys Tyr Ala Leu Pro Tyr Phe Lys Thr
945 950 955 960

Leu Phe Lys Arg Val Ser Val Glu Lys Gly Gly Val Thr Asp Ile Phe
965 970 975

Lys Lys Val Tyr Lys Val Gln Ser Arg Asn Ser Lys Lys Asp Arg Ser
980 985 990

Val His Ser His His Ala Gln Asp Ala Ala Ile Leu Thr Leu Ile Pro
995 1000 1005

Asn Ala Phe His Arg Glu Arg Ile Ile Lys Ala Tyr Glu Asn Glu
1010 1015 1020

Ile Asp Asn Arg Thr Gly Lys Thr Tyr His Glu Gln Pro Leu Asp
1025 1030 1035

Trp Glu Asn Phe Ser Glu Lys Tyr Ile Leu Glu Leu Gln Asp Lys
1040 1045 1050

Val Leu Ile Asn Asn Leu Thr Asp Asn Arg Thr Ile Thr Gln Thr
1055 1060 1065

Tyr Lys Thr Val Arg Lys Arg Gly Lys Val Val Tyr Asp Ile Asp
1070 1075 1080

Glu Asn Gly Asn Ser Leu Lys Arg Val Ser Lys Gly Asp Thr Ile
1085 1090 1095

Arg Gly Gln Leu His Gly Glu Thr Phe Tyr Gly Ala Ile Lys Gln
1100 1105 1110

Pro Val Arg Asp Glu Asp Asn Lys Ile Leu Phe Asp Glu Asn Lys
1115 1120 1125

Lys Met Ile Leu Lys Asp Glu Ile Tyr Leu Ala Val Arg Lys Pro

1130						1135					1140			
Leu	Val	Tyr	Lys	Lys	Asp	Ala	Asn	Ser	Pro	Gly	Phe	Lys	Thr	Leu
1145						1150					1155			
Asp	Glu	Val	Glu	Lys	Val	Ile	Val	Asp	Lys	Gly	Leu	Phe	Lys	Met
1160						1165					1170			
Ile	Lys	Lys	Gln	Val	Glu	Glu	Ser	Asp	Phe	Lys	Thr	Ala	Leu	Ile
1175						1180					1185			
Asp	Gly	Val	Tyr	Met	Leu	Asn	Glu	Ser	Gly	Glu	Lys	Val	Asn	Lys
1190						1195					1200			
Ile	Arg	Arg	Ile	Arg	Cys	Phe	Glu	Asn	Gly	Leu	Lys	Tyr	Thr	Thr
1205						1210					1215			
Ala	Ile	Lys	Val	His	Glu	His	Ser	Phe	Val	Ser	His	Lys	Glu	Tyr
1220						1225					1230			
Lys	Gln	Ser	Thr	Leu	Ala	Thr	Asn	Gly	Glu	Asn	Thr	Tyr	Cys	Leu
1235						1240					1245			
Phe	Tyr	Lys	Asn	Asp	Lys	Gly	Lys	Ala	Met	Lys	Ile	Leu	Ser	Ile
1250						1255					1260			
Val	Asp	Leu	Ala	Glu	Leu	Lys	Leu	Lys	Asn	Ile	Gln	Asn	Leu	Tyr
1265						1270					1275			
Asn	Glu	Pro	Glu	Phe	Ser	Ser	Phe	Glu	Val	Gly	Lys	Gly	Lys	Asn
1280						1285					1290			
Lys	Tyr	Lys	Ile	Pro	Leu	Tyr	Ser	Val	Leu	Lys	Ser	Gly	Asp	Lys
1295						1300					1305			
Val	Leu	Phe	Tyr	Lys	Glu	Asn	Ile	Tyr	Glu	Leu	Lys	Asp	Leu	Glu
1310						1315					1320			
Gln	Lys	Glu	Leu	Ser	Asp	Arg	Met	Phe	Lys	Met	Tyr	Gln	Phe	Glu
1325						1330					1335			

Lys Asp Gly Arg Ile Lys Phe Arg His His Leu Ala Ala Gly Ile
1340 1345 1350

Asp Thr Glu Leu Lys Lys Glu Asn Pro Glu Asn Ser Ser Val Asn
1355 1360 1365

Phe Glu Glu Lys Gln Val Phe Leu Arg Leu Ser Gln Gly Gln Trp
1370 1375 1380

Asn Phe Ala Ile Asp Gly Lys Asp Phe Glu Ile Ser Leu Asp Gly
1385 1390 1395

Lys Ile Asn Trp Lys Phe
1400

<210> 132
<211> 1092
<212> PRT
<213> Bacillus sp.

<223> APG09208_RGN

<400> 132

Met Arg Tyr Ser Ile Gly Leu Asp Ile Gly Thr Thr Ser Ile Gly Ser
1 5 10 15

Ala Val Ile Asn Asn Asp Leu Gln Arg Phe Glu His Leu Gly Val Arg
20 25 30

Ile Phe Asp Ala Ala Glu Asn Pro Lys Asp Gly Ser Ser Leu Ser Ala
35 40 45

Pro Arg Arg Leu Ala Arg Ser Ser Arg Lys Arg Leu Arg Arg Arg Lys
50 55 60

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

Thr Lys Thr Gln Val Lys Asn Leu Tyr Asn Ser Lys Asn Val Asn Leu
85 90 95

Asp Val Trp Asp Ile Arg Val Ser Gly Leu Asp Arg Lys Leu Phe Asn
100 105 110

Asp Glu Phe Ala Arg Val Leu Ile His Phe Ser Lys Asn Arg Gly Phe
115 120 125

Lys Ser Asn Arg Lys Ser Glu Leu Lys Glu Asp Asp Asn Gly Ala Ile
130 135 140

Leu Ser Ala Val Lys Glu Asn Lys Glu Leu Met Glu Glu Lys Gly Tyr
145 150 155 160

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

Lys Arg Asn Lys Asp Gly Asp Tyr Ser His Val Val Ser Arg Ser Asp
180 185 190

Ile Glu Asn Glu Ile Arg Leu Leu Phe Gln Lys Gln Arg Glu Tyr Gly
195 200 205

His Pro Phe Ala Ile Glu Glu Asn Glu Glu Ala Phe Leu Ala Ile Trp
210 215 220

Ser Ser Gln Arg Pro Phe Ala Thr Lys Asp Asp Ile Val Lys Lys Ile
225 230 235 240

Gly Thr Cys Thr Leu Glu Pro Arg Glu Lys Arg Ala Pro Lys Ser Thr
245 250 255

Tyr Thr Phe Glu Arg Phe Arg Ala Leu Asp Lys Leu Asn Arg Leu Arg
260 265 270

Ile Leu Ser Thr Thr Thr Pro Ser Arg Pro Leu Thr Asn Glu Glu Arg
275 280 285

Asp Ser Ile Leu Asn Ser Leu Phe Ser Lys Lys Glu Val Lys Tyr Lys
290 295 300

Glu Leu Arg Lys Met Leu Lys Leu Thr Asp Asp Gln Arg Phe Asn Glu
305 310 315 320

Ile Tyr Tyr Ser Pro Asp Glu Thr Ile Glu Lys Thr Glu Asn Arg Val
325 330 335

Phe Leu Ser Ser Glu Ser Gln Tyr Lys Ile Lys Lys Ile Ile Glu Lys
340 345 350

Thr Glu Ser Lys Asn Met Leu Ser Thr Tyr Cys Pro Ile Asp Tyr Asp
355 360 365

Thr Ile Gly Tyr Ala Leu Thr Val Phe Lys Asp Asp Lys Asp Ile Gln
370 375 380

His Tyr Leu Gln Asn Arg Tyr Val Asp Ser Lys Gly Lys Ala Ile Ser
385 390 395 400

Asn Met Ala Asn Arg Glu Tyr Asn Leu Glu Leu Ile Glu Glu Leu Leu
405 410 415

Gly Leu Ser Phe Ser Lys Phe Gly His Leu Ser Leu Lys Ala Leu Asn
420 425 430

Asn Ile Leu Pro Tyr Met Glu Glu Gly Glu Pro Tyr His Ile Ala Cys
435 440 445

Glu Met Ala Gly Tyr Gln Phe Ser Gln Arg Leu Thr Lys Glu Lys His
450 455 460

Arg Leu Leu Pro Pro Ile Pro Val Asp Glu Ile Pro Asn Pro Val Val
465 470 475 480

Val Arg Ala Leu Thr Gln Val Arg Lys Val Leu Asn Ser Ile Ile Lys
485 490 495

Lys Tyr Gly Pro Pro Ser Asp Ile Tyr Ile Glu Leu Ala Arg Glu Met
500 505 510

Gly Lys Pro Phe Lys Glu Arg Lys Ser Leu Glu Arg Glu Phe Asn Glu
515 520 525

Asn Arg Gln Val Asn Glu Lys Ala Lys Ala His Ile Ser Glu Leu Tyr
530 535 540

Arg Val Pro Asn Asp Pro Arg Pro His Asp Ile Leu Lys Phe Lys Leu
545 550 555 560

Trp Lys Glu Gln Asn Gly Ile Cys Pro Tyr Ser Phe Lys Pro Ile Ser
565 570 575

Val Asp Tyr Leu Phe Asn Ile Gly Tyr Thr Glu Val Asp His Ile Ile
580 585 590

Pro Tyr Ser Arg Ser Phe Asp Asp Ser Asn Gly Asn Lys Val Leu Val
595 600 605

Leu Thr Asn Glu Asn Arg Asn Lys Leu Asn Arg Thr Pro Tyr Glu Trp
610 615 620

Phe Gly His Asp Glu Asn Arg Trp Gly Asp Phe Val Ser Phe Val Arg
625 630 635 640

Thr Met Lys Val Gly Lys Lys Lys Lys Asp Met Leu Leu Lys Lys Asn
645 650 655

Phe Asp Glu Glu Gln Glu Glu Gln Ile Leu Ser Arg Asn Leu Asn Asp
660 665 670

Thr Arg Tyr Ile Thr Arg Tyr Ile Lys Ser Phe Ile Glu Asp Asn Leu
675 680 685

Glu Phe Arg Thr Glu Glu Asn Lys Lys Gln Tyr Val His Thr Val Asn
690 695 700

Gly Gly Tyr Thr Ser Leu Met Arg Lys Arg Trp Gly Leu Asn Lys Asp
705 710 715 720

Arg Arg Gly Asn Asp Leu His His Ala Val Asp Ala Ala Ile Ile Ala
725 730 735

Val Ser Leu Pro Phe Lys His Lys Val Asn Thr Tyr Phe Lys Arg Gln
740 745 750

Glu Ile Gly Leu Ser Asn Leu Ser Asn Asn Lys Lys Asp Ile Phe Pro
755 760 765

Glu Pro Trp Arg Asn Phe Ile Lys Glu Leu Glu Ala Arg Met Ile Gln
770 775 780

Asp Pro Glu Lys Met Lys Thr Ala Leu Glu Ser Leu Glu Leu Glu Thr
785 790 795 800

Tyr Gly Glu Ile Phe Leu Asn Lys Leu Lys Pro Ile Phe Val Ser Arg
805 810 815

Met Pro Lys His Ser Ile Lys Gly Gln Ile His Glu Glu Thr Ile Arg
820 825 830

Arg Val Arg Gly Leu Thr Glu Glu Gly Phe Leu Val Thr Val Lys Lys
835 840 845

Thr Arg Leu Asp Gln Ile Pro Phe Asp Lys Asn Gly Asp Phe Pro Met
850 855 860

Tyr Gly Lys Glu Thr Asp Ile Lys Thr Tyr Met Ala Ile Lys Gln Arg
865 870 875 880

Tyr Leu Glu Tyr Gly Gln Asp Lys Gln Lys Ala Phe Ala Ile Pro Leu
885 890 895

Arg Lys Pro Ser Lys Asn Pro Glu Asn Ala Pro Ile Val Arg Ser Val
900 905 910

Lys Ile Glu Gly Lys Ala Asn Arg Val Val Met Leu Asp Asp Lys Ala
915 920 925

Ala Ala Asp Asn Ala Ser Ile Ala Arg Thr Glu Val Phe Arg His Lys
930 935 940

Glu Thr Gly Glu Tyr Tyr Val Thr Pro Val Tyr Val Ala Asp Ile Val
945 950 955 960

Ser Asn Lys Ile Pro Asp Arg Leu Ile Thr Leu Lys Lys Pro Tyr Ser
965 970 975

Glu Trp Asp Arg Ile Thr Asp Glu Tyr Glu Tyr Leu Phe Ser Leu Tyr
980 985 990

Asn Asn Asp Leu Ile Lys Ile Val Leu Pro Lys Glu Lys Glu Thr Lys
995 1000 1005

Lys Tyr Thr Gly Gly Asn His Leu Trp Gln Glu Gly Phe Phe Tyr
1010 1015 1020

Phe Lys Gly Val Asp Ser Ser Asn Ala Ser Ile Lys Ile Ile Asp
1025 1030 1035

His Leu Asn Ser Phe Glu Val Arg Ile Gly Ile Arg Arg Leu Ile
1040 1045 1050

Ala Phe Glu Lys Tyr Gln Val Asn Pro Leu Gly Glu Ile Asn Lys
1055 1060 1065

Val His Gly Glu Lys Arg Pro Gly Glu Leu Leu Asn Glu Glu Glu
1070 1075 1080

Ile Lys Glu Asn Arg Lys Asn Ile Ser
1085 1090

<210> 133
<211> 1150
<212> PRT
<213> Enterococcus sp.

<223> APG05586_RGN

<400> 133

Met Tyr Ser Ile Gly Leu Asp Leu Gly Ile Ser Ser Val Gly Trp Ser
1 5 10 15

Val Ile Asp Glu Glu Thr Gly Lys Ile Val Asp Leu Gly Val Arg Leu
20 25 30

Phe Ser Ala Lys Asn Ser Glu Lys Asn Leu Glu Arg Arg Thr Ser Arg

35

40

45

Gly Ala Arg Arg Leu Ile Arg Arg Lys Thr Asn Arg Leu Lys Asp Ala
 50 55 60

Lys Lys Leu Leu Glu Ala Ile Gly Phe Tyr Glu Asp Lys Ala Leu Lys
 65 70 75 80

Asn Val Cys Pro Tyr Gln Leu Arg Val Lys Gly Leu Thr Glu Gly Leu
 85 90 95

Thr Lys Gly Glu Leu Tyr Lys Val Val Leu His Ile Val Lys Lys Arg
 100 105 110

Gly Ile Ser Tyr Leu Asp Glu Asp Asp Ala Glu Ala Ala Lys Glu Ser
 115 120 125

Gln Asp Tyr Lys Glu Gln Val Arg Lys Asn Ala Gln Leu Leu Thr Lys
 130 135 140

Tyr Thr Pro Gly Gln Ile Gln Leu Gln Arg Leu Lys Glu Asn Asn Arg
 145 150 155 160

Val Lys Thr Gly Ile Asn Gly Gln Gly His Tyr Gln Leu Asn Val Phe
 165 170 175

Lys Val Ser Ala Tyr Ala Asp Glu Leu Ala Thr Ile Leu Lys Thr Gln
 180 185 190

Gln Ala Leu Tyr Pro Asn Glu Leu Thr Asp Asp Trp Ile Ala Leu Phe
 195 200 205

Val Gln Pro Gly Ile Ala Glu Asn Ala Gly Leu Ile Tyr Arg Lys Arg
 210 215 220

Pro Tyr Tyr His Gly Pro Gly Asn Glu Ala Asn Asn Ser Pro Tyr Gly
 225 230 235 240

Arg Trp Ser Asp Phe Gln Lys Thr Gly Gln Pro Ala Ala Asn Ile Phe
 245 250 255

Asp Lys Leu Ile Gly Lys Asp Phe Gln Gly Glu Leu Arg Ala Ser Gly
260 265 270

Leu Ser Leu Ser Ala Gln Gln Tyr Asn Leu Leu Asn Asp Leu Thr Asn
275 280 285

Leu Lys Ile Asp Gly Glu Val Ser Leu Ser Pro Glu Gln Lys Glu Phe
290 295 300

Ile Leu Thr Glu Leu Met Thr Lys Glu Phe Ala Arg Phe Gly Val Asn
305 310 315 320

Asp Ile Ala Lys Leu Leu Gly Val Lys Lys Glu Gln Leu Ser Gly Trp
325 330 335

Arg Leu Asp Lys Lys Gly Lys Pro Glu Ile His Thr Leu Lys Gly Tyr
340 345 350

Arg Asn Trp Arg Lys Ile Phe Ala Glu Ala Gly Ile Asp Leu Ala Thr
355 360 365

Leu Pro Thr Glu Thr Ile Asp Cys Leu Ala Lys Val Leu Thr Leu Asn
370 375 380

Thr Glu Arg Glu Gly Val Glu Asn Thr Leu Ala Phe Glu Leu Pro Glu
385 390 395 400

Leu Ala Glu Pro Val Lys Ser Leu Val Leu Asp His Tyr Lys Glu Leu
405 410 415

Ser Gln Ser Ile Ser Thr Gln Ala Trp His Arg Phe Ser Leu Lys Thr
420 425 430

Leu His Leu Leu Ile Pro Glu Leu Ile Lys Ser Thr Ser Glu Gln Asn
435 440 445

Thr Leu Leu Glu Gln Phe Gln Leu Lys Ala Gly Val Arg Lys Arg Tyr
450 455 460

Ser Asp Tyr Lys Lys Leu Pro Thr Lys Glu Val Leu Ala Glu Ile Tyr

465		470		475		480									
Asn	Pro	Thr	Val	Asn	Lys	Thr	Val	Ser	Gln	Ala	Phe	Lys	Val	Met	Asp
				485					490					495	
Ala	Leu	Leu	Glu	Lys	Tyr	Gly	Lys	Asp	Gln	Ile	His	Tyr	Ile	Thr	Val
			500					505					510		
Glu	Met	Pro	Arg	Asp	Asp	Asn	Glu	Glu	Glu	Glu	Arg	Lys	Arg	Ile	Lys
		515					520					525			
Glu	Leu	Gln	Thr	Lys	Asn	Ser	Gln	Arg	Lys	Asn	Asp	Ser	Gln	Gln	Tyr
	530					535					540				
Phe	Leu	Gln	Lys	Ser	Gly	Trp	Ser	Gln	Glu	Lys	Phe	Gln	Ala	Thr	Ile
545					550					555					560
His	Lys	Asn	Arg	Arg	Phe	Leu	Ala	Lys	Leu	Leu	Tyr	Tyr	Phe	Glu	Gln
				565					570					575	
Asp	Gly	Val	Cys	Ala	Tyr	Thr	Gly	Asn	Pro	Ile	Ser	Pro	Glu	Leu	Leu
			580					585					590		
Val	Ser	Asp	Ser	Thr	Glu	Ile	Asp	His	Ile	Ile	Pro	Ile	Ser	Ile	Ser
		595					600					605			
Leu	Asp	Asp	Ser	Ile	Asn	Asn	Lys	Val	Leu	Val	Leu	Ser	His	Ala	Asn
	610					615					620				
Gln	Val	Lys	Gly	Gln	Gln	Thr	Pro	Tyr	Asp	Ala	Arg	Met	Ala	Gly	Ala
625					630					635					640
Phe	Asn	Lys	Ile	Asn	Gly	Lys	Phe	Ser	Asn	Trp	Asp	Glu	Tyr	Gln	Lys
				645					650					655	
Trp	Val	Glu	Ser	Arg	Pro	Phe	Ser	Arg	Lys	Lys	Val	Asn	Asn	Leu	Leu
			660					665					670		
Glu	Thr	Arg	Asn	Ile	Phe	Asp	Ser	Glu	Gln	Val	Gln	Lys	Phe	Leu	Ser
		675					680					685			

Arg Asn Leu Asn Asp Thr Arg Tyr Ala Ser Arg Leu Val Leu Asn Thr
690 695 700

Leu Gln Ser Phe Phe Glu Asn Gln Asp Thr Ile Val Arg Val Val Asn
705 710 715 720

Gly Ser Phe Thr His Thr Leu Arg Lys Lys Trp Gly Ala Asp Leu Asp
725 730 735

Lys Thr Arg Glu Thr His His His His Ala Val Asp Ala Thr Leu Cys
740 745 750

Ala Val Thr Pro Phe Val Lys Val Ser Arg Tyr His Tyr Ala Val Asn
755 760 765

Glu Glu Thr Gly Glu Lys Phe Met Arg Glu Ile Asp Val Glu Thr Gly
770 775 780

Glu Ile Leu Asp Glu Ile Pro Tyr Arg Glu Tyr Lys Lys Ala Lys His
785 790 795 800

Tyr Glu Arg Lys Thr Tyr Gln Val Lys Trp Ser Asn Phe Arg Glu Gln
805 810 815

Leu Lys Pro Ile Thr Ile His Pro Lys Ile Lys Phe Ser His Gln Val
820 825 830

Asp Arg Lys Ala Asn Arg Lys Leu Ser Asp Ala Thr Ile Tyr Ser Val
835 840 845

Arg Glu Lys Thr Glu Val Lys Thr Leu Lys Ser Gly Lys Glu Lys Ile
850 855 860

Thr Thr Asp Glu Tyr Thr Ile Gly Lys Ile Lys Asp Ile Tyr Thr Val
865 870 875 880

Asp Gly Trp Glu Ala Phe Lys Lys Lys Gln Asp Lys Leu Leu Met Lys
885 890 895

Glu Phe Asp Glu Lys Thr Tyr Glu Leu Leu Val Thr Ile Ala Ala Thr

900

905

910

Thr Pro Asp Phe Gln Glu Val Glu Glu Lys Asn Gly Lys Val Lys Arg
 915 920 925

Val Lys Arg Ser Pro Phe Ala Val Tyr Cys Glu Glu Asn Gly Ile Pro
 930 935 940

Ala Ile Arg Lys Tyr Ala Lys Lys Asn Asn Gly Pro Val Ile Arg Ser
 945 950 955 960

Leu Lys Tyr Tyr Asp Gly Lys Leu Asn Lys His Ile Asn Ile Thr Lys
 965 970 975

Asp Glu Lys Gly Arg Pro Val Glu Gln Thr Lys Asn Gly Arg Lys Val
 980 985 990

Thr Leu Gln Ser Leu Lys Pro Tyr Arg Tyr Asp Ile Tyr Gln Asp Leu
 995 1000 1005

Glu Thr Lys Ala Tyr Tyr Thr Val Gln Leu Tyr Tyr Ser Asp Leu
 1010 1015 1020

Arg Phe Val Glu Gly Glu Tyr Gly Ile Thr Glu Lys Glu Tyr Met
 1025 1030 1035

Lys Lys Val Ala Glu Gln Thr Lys Gly Gln Val Val Arg Phe Cys
 1040 1045 1050

Phe Ser Leu Gln Lys Asn Asp Gly Leu Glu Ile Glu Trp Lys Asp
 1055 1060 1065

Ser Gln Arg Tyr Asp Val Arg Phe Tyr Asn Phe Gln Ser Ala Asn
 1070 1075 1080

Ser Ile Asn Phe Lys Gly Leu Glu Gln Glu Met Ile Pro Ala Glu
 1085 1090 1095

Asn Gln Phe Lys Gln Lys Pro Tyr Asn Asn Gly Ala Ile Asn Leu
 1100 1105 1110

Asn Ile Ala Lys Tyr Gly Lys Glu Gly Lys Lys Leu Arg Lys Phe
1115 1120 1125

Asn Thr Asp Ile Leu Gly Lys Lys His His Leu Ser Tyr Glu Lys
1130 1135 1140

Glu Pro Lys Asn Ile Ile Lys
1145 1150

<210> 134
<211> 1150
<212> PRT
<213> Enterococcus sp.

<223> APG08770_RGN

<400> 134

Met Tyr Ser Ile Gly Leu Asp Leu Gly Ile Ser Ser Val Gly Trp Ser
1 5 10 15

Val Ile Asp Glu Glu Thr Gly Lys Ile Val Asp Leu Gly Val Arg Leu
20 25 30

Phe Ser Ala Lys Asn Ser Glu Lys Asn Leu Glu Arg Arg Thr Ser Arg
35 40 45

Gly Ala Arg Arg Leu Ile Arg Arg Lys Thr Asn Arg Leu Lys Asp Ala
50 55 60

Lys Lys Leu Leu Glu Ala Ile Gly Phe Tyr Glu Asp Lys Ala Leu Lys
65 70 75 80

Asn Val Cys Pro Tyr Gln Leu Arg Val Lys Gly Leu Thr Glu Gly Leu
85 90 95

Thr Lys Gly Glu Leu Tyr Lys Val Val Leu His Ile Val Lys Lys Arg
100 105 110

Gly Ile Ser Tyr Leu Asp Glu Asp Asp Ala Glu Ala Ala Lys Glu Ser
115 120 125

Gln Asp Tyr Lys Glu Gln Val Arg Lys Asn Ala Gln Leu Leu Thr Lys
130 135 140

Tyr Thr Pro Gly Gln Ile Gln Leu Gln Arg Leu Lys Glu Asn Asn Arg
145 150 155 160

Val Lys Thr Gly Ile Asn Gly Gln Gly Asn Tyr Gln Leu Asn Val Phe
165 170 175

Lys Val Ser Ala Tyr Ala Asp Glu Leu Ala Thr Ile Leu Lys Thr Gln
180 185 190

Gln Ala Leu Tyr Pro Asn Glu Leu Thr Asp Asp Trp Ile Ala Leu Phe
195 200 205

Val Gln Pro Gly Ile Ala Glu Asp Ala Gly Leu Ile Tyr Arg Lys Arg
210 215 220

Pro Tyr Tyr His Gly Pro Gly Asn Glu Ala Asn Asn Ser Pro Tyr Gly
225 230 235 240

Arg Trp Ser Asp Phe Gln Lys Thr Gly Gln Pro Ala Ala Asn Ile Phe
245 250 255

Asp Lys Leu Ile Gly Lys Asp Phe Gln Gly Glu Leu Arg Ala Ser Gly
260 265 270

Leu Ser Leu Ser Ala Gln Gln Tyr Asn Leu Leu Asn Asp Leu Thr Asn
275 280 285

Leu Lys Ile Asp Gly Glu Gln Pro Leu Ser Pro Thr Gln Lys Glu Phe
290 295 300

Ile Leu Thr Glu Leu Met Thr Lys Glu Phe Ala Arg Phe Gly Val Asn
305 310 315 320

Asp Ile Ala Lys Leu Leu Gly Val Lys Lys Glu Gln Leu Ser Gly Trp
325 330 335

Arg Leu Asp Lys Lys Gly Lys Pro Glu Ile His Thr Leu Lys Gly Tyr
340 345 350

Arg Asn Trp Arg Lys Ile Phe Ala Glu Ala Gly Ile Asp Leu Ala Thr
355 360 365

Leu Pro Thr Glu Thr Ile Asp Cys Leu Ala Lys Val Leu Thr Leu Asn
370 375 380

Thr Glu Arg Glu Gly Val Glu Asn Thr Leu Ala Phe Glu Leu Pro Glu
385 390 395 400

Leu Ala Glu Pro Val Lys Ser Leu Val Leu Asp His Tyr Lys Glu Leu
405 410 415

Ser Gln Ser Ile Ser Thr Gln Ala Trp His Arg Phe Ser Leu Lys Thr
420 425 430

Leu His Leu Leu Ile Pro Glu Leu Ile Lys Ser Thr Ser Glu Gln Asn
435 440 445

Thr Leu Leu Glu Gln Phe Gln Leu Lys Ala Gly Val Arg Lys Arg Tyr
450 455 460

Ser Asp Tyr Lys Lys Leu Pro Thr Lys Glu Val Leu Ala Glu Ile Tyr
465 470 475 480

Asn Pro Thr Val Asn Lys Thr Val Ser Gln Ala Phe Lys Val Met Asp
485 490 495

Ala Leu Leu Glu Lys Tyr Gly Lys Asp Gln Ile His Tyr Ile Thr Val
500 505 510

Glu Met Pro Arg Asp Asp Asn Glu Glu Glu Glu Arg Lys Arg Ile Lys
515 520 525

Glu Leu Gln Thr Lys Asn Ser Gln Arg Lys Asn Asp Ser Gln Gln Tyr
530 535 540

Phe Leu Gln Lys Ser Gly Trp Ser Gln Glu Lys Phe Gln Ala Thr Ile
545 550 555 560

His Lys Asn Arg Arg Phe Leu Ala Lys Leu Leu Tyr Tyr Phe Glu Gln
565 570 575

Asp Gly Val Cys Ala Tyr Thr Gly Asn Pro Ile Ser Pro Glu Leu Leu
580 585 590

Val Ser Asp Ser Thr Glu Ile Asp His Ile Ile Pro Ile Ser Ile Ser
595 600 605

Leu Asp Asp Ser Ile Asn Asn Lys Val Leu Val Leu Ser His Ala Asn
610 615 620

Gln Val Lys Gly Gln Gln Thr Pro Tyr Asp Ala Trp Met Ala Gly Ala
625 630 635 640

Phe Lys Lys Ile Asn Gly Lys Phe Ser Asn Trp Asp Glu Tyr Gln Lys
645 650 655

Trp Val Glu Ser Arg Pro Phe Ser Arg Lys Lys Ala Asn Asn Leu Leu
660 665 670

Glu Thr Arg Asn Ile Phe Asp Ser Glu Gln Val Gln Asn Phe Leu Ser
675 680 685

Arg Asn Leu Asn Asp Thr Arg Tyr Ala Ser Arg Leu Val Leu Asn Thr
690 695 700

Leu Gln Ser Phe Phe Glu Asn Gln Asp Thr Ile Val Arg Val Val Asn
705 710 715 720

Gly Ser Phe Thr His Thr Leu Arg Lys Lys Trp Gly Ala Asp Leu Asp
725 730 735

Lys Thr Arg Glu Thr His His His His Ala Ile Asp Ala Ser Leu Cys
740 745 750

Ala Val Thr Pro Phe Val Lys Val Ser Arg Tyr His Tyr Ala Val Asn
755 760 765

Glu Glu Thr Gly Glu Lys Phe Met Arg Glu Ile Asp Val Glu Thr Gly
770 775 780

Glu Ile Leu Asp Glu Ile Pro Tyr Arg Glu Tyr Lys Lys Ala Lys His
785 790 795 800

Tyr Glu Arg Lys Thr Tyr Gln Val Lys Trp Ser Asn Phe Arg Glu Gln
805 810 815

Leu Lys Pro Leu Thr Ile His Pro Lys Ile Lys Phe Ser His Gln Val
820 825 830

Asp Arg Lys Ala Asn Arg Lys Leu Ser Asp Ala Thr Ile Tyr Ser Val
835 840 845

Arg Glu Lys Thr Glu Val Lys Thr Leu Lys Ser Gly Lys Glu Lys Ile
850 855 860

Thr Thr Asp Glu Tyr Thr Ile Gly Lys Ile Lys Asp Ile Tyr Thr Val
865 870 875 880

Asp Gly Trp Glu Ala Phe Lys Lys Lys Gln Asp Lys Leu Leu Met Lys
885 890 895

Glu Phe Asp Glu Lys Thr Tyr Glu Leu Leu Val Thr Ile Ala Ala Thr
900 905 910

Thr Pro Asp Phe Gln Glu Val Glu Glu Lys Asn Gly Lys Val Lys Arg
915 920 925

Val Lys Arg Ser Pro Phe Ala Val Tyr Cys Glu Glu Asn Gly Ile Pro
930 935 940

Ala Ile Arg Lys Tyr Ala Lys Lys Asn Asn Gly Pro Val Ile Arg Ser
945 950 955 960

Leu Lys Tyr Tyr Asp Gly Lys Leu Asn Lys His Ile Asn Ile Thr Lys
965 970 975

Asp Glu Lys Gly Arg Pro Val Glu Gln Thr Lys Asn Gly Arg Lys Val
980 985 990

Thr Leu Gln Ser Leu Lys Pro Tyr Arg Tyr Asp Ile Tyr Gln Asp Leu
995 1000 1005

Glu Thr Lys Ala Tyr Tyr Thr Val Gln Leu Tyr Tyr Ser Asp Leu
1010 1015 1020

Arg Phe Val Glu Gly Glu Tyr Gly Ile Thr Glu Lys Glu Tyr Met
1025 1030 1035

Lys Lys Val Ala Glu Gln Thr Lys Gly Gln Val Val Arg Phe Cys
1040 1045 1050

Phe Ser Leu Gln Lys Asn Asp Gly Leu Glu Ile Glu Trp Lys Asp
1055 1060 1065

Ser Gln Arg Tyr Asp Val Arg Phe Tyr Asn Phe Gln Ser Ala Asn
1070 1075 1080

Ser Ile Asn Phe Lys Gly Leu Glu Gln Glu Met Met Pro Ala Glu
1085 1090 1095

Asn Gln Phe Lys Gln Lys Pro Tyr Asn Asn Gly Ala Ile Asn Leu
1100 1105 1110

Asn Ile Ala Lys Tyr Gly Lys Glu Gly Asn Lys Leu Arg Lys Phe
1115 1120 1125

Asn Thr Asp Ile Leu Gly Lys Lys His His Leu Ser Tyr Glu Lys
1130 1135 1140

Glu Pro Lys Asn Ile Ile Lys
1145 1150

<210> 135
<211> 1053
<212> PRT
<213> Staphylococcus sp.

<223> APG08167_RGN

<400> 135

Met Lys Arg His Tyr Ile Leu Gly Leu Asp Ile Gly Ile Thr Ser Val

1	5	10	15
Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Glu Val Ile Asp Ala Gly	20	25	30
Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg	35	40	45
Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Lys His Arg Leu	50	55	60
Gln Ser Val Lys Lys Leu Leu Phe Asp Tyr Asp Leu Leu Thr Asn His	65	70	75
Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu	85	90	95
Ser Gln Asn Leu Ser Glu Val Glu Phe Ser Gly Ala Leu Leu His Leu	100	105	110
Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr	115	120	125
Arg Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala	130	135	140
Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Thr	145	150	155
Asp Ser Glu Val Arg Gly Ala Ile Asn Arg Phe Lys Thr Ser Asp Tyr	165	170	175
Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Glu Tyr His Gln	180	185	190
Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg	195	200	205
Arg Thr Tyr Tyr Glu Gly Pro Gly Lys Gly Ser Pro Phe Gly Trp Lys	210	215	220

Asp Ile Lys Glu Trp Tyr Glu Met Leu Ile Gly His Cys Thr Tyr Phe
225 230 235 240

Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr
245 250 255

Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Ser Arg Asp Asp Asn
260 265 270

Glu Lys Leu Ala Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe
275 280 285

Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ser Lys Glu Ile Phe
290 295 300

Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys
305 310 315 320

Pro Glu Phe Thr Lys Phe Lys Ile Tyr His Asp Ile Lys Gly Ile Lys
325 330 335

Glu Arg Lys Glu Ile Ile Glu Asn Ala Glu Val Leu Asp Gln Ile Ala
340 345 350

Glu Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu
355 360 365

Ala Lys Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser
370 375 380

Lys Leu Thr Gly Tyr Thr Gly Thr His Ser Leu Ser Leu Lys Ala Ile
385 390 395 400

Asn Leu Ile Leu Asp Glu Leu Trp His Thr Ser Asp Asn Gln Met Ala
405 410 415

Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Met Asp Leu Ser Gln
420 425 430

Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro

435

440

445

Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile
 450 455 460

Ile Lys Lys Tyr Gly Leu Pro Ser Asp Ile Ile Ile Glu Leu Ala Arg
 465 470 475 480

Glu Lys Asn Ser Lys Asp Ala Gln Lys Leu Ile Asn Glu Met Gln Lys
 485 490 495

Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr
 500 505 510

Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asn
 515 520 525

Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu
 530 535 540

Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro
 545 550 555 560

Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys
 565 570 575

Gln Glu Glu Asn Ser Lys Lys Asp Asn Arg Thr Pro Phe Gln Tyr Leu
 580 585 590

Ser Ser Ser Asp Ser Lys Met Asp Tyr Glu Thr Phe Lys Lys His Ile
 595 600 605

Leu Asn Leu Ser Lys Gly Lys Asp Arg Ile Ser Lys Lys Lys Lys Val
 610 615 620

Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp
 625 630 635 640

Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Glu Leu
 645 650 655

Met Thr Leu Leu Arg Ser Tyr Phe Ser Val Asn Asn Leu Asp Val Lys
660 665 670

Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp
675 680 685

Lys Phe Lys Lys Glu Arg Asn Gln Gly Tyr Lys His His Ala Glu Asp
690 695 700

Ala Leu Ile Ile Ala Asn Ala Asp Tyr Val Phe Lys Glu Trp Lys Lys
705 710 715 720

Leu Asp Arg Ala Asn Lys Val Met Glu Asn Gln Thr Val Glu Glu Asn
725 730 735

Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu
740 745 750

Ile Phe Ile Thr Pro His Gln Ile Gln His Ile Lys Asp Phe Lys Asp
755 760 765

Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Ser Asn Arg Glu Leu Val
770 775 780

Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu
785 790 795 800

Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu
805 810 815

Lys Lys Leu Met Asn Gln Ser Pro Glu Lys Leu Leu Met Tyr His His
820 825 830

Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly
835 840 845

Gly Glu Lys Asn Pro Leu Tyr Lys Tyr His Glu Glu Thr Gly Asn Tyr
850 855 860

Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile

865		870		875		880
Lys Tyr Tyr Gly	Asn Lys Leu Asn Val	His Leu Asp Ile Thr	Asp Asp			
	885	890	895			
Tyr Pro Asn Ser Arg	Asn Lys Val Val	Lys Leu Ser Leu	Lys Pro Phe			
	900	905	910			
Arg Phe Asp Val Tyr	Leu Asp Asn Gly Val	Tyr Lys Phe Val	Thr Ile			
	915	920	925			
Lys Asn Leu Asp Val	Ile Lys Lys Glu Asp Tyr	Tyr Glu Val Asn Ser				
	930	935	940			
Lys Cys Tyr Lys Glu	Ala Lys Lys Leu Lys	Gln Ile Ser Asp Gln	Ala			
945	950	955	960			
Glu Phe Ile Ala Ser	Phe Tyr Asn Asn Asp	Leu Ile Lys Val	Asn Gly			
	965	970	975			
Glu Leu Tyr Arg Ile	Ile Gly Val Asn Asn Asp	Leu Leu Asn Arg	Ile			
	980	985	990			
Glu Val Asn Met Val	Asp Ile Thr Tyr Lys Glu	Tyr Leu Glu Thr Met				
	995	1000	1005			
Asn Asp Lys Arg Thr	Pro Arg Ile Phe Lys	Ala Ile Thr Gly	Lys			
	1010	1015	1020			
Thr Gln Ser Ile Lys	Lys Tyr Ser Thr Asp	Ile Leu Gly Asn	Leu			
	1025	1030	1035			
Tyr Glu Val Lys Ser	Lys Gln Asn Pro Gln	Met Ile Met Lys	Arg			
	1040	1045	1050			

<210> 136
 <211> 1052
 <212> PRT
 <213> Staphylococcus sp.
 <223> APG01604_RGN

<400> 136

Met Val Thr Lys Tyr Ile Leu Gly Leu Asp Ile Gly Ile Thr Ser Val
1 5 10 15

Gly Tyr Gly Ile Ile Asn Tyr Glu Asp Lys Thr Ile Ile Asp Ala Gly
20 25 30

Val Arg Leu Phe Pro Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg
35 40 45

Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Ile His Arg Leu
50 55 60

Asp Arg Ile Lys Gln Leu Leu Ser Glu Tyr Asn Leu Val Asp Leu Asp
65 70 75 80

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

Arg Glu Glu Leu Ser Lys Asp Glu Leu Val Ile Ala Leu Leu His Ile
100 105 110

Ala Lys Arg Arg Gly Ile His Asn Val Glu Ala Val Asp Glu Thr Gln
115 120 125

Asp Glu Gly Asn Glu Leu Ser Thr Lys Glu Gln Leu Ala Lys Asn Asn
130 135 140

Asn Leu Leu Lys Asp Lys Tyr Val Cys Glu Leu Leu Leu Glu Arg Leu
145 150 155 160

Lys Asp Gly Lys Val Arg Gly Glu Lys Asn Arg Phe Lys Thr Thr Asp
165 170 175

Ile Ile Lys Glu Val Lys Gln Leu Leu Glu Thr Gln Lys Glu Ala His
180 185 190

Gln Leu Asp Asp Asp Phe Ile Asn Arg Tyr Ile Asp Leu Ile Glu Thr
195 200 205

Arg Arg Glu Tyr Phe Glu Gly Pro Gly Lys Gly Ser Pro Phe Gly Trp
210 215 220

Gly Gly Asp Leu Lys Lys Trp Tyr Glu Thr Leu Met Gly His Cys Thr
225 230 235 240

Tyr Phe Pro Asn Glu Leu Arg Ser Val Lys Tyr Ser Tyr Ser Ala Asp
245 250 255

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

Gly Asn Ser Lys Leu Glu Tyr His Glu Lys Tyr His Ile Ile Glu Asn
275 280 285

Val Phe Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Asn Glu
290 295 300

Ile Gly Val Ser Pro Asp Asp Ile Lys Gly Phe Arg Ile Asn Lys Ser
305 310 315 320

Gly Lys Glu Thr Phe Thr Glu Phe Lys Leu Tyr His Asp Leu Lys Lys
325 330 335

Val Leu Ile Asp Gln Ser Ile Leu Glu Asn Val Gln Leu Leu Asp Gln
340 345 350

Ile Ala Glu Ile Leu Thr Ile Tyr Gln Asp Lys Glu Ser Ile Ile Asn
355 360 365

Glu Leu Asn Gln Leu Ser Glu Ile Ile Asn Glu Gln Asp Lys Glu Ser
370 375 380

Ile Ser Asn Leu Ser Gly Tyr Asn Gly Thr His Arg Leu Ser Leu Lys
385 390 395 400

Cys Ile Asn Leu Val Ile Glu Glu Leu Trp His Thr Ser Arg Asn Gln
405 410 415

Met Glu Ile Phe Thr Tyr Leu Asn Ile Lys Pro Lys Lys Ile Asp Leu
420 425 430

Ala Lys Thr Asn Lys Ile Pro Lys Asn Met Ile Asp Glu Phe Ile Leu
435 440 445

Ser Pro Val Val Lys Arg Thr Phe Gly Gln Ala Ile Asn Val Ile Asn
450 455 460

Lys Val Ile Glu Lys Tyr Gly Val Pro Glu Asp Ile Ile Ile Glu Leu
465 470 475 480

Ala Arg Glu Ser Asn Ser Lys Asp Lys Gln Lys Phe Ile Asn Ser Leu
485 490 495

Gln Lys Lys Asn Glu Thr Thr Arg Lys Arg Ile Asn Glu Ile Ile Gly
500 505 510

Gln Tyr Gly Asn Gln Asn Ala Lys Arg Leu Val Glu Lys Ile Arg Leu
515 520 525

His Asp Glu Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ser Ile Pro
530 535 540

Leu Glu Asp Leu Ile Asn Asn Pro Gln Tyr Tyr Glu Val Asp His Ile
545 550 555 560

Ile Pro Arg Ser Val Ser Phe Asp Asn Ser Tyr Gln Asn Lys Val Leu
565 570 575

Val Lys Gln Thr Glu Asn Ser Lys Lys Ser Asn Arg Thr Pro Tyr Gln
580 585 590

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

His Val Leu Asn Leu Ser Lys Ser Lys Asp Arg Ile Ser Lys Lys Lys
610 615 620

Lys Glu Tyr Leu Leu Glu Glu Arg Asp Ile Asn Lys Tyr Glu Val Gln
625 630 635 640

Lys Glu Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg
645 650 655

Glu Leu Thr Asn Tyr Leu Lys Ala Tyr Phe Ser Ala Asn Asp Met Asp
660 665 670

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

Val Trp Lys Phe Lys Lys Glu Arg Asn His Gly Tyr Lys His His Ala
690 695 700

Glu Asp Ala Leu Ile Ile Ala Asn Ala Asp Phe Leu Phe Lys Glu Asn
705 710 715 720

Lys Lys Leu Lys Lys Ala Asn Ala Ile Leu Glu Gln Pro Ser Leu Asp
725 730 735

Asn Gly Lys Ser Asp Ala Thr Val Glu Asn Asp Asn Glu Tyr Val Glu
740 745 750

Thr Phe Ser Ile Pro Lys Gln Val Asn Asp Ile Lys Glu Phe Arg Asp
755 760 765

Phe Lys Phe Ser His Arg Val Asp Lys Lys Pro Asn Arg Gln Leu Ile
770 775 780

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

Val Ser Pro Ile Thr Asn Ile Tyr Ser Lys Asp Asn Asp Glu Leu Lys
805 810 815

Lys Lys Phe Asn Lys Asn Pro Glu Lys Phe Leu Met Tyr Gln His Asp
820 825 830

Pro Lys Thr Phe Glu Lys Leu Glu Val Ile Met Lys Gln Tyr Ala Asn
835 840 845

Glu Lys Asn Pro Leu Ala Lys Tyr His Glu Glu Thr Gly Glu Tyr Leu
850 855 860

Thr Lys Tyr Ser Lys Lys Asn Asn Gly Pro Ile Val Lys Thr Ile Lys
865 870 875 880

Val Leu Gly Asp Lys Val Gly Lys His Leu Asp Val Thr His Lys Tyr
885 890 895

Lys Tyr Ser Asn Ser Lys Ile Val Lys Lys Thr Ile Asn Pro Tyr Arg
900 905 910

Phe Asp Val Tyr Leu Thr Asp Lys Gly Tyr Lys Phe Ile Thr Ile Ser
915 920 925

Tyr Leu Asp Val Leu Lys Lys Asp Asn Tyr Tyr Tyr Ile Leu Lys Glu
930 935 940

Lys Tyr Glu Glu Leu Lys Ile Lys Lys Ser Ile Ser Asp Thr Asp Gln
945 950 955 960

Phe Ile Gly Ser Phe Tyr Tyr Asn Asp Leu Ile Lys Ile Asn Asp Gln
965 970 975

Ile Phe Lys Val Val Gly Val Asn Asn Asp Leu Leu Asn Arg Ile Glu
980 985 990

Leu Asp Leu Leu Asp Ile Ser Tyr Lys Glu Tyr Cys Lys Ile Asn Asn
995 1000 1005

Ile Lys Thr Asn Arg Ile Ile Lys Gly Ile Thr Lys Lys Ile Thr
1010 1015 1020

Asn Ile Glu Lys Phe Ser Thr Asp Val Leu Gly Asn Leu Tyr Lys
1025 1030 1035

Ala His Ser Asn His Pro Gln Leu Ile Phe Lys Gln Arg Asp
1040 1045 1050

<210> 137
<211> 1129
<212> PRT
<213> Streptococcus sp.

<223> APG03021_RGN

<400> 137

Met Ser Asn Gly Lys Ile Leu Gly Leu Asp Ile Gly Val Ala Ser Val
1 5 10 15

Gly Val Gly Ile Ile Asp Ala Lys Thr Gly Asn Val Ile His Ala Asn
20 25 30

Ser Arg Leu Phe Ser Ala Ala Asn Ala Glu Asn Asn Ala Glu Arg Arg
35 40 45

Gly Phe Arg Gly Ala Arg Arg Leu Thr Arg Arg Lys Lys His Arg Val
50 55 60

Lys Arg Val Arg Asp Leu Phe Glu Lys Tyr Asp Ile Ser Thr Asp Phe
65 70 75 80

Arg Asn Leu Asn Leu Asn Pro Tyr Glu Leu Arg Val Lys Gly Leu Thr
85 90 95

Glu Gln Leu Thr Asn Glu Glu Leu Phe Ala Ala Leu Arg Thr Ile Ala
100 105 110

Lys Arg Arg Gly Ile Ser Tyr Leu Asp Asp Ala Glu Asp Asp Ser Thr
115 120 125

Gly Ser Ser Asp Tyr Ala Lys Ser Ile Asp Glu Asn Arg Arg Leu Leu
130 135 140

Lys Thr Lys Thr Pro Gly Gln Ile Gln Leu Glu Arg Leu Glu Lys Tyr
145 150 155 160

Gly Gln Leu Arg Gly Asn Phe Thr Val Tyr Asp Glu Asn Gly Glu Ala
165 170 175

His Arg Leu Ile Asn Val Phe Ser Thr Ser Asp Tyr Lys Asn Glu Ala
180 185 190

Arg Lys Ile Leu Glu Ala Gln Ser Asn Tyr Asn Lys Gln Ile Thr Asp

195

200

205

Glu Phe Ile Glu Asp Tyr Ile Glu Ile Leu Thr Gln Lys Arg Lys Tyr
 210 215 220

Tyr His Gly Pro Gly Asn Glu Lys Ser Arg Thr Asp Tyr Gly Arg Phe
 225 230 235 240

Arg Thr Asp Gly Thr Thr Leu Glu Asn Ile Phe Gly Ile Leu Ile Gly
 245 250 255

Lys Cys Ser Phe Tyr Pro Glu Glu Tyr Arg Ala Ser Lys Ala Ser Tyr
 260 265 270

Thr Ala Gln Glu Phe Asn Phe Leu Asn Asp Leu Asn Asn Leu Lys Val
 275 280 285

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

Asp Phe Ala Lys Lys Ser Lys Ala Leu Gly Ala Ser Lys Leu Leu Lys
 305 310 315 320

Glu Ile Ala Lys Ile Val Asp Cys Ser Val Asp Asp Ile Lys Gly Tyr
 325 330 335

Arg Val Asp Asn Lys Asp Lys Pro Asp Leu His Thr Phe Glu Pro Tyr
 340 345 350

Arg Lys Leu Lys Phe Asn Leu Ser Ser Ile Asp Ile Asp Glu Leu Ser
 355 360 365

Arg Glu Thr Leu Asp Lys Leu Ala Asp Ile Leu Thr Leu Asn Thr Glu
 370 375 380

Arg Glu Gly Ile Glu Asp Thr Ile Lys Arg Asn Leu Pro Ser Gln Phe
 385 390 395 400

Thr Glu Glu Gln Ile Ser Glu Ile Val Gln Ile Arg Lys Asn Gln Ser
 405 410 415

Ser Ala Phe Asn Lys Gly Trp His Ser Phe Ser Ala Lys Leu Met Asn
420 425 430

Glu Leu Ile Pro Glu Leu Tyr Val Thr Ser Glu Glu Gln Met Thr Ile
435 440 445

Leu Thr Arg Leu Glu Lys Phe Lys Val Asn Lys Lys Ser Ser Lys Asn
450 455 460

Thr Lys Thr Ile Asp Glu Lys Glu Ile Thr Asp Glu Ile Tyr Asn Pro
465 470 475 480

Val Val Ala Lys Ser Val Arg Gln Thr Ile Lys Ile Ile Asn Ala Ala
485 490 495

Val Lys Lys Tyr Gly Asp Phe Asp Lys Ile Val Ile Glu Met Pro Arg
500 505 510

Asp Lys Asn Ala Glu Asp Glu Lys Lys Phe Ile Asp Lys Lys Glu Lys
515 520 525

Glu Asn Lys Lys Glu Lys Asp Asp Ser Leu Lys Arg Ala Ala Phe Leu
530 535 540

Tyr Asn Gly Thr Asp Asn Leu Pro Asp Gly Val Phe His Gly Asn Lys
545 550 555 560

Glu Leu Lys Thr Lys Ile Arg Leu Trp Tyr Gln Gln Gly Glu Arg Cys
565 570 575

Leu Tyr Ser Gly Lys Leu Ile Ser Ile His Asp Leu Val His Asn Ser
580 585 590

Asn Lys Phe Glu Ile Asp His Ile Leu Pro Leu Ser Leu Ser Phe Asp
595 600 605

Asp Ser Leu Ala Asn Lys Val Leu Val Tyr Ala Trp Thr Asn Gln Glu
610 615 620

Lys Gly Gln Lys Thr Pro Tyr Gln Val Ile Asp Ser Met Asp Ala Ala

625					630						635				640
Trp	Ser	Phe	Arg	Glu	Met	Lys	Asp	Tyr	Val	Leu	Lys	Gln	Lys	Arg	Leu
				645					650					655	
Gly	Lys	Lys	Lys	Arg	Glu	Tyr	Leu	Leu	Thr	Thr	Glu	Asn	Ile	Asp	Lys
			660					665					670		
Ile	Glu	Val	Lys	Lys	Lys	Phe	Ile	Glu	Arg	Asn	Leu	Val	Asp	Thr	Arg
		675					680					685			
Tyr	Ala	Ser	Arg	Val	Val	Leu	Asn	Ser	Leu	Gln	Thr	Ala	Leu	Lys	Glu
	690					695					700				
Leu	Gly	Lys	Asp	Thr	Lys	Val	Ser	Val	Val	Arg	Gly	Gln	Phe	Thr	Ser
705					710					715					720
Gln	Leu	Arg	Arg	Lys	Trp	Asn	Ile	Asp	Lys	Ser	Arg	Glu	Thr	Tyr	His
				725					730					735	
His	His	Ala	Val	Asp	Ala	Leu	Ile	Ile	Ala	Ala	Ser	Ser	Gln	Leu	Lys
			740					745					750		
Leu	Trp	Gln	Lys	Gln	Glu	Asn	Pro	Met	Phe	Glu	Ser	Tyr	Gly	Glu	Asn
		755					760					765			
Gln	Val	Val	Asn	Lys	Glu	Thr	Gly	Glu	Ile	Leu	Ser	Ile	Ser	Asp	Asp
	770					775					780				
Lys	Tyr	Lys	Glu	Leu	Val	Phe	Gln	Pro	Pro	Tyr	Gln	Gly	Phe	Val	Asn
785					790					795					800
Thr	Ile	Ser	Ser	Lys	Gly	Phe	Glu	Asp	Glu	Ile	Leu	Phe	Ser	Tyr	Gln
				805					810					815	
Val	Asp	Ser	Lys	Phe	Asn	Arg	Lys	Val	Ser	Asp	Ala	Thr	Ile	Tyr	Ser
			820					825					830		
Thr	Arg	Lys	Ala	Lys	Leu	Gly	Lys	Asp	Lys	Lys	Asp	Glu	Thr	Tyr	Val
		835					840					845			

Leu Gly Lys Ile Lys Asp Ile Tyr Ser Gln Asp Gly Phe Asp Thr Phe
850 855 860

Ile Lys Arg Tyr Lys Lys Asp Lys Thr Gln Phe Leu Met Tyr Gln Lys
865 870 875 880

Asp Pro Leu Thr Trp Glu Asn Val Ile Glu Val Ile Leu Arg Asp Tyr
885 890 895

Pro Ser Glu Lys Leu Ser Glu Asp Gly Lys Lys Thr Val Lys Cys Asn
900 905 910

Pro Phe Glu Glu Tyr Arg Arg Glu Asn Gly Leu Ile Cys Lys Tyr Ser
915 920 925

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

Lys Leu Gly Asn Cys Ile Asp Ile Thr Pro Glu Lys Ser Lys Asn Arg
945 950 955 960

Val Val Leu Arg Gln Ile Ser Pro Trp Arg Ala Asp Ile Tyr Phe Asn
965 970 975

Leu Glu Thr Leu Lys Tyr Glu Leu Met Gly Leu Lys Tyr Ser Asp Leu
980 985 990

Ser Phe Glu Lys Gly Thr Gly Lys Tyr His Ile Ser Gln Glu Lys Tyr
995 1000 1005

Asp Ala Ile Arg Glu Lys Glu Gly Ile Gly Lys Lys Ser Glu Phe
1010 1015 1020

Lys Phe Thr Leu Tyr Arg Asn Asp Leu Ile Leu Ile Lys Asp Thr
1025 1030 1035

Leu Asn Asn Cys Glu Arg Met Leu Arg Phe Gly Ser Lys Asn Asp
1040 1045 1050

Thr Ser Lys His Tyr Val Glu Leu Lys Pro Leu Glu Lys Gly Thr

1055

1060

1065

Phe Asp Ser Glu Glu Glu Ile Leu Pro Val Leu Gly Lys Val Ala
 1070 1075 1080

Lys Ser Gly Gln Phe Ile Lys Gly Leu Asn Lys Thr Asn Ile Ser
 1085 1090 1095

Ile Tyr Lys Val Arg Thr Asp Val Leu Gly Asn Lys Phe Phe Ile
 1100 1105 1110

Lys Lys Glu Gly Asp Lys Pro Lys Leu Asp Phe Lys Asn Asn Asn
 1115 1120 1125

Lys

<210> 138
 <211> 1359
 <212> PRT
 <213> Pediococcus sp.

<223> APG06015_RGN

<400> 138

Met Glu Lys Val Pro Tyr Asn Ile Gly Leu Asp Ile Gly Thr Ser Ser
 1 5 10 15

Ile Gly Phe Ala Ala Thr Asp Asn Leu Asn Lys Pro Ile Arg Ala Lys
 20 25 30

Gly Lys Thr Val Ile Gly Val Arg Leu Phe Glu Glu Gly Lys Thr Ala
 35 40 45

Ala Asp Arg Arg Gly Phe Arg Thr Thr Arg Arg Arg Leu Ser Arg Arg
 50 55 60

Lys Trp Arg Leu Arg Leu Leu Asp Glu Ile Phe Asp Lys Glu Met Ala
 65 70 75 80

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

Pro Lys Asp Ala Asn Lys Lys Tyr Leu Gly Ser Leu Leu Phe Pro Glu
100 105 110

Lys Lys Asp Phe Lys Phe Tyr Glu Asp Tyr Pro Thr Ile Tyr His Leu
115 120 125

Arg Tyr Ala Leu Met His Glu Lys Arg Gln Phe Asp Ile Arg Glu Val
130 135 140

Tyr Leu Ala Met His His Ile Ile Lys Tyr Arg Gly Asn Phe Leu Asn
145 150 155 160

Ser Ala Pro Met Asn Ser Phe Lys Thr Gln Asp Phe Asp Phe Val Ala
165 170 175

Lys Phe Lys Lys Leu Asn Glu Leu Phe Glu Ser Ile Asp Ala Glu His
180 185 190

Glu Thr Lys Phe Asp Ile Glu Asn Ile Ser Lys Phe Arg Asp Ile Met
195 200 205

Leu Asn Gln Asp Ile Arg Lys Leu Asp Arg Lys Lys Gln Ala Ala Lys
210 215 220

Ile Leu Ile Leu Asp Ser Thr Asp Lys Thr Ala Lys Lys Ile Asn Asn
225 230 235 240

Lys Ile Ala Thr Ala Val Ala Asn Ser Ala Leu Gly Tyr Lys Phe Ala
245 250 255

Leu Asp Ala Ile Leu Lys Leu Asp Val Glu Glu Ser Lys Asp Trp Ser
260 265 270

Ile Ser Leu Asn Asp Glu Glu Ile Asp Ser Ile Leu Asp Asn Leu Thr
275 280 285

Ser Asp Leu Asp Ala Glu Arg Ile Glu Ile Ile Glu Ile Leu Arg Asp
290 295 300

Leu Tyr Ser His Ile Ala Leu Asn Glu Ile Val Pro Asn Gly Gln Ser
305 310 315 320

Leu Ser Lys Ser Met Met Asp Lys Tyr Asp Lys His His Ala Asp Leu
325 330 335

Asp Val Leu Lys Lys Val Ile Ser Asn Met Asp Asp Arg Lys Lys Ala
340 345 350

Lys Ser Leu Lys Asn Ile Tyr Asn Gln Tyr Val Gly Lys Thr Asn Asp
355 360 365

Lys Val Leu Asp Lys Asp Glu Phe Tyr Lys Gln Ile Gln Lys Asn Leu
370 375 380

Asp Lys Ser Glu Asp Ala Met Lys Ile Val Asn Glu Ile Glu Leu Asp
385 390 395 400

Gln Phe Met Pro Lys Gln Arg Thr Ser Gln Asn Gly Val Ile Pro His
405 410 415

Gln Leu His Gln Lys Glu Leu Asp Glu Ile Ile Glu Asn Gln Lys Gln
420 425 430

Tyr Tyr Pro Phe Leu Ala Glu Pro Asn Pro Asn Glu Lys Arg Lys Pro
435 440 445

His Ala Gln Phe Lys Leu Asp Glu Leu Ile Ala Phe Lys Ile Pro Tyr
450 455 460

Tyr Val Gly Pro Leu Ile Thr Lys Glu Glu Gln Gln Ala Gln Ser Gly
465 470 475 480

Ala Lys Phe Ala Trp Met Lys Arg Lys Gln Asp Gly Val Ile Thr Pro
485 490 495

Trp Asn Phe Asp Glu Lys Val Asp Arg Met Ala Ser Ala Asn Glu Phe
500 505 510

Ile Arg Arg Met Thr Thr Lys Asp Thr Tyr Leu Leu Gly Glu Asp Val
515 520 525

Leu Pro Asp Glu Ser Leu Ile Tyr Gln Lys Phe Lys Val Leu Asn Glu
530 535 540

Leu Asn Asn Val Lys Val Asn Asp Lys Lys Leu Thr Val Ala Asp Lys
545 550 555 560

Gln Asp Ile Phe Asn Asp Leu Phe Lys Lys Gln Lys Thr Val Ser Val
565 570 575

Ser Lys Leu Gln Lys Tyr Phe Val Thr Glu Lys His Tyr Leu Thr Glu
580 585 590

Pro Thr Ile Lys Gly Leu Ser Asp Thr Lys Lys Phe Ser Asn Ser Leu
595 600 605

Ser Thr Tyr Ile Asp Phe Glu Lys Ile Phe Gly Asn Glu Ile Leu Ala
610 615 620

Asp Gln Asn Lys Gln Asn Asp Leu Glu Lys Ile Ile Glu Trp Ser Thr
625 630 635 640

Ile Phe Glu Asp Arg Lys Ile Phe Glu Asp Lys Leu His Glu Ile Glu
645 650 655

Trp Leu Thr Glu Lys Gln Ile Lys Ala Ile Arg Arg Tyr Arg Gly Trp
660 665 670

Gly Arg Leu Ser Lys Lys Leu Leu Val Asp Leu Arg Asn Asn Glu Gly
675 680 685

Lys Ser Ile Leu Asp Glu Leu Trp Arg Thr Asn Asp Asn Phe Met Gln
690 695 700

Ile Gln Ala Arg Glu Glu Phe Ala Lys Ser Ile Val Glu Ala Asn Gln
705 710 715 720

Lys Leu Met Asn Leu Gly Gly Ala Ala Ser Val Gln Asn Thr Val Glu
725 730 735

Ser Thr Leu Glu Asp Ala Tyr Thr Ser Pro Gln Asn Lys Lys Ala Ile
740 745 750

Arg Gln Val Ile Lys Val Val Glu Asp Ile Val Lys Ala Val Gly Tyr
755 760 765

Ala Pro Glu Lys Ile Thr Ile Glu Phe Thr Arg Gly Ala Asp Lys Asn
770 775 780

Ser Arg Arg Thr Gln Asn Arg Gln Lys Gln Ile Met Glu Val Tyr Lys
785 790 795 800

Thr Ala Ala Lys Glu Ile Val Asp Ala Thr Leu Lys Gly Gln Leu Glu
805 810 815

Asn Glu Glu Lys Leu Thr Asp Lys Leu Tyr Leu Tyr Phe Thr Gln Leu
820 825 830

Gly Lys Asp Ile Tyr Ser Gly Glu Thr Ile Asn Ile Asp Gln Leu Asn
835 840 845

Asn Tyr Asp Ile Asp His Ile Leu Pro Gln Ala Phe Ile Lys Asp Asp
850 855 860

Ser Leu Asp Asn Arg Val Leu Thr Ser Arg Asp Leu Asn Asn Gly Lys
865 870 875 880

Ser Asp Ser Val Pro Val Lys Gln Phe Gly Ala Asn Met Lys Ser Phe
885 890 895

Trp Met Arg Leu Gln Ala His Gly Leu Ile Ser Lys Arg Lys Leu Asn
900 905 910

Asn Leu Met Thr Asp Pro Asp Ser Ile Gly Lys Tyr Thr Met Gln Gly
915 920 925

Phe Val Arg Arg Gln Leu Val Glu Thr Ser Gln Val Ile Lys Leu Thr
930 935 940

Ala Asn Ile Leu Gly Ala Val Tyr Gly Glu Asn Thr Asp Ile Val Glu
945 950 955 960

Ile Pro Ala Lys Leu Thr His Gln Met Arg Glu Lys Phe Asn Leu Tyr
965 970 975

Lys Val Arg Glu Val Asn Asp Tyr His His Ala Phe Asp Ala Tyr Leu
980 985 990

Thr Thr Phe Val Gly Asn Tyr Leu Phe Lys Arg Tyr Pro Lys Leu Arg
995 1000 1005

Pro Tyr Phe Val Tyr Gly Asp Phe Lys Met Thr Asp Asn Ala Leu
1010 1015 1020

Lys Gly Met Arg Arg Phe Asn Phe Leu His Asp Leu Lys Asp Asp
1025 1030 1035

Glu Val Leu Val Asp Asn Glu Thr Gly Glu Val Leu Trp Glu Gly
1040 1045 1050

Gln Lys Ser Ile Glu Glu Leu Lys Lys Ile Tyr Gly Tyr Lys Phe
1055 1060 1065

Met Leu Thr Thr His Glu Ala Tyr Ile Asn Asn Gly Ala Met Tyr
1070 1075 1080

Asp Gln Thr Val Tyr Pro Ala Lys Ser Asn Lys Lys Leu Ile Pro
1085 1090 1095

Ser Lys Asn Asn Arg Pro Thr Asn Ile Tyr Gly Gly Phe Ser Gly
1100 1105 1110

Asn Lys Asp Ala Tyr Leu Ala Ile Val Lys Ile Leu Gly Lys Lys
1115 1120 1125

Gly Asp Thr Tyr Lys Val Val Gly Val Pro Arg Arg Phe Ser Asp
1130 1135 1140

Glu Leu Leu Asn Ala Tyr Asn Asp Asp Tyr Ser Glu Tyr Leu Asn
1145 1150 1155

Lys Leu His Asp Val Leu Val Thr Gln Phe Thr Lys Lys Gln Arg
1160 1165 1170

Asn Lys Lys Thr Gly Lys Ile Ile Lys Lys Ile Glu Asn Phe Glu
1175 1180 1185

Ile Val Val Gly Lys Val Asn Tyr Arg Gln Arg Ile Ile Asp Arg
1190 1195 1200

Asn Gln Ser Phe Met Leu Gly Ser Ser Thr Tyr Lys Tyr Asn Thr
1205 1210 1215

Lys Gln Leu Val Leu Thr Glu Ser Ser Met Lys Ile Leu Ser Asn
1220 1225 1230

Phe Lys His Ala Ser Asp Asp Asn Leu Asn Leu Val Tyr Asp Glu
1235 1240 1245

Ile Leu Glu Lys Val Asn Glu Ser Phe Glu Leu Tyr Asp Lys Asn
1250 1255 1260

Gly Phe Arg Lys Lys Leu Asn Asn Asn Ile Asp Lys Phe Arg Lys
1265 1270 1275

Leu Pro Cys Lys Ser Ile Phe Glu Lys Asn Lys Leu Ile Ser Val
1280 1285 1290

Gly Lys Cys Glu Ile Leu Phe Gln Ile Leu Asn Gly Leu His Ala
1295 1300 1305

Asn Ala Thr Met Gly Asp Leu Lys Gly Ile Gly Phe Ser Thr Pro
1310 1315 1320

Phe Gly Phe Met Gln Asp Ala Ser Gly Ile Lys Leu Asn Lys Asp
1325 1330 1335

Ser Met Leu Leu Tyr Gln Ser Pro Thr Gly Leu Phe Glu Arg Lys
1340 1345 1350

Val Lys Leu Ser Asp Leu
1355

<210> 139
<211> 1488
<212> PRT
<213> Weissella sp.

<223> APG09344_RGN

<400> 139

Met Ala Lys Gln Lys Tyr Thr Val Gly Leu Asp Ile Gly Thr Asn Ser
1 5 10 15

Val Gly Trp Ala Val Ile Asn Lys Asp Phe Lys Leu Val Ser Gly Lys
20 25 30

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

Trp Gly Val Arg Leu Phe Asp Ala Ala Asp Thr Ala Ala Asp Arg Arg
50 55 60

Leu Lys Arg Gly Met Arg Arg Arg Ile Ala Arg Arg Lys Glu Arg Leu
65 70 75 80

Asn Tyr Leu Arg Gly Ile Phe Gln Ala Asp Ile Leu Ala Phe Asp Asp
85 90 95

Ser Phe Phe Ile Arg Met Asp Glu Ser Phe Leu Gln Asp Asp Asp Lys
100 105 110

Gln Ala Thr Ala Phe Ser Tyr Arg Asp Gln Ser Gly Ala Met Gln Thr
115 120 125

Arg Val Val Thr Glu Lys Glu Ala Val Lys Tyr Pro Leu Phe Asn Gly
130 135 140

Lys Ser Gly Ser Gly Glu Thr Tyr Thr Asp Glu Ser Ala Tyr Tyr Asn
145 150 155 160

Gln Tyr Pro Thr Ile Tyr His Leu Arg Gln Arg Leu Ile Asp Asn Pro
165 170 175

Ala Gln Ala Asp Leu Arg Leu Val Tyr Leu Ala Met His His Ile Leu
180 185 190

Lys Tyr Arg Gly His Phe Val Asn Gln Gly Gln Thr Phe Asp Leu Gly
195 200 205

Asn Ile Asn Ile Ala Gly Asn Leu Thr Ser Ala Leu Asp Lys Leu Asp
210 215 220

Glu Ala Ser Ser Phe Lys Phe Gly Leu Ala Asp Val Asp Ala Asp Gln
225 230 235 240

Ala Asn Ala Ile Leu Lys Asn Lys Gln Trp Ser Ala Ser Lys Lys Ala
245 250 255

Tyr Glu Leu Asn Thr Leu Tyr Glu Ile Ile Gln Asp Thr Val Tyr Arg
260 265 270

Arg Glu Asn Glu Thr Gln Ala Val Ala Tyr Asp Ser Leu Thr Val Lys
275 280 285

Gln Gln Glu Lys Trp Leu Ala Glu Lys Gln Lys Gln Val Lys Ala Phe
290 295 300

Leu Thr Gly Ile Val Gly Asn Thr Ile Ser Pro Lys Asp Ile Phe Ala
305 310 315 320

Asn Lys Asp Tyr Asp Arg Lys Gln Asn Glu Asp Phe Pro Glu Lys Ile
325 330 335

Lys Tyr Ser Asn Glu Asn Phe Glu Glu Gln Ile Gly Glu Leu Glu Lys
340 345 350

Tyr Leu Asp Glu Thr Glu Val Ala Ala Ile Val Ala Gly Lys Asp Val
355 360 365

Tyr Glu Ser Leu Val Leu Ser Asn Ile Leu Thr Lys Ser Thr Leu Ser
370 375 380

Gly Ser Met Ile Glu Lys Tyr Asn Val His Gly Thr Gln Leu Lys Ser

385		390		395		400									
Leu	Lys	Ser	Phe	Ser	Arg	Lys	Val	Ser	Ala	Glu	Phe	Tyr	Asp	Lys	Leu
				405					410					415	
Phe	Asn	Val	Val	Lys	Asp	Glu	Lys	Gly	Lys	Pro	Gln	Ile	Gly	Val	Tyr
			420					425					430		
Ala	Gln	Tyr	Ile	Asp	Gly	Val	Gly	Asn	Pro	Ala	Lys	Arg	Leu	Ala	Arg
		435					440					445			
Glu	Asp	Phe	Tyr	Asp	Ala	Leu	Lys	Lys	Val	Phe	Glu	Ser	Glu	Phe	Lys
	450					455					460				
Gly	Leu	Thr	Phe	Pro	Ala	Gly	Glu	Lys	Gly	Ile	Asp	Phe	Thr	Lys	Thr
465					470					475					480
Asp	Leu	Ser	Ala	Asp	Arg	Met	Ala	Phe	Ile	Arg	Glu	Met	Asn	Glu	Ala
				485					490					495	
Ile	Asn	Leu	Glu	Asn	Tyr	Leu	Pro	Lys	Gln	Arg	Gln	Ala	Asp	Asn	Gly
			500					505					510		
Ala	Ile	Pro	Tyr	Gln	Val	His	Glu	His	Glu	Leu	Ile	Arg	Ile	Ile	Glu
		515					520					525			
Asn	Gln	Lys	Gln	Tyr	Tyr	Pro	Phe	Leu	Gly	Asp	Lys	Val	Gln	Val	Thr
	530					535					540				
Tyr	Glu	Asp	Glu	Ser	Gly	Gln	Thr	Gln	Val	Arg	Ser	Glu	Tyr	Lys	Ile
545					550					555					560
Gln	Val	Leu	Phe	Lys	Phe	Arg	Ile	Pro	Tyr	Tyr	Val	Gly	Thr	Leu	Ala
				565					570					575	
Lys	Asn	Ser	Gly	Trp	Val	Asn	Glu	Asp	Gly	Lys	Lys	Leu	Val	Ala	Arg
			580					585					590		
Asn	Val	Ala	Ala	Lys	Asn	Ser	Trp	Val	Val	Arg	Asn	Ser	Asp	Glu	Arg
		595					600					605			

Val Thr Pro Trp Asn Phe Glu Gln Val Ile Asn Lys Glu Ala Ser Ser
610 615 620

Ile Asn Phe Ile Glu Arg Met Thr Asn Phe Asp Thr Tyr Leu Pro Asn
625 630 635 640

Glu Lys Val Leu Pro Lys Asn Ser Leu Leu Tyr Gln Glu Phe Ile Val
645 650 655

Asn Asn Glu Leu Ile Ser Ser Gly Tyr Tyr Leu Asn Gly Arg Lys Glu
660 665 670

Tyr Phe Leu Pro Glu Gln Arg Gln Arg Ile Ile Thr Arg Leu Phe Lys
675 680 685

Arg Tyr Arg Lys Val Ser Ala Lys Gln Met Leu Asp Phe Leu Lys Asn
690 695 700

Glu Tyr Gln Ile Lys Leu Asp Asp Asp Pro Lys Lys Leu Phe Gly Leu
705 710 715 720

Asp Thr Phe Val Lys Val Pro Ser Tyr Asn Gly Ser Leu Ser Thr Val
725 730 735

Val Asp Leu Thr Asn Lys Met Gly Ile Ser Glu Asp Leu Ile Ala Asn
740 745 750

Tyr Pro Glu Arg Phe Glu Glu Ile Ile Lys Trp Gln Thr Ile Phe Glu
755 760 765

Asp Lys Lys Ile Leu Lys Lys Thr Ile Arg Glu Val Asn His Asn Lys
770 775 780

Trp Ser Gly Met Leu Thr Ala Glu Gln Val Asn Lys Leu Ser Lys Arg
785 790 795 800

His Tyr Thr Gly Trp Gly Arg Leu Ser Gln Lys Leu Leu Asp Gly Val
805 810 815

Lys Thr Ser Gln Gly Arg Thr Ile Ile Glu Ser Leu Arg Glu Gly Ile

820

825

830

Cys Asp Asn Phe Met Arg Leu Ile Glu Asp Glu Lys Ile Ala Glu Phe
835 840 845

Ile Ala Asn Ala Gln Val Ala Gly Ala Lys Asp Gly Val Leu Ala Tyr
850 855 860

Asp Leu Val Asp Gln Leu Pro Gly Ser Pro Ala Ile Lys Lys Gly Ile
865 870 875 880

Trp Gln Ser Leu Arg Val Leu Gln Glu Leu Glu His Tyr Leu Gly Arg
885 890 895

Asp Ala Ile Gly Lys Val Val Ile Glu Met Ser Arg Asp Asp Glu Ser
900 905 910

Ser Arg Arg Thr Lys Ala Arg Lys Gln Gln Leu Glu Glu Phe Tyr Lys
915 920 925

Lys Phe Lys Glu Ser Thr Gly Ser Asp Val Thr Ala Glu Leu Lys Arg
930 935 940

Glu Leu Ser Glu Lys Asp Ala Lys Glu Ile Asp Asp Glu Lys Val Phe
945 950 955 960

Leu Tyr Phe Leu Gln Asn Gly Lys Ser Met Tyr Ser Gly Asp Gln Leu
965 970 975

Tyr Leu Ser Arg Leu Ala Asp Tyr Gln Val Asp His Ile Val Pro Gln
980 985 990

Thr Tyr Ile Lys Asp Asp Ser Leu Asp Asn Lys Val Leu Val Thr Gln
995 1000 1005

Arg Asp Asn Gln Asn Lys Gly Gly Asp Thr Pro Ser His Thr Val
1010 1015 1020

Val Thr Arg Met Asn Ser Tyr Trp Asp Met Leu Ala Lys Asn Gly
1025 1030 1035

Gln Val Ser Pro Arg Lys Leu Ala Asn Leu Lys Lys Gly Thr Ile
1040 1045 1050

Thr Leu Lys Gln Arg Glu Gly Phe Ile Asn Arg Gln Leu Val Glu
1055 1060 1065

Asn Arg Gln Ile Ile Lys His Val Ala Asn Ile Leu Thr Asn Tyr
1070 1075 1080

Tyr Thr Gly Thr Asp Thr Leu Val Leu Thr Pro Lys Ser Gly Leu
1085 1090 1095

Thr Ser Gln Leu Arg Ser Gly Leu Ile Tyr Glu Leu Asn Pro Ala
1100 1105 1110

Phe Asp Glu Ala Lys Ala Leu Asp Arg Gly Lys Ala Tyr Gln Val
1115 1120 1125

Glu Arg Tyr Thr Thr Val Lys Leu His Asp Arg Phe Val Lys Asn
1130 1135 1140

Arg Lys Leu Asn Asp Tyr His His Ala His Asp Ala Tyr Leu Asn
1145 1150 1155

Ala Phe Val Ala Gln Tyr Val Tyr Gln Glu His Pro Glu Trp Lys
1160 1165 1170

Asn Ala Trp Val Tyr Gly Lys Tyr Pro Arg Asn Gly Gln Ala Asp
1175 1180 1185

Phe Gly Lys Trp Ala Thr Gln Arg Lys Gln Lys Ser Leu Gln Leu
1190 1195 1200

Leu Ser Ser Met Ala Asn Asp Val Trp Asn Leu Glu Asp Pro Asp
1205 1210 1215

Thr Met Glu Lys Thr Ile Leu Asn Arg Asp Glu Thr Phe Glu Gln
1220 1225 1230

Met Arg Gln Thr Leu Cys Tyr Arg Asn Ile Asn Ile Val Lys Lys

1235						1240						1245			
Leu	Glu	Thr	Gln	Thr	Gly	Lys	Phe	Gly	Asp	Glu	Ser	Val	Tyr	Lys	
1250						1255					1260				
Lys	Gly	Asn	Lys	Ala	Asp	His	Tyr	Ser	Leu	Gly	Leu	Lys	Arg	Lys	
1265						1270					1275				
Tyr	Pro	Pro	Gln	Lys	Tyr	Gly	Gly	Thr	Lys	Gly	Ala	Ile	Ser	Ala	
1280						1285					1290				
Val	Thr	Ala	Leu	Val	Lys	Asp	Asn	Ser	Gly	Arg	Val	Ile	Pro	Val	
1295						1300					1305				
Ser	Ile	Ser	Ala	Ser	Asp	Tyr	Asp	Ser	Tyr	Gln	Asn	Ala	Ser	Asp	
1310						1315					1320				
Lys	Leu	Ala	Trp	Leu	Leu	Met	Arg	Lys	Lys	Asn	Ile	Ala	Glu	Ile	
1325						1330					1335				
Leu	Val	Ala	Ser	Leu	Pro	Lys	Tyr	Thr	Lys	Tyr	Glu	Leu	Pro	Ser	
1340						1345					1350				
Gly	Ala	Leu	Arg	Leu	Leu	Ala	Ser	Tyr	Gln	Glu	Ala	Gln	Ser	Ala	
1355						1360					1365				
Val	Glu	Leu	Pro	Met	Leu	Asn	Leu	Gly	Leu	Met	Asp	Asp	Glu	Ala	
1370						1375					1380				
Lys	Val	Leu	Asn	Leu	Tyr	Asp	Gln	Leu	Ala	Gln	Phe	Ile	Ala	Asp	
1385						1390					1395				
Asn	Lys	Leu	Phe	Thr	Asp	Lys	Lys	Ile	Ala	Leu	Leu	Lys	Gly	Glu	
1400						1405					1410				
Met	Arg	Val	Ala	Phe	Glu	Gly	Leu	Glu	Glu	Phe	Lys	Asp	Lys	Gln	
1415						1420					1425				
Lys	Val	Ile	Asp	Glu	Leu	Leu	Gly	Val	Thr	Asn	Gly	Ser	Asn	Gln	
1430						1435					1440				

Gly Leu Gln Ala Leu Asn Ser Ile Gly Leu Gly Thr Ser Asn Gln
1445 1450 1455

Arg Leu Lys Ser Gly Asn Thr Met Thr Asn Gly Val Thr Leu Ile
1460 1465 1470

Asn Glu Ser Val Thr Gly Leu Tyr Ser Thr Arg Thr Thr Tyr Asn
1475 1480 1485

<210> 140
<211> 1337
<212> PRT
<213> Enterococcus sp.

<223> APG07991_RGN

<400> 140

Met Lys Lys Asp Tyr Val Ile Gly Leu Asp Ile Gly Thr Asn Ser Val
1 5 10 15

Gly Trp Ala Val Met Thr Glu Asp Tyr Gln Leu Val Lys Lys Lys Met
20 25 30

Pro Ile Tyr Gly Asn Thr Glu Lys Lys Lys Ile Lys Lys Asn Phe Trp
35 40 45

Gly Val Arg Leu Phe Glu Glu Gly His Thr Ala Glu Asp Arg Arg Leu
50 55 60

Lys Arg Thr Ala Arg Arg Arg Ile Ser Arg Arg Arg Asn Arg Leu Arg
65 70 75 80

Tyr Leu Gln Ala Phe Phe Glu Glu Ala Met Thr Ala Leu Asp Glu Asn
85 90 95

Phe Phe Ala Arg Leu Gln Glu Ser Phe Leu Val Pro Glu Asp Lys Lys
100 105 110

Trp His Arg His Pro Ile Phe Ala Lys Leu Glu Asp Glu Val Ala Tyr
115 120 125

His Glu Thr Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Ala Asp
130 135 140

Ser Ser Glu Gln Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145 150 155 160

Ile Val Lys Tyr Arg Gly His Phe Leu Ile Glu Gly Lys Leu Ser Thr
165 170 175

Glu Asn Ile Ser Val Lys Glu Gln Phe Gln Gln Phe Met Ile Ile Tyr
180 185 190

Asn Gln Thr Phe Val Asn Gly Glu Ser Arg Leu Val Ser Ala Pro Leu
195 200 205

Pro Glu Ser Val Leu Ile Glu Glu Glu Leu Thr Glu Lys Ala Ser Arg
210 215 220

Thr Lys Lys Ser Glu Lys Val Leu Gln Gln Phe Pro Gln Glu Lys Ala
225 230 235 240

Asn Gly Leu Phe Gly Gln Phe Leu Lys Leu Met Val Gly Asn Lys Ala
245 250 255

Asp Phe Lys Lys Val Phe Gly Leu Glu Glu Glu Ala Lys Ile Thr Tyr
260 265 270

Ala Ser Glu Ser Tyr Glu Glu Asp Leu Glu Gly Ile Leu Ala Lys Val
275 280 285

Gly Asp Glu Tyr Ser Asp Val Phe Leu Ala Ala Lys Asn Val Tyr Asp
290 295 300

Ala Val Glu Leu Ser Thr Ile Leu Ala Asp Ser Asp Lys Lys Ser His
305 310 315 320

Ala Lys Leu Ser Ser Ser Met Ile Val Arg Phe Thr Glu His Gln Glu
325 330 335

Asp Leu Lys Asn Phe Lys Arg Phe Ile Arg Glu Asn Cys Pro Asp Glu
340 345 350

Tyr Asp Asn Leu Phe Lys Asn Glu Gln Lys Asp Gly Tyr Ala Gly Tyr
355 360 365

Ile Ala His Ala Gly Lys Val Ser Gln Leu Lys Phe Tyr Gln Tyr Val
370 375 380

Lys Lys Ile Ile Gln Asp Ile Ala Gly Ala Glu Tyr Phe Leu Glu Lys
385 390 395 400

Ile Ala Gln Glu Asn Phe Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly
405 410 415

Val Ile Pro His Gln Ile His Leu Ala Glu Leu Gln Ala Ile Ile His
420 425 430

Arg Gln Ala Ala Tyr Tyr Pro Phe Leu Lys Glu Asn Gln Lys Lys Ile
435 440 445

Glu Gln Leu Val Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ser
450 455 460

Lys Gly Asp Ala Ser Thr Phe Ala Trp Leu Lys Arg Gln Ser Glu Glu
465 470 475 480

Pro Ile Arg Pro Trp Asn Leu Gln Glu Thr Val Asp Leu Asp Gln Ser
485 490 495

Ala Thr Ala Phe Ile Glu Arg Met Thr Asn Phe Asp Thr Tyr Leu Pro
500 505 510

Ser Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Lys Phe Met
515 520 525

Val Phe Asn Glu Leu Thr Lys Ile Ser Tyr Thr Asp Asp Arg Gly Ile
530 535 540

Lys Ala Asn Phe Ser Gly Lys Glu Lys Glu Lys Ile Phe Asp Tyr Leu
545 550 555 560

Phe Lys Thr Arg Arg Lys Val Lys Lys Lys Asp Ile Ile Gln Phe Tyr
565 570 575

Arg Asn Glu Tyr Asn Thr Glu Ile Val Thr Leu Ser Gly Leu Glu Glu
580 585 590

Asp Gln Phe Asn Ala Ser Phe Ser Thr Tyr Gln Asp Leu Leu Lys Cys
595 600 605

Gly Leu Thr Arg Ala Glu Leu Asp His Pro Asp Asn Ala Glu Lys Leu
610 615 620

Glu Asp Ile Ile Lys Ile Leu Thr Ile Phe Glu Asp Arg Gln Arg Ile
625 630 635 640

Arg Thr Gln Leu Ser Thr Phe Lys Gly Gln Phe Ser Glu Glu Val Leu
645 650 655

Lys Lys Leu Glu Arg Lys His Tyr Thr Gly Trp Gly Arg Leu Ser Lys
660 665 670

Lys Leu Ile Asn Gly Ile Tyr Asp Lys Glu Ser Gly Lys Thr Ile Leu
675 680 685

Asp Tyr Leu Ile Lys Asp Asp Gly Val Ser Lys His Tyr Asn Arg Asn
690 695 700

Phe Met Gln Leu Ile Asn Asp Ser Gln Leu Ser Phe Lys Asn Ala Ile
705 710 715 720

Gln Lys Ala Gln Ser Ser Glu His Glu Glu Thr Leu Ser Glu Thr Val
725 730 735

Asn Glu Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile Tyr Gln Ser
740 745 750

Leu Lys Ile Val Asp Glu Leu Val Ala Ile Met Gly Tyr Ala Pro Lys
755 760 765

Arg Ile Val Val Glu Met Ala Arg Glu Asn Gln Thr Thr Ser Thr Gly
770 775 780

Lys Arg Arg Ser Ile Gln Arg Leu Lys Ile Val Glu Lys Ala Met Ala
785 790 795 800

Glu Ile Gly Ser Asn Leu Leu Lys Glu Gln Pro Thr Thr Asn Glu Gln
805 810 815

Leu Arg Asp Thr Arg Leu Phe Leu Tyr Tyr Met Gln Asn Gly Lys Asp
820 825 830

Met Tyr Thr Gly Asp Glu Leu Ser Leu His Arg Leu Ser His Tyr Asp
835 840 845

Ile Asp His Ile Ile Pro Gln Ser Phe Met Lys Asp Asp Ser Leu Asp
850 855 860

Asn Leu Val Leu Val Gly Ser Thr Glu Asn Arg Gly Lys Ser Asp Asp
865 870 875 880

Val Pro Ser Lys Glu Val Val Lys Lys Met Lys Ala Tyr Trp Glu Lys
885 890 895

Leu Tyr Ala Ala Gly Leu Ile Ser Gln Arg Lys Phe Gln Arg Leu Thr
900 905 910

Lys Gly Glu Gln Gly Gly Leu Thr Leu Glu Asp Lys Ala His Phe Ile
915 920 925

Gln Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys Asn Val Ala Gly
930 935 940

Ile Leu Asp Gln Arg Tyr Asn Ala Lys Ser Lys Glu Lys Lys Val Gln
945 950 955 960

Ile Ile Thr Leu Lys Ala Ser Leu Thr Ser Gln Phe Arg Ser Ile Phe
965 970 975

Gly Leu Tyr Lys Val Arg Glu Val Asn Asp Tyr His His Gly Gln Asp
980 985 990

Ala Tyr Leu Asn Cys Val Val Ala Thr Thr Leu Leu Lys Val Tyr Pro
995 1000 1005

Asn Leu Ala Pro Glu Phe Val Tyr Gly Glu Tyr Pro Lys Phe Gln
1010 1015 1020

Ala Phe Lys Glu Asn Lys Ala Thr Ala Lys Ala Ile Ile Tyr Thr
1025 1030 1035

Asn Leu Leu Arg Phe Phe Thr Glu Asp Glu Pro Arg Phe Thr Lys
1040 1045 1050

Asp Gly Glu Ile Leu Trp Ser Asn Ser Tyr Leu Lys Thr Ile Lys
1055 1060 1065

Lys Glu Leu Asn Tyr His Gln Met Asn Ile Val Lys Lys Val Glu
1070 1075 1080

Val Gln Lys Gly Gly Phe Ser Lys Glu Ser Ile Lys Pro Lys Gly
1085 1090 1095

Pro Ser Asn Lys Leu Ile Pro Val Lys Asn Gly Leu Asp Pro Gln
1100 1105 1110

Lys Tyr Gly Gly Phe Asp Ser Pro Val Val Ala Tyr Thr Val Leu
1115 1120 1125

Phe Thr His Glu Lys Gly Lys Lys Pro Leu Ile Lys Gln Glu Ile
1130 1135 1140

Leu Gly Ile Thr Ile Met Glu Lys Thr Arg Phe Glu Gln Asn Pro
1145 1150 1155

Ile Leu Phe Leu Glu Glu Lys Gly Phe Leu Arg Pro Arg Val Leu
1160 1165 1170

Met Lys Leu Pro Lys Tyr Thr Leu Tyr Glu Phe Pro Glu Gly Arg
1175 1180 1185

Arg Arg Leu Leu Ala Ser Ala Lys Glu Ala Gln Lys Gly Asn Gln
1190 1195 1200

Met Val Leu Pro Glu His Leu Leu Thr Leu Leu Tyr His Ala Lys
1205 1210 1215

Gln Cys Leu Leu Pro Asn Gln Ser Glu Ser Leu Ala Tyr Val Glu
1220 1225 1230

Gln His Gln Pro Glu Phe Gln Glu Ile Leu Glu Arg Val Val Asp
1235 1240 1245

Phe Ala Glu Val His Thr Leu Ala Lys Ser Lys Val Gln Gln Ile
1250 1255 1260

Val Lys Leu Phe Glu Ala Asn Gln Thr Ala Asp Val Lys Glu Ile
1265 1270 1275

Ala Ala Ser Phe Ile Gln Leu Met Gln Phe Asn Ala Met Gly Ala
1280 1285 1290

Pro Ser Thr Phe Lys Phe Phe Gln Lys Asp Ile Glu Arg Ala Arg
1295 1300 1305

Tyr Thr Ser Ile Lys Glu Ile Phe Asp Ala Thr Ile Ile Tyr Gln
1310 1315 1320

Ser Thr Thr Gly Leu Tyr Glu Thr Arg Arg Lys Val Val Asp
1325 1330 1335

<210> 141
<211> 1340
<212> PRT
<213> Enterococcus sp.

<223> APG01868_RGN

<400> 141

Met Lys Lys Glu Tyr Thr Ile Gly Leu Asp Ile Gly Thr Asn Ser Val
1 5 10 15

Gly Trp Ser Val Leu Thr Asp Asp Tyr Arg Leu Val Ser Lys Lys Met
20 25 30

Lys Val Ala Gly Asn Thr Glu Lys Ser Ser Thr Lys Lys Asn Phe Trp
35 40 45

Gly Val Arg Leu Phe Asp Glu Gly Gln Thr Ala Glu Ala Arg Arg Ser
50 55 60

Lys Arg Thr Ala Arg Arg Arg Leu Ala Arg Arg Arg Gln Arg Ile Leu
65 70 75 80

Glu Leu Gln Lys Ile Phe Ala Pro Glu Ile Leu Lys Ile Asp Glu His
85 90 95

Phe Phe Ala Arg Leu Asn Glu Ser Phe Leu Val Pro Asp Glu Lys Lys
100 105 110

Gln Ser Arg His Pro Val Phe Ala Thr Ile Lys Gln Glu Lys Ser Tyr
115 120 125

His Gln Thr Tyr Pro Thr Ile Tyr His Leu Arg Gln Ala Leu Ala Asp
130 135 140

Ser Ser Glu Lys Ala Asp Ile Arg Leu Val Tyr Leu Ala Met Ala His
145 150 155 160

Leu Leu Lys Tyr Arg Gly His Phe Leu Ile Glu Gly Glu Leu Asn Thr
165 170 175

Glu Asn Ser Ser Val Thr Glu Thr Phe Arg Gln Phe Leu Ser Thr Tyr
180 185 190

Asn Gln Gln Phe Ser Glu Ala Gly Asp Lys Gln Thr Glu Lys Leu Asp
195 200 205

Glu Ala Val Asp Cys Ser Phe Val Phe Thr Glu Lys Met Ser Lys Thr
210 215 220

Lys Lys Ala Glu Thr Leu Leu Lys Tyr Phe Pro His Glu Lys Ser Asn
225 230 235 240

Gly Tyr Leu Ser Gln Phe Ile Lys Leu Met Val Gly Asn Gln Gly Asn

245

250

255

Phe Lys Asn Val Phe Gly Leu Glu Glu Glu Ala Lys Leu Gln Phe Ser
260 265 270

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

Asp Asp Tyr Ile Asp Leu Phe Val Gln Ala Lys Asn Val Tyr Asp Ala
290 295 300

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

Lys Leu Ser Ala Gly Met Ile Arg Arg Tyr Asp Ala His Lys Glu Asp
325 330 335

Leu Val Leu Leu Lys Arg Phe Val Lys Glu Asn Leu Pro Lys Lys Tyr
340 345 350

Arg Ala Phe Phe Gly Asp Asn Ser Val Asn Gly Tyr Ala Gly Tyr Ile
355 360 365

Glu Gly His Ala Thr Gln Glu Asp Phe Tyr Lys Phe Val Lys Lys Glu
370 375 380

Leu Thr Gly Ile Arg Gly Ser Glu Val Phe Leu Thr Lys Ile Glu Gln
385 390 395 400

Glu Asn Phe Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Val Ile Pro
405 410 415

His Gln Ile His Leu Thr Glu Leu Arg Ala Ile Ile Ala Asn Gln Lys
420 425 430

Lys His Tyr Pro Phe Leu Lys Glu Glu Gln Glu Lys Leu Glu Ser Leu
435 440 445

Leu Thr Phe Lys Ile Pro Tyr Tyr Val Gly Pro Leu Ala Lys Lys Gln
450 455 460

Glu Asn Ser Pro Phe Ala Trp Leu Ile Arg Lys Ser Glu Glu Lys Ile
465 470 475 480

Lys Pro Trp Asn Leu Pro Glu Ile Val Asp Met Glu Asp Ser Ala Val
485 490 495

Arg Phe Ile Glu Arg Met Ile Asn Thr Asp Met Tyr Met Pro His Asn
500 505 510

Lys Val Leu Pro Lys Asn Ser Leu Leu Tyr Gln Lys Phe Ser Ile Tyr
515 520 525

Asn Glu Leu Thr Lys Val Arg Tyr Gln Asp Glu Arg Gly Gln Met Asn
530 535 540

Tyr Phe Ser Ser Ile Glu Lys Lys Glu Ile Phe His Glu Leu Phe Glu
545 550 555 560

Lys Asn Arg Lys Val Thr Lys Lys Asp Leu Gln Glu Phe Leu Tyr Leu
565 570 575

Lys Tyr Asp Ile Lys His Ala Glu Leu Ser Gly Ile Glu Lys Ala Phe
580 585 590

Asn Ala Ser Tyr Thr Thr Tyr His Asp Phe Leu Thr Met Ser Glu Asn
595 600 605

Lys Arg Glu Met Lys Gln Trp Leu Glu Asp Pro Glu Leu Ala Ser Met
610 615 620

Phe Glu Glu Ile Ile Lys Thr Leu Thr Val Phe Glu Asp Arg Glu Met
625 630 635 640

Ile Lys Thr Arg Leu Ser His His Glu Ala Thr Leu Gly Lys His Ile
645 650 655

Ile Lys Lys Leu Thr Lys Lys His Tyr Thr Gly Trp Gly Arg Leu Ser
660 665 670

Lys Glu Leu Ile Gln Gly Ile Arg Asp Lys Gln Ser Asn Lys Thr Ile

675

680

685

Leu Asp Tyr Leu Ile Asn Asp Asp Asp Phe Pro His His Arg Asn Arg
 690 695 700

Asn Phe Met Gln Leu Ile Asn Asp Asp Ser Leu Ser Phe Lys Lys Glu
 705 710 715 720

Ile Lys Lys Ala Gln Met Ile Thr Asp Thr Glu Asn Leu Glu Glu Ile
 725 730 735

Val Lys Glu Leu Thr Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln
 740 745 750

Ser Leu Lys Ile Val Asp Glu Ile Val Gly Ile Met Gly Tyr Glu Pro
 755 760 765

Ala Asn Ile Val Val Glu Met Ala Arg Glu Asn Gln Thr Thr Gly Arg
 770 775 780

Gly Leu Lys Ser Ser Arg Pro Arg Leu Lys Ala Leu Glu Glu Ser Leu
 785 790 795 800

Lys Asp Phe Gly Ser Gln Leu Leu Lys Glu Tyr Pro Thr Asp Asn Ser
 805 810 815

Ser Leu Gln Lys Asp Arg Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg
 820 825 830

Asp Met Tyr Thr Gly Ala Pro Leu Asp Ile His Arg Leu Ser Asp Tyr
 835 840 845

Asp Ile Asp His Ile Ile Pro Arg Ser Phe Thr Thr Asp Asn Ser Ile
 850 855 860

Asp Asn Lys Val Leu Val Ser Ser Lys Glu Asn Arg Leu Lys Lys Asp
 865 870 875 880

Asp Val Pro Ser Glu Lys Val Val Lys Lys Met Arg Ser Phe Trp Tyr
 885 890 895

Asp Leu Tyr Ser Ser Lys Leu Ile Ser Lys Arg Lys Leu Asp Asn Leu
900 905 910

Thr Lys Ile Lys Leu Thr Glu Glu Asp Lys Ala Gly Phe Ile Lys Arg
915 920 925

Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His Val Ala Gly Ile Leu
930 935 940

His His Arg Phe Asn Lys Ala Glu Asp Thr Asn Glu Pro Ile Arg Lys
945 950 955 960

Val Arg Ile Ile Thr Leu Lys Ser Ala Leu Val Ser Gln Phe Arg Asn
965 970 975

Arg Phe Gly Ile Tyr Lys Val Arg Glu Ile Asn Glu Tyr His His Ala
980 985 990

His Asp Ala Tyr Leu Asn Gly Val Ile Ala Leu Ala Leu Leu Lys Lys
995 1000 1005

Tyr Pro Gln Leu Ala Pro Glu Phe Val Tyr Gly Glu Tyr Leu Lys
1010 1015 1020

Phe Asn Ala His Lys Ala Asn Lys Ala Thr Val Lys Lys Glu Phe
1025 1030 1035

Tyr Ser Asn Ile Met Lys Phe Phe Glu Ser Asp Thr Pro Val Cys
1040 1045 1050

Asp Glu Asn Gly Glu Ile Phe Trp Asp Lys Ser Lys Ser Ile Ala
1055 1060 1065

Gln Val Lys Lys Val Ile Asn His His His Met Asn Ile Val Lys
1070 1075 1080

Lys Thr Glu Ile Gln Lys Gly Gly Phe Ser Lys Glu Thr Val Glu
1085 1090 1095

Pro Lys Lys Asp Ser Ser Lys Leu Leu Pro Arg Lys Asn Asn Trp

1100							1105					1110			
Asp	Pro	Ala	Lys	Tyr	Gly	Gly	Leu	Gly	Ser	Pro	Asn	Val	Ala	Tyr	
1115						1120					1125				
Thr	Val	Ala	Phe	Thr	Tyr	Glu	Lys	Gly	Lys	Ala	Arg	Lys	Arg	Thr	
1130						1135					1140				
Asn	Ala	Leu	Glu	Gly	Ile	Thr	Ile	Met	Glu	Arg	Glu	Ala	Phe	Glu	
1145						1150					1155				
Gln	Ser	Pro	Val	Leu	Phe	Leu	Lys	Asn	Lys	Gly	Tyr	Glu	Gln	Ala	
1160						1165					1170				
Glu	Ile	Glu	Met	Lys	Leu	Pro	Lys	Tyr	Ala	Leu	Phe	Glu	Leu	Glu	
1175						1180					1185				
Asn	Gly	Arg	Lys	Arg	Met	Val	Ala	Ser	Asn	Lys	Glu	Ala	Gln	Lys	
1190						1195					1200				
Ala	Asn	Ser	Phe	Leu	Leu	Pro	Glu	His	Leu	Val	Thr	Leu	Leu	Tyr	
1205						1210					1215				
His	Ala	Lys	Gln	Tyr	Asp	Glu	Ile	Ser	His	Lys	Glu	Ser	Phe	Asp	
1220						1225					1230				
Tyr	Val	Asn	Glu	His	His	Lys	Glu	Phe	Ser	Glu	Val	Phe	Ala	Arg	
1235						1240					1245				
Val	Leu	Glu	Phe	Ala	Gly	Lys	Tyr	Thr	Leu	Ala	Glu	Lys	Asn	Ile	
1250						1255					1260				
Glu	Lys	Leu	Glu	Lys	Ile	Tyr	Lys	Glu	Asn	Gln	Thr	Asp	Asp	Leu	
1265						1270					1275				
Ala	Lys	Leu	Ala	Ser	Ser	Phe	Val	Asn	Leu	Met	Gln	Phe	Asn	Ala	
1280						1285					1290				
Met	Gly	Ala	Pro	Ala	Asp	Phe	Lys	Phe	Phe	Asp	Val	Thr	Ile	Pro	
1295						1300					1305				

Arg Lys Arg Tyr Thr Ser Leu Thr Glu Ile Trp Gln Ser Thr Ile
1310 1315 1320

Ile His Gln Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Arg Met
1325 1330 1335

Gly Lys
1340

<210> 142
<211> 1327
<212> PRT
<213> Enterococcus sp.

<223> APG02998_RGN

<400> 142

Met Lys Lys Asn Tyr Thr Ile Gly Leu Asp Ile Gly Thr Ala Ser Val
1 5 10 15

Gly Trp Ala Val Leu Thr Glu Asp Tyr Asp Leu Val Lys Arg Lys Met
20 25 30

Lys Ile Ser Gly Asn Thr Gln Lys Lys Ala Val Lys Lys Asn Phe Trp
35 40 45

Gly Val Arg Leu Phe Glu Gln Gly Glu Thr Ala Glu Gly Arg Arg Leu
50 55 60

Lys Arg Thr Thr Arg Arg Arg Ile Ala Arg Arg Arg Gln Arg Ile Gln
65 70 75 80

Tyr Leu Arg Thr Ile Phe Asp Glu Ala Met Asn Gln Val Asp Ala Asn
85 90 95

Phe Phe Ala Arg Leu Asp Glu Ser Phe Ser Ile Thr Asp Glu Lys Glu
100 105 110

Asn Glu Arg His Pro Ile Phe Gly Asn Val Ala Glu Glu Val Ala Tyr
115 120 125

His Glu Gln Phe Arg Thr Ile Tyr Arg Leu Arg Glu Val Leu Ala Asn
130 135 140

Ser Ser Glu Gln Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145 150 155 160

Ile Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Glu Leu Asn Thr
165 170 175

Glu Asn Ser Ser Val Ser Gly Thr Phe Glu Gln Phe Ile Lys Val Tyr
180 185 190

Asn Glu Thr Phe Asn Val Glu Lys Val Leu Glu Leu Thr Val Asp Leu
195 200 205

Asp Glu Ile Ala Lys Gln Lys Val Ser Arg Met Lys Arg Ala Glu Leu
210 215 220

Ile Leu Ser Leu Phe Pro Glu Glu Lys Ser Thr Gly Asp Phe Ala Gln
225 230 235 240

Phe Ile Lys Met Ile Val Gly Asn Gln Gly Asn Val Lys Lys Thr Phe
245 250 255

Ser Leu Asn Glu Asp Ala Lys Ile Gln Phe Ser Lys Glu Glu Tyr Glu
260 265 270

Glu Asn Leu Glu Thr Leu Leu Ala Glu Ile Gly Glu Asp Phe Arg Gly
275 280 285

Val Phe Asp Ala Ala Lys Ser Val Tyr Asp Ala Ile Ser Leu Ala Asn
290 295 300

Ile Leu Lys Val Thr Asp Ala Ala Thr Arg Ala Lys Leu Ser Ser Ser
305 310 315 320

Met Val Ala Arg Phe Lys Glu His Lys Glu Asp Leu Lys Ala Leu Lys
325 330 335

Gln Phe Val Arg Asp Asn Leu Pro Asp Glu Tyr Asp Asp Leu Phe Lys
340 345 350

Asn Lys Lys Val Ala Gly Tyr Ala Gly Tyr Ile Asp Gly Asp Ala Thr
355 360 365

Gln Glu Glu Phe Tyr Lys Tyr Leu Lys Lys Ala Leu Ala Lys Thr Thr
370 375 380

Gly Ala Glu Tyr Phe Leu Ala Lys Ile Glu Gln Glu Asp Phe Leu Arg
385 390 395 400

Lys Gln Arg Thr Phe Asp Asn Gly Val Ile Pro Tyr Gln Leu His Leu
405 410 415

Glu Glu Leu Lys Ala Ile Ile Arg Asn Gln Lys Pro Tyr Tyr Pro Phe
420 425 430

Leu Asp Glu Glu Lys Ile Ser Gln Leu Met Thr Phe Arg Ile Pro Tyr
435 440 445

Tyr Val Gly Pro Leu Ala Gln Gly Gln Gly Asn Phe Ala Trp Leu Thr
450 455 460

Arg Lys Glu Ser Gly Lys Ile Thr Pro Trp Asn Leu Asn Glu Lys Val
465 470 475 480

Asp Ile Glu Lys Ser Ala Thr Asp Phe Val Glu Arg Met Thr Asn Asn
485 490 495

Asp Ser Tyr Leu Pro Met Glu Lys Val Leu Pro Lys His Ser Leu Leu
500 505 510

Tyr Glu Lys Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Arg Tyr Met
515 520 525

Asp Asp Asn Gly Arg Val Gln Asn Phe Ser Ser Lys Glu Lys Gln Gln
530 535 540

Ile Val Asn Asp Leu Phe Lys Gln Gln Arg Lys Val Lys Lys Glu Ile
545 550 555 560

Leu Glu Ala Phe Ile Lys Asn Glu Tyr Gly Ile Glu Asn Pro Lys Val
565 570 575

Glu Gly Ile Glu Lys Ala Phe Asn Ala Ser Leu Gly Thr Tyr His Asp
580 585 590

Leu Val Lys Leu Gly Ile Arg Pro Ala Leu Phe Glu Gln Pro Glu Tyr
595 600 605

Glu Gln Gln Phe Glu Gln Ile Val Lys Ile Leu Thr Val Phe Glu Asp
610 615 620

Arg Lys Met Arg Arg Lys Gln Leu Glu Gln Phe Ser Asp Leu Leu Thr
625 630 635 640

Glu Asp Glu Gln Lys Gln Leu Glu Arg Lys His Tyr Lys Gly Trp Gly
645 650 655

Arg Leu Ser Ala Lys Leu Ile His Gly Ile Val Asp Gln Lys Thr Gln
660 665 670

Lys Thr Ile Leu Asp Tyr Leu Ile Asp Asp Asp Asp Leu Pro Arg Asn
675 680 685

Arg Asn Arg Asn Phe Met Gln Leu Ile Asn Asp Glu Asn Leu Ser Phe
690 695 700

Lys Glu Glu Ile Glu Lys Ile Ala Phe Asp Asn Asp Lys Ser Thr Glu
705 710 715 720

Glu Ile Val Gln Glu Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile
725 730 735

Leu Gln Ser Leu Lys Ile Val Glu Glu Ile Ile Glu Ile Met Gly Glu
740 745 750

Leu Pro Thr Asn Ile Val Val Glu Met Ala Arg Glu Asn Gln Thr Thr
755 760 765

Ala Gln Gly Asn Arg Ala Ser Gln Ala Arg Met Lys Tyr Leu Glu Glu
770 775 780

Ser Ile Lys Lys Leu Gly Ser Ser Ile Leu Glu Asp Glu Pro Ile Ser
785 790 795 800

Lys Asp Ala Asn Leu Leu Arg Asn Asp Arg Leu Phe Leu Tyr Tyr Leu
805 810 815

Gln Asn Gly Arg Asp Met Tyr Thr Gly Asp Glu Leu Asp Ile Asn Asn
820 825 830

Leu Ser Ser Tyr Asp Ile Asp His Ile Ile Pro Gln Ser Phe Val Lys
835 840 845

Asp Asp Ser Ile Asp Asn Arg Val Leu Thr Thr Gln Arg Ile Asn Arg
850 855 860

Gly Lys Ser Asn Thr Val Pro Ala Glu Ser Val Val Lys Asp Met Arg
865 870 875 880

Ala Thr Trp Glu Arg Leu Tyr Ala Ser Asn Leu Ile Ser Lys Lys Lys
885 890 895

Leu Asn Asn Leu Met Lys Ala Val Asn Gly Gly Leu Thr Glu Glu Asp
900 905 910

Lys Ala Tyr Phe Ile Gln Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
915 920 925

Lys Asn Val Ala Gln Ile Leu His Gln Lys Tyr Asn Asn Glu Gln Ser
930 935 940

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

Gln Phe Arg Lys Asp Phe Ser Leu Tyr Lys Ile Arg Glu Leu Asn Asp
965 970 975

Tyr His His Ala Gln Asp Ala Tyr Leu Asn Gly Val Ile Ala Gln Ala
980 985 990

Leu Leu Lys Val Tyr Pro Lys Leu Glu Pro Glu Phe Val Tyr Gly Glu
995 1000 1005

Tyr Gln Lys Val Ser Ile Arg Ala Leu Asn Lys Ala Thr Ala Lys
1010 1015 1020

Lys Glu Thr Tyr Ser Asn Ile Met Lys Phe Phe Thr Asn Asp Glu
1025 1030 1035

Val Leu Thr Asn Lys Glu Thr Gly Glu Val Leu Trp Asn Arg Gln
1040 1045 1050

Arg Asp Ile Lys Thr Ile Lys Lys Val Met Asn Tyr His Gln Met
1055 1060 1065

Asn Ile Val Lys Lys Val Glu Ile Gln Thr Gly Arg Phe Thr Lys
1070 1075 1080

Glu Ser Ile Gln Pro Lys Gly Pro Ser Lys Lys Leu Ile Ala Arg
1085 1090 1095

Lys Asn Asn Trp Ala Pro Val Asn Tyr Gly Gly Phe Asp Ser Pro
1100 1105 1110

Thr Val Ala Tyr Ser Val Ile Ile Thr His Glu Lys Gly Lys Ala
1115 1120 1125

Ala Lys Val Val Gln Gln Leu Val Gly Ile Lys Ile Leu Glu Arg
1130 1135 1140

Gln Ala Phe Glu Gln Asn Glu Val Ala Phe Leu Glu Gly Lys Gly
1145 1150 1155

Phe Ile His Pro Lys Val Gln Leu Lys Leu Pro Lys Tyr Ser Leu
1160 1165 1170

Tyr Gln Phe Ala Asp Gly Arg Arg Arg Leu Leu Ala Ser Ala Asp
1175 1180 1185

Glu Ala Gln Lys Gly Asn Gln Met Val Leu Pro Thr Glu Leu Ile
1190 1195 1200

Glu Leu Leu Tyr His Ala Lys His Val Asn Asp Ser Ser Gly Glu
1205 1210 1215

Ser Leu Thr Tyr Leu Glu Asn His Arg Thr Lys Phe Ala Glu Leu
1220 1225 1230

Leu Glu Val Ile Leu Gln Phe Thr Asn Arg Tyr Val Gly Ala Glu
1235 1240 1245

Asp Arg Gln Lys Thr Val Arg Asp Leu Tyr Glu Lys Asn Gln Asp
1250 1255 1260

Ala Asp Met Lys Glu Leu Ala Pro Ser Phe Val Gln Leu Leu Gln
1265 1270 1275

Leu Asn Lys Gln Gly Ala Pro Ala Ser Phe Lys Phe Phe Gly Glu
1280 1285 1290

Thr Ile Glu Arg Ser Arg Tyr Lys Asn Met Thr Glu Leu Asn Lys
1295 1300 1305

Ala Thr Phe Ile Asn Gln Ser Ile Thr Gly Leu Tyr Glu Thr Gln
1310 1315 1320

Arg Arg Leu Val
1325